STAGE DURATION AS AN EMERGENT PROPERTY: INFLUENCE OF MORTALITY AND VARIATION IN ENVIRONMENT ON COPEPOD POPULATION DYNAMICS

by

Sandra Kitan

Submitted in partial fulfillment of the requirements
for the degree of Master of Science

at

Dalhousie University
Halifax, Nova Scotia
March 2018

Table of Contents

List of Tables ..... **v**

List of Figures ..... **vi**

Abstract ..... **viii**

List of Abbreviations and Symbols Used ..... **ix**

Acknowledgements ..... **xii**

Chapter 1: Introduction **1**

1.1 Motivation **1**

1.1.1 Thesis Objective 2

1.2 Study Species and Area **3**

1.2.1 *Calanus finmarchicus* 4

1.2.2 Northwest Atlantic: Newfoundland and Labrador 5

1.3 Stage Durations **7**

1.3.1 Defining and Determining a Stage Duration 7

1.3.2 Empirical Relationships for *C. finmarchicus* Stage Durations 8

1.3.3 Inherent Individual Variability and Stage Duration Distributions 9

1.3.4 Issues with Estimating Stage Durations using Empirical Relationships 10

1.3.5 Survivorship 11

1.4 Mortality **12**

1.4.1 Importance of Mortality 12

1.4.2 Mortality Rate Estimation using the Ratio Method 13

1.4.3 Violations of the Ratio Method 14

1.5 Simulating Population Dynamics **15**

1.5.1 Individual-Based Models (IBM) 15

1.5.2 Physical Ocean Models 16

1.6 Organization of Thesis **17**

Chapter 2: Mortality Estimation **19**

2.1 General **19**

2.2 Methods **20**

2.2.1 Data 20

2.2.2 Stage Duration Estimates 22
2.2.3 Mortality Rate Estimation using the Ratio Method 22
2.2.4 Calculating Average Mortality Rates and Survivorship 23
2.3 Results 27
2.3.1 Characterizing Regional Ecology 27
2.3.2 Methodological Analysis of Mortality Rate Estimation 30
2.3.2.1 Estimating Mortality with Original versus Restricted Data Subsets 30
2.3.2.2 Method Violations Detected with Negative Mortality Rate Estimates.. 34
2.3.2.3 Comparing DPA Mortality Rates and Survivorship Estimates..... 35
2.3.3 Seasonal Mortality Rates among Regions 36
2.4 Discussion 38
2.4.1 Characterizing Regional Ecology and Stage Duration Estimates..... 38
2.4.2 Methodological Analysis of Mortality Rate Estimation 40
2.4.2.1 Recommended Use of Restricted Subsets..... 40
2.4.2.2 Implications of Differences Among m_{DPA}..... 40
2.4.2.3 Comparison of m_{DPA} to Measures of Variability of MM and NLR 41
2.4.3 Seasonal Mortality Rates among Regions 42
Chapter 3: Development Rates..... 45
3.1 General..... 45
3.2 Methods..... 46
3.2.1 Existing Methods Re-derived for the Use of Development Rates..... 46
3.2.2 Modelling Variability in Development Rates 50
3.2.3 Single Stage Development Tests 51
3.3 Results 52
3.3.1 Test #1 – Temperature Effect..... 52
3.3.2 Test #2 – Individual Variability..... 53
3.3.3 Test #3 – Mortality..... 54
3.4 Discussion 55
Chapter 4: Simulating Spatial and Temporal Population Dynamics 57
4.1 General..... 57
4.2 Methods..... 58

4.2.1 Individual-Based Model (IBM) 58
4.2.1.1 Design Criteria and Desirable Characteristics for Model *58*
4.2.1.2 Characterization of Existing IBM *58*
4.2.1.3 Description of IBM Modifications and Enhancements *61*
4.2.2 Physical and Environmental Data Products 67
4.2.2.1 Particle Tracking Information *67*
4.2.2.2 Chlorophyll-a Data Retrieval *67*
4.2.3 Simulation Conditions for Population Dynamics Case Study 68
4.3 Results 71
4.3.1 Data Characterization 71
4.3.1.1 Particle Tracks and Temperature *71*
4.3.1.2 Chlorophyll-a *72*
4.3.2 Case Study: Simulating *C. finmarchicus* Population Dynamics 72
4.4 Discussion 76
Chapter 5: Conclusions 78
References 81
Appendix A 86
Appendix B 87
Appendix C 89
Appendix D 121
Appendix E 124

List of Tables

Table 2.1 (a) Sampling information: cruise identifiers with dates (2006) and number of stations sampled per region; station no. = original sample size per stage. (b) Restricted Data Subsets: sample sizes, gray shading shows subsets that required restrictions, with darker shading highlighting subsets with > 50% reduction in original sample size. 21

Table 2.2 Procedures of data processing approaches (DPAs) for determining "average" mortality rates. 26

Table 2.3 Regional average stage durations (days) for non-adult copepodites (C1-C5). Range (maximum-minimum) of station-specific stage durations are shown with shading: white (< 1 day), light gray (1-3 days), medium gray (3-7 days), and dark gray (> 7 days). 28

Table 3.1 Single stage development tests with various temperature scenarios (°C): T_{low} = 2, T_{avg} = 7, T_{high} = 12, and T_{dyn} = t/10+2. Components of simulations not included within a test are marked as "No". 51

Table 4.1 Major changes made to the base IBM and new additional components to create the modified IBM version for Chapter 4 objectives. Some examples ("e.g.,") are provided in MATLAB code. 62

List of Figures

Figure 1.1 *Calanus finmarchicus* copepod; photo credit Head (BIO, n.d.) 3

Figure 1.2 Copepod life cycle developing through stages: egg (E), naupliar (N1-N6), and copepodite (C1-C5), with adult being the last copepodite stage (C6) (adaptation, Kvile, 2015: sourced from NERC ZIMNES project) 4

Figure 1.3 Northwest Atlantic region focused on the Newfoundland and Labrador ocean region (Canada). 7

Figure 2.1 Locator map for cruise sampling stations for four regions off of Newfoundland and Labrador: NON (dark circles), NOFF (light circles), SON (dark triangles), and SOFF (light triangles) 20

Figure 2.2 Temperature (x - markers) and chlorophyll-a concentration (o - markers); cruise means (symbols) and minimum/maximum points (error bars) are displayed at mid-point date of cruise durations for all 4 regions. 27

Figure 2.3 Sample means of *Calanus finmarchicus* copepodite stage abundance observed in 2006 for 4 regions of study. The height of each stacked bar, plotted at the mid-point day of each cruise duration, is the total of all stage sample means. 30

Figure 2.4 Average mortality rates for North subsets estimated using various DPAs: MM (blue circles), MA (green left-pointing triangle), MSR (black 'x'), NLR (red right-pointing triangles). Measures of variability provided for MM, station-specific estimates (smaller blue circles), and NLR, 95% C.I. (red bars). The subset sample size (n) is provided at the top of each window. Maximum difference in DPA survivorship (%) estimated for positive C1-C5 m_{DPA} are provided beneath n-samples; cases with no positive m_{DPA} are labelled "N/A". Columns divide regions, while rows represent cruises. 32

Figure 2.5 Average mortality rates for South subsets estimated using various DPAs: MM (blue circles), MA (green left-pointing triangle), MSR (black 'x'), NLR (red right-pointing triangles). Measures of variability provided for MM, station-specific estimates (smaller blue circles), and NLR, 95% C.I. (red bars). The subset sample size (n) is provided at the top of each window. Maximum difference in DPA survivorship (%) estimated for positive C1-C5 m_{DPA} are provided beneath n-samples; cases with no positive m_{DPA} are labelled "N/A". Columns divide, while rows represent cruises. 33

Figure 2.6 Stage-pair mortality rates: averages of m_{DPA}, where all m_{DPA} were negative are denoted by "-", while those with mixed signs (negatives removed from the final average) are marked by "*". 38

Figure 3.1 The relationship between molt-cycle-fraction (MCF, n.d.) and development rate (dev, units of time) of stage i. **(a)** Examples of various

MCF relationships attaining MCF = 1 at the same time, **(b)** Similar examples as (a) but attaining MCF = 1 at various times, and **(c)** corresponding development rates to MCF curves in (b), i.e., derivatives. 47

Figure 3.2 Numerical approximations for MCF: **(a)** integration of development rate and, **(b)** "predicting" MCF using Euler method. 48

Figure 3.3. Example of molt-cycle-fraction estimate exceeding 1 within a time-step 49

Figure 3.4 Normalized stage duration results from Test #1, normalized by average stage duration result from T_{high} scenario. 52

Figure 3.5 Distributions of resultant stage duration results from Test #2, normalized by average stage duration result from T_{high} scenario. Statistics of normalized results provided for: emergent stage durations – dotted black vertical lines, median and variance. Results for each temperature simulation are shaded, while others are provided as outlines for relative comparison: (a) T_{low}, (b) T_{avg}, (c) T_{high}, and (d) T_{dyn}. 54

Figure 3.6 Distributions of resultant stage duration results from Test #3, normalized by average stage duration result from T_{high} scenario. Statistics of normalized results provided for: emergent stage durations – dotted black vertical lines, median and variance. Results for each temperature simulation are shaded, while others are provided as outlines for relative comparison: (a) T_{low} without mortality, results from Fig. 3.6a, (b) T_{low} with mortality, S = 68%, (c) T_{high} without mortality, results from Fig. 3.6b, and (d) T_{high} with mortality, S = 90%. 55

Figure 4.1 Conceptual flow chart of base IBM demonstrating (a) general structure of main code components, and (b) structure of main matrices: population, rates (individual and mean, i.e., "bar"), and environmental conditions (temperature and food). Note: n = total number of individuals.59

Figure 4.2 Conceptual flow chart of the modified IBM demonstrating: (a) general main code structure, and (b) main population matrix structure. 66

Figure 4.3 Simulated particle tracks (Day 0 = April 1st run for 180 days) derived from Han et al.'s (2008) 3D physical model. Only particles that remain mobile are displayed (554 particles). Deeper water depth is marked with darker gray scale contour. 68

Figure 4.4 Data of particle tracks A (yellow circles,) and B (orange triangles) over 180 days, Day 0 = April 1st. Temperature and chlorophyll-*a* data are displayed spatially (a, c) and temporally (b, d). 70

Figure 4.5 Transport of particle tracks A and B (1st column) and associated simulated stage abundances over 180 days. Abundances (# m^{-2}) are plotted on a logarithmic scale. 74

Figure 4.6 Transport of particle tracks A and B (1st column) and associated simulated stage proportions over 180 days. Stage proportions of population sizes are plotted on a logarithmic scale. 75

Abstract

A key species in the Northwest Atlantic food web is *Calanus finmarchicus*, which provides an important food source for many marine animals. To better understand its ecosystem roles, obtaining detailed knowledge of its life processes, such as development and mortality, and the factors controlling them is vital. Since pelagic systems are dynamic, variable, and advective, mortality cannot be measured *in situ*, and development is observed in constant laboratory conditions. Here, a series of modelling studies are conducted to investigate mortality and development separately, by applying and examining common methodologies and providing new approaches and recommendations, and explore the interaction of processes, by building an individual-based model to simulate the life history of *C. finmarchicus*. This model is presented with enhanced capabilities for coupling physical ocean data and tested with a case study that investigates the influence of environmental variability and mortality on simulated population dynamics in space and time.

List of Abbreviations and Symbols Used

3D	Three-dimensional
AZMP	Atlantic Zone Monitoring Program
C.I.	Confidence intervals
C.V.	Coefficient of variation
CDF	Cumulative distribution function
chl-*a*	Chlorophyll-*a*
C	Carbon
E	Egg stage of copepods
N1-N6	Naupliar stages of copepods
C1-C6	Copepodite stages of copepods
DPA	Data processing approach
EP	Egg production
m	Mortality rate in context
$m_{(DPA)}$	Average mortality rate results estimated by a specific DPA
MA	Mean abundance approach
MCF	Molt-cycle-fraction
m_{DPA}	Average mortality rate results estimated by all DPAs
MDT	Median development time
MM	Mean mortality approach
MSR	Mean stage ratio approach
NaN	Not a number (computing)
NLR	Nonlinear regression approach
NOFF	North off-shelf region
NON	North on-shelf region
PC_i	Percentage of group developed to stage *i*
PT	Particle track
S	Survivorship
SOFF	South off-shelf region
SON	South on-shelf region

θ	scale parameter for gamma distribution
σ^2_D	variance of stage durations to describe shape of CDF
σ^2_{dev}	variance of development rates to describe shape of CDF
μ_D	mean stage durations to describe shape of CDF
μ_{dev}	mean development rates to describe shape of CDF
∞	infinity
A_i	abundance of stage *i*
a_i	stage-dependent parameters in stage duration estimation
C1-C6	subscript for *i* denoting six copepodite stages
C6	subscript for *i* denoting adult stage
C6F	subscript for *i* denoting sex-specific adult stage, females
C6M	subscript for *i* denoting sex-specific adult stage, males
dev	development rate
D	resultant stage duration, i.e., of each individual
$\overline{D}$	emergent stage duration, i.e., arithmetic average of resultant stage durations
D_i	stage duration of stage *i*
$\overline{D}_i$	arithmetic average of *D* estimates for stage *i*
$D_{i\text{-}n}$	aggregate stage duration of stage *i* through stage *n*, reaching stage *n+1*
dt	continuous change in time
E	subscript for *i* denoting egg stage
F	food conditions
h	general form of independent variable increment, e.g., time-step
i	subscript identifying current stage
i+1	subscript identifying adjacent stage to *i*
k	shape parameter for gamma distribution
MCF	molt-cycle-fraction
MDT_i	median development time for stage *i*
m	mortality rate
m_i	average mortality rate for stage *i*
$m_{i,i+1}$	average mortality rate for stages *i* and *i*+1
Δm_{DPA}	max. difference among DPA average mortalities; $\max(m_{i,i+1}) - \min(m_{i,i+1})$

$\overline{m_{i,i+1}}$	arithmetic average of average mortality rate results of all DPAs
n	counter or sample size, in context
N1-N6	subscript for i denoting six naupliar stages
S_i	survivorship estimate through single stage i
ΔS	range of survivorship estimates for a given subset; $\max(S_{i,i+1}) - \min(S_{i,i+1})$
ΔS_{DPA}	max. difference among DPA survivorship estimates
$S_{i,i+1}$	survivorship estimate through aggregate stages i and i+1
R_i	recruitment rate into stage i
T	temperature conditions
t_n	time corresponding to discrete data point n
t_{molt}	time at molt out of stage i, into stage i+1
t_0	time at molt of stage i-1, into stage i
Δt	time interval or time-step

Acknowledgements

I would like to express many thanks to Dr. Wendy Gentleman, my research supervisor, for providing me with this opportunity from which I have learned how to apply theory into practice, always think "outside the box", and am particularly thankful for Dr. Gentleman's thoughtful guidance, much appreciated constructive criticism, and continuous support. I am grateful to Dr. Gordon Fenton for his thesis writing assistance, guidance and support, and to my committee members, Dr. Farzaneh Naghibi and Dr. Jinyu Sheng, for providing thoughtful insight for this work.

This work could not have been possible without the data and helpful information (and clarification when needed) from researchers at Fisheries and Oceans Canada (DFO): Pierre Pepin, Guoqi Han, Erica Head, Catherine Johnson, and Catherine Brennan.

This research was financially supported in part by a National Sciences and Engineering Research Council of Canada (NSERC) grant for Dr. Gentleman's lab, and Dalhousie University funding, awarded through graduate scholarships and bursaries.

I wish to extend my gratitude to faculty, staff, and fellow students within the Engineering Mathematics department for always providing a friendly work environment. A special thanks to Claire Chisholm for her immense generosity, thoughtfulness, and advice. I would like to particularly thank Danielle Dempsey, for her assistance with this work, but above all, sharing many laughs and keeping me grounded throughout this experience.

Finally, I would like to thank my family and friends, especially my parents, Dubravka and Selver Kitan, and partner, Andrew Lilly, for their continuous love, encouragement, and patience throughout my entire academic career.

This journey could not have been possible without any of you; from the bottom of my heart, thank you. ♥

Chapter 1: Introduction

1.1 Motivation

The northern portion of the North Atlantic Ocean is the subject of extensive ongoing multidisciplinary research projects, as its large latitudinal expanse characterized by circulation and water mass distribution, in combination with wide ranges of seasonally and geographically varying properties (e.g., temperature, salinity), provide a multitude of ecosystems for many interacting species (Melle *et al.,* 2014). A key component in the food web of the Northwest Atlantic is the species *Calanus finmarchicus*, a marine planktonic copepod, which acts as "secondary" production, providing the essential transfer of energy from primary production (e.g., phytoplankton) to high trophic levels (Pepin and Head, 2009). This species provides a nutritious food source for many ecologically and economically important species, such as larval fish, seabirds, and whales, but also provides a type of "top-down" control on lower trophic groups (i.e., relatively smaller plankton). To better understand and quantify the role of *C. finmarchicus* in the pelagic system, it is vital to obtain detailed knowledge of physiological and biological behaviours and processes, and the physical and biological factors controlling them (Hirche, 1996b). This species is associated with a complex life history, composed of a series of life stages with various characteristics and roles within ecosystems. The developmental timing of these life stages (i.e., stage durations) have been able to be studied in laboratory settings and empirical relationships to temperature and food have been found (e.g., Campbell *et al.,* 2001; Gentleman *et al.,* 2008). However, these relationships were derived from observations of copepods developing in constant conditions, which would not be appropriate to estimate stage durations for dynamic, variable conditions *in situ*. Another important process known to greatly influence the size and stage structure of copepod populations is mortality, which is primarily driven by predation, but may also include other "losses" (e.g., from starvation or senility). Due to the different characteristics of copepod life stages, the mortality rates (e.g., predators) are known to vary by stage, and also in space and time. Combining these variable factors with the continually changing and advective nature of pelagic (i.e., upper ocean) ecosystems makes measuring mortality *in situ* not possible, and even estimation quite

difficult. Various methods have been developed for estimating mortality rates, but because no method can be verified as the "most accurate", there are ongoing efforts to examine the underlying concepts and improve the methodology (e.g., Gentleman *et al.,* 2012). To investigate such questions not possible *in situ*, or those that may require costly (e.g., time, effort, financially) experiments, such as the case for plankton sampling, marine models have been used as powerful tools. Traditionally, ecosystem modelling approaches that include plankton often group them together by trophic level, which disregards any individual variation, and while many have improved to incorporate details, such as life stages, characterization of developmental timing and food intake are often unrealistic (Gentleman *et al.,* 2008; Gentleman and Neuheimer, 2008; Neuheimer *et al.,* 2010a). Individual-based models (IBM) were developed to simulate individuals and their characteristics separately, based on the concept of environmental factors affecting the physiology of each individual (i.e., possessing a measureable individual variability). IBMs have been greatly used in fisheries research but have only recently been gaining momentum for copepods over the last decade or so (Neuheimer *et al.,* 2010a). Therefore, characterization of processes (e.g., development) may be incorrect or important details overlooked, requiring ongoing efforts to build and improve upon existing models to increase confidence in simulations. To better understand the spatial variability of copepod population dynamics, the effect of spatially varying environments on their life history processes must be examined, which may be investigated by coupling IBMs with physical ocean models, or Lagrangian particle tracking. This coupling has been successfully completed for *C. finmarchicus* in various Northwest Atlantic regions, but not generally for the Newfoundland and Labrador ocean waters (e.g., indirectly, Pepin *et al.,* 2013).

1.1.1 Thesis Objective

The objective of this thesis study is to gain insight into the influence of dynamic environmental conditions and mortality on emergent stage durations and the overall temporal and spatial effects these factors have on *C. finmarchicus* population size and structure for Newfoundland and Labrador ocean waters. It is hypothesized that a dynamic environment will cause stage duration estimates to differ from empirical estimates, and that the addition of mortality will result in decreased average stage durations, and

variance, among individuals, due to a "survival of the fittest" effect, i.e., those that survive will have developed through the stages faster. The overall objective of this thesis is completed by a series of analyses: estimating mortality rates using commonly used methods in the literature, testing effects of variability in environment and mortality on individual stage durations using empirical relationships, and lastly, the development of a population dynamics model, representing the life history of *C. finmarchicus*, subject to variability in environment and mortality. Commonly used methodologies, applied from the literature, are also investigated on a quantitative and conceptual basis, providing new approaches and recommendations for use.

1.2 Study Species and Area

Marine plankton are classified as organisms that drift or float in the ocean (i.e., cannot swim against currents), and although most are microscopically small, the sizes greatly vary (e.g., including large jellyfish). "Copepods" are small plankton of the subclass "Copepoda" that are, typically, the dominate members of the zooplankton group, which defines those that feed on other plankton, as opposed to phytoplankton which feed by photosynthesis. The particular organisms used for this study are zooplankton from the genus *Calanus* and, specifically, the species *finmarchicus*, making up its scientific name of *Calanus finmarchicus* (or *C. finmarchicus,* shown in Figure 1.1).

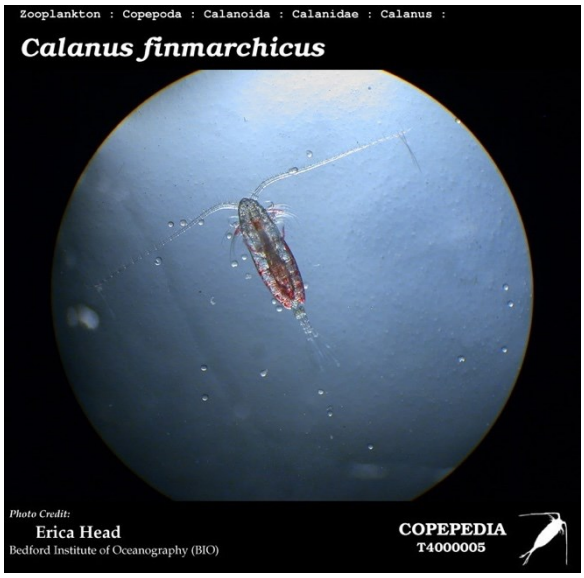


Figure 1.1 *Calanus finmarchicus* copepod; photo credit Head (BIO, n.d.)

1.2.1 *Calanus finmarchicus*

C. finmarchicus is considered a larger sized copepod as individuals can be seen with a naked eye, about 2-4 mm in body length (e.g., Head *et al.,* 2012). This species has a complex life history undergoing various stages of development (Figure 1.2). The first is an egg stage (E), which hatch into nauplii (six stages: N1-N6) and develop into copepodites (six stages: C1-C6). The final copepodite stage (C6) is considered to be the adult stage, where they are sex-specific (separated into females and males), for reproduction to occur and, at this point, the cycle is repeated by females spawning eggs (e.g., Hirche, 1996a). The developmental process from stage-to-stage is called "molting" and occurs at all stages except the final, adult, stage of development.

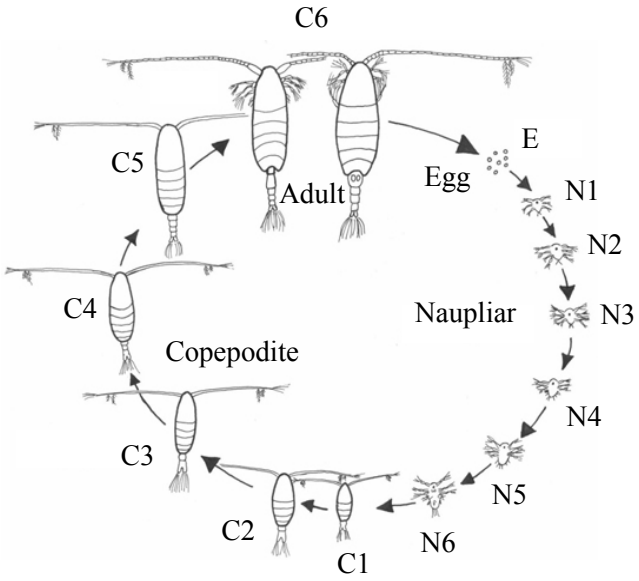


Figure 1.2 Copepod life cycle developing through stages: egg (E), naupliar (N1-N6), and copepodite (C1-C5), with adult being the last copepodite stage (C6) (adaptation, Kvile, 2015: sourced from NERC ZIMNES project)

The *C. finmarchicus* life cycle is typically on an annual timeline of events (Conover, 1988): reproduction and spawning of the "first generation" occurs in the spring time, these offspring develop through the naupliar and copepodite stages over spring and summer, reaching late copepodite stages (i.e., C4, C5) by summer and autumn, where most of the generation descends to deeper waters to enter "diapause" (Head and Pepin, 2008a). *C. finmarchicus* diapause is a type of overwintering strategy (Hirche, 1996b),

similar to bear hibernation, when temperatures and food are low during the late fall and winter months. During the diapause period, *C. finmarchicus* sinks to depths ranging 500 to 2000 m (depending on location; Heath *et al.,* 2004; Head and Pepin, 2008a), where growth is suppressed, metabolism is reduced (Hirche, 1996b), and individuals survive off of energy (lipids) that were stored internally over the previous months (Johnson *et al.,* 2008). Overwintering occurs mainly for individuals in stage C5, but C4 and females have also been found in samples (Hirche, 1996b; Head and Pepin, 2008a). Populations have been known to not enter diapause at the same time; in summer the individuals prepared for overwintering will sink to deeper waters, while those still developing remain in the surface, and potentially may never enter diapause, but continue molting and reproducing (i.e., producing "multiple generations", Hirche, 1996b; Head and Pepin, 2008a; Johnson *et al.,* 2008). The stages in diapause emerge in winter or spring (Conover, 1988), where they migrate back to surface waters to molt to sex-specific adults and reproduce (Hirche, 1996b). As with most marine animals, this species reproduces by "broadcast spawning", such that the fertilization occurs externally, and eggs are spawned freely in the water (Ohman *et al.,* 2002).

The main life processes that govern *C. finmarchicus* population sizes and structures throughout their life history are egg production (EP) and development (as described above), and *mortality*. Copepod mortality is known to be mainly due to predation (e.g., Hirst *et al.,* 2010), and some strategies (e.g., diapause) are believed, in part, to be adaptations to minimize their mortality risk (Ohman *et al.,* 1996). *C. finmarchicus* serve as major prey for many ecologically and economically important species, including larval fish (e.g., groundfish, Head and Pepin, 2008a), sea birds (e.g., Brown and Gaskin, 1988), and whales (e.g., endangered North Atlantic right whale, Michaud and Taggart, 2007), and apply "top-down" control on lower trophic levels by feeding on phytoplankton and microzooplankton.

1.2.2 Northwest Atlantic: Newfoundland and Labrador

C. finmarchicus is one of the most abundant copepod species found in the northern Atlantic Ocean, where it can make up to 70% or more of the zooplankton biomass on the

Northwest Atlantic continental shelf (Head and Pepin, 2008a). The focus area throughout this study will be ocean waters across Newfoundland and Labrador (NL; Figure 1.3). The waters in this region are affected by circulation, dominated by the equatorward flowing Labrador Current, where branches flow along the coast and shelf-edge, significantly influencing the dispersal of zooplankton (Pepin and Helbig, 1997; see patterns of currents in Head and Pepin, 2008a, Fig. 1). This advective system transports *C. finmarchicus* in the surface layers southward over the shelves in the spring and summer, while affecting overwintering stages in autumn and winter by moving southward in the slope waters (Head and Pepin, 2008a). It is expected that these "potentially complex interactions" between physical and biological drivers will result in both latitudinal and cross-shelf differences in *C. finmarchicus* population size and structure (Melle *et al.,* 2014).

In the Newfoundland-Labrador ocean region, the Atlantic Zone Monitoring Program (AZMP) started by Fisheries and Oceans Canada in the late 1990s regularly collects oceanographic data at both fixed stations and cruises (multiple yet various times of the year), providing physical, chemical, and biological samples (Therriault *et al.,* 1998); used throughout this thesis.

Although *C. finmarchicus* is known to be omnivorous, field data of chlorophyll-*a* (chl-*a*) concentrations are assumed to be a proxy of all "food" availability (e.g., Lynch *et al.,* 1998; Gentleman *et al.,* 2012), and further converted into levels of carbon, using C:chl-*a* mass ratio of 50 as typical for this area (Sathyendranath *et al.,* 2009), where necessary as the required unit of data for further analysis (see section 1.3.2).

Timing and causes of diapause for the Newfoundland Shelf region were explored by Johnson *et al.* (2008), determining a mean date of September 24 for onset and February 19 for emergence. Johnson *et al.* (2008) were unable to determine a "single environmental cue" to explain either, onset or emergence, dates. *C. finmarchicus* were found to enter and emerge from diapause earlier for relatively more southern regions than the Newfoundland Shelf (Scotian Shelf June 10 and January 10, respectively; Johnson *et al.,* 2008), which implies that the onset of diapause is later in the Labrador Sea, as compared to the Newfoundland shelf and slope waters.

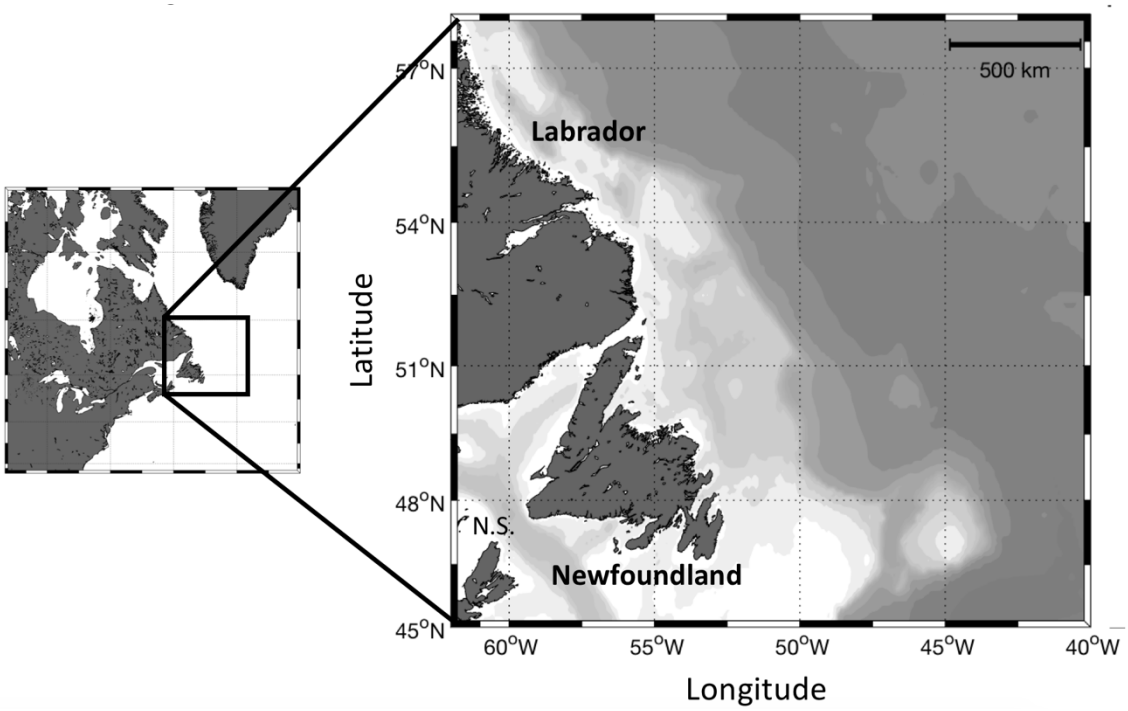


Figure 1.3 Northwest Atlantic region focused on the Newfoundland and Labrador ocean region (Canada).

1.3 Stage Durations

1.3.1 Defining and Determining a Stage Duration

An individual's stage duration is the length of time it takes to mature through a given stage over their life history (Figure 1.2), i.e., the amount of time it takes from molting "into" stage *i* (current stage) to molting "out", into subsequent stage *i*+1. For copepods, a stage duration is considered for all stages that are "successfully completed", which means that the individual did not die in stage *i* but survived and molted to stage *i*+1. Therefore, a stage duration does not exist for adult stages, as individuals do not "molt out" from their final stage of life.

Each stage has specific internal and physical characteristics that an individual must attain in order to molt into; for example, stage differs from human *age,* which progresses regardless of the person's body size or shape. The conditions that an individual experiences, such as temperature and food availability, play an important role in the development process, allowing for these necessary changes to happen. Favorable conditions, such as relatively warm temperature and/or satiated food conditions, may

result in rapid development through a stage, while colder temperatures and/or low food may slow down or even temporarily suspend development (e.g., Campbell *et al.,* 2001; Crain and Miller, 2001).

1.3.2 Empirical Relationships for *C. finmarchicus* Stage Durations

Stage durations (D_i for stage i, in days or d) of *C. finmarchicus* have been studied in laboratories and found to vary by stage, and to be dependent upon temperature and food availability. Many studies have described the temperature dependency of stage durations using the Belehrádek (1935) temperature function, i.e., $D_i = a_i(T - \alpha)^b$, where T is temperature and parameters a_i, α, and b are determined specific to the regions where samples of *C. finmarchicus* were obtained (e.g., Scotian Shelf: Corkett *et al.,* 1986; Georges Bank: Campbell *et al.,* 2001; Disko Bay: Jung-Madsen and Nielsen, 2015). Gentleman *et al.* (2008) expanded the study by Campbell *et al.* (2001) to include a food-dependency factor, i.e., (K/F + 1), and the parameters are as follows:

$$D_i = a_i(T + 9.11)^{-2.05}\left(\frac{17}{F} + 1\right) \tag{1.1}$$

where a_i are stage-dependent determined for stages i = E to C5, T is temperature (°C), and F is carbon-content of food conditions (mgC/m^3, to coincide with K = 17 mgC/m^3). The food factor (second term in parenthesis) is not included when estimating D_E to D_{N2}, as N3 is the first feeding stage. Here, it is noted, while the units of a_i, α, and b parameters are not explicitly stated in the literature, the following units are assumed to obtain D_i in days from T in °C: $[a_i]$ = day· °C$^{2.05}$, $[\alpha]$ = °C, and b is dimensionless. The parameter values of equation (1.1) were derived using *C. finmarchicus* copepods that were collected from Georges Bank (off of Cape Cod, Massachusetts, USA) and observed in a laboratory where various food and temperature conditions were tested (Campbell *et al.,* 2001; Gentleman *et al.,* 2008). There is no comparable study completed for the Newfoundland-Labrador region, so equation (1.1) has been previously applied for analyses involving Newfoundland-Labrador ocean waters (e.g., Neuheimer *et al.,* 2009, 2010a, 2010b; Gentleman *et al.,* 2012), and therefore, has also been chosen for this thesis study.

1.3.3 Inherent Individual Variability and Stage Duration Distributions

The parameters of equation (1.1) were determined from experiments involving tending a large group of *C. finmarchicus* eggs to adulthood and observing development patterns to each stage (Campbell *et al.,* 2001; Gentleman *et al.,* 2008). As described in Gentleman *et al.* (2008), evolution of the population's stage structure was observed by taking samples at frequent time intervals and tracking the "percentage of the group" that has matured to stage *i* (termed PC_i in Gentleman *et al.,* 2008, Fig. 1). From the PC_i time-series data, Campbell *et al.* (2001) determined median development times for each stage *i* (MDT_i) as the length of time from spawn until $PC_i \geq 50$ %. It is from the MDT_i data that observed stage durations were determined, by taking the difference of MDTs for successive stages (i.e., $D_i = MDT_{i+1} - MDT_i$), and parameters for equation (1.1) were determined by fitting relationships of stage durations to temperature and food. For more details regarding these experiments see Campbell *et al.* (2001) and Gentleman *et al.* (2008).

The relationship of PC*i* data *versus* time has been described by using curve fitting (e.g., linear regression; Gentleman *et al.,* 2008) or cumulative probability distributions (CDF; e.g., normal: Hu *et al.,* 2007; Gentleman *et al.,* 2008; gamma: Hu *et al.,* 2007). By obtaining a "fit" of PC_i to time, this relationship may be used in modelling studies to estimate stage duration when considering individual copepods with inherent variability among them and a variety of environmental conditions. For example, if the relationship is described by a normal CDF, the shape of the CDF for specific environmental conditions is defined by an estimate of mean and standard deviation of stage durations (e.g., as observed in Gentleman *et al.,* 2008), where the independent axis represents time. The dependent axis (i.e., PC_i) may then be considered as a measure of inherent individual "fitness" (i.e., level of "performance", valued between 0 and 1), and set such that individuals of higher fitness experience shorter stage durations and lower fitness corresponds to longer stage durations (i.e., for an individual, $PC_{i,ind} = 1 - fitness_{ind}$; for an application example, see Neuheimer *et al.,* 2010a). Individual stage durations are determined by finding the time on the independent axis that corresponds with provided individual fitness levels.

1.3.4 Issues with Estimating Stage Durations using Empirical Relationships

The studies determining relationships for estimating stage durations were completed under controlled laboratory conditions held constant, which would not reflect the dynamic and variable nature of stage durations *in-situ* (Aksnes and Ohman, 1996). A "simple illustration" is provided in Gentleman *et al.* (2008) that effectively describes how considering constant stage durations may be problematic in a dynamic scenario (here, D_i represents the estimated stage duration for an individual copepod entering stage i):

> …consider a situation, where environmental conditions were constant, the corresponding D_i = 10 days and an individual copepod had been developing through stage i for 5 days. Now, suppose conditions change at this 5-day point that correspond to a new D_i = 5 days. Because the copepod had only been in stage i for 50% of its physical development, one would not expect the copepod to molt to stage i+1 suddenly simply because its time spent in stage i now corresponds to the new D_i, i.e., the copepod is not "ready" to molt (adapted from Gentleman *et al.,* 2008, p. 407).

To avoid such scenarios in modelling practices (e.g., this issue may arise in age-within-stage models; see Neuheimer *et al.,* 2009, Lynch *et al.,* 1998), an alternative measure of the development process through a stage, rather than a set stage duration, is introduced as the dimensionless "molt-cycle-fraction" (MCF; Miller and Tande, 1993; see Neuheimer *et al.,* 2010, their Fig. 2). The MCF is a "measure" of proportional development for each individual copepod, equivalent to 0 upon entering stage i and increasing over time until reaching 1 which signifies that the individual is ready to molt to stage i+1 (reset to 0 for the subsequent stage). Due to stage duration observations, the MCF is assumed to also be dependent on environmental conditions (more in Chapter 3) and, therefore, MCF increments through stage i are determined by relating a discrete time-step value to the D_i estimate based on current conditions at "t", i.e., $MCF_{t+1} = MCF_t + \Delta t/D_{t,i}$. Hence, if environmental conditions are held constant from MCF = 0 to 1, the D_i estimate at each t will be the same, and the *time* it takes for the individual to molt will be equal to D_i. However, assigning stage durations becomes complicated in a scenario with dynamic environment, where D_i must be adjusted appropriately at each t to correspond to changing environmental conditions (Gentleman *et al.,* 2008), but yet the *time* it takes for MCF to develop from values of 0 to 1 is the "true" stage duration, not the estimates of D_i during

the development period. This confusion of what actually is the "stage duration" highlights the need to conceptually define a "development rate" to use with MCF throughout the development process, with *resultant* individual stage durations and the average of these individual stage durations considered an emergent property (more in Chapter 3).

Obtaining estimates of stage durations, or more often "aggregate" durations (i.e., "development time" through multiple stages), affected by environmental conditions are important for studying copepod populations dynamics. Often of interest are average larval duration ($D_{E\text{-}N6}$), time from egg spawn to C5 diapause ($D_{E\text{-}C4}$), and generation time ($D_{E\text{-}C5}$) (Gentleman *et al.,* 2008), as these development times provide insight into the timing, and/or location, of important life history events (e.g., egg spawning and diapause). Gaining more information about copepod stage durations not only enhances the understanding about their population dynamics and potential effects of climate change but may further provide implications about ecologically or economically important predators. However, more work must be done to include both inherent and environmental variability effects on individuals in current models, in order to increase the accuracy of life history timing (Gentleman *et al.,* 2008).

1.3.5 Survivorship

Obtaining average stage durations is an important factor for estimating survivorship of a population, i.e., a group of individual copepods. Survivorship (S_i) is the proportion of new stage i recruits that will survive for the duration of stage i (D_i), given a daily mortality rate for stage i (m_i, day^{-1} or d^{-1}), and molt into stage i+1. Survivorship of stage i is estimated using the relationship $S_i = e^{-m_i D_i}$, which is equivalent to the ratio of recruitment of individuals out of stage i, to those entering (i.e., R_{i+1}/R_i, see Gentleman *et al.,* 2012). The exponent of the survivorship formula, $m_i D_i$, is often considered separately as a "mortality risk" that provides a relative measure of the effect of that mortality rates on a group of individuals in stage i for the duration of that stage (more on mortality rates in section 1.4). When considering aggregate stages, survivorships of each stage are multiplied due to constant recruitment through the stages (i.e., the survivors of stage i are the new individuals into stage i+1; Gentleman and Head, 2017, Suppl. Sec. D), such that

survivorship through stages i and i+1 may be determined by $S_i \times S_{i+1}$. If mortality rates are constant for those stages, i.e., $m_{i,i+1} = m_i = m_{i+1}$ (more in section 1.4), and the law of exponents with a product of the same base is considered:

$$S_{i,i+1} = e^{-m_{i,i+1}(D_i + D_{i+1})} \tag{1.2}$$

$S_{i,i+1}$ is now the proportion of new stage i recruits that will survive for the total duration of both stages i and i+1 ($D_i + D_{i+1}$) and molt to stage i+2. Equation (1.2) shows that dependency of survivorship estimates on both the mortality rate and the duration of stages i and i+1. Essentially, the longer individuals spend in a stage, the more likely they are to die based on the mortality rate for stage i, which may have equal effects on survivorship for a shorter stage duration but relatively higher mortality rate.

1.4 Mortality

1.4.1 Importance of Mortality

Copepod population dynamics are known to be greatly influenced by mortality (Neuheimer *et al.,* 2009, 2010a, 2010b; Gentleman *et al.,* 2012), since neither food availability nor transport processes have been found to fully account for patterns in copepod distributions (Aksnes and Ohman, 1996), and certain life history and reproductive strategies are believed to be results of adaptations to minimize mortality risk (Ohman *et al.,* 1996; Gentleman *et al.,* 2012). While predation of higher trophic levels is known to be the primary source of mortality, "cannibalism" by older copepodites on earlier stages (e.g., females on eggs and nauplii) have been observed (Hirst and Kiorboe, 2002; Ohman *et al.,* 2008; Neuheimer *et al.,* 2009), as well as "losses" from starvation, non-hatching or sinking of eggs are also reported in the literature (e.g., Heath *et al.,* 2008; Gentleman and Head, 2017).

Not only do copepod stages have differences in sizes and behaviours, but they hold various roles within the ecosystem; for example, early stages provide food for many larval fish and larger copepods, while later stages prey on phytoplankton but also provide food for many larger fish, seabirds, and whales. These various roles cause mortality rates

to vary among the copepod stages (Heath *et al.,* 2008; Plourde *et al.,* 2009a, 2009b; Gentleman *et al.,* 2012), and also in space and time, dependent upon shifts in predation and environment (Gentleman, 2000; Neuheimer *et al.,* 2009, 2010b; Gentleman *et al.,* 2012). Because earlier stages are more vulnerable (e.g., predation, starvation, egg hatching failure), their mortality has generally been found to be higher than later stages (Hirst *et al.,* 2002; Ohman *et al.,* 2004; Bi *et al.,* 2011; Gentleman *et al*., 2012, 2017). In stages where mortality is higher, variability in development among individuals may give an advantage to those individuals that are able to optimize environmental conditions to reduce their stage durations and progress quickly (Miller and Tande, 1993; Gentleman *et al.,* 2008).

To increase understanding of copepod dynamics and their role in ecosystems, determining mortality rates is crucial. However, the variable and advective nature of pelagic ecosystems makes measuring mortality not possible and estimation quite difficult (Aksnes and Ohman, 1996). Meanwhile, the need for enhancing the accuracy of mortality estimates has been shown, since even the smallest of changes in mortality profoundly affects copepod dynamics simulations (e.g., Lynch *et al.,* 1998; Davis *et al.,* 2014). The most robust methods available for estimating mortality are believed to be the so-called "vertical-life table" (VLT) approaches, which use ratios of successive stage abundances at single time points, rather than requiring absolute abundances over time (which may be unavailable), as other methods do (Ohman *et al.,* 1996; Aksnes and Ohman, 1996; Ploude *et al.,* 2009b).

1.4.2 Mortality Rate Estimation using the Ratio Method

The most commonly applied VLT approach is the *Ratio* method (Mullin and Brooks, 1970; Aksnes and Ohman, 1996; Gentleman *et al.*, 2012; Gentleman and Head, 2017), which will be used here to estimate mortality rates (Chapter 2). The Ratio method is derived from a population dynamics model that incorporates several simplifying assumptions, including constant recruitment, development and mortality over a prolonged period (i.e., a "steady-state"), and negligible influence of transport on the population (Aksnes and Ohman, 1996; Gentleman *et al.,* 2012).

The formulae used for the Ratio method relates information between consecutive stages (i and i+1, termed a "stage pair") by equating the ratio of their abundances (A_i and A_{i+1}) observed at a given time and location, to a ratio of recruitment involving the respective stage durations (D_i and D_{i+1}), assuming a constant average mortality rate over the two stages (i.e., $m_i = m_{i+1}$, where the stage pair mortality rate is referred to as $m_{i,i+1}$):

$$\frac{A_i}{A_{i+1}} = \frac{e^{m_{i,i+1}D_i} - 1}{1 - e^{-m_{i,i+1}D_{i+1}}} \tag{1.3}$$

For this thesis study, equation (1.3) is used to provide an estimate of $m_{i,i+1}$ for pairs of non-adult copepodite stages (C1-C5; Chapter 2). When the latter stage (i+1) is an adult (C6), for which the duration is considered "infinite" (i.e., the copepods do not molt out, Gentleman *et al.,* 2012), the exponential term in the denominator becomes zero and it is possible to rearrange equation (1.3) to obtain an explicit relationship for $m_{C5,C6}$. To further estimate sex-specific adult mortality rates (i.e., when i+1 is separated into female, C6F, and male, C6M, stages separately), a sex ratio at molt must be specified (Gentleman *et al*., 2012). This ratio is variable (Hirst *et al.,* 2010), but a typical ratio of 1:1 is assumed (i.e., $A_{C6F} = A_{C6M} = 0.5A_{C6}$; Ohman *et al*., 2002), and equation (1.3) rearranges to become

$$m_{C5,C6F} = \frac{1}{D_{C5}} \ln\left(\frac{A_{C5}}{2A_{C6F}} + 1\right) \tag{1.4}$$

which is used to calculate the mortality for C5/C6F stage pair, or analogously for C5/C6M. For more details and derivations of these equations, refer to section 2.1.2 in Gentleman *et al.* (2012).

1.4.3 Violations of the Ratio Method

Mortality estimation methods become inaccurate when these underlying assumptions (e.g., steady-state, constant recruitment) are violated (e.g., Mullin and Brooks, 1970; Aksnes and Ohman, 1996). Method violations are most obvious when mortality rates are unable to be estimated (i.e., no numerical value is suitable, m_i = "NaN") or result in unrealistic mortality rate estimates, most commonly negatives but also detected with high

positive estimates (e.g., $m_i > 3$ d^{-1} in Heath *et al.,* 2008). Many recommendations have been made in the literature for dealing with negative mortality rate estimates, such as setting equal to zero (Aksnes and Ohman, 1996), retaining for further analysis (Hirst *et al.,* 2007), but most often they are disregarded along with other "useless" (i.e., NaN or unrealistically high) estimates (Heath *et al.,* 2008; Plourde *et al.,* 2009b). Various metrics for how to detect potential method violations directly from field data have been proposed (Hirst *et al.,* 2007; Plourde *et al.,* 2009b); a common indication of error being when an abundance (for stage *i, A_i*) is zero violating the assumption of constant recruitment (Gentleman *et al.,* 2012). Due to the Ratio method requiring *ratios* of consecutive stage abundances (i.e., A_i/A_{i+1} for stages i and i+1), zero abundances quantitatively affect calculation of these ratios. Zero abundances for stage i, i+1, or both produce "unusable" ratio results, and consequently lead to difficulties with estimating realistic mortality rates (e.g., $m_i = \pm \infty$ or are unable to be estimated; Gentleman *et al.,* 2012). Unlike the recommendations for unrealistic mortality rates provided, there has not been sufficient exploration into how to manage zero abundances in the data to potentially avoid method violations.

1.5 Simulating Population Dynamics

1.5.1 Individual-Based Models (IBM)

A model to simulate population dynamics is required to analyze the effect of dynamic environment and mortality on *C. finmarchicus* population size and structure, since observation of such effects *in situ* are not possible. Models have been used as powerful tools for investigating questions regarding ecosystem dependence on environment, but many "traditional" modelling approaches inaccurately define zooplankton life history characteristics and disregard individual variation by lumping them together (Neumeier *et al.,* 2010a). To avoid analyzing zooplankton together on a community-based scale, individual-based models (IBM) have been developed to track each individual as it is spawned, develops, and dies (e.g., Grunbaum, 1994; Neuheimer *et al.,* 2010a). IBMs were conceptually based on environmental factors affecting the physiology of each individual (i.e., possessing a measureable individual variability) and provide population-

level characteristics by considering the collection of individuals results (Grunbaum, 1994; Neuheimer *et al.,* 2010a).

An IBM describing the complex life history of copepods was built by Gentleman *et al.* (2008; also see Neuheimer *et al.,* 2010a), allowing for individual history to be simulated and results to be statistically compared for population-level observations. The IBM designates a "fitness" metric to individuals to describe inherent variability among them (see section 1.3.3), through which individual life history processes (e.g., stage duration) may be characterized using a CDF (e.g., normal, gamma; Gentleman *et al.,* 2008). Relationships describing life processes from empirical studies are included, so the IBM is highly accurate for environmental dependencies over a range of laboratory conditions (Campbell *et al.,* 2001; Gentleman *et al.,* 2008). With some modifications to assign stage durations, the IBM is also suitable for application to dynamic environmental conditions, due to individual development characterized by MCFs (see section 1.3.4, more in Chapter 3; Gentleman *et al.,* 2008). The IBM was originally built to describe *C. finmarchicus* and effects of environmental conditions on individuals (e.g., see Gentleman *et al.,* 2008; Neuheimer *et al.,* 2010a, b; modified for Gentleman *et al.,* 2012), but can be used for various applications (e.g., similar copepod species) and objectives (e.g., regarding diapause, effects of climate change on copepod populations; Gentleman *et al.,* 2008).

To provide a sense of the spatial dynamics through ocean transport that copepod populations might experience *in situ*, the IBM can be coupled with a type of physical ocean model, i.e., particle tracking, where each particle is representative of one or more individual copepods (Miller *et al.,* 1998; Pepin *et al.,* 2013).

1.5.2 Physical Ocean Models

Physical conditions (e.g., temperature and salinity) and processes (e.g., circulation and advection) of regional bodies of ocean waters can be simulated using various types of physical ocean models. Ability to simulate such processes provides an understanding of how they influence one another, and greatly enhances knowledge of their effects on conditions and dynamics of ecosystems.

Ocean waters, specifically off of NL, exhibit great seasonal variations, due to large-scale ocean circulation influences, changes in surface winds, and formation and melting of ice (Han *et al.,* 2008). Along with the AZMP physical observations, an ocean model for this region was developed by Han *et al.* (2008) to gain an understanding of physical impacts on biological processes. Their physical model is a three-dimensional (3D) unstructured-grid finite-element model, forced by buoyancy, wind, and tide (Han *et al.,* 2008; Pepin *et al.,* 2013). The wind forcings are monthly-means computed from 6-hourly data for 1990-1999, while only the most prominent tidal fields, M_2 tide, were included in the model; for more details see Han *et al.,* 2008. The physical model by Han *et al.* (2008) is able to provide simulated particle tracking information; for example, temperature and water depth, along tracks with latitudinal and longitudinal coordinates.

1.6 Organization of Thesis

The overall intent of this thesis is to enhance the understanding of mortality and environmental influence on emergent stage durations of *C. finmarchicus* copepods and how these influences on emergent stage durations may affect overall temporal and spatial dynamics of populations in ocean waters of NL. The overall thesis intent is completed through a series of analyses, including quantitative and conceptual investigation into commonly used methodologies, and providing new approaches and recommendations. A modified population dynamics model for *C. finmarchicus* is presented, allowing for environmental variability among individuals, conceptual application of development rates over stage durations, and coupling of spatial dynamics.

In Chapter 2, average mortality rates for copepodites of *C. finmarchicus* in Newfoundland-Labrador ocean waters are estimated using the commonly used Ratio method. Chapter 2 provides characterization of sampled AZMP data of copepodite abundances, temperature, and chl-*a*, as well as stage durations estimated from the environmental conditions using appropriate empirical relationships. Various data processing approaches (DPA) to obtain "average" mortality rates using the Ratio method are presented and tested by estimating mortality rates using sampled AZMP data. Ratio method violations are examined, and new data subset preparing techniques, to reduce useless mortality rate estimates, are presented and tested. Lastly, average mortality rates

are found, by averaging mortality rate results from all DPAs, to use as proxies for seasonal mortality rates of *C. finmarchicus* stage pairs for the Newfoundland-Labrador region.

Chapter 3 extends on the difficulties with using empirical stage duration estimates in dynamic environmental scenarios (introduced in section 1.3.4), by illustrating, and recommending, the use of development rates with MCF to define the developmental process, over stage durations. Various tests are conducted to demonstrate the differences in emergent stage durations due to variation in environmental conditions (e.g., constant laboratory *versus* dynamic conditions), inherent variability among individuals, and mortality.

In Chapter 4, the mortality rate estimates of Chapter 2 and stage development concepts from Chapter 3 are utilized to apply to and modify an existing *C. finmarchicus* population dynamics model (e.g., "original IBM"; see section 1.5.1), which is used to simulate the temporal and spatial effects of dynamic environmental conditions and mortality on population size and structure. The enhancements of the original IBM and additional components are outlined, presenting the new IBM with capabilities of including environmental variability among individuals and coupling to physical ocean model information. A case study is provided by simulating the development of an egg population to adulthood, subject to environmental conditions along various particle tracks, and the differences among population dynamics are discussed.

Chapter 5 summaries all findings, presents overall conclusions, and provides suggestions for future studies.

Chapter 2: Mortality Estimation

2.1 General

The objective of Chapter 2 is to ultimately estimate average mortality rates for *C. finmarchicus* copepodites in Newfoundland-Labrador ocean waters, and quantitatively analyze various data processing approaches for obtaining an average estimate using the commonly used Ratio method. Chapter 2 includes characterization of environmental conditions and zooplankton abundance samples collected by Fisheries and Oceans Canada as part of their Atlantic Zone Monitoring Program (AZMP). This field sample information is required to estimate mortality rates for *C. finmarchicus* in the Newfoundland-Labrador ocean region. Although the applied Ratio method for mortality rate estimation was developed for "single estimates" (i.e., using "station-specific" samples at one given time), it is often of interest to estimate *average* mortality rates over broader areas and/or time periods using multiple station samples. The Ratio method is conducted via a series of steps in which the necessary data may be manipulated in various ways to obtain an "average" mortality rate. Therefore, quantitative analysis of four "data processing approaches" (DPA) used in the literature has been conducted, where differences among the DPA mortality rate estimates and associated survivorship results are discussed. Many recommendations have been made in the literature for how to deal with *negative* mortality rate estimates (Aksnes and Ohman, 1996; Hirst *et al.,* 2007; Gentleman *et al.,* 2012; Head *et al.,* 2015), but there has not been sufficient exploration into how to manage zero abundances in the data to potentially avoid such method violations (see section 1.4.3). In Chapter 2, a technique to adjust (i.e., "restrict") the data prior to the mortality rate estimation process and eliminate the problematic effect of zeros is proposed and tested. Lastly, a proxy for regional and seasonal mortality rates experienced by *C. finmarchicus* copepodites over Newfoundland-Labrador ocean waters is determined by considering the averages of mortality rates estimated by all four DPAs.

2.2 Methods

2.2.1 Data

The data used in this study was sampled from ocean waters across the Newfoundland Shelf and western Labrador Sea (Figure 2.1). For this chapter, similar delineation as in Pepin *et al.* (2013) is followed describing four regions based on relative latitude, defining off of Labrador as "North" and off of Newfoundland as "South", and bottom water depth, < 300 m as "on-shelf" and ≥ 300 m as "off-shelf". These regions are hereafter referred to as NON (North on-shelf), NOFF (North off-shelf), SON (South on-shelf), and SOFF (South on-shelf) as shown in Figure 2.1.

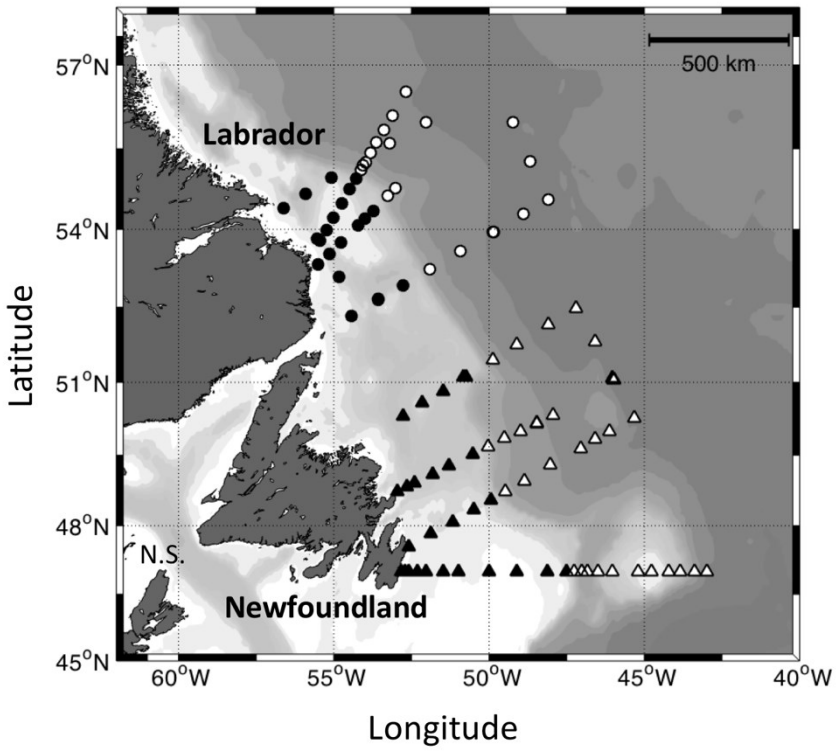


Figure 2.1 Locator map for cruise sampling stations for four regions off of Newfoundland and Labrador: NON (dark circles), NOFF (light circles), SON (dark triangles), and SOFF (light triangles)

The AZMP surveying was conducted on nine cruises between April and November of 2006, with four cruises covering the North and five cruises covering the South, sampling at a total of 180 stations (details in Table 2.1a).

Table 2.1 (a) Sampling information: cruise identifiers with dates (2006) and number of stations sampled per region; station no. = original sample size per stage. (b) Restricted Data Subsets: sample sizes, gray shading shows subsets that required restrictions, with darker shading highlighting subsets with > 50% reduction in original sample size.

(a) Sampling Information				(b) Restricted Data Subsets					
Cruise ID	**Cruise Dates**	**Shelf**	**Total No. of Stations**	**C1/C2**	**C2/C3**	**C3/C4**	**C4/C5**	**C5/C6F**	**C5/C6M**
Northern Data Subsets									
late May	May 26 – May 28	On	6	4	1	1	2	5	4
		Off	8	4	3	4	8	8	8
mid-June	June 16 – June 18	On	3	3	3	2	2	3	3
		Off	7	7	6	6	7	7	4
early August	Aug. 4 – Aug. 5	On	6	6	6	6	4	3	0
		Off	2	2	2	2	2	2	1
early Sept.	Sept. 2 – Sept. 5	On	8	8	8	8	8	7	1
		Off	7	6	4	5	7	6	4
Southern Data Subsets									
late April	Apr. 22 – May 2	On	16	11	6	3	4	9	4
		Off	16	15	14	10	11	15	9
mid-June	June 10 – June 15	On	9	9	9	8	5	6	1
		Off	12	12	12	11	11	10	4
late July	July 25 – July 31	On	17	17	15	14	11	6	1
		Off	14	13	14	14	13	12	1
late August	Aug. 25 – Aug. 31	On	9	9	9	9	9	8	1
		Off	11	6	9	10	11	10	2
late Nov.	Nov. 18 – Dec. 5	On	13	6	10	13	13	13	7
		Off	15	7	8	10	14	13	2

At each station, a vertical profile of water column properties was collected by lowering sensors at a speed of 1 m/s to the bottom, using a conductivity-temperature-depth sensor (Seabird 911) to obtain temperature measurements and a fluorometer (Chelsea AquaIII) for measurements of chl-*a* concentration (Mitchell *et al.,* 2002; Pepin *et al.,* 2011b). Average temperature (°C) for the top 25 m and average chl-*a* (mg/m^3) for the top 50 m were used to represent mixed layer values at each station, smoothing any in situ environmental patchiness (Pepin *et al.,* 2008). For this study, the data sampled during the late April cruise at one station in the SOFF region were excluded due to the temperature and chl-*a* values being extreme outliers, indicative of a potential database error. Zooplankton were collected in vertical net hauls towed from near bottom (maximum depth of 1000 m) to the ocean surface, raised at 1 m/s (Mitchell *et al.,* 2002), using a 0.75 m ring net fitted with 202 μm mesh. Once copepods of *Calanus* genus were identified and subsamples separated from other types of zooplankton, they were enumerated according to species (e.g., *finmarchicus*) and stages (Mitchell *et al.,* 2002; Head and

Pepin, 2008b). All further details of field sampling and laboratory protocols are available in Mitchell *et al.* (2002). The abundance data provided for this study was for *C. finmarchicus* copepodite stages (A_i, #/m^2, where stage i = C1 to C5, C6F and C6M), and were recorded for the following total number of stations in each region: 23 – NON, 24 – NOFF, 64 – SON, and 68 – SOFF (further separated into cruises in Table 2.1a). It is noted that the net size described above collects *C. finmarchicus* copepodites with "100% efficiency" (Head and Pepin, 2008b).

2.2.2 Stage Duration Estimates

Estimates of stage durations (D_i for stage i, d) are required inputs for mortality rate estimation methods (section 2.2.3) and are calculated here using equation (1.1). The environmental conditions used to estimate stage durations with equation (1.1) are from field samples (see section 2.2.1): the temperature measurements applied as T, while F was defined by converting the sampled chl-*a* to carbon (as described in section 1.2.2). Regional averages for stage durations are calculated by the arithmetic average of station-specific D_i estimates based on environmental samples from each cruise (Table 2.3).

2.2.3 Mortality Rate Estimation using the Ratio Method

Mortality rates (m, d^{-1}) of copepodite stages are estimated using the common Ratio method (as described in section 1.4.2), applying equation (1.3) for non-adults (C1-C5) and equation (1.4) for mortality rate estimates involving females and males (C6F and C6M). The abundances required are from field samples as described above (section 2.2.1), while stage durations are estimates calculated from environmental conditions (section 2.2.2).

Mortality rates cannot be calculated directly using equation (1.3) and, therefore, require numerical methods to solve. MATLAB's iterative root-finding algorithm, the *fzero* function, was used to solve for mortality rate, m, in equation (1.3).

2.2.4 Calculating Average Mortality Rates and Survivorship

Here, average mortality rates were estimated for copepodite *stage pairs* (6) to represent the conditions of each *region* and *cruise* (4 in the 2 North regions, and 5 in the 2 South regions; resulting in a total of 108 average mortality rates, see Table 2.1). Hereafter, the sampled (e.g., abundance, environmental conditions) or calculated (e.g., stage ratios, stage durations) suite of data required to obtain an "average" is referred to as a *subset* (see Table 2.1). The four considered DPAs are described below, with details of application and similar uses appearing in the literature outlined in Table 2.2. The names for each DPA have been chosen to reflect at which step in the mortality rate estimation process an *arithmetic average* is calculated but is simply referred to as "mean" (not to be confused with a "true mean"). Mortality rate results of all DPAs are simply referred to as "m_{DPA}" when discussed in the text, while estimates of a specific DPA are referred to as "$m_{(DPA)}$", where the abbreviation of the approach name replaces "DPA" in the parenthesis (Table 2.2, first column).

The first DPA, referred to as "mean mortality" (MM), is the most commonly used in the literature (Table 2.2). MM uses data from a single station (i.e., single location and time point) to generate ratios from consecutive stage abundances and stage durations from environmental conditions, where these are then used to estimate station-specific mortality rates and arithmetic averages are calculated to obtain "average" mortality rates for stage pairs to represent each subset. The second DPA, "mean abundance" (MA), has been used a few times in the literature (Table 2.2). MA firstly calculates average stage abundances and average environmental conditions for each by taking the arithmetic averages, from which stage ratios and stage durations are generated, respectively, and these are further used to estimate "average" mortality rates to represent each subset. The third DPA, "mean stage ratio" (MSR), rarely used in the literature (Table 2.2), follows the same initial step as MM by generating stage ratios of consecutive stage abundances for each station. MSR then differs from MM by calculating average stage ratios and average environmental conditions for each subset (by taking the arithmetic averages, similar to MA), and further uses these to estimate "average" mortality rates representing each subset. The fourth DPA, a nonlinear regression (NLR), mimics the first few steps of MSR

by finding station-specific stage ratios, but then uses an ensemble of ratios with a least-squares fit to compute an "average" mortality estimate, i.e., finds the best *m* "coefficient" for equation (1.3), representing each subset. For this DPA, MATLAB's *fzero* function was not used, but instead the *fitnlm* function was applied, which uses an iterative error minimization algorithm to find the regression.

Measures of variability for mortality rate estimates among subsets are attainable with MM, based on the station-specific estimates used to calculate average mortality rates for each subset, and also with NLR, due to assumptions about distribution of the residuals. Specifically, the full range of station-specific mortality rates estimated with MM and 95% confidence intervals (C.I.) provided by the NLR estimation were analyzed. Neither MA nor MSR are able to provide measures of descriptive statistics, beyond the averages, for mortality rate estimates, since data averaging occurs prior to mortality estimation and, therefore, only produces single mortality rate estimates for stage pairs of each subset. Note that when only one sample is present in a subset ($n = 1$), all m_{DPA} will be the same. In subsets with all abundances sampled as zeros for a given stage ($n = 0$), mortality rate estimates were indicated by "no data" in necessary figures and/or tables.

As was earlier discussed, zero abundances quantitatively affect mortality estimation by resulting in unusable stage ratios and, consequently, make estimating mortality rates using MM, MSR, or NLR problematic. However, MA is arguably robust in such situations, due to averaging zero and non-zero abundances together, providing usable stage ratios for effective mortality rate estimation. This was tested by Plourde *et al.* (2009b) who chose to use a form of MA for their main study, as opposed to a type of MM approach that resulted in many useless mortality rates due to the high number of zero abundances within their data. By using MA, their proportion of such useless mortality rate estimates substantially reduced (Plourde *et al.,* 2009b, Table 1, "no estimate" rows), and they "firmly believe" that the resulting MA mortality rates were "more-robust" in a manner similar to the well-known suggested use of "replicates" (i.e., several *m* estimates to obtain an average mortality rate, Aksnes and Ohman, 1996). A similar test was done to compare estimates of MA and MM using the AZMP data. For MM, an average mortality rate may still be obtained by discarding useless station-specific estimates of mortality

rates, however, with MSR or NLR, an average estimate is not attainable if even one stage ratio within a subset is unusable. As a workaround to discarding useful data, the original data subsets were preprocessed to create "restricted" subsets that only contain non-zero abundances for the stage pairs required for each mortality rate estimate (see example in Appendix A and subset details in Table 2.1b). Although most studies recommend disregarding negative mortalities (e.g., Heath *et al.,* 2008), negative station-specific mortalities for MM averages were included to ensure the same subset sample size for all DPAs to avoid any bias when comparing m_{DPA}. For example, consider the restricted subset for stage pair C1/C2 with 4 stations of abundance data (i.e., samples; similar to subset for late May on-shelf, Table 2.1), and suppose that the abundances from 1 of these stations provide a negative mortality estimate and is disregarded. The resultant average $m_{C1,C2}$ for MM would be determined using only 3 stations, while MA, MSR, and NLR averages would all use 4. Any negative m_{DPA} were also not disregarded to provide insight as to where, when, and for which stages method violations occur.

To put differences among m_{DPA} into context, corresponding aggregate survivorship using mortality rate estimates of non-adult stage pairs were evaluated using equation (1.2). Only positive average mortality results were used here, since negative mortalities lead to meaningless survivorship results (i.e., conceptually imply an increase in population size). Differences among m_{DPA} are examined in the context of estimated survivorship *ranges*, i.e., "range" here represents $\Delta S_{DPA} = S_{max} - S_{min}$, where S_{max} is a result of the lowest m_{DPA} for a given subset and S_{min} from the highest. Survivorship ranges less than (or equal to) 10% are considered negligible (based on sensitivity testing, not shown), while $\Delta S_{DPA} >$ 10% are considered "notable" and corresponding Δm_{DPA} are further discussed.

Lastly, "overall" average mortality rates are calculated by taking the arithmetic averages of m_{DPA} for each subset, i.e.,

$$\overline{m_{i,i+1}} = \frac{m_{i,i+1(MM)} + m_{i,i+1(MA)} + m_{i,i+1(MSR)} + m_{i,i+1(NLR)}}{4} \qquad (2.1)$$

to use as a proxy for describing regional and seasonal mortality rates experienced by copepodites in Newfoundland-Labrador ocean waters.

Table 2.2 Procedures of data processing approaches (DPAs) for determining "average" mortality rates.

DPA	Procedures for data processing approaches (as applied in present study)	
Mean Mortality (MM) $\overline{m\left(\frac{A_i}{A_{i+1}}\right)}$	*Step 1:*	• Using station-specific abundances, calculate corresponding stage ratios (A_i/A_{i+1}). • Using station-specific temperature and chl-*a*, calculate corresponding stage durations (Eqn. (1.1)).
	Step 2:	Using station-specific stage ratios and stage durations from *Step 1*, generate station-specific mortality rates, *m* (Eqn. (1.3) + MATLAB iterative solver, and Eqn. (1.4)).
	Step 3:	• Using station-specific mortality rate estimates from *Step 2*, calculate **average mortality** rates for each region/cruise subset, i.e., arithmetic average: $\overline{m}$ • Statistical measures of *m* are available for each subset (e.g., min., max., 95% C.I.)
	Ref.	Aksnes and Ohman, 1996; Ohman *et al.,* 2002; Mollmann and Koster, 2002; Ohman *et al.,* 2004; Hirst *et al.,* 2007; Gislason *et al.,* 2007; Hirst and Ward, 2008; Kvile *et al.,* 2016
Mean Abundance (MA) $m\left(\frac{\overline{A_i}}{\overline{A_{i+1}}}\right)$	*Step 1:*	• Using station-specific abundances, calculate **average abundances** for each stage within each region/cruise subset, i.e., arithmetic average: $\overline{A_i}$ • Using station-specific temperature and chl-*a*, calculate average temperature and chl-*a* for each region/cruise subset.
	Step 2:	• Using average abundances from *Step 1,* corresponding stage ratios. • Using average temperature and chl-*a* from *Step 1*, calculate corresponding stage durations (Eqn. (1.1)).
	Step 3:	Using average stage ratios and stage durations from *Step 2*, generate mortality rate estimates, *m* (Eqn. (1.1) + MATLAB iterative solver, and Eqn. (1.4)). • A single mortality rate estimate is provided for each subset (no statistical measures are available.
	Ref.	Plourde *et al.,* 2009a; Plourde *et al.,* 2009b
Mean Stage Ratio (MSR) $m\left(\overline{\left(\frac{A_i}{A_{i+1}}\right)}\right)$	*Step 1:*	• Using station-specific abundances, calculate corresponding stage ratios (A_i/A_{i+1}). • Using station-specific temperature and chl-*a*, calculate average temperature and chl-*a* for each region/cruise subset.
	Step 2:	• Using station-specific stage ratios from *Step 1,* calculate **average stage ratios** for each region/cruise subset, i.e., arithmetic average: $\overline{A_i/A_{i+1}}$ • Using average temperature and chl-*a* from *Step 1*, calculate corresponding stage durations (Eqn. (1.1)).
	Step 3:	Using average stage ratios and stage durations from *Step 2*, generate mortality rate estimates, *m* (Eqn. (1.1) + MATLAB iterative solver, and Eqn. (1.4)). • A single mortality rate estimate is provided for each subset (no statistical measures are available.
	Ref.	Mullin and Brooks, 1970
Nonlinear Regression (NLR) m_{fitted}	*Step 1:*	• Using station-specific abundances, calculate corresponding stage ratios (A_i/A_{i+1}). • Using station-specific temperature and chl-*a*, calculate corresponding stage durations (Eqn. (1.1)).
	Step 2:	Using station-specific stage ratios and stage durations from *Step 1*, generate mortality rate estimates, *m,* by **nonlinear regression** (Eqn. (1.1) + MATLAB nonlinear regression, and Eqn. (1.4)). • Statistical measures of *m* are available for each subset (e.g., 95% C.I.)
	Ref.	Bi *et al.,* 2011b

2.3 Results

2.3.1 Characterizing Regional Ecology

All regions exhibited very low variability for temperature and chl-*a* within each subset of samples in the summer and fall cruises (Figure 2.2; C.V. < 0.05), with greater variability occurring in the South during late April (on-shelf: $C.V._{chl\text{-}a} = 0.16$; off-shelf: $C.V._{chl\text{-}a} = 0.10$) and late November cruises (on-shelf: $C.V._{T} = 0.38$, $C.V._{chl\text{-}a} = 0.25$; off-shelf: $C.V._{T} = 0.10$, $C.V._{chl\text{-}a} = 0.29$). In all regions, sampled temperatures were typical for mid-latitude seasonal levels, highest in summer and lowest in spring (1.5 – 5.1 ºC), and for the South regions, dropping again to lower levels in autumn. Peak temperatures were colder in the North relative to the South (10 as compared to 14 ºC), and, across the seasons, on-shelf temperatures were often colder than off-shelf by differences of about 0.4 – 3.6 ºC between corresponding cruises. The highest average chl-*a* concentrations were 1.00 mg/m^3 in the North and 2.55 mg/m^3 in the South (both observed on-shelf). For more regions, maximum seasonal chl-*a* concentration occurred during the earliest cruise, while another rise in chl-*a* concentration is observed again in late summer.

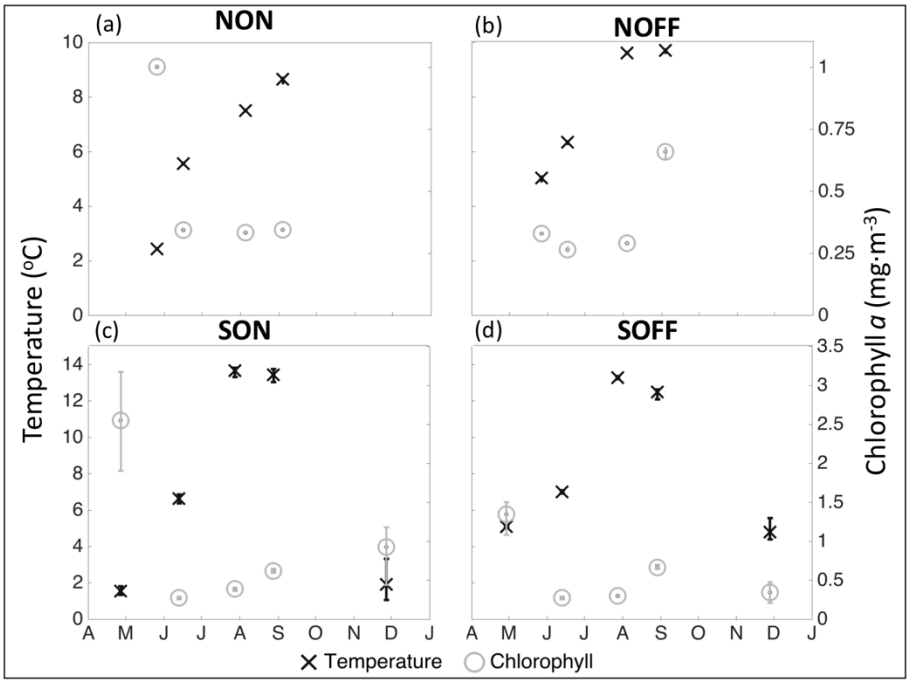


Figure 2.2 Temperature (x - markers) and chlorophyll-a concentration (o - markers); cruise means (x/o markers) and minimum/maximum points (error bars) are displayed at mid-point date of cruise durations for all 4 regions.

Station-specific stage durations were determined by using equation (1.1), and then the arithmetic average for each subset was calculated to represent an average value (Table 2.3). The average stage durations range from 2.5 to 41.6 days ($\overline{D_{C1}}$ and $\overline{D_{C5}}$, both estimated for the SON region). The ranges of station-specific D_i estimates within each subset (i.e., the difference between maximum and minimum D_i) are generally low, < 1 day for all stages in the North and most stages in the South. The largest D_i ranges within South subsets are for all stages in late November (1-7 days for SON subsets and > 3 days for SOFF subsets) and for C5s in late spring to summer (1-7 days). Seasonal patterns observed for temperature and chl-*a* samples were reflected in stage duration estimates; mostly influenced by temperature, but there are instances of notable chl-*a* effects. Within regions, longest stage durations are estimated for cruises with coldest sampled temperatures, and inversely shortest stage durations with warmest temperatures. However, for the North regions, chl-*a* has a strong effect on stage duration estimates in late May, where higher chl-*a* concentrations for the on-shelf result in equal stage durations to those estimated for the warmer, yet lower chl-*a*, off-shelf conditions at the same time (Figure 2.2; Table 2.3).

Table 2.3 Regional average stage durations (days) for non-adult copepodites (C1-C5). Range (maximum-minimum) of station-specific stage durations are shown with shading: white (< 1 day), light gray (1-3 days), medium gray (3-7 days), and dark gray (> 7 days).

Stages	NON				NOFF			
	late May	mid-June	early Aug.	early Sept.	late May	mid-June	early Aug.	early Sept.
C1	8.6	7.8	6.1	5.3	8.6	8.0	5.2	3.6
C2	10.1	9.2	7.2	6.2	10.1	9.5	6.1	4.2
C3	12.7	11.5	9.1	7.8	12.7	11.9	7.6	5.3
C4	19.2	17.5	13.8	11.8	19.2	18.0	11.6	8.0
C5	36.2	32.9	25.9	22.6	36.3	34.0	21.8	15.1

Stages	SON					SOFF				
	late Apr.	mid-June	late July	late Aug.	late Nov.	late Apr.	mid-June	late July	late Aug.	late Nov.
C1	8.6	7.6	3.0	2.5	9.8	5.3	7.2	3.5	2.7	9.1
C2	10.1	9.0	3.5	3.0	11.6	6.2	8.5	4.1	3.2	10.7
C3	12.7	11.2	4.4	3.7	14.5	7.8	10.7	5.2	4.0	13.5
C4	19.2	17.1	6.7	5.6	22.1	11.8	16.2	7.9	6.0	20.5
C5	36.3	32.1	12.6	10.6	41.6	22.2	30.5	14.8	11.3	37.1

Unlike the regional similarity in environmental conditions and stage durations, the seasonal total copepodite abundance and population structure dramatically differ among regions (Figure 2.3). Early stage copepodites in the South already dominate by late April, in contrast to the North, where abundances are still overwhelmingly dominated by adults in late May. C1s are present in June for the North and, similar to the South, remain through to September, with their relative proportion decreasing from summer to autumn months (a greater decline for off-shelf as compared to on-shelf). Although early to mid-stages are still somewhat active in later cruises (September for the North and November for the South), at this time the bulk of the regions' sampled populations are C5s (with the exception of samples from the NON region, still dominated by early stage copepodites at this time). The off-shelf regions contain overall greater proportions of C5s, which is consistent with studies finding dormant copepodites in deeper waters (Pepin and Head, 2009), whereas copepodites on-shelf have been found to remain active for multiple generations (Pepin *et al.,* 2011a). Average male abundance is extremely low across the year for all regions and although average female abundances are considerably larger than males, they are still quite low relative to other stages after the first cruise. Zero abundances were recorded for less than 20% of the samples within each region. Of the 126 total abundance data subsets, only one had all zeros recorded (C6M for the NOFF samples in early August), and half (63) had at least some zero abundances. The largest proportions of recorded zero abundances for early copepodite stages in the North are in late May, while for later stages are in the late summer. Similar sampling patterns of zero abundances are observed in the South for early to mid-stage occurring mostly in April and November, while zero abundances for later stages are prominent in early to mid-summer.

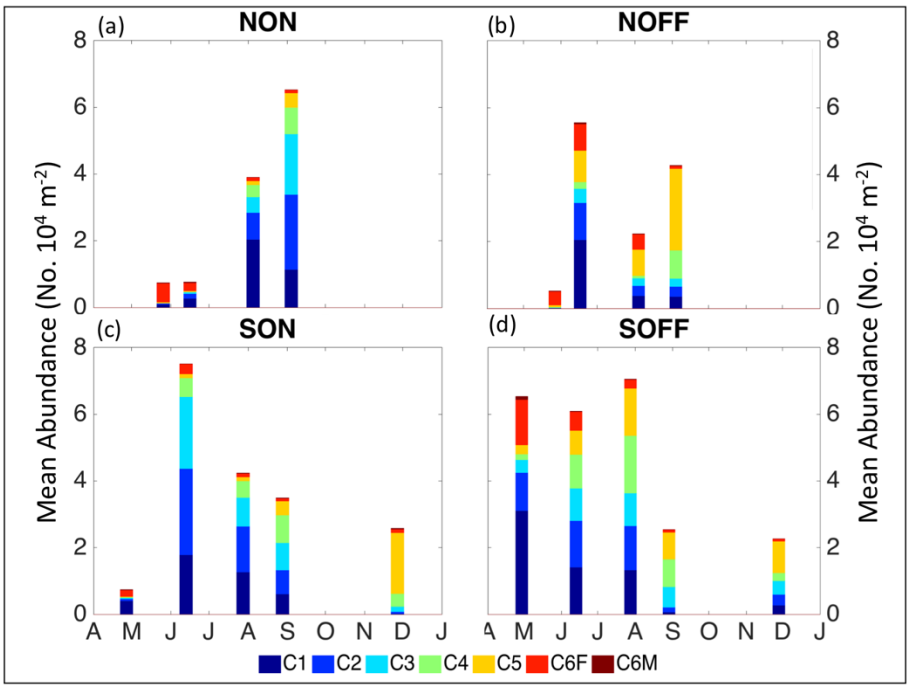


Figure 2.3 Sample means of *Calanus finmarchicus* copepodite stage abundance observed in 2006 for 4 regions of study. The height of each stacked bar, plotted at the mid-point day of each cruise duration, is the total of all stage sample means.

2.3.2 Methodological Analysis of Mortality Rate Estimation

In this section, sensitivity of DPAs (section 2.2.4) on average mortality rate estimates are explored by examining differences in m_{DPA} results with regard to estimated values and signs, corresponding survivorship estimates, and measures of variability where applicable. These "comparisons" are not assessed through standard statistical hypothesis testing, since only two of four DPAs are associated with measures of variability. General patterns and notable results are discussed below, and the complete set of results are available in Figure 2.4 and Figure 2.5.

2.3.2.1 Estimating Mortality with Original versus Restricted Data Subsets

Zeros in the abundance subsets consequently affect an even higher proportion of stage ratio subsets (70 of 108 subsets, Table 2.1), by causing quantitative issues for calculating ratios with corresponding abundances, and, therefore, the subsets were restricted (described in section 2.2.4) to avoid such issues. Very few of the restricted subsets required more than 50% of the original stage pair samples to be removed (Table 2.1),

with notable exceptions being the earliest on-shelf cruises and male abundances sampled mid-summer to fall, which had at least 50% of the original stage pair samples removed.

Following similar tests done by Plourde *et al.* (2009b), average mortality rates estimated with MM and MA were compared, using the original subsets, to test the sensitivity of mortality results to zero abundances. The results showed that by using MM, 20-36% of station-specific mortality estimates among regions were unable to be estimated, which affected all but one average mortality rate calculation. By using MA for the same subsets, all data was used, and only one (of 108) $m_{(MA)}$ was unable to be estimated due to all zero abundances in the associated subset (A_{C6M} for estimating $m_{C5,C6M}$). The test was extended to include MSR and NLR, both of which use the same stage ratio subsets for estimation. These DPAs were unable to estimate mortality rates for the same number of cases, 38-63% of mortality rate estimates regionally, due to the application of the same stage ratio subsets.

To further investigate the sensitivity of estimated average mortalities to zero abundances in the data, MA was applied to analyze the effect of using restricted subsets as compared to original subsets. Generally, the differences in mortality rate estimates generated by the two types of subsets were negligible (see Appendix B). In the few cases where, visually, differences in mortality estimates were apparent (e.g., Figs. B.1a, c), further investigation into the applied subsets showed signs of "disproportionate removal" of data. For example, restricted subsets for stages i and i+1 (estimating $m_{i,i+1}$) removed zeros in either subset (i or i+1), along with the "stage pair" abundances that happen to be relatively larger than the rest of the subset, skewing the average abundance. This scenario appeared to only have a notable effect on low mortality rate estimates ($-0.02 < m < 0.02$ d^{-1}), causing a shift in the signs (e.g., Fig. B.1b) This test had the most effect on mortality rates involving males, i.e., $m_{C5,C6M}$, because of the large number of zero sampled male abundances, causing all but two of the male, and corresponding C5, subsets to be restricted.

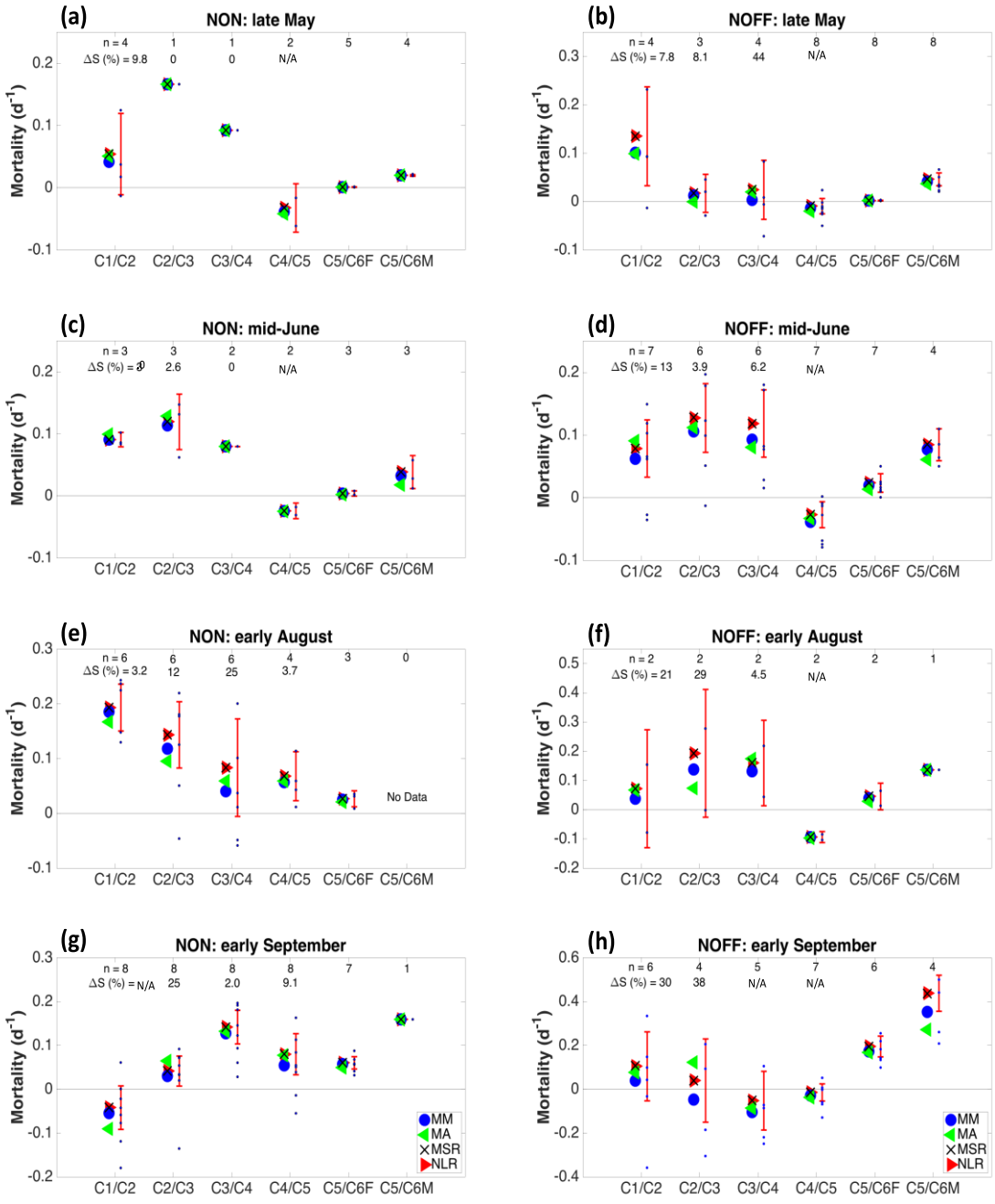


Figure 2.4 Average mortality rates for North subsets estimated using various DPAs: MM (blue circles), MA (green left-pointing triangle), MSR (black 'x'), NLR (red right-pointing triangles). Measures of variability provided for MM, station-specific estimates (smaller blue circles), and NLR, 95% C.I. (red bars). The subset sample size (n) is provided at the top of each window. Maximum difference in DPA survivorship (%) estimated for positive C1-C5 m_{DPA} are provided beneath n-samples; cases with no positive m_{DPA} are labelled "N/A". Columns divide regions, while rows represent cruises.

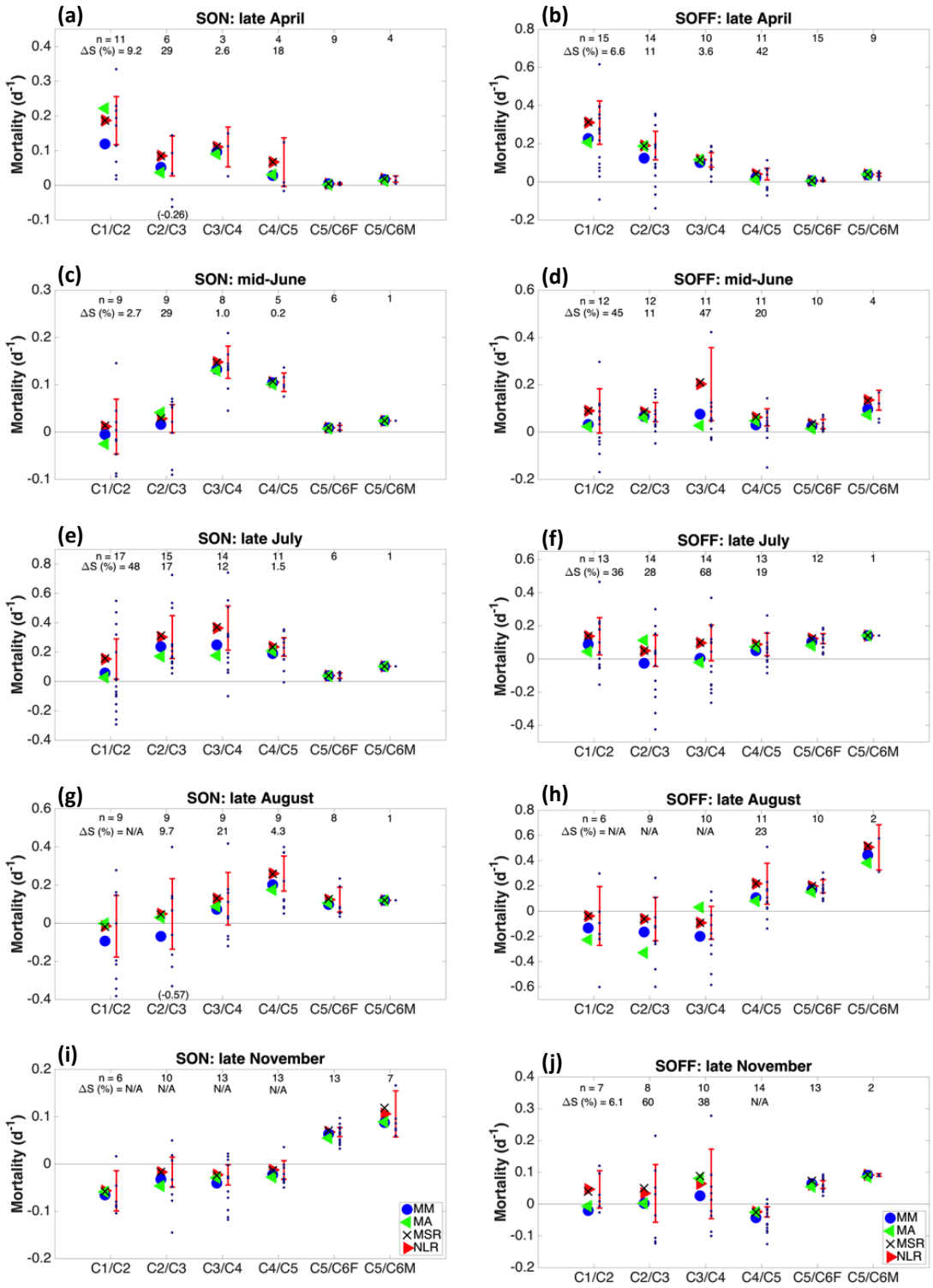


Figure 2.5 Average mortality rates for South subsets estimated using various DPAs: MM (blue circles), MA (green left-pointing triangle), MSR (black 'x'), NLR (red right-pointing triangles). Measures of variability provided for MM, station-specific estimates (smaller blue circles), and NLR, 95% C.I. (red bars). The subset sample size (n) is provided at the top of each window. Maximum difference in DPA survivorship (%) estimated for positive C1-C5 m_{DPA} are provided beneath n-samples; cases with no positive m_{DPA} are labelled "N/A". Columns divide, while rows represent cruises.

2.3.2.2 Method Violations Detected with Negative Mortality Rate Estimates

Although use of restricted subsets allows for more mortality rates to be estimated and for less data to be disregarded, the Ratio method may still be violated by measures of nonzero abundances. The results did not show any violations detected due to unusually high mortality rates, but many instances of negative average mortality rate estimates were observed, seemingly a result of abundance samples of stage *i*+1 being larger than those of stage *i*. This abundance scenario may be a result of specific life history timing coinciding with sampling or ocean currents bringing other populations to the sampling area, which provides misleading abundances and violates mortality estimation methods. The sensitivity of $m_{(MM)}$ to negative station-specific mortalities, recommended to be disregarded by the literature (see section 2.2.4), was tested by comparing averages calculated using all station-specific mortalities (i.e., $m_{(MM)}$) to mortality rate averages calculated with only positive station-specific mortalities, "$m_{(MM)}+$" (test results not illustrated, but see Figure 2.4 and Figure 2.5 for the range of MM station-specific mortalities). As expected, the removal of negative station-specific mortalities from subsets resulted in more positive $m_{(MM)}+$, due to the sign changing for all cases where the range of station-specific mortalities encompassed estimates of both signs, and only four (of 108) $m_{(MM)}$ were unable to be calculated where the entire range of station-specific mortalities were negative (Figure 2.4a, c, f). Negative $m_{(MM)}$ became *low* $m_{(MM)}+$ in cases where there was a greater (or equal) number and higher magnitude of negative station-specific mortalities for that subset, heavily weighting $m_{(MM)}$ to negative (e.g., see Figure 2.5i). For average mortality rate estimates involving adults, $m_{(MM)}$ and $m_{(MM)}+$ were the same, since all station-specific mortality rate estimates were positive. There was a high proportion of $m_{(MM)}$ for stages C1-C5 calculated to be positive (50 of 72), where removal of negative station-specific mortalities did not affect the *sign* of the average, as expected, and differences in survivorship estimates with $m_{(MM)}$ and $m_{(MM)}+$ were either negligible or experienced a decrease (i.e., $S(m_{(MM)}) \geq S(m_{(MM)}+)$. The cases where survivorship estimates decreased were a result of average mortality rate estimates being higher after negatives were excluded (i.e., $m_{(MM)} < m_{(MM)}+$), and while most survivorship decreases were not substantial, notable decreases in survivorship estimates for North subsets were $\Delta S = 66\%$ and for South subsets were $\Delta S = 85\%$.

2.3.2.3 Comparing DPA Mortality Rates and Survivorship Estimates

For the North subsets, m_{DPA} are all, generally, less than 0.2 d^{-1} (Figure 2.4) and differences of positive m_{DPA} within subsets are mostly small ($\Delta m_{DPA} < 0.05$ d^{-1}). The largest differences in m_{DPA} are for Augusts subsets for both shelf regions (max. $\Delta m_{DPA} = 0.12$ d^{-1}; Figure 2.4e, f). Estimated values and differences among m_{DPA} for South subsets are generally higher than those for the North, with most estimates less than 0.3 d^{-1} (Figure 2.5) and most differences < 0.10 d^{-1}. The largest differences among m_{DPA} within South subsets are in late July for on-shelf (max. $\Delta m_{DPA} = 0.19$ d^{-1}; Figure 2.5e) and mid-June off-shelf (max. $\Delta m_{DPA} = 0.18$ d^{-1}; Figure 2.5d, h). All *negative* m_{DPA} were estimated for few subsets (16 of 108) and estimated m_{DPA} with mixed signs occurred for even fewer cases (7 of 108). Results show that the largest amount of negative m_{DPA} were estimated by MM, with MA following, and MSR and NLR producing the least. Estimates of *positive* average mortality rates generated with MSR and NLR were very similar but differed from both MM and MA average mortality rate results (Figure 2.4 and Figure 2.5). The similar MSR and NLR average mortality rate estimates were mostly higher than those estimated with either MM or MA (i.e., generally: $m_{(MSR)} \approx m_{(NLR)} > m_{(MA)} > m_{(MM)}$), and, therefore, survivorship estimates were lower (i.e., generally: $S_{(MSR)} \approx S_{(NLR)} < S_{(MA)} < S_{(MM)}$). Average mortality rates by MM appear to be generally the lowest, due to a large proportion of station-specific negative estimates, while MA average mortality rates are lower than those estimated by MSR and NLR, due to the applied stage ratios for MA being smaller than those for MSR/NLR, a result of averaging abundance subsets first.

Overall, there were more notable survivorship differences among DPA estimates (i.e., $\Delta S_{DPA} > 10\%$) for South subsets (21) than North subsets (9), and similarly for off-shelf subsets (19) than on-shelf subsets (10; Figure 2.4 and Figure 2.5). Most of these cases of notable survivorship differences were for subsets from summer cruises, where the largest differences within the regions were 25% for NON and 38% for NOFF (both early September; Figure 2.4e, h), and 48% for SON and 67% for SOFF (both late July; Figure 2.5e, f).

The measures of variability provided for MM (range of station-specific mortality estimates) and NLR (95% confidence intervals) encompass most of the m_{DPA}, indicating that differences in mortality rate estimates due to DPAs are secondary to differences within the subsets. Both MM and NLR measures of variability are generally larger for subsets of: earlier than later stage pairs, the South than the North, and off-shelf than on-shelf (Figure 2.4 and Figure 2.5). The range of MM station-specific mortality rate estimates in some cases are quite broad and asymmetrical about the $m_{(MM)}$, often including both negative and positive results (e.g., Figure 2.5e). If *negative* station-specific MM mortality rate estimates are ignored (as was tested in the previous section), the ranges of *positive* station-specific mortality rate estimates are still quite wide, causing vastly different survivorship estimates within that subset (e.g., Figure 2.5e).

2.3.3 Seasonal Mortality Rates among Regions

In the following section, "overall" average mortality rates are analyzed, calculated using equation (2.1). Notable findings were highlighted regarding seasonal variability in mortality rates among stage pairs within each region, similarities among regional mortality rates, and cases where method violations are detected by negative average mortality rates.

Mortality rates for the NON region (Figure 2.6a) are highest for early copepodite stage pairs, C1/C2 and C2/C3, in early August and late May, respectively, where there is notable stage-to-stage variability for $m_{C1,C2}$ and $m_{C2,C3}$. While $m_{C1,C2}$ gradually increase over the cruises, peaking in early August, $m_{C2,C3}$ appear to decline over this same time, which makes interpreting mortality effects on C2 alone difficult. Method violations are detected by $m_{C1,C2}$ in early September and by $m_{C4,C5}$ for springtime cruises. Mortality rates for males and females show similar increase from late May to a substantial maximum in early September (with the exception of no male data in August), $m_{C5,C6M}$ being larger than $m_{C5,C6F}$.

The highest mortality rates for early stage copepodites for the NOFF region (Figure 2.6b) are in late May, comparable to mortality rate results for the NON region, while mortality rates for mid-stage copepodites are highest in early August. Interestingly, method

violations are detected by $m_{C4,C5}$ for all cruises. Mortality rates involving males and females have similar increasing values over the year, with maximum mortality rates in early September, $m_{C5,C6M}$ being larger than $m_{C5,C6F}$, and, overall, are larger than corresponding $m_{C5,C6M}$ and $m_{C5,C6F}$ for the NON region.

Early stage copepodite mortality rates for the SON region (Figure 2.6c) show great seasonal variability with large changes in magnitude from cruise to cruise. The highest mortality rates for most non-adult stage pairs occur in late July (C2/C3 to C4/C5), but in late April for C1/C2. Mortality rates for C3/C4 and C4/C5 show similar gradual increase in mortality rates from late April to late July, suggesting that mortality effects on C4 alone may be interpreted. Method violations appear for all non-adult stage pairs in late November and also for early stages in late August. Similar to both North regions, the male mortality rates shadow those involving females, with both mortality rates peaking in late August.

Mortality rates for early to mid-stage copepodites for the SOFF region (Figure 2.6d) show large seasonal variability, with maximums in late April and mid-June, respectively. Method violations are indicated by early stage mortality rates in late August and by $m_{C4,C5}$ in late November. Similar to mortality rates involving adults in the other three regions, seasonal increases in $m_{C5,C6F}$ and $m_{C5,C6M}$ are rather alike, with $m_{C5,C6M}$ larger than $m_{C5,C6F}$, and both mortality rates are highest in late summer.

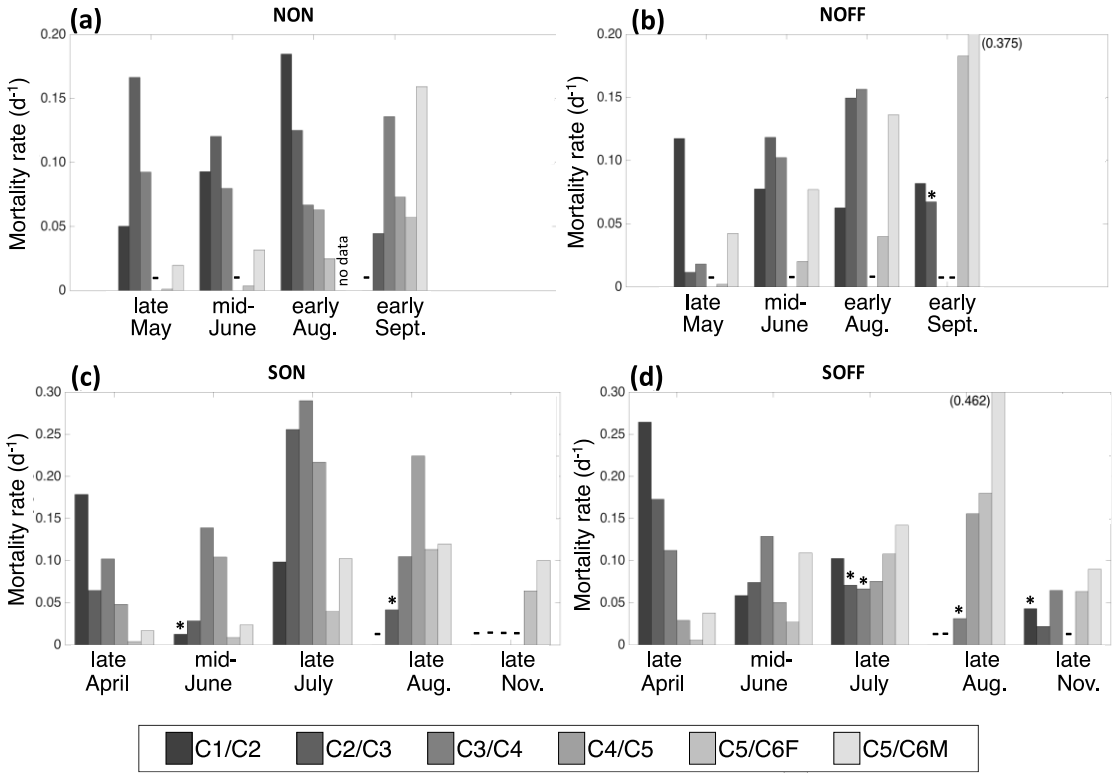


Figure 2.6 Stage-pair mortality rates: averages of m_{DPA}, where all m_{DPA} were negative are denoted by "-", while those with mixed signs (negatives removed from the final average) are marked by "*".

2.4 Discussion

2.4.1 Characterizing Regional Ecology and Stage Duration Estimates

The timing of cruise sampling in all four regions appeared to coincide with the chl-*a* bloom in the spring, and a second chl-*a* bloom in the autumn for most regions, except for the NON region. A chl-*a* autumn bloom may in fact occur in the NON region, but later than it occurs for conditions of the NOFF region and was, therefore, not captured by the sampling. The delay in peak temperatures among these two North regions suggests that a similar delay in chl-*a* bloom may occur. Average stage duration estimates and variability among subsets follow seasonal patterns observed with sampled temperature and chl-*a* as expected; i.e., stage durations are relatively longer for subsets where environmental conditions are lower, and shorter stage durations are due to higher conditions.

Regional differences for average copepodite abundance samples are quite pronounced but, in some cases, may be explained by considering timing of life history events. Recruitment into C1 appears to have begun earlier for both off-shelf regions, as compared to respective on-shelf regions, and therefore suggests that egg spawning and emergence from diapause also occurred earlier in the off-shelf regions, as compared to on-shelf. In the autumn, the off-shelf regions also appear to be entering diapause earlier than their neighbouring on-shelves, indicative of their respective earlier accumulation of C5s at this time. These observations suggest an overall earlier shift in the timing of life history events for deeper waters, as compared to shallower, on-shelf regions. However, the larger abundances of C5s observed in the off-shelf regions year-round may be partly due to a previous generation of C5s from the on-shelf sinking to greater depths (i.e., off the shelf) during diapause. These C5 abundances are highly reflective of diapause behavior for this region as discussed in Pepin *et al.* (2011a), where "prior observations" have found that, generally, C5s across the off-shelves emerge by April, begin to enter in September, and the bulk of the population are deep and dormant in November, while many on-shelf copepodites are still active in late autumn (Head and Pepin, 2008a; Pepin and Head, 2009). This may explain the earlier recruitment of copepods in off-shelf regions, as the waking C5s molt to adults and reproduce, while some on-shelf copepods may never have entered diapause and are not inclined to reproduce at this time. The relatively later development into early stage copepodites in the North regions (early summer), as compared to the South regions (mid-spring), agrees with studies of broader areas in Johnson *et al.* (2008), and may be explained by longer stage durations (i.e., due to overall colder temperatures and lower chl-*a* concentrations), which also suggests that emergence and entrance of diapause also occur later in the North, due to the combination of longer durations for each stage. Most instances of recorded zero abundances are not unexpected but coincide with life history timing when sampling variability would be high for that given stage. Large proportions of zero abundances are recorded for early stage copepodites in the springtime, which is indicative of a recruitment delay of the first generation to copepodite stages. Similarly, zero abundances recorded for late stage abundances are mostly in late summer and are presumably due to the onset of diapause, where the later stages have sunk to lower than sample depths.

The variability in regional abundances as compared to similarities of regional environmental conditions suggests that the physiological rates (e.g., EP, development) characterized by these environmental conditions are unable to fully explain the differences in copepod population size and structure, and, therefore, consideration of mortality rates is crucial to fill the knowledge gap.

2.4.2 Methodological Analysis of Mortality Rate Estimation

2.4.2.1 Recommended Use of Restricted Subsets

Mortality rates estimated with the original data subsets by applying MM and MA provided results that agree with those found in Plourde *et al.* (2009b). The use of MA, over MM (also MSR and NLR, for this chapter study), allows for less data to be discarded, since the amount of useless mortality rates arising from zero abundances (which result from application of MM, MSR, or NLR) are decreased. Generally, mortality rates estimated by MA with original and restricted subsets had smaller differences relative to the differences among m_{DPA}. While it is "unknown" as to which DPA is "most accurate" at estimating mortality rates, the use of restricted data subsets is highly recommended to utilize as much sampled data as possible and to provide the flexibility to use any of the DPAs and have minimal useless mortality estimates. Creating restricted subsets are especially recommended if comparing mortality rates estimated by various DPAs (as was done in this chapter), to ensure the results are not biased by applying different subsets.

2.4.2.2 Implications of Differences Among m_{DPA}

The similarity between MSR and NLR average mortality rate estimates is expected because both approaches use the same stage ratio subsets, but yet is rather interesting, given the DPAs different levels of programming complexity. The agreement between $m_{(MSR)}$ and $m_{(NLR)}$ suggests that average mortality rates may be obtained in a faster, and arguably, easier way and by eliminating the need for intricate software packages or so-called "black boxes" by using MSR over NLR. Unlike the complexity of creating a nonlinear regression program, MSR is easily implemented and allows total control over each step of the estimation process. The downside of using MSR is that it produces a

single average mortality rate estimate for each subset, and so variability is unable to be measured, as opposed to the suite of statistical measures available with NLR. Average mortality rates estimated with MM and MA are found to be generally lower than MSR and NLR, which implies that many studies in the literature are potentially underestimating average mortality rates, since MM and MA are the most commonly used (DPA references in Table 2.2). Underestimating average mortality rates leads to larger survivorship estimates which may further overestimate predicted populations.

The cases where m_{DPA} estimates had mixed signs highlight the importance of understanding that there are various approaches available for use and that negative results may potentially be avoided by simply applying a different approach. The results show that MM and MA produce the largest amount of negative m_{DPA}, as compared to MSR and NLR. This is an interesting find, again, given that MM and MA are the most commonly used in the literature.

Although most differences among m_{DPA} are small, the survivorship estimates show the importance of putting mortality rates into context. In some cases where survivorship differences were found to be notable, simply considering differences among m_{DPA} values could be negligible. For example, the differences in survivorship estimates for C2/C3, in the SOFF region, in late November was substantial (60%; Figure 2.5j), while the difference among m_{DPA} for the same subset was small ($\Delta m_{DPA} = 0.05$ d^{-1}). The large difference in survivorship estimates for the previous example are a result of stage durations being longest in late November, relative to the other cruises, which amplifies the effect of mortality rate differences on survivorship. By simply analyzing the value of the mortality rates, "important" differences in context may be overlooked.

2.4.2.3 Comparison of m_{DPA} to Measures of Variability of MM and NLR

The available MM and NLR measures of variability within subsets were found to be generally large for earlier stage pairs, the South regions, and off-shelf regions, which in some cases may be explained by presence of high sampling variability due to timing of life history events. MM and NLR variability in mortality rate estimates are greatest in the summer which may be reflective of sampling variability of early stages because

abundances are assumed to be relatively low at this time, since populations have most likely molted to later stages. The higher MM and NLR variability measures in the South subsets, as compared to North, are surprising given the larger sample sizes and the recommendation from the literature to use replicates with MM so that mortality rate estimates are "more robust" (Aksnes and Ohman, 1996; Plourde *et al.,* 2009b). The results here show that having more replicates does not necessarily provide more robust average mortality rates because, not only are the NLR 95% C.I. and MM range of mortality rate estimates larger, the differences among m_{DPA} for subsets with more samples are also greater. Therefore, more replicates do not necessarily result in "most robust" average mortality rates and users are cautioned about fully trusting the accuracy of mortality rate estimates solely based on larger sample sizes, as sampling variability may have considerable effects on mortality estimation. MM range of station-specific estimates were often asymmetric about corresponding $m_{(MM)}$ estimates and greater than NLR's 95% C.I., which suggests that using standard errors may not be an appropriate measure for mortality rate estimates, as it can lead to under- or overestimating the range of potential mortality results, depending on chosen DPA. Overall, although the use of various DPAs does result in different mortality rate estimates, these differences are not significant when compared to the ranges provided by measures of variability with MM and NLR, and more work needs to be done to manage sampling variability and provide more confidence in mortality estimation.

2.4.3 Seasonal Mortality Rates among Regions

Seasonal mortality rate estimates can provide information regarding potential predation risk for each copepodite stage pair within each region. Also, where method violations are detected by negative mortality rate estimates, the sampling variability is assumed to be high based on sampling dates coinciding with specific timing of copepod life history.

The greatest mortality risk for early staged copepodites appears to be in mid- to late spring, due to the combination of both highest mortality rates and longest stage durations. Given that in the springtime the first generation is expected to be recruiting into the copepodite stages from nauplii, this high mortality risk suggests that the timing of

predators, such as larval or young fish, are coinciding with this predominance of early stages in the population. However, the mortality risk of early stages in the NON region appears to be relatively lower in the spring, with low mortality rates coinciding with longer stage durations, which implies that predators may be choosing relatively more southern or off-shelf regions to feed. This may provide behavioral information of predators, suggesting that they may be migrating northward from regions further south, stopping to feed at the South regions, or perhaps may be choosing to feed in deeper waters (i.e., off-shelf regions), so as to minimize their own mortality risk from higher trophic predators.

Mid-staged copepodites appear to have the highest mortality rates in the summertime for most regions; however, their actual risk of mortality at this time is decreased, due to stage durations being the shortest. The seasonally higher mortality rates for some early and mid-staged copepodites at this time suggest that there may be shifts in predator guilds; i.e., copepodites developing to mid-stages may now be at risk to larger sized fish. Meanwhile, early stages are still at risk of consumption by developing and mature fish.

Stage pair mortality rates for C5s and adults appear highest in late summer for all regions, which may be indicative of diapause preparation causing bias on mortality estimation. The methods are estimating large mortality effects for the cases of relatively smaller proportion of adults to C5 abundances, which may in fact be due to the annual accumulation of C5s preparing for overwintering, and not a true reflection of C5 or adult mortality at this time. Nonetheless, the high mortality rates may provide information regarding timing and location of whale migration for feeding, as the storage of lipids in diapause-ready C5s are known to provide nutritious food source for North Atlantic right whales. The late stage copepodite mortality results would indicate that these whales are feeding in these regions in late summer, and, as suggested, with higher rates for off-shelves as compared to the on-shelves, preferably in deeper waters. The latter speculation may be validated with observations of higher abundances of C5s in the off-shelves as compared to the on-shelves, presumably an effect of sinking during diapause and a contribution from on-shelf populations.

Method violations detected by mortality rates for C4/C5 stage pair across the seasons for North regions, and autumn for South regions, are attributed to consistently higher C5 abundances. For summer and autumn, the high abundances are due to the accumulation of C5s in preparation for diapause. The overall higher C5 abundances in the NOFF regions may be indicative of a contribution of C5s from the on-shelf, suggesting variable timing of diapause in this region and sinking to deeper waters. However, the higher C5 abundances in the NON regions are unusual, and perhaps indicate large effects of transport and contribution of C5s from more northern areas. The method violations detected by early stages later in the year in the South regions are largely due to sampling variability, as early staged copepodites are not expected to be abundant at that time.

Where differences in seasonal estimates among adjacent stage pair mortality estimates arise (e.g., North region estimates for $m_{C1,C2}$ and $m_{C2,C3}$), contrasting survivorship conclusions regarding separate stages (e.g., C2 for the provided example) may be implied, and therefore, these methods may only be appropriate for making inferences regarding aggregate stages for each cruise duration (i.e., cannot separate stages to determine seasonal patterns; similar finding in Kimmerer, 2015).

Chapter 3: Development Rates

3.1 General

The objectives of this chapter are to expand on the general introduction of stage durations (section 1.3) by demonstrating the conceptual difference of assigning development rates, as opposed to fixed stage durations, and to illustrate the effects of dynamic environment and mortality on emergent stage durations.

The conceptual relationships between molt-cycle fraction (MCF), development rates, and stage durations are described using simple derivations and illustrated with corresponding figures. These relationships are expanded upon to demonstrate how they are used in modelling studies with use of environmental field samples and are modified to assign and apply development rates, rather than stage durations.

The effect of various factors on resultant stage durations are tested by simulating a group of individuals developing through the same generic stage, testing various temperature scenarios, the addition of inherent individual variability through a "fitness" metric, and the application of a constant mortality rate.

Emergent stage durations (i.e., "average", introduced in section 1.3.4) are expected to be shorter with relatively higher temperatures, than emergent stage duration results from lower temperatures, as found by empirical studies relating stage duration to environment (e.g., Campbell *et al.,* 2001; Gentleman *et al.,* 2008). Lower temperature scenarios with individual variability in development rates are expected to provide a larger distribution of resultant stage durations (e.g., higher variance), than those with higher temperatures (as observed in Gentleman *et al.* (2008)). Statistical measures of resultant stage durations are calculated for survivors (i.e., not all will survive when mortality is introduced) and compared among tests. Distributions of resultant stage durations are expected to display a *positive* skew (lean to the left; a likely characteristic of gamma distributions), which is hypothesized to occur *in situ*, anticipating individuals would optimize developmental "resources" (e.g., temperature) and result in stage durations that fall on the favourable (i.e., lesser) side of the average. Considering individual variability in development rates,

individuals with relatively faster development rates than others are likely to develop and molt out of that stage earlier (i.e., have shorter stage durations), and, therefore, will have less probability of mortality in that stage. This "survival of the fittest" notion hypothesizes that inclusion of mortality will cause a "shift" in the distribution of resultant stage durations of the surviving individuals by decreasing, both, the emergent stage duration and the variance.

3.2 Methods

The methodology provided here is not novel, as it has already been developed and used in modelling studies with *stage durations* (e.g., Gentleman *et al.,* 2008; Neuheimer *et al.,* 2010a). In this chapter however, the methods are redefined to directly use *development rates* and have stage durations considered as an emergent property. In this section, the relationships between MCF (introduced in section 1.3.4), development rates, and stage durations are demonstrated.

3.2.1 Existing Methods Re-derived for the Use of Development Rates

MCF is a useful metric for tracking an individual's progression through stage i by a measure of "readiness" to molt from 0 (into stage i) and 1 (out of stage i), rather than a set stage duration, particularly when conditions are variable (see section 1.3.4). The *time* it takes from entering stage i ($MCF = 0$) for an individual to attain $MCF = 1$ is the resultant stage duration (i.e., $D_i = t_{molt} - t_0$). However, relationships describing MCFs are unknown and, therefore, determining resultant stage durations is difficult. In an example where individuals each have a different increasing MCF pattern over time (Figure 3.1a), the actual MCF curves are unimportant if all reach $MCF = 1$, from 0, at the same *time*, resulting in a similar stage duration. In this example, the constant blue line (i), representing well understood constant conditions (e.g., in a laboratory setting), may be used as an approximation for all (see below, equation (3.2)).

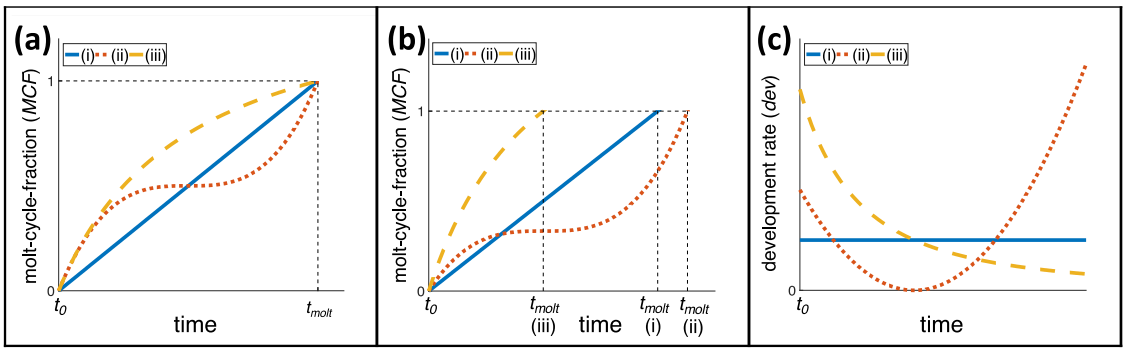


Figure 3.1 The relationship between molt-cycle-fraction (*MCF,* n.d.) and development rate (*dev,* units of time) of stage i. **(a)** Examples of various MCF relationships attaining *MCF* = 1 at the same time, **(b)** Similar examples as (a) but attaining *MCF* = 1 at various times, and **(c)** corresponding development rates to MCF curves in (b), i.e., derivatives.

However, given observed variability in development among individuals *and* dynamic, variable environmental conditions *in situ*, MCF relationships are expected to also be variable and, consequently, result in different resultant stage durations (Figure 3.1b).

To eliminate the confusion between the "actual" stage duration for stage *i* and D_i estimates for incrementing MCF at discrete times over the stage development, (introduced in section 1.3.4), this section offers a relationship between MCF and *development rates*, redefined from known methods involving D_i estimates. Note, the explanations in this section relate to individuals in stage *i,* but the subscript *i* has been dropped from variables *dev* or *D* for simplifying derivations. Conceptually, the rate of increase (i.e., derivative) of *MCF* may be defined as the development rate (*dev*), illustrated in Figure 3.1c.

Consider a laboratory situation, where conditions are held constant, the development rate would most likely also be constant and may resemble relationship (i) illustrated by the blue solid line in Figure 3.1c. For this constant scenario, the corresponding MCF relationship with time (Figure 3.1b) may be found by using a linear regression, assigning the "slope" to *dev(t)* and setting it as constant, and the independent variable intercept as *MCF* = 0:

$$MCF(t) = dev \cdot t \tag{3.1}$$

Therefore, a relationship between stage duration (*D*) and development rate may be found

by knowing that molting occurs when $MCF = 1$; rearranging equation (3.1) for *dev*:

$$dev = 1 / D \tag{3.2}$$

Equation (3.2) is well understood in the literature to describe development rates from stage durations, as empirical relationships for D, based on constant environmental conditions, have been determined (see section 1.3.2).

However, returning to dynamic *in situ* environments, MCF could resemble curves (ii) or (iii) in Figure 3.1b, and use of equations (3.1) and (3.2) as described above are no longer suitable, as *dev* is not constant over stage *i,* and an *approximation* for MCF progression is required. As introduced in section 1.3.4, by considering discrete time-steps, approximations of MCF increments at each time point are found, using the time-step and, as is typically done, the *stage duration* estimates, summing until $MCF = 1$. Here, to incorporate *development rate* estimates, equation (3.2) is rearranged for D and substituted for D_t in the MCF approximation from section 1.3.4 to give:

$$MCF_{t+1} = MCF_t + \Delta t \cdot dev_t \tag{3.3}$$

Equation (3.3) is analogous to approximating integration of development rate (e.g., illustrated using discrete points from example curve (iii) from Figure 3.1c in Figure 3.2a), or using a numerical method called the Euler method for estimating MCF, which "predicts" the next MCF for t+1 using the information known at t (Figure 3.2b).

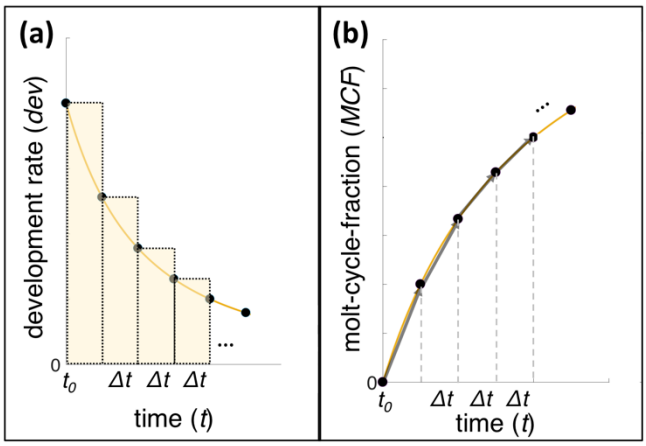


Figure 3.2 Numerical approximations for MCF: **(a)** integration of development rate and, **(b)** "predicting" MCF using Euler method.

Similar to the procedure explained in Gentleman *et al.* (2008; also see section 1.3.4) for adjusting D_i estimates at each time-point to correspond with changing environment, the development rate in equation (3.3) is to be updated at each time-step. This method estimates MCF_{t+1} with the assumption that the updated dev_t for t is held constant over the time interval, Δt. When MCF reaches 1, the duration of stage i is then determined by $D_i = t_{molt} - t_0$, or simply $D = t_{molt}$, if $t_0 = 0$.

The approximation for MCF could exceed 1 due to variable conditions (i.e., molt within Δt; Figure 3.3), and a time correction is required for determining the resultant stage duration and an adjustment for the MCF into stage i+1 (i.e., will not be zero). Implementing a time correction is necessary for simulating population dynamics, as neglecting the "over time" from each individual's stage duration could together add up to a substantial difference, which may have an impact on overall generational development times and conclusions made from simulated population dynamics.

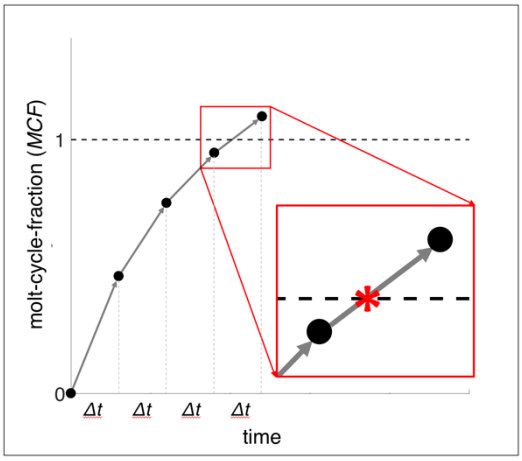


Figure 3.3. Example of molt-cycle-fraction estimate exceeding 1 within a time-step

Due to the assumption of constant *dev* across Δt, a linear interpolation may be used to determine at what point an individual "actually" molted within the time-step. Here, the discrete points (counted by n) outlined with the red box in Figure 3.3 are defined as: (t_{n+1}, MCF_{n+1}) for the new estimate above the MCF = 1 dotted line, (t_n, MCF_n) for the estimate below the line, and (t_{molt}, MCF_{molt}) = (t_{molt}, 1) for the point at molt (red star), and solved for t_{molt}:

$$t_{molt} = t_{n+1} - \frac{MCF_{n+1} - 1}{dev_n} \tag{3.4}$$

Again, if the time into the stage is $t_0 = 0$, then $D = t_{molt}$, otherwise $D = t_{molt} - t_0$. The difference between t_{n+1} and t_{molt} (i.e., the second term subtracted in equation (3.4) is the time spent in the new stage i+1 upon molting within the time-step (i.e., "residual time", Gentleman *et al.,* 2008), and the MCF in stage i+1 is corrected by multiplying this difference by the development rate corresponding to conditions at n for stage i+1.

3.2.2 Modelling Variability in Development Rates

To use these derivations in modelling studies, development rates for each individual at t is required. Inherent variability among individuals is simulated with a fitness metric, by assigning each individual a random value between 0 and 1, i.e., "weakest" to "strongest". Using the procedure introduced in section 1.3.3, the individual fitness values are used to provide individual probability of development rates, but unlike the relationship provided for probability and stage durations, for development rates *probability = fitness*, such that individuals with weaker fitness are more likely to experience slower than average development rates, while those with stronger fitness are more likely develop relatively faster. The probability (i.e., fitness) metrics are paired with a gamma CDF described for each stage i, as was done in Neuheimer *et al.* (2010b) to somewhat preserve the distribution of observed stage durations (e.g., Gentleman *et al.,* 2008) and to avoid negative values. In this chapter, individual *development rates* are directly determined from associated CDFs describing proposed development rates, rather than what has typically been done in the literature by finding individual *stage durations* from CDFs describing predicted stage durations and applying equation (3.3). "Scale" (θ) and "shape" (k) parameters are required to describe a gamma CDF for each stage i and are dependent on a development rate mean ($\mu_{dev,i}$) and variance ($\sigma^2_{dev,i}$), which are based on environmental conditions experienced by all individuals. For each stage i, due to the assumption of constant dev,i across each time-step, the required mean development rate is determined by equation (3.3) (i.e., $\mu_{dev,i} = 1/D_i$, where D_i is found by equation (1.1)). The variance of development rate is determined by the relationship $\sigma^2_{dev,i} = (\text{C.V.} \cdot \mu_{dev,i})^2$, with

a coefficient of variation (C.V.) of 0.30 for all stages, as suggested by Gentleman *et al.* (2012) to be "moderate". Development rates of individuals in stage *i* are then determined by using the assigned individual fitness with the respective CDF described by $\mu_{dev,i}$ and $\sigma^2_{dev,i}$ at each *t* (for more details and illustration of process refer to Appendix C.3.4).

3.2.3 Single Stage Development Tests

A large group of individuals (1×10^6) developing through the same stage were simulated and the individual (i.e., "resultant") stage durations were stored, based on various tests. Each single stage development test was run for a "generic" stage (i.e., no specified life stage) to observe the relative effects of four temperature (*T*) scenarios, with and without inherent individual variability (ind. var.), and later including the effect of mortality (details of each test in Table 3.1).

Table 3.1 Single stage development tests with various temperature scenarios (°C): T_{low} = 2, T_{avg} = 7, T_{high} = 12, and T_{dyn} = t/10+2. Components of simulations not included within a test are marked as "No".

Tests (T.)	Individual variability	Mortality	Temperature Scenarios			
			Constant			Dynamic
			T_{low}	T_{avg}	T_{high}	T_{dyn}
#1: Temperature effect	No	No	T. #1.1	T. #1.2	T. #1.3	T. #1.4
#2: Test #1 with ind. var.	Yes	No	T. #2.1	T. #2.2	T. #2.3	T. #2.4
#3: Test #2 with mortality	Yes	Yes	T. #3.1	No	T. #3.2	No

The temperature scenarios were chosen based on AZMP samples from Chapter 2 that *C. finmarchicus* would likely experience in Newfoundland-Labrador ocean waters: 2 °C for "low", 12 °C for "high", 7 °C for "average", and the "dynamic" scenario was set up as a constant increase from low to high conditions over a period of 100 days. The single stage development tests only considered temperature to demonstrate "environmental" effects on stage durations, as observed ranges of chl-*a* concentration do not have as great of an effect on stage durations as potential ranges in temperature (see section 2.3.1). The same assignment of fitness for individuals was used for each simulation. To analyze the *relative* stage duration differences among results for each test (rows in Table 3.1), the resultant stage durations were normalized by the shortest average stage duration result.

Doing so provides the shortest average stage duration (of all temperature scenarios) as 1, and all other results as relative stage duration values (i.e., "normalized" stage durations). The average of each simulation's set of normalized stage durations was calculated to represent emergent stage durations, along with variance and median for tests involving individual variability (Test #2 and #3).

3.3 Results

3.3.1 Test #1 – Temperature Effect

To demonstrate the effect of various temperature scenarios alone, all other factors were neglected for Test #1, and individuals were set to develop identically (i.e., same development rates were assigned at each point in time; Table 3.1, Figure 3.4).

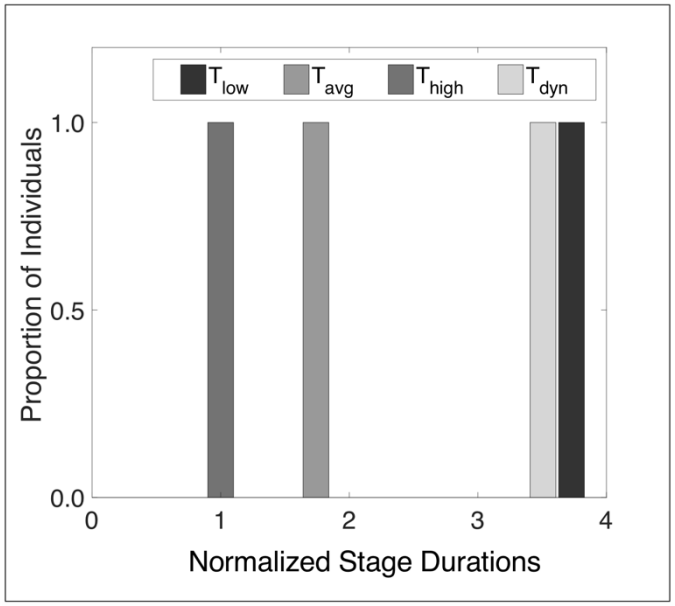


Figure 3.4 Normalized stage duration results from Test #1, normalized by average stage duration result from the T_{high} scenario (Test #1.3).

For each temperature scenario, all individuals resulted in the same stage duration due to having the same development rates, and the emergent stage duration results are as expected (Figure 3.4). The shortest stage durations are a result of individuals experiencing T_{high}, while T_{low} leads to the longest stage durations. The nonlinearity of the formula used for development rates (see section 3.2.2) is evident by the emergent stage duration with T_{avg} not being equal to the average of the emergent stage durations with T_{low}

and T_{high} simulations. The T_{avg} emergent stage duration is about 0.7 times greater than the emergent stage duration from the T_{high} simulation, while about 2.2 times less than the emergent stage duration from the T_{low} simulation. The individuals experiencing a dynamic temperature, T_{dyn} (Test #1.4, Table 3.1) develop through the stage slightly faster than with T_{low}, due to their development rates increasing over the simulation time, corresponding to the increase in temperature.

3.3.2 Test #2 – Individual Variability

Repeating the temperature scenario simulations in Test #1 but including individual variability in development rates (see section 3.2.2, Test #2 in Table 3.1) produced similar emergent stage duration results as found in Test #1 (Figure 3.5): T_{high} simulation results in the shortest emergent stage duration, T_{low} results in the longest, and T_{dyn} emergent stage duration is slightly less than the result with T_{low}.

The resultant stage durations are illustrated using histograms in Figure 3.5, to show the distributions of stage duration results from each temperature scenario (shaded histograms), as well as the relative results among the scenarios (outlined histograms display all results). The differences in stage duration variances among the temperature scenarios are similar to comparisons of emergent stage durations (Figure 3.5). The higher temperature simulation resulted in the shortest individual stage duration variance (0.1), which indicates that all individuals developed through the stage at very similar rates (i.e., with a high temperature fitness does not have as much of an affect; Figure 3.5c). In contrast, the T_{low} simulation leads to a wider spread of stage durations (variance of 1.5), and a longer right-side "tail" of the distribution (i.e., a "positive skew"; Figure 3.5a). The longest stage duration for this T_{low} simulation (i.e., low fitness or "weakest" individual) is about 4.2 times greater than the corresponding emergent stage duration. Similar to results found in Test #1, the T_{dyn} simulation has not only resulted in a smaller emergent stage duration as compared to the T_{low} results, but the individual stage duration variance is also smaller (from 1.5 with T_{low} to 1.1 with T_{dyn}; Figure 3.5a, d).

Here the median stage durations describe a positive skew in resultant stage duration distributions for all temperature scenarios, where each median value is smaller than the respective emergent stage duration.

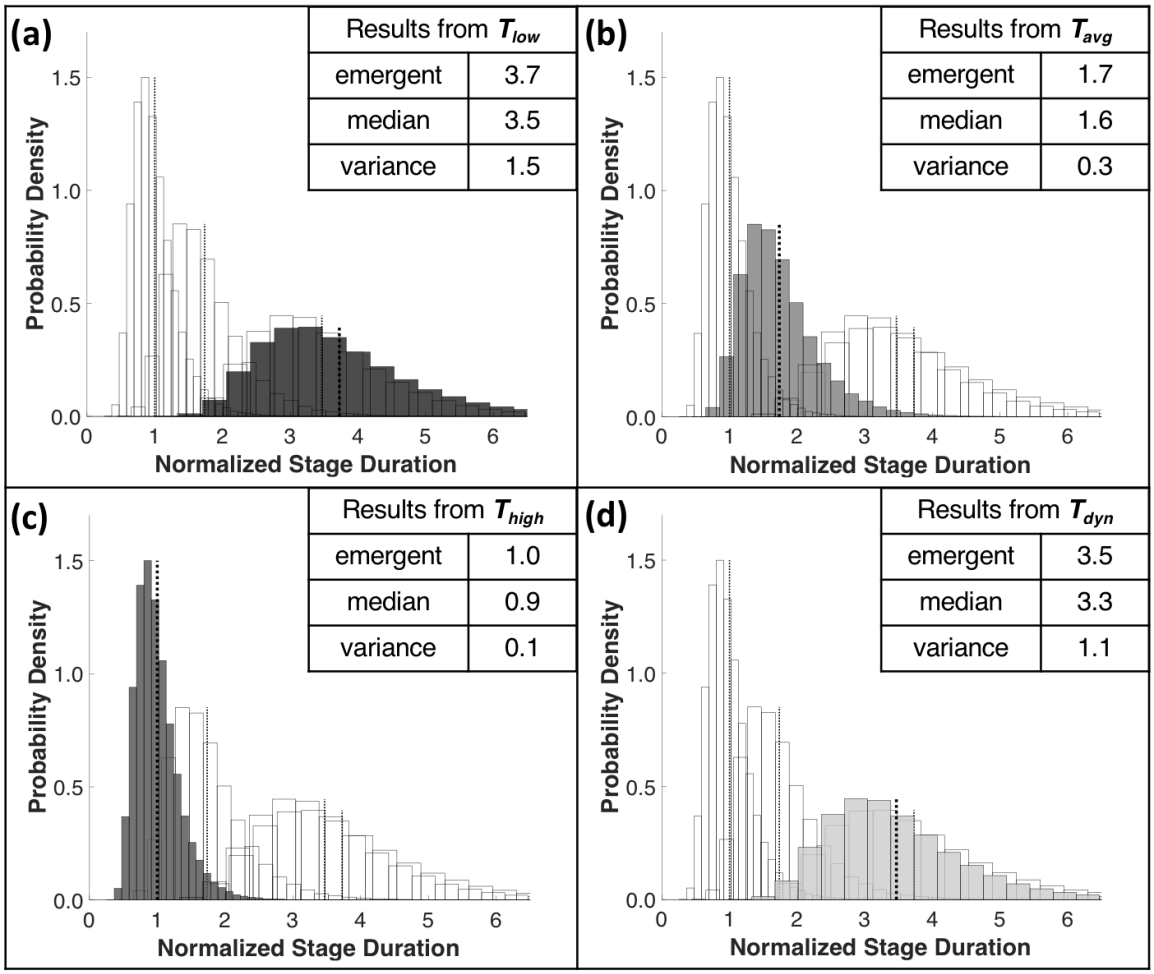


Figure 3.5 Distributions of resultant stage duration results from Test #2, normalized by average stage duration result from T_{high} scenario. Statistics of normalized results provided for: emergent stage durations – dotted black vertical lines, median and variance. Results for each temperature simulation are shaded, while others are provided as outlines for relative comparison: (a) T_{low}, (b) T_{avg}, (c) T_{high}, and (d) T_{dyn}.

3.3.3 Test #3 – Mortality

Repeating T_{low} and T_{high} simulations from Test #2 but with mortality resulted in shorter emergent stage durations and variances (Figure 3.6), as compared to results of Test #2 (Figure 3.5).

The resultant stage durations from less favourable temperature conditions (T_{low}) are more affected by mortality, as the emergent stage durations, median, and variance, all have a

greater decrease (Figure 3.6a, b) than those from T_{high} simulations (Figure 3.6c, d). Due to the killing of individuals with application of mortality, there are less resultant stage durations (i.e., not all survive): survivorship is 68% for the T_{low} simulations (Figure 3.6a, b) and 90% for the T_{high} simulations (Figure 3.6c, d).

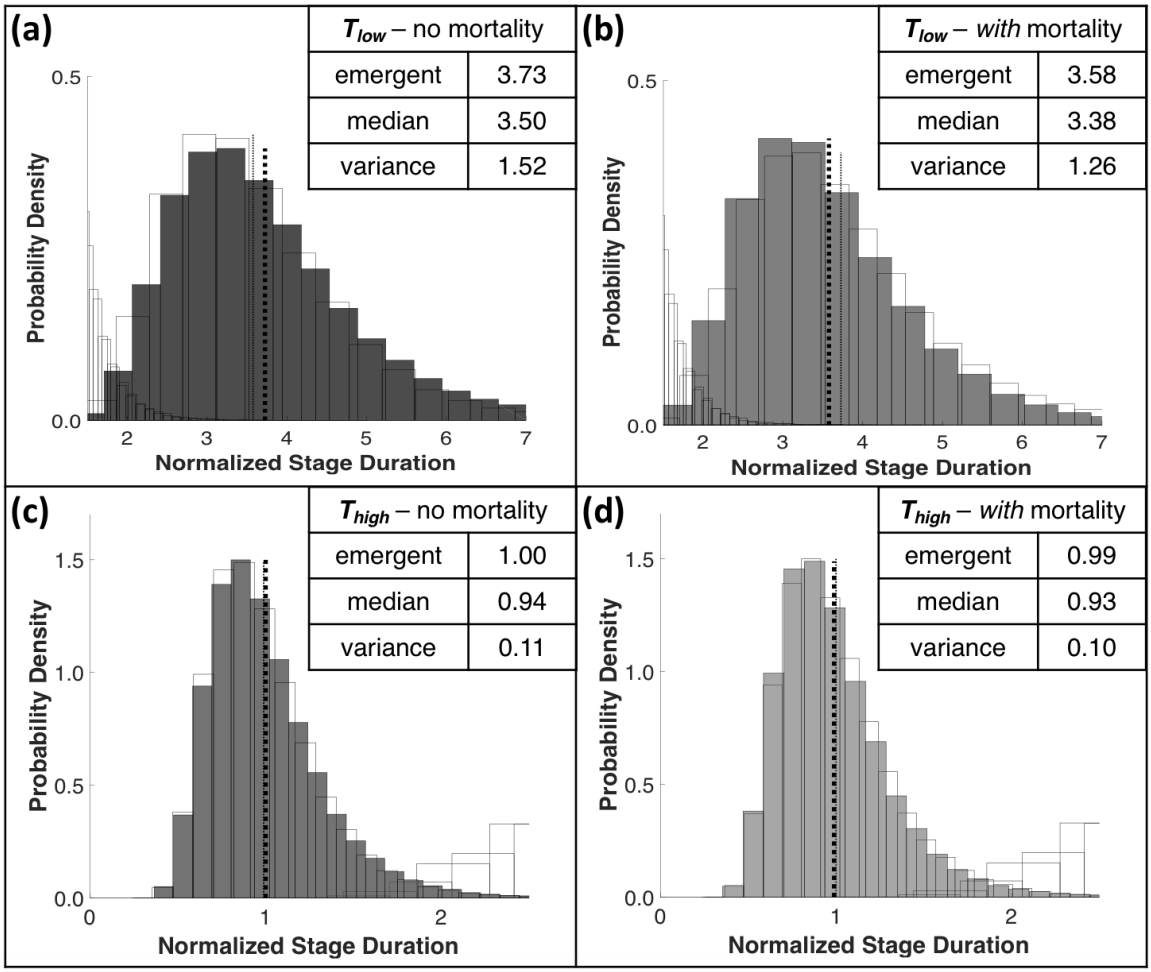


Figure 3.6 Distributions of resultant stage duration results from Test #3, normalized by average stage duration result from T_{high} scenario. Statistics of normalized results provided for: emergent stage durations – dotted black vertical lines, median and variance. Results for each temperature simulation are shaded, while others are provided as outlines for relative comparison: (a) T_{low} without mortality, results from Fig. 3.6a, (b) T_{low} with mortality, S = 68%, (c) T_{high} without mortality, results from Fig. 3.6b, and (d) T_{high} with mortality, S = 90%.

3.4 Discussion

The tests results have shown that variable temperature scenarios, inherent individual variability, and mortality, all have an influence on development rates, and, therefore, on resultant stage durations.

As was expected, based on findings in empirical studies, higher temperature conditions simulated shorter emergent stage durations and smaller variance among resultant stage durations from individuals with variability in development rates. The mortality test results suggest that higher temperatures are more favourable than lower, due to shorter stage durations and, therefore, greater survivorship.

The smaller emergent stage durations results with dynamic temperature simulation with individual variability (i.e., Test #2.4), as compared to T_{low} simulations (i.e., Test #2.1), suggests that using empirical relationships may overestimate stage durations from winter into summer seasons (i.e., when temperatures are increasing), when the first *C. finmarchicus* generations are developing. Overestimating stage durations at this time may result in an additive effect through the life stages and, thus, overestimate aggregate development times, such as larval duration, or diapause timing, providing false interpretation about the timing, and potentially location, of individuals at these stages. The opposite effect may occur with underestimating stage durations from summer into winter (i.e., decrease in temperatures), but this time of year is less of a concern as individuals are expected to have reached later stages and are entering, or have already entered, diapause.

The influence of mortality (i.e., Test #3) on resultant stage durations provides support for the "survival of the fittest" hypothesis. For both temperatures simulations, the emergent stage durations, median values, and variance, all decreased as compared to results without mortality, suggesting that those individuals that are surviving through the stage are those that have a higher "fitness" and able to optimize environmental conditions to develop through stages faster, while those that are "weaker" are subject to higher mortality risk, and may only survive by chance.

Chapter 4: Simulating Spatial and Temporal Population Dynamics

4.1 General

The objective of this chapter is to simulate the population dynamics of *C. finmarchicus* in the Newfoundland-Labrador waters involving inherent variability among individuals, mortality, and temporally and spatially dynamic environments. It is expected that these involved factors will affect stage durations of individuals, as was suggested with results of Chapter 3, and that population dynamics will be influenced. Building a suitable IBM to simulate *C. finmarchicus* population dynamics was a major component of the research presented in this chapter. An existing IBM was chosen to build upon and the enhancements and modifications made to this base IBM, along with any additional components, are described. The modified IBM version is coupled to physical ocean model data, to gain information about the influence of environmental conditions changing in both time and space. The physical model is a 3D circulation model for the Newfoundland-Labrador ocean region (Han *et al.,* 2008), from which particle tracks and associated temperatures were derived from and used for this study. Chl-*a* data was obtained from an online database, "GlobColour" (ACRI-ST, 2017) to create temporal and spatial food availability corresponding with the provided particle tracks. Temperature and chl-*a,* as well as simulated stage development solely dependent on environment, are characterized in the results of this chapter and illustrated using spatial maps provided in Appendix C. The final analysis of this chapter provides a case study that simulates the development of *C. finmarchicus* through the first annual generation, experiencing environmental conditions due to transport described by two chosen particle tracks. The two particle tracks have been chosen based on their similar southward movement, but yet very different temperature patterns experienced over the transport. The differences in population stage and structure simulation results among the two particle tracking scenarios are discussed in the results, and the implications of the population dynamics results on predation are offered in the discussion.

4.2 Methods

4.2.1 Individual-Based Model (IBM)

4.2.1.1 Design Criteria and Desirable Characteristics for Model

Analyzing the sensitivity of *C. finmarchicus* population dynamics to variability in environment and mortality are not possible with *in situ* observations, and, therefore, a mathematical model is required. The model is to simulate a copepod population and its development though stages over time, dependent upon the environmental conditions and mortality experienced. The specific life processes should be calibrated to reflect those of *C. finmarchicus* species, and individual variability in development rates is required to simulate realism, as variability in development among individuals has been observed in laboratory (section 1.3.3). To provide a sense of spatial population dynamics, the model must have the capability to be coupled to particle tracking information provided by a physical ocean model (section 4.2.2). Specifically, for this study, the model must have the capability of individual environmental conditions, or "particle track" conditions, such that individuals may be assigned to different tracks. Lastly, a record of each individuals' time spent in each stage is required, such that emergent stage durations may be analyzed post-simulation.

4.2.1.2 Characterization of Existing IBM

An IBM describing the life history of *C. finmarchicus* was previously built (Gentleman *et al.,* 2008; Neuheimer *et al.,* 2010a) and applied to studies regarding the Northwest Atlantic region (e.g., Neuheimer *et al.,* 2009, 2010b). Their IBM simulates individual life histories and the results regarding individuals may be combined to provide analysis on a population level. Their IBM includes inherent individual variability through an individual fitness metric (see section 3.2.2), by which other life history processes (e.g., mortality, stage duration) may be characterized probabilistically. Based on the capabilities of their IBM coinciding with many model requirements for this study, it was chosen as an appropriate "base" model to build upon and modify as necessary to suit the objectives of this chapter.

The basic flow of operations and main components involved in the base IBM are shown in Figure 4.1 and described below, as many components were maintained for the modified IBM (modifications described in section *4.1.1.3*). For more details regarding the base IBM, please refer to Neuheimer *et al.* (2010a). The base IBM simulates *C. finmarchicus* population dynamics by assigning a matrix to represent a population of individuals that undergo life processes, i.e., are set to spawn, develop, and die, where the elements of the population matrix are updated as necessary (e.g., stages and associated properties), simulating progression over time.

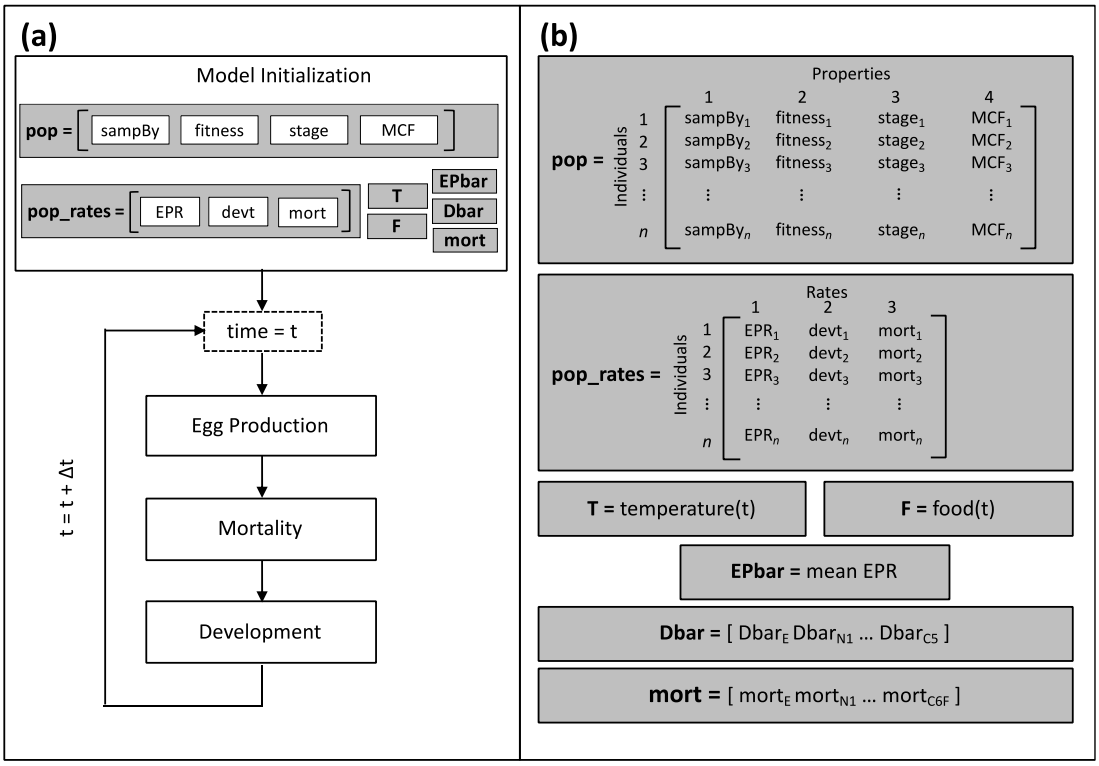


Figure 4.1 Conceptual flow chart of base IBM demonstrating (a) general structure of main code components, and (b) structure of main matrices: population, rates (individual and mean, i.e., "bar"), and environmental conditions (temperature and food). Note: n = total number of individuals.

Model Initialization: The model first initializes the simulation parameters, environmental conditions, and assigns values to variables representing average rates (i.e., EP and stage-specific mortality) and average stage durations based an empirical relationship which is dependent upon the environmental conditions. The model then initializes main population related matrices: a population matrix (Figure 4.1a), where rows are set to represent *living* individuals and columns represent individual properties or metrics (e.g., stage identity,

fitness level), and a separate corresponding matrix for individual rates (i.e., EP, development, and mortality; Figure 4.1b). These individual rates are determined from average rates (EP and mortality) and average stage durations (as described in section 3.2.1 with equation (3.2)) in combination with a chosen distribution and assigned individual fitness (see section C.3.4).

Time-stepping loop: The main code then performs iterative calculations of life history processes, at set time-steps for a prescribed simulation length. During each time-step, the population and rate matrices are run through the model components that simulate life processes, where the contents of the matrices are updated as necessary. At the end of each time-step, all conditions and average variables (i.e., rates and stage durations, based on conditions) are updated for the new time for the subsequent run. The basic operations of life history processes involved within the time-stepping loop are as follows (in order of operation):

- *Egg production:* This step simulates the production of new "eggs", based on the current number of females in the population and a prescribed egg production rate. The EP process has the option of "super-individuals", i.e., multiple eggs may be represented by a single matrix row, which is expressed through a "sample size" metric (a column in the population matrix). Once the matrix size is determined from the new egg abundance and sample size, the new egg matrix and corresponding rates matrix are created, and individual properties, metrics, and rates are determined, and assigned to matrix elements accordingly. The new egg matrix is added to the existing population matrix before the simulation continues.

- *Mortality:* The population matrix then "experiences" mortality, meaning that each individual is subject to a calculated mortality risk (introduced in section 1.3.5), where individuals are then chosen by chance to be "killed" (details in Appendix C) and are removed from the population by eliminating their corresponding rows in the population and rates matrices.

- *Development:* The surviving individuals are then advanced through their life stages by using MCFs, which are determined using the set time-step and individual development

rates (as described in section 3.2.1 determining *stage durations* from the assigned distribution and using equation (3.2)).

4.2.1.3 Description of IBM Modifications and Enhancements

Major changes made to the base IBM are detailed in Table 4.1, by comparing the capability and functionality of the base IBM to the modified IBM version used in this chapter, main topics include:

- fixing incorrect characterization of life history details,
- allowing for variability in environmental conditions among individuals (i.e., suitable for coupling to particle tracking information), and
- calculating and storing information regarding individuals' resultant stage durations.

Based on the modifications and enhancements, Figure 4.1 has been updated for the modified IBM version, outlining the general structure of components and main matrices in Figure 4.2. Full details regarding the modified IBM's components and coding, as well as examples of various IBM applications, have been documented in the form of a user's manual, provided in Appendix C.

Table 4.1 Major changes made to the base IBM and new additional components to create the modified IBM version for Chapter 4 objectives. Some examples ("e.g.,") are provided in MATLAB code.

Design Criteria/ Desirable Characteristic(s)	**Base IBM**	**Modified IBM**
Modularization and Clarity		
Main script: Simplify code for clarity	Main script contains a combination of sequential MATLAB commands and function calls.	Main script only calls functions; all sequential MATLAB commands were moved to new functions.
All model components: Produce identifiers (ID*)* for clarity of code and flexibility of model application	Matrix entries and life stages are assigned/called based on the indexing position of elements and corresponding values, respectively. (e.g., Assigning a *range* of individuals as females, with stages represented in the 1st column of the population matrix (*pop*) and females considered as the 13th stage: *pop*(*range*, 1) = 13)	Parameters, as IDs, were produced to represent matrix element positions and life stages and grouped using structures. (e.g., See left example; now defining structures *col* = matrix column positions and *stg* = life stage values, and IDs as *col.stage* = 1 and *stg.fem* = 13, reassigning as: *pop*(*range*, *col.stage*) = *stg.fem*)
	Different simulation objectives require code in all components to be updated. (e.g., See above example; changing stages to be represented in the 2nd column of *pop* requires all assigning/calling of stages to be updated: reassigning *pop*(*range*, 1) to *pop*(*range*, 2))	Parameter values are easily modified to adapt to various research needs, i.e., no need to update model contents, just ID values in model initialization step. (e.g., See left example; to move stages to the 2nd column of *pop* requires ID *col.stages* to be changed: reassigning *col.stages* = 1 to *col.stages* = 2)
All model components: Simplify main matrices for clarity of code and model application	Sample size (*sampBy*) property available in *pop* to allow for "super-individuals" (i.e., a row may represent multiple individuals), to improve run efficiency with larger population size.	Sample size property was removed (i.e., each row represents one individual), as preliminary testing determined population size would not be of considerable size to significantly decrease efficiency.

Design Criteria/ Desirable Characteristic(s)	Base IBM	Modified IBM
Modularization and Clarity (continued)		
All model components: Simplify main matrices for clarity of code and model application (continued)	Separate matrices are assigned with mean rate values (some also with standard deviation), and individual rates assigned/called using a population rates matrix (corresponding to individuals in *pop*).	No matrices assigned for mean rate values (calculated as necessary, see "New Simulations" category below), and individual rates are now stored in *pop* as individual properties (eliminated separate rates matrix).
All model components: Redefine calculations with *D* to use *dev* (section 3.2.1-2)	Formulae throughout model use assigned individual stage durations (with equation (3.3) whenever necessary. (e.g., Updating an individual's *MCF* uses their calculated *D* and time-step (Δt), such that the increment is: $\Delta MCF = \Delta t / D$)	Redefined formulae throughout model to use individual development rates whenever necessary; this is a conceptual change to promote prescribed development *rates* and stage durations as emergent properties. (e.g., See left example; updating an individual's *MCF* now uses their calculated *dev*, such that the increment is now: $\Delta MCF = \Delta t \cdot dev$)
All model components: Model documentation	Manuscript is available describing 2010 version of model and provides example of application; see Neuheimer *et al*., 2010a	User's manual has been developed to overview the current modified version, providing examples of model application (available in Appendix C).
Fix Incorrect Characterization of Life History Details		
Subroutines (development and mortality): Fix issues with super-individuals (i.e., *sampBy* property)	The following problems arise with *sampBy* > 1: • development to adult: issue with assigning males and females from molting C5s as 1:1 sex ratio • mortality: issue with rows determined as "dead" – question of remove entire row (i.e., multiple individuals) or select individuals and somehow reassign *sampBy*.	Fixed the issues listed (see left) with the elimination of the *rep* property (see "Simplify main matrices…" row above), i.e., each row represents one individual. Removed all incorrect code in subroutines attempting to handle *sampBy* > 1.

Design Criteria/ Desirable Characteristic(s)	Base IBM	Modified IBM
Improve Efficiency of Life History Processes		
*Subroutines (*new *molt info., development, and mortality):* Identify individual molt information prior to calling subroutines	Subroutines perform calculations to: update MCF and determine which individuals are ready to molt, whenever these metrics are necessary. (e.g., Development and mortality subroutines require the distinction between molters and non-molters, and updated MCF information; repeating calculations).	Produced new function to predetermine individuals' molt information regarding: • updated MCF • molting status (yes or no) • updated age-within-stage and store as a structure to use throughout code, reducing repetition of calculations among subroutines.
Subroutine (new eggs): Simplify overall egg production process	Egg production subroutine contains many subtle, but complex, life history factors for the new "eggs" to undergo development and mortality processes in the time-step. Within these subroutines, eggs require separate treatment from the rest of the *pop* and, therefore, calculations are repeated.	New egg "production" subroutine outputs matrix representing new egg *survivors* with properties. The new "eggs" do not undergo development and mortality during this time-step (i.e., time of spawn), and therefore, all properties are explicitly defined within the subroutine.
	Egg production rate is a property in the population rates matrix; value of zero for all non-female individuals.	Egg production rate was removed from population matrices (as it only relates to females during EP and may be misleading).
New Simulations (not possible with base IBM)		
All model components: Provide capability for individual variability in environmental conditions	Assigned environmental conditions (e.g., constant or dynamic over time) are experienced by the whole population. All average rates dependent upon environment are calculated using assigned conditions. (e.g., Individuals experience the same average rates, but individual rates are possible through the use of properties representing inherent individual variability, i.e., fitness levels.)	Added the capability for individual variability in environmental conditions, by storing as properties in *pop*, as well as introducing an individual "conditions ID*"* that assigns these conditions. (e.g., Simulations with particle tracking information are possible, where individuals may be assigned to separate paths and, therefore, experience different conditions for average rates; meanwhile inherent individual variability is possible with fitness levels.)

Design Criteria/ Desirable Characteristic(s)	Base IBM	Modified IBM
New Simulations (not possible with base IBM; continued)		
All model components: Modify code to account for individual variability in environmental conditions	Not applicable: individual variability in environmental conditions not available.	The following components were modified: • environmental forcings; determining individual conditions based on preset condition ID (also produced capability for various data types, e.g., constant or particle tracking, and extraction methods, e.g., interpolation) • mortality rates (based on individual conditions) development rates (based on individual conditions)
	The option for individual variability in rates is available. These are calculated as needed, by using stored general mean rate or stage duration values based on general environmental conditions, individuals' fitness levels and a distribution for sampling (section 3.2.2).	The option for individual variability in rates is available. These are calculated and stored in *pop* as properties, by using individual mean rate values based on individual environmental conditions, individuals' fitness levels and a distribution for sampling (section 3.2.2).

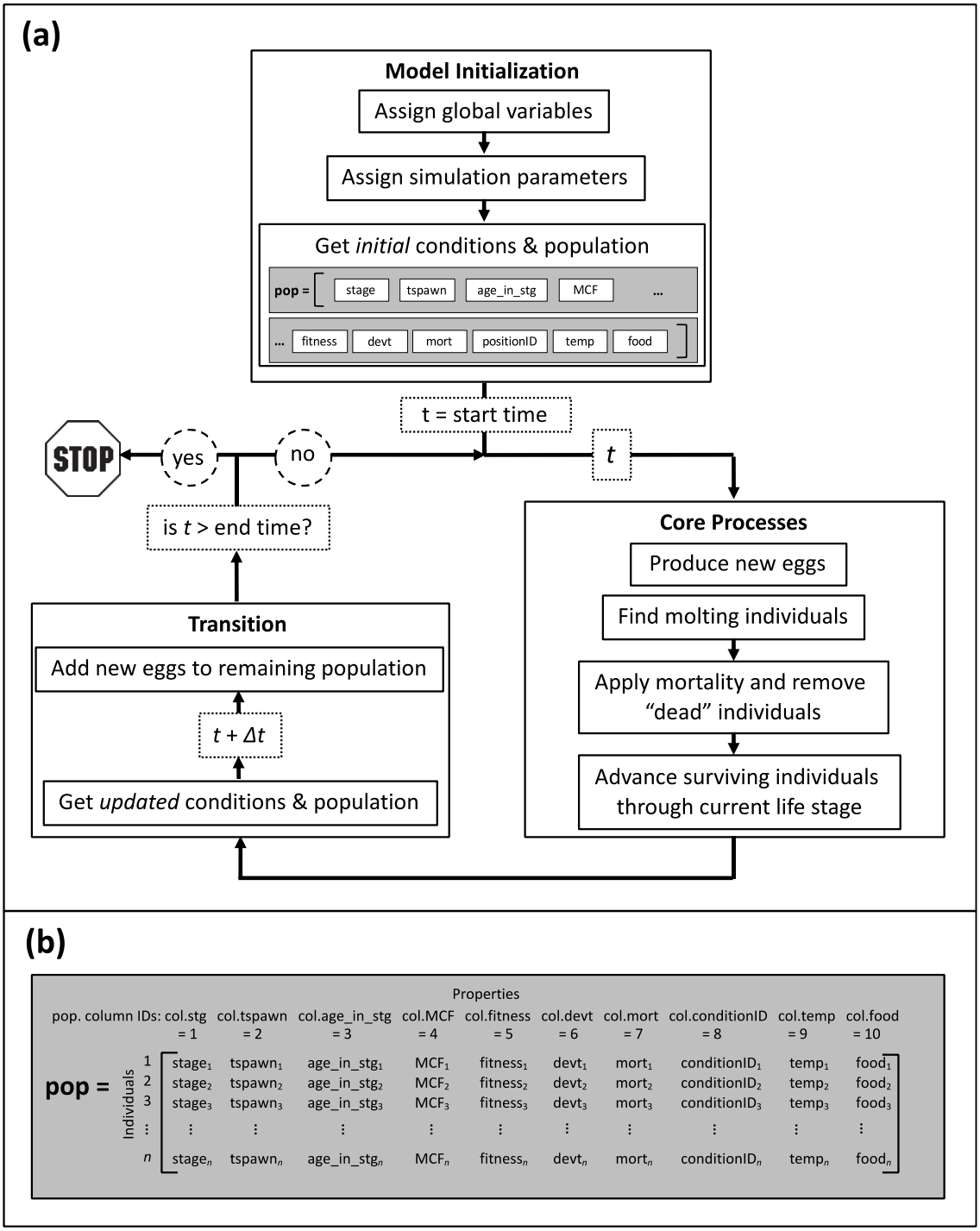


Figure 4.2 Conceptual flow chart of the modified IBM demonstrating: (a) general main code structure, and (b) main population matrix structure.

4.2.2 Physical and Environmental Data Products

4.2.2.1 Particle Tracking Information

The modified IBM is coupled to physical information derived from a 3D particle tracking model by Han *et al.* (2008; introduced in section 1.5.2). This particular physical ocean model appropriately meets the objectives of this chapter and was used by Pepin *et al.* (2013) for a similar analysis on *C. finmarchicus* in the Newfoundland-Labrador ocean region. The output data provided by Han *et al.*'s (2008) model consisted of longitudinal and latitudinal coordinates describing the particle tracks, with corresponding temperatures along the "tracks", representing climatological averages for 2000-2009 (spatial maps in Figure D.1). The model simulates a "release" of particles from the Labrador Shelf region for an April 1st start date (Figure 4.3a), tracking the particles daily for 180 days, at a constant water depth of 1 km (Figure 4.3). Many of the original provided particles had appeared to simulate a "collision" with land boundaries (i.e., hit the Newfoundland-Labrador coast and stopped moving), which would not provide a sense of realistic copepod transport. Therefore, the particle tracking data was preprocessed to only contain data of particles that remained mobile for the 180 days.

*4.2.2.2 Chlorophyll-*a *Data Retrieval*

To incorporate food availability along with the provided temperature, data was retrieved from the online database, the "GlobColour" project (ACRI-ST, 2017), due to the direct availability of chl-*a* concentrations (i.e., already converted from satellite "ocean colour") at spatial. Temporal and spatial ranges for chl-*a* data retrieval were chosen to correspond with the provided temperature data (section *4.1.1.4*). It is noted that although climatological average temperatures represent 2000-2009, averages for chl-*a* were calculated for 2003-2009, as there are no complete records prior to 2003 for the Newfoundland-Labrador ocean region.

The monthly averages of chl-*a* data retrieved represent mid-month conditions and are provided for specific latitude and longitude coordinates. In order to obtain chl-*a* data at spatial grids corresponding to the provided particle tracks, the chl-*a* data was further processed by applying 3D (i.e., 2D spatial grids and time) linear interpolation to obtain

daily chl-*a* estimates at matching spatial coordinates (complete chl-*a* maps provided in Figure D.2).

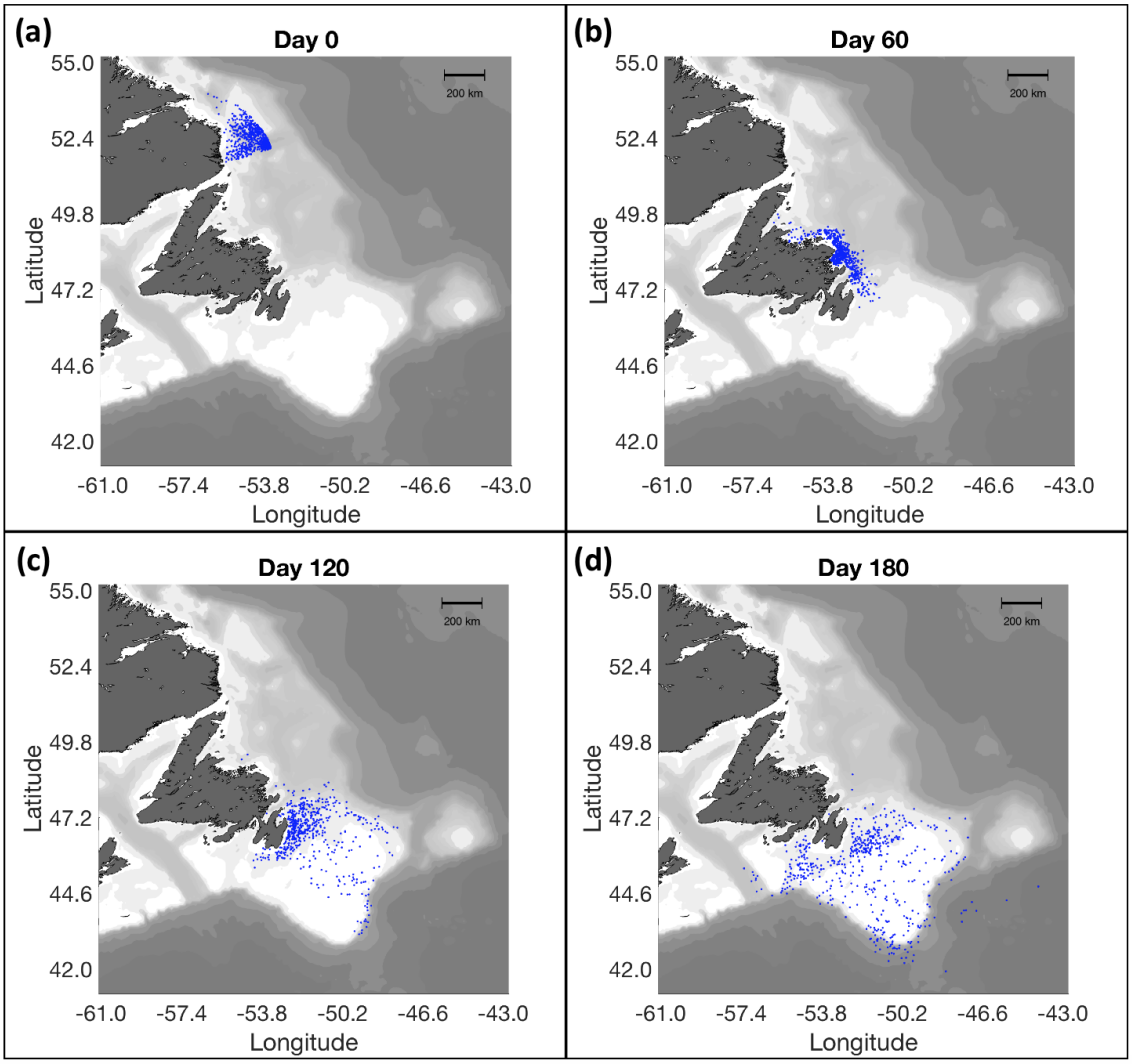


Figure 4.3 Simulated particle tracks (Day 0 = April 1st run for 180 days) derived from Han et al.'s (2008) 3D physical model. Only particles that remain mobile are displayed (554 particles). Deeper water depth is marked with darker gray scale contour.

4.2.3 Simulation Conditions for Population Dynamics Case Study

A case study is completed to investigate the influence of mortality and environmental variability on a single generation of *C. finmarchicus* developing over time, by coupling the IBM to various particle tracks.

Prior to including inherent individual variability and mortality, the temperature and chl-*a* particle track data are used to simulate the development of one individual per particle

over 180 days (development rates assigned as described in section 3.2.1). This simulation provides population dynamics results solely based on environmental conditions, useful for comparing to results of the case study by providing information on how inclusion of individual variability and/or mortality may influence population size and structure. The simulation is initialized with an egg staged individual (i.e., representing multiple identical eggs) on each particle at Day 0 and develops through life stages based on the environmental conditions experienced by that track.

The IBM is initialized with an approximate egg abundance (#/m^2) on the Labrador Shelf according to Head *et al.* (2015, Fig. 3c), representing the first annual generation of eggs spawned directly, bypassing simulation of egg production by females as this life process was not a focus of this research. Individual development rates were assigned as described in section 3.2.2: uniformly distributed random fitness metric (between 0 and 1) was assigned to represent inherent individual variability, determining individual rates using a gamma CDF described by associated parameters (section C.3.4). The onset of diapause is not simulated to provide a sense of generational development time (i.e., the population does not stop developing at C5s but continues to develop to adults). Mortality rates are assigned using average egg (m_E = 0.42 d^{-1}) and nauplii ($m_{N1\text{-}N6}$ = 0.10 d^{-1}) estimates from Head *et al.* (2015, Table 6) and overall averages for copepodites ($m_{C1\text{-}C6}$ = 0.05 d^{-1}) from estimates found in Chapter 2. Stage-specific abundances were stored at each time-step for analysis of simulated population size and stage structure over time. Resultant stage durations of individuals (i.e., survivors of each stage) were stored and stage averages were calculated, post-simulation, to represent emergent stage durations.

Two identically initialized populations were each assigned a physical scenario (i.e., particle track) to investigate the influence of environment on the developing populations, and to provide a demonstration of the IBM's capability to be coupled with physical information. For these physical scenarios, two particle tracks were chosen, termed "Particle Track A" (PT A) and "Particle Track B" (PT B; Figure 4.4), based on having similar transport patterns, but yet experience very different temperatures (Figure 4.4b). PT A and PT B experience similar southward movement, arriving near station 27, which

is a continuous sampling site off of Newfoundland and point of interest used in Pepin *et al.,* (2013; area shown by red outlined rectangle Figure 4.4a, c).

The model runs for time-steps of 0.25 days for the length of the available environmental data, 180 days. Due to the time-step being less than 1 day, the temperature and chl-*a* require interpolation between daily data points and are chosen to remain constant over each day for simulation efficiency, since the environmental conditions are not expected to be greatly variable within each day.

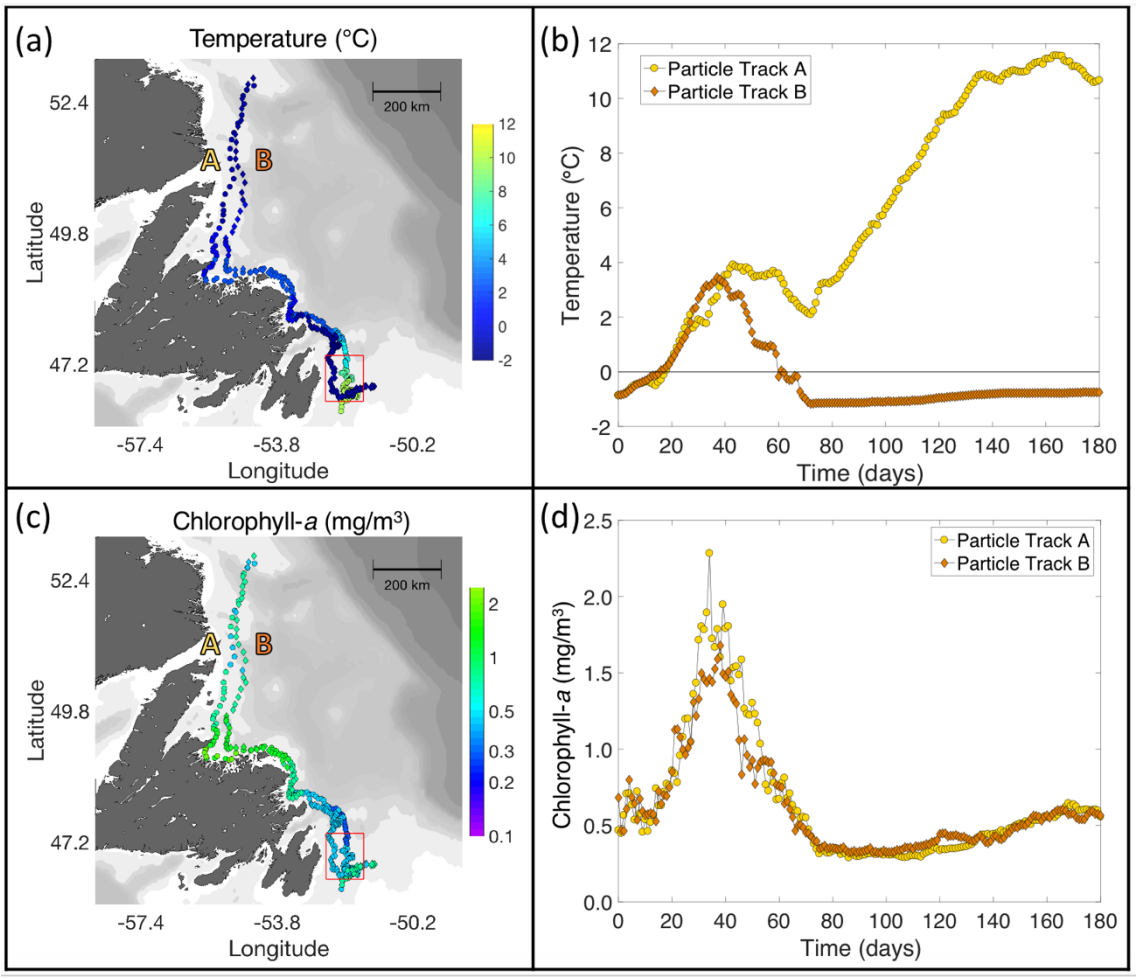


Figure 4.4 Data of particle tracks A (yellow circles,) and B (orange triangles) over 180 days, Day 0 = April 1st. Temperature and chlorophyll-*a* data are displayed spatially (a, c) and temporally (b, d).

4.3 Results

4.3.1 Data Characterization

4.3.1.1 Particle Tracks and Temperature

The movement of all particles in Figure 4.3 generally agree with particle tracks given by Pepin *et al.* (2013, Fig. 8), as the particles drift along the coastline, under the influence of the coastal and shelf-edge branches of the Labrador Current, arriving just above the Grand Banks at about Day 90 (end of June, Figure 4.3b). At about Day 150 (Figure 4.3c), majority of the particles are on the banks and no longer drift further southward, but disperse about the bank area, an effect of the encounter of the southwestern flowing Labrador Current with the northeastern flowing Gulf Stream.

Temperatures from particle tracks are characterized in this section, but illustrated with spatial maps in Appendix D. Sea surface temperature (SST) cycles, based on 2-week averages of satellite observations, from Pepin *et al.* (2013, Fig. 3), are used here to validate the temperature data provided by the physical ocean model, as the SST represents the same decade and region and are considered as representative of the temperature range which *C. finmarchicus* may have experienced during transport southward to station 27 (Pepin *et al.,* 2013). In the initial position of the particles, the temperatures range from -0.91 °C to -0.66 °C (Figure D.1a), coinciding with the range of SST for about April 1st in Pepin *et al.* (2013, Fig. 3 at "Day of year" = 90). After 150 days (early September), the majority of the particles have drifted to the Grand Banks area and over this time period experience increasing temperature conditions (at this point ranging from -0.92 °C to 17.42 °C; Figure D.1f). This increase in temperature conditions is expected, not only due to the southward movement into warmer waters, but also due to the summer season reaching the end. The temperature range is somewhat maintained over the remaining 30 days of data (-0.84 °C to 17.91 °C), due to the halt in southward drift. The particles flowing over deeper waters (i.e., darker contour shades; Figure D.1) appear to experience relatively colder temperatures than those over the shelves (i.e., lighter contour shades) over the 180-day transition, which is consistent with other findings in the literature for on-shelf conditions (Pepin *et al.,* 2013).

The temperature data for Particle Track A agrees with the general findings of the other particles, increases over time as it flows southward, and the seasons enter warmer months (Figure 4.4b). There is a change in the trend for PT A temperatures between Day 40 (mid-May) and 70 (mid-June), where the temperature slightly drops by about 2 ºC, but continues to increase after Day 70. Particle Track B was chosen as a contrast to PT A, where PT B temperatures experience a larger drop of about 4 ºC around Day 40 and does not recover above 0 ºC.

4.3.1.2 Chlorophyll-a

Similarly done with the temperature data, chlorophyll-*a* concentrations along the tracks are characterized here, and spatial maps provided in Appendix D. The temporal trends for overall chl-*a* tracks (Figure D.2) are expected, showing signs of an increase at Day 30 (early May, Figure D.2b), indicative of a spring bloom, and even perhaps a second smaller autumn bloom (i.e., Day 180, late September, Fig. D.2g), where concentrations appear to increase. Some signs of concerning low levels of chl-*a* for *C. finmarchicus* (i.e., chl-*a* < 0.34 mg/m^3) appear at Day 90 (late June) for particles over shallow waters (i.e., deep blue markers over lighter gray coloured contour, Fig. D.2d). The low levels of chl-*a* are observed for the remaining time, although by Day 180 are experienced by fewer particles that are farthest off of the Grand Banks (i.e., deeper waters).

Chl-*a* data for Particle Track A and B are very similar (Figure 4.4d), and follow the general patterns observed with all of the tracks. Both PT A and B chl-*a* data represent the spring bloom with a notable peak at about Day 40 (mid-May), and a small, gradual increase after Day 70 to 180. Both PT chl-*a* data experience a slight drop below 0.34 mg/m^3 between about Day 80 and 120 but is not expected to have a great effect on population dynamics simulations.

4.3.2 Case Study: Simulating *C. finmarchicus* Population Dynamics

The simulated population dynamics results by only experiencing particle track environmental conditions are provided as spatial maps in Appendix D. Results show that most of the population has developed through naupliar stages by Day 60 (Figure D.3e), and later stages by about Day 120 (Figure D.3i), suggesting that these are approximate

larval and generational times, respectively, and are used for comparison with the following case study results (see section 4.4).

Results of simulated population dynamics experiencing either PT A or PT B are shown by stage abundances in Figure 4.5 and stage structure in Figure 4.6, in periods of Day 0-30, 30-70, and 70-160, based on notable findings (full figures available in Appendix E).

The environmental conditions from both tracks provide nearly the same development of *C. finmarchicus* population size and structure up to Day 30 (Figure 4.5a and Figure 4.6a, respectively), where a difference in the structures of the two populations is evident after Day 50 (Figure 4.5b and Figure 4.6b). At about Day 70, almost half of the individuals experiencing PT A conditions have developed to stage C3 (Figure 4.5b), while the majority of the individuals experiencing PT B conditions are only at C2 and even signs of late naupliar stages (N6; Figure 4.6b). After 70 days, there appears to be a lag in stage development by almost two stages for the PT B population as compared to the PT A population. By the end of the simulation, the PT A population shows signs of development through all stages, resulting in emergent larval duration of 50 days, average time to C5 of 105 days, and generational time of 112 days, while the PT B population resulted in a similar emergent larval duration of 51 days, but individuals did not develop past C4 (average time to C4 of 133 days), and most only made it to C3 (all emergent stage duration results are found in Table E.1).

Size in both PT populations greatly decline to low abundances after about Day 70 (approx. 100 individuals), and while the PT B population "died" off around Day 145, a few individuals survive for the entire 180-day simulation with PT A conditions. The overall population sizes are similar among the two PTs, decreasing over time with a more rapid decrease early in the simulation with younger stages.

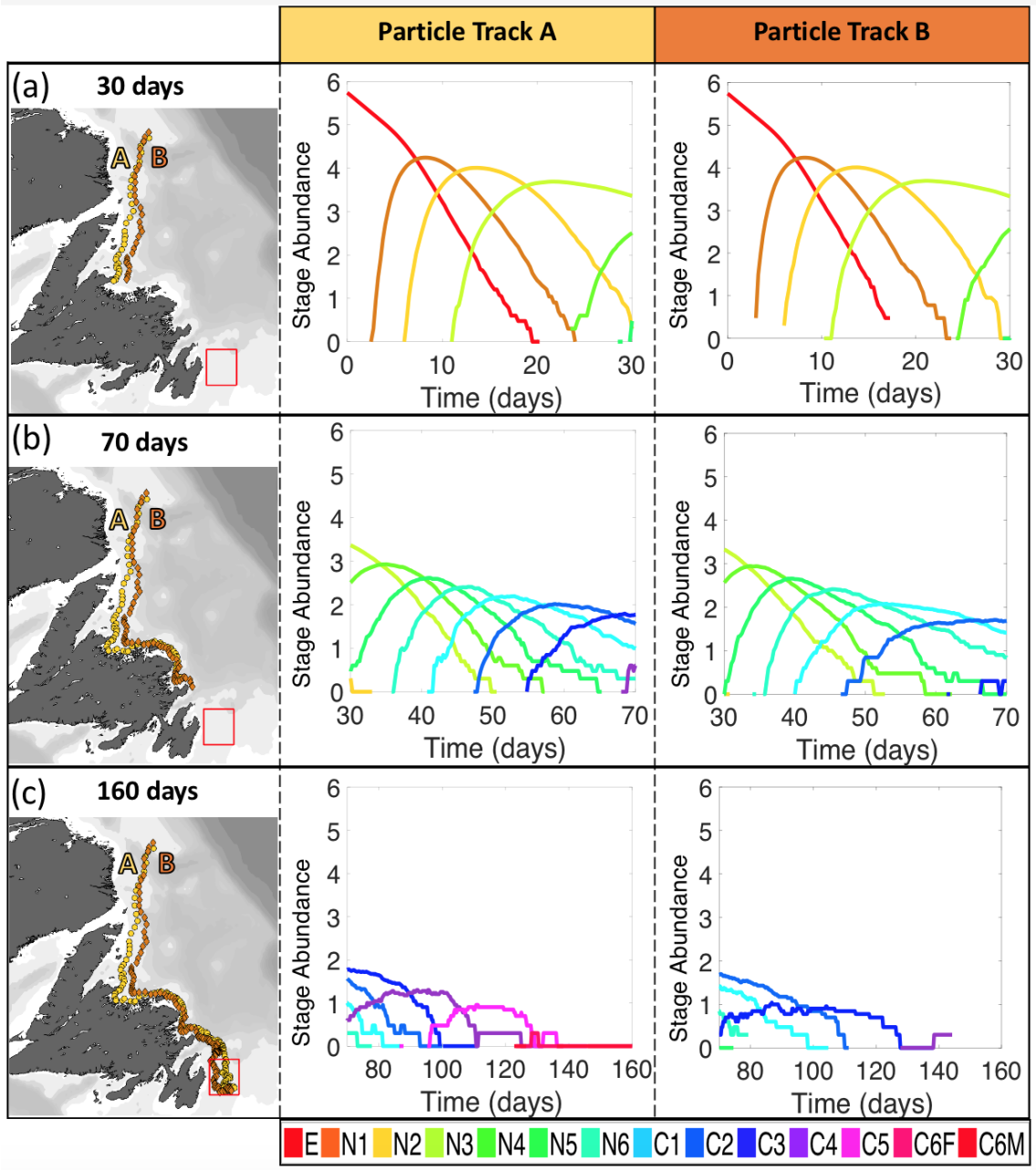


Figure 4.5 Transport of particle tracks A and B (1st column) and associated simulated stage abundances over 180 days. Abundances (# m^{-2}) are plotted on a logarithmic scale.

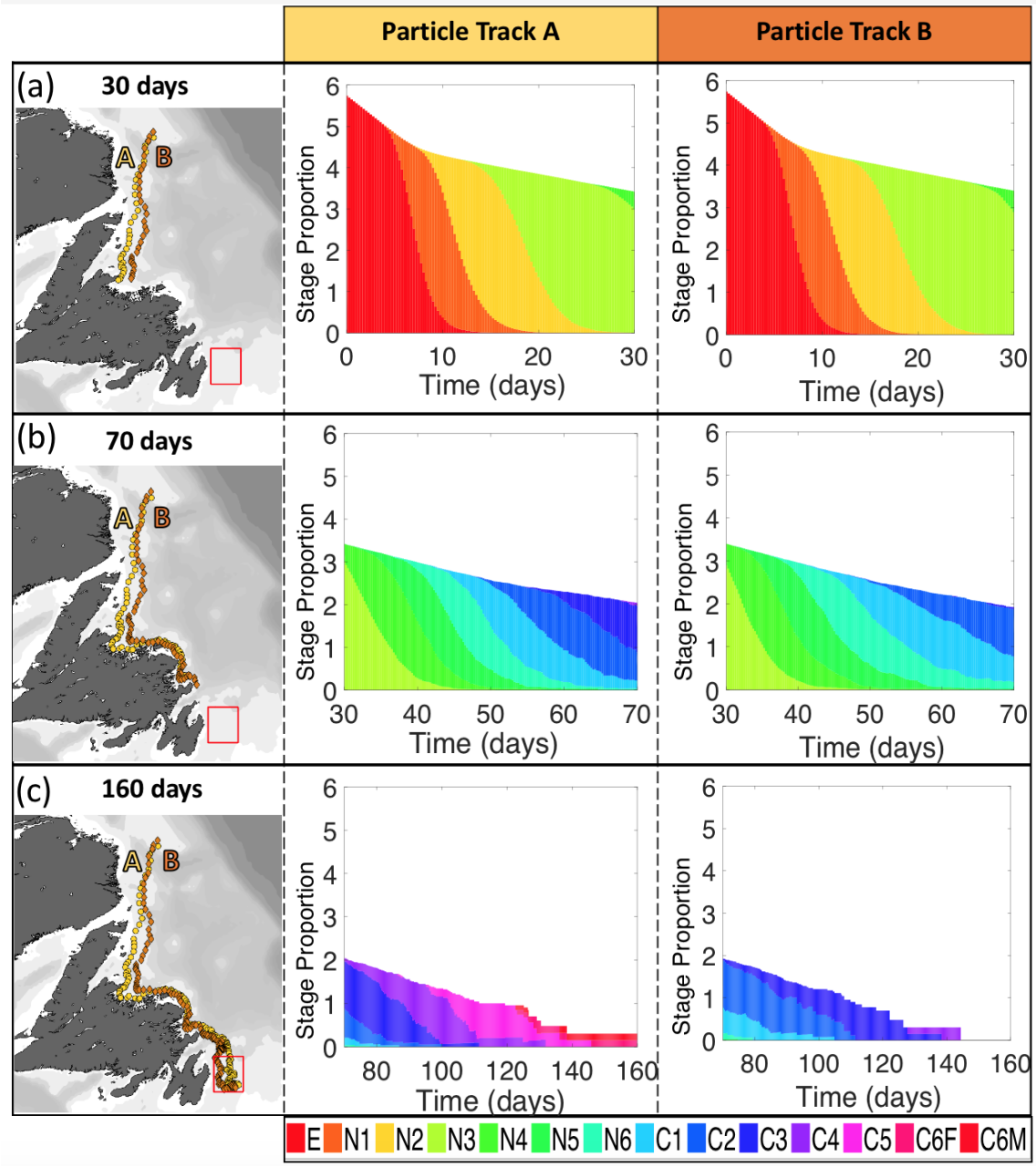


Figure 4.6 Transport of particle tracks A and B (1st column) and associated simulated stage proportions over 180 days. Stage proportions of population sizes are plotted on a logarithmic scale.

4.4 Discussion

The two chosen particle tracks appear to travel southward over similar water depths up to Day 30, which may explain the similarity in temperature conditions up to this time. After Day 30 to about 70, the relatively higher temperatures by PT A may be explained by the relatively closer transport to the coastline than PT B (by about 10 km at some points), where warmer waters are expected to be found. What is interesting is that PT B temperatures do not recover to above 0 ºC, even though the remaining transport is over shallower waters and is very similar to PT A. One would speculate that the low temperatures of PT B are a result of a strong influence of the cold Labrador Current, while PT A may have picked up warmer waters from shallow "pockets" along the coastline (e.g., Figure 4.4a, Figure D.1b) and carried along its track. Whether "real" or considered a demonstration of sampling "error", the great differences in temperature conditions, yet similar transport, among the two tracks demonstrate the effects of high temperature variability on simulating *C. finmarchicus* population dynamics and further inferences regarding population timing and location.

Aggregate durations of simulated populations experiencing PT A conditions and mortality, were shorter than those estimated with only environmental conditions, which is likely explained by the influence of mortality and "survival of the fittest" effect suggested by results in Chapter 3. The greater influence of mortality on longer stage durations, i.e., colder temperatures or later stages, is highlighted with PT B population dynamics results. As conditions of both PTs result in a similar decrease in population size up to Day 70, after this time the individuals experiencing PT B conditions are not only faced with relatively lower temperatures, but are entering copepodite stages with relatively longer durations, and although the "fittest" of the population may be surviving, these survivors are not able to withstand the effects of mortality and die around Day 145 (late August).

The bulk of the PT A population developing to C5s around Day 120 (late July), suggests that individuals of the first annual generation spawned in the mid-Labrador Shelf that experience the conditions of this particular track may prepare to enter diapause around the Grand Banks, or continue to molt into adults and reproduce for a second generation near station 27. In contrast, individuals spawned just above the starting point of PT A,

experience the conditions of PT B, and do not make it to diapause or adulthood (i.e., only develop to C3). The population dynamics results of PT A suggest that large predators dependent upon later staged *C. finmarchicus* for nutrients (e.g., the North Atlantic right whale) that may be migrating from further south are stopping to feed around the Grand Banks in late summer, as later stages of *C. finmarchicus* are simulated at this time and location. The delay in recruitment to later stages of populations from PT B may also provide a food source for predators but those smaller sized (e.g., economically important fish) or younger predators requiring nutrients for crucial growth periods.

Chapter 5: Conclusions

In this thesis study, the influences of environmental variability and mortality on the population dynamics of a highly important species of copepod in Newfoundland-Labrador ocean waters, *Calanus finmarchicus*, have been investigated through a series of analyses. Mortality rates for *C. finmarchicus* copepodite stages were estimated using common methods in the literature, where the underlying assumptions and data processing approaches were quantitatively analyzed through a variety of mortality rate estimation tests. While stage durations have been observed in laboratory studies, empirical relationships have been established based on constant conditions, but are difficult with estimating stage durations in a variable environment. The concept of considering development rates with variable environments and, therefore, stage durations as a resultant property, were introduced here and common methods re-examined. The effect of variable environment scenarios, inherent individual variability, and mortality on emergent stage durations and variance are investigated through a series of single stage development tests. The information gained regarding mortality and development rates is then applied to create an individual-based model for *C. finmarchicus*, with advanced capabilities to simulate life history processes coupled to physical ocean model data. The main features of the applied IBM are presented. Lastly, the IBM coupled to physical data is used to simulate *C. finmarchicus* population dynamics, experiencing mortality and environmental variability in both space and time.

In Chapter 2, mortality rates were estimated for copepodite stages of *C. finmarchicus* based on AZMP sampled environmental conditions in the Newfoundland-Labrador ocean region. Mortality rates were estimated using the most common method in the literature, the Ratio method, and four data processing approaches to determine "average" rates using this method were compared through results of mortality and survivorship. As no approach was able to be determined as the overall "best", recommendations are provided for how to choose an appropriate approach based on data and research requirements. Given the large number of zeros in the sampled abundance data, a technique for creating restricted data subsets for each mortality rate estimate is presented, in an effort to reduce useful data being discarded due to unrealistic (e.g., negative) mortality estimates. Lastly, a proxy

for mortality rates of *C. finmarchicus* in four ocean sub-regions of Newfoundland-Labrador are determined by averaging results of all four data processing approaches, and highest mortality appeared to be in the South regions (off of Newfoundland) for mid- to late staged copepodites in the summer months.

Chapter 3 expanded upon the known relationships between MCF and stage durations and presented the concept of considering development rates in a dynamic environmental scenario, where stage durations are emergent properties, and commonly used methods regarding stage duration are appropriately re-defined with development rates. This concept is applied to single stage development tests, where a group of individuals are simulated through a generic stage subject to various factors and the distribution of their resulting stage durations are analyzed. The simulations supported hypotheses based on observations in laboratory, such that higher temperature scenarios provide shorter emergent stage durations and variance among individual stage durations, and the notion that dynamically increasing temperature also results in smaller stage duration results was proven to be so. The application of mortality to a scenario with individual variability demonstrated that emergent stage durations and variance were also decreased, having a greater effect for less favorable conditions (e.g., low temperatures), providing support for the "survival of the fittest" hypothesis.

Lastly, in Chapter 4, an individual-based model for *C. finmarchicus* is presented with capabilities of coupling to physical ocean model data. The mortality rates found in Chapter 2 are applied to the model, along with individual variability in development rates using a fitness metric and gamma distribution, and dynamic environmental conditions based on particle tracks derived from a 3D circulation model. As no coupled model exists for this region, the IBM enhancements and modifications were a large portion of this study, and the major changes are outlined in this chapter, with a complete user's manual available upon request. The application of the model is demonstrated through a case study coupling two particle tracks (A and B), chosen for their similar transport and final destination around station 27, but yet different temperature conditions, to the IBM simulating a "first" annual generation, and resulting population size and structure are analyzed. The influence of Particle Track B's relatively lower temperature conditions is

evident with population stage structure, where the population develops through stages slower, and does not have individuals pass stage C4, before dying off earlier, as compared to the population experiencing the more favourable conditions of Particle Track A. The findings relating to influence of mortality were consistent with those of Chapter 3, such that a "survival of the fittest" effect was evident by shortened emergent stage durations, when compared to a simulation neglecting mortality or individual variability. The bulk of the population developing to at least C5 by mid-summer just above Georges Bank, with conditions of Particle Track A, suggest that this is a crucial time and location for dependent predators, such as North Atlantic right whales, which is similar to the findings of Chapter 2, where mortality is estimated to be highest for late stage copepodites in the summertime in regions off of Newfoundland. This case study demonstrates how the coupled IBM may be applied to effectively simulate *C. finmarchicus* life history subject to realistic conditions and make further inferences regarding potential migration patterns of predators based on results of population size and structure.

Future work is suggested to assess the sensitivity of population dynamics to parameter values, such as variability in food conditions or the C:Chl mass ratio used (as temperature here was a main focus). Similarly, sensitivity of mortality should be examined by estimating stage-specific mortality rates and testing effects of stage-varying mortalities on simulated population dynamics. Additionally, accurate characterization of egg production based on recent empirical studies (e.g., Head and Ringuette, 2017) is suggested to be incorporated to simulate multiple generations and analyze results on a longer timescale (e.g., complete year rather than single generation). This thesis study has demonstrated the effective use of the presented IBM coupled with physical data to simulate *C. finmarchicus* population dynamics in the Newfoundland-Labrador ocean region (Chapter 4), and is recommended to be applied for exploring a suite of questions regarding this species' life history (e.g., the timing and location of entrance of emergence of diapause), and to build off of the IBM by refining life history details as advanced field and laboratory studies are available, so that predictions regarding spatial and temporal *C. finmarchicus* population dynamics can be made.

References

Aksnes, D.L. and Ohman, M.D. (1996). A vertical life table approach to zooplankton mortality estimation. Limnol. Oceanogr. 41(7):1461–1469.

Belehrádek, J. (1935). Temperature and living matter. Protoplasma Monogr. 8:1–277.

Bi, H., Rose, K.A., and Benfield, M.C. (2011). Estimating copepod stage-specific mortality rates in open ocean waters: a case study from the northern Gulf of Mexico, USA. Mar. Ecol. Prog. Ser. 427:145–159.

Brown, R.G.B., and Gaskin, D.E. (1988). The pelagic ecology of the grey and red-necked phalaropes Phalaropus fulicarius and P. lobatus in the Bay of Fundy, eastern Canada. Ibis 130:234–250.

Campbell, R.G., Wagner, M.M., Teegarden, G.J., Boudreau, C.A., and Durbin, E.G. (2001). Growth and development rates of the copepod Calanus finmarchicus reared in the laboratory. Mar. Ecol. Prog. Ser. 221:161–183.

Crain, J.A. and Miller, C.B. (2001). Effects of starvation on intermolt development in Calanus finmarchicus copepodites: a comparison between theoretical models and field studies. Deep Sea Res. 48:551–566.

Conover, R.J. (1988). Comparative life histories in the genera Calanus and Neocalanus in high latitudes of the northern hemisphere. Hydrobiologia. 167/168:127–142.

Corkett, C.J., McLaren, I.A., and Sevigny, J.M. (1986). The rearing of the marine calanoid copepods Calanus finmarchicus (Gunnerus), C. glacialis Jaschnov and C. hyperboreus Krøyer with comment on the equiproportional rule. In: Schriever G, Schminke HK, Shih C-t (eds) Proceedings of the Second International Conference on Copepoda, 13–17 August 1984, Syllogeus 58. National Museums of Canada, Ottawa, p 539–546

Davis, K.T.A, Gentleman, W.C., DiBacco, C. and Johnson, C.L. (2014). Semi-annual spawning in marine scallops strengthens larval recruitment and connectivity on Georges Bank: a model study. Mar. Ecol. Prog. Ser. 516:209–227.

ACRI-ST, GlobColour (2017). *L3m_20040101-20091231__538758718_4_GSM-MERMODSWF_CHL1_MO_00* [Data file]. Retrieved from http://hermes.acri.fr/index.php?class=archive

Gentleman, W.C. and Head, E.J.H. (2017). Considering non-predatory death in the estimation of copepod early life stage mortality and survivorship. J. Plankton Res. 39(1):92–110.

Gentleman, W.C., Neuheimer, A.B., and Campbell, R.G. (2008). Modelling copepod development: current limitations and a new realistic approach. ICES J. Mar. Sci. 65:399–413.

Gentleman, W.C., Pepin, P., and Doucette, S. (2012). Estimating mortality: clarifying assumptions and sources of uncertainty in vertical methods. J. Mar. Syst., 1:105–108.

Gentleman, W.C. and Neuheimer, A.B. (2008). Functional responses and ecosystem dynamics: How clearance rates explain the influence of satiation, food-limitation and acclimation. J. Plank. Res. 30:1215–1231.

Gislason, A., Eiane, K., and Reynisson, P. (2007). Vertical distribution and mortality of Calanus finmarchicus during overwintering in oceanic waters southwest of Iceland. Mar. Biol. 150:1253–1263.

Grunbaum, D. (1994). Translating stochastic density-dependent individual behavior with sensory constraints to an Eulerian model of animal swarming. J. Math. Biol. 33:139–161.

Han, G., Lu, Z., Wang, Z., Helbig, J., Chen, N. and de Young, B. (2008). Seasonal variability of the Labrador Current and shelf circulation off Newfoundland. J Geophys. Res. 113(C10):13,1-23

Head, E.J.H. (Photographer), Bedford Institute of Oceanography (n.d.). *Calanus finmarchicus* [Digital image]. Retrieved from https://www.st.nmfs.noaa.gov/nauplius/media/copepedia/

Head, E.J.H. and Pepin, P. (2008a). Variations in overwintering depth distributions of Calanus finmarchicus in the slope waters of the NW Atlantic continental shelf and the Labrador Sea. J Northwest Atl. Fish Sci. 39:49–69.

Head, E.J.H. and Pepin, P. (2008b). Seasonal cycles of Calanus finmarchicus abundance at fixed time-series stations on the Scotian and Newfoundland shelves (1999 – 2006). AZMP Bull 7:17 – 20.

Head, E.J.H. and Ringuette, M. (2017). Variability in Calanus finmarchicus egg production rate measurements: methodology versus reality. J. Plankton Res. 39(4): 645– 663.

Head, E.J.H., Gentleman, W.C., and Ringuette, M. (2015). Variability of mortality rates for Calanus finmarchicus early life stages in the Labrador Sea and the significance of egg viability. J. Plankton Res. 37:1149–1165.

Heath, M.R., Boyle, P.R., Gislason, A., Gurney, W.S., Hay, S.J., Head, E.J., Holmes, S., et al. (2004). Comparative ecology of over-wintering Calanus finmarchicus in the northern North Atlantic, and implications for life-cycle patterns. ICES Journal of Marine Science 61:698–708.

Heath, M.R., Rasmussen, J., Ahmed, Y., et al. (2008). Spatial demography of Calanus finmarchicus in the Irminger Sea. Prog. Oceanogr. 76 (1):39–88.

Hirche, H.-J. (1996a). The reproductive biology of the marine copepod, Calanus finmarchicus – A review. Ophelia 44(1-3):111-128.

Hirche, H.-J. (1996b). Diapause in the marine copepod, Calanus finmarchicus – A review. Ophelia 44(1-3):129-143.

Hirst, A.G. and Kiorboe, T. (2002). Mortality of marine planktonic copepods: global rates and patterns. Mar. Ecol. Prog. Ser. 230:195–209.

Hirst, A.G. and Ward, P. (2008). Spring mortality of the cyclopoid copepod Oithona similis in polar waters. Mar. Ecol. Prog. Ser. 372:169–180.

Hirst, A.G., Bonnet, D., and Harris, R.P. (2007). Seasonal dynamics of Calanus helgolandicus and mortality rates of over two years at a station in the English Channel. Mar. Ecol. Prog. Ser. 340:189–205.

Hirst, A.G., Bonnet, D., Conway, D.V.P., and Kiorboe, T. (2010). Does predation control adult sex ratios and longevities in marine pelagic copepods? Limnol. Oceanogr. 55(5):2193–2206.

Hu, Q., Petrik, C.M. and Davis, C.S. (2007). Normal versus gamma: stochastic models of copepod molting rate. Journal of Plank. Res. 29:985–997.

Johnson, C., Leising, A., Runge, J., Head, E., Pepin, P., Plourde, S., and Durbin, E.G. (2008). Characteristics of Calanus finmarchicus dormancy patterns in the northwest Atlantic. ICES J. Mar. Sci.65:339–350. http://dx.doi.org/10.1093/icesjms/ fsm171.

Jung-Madsen, S. and Nielsen, T.G. (2015). Early development of Calanus glacialis and C. finmarchicus. Limnol. Oceanogr. 60:934–946.

Kimmerer, W.J. (2015). Mortality estimates of stage-structured populations must include uncertainty in stage duration and relative abundance. J. Plankton Res. 37(5):939–952.

Kvile, K.O. (2015). Under the surface: disentangling climate effects on Calanus finmarchicus dynamics in a high latitude system (doctoral dissertation). University of Oslo, Norway.

Kvile, K.O., Stige, L.C., Prokopchuk, I., and Langangen, O. (2016). A statistical regression approach to estimate zooplankton mortality from spatiotempiral survey data. J. Plankton Res. 38(3):624–635.

Lynch, D.R., Gentleman, W.C., McGillicuddy Jr, D.J., and Davis, C.S. (1998). Biological/physical simulations of Calanus finmarchicus population dynamics in the Gulf of Maine. Mar. Ecol. Prog. Ser. 169:189–210.

Michaud, J. and Taggart, C.T. (2007). Lipid and gross energy content of North Atlantic right whale food, Calanus finmarchicus, in the Bay of Fundy. Endang. Species. Res. 3:77–94.

Miller, C.B. and Tande, K.S. (1993). Stage duration estimation for Calanus populations, a modelling study. Mar. Ecol. Prog. Ser. 102:15–34.

Miller, C.B., Lynch, D.R., Carlotti, F., Gentleman, W.C. and Lewis, C.V.W. (1998). Coupling of an individual-based population model of Calanus finmarchicus to a circulation model for the Georges Bank region. Fish. Oceanogr. 7(3/4):219–234.

Mitchell, M.R., Harrison, G., Pauley, K., et al. (2002). Atlantic zone monitoring program sampling protocol. Canadian Data Report of Hydrography and Ocean Sciences. 23 pp.

Mollmann, C. and Koster, F.W. (2002). Population dynamics of calanoid copepods and the implications of their predation by clupeid fish in the Centric Baltic Sea. J. Plankton Res. 24(10):959–977.

Mullin, M.M. and Brooks, E.R. (1970). The ecology of plankton off La Jolla California in the period April through September, 1967. Part 7. Production of the planktonic copepod Calanus helgolandicus. Bull. Scripps Inst. Oceanogr. 17:89–103.

Natural Environment Research Council, (NERC), Zooplankton Identification Manual for Northern European Seas (ZIMNES), viewed 8 February 2018, <http://192.171.193.133/index.php>

Neuheimer, A.B., Gentleman, W.C., and Galloway, C.L. (2009). Modeling larval Calanus finmarchicus on Georges Bank: time-varying mortality rates and a cannibalism hypothesis. Fish. Oceanogr. 18(3):147–160.

Neuheimer, A.B., Gentleman, W.C., Pepin, P., and Head, E.J.H. (2010a). How to build and use individual-based models (IBMs) as hypothesis testing tools. J. Mar. Syst. 81:122–133.

Neuheimer, A.B., Gentleman, W.C., Pepin, P., and Head, E.J.H. (2010b). Explaining regional variability in copepod recruitment: implications for a changing climate. Prog. Oceanogr. 87:94–105.

Ohman, M.D., Eiane, K., Durbin, E.G., Runge, J.A., and Hirche, H.-J. (2004). A comparative study of C. finmarchicus mortality patterns at five localities in the North Atlantic. ICES J. Mar. Sci. 61:687–697.

Ohman, M.D., Aksnes, D.L. and Runge, J.A. (1996). The interrelationship of copepod fecundity and mortality. Limnol. Oceanogr. 41(7):1470–1477.

Ohman, M.D., Runge, J.A., Durbin, E.G., Field, D.B., and Niehoff, B. (2002). On birth and death in the sea. Hydrobiologia 480(1–3):55–68.

Pepin, P. and Head, E.J.H. (2009). Seasonal and depth- dependent variations in the size and lipid contents of stage 5 copepodites of Calanus finmarchicus in the waters of the Newfoundland Shelf and the Labrador Sea. Deep Sea Res. Part I Oceanogr. Res. Pap. 56:989–1002.

Pepin, P. and Helbig, J.A. (1997). Distribution and drift of Atlantic cod (Gadusmorhua) eggs and larvae on the northeast Newfoundland Shelf. Can. J. Fish. Aquat. Sci. 54:670–685.

Pepin, P., Colbourne, E., and Maillet, G. (2011b). Seasonal patterns in zooplankton community structure on the Newfoundland and Labrador shelf. Prog. Oceanogr. 91:273–285.

Pepin, P., Han, G., Head, E.J.H. (2013). Modelling the dispersal of Calanus finmarchicus on the Newfoundland Shelf: implications for the analysis of population dynamics from a high frequency monitoring site. Fish. Oceanogr. 22(5):371–387.

Pepin, P., Maillet, G.L., Fraser, S., Shears, T.H., and Redmond, G. (2008) Biological and chemical oceanographic conditions on the Newfoundland Shelf during 2007. Res Doc 2008/034. Canadian Science Advisory Secretariat, Ottawa

Pepin, P., Parrish, C., and Head, E.J.H. (2011a). Late autumn condition of Calanus finmarchicus in the northwestern Atlantic: evidence of size-dependent differential feeding. Mar. Ecol. Prog. Ser. 423:155–166.

Plourde, S., Maps, F., and Joly, P. (2009a). Mortality and survival in early stages controls recruitment in Calanus finmarchicus. J. Plankton Res. 31(4):371–388.

Plourde, S., Pepin, P., and Head, E.J.H. (2009b). Long-term seasonal and spatial patterns in mortality and survival of Calanus finmarchicus across the Atlantic Zone Monitoring Programme region, Northwest Atlantic. ICES J. Mar. Sci. 66(9):1942–1958.

Sathyendranath, S., Stuart, V., Nair, A., Oka, K., Nakane, T., Bouman, H., Forget, M.-H., Maass, and H., Platt, T. (2009). Carbon-to-chlorophyll ratio and growth rate of phytoplankton in the sea. Mar Ecol Prog Ser 383:73–84.

Therriault, J.-C., et al. (1998), Proposal for a northwest Atlantic zonal monitoring program, Can Tech. Rep. Hydrogr. Ocean Sci. 194, 57 pp., Ocean Sci. Div. of Fish. and Ocean Can., Dartmouth, N. S., Canada.

Appendix A

Restricted Subsets for Mortality Rate Estimation

The following Figure A.1 illustrates the procedure outlined in section 2.2.4 for creating restricted subsets from original data to avoid problematic stage ratio calculation with zero abundances. The example used in Figure A.1 demonstrates: (a) original abundance (#/m^2) data sampled from stations during the late May cruise on North on-shelf region, (b) separation of that data into abundance subsets required for each stage pair mortality rate estimation using the Ratio method (Eqns. 1.3 and 1.4), and (c) the removal of zeros and corresponding abundances.

(a) Original Data Set
(e.g. North on-shelf, late May)

day	station	C1	C2	C3	C4	C5	C6F	C6M
146	L3-2	1130	0	0	0	339	3051	0
146	L3-3	283	283	0	0	181	2622	90
146	L3-4	1085	904	0	0	633	9942	316
146	L3-5	565	0	0	90	271	5650	136
146	L3-6	1446	542	0	0	0	3978	0
146	L3-7	565	753	151	57	678	8193	283

(b) Original Subsets
Required for each stage pair mortality rate estimate

$m_{C1\text{-}C2}$

C1	C2
1130	0
283	283
1085	904
565	0
1446	542
565	753

$m_{C2\text{-}C3}$

C2	C3
0	0
283	0
904	0
0	0
542	0
753	151

$m_{C3\text{-}C4}$

C3	C4
0	0
0	0
0	0
0	90
0	0
151	57

$m_{C4\text{-}C5}$

C4	C5
0	339
0	181
0	633
90	271
0	0
57	678

$m_{C5\text{-}C6F}$

C5	C6F
339	3051
181	2622
633	9942
271	5650
0	3978
678	8193

$m_{C5\text{-}C6M}$

C5	C6M
339	0
181	90
633	316
271	136
0	0
678	283

(c) Restricted Subsets
Removal of stage pairs affected by zeros

$m_{C1\text{-}C2}$

C1	C2
283	283
1085	904
1446	542
565	753

$m_{C2\text{-}C3}$

C2	C3
753	151

$m_{C3\text{-}C4}$

C3	C4
151	57

$m_{C4\text{-}C5}$

C4	C5
90	271
57	678

$m_{C5\text{-}C6F}$

C5	C6F
339	3051
181	2622
633	9942
271	5650
678	8193

$m_{C5\text{-}C6M}$

C5	C6M
181	90
633	316
271	136
678	283

Figure A.1. Demonstration of steps required to create restricted subsets from original data for effective mortality rate estimation using the Ratio method.

Appendix B

MA Mortality Rate Estimates using Original and Restricted Subsets

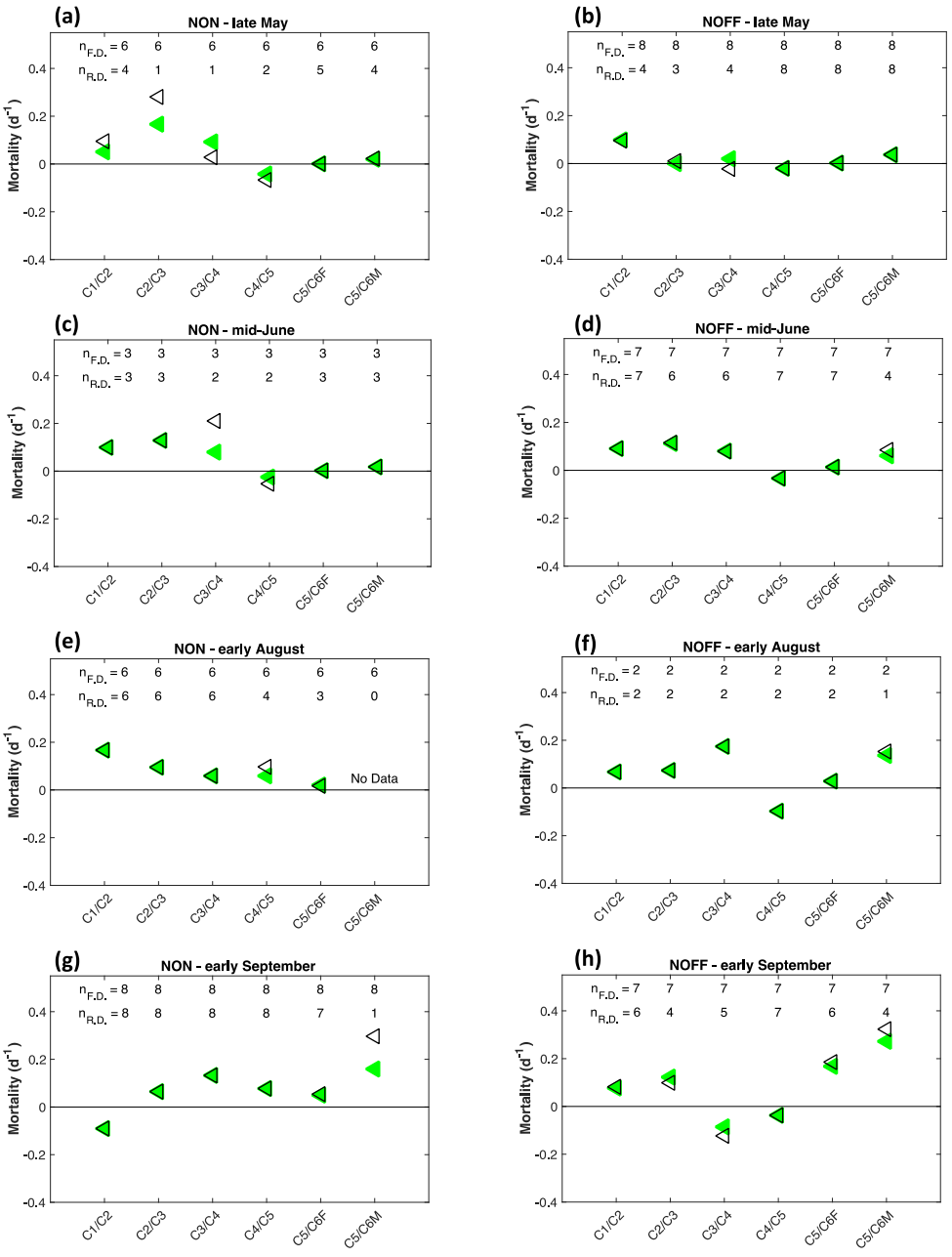


Figure B.1 Average mortality rates for North subsets, estimated by MA using original subsets (outlined triangles) and restricted subsets (green shaded triangles). The subset sample size (n) is provided at the top of each window for original ("full") data, "$n_{F.D.}$", and restricted data, "$n_{R.D.}$". Columns divide regions, while rows represent cruises.

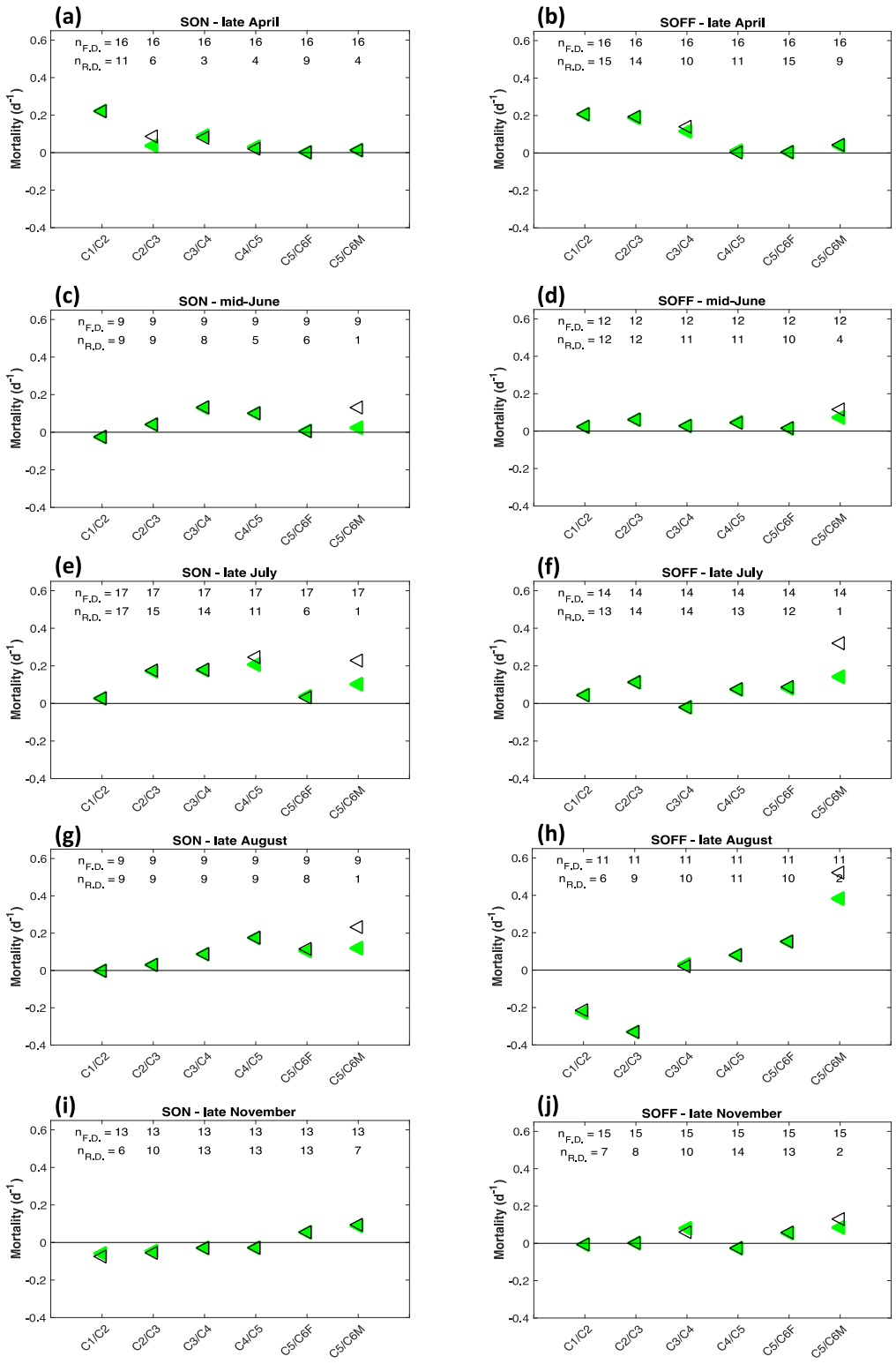


Figure B.2 Average mortality rates for South subsets, estimated by MA using original subsets (outlined triangles) and restricted subsets (green shaded triangles). The subset sample size (n) is provided at the top of each window for original ("full") data, "$n_{F.D.}$", and restricted data, "$n_{R.D.}$". Columns divide regions, while rows represent cruises.

Appendix C

Modified IBM User's Guide

An individual-based model (IBM) for simulating the population dynamics of broadcast spawning marine planktonic copepods (model species: *Calanus finmarchicus*).

C.1 Model Overview

The modified IBM as described in Chapter 4 is outlined here in detail with application and coding parameterization examples.

C.1.1 List of Model Functions

Throughout Appendix C:

- functions are provided in **bold** (written as **functionName.m**)
- variables and structures are *italicized*
- function input and output corresponds to variables and structures used in the code as: [*output*, …] = **functionName**(*input*, …).

The model package is provided with the following functions, in order of use (overview is provided in section C.2.1, while details follow in C.2.2 – C.2.5):

- **main.m**
- **getSimParam.m**
- **getStageIDs.m**
- **getPopColIDs.m**
- **getInitialConditions.m**
- **getTempFood.m**
- **getMortalityRates.m**
- **getDevelopmentInfo.m**
- **getGammaDist.m** (section C.4.4)
- **storeOutput.m**
- **newEggs.m**
- **getEggProductionInfo.m**
- **findMolters.m**
- **mortality.m**
- **development.m**
- **updatePop.m**

C.1.2 Glossary and Nomenclature

Table C.1 General terms to know (in alphabetical order, otherwise in order of use in table)

Terms (abbr.)	Definitions (as used throughout this document)
individual(s) (ind(s))	A single copepod with their own life history metrics (i.e. characteristics, rates, etc). This refers to a single row within the population matrix (i.e. see *pop* below).
population	The grouping of all individuals and their life history information; i.e. *pop* matrix.
subpopulation	Any given subset of the total population (e.g. considering only adults would require analysis of female and male subpopulations).
life history	The series of changes a copepod undergoes during its lifetime. This is affected by various aspects that can be broken down into: (1) characteristics: traits that affect the life of a copepod; e.g. investments in growth, reproduction, and survivorship (Wiki. definition) e.g. in the *pop* matrix: *col.stages, col.fitness, col.MCF* (2) descriptors: miscellaneous information regarding individual conditions and timing parameters. e.g. in the *pop* matrix: *col.tspawn, col.age_in_stg, col.conditionID, col.temp, col.food* (3) rates: life history processes that allow individual progress through life stages. e.g. in the *pop* matrix: *col.devt, col.mort*
identifier (ID)	A variable name (made up of letter and/or numerical characters) used to identify or refer to a property for matrix indexing; e.g. "column identifiers" define variables to represent column positions, while "stage identifiers" define variables for stage representation.
global variable	In MATLAB®: As each function has its own "local" variables, if several functions require the same particular variable, it may be declared as "global" and this will create a single copy of that variable in each function code. (For more information: https://www.mathworks.com/help/matlab/ref/global.html)
logical array	An array containing entries from the Boolean domain; i.e. 0 or 1 relating to false or true, respectively.
development rate	The rate (per day) at which individuals progress through their life stages
molt/molter	The process of an individual developing into the subsequent life stage, i.e. when MCF >= 1. This individual is referred to as a molter.
molt-cycle-fraction (MCF)	A dimensionless fraction of time, determined by the individual's development rate and simulated time step, advancing the individual through their life stages (i.e. MCF = 0 marks the beginning of the life stage, while MCF >= 1 marks the end of the stage/molt into the new stage).
mortality rate	The rate (per day) at which individuals are subject to death or are "removed" from the population.
stage duration (SD)	The time (days) it takes for an individual to fully complete a given stage; i.e. the time it takes for their MCF to progress from 0 to 1. Therefore, individuals who die cannot have a "stage duration" for that given stage.

egg production rate (EPR)	The rate (eggs/female/day) at which eggs are spawned by the female subpopulation.
structures	In MATLAB®: structures are a way to store variables in a group, these variables do not have to be the same size or even variable type. The structure name describes the group, while the variables are stored as "fields". To call or define a specific variable within a structure, the format is *structure.field*. (For more information: https://www.mathworks.com/help/matlab/structures.html)
structures: *col* (*col.*property)	Variables within the *"col"* structure describe the life history metric ('property') located in that column of *pop*; e.g. individual stage IDs are in the *pop* column defined as *col.stages*
structures: *n* (*n.*property)	Variables within the "*n"* structure describe the total number of the 'property' used in the simulation; e.g. "*n.columns"* is the total number of columns in *pop*.
structures: *stg* (*stg.*property)	Variables within the *"stg"* structure describe the life stage in use, where 'property' is the name of the stage; e.g. the ID number for egg stage is found using *stg.egg*.
structures:*sim* (*sim.*property)	Variables with the *"sim"* structure describe the simulation parameters ('property'); e.g. "*sim.deltat*" is the value for the time-step.
time-stepping loop	Iterations simulating progression through time where life history rates are applied to create a dynamic population.
mean	The value obtained by dividing the sum of several quantities by the number of quantities., i.e. referring to the arithmetic mean (units of x): $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$
variance	The expectation of the squared deviation of a random variable from its mean ((units of X)^2): $Var(X) = \sigma^2$

Table C.2 Descriptions of frequently used variables (in alphabetical order, otherwise in order of use in table)

Code Variable (Variable Size)	Description
A (*n.stages* x *sim.length*)	Total abundance of each life stage (i.e. sum of individuals) stored at every point in time specified by *time* (#/m^2).
*col.*property (1x1 each)	Column number in *pop* representing… a 'property': *stages:* life stage ID (see *stg.i*) *tspawn*: time of spawn, i.e. individual "birth" time *age_in_stg*: time spent of individual in its current stage; age-within-stage *MCF*: molt-cycle-fraction *fitness*: fitness level *devt*: development rates *mort*: mortality rates *conditionID*: a number that relates an individual to information about their position over time; such as environmental conditions (e.g. temperature, food), physical conditions (e.g. x-, y-, z-coordinates), etc. *temp*: temperature experienced by individuals (based on *conditionID*) *food*: food conditions experienced by individuals (based on *conditionID*)
*n.*property (1x1 each)	Total number of… 'property': *columns*: total number of columns (i.e. life history information) used in *pop* *stages:* total number of life stages used in simulation
*sim.*property (1x1 each)	Simulation parameters representing… 'property': *deltat*: Time interval (i.e. time-step) used for time-stepping loop *length*: Length of time for one simulation run (day)
*stg.*i (1x1 each)	ID numbers for life stages (may be separate or aggregate) involved in simulation (e.g. for a population involving all egg, nauplii, and copepodite stages, we have: *stg.egg* = 1, *stg.N1* = 2, …, *stg.fem* = 13, *stg.male* = 14).
molt_info.molters *.mcf_deltat* *.new_ageinstg* (*n_pop* x 1each)	*.molters* (n.d.)*:* Logical matrix used for indexing *pop* – identifying which individuals will be molting (i.e. 1 = molting, 0 = not molting) during the given time-step. *.mcf_deltat* (n.d.): The new incremented MCF over the time step (updated for all in **development.m**). *.new_ageinstg* (days): The new age-within-stage for molters (updated for molters in **development.m**).
*mu*_property (*sig_, var_, CV_*) (*n_pop* x 1 each)	Set value for mean (days; standard deviation, days; variance, $days^2$; coefficient of variation, n.d.) of… 'property', used for stochasticity in assigning individual rates (e.g. *mu_epr* represents the mean egg production rate).
newEggs_pop (n_newEggsAdded x 1)	Temporary storage for newly spawned eggs with life history information to be added to *pop* matrix (in **updatePop.m**)

pop (*n_pop* x *n.columns*)	Dynamic matrix containing life history information of all individuals; together these individuals make up the population
pop_in (n_pop_in x *n.columns*)	A subpopulation of *pop* used for function input when the full population is not required (i.e. improves model efficiency).
n_pop (1x1)	Length of *pop* (i.e. total number of individuals in population)
property_*inds* (n_inds x 1 each)	Individual values of a given property (e.g. *fitness_inds* is an array of random values to be set as the individualized fitness levels, while *devt_inds* is an array of individualized development rates, typically sampled with a distribution).
t (1x1)	"Current" time during time-stepping loop (simulation) (days)
time (1 x *sim.length*)	Points in time for time-series storage of simulated data (days)

C.2 Main Code Overview and Function Details

Details of how the model operates as well as information outlining the purpose of each function and the variables (and/or calculations) required.

C.2.1 Main Code Overview

The main code (**main.m**) is essentially the control center of the model (Figure C.1) and is responsible for declaring key variables and calling all functions necessary for simulation, divided into the following major steps: Initialization, Core Processes, Transition. A general overview of these steps is provided below, while specific function details are provided in sections C.2.2 – C.2.5.

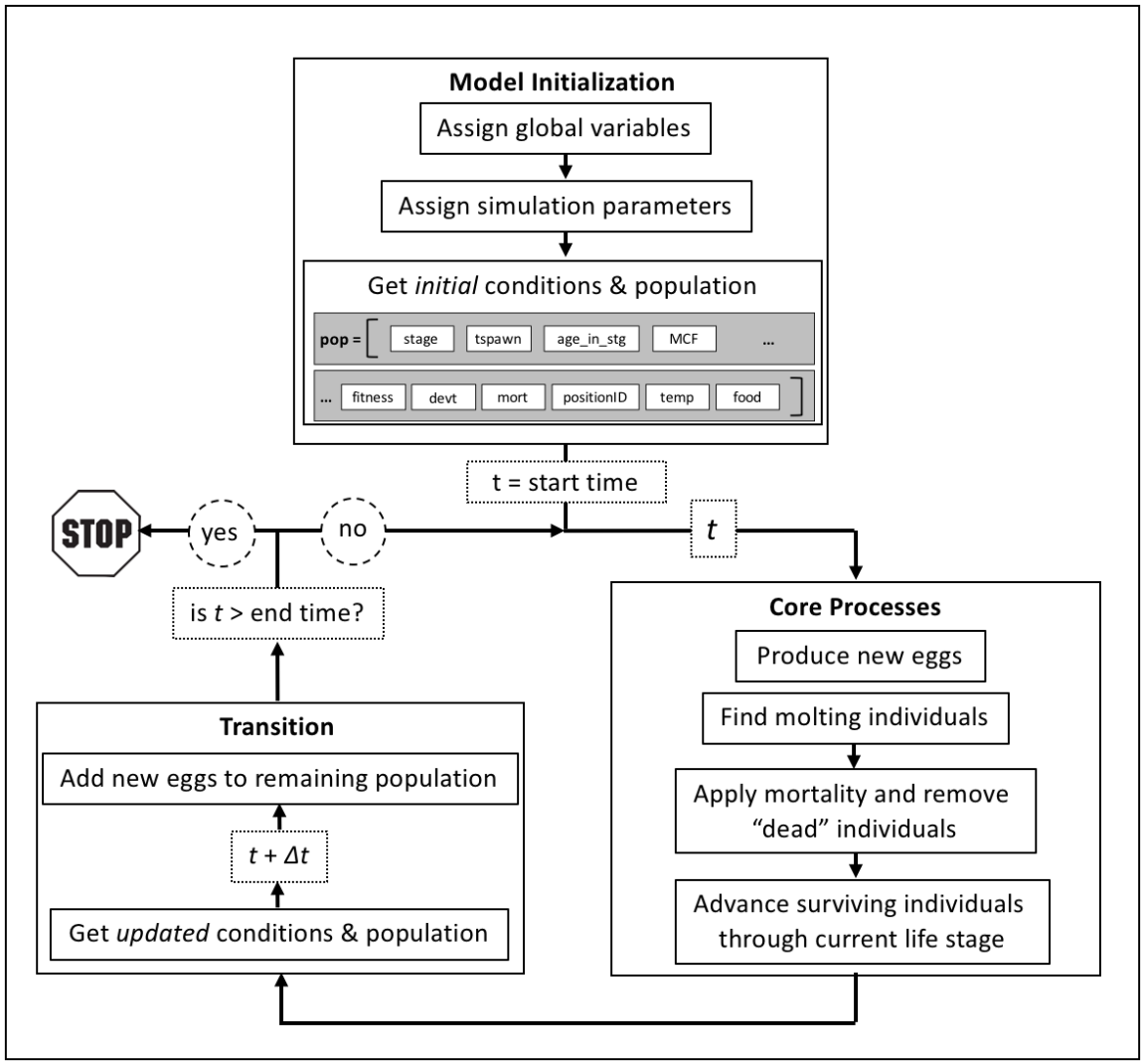


Figure C.1 Conceptual flow diagram representing the main code of the modified IBM

Initialization:

The initialization step is made up of two parts: (1) assigning parameters required for simulation and matrix setup, and (2) initialization of variables to prepare the population matrix with initial conditions.

Simulation Parameters and Setup:

The main code clears all previously defined variables and inputs, closes all open outputs (e.g. figures), and clears all displayed outputs in the "Command Window". This is good practice to eliminate possibility of interference with any previous simulation runs. The main code also assigns all "global variables", whose values are assigned by functions and are frequently used throughout the model. All functions titled "**getX.m**", where **X** is each function's descriptive name, are called to assign values to specific variables. These functions are used throughout the entire model wherever a new value for these specific variables is necessary (more details below).

First off, the main code calls the function **getSimParam.m** to assign parameters, pre-defined by the user, that are necessary for model simulation (i.e. simulation length, time-step value, start time). Next, the functions **getStageIDs.m** and **getPopColIDs.m** are called to assign values to variable names identifying life stages and matrix column positions necessary to assign and locate individual properties.

Initial Conditions:

The main code uses **getInitialConditions.m** to set up the initial population matrix (i.e. *pop*), where the number of rows (i.e. individuals) are defined by user-inputted abundances and columns (i.e. properties) are "filled in" either by calling forcing functions or hardcoded within the code. The properties that require functions are temperature and food (**getTempFood.m**), mortality (**getMortalityRates.m**), fitness and development rates (**getDevelopmentInfo.m**), which are all included in the population matrix, specific to each individual. If any of these should depend on the other, the order will matter as the independent property should be assigned first.

It is at this point in the simulation that any initial conditions that the user may need for further analysis is stored (e.g. see **storeOutput.m**).

Core Processes:

The simulation can now proceed with the time-stepping loop, where at each time-step...:

(i) Existing females produce new eggs – **newEggs.m** initially calls **getEggProductionInfo.m** to determine the number of new eggs to add to the population and creates a matrix representing the new subpopulation of eggs (i.e. *newEggs_pop*, Table C.2). Here, any properties related to spawning (e.g. timing or position) are set, where then the new egg matrix is returned to **newEggs.m** and remaining properties are assigned by hardcoding or calling appropriate functions (i.e. **getTempFood.m**, **getMortalityRates.m**, and **getDevelopmentInfo.m**). The output *newEggs_pop* should represent only "new eggs" being added to the population, i.e. survivors after spawning and potential "lost" eggs, and all properties regarding to "age" are set (e.g. age-within stage). Therefore, the new subpopulation matrix is stored and unused until the final step of the time-stepping loop (see Transition section below).

(ii) Find molting individuals – **findMolters.m** calculates each individual's new MCF for that time-step, and marks those that will be molting to the next stage. This information will be used throughout the remainder of the core processes, as molters undergo different mortality and development than non-molters.

(iii) Apply mortality & remove dead individuals – **mortality.m** determines which individuals are subject to mortality (i.e. death), and removes these individuals from the *pop* matrix, along with stored information in corresponding variables (i.e. *molt_info,* Table C.2).

(iv) Advance surviving individuals through life stages – **development.m** progresses each individual through their life stage using information regarding MCF and age-within-stage found in **findMolters.m**. Individuals that have molted require additional properties to be updated relating to their new stage.

Transition:

Towards the "end" of the time-stepping loop, the simulation is now required to increment the time-step for the next iteration and subsequently update the population matrix and its properties using **updatePop.m**:

(i) Add new eggs to remaining population – the new egg subpopulation is now added to the total population after mortality and development have taken place.

(ii) $t + \Delta t$ – the timing counter is incremented to allow for the next iteration.

(iii) Get updated conditions & population – environmental conditions are updated (i.e. **getTempFood.m**) based on the new time (e.g. if using temporally dynamic forcings) for each individual, as well as any potentially dependent properties (i.e. at least **getMortalityRates.m** and **getDevelopmentInfo.m**). These properties will also be updating for molters from the previous development step.

Completion:

As was done at the end of the population initialization step, the user may store any information that they may need for further analysis (e.g. see **storeOutput.m**).

The time-stepping loop will then compare the current time (t) with the assigned end time and will either return to the beginning of the loop for another iteration or will stop if the condition of the loop has been met.

C.2.2 Simulation Parameters and Setup

The following functions are required by the main code to provide parameters necessary for the simulation to run, as well as identifiers for matrix positions.

- The contents of these functions are to be specified by the user, prior to simulation, to suit their research needs and requirements.
- The running order of these functions within the main code may be rearranged (as well they may be compared into one function if desired).

getSimParam.m **(1)**
This function assigns variables necessary for simulation. No calculations.

Input: value of 0 (i.e. $t = 0$ for initial condition)
Output: *sim* (*sim.length, sim.deltat*), *t* (initial condition)

Note: Time step recommendation is that the value should be "small enough" to ensure that individuals are not "skipping" stages in development, i.e. do not jump to stage(i+2) from stage(i). The user should also take note of units of conditions (i.e. temperature or food) and choose a suitable time step (e.g. can use $\Delta t = 1$ hour if temperature provided is hourly, otherwise would have to decide how to interpolate if temperature is daily).

getStageIDs.m **(2)**
This function assigns variables to represent copepod life stages numerically; used for assigning stage-based individual life history metrics and vital rates. No calculations.

Input: value of 0 (i.e. $t = 0$ for initial conditions)
Output: *stg* (*stg*.i where i = stages involved, *stg.first, stg.last*), *n* (*n.stage*)

Note: The order of *stg*.i matters, as the stages involved must be numbered chronologically; e.g. including stages egg through to adult: *stg.egg* = 1, *stg.N1* = 2, …, *stg.adult* = 13, and therefore *stg.first* = 1, *stg.last* = 13 and *n.stage* = 13.

getPopColIDs.m **(3)**
This function assigns variables to represent matrix column positions (i.e. order of columns) and the total number of columns used in the population matrix. No calculations.

Input: value of 0 (i.e. $t = 0$ for initial conditions)
Output: *col* (*col.property*, user specific), *n* (*n.columns*)

Note: The order of *col.*property does not matter, as the code will set up the population matrix based on these set values. Each "property" is a stand-alone life history metric; e.g. if for stage IDs in the i^{th} column: *col.stages = i,* or for MCF in the n^{th} column: *col.MCF = n*.

C.2.3 Simulation Initialization

The following functions are required by the main code to initialize (i.e. at $t = 0$) variables necessary for the simulation.

- The contents of these functions are to be specified by the user, prior to simulation, to suit their research needs and requirements. Sections that need user-specific input are noted in each function by "Specify" either in the title and/or comments (to the right of code).
- The running order of these functions given below may NOT be rearranged, as subsequent functions rely on previously assigned variables.

getInitialConditions.m **(4)**
This function sets up the population matrix, *pop*, by user defined abundances and assigns variables for individual life history information using necessary functions (i.e. fills in *pop* matrix columns).

Input: *t* (i.e. value of 0)
Output: *pop*

The following functions are called *within* **getInitialConditions.m** to calculate and assign individual properties:

getTempFood.m **(5)**
This function assigns values for individual temperature and food over the simulation period. Data may be inputted through an external file (any format, e.g. .xlsx, .txt, .csv), hard coded (i.e. values assigned in the function code), or calculated (i.e. a form of hard coding using a relationship with already assigned variables).

Input: *t, pop* or *pop_in* (only required if using information regarding position; see C.2.4)
Output: *T* (*col.temp*), *F* (*col.food*)

Note: The units here should agree with any functions dependent on environmental conditions (e.g. development rate calculation in **getDevelopmentRates.m**).

getMortalityRates.m **(6)**
This function assigns values for individual mortality rates (set up for stage-based). Values for mortality may be inputted through an external file (any format, e.g. .xlsx, .txt, .csv), hard coded (i.e. values assigned in the function code), or calculated (i.e. a form of hard coding using a relationship with already assigned variables).

Input: *pop* or *pop_in* (required to use information regarding stage IDs; may also be used for mortality based on environmental conditions or position).
Output: *mort_inds* (*col.mort*)

Note: The mortality rate units should agree with the units of the time step and other vital rates.

getDevelopmentInfo.m **(7)**
This function assigns values for individual fitness and development rates (set up for stage-based). Values for fitness and development rates may be inputted through an external file (any format, e.g. .xlsx, .txt, .csv), hard coded (i.e. values assigned in the function code), or calculated (i.e. a form of hard coding using a relationship with already assigned variables.

Input: *pop* or *pop_in* (required to use information regarding environmental conditions)
Output: *fitness_inds* (*col.fitness*), *devt_inds* (*col.devt*)

Note: The development rate units should agree with the units of the time step and other vital rates.

The following function is called *after* **getInitialConditions.m** to store initial condition data:

storeOutput.m **(8)**
This function is run after initialization of matrices and variables to store the initial conditions necessary for post-simulation analysis. The inputs, outputs, and contents of this function may be tailored to the user's needs and requirements.

Input: *t* (i.e. 0), *pop, time, A*
Output: *time, A*

Note: The provided default function is set to illustrate an example of storing a data series of time and corresponding abundances of stages.

C.2.4 Core Functions and Transition

The following functions are required by the main code to simulate individual life history:

- The contents of these functions are NOT to be specified by the user (with the exception of **getEggProductionInfo.m**, see below).
- The running order of these functions (see numbering to left of titles) may NOT be rearranged as subsequent functions rely on previously assigned variables.

For efficiency of running the core functions, a binary matrix describing the molting status of the individuals that make up *pop* is implemented:

2. Molting individuals:
findMolters.m **(9)**
This function calculates/stores each individual's "new" incremented MCF (does not assign in *pop*) and identifies those individuals that *will* be molting over the current time step. This function also calculates/stores the time "spent" in the new stage of molters.

Input: *pop*
Output: *molt_info*

There are three core functions run during each time step, one for each of the vital life history processes (see section C.4 for more details).

1. Egg production:
newEggs.m **(10)**
This function calls **getEggProductionInfo.m** to create the subpopulation of newly spawned eggs. Here the columns of the subpopulation matrix are filled in with corresponding life metrics/rates and returned to the main code (later added to *pop*; see **updatePop.m**).

Input: *t, females* (i.e. female subpopulation of *pop*)
Output: *newEggs_pop*

getEggProductionInfo.m **(11)**
This function creates a matrix corresponding to the size of newly spawned eggs, based on egg production rates and life history characteristics at spawn are assigned, these are all specified by the user. Values for the egg production rates or total number of newly spawned eggs (if disregarding female subpopulation) may be inputted through an external file (any format, e.g. .xlsx, .txt, .csv), hard coded (i.e. values assigned in the function code), or calculated (i.e. a form of hard coding using a relationship with already assigned variables).

Input: *females* (i.e. female subpopulation of *pop*)
Output: *newEggs_pop*

3. Mortality:
mortality.m **(12)**
This function uses individual mortality rates to determine which individuals to remove from *pop* (i.e. die) over the time step.

Input: *pop, molt_info*
Output: *pop, molt_info*

Note: The information of individuals who "die" is removed from *pop* as well as all corresponding information is removed from *molt_info*.

4. Development:
development.m **(13)**
This function advances individuals in their life stages based on MCF/age-within-stage found in **findMolters.m**, and updates necessary life history information for molters.

Input: *pop, molt_info*
Output: *pop*

Note: Here, the Euler integration method is used for the developmental process.

5. Updating conditions and rates:
updatePop.m **(14)**

This function adds the newly spawned eggs from **newEggs.m** to *pop*, increments time, and updated individual conditions and rates (e.g. environment, mortality and development – by calling necessary functions).

Input: *t, pop, newEggs_pop*
Output: *t, pop*

C.2.5 Post Processing

The following function is called by the main code to store simulated data while the time stepping loop is running.

- The contents of this function are to be specified by the user to suit their research needs and requirements.

storeOutput.m **(8)**

This function is run after each time step run to store the simulated data for post-simulation analysis. The inputs, outputs, and contents of this function may be tailored to the user's needs and requirements.

Input: *t*, *pop, time, A*
Output: *time, A*

Note: The provided default function is set to illustrate an example of storing a data series of time and corresponding abundances of stages.

C.3 Subroutine Details

Derivation of calculations and variable assignments used in core process functions.

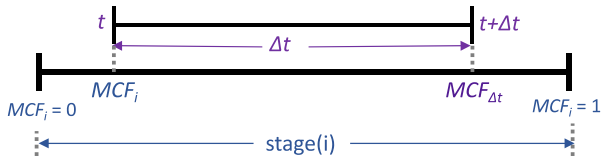
C.3.1 Determining molters

Find molting individuals
function: **findMolters.m**
description: finding which individuals *will be* molting during time-step

1. Calculate "new" MCF for the given time-step and find molters:

$$MCF_{\Delta t} = MCF_i + \Delta t \cdot devt_{t,i}$$

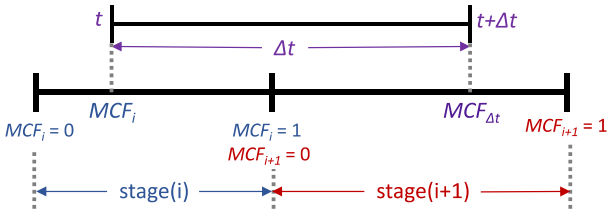
Non-molters:



$$MCF_{\Delta t} < 1$$

$\therefore$ *molters* = 0 (i.e. "false")

Molters:



$$MCF_{\Delta t} \geq 1$$

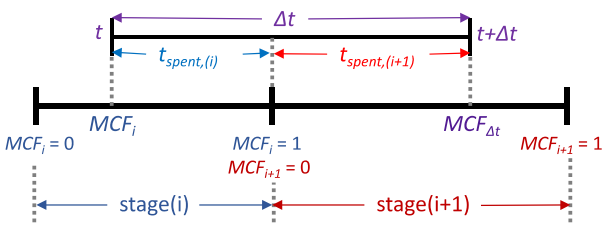
$\therefore$ *molters* = 1 (i.e. "true")

Find molting individuals
function: **findMolters.m**
description: finding which individuals *will be* molting during time-step

2. Calculate the time spent in stage(i+1) for molters (i.e. find age-within stage):

Molters:

*Below is an illustration of a molting situation at time *t* over *Δt*



Main Equations:

(1) $t = \frac{MCF}{devt} = \frac{\{n.d.\}}{\{d^{-1}\}} = \{d\}$

(2) $\Delta t = t_{spent,(i)} + t_{spent,(i+1)}$

(3) $MCF_{\Delta t} = MCF_i + \Delta t \cdot devt_{t,i}$

Method 1 (simple):
Find $t_{spent,(i+1)}$ using equations (1) & (3):

$$t_{spent,(i+1)} = \frac{MCF_{\Delta t} - 1}{devt_{t,(i)}}$$

$$= \frac{MCF_i + \Delta t \cdot devt_{t,(i)} - 1}{devt_{t,(i)}}$$

$$= \frac{MCF_i - 1}{devt_{t,(i)}} + \frac{\Delta t \cdot \cancel{devt_{t,(i)}}}{\cancel{devt_{t,(i)}}}$$

$$= \frac{MCF_i - 1}{devt_{t,(i)}} + \Delta t$$

$$= -t_{spent,(i)} + \Delta t$$

$$\therefore t_{spent,(i+1)} = \frac{MCF_{\Delta t} - 1}{devt_{t,(i)}}$$

Method 2 (extra):
Find $t_{spent,(i)}$ by rearranging eqn. (2) and plugging in eqn. (1) for $t_{spent,(i+1)}$ & using (3):

$$t_{spent,(i)} = \Delta t - t_{spent,(i+1)}$$

$$= \Delta t - \left(\frac{MCF_{\Delta t} - 1}{devt_{t,(i)}}\right)$$

$$= \Delta t - \left(\frac{(MCF_i + \Delta t \cdot devt_{t,(i)}) - 1}{devt_{t,(i)}}\right)$$

$$= \Delta t + \left(\frac{-MCF_i - \Delta t \cdot devt_{t,(i)} + 1}{devt_{t,(i)}}\right)$$

$$= \Delta t + \left(\frac{1 - MCF_i}{devt_{t,(i)}} - \frac{\Delta t \cdot \cancel{devt_{t,(i)}}}{\cancel{devt_{t,(i)}}}\right)$$

$$= \Delta t + \frac{1 - MCF_i}{devt_{t,(i)}} - \Delta t$$

$$t_{spent,(i)} = \frac{1 - MCF_i}{devt_{t,(i)}}$$

$$\therefore t_{spent,(i+1)} = \Delta t - t_{spent,(i)}$$

$$\therefore age_{instage(i+1),new} = t_{spent,(i+1)}$$

C.3.2 Mortality Process

Apply mortality & remove "dead" individuals
function: **mortality.m**
description: applying mortality to population and removing individuals who have "died"

1. Finding *mrisk:*

Mortality risk is the "amount of time the mortality rate is experienced for":

(4) $mrisk = m \cdot t$

Non-molters:

Simply use eqn. (4):

$$mrisk_i = m_i \cdot \Delta t$$

Molters:

Because mortality can be stage-dependent, molters require their risk to be calculated taking into consideration time spent in each stage and the mortality for that stage:

$$\therefore mrisk_{total} = mrisk_{(i)} + mrisk_{(i+1)}$$

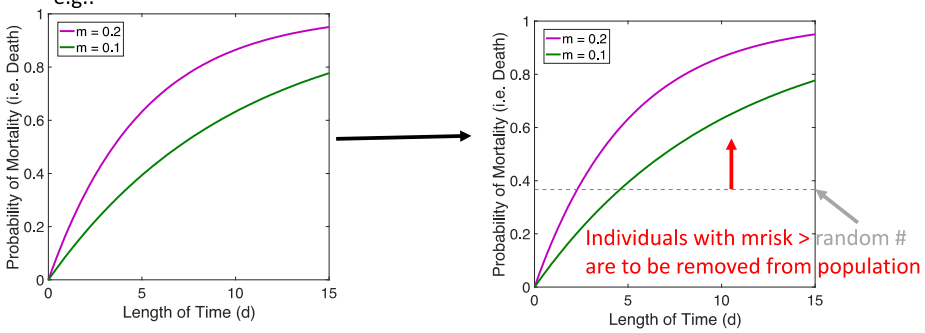
$$mrisk_{total} = m_i \cdot t_{spent,(i)} + m_{i+1} \cdot t_{spent,(i+1)}$$

2. Finding probability of "dying" and removing individuals:

(5) $mprob(ind) = 1 - e^{-mrisk(ind)}$

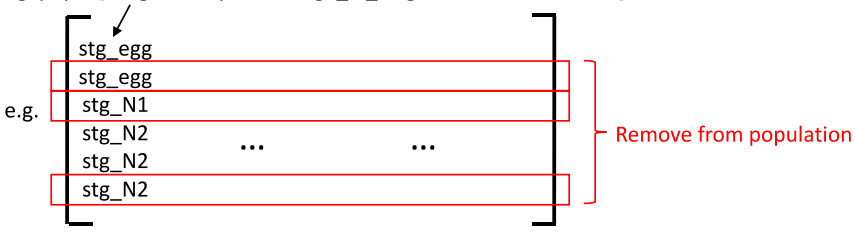
"Random chance" of death is applied to all individuals (i.e. each has own "random" #):

e.g.:



Individuals are removed (as well as corresponding molters' information):

e.g. pop = [stages tspawn age_in_stage MCF fitness …]



C.3.3 Development Process

Advance surviving individuals through life stages
function: **development.m**
description: advancing individuals' MCF/age-within-stage and updating properties for molters

1. Advance MCF/age-within-stage for *all* individuals and update corresponding information for molters:

Non-molters:

No updates; only increment MCF (see **findMolters.m** description for information) and age-within-stage (below):

$$age_{instage(i),new} = age_{instage(i),last} + t_{spent,(i)}$$

Molters:

Updating of new stage information (stage ID, fitness, development):

Updating time-within-stage:

$$age_{instage(i+1),new} = t_{spent,(i+1)}$$

Updating MCF, molting stages egg-C4:

and since: $t = \frac{MCF}{devt}$: $\quad MCF_{(i+1)} = t_{spent,(i+1)} \times devt_{t,(i+1)}$

Updating MCF, molting stages C5 (i.e. new adults - do not continue molt):

$$MCF_{(i+1)} = \text{NaN}$$

Splitting new adults into males and females
(assuming sex ratio at molt = males:females = 1:1):

Assign random number to each new adult (0 to 1):

$$\begin{bmatrix} \text{new C6} \\ \text{new C6} \\ \text{new C6} \\ \text{new C6} \end{bmatrix} = \begin{bmatrix} \text{rand(0-1)} \\ \text{rand(0-1)} \\ \text{rand(0-1)} \\ \text{rand(0-1)} \end{bmatrix}$$

Assign all adult individuals with random numbers > 0.5 (approx. half) as males:

e.g.

$$\begin{bmatrix} \text{new C6} \\ \text{new C6} \\ \text{new C6} \\ \text{new C6} \end{bmatrix} = \begin{bmatrix} 0.78 \\ 0.24 \\ 0.51 \\ 0.48 \end{bmatrix}$$

Update stage IDs

Note: most stage specific metrics in *pop* are updated in the next subroutine (**updatePop.m**) corresponding to the new time increment, *t*; therefore, only update stage ID, development, and new MCF here.

C.3.4 Determining Rates with Individual Variability

Add stochasticity using gamma distribution
function: **getGammaDist.m**
description: inversely samples a property value from a gamma distribution given a mean and standard deviation

Finding parameters necessary for gamma distribution:

We apply: $\mu = mean, \sigma^2 = variance$

We need: $k = shape, \theta = scale$

Known relationships:

(1) $\mu = \theta\, k$ (2) $\sigma^2 = k\theta^2$

Rearranging (1) and (2) for k and setting equal; solving for θ:

$$k = \frac{\sigma^2}{\theta^2} = \frac{\mu}{\theta}$$

$$\frac{\sigma^2}{\theta} = \mu$$

$$\therefore \theta = \frac{\sigma^2}{\mu}$$

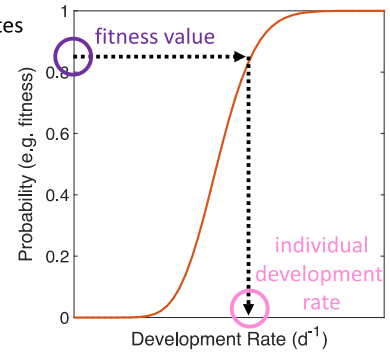
Adding equation for θ into equation (1) rearranged for k:

$$k = \frac{\mu}{\theta} = \frac{\mu}{\frac{\sigma^2}{\mu}}$$

$$\therefore k = \frac{\mu^2}{\sigma^2}$$

Inverse sampling from corresponding gamma distribution:

e.g., determining individual development rates



gamma CDF = $f(\theta, k)$

C.4 Application Examples

Examples of various scenarios that may be simulated using the modified IBM and how to parameterize.

C.4.1 Lab Scenario

Timing individual stage durations based on constant temperature and food conditions with inherent individual variability in development.

- This example demonstrates how to parameterize the modified IBM to replicate a laboratory experiment (à la: Campbell *et al*., 2001, Gentleman *et al.,* 2008).

Table C.4 Specification of variables in the IBM for Example C.4.1

Variables	Value/Formula	Units
getSimParam.m		
sim.length	100	d
sim.deltat	0.25	d
getStageIDs.m		
stg._: *egg,* *N1, N2, ..., C5*	1, 2, …, 12 (order matters; stages numbered chronologically)	n/a
getPopColIDs.m		
col._: *stages* *tspawn** *age_in_stg* *MCF* *fitness* *devt* *mort* *conditionID** *temp* *food*	1, 2, …, 10 (order does not matter; code sets up matrices dependant on these values)	n/a
getInitialConditions.m		
A	$A_E = 1000$ $A_{N1}, \ldots, A_{fem}, A_{male} = 0$	individuals m^{-2}
*col.tspawn**	NaN	n/a
*col.conditionID**	NaN	n/a
col.age_in_stg	0	d
col.MCF	$MCF_i = age\_within\_stage_i \times devt_i = 0$	n.d.

getTempFood.m		
F_applied (*col.food*)	see Table C.3 below	mgCarbon/m^3
T_applied (*col.temp*)	see Table C.3 below	°C
getMortalityRates.m		
mort_inds (*col.mort*)	$mort_E$, …, $mort_{fem}$, $mort_{male}$ = 0	d^{-1}
getDevelopmentInfo.m		
fitness_inds	stage-based: randomly chosen from uniform distribution	n.d.
dev	dev_E, …, $dev_{N2} = a_i^{-1}(T-b)^{-c}$ dev_{N3}, …, $dev_{C5} = a_i^{-1}(T-b)^{-c}(F/(d+F))$ dev_{fem}, dev_{male} = NaN (a_i, b, c, d: Campbell *et al.,* 2001)	d^{-1}
CV_dev	$CV_{dev,E}$, …, $CV_{dev,C5}$ = 0.15/0.3* $CV_{dev,adult}$ = NaN *approx. averages for low-med./high food from Gentleman *et al.,* 2008	n.d.
devt_inds (*col.devt*)	sampled using Gamma distribution	d^{-1}
newEggs.m		
newEggs_pop	"blank" = [ ] (i.e. no eggs produced)	n/a

**Are not required for this example but must have "placeholders" for the model to run. Should these be omitted the subroutine functions must be altered to remove corresponding sections.*

Table C.3 Combinations of constant food/temperature conditions for simulations (Gentleman et al., 2008)

Temperature (below)/ Food (right) Conditions	**Low (L, 25 mgC/m^3)**	**Medium (M, 40 mgC/m^3)**	**High (H, 350 mgC/m^3)**
T_1 = 4 °C	n/a	n/a	Sim. 1
T_2 = 8 °C	Sim. 5	Sim. 4	Sim. 2
T_3 = 12 °C	n/a	n/a	Sim. 3

Shown below in Figure C.2 are the relative values of model outputs of mean and standard deviation of development time from the egg stage to adult stage, as compared to experimental lab data found for five different scenarios (Table C.3). While the mean results are fairly similar to the data, the standard deviation results are off, which may be due to the fact that a constant approximate value for C.V. of development rate was used for all stages, while these may greatly differ in a true lab experiment.

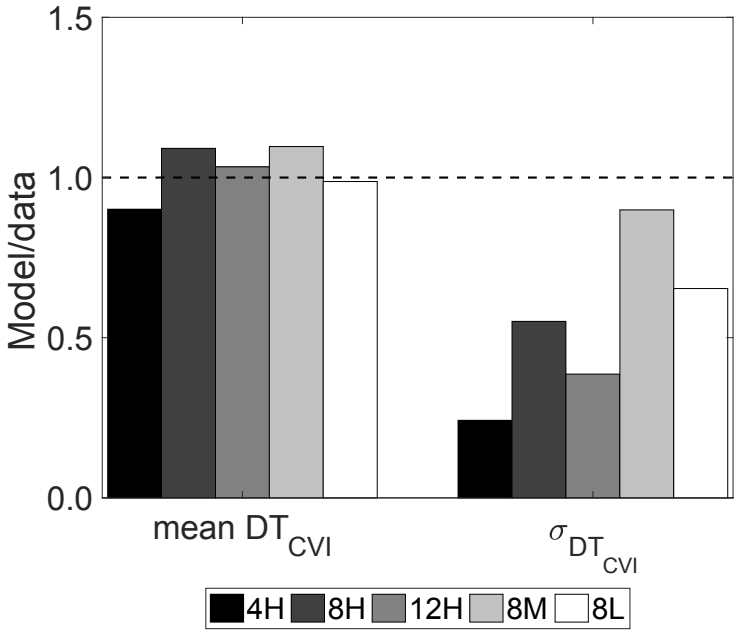


Figure C.2 Application example C.4.1: Comparison of model results vs. data for development time to adult stage (à la Gentleman et al., 2008, Fig. 11)

C.4.2 Dynamic Conditions (in Time)

Timing individual stage durations based on temporally dynamic temperature and food conditions with inherent individual variability in development.

- This example demonstrates how to parameterize the IBM in order to produce similar literature results (à la: Neuheimer *et al.,* 2010a).

Table C. 4 Specification of variables in the IBM for Example C.4.2

Variables	Value/Formula	Units
getSimParam.m		
sim.length	365	d
sim.deltat	0.25	d
getStageIDs.m		
stg._: *egg*, *N1*, *N2*, …, *N6* *C1, C2, ..., C5* *fem* (i.e. *C6*)	1, 2, …, 13 (order matters; stages numbered chronologically)	n/a
getPopColIDs.m		
col._: *stages* *tspawn** *age_in_stg* *MCF* *fitness* *devt* *mort* *conditionID** *temp* *food*	1, 2, …, 10 (order does not matter; code sets up matrices dependant on these values)	n/a
getInitialConditions.m		
A	A_E, A_{N1}, … A_{C1} = [82, 24, 20, 46, 16, 7, 4, 2] A_{C2}, A_{C3}, … A_{C5} = 0 A_{fem} = *f(t)*; see Figure C.4	individuals m^{-2}
*col.tspawn**	NaN	n/a
*col.conditionID**	NaN	n/a
col.age_in_stg	0	d
col.MCF	MCF_i = age_within_stage$_i$ x $devt_i$ = 0	n.d.
getTempFood.m		

F_applied (*col.food*)	$F_{chl\text{-}a} = f(t)$; see Figure C.4	mgC/m^3
T_applied (*col.temp*)	$T = f(t)$; see Figure C.4	°C
getMortalityRates.m		
mort_inds (*col.mort*)	$mort_E$, …, $mort_{fem}$ = [0.49, 0.49, 0.2156, 0.051, 0.0765, 0.1412, 0.1487, 0.1163, 0.0691, 0.061, 0.0672, 0.0578, 0] (derived from Ohman *et al.,* 2002)	d^{-1}
getDevelopmentInfo.m		
fitness_inds	stage-based: randomly chosen from uniform distribution	n/a
dev (*mu_dev*)	dev_E, …, $dev_{N2} = a_i^{-1}(T\text{-}b)^{-c}$ dev_{N3}, …, $dev_{C5} = a_i^{-1}(T\text{-}b)^{-c}(F/(d+F))$ dev_{fem}, dev_{male} = NaN (a_i, b, c, d: Campbell *et al.,* 2001)	d^{-1}
CV_dev	$CV_{dev,E}$, …, $CV_{dev,C5}$ = 0.15 (avg. see section C.4.1) $CV_{dev,adult}$ = NaN	n.d.
devt_inds (*col.devt*)	sampled using Gamma distribution	d^{-1}
getEggProductionInfo.m		
epr (*mu_epr*)	EPR = f(T, F) = (6.3T + 21)(0.0016F + 0.35) (as per Neuheimer *et al.,* 2010a)	eggs $female^{-1}$ d^{-1}
CV_epr	0	eggs $female^{-1}$ d^{-1}
epr_inds	*mu_epr* (since *CV_epr* = 0, see *epr*)	eggs $female^{-1}$ d^{-1}
col.age_in_stg	0	d

*Are not required for this example but must have "placeholders" for the model to run. Should these be omitted the subroutine functions must be altered to remove corresponding sections.

Figure C.4 shows the dynamic nature of temperature, food, and female abundance used as forcings for this example. These are climatological means from Neuheimer et al., 2010a, neglecting variance to simply demonstratig the model's ability. The simulated C1 abundance, Figure C.3 (right), seems slightly erratic as compared to Fig. 6 in Neuheimer et al., 2010a, which may be due to the abundance based off of averaged forcings, while the mentioned literature produced annual results to provide an overall average. However, the general seasonal trend and magnitude is similar, confirming that the modified IBM can successfully run with dynamic forcings.

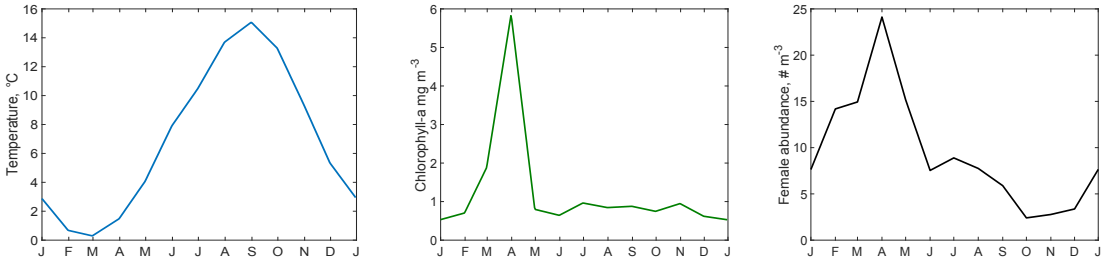


Figure C.4 Example C.4.2: Dynamic forcings; (left) temperature conditions, (middle) food in chlorophyll-a concentration, (right) female abundance over time (à la Fig. 4b, c, d Neuheimer et al., 2010a).

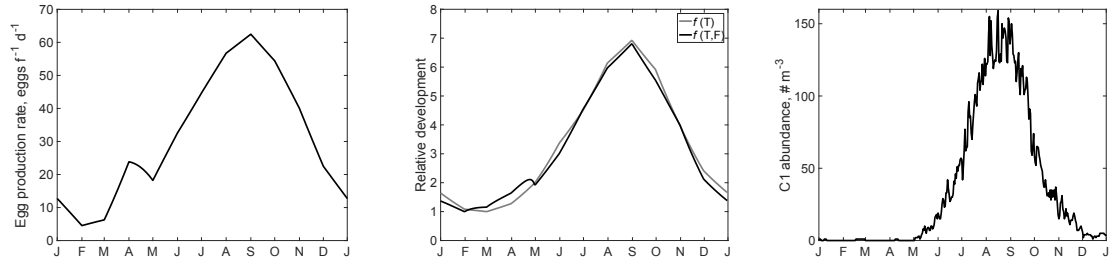


Figure C.3. Example C.4.2: Model results; (left) applied egg production rate, dependent on temperature and food, (middle) relative development, stage-independent relative to seasonal minimum in February, (right) simulated C1 abundance (à la Neuheimer et al., 2010a, Fig. 6 & 7).

C.4.3 Steady State Illustration

Simulating population dynamics over time subject to constant temperature and food conditions with no individual variability for vital rates

- This example details how to parameterize the modified IBM to illustrate a steady state situation (à la: Sim. 1 – 3.3 *IBM Simulations*, Gentleman et al., 2012).

Table C.5 Specification of variables in the IBM for Example C.4.3

Variables	Value/Formula	Units
getSimParam.m		
sim.length	100	d
sim.deltat	0.25	d
getStageIDs.m		
stg._: *EtoN6*, *C1, C2 ... C5* *fem* (i.e. *C6*)	1, 2, …, 7 (order matters; stages numbered chronologically)	n/a
getPopColIDs.m		
col._: *stages* *tspawn** *age_in_stg* *MCF* *fitness* *devt* *mort* *conditionID** *temp* *food*	1, 2, …, 10 (order does not matter; code sets up matrices dependant on these values)	n/a
getInitialConditions.m		
A	$A_{E\text{-}N6}$ (no data) = calculated (see code details below) A_{C1}, A_{C2}, … A_{fem}(data) = [6375, 4926, 3767, 2029, 1449, 2898]	individuals m^{-2}
*col.tspawn**	NaN	n/a
*col.conditionID**	NaN	n/a
col.age_in_stg	exponentially distributed across expected stage duration (see code details below)	d
col.MCF	exponentially distributed across stage (see code details below)	n.d.

getTempFood.m		
F_Chl	$F_{chl\text{-}a} = 1.45$	mgchl-*a*/m^3
C_to_Chl	$X_{C:chl\text{-}a} = 50$	C/chl-*a*
F_applied (*col.food*)	$F_C = X_{C:chl\text{-}a}F_{chl\text{-}a}$	mgC/m^3
T_applied (*col.temp*)	T = 5.1	°C
getMortalityRates.m		
mort_inds (*col.mort*)	stage-based: calculated using "Basic" mortality estimation method, see details below	d^{-1}
getDevelopmentInfo.m		
fitness_inds	stage-based: randomly chosen from uniform distribution	n/a
dev	$dev_{E\text{-}N6} = (1/dev_E + \ldots + 1/dev_{N6})^{-1}$ $dev_{C1}, dev_{C2}, \ldots dev_{C5}$ = as in section C.4.1 (see C.4.1 for details)	d^{-1}
CV_dev	0	n.d.
devt_inds (*col.devt*)	*mu_dev* (since *CV_dev* = 0, see "*dev*")	d^{-1}
getEggProductionInfo.m		
epr	EPR = f(T, F) = (6.3T + 21)(0.0016F + 0.35) (as per Neuheimer *et al.,* 2010b)	eggs female^{-1} d^{-1}
CV_epr	0	n.d.
epr_inds	*mu_epr* (since *CV_epr* = 0, see *epr*)	eggs female^{-1} d^{-1}
viability	0.97 a form of survivorship at spawn is applied, e.g. only 97% of the total number is added to the *pop*	n.d.
col.age_in_stg	spread over time-step: randomly chosen from uniform distribution	d

*Are not required for this example but must have "placeholders" for the model to run. Should these be omitted the subroutine functions must be altered to remove corresponding sections.

getInitialConditions.m code details

Examples of code required for *A*, *MCF*, and *age-within-stage* to ensure constant recruitment.

```
% The following example of code shows how to calculate abundance
for aggregate stage E-N6 based on copepodite data, specified
rates and steady state assumption (note: this is to be used
separately from the getInitialConditions.m script to calculate A_E-N6 where then the values for abundances are specified as an array
("A") within getInitialConditions.m):

      A = *initial copepodite abundance data*
      epr = *mean egg production rate*
      mort = *stage-based mean mortality rates*
      devt= *stage-based mean development rates*

      R_EtoN6 = epr * A(stg.fem);
      m_EtoN6 = mort(stg.EtoN6);
      devt_EtoN6 = devt(stg.EtoN6);

      A_EtoN6 = R_EtoN6 * (1 - exp(-m_EtoN6 /
devt_EtoN6))/m_EtoN6;

      A = round([A_EtoN6 A]);

% The following example of code shows how to define age-within-
stage and MCF so that individuals are "spread" across each stage
(i.e. at different levels of "development" within their assigned
stage, see figure below). This bit of code is to be placed after
individual mortality and developments have been assigned
(otherwise the stage-based rates must be defined prior to
calculation as "mort" and "devt", and the calculations for these
may be removed within the loop):

      for i = stg.EtoN6:stg.C5
      inds = pop(:,col.stages) == i;

      mort = mean(pop(inds, col.mort));
      devt = mean(pop(inds, col.devt));

      fr = exp(-mort/devt);
      inc = (fr - 1)/(A(i));
      p_age= 1:inc:(fr - inc);

      pop(inds, col.age_in_stg) = -log(p_age)/mort;
      pop(inds, col.MCF) = pop(inds, col.age_in_stg).*devt;
end

      females = pop(:, col.stages) = stg.fem;

    pop(females, col.MCF) = NaN;
```

getMortalityRates.m code details

Examples of code required for calculation of stage-based *mort* to ensure constant recruitment.

```
% Calculate mortality rates based on estimates from copepodite
data and specified rates (note: it is recommended to use this
separately from the getMortalityRates.m script to calculate
mortality rates where then the values are specified as an array
("mort" within getMortalityRates.m):

      A= *initial copepodite data*
      epr = *mean egg production rate*
      devt= *stage-based development rates*

% Mortality of E-N6 stage (no data) using Ratio Method, Eqn. 6 -
Gentleman et al., 2012:

      mguess = *make initial guess for mortality, suggestion: 0.1*

      A_i= A(stg.C1);% Get IC abundance for C1
      R_EtoN6 = epr * A(stg.fem);% Calc. "recruitment" into E-N6
      devt_EtoN6= devt(stg.EtoN6); % Get dev't rate for E-N6
      devt_i= devt(stg.C1); % Get dev't rate C1

      mort(stg.EtoN6) = fzero(@findRatio2Mort, mguess);

% Mortality of all other non-adult stages (data available) using
Basic Method, Eqn. 3 - Gentleman et al., 2012:

      mort(stg.C1) = mort(stg.EtoN6); % assuming m1 = m2

      R_i
      = R_EtoN6*exp(-mort(stg.EtoN6)*(1/(devt_EtoN6)+1/(devt_i)));

      for i = stg.C2:stg.C5
      A_i = A(i-1);
      devt_i = devt(i);

      mort(i) = fzero(@findBasicMort, mguess);

      R_i = R_i * exp(-mort(i)* 1/devt_i;
      end

% Mortality of females assuming 1:1 sex ratio at molt using Basic
Method, Eqn. 3 - Gentleman et al., 2012:
      mort(stg.fem)= 0.5*R_i/A(stg.fem);
```

Shown below in Figure C.4 are abundances of aggregate stages egg-nauplii and copepodites including females, relative to the total population size, as well as proportion of copepodite stages, over simulated time. These illustrate that the model may be parameterized to produce a steady state, as the overall population size stays relatively the same, as well as the stage proportions. In order to maintain a steady state, the amount of eggs produced each time step was limited with an "egg survivorship", because the eggs are not subject to mortality until the following time step (unlike the model used in Gentleman *et al.,* 2012).

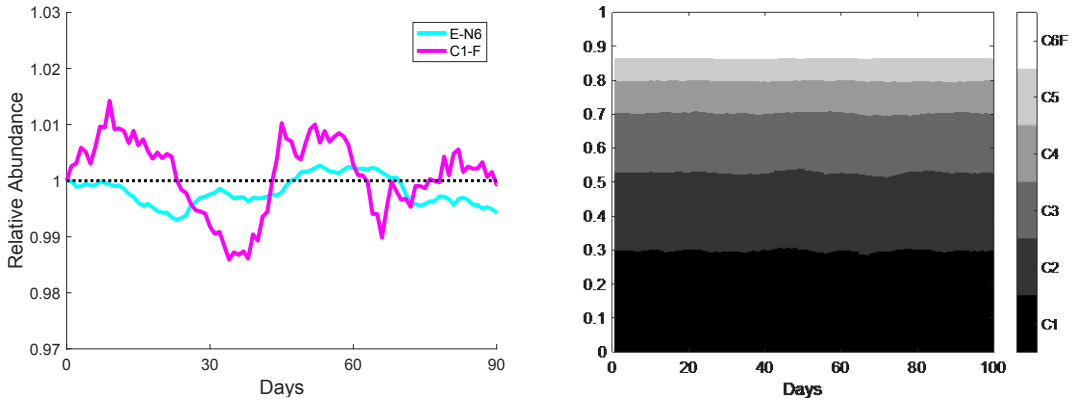


Figure C.4. Example C.4.3: Illustration of steady state situation; (left) relative aggregated abundances and (b) copepodite stage proportion over time (à la Gentleman et al., 2012, Fig. 6a, b)

C.4.4 Dynamic Conditions (in Time and Space)

Timing individual stage durations based on temporally and spatially dynamic temperature and food conditions with inherent individual variability in development.

- This example details how to parameterize the modified IBM using particle tracking data (or similar).

Table C.6. Specification of variables in the IBM for Example C.4.4

Variables	Value/Formula	Units
getPopColIDs.m		
col._: *stages* *tspawn** *age_in_stg* *MCF* *fitness* *devt* *mort* *conditionID* *temp* *food*	1, 2, …, 10 (order does not matter; code sets up matrices dependant on these values)	n/a
getInitialConditions.m		
*col.tspawn**	NaN	n/a
col.conditionID	Assign position ID numbers to new eggs. This corresponds to the "track" to be followed and conditions experienced over the simulation period.	n/a
getTempFood.m		
F_applied (*col.food*)	Read in data file (or hard-code in script) for full time-series, using the following format: [particle tracks (rows) x time (columns)] Additional code must be written to sample from the data series at given *t* over simulation period. (May need to interpolate if frequency of data and time step do not match).	mgC/m^3
T_applied (*col.temp*)	see *F_applied*; same requirements	°C
getEggProductionInfo.m		
col.conditionID (*newEggs_pop*)	Assign position ID numbers to new eggs. This corresponds to the "track" to be followed and conditions experienced over the simulation period.	n/a

*Are not required for this example but must have "placeholders" for the model to run. Should these be omitted the subroutine functions must be altered to remove corresponding sections.

Appendix D

Spatial Maps of Temperature, Chlorophyll-*a*, and Stage Development

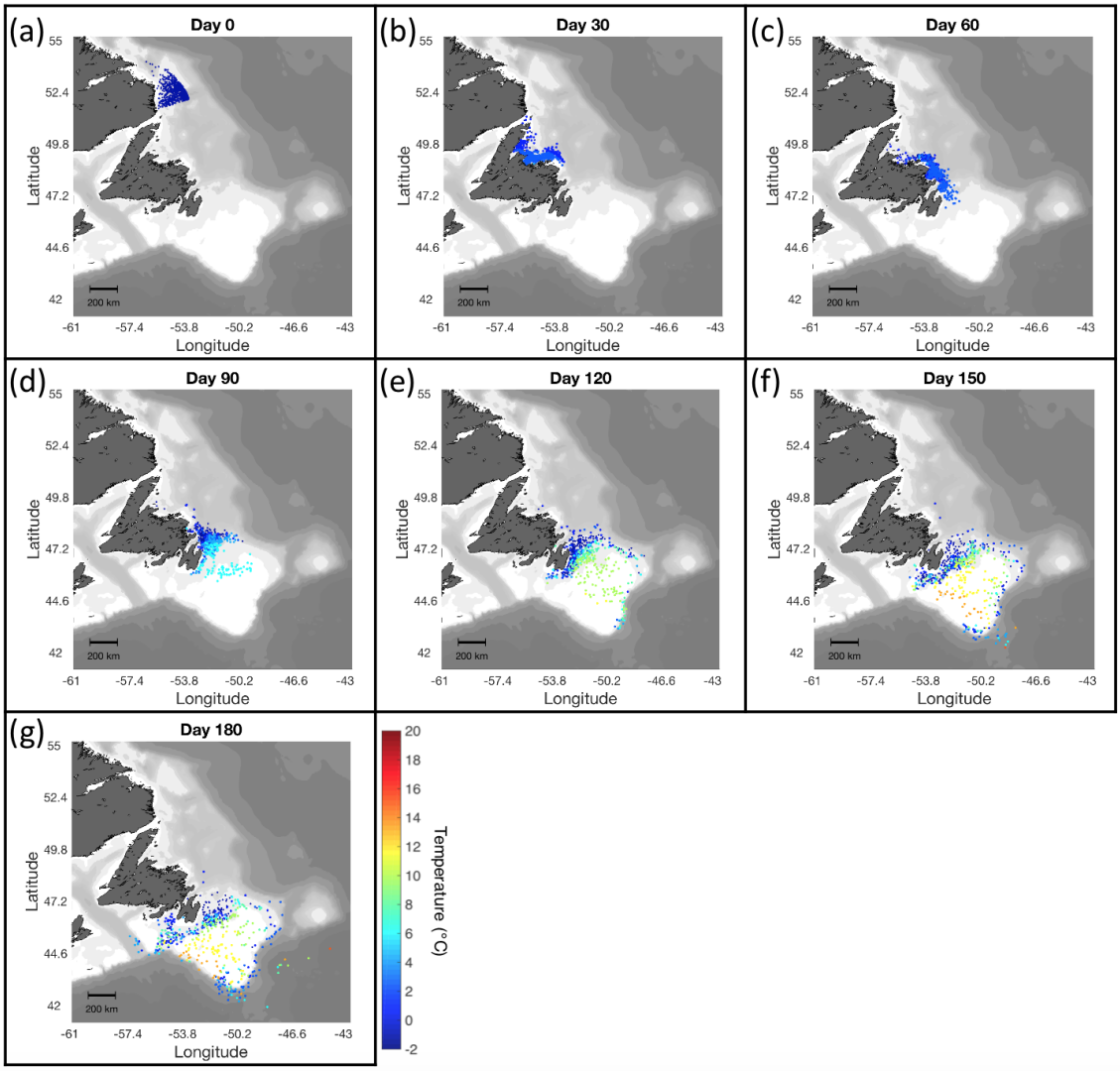


Figure D.1 Temperature particle tracks (Day 0 = April 1st run for 180 days) derived from 3D physical model (Han et al., 2008). Temperature level defined by colour bar in bottom row. Deeper water depth is marked with darker gray scale contour.

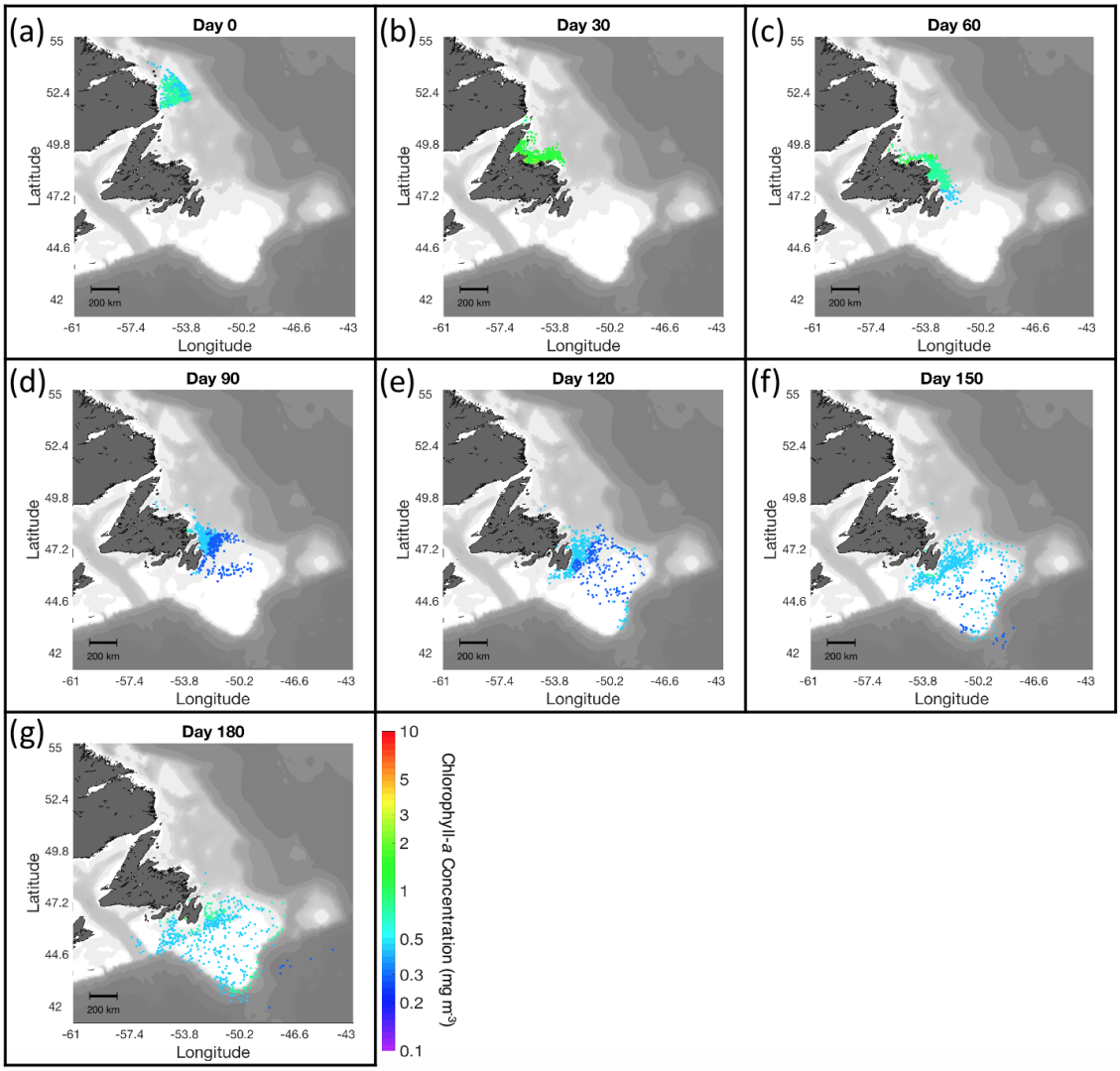


Figure D.2 Chlorophyll-*a* particle tracks (Day 0 = April 1st run for 180 days) produced by using data retrieved from the Glob Colour Project (ACRI-ST, 2017). Chlorophyll-*a* concentration defined by colour bar in bottom row. Deeper water depth is marked with darker gray scale contour.

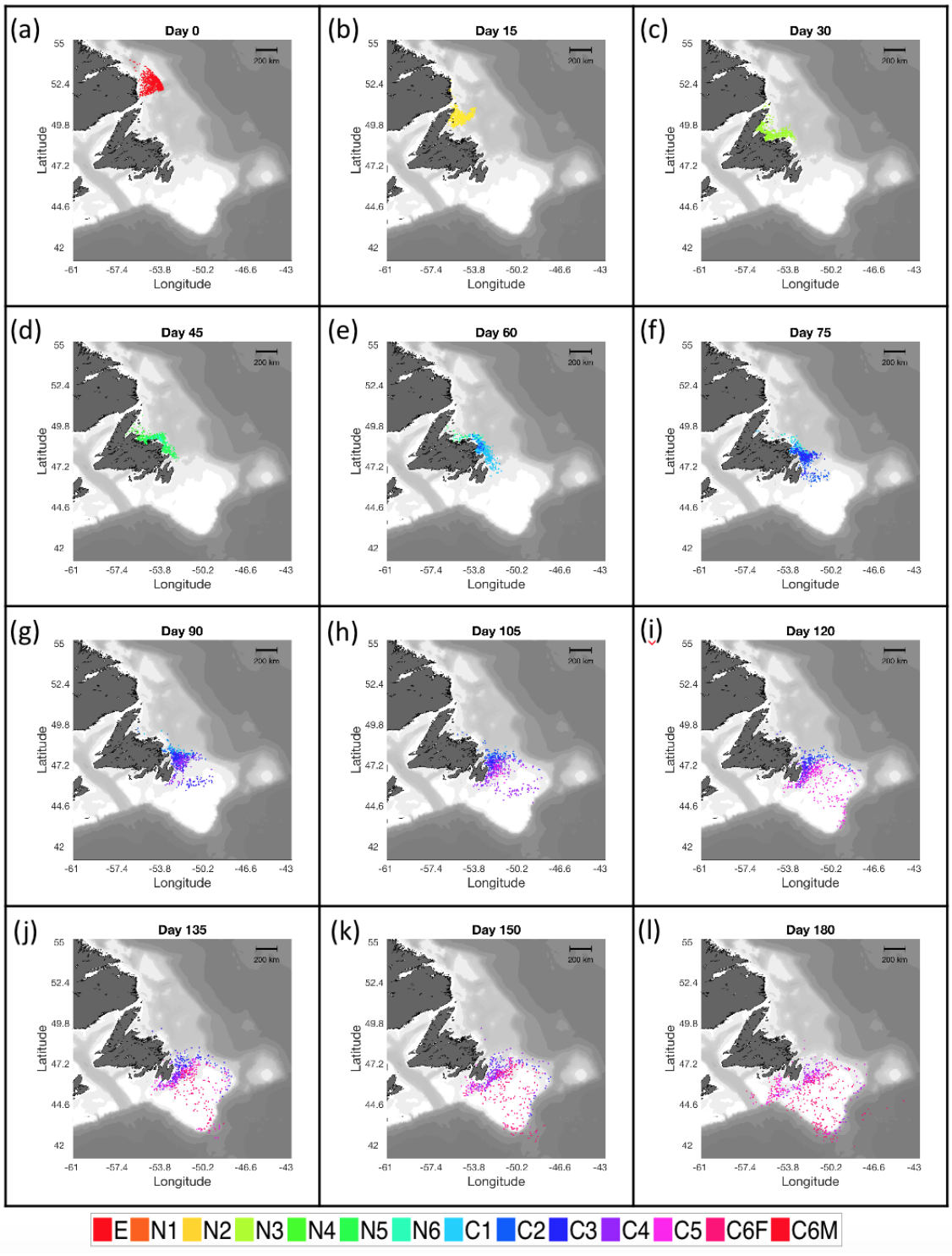


Figure D.3 Simulated Calanus finmarchicus population development over 180 days; initializing with eggs on each particle, neglecting individual variability and mortality. Development rates calculated using Eqn. 1.1 with temperature and chlorophyll-a tracks as shown in Figs. C.1-2. Stages are defined by colour bar in bottom row.

Appendix E

Simulated Population Dynamics using Particle Tracks

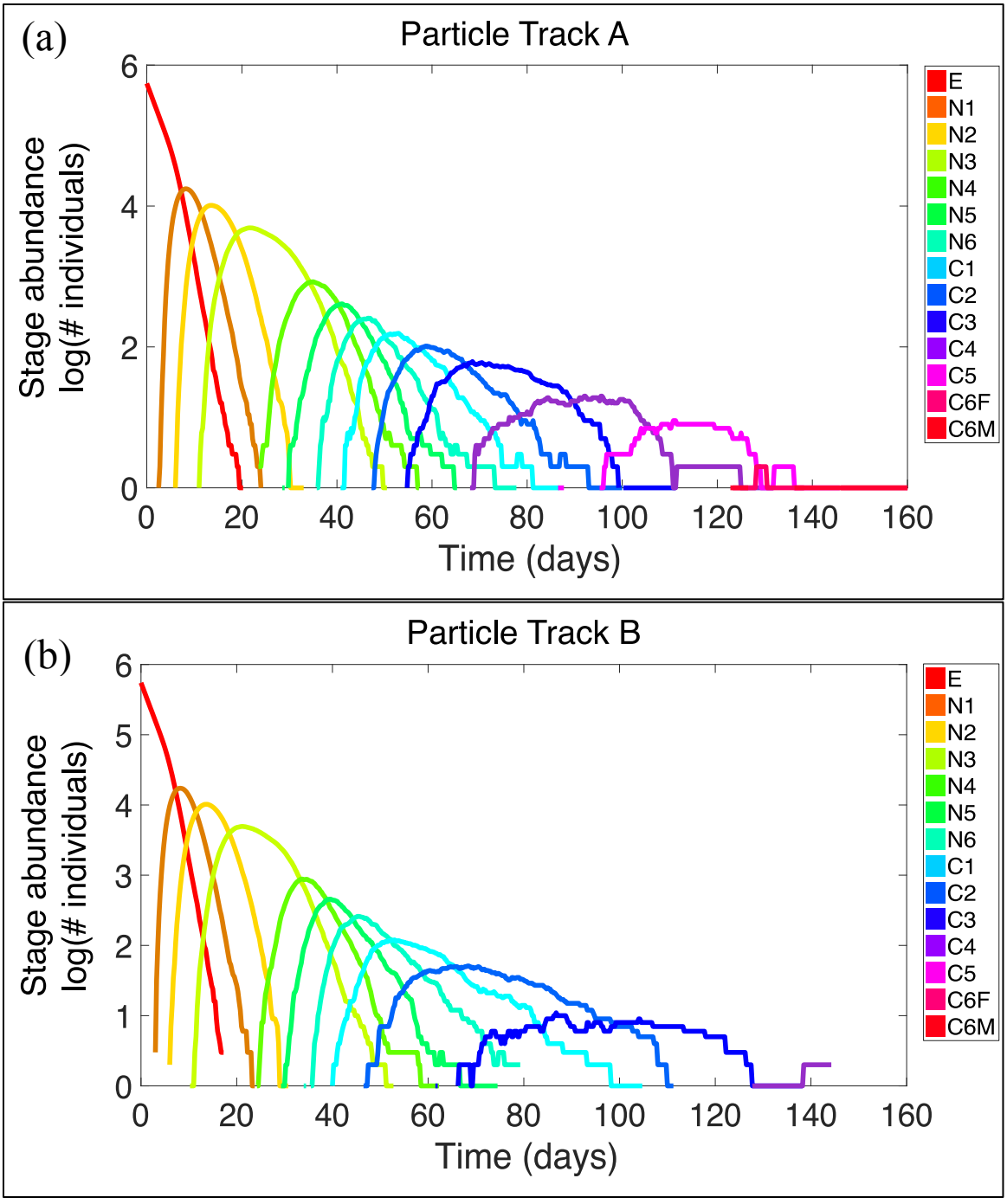


Figure E.1 Simulated stage abundance results (log scale) of Calanus finmarchicus population over time: based on environmental conditions from (a) Particle Track A, and (b) Particle Track B.

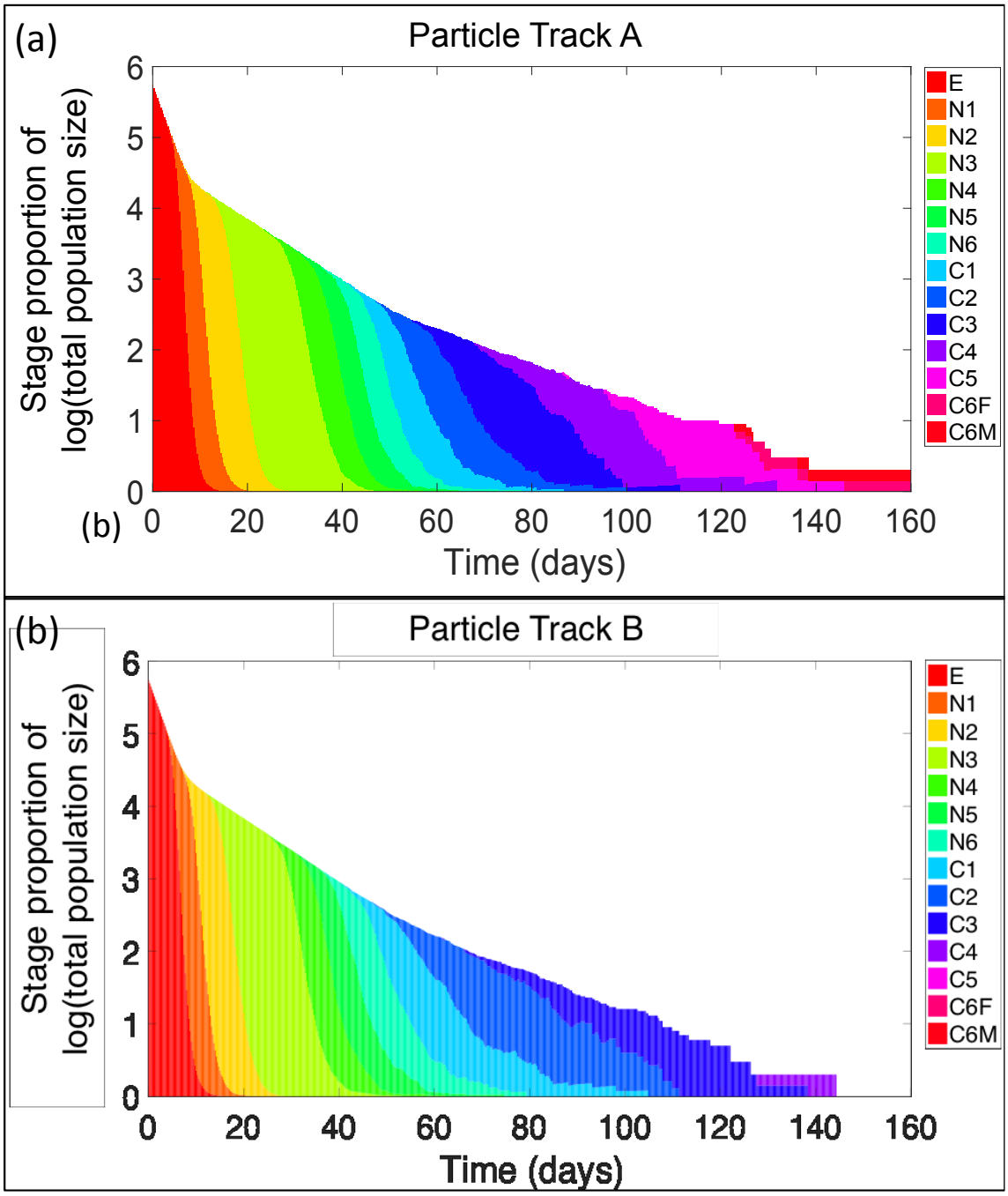


Figure E.2 Simulated stage proportion results of Calanus finmarchicus total population (log scale) over time: based on environmental conditions from (a) Particle Track A, and (b) Particle Track B.

Table E.1 Emergent stage durations (i.e., sample mean), sample variance, and number of individuals (i.e., sample size) results of population dynamics case study in Chapter 4. Stage durations unable to be calculated are marked by "n/a", as not available.

Stages	Particle Track A			Particle Track B		
	Sample Mean	Sample Variance	No. of Individuals	Sample Mean	Sample Variance	No. of Individuals
E	6.6	2.7	25841	6.6	2.6	25603
N1	4.8	2.1	15904	4.7	1.9	15764
N2	6.9	3.6	7911	6.6	3.4	8009
N3	14.4	12.2	1771	14.2	11.0	1804
N4	6.1	3.2	959	5.7	3.2	1013
N5	5.1	3.8	576	5.7	4.5	566
N6	5.7	3.9	317	7.8	8.8	232
Larval Duration	50.0	21.6	317	50.9	39.1	232
C1	7.6	9.6	204	12.5	31.7	105
C2	9.7	9.9	119	22.2	40.1	25
C3	16.5	34.9	44	44.3	54.9	2
C4	19.7	21.6	16	n/a	n/a	0
E to C5 Duration	105.2	123.8	16	n/a	n/a	0
C5	19.6	18.7	5	n/a	n/a	0
Generation Time	111.6	136.0	5	n/a	n/a	0