

Simulation of the Model Organism

Dosidicus gigas

Analysis of the Implications of an
Energy Driven Life History

Jörg Höhne

Dissertation to obtain the degree
Doctor rerum naturalium (Dr. rer. nat.)

Submitted to the Doctoral Committee of the
Faculty of Biology and Chemistry
of the University of Bremen in September 2021

Reviewers: **Dr. Hauke Reuter**
Leibniz Centre for Tropical Marine Research
Bremen, Germany

Prof. Dr. Matthias Wolff
Leibniz Centre for Tropical Marine Research
Bremen, Germany

Advisory Committee

Reviewers: **Dr. Hauke Reuter**
Leibniz Centre for Tropical Marine Research
Bremen, Germany

Prof. Dr. Matthias Wolff
Leibniz Centre for Tropical Marine Research
Bremen, Germany

Examiners: **Prof. Dr. Marko Rohlf**
University of Bremen Bremen, Germany
Bremen, Germany

Dr. Broder Breckling
University of Bremen Bremen, Germany
Bremen, Germany

Disputation colloquium on April 25, 2022

To the beauty of complexity

Contents

Abstract	xxi
Zusammenfassung	xxiii
1. Preface	1
1.1. Motivation, hypotheses and methods	1
1.2. Simulation model techniques applied in this thesis	5
1.3. Outline of the thesis structure	8
2. Objects of Study	11
2.1. <i>Dosidicus gigas</i> (<i>D. gigas</i>)	11
2.1.1. About the species, its importance and its basic characteristics . .	11
2.1.2. Distribution range, range extension and habitat	16
2.1.3. Food web	20
2.1.4. Life history	22
2.1.5. Intraspecific population structure	26
2.1.6. The suitability of <i>Dosidicus gigas</i> as a model organism for study	30
2.2. The functional triad migration-maturation-growth	31
2.3. Transferring mechanistic models to simulation models	33
3. Elements of an energy driven life history for <i>D. gigas</i>	35
3.1. Introduction	35
3.2. Developing the energy driven life history model (EDLHM)	36
3.2.1. The functional triad migration-maturation-growth	36
3.2.2. The reformulation at the individual level	37
3.2.3. The effects at population level	40
3.3. The energy driving elements at the individual level	41
3.4. The driving effects at school, cohort and population level in the literature	43
3.4.1. Decoupling energy demands of males and females by sex-specific maturation	43
3.4.2. Size-at-maturity energy optimization	43
3.4.3. The sex ratio change for energy optimization at population level	45

3.4.4. Cannibalism	46
4. Modeling the individual and its environment	49
4.1. Introduction	49
4.2. Modeling a <i>D. gigas</i> individual and its environment	51
4.2.1. Overview	51
4.2.2. Modeling a <i>D. gigas</i> individual	51
4.2.3. Modeling the environment	56
4.2.4. Modeling complex behavior: Prey seeking and swarming	58
4.3. Modeling basic traits	59
4.3.1. Overview	59
4.3.2. Age, growth, size and resource access	60
4.3.3. Size derived traits: Mass and surface	63
4.4. The energy model of the agent	66
5. Identifying spawning areas of <i>D. gigas</i>	73
5.1. Overview	73
5.2. Discussion of probable multiple spawning grounds	73
5.3. Previous identification of likely spawning areas	77
5.4. Predicting paralarvae presence probability by sea surface temperature data	81
5.4.1. Introduction	81
5.4.2. Computing the pp_{prob} by the averaged SST	84
5.4.3. Days of suitable SST for spawning development	90
5.4.4. Summary	92
5.5. Analyzing egg and paralarvae transport and dispersion	92
5.6. Computing paralarvae origin from attractors	101
5.7. Identifying spawning areas: Results and conclusions	105
6. Individual level traits and their computation	107
6.1. Introduction	107
6.2. The unisex growth function and dimorphic terminal size growth function	108
6.3. The energy aspect of the EDLHM	111
6.3.1. Energy needs, energy buffering capacity and shifts during <i>D. gi-</i> <i>gas</i> ' lifespan	111
6.3.2. Growing under energy constraints	114
6.3.3. Oxygen as a potentially limiting factor	120
6.3.4. The effects of energy constraints support the fTMMG	121

6.4.	Different preconditions as an explanatory factor to different size-at-maturity	123
6.4.1.	Overview	123
6.4.2.	Size differences in early ontogenetic stages may amplify exponentially	123
6.4.3.	Multiple spawning batches	125
6.5.	Summary	127
7.	Cannibalism as a possible survival strategy	129
7.1.	Overview	129
7.2.	Theoretical aspects of cannibalism	131
7.2.1.	Energy uptake as the driving force for cannibalism performing	131
7.2.2.	Classification criteria in respect of <i>D. gigas</i>	132
7.2.3.	<i>D. gigas</i> ' behavior in the context of cannibalistic effects	132
7.2.4.	Cannibalism as an evolutionary stable strategy	134
7.3.	The size disparity as a prerequisite for cannibalism	135
7.3.1.	Overview	135
7.3.2.	The size disparity and its effects	136
7.3.3.	The behavior of the DTSGF in respect of cannibalism	137
7.4.	Energy uptake by cannibalism	149
7.4.1.	Overview	149
7.4.2.	Analysis and discussion	150
7.5.	Conclusions on prerequisites on size disparity and energy uptake	153
7.6.	Simulation and exploration of cannibalism by an energetic view	154
7.6.1.	Model overview	154
7.6.2.	Comparison of cannibalism strategies and the growth functions	158
7.6.3.	Analysis of the cannibalism strategies	159
7.6.4.	Changing the start age and the age variation	165
7.6.5.	Varying the cannibalism energy threshold	168
7.6.6.	The validity of the model	170
7.7.	Survival of the females	172
7.8.	Discussion of an obsolete window of cannibalism	175
7.9.	Discussion of the results and conclusions	176
8.	Properties at the population level and their computation	183
8.1.	Introduction	183
8.2.	Reducing energy peak demand by asynchronous formation of female and male reproductive tissue	183

Contents

8.3. Energy optimization by smaller mantle length and higher generational turnover	189
8.4. Multiple spawning peaks to reduce peak energy demand	192
8.5. Optimization of energy use by sex ratio change	194
8.6. Discussion and conclusion	196
9. The simulation of ecological models: Enhanced approaches and techniques	199
9.1. Introduction	199
9.2. Individual- and agent-based-modeling	200
9.3. Agent-based-modeling Implementation: Structures, Methods and Techniques	202
9.3.1. General Systems Theory and agent-based-modeling	202
9.3.2. ABM implementation techniques	203
9.3.3. The cellular automaton, its universality and its equivalency to individual-based modeling	205
9.4. Enhanced techniques to implement ABM	206
9.4.1. Introduction	206
9.4.2. Essential object types for a spatio-temporally organized ABM	208
9.4.3. Object oriented programming as a supporting programming technique	209
9.4.4. The State-based modeling and programming paradigm	210
9.4.5. Ecological design patterns	216
9.5. Short comings and the path to a solution	218
9.5.1. Overview	218
9.5.2. Problems with the model definition	219
9.5.3. Implications of the lack of a common language	221
9.5.4. Desired features of ABM development	221
9.6. Simulation environments	222
9.6.1. Introduction	222
9.6.2. Implementation of simulation models in simulation environments	223
9.6.3. Discussion of available simulation environments	224
9.7. Frameworks in combination with a programming language	227
9.7.1. Introduction and overview	227
9.7.2. Frameworks and libraries	227
9.7.3. Choosing the simulation environment and programming language	229

9.8. An advanced approach to model to simulation model translation	230
9.8.1. Introduction	230
9.8.2. Model description using literate programming	231
9.8.3. The translation process	234
9.8.4. LP in the future practice	235
9.9. Summary and conclusions	238
10. Summary and outlook	241
10.1. Summary	241
10.2. Reflection	250
10.3. Outlook	253
A. Python scripts to compute SST-data and paralarvae distribution	283
A.1. Overview	283
A.2. Supporting script files	284
A.3. Preparation of directory structure	284
A.4. Display of the distribution range according to FAO	285
A.5. <i>D. gigas</i> sightings	285
A.6. Preprocessing of the SST-data	286
A.7. SST processing	287
A.8. Plotting the paralarvae sample counts	287
A.9. Computation and visualization of the egg and paralarvae dispersion . . .	288
B. Python scripts for cannibalism data evaluation	291
B.1. Overview	291
B.1.1. The @parameter-section	291
B.1.2. The @dataColumns-section	291
B.1.3. The @data-section	291
B.2. The scripts	291
B.2.1. Utility scripts	291
B.2.2. Preprocessing the files	292
B.2.3. MATHEMATICA processing	292
C. Mathematica-Scripts	293
C.1. GrowthFunctionsAndRelations.nb	293
C.1.1. Overview	293
C.1.2. Comments/Init	293

Contents

- C.1.3. Constants 293
- C.1.4. Basic functions and definition 294
- C.1.5. Growth, volume, surface and mass functions 294
- C.1.6. Modeling the individual: Growth 294
- C.1.7. Modeling the individual: The energy model 294
- C.1.8. Traits on individual level 294
- C.1.9. Cannibalism 294
- C.1.10. Properties on school and population level 295
- C.1.11. Objects of Study 295
- C.2. CannibalismEvaluation.nb 295

- D. Discrete event simulation with the Scala programming language 297**
- D.1. Overview 297
- D.2. The simulation model abstraction toolkit 297
 - D.2.1. General structure of a simulation model 297
 - D.2.2. The structure of the simulation model abstraction toolkit 298
- D.3. The simulation model 299
 - D.3.1. Package apps.common 299
 - D.3.2. Package apps.cannibalism 300
- D.4. Simulation details 301
 - D.4.1. Simulation parameters 301
 - D.4.2. Implementation details of the simulation model 301

- E. The simulation of ecological models: Enhanced approaches and techniques 305**
- E.1. Execution of finite state machines 305
- E.2. Simulation environments 306
 - E.2.1. NETLOGO 306
 - E.2.2. SeSAm 308
- E.3. Frameworks in combination with a programming language 310
 - E.3.1. Programming languages 310
 - E.3.2. The advantages of SCALA if compared to JAVA 312

List of Figures

2.1.	<i>D. gigas</i> drawing, Pfeffer 1912 in R. E. Young and Veccione (2013) . . .	11
2.2.	Annual landings in metric tons (FAO, 2021).	12
2.3.	Still images from video sequence showing feeding of <i>D. gigas</i>	14
2.4.	Anatomical images of <i>D. gigas</i>	15
2.5.	Current distribution of <i>D. gigas</i> (FAO, 2021).	17
2.6.	Referenced single occurrence of <i>D. gigas</i>	18
2.7.	Prey item relationship between <i>D. gigas</i> and other species.	21
3.1.	The functional triad migration-maturation-growth (fTMMG).	36
3.2.	Outline of energy driven life history model (EDLHM).	39
4.1.	The cause–effect diagram of the EDLHM for <i>D. gigas</i>	49
4.2.	The cause–effect diagram of the simulation model for a EDLHM of <i>D. gigas</i>	50
4.3.	Components of individual energy needs.	51
4.4.	Influences on the growth rate, locomotion, maturation rate and maturation level.	52
4.5.	A sigmoid mapping function for size development and daily growth rate.	60
4.6.	Segmentation of <i>D. gigas</i> to enable computation of its body mass development.	63
4.7.	Comparison of body mass estimating functions.	65
4.8.	The top-down hierarchy of the individual energy model and related parameters.	68
5.1.	Field sampling locations in Staaf et al. (2010a, 2010b).	76
5.2.	Plots of \log_{10} -transformed bongo and manta tow sampling counts of <i>Sthenoteuthis oualaniensis</i> and <i>Dosidicus gigas</i> complex (SD-complex) paralarvae.	79
5.3.	Proposed spawning areas of the SD-complex.	80
5.4.	SST and <i>D. gigas</i> pp_{prob} averaged over the years 2011–2015.	84
5.5.	Average SST in June to September.	85
5.6.	<i>D. gigas</i> pp_{prob} for June to September computed from SST data.	86
5.7.	Average SST in October to January.	87

List of Figures

5.8. <i>D. gigas</i> pp_{prob} for October to January computed from SST data.	88
5.9. SD-complex paralarvae log 10-transformed tow sampling counts overlaid on estimated pp_{prob}	89
5.10. Areas suitable for successful embryonic development.	91
5.11. Field sampling locations of ethanol preserved <i>D. gigas</i> paralarvae samples.	92
5.12. Material flow affecting the focal cell.	94
5.13. Surface current velocity from January 2011 to December 2016.	95
5.14. The initial setup for material flow computation of date 2013-05-28.	96
5.15. Emerging attractors and donators for egg/paralarvae dispersion in the northern hemisphere for the year 2013.	98
5.16. Emerging attractors and donators for egg/paralarvae dispersion in the southern hemisphere for the year 2012.	99
5.17. Locations of occurrences of <i>D. gigas</i> eggs and paralarvae used for back- tracking simulations.	101
5.18. Possible spawning areas of paralarvae samples identified by backtracking material transport flows over the 30 days prior to sampling.	103
5.19. Backtracking of dispersion of location 1.43°N/113.72°W for close dates.	104
6.1. Maximum ML plotted against initial daily growth rate assuming the logistic growth function.	109
6.2. The one year size development and daily growth rate for the DTSGF.	109
6.3. Daily absolute size and mass change using the DTSGF.	111
6.4. The sex-specific daily energy needs.	112
6.5. The sex-specific daily energy needs in relation to body mass.	113
6.6. Energy needs of males and females, specified by the growth function DTSGF over the course of the lifespan.	114
6.7. The effective relative body mass gain under different energy deprivation scenarios.	116
6.8. Sex-specific energy parameters of the and DTSGF in relation to ML.	117
6.9. Effects on body mass in relation to ML under different energy deprivation scenarios.	118
6.10. Volume and surface area in relation to ML.	120
6.11. Volume and surface ratio in relation to ML.	122
6.12. The effect of hatching sizes on ML_{terminal}	124
6.13. The time needed for smaller hatchlings to grow to 0.0013m.	124
6.14. The DTSGF ML development when spawning takes place over 20 days.	126

6.15. DTSGF ML differences development when spawning takes place over 20 days. 126

7.1. DTSGF: The male:female ratio plotted against relative age for different ML_{terminal} tuples. 138

7.2. DTSGF: Values of $\text{reqWOC}_{\text{UB}}$ for all female and male combinations. . . 139

7.3. DTSGF: Contour plot of values of $\text{reqWOC}_{\text{UB}}$ for all female and male combinations. 140

7.4. DTSGF: Values of $\Delta\text{reqWOC}_{\text{UB}}$ for females. 141

7.5. DTSGF: Values of $\Delta\text{reqWOC}_{\text{UB}}$ for males. 142

7.6. DTSGF: Values of $\text{reqWOC}_{\text{UB}}$ for all female and male combinations of cannibalism. 145

7.7. DTSGF: Contour plot of values of $\text{reqWOC}_{\text{UB}}$ for all female and male combinations of cannibalism. 146

7.8. DTSGF: Values of $\text{reqWOC}_{\text{UB}}$ for all female and male combinations of cannibalism 30d to 120d. 147

7.9. DTSGF: $\Delta\text{reqWOC}_{\text{UB}}$, showing differences between $\text{reqWOC}_{\text{UB}}$ values for cannibalization of individuals of the same sex and the same older generation and the minimum $\text{reqWOC}_{\text{UB}}$ values for alternative combinations of sex and age. 148

7.10. DTSGF: The number of cannibals whose energy requirements are met by the mass of the victim for all female and male combinations of cannibalism. 150

7.11. DTSGF: Contour plot of the number of cannibals whose energy requirements are met by the mass of the victim for all female and male combinations of cannibalism. 151

7.12. Simulation results of changes in school size for different cannibalism strategies and growth functions. 158

7.13. UGF: Changes in school size development with no and active cannibalism, partial view. 160

7.14. UGF, active cannibalism strategy: Changes in numbers of individuals and conspecifics available for cannibalism in a school. 161

7.15. Comparison between UGF and DTSGF growth models of simulated performance of cannibalism strategies. 162

7.16. UGF: Change in school size over time with active cannibalism. 165

7.17. UGF: Change in school size over time with passive cannibalism. 166

List of Figures

7.18. UGF: Changes in numbers of individuals and availability of conspecifics under an active cannibalism strategy in a school over time, with varying <i>start age variation</i>	166
7.19. UGF: Change in school size over time with active cannibalism and varying <i>start age variation</i>	167
7.20. Simulated changes over time in school size for different cannibalism strategies and growth functions.	168
7.21. UGF: Averaged changes in school size over time for active cannibalism with varied $threshold_{cann}$	169
7.22. UGF: Changes in school size over time for passive cannibalism with varied $threshold_{cann}$	170
7.23. UGF: Changes in the proportion of females over time for different cannibalism strategies.	173
7.24. DTSGF: Changes in the proportion of females over time for different cannibalism strategies.	174
7.25. The two-dimensional window of predation (WOP ²) as a function of prey item's relative size and intensity of its defensive response.	176
7.26. DTSGF: Changes in the value of $reqWOC_{UB}$ for female on female and male on male cannibalism, for an age difference $\omega=4$ d between predator and prey.	178
8.1. Schematic representation of the flattening of peak energy demand caused by time offset between maturation of females and males.	185
8.2. Schematic representation of how longer time offsets between the onset of female and male maturity reduce peak energy demand.	185
8.3. DTSGF: Mass development of females and males for different $ML_{terminal}$ over the lifespan.	186
8.4. DTSGF: Daily $total_{mr}$ of females and males over the lifespan.	187
8.5. DTSGF: Ratio of $growth_{mr}$ to $basal+locomotion_{mr}$ over the lifespan.	188
8.6. DTSGF: Accumulated daily $total_{mr}$ over a one-year lifespan for different $ML_{terminal}$	190
8.7. DTSGF: Accumulated $total_{mr}$ for female $ML_{terminal}=0.34$ m, male $ML_{terminal}=0.28$ m and different lifespans.	191
8.8. Schematic representation of multiple spawning events with shorter and long spawning offsets.	193

8.9.	The DTSGF estimated accumulated energy demand of a female and male for different ML_{terminal} over a one-year lifespan.	195
8.10.	The DTSGF estimated female:male ratio of daily and accumulated energy demand over a one-year lifespan.	196
9.1.	Graphical description of the firefly behavior.	212
9.2.	Development of a simulation model from a source model (a) following the standard procedure (left) and (b) using literate programming. Both methods incorporate the iterative refinement the simulation model until maximal correspondence with the source model is achieved.	232
E.1.	The interfaces of NETLOGO to access the model and an exemplary run of a contributed model.	307
E.2.	Full screen of a SeSAm-model.	308
E.3.	SeSAm activity (action) for taking food by a modeled ant.	309

List of Tables

2.1. Annual landings of <i>D. gigas</i> in metric tons.	13
2.2. Taxonomic classification of <i>D. gigas</i>	13
2.3. Ontogenesis of <i>D. gigas</i> adapted from C. Nigmatullin et al. (2001).	22
2.4. Estimated lifespan of <i>D. gigas</i> in studies by different authors.	24
2.5. Characteristics and geographical distribution of <i>D. gigas</i> size groups.	27
2.6. Size group classification by different authors.	28
5.1. Recorded occurrences of <i>D. gigas</i> eggs, paralarvae, juveniles and mature females.	75
5.2. Values of pp_{prob} of <i>D. gigas</i> for incremental values of SST.	82
5.3. Approximated locations for the backtracking of <i>D. gigas</i> eggs and paralarvae occurrence.	102
6.1. Values reported in the literature for lifespan, hatching size, ML_{terminal} and daily growth rate.	108
6.2. Unisex and dimorphic growth parameters.	110
7.1. Specified simulation parameters.	169
9.1. List of topics to be included in a (simulation) model in the most recent version of the ODDP (Grimm et al., 2010).	220
D.1. Parameter combinations controlling the cannibalism strategy.	300
D.2. Simulation model parameters.	301

Abbreviations

ABM agent-based-modeling.

ACS agent-based complex systems.

basal+locomotion_{mr} basal+locomotion metabolic rate.

basal_{mr} basal metabolic rate.

CA cellular automaton.

DES discrete-event simulation.

DFA deterministic finite automaton.

DSL domain specific language.

DTSGF dimorphic terminal size growth function.

DynLP dynamic literate programming.

EDLHM energy driven life history model.

EDP ecological design pattern.

eDSL ecological domain specific language.

EN El Niño.

event_{cann} cannibalism event.

event_{tee} terminal energy exhaustion event.

FAO Food and Agriculture Organization of the United Nations.

FP functional programming.

FSM finite state machine.

fTMMG functional triad migration-maturation-growth.

GIS geographic information system.

GOC Gulf of California.

growth_{mr} growth metabolic rate.

GST General Systems Theory.

IBM individual-based modeling.

IDE integrated developing environment.

Abbreviations

JVM Java Virtual Machine.

LN La Niña.

locomotion_{mr} locomotion metabolic rate.

LP literate programming.

ML mantle length.

ODDP Overview, Design concepts and Details Protocol.

OML oxygen minimum layer.

OO object oriented.

OOP object oriented programming.

OOSA object-oriented system analysis.

PDF portable document format.

POM pattern-oriented modeling.

pp_{prob} paralarvae presence probability.

reqWOC_{UB} required window of cannibalism upper bound.

SAM size-at-maturity.

SBM state-based modeling.

SD-complex *Sthenoteuthis oualaniensis* and *Dosidicus gigas* complex.

SeSAm shell for simulated agent system.

SMATK simulation model abstraction toolkit.

SST sea surface temperature.

total_{mr} total metabolic rate.

UGF unisex growth function.

UML Unified Modeling Language.

WOC window of cannibalism.

WOC_{LB} window of cannibalism lower bound.

WOC_{UB} window of cannibalism upper bound.

WOP window of predation.

WOP² two-dimensional window of predation.

Abstract

This thesis investigates the potential of computer modeling techniques to elucidate the life history and population-level features of *D. gigas*. Commonly known as giant squid or jumbo squid, *D. gigas* is endemic to the eastern Pacific Ocean. It is a fast-growing, opportunistic predator at the upper trophic level of the marine food web that also serves as a food source for other predators and is of economic importance for fisheries. *D. gigas* exhibits a capacity for rapid adaptation to environmental changes and could therefore serve as a model organism to study the impacts on marine species of changes to the marine ecosystem.

In addition to an extremely high metabolic rate, *D. gigas* has only a limited energetic buffer and individuals are therefore vulnerable to starvation when food resources are scarce. Moreover, as a semelparous species, loss of a single cohort could threaten the survival of the species. Despite these disadvantages, *D. gigas* has not only survived but has also recently expanded its geographic range. For this reason, adaptive mechanisms should exist to compensate for these disadvantageous traits. Proposed mechanisms include the existence of different size-at-maturity groups adapted to different environmental conditions, asynchronous gonadal development to reduce population-level peak energy demand, and cannibalism as a strategy to ensure the survival of populations during periods of food scarcity. Since *D. gigas* is difficult to observe in the field and does not survive under laboratory conditions, computer simulation is the most promising approach for this research into these and other hitherto unresolved questions about the physiology and life history of the species.

To this end, the thesis develops a radically simplified energy-driven life history model which is applied to investigate the following hypotheses:

1. The species *D. gigas* adheres to an energy driven life history that controls *D. gigas*'s phenotypic expression.
2. Cannibalism is an important survival strategy for *D. gigas* during periods of food scarcity.
3. The phenotypic extremes postulated in Keyl et al. (2008) serve to optimize survival and fitness under the (locally) prevailing energetic-environmental conditions.

Abstract

4. Methodological approaches currently used for ecological simulations require improvements in model specification, model formulation, model implementation and model documentation.

Mathematical analyses, individual- and agent-based simulation, cellular automata, and satellite-based environmental data are used to calculate the different stages in the life history of *D. gigas*, including the location of potential spawning grounds. The model calculations show that *D. gigas* adopts an energy-optimized life history and that observed phenotypical variation could be an adaptation to variations in food availability under different environmental conditions. Other assumptions, such as asynchronous formation of gonadal tissue, offer less potential for energetic optimization. Cannibalism was found to be effective in prolonging the survival of a school in times of food shortage. However, the most successful cannibalism strategy was found to be one where consumption of weakened conspecific formed part of the regular diet, rather than being a behavioral response to food shortages.

In addition to the creation of simulation models, the thesis critically examines the process used to develop computer simulation models from conceptual source models. This “translation” process is found to be prone to error. The thesis proposes an alternative procedure which avoids the shortcomings identified in currently available products and protocols and has the potential to become a new standard in ecological modeling.

Zusammenfassung

Diese Arbeit untersucht anhand des Modellorganismus *D. gigas* das Potenzial von Computermodellierung, um dessen Merkmale der life history sowie Eigenschaften auf Populationsebene zu untersuchen. *D. gigas*, allgemein als Humboldt-Kalmar bekannt, ist endemisch im östlichen Pazifik verbreitet und ein schnell wachsender, opportunistischer Räuber auf der oberen trophischen Ebene des marinen Nahrungsnetzes. Er dient gleichzeitig anderen Räubern als Nahrungsquelle und besitzt für die Fischerei ebenfalls eine ökonomische Bedeutung. Auf Umweltveränderungen zeigt *D. gigas* eine schnelle und hohe Anpassungsfähigkeit, weshalb er als Modellorganismus dienen könnte, um Auswirkungen von Veränderungen im marinen Ökosystem auf marine Arten zu untersuchen.

Neben seiner extrem hohen Stoffwechselrate verfügt *D. gigas* nur über einen begrenzten Energiepuffer, so dass die Individuen bei knappen Nahrungsressourcen anfällig für das Verhungern sind. Weil es sich um eine semelpare Spezies handelt, könnte der Verlust einer einzigen Kohorte zudem das Überleben der Art gefährden. Trotz dieser Nachteile hat *D. gigas* als Spezies nicht nur überlebt, sondern in jüngster Zeit auch sein geografisches Verbreitungsgebiet erweitert. Aus diesem Grund sollten Anpassungsmechanismen existieren, die diese nachteiligen Merkmale kompensieren. Zu den vorgeschlagenen Mechanismen gehören die Existenz verschiedener size-at-maturity-Gruppen, die an unterschiedliche Umweltbedingungen angepasst sind, eine asynchrone Entwicklung von gonadischem Gewebe zur Verringerung des Spitzenenergiebedarfs auf Populationsebene und Kannibalismus als Strategie zur Sicherung des Überlebens von Populationen in Zeiten der Nahrungsknappheit. Weil *D. gigas* in der Natur schwer zu beobachten ist und unter Laborbedingungen nicht überlebt, ist die Computersimulation der vielversprechendste Ansatz für die Erforschung dieser und anderer bisher ungelöster Fragen zur Physiologie und life history dieser Spezies.

Daher wird in dieser Arbeit ein radikal vereinfachtes energiegetriebenes life history-Modell entwickelt, das zur Untersuchung der folgenden Hypothesen herangezogen wird:

1. Die Spezies *D. gigas* folgt einer energiegetriebenen life history, welche die phänotypische Ausprägung von *D. gigas* steuert.

Zusammenfassung

2. In Zeiten der Nahrungsknappheit ist Kannibalismus für *D. gigas* eine wichtige Überlebensstrategie.
3. Die von Keyl et al. (2008) postulierten phänotypischen Extreme dienen der Optimierung von Überleben und Fitness unter den (lokal) vorherrschenden energetischen Umweltbedingungen.
4. Die derzeit für ökologische Simulationen verwendeten methodischen Ansätze erfordern Verbesserungen bei der Modellspezifikation, Modellformulierung, Modellimplementierung und Modelldokumentation.

Mit Hilfe mathematischer Analysen, individuen- und agentenbasierter Simulationen, zellulärer Automaten und satellitengestützter Umweltdaten werden die verschiedenen Stadien in der life history von *D. gigas* berechnet, einschließlich der Lage potenzieller Laichplätze. Die Modellrechnungen zeigen, dass *D. gigas* einer energieoptimierten life history unterliegt und dass die beobachtete phänotypische Variation eine Anpassung an die unterschiedliche Nahrungsverfügbarkeit aufgrund verschiedener Umweltbedingungen sein könnte. Andere Annahmen, wie die asynchrone Bildung von gonadischem Gewebe, bieten weniger Potenzial für eine Energieoptimierung. Kannibalismus verlängert das Überleben eines Schwarms in Zeiten der Nahrungsknappheit, jedoch war die erfolgreichste Kannibalismus-Strategie diejenige, bei der der Verzehr von geschwächten Artgenossen Teil der regulären Ernährung ist und nicht nur eine Verhaltensreaktion auf Nahrungsknappheit.

Neben der Erstellung von Simulationsmodellen beleuchtet diese Arbeit auch den Prozess der Entwicklung von computerbasierten Simulationsmodellen aus konzeptionellen Ausgangsmodellen kritisch. Dieser "Übersetzungsprozess" erweist sich als fehleranfällig, weshalb die Arbeit ein alternatives Verfahren vorschlägt, das die Mängel der derzeit verfügbaren Entwicklungswerkzeuge und Protokolle vermeidet und das Potenzial besitzt, ein neuer Standard in der ökologischen Modellierung zu werden.

1. Preface

1.1. Motivation, hypotheses and methods

Motivation

Rapid adaption to a changing environment, tolerance to temperature and oxygen level fluctuations, and a fierce opportunistic predatory behavior predestine *Dosidicus gigas* (*D. gigas*) to be a profiteer of the global climate change that is transforming the world's oceans and seas. The strong El Niño (EN) event in 1997/1998 demonstrated the capacity of *D. gigas* to take advantage of environmental change: since then *D. gigas* has extended its distribution range towards both poles (Field et al., 2007) and occupied new ecological niches (Keyl et al., 2008), while increased fishery landings suggest that population numbers have also increased, possibly to record levels (FAO, 2021; Keyl et al., 2008). As a highly abundant top-level predator that feeds on a wide range of prey and also serves as prey for other top-level predators, *D. gigas* interacts intensively with other elements of the food web, mediating a fast vertical energy transfer bridging several trophic levels between large-size marine mammals with low-trophic organism (Yu, Wen, et al., 2021).

In addition to its ecological importance, *D. gigas* is also of commercial importance (Yu, Chen, & Liu, 2021; Yu, Wen, et al., 2021), supporting a large cephalopod fishery on the one hand and predating on commercially fished species like hake on the other (Yu, Chen, & Liu, 2021; Yu, Wen, et al., 2021). Recent landings are at a high level, but these could collapse again, as happened during the strong 1982/1983 EN, most likely due to a population collapse (Arguelles et al., 2001).

The reasons for the range extension (Fang et al., 2018; Keyl et al., 2008) after the strong 1997/1998 EN event, when the landings dropped in 1998 to 18% of the landings of 1997, are unknown. Keyl et al. (2008) suggest the range expansion was enabled by a combination of reduced populations of the species' competitors and predators, partly caused by overfishing, and the occupation by *D. gigas* of ecological niches left vacant by other species adversely affected by environmental changes during the La Niña/El Niño period. These niches were "exceptionally accessible" for *D. gigas* following the return to cool-water regime nutrient rich waters (Keyl et al., 2008), not least because of its capacity for rapid population growth.

1. Preface

D. gigas displays extreme life history characteristics, i.e., rapid growth, a relatively short lifespan, high fertility, high energy demands (Bruno et al., 2021; Gilly et al., 2012; B. A. Seibel, 2011; Yatsu et al., 1999), and a growth-oriented metabolism with, as far as is known, limited energy buffering capability (Ibáñez & Keyl, 2010). These characteristics explain why *D. gigas* can profoundly influence the local environments through its voracious opportunistic feeding. The rapid growth of individuals with a wide window of predation (WOP) creates links between different trophic levels (Xavier et al., 2015) during their short period of development until maturity.

Rapid responses to environmental changes, e.g., changes in the reproductive pattern, may allow *D. gigas* to cope with fluctuations in environmental productivity (Rosa et al., 2013). These flexible life traits and *D. gigas*' sensitivity to temperature, oxygen level and ocean acidification (Anderson & Rodhouse, 2001; Argüelles & Tafur, 2010; Hu et al., 2019; Keyl et al., 2008; Tafur et al., 2010; J. W. Young et al., 2013), render this species a potential model organism for studying the effects of a changing environment on cephalopods and other marine organisms.

Despite *D. gigas*' high abundance offshore and near the coast, the species is difficult to observe in the wild and information is largely obtained from catch statistics (Markaida et al., 2005; Zeidberg & Robison, 2007; Zúñiga et al., 2008). Knowledge on many aspects of the species and its life history is still very limited and the object of ongoing research (Bazzino et al., 2010; Gilly, Markaida, et al., 2006; Gilly, Elliger, et al., 2006; Ibáñez et al., 2015; Li et al., 2017; Liu et al., 2020; Mejía-Rebollo et al., 2008; Staaf, 2010; W. Xu et al., 2019; Yatsu et al., 1999; Yu et al., 2015; Zepeda-Benitez et al., 2014).

At the population level, the existence of three different size-at-maturity (SAM)-groups is widely accepted, based on observational data. C. Nigmatullin et al. (2001) considered these groups to be genetically differentiated and to represent emerging species (in status nascendi). But the genetic origin of these groups is not supported by the weak genetic differences among them found in later studies (Ibáñez et al., 2011; Sanchez et al., 2020; Sandoval-Castellanos et al., 2009, 2010; Staaf et al., 2010b). Therefore, an alternative explanation is needed for this population level phenotypic plasticity.

Keyl et al. (2008) propose a functional triad migration-maturation-growth (fTMMG) as an explanatory approach whereby energy is the principal driver of *D. gigas*' life history trajectory. The fTMMG embeds the life history of *D. gigas* in a systemic vision: elements of the model are connected by cause-effect relations, whereby energetic flows give rise to physiological phenomena. To the knowledge of the author of this thesis,

the fTMMG represents the first and to date only attempt to explain the life history of *D. gigas* using a deterministic model.

The fTMMG was developed on the basis of analyzed field data and is a qualitative model, but the underlying cause–effect relations are susceptible to quantitative analysis by ecological simulation models informed by concepts from theoretical ecology. Placing the fTMMG on a quantitative basis would allow testing of its hypotheses on the life history of *D. gigas* using rigorous rigid mathematical methods and thereby identification of key drivers of a energy driven life history of *D. gigas* and the conditions under which these drivers come into play.

Baseline

D. gigas is a semelparous species and individuals die shortly after maturation and reproduction (Ibáñez & Keyl, 2010). In females, reproduction occurs soon after maturation (Tafur et al., 2010); so SAM presumably equates approximately to ML_{terminal} . Thus, for females, SAM is an important life history characteristic. The males mature approximately in the middle of their lifespan (Tafur et al., 2010); therefore male SAM corresponds less closely to ML_{terminal} .

The existence of a small, medium and large SAM-groups was first proposed in C. Nigmatullin et al. (2001) and is now generally accepted (Table 2.6). The groups are thought to be spatially distributed (Table 2.5), with the small SAM-group predominating in areas close to the equator, the medium SAM-group covering the entire species range except for the highest latitudes, and the large SAM-group occurring exclusively at higher latitudes (C. Nigmatullin et al., 2001). As mentioned above, genetic studies on *D. gigas* (Ibáñez et al., 2011; Sanchez et al., 2020; Sandoval-Castellanos et al., 2009, 2010; Staaf et al., 2010b) indicate that all three SAM-groups belong to single large population, comprising all groups. These studies find a weak genetic differentiation between individuals in the northern and southern hemispheres, which does not, however, explain the observed occurrence of SAM-groups within each hemisphere.

In the absence of evidence for the genetic origin of SAM-groups, Keyl (2009) proposes the fTMMG as an alternative explanation, whereby SAM-groups emerge as a functional response to environmental changes that is made possible by the phenotypic plasticity of *D. gigas*. Keyl (2009) also applies the fTMMG to explain the observed range expansion of *D. gigas* since the start of the present millennium, a topic which is not discussed in detail in this thesis. The fTMMG explains the emergence of SAM-groups as a response to two basic ecosystem conditions: One phenotypic extreme, i.e., the small SAM-group ensures population survival in warm periods characterized by low food availability,

1. Preface

while the other, i.e., the large SAM-group, maximizes the fitness of individuals in cool periods with high food availability.

The fTMMG essentially replaces the hypothesis of three spatially organized SAM-groups of genetic origin by the concept of a single population comprising two phenotypic extremes (small and large SAM) and their intermediate phenotypes, whose occurrence in all cases is determined by environmental conditions.

The focus on energy simplifies the view of this species since it relates phenotypical variation to environmental conditions, as a strategy for managing and/or optimizing the energy budget to ensure population survival and/or individual fitness. The high energy demand of *D. gigas* and its limited energy buffering capability (Ibáñez & Keyl, 2010; B. A. Seibel, 2007) naturally focus attention on energetic aspects of its life history, as well as the related optimization strategies discussed in this thesis, i.e., SAM-adaption, sex ratio change and cannibalism.

The original fTMMG provides an explanatory hypothesis, but this conceptual model is not susceptible to evaluation, without quantification of the embodied causal relationships. In this thesis, the fTMMG is extended to enable quantitative analysis of effects at the individual and population level reported in the literature.

Working hypotheses and methods

It is presumed that energy optimization is a crucial cornerstone in the life history of *D. gigas*. Without a careful balance between its energy demands and energy uptake, *D. gigas* individuals would probably die of terminal energy exhaustion, i.e., starvation during periods of food scarcity. Moreover, since *D. gigas* is a semelparous species, the simultaneous deaths of large numbers of individuals could threaten the survival of whole populations. Key characteristics of this species, and possibly other cephalopod species, such as the SAM or the presumed high cannibalism rate (Ibáñez & Keyl, 2010; Pecl & Jackson, 2008) may therefore be the expression of an energy driven life history (Pecl & Jackson, 2008). The pronounced expression of these traits in *D. gigas* facilitates elucidation of the main determining factors of its life history, which may also generate new insights into the life history of other species. The fTMMG initiated this approach by postulating cause–effect relations linking environmental drivers to physiological phenomena over the course of the lifespan. This thesis builds on the insights of the fTMMG to investigate the four following working hypotheses:

Hypothesis 1. *The species D. gigas adheres to an energy driven life history that controls D. gigas's phenotypic expression.*

1.2. Simulation model techniques applied in this thesis

Hypothesis 2. *Cannibalism is an important survival strategy for *D. gigas* during periods of food scarcity.*

Hypothesis 3. *The phenotypic extremes postulated in Keyl et al. (2008) serve to optimize survival and fitness under the (locally) prevailing energetic-environmental conditions.*

Hypothesis 4. *Methodological approaches currently used for ecological simulations require improvements in model specification, model formulation, model implementation and model documentation.*

The fTMMG was first formulated in Keyl et al. (2008). This thesis refers mainly to discussions and applications of the model in later published studies, specifically Argüelles et al. (2008), Tafur et al. (2010), Keyl et al. (2011), Ibáñez and Keyl (2010) and an unpublished article in Keyl (2009). The thesis develops a quantitative model that develops insights of the fTMMG to explain the physiological properties of *D. gigas*. It draws on an analysis of key features and properties of the model, as well as discussion and application of the model in the above-mentioned articles. The thesis examines both propositions focused on energetic drivers and the energy optimization, and those relating to emergent effects of these processes at a higher organizational level, i.e., at the population level, using a combination of mathematical analysis, individual-based modeling (IBM) and statistical analysis.

These approaches are applied to develop simulation models for *D. gigas* which apply the energy balance principle to explore possible energy management strategies and and their physiological manifestations at an individual, school, cohort and population level. This quantitative extension of the fTMMG creates a framework for rigorous mathematical and quantitative analysis of the hypotheses about *D. gigas* through the execution of simulation models and analysis of the results. The different simulation models developed vary in their degree of complexity and the simulation techniques employed.

1.2. Simulation model techniques applied in this thesis

Quantitative analysis of the qualitative relations in the fTMMG requires a multi-level approach to simulation. First, the physiology of *D. gigas* must be modeled at the individual level. For example, each individual hatchling exhibits a slightly different physiology, which may be amplified during development into distinct different phenotypic characteristics. Second, at the cohort level, local interactions of an individual with its

1. Preface

environment and other individuals can influence the development of the individual and of other individuals with which it interacts. These interactions and their outcomes are determined by the local distribution of individuals, other objects (such as prey items), and local environmental conditions. Third, at the population level, the distribution of large numbers of individuals over a range of environmental conditions that vary in space and time must also be modeled.

The exclusive use of (partial) differential functions for modeling and simulation is inadequate to meet these requirements. The use of such functions and equations in ecological modeling violates “two of the basic tenets of biology” (Huston et al., 1988). First, it ignores that each individual is different in behavior and physiology, assuming that all individuals, and aggregations of individuals, can be described using a single set of variable values. Second, the modeler presumes that, despite the inherently local and context-specific nature of interactions between individuals, each individual has the same effect on other individuals, with effects occurring only between pairs of individuals that experience direct contact, i.e., nearby individuals.

These limitations are overcome by individual-based modeling which is widely used in ecology to model interactions of individuals, aggregations of individuals (e.g., swarms) and environmental conditions (Breckling et al., 2006; Grimm & Railsback, 2005) and cellular automata (CA) (Margolus & Toffoli, 1987). An IBM focuses on individuals, whereby each individual is computed using behavior-describing rules, i.e., *algorithmic description*. The rules are (partially) applied at specific times during the simulation and, depending on the implementation, the rules may change, resulting in a highly dynamic model.

The use of IBM in biology took off in the early 1970s, when the required computational power became affordable (available) (Huston et al., 1988; Judson, 1994). Reuter et al. (2011) summarizes the advantages of IBMs for ecological simulation models as being the capacity “[...] to represent observation and knowledge in a form that is highly congruent with how we understand existing interaction.” In addition, an IBM embodies a low level of abstraction directly represents “elementary interactions that aggregate in the course of model execution in the same form as observable phenomena in empirical investigations.” IBMs are processed iteratively. Among other advantages, this allows logging of individuals and their states for “sampling” during simulation runs (Reuter et al., 2011). These features qualify IBMs as “a crucial tool in testing consistency of ecological knowledge”.

With reference to *D. gigas*, modeling of physiological and behavioral characteristics of individuals enables exploration of effects at school, cohort and population level,

1.2. Simulation model techniques applied in this thesis

generating synthetic knowledge that may help to fill gaps in existing knowledge of *D. gigas*, by comparing the results of theory-driven simulation models to data from field observations (see also the pattern-oriented modeling (POM) approach (Grimm & Railsback, 2005; S. F. Railsback & Grimm, 2012, 2019; S. F. Railsback & Harvey, 2020)).

The algorithmic rules describing the behavior of individuals are easier to transfer to a programming language than highly complex mathematical differential functions. This facilitates modeling and simulation and makes the computer models more accessible to non-specialists. It also allows the representation of highly heterogeneous individuals as types, whereby individuals of the same type start with different initial configurations, giving rise to highly heterogeneous interactions.

Furthermore, focusing on individual characteristics and behaviors enables more realistic calculation of possible effects of spatio-temporal heterogeneity in a population, compared to differential functions that can only describe the aggregated statistical behavior of large numbers of (identical) individuals.

To simulate heterogeneous environmental factors, the environment can be spatially organized, i.e., as a grid of cells, whereby each cell has a value or a state comprising values for multiple factors. Multiple layers may be defined in the grid, and the individuals can be programmed to access only subsets of these layers. If the environment is static, then fixed values are used to configure grid properties; the environment is then like a map. In case of a dynamic environment, the cells may either be continuously updated using precomputed values or computed according to a set of rules that determines the next state. The latter variant can equate to a cellular automaton (CA) (Subsection 9.3.3) and provides a powerful methodology for simulating the behavior of an environment and the (large number) of individuals it contains (Section 9.3).

Using a CA with an algorithmic description of rules to compute the next state of a cell also simplifies the modeling procedure, since it does not require the use of partial differential equations. Moreover external data, such as satellite temperature data, can be supplied to the cells as inputs for computation of the next state.

IBMs and CA are frequently used in ecological simulation models (Breckling et al., 2006; Breckling, Pe'er, et al., 2011; Grimm, 2018; Grimm et al., 2006), but there is potential for several improvements to current practice. Aspects where there is scope for improvement include the choice of programming language and library or simulation environment, the underlying programming structure of the model, model documentation, and model implementation, i.e., “translation” the conceptual source model into a computer simulation model.

1. Preface

In this thesis, the abbreviation IBM stands for individual-based modeling or individual-based model; the choice of meaning is made clear by the context, as it is for agent-based-modeling (ABM) and state-based modeling (SBM). The abbreviation CA stands for cellular automaton or cellular automata.

1.3. Outline of the thesis structure

The main objective of this thesis is the formulation of a testable quantitative model of an energy driven life history of *D. gigas*. The model is tested by comparing its outputs, for example in relation of energy use optimization strategies and their effects at individual, school and population level, with observational data reported in the literature. The fTMMG is used as a starting point; therefore the first step is the identification of quantifiable factors of the fTMMG, and the incorporation of these factors in a consistent conceptual model relates energy inputs to individual growth and behavior. This conceptual model is then implemented as a computer simulation model to explore the spatio-temporal effects of variations in energy availability on individuals and populations. A further goal of the thesis is to suggest, from the perspective of computer science, ways to improve the formulation and implementation of ecological simulation models.

To achieve these goals, Chapter 2 (“Objects of Study”) presents general aspects of the objects of interest, i.e., the species *Dosidicus gigas* and the fTMMG, as well as identifying imprecisions in the formulation of the fTMMG and its components which hinder quantitative verification of the model.

Chapter 3 (“Elements of an energy driven life history for *D. gigas*”) addresses these limitations. The basic relations of the fTMMG are extracted and reformulated into a new conceptual model, named energy driven life history model (EDLHM), which provides a structure for quantitative computation of an energy driven life history for *D. gigas*. The underlying methodology draws on General Systems Theory (GST) (Subsection 9.3.1), which decomposes the object of interest into a system of components and their interactions. It is known that the complexity of such a system increases rapidly with the number of components and understanding the behavior of the system becomes increasingly difficult. Therefore, a key aim of this chapter is to simplify the description of the hypothesized energy driven life history, whereby an individual optimizes its energy balance, in order to reduce the number of model components to a minimum.

Chapter 4 (“Modeling the individual and its environment”) specifies the EDLHM and sets out the framework for discussion of model components, energetic drivers and the implementation of the simulation model.

The life history of an individual *D. gigas* starts at the embryonic stage in the egg and thus with spawning by the female parent. Environmental conditions in the spawning grounds therefore determine the initial stages of development. However, the locations of these spawning grounds are largely unknown. Chapter 5 (“Identifying spawning areas of *D. gigas*”) addresses this information gap, analyzing available information on paralarva distribution, sea surface temperature and ocean currents. This chapter is distinct from the other chapters because, while it generates important information about initial environmental conditions, it only contributes indirectly to the subsequent development of the simulation model.

Chapter 6 (“Individual level traits and their computation”) focuses on the individual, its traits and its development. The energy needs and effects of energy availability on development are discussed with reference to different hypothetical growth functions. This discussion identifies constraints on individual development, as well as inconsistencies in several hypothesized aspects of the life history of *D. gigas*. This chapter provides a computational framework for the quantitative analysis of energy-dependent physiology.

The incapability of an individual to store energy requires a different strategy to ensure the continued availability of energy for metabolic needs in times of food deprivation. Cannibalism is one such strategy, as proposed by (Ibáñez & Keyl, 2010). Chapter 7 (“Cannibalism as a possible survival strategy”) applies the energy model developed in this thesis and the results of preparatory computations in Chapter 6 to analyze the effects of different cannibalism strategies under different scenarios.

The expression of individual traits in response to variations in energy availability may give rise to varied or even contradictory emerging effects at school, cohort and population level. These effects, are discussed in Chapter 8 (“Properties at the population level and their computation”), drawing on the previous discussion of energy-driven population-level effects in Chapter 3.

Chapter 9 (“The simulation of ecological models: Enhanced approaches and techniques”) addresses different topics related to the implementation of simulation models. It reviews recent advances and identifies opportunities for further improving all stages of the modeling process, including formulation of the conceptual source model, documentation, and implementation of the simulation model.

Chapter 10 (Summary and outlook) summarizes the results and presents an outlook on the future development of ecological simulation models.

1. Preface

The appendix contains additional material relevant to individual chapters or across multiple chapters. Appendix A provides the overview and description of the PYTHON scripts used for visualization of the results, mainly relating to sea surface temperature in Chapter 5. Appendix B presents a brief description of the scripts used to prepare the cannibalism simulation data for subsequent processing in MATHEMATICA scripts.

The MATHEMATICA scripts, described in Appendix C, provide the basis for the static calculation, analysis and visualization of the results in various chapters including the cannibalism simulations in Chapter 7.

Appendix D briefly describes the development of an abstraction layer for development of simulations without recourse to an off-the-peg simulation framework. This appendix also provides further information on the specific simulations and computations presented in different chapters. Appendix E contains further information about the topics discussed in Chapter 9, including the execution unit for a state-based modeling object (Appendix E.1), development environments for simulation models (Appendix E.2), and the use of simulation frameworks in combination with appropriate programming languages (Appendix E.3). The individual chapters are self-contained and employ a variety of methodologies, including static mathematical analysis and simulation techniques. Each chapter explores a different aspect or group of related aspects in the energy driven life history of *D. gigas* using a methodology appropriate to the topic under discussion.

2. Objects of Study

2.1. *Dosidicus gigas* (*D. gigas*)

2.1.1. About the species, its importance and its basic characteristics

About. *Dosidicus gigas* (*d'Orbigny, 1835*) is a little understood cephalopod species of high commercial and ecological importance. *D. gigas*, also known as the *Humboldt squid*, *jumbo squid*, *jumbo flying squid*, *red demon*, *red devil* (*diablo rojo*), is the only species of the genus *Dosidicus*. It is a pelagic, opportunistic, predatory squid, endemic in the Eastern Pacific Ocean, especially in the waters of the Humboldt Current (Peru Current) and off the Coast of Mexico and California (Argüelles & Tafur, 2010; Camarillo-Coop et al., 2010; Goicochea-Vigo et al., 2019; C. Nigmatullin et al., 2001; Staaf et al., 2010b; W. Xu et al., 2019). Starting with a paralarva size of approx. one millimeter after hatching, individuals reach a mantle length of up to 1.2m and a weight of up to 65kg in a short time, typically less than two years (C. Nigmatullin et al., 2001; Portner et al., 2020; Rosa et al., 2013). This fast growth rate, combined with its high abundance, mean that the species draws a significant amount of energy from the (marine) food web, which in turn impacts and shapes the ecosystem (Bruno et al., 2021; Gong et al., 2020; Liu et al., 2016; Liu et al., 2020).

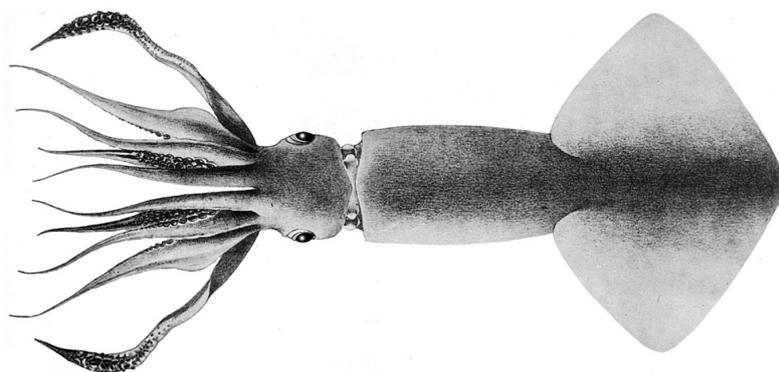


Figure 2.1.: *D. gigas* drawing, Pfeffer 1912 in R. E. Young and Vecchione (2013)

2. Objects of Study

Commercial and ecological importance. *D. gigas* supports a large (local) cephalopod fishery, mainly in Peru, Chile and Mexico, and also an industrial fishery (off the coast of Peru between 4°S and 10°S and in adjacent international waters between 3°S and 18°S; in the Chilean Exclusive Economic Zone at 36°S to 38.5°S off Biobio, and at 29°S to 34°S off Coquimbo and Valparaiso; and off the Central American coast between 5°N and 10°N) (Arkhipkin et al., 2014; Frawley et al., 2019; Gong et al., 2020; Liu et al., 2016; Portner et al., 2020; P. G. Rodhouse et al., 2006; Sanchez et al., 2020; Trasviña-Carrillo et al., 2018; Waluda & Rodhouse, 2005; H. Xu et al., 2021; Yu, Chen, & Liu, 2021; Yu, Wen, et al., 2021). In the Eastern Pacific Ocean, *D. gigas* is the most important species in terms of catch and economic value (Yu, Chen, & Liu, 2021).

The commercial *D. gigas* fishery uses automated squid jigs of different sizes and strong lights to attract specimens when fishing at night. Commercial fishing of *D. gigas* began in 1974. It collapsed during and after strong El Niño (EN) events in 1982/1983 and 1997/1998 (Arguelles et al., 2001; Markaida et al., 2008), but has intensified since 2002. Portner et al. (2020) also report a

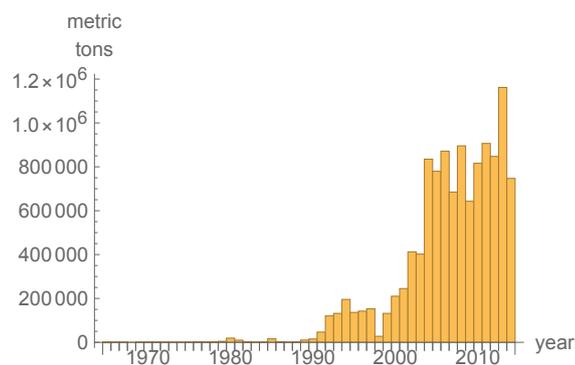


Figure 2.2.: Annual landings in metric tons (FAO, 2021).

drastic reduction in mantle length (ML) of mature individuals during the three strong El Niño events in 1997/1998, 2009/2010 and 2015/2016. Since 2004, landings have remained—with annual fluctuations—at a high level (Figure 2.2), even during the strong EN events in 2009/2010 and 2015/2016.

Growth in size across four orders of magnitude during rapid ontogenesis (Staaf, 2010) provides *D. gigas* access to different trophic levels in the food web in a short time. The species has a large daily food uptake, both in order to obtain energy to sustain its rapid growth (Arkhipkin et al., 2014; Bruno et al., 2021; Chen et al., 2020; Frawley et al., 2019; Goicochea-Vigo et al., 2019; Ibáñez et al., 2011; Ibáñez & Keyl, 2010; Robinson et al., 2013) and due to an inefficient propulsion drive for locomotion (Staaf, 2010). *D. gigas* is a voracious and opportunistic predator (Bazzino et al., 2010; Bruno et al., 2021; Field et al., 2007; Gilly, Markaida, et al., 2006; Gong et al., 2020; Ibáñez & Keyl, 2010; Liu et al., 2020; Markaida et al., 2008; R. I. Ruiz-Cooley et al., 2013; H. Xu et al., 2021) that impacts the local fish population by feeding on a wide variety of prey

2.1. *Dosidicus gigas* (*D. gigas*)

year	tons	year	tons	year	tons	year	tons	year	tons
1965	100	1976	1614	1987	309	1998	27471	2009	642855
1966	200	1977	659	1988	1737	1999	131868	2010	815978
1967	100	1978	1642	1989	10372	2000	210064	2011	906310
1968	0	1979	4581	1990	15103	2001	244955	2012	950630
1969	400	1980	19068	1991	46585	2002	412431	2013	847292
1970	600	1981	9787	1992	120459	2003	402045	2014	1161690
1971	500	1982	1152	1993	131832	2004	834754	2015	1003774
1972	200	1983	91	1994	195434	2005	779680	2016	747010
1973	100	1984	380	1995	136288	2006	871359		
1974	90	1985	15886	1996	142188	2007	684860		
1975	365	1986	1233	1997	152546	2008	895365		

Table 2.1.: Annual landings of *D. gigas* in metric tons (FAO, 2021).

species, including commercially important fish such as Chilean hake (*Merluccius gayi gayi*) and Pacific hake (*Merluccius productus*) (Field et al., 2007; Field et al., 2013; Holmes et al., 2008). *D. gigas* competes for the same resources as commercial fishing, so its commercial importance is twofold, in that it is both fished and competes with other fisheries.

A high potential fecundity up to 32 million eggs per female, a short generation time, opportunistic feeding behavior, and temperature and hypoxia tolerance (Gilly, Markaida, et al., 2006; C. M. Nigmatullin & Markaida, 2009; Trübenbach, Teixeira, et al., 2013) allow *D. gigas* to reproduce rapidly and adapt quickly to environmental changes, for example in oxygen and temperature levels (Frawley et al., 2019; Sánchez-Velasco et al., 2016; Yu, Chen, & Liu, 2021). These factors have enabled the recent expansion of the range of *D. gi-*

Kingdom:	Animalia
Phylum:	Mollusca
Class:	Cephalopoda
Order:	Teuthida
Family:	Ommastrephidae
Subfamily:	Ommastrephinae
Genus:	<i>Dosidicus</i> (Steenstrup, 1857)
Species:	<i>D. gigas</i>

Table 2.2.: Taxonomic classification of *D. gigas*.

gas toward both poles, which in turn has impacted on the food web structure in newly colonized areas (Keyl et al., 2008). Through its roles as a key predator and key prey item, *D. gigas* links lower trophic levels (the massive biomass of micronekton) to upper trophic levels (Alegre et al., 2014; Gong et al., 2020; J. W. Young et al., 2013).

2. Objects of Study

Taxonomy. *D. gigas* belongs to the group of cephalopods, which has existed for about 500 million years since the early Cambrian. The cephalopod group is part of the molluscan class and contains the three subclasses *Nautiloidea*, *Coleoidea*, and the extinct *Ammonoidea*. The *Nautiloidea* are distinguished by the possession of an external shell and the subclass consists of five recent species (Keyl, 2009). The *Coleoidea* groups 650–1000 recent species and includes the genus *Dosidicus* with its only known species *Dosidicus gigas* (Table 2.2). The survival of such a large number of species to this day is evidence of the adaptability of this subclass to environmental change.

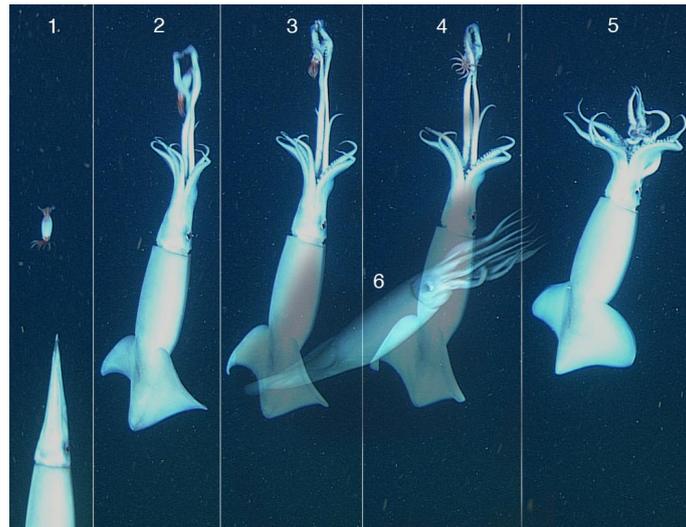


Figure 2.3.: Still images from video sequence showing feeding of *D. gigas* (R. E. Young & Vecchione, 2013). See also Zeidberg and Robison (2007) and supporting video material (<http://www.pnas.org/content/suppl/2007/07/19/0702043104.DC1>, date accessed 2021-07-15).

Anatomical characteristics. Members of the cephalopod family all possess the head-arm complex (*cephalopodium*) that gives the family its name, in which the arms are modified molluscan feet used for both locomotion and predation. Other characteristics of cephalopods include bilateral body symmetry (Figure 2.1), a siphon for jet propulsion locomotion, a beak, a well-developed eye, and the ability to squirt ink. The body mass consists of 75–90% muscle tissue, located mainly in the mantle that encloses the inner organs (Rosa & Seibel, 2010; Semmens et al., 2004). A distinctive characteristic of *D. gigas* is the possession of photophores in the skin that enable it to change color, giving rise to the names “red demon” and “red devil”.

2.1. *Dosidicus gigas* (*D. gigas*)

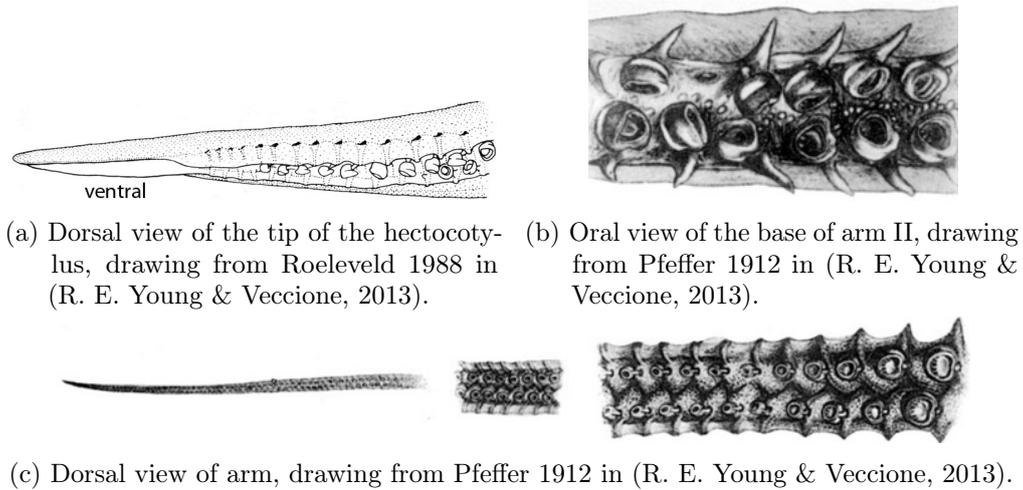


Figure 2.4.: Anatomical images of *D. gigas*.

The well-developed head-arm complex is grouped around the parrot-shaped beak that defines the mouth and contains a set of eight arms and two tentacles. The fourth left or right arm of the male is hectocotylyzed (Roper et al., 1984) and contains spermatophores. The arms are used to grasp and swim, while the tentacles are used to grab and pull prey items towards the beak (Figure 2.3).

The distal ends of the arms and tentacles are covered with teeth to hold prey items (Figure 2.4b) and 100–200 suckers with toothed chitinous rings that prevent prey items from sliding away (Keyl, 2009; Roper et al., 1984). This combination of arms, tentacles and strong beak enable efficient feeding. The beak tears the prey items into chunks, a capability that extends the spectrum of prey size to include prey items larger than the mouth opening. The radula, a rough structure inside the beak, transports the chunks through the oesophagus into the stomach.

The mantle encloses but does not encapsulate the inner organs, which comprise the gills, stomach, kidneys, reproductive organs (female: nidamental glands, oviducal glands, ovary; male: spermatophoric complex) and a digestive gland (hepatopancreas). A funnel (siphon) extrudes from the inside of the mantle and acts as a propulsion drive when water is pressed through the funnel by the contraction of the mantle. The orientation of the orifice of the funnel controls the direction of locomotion. A flap that can be opened or closed directs the seawater out of the mantle or through the funnel and the mantle contraction links locomotion to oxygen consumption by pushing water through the gills.

2. Objects of Study

Metabolism, energy needs and energy buffering. Compared to other species such as mammals, cephalopods have a high energy demand (B. A. Seibel, 2007): “*D. gigas* has among the highest metabolic rates of any animal on this planet” (B. A. Seibel, 2011). *D. gigas* grows from a millimeter sized paralarvae to a full grown adult with a ML of approx. 400mm within a few months (Table 2.3); larger specimen are presumed to need more time to achieve their terminal ML. The reproductive complex, especially in the females, occupies a considerable portion of the inner organs and requires large amounts of energy to build and maintain. Rapid growth and the energy demands of the reproductive tissue, as well as inefficient locomotion (O’Dor & Webber, 1991; Rosa & Seibel, 2008; Staaf, 2010), all mean that *D. gigas* requires a large daily uptake of energy from feeding and a correspondingly high digestion rate. Daily food uptake is estimated at about 3–13% of body weight (Ehrhardt, 1991; Gilly, Markaida, et al., 2006; C. Nigmatullin et al., 2001; Rosa & Seibel, 2010).

The metabolism is adapted to temperature and oxygen levels in the different layers of the water column. *D. gigas* tolerates temperatures of 4 – 32°C (C. Nigmatullin et al., 2001) and oxygen saturation levels encountered from the oxygen minimum layer (OML) up to the surface (Rosa & Seibel, 2010).

D. gigas has limited energy buffering capability (Ibáñez & Keyl, 2010). Given its high energetic needs, this makes *D. gigas* susceptible to starving and this is the principal threat to the survival of individual animals and populations in times of food scarcity. Since *D. gigas* is semelparous, it is likely that the species has evolved a special survival strategy that addresses this energy issue, otherwise it is unlikely to have survived to the present day.

2.1.2. Distribution range, range extension and habitat

Distribution range. *D. gigas* has significantly extended its distribution area since the start of the present century. The distribution area previously extended from Tierra del Fuego (south of Chile) to California (North America), from 30°N to 40°S, with highest abundances in the Gulf of California (GOC) and in the Humboldt Current in the waters off Peru (C. Nigmatullin et al., 2001). Since then *D. gigas* has extended its distribution range northwards to Canada/Alaska and further towards the southwest (Bazzino et al., 2010; Field et al., 2013; Gilly, 2006b; Keyl et al., 2008). The interactive species distribution map of the Food and Agriculture Organization of the United Nations (FAO) (FAO, 2021) shows a distribution range from 59°N to 53°S for 2021 (Figure 2.5); this is similar to the range respectively 60°N to 50°S given by Rosa et al. (2013). The

2.1. *Dosidicus gigas* (*D. gigas*)

longitudinal range to the west is still uncertain (FAO, personal communication) but is presumed to extend to a maximum of between 125°W and 140°W (Rosa et al., 2013).

The abundance of *D. gigas* is highly variable. For example, the species appeared in large numbers in the northern hemisphere in 1997–1998 and disappeared thereafter until 2002–2003; similar variations have been reported in the southern hemisphere (Keyl et al., 2008; Rosa et al., 2013). The changes in abundance may be linked to environmental conditions such as the strong EN event of 1997/1998 (Portner et al., 2020).

The distribution range may be larger than documented by the FAO. Figure 2.6 shows two occurrence events outside the known distribution range (“GBIF Occurrence Download”, 2021): a specimen recovered off Greenland and an unconfirmed sighting near Thailand. Both events, especially the Greenland specimen, may be indicators for a broader distribution range and further evidence of the adaptive potential of *D. gigas* to different environments.

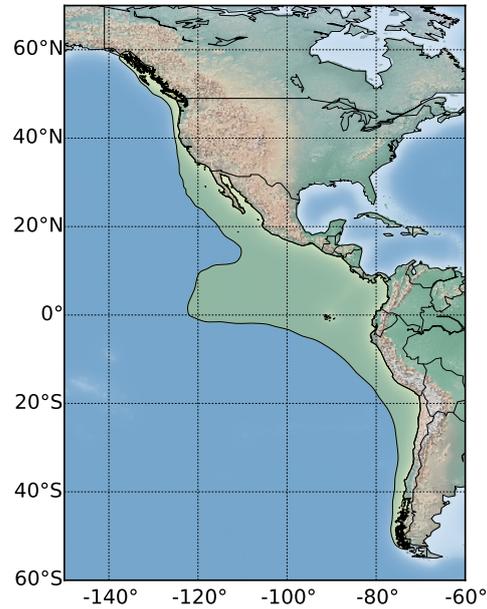


Figure 2.5.: Current distribution of *D. gigas* (FAO, 2021).

Range extension. *D. gigas* shows no marked genetic population structure within its range (Rosa et al., 2013; Sanchez et al., 2020; Sanchez et al., 2016; Staaf, 2010), which suggests a high frequency of migration between them (Rosa et al., 2013). Migration within the GOC over small distances, i.e., 130km, has been documented (Markaida et al., 2005), but the occurrence of large-scale migration is still in discussion. Local movements include, for example, the seasonal migration from southern and central California to the Queen Charlotte Islands at 53°N. Small individuals appear in late spring/early summer in central California and migrate northwards until late summer/early fall (Field et al., 2013); the by then large specimens migrate back to central California in late fall/early winter, possibly spawning along the way. *D. gigas* spawns in the northern hemisphere at least in GOC and the Pacific Ocean off Baja California, and in the southern hemisphere off the coast of Peru, mainly at 3°S to 8°S and 12°S to 17°S.

2. Objects of Study

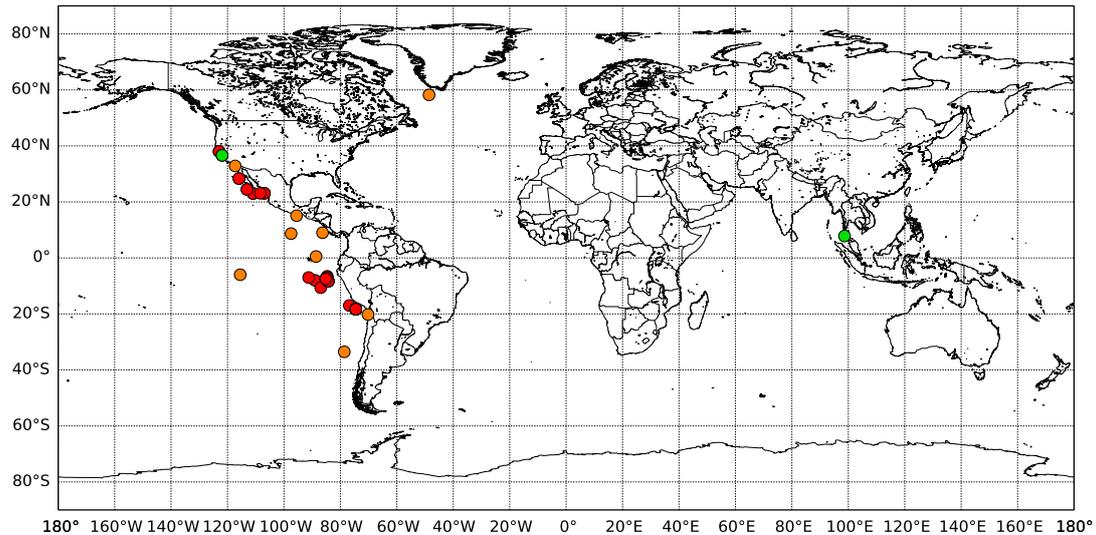


Figure 2.6.: Referenced single occurrences of *D. gigas* (“GBIF Occurrence Download”, 2021); red markers for preserved specimen, orange for specimen of undocumented state, green for sight record.

The ability to migrate over long distances, the flexible diet obtained through opportunistic feeding, tolerance to adverse environmental conditions such as hypoxia and high temperatures, and rapid growth combined with high fecundity are all considered to be important factors that have enabled the recent range expansion of *D. gigas* (Ibáñez & Cubillos, 2007; Keyl et al., 2008; Rosa et al., 2013). This may also be the result of exogenous factors such as the large-scale sea surface warming, expansion and shoaling of the OML (Hoving et al., 2013; R. I. Ruiz-Cooley et al., 2013; Sakai et al., 2008; Stewart et al., 2014; Xavier et al., 2015), and the removal of competing predators. It has been suggested that *D. gigas* was able to expand its distribution range during the La Niña (LN) event in 1995/1996 and the subsequent strong EN event in 1997/1998, as a result of an “ecosystem reset” that reduced abundances of competing teleost species, e.g., mackerel and hake. Its opportunistic feeding behavior and the absence of physiological limitations allowed *D. gigas* to occupy the ecological niches of these competitors (Keyl et al., 2008; Portner et al., 2020).

Rosa et al. (2013) ruled out temperature as the only factor driving the recent range expansion, pointing out that *D. gigas* is regularly exposed to large temperature extremes during daily vertical migrations and therefore temperature alone is unlikely to trigger long-distance migrations. Instead, climate change in the Eastern Pacific Ocean might confer a competitive advantage on *D. gigas* due to its temperature and hypoxia tolerance.

Assertion of range expansion. The most recent range extension may have been triggered by a single strong EN event, but a number of long-term changes in oceanic climate may favor the permanent establishment of *D. gigas* in the newly colonized areas. These include the hypothetically important effect of a projected horizontal and vertical expansion of the OML (Gilly et al., 2013), which provides favorable conditions for predation by *D. gigas* on organisms in the overlying deep scatter layer as explained below (Gilly et al., 2012). However, permanent latitudinal range extension towards the north and south poles is likely limited “[...] by the sustainability of foraging potential within an energetically acceptable distance of spawning grounds” (Rosa et al., 2013).

Habitat. *D. gigas* occurs between the surface and a depth of 1200m in the Pacific Ocean and nearshore waters (Keyl, 2009). Its occurrence within the upper layer (0–100m) coincides to areas within the isoline of average phosphate concentration of $0.8\text{mg P-PO}_4^{3-}/\text{m}^3$, which is presumed to be the boundary of the highly productive waters (Keyl, 2009; C. Nigmatullin et al., 2001; Staaf, 2010).

As a general rule of thumb, *D. gigas* inhabits the 200–1000m water column during daylight hours and undertakes vertical migrations up to the surface to the 0–200m water column during the night (Gilly, Markaida, et al., 2006; C. Nigmatullin et al., 2001; Stewart et al., 2013). The 200–1000m water column corresponds —depending on local conditions— to the OML and an oxygen saturation of $<0.5\text{mL L}^{-1}$ ($<20\mu\text{mol kg}^{-1}$). Above and below the OML is the oxygen limited zone with an oxygen saturation of $0.5\text{--}1.4\text{mL L}^{-1}$ ($20\text{--}60\mu\text{mol kg}^{-1}$). In tagging experiments, some specimens in the California Current System showed a preference for the oxygen limited zone over the OML, while specimens of the GOC preferred the OML (Gilly et al., 2012; Gilly et al., 2013; Stewart et al., 2013). The upper boundary of the OML varies with latitude, ranging from approx. 100m in the southern GOC to approx. 600m off British Columbia, while the lower boundary, independent of location, is at approx. 1200m (Rosa et al., 2013). Using aerobic metabolism suppressing techniques (Trübenbach et al., 2014) and changes in swimming behavior, *D. gigas* survives within the OML over an extended period of time ($>9\text{h}$). *D. gigas* is presumed to utilize the OML in several different ways:

Protection from predatory fishes The OML provides *D. gigas* protection from predatory fishes (Gilly, Markaida, et al., 2006; Keyl et al., 2011; Sakai et al., 2017; Stewart et al., 2013) (see food web discussion in Subsection 2.1.3), whose high oxygen demand is not met inside the OML (B. A. Seibel, 2011; Stewart et al., 2013). The OML does not provide protection of from predatory deep-sea mammals such as the sperm whale as they are not affected by levels of dissolved oxygen

2. Objects of Study

(Stewart et al., 2013). Since *D. gigas* moves more slowly within the OML and a predator like the sperm whale may be able to take advantage of this fact; however escape responses have not been studied under OML conditions and they might not be impaired (Gilly et al., 2012).

Active feeding Individual specimens of *D. gigas* can tolerate, remain active and feed actively in the cold, hypoxic conditions of the OML for several hours.

Active feeding above the zone The OML “concentrates” potential prey items (Stewart et al., 2013) above its boundary, because individuals of many species try to avoid visual hunting predators near the surface layers, but cannot move into the OML for an extended period of time. Prey items therefore assemble just above the OML and this allows *D. gigas* to forage efficiently (Sakai et al., 2017), unhindered by the dim light, in areas where it has few competitors, i.e., in the acoustic deep scattering layer where it overlaps with the OML (Rosa et al., 2013). Stewart et al. (2013) identifies the opportunity to forage from below on potentially lethargic animals within the upper OML as a further benefit (Rosa et al., 2013).

Thermal stress relief Despite its wide temperature range tolerance, it appears that *D. gigas* suffers thermal stress at high temperatures $> 23^{\circ}\text{C}$ and tends to avoid waters where temperatures exceed this threshold (Rosa et al., 2013). However these high temperatures are characteristic of near-surface waters that *D. gigas* enters during the night i.e., for high activity feeding. The lower temperatures in the OML may provide necessary thermal stress relief (Davis et al., 2007; Gilly, 2006b).

As noted above, during the night *D. gigas* generally inhabits the layer near the surface where the low light level protects it from visual hunting predators. However, the actual vertical migration pattern is more complex (Gilly, Markaida, et al., 2006; Sakai et al., 2017; Stewart et al., 2013). Even at night, when *D. gigas* is generally found near the surface, downward vertical migration has been observed in the course of active feeding and for possible thermal stress relief (Davis et al., 2007; Gilly, 2006b).

2.1.3. Food web

The high growth rate of *D. gigas* provides access to different trophic levels in a short time. The diet changes rapidly (Bruno et al., 2021; Liu et al., 2020); for example young squid (15–100mm ML) feed on mesozooplankton, medium sized specimens (150–330mm ML) feed mainly on myctophids, and large squid ($>400\text{mm ML}$) feed

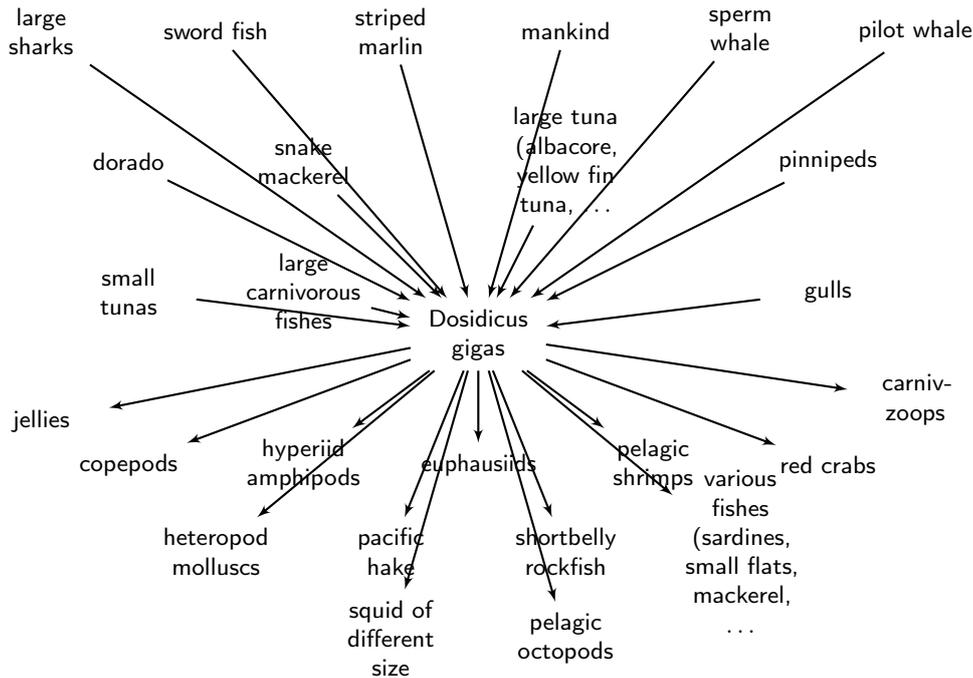


Figure 2.7.: Prey item relationship between *D. gigas* and other species, according to Field et al. (2006), Field et al. (2007), C. Nigmatullin et al. (2001), and R. Ruiz-Cooley and Markaida (2006). Carniv-zoops: “Functional group including pasiphaid, sergestid and other pelagic shrimps, chaetognaths, pelagic polychaetes, and the pelagic stages of many invertebrates, such as crab megalopae.” (Field et al., 2006) Arrows point from predator to prey item.

on myctophids, squids and larger fishes (Rosa et al., 2013). Figure 2.7 summarizes the geographical variation of predator-prey relationships between *D. gigas* and other species.

D. gigas affects the local populations of its prey species due to its high feeding demands and plays a significant role in the vertical energy flow of Eastern Pacific ecosystems (Rosa et al., 2013; Yu, Wen, et al., 2021). *D. gigas* is an opportunistic predator (Chen et al., 2020; Liu et al., 2020; Neira & Arancibia, 2013; Rosa et al., 2013) and the window of predation (WOP) (Ibáñez & Keyl, 2010) may be the principal criterion for prey selection. The typical WOP, with WOP in percent given by $ML_{\text{prey}}/ML_{\text{predator}} \cdot 100$, is estimated at 5–15% ML with an observed maximum of approx. 60% ML (Field et al., 2007; Field et al., 2013). This wide WOP allows *D. gigas* to feed on various, locally abundant species. If cannibalism is considered a special case of predation, then the

2. Objects of Study

WOP ranges from 4% to 87% ML (Bazzino et al., 2010; Markaida & Sosa-Nishizaki, 2003).

2.1.4. Life history

Introduction. There is no known study of the growth and maturation of *D. gigas* under controlled conditions over an extended period of time. Paralarvae have been studied (Staaf, 2010; Staaf et al., 2008; Yatsu et al., 1999) and mature *D. gigas* observed e.g., by tagging or fishing. A specimen of *D. gigas* is reported to have survived in captivity for twelve days (Keyl et al., 2008). Even basic traits, such as the lifespan (Table 2.4), are still subject to discussion, highlighting the scarcity of knowledge about *D. gigas*.

Ontogenesis. Table 2.3 shows the ontogenesis of *D. gigas*. Due to the scarcity of observational data, authors propose different timespans for the life stages. The first systematic study of the paralarval stage was conducted in the Gulf of Mexico in 2008 (Camarillo-Coop et al., 2010). The hatching time is temperature dependent; the optimum water temperature for hatching was determined by *in vitro* experiments to be 17–25°C, while development of embryos ceases above 30°C (Staaf et al., 2011).

Stage	Size (typical)	Information
egg	–	fertilized, spawned by female adult
embryo	–	6–9d at 18°C
paralarva	1 to 10mm ML	typical rhynchoteuthion
transient	10 to 15mm ML	change of morphology, food spectrum and ecological status
juvenile	15 to 100mm ML	
transient	100 to 150mm ML	change of morphology, food spectrum and ecological status
subadult	150 to 300–350mm ML	
transient	300–350 to 400–450mm ML	change of morphology, food spectrum and ecological status
adult	400–1000mm ML	

Table 2.3.: Ontogenesis of *D. gigas* adapted from C. Nigmatullin et al. (2001). The sizes may vary to geographical locations where warmer regions may result in smaller individuals in each stage and colder waters larger individuals in each stage.

The literature gives inconsistent reports on both the time required to reach the adult stage and the lifespan of individuals. Estimates range from three months to almost one year, and from approx. one year up to four years, respectively. For further discussion of this point, see also Subsection 2.1.5.

Controlling exogenous factors, growth function and growth. The maximum size of a monocyclic (semelparous) species, in case of *D. gigas* the maximum terminal ML (ML_{terminal}), is linked to the size-at-maturity (SAM) (Rosa et al., 2013). *D. gigas* individuals presumably die shortly after maturity, so the time to maturity limits the period available for growth. Growth is assumed to be determined by *food-uptake*, *temperature*, *day-length* and *light intensity* and perhaps other unknown factors (Anderson & Rodhouse, 2001; Argüelles & Tafur, 2010; Argüelles et al., 2008; Bazzino et al., 2007; Keyl, 2009; Keyl et al., 2011; Keyl et al., 2008; Markaida, 2006b; Markaida et al., 2004; Mejía-Rebollo et al., 2008).

There is consensus in the literature that these factors influence the growth rate and onset of maturity. However, the relative importance of these factors is controversial (Portner et al., 2020); moreover, the relationship between growth and factors favoring growth may be non-linear and, in some cases, influenced by threshold effects and interactions with other factors (Keyl, 2009). Keyl et al. (2008) summarize the relationship as: “High nutrition levels and low temperatures are known to cause late maturation and low nutrition levels and high temperatures cause early maturation.”

Lifespan. *D. gigas* is usually referred as “short-lived” (Argüelles & Tafur, 2010; Field et al., 2007; Frawley et al., 2019; Holmes et al., 2008; Ibáñez et al., 2011; Keyl, 2009; Mejía-Rebollo et al., 2008; Olson et al., 2006; Zúñiga et al., 2008) with a assumed lifespan of one to four years. Ibáñez et al. (2011) cites a lifespan of about one year, but larger specimens may live up to three years. However, Chen et al. (2011) found only specimens with a lifespan of one year. Different lifespans were found in different geographic regions, with individuals in warmer waters tending to have smaller ML with shorter lifespans (Portner et al., 2020). Table 2.4 shows the lifespans estimated by different authors.

The differences in measured lifespans are caused by the use of different methodologies. Some authors use statolith analysis based on the assumption of “one ring per day” (P. G. Rodhouse & Hatfield, 1990; Yatsu, 2000). However, according to Keyl et al. (2011), this method underestimates lifespans. The underlying assumptions have been proven for some Ommastrephidae, but not specifically for *D. gigas* (Mejía-Rebollo et al., 2008;

2. Objects of Study

	lifespan			
	1 yr	2 yrs	3 yrs	4 yrs
Portner et al. (2020)	+	+	.	.
Frawley et al. (2019)	+	+	.	.
Yu et al. (2019)	+	.	.	.
Yu and Chen (2018)	+	.	.	.
Lin et al. (2017)	+	+	.	.
Rosa et al. (2013)	+	+	.	.
Chen et al. (2013)	+	+	.	.
Ibáñez et al. (2011)	+	+	+	.
Chen et al. (2011)	+	+	.	.
Keyl (2009)	+	+	+	+
Keyl et al. (2008)	+	.	.	.
Holmes et al. (2008)	.	+	.	.
Mejía-Rebollo et al. (2008)	+	+	.	.
Field et al. (2007)	+	+	.	.
Gilly, Markaida, et al. (2006)	+	+	.	.
Waluda and Rodhouse (2006)	+	.	.	.
Markaida et al. (2005)	.	+	.	.
Markaida et al. (2004)	+	+	.	.
Anderson and Rodhouse (2001)	+	.	.	.
C. Nigmatullin et al. (2001)	+	+	.	.
Arguelles et al. (2001)	+	.	.	.
Yatsu (2000)	+	.	.	.
Yatsu et al. (1999)	+	+	.	.
Hernandez-Herrera et al. (1998)	+	+	.	.

Table 2.4.: Estimated lifespan of *D. gigas* in studies by different authors. “+” indicates an affirmative statement.

C. Nigmatullin et al., 2001). The formation of rings may not in fact occur regularly, i.e., on a daily basis, but irregularly in response to changes in feeding behavior, temperature and metabolism. Therefore, it is argued, age is systematically underestimated (Keyl et al., 2011; Markaida et al., 2005; Rosa et al., 2013).

An alternative method is to combine results of growth analysis obtained from tagged squid and modal progression analysis which results are sensitive to prior assumptions that affect the interpretation of the data (Keyl et al., 2011). Studies using this method suggest that *D. gigas* has a size-dependent lifespan of 1–3 years (Keyl et al., 2011; Markaida et al., 2005). Taking all the results shown in Table 2.3 into account, a lifespan of 12–23 months (Rosa et al., 2013) may be a reasonable assumption.

It should be noted that no studies combined the use of both methods, and no controlled experiments have been conducted to verify any of these hypotheses. In general, larger specimens are presumed to be older. However, specimens may grow more slowly in warmer areas due to reduced food availability and oxygen saturation and therefore mature (and die) earlier.

Reproduction, spawning and hatching. *D. gigas* has the highest known potential fecundity of all cephalopods, with an estimated average of 18–21 million oocytes per female and an estimated maximum of 32 million oocytes (C. M. Nigmatullin & Markaida, 2009). Females spawn “no less than half of the initial potential fecundity” (Rosa et al., 2013) during multiple (8–14) batches (Hernández-Muñoz et al., 2015; Ibáñez et al., 2011; Rosa et al., 2013; R. I. Ruiz-Cooley et al., 2013; Staaf et al., 2008; Tafur et al., 2010). A larger SAM is associated with a higher potential fecundity (C. M. Nigmatullin & Markaida, 2009). The male has 300–1200 spermatophores of a size of ≈ 25 – 35 mm which are stored for later egg fertilization inside the female during mating (Markaida & Sosa-Nishizaki, 2001; Tafur et al., 2010).

Spawning take place throughout the year in both hemispheres. In the southern hemisphere, there is a reported spawning peak between October and January (especially in November) and a less prominent one in July to August (Ibáñez et al., 2011; Jereb & Roper, 2010; Keyl et al., 2011; Rosa et al., 2013). Similarly, spawning peaks have been reported for the northern hemisphere from October to November and from April to May (Rosa et al., 2013). These observations indicate an environmentally dependent trigger for spawning, unless the peaks are artifacts of the methodologies used to measure spawning rates.

The spawning areas are largely unknown. *D. gigas* likely spawns in the “relatively inaccessible open sea and extrudes its eggs in a large and fragile pelagic gelatinous mass” (Rosa et al., 2013). Unmated mature females are rarely found and mature specimens occur throughout the year in both hemispheres (Rosa et al., 2013). But despite the high potential fecundity of female individuals, to date naturally deposited eggs have rarely been observed. Staaf et al. (2008) found neutrally buoyant eggs at 16 m depth at $\approx 22^\circ\text{C}$ in the GOC in 2006, while Birk et al. (2016) report six findings at 9–14 m depth at 24 – 25°C within a 5 km radius, also in the GOC.

Paralavae have been used as an indirect indicator for spawning ground detection, but the paralarvae of *D. gigas* can easily be confused with those of purple back flying squid (*Sthenoteuthis oualaniensis*) if identified by morphology alone (Ramos-Castillejos et al., 2010; Staaf, 2010). The geographical distributions of both species overlap

2. Objects of Study

(Gilly & Markaida, 2006; Ramos-Castillejos et al., 2010; Staaf et al., 2010b); therefore identification of paralarvae should be based on DNA analysis (Ramos-Castillejos et al., 2010; Staaf et al., 2013).

The few known spawning locations and regions are listed in Table 5.1 (Section 5.2). In general, spawning in the southern hemisphere seems to be widespread and is known to occur along the entire Peruvian coast, and particularly at 3°S to 8°S and 12°S to 17°S. In the northern hemisphere spawning is known to occur in GOC and the Pacific Ocean off Baja California (Rosa et al., 2013).

In vitro experiments demonstrated a temperature window of 15–25°C for successful embryonic development and hatching, but with slower development and higher mortality at 15°C, and failure of embryonic development over 30°C (Staaf et al., 2008; Staaf et al., 2011). Yatsu et al. (1999) reported a hatchling size of 0.9–1.3mm at 18°C six to nine days after fertilization and a growth rate of about 4% after hatching. In general, only limited information is available about successful hatching and required environmental conditions.

2.1.5. Intraspecific population structure

D. gigas changes its biological characteristics rapidly in response to environmental triggers such as EN and LN (Portner et al., 2020) events in both Humboldt Current System and California Current System and therefore its intra-specific population structure is considered to be complex.

Dimorphism. *D. gigas* exhibits a weak sexual dimorphism. The fourth left or right arm of the males is hectocotylized (Roper et al., 1984) and contains the spermatophores.

Females in the cephalopod family are generally larger than males (Markaida et al., 2004; C. Nigmatullin et al., 2001), see Table 2.5. In the case of *D. gigas* females may mature later and grow for longer due to delayed spawning. Different growth functions for the sexes may also apply (Chen et al., 2011; Markaida et al., 2004; Tafur et al., 2010) but, as is the case for estimates of life span, calculated values of the growth function in the literature vary widely (Rosa et al., 2013).

Keyl (2009) describes minor differences in body shape: the female's body is more barrel-like (wider in the centre and thinner at the ends), and the female beak tends to be stronger.

Size-at-maturity-groups and genetics. Different SAM are observed in specimens in the field at different locations. The occurrence of these differences in SAM led, in 1986,

Group	male ML in mm	female ML in mm	distribution
small-size	130–260	140–340	predominantly near-equatorial area
medium-size	240–420	280–600	whole species range, except higher latitudes
large-size	> (400–500)	550–650 to 1000–1200	northern 10–15°N/southern 10–15°S peripheries of range

Table 2.5.: Characteristics and geographical distribution of *D. gigas* size groups (C. Nigmatullin et al., 2001).

to the classification of *D. gigas* into two groups (Anderson & Rodhouse, 2001; Arguelles et al., 2001). Later observations and statistical analysis led to a classification of small-, medium- and large-sized SAM-groups (C. Nigmatullin et al., 2001). These groups have a different geographic distribution (see Table 2.5); however there is considerable overlap between the distributions of both the small- and medium-size groups, and the medium- and large-sized groups. The large-size group is mainly confined to the northern and southern peripheries of the range, but specimens sometimes occur near the equator (C. Nigmatullin et al., 2001).

The classification into three groups and the corresponding distributions are widely accepted by researchers (Keyl, 2009), but have been called into question by some more recent studies (Keyl, 2009; Sandoval-Castellanos et al., 2007; Staaf et al., 2010b). C. Nigmatullin et al. (2001) speculated that the three groups could be genetically distinct and even represent separate species *in status nascendi*. This hypothesis could not be confirmed by later genetic analysis (Ibáñez et al., 2011; Sanchez et al., 2016; Sandoval-Castellanos et al., 2009, 2010; Staaf et al., 2010b), which found only small genetic differences between specimens of the three SAM-groups. These differences were below the 0.5–2% range considered necessary (regardless of species) to confirm the existence of a different species (Roux et al., 2016).

Instead, a weak but significant differentiation has been found between the populations of the northern and southern hemispheres (Ibáñez et al., 2011; Sandoval-Castellanos et al., 2009, 2010; Staaf et al., 2010b). This result divides this species into two groups, each roughly confined to one or other of the two hemispheres. Staaf et al. (2010b) identified a biogeographic break at 5–6°N separating the two groups. The equatorial currents and counter-currents may be the physical phenomena separating these two populations; currents analysis (Section 5.5) shows significantly stronger currents at 5–6°N, which might separate spawning grounds and reduce exchange between these populations. This

2. Objects of Study

Authors	length-at-maturity groups			S/M/L ²
	one	two	three ¹	
Keyl et al. (2011)	·	+	+	+
Argüelles and Tafur (2010)	+	+	+	+
Staaf (2010)	·	·	+	+
Staaf et al. (2010b)	·	·	+	+
Sandoval-Castellanos et al. (2009)	·	+	+	+
Keyl et al. (2008)	·	+	+	·
Argüelles et al. (2008)	·	·	+	+
Bazzino et al. (2007)	·	·	+	+
Field et al. (2007)	·	·	+	+
Markaida (2006b)	·	·	+	·
Markaida (2006a)	·	·	+	·
R. Ruiz-Cooley and Markaida (2006)	·	·	+	·
Markaida et al. (2004)	·	·	+	+
Markaida and Sosa-Nishizaki (2003)	·	·	+	·
Anderson and Rodhouse (2001)	·	+	·	·
Markaida and Sosa-Nishizaki (2001)	·	·	+	+
C. Nigmatullin et al. (2001)	·	·	+	·
Arguelles et al. (2001)	·	+	+	+
Tafur et al. (2001)	·	+	+	·
Tafur and Rabí (1997)	·	·	+	·

¹ three groups: small, medium and large

² Reference to the small/medium/large classification in (C. Nigmatullin et al., 2001)

Table 2.6.: Size group classification by different authors. Some authors give more than one group classification.

weak genetic differentiation between a northern and southern population is not sufficient for the definitive classification of two distinct populations. For example, Rosa et al. (2013) considers that *D. gigas* comprises a single large population with considerable range expansion in the Humboldt Current System. This concurs with the results of mitochondrial DNA analysis by Sanchez et al. (2020), who found three maternal lineages and two genetically differentiated stocks, with all lineages and stocks homogeneously distributed across the northern and southern hemispheres.

Table 2.6 shows the group classifications of different authors and whether their classification follows that proposed by C. Nigmatullin et al. (2001). As can be seen from the table, most authors follow C. Nigmatullin et al. (2001)'s classification. However re-analysis of catches suggests an alternative classification, into a small and a large

SAM-group (Keyl et al., 2008) which are spatially and temporally distinct (Argüelles et al., 2008). The two SAM-groups are presumed to be phenotypic extremes that have evolved in response to environmental conditions (see the overview of the functional triad migration-maturation-growth (fTMMG) in Section 2.2). In addition, the ML size of fished specimens is highly dependent on jig sizes and the ML of the caught specimen follows a Gaussian distribution (Keyl, 2009; Keyl et al., 2011; Markaida & Sosa-Nishizaki, 2001). Therefore, even with a uniform ML distribution, individuals of a certain size are preferred for fishing, giving an observer the impression of a certain, predominant ML based on the individuals caught.

Female:male sex ratio. There are usually more females than males (Goicochea-Vigo et al., 2019; Tafur et al., 2001), but with exceptions and considerable changes over time. The female:male ratio ranges from 1:1 up to 24:1 with high monthly variability (Bazzino et al., 2007; Chen et al., 2008; Field et al., 2007; Ichii et al., 2002; Keyl, 2009; Markaida & Sosa-Nishizaki, 2001; Markaida & Sosa-Nishizaki, 2003; Rosa et al., 2013; Tafur et al., 2010; Tafur et al., 2001). Velzquez-Abunader et al. (2012) found different sex ratios in the GOC, with females exhibiting higher abundance over most of the period 2000–2009.

The reason for the female-heavy sex ratio may be related to the fact that *D. gigas* is a monocyclic species that dies shortly after reproduction. The delayed reproduction cycle in females means that the males transfer the spermatophores to the females and die shortly afterwards, while the females continue growing and spawn the fertilized eggs later (Markaida & Sosa-Nishizaki, 2001; Tafur et al., 2010). Another reason could be cannibalism, i.e., predation by larger females on smaller male individuals. It has also been suggested that the female-heavy sex ratio is an adaptation mechanism to low energy environments, since the survival of females as the main reproductive part of the population would favor the survival of the population under these conditions (Tafur et al., 2010). Changes in the sex ratio may depend on environmental productivity and *D. gigas* may change its reproductive pattern in response to different levels of energy supply in the environment (Rosa et al., 2013).

Cannibalism. Cannibalism is common among cephalopods (Roper et al., 1984) and high cannibalism rates appear to be particularly common among *D. gigas* (Ehrhardt, 1991; Ibáñez et al., 2008; Markaida et al., 2008; Markaida & Sosa-Nishizaki, 2003; C. Nigmatullin et al., 2001). In the case of *D. gigas*, the rate of cannibalism increases with increasing ML (Alegre et al., 2014; Bruno et al., 2021; Liu et al., 2020; Markaida

2. Objects of Study

et al., 2008; Markaida & Sosa-Nishizaki, 2003). Since females attain greater MLs than males, this probably explains why they exhibit cannibalistic behavior more frequently than males (Markaida & Sosa-Nishizaki, 2003).

A number of explanations for these observed high levels of cannibalism have been proposed. Some authors suggest that they might be partly artificial, due to fishing stress (Bruno et al., 2021; Ibáñez & Keyl, 2010; Ibarra-García et al., 2014; Keyl et al., 2008; Markaida & Sosa-Nishizaki, 2003; Portner et al., 2020; Rosas-Luis & Chompoy-Salazar, 2016), or to a systematic sampling bias because only larger individuals, that are more likely to exhibit cannibalistic behavior, are caught by fisheries (Keyl et al., 2008).

In contrast, other authors suggest that cannibalism may be underestimated (Alegre et al., 2014). They point out that cannibalism could be a beneficial survival strategy that uses the population as an energy storage buffer to support a cohort in times of food scarcity (Ibáñez & Keyl, 2010; Keyl et al., 2008). In other words, cannibalism might have evolved as a strategy to ensure the survival of at least some specimens of a cohort in times of adverse environmental conditions. Extinction is an ever-present threat for populations of monocyclic species such as *D. gigas*, since the survival of the population depends entirely on the reproductive success of the current generation (Pecl & Jackson, 2008).

2.1.6. The suitability of *Dosidicus gigas* as a model organism for study

Xavier et al. (2015) sees the group of cephalopods as potential model organisms for studying the impacts of environmental changes, such as global climate change, ocean warming, sea level rise, biodiversity loss, overfishing, ocean acidification and expanding hypoxia. The commercial importance of cephalopods also renders them suitable as model organisms for the development of fishery management strategies (see also Yu, Chen, and Liu (2021)).

In general, cephalopods (except for nautilus) are characterized by a short lifespan, rapid growth, high fertility, semelparous maturation patterns and high variability in life history strategies (Ibáñez et al., 2011; Pecl & Jackson, 2008; Xavier et al., 2015). These characteristics are important in the context of external phenomena, e.g., environmental changes or overfishing, to which cephalopods show a greater sensitivity and respond more rapidly (Ibáñez et al., 2011; Xavier et al., 2015; Yu, Wen, et al., 2021) than other, long-lived and slow-growing species (Ibáñez et al., 2011).

Among the Cephalopoda, *D. gigas* is notable for its extreme expression of traits shared with other members of the class. *D. gigas* possesses extreme phenotypic plasticity in response to changing environmental conditions, for example with respect to ML at

2.2. The functional triad migration-maturation-growth

maturity, which ranges between 0.2m to 1.2m. Attaining the latter ML over the species' short lifespan requires a capacity for extremely rapid growth, which is made possible by *D. gigas*' extremely high metabolism rate and opportunistic feeding behavior. However, the species also has the (contrasting) capacity for metabolism suppression under hypoxic conditions, which it is able to tolerate.

Cephalopods respond rapidly to environmental changes, first at the individual and then at the population level. *D. gigas* in particular is known to be highly sensitive to environmental changes (Yu, Wen, et al., 2021). The species may be expected to show exceptionally rapid and pronounced responses to environmental change at the individual level, given its short generational turnover, probably ranging from six months to a maximum of about two years, as well as its extremely high metabolism rate.

In summary the abundance of the species and its voracious feeding behavior mean that it occupies an important position in marine food webs. The wide and pronounced phenotypical variation of the species, as well as dramatic changes in its abundance, are likely consequences of its extreme life history traits, and the effects on these of external environmental factors. Changes in these traits at individual and population levels are both visible and measurable (Ibáñez et al., 2011; Tafur et al., 2010). Therefore, *D. gigas* should be suitable as a model organism to study the effects of environmental changes on marine organisms and ecosystems.

However, *D. gigas* and other cephalopods are difficult to observe in the field over prolonged periods or to study under laboratory conditions. To date no individual *D. gigas* has survived more than 12 days in captivity (Keyl et al., 2008). This thesis addresses these difficulties by developing a computer model of individual traits that can generate predictions of responses to environmental changes at both individual and population level. These predictions can then be verified by comparing them with data from field observations, and necessary adjustments can then be made to the model. The model can then — and this is the overarching aim of the thesis — be used to predict long-term developments affecting *D. gigas* and other components of marine ecosystems in response to changes in environmental marine conditions, including those brought about by human interventions.

2.2. The functional triad migration-maturation-growth

Keyl et al. (2008) proposed the functional triad migration-maturation-growth (fTMMG) as a conceptual explanation for the recent range expansion and the existence of size-at-maturity-groups (SAM-groups). This concept proposes that physiological and energetic

2. Objects of Study

mechanisms are sufficient to explain the migration, maturation and growth of an individual. It discards the notion of genetic differentiation and instead uses the exposure of an individual to exogenous factors during its lifespan to explain the existence of SAM-groups. In the fTMMG, the small and large SAM-groups are the expression of two phenotypic extremes. These represent contrasting life history responses to changes in two basic ecosystem conditions: The small SAM-group is a response to warm periods with low food availability and ensures survival under these conditions, while the large SAM-group maximizes individual fitness in cool periods with high food availability.

The fTMMG provides a simple explanation for the existence of SAM-groups with energy as the driving factor in the life history of *D. gigas* and particularly the optimization of energy use under low energy environmental conditions. The starting point for fTMMG is the high energy demand of the species and the consequent need for a large food uptake. The availability of food changes in line with the changing requirements for food (for example in terms of quality and size) of the rapidly growing individual. In response to these changing food requirements, an individual is obliged to migrate and follow the spatio-temporal occurrence of the preferred prey items.

These migrations create a unique temperature and nutrition history for each individual. Changes to these environmental variables (i.e., temperature and food availability) may also affect the onset of maturation, which in turn determines the SAM and thus the terminal size ML_{terminal} . It is therefore assumed that the different SAM-groups have experienced different environmental conditions over the course of their life history. Rosa et al. (2013) consider that long-distance migration is probably an important element in the life history of *D. gigas* and a cause of differences in SAM.

Consequently, following this line of argument, the combination of high nutrition levels and low temperatures will delay the individual onset of maturation, thus resulting in older individuals with larger ML_{terminal} . The combination of low food availability and high temperatures will bring forward the onset of maturation and will result in younger individuals with smaller ML_{terminal} , in accordance with Arguelles et al. (2001) and Keyl et al. (2008).

Therefore, according to the fTMMG, the small SAM-group is comprised of individuals that have optimized their energy use in accordance with conditions in low energy environments. Keyl et al. (2008) question this conclusion and argue that these individuals would be disadvantaged since, despite their higher relative energy investment (i.e., per unit of growth), they would have a lower reproductive output and lack access to larger prey items because of their size. However, smaller individuals invest less energy in absolute terms into reproduction, which is beneficial under conditions of low food

2.3. Transferring mechanistic models to simulation models

availability, and have access to more numerous and more diverse food items at lower trophic levels (Keyl et al., 2008). Furthermore, the shorter generation time of the small SAM-group leads to a higher generational turnover and therefore higher reproductive output of the whole population per unit of time (Keyl et al., 2008), which compensates for the lower individual reproductive output. Due to the high generational turnover, the predominance of smaller individuals decreases the mean trophic level of the population, which is able to feed on the more productive lower trophic levels (Keyl et al., 2008).

However, Tafur et al. (2010) demonstrate that there is only a loose, if any, coupling of age and ML_{terminal} . This contrasts with the line of argument in Keyl (2009) that postulates energy optimization as the reason for the adoption of life history strategies that give rise to existence of small SAM groups with short generation times. Furthermore, observed sex ratio changes are also expressions of an energy optimization strategy (Rosa et al., 2013; Tafur et al., 2010), but the fTMMG cannot explain this phenomenon because it does not distinguish between the sexes and possible sex-related differences in energy budgets.

In summary, the qualitative formulation of the fTMMG does not allow a definite conclusion to be drawn regarding the existence of different SAM groups in *D. gigas*. The operational mechanism of energy optimization remains open to discussion and there are a number of explanatory gaps. The modeling of the actual energy budgets of an individual and a population would help to fill these gaps and provide insights into the effects of energy optimization. Specifically, the conceptual foundations of fTMMG could be strengthened by using a computational model to generate quantitative estimates of energy flows under different environmental conditions. Such a simulation model would formalize the qualitative relations of the original concept and, by computing energy flows, allow the validity of the hypothesis to be tested. To this end, the underlying principles of energy optimization in the fTMMG are discussed, further developed, and reformulated as the energy driven life history model (EDLHM) in Chapter 3.

2.3. Transferring mechanistic models to simulation models

Information about *D. gigas* is scarce because it is difficult to make observations of the species under controlled conditions. Short-term experiments and observations, such as metabolic rate experiments (Trueblood & Seibel, 2013), may be performed, but in fact much basic information about the species remains inaccessible and theories cannot be directly tested. This thesis develops a mechanistic model of *D. gigas* and postulates that this is a feasible approach towards filling the experimental gap.

2. *Objects of Study*

In a mechanistic model, all parameters and variables have a biological definition and the model specifies hypothesized relationships among these parameters and variables that simulate biological processes (see Bolker (2008)). Mechanistic and phenomenological model form a dualism, in which phenomenological models focus on observed patterns in data and use these patterns (knowledge) to generate formulas that match known data, while mechanistic models are “more concerned with underlying processes, using functions and distributions based on theoretical expectations” (Bolker, 2008). Comparing mechanistic and phenomenological models, Bolker (2008) states:

All other things being equal, mechanistic models are more powerful since they tell you about the underlying processes driving patterns. They are more likely to work correctly when extrapolating beyond the observed conditions. Finally, by making more assumptions, they allow you to extract more information from your data — with the risk of making the wrong assumptions.

The translation of a mechanistic model into programming language to create a simulation model is a decisive step that is independent of the mechanistic validity of the model. A (mechanistic) model is first developed and composed using natural language, sketches, diagrams and other explanatory devices. The next step is to translate them into, for example, a text-based programming language or a simulation environment to create the simulation model. This procedure gives rise to a mechanistic model and the simulation model incorporating different sets of expressions and features that most likely do not match. However, in order to provide useful results, the features of interest and expressions the mechanistic model must be reproduced by the simulation model. The discrepancy between the two sets of expressions is a thorny problem that cannot be easily solved by testing the model. Testing, in this case software verification, is generally possible in principle. However, since the verification process is not fully automated it takes a long time and requires many skills. In practice, verification is limited to certain fields of application and/or models incorporating a limited degree of complexity.

In order to avoid a lengthy verification procedure, there are ways to simplify the translation process from a mechanistic model into a simulation model by using similar structures for both types of model. This thesis explores these options and proposes some guidelines for future simulation models (see Section 9.4ff). Additional verification issues, such as the validation and boundaries of the mechanistic model, are unaffected by the procedure adopted for translation process of the mechanistic model to an equivalent simulation model in another “language expression”.

3. Elements of an energy driven life history for *D. gigas*

3.1. Introduction

The functional triad migration-maturation-growth (fTMMG) is a conceptual model that shows how physiological and energetic processes can explain both phenotypic characteristics of *D. gigas* at the individual and population level, and its migrational behavior. According to the fTMMG, environmental factors that control the endogenous traits of a *D. gigas* individual during its lifetime, rather than genetic differences, explain differences in individual size-at-maturity (SAM) and the existence of SAM-groups. At the population level, the fTMMG explains the recent range expansion of *D. gigas* as the result of migratory behavior driven by nutritional requirements and favorable, supportive environmental conditions in the newly colonized areas. However, the qualitative formulation the fTMMG does not allow evaluation of the interaction of environmental factors with endogenous energy-optimizing traits, such as cannibalism, sex-specific maturation, size reduction and a changing sex ratio.

In this chapter, the individual-level analysis of the fTMMG is re-conceptualized in a form that can be used to develop a simulation model. The essential elements of the new concept are identified to set the scene for further discussion and analysis in the following chapters.

Section 3.2 describes the basic design, implications and limitations of the fTMMG. The fTMMG is reformulated as a new, internally consistent model derived from General Systems Theory (GST), the “energy driven life history model” (EDLHM), which provides the basis for the development of a simulation model for evaluation of energy effects at the individual, school, cohort and population level. The fTMMG was originally formulated in Keyl et al. (2008) but revised versions of some of the chapters in this thesis were published later as articles, and description of the fTMMG presented in this thesis is based on on these updated articles.

Section 3.3 lists individual-level factors to be assessed, and Section 3.4 lists the factors and their effects that emerge from the interaction of individuals, at school, cohort and population level.

3.2. Developing the energy driven life history model (EDLHM)

3.2.1. The functional triad migration-maturation-growth

The functional triad migration-maturation-growth (fTMMG) (Keyl et al., 2008) (Figure 3.1) connects the migration, maturation and growth of an individual *D. gigas* through a feedback system with a physiological and energetic basis. The fTMMG identifies energy as the only driving factor required to explain the life history of *D. gigas* and its energy optimization under low-energy environmental conditions.

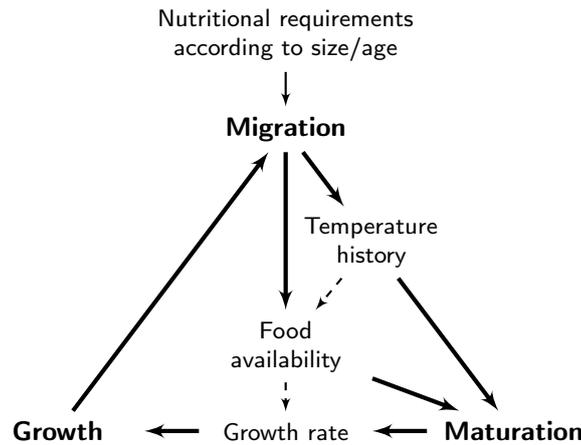


Figure 3.1.: The functional triad migration-maturation-growth (fTMMG) by Keyl et al. (2008) describes a feedback cycle.

The fTMMG takes the high energy demand of the species and thus the need for large food uptake by an individual *D. gigas* as its starting point. The individual grows rapidly and, as it grows, the quality and size of food required changes. The individual undertakes Migration in response to the Nutritional requirements according to size/age and searches for preferred prey items, whose availability is determined by their spatio-temporal distribution, i.e., squid in general respond quickly to changing sea surface temperature (SST) and move if necessary to new areas where SSTs support greater prey abundance (Sanchez et al., 2020; Yu, Chen, & Liu, 2021). Alegre et al. (2014) suggest that migratory behavior is driven by food (i.e., energy) uptake requirements and may occur because *D. gigas* actively seeks out more energetic prey items, rather than simply relying on those that are available. Liu et al. (2016) consider food availability as the “most fundamental factor influencing the active movement of cephalopods”. Migration incurs increased energetic needs caused by locomotion, which

3.2. Developing the energy driven life history model (EDLHM)

in turn determine the different locations with specific local temperatures and food availability that are accessible to *D. gigas* over time.

In the fTMMG, these local temperatures generate a factor Temperature history that controls the Food availability, e.g., cold waters are more productive than warm waters (Tafur et al., 2010; Yu et al., 2015). Thus the migration pattern in the conceptual feedback system of the fTMMG creates a unique temperature and nutrition history for an individual, which in turn may govern the onset of maturation. Rosa et al. (2013) consider long-distance migration as an important element in the life history of *D. gigas* and associate different patterns of migratory behavior with different SAM. In the fTMMG, the Food availability in combination with the factor Temperature history determine the onset of Maturation. High nutrition levels and low temperatures delay the individual onset of maturation and thus result in older individuals with larger ML_{terminal} . The combination of low food availability and high temperatures advances the onset of maturation and results in younger individuals with smaller ML_{terminal} (Arguelles et al., 2001; Frawley et al., 2019; Keyl et al., 2008; Portner et al., 2020).

Maturation and Food availability control the Growth rate, which in turn controls Growth. The parameter Growth controls the onset of Maturation and closes the feedback cycle.

At the population level, it is presumed that each SAM-group experiences similar environmental conditions. According to the fTMMG (Keyl et al., 2008), the SAM-groups are phenotypic expressions of energy optimization and adaption to the experienced environment. Specifically, the small SAM group ensures population survival under adverse energy conditions, while the large SAM group maximizes individual fitness in high-energy environments. Other authors support this explanatory approach, but focus on the effects of temperature on, for example, the caloric content and availability of food and thus on growth and maturation to explain the existence of SAM-groups (Arkhipkin et al., 2014; Frawley et al., 2019; Portner et al., 2020). Chen et al. (2020) note that “[...] an optimal trade-off between investment in reproduction and somatic growth has been found to maximize reproductive success, and ultimately determine population size and stability over time”.

3.2.2. The reformulation at the individual level

The reformulation differentiates between exogenous and endogenous factors, whereby an individual is considered exposed to the exogenous factors temperature (Temperature history) or food availability (Food availability), which are outside of its control. An individual controls the endogenous factors Migration, Maturation, Growth rate, Growth and Nutritional requirements of the fTMMG, but these factors are the result of responses to

3. Elements of an energy driven life history for *D. gigas*

the exogenous factors. A location change links to a new temporal set of local exogenous factor values, which in turn affect the endogenous factors.

The new energy driven life history model (EDLHM) requires quantifiable factors that form a coherent back-coupling or feedback system. In such a system, an *observed variable* receives information from the past and the present and, using these inputs, employs a *specified mechanism* to calculate a *future value*. The feedback mechanism controls the variable and amplifies (positive feedback) or inhibits (negative feedback) the observed dynamics.

In the fTMMG, the parameter Nutritional requirements according to size/age is not part of the feedback system and, since age is treated as an independent variable, does not influence its behavior. These nutritional requirements describe the general quality of the prey item in terms of size, which are presumed to be controlled by Growth rate and Growth. By contrast, in the EDLHM, the factors Migration, Maturation, Growth rate and Growth contribute to a quantifiable energy budget of an individual and are coupled to the feedback system via the parameter Food availability: Migration generates a quantifiable energy need for locomotion, Maturation a quantifiable energy need for (additional) growth of gonadic tissue and Growth and Growth rate a quantifiable energy need for increasing the size and thus the body mass of the individual. Given the opportunistic predation behavior of *D. gigas*, the Food availability is assumed to be equivalent to the quantifiable energy uptake, which contributes the necessary energy input in an energy balance model.

In the fTMMG, the factors Temperature history and Migration determine the spatio-temporal structure of the system. The parameter Migration determines a location, which in turn defines the local environmental conditions, i.e., the temperature. However, the Temperature history is defined as a list of past values. It thus has a different “integration level” in the feedback system compared to other parameters such Growth that define a current value. For this reason, Temperature history is unsuitable for use in a quantitative model of system behavior. In addition, the exogenous parameter Food availability depends on local environmental factors such as temperature rather than an individual Temperature history that most likely does not match the temperature history of a location. By contrast, as explained above, in the EDLHM, local environmental conditions are specified by the exogenous factor Food availability — which is in turn influenced by the exogenous factor temperature.

In summary, the reformulated individual level EDLHM, as shown in Figure 3.2, consists of factors that are quantifiable in terms of their contribution to the energy budget, have the same integration level, and form a coherent feedback system. Note

3.2. Developing the energy driven life history model (EDLHM)

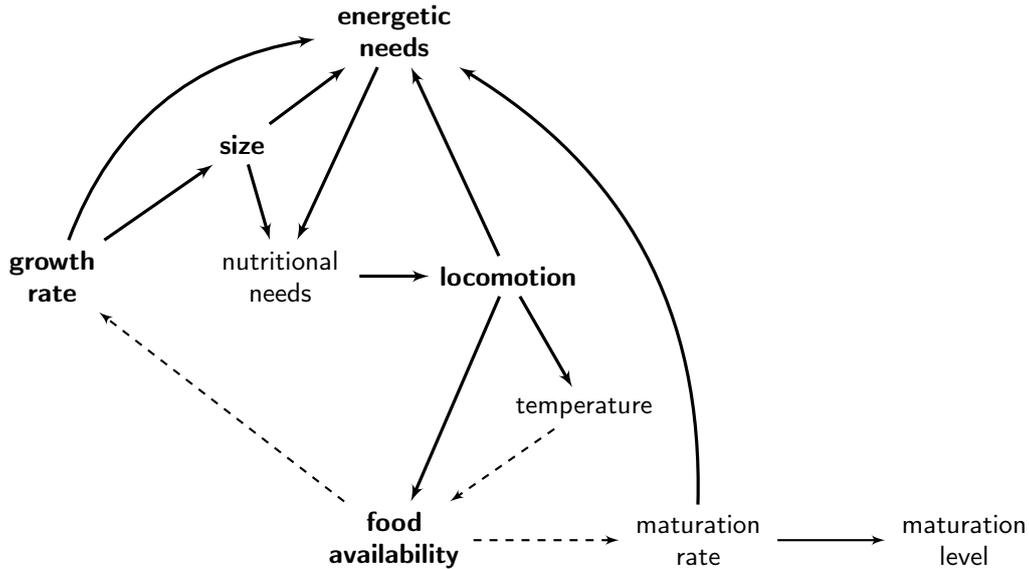


Figure 3.2.: Outline of the energy driven life history model (EDLHM) as a modification to the fTMMG in Figure 3.1. Dashed lines reference impacts of exogenous factors on the targeting factors, the solid ones reference the impact of endogenous factors.

that EDLHM factors are notated with initial lower case, to differentiate them from initial upper case factors of the fTMMG.

The parameter *energetic needs* specifies energy needs of an individual in the EDLHM and acts as the driver of individual behavior. This energy demand must be covered by food uptake with suitable prey items whose energy amount (quantity) and type (quality) is represented by the parameter *nutritional needs*. To fulfil the nutritional needs, the prey items have to be followed and this requires locomotion, which in turn influences the parameter *energetic needs*, forming a feedback loop, as shown in (Figure 3.2).

After a finite time, locomotion sets a (new) location, which in turn determines the spatio-temporal factors *temperature* and *food availability* in the EDLHM. The factor *temperature* is considered to drive the local *food availability*. In the EDLHM, there is no direct link between *temperature* and *growth rate*, i.e., the co-control of growth rate by the temperature as in Keyl (2009) and Pecl and Jackson (2008) is discarded in favor of a growth function that is dominated by physiological processes. This function sets the achievable maximum growth rate based on the current *size* and metabolism, which in turn depends on energy uptake and external factors such as (temperature dependent) oxygen saturation, which may limit metabolism through reduced oxygen availability (Subsection 4.2.2).

3. Elements of an energy driven life history for *D. gigas*

The factor **maturation rate** determines the maturation level that is outside the energy driven feedback loop but is still considered to limit the lifetime. However, the mantle length (ML) analysis in Tafur et al. (2010), one of the main articles on which the fTMMG is based, describes a sex-specific maturing pattern, with males maturing in the middle and females towards their end of lifespan. Instead, the end of lifetime of the males may be described after the period of mating of the cohort. The energy aspects of maturation rate and maturation level in the EDLHM are discussed in Subsection 4.2.2 on page 54.

The **growth rate** is modeled so that it depends only on food availability, i.e., energy uptake. The **temperature** is presumed to have an impact (Keyl, 2009; Pecl & Jackson, 2008) but the exact contribution is unknown (Portner et al., 2020); temperature is thus not part of feedback system described by the EDLHM. If necessary, the **temperature** may indirectly control the **growth rate** as a result of its effect on food availability (Frawley et al., 2019; Portner et al., 2020).

The feedback loop is closed by adding the factor **size**, which is determined by the energy controlled **growth rate** and contributes to the **energetic needs** and **nutritional needs**. The **size** determines the energy required to maintain the basal metabolic rate (basal_{mr}) and sets **nutritional needs**, for example required prey item size by the window of predation (WOP).

This model outline specifics all elements of the model and the relations among them. However, parameterization of these elements is still required before computer simulations using individual-based modeling (IBM) (Section 4.2) can be performed (as described in Chapter 6).

3.2.3. The effects at population level

The implications of the fTMMG are given by the formulation of different energy optimization strategies for *D. gigas*. At the population level, the energy-driven life history may allow a strategy that optimizes the response to environmental conditions, i.e., available energy. Specifically, such a strategy would prioritize individual survival by reducing energy demand (leading to small SAM) under low energy conditions (Frawley et al., 2019; Portner et al., 2020) and maximize fitness to increase fertility (the consequence of large SAM) under energy-supporting conditions (Keyl et al., 2008).

The adoption by *D. gigas* such an optimization strategy rests on certain assumptions, which must be fulfilled: (1) *D. gigas* must inherently have the phenotypic ability to express different SAM without genetic differentiation; (2) the small- and large-sized specimens mark the lower and upper bounds of phenotypic expressible sizes; and (3)

3.3. The energy driving elements at the individual level

this *inherent endogenous phenotypic plasticity* is controlled by exogenous factors such as temperature, daylight, nutrition and others.

All three assumptions are likely to be fulfilled, as different SAM-groups are observed, but there is no evidence of genetic differentiation between them. The sizes of observed specimens range within definable upper and lower ML_{terminal} limits. Several exogenous factors are believed to control the maturation process and terminal size, apparently based on responses to (unknown) thresholds; although responses to combinations of thresholds are non-linear and difficult to predict (Keyl, 2009).

It is noteworthy that optimal energy use by adapting SAM at the individual level most probably scales up to fitness maximization at the population level. Other observed life history traits that are presumed to be responses to exogenous drivers, such as changes to the sex ratio, cannibalism, and sex-specific maturation rates, may also serve to maximize population-level fitness. Again, the qualitative formulation of the fTMMG does not allow the evaluation of these probable optimization strategies, but the EDLHM does.

Furthermore, the explanatory, qualitative approach of the fTMMG cannot be used to assess the effects of feedback processes that become apparent when time dependencies are considered. For example, energy depleted environments shorten the time until maturation and reproduction, which reduces the amount of energy consumed during the reproductive cycle and, as such, could be considered beneficial from a survival point of view. However, the shortened lifespan increases the population turnover rate (Keyl et al., 2008), which can counteract energy-saving effects by increasing population size. The EDLHM enables analysis of population-level effects of this and other strategies to optimize energy use, by computing energetic needs at the individual and population level.

3.3. The energy driving elements at the individual level

The factors energetic needs, size, locomotion, growth rate and food availability control the direction and rate of energy flows; the remaining factors, i.e., nutritional needs, temperature, maturation rate and maturation level (Figure 3.2), are not considered to affect the energy flow or are duplicated by other factors.

The temperature may influence the growth rate, but its effect may vary at different ontological stages (Arkhipkin et al., 2014). For example, eggs and paralarvae may be more influenced than juveniles or adults by temperature changes, since these more rapidly affect the metabolism of smaller individuals. Thus temperatures at the so far

3. Elements of an energy driven life history for *D. gigas*

unidentified spawning grounds would be an important component of the Temperature history of the fTMMG. This issue is discussed in Chapter 5. In this thesis, evaluation and discussion of the factor temperature are limited to its role as an exogenous factor.

The growth rate is an influencing factor in the EDLHM, and will depend on the growth function that is used. Sex-specific growth functions are likely to be required to obtain the observed dimorphic ML_{terminal} . These will lead to different patterns of growth, which in turn may influence size-dependent factors such as energy demand and, of course, cannibalism. Section 6.2 discusses the growth function and how it affects other life history traits, e.g., food uptake and growth rate capacity (Subsection 6.3.2).

The driving factor energetic needs, which specifies the energy debt that needs to be satisfied by food uptake, and is considered in the EDLHM to be equivalent to the parameter food availability, is defined in Subsection 6.3.1.

Failure to meet factor energetic needs has an effect on growth rate and consequently on size. In the longer term, size affects energetic needs, as indicated by the feedback loop in Figure 3.2. The effects of energy constraints on the growth rate and size are evaluated in Subsection 6.3.2. Subsection 6.3.3 discusses the effects of temperature on the oxygen saturation required to metabolize the food uptake into growth.

The parameter nutritional needs defines the quantity of food required to satisfy the energetic needs, as well as aspects of food quality, e.g., the correspondence between WOP and available prey items, that in turn control the locomotion. However since, in the EDLHM, the factor food availability covers the energetic aspects of the food, the factor nutritional needs is excluded from further evaluation.

The factor locomotion controls the location of an individual, which in turn controls food availability and temperature, which are both considered as exogenous factors and, if necessary, could be calculated by an additional model. The parameter locomotion generates an energy requirement which is estimated and added to the factor energetic needs (Subsection 6.3.1).

The maturation rate is excluded from the evaluation since its relationship to temperature and food availability is still unknown (Keyl et al., 2008; Portner et al., 2020). Tafur et al. (2010) propose a sex-specific maturation process, in which different stages of maturity occur at fixed points during the life history and might not be triggered by any further external factors. Consequently, maturation rate and maturation level are excluded from further evaluation.

The factor food availability maps the individual's energy uptake, since it is assumed an individual exploits the factor food availability to the available maximum or less, e.g., only

3.4. *The driving effects at school, cohort and population level in the literature*

the individual's needs. Since energy uptake is considered to be equal to the factor food availability, no separate evaluation of factor food availability is conducted.

3.4. The driving effects at school, cohort and population level in the literature

Section 1.1 proposes additional types of energy-driven behavior that emerge from the interaction of individuals searching for food. These emerging effects are discussed below in relation to energy optimization and thus as probable driving elements in the life history of *D. gigas* at school, cohort and population level. The prior quantification of driving elements at individual level enables these further effects to be computed.

3.4.1. Decoupling energy demands of males and females by sex-specific maturation

Tafur et al. (2010) describe sex-specific patterns of maturation, with males maturing in the middle of their lifespan at approx. half of their terminal ML, and females shortly before the end of their lifespan, but in both cases during their exponential growth phase. The energy-intensive gonadosomatic growth follows an s-shaped growth function for males and a j-shaped growth function for females. Thus if their life expectancy is the same, males mature before females. Since it is assumed that growth of reproductive tissue requires additional energy (Tafur et al., 2010), energy requirements peak at different points in the life history of males and females. Taken together, energy requirements of males and females determine the energy demand of a cohort.

Furthermore, mature males store their spermatophores in the immature females and die shortly after reproduction. This removes males from competition for energy with females, giving the females access to additional energy resources for their spawning efforts. The effects of sex-specific maturation proposed by Tafur et al. (2010) are examined in Section 8.2.

3.4.2. Size-at-maturity energy optimization

The small and large SAM-groups are assumed to express two opposite life history responses to two basic ecosystem conditions (see Section 3.2). This hypothesized response to temporal variation in environmental productivity is substantiated by ML analyses of female *D. gigas* individuals in Peruvian waters between 1989 and 2004 (Argüelles et al., 2008) and the Peruvian Exclusive Economic Zone between 1991 and 2007 (Tafur et al., 2010). Both studies report that the mean ML of mature individuals

3. Elements of an energy driven life history for *D. gigas*

increased with prevalent cooler water conditions towards the end of the observation period.

Medium SAM individuals dominated *D. gigas* populations in Peruvian waters between 1989 and 1999; mainly medium and large SAM individuals were present during a transition period in 2000; after 2001 the large SAM-group prevailed and a dramatic increase in mature female ML was observed between 1989 and 2004 (Argüelles et al., 2008). Tafur et al. (2010) observed the same pattern, with a transition period between 1999 and 2000, separating the medium SAM individuals prior to this period from the large SAM individuals that predominated thereafter. It is worth mentioning that the “small” SAM-group individuals in Tafur et al. (2010) actually belong to the medium SAM-group according to Table 2.5. A similar transition period was found for the $ML_{50\%}$ -group, i.e., the ML at which 50% of the individuals are mature, between the austral winter 1999 and the austral summer 2001 (Tafur et al., 2010).

In the year 1999, the strong El Niño (EN) event of 1997/1998 was followed by a small increase in the ML, which in 2000 dropped to the level during the EN event of 1997/1998, but then from 2001 significantly higher ML were observed (Argüelles et al., 2008, Fig. 2). However an EN event is expected to decrease the ML due to the occurrence of less productive warm waters (Frawley et al., 2019; Portner et al., 2020), but the ML was already at a similar level before the EN event of 1997/1998 (Argüelles et al., 2008, Fig. 2). Individuals of the northern hemisphere generally belong to the large SAM-group. However, in 1997–1998 their SAM dropped to the same level as in the Peruvian waters during early 1990 (Tafur et al., 2010); similarly, the SAM in the Gulf of California (GOC) shifted from large to medium three years after the 1997/1998 EN event.

The different life histories described link the occurrence of SAM-groups to temporal variations in environmental productivity and food availability (Argüelles et al., 2008; Frawley et al., 2019; Keyl et al., 2011; Keyl et al., 2008; Portner et al., 2020; Tafur et al., 2010). This evidence counters both the hypothesis of three spatially distinct populations proposed by C. Nigmatullin et al. (2001) and the idea that SAM-groups have a genetic origin. However, there is no evidence of a direct correlation between SST and SAM (Argüelles et al., 2008; Portner et al., 2020), nor of a significant relationship between SST anomalies and mean growth rate, longevity, or maximum cohort size (Arkhipkin et al., 2014; Keyl et al., 2011).

The energy-saving benefits of reduced SAM can be estimated by calculating the overall energy requirement for terminal size and a given lifespan, as discussed in Section 8.3. This discussion draws on the observation in Keyl et al. (2008) that small SAM individuals

3.4. *The driving effects at school, cohort and population level in the literature*

and short generation times are associated with low-energy environments. Alternative explanations for the occurrence of SAM-groups are discussed in Section 6.4.

3.4.3. **The sex ratio change for energy optimization at population level**

Tafur et al. (2010) hypothesize that changing sex ratios are a means of energy optimization at the population level. It is assumed that males do not contribute to reproduction, but compete for resources with the females, who are considered to be the active reproductive part of the population. Males, however, provide additional genetic diversification in a population. A mixed paternity of the progeny, as reported for some cephalopods (Boyle & Rodhouse, 2005), could be a strategy to maintain high genetic diversification, even if the number of males is reduced to reduce the overall investment of energy per progeny (Tafur et al., 2010).

Tafur et al. (2010) report a high monthly variation in the female:male sex ratio, with more females than males except during a few very short periods, ranging from 1:1 to 24:1. In the Peruvian Exclusive Economic Zone, the mean female:male ratio was $\approx 3:11$ between the 1990s and about 2002, with the proportion of females slowly declining to a ratio of $\approx 2:1$ (mean value) after 2002 (Tafur et al., 2010). This decline in the proportion of females coincides with an increase in food availability (see Keyl et al. (2008)). In general, changes in sex ratio may be the result of changes in the reproductive pattern to cope with changes in environmental productivity (Rosa et al., 2013).

In addition, the sex ratio changes during the reproduction period due to female cannibalism of males (Ibáñez & Keyl, 2010), presumably in response to females' higher energy demand during reproduction. Size dimorphism makes it easier for females to cannibalize the smaller male individuals; while the delayed maturation of females may be a means of ensuring that the higher energy demands of females during reproduction can be satisfied by devouring males. At an individual level, cannibalism is a strategy to increase energy uptake; it might also be a strategy for energy optimization at the population level that is enabled by this individual trait.

The hypothesis that males are more energy expensive than females is tested in Section 8.5. The overall energy requirements and the daily energy requirements of males and females are compared to obtain estimation of the overall energy costs of the males over their lifespan in comparison to the females under different growth functions. Cannibalism as a possible tool to control the sex ratio is discussed in Section 7.3 and evaluated using simulation runs in Section 7.7.

3. Elements of an energy driven life history for *D. gigas*

3.4.4. Cannibalism

Cannibalism is a multifaceted behavioral trait that affects other driving factors. The rate of cannibalism in *D. gigas* is considered high, with conspecifics reported to represent up to 75% of the weight of stomach contents (Keyl et al., 2008). However, these high cannibalism rates might be induced by fishing stress (Bruno et al., 2021; Ibáñez et al., 2008; Keyl et al., 2008; Markaida & Sosa-Nishizaki, 2003; Rosas-Luis & Chompoy-Salazar, 2016) or the effects of fishing operations, or be the result of sampling bias, since larger specimens of a population, which display higher rates of cannibalism (Ibáñez & Keyl, 2010), are more likely to be captured (Keyl et al., 2008). Cannibalism is examined in Chapter 7 with a focus on its role as survival strategy, drawing on discussions of this topic in the literature.

At the individual level, *D. gigas* has an energy driven life history characterized by high energy demand (Keyl et al., 2008) combined with limited energy buffering capability (Ibáñez & Keyl, 2010). At the population level, low feeding levels may lead to starvation and endanger the survival of the population, which, since *D. gigas* as a semelparous species, depends on the reproductive success of the current generation (Pecl & Jackson, 2008). From this perspective, cannibalism could represent a population-level energy storage strategy, in which energy is built up under favorable conditions, i.e., stored by smaller individuals, as a reserve to be drawn on by larger individuals through the consumption of these smaller individuals under conditions of ecological stress (Keyl et al., 2008). This process would increase the fitness of the mostly larger individuals that are known to engage in cannibalism (Ibáñez & Keyl, 2010); however if large individuals predate on younger ones this might endanger the species' survival, but Chen et al. (2020) note that life history theory predicts that individuals should trade off energy allocation for reproduction against energy allocation for somatic growth or survival to maximize lifetime reproduction success.

The role of cannibalism in bringing about a changing in the sex ratio in response to reduced availability of energy has already been mentioned. However cannibalism may also change the sex ratio in response to females' increased energy demand during reproduction. Ibáñez and Keyl (2010) state that "in [the] case of *D. gigas*, cannibalism is intensified during reproductive period". Sexual dimorphism, in which females are larger than males, combined with the strong correlation between size and cannibalism, i.e., larger individuals consume smaller conspecifics (Ibáñez & Keyl, 2010), probably also contributes to reducing the number of males in the population.

The hypothesis of cannibalism as a survival strategy is evaluated and the different effects of cannibalism are examined in Chapter 7. The hypothesis that size dimorphism

3.4. The driving effects at school, cohort and population level in the literature

in favor of females may promote such a strategy is tested in Section 7.3. The aspect of energy uptake is examined in Section 7.4 and its effects on the partial survival of a school (Ibáñez & Keyl, 2010). The simulation runs in Section 7.6 and Section 7.7 complement the static analysis in previous sections.

4. Modeling the individual and its environment

4.1. Introduction

This chapter prepares the ground for the implementation of a *D. gigas* simulation model by defining the general relations of a model of a *D. gigas* individual. For an explanatory approach to an energy driven life history of *D. gigas*, the development of an energy model with quantifiable parameters is of utmost importance. Such a model must enable the estimation of the energy needs and their satisfaction by converting body mass into energy in times of food deprivation; this is of key importance because of the high energy needs of *D. gigas* and its lack of energy buffering capacity.

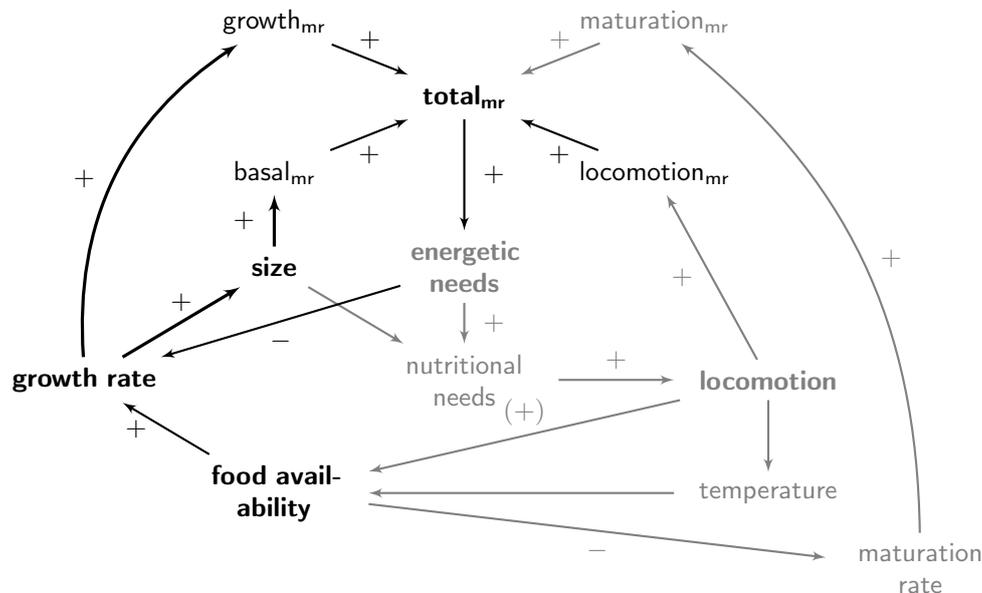


Figure 4.1.: The cause-effect diagram of the energy driven life history model (EDLHM) for *D. gigas* in this thesis. Elements shown in grey are removed or substituted for by other elements in the simulation model. The subscripted suffix “mr” stands for metabolic rate.

In this chapter, the previously identified energy driving elements of the outlined energy driven life history model (EDLHM) (Section 3.2) are further developed into a model with quantifiable factors (Figure 4.1) described using formulas and algorithms,

4. Modeling the individual and its environment

i.e., as computable elements that can be incorporated into a simulation model of an energy driven life history. As described below, an important aspect of the development process is the simplification of the model, reducing the number of components in order to facilitate the systematic evaluation of the driving factors in simulation runs.

In Section 4.2 the general cause–effect relations are developed by removing redundant factors or substitution of equivalent factors of the outline EDLHM. The energy model is based on the fish energy balance model (Winberg, 1960), which requires the specification of energy needs and energy input. Here, energetic needs are defined by modeling the individual total metabolic rate (total_{mr}) as the sum of energetic requirements of activities identified in the outline EDLHM. Energy uptake is reduced to the food availability, i.e., it is assumed that an individual *D. gigas* will consume all of the available food, and therefore food availability regulates energy uptake, as explained in the previous chapter. The value of food availability is set, instead of being retrieved by an additional model. This procedure excludes redundant input factors from the simulation model.

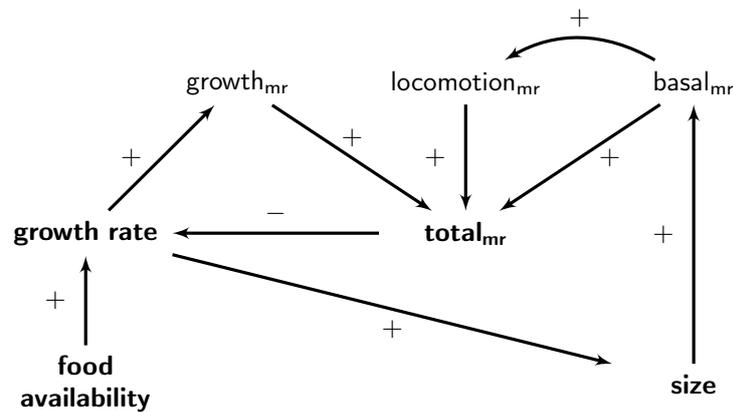


Figure 4.2.: The cause–effect diagram of the simulation model for an energy driven life history model (EDLHM) of *D. gigas*. The subscripted suffix “mr” stands for metabolic rate.

Section 4.3 defines the general relations linking an individual’s age, growth function, size, volume and surface area. This set of formulas specifies the energy balance model, including the conversion of body mass into energy and vice versa (Section 4.4), for subsequent implementation as a simulation model.

The relationships between factors in the model are summarized in the cause–effect diagram in Figure 4.2. This shows that food availability regulates the energy input while total_{mr} (the subscripted suffix “mr” stands for metabolic rate), generates the energy

needs. The balance of food availability and total_{mr} controls the growth rate; this in turn defines the size, whose development over time can be shown as a growth curve.

4.2. Modeling a *D. gigas* individual and its environment

4.2.1. Overview

The simulation model consists of *D. gigas* individuals and an environment. A *top-down* approach is used to identify a factor’s incoming parameters and factors. Values or formulas are then defined to quantify these incoming factors. These endogenous and exogenous factors require different modeling approaches, for example agent-based-modeling (ABM) for the individuals with their internal states, and a cellular automaton (CA) for the environmental spatial information.

Subsection 4.2.2 describes the modeling of a *D. gigas* individual, Subsection 4.2.3 modeling of an environment, and Subsection 4.2.4 modeling of more complex individual behavior, i.e., prey seeking and swarming. Section 4.3 specifies the relations identified in the models by supplying the necessary formulas.

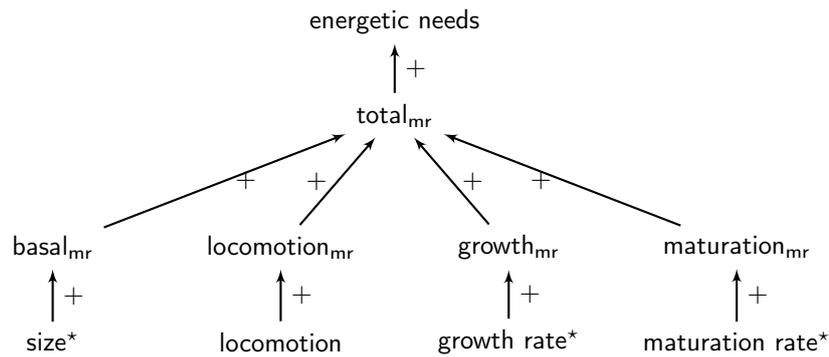


Figure 4.3.: Components of individual energy needs. The index “b” is “basal”, “lm” the “locomotion”, “gr” the “growth rate” and “mat” the “maturation”. An asterisk denotes a composite factor and the subscripted suffix “mr” stands for metabolic rate.

4.2.2. Modeling a *D. gigas* individual

Modeling the energetic needs. In the EDLHM, the total_{mr} (Figure 4.3) is equivalent to the factor energetic needs that drives the model by generating an energy deficit that must be fulfilled by energy (food) uptake, which in the model is equivalent to the factor food availability (see Section 3.2). The total_{mr} is calculated as the sum of the factors

4. Modeling the individual and its environment

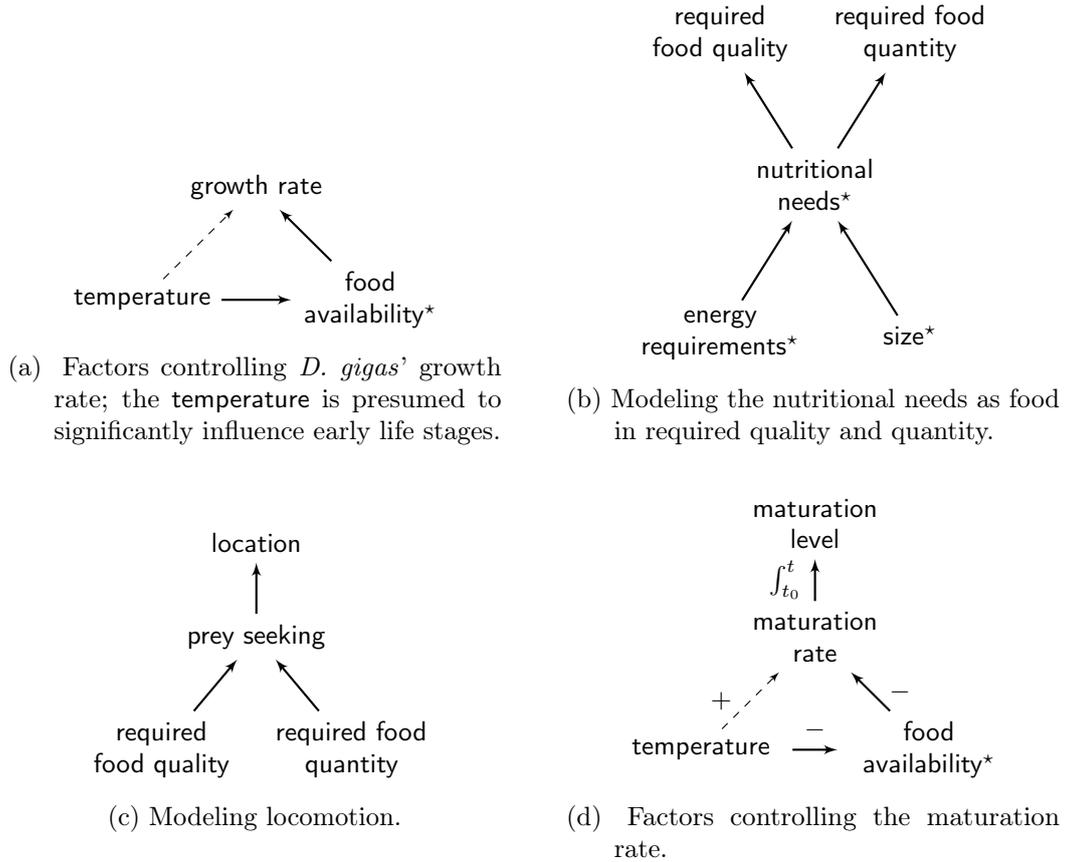


Figure 4.4.: Influences on the growth rate, locomotion, maturation rate and maturation level; an asterisk denotes a composite factor.

basal metabolic rate (basal_{mr}) given by the size-dependent body mass (Subsection 4.3.3), locomotion metabolic rate ($\text{locomotion}_{\text{mr}}$), growth metabolic rate ($\text{growth}_{\text{mr}}$) and maturation metabolic rate ($\text{maturation}_{\text{mr}}$):

$$\text{total}_{\text{mr}} = \text{basal}_{\text{mr}} + \text{locomotion}_{\text{mr}} + \text{growth}_{\text{mr}} + \text{maturation}_{\text{mr}} \quad (4.1)$$

Modeling the size. In each simulation time step, the new size size' is determined by multiplying the current size by the growth rate:

$$\text{size}' = \text{size} \cdot \text{growth rate} \quad (4.2)$$

4.2. Modeling a D. gigas individual and its environment

Modeling the growth rate. A growth function $*f_{gr}$ (Subsection 4.3.2) models the factor growth rate and sets the achievable growth rate for each age until the next simulation step:

$$\text{growth rate} = *f_{gr}(\text{age}) \quad (4.3)$$

Alternatively, the achievable growth rate may be determined by the current mantle length (ML) (since factor growth rate is a function of the ML):

$$\text{growth rate} = *f_{gr}(\text{ML}) \quad (4.4)$$

Instead of a pre-set age-related growth, the relation in Equation (4.4) may model growth more realistically as being based on physiological processes and constraints, for example metabolism, oxygen limitation, oxygen transport, maximum cell growth etc.

The single incoming factor food availability limits the growth rate, since this is dependent on the available energy. This is why the function $*f_{gr}$ does not define the actual growth rate of an individual, but rather the maximum growth rate achievable in a simulation step for an individual of a specified age or size. Tafur et al. (2010) report no relation between size-at-maturity (SAM) and age, so the growth rate must depend on other factors, i.e., food-uptake and (indirectly) temperature (Arkhipkin et al., 2014; Frawley et al., 2019; Portner et al., 2020; Sanchez et al., 2020), as well as age, or be independent of age. If age were the single factor determining growth rate, then both age and growth rate would have to be preset to create the observed SAM-groups. Furthermore, the maximum size is achievable under low temperature conditions, so it can be deduced that an elevated temperature does not increase physiological processes that control the growth rate; in fact, higher temperatures may (indirectly) limit the growth rate. Taking account of these considerations, in the model the actual growth rate is controlled by the available energy (and limited by size) and the factor growth rate controls the development of body mass (growth metabolic rate (growth_{mr})).

Modeling locomotion. Locomotion results in a *location* and does not directly affect ontogenetic development. In the model, factor locomotion is mainly undertaken to enable energy uptake by predation and to locate food resources by prey seeking. The new location determines the exogenous factors temperature and food availability. The factor prey seeking is actually an algorithm (sub-model) that determines the direction (vector) to the next location. The factor locomotion generates locomotion-related energy needs, denoted locomotion_{mr} .

4. Modeling the individual and its environment

Modeling the maturation rate and maturation level. In a semelparous species, maturation limits the lifespan and, since *D. gigas* individuals continue to grow through the lifespan, maximum maturation level, i.e., full maturity, should approximately coincide with the ML_{terminal} in the growth function. However, full maturation of males occurs towards the middle of their lifespan (Tafur et al., 2010); thus maturation cannot be used as an indicator for the end of their lifespan. Instead, a suitable indicator for males might be the time at which mating occurs and for the females the end of the spawning period.

The growth rate was described in the previous text as a maximum (available) growth rate limited by external factors, i.e., availability of energy. In males, the process of maturation and thus the growth of the reproductive organs begins approximately in the middle of the lifetime (Tafur et al., 2010). To achieve maturity, gonadic growth must be prioritized. There are (at least) two ways in which these energy needs can be met: (1) the maturation rate, the progress of which in turn defines the maturation level, distributes the available energy (food availability), prioritizing the growth of reproductive tissue over ML growth, but without elevating overall energy needs, leading to a decreased growth rate when there is insufficient energy; or (2) the reproductive tissue grows in addition to the ML growth, thereby elevating the energy requirements.

Tafur et al. (2010) considers “[...] any increase in the already high feeding rate [...]” as difficult, so gonadic growth at the expense of ML should be expected. In this case, growth is impaired, and this impairment would have sex-specific effects. However, Chen et al. (2020) suggests a “mixed income-capital breeding strategy”, whereby a limited somatic reserve might be available to support gonadic growth but the energy for reproduction is mainly derived from food uptake.

Females mature to the end lifespan shortly before attaining their ML_{terminal} . Moreover, the energy requirement for growth is relatively low compared to energy requirements for other purposes (i.e., metabolism and locomotion; see the later discussion in Subsection 6.3.2). Thus, in the case of females, it is likely that the overall effects of maturation on growth are limited and, therefore, the maturation rate is an insignificant factor in the EDLHM.

In contrast, for males, prioritization of maturation in the middle of their lifetime will lead to a noticeable delay in growth if the additional energy demand required by maturation is unavailable. Males have a relatively high growth rate at the onset of maturation (see Subsection 6.3.1), which may be difficult to maintain under low energy conditions.

4.2. Modeling a *D. gigas* individual and its environment

If males are unable to compensate for this growth deficit during maturation, they will eventually have a shorter ML_{terminal} than females; this could explain the sexual dimorphism on an energetic basis.

For the reasons stated above, the EDLHM therefore uses a sex-specific growth function which calculates energy demand (only) from ML and the derived body mass. It does not consider maturation since this contributes only marginally to the total energy requirement.

Modeling nutritional needs. The factor nutritional needs is split into the non-cumulative factors required food quality and required food quantity to define an index value, where quantity is the amount of energy uptake and quality is the type of prey. Type is defined as size, which is limited by the window of predation (WOP). Thus the parameter nutritional needs maps the “abstract” energetic needs to a specific amount and type (size) of prey. The parameter energetic needs ($total_{mr}$) is used to determine the required quantity of food. The definition of quality as size may be an oversimplification because prey items of the same size may provide different amounts of energy (Alegre et al., 2014; Arkhipkin et al., 2014; Portner et al., 2020). However, in the model, for an opportunistic predator, it seems reasonable to use WOP to define the quality of the food.

In practice, the factor nutritional needs can be substituted by a specified food uptake in the simulation model, unless an environment is modeled in detail, in which case detailed calculation of nutritional needs is required.

Modeling the temperature and food availability. The temperature is an environmental, location-dependent, exogenous factor. An implemented environment includes a temperature model that provides the temperature value for any location and time, i.e., a function $f_{\text{temperature}}(\text{time}, \text{location})$.

Similarly, the food availability is a spatio-temporal exogenous environmental factor (Arkhipkin et al., 2014; Frawley et al., 2019; Portner et al., 2020; Sanchez et al., 2020) that contributes information about food availability: $f_{\text{food availability}}(\text{time}, \text{location})$. The abundance of prey items is spatially distributed and depends on the productivity of the local ecosystem, which is considered to depend on the local temperature. In general, colder waters (lower temperature) are considered more productive than warmer waters, i.e., food availability decreases with increasing temperature values, see Figure 4.4a and Figure 4.4d.

The model in this thesis does not require an elaborated model for food availability, since food uptake (energy uptake), which is proportional (up to a set maximum) to

4. Modeling the individual and its environment

food availability, is an altered input variable whose values are manipulated to explore the reaction of the model.

Modeling prey seeking. The optional parameter `prey seeking` enables a modeled individual to locate and move towards its required prey items; `prey seeking` is implemented by “sensing” (scanning) over a limited area and moving towards the location where a prey item is most likely to occur (Subsection 4.2.4). This is a pragmatic approach to prey seeking, implemented over a limited range and assuming “global knowledge” about the occurrence of prey items in the given environment.

The factors `food quality` and `food quantity` of nutritional needs control the process of prey seeking, as prey items that correspond to these factors have to be followed to a new location.

4.2.3. Modeling the environment

Data organization and information access of an individual. In the simulation model, the environment is an information container that stores arbitrary types of individuals and spatio-temporally organized abiotic information in cells or cubes in, respectively, a two-dimensional or three-dimensional environment. The individuals (agents) that are part of the environment interact with each other and with their environment. To compute their internal states, individuals may query the environment for abiotic factor values and available individual interaction options by using their location and the time t_i as input parameters.

Reducing computational effort caused by the spatio-temporal structure and the ABM approach. The computational effort required to determine the state of an environmental at time t_i increases with the numbers of objects and cells/cubes. If the determination of a state $_i$ is associated with a previous state $_k$, where $k < i$, then the previous state $_k$ also has to be computed. This involves (re-)computation of values for all iteratively computed objects that make up the previous state $_k$, in the same way as for a CA. Such iterative processing may be computationally costly and therefore optimization procedures are required.

A one-directional time flow with $t_{i+1} > t_i$ offers considerable potential for optimization in simulation processes, since access to the current t_i values is sufficient to compute values for t_{i+1} . The environmental data is computed using previously computed and saved data from the previous state or, alternatively, taken from existing data; a re-computation of past values is not necessary. If individuals need access to previous (older)

information, such as temperature for a temperature history, then these individuals can save the history themselves to avoid the need for re-computation of past environmental states. However, this requires increased memory storage capacity. There is a trade-off between a higher memory requirement versus computational effort.

The individual-based modeling (IBM) approach enables individuals to interact with other individuals. During a simulation step, each individual must check *all* other objects/individuals to see if any of them are in interaction range. In each time step, theoretically, each of the n individuals performs at least $n-1$ checks, resulting in a total of $n^2 - n \approx n^2$ checks for possible interactions; such quadratic complexity renders simulation runs with many objects costly. The spatial organization of objects within distinct cells (regions) may drastically reduce the effort (Luke, 2019). To find another object for interaction, an object queries only those cells and respectively the objects they organize, that intersect with the object's interaction range. Such an optimization approach is successful if the granularity, the size and number of the organizing cells, and the interaction range of the object are well balanced (Luke, 2019) and reduces the “expensive” n^2 effort.

As another optimization approach, a group of individuals can be merged to a single object, a “super-individual” (Breckling et al., 2006; Gallego, 2011; Grimm & Railsback, 2005; Scheffer et al., 1995), which retains the original characteristics of these individuals. This reduces the computational effort by decreasing the number of individuals, instead of optimizing the interactions to reduce the computational effort required (as described in the previous paragraph). This approach is less trivial because it requires a set of rules for merging individuals, preserving information, group development (of the merged individuals), and for splitting up the merged object to recover the original individuals.

Such super-individuals reduce the number of individuals in the simulation system and thus the number of computationally costly interactions, but this procedure might distort the emergent effects that the IBM approach is intended to reveal.

Modeling environmental information. The temperature model divides the space into a CA-like structure, where the spatio-temporal resolution determines the accuracy and the amount of computational effort required. Historical data, temperature modeling, and a mixture (hybrid) of historical data and modeling can be used for each cell to determine the temperature in a simulation step.

In a hybrid approach, for example, a historical dataset provides temperature data, i.e., data fix points, and the model uses an algorithm to compute the intermediate temperature data. A model allows the extrapolation of data beyond the actual dataset,

4. Modeling the individual and its environment

but this affects the validity of the data. Extrapolation or interpolation reduces accuracy, but historical datasets of large areas are likely to be inaccurate *per se*, since these datasets contain already modeled information, e.g., to fill in gaps in satellite-based raw sea surface temperature (SST) data caused by occluding clouds.

Environmental productivity is spatially organized. It is presumed to be temperature dependent and can be derived from a function $f_{\text{prod}}(\text{time}, \text{temperature})$, calculated using temperature information from the CA. The inclusion of prey items in an IBM significantly increases the number of processed objects and increases the computational costs. Where there are a large number of individual prey items, prey item abundance can be expressed as the statistical probability of occurrence for any given value of environmental productivity.

4.2.4. Modeling complex behavior: Prey seeking and swarming

Prey seeking. In the model, the prey seeking process controls the locomotion of the individual, which “senses” the location of prey items and moves towards them. Such a capability of *D. gigas* is presumed to exist, because random detection of prey items in the vast spaces of marine systems seems unlikely; the actual process is unknown but environmental variables might provide information to *D. gigas* on the probability of higher prey abundance.

Swarm modeling. The generic term *swarm* is used to encompass other terms like schools or shoals. Swarm modeling is the aggregation of *similar* individuals into a swarm, which is represented in the simulation as a single object, similar to the “super-organism” or “super-individual” (Breckling et al., 2006; Gallego, 2011; Grimm & Railsback, 2005; Scheffer et al., 1995), to reduce the computational effort. Such swarms possess the key traits of the represented individuals and allow the deterministic extraction of the component individuals at any time.

The aggregation of similar individuals enables *information compression*, with lossless compression, which allows accurate and complete restoration of the input data (while lossy compression only allows reconstruction a sub-set of the original information with some loss of accuracy).

Swarm modeling involves a shift to a higher level of organizational order, with emergent characteristics that are different to those of the individuals. The behavior of the swarm is different from the behavior of individuals and is described by different rules; it cannot be deduced from the aggregate behavior of individuals. Swarm modeling

provides new types of information, i.e., “meta-information” about the swarm, that must be handled properly.

Swarm modeling rules (Reynolds, 1987) provide simple algorithmic descriptions of swarming behavior such as collision avoidance, coordinate locomotion, joining of individuals etc. But these rules do not describe the underlying mechanisms for the formation of an intra-species swarm, such as the recognition by an individual of other individuals its own kind. This is achieved in the model by assuming that continuously operating imprinting processes require swarm partners to stay together. The merging of swarms happens only if both swarms are similar, i.e., comprise the same or very similar species and individuals of approximate the same size. As long as an individual cannot inspect itself and has no capacity of self-awareness, the information about its own kind must be provided, e.g., by genetic coding, imprinting, a bias towards association with objects with similar swarming behavior, etc. Swarming enables continuous recognition by individuals of other individuals of the same kind and might allow smaller individuals to avoid larger cannibalizing conspecifics who do not conform to the scheme (Section 7.2).

4.3. Modeling basic traits

4.3.1. Overview

The basic model is designed to achieve simplicity in modeling and efficiency in computation. Special cases in definitions have been removed to enable implementation optimization, for example by using lookup tables. Whenever possible, parameter values are mapped to a dimensionless interval of $[0, 1]$. These values can be scaled and extended by a unit to represent values in the field, e.g., a lifespan of $[0, 1]$ is multiplied by 365d (days) to obtain a one-year lifespan. In a $[0, 1]$ lifespan interval, the value 0.0 marks the birth (instantiation) and the value 1.0 the death (termination). The lifetime of an individual is always inside its lifespan.

Unless otherwise stated, the dimensions are given in SI units, i.e., length in meters and mass in kilograms. The time is usually given as years (yrs) or days (d). The mapping to a dimensionless, universal $[0, 1]$ interval allows the *precomputation* of value tables for many relations. Precomputed lookup tables reduce the need for computationally expensive functions, such as the sigmoid size function, thereby reducing the computational cost of each simulation step. Derived values, such as the size-derived metabolism rate (Equation (4.37)), can also be precomputed.

The trade-off involved in the use of lookup tables is (depending on the number of tables, sizes and data types) the requirement for an increased memory allocation. There

4. Modeling the individual and its environment

is also a trade-off to be made between the error resulting from the granularity of the data in the table and the access algorithm if the granularity of the table is reduced and the possible impacts on both computation effort and the simulation results must be considered.

4.3.2. Age, growth, size and resource access

The age model. The central property age increases with each simulation step up to a threshold (death) at which point the agent is removed from simulation:

$$\text{age}_{i+1} = \text{age}_i + \Delta t \quad (4.5)$$

The life history defines the ontological development of an agent (growth, body mass etc.) and the actions it can perform (maturing, reproduction etc.). A model based on predetermined development can retrieve the expressed traits of an individual purely using the function age, without the need for a sophisticated model incorporating additional physiological and exogenous factors. The use of physiological factors to model growth rate was briefly discussed in Subsection 4.2.2.

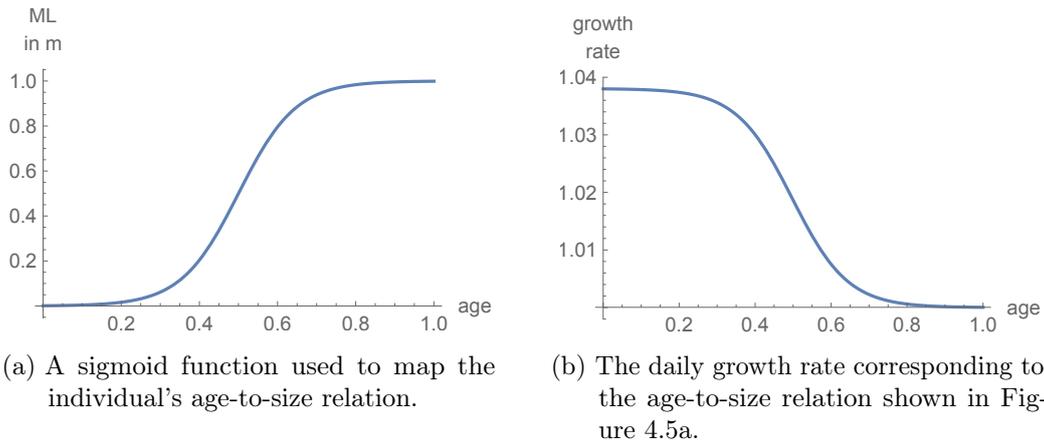


Figure 4.5.: A sigmoid mapping function for size development and daily growth rate.

The significance of the growth function in the model and use of values from the literature. Different growth functions were used to describe subadult and adult growth, i.e., “von Bertalanffy, exponential, linear and logistic” (Rosa et al., 2013), as explained in more detail below. In biology, the widely-used logistic growth function describes an exponential growth that is limited by an upper bound (G). For example,

the upper bound for an individual is the maximum size and for a population is the maximum population size supported by the environment. At both levels, a supporting element, i.e., a physiological process at individual level and an environmental factor at population level, sets a limit that is logarithmically approached to towards the end of growth. In this thesis, the logistic function is used to describe the universal features of basic growth processes, in accordance with Tafur et al. (2010)'s description of a logistic growth function for *D. gigas*. The equation:

$$f(t) = G \cdot \frac{1}{1 + e^{-k \cdot G \cdot t} \cdot \left(\frac{G}{f(0)} - 1\right)} \quad (4.6)$$

uses t as the time (age), G as upper bound (maximum ML), k as growth rate, i.e., steepness of the curve, and $f(0)$ as initial value at $t=0$, e.g., the hatching size. The relation of the exponential function and the hyperbolic tangent (\tanh)

$$\tanh x = \frac{\sinh x}{\cosh x} = \frac{e^{2x} - 1}{e^{2x} + 1} = 1 - \frac{2}{e^{2x} + 1} \quad (4.7)$$

allows the reformulation of Equation (4.6) to

$$f(t) = \frac{G}{2} \cdot \tanh \left(\frac{k \cdot G}{2} \cdot (t - t_{ip}) + 1 \right) \quad (4.8)$$

with the highest gradient inflection point t_{ip} at

$$t_{ip} = \frac{\ln \left(\frac{G}{f(0)} - 1 \right)}{k \cdot G} \quad (4.9)$$

and k for a given inflection point t_{ip} determined by

$$k = \frac{\ln \left(\frac{G}{f(0)} - 1 \right)}{G \cdot t_{ip}} \quad (4.10)$$

The inflection point t_{ip} indicates the time when the growth rate is highest. The $f(\text{age}) \rightsquigarrow \text{size}$ function maps the age to a current size, assuming that each individual follows the same growth pattern. The exponential growth at the inflection point in a logistic growth function decreases towards the end of life (Figure 4.5b) resulting in a typical bell-shaped sigmoid curve. Substituting ML_{terminal} for G and age for the input

4. Modeling the individual and its environment

parameter t gives a size-to-age function:

$${}^*f_{\text{size}}(\text{age}) = \frac{\text{ML}_{\text{terminal}}}{2} \cdot \tanh\left(\frac{k \cdot \text{ML}_{\text{terminal}}}{2} \cdot (\text{age} - t_{\text{ip}})\right) + 1 \quad (4.11)$$

The daily growth rate is retrieved by dividing the size of the next day by the size of the current day:

$${}^*f_{\text{growth rate}}(\text{age}) = \frac{{}^*f_{\text{size}}\left(\text{age} + \frac{1.0}{\text{lifespan in days}}\right)}{{}^*f_{\text{size}}(\text{age})} \quad (4.12)$$

The logistic growth function (Figure 4.5a), supported by analysis of available data for *D. gigas* (Tafur et al., 2010), utilizes capacity (terminal size) as a limiting factor and produces a slowing down in the growth rate towards the end of lifespan (Figure 4.5b). A lifespan of 365d (1.0), $G = 1.2$ m and a hatching size $f(0) = 0.0011$ m gives the value of k in Equation (4.10) as $k \approx 13.6227$. These parameters yield an initial growth rate of $\approx 3.8\%$, which is below the proposed 4–8% paralarval growth rate (Gilly, Elliger, et al., 2006; C. Nigmatullin et al., 2001; Yatsu et al., 1999) but consistent with a daily post-hatching growth rate around 4% ML in Rosa et al. (2013). This is important because subtle changes in the growth function or in the initial daily growth rate (see Section 6.2) result in significant different $\text{ML}_{\text{terminal}}$.

A logistic growth function also has the potential to represent the principal properties of other growth functions observed in the field; for example, it can be fitted to the asymptotic growth functions identified by Goicochea-Vigo et al. (2019) and Zepeda-Benitez et al. (2014) from sampled specimens. This justifies the choice of the logistic growth function to model growth in this thesis.

The above function describes the size development from beginning to end of the lifespan $[0, 1]$, but in field conditions age may not be the only factor determining the size; this is further discussed and used as the starting point for development of an energy-based growth function in Section 4.4.

Resource access. A resource in the model is a spatio-temporally organized model property defined by one or more individuals, a CA (Subsection 9.3.3) or a singleton (Subsection 9.4.2). In the model, a numeric resource interval determines whether an individual has access to a specific resource based on a size relation. If a resource fits into the size window, for example the WOP for prey items, then this resource can be accessed. Some resources are defined as unrestricted and their accessibility is determined in the model without reference to a resource interval. The size range of consumable prey items for both the WOP (Subsection 2.1.3) and the window of cannibalism (WOC)

(Claessen et al., 2000; Ibáñez & Keyl, 2010)) is given by the minimum and maximum size of prey in relation to the size of the predator, expressed as a percentage, i.e., in the case of the WOC as $ML_{\text{conspecific}}/ML_{\text{cannibal}} \cdot 100$.

4.3.3. Size derived traits: Mass and surface

Mass estimation. In order to simplify estimation of the body mass, the body of the individual is assumed to consist of a single, homogenous “material”. In this section, the mass is calculated from the size (ML), which in turn is derived from the age (Equation (4.11)). Mass therefore is a function of age in an age-based model, or a function of ML in a ML-based model (Subsection 4.2.2).

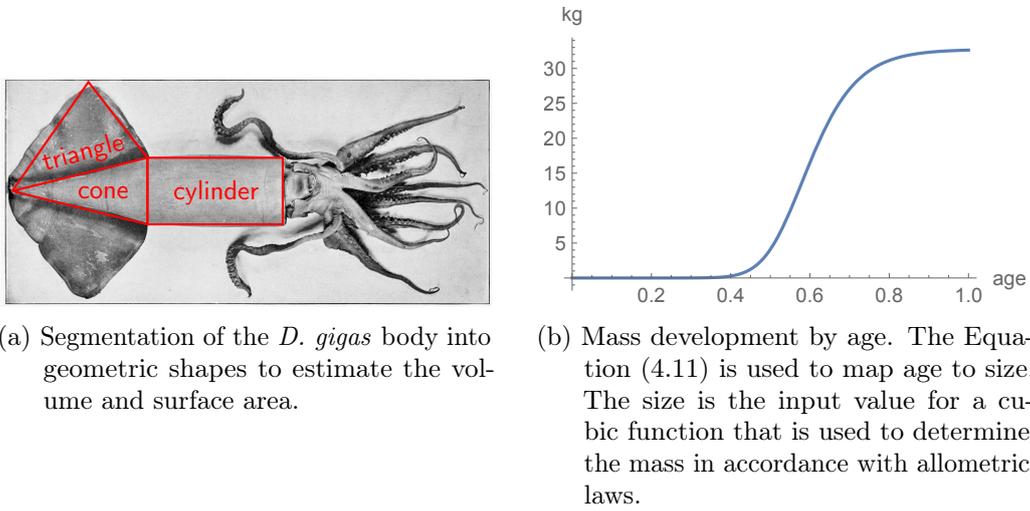


Figure 4.6.: Segmentation of *D. gigas* to enable computation of its body mass development using a logistic growth function.

A cylinder and a cone approximate the shape of a *D. gigas* individual for the purposes of volume and mass computation (Figure 4.6a). At the midpoint of its mantle length, the mantle is divided into a right circular cylinder and a right circular cone. The volume of the right circular cylinder V_{cyl} and surface area of its side (without the end face areas) L_{cyl} are given by:

$$V_{\text{cyl}} = \pi \cdot r^2 \cdot h \quad (4.13)$$

$$L_{\text{cyl}} = 2 \cdot \pi \cdot r \cdot h \quad (4.14)$$

4. Modeling the individual and its environment

The volume of the right circular cone V_{cone} and the surface area of its side L_{cone} (without the base area) are given by:

$$V_{\text{cone}} = \frac{1}{3} \cdot \pi \cdot r^2 \cdot h \quad (4.15)$$

$$L_{\text{cone}} = \pi \cdot r \cdot \sqrt{h^2 + r^2} \quad (4.16)$$

The parameter h equals half of the ML and is substituted for ML in the above equations:

$$h = \frac{\text{ML}}{2} \quad (4.17)$$

The parameter r is half of the mantle diameter, which is taken to be equal to one-quarter of the ML:

$$\text{mantle diameter} = \frac{\text{ML}}{4} \quad (4.18)$$

$$r = \frac{\text{mantle diameter}}{2} \quad (4.19)$$

$$r = \frac{\text{ML}}{8} \quad (4.20)$$

The fins contribute a significant amount of surface area. The area is approximated by an triangle where the base is the slant height of the cone ($\sqrt{r^2 + h^2}$) and the height is $\approx \frac{5}{7}$ of the cone's slant height.

$$L_{\text{single side fin}} = 0.5 \cdot \frac{5}{7} \cdot \sqrt{h^2 + r^2} \cdot \sqrt{h^2 + r^2} = \frac{5}{14} \cdot (h^2 + r^2) \quad (4.21)$$

Thus the total surface area of the fins is

$$L_{\text{fins}} = 4 \cdot \frac{5}{14} \cdot (h^2 + r^2) = \frac{10}{7} \cdot (h^2 + r^2) \quad (4.22)$$

since there are two fins, each with two sides. The volume of an individual is defined based on Equation 4.13 and Equation 4.15 as:

$$V_{\text{individual}} = \pi \cdot r^2 \cdot h + \frac{1}{3} \cdot \pi \cdot r^2 \cdot h = \pi \cdot r^2 \cdot h \cdot \left(1 + \frac{1}{3}\right) = \frac{4}{3} \cdot \pi \cdot r^2 \cdot h \quad (4.23)$$

A volume function of ML in meters $f_{V_{\text{individual}}}(\text{ML})$ is retrieved through substituting Equation 4.17 for h and Equation 4.20 for r :

$$f_{V_{\text{individual}}}(\text{ML}) = \frac{4}{3} \cdot \pi \cdot \left(\frac{\text{ML}}{8}\right)^2 \cdot \frac{\text{ML}}{2} = \frac{\text{ML}^3 \cdot \pi}{96} \approx 0.0327249 \cdot \text{ML}^3 \quad (4.24)$$

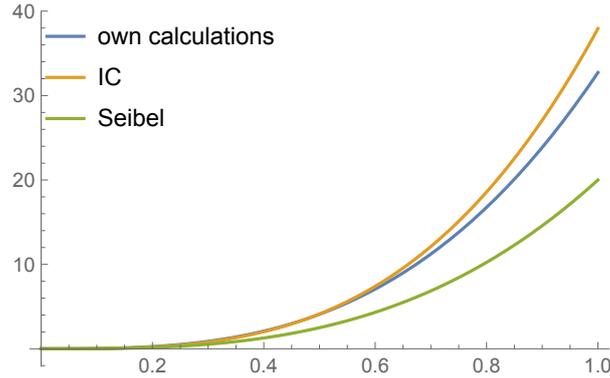


Figure 4.7.: Comparison of body mass estimating functions: “own calculations” based on Equation (4.25), “Seibel” (B. A. Seibel et al., 2000), “IC” (Ibáñez & Cubillos, 2007).

The cubic relation of size to volume amplifies the original sigmoid size function to a volume function with a higher slope in the mid-section of the function (Figure 4.6b). The body mass (BM) in kilograms is approximated by assuming that an individual consists only of proteins (muscle tissue) and not of 70–90% protein as estimated by (Keyl, 2009; Rosa & Seibel, 2010). A muscle specific density of $1.05\text{g}/\text{cm}^3$ (1.05kg L^{-1}) is close to 1 (1kg L^{-1}), so the volume in liters is presumed equal to mass in kilograms. Entering the ML in the SI unit meters in Equation (4.24) gives the weight in metric tonnes per cubic meter and is multiplied by factor 1000 to give the weight per cubic meter in the SI unit kilograms:

$$\text{BM in kg} \approx 32.7249 \cdot (\text{ML in m})^3 \quad (4.25)$$

Equation (4.25) differs from the mass estimate in B. A. Seibel et al. (2000) ($\text{BM} = 20 \cdot (\text{ML in m})^3$), but is closer to actual mass-length ratio reported in (Field et al., 2007), i.e., 30–50kg for a dorsal ML of 1.0–1.2m. A $\text{BM} = 0.0000151 \cdot (100 \cdot \text{ML})^{3.2}$ is described in Ibáñez and Cubillos (2007), who found that the average weight of females was greater than that of the males with the same ML when $\text{ML} > 0.5\text{ m}$. Figure 4.7 compares the mass function used in this study with those of B. A. Seibel et al. (2000) and Ibáñez and Cubillos (2007). Although the curves are relatively close, but the $\text{BM} \approx 68\text{ kg}$ is probably too high for a $\text{ML} = 1.2\text{ m}$ in the function of Ibáñez and Cubillos (2007), compared to the reported body mass of approx. 50kg for a 1.2m individual (Field et al., 2007).

4. Modeling the individual and its environment

The inverse function of that shown in Equation (4.25) determines for a given body mass in kg the ML in meter and is used in Section 4.4 to determine a growth rate for ML for a given body mass gain:

$$\text{ML} = \left(\frac{96 \cdot \text{BM}}{1000 \cdot \pi} \right)^{1/3} \approx (0.0305577 \cdot \text{BM})^{1/3} \quad (4.26)$$

Surface estimation. The surface of *D. gigas* is approximated by the sum of L_{cyl} , L_{cone} and L_{fins} to:

$$L_{\text{individual}} = 2 \cdot \pi \cdot r \cdot h + \pi \cdot r \cdot \sqrt{h^2 + r^2} + \frac{10}{7} \cdot (h^2 + r^2) \quad (4.27)$$

which when Equation (4.17) is substituted for h and Equation 4.20 for r is equal to:

$$L_{\text{individual}} = \pi \cdot \frac{\text{ML}^2}{8} + \pi \cdot \frac{\text{ML}}{8} \cdot \sqrt{\frac{\text{ML}^2}{4} + \frac{\text{ML}^2}{64}} + \frac{10}{7} \cdot \left(\frac{\text{ML}^2}{4} + \frac{\text{ML}^2}{64} \right) \quad (4.28)$$

and approximates to a surface function:

$$\begin{aligned} f_{L_{\text{individual}}}(\text{ML}) &= 0.379464 \cdot \text{ML}^2 + 0.202393 \cdot \text{ML}^2 + 0.392699 \cdot \text{ML}^2 \\ &= 0.974556 \cdot \text{ML}^2 \end{aligned} \quad (4.29)$$

The volume (mass) and surface approximations are used for metabolism-related computations (Section 4.4) and inform the discussion of oxygen consumption as a possible limiting factor (Subsection 6.3.3).

4.4. The energy model of the agent

Overview. The energy model is based on a fish energy balance model (Winberg, 1960) and the equivalence of mass and energy via the conversion factor $f_{\text{BM} \rightarrow \text{e}}$. This equivalency enables an energy focused approach, whereby the ML determines a related expected body mass (Equation (4.25)) and maps a related (virtual) age onto the growth function. The virtual age defines the potential (i.e., maximum available) growth rate determined by the growth function, while the actual growth rate (up to this maximum) is determined by the available energy. Conversion of the expected body mass gain to the equivalent energy value gives the energy requirement in order to attain the available growth rate as specified by the growth function. Stripped down to essentials, the energy

model consists of a growth function linked to the parameter virtual age. The model's simplicity allows the systematic exploration of its behavior in its parameter space.

The underlying fish energy balance model. The fish energy balance model of Winberg (1960) follows the law of energy conservation where F_T is total energy uptake through food, M is the energy expended on somatic functions and reproduction, G is energy expended on growth, and W is wasted energy via excretion (André et al., 2009):

$$F_T = M + G + W \quad (4.30)$$

The assimilation efficiency factor A defines the rate of energy uptake that is not lost through excretion. It substitutes for the factor W and determines the required energy uptake by food for energetic purposes F :

$$F = F_T \cdot A = M + G \quad (4.31)$$

The resulting energy balance E_R hence is given by:

$$E_R = F - M - G \quad (4.32)$$

The widely accepted Kleiber's law (Glazier, 2006; Kleiber, 1932) describes the metabolic rate m_r as a function of the body mass BM with $m_r = BM^{3/4} = BM^{0.75}$. This is a special case of the allometric scaling law of the form

$$m_r = q \cdot BM^p \quad (4.33)$$

where $q = 1$ and $p = 0.75$. All factors in Equation (4.32) are body mass dependent; therefore these factors are a function of the BM :

$$E_R(BM) = F(BM) - M(BM) - G(BM) \quad (4.34)$$

Substituting $F(BM)$ and $M(BM)$ by the allometric scaling law (Equation (4.33)), and $G(BM)$ by the linear relation that exists in the exponential growth phase between growth and BM (André et al., 2009), gives:

$$E_R(BM) = q_1 \cdot BM^{p_1} - q_2 \cdot BM^{p_2} - q_3 \cdot BM \quad (4.35)$$

4. Modeling the individual and its environment

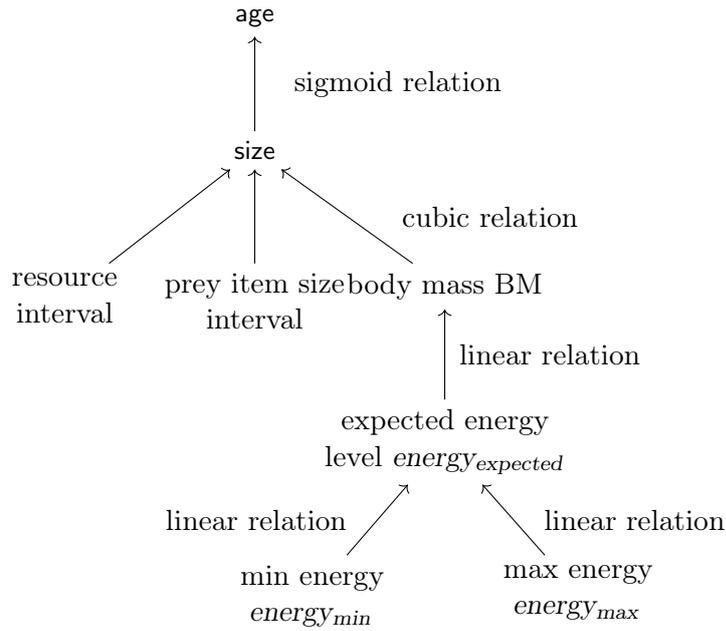


Figure 4.8.: The top-down hierarchy of the individual energy model and related parameters. The arrows indicate the direction of dependency.

The species-specific constants q_1 , q_2 and q_3 are obtained by experimental studies (André et al., 2009). $G(\text{BM})$ may change for $E_R(\text{BM}) < 0$ (Equation (4.35)) from exponential growth to linear or logarithmic growth to reduce an energy deficit. Any food limitation in an ecosystem can limit the increase of $F(\text{BM})$, so $F(\text{BM})$ functions as a limiting factor for $G(\text{BM})$.

Energy is the driving factor in the model through the equivalence of mass and energy. The equivalence of body mass and energy level identifies energy as the “universal currency” and allows energy to be set as the driving factor in the model. The energy flow starts with the reference value $energy_{expected}$, which is the energy level (calculated from its body mass) of a fully nourished individual of a given size (Equation (4.25)). The equation

$$energy_{expected} = f_{\text{BM} \rightarrow e} \cdot *f_{\text{BM}}(\text{size}) \quad (4.36)$$

converts the BM to energy by means of the body mass to energy conversion function $f_{\text{BM} \rightarrow e}$. Figure 4.8 illustrates the dependencies and the paths leading to the root node

“age”. Following the arrows with their applied functions up to the root node resolves the dependencies and maps $energy_{expected}$ to a function of (virtual) age.

The maximum energy level $energy_{max} = f_{energy_{max}} \cdot energy_{expected}$ is the threshold up to which an agent can store energy. Energy is stored in the body mass, so that larger individuals have a higher energy storing capacity. The minimum energy level $energy_{min} = f_{energy_{min}} \cdot energy_{expected}$ defines a threshold value below which an agent suffers from terminal energy exhaustion.

The energy balance E_R (Equation (4.32)) for an individual at each iteration of the model run is characterized by (1) the *lost energy* $total_{mr}$ (Subsection 4.2.2) and (2) the *energy uptake* by resource consumption, e.g., for predation. The $total_{mr}$ includes the energy expended on growth ($growth_{mr}$), but growth is limited if the energy uptake is insufficient. In this case, the increase in body mass is reduced to ensure availability of the energy required for basal metabolic rate ($basal_{mr}$) and locomotion metabolic rate ($locomotion_{mr}$).

Estimating the metabolism for the energy balance. An energy deficit i.e., $E_R(BM) < 0$ (Equation (4.35)), must be made up for by a body mass to energy conversion $f_{BM \rightarrow e}$ that results in a body mass loss BM_{loss} . The basal metabolic rate ($basal_{mr}$) and $locomotion_{mr}$ (Figure 4.3, basal metabolism and energy need for locomotion) define fundamental energy needs, since both factors are considered essential to maintain the role of *D. gigas* as an active and opportunistic predator. The BM_{loss} due to $basal_{mr}$ depends on the BM (Equation (4.25)):

$$BM_{loss} = \frac{BM^p}{f_{BM \rightarrow e}} = \frac{(32.7249 \cdot ML)^p}{f_{BM \rightarrow e}} = \frac{(32.7249 \cdot *f_{size}(age))^p}{f_{BM \rightarrow e}} \quad (4.37)$$

The $basal_{mr}$ energy loss is determined by the metabolism rate m_r and is estimated according to Kleiber’s law but with a different value of p . Dividing the metabolism rate by the BM defines the metabolism intensity rate m_i :

$$m_i = b_0 \cdot BM^b \quad (4.38)$$

The factor b_0 is a species-specific but mass-independent normalization constant, and $b = -0.25$ a scaling coefficient, that is related to the factor p in Kleiber’s law as $b = p - 1$; using the standard value $p = 0.75$ gives $b = -0.25$ (B. A. Seibel, 2007). However, the use of the standardized value $p = 0.75$ to described metabolic rate has been criticized since it is claimed that this is not universal among species (Glazier et al., 2015; Glazier, 2006). Specifically, oxygen consumption measurements for *D. gigas* give the values $b = -0.06$

4. Modeling the individual and its environment

(B. A. Seibel, 2007; B. A. Seibel et al., 2007) and $b_0 = 12.75$ (Rosa & Seibel, 2006). This gives $p = 0.94$, which is consistent with observed values of p for other cephalopod species such as *Octopus vulgaris* with approx. $p = 0.9$ (Katsanevakis et al., 2005).

B. A. Seibel (2007) uses the rate of oxygen consumption per unit of mass ($\mu\text{mol O}_2 \text{g}^{-1} \text{h}^{-1}$) as an indicator of the metabolism intensity rate m_i . This measure enables the mathematical conversion of energy to body mass and vice versa. The mass-specific oxygen consumption rate BM_{O_2} is equal to the metabolism intensity rate m_i :

$$\text{BM}_{\text{O}_2} = b_0 \cdot \text{BM}^b \quad (4.39)$$

Inserting the values $b_0 = 12.75$ and $b = -0.06$ gives:

$$\text{BM}_{\text{O}_2} = 12.75 \cdot \text{BM}^{-0.06} \quad (4.40)$$

Reformulating Equation (4.40) with the BM of an individual given in grams, taking the measurement period as 24h and applying the μmol to mol factor 10^{-6} gives:

$$\text{BM}_{\text{O}_2} = \text{BM} \cdot 24 \cdot 10^{-6} \cdot 12.75 \cdot \text{BM}^{-0.06} = 3.06 \cdot 10^{-4} \cdot \text{BM}^{0.94} \quad (4.41)$$

This gives the oxygen need of an individual to meet the requirements of basal_{mr} in mol d^{-1} . The volume of an ideal gas is 22.41L mol^{-1} at 101.325kPa and 273.15K , thus the volume in liters (L_{O_2}) of oxygen expended on basal_{mr} per day for a given body mass (in grams) is equivalent to

$$\text{L}_{\text{O}_2} = 22.41 \cdot 3.06 \cdot 10^{-4} \cdot \text{BM}^{0.94} = 6.85746 \cdot 10^{-3} \cdot \text{BM}^{0.94} \quad (4.42)$$

Converting mass in grams in Equation (4.42) to the SI unit kg gives:

$$\text{L}_{\text{O}_2} = 6.85746 \cdot 10^{-3} \cdot (\text{BM} \cdot 1000)^{0.94} = 4.53068 \cdot \text{BM}^{0.94} \quad (4.43)$$

The amount of energy expended on basal_{mr} is estimated assuming a protein-based metabolism and based on the amount of oxygen required to metabolize one gram of protein. Nelson et al. (2008) estimates this to be 0.94 liter oxygen per gram of protein. Thus the amount of protein in kg expended on basal_{mr} for an individual of a body mass BM in kg can be calculated by dividing the volume of expended oxygen (L_{O_2}) given by Equation (4.43) by 0.94 (to convert liters to grams protein) and by 1000 (gram to kg):

$$\text{kg protein spent on } \text{basal}_{\text{mr}} = \frac{4.53068 \cdot \text{BM}^{0.94}}{0.94 \cdot 1000} = 0.00481987 \cdot \text{BM}^{0.94} \quad (4.44)$$

4.4. The energy model of the agent

One kg protein provides 17.2MJ energy and this enables calculation of the energetic requirements per day to uphold the basal_{mr} for a *D. gigas* individual of known body mass in kg:

$$\begin{aligned}\text{basal}_{\text{mr}} \text{ d}^{-1} \text{ kJ} &= 17.2 \text{ MJ kg}^{-1} \cdot 0.00481987 \cdot \text{BM}^{0.94} \\ &= 82.9018 \cdot \text{BM}^{0.94}\end{aligned}\quad (4.45)$$

Converting between mass and energy. The equivalence of energy and protein mass allows the estimation of amount of energy invested into growth. Since the body mass of *D. gigas* has an 80% water content, the body mass of protein can be calculated as:

$$\text{BM}_{\text{protein}} = \text{BM} \cdot 0.2 \quad (4.46)$$

In other words, each unit of body mass contains only 0.2 mass units of protein, so the 17.2 MJ kg^{-1} energy requirement for protein upkeep is equivalent to 3.44 MJ per kg *D. gigas* body mass:

$$\text{kJ} = \text{BM}_{\text{protein}} \cdot 17\,200 \text{ kJ kg}^{-1} = \text{BM kg} \cdot 3440 \text{ kJ kg}^{-1} \quad (4.47)$$

Equation (4.47) defines the energy/mass equivalency for *D. gigas*. The formula is used to estimate the energy investment ΔE required for a given increase in body mass ΔBM (energy to mass conversion) and the mass loss ΔBM incurred when mass is converted to energy ΔE (i.e., mass to energy conversion) in times of food deprivation. To take account of the energy losses involved in these conversions, the equivalence between mass and energy is adjusted using an efficiency factor of 0.9, derived from fish metabolism (Brett & Groves, 1979), since the actual efficiency factor for *D. gigas* is not known:

$$\Delta E \text{ kJ} = \Delta \text{BM kg} \cdot 3440 \text{ kJ kg}^{-1} \cdot 0.9 \quad (4.48)$$

$$\Delta \text{BM kg} = \frac{\Delta E \text{ kJ}}{3440 \text{ kJ}} \cdot 0.9 \quad (4.49)$$

Summary. The above equations allow calculation of the basal_{mr} energy needs, the available energy based on body mass, the energy balance, and the conversion of body mass to energy and vice versa, to compute values for BM_{loss} and, in case of $\text{BM}_{\text{loss}} < 0$, body mass gain. The top (root) parameter **age** (Figure 4.8) defines the relevant body parameters at lower levels of the tree. The **age** is assigned to an individual and incremented in each simulation step.

4. *Modeling the individual and its environment*

The model in Figure 4.8 thus allows the basic parameter `age` to be used interchangeably with other related factors in implementing different sub-models, e.g., setting the `BM` as the main factor and retrieving the `size` using a function like that shown in Equation (4.26).

5. Identifying spawning areas of *D. gigas*

5.1. Overview

The exact locations of spawning areas of *D. gigas* have not yet been identified. Records of occurrences of eggs and paralarvae (Table 5.1 on page 75) point to the existence of several distinct or possibly a single large spawning area. Section 5.2 discusses the probable existence of multiple spawning areas over the distribution range of *D. gigas*. This discussion suggests that *D. gigas* spawns over a broad area, but the few findings of paralarvae and egg masses appear to contradict this conclusion. Section 5.3 discusses a prior work by Staaf et al. (2013), which describes probable spawning areas of the *Sthenoteuthis oualaniensis* and *Dosidicus gigas* complex (SD-complex) and identifies sea surface temperature (SST) as the main predictor variable of SD-complex paralarvae presence probability (pp_{prob}). Section 5.4 uses satellite-based SST data to predict possible spawning areas within the distribution range of *D. gigas* and reveals a broader potential spawning area than that described in Staaf et al. (2013).

In Section 5.5, an analysis of the surface currents of the Pacific within the distribution area of *D. gigas* explains the dispersed distribution of recorded occurrences of paralarvae through an egg or paralarvae transport mechanism to local “attractors” in this region. Based on the assumption of a such a material transport mechanism, Section 5.6 implements a reverse engineering approach to compute the origin of spawning areas by backtracking along the paths of currents from the locations of paralarvae occurrences. The results and conclusions of both the literature and new computations and models are summarized and discussed in Section 5.7.

This chapter uses a wide range of methodologies to process satellite data, compute the distributional flow of paralarvae and to backtrack paralarvae occurrences to their potential spawning areas. The computation of material transport is implemented using a cellular automaton (CA)-based simulation model.

5.2. Discussion of probable multiple spawning grounds

The large distribution range of *D. gigas* is compatible with the existence of multiple spawning grounds or a single large spawning area. Any range expansion to one or

5. Identifying spawning areas of *D. gigas*

both poles is “[...] limited by sustainability of foraging potential within an energetically acceptable distance of suitable spawning grounds [...]” (Rosa et al., 2013). The locations of suitable spawning grounds therefore determine the distribution range; conversely, the outer edges of the spawning grounds can be determined based on estimates of the “energetically acceptable distance” between spawning grounds and the known limits of the distribution range.

A hypothetical single, small and distinct spawning ground would require mature adults to migrate to this single spawning ground and, after spawning, the juveniles to radiate away from that area over the entire distribution range, which would constitute a very energy-intensive migration pattern. The cross-sectional distance between southern and northern limits of distribution range of approx. 112° degrees latitude (59°N to 53°S) can be computed, taking the Earth’s radius of approx. 6371km, as an arc d :

$$d = \frac{6371 \text{ km } 2\pi 112^\circ}{360^\circ} \approx 12\,454 \text{ km.} \quad (5.1)$$

A central spawning area within the distribution range would require some individuals to undertake an about 6000km migration to the northern or southern boundaries of the distribution range, and then return back to the spawning ground, a total migration of about 12 000km. This would imply a daily migration of approx. 30km per day during a one year lifespan, which is possibly beyond physical capabilities of *D. gigas*, quite apart from energetic considerations. Using electronic tagging, Gilly, Markaida, et al. (2006) estimate an average migration of approx. 30km per day for an adult specimen in the Gulf of California (GOC) with a mantle length (ML) greater than or equal to 77cm. But considering the small size of *D. gigas* during the first half of its lifespan, the outward journey to the edge of the distribution range seems unachievable and this makes a single central spawning area unlikely.

Given the inefficient propulsion locomotion of *D. gigas*, the energetic demands on an individual of such a long-distance migration would be enormous. Following prey items does not guarantee constant feeding, so individuals engaged in this kind of migration would be at risk from terminal energy exhaustion before they reached the spawning ground, which would prevent them from taking part in the reproduction process.

Furthermore, a single main (small) spawning area would attract predators and make it easy for them to feed on *D. gigas*, possibly jeopardizing reproductive success. Egg masses and paralarvae, in particular, are easy prey items. Even juveniles, if these were still in this area, would be susceptible to predators. Hence having a single spawning area would expose *D. gigas* to an increased risk of massive mortality due to predation.

5.2. Discussion of probable multiple spawning grounds

Considering the presumably high mortality of *D. gigas* (Boyle & v. Boletzky, 1996; Camarillo-Coop et al., 2013; Lipinski, 2002) during its early developmental stages, the ecosystem of a single, relatively small spawning area would have to support an enormous number of progeny; this might however be possible due to the low energy demand of each (small) individual.

reference	sample	date	location
Tafur et al. (2001)	mature females	1991–1995	3°S to 8°S, Peruvian coast
Tafur et al. (2001)	mature females	1991–1995	12°S to 17°S, Peruvian coast
Gilly, Elliger, et al. (2006)	paralarvae	2004-05-04	28.402°N / 112.454°W, GOC
Gilly, Elliger, et al. (2006)	juvenile	2004-05-06	28.192°N / 112.170°W, GOC
Gilly, Elliger, et al. (2006)	paralarvae	2004-05-07	28.169°N / 112.027°W, GOC
Staaf et al. (2008)	eggs	2006-06-19	27.12°N / 111.27°W ¹ , GOC
Birk et al. (2016)	eggs	2015-05-28 to 2015-05-30	27.24°N / 111.52°W ² , GOC
Argüelles and Tafur (2010)	mature females	1994–2006	3°S to 8°S, Peruvian coast
Ramos-Castillejos et al. (2010)	paralarvae	2005	24°N to 28°N, West Coast Baja California
Staaf et al. (2013)	paralarvae	2006-09-03 ³	8.85°N / 85.63°W, Pacific
Staaf et al. (2013)	paralarvae	2006-09-15 ³	1.43°S / 113.72°W, Pacific
Sakai et al. (2008)	paralarvae	2007-11-15	5.6°S / 86.6°W, Humboldt Current
Sakai et al. (2008)	paralarvae	2007	6°S to 8°S, Humboldt Current
Sakai et al. (2008)	paralarvae	2007	12°S to 17°S, Humboldt Cur- rent

¹ The original degrees, minutes and seconds location format 27°7.1'N/111°16'W was converted to decimal degrees.

² Location approximated, six findings in a 5km radius.

³ Date estimated.

Table 5.1.: Recorded occurrences of *D. gigas* eggs, paralarvae, juveniles and mature females.

The above discussion and the widely separated occurrences of eggs, paralarvae and mature females suggest the existence of multiple spawning grounds over a large area. If there is a single spawning area, the only natural record of an egg mass detection in the GOC (Birk et al., 2016; Staaf et al., 2008) would mark its location. In that case one would expect more egg masses to be found in this region, but there are no further reports of such occurrences. Moreover paralarvae have been found outside this area, as shown in Table 5.1. The occurrence of paralarvae, the next ontological stage after

5. Identifying spawning areas of *D. gigas*

hatching, indicates there is a spawning ground nearby, because paralarvae have limited swimming capability and there is only short incubation time of a few days between spawning and hatching (Table 2.3).

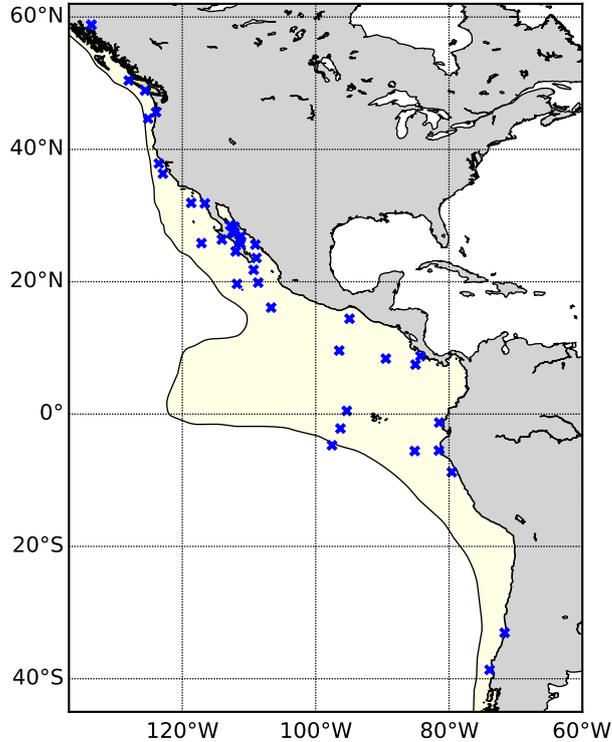


Figure 5.1.: Field sampling locations (blue cross) where tissue samples of *D. gigas* were collected by Staaf et al. (2010a, 2010b). Tissue samples of 802 *D. gigas* specimens were taken at 53 locations between 1999 and 2008.

The occurrence of eggs and paralarvae in the GOC (Birk et al., 2016; Gilly, Elliger, et al., 2006; Staaf et al., 2008), in the northern hemisphere off the coast of Mexico (Staaf et al., 2013) and in the southern hemisphere in a larger area off Peru (Sakai et al., 2008; Tafur et al., 2001) support the hypothesis of multiple spawning grounds. This is also consistent with the existence of two populations, weakly differentiated genetically, in the northern and southern hemispheres, separated by a biogeographical break at approx. 6°N (Staaf et al., 2010b). The existence of these two populations was inferred from tissue samples of 802 *D. gigas* specimens taken at 53 different locations (Figure 5.1) throughout the distribution range of the species between 199 and 2008 (Staaf et al., 2010a).

The hypothetical multiple spawning areas may not cover the entire distribution range of *D. gigas*. Stewart et al. (2012) suggest that some individuals of *D. gigas* may undertake a long-distance migration towards the outer edge of distribution range and return to the spawning area, involving, by a conservative estimate, a journey (from the Pacific Coast off Baja California to Vancouver Island in British Columbia) of “at least 2750km each way”. However, isotope analysis of beaks from middle-sized and large specimens (sampled between 2°S and 38°N) detected no evidence of long-distance latitudinal migration, and concluded that individuals migrate no more than 4° latitude during their lifetime (R. I. Ruiz-Cooley et al., 2010). Isotope analysis in the Northern

5.3. Previous identification of likely spawning areas

California Current System (R. I. Ruiz-Cooley et al., 2013) indicates that the sources of origin (spawning grounds) of specimens sampled in this region are in northern Baja California, Mexico (28°N–35°N), the waters of the southern Baja California Peninsula (22°N–23°N) and the much warmer offshore waters to the west near the boundary of the north Pacific Gyre system. *D. gigas* is considered to move in from these different sources to feeding grounds in the Northern California Current System (R. I. Ruiz-Cooley et al., 2013). Isotopic analysis of beaks by Hu et al. (2019) suggests that in El Niño (EN) years *D. gigas* migrates over shorter distances and these authors conclude that “the migration and foraging ecology of jumbo squid are substantially influenced by mesoscale oceanic oscillations.” Based on the results of isotope analysis of muscle tissue and beaks of specimens sampled in areas of the South Pacific adjacent to Ecuadorian, Peruvian and Chilean Exclusive Economic Zones, Liu et al. (2018) state that “squid off Chile, especially Peru, migrate over a large geographic range”. Also, isotope analyses of eye lenses suggest that *D. gigas* “live in wide areas with different isotopic baseline values” (Liu et al., 2020) in their early life stages.

In summary, the spawning grounds of *D. gigas* are probably distributed over large areas in both hemispheres. The existence of a single spawning area near the equator is not supported by the literature, nor is it likely individuals would be capable of making the journey to the outer regions of the species’ range and then returning to the spawning ground.

5.3. Previous identification of likely spawning areas

Overview. The auxiliary indicator “paralarvae occurrence” (Camarillo-Coop et al., 2010) provides inputs for the identification of spawning grounds. Paralarvae have no capacity for independent horizontal movement (Staaf et al., 2008); therefore higher paralarvae concentrations indicate probable spawning grounds. Occurrences of both paralarvae and eggs have been reported for the GOC (Camarillo-Coop et al., 2010; Gilly, 2006a; Staaf et al., 2008), and of paralarvae only for the Humboldt Current off Peru and Chile (Sakai et al., 2008) and for the Eastern Tropical Pacific (Staaf et al., 2013) (Table 5.1).

Methods. (Staaf et al., 2013) carried out an extensive study based on *in situ* oceanographic variables and abundance of paralarvae of the *Sthenoteuthis oualaniensis* and *Dosidicus gigas* complex (SD-complex) over eight years (1999–2008). Paralarvae of the SD-complex were collected using surface (manta) and subsurface (bongo) tows at

5. Identifying spawning areas of *D. gigas*

different locations and different times of the year. Two hours after sunset, with all deck lights off, the manta nets were towed for 15 minutes at 1–2 knots. The bongo nets were towed for a 15 minutes on a double oblique haul to a depth of approx. 200m at a ship speed of 1.5–2.0 knots. A flowmeter was used to estimate the volume of water filtered during the tows (Staaf et al., 2013). Using the Akaike information criterion, the *in situ* oceanographic variables SST, sea surface salinity, mixed-layer depth, temperature at thermocline and surface concentration of chlorophyll- α were stepwise combined to assess predictor variables against measured paralarvae concentrations. SST was identified as the strongest predictor variable, with minimal contribution from the other variables (Staaf et al., 2013).

The methodology applied by Staaf et al. (2013) provides a first insight into potential relations between paralarvae presence and oceanographic variables. However, for the identification of possible spawning grounds, additional aspects need to be considered:

- The data assembled by Staaf et al. (2013) cover a period of several years but do not consider the possible effects of year-on-year changes in values of oceanographic variables.
- Without an explicit underlying explanatory model, using “combinatorial guessing” to search for correlations is susceptible to pseudo-correlation (“Texas sharpshooter fallacy”).
- Only two of the few ethanol-preserved samples contained identified *D. gigas* paralarvae; in formalin-preserved samples, it is not possible to differentiate between the two species based on morphological properties. It is believed that the closely related species *S. oualaniensis* and *D. gigas* express the same behavior during ontological development (D. Staaf, personal communication), but the lack of differentiation between the two species in this study renders it more difficult to interpret aspects of the data that are relevant for the location of *D. gigas* spawning grounds.
- The distribution ranges of *S. oualaniensis* and *D. gigas* overlap, but *D. gigas* is the only adult ommastrephid found in the GOC (Staaf et al., 2013), which contradicts the assumption of shared behavior. Both species occur in the equatorial zone, but only *D. gigas* is present in the Humboldt Current System.
- The 2006 survey reported by Staaf et al. (2013), conducted aboard the two NOAA ships *David Starr Jordan* and *MacArthur II*, collected data from 28th July to 7th December. Starting out on the *David Starr Jordan* from San Diego, California

5.3. Previous identification of likely spawning areas

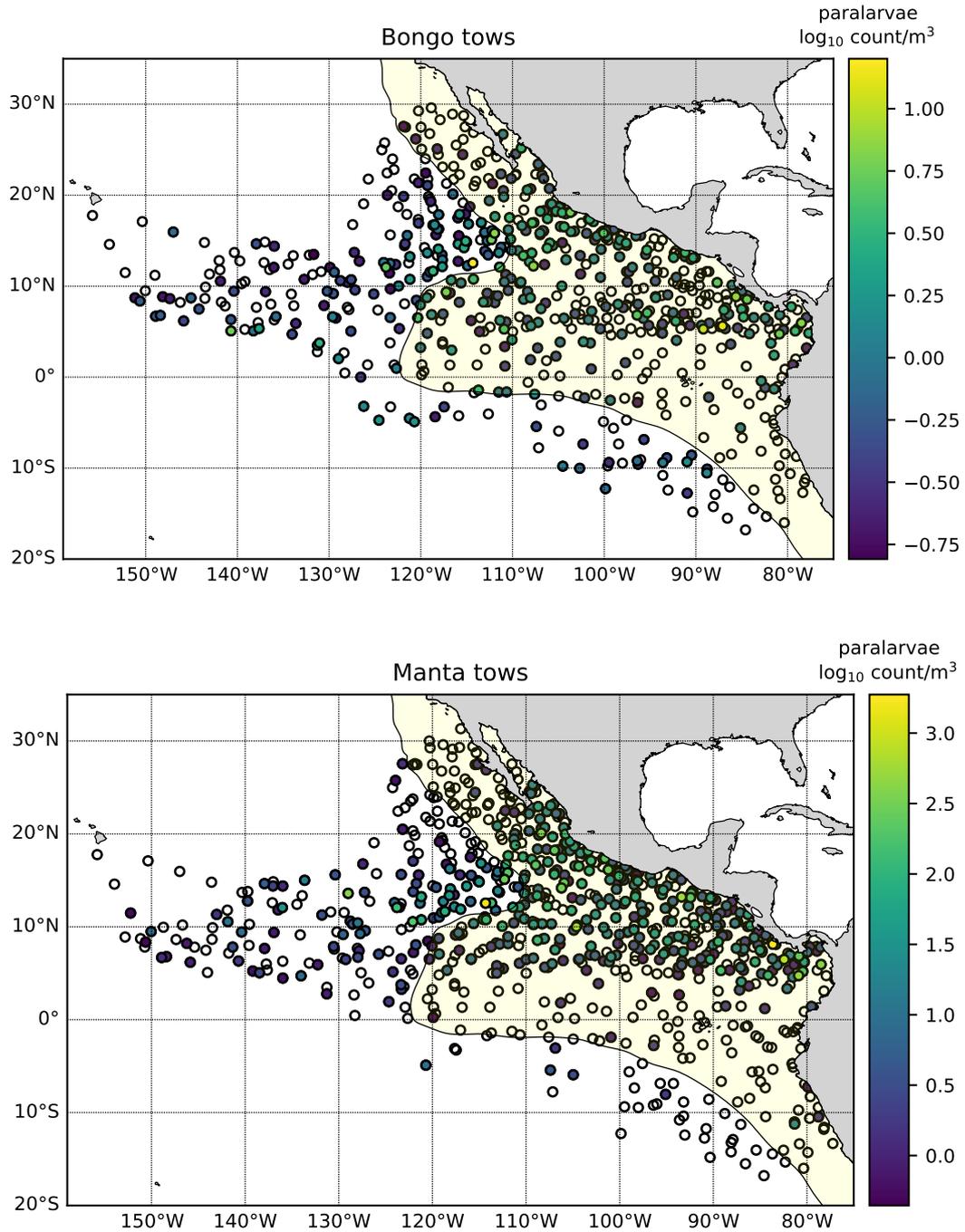


Figure 5.2.: Plots of log₁₀-transformed bongo (subsurface) and manta (surface) tow sampling counts of *Sthenoteuthis ovalaniensis* and *Dosidicus gigas* complex (SD-complex) paralarvae (Staaf et al., 2013), generated by data provided by D. Staaf (personal communication). Unfilled circles mark zero-count sample locations. The yellow area marks the distribution range of *D. gigas*.

5. Identifying spawning areas of *D. gigas*

and heading southwards, researchers found *D. gigas* paralarvae towards the end of the second leg from Manzanillo, Mexico to Puntarenas, Costa Rica (Kinzey et al., 2008), in samples estimated to have been collected between September 1st and 3rd. The second sample in which the presence of *D. gigas* paralarvae was confirmed was taken during the leg from Honolulu (Hawaii), to Manta (Ecuador), which corresponds to an estimated date of September 15th. Both sampling dates are towards the end of the presumed peak spawning season and, therefore, the occurrences of paralarvae may not correspond to the actual spawning areas.

Discussion and results. Despite these concerns, the study by Staaf et al. (2013) represents a feasible, interesting and undoubtedly the best approach so far towards filling data gaps with respect to the spawning grounds of the SD-complex. Staaf et al. (2013) processed the paralarvae counts (Figure 5.2) to locate likely spawning areas of the SD-complex (Figure 5.3) and identified SST as the main predictor variable of pp_{prob} for the SD-complex.

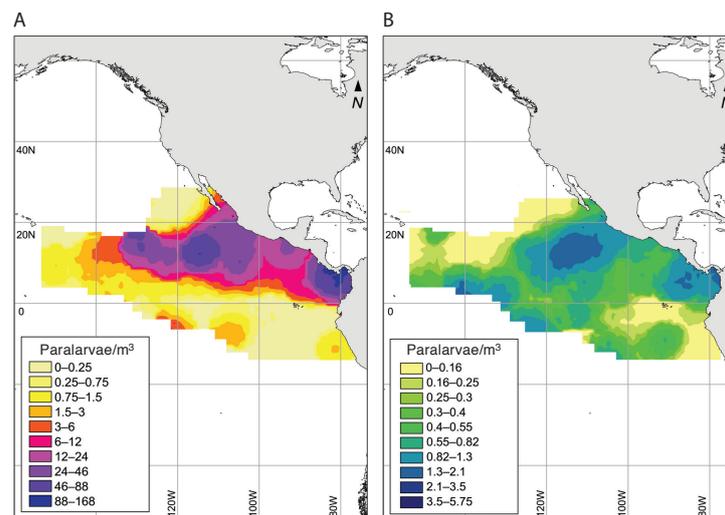


Figure 5.3.: Proposed spawning areas of *Sthenoteuthis oualaniensis* and *Dosidicus gigas* complex (SD-complex) (Staaf et al., 2013). The images show the interpolated paralarvae concentrations based on samples collected by (A) manta tow and (B) bongo tow (from Staaf et al. (2013)).

However, the continuous increase in pp_{prob} with increasing SST up to 32°C contrasts with the results of *in vitro* experiments, which indicate an optimal temperature range of 15–25°C (Staaf et al., 2011), hatching after 6–9d at 18°C after fertilization, a hatching size variation between 0.9–1.3mm ML and post-hatching growth rate around

5.4. Predicting paralarvae presence probability by sea surface temperature data

4% ML per day at that temperature (Yatsu et al., 1999). Staaf et al. (2011) observed a decrease in development *in vitro* above 28°C, while Rosa et al. (2013) expect successful embryonic development to occur only “in areas of the eastern Pacific where temperature at the pycnocline is within that thermal range [15–25°C]”. For the main spawning areas in the tropical Pacific, Arkhipkin et al. (2014) report a SST above 25°C.

The above discrepancy could be explained by (1) a lower temperature tolerance for *in vitro* eggs, (2) a below surface water temperature that is lower than SST, and (3) the fact that pp_{prob} values represent the average probability of occurrence of paralarvae up to a given temperature level and not the probability of hatching for that temperature level. The latter possibility highlights the need to take account of vertical temperature variation when interpreting data on the relation between SST and the probability of occurrence of paralarvae. Water column temperature measurements indicate that when $SST \approx 30^\circ\text{C}$ the temperature at the floating depth of the eggs is 2–3°C lower (Birk et al., 2016; Camarillo-Coop et al., 2010; Staaf et al., 2008).

Conversely, if one accepts the postulated strong relation between SST and pp_{prob} , this allows the prediction of the paralarvae presence probability for known SST using a reverse engineering approach. Yu and Chen (2018) adopt this approach to determine habitats where *D. gigas* could be fished sustainably using a model in which SST is correlated with the catch per unit effort. Yu et al. (2019) build on this approach by employing a more elaborate model incorporating SST, net primary production and sea surface height anomaly.

During processing of the sample counts, Staaf et al. (2013) may have ignored the low counts as these would show up as anomalies in the interpolated paralarvae concentrations. This is a reasonable procedure, because the sampling process probably returns some zero counts in areas of low paralarvae concentrations, reflecting the patchiness of paralarvae distribution and the low sampling rate. However, an alternative interpretation of the data, that does not require the removal of these low counts, is discussed in Subsection 5.4.2.

5.4. Predicting paralarvae presence probability by sea surface temperature data

5.4.1. Introduction

Concept. Staaf et al. (2013) postulate a relation between SST and paralarvae presence probability (pp_{prob}). Here a reverse engineering approach is applied, accepting the relation proposed by Staaf et al. (2013) and assuming that *S. oualaniensis* and *D. gigas*

5. Identifying spawning areas of *D. gigas*

express the same behavior in early developmental stages, to “predict” pp_{prob} from SST. The retrieved pp_{prob} values are displayed on a heat map, where the regions of higher pp_{prob} indicate possible spawning areas.

Methods. The Jet Propulsion Laboratory publishes the dataset MUR-SST (Multi-scale Ultra-high Resolution Sea Surface Temperature) (JPL MUR MEaSURES Project, 2015) giving daily $1\text{km}\times 1\text{km}$ high resolution SST data. These SST data files contain information about the temporal and spatial structure of the data values, the SST data itself and the corresponding error values. The satellite based SST data is of high accuracy, but adverse conditions, e.g., clouds, generate “data-voids” which are filled with data from other satellite-based sources (different sensors) and surface observations by ships and buoys (PO.DAAC, 2021a).

The PYTHON programming language (Python Software Foundation, 2020) was used to parse and process the files and to display the SST data, in combination with the library NETCDF4-PYTHON (Netcdf4-python, 2020), the scientific computing library NUMPY (SciPy.org, 2020b) for large array computations, and the MATPLOTLIB (SciPy.org, 2020a) and BASEMAP (Whitaker, 2018) libraries. The large amount of data retrieved was first checked for consistency and then pre-processed prior to extracting the statistical information. The “region of interest”, a rectangular area covering 60°N to 60°S and 159°W to 69°W , was extracted for each day and converted to a NUMPY-based data format for faster access and computation speed.

15°C	16°C	17°C	18°C	19°C	20°C	21°C	22°C	23°C
0.022	0.030	0.037	0.052	0.075	0.104	0.149	0.187	0.261
24°C	25°C	26°C	27°C	28°C	29°C	30°C	31°C	32°C
0.328	0.418	0.507	0.582	0.672	0.746	0.821	0.866	0.896

Table 5.2.: Values of pp_{prob} of *D. gigas* for incremental values of SST based on the interpolation of data in Staaf et al. (2013).

Staaf et al. (2013) hypothesize that paralarvae presence probability (pp_{prob}) is a function of temperature, i.e., $pp_{\text{prob}}(x) = f(x)$, with x in $^\circ\text{C}$. Here this function is represented using the interpolated linear values shown in Table 5.2 to associate each temperature value with a corresponding pp_{prob} . Table 5.2 itself was retrieved from the values for temperature and pp_{prob} shown in Figure 3 in Staaf et al. (2013).

5.4. Predicting paralarvae presence probability by sea surface temperature data

The relations between pp_{prob} and SST shown in the table are approximate, but inaccuracies arising during the construction of the table are considered marginal in comparison with those arising during processing and interpolation of the SST source data. In all cases, the degree of uncertainty is unknown. The results should therefore be interpreted as providing *qualitative* pp_{prob} information and indicating probable spawning areas; they should not be used to compute accurate probability values.

Further assumptions are made based on the limited knowledge of life history of *D. gigas* as the basis for the prediction of spawning areas:

- *Egg masses float at a depth of 9–16m.* The typical floating depth of the egg masses is set to 9–16m to correspond to the depth of natural *D. gigas* egg masses reported in Staaf et al. (2008) and Birk et al. (2016).
- *Successful embryonic development is bound to the range 15–25°C.* The examination of *in vitro* fertilized eggs from *D. gigas* indicates a strong relationship between temperature, successful hatching and embryonic development. Temperatures below 15°C and above 25°C inhibit successful hatching in *in vitro* experiments (Staaf et al., 2011).
- *SST is a strong indicator of temperature conditions for successful embryonic development.* Water temperatures in upper layers are correlated with the SST but temperature levels are lower (Camarillo-Coop et al., 2010; Staaf et al., 2008), so a high SST (>25°C) does not necessary mean that corresponding upper layer water temperatures are outside the bounds for successful embryonic development. The data in Figure 3 in Staaf et al. (2013) indicate that optimal conditions for paralarvae presence occur at SST of $\approx 28^\circ\text{C}$. Vecchione (1999) observed peak abundance of paralarvae (of different species, but including *S. oualaniensis*) at 27.5°C to 31°C SST.
- *D. gigas females spawn at locations where environmental conditions are suitable for successful embryonic development.* The survival of semelparous species depends on the successful reproduction of the current generation and therefore on the selection by the female of suitable spawning grounds.
- *Paralarvae abundance indicates suitable environmental conditions for spawning.* Paralarvae abundance likely indicates suitable temperature conditions for spawning because of (1) the limited horizontal swimming capability of paralarvae and (2) the short time period between spawning and the end of the paralarval stage,

5. Identifying spawning areas of *D. gigas*

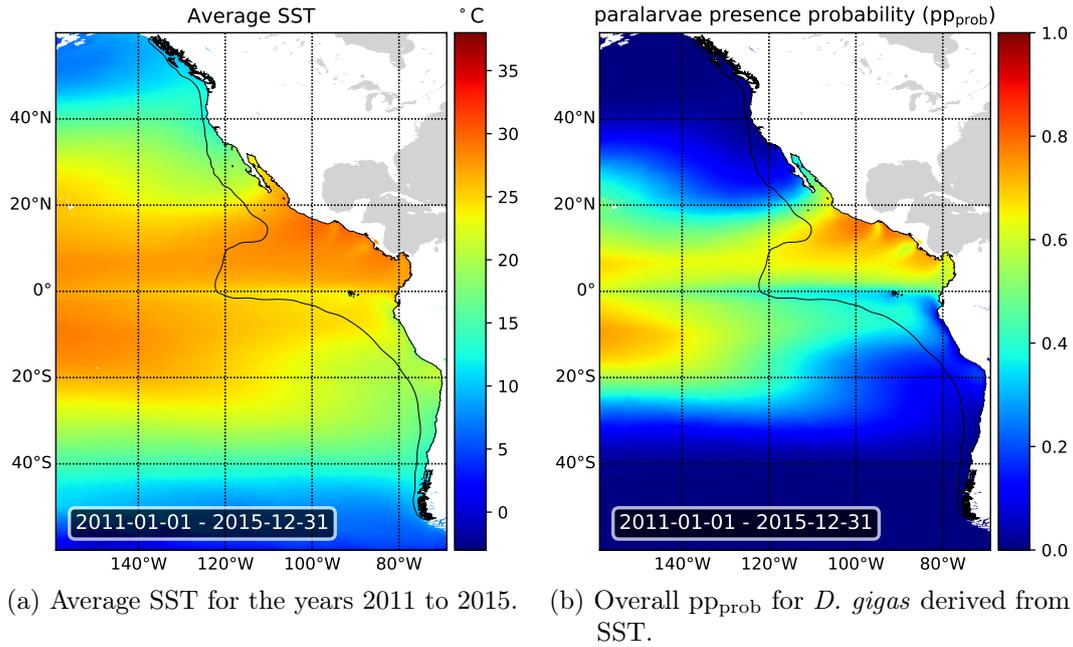


Figure 5.4.: SST and *D. gigas* pp_{prob} averaged over the years 2011–2015. The black line marks the distribution range of *D. gigas*.

which limits the potential for passive dispersal by marine currents. Factors that may reduce paralarvae abundance, like predation, are not considered.

5.4.2. Computing the pp_{prob} by the averaged SST

Method. Estimated pp_{prob} is calculated for each location using the probabilistic function in Staaf et al. (2013) and averaged SST data for the years 2011 to 2015 (Figure 5.4a). The average SST is higher in the presumed distribution area of *D. gigas* in the northern hemisphere than in the southern hemisphere, with highest values recorded in the GOC, and off the coast of Mexico and Central America down to the equator. A horizontal thermocline is associated with the North Equatorial Current, the Equatorial Countercurrent and the South Equatorial Current near the equator. The application of the inverse probabilistic function to the averaged SST data yields the distribution of pp_{prob} (Figure 5.4b).

Discussion. Compared to the northern hemisphere coast (0°N to 20°N), the pp_{prob} near the southern hemisphere coast (0°S to 20°S) is significantly lower except near the equator. Figure 5.5 shows annual averaged SST data for the years 2001–2015 during

5.4. Predicting paralarvae presence probability by sea surface temperature data

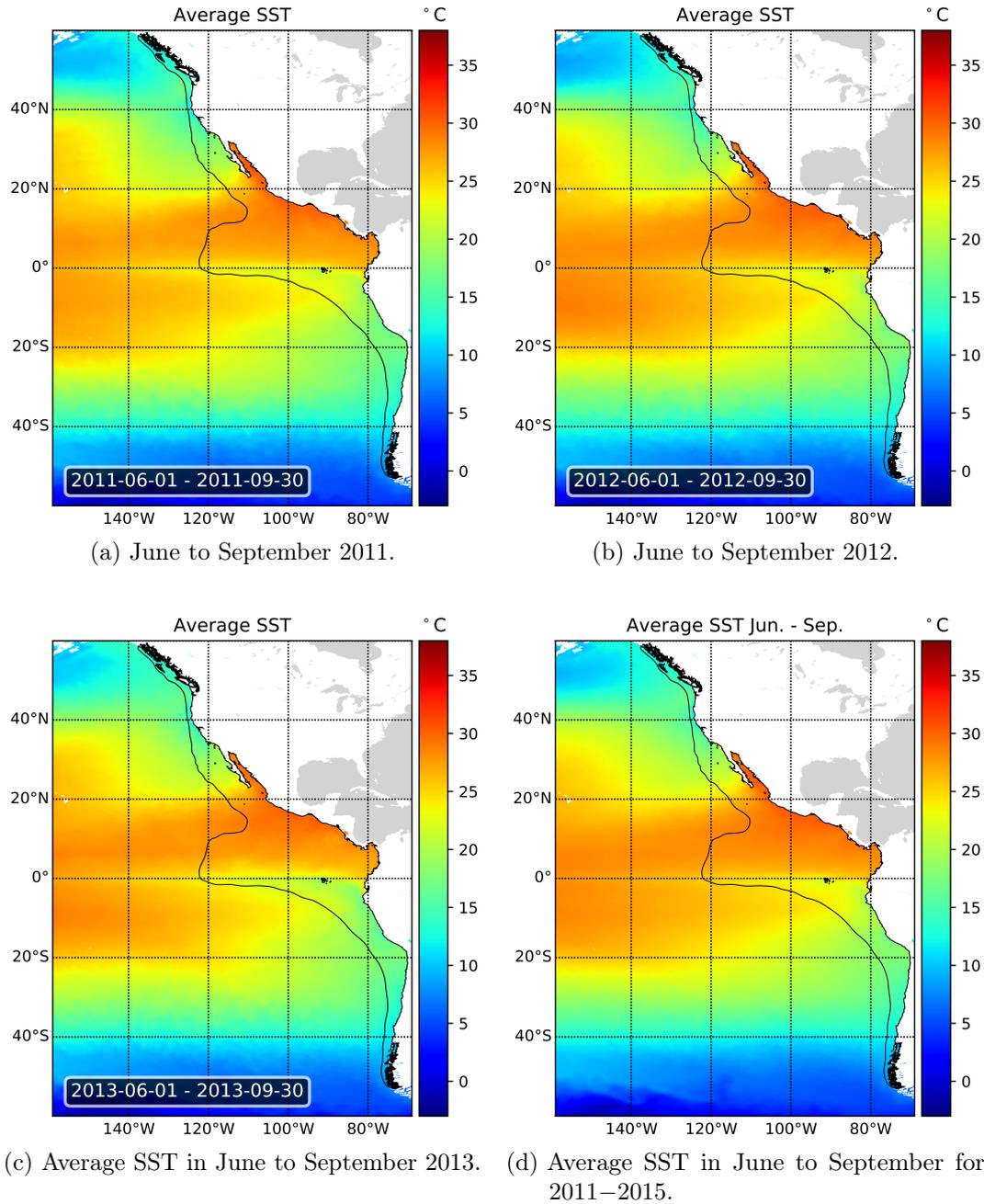


Figure 5.5.: Average SST in June to September. The black line marks the distribution range of *D. gigas*.

the spawning season in the northern hemisphere, i.e., June to September; Figure 5.7 shows the same data for the spawning season in the southern hemisphere, i.e., October to January. However, the SST is consistently higher in the northern hemisphere, even

5. Identifying spawning areas of *D. gigas*

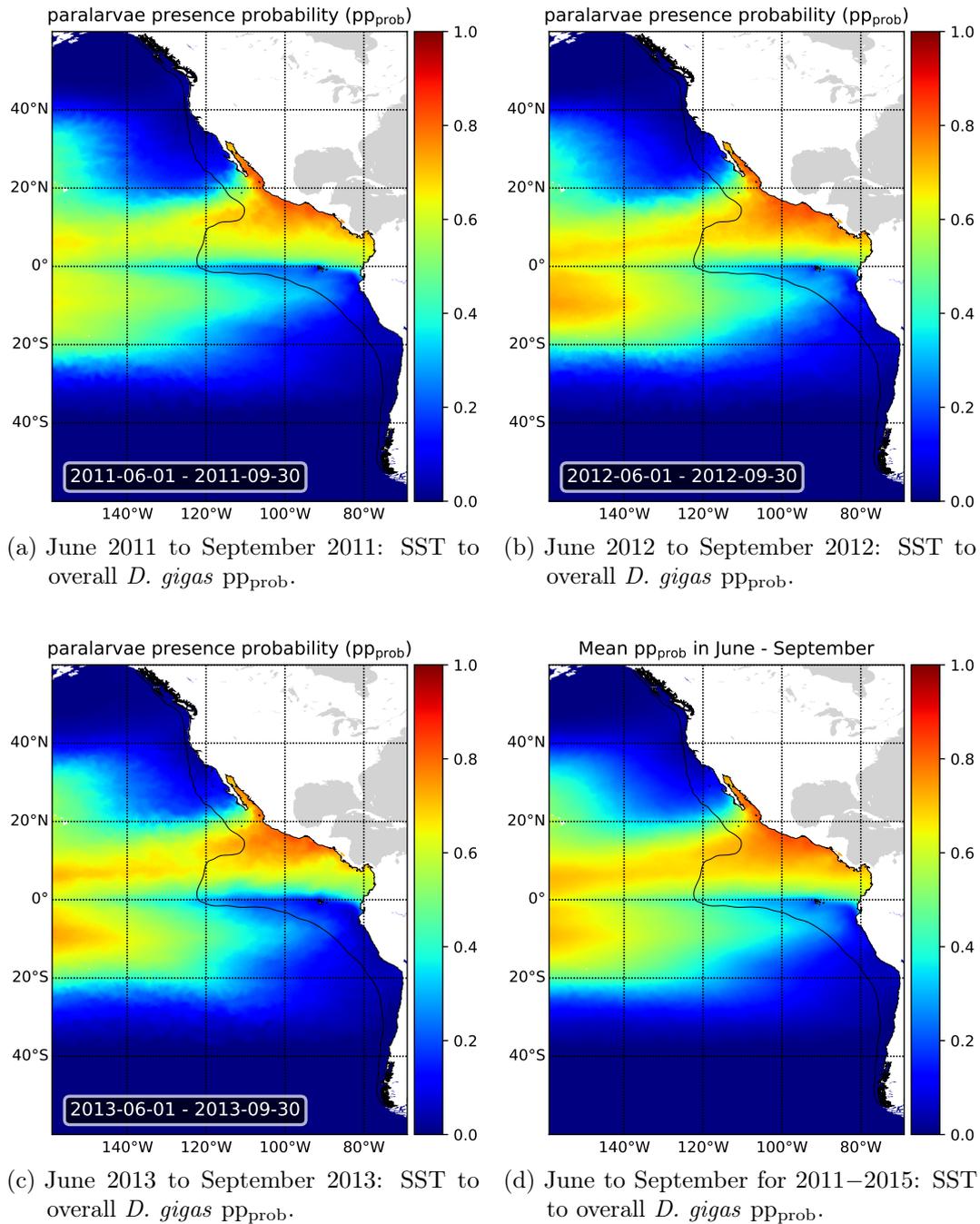


Figure 5.6.: *D. gigas* pp_{prob} for June to September computed from SST data. The black line marks the distribution range of *D. gigas*.

during the southern hemisphere spawning season. The SST in the southern hemisphere is more volatile and changes significantly during both hemispheric spawning seasons.

5.4. Predicting paralarvae presence probability by sea surface temperature data

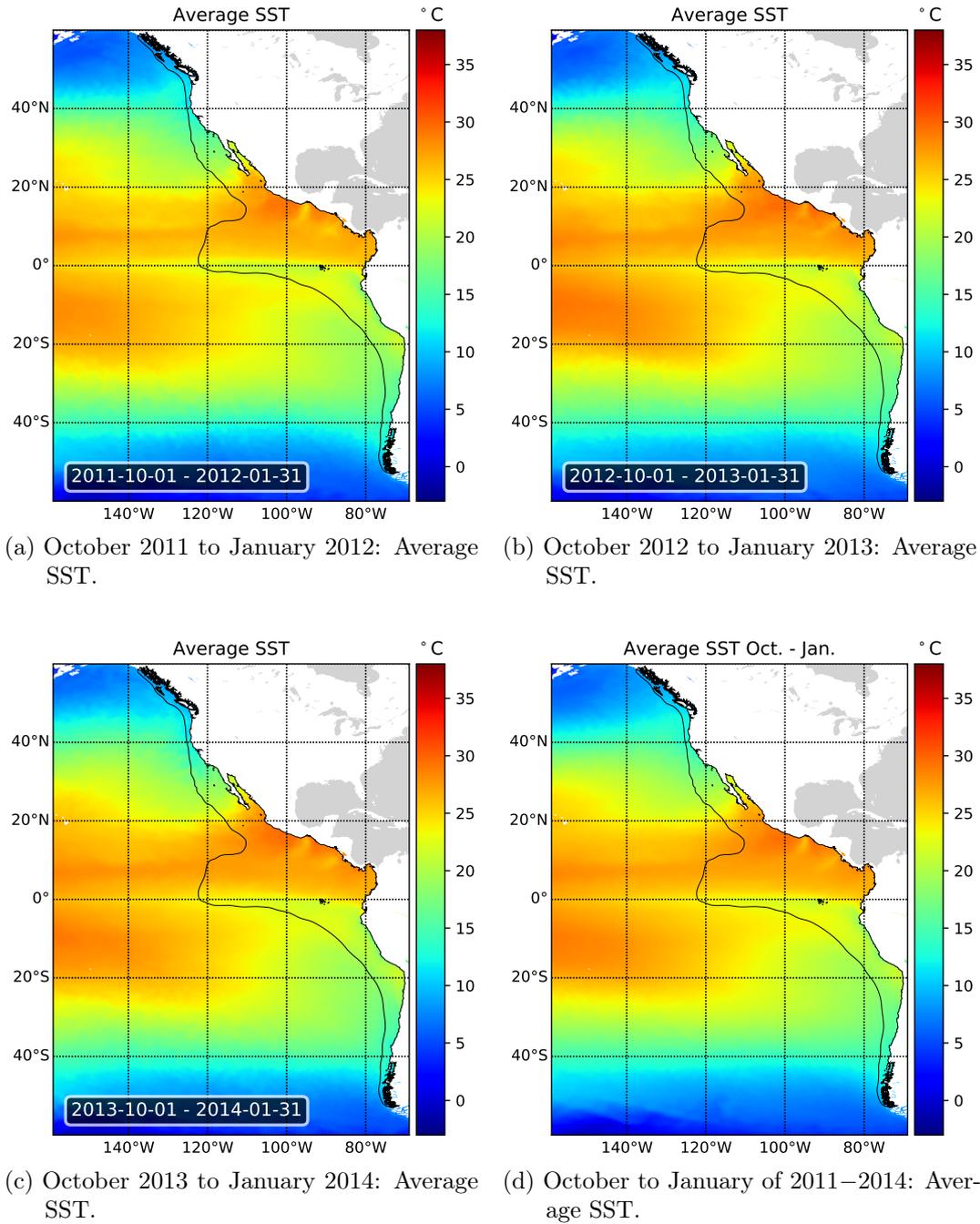


Figure 5.7.: Average SST in October to January. The black line marks the distribution range of *D. gigas*.

Figure 5.6 and Figure 5.8 show corresponding pp_{prob} values. In the northern hemisphere, pp_{prob} varies significantly across a latitudinal gradient, reaching a peak just north of the

5. Identifying spawning areas of *D. gigas*

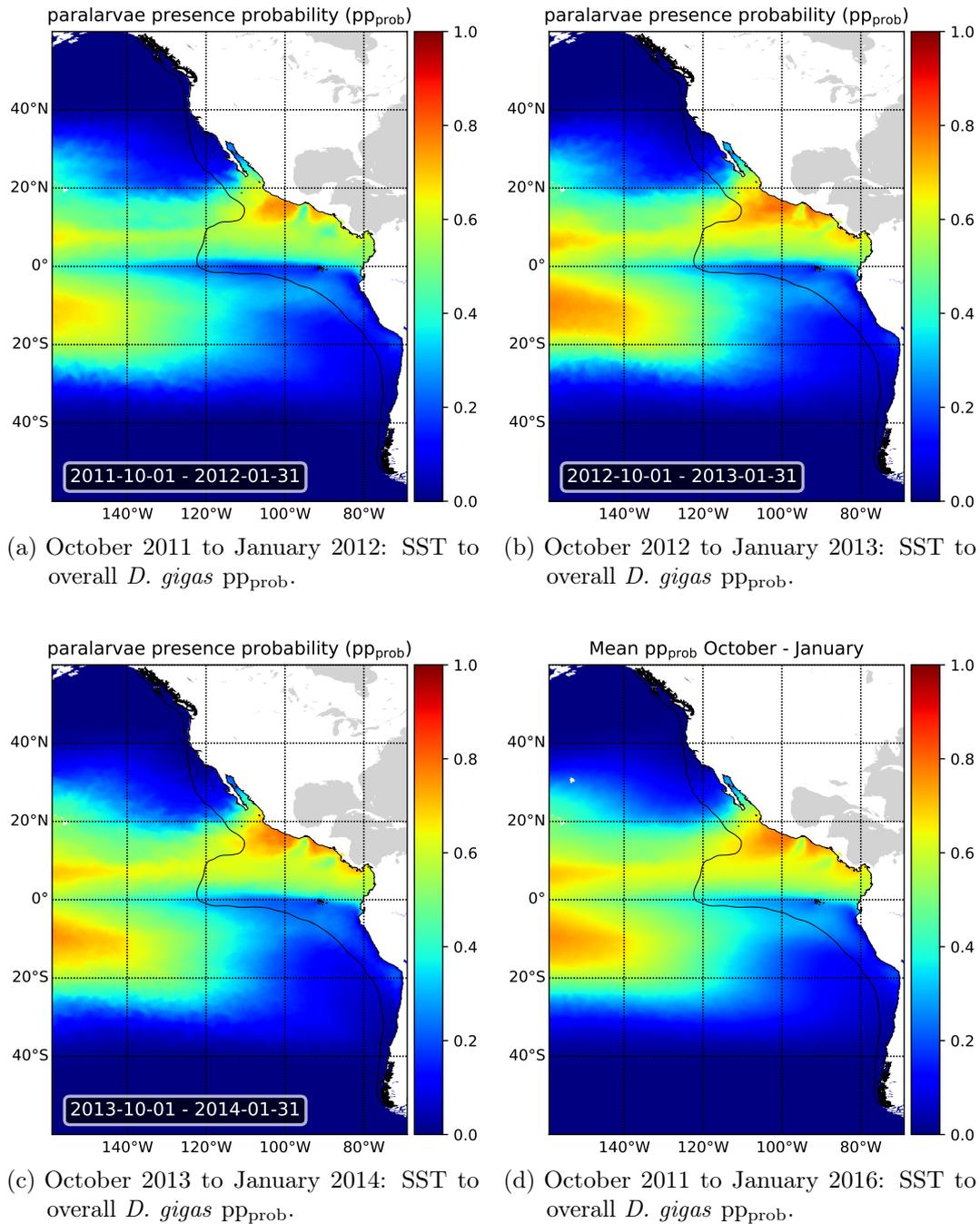


Figure 5.8.: *D. gigas* pp_{prob} for October to January computed from SST data. The black line marks the distribution range of *D. gigas*.

equator. In the southern hemisphere pp_{prob} exhibits low values along the entire South

5.4. Predicting paralarvae presence probability by sea surface temperature data

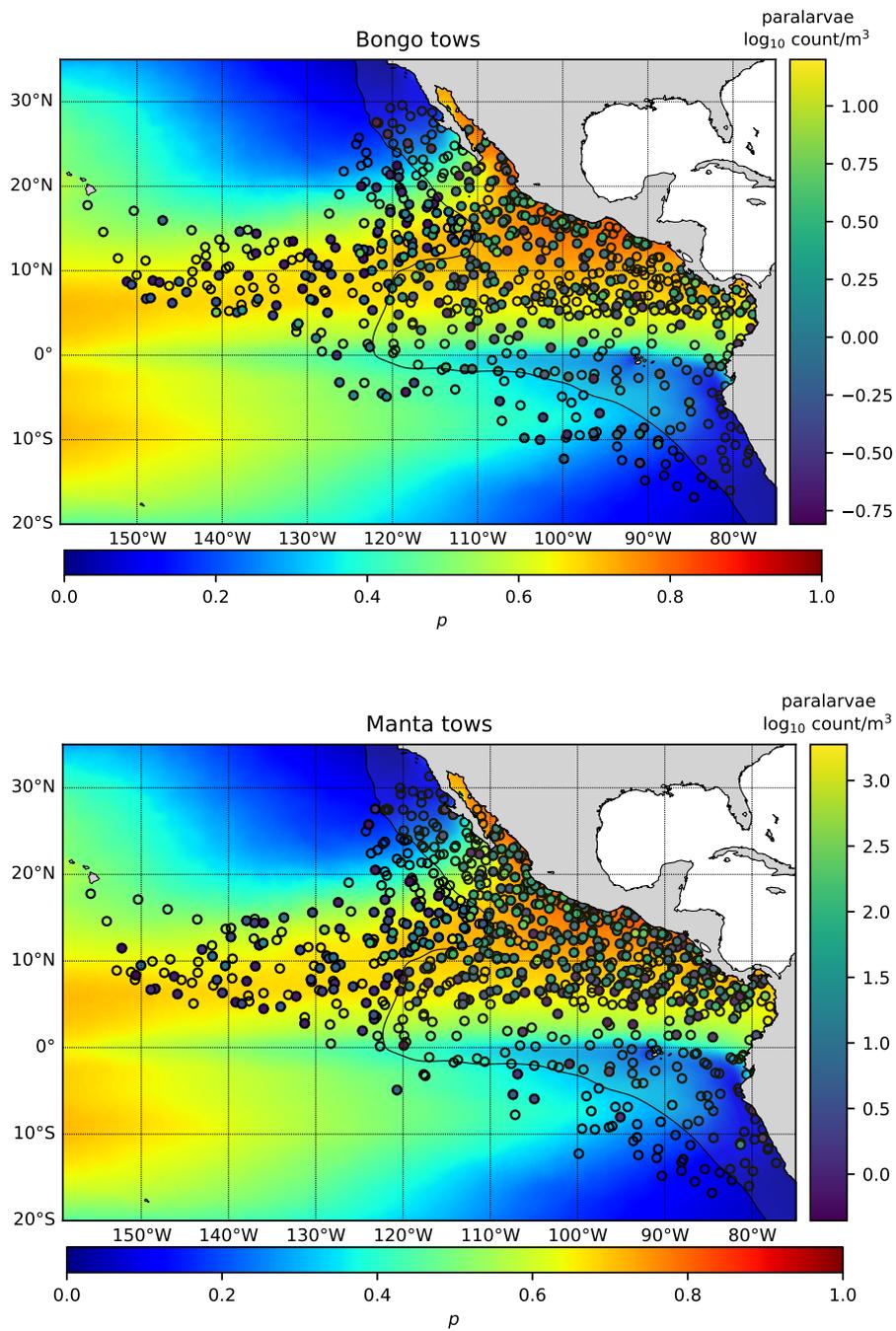


Figure 5.9.: SD-complex paralarvae bongo (subsurface) and manta (surface) log 10-transformed tow sampling counts (Figure 5.2) overlaid on estimated pp_{prob} calculated from SST data (Figure 5.6d). The black line marks the distribution range of *D. gigas*.

5. Identifying spawning areas of *D. gigas*

American coast. Estimated pp_{prob} is significantly higher in the northern hemisphere than in the southern hemisphere during both spawning seasons.

Figure 5.9 overlays actual sample counts onto computed pp_{prob} values. Low sample counts generally lie within areas with low pp_{prob} and high sample counts lie within the area of higher pp_{prob} . However, there are also concentrations of low sample counts close to high sample counts in adjacent regions with a high pp_{prob} . These low sample counts make it difficult to use interpolated data derived from sample counts to identify spawning areas; i.e., it was not possible to follow the procedure adopted by Staaf et al. (2013). Low sample counts close to high sample counts decrease interpolated concentrations in adjacent areas. When these low sample counts were included in the dataset, interpolation was not successful in identifying any spawning areas.

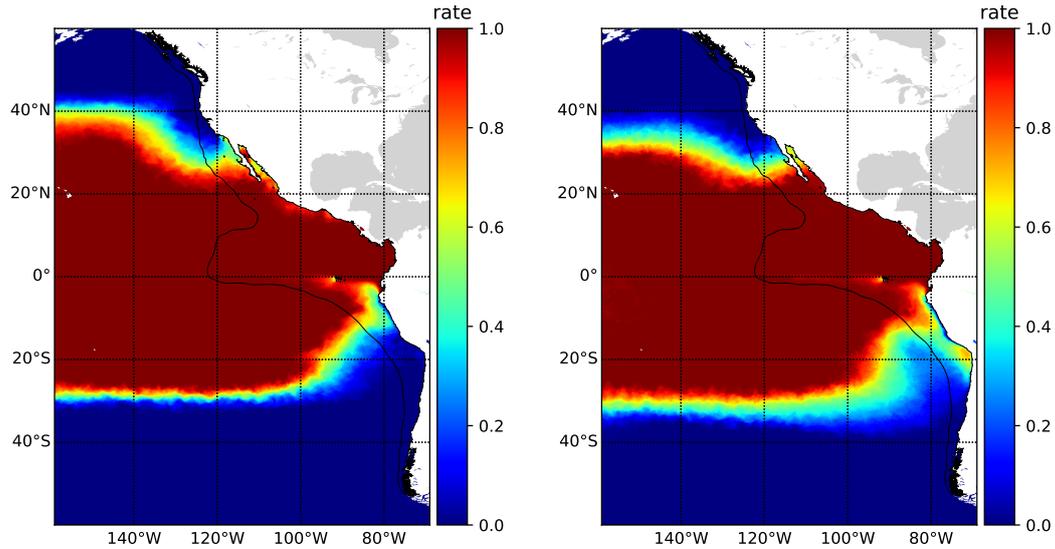
The SST data predicted possible spawning grounds over a wide area, i.e., 0°N to 20°N , but eggs or paralarvae were found only at a few locations within this area (Staaf et al., 2013). The single positively identified occurrence of *D. gigas* paralarvae (Figure 5.11) in the northern hemisphere was in a sample that also contained *Sthenoteuthis oualaniensis* paralarvae, but *D. gigas* paralarvae were not found in other samples containing *Sthenoteuthis oualaniensis*. Given, the high potential fecundity of *D. gigas*, one would expect to find many more paralarvae in the vicinity of a spawning ground.

The generally higher pp_{prob} above the equator implies a higher *D. gigas* abundance in the northern hemisphere. The low pp_{prob} south of the equator corresponds to the distribution of the samples (Figure 5.2). The postulated relation between SST and pp_{prob} matches the distribution of occurrences both hemispheres, but there are many samples with a zero count in regions with high probability of paralarval presence (Figure 5.9) including within the postulated spawning areas (Figure 5.3). Zero count samples near high count samples call into question the identification of these large areas as potential spawning grounds. The occurrence of high and low sample in close proximity to each other within the high pp_{prob} area is an inconsistency. This suggests that at least one additional explanatory factor in addition to SST is required to account for observed paralarval distribution.

5.4.3. Days of suitable SST for spawning development

The ratio $\frac{\text{days with suitable SST}}{\text{number of days in season}}$ may provide an indication of suitability of areas as spawning grounds if females select spawning grounds based on number of days where water temperatures are suitable for embryonic development. Based on data from 2011–2015, temperatures are suitable for embryonic development for most of the year in the northern hemisphere and in offshore areas in the southern hemisphere, but not

5.4. Predicting paralarvae presence probability by sea surface temperature data



(a) Areas suitable for successful embryonic development from June to September based on average SST in 2011–2015. (b) Areas suitable for successful embryonic development from October to January (2016) based on average SST in 2011–2015.

Figure 5.10.: Areas suitable for successful embryonic development during (a) northern and (b) southern hemisphere spawning seasons, computed from $\frac{\text{days with suitable SST}}{\text{number of days in season}}$. The suitable temperature range is 20.0°C to 31.0°C. The black line marks the distribution range of *D. gigas*.

in coastal waters in the southern hemisphere (Figure 5.10). Significantly higher ratio values occur off the coast near 20°S, but this area is relatively small compared to other suitable regions. In the southern hemisphere, the number of days with suitable SST temperatures for embryonic development increases during October to January when spawning is presumed to take place, (Figure 5.10b), but this effect is limited to a narrow strip near the coast extending southwards from the equator to approx. 20°S with a large gap to the west.

Analysis of the “suitable days ratio” provides support for the existence of a more productive spawning area in the northern hemisphere. It also shows that ratio of suitable to unsuitable days for embryonic development is lower in both hemispheres near the coast. The temperature in the pelagic zone far away from the coast is less volatile and this area appears to present more suitable conditions for spawning.

5. Identifying spawning areas of *D. gigas*

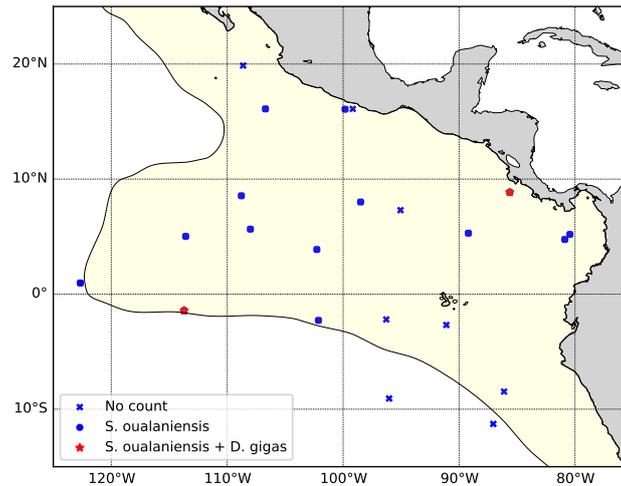


Figure 5.11.: Field sampling locations in Staaf et al. (2013), generated from data provided by D. Staaf (personal communication), of ethanol-preserved samples which allow the identification of *D. gigas* paralarvae. Blue crosses mark samples with a zero paralarvae count, blue circles mark samples with a *S. oualaniensis* paralarvae count and red circles mark samples with both *D. gigas* and *S. oualaniensis* paralarvae counts. The yellow area indicates the distribution range of *D. gigas*.

5.4.4. Summary

SST is an important factor for embryonic development and hatching, but may not be the only predictor of paralarvae occurrence, nor can the SST be used to delimit small areas as potential spawning grounds. The Tropical Eastern Pacific provides suitable SST for developing paralarvae over a large area and for several months of the year, but *D. gigas* paralarvae have only been found at a few locations with no special characteristics in terms of SST and their estimated pp_{prob} . Therefore, future surveys in this region can be expected to reveal more locations where paralarvae are present. Furthermore, regions with a high pp_{prob} yielded highly variable sample counts that cannot be explained exclusively by SST, which indicate the existence of a spatial organizing process, i.e., currents which aggregate paralarvae in specific areas within these regions.

5.5. Analyzing egg and paralarvae transport and dispersion

Problem description. A broad area matches the SST criterion for suitability for spawning, but data analysis of bongo and manta net samples, with only two identified paralarvae occurrences (Figure 5.11) over a period of eight years, indicates a heterogeneous

5.5. Analyzing egg and paralarvae transport and dispersion

distribution, with many small spawning areas sparsely distributed over a large region suitable for spawning. The Pacific current system may function as a spatial organizing process by transporting material, i.e., eggs, paralarvae and juveniles, to locations far away from their source of origin (R. I. Ruiz-Cooley et al., 2010). The passively floating early development stages of *D. gigas* (Boyle & v. Boletzky, 1996; Liu et al., 2016; Staaf, 2010) are exposed to Pacific currents from 9m to 16m depth, where floating eggs masses have been found (Section 5.4), to the surface. C. Nigmatullin et al. (2001) state that passive horizontal migration also transports juveniles and, sometimes, adults. An assumed average flow velocity of 0.5m s^{-1} (1.8km h^{-1}) near the equator would, for example, result in a displacement of 43.2km per day. Assuming 30 days between spawning and development of swimming ability, when active horizontal movement becomes possible, eggs and paralarvae could be passively transported to up to approx. 1300km from the spawning ground. This mechanism would facilitate dispersal of the eggs and allows young specimens to move to higher latitudes.

On a more granular level, the Pacific current system may aggregate eggs and paralarvae to certain attracting regions (*attractors*) from other regions (*donators*). If the Pacific currents function as a spatial organizing process in this way, then the attractors would have a higher pp_{prob} (i.e., eggs and paralarvae would occur more frequently) and could therefore be misidentified as spawning grounds. Furthermore, a closed system with attractors requires donators, i.e., areas from which material is transported away, giving rise to low or zero sample counts. This hypothesis aligns with the highly volatile sample counts reported in regions with similar SST characteristics.

Methods. To investigate this postulated phenomenon, (surface) material transport flows in the Eastern Pacific were analyzed in order to identify possible attractors and donators of eggs and paralarvae. Data on currents obtained from OSCAR (Ocean Surface Current Analysis) project (California Institute of Technology, 2009) were projected onto a CA-like structure for computation and analysis of the amount of incoming and outgoing material for each cell.

The OSCAR currents data (PO.DAAC, 2021b) comprise annually organized time-stamped datasets of water movements since 1993. The datasets have a resolution of $\frac{1}{3}^\circ$, are based on satellite data, models and local real-time measured currents, and are averaged over five days. The PYTHON language and tools were used as in Section 5.4 to process and visualize the data.

The datasets contain, for each $\frac{1}{3} \cdot \frac{1}{3}$ -sized cell, the vectors u (flow in west/east direction) and v (flow in north/south direction) that span the flow information $\vec{v} = u + v$

5. Identifying spawning areas of *D. gigas*

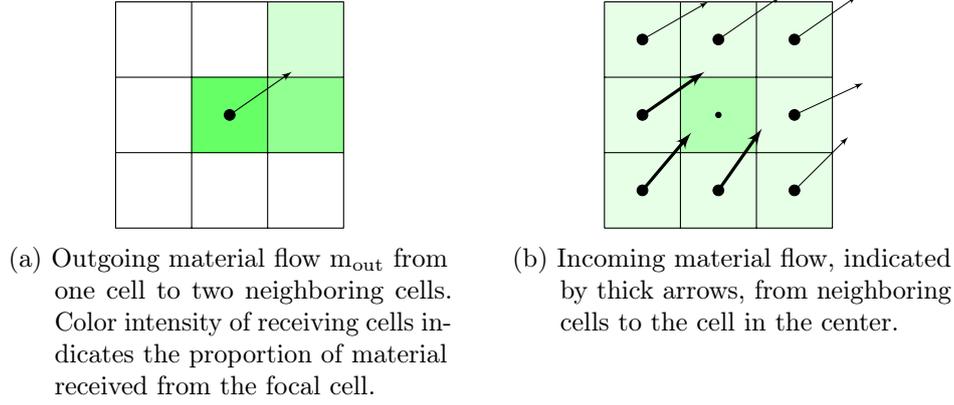


Figure 5.12.: Material flow affecting the focal cell, located in the center of the diagram.

in meters per second (m s^{-1}). The vector \vec{v} defines a point $P(u, -v)$ in the Cartesian coordinate system, but the north/south direction of v is reversed so that the sign of v is changed in computations. The equivalent polar coordinate $P(r, \angle\varphi)$ describes from the origin of a vector the point $P(u, -v)$ as a radius r and direction φ ; r and φ are given by

$$r = \sqrt{u^2 + v^2} \quad (5.2)$$

and

$$\tan \varphi = \frac{v}{u} \quad (5.3)$$

The radius r denotes the velocity and φ the direction of the flow. The velocity r determines the flow of outgoing material m_{out} per time unit of a cell and φ the identity of the two receiving cells (Figure 5.12) and the proportion of m_{out} of the incoming material that each cells receives.

The amount of outgoing material m_{out} is given by edge length of the cell $c_{\text{edge length}}$ in meters (m), flow velocity r in m s^{-1} and the time t in seconds (s). The units cancel down to the dimensionless material flow m_{out} :

$$m_{\text{out}} = t \frac{r}{c_{\text{edge length}}} \quad (5.4)$$

The $c_{\text{edge length}}$ becomes smaller from the equator to the poles, thus $c_{\text{edge length}}$ is calculated taking the Earth to be a perfect sphere and ϕ as the hemisphere latitude ($0-90^\circ$) by

$$c_{\text{edge length}} = \frac{2\pi \cdot r_{\text{Earth}} \cdot \cos \phi}{360} \quad (5.5)$$

5.5. Analyzing egg and paralarvae transport and dispersion

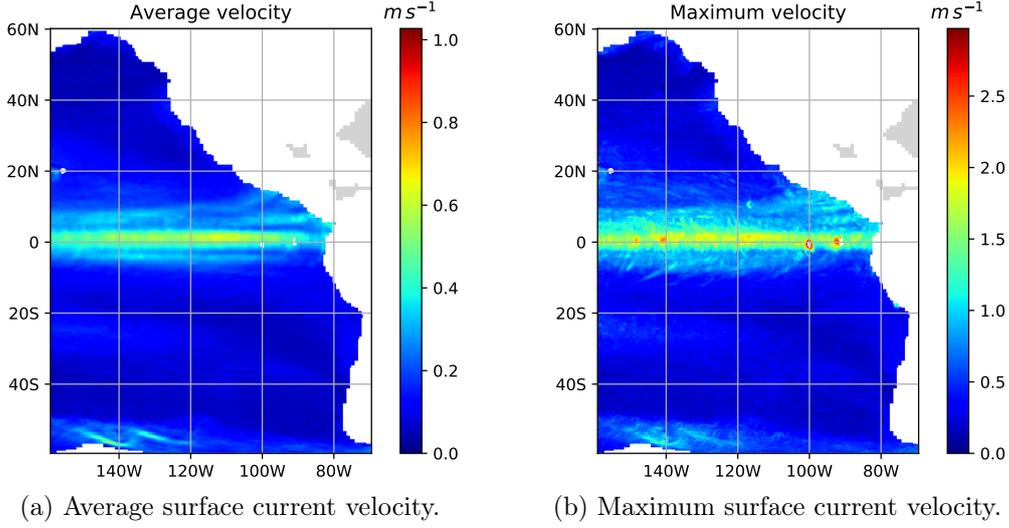


Figure 5.13.: Surface current velocity from January 2011 to December 2016. Maps are visualizations of average values for $\frac{1}{3} \cdot \frac{1}{3}^\circ$ cells in the OSCAR datasets.

and approximates with $r_{\text{earth}} = 6371 \text{ km}$ to

$$c_{\text{edge length}} \approx 111.2 \text{ km} \cdot \cos \phi \quad (5.6)$$

and a range of cell sizes from approx. 111.2km at the equator to approx. 55.6km at 60° latitude.

To evaluate possible long-distance material transport, the flow velocity in m s^{-1} of each cell is retrieved and displayed (Figure 5.13). To compute the material transport on a more granular level, each cell is initialized to the same value (amount of material) at start of the simulation (Figure 5.14) and then the attractors and donators are identified by means of iterative computations of the material flow. Attractors increase their values at the cost of the donators and a heat-map is used to visualize the amount of material lost or gained. The timespan, corresponding to the period from spawning to the time when an individual paralarvae is capable of active horizontal movement, is conservatively estimated to 30 days. Simulation runs are initiated from different starting times to show the effects of changing currents during the year. The non-associative calculations are performed in separate sequences of computation steps and more than one start time is necessary to cover the duration of each spawning season.

The OSCAR datasets have an update interval of five days, so each start day was rounded down to the beginning of its interval, e.g., the starting day 2013-06-01 was shifted to 2013-05-28. Sequences of 30-day material flows were computed for the starting

5. Identifying spawning areas of *D. gigas*

days of all intervals and, in a next step, selected sequences of material flow of the peak spawning seasons, June to September for the northern hemisphere and October to January for the southern hemisphere, were analyzed.

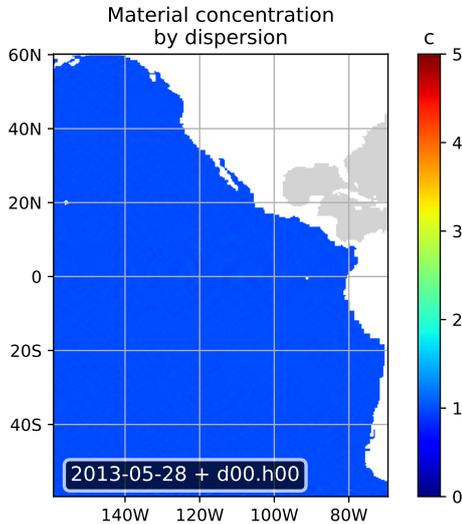


Figure 5.14.: The initial setup for material flow computation of date 2013-05-28 plus an offset of 0 days and 0 hours of the CA.

Each cell starts a simulation run with an initial value of 1.0 (Figure 5.14). The next state of a cell results from changes in its own state and those of neighboring cells. Thus, for each time step, it is essential that no directly neighboring cell is omitted from the analysis. A neighboring cell is skipped (i.e., omitted from the calculations) if Δt multiplied with the velocity of the current of this cell is larger than the c_{edge} length. Preliminary analysis indicated that a time step of $\Delta t > 8$ h would result in skipping at least one of the cells, so a time step resolution $\ll 8$ h is safe and $\Delta t = 1$ h was selected for the simulation runs. After the first simulation step, cell values > 1 indicate attractors and values < 1 indicate donators.

Model discussion in relation to other dispersion models. This particular approach benefits from the floating depth of the egg masses of *D. gigas*, which matches the given depth range of the satellite-sensed data. Thus this data can be used as input data for a hydrodynamic model to simulate egg and larvae dispersion. Literature on modeling currents and egg dispersion is sparse due to the complexity involved in implementing such models. Previous studies apply numerical methods to implement a hydrodynamic model (i.e., the water currents) and track the behavior of passive particles (i.e., the eggs/larvae) in the hydrodynamic model (Ådlandsvik et al., 2004; Han & Kulka, 2009; Mariani et al., 2010; Marinone et al., 2004; Stockhausen & Hermann, 2007). Mariani et al. (2010) use a mixture of partial differential functions for the hydrodynamic model and individual-based modeling (IBM) for passive floating particles to model the dispersal of eggs and larvae of the Atlantic bluefin tuna (*Thunnus thynnus thynnus*). Han and Kulka (2009) replace the partial differential functions by a finite-element circulation model, based on coarse spatial and temporal data. Stockhausen

5.5. Analyzing egg and paralarvae transport and dispersion

and Hermann (2007) precomputed the “currents fields” in the model using a numerical approximation and used particles to simulate rockfish (*Sebastes spp.*) eggs and larvae dispersal. Ådlandsvik et al. (2004) used a modified Princeton Ocean Model for the Greenland halibut (*Reinhardtius hippoglossoides*) and for particle tracking a Lagrangian Advection and Diffusion Model, which is not fully specified but, from its description, appears similar to an IBM.

All these approaches use a numerical approach to model currents and an IBM for modeling dispersal. This thesis uses spatio-temporally resolved satellite-sensed data on currents, and computes dispersal using a *directional diffusion model* that directly approximates to a partial differential model. The identification of donating and attracting areas is the purpose of the model; therefore it computes the change in concentration of material for each cell in each time step. The grid resolution of the datasets, i.e., the current data, limits the spatial resolution of the model, while the number of datasets per time limits the temporal resolution. An IBM approach would require tracking a huge number of individual particles in order to identify attracting and donating areas. This computationally expensive procedure is avoided here by representing the general material flow as a (quantitative) “particle concentration” change per cell and time. This procedure could also be applied to simplify the implementation of the models developed by authors cited in this paragraph.

The model presented here may not accurately predict actual attractors and donators because of the relatively coarse spatio-temporal resolution and other probable limitations of data quality. Rather, the intention is to assess the likelihood of the existence of attractors and donators, which allows qualitative conclusions to be drawn.

Discussion. At beginning of this section, in discussing the scenario of a passive transport of eggs or paralarvae over long distances, as suggested in previous studies (C. Nigmatullin et al., 2001; P. G. Rodhouse, 2008; R. I. Ruiz-Cooley et al., 2010), an average current velocity of 0.5m s^{-1} was assumed. This initial assumption likely overestimates the passive transport of eggs and paralarvae because the average current velocity only attains this value near the equator and drops to approx. 0.2m s^{-1} towards the poles (Figure 5.13a). The analysis of maximum current velocities (Figure 5.13b) yields a similar pattern with highest maximum velocities clustered around the equator. In addition, the high-velocity currents align with the North Equatorial Current and South Equatorial Current, flowing from the east to west, and the Equator Counter Current, flowing from west to east, so that the currents will not transport eggs, paralarvae or young passive floating individuals towards the poles (Anderson and Rodhouse, 2001,

5. Identifying spawning areas of *D. gigas*

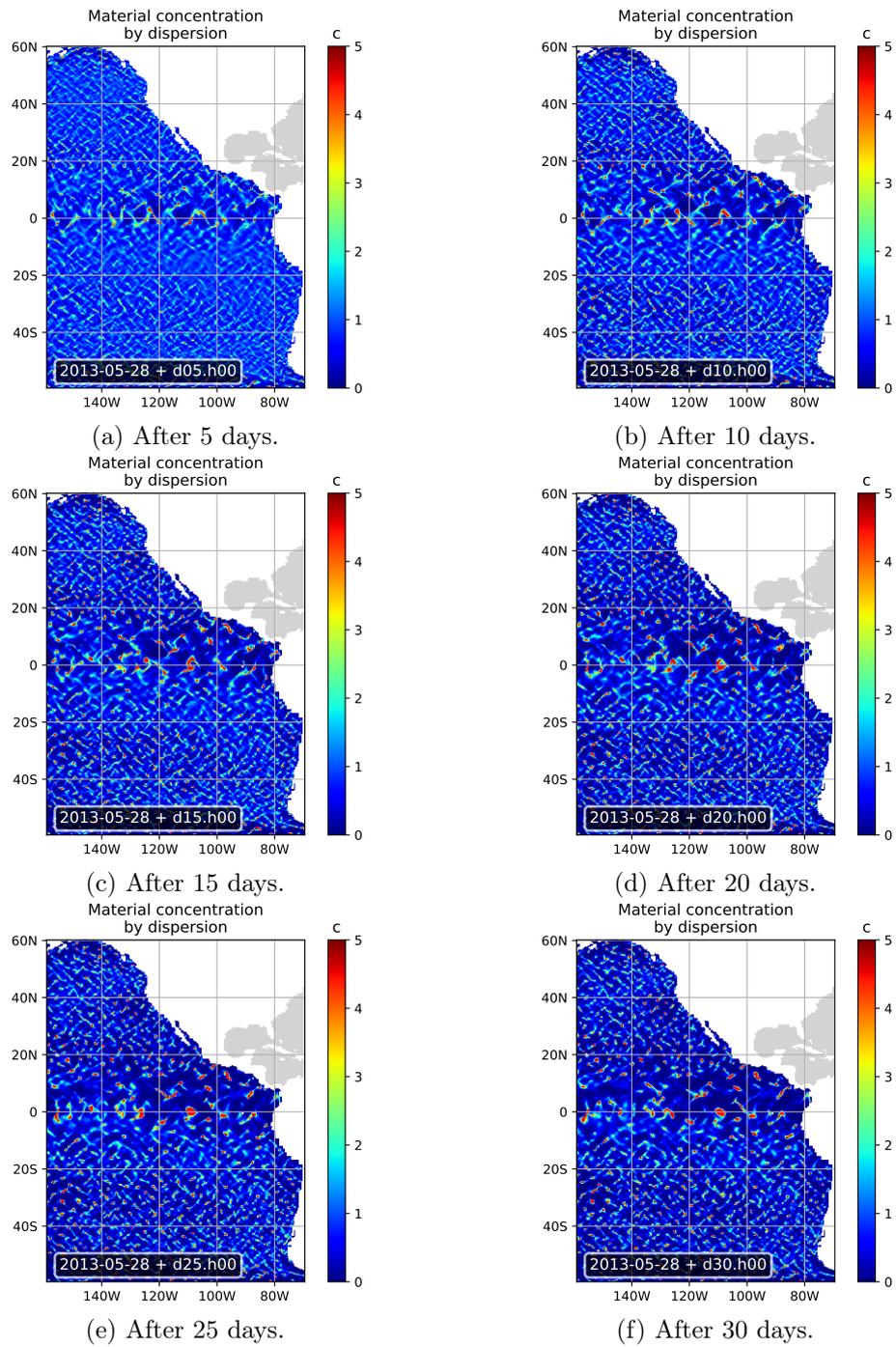


Figure 5.15.: Emerging attractors and donators for egg/paralarvae dispersion in the northern hemisphere for the year 2013.

5.5. Analyzing egg and paralarvae transport and dispersion

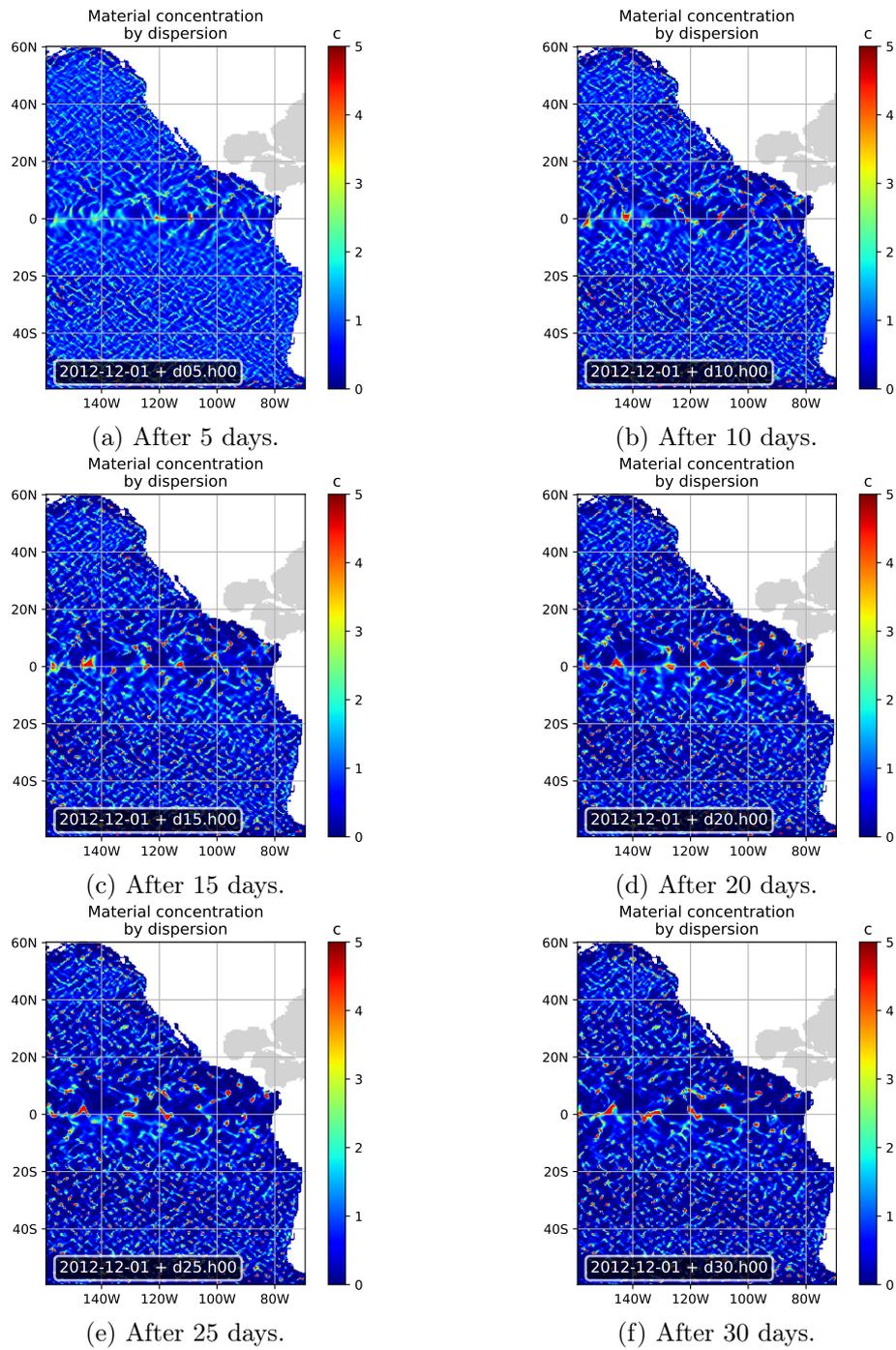


Figure 5.16.: Emerging attractors and donators for egg/paralarvae dispersion in the southern hemisphere for the year 2012.

5. Identifying spawning areas of *D. gigas*

Fig.4). Furthermore, the California Current and Peru Current transport material from the poles in direction of the North Equatorial Current and South Equatorial Current respectively (Anderson and Rodhouse, 2001, Fig.4). This combination of longitudinal and latitudinal flows would contain any passive floating material within an approximate latitudinal range of 10°N and 5°S.

Within these general current flows, areas of discriminative material concentration emerge, generating attractors and donators, which appear in dispersal simulations after a short time and develop more quickly in areas with higher velocity currents (Figure 5.15 and Figure 5.16). The generally extensive donating areas loose most of their material to relatively small attracting areas and this reduces the probability of egg/paralarvae presence in the donating areas. Attractors and donators are generally quite stable once established, exhibiting only minor changes their location and shape over a short period, but will of course change their location over a longer timescale (i.e., from year to year).

The coastline between approx. 20°N to 20°S is mainly a donating area and the donating areas become smaller in size (more granular) towards the upper and lower latitudes. The donating areas are wider and more homogeneous in the northern hemisphere compared to the same latitudes in the southern hemisphere. Emerging structures are more pronounced around the equator and in the northern hemisphere.

A self-organizing process creates these emerging structures because each cell transports material in a “preferred” direction, which is determined by the flow direction of the current (Figure 5.12), thereby increasing the amount of material in the neighboring receiving cells. Diffusion differs from directional material transport in that it transports material from the focal cell to *all* of its neighboring cells. If directional material transport prevails over diffusion and its direction does not change erratically over the course of the simulation (i.e., the direction of flow changes slowly over time), then material aggregation and donation is simply a matter of time.

The simulation cannot localize any potential spawning grounds, as the simulation runs show many attractors receiving material from large and possibly distant donating areas. On the other hand, the simulation highlight the difficulties involved in using sampling to locate *D. gigas* spawning grounds. Sampling outside these relatively small attractors may well fail to reveal eggs or paralarvae, even if the sample is taken in a spawning ground, since eggs/paralarvae are aggregated after spawning almost instantly to smaller regions. Moreover simulation runs cannot confirm the hypothesized spawning areas in Staaf et al. (2013), as their data are aggregated over an eight year period, but the attractors appear each year at different locations. The simulation runs do however explain the close proximity of highly different sample counts in Staaf et al. (2013).

5.6. Computing paralarvae origin from attractors

Concept. The material transport simulation runs indicate a process that creates locations of high paralarvae and eggs abundance (attractors) and regions of lower abundance (donators). If eggs/paralarvae are found at a specific location (coordinates) (Table 5.1), and assuming that such a specific location is an attractor, then the spawning grounds can be identified by backtracking based on flow analysis of the currents, i.e., by tracing the material transport back to its origin. This approach is similar to ray-tracing technology in computer graphics, where light rays are traced back from the destination, e.g., a computer screen, to their origin, the light source. The virtual scene objects reflect, refract and also split the light rays into multiple rays. The currents are similar to these objects in this model and the spawning locations are the light sources.

Methods. In the model, an attractor forms a cell c_a with a higher incoming m_{in} than outgoing m_{out} material flow. The outgoing material flow of the neighboring cells $C_{N(c)}$ defines the incoming material flow $m_{in}(c)$ of a cell c . The neighboring cells create a *hull* h to a cell c with $h(c) = C_{N(c)}$. The sum of all $m_{out}(c_n) \forall c_n \in h(c_a)$ defines the incoming material $m_{in}(c_a)$ of an attractor c_a . The cells in $h(c_a)$ that contribute a $m_{in}(c_a)$ must be identified for each time step to determine the origin.

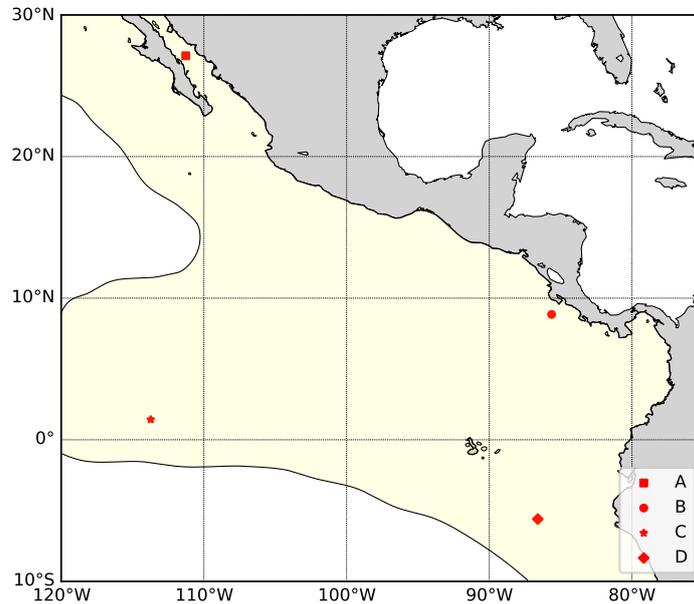


Figure 5.17.: Locations of occurrences of *D. gigas* eggs (A) and paralarvae (B,C,D) used for backtracking simulations.

5. Identifying spawning areas of *D. gigas*

All cells that create for each $c_n \in h(c_a)$ the $m_{in}(c_a) > 0$ are computed. The $m_{out}(c_n)$ of each $c_n \in h(c_a)$ depends for each $c_n \in h(c_a)$ on the $m_{in}(c_n)$, so the material transport of $h(h(c_a))$ must be determined. Each “outer” hull h_{i+1} contributes the m_{in} for each cell in the “inner” hull h_i . The total number of hulls depends on the maximum flow among all cells and the number of days to be tracked. If maximum distance of the material travel per day does not exceed the cell size, then the assignment of one hull per day is safe.

The hulls define all cells that contribute to c_a over a certain period of time. These entities are represented as a network graph where the nodes represent the neighboring cells and the edges contain the material flow. The avoidance of multiple ray (edge) creation, i.e., from the same node targeting the same destination cell, is a major challenge that must be tackled during graph construction.

label	type	date	sampling location	approx. location	Δ
A	eggs	2006-06-19	27.12°N/111.27°W	—	—
B	paralarvae	2006-09-03	8.85°N/85.63°W	8.67°N/85.67°W	0.02°/0.04°
C	paralarvae	2006-09-15	1.43°S/113.72°W	1.33°S/113.67°W	0.1°/0.05°
D	paralarvae	2007-11-15	5.6°S/86.6°W	5.67°S/86.67°W	0.07°/0.07°

Table 5.3.: Approximated locations used for the backtracking of occurrences of *D. gigas* eggs and paralarvae (Figure 5.17). The sampling location is taken from Table 5.1. The “approx. location” is the closest cell in the OSCAR dataset to the actual sampling location, and Δ gives the deviation (distance) from the sampling location to the “approx. location” in degrees. No approximate location is given for Sample A since the dataset for 2016 does not contain data on currents in the Gulf of California.

Discussion. For the simulation runs, the locations of reported egg/paralarvae findings were approximated to the nearest cell of the OSCAR currents dataset (Table 5.3); however the low resolution in the dataset of the year 2006 provides no flow information for the GOC, hence no backtracking was performed for sample A in Table 5.3.

According to the OSCAR data, there is only limited material transport near the coast, and no dispersion was found in the backtracking computations for sample B. So, according to the backtracking simulation, the location where these paralarvae were found is also the actual spawning location (not shown in Figure 5.18).

A vertically stretched contribution area was identified for sample C (Figure 5.18a), but this result is questionable, since a small change to the start date for backtracking

5.6. Computing paralarvae origin from attractors

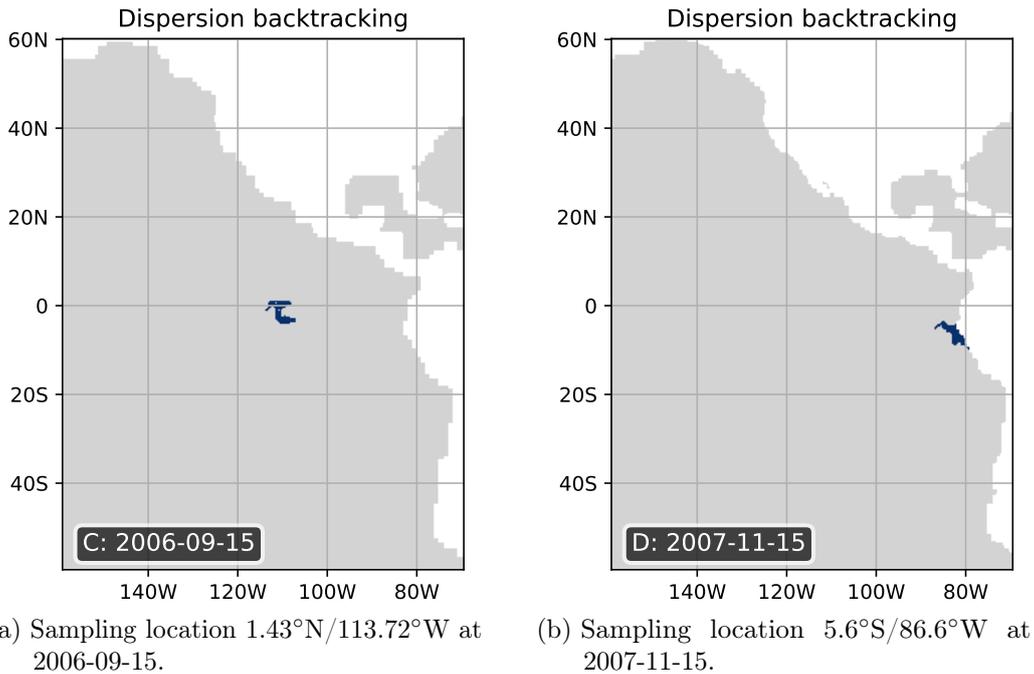


Figure 5.18.: Possible spawning areas of paralarvae samples identified by backtracking material transport flows over the 30 days prior to sampling.

(starting five days earlier or later than the sample date) results in different dispersion patterns and contribution areas, compare Figure 5.18a to Figure 5.19.

The area of origin of sample D is probably near the coast (Figure 5.18b), because it is likely that the Humboldt Current would transport eggs and paralarvae away from the coast. Calculations made using earlier and later sampling dates show similar patterns of material transport, and thus a similar area of origin to the one shown in Figure 5.18.

Recent data provides more information on current in the GOC. However, substituting missing data by recent data for backtracking is not feasible as even small changes in time (Figure 5.19) can lead to different results, so this procedure was not applied to backtrack from the historical sampling location A. Instead, an additional simulation run, similar to those described in Section 5.5, but with only the GOC initialized (filled with material) and using the higher resolution data from recent years, was used to simulate material transport out of the gulf. Some attractors emerged inside the gulf and only a small part of the material was exported (not shown). This indicates that the GOC is a self-contained ecosystem with restricted material exchange with surrounding areas.

5. Identifying spawning areas of *D. gigas*

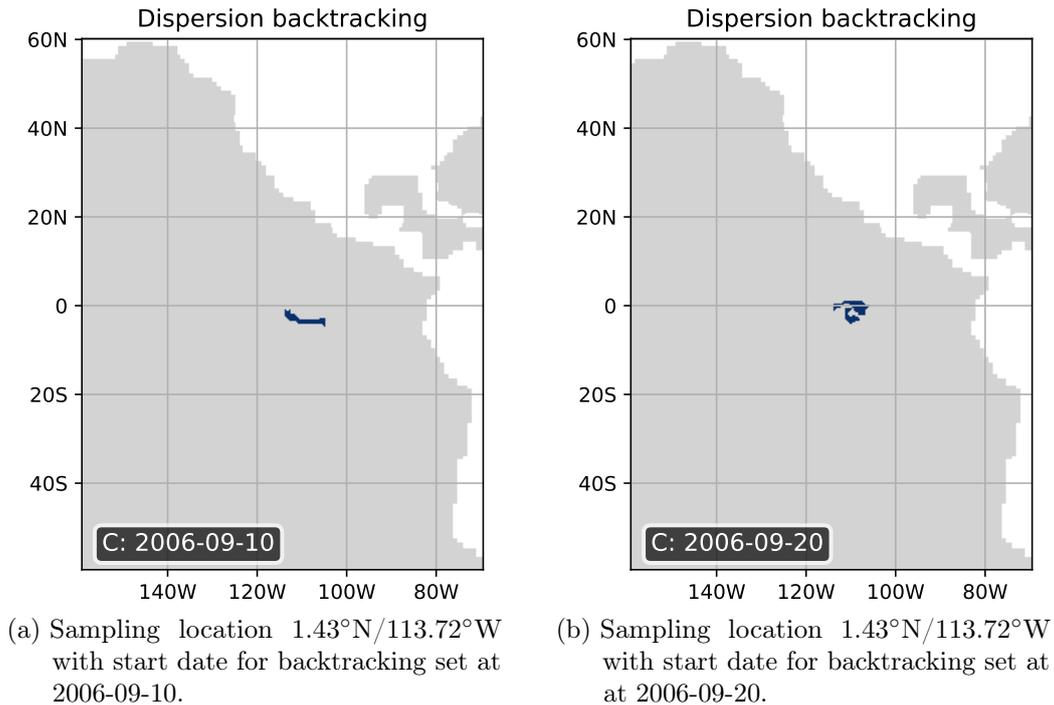


Figure 5.19.: Possible spawning areas of paralarvae sampled at 1.43°N/113.72°W identified by backtracking material transport flows over 30 days from 5 days prior to and after the actual sampling date. The changes in start date result in appreciable changes to the identified spawning location (compare with Figure 5.18a).

In summary, the results of backtracking known eggs/paralarvae location do not allow the reliable identification of spawning grounds, except in certain locations where material transport is low and/or in a constant direction. The 2006-09-03 (8.85°N/85.63°W) location near the coast (Sample B) is probably a spawning area due to the low material transport within this region. Simulation of flows in the GOC also shows restricted material exchange and suggests that the GOC was almost certainly the spawning location of eggs found in this area (Sample A).

The highly volatile dispersion results for the 2006-09-15 Sample C (Figure 5.19) indicate that the occurrence of paralarvae at this specific location is largely coincidental, due to the highly variable patterns of material transport. The hypothesized egg/paralarvae aggregation from distant donating regions would require a fine-meshed sampling to be of use for identification of possible spawning grounds, otherwise attractors could be erroneously identified as spawning areas. The biogeographic break found by genetic

markers analysis (Staaf et al., 2010b) is supported by the flow analysis showing currents forming a barrier near the equator.

5.7. Identifying spawning areas: Results and conclusions

The presumed co-occurrence of *D. gigas* and *Sthenoteuthis oualaniensis* renders the use of SD-complex data to model *D. gigas* spawning behavior plausible. The identification of SST as the main predictor variable of paralarvae presence probability (pp_{prob}) of the *Sthenoteuthis oualaniensis* and *Dosidicus gigas* complex (SD-complex) (Staaf et al., 2013) enabled computation of the SD-complex pp_{prob} inside the Pacific (60°N to 60°S and 159°W to 69°W) by applying a reversed engineering approach. The results revealed a generally higher pp_{prob} for the northern hemisphere than for the southern hemisphere.

The sample counts in Figure 5.2 are consistent with the computed pp_{prob} in Figure 5.9 to the extent that high counts are within regions of elevated pp_{prob} . However low sample counts also occur within elevated pp_{prob} regions, suggesting that a process of spatial aggregation leads to a patchy distribution of paralarvae presence; otherwise these low sample counts are inconsistent with the presumption that SST is the main predictor variable of pp_{prob} .

The analyses of surface currents support the hypothesized process of spatial aggregation, including within the sampling area of Staaf et al. (2013), whereby material, e.g., paralarvae, is transported from larger donator areas to mostly smaller attractor areas. Aggregation of paralarvae to an attractor area increases the number of paralarvae in this area, i.e., it increases the paralarva concentration, and therefore also the probability of a high sample count. This mechanism explains the highly variable sampling counts within a region with very similar SST and, therefore, predicted pp_{prob} . The donator and attractor areas are produced by currents that change over the time and are unspecific to any location. An enhanced version of the surface currents analysis uses a backtracking algorithm to determine the donating areas which may have been the source of sampled eggs or paralarvae. The available data indicate a relatively large amount of material transport over a 30 day period (Figure 5.18) and identify potential spawning areas. However, re-computation for one sample location with slightly different start dates (plus or minus 5 days, compared to the actual sampling date) results in appreciable shifts in the location of the spawning area (Figure 5.19). This result indicates that patterns of material transport are transient and highly variable.

In consequence, the postulated distinct SD-complex spawning grounds (Staaf et al., 2013) may be an artifact because the currents create small areas of increased paralarvae

5. Identifying spawning areas of *D. gigas*

concentration (attractors) and larger areas of decreased paralarvae concentration (donators). Coarse-grained sampling may fail to detect paralarvae concentrated in attractors and thereby misleadingly suggest the existence of distinct spawning areas. The data interpolation in this thesis, using the same sample data and including the zero sample counts, could not reproduce the results in Staaf et al. (2013) and is consistent with the above hypothesis.

The aggregation process supports the postulated relation between SST and pp_{prob} (Staaf et al., 2013), because the process provides an explanation for the low sample counts which would otherwise contradict the postulated relation. On the other hand, if spawning can take place successfully in areas with lower SST, i.e., within the temperature range of 15–25°C given by (Staaf et al., 2011), then the entire distribution area might be suitable for spawning. This is consistent with the observed relatively short-distance latitudinal migration (approx. 4°) by medium-sized and large specimens of *D. gigas* (R. I. Ruiz-Cooley et al., 2013). It would also be consistent with the hypothesis that *D. gigas* is likely to avoid long-distance migration to small distinct spawning grounds, since long-distance migration is a highly energetic process (R. I. Ruiz-Cooley et al., 2013) that would reduce available energy for reproduction. Furthermore, Rosa et al. (2013) link the recent northward range expansion to “the sustainability of foraging potential within an energetically acceptable distance to suitable spawning grounds”, which, if true, would make long-distance migrations unnecessary.

The hypothesis that *D. gigas* does not migrate or undertakes at most limited migrations (R. I. Ruiz-Cooley et al., 2013) is consistent with the observed existence of genetically similar (Staaf et al., 2010b) size-at-maturity (SAM) groups. In this case the existence of different SAM groups would be an expression of the phenotypic plasticity of *D. gigas* in response to changes in exogenous factors, such as temperature (Keyl et al., 2008; Markaida, 2006a; Staaf et al., 2010b), as postulated by the functional triad migration-maturation-growth (fTMMG). The genetic similarity of *D. gigas* specimens from different SAM groups and the relatively short-distance latitudinal migration undertaken by medium-sized and large specimens aligns with the presumption that individuals spawn close to their feeding grounds, where phenotypic and behavioral adaptation to prevailing local exogenous factors give rise to formation of the SAM groups.

6. Individual level traits and their computation

6.1. Introduction

In the energy driven life history model (EDLHM), the elements of the functional triad migration-maturation-growth (fTMMG) (Section 3.3) were reformulated into quantifiable elements (factors) at the individual level, but without undertaking a quantitative evaluation. This chapter provides an initiative quantitative evaluation of the model described in Chapter 4 with respect to observed data and the assumption of energy optimization at the population level that underpins the fTMMG.

Section 6.2 determines estimated values for the growth parameters used to set a standard (dimorphic) logistic growth function, denominated dimorphic terminal size growth function (DTSGF), which determines model parameters such as mass gain, energy requirements and energy buffering capability in relation to body mass. In the remainder of this chapter, the DTSGF is explored as an explanatory function, with respect both individual traits of *D. gigas* (Section 3.3) and properties at the population level (Section 3.4).

Section 6.3 applies the energy model in Section 4.4, to analyze individual energetic features of the DTSGF. This section evaluates the individual's resilience to energy deprivation using mathematical models based on estimation of the basal metabolic rate (basal_{mr}). Since oxygen-based metabolism is also restricted by the availability of oxygen, the mathematical model considers oxygen as a limiting factor determining energy needs. To explain the observed differences in size-at-maturity (SAM), Section 6.4 discusses differences in hatchling size and hatching time caused by a prolonged spawning period as possible factors influencing SAM, in addition to energy uptake.

The results of the analysis in this chapter indicates that daily growth rates are probably lower than reported in the literature and are highly correlated with energy uptake. The growth functions analyzed indicate different degrees of resilience to energy deprivation by males and females at different stages of the lifespan. Oxygen consumption can be a limiting factor with increasing mantle length (ML).

The chapter concludes that physiology-dependent growth, formulated as a standard logistic growth function, requires fewer preconditions than the alternative hypothesis of

6. Individual level traits and their computation

age-dependent growth in order to explain key features of the population-level growth and survival of *D. gigas*.

6.2. The unisex growth function and dimorphic terminal size growth function

As discussed in Subsection 4.3.2, development of *D. gigas* can assumed to follow a logistic growth function. In this thesis, the term dimorphic terminal size growth function (DTSGF) is used when discussing sex-specific growth functions, taking account of the dimorphic ML_{terminal} . These two functions for females and males are abbreviated $\text{females}_{\text{DTSGF}}$ and $\text{males}_{\text{DTSGF}}$ in tables and figures. The unisex growth function (UGF) defines a logistic growth function for both sexes with the same ML_{terminal} and is used for simplification and demonstration of general effects.

The logistic growth function is determined by setting values for four parameters: lifespan, hatching size, ML_{terminal} and daily growth rate, see Subsection 4.3.2. The initial daily growth rate of a paralarva decreases with age in accordance with the logistic growth function, so the initial daily growth rate is the highest growth rate during an individual's life. Relevant parameters reported in the literature (Table 6.1) are used to estimate the initial daily growth rate and thus the maximum growth rate. The following discussion uses the $[0, 1]$ standardization, where 0 equals birth and 1 equals death, introduced in Section 4.3.

property	value/range	references
lifespan	approx. 1–2.5yr	(Keyl et al., 2011)
average hatching size	1.1mm	(Staaaf et al., 2008; Yatsu et al., 1999)
terminal ML	1.2m	(Rosa et al., 2013)
daily growth rate	4–8%	(Gilly, Elliger, et al., 2006)

Table 6.1.: Values reported in the literature for lifespan, hatching size, ML_{terminal} and daily growth rate.

Figure 6.1 shows the relationship between initial daily growth rate and maximum ML, assuming the standard logistic growth function introduced in Subsection 4.3.2 with an inflection point at $t_{\text{ip}} = 0.5$, and a lifespan of one year. An initial 5% daily growth rate is sufficient to achieve a maximum ML of about 8m, which is not reported in the literature, while even an initial 4% daily growth rate achieves a maximum ML of 1.5m. Thus, initial daily growth rates $> 4\%$ are only compatible with observed ML if the

6.2. The unisex growth function and dimorphic terminal size growth function

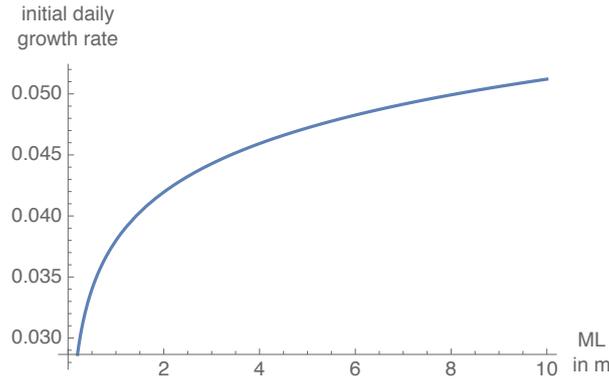


Figure 6.1.: Maximum ML plotted against initial daily growth rate assuming the logistic growth function in Subsection 4.3.2, a lifespan of one year and the range of initial daily growth rates reported in the literature as shown in Table 6.1. A growth rate of 0.04d^{-1} ($\approx 4\% \text{d}^{-1}$) would result in $\text{ML} \approx 1.5\text{ m}$, while a growth rate of $\approx 5\% \text{d}^{-1}$, would result in ML of approx. 8m which is not reported in the literature.

lifespan of individuals is significantly shorter than one-year, or if growth is regularly impaired during the lifespan, resulting in lower observed ML values compared to those calculated using the logistic growth function.

D. gigas's high phenotypic plasticity is an observable effect of local environmental conditions (Frawley et al., 2019; Portner et al., 2020; Sanchez et al., 2020; Yu, Chen, & Liu, 2021). Lin et al. (2017) report different variations in ML, lifespans and growth rates among individuals in different locations. For example, the lifespan in Ecuadorian waters is less than in Chilean waters while the ML growth rate and the sea surface temperature is higher in Ecuadorian waters.

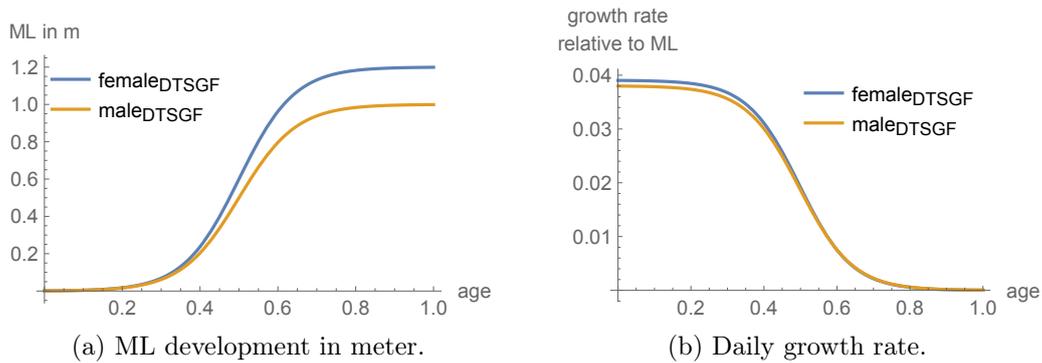


Figure 6.2.: The one year size development and daily growth rate for males and females given by the DTSGF, assuming a lifespan of 365d and with dimorphic parameters set to values shown in Table 6.2.

6. Individual level traits and their computation

The initial daily growth rate has to be set at approx. 3.8% (Table 6.2) in order for application of the logistic growth function (Section 4.3), which is based on analysis of available data (Tafur et al., 2010), to results in MLs consistent with observed values, see Figure 4.5.

D. gigas shows strongly dimorphic ML_{terminal} (Chen et al., 2011; Markaida et al., 2004; Tafur et al., 2010) and therefore sex-specific growth functions can be assumed to apply. In calculating sex-specific growth functions, the single record of 1.5m ML (Glaubrecht & Salcedo-Vargas, 2008) is considered to be an outlier and therefore excluded. Figure 6.2a shows the dimorphic growth function corresponding to observed values of ML_{terminal} .

These growth functions are based on an approx. 3.8% initial daily growth rate (Subsection 4.3.2), at the lower end of values reported in the literature (Table 6.1). However there are grounds for considering that the initial daily growth rate may be even lower. A $\approx 3.8\%$ initial daily growth rate would require a lifespan of only ≈ 0.53 (≈ 193 d) for specimens to attain approx. 0.6m ML, which would match observed MLs of female in the large SAM-group (Table 2.5), and would lead to high generation turnovers. Smaller individuals, e.g., ≈ 0.4 m ML, which matches the male large SAM-group (Table 2.5), would require approx. 3 weeks fewer, i.e., ≈ 0.47 (≈ 172 d) to reach their ML_{terminal} . These short lifespans and resulting high generation turnovers would suggest a different growth scheme to that postulated by the fTMMG, as the ecosystem may be incapable of supporting the energy needs of large specimens with such high generation turnover rates.

In summary, the initial daily growth rate must be limited to $<4\%$ if the logistic growth function is used, otherwise individuals would have a $ML_{\text{terminal}} > 1.2$ m after a one year lifespan. Smaller individuals, corresponding to observed MLs would require only approx. half a year for their terminal SAM, which implies a higher than observed generation turnover.

sex	lifespan	ML_{terminal}	hatching size f_0	initial daily growth rate
unisex	365d	1.0m	1.1mm	$\approx 3.8\%$
female _{DTSGF}	365d	1.2m	1.1mm	$\approx 3.9\%$
male _{DTSGF}	365d	1.0m	1.1mm	$\approx 3.8\%$

Table 6.2.: Unisex and dimorphic growth parameters used as settings for different logistic growth functions. A lifespan of 365d is assumed and, for dimorphic growth functions, a sex-specific ML_{terminal} is selected. The initial daily growth rate is the value required in order to achieve the corresponding ML_{terminal} .

6.3. The energy aspect of the EDLHM

This section focuses on the energy aspects of an individual's life history to identify critical time periods with increased energy needs and the resulting effects if these needs are not satisfied. The criteria are the mass dependent basal_{mr} , the locomotion metabolic rate ($\text{locomotion}_{\text{mr}}$) derived from basal_{mr} , and the growth function related growth metabolic rate ($\text{growth}_{\text{mr}}$) (Figure 4.3 in Subsection 4.2.2). In Section 4.3, size and mass development were computed using the equations in Section 4.4 to estimate the energy needs of basal_{mr} . These equations and relations enable identification of critical development stages in the lifespan of *D. gigas*, e.g., where high energy demands and limited buffering capability require the metabolization of body mass to meet energetic needs (Section 4.4).

Subsection 6.3.1 analyzes the energy needs of an individual during its lifespan to identify critical periods in which an individual is susceptible to energy exhaustion, Subsection 6.3.2 discusses growth potential under energy constraints and Subsection 6.3.3 discusses the oxygen consumption, the prerequisite to metabolism, as a potential limiting factor.

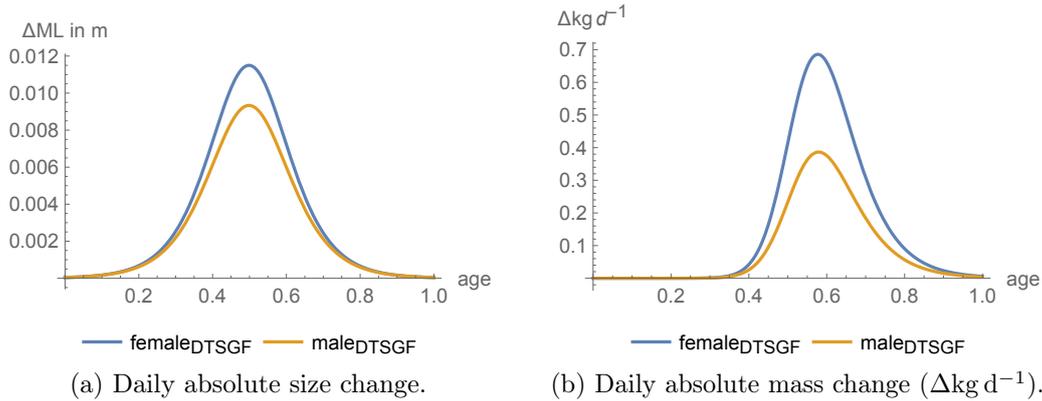


Figure 6.3.: Daily absolute size and mass change computed using the DTSGF shown in Figure 6.2. For computational parameters see Table 6.2.

6.3.1. Energy needs, energy buffering capacity and shifts during *D. gigas*' lifespan

The total metabolic rate (total_{mr}), the sum of basal metabolic rate (basal_{mr}), locomotion metabolic rate ($\text{locomotion}_{\text{mr}}$) and growth metabolic rate ($\text{growth}_{\text{mr}}$) (Figure 4.3), are used to define an individual's daily energy needs. The model uses the daily absolute mass change $\Delta\text{kg d}^{-1}$ (Figure 6.3) to estimate the daily energy requirements for growth

6. Individual level traits and their computation

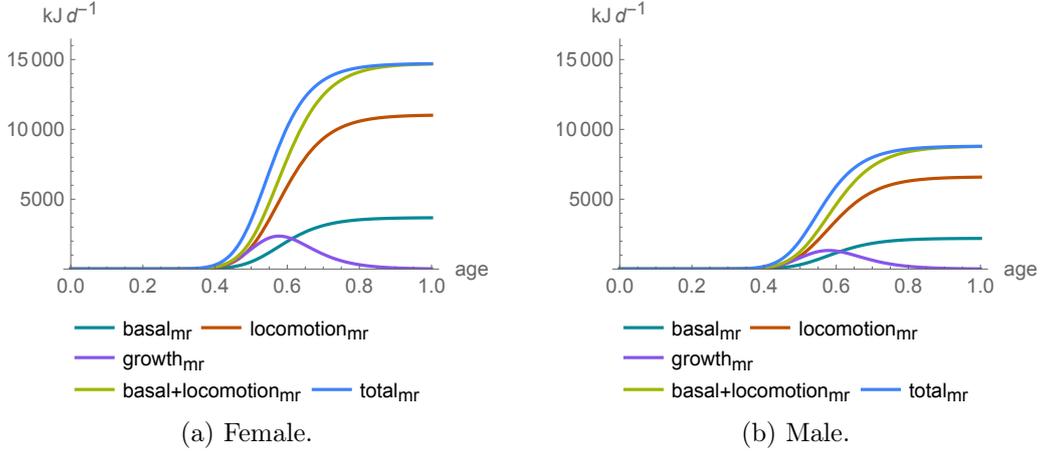


Figure 6.4.: Sex-specific daily energy needs calculated based on the DTSGF. For computational parameters see Table 6.2.

($\text{growth}_{\text{mr}}$) by the equivalence of mass and energy (Equation 4.48) and the resulting body mass defines the basal_{mr} (Section 4.4). For field metabolic rates, i.e., basal+locomotion metabolic rate ($\text{basal}+\text{locomotion}_{\text{mr}}$) in this thesis, Speakman (2000) gives an average 3.4 times basal_{mr} for mammals, and Karasov (1992) gives an average 4.1 times basal_{mr} during activity for mammals. In terrestrial vertebrates, 70% of daily energy expenditure is proportional to body mass and much of the remaining 30% is used for thermal regulation (Nagy, 2005), which is not considered in calculating energy expenditure by ectothermal species such as *D. gigas*. On this basis, Deutsch et al. (2015) assume a standard metabolic rate of 2 to 5 times basal_{mr} for marine species. Thus a factor of 3 times basal_{mr} can be considered as a conservative, lower estimate, especially when the ineffective jet propulsion of *D. gigas* is taken into account, which increases the energy requirement for locomotion, i.e., $\text{locomotion}_{\text{mr}}$, relative to basal_{mr} . Based on the above considerations, and estimated $\text{locomotion}_{\text{mr}}$ as $3 \cdot \text{basal}_{\text{mr}}$, total metabolic rate (total_{mr}) can be calculated as follows:

$$\text{growth}_{\text{mr}} = \Delta \text{kg d}^{-1} \quad (6.1)$$

$$\text{locomotion}_{\text{mr}} = 3 \cdot \text{basal}_{\text{mr}} \quad (6.2)$$

$$\text{basal}+\text{locomotion}_{\text{mr}} = \text{basal}_{\text{mr}} + \text{locomotion}_{\text{mr}} \quad (6.3)$$

$$\begin{aligned} \text{total}_{\text{mr}} &= \text{basal}_{\text{mr}} + \text{locomotion}_{\text{mr}} + \text{growth}_{\text{mr}} \\ &= 4 \cdot \text{basal}_{\text{mr}} + \text{growth}_{\text{mr}} \end{aligned} \quad (6.4)$$

6.3. The energy aspect of the EDLHM

The above equations allow estimation of the daily energy needs as shown in (Figure 6.4). In this model, the $\text{growth}_{\text{mr}}$ is the only metabolic process defined in the model independently of the basal_{mr} ; the energy requirement for growth decreases in later stages of the lifespan. Both sexes have a pronounced energy requirement, caused by increasing ML, which directly leads to a higher basal_{mr} . The energy requirements of other metabolic processes are proportional to basal_{mr} and increase throughout the lifetime of an individual, asymptotically approaching the limit defined by the basal_{mr} of $\text{ML} = \text{ML}_{\text{terminal}}$. The females have a more pronounced energy demand than males, and thus might be more likely to be affected by energy deprivation and prone to starvation.

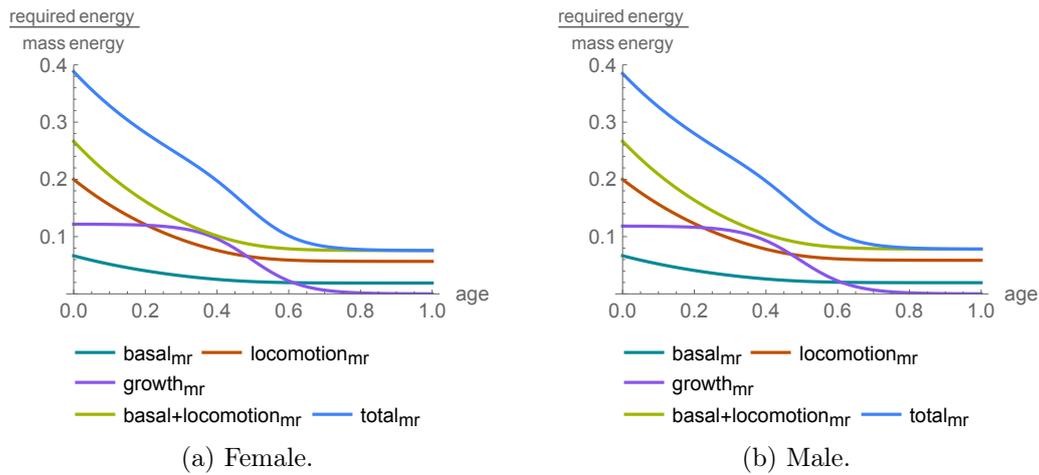


Figure 6.5.: Sex-specific daily energy needs in relation to body mass calculated for the DTSGF. For computational parameters see Table 6.2.

Figure 6.5 highlights the sex-specific energy pressure on an individual over the course of the lifespan and the overall high degree of susceptibility of *D. gigas* to food deprivation. In times of energy deprivation (i.e., when energy intake is less than energy needs), an individual must service the $\text{basal} + \text{locomotion}_{\text{mr}}$ energy needs by negative growth, i.e., by conversion of body mass to energy (mass energy in Figure 6.5). The *daily* body mass loss incurred in order to maintain $\text{basal} + \text{locomotion}_{\text{mr}}$ where energy uptake is zero is approx. 27% body mass at age = 0 and declines to approx. 9% at age ≈ 0.5 , which means an active, agile individual will lose approx. 9% of its body mass per day without taking into account conversion losses from body mass to energy. With increasing age and thus body mass, the buffering capability increases and body mass loss without energy intake asymptotically approaches $\approx 8\%$ body mass per day. Removing $\text{locomotion}_{\text{mr}}$ in the case of a lethargic individual (Trübenbach, Pegado, et al., 2013), reduces the energy requirement by 75% to the (still high) value of basal_{mr} .

6. Individual level traits and their computation

6.3.2. Growing under energy constraints

This section explores the role of energy deprivation, which may decouple the growth function from age and influence the ML_{terminal} of an individual, in more detail. Without energy deprivation, the standard (available) growth rate corresponding to the age-based growth function determines the ML_{terminal} reached by an individual's during its lifespan. However, growth also depends on energy uptake and this connection is incorporated in both the EDLHM and the fTMMG. Energy constraints can explain the different sizes of mature *D. gigas* individuals in the field, i.e., as being the result of impaired growth caused by food deprivation in food-scarce environments. The EDLHM handles these issues by assigning an order to claims on energy uptake for the fulfillment of energy needs: The basal_{mr} has a prior claim over $\text{locomotion}_{\text{mr}}$ in order to maintain basic body functionality, while $\text{locomotion}_{\text{mr}}$ has a prior claim over $\text{growth}_{\text{mr}}$ because locomotion is required for food uptake. The $\text{basal}+\text{locomotion}_{\text{mr}}$ therefore is the sum of these higher ranked claims.

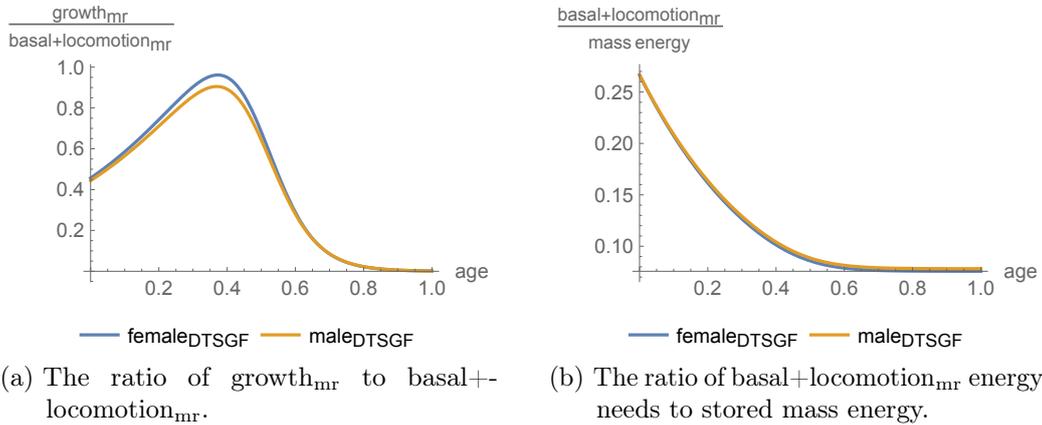


Figure 6.6.: Energy needs of males and females, specified by the growth function DTSGF over the course of the lifespan. The first diagram shows the ratio of $\text{growth}_{\text{mr}}$ to $\text{basal}+\text{locomotion}_{\text{mr}}$; the second relates $\text{basal}+\text{locomotion}_{\text{mr}}$ to energy stored as body mass and therefore provides an estimate of potential daily weight loss when energy uptake is insufficient to meet these needs. For computational parameters see Table 6.2.

The model assigns $\text{growth}_{\text{mr}}$ as the lowest ranked claim on energy uptake; the actual growth of an individual is thereby determined by the relationship between energy uptake and $\text{basal}+\text{locomotion}_{\text{mr}}$. Priority is given to the maintenance of $\text{basal}+\text{locomotion}_{\text{mr}}$, leaving an energy residue that can be used for growth. In periods of energy deprivation the energy used for growth will be less than the amount required to achieve the

available growth rate (i.e., as determined by the growth function). In the model, energy deprivation may slow the growth of an individual or even result in negative growth caused by the conversion of existing body mass into energy. The energy residue (positive or negative) is used to estimate an individual's effective mass gain (or loss) and thus its absolute mass. The mass is estimated from the ML using Equation (4.25), which also can be used to calculate the ML from a known mass, which in turn allows the effective (i.e., actual) growth rate to be determined for any given growth function.

Figure 6.6a tracks the ratio of $\text{growth}_{\text{mr}}$ to $\text{basal+locomotion}_{\text{mr}}$ over the lifespan, with the prioritized $\text{basal+locomotion}_{\text{mr}}$ greatly exceeding $\text{growth}_{\text{mr}}$ during most of the lifespan. Figure 6.6b shows the potential daily negative growth of body mass, i.e., the proportion of body mass that is lost when no energy uptake compensates for $\text{basal+locomotion}_{\text{mr}}$. Assuming that locomotion is maintained, energy deprivation results in the loss of a significant proportion of the body mass, especially in the earlier part of the lifespan. The actual loss of body mass e.g., the ratio of expected body mass based on the ML and the actual body mass of individuals sampled in the field, is not described in literature. However it is clear that, in the absence of energy uptake, individuals would be able to maintain their high $\text{basal+locomotion}_{\text{mr}}$ for only a short time before the onset of starvation. These observations highlight the high dependency of *D. gigas* on sufficient energy uptake for both growth and survival. One effect of this is that prioritization of $\text{basal+locomotion}_{\text{mr}}$ can decouple growth from age by impeding the growth rate in times of energy deprivation.

Figure 6.7 shows the effects of prioritized maintenance of $\text{basal+locomotion}_{\text{mr}}$ on relative body mass gain under different degrees of energy deprivation (Figure 6.7b to 6.7f), compared to unimpaired growth (Figure 6.7a). When energy uptake is less than total metabolic rate (total_{mr}), which includes $\text{growth}_{\text{mr}}$, daily relative body mass gain, i.e., $\frac{\text{body mass gain}}{\text{body mass}}$, Figure 6.7b to 6.7f, is reduced compared to maximum values in the growth functions. Since mass is assumed to be proportional to ML (see Equation (4.26) in Subsection 4.3.3) this also reduces the growth rate. Where a negative gain of relative body mass is shown, this indicates a loss of body mass due to sustain $\text{basal+locomotion}_{\text{mr}}$.

Any energy deprivation reduces the growth of body mass in early life stages and, in later life stages, can even reduce body mass through negative growth. Energy deprivation at any age reduces the body mass compared to values based on the growth function and this effect strengthens with increasing energy deprivation. Small impacts on the growth rate during the exponential growth phase in the first half of the lifespan

6. Individual level traits and their computation

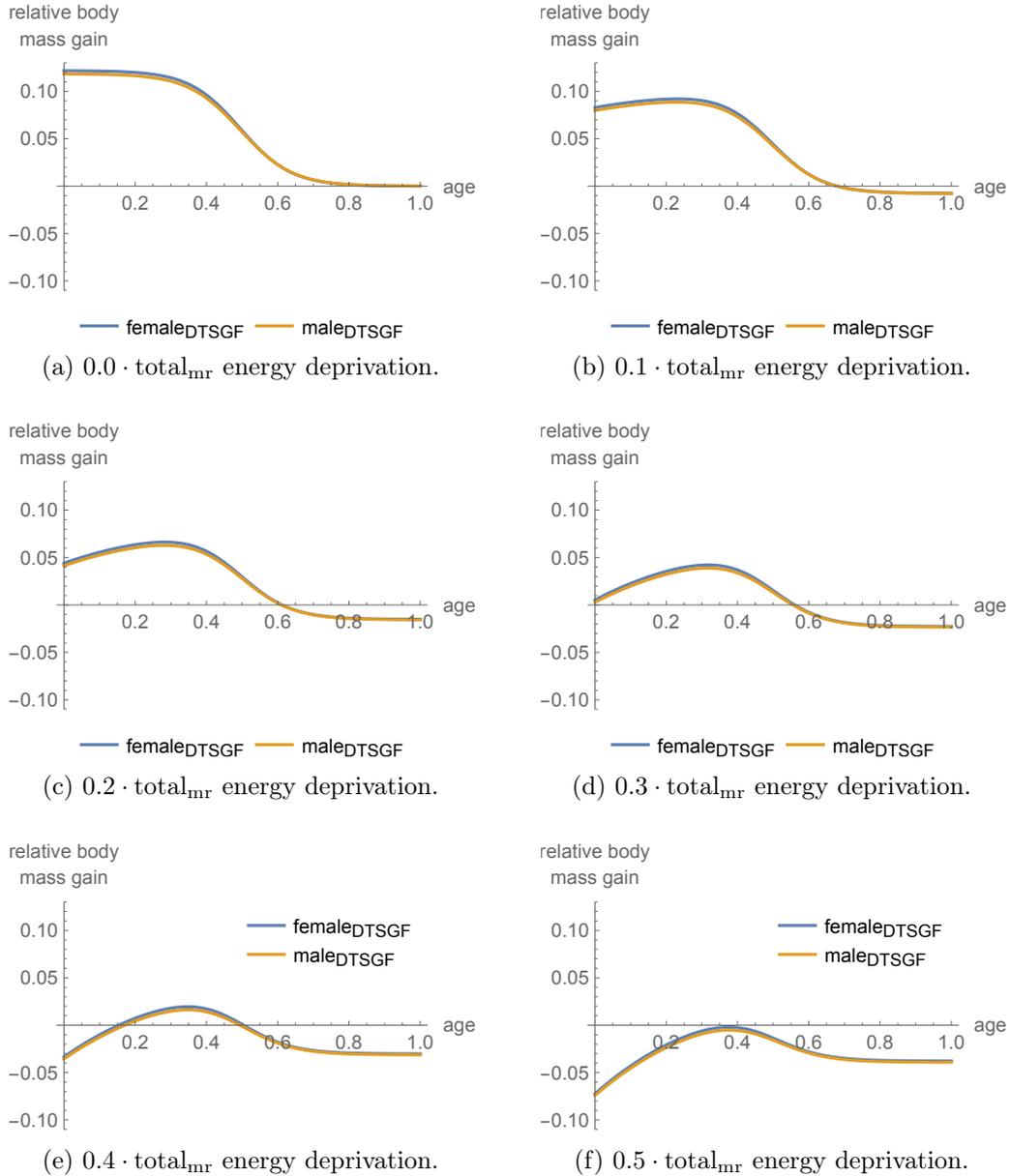


Figure 6.7.: The effective relative body mass gain, i.e., $\frac{\text{body mass gain}}{\text{body mass}}$ under different energy deprivation scenarios. The graphs show age-dependent body mass gain (or loss) by males and females for the growth function DTSGF, under the assumption that priority is given to $\text{basal} + \text{locomotion}_{\text{mr}}$ over total_{mr} . The value of “ $k \cdot \text{total}_{\text{mr}}$ energy deprivation” specifies the fraction of missing energy uptake compared to the energy required to maintain the total_{mr} including metabolism to achieve the maximum growth rate over the course of the lifespan in accordance with the growth function. For computational parameters see Table 6.2.

6.3. The energy aspect of the EDLHM

will be amplified over the remainder of the lifespan. Thus ML_{terminal} is likely to be highly variable due to the susceptibility of growth to energy deprivation.

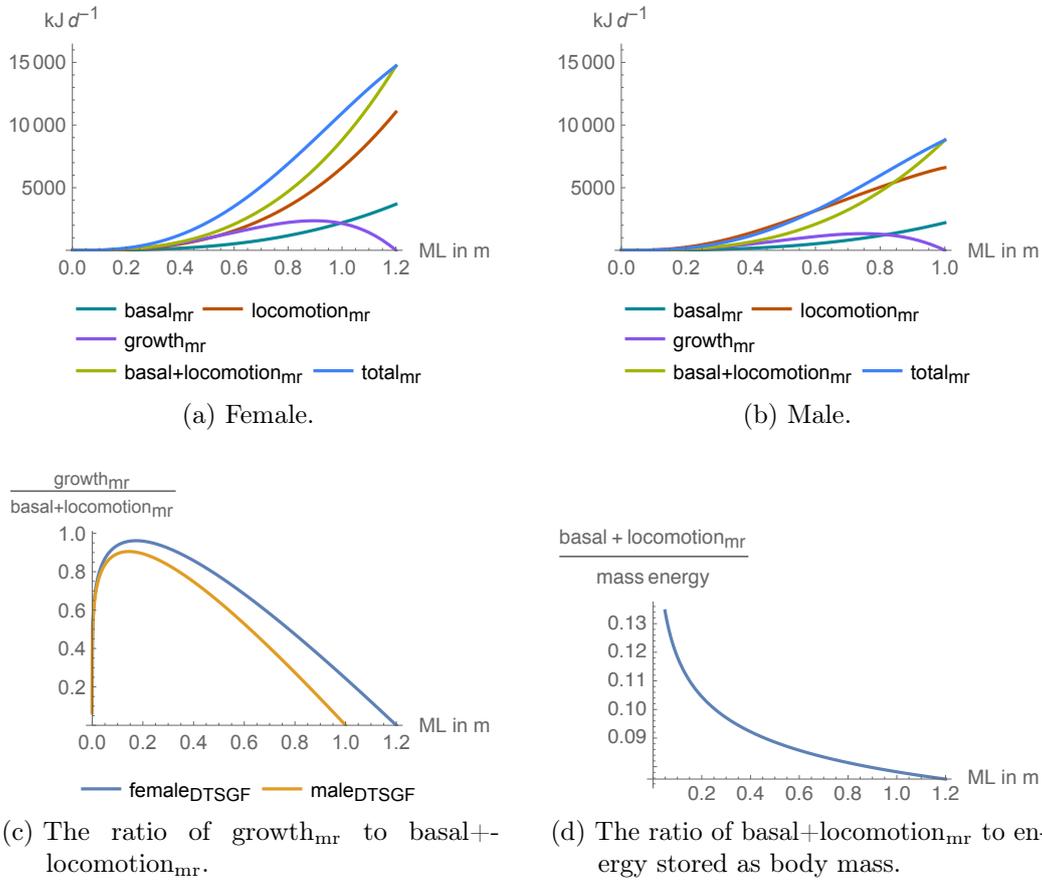


Figure 6.8.: Sex-specific energy parameters of the and DTSGF in relation to ML. The first two diagrams show daily energy needs. Diagram 6.8c shows the ratio of growth_{mr} to basal+locomotion_{mr} and indicates optimal ML at which a maximum proportion of energy is available for growth. Diagram 6.8d shows the ratio of basal+locomotion_{mr} to energy stored as body mass, i.e., the proportion of body mass for a given ML that would be required to be converted to energy to support basal+locomotion_{mr} under total energy deprivation. For computational parameters see Table 6.2.

If energy deprivation affects ML growth and the available growth rate depends on the current ML rather than age, then the current ML might be the most relevant parameter for the discussion of the relation of growth and energy. This hypothesis is supported by (Keyl et al., 2011), who found that older individuals do not necessarily have a larger ML, suggesting that ML is not dependent on age but determined by other, exogenous factors.

6. Individual level traits and their computation

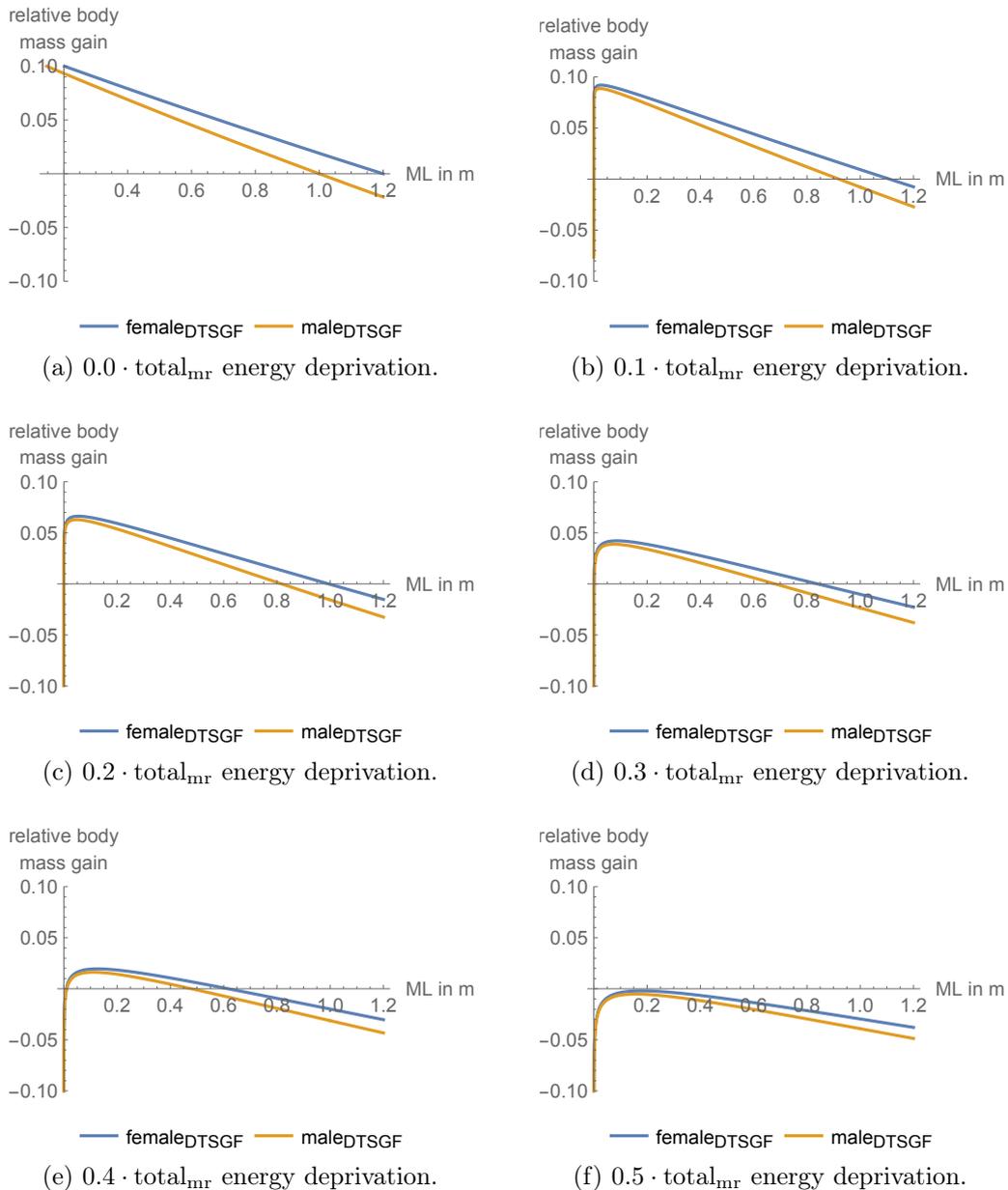


Figure 6.9.: Effects on body mass in relation to ML under different energy deprivation scenarios (Figure 6.9b to 6.9f) compared to under sufficient energy uptake (Figure 6.9a), see also Figure 6.7. The graphs show ML-dependent relative body mass gain (or loss) by males and females for the DTSGF, under the assumption that priority is given to $\text{basal} + \text{locomotion}_{\text{mr}}$ over total_{mr} .

The dominance of the $\text{basal} + \text{locomotion}_{\text{mr}}$ may impair growth under conditions of energy deprivation because growth requires relatively more energy at certain stages of

the individual lifespan, see Figure 6.6a. Over the course of the lifetime, the dominance of basal+locomotion_{mr} in relation to growth_{mr} decreases from ML = 0 m to a minimum at around ML = 0.2 m, i.e., the peak of the inverse ratio of growth_{mr} to basal+locomotion_{mr} in Figure 6.8c, and then steadily increases as ML increases.

If the ML of an individual is on the left side of the peak, then growth decreases the dominance of basal+locomotion_{mr}. Reduced dominance leaves more energy for growth even in times of moderate food deprivation, so growth is more probable and this effect increases the probability that growth will continue and reach this peak. Growth becomes less probable after this peak because basal+locomotion_{mr} leaves less energy for growth in times of (moderate) food deprivation.

Assuming low food availability in the ecosystem that impairs growth by leaving less energy for growth_{mr} due to prioritized basal+locomotion_{mr}, Figure 6.8c shows the relation between growth_{mr} and basal+locomotion_{mr} and indicates that optimal conditions for growth occur when ML is approx. 0.2m. It must be noted that the diagram should not be interpreted as providing information on probable values of ML_{terminal}, it simply indicates the increasing constraints on growth beyond this optimal ML, due to the increasing dominance of basal+locomotion_{mr} which leaves less energy available for growth_{mr} in times of food deprivation. The value of the optimal ML coincides with the ML_{terminal} for small SAM groups found in Tafur et al. (2010), see Table 2.5 p.27, and might support the validity of the EDLHM.

An increasing ML results in a smaller fraction of body mass being required for use in times of total energy deprivation to meet an individual's energy needs, see Figure 6.8d and Figure 6.9. But this effect is only pronounced over large differences in ML e.g., between 0.05m ML and 1.0m ML. By contrast, for example, a 0.2m ML individual requires conversion of approx. $0.105 \cdot$ body mass to cover its daily basal+locomotion_{mr} needs during total energy deprivation, compared with approx. $0.081 \cdot$ body mass in the case of a 0.8m ML individual. This difference might not have any relevance or observable effect in the field.

The dominant fraction of daily locomotion_{mr} in daily total_{mr} (Equation 6.4) might create a “loophole” for growth even under adverse food providing conditions. An individual could “save” a small part of locomotion_{mr}, i.e., by lethargic behavior, and invest this saved energy into growth_{mr}. This option would allow growth if the food uptake covered the basal_{mr}, the fraction of locomotion_{mr} required for feeding, while leaving some energy for growth_{mr}. In a mechanistic model, growth is therefore an optimization problem for an individual, i.e., balancing the trade-off inherent in reducing

6. Individual level traits and their computation

locomotion between increased availability of energy for growth and reduced energy uptake due to decreased feeding opportunities.

6.3.3. Oxygen as a potentially limiting factor

Aerobic metabolism, which is more powerful than an anaerobic metabolism, depends on oxygen consumption (Trübenbach, Pegado, et al., 2013). The oxygen consumption of *D. gigas* increases almost linearly with its body mass (Equation (4.41)), but the body mass (volume) is a cubic function of ML (Equation (4.25), Figure 6.10a), so the oxygen requirement increases cubically with ML.

Pauly (1981) discusses the disparity between the quadratic increase in gill surface area in fish as they grow and the cubic increase in body mass. For Pauly (1981), growth is the result of the two continuous processes: anabolism, which builds up body substances, and catabolism, which degrades them. Catabolism occurs in all living cells and is therefore proportional to body mass.

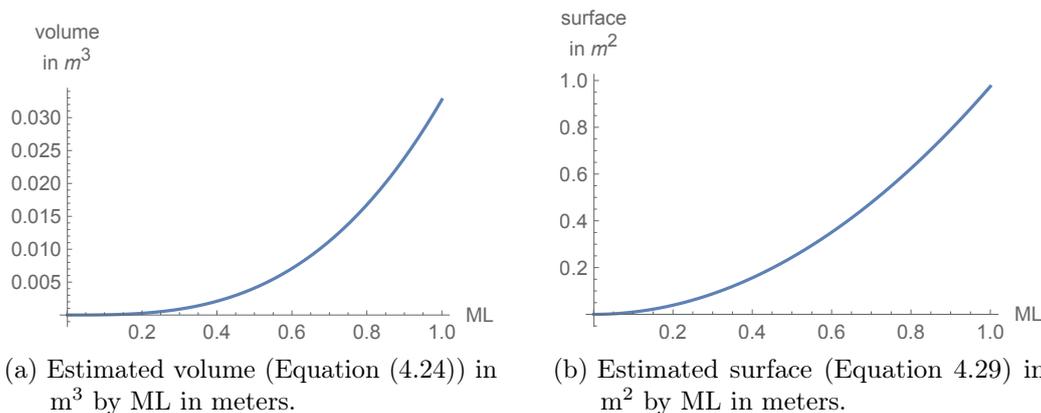


Figure 6.10.: Volume and surface area in relation to ML. Computational parameters are “unisex” values in Table 6.2.

By contrast, anabolism in fish is proportional to the respiratory rate, which is proportional to and, therefore, limited by a surface (i.e., the surface of the gills). Growth as the net result of anabolism and catabolism therefore becomes increasingly more difficult with increasing size, until the point when catabolism is equal to anabolism and growth ceases. Beyond this equilibrium is no further growth physiological possible, thus this equilibrium defines the maximum size of an individual.

Cephalopods breath by using a ventilatory system to pump water that is inside the mantle to force water over the gills. This ventilatory system is closely tied to the propulsion drive for locomotion, whereby mantle contraction forces the water inside the

mantle through the funnel to propel the individual in the opposite direction. This dual purpose system has low the oxygen extraction efficiency of 5–10%, compared 40–50% (Trübenbach, Teixeira, et al., 2013) in fish, so that “Squid are thought to live ‘chronically on the edge of oxygen limitation’.” (Trübenbach, Teixeira, et al., 2013). Rosa and Seibel (2010) estimate “larger squid consume nearly 10× as much oxygen as similar sized fishes [...]”. If oxygen is chronically limited, then the situation may become more adverse as a result of an increase in water temperature, since warmer water contains less dissolved oxygen. Oxygen limitation also affects large ML individuals more than smaller ones since their oxygen requirements are greater; only smaller individuals can thrive in lower oxygen saturated waters (Alegre et al., 2014).

In either case, *D. gigas* has to increase the volume of water passing over the gills in order to compensate for the higher oxygen demand and/or lower oxygen availability. Skin respiration could provide an additional source of oxygen (Trübenbach, Teixeira, et al., 2013), but the ratio of the quadratic growth skin surface area to the cubic growth of body mass (and therefore energy demand) in Figure 6.11 means that the amount of additional oxygen supplied by this means is inversely proportional to ML, i.e., doubling the ML halves the ratio of surface area to volume. Therefore skin respiration cannot compensate the oxygen limitation with increasing ML.

In summary, oxygen could be a critical factor limiting the maximum size of *D. gigas*, because the locomotion required for food uptake decreases the efficiency of oxygen extraction, while increasing locomotion_{mr}, and passive respiration is unlikely to extract sufficient additional oxygen to match the requirements of increasing ML. Low oxygen saturation levels are likely to further restrict growth as an individual has no recourse to mechanisms to compensate for low oxygen extraction efficiency of its ventilation system.

6.3.4. The effects of energy constraints support the fTMMG

Unconstrained growth is only possible with sufficient energy and oxygen uptake. The growth function (Subsection 4.2.2) in the model is assumed to describe the maximum available growth rate, that results for the set parameters in a ML_{terminal} of 1.2m. This maximum available growth rate and thus a ML_{terminal} of 1.2m is only achievable with sufficient energy supply, i.e., sufficient food uptake and oxygen uptake through respiration to metabolize the food.

The exact period of time (i.e., the lifespan) required to reach the ML_{terminal} is undetermined and probably not solely dependent on age, otherwise the small SAM-group should not exist. For this reason it would make sense to assume that the available

6. Individual level traits and their computation

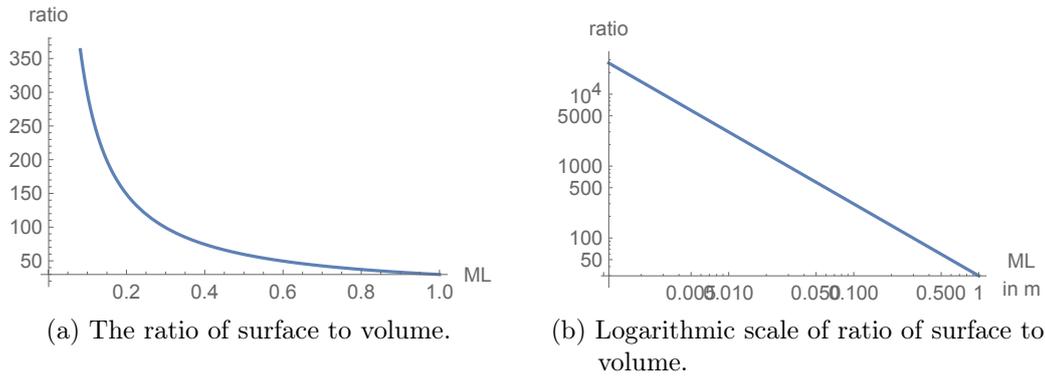


Figure 6.11.: Volume and surface ratio in relation to ML. The reduced increase of the surface may limit oxygen consumption hence also limit metabolism. Computational parameters are “unisex” values in Table 6.2.

growth rate is controlled by a size-dependent physiological growth capability, i.e., that growth capability at a given point in time largely depends on the current ML instead of the age (see Subsection 4.2.2).

If physiological processes determining the maximum growth rate are dependent on the ML, i.e., a given ML corresponds to maximum ΔML within a certain period e.g., one day, then this maximum ΔML can only be achieved if the energy supply is sufficient.

In case of energy deprivation, the prioritization of $\text{basal+locomotion}_{mr}$ impedes or suppresses growth. When ΔML is small (or zero), the ML changes little (or remains the same). This means that the ML-dependent available maximum growth rate, i.e., the maximum achievable ΔML , also changes little (or not at all). Thus the effect of reduced growth is to prolong the existing growth rate capability. In the case of a logistic growth function this will delay entry into the phase of reduced growth as ML approaches its maximum. If energy deprivation occurs often enough, then the ML remains small and the individual still has an high growth rate capability, so the individual remains in an exponential growth phase and it appears that the exponential growth phase is prolonged until maturation.

Alternatively, the ability to grow may be limited primarily by the age, whereby the current age defines the current available growth rate in accordance with a determined growth function. Limiting growth at any time for any period lowers the available ML_{terminal} . The preconditions for such a mechanism are a predetermined growth function and a predetermined lifespan. The predetermined lifespan is important because the physiological processes “need to know” where on which growth function they are currently located.

6.4. Different preconditions as an explanatory factor to different size-at-maturity

Delayed growth caused by energy deprivation is equivalent to a hatching offset ω in paralarvae, see Subsection 6.4.3. Every Δt period in which an individual does not grow can be interpreted as $\omega + \Delta t \rightarrow \omega$, i.e., as the expression of a paralarva that hatched at a later date and grew under optimal conditions. Therefore, a reduced growth rate is transformed into a “virtual” hatching offset ω , the effects of which are discussed in Subsection 6.4.3.

6.4. Different preconditions as an explanatory factor to different size-at-maturity

6.4.1. Overview

This section discusses factors excluded from both the fTMMG and the EDLHM, since these consider the effects of active individuals but not the previous stages, such as spawning and embryonic development. The effects of exogenous factors and/or small differences in early life stages are discussed as alternative or contributory factors to explain the variations in SAM in adults.

The life stage prior to hatching is distinctive in the sense that the embryo feeds on its yolk, has no requirement for an external energy supply and is exposed to the local environmental conditions due to lack of active movement. Hatching marks the transition from passive environmental exposure to an active phase with (initially limited) environmental interaction. The probable existence of multiple spawning locations within the distribution range (Chapter 5) exposes individuals in the (horizontally) immobile early developmental stages to location-specific environmental conditions.

If exogenous factors are relevant for ontogenesis, differences in ontogenesis (e.g., SAM), it may be the results of exposure of early life stages to local environmental factors, as well as or instead of being an expression of genetic differences. Subsection 6.4.2 discusses how exponential growth may amplify differences in hatching sizes until the adult stage, and Subsection 6.4.3 evaluates the extent to which the size differences can be explained by hatching offsets.

6.4.2. Size differences in early ontogenetic stages may amplify exponentially

The exponential growth phase in the early life stages can amplify initial, small differences, such as the hatching size f_0 , to significant ones in later stages, i.e., ML_{terminal} . Different hatching sizes may result from different egg sizes. A relation between egg size and female size is reported for other cephalopod species (Pecl & Jackson, 2008) and may also apply to *D. gigas* (Birk et al., 2016). Smaller eggs require less energy for their

6. Individual level traits and their computation

production and allow an increased number of eggs to be produced by females with smaller reproductive organs.

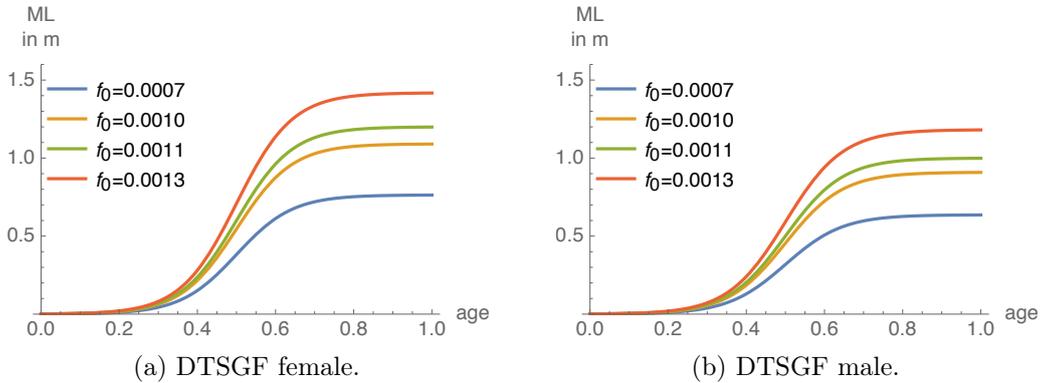


Figure 6.12.: The effect of hatching size ($f_0 = \{0.0007, 0.0010, 0.0011, 0.0013\}m$) on ML_{terminal} , for the DTSGF with the same growth rates. For computational parameters see Table 6.2.

The cubic ratio of volume to length means that small length (ML) differences correspond to large differences in volumes. A smaller egg volume contains significantly less yolk mass and supports only smaller embryos, so that *earlier hatching of smaller sized hatchlings* is mandatory. Higher temperatures during embryonic development can also result in earlier hatching and smaller hatchlings (Pecl & Jackson, 2008).

The occurrence of small SAM groups could then be explained as the result of a self-contained process: Smaller females as members of the small SAM-group (C. Nigmatullin et al., 2001) occur in the tropical zone, near the equator. Small females produce small eggs, small eggs can produce only small hatchlings which can only grow to small females.

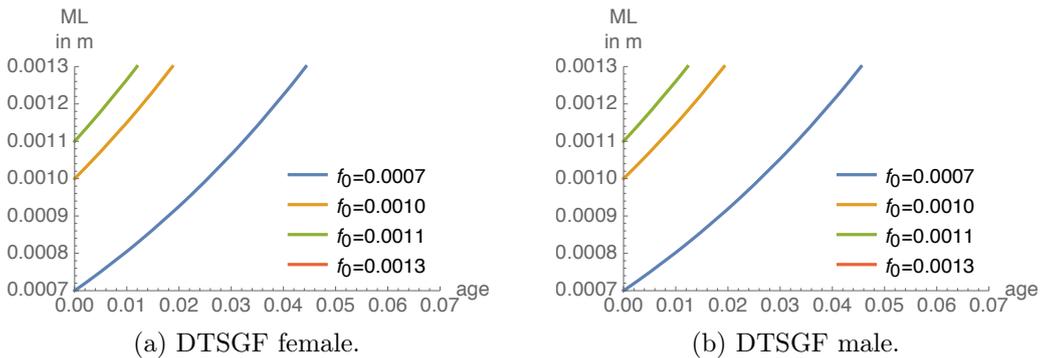


Figure 6.13.: The time needed for hatchlings of $f_0 = \{0.0007, 0.0010, 0.0011\}m$ to grow to $0.0013m$. For computational parameters see Table 6.2.

6.4. Different preconditions as an explanatory factor to different size-at-maturity

So if the factor temperature controls the hatchling size, then temperature also controls the size of SAM-groups.

The likelihood of such a self-contained process can be evaluated by applying the aged-based growth function (Subsection 4.3.2), since this function describes the optimal ML growth for individuals of a hatching size f_0 in accordance with the underlying maximum available growth rate (Subsection 6.3.2).

During early ontogenesis, *D. gigas* shows exponential growth that levels off at a later age (Figure 4.5b, p.60). Figure 6.12 illustrates how the growth of paralarvae of different initial sizes (Camarillo-Coop et al., 2010; Staaf et al., 2008; Yatsu et al., 1999) leads to pronounced differences in adult size through the amplifying effect of the exponential part of a logistic growth function. For example, applying an aged-based growth function, the continuous achievement of the maximum available growth rate by females would result in female hatchlings of $f_0 = 0.0013\text{m}$ attaining a $\text{ML}_{\text{terminal}}$ of approx. 1.5m. Camarillo-Coop et al. (2010) describe higher hatchling size deviations so a theoretically larger $\text{ML}_{\text{terminal}}$ might be possible according to the model.

Different hatchling sizes may explain a different $\text{ML}_{\text{terminal}}$ on basis of an age-based growth rate, but a difference in hatching size is technically equivalent to a time offset ω whereby some paralarvae hatch earlier than others. Assuming a lack of energy constraints, earlier hatching individuals growing at maximum available growth rate will eventually attain the same adult size as larger sized hatchlings. Figure 6.13 shows the time (equivalent to ω) paralarvae need to grow at maximum available growth rate to reach the hatching size of the largest $f_0 = 0.0013\text{m}$ hatchlings. A relative age of 0.01 is for a one year lifespan equivalent to 3.65d, and the ω is ≈ 0.044 (Figure 6.13a), giving an effective lead for growth of approx. 0.16d, equal to approx. 4 hours.

In conclusion, different hatching sizes are expected to level out under optimal energy condition within a day, making a self-contained process of SAM differentiation unlikely. Furthermore, the factor energy may be more important than the direct influence of the factor temperature during the whole of ontogenesis, in which case energy constraints will be the most significant cause of size differences, as described in Subsection 6.3.2.

6.4.3. Multiple spawning batches

It is believed that *D. gigas* spawns in multiple (10 to 14) batches over an unknown time period (Hernández-Muñoz et al., 2015; Ibáñez et al., 2011; R. I. Ruiz-Cooley et al., 2013; Staaf et al., 2008; Tafur et al., 2010). Distributing the spawning over the time may support the survival of the paralarvae because fast-growing paralarvae occupy different ecological niches as they grow.

6. Individual level traits and their computation

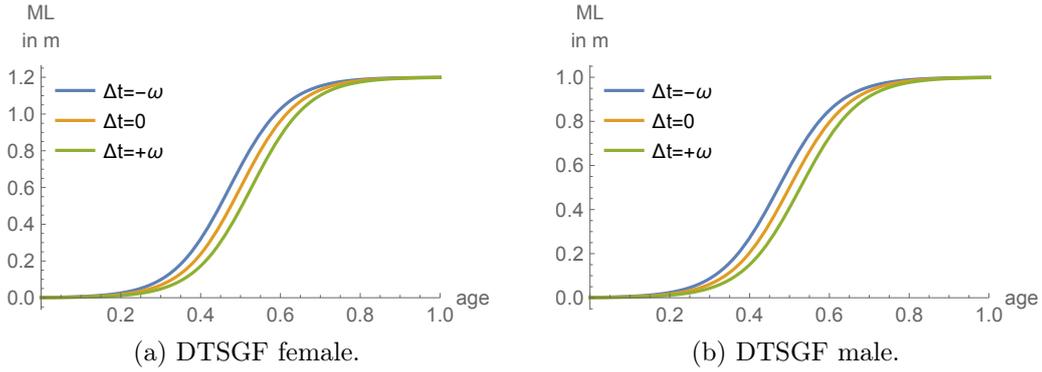


Figure 6.14.: The DTSGF ML development when spawning takes place over 20 days ($\omega = 10$ d). For computation parameters see Table 6.2.

Multiple spawning most likely requires formation of multiple of egg masses which takes time and increases the energy uptake needs, adding to already high basal+locomotion_{mr} energy needs (Tafur et al., 2010). This high energy demand may limit the growth rate of the egg masses inside the female and set a minimum time interval between each spawning occasion as well as the timespan for the entire spawning period.

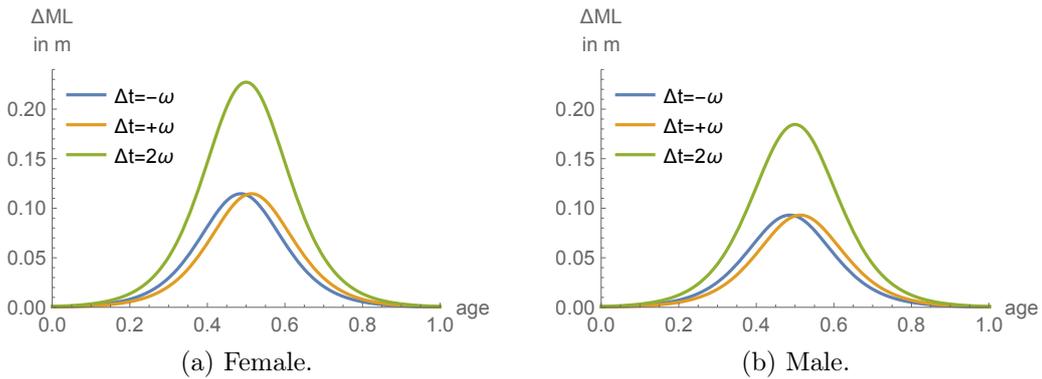


Figure 6.15.: DTSGF ML differences development when spawning takes place over 20 days ($\omega = 10$ d). For computational parameters see Table 6.2.

Figure 6.14 and Figure 6.15 show differences in ML for the theoretical scenario of a three-stage spawning period over 20 days where $\omega = \{-10, 0, +10\}$ days, with a negative offset representing earlier hatching. Figure 6.14 shows the size development for a first and last batch and Figure 6.15 shows the size differences, which are largest during the exponential growth phase, and occur at the same point in the lifespan, but are more marked in females because of their larger ML_{terminal} . The resulting differences in size (Figure 6.15) do not cover, even at the point of the highest growth rate at

which size differences are amplified, the range of sizes for the three group classification, see Table 2.5 p.27. Thus the hatching offset as a single factor does not provide an explanation for the occurrence of SAM groups.

6.5. Summary

The calculations presented in this chapter are based on a simplified energy balance model and some assumptions. Field observations, that could be used to validate the model, are rare. The results presented should therefore be interpreted as providing qualitative information on general effects and tendencies, rather than estimations of quantitative values. Accurate predictions of individual development under different environmental conditions are outside the scope of the model. However, the model allows investigation of the general effects and the order in which they occur.

According to the results, under the DTSGF females grow faster than males, and thus have higher energy requirements for growth_{mr}. However, as a consequence, they also have higher energy requirements for basal+locomotion_{mr}, which is prioritized over growth. This can hinder growth under scarce food conditions, resulting in a ML at maturity, i.e., SAM, smaller than that predicted by the DTSGF under conditions of unconstrained growth. The alternative hypothesis, that age largely determines the growth process, rather than physiological parameters, requires significantly more assumptions, including a predetermined ML_{terminal} and lifespan of an individual.

In summary, physiology-dependent growth, favored in this thesis, requires fewer preconditions than age-dependent growth in order to explain population-level growth (e.g., differences in SAM) and survival (e.g., through cannibalism). The DTSGF, which models both male and female growth based on a standard logistic growth function, is compatible with observed features of physiology-dependent growth, as demonstrated by the computational results presented in this chapter.

7. Cannibalism as a possible survival strategy

7.1. Overview

Cephalopods live “life in the fast lane” with fast growth rates and rapid rates of turnover at the population level (Pecl & Jackson, 2008). This lifestyle entails a high metabolism rate, high energy needs and consequently high food uptake needs. The short lifespan, little or no overlap between generations, and early semelparous breeding (Boyle & v. Boletzky, 1996), mean that the survival of the population depends on the reproductive success of the current generation (Pecl & Jackson, 2008).

In spite of this lifestyle (Alegre et al., 2014; Keyl et al., 2008; B. A. Seibel, 2007), *D. gigas* lacks an efficient energy buffering capability (Ibáñez & Keyl, 2010). This physiological deficiency might restrict the survival capability of populations under adverse conditions, e.g., in times of scarce food, but may be compensated for by an effective behavioral survival strategy. Cannibalism has been suggested as one such strategy for cephalopods (Ibáñez & Keyl, 2010; Keyl et al., 2008). Cannibalism widens the range of exploitable resources; it supports energy uptake and consequently population survival. Smith and Reay (1991) define cannibalism as “[...] the act of killing and consuming the whole, or major part, of an individual belonging to the same species, irrespective of its stage of development”.

D. gigas may strategically utilize its school or cohort as a population energy buffer that provides a high quality and easily accessible food resource even during migration (Alegre et al., 2014; Ibáñez & Keyl, 2010; Keyl et al., 2008). Compared to other cephalopods, *D. gigas* exhibits an extraordinarily high cannibalism rate, as found by the percentage of material from conspecifics in stomach analyses (Ehrhardt, 1991; Ibáñez et al., 2008; Markaida & Sosa-Nishizaki, 2003). Some authors (Ehrhardt, 1991; Ibáñez et al., 2008; Markaida & Sosa-Nishizaki, 2003) consider this high cannibalism rate as a population survival strategy, while other authors (Bruno et al., 2021; Field et al., 2013; Ibáñez et al., 2008; Ibáñez & Keyl, 2010; Ibarra-García et al., 2014; Keyl et al., 2008; Markaida et al., 2008; Portner et al., 2020) consider these high cannibalism rates to be artificial, i.e., as induced by fishing stress. In contrast, Alegre et al. (2014) hypothesize an underestimated cannibalism rate.

7. *Cannibalism as a possible survival strategy*

Cannibalism is considered size-related, as increasing size (mantle length (ML)) enhances the capability of an individual to cannibalize smaller individuals (Ibáñez & Keyl, 2010) and increases the cannibalism rate (Bruno et al., 2021; Ibarra-García et al., 2014; Liu et al., 2020). Bruno et al. (2021) found that while different factors influenced the cannibalism rate, ML was the best predictor. The window of cannibalism (WOC) defines the relative size of the prey item and is assumed to increase linearly with increasing size of the cannibal (Ibáñez & Keyl, 2010). Ibarra-García et al. (2014) found exactly this correlation in their studies, with females having a greater percentage of material from conspecifics in their stomach contents than males; this would be consistent with the average larger ML of females compared males of the same age.

Cannibalism is equivalent to the presence of an additional predator and therefore also has negative effects on the population. The benefits, in terms of increased survival of individuals in time of food must therefore outweigh the costs, in terms of additional deaths of individuals due to predation. This chapter analyzes the theoretical aspects of cannibalism. It assesses the contribution of cannibalism to the survival of individuals and populations, taking account of the growth functions analyzed in previous chapters and their effects on size relations among individuals within a population and, consequently, on the WOC. Furthermore, predicted effects of cannibalism on sex ratios are compared to observed sex ratios in the field, which have been interpreted as providing evidence for a population level survival strategy favoring the survival of females at the expense of males (Ibáñez & Keyl, 2010).

In this chapter, an individual-based modeling (IBM) based simulation model is used to test these theoretical aspects, with the aim of identifying viable parameter intervals for cannibalism to occur in combination with different growth functions and under a range of starting conditions.

To achieve these goals, Section 7.2 first discusses theoretical aspects of cannibalism and defines constraints on the parameter values used in the simulation runs. These constraints, which define the prerequisites of a cannibalism simulation model, are explored in relation to size relations in Section 7.3 and energy uptake in Section 7.4, and the results are summarized and discussed in Section 7.5. In Section 7.6, the theoretical aspects explored in the previous section are tested in simulation runs, whose results inform an evaluation of three alternative cannibalism strategies with respect to their impacts on energy uptake and energy buffering.

The results of these simulations contradict assumptions that cannibalism is a distinct feeding strategy to ensure population survival under adverse environmental conditions. The changing sex ratio observed in the field under low environmental energy conditions,

considered as a strategy to optimize energy use at the population level (Tafur et al., 2010), is tested in the context of cannibalism in Section 7.7. As an alternative, a “new” window of predation (WOP) based feeding strategy is proposed in Section 7.8, in which cannibalism is considered merely a “by-product” of generalized opportunistic predation. Section 7.9 summarizes the results.

In the following discussions, the term “size” is used to mean the size of an individual, i.e., the ML, and size of a school, i.e., the number of members. Where the meaning is not clear from the context, the precise definition is given. This avoids the re-definitions of terms used in previous chapters, e.g., size-at-maturity (SAM) to mantle-length-at-maturity. The term “energy exhaustion” describes a situation where an individual is starved of energy and presumably lethargic, but still alive. The term “terminal energy exhaustion” describes a situation where the individual has died of starvation.

7.2. Theoretical aspects of cannibalism

7.2.1. Energy uptake as the driving force for cannibalism performing

The discussion of energy needs in Subsection 6.3.1 shows *D. gigas*' proneness to metabolically driven energy exhaustion. To avoid this occurring, *D. gigas* requires a continuous energy supply and/or an energy buffering capacity. Ibáñez and Keyl (2010) suggestion that *D. gigas* has limited energy buffering capacity by means of lipid metabolism is still in discussion (Ibáñez & Keyl, 2010), but this may still be highly dependent on energy uptake.

Individuals compete for food resources within and among species. Cannibalistic competition, in which individuals of a species are also a food resource for that species, widens the range of exploitable food resources (Ibáñez & Keyl, 2010; Polis, 1981). Cannibalism is associated with high growth rates because conspecifics are a high-quality food source that provides vitamins, minerals and amino acids in “optimum proportions for maximal growth” (Smith & Reay, 1991). Thus the “main proximate advantage conferred by cannibalism is nutritional” (Smith & Reay, 1991). In times of food scarcity, cannibalism enables access to high-quality food, whereby “[...] no energy is lost due to conversion of ingested proteins to storable substances (fat, lipids) which is known to lead to losses of over 30% of energy at least in case of ureotelic animals” (Ibáñez & Keyl, 2010). In case of *D. gigas* and cephalopods in general, it is suggested that the outsourcing of an individual energy storage and buffering capability to the population and school level provides a means of meeting the requirement for continuous energy supply (Ibáñez & Keyl, 2010).

7. Cannibalism as a possible survival strategy

According to Ibáñez and Keyl (2010), Smith and Reay (1991) was the first author writing in English to describe cannibalism among teleosts. Cannibalism was documented in 36 of 410 teleost species surveyed by Smith and Reay (1991) and is presumed to occur in many other species. Cannibalism therefore might be more widespread and part of a regular diet, not limited to teleost species and, in summary, a common predatory behavior (Polis, 1981; Smith & Reay, 1991).

7.2.2. Classification criteria in respect of *D. gigas*

For teleost fish species, Smith and Reay (1991) propose a classification of cannibalism using three criteria that they consider to be generally applicable. The first criterion relates to the developmental of the prey and distinguishes between egg and post-hatching cannibalism. *D. gigas* engages in post-hatching cannibalism and there is no observational evidence for egg cannibalism.

The second classification criterion relates to the genetic relationship between cannibal and prey item, distinguishing between filial, sibling and non-kin cannibalism. In case of *D. gigas*, the genetic relation between cannibal and prey is unknown; however filial cannibalism can be ruled out because *D. gigas* is a semelparous species.

The third criterion relates to the relative age of cannibal and prey item and distinguishes between intra-cohort and inter-cohort cannibalism. For most fish species, the potential for cannibalism is “frequently determined by predator gape size” relative to the size of prey (Smith & Reay, 1991). For many species, the required size difference is found between individuals of different age cohorts. For semelparous species such as *D. gigas* the potential for cannibalism is determined “purely” by differences in size, which may be determined by other factors as well as age (Subsection 4.2.2, Section 3.2, Keyl et al., 2011). However, *D. gigas*’s ability to tear prey item into pieces decreases the need of a large size disparity.

7.2.3. *D. gigas*’ behavior in the context of cannibalistic effects

The shoaling of *D. gigas* individuals encourages cannibalism because cannibalism “tends to increase with increasing density and decreasing food availability” and “high density of either cannibal or prey item promotes cannibalism through increased encounter rate between the two” (Smith & Reay, 1991). For fish species, Smith and Reay (1991) report a widespread negative relationship between cannibalism and food intake, as cannibalism is likely triggered by food deprivation.

Cannibalism supports the immediate survival of larger individuals at the expense of smaller (younger) ones (Smith & Reay, 1991; van den Bosch et al., 1988). However in the case of intra-cohort cannibalism, it also reduces the number of individuals in the cohort, i.e., cohort size. Thus, it remains to be demonstrated whether cannibalism is an effective survival strategy at the population level, e.g., by increasing energy buffering capability. Possible benefits at the population level include the following:

- *Decreased intra-cohort competition for limited resources* by reducing the number of competitors (Polis, 1981; Smith & Reay, 1991). The immediate benefit for the individual is the energy uptake, and the long-term benefit at the population or school level is reduced intra-cohort competition for limited resources.
- *Increased per capita consumption* with resulting increased growth rates due to sufficient energy uptake (Subsection 6.3.2); larger individuals take advantage in inter-species and intra-cohort competition to maintain rapid growth rates.
- The *survival of part of a school* under conditions of food scarcity (Polis, 1981). The remaining (smaller) school still forms an intact organizational structure with all the concomitant benefits.
- *Increased size-weight coupled fecundity of females* due to additional energy uptake (Ibáñez & Keyl, 2010), thereby increasing fitness at the population level by supporting the gene transfer to the next generation, since females are considered as the main reproductive part of the population (Subsection 3.4.3).
- *Energy coupled growth*, as postulated in Keyl et al. (2008), increasing the survival rate of larger individuals by feeding on smaller conspecifics, thereby supporting their survival until reproduction.
- *Increased ecosystem carrying capacity* as a result of larger individuals feeding on smaller conspecifics, which themselves feed on lower trophic levels that larger individuals cannot access directly (Polis, 1981), thereby enabling faster energy propagation from lower to higher trophic levels. For example, large specimens of *D. gigas* gain access to energy resources in low oxygen saturation areas where smaller individuals feed (Alegre et al., 2014).
- *Support for reproduction* by the ensuring survival of larger individuals, i.e., the females due to the sexual dimorphism; in this case the cannibalism rate should be expected to intensify during reproductive periods (Ibáñez & Keyl, 2010).

7. *Cannibalism as a possible survival strategy*

A high incidence of cannibalism should promote the evolution of special cannibalism avoidance strategies on the part of the victim, but such adaptations are rarely observed. This is probably because such strategies would be “equally effective against interspecific predation” and therefore indistinguishable from general anti-predation strategies. So “most cases of cannibalism, can be thought of simply as intra-specific predation [...]” (Smith & Reay, 1991) and general anti-predator behavior should provide protection against cannibalism, provided that conspecifics are recognized as potential predators. Shoaling can be considered as “general antipredator behavior”, providing some defense at least against inter-cohort cannibalism (Smith & Reay, 1991), but its effects on the incidence of intra-cohort cannibalism are unknown.

One drawback of cannibalism might be the increased risk of infection by pathogens or parasites (Smith & Reay, 1991). This may apply to *D. gigas* (Ibáñez & Keyl, 2010), but is not considered in this thesis. Another potential drawback is a net decrease of fitness as a result of the consumption of close kin (Smith & Reay, 1991); this is unlikely to apply to *D. gigas*, since individuals die immediately after reproduction and therefore have no opportunity to predate on their progeny.

7.2.4. **Cannibalism as an evolutionary stable strategy**

Cannibalism “appears as a very effective competitive strategy, increasing an individual’s contribution of genes in the next generation through improved survival and reproductive success, while decreasing that of others” (Smith & Reay, 1991). For this reason, it can be considered as an evolutionary stable strategy (Ibáñez & Keyl, 2010), i.e., a strategy that, once fixed in a population, cannot be “invaded” by alternative, initially rare strategies.

The risk of retaliation by similar sized prey item is a general objection against an evolutionary stable strategy (Dawkins, 2006); but this does not apply to teleosts, because these require a size disparity due to the size of their mouths. In case of *D. gigas*, rapid and maybe dimorphic energy-dependent growth might create a size disparity between the sexes within a cohort (Subsection 6.3.2).

Size disparity among cohorts may be expected because *D. gigas* exhibits multiple spawning peaks throughout the year. Retaliation may also be avoided by predated on weak conspecifics; this possibility is incorporated into the concept of a modified WOP that is presented in Section 7.8.

7.3. The size disparity as a prerequisite for cannibalism

7.3.1. Overview

Predating on conspecifics of similar size is a hazardous strategy with the immanent risk of retaliation by the selected prey item, so a size disparity is a prerequisite for a high probability of a successful attack. The required size disparity sets a window of cannibalism upper bound (WOC_{UB}) for a prey item that determines the maximum relative size of the victim to the cannibal. But a victim should also provide enough energy to compensate for the energy expended in pursuing it, and this determines the minimum relative size, i.e., the window of cannibalism lower bound (WOC_{LB}). At the population level, the WOC_{LB} needs to be high enough to limit the incidence of cannibalism to levels that do not unduly reduce the population size and thereby endanger the survival of the population.

In this thesis, the energy driven life history model (EDLHM) (Figure 4.2, p.50) is driven by the energy supply, which limits growth when energy uptake is insufficient. If, on the other hand, one assumes that sufficient energy is always available during the lifetime of an individual, then the ML of an individual depends in the model only on age, sex and growth function (Subsection 4.2.2). The previous sections discussed the behavior of the model at the individual level. In this chapter, where the behavior of parameters is explored at a higher level e.g., school, a preset homogenous composition of the schools is assumed, i.e., that schools consist of individuals of the same age that have experienced the same environmental conditions. This assumption ignores the possibility of prior energy-dependent growth within schools, and all individuals of the same sex and age have the same size because they follow the same sex-specific growth function. An initial size disparity is introduced into the model by considering age distribution. Consideration is also given to the effects of different growth rates at the individual level, i.e., due to different energy uptake, on subsequent size distribution. The potential for cannibalism, and its possible effects, are analyzed for a range of size gradients, determined by different combinations of age differences and energy uptake rates. The special case of a uniform age distribution at population level is not considered, since it is known that *D. gigas* has a staggered reproductive season (Ibáñez et al., 2011; Jereb & Roper, 2010; Keyl et al., 2011; Rosa et al., 2013).

Subsection 7.3.2 sets out the calculations used to estimate the $size_{ratio}$ between victim and cannibal and the food uptake by cannibalism. Subsection 7.3.3 examines the parameters controlling of the dimorphic terminal size growth function (DTSGF) and their effects. It analysis the effects of cannibalism under a range of specific hatching

7. Cannibalism as a possible survival strategy

offsets which would be expected in the field, since these are expected to be the most important controlling factor.

7.3.2. The size disparity and its effects

Cannibalism in cephalopods is considered strongly size related (see Bruno et al. (2021) and Liu et al. (2020)), since physical capacity for cannibalism increases with size (ML), while the first ontological stages are unlikely to be capable of cannibalism (Ibáñez & Keyl, 2010). The WOC = [WOC_{LB}, WOC_{UB}] defines, similar to the window of predation (WOP), a range of relative sizes ($size_{ratio}$) within which a larger individual is able to predate on a smaller conspecific. The window of cannibalism lower bound (WOC_{LB}) and window of cannibalism upper bound (WOC_{UB}) define, respectively, the minimum and maximum relative size of the victim relative to the size of the cannibal. The value of the window of cannibalism upper bound (WOC_{UB}) is not easy to determine because the ability of *D. gigas* to tear off the prey's body tissue allows *D. gigas* to prey on relatively large conspecifics, compared to other species that consume their prey whole. Some estimates of WOC_{UB} for *D. gigas* are as high as 0.87 (87%) (Keyl et al., 2008; Markaida & Sosa-Nishizaki, 2003), but lower values such as 0.64 (Markaida et al., 2008) have also been proposed. In what follows, for ease of calculation, the WOC is pragmatically set at [0.1, 0.9]. The relatively high value of 0.9 was selected for WOC_{UB} in order to promote cannibalism in the simulations and to highlight its potential effects.

Assuming the age determines size for individuals that have experienced the same environmental conditions, then:

$$f_{size_{ratio}}(age, \Delta t) = \frac{*f_{size}(age)}{*f_{size}(age + \Delta t)} \quad (7.1)$$

calculates the $size_{ratio}$ of a younger individual to an older individual with a Δt age difference by applying the sex-specific $*f_{size}(age)$ function. The star “*” denotes the sex in the function if distinguished. The resulting $size_{ratio}$ must be lower than or equal to WOC_{UB} to enable the predation of the smaller (younger) individuals.

To examine this point in more detail the following sections, a required window of cannibalism upper bound (reqWOC_{UB}), semantically equal to $size_{ratio}$, is calculated, from modeled sizes of potential cannibals and victims within a population, to determine whether or not cannibalism is available as an energy source under a range of parameter settings:

$$reqWOC_{UB}(age, \omega) = f_{size_{ratio}}(age, \omega) = \frac{*f_{size}(age)}{*f_{size}(age + \omega)} \quad (7.2)$$

7.3. The size disparity as a prerequisite for cannibalism

The computation of $\text{reqWOC}_{\text{UB}}$ applies the sex-specific growth function $*f_{\text{size}}(\text{age})$ and specifies the relative age of individuals using a relative time offset ω , where a positive $\omega > 0$ describes an individual being ω fractions of the lifespan older than the other individual; the relative time offset ω therefore allows exploration of the effects of the growth function $*f_{\text{size}}(\text{age})$ on the WOC independently of the absolute lifespan. The $\text{reqWOC}_{\text{UB}}$ is a measure of the relative size of individuals (i.e., males and females) under different conditions (i.e., a range of hatching offsets) over the course of the lifespan. Cannibalism is available when $\text{reqWOC}_{\text{UB}}$ falls within the $[\text{WOC}_{\text{LB}}, \text{WOC}_{\text{UB}}]$ interval.

The potential mass uptake from cannibalism is estimated from the size of the victim according to the model in Subsection 4.3.3. The mass uptake defines the energy uptake available through cannibalism and allows estimation of the necessary number of victims to meet the energy requirements of the cannibal.

Unless otherwise stated, the computations are based on a maximum SAM of 1.2m and 1.0m for females and males respectively and a lifespan of one year for both sexes, see Table 6.2. The hatching offset ω is fractional value, calculated as the time that elapsed between hatching of the older and younger individual, relative to the lifespan of one year. The age_{rel} denotes the age of an individual within its lifespan. Both ω and age_{rel} are described by the interval $[0, 1]$, i.e., the sum of both values is capped at 1.0 because an individual does not grow larger beyond a relative age of 1.0, see Section 4.3.

7.3.3. The behavior of the DTSGF in respect of cannibalism

The controlling parameters

Overview. In the energy driven life history model (EDLHM), the current ML of an individual depends on (1) the sex-specific growth function (dimorphic terminal size growth function (DTSGF)), (2) the $\text{ML}_{\text{terminal}}$ and (3) age. In the case of cannibalism, the age difference between cannibal and prey, denoted by hatching offset ω , is an additional factor to be taken into account. In the following, the effects of varying $\text{ML}_{\text{terminal}}$ and ω are explored separately, due to the difficulty of visualizing the four-dimensional solution space that would be produced if the effects of three input variables (age, $\text{ML}_{\text{terminal}}$, ω) and size were considered simultaneously.

Varying the $\text{ML}_{\text{terminal}}$. Figure 7.1 shows the effect of different $\text{ML}_{\text{terminal}}$ -tuples constructed so that in all cases the $\text{ML}_{\text{terminal}}$ of females is 1.2 times longer than the males, in accordance with observed sexual dimorphism. The first and last tuple

7. Cannibalism as a possible survival strategy

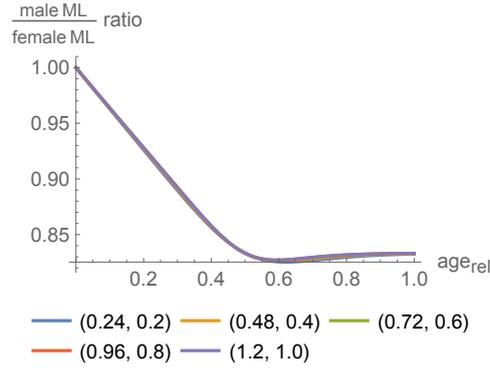


Figure 7.1.: DTSGF: The male:female ratio plotted against relative age for different ML_{terminal} tuples, e.g., $(0.24, 0.2)$, whereby the male:female ML_{terminal} remains 1.2 in all cases. The colored lines are difficult to distinguish because the different tuples have only a small effect on the resulting ratio.

represent assumed minimum and maximum ML_{terminal} values, the remaining tuples are equidistant intermediate values.

The DTSGF behaves robustly across all ML_{terminal} -tuples since varying the ML_{terminal} has almost no effect on the $size_{\text{ratio}}$. Therefore the *relative* lifespan $age_{\text{rel}} = [0, 1]$ can be used to represent any absolute lifespan.

With respect to cannibalism, the DTSGF provides the females with a size advantage over the males. However this may provide only limited opportunities for cannibalism in the field because the lowest $reqWOC_{\text{UB}} > 0.8$ at approx. 0.6 is in the upper range of the WOC bounds given in literature. The possibility of cannibalism may not be “binary”, i.e., there is no precise value of WOC_{UB} below which the success of a cannibalistic attack is assured.

Varying the hatching offset. The above analysis of the effects of varying ML_{terminal} assumes that all individuals hatch simultaneously, i.e., hatching offset $\omega = 0$. A $\omega > 0$ creates differences in age and thus —to a limited extent— length among individuals.

The exponential growth phase of the growth function may, due to the large absolute size changes during this phase, amplify initially small differences caused by hatching offsets to the extent that cannibalism among individuals born at different times becomes an available option. A hatching offset $\omega > 0$ may enable intra-sex cannibalism, in addition to inter-sex cannibalism, because size differences are expressed within as well as between the sexes. Therefore, effects of varying the hatching offset $\omega > 0$ over the lifespan have to be computed for four sex–age combinations in which an older individual predate on a younger individual. The computations for these combinations of sexes over the entire

7.3. The size disparity as a prerequisite for cannibalism

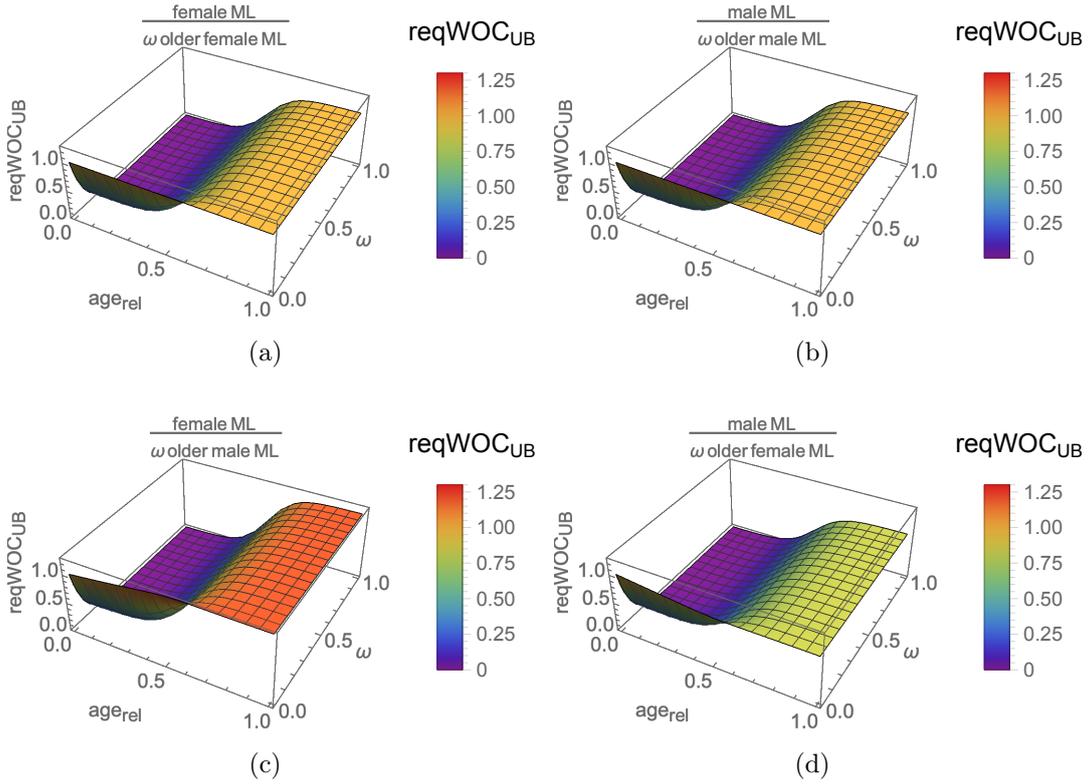


Figure 7.2.: DTSGF: Values of $\text{reqWOC}_{\text{UB}}$ for all female and male combinations, plotted against $\omega = [0, 1]$ and relative age $[0, 1]$. The resulting age by $\omega + \text{relative age}$ may exceed the upper bound of the $[0, 1]$ age-interval but the computed ML is still inside the $[0, \text{ML}_{\text{terminal}}]$ interval due to the asymptotic growth function. So the results will not change in cases (which are unlikely in the field) where $\omega + \text{relative age} > 1.0$.

lifespan and for $\omega = [0, 1]$ are shown in Figure 7.2. Figure 7.3 displays the same results but in the form of a modified contour plot where only values of $\text{reqWOC}_{\text{UB}} \leq 0.9$ are displayed; these figures show, for each combination of sexes and all values of ω , the relative age of the victim over its full lifespan when cannibalism by an ω older individual becomes possible, i.e., when $\text{reqWOC}_{\text{UB}} \leq 0.9$.

As shown in the diagrams, a higher ω increases the ability of individuals of both sexes to predate on conspecifics by lowering the $\text{reqWOC}_{\text{UB}}$. For the same ω , the $\text{reqWOC}_{\text{UB}}$ increases with the age of the victim (x-axis), since the size advantage of cannibal is reduced as growth slows towards the end of the lifespan. Towards the end, the $\text{ML}_{\text{terminal}}$ limits the size advantage by capping $\text{age} + \omega$, which is used to compute the ML of the cannibal, to 1.

7. Cannibalism as a possible survival strategy

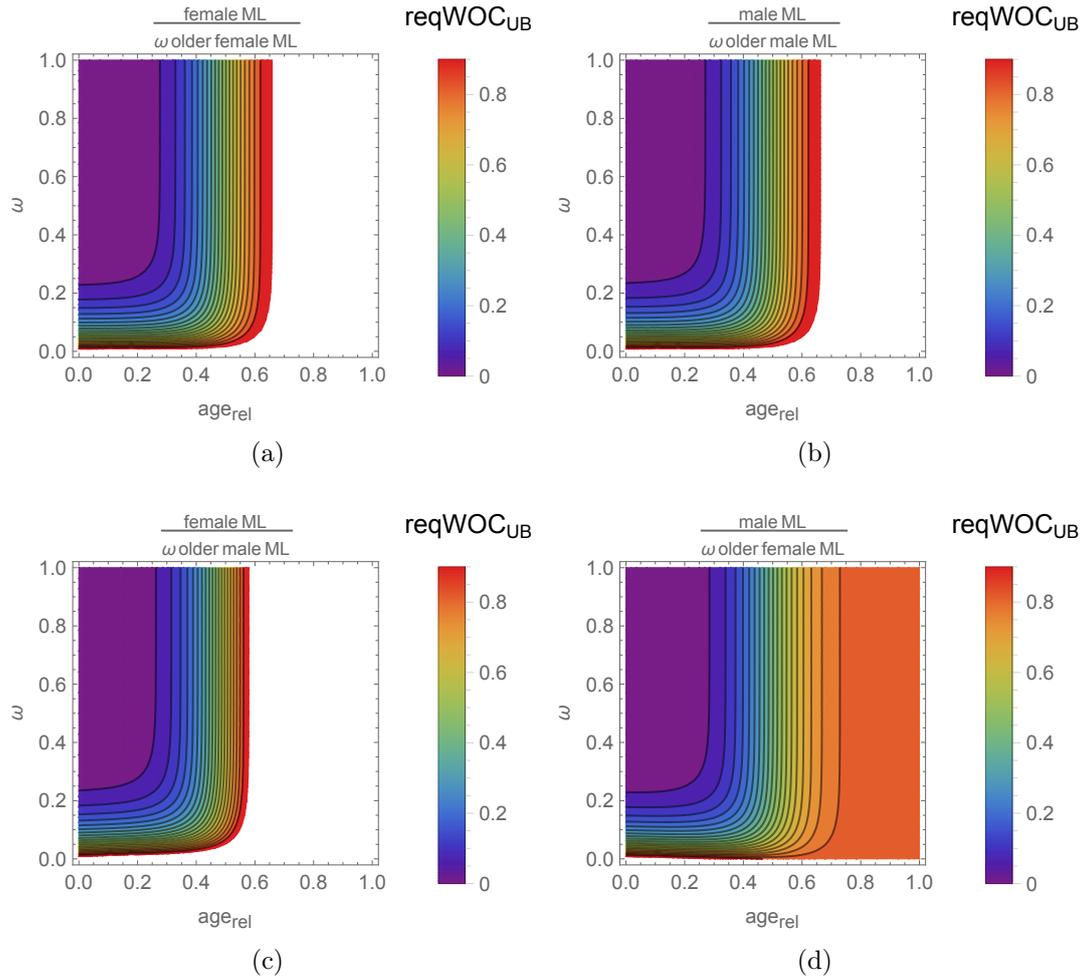


Figure 7.3.: DTSGF: Values of reqWOC_{UB} for all female and male combinations, plotted against $\omega = [0, 1]$ and relative age $[0, 1]$. The resulting age by $\omega +$ relative age may exceed the upper bound of the $[0, 1]$ age-interval but the computed ML is still inside the $[0, ML_{terminal}]$ interval due to the asymptotic growth function. So the results will not change in cases (which are unlikely in the field) where $\omega +$ relative age > 1.0 .

For intra-sex cannibalism, where an older individual preys on a younger conspecific of the same sex, values of reqWOC_{UB} approximate towards the shape of the growth function as values ω increase towards 1.0, since in this limiting case the denominator in Figure 7.2 equals 1.0 for all relative ages, and the ML ratio between victim and cannibal approximates towards 1.0 towards the end of the lifespan, as shown in Figure 7.2a and 7.2b. An increasing ω lowers the reqWOC_{UB} and may enable intra-sex cannibalism until a relative age of approximately 0.6 (Figure 7.3a and 7.3b). From the age of 0.5

7.3. The size disparity as a prerequisite for cannibalism

exponential growth begins to slow and, from an age of 0.6, values of $\text{reqWOC}_{\text{UB}}$ are too high to enable an older individual to prey on a younger conspecific of the same sex, even where ω values are high (and approach 1.0).

With respect to inter-sex cannibalism, an increasing ω also expands the WOC for cannibalism by an older male on a younger female conspecific (Figure 7.2c). For example, a $\omega > 0.1$ eliminates the female size advantage resulting in $\text{reqWOC}_{\text{UB}}$ values below the observed WOC_{UB} during approximately the first half of the males' lifespan. The female size advantage is regained after the exponential growth phase of the males comes to an end. From this point on, to the end of lifespan, the growth of females towards a larger $\text{ML}_{\text{terminal}}$ results in uniformly high $\text{reqWOC}_{\text{UB}}$, i.e., unfavorable values for male on female cannibalism.

For females (Figure 7.2d), the effects of increasing ω on the availability of inter-sex cannibalism are similar (Figure 7.2c), but in this case values of $\text{reqWOC}_{\text{UB}}$ remain favorable for female predation on younger male conspecifics until the end of the lifespan, because the increasing ω simply amplifies the inherent female size advantage as both sexes grow towards their respective $\text{ML}_{\text{terminal}}$, whereby the male:female size = $\frac{1.0}{1.2} = 0.8\bar{3}$, a value that is below the chosen WOC_{UB} of 0.9.

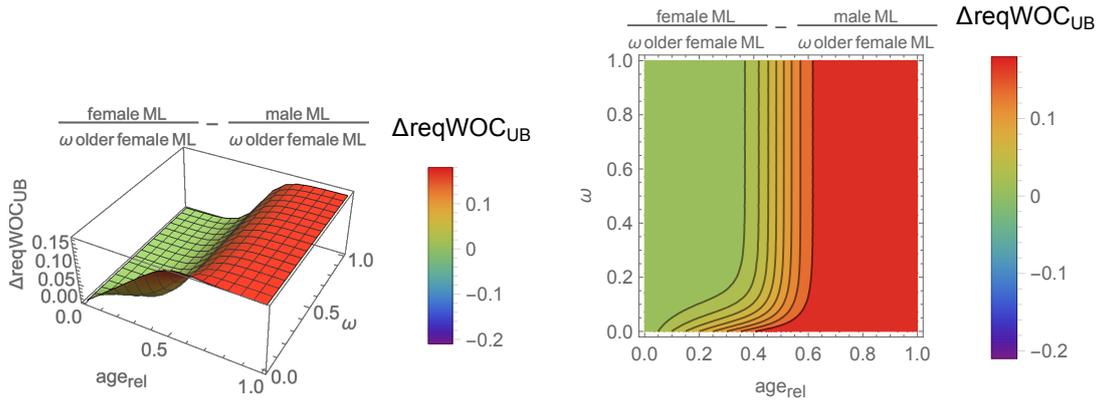


Figure 7.4.: DTSGF: Values of $\Delta \text{reqWOC}_{\text{UB}}$ for females, plotted against $\omega = [0, 1]$ and relative age $[0, 1]$, where positive values indicate size ratios that favor cannibalism by females on males instead of other females.

An increasing hatching offset ω amplifies the ability of both sexes to perform cannibalism, however it is unclear which sex benefits more from a rising ω . To address this issue, $\Delta \text{reqWOC}_{\text{UB}}$ was computed, representing the difference between $\text{reqWOC}_{\text{UB}}$ for the unfavorable case, in which one sex cannibalizes itself, and favorable case, in which one sex cannibalizes the other. Values of $\Delta \text{reqWOC}_{\text{UB}}$ are calculated by subtracting the $\text{reqWOC}_{\text{UB}}$ of the favorable case from the $\text{reqWOC}_{\text{UB}}$ of the unfavorable case. Thus, a

7. Cannibalism as a possible survival strategy

positive value indicates a more favorable size ratio for cannibalism on an individual of the opposite sex.

The comparison between the sexes shows that the females benefit considerably from an increasing ω (Figure 7.4). The positive values (or, in the worst case, zero value) of $\Delta\text{reqWOC}_{\text{UB}}$ show that, for the same value of ω , size ratios favor cannibalism by females on males rather than on other females.

The males, on the other hand, cannot take advantage of a hatching offset ω to prey on females (Figure 7.5), because the negative values (or, in the best case, zero value) of $\Delta\text{reqWOC}_{\text{UB}}$ indicate that size ratios favor cannibalism by males on other males rather than females.

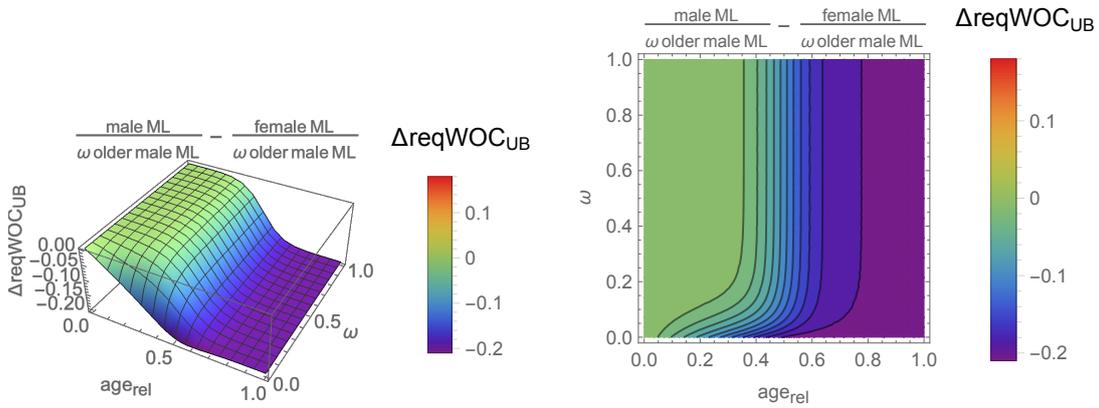


Figure 7.5.: DTSGF: Values of $\Delta\text{reqWOC}_{\text{UB}}$ for males, plotted against $\omega = [0, 1]$ and relative age $[0, 1]$, where positive values indicate size ratios that favor cannibalism by males on females instead of other males.

Conclusion. In general, the hatching offset ω is a significant factor controlling the availability of cannibalism as a source of energy. An increasing hatching offset ω amplifies the ability to engage in cannibalism for both sexes.

The observed sex ratio, which is generally in favor of the females, is compatible with the modeled effects of changing ω on intra- and inter-sex cannibalism. During periods of low environmental productivity with probable food scarcity, which could be expected to trigger cannibalism, the effect of cannibalism is a further shift of the sex ratio in favor of females. The sex ratio change in favor of females was proposed as a population-level response to adverse environmental conditions (Tafur et al., 2010) and cannibalism may be a mechanism (“tool”) by which this change occurs as a result of actions that are also effective on an individual level as a short-term response to adverse environmental conditions.

Specific hatching offsets and their effects

Overview. Spawning by *D. gigas* takes place throughout the year, however with distinct spawning peaks, while each female lays her eggs in 8–14 spawning batches (Subsection 2.1.4). As a result of spawning peaks, cohorts are likely to develop consisting of similarly aged individuals that experience the same environmental conditions and, therefore, have approximately the same size. This spawning pattern creates hatching offsets, both between spawning peaks and between spawning batches. This spawning pattern creates a range of potential constellations for cannibalism, i.e., intra-school, intra-cohort, inter-school, inter-cohort and inter-generation, each with a characteristic ω range.

In the model, size differences result from age differences, because the individuals are assumed to experience very similar environmental conditions (since *D. gigas* exhibits no significant latitudinal migration (R. I. Ruiz-Cooley et al., 2010)) and thus to exhibit uniform growth patterns. Therefore ω can be used to model these constellations.

Each single spawning batch may create a school because of the delay between each batch that also represents a hatching offset, i.e., $\omega > 0$. The original school size is greatly reduced over the time as mortality among young individuals is assumed to be high (Boyle & v. Boletzky, 1996; Camarillo-Coop et al., 2013; Lipinski, 2002). The school is postulated as the smallest level of aggregation of individuals of roughly the same size. It is assumed that hatchlings form a school remain in close proximity to each other after hatching, i.e., as a result of transport by currents, which determines the horizontal movement of hatchlings in their early life stages. Hence a school is very likely to consist of hatchlings of the same spawning batch. For evaluation purposes, it is assumed that hatchlings of the same spawning batch have equal size on hatching, while hatching offsets create size differences. In fact, it is unknown whether a single spawning batch consists of hatchlings of the same size at the time of hatching. However, if egg size depends on the ML of the mother as reported for *D. gigas* and other cephalopods (Birk et al., 2016; Pecl & Jackson, 2008), the hatchling size should be approximately the same for all individuals of a batch, since they are from eggs spawned by the same mother and develop under the same environmental conditions.

Intra-school cannibalism, compared to the other constellations, occurs among individuals within a narrow, clearly defined ω -range that is amenable to more precise theoretical evaluation. In the following analysis, a maximum hatching offset of 10 days within a school is assumed, since, over longer periods, currents are likely to disperse the egg masses and the already hatched vertically mobile paralarvae. Ages of individuals

7. Cannibalism as a possible survival strategy

that make up the school may be expected to be normally distributed, so the hatching offset ω between most pairs of individuals is probably $\omega < 10$ days.

It should be borne in mind, however, that currents may form a spatial aggregation system that draws free-floating paralarvae from different schools towards attractor regions (Section 5.5). Such a process may mix schools, thus creating schools with a more complex internal age structure and larger hatching offsets ω .

Intra-cohort cannibalism, i.e., among individuals from the same spawning peak, may be considered similar to intra-school cannibalism because the age structure of a cohort can be expected to be similar to that of a school, while ω values for intra-school cannibalism are likely to lie within the ranges for inter-school and inter-cohort cannibalism. For these reasons, the case of intra-cohort cannibalism is not considered in the following analyses.

In the simulation of the effects of cannibalism in this chapter (Section 7.6), the value of inter-school or inter-cohort hatching offsets is presumed to be 30d to 120d a one-year lifespan. This is assumed to be lower than the (unknown) generation time, which is likely to vary according to environmental conditions.

Inter-generation cannibalism includes all cases of cannibalism between individuals with hatching offset values $\omega \geq 0.5$, assuming that all individuals have approximate the same lifespan, which is equal to the generation period $\omega = 1$. Since the spawning season lasts several months, a difference $\omega \geq 0.5$ can be assumed between youngest (last hatched) individuals of one generation and the oldest (first hatched) individuals of the following generation. As explained above, the temporally organized spawning in combination with currents is presumed to result in the formation of distinct schools of different ages, but these schools may come into contact with each other. The hatching offset ω is the dominant factor controlling the size disparity and, therefore, the occurrence of cannibalism.

The following section examines the ω ranges of these different potential cannibalism constellations that can be expected to occur in the field more closely, in order to determine how they affect the potential for cannibalism and its outcomes.

Analysis. Until now ω has been defined as a number of days, i.e., an absolute time, while other aspects of life history have been defined in relative terms, i.e., as a fraction of the lifespan. The relative value of ω will therefore depend on the absolute value of the lifespan. Thus ten days (10d) corresponds to $\omega = 0.0558$ for a lifespan of half a year, $\omega = 0.0274$ for a one-year lifespan and $\omega = 0.0137$ for a two-year lifespan. To take all these possibilities into account, an interval of $\omega = [0, 0.0558]$ is considered in

7.3. The size disparity as a prerequisite for cannibalism

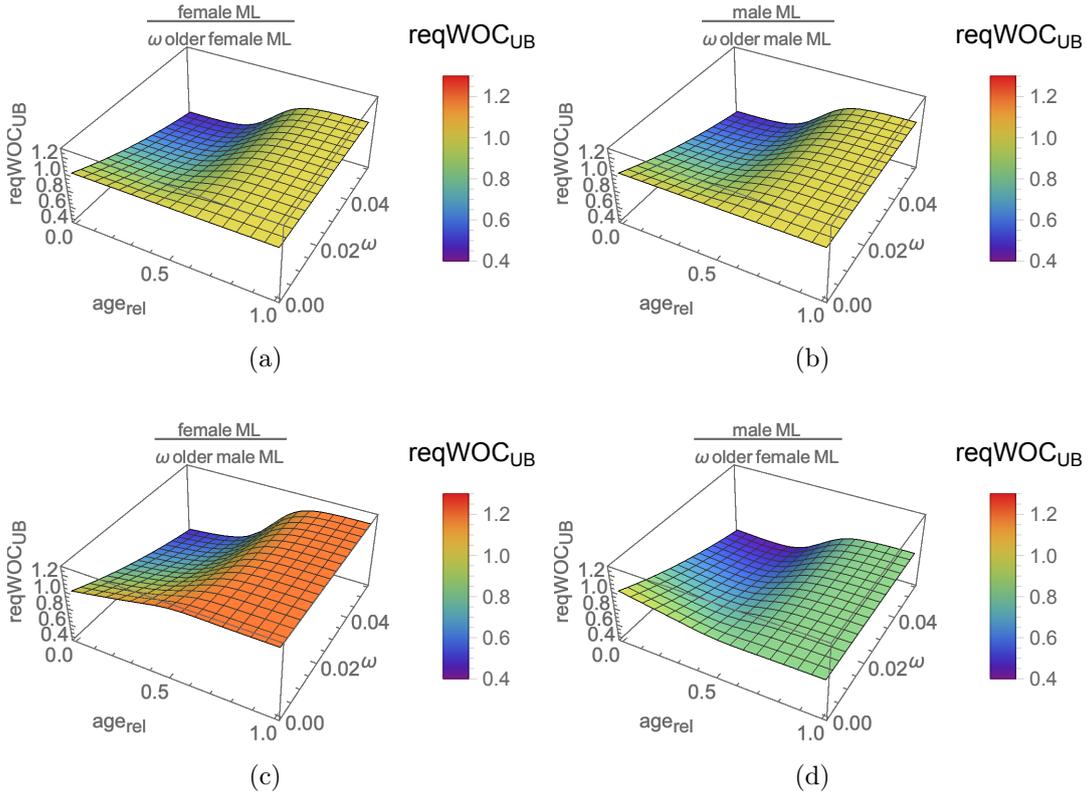


Figure 7.6.: DTSGF: Values of $\text{reqWOC}_{\text{UB}}$ for all female and male combinations of cannibalism, plotted against $\omega = [0, \approx 0.0558]$ (corresponds to $[0, 10]$ d in case of a half year lifespan) and relative age.

the following analysis. However, this range would correspond to a hatching offset ω of $[0, 40]$ days for a two year lifespan; these aspects should be borne in mind when interpreting the results of the analysis.

The Figure 7.6 and 7.7 show values of $\text{reqWOC}_{\text{UB}}$ for the DTSGF within the range of hatching offsets that can be expected in a school. In Figure 7.7, assuming $\text{WOC}_{\text{UB}} = 0.9$, cannibalism is only possible between the stated predator and prey at combinations of ω and relative age indicated by colored regions on the diagram.

In case of a small ω of 10d, the exponential phase of the DTSGF amplifies the size differences, thus lowers the $\text{reqWOC}_{\text{UB}}$ for an older individual to prey on a younger one, but not necessarily to $\text{reqWOC}_{\text{UB}} \leq 0.9$. With increasing hatching offset ω , values of $\text{reqWOC}_{\text{UB}}$ decrease in the intra-sex cases in favor of older individuals and between the sexes in favor of females. However, the magnitude of the effect of increasing ω depends on the sex of predator and prey and on the lifespan, since a shorter lifespan

7. Cannibalism as a possible survival strategy

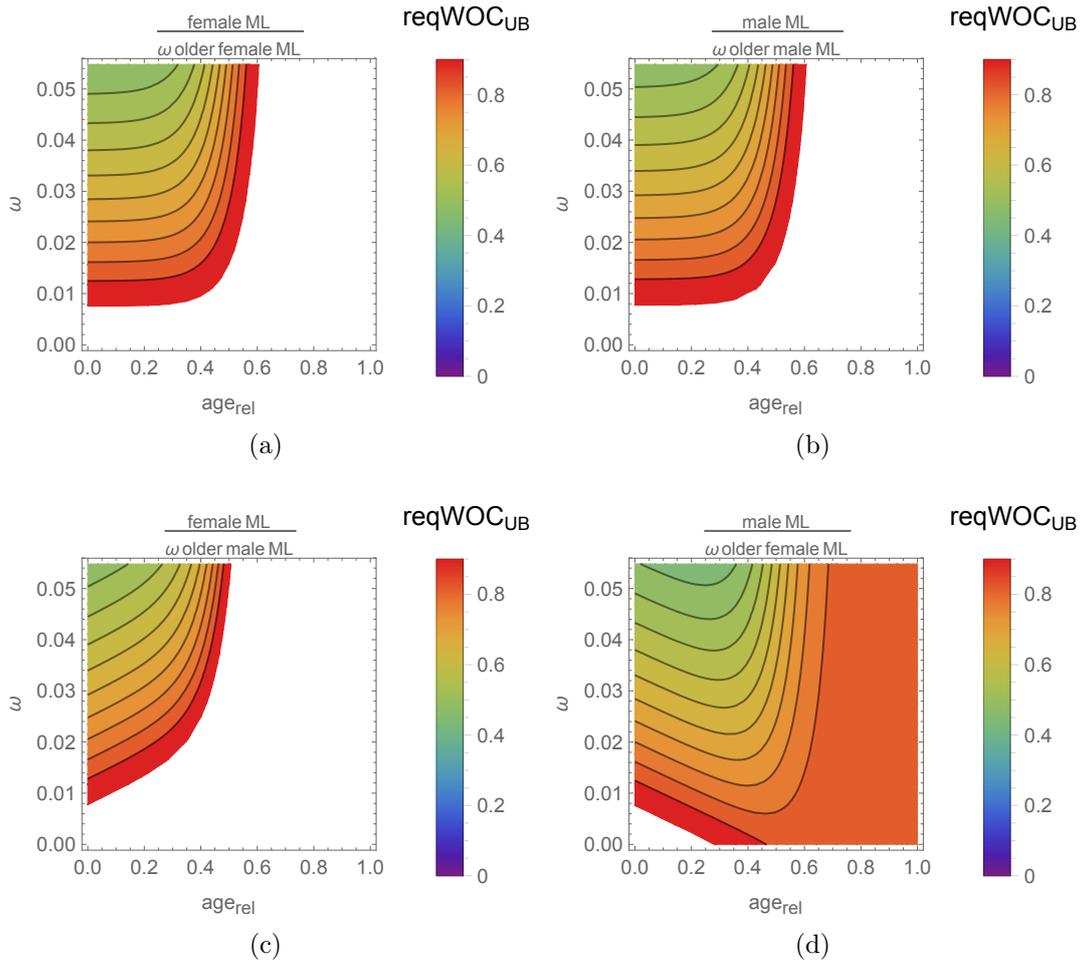


Figure 7.7.: DTSGF: Contour plot of values of reqWOC_{UB} for all female and male combinations of cannibalism, plotted against $\omega = [0, 0.0558]$ (corresponds to $[0, 10]$ d in case of a half year lifespan) and relative age.

may be considered to correspond to a higher relative ω value, as indicated by the white regions in Figure 7.7.

A hatching offset ω facilitates predation of younger females by older males during the first half of the lifespan, but the lower reqWOC_{UB} in the male intra-sex case makes predating on younger males a more favorable option. By contrast, the hatching offset allows older females to predate on younger males over the entire lifespan, since at all times reqWOC_{UB} values for female on male cannibalism are lower than for other cannibal-prey combinations. Not only is intra-sex cannibalism among females less likely, since females cannibals can be expected to select males as a prey item in preference to other females, but males will also be more vulnerable to attacks by other predators,

7.3. The size disparity as a prerequisite for cannibalism

which are more likely to select smaller individuals as prey. In summary, a hatching offset within a school elevates the probability of cannibalism by amplifying differences in size among individuals.

Higher hatching offsets, i.e., $\omega = [0.0822, 0.3288]$ representing the range of 30d to 120d for a one year lifespan, used in this thesis to model the cannibalism potential of inter-school and inter-cohort cannibalism, do not show different effects, compared to those for lower hatching offset (compare Figure 7.8 to Figure 7.7). An increasing

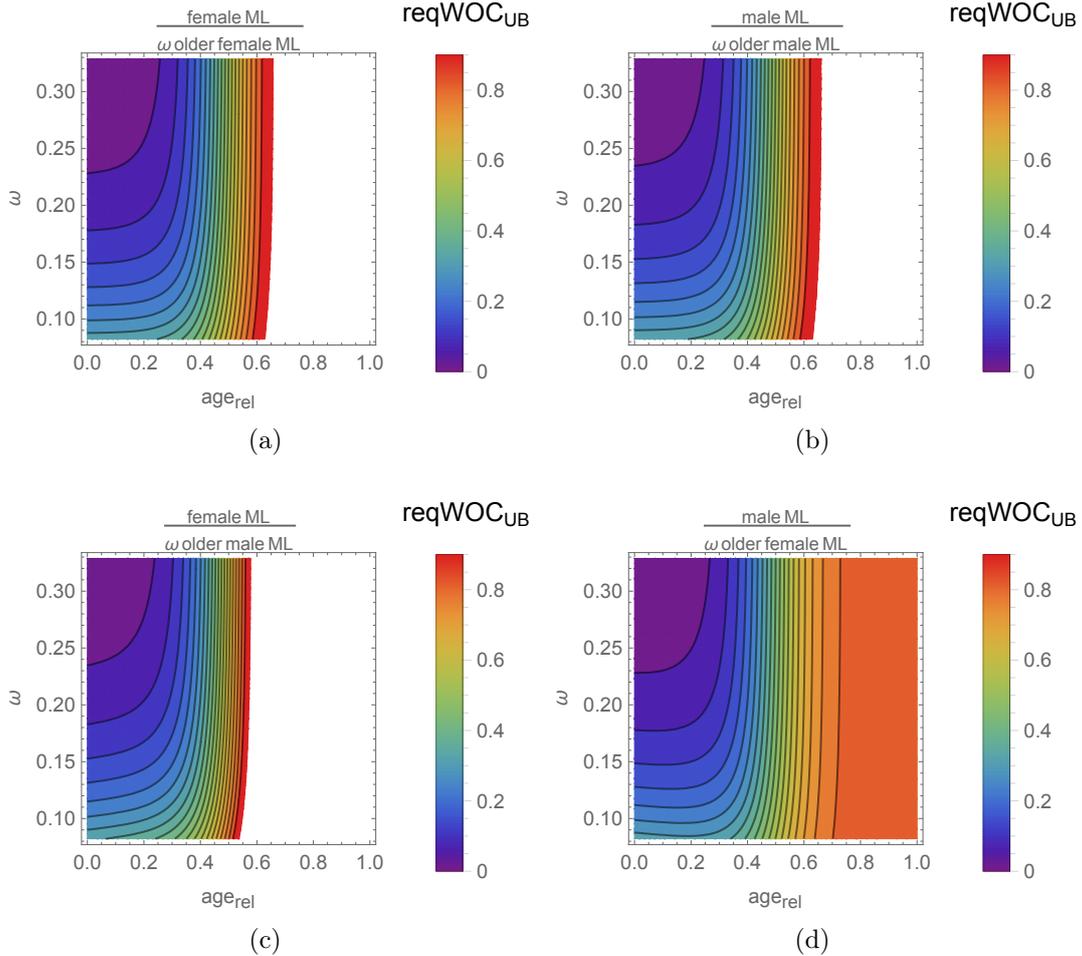


Figure 7.8.: DTSGF: Values of reqWOC_{UB} for all female and male combinations of cannibalism for $\omega = [0.0822, 0.3288]$ (30d to 120d), for a one year lifespan, plotted against relative age.

hatching offset lowers the reqWOC_{UB} for all sex combinations to values within the assumed WOC (Figure 7.8). The increased ω makes females unavailable for cannibalism

7. Cannibalism as a possible survival strategy

to males after the first half of their lifespan, while males are available to females over the entire lifespan.

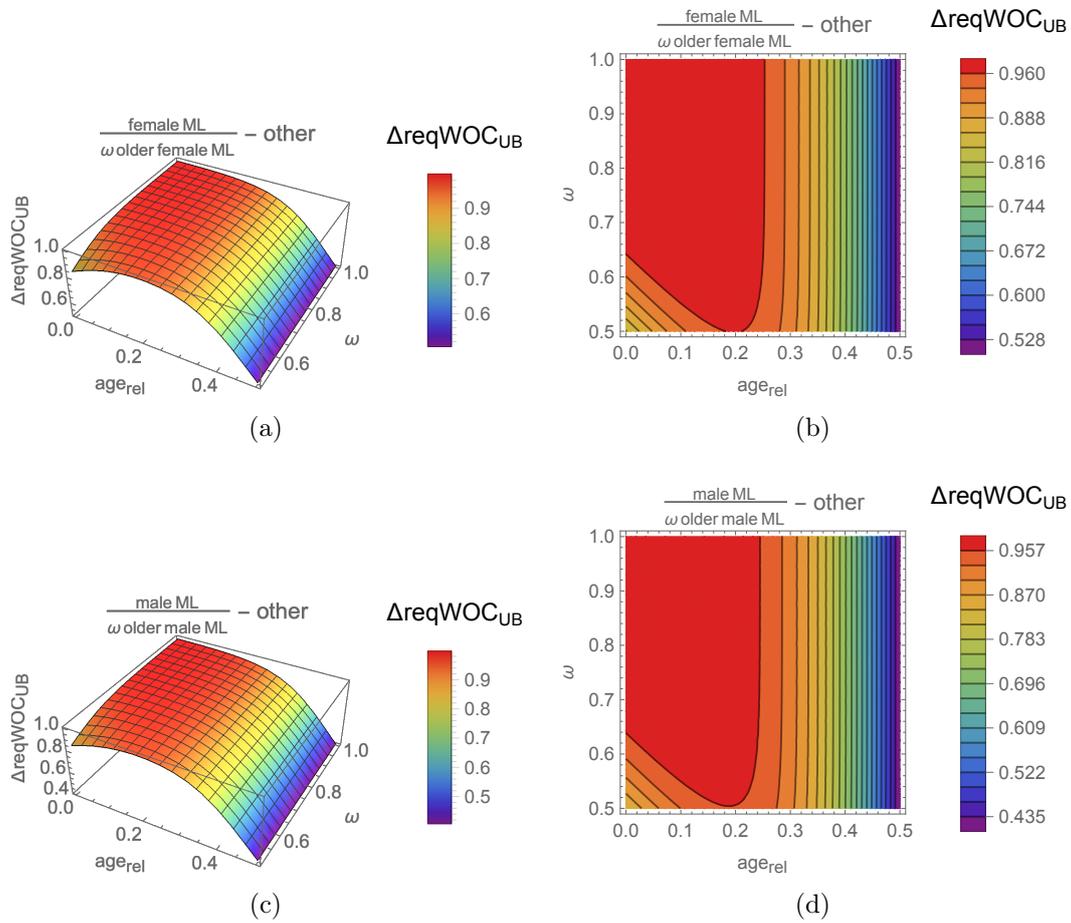


Figure 7.9.: DTSGF: $\Delta reqWOC_{UB}$, showing differences between $reqWOC_{UB}$ values for cannibalization of individuals of the same sex and the same older generation and the minimum $reqWOC_{UB}$ values for alternative combinations of sex and age, over the first half of the lifespan ($\omega = [0.5, 1]$).

The analysis for $\omega \geq 0.5$, which represents inter-generation cannibalism, is based on the results shown in Figure 7.3. For $\omega \geq 0.5$, i.e., inter-generation cannibalism, the potential for cannibalism is not more pronounced than for inter-school and inter-cohort cannibalism, with only moderate changes in $reqWOC_{UB}$ values compared to $\omega \geq 0.3$ for all sex combinations exhibit.

Figure 7.9 shows differences between $reqWOC_{UB}$ values, i.e., $\Delta reqWOC_{UB}$, for predation by an older individual of sex “a” on another older individual of the same sex and the minimum $reqWOC_{UB}$ from among the three alternative combinations of sex

and age, i.e., predation by an older individual of sex “a” on (1) an older individual of the opposite sex, (2) a younger individual of the same sex, and (3) a younger individual of the opposite sex. Values above 0 mean that one of the three cases has a lower $\text{reqWOC}_{\text{UB}}$, thus preying on the other sex or individuals of the younger generation is a more favorable option than preying of individuals of the same sex and the same older generation. The higher the (positive) difference, the more favorable is the alternative. Contrarily, a negative value means it is more favorable to prey on individuals of the same sex and the same older generation. The positive values in Figure 7.9 therefore show that it is relatively unfavorable to cannibalize conspecifics of the same sex from the same older generation.

Conclusion. The DTSGF is an advantageous growth function for females because it facilitates female cannibalism on males, as indicated by relatively low $\text{reqWOC}_{\text{UB}}$ for female on male cannibalism in the analyses. An increasing ω may encourage cannibalism that supports female energy uptake in times of food deprivation, presenting males as a more favorable prey item to other males, and is thus conducive to a change in the sex ratio within the population in favor of females.

7.4. Energy uptake by cannibalism

7.4.1. Overview

The estimation of the per capita energy uptake is an important criterion for evaluating the usefulness of cannibalism as a population energy buffer. A low energy uptake per cannibalized conspecific requires a high number of cannibalized conspecifics, which endangers the population and thus the species’ survival. To estimate the impact on the population, this section calculates the number of conspecifics which have to be sacrificed to satisfy the daily energy needs of a cannibal in times of complete energy uptake deprivation.

The victim’s body mass gives the cannibal’s energy uptake and the estimated number of conspecifics required to satisfy the daily energy needs. Equation (7.3) computes the fraction or number of ω -older cannibals, for a given value of ω , and a given energy conversion factor $f_{\text{m} \rightarrow \text{e}}$, whose daily total metabolic rate (total_{mr}) is satisfied by a victim as:

$${}^*f_{\text{cannibals}_{\text{fed}}}(\text{age}, \omega) = \frac{f_{\text{m} \rightarrow \text{e}} \cdot {}^*f_{\text{mass}}(\text{age})}{{}^*f_{\text{total}_{\text{mr}}}(\text{age} + \omega)} \quad (7.3)$$

7. Cannibalism as a possible survival strategy

The fulfillment of daily energy requirements corresponding to the total_{MR} allows regular growth of the cannibal; a partial fulfillment requires repeated cannibalism.

7.4.2. Analysis and discussion

The calculations were carried out under the premise that an individual follows its growth function and no restriction of growth takes place. These assumptions simplify the discussion, so that the hatching offset ω becomes the determining factor for the calculation of the size relation and thus the mass relation.

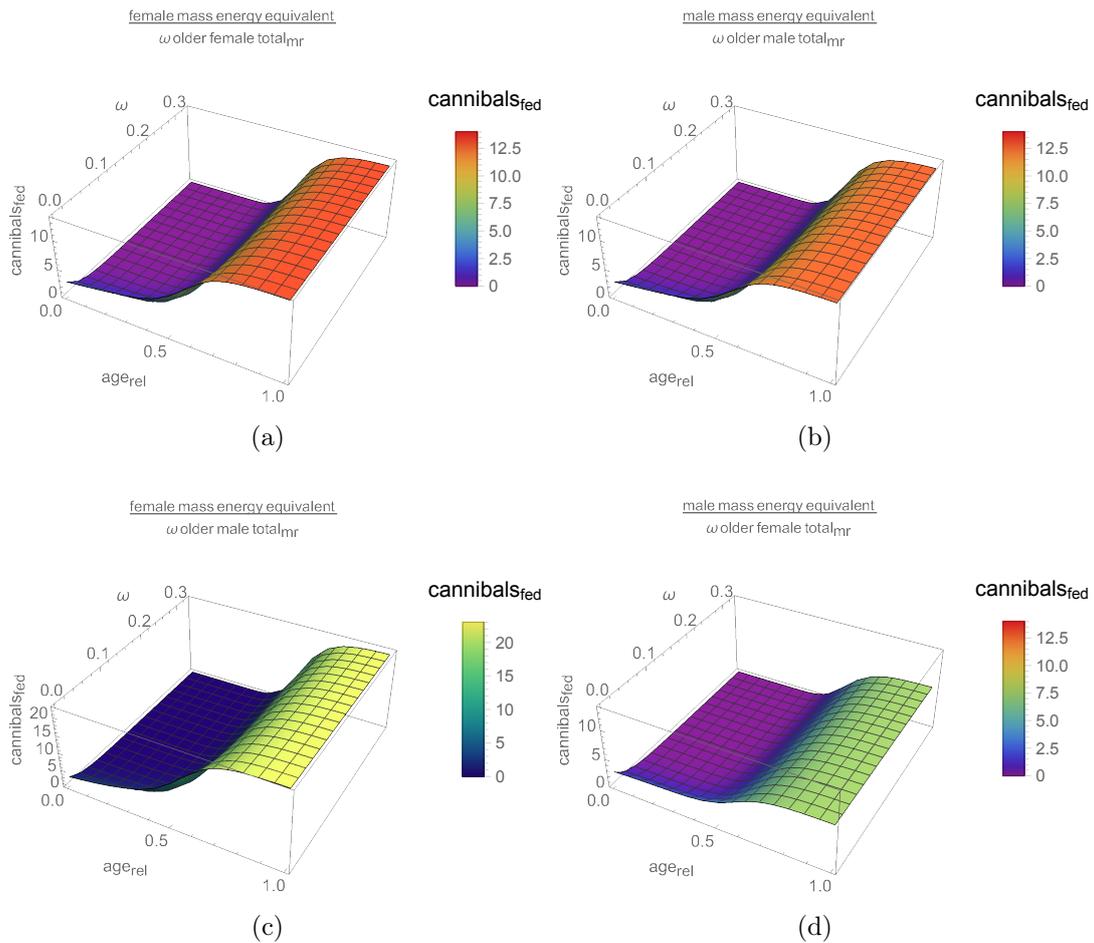


Figure 7.10.: DTSGF: The number of cannibals whose energy requirements are met by the mass of the victim for all female and male combinations of cannibalism, plotted against $\omega = [0.0, 0.3288]$ (0d to 120d), and relative age of the prey.

The cubic relation between mass and ML (Equation (4.25)) amplifies the mass differences between individuals of different ML. Since the mass is used to determine

7.4. Energy uptake by cannibalism

total_{mr} , the energy needs of a (cannibalistic) individual can be estimated from its mass. The mass ratio between the (usually larger) cannibalistic individual and prey indicates the proportion of the cannibal's energy needs that are satisfied by the prey item. Where the predator:prey mass ratio is high, this indicates the need for repeated cannibalism in order to meet the cannibal's energy needs.

The Figure 7.10 and 7.11 show the number of cannibals whose daily total_{mr} requirements are satisfied by a single (younger) individual for hatching offsets $\omega = [0.0, 0.3288]$ (0d to 120d for a one year lifespan) in case of intra-school, inter-school and inter-cohort cannibalism.

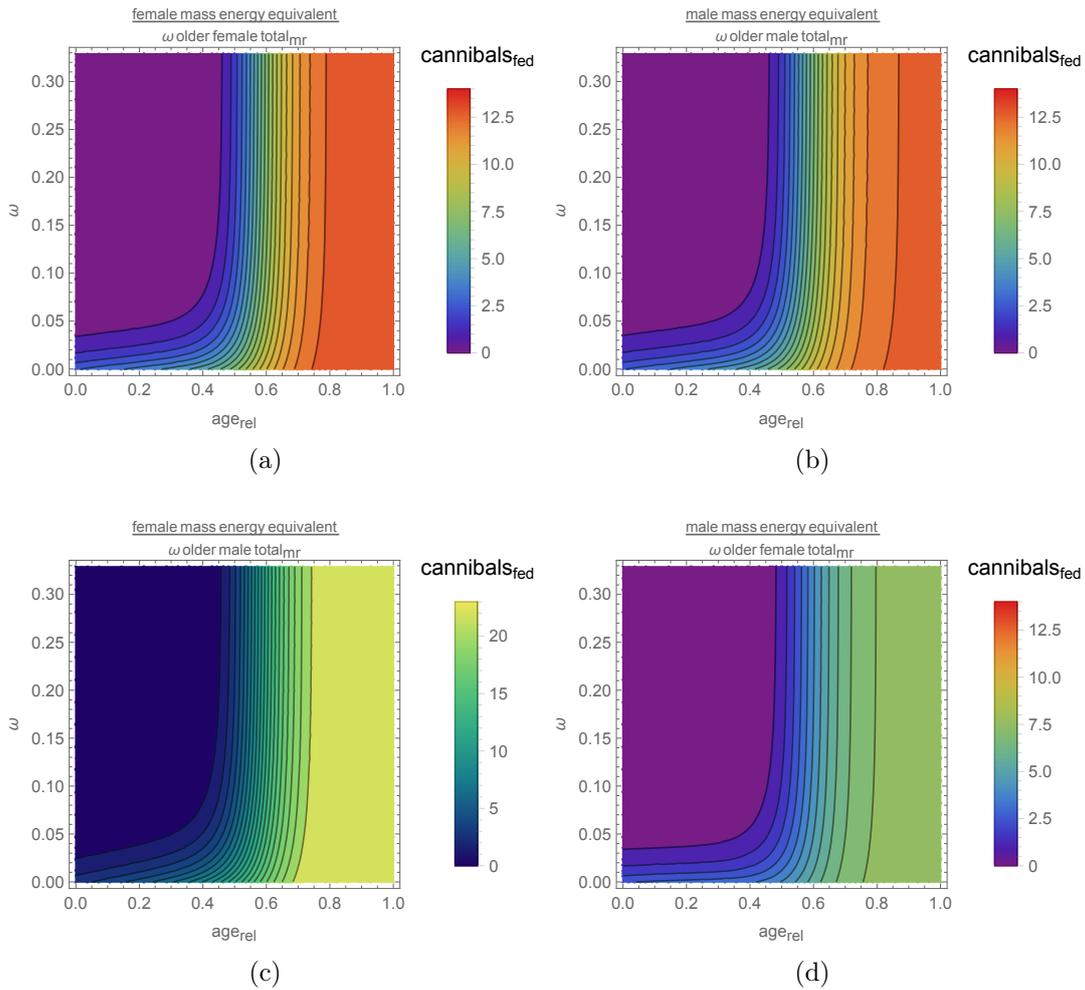


Figure 7.11.: DTSGF: Contour plot of the number of cannibals whose energy requirements are met by the mass of the victim for all female and male combinations of cannibalism, plotted against $\omega = [0.0, 0.3288]$ (0d to 120d), and relative age of the prey.

7. Cannibalism as a possible survival strategy

In all cases, an increasing ω reduces the number of cannibals whose daily energy needs are satisfied by a victim. Later on during growth, increasing age is a counteracting factor, because with increasing age the cannibal is in the asymptotic part of its growth curve so the ML ratio of cannibal to prey falls, the body mass of the prey approximates to the body mass of the cannibal and therefore the prey provides relatively more energy to the cannibal.

Males are the preferred prey, for cannibalism by either sex, therefore only cases where the male is victim are considered in the following discussion (Figure 7.11b and 7.11d). In these cases, a victim provides energy for more than one cannibal from an age of approximately 0.5, even with increasing ω , i.e., the hatching offset ω is no longer an influential factor from this age onwards. As a result, cannibalism will disproportionately reduce the number of males and change the sex ratio towards the observed prevailing female:male ratio reported by Tafur et al. (2010).

In the other cases, where females are the victims, (Figure 7.11a and 7.11c) fewer female conspecifics are required to satisfy the needs of the cannibal, especially when males consume females.

Similar results were obtained for $\omega > 0.5$ for all age-sex combinations, see Figure 7.11 for already smaller values of ω , so results for these ω combinations are not further discussed. For the completeness it should be noted that in theory younger and smaller cannibals may prey on older victims, whereby $\text{reqWOC}_{\text{UB}} > 1$ and, especially for large values of ω , obtain a high energy intake. This possibility was briefly examined but is not further discussed here because the results are obvious.

In conclusion, provided that cannibals and victims experience the same environmental conditions and thus have the same opportunities for growth prior to cannibalism taking place, males are disadvantaged by their relative lower body mass compared to females, since female cannibals need to consume relatively more males. Moreover, the number of individuals that need to be consumed is probably higher, since the conversion efficiency of prey biomass into cannibal biomass is considered in the calculations with an efficiency factor of 0.9. By contrast, Tafur et al. (2010) assume an efficiency factor of prey biomass conversion of 40% (0.4) while acknowledging that even this value may be too high.

If cannibalism is a strategy to use the population as an energy buffer (Ibáñez & Keyl, 2010), then the energy of this buffer should not be wasted. But, if the mass of a victim provides more energy than a single cannibal needs or can consume, then this would be a waste of energy. Therefore a variant of cannibalism, i.e., scavenging, could be a more efficient strategy to maintain energy uptake and increase the chance of survival in

7.5. Conclusions on prerequisites on size disparity and energy uptake

times of scarce food, as further discussed, under the heading of “passive cannibalism”, in Subsection 7.6.1.

7.5. Conclusions on prerequisites on size disparity and energy uptake

Cannibalism provides access to an additional food source through a trade-off against population reduction. It could be a strategy or a behavioral trait to promote the survival of the population at the expense of sacrificed conspecifics. But in the case of *D. gigas*, the evaluation of the effects of cannibalism in the previous sections indicates that these are highly dependent on the model and the growth functions used.

Cannibalism could be a tool to control the sex ratio and thereby lower the population energy demand. In the DTSGF, the effects of increasing age difference and thus also size generally favor cannibalism of males by females. This suggests that cannibalism is a viable strategy for ensuring population survival, since it favors the survival of the females that are needed for reproduction.

The DTSGF also indicates that larger prey, i.e., individuals above a certain ML, provide more energy than a single cannibal is able to consume. Cooperative cannibalism, in which two or more cannibals share the same victim, might be a solution to avoid wasteful use of this population energy buffer. Wasteful cannibalism would not contribute to survival of the population, since more victims than necessary would be sacrificed. The potential for a “cooperative” cannibalism strategy of this kind is a further feature of “passive cannibalism” (see Subsection 7.6.1).

The results presented above assume a ML_{terminal} of 1.2m for females and 1.0m for males. Computations for smaller ML_{terminal} (not presented in this thesis) give similar results, with regard to the influence of the growth functions on the WOC, and the effects of varying the hatching offset ω and relative age of prey.

It should be born in mind that in many cases, especially for small ML in early life stages, the computed and discussed effects may rely on small absolute ML differences. In view of these small absolute differences, it is questionable whether these effects would be significant in the field.

The next section uses a simulation model to evaluate the DTSGF in combination with different cannibalism strategies to identify emerging effects of different modalities of cannibalism under total energy deprivation.

7. Cannibalism as a possible survival strategy

7.6. Simulation and exploration of cannibalism by an energetic view

7.6.1. Model overview

This section evaluates the effects of varying model parameters on the number of individuals in a school. The schools are spatially separated and therefore do not interact, nor are external factors such as temperature included in the model. The absence of external factors means that no information about location is required. So the schools and the individuals they contain are not linked to any location information and the interaction of the individuals within the school is modeled stochastically.

The school is chosen as the smallest organizational unit for study of the effects of different cannibalism strategies in combination with the growth function; the effects observed within a school can be transferred to a population, since the population is essentially composed of schools.

An energy model specifies the individual's total metabolic rate (total_{mr}) as an inherent energetic deficit. The simulations assume that no other energy is available in the environment (i.e., a situation of total energy deprivation); thus the energy deficit must be covered by energy uptake through cannibalism, whereby the cannibal identifies conspecifics and selects one from these as its prey. In the model, starvation and cannibalism are triggered by energy level thresholds.

At each simulation step, information about the schools, such as school size and the number of cannibals, is logged and used to evaluate the different cannibalism strategies in relation to the growth functions. Table D.2, page 301, lists the controlling simulation model parameters; for more technical information see Appendix D.4.2 and the text below.

Individual. An individual uses the variables *age*, *ML* and its current *energy level* to assess its energy requirements based on the energy model derived from the discussion in Section 4.4, which incorporates a bidirectional convertibility of mass and energy and the estimation of basal+locomotion metabolic rate ($\text{basal+locomotion}_{\text{mr}}$) and growth metabolic rate ($\text{growth}_{\text{mr}}$). Surplus energy is converted into mass and the resulting *ML* is computed based on the body model in Subsection 4.3.3.

The *ML* defines the body mass *BM*, which in turn defines the *energy_{expected}* level, i.e., the energy equivalent of the *BM*:

$$\text{energy}_{\text{expected}} = \text{BM g} \cdot 17.2 \text{ J g}^{-1} \quad (7.4)$$

7.6. Simulation and exploration of cannibalism by an energetic view

This equivalence allows calculation of energy gain or loss, as a proportion of body mass, and simple modeling of starvation and energy deprivation, which in turn trigger cannibalistic behavior.

Behavior is controlled by three thresholds (1) $threshold_{max}$, the maximum energy level of the agent, (2) $threshold_{cann}$, that triggers cannibalism, and (3) $threshold_{min}$, below that an agent dies of starvation (terminal energy exhaustion). All thresholds are derived from $energy_{expected}$, as follows:

$$threshold_{max} = 1.1 \cdot energy_{expected} \quad (7.5)$$

$$threshold_{cann} = 0.9 \cdot energy_{expected} \quad (7.6)$$

$$threshold_{min} = 0.7 \cdot energy_{expected} \quad (7.7)$$

The general relation $threshold_{max} \geq threshold_{cann} \geq threshold_{min}$ is incorporated into the simulation model. Each simulation step t_i computes the current energy level by adding energy uptake and/or subtracting energy loss:

$$energy\ level_{t_i} = energy\ level_{t_{i-1}} + energy\ uptake_{t_i} - metabolic\ loss_{t_i} \quad (7.8)$$

Model variables are computed for each hour of simulation time using values from after cannibalism has taken place.

School modeling. A school consists of a set of individuals. Interactions between individuals are modeled stochastically as the the probability encountering a conspecific. Different hatching times are modeled by using a *start age* (in most cases 120 days) and a normally distributed *start age variation*:

$$age = start\ age + \frac{x}{2} \cdot start\ age\ variation \quad (7.9)$$

The algorithm draws x from a random number generator until $-2 < x < 2$ applies. The factor x is limited because, otherwise, a relative high *start age variation* with a low *start age* could result in a negative *age*.

Because a lack of energy can limit growth according to Equation (4.4) p.53, the virtual age of an individual is used to determine the point on the growth curve and thus the size of an individual. The size of an individual within a school is thus calculated from its virtual age by applying the growth function (see Subsection 4.2.2 and Section 4.4). The parameter *start age variation* initially defines this virtual age, whereby, according to Equation (7.9), $start\ age + 2 \cdot start\ age\ variation$ defines the maximum virtual age,

7. Cannibalism as a possible survival strategy

signifying that the individual attained the maximum size possible with unlimited food intake, i.e., the utilization of its maximum growth potential. Correspondingly, $start\ age - 2 \cdot start\ age\ variation$ describes the minimum size reached due to lack of food.

In combination with the growth functions, the $start\ age$ and $start\ age\ variation$ in the model allow the modeling of variable growth of individuals within a school due to different food intake.

Cannibalism classification and strategies. The model uses the control parameters, see Appendix D.4.1, to simulate no-cannibalism, active cannibalism, passive cannibalism and active+passive cannibalism strategies. Active cannibalism is the predation on living conspecifics where the WOC_{UB} is determined by the presumed defensive capabilities of the victim. Passive cannibalism, described in more detail below, is explored as an alternative, since initial results suggested that active cannibalism was less effective as a survival strategy than expected. The active+passive cannibalism strategy is introduced as a further working hypothesis to examine the potential contribution of a mixed cannibalism strategy to school survival in times of scarce food.

Passive cannibalism (Heinen & Abdella, 2005; Irvine, 1989; van Huis et al., 2008) is the consumption (scavenging) of already dead conspecifics. Dead individuals are referred as “terminally energy-exhausted” but the simulation model expands the potential prey of passive cannibalism by including lethargic conspecifics that are energy-exhausted but still alive. In the model, lethargy is a preliminary stage of terminal energy exhaustion, but is simulated as the same as death with regard to (the absence of) activity and defense actions. Thus $threshold_{min}$ determines the point at which a lethargic animal is indistinguishable from a dead individual (cadaver).

During energy deprivation, energetically expensive processes are usually reduced or shut down to conserve energy (Guppy & Withers, 1999), which is accompanied by a metabolic suppression. *D. gigas*’ metabolic rate drops by more than 75% during hypoxia (Trübenbach, Pegado, et al., 2013) and *D. gigas* also exhibits lethargic behavior under hypoxic conditions (Rosa & Seibel, 2010); thus energy exhaustion probably reduces the metabolic rate and causes lethargic behavior.

D. gigas exposes itself to hypoxia by deliberately entering the oxygen minimum layer (OML) during the day, and leaves the OML at night (Gilly, Markaida, et al., 2006; C. Nigmatullin et al., 2001; Stewart et al., 2013). *D. gigas* must therefore have a strategy to counter lethargy induced by hypoxia that enables individuals to leave the OML at the end of the day. By contrast, *D. gigas* may not possess a strategy to

7.6. Simulation and exploration of cannibalism by an energetic view

counter lethargy enforced by energy deprivation, so a lethargic individual may not be able to mobilize energy e.g., for defensive actions or predation.

The idea of passive cannibalism is based on the hypothesized ability of *D. gigas* to recognize energy-exhausted, lethargic conspecifics, for example by monitoring their reaction during encounters. Lethargic conspecifics exhibit suppressed defensive actions, or none at all, thus presenting themselves as helpless prey items. The cannibalism rate is known to increase in the vicinity of fishing operations, where cannibals prey on hooked and therefore helpless conspecifics (Alegre et al., 2014; Ibáñez et al., 2008; Ibáñez & Keyl, 2010). Based on this evidence, it is reasonable to assume that *D. gigas* may be able to spot conspecifics with impaired defense capability. However, in common with other cephalopod species (Smith & Reay, 1991), there is no evidence of adaptation to cannibalism by *D. gigas* through the development of anti-cannibalism strategies by potential victims; therefore cannibalism may simply arise as a “by-product” of generalized predation, rather than expressing the purposeful targeting of conspecifics.

Cadaver mode. The “cadaver mode” indicates terminal energy exhaustion that is the prerequisite for passive cannibalism. In the simulation model, a “cadaver” is an individual without metabolism. A cadaver is removed from simulation after three hours simulated time from the time that metabolism ceases, i.e., simulating a situation where the school moves on without any cannibal showing interest in the cadaver, or as soon as the cadaver contains $\leq 10\%$ of its original mass, when it assumed to become too small to be accessible by conspecifics.

Performing cannibalism. When its energy level falls below $threshold_{cann}$, an individual performs cannibalism and has full access to the school to locate suitable prey. Active cannibalism is performed by randomly choosing an individual whose size falls within the WOC. The cannibal’s energy level increases to $threshold_{max}$, or less if insufficient mass is consumed, while the victim becomes a cadaver or is removed from the simulation if wholly consumed.

Passive cannibalism is performed by randomly selecting an individual in cadaver mode regardless of the WOC. The cadaver/lethargic individual is (partially) consumed and the cannibals energy level raised up to $threshold_{max}$ or less if insufficient mass is provided.

7. Cannibalism as a possible survival strategy

7.6.2. Comparison of cannibalism strategies and the growth functions

Fertilized eggs hatch after three days (Staaf et al., 2011) and Staaf (2010) and Staaf et al. (2008) report no evidence of a significant variance in hatching time among eggs from the same batch. However, in the model, the parameter *start age variation* is used to simulate small differences in embryonic development that may well occur naturally.

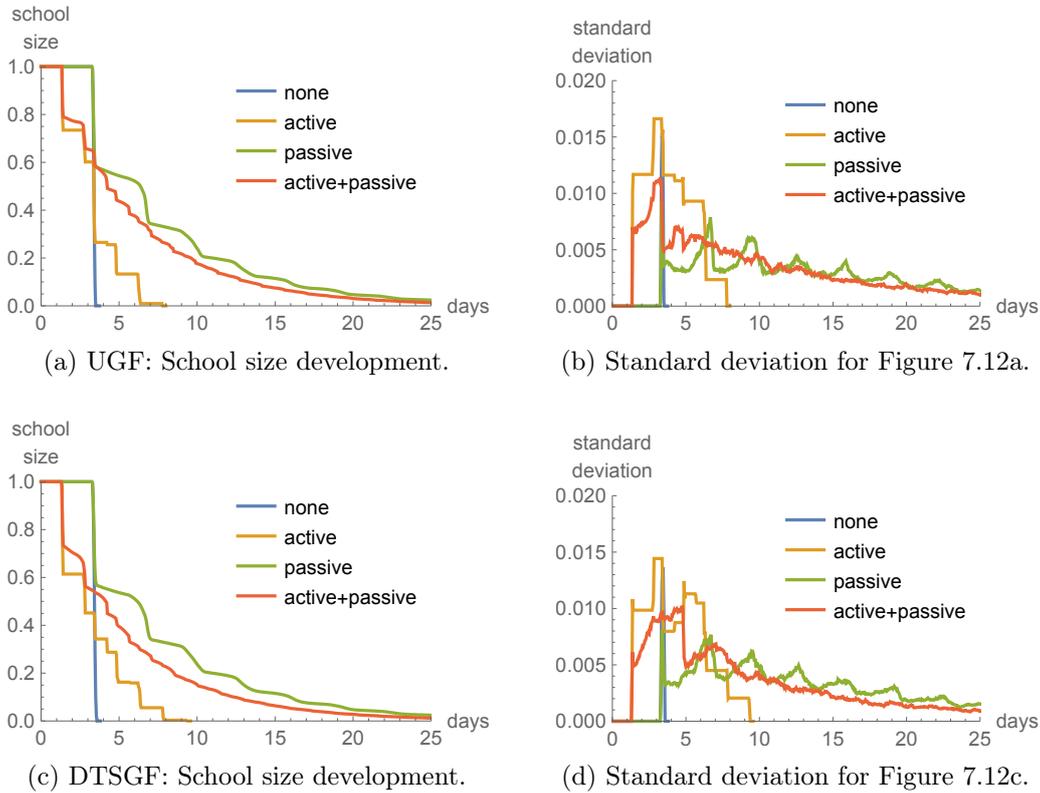


Figure 7.12.: Simulation results of school size development for different cannibalism strategies and growth functions under total food deprivation, averaged over 100 simulation runs. Simulation parameters: *start age* = 120 d, *start age variation* = 5 d.

Figure 7.12 shows the results of the different cannibalism strategies in combination with the different growth functions. The purpose of including the unisex growth function (UGF) is to reduce the number of variables for analysis, and specifically to explore the general effects of cannibalism in a population where the two sexes have identical growth parameters. In general, UGF and DTSGF perform similarly across all cannibalism strategies; results for the DTSGF reveal some additional dimorphic effects, as discussed below.

7.6. Simulation and exploration of cannibalism by an energetic view

In general, the small standard deviations over the simulation runs (right-hand panels in Figure 7.12) demonstrate the stable behavior of the cannibalism strategies in combination with both growth functions. The performance criterion used to evaluate the success of cannibalism as a survival strategy is the time period until the size of the school falls to a tenth of its original size. Using this criterion, the no-cannibalism strategy and active cannibalism strategy form one performance group, and passive cannibalism strategy and active+passive cannibalism strategy the other. The passive cannibalism and active+passive cannibalism strategies outperform the other group by maintaining the school size over 10% for a longer period of time. In the simulation, each strategy performs similarly for both the growth functions. In the following discussion, the notation “[growth function, cannibalism] strategy” specifies the combination of growth function and cannibalism strategy, e.g., [UGF, no-cannibalism] strategy.

In the simulation, in case of the [UGF, no-cannibalism] strategy, the school collapses within less than four days (≈ 3.67 d) due to the lack of energy uptake. The [UGF, active cannibalism] strategy only approximately doubles the time until the school collapses. Worse still, the school size is just above the criterion “10% of the original school size” after ≈ 5 days. Thus the [UGF, active cannibalism] strategy is only slightly more effective than no-cannibalism in maintaining the school size.

The [UGF, active+passive cannibalism] strategy preserves more than 10% school size over ≈ 13 days and the [UGF, passive cannibalism] strategy does so for over ≈ 16 days. These functions also preserve the school size at a higher level during the first days of energy deprivation. The following sections evaluate the different strategies in detail and compare their performance.

7.6.3. Analysis of the cannibalism strategies

The no-cannibalism and active cannibalism strategies

Figure 7.13 shows results for the UGF no-cannibalism and active cannibalism strategies, where day 0 (t_0) marks the start of energy deprivation for all individuals within the school, i.e., when energy uptake is set to zero. All individuals start with their maximum energy level at $threshold_{max}$; as time progresses the energy level of the individuals decreases rapidly (not shown) due to the high metabolic demands.

The first key effect of sustained energy deprivation in the simulation model is the terminal energy exhaustion event ($event_{tee}$) that is clearly depicted in the graphical representation of school size over time in the case of the [UGF, no-cannibalism] strategy. The energy level of individuals decreases from $threshold_{max}$ at t_0 to $threshold_{min}$ after

7. Cannibalism as a possible survival strategy

≈ 3.5 days ($t_{\approx 3.5}$), indicating that an individual at the age of 120 days starves to death under total energy deprivation after ≈ 3.5 days. School extermination follows shortly after $t_{\approx 3.5}$.

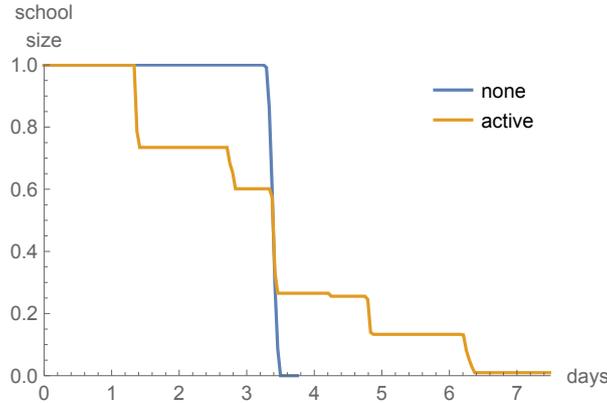


Figure 7.13.: UGF: Changes in school size development with no and active cannibalism, partial view of Figure 7.12a.

The behavior of the [UGF, active cannibalism] strategy differs from the [UGF, no-cannibalism] strategy as a result of the occurrence of $threshold_{cann}$ triggered cannibalism. The energy level of $\approx 37\%$ of individuals decreases from $threshold_{max}$ to $threshold_{cann}$ at $t_{\approx 1.33}$. This increase in the number of cannibalistic individuals is reflected in the reduction of school size at $t_{\approx 1.33}$ under the [UGF, active cannibalism] strategy, caused by the consumption of conspecifics, as shown in Figure 7.13. This decrease of school size marks the first cannibalism event ($event_{cann}$), which is the second key effect of energy deprivation in the simulation model.

In each $event_{cann}$ not all conspecifics fall within the WOC because of the ML distribution within the school. For the [UGF, active cannibalism] strategy, Figure 7.14 shows changes in the number of conspecifics available to cannibals (i.e., within the WOC) as a proportion of the number of cannibalistic individuals, and the proportions of cannibalistic individuals which do and not find victims. The distribution of ML among individuals in the schools and the constraints of the WOC lead to a shortage of victims, which prevents many cannibalistic individuals from performing cannibalism. This situation may change with increasing *start age variation*, as discussed in Subsection 7.6.4.

As shown in Figure 7.14, the proportion of unsuccessful cannibalistic individuals rises over the course of the simulation. This leads to a terminal energy exhaustion event ($event_{tee}$) in the [UGF, active cannibalism] strategy, when both the number of available victims and the number of successful cannibalistic individuals falls to zero (Figure 7.14).

7.6. Simulation and exploration of cannibalism by an energetic view

This occurs at $t_{\approx 3.5}$, which coincides with the event_{tee} in the [UGF, no-cannibalism] strategy.

In the simulation, event_{cann} at $t_{\approx 1.33}$ and event_{tee} at $t_{\approx 3.5}$ determine the change in school size over time in times of food deprivation, as shown in Figure 7.13. In the [UGF, active cannibalism] strategy, cannibalistic individuals that are unsuccessful in the first event_{cann} at $t_{\approx 1.33}$ die of total energy exhaustion (event_{tee}) at $t_{\approx 3.5}$, and cannibalistic individuals that are unsuccessful in the second event_{cann} at $t_{\approx 2.67}$ starve to death at $t_{\approx 4.83}$, an additional $t_{\approx 3.5}$ days after their last successful event_{cann} at $t_{\approx 1.33}$. The third cannibalism event at $t_{\approx 4.1}$ is marked by almost complete failure because by this time there are almost no suitable conspecifics available as prey, so the school becomes extinct $t_{\approx 3.5}$ days later at $t_{\approx 7.8}$ (Figure 7.12a). The times leading up to event_{cann} and event_{tee} increase slightly over the course of simulation time in line with an increase in the ratio of body mass to energy loss.

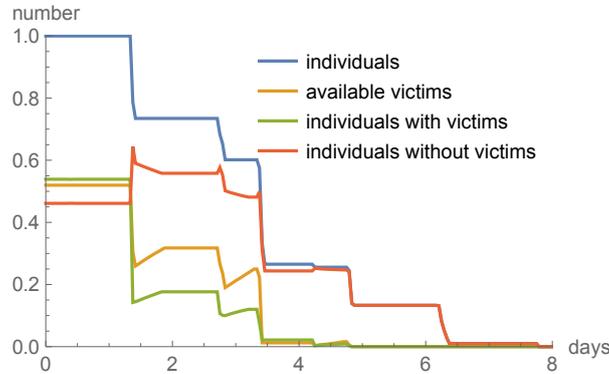


Figure 7.14.: UGF, active cannibalism strategy: Changes in numbers of individuals and conspecifics available for cannibalism in a school. Simulation parameters: *start age* = 120 d, *start age variation* = 5 d.

In case of the [UGF, active cannibalism] strategy, the small number of available victims (Figure 7.14) results in the “waste” of the majority of the smaller individuals. These individuals die due to total energy exhaustion, because they do not find conspecific prey within the WOC. As a population energy buffer, in the form of prey for larger individuals, they are lost before being consumed, because these individuals also do not fall within the WOC of the larger cannibalistic conspecifics. Therefore the [UGF, active cannibalism] strategy has little effect on school survival. At the point in time when the school becomes extinct under the [UGF, no-cannibalism] strategy, the school size has already fallen to only $\approx 27\%$ of its original size under the [UGF, active cannibalism] strategy. Overall, the [UGF, active cannibalism] strategy enables the school to survive

7. Cannibalism as a possible survival strategy

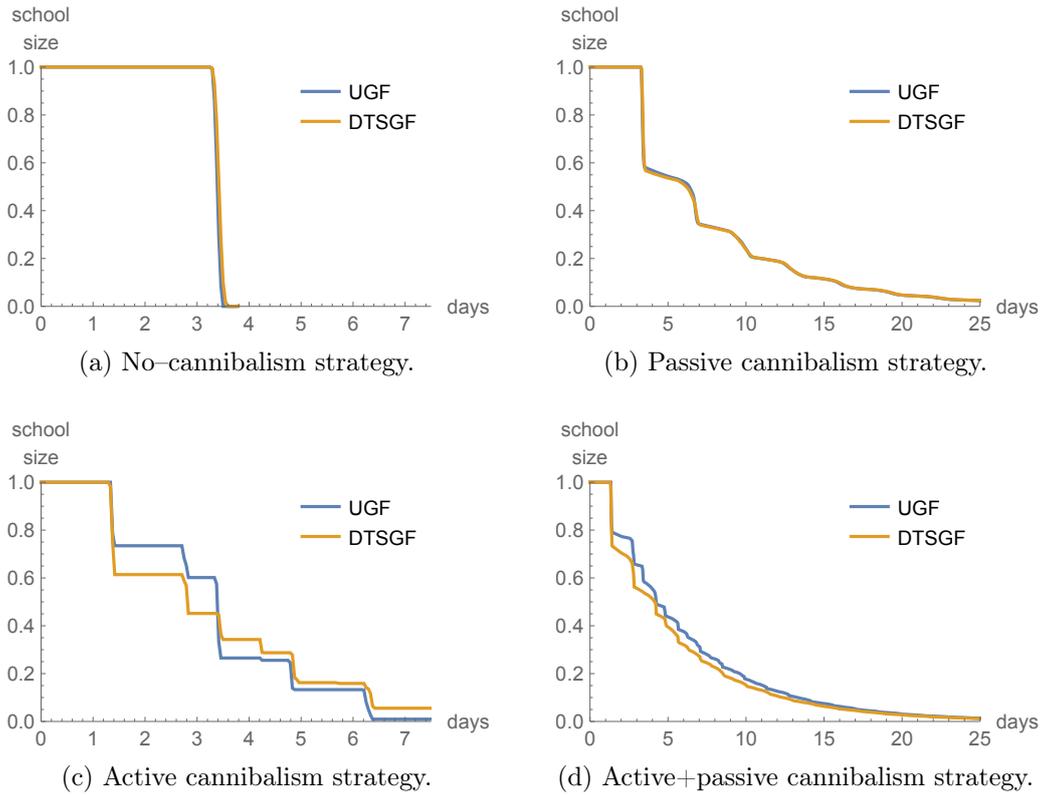


Figure 7.15.: Comparison between UGF and DTSGF growth models of simulated performance of cannibalism strategies shown in Figure 7.12

for about twice as long as [UGF, no-cannibalism] strategy, until at $t_{\approx 6.4}$ it becomes virtually extinct.

The UGF lacks a dimorphic growth pattern, so that the ML distribution of the individuals is less heterogenous than in the case of the DTSGF. The dimorphic growth of the DTSGF spreads the ML distribution and creates a more favorable environment for active cannibalism, i.e., active cannibalism strategy and active+passive cannibalism strategy, increasing the availability of suitably sized conspecifics. This more favorable environment allows a higher rate of cannibalism, especially in case of active cannibalism strategy, which is reflected in the relative improvement in performance of this strategy over time for the DTSGF, in comparison with the UGF (Figure 7.15c).

The passive cannibalism- and active+passive cannibalism strategies

The active cannibalism strategy represents an inefficient use of the population energy buffer and therefore may not perform well enough to ensure the survival of the school,

7.6. Simulation and exploration of cannibalism by an energetic view

or the wider population, during a period of energy deprivation. In contrast to the active cannibalism strategy, a cannibal in the passive cannibalism strategy consumes only energy-exhausted, therefore lethargic, or dead conspecifics.

In the field, different energy uptakes by individuals create a distribution of different energy levels within a school. This means that individuals are also affected by energy exhaustion due to energy deprivation at different times. The occurrence of energy exhaustion among individuals is thus distributed over time. Preying on energy-exhausted individuals resolves some issues associated with other cannibalism strategies, especially the active cannibalism strategy:

- A cannibal may not be able to completely consume a conspecific in a single feeding event, so it has either to hold on to the prey item, which might be difficult in the presence of other cannibalistic conspecifics, or abandon the partially consumed cadaver which represents wasted energy.
- A partly consumed lethargic or dead conspecific may provide enough food to enable several cannibals to restore their energy level up to $threshold_{max}$.
- A lethargic or dead prey item does not retaliate, so that the WOC_{UB} becomes insignificant as a constraint.
- Overall, the size distribution within a school becomes less important.

The beak of *D. gigas* enables the partial consumption of large prey items and thus provides the prerequisite for passive cannibalism. Stomach analyses indicate a high cannibalism rate for *D. gigas* (Keyl et al., 2008), but stomach analysis cannot distinguish between active and passive cannibalism or, therefore, rule out the occurrence of passive cannibalism.

Figure 7.12 (p.158) shows that the passive cannibalism strategy delays the initial reduction in school size until $t_{\approx 3.5}$ i.e., the point at which $event_{tee}$ occurs under the no-cannibalism strategy, for both growth functions. The exhausted conspecifics are most likely immediately utilized by cannibalistic conspecifics, because the energy level of many individuals, i.e., for the UGF 100% of the school at $t_{\approx 1.42}$, falls below $threshold_{cann}$ before any totally energy-exhausted individuals become available as prey for passive cannibalism.

In other words, no lethargic conspecifics are available when cannibalistic individuals first appear, but, over time, lethargic individuals are recruited from these cannibalistic individuals and are immediately consumed by their non-lethargic cannibalistic

7. Cannibalism as a possible survival strategy

conspecifics. Cannibalistic individuals that are close to terminal energy exhaustion will die unless they find some another conspecific to consume.

The process of waiting for energy-exhausted individuals to become available and then consuming them forms a negative cause-effect chain: an increased number of energetic exhausted individuals reduces the number of energy exhausting individuals. A successful cannibal fills its energy level up to $threshold_{max}$ or at least to $threshold_{cann}$, at which point the individual ceases to exhibit cannibalistic behavior, since the mass of a “fresh” lethargic or dead conspecific usually provides more energy than necessary. Therefore, partial consumption under passive cannibalism allows several cannibals to feed on the same prey and reduces the waste of population energy that occurs in the case of active cannibalism.

With a passive cannibalism strategy, waiting for energetically exhausted individuals exerts no additional pressure on the school size (unlike active cannibalism, which removes healthy individuals from the population), so the school size decreases more slowly compared with other cannibalism strategies. In case of the UGF, a passive cannibalism strategy enables almost 60% of the original school to survive beyond the time when the school becomes extinct in the case of no-cannibalism strategy ($t_{\approx 3.5}$) (Figure 7.12). This may be considered a worst-case scenario because the homogenous ML distribution and maximum energy level of each individual at the start of the simulation mean that all individuals become energy-exhausted at almost the same time. By contrast, in the field, a heterogenous energy level distribution is to be expected, resulting in a broader distribution of energetic exhaustion over time and thus a less marked expression of the key effects shown in Figure 7.12. A simulation with a heterogenous energy level distribution would not, however, provide new insights into the principle of operation of the passive cannibalism strategy.

Compared to the passive cannibalism strategy, the active+passive cannibalism strategy has the disadvantage of a premature reduction in school size caused by the “active” part of this combined strategy. The “passive” part prevails at the time of energetic exhaustion when the size distribution does not allow performance of active cannibalism.

The better performance of the passive cannibalism- and active+passive cannibalism strategy, compared to the active cannibalism strategy (Figure 7.15b and 7.15d), is also evident for the DTSGF. In the case of the passive cannibalism strategy a large proportion of the original school size is lost in the first event $_{cann}$, because all individuals are near the $threshold_{min}$ limit which is given by the factor 0.7 of the mass energy equivalent (Equation 7.7). Assuming each cannibals fills up its energy level to $threshold_{max}$ (Equation 7.5), the amount of energy required is

7.6. Simulation and exploration of cannibalism by an energetic view

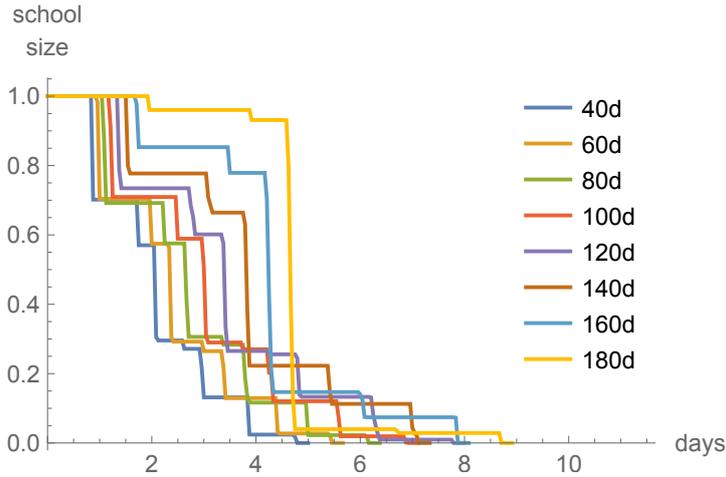


Figure 7.16.: UGF: Change in school size over time with active cannibalism averaged over 100 simulation runs with simulation parameters $start\ age = \{40, 60, \dots, 180\}$ d and $start\ age\ variation = 5$ d.

$(threshold_{max} - threshold_{min}) \cdot energy\ conversion\ factor$. Each average cannibal therefore needs about two-thirds of an average cadaver to satisfy its energy needs, which explains the initial rapid reduction in the school size.

In conclusion, the passive cannibalism strategy allows a school to maintain its size over a longer period of time in the event of food shortages compared to the other cannibalism strategies discussed.

7.6.4. Changing the start age and the age variation

Figure 7.16 shows the effects of a varied $start\ age$ on the active cannibalism strategy. The key effects appear as before but at different times, whereby an increasing $start\ age$ reduces the cannibalistic activity. With increasing $start\ age$, ML differences are reduced, meaning that fewer individuals within the WOC are available. As a result, the initial decrease of school size is less marked. In general, increasing the $start\ age$ reduces metabolic needs in relation to body mass, so that the onset of energy exhaustion is delayed (Subsection 6.3.1). However, an increasing $start\ age\ variation$ can be expected to increase the availability of conspecifics within the WOC and thus the availability of prey.

Passive cannibalism is largely unaffected by a changing $start\ age$, see Figure 7.17. The point of each $event_{cann}$ moves further along the time axis with increasing $start\ age$ due to the more favorable ratio of body mass to metabolic needs among older individuals,

7. Cannibalism as a possible survival strategy

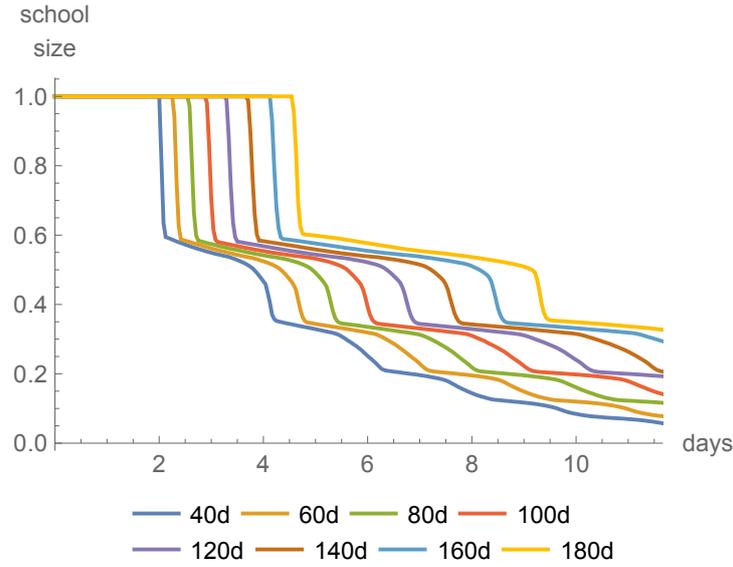
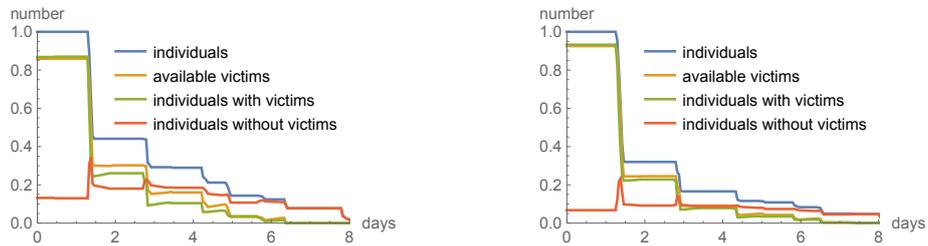


Figure 7.17.: UGF: Change in school size over time with passive cannibalism averaged over 100 simulation runs with simulation parameters $start\ age = \{40, 60, \dots, 180\}$ d and $start\ age\ variation = 5$ d.

which therefore benefit more from passive cannibalism strategy, compared to younger ones.



(a) Parameters: UGF, $start\ age = 120$ d, $start\ age\ variation = 10$ d. (b) Parameters: UGF, $start\ age = 120$ d, $start\ age\ variation = 15$ d.

Figure 7.18.: UGF: Changes in numbers of individuals and availability of conspecifics under an active cannibalism strategy in a school over time, with varying $start\ age\ variation$.

In case of active cannibalism, an increasing $start\ age\ variation$ creates larger ML differences within the school, so more conspecifics should be available for the cannibals than for small $start\ age\ variation$ values, see Figure 7.18 in comparison with Figure 7.14 p.161. However, as a corollary, the number of potential victims decreases more rapidly once cannibalism sets in, resulting in a more rapid decline in school size. In the short

7.6. Simulation and exploration of cannibalism by an energetic view

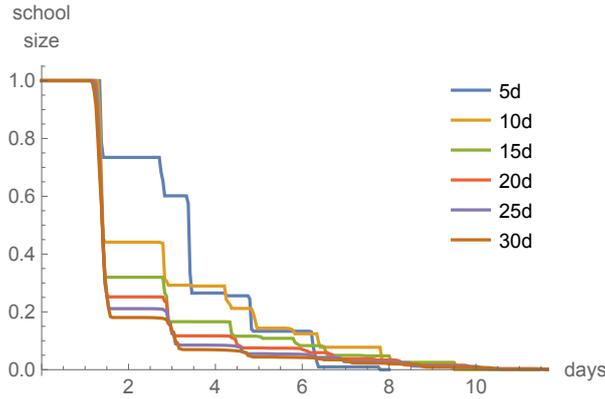


Figure 7.19.: UGF: Change in school size over time with active cannibalism and varying *start age variation*, averaged over 100 simulation runs. Simulation parameters: *start age* = 120 d and *start age variation* = {5, 10, 15, 20, 25, 30} d.

term active cannibalism performs best for *start age variation* = {5, 10} d, with survival of $\approx 30\%$ of the original school size at $t_{\approx 3.5}$. In the long term active cannibalism performs better with a *start age variation* = {10, 15} d, as shown in Figure 7.19. Energy exhaustion is greatly reduced for *start age variation* = 30 at $t_{\approx 3.5}$, because successful cannibals make up the majority of the school, but the school size is substantially reduced. Higher *start age variation* values generally result in less wastage of population energy compared to smaller values, but the trade-off is a more rapid reduction in the school size, to $\approx 20\%$ of original school size at the first cannibalism event at $t_{\approx 1.33}$. Figure 7.20 summarizes the performance of the different cannibalism strategies for a *start age variation* = {10, 15} d, where relative performance of the strategies is as before, with passive cannibalism the most effective strategy.

Additionally, in a cannibalism strategy that includes active cannibalism, a relatively dispersed ML distribution corresponding to high *start age variation* values promotes “hierarchical” cannibalism, whereby cannibals feed on already successful cannibals at the same event_{cann}, so that a cannibal may effectively feed on more than one conspecific at this time. The passive cannibalism strategy is less affected by changes in ML distribution because, in this case, “hierarchical” cannibalism may only occur when a successful cannibal becomes lethargic, which occurs some time after the current simulation step.

7. Cannibalism as a possible survival strategy

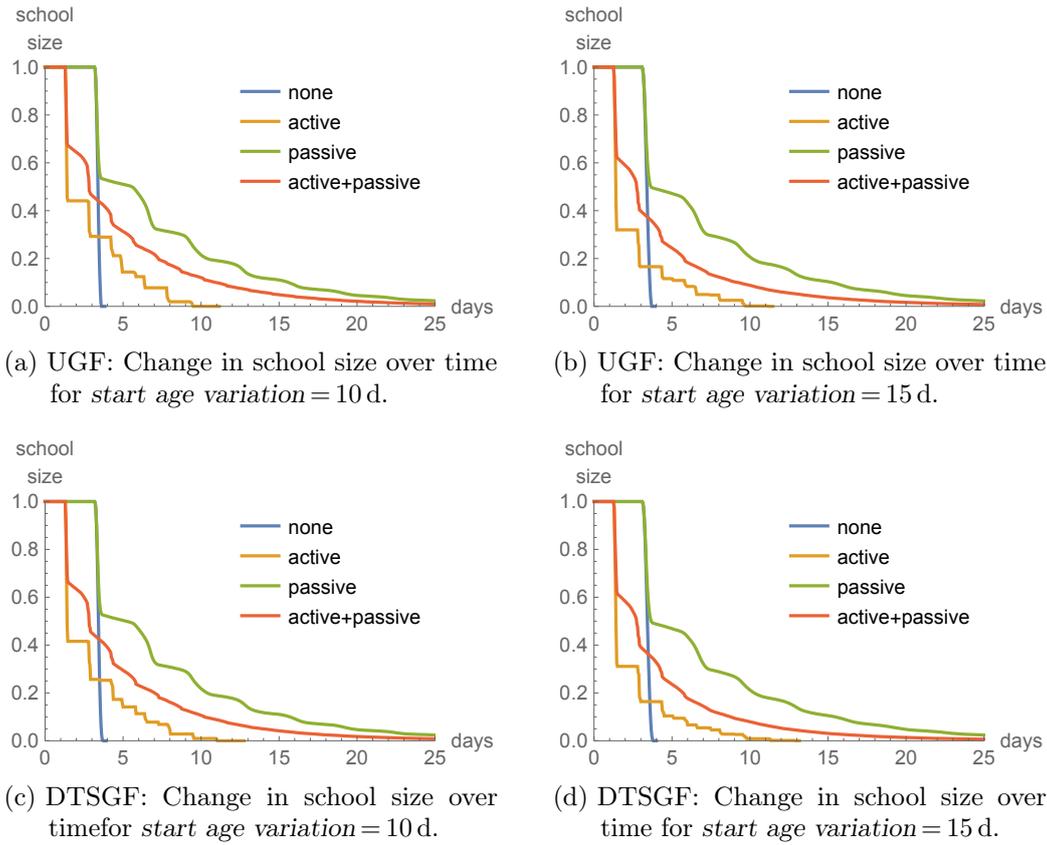


Figure 7.20.: Simulated changes over time in school size for different cannibalism strategies and growth functions under total food deprivation, averaged over 100 simulation runs. Simulation parameters: *start age* = 120 d, *start age variation* = {10, 15} d.

7.6.5. Varying the cannibalism energy threshold

For an individual, a lower $threshold_{cann}$ value delays the triggering of cannibalism and the search for consumable conspecifics until it is closer to its own terminal energy exhaustion ($threshold_{min}$). In contrast, increasing $threshold_{cann}$ values shift the cannibalism performing towards a default behavior and thus towards the predation of arbitrary prey items. This section discusses the effects of varying the $threshold_{cann}$ value within a $[threshold_{min}, threshold_{max}]$ interval of $[0.7, 1.1]$ on performance the active- and passive cannibalism-strategies.

For the active cannibalism strategy, a $threshold_{cann} = threshold_{min} = 0.7$ delays the triggering of cannibalism to the point of terminal energetic exhaustion with the consequent death of all individuals (Figure 7.21). Thus in this case the performance of active

7.6. Simulation and exploration of cannibalism by an energetic view

cannibalism strategy is equal to that of no-cannibalism strategy (compare Figure 7.21 to Figure 7.13). Overall, low $threshold_{cann}$ values delay the onset of cannibalism and thereby increase the risk of terminal energy exhaustion of a cannibal.

In contrast, a $threshold_{cann} = threshold_{max} = 1.1$, where each individual becomes cannibalistic if its energy level falls below the maximum energy level, cannibals will only attack and kill conspecifics which fall into the WOC but consume very little of them, because they require only a small amount of food to restore their energy levels to $threshold_{max}$. Therefore only a few individuals survive for a short time beyond the point of $event_{tee}$ at $t \approx 3.5$.

Increasing $threshold_{cann}$ values in case of the active cannibalism strategy trigger earlier cannibalism events, which delays terminal collapse of the school but with a more rapid initial reduction in school size. The $threshold_{cann} = 0.8$ value performs best on the school size criterion because it maintains the school size above 10% for the longest period time, but the school becomes extinct by the same time (approx. 6 than approx. 7 days) as for other $threshold_{cann}$ values. A $threshold_{cann} = 1.0$ shows a small but steady decrease of school size and maintains a small school size for the longest period of time.

Parameter	Value	Description
start age	120	The interval of chosen <i>start age</i> values.
start age variation	5	The interval of chosen <i>start age variation</i> values.
$threshold_{cann}$	{0.7, 0.8, 0.9, 1.0, 1.1}	The interval of chosen $threshold_{cann}$ values.

Table 7.1.: Specified simulation parameters.

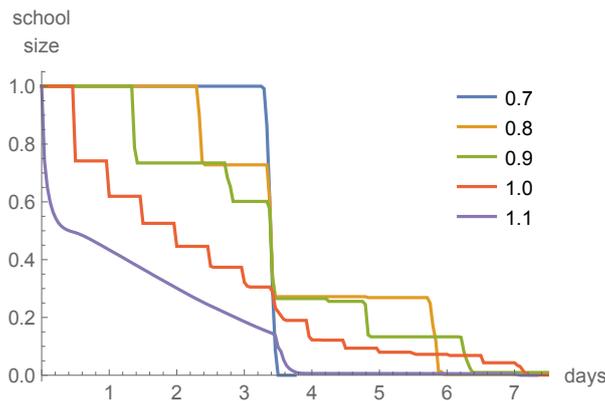


Figure 7.21.: UGF: Averaged changes in school size over time for active cannibalism with varied $threshold_{cann}$ (Table 7.1).

7. Cannibalism as a possible survival strategy

The reduction of the $threshold_{cann}$ values in the case of passive cannibalism reduces the “interest” in lethargic conspecifics and, theoretically, increases the probability of wasting the population energy, because there may be insufficient cannibals, due the low $threshold_{cann}$, to consume the already lethargic conspecifics. However, Figure 7.22 shows no significant change for different threshold values. In other words, there is no “oversupply” of lethargic individuals. Cannibalistic activity operates in a stable manner over a range of threshold values, controlled by the availability of energetic exhausted (lethargic) individuals.

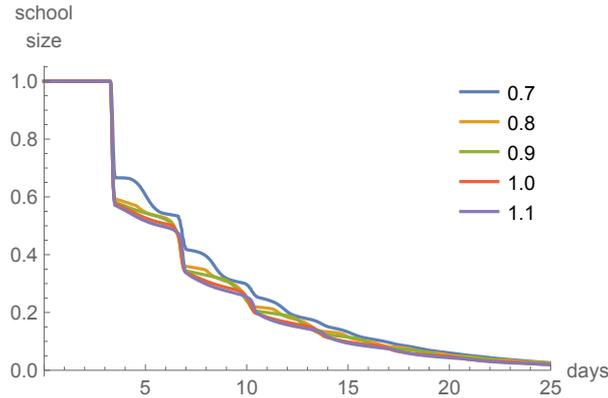


Figure 7.22.: UGF: Changes in school size over time for passive cannibalism with varied $threshold_{cann}$ (Table 7.1).

In the case of passive cannibalism, when $threshold_{cann}$ equals $threshold_{min}$, the passive cannibalism strategy would be expected to behave like the no-cannibalism strategy. However, in the sequential processing of actions in the simulation, a cannibalistic individual first looks for a lethargic conspecific and consumes it if available, then the individual’s current energy level is computed. As an effect, an individual that is, technically, already lethargic (i.e., with an energy level below $threshold_{min}$) may consume another lethargic individual whose lethargic state was set earlier in the same simulation step.

7.6.6. The validity of the model

The energy-centered model is likely to be quantitatively inaccurate, as values of basic parameters, such as the energy demand for locomotion or energy saving through metabolism suppression, are still open to discussion. However, these quantitative approximations should not affect the *qualitative* interpretation of the performance of the different cannibalism strategies, since effect of the energy needs on the results is

7.6. Simulation and exploration of cannibalism by an energetic view

clearly defined. Higher energy needs shift events to the left on the time axis, i.e., the effects appear faster. Lower energy needs shift the events to the right on the time axis, as shown in Figure 7.17, where the increasing *start age* and thus larger size provide the individuals with a lower daily energy need relative to body mass and therefore more time elapses before the occurrence of the event_{tee}, which shifts to the right on the x-axis.

Ibáñez and Keyl (2010) mention a hypothetical buffering capacity in the digestive gland that would as an effect increase an individual's energy buffer capacity through the metabolism of lipids. An increased buffering capacity of this kind would likewise affect the timing of terminal energetic exhaustion and/or the onset of cannibalism, moving these events to the right on the time axis.

D. gigas utilizes metabolism suppression (Trübenbach, Pegado, et al., 2013; Trübenbach, Teixeira, et al., 2013) during its stay in the oxygen minimum layer (OML), which in turn reduces the energy needs and shifts the key events to the right on the time axis. The low oxygen saturation of the OML supports only low metabolism rates, so that a hypothetical alternative oxygen source, e.g., an internal oxygen depot, would have little or no effect on energy needs and would cause little or no movement of key events to the right on the time axis. Alternatively, an “oxygen debt” may be created and/or an anaerobic metabolism utilized. An oxygen debt must be compensated for and would not affect the amount of energetic needs nor the time schedule of the key events.

Evaluation of the modeled performance of active cannibalism strategy in response to changes in the parameter *start age variation* shows optimum performance, in terms of school size decrease and school survival for *start age variation* = {10, 15} d (see Figure 7.19). These values also enable an optimal usage of the population energetic buffer, but this requires (1) that hatching events occur over a period of 20d to 30d, and (2) formation of schools during this time period. However eggs and hatchlings are exposed to variable currents that may disperse offspring spawned at the same location. Smaller values, such as the original *start age variation* = 5 d, are therefore more feasible than the optimal *start age variation* = {10, 15}. Figure 7.16 indicates optimal *start age variation* values for different *start age* values (i.e., age at the onset of energy deprivation) for active cannibalism. Higher *start age* values require higher *start age variation* values, a requirement which is unlikely to be fulfilled.

The parameter *start age variation* is used to simulate the ML distribution within a school, which is caused by differences in age or by variable growth due to limited food uptake. The model design does not differentiate between the different causes of ML

7. *Cannibalism as a possible survival strategy*

variation; effects are modeled based on the current size of individuals and the associated available growth potential.

Therefore, different levels of food intake can be represented by choosing an appropriate *start age variation*, where higher values indicate that food was less uniformly distributed over a longer period. The effects considered so far show no significant changes in their qualitative features between small and high values of *start age variation*.

In summary, all the parameters considered affect the model behavior linearly by shifting the key effects $\text{event}_{\text{tee}}$ and $\text{event}_{\text{cann}}$ on the time axis but the relative performance of the different cannibalism strategies is not affected.

7.7. **Survival of the females**

Because males are considered to be more energetically expensive than the females but contribute less to the reproductive process, a changing sex ratio in favor of the females (Subsection 3.4.3 and Section 8.5) is considered an adaptation to low-energy environmental conditions. Therefore, the effects of a cannibalism strategy triggered by low energy uptake in low-energy environmental conditions should be compatible with the observed sex ratio change. Cannibalism might be “the process” that adjusts the sex ratio in response to changing environmental conditions or “an additional process” that contributes to this change. Alternatively, the sex ratio may be determined during or shortly after spawning or by another selection process that favors the females. In any event, the cannibalism strategy should not affect an existing sex ratio by disadvantaging females in times of food deprivation.

Figure 7.23 and 7.24 show the changes over time, computed in 100 simulation runs per parameter set, in the proportion of females in the school for different start ages and growth functions during total food deprivation. In the case of no-cannibalism strategy, the survival of an individual depends on the energy buffering capability of the body mass. Larger individuals have higher buffering capability, therefore the growth function is the decisive factor, see Subsection 6.3.1. The simulation results for sex ratios when school sizes become small as simulation time progresses are not very informative, since random elements in the original population configuration, i.e., small differences in the absolute number of females in the school sizes, cause large variations. This “chaotic” behavior is evident in the high values for the standard deviation, e.g., in Figure 7.24d for the passive cannibalism, where standard deviation increases steeply at approx. day 13 when the school size is reduced to approx. 7% of its original size, see Figure 7.12c.

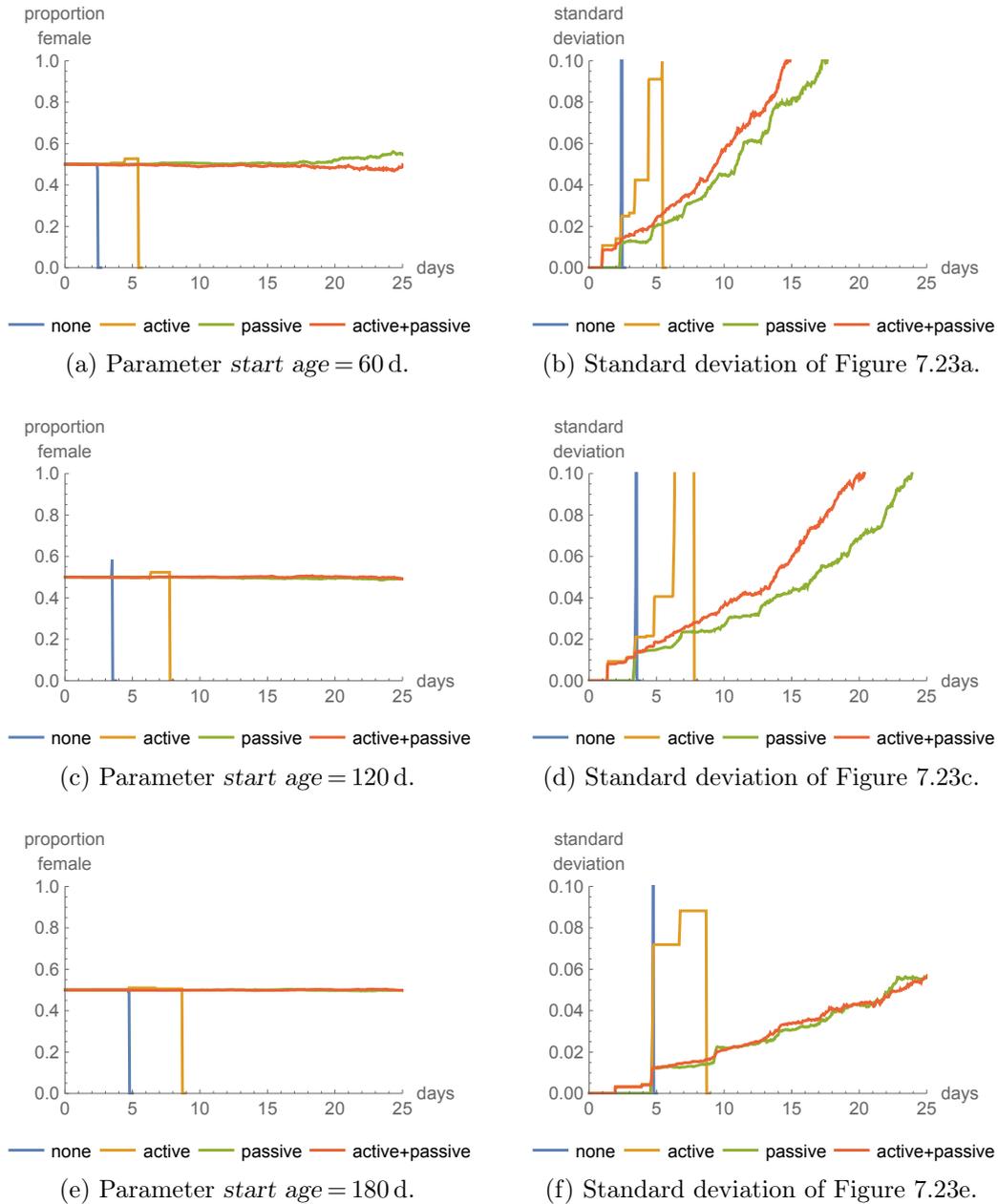


Figure 7.23.: UGF: Changes in the proportion of females over time for different cannibalism strategies. Simulation parameters: *start age* = {60, 120, 180} d, *start age variation* = 5 d.

The growth function is also the determining factor in the case of the active cannibalism strategy, because the current ML determines the ability to perform cannibalism by setting the bounds of the WOC. Cannibalism is performed by the larger individuals of a

7. Cannibalism as a possible survival strategy

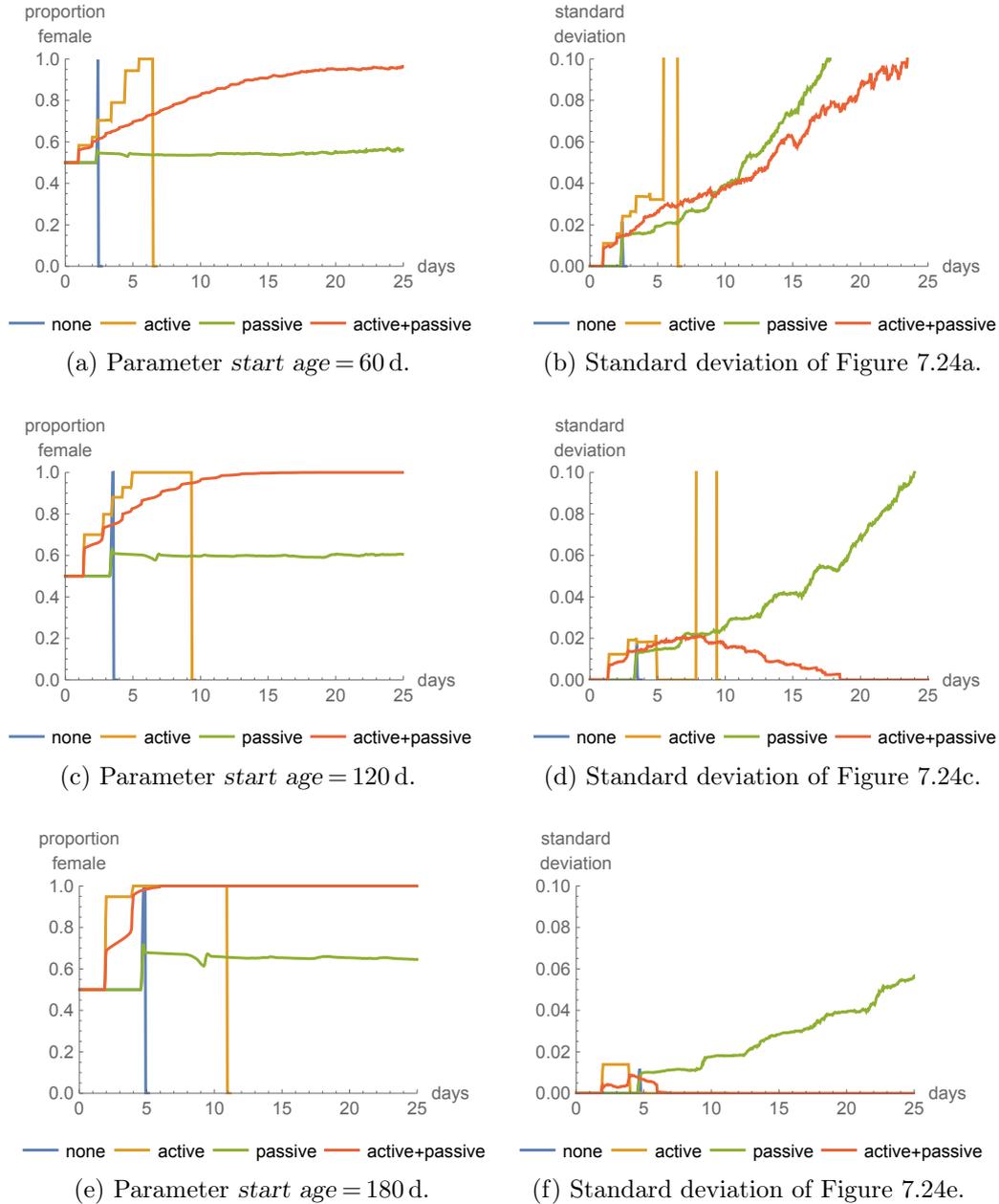


Figure 7.24.: DTSGF: Changes in the proportion of females over time for different cannibalism strategies. Simulation parameters: $start\ age = \{60, 120, 180\}\ d$, $start\ age\ variation = 5\ d$.

school who already have a higher buffering capability due to their body mass. For active cannibalism strategy, the UGF advantages neither males nor females, but the DTSGF favors the females, since they are always larger than the males. Indeed, Figure 7.23

7.8. Discussion of an obsolete window of cannibalism

indicates that, for the active cannibalism strategy, the population is composed entirely of females before it collapses. Moreover, the ML difference between females and males increases faster with increasing age, therefore the proportion of females in the population increases more rapidly as the starting age increases (compare Figure 7.23a with 7.23e).

In case of the UGF, passive cannibalism- and active+passive cannibalism strategy, do not affect the proportion of females in the school, except to the extent that the results are favorable to females compared to those for the no-cannibalism- and active cannibalism strategy. The sex ratio does not change appreciably except when school sizes become small towards the end of the simulation.

The DTSGF slightly favors females in combination with the passive cannibalism strategy and much more so in combination with the active+passive cannibalism strategy, which leads to a large increase in the proportion of females. The active part of the active+passive cannibalism strategy benefits females because they are larger than the males throughout the lifespan, thus female predation reduces the number of the males. This effect is more pronounced in higher ages, as size difference becomes more pronounced and more males fit into the WOC of the females. Their larger ML also means that females have a higher buffering capability than males, which represents a further advantage.

The DTSGF in combination with a active+passive cannibalism strategy appears to provide the optimal scenario for a sex ratio change in favor of females in response ongoing food deprivation or multiple food deprivation events caused by low environmental productivity, as postulated by Tafur et al. (2010).

7.8. Discussion of an obsolete window of cannibalism

The “strategic usefulness” of an energetic threshold that triggers cannibalism is debatable, since the performance of both active and passive cannibalism strategies in a “default” cannibalism mode, i.e., where $threshold_{cann} = threshold_{max}$, does not differ significantly to performance when cannibalism is triggered by a threshold, i.e., $threshold_{cann} < threshold_{max}$. Overall, the passive cannibalism strategy performed best and this suggests that consideration should be given to the possibility that passive cannibalism is in fact the default feeding strategy.

The reduced defensive activity of a lethargic conspecific renders it vulnerable (Ibarra-García et al., 2014) to *D. gigas*’s voracious and opportunistic feeding behavior and may trigger cannibalistic behavior under passive cannibalism even when other food sources are available. The WOC of *D. gigas* is based on observation and might actually be a

7. Cannibalism as a possible survival strategy

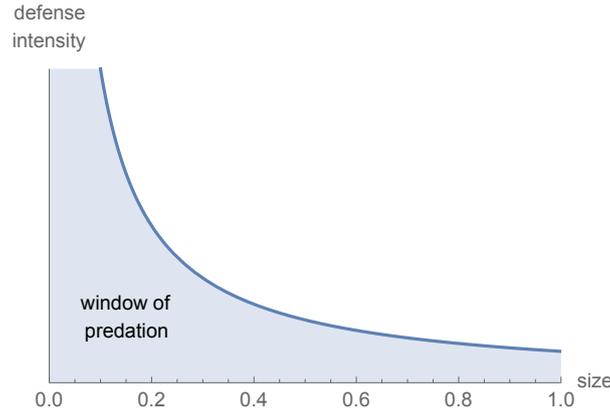


Figure 7.25.: Two-dimensional window of predation (WOP²): The diagram illustrates the concept of a two-dimensional WOP as a function of prey item's relative size and intensity of its defensive response. The shaded area marks the WOP.

WOP whereby the upper bound for cannibalism increases for lethargic or otherwise vulnerable prey items. In this case, the WOP would be based on a two-dimensional function (Figure 7.25), i.e., would depend on both the size and defensive activity of the prey item.

A two-dimensional window of predation (WOP²) of this kind would account for the ranges of the observed regular WOP and WOC, thereby further simplifying the feeding strategy of *D. gigas* to a simple opportunistic feeding strategy that is compatible with all kinds of cannibalism. The functional behavior of *D. gigas* might be based on an asymptotic function, since the body mass increases in a cubic relation to increasing size (i.e., ML) and physical strength is proportional to mass.

The function may be supplemented by a third parameter *energy level*, which takes into account the extent of the energy deficit, i.e., hunger. Low energy levels may promote cannibalism by increasing the urge to perform cannibalism even when the prey displays more intensive defensive activity and there is the risk of retaliation. However, the additional parameter might not be required by the passive cannibalism strategy, which is successful without this parameter.

7.9. Discussion of the results and conclusions

Cannibalism potentially reduces the fitness of a population because fewer individuals contribute to the next generation, which is why there must be other benefits at the population level. Such long-term benefits could include: partial school survival;

maintenance of female size-weight coupled fecundity, thus higher fitness; and access to resources in low oxygen saturation/low energy areas through the consumption of smaller individuals that feed on other species in these areas.

The high metabolism rates and short generation times of the cephalopods require fast and continuous energy uptake, otherwise the immediate survival of the individuals is endangered. Individual cannibals benefit primarily from energy uptake at a time when other food is scarce (Smith & Reay, 1991), which thereby ensures their immediate survival. During the pre-reproduction period, cannibalism supports the survival of the individual until maturity and the onset of the reproductive season. At the population level, cannibalism may help stabilize the population by providing “an energy loop that maintains calories within population” (Polis, 1981). In effect, this increases the carrying capacity of the ecosystem through an indirect extension of the food size spectrum, whereby larger individuals gain access to lower trophic levels by feeding on smaller conspecifics (Ibáñez & Keyl, 2010). This reduces competition for food within the population, but at the cost of a reduction in population size (Polis, 1981).

The hatching offset ω is a useful parameter for analysing the effects of cannibalism. When individuals experience the same exogenous factors and follow the same growth function, their different ages correspond to individuals of different sizes. An increasing hatching offset ω increases cannibalism capability of older individuals by lowering the $\text{reqWOC}_{\text{UB}}$, i.e., reducing the size of potential prey relative to the size of the cannibal, for the ω -older individual, but the nature and magnitude of the effect depends on the growth functions, the sex and the age of the individuals.

The DTSGF keeps the intra-sex $\text{reqWOC}_{\text{UB}}$ values at a high level throughout the whole lifespan and thus making intra-sex cannibalism more difficult, since there is no pronounced size advantage for the cannibal. But for inter-sex cannibalism, the DTSGF reduces the $\text{reqWOC}_{\text{UB}}$ for females to cannibalize the males to levels that would result in a change to the sex ratio in favor of females during periods of energy deprivation, which is consistent with observations reported by Tafur et al. (2010).

The active cannibalism strategy triggers the consumption of conspecifics soon after the onset of food deprivation. This “greedy energy harvesting strategy” enables individuals to satisfy their energy needs, but the trade-off is a rapid reduction of school size. Compared to the no-cannibalism strategy, under the modeled conditions of complete food deprivation, the active cannibalism strategy averts the extinction of the school in the first population collapse, stabilizing the school size at a low population level, but only for a short time as school extinction occurs shortly thereafter.

7. Cannibalism as a possible survival strategy

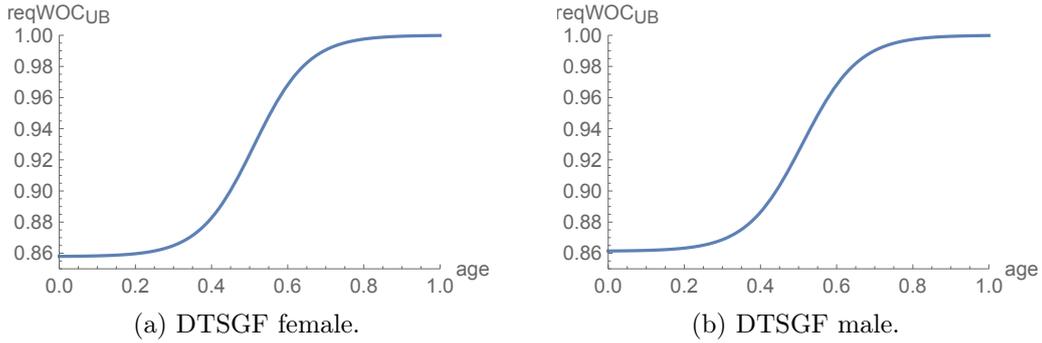


Figure 7.26.: DTSGF: Changes in the value of $\text{reqWOC}_{\text{UB}}$ for (a) female on female and (b) male on male cannibalism, for an age difference $\omega = 4$ d between predator and prey. These values of $\text{reqWOC}_{\text{UB}}$ approximate to those for cannibalism of an unsuccessful cannibal by a successful one, where failure to feed by the unsuccessful cannibalism leads to a growing size difference, equivalent to an age difference $\omega = 0$ d to 3.5 d, after which the unsuccessful cannibal dies of energy exhaustion.

Furthermore, the availability sufficient food after the onset of cannibalism (since cannibalism is triggered before the onset of total energy exhaustion), makes the active cannibalism strategy a clearly suboptimal strategy. The longer the period between the triggering of cannibalism and total energy exhaustion, the more probable it is that other food resources will become available during the intervening period, making cannibalism “unnecessary” to ensure school survival. The negative effects of this “strategic” failure of the active cannibalism strategy are exacerbated in a situation where several short food shortages occur one after the the other, each of which will trigger the “premature cannibalism”, provided that the ML distribution of the school supports cannibalism. The result is a premature and unnecessary reduction of school size and unfavorable changes in the ML distribution within the school. Survival of the school over a longer period of food deprivation is not achieved by by the active cannibalism strategy because there are an insufficient number of conspecifics within the WOC to serve as prey. The lack of conspecifics within the WOC means that a large proportion of the population energy buffer is inaccessible to the cannibalistic individuals. In conclusion, the constraining effect of the WOC is the decisive limiting factor for the effectiveness of active cannibalism.

An optimal *start age variation* value probably exists for the active cannibalism strategy, i.e., one that optimizes the ML distribution within the school to minimize the energy exhaustion of individuals, school size reduction and the waste of population energy. However, counterproductive behavior of active cannibals, and the negative effects of

the active cannibalism strategy remain unchanged. Moreover a high *start age variation* may result in “hierarchical” cannibalism that removes more than one conspecific per cannibal in a cannibalism event. To compensate for inherent adverse effects of the active cannibalism strategy, a ML distribution that is conducive to continued cannibalism has to be maintained. This could be achieved through:

- Immigration of smaller individuals or schools with smaller individuals. The latter case requires the existence of nearby schools with a different ML distribution, comprised of individuals that are not energetically exhausted. In addition, nearby schools with different energy level require locally very different food levels, otherwise similar energy uptake takes place which will result in similar energy levels and thus ML distribution in these schools. Nearby schools with similar ML distributions would not meet the cannibals’ need for small individuals to prey upon.

Moreover, smaller individuals may be unlikely to join a school of larger conspecifics, since they will probably recognize larger conspecifics as potential opportunistic predators and therefore avoid them as a defense mechanism against cannibalism (see Smith and Reay (1991)).

- The growth of cannibals to the extent that restores the ML distribution to one that supports ongoing active cannibalism strategy. Successful cannibals may gain a growth advantage during periods of food deprivation, thereby moving more smaller conspecifics into the WOC. This effect is equivalent to increasing the hatching offset (Section 7.3), but in reality the growth that is possible in the period of a few days between a cannibalism event and the need for further energy input is probably not sufficient to shift a sufficient number of conspecifics into the WOC. Figure 7.26 shows the changes in $\text{reqWOC}_{\text{UB}}$ for an age difference $\omega = 4 \text{ d}$ between predator and prey, which is larger than maximum size difference between a successful and an unsuccessful cannibal (equivalent to approx. 3.5d) before the unsuccessful cannibal dies of exhaustion. Therefore, unsuccessful cannibals will die of starvation in a short time and thereby become unavailable as victims for the growing cannibals.

It seems difficult to maintain a ML distribution that supports ongoing active cannibalism. Even if the ML distribution increases within a school and differences amplify as a result of cannibalism, i.e., differential food intake (see Hölker and Breckling (2005)), the change in size distribution is insufficient to support active cannibalism, as shown by the analysis of corresponding increased *start age variation* at the start of the simulation.

7. *Cannibalism as a possible survival strategy*

In contrast, the passive cannibalism strategy avoids the weaknesses inherent in the operating principle of active cannibalism and may be an advantageous strategy. First, passive cannibalism does not depend on a ML size distribution but only on the presence, within the school, of lethargic conspecifics. Second, lethargic conspecifics appear during periods of food deprivation, therefore offer a means of maintaining the maximum possible size of the school within the constraints set by the energy requirement during these critical periods. In case of the passive cannibalism strategy, lethargic individuals act as a “ready-made” food resource for non-lethargic cannibalistic conspecifics that enables them to avoid becoming lethargic. But if these cannibalistic conspecifics fail to perform cannibalism, then they themselves become lethargic if total food deprivation continues, thereby becoming available as a food source to other cannibalistic individuals. Thus a negative cause–effect chain emerges, where an increasing number of energy–exhausted (lethargic) individuals provide energy uptake to cannibalistic conspecifics thereby reducing the number of energy–exhausted individuals. This principle of operation continues during total energy deprivation.

A passive cannibalism strategy delays the onset of cannibalism until the appearance of energy–exhausted individuals, i.e., delays cannibalism until it is unavoidable because alternative food resources are missing. In contrast to the active cannibalism strategy, the passive cannibalism strategy is a *reactive* strategy. In addition, the incidence of “hierarchical” cannibalism is reduced because a successful cannibal will not exhibit lethargic behavior for some time afterwards, i.e., not during the same cannibalism event in the model, and will therefore be unavailable as prey. Of course, if food deprivation continues, a successful cannibal might itself be consumed at a later time.

The passive cannibalism strategy will not work if all individuals of a school have exactly the same ML, energy level and metabolism needs at the onset of energy deprivation. In this unlikely case, all individuals behave in exactly the same way and all will die of terminal exhaustion at the same moment. In reality, differences in energy levels are likely to exist within a school, whereby individuals are affected by energy exhaustion and become lethargic at different times; the passive cannibalism strategy exploits these differences. In summary, the passive cannibalism strategy is a resilient strategy that utilizes exhausted individuals very efficiently to stabilize school size over a longer period, by minimizing the number of sacrificed individuals.

In case of the active+passive cannibalism strategy, active cannibalism is triggered when $threshold_{cann} > threshold_{min}$, with all the adverse effects as discussed above. The passive cannibalism strategy part prevails in the case of absence of conspecifics within the WOC, and in this case the effects are the same as for a passive cannibalism strategy,

except possibly for a possible different ML distribution, since smaller individuals were removed in previous active cannibalism events. The active+passive cannibalism strategy has no advantage over the passive cannibalism strategy in terms of the school size and school survival criteria.

In summary, the passive cannibalism strategy performs best. However, a default passive cannibalism strategy is not distinguishable from an opportunistic feeding strategy whereby (lethargic) conspecifics are part of the regular diet. Crossland and Shine (2011) refers to a generalized predators as *passive cannibals* “[...] that treat conspecifics in the same way as they treat heterospecific prey” and thus not specifically targeting their conspecifics (“targeting cannibal”).

From this perspective, the two-dimensional window of predation (WOP²) is suggested, combining the prey size and the intensity of defensive activity, as a conceptual framework for effective passive cannibalism strategy and to avoid the restrictions imposed by the WOC on the substantially less effective active cannibalism strategy. Thus the WOP² combines the WOP and WOC and explains the opportunistic feeding strategy of *D. gigas* and every kind and rate of cannibalism.

The WOP² includes cannibalism as normal behavior within a feeding strategy that takes advantage of all available energetic resources. For a predator, this provides access to additional energy sources, i.e., through the consumption of larger prey items, while reducing the risk of retaliation. The idea that cannibalism is simply a means of accessing a high-quality food resource aligns with the lifestyle of *D. gigas*, which combines fast growth and high food uptake. Thus cannibalism is considered regular behavior and conspecific prey is part of a regular diet (Polis, 1981; Smith & Reay, 1991).

The WOP² feeding strategy substitutes for alternative explanations of cannibalism as a separate feeding strategy to support the survival of populations in times of food scarcity. The analysis in this chapter does not provide support for the idea of cannibalism as an evolutionary stable strategy in low energy environment as suggested by Ibáñez and Keyl (2010). Rather it seems likely that, in the case of *D. gigas*, a WOP² feeding strategy is a prerequisite for survival under the constraints of a high rate of metabolism and a small energy buffering capability. Apparently this feeding strategy was so successful that the species did not evolve an additional individual buffering capacity, since it could rely on its own populations as an effective population-level energy buffer.

8. Properties at the population level and their computation

8.1. Introduction

This chapter discusses energetic requirements of *D. gigas* at higher than individual level (Chapter 6), i.e., the effects on energy use resulting from the interaction of individuals at school, cohort and generation level. These interactions have already been discussed in relation to cannibalism; in this chapter the focus is on other drivers of energy use identified in Section 3.4.

This chapter applies the dimorphic terminal size growth function (DTSGF) to test the different mechanism that have been suggested for optimal use of available energy in the environment by *D. gigas* at the population level. Section 8.2 discusses the asynchronous formation of male and female reproductive tissue as a means of reducing peak energy demands on the ecosystem. The advantages and disadvantages of small size-at-maturity (SAM) as an adaptation to environments providing only limited amount of food resources (“low-energy environments”) are discussed in Section 8.3.

Section 8.4 considers the population-level effects of the distribution of spawning events over time. Section 8.5 examines whether changing the sex ratio to favor females increases population fitness in low energy environments. Section 8.6 summarizes the results.

8.2. Reducing energy peak demand by asynchronous formation of female and male reproductive tissue

Hypothesis. The process of maturation requires additional energy for development of reproductive organs, over and above that required for metabolism, growth and locomotion. In the case of *D. gigas*, the peak in energy demand during maturation is flattened at the population level by the time offset between female and male maturation: Males mature earlier and store spermatophores, which later are used to fertilize the eggs spawned by females (Chen et al., 2013; Chen et al., 2008; Field et al., 2013; Tafur et al., 2010).

The offset in the onset of maturity may thus be seen as a strategy to optimize survival at the population level, since simultaneous maturation of males and females might

8. Properties at the population level and their computation

require more energy than is available in the environment. Such a strategy would reduce intra-species competition for resources and thereby support the survival of females until spawning, as well as increasing population level fitness, since better-nourished females are expected to produce more eggs (Ibáñez & Keyl, 2010).

These hypothesized effects align with observations that the number of males in a cohort progressively decreases in the period following mating (Markaida & Sosa-Nishizaki, 2001; Tafur et al., 2010; Tafur et al., 2001). A increased female:male ratio following mating is reported by (Bazzino et al., 2007; Chen et al., 2008; Ichii et al., 2002), but not by Markaida and Sosa-Nishizaki (2003) who found that the sex ratio did not change. The removal of successfully mated males from a population is advantageous because it reduces the overall energy demand of the population, thereby reducing intra-species competition, flattening peaks in energy demand and supporting the survival of females until spawning. These and other effects of a changing sex ratio are explored in Section 8.5.

Tafur et al. (2010) note that additional energy is required during maturation for the production of gonadic tissue and discusses three options for increasing energy uptake or reducing energy consumption for other purposes: (1) additional food uptake, (2) less migration, and (3) reduced somatic metabolism, leading to reduced, or even negative somatic growth. However, the feasibility of these options is debatable (Tafur et al., 2010). Firstly, “any increase in the already high feeding rate is difficult”; secondly, little is known about the migration behavior of *D. gigas*, which may continue unaltered during maturation, as described for other cephalopod species; and, finally, “it has been presumed that growth during maturation continues with the same rate as before”.

Tafur et al. (2010) propose that a temporal organization of resource allocation that decouples the formation of female and male reproductive tissue allows *D. gigas* to make optimal use of available energy in the environment. The authors state that females mature towards the end of their lifespan, while males start forming reproductive tissue at “about half their maximum size“ (Tafur et al., 2010). Thus the male peak energy demand occurs before that of females. These sex-specific peak energy demands are presumed to be an adaptation to low-energy environments, i.e., the less productive warm waters where *D. gigas* is mainly found.

The working hypothesis explored in this chapter is that cohort and population level peak energy demands are reduced by the asynchronous onset of female and male maturity.

8.2. Reducing energy peak demand by asynchronous formation of female and male reproductive tissue

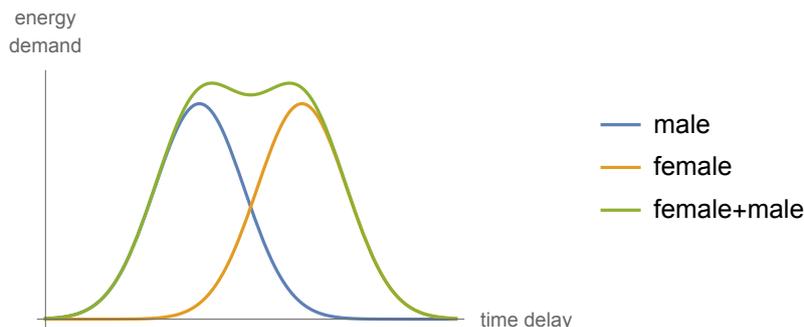


Figure 8.1.: Schematic representation of the flattening of peak energy demand caused by time offset between maturation of females and males.

Testing the hypothesis. The offset between the onset of maturity of females and males is expected to reduce the overall energy demand on the ecosystem at cohort level by prolonging the period of high energy demand for the formation of reproductive tissue, and reducing peak energy demand, as shown schematically in Figure 8.1. Smaller time offsets correspond to higher energy demand peaks, as shown in (Figure 8.2). An ecosystem is more likely to support prolonged demand for lower amounts of energy than a short high peak in energy demand. If energy needs are not satisfied by the ecosystem, population size may decrease, particularly in the case of *D. gigas* due to the lack of effective metabolic or population-level energy buffering capabilities (see the discussion of Subsection 6.3.1 and of cannibalism in Chapter 7).

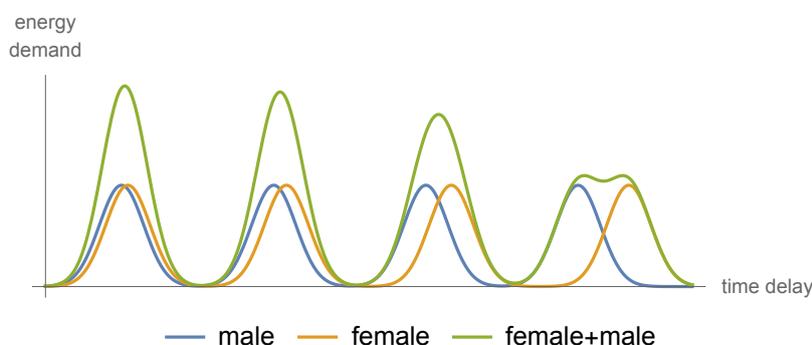


Figure 8.2.: Schematic representation of how longer time offsets between the onset of female and male maturity reduce peak energy demand.

Figure 8.2 illustrates how a longer time offset between maturation of females and males reduces peak energy requirements. However, longer time offsets may exceed the capacity of males of storing viable spermatophores, or increase overall population level

8. Properties at the population level and their computation

energy requirements by increasing the lifespan of females while they “wait” for males to mature.

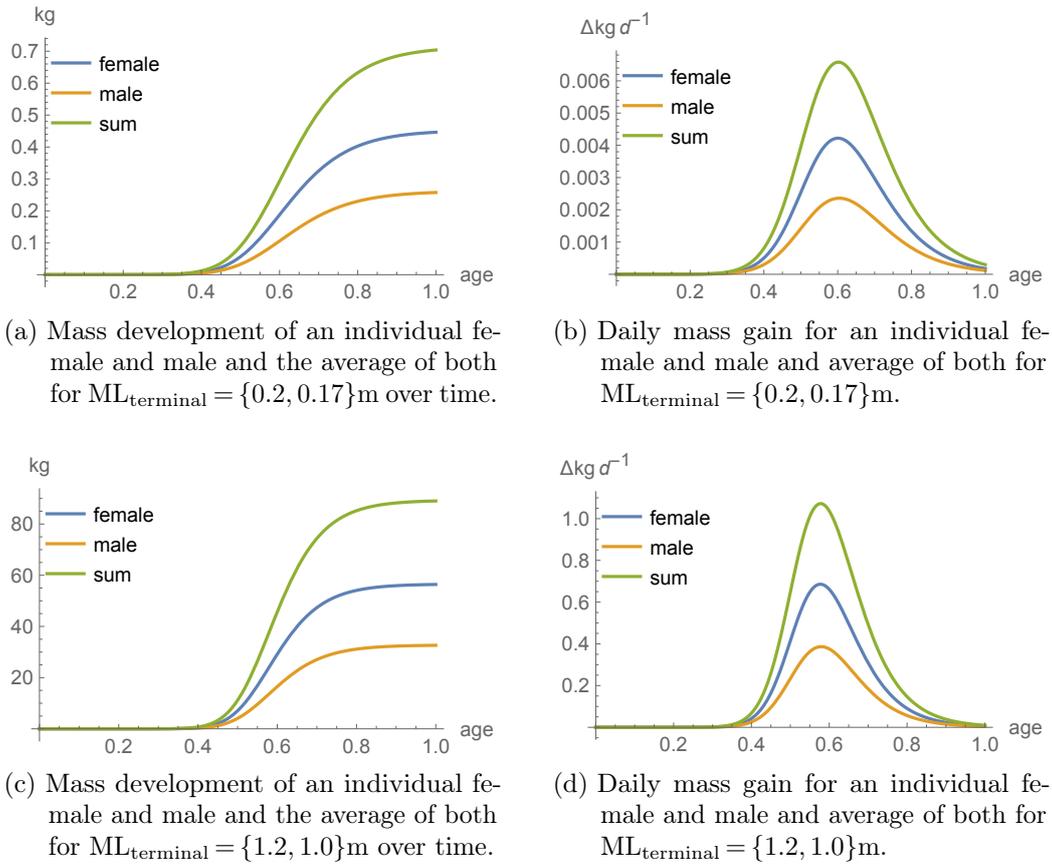


Figure 8.3.: DTSGF: Mass development of males and females computed for different ML_{terminal} over a one-year lifespan. The sum is the added mass of a female and male. The calculated values can be applied to a cohort with a 1:1 sex ratio.

The following analysis tests the hypothesis that an asynchronous onset on female and male maturity has a dampening effect on peak energy demand (Tafur et al., 2010) by estimating the energy requirements, based on the DTSGF, of individuals over the course of their lifespan. The DTSGF assumes that gonadic tissue growth involves no additional energy uptake because any additional energy requirements would be similar to the standard requirements for growth metabolic rate ($\text{growth}_{\text{mr}}$) and minor compared to basal+locomotion metabolic rate ($\text{basal+locomotion}_{\text{mr}}$). Gonadic tissue contributes to the body mass and is therefore included in the $\text{growth}_{\text{mr}}$, which is modelled based on the growth function and is therefore determinable.

8.2. Reducing energy peak demand by asynchronous formation of female and male reproductive tissue

In Figure 8.3, with parameter settings used in previous analyses, the DTSGF does not display an offset between male and female mass gain peaks (Figure 8.3b,d) for small and large ML_{terminal} ; also there is no flattening of the combined $\text{growth}_{\text{mr}}$ peak. In the following analyses, the equidistant values of $[0.2, 0.4, 0.6, 0.8, 1.0, 1.2]$ for the female ML_{terminal} are used, which covers a range of values from small to large SAM-group in Table 2.5; male ML_{terminal} values are calculated by dividing female ML_{terminal} by 1.2.

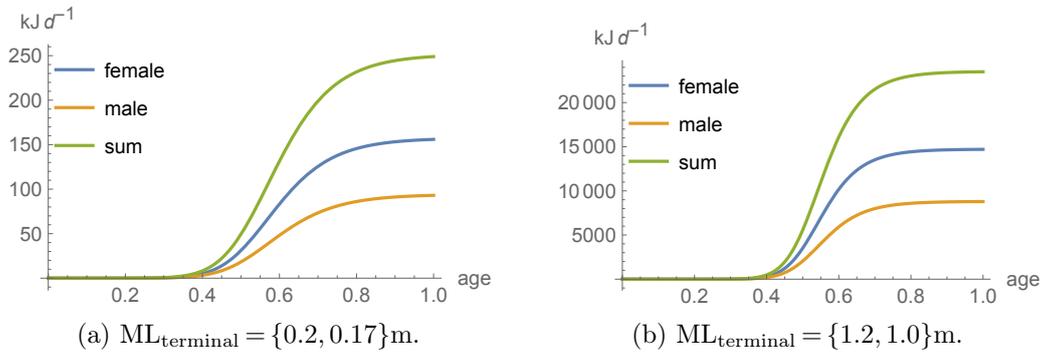


Figure 8.4.: DTSGF: Daily total_{mr} of females and males over the lifespan, for values of ML_{terminal} representing extremes of large and small SAM-groups.

In contrast to peak in daily energy demand for body mass gain, which is assumed to be proportional to the daily gain in body mass shown in Figure 8.3b,d, Figure 8.4 shows that total daily energy demand total metabolic rate (total_{mr}) increases continuously over the lifespan since growth continues until the end of the lifetime. The contrast between Figure 8.3 and Figure 8.4 reflects the fact that $\text{growth}_{\text{mr}}$ accounts for a varying proportion of total energy demand total_{mr} over the lifespan. As shown in Figure 8.5, the ratio of $\text{growth}_{\text{mr}}$ to other energy demands, represented as $\text{basal} + \text{locomotion}_{\text{mr}}$, rises from an initial value of approx. 0.4 to a maximum between approx. 0.75 to 0.95 at an age of approx. 0.4. Values over the lifespan are slightly higher for females than for males, and lower for the lower ML_{terminal} -tuples.

Conclusion. $\text{Basal} + \text{locomotion}_{\text{mr}}$ accounts for the largest part of total energy needs, i.e., total_{mr} , because the locomotion metabolic rate ($\text{locomotion}_{\text{mr}}$) is assumed to be proportional to the species-specific high basal metabolic rate (basal_{mr}), which increases with continuously increasing body mass. $\text{Growth}_{\text{mr}}$ represents a larger proportion of total_{mr} during the period of exponential growth of an individual. However, as the individual gains body mass during this phase, the proportion of energy required for growth decreases as body mass increases. In consequence, total energy needs over the

8. Properties at the population level and their computation

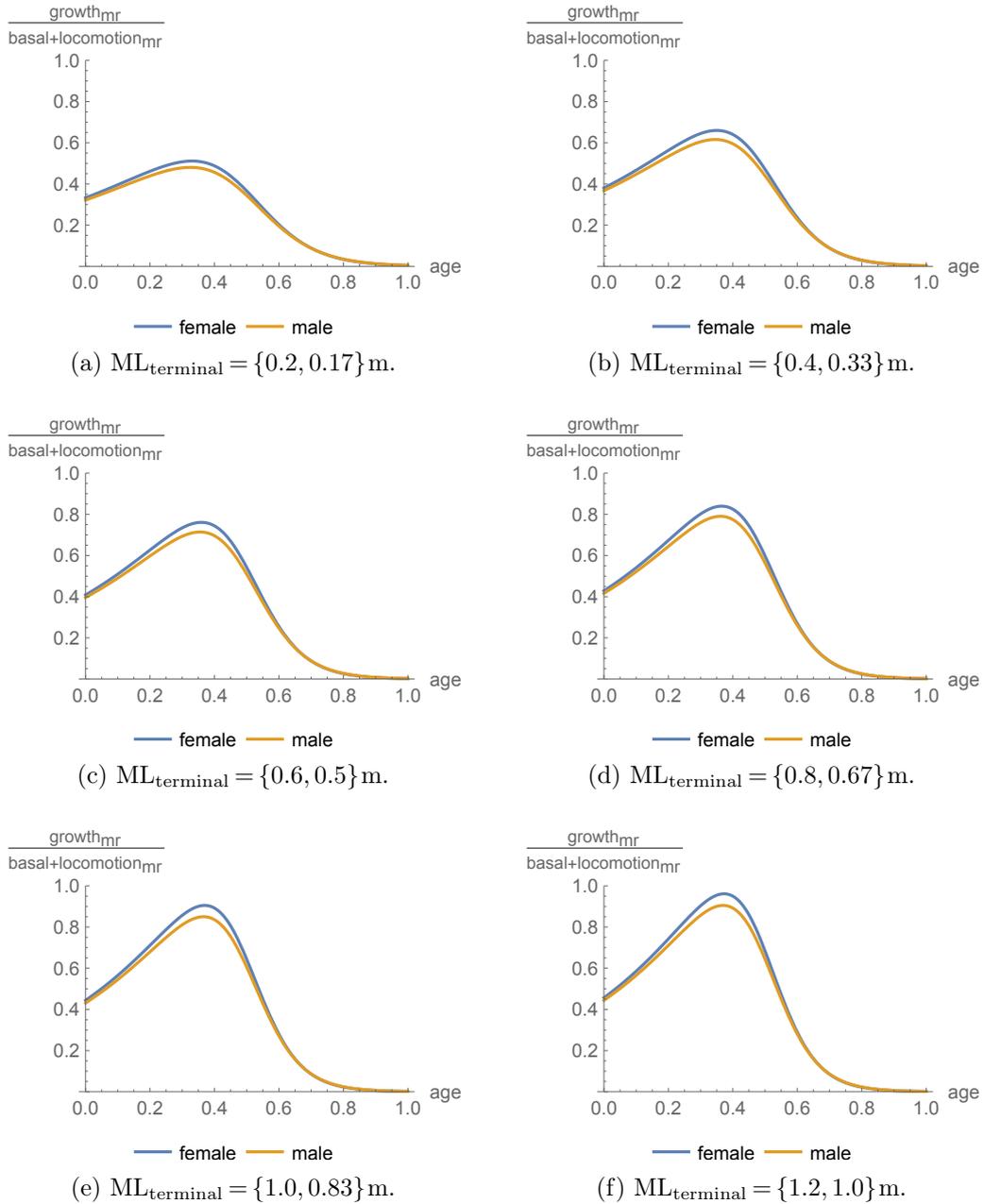


Figure 8.5.: DTSGF: Ratio of $\text{growth}_{\text{mr}}$ to $\text{basal}+\text{locomotion}_{\text{mr}}$ over the lifespan, for females and males and different values of $\text{ML}_{\text{terminal}}$.

course of the lifespan depend mainly on the absolute body mass of both sexes, which increases towards the end of lifespan.

8.3. Energy optimization by smaller mantle length and higher generational turnover

The extent to which a time offset between the onset of female and male maturity would reduce peak energy demand (i.e., for females plus males) would depend on growth rates. Thus populations of small ML_{terminal} individuals, which are present in low energy environments, would benefit less from such an offset than those composed of larger ML_{terminal} individuals, since lower $\text{growth}_{\text{mr}}$ means that there is less potential to save energy. In consequence, the temporal organization of energy needs, and therefore allocation of resources, is only possible if a means is found to reduce the mass-dependent energy needs, as occurs for example when basal and locomotive metabolism are suppressed during a stay in the oxygen minimum layer (OML); but it is questionable whether *D. gigas* would grow when other metabolic functions are suppressed in this way.

The results of the analysis are conclusive: There is no evidence for pronounced temporal organization of resource allocation by *D. gigas* (Figure 8.4) and therefore this possibility is not further explored.

8.3. Energy optimization by smaller mantle length and higher generational turnover

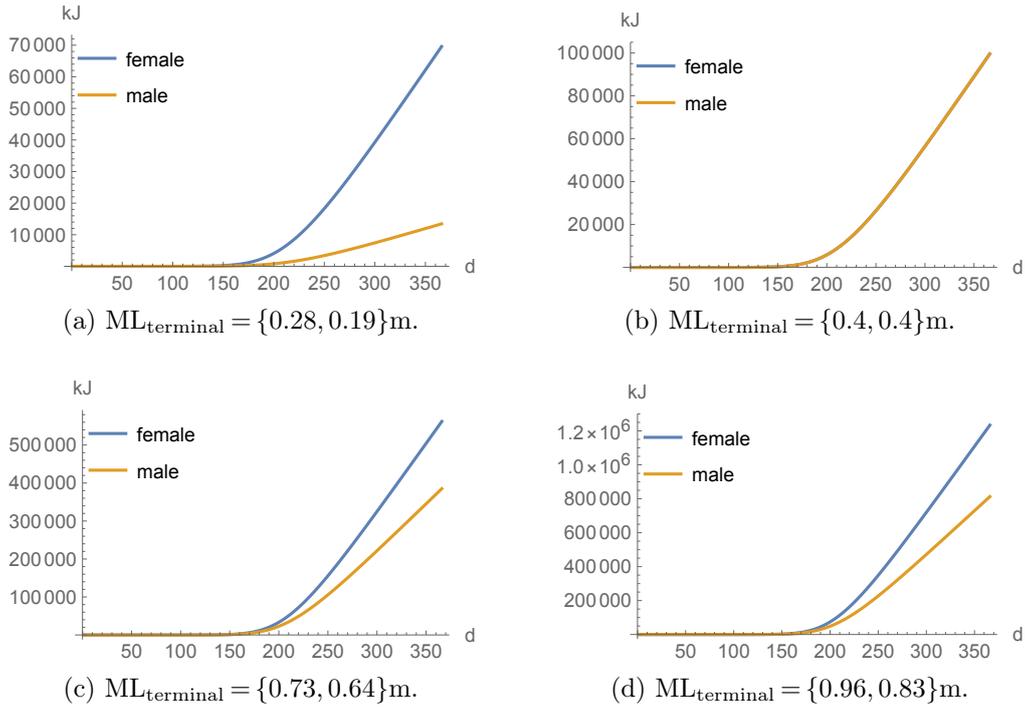
Hypothesis. Size-at-maturity-groups are postulated to be an adaptation to the capacity of the environment to provide food and a means of optimizing the use of available resources. The existence of SAM-groups is further hypothesized as resulting from differences in mantle length (ML) at the onset of maturity under varying environmental conditions, i.e., small ML and a short maturation times are a response to low energy environments (Keyl et al., 2011; Keyl et al., 2008; Tafur et al., 2010). Thus the life histories of the two phenotypic extremes, the small and large SAM-groups, are an adaptation to two basic ecosystem conditions, whereby the small SAM-group ensures population survival during warm periods with low food availability and the large SAM-group maximizes individual fitness during cool and nutrition rich periods (Keyl et al., 2008).

Presumed contribution to optimizing energy use. The above hypothesis relies on two presumed effects that enable optimal use of energy available in the environment: (1) Compared to larger individuals, small SAM-group individuals invest *relatively* higher amounts of energy in growth of reproductive organs and reproduction as a proportion of overall body mass and energy investment. However, because of their small body size, small SAM individuals invest less energy in reproduction in *absolute* terms (Keyl et al.,

8. Properties at the population level and their computation

2008). This is advantageous in a high-temperature environment with lower productivity and less available exploitable energy. Furthermore, small-sized individuals have access to lower trophic levels with higher productivity, containing more numerous but smaller prey items (Keyl et al., 2008).

(2) A shorter lifespan corresponds to a higher reproductive rate per unit of time, hence increased fitness. Early onset of maturity minimizes the absolute amount of required energy until reproduction.

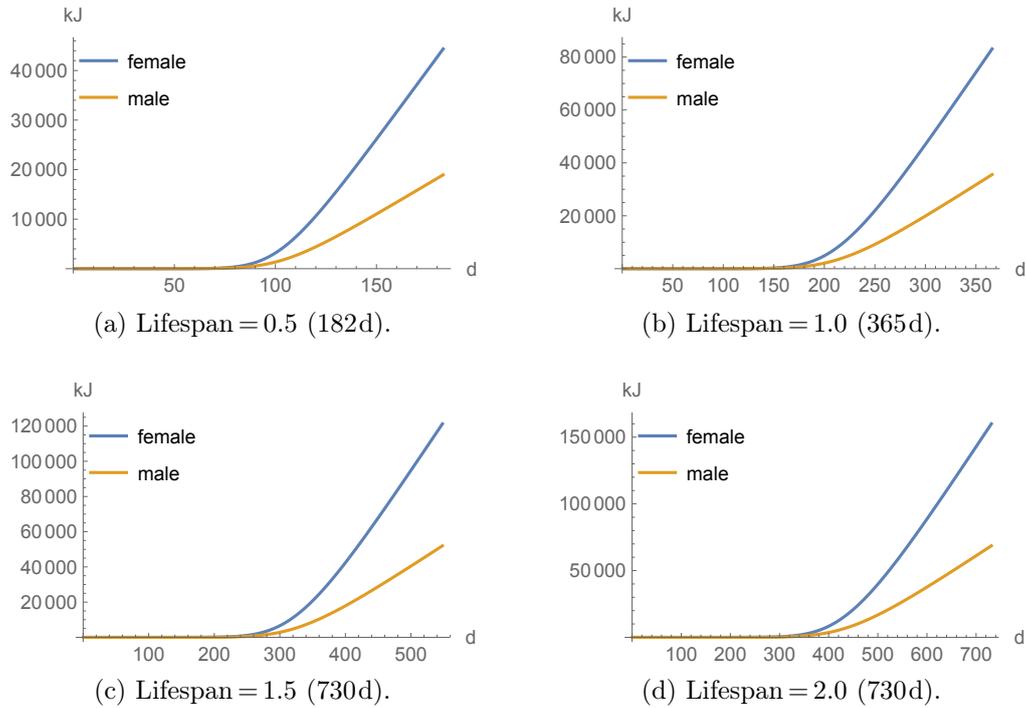


year	female ML_{terminal}	male ML_{terminal}	DTSGF female kJ	DTSGF male kJ
1998/1999	0.28m	0.19m	35 572	11 586
2000/2001	0.4m	0.4m	99 542	99 542
2001	0.73m	0.64m	561 999	385 038
2007	0.96m	0.83m	1 234 020	812 632

Figure 8.6.: DTSGF: Accumulated daily total_{mr} over a one-year lifespan for different ML_{terminal} , shown as {female, male}. The table shows corresponding total energy requirements over the lifespan. ML_{terminal} correspond to typical values in Fig. 2 B in Tafur et al. (2010) for the years shown in the left-hand column of the table.

8.3. Energy optimization by smaller mantle length and higher generational turnover

Testing the hypothesis. To test the hypothesis, the accumulated total metabolic rate (total_{mr}) is estimated for different SAM individuals by summing the daily total_{mr} from hatching until the end of lifespan. As shown in Figure 8.6, smaller ML individuals invest considerably less energy, in absolute terms, in building up body mass and sustaining body mass-dependent metabolism (i.e., $\text{basal}_{\text{mr}} + \text{locomotion}_{\text{mr}}$).



lifespan in years	DTS GF female kJ	DTS GF male kJ
0.5	44 360	18 948
1.0	83 078	35 572
1.5	121 243	51 956
2.0	159 961	68 581

Figure 8.7.: DTS GF: Accumulated total_{mr} for female $\text{ML}_{\text{terminal}}=0.34\text{m}$, male $\text{ML}_{\text{terminal}}=0.28\text{m}$ and different lifespans. The female $\text{ML}_{\text{terminal}}$ corresponds to the reported upper limit of the small SAM group (see Table 2.5), the male $\text{ML}_{\text{terminal}}$ is calculated by dividing the female $\text{ML}_{\text{terminal}}$ by 1.2. The table lists the final values of accumulated energy shown for different lifespans.

8. Properties at the population level and their computation

If a smaller ML_{terminal} is associated with a shorter lifespan, i.e., a higher generation turnover rate, then the shorter lifespan reduces $\text{basal+locomotion}_{\text{mr}}$ relative to total_{mr} . In Figure 8.7, the individuals attain the same ML_{terminal} over different lifespans; therefore, for longer lifespans, the energy required for mass growth is the same but $\text{basal+locomotion}_{\text{mr}}$ accounts for an increasing proportion of the total energy investment over the lifespan. Therefore, the doubling of the lifespan results in a slightly less than double of absolute energy requirement over the lifespan because $\text{growth}_{\text{mr}}$ does not double.

Thus, for individuals of the same ML_{terminal} , the energy required over the lifespan is roughly proportional to the lifespan. Therefore, a population composed of small SAM individuals with a higher generation turnover rate of should still require less energy over time than one composed of larger individuals. As shown in Figure 8.6 (compare panels b & c), an increase of about 80% in ML_{terminal} , i.e., from 0.4m to 0.74m, increases the total energy requirement for a single generational turnover by about four (for males) to five and a half times (for females).

Summary. These calculations support the hypothesis in Keyl et al. (2008) that life history differences in *D. gigas* are responses to low-energy and high-energy environments: (1) Compared to larger SAM individuals, small SAM individuals, place less burden on ecosystem resources over their lifetime. (2) The presumed higher generational turnover rate of the small SAM individual does not result in absolute energy requirements that are higher than those of a large SAM individual energy needs with a longer lifetime and slower generation turnover rate.

8.4. Multiple spawning peaks to reduce peak energy demand

Multiple intra-generation spawning peaks may optimize energy use at population level by distributing peak energy use over the time (Figure 8.8). If multiple females spawn at the same time, then these simultaneous spawning events create single peak in energy demand on the supporting ecosystem. The effect of multiple spawning peaks in reducing peak energy demand is quite conclusive; moreover, the more time that elapses between spawning events, the greater is the decrease in peak energy demand (Figure 8.8).

Simultaneous multiple spawning events place an energy burden on the supporting ecosystem. At a population level this is a counterproductive strategy, since some individuals may die due to energy exhaustion, or alternatively may be unable to obtain sufficient food to successfully complete the spawning process. Theoretically, an

8.4. Multiple spawning peaks to reduce peak energy demand

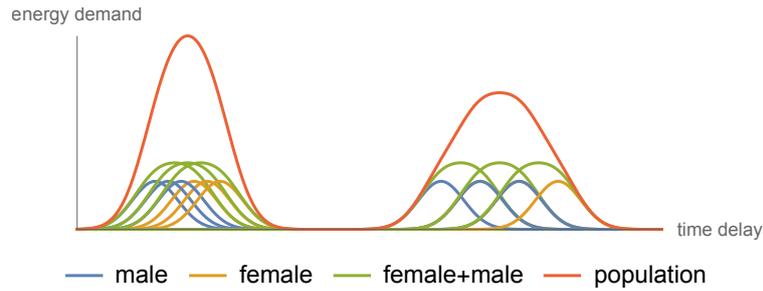


Figure 8.8.: Schematic representation of multiple spawning events with shorter (left) and long (right) spawning offsets, and corresponding change in overall energy demand over time.

optimal offset may exist between spawning events that minimizes energy demand at the population level. However, since both the number of spawns per year and the cohort lifespans are unknown variables (Keyl et al., 2011; Tafur et al., 2010), it is not possible to verify that *D. gigas* employs an energy optimization strategy of this kind.

Another strategy for optimizing energy use at the population level may be the spatial organization of spawning. Spawning at (distant) non-interfering locations would distribute total peak energy demand, creating multiple local peaks energy demands, which could more easily be met by ecosystems.

Finally, in certain circumstances, simultaneous spawning in a single location might actually favor successful spawning, if the peak energy demands of the cohort coincide with a peak in food availability in the area where spawning takes place.

In any case, when the total amount of individual energy needs are taken into account, it is by no means certain that asynchronous spawning would result in a reduction of peak energy demand as suggested by Tafur et al. (2010). In fact individual energy needs increase over the lifespan (Figure 8.4), and there is no evidence for the occurrence of peaks in energy demand as shown in Figure 8.1. Moreover, the energy needs of a cohort are given by the product of the number of individuals and their current energy needs. As time progresses, the absolute energy needs of individuals increase, but their number may decrease due to predation or cannibalism, for example. As an effect, each individual in a cohort needs more daily energy, but the smaller number of individuals decreases the total energy need of the cohort. In fact peak energy need at cohort level may occur earlier in the life cycle, when there are a larger number of smaller (younger) individuals.

In summary, asynchronous spawning theoretically reduces peak energy demand at the population or cohort level by distributing energy demand over time. However, the

8. Properties at the population level and their computation

computation of the individual daily energy calls into question the existence of energy demand peaks during spawning, which is the assumption that underlies this hypothesis. Moreover, the spatial distribution of spawning calls into question the relevance of energy peaks at population level, since the survival of a school will depend on the productivity of multiple local environments and not on the average productivity of a large regional ecosystem.

8.5. Optimization of energy use by sex ratio change

“[...] Males do not directly contribute to reproduction” but, from an energetic point of view, compete with the females, the actively reproducing part of the population (Tafur et al., 2010). A shift to a higher female:male ratio may be an adaptation to prioritize energy expenditure on reproduction in warm-water regimes with low productivity. According to Chen et al. (2020), an individual should trade off energy allocation against somatic growth and reproduction to achieve, when the trade-off is optimal, maximum reproductive success – and thus also success in terms of population size and stability. However, *D. gigas* also grows during maturation and attempts to maintain growth through continuous feeding, and cannibalism may be one way to achieve this. To be successful, this strategy would have to favor the survival of females, as the actively reproducing part of the population to which food resources should be preferentially allocated when food resources are scarce. This could be achieved, for example, by partially removing the males from the population. Tafur et al. (2010) found a change in the female:male sex ratio from an average of $\approx 3:1$ between 1991 and 2002 to $\approx 2:1$ in 2006. The sex ratio change coincided with increasing ML and decreasing water temperatures and, therefore, it is assumed, increasing environmental productivity.

The female:male sex ratio change followed the decrease in water temperature with a 2–3yr lag and the increase in ML with a 1–2yr lag (see Tafur et al. (2010) Fig. 2). In this case, ML might be a more plausible immediate trigger of an epigenetic shift in the sex ratio in favor of males, since these should be observable within the turnover time of a single generation, which is to be assumed shorter than 2–3yr. Large MLs are an indicator of currently high environmental productivity, i.e., of conditions where the existence of more males does not threaten the survival of females. The 1–2yr lag between increased ML and the changed sex ratio is explainable by the fact that this change occurs in the next generation, as a consequence of large ML females spawning relatively more males in comparison to small ML females. The longer time required

8.5. Optimization of energy use by sex ratio change

for large ML females to reach maturity, compared to short ML females, could further increase the delay before this effect is apparent.

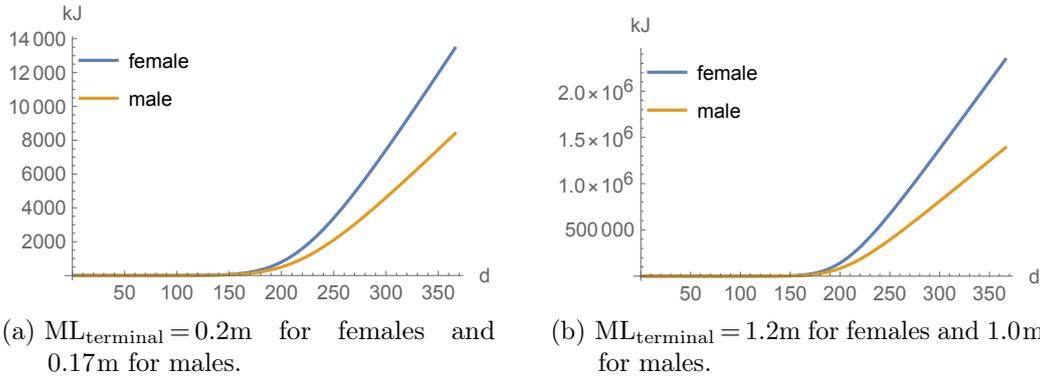


Figure 8.9.: The DTSGF estimated accumulated energy demand of a female and male for different ML_{terminal} over a one-year lifespan.

However, Tafur et al. (2010) also report a high degree of variation in sex ratios, i.e., between 6.55:1 and 0.24:1 before 2002 and between 3.86:1 and 1.09:1 after 2002. This variation suggest that there is no simple causal relation between ML and the sex ratio relation. However, the variation might also have been caused by sampling bias; e.g., the ML of individuals caught may depend on the jig size used for sampling (Keyl, 2009; Keyl et al., 2011; Markaida & Sosa-Nishizaki, 2001). The average sex ratio seems to be a reliable indicator of changes over time.

The energy saving potential of a sex ratio change is estimated by computing daily energy requirement of males and females, as well as their total energy needs over the course of the lifespan, as shown in Figure 8.9 for both phenotypic extremes. In the DTSGF-based model (Section 6.2), females use more energy than males, because of their then higher ML_{terminal} . As shown in Figure 8.10, both the daily and accumulated energy demands of females are much higher than those of males over the full lifespan and these do not change significantly between the phenotypic extremes.

A changing sex ratio in favor of the females by removing the males from a cohort reduces the total resource requirement to ensure reproduction of a cohort and thus may promote the survival of a cohort under adverse environmental conditions without losing too much genetic diversity. Of course, the removal of males is likely to affect the reproductive capacity of the current population, but this could be compensated for by the fertilization of several females by one male. In this case, only the genetic diversity would be affected.

8. Properties at the population level and their computation

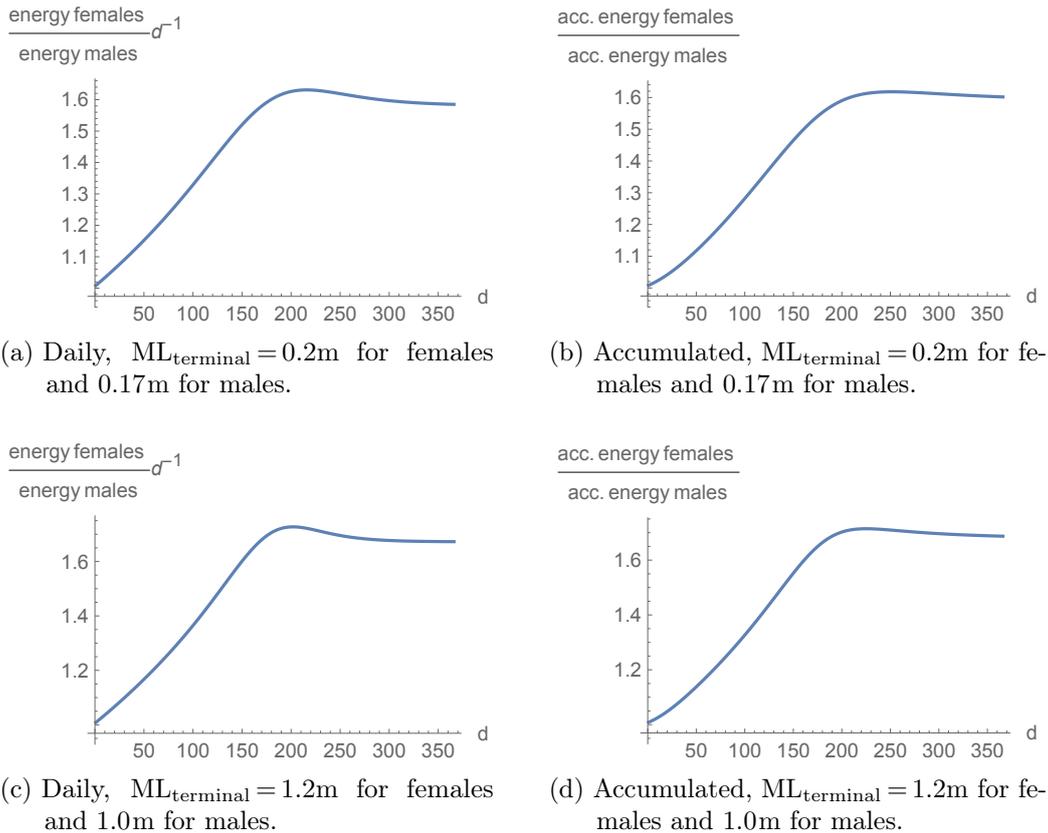


Figure 8.10.: The DTSGF estimated female:male ratio of daily and accumulated energy demand over a one-year lifespan.

In the model, males use less energy than females. Correspondingly, the females have a size advantage that allows them to predate on males (Figure 7.1, page 138) during periods of food deprivation and/or higher energy demand by females. Thus, cannibalism of males by females would not only provide females the energy required for reproduction but could also be the mechanism driving the change in the female:male ratio in favor of the females.

8.6. Discussion and conclusion

This chapter examined population-level effects based on the functional triad migration-maturation-growth (fTMMG) (Keyl et al., 2008), which was refined in Chapter 3 and Chapter 4, and the population-level effects and assumptions described in (Keyl et al., 2008) and Tafur et al. (2010). However, the population-level effects described are not

as straightforward as the calculated characteristics of individuals because they are the emergent result of multiple factors, including individual characteristics.

The analysis does not provide support for the concept of asynchronous formation of male and female reproductive tissue as a means to reduce peak energy demands on the the supporting ecosystem (Section 8.2). As individuals grow, i.e., ML increases, total metabolic rate (total_{mr}) is dominated by the $\text{basal+locomotion}_{\text{mr}}$ (Subsection 6.3.1), so no pronounced effects at population level energy use can be expected from sex-specific temporal offsets in the onset of maturity and growth of reproductive tissue. The DTSGF shows no prominent decoupling of male and female body body mass gain. However, modeled estimates of $\text{basal+locomotion}_{\text{mr}}$ have a significant degree of uncertainty and are based various untested assumptions, so the concept of asynchronous maturation as a means of optimizing energy use cannot be conclusively rejected.

The DTSGF-based model provides more conclusive support for the idea that smaller size, i.e., shorter ML, and a probable associated increase in generational turnover are an adaptation to low-energy environments. The accumulated total_{mr} is lower for a small ML individual than for a larger ML individual over the same timespan (Figure 8.6) for both sexes. A higher generational turnover rate does not increase the total population-level energy requirement because absence of large SAM individuals leads to a considerable saving in body mass-dependent $\text{basal+locomotion}_{\text{mr}}$ over the same period. Modeled estimations of energy use support the hypothesis in Keyl et al. (2008) that the existence of different SAM groups represents an adaptation of life history strategies to energy levels in the environment.

Female *D. gigas* individuals spawn at different times throughout the year (Chen et al., 2013; Keyl et al., 2011), thereby giving rise to a number of distinct cohorts within a population. This contributes to the optimal use of available energy in the environment, since peaks in energy demand by different cohorts are distributed over time. Thus peak energy demand on the environment will be lower than for a single cohort with the same number of individuals.

A changing sex ratio in favor of the females by reducing the number of the males in the population reduces the population size and therefore decreases the amount of energy required for reproduction of the population. This may also affect genetic diversity, but this aspect is not further considered here.

In the DTSGF, females use more energy than males, thus the removal of males and the reduction of population size is the only way to reduce the overall energy demand without reducing population-level fitness, i.e., without reducing the number of offspring recruited into the next generation. The DTSGF suggests that cannibalism of males

8. Properties at the population level and their computation

by larger females may constitute a mechanism for active control of the sex ratio. In this scenario, cannibalism would function as an effective survival strategy in response to low energy conditions both for individual females and for the whole population. Cannibalism provides a behavioral explanation for observed changes in the female:male ratio, suggesting that the sex ratio in a population is dynamic and not predetermined.

9. The simulation of ecological models: Enhanced approaches and techniques

9.1. Introduction

In the context of this thesis, simulations were developed, e.g., a flow simulation using a cellular automaton (CA) approach (see Chapter 5) and a simulation of cannibalistic individuals using a individual-based modeling (IBM) approach (see Chapter 7).

Modeling using differential equations was deliberately omitted because this is incompatible with the bottom-up approach of IBM, which is designed to elucidate effects arising from the interaction of individuals. Differential equations presuppose that the effects are already known, because they are already incorporated into formulation of the equations.

IBM and CA are well-established methodological approaches for the implementation of ecological simulation models. If these approaches are adopted, the structure of a simulation model has to be adequate for their use (see Jopp et al. (2011)). When developing the simulation model, the programmer may use off-the-peg tools, for example a simulation environment or a programming language with a framework, but these tools may require adaptation for implementation of IBM and CA.

For this thesis, it turned out that there was no freely available turnkey solution for the implementation of the simulation models. Moreover the flexibility and/or performance of available simulation environments appeared to be limited, and they were judged to be inadequate for full development of complex simulation models.

In the end, the implementation of the flow simulation in Chapter 5 used the programming language PYTHON and various freely available libraries for numerical computation, visualization and output; technically it was a combination of fast tools and a (slow) script language (PYTHON). The IBM simulation in Chapter 7 was intended to be much more complex; however, the original model was considerably simplified in the course of its development. In the context of this thesis, structures and procedures for the implementation of complex IBM were identified, but found to be unnecessary for the analysis.

Considerable problems became apparent during implementation of the model as the limits of the known approaches and existing tools emerged. To address these problems,

further approaches to modeling and simulation were considered and experimentally implemented in the thesis. The starting point was a mechanistic approach that is well established within the field of IBM: the individuals are regarded as “deterministic machines”, whose behavior is controlled by defined rules, which include pseudo-random behavior. Thus, an algorithmic description and therefore a computer executable model is possible, but at the same time there are fundamental problems in the representation of such machines in a discrete-event simulation (DES), because required features are not supported in widely used computer programming languages.

This chapter describes a modeling approach that addresses these problems, without claiming completeness, and proposes methodologies for its implementation. The aim is to simplify the formulation and implementation of simulation models and, to achieve this aim, both methodological and technical options are considered. As a first step, Section 9.2, the elementary (i.e., to be retained) building blocks of the simulation model are described and standardized and, where necessary, methodically combined. Section 9.3 discusses the general features of IBM and CA and how these affect model implementation. Section 9.4 continues the discussion of model implementation and suggests some ways to simplify the modeling process.

Section 9.5 discusses the problems involved in translating a source model into a simulation model written in programming language (Lorek & Sonnenschein, 1999). Section 9.6 assesses the suitability of selected simulation environments for the development of simulation models by non-specialists, with respect to documentation, validity and implementation. Section 9.7 discusses an alternative approach, i.e., the use of a programming language in combination with an extension in the form of a framework or library, and explores the complex issues involved in implementing this approach.

Section 9.8 extends the discussion to the modeling and translation procedure and presents literate programming (LP) and enhanced dynamic literate programming (DynLP) as approaches that enable *uniform model description and simulation model creation* and overcome the shortcomings of the Overview, Design concepts and Details Protocol (ODDP) (Grimm et al., 2006; Grimm et al., 2010; Grimm et al., 2020), proposed by (Grimm et al., 2020) as a standard procedure for model documentation.

9.2. Individual- and agent-based-modeling

IBM and agent-based-modeling (ABM) are established methods to model complex systems of actors, i.e., individuals, in an environment, e.g., socio-ecological or ecological systems. In reference to computer science, Grimm et al. (2005), Lorscheid et al. (2019),

and Mooij et al. (2014) refer to these as agent-based complex systems (ACS) because the modeled individuals pursue their goals independently and are therefore called agents. Modeling of agents (and their environment) now goes beyond simple algorithmic description, e.g., through the use of fuzzy-logic based algorithms, artificial neural networks, machine learning, and genetic algorithms (DeAngelis & Diaz, 2019; DeAngelis & Yurek, 2017; Manson et al., 2020; Marchini, 2011). In line with current literature, in which the term ABM is very often also used as a synonym for IBM, in this thesis IBM and ACS are both referred to as ABM; even in cases where the modeling of the individuals is relatively simple and therefore does not meet the standards of an “agent” in computer science, where the term usually refers to a reasoning entity rather than an individual that merely interacts with others, as in IBM. However, occasionally, for example when citing or quoting, the term IBM is used.

In general, ABM express a high degree of flexibility as a result of adopting a bottom-up approach. The modeling of the agents is first conducted locally on a small scale and is then extended by including more agents for execution on a larger scale (DeAngelis & Yurek, 2017; Grimm et al., 2017). In this way, ABM are able to “represent interaction of structural and functional features across different scales”, “quantitative transitions” and “qualitative or structural changes” (Reuter et al., 2011).

As a result of small-scale interactions at the local level, emergent effects, i.e., unpredictable patterns of behavior, arise at the global level (DeAngelis & Yurek, 2017). ABM consider, compute and analyze complex systems (ACS), but currently there is no unified theory to describe the behavior of these models at the level of agents or, at a higher level, as systems (Lorscheid et al., 2019; Zhang & DeAngelis, 2020). S. F. Railsback and Harvey (2020) state that execution of ABM, i.e., simulations, can reveal phenomena and their operating principles that help explain corresponding patterns found in nature, such as flocking (S. F. Railsback & Harvey, 2020). Pattern-oriented modeling (POM) (S. F. Railsback & Grimm, 2012, 2019; S. F. Railsback & Harvey, 2020) pursues this approach by testing hypothesis about patterns found in nature and making testable predictions.

Modeling the complexity of ACS is challenging in many ways. Complex systems, in this case ACS, elude formal analysis or a uniform description across all scales; thus, not all the interactions of a complex system can be incorporated into a model (which by definition is a simplification of reality). Difficulties arise, for example, during the design of models, their implementation, and their documentation. Documentation includes, for example, descriptions of the purpose, properties and behavior of a model to make it easier for others to understand, and render it amenable to scientific scrutiny. Solutions to the

above challenges have been proposed, e.g., the Overview, Design concepts and Details Protocol (ODDP), which specifies standard procedures for designing and documenting ABMs. However, as argued in Section 9.5, this approach lacks the assurance that the model, documentation, and implementation of the model do not differ semantically.

The chapter focuses on the design (structure), implementation and documentation of ABM, a field in which the author of the thesis, like Grimm et al. (2020), still sees (considerable) potential for improvement. The general problem of a unified description or theory on complex systems is not discussed explicitly or in depth. The focus is primarily on techniques, and especially implementation of ABM (see Jopp et al. (2011) and S. F. Railsback and Grimm (2012)), to provide inputs for further discussion on informatics-related aspects of these techniques; this is, like ABM itself, a bottom-up approach.

9.3. Agent-based-modeling Implementation: Structures, Methods and Techniques

9.3.1. General Systems Theory and agent-based-modeling

ABMs are based on General Systems Theory (GST). GST describes the object of interest as a system consisting of a set of *components*, *interactions*, and the *process of interactions* (Forrester, 1968). The first step in creating a system is to define the boundaries by separating the system from the surrounding, embedding context. Next, the components and their interactions, i.e., the driving elements and forces, are identified; this allows the construction of a cause–effect diagram that portrays the interactions (Breckling, Jopp, et al., 2011).

This procedure, still the starting point for developing an ABM (An et al., 2020; Cornell et al., 2019; Manson et al., 2020), derives from object-oriented system analysis (OOSA), an approach to modeling that is closely related to object oriented programming (OOP) languages (Breckling, Jopp, et al., 2011), which will form the basis for the discussion of how to improve the implementation of ABMs at the end of this chapter. In the OOSA approach, an individual is composed of components, each representing the quantifiable property of an individual, and a “flexible” structure that serves as a dynamic information store. The interaction among these components expresses the behavior and state of the individual. Components may be simple (e.g., size) or complex, (e.g., metabolism). It is the decision of the modeler to choose the components and define their degree of complexity. The OOSA approach also works at the population level by incorporating interactions of

individuals, which allows the representation of organisms in their environment (Reuter et al., 2011).

An ABM approach differs considerably from modeling using mathematical functions, i.e., differential equations, because these equations are distillations of knowledge that has already been extracted (from experimental or theoretical data) and analyzed. In this case, local-scale variations are not expected to give rise to emerging patterns on the larger scale. In the case of the ABM, however, the self-organizing interactions of the individual components can give rise to unforeseen patterns of system behavior at higher levels of organization. Since an ABM can be applied for the synchronous investigation of different aspects of a situation of interest, such as physiological processes, patterns of behavior, and environmental interactions, these emerging effects can yield new insights into ecological processes (Reuter et al., 2011). A further advantage of IBM is that of “joining structural, functional, quantitative and qualitative aspects in a way that closely conforms to observation data and conceptual knowledge representation” (Reuter et al., 2011).

9.3.2. ABM implementation techniques

Reuter et al. (2011) consider that implementation of an ABM consists of three essential steps; and this approach is still state-of-the-art for ABM implementation, for example in NETLOGO (Lippe et al., 2019; Manson et al., 2020; Murphy et al., 2020; S. F. Railsback & Harvey, 2020). Each step comprises methods for description of the system at a different level of abstraction: (1) “The representation of an individual entity as a class” that defines the individual components of the system; (2) “the layout of a structured presentation of the environment”, that defines the interactions among components, and (3) “the organization of the temporal model execution and interaction between the entities” that defines procedures for simulation of model behavior over time.

The representation of an individual entity as a class is an OOP-based approach for the implementation of a simulation model that is derived from OOSA. A class is a template that defines one or more properties of an individual and contains the code to access and manipulate these properties (see Subsection 9.4.3). Transferred to an ABM, the code defines the behavior and interaction of an individual while the data attached to the object specifies its properties.

An object (the individual entity in the simulation) is thus the instantiation of a class, whose properties (data) contain object-specific values that allow objects to be distinguished from each other (Reuter et al., 2011). Furthermore, in OOP, the only program code that is allowed to update the object’s properties is contained in the class,

9. *The simulation of ecological models: Enhanced approaches and techniques*

so no “external code” directly alters the object state; this *encapsulation* principle is an integral part of object oriented (OO) design in computer science. Thus, external events which require an object to alter its state, as may occur during a simulation, have to “ask” the object to determine that its state has been altered in response to the event (see also Subsection 9.4.3).

Internal states of the object are updated as a result of “activity procedures” (e.g., locomotion, reproduction, nutrition) and “physiological processes” (e.g., metabolism, aging). A practical approach is adopted, whereby each activity or process only manipulates its own set of variables (Reuter et al., 2011). This code structure avoids the complexity that would be caused by mixing functionality and the need to cross-reference between processes and properties. This conforms to good practice for programmers that a code segment should perform a single defined task.

The control of the activities is commonly implemented by a “loop control structure”, whereby the loop is iterated for as long as the object participates in the simulation, e.g., it has the internal status of an active (living) entity. Within this loop, “[...] the execution of any specific activity can be made dependent on distinctive conditions that relate to the internal state of the individual entity [...] and with the external situation [...]” (Reuter et al., 2011). An algorithm is used to select the activities to be performed from the pool of available activities. For example, a simple algorithm calls for the activities to be performed in a fixed order; a more sophisticated algorithm incorporates a “priority driven activity control”, whereby the priorities of the available activities change as the simulation progresses and those with the highest priority or those with a priority above a certain level are executed. The algorithm could also stipulate that an activity commences after the period in simulation time required to complete the preceding activity required. In this case, the system incorporates the principle of discrete-event simulation (DES), a system property which “[...] has to be considered from the start [of the design process]” (S. Railsback et al., 2017).

Most ABM simulations play out in a grid-based space. In the simplest case a CA without rules, i.e., that merely stores information, is used to lay out the environment, with each cell containing the information required (e.g., temperature, salinity, chlorophyll concentration etc.) for the modeling of spatially heterogeneous resources. A CA is a suitable and common way to represent a dynamic environment; some relevant aspects that need to be taken into account are briefly discussed in Subsection 9.3.3. A CA may be substituted by the integration of spatial data models, e.g., sensed and/or modeled data, so that the computation of the environment is externalized (Accolla et al., 2020; DeAngelis & Yurek, 2017).

9.3. Agent-based-modeling Implementation: Structures, Methods and Techniques

At a higher level, the objects are embedded in an execution environment that controls the updating of objects, i.e., operates the main iteration loop of an object, and determines the order in which events are executed. Parallel execution of events is difficult to perform, because for example (1) a single execution instance is required for each object during a simulation step, (2) blockages may be caused if more than one object attempts to access the same resource at the same time, and (3) modeled activities may require different completion times. The order of execution of events and their execution times may be determined by different components of the simulation model. Such process scheduling is not an integral part of OOP; instead, it is strongly dependent on the programming languages used and their extensions, i.e., libraries.

In summary, an ABM is based on the execution of discrete events which represent the interaction of agents with each other and their environment, whereby the current state of each agent, resulting from these interactions, is displayed for each specific point in simulation time (Reuter et al., 2011). The simulation environment contributes a scheduling mechanism and explicitly determines the order of execution of events.

The above description of the mechanisms for executing an IBM, adapted from Reuter et al. (2011), is still rather abstract. Expanding on this abstract representation, descriptions of individuals are augmented by specifying (alternative) states, which provide additional structuring support and can potentially simplify the development of models and simulation models, see Section 9.4. For the purposes of this discussion, this thesis differentiates between the development environments used to create and run ABM simulation models (see Section 9.6), and the programming language, with extensions, used to implement (and execute) ABM simulation models (see Section 9.7).

9.3.3. The cellular automaton, its universality and its equivalency to individual-based modeling

A cellular automaton (CA) consists of an arbitrarily sized finite regular grid of cells, each comprising an instance of the same automaton (cell) described by the same set of variables and subject to the same rules. To move from one system state to the next state, all cells are manipulated *concurrently*, with each cell following the same set of rules and taking account of the state of neighboring cells. In ecological simulations, a CA typically spans a two- or three-dimensional space, but any finite *dimension* $N \geq 1$ is allowed. The computational effort increases with the number of cells and dimensions.

In contrast to partial differential functions, a CA models spatio-temporal factors through its set of algorithmic rules (Boerlijst & Hogeweg, 1991; Breckling et al., 2005; Vincenot et al., 2011). The computation of the next cell state according to

9. *The simulation of ecological models: Enhanced approaches and techniques*

rules, including algorithms, is more expressive than mere mathematical functions. Furthermore, in an individual based simulation model that includes CA-modeled areas, the spatially distributed individuals may interact with the cells of the CA (Breckling, Pe'er, et al., 2011; Reuter et al., 2011).

A CA has a universal computability (Cook, 2004, 2009; Martínez et al., 2011) so that any problem, i.e., a system, can be computed applying the rules of physics at an appropriate resolution. With regard to resolution, the availability of computing resources imposes limits on the spatio-temporal precision of computed values, which should be borne in mind when interpreting the results. To be accurate, a CA would have to run at the scale of Planck units, with a resolution of Planck time (approx. $5.391\,06 \times 10^{-44}$ s), Planck length (approx. $1.616\,199 \times 10^{-35}$ m) and the reduced Planck constant \hbar as the smallest transferable energy amount.

IBM and CA differ in their algorithmic approach, but are functionally identical, since a group of spatially distributed individuals of the same type is the functional equivalent to the grid of cells in a CA. At the computational level the equivalence of IBM and CA depends critically on the resolution. At fine-scale resolutions, a CA may potentially contain an enormous number of cells, each of which has to be computed and mapped into the memory. Therefore an IBM is an optimization of a sparsely filled CA, because only cells neighboring an object (individual) need to be taken into account in the computations. As a corollary, compared to a CA, an IBM incorporates a more complex set of rules for its execution. However, the equivalence of IBM and CA means that statements about an IBM are transferable to a CA, and vice versa.

9.4. **Enhanced techniques to implement ABM**

9.4.1. **Introduction**

This section focuses on programming languages and programming techniques, because in the end it is these that determine the states and behaviors of an individual in a simulation model, even though simulation environments often use visual tools. Therefore, this discussion should also apply to graphically oriented simulation environments.

Reuter et al. (2011) state that “[...] IBM is a crucial tool in testing consistency of ecological knowledge”. Moreover, these authors “[...] advocate that ecologists, with the necessary knowledge at hand, should [...] learn to program their own models instead of relying on specialized programmers or pre-defined software tools which usually restrict the optimal adaptation to what the specific situation requires”. In other words, ecologists should do the programming themselves, using a suitable programming language, in

accordance with the requirements of their model. But, the implementation of simulation models without a predetermined simulation environment is, at least among ecologists, probably the exception rather than the rule. NETLOGO seems to be the preferred application (S. Railsback et al., 2017), with alternative platforms, such as MASON (Luke, 2019), only occasionally used (Zhu et al., 2015).

However, due “[...] to their potential to represent detailed biological knowledge and small-scale mechanisms, IBMs tend to have a complex model structure. This requires a particular attention to model documentation and evaluation [...]” (Reuter et al., 2011). Not surprisingly, many ecologists struggle with the programming of complex ABMs, as well as their documentation and evaluation, even when NETLOGO is employed as a user-friendly simulation environment (Ayllón et al., 2020; Grimm et al., 2020; Manson et al., 2020).

Discussions of the above issues in Reuter et al. (2011) in Jopp et al. (2011) highlight the mismatch between user requirements and available implementation tools. This mismatch is still evident in recent applications of ABMs (e.g., An et al. (2020) and Zhang and DeAngelis (2020)).

In the following discussion, approaches to these issues are addressed sequentially. The essential object types generally used for the development of spatio-temporally organized agent-based ecological simulation models are identified in Subsection 9.4.2. Subsection 9.4.3 differentiates object oriented programming (OOP) from state-based modeling (SBM), and the latter is suggested as an improved modeling approach in Subsection 9.4.4. This approach is based on an automaton and structures the modeling of an individual both in the design of the source model and in the implementation of the simulation model. Subsection 9.4.5 proposes using the ecological design pattern (EDP) approach, which can be combined with the use of any programming language, to facilitate model development by suggesting standardized solutions for reoccurring problems. State-based modeling (SBM) is a complementary approach that facilitates structuring of models implemented using OOSA and OOP approaches. Compared to these approaches, in SBM, the modeling process and the implementation of an individual is structurally more similar to a mechanistic model. SBM imposes an additional structure on the individual and its actions in a way that facilitates, similarly to OOP, the encapsulation of code and data. In this section, SBM and ecological design pattern (EDP) are suggested as approaches that can be used to ensure that model structure harmonizes with the user requirements of simulation models.

9. *The simulation of ecological models: Enhanced approaches and techniques*

Documentation, since it not strictly speaking part of the implementation process, is discussed separately in the following section, Section 9.5. (Documentation is considered in the context of the implementation process in Section 9.8.)

9.4.2. **Essential object types for a spatio-temporally organized ABM**

The spatio-temporally organized environment of an ecological ABM incorporates three basic object types:

- **Agent**, i.e., a (usually) mobile individual,
- **Singleton**, i.e., a single instantiation without a location
- **Resource**, i.e., an immobile spatial element.

All objects in an ABM are derived from one or other of these types by extending their basic functions; for a similar approach see SeSAm (2012a) and compare Breckling, Pe'er, et al. (2011), Castle and Crooks (2006), Grimm and Railsback (2005), and Reuter et al. (2011). The **Agent** represents the individual that interacts with other **Agent** objects and **Resource** objects. A **Singleton** is a read-only object accessible to all objects in the simulation model that does not interact with other objects; exactly one object exists of each **Singleton** type. A **Singleton** specifies globally relevant factors such as daylight length or a time value.

Data on **Resource** objects is organized in cells and an algorithm updates the information on each cell. A CA may be an appropriate way to organize resource objects as it is able to model one or more environmental factors by algorithmic computation. A CA is also able to model a vast amount of individual agents, e.g., paralarvae, by specifying only the number of individuals in each cell. Individuals may dissociate from the cells and transferred into an ABM in later simulation steps.

The (simulation) model-specific types **Agent**, **Singleton** and **Resource** would be implemented as classes in an OOP language and are considered as additional types to the existing types of a programming language, e.g., integer numbers, strings etc.

A **Resource** may also be driven by an external information source, e.g., a data base or geographic information system (GIS), where the state of a cell is determined by querying the data base or GIS (Accolla et al., 2020; Crooks & States, 2017; DeAngelis & Diaz, 2019; Schulze et al., 2017).

Agent, **Singleton** and **Resource** are the essential object types found in most environments and frameworks commonly used today, e.g., MASON, NETLOGO etc., and they continue to be recommended as basic components of an ABM (Crooks & States, 2017; DeAngelis

& Diaz, 2019; Fedriani et al., 2018; Grimm et al., 2020; Lippe et al., 2019; Liukkonen et al., 2018; Manson et al., 2020; Murphy et al., 2020; S. Railsback et al., 2017; S. F. Railsback & Grimm, 2019; S. F. Railsback & Harvey, 2020; Scherer et al., 2020; Schulze et al., 2017; Zhu et al., 2015).

9.4.3. Object oriented programming as a supporting programming technique

Reuter et al. (2011) considers object oriented programming (OOP) to be a practicable implementation technique for IBM, noting that “OOP has revolutionized computer programming due to its more flexible design structure and clear organization of programme code.”. This makes it easier to represent complex models, since “[...] OOP easily allows to handle the structurally complex interaction networks required for advanced ecological applications” (Reuter et al., 2011).

Reuter et al. (2011) describes the organization of the source code into distinct, delimited blocks as a characteristic of OOP. This is achieved by adhering to four important principles: (1) encapsulation, (2) data abstraction, (3) inheritance and (4) polymorphism.

(1) Encapsulation ensures that the object itself exclusively controls its own variables, and thus its own state. The internal variables are hidden to the outside and can only be accessed by the object’s accessors or manipulated (mutated) by the object’s mutators. Accessors and mutators hide the actual implementation details from the external caller. However, the degree of strictness of encapsulation depends on the design of the programming language used; for example it is less strictly available in JAVA and C++.

(2) Data abstraction is closely tied to encapsulation. It means that when accessors and mutators of an object create an interface to the internal state, they do not necessarily map a one-to-one relationship to the variable(s) of interest. Accessing a property, for example the property “health”, by an accessor may require the access to multiple (hidden) internal variables.

(3) The inheritance principle is an important characteristic of OOP that governs the reuse of source code; while (4) polymorphism provides a mechanism to override and overload operators and functions, i.e., through the use of functions of the same name but with a different signature (parameter set).

The main feature of OOP is the close relation between the encapsulated data (variables) and associated code for the access. This close relationship allows the creation of objects that represent autonomous individuals through their algorithmic description.

9. *The simulation of ecological models: Enhanced approaches and techniques*

The object-oriented system analysis (OOSA) approach and OOP are closely related, but the pre-structured object oriented (OO) way of formulating models by OOSA is only a first step towards implementing them in an OOP language. On its own, this procedure does not provide a solution to the problems of increasing (simulation) model complexity and its documentation. Additionally, it is more than 50 years since the introduction of the SIMULA67 programming language designed especially for the implementation of simulation models (Dahl et al., 1970; Nygaard & Dahl, 1978). Currently, no programming language with the appropriate features is in widespread use; nor has a standard for such a language been established that provides clearly defined methods and properties. As a result, the simulation landscape is fragmented.

9.4.4. **The State-based modeling and programming paradigm**

Benefits of implementing an automaton in ABM

State-based modeling (SBM) is based on automata theory and is considered as a “low-level” technical approach to simplifying the development and implementation of individuals (with complex behavior) in simulation models. An automaton comprises distinct defined states which may be seen as encapsulations of functional elements expressed as algorithms. Each of these elements comprises a set of functional behaviors. Each time an automaton is invoked, the algorithms expressing its current state are executed. This procedure provides regular updates of individual properties (Reuter et al., 2011) (see Subsection 9.4.5).

When an individual is represented by an automaton, an individual consists of manageable blocks of algorithmic expressions, each block representing a different behavior (state). Blocks are easier to manage when each represents a single behavior or trait, e.g., current phenotype, feeding, sleeping etc.

An automaton obtains a state through a defined sequence of events and thus exhibits a deterministic state change process. The expressed behavior is comprehensible due to this determinism.

In the following subsection, formal aspects of automata theory are explained using the example of a model of an individual firefly and its flashing. The execution logic required to run the model on a computer is also given. To facilitate conversion to a simulation model, the original conceptual model is already formulated as state based, i.e., as comprising clearly defined states and activities of individuals (and objects), occurring within clearly defined sets of circumstances.

States and transitions of an automaton

In a formal description (Hopcroft, 1979), a deterministic finite automaton (DFA) consists of a set Q of states that uses a transition function and input symbols to determine the next state. A 5-tuple $(Q, \Sigma, \delta, q_0, F)$, basically containing states, (transition) functions and events, represents a DFA where Q is a finite set Q of states, alphabet Σ the input symbols, $q_0 \in Q$ the initial state, $\delta(q, a) \rightarrow Q$ with $a \in \Sigma$ the transition function, and F with $F \subseteq Q$ the finite set of accepting states. A DFA starts in the initial state q_0 and stops in one of the accepting states in F . The DFA transits from its current state to a next state by an input symbol $a \in \Sigma$, e.g., an event.

Embedding state-dependent behavior, i.e., activities in states, and replacing the input symbols by events, would then create a finite state machine (FSM)-based object out of an individual DFA. Some simulation environments (Argonne National Laboratory, 2021; Lorek & Sonnenschein, 1999; SeSAm, 2012a) support a FSM-based modeling approach. The main advantage of this approach is probably that it allows the concise structuring of interactions based on the current state of an individual. The FSM principle can also be used to improve the object-oriented system analysis (OOSA) approach by including states as part of the modeled individual.

A finite state machine model by example: Synchronous flashing of fireflies

In this model, a single firefly flashes at a certain frequency while monitoring its surroundings for other flashes, which occur approximately at the same frequency. When a firefly perceives a flash, it advances its next flash by a small amount of time to synchronize its flash with the perceived flash. A minimum pause without flashing is forced after each flash. The same behavior of all individuals results in the synchronous flashing of the firefly swarm over time. This model is an adaption of the NETLOGO “Fireflies”-model (Wilensky, 1997) that is included in the NETLOGO simulation environment.

The firefly FSM consists of the three states $Q = \{\text{OBSERVE}, \text{FLASH}, \text{PAUSE}\}$ with the initial state $q_0 = \text{OBSERVE}$. Without an accepting state $F = \{\}$, the firefly cycles (transits) through the states forever. The alphabet Σ of “external events” is composed by the conditions $\text{flash urge} \geq \text{threshold}_{\text{urge}}$, $\text{flash elapsed time} \geq \text{flash duration}$ and $\text{pause elapsed time} \geq \text{pause duration}$, see Figure 9.1.

The firefly continuously increments variable flash urge and monitors the environment for external flashes in the state OBSERVE. For each perceived flash, an additional increment is added on flash urge . If $\text{flash urge} \geq \text{threshold}_{\text{urge}}$ then enter the next

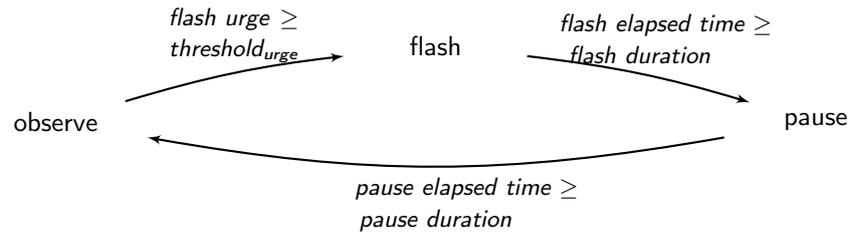


Figure 9.1.: Graphical description of the firefly behavior.

state FLASH. On entering state FLASH, the flash is turned on and the elapsed time in this state is counted in *flash elapsed time* until $flash\ elapsed\ time \geq flash\ duration$ applies. On leaving the state for the next state PAUSE, the flash is turned off. In the state PAUSE, the elapsed time in that state is counted in *pause elapsed time* until $pause\ elapsed\ time \geq pause\ duration$. The next state is again OBSERVE.

The state-based verbal description structures the model as a set of distinct states associated with state-dependent behavior. Since actions are state-dependent, a precise algorithmic description can be formulated that encapsulates state-specific variables and algorithms.

The definition of state-specific behaviors (activities) and variables minimizes the logical dependencies between the states, creates a more transparent program flow and reduces the overall complexity of the model by defining smaller and easier to manage units.

Figure 9.1 provides a visual representation of the dependencies between the states. Such diagrams are simplifications, designed to show the general structure of the automaton and its basic operations. Additional textual description is required to provide a formal specification of the algorithm corresponding to each state; however, these algorithms are not usually shown in the diagram for the sake of readability.

Extension and execution of the state model

Extending the states by phases. A SBM such as the above firefly model incorporate an enhanced code structure and program logic that is suitable for modeling less complex individuals, i.e., individuals with simple behavior. Each state serves a single purpose and exactly one state is active at a time. The program code assigned to a state encapsulates state-specific code and variables. In order to refine the algorithmic description and implementation, three phases, each with its associated code, are proposed (see also Klügl and Puppe (1998)):

- *onEnter*: invoked when the individual enters the state from a previous, different state, or as an initial state.
- *onIterate*: invoked to execute state-dependent actions in each simulation step.
- *onExit*: invoked when the individual leaves the state.

The phase `onEnter` provides state-dependent preparing code prior to invocation of `onIterate`. The phase `onIteration` executes the state-dependent actions and thus expresses state-dependent behavior. The phase `onExit` executes the state-dependent “clean-up”. The proposed phases define a superstructure or skeletal program code skeleton that supports that data and code encapsulation that define distinct states.

Defining a finite state machine for execution. To be executable, the states and transitions in a SBM must be defined by an algorithmic description, i.e., similar to the classes in OOP languages. Although ideally the model should consist of distinct, non-interfering states that encapsulate all required information and data for their execution, the superimposed structure of transitions from one state to another requires knowledge of all states to determine when and where state changes occur.

In this proposal, the transitions are defined inside each state, which requires a state to “know” of available next states. This is not a softening of the encapsulation principle of states, because technically the information about subsequent states must be defined somewhere. Placing outgoing transitions within a state has the advantage that this information is automatically deleted if a state is removed when modifications are made to the model. Incoming transitions still have to be removed manually by editing each state in the sequence of states because the action of leaving state affects the logic of the automaton. Thus, when an individual enters a new state, all the preceding states have to be modified anyway. This can be done by a compiler, who marks the outgoing transitions referencing non-existing states.

Exemplary implementation of states and phases. From the OO-perspective, an object itself decides when to change its state; all state variables and values are considered private information, known only to the state. However, technically an object is processed by a computing device that selects the next state. Algorithm 1-3 describe the firefly’s states and phases in a pseudo-code notation, while the state execution unit (Algorithm 6, Appendix E.1) controls the program flow. At the technical level, each object incorporates a method `nextState` to return its next state or, if no state change occurs, its current state.

Algorithm 1 State OBSERVE.

```
1: function ONENTER
2:   flash urge  $\leftarrow$  0
3: end function

4: function ONITERATE
5:   flash urge  $\leftarrow$  flash urge + 1
6:   if flash in neighborhood detected then
7:     flash urge  $\leftarrow$  flash urge + 1
8:   end if
9: end function

10: function ONEXIT
11: end function

12: function NEXTSTATE
13:   if flash urge  $\geq$  thresholdurge then return FLASH
14:   end if
15: end function
```

Algorithm 2 State FLASH.

```
1: function ONENTER
2:   flash  $\leftarrow$  on
3:   flash elapsed time  $\leftarrow$  0
4: end function

5: function ONITERATE
6:   flash elapsed time  $\leftarrow$  flash elapsed time + 1
7: end function

8: function ONEXIT
9:   flash  $\leftarrow$  off
10: end function

11: function NEXTSTATE
12:   if flash elapsed time  $\geq$  flash duration then return PAUSE
13:   end if
14: end function
```

Algorithm 3 State PAUSE.

```

1: function ONENTER
2:   pause elapsed time  $\leftarrow$  0
3: end function

4: function ONITERATE
5:   pause elapsed time  $\leftarrow$  pause elapsed time + 1
6: end function

7: function ONEXIT
8: end function

9: function NEXTSTATE
10:  if pause elapsed time  $\geq$  pause duration then return OBSERVE
11:  end if
12: end function

```

The `nextState` method implements transition functions inherent to each state, maintaining the strict encapsulation of logic and variables, so that each state functions independently of all other states. Potential side-effects, such as those caused by the sharing of variables, are minimized. Even when variables are shared, side-effects are limited as data may be accessed by only one state at a time.

The advantage of a state-based description and implementation. State-dependent code execution simplifies the tracing of program logic and flow, as only code relevant to the state in question needs to be considered. In the example of the fireflies, the controller turns on the flash on entering the state `FLASH` and turns it off on leaving. The turning off code is placed in the phase *onExit*, which is executed when the controller changes the state by invoking `Flash.onExit` and `Pause.onEnter`. The next state `PAUSE` does not need to worry about codes executed by previous states to ensure consistent object behavior. This is even more advantageous if a state has more than one possible preceding state.

If the flash is not turned off on leaving the state, i.e., if `Flash.onExit` is not incorporated in the state-dependent code, then all possible next states must implement the code to turn off the flash. Consequently, each next state must “know” all its possible preceding states and implement the dependent logic to maintain consistency. This leads to error-prone code duplication and more complex logic, including cross-references between states, resulting in a higher complexity. Moreover, the requirement for each state to know other states and their internal logic contradicts OO design principles, which are

9. The simulation of ecological models: Enhanced approaches and techniques

based on object-centered handling and control, without any form of external control, in order to avoid such cross-object logic problems. Logic, functionality, actions and data are encapsulated in states, thereby minimizing side-effects and facilitating the design of simulation models.

Implementing automata and states in a programming language

The implementation of a conceptual model by creating a computer simulation model in a programming language is not a trivial task, especially when this involves state-encapsulated programming. Popular and widely used programming languages, e.g., JAVA, C++, C# and C, currently do not incorporate language for state-encapsulated programming. Implementation of states in these programming languages requires auxiliary constructs, but these do not incorporate state inheritance as a necessary extension to the class inheritance that is a key feature of object oriented programming.

Some widely-used languages are more compatible with state-encapsulated programming. For example, an extension is available for the SCALA programming language (Scala Center, 2021) that supports state-based programming. SCALA allows the definition of constructs to create a domain specific language (DSL), which is a *custom* programming language designed especially to cover a designated field of application. The SCALA extension AKKA enables state-based modeling (SBM) using the FSM-model of the ERLANG programming language. AKKA contributes an underlying managing unit that does the necessary work to execute the states and determine the next state. Unfortunately, the AKKA extension is based on an actor model that increases the implementation effort for simulation models, while the ERLANG programming language, that supports this actor model at the language level, is not supported by common simulation frameworks.

In summary, the state inheritance required for OOP creates an additional layer of complexity, since the inheritance and overwriting of states and transitions must be defined and managed. Such a task may be expected to be far from trivial.

9.4.5. Ecological design patterns

The previous discussion outlined the practical issues involved with the implementation of a model at a programming level. This subsection discusses the potential of ecological design pattern (EDP) as an alternative approach to tackling these issues. This idea draws on software development that provides “reusable” solutions for known, recurring problems, in the form of design patterns that specify best practice to tackle a problem

but do not contribute specific source code. Some design patterns are bound to unique features of specific programming languages and, without modification, can therefore only be implemented in these languages.

As best practice, design patterns do not provide definitive solutions to problems based on objective analysis. Rather, each design pattern is the outcome of discussion of the problem it addresses. The development of design patterns resembles the “scientific method”, in which currently accepted solutions are continually tested and adapted and, in the future, may be discarded in favor of new solutions. If a design pattern often provides a solution to a problem, then this provides important clues as to the details of the problem. The study of EDPs may provide an ecologist with a set of analytical tools to identify recurring problems encountered in model development that may be solved by a standardized method.

Pattern-oriented modeling (POM) (S. F. Railsback & Grimm, 2012, 2019; S. F. Railsback & Harvey, 2020) differs from EDP in that it is applied during the initial development of a model rather than during the implementation of a (conceptual) source model. POM looks for (multiple) patterns in a real system, preferably across scales, as inputs for the design of a model of the system whose structure accurately represent these (multiple) patterns. A structurally correct model will reproduce the patterns of the real system for the right reasons (Grimm et al., 2005; S. F. Railsback & Grimm, 2012). “Each pattern serves as a filter of possible explanations so that after using, say, three to five patterns, we are more likely to have achieved structural realism.” (S. F. Railsback & Grimm, 2012). POM therefore considers validation of models instead their concrete implementation.

Similar to an EDP, Reuter et al. (2011) suggest a general program code structure for the implementation of an individual in a model that can be used “to update individual properties”. The suggested code structure resembles a schematic solution to a class of problems, whereby the elements of the code and their structure identify the general problem, as well as describing its scope and detailed manifestations, and thus enable the identification of “best practice” to solve any particular manifestation of the problem.

Taken together, EDPs should comprise solutions (or “updates”) to a collection of recurring problems in ecology models, e.g., related to properties of individuals, how these individuals sense the environment, simulation of diffusion and swarming, and the modeling of global properties like daylight etc. These precise nature of these problems and their schematic solutions (i.e., the design patterns) are likely to correspond to features of the chosen simulation environment (see the discussion in Section 9.6 and Section 9.7). This correspondence is inevitable because these features define the

9. *The simulation of ecological models: Enhanced approaches and techniques*

framework within which a simulation model can be created, constraining both the conceptualization and representation of solutions to problems in the model.

When a sufficient number of EDPs are identified and defined, these may evolve into a kind of “construction kit” covering many issues encountered during the implementation of ecological simulation models. Potentially, individual EDPs can function as building blocks, which can simply be lined up with “glue code” to form an almost complete simulation model. The MOBIDYC simulation environment (Subsection 9.6.3) adopts a similar approach, providing a set of configurable building blocks for model implementation, but confines the developer within a restricted range of predefined procedures.

Individual EDPs have a limited domain of application; therefore it is desirable that steps in their formulation should be standardized to make the design patterns more concise. Furthermore, in addition to their application for problem solving, EDPs can provide useful inputs for the development of the simulation model; their standardization may help the ecologists, who are usually not specialists in software development, to make the best implementation decisions. Standardization would reduce complexity of the modeling by limiting the choices available to modelers.

Candidates for standardization include EDPs related to objects, i.e., the data types Agent, Singleton and Resource (Subsection 9.4.2), which are presumed to be sufficient to map all objects required for an ABM. The classes comprise data fields whose primitive types and properties must also be defined, i.e., String, Float etc. The definition of these types helps identify “best practice” design procedures, for example explicit avoidance of Floats. Data structures, i.e., Maps, Lists, Trees etc., represent the next higher level of data. EDPs for these data structures would relate to their scope of application and certain characteristics such as memory usage or computing efforts.

The final stage of standardization could potentially involve development of a unified language to describe models and the ecological design pattern (EDP), as a framework for description and definition, that would facilitate reasoning and discussion of models by all parties in their implementation by removing ambiguity and reducing complexity.

9.5. Short comings and the path to a solution

9.5.1. Overview

The translation of a non-computational model, the hereby named “source model”, into a simulation model (formulated in programming language) is a major issue in ecological modeling because the formal description of the source model in a concise

language of a simulation model may differ in some respects from the source model. This section focuses on the “translation process” from source model to an executable simulation model. The “source model” may be purely conceptual (and is sometimes referred to below as “conceptual model”) or a fully developed model, which may be formulated in different ways, e.g., using words, diagrams or mathematical equations. A formal procedure is proposed that overcomes the problem of potential deviation of the simulation model from the source model and includes an automated verification process.

Methodical approaches are already available for structuring and documenting models, e.g., the Overview, Design concepts and Details Protocol (ODDP), and for documenting their execution, i.e., TRACE (TRANSPARENT and Comprehensive Ecological modeling documentation) (Grimm et al., 2014; Grimm et al., 2020). However, in addition to being rather abstract and offering high degrees of freedom in their use, these methods do not address the problems involved in implementing simulation models from non-computational models.

9.5.2. Problems with the model definition

Grimm et al. (2006) describes published ABMs as “[...] hard to read, incomplete, ambiguous, and therefore less accessible” and producing hard-to-reproduce results. To address these limitations, the Overview, Design concepts and Details Protocol (ODDP) (Table 9.1) was proposed as a protocol for development and description of ABMs (Grimm et al., 2006; Grimm et al., 2010). Grimm et al. (2006) criticize the lack of a standard structure for ABMs. This lack of standardization creates what they describe as an inefficient and cumbersome requirement to first read the full description of the model, in order to gain a general impression of its purpose, structure and processes.

The proposed ODDP tackles these issues by structuring the description of an ABM in the form of a defined sequence of topics (sections), with a checklist of information to be included under each topic. The NETLOGO simulation environment (Subsection 9.6.3) implements a similar approach, which is why several educational models were developed because of this structural similarity (S. F. Railsback & Grimm, 2012).

The ODDP is an attempt to reduce ambiguities in the description of the model, and to make models easier to read by defining a standard structure. However, the ODDP does not include a protocol for the translation of a non-computational model into a simulation model that incorporates a description of system dynamics. Since a static description in the form of a non-computational model may not adequately express the dynamics of a model, especially in the case of an ABM, the simulation model is often

Overview	Purpose
	Entities, state variables, scales
	Process overview and scheduling
Design concepts	Basic principles
	Emergence
	Adaption
	Objectives
	Learning
	Prediction
	Sensing
	Interaction
	Stochasticity
	Collectives
Observation	
Details	Initialization
	Input data
	Submodels

Table 9.1.: List of topics to be included in a (simulation) model in the most recent version of the ODDP (Grimm et al., 2010).

the only full description of the model (Grimm et al., 2020; Lorek & Sonnenschein, 1999). Without a clear protocol for deriving system dynamics from a static non-computational model, it may be difficult to verify the results of model simulations. This aligns with the general criticism of poor program documentation in published studies of ABM simulations, which is considered one of the biggest problems in understanding and maintaining source code (Shum & Cook, 1994).

Grimm et al. (2006) note that the Unified Modeling Language (UML) provides a specification scheme for addressing these issues, but consider UML to be too complex for use by non-software engineers. Grimm et al. (2006) propose that “ultimately, something similar to UML should be developed for individual-based and agent-based models: a visual declarative language that is easy to use and can directly be compiled to computer code [...]”.

The ODDP has been widely adopted (Grimm et al., 2020). However, the development of a readable, accurate and expressive programming language remains a crucial prerequisite for the readability and comprehensibility of simulations models.

In addition, the ODDP does not address the issue of the reproducibility of simulation results, the cornerstone of scientific work (Grimm & Railsback, 2005), because the

protocol does not consider interactions between the model and the simulation environment or framework. For example, if the source code contains a random element, then the repeated executions of a deterministic simulation model can still produce different results. A developer needs to be aware of such issues, identify the underlying algorithms, and provide accurate documentation.

9.5.3. Implications of the lack of a common language

The process of model implementation is the translation of a “human language” model into an executable simulation model expressed in another (programming) language. “One of the most important purposes of a platform is to provide a common language for thinking about and describing ABMs.” (S. F. Railsback et al., 2006). The ODDP, which was designed to facilitate the implementation of such models, does not address this need for a common language. In general, the use of different languages in model descriptions may result in discrepancies between model description, model implementation and model documentation.

Common tools for the implementation of a simulation model include libraries, frameworks and simulation environments. Limits on the expressiveness of these tools might constrain the development of the simulation model and limit its ability to reproduce all the features of the source (ecological) model. On the other hand, adequate tools may reduce the workload, because no workarounds, e.g., adaptations of the model to match features of the tool, have to be developed.

At the technical level, the environment or framework usually constrains the choice of programming language and vice versa. An appropriate combination of framework/environment and programming language(s) is key to the successful implementation of a simulation model and the ease of its subsequent execution. This combination determines the additional workload required to translate the non-computational model into a simulation model and the extent to which the source model has to be adapted to match the capabilities of the simulation model. The closer the language that describes the source model is to the language used in the technical system for implementation of the simulation model, the less translation effort is required.

9.5.4. Desired features of ABM development

Supporting features required by a simulation model of *D. gigas* include (1) an ABM architecture to support agent programming, (2) CA definition and programming for resource simulation, (3) discrete space and time handling, (4) a discrete event handling

9. *The simulation of ecological models: Enhanced approaches and techniques*

mechanism, (5) state-based agent programming support, and (6) data storage for ex-post analysis of the simulation data. Further requirements may include the visualization of interactions between agents and environment, which would provide insights into the inner workings of model and inputs for validation of the model (Lorek & Sonnenschein, 1999). A scene editor to set up the initial simulation state by populating the simulation space with agents, resources and environmental factors is another useful feature.

The feature wish list of S. F. Railsback et al. (2006) includes accurate documentation, examples and templates to facilitate setting up the model, integration of the applied framework into integrated developing environments (IDEs), integration of the applied framework into the simulation development flow, procedures for dealing with complex source models, tools for statistical output, tools for setting up and executing simulations, reduced trade-off between ease of use and generality, and technologies for testing, analyzing and understanding ABMs.

Lorek and Sonnenschein (1999) describes three general types of software tools to support ABM development: (1) Libraries or framework systems that function as extensions to general purpose programming languages, (2) easy-to-use but relatively inflexible tools for specific kinds of models, and (3) existing applications (simulators) developed for a specific domain or application. The next two sections examine the potential of the first two of these types of tools. Simulators are not considered in the following discussion because of their applicability is limited to a specific domain.

9.6. Simulation environments

9.6.1. Introduction

A simulation environment allows easy and fast implementation of a model if it provides the appropriate structure and flexibility (Subsection 9.5.4). However many such tools “do not speak the language of ecologists” and contain too many features that can only be used by experienced programmers (Lorek & Sonnenschein, 1999). Reuter et al. (2011) advocate not using simulation environments.

Subsection 9.6.2 discusses the restrictions and problems of implementing simulation models in simulation environments. Subsection 9.6.3 discusses the properties, including features such as state-based programming, of selected simulation environments and prepares the ground for the discussion for a program-based approach using frameworks and a regular programming language in Section 9.7.

9.6.2. Implementation of simulation models in simulation environments

Model configuration using existing simulation elements reduces the implementation effort and may help unexperienced programmers to set up a simulation model, but this approach constrains model capabilities within limits set by the simulation environment. Configuration provides a well-defined structure for simulation but can be disadvantageous when modeling complex situations, whose simulation requires model extensions that exceed the inbuilt capabilities of the predetermined system configuration. In general, simulation environments comprised solely of pre-configured elements are likely to have a limited field of application and to be too inflexible to support sophisticated simulation models (Reuter et al., 2011).

Where a simulation environment lacks features required by the model, either the model has to be adapted to be compatible with the existing features of the simulation environment, or the missing features are added to the environment by the modeler. In the latter case, an extension scheme is required that interfaces with internal data structures and the program flow of the simulation environment. The missing features have to be implemented using a programming language by an experienced modeler following procedures similar to framework-based model development (Section 9.7). To a great or lesser extent, this entails a switch from a pre-determined configuration to a model written in programming language, using a simulation framework provided by the simulation environment. Of course, these extensions have to be installed on all machines running the simulation.

A simulation environment constraints and binds the developer to its inbuilt functionality. When a simulation environment is no longer available, i.e., as a result of discontinuation of support or ongoing development, a simulation model may be lost, become inexecutable, or be incapable of further refinement. There is thus a risk that resources invested in model development may be lost; this risk makes simulation environments less suitable for developing of large and complex models.

A possible alternative would be to develop an open simulation model standard that defines a set of instructions, data and information structures that could support any simulation model. In this case, each simulation environment would need to support all encoded features to ensure full interchangeability. An essential requirement, to ensure reproducible output, is that each simulation environment should implement these features in the same way. The standard would include a programming language for agents and their environment, which would need to be carefully chosen to avoid future restrictions on model development. To minimize this risk, the potential for additional functionality may be integrated into the model, in the form of extensions that are

delivered with the model and are implemented within basic features of the simulation environment. EDPs may provide inputs for the development of such a standard.

9.6.3. Discussion of available simulation environments

An extensive survey by (Nikolai & Gregory, 2009) categorized ABM systems that existed at the time. Some of these are now obsolete, but the survey also considered tools that have stood the test of time and are still in use at the time of writing such as NETLOGO, ANYLOGIC and shell for simulated agent system (SeSAm). Manson et al. (2020) list open-source ABM toolkits with a focus on geographical explicit models. This list includes NETLOGO as well as some obsolete environments, and also MASON (Luke, 2019), which provides the underlying framework for the simulations in this thesis (see Subsection 9.7.2). Murphy et al. (2020) basically do the same (but list ANYLOGIC falsely as open source) and explicitly recommend NETLOGO for ecologists. The following discussion considers four simulation environments MOBIDYC, NETLOGO, SeSAm and ANYLOGIC, chosen as representative of different approaches to create ABM simulations and/or because they are supported by an active user community.

MOBIDYC. MOBIDYC (“MOdelling Based on Individuals for the DYnamics of Communities”) is a “[...] software project that aims to promote Individual-Based Modelling in the field of ecology, biology and environment.” (INRA, 2008). It allows people without computer skills to build and run ABMs (Ginot & Le Page, 1998). The project was designed to meet the needs of ecologists.

MOBIDYC works by configuration rather than by programming (Razavi et al., 2005). It takes the recurring actions of ecological models and structures them into so-called “primitives” (Ginot & Le Page, 1998). The configured primitives are assembled by users within tasks in the required order. The environment uses a unified object-based model and implements the essential object types listed in Subsection 9.4.2.

The MOBIDYC environment contributes the basic ingredients for an ecological ABM (Ginot & Le Page, 1998), but sets tight constraints for the definition of agents and their behavior (Ginot & Le Page, 1998). Advanced SMALLTALK skills are probably needed to program extensions.

MOBIDYC may be considered as an interesting experiment (Grimm & Railsback, 2005) that succeeds in creating a domain-specific simulation environment. However the tight constraints of the specific domain (S. F. Railsback et al., 2006), the lack of portability of the models created, and the fact that ongoing development has ceased disqualify MOBIDYC for further use. These points align with the arguments of Reuter

et al. (2011) (Subsection 9.4.1) against reliance on predefined software tools. However, it should be noted that while ongoing development of the original version of MOBIDYC has ceased, an open source alternative “ReMobidyc” is currently (July 2021) under development (Oda, 2021).

NetLogo. NETLOGO is a long-term project that is still in active development and is supported by huge user community. NETLOGO provides a complete simulation environment for ABM, including a programming language, a setup editor and a graphical user interface with controls, observers (inspector) and console output, but currently without a debugger. NETLOGO has been used to develop a wide range of models, including ecological ones.

NETLOGO evolved from STARLOGO. Like STARLOGO it was originally intended as an educational tool but “it has steadily been adopted by the research community” (Kornhauser et al., 2007). Positive reviews of NETLOGO mention its maturity, documentation and execution speed (Lytinen & Railsback, 2010; S. F. Railsback et al., 2006).

The straightforward graphical user interface consists of the panes (tabs) INTERFACE, INFO and CODE. The INTERFACE pane contains the (visual) output of the model and its controls. Controls and displays are addable and editable. The INFO pane displays the model description based on a contributed template (see Appendix E.2.1).

The documentation is edited in the MARKDOWN markup language (Gruber, 2004), that contributes basic text formatting capabilities; this is translated and displayed as a HTML (Hypertext Markup Language) document.

The NETLOGO programming language is a LOGO dialect that has been extended to include rudimentary OO-capabilities (Subsection 9.7.3). A unique name is assigned to each element of the INTERFACE pane and is automatically accessible in the code segment as a value or reference. Each INTERFACE element may trigger code segments in the CODE pane. Values set in the INTERFACE tab change the corresponding values in the program code and these values can be displayed in different formats, e.g., as a plot.

This close connectivity between INTERFACE and CODE panes makes it easy to create interactive simulations, which is why NETLOGO was chosen by S. F. Railsback and Grimm (2012) as the most practical approach available for ecological modeling. However, the origins of NETLOGO as an educational tool limit its suitability as a scientific tool. The creation of extensions to its predefined structure requires advanced programming skills, while the NETLOGO programming language is still rudimentary and lacking

9. *The simulation of ecological models: Enhanced approaches and techniques*

important OO-features (Subsection 9.7.3). Moreover the lack of a debugger renders the development of complex simulations tedious.

SeSAM. The shell for simulated agent system (SeSAM) (SeSAM, 2012a) was originally developed at the Department of Artificial Intelligence in Würzburg, Germany (Klügl & Puppe, 1998) and was hosted by the Interactive Simulation Research Group at Örebro University, Sweden, until the end of 2020. The powerful and flexible architecture of SeSAM provides a skeleton structure for the development, management and handling of plug-ins.

Without optional plug-ins, the basic configuration already incorporates the essential object types (Subsection 9.4.2). A GUI (graphical user interface) handles agent definition, the simulated environment, data logging and data analysis. SeSAM thus fulfills the above-mentioned feature wish list of S. F. Railsback et al. (2006).

The programming of the agent behavior is structured by nodes (states) and linking transitions, which form a SBM (Subsection 9.4.4). There is exactly one state active at a time and, for each node, routines for entering, leaving or periodically executing are defined (see Appendix E.2.2). But while the concept is advanced, the user interface is tedious to use.

The last update (May 2012) of the wiki documentation (SeSAM, 2012b) and the creation date of the current SeSAM version 2.5.2 (January 2012) indicate a cessation of development and maintenance. This, together with the tedious handling of model programming, makes SeSAM a poor choice for simulation model development; moreover, SeSAM advanced model features cannot be exported to other simulation environments.

In summary, SeSAM is more a research project than a simulation environment intended for widespread use. SeSAM is more advanced and powerful than NETLOGO, but with significantly less usability and as a result, lower levels of acceptance.

AnyLogic. The ANYLOGIC software is an advanced commercial product that runs on most popular computer platforms. ANYLOGIC can be used for a wide range of applications, providing ready-made libraries e.g., for pedestrian simulation or to support GISs. ANYLOGIC is a simulation environment for ABMs that provides discrete event simulation and system dynamics. UML-like diagrams as in SeSAM are used for programming. Extensions implemented in JAVA allow the extension of the ANYLOGIC capabilities. Its extensive debugging capabilities support the observation and manipulation of the current state of an object. The development, setup and control of simulations utilize a GUI with editor, setup editor, visual controls and visual elements.

The ANYLOGIC system provides full documentation, examples, tutorials and support for the large community of users.

The UML-like agent programming model includes a state-based programming paradigm similar to SeSAM, but with a more elaborate user interface. The agents, organized in a tree-like structure, consist of properties and functions and support inheritance. Educational versions are available but with limited features, so ANYLOGIC may not be the first choice for researchers. Despite its advanced features, ANYLOGIC is a proprietary tool with a model format that binds the developer to the predefined simulation environment.

9.7. Frameworks in combination with a programming language

9.7.1. Introduction and overview

As discussed above, all available pre-designed simulation environments have limitations and may not meet the requirements for the design of an ABM. This section considers the alternative approach, which is to design a tailor-made ABM using a programming language in combination with a framework or library.

Most programming languages do not support the discrete event simulation required for ABM. To be usable for this purpose, they require an extension in the form of a framework or a library. These options are discussed in Subsection 9.7.2.

Subsection 9.7.3 discusses the choice of programming language. It makes sense to look first at widely used programming languages with good community support and maintenance. However, these widely-used languages may not be suitable for the development of ecological simulation models.

9.7.2. Frameworks and libraries

A simulation framework should provide ABM features, i.e., an event handler, visualization, data logging, user interaction etc. Compared to a simulation environment, working with a programming language gives the modeler more control and more degrees of freedom (Appendix E.3.1). The trade-off is the requirement for sophisticated programming skills.

A framework interfaces with a limited number of programming languages, so an appropriate combination must be selected at the outset (Subsection 9.7.3). Similar to models created using a simulation environment, a framework-based simulation model may become inoperable if the framework it uses is discontinued. However, in this case, the effect is likely to be less severe, since the majority of programming is done in a

9. *The simulation of ecological models: Enhanced approaches and techniques*

common programming language, so “only” the functionality provided by the discontinued framework has to be substituted.

Several frameworks and libraries are reviewed in the literature (Bordini et al., 2006; Manson et al., 2020; Murphy et al., 2020; Nikolai & Gregory, 2009). However, a review undertaken for this thesis found that many of these have not been updated for years and are not supported by an active user community. Thus the current status of many ABM frameworks must be regarded as unknown. Most frameworks created at educational institutes lack continuity and are discontinued after project funding ends.

Frameworks are implemented in different programming languages and cover different application areas. To provide inputs for the discussion of programming languages in Appendix E.3.1, only the Java Virtual Machine (JVM)-supporting frameworks, representing about 42% of the environments reviewed by Nikolai and Gregory (2009), are considered here. S. F. Railsback et al. (2006) reviewed and analyzed the REPAST (Argonne National Laboratory, 2021) and MASON (Luke, 2019) frameworks, both supported by the JVM, using each framework to implement the same set of simulation models. The authors describe both these frameworks as “mature”. They find that the MASON framework is faster and offers more functionality, but consider that its complexity would be challenging for novice programmers (S. F. Railsback et al., 2006).

MASON, implemented in JAVA, is available under open-source license, so it is extensible if required, and its functionality is transferable to another framework. Any programming language that compiles to JVM should be able to interface with MASON (see Appendix E.3.1).

Different frameworks may produce different results, even for identical models and parameter sets. This is because frameworks may implement the same functions in different ways. In these circumstances, reproducibility, the cornerstone of scientific work (Grimm & Railsback, 2005), cannot be guaranteed. The solution to this problem may be to define (and agree on) the essential functionality of a simulation framework. This functionality would need to be precisely defined, including the details of how it is implemented, to guarantee deterministic results for the same parameter set across programming languages and computer platforms. For simulations in this thesis, an abstraction layer was developed, that hides details of the implementation of MASON and provides the programmer with a unified application programming interface to access the functionality that is essential for simulation model development.

The list of essential functions may be kept small, as only a few data structures, i.e., lists, hash tables and associated algorithms, have to be implemented in all models. Additional functions required for ecological models would include, for example, distance computing,

filtering and sorting by distance etc. Sophisticated data structures, like predefined agents and CA would complete the ecologist's daily programming requirements.

9.7.3. Choosing the simulation environment and programming language

NETLOGO lack certain key features like a single step debugger (Lytinen & Railsback, 2010). Moreover, NETLOGO's LOGO dialect has only limited OOP support and an advanced module concept to load functionality or split source code into multiple files, that is found in professional development environments like ECLIPSE (Eclipse Foundation, 2021) and supported by other programming languages, is not available. These missing features limit NETLOGO's field of application. For complex models, NETLOGO's limitations place additional demands on the developer, so this simulation environment may be only suitable for simple models.

For complex models, the combination of a framework with a mature OOP language and a more sophisticated development environment like ECLIPSE would probably be a more appropriate choice (Reuter et al., 2011). However, for most ecologists, this would require additional engagement with computer science (S. F. Railsback et al., 2006; Reuter et al., 2011) and, initially, a steep learning curve.

JAVA and the open source MASON library would be one such combination. State-based programming is currently supported neither by MASON nor by JAVA, but JAVA is a complete OOP language with a set of powerful development tools. JAVA compiles to the JVM, that provides convenient memory management and high execution speed, and has the major advantage of being widely available on many platforms.

Modern program language concepts, like functional programming (FP), incorporate more powerful features, only a few of which are integrated in JAVA (see the discussions of advantages of functional programming (FP) in Appendix E.3.1 and the disadvantages of JAVA in Appendix E.3.2). Thus, use of a FP language that compiles to the JVM, e.g., the SCALA programming language, which incorporates functional and imperative programming paradigms, may be a better choice than JAVA.

At the time of writing, the programming language JavaScript (ECMA, 2021) is supported by the major internet browsers and a range of computer platforms, including mobile devices. However, as a script programming language, JavaScript is designed for small programs. Moreover, extensive debugging capabilities are supported by only a few (commercial) IDEs. In summary, some feature of JavaScript make this language unsuitable for the larger programs required for ecological modeling.

Efforts are underway to overcome these limitations. E.g., TYPESCRIPT (Hejlsberg, 2021) is a superset of JavaScript that attempts to address the major issues while

maintaining full compatibility with JavaScript. As an alternative solution, special ports of programming languages such as SCALAJ (ScalaJS, 2021) or KOTLIN (Jetbrains, 2021) combine the strength of their programming language design and the portability of JavaScript by compiling to JavaScript and allowing execution in a browser.

9.8. An advanced approach to model to simulation model translation

9.8.1. Introduction

Model development and simulation model implementation are currently separated step performed using different tools and languages. The encoding process, i.e., the translation of the source model to the simulation model, has to ensure the equivalency of both models. If the translation process is flawless, then the simulation model is an executable version of the source model and a subsequent validation process confirms that the model works as expected. Thus, unambiguous definition of the source model is an indispensable prerequisite for successful implementation of the simulation model. In this case, the deterministic nature of the programming language avoids the introduction of ambiguities in the formulation of the simulation model.

Clearly, defining the source model using a description language that resembles programming languages would make the translation process into a simulation model more straightforward, or even automatic. An automatic process would be the gold standard as this would guarantee equivalency of source model and its executable version.

This idea of a unified description language has already been partially realized. Cornell et al. (2019) present an analytical tool executed in MATHEMATICA which can be used to generate (differential) equations to describe the stochastic and spatial (emergent) effects of an ABM. To this end, (1) the verbal source model is transformed into (2) a component-based model by means of an unambiguous computer-readable language. This in turn is transformed (3) into a “computer-readable model specification” in MATHEMATICA. From this computer-readable model specification, the system optionally generates (4) a simulation model (computer code) or (5) differential equations using computer symbolic algebra. These equations may be (6) numerically approximated, or (7) subjected to further computer-based analysis. This system uses an unambiguous model specification (Step 3) to generate an executable simulation model.

Mooij et al. (2014) adopt a similar approach for application to models defined by Lotka-Volterra equations. The authors process information from a database that holds information about the model and is augmented with additional information about the programming language and target platform. According to the authors, information

presented in tabular form in scientific papers is usually sufficient as a data base for augmentation. However, this method might not be powerful enough for application to complex ABM.

Knowing the past approaches, this section discusses the literate programming paradigm as an alternative way to tackle the problems of source model definition, model documentation and simulation model implementation.

9.8.2. Model description using literate programming

Shortcomings of the ODDP

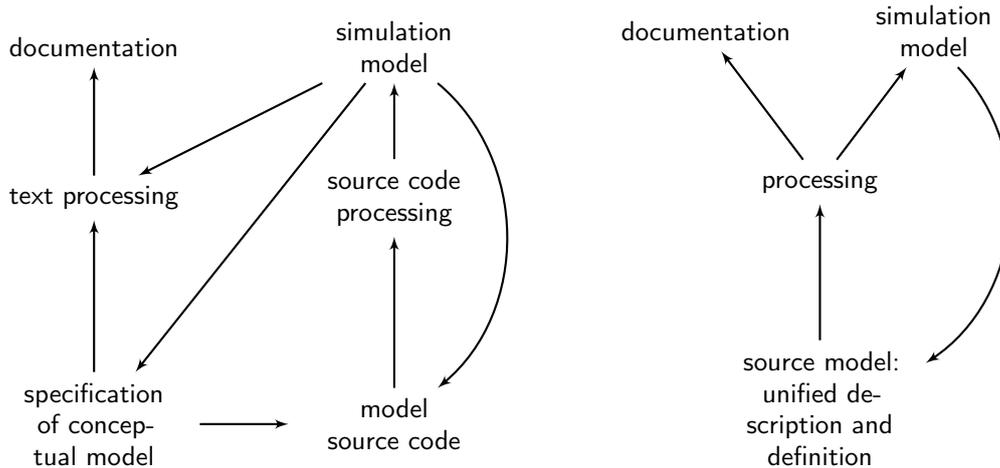
An unambiguous definition of an IBM or ABM describes all model components, interactions and properties, but not the resulting model dynamics during execution. The ODDP (Grimm & Railsback, 2005; Grimm et al., 2010; Grimm et al., 2020) fails to address this problem, since it does not incorporate rigorous procedures for describing model dynamics. Moreover, it would be difficult to describe model dynamics using static media of the kind employed by the ODDP, such as written language or diagrams.

An ODDP update (Grimm et al., 2010) suggests using a pseudo-code description of all essential parts of an IBM or ABM, including dynamics, but the author of this thesis is not aware of any commonly used standard for such a description. Grimm and Railsback (2005) explicitly do not recommend UML (that can describe model dynamics using diagrams), since this may be over-sophisticated for application to the simple rules of an ABM, while the computer-aided UML-notation requires special tools for creating the diagrams.

Another shortcoming of ODDP is the use of different independent sources (Figure 9.2a) to create the model documentation, e.g., using a word processor, and the simulation model, e.g., using a programming environment. (This limitation is addressed in Grimm et al. (2020), which was published while this thesis was being written.) Changes in one source may lead to discrepancies with other sources that are not automatically resolved. A unified source, from which model documentation and simulation model can be derived (Figure 9.2b) (also now proposed by Grimm et al. (2020)), solves this problem because it requires the developer to include the source code, or a description of the source code, into the model documentation. A unified source model may consist of several files but is still the only source from which both the simulation model and the documentation are generated.

A (new) standard for development and documentation of the source model should therefore focus on eliminating ambiguity in the model description. Ambiguity is excluded

9. The simulation of ecological models: Enhanced approaches and techniques



(a) The standard procedure from developing a simulation model from a conceptual source model incorporates multiple pathways and iterations.

(b) LP starts from a unified source model, which is used to model to create the documentation and the simulation model.

Figure 9.2.: Development of a simulation model from a source model (a) following the standard procedure (left) and (b) using literate programming. Both methods incorporate the iterative refinement the simulation model until maximal correspondence with the source model is achieved.

from the description of the subsequently derived simulation model by the use of source code; otherwise the results of executing the simulation model would be non-deterministic and not reproducible. The greater similarity of source model and simulation model leads to reduced ambiguity. The ODDP is deficient in this respect: it does not include a requirement for concise documentation because it incorporates too many degrees of freedom, specifying only the type of information required.

Literate programming and its benefits

Literate programming (LP) (Knuth, 1984) is an alternative approach that is compatible with the requirement for concise model definition and documentation. This approach involves first writing an “enriched” documentation (in this thesis named “source model”), that includes the ideas, concepts and embedded source code for the program (i.e., the simulation model). Then, in a second step, from that enriched documentation, the LP can automatically generate (1) the program (source) code, and (2) an explanatory documentation to help users to understand the program. The LP paradigm “[...]”

recognizes that two different target groups, human readers and compilers, will receive the program.” (Shum & Cook, 1994).

LP is applicable to many domains and is “especially useful in academic settings” (Spotnitz, 1998). The LP approach conforms to the idea of complete model documentation that also underpins ODDP by first defining and describing the source model and then implementing the simulation model. Transferred to ecological modeling, the advanced feature of LP is its capability to generate code for the executable model from a documentation that specifies a unified source model in natural language that, in combination with an embedded source code, can also be used for simulation modeling.

LP requires full program definition before creating the source code for the executable program, therefore LP “insists” on complete documentation that complies with program requirements. Based on the close similarity between the language used to describe the source model and programming language—such as in the definition of classes and data—a template for the model structure, definition and data required can be developed. It is advantageous if the documentation is written in a programming style syntax, which then allows the precise formulation of the documented model.

A programming style syntax allows the concise definition of agents and their behavior. Thus, LP is a meta-programming language for creating the source code that generates the final documentation and the code for the executable simulation model. It can be used with any programming language (for the executable simulation model) and the final documentation can include any kind of diagram, verbal description, or image, etc. Furthermore, this approach is not restricted to documentation using static media; dynamic documentation content could also be generated.

The literature programming text source code structure

The LP approach as originally developed used a single unified text file containing the verbal description (Knuth, 1984). The use of text files is not a dogma in LP, but use of other inputs, such as graphics, increase complexity because these inputs must be translated into literate programming source code. A text-based description facilitates program development, since text is closer to a programming language.

The distinction between text and program source code definition requires a domain specific language (DSL) to define these elements. For example, a DSL can operationalize the semantics of any text, e.g., program code, by embedding commands using a markup language such as \TeX , XML (Extensible Markup Language) or HTML. However, a text markup language is not a programming language, as in the original LP approach where the source code of an existing programming language is embedded in the program

description. Ideally, a DSL would be used to define and create the simulation model and documentation from a unified source that contains the full information. Otherwise, output from this unified source (i.e., documentation and the simulation model) will contain only a subset of the full information required to execute the simulation model. The DSL contributes an abstract but concise syntax to the source model definition, which is subsequently incorporated in the documentation and the simulation model. The retrieved output may need to be processed by a compiler to obtain the intended output (i.e., an executable simulation model). An ecology-focused LP therefore consists of a *text description scheme* for marking up the documentation text and a *model description language* that is used to define the conceptual model.

Tools like ANTLR (Parr, 2009, 2013, 2017) or XTEXT (Itemis, 2020) considerably simplify the non-trivial task of DSL development. ANTLR and XTEXT describe a (programming) language using a $LL(k)$ -grammar (left-to-right leftmost-derivation with lookahead k). $LL(1)$ -grammar based (programming) languages process only the next token (Hopcroft, 1979; Parr, 2009), where the token is a semantic unit such as a literal, keyword etc., from the input stream for the translation process. A $LL(1)$ grammar language is sufficient for the definition and documentation of simulation models, since programming languages such as SCALA, JAVA or C are based on a $LL(1)$ grammar. The DSL will require adaptation to be suitable for application to ecological source models. Such an adapted DSL is hereafter referred to as an ecological domain specific language (eDSL).

9.8.3. The translation process

The eDSL enables formulation of a meta-description of the model, i.e., a complete documentation of all aspects of the model, including all states, values, variables, definitions, calculations and descriptions. Without a syntactically complete description, an eDSL-based tool is unable to perform the translation of the source model into a simulation model. This requirement forces the developer to create a syntactically complete model description. However, for practical reasons, an eDSL-based tool should be able to process incomplete descriptions and use them for model documentation.

The output of the automated translation process (Figure 9.2b) consists of documentation and a simulation model. The output may require further processing by external programs, for example, if the documentation output is written as a \TeX project file or if the output for the simulation model requires compilation because it is defined in a programming language.

9.8. An advanced approach to model to simulation model translation

The use of a unified file, which contains all the information needed to create the documentation and a simulation model, excludes the possibility of interference or manipulation and is very different from conventional approaches: Use of a unified file means that the derived simulation model matches the model documentation because the model documentation is the only source. A model verification step is therefore unnecessary, since the simulation model is a direct translation of the source model. If the model definition (i.e., the model documentation) is incomplete, no translation is possible. However, model validation still has to be performed.

An automated translation process can target any output file format and any presentation style can be used for documentation (tables, graphs, references, etc.). Manual translation of the model description into the program source code is not necessary, since the program source code is already defined by the meta-description created by the eDSL.

9.8.4. LP in the future practice

Overview. Simulation modeling (in theoretical ecology) is a potential field for application of LP, due to the programmatic orientation of this approach. However, no LP-based process for ecological simulation modeling has been developed, even though both individual- and agent-based-modeling and LP have existed for decades. The author of this thesis is not aware of any simulation environments that employ LP. Previous and currently active LP projects are mainly focused on programming and program documentation.

Further, LP does not appear to be a current topic of discussion in theoretical ecology; LP is mentioned anecdotally in Grimm et al. (2020), but not discussed in detail, although the authors in Grimm et al. (2020) have been involved with simulation development and documentation for some time. It is possible that the potential of LP for agent-based-modeling has not yet been recognized.

If simulation models are to meet the requirements of scientific work in the future, then the models must be adequately documented, as is standard scientific practice; but additional requirements may also apply.

Adaptation of LP to simulation models. In its original version, LP, like ODDP, generates only static model documentation. Thus the dynamic component is missing, which is the essential feature of a simulation model.

Grimm et al. (2020) and Lorek and Sonnenschein (1999) note that the documentation attached to a simulation model is often the only documentation of the model (see

9. *The simulation of ecological models: Enhanced approaches and techniques*

Figure 9.2a). Therefore it would be logical and consistent to take a fully described simulation model as the starting point for documentation. Based on this, all necessary dynamic and static documentation could be generated, because all information would be available in one source, as provided in LP.

For example, MATHEMATICA (Wolfram Research Inc., 2021), MAPLE (Maplesoft, 2021) or JUPYTER (Jupyter Project, 2021) are tools which allow programming of dynamic calculations and texts in interactive “notebooks”. These tools can be extended via modules, so that they can be used for further applications. Currently these tools are in principle able to provide LP for simulations based on differential equations. Dynamic documentation would require development of a dynamic version of LP, i.e., DynLP.

Theses on LP. The extension of LP to create DynLP could transform the use of simulation models for ecological research, as summarized in the following theses:

Documentation of simulation models and models must be dynamic A simulation is dynamic, i.e., the state of the simulated system changes over time; thus the execution of the simulation expresses the characteristics of the model.

A static documentation thus lacks the expressiveness (complexity) required to document a simulation model. Even the program code, which describes the simulation model completely and unambiguously, is static until it is executed (on a computer) and does not represent the dynamic behavior. Documentation must therefore be dynamically executable.

Documentation and execution of a simulation model cannot be separated A simulation model consists of elements and their interactions, therefore the documentation also requires the description of these elements. The dynamics of these elements, i.e., the alteration of the values of their properties, could be clarified by means of supplementary material in the form of examples. These examples would also be simulations, showing the dynamic characteristics of an element under given scenarios.

The examples could contain interactive elements, e.g., as in MASON or NETLOGO, or integrate (predefined) events at certain points in time. The examples thus resemble test scenarios in software development, where, program segments are given certain inputs and the computed results are compared with expected results. Analogously, the behavior of individual model elements could be automatically

9.8. An advanced approach to model to simulation model translation

tested to see if it conforms to expected behavior. The (executable) model documentation then comprises the program code of the (tested) simulation model, test scenarios, and a (conceptual) description of how the model is intended or expected to behave. The documentation could also contain dynamic elements, i.e., variable text that changes in response to adjustments to test scenarios (examples), and inputs from readers.

Executable documentation requires scientific acceptance Gaining acceptance of such documentation in the scientific community could be challenging, because the dynamics of this form of documentation would be similar to that of an internet source and would require an approach to publishing of documentation of simulation models different in some respects from publishing in traditional (static) media. Dynamic documentation should therefore include a “publication configuration” with preset values for use in the published description of the model; this publication configuration would then correspond to a static text.

The documentation and the documentation system form a unit, so the documentation system itself would have to be clearly defined so that it can be moved to different platforms if necessary.

DynLP is beneficial to the scientific process in simulations Interactive documentation would be the basis for a new form of publication, which would open up new possibility for peer-review of the model, the generated data and the results. Through interaction with users (i.e., peer reviewers), the model can be thoroughly tested. In principle, users could also apply the model to test their own hypotheses. These possibilities for testing the model go far beyond what is possible in the traditional peer-review process.

Future challenges for DynLP. A dynamic documentation using DynLP includes an (interactive) simulation, which requires a standardized simulation system embedded in the document. This is similar in some ways to portable document format (PDF), because in both cases the document is prepared for display by a render engine.

In addition to the standardized simulation environment, there is also a need for an appropriate description language to define the model, i.e., the proposed eDSL that incorporates the essential (class) types, methods etc. required for describing and defining the source model. Furthermore, for the import of extension modules (if required by the document) a version management and documentation of these extension modules would be necessary.

9. *The simulation of ecological models: Enhanced approaches and techniques*

In summary, the development of a DynLP system would require considerable efforts and a high degree of technical expertise. It might also take time for the innovative approach to modeling and hitherto unfamiliar dynamic documentation to gain acceptance among the scientific community.

9.9. Summary and conclusions

The discussion in this chapter shows the potential for improvements in the development of simulation models. A number of different proposals are reviewed. For example, Reuter et al. (2011) argue that simulation models should be developed by ecologists without using off-the-peg simulation environments (such as NETLOGO) or programming experts, but also emphasize the need for complete documentation because of the complexity of ABMs. However, the increasing complexity of the models, the rigorous documentation requirements, and the technical challenges involved in the implementation of the simulation model mean that models developed by ecologists (and other non-specialists) are likely to be substandard in terms of accuracy, documentation and understandability. In this sense, the ODDP is not a workable solution; it simply places an additional workload on ecologists.

One suggested way to reduce the complexity of the process is to define individual states prior to the full implementation of the model. This approach starts by defining the different states of objects (individuals) and the corresponding state-dependent behavior. This allows the development of state-specific code and encapsulation of data as in OOP, which reduces possible side-effects. Furthermore, this approach obliges developers to ask questions about model structure and implementation at an early stage. By breaking down the definition of agents definition into small, manageable units, it can be helpful for structuring complex objects, both in the development of conceptual source model and subsequently when implementing the simulation model.

Such a state-based modeling (SBM) based approach is supported by the commercial ANYLOGIC simulation environment but the use of such off-the-peg tools is generally opposed by Reuter et al. (2011). The type of programming recommended by Reuter et al. (2011) requires the combination of a programming language and a framework to provide simulation model-specific functions. A conclusion of research for this thesis is that the programming language and framework used for ABM should be attached to the Java Virtual Machine (JVM) since the JVM is available on many different computer platforms. However, common programming languages and simulation frameworks to implement ABMs do not support the SBM approach.

Another approach that has been suggested is the development of ecological design patterns (EDPs). This strategy is based around the development of best practice solutions to commonly occurring problems in environmental modeling. The EDP approach is a procedural approach that can be applied to most programming languages, since the outputs are verbal descriptions rather than actual source code. In principle individual EDPs can be assembled into a sequence of “abstract” solutions for which solutions already exists, thereby providing a template for the structure of the simulation model.

However, despite these advances in programming techniques, the problem of documentation remains unsolved. The ODDP does not fill this gap, since it is procedure for carrying out the documentation as a process separate from implementation and using different tools. Since the documentation in ODDP is not part of the implementation process of the simulation model, changes on either side, whether to the model or the documentation, must be manually copied to the other side.

To address this issue, this chapter suggests using LP, enhanced by the development of DynLP for application to dynamic simulation models, which provides a meta-language for specification of the source model. From this specification, which is stored in a unified source, both the simulation model and the corresponding (dynamic) documentation can be generated. Changes to the unified source automatically generate corresponding changes in the documentation and the simulation model.

Application of this approach would require development of a ecological domain specific language (eDSL), a task which is by no means trivial. A lot of information would need to be provided, but this is only a reflection the inherent complexity of ecological simulation models and is therefore not a conceptual flaw of the proposed approach. The development of a display engine for dynamic documentation will also require a significant development effort, and ultimately the success of such a system will depend on its acceptance by the scientific community.

10. Summary and outlook

10.1. Summary

This thesis investigates the bioecology of *D. gigas* with a focus on its energy-driven traits. It uses the conceptual functional triad migration-maturation-growth (fTMMG) (Keyl et al., 2008) as a hypothesis to explain properties of the species as manifestations of individual physiological and energetic traits. It applies a refinement of the fTMMG, the quantifiable energy driven life history model (EDLHM), as an executable simulation model. Model outputs are calculated using mathematical programs as well as an implemented simulation model and analyzed based on the assumption of an energy driven life history to evaluate the role of energy as a driver of individual traits, such as growth or cannibalism, and characteristics at the population level, such as the existence of size-at-maturity (SAM)-groups. The thesis explores ways to improve ecological modeling and simulation, specifically with reference to the technical aspect of transferring conceptual models to documented simulation models. The potential distribution area of *D. gigas* is reviewed using satellite data with sea surface temperature and flow data, and a corresponding flow model was developed, which enables the backtracking of observed paralarvae samples to potential spawning grounds.

Specifically, the thesis investigated the following hypotheses (Section 1.1, page 4f):

1. The species *D. gigas* adheres to an energy driven life history that controls *D. gigas*'s phenotypic expression.
2. Cannibalism is an important survival strategy for *D. gigas* during periods of food scarcity.
3. The phenotypic extremes postulated in Keyl et al. (2008) serve to optimize survival and fitness under the (locally) prevailing energetic-environmental conditions.
4. Methodological approaches currently used for ecological simulations require improvements in model specification, model formulation, model implementation and model documentation.

As presented in “Objects of Study” (Chapter 2), *D. gigas* is a species of extremes, including at the individual level. For example it has an extremely high metabolism

10. Summary and outlook

rate, among the highest known, combined with the capacity for metabolism suppression. *D. gigas* also shows extremely high phenotypic plasticity, and can attain sexual maturity, at a length of approx. 0.2m to approx. 1.2m, in a relatively short time.

In addition to a relatively short generation cycle (i.e., the time taken to reach sexual maturity) of one to two years, *D. gigas* displays rapid phenotypic adaptation to changing environmental conditions. In this context, semelparity ensures that sexually mature individuals in a population are all from the same generation and are thus representative of adaptation by the species to current environmental conditions.

These extreme traits at the individual level result in population-level fluctuations in abundance and SAM, so that *D. gigas* can serve as an indicator for the environmental conditions of the (local) marine ecosystem. The tight coupling of energy uptake to growth and maturation means that these traits respond to marine environmental conditions, so that the species *D. gigas* could be suitable as a model organism to study changes in the marine ecosystem.

D. gigas is an ecologically and commercially important species about which relatively little is known (Bruno et al., 2021; Gong et al., 2020; Sanchez et al., 2020; Yu, Chen, & Liu, 2021), which extended its distribution range after the strong 1997/1998 El Niño (EN)-event. It has high energy requirements and is an opportunistic predator that feeds on a variety of species that is also prey for other higher-level predators. It therefore occupies a key position in the marine food web (Bruno et al., 2021; Fang et al., 2018; Frawley et al., 2019; Gong et al., 2020; Keyl et al., 2008; Yu, Chen, & Liu, 2021). At the individual level, the high energy demand contrasts with a limited energy buffering capacity, so that the survival of an individual may be threatened under even short-lasting and/or moderately unfavorable environmental conditions that limit its food intake (Keyl et al., 2008). Widespread individual mortality could threaten the survival of populations, since semelparity means that survival of a population depends on the reproductive success of the current generation (Pecl & Jackson, 2008).

As a result of these characteristics, it might appear that the species has little chance of survival in the long run. Nevertheless it has survived, despite being intensively fished, and has expanded its range despite environmental changes that cause significant problems for other species, e.g., warmer waters and hypoxia. To explain this apparent contradiction, Keyl et al. (2008) postulate the “functional triad migration-maturation-growth” as an energy management strategy to ensure survival.

The fTMMG uses a conceptual model of physiology and energy use of *D. gigas* to explain individual traits such as reproductive tissue growth and size-at-maturity (SAM), as well as population level properties such as group structure and the existence of

SAM-groups. In the fTMMG, individual SAM is the result of an energetic optimization behavior and is located on a range between two phenotypic extremes. A small SAM enables individuals to cope with low-energy environments to ensure population survival under unfavorable environmental conditions, whereas the large SAM maximizes fitness (Keyl et al., 2008). The fTMMG focuses attention on the species' energy driven life history, but due to its qualitative formulation it can only postulate general effects, e.g., the hypothesized evolution of SAM-groups as a population-level energy optimization strategy in low-energy environments. In the absence of a quantitative formulation, it is difficult to test these and other hypotheses.

In this thesis, the energy driven life history model, which key elements are developed in “Elements of an energy driven life history for *D. gigas*” (Chapter 3), is proposed as a more detailed formulation and refinement of the original fTMMG. The quantitative formulation of the EDLHM means that it is susceptible to computational analysis and allows the exploration of the effects of energy-related drivers of change at individual and population level. Modeling of individual-level effects, and subsequent implementation and execution of an individual-based model incorporating these effects may help explain the population-level effects described in the literature by proponents of the original conceptual fTMMG.

The EDLHM includes an energy balance model (Chapter 4) of feedback among energy-related drivers of change at an individual level, i.e., the parameters growth rate, size, and food availability. Emerging population-level effects described in the literature, such as sexual maturation, sex ratio, and cannibalism, are explored in relation to individuals' energy-driven behavior. The effective of cannibalism is as a strategy for population-level survival, in response to the species' high energy requirements and limited energy buffering capacity (Fox, 1975; Ibáñez & Keyl, 2010; Keyl et al., 2008).

Chapter 4 (“Modeling the individual and its environment”) further develops EDLHM to quantify the cause–effect relations determining individual and population level effects. To this end, the following features were developed: a growth function, a geometric approximation of the *D. gigas* body form, formulas for size and body mass, an energy model, and procedures for the determination of basal metabolic rate (basal_{mr}), and the conversion between mass and energy. The result is an energy-based model of a *D. gigas* individual life history, expressed solely in terms of energy demand, energy intake, and growth.

The approach maps the individual energy demand to the current body mass, that itself is determined using the single input parameter size, and which links the growth function to the energy model. The size is obtained either from a growth function

10. Summary and outlook

that takes age as the input parameter or from a growth function that obtains the maximum available growth rate from the current size and actual growth by mapping the physiological processes that determine growth capacity and available energy. In this model, the determination of an individual's energy level is sufficient to model energy-constrained growth and the effects of a lack of energy buffer.

The spawning area is where the life history of *D. gigas* starts. The importance of the spawning area stems from its influence of temperature on feeding and probably the early developmental stages, as well as on horizontal migration behavior and therefore on the energy requirements for locomotion metabolic rate ($\text{locomotion}_{\text{mr}}$). However, only a few locations have been identified as potential spawning areas, which are probably insufficient to account for the presence of the species across the whole distribution range.

Based on an analysis of long-term sampling data, Staaf et al. (2013) find a relation between sea surface temperature (SST) and paralarvae presence probability (pp_{prob}) for the *Sthenoteuthis oualaniensis* and *Dosidicus gigas* complex (SD-complex) as well as identifying some possible spawning areas. In this thesis, this postulated relation was re-evaluated using satellite-based SST and currents data in “Identifying spawning areas of *D. gigas*” (Chapter 5). The analysis concludes that nearly the entire distribution range is suitable for spawning if the postulated relation of sea surface temperature (SST) and paralarvae presence probability (pp_{prob}) is valid. The pp_{prob} within the range may vary, but the calculations do not exclude extensive areas within the known distribution range as spawning grounds, since paralarvae are recorded from zones where pp_{prob} is low, such as off the coast of South America (Table 5.1). A large spawning area is consistent with minimization of energy requirements for migration to cover the entire known distribution range, including the high-latitude outer bounds. Widespread spawning would also mean that the observed range expansion is most likely consistent with the energetic capabilities of individuals.

However, this result contrasts with the identification of distinct spawning areas based on observational data Staaf et al. (2013). Comparison of the data in Staaf et al. (2013) with the results of the analysis in this thesis, found that observed sample counts were low in areas of expected higher paralarva concentration. As a possible explanation for this discrepancy, satellite-based currents data were used to create a model of material transport by ocean currents within the distribution range of *D. gigas*. It was hypothesized that the action of these currents would transport eggs and paralarvae from donating areas to aggregating areas. This spatial structuring process is reflected in the observed distribution of paralarvae, leading to their concentration in specific areas

which could be misinterpreted as spawning grounds. Both donating and aggregating areas are spatio-temporally variable, so the locations of none of these areas can be predicted.

Additionally, the idea that *D. gigas* spawns throughout its range supports the Hypothesis 3 that observed phenotypical variation is a response to environmental conditions, i.e., larger individuals occur in higher latitudes where more food resources are available, due to cooler water.

To test the hypothesis of energy-driven phenotypical traits at the individual level, a logistic dimorphic terminal size growth function (DTSGF) expressing sexual dimorphism was used to model growth in “Individual level traits and their computation” (Chapter 6). The dimorphic terminal size growth function (DTSGF) requires a maximum initial daily growth rate of approximately 4% in order for computed ML_{terminal} to not exceed observed ML_{terminal} ; however, this initial growth rate is at the lower bound of values reported in literature (Gilly, Elliger, et al., 2006; C. Nigmatullin et al., 2001; Rosa et al., 2013; Yatsu et al., 1999). Examination of DTSGF reveals a high dependence on energy intake to meet the energy requirements of basal+locomotion metabolic rate (basal+locomotion_{mr}) to support survival, and also restricted growth in the event of energy intake deprivation. In addition, during periods of food shortage, body mass is used to buffer the energy requirements of basal+locomotion_{mr} and starvation rapidly set in. Females appear to be at a disadvantage due to their lower buffering ability relative to their body mass, and therefore are more susceptible to starvation. Interestingly, according to the model, there is an optimum mantle length (ML) of approx. 0.2m that minimizes the negative effects of food deprivation. This coincides with the size of individuals in the small SAM group that is found in low-energy environments and may indicate that a small SAM is an adaptation to low-energy environments. The calculations in Chapter 6 show a pronounced effect of energy on the life history of *D. gigas* and provide support for Hypothesis 1 (page 241).

Exponential growth may amplify differences in hatching size to large differences in adults, so that hatching size is likely a key factor determining terminal size under optimal growing conditions (Section 6.4). The hatchling size is related to egg size, which in turn depends on the size of the female (Birk et al., 2016; Pecl & Jackson, 2008). A self-contained process may become established, whereby small females produce small hatchlings that can only grow into small adults. Cannibalism is an individual trait that gives rise to emergent effects at the school, cohort, and population levels. The preconditions for cannibalism and its effects are discussed in “Cannibalism as a possible survival strategy” (Chapter 7). The “Properties at the population level

10. Summary and outlook

and their computation” (Chapter 8) explores the hypothesized mechanisms by which population effects emerge from individual behavior and characteristics.

This analysis identifies several features of *D. gigas*’s cannibalism that do not correspond with traditional concepts of cannibalism (Smith & Reay, 1991; van den Bosch et al., 1988), see Section 7.2, including the idea of a window of cannibalism (WOC) (Claessen et al., 2000; Ibáñez & Keyl, 2010). In principle cannibalism is advantageous for the survival of monocyclic species such as *D. gigas* that depend on the reproductive success of the present generation. The physiological constraints of *D. gigas*, i.e., a high metabolic rate and lack of energy storage capacity, create adverse conditions for reproductive success. If food uptake from cannibalism (Smith & Reay, 1991) supports cannibal’s gene transfer to the next generation, then cannibalism both increases the fitness of the individual cannibal and (primarily) supports the survival of the species. However, looking more closely at cannibalism in *D. gigas*, a contradictory picture emerges.

Cannibalism is thought to be size-dependent, i.e., cannibalism rates increase with increasing ML of the cannibal and the smaller relative size $size_{ratio}$ of the victim (Keyl, 2009). Based on the DTSGF, females have an advantage over males in this respect, due to their larger ML relative to male conspecifics of the same age. In this sense, given the importance of the females for reproduction, the DTSGF appears to be compatible with cannibalism.

However, simulation of the population-level effects of cannibalism under conditions of absolute food scarcity shows that a $size_{ratio}$ -based cannibalism strategy (active cannibalism) is less successful than other cannibalism strategies in ensuring the survival of a school, since it leads to a premature reduction in school size (i.e., before starvation of individuals sets in). In the simulation, active cannibalism doubles the length of survival of the school, compared to a no-cannibalism strategy. But even this doubling is based on optimistic assumptions in the simulation, e.g., the high conversion factor for food intake. If these assumptions are not met, the strategy would be even less successful. The disadvantages of the active cannibalism strategy, including loss of biomass due to partial consumption of conspecifics as well the possible premature reduction in school size, lead to consideration of an alternative “passive cannibalism strategy”. Under this strategy, the performance of cannibalism is delayed until the appearance of energetically depleted conspecifics of either sex. Consumption of energetically depleted individuals transfers the energy that otherwise be lost when these individuals die in the form of biomass to the remaining population. However, unlike the active cannibalism strategy, passive cannibalism does not exert additional pressure on the population and sacrifices

only the bare minimum of conspecifics necessary for school survival. Compared to the active cannibalism strategy (and the other cannibalism strategies explored), the passive cannibalism strategy is more successful in prolonging the survival of a school because it makes more efficient use of the school as an energy buffer.

Under the passive cannibalism strategy, cannibalism occurs whether or not individuals are threatened by food scarcity; it is part of an opportunistic feeding strategy in which feeding on lethargic conspecifics is part of the regular diet. Taking this into account, a new two-dimensional window of predation (WOP^2) is proposed, combining the window of predation (WOP) and WOC, to provide a theoretical explanation for an opportunistic feeding strategy that includes cannibalism, whereby the availability of prey items is the result of a reasonable trade-off between the energy gained through consumption of prey items and the risk of retaliation. Such a feeding strategy encompasses every type of cannibalism strategy, so a separate cannibalism strategy is redundant. From this perspective, cannibalism is not an evolutionary stable strategy, as hypothesized in Ibáñez and Keyl (2010), but simply part of a WOP^2 -based feeding strategy, in which no distinction is made between conspecifics and other species. Hypothesis 2 (page 241) was only partially confirmed because cannibalism in *D. gigas* is better explained as being the result of an opportunistic feeding strategy.

Chapter 8 evaluates population-level properties, and particularly the postulated effect of non-synchronous reproductive tissue growth among members of a cohort on peak energy demand. The calculations show that non-synchronous reproductive tissue growth does not cause a significant reduction in the energy demand of a cohort, since $basal+locomotion_{mr}$ becomes increasingly dominant over growth metabolic rate ($growth_{mr}$) as ML increases. This result applies across all $ML_{terminal}$, but is more pronounced for larger $ML_{terminal}$.

Multiple spawning peaks create different cohorts and distribute energy demand peaks of these cohorts throughout the year, so that maximum population energy demand at any given moment in time is lower than it would be for a single-cohort population. The dominance of $basal+locomotion_{mr}$ means that absolute energy demand (i.e., total metabolic rate ($total_{mr}$)) is lower in small SAM individuals and populations, even allowing for the increased energy requirements associated with a shorter generational turnover. Small SAM groups are therefore a likely adaptation that optimizes energy use in low-energy environments.

A further finding was that changing the sex ratio in favor of females saves energy by reducing the number of males in the population and thereby decreasing overall population size. Based on the DTSGF, females consume more energy than males.

10. Summary and outlook

However, since females are essential for reproduction, only the removal of some of the males can reduce the absolute energy demand of the population without negative effects on its reproductive capacity. The sex ratio may be actively altered by cannibalism of males by females, but this would require females adopting an active cannibalism strategy, which is not supported by the findings of the cannibalism simulations.

In summary, as a cephalopod, *D. gigas* shows typical traits of the class, such as semelparity, high fertility, short lifespan and rapid growth. Some traits common to cephalopods are more extreme in *D. gigas*, such as its extremely high fertility, and extremely high phenotypic plasticity in combination with the very rapid growth. The resulting extremely high energy requirements must be met through food-uptake, since *D. gigas* has only a limited energy buffering capacity, if any, to enable it to survive periods of food scarcity. However, *D. gigas* has a high temperature and hypoxia tolerance and shows a voracious opportunistic feeding behavior, which allows it to expand its distribution range as opportunities arise. The expression of all these traits is altered by environmental factors, including negative anthropogenic changes such as overfishing and ocean acidification, rising water temperatures and extension of oxygen minimum zones due to global climate change. Its rapid responses to environment changes qualifies *D. gigas* as a model organism for study of their effects on marine ecosystems.

D. gigas' extremely high energy requirements in combination with a restricted energy buffering capacity impose to an energy driven life history, whereby the satisfaction of the energy requirements is paramount throughout the lifespan. This amplifies the effects of *D. gigas*' traits at the individual level, giving rise observable population-level characteristics, for example, variations in size-at-maturity (SAM) and extreme temporal variations in abundance. For example the existence of small SAM groups favors population survival in low energy environments, since small SAM individuals require less food uptake to reach maturity. Thus a population of small SAM individuals requires in total less energy until reproduction than a population (of the same size) of larger SAM individuals. Conversely, larger individuals are better adapted to high-energy environments since they are more fertile.

The energy driven life history of *D. gigas* can be modeled using an energy driven life history model (EDLHM) that is reduced to few parameters, relations and a standard logistic growth function. This model reproduces key traits of the species and explains how an energy driven life history enhances the survival of *D. gigas* at the individual and population level. However, there are many gaps in the current understanding of *D. gigas*, including of factors affecting energy requirements such as metabolic suppression. When

more is known some parameters of the EDLHM may need to be adjusted to take account of these aspects, without however altering the basic model design. New knowledge is not expected to affect the evaluation of cannibalism strategies in this thesis; nor the key finding from applying the EDLHM-based Hypothesis 3 (page 241) that the two phenotypic extremes (small and large SAM-group) are an adaptation to the environment to optimize fitness under different levels of food availability.

In the process of formulating the models and converting them into simulation models, some significant problems arose that have not yet been solved. The “The simulation of ecological models: Enhanced approaches and techniques” (Chapter 9) addresses these problems; it proposes state-based modeling (SBM) as an approach to structuring and implementation of individuals (agents) in simulation models. State-based modeling superimposes an additional structuring layer that forces the modeler to structure the model into distinct states. Each state is specific to a particular component of the model; it is functionally independent of the other states and does not interfere with them. State-based modeling should simplify model development, compared to traditional approaches, and reduce the occurrence of unanticipated side effects (due to interactions among components) during model execution.

If such a state model is supported by a programming language, then state-based modeling would enable unification of the conceptual model and the implemented simulation model in a single source model, avoiding the errors associated with the translation of a source model expressed (for example) in text and diagrams into a simulation model formulated in a programming language.

Therefore, as a next step, an ecological domain specific language (eDSL) should be developed as a common language to enable concise formulation of the model using literate programming (LP) (Knuth, 1984). LP enables parallel generation of the model documentation and the simulation model from a unified source, so that the source model and the simulation model should be functionally identical. The original LP approach is considered insufficient to cover the complexity of dynamic models, so an extended LP approach, the dynamic literate programming (DynLP), is proposed. A DynLP system would allow (smaller) simulation models and test scenarios to be embedded in the model description, thereby enabling documentation and testing of dynamic components of the simulation model as well static ones. This approach is expected to overcome limitations of existing static documentation procedures, such as the text-based descriptions required by Overview, Design concepts and Details Protocol (ODDP) (Grimm et al., 2006; Grimm et al., 2010; Grimm et al., 2020). However, a

10. Summary and outlook

DynLP would represent a novel approach to model development and documentation and require acceptance by the scientific community.

The ecological design pattern (EDP) is another (additional) approach to simplify the creation of models and simulation models, based on the identification of best-practice solutions to common modeling problems. Such an approach can be implemented without much effort, since these solutions are rather abstract proposals, described in human language, that are essentially independent of technical environments.

In summary, since the introduction of object oriented programming, few advances in the development of ecological simulation models have been adopted or even widely discussed. Thus, the Hypothesis 4 (see page 241) is also confirmed.

10.2. Reflection

The present work combines qualitative and quantitative methods to develop a physiological energetic model of *D. gigas* to explore a postulated energy driven life history of *D. gigas* and the associated working hypotheses (Section 1.1, page 4f).

The relatively sparseness of information available on *D. gigas* requires working with assumptions, creating a degree of uncertainty at least for the quantitative statements. Qualitative statements, i.e., confirmation of hypotheses, are less affected, so these conclusions are presumably valid.

In Chapter 3 and Chapter 4, a model was developed with the available information and the calculation of the model parameters was specified. The impact of undetected errors in the assumptions, in the model, or in the literature will run through all conclusions from this moment on, possibly amplifying small errors into large deviations over successive iterated calculations. Therefore these chapters include discussions of the possible implication of quantitative errors on the qualitative conclusions.

Results of the calculations describe a species with a high energy consumption, lack of energy buffering capacity, and a high cannibalism rate. The results are consistent within the framework EDLHM, but due to the uncertainty factor present, the extent the quantitatively results can be directly transferred to *D. gigas* is unclear. The calculated results are comparable with each other but may have to be subsequently recalculated as new data on *D. gigas* becomes available; methodically, however, nothing changes.

The confirmation of Hypothesis 1 might be considered questionable, to the extent that an energy driven life history may not account for all the observed behavior of *D. gigas* in the field. This limitation is a consequence the modeling approach adopted, which focuses on the energetic aspects and excludes consideration of other factors. In

any case, all the calculated effects require verification in the field, in accordance with the principle that theoretical ecology and simulation modeling can do no more than suggest possible effects.

The confirmation of Hypothesis 3 that observed phenotypic extremes are an adaptation to low-energy or high-energy environments results from the calculations carried out to test Hypothesis 1; the validity of this result is therefore subject to the same limitations as those related to Hypothesis 1.

The results on cannibalism (Hypothesis 2) should be sound with respect to the evaluation of alternative cannibalism strategies (or the lack of an explicit strategy), because the effects observed during the simulation (e.g., extinction of the school under conditions of absolute food scarcity) will occur in any case, with only the timing depending on the underlying assumptions. Reducing cannibalism to a component of an opportunistic feeding strategy could, if the result has general validity, relegate the importance of cannibalism as a separate survival strategy; restricting it to culture-creating species, where cannibalism is a conscious choice.

All calculations of the characteristics of *D. gigas*, as well as calculations of the model results, were strongly mechanistic, i.e., based a reduction to the essential (energetic) aspects of *D. gigas*'s life history. The model is reduced to its essence, which involves the risk of oversimplification. However, since the focus of the thesis was on the energetic components, the simplification is reasonable.

The computations of spawning grounds are methodologically and substantively a separate, distinct part of this thesis. Localized spawning grounds are of interest in the sense that the energy expenditure required for long-distance migrations to higher latitudes would have to be explained. The methodological approach is different from the other chapters because this chapter analyzes different data (i.e., satellite-based SST data) and models different phenomena (i.e., ocean currents). The novel approach adopted yields new insights regarding spawning grounds that are very different from those obtained from standard statistical analysis.

The thesis refers only briefly and rather speculatively to epigenetics, which, however, is an important topic that may be of relevance to the phenomena under investigation. Epigenetics could explain variations in SAM as the result of altered gene activity patterns developed by one or both parents in response the environmental conditions and subsequently passed down to the next generation (Lamarckism inheritance). Epigenetics would thus be responsible for short-term adaptation to changing environmental conditions, and regular genetics for long-term adaptation.

10. Summary and outlook

The application of epigenetics could contradict Hypothesis 1 and Hypothesis 3 by providing a genetic explanation for variations in SAM. The epigenetic hypothesis could be substantiated, for example, by using satellite-based temperature analyses and trapping statistics to look for correlations between the SAM and the energetic environmental conditions of the previous generation and/or the gene activity patterns at early developmental stages.

The hypothesis of restricted spawning grounds postulated by Staaf et al. (2013) is not supported by results obtained using the more comprehensive methodological approach employed in this thesis. The analysis of material transport provides a convincing explanation of close proximity of areas of high and low paralarva concentrations, as the result of a spatial organizing process driven by ocean currents, that transports paralarvae from extensive spawning grounds to relatively small “paralarvae gathering areas”.

As part of the modeling process, an abstraction layer, see Appendix D.2 for a brief description, was developed between MASON (Luke, 2019) and the actual simulation model. This abstraction layer is intended to provide a unified framework with the functionality of MASON.

The abstraction layer can —unlike MASON— easily switch between 2D and 3D models and provides the elementary basic types `Agent`, `Singleton` and `Resource` (see Subsection 9.4.2) required for model development. At the same time, it allows parallel execution of multiple simulation models, while capturing relevant data and writing it to a log file before and after each time step. This abstraction layer is intended to improve compliance with quality standards for scientific work, which include reproducibility and data collection.

The development of the abstraction layer was initiated in response to the theoretical considerations outlined in the previous paragraph at an early stage of work on the thesis. In retrospect, it is apparent that the features of the abstraction layer were only partially made use of in this thesis. Thus much of the considerable effort spent on developing the simulation environment could have been avoided, and the overall time required to develop the simulation models for cannibalism reduced. However, the effort was not without value, because it highlighted the difficulty and amount of effort that would be involved in the development of such an abstraction layer, which is certainly not yet complete in its present form. Completion of the abstraction layer would be the topic for a separate, self-contained thesis.

At the same time, this experience highlighted the need for implementation or specification of a freely accessible uniform development environment, which includes the

features of modern programming languages, e.g., functional programming. NETLOGO, widely used for ecological simulation, does not meet these requirements, in the opinion of the author of this thesis. SIMULA67, on the other hand, has shown that a specially developed programming language can have a significant impact.

The development of a formal specification language for ecological simulation models, as well as standards for dynamic documentation and procedures for the automatic implementation to a simulation model from a conceptual source model, are logical next steps. Unfortunately, it was only possible to outline criteria for achieving these objectives and suggest some approaches through which they could be achieved. The complexity and effort that would be involved in the development, or even specification of such tools quickly became apparent.

The confirmation of Hypothesis 4 is not based on a fully systematic examination of the prevailing “simulation landscape” in theoretical ecology, so a bias is certainly possible. However, in the opinion of the author of this thesis, it would be surprising if important developments in the field escaped his observation.

10.3. Outlook

This thesis has not identified any new characteristics in the life history of *D. gigas*, merely confirmed or rejected existing assumptions. The species *D. gigas* remains enthralling because it is a cephalopod that has existed for about 500 million years, successfully adapting to changing marine environments and ecosystems. Possibly this survival is based on the radically simple “design” of the species, which goes through generational change at high speed without superfluous bells and whistles, reducing its existence to the essence of life: feeding, growing, maturing, reproduction and death.

A mechanistic approach seems appropriate for analysis of this radically simple set of principles. In this sense, *D. gigas* commends itself as a simple model organism for further research on environmental change and organismal adaptation, allowing simplification of assumptions made with respect to other, more complex species. For example, as demonstrated in this thesis, in the case of *D. gigas*, there is no need to postulate a distinct cannibalism strategy, since cannibalism can be explained as a component of an opportunistic feeding strategy.

The difficulties involved in laboratory studies and the generally rather limited observational possibilities in the oceans, complicate the work of marine biologists, who frequently resort to alternative approaches such as simulation modeling of organisms and ecosystems. However, studies in this field, as in others, adopt a wide variety

10. Summary and outlook

of approaches to the development and application of simulation models. Models are frequently developed by one research group working in isolation and remain within the group as proprietary developments. It is evident that procedures for the development of conceptual models and their implementation as simulation models require improvement. The latter need has been addressed using different programming languages, simulation frameworks and simulation environments. Since, in the recent past, there has been no breakthrough in systematic aspects of simulation model implementation, these efforts are generally focused on improving implementation techniques, discussing the relative advantages of programming languages or available frameworks, and other programming-related issues.

Potential improvements include the formulation of ecological design patterns (EDPs) that provide reusable solutions to recurring problems encountered in ecological modeling. This is a generally applicable approach that is largely independent of programming languages, but, in practice, application of the approach will vary depending on the simulation framework that is used. This should not be a major problem provided that the frameworks share a set of common functions, which would make it possible for ecological simulation models to be shared across frameworks. The formulation of EDPs is a collective scientific endeavor to which many researchers can contribute.

Another, complementary possibility is use of state-based modeling (SBM) both for formulating a conceptual model and implementing the corresponding simulation model. State-based modeling structures an individual's behavior into distinct states, each of which contains state-specific code that is executed when an individual enters that state. This additional structure encapsulates different sections of code, reduces dependencies in the code structure, and provides a concise definition of an individual's behavior. This is a practical approach to managing the recurring structural elements in model formulation and model implementation.

State-based modeling marks an important step towards the structural alignment of a conceptual model and the corresponding simulation model, but different tools and languages are still used in their development. In addition, there is no requirement for model documentation, which is essential for understanding the model concerned. A separate approach to documentation, as in ODDP, attached to the modeling process, seems clumsy; moreover it does not solve the problem of possible discrepancies between the ecologist's conceptual model and the implemented simulation model, which means that ecologists could end up analyzing results from a model that is different from what they intended.

To address existing shortcomings, using literate programming (LP) to develop a unified meta-language would represent a feasible solution for model formulation, model documentation and simulation model generation. Previous approaches and their tools use different (programming) languages and create redundancy, as well as being prone to mismatches between the conceptual model and the simulation model. In a LP approach, changes to the unified source automatically generate corresponding changes to source model formulation, documentation, and specification of the simulation model, relieving the ecologist of additional work.

The advantages of this approach are obvious: the simulation model is an exact definition of the source model. The simulation model expresses a source model, that is largely specified in a human language which contains semantic ambiguity, in a semantically precise language from which ambiguity is excluded. However, the program code is hard for humans to read and understand and only unfolds its expressiveness when executed on a computer. The use of a unified language that is capable of generating both the simulation model and the documentation implicitly requires a specification language with the same accuracy as a programming language.

Taking this idea further, a model specification system like literate programming (LP) or, in the future, dynamic literate programming (DynLP) could provide a new gold standard for ecological modeling in the form of a single language for model description, model formulation, model documentation and model implementation. This simplified approach to modeling would seem to offer considerable benefits for the scientific community, but the need for such a system has not been widely articulated. Either there is no need for such a system, i.e., scientists are comfortable with existing methodologies for model development, or there is a need for more thinking outside the box.

Many improvements remain to be made in the development and application of ecological models and simulations. First, more use should be made of the potential of models to challenge established assumptions, such as that cannibalism is a survival strategy, by revealing simpler mechanisms to explain observed phenomena. With respect to the development of models and simulation models, there have been no game-changing advances with the potential to revolutionize the genre and simplify the work of ecologists since the development of SIMULA67. In their absence, the only options are for ecologist to engage in the intensive study of computer science, and/or for computer scientists to become more familiar with ecology. This “meeting of minds” is possibly the only way that current problems affecting the development of ecological simulations are going to find a solution.

Bibliography

- Accolla, C., Vaugeois, M., Grimm, V., Moore, A. P., Rueda-Cediel, P., Schmolke, A., & Forbes, V. E. (2020). A review of key features and their implementation in unstructured, structured, and agent-based population models for ecological risk assessment. *Integrated Environmental Assessment and Management*. <https://doi.org/10.1002/ieam.4362>
- Ådlandsvik, B., Gundersen, A. C., Nedreaas, K. H., Stene, A., & Albert, O. T. (2004). Modelling the advection and diffusion of eggs and larvae of Greenland halibut (*Reinhardtius hippoglossoides*) in the north-east Arctic. *Fisheries Oceanography*, *13*(6), 403–415. <https://doi.org/10.1111/j.1365-2419.2004.00303.x>
- Alegre, A., Ménard, F., Tafur, R., Espinoza, P., Argüelles, J., Maehara, V., Flores, O., Simier, M., & Bertrand, A. (2014). Comprehensive Model of Jumbo Squid *Dosidicus gigas* Trophic Ecology in the Northern Humboldt Current System. *PLoS ONE*, *9*(1), e85919, 1–11. <https://doi.org/10.1371/journal.pone.0085919>
- An, L., Grimm, V., & Turner, B. L. (2020). Editorial: Meeting grand challenges in agent-based models. *JASSS*, *23*(1). <https://doi.org/10.18564/jasss.4012>
- Anderson, C. I., & Rodhouse, P. G. (2001). Life cycles, oceanography and variability: Ommastrephid squid in variable oceanographic environments. *Fisheries Research*, *54*(1), 133–143. [https://doi.org/10.1016/S0165-7836\(01\)00378-2](https://doi.org/10.1016/S0165-7836(01)00378-2)
- André, J., Grist, E. P., Semmens, J. M., Pecl, G. T., & Segawa, S. (2009). Effects of temperature on energetics and the growth pattern of benthic octopuses. *Marine Ecology Progress Series*, *374*, 167–179. <https://doi.org/10.3354/meps07736>
- Argonne National Laboratory. (2021). RePast. Retrieved July 15, 2021, from <https://repast.github.io/index.html>
- Argüelles, J., Rodhouse, P. G. K., Villegas, P., & Castillo, G. (2001). Age, growth and population structure of the jumbo flying squid *Dosidicus gigas* in Peruvian waters. *Fisheries Research*, *54*(1), 51–61. [https://doi.org/10.1016/S0165-7836\(01\)00380-0](https://doi.org/10.1016/S0165-7836(01)00380-0)
- Argüelles, J., & Tafur, R. (2010). New insights on the biology of the jumbo squid *Dosidicus gigas* in the Northern Humboldt Current System: Size at maturity, somatic and reproductive investment. *Fisheries Research*, *106*(2), 185–192. <https://doi.org/10.1016/j.fishres.2010.06.005>

Bibliography

- Argüelles, J., Tafur, R., Taipe, A., Villegas, P., Keyl, F., Dominguez, N., & Salazar, M. (2008). Size increment of jumbo flying squid *Dosidicus gigas* mature females in Peruvian waters, 1989–2004. *Progress in Oceanography*, *79*(2–4), 308–312. <https://doi.org/10.1016/j.pocean.2008.10.003>
- Arkhipkin, A., Argüelles, J., Shcherbich, Z., & Yamashiro, C. (2014). Ambient temperature influences adult size and life span in jumbo squid (*Dosidicus gigas*). *Canadian Journal of Fisheries and Aquatic Sciences*, *72*(3), 400–409. <https://doi.org/10.1139/cjfas-2014-0386>
- Ayllón, D., Railsback, S. F., Gallagher, C., Augusiak, J., Baveco, H., Berger, U., Charles, S., Martin, R., Focks, A., Galic, N., Liu, C., van Loon, E. E., Nabe-Nielsen, J., Piou, C., Polhill, J. G., Preuss, T. G., Radchuk, V., Schmolke, A., Stadnicka-Michalak, J., ... Grimm, V. (2020). Keeping modelling notebooks with TRACE: Good for you and good for environmental research and management support. *Environmental Modelling & Software*, 104932. <https://doi.org/10.1016/j.envsoft.2020.104932>
- Bazzino, G., Salinas-Zavala, C., & Markaida, U. (2007). Variability in the population structure of jumbo squid (*Dosidicus gigas*) in Santa Rosalía, central Gulf of California | Variabilidad en la estructura poblacional del calamar gigante (*Dosidicus gigas*) en Santa Rosalía, región central del Golfo de California. *Ciencias Marinas*, *33*(2), 173–186. <https://doi.org/dx.doi.org/10.7773/cm.v33i2.1055>
- Bazzino, G., Gilly, W. F., Markaida, U., Salinas-Zavala, C. A., & Ramos-Castillejos, J. (2010). Horizontal movements, vertical-habitat utilization and diet of the jumbo squid (*Dosidicus gigas*) in the Pacific Ocean off Baja California Sur, Mexico. *Progress in Oceanography*, *86*(1–2), 59–71. <https://doi.org/10.1016/j.pocean.2010.04.017>
- Birk, M. A., Paight, C., & Seibel, B. A. (2016). Observations of multiple pelagic egg masses from small-sized jumbo squid (*Dosidicus gigas*) in the Gulf of California. *Journal of Natural History*, *51*(43–44), 2569–2584. <https://doi.org/10.1080/00222933.2016.1209248>
- Boerlijst, M., & Hogeweg, P. (1991). Spiral wave structure in pre-biotic evolution: Hypercycles stable against parasites. *Physica D: Nonlinear Phenomena*, *48*(1), 17–28. [https://doi.org/10.1016/0167-2789\(91\)90049-f](https://doi.org/10.1016/0167-2789(91)90049-f)
- Bolker, B. M. (2008). *Ecological Models and Data in R*. Princeton University Press. <https://doi.org/10.1016/j.phrs.2010.10.003>
- Bordini, R. H., Braubach, L., Dastani, M., El Fallah Seghrouchni, A., Gomez-Sanz, J. J., Leite, J., O'Hare, G., Pokahr, A., & Ricci, A. (2006). A Survey of Programming

- Languages and Platforms for Multi-Agent Systems. *Informatica*, 30(1), 33–44. <http://www.informatica.si/index.php/informatica/article/view/71>
- Boyle, P. R., & v. Boletzky, S. (1996). Cephalopod Populations: definition and dynamics. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 351(1343), 985–1002. <https://doi.org/10.1098/rstb.1996.0089>
- Boyle, P. R., & Rodhouse, P. G. K. (2005). *Cephalopods: Ecology and Fisheries* (1st, Vol. 22). Blackwell Science Ltd. <https://doi.org/10.3354/meps022077>
- Breckling, B., Jopp, F., & Reuter, H. (2011). System Analysis and Context Assessment. In F. Jopp, H. Reuter, & B. Breckling (Eds.), *Modelling complex ecological dynamics* (1st, pp. 43–54). Springer. <https://doi.org/10.1007/978-3-642-05029-9>
- Breckling, B., Middelhoff, U., & Reuter, H. (2006). Individual-based models as tools for ecological theory and application: Understanding the emergence of organisational properties in ecological systems. *Ecological Modelling*, 194(1), 102–113. <https://doi.org/10.1016/j.ecolmodel.2005.10.005>
- Breckling, B., Müller, F., Reuter, H., Hölker, F., & Fränzle, O. (2005). Emergent properties in individual-based ecological models – introducing case studies in an ecosystem research context. *Ecological Modelling*, 186(4), 376–388. <https://doi.org/10.1016/j.ecolmodel.2005.02.008>
- Breckling, B., Pe'er, G., & Matsinos, Y. G. (2011). Cellular Automata in Ecological Modelling. In F. Jopp, H. Reuter, & B. Breckling (Eds.), *Modelling complex ecological dynamics* (pp. 105–117). Springer. <https://doi.org/10.1007/978-3-642-05029-9>
- Brett, J. R., & Groves, T. D. D. (1979). Physiological Energetics. In W. Hoar, D. J. Randall, & J. R. Brett (Eds.), *Fish physiology – bioenergetics and growth* (pp. 279–352). Academic Press.
- Bruno, C., Cornejo, C. F., Riera, R., & Ibáñez, C. M. (2021). What is on the menu? Feeding, consumption and cannibalism in exploited stocks of the jumbo squid *Dosidicus gigas* in south-central Chile. *Fisheries Research*, 233(September 2020). <https://doi.org/10.1016/j.fishres.2020.105722>
- California Institute of Technology. (2009). Ocean Surface Current Analysis (OSCAR) Third Degree Resolution User's Handbook.
- Camarillo-Coop, S., Salinas-Zavala, C. a., Lavaniegos, B. E., & Markaida, U. (2013). Food in early life stages of *Dosidicus gigas* (Cephalopoda: Ommastrephidae) from the Gulf of California, Mexico. *Journal of the Marine Biological Association of the United Kingdom*, 93(07), 1903–1910. <https://doi.org/10.1017/S0025315413000398>

Bibliography

- Camarillo-Coop, S., Salinas-Zavala, C. A., Manzano-Sarabia, M., & Aragón-Noriega, E. A. (2010). Presence of *Dosidicus gigas* paralarvae (Cephalopoda: Ommastrephidae) in the central Gulf of California, Mexico related to oceanographic conditions. *Journal of the Marine Biological Association of the United Kingdom*, *91*(04), 807–814. <https://doi.org/10.1017/S0025315410001517>
- Castle, C. J. E., & Crooks, A. T. (2006). Principles and Concepts of Agent-Based Modelling for Developing Geographical Simulations. <https://www.ucl.ac.uk/bartlett/casa/publications/2006/sep/casa-working-paper-110>
- Chen, X., Han, F., Zhu, K., Punt, A. E., & Lin, D. (2020). The breeding strategy of female jumbo squid *Dosidicus gigas*: energy acquisition and allocation. *Scientific reports*, *10*(1), 9639. <https://doi.org/10.1038/s41598-020-66703-5>
- Chen, X., Li, J., Liu, B., Chen, Y., Li, G., Fang, Z., & Tian, S. (2013). Age, growth and population structure of jumbo flying squid, *Dosidicus gigas*, off the Costa Rica Dome. *Journal of the Marine Biological Association of the United Kingdom*, *93*(02), 567–573. <https://doi.org/10.1017/S0025315412000422>
- Chen, X., Liu, B., & Chen, Y. (2008). A review of the development of Chinese distant-water squid jigging fisheries. *Fisheries Research*, *89*(3), 211–221. <https://doi.org/10.1016/j.fishres.2007.10.012>
- Chen, X., Lu, H., Liu, B., & Chen, Y. (2011). Age, growth and population structure of jumbo flying squid, *Dosidicus gigas*, based on statolith microstructure off the Exclusive Economic Zone of Chilean waters. *Journal of the Marine Biological Association of the United Kingdom*, *91*(01), 229–235. <https://doi.org/10.1017/S0025315410001438>
- Claessen, D., De Roos, A. M., & Persson, L. (2000). Dwarfs and Giants: Cannibalism and Competition in Size-Structured Populations. *American Naturalist*, *155*(2), 219–237. <https://doi.org/10.1086/303315>
- Colmerauer, A. (1996). Prolog. Retrieved July 15, 2021, from <http://alain.colmerauer.free.fr>
- Cook, M. (2004). Universality in Elementary Cellular Automata. *Complex Systems*, *15*(1), 1–40.
- Cook, M. (2009). A Concrete View of Rule 110 Computation. *Electronic Proceedings in Theoretical Computer Science*, *1*, 31–55. <https://doi.org/10.4204/EPTCS.1.4>
- Cornell, S. J., Suprunenko, Y. F., Finkelshtein, D., Somervuo, P., & Ovaskainen, O. (2019). A unified framework for analysis of individual-based models in ecology and beyond. *Nature Communications*, *10*(1), 1–14. <https://doi.org/10.1038/s41467-019-12172-y>

- Crooks, A., & States, U. (2017). Agent-Based Modeling. In T. J. Cova & M.-H. Tsou (Eds.), *Comprehensive geographic information systems* (First Edit, pp. 218–243). Elsevier. <https://doi.org/10.1016/B978-0-12-409548-9.09704-9>
- Crossland, M. R., & Shine, R. (2011). Cues for cannibalism: Cane toad tadpoles use chemical signals to locate and consume conspecific eggs. *Oikos*, *120*(3), 327–332. <https://doi.org/10.1111/j.1600-0706.2010.18911.x>
- Dahl, O.-J., Nugaard, B., & Kristian, M. (1970). *SIMULA Information – Common Base Language* (Revised ed). Nowegian Computing Center.
- Davis, R. W., Jaquet, N., Gendron, D., Markaida, U., Bazzino, G., & Gilly, W. (2007). Diving behavior of sperm whales in relation to behavior of a major prey species, the jumbo squid, in the Gulf of California, Mexico. *Marine Ecology Progress Series*, *333*, 291–302. <https://doi.org/10.3354/meps333291>
- Dawkins, R. (2006). *The Selfish Gene* (30th anniv). Oxford University Press.
- DeAngelis, D. L., & Diaz, S. G. (2019). Decision-Making in Agent-Based Modeling: A Current Review and Future Prospectus. *Frontiers in Ecology and Evolution*, *6*(JAN), 1–15. <https://doi.org/10.3389/fevo.2018.00237>
- DeAngelis, D. L., & Yurek, S. (2017). Spatially Explicit Modeling in Ecology: A Review. *Ecosystems*, *20*(2), 284–300. <https://doi.org/10.1007/s10021-016-0066-z>
- Deutsch, C., Ferrel, A., Seibel, B., Pörtner, H.-O., & Huey, R. B. (2015). Climate change tightens a metabolic constraint on marine habitats. *Science*, *348*(6239), 1132–1136. <https://doi.org/10.1126/science.aaa1605>
- Eclipse Foundation. (2021). Eclipse IDE. Retrieved July 15, 2021, from <https://www.eclipse.org>
- ECMA. (2021). ECMA-262. Retrieved July 15, 2021, from <https://www.ecma-international.org/publications-and-standards/standards/ecma-262/>
- Ehrhardt, N. (1991). Potential Impact of a Seasonal Migratory Jumbo Squid (*Dosidicus gigas*) Stock on a Gulf of California Sardine (*Sardinops Sagax Caerulea*) Population. *Bulletin of Marine Science*, *49*(1–2), 325–332. <http://www.ingentaconnect.com/content/umrsmas/bullmar/1991/00000049/F0020001/art00028>
- Fang, Z., Chen, X., Su, H., Staples, K., & Chen, Y. (2018). Exploration of statolith shape variation in jumbo flying squid, *Dosidicus gigas*, based on wavelet analysis and machine learning methods for stock classification. *Bulletin of Marine Science*, *94*(4), 1465–1482. <https://doi.org/10.5343/bms.2017.1176>
- FAO. (2021). *Dosidicus gigas* Global Capture Production. Retrieved May 14, 2021, from <http://www.fao.org/fishery/species/2721/en>

Bibliography

- Fedriani, J. M., Wiegand, T., Ayllón, D., Palomares, F., Suárez-Esteban, A., & Grimm, V. (2018). Assisting seed dispersers to restore oldfields: An individual-based model of the interactions among badgers, foxes and Iberian pear trees. *Journal of Applied Ecology*, *55*(2), 600–611. <https://doi.org/10.1111/1365-2664.13000>
- Field, J. C., Francis, R. C., & Aydin, K. (2006). Top-down modeling and bottom-up dynamics: Linking a fisheries-based ecosystem model with climate hypotheses in the Northern California Current. *Progress in Oceanography*, *68*(2), 238–270. <https://doi.org/10.1016/j.pocean.2006.02.010>
- Field, J. C., Baltz, K. E. N., & Walker, W. A. (2007). Range Expansion and Trophic Interactions of the Jumbo Squid, *Dosidicus Gigas*, in the California Current. *California Cooperative Oceanic Fisheries Investigations Report*, *48*, 131–146. http://calcofi.org/~%7Dcalcofi/publications/calcofireports/v48/Vol%7B%5C_%7D48%7B%5C_%7DField.pdf
- Field, J. C., Elliger, C., Baltz, K., Gillespie, G. E., Gilly, W. F., Ruiz-Cooley, R., Pearse, D., Stewart, J. S., Matsubu, W., & Walker, W. A. (2013). Foraging ecology and movement patterns of jumbo squid (*Dosidicus gigas*) in the California Current System. *Deep Sea Research Part II: Topical Studies in Oceanography*, *95*, 37–51. <https://doi.org/10.1016/j.dsr2.2012.09.006>
- Forrester, J. W. (1968). *Principles of Systems* (Vol. 2). Wright Allen Press.
- Fox, L. R. (1975). Cannibalism in Natural Populations. *Annual Review of Ecology and Systematics*, *6*(1), 87–106. <https://doi.org/10.1146/annurev.es.06.110175.000511>
- Frawley, T. H., Briscoe, D. K., Daniel, P. C., Britten, G. L., Crowder, L. B., Robinson, C. J., Gilly, W. F., & Arkhipkin, A. (2019). Impacts of a shift to a warm-water regime in the Gulf of California on jumbo squid (*Dosidicus gigas*). *ICES Journal of Marine Science*, *76*(7), 2413–2426. <https://doi.org/10.1093/icesjms/fsz133>
- Gallego, A. (2011). Biophysical Models: An Evolving Tool in Marine Ecological Research. In F. Jopp, B. Breckling, & H. Reuter (Eds.), *Modelling complex ecological dynamics* (pp. 279–289). Springer. https://doi.org/10.1007/978-3-642-05029-9_20
- GBIF Occurrence Download. (2021). <https://doi.org/DOI10.15468/dl.6ub4w8>
- GeospatialPython. (2017). Python Shapefile Library (pyshp). Retrieved October 31, 2017, from <https://github.com/GeospatialPython/pyshp>
- Gilly, W. F., & Markaida, U. (2006). Perspectives on *Dosidicus gigas* in a changing world. In R. J. Olson & J. W. Young (Eds.), *The role of squid in open ocean ecosystems* (pp. 81–90).

- Gilly, W. F., Markaida, U., Baxter, C. H., Block, B. A., Boustany, A., Zeidberg, L., Reisenbichler, K., Robison, B., Bazzino, G., & Salinas, C. (2006). Vertical and horizontal migrations by the jumbo squid *Dosidicus gigas* revealed by electronic tagging. *Marine Ecology Progress Series*, *324*, 1–17. <https://doi.org/10.3354/meps324001>
- Gilly, W. F., Zeidberg, L. D., Booth, J. A. T., Stewart, J. S., Marshall, G., Abernathy, K., & Bell, L. E. (2012). Locomotion and behavior of Humboldt squid, *Dosidicus gigas*, in relation to natural hypoxia in the Gulf of California, Mexico. *Journal of Experimental Biology*, *215*(18), 3175–3190. <https://doi.org/10.1242/jeb.072538>
- Gilly, W. F. (2006a). Horizontal and vertical migrations of *Dosidicus gigas* in the Gulf of California revealed by electronic tagging. http://www.soest.hawaii.edu/PFRP/nov06mtg/gilly%7B%5C_%7Dgigas.pdf
- Gilly, W. F. (2006b). *Horizontal and vertical migrations of Dosidicus gigas in the Gulf of California revealed by electronic tagging* (tech. rep. June 2005). Hopkins Marine Stations, Department of Biological Sciences, Stanford University. Pacific Grove.
- Gilly, W. F., Beman, J. M., Litvin, S. Y., & Robison, B. H. (2013). Oceanographic and Biological Effects of Shoaling of the Oxygen Minimum Zone. *Annual Review of Marine Science*, *5*(1), 393–420. <https://doi.org/10.1146/annurev-marine-120710-100849>
- Gilly, W. F., Elliger, C. A., Salinas, C. A., Camarilla-Coop, S., Bazzino, G., & Beman, M. (2006). Spawning by jumbo squid *Dosidicus gigas* in San Pedro Mártir Basin, Gulf of California, Mexico. *Marine Ecology Progress Series*, *313*, 125–133. <https://doi.org/10.3354/meps313125>
- Ginot, V., & Le Page, C. (1998). Mobidyc, a Generic Multi-Agents Simulator for Modeling Populations Dynamics. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *1416*, 806–814. https://doi.org/10.1007/3-540-64574-8_467
- Glaubrecht, M., & Salcedo-Vargas, M. A. (2008). The Humboldt squid *Dosidicus gigas* (Orbigny, 1835): History of the Berlin specimen, with a reappraisal of other (bathy-)pelagic “gigantic” cephalopods (Mollusca, Ommastrephidae, Architeuthidae). *Zoosystematics and Evolution*, *80*(1), 53–69. <https://doi.org/10.1002/mmzn.20040800105>
- Glazier, D. S., Hirst, A. G., & Atkinson, D. (2015). Shape shifting predicts ontogenetic changes in metabolic scaling in diverse aquatic invertebrates. *Proceedings of the*

Bibliography

- Royal Society B: Biological Sciences*, 282(1802), 1–9. <https://doi.org/10.1098/rspb.2014.2302>
- Glazier, D. S. (2006). The 3/4-Power Law Is Not Universal: Evolution of Isometric, Ontogenetic Metabolic Scaling in Pelagic Animals. *BioScience*, 56(4), 325. [https://doi.org/10.1641/0006-3568\(2006\)56\[325:TPLINU\]2.0.CO;2](https://doi.org/10.1641/0006-3568(2006)56[325:TPLINU]2.0.CO;2)
- Goicochea-Vigo, C., Morales-Bojórquez, E., Zepeda-Benitez, V. Y., Hidalgo-De-La-Toba, J. Á., Aguirre-Villaseñor, H., Mostacero-Koc, J., & Atoche-Suclupe, D. (2019). Age and growth estimates of the jumbo flying squid (*Dosidicus gigas*) off Peru. *Aquatic Living Resources*, 32. <https://doi.org/10.1051/alr/2019007>
- Gong, Y., Li, Y., Chen, X., & Yu, W. (2020). Trophic Niche and Diversity of a Pelagic Squid (*Dosidicus gigas*): A Comparative Study Using Stable Isotope, Fatty Acid, and Feeding Apparatuses Morphology. *Frontiers in Marine Science*, 7(July), 1–10. <https://doi.org/10.3389/fmars.2020.00642>
- Grimm, V., & Railsback, S. F. (2005). *Individual-based Modeling and Ecology*. Princeton University Press. <https://doi.org/10.1111/j.1467-2979.2008.00286.x>
- Grimm, V. (2018). Ecological models: Individual-Based Models (2nd ed.). *Encyclopedia of Ecology*, (March), 65–73. <https://doi.org/10.1016/B978-0-12-409548-9.11144-3>
- Grimm, V., Augusiak, J., Focks, A., Frank, B. M., Gabsi, F., Johnston, A. S., Liu, C., Martin, B. T., Meli, M., Radchuk, V., Thorbek, P., & Railsback, S. F. (2014). Towards better modelling and decision support: Documenting model development, testing, and analysis using TRACE. *Ecological Modelling*, 280, 129–139. <https://doi.org/10.1016/j.ecolmodel.2014.01.018>
- Grimm, V., Ayllón, D., & Railsback, S. F. (2017). Next-Generation Individual-Based Models Integrate Biodiversity and Ecosystems: Yes We Can, and Yes We Must. *Ecosystems*, 20(2), 229–236. <https://doi.org/10.1007/s10021-016-0071-2>
- Grimm, V., Berger, U., Bastiansen, F., Eliassen, S., Ginot, V., Giske, J., Goss-Custard, J., Grand, T., Heinz, S. K., Huse, G., Huth, A., Jepsen, J. U., Jørgensen, C., Mooij, W. M., Müller, B., Pe'er, G., Piou, C., Railsback, S. F., Robbins, A. M., ... DeAngelis, D. L. (2006). A standard protocol for describing individual-based and agent-based models. *Ecological Modelling*, 198(1–2), 115–126. <https://doi.org/10.1016/j.ecolmodel.2006.04.023>
- Grimm, V., Berger, U., DeAngelis, D. L., Polhill, J. G., Giske, J., & Railsback, S. F. (2010). The ODD protocol: A review and first update. *Ecological Modelling*, 221(23), 2760–2768. <https://doi.org/10.1016/j.ecolmodel.2010.08.019>

- Grimm, V., Railsback, S. F., Vincenot, C. E., Berger, U., Gallagher, C., DeAngelis, D. L., Edmonds, B., Ge, J., Giske, J., Groeneveld, J., Johnston, A. S., Milles, A., Nabe-Nielsen, J., Polhill, J. G., Radchuk, V., Rohwäder, M. S., Stillman, R. A., Thiele, J. C., & Ayllón, D. (2020). The ODD protocol for describing agent-based and other simulation models: A second update to improve clarity, replication, and structural realism. *JASSS*, *23*(2). <https://doi.org/10.18564/jasss.4259>
- Grimm, V., Revilla, E., Berger, U., Jeltsch, F., Mooij, W. M., Railsback, S. F., Thulke, H.-h., Weiner, J., Wiegand, T., & Deangelis, D. L. (2005). Pattern-Oriented Modeling of Agent-Based Complex Systems: Lessons from Ecology. *Science*, *310*(November), 987–991. <https://doi.org/10.1126/science.1116681>
- Gruber, J. (2004). Markdown. Retrieved November 18, 2017, from <https://daringfireball.net/projects/markdown/>
- Guppy, M., & Withers, P. (1999). Metabolic depression in animals: physiological perspectives and biochemical generalizations. *Biological reviews of the Cambridge Philosophical Society*, *74*(1), 1–40. <https://doi.org/10.1017/s0006323198005258>
- Han, G., & Kulka, D. W. (2009). Dispersion of Eggs, Larvae and Pelagic Juveniles of White Hake (*Urophycis tenuis*) in Relation to Ocean Currents of the Grand Bank: A Modelling Approach. *Journal of Northwest Atlantic Fishery Science*, *41*, 183–196. <https://doi.org/10.2960/J.v41.m627>
- Heinen, J. T., & Abdella, J. a. (2005). On the Advantages of Putative Cannibalism in American Toad Tadpoles (*Bufo a. americanus*): Is it Active or Passive and Why? *The American Midland Naturalist*, *153*(2), 338–347. [https://doi.org/10.1674/0003-0031\(2005\)153\[0338:OTAOPC\]2.0.CO;2](https://doi.org/10.1674/0003-0031(2005)153[0338:OTAOPC]2.0.CO;2)
- Hejlsberg, A. (2021). Typescript. Retrieved July 15, 2021, from <https://www.typescriptlang.org>
- Hernandez-Herrera, A., Morales-Bojorquez, E., Cisneros-Mata, M. A., Nevarez-Martinez, M. O., & Rivera-Parra, G. I. (1998). Management strategy for the giant squid (*Dosidicus gigas*) fishery in the Gulf of California, Mexico. *California Cooperative Oceanic Fisheries Investigations Reports*, *39*, 212–218. http://www.calcofi.org/publications/calcofireports/v39/Vol%7B%5C_%7D39%7B%5C_%7DHernandez-Herrera%7B%5C_%7Detal.pdf
- Hernández-Muñoz, A. T., Rodríguez-Jaramillo, C., Mejía-Rebollo, A., & Salinas-Zavala, C. A. (2015). Reproductive strategy in jumbo squid *Dosidicus gigas* (D’Orbigny, 1835): A new perspective. *Fisheries Research*, *173*, 145–150. <https://doi.org/10.1016/j.fishres.2015.09.005>

Bibliography

- Hölker, F., & Breckling, B. (2005). A spatiotemporal individual-based fish model to investigate emergent properties at the organismal and the population level. *Ecological Modelling*, *186*(September), 406–426. <https://doi.org/10.1016/j.ecolmodel.2005.02.010>
- Holmes, J., Cooke, K. E. N., & Cronkite, G. (2008). Interactions Between Jumbo Squid (*Dosidicus Gigas*) and Pacific Hake (*Merluccius Productus*) in the Northern California Current in 2007. *California Cooperative Oceanic Fisheries Investigations Reports*, *49*, 129–141.
- Hopcroft, J. E. (1979). *Introduction to Automata Theory, Languages and Computation* (3rd). Addison-Wesley Longman Publishing Co., Inc.
- Hoving, H. J. T., Gilly, W. F., Markaida, U., Benoit-Bird, K. J., -Brown, Z. W., Daniel, P., Field, J. C., Parassenti, L., Liu, B., & Campos, B. (2013). Extreme plasticity in life-history strategy allows a migratory predator (jumbo squid) to cope with a changing climate. *Global Change Biology*, *19*(7), 2089–2103. <https://doi.org/10.1111/gcb.12198>
- Hu, G., Boenish, R., Gao, C., Li, B., Chen, X., Chen, Y., & Punt, A. E. (2019). Spatio-temporal variability in trophic ecology of jumbo squid (*Dosidicus gigas*) in the southeastern Pacific: Insights from isotopic signatures in beaks. *Fisheries Research*, *212*(July 2018), 56–62. <https://doi.org/10.1016/j.fishres.2018.12.009>
- Huston, M., DeAngelis, D., & Post, W. (1988). New Computer Models Unify Ecological Theory. *BioScience*, *38*(10), 682–691. <https://doi.org/10.2307/1310870>
- Ibáñez, C. M., Cubillos, L. A., Tafur, R., Argüelles, J., Yamashiro, C., & Poulin, E. (2011). Genetic diversity and demographic history of *Dosidicus gigas* (Cephalopoda: Ommastrephidae) in the Humboldt Current System. *Marine Ecology Progress Series*, *431*, 163–171. <https://doi.org/10.3354/meps09133>
- Ibáñez, C. M., Arancibia, H., & Cubillos, L. A. (2008). Biases in determining the diet of jumbo squid *Dosidicus gigas* (D'Orbigny 1835) (Cephalopoda: Ommastrephidae) off southern-central Chile (34°S–40°S). *Helgoland Marine Research*, *62*(4), 331–338. <https://doi.org/10.1007/s10152-008-0120-0>
- Ibáñez, C. M., & Cubillos, L. A. (2007). Seasonal variation in the length structure and reproductive condition of the jumbo squid *Dosidicus gigas* (d'Orbigny, 1835) off central-south Chile. *Scientia Marina*, *71*(1), 123–128. <https://doi.org/10.3989/scimar.2007.71n1123>
- Ibáñez, C. M., & Keyl, F. (2010). Cannibalism in cephalopods. *Reviews in Fish Biology and Fisheries*, *20*(1), 123–136. <https://doi.org/10.1007/s11160-009-9129-y>

- Ibáñez, C. M., Sepúlveda, R. D., Ulloa, P., Keyl, F., & Pardo-Gandarillas, M. C. (2015). The biology and ecology of the jumbo squid *Dosidicus gigas* (Cephalopoda) in Chilean waters: a review. *43*(3), 402–414. <https://doi.org/10.3856/vol43-issue3-fulltext-2>
- Ibarra-García, L. E., Camarillo-Coop, S., & Salinas-Zavala, C. A. (2014). Cannibalism assessment of jumbo squid *Dosidicus gigas* from the Gulf of California. *Hidrobiológica*, *24*(1), 51–56.
- Ichii, T., Mahapatra, K., Watanabe, T., Yatsu, A., Inagake, D., & Okada, Y. (2002). Occurrence of jumbo flying squid *Dosidicus gigas* aggregations associated with the countercurrent ridge off the Costa Rica Dome during 1997 El Niño and 1999 La Niña. *Marine Ecology Progress Series*, *231*, 151–166. <https://doi.org/10.3354/meps231151>
- INRA. (2008). MOBIDYC. Retrieved November 18, 2017, from http://w3.avignon.inra.fr/mobidyc/index.php/English%7B%5C_%7Dsummary
- Irvine, W. B. (1989). Cannibalism, Vegetarianism, and Narcissism. *Between the Species: An Online Journal for the Study of Philosophy and Animals*, *5*(1), 11–17. <https://doi.org/10.15368/bts.1989v5n1.2>
- Itemis. (2020). XText. Retrieved November 14, 2020, from xtext.org
- Jereb, P., & Roper, C. F. E. (2010). Cephalopods of the World. *FAO Species Catalogue for Fishery Purposes*, *2*, 649. <http://www.vliz.be/imisdocs/publications/255659.pdf>
- Jetbrains. (2021). Kotlin. Retrieved July 15, 2021, from <https://kotlinlang.org>
- Jopp, F., Reuter, H., & Breckling, B. (2011). *Modelling Complex Ecological Dynamics - An Introduction into Ecological Modelling for Students, Teachers & Scientists* (F. Jopp, B. Breckling, & H. Reuter, Eds.; 1st). Springer-Verlag. <https://doi.org/10.1007/978-3-642-05029-9>
- JPL MUR MEaSURES Project. (2015). GHRSSST Level 4 MUR Global Foundation Sea Surface Temperature Analysis. <https://doi.org/10.5067/GHGMR-4FJ04>
- Judson, O. P. (1994). The rise of the individual-based model in ecology. *Trends in Ecology & Evolution*, *9*(1), 9–14. [https://doi.org/10.1016/0169-5347\(94\)90225-9](https://doi.org/10.1016/0169-5347(94)90225-9)
- Jupyter Project. (2021). Jupyter. Retrieved July 15, 2021, from <https://jupyter.org>
- Karasov, W. H. (1992). Daily Energy Expenditure and the Cost of Activity in Mammals. *American Zoologist*, *32*(2), 238–248.
- Katsanevakis, S., Stephanopoulou, S., Miliou, H., Moraitou-Apostolopoulou, M., & Verriopoulos, G. (2005). Oxygen consumption and ammonia excretion of Octopus

Bibliography

- vulgaris (Cephalopoda) in relation to body mass and temperature. *Marine Biology*, 146(4), 725–732. <https://doi.org/10.1007/s00227-004-1473-9>
- Keyl, F. (2009). *The Cephalopod Dosidicus gigas of the Humboldt Current System Under the Impact of Fishery and Environmental Variability* (Dissertation April). Bremen. <https://media.suub.uni-bremen.de/handle/elib/2714>
- Keyl, F., Argüelles, J., & Tafur, R. (2011). Interannual variability in size structure, age, and growth of jumbo squid (*Dosidicus gigas*) assessed by modal progression analysis. *ICES Journal of Marine Science*, 68(3), 507–518. <https://doi.org/10.1093/icesjms/fsq167>
- Keyl, F., Wolff, M., Argüelles, J., Mariátegui, L., Tafur, R., & Yamashiro, C. (2008). A hypothesis on range expansion and spatio-temporal shifts in size-at-maturity of jumbo squid (*Dosidicus gigas*) in the Eastern Pacific Ocean. *California Cooperative Oceanic Fisheries Investigations Reports*, 49, 119–128.
- Kinzey, D., Gerrodette, T., Barlow, J., Dizon, A., Perryman, W., & Olson, P. (2008). Marine Mammal Data Collected during a Survey in the Eastern Tropical Pacific Ocean Aboard the NOAA Ships McArthur and David Starr Jordan, July 28 – December 9, 1999. *NOAA Technical Memorandum*, 293, 89. <http://137.110.142.7/publications/TM/SWFSC/NOAA-TM-NMFS-SWFSC-421.pdf>
- Kleiber, M. (1932). Body size and metabolism. *Hilgardia*, 6, 315–352.
- Klügl, F., & Puppe, F. (1998). The Multi-Agent Simulation Environment SeSAM. *In Proceedings of the Workshop "Simulation and Knowledge-Based Systems", tr-ri.98-1*.
- Knuth, D. E. (1984). Literate Programming. *Computer*, 27(1), 97–111. <https://doi.org/10.1093/comjnl/27.2.97>
- Kornhauser, D., Rand, W., & Wilensky, U. (2007). *Visualization Tools for Agent-Based Modeling in NetLogo* (tech. rep.). <https://ccl.northwestern.edu/2007/Kornhauser-Agent-2007.pdf>
- Li, Y., Gong, Y., Zhang, Y., & Chen, X. (2017). Inter-annual variability in trophic patterns of jumbo squid (*Dosidicus gigas*) off the exclusive economic zone of Peru, implications from stable isotope values in gladius. *Fisheries Research*, 187(March), 22–30. <https://doi.org/10.1016/j.fishres.2016.11.005>
- Lin, B., Xin, L., Chen, J., Chen, Y., Yu, G., Wei, H., Tao, J., Jing, W., & Lin, Y. (2017). Periodic increments in the jumbo squid (*Dosidicus gigas*) beak: a potential tool for determining age and investigating regional difference in growth rates. *Hydrobiologia*, 790(1), 83–92. <https://doi.org/10.1007/s10750-016-3020-3>

- Lipinski, M. (2002). Growth of Cephalopods: Conceptual Model. In H. Summesberger, K. Histon, & A. Daurer (Eds.), *Cephalopods – present and past* (pp. 133–138). Abhandlungen der geologischen Bundesanstalt. https://www.zobodat.at/pdf/AbhGeolBA%7B%5C_%7D57%7B%5C_%7D0133-0138.pdf
- Lippe, M., Bithell, M., Gotts, N., Natalini, D., Barbrook-Johnson, P., Giupponi, C., Hallier, M., Hofstede, G. J., Le Page, C., Matthews, R. B., Schlüter, M., Smith, P., Teglio, A., & Thellmann, K. (2019). Using agent-based modelling to simulate social-ecological systems across scales. *GeoInformatica*, *23*(2), 269–298. <https://doi.org/10.1007/s10707-018-00337-8>
- Liu, B. L., Cao, J., Truesdell, S. B., Chen, Y., Chen, X. J., & Tian, S. Q. (2016). Reconstructing cephalopod migration with statolith elemental signatures: a case study using *Dosidicus gigas*. *Fisheries Science*, *82*(3), 425–433. <https://doi.org/10.1007/s12562-016-0978-8>
- Liu, B. L., Lin, J. Y., Chen, X. J., Jin, Y., & Wang, J. T. (2018). Inter- and intra-regional patterns of stable isotopes in *Dosidicus gigas* beak: Biological, geographical and environmental effects. *Marine and Freshwater Research*, *69*(3), 464–472. <https://doi.org/10.1071/MF17144>
- Liu, B. L., Xu, W., Chen, X. J., Huan, M. Y., & Liu, N. (2020). Ontogenetic shifts in trophic geography of jumbo squid, *Dosidicus gigas*, inferred from stable isotopes in eye lens. *Fisheries Research*, *226*(September 2019), 105507. <https://doi.org/10.1016/j.fishres.2020.105507>
- Liukkonen, L., Ayllón, D., Kunnasranta, M., Niemi, M., Nabe-Nielsen, J., Grimm, V., & Nyman, A. M. (2018). Modelling movements of Saimaa ringed seals using an individual-based approach. *Ecological Modelling*, *368*, 321–335. <https://doi.org/10.1016/j.ecolmodel.2017.12.002>
- Lorek, H., & Sonnenschein, M. (1999). Modelling and simulation software to support individual-based ecological modelling. *Ecological Modelling*, *115*(2–3), 199–216. [https://doi.org/10.1016/S0304-3800\(98\)00193-8](https://doi.org/10.1016/S0304-3800(98)00193-8)
- Lorscheid, I., Berger, U., Grimm, V., & Meyer, M. (2019). From cases to general principles: A call for theory development through agent-based modeling. *Ecological Modelling*, *393*, 153–156. <https://doi.org/10.1016/j.ecolmodel.2018.10.006>
- Luke, S. (2019). *Multiagent Simulation and the MASON Library* (20th ed.). George Mason University. <http://cs.gmu.edu/~%7B~%7Ddeclab/projects/mason/manual.pdf>

Bibliography

- Lytinen, S. L., & Railsback, S. F. (2010). The Evolution of Agent-based Simulation Platforms: A Review of NetLogo 5.0 and ReLogo. *European Meetings on Cybernetics and Systems Research*, 1–11.
- Manson, S., An, L., Clarke, K. C., Heppenstall, A., Koch, J., Krzyzanowski, B., Morgan, F., O'Sullivan, D., Runck, B. C., Shook, E., & Tesfatsion, L. (2020). Methodological issues of spatial agent-based models. *JASSS*, 23(1). <https://doi.org/10.18564/jasss.4174>
- Maplesoft. (2021). Maple. Retrieved July 15, 2021, from <https://www.maplesoft.com/products/Maple/>
- Marchini, A. (2011). Modelling Ecological Processes with Fuzzy Logic Approaches. In Fred Jopp, Hauke Reuter, & Broder Breckling (Eds.), *Modelling complex ecological dynamics* (pp. 133–145). Springer. <https://doi.org/10.1007/978-3-642-05029-9>
- Margolus, N., & Toffoli, T. (1987). Cellular Automata Machines. *Complex Systems*, 1, 967–993.
- Mariani, P., MacKenzie, B. R., Iudicone, D., & Bozec, A. (2010). Modelling retention and dispersion mechanisms of bluefin tuna eggs and larvae in the northwest Mediterranean Sea. *Progress in Oceanography*, 86(1–2), 45–58. <https://doi.org/10.1016/j.pocean.2010.04.027>
- Marinone, S. G., Gutiérrez, O. Q., & Parés-Sierra, A. (2004). Numerical simulation of larval shrimp dispersion in the Northern Region of the Gulf of California. *Estuarine, Coastal and Shelf Science*, 60(4), 611–617. <https://doi.org/10.1016/j.ecss.2004.03.002>
- Markaida, U., & Sosa-Nishizaki, O. (2001). Reproductive biology of jumbo squid *Dosidicus gigas* in the Gulf of California, 1995–1997. *Fisheries Research*, 54(1), 63–82. [https://doi.org/10.1016/S0165-7836\(01\)00373-3](https://doi.org/10.1016/S0165-7836(01)00373-3)
- Markaida, U. (2006a). Food and feeding of jumbo squid *Dosidicus gigas* in the Gulf of California and adjacent waters after the 1997–98 El Niño event. *Fisheries Research*, 79(1–2), 16–27. <https://doi.org/10.1016/j.fishres.2006.02.016>
- Markaida, U. (2006b). Population structure and reproductive biology of jumbo squid *Dosidicus gigas* from the Gulf of California after the 1997–1998 El Niño event. *Fisheries Research*, 79(1–2), 28–37. <https://doi.org/10.1016/j.fishres.2006.02.009>
- Markaida, U., Gilly, W. F., Salinas-Zavala, C. A., Rosas-Luis, R., & Booth, J. A. (2008). *Food and feeding of jumbo squid *Dosidicus gigas* in the central Gulf of California during 2005–2007* (tech. rep.). California Cooperative Oceanic Fisheries Investigations.

- Markaida, U., Quiñónez-Velázquez, C., & Sosa-Nishizaki, O. (2004). Age, growth and maturation of jumbo squid *Dosidicus gigas* (Cephalopoda: Ommastrephidae) from the Gulf of California, Mexico. *Fisheries Research*, *66*(1), 31–47. [https://doi.org/10.1016/S0165-7836\(03\)00184-X](https://doi.org/10.1016/S0165-7836(03)00184-X)
- Markaida, U., Rosenthal, J. J. C., & Gilly, W. F. (2005). Tagging studies on the jumbo squid (*Dosidicus gigas*) in the Gulf of California, Mexico. *Fishery Bulletin*, *103*(1), 219–226.
- Markaida, U., & Sosa-Nishizaki, O. (2003). Food and feeding habits of jumbo squid *Dosidicus gigas* (Cephalopoda: Ommastrephidae) from the Gulf of California, Mexico. *Journal of the Marine Biological Association of the UK*, *83*(3), 507–522. <https://doi.org/10.1017/S0025315403007434h>
- Martínez, G. J., Mcintosh, H. V., Mora, J. C. T., & Vergara, S. V. (2011). Reproducing the cyclic tag system developed by Matthew Cook with rule 110 using the phases $f_i - 1$. *Journal of Cellular Automata*, *6*(2–3), 121–161.
- Mejía-Rebollo, A., Salinas-Zavala, C., Quiñónez-Velázquez, C., & Markaida, U. (2008). Age, growth and maturity of jumbo squid (*Dosidicus gigas* D’Orbigny, 1835) off the western coast of the Baja California Peninsula. *California Cooperative Oceanic Fisheries Investigations Report*, *49*, 256–262. [https://doi.org/10.1016/S0165-7836\(03\)00184-X](https://doi.org/10.1016/S0165-7836(03)00184-X)
- Mooij, W. M., Brederveld, R. J., de Klein, J. J., DeAngelis, D. L., Downing, A. S., Faber, M., Gerla, D. J., Hipsey, M. R., ’t Hoen, J., Janse, J. H., Janssen, A. B., Jeuken, M., Kooi, B. W., Lischke, B., Petzoldt, T., Postma, L., Schep, S. A., Scholten, H., Teurlinckx, S., ... Kuiper, J. J. (2014). Serving many at once: How a database approach can create unity in dynamical ecosystem modelling. *Environmental Modelling and Software*, *61*, 266–273. <https://doi.org/10.1016/j.envsoft.2014.04.004>
- Murphy, K. J., Ciuti, S., & Kane, A. (2020). An introduction to agent-based models as an accessible surrogate to field-based research and teaching. *Ecology and Evolution*, *10*(22), 12482–12498. <https://doi.org/10.1002/ece3.6848>
- Nagy, K. A. (2005). Field metabolic rate and body size. *Journal of Experimental Biology*, *208*(9), 1621–1625. <https://doi.org/10.1242/jeb.01553>
- Neira, S., & Arancibia, H. (2013). Food web and fish stock changes in central Chile: comparing the roles of jumbo squid (*Dosidicus gigas*) predation, the environment, and fisheries. *Deep Sea Research Part II: Topical Studies in Oceanography*, *95*, 103–112. <https://doi.org/10.1016/j.dsr2.2013.04.003>

Bibliography

- Nelson, P., Radosavljević, M., & Bromberg, S. (2008). *Biological Physics: Energy, Information, Life* (updat. 1st). W. H. Freeman; Co.
- Netcdf4-python. (2020). netcdf4-python – Python/numpy interface to the netCDF C library. Retrieved November 18, 2020, from <https://github.com/Unidata/netcdf4-python>
- Nigmatullin, C. M., & Markaida, U. (2009). Oocyte development, fecundity and spawning strategy of large sized jumbo squid *Dosidicus gigas* (Oegopsida: Ommastrephinae). *Journal of the Marine Biological Association of the United Kingdom*, 89(4), 789–801. <https://doi.org/10.1017/S0025315408002853>
- Nigmatullin, C., Nesis, K., & Arkhipkin, A. (2001). A review of the biology of the jumbo squid *Dosidicus gigas* (Cephalopoda: Ommastrephidae). *Fisheries Research*, 54(1), 9–19. [https://doi.org/10.1016/S0165-7836\(01\)00371-X](https://doi.org/10.1016/S0165-7836(01)00371-X)
- Nikolai, C., & Gregory, M. (2009). Tools of the Trade: A Survey of Various Agent Based Modeling Platforms. *Journal of Artificial Societies and Social Simulation*, 12(2), 1–37.
- Nygaard, K., & Dahl, O.-J. (1978). *The development of the SIMULA languages*. Association for Computing Machinery. <https://doi.org/10.1145/800025.1198392>
- Oda, T. (2021). ReMobidyc. Retrieved July 18, 2021, from <https://github.com/tomooda/ReMobidyc>
- O’Dor, R. K., & Webber, D. M. (1991). Invertebrate Athletes: Trade-Offs between Transport Efficiency and Power Density in Cephalopod Evolution. *Journal of Experimental Biology*, 160(1), 93–112. <https://doi.org/10.1242/jeb.160.1.93>
- Olson, R. J., Román-Verdesoto, M. H., & Macías-Pita, G. L. (2006). Bycatch of jumbo squid *Dosidicus gigas* in the tuna purse-seine fishery of the eastern Pacific Ocean and predatory behaviour during capture. *Fisheries Research*, 79(1–2), 48–55. <https://doi.org/10.1016/j.fishres.2006.02.012>
- Parr, T. (2009). *Language Implementation Patterns: Create Your Own Domain-Specific and General Programming Languages* (1st). Pragmatic Bookshelf.
- Parr, T. (2013). *The Definitive ANTLR 4 Reference* (2nd). Pragmatic Bookshelf.
- Parr, T. (2017). ANTLR. Retrieved November 15, 2017, from <http://www.antlr.org>
- Pauly, D. (1981). The relationships between gill surface area and growth performance in fish: a generalization of von Bertalanffy’s theory of growth. *Berichte der Deutschen Wissenschaftlichen Kommission für Meeresforschung*, 28(4), 251–282.

- Pecl, G. T., & Jackson, G. D. (2008). The potential impacts of climate change on inshore squid: biology, ecology and fisheries. *Reviews in Fish Biology and Fisheries*, 18(4), 373–385. <https://doi.org/10.1007/s11160-007-9077-3>
- PO.DAAC. (2021a). MEaSUREs - Multi-scale Ultra-high Resolution (MUR) Sea Surface Temperature (MEaSUREs-MUR). <https://doi.org/10.5067/GHGMR-4FJ04>
- PO.DAAC. (2021b). OSCAR third degree resolution ocean surface currents. <https://doi.org/10.5067/OSCAR-03D01>
- Polis, G. A. (1981). The Evolution and Dynamics of Intraspecific Predation. *Annual Review of Ecology and Systematics*, 12(1), 225–251. <https://doi.org/10.1146/annurev.es.12.110181.001301>
- Portner, E. J., Markaida, U., Robinson, C. J., & Gilly, W. F. (2020). Trophic ecology of Humboldt squid, *Dosidicus gigas*, in conjunction with body size and climatic variability in the Gulf of California, Mexico. *Limnology and Oceanography*, 65(4), 732–748. <https://doi.org/10.1002/lno.11343>
- Python Package Authority. (2017). Pip installs packages (pip). Retrieved October 31, 2017, from <https://pip.pypa.io/en/stable/>
- Python Software Foundation. (2020). Python. Retrieved November 18, 2020, from <https://www.python.org>
- Railsback, S., Ayllón, D., Berger, U., Grimm, V., Lytinen, S., Sheppard, C., & Thiele, J. (2017). Improving Execution Speed of Models Implemented in NetLogo. *JASSS*, 20(1). <https://doi.org/10.18564/jasss.3282>
- Railsback, S. F., & Grimm, V. (2012). *Agent-Based and Individual-Based Modeling: A Practical Introduction*. Princeton University Press.
- Railsback, S. F., & Grimm, V. (2019). *Agent-Based and Individual-Based Modeling: A Practical Introduction* (Second). Princeton University Press.
- Railsback, S. F., & Harvey, B. C. (2020). *Modeling Populations of Adaptive Individuals* (Vol. 44). Princeton University Press. <https://doi.org/10.1088/1751-8113/44/8/085201>
- Railsback, S. F., Lytinen, S. L., & Jackson, S. K. (2006). Agent-based Simulation Platforms: Review and Development Recommendations. *SIMULATION: Transactions of The Society for Modeling and Simulation International*, 82(9), 609–623. <https://doi.org/10.1177/0037549706073695>
- Ramos-Castillejos, J. E., Salinas-Zavala, C. A., Camarillo-Coop, S., & Enríquez-Paredes, L. M. (2010). Paralarvae of the jumbo squid, *Dosidicus gigas*. *Invertebrate Biology*, 129(2), 172–183. <https://doi.org/10.1111/j.1744-7410.2010.00194.x>

Bibliography

- Razavi, R., Perrot, J. F., & Guelfi, N. (2005). Adaptive modeling: An Approach and a Method for Implementing Adaptive Agents. *Lecture Notes in Artificial Intelligence*, 3446, 136–148. <https://doi.org/10.1007/b138262>
- Reuter, H., Breckling, B., & Jopp, F. (2011). Individual-Based Models. In F. Jopp, H. Reuter, & B. Breckling (Eds.), *Modelling complex ecological dynamics* (pp. 163–178). Springer. <https://doi.org/10.1007/978-3-642-05029-9>
- Reynolds, C. W. (1987). Flocks, herds and schools: A distributed behavioral model. *ACM SIGGRAPH Computer Graphics*, 21(4), 25–34. <https://doi.org/10.1145/37402.37406>
- Robinson, C. J., Gómez-Gutiérrez, J., & de León, D. A. S. (2013). Jumbo squid (*Dosidicus gigas*) landings in the Gulf of California related to remotely sensed SST and concentrations of chlorophyll a (1998–2012). *Fisheries Research*, 137, 97–103. <https://doi.org/10.1016/j.fishres.2012.09.006>
- Rodhouse, P. G., & Hatfield, E. M. (1990). Age Determination in Squid Using Statolith Growth Increments. *Fisheries Research*, 8(4), 323–334. [https://doi.org/10.1016/0165-7836\(90\)90002-D](https://doi.org/10.1016/0165-7836(90)90002-D)
- Rodhouse, P. G. (2008). Large-scale range expansion and variability in ommastrephid squid populations: A review of environmental links. *CalCOFI Report*, 49, 83–89.
- Rodhouse, P. G., Waluda, C. M., Morales-Bojórquez, E., & Hernández-Herrera, A. (2006). Fishery biology of the Humboldt squid, *Dosidicus gigas*, in the Eastern Pacific Ocean. *Fisheries Research*, 79(1–2), 13–15. <https://doi.org/10.1016/j.fishres.2006.02.008>
- Roper, C. F. E., Sweeney, M. J., & Nauen, C. E. (1984). *FAO species catalogue - Cephalopods of the world* (Vol. 3). Food; Agriculture Organization of the United Nations.
- Rosa, R., & Seibel, B. (2006). Effect of high CO₂ on the metabolism of jumbo squid *Dosidicus gigas*.
- Rosa, R., & Seibel, B. A. (2008). Synergistic effects of climate-related variables in a top oceanic predator metabolic rate. *Proceedings of the National Academy of Sciences*, 105(52), 20776–20780. <https://doi.org/www.pnas.org/cgi/doi/10.1073/pnas.0806886105>
- Rosa, R., & Seibel, B. A. (2010). Metabolic physiology of the Humboldt squid, *Dosidicus gigas*: Implications for vertical migration in a pronounced oxygen minimum zone. *Progress in Oceanography*, 86(1), 72–80. <https://doi.org/10.1016/j.pocean.2010.04.004>

- Rosa, R., Yamashiro, C., Markaida, U., Rodhouse, P. G. K., Waluda, C. M., Salinas-Zavala, C. A., Keyl, F., O'Dor, R., Stewart, J. S., & Gilly, W. F. (2013). *Dosidicus gigas*, Humboldt Squid. In R. Rosa, G. Pierce, & R. O'Dor (Eds.), *Advances in squid biology, ecology and fisheries. Part II Oegopsid Squids* (p. 39). Nova Science Publishers, Inc. <https://doi.org/10.13140/2.1.4154.2083>
- Rosas-Luis, R., & Chompoy-Salazar, L. (2016). Description of food sources used by jumbo squid *Dosidicus gigas* (D'Orbigny, 1835) in Ecuadorian waters during 2014. *Fisheries Research*, *173*(2009), 139–144. <https://doi.org/10.1016/j.fishres.2015.08.006>
- Roux, C., Fraïsse, C., Romiguié, J., Anciaux, Y., Galtier, N., & Bierne, N. (2016). Shedding Light on the Grey Zone of Speciation along a Continuum of Genomic Divergence (C. Moritz, Ed.). *PLoS Biology*, *14*(12), 1–22. <https://doi.org/10.1371/journal.pbio.2000234>
- Ruiz-Cooley, R. I., Villa, E. C., & Gould, W. R. (2010). Ontogenetic variation of $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$ recorded in the gladius of the jumbo squid *Dosidicus gigas*: Geographic differences. *Marine Ecology Progress Series*, *399*, 187–198. <https://doi.org/10.3354/meps08383>
- Ruiz-Cooley, R., & Markaida, U. (2006). Use of stable isotopes to examine foraging ecology of jumbo squid (*Dosidicus gigas*). *Global Ocean Ecosystem Dynamics*, *62*. <https://doi.org/10.1.1.134.6845>
- Ruiz-Cooley, R. I., Ballance, L. T., & McCarthy, M. D. (2013). Range Expansion of the Jumbo Squid in the NE Pacific: $\delta^{15}\text{N}$ Decrypts Multiple Origins, Migration and Habitat Use. *PLoS ONE*, *8*(3), e59651, 1–7. <https://doi.org/10.1371/journal.pone.0059651>
- Sakai, M., Mariátegui, L., Wakabayashi, T., & Tuchiya, K. (2008). Distribution and abundance of jumbo flying squid paralarvae (*Dosidicus gigas*) off Peru and in waters west of the Costa Rica Dome during the 2007 La Niña. *4th Int Symp Pacific Squids*, 95–97.
- Sakai, M., Tuchiya, K., Mariategui, L., Wakabayashi, T., & Yamashiro, C. (2017). Vertical Migratory Behavior of Jumbo Flying Squid (*Dosidicus gigas*) off Peru: Records of Acoustic and Pop-up Tags. *Japan Agricultural Research Quarterly*, *51*(2), 171–179. <https://doi.org/10.6090/jarq.51.171>
- Sanchez, G., Kawai, K., Yamashiro, C., Fujita, R., Wakabayashi, T., Sakai, M., & Umino, T. (2020). Patterns of mitochondrial and microsatellite DNA markers describe historical and contemporary dynamics of the Humboldt squid *Dosidicus*

Bibliography

- gigas in the Eastern Pacific Ocean. *Reviews in Fish Biology and Fisheries*, 30(3), 519–533. <https://doi.org/10.1007/s11160-020-09609-9>
- Sanchez, G., Tomano, S., Yamashiro, C., Fujita, R., Wakabayashi, T., Sakai, M., & Umino, T. (2016). Population genetics of the jumbo squid *Dosidicus gigas* (Cephalopoda: Ommastrephidae) in the northern Humboldt Current system based on mitochondrial and microsatellite DNA markers. *Fisheries Research*, 175(March), 1–9. <https://doi.org/10.1016/j.fishres.2015.11.005>
- Sánchez-Velasco, L., Ruvalcaba-Aroche, E. D., Beier, E., Godínez, V. M., Barton, E. D., Díaz-Viloria, N., & Pacheco, M. R. (2016). Paralarvae of the complex *Sthenoteuthis oualaniensis*-*Dosidicus gigas* (Cephalopoda: Ommastrephidae) in the northern limit of the shallow oxygen minimum zone of the Eastern Tropical Pacific Ocean (April 2012). *Journal of Geophysical Research: Oceans*, 121(3), 1998–2015. <https://doi.org/10.1002/2015JC011534>
- Sandoval-Castellanos, E., Uribe-Alcocer, M., & Díaz-Jaimes, P. (2007). Population genetic structure of jumbo squid (*Dosidicus gigas*) evaluated by RAPD analysis. *Fisheries Research*, 83(1), 113–118. <https://doi.org/10.1016/j.fishres.2006.09.007>
- Sandoval-Castellanos, E., Uribe-Alcocer, M., & Díaz-Jaimes, P. (2009). Lack of genetic differentiation among size groups of jumbo squid (*Dosidicus gigas*). *Ciencias Marinas*, 35(4), 419–428. <https://doi.org/10.3892/ijmm.11.2.139>
- Sandoval-Castellanos, E., Uribe-Alcocer, M., & Díaz-Jaimes, P. (2010). Population genetic structure of the Humboldt squid (*Dosidicus gigas* d'Orbigny, 1835) inferred by mitochondrial DNA analysis. *Journal of Experimental Marine Biology and Ecology*, 385(1–2), 73–78. <https://doi.org/10.1016/j.jembe.2009.12.015>
- Scala Center. (2021). Scala Programming Language. Retrieved May 4, 2021, from <https://www.scala-lang.org>
- ScalaJS. (2021). Scala.js. Retrieved July 15, 2021, from <https://www.scala-js.org>
- Scheffer, M., Baveco, J. M., DeAngelis, D. L., Rose, K. A., & van Nes, E. H. (1995). Super-individuals a simple solution for modelling large populations on an individual basis. *Ecological Modelling*, 80(2–3), 161–170. [https://doi.org/10.1016/0304-3800\(94\)00055-M](https://doi.org/10.1016/0304-3800(94)00055-M)
- Scherer, C., Radchuk, V., Franz, M., Thulke, H. H., Lange, M., Grimm, V., & Kramer-Schadt, S. (2020). Moving infections: individual movement decisions drive disease persistence in spatially structured landscapes. *Oikos*, (December 2019), 651–667. <https://doi.org/10.1111/oik.07002>

- Schulze, J., Müller, B., Groeneveld, J., & Grimm, V. (2017). Agent-based modelling of social-ecological systems: Achievements, challenges, and a way forward. *JASSS*, *20*(2). <https://doi.org/10.18564/jasss.3423>
- SciPy.org. (2020a). Matplotlib. Retrieved November 18, 2020, from <https://www.matplotlib.org>
- SciPy.org. (2020b). NumPy. Retrieved November 18, 2020, from <https://www.numpy.org>
- Seibel, B. A. (2007). On the depth and scale of metabolic rate variation: scaling of oxygen consumption rates and enzymatic activity in the Class Cephalopoda (Mollusca). *Journal of Experimental Biology*, *210*(1), 1–11. <https://doi.org/10.1242/jeb.02588>
- Seibel, B. A. (2011). Critical oxygen levels and metabolic suppression in oceanic oxygen minimum zones. *Journal of Experimental Biology*, *214*(2), 326–336. <https://doi.org/10.1242/jeb.049171>
- Seibel, B. A., Rosa, R., & Trueblood, L. (2007). Cephalopod metabolism as a function of body size. *The Role of Squid in Open Ocean Ecosystems*, 16–17(11), 9–10.
- Seibel, B. A., Thuesen, E. V., & Childress, J. J. (2000). Light-Limitation on Predator-Prey Interactions: Consequences for Metabolism and Locomotion of Deep-Sea Cephalopods. *Biological Bulletin*, *198*(2), 284–298. <https://doi.org/10.2307/1542531>
- Semmens, J. M., Pecl, G. T., Villanueva, R., Jouffre, D., Sobrino, I., Wood, J. B., & Rigby, P. R. (2004). Understanding octopus growth: patterns, variability and physiology. *Marine and Freshwater Research*, *55*(4), 367–377. <https://doi.org/10.1071/MF03155>
- SeSAM. (2012a). Multiagent Simulation | SeSAM. Retrieved December 8, 2015, from <http://130.243.124.21/sesam/>
- SeSAM. (2012b). SeSAM-Wiki. Retrieved December 8, 2015, from http://130.243.124.21/mediawiki/index.php/Main%7B%5C_%7DPage
- Shum, S., & Cook, C. (1994). Using literate programming to teach good programming practices. *ACM SIGCSE Bulletin*, *26*(1), 66–70. <https://doi.org/10.1145/191033.191059>
- Smith, C., & Reay, P. (1991). Cannibalism in teleost fish. *Reviews in Fish Biology and Fisheries*, *1*(1), 41–64. <https://doi.org/10.1007/BF00042661>
- Speakman, J. R. (2000). The Cost of Living: Field Metabolic Rates of Small Mammals. *Advances in Ecological Research*, *30*, 177–297. [https://doi.org/10.1016/S0065-2504\(08\)60019-7](https://doi.org/10.1016/S0065-2504(08)60019-7)

Bibliography

- Spotnitz, R. (1998). Literate programming. *Computers and Chemical Engineering*, 22(12), 1745–1747. [https://doi.org/10.1016/S0098-1354\(98\)00029-5](https://doi.org/10.1016/S0098-1354(98)00029-5)
- Staaf, D. J. (2010). *Reproduction and early life of the Humboldt squid* (Dissertation). Stanford University. https://stacks.stanford.edu/file/druid:cq221nc2303/Staaf%7B%5C_%7DThesis-augmented.pdf
- Staaf, D. J., Camarillo-Coop, S., Haddock, S. H., Nyack, A. C., Payne, J., Salinas-Zavala, C. A., Seibel, B. A., Trueblood, L., Widmer, C., & Gilly, W. F. (2008). Natural egg mass deposition by the Humboldt squid (*Dosidicus gigas*) in the Gulf of California and characteristics of hatchlings and paralarvae. *Journal of the Marine Biological Association of the United Kingdom*, 88(04), 759–770. <https://doi.org/10.1017/S0025315408001422>
- Staaf, D. J., Redfern, J. V., Gilly, W. F., Watson, W., & Ballance, L. T. (2013). Distribution of ommastrephid paralarvae in the eastern tropical Pacific. *Fishery Bulletin*, 111(1), 78–89. <https://doi.org/10.7755/FB.111.1.7>
- Staaf, D. J., Ruiz-Cooley, R. I., Elliger, C., Lebaric, Z., Campos, B., Markaida, U., & Gilly, W. F. (2010a). Ommastrephid squids *Sthenoteuthis oualaniensis* and *Dosidicus gigas* in the eastern Pacific show convergent biogeographic breaks but contrasting population structures. <https://doi.org/10.3354/meps08829>
- Staaf, D. J., Ruiz-Cooley, R. I., Elliger, C., Lebaric, Z., Campos, B., Markaida, U., & Gilly, W. F. (2010b). Ommastrephid squids *Sthenoteuthis oualaniensis* and *Dosidicus gigas* in the eastern Pacific show convergent biogeographic breaks but contrasting population structures. *Marine Ecology Progress Series*, 418, 165–178. <https://doi.org/10.3354/meps08829>
- Staaf, D. J., Zeidberg, L. D., & Gilly, W. F. (2011). Effects of temperature on embryonic development of the Humboldt squid *Dosidicus gigas*. *Marine Ecology Progress Series*, 441(Arnold 1974), 165–175. <https://doi.org/10.3354/meps09389>
- Stewart, J. S., Hazen, E. L., Foley, D. G., Bograd, S. J., & Gilly, W. F. (2012). Marine predator migration during range expansion: Humboldt squid *Dosidicus gigas* in the northern California Current System. *Marine Ecology Progress Series*, 471, 135–150. <https://doi.org/10.3354/meps10022>
- Stewart, J. S., Field, J. C., Markaida, U., & Gilly, W. F. (2013). Behavioral ecology of jumbo squid (*Dosidicus gigas*) in relation to oxygen minimum zones. *Deep-Sea Research Part II: Topical Studies in Oceanography*, 95, 197–208. <https://doi.org/10.1016/j.dsr2.2012.06.005>
- Stewart, J. S., Hazen, E. L., Bograd, S. J., Byrnes, J. E. K., Foley, D. G., Gilly, W. F., Robison, B. H., & Field, J. C. (2014). Combined climate- and prey-mediated

- range expansion of Humboldt squid (*Dosidicus gigas*), a large marine predator in the California Current System. *Global Change Biology*, 20(6), 1832–1843. <https://doi.org/10.1111/gcb.12502>
- Stockhausen, W., & Hermann, A. (2007). Modeling Larval Dispersion of Rockfish: A Tool for Marine Reserve Design? *Biology, Assessment, and Management of North Pacific Rockfishes*, 251–273. <https://doi.org/10.4027/bamnpr.2007.15>
- Tafur, R., Keyl, F., & Argüelles, J. (2010). Reproductive biology of jumbo squid *Dosidicus gigas* in relation to environmental variability of the northern Humboldt Current System. *Marine Ecology Progress Series*, 400(Table 1), 127–141. <https://doi.org/10.3354/meps08386>
- Tafur, R., Villegas, P., Rabí, M., & Yamashiro, C. (2001). Dynamics of maturation, seasonality of reproduction and spawning grounds of the jumbo squid *Dosidicus gigas* (Cephalopoda: Ommastrephidae) in Peruvian waters. *Fisheries Research*, 54(1), 33–50. [https://doi.org/10.1016/S0165-7836\(01\)00379-4](https://doi.org/10.1016/S0165-7836(01)00379-4)
- Tafur, R., & Rabí, M. (1997). Reproduction of the jumbo flying squid, *Dosidicus gigas* (Orbigny, 1835) (Cephalopoda: Ommastrephidae) off Peruvian coasts. *Scientia Marina*, 61, 33–37. <http://www.icm.csic.es/scimar/pdf/61/sm61s2033.pdf>
- Trasviña-Carrillo, L. D., Hernández-Herrera, A., Torres-Rojas, Y. E., Galván-Magaña, F., Sánchez-González, A., & Aguñiga-García, S. (2018). Spatial and trophic preferences of jumbo squid *Dosidicus gigas* (D'Orbigny, 1835) in the central Gulf of California: ecological inferences using stable isotopes. *Rapid Communications in Mass Spectrometry*, 32(15), 1225–1236. <https://doi.org/10.1002/rcm.8147>
- Trübenbach, K., Pegado, M. R., Seibel, B. A., & Rosa, R. (2013). Ventilation rates and activity levels of juvenile jumbo squids under metabolic suppression in the oxygen minimum zones. *Journal of Experimental Biology*, 216(3), 359–368. <https://doi.org/10.1242/jeb.072587>
- Trübenbach, K., da Costa, G., Ribeiro-Silva, C., Ribeiro, R. M., Cordeiro, C., & Rosa, R. (2014). Hypoxia-driven selective degradation of cellular proteins in jumbo squids during diel migration to the oxygen minimum zones. *Marine Biology*, 161(3), 575–584. <https://doi.org/10.1007/s00227-013-2360-z>
- Trübenbach, K., Teixeira, T., Diniz, M., & Rosa, R. (2013). Hypoxia tolerance and antioxidant defense system of juvenile jumbo squids in oxygen minimum zones. *Deep-Sea Research Part II: Topical Studies in Oceanography*, 95, 209–217. <https://doi.org/10.1016/j.dsr2.2012.10.001>
- Trueblood, L. A., & Seibel, B. A. (2013). The jumbo squid, *Dosidicus gigas* (Ommastrephidae), living in oxygen minimum zones I: Oxygen consumption rates and

Bibliography

- critical oxygen partial pressures. *Deep-Sea Research Part II: Topical Studies in Oceanography*, 95, 218–224. <https://doi.org/10.1016/j.dsr2.2012.10.004>
- van den Bosch, F., de Roos, A. M., & Gabriel, W. (1988). Cannibalism as a life boat mechanism. *Journal of Mathematical Biology*, 26(6), 619–633. <https://doi.org/10.1007/BF00276144>
- van Huis, A., Woldewahid, G., Toleubayev, K., & van der Werf, W. (2008). Relationships between food quality and fitness in the desert locust, *Schistocerca gregaria*, and its distribution over habitats on the Red Sea coastal plain of Sudan. *Entomologia Experimentalis et Applicata*, 127(2), 144–156. <https://doi.org/10.1111/j.1570-7458.2008.00682.x>
- Vecchione, M. (1999). Extraordinary abundance of squid paralarvae in the tropical eastern Pacific Ocean during El Niño of 1987. *Fish. Bull.*, 97(November), 1025–1030.
- Velquez-Abunader, J. I., Hernández-Herrera, A., Martínez-Aguilar, S., Díaz-Uribe, J. G., & Morales-Bojórquez, E. (2012). Interannual variability in mantle length structure, recruitment, and sex ratio of jumbo squid, *Dosidicus gigas*, in the Central Gulf of California, Mexico. *Journal of Shellfish Research*, 31(1), 125–134. <https://doi.org/10.2983/035.031.0116>
- Vincenot, C. E., Giannino, F., Rietkerk, M., Moriya, K., & Mazzoleni, S. (2011). Theoretical considerations on the combined use of System Dynamics and individual-based modeling in ecology. *Ecological Modelling*, 222(1), 210–218. <https://doi.org/10.1016/j.ecolmodel.2010.09.029>
- Waluda, C. M., & Rodhouse, P. G. (2005). *Dosidicus gigas* fishing grounds in the eastern pacific as revealed by satellite imagery of the light-fishing fleet. *Systems Research*, 328(February), 321–328.
- Waluda, C. M., & Rodhouse, P. G. (2006). Remotely sensed mesoscale oceanography of the Central Eastern Pacific and recruitment variability in *Dosidicus gigas*. *Marine Ecology Progress Series*, 310, 25–32. <https://doi.org/10.3354/meps310025>
- Whitaker, J. (2018). Basemap Matplotlib Toolkit. Retrieved March 25, 2018, from matplotlib.org/basemap/
- Wilensky, U. (1997). Fireflies. Retrieved July 15, 2021, from <http://ccl.northwestern.edu/netlogo/models/Fireflies>
- Winberg, G. (1960). *Rate of metabolism and food requirements of fishes* (Vol. 194). Fisheries Research Board of Canada.
- Wolfram Research Inc. (2021). Mathematica, Version 12.3.1. Retrieved July 15, 2021, from <https://www.wolfram.com/mathematica>

- Xavier, J. C., Allcock, A. L., Cherel, Y., Lipinski, M. R., Pierce, G. J., Rodhouse, P. G., Rosa, R., Shea, E. K., Strugnell, J. M., Vidal, E. A., Villanueva, R., & Ziegler, A. (2015). Future challenges in cephalopod research. *Journal of the Marine Biological Association of the United Kingdom*, *95*(05), 999–1015. <https://doi.org/10.1017/S0025315414000782>
- Xu, H., Wang, C., Liu, Y., Liu, B., & Li, G. (2021). Development and characterization of 101 SNP markers in jumbo flying squid, *Dosidicus gigas*. *Conservation Genetics Resources*, *13*(1), 13–20. <https://doi.org/10.1007/s12686-020-01177-1>
- Xu, W., Chen, X., Liu, B., Chen, Y., Huan, M., Liu, N., & Lin, J. (2019). Inter-individual variation in trophic history of *Dosidicus gigas*, as indicated by stable isotopes in eye lenses. *Aquaculture and Fisheries*, *4*(6), 261–267. <https://doi.org/10.1016/j.aaf.2019.05.001>
- Yatsu, A. (2000). Age Estimation of Four Oceanic Squids, *Ommastrephes bartramii*, *Dosidicus gigas*, *Sthenoteuthis oualaniensis*, and *Illex argentinus* (Cephalopoda, Ommastrephidae) Based on Statolith Microstructure. *Japan Agricultural Research Quarterly*, *34*(1), 75–80.
- Yatsu, A., Tafur, R., & Maravi, C. (1999). Embryos and Rhynchoteuthion Paralarvae of the Jumbo Flying Squid *Dosidicus gigas* (Cephalopoda) Obtained through Artificial Fertilization from Peruvian Waters. *Fisheries Science*, *65*(6), 904–908.
- Young, J. W., Olson, R. J., & Rodhouse, P. G. (2013). The role of squids in pelagic ecosystems: An overview. *Deep-Sea Research Part II: Topical Studies in Oceanography*, *95*, 3–6. <https://doi.org/10.1016/j.dsr2.2013.05.008>
- Young, R. E., & Vecchione, M. (2013). *Dosidicus gigas*. Retrieved November 18, 2020, from http://tolweb.org/Dosidicus%7B%5C_%7Dgigas/19945/2013.01.08
- Yu, W., & Chen, X. (2018). Ocean warming-induced range-shifting of potential habitat for jumbo flying squid *Dosidicus gigas* in the Southeast Pacific Ocean off Peru. *Fisheries Research*, *204* (August), 137–146. <https://doi.org/10.1016/j.fishres.2018.02.016>
- Yu, W., Chen, X., & Liu, L. (2021). Synchronous Variations in Abundance and Distribution of *Ommastrephes bartramii* and *Dosidicus gigas* in the Pacific Ocean. *Journal of Ocean University of China*, *20*(January), 11. <https://doi.org/10.1007/s11802-021-4644-0>
- Yu, W., Chen, X., & Zhang, Y. (2019). Seasonal habitat patterns of jumbo flying squid *Dosidicus gigas* off Peruvian waters. *Journal of Marine Systems*, *194*(February), 41–51. <https://doi.org/10.1016/j.jmarsys.2019.02.011>

Bibliography

- Yu, W., Wen, J., Chen, X., & Liu, B. (2021). El Niño–Southern Oscillation impacts on jumbo squid habitat: Implication for fisheries management. *Aquatic Conservation: Marine and Freshwater Ecosystems*, (November 2020), 1–12. <https://doi.org/10.1002/aqc.3584>
- Yu, W., Yi, Q., Chen, X., & Chen, Y. (2015). Modelling the effects of climate variability on habitat suitability of jumbo flying squid, *Dosidicus gigas*, in the Southeast Pacific Ocean off Peru. *ICES Journal of Marine Science*, *73*(2), 239–249. <https://doi.org/10.1093/icesjms/fsv223>
- Zeidberg, L. D., & Robison, B. H. (2007). Invasive range expansion by the Humboldt squid, *Dosidicus gigas*, in the eastern North Pacific. *Proceedings of the National Academy of Sciences*, *104*(31), 12948–12950. <https://doi.org/10.1073/pnas.0702043104>
- Zepeda-Benitez, V. Y., Quiñonez-Velázquez, C., Morales-Bojórquez, E., & Salinas-Zavala, C. A. (2014). Age and Growth Modelling for Early Stages of the Jumbo Squid *Dosidicus Gigas* Using Multi-Model Inference. *CalCOFI Rep.*, *55*, 197–204.
- Zhang, B., & DeAngelis, D. L. (2020). An overview of agent-based models in plant biology and ecology. *Annals of Botany*, *126*(4), 539–557. <https://doi.org/10.1093/aob/mcaa043>
- Zhu, L., Qualls, W. A., Marshall, J. M., Arheart, K. L., DeAngelis, D. L., McManus, J. W., Traore, S. F., Doumbia, S., Schlein, Y., Müller, G. C., & Beier, J. C. (2015). A spatial individual-based model predicting a great impact of copious sugar sources and resting sites on survival of *Anopheles gambiae* and malaria parasite transmission. *Malaria Journal*, *14*(1), 1–15. <https://doi.org/10.1186/s12936-015-0555-0>
- Zúñiga, M. J., Cubillos, L. A., & Ibáñez, C. M. (2008). A regular pattern of periodicity in the monthly catch of jumbo squid (*Dosidicus gigas*) along the Chilean coast (2002–2005). *Ciencias Marinas*, *34*(1), 91–99. <https://doi.org/10.7773/cm.v34i1.1138>

A. Python scripts to compute SST-data and paralarvae distribution

A.1. Overview

In this thesis, several scripts in various programming languages are used to process data from different sources. For a setup to process the data and retrieve the results, these scripts and actions must be executed in the following order :

1. Use pip (Python Package Authority, 2017) to install the libraries `pysnp` (GeospatialPython, 2017), `basemap` and `netcdf4` in the PYTHON programming environment.
2. Run `Prepare.py` to create the directory structure (Appendix A.3) required for script execution.
3. Copy the files `DIST.cst`, `DIST.dbf`, `DIST.prj`, `DIST.shp` and `DIST.shx` to the directory `"/originalData/GIST/"`.
4. Copy the files `bongo.csv`, `ethanol.csv` and `manta.csv` to the `"/originalData/Staaf/"` directory.
5. Copy the files `20030101-JPL-L4UHfnd-GLOB-v01-fv04-MUR.nc` until `20161231-JPL-L4UHfnd-GLOB-v01-fv04-MUR.nc`, about 550GB, to the directory `"/originalData/MUR/"`.
6. Copy the files `oscar_vel1992.nc` to `oscar_vel2015.nc`, about 31GB, to the directory `"/originalData/OSCAR/"`.
7. Optional: Execute the script `SSTPreprocessMURMain.py` to convert the `20030101-JPL-L4UHfnd-GLOB-v01-fv04-MUR.nc` to `20161231-JPL-L4UHfnd-GLOB-v01-fv04-MUR.nc` for faster computations to equivalent `.npy`-files (`binMURSST2003-01-01.npy` to `binMURSST2016-12-31.npy`). The script writes longitudinal and latitudinal information to the `"[Magic.dataRootPath]/results/MUR/"` directory.

The scripts share a common code basis by importing supporting scripts which will not be explained in detail. Most notably, the file `Magic.py` contains most "magic" values, e.g., hard encoded constants like paths and numbers.

A. PYTHON scripts to compute SST-data and paralarvae distribution

The scripts are written in PYTHON version 3.x. To execute a python script, the PYTHON programming language has to be installed on the target system, for example the JETBRAINS PYCHARM-IDE.

A.2. Supporting script files

The supporting script files provide functionality used among other script files:

1. `Animation.py`: Provides basic functionality for the display and saving of mapped data or video sequences (movies).
2. `Compute.py`: General computations required by different scripts.
3. `DrawDistributionRange.py`: General functionality to draw the distribution range of *D. gigas*.
4. `FileUtils.py`: General functionality to access files, directory structures, list of files etc.
5. `Magic.py`: The definition of “magic” constants, literals, paths etc. This file maps values to semantic values by providing a name.
6. `SSTReadData.py`: Functions and definition related to read and write SST data.
7. `Utils.py`: General utility functions which are semantically not grouped.

A.3. Preparation of directory structure

Objective. Creation of the necessary directory structure for storing the data.

Method. The script creates the directory structure on the destination storage device, defined by the `Magic.dataRootPath` in `Magic.py`. No files are deleted or altered; no directories are removed.

Script files.

1. `Prepare.py`
2. imported `FileUtils.py`
3. imported `Magic.py`

A.4. Display of the distribution range according to FAO

Results. The execution of the script results in a directory tree structure

```
[Magic.dataRootPath]/
├── modifiedData
│   ├── MUR
│   └── OSCAR
├── originalData
│   ├── GIS
│   ├── MUR
│   ├── OSCAR
│   └── Staaf
└── results
    ├── MUR
    └── OSCAR
```

below the global `Magic.dataRootPath`.

A.4. Display of the distribution range according to FAO

Objective. Display of the distribution range of *D. gigas* with geographic information system (GIS) data provided by Food and Agriculture Organization of the United Nations (FAO).

Script files.

1. FAODistributionMain.py
2. imported DrawDistributionRange.py
3. imported Magic.py

Method. The FAODistributionMain.py reads the GIS-information about the distribution range contributed by the FAO. The GIS-information is located in the directory “[`Magic.dataRootPath`]/originalData/GIS/”.

The data of the GIS-file is analyzed, the polygons extracted and projected onto a map contributed by the basemap library; the result is displayed in Figure 2.5, page 17.

Results. The script writes the distribution range to the portable document format (PDF)-file [`Magic.dataRootPath`]/results/DosidicusDistribution.pdf.

A.5. *D. gigas* sightings

Objective. Display of documented sightings of *D. gigas*.

A. PYTHON scripts to compute SST-data and paralarvae distribution

Script files.

1. OccurrencesMain.py
2. imported Magic.py

Method. The OccurrencesMain.py script reads information about documented *D. gigas* sightings and displays these in a geographical map, see Figure 2.6, page 18.

Results. The script writes these spottings to the PDF-file [Magic.dataRootPath]/results/DosidicusSpotted.pdf.

A.6. Preprocessing of the SST-data

Objective. The processing of .nc-files is time expensive because the data has to be converted for the computations. Therefore, especially during testing, the pre-computation to the target data format (.npy) saves time.

Script files.

1. SSTPreprocessMURMain.py
2. imported FileUtils.py
3. imported Magic.py
4. imported Utils.py

Method. The SSTPreprocessMURMain.py script converts all files in the directory “[Magic.dataRootPath]/originalData/MUR/” to .npy-files in the directory “[Magic.dataRootPath]/results/MUR/”. The .npy-files have the file name format binMURSSTYYYY-MM-DD.

The script creates a number of processes, controlled by Magic.processes, which concurrently read, process and write the files. The bottleneck of processing is usually the data transfer rate of the medium that provides the files.

Results. The .npy-files are larger in size as the original .nc-files and written to the “[Magic.dataRootPath]/results/MUR/” directory. The resulting bottleneck is for later computations the transfer rate of the storage device to the computer.

A.7. SST processing

Objective. The script processes the MUR-SST (JPL MUR MEaSURES Project, 2015) data to compute the average SST (e.g., Figure 5.4a, page 84), the pp_{prob} (e.g., Figure 5.4b, page 84) and the days of suitable temperature for paralarvae development (e.g., Figure 5.10, page 91).

Script files.

1. EvaluationSSTMain.py
2. imported Animation.py
3. imported Compute.py
4. imported FileUtils.py
5. imported Magic.py
6. imported SSTReadData.py
7. imported Utils.py

Method. The EvaluationSSTMain.py reads the MUR-SST data from the “[Magic.dataRootPath]/originalData/MUR/” directory and computes for different periods the average SST, the pp_{prob} and the ratio of suitable days for paralarvae development.

Results. The scripts evaluates different time periods and saves the resulting files in the “[Magic.dataRootPath]/results/MUR/” directory. The PDF-files contain the graphical evaluation while the .npy-files contain the computed information. These .npy-files will be used for later computation to avoid the re-computation of the data, see Appendix A.8.

A.8. Plotting the paralarvae sample counts

Objective. Based on analyzed samples, Staaf et al. (2013) postulates a relation of SST and pp_{prob} . These sample counts will be merged with MUR-SST data to demonstrate the postulated SST to sample count relation.

A. PYTHON scripts to compute SST-data and paralarvae distribution

Script files.

1. SpawningGroundsMain.py
2. imported Compute.py
3. imported DrawDistributionRange.py
4. imported Magic.py
5. imported Utils.py

Method. D. Staaf provided the dataset used in Staaf et al. (2013) to project possible spawning grounds of the *Sthenoteuthis oualaniensis* and *Dosidicus gigas* complex (SD-complex) (Section 5.3). The script reads the provided CSV-files and plots the locations, the color-encoded counts, their location and the adjacent MUR-SST on a map. The previously computed MUR-SST results are used to avoid the time-consuming analysis of the MUR-SST files.

Results. Various PDF-files are written to the “[Magic.dataRootPath]/results/Staaf/” directory and show the occurrences, sample counts and sample counts merged with SST data.

A.9. Computation and visualization of the egg and paralarvae dispersion

Objective. Determination of possible places of origin for paralarvae. If paralarvae have been found, the possible spawning grounds can be computed by tracing the material flow by the sea surface currents.

Script files.

1. SpawningGroundsMain.py
2. imported DispersionAnalyze.py
3. imported DispersionAnalyzeResults.py
4. imported DispersionData.py
5. imported DispersionDisplay.py
6. imported DispersionTest.py

A.9. Computation and visualization of the egg and paralarvae dispersion

7. imported DispersionUtils.py
8. imported Animation.py
9. imported Compute.py
10. imported FileUtils.py
11. imported Magic.py
12. imported Utils.py

Method. The DispersionMain.py-script analyzes the currents of the sea-surface and computes the possible dispersion of floating eggs and paralarvae.

If required, the script reads the OSCAR data and performs some consistency checks.

Results. The consistency checks show some deviations regarding the mask, so the data may not be consistently available for all cells over in the same time period. The deviations are mainly near the poles and the values mainly to the east are impacted; these regions are of lesser interest for the actual computations.

Possible places of origin were computed during evaluation and demonstrate that the principle of operation works, but the results are inclusive, see Section 5.6.

B. Python scripts for cannibalism data evaluation

B.1. Overview

The execution of the cannibalism strategy simulations generates .tsv-files parted into three sections. Two sections contain additional information for categorizing the data and one section contains the data itself. The file is ASCII-encoded and the entries are tab-delimited.

Each sections starts at the beginning of the line with an '@' plus the section name, i.e., @parameter, @dataColumns and @data.

B.1.1. The @parameter-section

The first @parameter section contains the parameter set information. In each line, each parameter is documented by a tab-delimited triple of the form type, name and value.

The parameter set is needed to interpret the results in the @data section.

B.1.2. The @dataColumns-section

The @dataColumns sections defines the names of the data columns to identify the entries. Each column name entry is separated by a tab character.

B.1.3. The @data-section

The @data contains the data for the columns, where the columns are separated by the tab character and each line marks a new entry.

B.2. The scripts

B.2.1. Utility scripts

The directory contains the utility scripts TSVCollector.py and TSVSegegmentReader.py. The TSVSegegmentReader.py reads a .tsv file and extracts the data into internal data structures. The script returns an object with the structured data of the .tsv-file.

The TSVCollector.py script defines the classes Results and ParameterList. The Results-object contains all segments of the file path after processing, i.e., all results. The

B. PYTHON scripts for cannibalism data evaluation

ParameterList contains after parsing all parameter sets and the corresponding file names of the path.

B.2.2. Preprocessing the files

The script MainPreProcess.py preprocesses the files for evaluation by removing the sections @parameter, @dataColumns and the header of section @data.

The files are renamed according to the represented cannibalism strategy, where “a” denotes the active cannibalism strategy, “b” (both) denotes the active+passive cannibalism strategy, “n” denotes the no-cannibalism strategy and “p” denotes the passive cannibalism strategy.

The MainPreProcess.py uses the TSVSegmentReader.py script to extract the @data segment data.

B.2.3. Mathematica processing

The script MainMathematica.py filters and rearranges the data and processes them. This rearrangement is performed over the different result files to evaluate for example the changes caused by parameter set changes.

C. Mathematica-Scripts

C.1. GrowthFunctionsAndRelations.nb

C.1.1. Overview

The MATHEMATICA-script `GrowthFunctionsAndRelations.nb` is used for statical analysis of individual traits, the properties at population level, and the display of the results. The script implements different growth functions (see Section 4.3 and Section 6.2) and provides volume, mass and surface estimations through a geometric model of *D. gigas* (Subsection 4.3.2 and Subsection 4.3.3), and the metabolisms (see Section 4.4).

The script is divided into “chapters”, according to the order of calculations in this thesis.

The execution of the script exports PDF files to a directory structure with a relative path to the location of the script:

```
/[ScriptLocation]/graphics/  
├─ cannibalism  
│ └─ energy  
│   └─ offset  
├─ model  
├─ objects  
├─ traitsHigher  
└─ traitsIndividual
```

C.1.2. Comments/Init

This part of the script contains basic comments and information. The initialization performs certain settings for the MATHEMATICA-environment, especially setting the working directories and graphical output to PDF-files.

C.1.3. Constants

This script-chapter first defines the elementary constants and also some constants for display and printing. In addition to the fundamental meaning of these constants, functions are used to set a consistent output of the graphics to be included in the chapters of this thesis.

C.1.4. Basic functions and definition

This chapter defines the exponential, logistic and dimorphic growth functions and the required constants. Some constants are especially defined for the dimorphic growth functions to represent the characteristics such as the modified sigmoid growth functions and different ML_{terminal} .

These definitions are the building blocks for enhanced computations, so changes in this section propagates through all other computations.

C.1.5. Growth, volume, surface and mass functions

This chapter defines basic functions regarding the growth and adjoined traits like mass and surface, see Subsection 4.3.2. Growth is basically age-related by using the age as the input parameter for the sex-dependent growth functions.

The surface, volume and mass functions build on the geometric model of *D. gigas* and the retrieved ML as input parameter (Subsection 4.3.3).

C.1.6. Modeling the individual: Growth

The modeled growth functions provide the basic results in this chapter of the script.

C.1.7. Modeling the individual: The energy model

The energy model (Section 4.4) defined in this chapter of the script bases on the mass of an individual (Subsection 4.3.3). The script models the dependencies, but does not perform any computations.

C.1.8. Traits on individual level

Computations and evaluations regarding the growth, metabolism, mass gain and surface to mass plots are performed in this chapter. The results are displayed and discussed in Chapter 6 of this thesis.

This script chapter mainly evaluates the dependencies and formulas of the previous script chapters and does not introduce any new models.

C.1.9. Cannibalism

This chapter of the script performs a statical evaluation of cannibalism related traits, such as the sex-specific ML-development. These evaluations give a first insight into

the performance of cannibalism and prepare the basic setup of cannibalism simulation through a discrete event model.

This chapter mainly evaluates the dependencies and formulas of the preceding script chapters and does not introduce any new models.

C.1.10. Properties on school and population level

This chapter of the script contains the evaluations at school and population level, e.g., the effects of decoupled growth of reproductive tissue. Traits at the individual level are computed and scaled to swarm, cohort and population level.

As before, the chapter mainly evaluates the dependencies and formulas of the preceding script chapters and does not introduce any new models.

C.1.11. Objects of Study

The MATHEMATICA script code to display the graphical plot of the landing data.

C.2. CannibalismEvaluation.nb

The MATHEMATICA-script `CannibalismEvaluation.nb` takes the output from the script in Appendix B.2.3, evaluates and visualizes the data for Section 7.6.

D. Discrete event simulation with the Scala programming language

D.1. Overview

The simulation model was implemented in the SCALA programming language and uses the JAVA MASON library (Luke, 2019) to handle the basic discrete-event simulation (DES). The simulation model evaluates various aspects of cannibalism.

The simulation model was first developed specific to the library MASON and its programming model. Then the simulation model abstraction toolkit (SMATK) was developed to abstract from the MASON specifics. SMATK encapsulates the MASON library and provides a generalized and simplified access to DES, see the discussion in Section 9.8.

Furthermore, the SCALA programming language supports some features, e.g., traits, which were not found in more popular programming languages like JAVA at the time of implementation. These features enable a more flexible implementation of the simulation models, e.g., the mix-in of traits during run-time to choose for example either 2D or 3D simulations with the same the source code. Additionally, SCALA's functional approach allows a more concise implementation and error checking at compile time.

The SMATK provides the basic agent types (Subsection 9.4.2) for a simplified and unified implementation and allows easy scripting of the variables to run the simulations; it thus forms a unified approach to DES.

D.2. The simulation model abstraction toolkit

D.2.1. General structure of a simulation model

The SMATK was first developed in the programming language JAVA and then again in SCALA, resulting in much more compact source code due to the advanced features of SCALA. In the SMATK, DES simulation models are defined within a defined structure that strictly distinguishes between the simulation model and its visualization. The visualization may encapsulate the simulation model, but the simulation model may be run as a stand-alone application in the terminal. The SMATK defines several base classes to ease the development of DES.

D. Discrete event simulation with the SCALA programming language

Universe. The (abstract) class `Universe` establishes the basic mechanisms of the simulated world by providing a spatio-temporal structure and handling of events. The agents are added to the `Universe` and are provided with these mechanisms so the agents can interact.

The `Universe` extends the original class `SimState` of MASON, which is responsible for discrete event handling, by additional functionality, e.g., the adding of agents and an extended discrete event scheduler for evaluation purposes.

The agent types `Resource` and `Singleton` are directly handled by the class `Universe`, instead the `Agent` is handled by the additional class `Space` that provides an abstract space in which the agents of type `Agent` may interact.

Space. The class `Space` provides the agents a structured and bounded space and provides some utility methods to find other agents within the same `Space` object.

Agents of type `Agent` may interact in the grouping class `Space` but agents of different `Space` objects, even if these `Space` objects overlap, cannot access one another. However, all agents have access to the same `Resource` and `Singleton` agents of the same `Universe` object. `Agent` typed objects may be grouped inside a class `Space` object to optimize the processing since some `Agent` types will never interact.

UniverseGUI. The `UniverseGUI` provides a basic visualization and control scheme by embedding a `Universe` object into this visualization class. The visualization class `UniverseGUI` is only needed for visualization while the `Universe` performs the actual DES.

The `UniverseGUI` interacts with the underlying MASON library by creating the actual viewport, i.e., the 2D or 3D variant, plus any diagrams derived from the class `Diagram`.

D.2.2. The structure of the simulation model abstraction toolkit

The programming language `SCALA` enabled a relatively compact rudimentary implementation of this abstraction layer on top of the MASON library. The `SMATK` defines four packages with additional classes. At the root level there are the classes `Universe`, `UniverseGUI` and `Parameters`, which defines the parameters for the simulation.

Package `smatk.agents`. The three agents types `Agent`, `Resource` and `Singleton` are defined according to the description in Subsection 9.4.2. These three types derive the basic DES functionality from the basic agent class `BasicObject` defined in the sub-package `smatk.agents.base`. Agents with a location (`Agent` and `Resource`) derive a location from the

abstract class `Location`. To enable an `Agent` object's mobility, the `Location` is extended by `Moving`.

Package `smatk.discrete`. This package defines the spatial space in class `Space` in which `Agent` type agents are placed. The class `Space` maps directly to the MASON functionality of dividing space and fast object mapping, including the important functions for finding agents in the vicinity.

Package `smatk.system`. This package defines direct extensions to the MASON library, i.e., the class `ExtendedSchedule` that contributes a refined schedule handling of events.

Package `smatk.utils`. The package `smatk.utils` defines utility classes: Class `DataLogStream` helps to log data into a file, `Diagram` implements a visualization, `IndexToValue` allows to map indices to precomputed function values, class `Interval` implements an interval class with various functionality for setting and comparing, class `Misc` that defines supporting helper functions, `ParameterLooper` allows the systematic (scripted) variation of parameters and `Sigmoid` contributes a basic logistic function that may be combined with `IndexToValue` to minimize expensive computations.

The `Processors` defines two classes of which the `onCurrent` is invoked before or after each simulation step. These classes are suitable for collecting information about the state of the simulation system and can be used for display or logging purposes.

D.3. The simulation model

The simulation model in the package `apps` uses only some parts of the SMATK because the simulated individuals do not need a location for the cannibalism simulation. The cannibalism simulation model contains the two packages `apps.cannibalism` and `apps.common`.

D.3.1. Package `apps.common`

The package `apps.common` defines the common source code used by any kind of simulation. This includes the basic traits of an individual, e.g., its growth function and metabolism, but also the parameters for the cannibalism simulation model in `UniversalSimulationParameters`.

The common (abstract) class `UniversalDgAgent` makes it possible to dynamically extend the agent during runtime by the appropriate growth and energy functions. The `Biometrics` file defines several basic elements of the biometrics of an agent. During

D. Discrete event simulation with the SCALA programming language

runtime, these abstract methods are replaced by the implemented growth functions, i.e., by `BioMetricsDTSGF` and `BioMetricsUnisex`.

D.3.2. Package `apps.cannibalism`

The actual simulation model is defined in the package `apps.cannibalism`. The class `EDCannSim` extends the abstract class `Universe` to set up the simulation. The setup defines several distinct `Space` objects, populated with `Cannibal` objects, within the `Universe`.

The abstract class `Cannibal` is instantiated in `EDCannSim` by dynamically extending the abstract methods of class `Cannibal` with SCALA-defined traits, which provide the actual methods. With this dynamic extension of an abstract class, the `Cannibal` object provides a unified interface for access, but does not implement specific characteristics such as the growth function of an individual.

The class `CannSimParameters` defines the current simulation parameter set and the file `AgentCounting` provides class `smatk.utils.AfterIteration` and class `smatk.utils.BeforeIteration` to log data during the simulation runs. This data, written as a text file, is processed by additional PYTHON and MATHEMATICA scripts.

The actual simulation model defines several `Space` objects with non-interfering populations. The objects in `AgentCounting` collect the data of these distinct populations and compute statistical data, for example the standard deviation.

The class `EDCannSimGUI` defines a visual interface to the simulation to display diagrams (`CannDiagrams`). The class `FoodLevel` was used for experimental purposes of the `EDCannSimGUI` to implement a Singleton object that defines a dynamically changing feeding level.

active	passive	Classification
False	False	no-cannibalism
True	False	active cannibalism
False	True	passive cannibalism
True	True	active+passive cannibalism

Table D.1.: Parameter combinations from controlling the cannibalism strategy. See Table D.2 for the simulation parameters.

Parameter	Type/Unit	Value	Description
number of worlds	Integer	100	Number of parallel simulation runs.
active cannibalism	Boolean	[True, False]	Enables active cannibalism.
passive cannibalism	Boolean	[True, False]	Enables passive cannibalism.
school size	Integer	1000	The initial school size.
sex ratio factor	Double	$[0 \dots \infty]$	The ratio of females to males.
growth function	String	[GF.DTSGF, GF.Unisex]	The growth function for the individuals.
start age	Double/days	$[30, \dots, 180]$	The interval of chosen <i>start age</i> values.
start age variation	Double/days	$[5, \dots, 30]$	The interval of chosen <i>start age variation</i> values.
WOC	[Double, Double]	$[0.2 \dots 0.85]$	The window of cannibalism (WOC).
threshold max. energy	Double	1.1	$threshold_{max}$ (Equation 7.5)
threshold cannibalism	Double	0.9	$threshold_{cann}$ (Equation 7.6)
threshold min. energy	Double	0.7	$threshold_{min}$ (Equation 7.7)

Table D.2.: Simulation model parameters.

D.4. Simulation details

D.4.1. Simulation parameters

The parameters *activeCannibalism* and *passiveCannibalism* control the current cannibalism strategy in the simulation model, see Table D.1.

D.4.2. Implementation details of the simulation model

The cannibalism simulation model is implemented in SCALA and uses the SMATK. An abstraction layer has been added to the toolkit to automate the simulation runs with different parameter sets. In addition, the abstraction layer contributes a set of services, such as the identification of neighbors, which should be common for ABMs, but independent of a particular simulation toolkit.

Algorithm 4 Initialization.

- 1: **function** ONINIT
 - 2: $age \leftarrow start\ age + \frac{x}{2} \cdot start\ age\ variation$
 - 3: $ML \leftarrow getMLByAge(age)$
 - 4: $massEnergyEquivalent \leftarrow getEnergyByML(ML)$
 - 5: **end function**
-

D. Discrete event simulation with the SCALA programming language

Usually 100 schools with the same initializing parameter set are run concurrently during a simulation. The cannibalism simulation model organizes the individuals in separate, non-interacting schools, but without location information for the schools and individuals. A list per school organizes all actions and interactions of these individuals. The granularity of time is set to one hour simulation time.

The cannibalism simulation model logs the state determining parameters of each simulation step, averages these results over all schools and computes the standard deviation of some parameters for a quality criterion. These logged results, such as school size, the female:male ratio or the cannibals ratio, are processed by a PYTHON script to evaluate the performance of the growth function in combination with cannibalism strategies after the simulation runs.

At the individual level, an individual is set with a sex, an age, an initial ML according to the (virtual) age and growth function (Subsection 4.2.2), and an energy level up to the maximum energy available $threshold_{max}$, see Algorithm 4. All individuals are immediately exposed to energy deprivation when the cannibalism simulation model starts.

Each individual runs through a cycle in which the energy level is computed and compared to an energy level expected from the current ML. The ratio of the current energy level and the expected energy level is compared to $threshold_{max}$, $threshold_{cann}$ and $threshold_{min}$ to trigger actions like cannibalism or setting the state like energy exhaustion, see Subsection 7.6.1. The simulation model is reduced to a few factors to allow their evaluation.

The cycle of an individual first computes the ML increase that an individual can attain under optimal conditions until the next simulation step (line 2). The current ML is taken and the corresponding virtual age, which represents the point of time on the growth function, is calculated. Based on this virtual age, the available growth rate, the available next ML (line 3) and the required energy $growth_{mr}$ (line 9) are determined. The energy model first reduces the mass energy equivalent by the prioritized $basal+locomotion_{mr}$ (Subsection 6.3.2) and adds the energy uptake, which is zero due to energy deprivation.

The resulting mass energy equivalent is then compared to the expected energy of the current ML. If the mass energy equivalent is below $threshold_{cann}$, the cannibalism action is triggered. When cannibalism occurs, a cannibal attempts to attain the intended energy uptake up to $maxEnergyUptake$. The actual energy uptake is used to fill up the mass energy equivalent which is then checked for energy exhaustion. If energy

Algorithm 5 Iteration.

```

1: function ONITERATE
2:   virtual age  $\leftarrow$  getAgeByML(ML)
3:   nextML  $\leftarrow$  getMLByAge(virtual age +  $\Delta t_{\text{next time step}}$ )
4:   energyExpected  $\leftarrow$  getEnergyByML(ML)
5:   actualMass  $\leftarrow$  getRawMassByEnergy(massEnergyEquivalent)
6:   energyMax  $\leftarrow$  1.1  $\cdot$  getEnergyByML(nextML)
7:   energyMaxUptake  $\leftarrow$  energyMax - massEnergyEquivalent
8:   basal+locomotionmr  $\leftarrow$  4  $\cdot$  getBasalMetabolism(actualMass)  $\cdot$   $\Delta t_{\text{next time step}}$ 
9:   growthmr  $\leftarrow$  4  $\cdot$  getBasalMetabolism(actualMass)  $\cdot$   $\Delta t_{\text{next time step}}$ 
10:  energyMax  $\leftarrow$  1.1  $\cdot$  getEnergyByML(nextML)
11:  massEnergyEquivalent  $\leftarrow$  massEnergyEquivalent - basal+locomotionmr
12:  massEnergyEquivalent  $\leftarrow$  massEnergyEquivalent + energy uptake
13:  if massEnergyEquivalent <  $threshold_{cann}$  then
14:    perform cannibalism
15:  end if
16:  if massEnergyEquivalent <  $threshold_{min}$  then
17:    go into cadaver state
18:  else
19:    if massEnergyEquivalent > energyExpected then
20:      ML  $\leftarrow$  max(ML, min(nextML, getMLByEnergy(massEnergyEquivalent)))
21:    end if
22:    energyExpected  $\leftarrow$  getEnergyByML(ML)
23:  end if
24: end function

```

exhaustion has not occurred, growth will progress to the next available ML or, in case of insufficient energy uptake, less.

The model bases on an energy dependent growth rate, where growth is only possible up to the point where energy is available for growth. An enforced growth, where an individual growth by the maximum growth rate each time step and gain the maximum ML_{terminal} by the end of lifespan, would not allow to establish different SAM groups in terms of energy aspects.

E. The simulation of ecological models: Enhanced approaches and techniques

E.1. Execution of finite state machines

In SBM, the invocation of phases of a state (Subsection 9.4.4) requires an “external execution engine”. This engine would be based on a simple execution scheme like the one in Algorithm 6. This execution scheme defines two variables containing the current state and the next state. The invocation of `currentState.nextState()` determines the next state of an object and the execution engine algorithm performs the state changes until it obtains an accepting state $currentState \in F$.

Algorithm 6 Execution of a FSM.

```
1: currentState ← ∅
2: nextState ← q0
3: while currentState ∉ F do
4:   if currentState ≠ nextState then
5:     currentState.onExit()
6:     nextState.onEnter()
7:     currentState ← nextState
8:   end if
9:   currentState.onIterate()
10:  nextState ← currentState.nextState()
11: end while
12: currentState.onExit()
```

If $currentState \neq nextState$, then the state change first invokes the methods `currentState.onExit` and `nextState.onEnter`, and then sets $currentState$ to $nextState$. Invoking the `currentState.onIterate` calls the code of the current state, and then $currentState$ is set to the next state. When an accepting state is obtained, the loop is exited by invoking the `currentState.onExit` (of the accepting state) as the (individual’s) last action.

Organizing program code into these phases helps to structure the code, since code segments, that are applied to specific states and phases, are executed exclusively in distinct states. If the state variables are encapsulated, then each state acts independently of other states. This is a less error-prone execution scheme, similar to object oriented

programming (OOP), or more strict in the sense of functional programming (FP), where data and code are encapsulated to prevent external access and its potential side effects. A state encapsulates its behavior and reduces dependencies on other states, since the state-dependent operations are placed within the state. States do not require prior knowledge of the other states, except for the next states available to them.

When a state performs its assigned task, it may need to alter the properties of an object (individual). These properties are defined outside the states and accessible among all states. This weakens the independency between the states; however, since only one state is active at a time during execution, side-effects are not expected since the active state exclusively modifies the properties of an object.

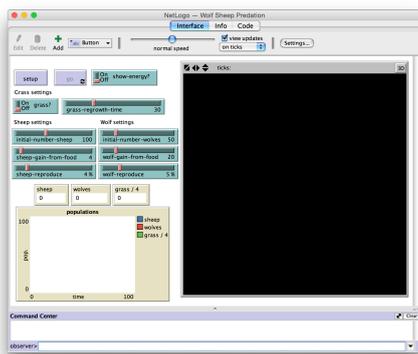
E.2. Simulation environments

E.2.1. NetLogo

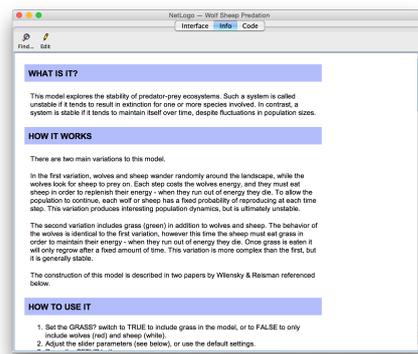
The simulation environment NETLOGO is the successor and further development of STARLOGO. The NETLOGO GUI (graphical user interface) consists of the panes (tabs) INTERFACE, INFO and CODE. The INTERFACE pane contains the (visual) output of the simulation model and the controls. Controls and displays are addable and editable. The INFO pane displays the model description through a contributed template with the topics:

- 1) “WHAT IS IT?: (a general understanding of what the model is trying to show or explain)”,
- 2) “HOW IT WORKS: (what rules the agents use to create the overall behavior of the model)”,
- 3) “HOW TO USE IT: (how to use the model, including a description of each of the items in the Interface tab)”,
- 4) “THINGS TO NOTICE: (suggested things for the user to notice while running the model)”,
- 5) “THINGS TO TRY: (suggested things for the user to try to do (move sliders, switches, etc.) with the model)”,
- 6) “EXTENDING THE MODEL: (suggested things to add or change in the Code tab to make the model more complicated, detailed, accurate, etc.)”,

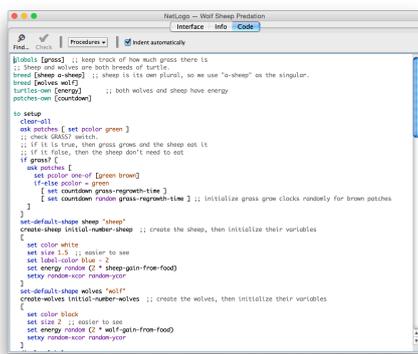
E.2. Simulation environments



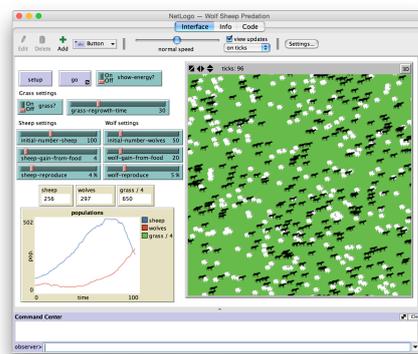
(a) Selected NETLOGO's INTERFACE.



(b) Selected NETLOGO's INFO.



(c) Selected NETLOGO's CODE.



(d) A NETLOGO exemplary run.

Figure E.1.: The interfaces of NETLOGO to access the model and an exemplary run of a contributed model.

- 7) “NETLOGO FEATURES: (interesting or unusual features of NetLogo that the model uses, particularly in the Code tab; or where workarounds were needed for missing features)”
- 8) “RELATED MODELS: (models in the NetLogo Models Library and elsewhere which are of related interest)”
- 9) “CREDITS AND REFERENCES: (a reference to the model's URL on the web if it has one, as well as any other necessary credits, citations, and links)”

The code pane contains the source code of the simulation model. The user interface elements are referenceable and accessible by the source code and allow the configuration of the model and visualizing the results of simulation runs.

E. The simulation of ecological models: Enhanced approaches and techniques

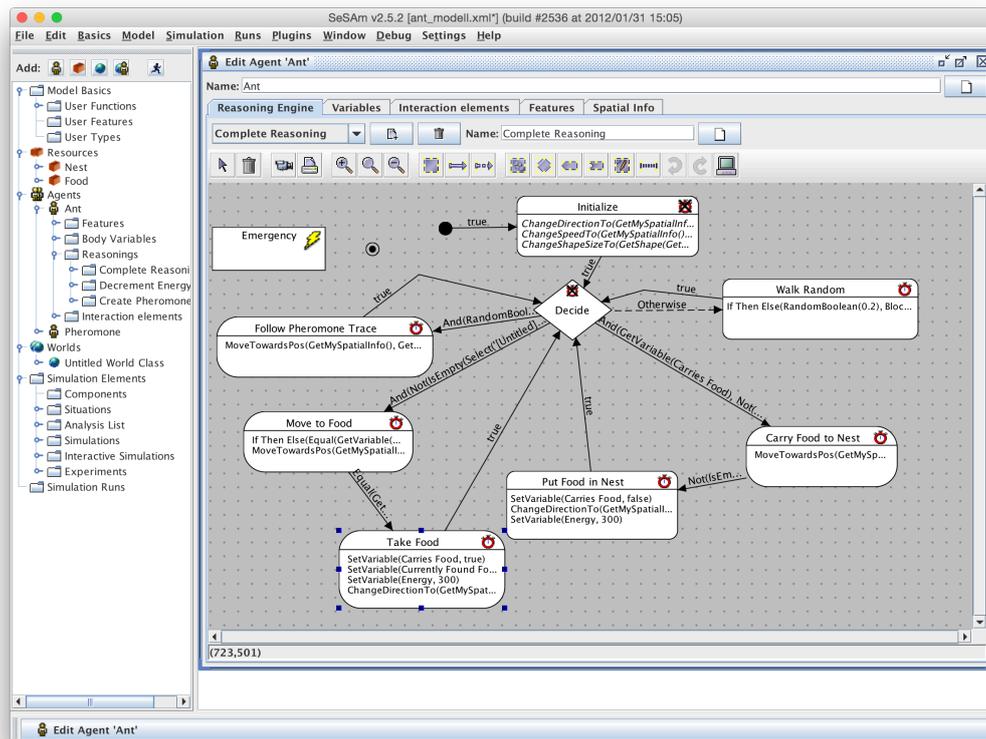


Figure E.2.: Full screen of a SeSAM-model. The left window contains the tree-like structure with the simulation model, setups and results. The center window is the reasoning window for a selected agent type.

The code area contains the source code of the simulation model. The elements of the user interface can be referenced and are accessible via the source code and allow the configuration of the model during the simulation runs.

E.2.2. SeSAM

In SeSAM, a treelike structure, see left panel in Figure E.2, contains the simulation model, setups and results. When selecting an item of the tree, an editor window opens for the properties of the object. The *Model Basics* of the simulation, including the User functions, User Features and User Types, are global definitions of variables and functions, that are accessible to agents of the simulation model. The entry *Resources* defines the read-only resource entities, e.g., a parameter like the global temperature that is available to all other agent types. The entry *Agents* contains localized mobile agent

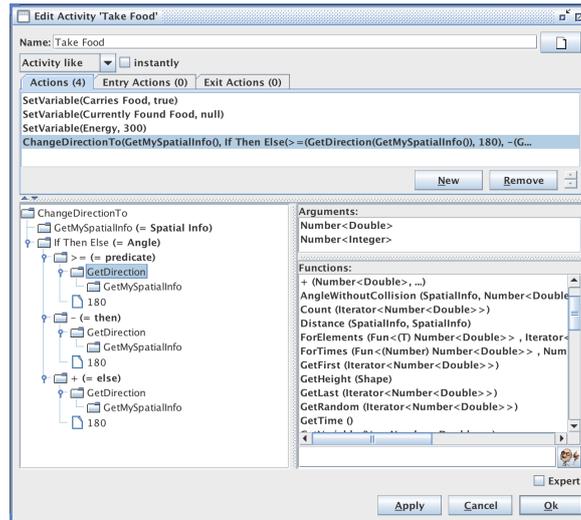


Figure E.3.: SeSAM activity (action) for taking food by a modeled ant.

entities that move through the universe and may interact with all other agents. The entry *Worlds* contains world localized immobile agent entities that span a discrete space in which each cell contains the same set of variables but cell specific values. The entry *Simulation Elements* specifies the setup for a simulation run, while the entry *Simulation Runs* contains the results of a simulation run. These elements in SeSAM meet the requirements for later analysis and post-processing of generated simulation data, see Subsection 9.5.4.

All agent types contain the groups *Features*, *Body Variables* and *Interaction Elements*. The agent types *agent* and *world* include the group *Reasoning* and *agent* also includes *Interaction Elements*:

- *Features* add agent type specific functions and properties through a feature item. The basic installation includes a set of predefined features.
- *Body Variables* define the agent properties (and their types); similar to an object field definition in an OOP language.
- *Reasoning* (excluded for resource) defines periodically executed actions. Multiple entries may be defined and are executed. The reasoning determines the next state of an agent.
- *Interaction Elements* define the interaction between *agent* and the other agent types. The interaction with *world* allows the modification of the world's cells.

The shell for simulated agent system (SeSAM) agent definition is a typical object oriented (OO)-paradigm based class definition that relies on a scheduler to perform certain tasks by invoking the agent's reasoning (methods). The agent's reasoning (Figure E.2) is organized by a node diagram and linking transitions, which specify a SBM (Subsection 9.4.4). Exactly one state is active at a time, and the next active state is selected randomly among the outgoing transitions whose condition is true. The node's *Entry Actions* and *Exit Actions* are performed on entering and leaving a state. This execution scheme complies to SBM and its execution.

The SeSAM programming model meets the ecological modeling requirements by providing functionality to define agent-based-modeling (ABM), including setups and analyses, see Subsection 9.5.4. The simulation is structured in an object-centric view and accessible through a treelike structure. Object oriented programming (OOP) is supported but lacks an inheritance model. This qualifies SeSAM only for small models with few different agent types. The state-based programming feature is a fundamental extension to the standard OOP by superimposing a beneficial additional structure on the model, see the discussion in Subsection 9.4.4.

SeSAM is conceptual advanced, but its strict GUI-centered workflow is a short-coming. With an increasing number of agent properties (body variables), the fully dialog-based definition of these properties renders programming to a tedious task. A text-based definition is more convenient, especially when considering the dialog-controlled editing of field name, type and modifiers per property. Programming any mathematical term is of similar difficulty: a tree structure maps the mathematical expression in a functional approach, where each node constitutes a value by evaluating its nodes below. The tree view makes it difficult to read and edit an expression (Figure E.3), so an expression in a text-based programming language would be an alternative worth considering.

E.3. Frameworks in combination with a programming language

E.3.1. Programming languages

Imperative and functional programming languages

The imperative and the functional programming principles define two common types of programming languages. The most common programming languages belong to imperative programming. Its basic paradigm is the control of data flow and command execution. Technically, a significant part of code and its logic ensures (controls) the

availability of data at a given time in execution flow. The design of these programming languages maps the von-Neumann computer hardware architecture.

In contrast, the functional programming (FP) languages concept defines only the logical order of execution of instructions where the actual execution is triggered by the *availability of data*. This data-driven concept focuses on the data and the execution of adjoined functions (program code) on the data, instead of focusing on the execution of program code and the adjoined data as in the control-based concept. Declarative languages like PROLOG (Colmerauer, 1996) may span a different branch but these are used in very special domains only.

OOP is an extension of the imperative and FP language paradigms and does not define an extra type of programming language. Imperative and FP languages are computationally universal (Turing-complete), therefore both types are equally suitable for computing, but not necessarily for the implementation of simulation models. The FP style exhibits some features that help avoid typical pitfalls of imperative programming languages:

- In FP, the functions are supposed to be pure (self-contained and stateless), they do not rely on external states and always return a value. The returned value is only determined by the ingoing parameters.
- The stateless design of a function allows the developer (and compiler) to reason on a function only by referring to the input and output of a function.
- Pure functions can be altered without side-effects to other pure functions.
- FP is implicit concurrent because side-effects-free pure functions allow parallel execution without any additional effort or special code structure.
- Data is immutable and prevents changing information in an unpredicted way. If data needs to be changed, it is copied before changing and a new set of immutable data is created.
- FP tends to formulate source code more concisely by applying functions on data; in imperative programming languages functions are called with data.

The paradigm of immutable of data avoids hard-to-track errors during runtime. Since standard imperative programming may alter objects during function calls (side-effect), all programmers in a project must keep track of possible side-effects and the effects on the entire program flow and data consistency.

A second advantage is the non-discriminative handling of data and code. Executable code in form of code segments or functions can be provided as function parameters (λ -function) or returned as a function. This feature makes it possible to generate superstructures with code that are shared by many functions, but apply the actual code as a function parameter. Indiscrimination allows FP languages to extend their language specification, e.g., new language features may be added by the developer.

Pure functions, such as the concept of immutable data, enable simplified analysis of source code, because the results of function execution are broken down to the input and output of the functions. Instead, in imperative programming languages, a procedure does not return any value, so a procedures must rely—from a functional point of view—on side-effects (external states, altering something outside the procedure), otherwise a procedure call is redundant because no result is returned. Despite these advantages of FP, the different program formulation and techniques may initially lead to an increase of development time. Recent versions of imperative programming languages include some features of FP, for example the λ -expression in JAVA.

E.3.2. The advantages of Scala if compared to Java

The recent developments in program language designs render JAVA outdated, even in the light recent modifications to the JAVA programming language. Some issues still persist, for example:

- Lack of concurrent execution on multicore systems at language level. The support of concurrent execution through a library in combination with an imperative programming language may cause unpredictable errors during concurrent execution and the adjoined alteration of states, e.g., variables or objects.
- Unintentional alteration of states by design: When the program flow is changed during the development process, e.g., different sequence of procedure calls, this may change objects in an unintended way because of the inherent side-effects. This might be avoided by experienced programmers, but ecologists in general may not have the necessary experience and programming skills to take explicit precautions.
- The inconsequent object oriented design due to the co-existence of non-object primitive data types and objects that requires separate code for handling primitive data types and objects.

E.3. Frameworks in combination with a programming language

- Missing default detection mechanisms for unintentional overriding of methods or object properties. Such semantic errors are syntactical correct, but may be revealed only at runtime and extensive testing.

The JAVA programming language requires a lot of responsibility from the developer to write semantic correct program code. Instead, the FP-based design of the object oriented programming language SCALA minimizes these issues. SCALA interacts seamlessly with JAVA and combines the concepts of imperative and FP. The elaborate strict OO-model of SCALA, traits and closures, minimizes redundancies. The available FP style (as the non-encouraged usage of the available imperative) avoids the pitfalls of imperative programming. The object model of SCALA allows the modeling of weak typing as in OBJECTIVE-C, which was preferred to strong typing of JAVA in S. F. Railsback et al. (2006).

Furthermore, the FP design of SCALA allows self-extension of the programming language to contribute basic support for state-based-programming, e.g., as in the AKKA library. The AKKA library does not meet all requirements (Subsection 9.4.4), but the AKKA open source may be helpful to develop a better implementation.

Acknowledgements

I would like to thank all the people who have contributed to a successful completion of this thesis. First and foremost, I would like to thank Hauke Reuter and Broder Breckling for accepting and supporting me in my efforts to develop and explore ecological models and their simulation, despite my educational background in computer science. And for not turning me away at the very first meeting.

Special thanks are due to Hauke, who accepted the challenge of supervising my research, reading and discussing all my “superb” ideas, even if not all of them were accepted or survived the scientific review process. However, the discussions were productive and hopefully broadened my horizons and focused my mind. It was a great experience.

My thanks also go to Matthias Wolff for reading and evaluating this dissertation and Marko Rohlf and Broder Breckling for their time and participation in the examination committee. Many thanks to Andreas Kubicek and Maren Kruse for biological insights into modeling individuals and ecosystems.

Finally, I would also like to thank my family Nicole, Jasper, Ella, and Siri for supporting me in this work.

Ehrenwörtliche Erklärung

Hiermit versichere ich, dass ich die hier vorliegende Arbeit

1. ohne unerlaubte fremde Hilfe angefertigt habe,
2. keine anderen als die von mir im Text angegebenen Quellen und Hilfsmittel verwendet habe,
3. die den Werken wörtlich oder inhaltlich übernommenen Stellen als solche kenntlich gemacht habe.

Ottersberg, 15.September 2021

Jörg Höhne