Environmental Change Impacts on Marine Calcifiers: Spatial and Temporal Biomineralisation Patterns in Mytilid Bivalves

Luca Telesca

Gonville and Caius College

University of Cambridge

This dissertation in submitted for the degree of

Doctor of Philosophy

Department of Earth Sciences
&
British Antarctic Survey

September 2018

Environmental Change Impacts on Marine Calcifiers: Spatial and Temporal Biomineralisation Patterns in Mytilid Bivalves

Luca Telesca

Abstract

Environmental change is a major threat to marine ecosystems worldwide. Understanding the key biological processes and environmental factors mediating spatial and temporal species' responses to habitat alterations underpins our ability to forecast impacts on marine ecosystems under any range of scenarios. This is especially important for calcifying species, many of which have both a high climate sensitivity and disproportionately strong ecological impacts in shaping marine communities. Although geographic patterns of calcifiers' sensitivity to environmental changes are defined by interacting multiple abiotic and biotic stressors, local adaptation, and acclimation, knowledge on species' responses to disturbance is derived largely from short- and medium-term laboratory and field experiments. Therefore, little is known about the biological mechanisms and key drivers in natural environments that shape regional differences and long-term variations in species vulnerability to global changes.

In this thesis, I examined natural variations in shell characteristics, both morphology and biomineralisation, under heterogeneous environmental conditions i) across large geographical scales, spanning a 30° latitudinal range (3,334 km), and ii) over historical times, using museum collections (archival specimens from 1904 to 2016 at a single location), in mussels of the genus *Mytilus*. The aim was to observe whether plasticity in calcareous shell morphology, production, and composition mediates spatial and temporal patterns of resistance to climate change in these critical foundation species.

For the morphological analyses, the combined use of new statistical methods and multiple study systems at various geographical scales allowed the uncoupling of the contribution of development, genetic status, and environmental factors to shell morphology. I found salinity had the strongest effect on the latitudinal patterns of *Mytilus* shape. Temperature and food supply, however, were the main predictor of mussel shape heterogeneity. My results suggest the potential of shell shape plasticity in *Mytilus* as a powerful indicator of rapid environmental changes.

I found decreasing shell calcification towards high latitudes. Salinity was the best predictor of regional differences in shell deposition, and its mineral and organic composition. In polar, low-salinity environments, the production of calcite and organic shell layers was increased, while aragonite deposition was enhanced under temperate, higher-salinity regimes. Interacting strong effects of decreasing salinity and increasing food availability on compositional shell plasticity predict the deposition of a thicker external organic layer (periostracum) at high latitudes under forecasted future

conditions. This response potential of *Mytilus* shell suggests an enhanced protection of temperate mussels from predators and a strong capacity for increased resistance of polar and subpolar individuals to dissolving water conditions.

Analyses of museum specimens indicated increasing shell calcification during the last century. Deposition of individual shell layers was more closely related to temporal changes in the variability of key environmental drivers than to alterations of mean habitat conditions. Calcitic layer and periostracum showed marked responses to alterations of biotic conditions, suggesting the potential of mussels to trade-off between the deposition of calcareous and organic layers as a compensatory response to strategy-specific predation pressure. These changes in biomineralisation indicated a marked resistance to environmental change over the last century in a species predicted to be vulnerable, and how locally heterogeneous environments and predation levels can have a stronger effect on *Mytilus* responses than global environmental trends.

My work illustrates that biological mechanisms and local conditions, driving plastic responses to the spatial and temporal structure of multiple abiotic and biotic stressors, can define geographic and temporal patterns of unforeseen species resistance to global environmental change.

A Vito e Carlo

My mama always said
"Life was like a box of chocolates.
You never know what you're going to get."

Forrest Gump (1994)

Acknowledgment

Declaration

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except where specifically indicated in the text.

It is not substantially the same as any that I have submitted, or is being concurrently submitted, for a degree or other qualification at the University of Cambridge or any other. I further state that no substantial part of my dissertation has already been submitted, or is being concurrently submitted, for any such degree or other qualification at the University of Cambridge or any other University or similar institution.

It does not exceed the prescribed page limit specified by the Earth Sciences and Geography Degree Committee.

Luca Telesca

September 2018

Table of Contents

List of Figures

List of Tables

Chapter One

1 Introduction

1.1 Environmental change

Human influence on the climate system is clear. Since the industrial revolution, anthropogenic greenhouse gas emissions have increased, driven largely by economic and population growth, and have reached the highest level in history [1]. This has led to faster increases in the atmospheric concentrations of carbon dioxide, methane, and nitrous oxide than Earth has ever previously experienced (40%, 150% and 20%, respectively, Figure 1.1) [2–4]. Their effects, together with those of other anthropogenic drivers (e.g. land use/cover, ozone depletion), have been detected throughout the climate system and are extremely likely to have been the dominant cause of the observed changes since the mid-20th century, among which are climate warming and acidification of oceans [1, 2, 5]. Atmospheric carbon dioxide (CO_2) has increased from 278 to 400 parts per million (ppm) over the industrial period and, together with the other greenhouse gases, has driven a series of major environmental changes [1, 6]. About half of the anthropogenic CO_2 emissions between 1750 and 2011 have occurred in the last 40 years [1]. Forty per cent of these emissions have remained in the atmosphere, 30-32% was removed from the atmosphere and stored on land (i.e. in plants and soils), and the rest are in the ocean [1].

The ocean covers about 71% of Earth's surface and represents more than 90% of its habitable space. It hosts 25% of all eukaryotic species [7], provides 11% of animal protein consumed worldwide [8], and more. Overall, oceans play particularly important ecological and economical roles, providing invaluable services [9]. The global ocean (including enclosed seas) plays a major role in global climate dynamics acting as a climate integrator [1, 6]. Indeed, it has absorbed 93% of Earth's additional heat since the 1970s, offsetting much atmospheric warming, but increasing ocean temperature and sea level with deleterious effects for most ecosystems [9, 10]. It also captured 28-30% of the anthropogenic CO_2 emissions produced over the last 250 years by fossil fuel burning and other human activities [1, 11, 12]. Plankton converts some of that CO_2 into organic matter, part of which is exported into the deeper ocean, while the remaining CO_2 causes progressively increasing acidification of the ocean [2, 6, 11–14]. Moreover, it accumulated nearly all water resulting from melting glaciers and ice sheets, hence furthering the rise in sea level and the decrease of local water salinity [1, 15–17]. Thus, the ocean moderates anthropogenic climate change at the cost of major changes in its fundamental chemistry and physics [6]. These alterations of ocean properties profoundly affect species' biogeography and
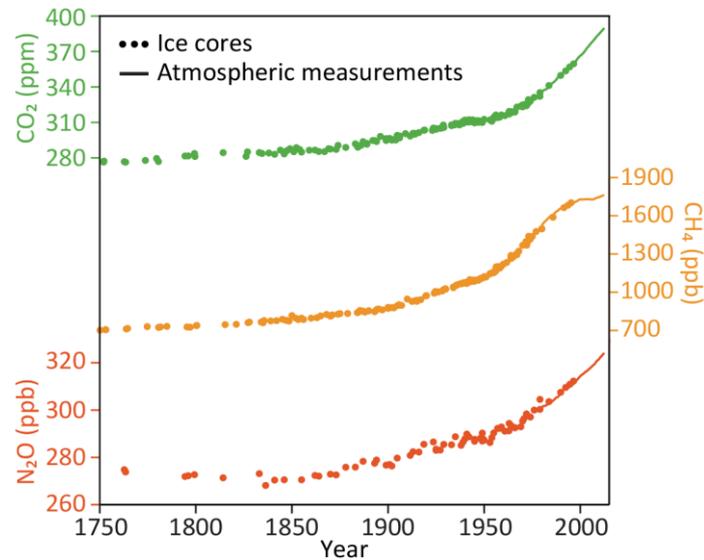
Figure 1.1 Globally averaged greenhouse gas concentrations
Observed increase in atmospheric greenhouse gases concentration since the industrial revolution [1]. Atmospheric concentrations of carbon dioxide (CO_2, green), methane (CH_4, orange), and nitrous oxide (N_2O, red). Data from ice cores (dots) and direct atmospheric measurements (lines) are overlaid.

phenology [18–22], community structure, and ecosystem dynamics, as well as biogeochemical cycling [1, 9, 23–26]. Moreover, such changes inevitably affect marine and coastal ecosystems and the services they provide [6, 26].

1.1.1 Physical and chemical changes

Anthropogenic emissions are changing the physical and chemical properties of the ocean. These alterations include increasing surface temperature and sea level, decreasing pH, lower saturation state of calcium carbonate, as well as alterations of surface salinity and dissolved oxygen concentration (Figure 1.2) [1].

Warming of the climate system is unequivocal [1]. Each of the last three decades has been successively warmer at the Earth's surface with the period from 1983 to 2012 being the warmest of the last 1400 years in the Northern Hemisphere (Figure 1.2a) [1]. The globally averaged combined (both land and ocean) surface temperature data indicates a mean linear warming of 0.85°C over the 1880 - 2012 period [1]. Ocean warming dominates the increasing energy stored in the climate system, accounting for more than 90% of the total heat accumulated between 1971 and 2010 with only about 1% stored in the atmosphere [1]. On a global scale, the ocean warming is largest near the surface, with a mean warming of 0.11°C per decade [1]. Temperature is projected to rise over the 21st century under all assessed emission scenarios (Figure 1.2a) [1] with a series of likely consequences on acidification, water cycle, surface salinity, and sea level rise.
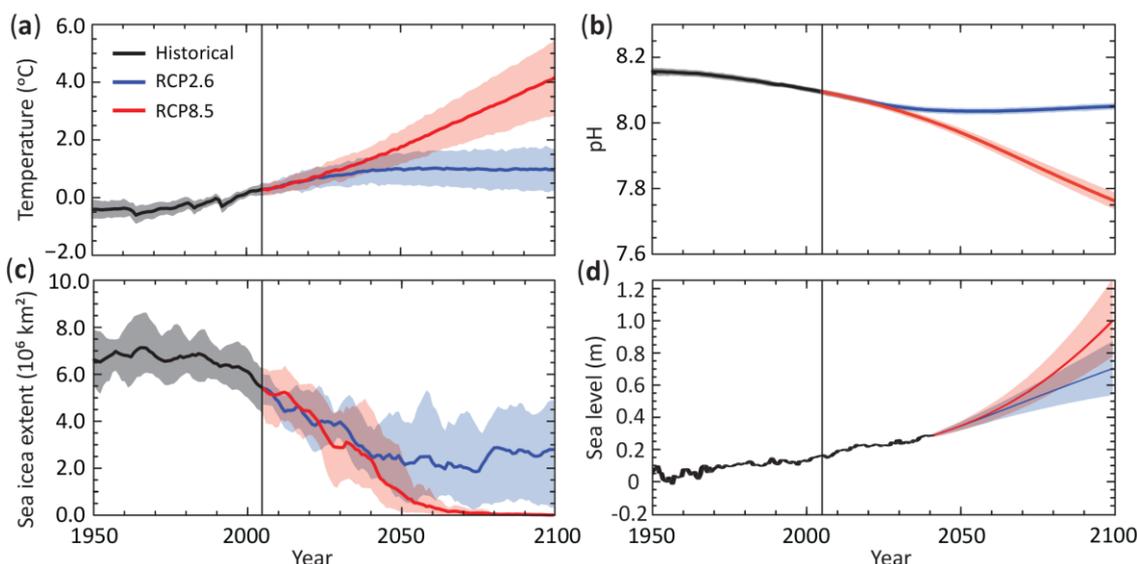
Figure 1.2 Environmental changes over the industrial period and 21st century
Time series from 1950 to 2100 from the IPCC report [1] for (**a**) change in global annual mean surface temperature, (**b**) global mean ocean surface pH, (**c**) Northern Hemisphere September sea ice extent, and (**d**) global mean sea level variation. Time series of projections and 95% prediction intervals (shaded areas) are shown for the stringent emission-mitigation scenario RCP2.6 (blue) and the business-as-usual high-emission scenario RCP8.5 (red). Black (grey shading) is the modelled historical evolution.

Among the various contributors to radiative forcing (i.e. the difference between solar energy absorbed and radiated back by the Earth) between 1750 and 2012 (e.g. ozone, methane, and nitrous oxide), CO_2 has the strongest effects on the ocean [27]. Indeed, CO_2 accounts for two or more times the warming attributed to the non-CO_2 greenhouse gases [27] and causes a process known as **ocean acidification** [12, 28]. Specifically, the ocean carbonate system is governed by a series of chemical equilibria:

$$CO_{2(atm)} \leftrightarrow CO_{2(aq)} + H_2O \leftrightarrow H_2CO_3 \leftrightarrow H^+ + HCO_3^- \leftrightarrow 2H^+ + CO_3^{2-}$$

Equation 1.1

The uptake of excess anthropogenic CO_2 by the ocean increases the partial pressure of carbon dioxide (pCO_2) and the formation of carbonic acid (H_2CO_3), which can dissociate to form hydrogen ions H^+, bicarbonate HCO_3^- and carbonate ions CO_3^{2-} (Equation 1.1). These reactions are reversible and are near equilibrium [2]. Increasing $CO_{2(atm)}$ increases $CO_{2(aq)}$, shifting the equilibrium towards higher concentrations of carbonic acid, with formation of dissolved carbon species (Equation 1.1). Most of the H^+ which are produced neutralise when they react with CO_3^{2-}, however excess of acid increases $[H^+]$ which lowers seawater pH (-$\log[H^+]$) (Figure 1.2b) [14]. Variability of pH in coastal waters is considerably larger than that in the open ocean, partly driven by upwelling [29], freshwater input [30], eutrophication [31], and biogeochemical processes [32].

Among the diverse consequences of warming there is the progressive **melting of ices sheets** [16]. Over the last two decades, the Greenland ice sheet has been shrinking and losing mass (Figure 1.2c) [1, 16].

The annual mean Arctic sea ice extent decreased over the 1979 – 2012 period proceeding at a rate of 3.5 - 4.1% per decade [1]. This lost mass has also contributed to sea level rise and local decrease in water salinity [1, 15]. While in the Antarctic sea ice extent increased by 1.2 - 1.8% per decade between 1979 and 2012 [1]. However, there are strong regional differences in Antarctica, with extent increasing in some regions and decreasing in others [1].

Over the 1901 - 2010 period, global mean **sea level** rose by a mean of 0.19 m (Figure 1.2d), at a rate of change that has been larger than the mean rate during the previous two millennia [1]. Since the early 1970s, glacier mass loss and ocean thermal expansion from warming together have explained about 75% of the observed global mean sea level rise [1, 5]. Rates of sea level rise over broad regions can be several times larger or smaller than the global mean sea level rise (e.g. Western vs Eastern Pacific) for periods of several decades, due to fluctuations in ocean circulation [1].

In addition, a number of studies have also considered surface **salinity** changes on a global scale during the 20th century [33, 34]. It has long been noted that mean sea surface salinity patterns are highly correlated with intensification of the water cycle [1, 34]. Indeed, these provide indirect evidence for changes in evaporation and precipitation over the oceans [1, 17]. Documented surface patterns are following the so called "rich get richer" mechanisms [35], where salty ocean regions (compared to the global mean) are getting saltier, whereas fresh regions are getting fresher. In the broad scale, these changes suggest that, since the 1950s, surface water salinity has increased in evaporation-dominated mid-latitudes and decreased in the rainfall-dominated areas, such as tropical atmospheric convergence zone and polar regions [34, 35].

Moreover, **dissolved oxygen** concentration levels are decreasing worldwide with ocean warming being a potential driver [1]. It is very likely that the dissolved oxygen content of the ocean will decrease by a few percent during the 21st century in response to increasing warming, predominantly in the subsurface mid-latitude oceans. However, there is no consensus on the future oxygen volume in the open ocean because of large uncertainties in potential biogeochemical effects, and in the evolution of tropical and temperate ocean dynamics [1, 5].

In summary, the carbon emitted will change the earth system irreversibly and the condition of the future ocean will depend on the amount of carbon emitted in the coming decades with potential different more or less deleterious cascade effects on various chemical and physical aspects of the ocean (Figure 1.2) [6].

1.1.2 Biological and ecosystem effects of environmental change

Environmental change is a major force shaping the future of our oceans [6, 26]. Organisms and ecosystems are changing in response to rapid environmental alterations. The inherent difficulty of distinguishing climate signals from natural variability [36], and of accounting for genetic adaptation [37], makes these shifts difficult to document, but nevertheless broad anthropogenic impacts are evident. In addition, forecasting emergent consequences of climate change on marine communities and ecosystems remains difficult [38]. Ecosystem-wide projections are severely constrained by heterogeneous patterns of ocean warming and acidification [6], multiple interacting stressors [39], species interactions, and species-specific effects [21]. Indeed, existing knowledge predominantly stems from short- to long-term experimentally induced responses in model organisms or simplified "communities" [21, 23, 24, 26, 40–42], while complex variations under multiple stressors have rarely been investigated in natural environments [23, 24, 43, 44]. However, such conclusions may not necessarily translate in long-term acclimated or adapted natural populations within the complexity of a dynamic ecological system [38, 45–47].

1.1.3 Multiple drivers

Investigations of single drivers can produce misleading inferences about organismal responses in a multivariate environment because interactive (additive, synergistic, or antagonistic) effects often are not predictable from single driver studies [38, 48]. This is a major source of uncertainty for projections [37, 38], but recent studies have better characterised interactions among some drivers [43, 49]. Changes in temperature and pH, such as those projected for the year 2100, can have synergistic negative effects on species growth, survival, fitness, calcification, and development [42, 50, 51]. Growing evidence also suggests that interactions with other environmental factors, such as irradiance, nutrient availability, geographic location, and species community composition, can strongly impact the biological effects of warming and ocean acidification [24, 52–56]. Other direct human impacts can reduce the adaptive capacity of marine species and ecosystems to CO_2-related effects. For example, fishing reduces species diversity, simplifies the trophic food web, and increases ecosystem sensitivity to climate change [57–59]. Because relatively little is known on the interacting effects of environmental factors and the complexity of the marine food web, it is premature to make ecosystem-wide projections [6, 55]. However, impacts on keystone species and ecosystem engineers are likely to shift whole communities [24, 25, 56, 60–62].

1.1.4 Environmental change study approaches

Various approaches have been developed for environmental change research, including laboratory experiments, mesocosm studies, field experiments, and analyses of archival collections. However, all these approaches have strengths and weaknesses with no single ideal method. Therefore, a combination of approaches is need for understanding impacts and for predicting responses to future change [47].

Laboratory experiments have the advantage to accurately manipulate environmental conditions in highly controlled settings and to allow observing responses to specific parameters and their interactions [47]. A limitation of this approach is its relatively short duration, although, longer-term experiments can inform on the acclimation and/or adaptation potential in organisms with short generation times [20, 24, 40, 63, 64]. Most laboratory experiments have also not incorporated the role of intraspecific variation in responses to change [47, 63]. Moreover, few studies only have also included effects on species interactions (e.g. predation and competition), communities [23, 65], as well as multistressors impacts [53, 66].

Field experiments, such as *in situ* mesocosm or CO_2 vents experiments, were developed to increase the degree of realism and allow the study of species and population/community responses in (or a deviation from) their natural context [24, 56, 67]. Limitations of *in situ* studies include the difficulty of maintaining constant (both in their mean value or variation) experimental conditions, the acclimation and adaptation potential of organisms is not assessed due to their relatively short duration times, as well as being logistically challenging [21, 47, 67].

While laboratory and field experiments provide insights into species' responses to future changes, the analysis of archival specimens form **museum collections** can demonstrate how organisms have already responded to abiotic and biotic alterations [68–71]. Museum collections can be used as a reference for future responses by providing unique data on rate and magnitude of change we might expect in natural populations. Historical specimens can also provide long-term dataset for the assessment of possible adaptations that have already occurred in marine species [69, 70]. This approach in addition to field and laboratory experiments might provide greater understanding of species responses to environmental change.

1.1.5 Effects of change on calcifying organisms

Environmental change poses a major threat to marine organisms, among which species producing calcium carbonate shell and skeletons are possibly experiencing the strongest impact from rapid habitat alterations [14, 21, 23, 25, 26, 44, 72, 73]. Calcifiers are recognised as key species at the ecosystem level, as they have the potential to impact both community structure and ecosystem functioning [74].

Calcifying organisms are major producers of calcium carbonate ($CaCO_3$) in marine and estuarine ecosystems [75–77]. Studies of their responses to environmental change across a number of life-history stages suggest that larvae and adults will find it more difficult to build and maintain their $CaCO_3$ structures [73, 78–83]. Moreover, calcifiers are predicted to experience a range of negative impacts of ocean warming and acidification, including changes in metabolism, acid-base status, reduced reproduction, and survival in a climate-changed ocean [20, 40, 44, 70, 72, 81, 84–86].

One of the key consequences of warming on calcifiers is represented by potential species' range shifts, usually following a shift in isotherms or temperature extremes [9, 18, 25, 87]. Recent studies strongly reiterate that many species, including calcifiers, are undergoing phenological and geographical shifts as a result of warming [19, 87–89]. These will cause potentially permanent changes to ecosystems, including local extinctions, while simultaneously producing novel assemblages [24, 25, 56, 90]. However, calcifiers' responses to changing temperature will depend on species-specific windows of thermal tolerance [64, 87]. Indeed, there is medium confidence that animals adapted to a wide range of temperatures will cope better with future conditions, whereas tropical and polar specialists are at greatest risk [18, 91–93]. A good example of reef-building calcifiers that are extremely vulnerable to warming are corals [9, 85, 91]. Warming causes mass mortality of warm-water zooxanthellate corals through bleaching, as well as through biotic diseases, resulting in declines in coral abundance and biodiversity. Coral reefs can recover if thermal stress is minimal and of short duration [94]. However, ocean warming and acidification are expected to act synergistically to push corals and coral reefs into conditions that are unfavourable for coral reef ecosystems [95]. Recent studies have shown short-term acclimation and adaptation in some fast-growing species [96] and suggested that some genetic mechanisms may allow faster rates of change [97]. Another problem associated to warming is represented by the expanding oxygen minimum zones [9]. Higher temperatures increase species' sensitivity to hypoxia [53, 98], limiting the depth distribution of species not adapted to hypoxic conditions and leading to community-level shifts [91].

Ocean acidification is also documented to have series of deleterious effects on calcifying organisms [14]. These include reduced calcification, lower rates of repair, and weakened calcified structures, but responses are species-specific [23, 40, 74, 84, 99, 100]. Reproductive success, early life-stage survival, feeding rate, and stress-response mechanisms may also be affected [72, 74, 101]. For calcifiers, the predicted reduction in carbonate ions would make it more difficult to produce calcium carbonate structural components [2, 21, 44, 51, 102]. Specifically, rates of $CaCO_3$ formation and dissolution vary with saturation state (Ω) which is defined as the product of calcium and carbonate ion concentrations [2, 103]:

$$\Omega = [Ca^{2+}][CO_3^{2-}]/K'_{sp}$$

Equation 1.2

The solubility product K'$_{sp}$ depends on temperature, salinity, pressure and the mineral phase [13]. Generally, shell formation occurs in supersaturated conditions ($\Omega > 1$) and dissolution occurs in under-saturated waters ($\Omega < 1$). Therefore, the magnitude of the effect of ocean acidification on calcifiers depends largely on the $CaCO_3$ polymorph used to construct their skeleton/shell, with aragonite being more soluble than calcite [104], as well as the amount of organic layers protecting calcareous structures [81, 105]. The forecasted changes in seawater conditions can have, therefore, profound influence on ocean carbonate chemistry [2, 79]

Most studies have investigated these effects on isolated organisms; far less is known about the effects on communities and ecosystems. Indeed, more recent evidences suggest how calcifiers sensitivity is influenced by a series of direct or indirect effects, species interactions, and compensatory mechanisms [23, 24, 41, 54, 66, 106, 107]. These studies, including results from this thesis (Chapter 4 and 5) [108, 109], suggest how calcifiers' responses can vary significantly depending on the geographic regions analysed [44, 62], the potential for transgenerational adaptation [110], the indirect effect of altered food availability [53, 66], and potential compensatory mechanisms, such as increased feeding rate under increased food supply [65, 107].

1.2 Molluscs as study organisms

Among the organisms producing calcium carbonate structures, the phylum of Mollusca is one of the largest, most diverse, and important taxon in the animal kingdom, second only to Arthropoda in number of living species. It includes familiar animals such as clams, mussels, snails, slugs, octopods, and squids. Shelled molluscs have a significant contribution to local and global carbon cycling [74]. For example, for pteropods gastropods this contribution to the total global $CaCO_3$ flux can vary regionally between 10% and 50% on a regional basis [74]. There are about 85,000 extant species of molluscs [111], with estimates ranging from 50,000 to a maximum of 120,000 species (due to the large number of species described multiple times). The number of fossil species is estimated between 60,000 and 100,000 additional species [112]. The phylum has a long evolutionary history and a rich fossil record going back to the early Cambrian about 540 million years ago [113]. Mollusca are highly diverse, not just in size and in anatomical structure, but also in behaviour and in habitat [114, 115]. Molluscs are chiefly marine, comprising about 23% of all the named marine organisms. Some bivalves and gastropods live in freshwaters, but only gastropods are present on terrestrial habitats [114]. Seven classes of molluscs are recognised formally, mostly based on phylogenetic analysis and morphological characters in extant and fossil taxa: Aplacophora (~150 species), Polyplacophora (~1000 species), Monoplacophora (~30 species), Gastropoda (> 100,000 species), Bivalvia (> 20,000 species), Cephalopoda (~1,000 species), and Scaphopoda (~800 species) [113].

1.2.1 Bivalves

The focus of this thesis is on Bivalvia, which is the second largest class within the phylum of Mollusca. Extant species are an important component of marine and freshwater ecosystems, with more than 80% of molluscan species living in marine habitats [115]. Modern representatives include the common mussels, cockles, clams, and oysters, which have both significant ecological and economical value. The bivalves are in some ways the most highly modified of all the molluscs. They comprise two lateral calcareous shell valves, secreted by the mantle, hinged dorsally by a partially calcified elastic ligament, enclosing the soft tissue. One or two adductor muscles control the opening and closing of the shell valves. The foot is laterally flattened. Unique among the molluscs, bivalves have lost the radula and almost all rely on the enlarged gills representing one of the most efficient systems of ciliary feeding (Figure 1.3) [114, 115].

Bivalves have conquered a range of habitats from the deep sea to freshwater, exploited a wide range of life habits, and undergone a near-exponential taxonomic proliferation, an astonishing example of adaptive radiation [115]. Bivalves can be sessile epifaunal, infaunal burrowers (around 50% of all modern families) of soft and hard substrata, or even free-living. Most bivalves are marine, exploiting niches from the abyssal depths to intertidal zone, but successful groups have also invaded more brackish and freshwater conditions where some can form extensive beds and biostructures, providing foundation for other organisms and supporting significant amount of biodiversity [116, 117].

In addition to their wide distribution and ecological importance, bivalves have also a significant economical vale for fisheries and aquaculture industries worldwide. An examination of the Food and Agriculture Organization (FAO), Global Aquaculture Production data for 1950–2015 [8] shows that the number of bivalve species being cultured and commercial harvest have almost doubled over the last 10-year period [115].

In this thesis, the study organisms used for the analyses in Chapters 4, 5 and 6 are the pervasive Mytilids mussels of the family Mytilidae. In the following Sections 1.2.2 - 1.2.12 I provide information on their taxonomy, geographical distribution, ecology, anatomy and shell structure, as well as their forecasted responses to environmental changes.

(a)

Posterior byssal retractor muscle

Posterior byssal retractor muscle

Ligament

Posterior adductor muscle

Exhalant siphon

Inhalant aperture

Outer labial palp

Inner labial palp

Foot

Ventral pallial rejectory tract

Ventral marginal food groove

Byssus

(b)

Pericardium

Posterior byssal retractor muscle

Auricle   Ventricle

Rectum

Posterior byssal retractor muscle

Dorsal septum

Midgut

Posterior pedal retractor muscle

Posterior adductor muscle

Ligament

Visceral mass

Digestive diverticule

Ovary

Exhalant siphon

Outer labial palp

Siphonal septum

Anterior byssal retractor muscle

Inhalant aperture

Umbo

Inner demibranch

Anterior adductor muscle

Outer demibranch   Byssal gland

Inflected periostracum

Byssus   Foot   Mantle margin

Figure 1.3 Anatomy of a bivalve (mussel)
(**a**) Interior of the left side with the left valve and the left mantle skirt removed. (**b**) Partial dissection, showing part of the visceral mass. The ciliary currents are also shown (red arrows). Image redrawn after Morton [118].

1.2.2 Mytilidae

The Mytilidae is a family of considerable antiquity dating back to the Devonian times [119] and includes many important byssal attached genera such as *Mytilus*, *Perna*, *Modiolus*, and *Choromytilus*. The neotenous retention of the larval byssus by adult individuals made mytilids dominant space occupiers worldwide [119]. This allowed mussels to successfully exploit hard or semi-consolidated substrata [120, 121], and to dominate rocky shore habitats on all continents, especially at temperate latitudes [122].

1.2.2.1 Taxonomy

The focus of this thesis is on the species of the genus *Mytilus*, from the *Mytilus edulis* species-complex, specifically the mussels *M. edulis*, *M. trossulus*, and *M. galloprovincialis* (Figure 1.4) [123]. The classification of these species used here is based on Newell [124] and Carter *et al.* [125]:

Class: Bivalvia Linnaeus, 1758

  Subclass: Pteriomorphia Beurlen, 1944

    Order: Mytilida Férussac, 1822

      Superfamily: Mytiloidea Rafinesque, 1815

        Family: Mytilidae Rafinesque, 1815

          Subfamily: Mytilinae Rafinesque, 1815

            Genus: *Mytilus* Linnaeus, 1758

              Species: *Mytilus californianus* Conrad, 1837; California mussel

              Species: *Mytilus coruscus* Gould, 1861

and the so called *Mytilus edulis* species-complex:

              Species: <u>*Mytilus edulis*</u> Linnaeus, 1758; blue mussel

              Species: <u>*Mytilus galloprovincialis*</u> Lamarck, 1819; Mediterranean mussel

              Species: <u>*Mytilus trossulus*</u> Gould, 1850; foolish mussel

              Species: *Mytilus planulatus* Lamarck, 1819; Australian blue mussel

              Species: *Mytilus chilensis* Hupé, 1854; Chilean blue mussel
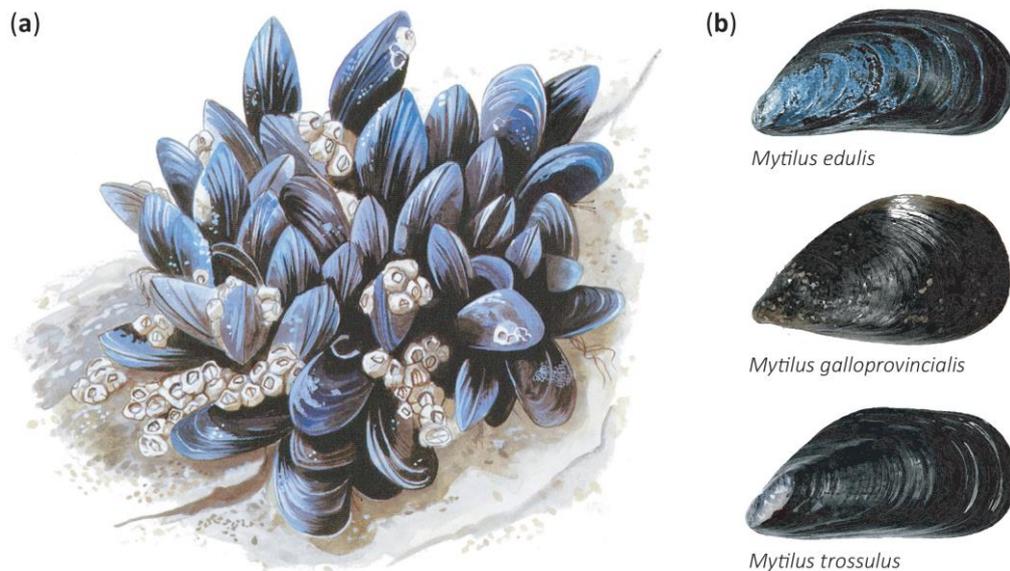
Figure 1.4 North Atlantic *Mytilus* species
(**a**) Example of a typical cluster of marine mytilids (*Mytilus* spp.) on a hard substratum from the eulittoral zone. (**b**) *Mytilus* species form the *Mytilus edulis* complex used as study organisms. Mussel cluster drawing courtesy of Gulfstream Editor.

1.2.3 Global and local distribution patterns

Marine mussels belonging to the genus *Mytilus* are widely distributed throughout the northern and southern hemispheres, and have been proved to be ideal model organisms for various physiological, biochemical, and genetic investigations [66, 126–129]. Geographically, mussels of the genus *Mytilus* dominate the eulittoral and sublittoral zones from mild-subtropical to polar waters around the world

(Figure 1.5). They settle in dense populations on any substratum that is firm enough to provide a secure anchorage, such as rocks, pebbles, concrete, shells, and wood [115, 126, 130].

The blue mussel *Mytilus edulis* has a wide geographical distribution in Norther hemisphere, spanning a 47°N to 81°N latitudinal range. Blue mussels occur along European coastlines from western France [131] to the North Sea [132–134], and on the Atlantic coast of North America from North Carolina northward to the Canadian Maritimes [116, 130]. Blue mussels are also present along the coasts of Iceland [135], Svalbard [132, 136, 137], and the western coastline of Greenland [62, 136, 138, 139]. On a local scale, *M. edulis* dominates the intertidal to subtidal regions of rocky shores. It is a quite eurytopic organism [62, 126, 130] and, therefore, has the capability to occupy a wide range of microhabitats. Blue mussels have the widest local distribution pattern in the genus, extending from high intertidal to subtidal zones from brackish (salinity of 4-6 psu) to fully marine environments (salinity 15-30 psu) [126, 130], areas with different temperatures regimes (from -13°C to 32°C) [126, 127, 140], and from sheltered to extremely wave-exposed shores.

Mytilus galloprovincialis has a relatively wide geographical distribution. It occurs in the Mediterranean Sea, Black Sea, continental Europe, from southern Spain to northern France, south-west Britain and Ireland, Arctic region, and North Africa [116, 123, 126, 131, 136, 141–144]. It is also present in South Africa, on the Pacific coast of North America, Japan, Hong Kong, south Australia, and New Zealand [130, 145, 146]. The Mediterranean mussel is also a very adaptable species, surviving in water with salinity values ranging from 12 psu to 38 psu [126] and temperature variations between 7°C and 27°C [147]. It is commonly found in intertidal and estuarine habitats, and tolerate high levels of wave exposure [121, 126].

Mytilus trossulus is restricted to cool water regions of the Northern hemisphere. Our understanding of the regional distribution of this species has significantly changed after a multitude of recent molecular studies revealing new *M. trossulus* populations previously thought to be *M. edulis* [131, 148]. *Mytilus trossulus* occurs, along the European coastline in the Baltic Sea, Barents Sea and White Sea [128, 131, 134, 136, 148–152]. There are records along the Norwegian [133, 136, 153] and northwest Greenlandic coastlines [136, 138]. It also inhabits the Atlantic coasts of Canada [151], and the Pacific coast of North America and Russia [115, 116]. This mussel is usually more abundant in brackish or sheltered coastal acres, where it tolerates low salinities (up to 4 psu, i.e. Baltic Sea). In these areas *M. trossulus* is usually very abundant contributing up to 90% to local hard bottom biomass, and thus it is an important habitat builder [51, 154].
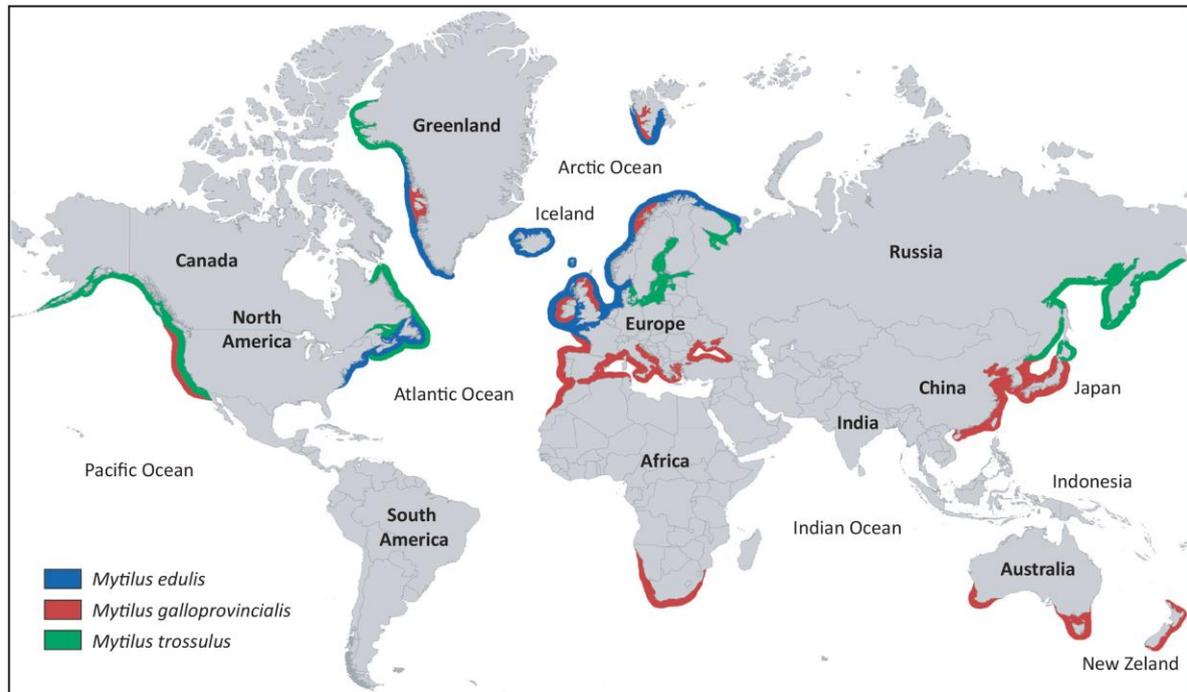
Figure 1.5 Approximate distribution of *Mytilus* spp.
Global distribution of *Mytilus edulis* (blue), *M. galloprovincialis* (red), and *M. trossulus*
(green). Image redrawn after Gosling [116].

1.2.4 Hybridisation

It is important to note that in the northern hemisphere, where the ranges of *M. edulis*, *M. trossulus*, and
M. galloprovincialis overlap (Figure 1.5), variable amounts of hybridisation occur between species
pairs [128, 142–144, 155, 156]. A pronounced hybrid zone of *M. edulis* × *M. trossulus* can be found in
the Baltic Sea [128, 149, 150]. *Mytilus edulis* and *M. galloprovincialis* hybridise in Europe producing
a mosaic pattern of both pure and hybrid populations [157] along the west coast of France [143],
Cornwall [123], and Ireland [141]. Hybridisation among all the three *Mytilus* taxa has also been
documented in an aquaculture context [155, 156, 158], indicating the potential of rope culturing to
increase local hybridisation [155].

1.2.5 Ecology

On a worldwide basis, mussels of the genus *Mytilus* form the foundation, both in terms of percent cover
and organic production, for most exposed rocky shore communities within the temperate zone [121,
122]. On a local scale, *Mytilus* mussels dominate the eulittoral and sublittoral regions [115, 126]. They
occur in a variety of shore habitats, from the sediment shores of protected bays and estuaries, through
gravel and pebble shores in semi-exposed condition, to true rocky shores which are exposed to

considerable wave action [126, 130]. In addition, mussel populations have the ability to colonise costal infrastructures, such as pier pilings, breakwaters, harbour walls, and oil platforms.

Mytilus species typically occur in intertidal habitats [121]. This distribution appears mostly controlled by predation pressure (i.e. both intensity and type of predators), and competition for space and food [121, 159]. Upper distribution limits for *Mytilus* are rather constant over long period of time [122]. In the high intertidal, emersion times, temperature extremes, and desiccation represent the most important physiological stressors in the determination of the upper limits in rocky intertidal sites [160, 161]. Lower zonation limits for *Mytilus* have been shown to be under the strong influence of predation, among which sea stars have been recognised as the most important predators [121, 122]. Other predators, such as crabs and dog whelks, and competitors for space, such as barnacles, other mussels or encrusting species, can have a significant influence depending on the geographic location [121, 159].

Mytilus communities represent dominant biomass structural components [122]. The presence of mussel drastically modifies the local environment [122, 126]. On rocky shores mussel beds provide a habitat for a few but abundant macroinvertebrate [121]. Moreover, the shells themselves provide secondary substratum for colonisation by many epibionts and epiphytes [117, 122, 126]. As *Mytilus* beds age and grow they increase not only the biological component, but they also enlarge their physical component, producing structurally complex entities that are capable to provide foundation to a diverse assemblage of associated fauna and flora [117, 121]. Bed thickness, connectedness between individuals, and sediment load, are all increased within the bed, changing dramatically the microhabitats under, between, and around the mussels, that may provide refuge for a myriad of associated fauna [117]. Because of their ecological importance, *Mytilus* assemblages have long served as the focal point for numerous studies on intertidal population and community ecology.

1.2.6 Economic value

In addition to the high level of biodiversity supported, mussel of the genus *Mytilus* have a significant economic value for fisheries and the aquaculture industry worldwide [8]. In 2016, the production from *Mytilus* aquaculture worldwide was of ~590,000 tonnes worth ~2,6 billion USD in 2016 [8]. At a global scale, *Mytilus chilensis* represents the most harvested species with ~301,000 t worth ~2,195 million USD in 2016. At a European scale, *M. edulis* represents the most harvested species with ~155,000 t worth 234 million USD, followed by *M. galloprovincialis* with ~103,000 t worth 90 million USD.

1.2.7 Key abiotic factors for growth

In general, it is important to note that key abiotic factors foremost temperature, salinity, and food supply not only set limits on the spatial distribution of bivalves, but also affect key aspect of mussel biology [121, 162]. Water **temperature** is widely acknowledged as an important factor in controlling growth rate and level of physiological activity in marine organisms. Temperature varies with latitude and there is a general consensus that growth in calcifiers, including bivalves, increases with rising temperature over the ecological range of the species [44, 102, 109, 126, 127, 163, 164]. Almada-Villela *et al.* [165] observed the effect of temperature on *M. edulis* and found that between 3°C and 20°C growth increased logarithmically; above 20°C growth declined sharply, while at lower temperatures (3°C and 5°C) it proceeded very slow but at a constant rate. Further results showed how *M. edulis* is well-adapted to life in different or fluctuating temperature regimes [126, 127, 166, 167].

Water **salinity** is a key factor regulating distribution, growth, physiological processes, and so survival of marine organisms [168]. Brackish estuaries and lagoons are known to be favourable habitats for mussel growth, but this probably reflects the higher food levels associated with these environments [121, 162]. Indeed, studies have demonstrated that low salinity concentrations adversely affect growth [51, 169, 170], strongly reduce calcification [51], and may even be lethal to *Mytilus* under extreme conditions [162, 171]. However, *Mytilus* can survive considerably reduced salinities, showing a marked decline in size at salinities down to 4-5 psu, such as in the inner part of the Baltic Sea [172].

Food supply, within the physiological limits for long-term survival, including phytoplankton, bacteria, organic detritus, and material of inorganic origin, is probably the single most important factor in determining growth rate, since if food in unavailable, sustained growth cannot occur. Mussels, like most bivalves, are efficient filters feeders, removing particles down to 2-3 μm with 80-100% efficiency [173]. Food supply has been shown to be have a critical role in bivalve growth, calcification, and stress resistance [53, 66, 126, 129, 130].

1.2.8 Key predators

Predation represents the single most important source of natural mortality in *Mytilus* [121, 174]. Many species predate on mussels, among which are crabs, gastropods, starfish and sea/shorebirds [115, 121, 174, 175]. Gastropods are significant predators of mussels worldwide. The **dogwhelk** *Nucella lapillus* is widely distributed on exposed shores in northern Europe and on the east coast of North America. It is especially abundant on rocky, wave-exposed shore [176, 177]. The dogwhelk uses dissolving secretions from the accessory bring organ, assisted by the rasping action of the radula, to drill a small hole through the shell around the umbo or adductor muscle insertion regions [126, 178–180]. Although levels of consumption may appear low, the high density of dog whelks has a serious impact on mussel

coverage on exposed shores [176], with the potential to consume 60-90% of mussels at a given site in one season [181].

Starfish are major mussel predators in many areas. These are clearly a major factor influencing the distribution and abundance of *Mytilus* on the lower shore and in the sublittoral zone [176]. Starfish predate on mussels and other bivalves by using force and causes the bivalve to gape just sufficiently to allow the starfish to extrude its stomach into the shell opening and digests the prey. Starfish have been regarded as "keystone" predators, since they exerts a disproportionate influence on community structure, in this case setting the lower limits of mussel distribution [174].

Crabs, particularly *Cancer*, *Carcinus* and *Pachygrapsus*, are also significant predators of mussels [176, 177]. Mortality from crabs is usually most intense in the lower shore, where crabs are able to forage for longer periods of time [176, 177]. Crabs employ size selection of prey, with the upper size limit that can be successfully attacked being directly related to the size of the crab [121, 127]. Indeed, small mussels are especially vulnerable and easy to crush by virtually all size ranges of crabs [182], whereas larger mussels are available only to large crabs with strong chelae. Therefore, vulnerability to crab predation generally declines with increasing body size during growth.

Several **bird** species predate on mussels, among which oystercatchers, gulls and ducks. Seabirds can determine significant mortalities in wave protected environments and mussel cultures [175, 183]. They usually select thin-shelled mussels because they are easier to crack or open than thick-shelled individuals. In sheltered bays, birds have a significant impact on wild populations with the potential to remove between 30% and 54% of the standing mussel stock [115, 121]. Mussel farms, with their very high densities of small thin-shelled mussels, can also be foraging hot spots for diving ducks, particularly during spring and autumn [121].

1.2.9 Anatomy

Mussels consist of a soft part and a hard shell. The soft part consists of the mantle, a visceral mass, gills, foot, and muscles, which are surrounded by a protective calcareous shell (Figure 1.6). The mantle covers the visceral mass dorsally and extends laterally forming two lobes which cover the inner surface of the valves. The mantle lobes enclose the body and delimits the mantle cavity. At the shell margin, the mantle margin divides into three longitudinal folds, separated by two grooves (Figure 1.7a). The inner fold is muscular and contains the pallial muscles, which are inserted along a curved pallial line that extends from the anterior to posterior parallel to the margin of the shell (Figure 1.6b). The middle fold is sensory and may bear different types of receptors. The outer fold, in conjunction with the entire outer mantle epithelium, secretes the shell. A large posterior and a small anterior adductor muscles are attached to the inner shell surface and in combination with the elastic shell ligament regulate shell
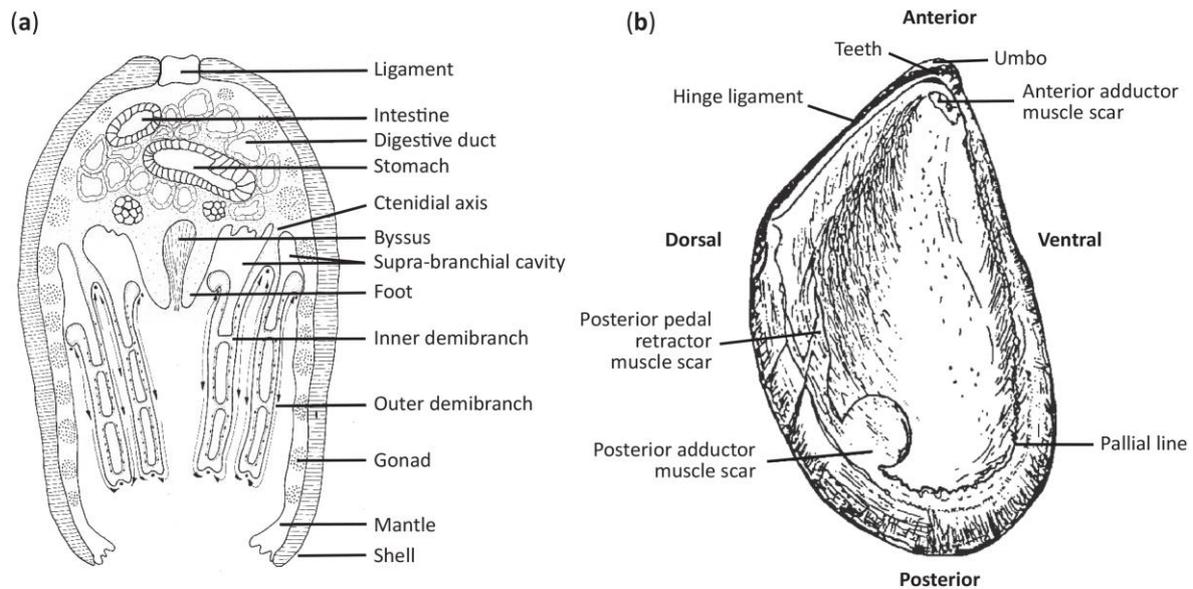
Figure 1.6 *Mytilus* **anatomy and shell**
(**a**) Diagrammatic traverse section through *Mytilus edulis* to show the visceral mass, gills form as well as the direction of the main ciliar currents (arrows, redrawn after Bayne [126]).
(**b**) Internal shell view of a blue mussel (redrawn after Gosling [115]).

opening and closure. The foot is laterally compressed, directed anteriorly, and secretes the byssus. Gills are wide with long folded filaments, which are held tougher by interlamellar junctions and are highly specialised for suspension feeding. A very small space, containing extra-pallial fluid, separates the mantle from the shell between the inner shell region delimited by the pallial line, and between the pallial line and the shell margin (Figure 1.7a). Into this space calcareous and organic material for shell formation is deposited.

1.2.10   Shell structure

Mussels consist of two convex, bilaterally symmetrical calcareous vales, similar in size (equivalve) and roughly triangular in shape (Figure 1.6b). The valves articulate with each other along the dorsal hinge. The two valves are joined dorsally by a two-part, elastic (proteinaceous) hinge ligament. Beside the hinge each valve bears a dorsal protuberance called the umbo (or beak). The pallial line runs along a small distance away from the shell margin more or less paralleling it. On the inside of each valve are two adductor muscle scars: the attachments for the large posterior and reduced anterior adductor muscles, and attachments for the foot retractor muscles. The exterior of the shell is characterised by the presence of concentric lines as a result of shell growth. Shell grows both **i**) in circumference by marginal accretion from the edge of the mantle lobe, and **ii**) in thickness by deposition from the general mantle surface.

The shell of the *Mytilus* species-complex includes three layers, one organic and two calcareous: (1) a thin outermost organic layer, the periostracum, (2) a middle calcareous layer, the fibrous prismatic layer, and (3) an inner calcareous layer, the nacreous layer (Figure 1.7). Calcareous shell layers are composed of calcium carbonate crystals in different forms (calcite prisms or aragonite tablets) shrouded by an envelope of organic matrix, which constitutes between 1% and 3% of shell weight [109].

The **periostracum** (Figure 1.7) is an organic layer consisting of the quinone tanned protein, conchiolin (a fibrous insoluble protein) [184, 185]. The periostracum is the first part of the shell to be formed. The "forming periostracum" is initiated by cells at the base of the periostracal groove [185], which lies between the outer and middle mantle fold (Figure 1.7a). The newly formed periostracum is thickened and matured (i.e. sclerotinisation by quinone tanning) by epithelial cells of the middle fold [186]. It then detaches from the middle lobe (forming the "free periostracum"), and then doubles back upon itself, dorsally, to cover the outer side of the prismatic layer (Figure 1.7a). Exposed on the outside of the shell, the "outer periostracum" undergoes decay, physical and biotic abrasion over time; this is why this layer is often lacking on the oldest part of the shell. The periostracum has two main functions: **i**) it provides an isolated compartment from the external aqueous environment where the deposition of the calcareous shell components is initiated [187], and **ii**) it provides protection against the corrosive effect of acidic waters, infestation by epibionts, predatory and endolithic borers, as well as a potential role in immunity [179, 185, 188, 189].

The **fibrous prismatic layer** (Figure 1.7) is a dark blue/purple coloured layer which consists in well-structured and ordered simple calcite crystals deposited in columnar prisms and surrounded by organic matrix (conchiolin) [190, 191] (Figure 1.7c). Calcite prisms progressively rotate during growth, forming an angle of ~30-45° to the shell surface with external crystals' edges oriented towards the umbo [192, 193]. The calcitic layer is formed by marginal accretion through nucleation of calcite prisms on the inner surface of the periostracum [194–196]. This layer has a finite thickness, since its deposition stops when the nacreous layer is laid down. Since calcite represents the more thermodynamically stable phase of calcium carbonate this may provide some degree of protection against dissolving conditions, although this might not be its primary role [197]. This layer has also the potential to provide protection against borers, drilling and shell-breaking predators [198, 199], and its thickness largely determine vulnerability of mussel shells to seabird predation [200]. A simple prismatic layer, called pallial myostracum, separates the calcitic layer and the innermost nacreous layer.

The **nacreous layer** (Figure 1.7) consists of layers of aragonitic hexagonal platelets separated by inter-lamellar layers of organic matrix [201–203]. Generally, the platelets are around 10-20 μm wide and are arranged in laminae, parallel to the inner surface of the valve, with a so called "brick wall structure", consisting in tablets in each layer that offset with respect to those in the layers above and below them [201, 203, 204] (Figure 1.7c). This layer is secreted by deposition from the general mantle surface and
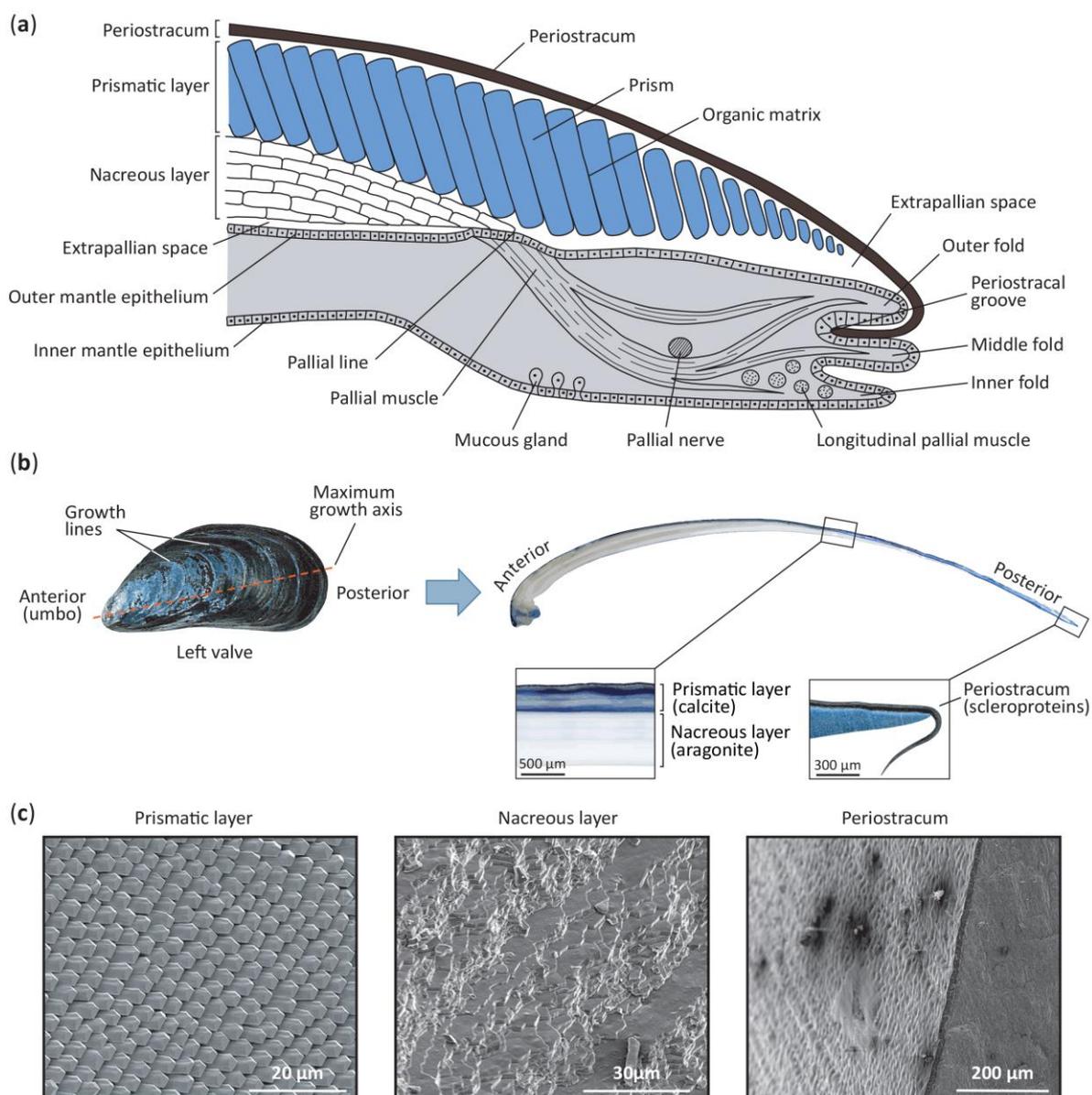
Figure 1.7 *Mytilus* shell structure: margin, section, and microstructure
(**a**) Diagrammatic example of transverse section of the edge of a mussel valve and mantle margin showing different shell layers and mantle edge anatomy. (Redrawn and modified after Ruppert *et al.* [114]). (**b**) *Mytilus* shell valve morphology and anteroposterior cross-section of shell valve along the axis of maximum growth (from umbo to posterior commissure, red dashed line) showing internal structure and arrangement of individual mineral (prismatic and nacreous layers) and organic (periostracum) shell layers. (**c**) Scanning electron micrographs showing the microstructure of shell layers in *Mytilus*: prismatic layer, nacreous layer, and periostracum.

grows continuously in thickness during aging. Aragonite is stabilised in biological microstructures, such as the nacreous layer, conferring a high mechanical strength to the shell [99].

The shape of shells in *Mytilus* is very heterogeneous depending on a range of biotic and abiotic factors among which temperature, salinity, food availability, age, population density, predation, and genotype [108, 159, 176, 205–208]. For example, mussel shell becomes progressively more elongated and

ventrally concave during growth [108, 159, 209]. The influence of genotype can be also evident in mussels from stable environmental conditions [134, 205, 210]. However, uncoupling the relative contribution of environmental vs genetic factors to shell shape may be difficult depending on both the geographical scale of analysis (local vs regional), environmental variability, and genetic data available [108, 205, 207, 211].

1.2.11   Bivalves and calcification

In organisms, calcification is the formation of calcium carbonate minerals requiring calcium ions (Ca^{2+}) and bicarbonate (HCO_3^-) to form $CaCO_3$ [212] *via* the reaction:

$$Ca^{2+} + HCO_3^- \rightarrow CaCO_3 + H^+$$
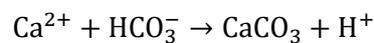
Equation 1.3

Experimental studies have suggested that seawater is the primary source of Ca^{2+} in calcifying bivalves [213, 214], accounting for ~90% of inorganic carbon substrate with the other ~10% from metabolically excreted CO_2 [215]. The mantle most likely controls the mechanisms of $CaCO_3$ deposition in molluscs [216]. Biomineralisation has been revealed to be a highly complex process with evidence suggesting both intracellular and extracellular pathways for calcification. Intracellular calcification models suggest $CaCO_3$ is initially formed as amorphous calcium carbonate (ACC), which is transported to the site of calcification, the organic matrix, where nucleation and crystal formation occurs [216–218]. Recent work, however, suggests an extracellular mineral formation process with substrates for mineralisation being transported to the calcifying space [196]. Furthermore, these findings are supported by evidences suggesting no major role of an ACC precursor molecule in biomineralisation [219, 220]. However, it cannot be ruled out that ACC containing vesicles are present on nanometre scales and rapidly transform into crystalline structures [221].

1.2.12   Environmental change impacts on Mytilidae

Climate change is considered a major force shaping marine ecosystems worldwide [21–23, 26], with ocean warming and acidification [5, 12] profoundly affecting species life history and ecology [18–22, 72], marine communities and ecosystems [6, 23–26, 222]. A number of studies and meta-analyses have suggested species producing $CaCO_3$ structures to be particularly vulnerable to rapid environmental changes [14, 21, 72, 74, 84, 101, 104]. However, responses of marine calcifiers and their sensitivity to disturbance can vary significantly depending on the taxonomic group, developmental stage, and geographic region analysed [18, 21, 46, 72, 223].

Atlantic mussels of the genus *Mytilus* are important bed-forming foundation species in eulittoral ecosystems worldwide (up to 90% of epibenthic biomass) [51], and represent valuable resources for aquaculture (192,000 tonnes produced in 2015 worth 325 million USD) [8]. Growing awareness of the consequences of environmental change on biodiversity and industry that *Mytilus* species support has stimulated a number of studies to estimate the response potential of these habitat-forming calcifiers to changing ocean conditions [51, 62, 83, 99, 109]. Given their wide distribution, ecological and economic value, *Mytilus* spp. represent exceptional model organisms to understand the regional differences in response potential of calcifiers to rapid climate changes and their potential effects on supported communities.

Mytilus growth, shell formation, and fitness are linked to multiple drivers, including water temperature, salinity, and food supply [51, 53]. Although *Mytilus* is a eurytopic species, a series of works have demonstrated a range of direct deleterious effects of change on different aspects on mussel biology, such as alterations of calcification and growth [16, 42, 103, 224–230, 70, 71, 74, 83, 86, 92, 99, 100], as well as ecology and larval development [43, 50, 78, 223, 231]. There is also a widespread consensus that adult individuals of marine organisms are less vulnerable to change than larval stages [21, 72, 223, 232], although recent studies suggest a potential tolerance in the latter [110, 196, 233]. However, existing knowledge on mussel sensitivity predominantly stems from experimental studies on single species or simplified experimental setup that cannot fully reproduce complex conditions in functioning ecosystems [25, 65, 107, 234–236].

Many recent studies suggest how the vulnerability of *Mytilus* (and of calcifiers in general), significantly vary depending on the geographic region analysed [44, 62, 102, 108, 132], the presence of multiple interacting stressor [43, 51, 109, 163], such as positive indirect effects of food supply [53, 66, 82, 129], as well as presence of compensatory mechanisms [109, 196]. All these factors have the strong potential to buffer indirectly the direct negative effects of change showing the strong potential of mussels to resist rapid environmental disturbances.

1.3 Aim and outline of thesis

Species producing calcium carbonate shell and skeleton, among them bivalves, are experiencing the strongest impacts of rapid environmental changes. Calcareous shells perform a range of vital functions in mytilids, including structural support, protection against predators and adverse environmental conditions. Because shell integrity determines survival, shell traits are subject to strong selection pressure, with functional success or failure a fundamental evolutionary driver. Therefore, any shift in environmental gradients impacting shell production, composition, and structure may alter mussel sensitivity to disturbance and affect its survival. However, existing knowledge on calcifiers' response potential predominantly stems from short- to long-term studies on single species or simplified experimental "communities", which may not necessarily translate to populations within the complexity of a dynamic ecological system.

Understanding key mechanisms and drivers mediating responses to disturbance in natural environments is, therefore, essential to project impacts of climate change to calcifiers under any range of scenarios. The aim of this thesis is to examine variations of shell characteristics (both morphology and biomineralisation) in *Mytilus* under heterogeneous environmental conditions across large geographical scales and over historical times to understand the potential of shell plasticity in mediating spatial and temporal impacts of global environmental change on these critical foundation species.

In Chapter 2, I provide information on the sampling design and the methodologies used for the characterisation of *Mytilus* shells, including thickness, organic content, and morphometrics analyses.

In Chapter 3, a detailed description of statistical methods used, including data exploration protocols and modelling approaches, is presented.

In Chapter 4, I investigate spatial patterns of shell morphology in *Mytilus* with latitudinal environmental gradients. I show how the used of new statistical methods and multiple study systems at various geographical scales made it possible to uncouple the contribution of environmental factors, development, and genetic status to shell morphology. Overall, the chapter describes the marked shell shape plasticity in mytilids and its potential as a powerful indicator of rapid environmental changes.

In Chapter 5, I analyse large-scale spatial patterns in shell biomineralisation under heterogeneous environmental gradients (30° latitudinal range), to test whether plasticity in shell deposition and composition mediates geographic patterns of resistance to climate change in *Mytilus*. I describe latitudinal changes in shell calcification with latitude and how salinity is the best predictor of regional differences in shell deposition, mineral and organic composition, suggesting a capacity for increased resistance of high-latitude mussel populations to future changes. This chapter illustrates how

mechanisms driving plastic responses to the spatial structure of multiple stressors can define geographic patterns of calcifiers resistance to environmental change.

In Chapter 6, I examine temporal patterns of shell morphology and biomineralisation over historical times (from 1903 to 2016 at a single location) using archival *Mytilus* specimens from museum collections (provided by the Royal Belgian Institute of Natural Sciences). I describe how shell calcification surprisingly increased during the last century, with the shell deposition being more closely related to temporal changes in the variability of environmental drivers than to alterations of mean habitat conditions. I also analyse how calcareous and organic shell layers respond to different level and types of predation and local abiotic conditions. This chapter illustrates a marked resistance to change over the last century in a species predicted to be vulnerable, and how locally heterogeneous environments and predation levels can have a stronger effect on *Mytilus* responses than global environmental trends.

In Chapter 7, I summarise the aim of my research and key findings, explaining their potential applications and original contribution to the field, and the possible areas for future research.

Overall, this thesis illustrates that biological factors and local conditions, driving plastic responses to the spatial and temporal structure of multiple abiotic and biotic stressors, can define geographic and temporal patterns of species resilience to global environmental change.

Chapter Two

2 Materials and methods for shell characterisation

In this thesis, I have analysed a variety of shell characteristics including the thickness of the individual layers forming the shell and their organic content, as well as the shape, or outline, of the shell valves, and their variations across wide geographical scales and over historical times. For these analyses, I have used *Mytilus* samples collected at different geographic and temporal scales, a range of different methodologies for characterising their shells, as well as a set of statistical and modelling approaches (Chapter 3). The details of samples collected and of each of these approaches are given below.

2.1 *Mytilus* samples collection

For this study, I collected current (both wild and cultured) and archival *Mytilus* specimens (*Mytilus edulis* and *M. trossulus*), for a total of 868 individuals, from 17 locations from different study systems in the North Atlantic area along European and Greenlandic coastlines (Figure 2.1).

- Wild *Mytilus* specimens (*M. edulis* and *M. trossulus*) were collected from a total of 17 populations along the North Atlantic, Artic and Baltic coastlines from four distinctive climatic regions (warm-temperate, cold-temperate, subpolar, and polar) (Figure 2.1a). Sampled populations covered a latitudinal range of 30° (a distance of 3,334 km), from Western European (Brest, North-West France, 48°N, site 1) to Northern Greenlandic (Qaanaaq, North-West Greenland 78°N, site 17) coastlines. Mussels were collected by hand from the eulittoral (low intertidal) zone on rocky shores between December 2014 and September 2015. I collected *Mytilus* specimens of various size classes (shell length of 26 - 81 mm) from juveniles to large adults from each site, for a total of 480 individuals. Different wild specimens sampled from different locations were used for both morphometric and shell thickness analyses in Chapters 3 and 4.

- Cultured *M. edulis* specimens were obtained from a traditional mussel farm on the Scottish west coast in Loch Leven (Figure 2.1b, site 10) as part of a long-term monitoring programme between December 2014 and January 2016. A total of 120 specimens from four batches of mussels (30 individuals each), originating from a natural spatfall, were hand collected at 1, 3, 5, and 7 meters depth along the cultivation rope (batches I, III, V, and VII, Figure 2.1b). These samples were representative of the natural distribution of mussels at the cultivation site.

Cultured mussels, along with wild specimens, were used for morphometric analyses in Chapter 3.

- Archival specimens of *M. edulis* were obtained from museum collections donated by the Royal Belgian Institute of Natural Sciences (RBINS, Brussels, Belgium). Museum collections (both wet and dry collections) consisted of mussel specimens sampled with a near decadal frequency between 1904 and 2016 along the Belgian coastline at 13 locations between Oostende (51°14′16.27″N - 2°55′03.09″E) and Nieuwpoort (51°09′14.14″N - 2°43′23.62″E) (Figure 2.1, site 3). I selected a total of 238 adult individuals with a shell length of 39 - 66 mm, which were collected from the eulittoral zone on stone breakwaters. Archival specimens were used in Chapter 5 for analysis of historical patterns in shell morphometrics and thickness.
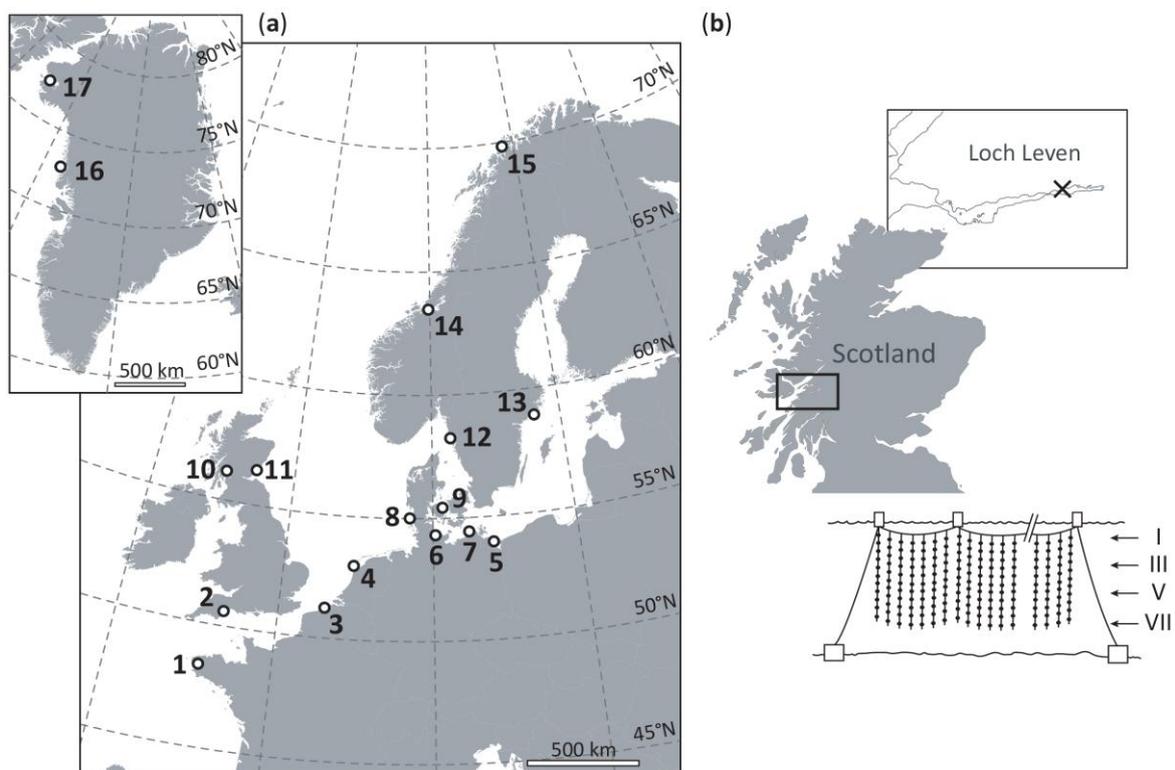
Figure 2.1 *Mytilus* spp. collection sites
(**a**) Locations where *Mytilus* was collected across the Eastern European and Greenlandic coastlines: (1) Brest, France, (2) Exmouth, England, (3) Oostende, Belgium, (4) Texel, Netherlands, (5) Usedom, (6) Kiel, (7) Ahrenshoop, (8) Sylt, all Germany, (9) Kerteminde, Denmark, (10) Tarbet (Kintyre), Scotland, (11) St. Andrews, Scotland, (12) Kristineberg, Sweden, (13) Nynäshamn, Sweden (14) Trondhiem, Norway, (15) Tromsø, Norway, (16) Upernavik, Greenland and (17) Qaanaaq, Greenland. (**b**) Collection site of cultured *Mytilus* specimens obtained from a mussel farm (Glencoe Shellfish Ltd.) in Loch Leven, west Scotland; schematic example of a traditional long-line, with four sampling depths (I, III, V, and VII meters). Maps created with ArcMap 10.5 (ArcGIS software by Esri, http://esri.com), background image courtesy of OpenStreetMap (http://www.openstreetmap.org).

2.2 Thickness analysis

In Chapters 5 and 6 I analysed spatial and temporal patterns of shell deposition in *Mytilus* with focus on the individuals organic and inorganic shell layers: the calcareous prismatic layer, the calcareous nacreous layer, and the organic periostracum. Deposition of each individual shell layer was expressed as a measure of the thickness of the individual layers.

To measure the thickness of each shell layers I embedded left shell valves in polyester resin (Kleer-Set FF, MetPrep, Coventry, U.K.) blocks using silicon moulds. Embedded shell dried 24 hours prior to processing. Specimens were sliced longitudinally along their axis of maximum growth using a diamond saw. Sections were then progressively polished through successive grades of silicon carbide paper (grit size from P800 to P2500) and multiple steps of diamond paste (grading from 9 µm to 1 µm). Polished section were rinsed with mill-Q water and air dried for 24 hours prior to measuring (Figure 2.2).

Photographs of polished sections were acquired with a stereo-microscope (Leica M165 C equipped with a DFC295 HD camera, Leica, Wetzlar, Germany) and shell thickness was measured using the Fiji software (v1.51u) [237] (Figure 2.2). Since larger individuals had undergone evident environmental abrasion or dissolution which had removed the periostracum and prismatic layer closer to the umbo, I estimated the thickness of the whole-shell, prismatic and nacreous layers at the midpoint along the shell cross-section. Periostracum thickness was measured at the posterior edge where the outer periostracum attaches to the external side of the prismatic layer, to estimate the fully formed organic layer that was unaffected by decay or abrasion [185].
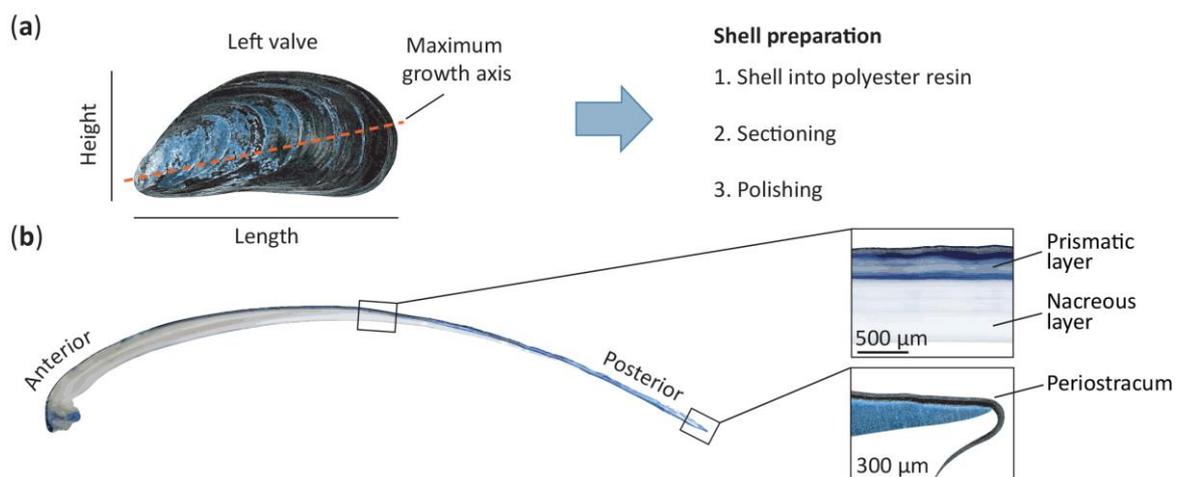
Figure 2.2 Measurement of individual shell layers in *Mytilus*
Schematic representation of steps for shell processing and individual shell layers measurement. (**a**) *Mytilus* left shell valve morphology and section plane (red dashed line), and step for preparation of shell sections. (**b**) Anteroposterior cross-section of shell valve along the axis of maximum growth showing individual mineral and organic shell layers measured for analysis.

2.2.1 Shell dimensions

For all the analysed *Mytilus* specimens, shell dimensions (i.e. shell length, height, and width) were measured using digital Vernier callipers at the nearest 0.01 mm (Figure 2.12b). Shell length was used as a within- population proxy for age [162, 209].

2.3 Organic content analysis

2.3.1 Thermal methods

The observation of the behaviour of a sample of material and the quantitative measurement of its changes on heating can yield useful information on the nature of the material. Thermal analyses have developed out of the scientific study of the changes in the properties of a sample that occur during heating [238, 239]. Some alterations in material properties may be straightforward to measure, such as changes of density or mechanical strength. Specific thermal approaches can provide useful insight into material properties depending on the bonding, molecular structure, and nature of the material. These include thermodynamic properties, such as heat capacity, enthalpy, and entropy as well as structural properties.

When a system (e.g. a material) is heated, heat energy will increase its temperature and determine the change of that system into a more stable state. An example is the thermal decomposition of calcium carbonate ($CaCO_3$, Equation 2.1). Upon heating $CaCO_3$ dissociates to yield calcium oxide (CaO) and carbon dioxide gas (CO_2). If CO_2 is not removed, these two components will recombine on cooling, indicating how the system reaches the most appropriate equilibrium state at a particular temperature.

$$CaCO_{3(solid)} \underset{cool}{\overset{heat}{\rightleftarrows}} CaO_{(solid)} + CO_{2(gas)}$$
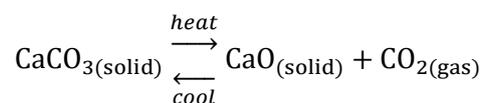
Equation 2.1

Temperature-induced transformations that change the materials in a system will alter one or more of their properties. These changes may be physical, such as melting, crystalline transition or vaporisation, or they may be chemical, involving reactions that alter the molecular structure of the material. The temperature at which a particular event occurs, or the temperature range over which a reaction happens, are characteristic of the nature and history of the sample. These transitions on a molecular level are fundamental, and their study enables the analyst to draw valuable conclusions about the sample of material, its preparation, chemical nature (organic or inorganic), and the likely behaviour during its proposed use.

Thermal methods, as described in the following Sections 2.3.2 - 2.3.5, include a group of techniques in which one (or more) property of a sample is studied while the sample is subjected to a controlled temperature programme [238]. The temperature programme may take many forms: **i**) the sample may be subjected to a constant heating (or cooling) rate, **ii**) it may be held isothermally, **iii**) the sample may be subjected to a modular temperature (e.g. temperature alterations superimposed onto an underlying heating rate), or **iv**) the heating may be controlled by the response of the sample itself. These methods represent useful tools for the identification of the temperature of phase change, the temperature of decomposition, and the products of the reaction of the material analysed. In Chapters 5 and 6, I used thermal methods, including thermogravimetry and derivative thermogravimetry (Sections 2.3.2 - 2.3.5), to quantify the proportion of organic matter within the fibrous prismatic layer of *Mytilus* shells.

2.3.2 Thermogravimetry and derivative thermogravimetry

Thermogravimetric analysis, or thermal gravimetric analysis, (TGA) is defined as a thermal technique in which the mass of a substance is measured as a function of temperature as the sample is subjected to a controlled temperature programme in a controlled atmosphere [238, 239]. TGA results are presented as a plot of the mass (m) of the sample, or the proportion of weight loss (wt%), against temperature (T) or time (t). The mass loss appears as a step as in Figure 2.3a. It should be noted that the shape is sigmoid in nature, and although most mass loss occurs around one temperature (the steepest part of the line), some reaction starts well before the main reaction temperature. Similarly, there is still some residual mass loss well after the main reaction. An alternative presentation of the results is to take the derivative of the original experimental curve to give dm/dT, or rate of mass loss against temperature, and to plot that against, temperature. Alternatively, the derivative may be against time t, giving dm/dt. The
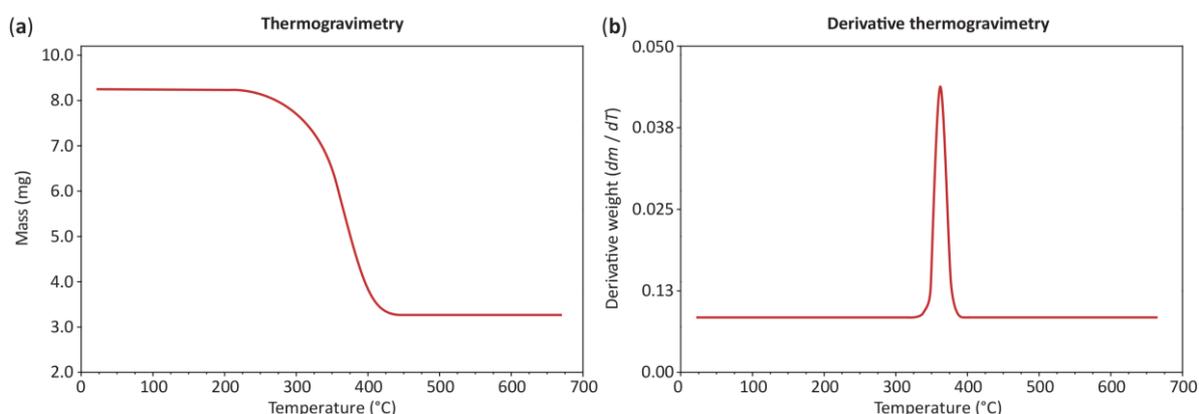
Figure 2.3 Typical thermogravimetry results
(**a**) Thermogravimetric curve representing the weight changes with increasing treatment temperature. (**b**) Derivative thermogravimetric curve representing the derivative of the thermogravimetric curve (dm/dT) and showing the rate of weight loss during heating.

production of such curves is called derivative thermogravimetry (DTG) which represents the variation in rate of sample weight loss during heating (Figure 2.3b). The spread of the reaction over a specific

temperature range is indicated by the width of the peak. The DTG curve is of assistance if there are overlapping reactions, such as double peaks or a shoulder on a main peak. The area under the DTG peak is proportional to the mass loss, so relative mass losses may be compared. The position of the peak indicates where mass loss is fastest and may not be indicative of any characteristic point in the mechanism of the reaction. TGA has many applications in both industrial and biological contexts including the quantification of water within a material, the relative amount of organic and inorganic components, as well as identification of decarboxylation, oxidation, and decomposition processes.

2.3.3 Thermogravimetric analyser

In its essential form, a thermogravimetric analyser consists of **i**) precision balance which supports a sample pans and monitors its mass, **ii**) a furnace that heats (or cools) the sample, and **iii**) a purge gas that controls the furnace environment (Figure 2.4).

The **microbalance** that holds the sample and monitors its weight variation during the experiment. The balance can usually measure sample weights ranging from less than 1 mg up to 10 - 100 mg (Figure 2.4a). The **furnace** surrounds the sample and can work from -160°C up to 1000 - 1600°C. This usually consists in a resistive wire, wound on a ceramic or silica tube, which is mounted in a metal container and is surrounded by insulating material. The furnace is capable of being moved away from the balance case through a sliding support to allow access to the sample (Figure 2.4a). The system of balance plus furnace is called the **thermobalance** and a typical example is shown in detail in Figure 2.4b. The simplest TG experiment would be to heat the sample in static air. However, the sample may react with this atmosphere in oxidising and burning. Therefore, most TGA experiments use an **inert atmosphere** made of sample purge gas, such as nitrogen or argon. This is done so the sample only reacts to temperature during decomposition in an inert atmosphere (pyrolysis). A flowing purge gas is fed over the balance mechanism first, it flows over the sample, and then exits through an exhaust (Figure 2.4b). Sample **crucibles** (or pan) are usually made of platinum that is inert with respect to most gases and inorganic materials, and only melt at 1769°C. They are made of thin platinum to keep the mass low so that they have low heat capacity and follow the furnace temperature without any lag. The temperature in the system is measured by **thermocouples** to monitor system and sample temperature (Figure 2.4b). These consist of two different metals fused into a junction, which produces and electromagnetic field varying with temperature. As well as the thermocouple system for measurement, a separate thermocouple is provided to sense the furnace temperature and is positioned as near to the sample as possible (Figure 2.4b).
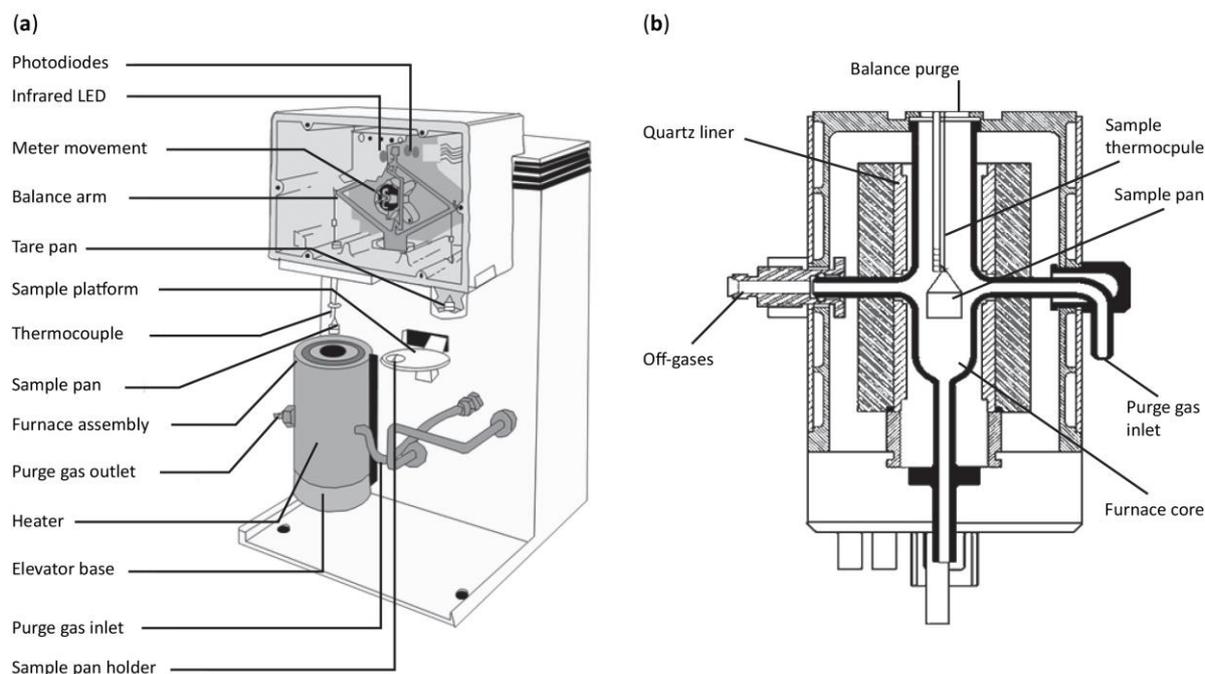
Figure 2.4 Scheme of a typical commercial thermogravimetric analyser
(**a**) Components of a thermogravimetric analyser, TA Q500 series. (**b**) Detailed scheme of the thermobalance (balance system + furnace). Image courtesy of TA Instruments Ltd., New Castle, DE, US.

2.3.4 TGA of blue mussel shells

In Chapter 5 and 6 TGA was used to estimate the weight proportion (wt%) of organic matrix within the prismatic (calcite) layer of *Mytilus* mussels. This method allows the identification of **i**) the proportion of physically adsorbed water within the shell microstructure, **ii**) the decomposition, or onset, temperature and the wt% of organic matter within the shell layer, and **iii**) the decomposition temperature of calcium carbonate [240–244]. The wt% of organic matter within the prismatic layer was estimated as the proportion of weight loss during the thermal treatment between 150°C and 550°C (Figure 2.5).

Although the shell sample may be decomposing at a temperature that is characteristic of the components, the shape of the decomposition curve will be affected by many factors, among these: particle size, material and size of crucibles (controlling the rate of heat flow), the atmosphere used, product gas of the reactions, and the rate of heating. Therefore, to enable reproducibility and recreation of adequate experimental conditions across the different studies, I produced and consistently applied a detailed protocol for the analysis of each sample of shell material.
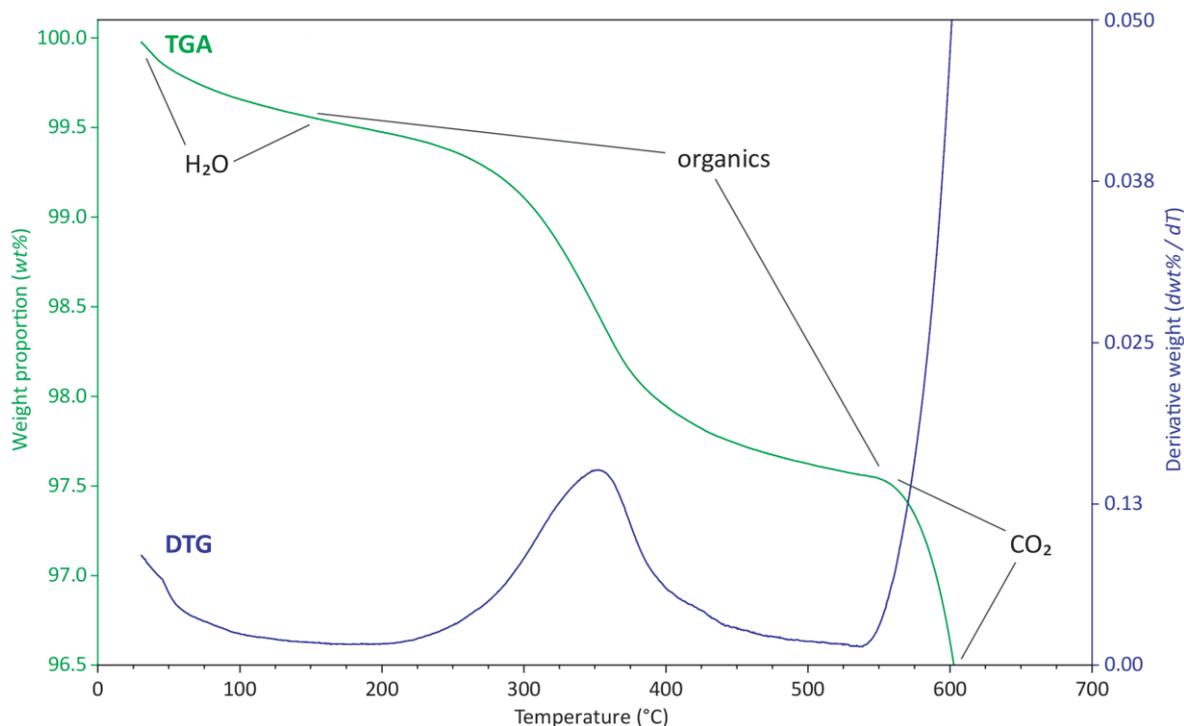
Figure 2.5 Example of thermogravimetric analysis (TGA) and derivative thermogravimetry (DTG) of mussel shell

The TGA curve (green line) represents the weight changes with increasing treatment temperature for the prismatic layer of *Mytilus edulis*. The sample was exposed to a constant heating, from ~25°C to 700°C at a linear rate of 10°C min^{-1}. Three known regions of weight loss with increasing temperature are highlighted [241]: **i**) the evaporation of physically adsorbed water at 30-150°C, **ii**) the degradation of organics at 150-550 °C, and **iii**) the rapid decomposition of $CaCO_3$ into CaO and CO_2 starting at ~550°C. The DTG curve (blue line) represents the derivative of the thermal curve and shows the rate of weight loss during heating. The peak indicates the temperature at which the organic mass loss was fastest.

2.3.5 TGA protocol

Thermogravimetric analyses are reported following the guidelines made by the Committee on Standardisation of the International Confederation for Thermal Analysis and Calorimetry (ICTAC) and appeared in standards as ASTM E 472 (1991) [238, 239].

A. Properties of the sample

o *Source of material and identification*

Shell of wild Atlantic blue mussel (*Mytilus* spp.). Prismatic layer composed of $CaCO_3$ (calcite), variable amount of organics (~1 - 3%) and other components, such as quartz (SiO_2) and magnesium (Mg).

o **Sample preparation**

Shells were cleaned, rinsed with mill-Q water, and air-dried at room temperature for seven days. The periostracum was removed by sanding and tiles (8×5 mm) of outer prismatic layer isolated with a Dremel rotary tool (Dremel 300/395RD MultiPro, Racine, Wisconsin, USA). Samples were cleaned in an ultrasonic bath (Ultrasonic Cleaner CD-4800, Practical Systems Inc., Odessa, FL, USA) with mill-Q water, air-dried and finely ground with an agate mortar. Additional, oven drying (30°C for 24h, convection oven) was performed to remove residual pre-treatment water.

o **Physical properties of samples**

Fine grade powder.

B. Experimental conditions

o **Apparatus used**

Thermogravimetric Analyser TGA Q500, TA instrument (New Castle, DE, USA) Q series was used, located in the Department of Materials Science and Metallurgy, University of Cambridge.

o **Thermal treatment**

Samples were subjected to constant heating from an initial temperature of ~25°C to a final temperature of 700°C at a linear heating rate of 10°C min^{-1}.

o **Sample atmosphere**

Dynamic (flowing) atmosphere of nitrogen, "white spot" (~0.01% trace of oxygen). Flow rates of 40 ml min^{-1} for the balance and 60ml min^{-1} for the sample were used.

o **Sample holder**

Platinum crucible, cylindrical: diameter 10 mm and height 1.5 mm. Sample was tipped and spread (not tamped) to cover the bottom of the crucible.

o **Sample mass**

10 mg of powder for each sample were weighted on a separate micro-balance (Ultramicro 4504 MP8, Sartorius, Göttingen; readability 0.1 µg).

C. Data acquisition and manipulation methods

o **Software version**

Universal Analysis 2000, version 4.5A, TA instrument (New Castle, DE, USA).

2.4 Morphometrics analysis

The "pipeline" developed for the shape analysis of blue mussel shell outlines as described in this section has been published in:

Comparing the anatomical features of an organism and their shape has been a central element of biology for centuries [245, 246]. A variety of biological processes produce differences in shape between individuals and their parts, such as ontogenetic development, growth, adaption to local habitats, disease, or long-term evolutionary diversification [247, 248]. Differences in shape may signal different functional roles played by the same parts, different responses to the same selective pressure (or differences in selective pressure themselves), as well as heterogeneous growth processes [108, 159, 247–251]. In systematics, analysing shape variation among individuals helps to quantify and describe differences between taxa or populations [252]. Moreover, in many biological fields, an understanding of the variation in organismal traits can bring deep insights into the underlying mechanisms of change and the physical constraints involved.

During the early 20th century, biology began the transition from a descriptive field to a quantitative science, and the analysis of morphology saw a similar "quantification revolution" [246, 253]. Morphological studies included quantitative data for one or multiple measurable traits that were compared among groups. The development of statistical methods, such as the correlation coefficient [254], analysis of variance [255], and principal components analysis [256, 257], further advanced these quantitative approaches. By the mid-twentieth century, quantitative description of morphological shape was combined with newly developed statistical methods, describing patterns of shape variation within and among groups, and the modern field of morphometrics began [246, 258].

Morphometrics is the study of variation and covariation of shape and size with other variables [258, 259]. Hereafter, I will consider shape as defined by Kendall [260] and Small [261] that it is "the total of all information (or geometric properties) of an object invariant under translation, rotation, and isotropic re-scaling". Until recently, shapes and their variation were estimated using *ad hoc* distances, angles, or proportions. Shape analysis was improved during the 20th century with the development of methods aiming to describe shapes objectively and quantitatively [262, 263]. This "traditional morphometrics" approach consisted in the application of multivariate statistical analyses to set of quantitative variables [264–266]. In the late 1980's and early 1990's, however, a shift occurred in the
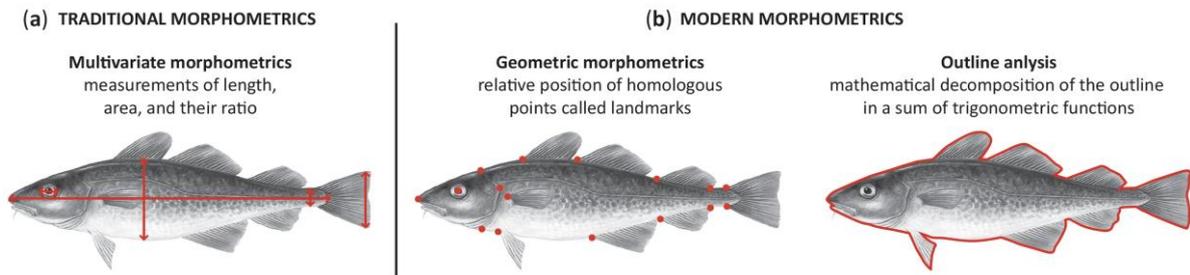
Figure 2.6 Traditional and modern morphometrics approaches
Simplified representation of how the shape of a biological object (Atlantic cod, *Gadus morhua*) can be described through (**a**) traditional and (**b**) modern morphometrics approaches. Cod drawings courtesy of Scandinavian Fishing Year Book.

way morphological data were acquired and analysed. This shift emphasised methods capturing the geometry of the morphological structures of interest and preserving this information throughout the analyses. In 1993, this new morphometric approach was defined as "geometric morphometrics" and Rohlf & Marcus [263] suggested that this paradigm shift signalled a "revolution in morphometrics" [246, 267]. Since then, an increasing understanding of the theoretical underpinnings of geometric morphometric methodology has led to much progress in this field and its applications.

In this thesis, geometric morphometrics analyses of outlines were performed for analysing shape variations of *Mytilus* shells with environmental factors across different geographic scales and over historical times. In Sections 2.4.1 - 2.4.10, I describe traditional and modern morphometrics methods, the mathematical background of Fourier-based approaches and their advantages, as well as the procedures used for analysing *Mytilus* shell shapes variations in Chapters 4 and 6.

2.4.1 Traditional morphometrics

A quantitative framework for shape analysis was introduced in the 1960's using the full set of multivariate statistical tools to describe within- and among-individuals patterns of shape variation [263]. This approach, subsequently called "traditional morphometrics" [264, 265] or "multivariate morphometrics" [266], consisted of applying multivariate statistical analyses to sets of morphological variables (Figure 2.6a). For instance, linear distance measurements (i.e. length, width, and height) and areas, as well as counts, scalar indices, and angles were used. These approaches allowed biometricians to quantify covariation between morphological measurements and to assess patterns of variation between and among samples using univariate or multivariate approaches.

Many studies started also investigating allometry of traits in organisms [250, 251]. Because linear distances are usually highly correlated with size [268], much effort was spent developing methods for size correction, so that size-independent variables could be used and patterns of shape variation properly

Figure 2.7 Graphic explanation of the Thompson's deformation grids
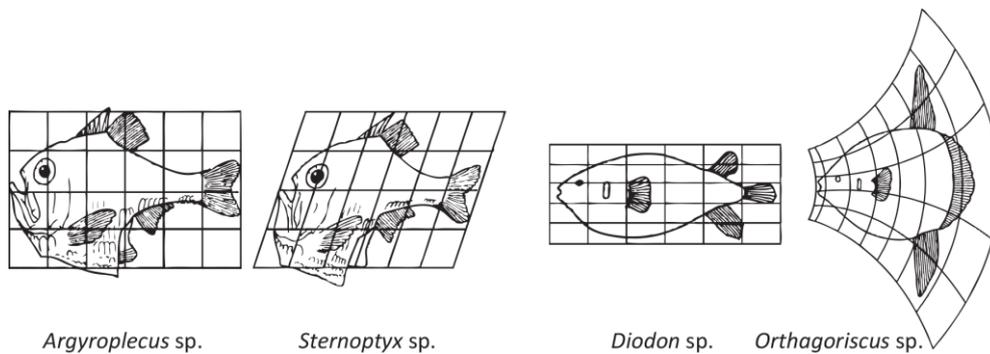
Figure 2.7 Graphic explanation of the Thompson's deformation grids
From a morphological point of view, these fishes seem very different, but little changes in the growth direction (left) or in the development of the caudal part (right), can lead to dramatic changes in the final form. This represents the key concepts of the Thompson's shape analysis approach [245]: **i**) some changes in the developmental process of living organisms, while minor, can lead to dramatic morphological changes, and **ii**) physical constraints such as growing mechanisms are of primary importance in the final form of an organism. Redrawn after Thompson [245] and Bonhomme [274].

descried [269–272]. However, several difficulties remained. For instance, there is little agreement on which methods of size correction should be used, since different methods usually yield different results. Some studies have analysed ratios, but these pose serious statistical problems [248, 273]. The homology of linear measurements is also difficult to assess. Moreover, it is not usually possible to generate a graphical representation of shape from linear distances because the geometric relationships among the variables are not preserved.

Thomson [245], in his seminal book *On Growth and Form*, was the first to offer new ways of understanding shape variations in what could be called today a "modern morphometric" approach. The development of the concept of deformation grid changed the perception of shape variation in terms of mathematical transformations (Figure 2.7). However, it was only at the end of the 20th century that computerised data acquisition and treatment arose synchronously with the development of the mathematical framework for constructing these grids (e.g. Fourier and thin plate spline analyses). These techniques of "modern morphometrics" revolutionised the historical scope of this field [246, 263].

2.4.2 Modern morphometrics

"Modern morphometrics", or "geometric morphometrics", considers shape as a whole and encompasses an array of techniques capturing all the geometrical relationships of morphological structures and preserving their information throughout the analysis [246, 263]. The development of rigorous statistical theory for shape analysis [260] made it possible to combine the use of multivariate statistics with methods for the direct visualisation of biological forms. The two main approaches in use are the study of landmark configurations [259] and outline analysis [275, 276] (Figure 2.6b). Both of them preserve

the geometrical information (i.e. relative position between landmark or outline's points) [259, 277] and, therefore, allow shape reconstruction from the numerical signature providing great advantages compared to more traditional approaches [246, 274, 278].

Landmark-based methods analyse the relative position (two- or three-dimensional coordinates) of a set of biologically definable points, called landmarks, through a matrix of their pair-wise Euclidean distances [259, 279, 280]. The selected landmarks should be structurally similar (i.e. homologous points) between individuals. Outline analysis considers outlines as a whole. Here, an outline is defined as the closed polygon formed by the (*x*; *y*) coordinates of the pixels defining it. Since outlines can be described as periodic functions, Fourier-based methods are typically used to extract their geometric information. This approach fits Fourier series separately on the *x* and *y* coordinates on an outline, projected on a plane [258, 274, 276].

2.4.3 Mathematical background of Fourier-based outline analyses

Different from landmark configurations analysis, examining the relative position of landmarks after superimposition, approaches for analysing outlines estimate parameters of functions by fitting Fourier series to shape descriptors [280]. Fourier-based approaches are powerful methods to extract geometric information. They have been implemented on the concept of Fourier series: to decompose a periodic function into a sum of more simple trigonometric functions, such as sine and cosine [258, 274]. These simple functions have frequencies that are integer multiples, therefore they are harmonics of one another. Low-frequency harmonics provide approximation of the coarse-scale trends in the original periodic function, while high-frequency harmonics fit its fine-scale variations.

Among other applications, Fourier series can be used in morphometrics to decompose outlines (closed curves) since these can be described as periodic functions[1]. Specifically, the outline geometry can be described as a function of **i**) the distance of each point on the outline to the geometric centre, or centroid, of the shape, **ii**) the variation of the tangent angle for any point, or **iii**) the difference between the abscissa/ordinate of the first and all successive point along the outline. These three different methods available are called "radius variation", "tangent angle", and "elliptical analysis", respectively [281] (Figure 2.8). With these methods, one or more periodic functions are then obtained for each outline, and these can be described through Fourier series. The Fourier transform then decomposes them into the harmonic sum of trigonometric functions (Figure 2.9), weighted with coefficients called harmonic coefficients. These are usually normalised to remove homothetic, translational, and rotational

---

[1] Note that the principle of Fourier series applies to continuous functions. Since the shape of an outline is based on a finite number of points, typically coordinates on a plane (or a space), a discrete equivalent of Fourier series is used in morphometrics [423].
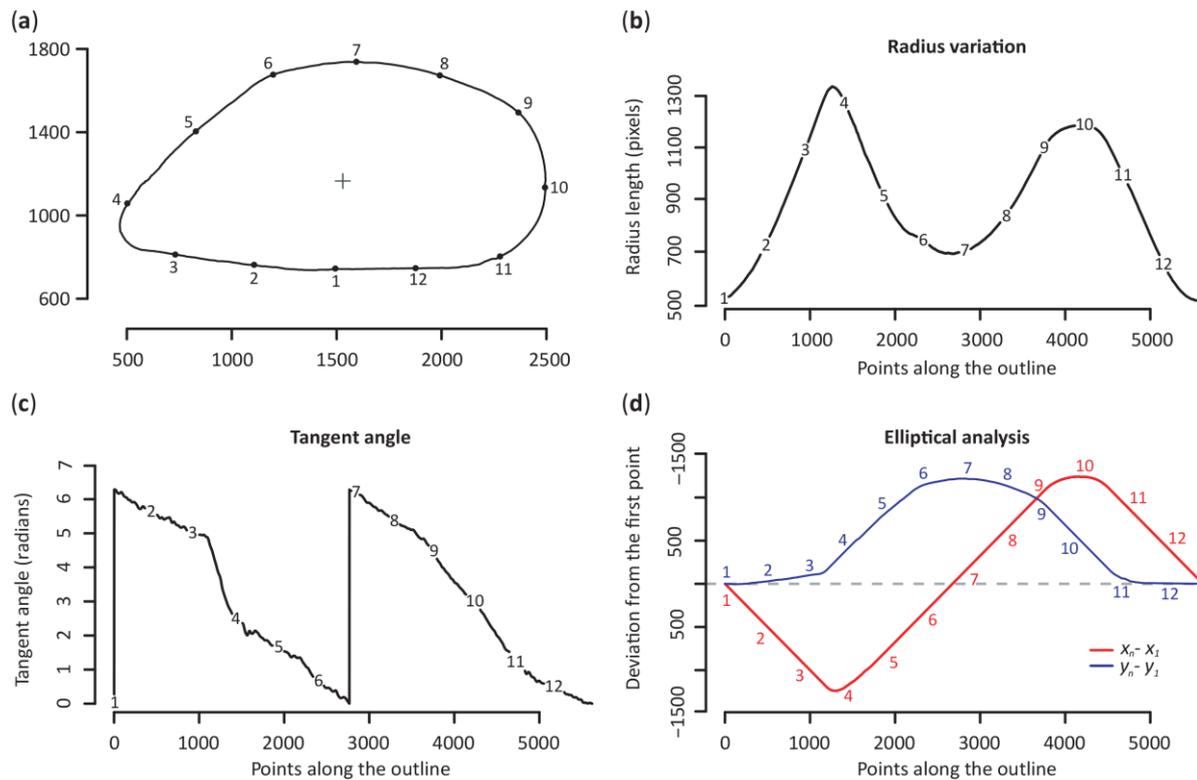
Figure 2.8 Fourier-based approaches in outline analysis
Description of the outline as one, or multiple, periodic functions using Fourier-based methods.
(**a**) Twelve equidistant points have been sampled, starting from the centre of the ventral side,
clockwise and along the outline of a blue mussel shell, *Mytilus edulis* (lateral view, left valve).
(**b**) Radius variation illustrates the variation of the length of the radius, here considered as the
distance between the centroid of the shape (+ sign) and the point along the outline. (**c**) Tangent
angle illustrates the variation of the tangent angle along the outline. (**d**) Elliptical analysis
shows the two curves corresponding to $x_n - x_1$ (red line) and $y_n - y_1$ (blue line).

differences between shapes [278]. Two or four coefficients, depending on the approach used, are then
calculated from each harmonic and can be considered as quantitative variables. The geometrical
information (shape) contained in the outlines is thus quantified by harmonic coefficients and they can
be analysed with classical multivariate tools.

2.4.3.1   *Fourier radius variation*

For a closed outline, the radius r, taken as the distance from the outline centroid and a given point of
the outline, can be expressed as a periodic function of the angle θ [282] (Figure 2.8b). Harmonics from
0 to k approximate the function:

$$r(\theta) = \frac{1}{2}a_0 + \sum_{n=1}^{k} a_n \cos(w_n\theta) + b_n \sin(w_n\theta)$$

Equation 2.2
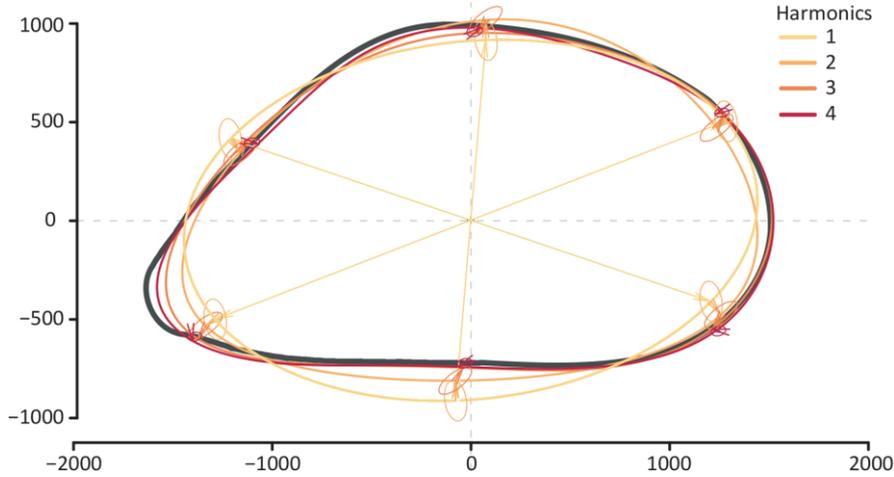
Figure 2.9 Decomposition of an outline into the harmonic sum of trigonometric functions

Example of Fourier decomposition of an outline projected on the plane. The first harmonic defines the best-fitting ellipse. Each harmonic, or ellipse, is defined by the *x* and *y* projections of the sum of a cosine and sine curve with different amplitudes and harmonically related frequencies. Here, four harmonics are shown at 6 locations on the original outline.

where *w* refers to the pulse and *p* is the number of sampled points along the outline (equivalent to the number of sampled radii). The harmonic coefficients a_n and b_n are extracted for every individual shape and are estimated as follows:

$$a_n = \frac{2}{p}\sum_{i=1}^{p} r_i \cos(n\theta_i) \qquad\qquad b_n = \frac{2}{p}\sum_{i=1}^{p} r_i \sin(n\theta_i)$$

Equation 2.3

2.4.3.2   Fourier tangent angle

Radius variation may fail to fit some complex outlines especially when they present convexities and concavities. The Fourier tangent angle [282] fits the cumulative change in the angle of a tangent vector ($\phi(t)$), as a function of the cumulative curvilinear distance *t* along the outline (Figure 2.8c). Given a closed outline, previously scaled to 2π, $\phi(t)$ can be expressed as $\phi(t) = \theta(t) - \theta(0) - t$ where *t* is the distance along the outline, $\phi(t)$ is the angle of the tangent vector at *t*, and $\phi(0)$ the angle of the tangent vector taken for the first point. Two coefficients per harmonic can be estimated as follows:

$$a_n = \frac{2}{p}\sum_{i=1}^{p} \phi(t) \cos(n\theta_i) \qquad\qquad b_n = \frac{2}{p}\sum_{i=1}^{p} \phi(t) \sin(n\theta_i)$$

Equation 2.4

2.4.3.3   Elliptic Fourier analysis

Elliptic Fourier analysis (EFA) [275, 276] fits separately the *x* and *y* coordinates of the outlines projected on the Cartesian plane, as a function of the curvilinear abscissa. This method has become very popular since it has great advantages over the other Fourier-based approaches [274, 281]: equally-spaced points

are not required, virtually any outline can be fitted, and the coefficients can be made independent of homothetic, rotational and translational differences between outlines [246, 258, 274, 278, 281, 283, 284]. The original signal $f(t)$ (both $x(t)$ and $y(t)$) can be expressed as follows:

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{+\infty} \left[a_n \cos\left(n\frac{2\pi}{T}t\right) + b_n \sin\left(n\frac{2\pi}{T}t\right) \right]$$

Equation 2.5

where T is the perimeter of a given closed outline, here considered as the period of the signal and t the curvilinear abscissa that varies from 0 to T. The harmonic coefficients a_n and b_n can be expressed as follows:

$$a_n = \frac{2\pi}{T} + \int_0^T f(t) \cos\left(n\frac{2\pi}{T}t\right) dt \qquad b_n = \frac{2\pi}{T} + \int_0^T f(t) \sin\left(n\frac{2\pi}{T}t\right) dt$$

Equation 2.6

2.4.4 EFA of mussel shell outlines

Analysis of spatial and temporal patterns of shell morphology in *Mytilus* spp. were carried out through a geometric morphometrics approach (Figure 2.10a-c). I used an EFA of outline to examine shell shape variation both within and between mussel populations from different geographic areas (Chapter 4) and to describe between-individual shell shape variations over the last century (Chapter 6). Outline processing and EFA were carried out using the package Momocs (v1.2.9) [274] in the software R (v3.4.1) [285]; for details on the other packages used for graphing see Supporting Table A.1, in Appendix A,.

2.4.5 Outline extraction

Digital images of orthogonal lateral and ventral shell views of left valves were acquired with a high-resolution digital camera (Nikon D3300, fitted with Sigma 105mm f/28 EX DG Macro lens). Prior to outline digitisation, photographs were processed using image analysis software (©Adobe Photoshop), centred and consistently aligned. Only the shapes of intact specimens were selected for analysis. Shell photographs were then converted into black masks on a white background (8-bit, grey-scale mode images with no level of compression). Outlines were isolated, using an algorithm implemented by Claude [258], converted into a list of (x; y) pixel coordinates sampled along the outline, and used as input data for the following morphometrics analyses (Figure 2.10c). Outlines of lateral and ventral shell views (Figure 2.10b) were then processed independently and later re-combined for analysis.

(a) Morphology of a mussel shell
Dorsal margin
Ligament area
Growth lines
Lateral view
Posterior margin
Umbo
Left valve
Ventral margin
Ventral view
Right valve

(b) Shell dimensions
Width
Height
Length
Shell outline

(c) Geometric morphometrics: elliptic Fourier analysis (EFA) of outlines

Step 1: digitisation of orthogonal shell views

Step 2: alignment and conversion to black mask

Step 3: outline isolation

Step 4: outline normalisation (smoothing, centreing, scaling)

Step 5: pseudo-landmark configurations (1000)

Step 6: Procrustes superimpostionon and normalisation of starting points

Step 7: EFA and calibration

Step 8: calculation of harmonic coefficients (4 × harmonic × outline)

Step 9a: multivariate analysis (PCA)

Step 9b: mean shape and TPS analysis

Step 10a: multivariate analysis (MANOVA, LDA)

Step 10b: modelling (GAMMs)

Figure 2.10 Blue mussel shell shape morphology and shape analysis
(**a**) Terminology used to describe the morphology of *Mytilus* shells. (**b**) Main shell linear dimensions. (**c**) Summary of the EFA of blue mussel shell outlines with details and graphical steps on the shape acquisition, processing, and analysis of both lateral and ventral shell views.

2.4.6 Outline normalisation

The concept of normalisation is central in EFA and there are two ways of normalising outlines. The first, and by far the most used, is to use a "numerical" alignment directly on the matrix of coefficients. The coefficients of the first harmonic are "consumed" by this process, but higher-frequency harmonics are normalised for size and rotation [258]. This is sometimes referred as "using the first harmonic/ellipse", since the first harmonic, which defines the best-fitting ellipse in the plane, is removed during this process. However, biases can be introduced in the numerical adjusted if shapes are prone to bad alignment (i.e. usually rounded or elliptic outlines with a bilateral symmetry). This can determine a poor numerical alignment of the matrix of coefficients, resulting in not homologues elliptic Fourier descriptors [278, 281]. When this happens, a second method of normalisation should be used that involves an alignment through geometric operations. For my analyses, an outline normalisation through geometric operations was directly performed on the list of (*x*; *y*) pixel coordinates prior to calculation of elliptic Fourier transforms. This *a priori* normalisation was required to avoid bias due to bad numerical alignment of mussel shell outlines (Figure 2.10c). Outlines were first smoothed in order to remove any noise introduced during the digitisation process, centred, and coordinates were re-scaled by their centroid size (Figure 2.11a, b). Equal numbers of points were sampled along each outline (1000 pseudo-landmark) (Figure 2.11c). Point configurations were then aligned through a Procrustes superimposition [258, 259] and starting point normalised (made homologous) (Figure 2.11d). Shapes invariant to shell size, rotation, and outline position were obtained and an EFA was computed on the resulting coordinates.
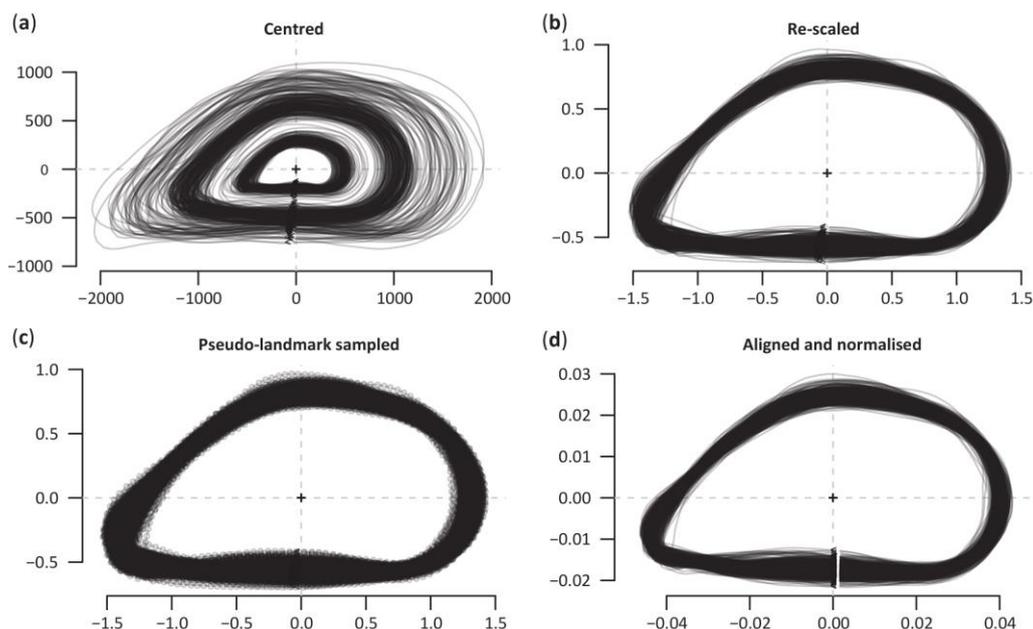
Figure 2.11 Outline normalisation trough geometric operations
(**a**) Outlines are centred, (**b**) re-scale, (**c**) pseudo-landmark are sampled, and (**d**) point configurations are aligned through Procrustes superimposition and starting point aligned.
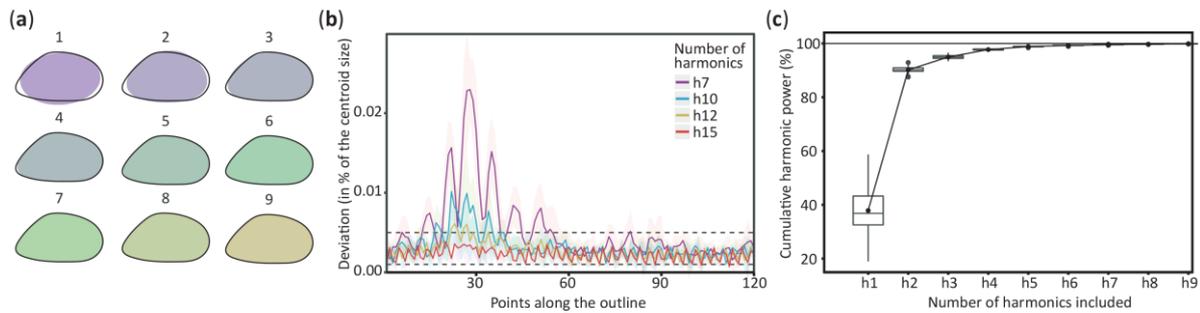
Figure 2.12 EFA calibration methods
(**a**) Mean shape reconstructed (coloured background) for an increasing number of harmonics. (**b**) Deviation between the best possible fit and a given number of harmonics (7, 10, 12, and 15) for each point (continuous line) and standard error (shaded area). The *y*-axis represents the deviation for every sampled point (on the *x*-axis). Dotted lines represent average deviation of 0.001 and 0.005 respectively. (**c**) Cumulative harmonic Fourier power for mussel outlines. The first 7 harmonics gather nearly 100% of the harmonic power.

2.4.7 Calibration of outline analysis

The advantage of the Fourier-based approaches is that they can fit virtually any shapes, however the ratio signal/noise can be quite small for higher levels harmonics. This means that the details described by the high frequency harmonics may be due to many things, such as digitisation errors, and not to real difference among shapes. Therefore, a critical point in Fourier-based approaches is determining the right number of harmonics to retain. So far, there is no objective criterion, since it depends uniquely of the scope of the study and the level of details required. However, different approaches are available (Figure 2.12). Before carrying out a Fourier analysis, I assessed the most appropriate number of harmonics to retain through inspection of **i**) the outline reconstruction efficiency, **ii**) the deviation from optimal fit, and **iii**) the spectrum of harmonic Fourier power.

2.4.7.1 *Shape reconstruction*

This method provides a qualitative estimation of the number of harmonics to be used. The approach consists in comparing the reconstructed mean shapes for a given range of harmonics (Figure 2.12a). The number of harmonics providing a satisfying, or almost perfect, reconstruction of the outline can be selected [258, 274].

2.4.7.2 *Deviation*

The idea of the deviation approach is to define the best possible fit for a given number of sampled points along the outline and then calculate the deviation, in terms of Euclidean distance, between fits with a different range of harmonics and the best fit for every sampled point (Figure 2.12b). One can, for instance, chose the minimal number of harmonics that lead to an average deviation of 1 pixel, or a distance expressed as a proportion of the centroid size [258, 274].

2.4.7.3 Harmonic power

The number of harmonics is also estimated after examining the spectrum of harmonic Fourier power. This is considered as a measure of the shape information explained by a harmonic function. Harmonic power is proportional to the harmonic amplitude and can be considered as a measure of shape information. As the harmonic rank increases the power decreases, adding less information (Figure 2.12c). The number of harmonics to select can be evaluated so that their cumulative power gather 99% of the total cumulative harmonic power [283]. The power for a give harmonic is calculated as follows:

$$Power_n = \frac{a_n^2 + b_n^2 + c_n^2 + d_n^2}{2}$$

Equation 2.7

where a_n, b_n, c_n, and d_n are the harmonic coefficients for the n harmonic.

2.4.8 Computing elliptic Fourier coefficients

Once the right number of harmonics was determined, an EFA was then computed on the resulting list of coordinates to extract the geometric information contained in the outlines. The Fourier transform decomposes the outlines into the harmonic sum of trigonometric functions, weighted by harmonic coefficients (Figure 2.9). Four coefficients per harmonic were extracted for each outline (a matrix of size [$4 \times n$ of harmonics $\times n$ of outlines]) and used as variables quantifying the geometrical information [258, 281]. So, the shape information contained in the outlines was quantified and analysed with classical multivariate tools.

2.4.9 Thin plate splines

Deformation grids such as those proposed by Thompson [245] can be obtained using thin plate splines (TPS) mathematical formalisation. The notion of TPS has been borrowed from mechanics and involves the graphical representation of the forces applied for bending of a thin sheet of metal [259, 274]. The purpose of TPS is to compare shapes and visualise the outline deformation required to pass from a shape to another. For example, the deformations required to pass from the mean shape to the shapes at extremes of the factorial space can be calculated and displayed through different methods, including deformation grids, iso-deformation lines/fields, and vector fields.

2.4.10   Multivariate analyses for shape variables

The calculated harmonic coefficients were considered as quantitative variable and were analysed with classical multivariate frameworks, such as principal component analysis, linear discriminant analysis, and multivariate analysis of variance.

2.4.10.1 Principal component analysis

Geometric shape variables are neither biologically nor statistically independent. For example, the harmonics calculated from the Fourier transform and the extracted harmonic coefficient are not statistically independent and are highly collinear since they describe variations in overlapping regions of an organisms or structure. Moreover, morphometric variables are expected to be correlated because they describe features of the organism that are functionally, developmentally, or genetically linked. Therefore, their patterns of variation and covariation are often complex and difficult to interpret [248]. The purpose of the principal component analysis (PCA) [256, 257] is to simplify those patterns and to make them easier to interpret by replacing the original variables with new ones (principal components, PCs) that are linear combinations of the original variables and independent of each other. The PCA constructs variables that can be used to examine **variations among individuals** within a sample. The produced PCs are a function of the covariances among the original variables, and these can provide insights into the identity of the casual factors underlying those covariances. A useful aspect of the PCA is that most of the variation in the sample usually can be described with only a few PCs, simplifying modelling and the presentation of the results. An indirect aspect of the PCA is also that it summarises the description of the differences among individuals, making it easier to identify clusters of individuals across the morphospace, the theoretical shapes drawn on the factorial plane, than it would be by simply plotting the original morphological variables (i.e. harmonic coefficients).

For my analyses, a PCA with a singular value decomposition method was carried out on the matrix of coefficients to explore the among-individual morphological variation and define variables capturing most of the shape variance. Since high-order harmonics contributed less than low-order harmonics, the PCA was performed without re-scaling to preserve this differential contribution to the shape variability [252, 258]. Each calculated PC captured an independent and synthetic aspect of shape variation. Reconstruction of the morphospace and extreme morphologies along each PC helped to interpret the contribution of individual PCs to the total shape variation. The top PCs capturing most of the shell shape variance were considered as new shape variables and used for statistical modelling.

2.4.10.2 Linear discriminant analysis

The purpose of the linear discriminant analysis (LDA) [286] is to simplify the description of **differences among groups**. A group is a set of individuals (a class) defined as sharing a particular state of a

discontinuous trait. For examples, "sex", "species", and supra-specific categories are examples of discontinuous traits that have two or multiple classes. To be analysed with a LDA the groups must be mutually exclusive, meaning that they cannot comprise nested or intersecting sets. The groups usually differ by a categorical variable that is called "qualitative trait" or "grouping variable". These names refer to the fact that the states of the trait do not have intrinsic numerical values or an inherent order, but they can be used to sort individuals into groups or categories. Therefore, if individuals in a study can be sorted into mutually exclusive sets, LDA can be used to describe the differences among those sets. However, LDA cannot be used to test the statistical significance of the differences among sets; for that, a multivariate analysis of variance is needed (see Section 2.4.10.3).

There are many similarities between LDA and PCA. In fact, LDA constructs a new coordinate system (the linear discriminants, LDs) and determines the score on those axes for all individuals in the study. The LDs are also linear combinations of the original variables and are considered to be mutually orthogonal. However, whereas the PCA is used to describe differences among individuals, the LDA is used to describe difference among group means. In fact, the LDA does not simply rotate the original coordinate systems to the axes that maximise the group differences (if it did, it would be exactly equivalent to a PCA on the group mean). In addition, the optimisation process in LDA involves re-scaling such that the new axes are scaled differently from the original axes and scaled differently from each other. Consequently, distances on the LD space can be quite different from distances in the original coordinate space. As a result, LDA finds the axes that optimise between-group differences relative to within-group variation [248]. Therefore, a LD represents the direction in which groups are most effectively discriminated, which is not necessarily the direction in which individuals are most different.

In Chapter 4, LDA was performed on the shape variables (PCs calculated from harmonic coefficients) to identify the linear combination of shell shape features that were able to discriminate between *Mytilus* species.

2.4.10.3 Multivariate analysis of variance

There are many situations where we record more than one response variable from each sampling or experimental unit and where these units occur in treatment groups. This is usually the case in morphometrics analyses where multiple morphological variables (i.e. harmonic coefficients or PCs) describe altogether difference among individuals or groups on the morphospace. With multiple response variables, we might be more interested in whether there are significant group differences on all the response variables considered simultaneously. This is the aim of multivariate analysis of variance (MANOVA), the analogue of univariate analysis of variance (ANOVA) when we have multiple response variables for each experimental or sampling unit. In the simplest MANOVA, we have a multivariate dependant variable (i.e. shape coefficients or PCs) and one or more independent categorical variables. The hypothesis tested is about group effects on a combination of the response variables, and

instead of comparing group means on a single variable, we now compare group centroids for two or more variables. In my analyses, MANOVAs were performed on the new shape variable (PCs) to test for significant differences between subsets of shapes, or for significant effect of location of origin or year of collection on shape variance.

Chapter Three

3   Statistical analysis

The statistical analysis of ecological data often presents unique challenges. Ecology is a fast moving field that in the past three decades has seen a rapid development of statistical techniques to analyse increasingly complex datasets [287–289]. Ecological datasets may be characterised by outliers, correlation between multiple covariates, non-linear relationships among variables, many zero observations, count or proportional observations, and a variety of dependency structures in the data, such as spatial and temporal dependency structures. Given this complexity, the ability to understand the data and select appropriate statistical models, verify all their assumptions, check that the model employed are appropriate, and interpret their output correctly is crucial for making rigorous inferences. Moreover, a clear presentation of methods and analytical approaches employed is key for reproducibility of the analyses.

Here, I provide a detailed explanation of the statistical approaches used in Chapters 4, 5 and 6, for modelling variations in *Mytilus* shell deposition, composition, and morphology across space and time, and for quantifying the effects of key biotic and abiotic descriptors on the variation of shell characteristics. In Sections (3.1 - 3.7), I describe **i**) the protocol used to conduct data exploration prior to analysis, **ii**) the theoretical basis of the main statistical models used (e.g. mixed and additive models), **iii**) the approaches used for model development and selection, **iv**) details on model validation and interpretation of model results. All data processing, exploration and analysis were carried out with the free software environment for statistical computing R (v3.4.1) [285]; details on the packages used are given in Supporting Table A.1.

3.1   Data exploration

As already stated, in the last decades there has been an enormous expansion of statistical tools available to ecologists. Among these techniques, we have generalised linear (mixed) models, GL(M)Ms, [289, 290], generalised additive (mixed) models, GA(M)Ms, [104, 291, 292], classification trees, neural networks [293], structural equation modelling, and multivariate analysis with all its main methods, such as PCA, LDA, various time series and spatial techniques. Although many of these methods have been available for some time, the development of freely available software, such as R [285], makes it possible

to apply routinely sophisticated statistical methods on any type of ecological data. However, less attention is usually paid to the vital steps that should precede the application of these methods [294].

All statistical techniques have in common the same problem, which is that if the input data are not properly formatted, violating the underlying assumptions of the method used, the output results are invalidated leading to wrong conclusions [287, 294]. In some methods, for example, a single outlier may drive the results. Heterogeneity of variances may cause serious problems in linear regression and analysis of variance [295], and with certain multivariate methods [296]. When determining which covariates are driving a system, collinearity (correlation between covariates) is the most problematic aspect, since increases type II errors (i.e. failure to reject the null hypothesis when it is untrue) [297]. In univariate analysis, techniques like generalised linear modelling (GLM) for count data, zero inflation of the response variable causes biased parameter estimates [298]. Because of these potential pitfalls, avoiding a false covariate effect (type I error), wrongly dismissing a model with a meaningful covariate (type II error), or producing results determined by only a few outliers should be avoided. Some violations have little impact on the results; yet others increase type I or type II errors, potentially resulting in wrong ecological conclusions. Most of these violations can be avoided by applying a detailed data exploration before any statistical analysis.

3.1.1 Six-steps exploration protocol

In this section, I provide the 6-step protocol used for data exploration, and discuss the methods and tools used to identify potential problems prior to statistical analyses (Table 3.1) following from the approach of Zuur *et al.* [287, 294, 299]. In this thesis I mainly focused on the use of graphical tools [300, 301], rather than tests for normality or homogeneity. In fact, the statistical literature warns against certain tests, which are influenced by sample size and have by their own assumptions, and advocates graphical tools [302, 303].

Table 3.1 Six-steps data exploration protocol
Details on the data exploration steps, variables (dependant, *Y*, or independent, *X*) interested, and main graphical methods used.

Data exploration steps	Variable	Main methods
Outliers	*Y & X*	Boxplot, Cleveland dotplots
Normality	*Y*	Histogram, QQ-plot
Homogeneity of variance	*Y*	Conditional boxplots
Collinearity	*X*	VIF, scatterplots, correlation coefficients
Dependence structure	*Y*	ACF, pACF, variogram
Relationships and Interactions	*Y & X*	Pairwise scatterplot, Coplot

3.1.1.1 Step 1: Outliers in Y and X

In some statistical techniques the results can be dominated by outliers, while others treat them like any other value. An outlier is defined as an observation that has a relatively large or small value compared to most of data. A graphical tool used for outlier detection is the boxplot. It represents the median and spread of a univariate variable along with the 25% and 75% quartiles defining the edges of a box. Lines are drawn from both the edges of the box up to 1.5 times the interquartile distance. Any points beyond these lines are usually (but not always correctly) considered as outliers. Another useful tool I used to visualise outliers (and homogeneity) was the Cleveland dotplot [304]. In this graph, the row number of an observation (as in the original spreadsheet) is plotted against the observation value. Any isolated points to the right- or left-hand side of the main "cluster" of points are considerably larger, or smaller, than most of the observations. These require further investigation checking for potential errors and assessing whether the observed values are reasonable. Different approaches were used to treat observations with a considerable larger, or smaller, value that were found in the response variable and covariates. For the latter, if extreme values were likely measurement errors that could not be measured again and the sample was size large enough, they were dropped because their presence would have dominated the analyses [294]. When this was not reasonable I transformed the explanatory variable. For outliers found in the response variable transforming the data was an option, but as the response was of primary interest, I preferred to work with the original data and to choose a statistical method that uses a probability distribution that allows greater variation for large mean values (e.g. gamma and Poisson for continuous and count data, respectively).

3.1.1.2 Step 2: Normal distribution of Y

When using a wide range of statistical approaches, it is important to know whether the statistical technique to be used assumes normality, and what exactly is assumed to be normally distributed. For example, PCA does not require normality [305]. Linear regression assumes normality, but is reasonably robust against it violation [295]. However, LDA assumes normality of the observations of a particular variable within each group [296]. Several authors argue that violation of normality in not a serious problem, or that it is not even needed provided the sample size is large enough [306]. For instance, Läärä [307] enumerates several reasons for not testing for normality, including: **i**) most statistical techniques are robust against violation of normality, **ii**) for larger data sets the central limit theory implies approximate normality, **iii**) for small samples the power of the normality tests is low, and **iv**) for larger datasets normality tests are sensitive to small deviations (contradicting the central limit theory). Normality at each *X* was checked by making histograms and QQ-plots of all observations at each *X* value (by category, if categorical variables were included). The QQ-plot is a particularly useful tool for comparing the distribution of the quantiles, dividing the observations of a sample into intervals of the same size, to a theoretical Gaussian distribution.

In linear regression, we theoretically assume normality of all the replicate observations at a particular covariate value [302]. However, assumption of normality (and homogeneity) at each covariate value cannot be verified unless many (> 25) replicates per *X* value are taken, which is seldom the case in ecological studies. In practice, we can check normality after having run a linear regression. Specifically, we can calculate the model residuals, and we can then make a histogram or QQ-plot of the pooled residuals to get an estimate of normality [294, 299, 303], since normality of the raw data implies normality of the model residuals. Normality of the pooled residuals is reassuring, but this does not provide conclusive evidence for normality of the population data.

3.1.1.3   *Step 3: Homogeneity of variance in Y*

Homogeneity of variance is an important assumption in analysis of variance, regression-related models and in multivariate techniques like LDA. Violation of homogeneity, also called heteroscedasticity, happens if the spread of the data in the response is not the same at each value of the predictor (*X*) [294]. Minor violations of homogeneity are not too problematic, but serious heteroscedasticity leads to poor estimation of coefficients. If multiple observations for each *X* value were available, homogeneity was checked comparing the spread of the response for each value or category of the predictors using box-plots or Cleveland dotplots. If the model includes both continuous and categorical predictors, homogeneity was checked with conditional boxplots to observe the spread of the response for each covariate value by category. If multiple observations were not available, verification of homogeneity was done using the standardised residuals of the model, by plotting residuals vs fitted values, and making conditional boxplots for the residuals. The solution to heterogeneity of variance was either a transformation of the response variable to stabilise the variance, or applying statistical techniques that do not require homogeneity (e.g. generalised least squares) [287, 289].

3.1.1.4   *Step 4: Collinearity in X*

Collinearity is the existence of correlation between covariates. If the underlying question is to select which covariates are driving the response variable in a system, then a critical problem to overcome is often collinearity [294, 297]. If collinearity is ignored, the result is likely to be a confusing statistical analysis in which no term is significant, but where dropping one covariate can make the others significant, or even change the sign of estimated parameters [294]. In multiple linear regression (i.e. $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_j X_j + \varepsilon$), an expression for the variances of the parameters β_j is given by [295, 308]:

$$Variance(\beta_j) = \frac{1}{1 - R_j^2} \times \frac{\sigma^2}{(n-1)S_j^2}$$

Equation 3.1

The second expression is a constant, depending on the datasets and covariate considered, where the term S_j depends on covariate values, *n* is the sample size, and σ^2 is the variance of the residuals. In the

first expression, the term R_j^2 is the R^2 from a linear regression model in which covariate X_j is used as a response variable, and all other covariates as explanatory variables. A high R^2 means that most of the variation in covariate X_j is explained by all other covariates, which indicates collinearity. As a consequence the standard errors (SE) of the parameters are inflated with the square root of $1/(1 - R_j^2)$ [since $SE(\beta_j) = \sqrt{Variance(\beta_j)}/\sqrt{n}$], also called the variance inflation factor (VIF), determining larger *p*-values. Other approaches used to detect collinearity were creation of pairwise scatterplots among covariates and calculation of pairwise correlation coefficients.

The strategy used for addressing collinearity was to sequentially drop from a model the covariate with the highest VIF, recalculate the VIFs, and repeat this process until all VIFs are smaller than a preselected threshold [302], usually a value of 3 or even 2 for a more stringent approach [287, 294]. The easiest way to solve collinearity was by dropping collinear covariates. Other available methods to solve collinearity include principal components regression, residual and sequential regression [297].

3.1.1.5 Step 5: Independence of Y values

A crucial assumption of most statistical techniques is that observations are independent of one another [309]. Violation of independence is the most serious problem as it invalidates important tests, such as F-test and *t*-test. There is violation of independence if the *Y* value at X_i is influenced by other X_i [303], meaning that information from any one observation should not provide information on another after the effects of other variables have been accounted for. For example, if individuals at locations close to each other have characteristics that are more similar to each other than to individuals from locations separated by larger distances, then we would violate the independence assumption. There are two ways independence can be introduced: either an improper model specification (e.g. use of a linear rather than non-linear model) or a dependence structure due to the nature of data [287]. When such dependence arises, the statistical model used needs to account for it by including, for instance, any spatial or temporal relationships, or by nesting data in a hierarchical structure. There were many ways used to include a temporal or spatial dependence structures such as using mixed-effects modelling [289, 310], incorporating a temporal residual correlation structure using generalised least squares [287, 289], the inclusion of a smoothing function of time or a two-dimensional smoother of spatial coordinates [288, 291].

Independence was checked in the raw data before the analysis by plotting the response variable vs time or spatial coordinates, and in residuals afterwards by plotting residuals vs fitted values and any covariate included in the model. Detection of residual patterns (e.g. sinusoidal trends or clustering above or below zero) indicated sign of dependence. Alternatively models with and without a correlation structure were compared through selection criteria or hypothesis-based tests to check improvement in model fit [289]. A more formal way to assess the presence of temporal dependence was to plot auto-correlation (ACF) and partial autocorrelation (pACF) functions for regularly spaced time series, or variograms for

irregularly spaced time series and spatial data [311]. An ACF calculates the Pearson correlation between a time series and the same time series shifted by 1, 2, …, k time units. This allowed me to estimate the correlation between any two points at time t_k and t_{k-1}, if significant correlations were found it meant observations were not independent.

3.1.1.6   Step 6: Relationships and interactions

Another essential part of data exploration was plotting the response variable against each covariate [294]. Scatterplots were used to **i**) identify the type of potential relationship (e.g. linear vs non-linear) between dependant and independent variables, **ii**) detect observations that did not comply with the general patterns between two variables, and **iii**) identify potential interaction between continuous variables. Coplots, plotting the relationships between dependant and independent variables for the groups of categorical covariates, where used when categorical predictors were included in the model.

3.2 Linear model (LM)

The cause-effect relationship in an empirical model is often represented as a linear statistical model. The linear model (LM) is the basis of many regression-type methods for analysing data. The aim of a linear model is to describe the variation of one continuous variable, the dependant (or response) variable has a function of one or multiple variables, the independent (or explanatory) variables. The simplest form of linear model is the bivariate linear regression model that is used to describe the relationships between two variables. The bivariate linear regression is defined by:

$$Y_i = \beta_0 + \beta_1 \times X_i + \varepsilon_i \qquad \varepsilon_i \sim N(0, \sigma^2)$$

Equation 3.2

where Y_i is the value of the response variable for the ith observation, which is measured on a population of size N (i = 1, …, N), and X_i is the explanatory variable. The parameter β_0 is the population intercept, and β_1 is the population slope, both of which are unknown. The unexplained variation in the data is captured by the residual error ε_i that is assumed to follow a Normal distribution (N, or Gaussian) with expectation (mean) 0 and a variance σ^2 that is unknown. Our objective here is to calculate β_0, β_1, and σ^2.

In practice, in an experiment, what we do first is to take a sample of size n from the population. We use this sample to calculate the value of the estimates β_0, β_1, and σ^2 for the sample plus confidence intervals, which inform us about the variation in the estimated parameters if we repeat the same experiment a large number of times. If the model meets a series of assumptions, we can use the estimated parameters to make inferences about the population parameters.

What the model in Equation 3.2 means for our sample is usually summarised by the **matrix notation** of the model. Given a population sample of size n, with the index i ranging from 1 to n, the above Equation 3.2 can be written as:

$$Y_1 = \beta_0 + \beta_1 \times X_1 + \varepsilon_1 \qquad \varepsilon_1 \sim N(0, \sigma^2)$$
$$Y_2 = \beta_0 + \beta_1 \times X_2 + \varepsilon_2 \qquad \varepsilon_2 \sim N(0, \sigma^2)$$
$$\vdots \qquad\qquad \vdots \qquad\qquad\qquad \vdots$$
$$Y_n = \beta_0 + \beta_1 \times X_n + \varepsilon_n \qquad \varepsilon_n \sim N(0, \sigma^2)$$

Equation 3.3

representing the linear relationship (Equation 3.2) for each observation i in the sample. If we define the following vectors:

$$\mathbf{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} \quad \mathbf{X} = \begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ \vdots & \vdots \\ 1 & X_n \end{bmatrix} \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} \quad \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

Equation 3.4

the linear model expressed in Equation 3.3 can be simplified with its matrix notation (note the use of bold characters for expressing vectors and matrices):

$$\mathbf{Y} = \mathbf{X} \times \boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

Equation 3.5

$\mathbf{Y}$ is a vector including the observations for the response variable and $\boldsymbol{\varepsilon}$ is a vector of residuals, both of length n (the sample size). $\boldsymbol{\beta}$ is a vector comprising the regression parameters, the intercept and the slope (to be estimated) with length p (the number of parameters). $\mathbf{X}$ is called the design matrix, or model matrix, and it is a $n \times p$ matrix of known constants. Here, the design matrix contains two columns; all the values of the first column are equal to 1, and these are multiplied by β_0 for the intercept. The second column contains the covariate values for all the observations that are multiplied by β_1. This matrix notation is the basis for the calculation of the parameters estimates, which can be obtained with different estimation methods, such as ordinary least squares (OLS), maximum likelihood (ML), and restricted maximum likelihood (REML). For examples, in OLS the estimated vector of regression parameters ($\hat{\boldsymbol{\beta}}$) is calculated as:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y}$$

Equation 3.6

where $\mathbf{X}^T$ is the transpose[2] of the $\mathbf{X}$ matrix, and $(\mathbf{X}^T\mathbf{X})^{-1}$ is the inverse[3] of the $\mathbf{X}^T\mathbf{X}$ matrix.

---

[2] The transpose of a matrix $\mathbf{X}$ is a matrix, denoted $\mathbf{X}'$ or $\mathbf{X}^T$, whose rows are the columns of $\mathbf{X}$ and whose columns are the rows of $\mathbf{X}$, all in the same order.
[3] The inverse of $\mathbf{X}$ is the matrix $\mathbf{X}^{-1}$ that you have to multiply $\mathbf{X}$ by in order to obtain the identity matrix ($\mathbf{I}$), such that $\mathbf{X}^{-1}\mathbf{X} = \mathbf{I}$. The identity matrix is a matrix with 1's on the diagonal and 0's elsewhere with the same dimension of $\mathbf{X}$.

The cause-effect relationship between a response variable and multiple explanatory variables can be expressed with a multiple linear regression. A multiple linear regression model is an extension of the bivariate liner regression model (Equation 3.2) in the sense that *m* multiple predictors are used.

$$Y_i = \beta_0 + \beta_1 \times X_{1i} + \beta_2 \times X_{2i} + \cdots + \beta_m \times X_{mi} + \varepsilon_i \qquad \varepsilon_i \sim N(0, \sigma^2)$$

Equation 3.7

In the matrix notation, the design matrix **X** would contain *m* + 1 columns, where all the value of the first column are 1. The second and other columns contain the covariates' values for all the observations. **β** is a vector comprising all the *m* + 1 regression parameters.

3.3 Generalised linear model (GLM)

A generalised linear model (GLM) [312] is a flexible generalisation of the ordinary linear model. The GLM generalises linear regression by allowing **i**) response variables to have error distributions other than Normal, **ii**) the linear model to be related to the response variable *via* a link function, and **iii**) the magnitude of the variance of each measurement to be a linear function of its predicted value. Generalised linear modelling unifies several statistical and modelling techniques, among which linear regression, multiple linear regression, analysis of variance and covariance, logistic regression, and Poisson regression [287, 312].

The GLM consists of three elements: (1) the **linear predictor**, (2) the **distribution of the response error**, and (3) the **link function**. Any linear model can be obtained using the appropriate combination of these three properties. Compared to the classical linear model assuming that errors, or residuals, are normally distributed, the GLMs cover a wider range of situations by allowing the error to have an arbitrary distribution (rather than simply Gaussian), and for an arbitrary function of the response variable (the link function) to vary linearly with the predictor values (rather than assuming that the response itself must vary linearly).

Prior to the development of GLMs, researchers often applied transformations to data where the errors were not normally distributed, such as count and proportional data. Although transformed data with Normal errors may lead to better fitting and simpler models than GLMs, Warton & Hui [313] and O'Hara & Kotze [314] were the first to start advocating strongly the use of GLMs rather than transformations for count, proportion, and other data where Normal errors may not be appropriate. Much of the data ecologists commonly collect do not present normally distributed errors [315]. For example, count data (e.g. number of species, individuals) are discrete values constrained to be positive. Binary data (e.g. female/male, presence/absence) are discrete and bounded at 0 and 1. Proportional data with a meaningful denominator (e.g. proportion of live/dead individuals) and without a meaningful denominator (e.g. % of plant cover) are continuous but bounded at 0 and 1. The bounded and/or discrete

nature of biological data leads to errors that are not normally distributed, and distributions such as Poisson, Binomial, Negative Binomial, may be appropriate for modelling the error [287, 299].

If count or proportional data are handled improperly, results can be biased, underpowered, and difficult to interpret [313, 314]. Maximum likelihood estimates of the parameters, for example, depend on the specified error distribution; therefore, using different error distributions can lead to changes in the significance of explanatory variables. If we are to move beyond determining statistical significance alone (*p*-values) to appropriately modelling a biological process, it would be more appropriate to use a model that captures more of the properties of the processes we are trying to understand [315]. Often, we are not only interested in the fitted relationship, or mean response, but also in the unexplained variation that surrounds the response. For this, it is necessary to describe the nature and magnitude of this error variation appropriately, including its distribution around the fitted values, as it may not be homogeneous.

Moreover, proportion, count, and binary data by their nature tend to have fundamentally non-linear relationships with explanatory variables [287, 316]. GLMs enable the inclusion of inherent non-linearity through the link function. The link function relates the linear predictor, which is not bounded and has homogenous variance, to the expected values of the data, which may be bounded and have heterogeneous variance [316]. Link functions allow fitted values from the model of the linear predictor to be converted to non-linear predictions on the same scale that the data were collected. In a GLM, for example, a logit link function enables fitting and predictions of probabilities bounded to 0 and 1 [315].

3.3.1 GLM structure

In a GLM the response variable Y_i is assumed to follow an exponential family distribution (a large range of probability distribution including Normal, Binomial, Poisson, and Gamma distributions). In the following sub-sections, I will show the simple case of the formulation of a Gaussian and a Poisson GLM. A Gaussian GLM, is a GLM with the response Y_i following a Normal distribution which is assumed to have mean μ_i and variance σ^2.

$$E(Y_i) = \mu_i \quad \text{and} \quad var(Y_i) = \sigma^2$$

Equation 3.8

Where E(Y_i) is the expected, or mean, value for Y_i, and var(Y_i) is the variance for Y_i. The formulation of a GLM consists of three steps:

1. Specification of the systematic part, or linear predictor, as a function of the explanatory variables.
2. An assumption on the random part that is the distribution of the response variable Y_i (mean μ_i and variance σ^2 in this case).

3. Definition of the relationship between the mean value of Y_i and the systematic part, called the link function.

3.3.1.1 Linear predictor

First, we need to define the systematic part of the model, the linear predictor, which is represented by the explanatory variables. The linear predictor η_i, for each data point i, is a linear function of the explanatory variables of a model

$$\eta_i = \eta(X_{1i}, \ldots, X_{pi}) = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_p X_{pi}$$

Equation 3.9

Where X_{1i}, …, X_{pi} are the explanatory variables, β_0 is the intercept, and β_p is the slope of the *p*th explanatory variable. Note that no error is included in the linear predictor, it just contains the systematic part of the model. In matrix notation (see Equation 3.5) the linear predictor is:

$$\boldsymbol{\eta} = \mathbf{X} \times \boldsymbol{\beta}$$

Equation 3.10

3.3.1.2 The error structure

Once the initial model for the linear predictor is constructed, an appropriate structure for the error of the response variable should be selected (Table 3.2). The error distribution describes how the variation, which is not explained as part of the linear predictor, is distributed. There are usually good *a priori* reasons for choosing a particular error distribution for a particular type of data, such as discrete distributions (e.g. Poisson, Binomial) for count data and continuous distribution (e.g. Gaussian, invers-Gaussian, Gamma) for continuous data. In this case, for a Normal error distribution model we assume that the response Y_i is normally distributed with mean μ_i and variance σ^2:

$$Y_i \sim N(\mu_i, \sigma^2) \qquad \text{or} \qquad \mathbf{Y} \sim N(\boldsymbol{\mu}, \sigma^2)$$

Equation 3.11

The observations for the vector **Y** are estimated to come from a Normal distribution with mean $\boldsymbol{\mu}$ and constant variance σ^2. The mean, $\boldsymbol{\mu}$, is a vector of fitted (predicted) values describing the relationship between the response variables and the explanatory variables. Differently from the variance, the mean value changes with the value of the explanatory variables.

Table 3.2 Error distributions from the exponential family
Examples of error distributions, canonical link functions, and type of commonly modelled response variables.

Error distribution	Canonical link function (name)		Type of response variable
Normal/Gaussian	μ	(identity)	Continuous positive and negative values
Poisson	$\log(\mu)$	(log)	Count data
Quasi-Poisson	$\log(\mu)$	(log)	Count data
Negative Binomial	$\log(\mu)$	(log)	Count data
Binomial	$\log(\mu/\mu\text{-}1)$	(logit)	Proportions data
Binomial (Bernoulli)	$\log(\mu/\mu\text{-}1)$	(logit)	Binary data
Quasi-Binomial	$\log(\mu/\mu\text{-}1)$	(logit)	Proportions and binary data
Beta	$\log(\mu/\mu\text{-}1)$	(logit)	Proportions (continuous)
Gamma	$1/\mu$	(inverse)	Continuous positive values
Inverse Gaussian	$1/\mu^2$	_	Continuous positive values

3.3.1.3 The link function

The last step of a GLM is the definition of the link function. The link function provides the relationship between linear predictor, η_i or $\eta(X_{1i}, \ldots, X_{pi})$, and the mean (or expected) value of the response, Y_i (which is μ_i). In general, $\boldsymbol{\eta} = g(\boldsymbol{\mu})$, where g() is called the link function. In the case of the Normal distribution of the error, the appropriate link function which connects the predictor function to the mean value of the Normal distribution is the identity link $\boldsymbol{\eta} = \boldsymbol{\mu}$. This means that the liner predictor predicts on the same scale of the original response variable.

Usually we have the linear predictor $\boldsymbol{\eta}$ (fitted values on the scale of the linear predictor) and we want to convert these to the predicted means $\boldsymbol{\mu}$ (equivalent to the fitted values on the scale of the response variable **Y**). Therefore, we apply the inverse of the link function g()$^{-1}$, and the inverse of the identity link is still the identity link. The expectation of the response variable **Y** is therefore equivalent to the fitted values of both the linear predictor and response scales, $E(\mathbf{Y}) = \boldsymbol{\mu} = g(\boldsymbol{\eta})^{-1} = \boldsymbol{\eta}$.

In general, GLMs allow the error distribution to come from the exponential family other than the Normal and the link function may be any monotonic differentiable function. Canonical link functions are well specified for each family of error distribution (Table 3.2); these are statistically convenient, but are not necessarily the most appropriate to use, so alternative link functions can be used for each family.

3.3.1.4 Gaussian GLM

The three steps above give the following GLM:

$$Y_i \sim N(\mu_i, \sigma^2)$$
$$E(Y_i) = \mu_i \quad \text{and} \quad var(Y_i) = \sigma^2$$
$$\mu_i = \eta(X_{1i}, \dots, X_{pi})$$

Equation 3.12

These elements (Equation 3.12) describe a GLM with a Normal distribution and an identity link function

$$E(Y_i) = \eta(X_{1i}, \dots, X_{pi}) = \beta_0 + \beta_1 X_{1i} + \dots + \beta_p X_{pi}$$

Equation 3.13

The matrix notation of Equation 3.13 is

$$E(\mathbf{Y}) = \boldsymbol{\eta} = \mathbf{X} \times \boldsymbol{\beta} \quad \text{or} \quad \mathbf{Y} \sim N(\boldsymbol{\eta}, \sigma^2)$$

Equation 3.14

3.3.1.5 Poisson GLM

The creation of any GLM with different error family distribution and link function can be easily implemented for example when analysing count data through a Poisson GLM. Here, Y_i is Poisson distributed (P) with mean μ_i. For a Poisson distribution the mean is equal to the variance μ_i. The systemic part of the model is still $\eta(X_{1i}, \dots, X_{pi})$. While there is a logarithmic link function (natural logarithm) which relates the mean of Y_i and the linear predictor. The log-link function ensures that the fitted values are always non-negative (useful for predicting count data).

As result we define the following Poisson GLM:

$$Yi \sim P(\mu_i)$$
$$E(Y_i) = \mu_i \quad \text{and} \quad var(Y_i) = \mu_i$$
$$\log(\mu_i) = \eta(X_{1i}, \dots, X_{pi}) \quad \text{or} \quad \mu_i = e^{\eta(X_{1i}, \dots, X_{pi})}$$

Equation 3.15

Which gives the following model

$$\log(E(\mathbf{Y})) = \mathbf{X} \times \boldsymbol{\beta} \quad \text{or} \quad \mathbf{Y} \sim P(\log(\boldsymbol{\eta})^{-1})$$

Equation 3.16

3.4 Generalised linear mixed model (GLMM)

Many common statistical models can be expressed as linear models that incorporate both **fixed** effects, which are parameters associated with an entire population or with certain repeatable levels of experimental factors, and **random** effects, which are associated with individual experimental units drawn at random from a population. A model with both fixed effects and random effects is called a

mixed-effects model. Mixed-effects models are primarily used to describe relationships between a response variable and some covariates in data that are grouped according to one or more classification factors. Examples of such grouped data include repeated measures data, multilevel data, and block designs. By associating common random effects to observations sharing the same level of a classification factor, mixed-effects models flexibly represent the covariance structure induced by the grouping of the data [289].

Generalised linear mixed models (GLMMs)[4] are a powerful class of statistical models that combine the characteristics of GLMs (Section 3.3) and mixed-effects models in which both the fixed and the random effects occur linearly in the model function [287, 310]. They extend linear models handling fixed effect (the typical way to compare effects of explanatory variables) and random effects, which can be regarded as additional error terms to account for correlation among observations within the same group. They can also handle a wide range of error distribution families and link functions. The complete specification of a GLMM includes the distribution of the response variable, the link function, the definition of continuous or categorical fixed-effect predictors, and the definition of the random effects, which specify how some model parameters vary randomly across groups.

3.4.1.1   *Random effects*

The traditional view of random effects is as a way to do correct statistical tests when some observations are correlated. When samples are collected in groups, we violate the assumption of independence. Indeed, there will be some variation within groups (σ^2_{within}) and some among groups (σ^2_{among}). The total variance is $\sigma^2_{total} = \sigma^2_{within} + \sigma^2_{among}$ and therefore the correlation between any two observations in the same group is $\rho = \sqrt{\sigma^2_{among}/\sigma^2_{total}}$ (observations that come from different groups are uncorrelated). Sometimes one can solve this problem easily by taking group averages, and test for significant difference among groups using the calculated mean values [310, 315].

We can also think of random effects as a way to combine information from different groups (levels) within a grouping variable. If we had only a few observations from few levels, we might have to pool the data, ignoring the differences among levels. Pooling assumes that σ^2_{among} (the variance in slopes among levels) is effectively zero and observations are uncorrelated ($\rho = 0$). On the other hand, if we had many observations from each level, and especially if we had a small number of levels, we might want to estimate the slope for each level individually, or in other words to estimate a fixed effect of the explanatory variable for each level. Treating the grouping factor as a fixed effect assumes that information about one level gives us no information about the slope at any other level. Treating a

---

[4] GLMMs are part of the statistical frontier and are still being implemented. GLMMs may result difficult to use and, therefore, should be used with care [315].

grouping variable as a random effect compromises between the extremes of pooling all the observations and calculating a single slope, and estimating separate slopes for each level. In a random term grouping variable the deviation of each level's value from the population average are called conditional modes, rather than estimates [315].

Random effects are especially useful when we have **i)** many levels (e.g., many species or blocks), **ii)** relatively little data on each level (although multiple samples from most of the levels are needed), and **iii)** uneven sampling across levels. Random effects can also be described as predictors where we are interested in making inferences about the distribution of values (i.e. the variance among the values of the response at different levels) rather than in testing the differences of mean values between particular levels. Choosing a random effect trades the ability to test hypotheses about differences among levels for the ability to quantify the variance among levels and generalise to levels that were not measured in the experiment. If we treat, for example, "species" as a fixed effect, we cannot make any inference about an unmeasured species. While, if we use it as a random effect, then we can state that an unmeasured species will have a value equal to the population mean estimated from the species that were measured. As with all statistical generalisation, levels must be chosen in a way that is representative of the population we want to generalise to [287, 315, 316].

Sometimes random effects are used to control for variation among sites when they represents "nuisance" that it is often the case in ecological experiments [287]. However, random effects are generally ineffective when the grouping variable has too few levels [310]. These should not be used when the grouping variable has fewer than five (or better eight) levels [317], since variance estimates are calculated from a small sample and are unstable.

3.4.2 GLMM definition and use

The general structure of a GLMM as proposed by Laird & Ware [318] is represented by its single-level formulation. The original single-level formulation expresses the n_i-dimensional response vector $\mathbf{Y}_i$ for the ith group ($i = 1, \ldots, m$) as:

$$\mathbf{Y}_i = \mathbf{X}_i \times \boldsymbol{\beta} + \mathbf{Z}_i \times \boldsymbol{b}_i + \boldsymbol{\varepsilon}_i$$

$$\boldsymbol{b}_i \sim N(0, \boldsymbol{\Psi}) \qquad \boldsymbol{\varepsilon}_i \sim N(0, \boldsymbol{\Lambda}\sigma^2)$$

Equation 3.17

$\mathbf{Y}_i$ is a vector including the observations for the dependant variable for the ith group; $\boldsymbol{\beta}$ is a p-dimensional vector of fixed effects; $\boldsymbol{b}_i$ is the q-dimensional vector of random effects; $\mathbf{X}_i$ is the design matrix of fixed effects of dimension $n_i \times p$; $\mathbf{Z}_i$ is the design matrix of the random effect of dimension $n_i \times q$; $\boldsymbol{\varepsilon}_i$ is the n_i-dimensional within-group error vector with a spherical Gaussian distribution and variance covariance matrix $\boldsymbol{\Lambda}\sigma^2$. $\boldsymbol{\Lambda}$ is a positive definite matrix, which is typically used to model residual
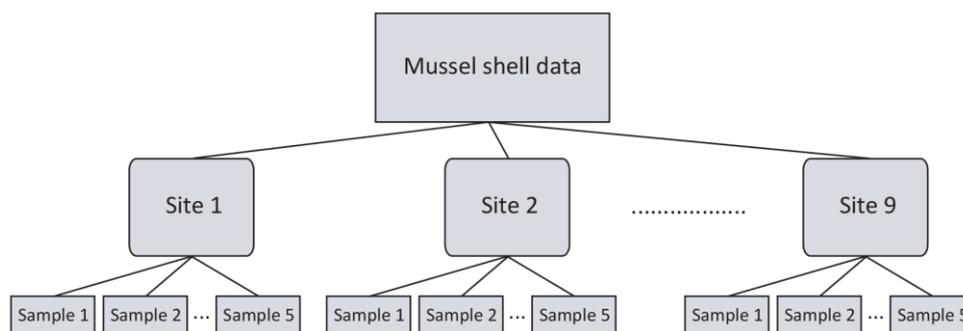
Figure 3.1 Hierarchical structure of the data
Illustration of the nested nature of the dataset. We have multiple samples (observations of shell weight) from each of the sites. Observation values from samples on the same site are likely to be more similar to each other than to values from different sites.

correlation, and its elements are usually determined by simple models. Often $\mathbf{\Lambda}\sigma^2$ is simply the identity matrix ($\mathbf{I}\sigma^2$, a matrix with a diagonal of 1 and all 0).

The random effect $\boldsymbol{b}_i$ and the within-group errors $\boldsymbol{\varepsilon}_i$ are assumed to be independent for different groups and to be independent of each other in the same group. $\boldsymbol{b}_i$ is allowed to have a different value for each level of the random effect grouping variable, while $\boldsymbol{\beta}$ is the same for all the level of the random term component. Because the distribution of the random effect vector $\boldsymbol{b}_i$ is assumed to be Normal with mean 0, it is completely characterised by its variance-covariance matrix $\mathbf{\Psi}$. The fixed $\mathbf{X}_i \times \boldsymbol{\beta}$ term and the random $\mathbf{Z}_i \times \boldsymbol{b}_i$ term defines together the mixed model. This extension allows to model more complex stochastic structures than the ordinary linear model and implies that the elements of the response vector are no longer independent.

3.4.2.1 *Random intercept and slope model: a working example*

The easiest way to explain the formulation of the GLMM in Equation 3.17 and its applications is by using an example which is similar to the analyses performed in Chapters 4, 5, and 6. In this example, our objective is to estimate the coefficients for the relationships between the weight (dependant variable) and the length (continuous covariate) for the shells of blue mussels collected from multiple geographic areas. For this experiment, five mussel specimens ($i = 1, \ldots, 5$) from nine different sites ($j = 1, \ldots, 9$) were collected, and for each individual the shell weight (Y_i) and length (X_i) were measured (Figure 3.1). The question being asked is whether there are any differences between the weight-length relationships at these nine sites. As explained, GLMM is the mixed modelling extension of a GLM. So, we start with a simple GLM using all 45 observations pooled from all nine sites. This regression line can be written:

$$Y_i = \beta_0 + \beta_1 X_j + \varepsilon_i \quad \text{and} \quad \varepsilon_i \sim N(0, \sigma^2)$$

Equation 3.18

This model contains only three unknown parameters: the two regression parameters (intercept β_0 and slope β_1) and the variance of the error σ^2. This model produces one line and assumes that the weight-

length relationship is the same at all the nine sites (Figure 3.2a). If we want to model among-site differences in the weight-length relationship, we should define a model allowing the intercept and slope to differ per site:

$$Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + \varepsilon_{ij} \quad \text{and} \quad \varepsilon_{ij} \sim N(0, \sigma^2)$$

Equation 3.19

where $j = 1, \ldots, 9$, and $i = 1, \ldots, 5$ (there are five observations per site). This model provides multiple regression lines (a different prediction for the jth site, Figure 3.2b) in one regression analysis using site as a factor, shell length as a continuous explanatory variable, and a the site-length interaction term. The total number of parameters to estimate are a slope and intercept per site ($9 + 9$ parameters) plus the error variance, for a total of 19 estimates. The total number of observations is only 45, so proportionally this model has a worryingly large number of parameters. The model with only one regression line (Equation 3.18) is the most basic model, and the model with nine regression lines in which slope and intercept are allowed to differ per beach (Equation 3.19) is the most complex model. Moreover, there are two intermediate models that we can build: one where the intercept is allowed to differ between the beaches but with equal slopes, and one where the intercepts are kept the same and the slopes are allowed to differ. This reduces the number of parameters to 11. These models can be written as:

$$Y_{ij} = \beta_{0j} + \beta_1 X_{ij} + \varepsilon_{ij} \quad \text{and} \quad \varepsilon_{ij} \sim N(0, \sigma^2)$$

Equation 3.20

$$Y_{ij} = \beta_0 + \beta_{1j} X_{ij} + \varepsilon_{ij} \quad \text{and} \quad \varepsilon_{ij} \sim N(0, \sigma^2)$$

Equation 3.21

If we are only interested in the general relationship between shell weight and length, and do not care about differences between sites, then we could ignore the nominal variable site. However, this means that the variance component might contain between-site variation, and not taking this into account might affect standard errors and p-values of the fixed effects, but the price of 16 extra regression parameters can be rather large. To avoid this, mixed modelling can be used. In the model in Equation 3.20, each beach has a different intercept but the same slope. We extend it by using site as a random grouping variable:

$$Y_{ij} = \beta_0 + \beta_1 X_{ij} + b_{0j} + \varepsilon_{ij}$$
$$b_{0j} \sim N(0, \sigma_{b_0}^2) \quad \text{and} \quad \varepsilon_{ij} \sim N(0, \sigma^2)$$

Equation 3.22

Again, the index are $j = 1, \ldots, 9$ (representing site), and $i = 1, \ldots, 5$ (representing samples on a site). Rather than estimating nine intercept and a variance term as in Equation 3.20, in Equation 3.22, we assume there is only one overall regression line with a single intercept and a single slope. The single intercept β_0 and single slope β_1 are the fixed parameters, and the additional parameter, called random intercept, b_0, adds a certain amount of random variation to the intercept at each site (the intercept is
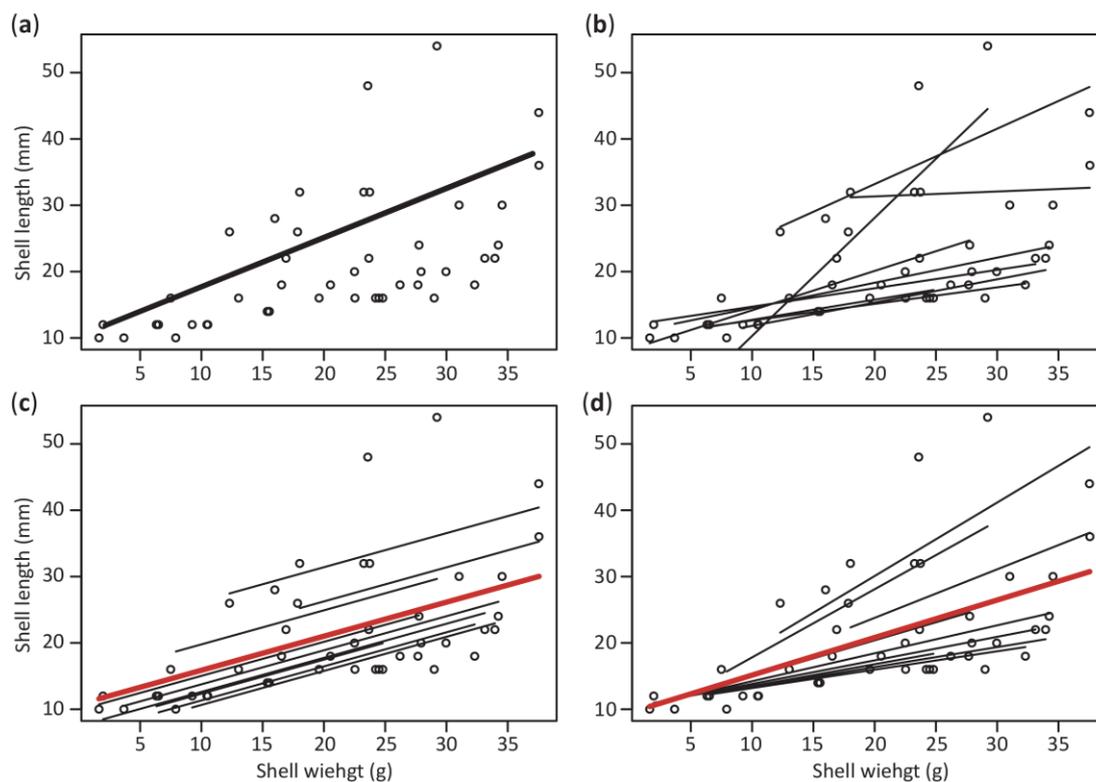
Figure 3.2 Comparison of GLM and GLMM predictions
Scatterplots of shell weight and length for the mussel experiment and model predictions.
Predictions from GLMs showing (**a**) a single regression line on the pooled data, and (**b**)
multiple regression lines, one for each site. Fitted value obtained by a GLMM (**c**) with a
random intercept, and (**d**) with a random intercept and slope. The thick red line represents
the fitted values obtained by the fixed component of the model, also called the population
model, while the other lines represent the within-group predictions and are obtained by
adding the contribution of the of the random component for each site *j* to the population
fitted curve.

increased or decreased by a random value, Figure 3.2c). The random intercept is assumed to follow a
Normal distribution with expectation 0 and variance $\sigma_{b_0}^2$. Hence, the unknown parameters in the model
are β_0, β_1 the variance of the noise σ^2, and the variance of the random intercept $\sigma_{b_0}^2$, only 4 parameters.
So, basically, we have the same type of model fit as Equation 3.20, but instead of nine estimated levels
for each intercept, we now have nine realisations au b_{01}, …, b_{09} from which we assume that they follow
a normal distribution and we only need to estimate the variance of this distribution. This type of model
is called a mixed-effects model with **random intercept**.

If we extend Equation 3.22 to get the mixed modelling equivalent Equation 3.21, we need a mixed
model with a random intercept that allows the regression line to change randomly:

$$Y_{ij} = \beta_0 + \beta_1 X_{ij} + b_{0j} + b_{1j}X_{ij} + \varepsilon_{ij}$$

$$b_{0j} \sim N\left(0, \sigma_{b_0}^2\right) \quad \text{and} \quad b_{1j} \sim N\left(0, \sigma_{b_1}^2\right) \quad \text{and} \quad \varepsilon_{ij} \sim N(0, \sigma^2)$$

Equation 3.23

This is the same model, except for the term $b_{1j}X_{ij}$. This new term allows for random variation of the slope at each site j (Figure 3.2d). The model fit will look similar to the one in Equation 3.19, except that considerably fewer parameters are used: the fixed intercept β_0 and β_1 slope, and three random variances $\sigma_{b_0}^2$, $\sigma_{b_1}^2$ and σ^2, for a total of five parameters against the 19 of model Equation 3.19. This model is called random intercept and slope model.

3.5 Generalised additive model (GAM)

A generalised additive model (GAM) [288, 291, 292] is a generalised linear model with a linear predictor involving a sum of smooth functions of covariates. In general, the model has a structure like:

$$E(Y_i) = \mu_i \quad \text{and} \quad Y_i \sim EF(\mu_i, \varphi)$$
$$g(\mu_i) = \mathbf{X}_i^* \times \boldsymbol{\theta} + f_1(X_{1i}) + f_2(X_{2i}) + \cdots + f_j(X_{ji}) + \varepsilon_{ij}$$

or

$$g(\mu_i) = \mathbf{X}_i^* \times \boldsymbol{\theta} + \sum_j f_j(X_{ji})$$

Equation 3.24

where Y_i is the response variable, which follows an exponential family (EF) distribution with mean μ_i and scale parameter φ; $g()$ is a specified link function; $\mathbf{X}_i^*$ is a row of the design matrix $\mathbf{X}_i$ for any strictly parametric (linear) model components, and $\boldsymbol{\theta}$ is the corresponding parameter vector (so that $\mathbf{X}_i^* \times \boldsymbol{\theta}$ is the linear predictor); f_j are smooth functions of the covariate X_j; ε_{ij} is the error that is assumed to be normally distributed with expectation 0 and variance σ^2.

The model allows for flexible specification of the dependence of the response on the covariates, but by specifying the model only in terms of "smooth functions", rather than detailed parametric relationships. This flexibility and convenience comes at the cost of two new theoretical problems: it is necessary both to represent the smooth functions in some way and to choose how smooth they should be [288]. GAMs produce smooth lines that are visually similar to those produced with older smoothing approaches such as locally estimated (or weighted) scatterplot smoothing, LO(W)ESS, but they use a different method. The smooth functions used in GAMs are called splines. The word spline defines a flexible strip that can be fixed at certain points, called knots, and then bent in a smooth curve around these points. Briefly, the X gradient (interval of observations) is divided into a certain number of intervals by knots. Then a series of cubic polynomial is fitted to each interval of observations, and the fitted values per segment are joint together at the knots to form a smoothing curve. Statistical splines improve on LOESS and related smoothers by **i**) having a stronger analytic basis, **ii**) being better at preventing over-smoothing, **iii**) having superior software implementations, and **iv**) being easier to make part of GA(M)Ms (see Section 3.6).

3.5.1.1 *Smoothing splines*

The representation and estimation of component functions of a model is best introduced by considering a model containing one function of one covariate, the univariate smoother:

$$Y_i = \beta_0 + f(X_i) + \varepsilon_{ij} \quad \text{and} \quad \varepsilon_{ij} \sim N(0, \sigma^2)$$

Equation 3.25

Y_i is a response variable, X_i a covariate, and f a smooth function, or spline. The family of splines is rather large including cubic splines, natural splines, spline with shrinkage and smoothing splines. Here, I will show a simple case of a cubic regression spline and how this can be defined.

To estimate f requires $f(X_i)$ to be represented in such a way that it becomes a linear model. This can be done by choosing a *basis* that defines the space of functions of which $f(X_i)$ is an element. Choosing a basis corresponds to choosing a series of *basis functions* which are the basic components of any continuous function[5]. If we define $b_j(X_i)$ as the jth basis function, we can represent $f(X_i)$ as:

$$f(X_i) = \sum_{i=1}^{p} \beta_j \times b_j(X_i)$$

Equation 3.26

substituting Equation 3.26 in Equation 3.25 clearly yields to a linear model. We can notice similarities with the GLMM in Equation 3.17, suggesting how a smoothing function can be represented by linear combination of fixed, β_j, and random, b_j, elements. Equation 3.26 can be used to generate a range of polynomial functions depending on the value of p (by defining the space of polynomials of order p). As a simple example, suppose that f is a 4th order polynomial as:

$$f(X_i) = \beta_1 \times b_1(X_i) + \beta_2 \times b_2(X_i) + \beta_3 \times b_3(X_i) + \beta_4 \times b_4(X_i)$$

Equation 3.27

A basis for this space is $b_1f(X_i) = 1$, $b_2f(X_i) = X_i$, $b_3f(X_i) = X_i^2$, $b_4f(X_i) = X_i^3$, then Equation 3.27 becomes a simple polynomial function:

$$f(X_i) = \beta_1 + \beta_2 \times X_i + \beta_3 \times X_i^2 + \beta_4 \times X_i^3$$

Equation 3.28

This is a basis function representation of f, using a polynomial basis, as a cubic spline, and it can produce a wide range of possible shape depending on the values of $\beta_{1, ..., } \beta_4$. However, when modelling Y_i, we do not know the values β_1, β_2, β_3, and β_4, and the shapes the function f can assume are not flexible enough to model too complicated patterns (such as higher order polynomials).

---

[5] A basis function is an element of a particular basis for a function space. Every continuous function in the function space can be represented as a linear combination of basis functions.

GAMs solve these two problems by first dividing the dependant variable gradient into several segments (interval of data). The segments are delimited by the knots, which are defined as the number of splits and the two end points of the variable space. The number of knots to be selected could be decided manually or through optimisation processes, depending on patterns complexity and avoiding overfitting (Figure 3.3a). The model in Equation 3.28 is then fitted on each segment using a specific smoother (Figure 3.3b). There is actually a large collection of smoothers, and their main difference is in the definition of the basis functions b_js, and also a few differ with respect to the optimisation criterion [288]. Different smoothers can be used in a variety of situations to provide the best fit to a wide range of complex continuous or cyclic trends. For example, a useful smoother is the cyclic cubic regression spline. It ensures that the value of the smoother at the far-left point of the gradient is the same as at the far-right point of the gradient. This is useful if we have a smoother for month (with 12 values); it would not make sense to have a big jump between the January value and the December value for the month smoother. Shrinkage smoothers are also useful since they can have 0 smoothing. This means that then we perform a backwards selection to find the optimal model, all smoothers with 0 amount of smoothing can be dropped simultaneously from the model.

Fitting a smoother provides the βs for each segment, and because we used multiple segments, more complicated patterns than a 4th order polynomial can be fitted. However, there would be discontinuities between the different curves, and each fitted segment would not be nicely joint with the adjacent one (Figure 3.3b). In addition to estimate the optimal degree of smoothness, a cubic regression spline also ensures that the line will look smooth at the points where the individual lines (from the segments) connect through the use relatively complex first-order and second-order derivatives (Figure 3.3c). The amount of smoothing (or the degree of non-linearity) of an estimated smooth term is expressed using the effective degrees of freedom (edf). These are roughly equivalent to the polynomial order of the smoother plus one [288, 319]; however this rule becomes very approximate for edf smaller than 3. A high value (8-10) means that the curve is highly non-linear, whereas a smoother with 1 edf is a straight line.

3.6 Generalised additive mixed model (GAMM)

A generalised additive mixed model (GAMM) is just a GLMM (Section 3.4) with a linear predictor involving a sum of smooth functions of covariates [291]. In general, the model has a structure like:

$$E(Y_i) = \mu_i \quad \text{and} \quad Y_i \sim EF(\mu_i, \varphi)$$
$$g(\mu_i) = \mathbf{X}_i^* \times \boldsymbol{\theta} + f_1(X_{1i}) + f_2(X_{2i}) + \cdots + f_j(X_{ji}) + \mathbf{Z}_i \times \boldsymbol{b}_i + \varepsilon_{ij}$$
$$\boldsymbol{b}_i \sim N(0, \boldsymbol{\Psi}) \qquad \boldsymbol{\varepsilon}_i \sim N(0, \boldsymbol{\Lambda}\sigma^2)$$

Equation 3.29
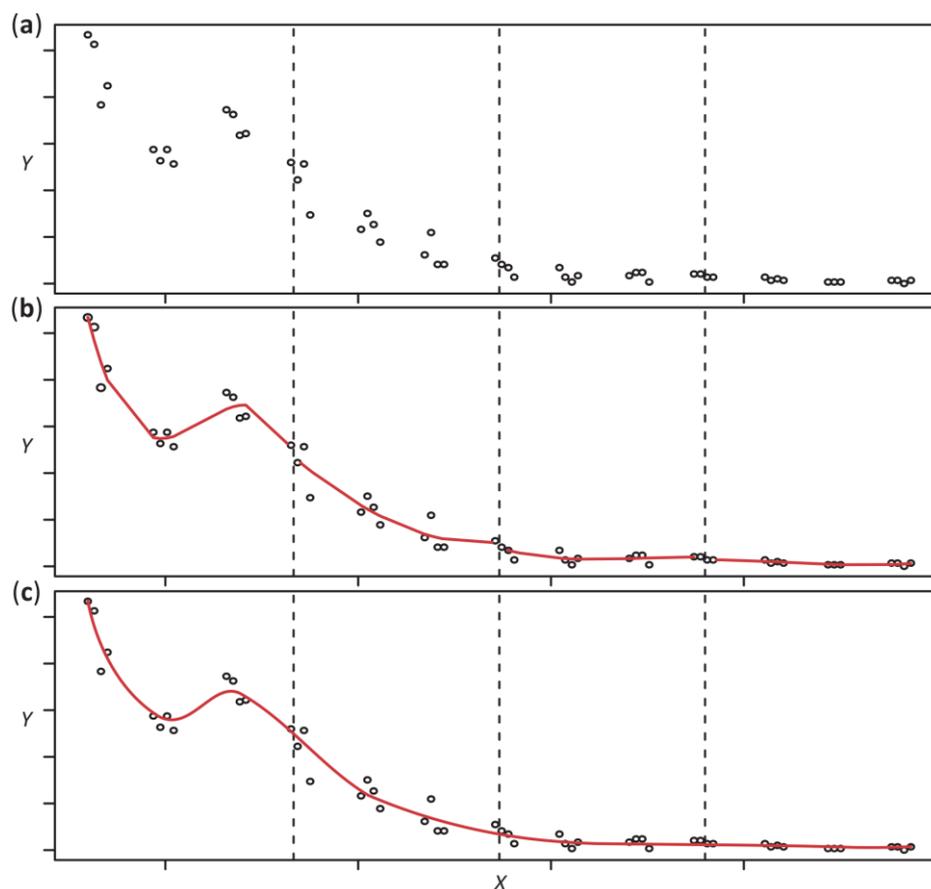
Figure 3.3 Smoothing spline fitting in GAM
Illustration of fitting a cubic polynomial on four segments of data. (**a**) First the dependant variable (*X*) gradient is divided into four intervals (segments) delimited by knots (5 knots: dashed lines plus ending points). (**b**) A GAM with a cubic regression spline is fitted to each segment; each line is the fit of a cubic polynomial model. (**c**) Smoothing spline ensures than discontinuities between different curves are removed by making the prediction smooth at the points where the individual lines connect.

Y_i is the response variable, which follows an exponential family (EF) distribution with mean μ_i and scale parameter φ; $g()$ is a specified link function; $\mathbf{X}_i^*$ is a row of the design matrix $\mathbf{X}_i$ for any strictly parametric model components, $\boldsymbol{\theta}$ is the corresponding parameter vector (so that $\mathbf{X}_i^* \times \boldsymbol{\theta}$ is the linear predictor); f_j are smooth functions of the covariate X_j; $\mathbf{Z}_i$ is a row of random effect design matrix; $\mathbf{b}_i$ is a vector of random effects coefficients with unknown positive definite covariance matrix $\boldsymbol{\Psi}$; ε_{ij} is a residual error vector which is assumed to be normally distributed with expectation 0 and a covariance matrix $\boldsymbol{\Lambda}\sigma^2$. This approach combines both the advantages of GAMs, with their smooth spline definition and flexible estimation of dependence structures, with the GLMMs nature, allowing the experimenter to deal with independence and model correlations among observations from the same block within an experimental unit.

3.7   Model building and optimisation

After having built and applied a statistical model, one must choose among philosophies of optimisation and inference. We may have that diagnostic plots present some residual patterns (e.g. heteroscedasticity, violation of independence), therefore we may be interested in refining our model to optimise its fit of the data. We may also have explanatory variables that have a non-significant effect (e.g. through inspection of *p*-values, F-values, or confidence intervals), and if our aim is to understand which covariates are driving the system we may decide to select and retain only the significant variables. These problems can be solved through model selection, which represents one of the most controversial and important topic in statistical analysis [310]. Deciding which model among a selection of potentially valid models is the best one is complex. Unfortunately, there is not a single, universal procedure because different selection methods are appropriate for different problems and how one analyses data depends strongly on one's philosophical approach (e.g. Bayesian or frequentist). In any case, care is needed.

Model development and selection can be relatively complex and iterative. Here, I present the protocol used for model optimisation and selection procedures used in the following chapters, following the guidance given by Zuur *et al*. [287, 316] and Bolker [320]:

1. Start with a model, the beyond optimal model, that contains all explanatory variables (fixed effects) and interactions of interest. Based on prior knowledge of the dependency structure of the data select the random structure (i.e. random intercept and/or slopes) *a priori*. Many statistical newsgroups have long threads on the subject of which interaction to include in the beyond optimal model [287]. The approaches I used, depending on specific datasets and research question, fell into three categories:

   a. Start with a model with no interactions. Apply the model selection, and model validation, and if there are patterns in the residuals, investigate why. Adding interactions may be an option to improve the model.

   b. Use biological knowledge to decide which, if any, interactions are sensible to add.

   c. Apply a good data exploration to see which interactions may be important.

   Investigate whether the assumptions of normality, homogeneity and independence are valid by making histograms/QQ-plots of standardised residuals, plotting the residuals vs fitted values, and by plotting the residuals versus each individual explanatory variable. If violation of normality, homogeneity, or independence is found go on to step 2. If there is no clear violation of homogeneity, just continue with a model selection on the explanatory variables, step 4.

The beyond optimal model was then optimised by first selecting the optimal **random** structure and then the optimal **fixed** components. The principal tools for model comparison used were the corrected

Akaike Information Criterion (AICc) [321] and bootstrapped likelihood ratio tests [322] (see Section 3.7.1for details) .

2. Fit the model with a restricted maximum likelihood (REML) estimation method. If spatial or temporal covariate are included, make ACF, pACF and variograms to check for temporal or spatial autocorrelation in model residuals. If no autocorrelation is detected go on to step 3. If residual patters are found, select the most appropriate spatial or temporal autocorrelation structure to use for the residuals by comparing different models with different correlations structures, while keeping unchanged the fixed part of the model.

3. Depending on the graphical model validation in step 1, select an appropriate variance structure using a generalised least square method [287, 289] to allow model residuals to vary with respect to predictors (both continuous and categorical). The goal is to identify differences in variance as a function of either one or more continuous or categorical predictors. Compare models with different variance structures to find the most suitable one to use, while keeping unchanged the fixed part of the model. Inspect the residuals for homogeneity (using the same tools as in step 1). If the homogeneity assumption is violated, reiterate this step to find the optimal variance structure. Alternatively try different error distributions (e.g. Poisson or Negative Binomial) or consider a transformation on the response variable as a last resort.

4. After the optimal random component is found (spatial/temporal autocorrelation structure and variance structure) using a REML estimation method, it is time to look for the optimal fixed component of the model. Again, there is an ongoing discussion on how to perform model selection, and different approaches for finding the "optimal" model are available:

    a. Apply a classical model selection using backward (or forward) selection through Information Criteria, such as AIC or AICc.

    b. Apply model selection based on hypothesis testing (i.e. drop the least significant terms).

    c. Apply a model selection on the interaction term, but not on the main fixed terms.

    d. Adopt the Information Theoretic approach following Burnham & Anderson [321], averaging a pre-defined set of candidate models which should be selected *a priori*.

The model is then fitted with a maximum likelihood (ML) estimation method to determine the significance of the fixed effects, while keeping unchanged the random part of the model. The full model is then fitted and compared with nested models[6] (reduced models in which all terms of a smaller model occur in a larger model).

---

[6] A simpler model (with fewer parameters) is **nested** in another, more complex, model (with more parameters) if the complex model reduces to the simpler model by setting some parameters to particular values (often zero).

5. Depending on the specific questions, use a combination of point **a** and **c** (step 4), to remove non-significant interaction terms and retain the main fixed terms.

6. The optimal model is then refitted with REML and validated.

3.7.1 Model comparison

Deciding what models to use and how to use them is fundamentally difficult. In one form or another, this debate goes all the way back to the early Bayesian vs frequentist divide. While statisticians have come a long way in exploring the possible approaches and in providing practical recipes for applying them to a range of situations, there is still (and there never will be) no single best method [320]. Although the approaches for model composition and selection to use are more a combination of one's philosophical approach, statistical common-sense, and lots of care, model selection is based on the concept of parsimony. Parsimony, or "Occam's razor", is a general argument favouring the choice of simpler models even though we know the natural world is complex. In general, we should prefer a simpler model to a more complex one. This is based on the concept that model complexity affects our predictive ability. As we add more parameters to a model, we necessarily get an increasingly accurate fit to the particular data we have observed, but our precision for predicting future observations decreases as well (i.e. the variance of our predictions increases). Data contain a fixed amount of information, and as we estimate more parameters we spread the data "thinner and thinner". Eventually the gain in accuracy from having more details in the model is outweighed by the loss in precision from estimating the effect of each of those details more poorly [320]. Therefore, model selection approaches typically go beyond parsimony to say that a more complex model must be not just better than a simpler model. If the more complex model does not exceed a threshold of improvement in fit, we typically reject it in favour of the simpler model. This threshold can be measured with different methods depending on the approach for model selection used.

Hypothesis testing based on the **Likelihood Ratio test** (LRT) is well-established and widely used. Likelihood methods are based on the concept of maximum likelihood and are used for comparing nested model. The likelihood (L) is a measure of the fit of a model and approximately expresses the probability of a model being a true representation of reality (with higher values indicating a better approximation). This is usually expressed in terms of log-likelihood ($\log L$). There are times when we want a yes-or-no answer when selecting between nested models. We may want to know if some ecological factors are affecting the system in a way that is distinguishable from randomness, and the LRT is appropriate here. The LRT becomes unwieldy when there are many possibly interacting factors or if the effective sample size in mixed-modelling is not large enough (e.g. the smallest number of levels of any random grouping variable in the model is < 40). If inconsistent results from a LRT analysis are found (e.g. if some

parameters are only significant when other parameters are included in the model), we should use another selection approach.

Other approaches are bootstrapping-based methods. Bootstrapping means resampling data with replacement to derive new pseudo-data sets, from which we can estimate confidence intervals. **Parametric bootstrapping** (PB) instead simulates pseudo-data from the fitted model (or from nested models that omit a parameter we are interested in making inferences about). The model is then refitted to these pseudo-datasets to get reliable *p*-values or confidence intervals. PB is very slow (taking hundreds or thousands of times as long as fitting the original model), and it assumes that the model structure is appropriate, and that the estimated parameters are close to the true parameters, but it is an accurate way we can compute *p*-values and confidence intervals for complex models, such as GLMMs. For models with more complex random structures (e.g. crossed random effects) appropriate bootstrapping may be difficult.

One way to avoid having to make pairwise model comparisons is to select models based on **Information Criteria** (IC), which compare all candidate models that are not required to be nested. These relatively recent alternatives to likelihood ratio tests are based on the expected distance between a particular model and the "true" model [321]. In practice, these methods compare the fit and complexity of a model, to find the model with the best predictive ability. Their aim is to find the model that minimises some criterion, in this case the sum of a term based on the likelihood (usually $\log L$), and a penalty term which is different for different information criteria. The Akaike Information Criterion, or AIC, is the most used criterion, and is defined as:

$$\text{AIC} = -2\log L + 2k$$

Equation 3.30

This is defined as twice the difference between the value of $\log L$ (measure of fit) and the number of parameters k used in the model. As with all information criteria, small values represent better overall fits. Essentially, information criteria as the AIC describe as the "best" model, the model that explains most of the variation of the data with the smallest number of parameters. For small sample sizes (n) such as when $n / k < 40$ [321] we should use a finite-size correction and apply the corrected AIC (AICc) instead:

$$\text{AIC}_{\text{c}} = -2\log L + 2k + \frac{2k(k + 1)}{n - k - 1}$$

Equation 3.31

As *n* grows large, the correction term vanishes and the AICc matches the AIC. The AICc was originally derived from linear models with normally distributed errors, so it may apply to a smaller range of models than the AIC, but this is really an open question [320]. Information criteria do not allow significance tests. With IC-method, candidate models are not compare in terms of statistical

significance, such as by using *p*-values as in LRT. Instead, there are commonly used rules of thumb: models with ICs less than 2 apart ($\Delta IC < 2$) are more or less equivalent; those with $3 < \Delta IC < 9$ are clearly distinguishable; and models with $\Delta IC > 10$ are definitely different. One big advantage of IC-based approaches is that they do not require nested models. We can compare all models to each other, rather than using a stepwise approach which is usually subject to a certain degree of subjectivity on which term to drop or retain. In IC-based methods, we simply compute the likelihood and IC for all the candidate models and rank them in order of increasing IC. The model with the lowest IC is the best fit to the data; those models with ICs within 6-8 units of the minimum IC are worth considering [320, 321].

3.7.2 Model validation

Once the optimal model has been found, it is time to apply model validation. Model validation is applied to confirm that the model complies with its underlying assumptions [323]. Violation of these assumptions may result in biased parameter estimates and type I errors [303]. As for the data exploration steps, particular emphasis was given to graphical methods for checking model assumptions. Care is needed with tests as some tests used to assess homogeneity heavily depend on normality. In my analyses, model validation was conducted through the analysis of patterns in model residuals (i.e. standardised or Pearson residuals, depending on the statistical model used). This process consists (as a minimum) of these steps:

- Creation of a histogram or QQ-plot of model residuals to verify the normality assumption. (Note that for Pearson residuals approximate normality is expected, since these residuals are not assumed to follow a Gaussian distribution).

- Plot residuals against fitted values to verify homogeneity of variance. The spread of the residuals around 0 should be consistent along the horizontal axis with no evident pattern (e.g. increasing spread or grouping above/below 0).

- Plot the residuals against each explanatory variable used and against each explanatory variable not used (dropped) in the model to verify independence (or misfit). If patterns are detected there is violation of independence. If omitted explanatory variable are responsible for the pattern, they should be re-included, and the model re-fitted.

- If data include spatial or temporal variables (e.g. measurements were taken over time or at multiple spatial locations), ACF/pACF or variograms should be used to assess independence of residuals.

- Check the model for influential observations using Cook distance function [295].

Additional steps may be required depending on the statistical methods applied. When model validation was completed, and there was no need to refit or implement the model, optimal models were used to make inferences, plotting results, predicting patterns, or simulating from the data.

3.7.3   Software and packages used

Data processing, exploration and statistical analyses were carried out with R [285]. R is a language and environment for statistical computing and graphics which provides a wide variety of statistical and graphical techniques, and is highly extensible. A list of key packages used for statistical analysis and graphing is provided in Supporting Table A.1, in Appendix A.

Chapter Four

4 Blue mussel shell shape plasticity and natural environments: a quantitative approach

4.1 Introduction

Exploring shape variability and uncovering its underlying causes is essential to understand the diversity of life, as well as to appreciate the great heterogeneity of forms that exist in nature [245–247]. Physical constraints are of primary importance in determining the form of an organism as minor variations in growth processes can lead to dramatic shape alterations [245, 274]. Therefore, developing rigorous methods to quantify shapes and describe their natural variation is of critical interest to provide a better understanding of the underlying mechanisms driving the diversity of biological forms.

Bivalves constitute a substantial component of coastal benthic communities [116]. Among them, blue mussels of the genus *Mytilus* are important foundation species throughout the temperate and polar littoral zones of the northern and southern hemispheres [130, 136], and represent an important economic resource for the aquaculture industry [8]. A number of studies have shown a variable distribution of blue mussel species at a North Atlantic scale [130, 131]. In the *Mytilus edulis* species-complex (*Mytilus edulis*, *M. trossulus*, and *M. galloprovincialis*), an extensive hybridisation pattern has been documented

wherever the ranges of these three species overlap [130, 151] and in an aquaculture context [156]. This has a potentially complicated influence on mussel shapes [205, 211, 324].

During the last two decades, much attention has been paid to climate change and its evident effects on calcifying marine organisms [14, 21]. Heterogeneous patterns of environmental variation and increasing anthropogenic pressures have highlighted limitations in our ability to forecast emergent ecological consequences of environmental change [39]. There is, therefore, a clear need for knowledge on the processes regulating marine ecosystems and their resilience [38]. These issues are creating new challenges for understanding organismal responses to key environmental drivers, which is essential for predicting sensitivity to multiple stressors and improving our ability to forecast alterations at higher levels of organisation [38].

Atlantic *Mytilus* spp. have been widely used as model organisms for studying ecological and physiological responses to different environmental conditions [53, 66, 126]. Growing awareness of climate change and its consequences for the considerable biodiversity that blue mussels support [25, 154] have sparked interest in predicting sensitivity of these habitat-forming species [25, 43]. Indeed, the understanding of the significance of morphological variation in *Mytilus* [99, 129] is increasing in parallel with the development of statistical tools to predict species-specific responses [164].

Growth and shape of mussels and the degree to which they vary with respect to environmental factors have been documented for numerous species and habitats [43, 162, 164]. In fact, *Mytilus* shell changes can reflect responses to conditions selecting for specific traits [99, 159, 162] and the level of shape variation may be used as a good indicator of habitat change. Documented shell modifications under forecasted conditions could potentially increase mussel sensitivity to biotic and abiotic drivers [43, 99] and have profound indirect impacts on this foundation species with cascade consequences for supported communities and ecosystems [25, 56, 325]. Therefore, multi-population studies across broad geographical areas, spanning a range of environmental conditions, are critical to identify organismal responses to drivers in a multivariate natural environment [38, 43].

A range of qualitative [159] and quantitative [99, 164, 205, 211] methods have been used to describe the variation in shell traits (morphometrics) and outline (shape) of *Mytilus* in relation to environment and genotype. Standard approaches constitute traditional morphometrics and regression-type analyses [324, 326]. However, their application can result in predictions with poor accuracy of the factors driving shell shape [287] and have implications for the understanding of plasticity.

Traditional morphometrics, which involves applying multivariate analysis to sets of linear descriptors, can mask phenotypic responses [326]. Indeed, mussels can be characterised by variations in shell features that are difficult to quantify (e.g. umbo orientation, convexity of the ventral margin) [207] showing fine-scale shape patterns without alterations of linear shell dimensions. In contrast, the development of geometric morphometrics has emphasised the potential to capture the geometry of the

features of interest [246] and to provide powerful analyses of bivalve shape variation [205, 208, 211, 283]. Unlike ordinary least square methods, newly developed generalised additive mixed models (GAMMs) [288, 291] offer ways to account for the hierarchical structure of ecological datasets, and are powerful tools for defining flexible dependence structures as well as dealing with heterogeneous distributions [327]. However, a combination of these methods and their inferential advantages have never been applied to the study of heterogeneous patterns of organismal shapes in natural environments.

The aims of this study are to **i**) quantify shell shape variation in North Atlantic and Arctic *Mytilus* species from different geographical regions through an elliptic Fourier analysis (EFA) of outlines, **ii**) identify general and local environmental effects on shell shape mean and heterogeneity, through the use of GAMMs and study systems at various geographical scales, **iii**) show how the use of new methods allows the uncoupling of environmental, developmental (age) and genetic (species) contributions to *Mytilus* shape and the description of relationships between blue mussel shape variation and environment that are independent of age and species influences, and **iv**) test the hypotheses that environmental covariates drive the among-individual shell shape variation and environmental stressors can induce the formations of similar shapes at the different geographical scales of analysis. This work further aims to reveal previously unrecognised fine-scale shape responses in Atlantic blue mussels and to estimate effect sizes of different drivers on shape variation. By providing a representative sample for the distribution of blue mussels as well as powerful methods to identify factors influencing shell shape plasticity, it would become possible to appreciate the great variation in mussel forms that exist in nature [159].

4.2 Material and methods

4.2.1 Mussel collection

I collected a total of 16 *Mytilus* spp. populations living along the North Atlantic, Arctic and Baltic Sea coastlines from three study systems at different geographical scales (large-, medium- and small-scale, Supporting Table A.2). I analysed shell shape variation among habitats across a large geographical scale, System 1, on ten wild blue mussel populations (sites 1-10; Figure 4.1a) sampled at different latitudes from four distinct climatic regions (warm temperate, cold temperate, subpolar and polar). Mussel specimens were collected from Western European (Exmouth, England, 50°N) to Northern Greenlandic (Qaanaaq, 78°N) coastlines, covering a latitudinal range of 28° (a distance of 3,980 km). Environmental influence on a medium spatial scale, System 2, was investigated using five wild mussel populations (sites A - E; Figure 4.1b) collected from the North Sea (Sylt, Germany) to the innermost part of the Baltic Sea (Nynäshamn, Sweden). In addition, I studied shell shape variation on a small
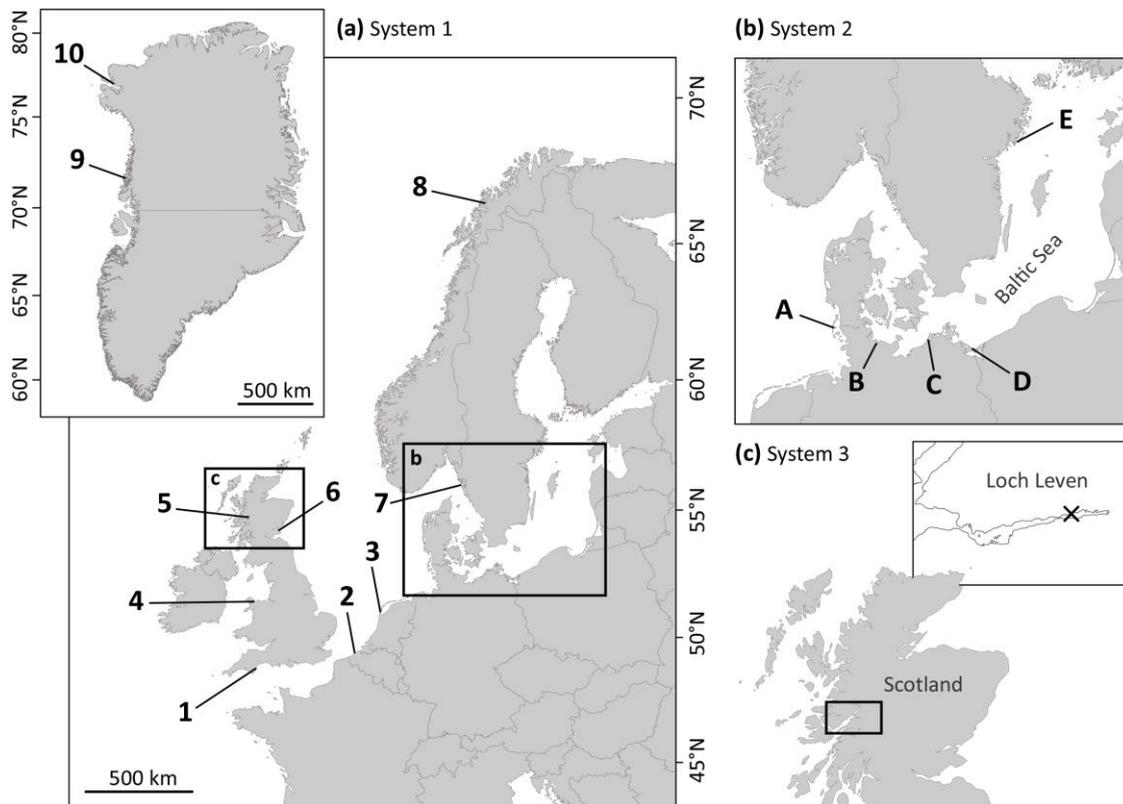
Figure 4.1 Blue mussels collection sites

Study systems and locations where *Mytilus* populations were collected. (**a**) System 1, large-scale North Atlantic and Arctic regions: (1) Exmouth, south-west England, (2) Oostende, Belgium, (3) Texel, north Netherlands, (4) Menai Bridge, north Wales, (5) Tarbet, Kintyre, west Scotland, (6) St. Andrews, east Scotland, (7) Kristineberg, west Sweden, (8) Tromsø, north Norway, (9) Upernavik and (10) Qaanaaq, west Greenland. (**b**) System 2, medium-scale Baltic region: (A) Sylt, (B) Kiel, (C) Ahrenshoop, (D) Usedom, all Germany and (E) Nynäshamn, east Sweden. (**c**) System 3, small-scale: (X) long-line mussel farm (Glencoe Shellfish Ltd.) in Loch Leven, west Scotland, with four sampling depths (I, III, V and VII meters). Map created with ArcMap 10.3 (ArcGIS software by Esri, http://www.esri.com), background image courtesy of OpenStreetMap (http://www.openstreetmap.org).

geographical scale, System 3, using specimens obtained from a traditional longline mussel farm on the Scottish west coast (Loch Leven, UK; Figure 4.1c). Four batches of mussels, originating from a natural spatfall, were collected at one, three, five and seven metres depth (batches I, III, V and VII, respectively), representing the natural distribution of mussels along the cultivation ropes.

During December 2014 and January 2016, I collected mussels of various size classes for each population (shell length 25 - 81 mm) for a total of 555 individuals. Wild adult mussels (System 1 and 2) were sampled from the eulittoral zone and cultured specimens (System 3) were harvested as part of a long-term monitoring programme. For each specimen, I measured shell dimensions with digital callipers (0.01 mm precision) (Figure 2.10a, b), among which shell length was used as a within-population proxy for age [116, 162, 209].

I examined *Mytilus* populations with available information on their genotype, with a particular focus on species identity and documented hybridisation (*Mytilus edulis*, *M. trossulus*, and *M. edulis* × *M. trossulus* hybrids). Blue mussels used were from populations recently analysed in genetic investigations, sites routinely employed in regional monitoring programmes, and specimens already used for genetic analyses (Supporting Table A.2). Therefore, I used populations with a known genetic status.

Reference populations of *Mytilus edulis* and *M. trossulus* were selected from two sites in western Europe, one site in Greenland and one Baltic location (populations 1, 4, 10 and E, respectively; Figure 4.1a, b). According to genetic analyses of these populations, which are based on multiple genetic markers (multi-locus genotyping) or SNP analyses, these samples are representative of these two species [136, 144, 149]. Although molecular studies have revealed various episodes of introgression and hybridisation, which increases the evidence that no completely pure reference populations exist in the North Atlantic and Baltic Sea [151, 205], these reference samples are as representative as possible considering the geographical range of the study. Therefore, these provide a solid starting point for the following among-species shell shape comparisons. Given the absence of representative populations of *M. galloprovincialis* at the analysed spatial scale, I avoided areas where this species was either present or there was a high degree of hybridisation (e.g. south-central Norway, parts of continental European and Ireland's coastlines) [133, 141]. I did, however, sample sites where very low proportions of *M. edulis* × *M. galloprovincialis* hybrids have been reported [131].

4.2.2 Environmental parameters

I selected environmental covariates according to the availability of data for the investigated areas and their known effects on growth, development, and mussel energy budget [43, 127, 162]. Given the high collinearity of many physical and biogeochemical descriptors at the geographical scale considered, I chose three key parameters: water surface temperature, salinity and chlorophyll-a (Chl-*a*) concentration, the latter being validated as a proxy for food availability [53, 164]. Predictors for the large- and medium-scale systems (System 1 and 2) were generated using the Copernicus Marine Environment Monitoring Service (CMEMS) [328]. These datasets are composed of high-resolution physical and biogeochemical analyses of assimilated (integration of observational and predicted information) daily data ($n = 730$ per parameter; Appendix B, Environmental Datasets). For each parameter, mean values per site for the 2014 - 2015 period were used as predictors. For the large- and medium-scale systems, remote-sensing and assimilated data presented potential advantages compared to traditional measurements [164, 329] due to their known high spatial and temporal resolution, advanced calibration and validation (i.e. high correlation with discrete field measurements) [328, 330]. Environmental parameters for the small-scale

system (System 3) were calculated from samples collected fortnightly over the course of a year and expressed as annual mean values for each depth.

4.2.3 Elliptic Fourier analysis of shell outlines

Shape analyses of *Mytilus* shells were performed through a geometric morphometrics approach [246]. I used an elliptic Fourier analysis (EFA) of outlines [274, 276] to examine shell shape variation both within and between populations from different study systems. This elliptic Fourier method presents several advantages compared to older approaches [274]: complex shapes can be fitted, outlines smoothed, starting points normalised, and homothetic, translational and rotational differences removed [258, 274, 278, 283] (Section 2.4.3).

Outlines of orthogonal lateral and ventral views of the left valves were digitised, converted into a list of *x-y* pixel coordinates and used as input data. The outlines for both views were then processed independently, geometrically aligned and later combined for analysis following the protocol in Sections 2.4.4-2.4.8, (Figure 2.10c). I then computed an EFA on the resulting coordinates from shapes invariant to outline size and rotation. After preliminary calibration, I chose seven harmonics, encompassing 98% of the total harmonic power [283]. Four coefficients per harmonic were extracted for each shell outline (28 descriptors) and used as variables quantifying the geometric information [258].

A principal component analysis (PCA), with a singular value decomposition method, was performed on the matrix of coefficients to observe shape variation among individuals and populations from the different study systems. Calculated principal components (PCs) were considered as new shape variables. To understand the contribution of individual variables to shell shape, I reconstructed extreme outlines along each PC. The first 10 PCs, accounting for 97% of outline variation, were analysed with a multivariate analysis of variance (MANOVA) to test for significant effects of the location of origin and shell length (size) on shape variances. To visualise differences at the extremes of the morphospace, I generated deformation grids [245] and iso-deformation lines through mathematical formalisation of thin plate splines (TPS) analysis [259].

For the reference populations of *M. edulis* and *M. trossulus*, I performed a linear discriminant analysis (LDA) based on the new shape variables (PCs), with a leave-one-out cross-validation procedure, to identify the linear combination of shape features that was able to discriminate between *Mytilus* species. Standardised coefficients from the calculated discriminant function were used to compare the relative importance of each shape variables (PCs) at discriminating between species. I set *a priori* classification probabilities to be proportional to group sizes and Wilks' λ were calculated to test for significant discrimination. I estimated discriminant coefficients to identify shell shape features that optimised the between-species differences "relative" to the within-species variation [258].

4.2.4 Data exploration and statistical modelling

I selected for analysis the first five PCs, capturing 91% of shape variance and describing distinguishable features along the outline. I used generalised additive mixed models (GAMMs) [291, 327] to explain shape variance with respect to mean environmental parameters and shell size, and to compare between individual shape features (PCs). Given the dependency on the same set of predictors, I analysed all the PCs for each study system within the same model. This new approach allows **i**) accounting for the dependence of multiple shape variables, which describe synergistically the shell outline as a whole (as implied by the adopted EFA method), and **ii**) defining combinations of linear and non-linear relationships simultaneously. The among-individual shape variances were then analysed together without losing descriptive power or increasing the probability of Type I error [287].

I carried out data exploration following the protocol in Section 3.1.1. Initial inspection revealed no outliers. Conditional boxplots showed heterogeneous shape variances (PCs eigenvalues) as a procedural consequence of the PCA. This required standardisation prior to analysis since I was not interested in between-feature heterogeneity [287]. Response variables did not require any transformation. Pairwise scatterplots and calculation of variance inflation factors (VIFs) [294] indicated low collinearity between predictors for System 1 and 2. For these systems, the effects of multiple environmental covariates on shape variance were modelled simultaneously only if VIFs < 3 [294]. I detected high collinearity among environmental predictors for System 3. Therefore, I performed a PCA on these explanatory variables to calculate new linear combinations of covariates accounting for the greatest variation in the original values [297]. I then used scores of orthogonal PCs (enviro-PC1, PC2 and PC3) as new independent environmental predictors to model shape variance. In addition, potential interactions between continuous covariates and shape features, and clear non-linear patterns were detected.

I used GAMMs to model shape variance for **i**) large-scale (System 1), **ii**) medium-scale (System 2), **iii**) small-scale (System 3) study systems and **iv**) the pooled mussel populations (Atlantic system; Equation 4.1). I employed a combination of a single question approach (individual systems) and an analysis of the pooled populations (Atlantic system) to model and differentiate local environmental effects, being more dependent on the geographical scales considered, from the general effects of environmental variation, having a more consistent influence on the shell shape of blue mussels from different regions.

To model shape variance as a function of environmental covariates for the Atlantic system, I used a GAMM with a normal distribution (Equation 4.1). Fixed continuous covariates used were water *temperature*, *salinity*, and *Chl-a* concentration all fitted as smoothers, in addition to *shell length* (continuous), shape features (*PC*, categorical with five levels: PC1, …, PC5), and their interactions with continuous predictors. To incorporate the dependency among specimens from the same site of collection, I used *site* as a random effect. The final model was of the form:

$$ShapeVar_{ijk} \sim N(\mu_{ijk}; \sigma^2)$$
$$\mu_{ijk} = f(Temperature_i) \times PC_j + f(Salinity_i) \times PC_j + f(Chl\text{-}a_i) \times PC_j + Length_{ik}$$
$$+ PC_j + Length_{ik} \times PC_j + Site_i$$
$$Site_i \sim N(0; \sigma^2_{Site})$$

Equation 4.1

where *ShapeVar*$_{ijk}$ is the kth observation for j^{th} PC (j = 1, …, 5 levels) and i^{th} site (i = 1, …, 15 levels). f is the smoothing function and *Site*$_i$ is the random intercept, which is assumed to be normally distributed with expectation 0 and variance σ^2_{Site}. The *f(continuous predictor) $\times$ PC* interaction applies a smoother on the data for each PC.

I manually selected the optimal amount of smoothing and a cubic regression spline was applied [291]. Variograms indicated no spatial or temporal autocorrelation. Statements about trends of shape variance with environmental gradients are based on the significance (at α = 0.01) of individual interaction terms between predictors and PCs. Models were optimised by first selecting the random structure and then the optimal fixed component [287, 323], following the protocol in Section 3.7. Visual inspection of residual plots indicated a violation of homogeneity in most cases. This required the use of specific variance structures (generalised least squares) allowing the residual spread to vary with respect to continuous predictors and shape features [287]. Once I found the optimal model (in terms of the random structure), I applied further selection by rejecting any non-significant interaction term between the explanatory variables. The principal tools for model comparisons were the corrected Akaike Information Criterion (AICc) and likelihood ratio tests for each nested model. Final models (Table 4.1) were validated by inspection of standardised residual patterns to verify the assumptions of normality, homogeneity and independence [287], as in Section 3.7.2. I used models to predict trends with environmental gradients and estimate the mean effect sizes (same measurement unit) of standardised environmental parameters. For standardisation, I subtracted the sample mean from the variable values and divided them by the sample standard deviation [$z_i = (x_i - \bar{x}) / \sigma_x$]. Confidence intervals (95% CI) and mean effect sizes were estimated to compare the magnitude of the effect of individual covariates on the responses. If the confidence intervals did not overlap with zero, the effect size was considered significant.

Table 4.1 Optimal models
Best models after selection of random part (variance structures) and fixed components for each study systems are reported.

Selected final models

System 1

$$ShapeVar_{ijk} = f_j(Temp_i) + f_j(Sal_i) + \ + \beta_{1j}Chla_i + \beta_{2j}Length_{ik} + Site_i + \varepsilon_{ijk}$$
$$Site_i \sim N(0; \sigma^2_{Site})$$
$$\varepsilon_{ijk} \sim N(0; \sigma^2 \times e^{2\delta_j \, Temp_i})$$

System 2

$$ShapeVar_{ijk} = f_j(Sal_i) + \beta_1 Chla_i + \beta_2 Length_{ik} + Site_i + \varepsilon_{ijk}$$
$$Site_i \sim N(0; \sigma^2_{Site})$$
$$\varepsilon_{ijk} \sim N(0; \sigma^2 \times e^{2\delta_j \, Chla_i})$$

$$ShapeVar_{ijk} = f_j(Temp_i) + \beta_{1j}Chla_i + \beta_2 Length_{ik} + Site_i + \varepsilon_{ijk}$$
$$Site_i \sim N(0; \sigma^2_{Site})$$
$$\varepsilon_{ijk} \sim N(0; \sigma^2 \times e^{2\delta_j \, Chla_i})$$

System 3

$$ShapeVar_{ijk} = f_j(enviro.PC1_i) + Batch_i + \varepsilon_{ijk}$$
$$Batch_i \sim N(0; \sigma^2_{Batch})$$
$$\varepsilon_{ijk} \sim N(0; \sigma^2 \times e^{2\delta_j \, enviro.PC1_i})$$

Atlantic system

$$ShapeVar_{ijk} = f_j(Temp_i) + f_j(Sal_i) + f_j(Chla_i) + \beta_{1j}Length_{ik} + Site_i + \varepsilon_{ijk}$$
$$Site_i \sim N(0; \sigma^2_{Site})$$
$$\varepsilon_{ijk} \sim N(0; \sigma^2 \times e^{2\delta_j \, Temp_i})$$

4.3 Results

4.3.1 Mussel geometric morphometrics

The first two PCs, from PCAs performed on harmonic coefficients, accounted for 68.7%, 70.0% and 66.1% of the shape variation among individuals from Systems 1, 2 and 3 respectively, and 70.2% of variance for the Atlantic system. Scatterplots of PC1 and PC2 showed a clear separation among groups across the morphospace (Figure 4.2a - c, 4.3a), revealing marked among-individual variation for both lateral and ventral views. For the Atlantic system, MANOVAs revealed significant effects of collection site (Wilk's $\lambda = 0.032$, approximate $F_{1,419} = 12.611$, $p < 0.0001$) and shell length (Wilk's $\lambda = 0.873$, approximate $F_{1,419} = 5.95$, $p < 0.0001$) on shape variance. Additionally, significant influences of location of origin and shell length at different geographical scales were identified (Table 4.2). Mean shapes and TPS analyses indicated the main outline deformations required to pass from one extreme of the morphospace to another (Figure 4.2d, 4.3b, 4.4).
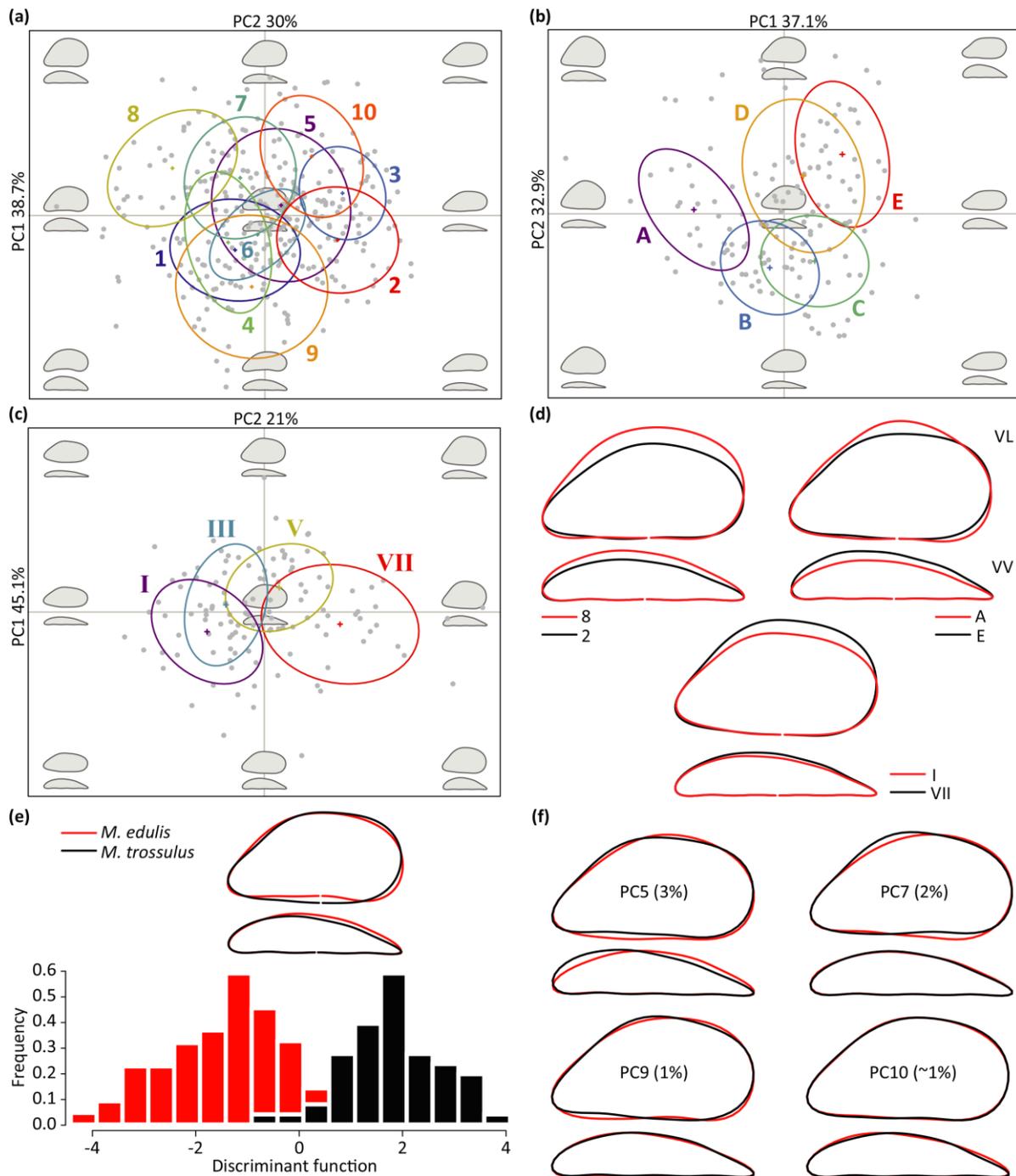
Figure 4.2 Among-population variations in outlines and shape features

Scatterplots of the first two PCs from PCAs performed on the elliptic Fourier coefficients, of lateral and ventral shell views, showing a clear separation and marked shape variation among specimens from (**a**) System 1 (large-scale), (**b**) System 2 (medium-scale) and (**c**) System 3 (small-scale). Confidence intervals for each group of origin (ellipses) and the reconstructed morphospace (background) are shown. (**d**) Mean shape differences of lateral (VL) and ventral (VV) views between populations or batches at the extremes of the morphospace: System 1 populations 8 - 2, System 2 locations A - E, and System 3 batches I - VII. Population 8 had rounder and wider shells, with higher and more convex ventral sides than population 2 (elongated and narrow shells). Location A was characterised by rounder and higher shells with a more convex ventral side than site E (elongated and wide

shells with almost parallel dorsoventral margins). Mussels from batch I displayed more elongated shells, with a smaller height, ligament area and width than mussel from batch VII (round shells). All the reconstructed mean outlines showed a consistent variation in distinct shell features: shell height, ventral side shape, ligament area and shell width. (**e**) Discriminant function calculated from a LDA on the new shape variables showing a significant separation between species (Wilk's $\lambda = 0.264$, approximate $F_{1,113} = 29.00$, $p < 0.0001$) and differences between the mean shell outlines for the individual groups (red: *Mytilus edulis*; black: *M. trossulus*). (**f**) Shape variables (PC5, 7, 9 and 10) contributing the most to discriminate between species. These PCs captured the species-contribution to the shell shape of *Mytilus* (7% of total shape variance). Individual contributions to the mean outline are represented through the reconstruction of mean shapes for high (red) and low (black) PCs values (Mean ± 3 SD, respectively).

I identified shape features that contributed the most to the observed patterns of shape variation for different systems through comparison of reconstructed outlines at the extremes of the morphospace along each axis. The first five PCs, depicting the variation in specific shell features, were described through their individual contribution to the outline reconstruction for increasing PC values, for individual study systems (Figure 4.5, 4.6a, Supporting Table A.3) and the Atlantic system (Figure 4.6b, Supporting Table A.3).

A LDA based on the new shape variables allowed the identification of the shape features that discriminate most between the two *Mytilus* species and to isolate the species-contribution to the shape variance. Ninety-seven percent of individuals were correctly reclassified by the new discriminant function (Figure 4.2e). The LDA on the reference populations produced an efficient separation between groups and a cross-validation (leave-one-out) at species level showed a high confidence in the reclassification (98.3% and 94.6% of correct reattribution for *M. edulis* and *M. trossulus*, respectively). Standardised discriminant coefficients indicated PC5 (3%), PC7 (2%), PC9 (1%) and PC10 (~1%) had the highest contribution to the separation between species. I selected these PCs as the variables capturing most of the shape information on the species-contribution to *Mytilus* shell shape. The identified variables contributed to subtle variations in shell outlines (Figure 4.2f) and showed limited overlap with the shape features described by the PCs capturing the most of shape variance among individuals (PC1 - PC4).

4.3.2 Shell shape variation and environmental factors

GAMMs indicated highly significant relationships between the axes capturing most of the shape variation and environmental parameters for all the study systems, with associations depending on the shape features (PCs) and system considered. Only significant relationships ($p < 0.01$) are presented in the following section.

Table 4.2 MANOVA output
Summary of MANOVA models on the first 10 calculated shape variables (PCs) for each study system, showing the effects of location of origin and shell length (size) on shape variance.

Factor	*df*	Wilk's λ	F	num. *df*	den. *df*	*p*-value
System 1						
Site	9	0.059	10.88	90	1889	**<0.0001**
Length	1	0.849	4.94	10	277	**<0.0001**
System 2						
Site	4	0.049	14.20	40	468	**<0.0001**
Length	1	0.872	1.80	10	123	0.068
System 3						
Batch	3	0.186	2.17	30	311	**<0.0001**
Length	1	0.830	8.04	10	106	0.025
Atlantic system						
Site	14	0.032	12.61	140	3381	**<0.0001**
Length	1	0.873	5.95	10	410	**<0.0001**

4.3.2.1 *System 1 – Large geographical scale*

The model (Supporting Table A.4) showed positive and negative non-linear relationships of PC1 with temperature and salinity, respectively. I detected associations of PC2 with salinity, shell length and Chl-*a*. PC3 and PC4 showed marginal negative relationships with temperature and Chl-*a*, respectively. Overall, I observed the formation of elongated and narrow shells with decreasing temperature and salinity (Figure 4.7a), and a transition from elliptical to more elongated, curved, and wider profiles with increasing food supply and shell length (Figure 4.8). An exponential variance structure indicated a negative effect of water temperature ($df = 5$, L = 39.82, $p < 0.0001$) on shape variance.

4.3.2.2 *System 2 – Medium geographical scale*

GAMMs (Supporting Table A.4) indicated a non-linear association between PC1 and salinity only. PC2 and PC5 showed negative relationships with salinity and temperature. The model revealed a general positive effect of shell length ($df = 1$, $F_{1,126} = 7.75$, $p = 0.0055$). Overall, I observed more elongated, wide shells and more squared margins with decreasing salinity and temperature (Figure 4.9a). Round mussels with large ligament areas were associated with high salinities (~30 psu), while elongated, wide shells were characteristic of low salinities (~6 psu) (Figure 4.7b). An exponential variance structure indicated a positive effect of Chl-*a* concentration ($df = 5$, L = 52.05, $p < 0.0001$) on shape variance.

4.3.2.3 *System 3 – Small geographical scale*

Model selection reported significant effects of enviro-PC1 only along the cultivation rope (Supporting Table A.4). PCA indicated an equal positive contribution of water temperature and salinity, and a negative contribution of Chl-*a* concentration to the enviro-PC1 loadings. The optimal model showed a positive non-linear relationship with PC1, a marginal increasing trend with PC2, and a non-linear
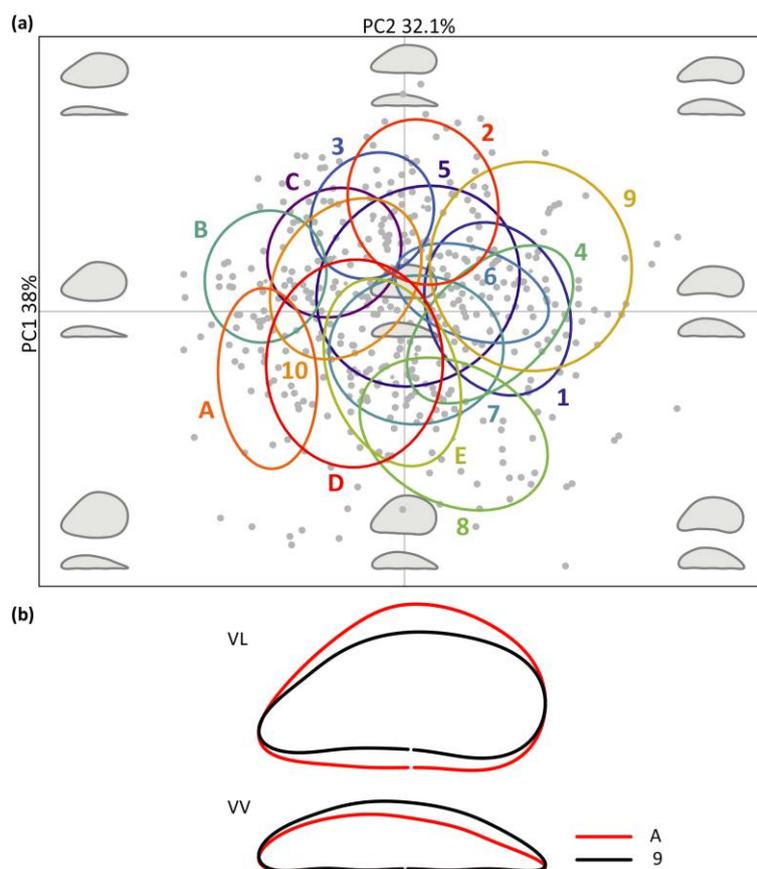
Figure 4.3 Variation in outlines and shape features (Atlantic system)
(**a**) Scatterplots of the first two PCs from a PCA performed on the elliptic Fourier coefficients, of lateral and ventral shell views, showing a clear separation and marked shape variation among specimens from the pooled locations. Confidence intervals for each group (ellipses) and the reconstructed morphospace (background) were represented. (**b**) Mean shape differences of lateral (VL) and ventral (VV) views between populations at the extremes of the morphospace. Population A showed rounder and narrower shells, with bigger height and more convex ventral sides than population 9 with elongated, curved and wide shells.

association with PC5 (Figure 4.9b). Overall, shell height, width, and ligament length progressively increased with increasing values of enviro-PC1, showing a transition from elongated and narrow to round and wide mussel shells with increasing temperature, salinity, and decreasing food availability (Figure 4.7c).

4.3.2.4 Atlantic system

Equation 4.1 indicated relationships between blue mussel shape and all the modelled predictors (Table 4.3, Figure 4.10a). PC1 showed positive relationships with temperature and shell length, and non-linear patterns with salinity and food availability. PC2 indicated non-linear relationships with temperature and Chl-*a*, a negative association with salinity and a positive one with shell length. I detected positive associations of PC3 with both temperature and salinity. PC4 was characterised by a positive relationship with temperature and non-linear association with salinity and food availability. Overall, I identified the

formation of elongated, narrow shells and more squared margins with decreasing salinity, an increasing shell height and width with increasing Chl-*a* and a transition from elliptical to elongated, curved and wider shells with increasing temperature and shell length (Figure 4.10a). I specified exponential variance structures [287] allowing residuals to vary with respect to surface temperature and shape features (PCs). The best variance structure indicated a negative exponential effect of temperature (df = 5, L = 59.65, p < 0.0001) on shape variance. Mean effect sizes revealed differences in the relative contribution of modelled covariates (Table 4.3, Figure 4.10b). PC1 was characterised by a marked effect of shell length and environmental influences of temperature and Chl-*a*. Water salinity had the strongest effect on PC2, being about three times bigger than the effect of shell length. I also identified a marked influence of temperature, salinity and a weak effect of length on PC3, while PC4 was strongly influenced by all the environmental descriptors. I detected no effect on PC5.
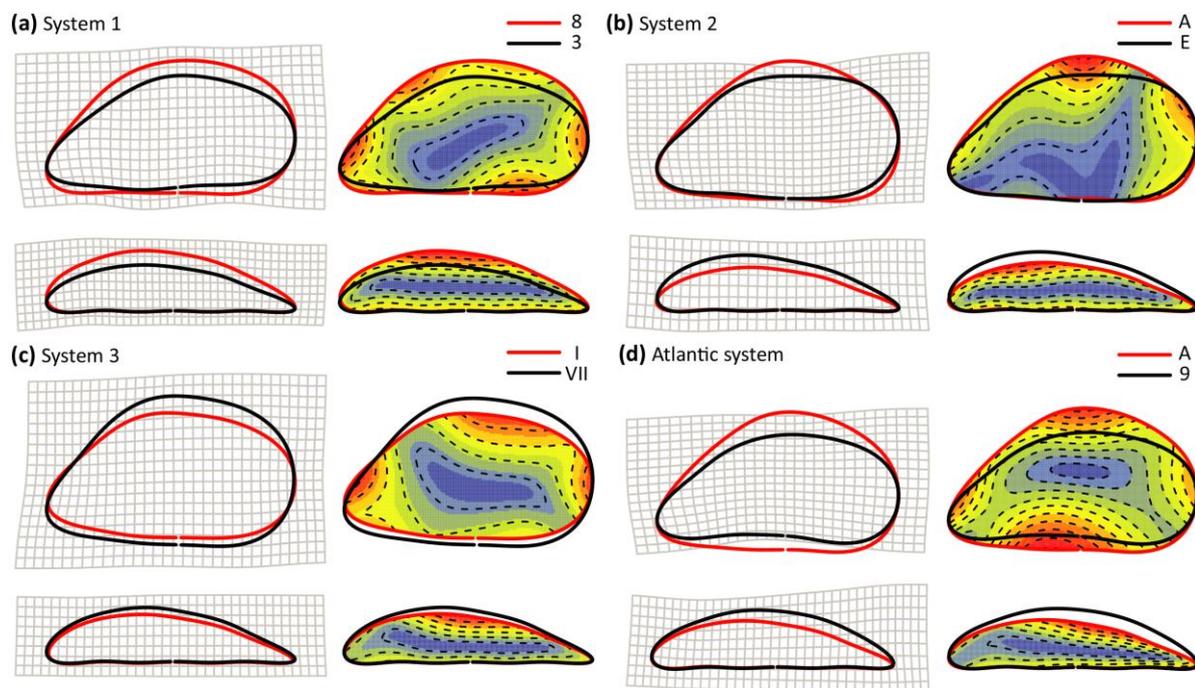
Figure 4.4 Mean shape differences at the extremes of the morphospace
Deformation grids (left), depicting the bindings required to pass from an extreme of the morphospace to another, and iso-deformation lines (right), representing the outline regions subjected to different degrees of change (blue: low deformation; red: strong deformation), for (**a**) System 1, populations 8 and 3, (**b**) System 2, populations A and E, (**c**) System 3, batches I and VII, and (**d**) Global system, populations A and 9.
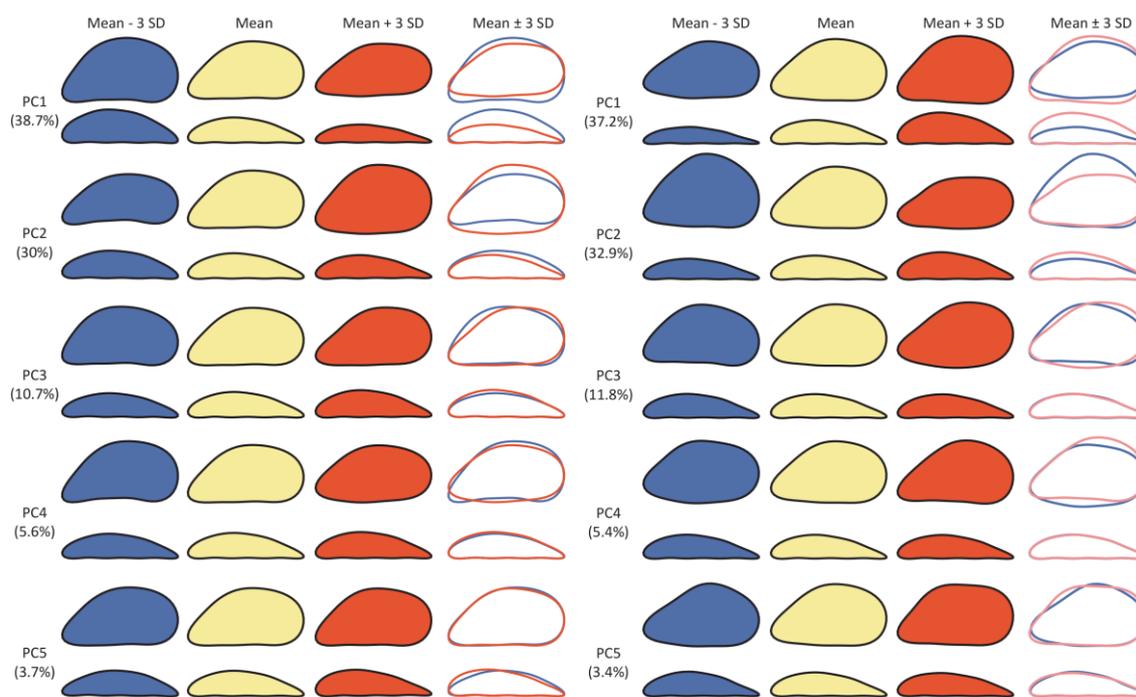
Figure 4.5 PCs contribution to shape reconstruction (System 1 and 2)
Contribution of the first five shape variables (PCs) to shape variation (large-scale and medium-scale study systems). The average shell shapes, for both lateral and ventral views, were represented for increasing values along each PC (Mean − 3 SD, Mean, Mean + 3 SD) and extreme shapes were compared (Mean ± 3 SD).
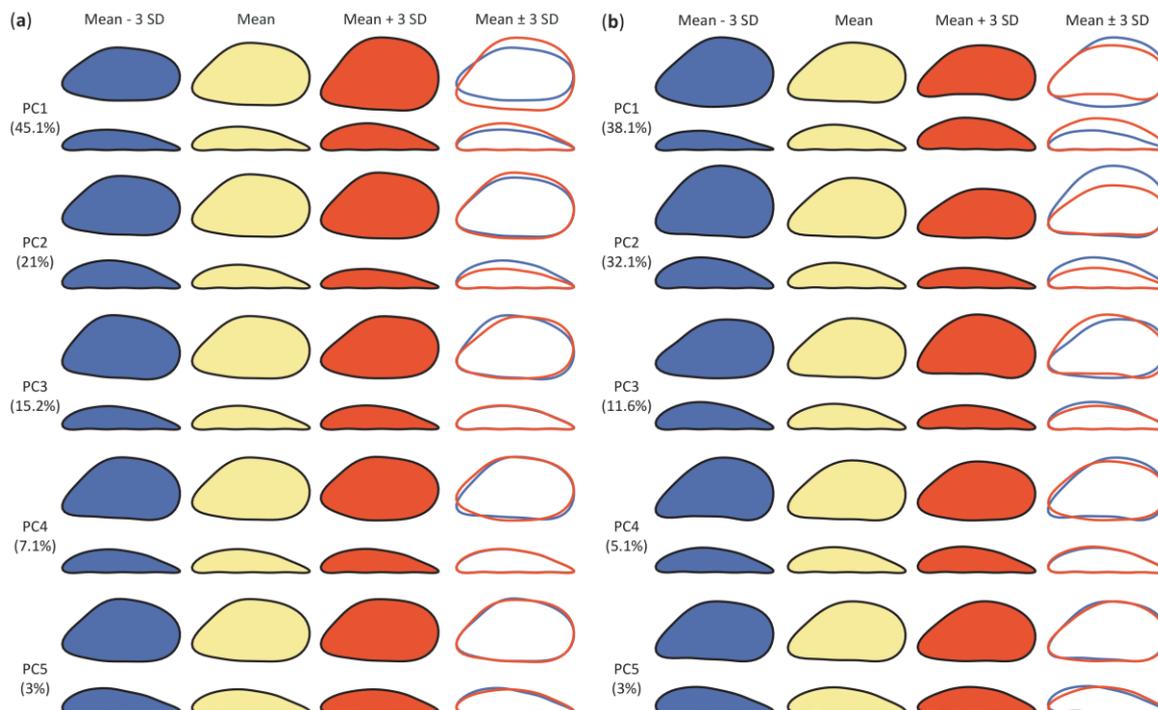
Figure 4.6 PCs contribution to shape reconstruction (System 3 and Atlantic system)
Contribution of the first five shape variables (PCs) to shape variation (small-scale and Atlantic study systems). The average shell shapes, for both lateral and ventral views, were represented for increasing values along each PC (Mean − 3 SD, Mean, Mean + 3 SD) and extreme shapes were compared (Mean ± 3 SD).

4.4 Discussion

Shape analysis is a fundamental component of several areas of biological research [246]. In ecology, it can allow discrimination of shapes of organisms from specific habitats and understanding of the underlying mechanisms leading to variation of morphological structures [274]. This is especially important for economically and ecologically valuable taxa, such as blue mussels [126, 130]. With regards to aquaculture, shape variations under changing environments could produce fragile shelled mussels [99]. These are less valuable economically [331], being easily damaged during harvest, grading, and transport processes, and may lead to significant financial losses for the industry [155]. In natural habitats, changes in shapes and structural integrity of shells can increase their vulnerability to predation [83, 99], with potential profound cascade impacts on whole ecosystems [25].

Geographical variation in *Mytilus* shell shape is confounded by marked shell modifications during growth [159] and among-species differences [205, 211], on top of which environmental heterogeneity strongly influences spatial shape patterns [162]. Several studies have explored the effects of these individual factors on natural shape variations in different mussel species. Seed [159, 209] investigated the influence of growth rate and age on *M. edulis* form, providing a qualitative baseline for the interpretation of its developmental changes. The effect of genotype on shell shape and morphology has also been explored for the Atlantic mussels, showing differences among taxa in various geographical regions [155, 205, 324]. Modelling was used to identify the relationships between mussel growth and environmental factors across relatively small spatial scales [164], while broad-scale studies have highlighted consistent morphometric patterns along latitudinal gradients in the South Pacific [207, 208]. In addition, experimental induced phenotypic shape responses have indicated potential deleterious effects of future increases in temperature and pCO_2 on shell integrity of *M. edulis* [99]. Moreover, a body of research showed many factors have more local influences on shape, such as hydrodynamic regimes, ice cover and parasitic diseases [130, 324, 332]. However, our ability to forecast heterogeneous patterns of mussel shape responses to altered environmental conditions in multi-population studies is limited by our ability to uncouple the contributions of developmental (age) and genetic (species) factors from shell shape variations. Specifically, heterogeneous size classes and multiple species prevent us from identifying general relationships between *Mytilus* shape variation and local drivers without selectively controlling for these two confounding factors (i.e. analysing similarly sized individuals and/or different species separately).

In this study, the combination of EFA, GAMMs, and multiple systems on different spatial scales allowed the identification of shell features under control of age and species factors and allowed me to uncouple these from the modelled shape variance to describe independent general and more local relationships between Atlantic *Mytilus* shape and natural environments.
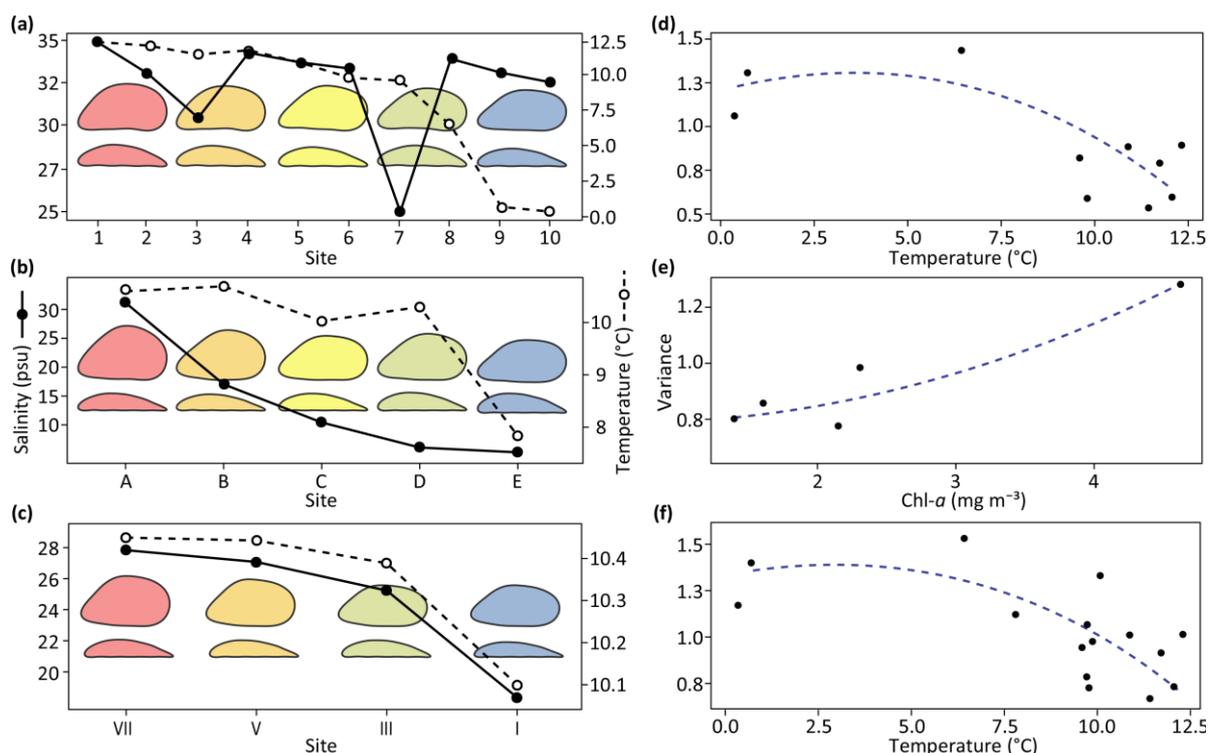
Figure 4.7 Response of *Mytilus* shape and heterogeneity to environmental variation
Left: graphs show a marked convergence of average shell shapes for the individual study systems. (**a**) System 1, large-scale North Atlantic and Arctic populations (only populations 1, 4, 5, 9 and 10 are shown in the background), (**b**) System 2, medium-scale Baltic region and (**c**) System 3, small-scale Loch Leven. The convergence of mean shell outlines at different geographical scales indicates the consistent formation of elongated, narrow shells and more parallel dorsoventral margins under lower temperature and salinity. Right: graphs represent system-wise patterns of shape heterogeneity with habitat conditions estimated from optimal variance structures within individual GAMMs. Loess smoothers (dashed lines) are fitted for visual interpretation. (**d**) The range of shell variation in North Atlantic and Arctic populations (System 1) showing the formation of more heterogeneous shapes in colder waters. (**e**) Positive trend of shape variance in the Baltic region (System 2) depicting more heterogeneous shell shapes with increasing food levels. (**f**) Pattern of shape variance for the Atlantic system showing an increase of shape heterogeneity with decreasing water temperature.

4.4.1 Quantifying environmental effects on shell shape

Environmental influence on mussels is complex, with interacting factors that may result in a variety of shape patterns [162]. These interactions make it problematic to isolate effects of single drivers in natural environments and constrain predictive power [38, 333]. Among these drivers, population density and predation influence responses in blue mussels including changes in shell proportions [159] and structure [198, 334]. Genotypic differences and hybridisation within the *Mytilus edulis* species-complex are also known to influence spatial patterns of shell variation [130, 151, 205, 211]. Moreover, although I considered the annual mean of environmental parameters, other factors could have substantial effects on mussel growth and shape, such as seasonality and the environmental conditions during specific life-stages [43, 78]. However, it is not always possible to include all the interacting drivers at the different
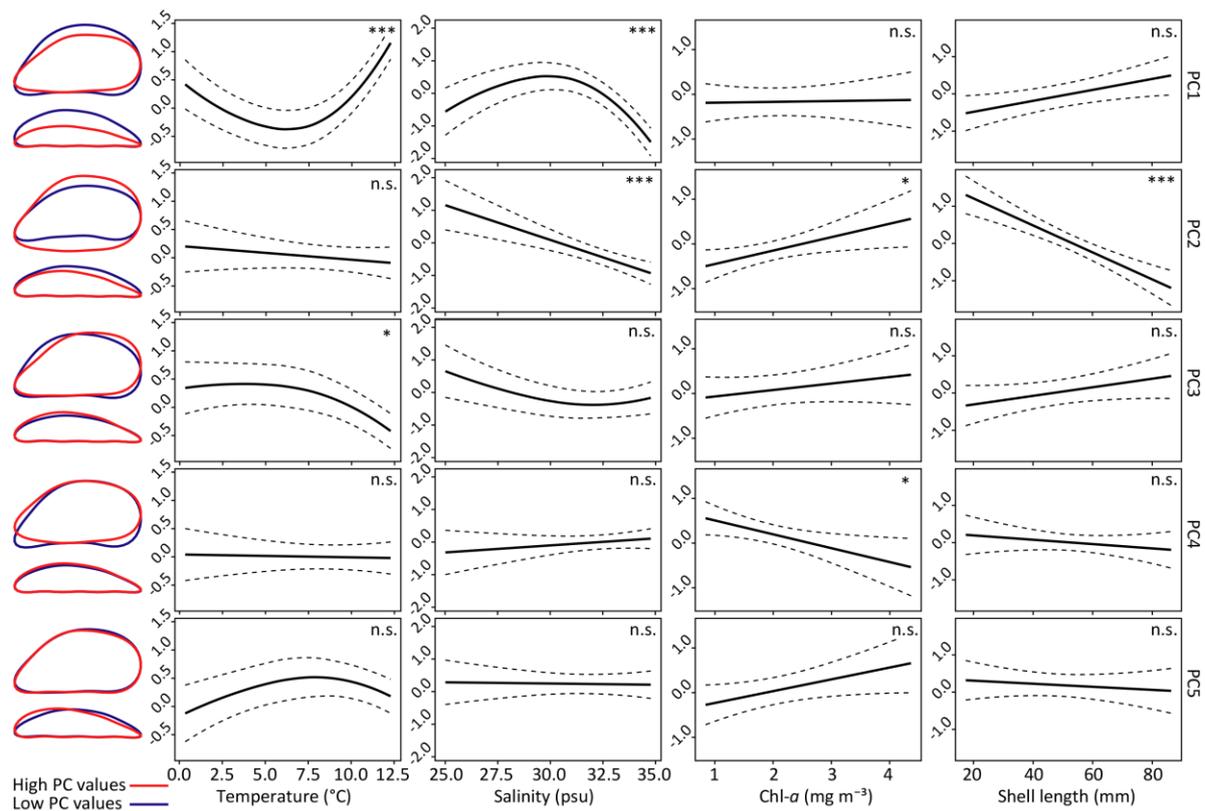
Figure 4.8 *Mytilus* shell shape patterns (System 1)
Modelled shape trends of individual shell features (PC1 - PC5) with environmental descriptors (surface waters temperature, salinity and Chl-*a* concentration) and shell length (size) for the large-scale study system. Predicted values (continuous lines), 95% CIs (dashed lines) and significance level of each fitted smoother and linear predictor are shown. Mussel shape variations described by each shape variable are represented through comparison of mean outlines reconstructed for low and high PC values (blue: Mean − 3 SD; red: Mean + 3 SD). (Significance, n.s. $p > 0.01$, $*$ $p < 0.01$, $**$ $p < 0.001$, $***$ $p < 0.0001$)

scales of analysis. I overcame these limitations through the study of blue mussel populations collected from systems on various geographical scales and with known genetic status, overall providing different degrees of control on regional confounding factors.

Specifically, in the aquaculture system (small-scale, System 3) the cultivation technique considerably reduces accessibility of predators [158] and densities are often actively controlled [335]. In mussel farms on the Scottish west coast, multiple *Mytilus* species and hybrids generally occur in relatively low frequencies and are geographically restricted [158]. Therefore, I used cultured mussels to identify fine-scale shape responses to different environmental exposures (depending on cultivation depth) in a habitat offering ideal conditions for rapid growth. In the Baltic Sea (medium-scale, System 2), mussels constitute 80-90% of the coastal animal biomass [152] and have a strong advantage over competitors for space [336]. This dominance is attributed to an almost complete absence of predators [152, 337]. Here, an increasing *M. edulis* × *M. trossulus* hybridisation with decreasing salinity has been reported [149, 151]. Overall, this region offered low competition and predation across different salinities ranging
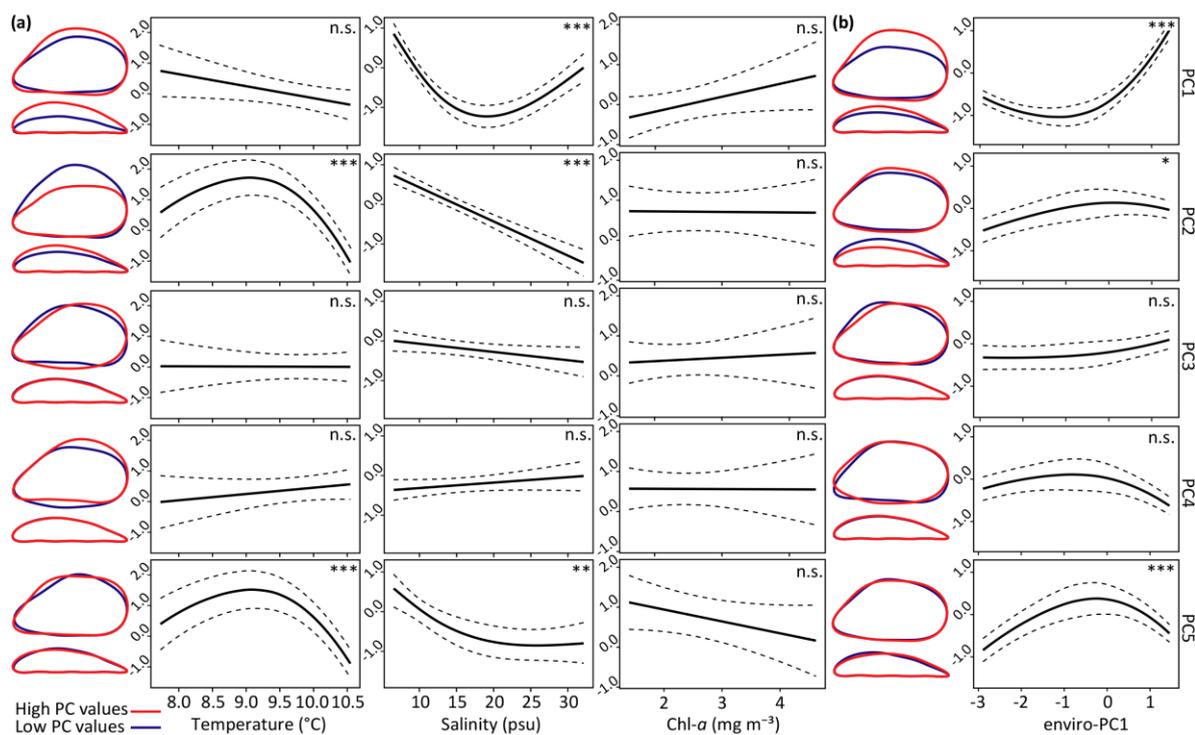
Figure 4.9 Mytilus shell shape patterns (System 2 and 3)
Modelled shape trends of individual shell features (PC1 - PC5) with environmental descriptors (surface waters temperature, salinity and Chl-*a* concentration) and shell length (size) for the (**a**) medium- and (**b**) small-scale study systems. Predicted values (continuous lines), 95% CIs (dashed lines) and significance level of each fitted smoother and linear predictor are shown. Mussel shape variations described by each shape variable are represented through comparison of mean outlines reconstructed for low and high PC values (blue: Mean − 3 SD; red: Mean + 3 SD). (Significance, n.s. $p > 0.01$, * $p < 0.01$, ** $p < 0.001$, *** $p < 0.0001$)

from marine (~32 psu, outer basin) to brackish waters (~6 psu, inner basin). Conversely, Northern Atlantic and Arctic *Mytilus* populations (large-scale, System 1) face variable predation pressures with latitude [338] and competition for space [159, 207]. These confounding factors are generally difficult to quantify directly across a wide geographical scale. However, the possibility to demonstrate broad-scale shape patterns across latitudinal gradients and to compare these variations with more local trends (System 2 and 3) provided complementary information on the factors regulating mussel form and made it possible to draw more general conclusions on shell shape plasticity.

4.4.2 Local shape variation

In System 1 (Figure 4.7a, 4.8), I observed a strong environmental influence on the shape variation captured by PC1, and additional temperature and Chl-*a* effects on PC3 and PC4, respectively. According to documented growth trends of *Mytilus* [159], PC2 indicated a strong developmental (age) effect on shape variance, describing differences between young (round) mussels and old (curved with

Table 4.3 GAMM summary results for smooth and linear terms (Equation 4.1)
Estimated degrees of freedom, F statistics, significance values (upper table), mean effect size of predictors for each response variable (PCs) and 95% CIs (lower table), for each term from the interactions between environmental covariates and PCs are reported.

PCs	*edf*	Estimate	F	*p*-value	95% CI
***f*(Temperature) × PCs**					
PC1	1.92	-1.11	13.89	**<0.0001**	**-1.722; -0.500**
PC2	1.98	0.40	26.67	**<0.0001**	-0.171; 0.975
PC3	1.88	-0.51	6.38	**0.0012**	**-0.908; -0.120**
PC4	1	-1.12	11.38	**0.00076**	**-1.716; -0.527**
PC5	1.87	-0.12	2.95	0.038	-0.691; 0.444
***f*(Salinity) × PCs**					
PC1	2.00	0.30	101.32	**<0.0001**	-0.161; 0.756
PC2	1	0.82	13.13	**0.0003**	**0.382; 1.268**
PC3	1.26	-0.95	13.36	**0.0001**	**-1.378; -0.514**
PC4	1.98	0.83	23.71	**<0.0001**	**0.389; 1.271**
PC5	1.86	-0.28	3.65	0.059	-0.736; 0.173
***f*(Chl-*a*) × PCs**					
PC1	1.94	0.99	11.79	**<0.0001**	**0.351; 1.631**
PC2	1.88	0.53	5.77	**0.0056**	-0.095; 1.152
PC3	1	0.39	1.57	0.21	-0.228; 1.016
PC4	1.90	1.34	9.75	**<0.0001**	**0.702; 1.984**
PC5	1.70	0.22	1.79	0.27	-0.425; 0.855
Length × PCs					
PC1	1	0.55	70.64	**<0.0001**	**0.406; 0.696**
PC2	1	0.26	13.60	**0.0002**	**0.119; 0.406**
PC3	1	-0.33	18.75	**<0.0001**	**-0.474; -0.190**
PC4	1	-0.10	1.74	0.19	-0.244; 0.043
PC5	1	-0.17	3.80	0.051	-0.312; 0.022

wider shells) individuals. Exponential variance structures indicated formation of more heterogeneous average shapes with decreasing temperature (Figure 4.7d).

Mussels in System 2 (Figure 4.7b, Figure 4.9a) showed marked environmental effects on the shape captured by PC1 and PC2, although no correlation between salinity gradients and shell traits was found previously using traditional morphometrics [339]. I observed increasing elongation and shell width with decreasing salinity, indicating a stronger effect of this factor on shape than increasing mussel densities in the inner Baltic as previously thought [172]. The formation of more heterogeneous shapes for higher Chl-*a* concentrations (Figure 4.7e) suggests a strong effect of food availability on mussels growing at low salinities, especially near Baltic coastal lagoons, where food-enriched water inputs are markedly seasonal [340].

In System 3, I detected a strong environmental effect on shell shape (PC1 - PC2; Figure 4.7c, Figure 4.9b). This highlights how altered growth rates [162] as well as decreasing stocking densities with depth are likely to contribute to the shape variations along the cultivation rope.

Overall, I observed similar shell shape patterns at different geographical scales, consisting of the formation of elongated, narrow shells and more parallel dorsoventral margins with decreasing temperature, salinity and food supply. There was also a consistent overlap among PCs for the different study systems, except for PC2 from System 1, describing a strong age effect on mussel shapes due to the wide range of size classes available.

4.4.3 General shape variation

The optimal model for the Atlantic system showed more general environmental effects on shell shape and confirmed some of the detected local relationships (Figure 4.10a). As with PC2 from System 1, PC1 revealed a strong age contribution. I detected a marked environmental influence on PC2, PC3 and PC4, demonstrating a strong effect of salinity on the shape responses in *Mytilus* (Figure 4.10b). The absence of environmental or age effects on PC5 indicated a genetic (species) influence on the captured shape variance. I also detected increased shape heterogeneity in colder waters corroborating the documented variance structure in System 1 (Figure 4.7f).

Exponential variance structures revealed new patterns of shape heterogeneity depending on the spatial scale analysed (Figure 4.7d - f). The local trend observed for System 2 should be considered more of an independent case, showing how the strong salinity effect can be altered locally by increased food supply. On larger geographical scales, however, temperature had a stronger effect on shape heterogeneity. I observed heterogeneous mussel responses in colder waters, creating generally less favourable conditions for mussel growth [162, 166]. Specifically, identified growth alterations might be more evident due to potential competitive advantages of some individuals under environments selecting for specific shapes. On the other hand, in warmer waters, among-individual shape differences may be less conspicuous due to generally more favourable conditions and higher growth potential [162].
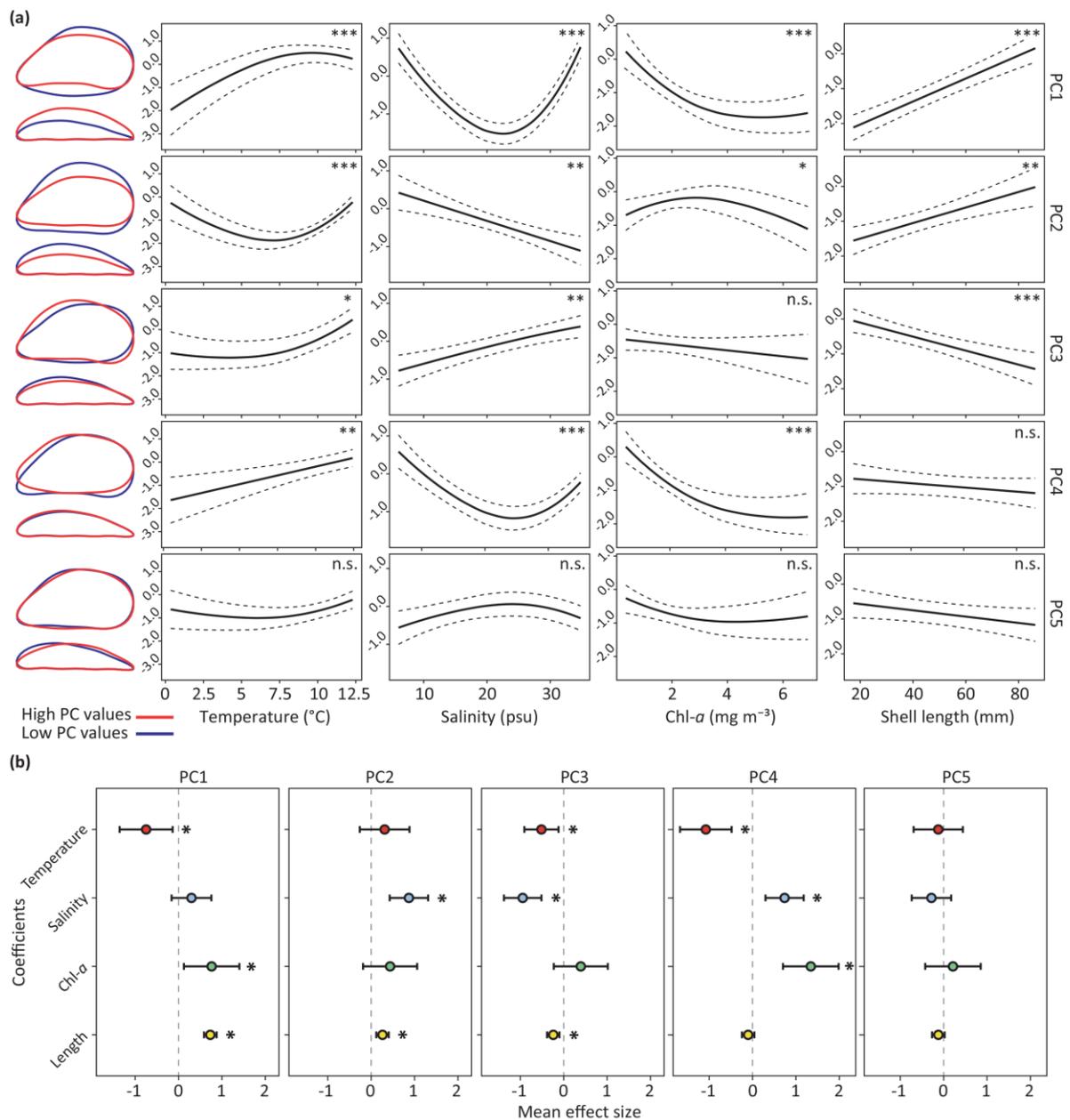
Figure 4.10 *Mytilus* shell shape patterns and effect sizes
(**a**) Modelled shape trends of individual shell features (PC1 - PC5) with environmental descriptors (surface waters temperature, salinity and Chl-*a* concentration) and shell length (size) for the mussel populations from the Atlantic system (Equation 4.1). Predicted values (continuous lines), 95% CIs (dashed lines) and significance level (accuracy of estimated standard errors) of each fitted smoother and linear predictor are shown. Mussel shape variations described by each shape variable are represented through the comparison of mean outlines reconstructed for low and high PC values (blue: Mean -3 SD; red: Mean $+ 3$ SD). (Significance, n.s. $p > 0.01$, * $p < 0.01$, ** $p < 0.001$, *** $p < 0.0001$). (**b**) Mean effect sizes of temperature [Mean (SD) = 8.87°C (5.55)], salinity [Mean (SD) = 26.15 psu (10.54)], Chl-*a* [Mean (SD) = 2.13 mg m^{-3} (1.50)], shell length [Mean (SD) = 50.35 mm (17.03)] for individual shape variables (PCs) and their significance. Error bars represent 95% CIs. Significance is determined when the confidence interval does not cross zero (* $p < 0.01$).

4.4.4 Trends in shell shape

An analysis of the Atlantic system and the comparative use of smaller-scale study systems allowed the identification of general patterns of shape variation as well as independent local trends. Few differences were detected between the individual levels of analysis, while the explained shape features and associations were generally consistent.

The definition of independent variables (PCs) allowed the uncoupling of the individual components of shape variance affected by environmental, developmental (age) and genetic (species) factors, and the identification of the shell features characterised by the strongest shape alterations (Figure 4.11). Specifically, PC1 captured a significant proportion of shape variance related to age modifications of shell outlines during growth (Figure 4.11a). PC2 expressed the largest component of the environmental contribution to shape, describing shell variations under a marked salinity effect (Figure 4.11b). Additional environmental contributions were described by PC3 and PC4 affected by temperature, salinity and strongly by food availability (Figure 4.11c, d). PC5 (+ PC7, PC9, PC10) described the shape variance controlled by species identity and, therefore, the shell features discriminating between *M. edulis*- and *M. trossulus*-like specimens (Figure 4.11e).

Overall, environmental variation influenced a larger proportion of shape variance (PC2 + PC3 + PC4: 49%) and exerted a stronger effect than age (PC1: 38%) and species identity (PC5 + PC7 + PC9 + PC10: 7%) on the shell shape variation in Atlantic *Mytilus*. I detected similar *Mytilus* shape responses to less favourable conditions at the different scales of analysis, indicating the formation of elongated and narrow shells, with more parallel dorsoventral margins (Figure 4.7). These variations could be explained by shapes being driven by the maintenance principle of a physiologically favourable surface-area to volume ratio [162], which increases in elongated shells. The observed shapes, along with physiological acclimatisation [53], could represent an important component of mussel adaptation to environmental stressors.

GAMMs demonstrated water salinity to have a stronger influence (effect size) than other predictors on mussel shape variation than previously reported [164, 205, 339]. Results suggest this physical parameter can lead to dramatic shape changes under sub-optimal conditions to cope with increased metabolic costs resulting from osmotic stress in low saline waters [129]. My models also identified previously unrecognised *Mytilus* shape patterns revealing the formation of less heterogeneous outlines with increasing water temperature as well as more local effects of food supply on the variability of shape responses. My method allows the identification of specific environmental effects on shape variation and the use of contemporary mussels to forecast the responses of benthic communities to near-future climate changes at different regional scales. These findings could also be applied to the study of population responses with temporal clines and to promote the use of shell shape from fossils bivalves to understand past climates and environments.
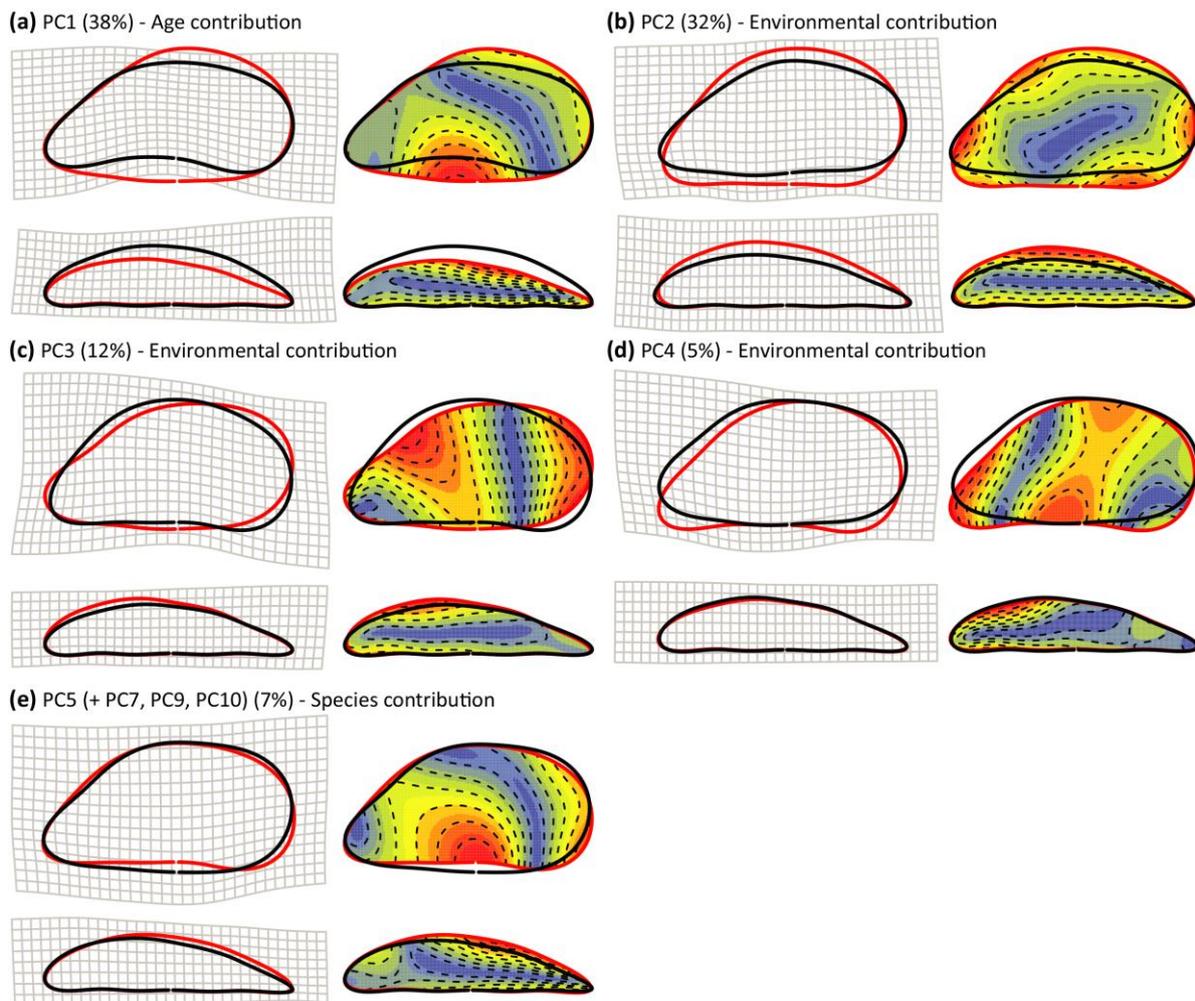
Figure 4.11 Effects of shape variance components on the outline reconstruction
Contributions to mean shapes for the individual components of shape variance regulated by age, environmental and species factors. The influence of each PC on the average *Mytilus* shapes, for both lateral and ventral views, were visualised with: deformation grids (left), depicting the bindings required to pass from the average shape for low (black) to high (red) PC values, and iso-deformation lines (right), representing the outline regions subject to different degrees of change (continuous scale from blue, low deformation, to red, strong deformation). (**a**) PC1 (38%, age contribution) showed a progressive elongation of the shell, with the formation of convex ventral margins, giving a curved aspect to the outline, and an increase of shell width with increasing shell length [159]. (**b**) PC2 (32%, environmental contribution) illustrated the formation of elongated and narrow shells with decreasing salinity. (**c**) PC3 (12%, environmental contribution) explained a progressive rounding of dorsal margins and concaving of ventral profiles for increasing temperature and salinity. (**d**) PC4 (5%, environmental contribution) reported an increase of ventral concaveness with increasing food availability. (**e**) PC5 (+ PC7, 9, 10) (7%, species contribution) indicated the development of concave ventral sides and round dorsal margins in *M. edulis*-like individuals (red) and elongated shells with parallel dorsoventral margins in *M. trossulus*-like specimens (black).

4.5 Conclusions

The combined use of EFA and GAMMs, and the employment of multiple study systems on a wide geographical scale made it possible to describe general relationships between shell shape variation in Atlantic *Mytilus* species and environmental drivers that are independent of developmental (age) and genetic (species) contributions to mussel shape. New methods allowed the identification of previously unrecognised patterns of mussel form and variations in specific shell features at a much finer scale than possible previously.

1. EFA of outlines on blue mussel populations covering a wide latitudinal range (28°, about 3,980 km) allowed an in-depth quantification of shell shape and the definition of new independent variables expressing shape variations at different regional scales.

2. GAMMs and multiple levels of analysis (from small to large geographical scale) described general patterns as well as more local trends of natural shell shape variation and heterogeneity in blue mussels.

3. Powerful statistical methods allowed the identification of shell features under control of environmental, age and species (*M. edulis* and *M. trossulus*) factors. The ability to uncouple these individual components from the modelled shape variance made it possible to describe independent relationships between blue mussel shape variation and environment.

4. Models demonstrated that salinity has the strongest effect on the spatial patterns of shell shape variation, while temperature and food supply are the main predictors of shape heterogeneity, predicting potentially dramatic shape modifications in blue mussels under future environmental conditions.

5. Blue mussels showed similar shell shape responses to less favourable environmental conditions at different geographical scales, with the formation of elongated, narrow shells and more parallel dorsoventral margins, suggesting shell shape variability represents an important adaptive component to environmental stressors.

Although relationships between mussel shape and environmental factors were identified, more studies are needed to understand the adaptive significance of the observed alterations and their underlying causes. Therefore, by providing appropriate study systems and accurate ways to quantify animal shape and diversity, morphological variation can represent a powerful indicator for understanding the adaptation of organisms and help to predict their responses in a rapidly changing environment.

Chapter Five

5 Biomineralisation plasticity and environmental heterogeneity predict geographic resistance patterns of blue mussels to future change

External collaborators contribution

Dr Jakob Thyrring and Dr Mikael K. Sejr (Aarhus University, Aarhus, Denmark) provided mussel specimens from Greenland. Trystan Sanders (GEOMAR Helmholtz Centre for Ocean Research, Kiel, Germany) provided mussel specimens from the Baltic Sea.

5.1 Introduction

Unprecedented global environmental changes are driving scientists to conduct research that will provide increased understanding of the mechanisms underlying geographic variation in species' responses to future environmental conditions [26, 341]. However, our ability to forecast emergent ecological consequences of climate change on marine populations, communities and ecosystems remains limited [38]. Ecosystem-wide projections are severely constrained by heterogeneous patterns of ocean warming and acidification [6], multiple interacting stressors [39], and species-specific effects [21], as well as predictive models which often exclude important biological mechanisms when projecting changes to species and ecosystems in response to climate change [341]. A better mechanistic understanding of the biological processes and environmental sources mediating species' responses to disturbances is critical for building the theoretical baseline necessary to forecast the combined effects of multiple emerging stressors [38, 341].

Advances in macroecology suggest that permanent environmental mosaics, defined by spatial overlaps of non-monotonic environmental gradients [43], as well as regional adaption or acclimatisation [18, 22, 46], dictate geographic variations in species performance and sensitivity to environmental change in

marine ecosystems. Key to these works is that responses vary among populations and individual taxa [18, 21, 108], which often play disproportionately strong roles in structuring benthic communities [24]. Thus, species-specific biological mechanisms driving organismal variability may shape differential regional responses of foundation species to co-occurring multiple drivers. This can establish spatial patterns of unexpected susceptibility of marine communities to future conditions.

Climate change is considered a major threat to marine ecosystems worldwide, with ocean warming and acidification profoundly affecting species life history and ecology [21, 22], as well as community structure and ecosystem dynamics [23, 24]. Species producing calcium carbonate ($CaCO_3$) shells and skeletons are possibly experiencing the strongest impacts of rapid environmental changes [21]. Knowledge of their sensitivity is derived largely from experimentally induced responses in model organisms [21, 26], while complex variations under multiple stressors have rarely been investigated in natural environments [23, 24, 43, 44]. Therefore, inferences made from experimental studies can be misleading and not fully applicable to marine ecosystems [46]. Indeed, species-specific mechanistic responses to habitat alterations [342] on top of mixed outcomes of environmental interactions (additive, synergistic or antagonistic) make future ecosystem predictions extremely challenging. This leaves open the question: do differences in biological mechanisms, shaping regional differences in responses of calcifiers to interacting environmental stressors, define geographic patterns of unforeseen species sensitivity or resistance to global environmental change?

A body of research has focused on responses of marine calcifiers to altered water chemistry [21, 26], but studies have rarely considered changes in biogeochemical cycles that strongly mediate biological responses to environmental alterations [6]. Among those, a marked intensification of the global water cycle in response to warming (+ 4% for + 0.5°C) has been documented over recent decades through changes in ocean salinity [34]. Salinity is a major ecological factor dictating survival of aquatic organisms and ecosystem functioning. Multidecadal studies have revealed a global salinity pattern following the "rich-get-richer" mechanism, where salty ocean regions (compared to the global mean) are getting saltier (mid-latitudes), whereas low salinity regions are getting fresher (tropical convergence zones and polar regions) [34]. In a future 2 - 3°C warmer world [5], a substantial 16 - 24% intensification of the global water cycle is predicted to occur making latitudinal gradients of salinity much sharper [34]. However, emergent ecological effects of changing salinity on calcifying species and marine communities are largely unknown.

Atlantic mussels, *Mytilus edulis* and *M. trossulus*, are important bed-forming foundation species throughout the eulittoral ecosystems of the northern hemisphere (up to 90% of epibenthic biomass), and represent valuable resources for aquaculture (192,000 tonnes produced in 2015 worth 325 million USD) [8]. Growing awareness of the consequences of climate change on biodiversity and industry that *Mytilus*

species support has stimulated a number of studies to estimate the response potential of these habitat-forming calcifiers to changing ocean conditions [62, 83, 99, 108, 110].

Calcareous shells perform a range of vital functions in *Mytilus*. Because shell integrity determines survival, shell traits are subject to strong selection pressure with functional success or failure a fundamental evolutionary driver. *Mytilus* shell consists of three layers (see Section 1.2.10, Figure 1.7): (1) the outer organic periostracum, (2) the calcareous prismatic layer, and (3) the calcareous nacreous layer. These calcareous layers are characterised by different microstructures and more (e.g. aragonite) or less (e.g. calcite and organics) soluble components the combination of which determines chemical and mechanical shell properties [83, 99]. Differences in energetic costs of making shell components [44] combined with future shifts in environmental gradients [6] may influence variations in shell production, composition and structure, shaping regional patterns of shell strength and resistance to acidification.

Mytilus growth, biomineralisation and fitness are linked to multiple drivers, including water temperature, salinity and food supply [chlorophyll-*a* (Chl-*a*) concentration] [51, 53] (see Section 1.2.7). In the North Atlantic and Arctic Oceans, these key environmental factors vary heterogeneously with latitude (Figure 5.1a, b), encompassing a range of conditions predicted under different future climate change scenarios [5]. Here I hypothesise that biological mechanisms driving spatial variations in shell production, mineral and organic composition: **i**) shape regional differences in the responses of *Mytilus* species to interacting environmental drivers, and **ii**) define geographic patterns of unanticipated mussel vulnerability in the face of global environmental changes.

Despite projected environmental alterations [6, 34], salinity gradients have been overlooked in large-scale models predicting emergent effects of climate changes on marine organisms. This knowledge is essential to predict whether environmental changes affect shell variability (i.e. thickness, mineral and organic content) and its properties, especially in calcifying foundation species such as *M. edulis* and *M. trossulus*. These factors are crucial for understanding species regional susceptibility patterns to other rapidly emergent stressors, such as warming and acidification [38].

In this study, I examine the relationships between the plasticity in *Mytilus* shell production and composition (from juveniles to large adults) and interactive environmental gradients of surface water temperature, salinity and Chl-*a* concentration in 17 populations spanning a latitudinal range of 30° (3,334 km) across the Atlantic-European and Arctic coastline (Figure 5.1a, b). In particular, I test for a latitudinal effect on *Mytilus* shell calcification (variation in shell thickness and organic content) that I hypothesise will show a general decrease from temperate to polar regions. I also identified environmental sources and magnitude of regional variations in shell deposition, to test whether salinity affects shell production and mineral composition during growth, driving changes of mechanical and chemical shell properties. Finally, I modelled spatial trends in the production of individual shell layers

with environmental gradients, to test whether biological mechanisms, driving variations in shell structure and properties, shape regional responses of *Mytilus* to interacting stressors (especially salinity) and define geographic patterns of sensitivity to future changes.

5.2 Materials and methods

5.2.1 *Mytilus* collection

I sampled individuals from a total of 17 *Mytilus* (*Mytilus edulis* and *M. trossulus*) populations along the North Atlantic, Arctic and Baltic Sea coastlines from four distinctive climatic regions (warm-temperate, cold-temperate, subpolar, and polar) covering a latitudinal range of 30° (a distance of 3,334 km), from Western European (Brest, North-West France, 48°N) to Northern Greenlandic (Qaanaaq, North-West Greenland 78°N) coastlines (Figure 5.1a). During December 2014 - September 2015, mussels of various size classes for each site (shell length of 26 - 81 mm) were sampled from the eulittoral (low intertidal) zone on rocky shores for a total of 424 individuals (Supporting Table A.5). For each specimen, shell length was measured with digital callipers (0.01 mm precision) and used as a within-population proxy for age. I analysed *Mytilus* populations of which the genetic structure was known, with particular focus on species identity and hybrid status (*M. edulis* × *M. trossulus*). *Mytilus* shells used were either from specimens already evaluated in genetic investigations or mussels obtained from sites routinely used in regional monitoring programs that provided information on species identity (Supporting Table A.5). Areas where the Mediterranean mussel, *Mytilus galloprovincialis*, was present were avoided. I did, however, sample few (3) sites with very low levels of *M. edulis* × *M. galloprovincialis* hybridisation.

5.2.2 Mussel shell preparation

Left shell valves (*n* = 424) were set in polyester resin (Kleer-Set FF, MetPrep, Coventry, U.K.) blocks and prepared following the procedure in Section 2.2. Embedded specimens were sliced longitudinally along their axis of maximum growth (Figure 2.2) using a diamond saw and then progressively polished with silicon carbide paper (grit size: P800 - P2500) and diamond paste (grading: 9 - 1 µm). Photographs of polished sections (Figure 2.2) were acquired with a stereo-microscope (Leica M165 C equipped with a Leica DFC295 HD camera, Leica, Wetzlar, Germany) and shell thickness was measured using the Fiji software (v1.51u). Since larger individuals had undergone evident environmental abrasion or dissolution which removed the periostracum and prismatic layer closer to the umbo, I estimated the thickness of the whole-shell, prismatic and nacreous layers at the midpoint along the shell cross-section. The proportion of calcite was estimated as:
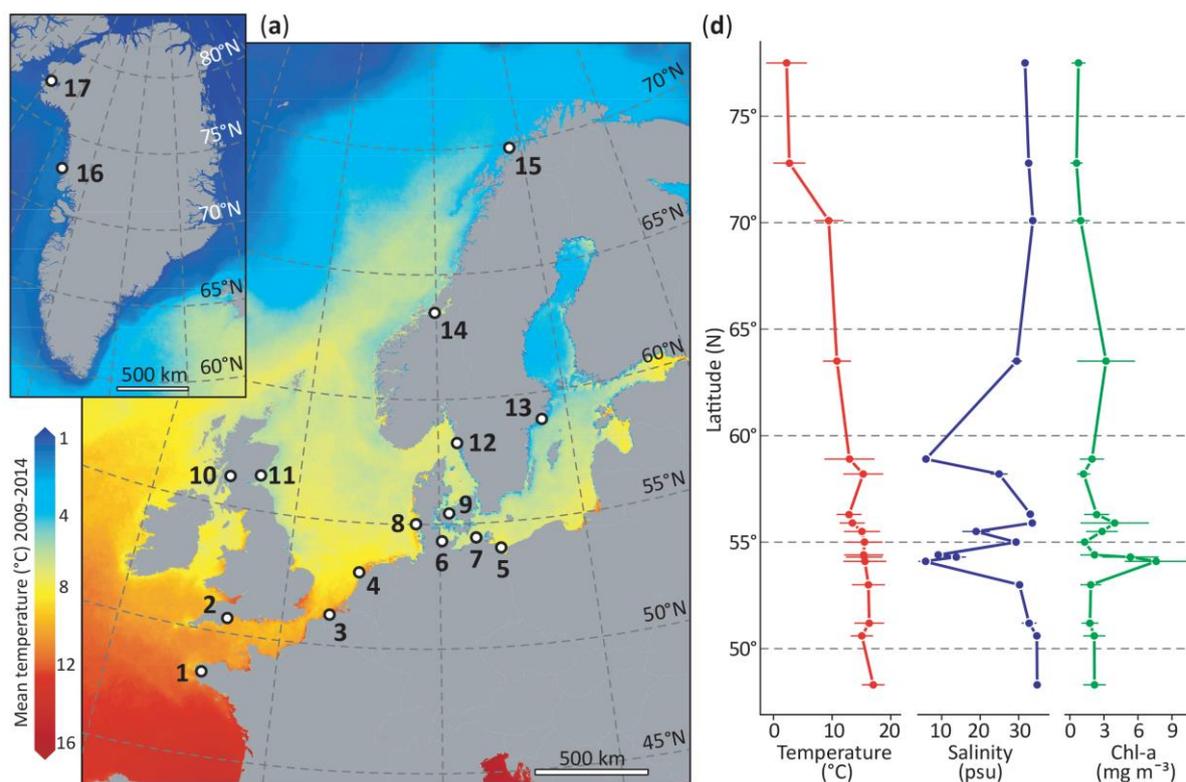
Figure 5.1 *Mytilus* spp. collection sites and environmental heterogeneity
(**a**) Thermal map of North-East Atlantic and Arctic surface waters from the CMEMS (http://marine.copernicus.eu/) biogeochemical datasets showing locations (open circles) where *Mytilus* was collected from across the Eastern European and Greenlandic coastlines (from 48°N to 78°N): (1) Brest, France, (2) Exmouth, England, (3) Oostende, Belgium, (4) Texel, Netherlands, (5) Usedom, (6) Kiel, (7) Ahrenshoop, (8) Sylt, all Germany, (9) Kerteminde, Denmark, (10) Tarbet, Kintyre, Scotland, (11) St. Andrews, Scotland, (12) Kristineberg, Sweden, (13) Nynäshamn, Sweden (14) Trondhiem, Norway, (15) Tromsø, Norway, (16) Upernavik, Greenland and (17) Qaanaaq, Greenland. Map created with ArcMap 10.5 (ArcGIS software by Esri, http://esri.com), background image courtesy of OpenStreetMap (http://www.openstreetmap.org). (**b**) Latitudinal gradients for sea surface temperature, salinity and chlorophyll-*a* (Chl-*a*) concentration, showing environmental heterogeneity across the study regions. Mean values (May - October, filled circles) and SD (horizontal lines) for the 6-year period 2009 - 2014 were estimated from CMEMS datasets.

$$alcite\% = \left(\frac{prismatic\ thickness}{whole\text{-}shell\ thickness} \right) \times 100$$

Equation 5.1

Periostracum thickness was measured at the posterior edge where it attaches to the external side of the prismatic layer, to estimate the fully formed organic layer that was unaffected by decay or abrasion [185].

5.2.3 Organic content analyses

I performed thermogravimetric analyses (TGA) to estimate the weight proportion (wt%) of organic matrix within the prismatic layer. *Mytilus edulis* specimens were selected from four populations (sites 1, 11, 15, 16, Figure 5.1a) to explore differences in shell organic content under temperate and polar regimes. I removed the periostracum by sanding, and prismatic layer tiles (8×5 mm, $n = 20 \times 4$ sites) were isolated along the posteroventral shell margin. Tiles were cleaned, air-dried and then finely ground. Ten milligrams of this powdered shell were tested with a thermogravimetric analyser (TGA Q500, TA Instruments, New Castle, DE, U.S.A.). Samples were subjected to constant heating from ~25°C to 700°C at a linear rate of 10°C min^{-1} under a dynamic nitrogen atmosphere and weight changes were recorded (see protocol in Section 2.3.5). I estimated the wt% of organic matter within the shell microstructure as the proportion of weight loss during the thermal treatment between 150°C and 550°C (Figure 2.5).

5.2.4 Environmental characterisation

I selected three key environmental drivers based on their known influence on mussel growth, their level of collinearity across the geographic scale investigated and the forecasted major ocean alterations under climate change [5]. For each site, measurements of sea surface temperature, salinity and Chl-*a* concentration, the latter being used as a proxy for food supply [53], were generated using the Copernicus Marine Environment Monitoring Service (CMEMS) (http://marine.copernicus.eu/). These climate datasets are composed of high-resolution physical and biogeochemical assimilated (integration of observational and predicted information) daily data ($n = 2,191$ per parameter) (Appendix B, Environmental Datasets). To provide a first order approximation of the water conditions prevailing during the near-maximum rates of shell deposition [343], I expressed parameters as mean May - October values averaged over the 6-year period 2009 - 2014 and used these as input variables (Figure 5.1b, Table 5.1).

Direct environmental monitoring for each site was not feasible due to the number of populations analysed, their geographic range (> 3,300 km) and the temporal resolution (6 years) required to estimate the average growth conditions during the lifespan of sampled specimens. For this large-scale study, remote-sensing and assimilated data presented potential advantages compared to traditional measurements due to their high spatio-temporal resolution, advanced calibration and validation (i.e. high correlation with discrete field measurements) [329, 330].

Table 5.1 Environmental covariates

Summary statistics (mean value and SD) of environmental conditions at each study site. Site codes as in Figure 5.1c.

Site	Temperature (SD) (°C)	Salinity (SD) (psu)	Chl-*a* (SD) (mg m^{-3})
1	17.01 (1.94)	34.79 (0.33)	2.16 (1.00)
2	15.08 (1.92)	34.70 (0.39)	2.14 (0.98)
3	16.33 (2.55)	32.70 (1.92)	1.76 (0.75)
4	16.19 (2.82)	30.20 (2.28)	1.84 (0.93)
5	15.59 (3.67)	5.39 (1.45)	7.62 (3.02)
6	15.46 (3.11)	13.77 (2.54)	5.34 (7.10)
7	15.39 (3.34)	9.09 (1.24)	2.16 (1.25)
8	15.51 (3.10)	29.34 (1.14)	1.29 (0.70)
9	15.08 (3.09)	18.92 (3.57)	2.83 (1.40)
10	13.43 (2.15)	33.51 (0.49)	3.95 (3.72)
11	12.89 (2.14)	33.13 (0.57)	2.36 (1.12)
12	15.30 (3.39)	24.90 (2.20)	1.21 (0.59)
13	12.95 (4.28)	5.92 (0.45)	1.94 (1.07)
14	10.82 (2.39)	29.47 (1.35)	3.18 (2.55)
15	9.42 (2.49)	33.67 (0.37)	0.94 (0.81)
16	2.70 (2.74)	32.60 (0.53)	0.58 (0.54)
17	2.23 (3.46)	31.64 (0.97)	0.75 (0.61)

5.2.5 Statistical analysis

Generalised linear mixed models (GLMMs) were applied to account for the hierarchical structure of the dataset consisting of multiple specimens (*n* = 24 - 26 replicates) from each collection site and to generalise my results to *Mytilus* populations beyond the study sample.

I carried out data exploration following the protocol in Section 3.1.1. Initial inspection revealed no outliers. Pairwise scatterplots and variance inflation factors (VIFs) were calculated to check for collinearity between input variables. VIF values < 2 indicated an acceptable degree of correlation among covariates to be included within the same model. I applied residual regression to uncouple the unique from the shared contribution of temperature and Chl-*a* concentration to the response [297]. This allowed me to account for the existing causal link between these two parameters and to avoid inferential problems from modelling non-independent covariates without losing explanatory power [297]. To directly compare model estimates from predictors on different measurement scales, estimate biologically meaningful intercepts, and interpret main effects when interactions are present, I standardised all the input variables (environmental parameters and shell length). For standardisation, I subtracted the sample mean from the variable values and divided them by the sample standard deviation [$z_i = (x_i - \bar{x}) / \sigma_x$]. Preliminary inspection of residual patterns showed heteroscedasticity in most

models. The use of different continuous probability distributions (i.e. gamma) and link functions did not stabilise the variance, therefore a ln-transformation of the response was required, except for calcite% and wt% measurements. Response variables did not require further transformations.

Separated GLMMs were used to explore patterns of shell thickness and composition, from juvenile to large adults, with respect to latitude and shell length (size), and to compare between individual shell layers. The proportion (wt%) of organic matrix ($n = 80$) was modelled with a generalised linear model (GLM) as a function of collection site (categorical, four levels: site 1, 11, 15 and 16) and prismatic thickness (continuous) to test for differences between polar and temperate regions and association with shell thickness. The response variable was coded as a value from 0 to 1. Therefore, to ensure that the fitted values range from 0 to 1, I could not use a Gaussian linear model and instead applied a model with a beta distribution, which can be used if the response variable is a continuous variable ranging from x_1 and x_2, with a logistic link function. Pair-wise contrasts with a Bonferroni correction were then used to test for differences in wt% among sites within and between climatic regions.

Different approaches were used to investigate the relationships between shell thickness and environmental gradients. Individual GLMMs were fitted to explain spatial variations in whole-shell thickness, periostracum thickness, and calcite% with environmental conditions during shell growth ($n = 424$ each). Prismatic and nacreous layer thickness were analysed within the same GLMM, allowing the simultaneous prediction of common and divergent environmental effects on both layers and to reduce the probability of type I error. To model shell thickness ($n = 424 \times 2$ layers) as a function of the environmental predictors, I used a GLMM with a normal distribution (Equation 5.2). Fixed continuous covariates of the optimal models were standardised *temperature*, *salinity*, and *Chl-a* in addition to shell *layer* (categorical, two levels: prismatic and nacreous) and their two-way interactions. Shell *length* (continuous) was included to control for possible effects of within-population age on layer thickness. To incorporate the dependency among observations for a specific layer from the same collection site, I used *site* as a random intercept. The optimal model was of the form:

$$\ln(Thickness_{ijk}) \sim N(\mu_{ijk}; \sigma_j^2)$$
$$\mu_{ijk} = Temperature_{ijk} + Salinity_{ijk} + Chl\text{-}a_{ijk} + Length_{ik} + Layer_j$$
$$+ \, Temperature_{ijk} \times Layer_j + Length_{ik} \times Layer_j + Site_{ij}$$
$$Site_{ij} \sim N(0; \sigma_{Site}^2)$$

Equation 5.2

where *Thickness_{ijk}* is the *k*th thickness observation from layer *j* (*j* = prismatic, nacreous) and site *i* (*i* = 1, …, 17). *Site_{ij}* is the random intercept for layer *j*, which is assumed to be normally distributed with expectation 0 and variance σ_{Site}^2.

Models were optimised by first selecting the random structure and then the optimal fixed component following the protocol in Section 3.7. The principal tools for model comparison were the corrected Akaike Information Criterion (AICc) [321] and bootstrapped likelihood ratio tests [322]. Random terms were selected on prior knowledge of the dependency structure of the dataset. Visual inspection of residual patterns indicated violation of homogeneity in most cases. This required the use of variance structures (generalised least squares) allowing the residual spread to vary with respect to shell layer. The fixed component was optimised by rejecting only non-significant interaction terms that minimised the AICc value. For all model comparisons, variation of AICc between the optimal (lowest AICc value) and competing models were greater than 8, and fixed-effect estimates were nearly identical, indicating that competing models were very unlikely to be superior [321]. The proportion of variance explained by the models was quantified with conditional or pseudo determination coefficients (cR^2 or $pseudoR^2$). I used variograms to assess the absence of spatial autocorrelation. Final models were validated by inspection of standardised residual patterns to verify GLMM assumptions of normality, homogeneity and independence as in Section 3.7.2. I used optimal models to estimate the mean effect sizes (same measurement scale) of environmental drivers on the response. Ninety-five per cent confidence intervals (95% CI) for the regression parameters were generated using bias-corrected parametric bootstrap methods (10,000 iterations). 95% CIs were used for statistical inference due to estimation of approximated significance values (*p*-value) in mixed-modelling [287, 310, 316]. If the confidence intervals did not overlap zero, then the effect was considered significant.

5.3 Results

5.3.1 Latitudinal patterns of shell deposition

GLMMs indicated a general decrease of *Mytilus* whole-shell thickness with increasing latitude from warm-temperate to polar regions (Figure 5.2a). I detected a significant negative relationship between the prismatic layer thickness and latitude (Figure 5.2a), while there was no variation in nacreous thickness, periostracum thickness, and relative proportion of prismatic layer thickness (calcite%) (Table 5.2). Shell length was positively correlated with thickness in all layers indicating thickening during growth (Table 5.2).

Prismatic layers were characterised by a significantly higher organic content (lower proportion of $CaCO_3$) in mussel shells from polar than temperate regions, indicating decreased shell calcification at higher latitudes (Figure 5.2b). Polar shells [sites 15, 16; mean (SD) = 1.8 wt% (0.31)] were characterised by an average of 29% more organic content compared to temperate mussels [sites 1, 11; mean (SD) =
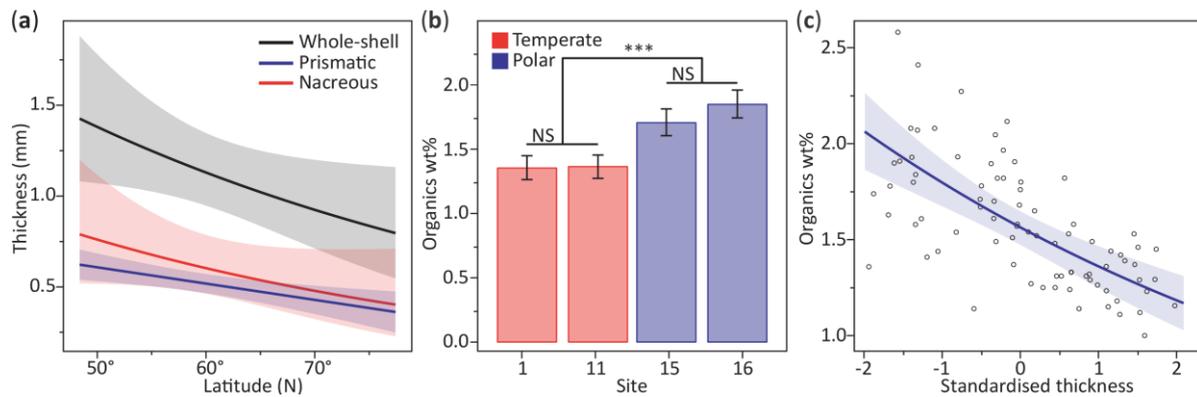
Figure 5.2 Latitudinal patterns of shell thickness, organic content and calcification
(**a**) Relationships between the thickness of whole-shell (black), prismatic (blue) and nacreous (red) layers and latitude. Whole-shell thickness decreased poleward (95% CI = -0.36 to -0.01, $cR^2 = 0.81$). The prismatic layer was significantly related to latitude (95% CI = 4.70 to 5.73, $cR^2 = 0.72$). Predicted values (continuous lines) and confidence intervals (shaded areas) were estimated for mussels of mean shell length (47.42 mm). Parameters' significance is determined when the bootstrapped 95% CI does not include zero. (**b**) Variations in organic content among shells from temperate (sites 1, 11, white bars) and polar (sites 15, 16, grey bars) climates. Pair-wise contrasts indicated significantly higher proportions of organics in high-latitude than low-latitude specimens [mean difference = 0.44%; $z = 8.27$, $p < 0.0001$ (***), pseudo$R^2 = 0.49$], in addition to non-significant differences (NS) among temperate (mean difference = 0.002%; $z = 0.12$, $p = 0.91$) and polar (mean difference = 0.13%, $z = 1.86$, $p = 0.063$) populations. (**c**) Relationship between the proportion of organics and standardised thickness of the prismatic [mean (SD) = 529 µm (174)] (sites 1, 7, 10 and 11), indicating a negative association between layer thickness and calcification level ($z = -7.10$, $p < 0.0001$, pseudo$R^2 = 0.40$).

1.4 wt% (0.16)]. The organics wt% was negatively correlated with prismatic thickness (Figure 5.2c), indicating a lower proportion of $CaCO_3$ and thinner, so less calcified, shells at polar latitudes.

5.3.2 Environmental influence on shell production and composition

I identified significant trends in shell thickness with environmental gradients depending on the shell measurement considered (Figure 5.3a - e, Table 5.3, Supporting Figure A.1). Whole-shell thickness was positively related to temperature, salinity and shell length, but there was no influence of Chl-*a* ($cR^2 = 0.93$; Figure 5.3a). Salinity had an effect on shell thickness that was 3.4 and 2.1 times larger than temperature and length, respectively (Figure 5.3a, Table 5.3). I detected a negative relationship between calcite% and salinity (95% CI = -12.03 to -2.38, $cR^2 = 0.56$) (Figure 5.3b, Table 5.3), with none of the other drivers having a significant effect.

Sea surface temperature, salinity and shell length all successfully predicted ($cR^2 = 0.93$, Equation 5.2) variations in the thickness of prismatic and nacreous layers, while no influence of Chl-*a* was detected

Table 5.2 Latitudinal GLMMs summary
Estimated statistics and bootstrapped 95% CIs for regression parameters are reported for the modelled relationships between individual shell measurements, standardised latitude, and shell length.

Parameter	Estimate	SE	95% CI	*t*-value	*p*-value	Random effects	SD
Whole-shell							
(Intercept)	7.036	0.089	**6.863; 7.210**	78.89	**<0.0001**	Site	0.293
Latitude	-0.188	0.089	**-0.316; -0.011**	-2.10	0.067	Residuals	0.192
Length	0.188	0.036	**0.115; 0.259**	5.18	**<0.0001**		
Prismatic layer							
(Intercept)	6.211	0.267	**6.118; 6.302**	133.06	**<0.0001**	Site	0.867
Latitude	-0.839	0.268	**-1.373; -0.305**	-3.13	**0.013**	Residuals	0.906
Length	-0.786	0.155	**0.478; 1.106**	4.83	**<0.0001**		
Nacreous layer							
(Intercept)	6.410	0.135	**6.148; 6.674**	47.63	**<0.0001**	Site	0.442
Latitude	-0.217	0.135	-0.480; 0.050	-1.61	0.14	Residuals	0.321
Length	0.210	0.060	**0.089; 0.332**	3.39	**0.00088**		
Periostracum							
(Intercept)	3.709	0.053	**3.605; 3.814**	70.02	**<0.0001**	Site	0.170
Latitude	-0.031	0.053	-0.136; 0.073	-0.59	0.58	Residuals	0.224
Length	0.283	0.035	**0.213; 0.353**	7.65	**<0.0001**		
Calcite%							
(Intercept)	45.255	2.476	**40.327; 50.134**	18.28	**<0.0001**	Site	8.019
Latitude	1.097	2.484	-3.718; 6.011	0.44	0.67	Residuals	8.829
Length	-0.510	1.486	-3.444; 2.515	-0.33	0.75		

(Table 5.3). The mean effect size of salinity on the response was twice as large as the effect of shell length, while it was 2.9 and 4.7 times larger than the effect of temperature on the prismatic and nacreous layers, respectively (Equation 5.3, Figure 5.3c, d). This indicates salinity had a stronger contribution to predicting shell structure than the effects of temperature, Chl-*a*, and shell length combined (Figure 5.4).

$$\mu_{ijk} =$$
$$\begin{cases} 5.907 + 0.138 \times Temperature + 0.396 \times Salinity + 0.028 \times Chl\text{-}a + 0.197 \times Length & \text{Prismatic} \\ 5.853 + 0.138 \times Temperature + 0.654 \times Salinity + 0.028 \times Chl\text{-}a + 0.308 \times Length & \text{Nacreous} \end{cases}$$

Equation 5.3

Interactions between shell layer and both salinity and shell length (Equation 5.3) indicate deposition of proportionally thicker prismatic layers under low salinities and proportionally thicker nacreous layers under higher salinities across the entire range of shell lengths (Figure 5.4).
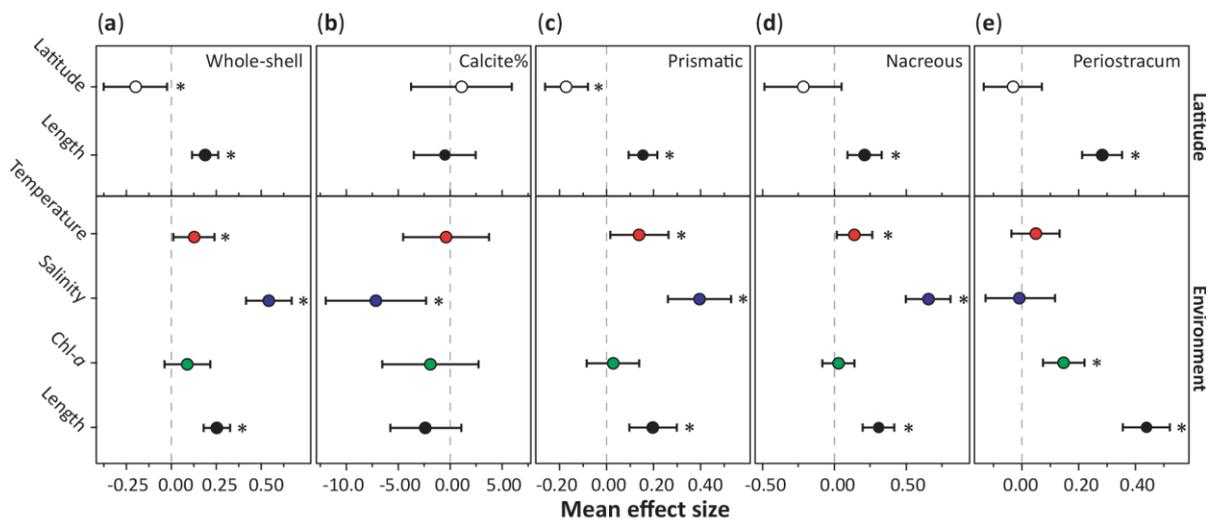
Figure 5.3 Mean effect size of predictors on *Mytilus* shell measurements
Effect sizes were estimated from individual latitudinal (left panels) and environmental (right panels) GLMMs. Mean effect sizes and direction of impacts of latitude (white), shell length (black), sea surface temperature (red), salinity (blue) and Chl-*a* concentration (green) on layer ln-thickness (µm) measurements and calcite% are reported: (**a**) whole-shell, (**b**) calcite%, (**c**) prismatic layer, (**d**) nacreous layer, and (**e**) periostracum. Note the different scales on the y-axis to highlight variations among layers. Significance of regression parameters is determined when the bootstrapped 95% CI (error bars) does not cross zero (* denotes a significant difference from zero).

5.3.3 Periostracum thickness plasticity

Models of periostracum thickness revealed significant exponential relationships with Chl-*a* and shell length ($cR^2 = 0.81$) (Table 5.3). Length had a mean effect that was 3 times larger than Chl-*a* (Figure 5.3e), showing a rapid thickening of the periostracum during shell growth. The interactions between shell length and both salinity and temperature indicate that the effects of these variables on periostracum were interdependent. At low salinities, the higher values of shell length had a greater positive effect on periostracum thickness, while the reverse was true for higher temperatures having a marginal effect only on thickening rates (Figure 5.5a, b). This suggests that increasing shell size was a more important factor for periostracum growth in fresher waters than in relatively saltier conditions.

5.3.4 Among-site shell variation

GLMMs showed no difference in collection site-level effects (conditional modes) on each thickness measurement (Figure 5.6). This indicated no residual effect of species identity or hybridisation on the thickness of individual shell layers at different sites after accounting for the effects of environmental factors and shell length.
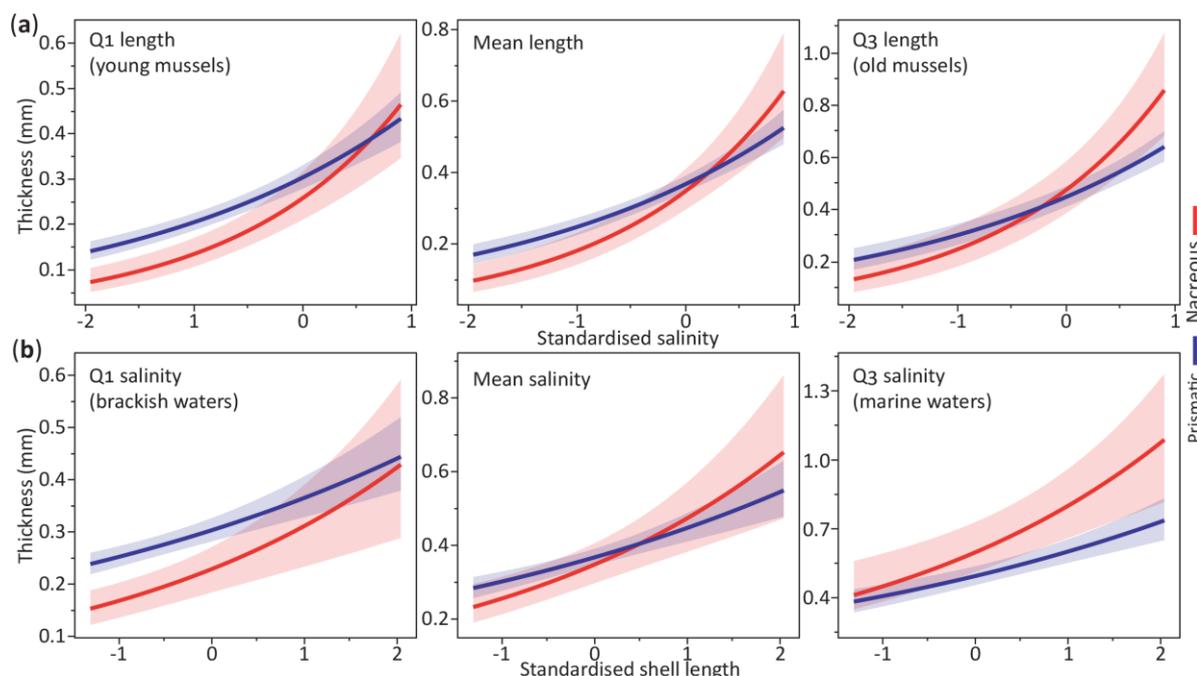
Figure 5.4 Environmental influence on shell production and composition
Predicted relationships between thickness of prismatic (blue) and nacreous (red) layers, and standardised salinity [mean (SD) = 25.52 psu (10.29)], shell length [mean (SD) = 47.42 mm (16.20)] and their interactions. (**a**) Shell thickness is modelled as a function of salinity for the 1st quartile (Q1 = 31.50 mm), mean value (47.42 mm) and 3rd quartile (Q3 = 63.90 mm) of the shell lengths sampled. For medium-sized mussels, I detected a decreasing proportion of the calcitic component with increasing salinity and the deposition of relatively thicker aragonitic layers at salinities > 27.67 psu. (**b**) Thickness is modelled as a function of length for the 1st quartile (Q1 = 18.92 psu), mean value (25.52 psu) and 3rd quartile (Q3 = 33.13 psu) of salinity. At mean salinity, I detected an inversion of the relative layers' thickness for shell length > 55.30 mm. Across the entire range of shell lengths, the model predicts formation of calcite- and nacreous-dominated shells under low- and high-salinities, respectively. Mean values (continuous lines) and confidence intervals (shaded areas) are predicted while controlling for temperature (13.03°C) and Chl-*a* (2.48 mg m^{-3}).

5.4 Discussion

My results demonstrate that plasticity in shell production in *Mytilus* species and the spatial structure of environmental conditions drive geographic variations in shell responses shaping regional differences in the resistance of these foundation species to global environmental change. An understanding of the biological mechanisms driving regional differences in species' responses to multiple interacting stressors is crucial for improving predictive accuracy and informing more realistic projections of species and ecosystem resistance to climate change [341]. Heterogeneous population-level responses from different climates suggest that environmental stressors, especially salinity, predict regional variations in *Mytilus* shell production, mineral (prismatic and nacreous) and organic (periostracum) composition during growth, which is reflected in the relative proportion of each shell layer. Variations in shell
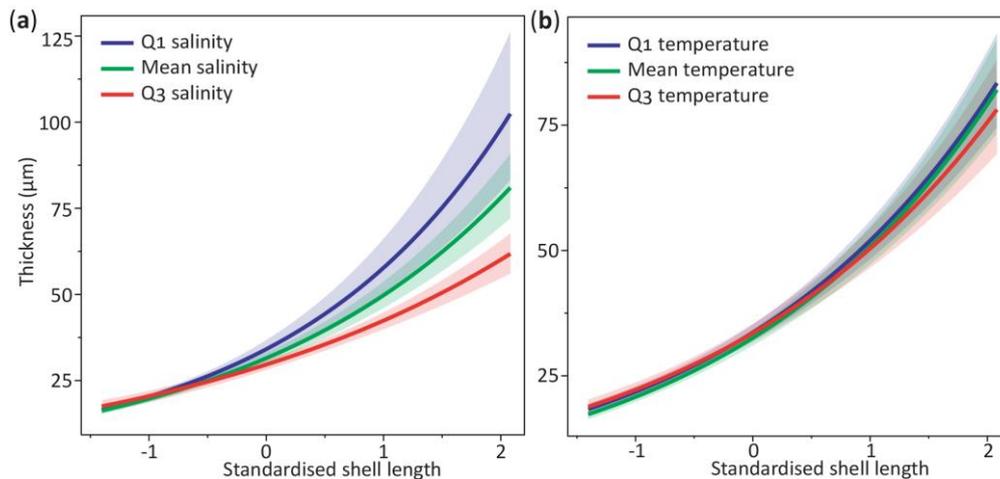
Figure 5.5 Periostracum plasticity

Interacting effects of salinity, temperature and shell length on shell periostracum. (**a**) Periostracum thickness is modelled as a function of shell length [mean (SD) = 47.42 mm (16.20)] for the 1st quartile (Q_1 = 18.92 psu, blue line), mean (25.52 psu, black line) and 3rd quartile (Q_3 = 33.13 psu, red line) of water salinity. Predicted values (continuous lines) and confidence intervals (shaded areas) indicate higher rates of exponential periostracal thickening with decreasing salinity. Smaller individuals (shell length < 48.38 mm) were characterised by non-significant thickness differences under different salinity regimes. (**b**) Thickness is modelled for the 1st quartile (Q_1 = 12.89°C, blue line), mean (13.03°C, black line) and 3rd quartile (Q_3 = 15.51°C, red line) of water temperature. Predicted values indicate a marginal influence of temperature on periostracal thickening.
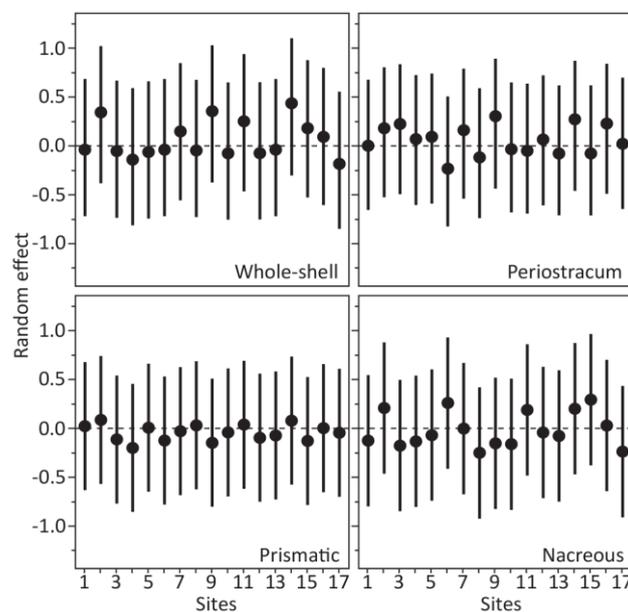
Figure 5.6 Among-site shell variation

GLMMs' conditional modes (filled circles) and variances (continuous lines) of the random effect estimated for individual shell layers. Modes represent the difference between the average predicted response (layer thickness) for a given set of fixed-effects values (mean environmental covariates and shell length) and the response predicted at a particular site. These indicate no detectable residual effect of species (*Mytilus edulis* or *M. trossulus*) and level of hybridisation on shell thickness for each site.

production and composition determine geographic differences in chemical and mechanical protection of shells, shaping the vulnerability of these habitat-forming species to future conditions.

5.4.1 Spatial variations in shell calcification

Decreasing shell calcification (increasing organic content and thinner shells) towards high latitudes (Figure 5.2) supports documented patterns of skeletal production [44, 102]. Two explanatory paradigms exist for decreased skeletal size at higher latitudes: increased calcification costs [44] and reduced predation pressure [338]. Given the higher production cost of organics than $CaCO_3$ deposition [44] and problematic protein production at polar temperatures [344], we might expect a reduced proportion of organic matrix. Moreover, decreasing predation pressure [338] should result in thinner shells of the same composition irrespective of geographic area. However, the wt% of organic matrix was higher at Arctic latitudes. This could suggest either (or a combination of) a marked increase in the cost of calcification in polar regions [44], altering significantly the relative costs of $CaCO_3$ and organics production, or a decreased saturation state (increased dissolution) of $CaCO_3$ due to low temperatures and, more importantly, salinity (low $[Ca^{2+}]$ availability) [51]. In either case, these underlying effects would result in decreased shell calcification at high latitudes. This increased proportion of organic matrix could protect the calcified shell components from dissolution and have an adaptive beneficial effect in more corrosive conditions.

5.4.2 Environmental effect on shell deposition and composition

These results illustrate that different drivers significantly affect both shell thickness and composition in *Mytilus* (Figure 5.3). For over 60 years, temperature and shell size have been considered key predictors of $CaCO_3$ shell mineralogy across latitudes, dictating the formation of predominantly aragonitic structures in temperate regions and increased calcite precipitation in cold climates [343, 345, 346]. Although this study partly supports previous findings, I demonstrate that salinity has the strongest influence on shell production and composition in *Mytilus*, which is contrary to the general assumption of temperature and shell size being the primary predictors of shell plasticity.

The interaction between shell layer, salinity and shell size indicates heterogeneous, age-related compositional changes in *Mytilus* shells across different salinities (Figure 5.4a). Shifts in shell properties from juveniles to large adults are strongly modulated by salinity, which leads to the formation of exclusively prismatic-dominated shells in brackish waters and nacreous-dominated structures under marine conditions (Figure 5.4b). These patterns, which I show were independent of species or hybrid status (Figure 5.6), indicate that mussel shell plasticity during growth (the *Length* × *Layer* interaction,

Table 5.3 Environmental GLMMs summary
Estimated statistics and bootstrapped 95% CIs for regression parameters are reported for the
modelled relationships between individual shell measurements, standardised environmental
covariates and shell length.

Parameter	Estimate	SE	95%CI	*t*-value	*p*-value	Random effects	SD
Whole-shell							
(Intercept)	6.617	0.051	**6.517; 6.717**	128.62	**<0.0001**	Site	0.209
Temperature	0.156	0.054	0.014; 0.240	2.89	**0.013**	Residual	0.188
Salinity	0.525	0.060	**0.411; 0.672**	8.69	**<0.0001**		
Chl-*a*	0.074	0.054	-0.042; 0.216	1.37	0.20		
Length	0.248	0.037	**0.181; 0.327**	6.44	**<0.0001**		
Prismatic and nacreous							
(Intercept)	5.907	0.031	**5.774; 6.036**	188.31	**<0.0001**	Site (Pr)	0.123
Temperature	0.138	0.033	**0.013; 0.260**	4.17	**0.0011**	Site (Na)	0.310
Salinity	0.396	0.039	**0.264; 0.531**	10.22	**<0.0001**	Residual	0.198
Chl-*a*	0.028	0.033	-0.087; 0.139	0.86	0.41		
Length	0.197	0.031	**0.096; 0.297**	6.39	**<0.0001**		
Layer(Na)	-0.054	0.078	-0.247; 0.134	-0.70	0.49		
Salinity × Layer(Na)	0.259	0.088	**0.065; 0.443**	2.95	**0.0033**		
Length × Layer(Na)	0.111	0.067	-0.036; 0.258	1.66	0.096		
Periostracum							
(Intercept)	3.500	0.048	**3.406; 3.596**	71.03	**<0.0001**	Site	0.130
Temperature	0.049	0.043	-0.036; 0.134	1.12	0.28	Residual	0.230
Salinity	-0.009	0.061	-0.131; 0.111	-0.14	0.89		
Chl-*a*	0.147	0.038	**0.071; 0.221**	3.88	**0.002**		
Length	0.439	0.041	**0.357; 0.522**	10.25	**<0.0001**		
Temperature × Length	-0.064	0.035	-0.135; 0.006	-1.77	0.082		
Salintiy × Length	-0.151	0.061	**-0.271; -0.029**	-2.38	**0.020**		
Calcite%							
(Intercept)	51.092	1.903	**47.338; 54.927**	26.84	**<0.0001**	Site	7.605
Temperature	-4.003	2.112	-4.508; 3.791	-0.19	0.85	Residual	9.656
Salinity	-7.163	2.476	**-12.03; -2.382**	-2.87	**0.012**		
Chl-*a*	-1.908	2.392	-6.662; 2.764	-0.80	0.44		
Length	-2.404	1.690	-5.807; 1.071	-1.367	0.18		

Equation 5.2) has an indirect effect on layer thickness by allowing salinity-induced compositional
changes and, therefore, the production of the most appropriate shell structure for specific environmental
conditions.

Under current scenarios, plasticity in shell production could confer *Mytilus* species an advantage when facing different water chemistries and predation levels. In fact, at high-latitudes and in the Baltic region, where durophagous (shell-breaking) predators are rare or absent and the water is more corrosive [44, 338], mussels are characterised by thinner, prismatic-dominated shells, providing a generally higher protection from dissolution. Conversely, at mid-latitudes, where durophagous predators are more abundant [338] and the $CaCO_3$ solubility of the water is lower [44], mussels display thicker, nacreous-dominated shells suggesting higher mechanical resistance.

5.4.3 Biomineralisation plasticity and future changes

Despite rapid global changes in the water cycle and salinity gradients [34], *Mytilus* species show a strong capacity to respond to heterogeneous environments. This plasticity in shell production could help to mitigate the emergent negative effects of changing water chemistry. In fact, the interacting effects of salinity and shell length, as well as a minor influence of temperature, on the periostracum (Figure 5.5a, b), which represents a strong chemical barrier to dissolution in molluscs [185, 188, 347], suggest enhanced periostracal thickness under decreasing salinities could mediate impacts of ocean acidification.

Although populations in high-latitude ecosystems will experience globally the most rapid acidification [6], the concurrent decrease in salinity predicts thicker prismatic layers and periostraca will be produced which increase protection from higher solubility conditions. Conversely, in temperate areas, increasing salinity would determine deposition of thicker shells and a relatively thicker nacreous layer and thinner periostracum, favouring mechanical shell resistance. However, predicted changes in periostracal thickening rate under different salinities depend on shell size and would be more evident in larger individuals (length > 48 mm) (Figure 5.5a).

In Greenland, where the rate of melting of the ice sheet has doubled in the last decade [16], low salinities during summer (< 20 psu) and high productivity (food supply) in coastal areas and fjords [348] could drive formation of thicker periostraca and increased relative thickness of organic-enriched (high wt%) calcitic layers. These changes could make Arctic *Mytilus* populations more resistant to future acidification. Differently, in the Baltic Sea, the forecasted decrease in salinity (maximum 45% reduction) [349], combined with a considerable physiological stress, would be particularly critical for mussels inhabiting already unfavourable conditions for calcification (salinity from 22 psu to 3 psu, low water $[Ca^{2+}]$ and $CaCO_3$ saturation state) [51]. Moreover, the reduced shell size of Baltic *Mytilus* does not predict formation of thicker periostraca, which would further increase vulnerability to dissolution. Impacts of changing salinity on this habitat-forming species, which contributes up to 90% of the Baltic

benthic biomass, could strongly affect the ecosystem, most likely resulting in substantial range restrictions towards higher salinity areas.

Although my results strongly support the hypothesis that biological mechanisms for variations in shell production can shape regional responses in *Mytilus*, changes of other biological drivers, such as predation pressure and primary production, could also have profound influences [44, 350]. In fact, as temperature rises, durophagous predators could expand their ranges towards polar regions [351], suggesting an increased vulnerability of thin-shelled individuals. However, predicted northward phytoplankton expansions and an overall increase in primary production at high latitudes [350], could favour periostracal growth potential in *Mytilus* and, thus, increased resistance to dissolution for all the size classes in polar and subpolar regions.

Mytilus shells have a thick periostracum and a marked compositional plasticity compared to other calcifiers that often compete with it for space (e.g. barnacles and spirorbid polychaetes). This layer provides a strong defence against shell dissolution allowing mytilids to survive in oligohaline waters (~ 5 psu) and extremely acidified conditions (e.g. hydrothermal vents) [347]. These factors may shift the ecological balance and community structure in favour of species with stronger resistance to corrosive conditions, such as mussels, when ocean waters become fresher and more acidic in future decades.

5.5 Conclusions

As hypothesised, plasticity in shell production and the spatial structure of environmental conditions drive regional differences in *Mytilus* shell deposition and composition, shaping spatial patterns of chemical and mechanical shell properties. Overall, mussel shell calcification decreased towards high latitudes, with salinity being the major predictor of geographical variations in shell production, mineral and organic composition. The marked compositional plasticity in calcareous shell components (prismatics and nacreous layers) suggests a higher resistance against dissolution for mussels in polar, low-salinity environments, and an enhanced mechanical shell protection from predators in temperate, higher-salinity regions. The strong response potential of *Mytilus* shell periostracum to heterogeneous environments suggests an increased resistance to ocean acidification in polar and sub-polar mussels, and a higher sensitivity of Baltic populations under future environmental conditions.

My findings demonstrate that biological mechanisms, driving spatial variability of mussel responses to interacting environmental factors, shape the complex geographic pattern of shell deposition and properties, dictating regional differences in *Mytilus* species sensitivity to future environmental change. As the magnitude of anthropogenic impacts continue to increase, further studies are needed to better understand the key biological processes mediating species' response to habitat alterations, especially

for those having both high climate sensitivity and disproportionately strong ecological impacts in shaping marine communities. This knowledge underpins our ability to predict accurately and reduce the damaging effect of climate change on future biodiversity under any range of scenarios [341]. This study has important implications because it clarifies the links between **i**) the mechanisms of biological variation, **ii**) the predicted shift in spatial co-occurrence of multiple environmental drivers, and **iii**) regional differences in the plastic responses and sensitivity of calcifying, foundation species to changing habitats. This understanding is of critical importance for making realistic projections of emergent ecological effects of global environmental changes, such as altered salinity regimes, and to improve our predictive accuracy for impacts on marine communities and ecosystems, and the services they provide.

Chapter Six

6 A century of resistance: blue mussels withstand environmental change *via* compensatory biomineralisation

External collaborators contribution

Dr Thierry Backeljau and Yves Samyn (Royal Belgian Institute of Natural Sciences, RBINS, Brussel, Belgium) provided access to archival *Mytilus edulis* specimens' collections at the RBINS; Mario Heinig (Ludwig-Maximilians University of Munich, Munich, Germany) helped with thermogravimetric analyses (TGA).

6.1 Introduction

Environmental change is a major force shaping the future of our oceans and the ecological services they provide [6, 26]. The complex interplay of unprecedented physical and chemical changes [5, 12], in addition to pervasive anthropogenic pressures [352], profoundly affects marine species physiology and phenology [18–21, 64, 72, 232], as well as community structure and ecosystem functions [6, 23–26, 222].

However, our ability to project how emergent ecological consequences of change will scale-up from species to community and ecosystems dynamics is limited [38, 341, 353]. Ecosystem-wide projections are strongly constrained by multiple direct and indirect environmental effects [25, 234], altered species interactions [56, 61, 354], and limited insights into compensatory dynamics that stabilise communities [55, 65]. Indeed, existing knowledge predominantly stems from short- to long-term experimental studies on single species or simplified "communities" [21, 23, 24, 40, 41, 44]. However, such conclusions may not necessarily translate to long-term acclimated or adapted natural populations within the complexity of a dynamic ecological system [38, 45–47]. Thus, a better mechanistic understanding of ecological driving processes in functioning ecosystems, across a relevant time scale to account for transgenerational plasticity and genetic change, is needed for building the theoretical framework necessary to anticipate the scope of ecosystem responses [38, 47, 64, 341, 353, 355].

To date, most scientific thinking has focused on climate change as a stressor for marine organisms [21, 26]. Whilst, environmental disturbances can constrain some species (i.e. stressors), they can provide advantages for others (i.e. resource), with the potential to buffer indirectly (e.g. *via* provision of more food [234]) the negative effects of abiotic change [23, 54]. Only recently, we have begun to appreciate the strength and ubiquity of indirect effects [234, 235, 356]. Their impacts on the key drivers of ecosystem dynamics highlight the need to incorporate species interactions [56], altered predation [61, 174], and compensatory processes [54, 65] into predictive models. This is especially so given their potential to accelerate [25, 56], attenuate [65, 234], or even reverse [61, 107] the direct effects of changing climate.

Compensatory dynamics have long been regarded as important stabilising mechanisms through which natural systems respond to environmental change [65, 357, 358]. However, these have been investigated almost entirely through the perspective of diversity loss, changes in density, and trophic adjustments [55, 65, 357]. Despite being straightforward to measure, these aspects only partially account for stability [55]. Indeed, compensation can arise from a broader range of responses at population [54, 359] or organism level [54, 106], acting through physiological adjustments to withstand disturbance (i.e. resistance) [23, 40, 357]. These mechanisms are, however, difficult to detect, because they may ultimately produce no net change in ecological functions of species and higher levels of organisation [65]. Despite an increasing focus on scaling compensatory responses to broader ecological networks [25, 38, 55], these have seldom been examined in species with key ecological roles, such as those forming structural habitats supporting marine communities [25, 61].

Many calcifying organisms play conspicuous and essential roles for healthy ecosystem functioning. Knowledge of detrimental effects of environmental change on calcifiers is largely drawn from laboratory experiments [21, 26, 40, 41], mesocosm studies [67], and *in situ* observations at CO_2 vents [56, 107], with molluscs described as being particularly vulnerable [21, 101]. These experimental approaches are, however, logistically constrained by inevitable trade-offs between the accuracy in manipulating single or multiple stressors, the recreation of realistically complex system conditions, and the duration of the study [24, 45–47, 355]. As a consequence, experimental observations, lasting from days to a few years, cannot incorporate very long-term acclimation and transgenerational genetic adaption, the processes identified as critical for conferring resistance [40, 41, 110, 360–362]. The ability of an organism to use plasticity of the phenotype, both within and across generations, in combination with long-term adaption of the gene complement, represent the most important factors for responding to altered environments [18, 54, 64, 110, 361, 362]. However, very long-term data, from years to decades, incorporating their effects on calcifiers' response potential, have been absent in the overwhelming majority of studies to date [41, 69, 73]. Thus, studies are needed in functioning ecosystems across very long times-scales accounting for acclimation and adaption [69, 355, 363], if calcifiers' resistance is to be better incorporated into the conceptual framework of global change.

A rarely used approach consists of evaluating changes over decades using museum collections [68–70]. By providing extensive records of net species responses to change [68], this approach is especially useful to **i**) assess ecological effects of stressors as they are gradually (or abruptly) introduced over time [73], and **ii**) account for potential long-term acclimation and local adaption [69], complementing laboratory and *in situ* experiments.

Marine mussels of the genus *Mytilus* are bed-forming foundation species throughout the world's eulittoral ecosystems [130]. Owing to their ecological importance [121], economic value for aquaculture (with ~590,000 t harvested worth 2,6 billion USD in 2016) [8], and projected sensitivity to ocean change [99, 110], the fate of mussels might be closely tied to those of productive coastal systems worldwide. As both calcifiers and habitat-formers, they have received much attention as primary indicators species [43, 108], especially the Atlantic blue mussel *Mytilus edulis*. However, because blue mussels are common and abundant, these are not generally preserved in museums and extensive archival collections of *Mytilus* from the last century are very rare.

Blue mussel shells perform a range of structural and protective functions. They are composed of three layers (see Section 1.2.10, Figure 1.7): (1) the organic periostracum, (2) the calcareous prismatic (calcite) layer, and (3) the calcareous nacreous (aragonite) layer. Growth and shell traits that influence ecological outcomes confer greater fitness on individuals with faster growth, rounder and thicker shells [43, 83, 99]. Shell growth and shape are major determinants in competition for space and predation [99, 108, 159]. The periostracum provides a substratum for shell formation, and a strong chemical protection against predatory and endolithic borers, as well as from corrosive waters [179, 185, 188]. Deposition of calcareous layers favours thicker-shelled individuals against shell-breaking (durophagous) predators [43, 83, 198]. Considering their projected vulnerability and plastic shell responses [70, 83, 99, 109], compensatory trade-offs in shell deposition and composition over time may be central in maintaining calcification, so fitness, in changing selective environments. Understanding compensatory adjustments in shell biomineralisation to withstand disturbance over historical times is essential to anticipate factors that will influence the persistence and ecological functions of these calcifiers in the near future. This is especially important given their potential to accelerate or stabilise against changes to supported communities and the broader ecological network.

Here, I quantify temporal trends in shell deposition, composition and morphology of the intertidal blue mussel, *Mytilus edulis*, over the last century, using unique museum collections of specimens sampled regularly from a single location along the Belgian coastline on the same substratum (i.e. stone breakwaters) between 1904 and 2016. Combining historical patterns of shell biomineralisation, with high-resolution, long-term datasets of key environmental descriptors (physical and biogeochemical) and abundance of keystone predators (different taxa and predation strategy), across a time-scale accounting for long-term acclimation and adaption, I tested the hypothesis that compensatory mechanisms in shell

biomineralisation can mitigate the predicted negative effects of change and altered species interactions within a complex ecological system.

6.2 Materials and methods

6.2.1 Physical environment

The Belgian coastline is a 65-km long, southwest to northeast directed, almost linear sandy shoreline between 51°05′N - 02°32′E and 51°05′N - 03°22′E, stretching from Calais in France to the Westerschelde estuary in the Netherlands (Figure 6.1a). The supralittoral zone is bordered by small dunes and the sublittoral habitat is a shallow sandy bottom, making this coastal profile unsuitable for mussel beds or other rocky shore organisms. However, intensive human activities over the last two centuries have provided suitable substrata for hard-bottom benthic assemblages, such as dikes, harbour-related infrastructures, and breakwaters [364, 365]. Among these, a whole series of regularly placed stone breakwaters (Figure 6.1b) was built to stabilise Belgian beaches against high levels of coastal erosion. Breakwaters are perpendicular to the coastlines and are made of basalt rocks or concrete. The length or each is about 400 m with a large part situated in the intertidal zone. The distance between two consecutive breakwaters is between 200 and 500 m (Figure 6.1c). According to Becuwe [366], the first breakwaters were built between 1815 and 1830, and many of those still in place were built around 1880.

6.2.2 Museum collections

I evaluated a total of 268 specimens of the blue mussel, *Mytilus edulis*, collected with an almost decadal frequency between 1904 and 2016 along the Belgian coastline at 13 sites between Oostende (51°14′16.27″N - 2°55′03.09″E) and Nieuwpoort (51°09′14.14″N - 2°43′23.62″E) (Figure 6.1b). Details on collection data for each sample are reported in Table 6.1.

Specimens collected during the 1904 - 1987 period were obtained from archival shell collections donated by the Royal Belgian Institute of Natural Sciences (RBINS, Brussels, Belgium). This unique collection of a single species is composed of both wet (body tissues and shells) and dry (shells only) specimens that were donated to the RBINS by private collectors and monitoring programmes during the last century. Only archival specimens with detailed information on species identity, location, depth and date of sampling were evaluated. Depending on availability, I selected adult individuals of *M. edulis* with a shell length of 39 - 66 mm, which were collected from the eulittoral zone on hard substrata (i.e. stone breakwaters). Thus avoiding any bias to the analysis of shell characteristics due to the different ranges of size classes available for each collection year. In all, 30 intertidal individuals were collected
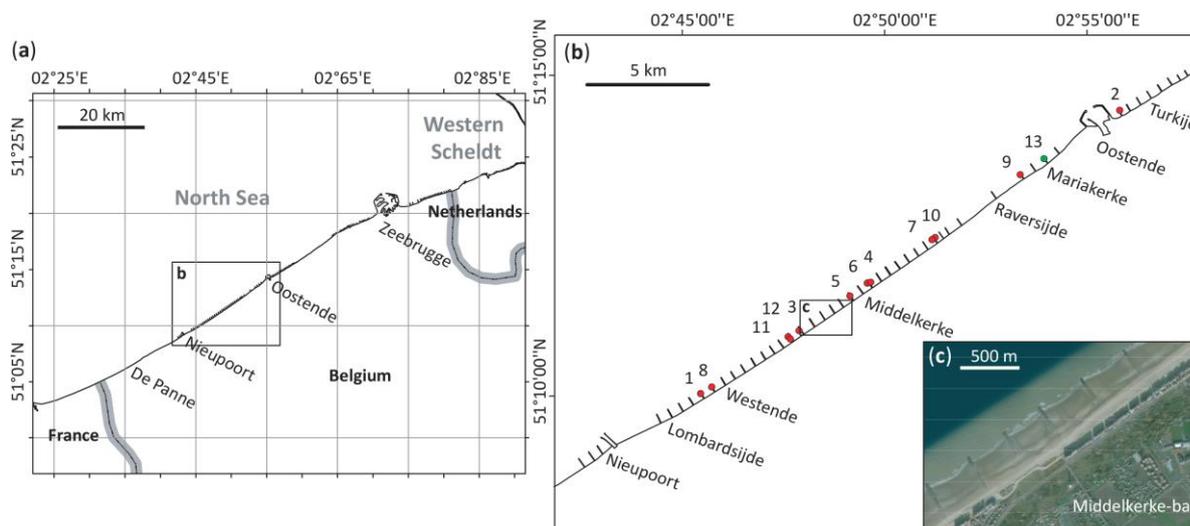
Figure 6.1 *Mytilus edulis* **collection sites: the Belgian coasts and its breakwaters system**

(**a**) The Belgian coastal system and collection area. (**b**) Map of the sampling locations and breakwater systems between Oostende and Nieuwpoort (51°09′14.14″N - 51°14′16.27″N and 2°43′23.62″E - 2°55′03.09″E) showing the sites where archival (1 - 12, red circles) and modern (13, green circle) specimens of *Mytilus edulis* were collected. (**c**) Detail of the regular series of stone breakwaters (low-tide) characterising the coastline of the sampling locations. Map created with ArcMap 10.5 (ArcGIS software by Esri, http://esri.com), background image and topographic details courtesy of OpenStreetMap (http://www.openstreetmap.org).

in 2016 by hand from the same location and substratum (Mariakerke). For each specimen, shell length was measured with digital callipers (0.01 mm precision) and used as a within-year proxy for age [108].

6.2.3 Mussel shell preparation

Left shell valves (*n* = 256) were set in polyester resin (Kleer-Set FF, MetPrep, Coventry, U.K.) blocks and prepared following the procedure in Section 2.2. Embedded specimens were sliced longitudinally along their axis of maximum growth (Figure 2.2) using a diamond saw and then progressively polished with silicon carbide paper (grit size P800 - P2500) and diamond paste (grading 9 - 1 µm). Photographs of polished sections (Figure 2.2) were acquired with a stereo-microscope (Leica M165 C equipped with a Leica DFC295 HD camera, Leica, Wetzlar, Germany) and shell thickness was measured using the Fiji software (v1.51w) [237]. Since many individuals had undergone evident environmental abrasion or dissolution, which removed the periostracum and prismatic layer closer to the anterior shell side (umbo), the thickness of the whole-shell, prismatic and nacreous layers was estimated at the midpoint along the shell cross-section.

Table 6.1 Collection details for archival and modern *Mytilus edulis* specimens used for the study

For each sampling site (site code as in Figure 6.1b), the year of collection, geographic location, site coordinates (latitude and longitude), sample size (*n*), average sample shell length, and collection type (wet: archival shell + tissue; dry: archival shell only; live: hand-collected mussels) are reported. All the blue mussel specimens were collected from the upper intertidal zone on stone breakwaters.

Site	Year	Location	Latitude (N)	Longitude (E)	*n*	Length ± SD	Collection type
1	1904	Westende	51.168751	2.762271	25	55.70 ± 4.01	Wet
2	1907	Oostende - Turkije	51.240798	2.929859	13	47.64 ± 6.71	Wet/Dry
3	1922	Westende - Middelkerke	51.183889	2.798636	25	53.56 ± 5.72	Dry
4	1936	Middelkerke	51.194368	2.823377	15	54.08 ± 5.58	Wet/Dry
5	1937	Middelkerke	51.190883	2.815409	12	55.38 ± 4.78	Wet
6	1938	Middelkerke	51.194658	2.822695	15	53.97 ± 3.44	Wet/Dry
7	1950	Raversijde	51.206001	2.849758	25	55.59 ± 3.43	Wet
8	1966	Westende	51.170646	2.766574	13	54.17 ± 5.61	Wet
9	1971	Mariakerke	51.224318	2.889219	22	51.21 ± 3.89	Wet
10	1974	Raversijde	51.206325	2.849672	26	54.31 ± 3.15	Wet
11	1986	Westende - Middelkerke	51.182291	2.794256	26	55.35 ± 2.84	Wet
12	1987	Westende - Middelkerke	51.181848	2.794758	21	51.97 ± 4.55	Wet
13	2016	Mariakerke	51.228793	2.898682	30	47.38 ± 2.03	Live

The proportion of calcite was estimated as:

$$calcite\% = \left(\frac{prismatic\ thickness}{whole\text{-}shell\ thicknes}\right) \times 100$$

Equation 6.1

Periostracum thickness was measured at the posterior edge where it attaches to the external side of the prismatic layer, to estimate the fully formed organic layer that was unaffected by any degree of decay or abrasion [185]

6.2.4 Organic content analysis

Thermogravimetric analyses (TGA) were performed to estimate the weight proportion (wt%) of organic matrix within the calcitic prismatic layer. *Mytilus edulis* specimens were selected from three years 1904,

1950 and 2016 (n = 10, 10 and 11, respectively) to explore differences in shell organic content over the last century. The periostracum was removed by sanding, and prismatic layer tiles (8×5 mm, n = 31) were isolated along the posteroventral shell margin (where there was no nacreous aragonitic layer underneath). Tiles were cleaned, air-dried and then finely ground. Ten milligrams of this powdered shell were tested with a thermogravimetric analyser (TGA Q500, TA Instruments, New Castle, DE, U.S.A.). Samples were subjected to constant heating from ~25°C to 700°C at a linear rate of 10°C min^{-1} under a dynamic nitrogen atmosphere and weight changes were recorded (see protocol in Section 2.3.5). The wt% of organic matter within the prismatic layer was estimated as the proportion of weight loss during the thermal treatment between 150°C and 550°C [241] (Figure 2.5).

6.2.5 Shell shape analysis

Shape analyses of *M. edulis* shells were carried out through a geometric morphometrics approach [108, 246]. I performed an elliptic Fourier analysis (EFA) of outlines (Section 2.4.3) [258, 276] to examine between-individuals variations in shell shape over the last century.

Outlines of the lateral view of left shell valves (n = 268) were digitised and converted into a list of *x-y* pixel coordinates, which were used as input data. Outlines were smoothed to remove digitisation noise, and geometrically aligned [108] following the protocol in Section 2.4.4-2.4.8. An EFA was then computed on the resulting coordinates from shapes invariant to size and rotation. After calibration (Figure 6.2), I chose seven harmonics, encompassing 99% of the total harmonic power. Four coefficients per harmonic (28 descriptors) were extracted for each shell outline and used as variables quantifying geometric information [108, 258].

A principal component analysis (PCA), with a singular value decomposition method, was performed on the matrix of coefficients to define axes capturing the most of shape variation among individuals. Calculated principal components (PCs) were considered as new shape variables. The first two PCs capturing 80% of the shape variance were selected for modelling [108]. To understand the contribution of individual variables to shell shape, I reconstructed extreme outlines along each PC (Figure 6.3a). The first eight PCs (99% of outline variation) were analysed with a multivariate analysis of variance (MANOVA) to test for significant effects of collection year and shell length (size) on shape variances. To visualise outline differences among mean shapes from each collection year and shapes at the extremes of the morphospace, I generated deformation grids and iso-deformation lines through mathematical formalisation of thin plate splines (TPS) analysis [259] (Figure 6.3b, c).
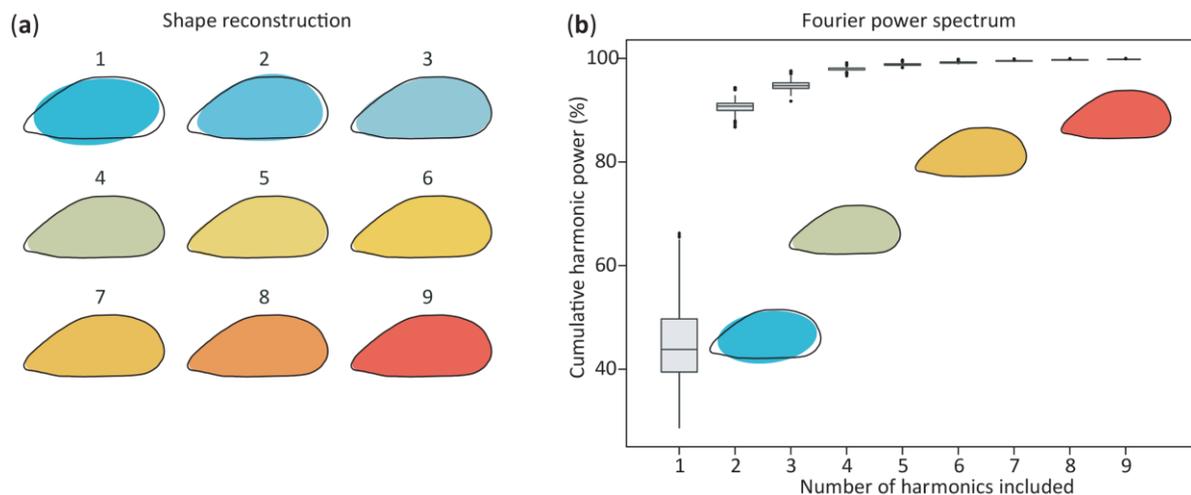
Figure 6.2 Calibration methods for EFA of shell outlines
(**a**) Average shell shape reconstruction for different numbers of harmonics (from 1 to 9). Six-seven harmonics gave a satisfactory reconstruction of shell outlines and for nine harmonics the approximation was almost perfect. (**b**) Cumulative spectrum of harmonic Fourier power. The power is proportional to the harmonic amplitude and can be considered as a measure of shape information. I evaluated the appropriate number of harmonics to retain (7), so that their cumulative power gathered 99% of the total cumulative power [283].

6.2.6 Environmental dataset

I selected key environmental drivers based on their known influence on blue mussel growth, their spatio-temporal availability, and the forecasted major ocean alterations under climate change [5]. I evaluated long-term datasets of surface water temperature, salinity, dissolved oxygen, water pH and chlorophyll-*a* concentration (Chl-*a*), the latter being used as a proxy for food supply [53].

Time series of daily measurements for surface water temperature (1900 - 1984) and salinity (1904 - 1984) in the area 51°14′N - 51°05′N, 02°55′E - 02°32′E were obtained from the International Council for the Exploration of the Sea Data Centre (ICES, http://www.ices.dk/marine-data/) [367] and the Integrated   Marine   Environmental   Readings   and   Samples   (IMERS, http://www.vliz.be/vmdcdata/imers/) [368]. Chlorophyll-*a* (Chl-*a*) concentration (1971 - 1984) and pH (1951 - 2016) datasets from the same area were obtained from the Management Unit of the Mathematical Model of the North Sea (MUMM, at the RBINS, http://www.mumm.ac.be/datacentre/) and ICES Data Centre.

Environmental datasets of water temperature, salinity, Chl-*a* and dissolved oxygen for the period 1985 - 2016 were generated using the Copernicus Marine Environment Monitoring Service (CMEMS, http://marine.copernicus.eu/) [328]. These climate data are composed of high-resolution physical and biogeochemical assimilated (integration of observational and predicted information) daily observations (Appendix B, Environmental Datasets). These assimilated data presented several advantages compared to traditional measurements due to their high spatio-temporal resolution, both coverage and frequency
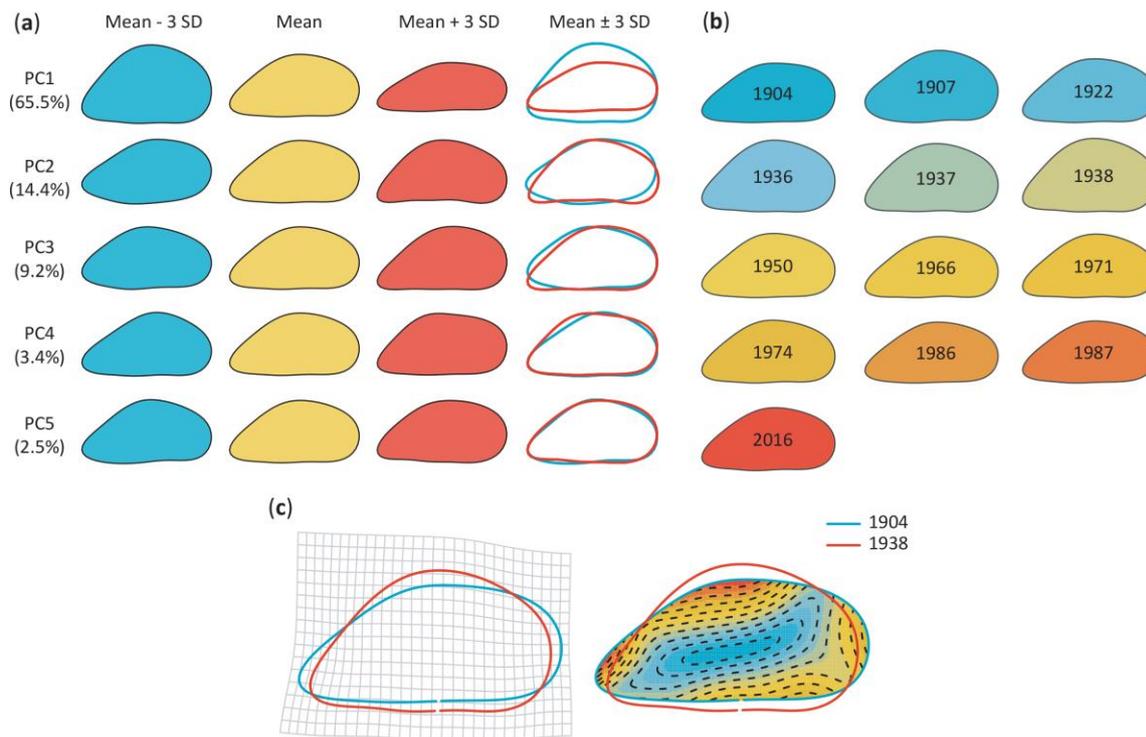
Figure 6.3 PCs contribution and mean shape variation across the morphospace
(**a**) Contribution of the first five shape variables (PCs) to shape variation. The average shell shapes for the lateral view were represented for increasing values along each PC (Mean − 3SD, Mean, Mean + 3SD), and shapes at the extremes of each variable were compared (Mean ± 3 SD). (**b**) *Mytilus edulis* average shell shapes for each collection year. (**c**) Differences between mean shapes at the extremes of the morphospace were observed with deformation grids (left), depicting the bindings required to pass from an extreme (1904) to another (1938), and iso-deformation lines (right), representing the outline regions subjected to different degrees of change (blue: low deformation; red: strong deformation).

(continuous daily measurements across 31 years), advanced calibration and validation (i.e. high correlation with discrete field measurements) [329, 330].

6.2.7 Mussel-predators dataset

To understand temporal changes in predation pressure on intertidal *M. edulis* beds and its influence on shell characteristics, I obtained long-term datasets on the abundance of decapods (macrobenthos and their planktonic larvae), sea birds (gulls), and gastropods (dog whelks). I selected these key taxa based on their predation strategy (i.e. durophagy and drillers), influence upon the intertidal zone and known effects on mussel shell [116, 121, 126, 174, 369].

Information on abundance of decapods (number of individuals per sample) in the Belgian coastal area (51°20′N - 51°05′N, 03°12′E - 02°32′E) were obtained from the ICES Data Centre and European Environment Agency (EEA, https://www.eea.europa.eu/). Macrobenthic decapod datasets for the 1978 - 2017 period included analyses based on fisheries data, the ICES North Sea Benthos Survey (1986 -

2000) [370], the North Sea Benthos Project 2000 (1999 - 2001) [371], and literature (1978 – 2017) [58, 59, 379–381, 89, 372–378]. Decapod larvae datasets for the 1958 - 2009 period were obtained from the ICES Data Centre, MUMM, and EEA. These include data from the Continuous Plankton Recorder (CPR) survey [382], operating in the North Sea on a monthly basis since 1946. This represents the largest marine biological time series available, consisting of monthly abundance data of holozooplankton and merozooplanktonic larvae of decapods, and other invertebrate taxa.

Numbers and location of breeding pairs of seagulls along the Belgian coastline (Nieupoort - Zeebrugge) were obtained from literature for four dominant gull species: *Larus ridibundus* (black-headed, 1969 - 2007), *L. fuscus graellsii* (lesser black-backed, 1985 - 2007), *L. argentatus* (herring, 1960 - 2007), and *L. canus* (common, 1975 - 2007) [89, 383–390]. Local reports were used to assess the presence of the common dog whelk, *Nucella lapillus*, along the Belgian shores [391, 392].

6.2.8 Statistical analysis

Generalised additive mixed models (GAMMs) [288] were used to account for the hierarchical structure of the datasets, consisting of multiple *M. edulis* specimens ($n = 12$ - 30 replicates) from each collection site, and to allow for flexible specification of the dependence structure of the response on the covariates.

I carried out data exploration following the protocol in Section 3.1.1. Initial inspection revealed no outliers. Pairwise scatterplots and variance inflation factors (VIFs) were calculated to check for collinearity between input variables. VIF values < 2 indicated an acceptable degree of correlation among covariates to be included within the same model. Preliminary inspection of residual patterns showed heteroscedasticity in most models. This required using different combinations of link functions and probability distributions, allowing greater variation for large mean values (i.e. gamma or inverse Gaussian for continuous data, and Poisson or negative binomial for count data) for individual responses. When changing the underlying distribution did not stabilise the variance, a ln-transformation of the response was required, except for calcite% and wt% measurements.

6.2.9 Time series analysis

To understand long-term and seasonal changes in environmental conditions and predation pressure over the last decades, I used GAMMs to model time series of environmental descriptors (surface temperature, salinity, Chl-*a*, dissolved oxygen and pH) and the abundance of predators. Time series, consisting of discontinuous daily observations per year, were expressed as continuous monthly-averaged measurements. Fixed covariates were *month* (seasonal variation, expressed as numeric 1, …, 12 indicator), *year* (trend), both fitted as smoothers, and their interaction that was defined through a smooth

tensor product interaction [288] (Equation 6.2). The tensor product interaction is used to represent functions of covariates, which are measured in different units (i.e. different magnitude of change), by allowing the smooth effect of one variable (i.e. seasonal variation) to vary as a smooth function of the second variable (i.e. trend). This allowed the within-year spline (effect) to change smoothly with the between-year effect.

$$Series_i = \beta_0 + f_{seasonal}(Month_i) + f_{trend}(Year_i) + f(Month_i, Year_i) + \varepsilon_i$$
$$\varepsilon_i \sim N(0; \Lambda\sigma^2)$$

Equation 6.2

where *Series_i* is the ith observation of the time series, β_0 is the intercept, $f_{seasonal}$ and f_{trend} are smooth functions for the seasonal and trend features of interest, *f* is the smooth function (interaction) of the two time variables, and ε_i the normally distributed and correlated (Λ) error. By allowing the seasonal component to change in time along with the trend (Equation 6.2), I modelled **i**) any seasonal, or within-year variation, **ii**) any trend, or long-term change, in the mean level of the time series, and **iii**) any interaction between seasonal and trend features of the data.

Boundary and numbers of knots (limits and dimensions of the basis used for the splines) were manually selected, while effective degrees of freedom were estimated by the smooth function [291]. A cubic regression spline basis was used for the trend smooth term, while a cyclic cubic spline basis was used for the seasonal smooth term to avoid discontinuity between January and December values [288]. Models indicated significant within-year residuals autocorrelation, which required the use of different residual correlation structures [287]. Autoregressive (AR, for equally spaced time series) or conditional autoregressive (CAR, for unequally spaced observations) models, nested within each *year*, were fitted to the residuals to account for temporal autocorrelation.

A negative binomial GAM with a log link function was used to model the numbers of seagull breeding pairs as a function of year and to compare between the four species (Equation 6.3). The log link function ensures positive fitted values, and the negative binomial distribution is typically used for count data that are overdispersed with respect to a Poisson distribution. Covariates were *year* (smoothed with a cubic spline), gull *species* (categorical, four levels: black-headed, lesser black-backed, herring, and common), and their interaction.

$$BreedingPairs_{ij} \sim NB(\lambda_{ij}; \theta)$$
$$\ln(\lambda_{ij}) = \beta_0 + f(Year_j) \times Species_i + \varepsilon_{ij}$$
$$\varepsilon_{ij} \sim N(0; \sigma^2)$$

Equation 6.3

where *BreedingPairs_{ij}* is the jth observation for the response variable (number of breeding airs) for the ith species (*i = black-headed, lesser black-backed, herring, common*). The response follows a negative binomial distribution (*NB*) described by two parameters, λ_{ij} and θ, where λ_{ij} is the mean value and θ is

the shape parameter (overdispersion parameter) of the negative binomial distribution. $\ln(\lambda_{ij})$ is the latent value for the jth observation of the ith species (with a log link function), f is the smoothing function (cubic regression spline), β_0 is the intercept, ε_{ij} is independently distributed error with expectation 0 and variance σ^2.

6.2.10 Modelling

Individual GAMMs were used to describe shell thickness ($n = 256$) and shape ($n = 268$) with respect to time, and to compare between shell measurements (thickness of whole-shell, prismatic layer, nacreous layers or periostracum, and calcite%) and morphological traits (PC1 and PC2 from EFA coefficients) (Equation 6.4). Fixed covariates were collection *year* and shell *length* (both fitted as smoothers). To estimate biologically meaningful intercepts, I centered the input variables. *Year* of collection was centered to 1900, while *length* to the mean shell length of the sample. To incorporate the dependency among observations from the same sampling site, *site* was used as a random intercept.

$$Response_{ij} = \beta_0 + f_1(Year_i) + f_2(Length_{ij}) + Site_i + \varepsilon_{ij}$$

$$Site_i \sim N(0; \sigma^2_{Site})$$

$$\varepsilon_{ij} \sim N(0; \sigma^2)$$

Equation 6.4

where $Response_{ij}$ is the jth observation for the response variable (thickness or shape variable) in site i ($i = 1, \ldots, 13$), f is the smoothing function (cubic regression spline), β_0 is the intercept, ε_{ij} is the noise and $Site_{ij}$ is the random intercept, which are assumed to be independently distributed with expectation 0 and variance σ^2.

I modelled the weight proportion (wt%) of organic matrix ($n = 31$) as a function of year of collection (categorical, three levels: 1904, 1950, and 2016) and prismatic layer thickness (continuous), to test for differences among time periods and associations with shell deposition. The response variable was coded as a value from 0 to 1; therefore, a generalised linear model (GLM) with a beta distribution and a logistic link function was used. Pair-wise contrasts with a Bonferroni correction were then used to test ($\alpha = 0.017$) for differences in the wt% of organics among collection years. To estimate the potential interactions of local environmental conditions on mussel shell biomineralisation over decades, I coupled shell thickness measurements with comprehensive long-term environmental datasets (i.e. temperature and salinity). First, descriptive statistics for individual factors per year (i.e. mean, median, standard deviation, 10th, 25th, 75th and 90th percentiles, minimum and maximum annual values) were calculated. To provide a first order approximation of the water conditions prevailing during the lifespan of sampled specimens, I expressed the descriptive statistics as average values for the 6-year period prior to
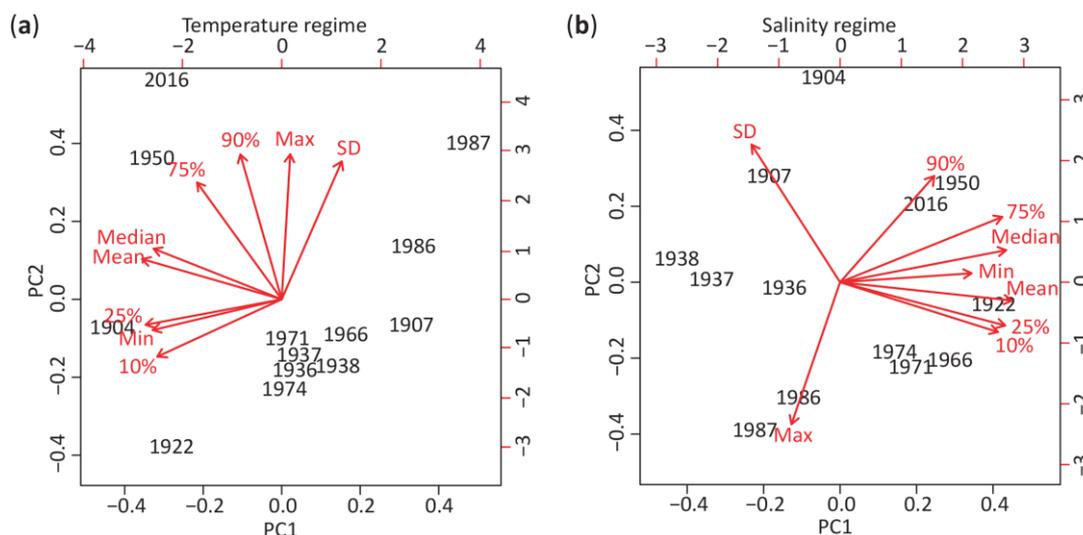
Figure 6.4 Environmental regimes
Biplots of PCAs performed on (**a**) temperature-related and (**b**) salinity-related descriptors.
Top and right axes report the loadings scores. Red arrows represent the loadings vectors of
the original variables on the PCs, showing how each environmental descriptor is related to
the plotted PCs (Table 6.2). The direction of the arrows indicates the direction of increase
of the variable values. Mean, median = annual mean and median; Min, Max = minimum
and maximum values; SD = standard deviations; % = parameter percentile. The position of
different collection years on the plane is determined by their relative scores on the plotted
PCs.

Table 6.2 PCA summary
Environmental descriptors and their loadings (correlations) on the first two principal
components from PCAs performed on temperature-related and salinity-related variables.

| | **Temperature** | | **Salinity** | |
Descriptor	PC1	PC2	PC1	PC2
Mean	**-0.44**	0.16	**0.42**	-0.08
Median	**-0.40**	0.19	**0.40**	0.13
SD	0.23	**0.46**	-0.24	**0.57**
Max	0.05	**0.49**	-0.10	**-0.58**
Min	**-0.40**	-0.13	**0.29**	0.04
90%	-0.11	**0.51**	0.26	**0.41**
75%	-0.29	**0.38**	**0.38**	0.27
25%	**-0.43**	-0.11	**0.40**	-0.19
10%	**-0.39**	-0.23	**0.38**	-0.22

collection (4-years for individuals from 1904 due to lack of environmental data prior to 1900). I then
standardised all the descriptors to account for differences in measurement units. For standardisation, I
subtracted the sample mean from the variable values and divided them by the sample standard deviation
$[z_i = (x_i - \bar{x}) / \sigma_x]$. Descriptive statistics for surface temperature and salinity were analysed in
separate PCAs for key environmental variables (i.e. I performed separate PCA for all the **i**) temperature-

related and **ii**) salinity-related variables) (Figure 6.4, Table 6.2). I then used the first two PCs from temperature and salinity PCAs as input variables in GAMMs (all fitted as smoothers) to quantify the proportion of shell variation explained by abiotic conditions, and to describe relationships between shell thickness and environmental regimes. PCs from EFAs (Figure 6.3a) were also included as a proxies for mussel age (PC1) and food supply (PC2) [108]. Collection site was included as a random intercept.

6.2.10.1 Model optimisation

Models were optimised by first selecting the random structure and then the optimal fixed component following the protocol in Section 3.7. The principal tools for model comparison were the corrected Akaike Information Criterion (AICc) and likelihood ratio tests. Random terms were selected on prior knowledge of the dependency structure of the dataset. Visual inspection of residual patterns indicated violation of homogeneity in most cases. This required the use of variance structures (generalised least squares) allowing the residual spread to vary with respect to predictors. To assess the presence of temporal dependence in model residuals, I used (partial) autocorrelation functions, pACF and ACF, for regularly spaced time series, or variograms, for irregularly space ones. When temporal autocorrelation was found, the best residual correlation structure was selected through comparisons of models with differing stochastic trend models in the residuals. The fixed component was optimised by rejecting non-significant interaction terms only that minimised the AICc value. For environmental GAMMs with multiple smooth terms, selection was carried out using cubic regression splines with shrinkage and a restricted maximum likelihood (REML) estimation method [288, 393]. Shrinkage smoothers are useful since they can have 0 smoothing. This made it possible to perform a backwards selection of all non-significant terms in a single step. Indeed, all smoothers with 0 amount of smoothing (non-significant terms) can be dropped simultaneously from the model, without the need to re-fit and compare multiple nested models through a step-wise procedure. Final models were validated by inspection of standardised residual patterns to verify assumptions of normality, homogeneity and independence, as in Section 3.7.2. The proportion of variance explained by the models was quantified with conditional determination coefficients (cR^2) [394].

6.2.10.2 Model predictions

Optimal models were used to describe changes in mussel shell characteristics over the last century. To identify periods of statistically significant change in the time series, I first used the method of finite differences to compute the first derivative of the fitted splines. Without an equation for the spline, a derivative cannot be calculated analytically. Hence, this method samples a number of infinitesimally distant points on the fitted spline and estimates the slope between pairs of points throughout the trend.
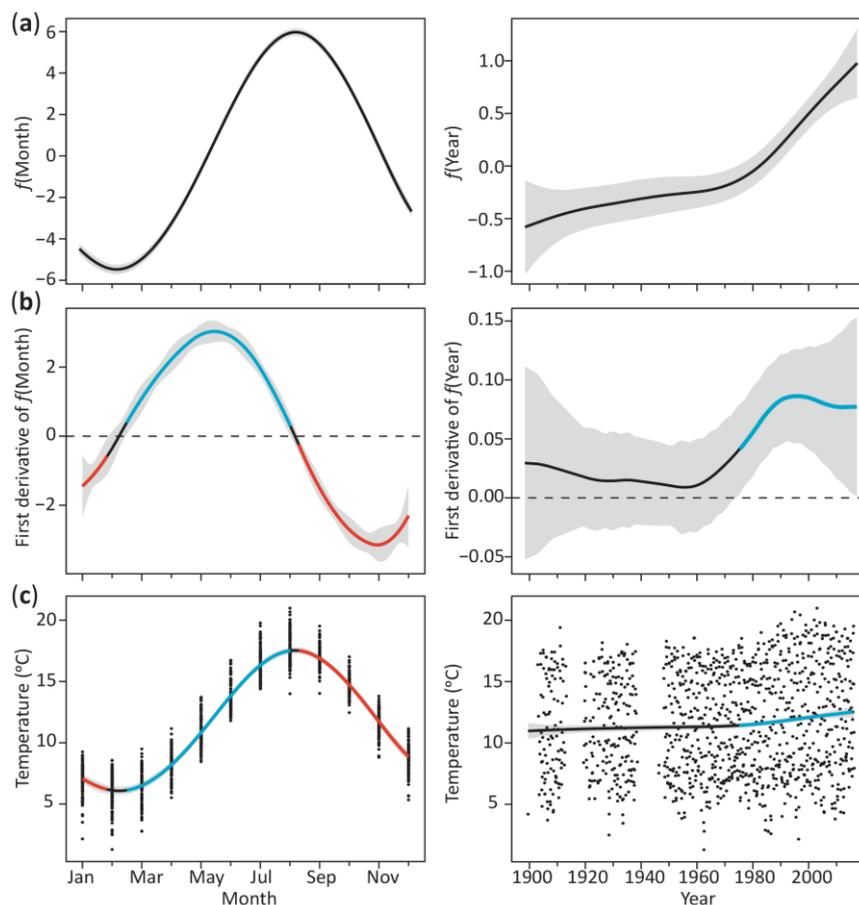
Figure 6.5 Finite difference methods and periods of change in the time series
(**a**) Smooth terms for the optimal sea surface temperature model. The seasonal term (cyclic cubic regression spline) (left), indicating within-year changes, and the trend term (cubic regression spline) (right), indicating long-term changes in the mean level of the time series between 1900 and 2016. Note the very different scales of the two splines, which illustrate the relative degrees of variation in the seasonal (~12°C, within-year variation) and trend term (~2°C, between-year variation). (**b**) First derivatives for the fitted seasonal and trend splines estimated with the method of finite differences. This method approximates the first derivative of a fitted spline by **i**) choosing a set of points p on the function and another set p' positioned at a very small distance from the first set, **ii**) evaluating the fitted trend spline at the location p and p', and **iii**) computing the rate of change (slope) in the function between the pair of points. The shaded areas represent simultaneous 95% confidence intervals, reflecting the uncertainty of the fitted functions. These were calculated through a posterior simulation-based approach following Ruppert *et al.* [319]. Sections of the spline derivatives where zero is not included in the simultaneous 95% confidence interval represent period of significant change in the time series (blue: increase; red: decrease). (**c**) Intervals of significant change are then superimposed to the fitted trends to show period of significant increase or decrease for the seasonal and long-term surface temperature patterns.

The first derivative of the spline is then approximated to the change in the slopes between these points. Following Ruppert *et al.* [319], I then used a posterior simulation-based approach to generate simultaneous 95% confidence intervals drawing from the Bayesian covariance matrix of the model derivative (10,000 draws). This approach calculates the appropriate critical value (scaling factor of the

standard error) to estimate a 95% confidence interval reflecting the uncertainty of the fitted function. Estimated points of the spline's derivative at which the simultaneous 95% confidence interval does not include zero represent periods of significant change in the time series. This information was used to identify which part of the predicted trend were significantly increasing or decreasing over time (Figure 6.5).

6.3 Results

6.3.1 Historical patterns of shell deposition and calcification

I identified different linear and non-linear temporal patterns of shell biomineralisation in adult *M. edulis* ($n = 256$) along the Belgian coastlines, depending on the shell layer analysed (Table 6.3). The overall shell thickness of blue mussels increased by 57% over the last 112 years (Figure 6.6a). The deposition of calcareous shell components increased over time, with prismatic and nacreous layers from modern shells being on average 48% and 74% thicker, respectively, than those of the oldest archival specimens (Figure 6.6a). Periostracum thickness changed non-linearly over the last century (Figure 6.6a), while there was no variation in the proportion of calcitic layer thickness.

The wt% of organic content in the prismatic layer ($n = 31$) of modern mussels [2016; mean (SD) = 2.03 wt% (0.21)] was 13% lower (larger proportion of $CaCO_3$) than that of archival shells [1904; mean (SD) = 2.28 wt% (0.19)], indicating an increase in shell calcification over time (Figure 6.6b). A negative correlation between wt% of organics and prismatic thickness was detected (Figure 6.6c), revealing higher proportions of $CaCO_3$ and thicker, so more calcified, shells in modern *M. edulis*.

6.3.2 Among-years shell shape variations

A principal component analysis (PCA), performed on the harmonic coefficients from an EFA of shell outlines ($n = 268$), revealed marked variations in *M. edulis* shell height, ligament (PC1), and ventral margin shape (PC2) among collection years (Figure 6.3b, 6.7a). PC1 (65.5%) was significantly correlated with shell length (Table 6.3), indicating shape changes during growth (age), while PC2 (14.4%) showed shell variations observed under changeable food regimes (chlorophyll-*a* concentration) [108] (Figure 6.7b). There was no detectable relationship between PCs and year of collection (Table 6.3).
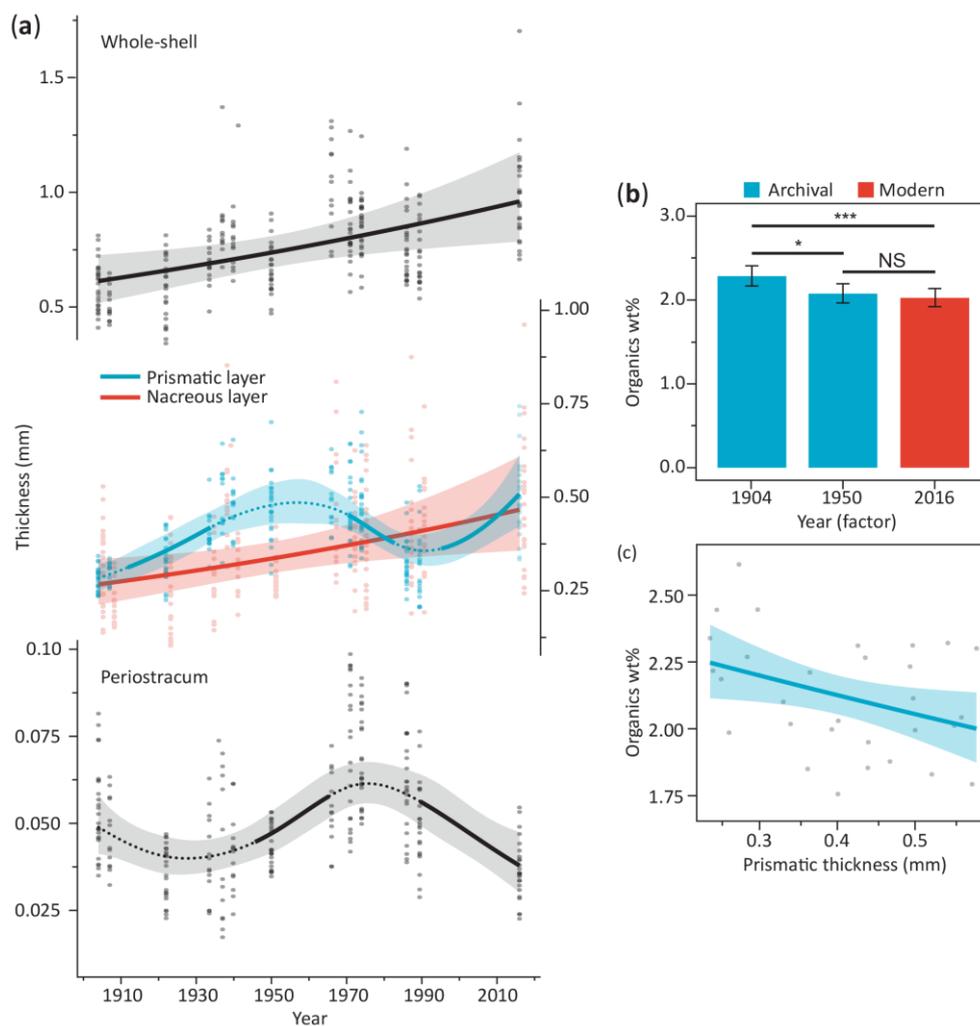
Figure 6.6 Historical patterns of blue mussel shell thickness, organic content, and calcification between 1904 and 2016

(**a**) Predicted trends in the thickness of whole-shell, calcareous, and organic layers for the mean shell length of the sample (53.13 mm) over the last century. Whole-shell ($df = 1$, F = 2.57, $p = 0.0059$) and nacreous layer ($df = 1$, F = 2.58, $p = 0.01$) thickness increased linearly over time. Prismatic layer deposition significantly decreased between 1970 and 1983, and increased in the 1911-1934 period and after 1995 ($edf = 4.09$, F = 11.50, $p < 0.0001$). Periostracum thickness increased between 1945 and 1962, and showed a 29% decrease after 1985 ($edf = 3.51$, F = 17.58, $p < 0.0001$). Simultaneous 95% confidence intervals (shaded areas) and periods of significant change (continuous lines) are reported. (**b**) Variations in organic content between archival (blue bars) and modern (red bar) mussel shells. Pair-wise contrasts indicated lower proportions of organics in modern that archival specimens [mean = 0.25 wt%; $z = 3.15$, $p = 0.0017$ (***)], a marginally significant difference between 1904 and 1950 [mean = 0.21 wt%; $z = -2.46$, $p = 0.014$ (*)], and no difference between 1950 and 2016 [mean = 0.04 wt%; $z = 0.62$, $p = 0.53$ (NS)]. (**c**) Relationship between the organic wt% and prismatic thickness, indicating a positive association between layer deposition and calcification level ($z = -2.17$, $p = 0.03$).
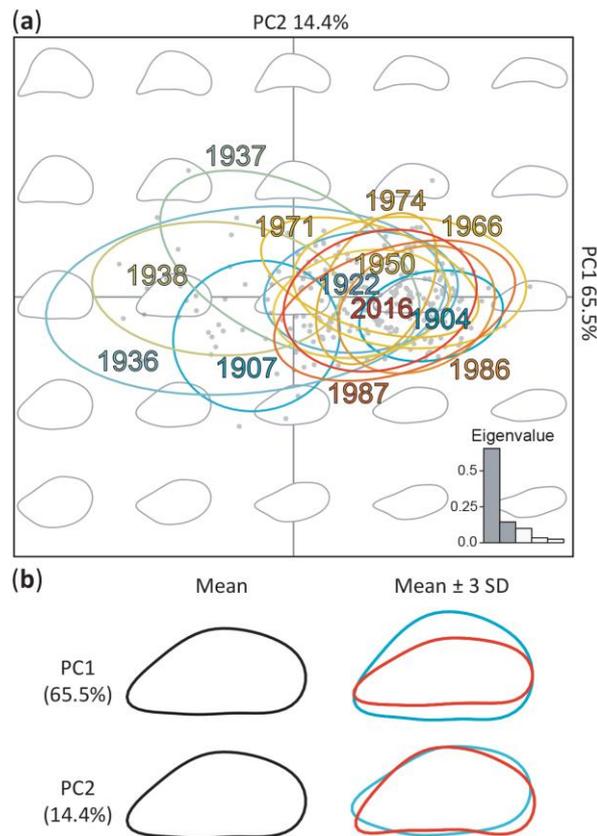
Figure 6.7 Temporal variation in blue mussel shell shape
(**a**) Scatterplot of the first two principal components (PCs) from a PCA performed on elliptic Fourier coefficients of lateral shell views, showing significant among-year differences across the morphospace (background) (MANOVA: Wilk's λ = 0.094, approximate $F_{12,\,242}$ = 6.64, $p < 0.0001$). Mussels from 1907 and 1936 - 1938 had rounder shells than other collection years with more elliptic profiles (Figure 6.3b, c). Confidence intervals (95%, continuous lines) and the proportion of variance explained by the first five PCs (histogram) are reported. (**b**) PCs contribution to the mussel shape variation for mean, + 3SD (red) and – 3SD (blue) PC values. PC1 (65.5%) captured variations in shell height and ligament angle, with a transition from round to elongated shell for increasing values. PC2 (14.4%) indicated more concave ventral shell profiles for increasing values.

6.3.3 Long-term changes in environment and predation

I identified significant long-term changes and variations of the seasonal patterns over time in key environmental conditions and the abundance of key mussel predators along the Belgian coastlines across the last century (Figure 6.8a, Table 6.4).

6.3.3.1 *Environmental conditions*

Surface water temperature significantly increased by 1.10°C between 1975 and 2016, with large seasonal changes (13.6°C within-year variation). Temperature increased in spring and summer (April - September) with maximum temperatures increasing by 2.0 - 3.6°C in May and August, and no variation in winter and autumn, revealing a progressive shift towards an earlier seasonal peak (Figure 6.9a, b).

Table 6.3 GAMMs summary statistics of shell characteristics over time
M. edulis shell thickness (different layers) and shape (PC1-PC2; Figure 6.4) variation with collection year and shell length (size) is reported (Equation 6.4). (*edf*: estimated degrees of freedom)

Parameter	*edf*	Estimate	SE	F	*p*-value	Random effect	SD
Whole-shell							
(Intercept)	1	6.402	0.091	70.26	**<0.0001**	Site	0.163
Year	1	0.004	0.001	2.75	**0.0059**	Residual	0.190
Length	1	0.002	0.003	0.58	0.56		
Periostracum							
(Intercept)	1	3.895	0.025	156.74	**<0.0001**	Site	0.066
f(Year)	3.51	_	_	17.58	**<0.0001**	Residual	0.261
Length	1	0.001	0.004	0.21	0.84		
Prismatic layer							
(Intercept)	1	5.958	0.026	227.49	**<0.0001**	Site	0.095
f(Year)	4.09	_	_	11.50	**<0.0001**	Residual	0.009
Length	1	0.009	0.003	3.41	**0.001**		
Nacreous layer							
(Intercept)	1	5.568	0.121	45.98	**<0.0001**	Site	0.210
Year	1	0.005	0.002	2.58	**0.010**	Residual	0.335
Length	1	-0.003	0.005	-0.67	0.50		
Calcite%							
(Intercept)	1	0.571	0.028	20.02	**<0.0001**	Site	0.048
Year	1	0.000	0.000	-0.98	0.33	Residual	0.091
Length	1	0.003	0.001	2.17	**0.030**		
Shape-PC1							
(Intercept)	1	-0.001	0.001	-1.72	0.11	Site	0.001
Year	1	1.6E-05	9.3E-06	1.69	0.12	Residual	0.001
Length	1	4.1E-05	1.4E-05	3.01	**0.0029**		
Shape-PC2							
(Intercept)	1	-1.4E-04	1.6E-04	-0.89	0.39	Site	0.001
Year	1	2.6E-06	2.5E-06	1.05	0.32	Residual	0.001
Length	1	6.9E-06	8.9E-06	0.77	0.44		

There was no identifiable long-term variation in mean salinity between 1904 and 2016, while the seasonal pattern changed over time (0.6 psu increase in within-year variation). Salinity increased most in autumn (October - November) and early spring (February - April), with maximum mean change of 0.25 psu in March and April, indicating more gradual and defined within-year variations over time (Figure 6.9c, d). Chl-*a* mean concentration significantly increased by 0.17 mg/m^3 between 1971 and 2016. The seasonal concentration peaks (2.5 mg m^{-3} within-year variation) progressively shifted from spring and summer (April, August) to winter and spring (February, May), with the gradual formation
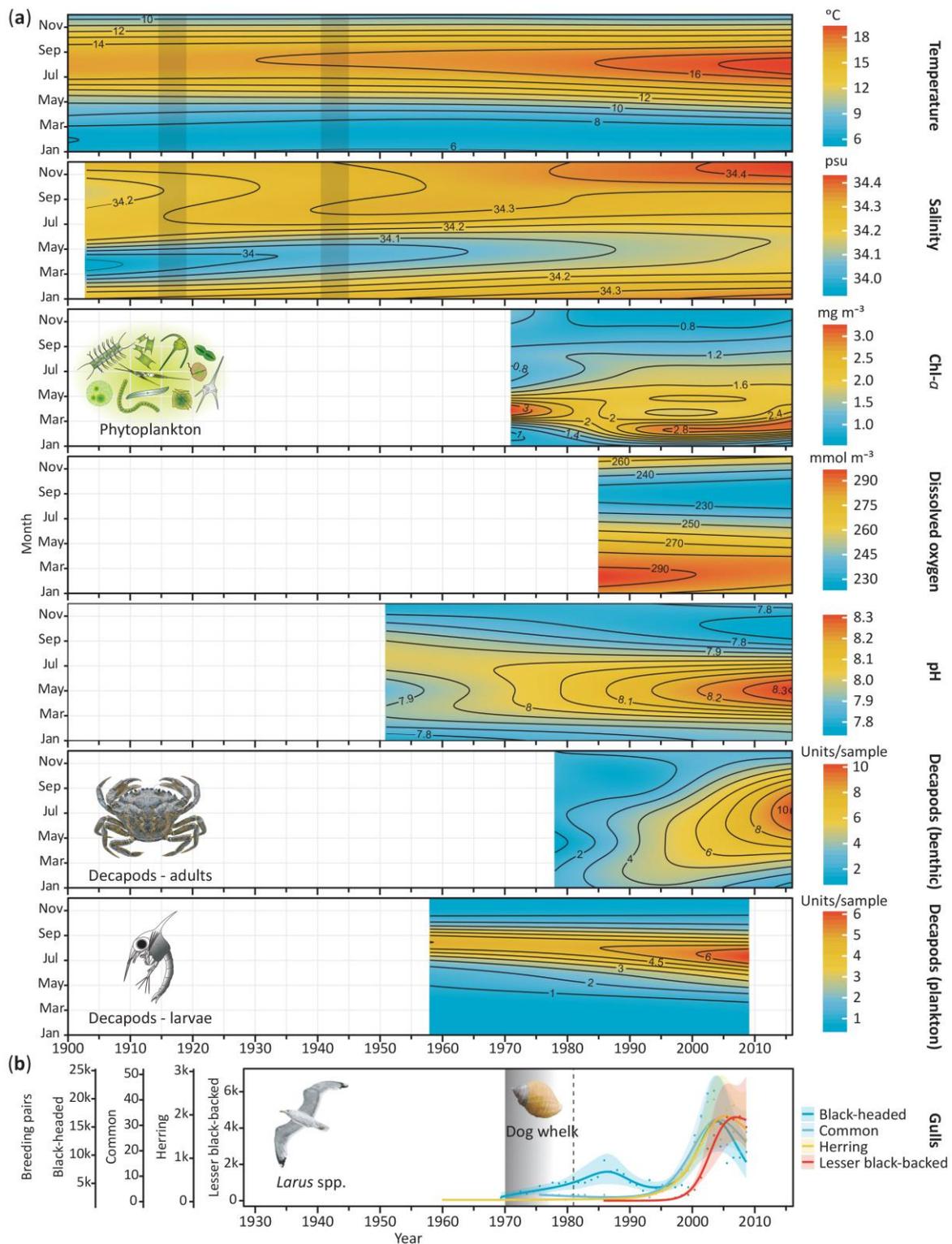
Figure 6.8 Historical trends (long-term change) and seasonal patterns (within-year variations) in local abiotic and biotic conditions over the last century

(**a**) (Left panels) Contour plots showing predictions of long-term changes (between-year) and variations in the seasonal (monthly) patterns of key abiotic and biotic descriptors over the last 116 year. Variations in the response are represented through individual colour-scales (red: high values; blue: low values) and isolines (continuous lines). Shaded areas (temperature and salinity) represent historical periods with missing data (1914 - 1918 and 1940 - 1945) for which I extrapolated my predictions. (Right panels) Perspective plots of the prediction plane as a result of the tensor product interaction [288] between trend and seasonal components of the

data. Note the different scales of the ordinate (response, marks as in the colour-scale) and abscissa (year axis), highlighting differences in the magnitudes of change and temporal availability for each parameter. (**b**) Long-term changes in the numbers of breeding pairs in four dominant seagull species. Note the marked among-species differences in abundance. Progressive decrease (blended area) and complete disappearance in 1981 (vertical dashed line) of *Nucella lapillus* from the study location.

of a unique and earlier peak in late winter (February – March) after 1990 (Figure 6.9e, f). Dissolved oxygen concentration decreased by 6.64 mmol m^{-3} between 1985 and 2016, with no change in the seasonal pattern (65.82 mmol m^{-3} within-year variation). Dissolved oxygen decreased all-over the year, except in March - April, with a maximum of -11.9 mmol m^{-3} in winter (February) (Figure 6.10a, b). Surface pH changed significantly in seasonality (0.29 within-year variation) between 1950 and 2016, but no overall long-term change was detected. pH increased in spring and summer (March - June) with a maximum of 0.45 in May, and decreased in late summer and autumn (August - November), resulting in a more variable pH regime and sharper within-year gradients (Figure 6.10c, d).

6.3.3.2   Predator abundance

The abundance of macrobenthic decapods (1978 - 2017) and their planktonic larvae (1958 - 2009) increased 2.5 and 1.9 times, respectively over the last 60 years (Figure 6.11a, b). Benthic decapods increased significantly between spring and autumn during each of the last 30 years, while planktonic larvae increased between late winter and summer, with the seasonal abundance peak shifting from late to early summer after 1988 (Figure 6.11c, d), indicating decapods appearing earlier in the year. Gull numbers significantly increased between 1960 and 2007 (Figure 6.8b) showing an exponential growth of breeding pairs, up to 200 - 1,500 times, in four dominant Belgian species: *Larus ridibundus* (black-headed), *L. canus* (common), *L. argentatus* (herring), and *L. fuscus graellsii* (lesser black-backed) (Table 6.5). While, the keystone predatory gastropod *Nucella lapillus* began to decrease in numbers in the 1970's and disappeared from Belgium in 1981 (Figure 6.8b).

6.3.4   Environmental influence on mussel shell deposition

Individual PCAs performed on descriptive statistics for surface temperature and salinity allowed **i**) exploring whether mussel characteristics were more closely related to estimates of mean conditions vs the variability of these parameters (Supporting Table A.6), and **ii**) creating from collinear descriptors a set of independent variables (PCs), characterising temperature and salinity "regimes" (Figure 6.4, Table 6.2).

Temperature variability (PC2) had a stronger positive effect on shell deposition than food supply (shape PC2) over time (cR^2 = 0.41, Figure 6.12a). Prismatic thickness increased with more variable temperature conditions (PC2) (cR^2 = 0.49, Figure 6.12b), while the nacreous thickness increased with age (shape

Table 6.4 GAMMs summary statistics (time-series analyses) of environmental descriptors and abundance of predators

The time range of the data, and the significance of the trend term [*f(Year)*, long-term change in the mean level of the time series], seasonal pattern [*f(Month),* within-year variation], and the interaction between seasonal and trend terms [*f(Year, Month),*] are reported (Equation 6.2).

Parameter	*edf*	Estimate	SE	F	*p*-value
Temperature (1900 - 2016)					
(Intercept)	1	-0.153	0.058	7.06	**0.0080**
f(Year)	2.90	_	_	17.39	**<.0001**
f(Month)	8.97	_	_	894.05	**<.0001**
f(Year, Month)	39.81	_	_	0.97	**<.0001**
Salinity (1903 - 2016)					
(Intercept)	1	32.420	1.654	384.32	**<.0001**
Year	1	0.001	0.001	1.19	0.28
f(Month)	4.98	_	_	5.35	**<.0001**
f(Year, Month)	7.95	_	_	0.05	0.12
Chl-*a* (1971 - 2016)					
(Intercept)	1	-7.308	2.286	10.22	**0.0015**
Year	‡1	0.004	0.001	11.11	**0.00094**
f(Month)	8.71	_	_	143.24	**<.0001**
f(Year, Month)	28.62	_	_	0.89	**<.0001**
Dissolved oxygen (1985 - 2016)					
(Intercept)	1	676.238	116.358	33.78	**<.0001**
Year	‡1	-0.209	0.058	12.90	**0.00037**
f(Month)	8.14	_	_	225.60	**<.0001**
f(Year, Month)	0.01	_	_	0.01	0.72
pH (1951 - 2016)					
(Intercept)	1	5.598	3.682	2.31	0.13
Year	1	0.001	0.002	0.63	0.53
f(Month)	3.28	_	_	4.39	**<.0001**
f(Year, Month)	9.75	_	_	0.63	**<.0001**
Decapoda - adults (1978 - 2017)					
(Intercept)	1	4.466	0.343	169.00	**<.0001**
f(Year)	2.36	_	_	19.87	**<.0001**
f(Month)	3.64	_	_	1.33	**0.00016**
f(Year, Month)	4.18	_	_	0.12	**0.0072**
Decapoda - larvae (1958 - 2009)					
(Intercept)	1	-26.565	7.378	12.97	**0.00034**
Year	‡1	0.013	0.004	12.54	**0.00043**
f(Month)	7.46	_	_	92.46	**<.0001**
f(Year, Month)	9.48	_	_	0.12	**0.00070**

‡ Linear significant effect of the covariate on the response

Table 6.5 GAM summary statistics (parametric and smooth terms) of breeding pairs variation with seagull species and collection year, and their interaction (Equation 6.3)

Parameter	*edf*	Estimate	SE	F	*p*-value
Parametric coefficients					
(Intercept) [+]	1	8.651	0.081	11344.38	<.0001
SpeciesHerring	1	-4.524	0.100	2058.26	<.0001
SpeciesLesserBlackBacked	1	-5.316	1.463	13.21	0.0004
SpeciesCommon	1	-7.137	0.230	963.23	<.0001
Approximate significance of smooth terms					
f(Year) × SpeciesBlackHeaded	6.02	_	_	11.74	<.0001
f(Year) × SpeciesHerring	7.65	_	_	624.91	<.0001
f(Year) × SpeciesLesserBlackBacked	3.96	_	_	399.31	<.0001
f(Year) × SpeciesCommon	5.12	_	_	75.06	<.0001

[+]SpeciesBlackHeaded is used as the intercept.

PC1), food supply (shape PC2), and less variable salinity regimes (PC2; $cR^2 = 0.34$, Figure 6.12c). Periostracum thickness increased with less variable temperature (PC2) and salinity regimes (PC2; $cR^2 = 0.43$, Figure 6.12d). This suggests changes in environmental variability were generally more important in supporting shell deposition than altered mean water conditions.

6.4 Discussion

Using a unique museum collection of a single species from a single site of regular sampling across 112 years, I showed how a calcifier predicted to be vulnerable to climate change could thrive under physiologically stressful conditions (i.e. greater energetic demand associated with variable environments and increased predation) through plastic compensatory responses of shell biomineralisation. These results demonstrate that a detailed understanding of temporal responses in foundation species to changing environments, within the complexity of a human-impacted ecological system, is critical when making projections of changes in resistance and ecological functions of calcifiers. It is further important for analyses aimed at anticipating emergent effects of change on supported communities and the broader ecological network.

The Belgian coast represents a complex and dynamic littoral system. Although its geomorphology has been profoundly shaped by rapid urbanisation and anthropogenic pressure (i.e. harbour development, beach nourishment and re-profiling) [364], its regular series of stone breakwaters has been substantially preserved over the last 140 years [364, 366]. By providing an ideal control in terms of substratum, and
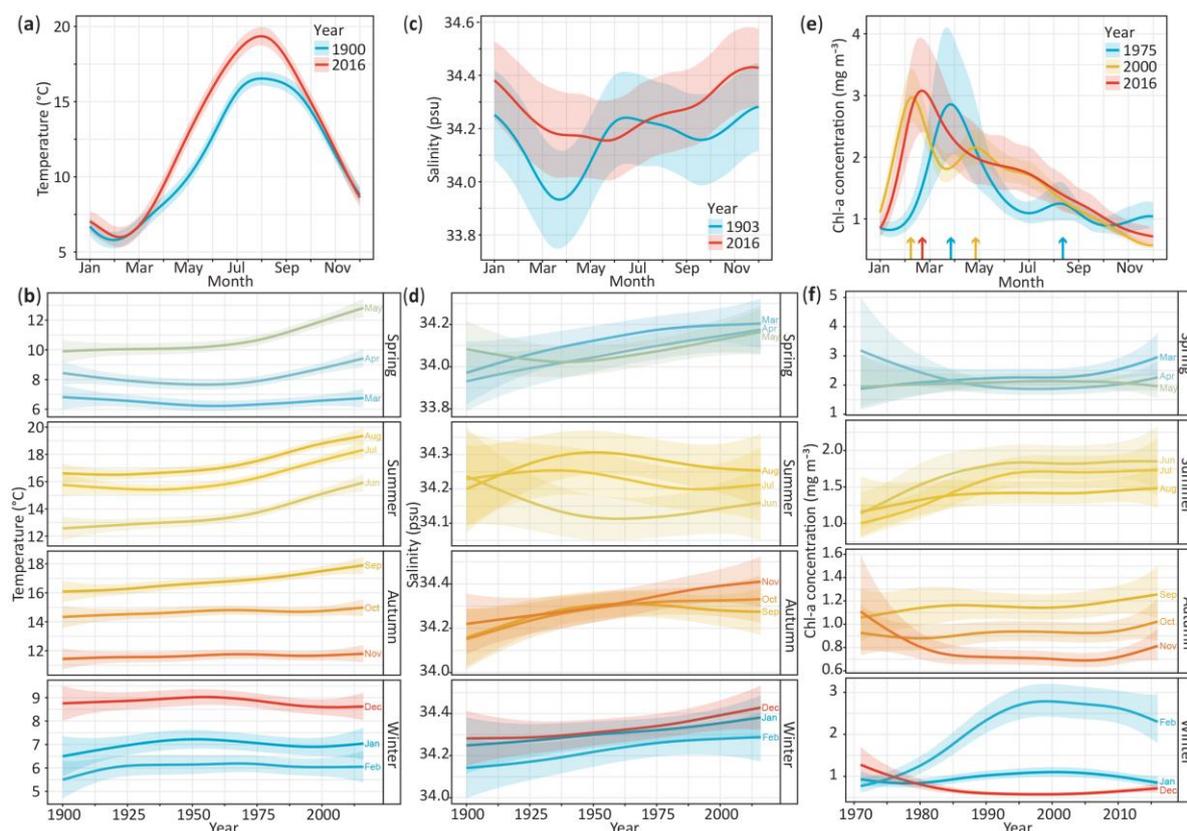
Figure 6.9 Historical variation in the seasonal pattern (within-year variation) for water temperature, salinity, and Chl-*a* concentration

Predicted seasonal (monthly) variation for (**a**) surface water temperature for 1900 and 2016, (**c**) salinity for 1904 and 2016, and (**e**) Chl-*a* concentration for 1975, 2000 and 2016. Trends (long-term variations) for each month grouped by quarter/season for (**b**) surface water temperature over the 1900 - 2016 period, (**d**) salinity over the 1904 - 2016 period, and (**f**) Chl-*a* concentration over the 1975 - 2016 period. Simultaneous 95% confidence intervals (shaded areas) are reported for each prediction.

exposure to tidal, hydrological, and sedimentary coastal dynamics, this breakwater system represented an ideal long-term field experimental location to evaluate historical changes in hard-bottom communities. The use of archival specimen of *M. edulis* collected from the same location and substratum over the last century, coupled with long-term datasets of key environmental descriptors and predators, allowed the observation of temporal responses of a calcifying foundation species under "controlled" conditions (relative to previous historical studies) to climatic (i.e. dynamism and predation) and local anthropogenic pressures (i.e. eutrophication and biodiversity loss) within a naturally complex "experimental" system.
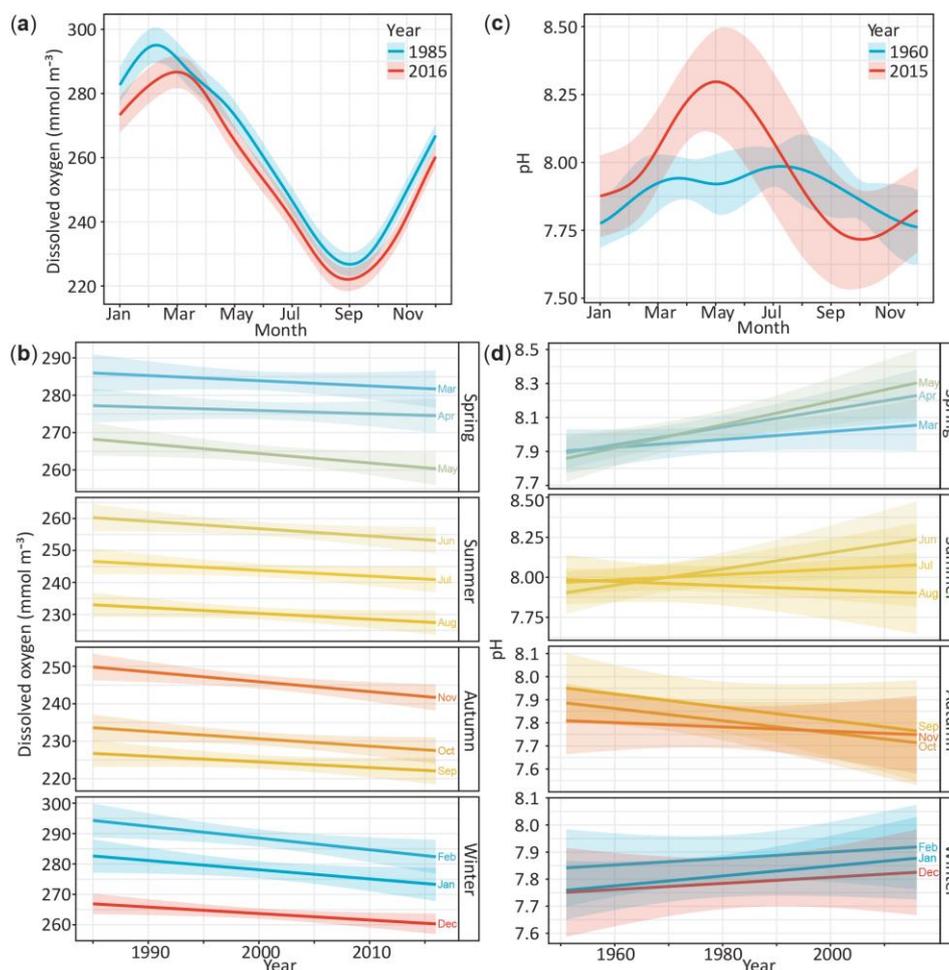
Figure 6.10 Historical variation in the seasonal pattern (within-year variation) for dissolved oxygen concentration and water pH
Predicted seasonal (monthly) variation for (**a**) surface dissolved oxygen for 1985 and 2016, and (**c**) pH for 1955 and 2016. Trends for each month grouped by quarter/season for (**b**) surface dissolved oxygen over the 1985 - 2016 period, (**d**) surface water pH over the 1955 - 2016 period. Simultaneous 95% CIs (shaded areas) are reported for each prediction.

6.4.1 Historical mussel shell patterns

Contrary to predictions [70, 83, 99], blue mussel shell calcification markedly increased over the last 112 years (i.e. 57% thicker shells and 13% less organics, Figure 6.6). This observation is in contrast with historical and experimental evidences on marine calcifiers suggesting either a depression of calcification [21, 70, 99], trade-off between skeletal deposition and growth [73], or general stability [69]. Historical studies have indicated shells of modern *Mytilus californianus* (2009 - 2011 periods) to be 42% thinner per unit size than shells from Native American Middens (~1000 - 2400 years BP) [70]. *Mytilus* larvae and adults produced thinner, weaker, and smaller shells under laboratory-simulated carbonate conditions predicted for 2100 than individuals grown under present-day conditions [83, 99, 223]. In coralline algae (*Lithophyllum*, *Pseudolithophyllum*, and *Lithothamnion* spp.), thallus thickness decreased by 2 - 2.3 times between 1981 and 2012 in species generally regarded as thick with no change
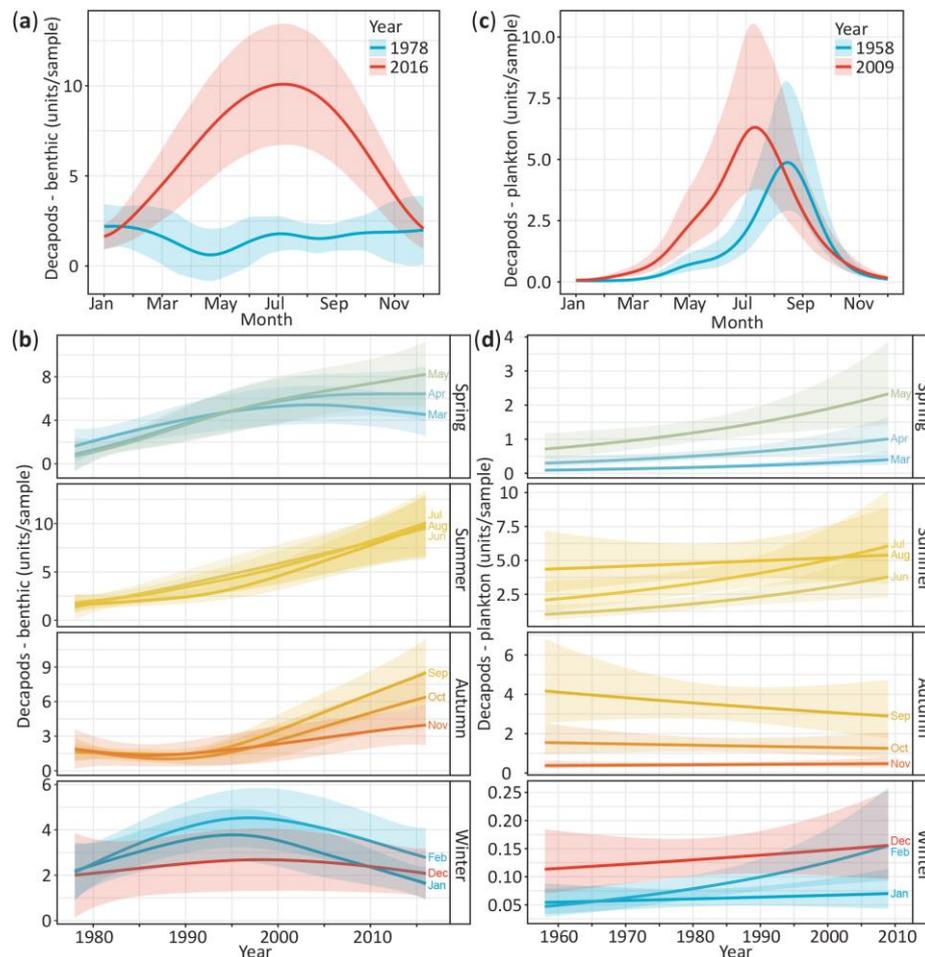
Figure 6.11 Historical variation in the seasonal pattern (within-year variation) for benthic decapods and their planktonic larvae

Predicted seasonal (monthly) variation for (**a**) the abundance of macrobenthic decapods for 1978 and 2016 (10 units/sample within-year variation), and (**c**) decapod's planktonic larvae for 1978 – 2009 (6 units/sample within-year variation). Trends for each month grouped by quarter/season for (**b**) the abundance of macrobenthic decapods over the 1978 - 2016 period, and (**d**) decapod planktonic larval abundance over the 1978 - 2009 period. Simultaneous 95% confidence intervals (shaded areas) are reported for each prediction.

in internal skeletal metrics, while the reverse was observed for species generally regarded as thin [73], suggesting trade-offs between morphological type and internal calcification. Meanwhile, shell density increased by 3.4% in the brachiopod *Calloria inconspicua* from New Zealand over the last century, but with no change in other shell metrics [69]. Reported differences in susceptibility suggest how species-specific types and magnitude of compensatory processes might be responsible for the wide range of observed responses (change vs stability) among calcifiers, despite the ubiquitous presence of multiple stressors.

As expected, blue mussel shells increased elongation and concavity of the ventral margin during growth [108], while there was no identifiable trend in shell shape over the last century (Figure 6.7). This is in contrast with laboratory experiments reporting morphological alterations in *M. edulis* as a compensatory
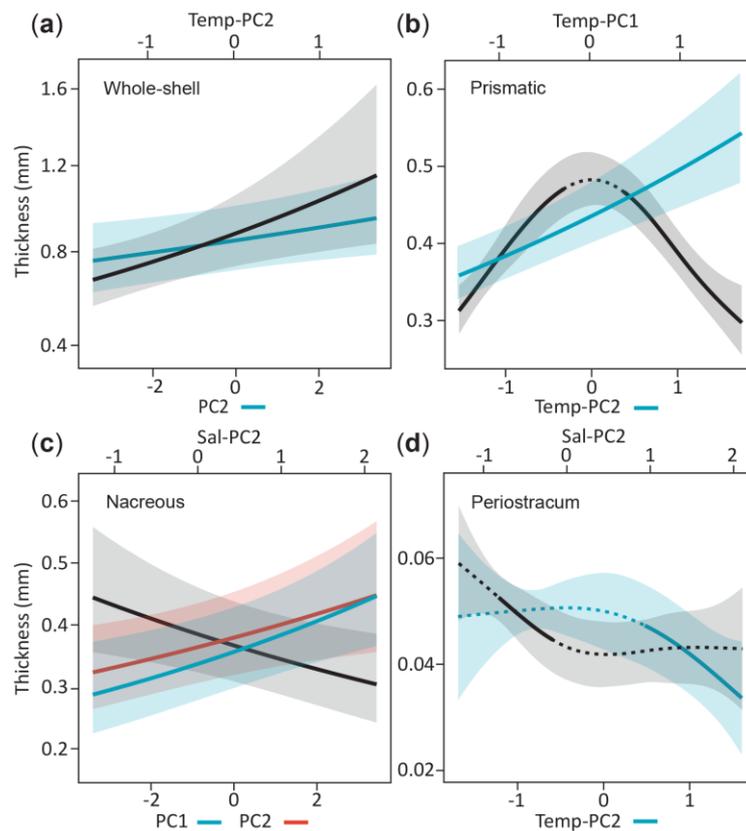
Figure 6.12 Environmental effects on the deposition of blue mussel shell layers
(**a**) Whole-shell thickness was positively related to PC2 of shape ($df = 1$, F = 5.29, $p = 0.02$) and temperature ($edf = 2.86$, F = 5.32, $p = 0.002$). (**b**) Significant relationships between prismatic thickness, PC1 and PC2 of temperature ($edf = 2.37$, F = 26.08, $p < 0.0001$; $df = 1$, F = 34.81, $p < 0.0001$, respectively) were detected. (**c**) The nacreous layer was positively correlated with shape PC1 ($df = 1$, F = 8.88, $p = 0.0031$) and PC2 ($df = 1$, F = 5.81, $p = 0.017$), and negatively with salinity PC2 ($df = 1$, F = 6.20, $p = 0.013$). (**d**) Periostracum thickness had a non-linear relationship with temperature PC2 ($edf = 2.12$, F = 4.23, $p = 0.008$) and salinity PC2 ($edf = 2.23$, F = 3.36, $p = 0.04$).

adjustment to enhance protection against predators in thin-shelled individuals under physiologically stressful conditions [99]. The observed lack of changes in shell shape of *M. edulis* in a natural system over time supports the potential for an increased resistance of mussels to altered environments through a compensatory increase of shell deposition.

6.4.2 Long-term environmental patterns

The long-term changes of surface temperature are in line with both global ocean trends [5] and local North Sea trends [89]. Seasonal alteration of salinity mirrored Chl-*a* surface patterns (increasing salinity corresponding to lower food levels), reflecting the projected alterations of local water cycles and fresh river inputs [17]. Increased Chl-*a* concentration has been widely documented and correlated to increasing human activity and land use over the last 60 years, enhancing inorganic nutrients and organic

carbon river loads [52, 395]. These processes have led to a significant coastal eutrophication, boosting phytoplankton abundance and, therefore, primary production (food supply) [52, 396]. Decreased oxygen levels are in line with higher temperatures (lower oxygen solubility) and primary production, increasing biological activity. Contrary to global trends, pH showed no long-term decrease and an increasingly marked seasonality over the last 75 years. This observation supports the strong impacts of quantitative and qualitative alterations of phytoplankton communities on coastal carbon cycling and sequestration [52], as well as eutrophication significantly buffering the effects of ocean acidification on the carbonate chemistry of surface waters [32].

These results suggest the formation of generally more variable surface conditions (Figure 6.8a) could lead to energetically demanding regimes for intertidal mussels [127, 397]. These factors may, however, determine indirectly, *via* provision of more food [107] and buffered ocean acidification [32], less unfavourable energetic conditions for calcification (Figure 6.13). From an historical perspective, this suggests current field populations have already made energetic adjustments to cope with increased physiological stress in this region, with no evident alteration of shell growth.

6.4.3 Historical change of predation regime

These results indicate an overall increase of predation pressures on *M. edulis* with a rapid shift to shell-breakers (durophagy) dominated regimes (i.e. increasing numbers of decapods and seagulls) after the local extinction of a keystone predatory driller (i.e. dog whelk) (Figure 6.13). Long-term increases in numbers of decapods in the North Sea, which have doubled over the last 30 years (Figure 6.8a) [398, 399], have been explained by a temperature-driven, abrupt ecosystem shift during the 1980's [59]. This has driven profound changes in benthic and pelagic assemblages, favouring greater abundance and range of warm-water decapod taxa [59, 398]. These changes, which are correlated with surface temperature [58, 398], could reflect increasing predation pressure of decapods on fish recruits and benthic bivalves, both decreasing in numbers over recent decades [59] (Figure 6.13). Increasing planktonic decapod larvae and shifting seasonal peak abundance suggest an increased and earlier recruitment [398]. Recruitment could be favoured by boosted primary production and decreasing predation pressure on meroplanktonic larvae due to increased overfishing over the last 40 years [58, 59, 400], decimating white-fish stocks in the North Sea [401, 402]. The key role of decapods as predators of bivalves, and other benthic organisms, suggests a strong influence of their increasing abundance on the coastal trophic linkages of blue mussels (Figure 6.13), and the potential to amplify effects of abiotic change at different trophic [59] and organisational levels [55, 89].

European populations of the dog whelk *Nucella lapillus* started declining in concomitance with the widespread employment of tributyltin (TBT) in the 1970s [403], leading to its complete eradication

from Belgium in 1981 [391]. Dog whelks represent keystone predators of mussels in the intertidal where sea star pressure is significantly reduced [174], and are expected to have a stronger selective impact on *M. edulis* under altered carbonate conditions [369]. However, their disappearance could have changed significantly predator-prey coastal dynamics [174] (Figure 6.13), with potential benefits for mussel fitness.

The abundance of seagulls increased exponentially in four dominant species along the Belgian coastlines between the mid 1990's and 2007 (Figure 6.8b). This has been linked to an amplification of the temperature signal, through decapod larvae and adults abundance, boosting food supplies and, therefore, breeding success for multiple gull colonies in the North Sea [89] (Figure 6.13). Seabird breeding success is partially controlled by abundance and nutritional quality of the prey items that the parents feed to their chicks [404], with decapods, pelagic fishes, and fishing discard being dominant components [401]. A greater abundance of decapods [398] may, therefore, have positive implications for the breeding success of seagulls [89], by providing food and a significant $CaCO_3$ supply for both eggshell and bone development in chicks during the breeding season (May - June) [183], matching with decapod peak abundance (Figure 6.8a). The North Sea supports many important fisheries, generating substantial quantities of bycatch and discard for seabirds [405]. However, the general decline in volume of discard [401] coupled with deleterious effect of overfishing on fish stocks [58, 402], overall reducing food supply for coastal seabirds, has been documented to exacerbate prey-switching tendencies in generalist scavenging birds, such as seagulls, resulting in an increased predation on coastal taxa [401].

Seagull influence on blue mussel populations has been widely documented in Europe with the potential to remove from 30% up to 70% of the natural and cultured intertidal stock [116, 406]. The increase in Belgian gull numbers [89], coupled with enhanced prey-switching behaviour [401], are likely to increase predation on other coastal resources, among which are the dominant and easily accessible intertidal blue mussel beds. Seagulls represent an important link between changes in marine and terrestrial ecosystems [89] (Figure 6.13). This suggests how climate- and anthropogenic-driven changes in marine organisms can extend to the avian fauna, and so propagate to intertidal and terrestrial food-webs around seabirds colonies [89, 401].

Terrestrial

Eulittoral

Subittoral

Overfishing [399, 401]
- Reduction fish stock & discard [400] →
less food supply →
exacerbated prey-switching behaviour →
from scavengers to predators [400]
- Increased nesting sites [369]

AI

Breeding success
- Increasing decapods [397, 398] →
boosted food & CaCO₃ supply →
increased breeding success [89, 182]
- Increased decapods predation →
decreased fish stock resources [58, 59]

Temperature-driven
ecosystem shift [59, 397]
- Greater abundance & range
warm-water decapods [58, 397]
- Warming →
intensified predation on bivalves
& fish recruits [58, 59, 397]

ENV

- Faster development →
quicker reproductive cycle →
greater & earlier recruitment [397, 398]
- Ecosystem shift → change composition
& abundance holozooplankton [59] →
expansion decapod larvae population →
increased predation, feeding on
phyto/merozooplankton [59, 397]

ENV

Overfishing [399, 401]
- Decreased predation pressure
meroplanktonic larvae [58] →
grater range & abundance
decapods larvae [59, 397]

AI

Warming
- Increased metabolism →
energy demand →
intensified predation [221]
- Faster development →
quicker reproductive cycle [231]
- Expanding predator populations [89]

- Impaired larval navigation [231]
- Impaired movement,
hunting behaviour [19]

ENV

TBT contamination
- Imposex [402]

AI

Shifting surface regimes
- Increased temperature →
boosted metabolisms & feeding [127]
- Dynamic pH & salinity regimes,
decreased oxygen →
higher energetic demand [396],
dissolving conditions [99]
- Carbon supply →
increased phytoplankton abundance
& composition shifts [52,394] →
increased primary production
[107,233]

Urbanisation, land use
- Increased inorganic nutrients &
organic carbon river loads [394] →
coastal eutrophication [52, 394] →
phytoplankton community
increase & shift →
boosted primary production [52]

Decapoda (adults)

Increased recruitment [397]

Decapoda (larvae)

Boosted food quantity & quality [89]

Increased predation pressure [89]

Increased predation [59]

Increased predation pressure

Increased predation larval stages [59]

Boosted food supply

Dampened food supply

Seagulls (*Larus* spp.)

Increased abundance & prey-switching [400] → increased predation

Mytilus edulis (blue mussel)

Boosted food supply

Dampened food supply

Phytoplankton

Potential target of prey-switching behaviour

Increased predation pressure [368]

Control on carbon cycling and sequestration [52]

Carbon supply

Boosted nutrient & carbon [52]

Nucella lapillus (dog whelk)

Local extinction (1981) [390]

Eutrophication buffers ocean acidification [32]

ENV

AI

Figure 6.13 Belgian mussel ecosystem interactions
Conceptual representation of end-to-end ecosystem interactions (considered in the current study) governing blue mussels beds along the Belgian coastal zone, including primary producers, and primary, secondary, and tertiary consumers from both marine (sublittoral and eulittoral) and terrestrial systems. The graph represents top-down and bottom-up pathways (blue arrows) through which abiotic and biotic interactions within the *M. edulis* ecological network can stimulate compensatory responses of shell biomineralisation. Signs indicate positive (+) or negative (-) responses. Arrows indicated direct (continuous lines) and indirect (dashed lines) effects. Red arrows indicate external environmental (ENV) and anthropogenic (AI) impacts. Compensatory adjustments of blue mussel shells represent the shifting balance between responses to propagating anthropogenic impacts through the durophagous linkages (increasing decapod larvae → adult decapods → seagulls), altered seabird predation behaviour (prey-switching in seagulls), the disappearance of drillers (dog whelk), and more energetically demanding environmental regimes (increasing temperature, decreasing dissolved oxygen, and fluctuating pH and salinity) in an increasingly productive coastal system (boosted Chl-*a*); and their chain of direct and indirect feedbacks strengthened or weakened by climatic- and anthropogenic-driven stressors.

6.4.4 Anthropogenic triggers of shell compensation

Over the last century, human impact represented a major driver of Belgian coastal dynamics with an end-to-end ecosystem effect (i.e. from the physical environment to tertiary producers), acting through different direct and indirect pathways on the various trophic levels and species interactions governing the blue mussels' ecological network (Figure 6.13). A relatively limited influence of abiotic regimes (temperature and salinity, and collinear oxygen levels Figure 6.8a) on temporal patterns of shell deposition ($0.34 < cR^2 < 0.49$) suggests anthropogenic-altered local predator-prey dynamics [174] in a highly productive system [32] to trigger compensatory responses in shell biomineralisation to physiological stress.

Mytilus edulis fitness is tightly linked to shell integrity and the potential to regulate its deposition and composition in selecting environments [83, 99, 109]. Given the marked differences in energetic costs of $CaCO_3$ vs organics production [44, 407], mechanical and chemical protection of shell layers [83], as well as increased energy demand and predation [369, 397], mussels would be under strong selection pressure to modify their shell to counter disturbance.

The historical increase of calcareous layers deposition (Figure 6.6a) is in line with quantitative and qualitative temporal changes of predation pressure [198]. Prismatic thickness has been documented as the factor mostly affecting *M. edulis* vulnerability to seabird predation [200]. This suggests the rapid prismatic thickening after 1995 likely reflects a compensatory response to the exponential growth of seagull colonies (starting in 1990 - 1995, Figure 6.8b). Similarly, the rapid periostracal thinning after 1984 and a lower proportion of organic matrix in modern shells (increasing solubility) (Figure 6.6a, b), suggest a compensatory response to the disappearance of *N. lapillus* in 1981 (Figure 6.8b), selecting for

shells with high chemical resistance (i.e. thick periostracum and abundant organic matrix) [179], and to buffered ocean acidification effects [32] (Figure 6.13). Although the limited availability of environmental datasets does not allow explanation of the older shell patterns, these findings support the potential of mussels to trade-off between the deposition of organic and calcareous shell layers as a compensatory response to strategy-specific predation pressures.

These results highlight how complex, local-scale conditions, triggering compensatory mechanisms, can led to unanticipated outcomes (i.e. increase calcification and variable surface regimes) contrary to global-scale projections. In this regional scenario, by abruptly removing a keystone predator, selecting for energetically "expensive" responses and boosting food supply *via* eutrophication, local anthropogenic pressure may have shifted the mussel energy balance in favour of the formation of "cheaper" calcareous-dominated shells to compensate increased durophagous predation.

6.5 Conclusions

My findings suggest a strong capacity of *Mytilus* to withstand a wide range of environmental perturbations through plastic compensatory responses in shell biomineralisation. This naturally complex historical scenario supports the hypothesis that compensatory processes acting through qualitative and quantitative change in shell formation can counter predicted negative ramifications of environmental change on mussel persistence. Increasing evidences support the potential for resistance in calcifiers [69, 106, 110], suggesting that compensatory mechanisms of shell biomineralisation to disturbance may not be unique to mussels and potentially act in a range of other organisms, with their net outcomes being driven by species-specific heterogeneity in skeletal structure and physiological plasticity. This highlights the adaptive value of compensatory mechanisms to maintain shell calcification, and so fitness, in rapidly changing environments *via* formation of shell structures reflecting the equilibrium between co-acting climatic and anthropogenic selective forces, and the energetic costs of compensation.

Compensatory biomineralisation potentially represents an important, but relatively unexplored, mechanisms of resistance to change that complements documented physiological, population and community level adjustments [23, 40, 43, 53]. A better understanding of type and magnitude of compensatory species' responses is critical to inform more realistic projections of calcifiers' resistance to rapidly changing environments. In this regard, historical studies and the use of archival records represent a rarely used but powerful approach to complement inferential limitations of experimental ecology, providing a more integrative assessment of the predominant direct and indirect pathways of ecological change or stability.

Chapter Seven

7 Conclusions

7.1 Overview

Environmental change represents a major force shaping marine ecosystems worldwide [6, 26]. The complex interplay of a wide range of physical and chemical changes, including warming, variations of salinity, deoxygenation, and acidification [5, 12], in addition to pervasive anthropogenic pressures [352], are profoundly affecting species biology and ecology [18, 21], as well as community structures and ecosystem functioning [25, 56]. However, the rapid evolution of abiotic and biotic environmental conditions is outpacing our ability to forecast their consequences on marine populations, communities, and ecosystems [38]. Indeed, ecosystem-wide projection are severely constrained by heterogeneous patterns of change [6], multiple interacting stressors [39], and altered species interactions [56, 354], as well as limited insights into compensatory mechanisms driving responses at species or community levels [54, 65].

In the marine environment, species producing shell and skeletons, among them shelled molluscs, are experiencing the strongest impacts of environmental change [21, 72, 105]. Studies on their responses to disturbance suggest that calcifiers will find it more difficult to build and maintain their $CaCO_3$ structures, and they will potentially suffer a range of negative impacts including changes in metabolism, acid-base status, and survival under future ocean conditions [72, 74, 84, 101].

Among molluscs, bivalves constitute a substantial component of coastal benthic communities [115], and these include mussels which have been widely used as model organisms for studying ecological and physiological responses to disturbance [66, 110, 126]. The Atlantic mussels of the *Mytilus edulis* species-complex (*M. edulis*, *M. trossulus*, and *M. galloprovincialis*) are important bed-forming foundation species throughout the temperate and polar littoral zones worldwide. These have a significant ecological value supporting benthic communities and local biodiversity [121], and represent an important economic resource for the aquaculture industry [8]. As both calcifiers and habitat-formers, they have received much attention as primary indicators organisms [43, 108].

Calcareous shells perform a range of vital functions in *Mytilus*, including structural support and protection against predators. Because shell integrity determines survival, shell characteristics (i.e. form, thickness, and composition) are subject to strong selection pressure, with functional success or failure

a fundamental evolutionary driver. *Mytilus* shell consists of both organic and calcareous layers (Section 1.2.10). These layers are characterised by different energetic costs [44, 407], microstructures, and more or less soluble components the combination of which determines differential chemical and mechanical shell protection against dissolving conditions, predatory and endolithic borers, as well as durophagous predation [83, 99, 185]. Therefore, future shifts in environmental gradients [6] may influence variations in shell production, composition, and structure, shaping regional patterns of mussel sensitivity and survival to future conditions. However, existing knowledge on calcifiers' response potential predominantly stems from short- to long-term studies on single species or simplified experimental "communities", which may not necessarily translate to natural populations within a dynamic ecological system [69, 107]. Therefore, little is known about the biological mechanisms and key drivers in natural environments that shape regional differences and long-term variations in species vulnerability to global changes.

In this thesis, I examined variations in shell characteristics, including morphology and biomineralised microstructures, in natural populations of *Mytilus* across a large geographical scale (30° latitudinal range) along the North Atlantic coastline, and over historical times (last century). The aim was to understand the response potential of mussel shells to heterogeneous environmental gradients and local conditions. I found a marked response potential of shell morphology and biomineralisation over space and time; *Mytilus* species showed a wide range of plastic responses in shell deposition and composition to variations in abiotic and biotic environmental factors at a North Atlantic scale and over the last century. These findings indicate the marked potential of mussels to trade-off between **i**) different morphological traits, and **ii**) the deposition of individual calcareous and organic shell layers as a response to specific environmental conditions and variations in strategy-specific predation pressures. This suggests a strong capacity for an increased resistance of *Mytilus* to forecasted climate changes. Overall, this thesis illustrates how biological mechanisms and variations in local environmental conditions, driving plastic shell responses to the spatial and temporal structure of multiple abiotic and biotic stressors, defining geographic and historical patterns of unforeseen species resilience to global environmental changes.

7.1.1 Shell morphology across space

The use of geometric morphometrics and generalised additive mixed models allowed me to identify spatial patterns of natural shell shape variation with environmental gradients (i.e. temperature, salinity and food availability) at different geographical scales in North Atlantic and Arctic *Mytilus* populations. I showed how the combined use of new statistical methods provided some unprecedented methodological advantage in the analysis of shape variation, and how the employment of multiple study systems at various geographical scales made it possible to uncouple the contribution of development

and genetic status to shell morphology. Specifically, an elliptic Fourier analysis of shell outlines allowed an in-depth quantification of shell shape and the definition of new independent variables expressing shape variations at different regional scales. Meanwhile, additive modelling of shape variables from multiple levels of analysis (from small to large geographical scale) allowed the identification of general patterns as well as more local trends of natural shape variation and heterogeneity. This made it possible to describe general relationships between shell shape variation in Atlantic *Mytilus* and environmental drivers that were independent of developmental (age) and genetic (species) influence to mussel shape.

This new methodological approach allowed the identification of previously unrecognised patterns of mussel form and the variation in specific shell features at a much finer scale than possible previously. I found salinity had the strongest effect on the latitudinal patterns of *Mytilus* shape, producing shells that were more elongated, narrower and with more parallel dorsoventral margins at lower salinities; variations that could be explained by the maintenance principle of a physiologically favourable surface-area to volume ratio. Temperature and food supply, however, were the main predictors of mussel shape heterogeneity. Overall, *Mytilus* showed similar shape responses to less favourable conditions across the different geographical scales analysed. These results indicate shell shape variability is a potentially important adaptive component to environmental stressors and suggest potentially marked shape modifications in mussels under future sub-optimal environmental conditions.

These findings show shell shape plasticity represents a powerful indicator to understand the alterations of blue mussel communities in rapidly changing environments. Climatic changes are projected to affect all areas of the ocean, with different areas changing at different rates. Therefore, understanding the links between calcifiers' responses and evolving conditions is essential to forecast accurately spatial sensitivity patterns of this foundation species. My work illustrates that the use of new statistical tools combined with robust, more traditional, methods of analysis can improve our understanding of mussel ecology, by disentangling the different factors affecting shell shape and allowing more accurate predictions of its variations and their potential implications in a changing multivariate environment.

7.1.2 Shell biomineralisation across large geographical scales

The analysis of large-scale spatial variations in shell microstructure of *Mytilus* under heterogeneous environmental gradients across a 30° latitudinal range, allowed me to test whether plasticity in calcareous shell production and composition, from juveniles to large adults, mediates geographic patterns of resistance to climate change in this critical foundation species. As expected, I find mussels showed lower calcification levels in polar than temperate regions, which supports the general paradigm of decreasing skeletal production towards high latitudes. I found salinity was the major predictor of regional differences in mussel shell deposition, and its mineral and organic composition. In low-salinity

environments, the production of calcite and organic shell layers was increased, suggesting higher chemical resistance to corrosive water conditions. Conversely, under higher-salinity environments, increased aragonite deposition suggests enhanced mechanical protection. This marked compositional plasticity in calcareous shell components suggests a potentially higher resistance of *Mytilus* to dissolution in polar, low-salinity environments, than mussels in temperate, higher-salinity regions, which would be characterised by an enhanced protection from predators. Interacting strong effects of decreasing salinity and increasing food availability on the compositional shell plasticity in polar and subpolar mussels during growth predict the deposition of a thicker external organic layer (periostracum) under forecasted future conditions. The response potential of the periostracum in larger *Mytilus* to heterogeneous environments suggests a capacity for increased protection from ocean acidification in polar and sub-polar populations. Meanwhile a higher sensitivity of the Baltic populations is expected under future changes due to their reduced size. These results support the hypothesis that plasticity in shell production and the spatial structure of environmental conditions shape regional differences in *Mytilus* shells deposition and composition, and, therefore, their chemical and mechanical shell properties, mediating geographic patterns of resistance to environmental disturbance.

Overall, I demonstrated how biological mechanisms as shell plasticity, which drive spatial variability of mussel responses to interacting environmental factors, has the potential to determine complex regional patterns of shell deposition and properties, and to dictate spatial differences in species sensitivity to projected habitat alterations.

7.1.3 Shell characteristic over the last century

The use of museum collections of *Mytilus edulis* specimens sampled regularly from a single location along the Belgian coastline on stone breakwaters between 1904 and 2016, provided the unique opportunity to examine temporal trends in shell deposition, composition, and morphology in intertidal mussels over the last century. Availability of historical datasets of abiotic and biotic conditions made it possible to examine relationships between historical patterns of shell biomineralisation, and high-resolution, long-term changes in key environmental descriptors and abundance of keystone predators (different taxa and predation strategy), across a time-scale accounting for long-term acclimation and adaption. This allowed me to test the hypothesis that compensatory mechanisms in shell biomineralisation to natural and anthropogenic pressures can mediate the predicted negative effects of change and altered species interactions within a complexity ecological system.

Contrary to predictions calcification in *Mytilus* markedly increased during the last 112 years. This is in contrast with previous historical and experimental evidences documenting decreasing, or a general stability, in calcification for mussels or other calcifiers. Analysis of long-term datasets indicated the

formation of generally more variable surface conditions with formation of more energetically demanding coastal regimes for mussels. I also illustrated how the deposition of individual shell layers was more closely related to temporal changes in the variability of key environmental predictors than to alterations of mean habitat conditions. This suggests current field populations may have already made energetic adjustments or adapted to increased physiological stress in this region, with no negative effects on shell deposition. Moreover, I document an increasing predation pressure over time with a rapid shift to shell-breakers (durophagy) dominated regimes (i.e. decapods and seagulls) after the local extinction of a keystone predatory driller (i.e. dog whelk). Historical patterns of shell deposition and composition are in line with different level and types of predation. These findings support the potential of *Mytilus* to regulate the deposition of organic and calcareous shell layers as a compensatory response to strategy-specific predation pressures. This naturally complex historical scenario supports the hypothesis that compensatory processes acting through qualitative and quantitative change in shell formation can counter, or even reverse, predicted negative impacts of environmental change on mussels. My findings suggest **i**) a strong capacity of *M. edulis* to withstand a wide range of environmental perturbations through plastic compensatory responses in shell biomineralisation, and **ii**) how local-scale conditions, triggering compensatory mechanisms, can led to unanticipated outcomes which are in contrast with global-scale projections.

This work highlights how historical studies and the use of archival records represent a powerful tool to study very long-term changes is species responses, with potential for acclimation or adaption, to the evolution of local abiotic and biotic conditions. Furthermore, I illustrate the potential adaptive value of compensatory mechanisms, to maintain shell calcification, and so fitness, in rapidly changing environments *via* formation of shell structures reflecting the equilibrium between co-acting climatic and anthropogenic selective forces, and the energetic costs of compensation.

7.2 Future directions

In this thesis, I examined variations in shell morphology and biomineralisation in natural populations of *Mytilus* across a large geographical scale and over historical times to address the question on whether marine mussels have and will be affected by environmental change. A wide range of shell characteristics where investigated, indicating different spatial and temporal patterns of resistance to change in species predicted to be vulnerable, as well as differences between large and local scale abiotic and biotic drivers of compensatory shell responses in *Mytilus*.

Global change research is rapidly becoming an increasingly multidisciplinary field bringing together and trying to integrate numerous techniques and approaches to understand how populations of marine organisms will respond to the complex, multiple stressors scenario emerging in the contemporary world

ocean. Although I have presented spatial and temporal response patterns in a range of shell characteristics, there are other aspects of mussel biology and ecology that deserve further investigations to project the impacts of climate change on this calcifying foundation species. Biological processes such as metabolic rate, respiration and reproduction, as well as ecological factors such as interspecific and predator-prey interactions, in addition to genetics and epigenetics are key to fully understand the response potential of *Mytilus* to unprecedented future environmental alterations.

In Chapter 4, I illustrated how the uses of powerful statistical methods and multiple study systems on a wide geographical scale made it possible to identify previously unrecognised, fine-scale patterns of mussel form. These methods and approaches have the strong potential to be used with a wide range of other marine taxa beyond *Mytilus*. For example, these have been already employed by Vendrami *et al.* [408] with the commercially important king scallops, *Pecten maximus*, in combination with genetic investigations, including classical genetic markers (microsatellites), or modern high-throughput sequencing (RAD sequencing), for understanding fine-scale population structures and identifying which shell traits are phenotypically plastic. This suggests how RAD sequencing in combination with the morphometric approaches I implemented are powerful tools for studying phenotypic plasticity in natural populations. Furthermore, despite the identified wide range of responses, more studies are needed to understand the adaptive significance of the observed alterations and their underlying causes. Indeed, shell morphology represents an important aspect influencing bivalves' ecology, such as interspecific completions for space and predation vulnerability. Only an in depth understanding of these factors can make morphological variation a powerful tool for understanding the adaptation of organisms and help to predict their vulnerability in a rapidly changing environment.

Analysis of spatial and temporal patters in *Mytilus* biomineralisation indicated how shell deposition and composition are controlled by different set of drivers depending on the geographical and temporal scale analysed. Environmental factors were shown to be primary predictors of biomineralisation across large geographical scales, while type and level of predation pressure were the most influential local drivers. These results indicate how future experiments investigating the impact of predicted climate change on marine organisms should utilise environmental gradients and naturally occurring locally adapted populations. The shell, as well as calcified structures in other marine taxa, is a critical components of mussel sensitivity to change, and its alterations are key to understand sensitivity to a range of direct abiotic and biotic impacts, such as dissolving conditions and altered predation pressure. As global climate change accelerates, integrating data on population dynamics with information on shell alterations at a structural level (both qualitative and quantitative changes), as well as information of altered physiological status, is critical to better understand processes and mechanisms mediating species' responses to disturbance. This knowledge is essential to develop accurate predictions about species-specific responses and vulnerability to change, especially for those organisms having both high climate sensitivity and disproportionately strong ecological impacts in shaping marine communities.

Compensatory mechanisms have long been regarded as important stabilising mechanisms in natural systems. However, these have been investigated almost entirely in terms of diversity loss, changes in density, and trophic adjustments. Despite being straightforward to measure, these aspects only partially account for resistance at a community level. In fact, compensation in calcifiers can arise from a wide range of responses at population or species levels, acting through physiological adjustments, and potentially resulting in alterations of biomineralisation processes and, therefore, the final shell produced. However, quantitative and qualitative changes in shell deposition and composition, and their potential influence on inter- and intra-specific interactions have been generally overlooked. Compensatory mechanisms, such as trade-offs in biomineralisation, represents important mechanisms of resistance to change that complements documented physiological and community level adjustments. These processes deserve further investigation in both experimental and natural contexts, since an understanding of type and magnitude of compensatory responses is essential to inform realistic projections of calcifiers' resistance in a rapidly changing environment.

Large-scale analyses of natural populations and historical studies of archival records represent rarely used, but powerful, approaches to complement inferential limitations of experimental ecology, with the strong potential to provide a more integrative assessment of the predominant direct and indirect pathways of ecological change or stability. This understanding is of critical importance for making realistic projections of emergent ecological effects of environmental change at global and regional scales, and for improving our predictive accuracy for impacts on marine ecosystems, and the services they provide.

References

[1] IPCC, *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge University Press, Cambridge, UK, 2013).

[2] S. Doney, W. Balch, V. Fabry, R. Feely, Ocean acidification: a critical emerging problem for the ocean sciences. *Oceanography*. **22**, 16–25 (2009).

[3] S. C. Doney, D. S. Schimel, Carbon and climate system coupling on timescales from the Precambrian to the Anthropocene. *Annu. Rev. Environ. Resour.* **32**, 31–66 (2007).

[4] P. N. Pearson, M. R. Palmer, Atmospheric carbon dioxide concentrations over the past 60 million years. *Nature*. **406**, 695–699 (2000).

[5] B. Kirtman, S. B. Power, J. A. Adedoyin, G. J. Boer, R. Bojariu, I. Camilloni, F. J. Doblas-Reyes, A. M. Fiore, M. Kimoto, G. A. Meehl, M. Prather, A. Sarr, C. Schär, R. Sutton, G. J. van Oldenborgh, G. Vecchi, H. J. Wang, in *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, T. F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, P. M. Midgley, Eds. (Cambridge University Press, Cambridge, UK, 2013), pp. 953–1028.

[6] J. P. Gattuso, A. Magnan, R. Bille, W. W. L. Cheung, E. L. Howes, F. Joos, D. Allemand, L. Bopp, S. R. Cooley, C. M. Eakin, O. Hoegh-Guldberg, R. P. Kelly, H. O. Portner, A. D. Rogers, J. M. Baxter, D. Laffoley, D. Osborn, A. Rankovic, J. Rochette, U. R. Sumaila, S. Treyer, C. Turley, Contrasting futures for ocean and society from different anthropogenic CO_2 emissions scenarios. *Science*. **349**, aac4722 (2015).

[7] C. Mora, D. P. Tittensor, S. Adl, A. G. B. Simpson, B. Worm, How many species are there on earth and in the ocean? *PLoS Biol.* **9**, e1001127 (2011).

[8] FAO, *FAO Yearbook. Fishery and Aquaculture Statistics. 2015* (FAO, Rome, 2017).

[9] IPCC, *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge University Press, Cambridge, UK, 2014).

[10] P. J. Durack, S. E. Wijffels, P. J. Gleckler, Long-term sea-level change revisited: the role of salinity. *Environ. Res. Lett.* **9**, 114017 (2014).

[11] C. L. Sabine, R. A. Feely, N. Gruber, R. M. Key, K. Lee, J. L. Bullister, R. Wanninkhof, C. S. Wong, D. W. R. Wallace, B. Tilbrook, F. J. Millero, T.-H. Peng, A. Kozyr, T. Ono, A. F. Rios, The oceanic sink for anthropogenic CO_2. *Nature*. **305**, 367–371 (2004).

[12] K. Caldeira, M. E. Wickett, Anthropogenic carbon and ocean pH. *Nature*. **425**, 365–365 (2003).

[13] R. A. Feely, C. L. Sabine, K. Lee, W. Berelson, J. Kleypas, V. J. Fabry, F. J. Millero, Impact of anthropogenic CO_2 on the $CaCO_3$ system in the oceans. *Science*. **305**, 362–366 (2004).

[14] J. C. Orr, V. J. Fabry, O. Aumont, L. Bopp, S. C. Doney, R. A. Feely, A. Gnanadesikan, N. Gruber, A. Ishida, F. Joos, R. M. Key, K. Lindsay, E. Maier-Reimer, R. Matear, P. Monfray, A. Mouchet, R. G. Najjar, G.-K. Plattner, K. B. Rodgers, C. L. Sabine, J. L. Sarmiento, R. Schlitzer, R. D. Slater, I. J. Totterdell, M.-F. Weirig, Y. Yamanaka, A. Yool, Anthropogenic ocean acidification over the twenty-first century and its impact on calcifying organisms. *Nature*. **437**, 681–686 (2005).

[15] M. K. Sejr, C. A. Stedmon, J. Bendtsen, J. Abermann, T. Juul-Pedersen, J. Mortensen, S. Rysgaard, Evidence of local and regional freshening of Northeast Greenland coastal waters. *Sci. Rep.* **7**, 13183 (2017).

[16] K. K. Kjeldsen, N. J. Korsgaard, A. A. Bjørk, S. A. Khan, J. E. Box, S. Funder, N. K. Larsen, J. L. Bamber, W. Colgan, M. van den Broeke, M.-L. Siggaard-Andersen, C. Nuth, A. Schomacker, C. S. Andresen, E. Willerslev, K. H. Kjær, Spatial and temporal distribution of mass loss from the Greenland Ice Sheet since AD 1900. *Nature*. **528**, 396–400 (2015).

[17] P. Durack, Ocean salinity and the global water cycle. *Oceanography*. **28**, 20–31 (2015).

[18] P. Calosi, S. Melatunan, L. M. Turner, Y. Artioli, R. L. Davidson, J. J. Byrne, M. R. Viant, S. Widdicombe, S. D. Rundle, Regional adaptation defines sensitivity to future ocean acidification. *Nat. Commun.* **8**, 13994 (2017).

[19] S.-A. Watson, S. Lefevre, M. I. McCormick, P. Domenici, G. E. Nilsson, P. L. Munday, Marine mollusc predator-escape behaviour altered by near-future carbon dioxide levels. *Proc. R. Soc. B Biol. Sci.* **281**, 20132377 (2014).

[20] S. A. Morley, C. C. Suckling, M. S. Clark, E. L. Cross, L. S. Peck, Long-term effects of altered pH and temperature on the feeding energetics of the Antarctic sea urchin, *Sterechinus neumayeri*. *Biodiversity*. **17**, 1–2 (2016).

[21] K. J. Kroeker, R. L. Kordas, R. Crim, I. E. Hendriks, L. Ramajo, G. S. Singh, C. M. Duarte, J.-P. Gattuso, Impacts of ocean acidification on marine organisms: quantifying sensitivities and interaction with warming. *Glob. Change Biol.* **19**, 1884–1896 (2013).

[22] L. S. Peck, in *Oceanography and Marine Biology: an Annual Review*, S. J. Hawkins, A. J. Evans, A. C. Dal, L. B. Firth, I. P. Smith, Eds. (CRC Press, 2018), vol. 56, p. 510.

[23] L. S. Peck, M. S. Clark, D. Power, J. Reis, F. M. Batista, E. M. Harper, Acidification effects on biofouling communities: winners and losers. *Glob. Change Biol.* **21**, 1907–1913 (2015).

[24] G. V. Ashton, S. A. Morley, D. K. A. Barnes, M. S. Clark, L. S. Peck, Warming by 1°C drives species and assemblage level responses in Antarctica's marine shallows. *Curr. Biol.* **27**, 2698–2705.e3 (2017).

[25] J. M. Sunday, K. E. Fabricius, K. J. Kroeker, K. M. Anderson, N. E. Brown, J. P. Barry, S. D. Connell, S. Dupont, B. Gaylord, J. M. Hall-Spencer, T. Klinger, M. Milazzo, P. L. Munday, B.

D. Russell, E. Sanford, V. Thiyagarajan, M. L. H. Vaughan, S. Widdicombe, C. D. G. Harley, Ocean acidification can mediate biodiversity shifts by changing biogenic habitat. *Nat. Clim. Change*. **7**, 81–85 (2016).

[26]   I. Nagelkerken, S. D. Connell, Global alteration of ocean ecosystem functioning due to increasing human CO_2 emissions. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 13272–13277 (2015).

[27]   K. M. Strassmann, G.-K. Plattner, F. Joos, CO_2 and non-CO_2 radiative forcings in climate projections for twenty-first century mitigation scenarios. *Clim. Dyn.* **33**, 737–749 (2009).

[28]   K. Caldeira, Seawater pH and atmospheric carbon dioxide. *Science*. **286**, 2043a–2043 (1999).

[29]   R. A. Feely, C. L. Sabine, J. M. Hernandez-Ayon, D. Ianson, B. Hales, Evidence for upwelling of corrosive "acidified" water onto the continental shelf. *Science*. **320**, 1490–1492 (2008).

[30]   J. Salisbury, M. Green, C. Hunt, J. Campbell, Coastal acidification by rivers: a threat to shellfish? *Eos, Trans. Am. Geophys. Union*. **89**, 513–513 (2008).

[31]   W.-J. Cai, X. Hu, W.-J. Huang, M. C. Murrell, J. C. Lehrter, S. E. Lohrenz, W.-C. Chou, W. Zhai, J. T. Hollibaugh, Y. Wang, P. Zhao, X. Guo, K. Gundersen, M. Dai, G.-C. Gong, Acidification of subsurface coastal waters enhanced by eutrophication. *Nat. Geosci.* **4**, 766–770 (2011).

[32]   A. V. Borges, N. Gypens, Carbonate chemistry in the coastal zone responds more strongly to eutrophication than ocean acidification. *Limnol. Oceanogr.* **55**, 346–353 (2010).

[33]   P. J. Durack, T. Lee, N. T. Vinogradova, D. Stammer, Keeping the lights on for global ocean salinity observation. *Nat. Clim. Change*. **6**, 228–231 (2016).

[34]   P. J. Durack, S. E. Wijffels, R. J. Matear, Ocean salinities reveal strong global water cycle intensification during 1950 to 2000. *Science*. **336**, 455–458 (2012).

[35]   C. Chou, J. D. Neelin, C.-A. Chen, J.-Y. Tu, Evaluating the "rich-get-richer" mechanism in tropical precipitation change under global warming. *J. Clim.* **22**, 1982–2005 (2009).

[36]   P. W. Boyd, S. T. Lennartz, D. M. Glover, S. C. Doney, Biological ramifications of climate-change-mediated oceanic multi-stressors. *Nat. Clim. Change*. **5**, 71–79 (2015).

[37]   U. Riebesell, J.-P. Gattuso, Lessons learned from ocean acidification research. *Nat. Clim. Change*. **5**, 12–14 (2015).

[38]   K. J. Kroeker, R. L. Kordas, C. D. G. Harley, Embracing interactions in ocean acidification research: confronting multiple stressor scenarios and context dependence. *Biol. Lett.* **13**, 20160802 (2017).

[39]   D. L. Breitburg, J. Salisbury, J. Bernhard, W.-J. Cai, S. Dupont, S. Doney, K. Kroeker, L. Levin, W. C. Long, L. Milke, S. Miller, B. Phelan, U. Passow, B. Seibel, A. Todgham, A. Tarrant, And on top of all that… Coping with ocean acidification in the midst of many stressors. *Oceanography*. **28**, 48–61 (2015).

[40]   C. C. Suckling, M. S. Clark, C. Beveridge, L. Brunner, A. D. Hughes, E. M. Harper, E. J. Cook, A. J. Davies, L. S. Peck, Experimental influence of pH on the early life-stages of sea urchins II:

increasing parental exposure times gives rise to different responses. *Invertebr. Reprod. Dev.* **58**, 161–175 (2014).

[41] E. L. Cross, L. S. Peck, M. D. Lamare, E. M. Harper, No ocean acidification effects on shell growth and repair in the New Zealand brachiopod *Calloria inconspicua* (Sowerby, 1846). *ICES J. Mar. Sci.* **73**, 920–926 (2016).

[42] C. L. Mackenzie, G. A. Ormondroyd, S. F. Curling, R. J. Ball, N. M. Whiteley, S. K. Malham, Ocean warming, more than acidification, reduces shell strength in a commercial shellfish species during food limitation. *PLoS One*. **9**, e86764 (2014).

[43] K. J. Kroeker, E. Sanford, J. M. Rose, C. A. Blanchette, F. Chan, F. P. Chavez, B. Gaylord, B. Helmuth, T. M. Hill, G. E. Hofmann, M. A. McManus, B. A. Menge, K. J. Nielsen, P. T. Raimondi, A. D. Russell, L. Washburn, Interacting environmental mosaics drive geographic variation in mussel performance and predation vulnerability. *Ecol. Lett.* **19**, 771–779 (2016).

[44] S.-A. Watson, S. A. Morley, L. S. Peck, Latitudinal trends in shell production cost from the tropics to the poles. *Sci. Adv.* **3**, e1701362 (2017).

[45] T. Wernberg, D. A. Smale, M. S. Thomsen, A decade of climate change experiments on marine organisms: procedures, patterns and problems. *Glob. Change Biol.* **18**, 1491–1498 (2012).

[46] C. A. Vargas, N. A. Lagos, M. A. Lardies, C. Duarte, P. H. Manríquez, V. M. Aguilera, B. Broitman, S. Widdicombe, S. Dupont, Species-specific responses to ocean acidification should account for local adaptation and adaptive plasticity. *Nat. Ecol. Evol.* **1**, 0084 (2017).

[47] P. W. Boyd, S. Collins, S. Dupont, K. Fabricius, J.-P. Gattuso, J. Havenhand, D. A. Hutchins, U. Riebesell, M. S. Rintoul, M. Vichi, H. Biswas, A. Ciotti, K. Gao, M. Gehlen, C. L. Hurd, H. Kurihara, C. M. McGraw, J. M. Navarro, G. E. Nilsson, U. Passow, H.-O. Pörtner, Experimental strategies to assess the biological ramifications of multiple drivers of global ocean change - A review. *Glob. Change Biol.* **24**, 2239–2261 (2018).

[48] C. M. Crain, K. Kroeker, B. S. Halpern, Interactive and cumulative effects of multiple human stressors in marine systems. *Ecol. Lett.* **11**, 1304–1315 (2008).

[49] B. B. Hughes, S. C. Lummis, S. C. Anderson, K. J. Kroeker, Unexpected resilience of a seagrass system exposed to global stressors. *Glob. Change Biol.* **24**, 224–234 (2018).

[50] R. K. Bechmann, I. C. Taban, S. Westerlund, B. F. Godal, M. Arnberg, S. Vingen, A. Ingvarsdottir, T. Baussant, Effects of ocean acidification on early life stages of shrimp (*Pandalus borealis*) and mussel (*Mytilus edulis*). *J. Toxicol. Environ. Heal. Part A*. **74**, 424–438 (2011).

[51] J. Thomsen, K. Ramesh, T. Sanders, M. Bleich, F. Melzner, Calcification in a marginal sea – influence of seawater $[Ca^{2+}]$ and carbonate chemistry on bivalve shell formation. *Biogeosciences*. **15**, 1469–1482 (2018).

[52] N. Gypens, A. V. Borges, C. Llancelot, Effect of eutrophication on air-sea CO_2 fluxes in the coastal Southern North Sea: a model study of the past 50 years. *Glob. Change Biol.* **15**, 1040–1056 (2009).

[53]   J. Thomsen, I. Casties, C. Pansch, A. Körtzinger, F. Melzner, Food availability outweighs ocean acidification effects in juvenile *Mytilus edulis*: laboratory and field experiments. *Glob. Change Biol.* **19**, 1017–1027 (2013).

[54]   G. Ghedini, S. D. Connell, Organismal homeostasis buffers the effects of abiotic change on community dynamics. *Ecology.* **97**, 2671–2679 (2016).

[55]   G. Ghedini, S. D. Connell, Moving ocean acidification research beyond a simple science: Investigating ecological change and their stabilizers. *Food Webs.* **13**, 53–59 (2017).

[56]   K. J. Kroeker, F. Micheli, M. C. Gambi, Ocean acidification causes ecosystem shifts via altered competitive interactions. *Nat. Clim. Change.* **3**, 156–159 (2012).

[57]   R. I. Perry, R. E. Ommer, M. Barange, F. Werner, The challenge of adapting marine social–ecological systems to the additional stress of climate change. *Curr. Opin. Environ. Sustain.* **2**, 356–363 (2010).

[58]   R. R. Kirby, G. Beaugrand, J. A. Lindley, Synergistic effects of climate and fishing in a marine ecosystem. *Ecosystems.* **12**, 548–561 (2009).

[59]   R. R. Kirby, G. Beaugrand, Trophic amplification of climate warming. *Proceedings. Biol. Sci.* **276**, 4095–4103 (2009).

[60]   T. Wernberg, S. Bennett, R. C. Babcock, T. de Bettignies, K. Cure, M. Depczynski, F. Dufois, J. Fromont, C. J. Fulton, R. K. Hovey, E. S. Harvey, T. H. Holmes, G. A. Kendrick, B. Radford, J. Santana-Garcon, B. J. Saunders, D. A. Smale, M. S. Thomsen, C. A. Tuckett, F. Tuya, M. A. Vanderklift, S. Wilson, Climate-driven regime shift of a temperate marine ecosystem. *Science.* **353**, 169–72 (2016).

[61]   C. D. G. Harley, Climate change, keystone predation, and biodiversity loss. *Science.* **334**, 1124–1127 (2011).

[62]   J. Thyrring, M. Blicher, J. Sørensen, S. Wegeberg, M. Sejr, Rising air temperatures will increase intertidal mussel abundance in the Arctic. *Mar. Ecol. Prog. Ser.* **584**, 91–104 (2017).

[63]   A. Andersson, D. Kline, P. Edmunds, S. Archer, N. Bednaršek, R. Carpenter, M. Chadsey, P. Goldstein, A. Grottoli, T. Hurst, A. King, J. Kübler, I. Kuffner, K. Mackey, B. Menge, A. Paytan, U. Riebesell, A. Schnetzer, M. Warner, R. Zimmerman, Understanding ocean acidification impacts on organismal to ecological scales. *Oceanography.* **25**, 16–27 (2015).

[64]   L. S. Peck, S. A. Morley, J. Richard, M. S. Clark, Acclimation and thermal tolerance in Antarctic marine ectotherms. *J. Exp. Biol.* **217**, 16–22 (2014).

[65]   G. Ghedini, B. D. Russell, S. D. Connell, Trophic compensation reinforces resistance: herbivory absorbs the increasing effects of multiple disturbances. *Ecol. Lett.* **18**, 182–187 (2015).

[66]   F. Melzner, P. Stange, K. Trübenbach, J. Thomsen, I. Casties, U. Panknin, S. N. Gorb, M. a Gutowska, Food supply and seawater pCO_2 impact calcification and internal shell dissolution in the blue mussel *Mytilus edulis*. *PLoS One.* **6**, e24223 (2011).

[67]   I. Nagelkerken, P. L. Munday, Animal behaviour shapes the ecological effects of ocean

acidification and warming: moving from individual to community-level responses. *Glob. Change Biol.* **22**, 974–989 (2016).

[68]   A. M. Lister, Natural history collections as sources of long-term datasets. *Trends Ecol. Evol.* **26**, 153–154 (2011).

[69]   E. L. Cross, E. M. Harper, L. S. Peck, A 120-year record of resilience to environmental change in brachiopods. *Glob. Change Biol.* **24**, 2262–2271 (2018).

[70]   C. A. Pfister, K. Roy, J. T. Wootton, S. J. McCoy, R. T. Paine, T. H. Suchanek, E. Sanford, Historical baselines and the future of shell calcification for a foundation species in a changing ocean. *Proc. R. Soc. B Biol. Sci.* **283**, 20160392 (2016).

[71]   S. J. McCoy, N. A. Kamenos, P. Chung, T. J. Wootton, C. A. Pfister, A mineralogical record of ocean change: decadal and centennial patterns in the California mussel. *Glob. Change Biol.* **24**, 2554–2562 (2018).

[72]   K. J. Kroeker, R. L. Kordas, R. N. Crim, G. G. Singh, Meta-analysis reveals negative yet variable effects of ocean acidification on marine organisms. *Ecol. Lett.* **13**, 1419–1434 (2010).

[73]   S. J. McCoy, F. Ragazzola, Skeletal trade-offs in coralline algae in response to ocean acidification. *Nat. Clim. Change*. **4**, 719–723 (2014).

[74]   F. Gazeau, L. M. Parker, S. Comeau, J.-P. Gattuso, W. A. O'Connor, S. Martin, H.-O. Pörtner, P. M. Ross, Impacts of ocean acidification on marine shelled molluscs. *Mar. Biol.* **160**, 2207–2245 (2013).

[75]   J. L. Gutierrez, C. G. Jones, D. L. Strayer, O. O. Iribarne, Mollusks as ecosystem engineers: the role of shell production in aquatic habitats. *Oikos*. **101**, 79–90 (2003).

[76]   S. Comeau, R. Jeffree, J.-L. Teyssié, J.-P. Gattuso, Response of the arctic pteropod *Limacina helicina* to projected future environmental conditions. *PLoS One*. **5**, e11362 (2010).

[77]   S. V Smith, D. W. Kinsey, Calcium carbonate production, coral reef growth, and sea level change. *Science*. **194**, 937–9 (1976).

[78]   A. Ventura, S. Schulz, S. Dupont, Maintained larval growth in mussel larvae exposed to acidified under-saturated seawater. *Sci. Rep.* **6**, 23728 (2016).

[79]   S. Dupont, H. Pörtner, Get ready for ocean acidification. *Nature*. **498**, 429–429 (2013).

[80]   G. De'ath, J. M. Lough, K. E. Fabricius, Declining coral calcification on the Great Barrier Reef. *Science*. **323**, 116–119 (2009).

[81]   J. B. Ries, A. L. Cohen, D. C. McCorkle, Marine calcifiers exhibit mixed responses to CO_2-induced ocean acidification. *Geology*. **37**, 1131–1134 (2009).

[82]   J. Thomsen, F. Melzner, Moderate seawater acidification does not elicit long-term metabolic depression in the blue mussel *Mytilus edulis*. *Mar. Biol.* **157**, 2667–2676 (2010).

[83]   S. C. Fitzer, W. Zhu, K. E. Tanner, V. R. Phoenix, N. A. Kamenos, M. Cusack, Ocean acidification alters the material properties of *Mytilus edulis* shells. *J. R. Soc. Interface*. **12**, 20141227 (2015).

[84]   J. Thomsen, K. Haynert, K. M. Wegner, F. Melzner, Impact of seawater carbonate chemistry on the calcification of marine bivalves. *Biogeosciences*. **12**, 4209–4220 (2015).

[85]   L. Georgiou, J. Falter, J. Trotter, D. I. Kline, M. Holcomb, S. G. Dove, O. Hoegh-Guldberg, M. McCulloch, pH homeostasis during coral calcification in a free ocean CO_2 enrichment (FOCE) experiment, Heron Island reef flat, Great Barrier Reef. *Proc. Natl. Acad. Sci.* **112**, 13219–13224 (2015).

[86]   J. A. Berge, B. Bjerkeng, O. Pettersen, M. T. Schaanning, S. Øxnevad, Effects of increased sea water concentrations of CO_2 on growth of the bivalve *Mytilus edulis* L. *Chemosphere*. **62**, 681–687 (2006).

[87]   J. M. Sunday, A. E. Bates, N. K. Dulvy, Thermal tolerance and the global redistribution of animals. *Nat. Clim. Change*. **2**, 686–690 (2012).

[88]   L. E. Chambers, R. Altwegg, C. Barbraud, P. Barnard, L. J. Beaumont, R. J. M. Crawford, J. M. Durant, L. Hughes, M. R. Keatley, M. Low, P. C. Morellato, E. S. Poloczanska, V. Ruoppolo, R. E. T. Vanstreels, E. J. Woehler, A. C. Wolfaardt, Phenological changes in the southern hemisphere. *PLoS One*. **8**, e75514 (2013).

[89]   C. Luczak, G. Beaugrand, J. A. Lindley, J.-M. Dewarumez, P. J. Dubois, R. R. Kirby, North Sea ecosystem change from swimming crabs to seagulls. *Biol. Lett.* **8**, 821–824 (2012).

[90]   O. Hoegh-Guldberg, J. F. Bruno, The impact of climate change on the world's marine ecosystems. *Science*. **328**, 1523–8 (2010).

[91]   H.-O. Pörtner, D. M. Karl, P. W. Boyd, W. W. L. Cheung, S. E. Lluch-Cota, Y. Nojiri, D. N. Schmidt, P. O. Zavialov, in *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, C. B. Field, V. R. Barros, D. J. Dokken, K. J. Mach, M. D. Mastrandrea, T. E. Bilir, M. Chatterjee, K. L. Ebi, Y. O. Estrada, R. C. Genova, B. Girma, E. S. Kissel, A. N. Levy, S. MacCracken, P. R. Mastrandrea, L. L. White, Eds. (Cambridge University Press, Cambridge, UK, 2014), pp. 411–484.

[92]   M. L. Zippay, B. Helmuth, Effects of temperature change on mussel, *Mytilus*. *Integr. Zool.* **7**, 312–327 (2012).

[93]   S. Bennett, T. Wernberg, B. Arackal Joy, T. de Bettignies, A. H. Campbell, Central and rear-edge populations can be equally vulnerable to warming. *Nat. Commun.* **6**, 10280 (2015).

[94]   N. A. J. Graham, S. Jennings, M. A. MacNeil, D. Mouillot, S. K. Wilson, Predicting climate-driven regime shifts versus rebound potential in coral reefs. *Nature*. **518**, 94–97 (2015).

[95]   B. C. C. Hume, C. D'Angelo, E. G. Smith, J. R. Stevens, J. Burt, J. Wiedenmann, *Symbiodinium thermophilum* sp. nov., a thermotolerant symbiotic alga prevalent in corals of the world's hottest sea, the Persian/Arabian Gulf. *Sci. Rep.* **5**, 8562 (2015).

[96]   S. R. Palumbi, D. J. Barshis, N. Traylor-Knowles, R. A. Bay, Mechanisms of reef coral resistance to future climate change. *Science*. **344**, 895–898 (2014).

[97] M. Schweinsberg, L. C. Weiss, S. Striewski, R. Tollrian, K. P. Lampert, More than one genotype: how common is intracolonial genetic variability in scleractinian corals? *Mol. Ecol.* **24**, 2673–2685 (2015).

[98] C. J. Gobler, E. L. De Pasquale, A. W. Griffith, H. Baumann, Hypoxia and acidification have additive and synergistic negative effects on the growth, survival, and metamorphosis of early llfe stage bivalves. *PLoS One.* **9**, e83648 (2014).

[99] S. C. Fitzer, L. Vittert, A. Bowman, N. A. Kamenos, V. R. Phoenix, M. Cusack, Ocean acidification and temperature increase impact mussel shell shape and thickness: problematic for protection? *Ecol. Evol.* **5**, 4875–4884 (2015).

[100] F. Gazeau, C. Quiblier, J. M. Jansen, J.-P. Gattuso, J. J. Middelburg, C. H. R. Heip, Impact of elevated CO_2 on shellfish calcification. *Geophys. Res. Lett.* **34**, L07603 (2007).

[101] L. Parker, P. Ross, W. O'Connor, H. Pörtner, E. Scanes, J. Wright, Predicting the response of molluscs to the impact of ocean acidification. *Biology.* **2**, 651–692 (2013).

[102] S.-A. Watson, L. S. Peck, P. A. Tyler, P. C. Southgate, K. S. Tan, R. W. Day, S. A. Morley, Marine invertebrate skeleton size varies with latitude, temperature and carbonate saturation: implications for global change and ocean acidification. *Glob. Change Biol.* **18**, 3026–3038 (2012).

[103] V. J. Fabry, B. A. Seibel, R. A. Feely, J. C. Orr, Impacts of ocean acidification on marine fauna and ecosystem processes. *ICES J. Mar. Sci.* **65**, 414–432 (2008).

[104] J. P. Gattuso, L. Hansson, *Ocean Acidification* (Oxford University Press, Oxford, UK, 2011).

[105] F. Gazeau, L. M. Parker, S. Comeau, J.-P. Gattuso, W. A. O'Connor, S. Martin, H.-O. Pörtner, P. M. Ross, Impacts of ocean acidification on marine shelled molluscs. *Mar. Biol.* **160**, 2207–2245 (2013).

[106] J. Y. S. Leung, B. D. Russell, S. D. Connell, Mineralogical plasticity acts as a compensatory mechanism to the impacts of ocean acidification. *Environ. Sci. Technol.* **51**, 2652–2659 (2017).

[107] S. D. Connell, Z. A. Doubleday, S. B. Hamlyn, N. R. Foster, C. D. G. Harley, B. Helmuth, B. P. Kelaher, I. Nagelkerken, G. Sarà, B. D. Russell, How ocean acidification can benefit calcifiers. *Curr. Biol.* **27**, R95–R96 (2017).

[108] L. Telesca, K. Michalek, T. Sanders, L. S. Peck, J. Thyrring, E. M. Harper, Blue mussel shell shape plasticity and natural environments: a quantitative approach. *Sci. Rep.* **8**, 2865 (2018).

[109] L. Telesca, L. S. Peck, T. Sanders, J. Thyrring, M. K. Sejr, E. M. Harper, Plasticity and environmental heterogeneity predict geographic resilience patterns of foundation species to future change. **In review** (2018).

[110] J. Thomsen, L. S. Stapp, K. Haynert, H. Schade, M. Danelli, G. Lannig, K. M. Wegner, F. Melzner, Naturally acidified habitat selects for ocean acidification – tolerant mussels. *Sci. Adv.* **3**, e1602411 (2017).

[111] G. Rosenberg, A new critical estimate of named species-level diversity of the recent Mollusca.

Am. Malacol. Bull. **32**, 308–322 (2014).

[112] P. D. Taylor, D. N. Lewis, *Fossil Invertebrates* (Harvard University Press, Cambridge, MS, USA, 2005).

[113] W. F. Ponder, D. R. Lindberg, *Phylogeny and Evolution of the Mollusca* (University of California Press, Berkeley, CA, USA, 2008).

[114] E. E. Ruppert, R. S. Fox, R. D. Barnes, *Invertebrate Zoology : A Functional Evolutionary Approach* (Thomson-Brooks/Cole, Pacific Grove, CA, USA, 2004).

[115] E. Gosling, *Marine Bivalve Molluscs* (John Wiley & Sons, Ltd, Chichester, UK, 2015).

[116] E. M. Gosling, *Bivalve Molluscs: Biology, Ecology, and Culture* (Wiley-Blackwell, New Jersey, US, 2003).

[117] T. H. Suchanek, Extreme biodiversity in the marine environment: mussel bed communities of *Mytilus californianus*. *Northwest Environ. J.* **8**, 150–152 (1992).

[118] B. Morton, in *Limnoperna Fortunei - The Ecology, Distribution and Control of a Swiftly Spreading Invasive Fouling Mussel* (Springer International Publishing, Cham, 2015), pp. 3–41.

[119] S. M. Stanley, Functional morphology and evolution of byssally attached bivalve mollusks. *J. Paleontol.* **46**, 165–212 (1972).

[120] C. Richardson, R. Seed, E. Naylor, Use of internal growth bands for measuring individual and population growth rates in *Mytilus edulis* from offshore production platforms. *Mar. Ecol. Prog. Ser.* **66**, 259–265 (1990).

[121] R. Seed, T. H. Suchanek, in *The Mussel Mytilus: Ecology, Physiology, Genetics and Culture*, E. M. Gosling, Ed. (Elsevier, Amsterdam, 1992), pp. 87–170.

[122] T. H. Suchanek, in *The Ecology of Rocky Shores*, P. G. Moore, R. Seed, Eds. (Hodder & Stoughton, Sevenoaks, 1985), pp. 70–96.

[123] J. P. A. Gardner, The *Mytilus edulis* species complex in southwest England: effects of hybridization and introgression upon interlocus associations and morphometric variation. *Mar. Biol.* **125**, 385–399 (1996).

[124] N. D. Newell, *Classification of the Bivalvia* (New York, N.Y. : American Museum of Natural History, 1965).

[125] J. G. Carter, C. R. Altaba, L. C. Anderson, R. Araujo, A. S. Biakov, A. E. Bogan, D. Campbell, M. Campbell, J. Chen, J. C. W. Cope, G. Delvene, H. H. Dijkstra, Z. Fang, R. N. Gardner, V. A. Gavrilova, I. A. Goncharova, P. J. Harries, J. H. Hartman, M. Hautmann, W. R. Hoeh, J. Hylleberg, B. Jiang, P. Johnston, L. Kirkendale, K. Kleemann, J. Koppka, J. Kříž, D. Machado, N. Malchus, A. Márquez-Aliaga, J.-P. Masse, C. A. McRoberts, P. U. Middelfart, S. Mitchell, L. A. Nevesskaja, S. Özer, J. J. Pojeta, I. V. Polubotko, J. M. Pons, S. Popov, T. Sánchez, A. F. Sartori, R. W. Scott, I. I. Sey, J. H. Signorelli, V. V. Silantiev, P. W. Skelton, T. Steuber, J. B. Waterhouse, G. L. Wingard, T. Yancey, A synoptical classification of the Bivalvia (Mollusca). *Paleontol. Contrib.* **4**, 1–47 (2011).

[126] B. L. Bayne, *Marine Mussels: Their Ecology and Physiology* (Cambridge University Press, Cambridge, 1976).

[127] J. Widdows, B. L. Bayne, Temperature acclimation of *Mytilus edulis* with reference to its energy budget. *J. Mar. Biol. Assoc. UK.* **51**, 827–843 (1971).

[128] H. Stuckas, K. Stoof, H. Quesada, R. Tiedemann, Evolutionary implications of discordant clines across the Baltic *Mytilus* hybrid zone (*Mytilus edulis* and *Mytilus trossulus*). *Heredity.* **103**, 146–156 (2009).

[129] J. Thomsen, M. A. Gutowska, J. Saphorster, A. Heinemann, K. Trubenbach, J. Fietzke, C. Hiebenthal, A. Eisenhauer, A. Kortzinger, M. Wahl, F. Melzner, Calcifying invertebrates succeed in a naturally CO_2-rich coastal habitat but are threatened by high levels of future acidification. *Biogeosciences.* **7**, 3879–3891 (2010).

[130] E. M. Gosling, *The Mussel Mytilus: Ecology, Physiology, Genetics, and Culture* (Elsevier, Amsterdam, 1992).

[131] T. Kijewski, B. Śmietanka, M. Zbawicka, E. Gosling, H. Hummel, R. Wenne, Distribution of *Mytilus* taxa in European coastal areas as inferred from molecular markers. *J. Sea Res.* **65**, 224–234 (2011).

[132] J. Berge, G. Johnsen, F. Nilsen, B. Gulliksen, D. Slagstad, Ocean temperature oscillations enable reappearance of blue mussels *Mytilus edulis* in Svalbard after a 1000 year absence. *Mar. Ecol. Prog. Ser.* **303**, 167–175 (2005).

[133] S. J. Brooks, E. Farmen, The distribution of the mussel *Mytilus* species along the Norwegian coast. *J. Shellfish Res.* **32**, 265–270 (2013).

[134] M. Katolikova, V. Khaitov, R. Väinölä, M. Gantsevich, P. Strelkov, Genetic, ecological and morphological distinctness of the blue mussels *Mytilus trossulus* Gould and *M. edulis* L. in the White Sea. *PLoS One.* **11**, e0152963 (2016).

[135] S.-L. Varvio, R. K. Koehn, R. Väinölä, Evolutionary genetics of the *Mytilus edulis* complex in the North Atlantic region. *Mar. Biol.* **98**, 51–60 (1988).

[136] S. S. Mathiesen, J. Thyrring, J. Hemmer-Hansen, J. Berge, A. Sukhotin, P. Leopold, M. Bekaert, M. K. Sejr, E. E. Nielsen, Genetic diversity and connectivity within *Mytilus* spp. in the subarctic and Arctic. *Evol. Appl.* **10**, 39–55 (2017).

[137] M. Blicher, M. Sejr, S. Høgslund, Population structure of *Mytilus edulis* in the intertidal zone in a sub-Arctic fjord, SW Greenland. *Mar. Ecol. Prog. Ser.* **487**, 89–100 (2013).

[138] R. Wenne, L. Bach, M. Zbawicka, J. Strand, J. H. McDonald, A first report on coexistence and hybridization of *Mytilus trossulus* and *M. edulis* mussels in Greenland. *Polar Biol.* **39**, 343–355 (2016).

[139] A. D. Wanamaker, K. J. Kreutz, H. W. Borns, D. S. Introne, S. Feindel, S. Funder, P. D. Rawson, B. J. Barber, Experimental determination of salinity, temperature, growth, and metabolic effects on shell isotope chemistry of *Mytilus edulis* collected from Maine and Greenland.

Paleoceanography. **22**, PA2217 (2007).

[140] S. J. Jones, F. P. Lima, D. S. Wethey, Rising environmental temperatures and biogeography: poleward range contraction of the blue mussel, *Mytilus edulis* L., in the western Atlantic. *J. Biogeogr.* **37**, 2243–2259 (2010).

[141] E. M. Gosling, S. Doherty, N. Howley, Genetic characterization of hybrid mussel (*Mytilus*) populations on Irish coasts. *J. Mar. Biol. Assoc. UK*. **88**, 341–346 (2008).

[142] C. Daguin, F. Bonhomme, P. Borsa, The zone of sympatry and hybridization of *Mytilus edulis* and *M. galloprovincialis*, as described by intron length polymorphism at locus mac-1. *Heredity*. **86**, 342–354 (2001).

[143] N. Bierne, P. Borsa, C. Daguin, D. Jollivet, F. Viard, F. Bonhomme, P. David, Introgression patterns in the mosaic hybrid zone between *Mytilus edulis* and *M. galloprovincialis*. *Mol. Ecol.* **12**, 447–461 (2003).

[144] T. Hilbish, E. Carson, J. Plante, L. Weaver, M. Gilg, Distribution of *Mytilus edulis*, *M. galloprovincialis*, and their hybrids in open-coast populations of mussels in southwestern England. *Mar. Biol.* **140**, 137–142 (2002).

[145] J. P. A. Gardner, K. M. Westfall, Geographic distribution and molecular identification of a metapopulation of blue mussels (genus *Mytilus*) in northeastern New Zealand. *J. Molluscan Stud.* **78**, 66–73 (2012).

[146] G. M. Branch, C. Nina Steffani, Can we predict the effects of alien species? A case-history of the invasion of South Africa by *Mytilus galloprovincialis* (Lamarck). *J. Exp. Mar. Biol. Ecol.* **300**, 189–215 (2004).

[147] O. Aral, Growth of the Mediterranean mussel (*Mytilus galloprovincialis* Lam., 1819) on ropers in the Black Sea, Turkey. *Turkish J. Vet. Anim. Sci.* **23**, 183–189 (1999).

[148] R. Väinölä, P. Strelkov, *Mytilus trossulus* in Northern Europe. *Mar. Biol.* **158**, 817–833 (2011).

[149] H. Stuckas, L. Knöbel, H. Schade, C. Breusing, H.-H. Hinrichsen, M. Bartel, K. Langguth, F. Melzner, Combining hydrodynamic modelling with genetics: can passive larval drift shape the genetic structure of Baltic *Mytilus* populations? *Mol. Ecol.* **26**, 2765–2782 (2017).

[150] T. K. Kijewski, M. Zbawicka, R. Väinölä, R. Wenne, Introgression and mitochondrial DNA heteroplasmy in the Baltic populations of mussels *Mytilus trossulus* and *M. edulis*. *Mar. Biol.* **149**, 1371–1385 (2006).

[151] C. Riginos, C. W. Cunningham, Local adaptation and species segregation in two mussel (*Mytilus edulis* × *Mytilus trossulus*) hybrid zones. *Mol. Ecol.* **14**, 381–400 (2005).

[152] N. Kautsky, K. Johannesson, M. Tedengren, Genotypic and phenotypic differences between Baltic and North Sea populations of *Mytilus edulis* evaluated through reciprocal transplantations. I Growth and morphology. *Mar. Ecol. Prog. Ser.* **59**, 203–210 (1990).

[153] G. Ridgway, G. Nævdal, Genotypes of *Mytilus* from waters of different salinity around Bergen, Norway. *Helgol. Mar. Res.* **58**, 104–109 (2004).

[154] P. Enderlein, M. Wahl, Dominance of blue mussels versus consumer-mediated enhancement of benthic diversity. *J. Sea Res.* **51**, 145–155 (2004).

[155] A. R. Beaumont, M. P. Hawkins, F. L. Doig, I. M. Davies, M. Snow, Three species of *Mytilus* and their hybrids identified in a Scottish Loch: natives, relicts and invaders? *J. Exp. Mar. Biol. Ecol.* **367**, 100–110 (2008).

[156] K. Michalek, A. Ventura, T. Sanders, *Mytilus* hybridisation and impact on aquaculture: a minireview. *Mar. Genom.* **27**, 3–7 (2016).

[157] A. R. Beaumont, G. Turner, A. R. Wood, D. O. . Skibinski, Hybridisations between *Mytilus edulis* and *Mytilus galloprovincialis* and performance of pure species and hybrid veliger larvae at different temperatures. *J. Exp. Mar. Biol. Ecol.* **302**, 177–188 (2004).

[158] P. J. Dias, A. Dordor, D. Tulett, S. Piertney, I. M. Davies, M. Snow, Survey of mussel (*Mytilus*) species at Scottish shellfish farms. *Aquac. Res.* **40**, 1715–1722 (2009).

[159] R. Seed, Factors influencing shell shape in the mussel *Mytilus edulis*. *J. Mar. Biol. Assoc. UK*. **48**, 561–584 (1968).

[160] L. Tomanek, B. Helmuth, Physiological ecology of rocky intertidal organisms: a synergy of concepts. *Integr. Comp. Biol.* **42**, 771–775 (2002).

[161] B. Helmuth, Thermal biology of rocky intertidal mussels: quantifying body temperatures using climatological data. *Ecology*. **80**, 15–34 (1999).

[162] R. Seed, C. A. Richardson, in *The Neurobiology of Mytilus edulis*, G. B. Stefano, Ed. (Manchester University Press, Manchester, UK, 1990), pp. 1–37.

[163] K. J. Kroeker, B. Gaylord, T. M. Hill, J. D. Hosfelt, S. H. Miller, E. Sanford, The role of temperature in determining species' vulnerability to ocean acidification: a case study using *Mytilus galloprovincialis*. *PLoS One*. **9**, e100353 (2014).

[164] P. Bergström, M. Lindegarth, Environmental influence on mussel (*Mytilus edulis*) growth – a quantile regression approach. *Estuar. Coast Shelf Sci.* **171**, 123–132 (2016).

[165] P. C. Almada-Villela, J. Davenport, L. D. Gruffydd, The effects of temperature on the shell growth of young *Mytilus edulis* L. *J. Exp. Mar. Biol. Ecol.* **59**, 275–288 (1982).

[166] J. Thyrring, S. Rysgaard, M. E. Blicher, M. K. Sejr, Metabolic cold adaptation and aerobic performance of blue mussels (*Mytilus edulis*) along a temperature gradient into the High Arctic region. *Mar. Biol.* **162**, 235–243 (2015).

[167] A. A. Sukhotin, D. Abele, H.-O. Pörtner, Ageing and metabolism of *Mytilus edulis*: populations from various climate regimes. *J. Shellfish Res.* **25**, 893–899 (2006).

[168] M. Solan, N. Whiteley, *Stressors in the Marine Environment* (Oxford University Press, Oxford, UK, 2016).

[169] H. U. Riisgård, L. Bøttiger, D. Pleissner, Effect of salinity on growth of mussels, *Mytilus edulis*, with special reference to Great Belt (Denmark). *Open J. Mar. Sci.* **2**, 167–176 (2012).

[170] T. Sanders, L. Schmittmann, J. C. Nascimento-Schulze, F. Melzner, High calcification costs

limit mussel growth at low salinity. *Front. Mar. Sci.* **5**, 352 (2018).

[171] P. C. Almada-Villela, The effect of reduced salinity on the shell growth of small *Mytilus edulis*. *J. Mar. Biol. Assoc. UK*. **64**, 171–182 (1984).

[172] M. Westerbom, M. Kilpi, O. Mustonen, Blue mussels, *Mytilus edulis*, at the edge of the range: population structure, growth and biomass along a salinity gradient in the north-eastern Baltic Sea. *Mar. Biol.* **140**, 991–999 (2002).

[173] F. Møhlenberg, H. U. Riisgård, Efficiency of particle retention in 13 species of suspension feeding bivalves. *Ophelia*. **17**, 239–246 (1978).

[174] S. A. Navarrete, B. A. Menge, Keystone predation and interaction strength: Interactive effects of predators ontheir main prey. *Ecol. Monogr.* **66**, 409–429 (1996).

[175] S. E. a. L. V. Dit Durell, J. D. Goss-Custard, Prey selection within a size-class of mussels, *Mytilus edulis*, by oystercatchers, *Haematopus ostralegus*. *Anim. Behav.* **32**, 1197–1203 (1984).

[176] R. Seed, The ecology of *Mytilus edulis* L. (Lamellibranchiata) on exposed rocky shores. *Oecologia*. **3**, 317–350 (1969).

[177] B. A. Menge, Components of predation intensity in the low zone of the New England rocky intertidal region. *Oecologia*. **58**, 141–155 (1983).

[178] R. N. Hughes, S. de B. Dunkin, Effect of dietary history on selection of prey, and foraging behaviour among patches of prey, by the dogwhelk, *Nucella lapillus* (L.). *J. Exp. Mar. Biol. Ecol.* **79**, 159–172 (1984).

[179] E. M. Harper, P. W. Skelton, A defensive value of the thickened periostracum in the Mytiloidea. *Veliger*. **36**, 36–42 (1993).

[180] M. R. Carriker, Shell penetration and feeding by naticacean and muricacean predatory gastropods: A synthesis. *Malacologia*. **20**, 403–422 (1981).

[181] M. Carroll, R. Highsmith, Role of catastrophic disturbance in mediating *Nucella-Mytilus* interactions in the Alaskan rocky intertidal. *Mar. Ecol. Prog. Ser.* **138**, 125–133 (1996).

[182] B. Morton, E. M. Harper, Predation upon *Mytilus galloprovincialis* (Mollusca: Bivalvia: Mytilidae) by juvenile *Carcinus maenas* (Crustacea: Decapoda) using mandibular chipping. **88**, 563–568 (2008).

[183] P. Schwemmer, S. Garthe, At-sea distribution and behaviour of a surface-feeding seabird, the lesser black-backed gull *Larus fuscus*, and its association with different prey. *Mar. Ecol. Prog. Ser.* **285**, 245–258 (2005).

[184] J. D. Taylor, W. J. Kennedy, The influence of the periostracum on the shell structure of bivalve molluscs. *Calcif. Tissue Res.* **3**, 274–283 (1969).

[185] E. M. Harper, The molluscan periostracum: an important constraint in bivalve evolution. *Palaeontology*. **40**, 71–97 (1997).

[186] J. H. Waite, S. O. Andersen, 3,4-Dihydroxyphenylalanine (Dopa) and sclerotization of periostracum in *Mytilus edulis* L. *Biol. Bull.* **158**, 164–173 (1980).

[187] D. J. Taylor, W. J. Kennedy, A. Hall, *The Shell Structure and Mineralogy of the Bivalvia II, Lucinacea-Clavagellacea* (London Bulletin of the British Museum National History Zoology, London, 1973).

[188] V. L. Peck, G. A. Tarling, C. Manno, E. M. Harper, E. Tynan, Outer organic layer and internal repair mechanism protects pteropod *Limacina helicina* from ocean acidification. *Deep-Sea Res. Pt. II*. **127**, 41–52 (2016).

[189] J. Arivalagan, B. Marie, V. A. Sleight, M. S. Clark, S. Berland, A. Marie, Shell matrix proteins of the clam, *Mya truncata*: roles beyond shell formation through proteomic study. *Mar. Genom.* **27**, 69–74 (2016).

[190] A. G. Checa, C. M. Pina, A. J. Osuna-Mascaró, A. B. Rodríguez-Navarro, E. M. Harper, Crystalline organization of the fibrous prismatic calcitic layer of the Mediterranean mussel *Mytilus galloprovincialis*. *Eur. J. Mineral.* **26**, 495–505 (2014).

[191] C. Gregoire, Structure of the conchiolin cases of the prisms in *Mytilus edulis* Linne. *J. Cell Biol.* **9**, 395–400 (1961).

[192] M. Cusack, A. Pérez-Huerta, P. Dalbeck, Common crystallographic control in calcite biomineralization of bivalved shells. *CrystEngComm*. **9**, 1215–1218 (2007).

[193] Q. Feng, H. Li, G. Pu, D. Zhang, F. Cui, Crystallographic alignment of calcite prisms in the oblique prismatic layer of *Mytilus edulis* shell. *J. Mater. Sci.* **5**, 3337–3340 (2000).

[194] A. Checa, A new model for periostracum and shell formation in Unionidae (Bivalvia, Mollusca). *Tissue Cell*. **32**, 405–416 (2000).

[195] D. E. Jacob, R. Wirth, A. L. Soldati, U. Wehrmeister, A. Schreiber, Amorphous calcium carbonate in the shells of adult Unionoida. *J. Struct. Biol.* **173**, 241–249 (2011).

[196] K. Ramesh, M. Y. Hu, J. Thomsen, M. Bleich, F. Melzner, Mussel larvae modify calcifying fluid carbonate chemistry to promote calcification. *Nat. Commun.* **8**, 1709 (2017).

[197] E. M. Harper, Are calcitic layers an effective adaptation against shell dissolution in the Bivalvia? *J. Zool.* **251**, 179–186 (2000).

[198] A. S. Freeman, J. E. Byers, Divergent induced responses to an invasive predator in marine mussel populations. *Science*. **313**, 831–833 (2006).

[199] L. D. Smith, J. A. Jennings, Induced defensive responses by the bivalve *Mytilus edulis* to predators with different attack modes. *Mar. Biol.* **136**, 461–469 (2000).

[200] A. P. Le Rossignol, S. G. Buckingham, S. E. G. Lea, R. Nagarajan, Breaking down the mussel (*Mytilus edulis*) shell: which layers affect Oystercatchers' (*Haematopus ostralegus*) prey selection? *J. Exp. Mar. Biol. Ecol.* **405**, 87–92 (2011).

[201] A. G. Checa, H. Mutvei, A. J. Osuna-Mascaró, J. T. Bonarski, M. Faryna, K. Berent, C. M. Pina, M. Rousseau, E. Macías-Sánchez, Crystallographic control on the substructure of nacre tablets. *J. Struct. Biol.* **183**, 368–376 (2013).

[202] A. G. Checa, J. H. E. Cartwright, M.-G. Willinger, Mineral bridges in nacre. *J. Struct. Biol.* **176**,

330–339 (2011).

[203] J. H. E. Cartwright, A. G. Checa, The dynamics of nacre self-assembly. *J. R. Soc. Interface*. **4**, 491–504 (2007).

[204] Q. Feng, H. Li, F. Cui, H. Li, T. Kim, Crystal orientation domains found in the single lamina in nacre of the *Mytilus edulis* shell. *J. Mater. Sci. …*. **8**, 1547–1549 (1999).

[205] J. P. A. Gardner, R. J. Thompson, Influence of genotype and geography on shell shape and morphometric trait variation among North Atlantic blue mussel (*Mytilus* spp.) populations. *Biol. J. Linn. Soc.* **96**, 875–897 (2009).

[206] C. Steffani, G. Branch, Growth rate, condition, and shell shape of *Mytilus galloprovincialis*: responses to wave exposure. *Mar. Ecol. Prog. Ser.* **246**, 197–209 (2003).

[207] A. Valladares, G. Manríquez, B. Suárez-Isla, Shell shape variation in populations of *Mytilus chilensis* (Hupé 1854) from southern Chile: a geometric morphometric approach. *Mar. Biol.* **157**, 2731–2738 (2010).

[208] S. Krapivka, J. E. Toro, A. C. Alcapán, M. Astorga, P. Presa, M. Pérez, R. Guiñez, Shell-shape variation along the latitudinal range of the Chilean blue mussel *Mytilus chilensis* (Hupé 1854). *Aquac. Res.* **38**, 1770–1777 (2007).

[209] R. Seed, Absolute and allometric growth in the mussel, *Mytilus edulis* L . (Mollusca Bivalvia). *Proc. Malacol. Soc. Lond.* **40**, 343–357 (1973).

[210] J. H. McDonald, R. Seed, R. K. Koehn, Allozymes and morphometric characters of three species of *Mytilus* in the Northern and Southern Hemispheres. *Mar. Biol.* **111**, 323–333 (1991).

[211] A. Illesca, P. A. Oyarzún, J. E. Toro, J. P. A. Gardner, Morphometric variability of smooth-shelled blue mussels from the Pacific coast of South America. *Biol. J. Linn. Soc.* **20**, 1–16 (2018).

[212] T. A. McConnaughey, J. F. Whelan, Calcification generates protons for nutrient and bicarbonate uptake. *Earth-Science Rev.* **42**, 95–117 (1997).

[213] G. Bevelander, Calcification in molluscs. III. Intake and deposition of Ca^{45} and P^{32} in relation to shell formation. *Biol. Bull.* **102**, 9–15 (1952).

[214] J. K. Sillanpää, K. Ramesh, F. Melzner, H. Sundh, K. Sundell, Calcium mobilisation following shell damage in the Pacific oyster, *Crassostrea gigas*. *Mar. Genom.* **27**, 75–83 (2016).

[215] T. A. McConnaughey, D. P. Gillikin, Carbon isotopes in mollusk shell carbonates. *Geo-Marine Lett.* **28**, 287–299 (2008).

[216] L. Addadi, D. Joester, F. Nudelman, S. Weiner, Mollusk shell formation: a source of new concepts for understanding biomineralization processes. *Chem. - A Eur. J.* **12**, 980–987 (2006).

[217] S. Weiner, L. Hood, Soluble protein of the organic matrix of mollusk shells: a potential template for shell formation. *Science*. **190**, 987–989 (1975).

[218] S. Weiner, L. Addadi, Crystallization pathways in biomineralization. *Annu. Rev. Mater. Res.* **41**, 21–40 (2011).

[219] K. Saruwatari, T. Matsui, H. Mukai, H. Nagasawa, T. Kogure, Nucleation and growth of aragonite crystals at the growth front of nacres in pearl oyster, *Pinctada fucata*. *Biomaterials*. **30**, 3028–3034 (2009).

[220] K. Ramesh, F. Melzner, A. W. Griffith, C. J. Gobler, C. Rouger, D. Tasdemir, G. Nehrke, *In vivo* characterization of bivalve larval shells: a confocal Raman microscopy study. *J. R. Soc. Interface*. **15**, 20170723 (2018).

[221] R. T. DeVol, C.-Y. Sun, M. A. Marcus, S. N. Coppersmith, S. C. B. Myneni, P. U. P. A. Gilbert, Nanoscale transforming mineral phases in fresh nacre. *J. Am. Chem. Soc.* **137**, 13325–13333 (2015).

[222] S. U. Goldenberg, I. Nagelkerken, C. M. Ferreira, H. Ullah, S. D. Connell, Boosted food web productivity through ocean acidification collapses under warming. *Glob. Change Biol.* **23**, 4177–4184 (2017).

[223] B. Gaylord, T. M. Hill, E. Sanford, E. A. Lenz, L. A. Jacobs, K. N. Sato, A. D. Russell, A. Hettinger, Functional impacts of ocean acidification in an ecologically critical foundation species. *J. Exp. Biol.* **214**, 2586–2594 (2011).

[224] S. C. Fitzer, M. Cusack, V. R. Phoenix, N. a. Kamenos, Ocean acidification reduces the crystallographic control in juvenile mussel shells. *J. Struct. Biol.* **188**, 39–45 (2014).

[225] S. Hahn, R. Rodolfo-Metalpa, E. Griesshaber, W. W. Schmahl, D. Buhl, J. M. Hall-Spencer, C. Baggini, K. T. Fehr, A. Immenhauser, Marine bivalve shell geochemistry and ultrastructure from modern low pH environments: environmental effect versus experimental bias. *Biogeosciences*. **9**, 1897–1914 (2012).

[226] S. Hahn, E. Griesshaber, W. W. Schmahl, R. D. Neuser, A.-C. Ritter, R. Hoffmann, D. Buhl, A. Niedermayr, A. Geske, A. Immenhauser, Exploring aberrant bivalve shell ultrastructure and geochemistry as proxies for past sea water acidification. *Sedimentology*. **61**, 1625–1658 (2014).

[227] B. Michaelidis, C. Ouzounis, A. Paleras, H. Pörtner, Effects of long-term moderate hypercapnia on acid-base balance and growth rate in marine mussels *Mytilus galloprovincialis*. *Mar. Ecol. Prog. Ser.* **293**, 109–118 (2005).

[228] A. K. Hüning, F. Melzner, J. Thomsen, M. A. Gutowska, L. Krämer, S. Frickenhaus, P. Rosenstiel, H.-O. Pörtner, E. E. R. Philipp, M. Lucassen, Impacts of seawater acidification on mantle gene expression patterns of the Baltic Sea blue mussel: implications for shell formation and energy metabolism. *Mar. Biol.* **160**, 1845–1861 (2013).

[229] A. Beesley, D. Lowe, C. Pascoe, S. Widdicombe, Effects of CO_2-induced seawater acidification on the health of *Mytilus edulis*. *Clim. Res.* **37**, 215–225 (2008).

[230] R. Bibby, S. Widdicombe, H. Parry, J. Spicer, R. Pipe, Effects of ocean acidification on the immune response of the blue mussel *Mytilus edulis*. *Aquat. Biol.* **2**, 67–74 (2008).

[231] H. Kurihara, T. Asai, S. Kato, A. Ishimatsu, Effects of elevated pCO_2 on early development in the mussel *Mytilus galloprovincialis*. *Aquat. Biol.* **4**, 225–233 (2008).

[232] T. Rossi, I. Nagelkerken, J. C. A. Pistevos, S. D. Connell, Lost at sea: ocean acidification undermines larval fish orientation via altered hearing and marine soundscape modification. *Biol. Lett.* **12**, 20150937 (2016).

[233] L. S. Peck, T. Souster, M. S. Clark, Juveniles are more resistant to warming than adults in 4 species of Antarctic marine invertebrates. *PLoS One*. **8**, e66033 (2013).

[234] S. D. Connell, K. J. Kroeker, K. E. Fabricius, D. I. Kline, B. D. Russell, The other ocean acidification problem: CO_2 as a resource among competitors for ecosystem dominance. *Philos. Trans. R. Soc. B Biol. Sci.* **368**, 20120442 (2013).

[235] J. T. Wootton, The nature and consequences of indirect effects in ecological communities. *Annu. Rev. Ecol. Syst.* **25**, 443–466 (1994).

[236] J. Lord, J. Barry, D. Graves, Impact of climate change on direct and indirect species interactions. *Mar. Ecol. Prog. Ser.* **571**, 1–11 (2017).

[237] J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J.-Y. Tinevez, D. J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, A. Cardona, Fiji: an open-source platform for biological-image analysis. *Nat. Methods*. **9**, 676–682 (2012).

[238] P. J. Haines, *Principles of Thermal Analysis and Calorimetry* (Royal Society of Chemistry, Cambridge, UK, 2002).

[239] S. Gaisford, V. Kett, P. J. Haines, *Principles of Thermal Analysis and Calorimetry: Edition 2* (Royal Society of Chemistry, Cambridge, UK, 2016).

[240] F. D. Fleischli, M. Dietiker, C. Borgia, R. Spolenak, The influence of internal length scales on mechanical properties in natural nanocomposites: a comparative study on inner layers of seashells. *Acta Biomater.* **4**, 1694–1706 (2008).

[241] C. M. Zaremba, D. E. Morse, S. Mann, P. K. Hansma, G. D. Stucky, Aragonite−hydroxyapatite conversion in gastropod (Abalone) nacre. *Chem. Mater.* **10**, 3813–3824 (1998).

[242] N. M. Neves, J. F. Mano, Structure/mechanical behavior relationships in crossed-lamellar sea shells. *Mater. Sci. Eng. C*. **25**, 113–118 (2005).

[243] J. Balmain, B. Hannoyer, E. Lopez, Fourier transform infrared spectroscopy (FTIR) and X-ray diffraction analyses of mineral and organic matrix during heating of mother of pearl (nacre) from the shell of the mollusc *Pinctada maxima*. *J. Biomed. Mater. Res.* **48**, 749–754 (1999).

[244] A. G. Checa, E. Macías-Sánchez, E. M. Harper, J. H. E. Cartwright, Organic membranes determine the pattern of the columnar prismatic layer of mollusc shells. *Proc. R. Soc. B Biol. Sci.* **283**, 20160032 (2016).

[245] D. W. Thompson, *On Growth and Form* (Cambridge University Press, Cambridge, 1917).

[246] D. C. Adams, F. J. Rohlf, D. E. Slice, Geometric morphometrics: ten years of progress following the "revolution." *Ital. J. Zool.* **71**, 5–16 (2004).

[247] K. Schmidt-Nielsen, *Scaling. Why is Animal Size so Important?* (Cambridge University Press,

New York, NY, 1984).

[248] M. L. Zelditch, D. L. Swiderski, H. D. Sheets, *Geometric Morphometrics for Biologists: a Primer* (Elsevier, ed. 2nd, 2012).

[249] R. Seed, in *Skeletal Growth of Aquatic Organisms*, D. C. Rhoads, R. A. Lutz, Eds. (Plenum Press, New York, NY, USA, 1980), pp. 23–67.

[250] J. F. White, S. J. Gould, Interpretation of the coefficient in the allometric equation. *Am. Nat.* **99**, 5–18 (1965).

[251] J. S. Huxley, G. Teissier, Terminology of relative growth. *Nature*. **137**, 780–781 (1936).

[252] M. Claude, Log-shape ratios, Procrustes superimposition, elliptic Fourier analysis: three worked examples in R. *Hystrix*. **24**, 94–102 (2013).

[253] F. L. Bookstein, A hundred years of morphometrics. *Acta Zool. Acad. Sci. Hungaricae*. **44**, 7–59 (1998).

[254] K. Pearson, Note on regression and inheritance in the case of two parents. *Proc. R. Soc. London*. **58**, 240–242 (1895).

[255] R. A. Fisher, The logic of inductive inference. *J. R. Stat. Soc.* **98**, 39 (1935).

[256] K. Pearson, On lines and planes of closest fit to systems of points in space. *Philos. Mag.* **2**, 559–572 (1901).

[257] H. Hotelling, Analysis of a complex of statistical variables into principal components. *J. Educ. Psychol.* **24**, 417–441 (1933).

[258] J. Claude, *Morphometrics with R* (Springer, New York, USA, 2008).

[259] F. L. Bookstein, *Morphometric Tools for Landmark Data: Geometry and Biology* (Cambridge University Press, Cambridge, 1991).

[260] D. G. Kendall, A Survey of the Statistical Theory of Shape. *Stat. Sci.* **4**, 116–120 (1989).

[261] C. G. Small, *The Statistical Theory of Shape* (Springer, 1996), vol. 93.

[262] I. L. Dryden, K. V. Mardia, *Statistical Shape Analysis* (John Wiley & Sons, New York, NY, USA, 1988).

[263] F. J. Rohlf, L. F. Marcus, A revolution in morphometrics. *Trends Ecol. Evol.* **8**, 129–132 (1993).

[264] L. F. Marcus, in *Proceedings of the Michigan Morphometrics Workshop*, F. J. Rohlf, F. L. Bookstein, Eds. (University of Michigan Museum of Zoology, Ann Arbor, 1990), pp. 77–122.

[265] R. A. Reyment, *Multidimensional Palaeobiology* (Pergamon Press, Oxford, NY, USA, 1991).

[266] R. E. Blackith, R. A. Reyment, *Multivariate Morphometrics* (Academic Press, New York, NY, USA, 1971).

[267] D. C. Adams, F. J. Rohlf, D. E. Slice, A field comes of age: geometric morphometrics in the 21st century. *Hystrix*. **24**, 7–14 (2013).

[268] F. L. Bookstein, *Morphometrics in Evolutionary Biology: The Geometry of Size and Shape Change, with Examples from Fishes* (Academy of Natural Sciences of Philadelphia, Philadephila, PE, USA, 1985).

[269] J. D. Reist, An empirical evaluation of coefficients used in residual and allometric adjustment of size covariation. *Can. J. Zool.* **64**, 1363–1368 (1986).

[270] J. Lleonart, J. Salat, G. J. Torres, Removing allometric effects of body size in morphological analysis. *J. Theor. Biol.* **205**, 85–93 (2000).

[271] R. S. Thorpe, Quantitative handling of characters useful in snake systematics with particular reference to intraspecific variation in the Ringed Snake *Natrix natrix* (L.). *Biol. J. Linn. Soc.* **7**, 27–43 (1975).

[272] J. D. Reist, An empirical evaluation of several univariate methods that adjust for size variation in morphometric data. *Can. J. Zool.* **63**, 1429–1439 (1985).

[273] P. Dodson, On the use of ratios in growth studies. *Syst. Zool.* **27**, 62–67 (1978).

[274] V. Bonhomme, S. Picq, C. Gaucherel, J. Claude, Momocs: outline analysis using R. *J. Stat. Softw.* **56**, 1–24 (2014).

[275] C. R. Giardina, F. P. Kuhl, Accuracy of curve approximation by harmonically related vectors with elliptical loci. *Comput. Graph. Image Process.* **6**, 277–285 (1977).

[276] F. P. Kuhl, C. R. Giardina, Elliptic Fourier features of a closed contour. *Comput. Graph. Image Process.* **18**, 236–258 (1982).

[277] H. Moellering, J. N. Rayner, The harmonic analysis of spatial shapes using Dual Axis Fourier Shape Analysis (DAFSA). *Geogr. Anal.* **13**, 64–77 (1981).

[278] A. J. Haines, J. S. Crampton, Improvements to the method of Fourier shape analysis as applied in morphometric studies. *Palaeontology.* **43**, 765–783 (2000).

[279] J. T. Richtsmeier, V. Burke Deleon, S. R. Lele, The promise of geometric morphometrics. *Am. J. Phys. Anthropol.* **119**, 63–91 (2002).

[280] J. T. Richtsmeier, J. M. Cheverud, L. Subhash, Advances in anthropological morphometrics. *Annu. Rev. Anthropol.* **21**, 283–305 (1992).

[281] F. J. Rohlf, J. W. Archie, A comparison of Fourier methods for the description of wing shape in mosquitoes (Diptera: Culicidae). *Syst. Zool.* **33**, 302 (1984).

[282] C. T. Zahn, R. Z. Roskies, Fourier descriptors for plane closed curves. *IEEE Trans. Comput.* **21**, 269–281 (1972).

[283] J. S. Crampton, Elliptic Fourier shape analysis of fossil bivalves: some practical considerations. *Lethaia.* **28**, 179–186 (1995).

[284] S. Renaud, J. R. Michaux, Adaptive latitudinal trends in the mandible shape of *Apodemus* wood mice. *J. Biogeogr.* **30**, 1617–1628 (2003).

[285] R Core Team, R: a language and environment for statistical computing. (2017), (available at https://www.r-project.org/).

[286] R. A. Fisher, The use of multiple measurements in taxonomic problems. *Ann. Eugen.* **7**, 179–188 (1936).

[287] A. F. Zuur, E. N. Ieno, N. Walker, A. A. Saveliev, G. M. Smith, *Mixed Effects Models and*

Extensions in Ecology with R (Springer New York, New York, NY, 2009).

[288] S. N. Wood, *Generalized Additive Models : an Introduction with R, Second Edition* (Chapman & Hall/CRC, London, 2017).

[289] J. Pinheiro, D. Bates, *Mixed-Effects Models in S and S-PLUS* (Springer-Verlag, New York, USA, 2000).

[290] D. Bates, M. Mächler, B. Bolker, S. Walker, Fitting linear mixed-effects models using *lme4*. *J. Stat. Softw.* **67**, 1–48 (2015).

[291] S. N. Wood, *Generalized Additive Models: an Introduction with R* (Chapman and Hall, London, 2006).

[292] T. Hastie, R. Tibshirani, *Generalized Additive Models* (Chapman and Hall, London, 1990).

[293] B. D. Ripley, in *Networks and Chaos: Statistical and Probabilistic Aspects*, O. E. Barndorff-Nielsen, J. L. Jensen, W. . Kendall, Eds. (Chapman & Hall, London, UK, 1993), pp. 40–123.

[294] A. F. Zuur, E. N. Ieno, C. S. Elphick, A protocol for data exploration to avoid common statistical problems. *Methods Ecol. Evol.* **1**, 3–14 (2010).

[295] J. Fox, *Applied Regression Analysis and Generalized Linear Models* (SAGE Publications, CA, 2015).

[296] C. J. Huberty, *Applied Discriminant Analysis* (Wiley, New York, NY, USA, 1994).

[297] M. H. Graham, Confronting multicollinearity in ecological multiple regression. *Ecology.* **84**, 2809–2815 (2003).

[298] A. C. Cameron, P. K. Trivedi, *Regression Analysis of Count Data* (Cambridge University Press, Cambridge, UK, 2013).

[299] A. F. Zuur, E. N. Ieno, G. M. Smith, *Analyzing Ecological Data* (2007).

[300] C. Chatfield, *Problem Solving: A Statistician's Guide* (Chapman and Hall/CRC, New York, NY, USA, 1995).

[301] A. Gelman, C. Pasarica, R. Dodhia, Let's practice what we preach. *Am. Stat.* **56**, 121–130 (2002).

[302] D. C. Montgomery, E. A. Peck, G. G. Vining, *Introduction to Linear Regression Analysis* (Wiley, 2012).

[303] G. P. Quinn, M. J. Keough, *Experimental Design and Data Analysis for Biologists* (Cambridge University Press, 2002).

[304] W. S. Cleveland, *Visualizing Data* (At & T Bell Laboratories, Murray Hill, NJ, USA, 1993).

[305] I. T. Jolliffe, *Principal Component Analysis* (Springer-Verlag, New York, NY, USA, 2010).

[306] R. R. Sokal, F. J. Rohlf, *Biometry: the Principles and Practice of Statistics in Biological Research. 3rd Edition* (W.H. Freeman, New York, NY, USA, 1995).

[307] E. Läärä, Statistics: reasoning on uncertainty, and the insignificance of testing null. *Ann. Zool. Fennici.* **46**, 138–157 (2009).

[308] N. R. Draper, H. Smith, *Applied Regression Analysis* (Wiley, New York, NY, USA, 1998).

[309] S. H. Hurlbert, Pseudoreplication and the design of ecological field experiments. *Ecol. Monogr.*

54, 187–211 (1984).

[310] B. M. Bolker, M. E. Brooks, C. J. Clark, S. W. Geange, J. R. Poulsen, M. H. H. Stevens, J.-S. S. White, Generalized linear mixed models: a practical guide for ecology and evolution. *Trends Ecol. Evol.* **24**, 127–135 (2009).

[311] O. Schabenberger, F. J. Pierce, *Contemporary Statistical Models for the Plant and Soil Sciences* (CRC Press, Boca Raton, FL, USA, 2002).

[312] J. A. Nelder, R. W. M. Wedderburn, Generalized linear models. *J. R. Stat. Soc. Ser. A.* **135**, 370–384 (1972).

[313] D. I. Warton, F. K. C. Hui, The arcsine is asinine: the analysis of proportions in ecology. *Ecology.* **92**, 3–10 (2011).

[314] R. B. O'Hara, D. J. Kotze, Do not log-transform count data. *Methods Ecol. Evol.* **1**, 118–122 (2010).

[315] G. A. Fox, S. Negrete-Yankelevich, V. J. Sosa, *Ecological Statistics : Contemporary Theory and Application* (Oxford University Press, Oxford, UK, 2015).

[316] A. F. Zuur, J. M. Hilbe, E. N. Ieno, *A Beginner's Guide to GLM and GLMM with R* (Highland Statistics Ltd, Newburgh, UK, 2015).

[317] M. J. Crawley, *The R Book* (John Wiley & Sons, Ltd, Chichester, UK, 2007).

[318] N. M. Laird, J. H. Ware, Random-effects models for longitudinal data. *Biometrics.* **38**, 963–974 (1982).

[319] D. Ruppert, M. P. Wand, R. J. Carroll, *Semiparametric Regression* (Cambridge University Press, Cambridge, 2003).

[320] B. M. Bolker, *Ecological Models and Data in R* (Princeton University Press, Princeton, NJ, USA, 2008).

[321] K. P. Burnham, D. R. Anderson, *Model Selection and Multimodel Inference : a Practical Information-Theoretic Approach* (Springer-Verlag, New York, NY, USA, 2002).

[322] U. Halekoh, S. Højsgaard, A Kenward-Roger approximation and parametric bootstrap methods for tests in linear mixed models - The R package pbkrtest. *J. Stat. Softw.* **59**, 1–32 (2014).

[323] A. F. Zuur, E. N. Ieno, A protocol for conducting and presenting results of regression-type analyses. *Methods Ecol. Evol.* **7**, 636–645 (2016).

[324] D. J. Innes, J. A. Bates, Morphological variation of *Mytilus edulis* and *Mytilus trossulus* in eastern Newfoundland. *Mar. Biol.* **133**, 691–699 (1999).

[325] K. J. Kroeker, F. Micheli, M. C. Gambi, T. R. Martz, Divergent ecosystem responses within a benthic marine community to ocean acidification. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 14515–14520 (2011).

[326] N. Caill-Milly, N. Bru, K. Mahé, C. Borie, F. D'Amico, Shell shape analysis and spatial allometry patterns of Manila clam (*Ruditapes philippinarum*) in a mesotidal coastal lagoon. *J. Mar. Biol.* **2012**, 1–11 (2012).

[327] A. F. Zuur, A. A. Saveliev, E. N. Ieno, *A Beginner's Guide to Generalized Additive Mixed Models with R* (Highland Statistics Ltd, Newburgh, 2014).

[328] EU Copernicus Marine Service, Copernicus marine environment monitoring service - CMEMS (2018), (available at http://marine.copernicus.eu/).

[329] Y. Thomas, J. Mazurié, M. Alunno-Bruscia, C. Bacher, J.-F. Bouget, F. Gohin, S. Pouvreau, C. Struski, Modelling spatio-temporal variability of *Mytilus edulis* (L.) growth by forcing a dynamic energy budget model with satellite-derived environmental data. *J. Sea Res.* **66**, 308–317 (2011).

[330] IOCCG, "Phytoplankton Functional Types from Space. Reports of the International Ocean-Colour Coordinating Group, No. 15" (IOCCG, Dartmouth, 2014).

[331] R. W. Penney, M. J. Hart, N. D. Templeman, Shell strength and appearance in cultured blue mussels *Mytilus edulis*, *M. trossulus*, and *M. edulis* × *M. trossulus* hybrids. *N. Am. J. Aquacult.* **69**, 281–295 (2007).

[332] R. Guiñez, J. C. Castilla, A tridimensional self-thinning model for multilayered intertidal mussels. *Am. Nat.* **154**, 341–357 (1999).

[333] F. Valladares, S. Matesanz, F. Guilhaumon, M. B. Araújo, L. Balaguer, M. Benito-Garzón, W. Cornwell, E. Gianoli, M. van Kleunen, D. E. Naya, A. B. Nicotra, H. Poorter, M. A. Zavala, The effects of phenotypic plasticity and local adaptation on forecasts of species range shifts under climate change. *Ecol. Lett.* **17**, 1351–1364 (2014).

[334] O. Reimer, M. Tedengren, Phenotypical improvement of morphological defences in the mussel *Mytilus edulis* induced by exposure to the predator *Asterias rubens*. *Oikos*. **75**, 383 (1996).

[335] A. M. Cubillo, L. G. Peteiro, M. J. Fernández-Reiriz, U. Labarta, Influence of stocking density on growth of mussels (*Mytilus galloprovincialis*) in suspended culture. *Aquaculture*. **342**–**343**, 103–111 (2012).

[336] S. Dürr, M. Wahl, Isolated and combined impacts of blue mussels (*Mytilus edulis*) and barnacles (*Balanus improvisus*) on structure and diversity of a fouling community. *J. Exp. Mar. Biol. Ecol.* **306**, 181–195 (2004).

[337] J. Norberg, M. Tedengren, Attack behaviour and predatory success of *Asterias rubens* L. related to differences in size and morphology of the prey mussel *Mytilus edulis* L. *J. Exp. Mar. Biol. Ecol.* **186**, 207–220 (1995).

[338] E. M. Harper, L. S. Peck, Latitudinal and depth gradients in marine predation pressure. *Glob. Ecol. Biogeogr.* **25**, 670–678 (2016).

[339] C. Breusing, *Population Genetics and Morphometric Variation of Blue Mussels in the Western Baltic Sea.* (Master Thesis. GEOMAR Helmholtz Centre for Ocean Research Kiel, 2012).

[340] U. Schiewer, *Ecology of Baltic Coastal Waters* (Springer-Verlag Berlin Heidelberg, Berlin, Heidelberg, 2008).

[341] M. C. Urban, G. Bocedi, A. P. Hendry, J.-B. Mihoub, G. Peer, A. Singer, J. R. Bridle, L. G.

Crozier, L. De Meester, W. Godsoe, A. Gonzalez, J. J. Hellmann, R. D. Holt, A. Huth, K. Johst, C. B. Krug, P. W. Leadley, S. C. F. Palmer, J. H. Pantel, A. Schmitz, P. A. Zollner, J. M. J. Travis, Improving the forecast for biodiversity under climate change. *Science*. **353**, aad8466 (2016).

[342] M. S. Clark, U. Sommer, J. K. Sihra, M. A. S. Thorne, S. A. Morley, M. King, M. R. Viant, L. S. Peck, Biodiversity in marine invertebrate responses to acute warming revealed by a comparative multi-omics approach. *Glob. Change Biol.* **23**, 318–330 (2017).

[343] J. G. Carter, R. Seed, in *Bivalves: an Eon of Evolution*, P. A. Johnston, J. W. Haggart, Eds. (University of Calgary Press, Vancouver, 1998), pp. 87–117.

[344] L. S. Peck, A cold limit to adaptation in the sea. *Trends Ecol. Evol.* **31**, 13–26 (2016).

[345] H. A. Lowenstam, Factors affecting the aragonite:calcite ratios in carbonate-secreting marine organisms. *J. Geol.* **62**, 284–322 (1954).

[346] L. Ramajo, A. B. Rodriguez-Navarro, C. M. Duarte, M. A. Lardies, N. A. Lagos, Shifts in shell mineralogy and metabolism of *Concholepas concholepas* juveniles along the Chilean coast. *Mar. Freshw. Res.* **66**, 1147–1157 (2015).

[347] V. Tunnicliffe, K. T. A. Davies, D. A. Butterfield, R. W. Embley, J. M. Rose, W. W. Chadwick Jr, Survival of mussels in extremely acidic waters on a submarine volcano. *Nat. Geosci.* **2**, 344–348 (2009).

[348] L. Meire, J. Mortensen, P. Meire, T. Juul-Pedersen, M. K. Sejr, S. Rysgaard, R. Nygaard, P. Huybrechts, F. J. R. Meysman, Marine-terminating glaciers sustain high productivity in Greenland fjords. *Glob. Change Biol.* **23**, 5344–5357 (2017).

[349] U. Gräwe, R. Friedland, H. Burchard, The future of the western Baltic Sea: two possible scenarios. *Ocean Dyn.* **63**, 901–921 (2013).

[350] S. Dutkiewicz, J. J. Morris, M. J. Follows, J. Scott, O. Levitan, S. T. Dyhrman, I. Berman-Frank, Impact of ocean acidification on the structure of future phytoplankton communities. *Nat. Clim. Change*. **5**, 1002–1006 (2015).

[351] J. García Molinos, B. S. Halpern, D. S. Schoeman, C. J. Brown, W. Kiessling, P. J. Moore, J. M. Pandolfi, E. S. Poloczanska, A. J. Richardson, M. T. Burrows, Climate velocity and the future global redistribution of marine biodiversity. *Nat. Clim. Change*. **6**, 83–88 (2016).

[352] B. R. Scheffers, L. De Meester, T. C. L. Bridge, A. A. Hoffmann, J. M. Pandolfi, R. T. Corlett, S. H. M. Butchart, P. Pearce-Kelly, K. M. Kovacs, D. Dudgeon, M. Pacifici, C. Rondinini, W. B. Foden, T. G. Martin, C. Mora, D. Bickford, J. E. M. Watson, The broad footprint of climate change from genes to biomes to people. *Science*. **354**, aaf7671 (2016).

[353] A. M. Queirós, J. A. Fernandes, S. Faulwetter, J. Nunes, S. P. S. Rastrick, N. Mieszkowska, Y. Artioli, A. Yool, P. Calosi, C. Arvanitidis, H. S. Findlay, M. Barange, W. W. L. Cheung, S. Widdicombe, Scaling up experimental ocean acidification and warming research: from individuals to the ecosystem. *Glob. Change Biol.* **21**, 130–143 (2015).

[354]   J. T. Wootton, C. A. Pfister, J. D. Forester, Dynamic patterns and ecological impacts of declining ocean pH in a high-resolution multi-year dataset. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 18848–18853 (2008).

[355]   B. Gaylord, K. J. Kroeker, J. M. Sunday, K. M. Anderson, J. P. Barry, N. E. Brown, S. D. Connell, S. Dupont, K. E. Fabricius, J. M. Hall-Spencer, T. Klinger, M. Milazzo, P. L. Munday, B. D. Russell, E. Sanford, S. J. Schreiber, V. Thiyagarajan, M. L. H. Vaughan, S. Widdicombe, C. D. G. Harley, Ocean acidification through the lens of ecological theory. *Ecology*. **96**, 3–15 (2015).

[356]   N. Ockendon, D. J. Baker, J. A. Carr, E. C. White, R. E. A. Almond, T. Amano, E. Bertram, R. B. Bradbury, C. Bradley, S. H. M. Butchart, N. Doswald, W. Foden, D. J. C. Gill, R. E. Green, W. J. Sutherland, E. V. J. Tanner, J. W. Pearce-Higgins, Mechanisms underpinning climatic impacts on natural populations: altered species interactions are more important than direct effects. *Glob. Change Biol.* **20**, 2221–2229 (2014).

[357]   M. Loreau, C. de Mazancourt, Biodiversity and ecosystem stability: a synthesis of underlying mechanisms. *Ecol. Lett.* **16**, 106–115 (2013).

[358]   S. J. McNaughton, Diversity and stability of ecological communities: a comment on the role of empiricism in ecology. *Am. Nat.* **111**, 515–525 (1977).

[359]   T. E. Reed, V. Grotan, S. Jenouvrier, B.-E. Saether, M. E. Visser, Population growth in a wild bird is buffered against phenological mismatch. *Science*. **340**, 488–491 (2013).

[360]   E. L. Cross, L. S. Peck, E. M. Harper, Ocean acidification does not impact shell growth or repair of the Antarctic brachiopod *Liothyrella uva* (Broderip, 1833). *J. Exp. Mar. Biol. Ecol.* **462**, 29–35 (2015).

[361]   G. N. Somero, The physiology of climate change: how potentials for acclimatization and genetic adaptation will determine "winners" and "losers." *J. Exp. Biol.* **213**, 912–920 (2010).

[362]   L. S. Peck, Organisms and responses to environmental change. *Mar. Genom.* **4**, 237–243 (2011).

[363]   C. D. G. Harley, A. Randall Hughes, K. M. Hultgren, B. G. Miner, C. J. B. Sorte, C. S. Thornber, L. F. Rodriguez, L. Tomanek, S. L. Williams, The impacts of climate change in coastal marine systems. *Ecol. Lett.* **9**, 228–241 (2006).

[364]   J. Speybroeck, D. Bonte, W. Courtens, T. Gheskiere, P. Grootaert, J.-P. Maelfait, S. Provoost, K. Sabbe, E. W. M. Stienen, V. Van Lancker, W. Van Landuyt, M. Vincx, S. Degraer, The Belgian sandy beach ecosystem: a review. *Mar. Ecol.* **29**, 171–185 (2008).

[365]   T. Warmose, T. Backeljau, L. De Bruyn, The littorinid fauna of the Belgian coast (Mollusca, Gastropoda). *Bull. L'institut R. Des Sci. Nat. Belgique*. **58**, 51–70 (1988).

[366]   M. Becuwe, Het voorkomen van de Steenloper, *Arenaria interpres*, en de Paarse Strandloper, *Calidris maritima*, in België en Zeeuws-Vlaanderen (Nederland). *Le Gerfaut*. **61**, 175–223 (1971).

[367]   ICES, Dataset on Ocean Hydrography. The International Council for the Exploration of the Sea,

Copenhagen (2014), (available at http://www.ices.dk/marine-data/dataset-collections/Pages/default.aspx).

[368] IMERS, Integrated Marine Environmental Readings and Samples. Flanders Marine Institute. (2018), (available at http://www.vliz.be/vmdcdata/imers/imers.php).

[369] D. E. Sadler, A. J. Lemasson, A. M. Knights, The effects of elevated CO_2 on shell properties and susceptibility to predation in mussels *Mytilus edulis*. *Mar. Environ. Res.* **139**, 162–168 (2018).

[370] I. Kröncke, H. Reiss, J. D. Eggleton, J. Aldridge, M. J. N. Bergman, S. Cochrane, J. A. Craeymeersch, S. Degraer, N. Desroy, J.-M. Dewarumez, G. C. A. Duineveld, K. Essink, H. Hillewaert, M. S. S. Lavaleye, A. Moll, S. Nehring, R. Newell, E. Oug, T. Pohlmann, E. Rachor, M. Robertson, H. Rumohr, M. Schratzberger, R. Smith, E. Vanden Berghe, J. van Dalfsen, G. van Hoey, M. Vincx, W. Willems, H. L. Rees, Changes in North Sea macrofauna communities and species distribution between 1986 and 2000. *Estuar. Coast Shelf Sci.* **94**, 1–15 (2011).

[371] H. Reiss, K. Meybohm, I. Kröncke, Cold winter effects on benthic macrofauna communities in near- and offshore regions of the North Sea. *Helgol. Mar. Res.* **60**, 224–238 (2006).

[372] D. Basford, A. Eleftheriou, The benthic environment of the North Sea (56° to 61°N). *J. Mar. Biol. Assoc. UK*. **68**, 125–141 (1988).

[373] A. Eleftheriou, D. J. Basford, The macrobenthic infauna of the offshore northern North Sea. *J. Mar. Biol. Assoc. UK*. **69**, 123–143 (1989).

[374] D. J. Basford, A. Eleftheriou, I. M. Davies, G. Irion, T. Soltwedel, The ICES North Sea benthos survey: the sedimentary environment. *ICES J. Mar. Sci.* **50**, 71–80 (1993).

[375] J. A. Craeymeersch, C. H. R. Heip, J. Buijs, "Atlas of the North Sea Benthic Infauna, ICES Cooperative Research Report 218" (1997).

[376] K. Ghertsos, C. Luczak, J.-M. Dewarumez, J.-C. Dauvin, Influence of spatial scales of observation on temporal change in diversity and trophic structure of fine-sand communities from the English Channel and the southern North Sea. *ICES J. Mar. Sci.* **57**, 1481–1487 (2000).

[377] C. Heip, J. A. Craeymeersch, Benthic community structures in the North Sea. *Helgoländer Meeresuntersuchungen*. **49**, 313–328 (1995).

[378] I. Kröncke, C. Bergfeld, North sea benthos: a review. *Senckenbergiana maritima*. **33**, 205–268 (2003).

[379] A. Kunitzer, D. Basford, J. A. Craeymeersch, J. M. Dewarumez, J. Dorjes, G. C. A. Duineveld, A. Eleftheriou, C. Heip, P. Herman, P. Kingston, U. Niermann, E. Rachor, H. Rumohr, P. A. J. de Wilde, The benthic infauna of the North Sea: species distribution and assemblages. *ICES J. Mar. Sci.* **49**, 127–143 (1992).

[380] J. A. Lindley, G. Beaugrand, C. Luczak, J.-M. Dewarumez, R. R. Kirby, Warm-water decapods and the trophic amplification of climate in the North Sea. *Biol. Lett.* **6**, 773–776 (2010).

[381] H. Reiss, G. Wieking, I. Kroencke, Microphytobenthos of the Dogger Bank: a comparison

between shallow and deep areas using phytopigment composition of the sediment. *Mar. Biol.* **150**, 1061–1071 (2007).

[382] S. D. Batten, R. Clark, J. Flinkman, G. Hays, E. John, A. W. . John, T. Jonas, J. A. Lindley, D. P. Stevens, A. Walne, CPR sampling: the technical background, materials and methods, consistency and comparability. *Prog. Oceanogr.* **58**, 193–215 (2003).

[383] D. De Groote, *Aantalsverloop, Verspreiding en Gedrag Vanwatervogels in en Rond de Ijzermonding te Nieuwpoort Inrelatie tot het Getij. M.Sc. thesis.*, University of Gent (2003).

[384] H. Engledow, G. Spanoghe, A. Volckaert, E. Coppejans, S. Degraer, M. Vincx, M. Hoffmann, "Onderzoek naar de Fysischekarakterisatie en de Biodiversiteit van Strandhoofden en Andereharde Constructies langs de Belgische Kust." (Gent, 2001).

[385] J. Seys, J. VanWaeyenberge, K. Devos, P. Meire, E. Kuijken, The recent expansion of breeding gulls along the Belgian North Sea coast. *Sula*. **12**, 209–216 (1998).

[386] G. Spanoghe, *Aantallen en Verspreiding, gedrag en Habi-Tatkeuze van Meeuwen (Laridae) aan de Vlaamse Kust in Hetwinterhalfjaar. M.Sc. thesis*, University of Ghent (1999).

[387] E. Stienen, J. Van Waeyenberge, J. . Vercruijsse, Zilvermeeuw *Larus argentatus* en Kleine Mantelmeeuw *Larus fuscus* als broedvogels in Vlaanderen. *Nattur Oriolus*. **68**, 104–110 (2002).

[388] V. Stuer, *Trofische Interacties in Relatie tot Zwinnen Vande Belgische kust (Schipgatduinen Koksijde): Epibenthos Enavifauna. M.Sc. thesis*, University of Gent (2002).

[389] G. Vermeersch, A. Anselin, K. Devos, "Bijzondere Broedvogels in Vlaanderen in de periode 1994-2005" (Brussels, 2007).

[390] G. Vermeersch, A. Anselin, "Broedvogels in Vlaanderen in 2006-2007" (Brussels, 2009).

[391] OSPAR, "Background Document for *Nucella lapillus* (Dog whelk) Biodiversity Series" (2009).

[392] F. Kerckhof, Over het verdwijnen van de Purperslak *Nucella lapillus* (Linnaeus, 1758) langs de Belgische kust. *De Strandvlo*. **8**, 82–85 (1988).

[393] G. Marra, S. N. Wood, Practical variable selection for generalized additive models. *Comput. Stat. Data Anal.* **55**, 2372–2387 (2011).

[394] S. Nakagawa, P. C. D. Johnson, H. Schielzeth, The coefficient of determination R^2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *J. R. Soc. Interface*. **14**, 20170213 (2017).

[395] N. Gypens, C. Lancelot, A. V. Borges, Carbon dynamics and CO_2 air-sea exchanges in the eutrophied coastal waters of the Southern Bight of the North Sea: a modelling study. *Biogeosciences*. **1**, 147–157 (2004).

[396] C. Lancelot, N. Gypens, G. Billen, J. Garnier, V. Roubeix, Testing an integrated river–ocean mathematical tool for linking marine eutrophication to land use: the *Phaeocystis*-dominated Belgian coastal zone (Southern North Sea) over the past 50 years. *J. Mar. Syst.* **64**, 216–228 (2007).

[397] S. Mangan, M. A. Urbina, H. S. Findlay, R. W. Wilson, C. Lewis, Fluctuating seawater pH/pCO$_2$

regimes are more energetically expensive than static pH/pCO$_2$ levels in the mussel *Mytilus edulis*. *Proc. R. Soc. B Biol. Sci.* **284**, 20171642 (2017).

[398] J. A. Lindley, G. Beaugrand, C. Luczak, J.-M. Dewarumez, R. R. Kirby, Warm-water decapods and the trophic amplification of climate in the North Sea. *Biol. Lett.* **6**, 773–776 (2010).

[399] J. Lindley, R. Kirby, Climate-induced changes in the North Sea Decapoda over the last 60 years. *Clim. Res.* **42**, 257–264 (2010).

[400] K. T. Frank, Trophic cascades in a formerly cod-dominated ecosystem. *Science.* **308**, 1621–1623 (2005).

[401] S. C. Votier, R. W. Furness, S. Bearhop, J. E. Crane, R. W. G. Caldow, P. Catry, K. Ensor, K. C. Hamer, A. V. Hudson, E. Kalmbach, N. I. Klomp, S. Pfeiffer, R. A. Phillips, I. Prieto, D. R. Thompson, Changes in fisheries discard rates and seabird communities. *Nature.* **427**, 727–730 (2004).

[402] ICES, "Report of the ICES Advisory Committee on Fishery Management and Advisory Committee on Ecosystems, 2004. ICES Advice. Volume 1, Number 2." (Copenhagen, 2004).

[403] S. K. Spence, G. W. Bryan, P. E. Gibbs, D. Masters, L. Morris, S. J. Hawkins, Effects of TBT contamination on *Nucella* populations. *Funct. Ecol.* **4**, 425–432 (1990).

[404] S. Wanless, M. Harris, P. Redman, J. Speakman, Low energy values of fish as a probable cause of a major seabird breeding failure in the North Sea. *Mar. Ecol. Prog. Ser.* **294**, 1–8 (2005).

[405] S. Garthe, K. Camphuysen, R. Furness, Amounts of discards by commercial fisheries and their significance as food for seabirds in the North Sea. *Mar. Ecol. Prog. Ser.* **136**, 1–11 (1996).

[406] G. Nehls, I. Hertzler, G. Scheiffarth, Stable mussel *Mytilus edulis* beds in the Wadden Sea - They're just for the birds. *Helgoländer Meeresuntersuchungen.* **51**, 361–372 (1997).

[407] A. R. Palmer, Calcification in marine molluscs: how costly is it? *Proc. Natl. Acad. Sci. U.S.A.* **89**, 1379–1382 (1992).

[408] D. L. J. Vendrami, L. Telesca, H. Weigand, M. Weiss, K. Fawcett, K. Lehman, M. S. Clark, F. Leese, C. McMinn, H. Moore, J. I. Hoffman, RAD sequencing resolves fine-scale population structure in a benthic invertebrate: implications for understanding phenotypic plasticity. *R. Soc. Open Sci.* **4**, 160548 (2017).

[409] F. Cribari-Neto, A. Zeileis, Beta regression in R. *J. Stat. Softw.* **34**, 1–24 (2010).

[410] J. Fox, S. Weisberg, *An R Companion to Applied Regression* (SAGE Publications, Thousand Oaks, CA, USA, 2011).

[411] S. Wood, F. Scheipl, gamm4: Generalized Additive Mixed Models using "mgcv" and "lme4" (2017), (available at https://cran.r-project.org/web/packages/gamm4/).

[412] H. Wickham, *Ggplot2 : Elegant Graphics for Data Analysis* (Springer, New York, NY, USA, 2016).

[413] A. Magnusson, H. Skaug, A. Nielsen, C. Berg, K. Kristensen, M. Maechler, K. van Bentham, B. Bolker, M. Brooks, glmmTMB: Generalized Linear Mixed Models using Template Model

Builder (2017), , doi:10.1101/132753.

[414] A. Loy, S. Steele, lmeresampler: bootstrap methods for nested linear mixed-effects models. (2016), (available at https://cran.r-project.org/package=lmeresampler).

[415] A. Kuznetsova, P. B. Brockhoff, R. H. B. Christensen, lmerTest package: tests in linear mixed effects models. *J. Stat. Softw.* **82**, 1–26 (2017).

[416] K. Barton, MuMIn: Multi-Model Inference. (2017), (available at https://cran.r-project.org/package=MuMIn).

[417] J. Pinheiro, D. Bates, S. DebRoy, D. Sarkar, R Core Team, nlme: linear and nonlinear mixed effects models (2017), (available at https://cran.r-project.org/package=nlme).

[418] F. Scheipl, S. Greven, H. Küchenhoff, Size and power of tests for a zero random effect variance or polynomial regression in additive and linear mixed models. *Comput. Stat. Data Anal.* **52**, 3283–3299 (2008).

[419] G. Simpson, tsgam: Utilities for Working with GAMs Fitted to Time Series (2018), (available at https://github.com/gavinsimpson/tsgam).

[420] T. Kijewski, J. W. M. Wijsman, H. Hummel, R. Wenne, Genetic composition of cultured and wild mussels *Mytilus* from The Netherlands and transfers from Ireland and Great Britain. *Aquaculture*. **287**, 292–296 (2009).

[421] H. Tyler-Walters, B. James, M. Carruthers, C. Wilding, O. Durkin, C. Lacey, E. Philpott, L. Adamas, P. D. Chaniotis, P. T. V. Wilkes, R. Seeley, M. Neilly, J. Dargie, O. T. Crawford-Avis, "Descriptions of Scottish Priority Marine Features (PMFs). Scottish Natural Heritage Commissioned Report No. 406." (2016).

[422] T. Yarra, K. Gharbi, M. Blaxter, L. S. Peck, M. S. Clark, Characterization of the mantle transcriptome in bivalves: *Pecten maximus*, *Mytilus edulis* and *Crassostrea gigas*. *Mar. Genom.* **27**, 9–15 (2016).

[423] S. Ferson, F. J. Rohlf, R. K. Koehn, Measuring shape variation of two-dimensional outlines. *Syst. Zool.* **34**, 59–68 (1985).

Appendix

A Supporting figures and tables

Supporting material for Chapter 2 and 3.

Table A.1 R packages list
Packages used for exploration, analysis and plotting. Package name, details on the version and bibliographic references are provided.

Package	Complete name	V	Year	Source	Use
betareg	Beta Regression	3.1	2016	[409]	Beta regression (GLM with beta distribution)
car	Companion to Applied Regression	3.0-0	2018	[410]	Type III ANOVA table
gamm4	Generalized Additive Mixed Models using 'mgcv' and 'lme4'	0.2-5	2017	[411]	Generalised additive (mixed) models (GA(M)Ms) using an lme4 method
ggplot2	Create Elegant Data Visualisations Using the Grammar of Graphics	2.2.1	2016	[412]	Graphing
glmTMB	Generalized Linear Mixed Models using Template Model Builder	0.2.0	2017	[413]	Mixed-effect models (GLMMs) using the template model builder
lme4	Linear Mixed-Effects Models using 'Eigen' and S4	1.1-17	2018	[290]	Mixed-effect models (GLMMs)
lmeresampler	Bootstrap Methods for Nested Linear Mixed-Effects Models	0.1.0	2016	[414]	Parametric bootstrap for nlme model
lmerTest	Tests in Linear Mixed Effects Models	3.0-1	2018	[415]	Wald approximated confidence intervals
mgcv	Mixed GAM Computation Vehicle with Automatic Smoothness Estimation	1.8-23	2018	[288]	Generalised additive (mixed) models (GA(M)Ms)
MuMIn	Multi-Model Inference	1.40.4	2018	[416]	Pseudo-R-squared for Generalised Mixed-Effect models
nlme	Linear and Nonlinear Mixed Effects Models	3.1-137	2017	[417]	Mixed-effect models (GLMMs), variance structure (GLS), spatial correlation
pbkrtest	Parametric Bootstrap and Kenward Roger Based Methods for	0.4-7	2017	[322]	Kenward-Roger approximations for degrees of freedom and parametric

	Mixed Model Comparison				bootstrap for model comparisons
RLRsim	Exact (Restricted) Likelihood Ratio Tests for Mixed and Additive Models	3.1-3	2016	[418]	Random effect simulation
tsgam	Utilities for Working with GAMs Fitted to Time Series	0.0-4	2018	[419]	Derivative, simultaneous confidence interval and simulations estimation with GAMMs

Supporting material for Chapter 4.

Table A.2 Origin and details of mussel specimens

Information on collection areas and taxonomic status of the *Mytilus* populations used for the EFA of outlines. For each group, the study system, sampling site (code as in Figure 4.1) geographical location, samples size (*n*), site coordinates (longitude and latitude), genotypic status (Me: *Mytilus edulis*; Mt: *M. trossulus*; Mg: *M. galloprovincialis* and hybrids), and reference to taxonomic status and/or previous use of the studied populations are reported.

Site	Location	*n*	Longitude	Latitude	Status of mussels	Reference to taxonomic status and/or previously published use
	System 1					
1	Exmouth (England, UK)	30	3°25'44.8"W	50°37'19.2"N	Me	[144]
2	Oostende (Belgium)	30	2°54'15.6"E	51°13'49.4"N	Me	[131, 420]
3	Texel (Netherlands)	30	4°47'46.2"E	53°0'19.3"N	94%Me, 6%Me×Mg	[131, 420]
4	Menai Bridge (Wales, UK)	30	4°15'20.8"W	53°10'22.7"N	Me	[420]
5	Tarbet, Kintyre (Scotland, UK)	30	5°24'40.8"W	55°51'56.0"N	90%Med, 1%Mt, 6%Me×Mg, 3%Me×Mt	[158]
6	St. Andrews (Scotland, UK)	30	2°46'53.3"W	56°20'22.6"N	Me	[421], MPA monitoring
7	Kristineberg (Sweden)	30	11°25'19.5"E	58°12'37.8"N	94%Me, 6%Me×Mt	[128, 131, 133, 148, 149, 339]
8	Tromsø (Norway)	31	19°8'36.8"E	70°4'28.8"N	Me	[133, 136, 148]
9	Upernavik (Greenland)	28	56°6'10.1"W	72°47'38.0"N	33% Me, 51%Mt, 14%Med×Mtr	[136]

10	Qaanaaq (Greenland)	28	69°14'25.1"W	77°27'54.0"N	Mt	[136]
	System 2					
A	Sylt (Germany)	28	8°26'09.4"E	55°01'32.0"N	Med	[128, 143, 148, 149, 339]
B	Kiel (Germany)	25	10°08'56.0"E	54°19'45.1"N	68% Med, 32% Me×Mt	[128, 149, 339]
C	Ahrenshoop (Germany)	30	12°25'37.0"E	54°23'12.7"N	80% Med, 20% Me×Mt	[128, 131, 149]
D	Usedom (Germany)	28	14°00'39.7"E	54°03'20.2"N	37% Med, 10% Mt, 53% Me×Mt	[149, 150]
E	Nynäshamn (Sweden)	27	17°55'45.4"E	58°52'38.5"N	Mt	[128, 131, 148, 149, 339]
	System 3					
-	Loch Leven (Scotland, UK)	120	5°01'41.19"W	56°42'35.1"N	90% Med, 1% Mt, 6% Me×Mg, 3% Me×Mt	[158]

Table A.3 PCs contribution to the shell shape

Proportion of shape variance captured by individual shape variables and description of their contributions to the shell features and mean shape reconstruction for each study system (Figure 4.5, Figure 4.6).

PC	%variance	Contribution to the shell shape
System 1		
PC1	38.7	Variation in shell height, ligament angle, ventral margin shape, and shell width: high values corresponded to elongated and narrow shells, with acute ligament angles and convex ventral margins; low values to round and wide shells, with big ligament angles and concave ventral margins.
PC2	30	Variation in the shape of ventral margin, ligament angle and shell height, with small contribution to the symmetry of ventral profile: low values corresponded to curved shells, with concave ventral margins, small ligament angles and symmetric ventral profiles; high values corresponded to round shells, with big ligament angles, convex ventral margins and less symmetric ventral profiles.
PC3	10.7	Variation in the shape of umbo, ligament and ventral margin: low values corresponded to "curved" shells with concave ventral margins and ligaments, and an umbo oriented towards the ventral side; high values described elliptical shells with convex ventral margins and ligaments, and an umbo oriented toward the anterior side.
PC4	5.6	Variation in the shape of ventral margin and differences between elliptical and "curved" shells, with a concave ventral margin.
PC5	3.7	Small variations in dorsoventral shape (more or less parallel margins) and the symmetry of the ventral view outlines.

System 2

PC1	37.2	Variation in shell height, convexity of the ventral margin and shell width, and differences between elongated and narrow shells, and round and wide specimens with flat ventral margins.
PC2	32.9	Variation in dorsal margin, ligament area and shell width: low values corresponded to large and narrow shells, with a round dorsoposterior commissure; high values indicated elongated and wide shells, with more parallel dorsoventral margins.
PC3	11.8	Variation in the shape of ventral and dorsal margins, and differences between "curved" and elliptical shells.
PC4	5.4	Variation in ligament length, ventral and dorsal margin shape, showing a progressive increase of ventral margin convexity and ligament length for higher values.
PC5	3.4	Variations in the shape of the dorsoventral margins (more or less parallel) and ligament length.

System 3

PC1	45.1	Variation in shell height, width and ligament: low values corresponded to elongated and narrow shells, and acute ligament angles; high values were associated with round and wide shells, with a big height and ligament angle.
PC2	21	Variation in shell height and width: increasing values corresponded to increasing shell height and decreasing width.
PC3	15.2	Variation in ligament area, dorsoposterior and ventral margins shape: low values corresponded to a big ligament angle; high values indicated increasing roundness of posterior margins and a more acute ligament angle.
PC4	7.1	Variation in the shape of ventral margin and differences between "curved" shells with concave ventral margins and elliptical shells.
PC5	3	Variation in dorsoposterior margins shape and the symmetry of ventral view.

Atlantic system

PC1	38.1	Variation of shell height, the shape of ventral and dorsal margins, ligament length and shell width: low values corresponded to elliptical shells, with concave ventral margins, long ligaments and narrow profiles, while high values were associated with "curved" and wide shells, with convex ventral margins and short ligaments.
PC2	32.1	Variation of shell height, ligament angle and width, describing a gradual transition from round and wide shells to elongated and narrow mussels for increasing values.
PC3	11.6	Variations in the shape of ventral margin, umbo and ligament with small variations of shell width: negative values corresponded to shells with concave ventral margins, convex ligaments and an umbo oriented toward the anterior side, while positive values to more "curved" shells with concave ventral margins, convex ligaments and an umbo pointing downwards.
PC4	5.1	Contribution to the shape of ventral margin and variability between "curved" (concave ventral margins) and elliptical shells.
PC5	3	Small variations in dorsoventral shape (more or less parallel margins) and the symmetry of ventral view.

Table A.4 GAMM summary results for individual smooth and linear terms (System 1, 2 and 3)
Estimated degrees of freedom, F statistics and significance values for each term from the interactions between environmental covariates and PCs are reported for individual study system.

	PC	*f*(Temperature) × PCs			*f*(Salinity) × PCs			Chl-*a* × PCs			Length × PCs		
		edf	F	*p*-value	*edf*	F	*p*-value	*df*	F	*p*-value	*df*	F	*p*-value
System 1	PC1	1.98	23.61	**<0.0001**	1.99	40.05	**<0.0001**	1	0.11	0.74	1	4.27	0.039
	PC2	1	0.99	0.32	1	19.32	**<0.0001**	1	6.20	**0.013**	1	33.70	**<0.0001**
	PC3	1.82	3.92	**0.013**	1.82	4.83	0.045	1	1.73	0.19	1	1.99	0.16
	PC4	1	0.05	0.83	1	0.70	0.40	1	5.52	0.019	1	1.43	0.23
	PC5	1.81	3.16	0.103	1	0.12	0.73	1	5.08	0.024	1	0.88	0.35
		f(Temperature) × PCs			*f*(Salinity) × PCs			Chl-*a* × PCs					
System 2	PC1	1	4.39	0.04	1.99	53.66	**<0.0001**	1	3.51	0.062			
	PC2	1.98	31.33	**<0.0001**	1	91.21	**<0.0001**	1	3.03	0.082			
	PC3	1	0.001	0.97	1	4.18	0.041	1	0.17	0.68			
	PC4	1	1.17	0.28	1	1.80	0.18	1	0.001	0.97			
	PC5	1.97	18.98	**<0.0001**	1.86	10.98	**0.00061**	1	2.61	0.11			
		f(enviro-PC1) × PCs											
System 3	PC1	1.97	32.23	**<0.0001**									
	PC2	1.72	7.52	**0.012**									
	PC3	1.48	1.82	0.10									
	PC4	1.85	2.59	0.048									
	PC5	1.94	15.12	**<0.0001**									

Supporting material for Chapter 5.
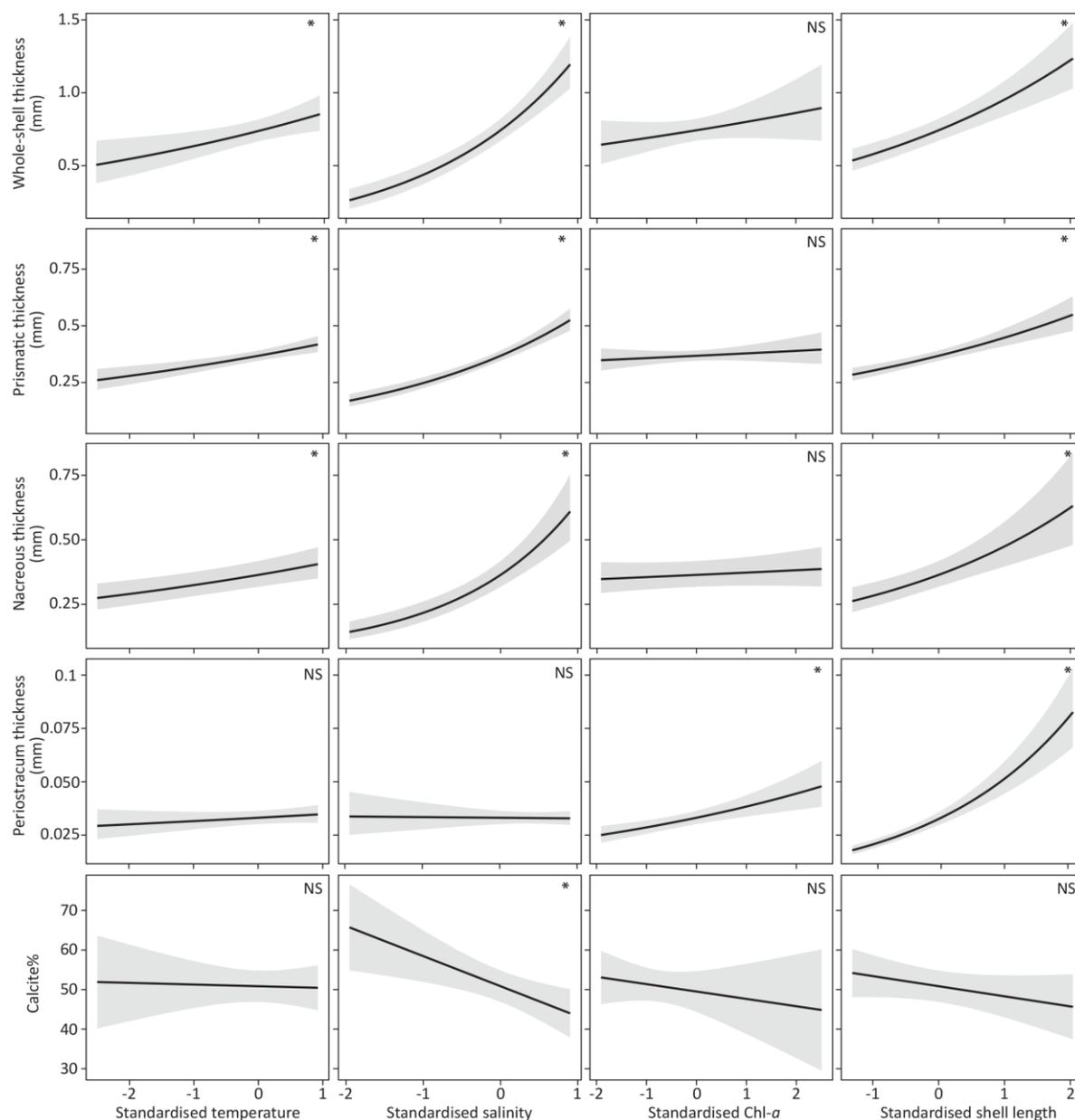
Figure A.1 Relationships between shell layers and modelled predictors
Predicted relationships between the whole-shell, prismatic layer, nacreous layer, periostracum thickness, and the calcite% with standardised water temperature [mean (SD) = 13.03°C (4.32)], salinity [mean (SD) = 25.52 psu (10.29)], Chl-*a* concentration [mean (SD) = 2.48 mg m^{-3} (1.41)], and shell length [mean (SD) = 47.42 mm (16.20)]. Predicted values (continuous lines) and confidence intervals (shaded areas) across the range of each predictor were estimated while controlling statistically for the other covariates (mean values). (NS $p > 0.05$, * $p < 0.05$)

Table A.5 Provenience and taxonomic status of the *Mytilus* populations used for the study

For each sampling site (site codes as in Figure 5.1a), geographic location, samples size (*n*), site coordinates (longitude and latitude), genotypic status [proportion of *Mytilus edulis* (Me), *M. trossulus* (Mt), *M. galloprovincialis* (Mg) and hybrids], and reference and/or previous use of the studied populations are reported.

Site	Location	*n*	Longitude	Latitude	Status of mussels	Reference to taxonomic status and/or previous use of the shells
1	Brest (France)	25	4°22'9,75"W	48°20'16,81"N	96%Me, 4%Me×Mg	[131, 143]
2	Exmouth (England, UK)	26	3°25'44.8"W	50°37'19.2"N	Me	[144]
3	Oostende (Belgium)	25	2°54'15.6"E	51°13'49.4"N	Me	[131, 420]
4	Texel (Netherlands)	25	4°47'46.2"E	53°0'19.3"N	94%Me, 6%Me×Mg	[131, 420]
5	Usedom (Germany)	25	14°00'39.7"E	54°03'20.2"N	37%Me, 10%Mt, 53%Me×Mt	[149, 150]
6	Kiel (Germany)	25	10°08'56.0"E	54°19'45.1"N	68%Me, 32%Me×Mt	[128, 149, 339]
7	Ahrenshoop (Germany)	25	12°25'37.0"E	54°23'12.7"N	80%Me, 20%Me×Mt	[128, 131, 149]
8	Sylt (Germany)	25	8°26'09.4"E	55°01'32.0"N	Me	[128, 143, 148, 149, 339]
9	Kerteminde (Denmark)	25	10°40'04.7"E	55°27'04.5"N	57%Me, 5%Mt, 38%Me×Mt	[149, 150]
10	Tarbet, Kintyre (Scotland, UK)	25	5°24'40.8"W	55°51'56.0"N	Me	[422]
11	St. Andrews (Scotland, UK)	25	2°46'53.3"W	56°20'22.6"N	Me	[421], MPA monitoring
12	Kristineberg (Sweden)	24	11°25'19.5"E	58°12'37.8"N	94%Me, 6%Me×Mt	[128, 131, 133, 148, 149, 339]
13	Nynäshamn (Sweden)	25	17°55'45.4"E	58°52'38.5"N	Mt	[128, 131, 148, 149, 339]
14	Trondheim (Norway)	25	10°26'52.4"E	63°27'25.9"N	70%Me, 5%Mt, 25%Me×Mt, 5%Me×Mg	[133, 148]
15	Tromsø (Norway)	24	19°8'36.8"E	70°4'28.8"N	Me	[133, 136, 148]
16	Upernavik (Greenland)	25	56°6'10.1"W	72°47'38.0"N	33%Me, 51%Mt, 14%Me×Mt	[136]
17	Qaanaaq (Greenland)	25	69°14'25.1"W	77°27'54.0"N	Mt	[136]

Supporting material for Chapter 6.

Table A.6 GAMMs summary statistics of shell characteristic with environmental regimes
    M. edulis shell thickness (different layers) variation with environmental regimes (temperature and salinity: Temp-PC1/PC2, Sal-PC1/PC2) and shell shape (EFA: Shape-PC1/PC2) is reported.

Parameter	*edf*	Estimate	SE	F	*p*-value	Random effects	SD
Whole-shell							
(Intercept)	1	6.579	0.040	164.34	<0.0001	Site	0.135
Shape-PC2	1	45.166	19.636	5.29	0.022	Residual	0.189
Temp-PC1	2.9	_	_	5.32	0.002		
Temp-PC2	1	0.076	0.027	5.29	0.0054		
Periostracum							
(Intercept)	1	50.631	1.048	48.31	<0.0001	Site	2.311
Temp-PC2	2.12	_	_	4.23	0.008	Residual	13.011
Sal-PC1	3.74	_	_	7.95	0.00017		
Sal-PC2	2.2	_	_	3.36	0.045		
Prismatic layer							
(Intercept)	1	5.973	0.015	406.51	<0.0001	Site	0.027
Temp-PC1	2.4	_	_	26.08	<0.0001	Residual	0.198
Temp-PC2	1	0.066	0.011	34.81	<0.0001		
Sal-PC2	2.4	_	_	3.88	0.0093		
Length	1	0.008	0.003	8.07	0.0048		
Nacreous layer							
(Intercept)	1	5.778	0.046	125.70	<0.0001	Site	0.145
Shape-PC1	1	62.541	20.965	8.88	0.0031	Residual	0.337
Shape-PC2	1	84.089	34.939	5.81	0.017		
Temp-PC2	3.77	_	_	4.70	0.0021		
Sal-PC2	1	-0.096	0.038	6.20	0.013		
Calcite%							
(Intercept)	1	0.542	0.011	50.04	<0.0001	Site	0.033
Shape-PC1	1	-26.611	5.516	23.23	<0.0001	Residual	0.089
Sal-PC1	1	0.013	0.005	6.40	0.028		
Length	1	0.004	0.001	10.11	0.0020		

B Environmental datasets

List of the datasets used for the calculation of mean annual values of environmental descriptors. Water temperature, salinity and chlorophyll-*a* concentrations. This study has been conducted using the Copernicus Marine Service Products: COPERNICUS - Marine Environment Monitoring System (http://marine.copernicus.eu/).

DATASET #1

Product identifier:	**GLOBAL_ANALYSIS_FORECAST_PHY_001_024**
Link (last accessed on 11-09-2018):	http://marine.copernicus.eu/services-portfolio/access-to-products/?option=com_csw&view=details&product_id=GLOBAL_ANALYSIS_FORECAST_PHY_001_024
Short description:	The Operational Mercator global Ocean analysis and forecast system at 1/12 degree is providing 7 days of 3D global ocean forecasts updated daily and ocean analysis updated weekly. The time series start on January 1st 2013 and is aggregated in time in order to reach a two full years' time series sliding window. This product includes daily mean files of temperature, salinity, currents, sea level and ice parameters from the top to the bottom of the Ocean over the Global Ocean. It also includes 2-hourly mean surface fields for temperature, currents and sea level.
Spatial resolution:	0.08 degree
Vertical coverage:	from -5500.0 m to 0.0 m
Temporal resolution:	Daily mean, hourly mean
Update frequency:	Daily
Production unit:	GLO-MERCATOR-TOULOUSE-FR

DATASET #2

Product identifier:	**NORTHWESTSHELF_ANALYSIS_FORECAST_BIO_004_002_b**
Link (last accessed on 11-09-2018)	http://marine.copernicus.eu/services-portfolio/access-to-products/?option=com_csw&view=details&product_id=NORTHWESTSHELF_ANALYSIS_FORECAST_BIO_004_002_b
Short description:	The Forecasting Ocean Assimilation Model, Atlantic Margin model (FOAM AMM7) is a coupled hydrodynamic-ecosystem model, nested in a series of one-way nests to the Met Office global ocean model. The hydrodynamics are supplied by the Nucleus for European Modelling of the Ocean (NEMO) with the 3DVar

NEMOVAR system used for the assimilation of sea surface temperature data. This is coupled to the European Regional Seas Ecosystem Model (ERSEM), developed at Plymouth Marine Laboratory (PML). ERSEM based models have been used operationally to forecast biogeochemistry in the region for a number of years.

Spatial resolution: 0.11 x 0.7 degree

Vertical coverage: from -5000 m to 0 m

Temporal resolution: Daily mean

Update frequency: Daily

Production unit: NWS-METOFFICE-EXETER-UK

DATASET #3

Product identifier: **GLOBAL_REANALYSIS_BIO_001_018**

Link (last accessed on 11-09-2018) :   http://marine.copernicus.eu/services-portfolio/access-to-products/?option=com_csw&view=details&product_id=GLOBAL_REANALYSIS_BIO_001_018

Short description: Biogeochemistry simulation over period 1998 - 2015. Outputs are delivered as monthly mean files with .netcdf format (CF/COARDS 1.5 convention) on the native tri-polar grid (ORCA025) at ¼° resolution with 75 vertical levels. This simulation is based on the PISCES biogeochemical model. It is forced offline at a daily frequency by the equivalent of the GLOBAL-REANALYSIS-PHYS-001-009 physics product.

Spatial resolution: 0.11 degree

Vertical coverage: from -5500.0 m to 0.0 m

Temporal resolution: Daily mean

Update frequency: Daily

Production unit: GLO-MERCATOR-TOULOUSE-FR

DATASET #4

Product identifier: **NORTHWESTSHELF_REANALYSIS_BIO_004_011**

Link (last accessed on 11-09-2018) :   http://marine.copernicus.eu/services-portfolio/access-to products/?option=com_csw&view=details&product_id=NORTHWESTSHELF_REANALYSIS_BIO_004_011

Short description: This is a hydrodynamic model of the North West European shelf forced at the surface by ERA-interim winds, atmospheric temperature, and precipitation fluxes. Horizontal boundary

conditions were provided by the NOC global reanalysis prior to 1989 and by the GloSea reanalysis thereafter. Boundary conditions in the Baltic Sea came from the IOM-GETM model and CMEMS-BALTICSEA ANALYSIS_FORECAST_PHYS_003_006. E-Hype data were used for river inputs. Hydrodynamic calculations were performed by the Nucleus for European Modelling of the Ocean (NEMO) system, while the 3DVar NEMOVAR system was used for the assimilation of sea surface temperature data. Physical outputs are provided both as monthly means and as daily 25 hours, edited, averages. The reanalysis was conducted in four sections.

Spatial resolution:	0.11 x 0.7 degree
Vertical coverage:	from -5000.0 m to 0.0 m
Temporal resolution:	Daily mean
Update frequency:	Daily
Production unit:	NWS-METOFFICE-EXETER-UK

DATASET #5

Product identifier:	**BALTICSEA_ANALYSIS_FORECAST_PHYS_003_006**
Link (last accessed on 11-09-2018) :	http://marine.copernicus.eu/services-portfolio/access-to-products/?option=com_csw&view=details&product_id=BALTICSEA_ANALYSIS_FORECAST_PHY_003_006

Short description: This Baltic Sea physical model product provides forecasts for the physical conditions in the Baltic Sea. The Baltic forecast is updated twice daily providing a new two days forecast with hourly data for sea level variations, ice concentration and thickness at the surface, and temperature, salinity and horizontal velocities for the 3D field. The product is based on the 3D ocean model code HBM developed within the Baltic ocean community.

Spatial resolution:	2 km
Vertical coverage:	from -5500.0 m to 0.0 m
Temporal resolution:	Daily mean, hourly mean
Update frequency:	Daily
Production unit:	GLO-MERCATOR-TOULOUSE-FR

DATASET #6

Product identifier:	**BALTICSEA_ANALYSIS_FORECAST_BIO_003_007**

Link (last accessed on 11-09-2018):	http://marine.copernicus.eu/services-portfolio/access-to-products/?option=com_csw&view=details&product_id=BALTICSEA_ANALYSIS_FORECAST_BIO_003_007
Short description:	This Baltic Sea biogeochemical model product provides forecasts for the biogeochemical conditions in the Baltic Sea. The Baltic forecast is updated twice daily providing a new two days forecast with hourly data for the parameters dissolved oxygen, nitrate, phosphate, Chl-a. The product is produced by the biogeochemical model ERGOM one way coupled to the Baltic 3D ocean model HBM, which provides the CMEMS Baltic physical ocean forecast product.
Spatial resolution:	2 km
Vertical coverage:	from -400 m to 0 m
Temporal resolution	Daily mean, hourly instantaneous
Update frequency:	Daily
Production unit:	BAL-DMI-COPENHAGEN-DK