

Nonlinear mixed-effects modelling of *Ciona intestinalis*  
population growth, dependent upon abiotic conditions

By

Kieran J. Murphy

A thesis submitted to the Department of Biology and the Faculty of Graduate Studies  
Committee, in partial fulfillment of the requirements for the degree of Master of Science

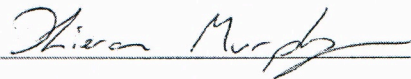
St. Francis Xavier University

Antigonish, Nova Scotia


© 2016

## Copyright Permission

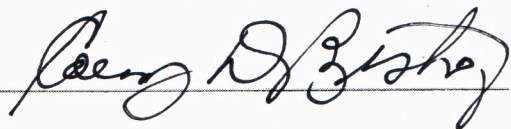
Permission is herewith granted to St. Francis Xavier University to have copied for non-commercial purposes, at its discretion, this thesis upon request of individuals and institutions. The author reserves other publication rights and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.



Kieran Murphy (Author)



Dr. Russell C. Wyeth (Co-Supervisor)



Dr. Cory D. Bishop (Co-Supervisor)

7 October 2016

Date

## Abstract

The tunicate *Ciona intestinalis* is a nuisance biofouling species due to its negative effects on the aquaculture industry, particularly mussel farming. I conducted a two-year observational study on the Atlantic coast of Nova Scotia to ascertain the relationship between abiotic conditions and the heterogeneous distribution and growth of *C. intestinalis* populations. Temperature, salinity, pH, and water flow were recorded with *in situ* data loggers at thirteen sites from May to October in 2014 and 2015, while *C. intestinalis* abundance was monitored on settlement plates. Nonlinear mixed-effects (NLME) modelling fit the most likely models to these data using a logistic growth function with parameter estimates determined by the abiotic variables using maximum likelihood estimation methods. *C. intestinalis*' response to the abiotic variables exhibited inter-annual and inter-site variation. The best model fit to the 2014 data contained salinity variables only, while the best model in 2015 contained the variables temperature, salinity, pH, and water flow. Intra-annual model validation and nested cross-validation resulted in c. 90 % model efficiency in prediction. The inter-annual difference in *C. intestinalis* response meant that model predictions of alternate years were variable with reduced efficiency. The observed variability in this system has underlined the importance of continued long-term monitoring. All of the abiotic variables added valuable information to the models, possibly due to the lack of a dominant variable that would reduce the value of variables contributing less information. The use of random effects improved model fit compared to fixed effects alone, exhibiting that NLME modelling is an effective tool to explain the growth of *C. intestinalis* in stochastic abiotic environments.

## Acknowledgements

This thesis is the sum of many parts, of which a great amount would not have been possible without the help of others. I have a lot of people to acknowledge and thank for their significant contributions. To my co-supervisors Dr. Russell Wyeth and Dr. Cory Bishop I thank you both for your guidance and contributions, without which the completion of this work would be a mere dot on the horizon. Thank you Dr Ricardo Scrosati for sitting on my committee, showing a keen interest in the project and always being ready to listen and advice. Dawn Sephton has been at the heart of this work from the beginning and I express my sincere gratitude to her for always being available to lend a helping hand and the guide me (as well as being my fix for batteries and cable-ties!).

This research was possible through funding from the Department of Fisheries and Oceans Canada and the Aquaculture Association of Nova Scotia. I'm sure there was a lot going on behind the scenes at the Bedford Institute of Oceanography and within the DFO that I wasn't aware of, but to name a few people, thank you to Edward Horne, Benedict Vercaemer, and John O'Reilly. A massive thank you is in order for everyone involved at the AANS, Danielle Goodfellow and Dr. Vikki Swan in particular. I also would not have been able to complete the work in the manner I would have liked without receiving a Nova Scotia Research and Innovation Graduate Scholarship.

I owe a huge thank you to everyone involved at the mussel farms I worked at. In particular, Peter Darnell at Indian Point Marine Farm, Darlene Meade, John Stairs, Kaija Lind and the whole crew at AquaPrime Mussel Ranch Ltd., and last but not least all the guys in Englishtown involved with Cape Breton Bivalves, especially Mac and Aaron.

I can't thank everyone at Hoskin Scientific Ltd. enough for going above and beyond to secure the improved replacement data loggers for the 2015 field season. I have to single out Natalia Lecki and Jon Matheson for their efforts, for which I will always be grateful.

A substantial debt is owed to Dr. Kerenaftali Klein for helping with the growth model development and her continued assistance when I was doing a lot of head scratching.

Of course I have a massive thank you for my amazing lab mate, Natalia Filip, despite her deserting me to follow her heart back to the mountains. There is nothing like a hands-on experience to understand what an introduced species feels like in a new environment, and I feel like Nat and I put each other through this experience with our thermostat wars (sorry

Russell). We had some wonderful company in the Wyeth lab from summer visitors. Thank you to Rebecca Kennedy, who helped me with all thing fieldwork related in 2014 and to Veronica Ells for her assistance in 2015!

Of course none of this would have been the same without the great people in the biology department, thank you all for everything. Thank you for making me feel so welcome here.

To any one not named above that I have chatted to about ideas, sought advice from, or had anything to do with the project, thank you so much.

## **Dedication**

*To my friends and family who have supported and encouraged me through thick and thin, I can always rely on you.*

*To Robert M. Thomas and Hobart D. Betts, the inventors of the cable tie. Need I say more?*

*To this thesis, I promise to love and to cherish you, for better or worse, until defence do us part.*

## Table of Contents

Copyright Permission.....	i
Abstract.....	ii
Acknowledgements.....	iii
Dedication.....	v
Table of Contents.....	vi
List of Tables.....	xi
List of Figures.....	xv
List of Abbreviations.....	xxiii
1. Introduction.....	1
1.1 Overview.....	1
1.2. Invasion ecology.....	2
1.2.1. The history and terminology.....	3
1.2.2. Introduced species: a precedent for invasion?.....	4
1.2.3. Transport and introduction of marine benthic invasive species.....	6
1.2.4. Gaining and maintaining dominance in a new environment.....	8
1.3. <i>Ciona intestinalis</i> .....	12
1.3.1. The colourful history of <i>C. intestinalis</i> .....	12
1.3.2. Distribution of <i>C. intestinalis</i> in Atlantic Canada.....	14
1.3.3. <i>C. intestinalis</i> biology.....	18
1.3.4. Why is <i>C. intestinalis</i> a successful invader?.....	23
1.3.5. <i>C. intestinalis</i> impact.....	25
1.4. Prevention and mitigation efforts.....	28
1.4.1. Prevention of introduction and establishment.....	28
1.4.2. Reduce spread and impact post-introduction.....	28
1.4.3. Long-term mitigation and control.....	29
1.5. Research objectives.....	30

2. Materials and methods .....	32
2.1. Field monitoring sites.....	32
2.1.1. Site selection.....	32
2.1.2. Site groupings for analyses.....	33
2.2. Biotic monitoring .....	34
2.2.1. Settlement collector design.....	34
2.2.2. Collector deployment: intra-site location and depth.....	35
2.2.3. Settlement plate monitoring.....	36
2.2.4. Settlement plate holder .....	36
2.2.5. Light-box .....	37
2.2.6. Image processing and analysis .....	38
2.3. Abiotic monitoring .....	39
2.3.1. Environmental data loggers .....	39
2.3.2. Data-logger calibration .....	40
2.3.3. Data-logger losses and repairs.....	40
2.5. Abiotic metrics .....	41
2.5.1. Mean metrics .....	41
2.5.2. Maximum and minimum metrics .....	41
2.6. Biotic and abiotic data plots .....	42
2.6.1. Biological data plots .....	42
2.6.2. Abiotic data plots.....	42
2.7. Statistical analysis .....	42
2.7.1. Correlation analysis .....	42
2.7.2. Information-theoretic inference.....	43
2.7.3. Preselection of abiotic variables .....	44
2.7.4. Mixed-effects modelling: fixed and random effects.....	44
2.7.5. Nonlinear mixed-effects (NLME) model .....	45



3. Results.....	57
3.1. Cumulative <i>C. intestinalis</i> abundance.....	57
3.1.1. Cumulative <i>C. intestinalis</i> growth in 2014.....	58
3.1.2. Cumulative <i>C. intestinalis</i> growth in 2015.....	58
3.1.3. Comparison between cumulative abundance in 2014 and in 2015.....	60
3.2. Monthly <i>C. intestinalis</i> abundance.....	61
3.2.1. Monthly <i>C. intestinalis</i> growth in 2014.....	61
3.2.2. Monthly <i>C. intestinalis</i> growth in 2015.....	61
3.2.3. Comparison between monthly abundance in 2014 and in 2015.....	64
3.3. Abiotic variables.....	64
3.3.1. Temperature.....	64
3.3.2. Salinity.....	68
3.3.3. pH.....	72
3.3.4. Water motion.....	75
3.4. Correlation Analysis Results.....	78
3.5. NLME model: 2014.....	79
3.5.1. Preselection of abiotic variables.....	79
3.5.2. Bottom-up and top-down model candidate sets.....	80
3.5.3. Observed values vs. fitted values.....	83
3.6. NLME model: 2015.....	90
3.6.1. Preselection of abiotic variables.....	90
3.6.2. Bottom-up and top-down model candidate sets.....	90
3.6.3. Observed values vs. fitted values.....	94
3.7. Model accuracy.....	103
3.8. Nested cross-validation.....	104
3.9. Inter-annual model predictions.....	105
4. Discussion.....	107

4.1. Summary of findings .....	107
4.2. Abiotic determination of population growth .....	108
4.2.1. Temperature .....	108
4.2.2. Salinity .....	108
4.2.3. Water flow .....	109
4.2.4. pH .....	110
4.3. Modelling approach.....	110
4.3.1. Mixed-effects use in marine modelling .....	110
4.3.2. Was NLME appropriate?.....	111
4.3.3. Bottom-up versus top-down approach.....	111
4.3.4. Model reliability: inference versus prediction .....	112
4.3.5. Sources of potential experimental error.....	112
4.4. Inter-site and inter-annual stochasticity .....	114
4.4.1. Genetic diversity .....	114
4.4.1.1. Suitability to the environment .....	115
4.4.1.2. Response to stress .....	116
4.4.2. Biotic interactions.....	117
4.5. Climate Change .....	119
4.6. Moving forward.....	120
4.6.1. Genetic sequencing and analysis .....	120
4.6.2. Improved NLME model input .....	121
4.6.3. Future experiments .....	123
4.6.4. Abiotic data calibration of online databases .....	124
4.7. Implications for management.....	124
5. Conclusions.....	127
References.....	129
Appendices.....	142

Appendix A: Deployment and monitoring schedules .....	142
Appendix B: Additional sampling procedures .....	143
B.1. Chlorophyll-a point-sample data.....	143
B.2. <i>Ciona intestinalis</i> sample collection for genetic analysis .....	146
Appendix C: Multivariate community analysis .....	147
C.1. Introduction.....	147
C.2. Methods.....	148
C.3. Results.....	150
C.4. Discussion .....	158
Appendix D: Raw abiotic data plots .....	159
Appendix E: 2014 and 2015 combined NLME model.....	174
E.1. Preselection of abiotic variables.....	174
E.2. Bottom-up and top-down candidate sets .....	176
E.3. Observed values vs. fitted values .....	179
Appendix F: Residual plots .....	181
F.1. Best temperature-only model (BTO) 2014.....	181
F.2. Best bottom-up model (BBU) 2015 .....	182
F.3. Best temperature-only model (BTO) 2015.....	183
F.4. Best bottom-up model (BBU) combined.....	184
F.5. Best top-down model (BTD) combined .....	185
F.6. Second best top-down model (SBTD) combined.....	186
Appendix G: Settlement collector images.....	187
Appendix H: ImageJ macro and R codes .....	189
H.1. ImageJ.....	189
H.2. R Code .....	190

## List of Tables

Table 2.1: The 16 sites monitored in 2014 and 2015 (sites in italics were only monitored in 2015), including the name, site code, type of site, and GPS co-ordinates. ....	32
Table 2.2: The number of collectors deployed in 2014 and in 2015. ....	36
Table 2.3: A description of all the environmental data-loggers used in 2014 and in 2015. The data loggers used to measure each abiotic variable are specified by year. The data loggers were used at all sites unless specified in brackets. The full logger names follow: WQL = WTW WQL-pH Logger; CT2X = INW Aquistar® CT2X Conductivity Smart Sensors; HOBO-G Accelerometer = Onset HOBO Pendant® G Data Logger; HOBO-Temp. = Onset HOBO Pendant® Temperature/Light Data Logger; and YSI 600XLM = YSI 6-Series 600XLM Multiparameter Water Quality Sonde. *The default temperature data logger in 2014 was the CT2X, but WQL recordings were used when CT2X was not available. ....	39
Table 3.1: Summary temperature values for all 15 sites (in alphabetical order) and both years (where applicable). The 2014 values, where available are shaded in light grey, while IN and LB are shaded in dark grey. Mean = Overall mean of daily mean values. Minimum = Minimum daily mean value. Maximum = Maximum daily mean value. Variance = Variance of daily mean values. ....	65
Table 3.2: Summary salinity values for all 15 sites (in alphabetical order) and both years (where applicable). The 2014 values, where available are shaded in light grey, while IN and LB are shaded in dark grey. Mean = Overall mean of daily mean values. Minimum = Minimum daily mean value. Maximum = Maximum daily mean value. Variance = Variance of daily mean values. ....	69
Table 3.3: Summary pH values for all 15 sites (in alphabetical order) and both years (where applicable). The 2014 values, where available are shaded in light grey, while IN and LB are shaded in dark grey. Mean = Overall mean of daily mean values. Minimum = Minimum daily mean value. Maximum = Maximum daily mean value. Variance = Variance of daily mean values. ....	72
Table 3.4: Summary water motion values for all 15 sites (in alphabetical order) and both years (where applicable). The 2014 values, where available are shaded in light grey, while IN and LB are shaded in dark grey. Mean = Overall mean of daily variance. Minimum = Minimum daily variance. Maximum = Maximum daily variance. Variance = Variance of daily variance. ....	75

Table 3.5: Results from Kendall’s Rank Correlation Coefficient ( $\tau$ ) tests between site-specific mean *C. intestinalis* for each measurement period and the corresponding site-specific abiotic variable measurement period values. .... 78

Table 3.6: Stepwise removal of a single abiotic metric from the most correlated pair using the 2014 abiotic metrics. The pair of metrics most correlated at each step is highlighted. The light grey shaded metric was kept (Keep), while the metric shaded with dark grey was removed (Rem.) from the process. There was a selection rule to always have one metric from each variable remaining. .... 79

Table 3.7: The ten best models (30 models in total) from the bottom-up model development approach (2014), ranked by lowest AICc value. \*Interaction term included between the variables in this model parameter. The site-level effect for all candidate models was specified as MaxPop only. K = no. model parameters.  $\Delta_i$  = AICc difference. Model Lik. = model likelihood.  $w_i$  = Akaike weight. The site-level effects for all candidate models was specified as MaxPop only..... 81

Table 3.8: The ten best models (total N = 29) from the top-down model fitting approach for the 2014 data only. The models are ranked by their AICc value, a lower value being better. The site-level effects for all candidate models were specified as MaxPop only. K = no. model parameters.  $\Delta_i$  = AICc difference. Mod. Lik. = model likelihood. Cumul.  $w_i$  = cumulative Akaike weight. .... 82

Table 3.9: Estimates, with their associated standard error (SE), for each of the species-wide and site-level model parameters from the best model in 2014 (bottom-up and top-down model building methods provided the same best model)..... 83

Table 3.10: Stepwise removal of a single abiotic variable metric from the most highly correlated pair using the 2015 abiotic data. The pair of metrics most correlated at each step is highlighted. The light grey shaded metric was kept (Keep) while the metric shaded with dark grey was removed (Rem.) from the process. There was a selection rule to always have one metric from each variable remaining. Only pairwise complete observations were used in the correlation. .... 90

Table 3.11: The ten best models (n = 34) from the bottom-up model candidate set from the 2015 data, ranked by lowest AICc value. The abiotic metric abbreviations are explained as: Temp. = mean temperature; Sal. = max salinity; pH = minimum pH; and Accel. = maximum variance in acceleration. K = number of model parameters.  $\Delta_i$  = AICc difference.  $w_i$  = Akaike weight..... 92

Table 3.12: The ten best models from the 2015 top-down candidate set (n = 29), ranked by AICc value. Each model parameter (MaxPop, D50, and Lag) could have all four of the abiotic metrics in it in any given model. The metrics that were included in each model are shown in the table, with any grey filled cells showing which metrics were excluded from a given model. The abiotic metrics in the table have been assigned shorthand names as follows: Temp = mean temperature; Sal = maximum salinity; pH = minimum pH; and Accel = maximum variance in acceleration. K = number of model parameters.  $\Delta_i$  = AICc difference.  $w_i$  = Akaike weight. \*MaxPop is the only site-level effect in all models within this candidate set. ....93

Table 3.13: Estimates (est.), with associated standard error (SE), for each of the species-wide (Sp.) and site-level (Site) model parameters from the best model, for both the bottom-up (BBU) and the top-down (BTD) model building approaches using 2015 data. The absolute scaled effect size (ASES) is provided for each abiotic metric parameter. Temp. = mean temperature - 14; Sal. = maximum salinity - 30; pH = minimum pH - 8; Accel. = maximum variance in acceleration.....94

Table 3.14: Model accuracy results for the best bottom-up (BBU) and best top-down (BTD) models for 2014 and 2015. The four methods of model assessment are: mean absolute error (MAE), measured in units of *C. intestinalis* %; root mean square error (RMSE), measured in units of *C. intestinalis* %; model efficiency (MEF), a measure of the variance explained by the model [similar to R<sup>2</sup> (Amaro et al., 1998)]; and the Spearman rank correlation coefficient ( $\rho$ ). Full = The model with both species-wide and site-level effects. SW = Species-wide effects only. BTO = Best temperature-only model. BSO = Best salinity-only model. Best BU & TD 2014 model is also the BSO for 2014..... 103

Table 3.15: Model fit assessment results from the nested cross-validation procedure, as measured by MAE, RMSE, MEF, and  $\rho$ . .... 105

Table 3.16: Assessment of the results from model predictions carried out on ‘test’ datasets that were wholly or partially not included in the initial model fitting. .... 106

Table B.1: Chl-*a* measurement results from fluorometry analysis presented for each site and date of measurement. .... 144

Table C.1: Best abiotic variable combinations at explaining the cumulative biological community data from 2014, with measurement period as a within-level factor in the BIOENV procedure. Dataset contains missing abiotic values..... 150

Table C.2: Best abiotic variables at explaining the cumulative biological community data from 2014 with measurement period as a factor within the BIOENV procedure. Dataset contains no missing abiotic values (replacement via EM algorithm). ..... 152

Table C.3: Marginal test results from the DistLM analysis in 2014..... 153

Table C.4: Best models in 2014 from the DistLM procedure according to AICc ranking.... 154

Table C.5: Best abiotic variable combinations at explaining the cumulative biological community data from 2015, with measurement period as a within-level factor in the BIOENV procedure..... 155

Table C.6: Marginal test results from the DistLM analysis in 2015..... 156

Table C.7: Best models in 2015 from the DistLM procedure according to AICc ranking.... 157

Table E.1: Stepwise removal of a single abiotic variable metric from the most highly correlated pair using the combined 2014 and 2015 abiotic data. The pair of metrics most correlated at each step is highlighted. The light grey shaded metric was kept (Keep) while the metric shaded with dark grey was removed (Rem.) from the process. There was a selection rule to always have one metric from each variable remaining. Only pairwise complete observations were used in the abiotic metric correlations. .... 175

Table E.2: The seven best models (total no. models = 32) from the bottom-up model approach (combined), ranked by lowest AICc value. Each model parameter (MaxPop, D50, and Lag) could have all four of the abiotic metrics in it in any given model. The abiotic metrics in the table have been assigned shorthand names as follows: Mean temp. = mean temperature; Temp = minimum temperature; Sal = minimum salinity; pH = maximum pH; and Accel = maximum variance in acceleration. K = no. model parameters. .... 177

Table E.3: The six best models (total no. models = 29) from the top-down model approach (combined), ranked by lowest AICc value. Each model parameter (MaxPop, D50, and Lag) could have all four of the abiotic metrics in it in any given model. The abiotic metrics in the table have been assigned shorthand names as follows: Temp = minimum temperature; Sal = minimum salinity; pH = maximum pH; and Accel = maximum variance in acceleration. K = no. model parameters. .... 178

## List of Figures

Figure 1.1: A simple model of the invasion process showing the distinct stages, reproduced from Lockwood et al. (2013) with the permission of John Wiley and Sons. ....	4
Figure 1.2: Description of potential mechanisms involved in the rise to, establishment of, and loss of dominance reproduced from Wootton (2002) with the permission of Springer. ....	9
Figure 1.3: Abundance (represented as % cover on standard monitoring collectors) of <i>C. intestinalis</i> at NS coastal and inland monitoring sites in 2010, obtained through DFO AIS surveys (Reproduced from Sephton et al. (2014) with permission). ....	16
Figure 1.4: Abundance (represented as % cover on standard monitoring collectors) of <i>C. intestinalis</i> at NS coastal and inland monitoring sites in 2011, obtained through DFO AIS surveys (Reproduced from Sephton et al. (2015) with permission). ....	16
Figure 1.5: Abundance (represented as % cover on standard monitoring collectors) of <i>C. intestinalis</i> at NS coastal and inland monitoring sites in 2012, obtained through DFO AIS surveys (Pers. comm. Sephton D., with permission from the DFO). ....	17
Figure 1.6: Abundance (represented as % cover on standard monitoring collectors) of <i>C. intestinalis</i> at NS coastal and inland monitoring sites in 2013, obtained through DFO AIS surveys (Pers. comm. Sephton D., with permission from the DFO). ....	17
Figure 1.7: Abundance (represented as % cover on standard monitoring collectors) of <i>C. intestinalis</i> at NS coastal and inland monitoring sites in 2014, obtained through DFO AIS surveys (Pers. comm. Sephton D., with permission from the DFO). ....	18
Figure 1.8: Left - <i>C. intestinalis</i> individuals finding kelp to be a suitable substrate. Right – <i>C. intestinalis</i> completely covering a settlement plate, out of the water. ....	26
Figure 2.1: Distribution map of the sites monitored during this project. The 13 sites with grey markers were monitored in 2014 and 2015 (Core Sites), while the 3 sites with black markers were only monitored in 2015 (New Sites). ....	33
Figure 2.2: Left – The setup of the individual settlement plates; Right – Examples of the entire collector. ....	34
Figure 2.3: Top left – Exterior of the lightbox; top right – Interior of the lightbox; bottom right – End-on view of the settlement plate holder; and bottom left – Side-on view of the settlement plate holder. ....	37
Figure 2.4: An example of a settlement plate image, cropped, scaled and ready for percentage cover analysis with the grid overlay applied in ImageJ. ....	38



Figure 2.5: A simplified description of the NLME model development process from data collection to model predictions, represented by a flow diagram. Grey dashed arrows and grey-fill boxes indicate stochastic processes that can vary within a given candidate set of models, while black solid arrows and white-fill boxes are invariant processes within the given candidate set of models. TD = top-down and BU = bottom-up.....47

Figure 2.6: Nine example empirical plots that describe how the three model parameters can change the size and shape of the growth curve. D50 has a different value in each of the three columns of plots, while MaxPop has a different value in each of the three rows of plots. Each growth curve line in the individual plots represents a different Lag value.....49

Figure 3.1: *C. intestinalis* growth curves for all 13 sites throughout the 2014 monitoring period. The individual plate abundance is displayed (black circles), with a growth curve (black line) plotted through the mean of each measurement period. ....57

Figure 3.2: *C. intestinalis* growth curves for all 15 sites throughout the 2015 monitoring period. The individual plate abundance is displayed (black circles), with a growth curve (black line) plotted through the mean of each measurement period. ....59

Figure 3.3: The monthly *C. intestinalis* abundance for each measurement period in 2014, recorded independently of all other measurement periods on new settlement plates deployed at each site every visit. Sites ordered with the highest overall-mean *C. intestinalis* abundance at the bottom. ....62

Figure 3.4: The monthly *C. intestinalis* abundance for each measurement period in 2015, recorded independently of all other measurement periods on new settlement plates deployed at each site every visit. Sites ordered with the highest overall-mean *C. intestinalis* abundance at the bottom. ....63

Figure 3.5: Left - The temperature profiles of the core sites monitored in 2014 are presented by a heatmap of the daily mean values. The sites are ordered by the overall mean of daily means, with the highest overall mean at the top. \* The white sections at the start of ZZ and VC plots are missing data points. Right – The mean *C. intestinalis* abundance for each measurement period in 2014 represented as a heatmap, with sites corresponding to the appropriate site from the temperature plot on the left. \* The white section at the start of ZZ is a missing datum point. ....66

Figure 3.6: Left - The temperature profiles of the expanded sites monitored in 2015 are presented by a heatmap of the daily mean temperature values. The sites are ordered by their overall mean temperature, with the highest overall mean at the top. Right – The mean *C.*

*intestinalis* abundance for each measurement period in 2015 represented as a heatmap, with sites corresponding to the appropriate site from the temperature plot on the left.....67

Figure 3.7: Left - The salinity profiles of the core sites monitored in 2014 are presented by a heatmap of the daily mean values. The sites are ordered by their overall mean salinity, with the highest overall mean at the top. \* The white sections present for all site profiles are missing data points. Right – The mean *C. intestinalis* abundance for each measurement period in 2014 represented as a heatmap, with sites corresponding to the appropriate site from the salinity plot on the left. \* The white section at the start of ZZ is a missing datum point. .70

Figure 3.8: Left - The salinity profiles of the expanded sites monitored in 2015 are presented by a heatmap of the daily mean salinity values. The sites are ordered by their overall mean salinity, with the highest overall mean at the top. \* The white sections present in the LB and IN profiles are missing data points. Right – The mean *C. intestinalis* abundance for each measurement period in 2015 represented as a heatmap, with sites corresponding to the appropriate site from the salinity plot on the left. ....71

Figure 3.9: Left - The pH profiles of the core sites monitored in 2014 are presented by a heatmap of the daily mean pH values. The sites are ordered by their overall mean pH, with the highest overall mean at the top. \* The white sections present for multiple site profiles are missing data points. Right – The mean *C. intestinalis* abundance for each measurement period in 2014 represented as a heatmap, with sites corresponding to the appropriate site from the pH plot on the left. \* The white section at the start of ZZ is a missing datum point. ....73

Figure 3.10: Left - The pH profiles of the expanded sites monitored in 2015 are presented by a heatmap of the daily mean pH values. The sites are ordered by their overall mean pH, with the highest at the top. \* The white sections present in the LB and IN profiles are missing data points. Right – The mean *C. intestinalis* abundance for each measurement period in 2015 shown in a heatmap, with sites corresponding to the appropriate site from the pH plot on the left. ....74

Figure 3.11: Left - The water motion profiles of the core sites monitored in 2014 are presented by a heatmap of the daily variance in water motion values. The sites are ordered by the overall mean variance in water motion, with the highest at the top. \* The white sections at the start of the profiles for all sites (except PB) are missing data points. Right – The mean *C. intestinalis* abundance for each measurement period in 2014 represented as a heatmap, with sites corresponding to the appropriate site from the water motion plot on the left. \* The white section at the start of ZZ is a missing datum point. ....76

Figure 3.12: Left - The water motion profiles of the expanded sites monitored in 2015 are presented by a heatmap of the daily variance in water motion values. The sites are ordered by the overall mean variance in water motion, with the highest at the top. \* The white sections at the start of the VC profile are missing data points. Right – The mean *C. intestinalis* abundance for each measurement period in 2015 shown in a heatmap, with sites corresponding to the appropriate site from the water motion plot on the left.....77

Figure 3.13: Full-model fitted *C. intestinalis* relative abundance (black dashed line) for the 11 sites, fit by the 2014 best model (same model for BU and TD), and the fitted values using species-wide effect only (black dotted line). The observed *C. intestinalis* values (red line and circle markers) are overlaid. ZZ had a delayed deployment.....84

Figure 3.14: Full-fit from the 2014 best model (solid green line), with 95 % confidence intervals (dashed black lines), and 95 % prediction intervals (dotted black lines).....85

Figure 3.15: BBU and BTD 2014 residual boxplots by site, ordered alphabetically. The full-model residuals are displayed in the upper plot, while the species-wide effects residuals in isolation are displayed in the lower plot. ....86

Figure 3.16: The full-model residuals for the BBU and BTD 2014 plot against days after deployment.....87

Figure 3.17: Mean salinity, the independent variable used in the BBU and BTD 2014 plotted against the site-level effects to inspect their relationships in respect to homogeneity, with the correlation coefficient presented at the top of the plot. ....88

Figure 3.18: Full-model fitted *C. intestinalis* relative abundance (black dashed line) for the 11 sites, fit by the 2014 best temperature-only model (BTO), and the fitted values using species-wide effect only (black dotted line). The observed *C. intestinalis* values (red line and circle markers) are overlaid. ZZ had a delayed deployment.....89

Figure 3.19: Full-model fitted *C. intestinalis* relative abundance (black dashed line) for the 11 sites, fit by the 2015 best bottom-up model (BBU), and the fitted values using species-wide effect only (black dotted line). The observed *C. intestinalis* values (red line and circle markers) are overlaid. ....95

Figure 3.20: Full-model fitted *C. intestinalis* relative abundance (black dashed line) for the 11 sites, fit by the 2015 best top-down model (BTD), and the fitted values using species-wide effect only (black dotted line). The observed *C. intestinalis* values (red line and circle markers) are overlaid. ....97

Figure 3.21: Full-fit by the 2015 BTD (solid green line), with 95 % confidence intervals (dashed black lines), and 95 % prediction intervals (dotted black lines).....98

Figure 3.22: BTD 2015 residual boxplots by site, ordered alphabetically. The full-model residuals are displayed in the upper plot, while the species-wide effects residuals in isolation are displayed in the lower plot. ....99

Figure 3.23: The full-model residuals for the BTD 2015 plotted against days after deployment..... 100

Figure 3.24: Predictor variables used in the BTD 2015 plotted against the site-level effects to inspect their relationships, with the correlation coefficient presented at the top of each plot. .... 101

Figure 3.25: Full-model fitted *C. intestinalis* relative abundance (black dashed line) for the 11 sites, fit by the 2015 best temperature-only model (BTO), and the fitted values using species-wide effect only (black dotted line). The observed *C. intestinalis* values (red line and circle markers) are overlaid. .... 102

Figure C.1: The distribution of rho values from the 999 permutation tests carried out for the 2014 BIOENV procedure using the dataset with missing values. The rho value (dotted vertical line) of the best predictor (Mean salinity, Table C.1) variable and its significance are displayed. .... 152

Figure C.2: The distribution of rho values from the 999 permutation tests carried out for the 2014 BIOENV procedure using the dataset with no missing values (missing data replaced using EM algorithm). The rho value (dotted vertical line) of the best predictor variables (Mean salinity + Mean temp, Table C.2) and its significance are displayed..... 153

Figure C.3: Ordination of the 2014 cumulative biological data according to distance-based redundancy analysis using the abiotic data as predictor variables. A vector overlay represents the distance and proportion of the abiotic predictor effects from the best model (Table C.3). .... 154

Figure C.4: The distribution of rho values from the 999 permutation tests carried out for the 2015 BIOENV procedure using the dataset. The rho value (dotted vertical line) of the best predictor variable (Mean temp, Table C.5) and its significance are displayed..... 156

Figure C.5: Ordination of the 2015 cumulative biological data according to distance-based redundancy analysis using the abiotic data as predictor variables. A vector overlay represents the distance and proportion of the abiotic predictor effects from the best model (Table C.5). .... 157

Figure D.1: 5-minute interval data for each of the four abiotic variables measured in 2014 and in 2015 at Camp Cove (CC). The 2014 plots are in the first column, while 2015 plots are in

the second. From the top row to the bottom, the abiotic variables are as follows: temperature; salinity; pH; and water motion..... 159

Figure D.2: 5-minute interval data for each of the four abiotic variables measured in 2014 and in 2015 at Cape Canso (ZZ). The 2014 plots are in the first column, while 2015 plots are in the second. From the top row to the bottom, the abiotic variables are as follows: temperature; salinity; pH (2014 missing); and water motion..... 160

Figure D.3: 5-minute interval data for each of the four abiotic variables measured in 2014 and in 2015 at Dingwall (DW). The 2014 plots are in the first column, while 2015 plots are in the second. From the top row to the bottom, the abiotic variables are as follows: temperature; salinity; pH; and water motion..... 161

Figure D.4: 5-minute interval data for each of the four abiotic variables measured in 2014 and in 2015 at Fall’s Point (FP). The 2014 plots are in the first column, while 2015 plots are in the second. From the top row to the bottom, the abiotic variables are as follows: temperature; salinity; pH; and water motion..... 162

Figure D.5: 5-minute interval data for each of the four abiotic variables measured in 2015 at Ingomar (IN). From the top row to the bottom, the abiotic variables are as follows: temperature; salinity; pH; and water motion. .... 163

Figure D.6: 5-minute interval data for each of the four abiotic variables measured in 2014 and in 2015 at Indian Point (IP). The 2014 plots are in the first column, while 2015 plots are in the second. From the top row to the bottom, the abiotic variables are as follows: temperature; salinity; pH; and water motion..... 164

Figure D.7: 5-minute interval data for each of the four abiotic variables measured in 2015 at Louisbourg (LB). From the top row to the bottom, the abiotic variables are as follows: temperature; salinity; pH; and water motion. .... 165

Figure D.8: 5-minute interval data for each of the four abiotic variables measured in 2014 and in 2015 at Little River (LR). The 2014 plots are in the first column, while 2015 plots are in the second. From the top row to the bottom, the abiotic variables are as follows: temperature; salinity; pH; and water motion..... 166

Figure D.9: 5-minute interval data for each of the four abiotic variables measured in 2014 and in 2015 at Port Bickerton (PB). The 2014 plots are in the first column, while 2015 plots are in the second. From the top row to the bottom, the abiotic variables are as follows: temperature; salinity; pH; and water motion..... 167

Figure D.10: 5-minute interval data for each of the four abiotic variables measured in 2014 and in 2015 at Petit-de-Grat (PG). The 2014 plots are in the first column, while 2015 plots are

in the second. From the top row to the bottom, the abiotic variables are as follows: temperature; salinity; pH; and water motion. ....	168
Figure D.11: 5-minute interval data for each of the four abiotic variables measured in 2014 and in 2015 at St. Ann’s Bay (SA). The 2014 plots are in the first column, while 2015 plots are in the second. From the top row to the bottom, the abiotic variables are as follows: temperature; salinity; pH; and water motion. ....	169
Figure D.12: 5-minute interval data for each of the four abiotic variables measured in 2014 and in 2015 at Ship Harbour (SH). The 2014 plots are in the first column, while 2015 plots are in the second. From the top row to the bottom, the abiotic variables are as follows: temperature; salinity; pH; and water motion. ....	170
Figure D.13: 5-minute interval data for each of the three abiotic variables measured in 2015 at St Peters (SP). From the top row to the bottom, the abiotic variables are as follows: temperature; salinity; and water motion. ....	171
Figure D.14: 5-minute interval data for each of the four abiotic variables measured in 2014 and in 2015 at Venus Cove (VC). The 2014 plots are in the first column, while 2015 plots are in the second. From the top row to the bottom, the abiotic variables are as follows: temperature; salinity; pH; and water motion. ....	172
Figure D.15: 5-minute interval data for each of the four abiotic variables measured in 2014 and in 2015 at Wedgeport (WP). The 2014 plots are in the first column, while 2015 plots are in the second. From the top row to the bottom, the abiotic variables are as follows: temperature; salinity; pH; and water motion. ....	173
Figure D.16: 5-minute interval data for each of the four abiotic variables measured in 2014 and in 2015 at Yarmouth Bar (YB). The 2014 plots are in the first column, while 2015 plots are in the second. From the top row to the bottom, the abiotic variables are as follows: temperature; salinity; pH; and water motion. ....	174
Figure E.1: Red line and circle point markers represent the observed values. Dashed black line and circle markers represent the fitted values. The black dotted line and triangle markers represent the species-wide effect only fitted values. ....	180
Figure E.2: Red line and circle point markers represent the observed values. Dashed black line and circle markers represent the fitted values. The black dotted line and triangle markers represent the species-wide effect only fitted values. ....	181
Figure F.1: BTO 2014 residual boxplots by site, ordered alphabetically. The full-model residuals are displayed in the upper plot, while the species-wide effects residuals in isolation are displayed in the lower plot. ....	182

Figure F.2: BBU 2015 residual boxplots by site, ordered alphabetically. The full-model residuals are displayed in the upper plot, while the species-wide effects residuals in isolation are displayed in the lower plot. .... 183

Figure F.3: BTO 2015 residual boxplots by site, ordered alphabetically. The full-model residuals are displayed in the upper plot, while the species-wide effects residuals in isolation are displayed in the lower plot. .... 184

Figure F.4: BBU combined residual boxplots by site, ordered alphabetically. The full-model residuals are displayed in the upper plot, while the species-wide effects residuals in isolation are displayed in the lower plot. .... 185

Figure F.5: BTD combined residual boxplots by site, ordered alphabetically. The full-model residuals are displayed in the upper plot, while the species-wide effects residuals in isolation are displayed in the lower plot. .... 186

Figure F.6: SBTD combined residual boxplots by site, ordered alphabetically. The full-model residuals are displayed in the upper plot, while the species-wide effects residuals in isolation are displayed in the lower plot. .... 187

Figure G.1: Representative settlement plate images for each measurement period in both years for the expanded sites, showing the different biofouling communities present at that point in time. Ranked from highest 2015 mean *C. intestinalis* abundance. .... 188

## List of Abbreviations

°C = Degrees Celsius

$\Delta_i$  = AIC difference

$\theta_i$  = Parameter vector

ACRDP = Aquaculture Collaborative Research and Development Program

ASES = Absolute scaled effect size

BBU = Best bottom-up model

BIO = Bedford Institute of Oceanography

BRH = Biotic resistance hypothesis

BTD = Best top-down model

CC = Camp Cove

Chl-*a* = Chlorophyll-*a*

Days50 (D50) = Time in days for half of the maximum population to occur

DFO = Department of Fisheries and Oceans Canada

DW = Dingwall

ERH = Enemy release hypothesis

FP = Falls Point

*g* = Acceleration due to gravity

GFF = Glass fibre filter

GM = Global model

IMH = Invasional meltdown hypothesis

IN = Ingomar

IP = Indian Point



Lag = Lag phase

LB = Louisbourg

LR = Little River

MAE = Mean absolute error

MaxPop = Maximum population

MEF = Modelling efficiency

ML = Maximum likelihood

mS/cm = Millisiemen per centimetre

NIS = Non-indigenous species

NLME = Nonlinear mixed-effects

NS = Nova Scotia

PB = Port Bickerton

PEI = Prince Edward Island

PG = Petit-de-Grat

PSU = Practical salinity units

ppt = Parts per thousand

QTL = Quantitative trait loci

RMSE = Root mean square error

SA = St Ann's Bay

SH = Ship Harbour

SP = St Peters

VC = Venus Cove

WP = Wedgeport

YB = Yarmouth Bar

ZZ = Cape Canso

# 1. Introduction

## 1.1 Overview

Invasive species spread by human activities are a major problem in today's marine ecosystems (Grosholz, 2002). Invaders can out-compete native species for food, space or other resources, disrupting natural communities and compromising the ecosystem services important for human populations. Several tunicatae species are among the most invasive marine species documented (Adams et al., 2011; Carman et al., 2010), with populations spread worldwide through ballast water, hull fouling, and improperly handled aquaculture materials (Ramsay et al., 2008a). These animals with short-lived planktonic larvae and sessile, filter-feeding adults proliferate on many subtidal surfaces, including man-made substrata. The establishment of *Ciona intestinalis*, commonly referred to as the vase tunicate, in Atlantic Canadian waters has negatively impacted mussel farm productivity (ACRDP, 2010). *C. intestinalis* grows in dense aggregations on ropes, nets and the cultured mussels themselves. Tunicate fouling affects many types of aquaculture, but for mussel farmers in particular, managing vase tunicates can increase costs to the point of zero profit. The Department of Fisheries and Oceans Canada (DFO) has monitored *C. intestinalis* in Nova Scotia (NS) since 2005, and the distribution throughout coastal NS was spatially and temporally heterogeneous (Sephton et al., 2011). This variability in invasive success has not yet been properly explained. An improved understanding of what determines *C. intestinalis* distribution and abundance around the province may help to mitigate current invasions, limit further spread, and guide aquaculture management decisions.

The overall goal of this research is to explain why *C. intestinalis* populations are spatially and temporally heterogeneous. Heterogeneity could be due to: (i) patterns of anthropogenic activity, such as aquaculture practices and vessel movement; (ii) genetic diversity that causes different responses to similar environmental conditions; (iii) variable biotic interactions (competition, predation, facilitation); or (iv) variable environmental conditions throughout the province. For this study, I investigated the fourth possibility, to determine whether heterogeneous *C. intestinalis* abundances can be linked to one or more environmental parameters that can regulate population growth. The basic objectives were: (i) to generate a detailed record of *C. intestinalis* population growth and of temperature, salinity, pH and water movement conditions at a range of sites; (ii) to determine if there are associations between the environmental conditions and *C. intestinalis* population growth; (iii) and to develop and

improve upon existing methods to monitor macroscale population growth of a sessile marine invertebrate. Relative abundance was monitored on standard invasive species monitoring collectors deployed at harbours and aquaculture sites with varying *C. intestinalis* abundance, according to historical records. Environmental monitoring was accomplished with *in-situ* data-loggers that allowed long-term, high-frequency measurements, improving the ability to capture potentially relevant characteristics that either negatively or positively affect *C. intestinalis* growth. Temperature (Dybern, 1965), salinity (Dybern, 1967) and water movement (Havenhand and Svane, 1991) are all known to affect *C. intestinalis* at various life-stages. Although very little is known about specific pH effects on tunicates in general, the limited information suggests tunicates (*C. intestinalis* among the handful tested) may find more acidic conditions beneficial (Dupont et al., 2009, 2008). Given the changing state of global oceanic chemistry due to increasing carbon dioxide uptake, this is a key area of interest (Doney et al., 2009). The overall objective was to develop a predictive model based on these data that inputs the abiotic determinants (if present) of *C. intestinalis* population growth, thus producing a robust and reliable guide to understanding and perhaps predicting *C. intestinalis* distribution and abundance.

Aquaculture is likely to be an important industry for future global food security (Gjedrem et al., 2012). While my research is fundamentally addressing *C. intestinalis* population ecology, the results may be useful for mussel farming and other activities that are affected by biofouling caused by this nuisance species. Links (or lack thereof) between *C. intestinalis* abundance and environmental conditions could help to refine assessment of aquaculture sites for their risk of infestation, possibly affecting profitability of existing and potential sites. My work could improve general knowledge of the spread and success of a prominent invasive species, with potential to inform future prevention and mitigation strategies. In addition, invasive species, in general, pose a threat to ecosystems, making it desirable to understand factors that lead to population changes in species of interest. The methods used in my study to model associations between abiotic conditions and population growth could be transferable to other troublesome biofouling species, providing a framework to approach a similar research project. Finally, with increasing pressure on the environment from global change, baseline monitoring provided by research, such as this study, is essential for assessing the scope of such changes.

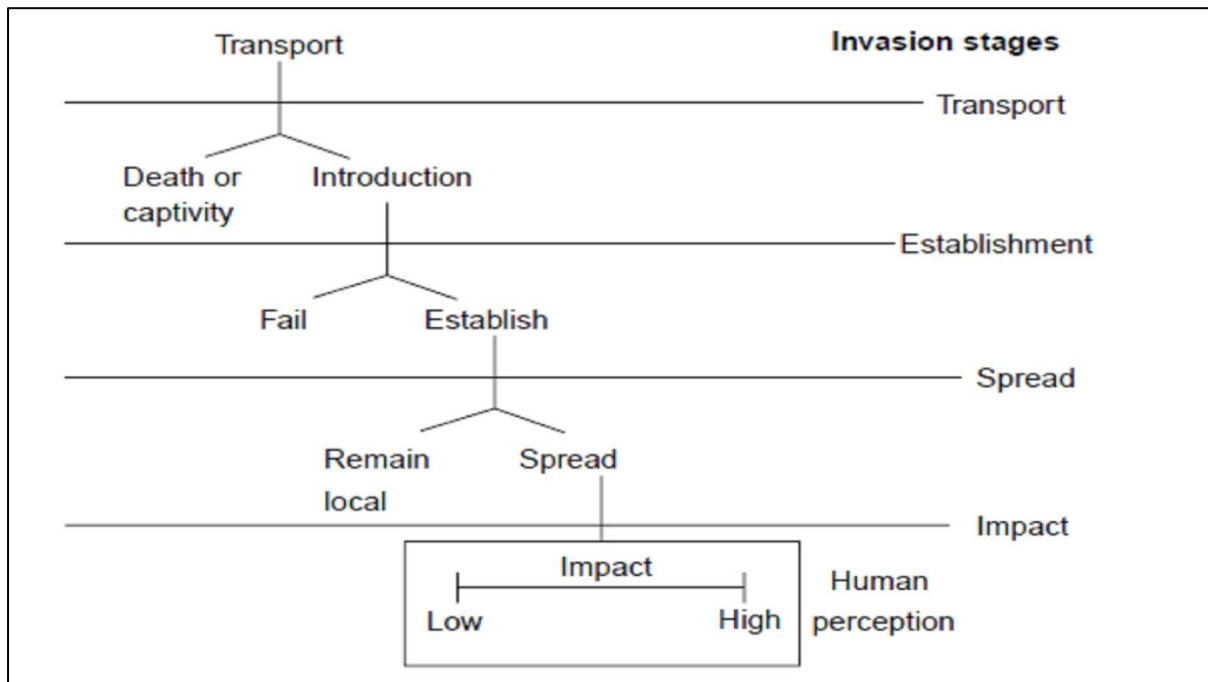
## **1.2. Invasion ecology**

### 1.2.1. The history and terminology

The first issue encountered while reading about invasion ecology, regardless of the environment in question, is the complex and often contradictory vernacular within the discipline. Some authors use words like ‘introduced’ and ‘invasive’ interchangeably (Williams and Smith, 2007), while others have clear-cut definitions for important terminology. Introduced species, often referred to as alien, non-native, exotic or non-indigenous species (NIS), are ones that have been transported by humans, intentionally or not, to an area outside of their native range, often across a major geographical barrier (Richardson et al., 2000). The concept of ‘native’ and ‘non-native’ species was introduced by Henslow (1835) and these are widely accepted terms. The perception of an introduced species is relatively ubiquitous: it is trying to define what an invasive species is that poses a challenge. There appear to be two general opinions as to what the correct definition of an invasive species should be. Some authors, Richardson et al. (2000) among them, argue that the term ‘invasive’ should only be applied to introduced species that have become abundant and expanded their geographic range, linking the term to a clear biological phenomenon, geographic range expansion. The second general understanding, advocated for by Lockwood et al. (2013) and many others, is that ‘invasive’ describes a species that has a demonstrable negative ecological or economic impact. Lockwood et al. (2013) argue that invasion is truly the end product of a series of stages (Fig. 1.1), each one providing a barrier to the potentially invasive species. This proposes the view of invasion as a stepwise process, leading to a better ability to manage invaders through a clearer view of when anthropogenic and ecological processes come into play, rather than viewing invasion as a dichotomous classification, i.e. invasive or not.

There have been numerous attempts to end the inconsistencies in terminology used in invasion ecology and to construct a common framework that could be employed universally in the discipline (Blackburn et al., 2011; Inderjit et al., 2006; Richardson et al., 2000). Despite these efforts, to satisfy disagreements over terminology it’s common for authors to lay out their definitions of key terminology (Inderjit et al., 2006; Lockwood et al., 2013). Based on the literature and for the purposes of this thesis, I will be using the Lockwood et al. (2013) definition of invasive species. An invasive species is one that causes an evident negative ecological or economic impact. The detrimental economic impact imposed on societal sectors of interest by *C. intestinalis* biofouling is clear (ACRDP, 2010; Ramsay et al.,

2008b). This, on top of the likely negative impacts it has for a number of native species (Blum et al., 2007), provides a good basis to consider *C. intestinalis* as an invasive species.



**Figure 1.1: A simple model of the invasion process showing the distinct stages, reproduced from Lockwood et al. (2013) with the permission of John Wiley and Sons.**

### 1.2.2. Introduced species: a precedent for invasion?

Biological invasions are recognised as a major threat to worldwide terrestrial, marine and freshwater biodiversity, as well as causing substantial economic losses across the globe (Colautti et al., 2006; Kelly et al., 2009; Marquet et al., 2012; Prenter et al., 2004; Sylvester et al., 2011). The majority of the literature, when addressing the issue of biological invasion, opens by stating the current and impending threats that this process poses to biodiversity and the economy. Prenter et al. (2004) go as far as to say we are facing an “invasional meltdown”, hence why it is so important to gain a greater insight into the processes of biological invasion, from beginning to end. This is not to say that all introduced species pose an invasion risk. In fact, it is estimated that roughly 0.1% of original introductions reach a point where they end up categorized as invasive (Williamson and Fitter, 1996). Despite the seemingly negligible amount of introductions that pose an invasion risk, they have an impact disproportionate to the number of species.

However, in recent times there have been voices rising in opposition to the status-quo, instead advocating for a more flexible approach when viewing ‘introduced’ species in the

increasingly connected world we are faced with. Davis et al. (2011) argue that the predisposition to view introduced species as bad is biased and leads to poor management decisions and wasted resources on needless control efforts. Debate and subjectivity surrounding the true definition and importance of biological invasion is not a new occurrence. The divide in opinions has been building since the inception of the discipline of invasion ecology in the mid twentieth century. Despite being such a long standing issue, this field of discussion is far from over with multiple avid proponents on either side of the debate. However, there was a decisive rebuttal to Davis et al. (2011) in the shape of a series of correspondents in *Nature* (Alyokhin, 2011; Lerdaun and Wickham, 2011; Lockwood et al., 2011; Simberloff, 2011) all addressing separate flaws in the argument. While it is becoming more popular to condemn the segregation of species based on their native and non-native status as a precursor to judgment of invasive likelihood, there is a solid foundation to the discipline and reasoning behind these practices. While there may be benefits to some introduced species, the ramifications introduction are largely unpredictable. History has shown us that in the grand scheme of things, an increasing number of introductions lead to a greater number of troublesome invasive species. As to the worry that resources are wasted on needless management: limited starting resources are inherent in the selection process for control and management, meaning there is already a rigorous screening system to focus only on species with a high potential negative impact (Simberloff, 2011). These high risk species are often outlined on a national or regional basis in terms of lists, such as ‘species of most concern’ or ‘least wanted’ that are often a part of public awareness campaigns. While Davis et al. (2011) challenge the militaristic terminology implemented in invasion ecology, the attention of the public has been caught in the case of many well-known invasive species, and it may be best to capitalize on this awareness to reduce overall introductions, in the hope of preventing some invasions. Adopting a framework à la Davis et al. (2011) whereby management bodies, public and private, accept a new age of global connectivity and the ecological and biological changes that occur as a result would be fine, if resources that we as humans rely on weren’t detrimentally impacted. A balanced view is required when assessing introduced species. Not every introduction causes problems, and the majority of invasion ecologists explicitly proclaim this (Lockwood et al., 2011). Focus on the tiny proportion of introductions that do go on to be classed as invasive could present an apparent bias in the literature, but the “vilification” that Davis et al. (2011) believe is pervasive within the discipline is justified for the small proportion of introductions that do wreak environmental and economic havoc. The negative impacts caused by invasive species are well documented,

as well as being very difficult to predict *a priori* (Simberloff, 2011). But for the small number of introduced species that do have a large negative impact, it is more effective to err on the side of caution and implement prevention management and rapid response management (Lerdau and Wickham, 2011).

### *1.2.3. Transport and introduction of marine benthic invasive species*

Introductions in the marine environment have been greatly facilitated by human mediated transport and the aquaculture industry (Rigal et al., 2010). Murray et al. (2012) suggest that the transport of marine NIS and thus, the first step of the invasion process, is the uptake by a human mediated vector, either accidentally through hull fouling of ships, intake of species via ballast water or sea chests (intake chambers in vessel hulls), or intentionally through the trade of resources via the aquaculture industry. Hull fouling and the transportation of organisms in ballast water tanks and sea chests are recognised as the main vectors for the inadvertent introduction of the majority of marine benthic introductions in all oceans and estuaries, despite current technology and preventative measures (Farrapeira et al., 2011; Murray et al., 2013). Successful introductions then rely on the survival of such species transported by any of these means.

The study of biofouling communities on submerged surfaces in the marine environment is not a recent area of study. A study dating back to the early 20<sup>th</sup> Century, entirely focuses on the cause and effects of biofouling on ships submerged surfaces, while the US Navy also carried out research in this area (Visscher, 1927). However, the focus of such studies has changed over the years. While Visscher (1927) and the US Navy were primarily interested in the impact of biofouling on fuel consumption, recent studies have paid more attention to the ability of ships to act as vectors for global marine biological invasion. One of the first publications in a leading journal, *Nature*, to pay attention to the spread of a marine benthic species outside of its historically native range was written by Sandison (1950). He was interested in the appearance of *Elminius modestus* Darwin, an Australasian barnacle, in the northern hemisphere for the first time. Only a few years later, Allen (1953) presented his findings that a probable cause for the discovery of several species of Bryozoa never previously recorded in Australian waters, was the movement of sessile marine organisms attached to the submerged surfaces of ships. More recently, there has been an explosion in the interest shown towards biological invasion, with a marked increase in literature referencing ‘invasion ecology’ in their keywords from the early 1990s (Lockwood et al., 2013). While it



is a common comment in the introduction of articles that the initial processes of marine biological invasion are understudied and poorly understood, especially at a global level (Davidson et al., 2010; Lacoursière-Roussel et al., 2012a; Murray et al., 2012; Ruiz et al., 2013; Seebens et al., 2013), I believe that there has been a significant amount added to the literature recently. Particularly in attempts to create predictive models for invasion based on environmental conditions, biogeographic distribution information and vector based information. This work is helping to vastly improve the understanding of the initial process of marine benthic biological invasion. Despite the greater attention to this area of late, there is still work to be done, given the ever increasing global trend of invasion.

Verling et al. (2005) described the factors that most commonly determine the success of a marine invasion. The spread of marine benthic invasive species (MBIS) is subject to: the availability of suitable transport vectors; the rate of vector movements between the source and the receiving environment; the quantity and quality of propagules released in individual inoculation events; the timing of any introduction event relative to the species' life history; and the availability of limiting resources in the receiving environment. The vast variation between any one of these factors at any one time, make the spread of MBIS very difficult to predict (Floerl et al., 2009). However, predictive models based on the factors outlined above (Verling et al., 2005) have been developed in order to predict where sources of introductions are most likely to occur and where is more susceptible to introductions of MBIS. In one such model, developed by Floerl et al. (2009), the results suggested that locations with a high degree of connectivity to other locations, through a high volume of traffic, were four times as likely to become invaded as locations with low traffic and connectivity, as well as experiencing invasion events earlier. A number of other models like this have been devised to try and increase the success of vector management, using predictors like: time since last antifouling treatment; time since last cleaning; type of vessel; length of voyage; speed of voyage; and time spent in each port (Lacoursière-Roussel et al., 2012b; Sylvester et al., 2011). All of the aforementioned can contribute to the probability of introduction and the potential success of invasion.

The initial introduction of MBIS is dependent on their survival during the transportation process. Depending on the vector, they may have to resist a variety of physical and physiological stressors connected to their journey (Murray et al., 2012). The attachment strength and drag coefficient are attributes that are very important in the assessment of hull fouling species. Murray et al. (2012) hypothesise that a successful invader that utilises hull

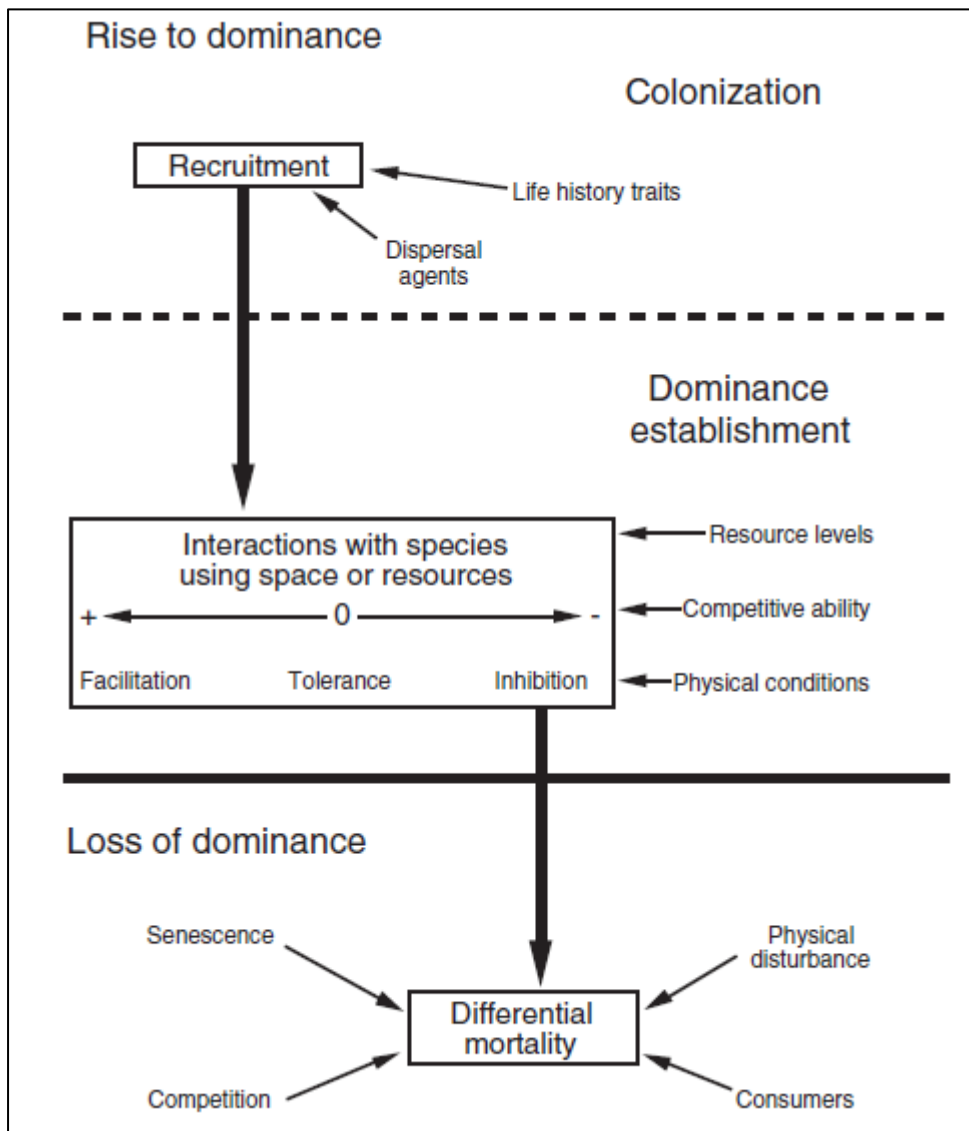
fouling as a means of introduction and spread, would have a strong attachment strength and low drag coefficient. When compared to a native counterpart it is suggested that they would possess superior biomechanical properties, allowing them to remain attached more successfully. Testing the attachment strength and drag coefficient, they found that *Styela clava*, a well-known invasive solitary tunicate (like *C. intestinalis*), has superior attachment strength and a lower drag coefficient than a number of native tunicates. It was the only tested species to have a dislodgment velocity well above that of a fast moving vessel. Their results show that different vectors use by different types of species is likely, with colonial tunicates such as *Didemnum vexillum* and *Botrylloides violaceus* more likely to be found on slow moving vessels such as sail boats and barges. In saying that, colonial tunicates can utilize a different strategy likely to be employed by a number of species. Fragmentation of the organism can occur when drag becomes too great (or when pressure washed as a biofouling treatment), possibly leaving remains of the individual on the boat/substrate while the fragmented portion can float with the ability to reattach elsewhere (Paetzold and Davidson, 2010; Reinhardt et al., 2012).

Species-specific traits often determine what vector will be more likely to cause an introduction for that given species. The transportation and release of ballast water is regarded as one of the most important vectors for the introduction of benthic marine species around the globe (Chan et al., 2013; DiBacco et al., 2012; Seebens et al., 2013). Despite this, tunicates have rarely been recorded in ballast water (Lacoursière-Roussel et al. 2012b). The fouling of hulls however is a well-recognised vector for tunicate introduction (Davidson et al., 2010). Understanding the most likely vector of introduction provides a vital step in detection and prevention. Not only is the vector-type important, but so too is understanding the given vector in a specific region. Seasonal trends and boating patterns can influence the spread of fouling species greatly. Differences in the seasonal peak of tunicate infestation and boating usage vary between Canada and New Zealand for example, displaying that one method of control does not work across the board, due to varying spatial and temporal factors (Lacoursière-Roussel et al. 2012b).

#### *1.2.4. Gaining and maintaining dominance in a new environment*

The chronic and stochastic challenges facing introduced species can be both abiotic and biotic in nature, with a combination of these forces generally acting in concert with the particular characteristics of the species at hand, determining the success of establishment and

potential range expansion (Lockwood et al., 2013; Mack et al., 2000). While my research is focused primarily on the abiotic forces at play in determining the growth of *C. intestinalis*, it is important to outline some of the key invasion ecology theory underlying potential biotic forces at play in the system. The fundamental processes that occur during succession can help us to understand the processes of species invasion, as the initial rise to dominance through succession can be compared to the initial stages of biological invasion (Wootton, 2002).



**Figure 1.2: Description of potential mechanisms involved in the rise to, establishment of, and loss of dominance reproduced from Wootton (2002) with the permission of Springer.**

Generally, a species' dominance in a given community is related to the amount of biomass present (Wootton, 2002), particularly for sessile assemblages such as macroalgae and many marine invertebrates. Fig. 1.2 outlines the general stages that occur in succession that can

eventually lead to species dominance and in the case of biological invasions, range expansion. The colonization process outlined in Fig. 1.2 is comparable to the transport and introduction phase of biological invasion discussed earlier in sec. 1.2.2., as well as survival to recruitment. It is important to take the initial process of establishment into account when considering the broader temporal issue of maintained dominance, as the initial propagule pressure and colony pressure imposed by an introduction have been suggested to play a role in post-establishment growth (Ricciardi et al., 2013). The other parallel between succession and invasion is the interaction and effects of the biotic and abiotic forces acting on a given organism or system. Fig. 1.2 describes the processes and factors involved in the rise and fall of dominance that determine succession. It is only recently that the attention of researchers has been drawn to the biotic interactions that mediate the persistence of invasive species after establishment (Scheibling and Gagnon, 2006). The majority of interactions described in the literature have previously been inferred from correlative evidence rather than rigorous experimental testing of mechanisms. Mortality is the fundamental process that leads to the demise of a dominant species and can occur for a number of different reasons specified in Fig. 1.2. Therefore, once an introduced species has become established and invasive, it can prevent the loss of dominance described in Fig. 1.2 by overcoming factors such as competition, consumption and physical disturbance. The abiotic conditions a species is subjected to can play a role in their ability to deal with and overcome such factors. Comparing the processes of succession to biological invasion can help us to understand what species' traits and interactions lead to the establishment and maintenance of dominance in a biological invasion.

#### 1.2.4.1. Current hypotheses explaining the success of MBIS

A variety of ecological mechanisms have been offered to explain invasion success through hypotheses such as the 'enemy-release hypothesis' (ERH), 'novel weapon hypothesis', 'evolution of increased competitive ability', 'ecosystem engineering' and the 'invasional meltdown hypothesis (IMH)' (Kimbrow et al., 2013). This area is severely understudied in relation to MBIS and would certainly classify as a knowledge gap. The majority of work on these hypotheses has been carried out for terrestrial plants and is theoretical on the whole (Ricciardi et al., 2013). But there has been a recent appearance of empirical studies on algae (Enge et al., 2013; Svensson et al., 2013) and to a lesser extent, invertebrates (Raw et al., 2013).

#### 1.2.4.2. The role of predation in biological invasion

As highlighted in Fig. 1.2, consumption is one of the main contributors to population mortality and if experienced at high enough levels can lead to the loss of dominance of a species in a given community. One of the explanations for invasive species success in establishment, persistence, and spread is the ERH, originally conceived in the field of plant invasions (Keane and Crawley, 2002). It is the theory that invaders escape their natural predators and are more successful in the invaded range as the native predators do not consume them, or to a lesser extent than native prey (Dumont et al., 2011). There is contrasting evidence throughout the literature about the validity of this hypothesis, not to mention the conflicting biotic resistance hypothesis (BRH), which predicts that introduced species are poorly adapted to deter native consumers, limiting invasibility (Elton, 1958). Both hypotheses are supported and it is likely that both are important in different situations. This is another area where the abiotic conditions can influence the outcome these biotic forces. There is a major issue in that the majority of work carried out on these hypotheses has been carried out on terrestrial plant communities, results of which are then being transplanted and compared to the limited work carried out in the marine environment. Further focused experimental work is required in this area before a consensus can be formed on what role predation plays in invasion success.

#### 1.2.4.3. The role of competition in marine benthic invertebrate invasion

Competition has traditionally been viewed as the key mechanism determining NIS impacts (Kelly et al., 2009). It is a key factor in NIS displacing native species and remaining dominant in the invaded region (Byers, 2002) and Fig. 1.2 displays competitive ability as a key component of the establishment and maintenance of dominance (invasion). Superior competitive ability may be explained by a superior ability for resource acquisition, a rapid rate of both reproduction and growth, greater tolerance to environmental change, and adaptation to disturbance (Enge et al., 2013). Anthropogenic activity, particularly environmental disturbance, is altering the interspecific interactions that occur in marine communities, often to the benefit of opportunistic species (Harris and Tyrrell, 2001). The overfishing of top predators such as fish and sea urchins, as well as the physical disturbance of activities such as dredging are providing an environment that many invasive species can flourish in, and combined with the absence of top predators, is allowing for invasive species to prosper and outcompete native species (Harris and Tyrrell, 2001).

### 1.3. *Ciona intestinalis*

#### 1.3.1. *The colourful history of C. intestinalis*

Commonly referred to as the vase tunicate, *C. intestinalis* (Linnaeus 1767) is a solitary ascidian and member of the class Ascidiacea, first described as *Ascidia testinalis* by Linnaeus in 1767. Not so long ago, *C. intestinalis* was believed to be a single globally distributed species. However, genetic and genomic studies carried out on individuals from different populations around the world revealed the true diversity present, with at least four distinct entities discovered (Caputi et al., 2007; Suzuki et al., 2005). Whether it was a cryptic species-complex (Caputi et al., 2007) or a series of sub-species (Suzuki et al., 2005) was debated. Of these four entities, there were two in particular, referred to at the time as *C. intestinalis* type A and type B, that were widespread and also causing the biofouling and invasive issues. Morphological and genetic studies continued at an increasing rate. Sato et al. (2012) published reasonably reliable morphological metrics capable of identifying the two in the field, while a host of genetic studies provided more evidence underlying the distinctions between the entities (Nydam and Harrison, 2011, 2010, 2007; Roux et al., 2013; Zhan et al., 2012, 2010). What was previously thought of as a single species with a remarkable global distribution was in fact comprised of “two distinct, highly divergent entities” (Nydam and Harrison, 2010). Findings of morphological differences between adults and larvae of the two widely distributed types led Brunetti et al. (2015) and Pennati et al. (2015) to recommend the two entities be officially recognised as separate species. They propose that *C. robusta* (already described as a unique species (Hoshino and Tokioka, 1967), until it was redefined as a synonym of *C. intestinalis*) be adopted for the formerly referred to *C. intestinalis* type A, while type B shall remain *C. intestinalis*. The classification recommendations have already been adopted by some (Bouchemousse et al., 2015), while other fields of research are not consistent or up-to-date. All work prior to 2005 presented work regarding *C. intestinalis* as relating to a single global entity, understandably so. Even for a few years thereafter, it is reasonable to accept that the distinctions that were becoming apparent were not widely adopted, in any field of research. However, the fact that it is still being overlooked to this day, for whatever reason, is unreasonable.

A huge number of proteomics and development studies carried out on what is called ‘*C. intestinalis*’ in the publications are actually using samples collected from areas where *C. robusta* is exclusively present, while others are using samples from areas where sympatry

occurs (Goff et al., 2015; Matsunobu and Sasakura, 2015; Nakamura et al., 2015; Vizzini et al., 2015). The majority of these studies do not provide any account of the presence of two distinct species, let alone offer a description of the molecular divergence of the two species. It is the rare case for a publication in this field to distinguish between the two species. In one proteomics study (Nakazawa et al., 2015), the study organism is referred to as “*Ciona intestinalis* type A (*C. robusta*)” in the abstract and material and methods, but then throughout the remainder of the publication it’s referred to solely as *C. intestinalis*. Having recognised the recent recommendation that *C. intestinalis* type A is actually *C. robusta*, it is unusual that it is then reverted to *C. intestinalis*. Then there are studies that correctly use the nomenclature of *C. intestinalis* throughout when referring to specimens collected in an allopatric region, but have failed to distinguish the varying regions of *C. intestinalis* and *C. robusta* in the global distribution when described in the introduction (Zhao et al., 2015). A far smaller portion of recent works are fully clarifying where the sample organisms originated, appropriately recognizing the *Ciona* type, and consequently referring to it appropriately throughout the publication (Jeffery, 2015).

The consequences of carrying out research on a given species and then attributing the results to a closely related species without distinguishing the potential genetic and phenotypic variation between the two are difficult to quantify. I am by no means admonishing any of the previous works, or suggesting they are less relevant, but the question remains: what are the true consequences of misattribution of result? It is also possible that the degree of importance of such misattributions is variable depending on the field and context of a given study. Transferability of genome-wide microsatellite markers from *C. robusta* and *C. intestinalis* is lower than might be expected, while all transferable markers are polymorphic in *C. intestinalis* (Lin et al., 2015). While the true meaning of this in terms of the phenotypic differences that might occur as a result of the genome divergence is not clear, there is a correlation between low transferability and genetic divergence across many taxa, with potential developmental and ecological implications (Lin et al., 2015).

Closer to home, multiple ecological publications have grouped data obtained from the two species into single predictive models, risk assessments or physiological experiments, without a caveat mentioning the distinction and possible implications of the differences (Kanary et al., 2011; Patanasatienkul et al., 2014; Serafini et al., 2011; Therriault and Herborg, 2008a, 2008b). This could be quite important, given some of the information input into models about environmental tolerances or reproductive and development characteristics being quite

different. The difference between c.12,000 eggs in a lifetime in individuals sampled in Canada compared to c.100,000 eggs in the lifetime of *C. robusta* in Japan is a starting estimate that could have significant bearing on a predictive dispersal or life-history model. The fundamentals of designing experiments to answer questions about a specific species can potentially be erroneous if the most recent information regarding such nuances is not accounted for.

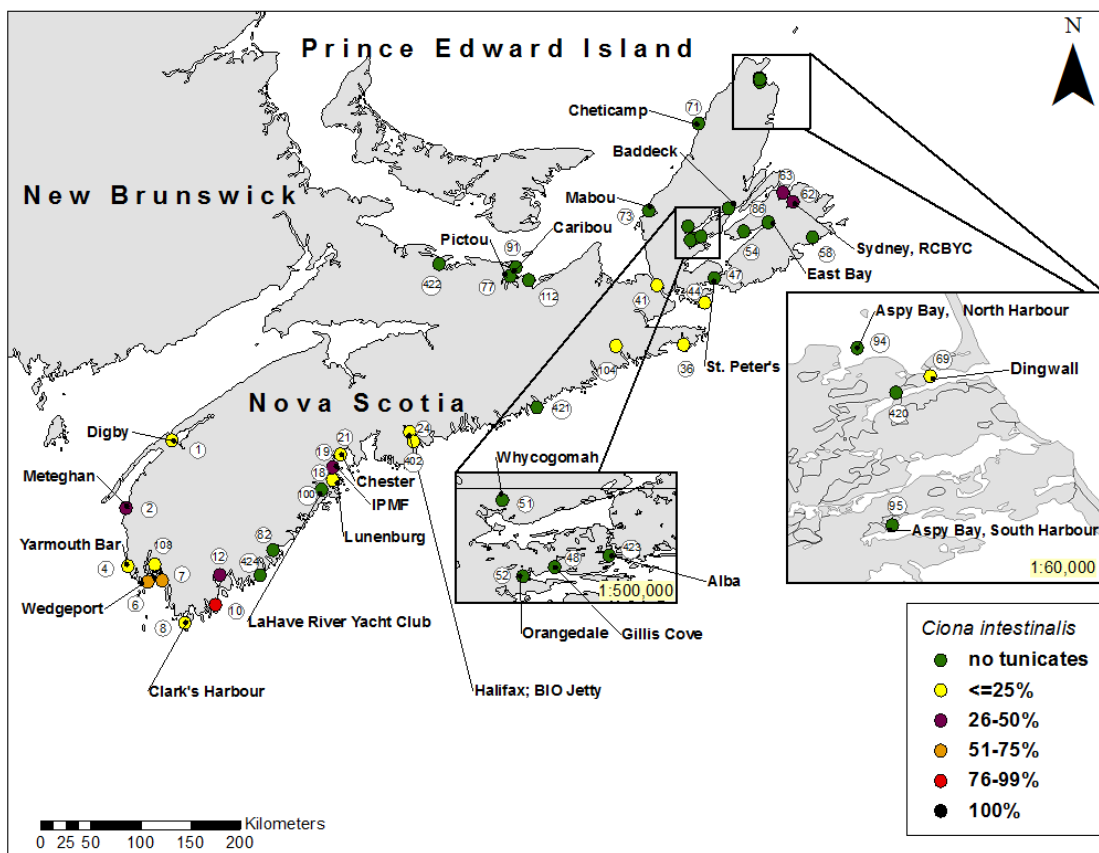
These revelations in the field of *C. intestinalis* studies have shed light on the distribution and potential origins of the two species previously grouped under the same name. It appears as if *C. intestinalis* is native to the Northern Atlantic Ocean, and was introduced to the Western Atlantic Ocean, displaying a preference for a cold temperate climate (Nydam and Harrison, 2011). The Northwestern Pacific region is most likely the indigenous range of *C. robusta* (Nydam and Harrison, 2011), which has been introduced to the east coast of N. America, the Mediterranean Sea, both Atlantic and Indian Ocean coasts of South Africa, the Atlantic coast of South America, Australia, and New Zealand (Caputi et al., 2007; Rius et al., 2014; Zhan et al., 2012, 2010). What was previously considered as *C. intestinalis*' native and invasive ranges (Therriault and Herborg, 2008a), now loosely translates to the distribution of *C. intestinalis* and *C. robusta* respectively. The only known region of sympatry is the English Channel and French Atlantic coast (Nydam and Harrison, 2011; Roux et al., 2013).

### 1.3.2. Distribution of *C. intestinalis* in Atlantic Canada

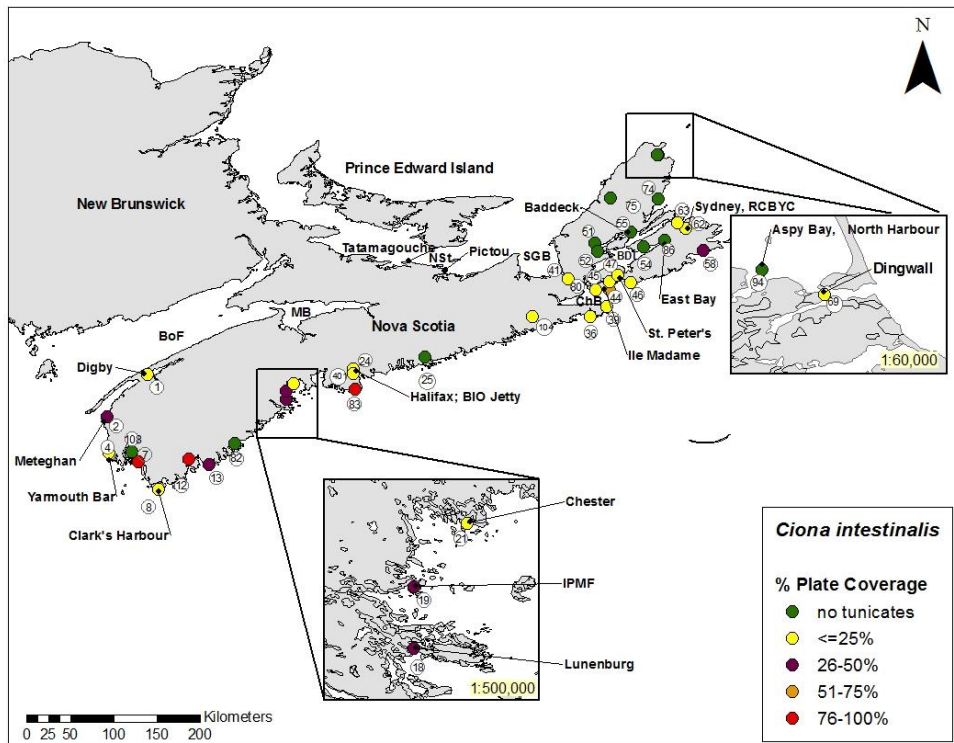
*C. intestinalis* is currently classed as a cryptogenic species (of unknown origin) in Canadian waters, but was most likely introduced from northern Europe via ship hull fouling. Despite this, there is a degree of uncertainty due to records of its presence in New Brunswick dating back to 1852 (Carver et al., 2006). The recorded progression of *C. intestinalis* around NS, appearing to stem from a locus on the southern shore, appears more like an introduction and spread rather than a species that has been present in the coastal waters of NS for a long period of time. Given the long-running confusion regarding the taxonomic classification of species within the *Ciona* genus, I would advise caution when using a single observation of *C. intestinalis* in the mid-1800s to establish its historical record in Atlantic Canada. It is possible that the recording in New Brunswick was a single introduction occurrence or even a different species misidentified. There is a sufficient weight of evidence to progress with the view that *C. intestinalis* was introduced to NS via some mode of unintentional vector transport.



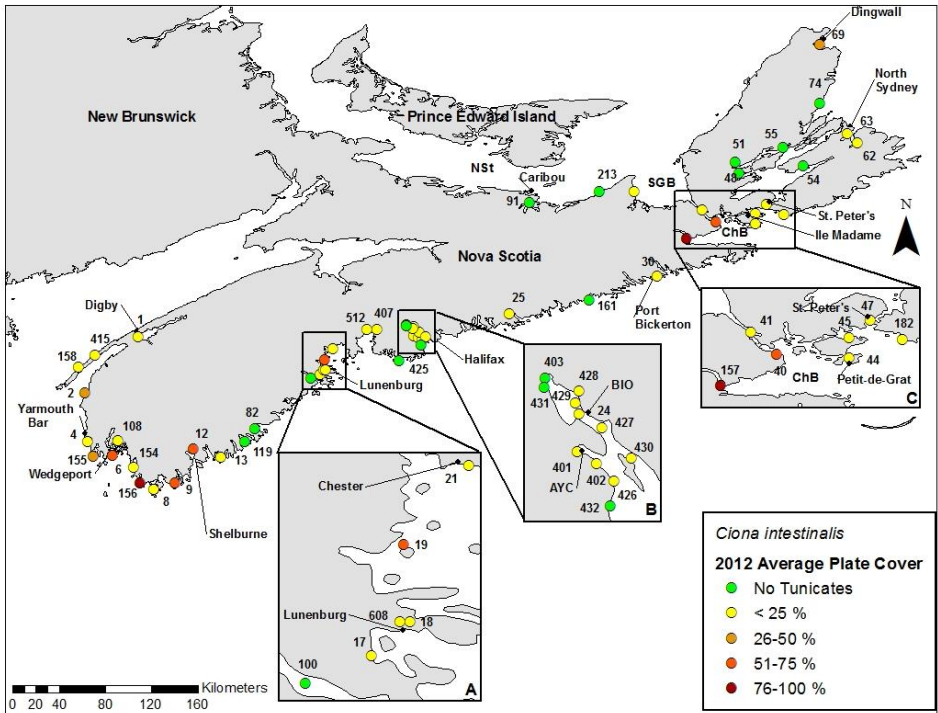
*C. intestinalis* is heterogeneously distributed on the NS coastline (Sephton et al., 2011), a relatively common occurrence among populations of this species. In what is believed to be the native range of *C. intestinalis*, Northern Europe, heterogeneity has been recorded both spatially and temporally (Petersen and Svane, 1995). Spatial abundance and heterogeneity of *C. intestinalis* populations in Norway have been linked to predation by a number of fishes and particularly the sea star *Asterias rubens* (Gulliksen and Skjæveland, 1973). The mechanisms regulating the temporal variation in abundance representative of *C. intestinalis* populations in Norway not certain, but it is believed to be a long-term cycle linked to recruitment that causes it (Petersen and Svane, 1995). The heterogeneous distribution pattern in NS (Fig. 1.3 – 1.7), spatially and temporally, is similar to other *C. intestinalis* populations around the world, but the reasons behind it are as of yet unknown. Temperature and salinity do influence *C. intestinalis* survival and growth, but to what degree local conditions influence the establishment and spread of individual *C. intestinalis* populations remains unknown (Vercaemer et al., 2011). There are also a wide range of additional environmental variables that could act individually or together to explain the variability of establishment and spread of *C. intestinalis* in NS.



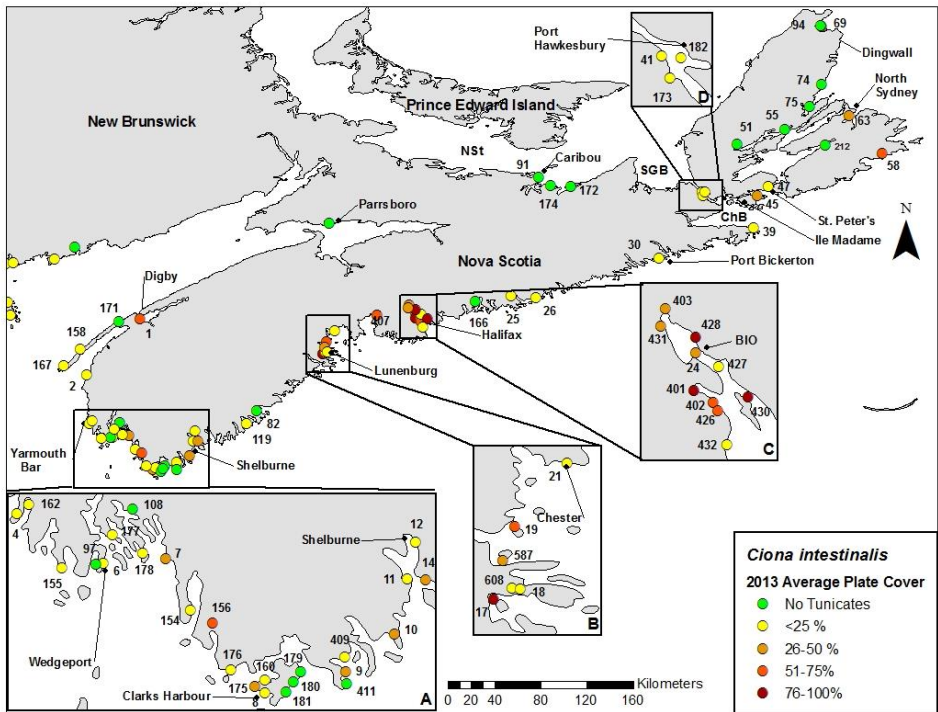
**Figure 1.3: Abundance (represented as % cover on standard monitoring collectors) of *C. intestinalis* at NS coastal and inland monitoring sites in 2010, obtained through DFO AIS surveys (Reproduced from Sephton et al. (2014) with permission).**



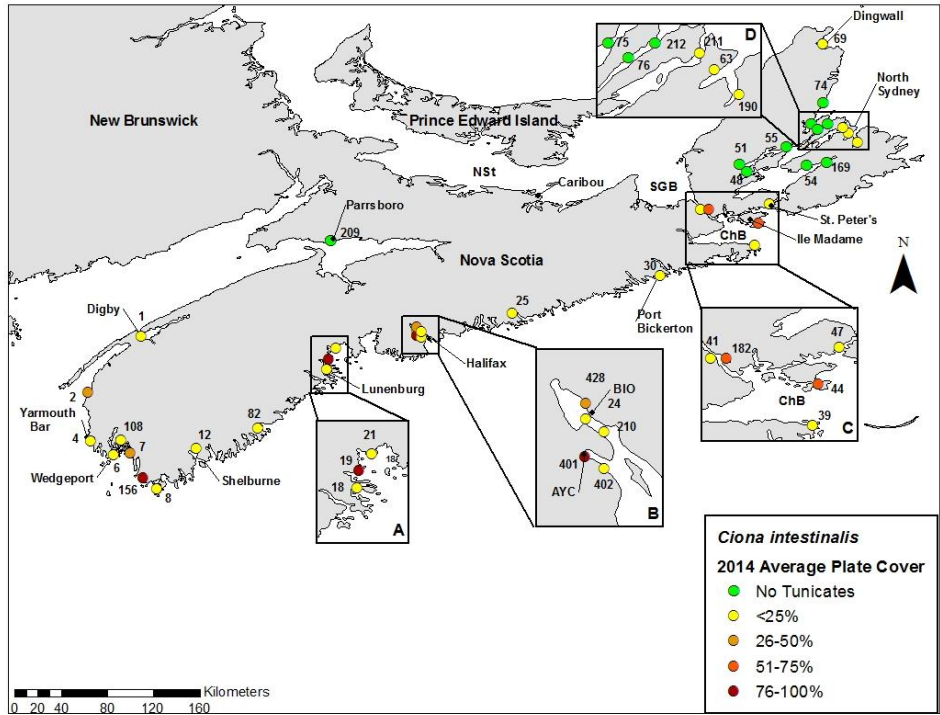
**Figure 1.4: Abundance (represented as % cover on standard monitoring collectors) of *C. intestinalis* at NS coastal and inland monitoring sites in 2011, obtained through DFO AIS surveys (Reproduced from Sephton et al. (2015) with permission).**



**Figure 1.5: Abundance (represented as % cover on standard monitoring collectors) of *C. intestinalis* at NS coastal and inland monitoring sites in 2012, obtained through DFO AIS surveys (Pers. comm. Sephton D., with permission from the DFO).**



**Figure 1.6: Abundance (represented as % cover on standard monitoring collectors) of *C. intestinalis* at NS coastal and inland monitoring sites in 2013, obtained through DFO AIS surveys (Pers. comm. Sephton D., with permission from the DFO).**



**Figure 1.7: Abundance (represented as % cover on standard monitoring collectors) of *C. intestinalis* at NS coastal and inland monitoring sites in 2014, obtained through DFO AIS surveys (Pers. comm. Sephton D., with permission from the DFO).**

1.3.3. *C. intestinalis* biology

Abiotic conditions have the potential to influence *C. intestinalis* survival, growth, and reproduction throughout its life cycle. In the following sections I will introduce an overview of *C. intestinalis* biology in relation to the environment it inhabits, with an onus on studies conducted in Atlantic Canada, especially related to temperature, salinity, water flow, and pH.

1.3.3.1 Reproduction

*C. intestinalis* is a hermaphroditic broadcast spawner, releasing eggs and sperm into the water column where fertilization occurs (Byrd and Lambert, 2000). At the onset of gamete production, sperm is produced in advance of the appearance of eggs (protandric hermaphroditism). The production of gametes is believed to be mainly temperature dependent, while their release is influenced to some degree by light (Lambert and Brandt, 1967). A study carried out in NS by Carver et al. (2003) assessed the reproductive capabilities of *C. intestinalis*, relating them to the seasonal changes in water temperature. They found that from November (7°C) to January (1°C) ripe eggs were present in the ovary, while egg resorption occurred in February and March (-1°C). The production of gametes

began again in April when water temperature increased to 4°C. *C. intestinalis* populations in NS are appear to commence spawning when water temperatures reach about 8°C (Ramsay et al., 2008a). While *C. intestinalis* has been observed reproducing continuously throughout the year, provided the water temperatures are suitable (Yamaguchi, 1975), the cited study was carried out on populations in Japan, most likely composed of *C. robusta*.

Eggs are viable for about 30 h and release a sperm chemoattractant until fertilization occurs (Carver et al., 2006). Upon fertilization, embryonic development varies with temperature, generally taking ~ 24 h before larval hatching occurs, resulting in a non-feeding larvae about 1mm in length. They tend to be short-lived and have the ability to swim, if relatively poorly. The duration of the pelagic larval stage is variable, again dependent upon temperature, while other factors are likely to influence this. Estimates tend to range between 1 and 5 days for maximum lifespan of the larvae. The salinity of the environment inhabited by the adult *C. intestinalis* has been shown to have an effect on the fertilization and larval development of their progeny. Regardless of the long-term salinity of an environment, parents exhibit high levels of phenotypic plasticity, resulting in improved larval tolerance and metamorphosis if they are present in a similarly saline environment to which their parents were subjected to in the recent past (Renborg et al., 2013). It is also suggested that lower pH values improve the survival and development of *C. intestinalis* larvae (Dupont et al., 2009).

#### 1.3.3.2. Dispersal

When considering an invasive species, not only is the primary means of introduction an important factor, but so too is the dispersal that follows on after it has become established in one area. In this case, the dispersal of *C. intestinalis* throughout NS and Atlantic waters is most likely to be a combination of natural and anthropogenic processes.

##### 1.3.3.2.1. Natural dispersal

Natural dispersal of planktonic invertebrate larvae is a major factor influencing the structure of benthic adult populations. It is common for recent studies to question the often assumed, but rather simplistic, theory that species with early planktonic life-stages have a good dispersive capability and ‘open’ populations (Ayre et al., 2009). The ability to accurately determine the scale at which marine larvae actually disperse and successfully settle is still poor, and remains unknown for many marine organisms (Miller and Morgan, 2013; Salinas-de-León et al., 2012). Understanding pelagic larval dispersal distance and subsequent

settlement is critical given the import role it can play in the formation of new benthic populations, as well as maintaining current ones (Lloyd et al., 2012).

For the majority of ascidians, dispersal takes place during the pelagic life-stages, the extent of which is generally limited to a scale of metres to kilometres. Due to their non-feeding, short-lived larvae, ascidians cannot disperse large distance (Lambert and Lambert, 1998). In regard to *C. intestinalis*, limited dispersal capabilities can be explained by the physical environment in which *C. intestinalis* tends to frequent, as well as biological processes exhibited by the organism. *C. intestinalis* tends to inhabit semi-enclosed or sheltered water bodies (Petersen and Svane, 1995), such as inlets and bays. This tends to reduce the exposure of planktonic larvae to hydrodynamic conditions that could increase dispersal of the planktonic life-forms. Indeed, a study in Nova Scotia (Howes et al., 2007) found that sheltered sites experienced greater *C. intestinalis* recruitment, in successive years, when compared to nearby sites exposed to stronger hydrodynamic forces, such as waves. Dispersal models have demonstrated that without the aid of anthropogenic vectors, the limit of dispersal for *C. intestinalis* does not tend to extend beyond the bay within which the adult population is located (Kanary et al., 2011; Petersen and Svane, 1995). This suggests that population establishment and spread may be particularly susceptible to local abiotic conditions. As for a dispersal-limiting biological processes exhibited by *C. intestinalis* itself, eggs are often released in mucus strings that can attach directly to the adult or drift off and attach to substrate nearby (Svane and Havenhand, 1993), a factor that may lead to aggregated settlement. Eggs that are released without mucus strings will sink in still water due to their negative buoyancy, adhering to available substrate (Berrill, 1947). This reinforces the limited dispersal already dictated by the surrounding hydrodynamic conditions.

Predation has also been suggested as a reason for spatial constriction of *C. intestinalis* populations. It has been shown that fishes and sea-stars rely on *C. intestinalis* as a food-source in Norway (Gulliksen and Skjæveland, 1973), while crabs have also been documented as using *C. intestinalis* as a food source in Atlantic Canada (Carver et al., 2003). Micro-predators have also been described to feed on a variety of juvenile ascidians, *C. intestinalis* amongst them (Osman and Whitlatch, 2004, 1995).

The total length of time from gamete release to larval settlement is an important factor in determining the potential dispersal distance of *C. intestinalis*. The longer embryonic and larval development takes, the greater the potential for dispersal is. These development times vary with temperature. Embryonic development time consistently decreases with an increase

in temperature, with a reported range for North Atlantic populations anywhere between < 24 h to 63 h, depending on the temperature and location (Dybern, 1965; Svane and Havenhand, 1993). Larval development and survival is also temperature dependent, with variable times reported in the literature. Under laboratory conditions, the upper time limit of the larval stage tends to be reported approximately as five days (Collin et al., 2013). Despite this, it is thought that under natural conditions the larval stage is likely to be closer to hours than days, with most settlement occurring within 24 hours (Havenhand and Svane, 1991; Svane and Havenhand, 1993). Salinity and pH conditions have also been shown to affect larval metamorphosis and survival (Dupont et al., 2008; Renborg et al., 2013), playing a potential role in the length of time larvae are in the pelagic phase.

#### 1.3.3.2.2. Anthropogenic dispersal

Ascidians are prominent in the overarching global trend of a significant increase in introductions and spread of NIS (Ricciardi et al., 2013), thus increasing the number of biological invasions. This trend can be predominantly attributed to the simultaneous increase in the quantity and size of commercial vessels, as well as recreational boating, rather than a change in the natural dispersal mechanisms of ascidian species (Lambert and Lambert, 1998). As predicted by Carver et al. (2003), the range expansion of *C. intestinalis* in NS (Fig. 1.3 - Fig. 1.7), evident throughout field monitoring carried out between 2006-2009, is most likely due to the translocation of adult individuals attached to boat hulls from one area of infestation to a suitable location previously not subjected to infestation (Sephton et al., 2011). Zhan et al. (2012) added further weight to this argument, finding that the haplotype diversity of neighbouring populations in NS and PEI was vastly different at times, while very similar genetic patterns were present in populations much further apart than is reasonable to expect resulted from natural dispersal. The most likely cause for these patterns is the movement of individuals via recreational and commercial boats. In addition to anthropogenic vectors, anthropogenic influences are increasing the frequency and intensity of storm events (Pachauri et al., 2014). This could potentially increase the number of extraordinary dispersal events due to storm surges.

#### 1.3.3.3. Settlement

Unlike a number of other ascidians, *C. intestinalis* does not demonstrate gregarious settlement, but the presence of adult aggregations do seem to affect the settlement of larvae due to their influence on water flow patterns (Havenhand and Svane, 1991). Rius et al. (2010)

came to a similar conclusion when they studied the effects of conspecifics on the settlement of larvae, finding that *C. intestinalis* adults do not appear to have any influence over larval behaviour due to chemical cues. Rather, the physical conditions either created by adults or environmental conditions play a greater role.

*C. intestinalis* larvae have been shown to sink to greater depths where light intensity is lower, and then proceed to swim upwards to find a darker settlement site (Howes et al., 2007; Rius et al., 2010). Despite the preference for settlement on shaded substrates, Rius et al. (2010) found that under laboratory conditions, when the larvae were provided with no choice between shade and no shade on experimental settlement surfaces, the same proportions settled on the bottom, sides and top of the experimental chambers, indicating that *C. intestinalis* larvae will settle at the same rate in light conditions when no dark substrate is available. The preference for dark conditions and the habit of sinking initially or swimming upwards leads to the majority of *C. intestinalis* individuals in the field found on downward facing substrate (Rius et al., 2010).

#### 1.3.3.4. Recruitment and adult stage

*C. intestinalis* recruitment in NS appears to experience two peaks per growing season, typically one in late spring/early summer and another in late summer/early autumn. Carver et al. (2003) observed a peak in recruitment in May-June, with a second event peak occurring in August-September of 2000 in Lunenburg Bay, NS. A similar two-generation recruitment cycle was noted by Howes et al. (2007) during a monitoring program in nearby Mahone Bay, NS from 2002-2005. Individuals mature rapidly, in a matter of weeks (8-10), with reproduction reliant on size rather than age, and exhibit a high reproductive output (Carver et al., 2006). *C. intestinalis* sampled in NS have a maximum body length of ~ 140 mm (Ramsay, 2009). Fecundity of NS individuals is suggested to be in the region of a total lifetime production of 12,000 eggs (Carver et al., 2006). This figure is relatively consistent with populations of sampled in Scandinavia by Petersen and Svane (1995) who estimated a total lifetime fecundity of 10,000 eggs per individual. In northern European waters, the life cycle of *C. intestinalis* is between 12-18 months, with the growth and longevity of individuals thought to be predominantly influenced by food supply and temperature (Carver et al., 2003). In southern Scandinavia most individuals frequent depths of c. 25m and shallower, but populations at depths of 100 m and deeper have been recorded (Dybern, 1967). The highest rate of recruitment in a study carried out in NS always occurred at 4.5 m with little or no recruitment at depths of 0.5, 12.5 or 16.5 m (Howes et al., 2007). This pattern of preferable



settlement and recruitment at a depth of 4.5 m is particularly relevant given that it's a common depth that mussel lines are suspended at. Most studies that have specifically observed or noted the type of hydrodynamic conditions that *C. intestinalis* tends to grow in have focused on the difference between wave exposed sites compared to more sheltered sites, finding that it prefers sheltered and low flow sites (Carver et al., 2003; Howes et al., 2007; Petersen and Svane, 1995). In my study, I am primarily monitoring growth within sheltered habitats and therefore exploring the effect of varying water flow within sheltered regions. Some studies have investigated the stress tolerance of *C. intestinalis* adult individuals at the gene expression level, generally focused on the heat shock protein chaperones and endoplasmic reticulum chaperones (Fujikawa et al., 2009; Kawaguchi et al., 2015; Sato et al., 2015; Serafini et al., 2011). *C. intestinalis* is suggested to be a species capable of responding to a wide range of stress via molecular mechanisms, but the degree to which this is variable amongst populations and how it relates to growth is not yet clear.

#### *1.3.4. Why is C. intestinalis a successful invader?*

Given many tunicates display what would appear to be a highly localized dispersal pattern, the aid provided by global shipping networks, aquaculture trade and recreational boating to overcome the barrier of dispersal has also greatly increased their prevalence (Ramsay et al., 2008a), where they otherwise would not have had an opportunity to establish populations (Rigal et al., 2010). Mechanisms presented to explain the success of tunicates at colonising and establishing substantial populations on both natural and artificial substrates are “early onset of reproduction”, “high reproductive rates”, and “fast colonization and growth rates” (Lutz-Collins, 2009). It takes a relatively small period of time before ascidians reach sexual maturity due to their rapid growth rate, which in unison with what can be a prolonged breeding season can lead to a rapidly established breeding population from only a handful of individuals (Lambert and Lambert, 1998). Tunicates (introduced and native) are strong spatial competitors and prominent biofouling organisms. At times they make up 100% of the biofouling community (Carman et al., 2010), and they are a problem on a global scale (Adams et al., 2011; Fletcher et al., 2013). Artificial habitat is provided by aquaculture in areas such as estuarine environments where they would normally be lacking in hard substrate naturally available for benthic species, such as tunicates (Woods et al., 2012). The theory that the availability of unused resources, such as space, can provide a suitable environment for

invasion (Davis et al., 2000) is a potential explanation for the success of *C. intestinalis* in artificial environments where there is an abundance of unoccupied space for individuals to establish. The provision of optimal substrate, at a suitable depth may have contributed to localised explosions of *C. intestinalis* populations (Howes et al., 2007).

But why is *C. intestinalis* the prevalent tunicate implicated in aquaculture biofouling in areas where multiple NIS and invasive tunicates have been recorded? In PEI, four introduced tunicates species of great concern have been recorded in the recent past (Ramsay et al., 2008a). Despite a well-known invasive solitary tunicate, *Styela clava*, being present in PEI waters since 1997, *C. intestinalis* experienced a rapid rise to being the dominant fouling species in PEI. It had all but eradicated the previously dominant *S. clava* within two years of the first documentation of its presence in the Brudenell estuary (Ramsay et al., 2008a). The breeding season of *C. intestinalis* in NS is not characteristic of the typical reproductive cycle as described from its native range by Dybern (1965). There it is relatively confined to a few months in the summer and generally occurs in mature individuals (> 1 year old). In NS it displays a prolonged reproductive season and individuals reach sexual maturity in weeks to months (Ramsay et al., 2008a). *C. intestinalis* gains an advantage over other tunicate species through earlier recruitment due to a commencement of the reproductive cycle at lower temperatures, e.g. *C. intestinalis* begins at 8°C, while *S. clava* begins its reproductive cycle when water temperatures reach 12°C (Ramsay et al., 2008a). This early start compared to other biofouling species is crucial in an environment where competition for space is high (Lambert, 2005). There is a possibility that *C. intestinalis* can experience three generations within a single season, while it is regular for new recruits to reproduce in their first year (Ramsay et al., 2008a). High propagule pressure is important for successful invasion, as highlighted in sec. 1.2.4., and *C. intestinalis* checks all the boxes in this regard. High fecundity coupled with the hydrodynamic patterns and short-lived planktonic life-stage that lead to localised settlement result in the formation of dense aggregations displaying a spatially heterogeneous but locally abundant population distribution (Howes et al., 2007). Rius et al. (2010) suggest that this behaviour, much like true gregarious behaviour can be a prerequisite or advantage when it comes to successfully invading when introduced. As well as high reproductive capacity and settlement early in the year in dense space occupying aggregations, *C. intestinalis* also seems to hinder the settlement of other species once it is established due to its unsuitable soft tunic and mucoidal surface (Ramsay et al., 2008a). In regard to the popular hypotheses explaining successful invasions discussed in sec. 1.2.4.,

there is still much we don't know about *C. intestinalis*. Little is known about the parasitic loading effect that could be introduced with *C. intestinalis*. Although, a risk assessment survey completed by experts in this field suggests that harmful parasite introduction with *C. intestinalis* was deemed to not be of major concern, in comparison to the other impacts *C. intestinalis* has on the introduced environment (Therriault and Herborg, 2008b). No work has been conducted on the potential positive effect of parasite release in the introduced range and how this may help *C. intestinalis* establishment and spread. It is possible that *C. intestinalis* benefits from reduced predation in the introduced range. Although feeding trials showed that individuals were preyed upon by Atlantic rock crabs, it is unclear how the rate of feeding compares to rates of feeding in the native range. Also, green crabs had a lower feeding rate on *C. intestinalis* compared to rock crabs, as well as eating primarily smaller individuals, leaving the larger adults (Carver et al., 2006). It is plausible that an 'invasional meltdown' type occurrence may occur in the field where the green crab invasion problem could assist *C. intestinalis* through the reduction of predators that would prey upon it at a greater rate. In addition to different predator types, the main observed habitat for *C. intestinalis* in introduced ranges tends to be artificially engineered environments such as harbours and aquaculture sites (Lambert, 2005). *C. robusta* gains an advantage in artificial environments on the west coast of Canada, as the created environment provided a refuge from predators, allowing the establishment of large populations (Dumont et al., 2011). A similar scenario could be occurring in NS for *C. intestinalis*, with man-made environments providing a biofouling 'playground'.

When considering a species with a wide tolerance range for many key abiotic variables in the marine environment, the conditions of the receiving environment may be considered as less important to invasion success. But the fact remains that we can observe this heterogeneous distribution of *C. intestinalis* populations around NS. The extent to which the abiotic conditions of the receiving environment influence the successful establishment and spread is not yet known. Variable abiotic conditions may be interacting with all the different life-stages of *C. intestinalis*, resulting in changeable spatial and temporal responses of the species.

### 1.3.5. *C. intestinalis* impact

#### 1.3.5.1. Ecosystem and biodiversity

A large component of biofouling species are strong spatial competitors, managing to reach very high densities or biomass over relatively short spaces of time (Blum et al., 2007). This

strong ability to outcompete native species for space is demonstrated by *C. intestinalis*. Blum et al. (2007) conducted manipulative experiments in San Francisco Bay, testing the effects of *C. intestinalis* (more likely *C. robusta*) on the sessile invertebrate community. Using fouling panels made of PVC, three treatment types were implemented: experimental removal of *C. intestinalis* on a weekly basis; a manipulated control that was identical to the experimental removal procedure, except no removal was carried out; and a control that remained in the water for the duration of the experiment.



**Figure 1.8: Left - *C. intestinalis* individuals finding kelp to be a suitable substrate. Right – *C. intestinalis* completely covering a settlement plate, out of the water.**

The presence of dense *C. intestinalis* aggregations reduced overall species richness. It also reduced the abundance of key species such as *Bugula neritina*. The communities of sessile invertebrates differed significantly depending on the presence or absence of *C. intestinalis*, with some species exhibiting strong associations with *C. intestinalis* dominant assemblages, with others largely absent when it is present. On the whole, the structure of the communities was altered significantly by the presence of dense *C. intestinalis* aggregations, with the majority of species negatively affected, with a small number of species having a positive association. As well as the effective space occupation of *C. intestinalis* described, it also provides a poor substrate for other settlers, through the use of antifouling chemical defences, strong anti-microbial compounds, and mucus and surface cell sloughing (Blum et al., 2007). While this key publication is most likely reporting results from experiments carried out on *C. robusta*, the similarity between the morphology and behaviour of the two species suggests that *C. intestinalis* could influence the surrounding environment in a similar fashion.

While this experiment clearly shows that high density *C. intestinalis* assemblages have the potential to vastly change community composition, it is unclear how the effects recorded in this experiment would translate on a larger spatial and temporal scale. Blum et al. (2007) hypothesise that other NIS of ascidians known to have achieved similar high density invasions in their introduced ranges have the potential to significantly affect community structure. Given the growing trend of tunicate invasion around the world, this could lead to drastic changes in community richness, with unknown consequences for ecosystem functioning.

The effects of *C. intestinalis* cover on mobile fauna is an area of limited research (Blum et al., 2007). A study carried out by Lutz-Collins (2009) did investigate the effect of *C. intestinalis* fouling on native species with different levels of mobility that are associated with mussel lines. While they found the fully sedentary species to be most affected, there is one obvious drawback in the study's design. It took advantage of pre-existing differences between biofouling levels rather than using a manipulative experimental approach. This could lead to different effects on the native groups between differing fouling intensities due to other factors, such as food availability, hydrodynamic conditions, abiotic influences and a number of other potential conflicting factors (Fletcher et al., 2013).

The nature of studies carried out using artificial panels replicates the process of *C. intestinalis* infestation on artificial substrates such as docks, aquaculture gear and boats. It is plausible that if *C. intestinalis* establishes assemblages on natural substrate, the impacts would be consistent with those documented on artificial substrates. What is presumably *C. robusta* has been recorded as an introduced species in California as early as 1917, and remains abundant to this day, while some of the native benthic species previously documented as being abundant are now considered rare or have disappeared from harbours completely (Lambert and Lambert, 1998). This pattern describes the potential for long-term structural changes in biological communities that co-occur with *C. intestinalis* infestations. It certainly holds the potential to contribute in the demise of some previously abundant native species.

#### 1.3.5.2. Social impact

The acknowledgement that ascidians are a key group of marine fouling organisms that have a significant effect on aquaculture operations is not a recent occurrence, with studies such as Yamaguchi (1975) investigating the growth and reproductive cycles of four ascidians well known for their disruptive characteristics. In particular, *C. intestinalis* fouling of shellfish and

aquaculture equipment has a significant economic impact on mussel farms around the world (Robinson et al., 2005), significantly increasing production and processing costs (Comeau et al., 2012). Biofouling communities can add significant weight to aquaculture systems, which in turn require additional buoyancy and anchoring, a further expense for the farmers (Sievers et al., 2013). The additional weight caused by *C. intestinalis* fouling can even lead to entire crops of mussels being lost, in severe cases (Comeau et al., 2012). Another major problem is the additional weight added to the boat when harvesting mussels, thereby reducing efficiency (Stairs, pers. comm.). These combined effects of *C. intestinalis* fouling can lead to mussel farm closure, resulting in a negative impact for local economies, but also jeopardising an important resource for future food security.

#### **1.4. Prevention and mitigation efforts**

##### *1.4.1. Prevention of introduction and establishment*

The only surefire way to prevent biological invasion problems posed by NIS is to stop introductions occurring. The scale of this task, especially in the marine environment is colossal and virtually impossible, even on a regional scale. In order to maximise the effectiveness of resources available to monitoring and invasive species managements, accurate risk assessments of potential introductions are essential. Such a risk assessment has recently been devised for the Canadian marine environment (Drolet et al., 2015) and continual improvements and adjustments are essential to maintain the efficacy of such methods. Using these risk assessments can minimize the waste of resources monitoring and preventing introductions of NIS that pose little risk to the introduced environment. In the marine environment, there has also been an emphasis on understanding the vectors that pose introduction risk and how to assess these vectors and model them in order to best allocate management resources (Floerl et al., 2009; Murray et al., 2013; Ruiz and Carlton, 2003; Sylvester et al., 2011).

While *C. intestinalis* has already been introduced in NS and other Atlantic Canadian waters, it is still vital that monitoring similar to that carried out by Sephton et al. (2015, 2014, 2011) continues in order to understand the extent of issue and how best to counter it.

##### *1.4.2. Reduce spread and impact post-introduction*

Some success has been achieved in limiting the spread of tunicates in PEI through the ‘Introductions and Transfers permitting process’, but this has no authority over vectors

outside of the aquaculture industry, such as recreational boating and can't prevent natural dispersal (Kanary et al., 2011). *C. intestinalis* dispersal modelling has been conducted in PEI in an effort to understand the potential for natural dispersal (Collin et al., 2013; Kanary et al., 2011). These dispersal estimates have assisted in developing the best possible management protocols to prevent further sites falling under infestation, a scenario that would have severe economic impacts. The results from the model emphasised the importance of regular maintenance of potential settlement nodes for *C. intestinalis*, which range from buoys to wharfs. This could minimize the risk of a stepping-stone process occurring, allow the spread of *C. intestinalis* within harbours and bays, especially ones that are currently not infested. Collin et al. (2013) stress the importance of microscale localized modelling in unison with macroscale species distribution models as a useful tool in combatting nascent populations of invasive species. Dispersal of invasive species is a complex myriad of biophysical factors including species-specific life-history characteristics that are often variable given abiotic factors interacting with physical processes, such as currents and tidal movements, combining to make it very difficult to establish the extent of possible dispersion with accuracy (Kanary et al., 2011). This emphasises the importance of implementing management plans based on the most recent and comprehensive research available, as well as striving to improve the available knowledge about dispersal and population establishment. Collin et al. (2013) suggest that a stricter management of vector movement is vital in areas where known problematic species are discovered. They suggest that the problem could have been limited further by immediate vector management in PEI, due to natural dispersal being less likely to transport *C. intestinalis* from bay to bay.

#### *1.4.3. Long-term mitigation and control*

Biological control is an option that has been investigated to attempt to control *C. intestinalis*, again in PEI. Approximately 73,000 Atlantic rock crab, *Cancer irroratus*, were caught and transplanted to an aquaculture site in an estuary of PEI, as this is a species that has been shown to prey upon *C. intestinalis* (Comeau et al., 2012). A number of the transplanted crabs were fitted with tracking equipment to investigate their movements. The transplanted individuals did not remain in the area they were moved to and Comeau et al. (2012) do not believe this is a viable means on biological control for *C. intestinalis* infestation.

Various efforts have been made to investigate potential management plans that mussel farmers could implement to limit the loss of mussels, as well as reducing mitigation costs.

One such study was carried out by Ramsay et al. (2008b), who considered the effect of mussel stocking density on *C. intestinalis* infestation in PEI. Using three levels of mussel density (high, medium, and low), several characteristics of *C. intestinalis* growth were monitored at several points in the growing season. The hypotheses being tested were: higher mussel density stocking would reduce the number of viable *C. intestinalis* larvae for settlement due to a larger amount of feeding; the surface area per mussel is lower at a high density allowing less space for *C. intestinalis* to utilise; and that as mussels display self-thinning, the mussels that are subject to *C. intestinalis* settlement fall off due to stress on the byssal attachment leaving unfouled mussels attached. The study highlights the temporal differences exhibited in *C. intestinalis* populations, with significant differences in *C. intestinalis* biomass between different stocking densities in August, but by October no differences were evident. The length of *C. intestinalis* individuals was also measured, with lower density stocked mussel socks displaying individuals of the greatest length, both in August and October. The results of this study did support the self-thinning hypothesis, as by the end of the study period there was no difference in mussel length or condition between different densities. The expectation is that the mussels that are present in higher density will experience lower growth rates and worse condition than those of stocked at a lower density. It is possible that the effects of *C. intestinalis* fouling caused the unexpected results (Ramsay et al., 2008b). This study emphasises that the timing of mitigation efforts are likely to be important, while a combination of factors such as density of stocking and timing of treatment to eradicate or limit *C. intestinalis* growth could be key (Ramsay et al., 2008b). A range of other treatment regimens have been tested, including desiccation trials, pressure washing, aeration/turbidity treatments, chemical sprays, and other chemical applications (Carver et al., 2003; Hopkins et al., 2016; Locke et al., 2009; Lowen et al., 2016; Paetzold and Davidson, 2011). The treatment of choice depends on a variety of factors such as effects on the cultured organism, environmental concerns, cost and timing of treatment, with some treatments, particularly in combination being effective, if costly and time consuming.

### **1.5. Research objectives**

The underlying outcome of this research, regardless of the success in explaining abiotic and *C. intestinalis* population relationships was establishing baseline information relating to the degree of variability between populations of a species of importance on a large spatial scale. Whether *C. intestinalis* is viewed as a problem or perhaps as a potentially useful species (the first thing anyone asks is “What can you use it for?”), the fact remains that little is known



about its population dynamics and this needs to be rectified. In addition to this, a comprehensive abiotic dataset was collected for important variables in the marine environment. To ascertain what relationships were observable given the abiotic variables being monitored, I tested a range of hypotheses (in the form of models) to determine the most likely explanatory combination of abiotic variables to account for the observed *C. intestinalis* data. Any relationships between *C. intestinalis* abundance and environmental conditions may help to refine assessment of aquaculture sites for their risk of infestation, affecting profitability of both existing and potential future sites.

Invasion ecology is an expanding field making it desirable to understand factors that lead to population increases in species of interest. This species-specific work in turn could contribute to general understanding of the spread and success of other invasive species, therefore informing future prevention and mitigation strategies. If successful, this approach to modelling could be transferable to other troublesome biofouling species, providing a framework to approach a similar research project.

Finally, with increasing pressure on the environment from global change, baseline monitoring provided by research, such as this study, is essential for assessing the scope of such changes.

## 2. Materials and methods

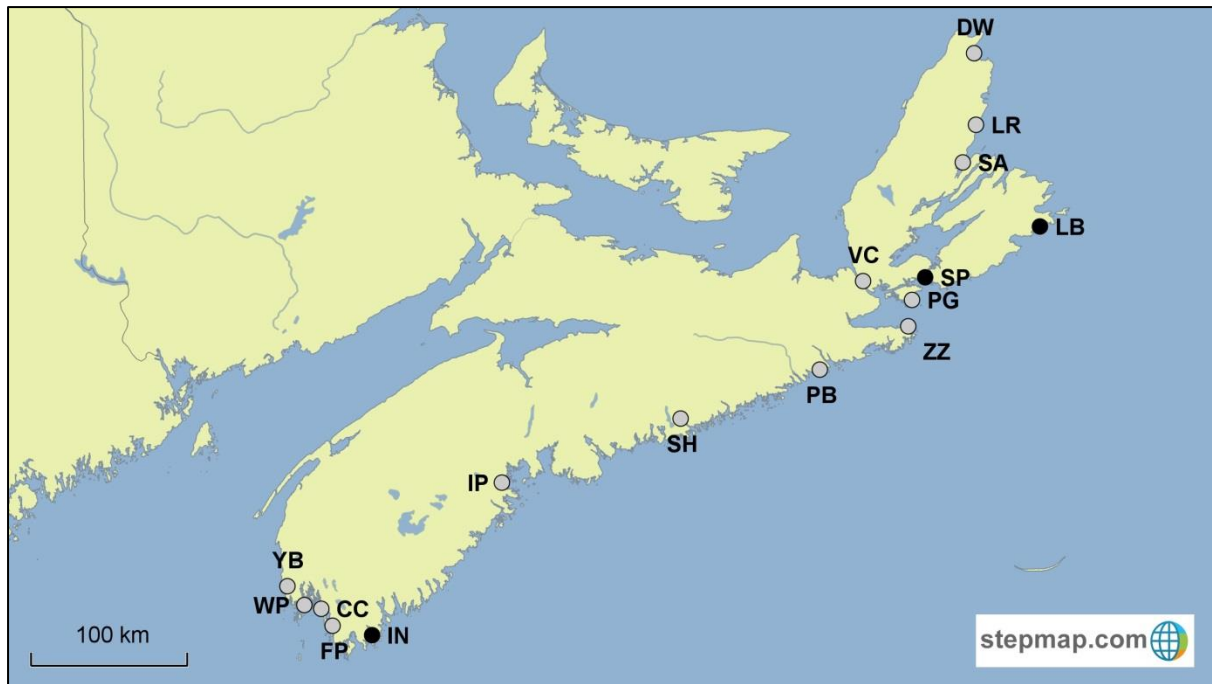
### 2.1. Field monitoring sites

#### 2.1.1. Site selection

I selected monitoring sites aiming to evenly represent a spectrum of *C. intestinalis* abundance from absent to high abundance. I decided to limit the selection to sites that are predominantly on the Atlantic coastline of NS. Testing within one major oceanographic region was an attempt to limit potential confounding effects outside of the abiotic variables selected for monitoring. The main contributing factor in the selection process was the historical presence and abundance records of *C. intestinalis*. Historical records were provided by the DFO (Sephton, Pers. Com.). I also wanted sites to span the entire NS Atlantic coastline, as evenly as possible. The location (Fig. 2.1) and additional information (Table 2.1) are provided for each site.

**Table 2.1: The 16 sites monitored in 2014 and 2015 (sites in italics were only monitored in 2015), including the name, site code, type of site, and GPS co-ordinates.**

Name	Code	Type	Latitude	Longitude
Camp Cove	CC	Harbour	43.724171	-65.839937
Cape Canso	ZZ	Marina	45.334817	-60.985151
Dingwall	DW	Harbour	46.903055	-60.460756
Falls Point	FP	Harbour	43.531117	-65.740845
Indian Point	IP	Mussel lease	44.457128	-64.306808
<i>Ingomar</i>	<i>IN</i>	<i>Harbour</i>	<i>43.563249</i>	<i>-65.364162</i>
Little River	LR	Harbour	46.446528	-60.458987
<i>Louisbourg</i>	<i>LB</i>	<i>Harbour</i>	<i>45.917759</i>	<i>-59.971003</i>
Petit-de-Grat	PG	Harbour	45.507234	-60.960693
Port Bickerton	PB	Harbour	45.104438	-61.721917
Ship Harbour	SH	Mussel lease	44.808099	-62.846649
St Ann's Bay	SA	Mussel lease	46.265983	-60.583567
<i>St Peters</i>	<i>SP</i>	<i>Marina</i>	<i>45.661397</i>	<i>-60.874862</i>
Venus Cove	VC	Sheltered Dock	45.614638	-61.390123
Wedgeport	WP	Harbour	43.713973	-65.968626
Yarmouth Bar	YB	Harbour	43.815751	-66.148124



**Figure 2.1: Distribution map of the sites monitored during this project. The 13 sites with grey markers were monitored in 2014 and 2015 (Core Sites), while the 3 sites with black markers were only monitored in 2015 (New Sites).**

### 2.1.2. Site groupings for analyses

Thirteen sites were selected in 2014, while three additional sites on top of the original 2014 sites were chosen for the 2015 monitoring season. To work around the missing data constraints for some analyses, as well as for theoretical reasons, it was necessary to categorise sites into different groups. The first group of thirteen sites monitored in both 2014 and 2015 are referred to as the ‘core sites’. Three additional sites were monitored in 2015, Louisbourg, Ingomar, and St Peters. St Peters’ (SP) data was retroactively removed from the dataset before any analyses were conducted. This is due to the major difference in waterbody type between SP and all other sites. SP is located in the southern region of the Bras d’Or Lakes in Cape Breton. All other sites are harbours, mussel leases or sheltered coastal sites. These are all nearshore marine environments compared to the saline lake environment of the Bras d’Or Lakes.

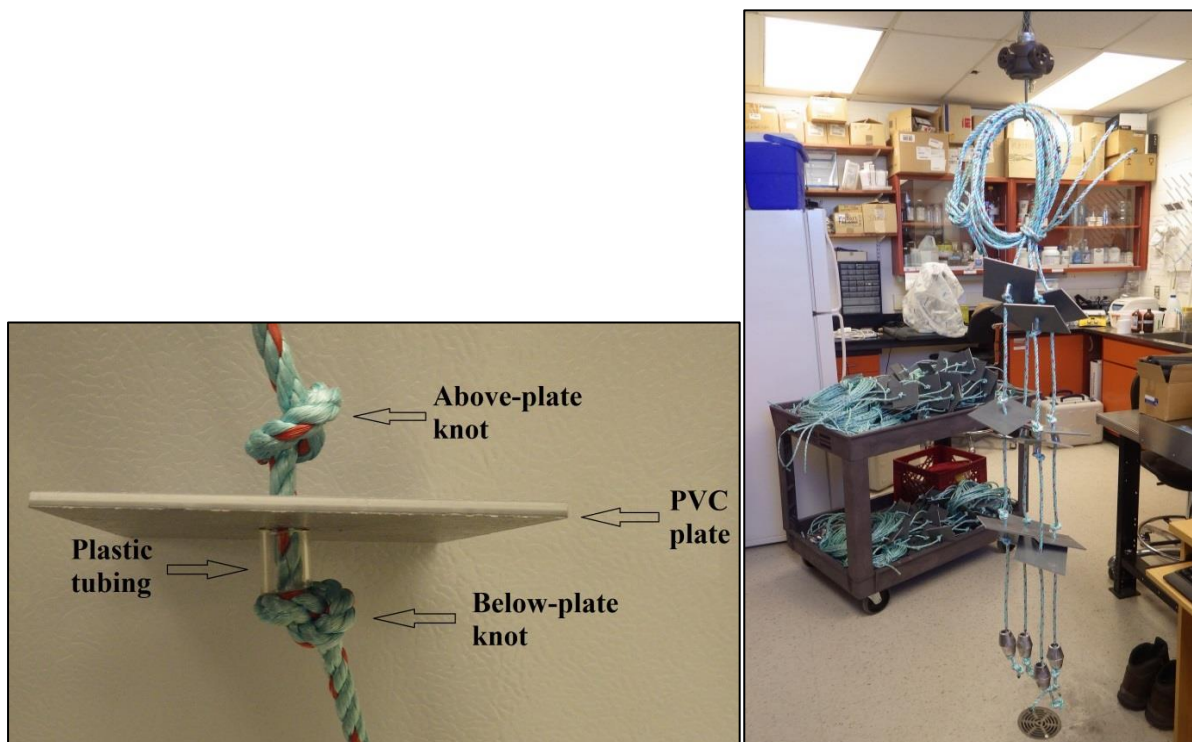
The expanded group of 15 sites in 2015 is referred to as ‘expanded sites’. I also categorised a group excluding sites with no *C. intestinalis* presence. This is due to the difficulty in distinguishing the difference between a site where *C. intestinalis* is absent because it has simply never been introduced into that site, or whether an introduction has taken place in the past but environmental conditions have resulted in the inability of the introduced individuals

to establish a population. When working exclusively with sites at which *C. intestinalis* is present (regardless of abundance), the groups are referred to as either ‘core sites - minus absent’ or ‘expanded sites - minus absent’.

## 2.2. Biotic monitoring

### 2.2.1. Settlement collector design

The presence and abundance (percent cover) of *C. intestinalis*, as well as any other sedentary biofouling species, was monitored using standard settlement collectors consistent with previous monitoring efforts carried out by DFO (Sephton et al., 2011). Each settlement collector (Fig. 2.2) consists of three polyvinyl chloride (PVC) settlement plates, polypropylene rope (8 mm diameter), and a lead weight (250 gm). Each PVC plate had dimensions approximate to 10 cm x 10 cm x 3 mm, with a hole drilled in the centre for the collector rope to thread through. All PVC plates were all lightly sanded in a consist manner.



**Figure 2.2:** Left – The setup of the individual settlement plates; Right – Examples of the entire collector.

The three plates were evenly spaced at intervals about 25 cm in length along the rope. A knot in the rope above and below each plate prevented them from moving freely up or down the rope. The lead weight was tied onto the bottom end of the rope about 25 cm below the

bottom-most plate. This was to weigh the collectors down once deployed in the water, ensuring the collectors hung vertically in the water column. The knots in the rope were positioned so that plates were not entirely constricted in their movement. The plates were free to spin around the rope they were attached to, as well as having the potential to move roughly 1 cm along their vertical axis. A small piece of flexible plastic tubing (Fig. 2.2) was placed between each of the plates and the bottom knots in order to minimize friction caused by the plate movement around the rope. Once the plates were tied onto the rope with the plastic tubing and knot below them, their orientation when in the water column was horizontal, with an acceptable degree of variation depending on the given knot, plastic tube and plate.

### *2.2.2. Collector deployment: intra-site location and depth*

Permanent collectors (cumulative growth) that remained in the water the entire monitoring period, as well as monthly collectors (monthly growth) that were replaced during each site visit, were deployed. The quantity of collectors deployed at each site can be found in Table 2.2. When deployed, the collector plates were suspended horizontally in the water at 1.5, 1.75 and 2 m depth, with the lead weight not touching the substrate (with three exceptional sites – see below). The collectors were deployed relatively close together at each site since my goal was to test inter-site variability and minimize intra-site variability. In addition, this approach minimized the distance between the single set of data loggers at each site and the sampling collectors, improving the correspondence between the abiotic and biotic data collected. Typically, collectors were hung from the side or between docks (tied to moorings or suitable fixtures), and spaced every 2 m to minimize the chances of two collectors tangling with each other. Deployment was sometimes constrained by dock geometry or minimizing interference with other dock users. At Indian Point and St Ann’s Bay, the collectors were tied directly to the long-lines used in mussel aquaculture. At the remaining mussel lease, Ship Harbour, an independent float was used to suspend all collectors (separated by 0.5 m due to space constraints). At Yarmouth Bar, Wedgeport, and Ingomar, the shallow harbour did not allow for settlement collectors to be submerged to a depth of 1.5 m at the uppermost settlement plate. Instead, the uppermost settlement plate was submerged at a depth of 1 m or 0.5 m depth.

**Table 2.2: The number of collectors deployed in 2014 and in 2015.**

<b>Year</b>	<b>Permanent Collectors(PC) (per site)</b>	<b>Monthly Collectors (MC) (per site)</b>	<b>PC Exceptions</b>	<b>MC Exceptions</b>
2014	3	2	WP (4)	Not applicable
2015	5	2	IP (4), SA (4)	SA (1)

### *2.2.3. Settlement plate monitoring*

Settlement plates were photographed during each monthly site visit (see Appendix A for dates of site visits). These images were used to track the growth of the biofouling community on each plate over the monitoring period (representative plate images for each site in Appendix F). Sites were visited at approximately four week intervals. The time varied from month to month due to circumstances such as inclement weather conditions and availability of personnel at mussel leases to provide assistance. In general, the order in which the sites were visited was kept as consistent as possible. The large spatial range over which monitoring was being conducted resulted in a lag between the first and last site visit in any given measurement period. I aimed to minimise the time between the first site sampled and the last site sampled in any given measurement period, while keeping the time between successive visits to each site as consistent as possible. The southern shore of the NS mainland experiences higher water temperatures earlier in the year when compared to Cape Breton, in particular. This is why I decided to commence the site visits in the southern region of mainland NS and proceeding north.

### *2.2.4. Settlement plate holder*

The settlement plate holder (Fig. 2.3) was manufactured in the St FX University Machine Shop by Steven MacDonald. Its purpose was to hold the settlement plates while they were being photographed, so that the plates were horizontal and could be orientated consistently for every image. It is made up of three identical PVC sections, each of which was just large enough for the settlement plates to sit on top of. All four sides of each section had a groove

cut out so that the collector rope could rest in them so as not to interfere with the settlement plates' orientation.

### 2.2.5. Light-box

A light-box was used while taking pictures of the settlement plates in order to improve image quality and provide consistency between images. The main purpose of a light-box is to eliminate all ambient light from the environment in which the images are being captured. This provides the user with an ability to apply a consistent light source to illuminate the object of interest to a desired level. The light-box used in this project was made from a rectangular window panel plant pot (Fig. 2.3). The internal dimensions of the light-box were 58 cm x 25 cm x 16 cm. The light source of choice was battery-powered LED lights, attached to the interior walls of the lightbox (Fig. 2.3). Given the aqueous nature of the working environment, glare caused by the reflection of light on the wet surfaces of both biological and inorganic objects was an issue that needed to be dealt with.



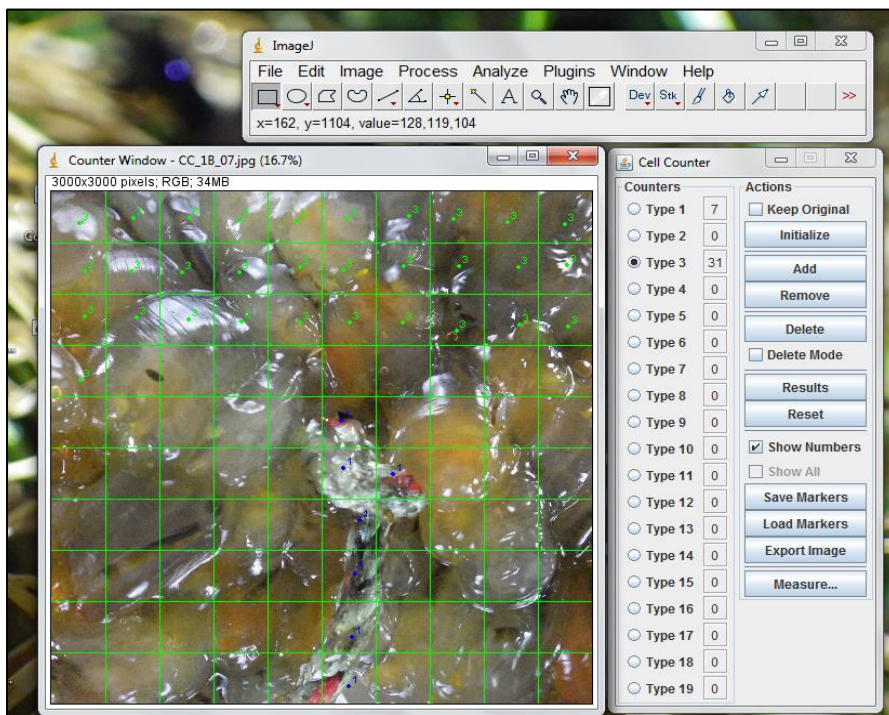
**Figure 2.3: Top left – Exterior of the lightbox; top right – Interior of the lightbox; bottom right – End-on view of the settlement plate holder; and bottom left – Side-on view of the settlement plate holder.**

This issue was countered by covering the LED lights with a semi-transparent white material. Both waterproof labels and white garbage bags were used for this purpose, folded over until

the desired light intensity was found. This covering acted as a light diffuser, thus reducing glare on wet surfaces and improving the overall quality of the images obtained. When images of settlement plates were being taken, a foam board acted as a base underneath the light-box. These foam boards were spray painted black so that light did not reflect off the surface of the foam. The settlement plate holder would be placed on top of the foam board. The settlement plates would be placed appropriately onto the holder. Then the light-box was placed over the top and images of each plate were taken through the holes cut in the top of the light-box.

### 2.2.6. Image processing and analysis

Seascape (Teixidó et al., 2011) and photoQuad (Trygonis and Sini, 2012) are free software packages specifically designed for the image analysis of benthic marine communities. Both can be useful for a range of ecological purposes and they were considered and tested for use in this project. However, the semi-automated procedures that each program uses were not ideal for the images in this project. Given the translucent nature of *C. intestinalis*, the region-of-interest selection implemented by specialized algorithms within each program was inaccurate beyond an acceptable level. A manual approach for calculating species percent cover was chosen instead.



**Figure 2.4: An example of a settlement plate image, cropped, scaled and ready for percentage cover analysis with the grid overlay applied in ImageJ.**



All images were cropped so that the extremities of the settlement plates were the new border of the images. Using a batch conversion in ImageJ, all of the cropped images were scaled so that their dimensions were all identical (3000 x 3000 pixels). Using a custom macro in ImageJ, a grid consisting of 100 grid cells was overlaid onto each image (Fig. 2.4). Given the approximate size of each settlement plate (c.10 x 10cm), each grid cell is approximately 1 x 1 cm<sup>2</sup>. The selection rule chosen for specifying the species present in each grid-cell was as follows: if 50% or more of a given cell is occupied by a single species, then that cell is counted as one cell for that species using the ‘Cell Counter’ application in ImageJ. Only sessile species that attach directly to the settlement plates were considered for this percentage cover calculation. Mobile invertebrates and epibionts were not considered.

When there were no clear, distinguishable species visible on the settlement plates, but it was clear that there was a layer of unidentifiable growth (most likely a biofilm), the cells were assigned a generic term of ‘marine slime’. If no single species occupied 50% or more of the space in a given cell, when there were three or more species in a single cell, these cells were categorized as ‘Other’. The underlying logic of the 50% or greater rule is that the number of cells with less than 50% of a species will on average balance out with the cells that have 50% or more. All grid-cells obscured by the collector rope were discounted from the total number of grid-cells used to calculate the final percentage values.

### 2.3. Abiotic monitoring

#### 2.3.1. Environmental data loggers

**Table 2.3: A description of all the environmental data-loggers used in 2014 and in 2015. The data loggers used to measure each abiotic variable are specified by year. The data loggers were used at all sites unless specified in brackets. The full logger names follow: WQL = WTW WQL-pH Logger; CT2X = INW Aquistar® CT2X Conductivity Smart Sensors; HOBO-G Accelerometer = Onset HOBO Pendant® G Data Logger; HOBO-Temp. = Onset HOBO Pendant® Temperature/Light Data Logger; and YSI 600XLM = YSI 6-Series 600XLM Multiparameter Water Quality Sonde. \*The default temperature data logger in 2014 was the CT2X, but WQL recordings were used when CT2X was not available.**

	Temperature	Salinity	pH	Water motion
2014	CT2X & WQL*	CT2X	WQL (Except ZZ: No pH in 2014)	HOBO-G Accelerometer

2015	YSI 600XLM (Except IN & LB: HOBO-Temp.)	YSI 600XLM (Except IN & LB: WQL-pH)	YSI 600XLM (Except IN & LB: WQL-pH)	HOBO-G Accelerometer
------	---	---	---	-------------------------

### 2.3.2. Data-logger calibration

#### 2.3.2.1. INW AquiStar<sup>®</sup> CT2X Conductivity Smart Sensors

A two-point (30 mS/cm and 70 mS/cm conductivity solution) calibration was performed immediately prior to deployment in 2014. AquiStar’s technical support advice was that one calibration at the start of the monitoring was sufficient. In hindsight, a monthly calibration during site visits would have been better. In 2015, one-point (50 mS/cm conductivity solution) calibrations were performed during the site visits for the three sites where CT2X loggers were used. Product code “HI 7030” Conductivity Calibration Solution. 12.88 mS/cm @ 25°C. Hanna Instruments Inc.

#### 2.3.2.2. WTW WQL-pH Logger

A two-point calibration (pH 7 and pH 10 buffer solutions) was carried out prior to deployment and then repeated during every site visit.

#### 2.3.2.3. Onset HOBO Pendant<sup>®</sup> G Data Logger

Informal water flow tests were conducted on a number of the accelerometers in controlled flow tanks prior to deployment to ensure that the measurement of acceleration was providing a relative measure of degree of flow and variability in water flow.

#### 2.3.2.4. YSI 600XLM

For the pH probe, a two-point calibration (pH 7 and pH 10 buffer solutions) was carried out prior to deployment and then repeated during every site visit.

For the conductivity probe, a one-point (50 mS/cm conductivity solution) calibration was carried out prior to deployment and then repeated during every site visit.

### 2.3.3. Data-logger losses and repairs

All but one of the accelerometers were lost during the first month of 2014 fieldwork. Timing of site visits and availability of replacements determined the amount of data lost for each site. Two measurement periods of water motion recordings were lost for about half of the sites. One measurement period of recordings was lost for the other sites (predominantly the Cape Breton based sites).

Leaking and malfunctions caused a degree of lost data for the WQL pH loggers in 2014, as seen in Fig. 3.9 (and for the latter part of 2015 for IN and LB, Fig. 3.10).

Biofouling of the CT2X data loggers, as well as some cases where calibration was unreliable also resulted in some missing and unusable salinity data in 2014 (Fig. 3.7 & Fig. 3.8 ).

## **2.5. Abiotic metrics**

Three metrics were devised for each of the four abiotic variables. A minimum (min.), mean, and maximum (max.) metric for each of temperature, salinity, and pH. In the case of water motion, the variance in acceleration is the value used. So the min. variance in acceleration (var. accel.), mean var. accel., and max. var. accel. are the water motion metrics. A value for each metric was calculated for every measurement period at every site, so that it matched the corresponding *C. intestinalis* abundance value for the same measurement period and site.

### *2.5.1. Mean metrics*

The mean metric for temperature, salinity and pH is simply the mean value of all the 5-min interval recordings for a given measurement period. For the mean var. accel. metric, the variance between all of the raw readings of the sum of the acceleration across the X, Y, and Z axes was calculated and then the mean of the variance was used. This is because the absolute value of water motion was measured in acceleration (g) and this value for perfectly still water could differ from site to site depending on the depth or angle of the accelerometer. The variance between each of the raw measurements provides the relative measure of water motion at a site.

### *2.5.2. Maximum and minimum metrics*

The potential for introducing error when calculating min./max. values for each abiotic variable based on a single raw point measure is high. In addition, it is unlikely that a single 5-minute period of a low or high abiotic value would unduly affect *C. intestinalis*. To calculate min/max metrics for each abiotic variable for a given measurement period, daily-mean values were calculated and subsequently used to determine the max/min metrics. Using the min.

daily mean and max. daily mean as the min/max metrics also allowed for the min/max daily variance in acceleration to be calculated for each measurement period. If a single measure was used for min/max metrics, then no variance could be calculated and only the mean variance in acceleration could be used.

## **2.6. Biotic and abiotic data plots**

All biotic and abiotic plots were implemented using R, Version 3.2.2 (R Core Team, 2015), an open source software run through RStudio, Version 0.99.489 – © 2009-2015 RStudio, Inc. (RStudio Team, 2015).

### *2.6.1. Biological data plots*

The cumulative *C. intestinalis* abundances were plotted at settlement plate unit of measurement, by site over time, with the average value for each measurement period plotted with a line, using the lattice package in R (Sarkar, 2008). Each year was plot separately. This method was chosen given the continuous nature of these data. For the monthly *C. intestinalis* data, the mean abundance for each measurement period by site and year is displayed by bar charts using the ggplot2 package in R (Wickham, 2009). This method was chosen as these data are independent measures but still represent a progression though time at each site.

### *2.6.2. Abiotic data plots*

Heat maps of each of the four abiotic variables were created for both years using the lattice package in R (Sarkar, 2008). The daily mean values were used for temperature, salinity and pH, while daily variance was used for water flow. Each year was plotted separately. In each plot, the sites are ordered from top to bottom by highest overall mean to lowest. Alongside each abiotic plot is a heat map plot of the cumulative *C. intestinalis* abundance, so the visual inspection for trends in these data is facilitated.

## **2.7. Statistical analysis**

### *2.7.1. Correlation analysis*

I used Kendall's rank correlation coefficient ( $\tau$ ) to test the strength of the relationships between the mean values of each abiotic variable metric for all the measurement periods and the corresponding mean *C. intestinalis* abundance. This non-parametric correlation was chosen above Spearman's rank correlation coefficient, as while they are perceived to be very

similar, it is suggested that Kendall's tau has some advantages and will be more conservative on average (Newson, 2002). The more conservative approach is favoured based on the following reasoning. For these data, the results of the correlation analyses should be viewed with the large caveat that I recognise it as a very basic form of exploratory analysis to highlight the strongest relationships between abiotic variables and *C. intestinalis* abundance. One obvious example of why it is not entirely appropriate is the likely strong relationship that would occur between any growth and a natural change in an abiotic variable over time, like the increase in temperature over time. Correlation analysis does not adequately account for inter-site or inter-measurement period variation, but rather provides a blanket result for the general association between the two variables.

### 2.7.2. Information-theoretic inference

I used an information-theory approach to analyse these empirical data, as it is suggested to be a sound and appropriate method of inference for observational studies (Burnham and Anderson, 2003). (Burnham and Anderson, 2003) have probably done more to promote and educate the wider community of researchers about information-theory than just about anyone in the last two decades (their classic text having ~ 31,000 citations according to Google Scholar). They go as far as to say that when used to analyse observational studies, virtually all hypothesis-testing approaches, a more classically recognised form of statistical analysis, “have no theoretical justification and may often perform poorly”. This is the reasoning behind my decision to focus on fitting explanatory models to the observed data, accounting for model parsimony and the size and precision of effects, with less focus on testing for classical ‘significance’ (Beninger et al., 2012). This approach does not put any emphasis in the search for a ‘true’ model, but rather focuses on selecting the ‘best’ model based on the information provided by the observed data (Beninger et al., 2012). This seems appropriate, as the concept of a ‘true’ model in such a stochastic environment (marine) appears unlikely. The most prominent measure of models’ attributes used in information theory is called the Akaike Information Criterion (AIC) (Akaike, 1973). The AIC is a method of ranking the likelihood (goodness-of-fit) of a model based on the observed data, and it can be used to evaluate the probability of it being the best fit in comparison to other candidate models. The AIC (Eq. 1) is underlined by the process of maximizing the expected log-likelihood of a given model through maximum likelihood (ML) methods of estimation (Akaike, 1973). The major components that contribute to optimal inference from models fitted via information theoretical approaches are: appropriate model specification in relation to the parameters it

contains; the estimation of these model parameters; and assessing the parameter estimate accuracy (Burnham and Anderson, 2003).

$$AIC = -2 \log(L(model|data)) + 2K \quad (1)$$

L in Eq. 1 is the likelihood of the model, given the data, and K is the number of model parameters.

I used the second order AIC (AICc) for model comparison, Eq. 2, a modified version of the AIC for small sample sizes.

$$AICc = -2 \log(L(model|data)) + 2K + \frac{2K(K+1)}{n-K-1} \quad (2)$$

### *2.7.3. Preselection of abiotic variables*

In order to reduce cases of high collinearity between abiotic variable metrics, as well as making the number of metrics more manageable for model development, a preselection process was used. This involved a step-wise removal of one metric from the pair with the highest correlation, using Pearson's r. This process was repeated until there were only four metrics remaining and using a selection rule that one metric from each variable was left in the final selection of metrics. Pairwise complete observations were used in this correlation analysis, meaning the correlation between each pair of variables is computed using only complete pairs of observations for the given variables (R Core Team, 2015).

### *2.7.4. Mixed-effects modelling: fixed and random effects*

Mixed-effects models include both fixed and random effects and are recognised as a good solution when dealing with repeated-measure analysis, as I was (Laird and Ware, 1982). One major advantage they have over other modeling techniques is that if there are missing values associated with any individual, just the measurement that has missing values can be removed, without discounting the entire set of measurements for that particular individual. This was especially useful for the 2014 modeling, as there were a lot of missing data. They are also particularly adept at accounting for heterogeneous variability at a variety of levels, caused by random effects (Pinheiro and Bates, 2000). Not long ago, the fitting of a model with fixed and random effects, required a two-step process, but now it is possible to fit the two simultaneously, providing a tool capable of explaining intra-individual variability associated with repeated-measurement data (Adame et al., 2008a). Random effects are particularly

useful in a model when the aim is to extrapolate results from a given study on a wider spatial scale, encompassing areas that were not monitored or tested in the study, as the random effect provides a measure of the random variability that is not accounted for by classic fixed parameters. This is of importance for us, as any model developed is aimed at a much greater area than just the eleven sites included in the model.

Working in this area is not without its perils though. There are countless definitions (and indeed many synonyms) for these two effects throughout the literature. Not only that, there are different types of models that include terms with the same name, i.e. a fixed effect, that have different fundamental workings and applications. A good example of this is that a fixed effect in a mixed-effects model is not the same as a fixed effect in a fixed-effects model (sometimes called an unobserved effects model), used commonly in economics (Gunasekara et al., 2014). Also, a random effect in a mixed-effects model can mean a very different thing to a random effect in common analyses from genetic studies (Skrondal and Rabe-Hesketh, 2004). As a result of this scope for confusion, I find it useful to assign appropriate terms to the effects in mixed-effects models, as it allows for a more intuitive understanding of exactly what each of the fixed and random effects are doing.

The fixed effect is often classically referred to as a population effect, i.e. it is the average response (*C. intestinalis* abundance) that can be accounted for by observed variables (the monitored abiotic data) for the entire population (Subedi and Sharma, 2011). Given the potential confusion that could arise when discussing multiple *C. intestinalis* populations represented by the different sites in this design, I am going to refer to this fixed effect as the ‘species-wide effect’. This is to say that this effect is consistent for the study species, *C. intestinalis*, regardless of individual or site. In addition to variation in the response that can be accounted for by the observed abiotic variables, there are unknown sources of variance in *C. intestinalis* abundance that vary from site to site. The random effect (or modelled varying intercepts (Bafumi and Gelman, 2007)) in this case is accounting for the unknown variability that occurs at the different levels of the design, i.e. the site level, in a similar fashion to the plot random effect in Adame et al. (2008). As a result, the random effect is called the ‘site-level’ effect in my models.

#### 2.7.5. *Nonlinear mixed-effects (NLME) model*

I liken this experimental design to an observational epidemiological study in the sense that the main aim is to discover and understand relationships between exposure and outcomes

(Gunasekara et al., 2014). In this case, I was interested in *C. intestinalis* exposure to abiotic oceanographic conditions, with the response outcome being abundance of *C. intestinalis*. A logistic or sigmoidal S-shaped curve is a very common growth trajectory within biology (Wu et al., 2002) and it is the growth curve I decided to model *C. intestinalis* with in this study. A repeated-measure design was used in this observational study, whereby the same settlement plates are repeatedly measured through time. The site-level effects discussed in the previous section can incorporate within-individual variance between the repeated-measurements, accounting for such complexity.

I used a nonlinear mixed-effects (NLME) model fit by maximum likelihood (ML) estimation for all model development (Pinheiro et al., 2015). The model parameter estimates fit by ML are the values that present the most likely model given the observed data (Bolker, 2008). The *likelihood* of any given model is our goodness-of-fit metric, the likelihood being the probability that these data would occur from the model prediction. The overall approach to the growth modelling can be seen in Fig. 2.5, starting from the raw data collection and finishing with model validation and predictions.

All NLME model fitting was implemented in R, Version 3.2.2 (R Core Team, 2015), an open source software run through RStudio, Version 0.99.489 – © 2009-2015 RStudio, Inc. (RStudio Team, 2015).



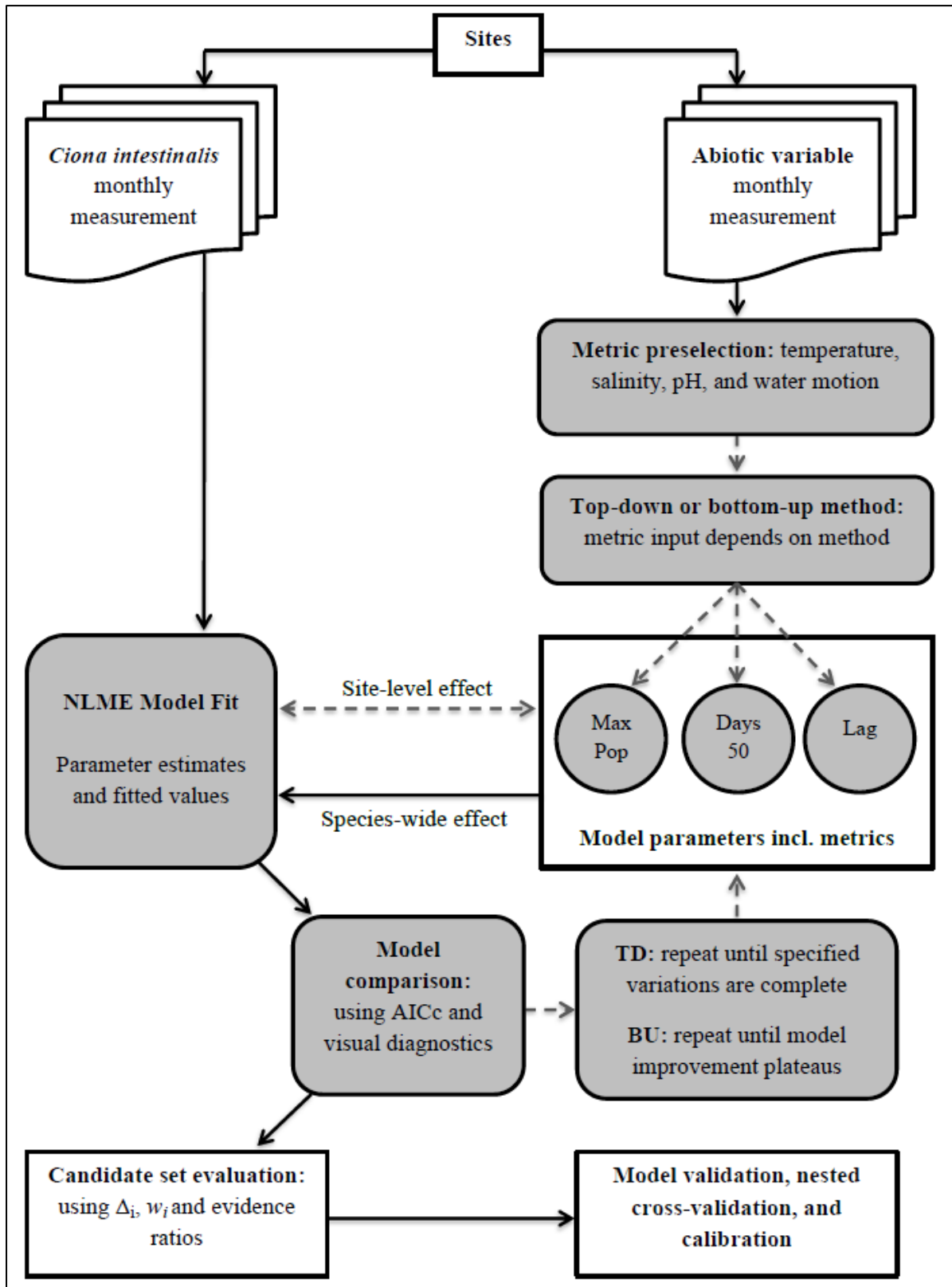


Figure 2.5: A simplified description of the NLME model development process from data collection to model predictions, represented by a flow diagram. Grey dashed arrows and grey-fill boxes indicate stochastic processes that can vary within a given candidate set of models, while black solid arrows and white-fill boxes are invariant processes within the given candidate set of models. TD = top-down and BU = bottom-up.

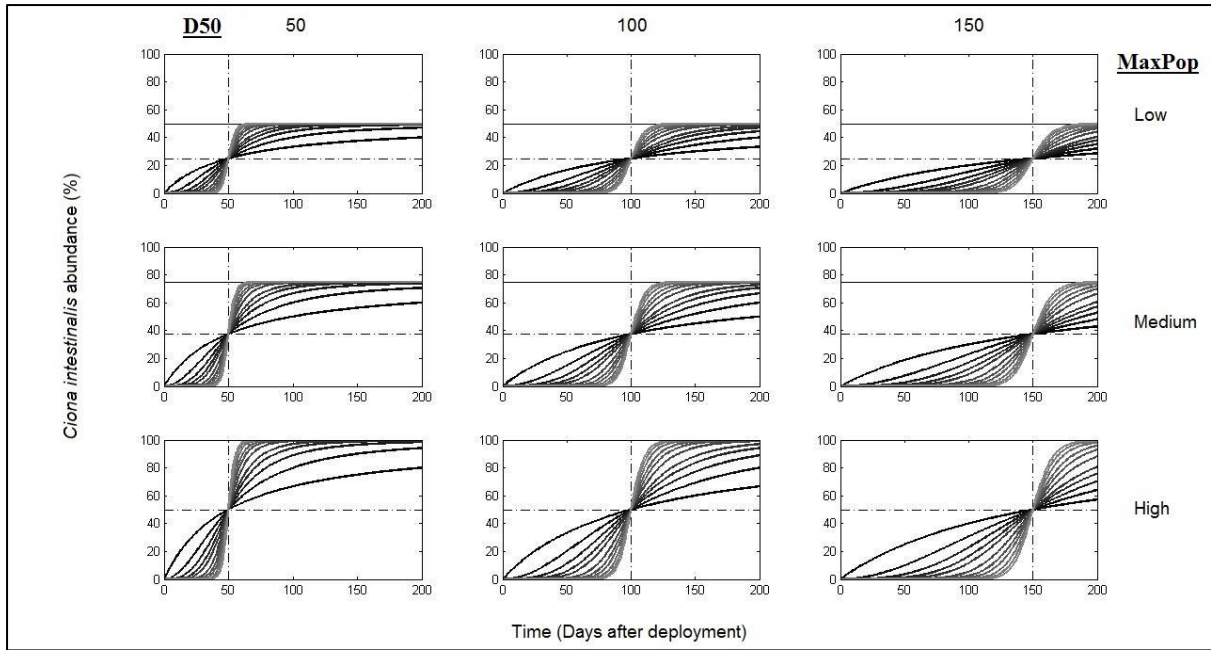
### 2.7.5.1. Model parameters and formulation

The nonlinear growth curve model selected for use is comprised of three model parameters. The Maximum Population (MaxPop) model parameter determines the maximum potential *C. intestinalis* abundance at a given site. This is the asymptote of the growth curve and in the case of this study is bounded by 0% and 100%. The Days50 (D50) model parameter determines the time (in days) that it takes for 50% of the maximum population to be achieved. The Lag Phase (Lag) parameter influences the shape of the growth curve, but not the asymptote (it does alter the time it takes to reach MaxPop) or Days50. It changes the rate of growth following its onset. The three aforementioned parameters are unknown and are estimated in the NLME models by maximum likelihood (ML) estimation. The remaining inputs into the model are all known: the time component; the cumulative *C. intestinalis* abundance; and the abiotic data. The time component in the model is days after deployment. So if deployment occurred at a given site on the 120<sup>th</sup> day of the year (Day N = 120) and the first measurement of *C. intestinalis* occurred at that site on Day N = 150, the days after deployment (DaD) value that is input into the model for that datum point is ‘DaD’ = 30. The abiotic data are the constituents of the three unknown model parameters, MaxPop, Days50, and Lag. Each abiotic metric input into each model parameter technically adds an additional parameter estimate to the model. So if temperature and salinity for example were chosen as input metrics for MaxPop in a model, then that counts as three parameters. The estimate of the intercept for MaxPop, as well as the estimate associated with the temperature and salinity parameters within the overall MaxPop parameter. How each of the model parameters theoretically determines the shape and size of the growth curve is displayed in Fig. 2.6,

The NLME model underlying the process is laid out in detail in Lindstrom and Bates (1990), but the general form is described here in Eq. (3) as

$$y_{ij} = f(\theta_i, x_{ij}) + e_{ij}, \quad (3)$$

where  $y_{ij}$  is *C. intestinalis* percent cover for the  $j$ th measurement period of the  $i$ th plate,  $x_{ij}$  is the abiotic metric vector for the  $j$ th measurement period of the  $i$ th plate,  $f$  is a nonlinear function of the abiotic metric vector and the parameter vector  $\theta_i$  of length  $r$ , and  $e_{ij}$  are the normally distributed residuals for the  $j$ th measurement period of the  $i$ th plate.



**Figure 2.6: Nine example empirical plots that describe how the three model parameters can change the size and shape of the growth curve. D50 has a different value in each of the three columns of plots, while MaxPop has a different value in each of the three rows of plots. Each growth curve line in the individual plots represents a different Lag value.**

The ability of the model to assign a different parameter estimate per site level is through the incorporation of the random effects in  $\theta_i$ , as seen in Eq. (4),

$$\theta_i = A_i\beta + B_ib_i, \quad b_i \sim N(0, \sigma^2D), \quad (4)$$

Where  $\beta$  is a  $p$ -vector of fixed population parameters,  $b_i$  is a  $q$ -vector of random effects associated with individual  $i$ , the matrices  $A_i$  and  $B_i$  are design matrices of size  $r \times p$  and  $r \times q$  for the fixed and random effects, respectively, and  $\sigma^2D$  is a covariance matrix.

A detailed description of how the fixed and random effects matrices can be specified to suit any NLME model can be found in (Lindstrom and Bates, 1990). Instead of getting caught up in the complex formula derivations, I will explain how the actual NLME model used was derived. The predicted values of *C. intestinalis* were calculated using the basic formula described in Eq. (5),

$$C. \textit{intestinalis} \% = \frac{\text{MaxPop} \times \text{DaD}^{\text{Lag}}}{\text{Days50}^{\text{Lag}} + \text{DaD}^{\text{Lag}}} \quad (5)$$

where the estimated (based on the abiotic vector  $x_{ij}$ ) values of MaxPop, Days50, and Lag make up the parameter vector  $\theta_i$  from Eq. (3) and DaD is the known time value. Eq. (6) is assigning a general term, X, to the expanded version of the right hand side of Eq. (5):

$$X = \frac{\exp(\text{MaxPop}) \times \text{DaD}^{\text{Lag}}}{\exp(\text{Days50})^{\text{Lag}} + \text{DaD}^{\text{Lag}}} \quad (6)$$

Eq. (7) is simply denoting *C. intestinalis* as the term Y:

$$Y = C. \textit{intestinalis} \% \quad (7)$$

Eq. (8) is the full equation used in the NLME package in R to fit the model based on the input data of observed *C. intestinalis* and abiotic variables:

$$Y = \text{nlme}(100 \times ((\text{inv.logit}(X) - 0.5) \times 2)) \quad (8)$$

Inv.logit(X) is limited to output values between 0.5 and 1, hence why it is readjusted in the equation to output a value between 0 and 1, so that the total potential range of fitted values is between 0 and 100 after it has been multiplied.

#### 2.7.5.2. Bottom-up model selection

Only the core sites – minus absent were used in the development of the NLME models. Ideally, a limited set of *a priori* models would be devised, only including models that have pure theoretic reasoning behind every inclusion of a given predictor variable in a model parameter. The model fitting I carried out was not exactly according to this *a priori* approach, but instead was a combination of theory and exploration, while also being dependent upon some limitations in achieving model convergence. The abiotic predictor variables were chosen *a priori*, but the specific metrics used to represent each one, as well as which model parameter each metric would be used in was not entirely *a priori*. The bottom-up approach began with the development of a ‘null model’, which was composed of the model parameters not including any abiotic metrics. Once convergence of the ‘null model’ was achieved, the inclusion of site-level effects was then tested. I started with the MaxPop parameter alone as a site-level effect, then proceeding to test D50 alone and Lag alone. Providing model convergence was achievable, I then included multiple combinations of two or three of the model parameters as site-level effects until the best model fit was found, according to AICc values. Once the null model with optimal site-level effects was found, the first abiotic metric was introduced into the parameters. The metrics that would be used in the development of any given candidate set were chosen during the preselection process described in section 2.7.3. The decision of what metric to introduce first was based on the literature and prior knowledge of the system, through visually assessing trends in the biotic and abiotic plots, and through some exploratory analyses (correlation). For an example, in 2015 temperature was

always introduced first, as it appeared to be the most important variable both through visual assessment and existing knowledge. However, in 2014 salinity appeared to have a stronger association with *C. intestinalis* from visual inspection of these data, so it was introduced first. This somewhat exploratory procedure was carried out, all the while conducting model fit comparisons and diagnostic checks (residual plots and observed vs. fitted plots) of each model. This continued until one model was continually coming out as the most likely.

It is worth noting now that the absolute value of AICc is not of importance in and of itself, but the relative values (and differences in particular) of the AICc compared to other models is the useful aspect. Also, AICc can only be used to compare nested models, all of which were fit using the same number of observations in the same dataset. If AICc values were very similar, the more parsimonious model was chosen to continue model development with. Once it was deemed unlikely that a more likely model could be fit to these data, then the entire set of models developed (candidate set) was compared. The first inter-model assessment criterion is the AICc difference,  $\Delta_i$ , where,

$$\Delta_i = \text{AICc}_i - \text{AICc}_{\min}, \quad (10)$$

the  $\Delta_i$  being the difference between a given model's AICc ( $\text{AICc}_i$ ) and the best model's AICc ( $\text{AICc}_{\min}$ ). How to interpret  $\Delta_i$  values are described by (Burnham and Anderson, 2003) as: when  $\Delta_i \sim 0-2$ , there is substantial empirical evidence that supports the model and it can be considered as likely as the best model;  $\Delta_i$  of 4-7 means models have considerably less empirical support than the best model; and  $\Delta_i > 10$  means that the model has essentially no empirical support. While there are some gaps in this system (what does a  $\Delta_i$  of 3-4 mean?), it is a good guiding set of principles and the grey areas can be somewhat context dependent. The next model comparison method used a weighting system of all of the models within the candidate set. This system assigns each model within a candidate set with an Akaike weight ( $w_i$ ) described as,

$$w_i = \frac{\exp(-\frac{1}{2}\Delta_i)}{\sum_{r=1}^R \exp(-\frac{1}{2}\Delta_r)}, \quad (11)$$

where  $w_i$  depends on the entire candidate set and all add up to 1. The  $w_i$  is considered as the weight of evidence in favour of model  $i$  being the actual best model for the situation at hand given that one of the  $R$  models must be the best of the candidate set (Burnham and Anderson, 2003). A model that is clearly the best of a given candidate set of models would have a  $w_i$  of

$> 0.9$  (Anderson et al., 2001). If the best model does not have a large  $w_i$  and the top models are all quite likely, this means that model selection uncertainty is likely to be high (Burnham and Anderson, 2003). The final method of comparison is called the evidence ratio, whereby the likelihood of one model is compared to another invariant of all other models in the candidate set. The evidence ratio is simply calculated as the ratio between the Akaike weights of the two models,  $w_i/w_j$ . The model comparisons for either the top ten models or all the models with a  $\Delta_i < 10$  are presented for each candidate set are presented in the results section.

#### 2.7.5.3. Top-down model selection

The top-down model selection process can be likened to the ‘global model’ in Burnham and Anderson (2003), where the most parameterized model with all variables thought to be important are included. I did however use the preselection process (Section 2.7.2.) before establishing the abiotic variable metrics to be used in the top-down (TD) model building approach. Once the four metrics were determined via preselection, they are all introduced into the model. This led to four metrics in each model parameter (technically, 12 abiotic parameters and 3 model parameters, a total of 15 parameters, but for the purposes of clarity there is a distinction drawn in my thesis between the model parameters [e.g. MaxPop] and the abiotic parameters that are coefficients within the greater model parameters). Assuming convergence was achieved, this top-down model is now referred to as the global model (GM), and all others developed are nested (or reduced) versions of the GM. Running every possible abiotic metric combination in all three parameters was not feasible (~14,000 possible combinations) or good practice (risk of data-dredging). To counter these risks, a handful of model designs were decided beforehand. Each metric was systematically removed from the GM, one parameter at a time. For example, when considering the salinity metric included in the GM, the next model run would be the GM excluding the salinity metric from MaxPop. This would be repeated for each of the other model parameters. Once the three models had been run excluding only the single salinity metric from each parameter at a time, this would be repeated removing the salinity metric from two of the parameters at a time. This was repeated until all three combinations were completed. Finally, the GM was run without the salinity metric in any of the model parameters. This whole process was repeated for each abiotic variable metric. There was a predefined number of model variations to run in the TD approach, so all TD candidate sets have twenty nine models (one GM and then seven variations of metric removal for each abiotic variable metric), as long as convergence could

be achieved for all model variants (it could). The same methods described above in the bottom-up section, 2.7.4.2., were used to compare models within a candidate set.

#### 2.7.5.4. Best temperature-only and salinity-only models

In the interest of model parsimony and to counter the large portions of missing data in 2014, I developed models using either temperature only or salinity only data. These model fitting procedures were performed post hoc after the initial bottom-up and top-down model development. The temperature or salinity metric chosen for development was kept consistent with the metric chosen from the preselection process for the given year. Models were developed and assessed in the same manner as described in section 2.7.5.2. Best salinity-only model (BSO) for 2015 are only presented in the model accuracy assessment, nested cross-validation, and inter annual predictions (section 3.7. onwards).

#### 2.7.5.5. Parameter estimates and effect sizes

Coefficient estimates for the predictor variable parameters, as well as the standard deviations of the site-level effects from the species-wide intercept are presented for the ‘best’ model within each year are presented. Effect sizes are not easily obtained (and can be nonsensical) from multivariate models due to multicollinearity between model parameters, as well as inconsistent scaling and a lack of a common denominator between model parameters (Cade, 2015). The results from the NLME modelling I carried out are best interpreted holistically, assessing the collective output of any given model as opposed to assessing the individual parameters’ regression coefficients (Klein, Pers. Com.). As a purely qualitative guide to assist in the interpretation of parameter estimates, I calculate the absolute scaled effect size (ASES) for each of the abiotic covariate model parameters, which can be used to compare abiotic covariates within a given model parameter, but not between model parameters. The scaling is simply to present the estimated coefficient in a comparable scale to the other abiotic metrics. The estimate for a given abiotic covariate from the best model was multiplied by the observed range for that same abiotic variable within these data, these producing a scaled estimate. The ASES metric is absolute, as it does not describe the direction or rate of change in the response, but rather the absolute range over which a given abiotic covariate can influence the response.

#### 2.7.5.6. Observed vs. fitted plot

The observed *C. intestinalis* data were plotted alongside the best models' full-fit (species-wide effects fit and site-level effects fit combined) values from each of the models to visualize the fitted growth for each site. The species-wide effects fit was also plot in isolation, in a similar fashion to Meng et al. (2008) and Calegario et al. (2005), alongside the full-fit in order to display the fit accounted for by the abiotic variables alone, without site-level effects.

#### 2.7.5.7. Residual plots

Residual plots are only presented for the best model (based on AICc) within each year in the results. The residual plots for all other models can be found in Appendix F. The residual variance was plotted using boxplots and separated at the site-level. This was conducted for the full-model (species-wide effects and site-level effects), as well as for the species-wide effects in isolation (Buqui et al., 2015; Calegario et al., 2005). An area of concern when using random effects in a model is the bias that can be incorporated into the estimate if the random effects are correlated with the covariates/predictor variables. If such correlations are present, they violate model assumptions, introducing bias into predictions throughout the measured range of the covariates, as well as compromising parameter estimates (Bafumi and Gelman, 2007). This issue is easily assessed by visually inspecting the relationship by plotting the random effects with the predictor variables to inspect the homogeneity of the relationship. Solutions, if needed, are readily available to counter correlation between predictors and random effects (Bafumi and Gelman, 2007; Bell and Jones, 2015).

#### 2.7.5.8. Confidence intervals and prediction intervals

The 95 % confidence intervals of the population median and the 95 % prediction intervals were calculated for the best model in both years using a bootstrapping technique using the plyr R package (Wickham, 2011). The confidence intervals provide a measure of how well the population median was determined by the models, with 95 % confidence. The confidence interval median was used, rather than the mean, as the median is less sensitive to outliers and is suitable for the initial growth period when zero growth is common. The prediction intervals determine a range in which fitted values are expected to be found 95 % of the time. A bootstrapping technique can provide better coverage properties for the likely range of the true parameter value than confidence intervals calculated on the observed data alone, particularly for small sample sizes or non-normal distributions (Beninger et al., 2012).

#### 2.7.5.9. Model accuracy



Model accuracy was assessed by comparing the model output (fitted values) with the corresponding observed values, using a range of methods to assess the error rate of the fitted values and the variance in the observed data explained by the model. It is all too common for the measure of significance and explanatory power of a model to be provided by the coefficient of determination ( $R^2$ ) and a  $p$ -value, especially in marine ecology (Boldina and Beninger, 2016). Neither of these provides a measure of whether the model is appropriately specified, or whether assumptions have been met. In nonlinear modeling an  $R^2$  value is even less meaningful, as the value is not constrained between 0 and 1, with random effects contributing explanatory power at multiple levels of a given effect. To carry out appropriate model accuracy assessment, I used four methods recommended from a comprehensive comparison of techniques using an ecosystem model example (Olsen et al., 2016). Olsen et al. (2016) recommend using Spearman's rho as the correlation method for model assessment. A correlation coefficient provides a measure of the tendency of the fitted values and the observed values to vary together. This provides a measure of the similarity in which the two vary, but it does not provide a measure of accuracy (Stow et al., 2003). Two measures of model error are recommended by Olsen et al. (2016). Mean absolute error (MAE) is a particularly useful method, as it provides a measure of error in the same units as the dependent variable for a given model (optimal value = 0). The root mean square error (RMSE) is another well-known measure used to assess the level of error in the fitted values of a model (Meng et al., 2008). While it is similar to the MAE, RMSE is a method that is more sensitive to outliers, especially rare, extreme outliers, as the squaring procedure accentuates these (optimal value = 0). In this case, both MAE and RMSE are presented in units of percent *C. intestinalis*. This provides an intuitive measure of the accuracy of the model for any observer. MAE and RMSE are favourable compared to a measure of error such as the average error. The MAE and RMSE measure the magnitude of error, not the direction, so positive and negative fitted values don't cancel each other out, as would be the case for the average error (Stow et al., 2003). MAE and RMSE were calculated using the hydroGOF R package (Zambrano-Bigiarini, 2014). Modelling efficiency (MEF) is commonly used to assess the predictive capability of a model by measuring the proportion of variability that is explained (ranges from -1 to 1, optimal value = 1) (Amaro et al., 1998; Lehuta et al., 2013; Shimoda and Arhonditsis, 2016). It is similar to an  $R^2$  measure, but more appropriate and commonly used for NLME models (Amaro et al., 1998). MEF was calculated using the JBTools R package (Buttlar, 2015).

#### 2.7.5.10. Nested cross-validation and model prediction

A nested cross-validation technique was used to assess the predictive power of a given model within the observed dataset. The procedure involves separating the observed data into  $n$  folds, training the model on  $n-1$  folds (training dataset) and then using this trained model to predict the observed values in the single fold excluded (test fold) from the training dataset. This process is repeated by successively excluding all  $n$  folds and the results of all test folds were averaged for each model. The variance explained in the fitted values of the test fold was then assessed using MAE, RMSE, MEF, and the Spearman rank correlation coefficient, as in model validation. The nested cross-validation was carried out on the annual datasets in isolation. An additional step to test the predictive power was assessed by using models developed in a single year to predict the observed values in the second year based on the observed abiotic data. The error and predictive power was again measured using MAE, RMSE, MEF, and Spearman's  $\rho$ .

### 3. Results

#### 3.1. Cumulative *C. intestinalis* abundance

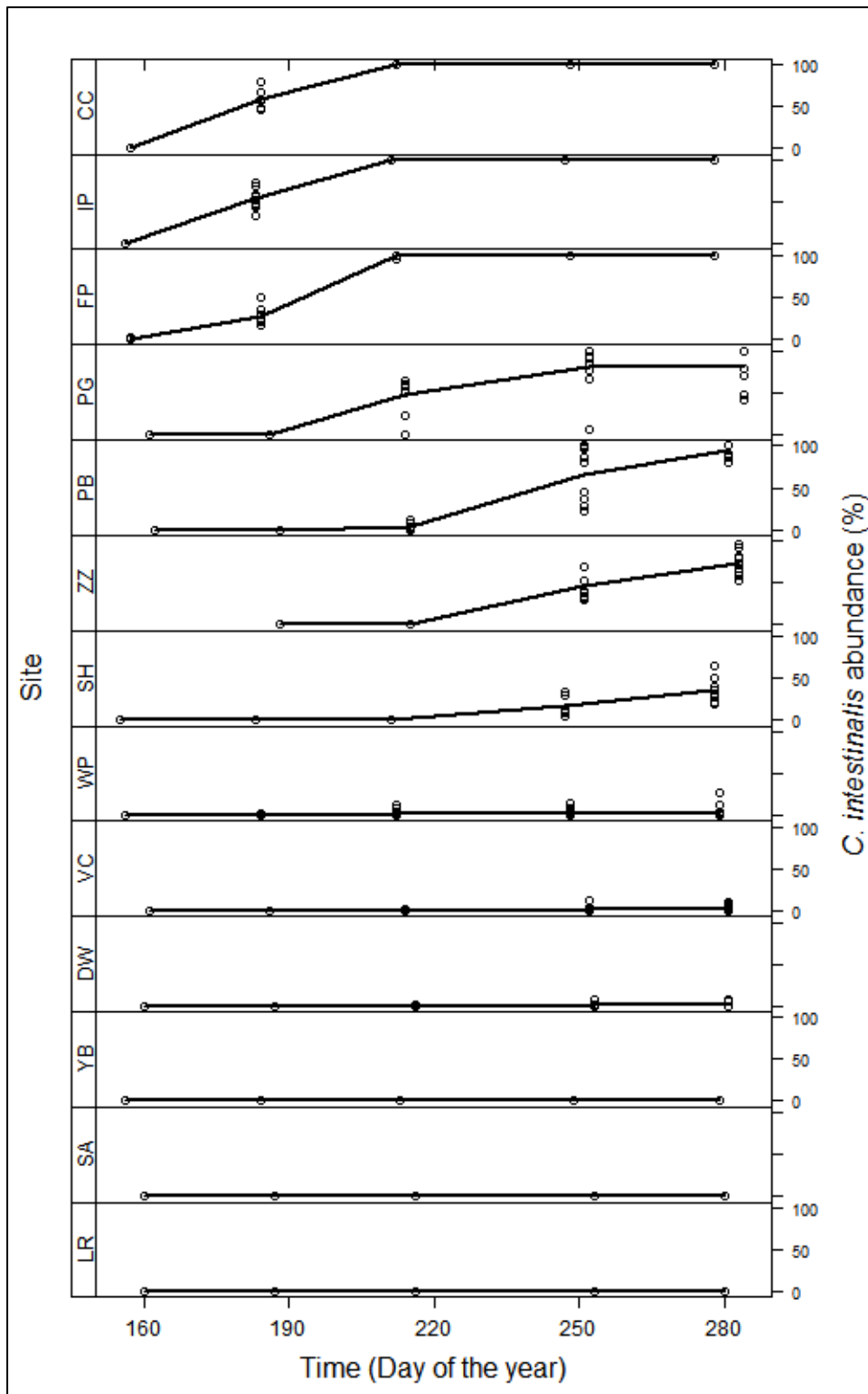


Figure 3.1: *C. intestinalis* growth curves for all 13 sites throughout the 2014 monitoring period. The individual plate abundance is displayed (black circles), with a growth curve (black line) plotted through the mean of each measurement period.

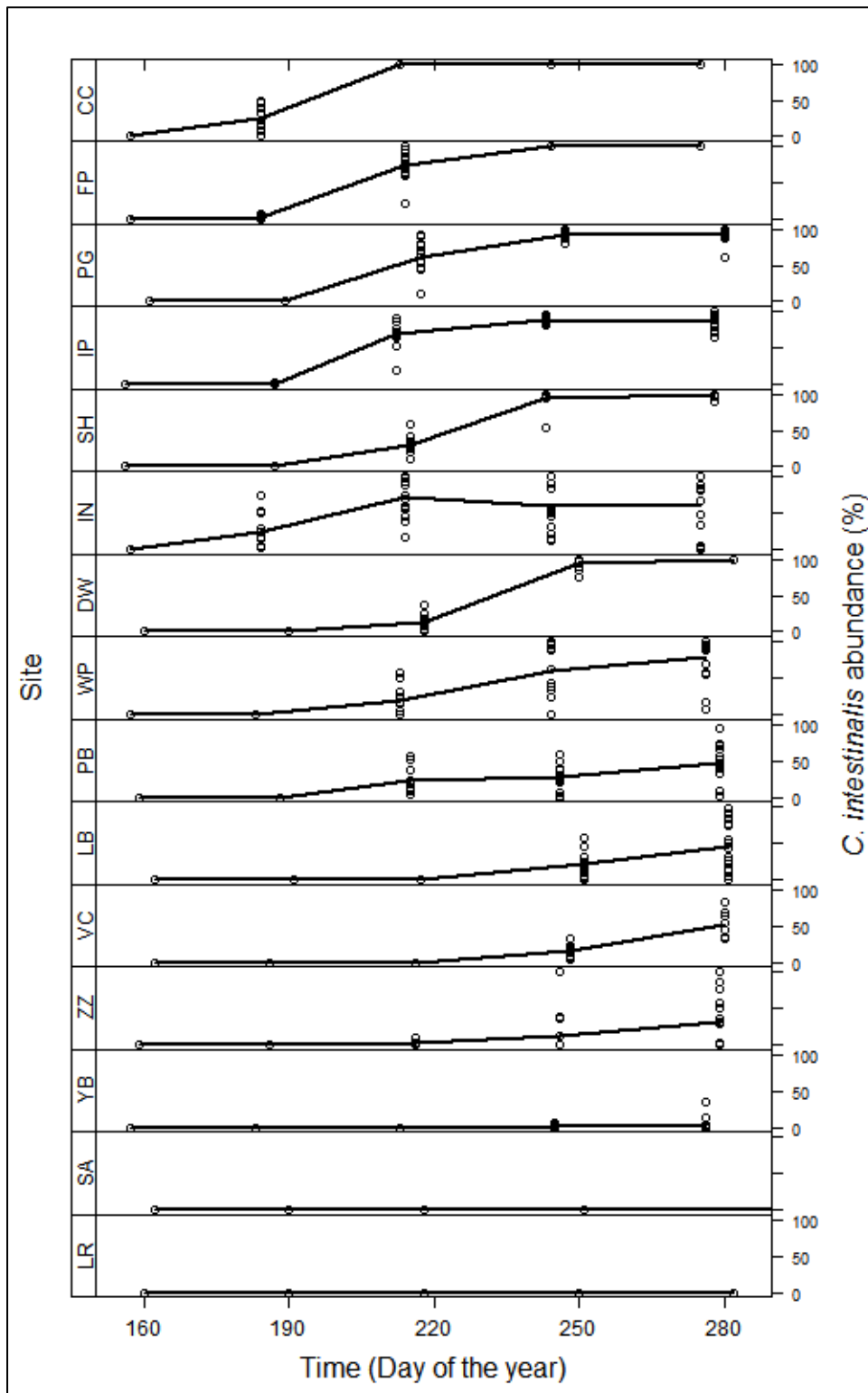
### 3.1.1. Cumulative *C. intestinalis* growth in 2014

No *C. intestinalis* was recorded on any settlement plate during the first measurement period in 2014 (Fig. 3.1). Anecdotal observations suggest settlement occurred at sites that would go on to exhibit high abundance in later measurement periods (pers. obs.), but these individuals were too small to be positively identified and measured. The sites with the highest mean abundance of *C. intestinalis* in 2014 were CC, IP, and FP respectively (Fig. 3.1). The *C. intestinalis* abundance had reached saturation at these three sites by the third measurement period, equating roughly to the first week in August. The growth of *C. intestinalis* at PG lags behind these three sites by about a month, as it is measurable in the third measurement period with growth increasing until plateauing in the fifth measurement period before saturation is reached. PB and ZZ (deployed c. two months later than all other sites) exhibit a further lag in growth in comparison with PG, as no *C. intestinalis* was measured in the third measurement period, except for the odd individual at PB. Increased abundance was recorded in the fourth measurement period and growth continued until the final measurement period at both sites. SH followed a similar pattern to that of PB and ZZ in terms of when the first recruits were present, but the abundance was never as high. This resulted in a relatively stable abundance once *C. intestinalis* was present. *C. intestinalis* was present at WP, VC, and DW in very low abundance. Despite one recruit on a plate at WP in the second measurement period, this did not result in high number of individuals or a high abundance later on in the monitoring. The first *C. intestinalis* presence was measured at both VC and DW during the third measurement period, but again, very little growth occurred in the following measurement periods. No *C. intestinalis* presence was recorded on the settlement plates at any of YB, SA or LR. It should be noted though that three individuals were discovered on the floating dock infrastructure at YB. So despite no recording of *C. intestinalis* on the settlement plates, its presence (potentially at very low abundance) in the water of YB harbour can be confirmed. The same cannot be said of SA and LR where there is no evidence to suggest *C. intestinalis* is present in the waters at these two sites.

### 3.1.2. Cumulative *C. intestinalis* growth in 2015

No *C. intestinalis* presence was recorded on settlement plates at any site during the first measurement period (Fig. 3.2). CC and IN exhibited similar *C. intestinalis* abundance in the second measurement period. CC reached saturation point in measurement period three, while IN displayed a high level of variance, abundance increasing in measurement period three

before decreasing in measurement period four and then stabilizing until the final measurement (Fig. 3.2).



**Figure 3.2:** *C. intestinalis* growth curves for all 15 sites throughout the 2015 monitoring period. The individual plate abundance is displayed (black circles), with a growth curve (black line) plotted through the mean of each measurement period.

*C. intestinalis* abundance began to grow at FP, PG, IP, SH, DW, WP, and PB in the third measurement period (barring the odd individual at FP and IP in the second measurement period). FP reached saturation point in the fourth measurement period, while PG and IP followed similar patterns of growth without ever reaching saturation point. SH, DW, WP, and PB had lower abundance in the third measurement period compared to FP, PG, and IP. SH and DW displayed rapid growth in measurement period four, with saturation occurring in measurement period five at DW, while near saturation occurred at SH. WP and PB exhibited a slower rate of growth than the sites mentioned above, while also displaying a high degree of variability between settlement plates. Neither site reached saturation point. *C. intestinalis* presence was measured at LB, VC, and ZZ in the fourth measurement period (except for the odd individual at ZZ in the third measurement period), with all site exhibiting slow and variable growth until the final measurement period. *C. intestinalis* was measured at YB in the fourth measurement period at very low abundance and remained very low in the final measurement period. No *C. intestinalis* presence was recorded on the settlement plates at any SA or LR.

### 3.1.3. Comparison between cumulative abundance in 2014 and in 2015

CC had the highest mean *C. intestinalis* abundance in both years. Qualitatively it had the highest biomass of any site on settlement plates and on surrounding harbour infrastructure for both years (pers. obv.). FP and IP had markedly lower *C. intestinalis* abundance during the second measurement period in 2015 compared to 2014. Notably, saturation was never reached at IP in 2015, while it took a measurement period longer for FP to reach saturation in 2015. The growth at PG was similar in both years in terms of onset and rate of growth, with a marginally higher abundance in 2015. Two low abundance (WP and DW) sites and one moderate abundance (SH) site in 2014 had markedly higher abundance in 2015. SH settlement plates were practically saturated by the fourth measurement period, which is a stark contrast to 2014. *C. intestinalis* presence at DW was first recorded in the third measurement period, from which there is a steep increase in abundance to near saturation cover in the fourth measurement period (saturation point in the 5<sup>th</sup>). Growth at WP was less pronounced than SH and DW, having an earlier onset and more gradual increase with higher plate variability too. Two moderate to high (PB and ZZ) abundance sites in 2014 had lower abundance in 2015, with greater variability between settlement plates. VC had higher abundance in 2015, but the difference was not as great as that seen at SH, DW and WP. YB had no *C. intestinalis* present on settlement plates in 2014 but had a low abundance in 2015.

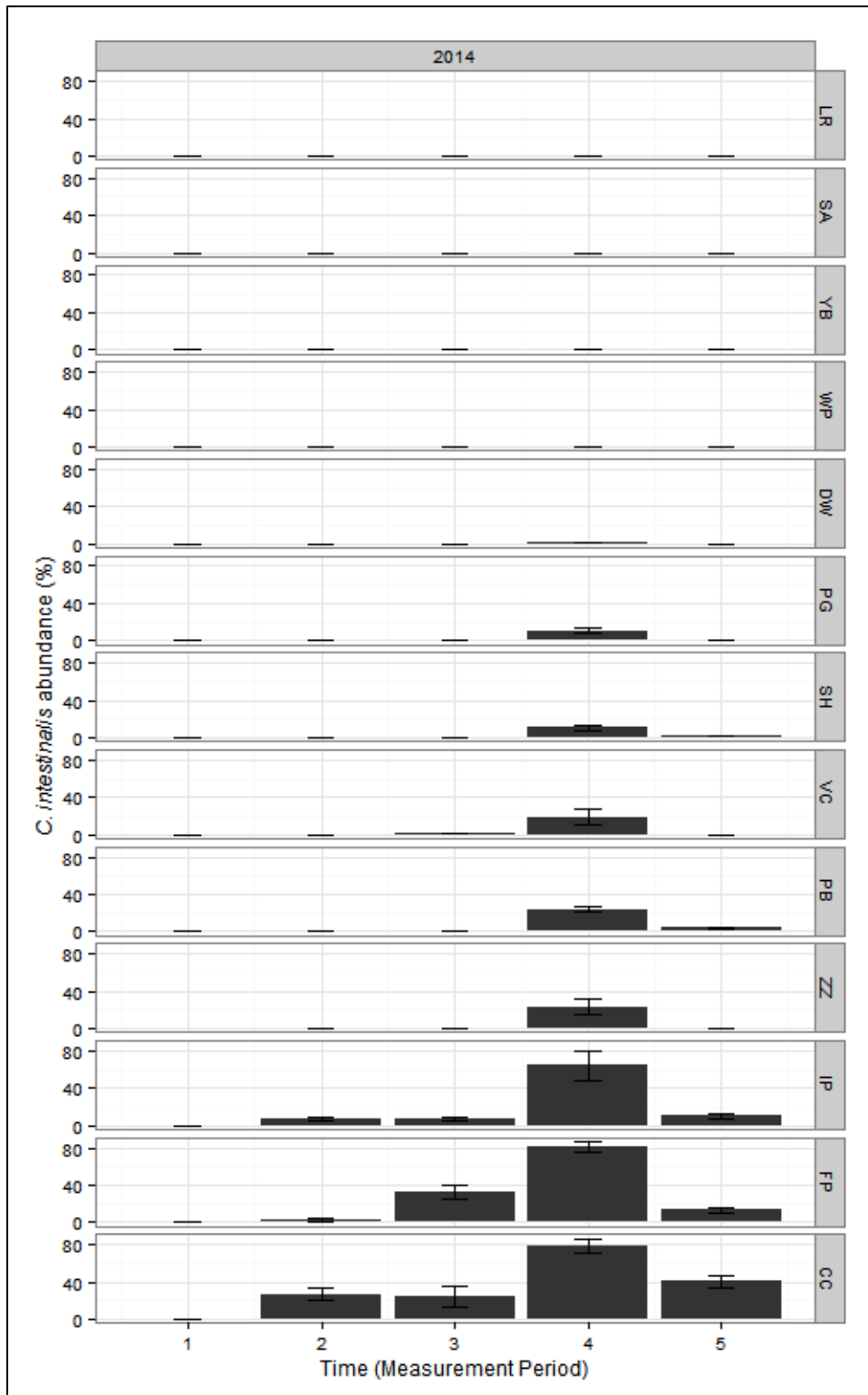
## 3.2. Monthly *C. intestinalis* abundance

### 3.2.1. Monthly *C. intestinalis* growth in 2014

Providing the presence of *C. intestinalis* at a site, in 2014 the abundance was always highest in the fourth measurement period (sometimes the only period with any growth), which equates roughly to the month of August (Fig 3.3). Very few sites displayed any measurable growth in the other measurement periods. LR, SA, and YB had no growth during any of the measurement periods. WP and DW had such low monthly abundance that it appears negligible. PG and SH had low abundance in measurement period four. In measurement period four, VC, PB, and ZZ had roughly double the abundance that PG and SH had. CC, FP, and IP (overall mean abundance decreasing respectively) had high *C. intestinalis* abundance in the fourth measurement period and growth occurred in all but the first measurement period.

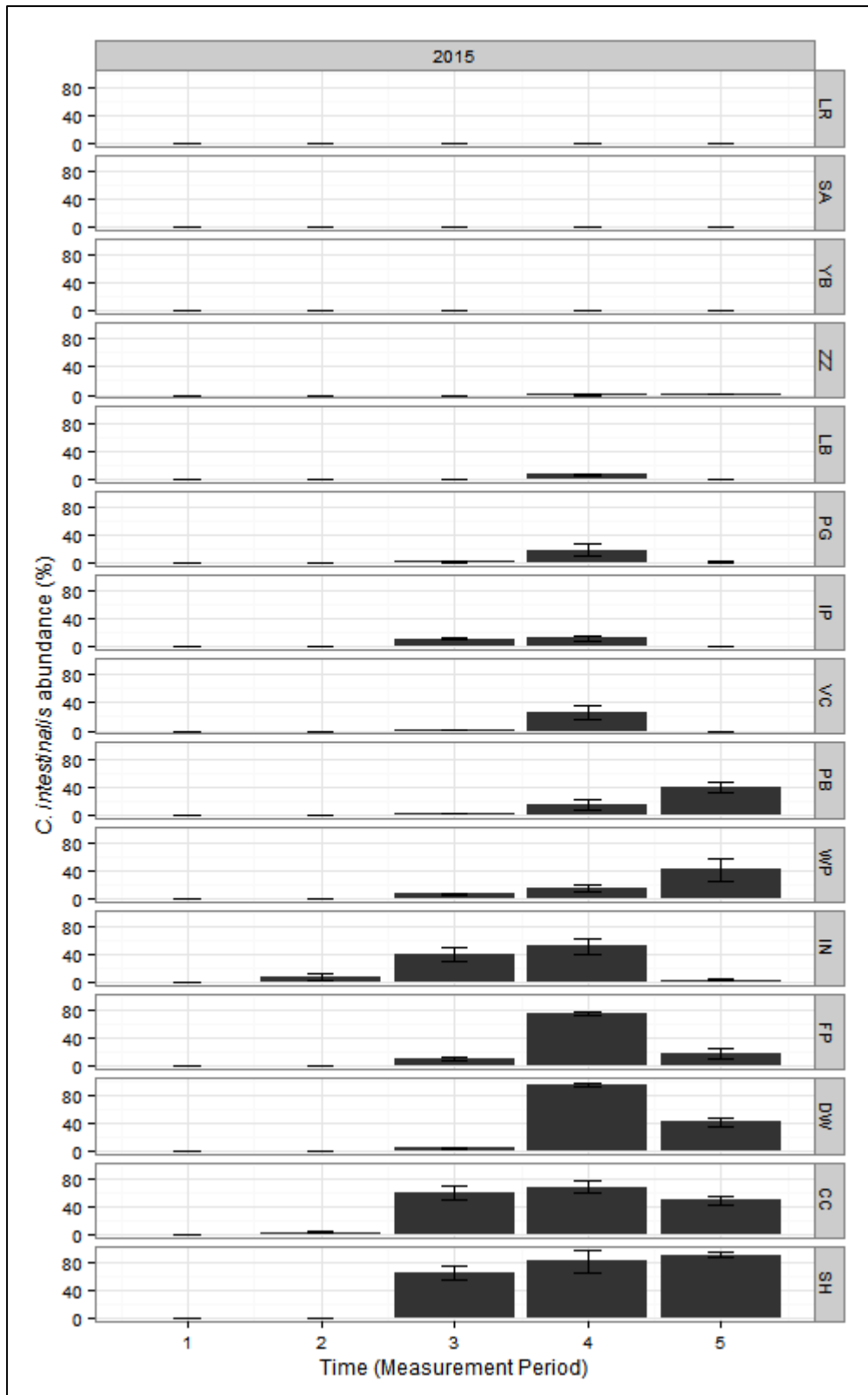
### 3.2.2. Monthly *C. intestinalis* growth in 2015

There is a range of different *C. intestinalis* monthly abundance patterns displayed throughout the thirteen sites in 2015 (Fig. 3.4). Providing the presence of *C. intestinalis* at a site, in 2015 the abundance was always highest in the fourth or fifth measurement, which equates roughly to the months of August and September respectively. LR, SA, and YB had no *C. intestinalis* growth on any of the monthly settlement plates, while ZZ had very low growth during the fourth and fifth measurement periods. LB, PG, IP, and VC all had low growth in the fourth measurement period, with even lower growth in either the third or fifth period. PB and WP had moderate *C. intestinalis* abundance measurement period five (their highest abundance period), with low abundances in the third and fourth measurement periods. IN had *C. intestinalis* present in low abundance in measurement period two, with each period progressively increasing to a moderate abundance in the fourth period, before dropping to near zero in the final period. FP, DW, and CC all had high abundance in the fourth measurement period, with CC the only one of the three to also exhibit quite high abundances in the third and fifth periods as well. SH had very similar abundance levels to CC in both the third and fourth periods, while the fifth period had the highest.



**Figure 3.3: The monthly *C. intestinalis* abundance for each measurement period in 2014, recorded independently of all other measurement periods on new settlement plates deployed at each site every visit. Sites ordered with the highest overall-mean *C. intestinalis* abundance at the bottom.**





**Figure 3.4: The monthly *C. intestinalis* abundance for each measurement period in 2015, recorded independently of all other measurement periods on new settlement plates deployed at each site every visit. Sites ordered with the highest overall-mean *C. intestinalis* abundance at the bottom.**

### 3.2.3. Comparison between monthly abundance in 2014 and in 2015

Considerably different growth patterns were recorded between the two years for a number of sites (Fig. 3.3 and Fig. 3.4). The most noticeable pattern change was that in 2015, peak abundance for each site was not always in the fourth measurement period. The greatest shift in growth saw SH go from a low abundance in 2014 to the highest overall abundance in 2015. Another dramatic increase in abundance between the years was displayed at DW, where the abundance in the fourth and fifth measurement periods in 2015 was a lot higher. PB and WP also had higher abundance in 2015, but not the same degree of increase as seen at SH and DW. The biggest shift in abundance in the opposite direction was seen at IP, where the abundance in the fourth measurement period went from high in 2014 to low in 2015. ZZ 2015 abundance also saw a decrease compared to 2014, but it was a relatively small change compared to IP. Abundances at PG, VC, FP, and CC were relatively similar between the years, apart from a few differences between measurement periods. LR, SA, and YB had no *C. intestinalis* growth on any of the monthly settlement plates in either year.

## 3.3. Abiotic variables

### 3.3.1. Temperature

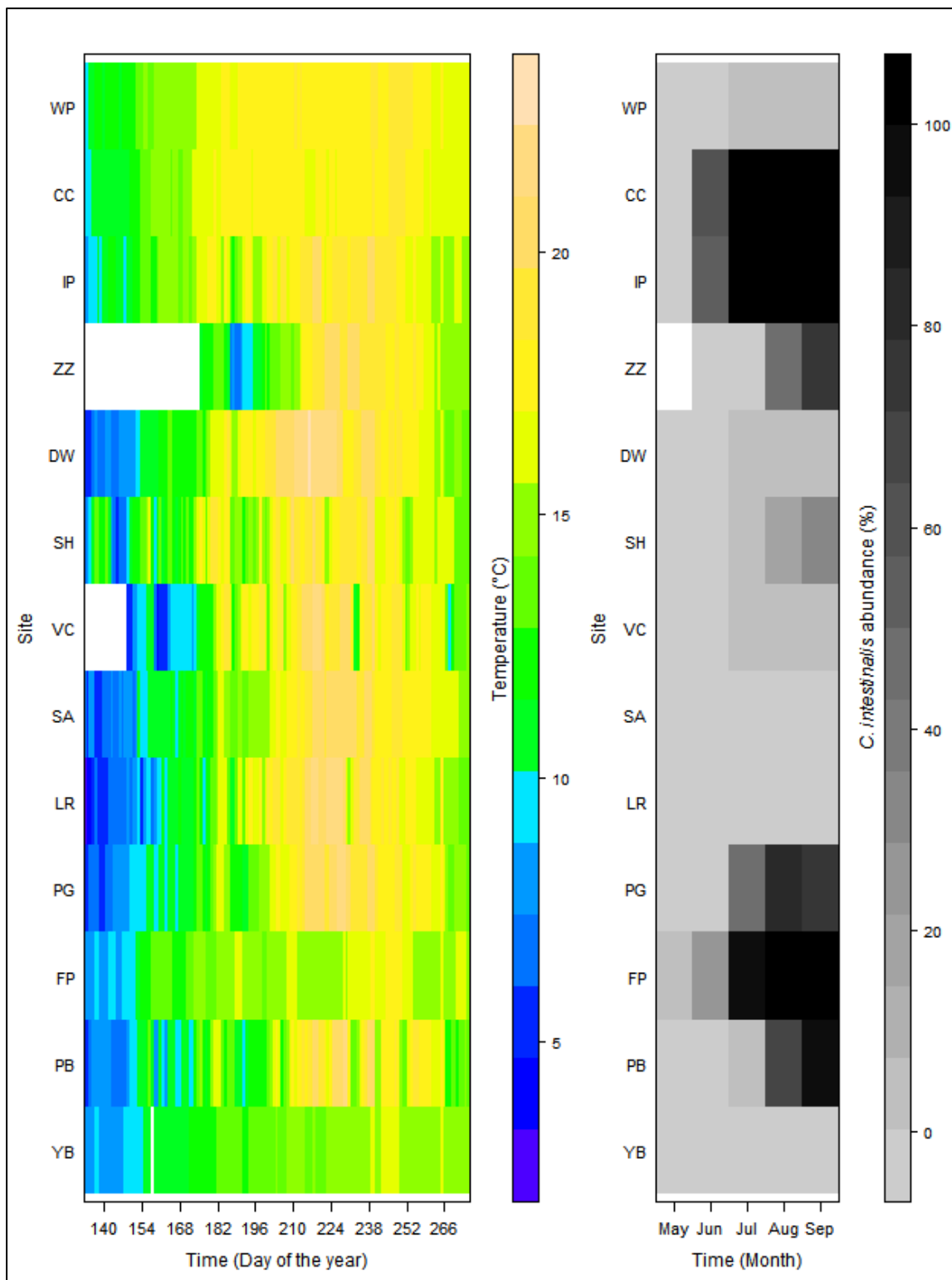
In cases where temperature was recorded in both years, the minimum daily mean over the monitoring period was always lower in 2015, while the maximum daily mean over the same period was always higher in 2015 (Table 3.1). In some cases these differences were slight, like at PG (~ 1°C in both cases), while the differences were noticeable in others, IP for example (~ 4°C in both cases). On average, 2015 was a colder year, with the exceptions of IP and PB (Table 3.1). In 2014, the delay in monitoring commencement at VC and ZZ lead to the absence of data for the coldest portion of the monitoring period. This resulted in them having a higher overall mean temperature than what is truly representative to compare to the other sites. This can be seen in Fig. 3.5 where they appear quite high in the rank of hottest to coldest sites. A more representative rank for these sites is displayed in Fig. 3.6. When monitored during the same period of time, of the core sites, only YB is colder on average in 2015 than the pair.

A clear trend is visible in 2015, where the sites with a higher mean temperature also have higher *C. intestinalis* abundance (Fig. 3.6). This was not as clear in 2014, although the artificial elevation of ZZ and VC may have confounded this slightly (Fig. 3.5). The main

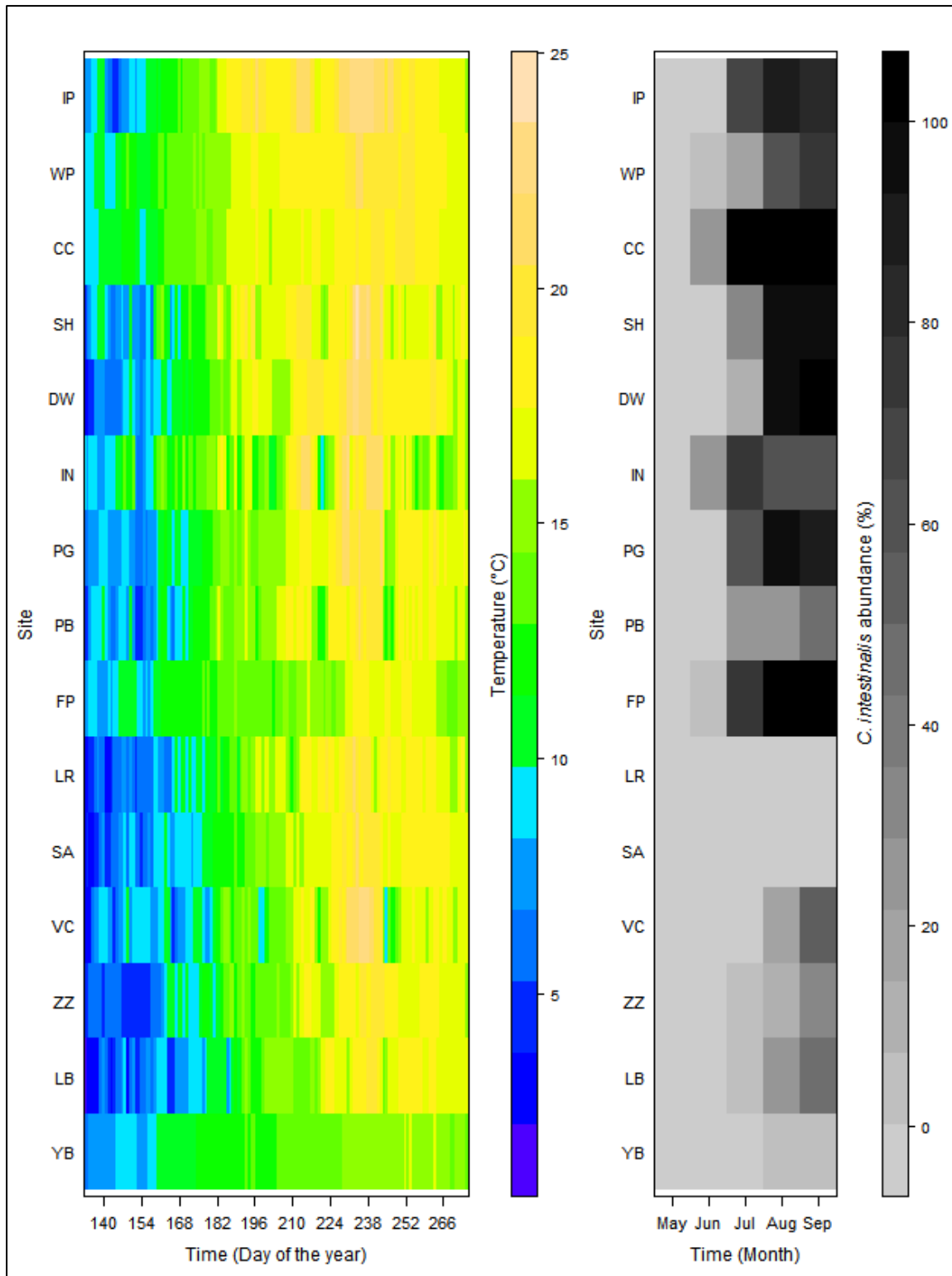
differences between the years that resulted in a clear temperature trend in 2015 is the marked increase in *C. intestinalis* abundance at WP, DW, and SH. In addition, SA and LR were ~ 1°C colder on average in 2015, while PB was slightly warmer and FP was only ~ 0.5°C colder on average in 2015 (Table 3.1), which shifted the order based on the mean (Fig. 3.6).

**Table 3.1: Summary temperature values for all 15 sites (in alphabetical order) and both years (where applicable). The 2014 values, where available are shaded in light grey, while IN and LB are shaded in dark grey. Mean = Overall mean of daily mean values. Minimum = Minimum daily mean value. Maximum = Maximum daily mean value. Variance = Variance of daily mean values.**

Site	Year	Minimum	Mean	Maximum
CC	2014	9.85	15.99	19.09
CC	2015	8.03	15.62	20.22
DW	2014	5.02	15.53	22.42
DW	2015	2.24	14.74	23.67
FP	2014	7.74	14.07	17.79
FP	2015	7.04	13.65	19.28
IN	2015	5.74	14.46	23.84
IP	2014	8.05	15.82	20.74
IP	2015	3.91	16.13	24.14
LB	2015	1.43	12.47	21.21
LR	2014	3.28	14.48	21.91
LR	2015	2.34	13.53	22.65
PB	2014	4.94	13.81	21.24
PB	2015	2.94	13.94	21.92
PG	2014	5.01	14.47	21.29
PG	2015	4.14	14.19	22.31
SA	2014	5.28	14.67	20.94
SA	2015	1.48	13.46	21.33
SH	2014	5.62	15.14	20.52
SH	2015	3.40	15.07	25.11
VC	2014	4.07	15.10	21.57
VC	2015	1.69	13.29	23.58
WP	2014	9.86	16.29	19.05
WP	2015	7.82	16.04	22.16
YB	2014	7.48	13.05	16.02
YB	2015	5.77	12.45	17.92
ZZ	2014	7.05	15.61	20.36
ZZ	2015	3.50	13.13	20.73



**Figure 3.5:** Left - The temperature profiles of the core sites monitored in 2014 are presented by a heatmap of the daily mean values. The sites are ordered by the overall mean of daily means, with the highest overall mean at the top. \* The white sections at the start of ZZ and VC plots are missing data points. Right – The mean *C. intestinalis* abundance for each measurement period in 2014 represented as a heatmap, with sites corresponding to the appropriate site from the temperature plot on the left. \* The white section at the start of ZZ is a missing datum point.



**Figure 3.6:** Left - The temperature profiles of the expanded sites monitored in 2015 are presented by a heatmap of the daily mean temperature values. The sites are ordered by their overall mean temperature, with the highest overall mean at the top. Right – The mean *C. intestinalis* abundance for each measurement period in 2015 represented as a heatmap, with sites corresponding to the appropriate site from the temperature plot on the left.

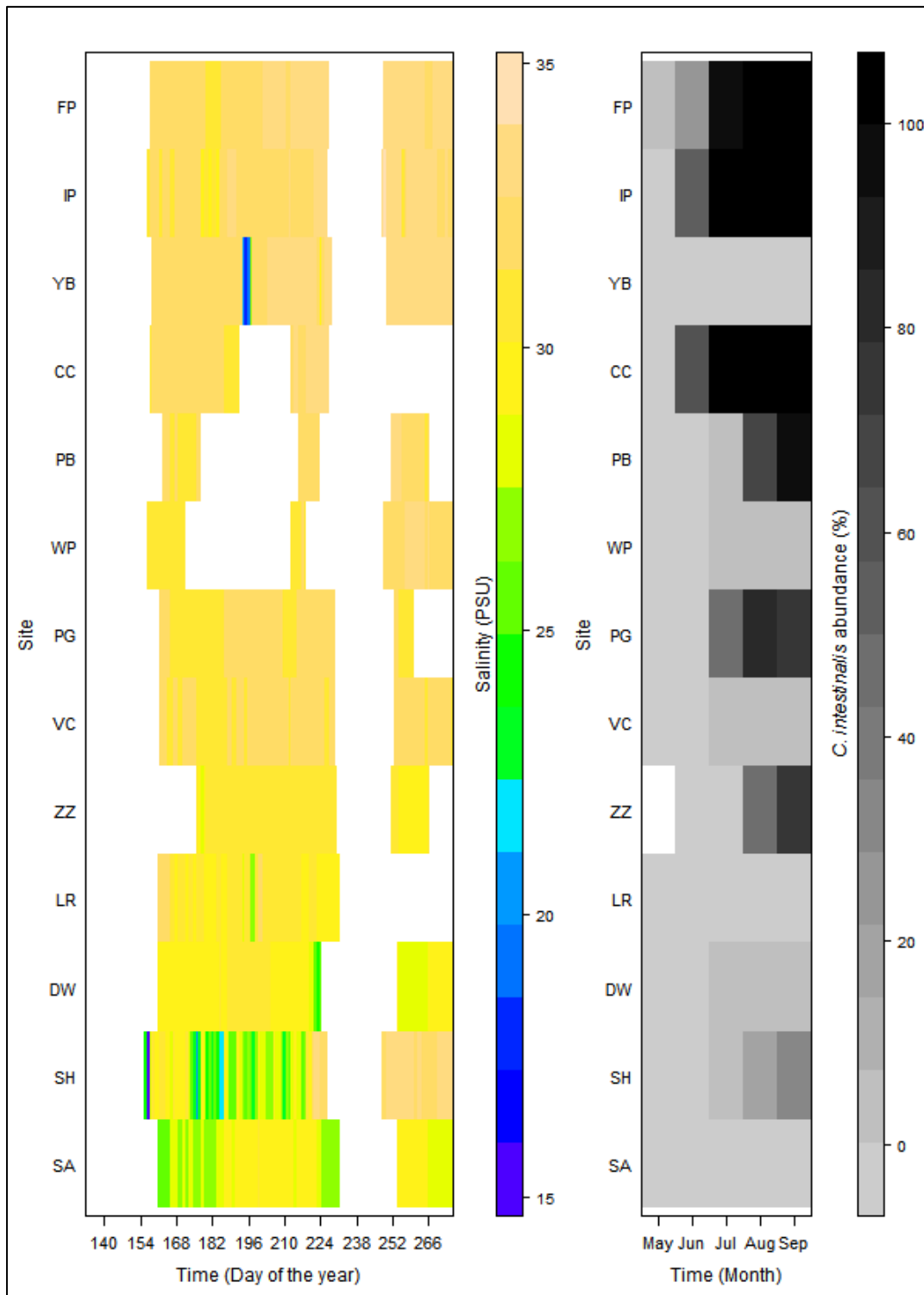
### 3.3.2. Salinity

The overall mean salinity and maximum daily-mean salinity across all sites was higher in 2014, while the minimum daily-mean salinity was higher at all sites in 2014, with the exception of YB (Table 3.2). This trend could be a result of a number of contributing factors: naturally occurring inter-annual variation could, as a result of differences in freshwater input, temperature, and oceanic currents; the large quantities of missing data leading to a loss of vital information; or differences in equipment. If the first is true, then this is a real effect in the variation in salinity. If either the second or third are true, it at least appears as if it would be a systematic error, which could result in the relative trend in 2014 still being valuable information, but it complicates comparisons between the two years. There is also a chance that the missing data or equipment differences would lead to error that is not systematic, but instead varies within 2014. I think this is less likely, but still important to consider.

Comparing the ranked sites according to overall mean salinity, there appears to be a pattern in 2014, where the sites with higher overall mean salinity also had higher mean *C. intestinalis* abundances (Fig. 3.7). The three highest abundance sites are all in the top four sites ranked by overall mean salinity. This same pattern is not evident in 2015, where high abundance sites are scattered throughout the ranking by overall mean salinity (Fig. 3.8). While there are some inter-annual differences between the site rankings according to overall mean salinity, they do not appear to be dramatic. Three of the four lowest mean salinity sites in 2014 were again the three lowest (with the exception of the new site LB) in 2015 (Fig. 3.7 & Fig. 3.8). WP is the exception, as it moved up the site rankings according to overall mean salinity. Again, three of the top four mean salinity sites in 2014 were the same in 2015, the exception being IP.

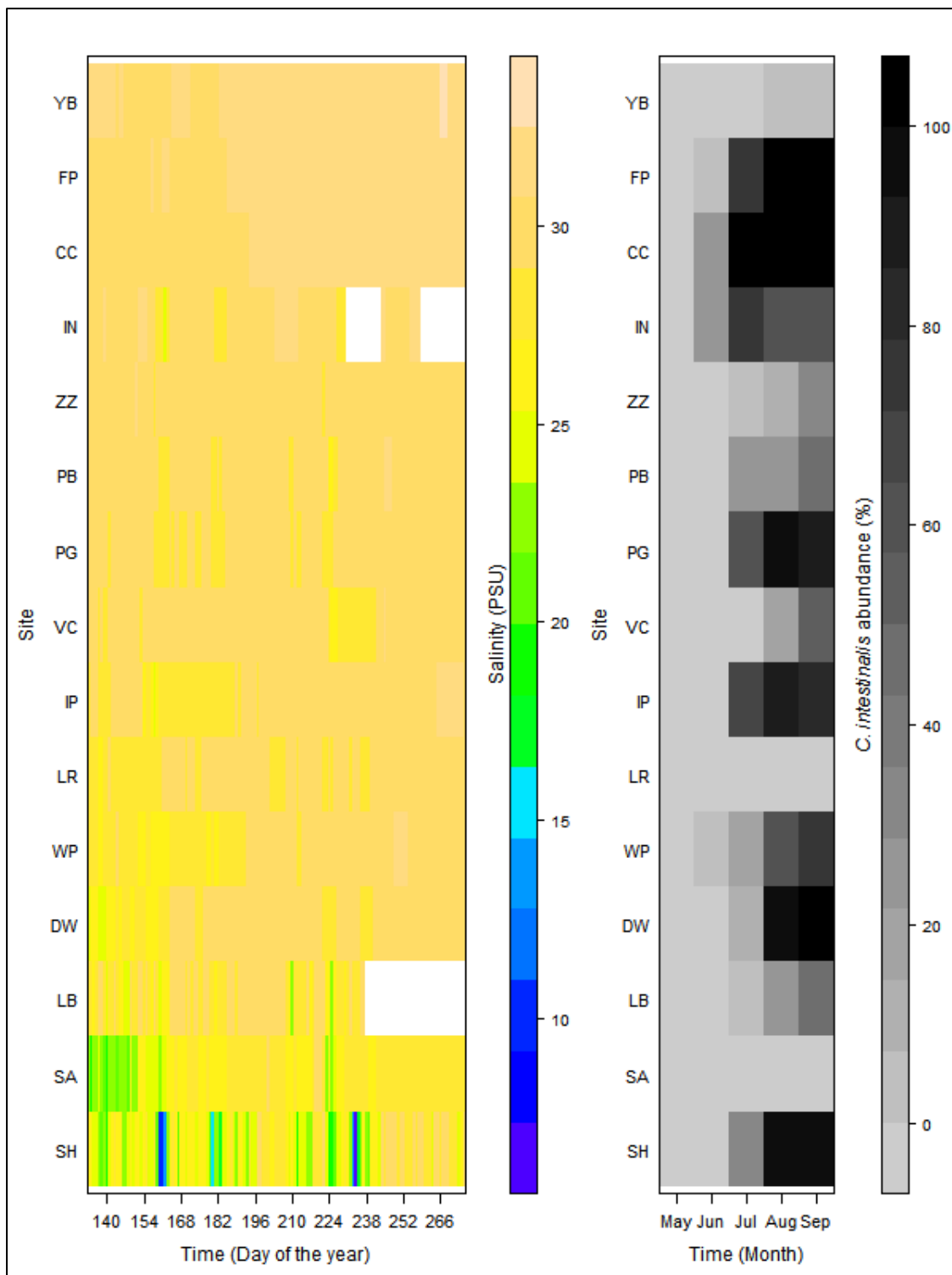
**Table 3.2: Summary salinity values for all 15 sites (in alphabetical order) and both years (where applicable). The 2014 values, where available are shaded in light grey, while IN and LB are shaded in dark grey. Mean = Overall mean of daily mean values. Minimum = Minimum daily mean value. Maximum = Maximum daily mean value. Variance = Variance of daily mean values.**

<b>Site</b>	<b>Year</b>	<b>Minimum</b>	<b>Mean</b>	<b>Maximum</b>
CC	2014	30.48	31.93	33.36
CC	2015	28.91	30.71	32.24
DW	2014	24.90	29.32	30.46
DW	2015	20.98	28.67	30.37
FP	2014	30.20	32.50	33.65
FP	2015	29.62	31.11	32.31
IN	2015	23.00	30.02	32.17
IP	2014	29.71	32.23	33.94
IP	2015	22.71	29.36	31.04
LB	2015	16.84	28.39	30.84
LR	2014	26.31	30.38	32.01
LR	2015	23.70	29.07	30.97
PB	2014	30.70	31.90	33.06
PB	2015	25.50	29.85	31.36
PG	2014	30.68	31.55	32.30
PG	2015	25.82	29.60	31.33
SA	2014	25.54	28.48	30.07
SA	2015	10.96	26.70	29.64
SH	2014	15.92	29.00	33.53
SH	2015	3.83	25.29	30.83
VC	2014	30.12	31.50	32.42
VC	2015	24.36	29.50	31.26
WP	2014	30.24	31.64	33.01
WP	2015	23.07	29.04	31.45
YB	2014	17.94	32.15	33.79
YB	2015	26.02	31.33	32.73
ZZ	2014	27.79	30.40	31.17
ZZ	2015	21.65	29.90	30.90



**Figure 3.7:** Left - The salinity profiles of the core sites monitored in 2014 are presented by a heatmap of the daily mean values. The sites are ordered by their overall mean salinity, with the highest overall mean at the top. \* The white sections present for all site profiles are missing data points. Right – The mean *C. intestinalis* abundance for each measurement period in 2014 represented as a heatmap, with sites corresponding to the appropriate site from the salinity plot on the left. \* The white section at the start of ZZ is a missing datum point.





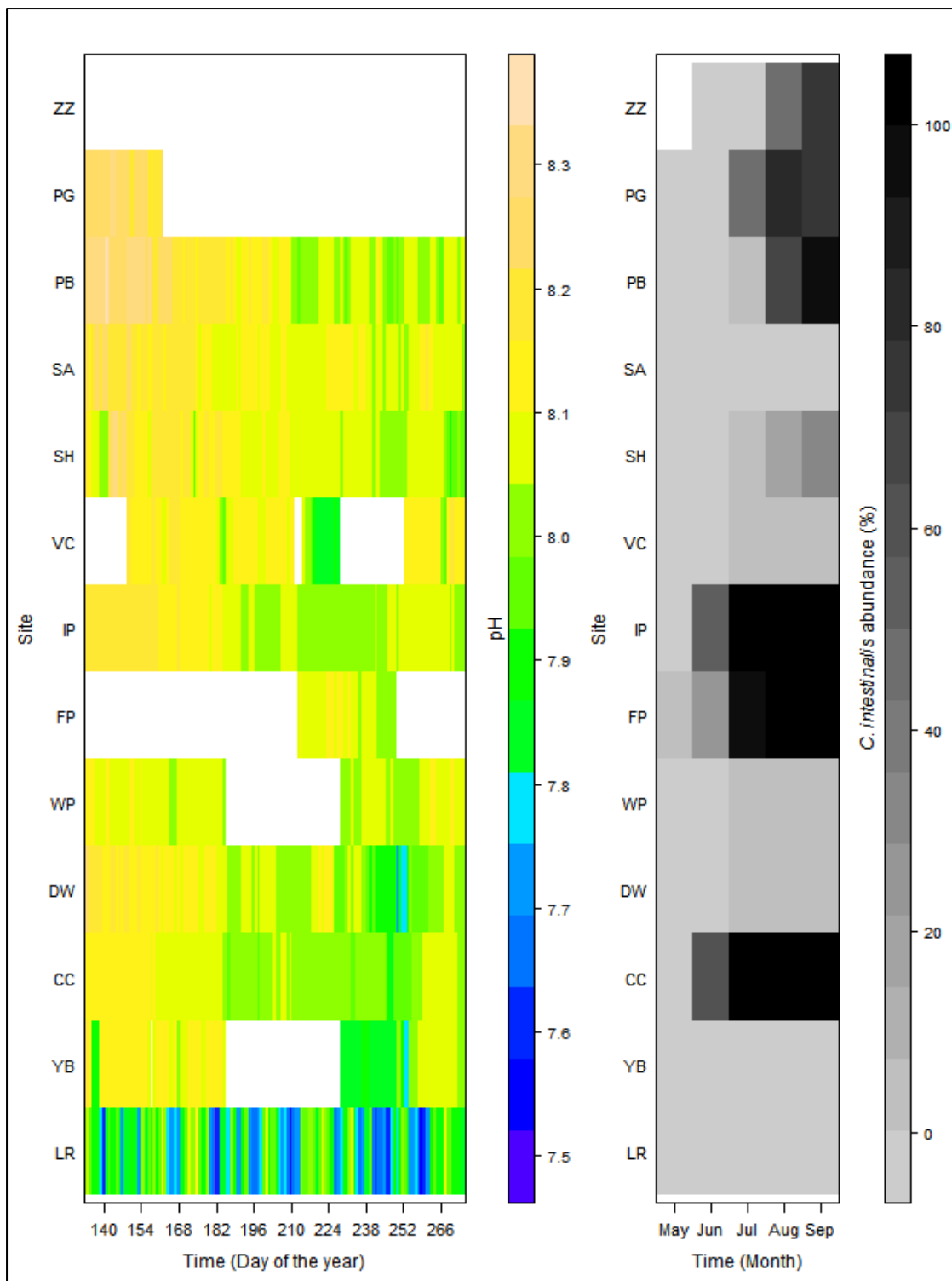
**Figure 3.8: Left - The salinity profiles of the expanded sites monitored in 2015 are presented by a heatmap of the daily mean salinity values. The sites are ordered by their overall mean salinity, with the highest overall mean at the top. \* The white sections present in the LB and IN profiles are missing data points. Right – The mean *C. intestinalis* abundance for each measurement period in 2015 represented as a heatmap, with sites corresponding to the appropriate site from the salinity plot on the left.**

### 3.3.3. pH

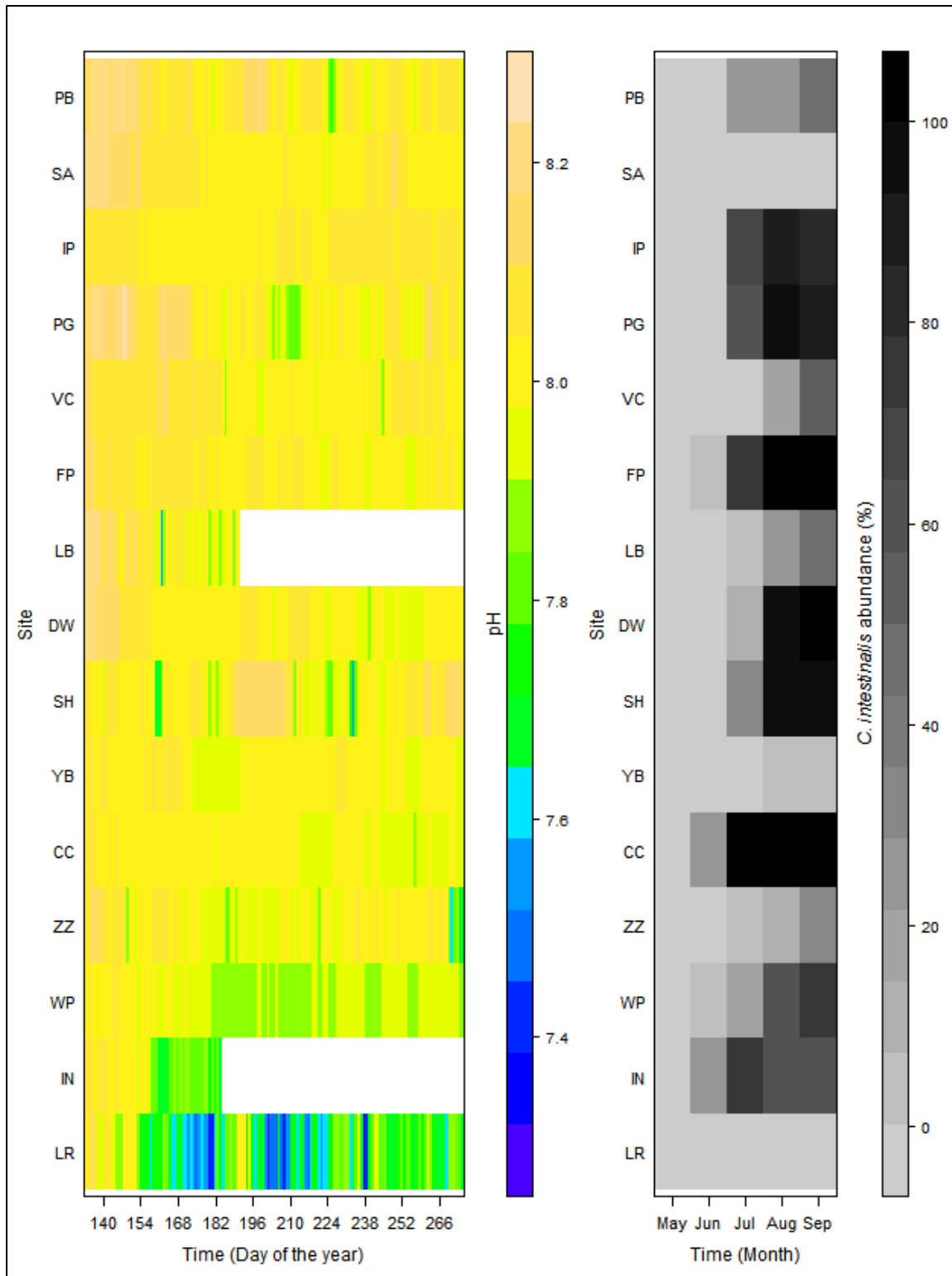
The overall mean pH was lower in 2015 across all sites (Table 3.3). The maximum and minimum daily mean values do not follow as consistent a pattern. The minimum was higher in 2015 at CC, DW, and YB, by relatively small margins (Table 3.3). The maximum was higher in 2015 at DW, FP, LR, PB, and SA, mostly by relatively slight differences again. No notable patterns are evident in either year between the overall mean pH and the *C. intestinalis* mean abundance (Fig. 3.9 & Fig. 3.10).

**Table 3.3: Summary pH values for all 15 sites (in alphabetical order) and both years (where applicable). The 2014 values, where available are shaded in light grey, while IN and LB are shaded in dark grey. Mean = Overall mean of daily mean values. Minimum = Minimum daily mean value. Maximum = Maximum daily mean value. Variance = Variance of daily mean values.**

Site	Year	Minimum	Mean	Maximum
CC	2014	7.84	8.04	8.14
CC	2015	7.88	7.99	8.11
DW	2014	7.73	8.05	8.22
DW	2015	7.81	8.02	8.25
FP	2014	7.99	8.07	8.11
FP	2015	7.90	8.03	8.18
IN	2015	7.51	7.91	8.18
IP	2014	8.01	8.08	8.20
IP	2015	7.95	8.04	8.15
LB	2015	7.44	8.03	8.55
LR	2014	7.52	7.85	8.13
LR	2015	7.21	7.75	8.39
PB	2014	7.96	8.12	8.33
PB	2015	7.70	8.06	8.38
PG	2014	8.16	8.23	8.29
PG	2015	7.69	8.04	8.28
SA	2014	8.01	8.12	8.26
SA	2015	7.79	8.05	8.28
SH	2014	7.92	8.09	8.30
SH	2015	7.24	8.02	8.28
VC	2014	7.82	8.09	8.18
VC	2015	7.67	8.04	8.17
WP	2014	7.99	8.07	8.13
WP	2015	7.57	7.94	8.10
YB	2014	7.79	8.03	8.17
YB	2015	7.86	8.00	8.16
ZZ	2014	NA	NA	NA
ZZ	2015	6.97	7.99	8.30



**Figure 3.9:** Left - The pH profiles of the core sites monitored in 2014 are presented by a heatmap of the daily mean pH values. The sites are ordered by their overall mean pH, with the highest overall mean at the top. \* The white sections present for multiple site profiles are missing data points. Right – The mean *C. intestinalis* abundance for each measurement period in 2014 represented as a heatmap, with sites corresponding to the appropriate site from the pH plot on the left. \* The white section at the start of ZZ is a missing datum point.



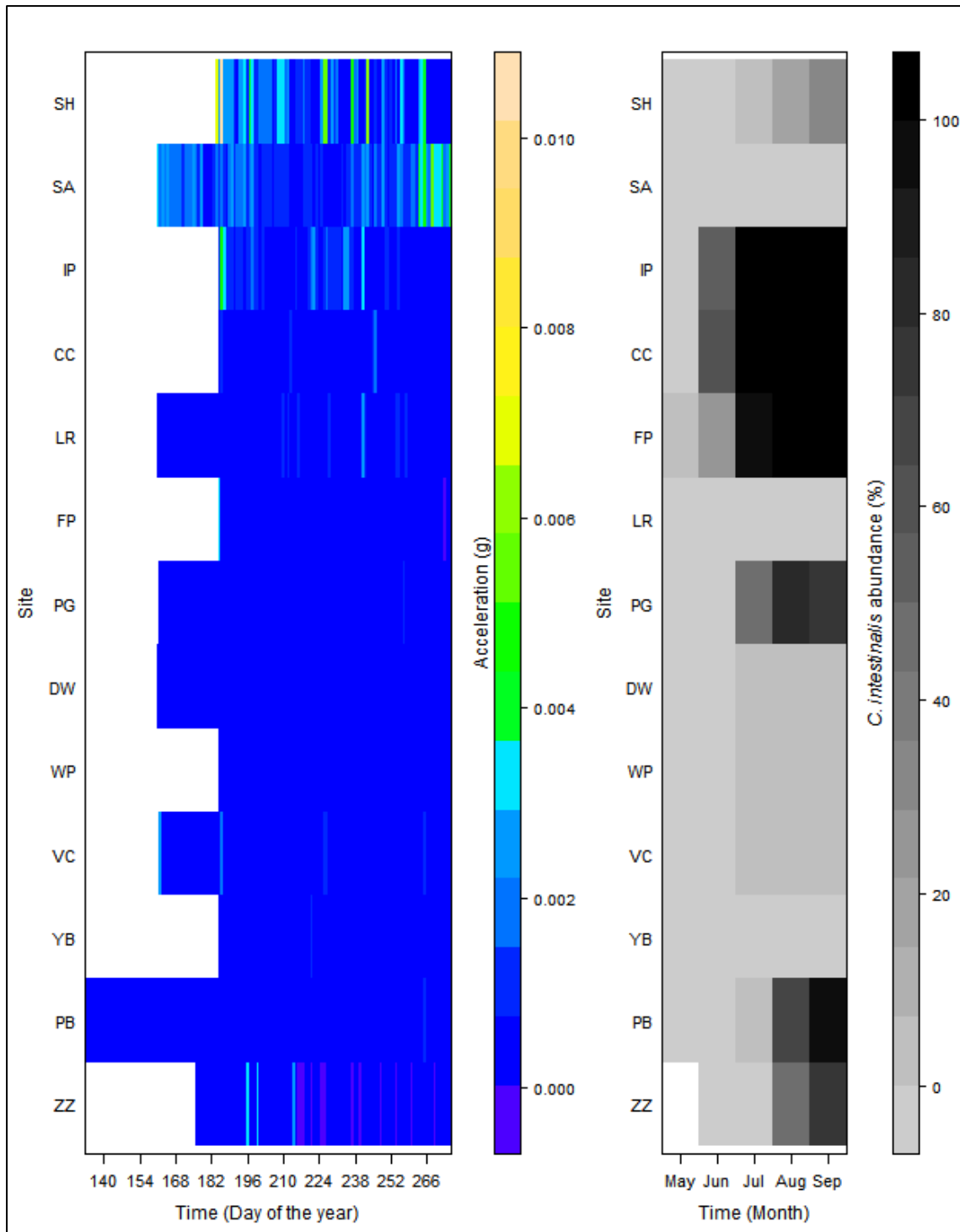
**Figure 3.10: Left - The pH profiles of the expanded sites monitored in 2015 are presented by a heatmap of the daily mean pH values. The sites are ordered by their overall mean pH, with the highest at the top. \* The white sections present in the LB and IN profiles are missing data points. Right – The mean *C. intestinalis* abundance for each measurement period in 2015 shown in a heatmap, with sites corresponding to the appropriate site from the pH plot on the left.**

### 3.3.4. Water motion

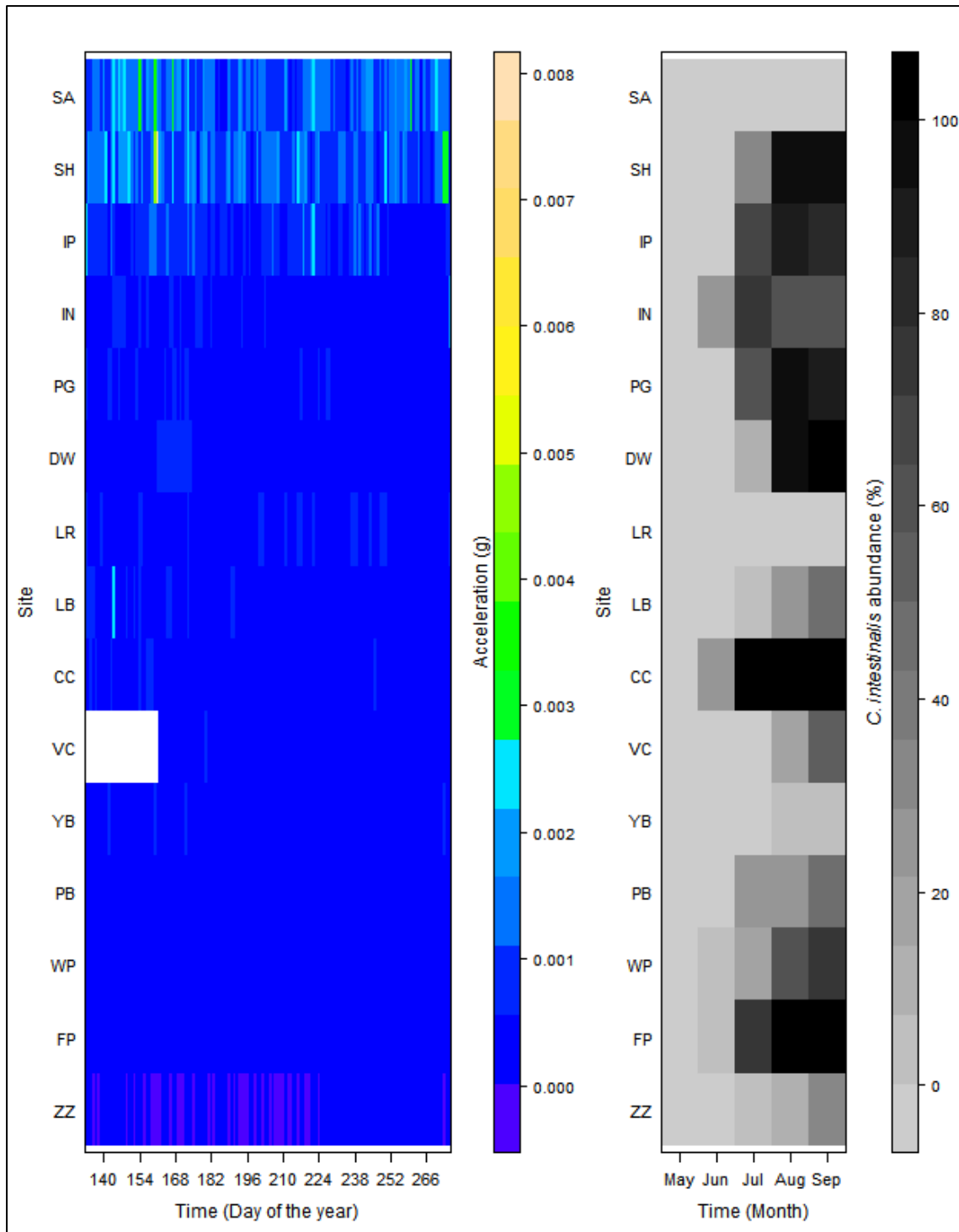
The overall mean values of the daily variance in acceleration were lower in 2015, the exceptions to which are PB and YB (Table 3.4). SA and SH, stand out as having higher water flow values, regardless of year. To a lesser extent, the same can be said for IP. All three of these sites are mussel leases, and are in sheltered bays, rather than man-made harbours and marinas. More sites with high *C. intestinalis* abundance are ranked near the top of the higher acceleration values in 2014, while no discernible trend is visible in 2015 (Fig. 3.11 & 3.12).

**Table 3.4: Summary water motion values for all 15 sites (in alphabetical order) and both years (where applicable). The 2014 values, where available are shaded in light grey, while IN and LB are shaded in dark grey. Mean = Overall mean of daily variance. Minimum = Minimum daily variance. Maximum = Maximum daily variance. Variance = Variance of daily variance.**

Site	Year	Minimum	Mean	Maximum
CC	2014	$3.14 \times 10^{-4}$	$6.1 \times 10^{-4}$	$1.48 \times 10^{-3}$
CC	2015	$1.03 \times 10^{-4}$	$3.06 \times 10^{-4}$	$7.73 \times 10^{-4}$
DW	2014	$1.2 \times 10^{-4}$	$4.03 \times 10^{-4}$	$6.1 \times 10^{-4}$
DW	2015	$7.27 \times 10^{-5}$	$3.35 \times 10^{-4}$	$6.12 \times 10^{-4}$
FP	2014	$1.94 \times 10^{-5}$	$4.61 \times 10^{-4}$	$3.14 \times 10^{-4}$
FP	2015	$8.8 \times 10^{-5}$	$1.62 \times 10^{-4}$	$2.95 \times 10^{-4}$
IN	2015	$1.31 \times 10^{-4}$	$4.57 \times 10^{-4}$	$1.77 \times 10^{-3}$
IP	2014	$3.33 \times 10^{-4}$	$9.1 \times 10^{-4}$	$4.77 \times 10^{-4}$
IP	2015	$1.98 \times 10^{-4}$	$7.04 \times 10^{-4}$	$2.34 \times 10^{-3}$
LB	2015	$1.01 \times 10^{-4}$	$3.15 \times 10^{-4}$	$2.23 \times 10^{-3}$
LR	2014	$1.4 \times 10^{-4}$	$4.62 \times 10^{-4}$	$2.46 \times 10^{-3}$
LR	2015	$6.35 \times 10^{-5}$	$3.23 \times 10^{-4}$	$9.21 \times 10^{-4}$
PB	2014	$6.75 \times 10^{-5}$	$2.19 \times 10^{-4}$	$8.21 \times 10^{-4}$
PB	2015	$1.56 \times 10^{-4}$	$2.38 \times 10^{-4}$	$4.41 \times 10^{-4}$
PG	2014	$2.02 \times 10^{-4}$	$4.33 \times 10^{-4}$	$8.47 \times 10^{-4}$
PG	2015	$1.17 \times 10^{-4}$	$3.98 \times 10^{-4}$	$1.02 \times 10^{-3}$
SA	2014	$4.18 \times 10^{-4}$	$1.53 \times 10^{-3}$	$5.78 \times 10^{-3}$
SA	2015	$3.46 \times 10^{-4}$	$1.29 \times 10^{-3}$	$3.55 \times 10^{-3}$
SH	2014	$8.92 \times 10^{-5}$	$1.7 \times 10^{-3}$	$1.02 \times 10^{-2}$
SH	2015	$4.16 \times 10^{-4}$	$1.29 \times 10^{-3}$	$7.63 \times 10^{-3}$
VC	2014	$9.36 \times 10^{-5}$	$3.03 \times 10^{-4}$	$2.37 \times 10^{-3}$
VC	2015	$1.07 \times 10^{-4}$	$3 \times 10^{-4}$	$6.81 \times 10^{-4}$
WP	2014	$2.24 \times 10^{-4}$	$3.35 \times 10^{-4}$	$7.19 \times 10^{-4}$
WP	2015	$8.89 \times 10^{-5}$	$2.36 \times 10^{-4}$	$5.29 \times 10^{-4}$
YB	2014	$7.31 \times 10^{-5}$	$2.34 \times 10^{-4}$	$8.95 \times 10^{-4}$
YB	2015	$4.69 \times 10^{-4}$	$2.86 \times 10^{-4}$	$6.95 \times 10^{-4}$
ZZ	2014	$9.75 \times 10^{-6}$	$1.46 \times 10^{-4}$	$3.12 \times 10^{-3}$
ZZ	2015	$1.27 \times 10^{-6}$	$3.75 \times 10^{-5}$	$4.21 \times 10^{-4}$



**Figure 3.11: Left - The water motion profiles of the core sites monitored in 2014 are presented by a heatmap of the daily variance in water motion values. The sites are ordered by the overall mean variance in water motion, with the highest at the top. \* The white sections at the start of the profiles for all sites (except PB) are missing data points. Right – The mean *C. intestinalis* abundance for each measurement period in 2014 represented as a heatmap, with sites corresponding to the appropriate site from the water motion plot on the left. \* The white section at the start of ZZ is a missing datum point.**



**Figure 3.12: Left - The water motion profiles of the expanded sites monitored in 2015 are presented by a heatmap of the daily variance in water motion values. The sites are ordered by the overall mean variance in water motion, with the highest at the top. \* The white sections at the start of the VC profile are missing data points. Right – The mean *C. intestinalis* abundance for each measurement period in 2015 shown in a heatmap, with sites corresponding to the appropriate site from the water motion plot on the left.**

### 3.4. Correlation Analysis Results

Results from exploratory correlation analyses suggest that throughout the two years of monitoring, temperature displayed the strongest relationship with *C. intestinalis* abundance (Table 3.5). In 2014, mean *C. intestinalis* abundance was significantly correlated ( $p < 0.0001$ ) to the minimum and mean temperature metrics, as well as the minimum and mean salinity metrics (Table 3.5). In 2015, mean *C. intestinalis* abundance was significantly correlated (at  $p < 0.0001$ ) to all three of the temperature metrics (Table 3.5). There are a number of additional significant correlations between mean *C. intestinalis* abundance and abiotic metrics, but at a lower significance threshold.

**Table 3.5: Results from Kendall’s Rank Correlation Coefficient ( $\tau$ ) tests between site-specific mean *C. intestinalis* for each measurement period and the corresponding site-specific abiotic variable measurement period values.**

Abiotic variable metric	Kendall’s tau:		Kendall’s tau:	
	2014 <i>C. intestinalis</i>	<i>p</i> -value	2015 <i>C. intestinalis</i>	<i>p</i> -value
Min. daily-mean temp.	0.42	< 0.0001	0.56	< 0.0001
Mean temperature	0.41	< 0.0001	0.55	< 0.0001
Max. daily-mean temp.	0.32	< 0.001	0.49	< 0.0001
Min. daily-mean sal.	0.45	< 0.0001	0.28	< 0.01
Mean salinity	0.43	< 0.0001	0.29	< 0.01
Max. daily-mean sal.	0.36	< 0.001	0.28	< 0.01
Min. daily-mean pH	-0.05	> 0.05	-0.10	> 0.05
Mean pH	-0.27	< 0.05	-0.17	> 0.05
Max. daily-mean pH	-0.40	< 0.001	-0.22	< 0.05
Min. daily-var. accel	0.10	> 0.05	-0.17	> 0.05
Var accel	0.02	> 0.05	-0.19	< 0.05
Max. daily-var. accel	0.01	> 0.05	-0.15	> 0.05



### 3.5. NLME model: 2014

#### 3.5.1. Preselection of abiotic variables

The four metrics remaining after the stepwise removal process was carried out on the 2014 abiotic data were: mean temperature; mean salinity; minimum pH; and minimum daily-variance in acceleration (Table 3.6). There were a number of very highly correlated ( $r > 0.9$ ) metrics (Table 3.6). All of the very highly ( $> 0.9$ ) correlated metrics were within the same abiotic variable, e.g. mean and maximum temperature ( $r = 0.92$ ).

**Table 3.6: Stepwise removal of a single abiotic metric from the most correlated pair using the 2014 abiotic metrics. The pair of metrics most correlated at each step is highlighted. The light grey shaded metric was kept (Keep), while the metric shaded with dark grey was removed (Rem.) from the process. There was a selection rule to always have one metric from each variable remaining.**

Variable	Metric	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7	Step 8*	Step 9
Temp.	Mean	Keep		Keep				Keep	Keep	
	Minimum			Rem.						
	Maximum	Rem.								
Salinity	Mean					Keep	Keep			Keep
	Minimum						Rem.			
	Maximum					Rem.				
pH	Mean		Rem.							
	Minimum								Rem.	
	Maximum		Keep					Rem.		
Water Motion Accel.	Variance				Rem.					
	Minimum									
	Maximum				Keep					Rem.
Pearson's r		0.92	0.91	0.9	0.89	0.83	0.75	-0.74	-0.53	-0.36

As the highly collinear metrics were reduced, some moderate to high (0.7 to 0.89) correlations were present between metrics from different abiotic variables, e.g. mean temperature and maximum pH ( $r = -0.74$ ). If the selection rule requiring the presence of one metric from each abiotic in the final four was not implemented, then all metric pairings with

values higher than a moderate correlation ( $> 0.5$ ) would have been removed. In the last removal step (Step 8, Table 3.6) neither of the most correlated pair could be removed, as they were the last representative for their respective variable. The second most correlated pair was used instead (Step 9, Table 3.6).

### 3.5.2. Bottom-up and top-down model candidate sets

The best model in 2014 was the same for the bottom-up (BU) and the top-down (TD) methods (Table 3.7 and Table 3.8). The best model (parameter estimates found in Table 3.9) contained only the mean salinity metric in each of the three model parameters and had an Akaike weight ( $w_i$ ) of 0.48 out of the total ( $N = 30$ ) set of compared models (candidate set) for the BU approach (Table 3.7). The  $w_i$  for the same model from the TD candidate set ( $N = 29$ ) was 0.3 (Table 3.8). In the BU candidate set of models, the second best model had an AICc difference ( $\Delta_i$ ) of  $\sim 2$  compared to the best model (Table 3.7). This means that this second-best model has substantial evidence supporting its likelihood of being the best model within the candidate set for these data (Burnham and Anderson, 2003). Once the  $\Delta_i$  increases to the 4 – 7 range, these candidate models are considered to have considerably less support and are not as likely to be the best model, based on these data. The remaining eight models displayed in Table 3.7 from the BU candidate set are in the 4 – 7 range. In the TD candidate set of models there were four models within the 2  $\Delta_i$  range of the best model. In this case, the TD candidate set has a higher level of uncertainty related to the selection of a best model. The remaining five models presented in Table 3.8 are in the 4 – 7  $\Delta_i$  range. The total  $w_i$  of any candidate set equals 1 and in the BU candidate comparison, the bottom twenty models not presented in Table 3.7 only have a cumulative weight of 0.01 between them. The bottom nineteen models not presented in the TD candidate set in Table 3.8 have a cumulative weight of 0.03 between them. None of these models are worth considering any further. The second-best model in the BU candidate set has the mean temperature metric included in the Days 50 parameter, in addition to the salinity metric in all three parameters like the best model. The evidence ratio ( $w_i / w_j$ ) between this and the best BU model is 3, roughly meaning the best model is three times as likely as the second best to be the most likely model given the data. In the TD candidate set, the four other models with substantial support are all like the full TD model in terms of the temperature, salinity, and acceleration metrics in all three parameters, except either one or two of the temperature metrics have been excluded. The evidence ratios for these four models in the TD set range between 1.5 and 2.5.

**Table 3.7: The ten best models (30 models in total) from the bottom-up model development approach (2014), ranked by lowest AICc value. \*Interaction term included between the variables in this model parameter. The site-level effect for all candidate models was specified as MaxPop only. K = no. model parameters.  $\Delta_i$  = AICc difference. Model Lik. = model likelihood.  $w_i$  = Akaike weight. The site-level effects for all candidate models was specified as MaxPop only.**

Model rank	K	AICc	$\Delta_i$	Model Lik.	Model $w_i$	Cumulative $w_i$	Maximum population (MaxPop)		Days 50 (D50)		Lag phase (Lag)	
1	8	2524.76	0.00	1.00	0.48	0.48	Salinity		Salinity		Salinity	
2	9	2526.93	2.16	0.34	0.16	0.65	Salinity		Salinity	Temp	Salinity	
3	10	2528.87	4.10	0.13	0.06	0.71	Salinity		Salinity	*Temp	Salinity	
4	10	2529.02	4.26	0.12	0.06	0.77	Salinity	*Temp	Salinity		Salinity	
5	10	2529.04	4.27	0.12	0.06	0.82	Salinity		Salinity		Salinity	*Temp
6	10	2529.14	4.38	0.11	0.05	0.88	Salinity	Temp	Salinity		Salinity	Temp
7	9	2529.49	4.73	0.09	0.05	0.92	Salinity		Salinity		Salinity	Temp
8	9	2529.86	5.10	0.08	0.04	0.96	Salinity	Temp	Salinity		Salinity	
9	10	2531.32	6.55	0.04	0.02	0.98	Salinity		Salinity	Temp	Salinity	Temp
10	12	2532.09	7.32	0.03	0.01	0.99	Salinity	Temp	Salinity		Salinity	*Temp

**Table 3.8: The ten best models (total N = 29) from the top-down model fitting approach for the 2014 data only. The models are ranked by their AICc value, a lower value being better. The site-level effects for all candidate models were specified as MaxPop only. K = no. model parameters.  $\Delta_i$  = AICc difference. Mod. Lik. = model likelihood. Cumul.  $w_i$  = cumulative Akaike weight.**

Model rank	K	AICc	$\Delta_i$	Mod. Lik.	Model $w_i$	Cumul. $w_i$	Maximum population (MaxPop)			Days 50 (D50)			Lag phase (Lag)		
1	8	2524.78	0.00	1.00	0.3	0.3		Sal			Sal			Sal	
2	13	2525.56	0.78	0.68	0.20	0.5	Temp	Sal	Accel	Temp	Sal	Accel		Sal	Accel
3	13	2526.02	1.24	0.54	0.16	0.66		Sal	Accel	Temp	Sal	Accel	Temp	Sal	Accel
4	12	2526.58	1.8	0.41	0.12	0.78		Sal	Accel	Temp	Sal	Accel		Sal	Accel
5	13	2526.58	1.8	0.41	0.12	0.9	Temp	Sal	Accel		Sal	Accel	Temp	Sal	Accel
6	13	2529.11	4.33	0.11	0.03	0.93	Temp	Sal	Accel	Temp		Accel	Temp	Sal	Accel
7	13	2530.64	5.86	0.05	0.02	0.95	Temp	Sal		Temp	Sal	Accel	Temp	Sal	Accel
8	12	2531.87	7.09	0.03	0.01	0.95	Temp	Sal	Accel	Temp	Sal		Temp	Sal	
9	12	2531.90	7.12	0.03	0.01	0.96	Temp	Sal	Accel	Temp		Accel	Temp		Accel
10	11	2532.01	7.23	0.03	0.01	0.97		Sal	Accel		Sal	Accel		Sal	Accel

**Table 3.9: Estimates, with their associated standard error (SE), for each of the species-wide and site-level model parameters from the best model in 2014 (bottom-up and top-down model building methods provided the same best model).**

Parameter	Type	Estimate (SE)
MaxPop Intercept	Species-wide	-0.48 (0.66)
MaxPop Intercept sd	Site-level	2.08 (10.13)
MaxPop (Avg Sal – 31)	Species-wide	0.65 (0.12)
D50 Intercept	Species-wide	4.48 (0.01)
D50 (Avg Sal – 31)	Species-wide	0.30 (0.01)
Lag Intercept	Species-wide	21.26 (5.60)
Lag (Avg Sal – 31)	Species-wide	-20.55 (6.29)

### 3.5.3. Observed values vs. fitted values

#### 3.5.3.1. Best bottom-up (BBU) and best top-down (BTD) 2014 model

The proportion of the fitted values attributable to the site-level effect and the species-wide effect varies from site to site in the best BU and TD model in 2014 (Fig. 3.13). The fitted values predicted by the best model based on the 2014 data are almost a perfect fit, as can be seen by the black dashed line (full-model fitted values) fit almost directly over the observed mean *C. intestinalis* relative abundance (red solid line). The degree to which the site-level effects improve the full-model fit is displayed in Fig. 3.13, where some sites like ZZ and YB are fit largely by site-level effects. Conversely, SH and DW are fit almost entirely by species-wide effects. As days after deployment was used as the known time component in the models, and ZZ was deployed much later than other sites in 2014, it resulted in the odd plot alignment of the growth at ZZ. The plotted growth (fitted and observed) also shows how much of the data is missing from the 2014 models, as missing values cannot be accommodated. The 95% confidence and prediction intervals for the BBU and BTD 2014 display the high level of confidence in the model fit (Fig. 3.14).

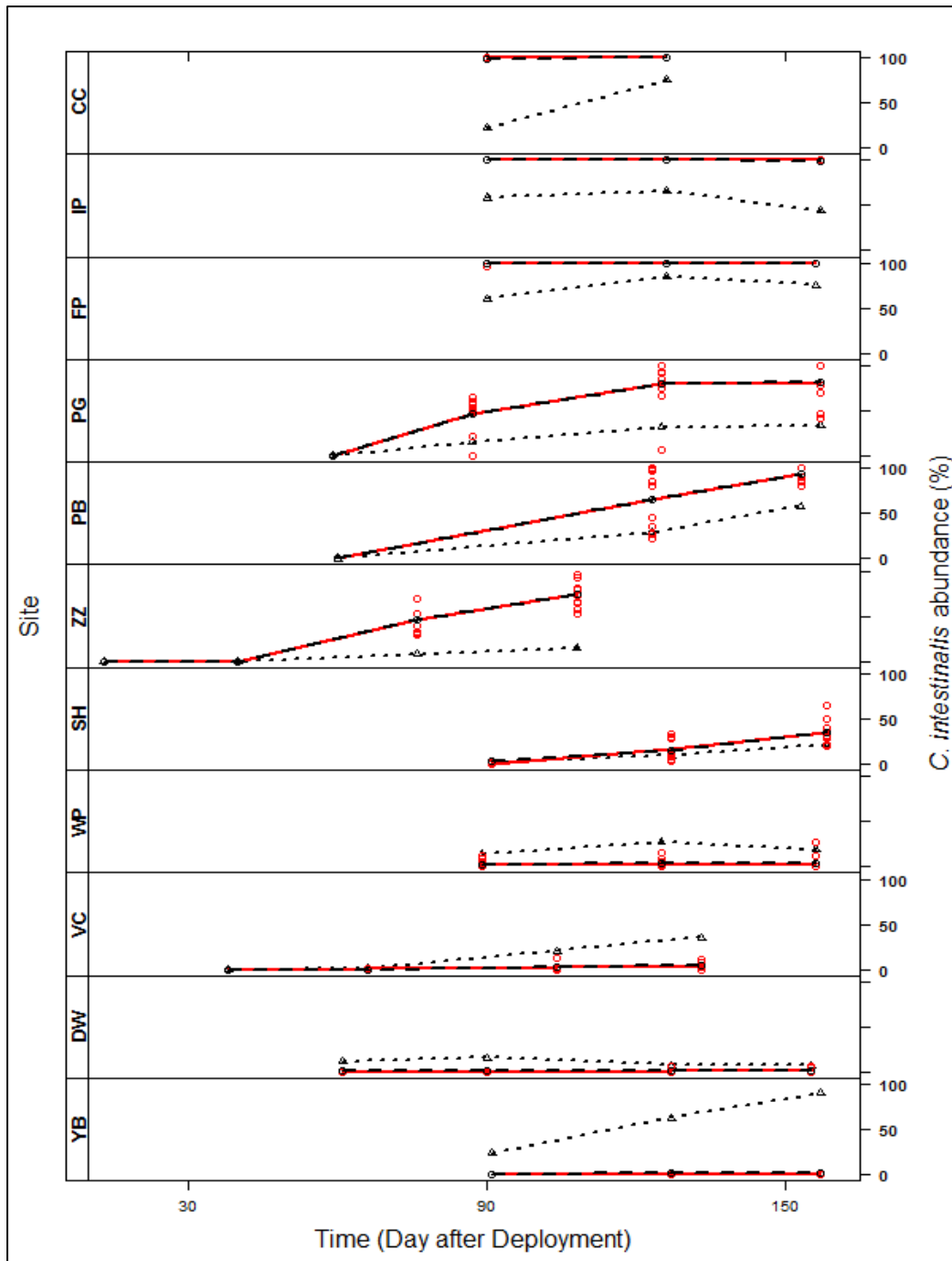
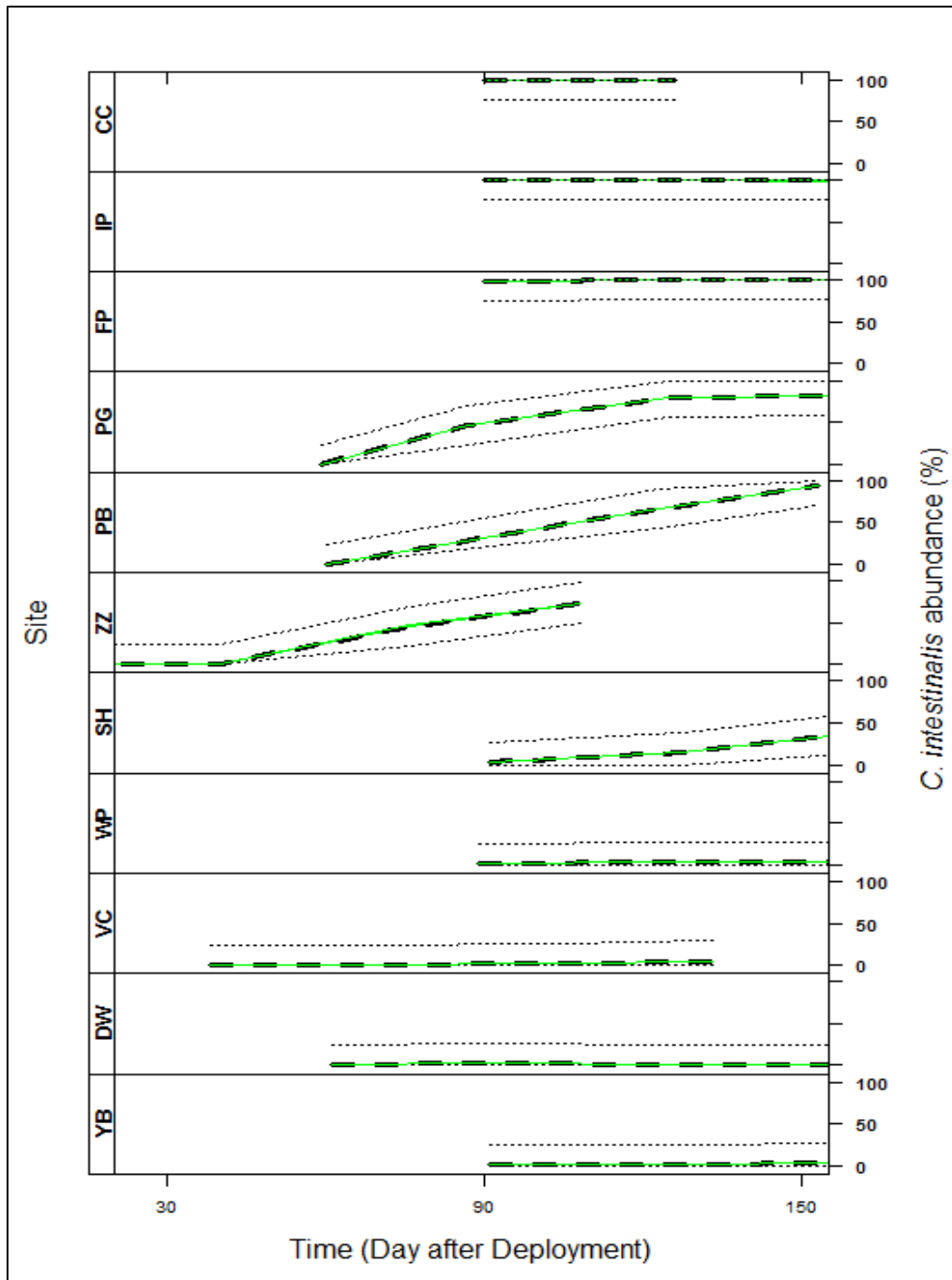


Figure 3.13: Full-model fitted *C. intestinalis* relative abundance (black dashed line) for the 11 sites, fit by the 2014 best model (same model for BU and TD), and the fitted values using species-wide effect only (black dotted line). The observed *C. intestinalis* values (red line and circle markers) are overlaid. ZZ had a delayed deployment.

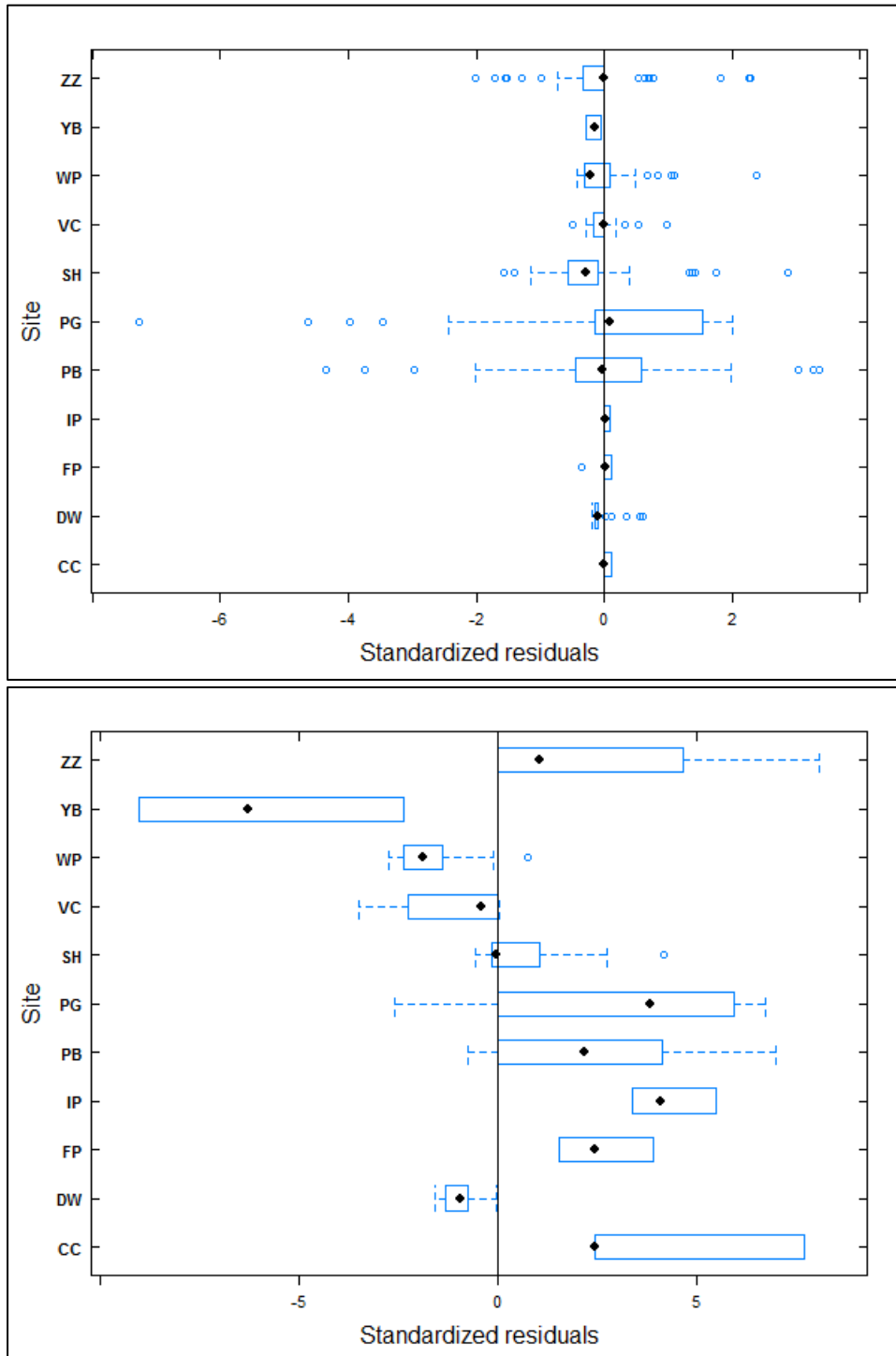


**Figure 3.14: Full-fit from the 2014 best model (solid green line), with 95 % confidence intervals (dashed black lines), and 95 % prediction intervals (dotted black lines).**

### 3.5.3.2. Residual plots: BBU and BTD 2014 model

On the whole, the residuals from the full-fit best BU and TD model in 2014 are good, not seeming to violate any assumptions egregiously (Fig. 3.15). The means of the residuals at the site-level are all centred near zero, while their variance around the mean is homogeneously distributed, with some exceptions. There are some outlying residuals; the most obvious of which can be seen at PB (positive and negative outliers) and PG (negative outliers). The

residuals from the species-wide effects of the best BU and TD model in 2014, when inspected in isolation, appear to violate a number of assumptions. Bias (heterogeneity of variance around the mean) is introduced into the residuals, while the means deviate from zero, substantially in some cases, at the site-level (Fig. 3.15).

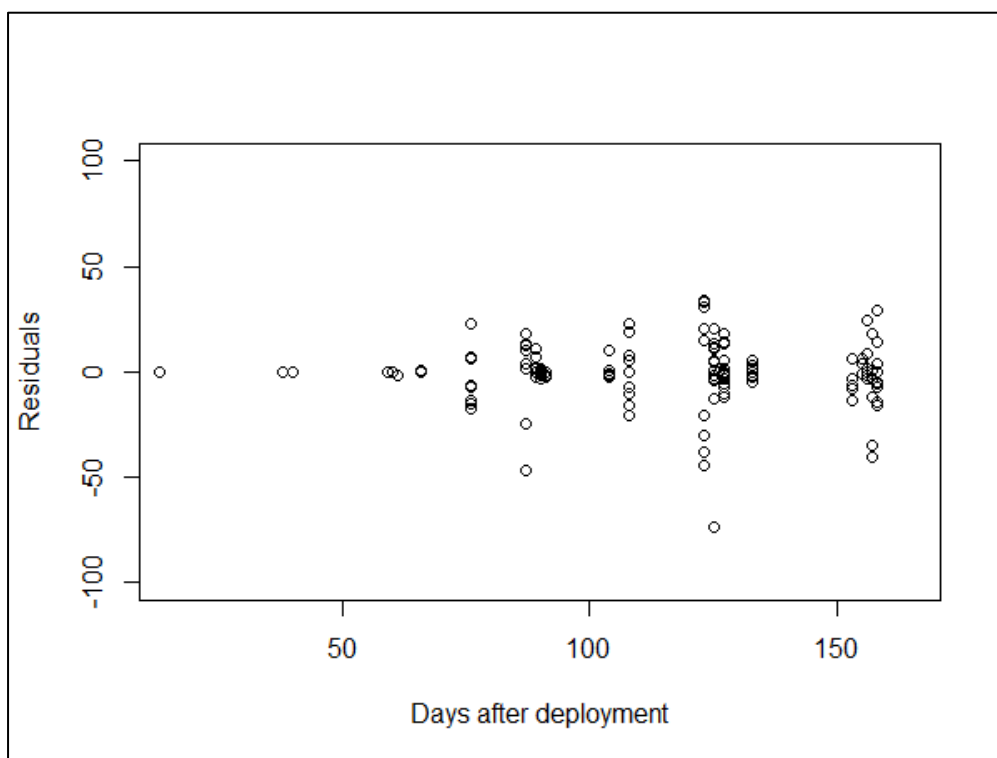


**Figure 3.15: BBU and BTD 2014 residual boxplots by site, ordered alphabetically. The full-model residuals are displayed in the upper plot, while the species-wide effects residuals in isolation are displayed in the lower plot.**

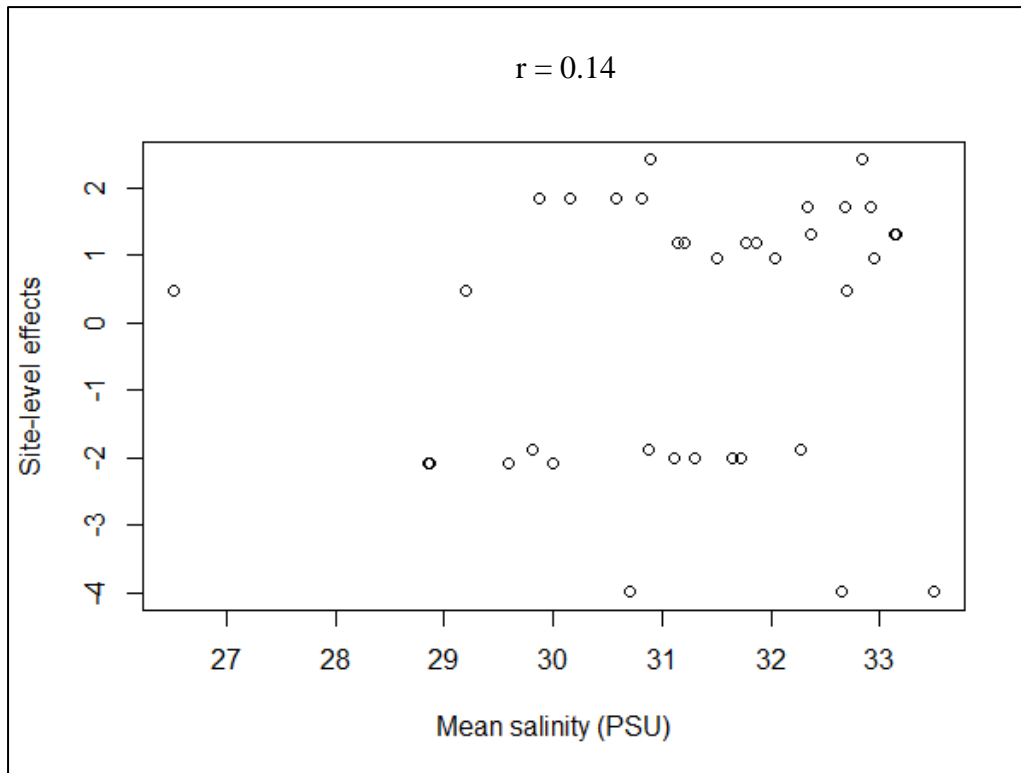


Heteroskedasticity is apparent in the residuals of the full-fit for the BBU and BTM in 2014 (Fig. 3.16). While this is a violation of model assumptions, there is no bias in the residuals and heteroscedasticity is a very common feature in growth models, where the observed values are increasing rapidly through time. When a covariance and correlation structure was introduced into the model to account for variability within individual settlement plates through time, it did not improve the model or the heteroscedasticity of residuals, so the more parsimonious model without an autocorrelation structure was used.

There is no relationship between the site-level effects from the best BU and TD model in 2014 and the independent variable, mean salinity (Fig. 3.17).



**Figure 3.16: The full-model residuals for the BBU and BTM 2014 plot against days after deployment.**



**Figure 3.17: Mean salinity, the independent variable used in the BBU and BTB 2014 plotted against the site-level effects to inspect their relationships in respect to homogeneity, with the correlation coefficient presented at the top of the plot.**

### 3.5.3.3. Best temperature-only model (BTO) 2014

The best temperature-only model (BTO) contained the mean temperature metric in the MaxPop and D50 model parameters, while MaxPop was set as a site-level effect in addition to all model parameters being species-wide effects by default. The *C. intestinalis* values fit by the full BTO model are almost all a direct fit with the mean *C. intestinalis* percent cover observed (Fig. 3.18), except in a couple of cases, PB, PG, and SH, where the fitted growth is slightly different to observed. The proportion of the fit associated with species-wide effect and site-level effects can be visually assessed, with a wide degree of inter-site variation. The species-wide effects at some sites are consistently and markedly different to the observed values, like CC and IP. Some sites (FP) follow a similar pattern, only for the species-wide effects to then contribute to a closer fit compared to the observed values nearer the end of the monitoring period. The species-wide effects fitted values can also be very close to the observed values at the start of the monitoring period, with the fitted values deviating further from the observed values as time progresses: this is evident for VC, DW, and YB. A couple of sites, namely PB and ZZ are fit quite well by the species-wide effects.

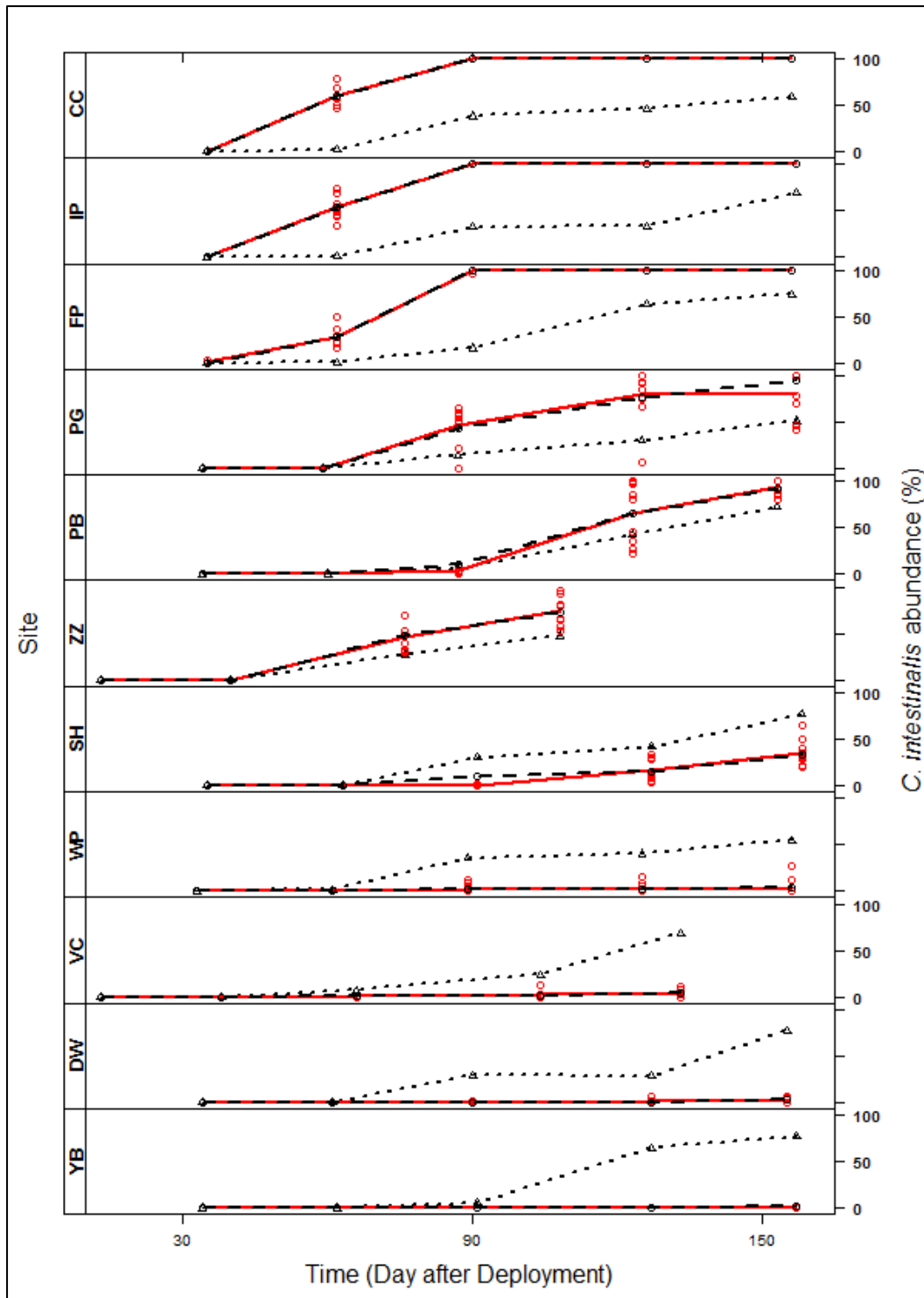


Figure 3.18: Full-model fitted *C. intestinalis* relative abundance (black dashed line) for the 11 sites, fit by the 2014 best temperature-only model (BTO), and the fitted values using species-wide effect only (black dotted line). The observed *C. intestinalis* values (red line and circle markers) are overlaid. ZZ had a delayed deployment.

### 3.6. NLME model: 2015

#### 3.6.1. Preselection of abiotic variables

**Table 3.10: Stepwise removal of a single abiotic variable metric from the most highly correlated pair using the 2015 abiotic data. The pair of metrics most correlated at each step is highlighted. The light grey shaded metric was kept (Keep) while the metric shaded with dark grey was removed (Rem.) from the process. There was a selection rule to always have one metric from each variable remaining. Only pairwise complete observations were used in the correlation.**

Variable	Metric	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7	Step 8
Temp.	Mean	Keep		Keep					Keep
	Minimum			Rem.					
	Maximum	Rem.							
Salinity	Mean		Rem.						
	Minimum		Keep					Rem.	
	Maximum								
pH	Mean				Rem.				
	Minimum								
	Maximum				Keep				Rem.
Water	Variance					Keep	Rem.		
Motion	Minimum					Rem.			
Accel.	Maximum						Keep	Keep	
Pearson's r		0.95	0.93	0.93	0.9	0.9	0.86	-0.74	-0.52

The four abiotic metrics remaining after the stepwise removal process was carried out on the pairwise-complete observations of the 2015 abiotic data were mean temperature, maximum salinity, minimum pH, and maximum variance in acceleration (Table 3.10). All of the highly correlated ( $r > 0.9$ ) metric pairings were within the same abiotic variable (e.g. mean and maximum temperature [ $r = 0.95$ ]). As the highly correlated metrics were removed, some moderate-to-high correlations were present between metrics from two different abiotic variables (e.g. minimum salinity and maximum variance in acceleration [ $r = -0.74$ ]).

#### 3.6.2. Bottom-up and top-down model candidate sets

In 2015 the best bottom-up model (BBU) was more parsimonious than the best top-down model (BTD), but the BTD had a much lower AICc value (Table 3.11 & 3.12 respectively). The BBU (parameter estimates found in Table 3.13) contained the maximum salinity metric in the MaxPop, Days50, and Lag model parameters. It also contained the maximum variance in acceleration and minimum pH metrics in the lag parameter. The BBU's Akaike weight ( $w_i$ ) of 0.25 out of the total ( $N = 34$ ) set of compared models (candidate set) for the bottom-up (BU) approach is quite low (Table 3.11). In the BU candidate set, the models ranked from 2 – 5 in the candidate set were within an AICc difference ( $\Delta_i$ ) of two compared to the BBU, while models ranked 6 and 7 were within and  $\Delta_i$  of 4 (Table 3.11). The low  $w_i$  and the high number of models within an  $\Delta_i$  of 4 suggests that there is a high level of uncertainty as to what the 'best' model is in this set.

The BTD was similar to the global model (GM), with all four of the abiotic metrics from the preselection process (Table 3.10) in each of the three model parameters, except for maximum variance in acceleration, which was not included in the MaxPop model parameter (Table 3.12). The BTD (parameter estimates found in Table 3.13) has a relatively high number of parameters ( $K = 17$ ), but AICc (Eq. 2) accounts for parameterization, effectively penalizing over parameterization if it occurs. In the TD candidate set of models there were two models within a  $\Delta_i$  of 2 from the 'best' model and one additional model that had a  $\Delta_i$  of  $< 4$  (Table 3.12). One model (5<sup>th</sup> ranked model) in the TD candidate set was in the 4 – 7  $\Delta_i$  range (Table 3.12). When accounting for rounding, the bottom nineteen models not presented from the TD candidate set in Table 3.12 have no cumulative weight. None of these models are worth considering any further. The abiotic metrics included in the parameters of the models with a  $\Delta_i$  of 7 (also likely to be the best model) of the BTD are all quite similar, mostly missing a salinity metric or two in the model parameters, while the GM is also within a  $\Delta_i$  of 2 from the BTD. While a  $w_i$  of 0.47 for the BTD does not suggest that it is an overwhelmingly better model than any other in the candidate set, it is a more likely 'best' model within the set than was provided by the BU approach. Despite this, there is still a relatively high degree of uncertainty as to what the best model is in the TD candidate set. The  $\Delta_i$  between the BTD and BBU is  $> 11$ , strongly suggesting that the BTD is much more likely to be the best 2015 model, given these data.

**Table 3.11: The ten best models (n = 34) from the bottom-up model candidate set from the 2015 data, ranked by lowest AICc value. The abiotic metric abbreviations are explained as: Temp. = mean temperature; Sal. = max salinity; pH = minimum pH; and Accel. = maximum variance in acceleration. K = number of model parameters.  $\Delta_i$  = AICc difference.  $w_i$  = Akaike weight.**

Model rank	K	AICc	$\Delta_i$	Model Likelihood	Model $w_i$	Cumulative $w_i$	Maximum population* (MaxPop)			Days 50* (D50)			Lag phase (Lag)			
1	12	6111.7	0	1	0.25	0.25		Sal			Sal			Sal	Accel	pH
2	11	6111.9	0.25	0.88	0.22	0.47		Sal			Sal			Sal		pH
3	12	6112.7	1.01	0.60	0.15	0.62		Sal			Sal	Accel		Sal		pH
4	13	6113.5	1.81	0.41	0.10	0.72		Sal			Sal	Accel		Sal	Accel	pH
5	12	6113.6	1.89	0.39	0.10	0.82		Sal	Accel		Sal			Sal		pH
6	12	6114.7	2.97	0.23	0.06	0.88	Temp	Sal			Sal			Sal		pH
7	12	6115.7	4.00	0.14	0.03	0.91		Sal		Temp	Sal			Sal		pH
8	11	6116.5	4.78	0.09	0.02	0.93	Temp			Temp		Accel	Temp			
9	13	6118	6.29	0.04	0.01	0.94		Sal			Sal		Temp	Sal	Accel	pH
10	10	6118.7	7.02	0.03	0.01	0.95		Sal			Sal			Sal		

**Table 3.12: The ten best models from the 2015 top-down candidate set (n = 29), ranked by AICc value. Each model parameter (MaxPop, D50, and Lag) could have all four of the abiotic metrics in it in any given model. The metrics that were included in each model are shown in the table, with any grey filled cells showing which metrics were excluded from a given model. The abiotic metrics in the table have been assigned shorthand names as follows: Temp = mean temperature; Sal = maximum salinity; pH = minimum pH; and Accel = maximum variance in acceleration. K = number of model parameters.  $\Delta_i$  = AICc difference.  $w_i$  = Akaike weight. \*MaxPop is the only site-level effect in all models within this candidate set.**

Model rank	K	AICc	$\Delta_i$	Mod. Lik.	Model $w_i$	Cumulative $w_i$	Maximum Population* (MaxPop)				Days 50 (D50)				Lag phase (Lag)			
							Temp	Sal	Accel	pH	Temp	Sal	Accel	pH	Temp	Sal	Accel	pH
1	16	6100.2	0	1	0.47	0.47	Temp	Sal		pH	Temp	Sal	Accel	pH	Temp	Sal	Accel	pH
2	16	6101.5	1.3	0.52	0.24	0.71	Temp	Sal	Accel	pH	Temp		Accel	pH	Temp	Sal	Accel	pH
3	17	6102.2	2.0	0.36	0.17	0.88	Temp	Sal	Accel	pH	Temp	Sal	Accel	pH	Temp	Sal	Accel	pH
4	15	6103.3	3.2	0.20	0.09	0.97	Temp		Accel	pH	Temp	Sal	Accel	pH	Temp		Accel	pH
5	15	6106.5	6.4	0.04	0.02	0.99	Temp	Sal	Accel	pH	Temp		Accel	pH	Temp		Accel	pH
6	15	6109.0	8.8	0.01	0.01	1	Temp		Accel	pH	Temp		Accel	pH	Temp	Sal	Accel	pH
7	16	6114.5	14.4	0	0.00	1	Temp	Sal	Accel		Temp	Sal	Accel	pH	Temp	Sal	Accel	pH
8	14	6115.5	15.4	0	0.00	1	Temp	Sal	Accel		Temp	Sal	Accel		Temp	Sal	Accel	
9	16	6115.5	15.4	0	0.00	1	Temp	Sal	Accel	pH	Temp	Sal		pH	Temp	Sal	Accel	pH
10	15	6117.4	17.3	0	0.00	1	Temp	Sal	Accel		Temp	Sal	Accel		Temp	Sal	Accel	pH

**Table 3.13: Estimates (est.), with associated standard error (SE), for each of the species-wide (Sp.) and site-level (Site) model parameters from the best model, for both the bottom-up (BBU) and the top-down (BTD) model building approaches using 2015 data. The absolute scaled effect size (ASES) is provided for each abiotic metric parameter. Temp. = mean temperature - 14; Sal. = maximum salinity - 30; pH = minimum pH - 8; Accel. = maximum variance in acceleration.**

Parameter	Type	BBU est. (SE)	BBU ASES	BTD est. (SE)	BTD ASES
MP Int.	Sp.	1.06 (0.28)	-	27.48 (15.97)	-
MP Int. sd	Site	0.74 (12.42)	-	2.12 (12.41)	-
MP Temp.	Sp.	-	-	-3.76 (2.33)	57.57
MP Sal.	Sp.	-0.13 (0.20)	0.69	-2.54 (2.21)	13.08
MP pH	Sp.	-	-	18.10 (13.62)	9.51
MP Accel.	Sp.	-	-	-	-
D50 Int.	Sp.	4.56 (0.07)	-	11.29 (3.62)	-
D50 Int. sd	Site	0.20 (12.42)	-	-	-
D50 Temp.	Sp.	-	-	-0.92 (0.54)	14.07
D50 Sal.	Sp.	0.17 (0.07)	0.88	-0.31 (0.48)	1.57
D50 pH	Sp.	-	-	5.03 (3.18)	2.64
D50 Accel.	Sp.	-	-	-674.31 (236.48)	5.07
Lag Int.	Sp.	14.88 (3.48)	-	3.64 (0.51)	-
Lag Temp.	Sp.	-	-	-0.10 (0.06)	1.47
Lag Sal.	Sp.	-2.77 (2.51)	14.22	-0.01 (0.10)	0.07
Lag pH	Sp.	52.64 (16.95)	27.64	-1.90 (1.18)	1.00
Lag Accel.	Sp.	6352.57 (3766.51)	47.80	1901.05 (1182.88)	14.30

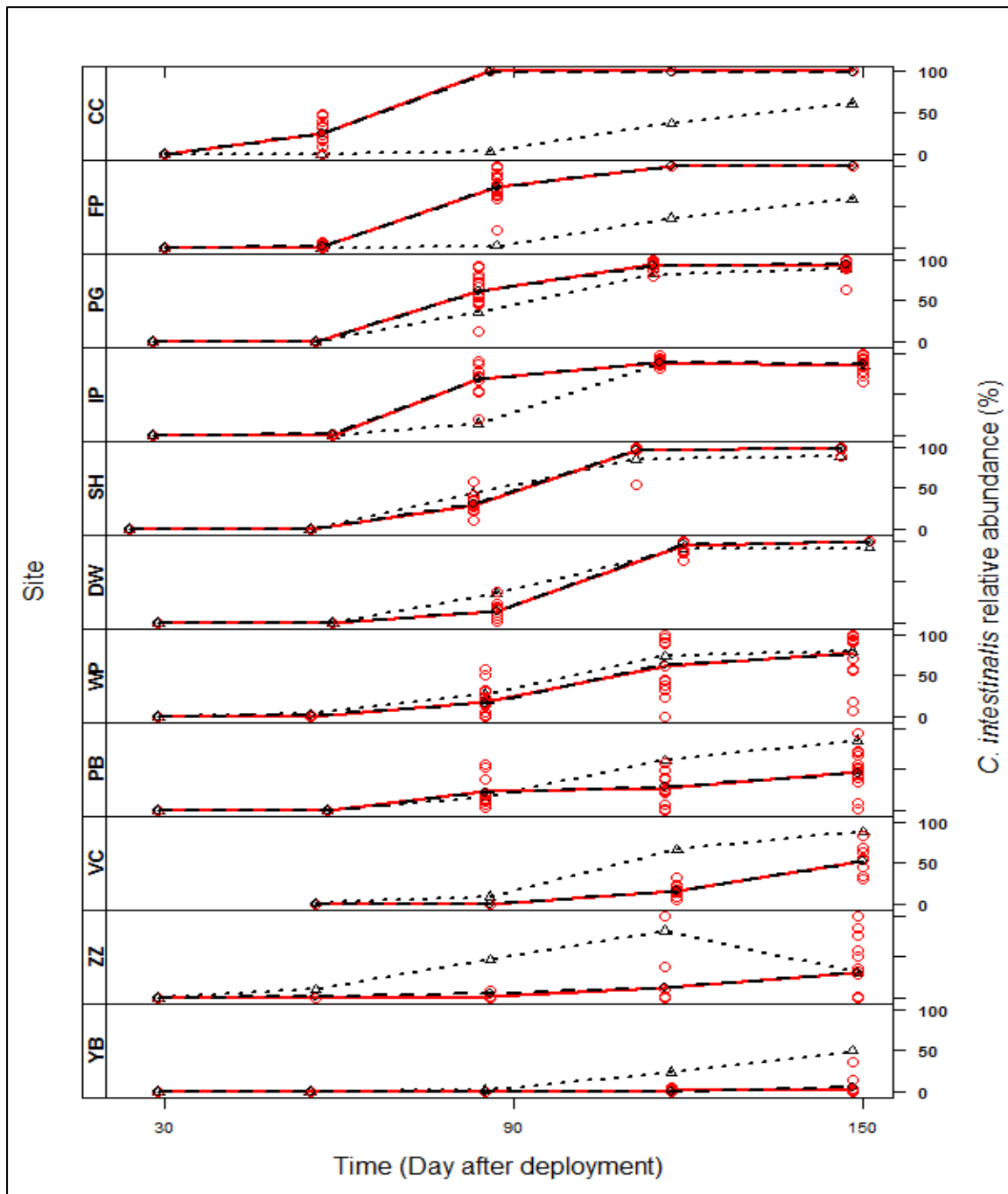
### 3.6.3. Observed values vs. fitted values

#### 3.6.3.1. Best bottom-up (BBU) 2015 model

A purely qualitative visual inspection of the fitted values against observed according to the BBU model in 2015 (Fig. 3.19) suggests that this model has a greater proportion of fit associated with species-wide effects than the best model fit in 2014 (Fig. 3.13). The full-fit (species-wide and site-level) is very good for all sites, with almost perfect fit to the average observed growth. Only at PB and PG does the full fit slightly deviate from the average observed growth. Three sites, IP, DW, and WP have a good species-wide effect fit. SH and



PB have a good fit attributed to the species-wide effect until the end of the growth period when it deteriorates.



**Figure 3.19: Full-model fitted *C. intestinalis* relative abundance (black dashed line) for the 11 sites, fit by the 2015 best bottom-up model (BBU), and the fitted values using species-wide effect only (black dotted line). The observed *C. intestinalis* values (red line and circle markers) are overlaid.**

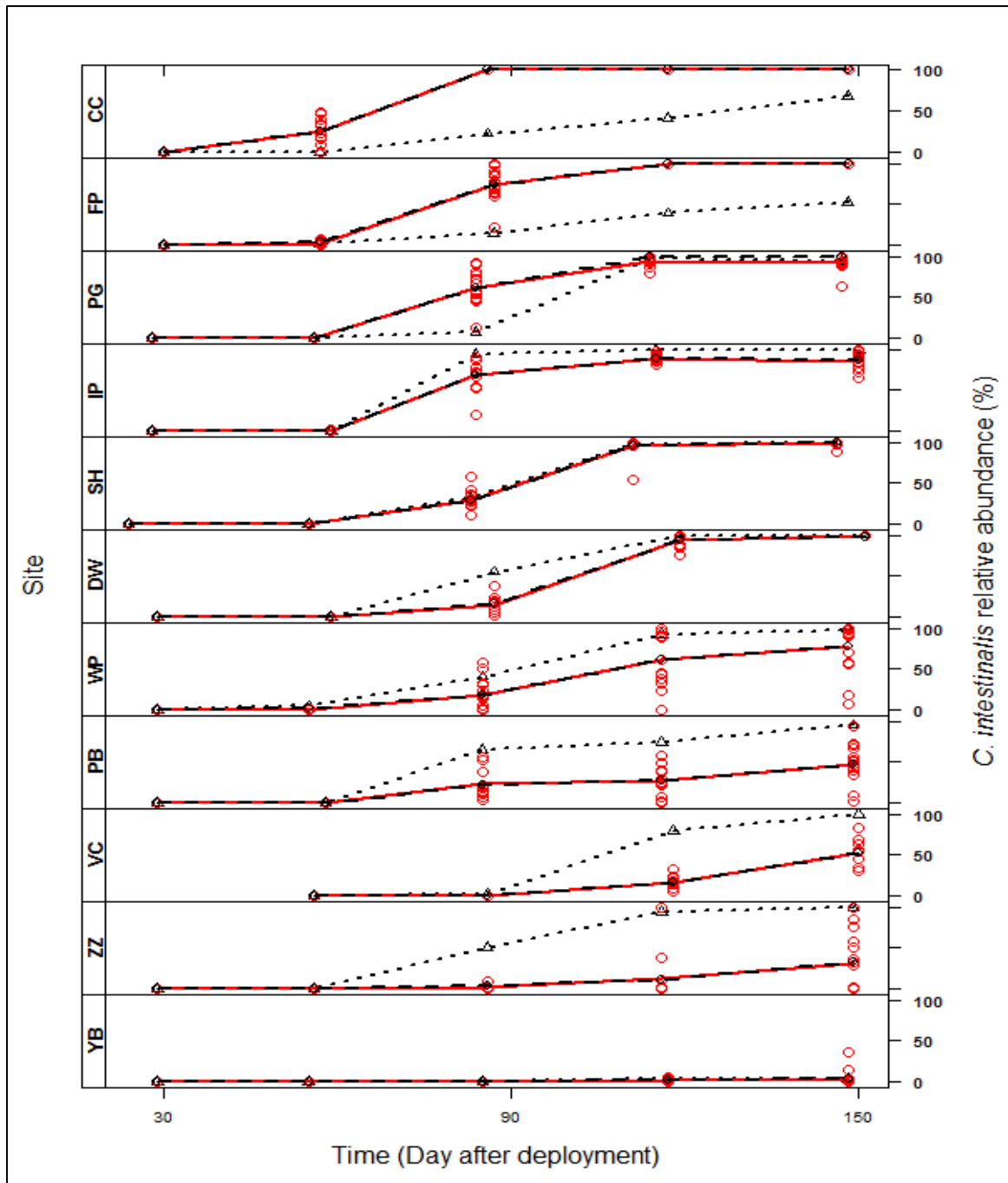
CC, FP, and PG all have a poor fit associated with the species-wide effect in the mid-section of the growth period, before it improves greatly by the end. In these cases, the species-wide

effect has accurately predicted the maximum population, but did not accurately fit the rate at which it got to the maximum point. VC, ZZ, and YB all have a good species-wide effect fit during the early period while no growth had actually occurred, before the species-wide effect over projects the growth rate and maximum population near the latter stages.

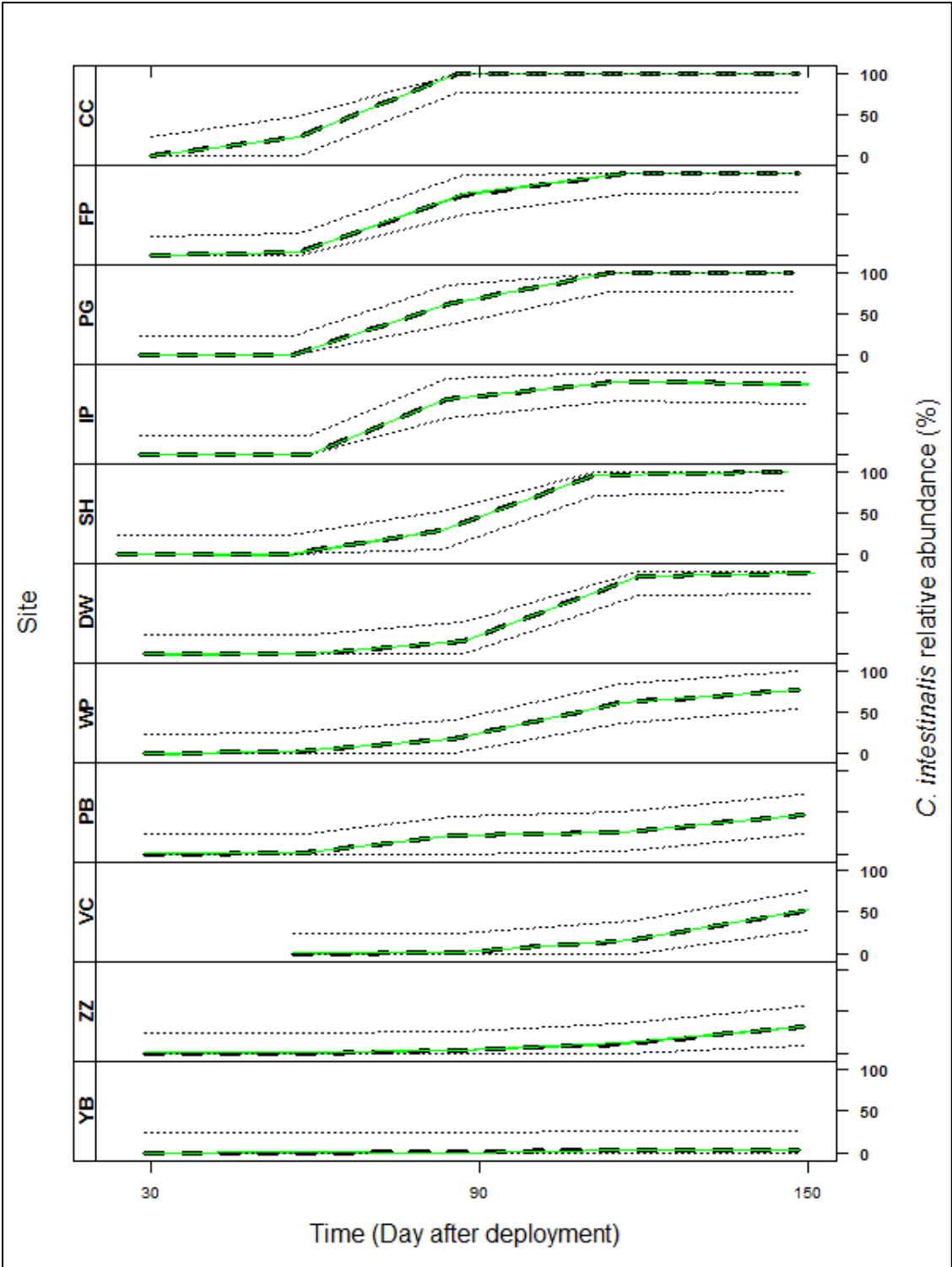
#### 3.6.3.2. Best top-down (BTD) 2015 model

The full model fit for the ‘best’ top-down model (BTD) is very good, with almost complete parity with the average observed growth, except for a slight deviation at PG (Fig. 3.20).

There are some noticeable differences between the species-wide effects fit in the BBU model (Fig. 3.19) and the BTD model (Fig. 3.20). The rate of growth and the maximum population size at CC according to the species-wide effects fit was lower in the BTD, leading to a poorer species-wide effect fit. Conversely, the maximum population predicted for YB by the BTD was much lower compared to the BBU, but this led to a much better species-wide effect fit. FP and PG have very similar species-wide effect fitting in both BBU and BTD, except for a slight lag in the time it takes for the FP predictions to come on par with the observed values in the BTD. IP, SH, DW, and WP all have very good species-wide effect fitting in the BTD, arguably all better than the BBU, with the exception of WP, that is slightly worse. The BTD species-wide effect fit for PB is obviously worse, while VC and ZZ are about the same in both models. The common trend in both BBU and BTD is that the sites that are in the mid-range of *C. intestinalis* abundance, having high maximum population size but taking slightly longer to reach that point than the highest abundance site, are all fit well by the species-wide effect. It is the sites that are the highest and lowest abundance that require the site-level effect to account for the variation not unexplained by the abiotic variables. There is a high level of confidence in the model fit from the BTD 2015 (Fig. 3.21).

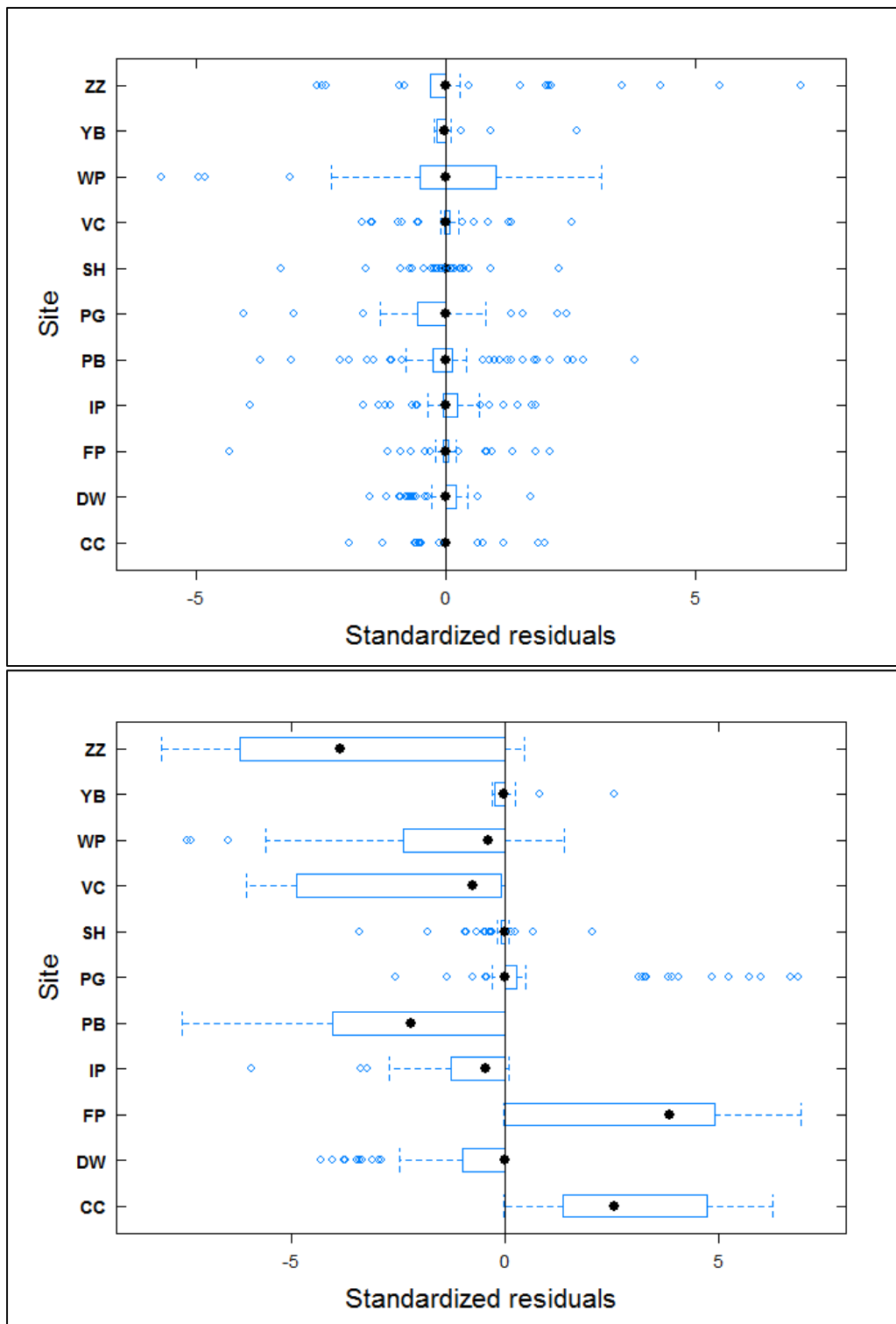


**Figure 3.20: Full-model fitted *C. intestinalis* relative abundance (black dashed line) for the 11 sites, fit by the 2015 best top-down model (BTD), and the fitted values using species-wide effect only (black dotted line). The observed *C. intestinalis* values (red line and circle markers) are overlaid.**



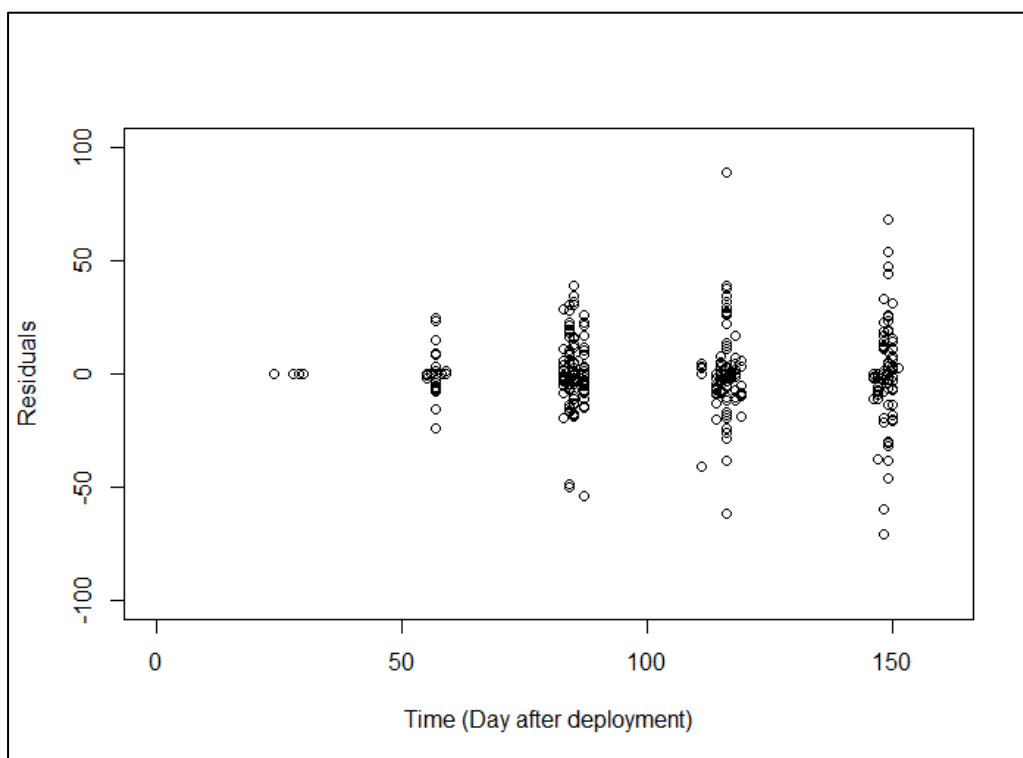
**Figure 3.21: Full-fit by the 2015 BTD (solid green line), with 95 % confidence intervals (dashed black lines), and 95 % prediction intervals (dotted black lines).**

### 3.6.3.3. Residual plots: BTD 2015 model



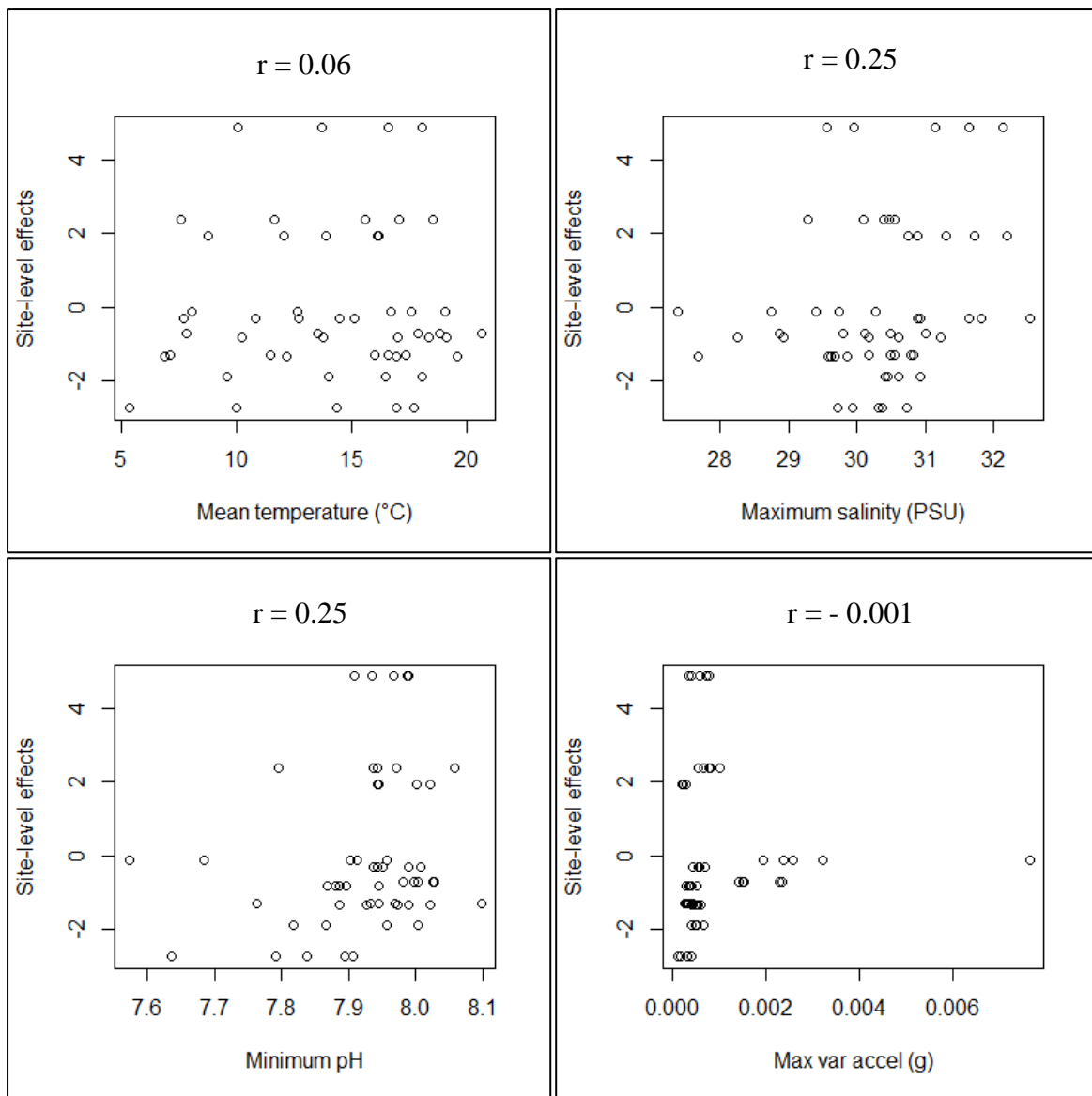
**Figure 3.22: BTD 2015 residual boxplots by site, ordered alphabetically. The full-model residuals are displayed in the upper plot, while the species-wide effects residuals in isolation are displayed in the lower plot.**

On the whole, the residuals from the full-fit BTD in 2015 are good, not seeming to violate any assumptions egregiously (Fig. 3.22). The means of the residuals at the site-level are all centred near zero, while their variance around the mean is homogeneously distributed, with some exceptions. There are some outlying residuals; the most obvious of which can be seen at ZZ (positive outlier) and WP (negative outliers). The residuals from the species-wide effects of the BTD in 2015, when inspected in isolation, appear to violate a number of assumptions. Bias (heterogeneity of variance around the mean) is introduced into the residuals, while the means deviate from zero, substantially in some cases, at the site-level (Fig. 3.22).



**Figure 3.23: The full-model residuals for the BTD 2015 plotted against days after deployment.**

Heteroskedasticity is apparent in the residuals of the full-fit for the BTD in 2015 (Fig. 3.23). While this is a violation of model assumptions, there is no bias in the residuals and heteroscedasticity is a very common feature in growth models where the observed values are increasing rapidly through time. When a covariance and correlation structure was introduced into the model to account for variability between and within individual settlement plates through time, it did not improve the model or the heteroscedasticity, so the more parsimonious model without an autocorrelation structure was used. There are no relationships between site-level effects and the independent variables in the BTD 2015 (Fig. 3.24).

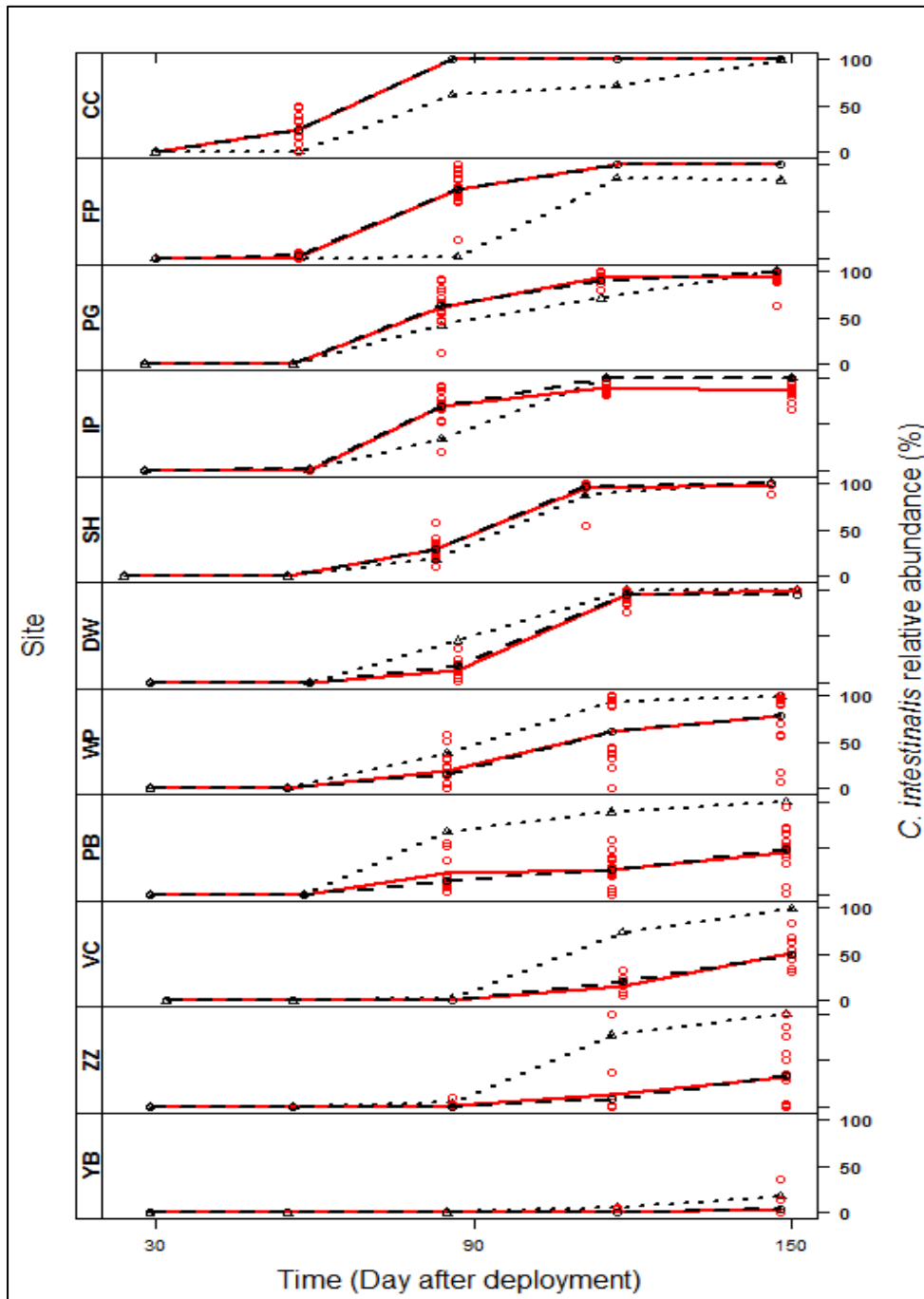


**Figure 3.24: Predictor variables used in the BTD 2015 plotted against the site-level effects to inspect their relationships, with the correlation coefficient presented at the top of each plot.**

#### 3.6.3.4. Best temperature-only model (BTO) 2015

The best temperature-only model (BTO) in 2015 contained the mean temperature metric in the MaxPop, D50, and Lag model parameters, while MaxPop was specified as a site-level effect in addition to all model parameters being species-wide effects by default. The *C. intestinalis* values fit by the full BTO model are almost all a direct fit with the mean *C. intestinalis* percent cover observed (Fig. 3.25). The proportion of the fit associated with species-wide effect and site-level effects are clearly visible, with a wide degree of inter-site variation. The visual inspection of the fitted vs observed values for the BTO 2015 indicates

that a fair portion of the observed growth is fit well by the model. The observed growth at five sites (PG, IP, SH, DW, WP, and YB) is matched very closely by the fitted values (Fig. 3.25). The fitted values for CC and FP are not quite as good, but they do improve nearer the end of the monitoring period. The fitted values at PB, VC, and ZZ start well, before deviating substantially from the observed growth.



**Figure 3.25:** Full-model fitted *C. intestinalis* relative abundance (black dashed line) for the 11 sites, fit by the 2015 best temperature-only model (BTO), and the fitted values using species-wide effect only (black dotted line). The observed *C. intestinalis* values (red line and circle markers) are overlaid.



### 3.7. Model accuracy

Models developed using data from the same year, regardless of whether they are fit by BU, TD, or another approach (BTO etc.), tend to have very similar error rates and predictive power when considering the full (species-wide effects and site-level effects) model fit (Table 3.14). However, model performance can vary substantially when comparing the species-wide effects alone (Table 3.14). Inspecting the full-fit of models from 2014 and 2015, the lowest error rate (MAE and RMSE) and joint-highest modelling efficiency (MEF) is provided by the best temperature-only model (BTO) from 2014, although the Spearman rank correlation is slightly lower than the BBU/BTD/BSO 2014 model.

**Table 3.14: Model accuracy results for the best bottom-up (BBU) and best top-down (BTD) models for 2014 and 2015. The four methods of model assessment are: mean absolute error (MAE), measured in units of *C. intestinalis* %; root mean square error (RMSE), measured in units of *C. intestinalis* %; model efficiency (MEF), a measure of the variance explained by the model [similar to R<sup>2</sup> (Amaro et al., 1998)]; and the Spearman rank correlation coefficient ( $\rho$ ). Full = The model with both species-wide and site-level effects. SW = Species-wide effects only. BTO = Best temperature-only model. BSO = Best salinity-only model. Best BU & TD 2014 model is also the BSO for 2014.**

Model	Year	MAE	MAE	RMSE	RMSE	MEF	MEF	$\rho$	$\rho$
		Full	SW	Full	SW	Full	SW	Full	SW
Null <sup>a</sup>	2014	5.95	34.41	10.93	42.18	0.94	0.06	0.88	0.35
BTO <sup>a</sup>	2014	3.85	26.16	8.86	37.02	0.95	0.16	0.89	0.56
Best BU & TD <sup>a</sup>	2014	4.86	26.33	9.97	35.66	0.95	0.33	0.9	0.6
Null <sup>a</sup>	2015	6.57	19.87	13.08	34.38	0.90	0.34	0.90	0.72
Null <sup>b</sup>	2015	6.00	20.61	12.44	31.97	0.91	0.43	0.92	0.72
BTO <sup>a</sup>	2015	5.92	16.00	12.52	28.63	0.91	0.54	0.91	0.81
BSO <sup>b</sup>	2015	5.88	20.46	12.33	32.17	0.91	0.42	0.92	0.71
BBU <sup>b</sup>	2015	5.88	20.16	12.278	32.24	0.92	0.42	0.92	0.73
BTD <sup>a</sup>	2015	5.83	20.08	12.32	34.17	0.91	0.34	0.92	0.77
BTD <sup>a*</sup>	2015	5.81	18.2	12.31	31.27	0.91	0.45	0.92	0.78

\* BTD model in 2015 developed using MaxpH as the metric for pH variable, instead of MinpH. This was a serendipitous finding as a part of model building exploration, due to a lack of model convergence using MinpH.

<sup>a</sup> Models that contain a site-level effect for MaxPop parameter only.

<sup>b</sup> Models that contain a site-level effect for the MaxPop and D50 parameters.

Upon closer inspection of the species-wide effects fit by the 2014 BTO, the MEF drops noticeably compared to the BBU/BTD/BSO model in 2014, which explains ~ twice the amount variability within the 2014 dataset.

The variation in *C. intestinalis* relative abundance explained by the best 2014 model, which includes the mean salinity metric in all three model parameters, is a marked improvement from the variation explained by a basic growth model devoid of any abiotic information, depicted by the 2014 null model. The null model in 2014 has a very low MEF, only explaining about 6 % of the observed variability in *C. intestinalis* relative abundance. The best 2014 model improved this basic model MEF by ~ 25 %, suggesting salinity explains about a quarter of the variability in the observed data in 2014. In contrast to 2014, a basic growth model in 2015 can explain between 34 – 43 % of the observed variation in *C. intestinalis* relative abundance, depending on the site-level effects specification (2015: Null<sup>a</sup> and Null<sup>b</sup> - Table 3.14). The addition of abiotic metrics and the resultant best models from the BU and TD approaches in 2015 did not improve the MEF of the models, suggesting that associations between abiotic metrics and *C. intestinalis* relative abundance were not as strong in 2015 when compared to 2014. Models that only contained temperature metrics did not rank well in the BU and TD approaches in 2015, according to AICc, (Table 3.11 & 3.12), but the BTO in 2015 has the highest MEF when considering the species-wide effects alone. Despite the strong association between mean salinity and *C. intestinalis* relative abundance in 2014, the BSO in 2015 did not dominate the candidate sets of models developed in either the BU or TD approach (Table 3.7 & 3.8 for 2014 compared with Table 3.11 & 3.12 for 2015), while the species-wide effects associated with the BSO 2015 were very similar to a null model with no abiotic information (Table 3.14).

### **3.8. Nested cross-validation**

All of the model assessment methods yielded very similar results in the nested cross-validation tests, when compared to the results from the full-model accuracy results (Table 3.14 & 3.15). This indicates strongly that the models have a high predictive power when used to predict *C. intestinalis* abundance based on a portion of the intra-annual or combined dataset that was excluded from the model fitting process.

**Table 3.15: Model fit assessment results from the nested cross-validation procedure, as measured by MAE, RMSE, MEF, and  $\rho$ .**

Model	Year	MAE (%)	RMSE (%)	MEF	$\rho$
BTO	2014	3.85	8.51	0.95	0.88
BBU/BTD/BSO	2014	4.86	9.53	0.94	0.9
BTO	2015	5.92	12.26	0.91	0.91
BSO	2015	5.87	12.13	0.91	0.92
BBU	2015	5.88	12.03	0.91	0.92
BTD	2015	5.83	12.08	0.91	0.92
BTD *	2015	5.81	12.06	0.91	0.92

### 3.9. Inter-annual model predictions

All models, be they fit on 2014 or 2015 data performed relatively well when predicting the combined dataset (Table 3.16). This would be expected, as the nested-cross validation showed that the models are very good at predicting the intra-annual variation (Table 3.15), and the combined dataset contains the data that were used to fit the given model carrying out the predictions, as well as the additional ‘new to the model’ data from the other year that was not involved in the model fitting process. The predictive power of models varies much more when they are used to predict the observed values from the other year in isolation, with some very poor predictions when the 2015 models were predicting 2014 data, in particular. The 2014 BTO was able to predict about 30 % of the variation in the 2015 data (Table 3.16), while it could predict almost 50 % of variation in the combined dataset. There are probably two major contributing factors that led to the very poor (high error rates and negative MEF) predictive power of the 2015 BBU and BTD predicting 2014 data. First, it is likely as a result of the differing inter-annual abiotic trends, where the role of salinity in fitting 2014 models, as well as being evident upon visual inspection of the abiotic plot alongside *C. intestinalis* abundance (Fig. 3.7 and Fig. 3.8), was more important than in 2015. Given that the 2015 models are very accurate based on the data they were fit with, they do not perform well when predicting a set of data with a different trend. While this is most likely a major player in the poor prediction, it was almost certainly exasperated by the small number of data in 2014 for which the 2015 models could predict. This small dataset could lead to a high frequency of erroneous predictions, especially due to the absence of data from the beginning of the monitoring period where overall variability in observed *C. intestinalis* values was lower. This

is highlighted by the fact that when the 2014 models were used to predict the 2015 data (a substantially larger dataset, with more measurement points for each site-level and the start of the monitoring period), they perform much better predicting ~30 % of variability in these data, despite the variable inter-annual trends being proportional for each respective year. The most parsimonious model in 2015, the BTO, provided better inter-annual prediction than the 2015 BBU or BTM (Table 3.16).

**Table 3.16: Assessment of the results from model predictions carried out on ‘test’ datasets that were wholly or partially not included in the initial model fitting.**

<b>Model</b>	<b>Test data</b>	<b>MAE (%)</b>	<b>RMSE (%)</b>	<b>MEF</b>	<b>S rho</b>
BTO 2014	2015	18.57	35.1	0.31	0.79
BTO 2014	Combined	15.23	31.13	0.47	0.8
BBU/BTD/BSO 2014	2015	21.4	35.67	0.27	0.67
BBU/BTD/BSO 2014	Combined	17.4	31.70	0.45	0.71
BTO 2015	2014	19.11	36.46	0.18	0.73
BTO 2015	Combined	11.11	24.73	0.64	0.84
BSO 2015	2014	25.40	39.68	0.10	0.64
BSO 2015	Combined	12.37	24.97	0.65	0.83
BBU 2015	2014	32.96	48.49	- 0.22	0.14
BBU 2015	Combined	11.92	25.31	0.65	0.78
BTD 2015	2014	36.10	53.65	- 0.50	0.40
BTD 2015	Combined	12.56	27.54	0.58	0.80
BTD 2015 *	2014	33.47	49.59	- 0.28	0.42
BTD 2015 *	Combined	11.96	25.79	0.63	0.81

## 4. Discussion

### 4.1. Summary of findings

I carried out *in situ* monitoring of four abiotic variables in the nearshore marine environment, in tandem with biological monitoring of *C. intestinalis* population growth, across the entire Atlantic shore of NS. The scale of this sampling regime was novel in regard to the high-resolution *in situ* abiotic recordings that complement the invasive tunicate monitoring. There was an association between *C. intestinalis* growth and salinity in 2014: this same trend was not evident in 2015. With a more comprehensive dataset in 2015, the best models, ranked via their AICc value, include a metric from each of the four abiotic variables in at least one model parameter. This could suggest that they all contribute useful information to explain the variability in *C. intestinalis* growth. However, similar model accuracy assessment results between the null models and the BBU and BTD in 2015 indicate that the lack of any clear associations between the abiotic variables and *C. intestinalis* growth is a more probable cause for all abiotic variables being included in the best models. The species-wide effects from the BTD\* 2015 accounted for approximately half of the total variability observed in *C. intestinalis* growth (~ 45 % MEF, Table 3.14). Despite this, the null models from 2015 have MEF values of 0.34 and 0.43, depending on the parameters assigned as site-level effects (Table 3.14). This suggests that in 2015, a basic growth model with no abiotic input can account for between 34 – 43 % of the variability in the observed *C. intestinalis* response. Including abiotic information into the model does improve AICc substantially (Table 3.11 & 3.12), but does not improve the model accuracy assessment results greatly (Table 3.14). This suggests that despite the lack of strong associations between abiotic conditions and growth, the information does improve model fitting. This is in contrast to the trends observed in 2014, where the null model explained very little variability in the observed *C. intestinalis* growth, but there was a strong association with salinity that improved model fitting and variability explained (Table 3.7, 3.8 & 3.14). On the whole, the population response across the different sites displayed a more consistent pattern in 2015, being fit much better than 2014 with a basic ‘null model’ containing no abiotic information. However, when the response is broken down on a site by site basis, the pattern exhibited was remarkably variable, as can be seen in the observed vs. fitted plots (Fig. 3.13 & Fig. 3.20). This inter-annual variation suggests that none of the monitored abiotic variables can be excluded from future discussions about the potential abiotic contributors to *C. intestinalis* population growth, but salinity and

temperature do appear to be the most prominent, despite being modest and changeable in their predictive power. These results provide novel insight into the proportion of population variability that can be explained by these four abiotic variables. This exploratory study offers theoretical backing to pursue a variety of hypotheses that could further explain observed variability in *C. intestinalis* growth.

## **4.2. Abiotic determination of population growth**

### *4.2.1. Temperature*

Temperature appears to have an effect on *C. intestinalis* population growth (Table 3.16), with its absence in the 2014 best model possibly due to data exclusion from the model as a result of the missing data for other variables. In general, the warmer the water is at a site, the more likely *C. intestinalis* is to grow well. Temperature influences physiological processes and this general effect is universal (Serafini et al., 2011). There are many ways that warmer temperatures could enhance physiological processes, resulting in increased *C. intestinalis* growth, such as earlier onset of gametogenesis, earlier spawning and recruitment, and faster growth rate, to name a few (Ramsay, 2009). My results support the existing field (Carver et al., 2003; Howes et al., 2007; Ramsay, 2009; Sephton et al., 2011) and laboratory (Vercaemer et al., 2011) studies based in Atlantic Canada that describe the importance of temperature on *C. intestinalis* growth.

### *4.2.2. Salinity*

Salinity is well touted in the literature as a factor that explains tunicate distribution (Auker and Oviatt, 2008; Epelbaum et al., 2009; Lambert and Lambert, 1998), *C. intestinalis* included (Therriault and Herborg, 2008a; Vercaemer et al., 2011). This makes sense, as ascidians are largely considered to be stenohaline osmotic conformers and are largely immobile, so they cannot move when abiotic conditions fluctuate and become unfavourable (Shumway, 1978). In 2014 this held true, as there was a strong relationship between salinity and *C. intestinalis* abundance (Table 3.7. and 3.8.). The tolerance pattern described by Vercaemer et al. (2011) was evident, with lower salinity sites tending to have the lowest *C. intestinalis* abundance. For the periods where data were available for all abiotic metrics in 2014, salinity explained the most variation in the maximum population size and the rate of growth. In 2015, the association between salinity and *C. intestinalis* growth changed, with a negative relationship between *C. intestinalis* growth and increasing salinity (data not shown:

coefficients from models with a single salinity metric only describe this association). Compared to 2014 when it was the sole variable in the best model, it appears as if salinity was not as influential in explaining the observed variability in 2015, but still provided useful information for the best model. The difference in the association between salinity and *C. intestinalis* between the years could be due to factors that were not measured that influenced the relationship between salinity and *C. intestinalis* growth, causing it to change, or the relationships in both years could be caused by unknown factors that changed but happen to be correlated with salinity. The timing and duration of low salinity periods may also have a bearing on the overall impact of salinity on a given population of *C. intestinalis*. Exposure to water ~ 20 ppt caused the siphons of individuals in a laboratory experiment to close tightly, preventing any fluid clearance, thereby leading to no oxygen or food intake (Shumway, 1978). The individuals used in this experiment were collected from the mid-west coast of the UK, so it is more likely than not that they were *C. intestinalis*, as the region of sympatry with *C. robusta* has not been shown to extend further north than the south coast of England. The availability of food and oxygen prior to low salinity conditions and the duration of low salinity, hence influencing the length of time the individuals closed their siphons, could influence the affect that low salinity has on overall respiration, growth, and reproduction. Short periods of low salinity due to sporadic freshwater runoff from heavy precipitation may not have a negative effect if the individuals exposed could tightly close their siphons and cease intake of water for the period. Conversely, a sustained lower salinity could hinder individuals more so. The timing of extremes for any abiotic variable is important in regard to the life-stage of the majority of a population. To explain that, take the example of pulse event of low salinity. This, or an extreme high or low of any abiotic variable, may have a greater impact on the pelagic larval stage of *C. intestinalis* (Pineda et al., 2012a). If a period of low salinity were to coincide with a peak spawning event, this could have serious implications for the population through reduced fertilization, settlement and recruitment, especially if the parents had been subjected to relatively high salinity conditions until then (Renborg et al., 2013).

#### 4.2.3. Water flow

My results suggest that in 2014, water flow data did not contribute useful information to the best model, while in 2015 it was included in the BTD, suggesting the information improved model fitting. It is possible that increased water flow conditions affect fertilization, larval behaviour and ability to swim and settle, as well as physiologically stressing adult

individuals. *C. intestinalis* tends to frequent sheltered habitats (Petersen and Svane, 1995), not being known to tolerate wave exposure (Howes et al., 2007).

#### 4.2.4. pH

Lower pH is potentially a beneficial factor for *C. intestinalis* (Dupont et al., 2009, 2008). While the inclusion of pH was not possible in 2014 due to the missing data and small sample to fit models with, in 2015 it was a contributing variable in all parameters of the BTM. This suggests that pH is associated with the population size and growth rate of *C. intestinalis*. It is difficult to discern a strong relationship from a purely visual inspection of abiotic and abundance plots in the results (Fig. 3.9 and 3.10), but the results from the models which include only a single maximum pH metric in a single model parameter suggest that pH does have a negative association with *C. intestinalis* growth (data not shown). These results agree with the minimal information available from the literature (Dupont et al., 2009, 2008). The inability to model pH adequately in 2014 combined with sparse available information preceding my study makes it all the more difficult to infer relationships in regard to pH. But, given the suggestion in the literature that lower pH is beneficial for *C. intestinalis* larval development in combination with its contribution to the best model in 2015 mean that it cannot be ruled out as having an effect on *C. intestinalis* abundance. If lower pH does indeed improve the growth and abundance of *C. intestinalis*, this is not good news as lower pH could negatively affect the growth of mussels (Bechmann et al., 2011), while it is largely considered to be detrimental for the majority of marine invertebrates, especially calcifying organisms (Dupont et al., 2009). This is an abiotic relationship that is ripe for further study, as is discussed later in section 4.6.3.

### 4.3. Modelling approach

#### 4.3.1. Mixed-effects use in marine modelling

Most mixed-effects modelling research based in the marine environment uses morphological measures of the species in question at the individual organism level, occasionally with some form of environmental model component, particularly fisheries modelling where individual size is of particular interest (Averbuj et al., 2015; Escati-Peñaloza et al., 2010; Hernández-Otero et al., 2014). In these cases, the random effects are introduced into the models to estimate the degree of variability within and between individuals for a variety of spatial and temporal factors. Others use a combination of individual morphological measurements in



tandem with abiotic variables in the mixed-effects models. I am yet to find a marine based comparison example to this study, where abiotic variables are exclusively used to predict parameter estimates for a relative population growth measure at the species-wide (fixed effect) level, as well as at a site-level (random effect). It also seems to be quite rare to try and ascertain a predictive growth model for a species with a short life-span and a rather stochastic growth pattern. Most studies that are similar in their design measure growth over a much longer period, like in forestry research where the use of mixed-effects modelling is widespread (Adame et al., 2008b; Amaro et al., 1998; Calegario et al., 2005; Subedi and Sharma, 2011).

#### 4.3.2. Was NLME appropriate?

I managed to produce models within each year that were very accurate (Table 3.15 and 3.16). Using the BU and TD model development approaches, a sizeable component of the model fit (33 % in 2014 and 45 % in 2015) was associated with species-wide effects. The use of site-level effects in addition to the species-wide effects assisted in estimating the unexplained variation observed, improving the modelling process substantially compared to species-wide effects alone (Table 3.14). Confidence in the model parameter estimates is high, as shown in the confidence and prediction interval plots (Fig. 3.14 & 3.21). Not only did models have good explanatory power via the species-wide and site-level effects when fit to the whole dataset, they were robust under nested cross-validation tests too (Table 3.16), as well as displaying a relatively good measure of predictive power in some cases (Table 3.17). One apparent downside of the AICc ranking used to assess the best models is the lack of an easy assessment of the model fit attributable to the species-wide effects. I was interested in the ability of abiotic variables as predictors of *C. intestinalis* growth, so while the NLME model provided well-fit models, it does not have an inbuilt method to favour models that are more reliant on species-wide effects, rather than the site-level effects. Selecting models with a higher proportion of species-wide effects in the fit was desirable, hence why the additional model accuracy assessments, nested cross-validation, and inter annual predictions were important.

#### 4.3.3. Bottom-up versus top-down approach

Overall, bottom-up and top-down modelling approaches provided relatively similar ‘best’ models. In 2014, both approaches provided an identical conclusion, that the best model contained only salinity. In 2015, the TD approach suggested that the global model, minus one

water acceleration metric, was the best model, while the best model resulting from the BU approach was a more parsimonious (but less likely based on AICc) model without any temperature metrics at all. While determining a set of candidate models (hypotheses) established *a priori* before any model fitting is best practice, it is not always feasible or even entirely sensible for some exploratory studies. In a sense, both BU and TD approaches are theoretical in nature, as the abiotic variables monitored were carefully selected and the preselection of metrics limits potential overfitting with collinear components. The TD approach also ensures that no combination of abiotic metrics was overlooked due to human error. For these reasons, in addition to the exploratory nature of this study, the combined BU and TD approach used appears to be justified. In this case there was a relatively high degree of complexity when fitting the NLME model (difficulty in determining starting values for parameter estimates and cases of models not converging) so this dual-approach provided a good solution to test a range of candidate models for years in isolation, as well as for the combined dataset (see Appendix E).

#### *4.3.4. Model reliability: inference versus prediction*

The fitted models allow us to infer that a substantial amount of observed variability in *C. intestinalis* growth can be explained by the null growth model equation, the observed abiotic variables, or via the incorporation of a measure of the unexplained variability via site-level effects. A major issue for consideration is whether or not these inferences actually lead to accurate predictions (Shmueli, 2010). The results suggest that until further data is collected to ascertain the long-term variation that occurs in *C. intestinalis* growth in response to the abiotic variables like salinity, the best models at explaining variation in these data are not necessarily the best model for predictive purposes (Table 3.17). With only two years of observed data currently available, it is impossible to say whether this variation is going to continue to fluctuate, or whether 2014 or 2015 was an aberration. Over time, the ‘true’ trend could emerge from the currently variable trends observed for *C. intestinalis*’ response to abiotic conditions. This could mean that with the addition of more information, the best models incorporating all four abiotic variables could produce quite accurate predictions consistently.

#### *4.3.5. Sources of potential experimental error*

The exclusive inclusion of salinity in the best 2014 model is most likely due to the quantity of missing data that led to large portions of the temperature data at the start of the monitoring

period to be excluded from model fitting. The late spring/early summer temperature is correlated to the subsequent growth rate of *C. intestinalis* (data not shown), so the necessity to remove these data from the 2014 model to allow the inclusion of salinity, pH and water flow variables is likely to have influenced the final model.

The predation pressure for any given site is an unknown factor in this study and the potential effects of which are discussed later in section 4.4. An area where this unknown factor may have been introduced error or imbalance to the dataset due to the experimental design is if increased predation was experienced by *C. intestinalis* populations at sites where the water column was shallow and the bottom of collectors (rope or weight) was resting on the bottom of the harbour during a portion of low tide. Likely primary predators of the adult *C. intestinalis* individuals, such as crabs and sea-stars may have had easier access to their prey when the bottom of the collectors were touching the benthic substrate, even if only during the lowest tidal period. This potential source of error could be reduced in future studies by ensuring deployment takes place when the tide is at its lowest point for all sites and reducing the depth at which collectors are submerged. If the water column is too shallow, then submerging the collectors at a very shallow depth could be introducing a new source of potential error, so at that point if both of those are a risk, a new site may need to be considered. Alternatively, a different settlement collector design could be used, such that there is only one plate per collector, thus reducing the overall depth required and instead increasing the number of single-plate collectors as compensation.

There are also some sources of abiotic and population variability that were purely anthropogenic in nature that may have influenced my observations. During the spring of 2015 there was substantial infrastructural work carried out at Port Bickerton that involved some dredging. There is some evidence that turbidity can influence *C. intestinalis* (McLaughlin et al., 2013), so the possibility of dredging creating turbid conditions beyond those naturally occurring at that site are unknown. The growth in 2015 at Port Bickerton was reduced compared to 2014, but there is no way to know if that is related to a potential turbidity effect, as I didn't monitor that. Another source of potential anthropogenic influence occurred at Indian Point, where large quantities of adults were removed due to the removal of biofouled mussel farm long-lines before the 2015 monitoring period. This is important as the propagule pressure may have been reduced at this site in 2015 as a result, with the potential to effect growth. *C. intestinalis* growth was lower on the settlement plates in 2015 at Indian Point, as well as the overall biomass being lower (pers. obs.).

#### 4.4. Inter-site and inter-annual stochasticity

My results show that all four of the abiotic variables measured can contribute to explaining the variability observed in *C. intestinalis* growth. In addition, temperature, salinity, and water flow (wave exposure) are regarded in the literature as primary abiotic variables that determine the distribution and growth of *C. intestinalis*, while pH has been suggested as playing a role in larval physiology too. Despite this, my results also show that there is clearly much more going on than just abiotic control of *C. intestinalis* population growth.

Determining the influence of abiotic variables on the success of invasive species in the marine environment can be very difficult (Dijkstra et al., 2008). There are numerous factors in addition to abiotic conditions, both biotic and anthropogenic in nature, that could influence the establishment, growth, survival, reproduction, and range expansion of invasive species. In the context of *C. intestinalis* growth, the genetic diversity within and between sites could influence the variable response to abiotic conditions, both between sites within a single year, as well as influencing inter-annual variability in growth response to abiotic conditions. Biotic interactions, both resistant and facilitative in nature (predation and competition), could also cause inter-site and inter-annual fluctuations in the *C. intestinalis* growth response to abiotic conditions (Lockwood et al., 2013). There are examples where the abiotic variables I recorded explain, through species-wide effects, almost all of the growth variability at a specific site (Fig. 3.19, IP, DW, and SH), while at other sites the observed growth is not influenced by the abiotic conditions to any great degree (largely fit by site-level effects). There did not appear to be an obvious or consistent geographic trend (that could indicate a regional abiotic or biotic effect not monitored in the study) in the response of *C. intestinalis* growth.

##### 4.4.1. Genetic diversity

The heterogeneous influence of abiotic conditions on *C. intestinalis* growth could be due to a host of factors, acting in isolation or in combination. The phenotypic response to any given environmental exposure is due to the molecular make-up of a given organism (Serafini et al., 2011). Therefore, genetic diversity resulting in variable gene expression when individuals are exposed to the same abiotic conditions could result in different phenotypic trends between sites and inter-annually at the same site. There could be substantial genetic diversity on the provincial scale, but also within sites, with a combination of factors determining the genetic composition of introduced populations. Diversity on the macro-provincial scale is largely due

to anthropogenic vector transport (Zhan et al., 2012). On the meso- and micro-scale, the proximity of populations to each other is generally a good indicator of the genetic relationship, within the benthic marine community (López-Legentil et al., 2014). But the level of connectivity between local sites may vary from region to region depending on geographic barriers and hydrodynamic conditions affecting dispersal, as well as vector movement through recreational boating and aquaculture trade. Genetic diversity within a site or population could be influenced intra- or inter-annually by multiple introductions from different source populations, as well as being affected by the size of the introductions (propagule pressure). There is no easy way to tell how related a given population is to another in the context of this study without carrying out genetic sequencing. To survive and thrive in a new environment, you can either be well adapted to the introduced environment from the off, or you can have the mechanisms to tolerate or adapt to stress. These two options are most likely represented by gradients of response. Some of the ways this dynamic could play out for *C. intestinalis* in NS are discussed next.

#### 4.4.1.1. Suitability to the environment

Given the wide tolerance range that *C. intestinalis* is reputed to possess, it means that it probably find a relatively large number of areas it is introduced to as suitable environments, compared to other less-tolerant species. *C. intestinalis* also has the ability to rapidly mature and become reproductively viable, while also possessing the ability to produce multiple generations per year (Lambert and Lambert, 1998). Hence if a sizeable introduction of individuals occurs, a portion of which are well suited to the new environment, and if there are no other major contributing factors, then growth is likely to increase quickly. The diversity of introduced individuals is described by Lockwood et al. (2013) as the “genetic lottery”, and it can influence establishment success, as well being a key determinant in the rate of population expansion. If an introduction were to occur and the collection of individuals as a whole are not that well suited to the environment, due to low numbers or genetic homogeneity, establishment is less likely, while the rate of growth of that population if it does become established will be subsequently reduced. This may result in a lag in population growth for some time, perhaps lasting months or years depending on other contributing abiotic and biotic factors. An example where this may have happened would be in 2014 at Dingwall and Wedgeport, where the low salinity tolerant individuals survived and reproduced the most, leading to an overwintering population that were tolerant to low salinity. While the population was small in 2014, the resident population could have been better suited to the

conditions in 2015. This in turn could have resulted in a population capable of enhanced growth and reproduction under similarly low salinity conditions in 2015. This type of prominent and rapid natural selection initially after a population is introduced is thought to be an important dynamic in invasion ecology (Lavergne and Molofsky, 2007). If this type of selection did occur, the extent to which it did could have been variable from site to site, offering one possible explanation for variability in observed *C. intestinalis* growth not accounted for by the monitored abiotic variables.

The abiotic conditions that parents live in affect the survival of *C. intestinalis* offspring (Renborg et al., 2013). Individuals that survived in lower salinity conditions in 2014 may have produced offspring that were able to tolerate lower levels of salinity in 2015, as opposed to the individuals in areas of relatively higher salinity in 2014, resulting in offspring not as well suited to the lower salinity conditions in 2015. It is also possible that the source population that produced the larvae that settled on the plates I monitored were relatively far away (natural dispersal) and were subjected to different abiotic conditions to those present at the monitored site inhabited by their progeny. Given the findings of Renborg et al. (2013), this could suggest that larvae would not do as well in these unfamiliar conditions. The extent to which this could occur would depend on site specific abiotic conditions as well as the genetic diversity of the populations.

#### 4.4.1.2. Response to stress

The balance between stress management and energy efficient growth could also be important in the phenotypic variation observed amongst populations of *C. intestinalis* (López-Maury et al., 2008). An analysis focused primarily on the genetic reasoning behind variable temperature tolerance could help to explain the potential *C. intestinalis* has to respond to a range of abiotic stresses. Although the study was most likely conducted on *C. robusta*, as they collected samples from California, Serafini et al. (2011) carried out molecular examination of individuals from two members of the *Ciona* genus after manipulative temperature treatments. They found that *C. intestinalis* (most likely *C. robusta*) maintained higher levels of molecular chaperones compared to another species *C. savignyi*. They suggest this is a possible reason why *C. intestinalis* (although distinct to *C. robusta* they are the closest relatives within the *Ciona* genus) seems to have a good capacity to respond quickly to heat stress. This high steady-state level of chaperones could also have implications for the response to other physiological stress, not just temperature. The ability to respond to stress through molecular mechanisms could vary between populations of *C. intestinalis*, depending

on their molecular differences. While temperatures in NS would not appear to often reach a point where it would be subjecting *C. intestinalis* individuals to upper thermal limits, adaptation to a colder environment, hence resulting in differing ability to commence reproduction and growth earlier in the year could be an important factor. The molecular flexibility that *C. robusta* has been shown to possess (Serafini et al., 2011) is likely to be shared to some extent by *C. intestinalis* as well, providing a potential mechanism for *C. intestinalis*' ability to tolerate a wide range of abiotic conditions that may induce degrees of stress. If an individual is able to account for stress more easily than others, it may result in an ability to allocate greater resources to growth and reproduction. If the molecular make-up of populations at different sites observed here in NS is different, it could provide an explanation for some of the variability in *C. intestinalis* growth not explained by the observed abiotic variables. Likewise, if two sites have quite a similar molecular composition, then this could be an explanation for consistency in response, either explained or not explained by the observed abiotic variables.

#### 4.4.2. Biotic interactions

##### 4.4.2.1. Predation

It is widely documented that *C. intestinalis* (Petersen and Svane, 1995) and the closely related *C. robusta* (Caputi et al., 2015) can experience quite abrupt population decline. It has been suggested that a long-term cyclical recruitment phenomenon underlies the pattern, but what regulates it is still uncertain. In the native range of *C. intestinalis* there are observational data that describe a shared seven year cyclical pattern of *C. intestinalis* and its dominant predators (Lundälv and Christie, 1987). Much like the classic cyclical population pattern exhibited by snowshoe hare and Canadian lynx (Elton, 1942), it's possible that *C. intestinalis* exhibits a similar type of coupled boom-bust cyclical pattern in tandem with native predators. Lundälv and Christie (1987) do note however that there were local deviations from the overall common cyclical pattern displayed by predators and prey, while they suggest that fluctuations of hydrodynamics conditions in tandem with large-scale climatic events could contribute to the pattern. Cyclical predator-prey relationships that we are unaware of in NS may have substantial effects dependent upon the peaks and troughs of both the predator and prey population abundance. Given the dynamic nature of introduced species' relationship with predators in the introduced range, this could be a highly variable pattern if present. Predation of larvae by fishes is also a suggested mechanism of spatial limitation, due to suppressed

dispersal, in the native range (Petersen and Svane, 1995). If variable predation pressure is occurring in different areas of the province, this in turn could affect the population connectivity on a local scale (through dispersal limitation), which ties into the discussion about genetic diversity in section 4.4.1. The occurrence of *C. intestinalis* predation to some degree in NS is almost a certainty. The feeding trials outlined in the introduction (section 1.3.4.) describe the potential predators in question, but the extent to which native predators prey upon *C. intestinalis* compared to native prey is not clear, especially in a field setting. Explicit testing of the enemy release hypothesis (ERH) is required in the future to be able to come to any conclusion about the ramifications for variable *C. intestinalis* growth. Other invasive species, such as the green crab may in fact aid *C. intestinalis* by not feeding on them as much as other species, while competing with and reducing numbers of native predators that would eat *C. intestinalis*. This invasional meltdown hypothesis (IMH) has not been tested in relation to *C. intestinalis*, or many marine invertebrates for that matter. The lack of manipulative experiment in this field of research severely limits the avenues for discussion that follow a firm footing, with any suggestions about the predator-prey dynamics occurring in NS being rather subjective. In saying that, predation is certainly a viable candidate to potentially explain some of the observed variability in this study that was not accounted for by the monitored abiotic variables. A study that was conducted in relation to *C. intestinalis* predation in NS was actually based on testing the contribution of an NIS, *Caprella mutica*, to the biotic resistance hypothesis (BRH) of the receiving environment. Collin and Johnson (2014) did indeed find that *C. intestinalis* recruitment was dramatically reduced by *C. mutica*, as a result of predation of *C. intestinalis* larvae. The scope for variable interactions related to the ERH, BRH via native and other NIS or invasive species, or the invasional meltdown hypothesis lead me to conclude that there is a strong possibility that these dynamic and variable interactions could be a contributor to the unexplained variability observed in *C. intestinalis* growth during this study. It is also possible that all these factors contribute to variable levels with the potential to cause localized large-scale die-offs (Collin and Johnson, 2014; Dumont et al., 2011). The recovery of a population then depends on a number of factors that tie into the initial establishment of an invasive population, as were previously discussed. To touch on one, propagule pressure is thought to be pivotal in successful invasion (Drolet and Locke, 2016; Ricciardi et al., 2013; Verling et al., 2005), so variable levels of predation on spawning adults and larvae, as well as different abiotic effects could lead to variable inter and intra-annual population responses.



#### 4.4.2.2. Competition

Another factor that could cause intra- and inter-annual variation in *C. intestinalis* population growth is competition. Interspecific competition with *C. intestinalis* was tested in PEI using settlement plates that were pre-settled by the invasive colonial tunicates *Botryllus schlosseri* and *Botrylloides violaceus* (Paetzold et al., 2012). While this study showed that plates previously occupied by one of the colonial tunicates were less abundant in *C. intestinalis* compared to the plates that had no previous settlement on them, it was over carried out over a relatively short period of time and did not account for abiotic conditions. One such study that investigated abiotic conditions and interspecific competition was conducted on *B. violaceus*. It showed that interspecific competition appeared to have less of an effect on *B. violaceus* population growth rate compared to variable spatial environmental conditions (Grey, 2011). Further manipulative experiments that address abiotic conditions and interspecific competition between *C. intestinalis* and a range of native, NIS and invasive species are required to improve the understanding of these complex interactions.

The stochastic nature of species introductions and invasion dynamics could have resulted in a piecemeal pattern of effects determining population growth throughout NS. A multitude of effect ‘regimes’ determining patterns of population growth are likely, made up of different levels of abiotic and biotic contributions. If this is the case, long and short-term fluctuations in abiotic conditions, predator abundances, competitor abundance, and resource availability would affect populations in variable ways.

#### 4.5. Climate Change

With the consensus that global temperatures are rising at an unprecedented rate and predicted to continue on this trend (Pachauri et al., 2014), the issue of climate change is foremost in the minds of many ecologists. Marine benthic invertebrate communities are heavily influenced by temperature, while it is the widely accepted that invasive species are well adapted for environmental change, fortified by their tendency to have a wide range of environmental tolerance (Lord et al., 2015). Not only is temperature of concern, so too are pH due to ocean acidification as a result of CO<sub>2</sub> uptake, salinity changes as a result of changes in freshwater input as well as alterations in oceanic currents, while a projected increase in the frequency and intensity of extreme weather events certainly has the potential to alter water flow, particularly high flow events. All four of the abiotic variables I monitored have to be considered in the discussion about climate change, while all four have the potential to

influence *C. intestinalis* population dynamics. The way in which future changes will actually effect the distribution and abundance of *C. intestinalis* in NS and elsewhere remains largely unknown. Not only do these variables have the potential to shape sessile invertebrate communities on a short timescale and in relation to rapid environmental stochasticity, but they are also likely to shift in the long-term in relation to anthropogenic forcing. Sea surface temperature, perhaps the most prominent abiotic condition that can shape the changes of future *C. intestinalis* populations, is going to rise (Pachauri et al., 2014). The degree to which it will increase is predominantly dependent upon the atmospheric CO<sub>2</sub> concentration in the future, but a range of models suggest it could be from c.1.5°C – 6°C by 2100, without being able to rule out extreme climate sensitivity of > 10°C (Pachauri et al., 2014; Thompson, 2015). These predictive models tend to be globally administered, with a relatively coarse resolution of ~100 km. High resolution (~10 km) models carried out for the Northwest Atlantic suggest that warming in the region of the Gulf of Maine and the Scotian Shelf is likely to be 3 – 4 °C by 2100, with resultant ecosystem effects expected (Saba et al., 2016). Salinity, especially in the nearshore is likely to be influenced by the predicted increase in precipitation, especially the predicted rise of major precipitation events. Absorption of CO<sub>2</sub> by the oceans will only continue to lower the pH and create a more acidic environment. With an increase in major weather events, water motion, especially in wave exposed sites has the potential to become even more influential on benthic marine environments and biological communities. The complex nature of abiotic interactions with the population dynamics of *C. intestinalis* currently suggest that it is very difficult to predict how global change may influence the species in the future. Viewing any one abiotic variable in isolation is unlikely to provide an accurate picture of the scope of change under climate change (Todgham and Stillman, 2013), while the combined picture is ultimately very complicated. My results do suggest that a portion of *C. intestinalis* response in the future will be influenced by the changing abiotic environment.

## **4.6. Moving forward**

### *4.6.1. Genetic sequencing and analysis*

Completing the genetic sequencing of the individuals sampled in 2015, as well as commencing analysis on the sequences in hand for the 2014 samples is a high priority. It is important to discern the genetic diversity of the province-wide population, the molecular similarity between sub-populations, as well as getting an insight into the pattern of

introduction and spread, all of which can have knock on effects relating to potential range expansion (Lavergne and Molofsky, 2007). I will also continue to explore the methods by which a genetic similarity factor can be incorporated into the growth model. The identification of major genes responsible for differential growth trajectories within a population is well understood in sectors such as forestry and agriculture, and are often expressed using quantitative trait loci (QTL) (Wu et al., 2003, 2002). A method is presented by Malosetti et al. (2006) to incorporate growth modelling with QTL via nonlinear mixed modelling. This could be a suitable method to develop a new model, incorporating genetic information, to further explain the variable growth in *C. intestinalis* populations.

#### 4.6.2. Improved NLME model input

An element not included in the model is the size of individuals compared to the number of individuals at a given site, both of which are currently summarized together using the single percent cover measure. This information, in addition to the measure of relative abundance using the percentage cover could improve model fitting in the future. Considering logistical and theoretical arguments, the percent cover abundance measure was the best choice for this study and provides a good measure of relative abundance at each site, but it is not flawless. Once the settlement plates had saturated at a site, like what happened at CC and FP in particular, there is then no additional measure of how abundance increases later on in the growing season at those sites. The rate at which the plates reach saturation provides information in itself, but it is not a perfect system of measurement. In addition, adult fecundity is size dependent, so reproductive output and potential subsequent recruitment from local individuals may vary from site to site depending on the size of individuals.

Including a measure of the food availability, via chl-*a* data, in the model is desirable. Point-samples of one measurement per measurement period are not an ideal data type to include in a model that contains high resolution abiotic data for the other variables. An alternative option is to use chl-*a* data obtained from satellite images (Antoine et al., 1996; Longhurst et al., 1995). However, obtaining chl-*a* data for coastal regions via satellite imagery can be problematic due to confounding effects on water colour (Barnes et al., 2014; Craig et al., 2012; Gons et al., 2002). The point-sample chl-*a* data collected in 2014 (see Appendix B) will be used to assess the suitability of chl-*a* data obtained from satellite imagery. Extracting satellite image chl-*a* data for the corresponding periods of time that point-samples were taken during 2014, and establishing the relationship between the two, will allow us to assess the

accuracy of satellite image data. If there is a reasonable relationship between the point-samples and the satellite image data, that would provide a good source of chl-*a* data to include in the model. It is also possible to build a local chl-*a* model, similar to that of Barnes et al. (2014) and Craig et al. (2012), using the *in situ* point-samples and available satellite data. This model tuned specifically to our study area is likely to provide superior estimates compared to generic models used for satellite data (Craig, pers. comm.).

Incorporating abiotic and biotic information from the winter of both years would be a potential way of improving the models. Unfortunately, there is no available data about the overwintering *C. intestinalis* populations. There is however winter temperature data available for the 2013/2014 winter, while winter temperature data could be accessed through online databases for the 2014/2015 winter period. Another notable factor of interest for the winter period is the extent of ice cover and potential ice scour for each site. Ice scour has been shown to have an influential role on rocky intertidal communities in NS (Belt et al., 2009), but to my knowledge there have been no studies documenting the effect of ice scour on biofouling species inhabiting artificial substrate in harbours or bays. It is possible that the variation in pack ice levels from year to year and duration of ice occupation of harbours could have an effect on the biofouling community. As usual, it is not quite as simple as increased ice quantity resulting in more disruption to the biological community. Ice cover can have an insulating effect (Scrosati and Eckersley, 2007), so depending on the nature of the ice cover and extent of scour, it may provide an advantage for overwintering *C. intestinalis* populations by increasing water temperature over the harshest months of the year. The level of disturbance over winter could also leave substrates more vulnerable to new introductions the following spring and summer, so once again there is a catalogue of potential interactions and complications when assessing the effect of ice on the biofouling communities.

There are a number of abiotic variables that were not monitored in this study, for logistical reasons, that could influence *C. intestinalis* growth. Different exposure to irradiance is believed to influence *C. intestinalis* settlement and growth. A study carried out in Chile (possibly *C. robusta*, although displaying temperature preferences similar to those of known *C. intestinalis* populations) found that full exposure of the study species to direct solar radiation had a negative effect on survival (Madariaga et al., 2014). I didn't monitor irradiance due to the difficulty in obtaining accurate results with the length between site visits. The settlement plates provide cover from direct solar radiation, while the floating docks and general harbour substrates provided some shelter from direct exposure to solar

radiation. Consistency in placement of collectors in relation to degree of solar radiation exposure was a component of location choice. In addition, two of the sites with the least amount of shade provided by surrounding infrastructure, the mussel leases IP and SH had high *C. intestinalis* abundance. It is also suggested that turbidity of the surrounding environment can influence settlement and growth of *C. intestinalis* (McLaughlin et al., 2013), while dissolved oxygen concentration, nutrient levels, as well as environmental pollutants may all have an effect on the growth of *C. intestinalis*.

#### 4.6.3. Future experiments

Controlled laboratory survival and growth experiments can confirm the presence of ‘true’ abiotic effects, but it could be the case that these ‘true’ effects are severely altered in the reality of a natural habitat, due to a multitude of additional contributing factors (Todgham and Stillman, 2013). My study emphasises how variable the response of *C. intestinalis* populations to abiotic effects is. Some *C. intestinalis* growth is determined by major abiotic players in the marine environment, like temperature and salinity. However, a large component of the response is due to other factors that remain unknown for now. Controlled growth experiments are commonly used. The environmental tolerance of a widely distributed tunicate species, *Microcosmus exasperates*, in the Mediterranean (Nagar and Shenkar, 2016), was investigated in much the same way that laboratory experiments were carried out on *C. intestinalis* individuals from NS (Vercaemer et al., 2011). Both studies showed that there are important relationships between temperature and salinity, outlining the species’ tolerance and the response of the respective tunicates. While these studies are very important, in the future it is worth bearing in mind the truly heterogeneous response of *C. intestinalis*, and possibly other invasive benthic species with heterogeneous distributions, on a large spatial scale in an uncontrolled environment. For abiotic variables that are well established as important variables in determining the growth of a species, I would suggest that more intricate experimental designs should be tested. In the case of temperature and salinity, new laboratory controlled growth and reproductive experiments could be conducted with the addition of direct and indirect predation effects, as well as variable competition effects.

This study has indicated that both pH and water flow are influential in determining *C. intestinalis* growth. As the track record in the literature for these two abiotic variables is not as good compared to that of temperature and salinity, basic controlled laboratory experiments assessing pH and water flow effect on larval development, settlement and recruitment are

required. In addition to testing the effects on larvae, these two abiotic variables could be used to test the effect on adult individuals, in relation to their growth, reproductive capacity and stress response. In a similar fashion to experiments carried out by Pineda et al. (2012b) investigating the gene expression of *Styela plicata* to salinity and temperature stress, *C. intestinalis* could be tested for stress tolerance to pH and water flow changes. In addition to that, simple experiments testing the fecundity and growth rate of adult individuals under varies pH and water flow treatments could provide useful information about how these abiotic variables influence *C. intestinalis* biology.

As has been mentioned previously, and as is emphasised by this study, field studies investigating new abiotic effects that could explain *C. intestinalis* growth are important. Only through field studies is the true complexity of an effect, or lack thereof, observed. Also, field studies addressing biotic interactions in relation to abiotic variables are important to pursue. The effect of predation could be tested through manipulative experiments using caged settlement plates to exclude fishes, crabs, sea star, large predators. Competition effects could be tested using manipulated settlement plates, with previously plates submerged alongside controls.

#### 4.6.4. Abiotic data calibration of online databases

In a similar fashion as described above for assessing the appropriateness of satellite chl-*a* data, the other *in situ* abiotic measurements could be used to assess the accuracy of large-scale sources of data from online databases. If there are good relationships between the online databases and the fine-scale *in situ* measurements, this could results in much greater sampling potential, as the need to conduct abiotic monitoring could be reduced, leaving more time and resources for biological monitoring.

### 4.7. Implications for management

Efforts have been made to create predictive models for the life-history stages and potential distribution of *C. intestinalis* in the past (Patanasatienkul et al., 2014; Therriault and Herborg, 2008a). These prior efforts were not reliant on extensive field trials, instead relying on what could be found in the literature. While this is beneficial in terms of being less labour and time intensive, there are some concerns with models developed prior to this study. There has been a tendency to include (potentially erroneous) information obtained from studies conducted on a distinct species (*C. robusta*) in *C. intestinalis* models (Patanasatienkul et al., 2014;

Therriault and Herborg, 2008a). Temperature and salinity are commonly used in predictive modelling for *C. intestinalis* (Therriault and Herborg, 2008a), so it is somewhat problematic when the tolerances assigned to *C. intestinalis* for these key components in the model are based on information from two species with a known difference in temperature tolerance. While using temperature and salinity may indeed be the best approach for simplicity of the model and also in terms of availability of data, it is crucial to use starting information germane to the species at hand, as well as noting that there is potential for a very heterogeneous response of *C. intestinalis* to these abiotic variables. The true strength of a model could be considered as the knowledge of where it is weakest and has limitations. The famous phrase ‘all models are wrong; some are useful’ loosely coined by George Box (Box, 1976) is pertinent to this work. In this respect, we can move forward with a better understanding of the variable *C. intestinalis* growth response in relation to abiotic variables, while also having a well suited tool to assist in accounting for this heterogeneity through the mixed-effects modelling.

The likelihood-theory approach used here lends itself to aiding management decisions, as it is easy to identify how likely models are to be the best model in comparison to one another. This provides a template to guide model selection depending on a given set of management criteria or based on the information that is realistically available. In this study, all available information from four abiotic variables was used to create the best explanatory model for variable *C. intestinalis* growth. But with pre-set criteria that only temperature and salinity data are readily available, the approach can be easily tailored to suit these requirements and select the best model within these constraints. If all of the available independent predictor variables used in model development are attainable for the given spatial range that predictions are desired, then the most likely model can be chosen, with a knowledge of the degree of uncertainty surrounding such a model. When working with a complex system that is difficult to obtain information for, then ultimately the more parsimonious models that would be easier to obtain good abiotic data for (e.g. temperature and salinity via online databases) should be chosen above highly parameterized models containing variables that it is more difficult to obtain data for (pH in this case). Ease of use for a predictive model should surely govern model selection when the difference between a ‘better’ more parametrized model and a more parsimonious model is not substantial.

While the best-temperature only models provided the most consistent predictions across both years, the models containing more abiotic variables have the potential to be better predictive

models. In order to avoid error in prediction due to the variable response of *C. intestinalis* to abiotic variables inter-annually, a sound predictive framework is required. I propose a Venn diagram predictive framework. In this framework, the abiotic variables that can be obtained are chosen. In this case temperature and salinity are obtainable for the province of NS. The best model from both 2014 and 2015 containing these two variables would then be used to predict *C. intestinalis* growth for the nearshore environment for all of NS. Then the two predictive maps would be compared. In areas that are forecast as low abundance sites by both models, these are the most likely to be unsuitable for *C. intestinalis*. In areas that both models predict as high abundance, these are the most likely to be suitable for *C. intestinalis* growth and most likely to result in a high abundance population. For areas that the two models do not agree on, proceed with caution, as the success of a potential population at these sites is the most uncertain, and would remain so until further data can be obtained to improve either the model, or the degree of belief in the ‘true’ response of *C. intestinalis* to abiotic conditions.

For management purposes, I would advise the use of models developed based on the more comprehensive 2015 dataset, or to use the Venn-diagram prediction approach outline above. I would also advise the use of a reduced model with only salinity and temperature to create a projection map for *C. intestinalis* growth in Atlantic Canada. This would maximise the area that predictions could be made for, due to the availability of data for these two abiotic variables. This study indicates that pH and water flow data would improve a predictive model, but in reality the lack of high resolution data for these two abiotic variables would severely limit the potential range for predictions. The use of the extended and most likely model according to this study, with all four abiotic variables, could be used for smaller scale predictions based on a specific location or multiple locations where obtaining the necessary abiotic data would be more manageable than for a large scale provincial prediction.

The results from this study emphasise the importance of continued long-term monitoring of *C. intestinalis* populations in NS. However, the results also suggest that increased sampling in a more localized area may be as effective at determining long-term population dynamics as compared to monitoring across the whole province. By testing alternative hypotheses to explain the variability in *C. intestinalis* growth that was not explained by the four abiotic variables here, as well as continued monitoring of growth in relation to abiotic variables using the NLME modelling approach, this may be sufficient to improve predictive ability of *C. intestinalis* population growth in the future.



## 5. Conclusions

Some of the heterogeneous response of *C. intestinalis* population growth can be attributed to environmental conditions, with growth responding to changes in surface water temperature, salinity, pH, and water flow. Salinity and temperature appear to be the most important abiotic determinants of population growth across both years. The influence of salinity is highly changeable, while temperature also had varying degrees of association with the observed population growth, based on this study. The extent to which water flow and pH influence population growth is unclear, but they did improve model fitting in 2015.

This study emphasised the variable and sometimes unpredictable dynamics of *C. intestinalis* population growth, highlighting the necessity to conduct more long-term studies. A key priority moving forward should be improving prevention of further introductions of *C. intestinalis* around the province and further afield. While variability of *C. intestinalis* growth response to abiotic conditions is high, this study showed that mixed-effects modelling can provide a well fitted model that performs consistently under standard model validation and nested cross-validation testing.

The models fit with data from either 2014 only or 2015 only predicted with variable accuracy when used to predict the observed data from the year not included in model fitting. The best model (BU and TD the same) from 2014 was able to predict 2015 data with ~30 % accuracy. On the other hand, the best model (TD) in 2015 performed very poorly at predicting 2014 observations, with a negative model efficiency score. This poor performance is likely a combination of a smaller dataset devoid of any observations from the beginning of the monitoring period, as well as the shift in *C. intestinalis* growth response to abiotic conditions between the two years. In the interest of model parsimony, temperature-only models were also assessed in the model validation, cross-validation, and prediction processes. While models containing temperature exclusively did not come out on top in the BU and TD model building approaches, they performed almost as well (BTO 2014) and much better (BTO 2015) at predicting inter-annual observations from data not used in the model fitting. This suggest that temperature is the most consistent abiotic predictor of *C. intestinalis* growth.

This exploratory study has laid out some fundamental baseline information about *C. intestinalis* growth in relation to a group of important abiotic variables in the nearshore

environment. These findings can be built upon in order to continue improving the understanding of *C. intestinalis* population growth.

## References

- ACRDP, 2010. Containment and mitigation of nuisance tunicates on Prince Edward Island to improve mussel farm productivity. Aquaculture Collaborative Research and Development Program (ACRDP) Fact Sheet 6.
- Adame, P., del Río, M., Cañellas, I., 2008a. A mixed nonlinear height–diameter model for pyrenean oak (*Quercus pyrenaica* Willd.). *For. Ecol. Manag.* 256, 88–98. doi:10.1016/j.foreco.2008.04.006
- Adame, P., Hynynen, J., Cañellas, I., del Río, M., 2008b. Individual-tree diameter growth model for rebollo oak (*Quercus pyrenaica* Willd.) coppices. *For. Ecol. Manag.* 255, 1011–1022. doi:10.1016/j.foreco.2007.10.019
- Adams, C.M., Shumway, S.E., Whitlatch, R.B., Getchis, T., 2011. Biofouling in Marine Molluscan Shellfish Aquaculture: A Survey Assessing the Business and Economic Implications of Mitigation. *J. World Aquac. Soc.* 42, 242–252. doi:10.1111/j.1749-7345.2011.00460.x
- Akaike, H., 1973. Information theory and an extension of the maximum likelihood principle, in: B.N. Petrov, F. Csaki (Eds.), *Second International Symposium on Information Theory*. Akademiai Kiado, Budapest, pp. 267–281.
- Allen, F., 1953. Distribution of Marine Invertebrates by Ships. *Mar. Freshw. Res.* 4, 307–316.
- Alyokhin, A., 2011. Non-natives: put biodiversity at risk. *Nature* 475, 36–36. doi:10.1038/475036b
- Amaro, A., Reed, D., Tomé, M., Themido, I., 1998. Modeling Dominant Height Growth: Eucalyptus Plantations in Portugal. *For. Sci.* 44, 37–46.
- Anderson, D.R., Link, W.A., Johnson, D.H., Burnham, K.P., 2001. Suggestions for Presenting the Results of Data Analyses. *J. Wildl. Manag.* 65, 373–378.
- Antoine, D., André, J.-M., Morel, A., 1996. Ocean primary production. 2. Estimation at global scale from satellite (coastal zone color scanner) chlorophyll. *Glob. Biogeochem. Cycles* 10, 57–69.
- Auker, L.A., Oviatt, C.A., 2008. Factors influencing the recruitment and abundance of *Didemnum* in Narragansett Bay, Rhode Island. *ICES J. Mar. Sci. J. Cons.* 65, 765–769. doi:10.1093/icesjms/fsm196
- Averbuj, A., Escati-Peñaloza, G., Penchaszadeh, P.E., 2015. Individual Growth in the Patagonian Gastropod *Buccinanops cochlidium* (Nassariidae): A Field Tagging-Recapture Experiment. *Malacologia* 59, 121–133. doi:10.4002/040.059.0108
- Ayre, D.J., Minchinton, T.E., Perrin, C., 2009. Does life history predict past and current connectivity for rocky intertidal invertebrates across a marine biogeographic barrier? *Mol. Ecol.* 18, 1887–1903. doi:10.1111/j.1365-294X.2009.04127.x
- Bafumi, J., Gelman, A., 2007. Fitting Multilevel Models When Predictors and Group Effects Correlate (SSRN Scholarly Paper No. ID 1010095). Social Science Research Network, Rochester, NY.
- Barnes, B.B., Hu, C., Cannizzaro, J.P., Craig, S.E., Hallock, P., Jones, D.L., Lehrter, J.C., Melo, N., Schaeffer, B.A., Zepp, R., 2014. Estimation of diffuse attenuation of ultraviolet light in optically shallow Florida Keys waters from MODIS measurements. *Remote Sens. Environ.* 140, 519–532.
- Bechmann, R.K., Taban, I.C., Westerlund, S., Godal, B.F., Arnberg, M., Vingen, S., Ingvarsdottir, A., Baussant, T., 2011. Effects of Ocean Acidification on Early Life Stages of Shrimp (*Pandalus borealis*) and Mussel (*Mytilus edulis*). *J. Toxicol. Environ. Health A* 74, 424–438. doi:10.1080/15287394.2011.550460

- Bell, A., Jones, K., 2015. Explaining Fixed Effects: Random Effects Modeling of Time-Series Cross-Sectional and Panel Data. *Polit. Sci. Res. Methods* 3, 133–153. doi:10.1017/psrm.2014.7
- Belt, K.M., Cole, S.W., Scrosati, R.A., 2009. Intertidal barnacles as indicators of the intensity of scour by sea ice. *Mar. Ecol. Prog. Ser.* 381, 183–187.
- Beninger, P.G., Boldina, I., Katsanevakis, S., 2012. Strengthening statistical usage in marine ecology. *J. Exp. Mar. Biol. Ecol.* 426–427, 97–108. doi:10.1016/j.jembe.2012.05.020
- Berrill, N.J., 1947. The Development and Growth of *Ciona*. *J. Mar. Biol. Assoc. U. K.* 26, 616–625. doi:10.1017/S0025315400013825
- Blackburn, T.M., Pyšek, P., Bacher, S., Carlton, J.T., Duncan, R.P., Jarošík, V., Wilson, J.R.U., Richardson, D.M., 2011. A proposed unified framework for biological invasions. *Trends Ecol. Evol.* 26, 333–339. doi:10.1016/j.tree.2011.03.023
- Blum, J.C., Chang, A.L., Liljesthrom, M., Schenk, M.E., Steinberg, M.K., Ruiz, G.M., 2007. The non-native solitary ascidian *Ciona intestinalis* (L.) depresses species richness. *J. Exp. Mar. Biol. Ecol.*, Proceedings of the 1st International Invasive Sea Squirt Conference 342, 5–14. doi:10.1016/j.jembe.2006.10.010
- Boldina, I., Beninger, P.G., 2016. Strengthening statistical usage in marine ecology: Linear regression. *J. Exp. Mar. Biol. Ecol.* 474, 81–91. doi:10.1016/j.jembe.2015.09.010
- Bolker, B., 2008. *Ecological models and data in R*. Princeton University Press, Princeton, NJ.
- Bouchemousse, S., Haag-Liautard, C., Bierne, N., Viard, F., 2015. Past and contemporary introgression between two strongly differentiated *Ciona* species as revealed by a post-genomic SNPs panel. *bioRxiv* 30346. doi:10.1101/030346
- Box, G.E.P., 1976. Science and Statistics. *J. Am. Stat. Assoc.* 71, 791–799. doi:10.2307/2286841
- Brunetti, R., Gissi, C., Pennati, R., Caicci, F., Gasparini, F., Manni, L., 2015. Morphological evidence that the molecularly determined *Ciona intestinalis* type A and type B are different species: *Ciona robusta* and *Ciona intestinalis*. *J. Zool. Syst. Evol. Res.* 53, 186–193. doi:10.1111/jzs.12101
- Buqui, G.A., Sy, S.K.B., Merino-Sanjuán, M., Gouvea, D.R., Nixdorf, S.L., Kimura, E., Derendorf, H., Lopes, N.P., Diniz, A., 2015. Characterization of intestinal absorption of C-glycoside flavonoid vicenin-2 from *Lychnophora ericoides* leaves in rats by nonlinear mixed effects modeling. *Rev. Bras. Farmacogn.* 25, 212–218. doi:10.1016/j.bjp.2015.04.001
- Burnham, K.P., Anderson, D.R., 2003. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. Springer Science & Business Media.
- Buttlar, J.V., 2015. *JBTools: Misc Small Tools and Helper Functions for Other Code of J. Buttlar*.
- Byers, J.E., 2002. Impact of non-indigenous species on natives enhanced by anthropogenic alteration of selection regimes. *Oikos* 97, 449–458. doi:10.1034/j.1600-0706.2002.970316.x
- Byrd, J., Lambert, C.C., 2000. Mechanism of the block to hybridization and selfing between the sympatric ascidians *Ciona intestinalis* and *Ciona savignyi*. *Mol. Reprod. Dev.* 55, 109–116. doi:10.1002/(SICI)1098-2795(200001)55:1<109::AID-MRD15>3.0.CO;2-B
- Cade, B.S., 2015. Model averaging and muddled multimodel inferences. *Ecology* 96, 2370–2382. doi:10.1890/14-1639.1
- Calegario, N., Daniels, R.F., Maestri, R., Neiva, R., 2005. Modeling dominant height growth based on nonlinear mixed-effects model: a clonal Eucalyptus plantation case study. *For. Ecol. Manag.* 204, 11–21. doi:10.1016/j.foreco.2004.07.051

- Caputi, L., Andreakis, N., Mastrototaro, F., Cirino, P., Vassillo, M., Sordino, P., 2007. Cryptic speciation in a model invertebrate chordate. *Proc. Natl. Acad. Sci.* 104, 9364–9369. doi:10.1073/pnas.0610158104
- Caputi, L., Crocetta, F., Toscano, F., Sordino, P., Cirino, P., 2015. Long-term demographic and reproductive trends in *Ciona intestinalis* sp. A. *Mar. Ecol.* 36, 118–128. doi:10.1111/maec.12125
- Carman, M.R., Morris, J.A., Karney, R.C., Grunden, D.W., 2010. An initial assessment of native and invasive tunicates in shellfish aquaculture of the North American east coast. *J. Appl. Ichthyol.* 26, 8–11. doi:10.1111/j.1439-0426.2010.01495.x
- Carver, C.E., Chisholm, A., Mallet, A.L., 2003. Strategies to mitigate the impact of *Ciona intestinalis* (L.) biofouling on shellfish production. *J. Shellfish Res.* 22, 621–631.
- Carver, C.E., Mallet, A.L., Vercaemer, B., 2006. Biological synopsis of the solitary tunicate *Ciona intestinalis*. Bedford Institute of Oceanography Dartmouth, Nova Scotia.
- Chan, F.T., Bailey, S.A., Wiley, C.J., MacIsaac, H.J., 2013. Relative risk assessment for ballast-mediated invasions at Canadian Arctic ports. *Biol. Invasions* 15, 295–308. doi:10.1007/s10530-012-0284-z
- Clarke, K., Gorley, R., 2015. PRIMER v7: User Manual/Tutorial. PRIMER-E, Plymouth.
- Colautti, R.I., Bailey, S.A., Overdijk, C.D.A. van, Amundsen, K., MacIsaac, H.J., 2006. Characterised and Projected Costs of Nonindigenous Species in Canada. *Biol. Invasions* 8, 45–59. doi:10.1007/s10530-005-0236-y
- Collin, S.B., Edwards, P.K., Leung, B., Johnson, L.E., 2013. Optimizing early detection of non-indigenous species: Estimating the scale of dispersal of a nascent population of the invasive tunicate *Ciona intestinalis* (L.). *Mar. Pollut. Bull.* 73, 64–69. doi:10.1016/j.marpolbul.2013.05.040
- Collin, S.B., Johnson, L.E., 2014. Invasive species contribute to biotic resistance: negative effect of caprellid amphipods on an invasive tunicate. *Biol. Invasions* 16, 2209–2219. doi:10.1007/s10530-014-0659-4
- Comeau, L.A., Sonier, R., Hanson, J.M., 2012. Seasonal movements of Atlantic rock crab (*Cancer irroratus* Say) transplanted into a mussel aquaculture site. *Aquac. Res.* 43, 509–517. doi:10.1111/j.1365-2109.2011.02856.x
- Craig, S.E., Jones, C.T., Li, W.K., Lazin, G., Horne, E., Caverhill, C., Cullen, J.J., 2012. Deriving optical metrics of coastal phytoplankton biomass from ocean colour. *Remote Sens. Environ.* 119, 72–83.
- Davidson, I., Zabin, C., Chang, A., Brown, C., Sytsma, M., Ruiz, G., 2010. Recreational Boats as Potential Vectors of Marine Organisms at an Invasion Hotspot. *Cent. Lakes Reserv. Publ. Present.*
- Davis, M.A., Chew, M.K., Hobbs, R.J., Lugo, A.E., Ewel, J.J., Vermeij, G.J., Brown, J.H., Rosenzweig, M.L., Gardener, M.R., Carroll, S.P., Thompson, K., Pickett, S.T.A., Stromberg, J.C., Tredici, P.D., Suding, K.N., Ehrenfeld, J.G., Philip Grime, J., Mascaro, J., Briggs, J.C., 2011. Don't judge species on their origins. *Nature* 474, 153–154. doi:10.1038/474153a
- Davis, M.A., Grime, J.P., Thompson, K., 2000. Fluctuating resources in plant communities: a general theory of invasibility. *J. Ecol.* 88, 528–534.
- DiBacco, C., Humphrey, D.B., Nasmith, L.E., Levings, C.D., 2012. Ballast water transport of non-indigenous zooplankton to Canadian ports. *ICES J. Mar. Sci. J. Cons.* 69, 483–491. doi:10.1093/icesjms/fsr133
- Dijkstra, J., Dutton, A., Westerman, E., Harris, L., 2008. Heart rate reflects osmotic stress levels in two introduced colonial ascidians *Botryllus schlosseri* and *Botrylloides violaceus*. *Mar. Biol.* 154, 805–811. doi:10.1007/s00227-008-0973-4

- Doney, S.C., Fabry, V.J., Feely, R.A., Kleypas, J.A., 2009. Ocean Acidification: The Other CO<sub>2</sub> Problem. *Annu. Rev. Mar. Sci.* 1, 169–192. doi:10.1146/annurev.marine.010908.163834
- Drolet, D., DiBacco, C., Locke, A., McKenzie, C.H., McKindsey, C.W., Moore, A.M., Webb, J.L., Therriault, T.W., 2015. Evaluation of a new screening-level risk assessment tool applied to non-indigenous marine invertebrates in Canadian coastal waters. *Biol. Invasions* 18, 279–294. doi:10.1007/s10530-015-1008-y
- Drolet, D., Locke, A., 2016. Relative importance of propagule size and propagule number for establishment of non-indigenous species: a stochastic simulation study.
- Dumont, C.P., Gaymer, C.F., Thiel, M., 2011. Predation contributes to invasion resistance of benthic communities against the non-indigenous tunicate *Ciona intestinalis*. *Biol. Invasions* 13, 2023–2034. doi:10.1007/s10530-011-0018-7
- Dupont, S., Havenhand, J., Thorndyke, M., 2009. Impact of CO<sub>2</sub>-driven ocean acidification on early life-history — what we know and what we need to know. *IOP Conf. Ser. Earth Environ. Sci.* 6, 462006. doi:10.1088/1755-1307/6/46/462006
- Dupont, S., Thorndyke, M.C., Commission Internationale pour l'Exploration Scientifique de la Mer Mediterranee- CIESM, M., Briand, F., 2008. Ocean acidification and its impact on the early life-history stages of marine animals, in: CIESM Workshop Monographs. CIESM, Monaco, pp. 89–97.
- Dybern, B.I., 1967. The distribution and salinity tolerance of *Ciona intestinalis* (L.) F. typica with special reference to the waters around Southern Scandinavia. *Ophelia* 4, 207–226. doi:10.1080/00785326.1967.10409621
- Dybern, B.I., 1965. The Life Cycle of *Ciona intestinalis* (L.) f. typica in Relation to the Environmental Temperature. *Oikos* 16, 109–131. doi:10.2307/3564870
- Elton, C.S., 1958. *The ecology of invasions by animals and plants*. Methuen & Co, London.
- Elton, C.S., 1942. *Voles, Mice and Lemmings: Problems in population dynamics*. Clarendon, Oxford.
- Enge, S., Nylund, G.M., Pavia, H., 2013. Native generalist herbivores promote invasion of a chemically defended seaweed via refuge-mediated apparent competition. *Ecol. Lett.* 16, 487–492. doi:10.1111/ele.12072
- Epelbaum, A., Herborg, L.M., Therriault, T.W., Pearce, C.M., 2009. Temperature and salinity effects on growth, survival, reproduction, and potential distribution of two non-indigenous botryllid ascidians in British Columbia. *J. Exp. Mar. Biol. Ecol.* 369, 43–52. doi:10.1016/j.jembe.2008.10.028
- Escati-Peñaloza, G., Parma, A.M., Orensanz, J.M. (Lobo), 2010. Analysis of longitudinal growth increment data using mixed-effects models: Individual and spatial variability in a clam. *Fish. Res.* 105, 91–101. doi:10.1016/j.fishres.2010.03.007
- Farrapeira, C.M.R., Tenório, D. de O., Amaral, F.D. do, 2011. Vessel biofouling as an inadvertent vector of benthic invertebrates occurring in Brazil. *Mar. Pollut. Bull.* 62, 832–839. doi:10.1016/j.marpolbul.2010.12.014
- Fletcher, L.M., Forrest, B.M., Bell, J.J., 2013. Impacts of the invasive ascidian *Didemnum vexillum* on the green-lipped mussel *Perna canaliculus* aquaculture in New Zealand. *Aquac. Environ. Interact.* 4, 17–30. doi:doi: 10.3354/aei00069
- Floerl, O., Inglis, G.J., Dey, K., Smith, A., 2009. The importance of transport hubs in stepping-stone invasions. *J. Appl. Ecol.* 46, 37–45. doi:10.1111/j.1365-2664.2008.01540.x
- Fujikawa, T., Munakata, T., Kondo, S., Satoh, N., Wada, S., 2009. Stress response in the ascidian *Ciona intestinalis*: transcriptional profiling of genes for the heat shock protein 70 chaperone system under heat stress and endoplasmic reticulum stress. *Cell Stress Chaperones* 15, 193–204. doi:10.1007/s12192-009-0133-x

- Gjedrem, T., Robinson, N., Rye, M., 2012. The importance of selective breeding in aquaculture to meet future demands for animal protein: A review. *Aquaculture* 350–353, 117–129. doi:10.1016/j.aquaculture.2012.04.008
- Goff, E.L., Martinand-Mari, C., Martin, M., Feuillard, J., Boublik, Y., Godefroy, N., Mangeat, P., Baghdiguian, S., Cavalli, G., 2015. Enhancer of zeste acts as a major developmental regulator of *Ciona intestinalis* embryogenesis. *Biol. Open* bio.010835. doi:10.1242/bio.010835
- Gons, H.J., Rijkeboer, M., Ruddick, K.G., 2002. A chlorophyll-retrieval algorithm for satellite imagery (Medium Resolution Imaging Spectrometer) of inland and coastal waters. *J. Plankton Res.* 24, 947–951. doi:10.1093/plankt/24.9.947
- Grey, E.K., 2011. Relative effects of environment and direct species interactions on the population growth rate of an exotic ascidian. *Oecologia* 166, 935–947. doi:10.1007/s00442-011-1931-2
- Grosholz, E., 2002. Ecological and evolutionary consequences of coastal invasions. *Trends Ecol. Evol.* 17, 22–27. doi:10.1016/S0169-5347(01)02358-8
- Gulliksen, B., Skjæveland, S.H., 1973. The sea-star, *Asterias rubens* L., as predator on the ascidian, *Ciona intestinalis* (L.), in Borgenfjorden, North-Trøndelag, Norway. *Sarsia* 52, 15–20. doi:10.1080/00364827.1973.10411228
- Gunasekara, F.I., Richardson, K., Carter, K., Blakely, T., 2014. Fixed effects analysis of repeated measures data. *Int. J. Epidemiol.* 43, 264–269. doi:10.1093/ije/dyt221
- Harris, L.G., Tyrrell, M.C., 2001. Changing Community States in the Gulf of Maine: Synergism Between Invaders, Overfishing and Climate Change. *Biol. Invasions* 3, 9–21. doi:10.1023/A:1011487219735
- Havenhand, J.N., Svane, I.B., 1991. Roles of hydrodynamics and larval behaviour in determining spatial aggregation in the tunicate *Ciona intestinalis*. *Mar Ecol Prog Ser* 68, 271–276.
- Henslow, J.S., 1835. Observations concerning the indigenoussness and distinctness of certain species of plants included in the British floras. *Mag. Nat. Hist.* 8, 84–88.
- Hernández-Otero, A., Gaspar, M.B., Macho, G., Vázquez, E., 2014. Age and growth of the sword razor clam *Ensis arcuatus* in the Ría de Pontevedra (NW Spain): Influence of environmental parameters. *J. Sea Res.* 85, 59–72. doi:10.1016/j.seares.2013.09.006
- Hopkins, G.A., Prince, M., Cahill, P.L., Fletcher, L.M., Atalah, J., 2016. Desiccation as a mitigation tool to manage biofouling risks: trials on temperate taxa to elucidate factors influencing mortality rates. *Biofouling* 32, 1–11. doi:10.1080/08927014.2015.1115484
- Hoshino, Z., 'ichiro, Tokioka, T., 1967. An unusually robust *Ciona* from the northeastern coast of Honsyu Island, Japan.
- Howes, S., Herbinger, C.M., Darnell, P., Vercaemer, B., 2007. Spatial and temporal patterns of recruitment of the tunicate *Ciona intestinalis* on a mussel farm in Nova Scotia, Canada. *J. Exp. Mar. Biol. Ecol., Proceedings of the 1st International Invasive Sea Squirt Conference* 342, 85–92. doi:10.1016/j.jembe.2006.10.018
- Inderjit, Chapman, D., Ranelletti, M., Kaushik, S., 2006. Invasive Marine Algae: An Ecological Perspective. *Bot. Rev.* 72, 153–178. doi:http://dx.doi.org/10.1663/0006-8101(2006)72[153:IMAAEP]2.0.CO;2
- Jeffery, W.R., 2015. Distal regeneration involves the age dependent activity of branchial sac stem cells in the ascidian *Ciona intestinalis*. *Regeneration* 2, 1–18. doi:10.1002/reg2.26
- Kanary, L., Locke, A., Watmough, J., Chasse, J., Bourque, D., Nadeau, A., 2011. Predicting larval dispersal of the vase tunicate *Ciona intestinalis* in a Prince Edward Island

- estuary using a matrix population model. *Aquat. Invasions* 6, 491–506.  
doi:10.3391/ai.2011.6.4.14
- Kawaguchi, A., Utsumi, N., Morita, M., Ohya, A., Wada, S., 2015. Application of the cis-regulatory region of a heat-shock protein 70 gene to heat-inducible gene expression in the ascidian *Ciona intestinalis*. *genesis* 53, 170–182. doi:10.1002/dvg.22834
- Keane, R.M., Crawley, M.J., 2002. Exotic plant invasions and the enemy release hypothesis. *Trends Ecol. Evol.* 17, 164–170. doi:10.1016/S0169-5347(02)02499-0
- Kelly, D.W., Paterson, R.A., Townsend, C.R., Poulin, R., Tompkins, D.M., 2009. Parasite spillback: A neglected concept in invasion ecology? *Ecology* 90, 2047–2056.  
doi:10.1890/08-1085.1
- Kimbro, D.L., Cheng, B.S., Grosholz, E.D., 2013. Biotic resistance in marine environments. *Ecol. Lett.* 16, 821–833. doi:10.1111/ele.12106
- Lacoursière-Roussel, A., Bock, D.G., Cristescu, M.E., Guichard, F., Girard, P., Legendre, P., McKindsey, C.W., 2012a. Disentangling invasion processes in a dynamic shipping–boating network. *Mol. Ecol.* 21, 4227–4241. doi:10.1111/j.1365-294X.2012.05702.x
- Lacoursière-Roussel, A., Forrest, B.M., Guichard, F., Piola, R.F., McKindsey, C.W., 2012b. Modeling biofouling from boat and source characteristics: a comparative study between Canada and New Zealand. *Biol. Invasions* 14, 2301–2314.  
doi:10.1007/s10530-012-0230-0
- Laird, N.M., Ware, J.H., 1982. Random-Effects Models for Longitudinal Data. *Biometrics* 38, 963–974. doi:10.2307/2529876
- Lambert, C.C., Brandt, C.L., 1967. The Effect of Light on the Spawning of *Ciona intestinalis*. *Biol. Bull.* 132, 222–228. doi:10.2307/1539890
- Lambert, C.C., Lambert, G., 1998. Non-indigenous ascidians in southern California harbors and marinas. *Mar. Biol.* 130, 675–688. doi:10.1007/s002270050289
- Lambert, G., 2005. Ecology and natural history of the protochordates. *Can. J. Zool.* 83, 34–50. doi:10.1139/z04-156
- Lavergne, S., Molofsky, J., 2007. Increased genetic variation and evolutionary potential drive the success of an invasive grass. *Proc. Natl. Acad. Sci.* 104, 3883–3888.  
doi:10.1073/pnas.0607324104
- Lehuta, S., Petitgas, P., Mahévas, S., Huret, M., Vermard, Y., Uriarte, A., Record, N.R., 2013. Selection and validation of a complex fishery model using an uncertainty hierarchy. *Fish. Res.* 143, 57–66. doi:10.1016/j.fishres.2013.01.008
- Lerdau, M., Wickham, J.D., 2011. Non-natives: four risk factors. *Nature* 475, 36–37.  
doi:10.1038/475036d
- Lin, Y., Chen, Y., Xiong, W., Zhan, A., 2015. Genomewide gene-associated microsatellite markers for the model invasive ascidian, *Ciona intestinalis* species complex. *Mol. Ecol. Resour.* n/a-n/a. doi:10.1111/1755-0998.12481
- Lindstrom, M.J., Bates, D.M., 1990. Nonlinear Mixed Effects Models for Repeated Measures Data. *Biometrics* 46, 673–687. doi:10.2307/2532087
- Lloyd, M., Metaxas, A., deYoung, B., 2012. Patterns in vertical distribution and their potential effects on transport of larval benthic invertebrates in a shallow embayment. *Mar. Ecol. Prog. Ser.* 469, 37–52. doi:10.3354/meps09983
- Locke, A., Doe, K.G., Fairchild, W.L., Jackman, P.M., Reese, E.J., 2009. Preliminary evaluation of effects of invasive tunicate management with acetic acid and calcium hydroxide on non-target marine organisms in Prince Edward Island, Canada. *Aquat. Invasions* 4, 221–236.
- Lockwood, J.L., Hoopes, M.F., Marchetti, M.P., 2013. *Invasion Ecology*. John Wiley & Sons.



- Lockwood, J.L., Hoopes, M.F., Marchetti, M.P., 2011. Non-natives: plusses of invasion ecology. *Nature* 475, 36–36. doi:10.1038/475036c
- Longhurst, A., Sathyendranath, S., Platt, T., Caverhill, C., 1995. An estimate of global primary production in the ocean from satellite radiometer data. *J. Plankton Res.* 17, 1245–1271. doi:10.1093/plankt/17.6.1245
- López-Legentil, S., Legentil, M.L., Erwin, P.M., Turon, X., 2014. Harbor networks as introduction gateways: contrasting distribution patterns of native and introduced ascidians. *Biol. Invasions* 17, 1623–1638. doi:10.1007/s10530-014-0821-z
- López-Maury, L., Marguerat, S., Bähler, J., 2008. Tuning gene expression to changing environments: from rapid responses to evolutionary adaptation. *Nat. Rev. Genet.* 9, 583–593. doi:10.1038/nrg2398
- Lord, J.P., Calini, J.M., Whitlatch, R.B., 2015. Influence of seawater temperature and shipping on the spread and establishment of marine fouling species. *Mar. Biol.* 1–12. doi:10.1007/s00227-015-2737-2
- Lowen, J.B., Deibel, D., McKenzie, C.H., Couturier, C., DiBacco, C., 2016. Tolerance of early life-stages in *Ciona intestinalis* to bubble streams and suspended particles.
- Lundälv, T., Christie, H., 1987. Comparative trends and ecological patterns of rocky subtidal communities in the Swedish and Norwegian Skagerrak area, in: *Long-Term Changes in Coastal Benthic Communities*. Springer, pp. 71–80.
- Lutz-Collins, V., 2009. Invasive tunicates fouling mussel lines: evidence of their impact on native tunicates and other epifaunal invertebrates. *Aquat. Invasions* 4, 213–220. doi:10.3391/ai.2009.4.1.22
- Mack, R.N., Simberloff, D., Mark Lonsdale, W., Evans, H., Clout, M., Bazzaz, F.A., 2000. Biotic Invasions: Causes, Epidemiology, Global Consequences, and Control. *Ecol. Appl.* 10, 689–710. doi:10.1890/1051-0761(2000)010[0689:BICEGC]2.0.CO;2
- Madariaga, D.J., Rivadeneira, M.M., Tala, F., Thiel, M., 2014. Environmental tolerance of the two invasive species *Ciona intestinalis* and *Codium fragile*: their invasion potential along a temperate coast. *Biol. Invasions* 16, 2507–2527. doi:10.1007/s10530-014-0680-7
- Malosetti, M., Visser, R.G.F., Celis-Gamboa, C., Eeuwijk, F.A. van, 2006. QTL methodology for response curves on the basis of non-linear mixed models, with an illustration to senescence in potato. *Theor. Appl. Genet.* 113, 288–300. doi:10.1007/s00122-006-0294-2
- Marquet, N., Nicastró, K.R., Gektidis, M., McQuaid, C.D., Pearson, G.A., Serrão, E.A., Zardi, G.I., 2012. Comparison of phototrophic shell-degrading endoliths in invasive and native populations of the intertidal mussel *Mytilus galloprovincialis*. *Biol. Invasions* 15, 1253–1272. doi:10.1007/s10530-012-0363-1
- Matsunobu, S., Sasakura, Y., 2015. Time course for tail regression during metamorphosis of the ascidian *Ciona intestinalis*. *Dev. Biol.* 405, 71–81. doi:10.1016/j.ydbio.2015.06.016
- McLaughlin, J., Bourque, D., LeBlanc, A., Fortin, G., 2013. Effect of suspended inorganic matter on fertilization success, embryonic development, larval settlement, and juvenile survival of the vase tunicate *Ciona intestinalis* (Linnaeus, 1767). *Aquat. Invasions* 8, 375–388. doi:10.3391/ai.2013.8.4.02
- Meng, S.X., Huang, S., Lieffers, V.J., Nunifu, T., Yang, Y., 2008. Wind speed and crown class influence the height–diameter relationship of lodgepole pine: Nonlinear mixed effects modeling. *For. Ecol. Manag.* 256, 570–577. doi:10.1016/j.foreco.2008.05.002
- Miller, S.H., Morgan, S.G., 2013. Interspecific differences in depth preference: regulation of larval transport in an upwelling system. *Mar. Ecol. Prog. Ser.* 476, 301–306.

- Murray, C.C., Therriault, T.W., Martone, P.T., 2012. Adapted for invasion? Comparing attachment, drag and dislodgment of native and nonindigenous hull fouling species. *Biol. Invasions* 14, 1651–1663. doi:10.1007/s10530-012-0178-0
- Murray, C.C., Therriault, T.W., Pakhomov, E., 2013. What Lies Beneath? An Evaluation of Rapid Assessment Tools for Management of Hull Fouling. *Environ. Manage.* 52, 374–384. doi:10.1007/s00267-013-0085-x
- Nagar, L.R., Shenkar, N., 2016. Temperature and salinity sensitivity of the invasive ascidian *Microcosmus exasperatus* Heller, 1878.
- Nakamura, J., Tetsukawa, A., Fujiwara, S., 2015. Chondroitin 4-O-sulfotransferases are required for cell adhesion and morphogenesis in the *Ciona intestinalis* embryo. *Dev. Growth Differ.* 57, 58–67. doi:10.1111/dgd.12188
- Nakazawa, S., Shirae-Kurabayashi, M., Otsuka, K., Sawada, H., 2015. Proteomics of ionomycin-induced ascidian sperm reaction: Released and exposed sperm proteins in the ascidian *Ciona intestinalis*. *PROTEOMICS* 15, 4064–4079. doi:10.1002/pmic.201500162
- Newson, R., 2002. Parameters behind “nonparametric” statistics: Kendall’s tau, Somers’ D and median differences. *Stata J.* 2, 45–64.
- Nydam, M.L., Harrison, R.G., 2011. Introgression Despite Substantial Divergence in a Broadcast Spawning Marine Invertebrate. *Evolution* 65, 429–442. doi:10.1111/j.1558-5646.2010.01153.x
- Nydam, M.L., Harrison, R.G., 2010. Polymorphism and divergence within the ascidian genus *Ciona*. *Mol. Phylogenet. Evol.* 56, 718–726. doi:10.1016/j.ympev.2010.03.042
- Nydam, M.L., Harrison, R.G., 2007. Genealogical relationships within and among shallow-water *Ciona* species (Ascidiacea). *Mar. Biol.* 151, 1839–1847. doi:10.1007/s00227-007-0617-0
- Olsen, E., Fay, G., Gaichas, S., Gamble, R., Lucey, S., Link, J.S., 2016. Ecosystem Model Skill Assessment. Yes We Can! *PLoS ONE* 11. doi:10.1371/journal.pone.0146467
- Osman, R.W., Whitlatch, R.B., 2004. The control of the development of a marine benthic community by predation on recruits. *J. Exp. Mar. Biol. Ecol.* 311, 117–145. doi:10.1016/j.jembe.2004.05.001
- Osman, R.W., Whitlatch, R.B., 1995. Predation on early ontogenetic life stages and its effect on recruitment into a marine epifaunal community. *Oceanogr. Lit. Rev.* 9, 772.
- Pachauri, R.K., Allen, M.R., Barros, V.R., Broome, J., Cramer, W., Christ, R., Church, J.A., Clarke, L., Dahe, Q., Dasgupta, P., Dubash, N.K., Edenhofer, O., Elgizouli, I., Field, C.B., Forster, P., Friedlingstein, P., Fuglestedt, J., Gomez-Echeverri, L., Hallegatte, S., Hegerl, G., Howden, M., Jiang, K., Jimenez Cisneroz, B., Kattsov, V., Lee, H., Mach, K.J., Marotzke, J., Mastrandrea, M.D., Meyer, L., Minx, J., Mulugetta, Y., O’Brien, K., Oppenheimer, M., Pereira, J.J., Pichs-Madruga, R., Plattner, G.-K., Pörtner, H.-O., Power, S.B., Preston, B., Ravindranath, N.H., Reisinger, A., Riahi, K., Rusticucci, M., Scholes, R., Seyboth, K., Sokona, Y., Stavins, R., Stocker, T.F., Tschakert, P., van Vuuren, D., van Ypserle, J.-P., 2014. *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.* IPCC, Geneva, Switzerland.
- Paetzold, C., Giberson, D., Hill, J., Davidson, J., Davidson, J., 2012. Effect of colonial tunicate presence on *Ciona intestinalis* recruitment within a mussel farming environment. *Manag. Biol. Invasions* 3, 15–23. doi:10.3391/mbi.2012.3.1.02
- Paetzold, S.C., Davidson, J., 2011. Aquaculture fouling: Efficacy of potassium monopersulphonate triple salt based disinfectant (Virkon® Aquatic) against *Ciona intestinalis*. *Biofouling* 27, 655–665. doi:10.1080/08927014.2011.594503

- Paetzold, S.C., Davidson, J., 2010. Viability of golden star tunicate fragments after high-pressure water treatment. *Aquaculture* 303, 105–107. doi:10.1016/j.aquaculture.2010.03.004
- Patanasatienkul, T., Revie, C., Davidson, J., Sanchez, J., 2014. Mathematical model describing the population dynamics of *Ciona intestinalis*, a biofouling tunicate on mussel farms in Prince Edward Island, Canada. *Manag. Biol. Invasions* 5, 39–54. doi:10.3391/mbi.2014.5.1.04
- Pennati, R., Ficetola, G.F., Brunetti, R., Caicci, F., Gasparini, F., Griggio, F., Sato, A., Stach, T., Kaul-Strehlow, S., Gissi, C., Manni, L., 2015. Morphological Differences between Larvae of the *Ciona intestinalis* Species Complex: Hints for a Valid Taxonomic Definition of Distinct Species. *PLoS ONE* 10, e0122879. doi:10.1371/journal.pone.0122879
- Petersen, J.K., Svane, I., 1995. Larval dispersal in the ascidian *Ciona intestinalis* (L.). Evidence for a closed population. *J. Exp. Mar. Biol. Ecol.* 186, 89–102. doi:10.1016/0022-0981(94)00157-9
- Pineda, M.C., McQuaid, C.D., Turon, X., López Legentil, S., Ordóñez, V., Rius, M., 2012a. Tough adults, frail babies: an analysis of stress sensitivity across early life-history stages of widely introduced marine invertebrates. *PLoS One* 2012 Vol 7 Num 10 P E46672.
- Pineda, M.C., Turon, X., López-Legentil, S., 2012b. Stress levels over time in the introduced ascidian *Styela plicata*: the effects of temperature and salinity variations on hsp70 gene expression. *Cell Stress Chaperones* 17, 435–444. doi:10.1007/s12192-012-0321-y
- Pinheiro, J., Bates, D., 2000. *Mixed-Effects Models in S and S-PLUS*. Springer, New York.
- Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., R Core Team, 2015. *nlme: Linear and Nonlinear Mixed Effects Models*.
- Prenter, J., MacNeil, C., Dick, J.T.A., Dunn, A.M., 2004. Roles of parasites in animal invasions. *Trends Ecol. Evol.* 19, 385–390. doi:10.1016/j.tree.2004.05.002
- R Core Team, 2015. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Ramsay, A., 2009. Recruitment patterns and population development of the invasive ascidian *Ciona intestinalis* in Prince Edward Island, Canada. *Aquat. Invasions* 4, 169–176. doi:10.3391/ai.2009.4.1.17
- Ramsay, A., Davidson, J., Landry, T., Arsenault, G., 2008a. Process of invasiveness among exotic tunicates in Prince Edward Island, Canada. *Biol. Invasions* 10, 1311–1316. doi:10.1007/s10530-007-9205-y
- Ramsay, A., Davidson, J., Landry, T., Stryhn, H., 2008b. The effect of mussel seed density on tunicate settlement and growth for the cultured mussel, *Mytilus edulis*. *Aquaculture* 275, 194–200. doi:10.1016/j.aquaculture.2008.01.024
- Raw, J.L., Miranda, N.A., Perissinotto, R., 2013. Chemical cues released by an alien invasive aquatic gastropod drive its invasion success. *PLoS One* 8, e64071.
- Reinhardt, J.F., Gallagher, K.L., Stefaniak, L.M., Nolan, R., Shaw, M.T., Whitlatch, R.B., 2012. Material properties of *Didemnum vexillum* and prediction of tendrill fragmentation. *Mar. Biol.* 159, 2875–2884. doi:10.1007/s00227-012-2048-9
- Renborg, E., Johannesson, K., Havenhand, J., 2013. Variable salinity tolerance in ascidian larvae is primarily a plastic response to the parental environment. *Evol. Ecol.* 28, 561–572. doi:10.1007/s10682-013-9687-2
- Ricciardi, A., Hoopes, M.F., Marchetti, M.P., Lockwood, J.L., 2013. Progress toward understanding the ecological impacts of nonnative species. *Ecol. Monogr.* 83, 263–282. doi:10.1890/13-0183.1

- Richardson, D.M., Pysek, P., Rejmanek, M., Barbour, M.G., Panetta, F.D., West, C.J., 2000. Naturalization and Invasion of Alien Plants: Concepts and Definitions. *Divers. Distrib.* 6, 93–107. doi:10.2307/2673320
- Rigal, F., Viard, F., Ayata, S.-D., Comtet, T., 2010. Does larval supply explain the low proliferation of the invasive gastropod *Crepidula fornicata* in a tidal estuary? *Biol. Invasions* 12, 3171–3186. doi:10.1007/s10530-010-9708-9
- Rius, M., Branch, G.M., Griffiths, C.L., Turon, X., others, 2010. Larval settlement behaviour in six gregarious ascidians in relation to adult distribution. *Mar. Ecol. Prog. Ser.* 418, 151–163.
- Rius, M., Clusella-Trullas, S., McQuaid, C.D., Navarro, R.A., Griffiths, C.L., Matthee, C.A., von der Heyden, S., Turon, X., 2014. Range expansions across ecoregions: interactions of climate change, physiology and genetic diversity. *Glob. Ecol. Biogeogr.* 23, 76–88. doi:10.1111/geb.12105
- Robinson, T.B., Griffiths, C.L., McQuaid, C.D., Rius, M., 2005. Marine alien species of South Africa — status and impacts. *Afr. J. Mar. Sci.* 27, 297–306. doi:10.2989/18142320509504088
- Roux, C., Tsagkogeorga, G., Bierne, N., Galtier, N., 2013. Crossing the Species Barrier: Genomic Hotspots of Introgression between Two Highly Divergent *Ciona intestinalis* Species. *Mol. Biol. Evol.* 30, 1574–1587. doi:10.1093/molbev/mst066
- RStudio Team, 2015. RStudio: Integrated Development for R. RStudio, Inc., Boston, MA.
- Ruiz, G.M., Carlton, J.T., 2003. *Invasive Species: Vectors and Management Strategies*. Island Press, Washington.
- Ruiz, G.M., Fofonoff, P.W., Ashton, G., Minton, M.S., Miller, A.W., 2013. Geographic variation in marine invasions among large estuaries: effects of ships and time. *Ecol. Appl.* 23, 311–320. doi:10.1890/11-1660.1
- Saba, V.S., Griffies, S.M., Anderson, W.G., Winton, M., Alexander, M.A., Delworth, T.L., Hare, J.A., Harrison, M.J., Rosati, A., Vecchi, G.A., Zhang, R., 2016. Enhanced warming of the Northwest Atlantic Ocean under climate change. *J. Geophys. Res. Oceans* n/a-n/a. doi:10.1002/2015JC011346
- Salinas-de-León, P., Jones, T., Bell, J.J., 2012. Successful Determination of Larval Dispersal Distances and Subsequent Settlement for Long-Lived Pelagic Larvae. *PLoS ONE* 7, e32788. doi:10.1371/journal.pone.0032788
- Sandison, E.E., 1950. Appearance of *Elminius modestus* Darwin in South Africa. *Nature* 165, 79–80. doi:10.1038/165079b0
- Sarkar, D., 2008. *Lattice: Multivariate Data Visualization with R*. Springer, New York.
- Sato, A., Kawashima, T., Fujie, M., Hughes, S., Satoh, N., Shimeld, S.M., 2015. Molecular basis of canalization in an ascidian species complex adapted to different thermal conditions. *Sci. Rep.* 5. doi:10.1038/srep16717
- Sato, A., Satoh, N., Bishop, J.D.D., 2012. Field identification of “types” A and B of the ascidian *Ciona intestinalis* in a region of sympatry. *Mar. Biol.* 159, 1611–1619. doi:10.1007/s00227-012-1898-5
- Scheibling, R.E., Gagnon, P., 2006. Competitive interactions between the invasive green alga *Codium fragile* ssp. *tomentosoides* and native canopy-forming seaweeds in Nova Scotia (Canada). *Mar. Ecol. Prog. Ser.* 325, 1–14.
- Scrosati, R., Eckersley, L.K., 2007. Thermal insulation of the intertidal zone by the ice foot. *J. Sea Res.* 58, 331–334. doi:10.1016/j.seares.2007.08.003
- Seebens, H., Gastner, M.T., Blasius, B., 2013. The risk of marine bioinvasion caused by global shipping. *Ecol. Lett.* 16, 782–790. doi:10.1111/ele.12111

- Sephton, D., Ouellette-Plante, J., Vercaemer, B., 2014. Biofouling monitoring for aquatic invasive species (AIS) in DFO Maritimes Region, Nova Scotia: May – December 2010 (No. 3034), Can. Tech. Rep. Fish Aquat. Sci.
- Sephton, D., Stiles, L., Vercaemer, B., 2015. Biofouling monitoring for aquatic invasive species (AIS) in DFO Maritimes Region, Nova Scotia: May – December 2011 (No. 3082), Can. Tech. Rep. Fish Aquat. Sci.
- Sephton, D., Vercaemer, B., Nicolas, J.M., Keays, J., 2011. Monitoring for invasive tunicates in Nova Scotia, Canada (2006-2009). *Aquat. Invasions* 6, 391–403. doi:10.3391/ai.2011.6.4.04
- Serafini, L., Hann, J.B., Kültz, D., Tomanek, L., 2011. The proteomic response of sea squirts (genus *Ciona*) to acute heat stress: A global perspective on the thermal stability of proteins. *Comp. Biochem. Physiol. Part D Genomics Proteomics* 6, 322–334. doi:10.1016/j.cbd.2011.07.002
- Shimoda, Y., Arhonditsis, G.B., 2016. Phytoplankton functional type modelling: Running before we can walk? A critical evaluation of the current state of knowledge. *Ecol. Model.* 320, 29–43. doi:10.1016/j.ecolmodel.2015.08.029
- Shmueli, G., 2010. To explain or to predict? *Stat. Sci.* 289–310.
- Shumway, S.E., 1978. Respiration, pumping activity and heart rate in *Ciona intestinalis* exposed to fluctuating salinities. *Mar. Biol.* 48, 235–242. doi:10.1007/BF00397150
- Sievers, M., Fitridge, I., Dempster, T., Keough, M.J., 2013. Biofouling leads to reduced shell growth and flesh weight in the cultured mussel *Mytilus galloprovincialis*. *Biofouling* 29, 97–107. doi:10.1080/08927014.2012.749869
- Simberloff, D., 2011. Non-natives: 141 scientists object. *Nature* 475, 36–36. doi:10.1038/475036a
- Skrondal, A., Rabe-Hesketh, S., 2004. Generalized Latent Variable Modeling: Multilevel, Longitudinal, and Structural Equation Models. CRC Press.
- Stow, C.A., Roessler, C., Borsuk, M.E., Bowen, J.D., Reckhow, K.H., 2003. Comparison of Estuarine Water Quality Models for Total Maximum Daily Load Development in Neuse River Estuary. *J. Water Resour. Plan. Manag.* 129, 307–314. doi:10.1061/(ASCE)0733-9496(2003)129:4(307)
- Subedi, N., Sharma, M., 2011. Individual-tree diameter growth models for black spruce and jack pine plantations in northern Ontario. *For. Ecol. Manag.* 261, 2140–2148. doi:10.1016/j.foreco.2011.03.010
- Suzuki, M.M., Nishikawa, T., Bird, A., 2005. Genomic Approaches Reveal Unexpected Genetic Divergence Within *Ciona intestinalis*. *J. Mol. Evol.* 61, 627–635. doi:10.1007/s00239-005-0009-3
- Svane, I., Havenhand, J.N., 1993. Spawning and Dispersal in *Ciona intestinalis* (L.). *Mar. Ecol.* 14, 53–66. doi:10.1111/j.1439-0485.1993.tb00364.x
- Svensson, J.R., Nylund, G.M., Cervin, G., Toth, G.B., Pavia, H., 2013. Novel chemical weapon of an exotic macroalga inhibits recruitment of native competitors in the invaded range. *J. Ecol.* 101, 140–148. doi:10.1111/1365-2745.12028
- Sylvester, F., Kalaci, O., Leung, B., Lacoursière-Roussel, A., Murray, C.C., Choi, F.M., Bravo, M.A., Therriault, T.W., MacIsaac, H.J., 2011. Hull fouling as an invasion vector: can simple models explain a complex problem? *J. Appl. Ecol.* 48, 415–423. doi:10.1111/j.1365-2664.2011.01957.x
- Teixidó, N., Albajes-Eizagirre, A., Bolbo, D., Le Hir, E., Demestre, M., Garrabou, J., Guigues, L., Gili, J., Píera, J., Prelot, T., Soria-Frisch, A., 2011. Hierarchical segmentation-based software for cover classification analyses of seabed images (Seascape). *Mar. Ecol. Prog. Ser.* 431, 45–53. doi:10.3354/meps09127

- Therriault, T.W., Herborg, L.-M., 2008a. Predicting the potential distribution of the vase tunicate *Ciona intestinalis* in Canadian waters: informing a risk assessment. *ICES J. Mar. Sci. J. Cons.* 65, 788–794. doi:10.1093/icesjms/fsn054
- Therriault, T.W., Herborg, L.-M., 2008b. A qualitative biological risk assessment for vase tunicate *Ciona intestinalis* in Canadian waters: using expert knowledge. *ICES J. Mar. Sci. J. Cons.* 65, 781–787.
- Thompson, R., 2015. Climate sensitivity. *Earth Environ. Sci. Trans. R. Soc. Edinb.* 106, 1–10. doi:10.1017/S1755691015000213
- Todgham, A.E., Stillman, J.H., 2013. Physiological Responses to Shifts in Multiple Environmental Stressors: Relevance in a Changing World. *Integr. Comp. Biol.* 53, 539–544. doi:10.1093/icb/ict086
- Trygonis, V., Sini, M., 2012. photoQuad: A dedicated seabed image processing software, and a comparative error analysis of four photoquadrat methods. *J. Exp. Mar. Biol. Ecol.* 424–425, 99–108. doi:10.1016/j.jembe.2012.04.018
- Vercaemer, B., Sephton, D., Nicolas, J.M., Howes, S., Keays, J., 2011. *Ciona intestinalis* environmental control points: field and laboratory investigations. *Aquat. Invasions* 6, 477–490.
- Verling, E., Ruiz, G.M., Smith, L.D., Galil, B., Miller, A.W., Murphy, K.R., 2005. Supply-side invasion ecology: characterizing propagule pressure in coastal ecosystems. *Proc. R. Soc. Lond. B Biol. Sci.* 272, 1249–1257. doi:10.1098/rspb.2005.3090
- Visser, P.J., 1927. Nature and extent of fouling of ships' bottoms. *Bull. Bur. Fish.* 43, 193–252.
- Vizzini, A., Di Falco, F., Parrinello, D., Sanfratello, M.A., Mazzarella, C., Parrinello, N., Cammarata, M., 2015. *Ciona intestinalis* interleukin 17-like genes expression is upregulated by LPS challenge. *Dev. Comp. Immunol.* 48, 129–137. doi:10.1016/j.dci.2014.09.014
- Wickham, H., 2011. The Split-Apply-Combine Strategy for Data Analysis. *J. Stat. Softw.* 40, 1–29.
- Wickham, H., 2009. *ggplot2: elegant graphics for data analysis*. Springer, New York.
- Williams, S.L., Smith, J.E., 2007. A Global Review of the Distribution, Taxonomy, and Impacts of Introduced Seaweeds. *Annu. Rev. Ecol. Evol. Syst.* 38, 327–359.
- Williamson, M., Fitter, A., 1996. The Varying Success of Invaders. *Ecology* 77, 1661–1666. doi:10.2307/2265769
- Woods, C.M., Floerl, O., Hayden, B.J., 2012. Biofouling on Greenshell™ mussel (*Perna canaliculus*) farms: a preliminary assessment and potential implications for sustainable aquaculture practices. *Aquac. Int.* 20, 537–557.
- Wootton, J.T., 2002. Mechanisms of successional dynamics: Consumers and the rise and fall of species dominance. *Ecol. Res.* 17, 249–260. doi:10.1046/j.1440-1703.2002.00484.x
- Wu, R., Ma, C.-X., Chang, M., Li, R.C., Wu, S.S., Yin, T., Huang, M., Wang, M., Casella, G., 2002. A logistic mixture model for characterizing genetic determinants causing differentiation in growth trajectories. *Genet. Res.* 79, 235–245. doi:10.1017/S0016672302005633
- Wu, R., Ma, C.-X., Yang, M.C.K., Chang, M., Li, R.C., Santra, U., Wu, S.S., Yin, T., Huang, M., Wang, M., Casella, G., 2003. Quantitative trait loci for growth trajectories in *Populus*. *Genet. Res.* 81, 51–64. doi:10.1017/S0016672302005980
- Yamaguchi, M., 1975. Growth and reproductive cycles of the marine fouling ascidians *Ciona intestinalis*, *Styela plicata*, *Botrylloides violaceus*, and *Leptoclinum mitsukurii* at Aburatsubo-Moroiso Inlet (central Japan). *Mar. Biol.* 29, 253–259. doi:10.1007/BF00391851

- Zambrano-Bigiarini, M., 2014. hydroGOF: Goodness-of-fit functions for comparison of simulated and observed hydrological time series.
- Zhan, A., Darling, J.A., Bock, D.G., Lacoursière-Roussel, A., MacIsaac, H.J., Cristescu, M.E., 2012. Complex genetic patterns in closely related colonizing invasive species. *Ecol. Evol.* 2, 1331–1346. doi:10.1002/ece3.258
- Zhan, A., Macisaac, H.J., Cristescu, M.E., 2010. Invasion genetics of the *Ciona intestinalis* species complex: from regional endemism to global homogeneity. *Mol. Ecol.* 19, 4678–4694. doi:10.1111/j.1365-294X.2010.04837.x
- Zhao, Y., Wang, M., Lindström, M.E., Li, J., 2015. Fatty Acid and Lipid Profiles with Emphasis on n-3 Fatty Acids and Phospholipids from *Ciona intestinalis*. *Lipids* 50, 1009–1027. doi:10.1007/s11745-015-4049-1

## Appendices

### Appendix A: Deployment and monitoring schedules

**Table A.1: Dates of deployment and site visits for 2014. Note that Venus Cove was deployed at the end of May and Cape Canso near the end of June. As a result, Cape Canso has one less site visit, the first of which corresponds to the second site visit at all other sites, and so on.**

Site	Deployment	Visit 1	Visit 2	Visit 3	Visit 4	Visit 5
Ship Harbour	1-May	5-Jun	3-Jul	31-Jul	5-Sep	6-Oct
Indian Point	2-May	6-Jun	3-Jul	31-Jul	5-Sep	6-Oct
Camp Cove	3-May	7-Jun	4-Jul	1-Aug	6-Sep	6-Oct
Fall's Point	3-May	7-Jun	4-Jul	1-Aug	6-Sep	6-Oct
Yarmouth Bar	3-May	6-Jun	4-Jul	2-Aug	7-Sep	7-Oct
Wedgeport	4-May	6-Jun	4-Jul	1-Aug	6-Sep	7-Oct
St Ann's Bay	7-May	10-Jun	7-Jul	5-Aug	11-Sep	8-Oct
Little River	7-May	10-Jun	7-Jul	5-Aug	11-Sep	8-Oct
Dingwall	7-May	10-Jun	7-Jul	5-Aug	11-Sep	9-Oct
Petit-de-Grat	8-May	11-Jun	6-Jul	3-Aug	10-Sep	12-Oct
Port Bickerton	9-May	12-Jun	8-Jul	4-Aug	9-Sep	9-Oct
Venus Cove	29-May	11-Jun	6-Jul	3-Aug	10-Sep	9-Oct
Cape Canso	25-Jun	NA	8-Jul	4-Aug	9-Sep	11-Oct

**Table A.2: Dates of deployment and site visits for 2015. Note the three new sites, SP having since been removed from consideration in the analyses. Deployment was roughly a week later in 2015 compared to 2014, due to ice-cover in harbours later in the year.**

Site	Deployment	Visit 1	Visit 2	Visit 3	Visit 4	Visit 5
Ship Harbour	12-May	5-Jun	6-Jul	3-Aug	31-Aug	5-Oct
Indian Point	8-May	5-Jun	6-Jul	31-Jul	31-Aug	5-Oct
Camp Cove	7-May	6-Jun	3-Jul	1-Aug	1-Sep	2-Oct
Ingomar	7-May	6-Jun	3-Jul	2-Aug	1-Sep	2-Oct
Fall's Point	7-May	6-Jun	3-Jul	2-Aug	1-Sep	2-Oct
Yarmouth Bar	8-May	6-Jun	2-Jul	1-Aug	2-Sep	3-Oct
Wedgeport	8-May	6-Jun	2-Jul	1-Aug	1-Sep	3-Oct
St Ann's Bay	13-May	11-Jun	9-Jul	6-Aug	8-Sep	26-Oct



**Table A.2. cont.**

Little River	11-May	9-Jun	9-Jul	6-Aug	7-Sep	9-Oct
Dingwall	11-May	9-Jun	9-Jul	6-Aug	7-Sep	9-Oct
Petit-de-Grat	13-May	10-Jun	8-Jul	5-Aug	4-Sep	7-Oct
Port Bickerton	10-May	8-Jun	7-Jul	3-Aug	3-Sep	6-Oct
Venus Cove	10-May	11-Jun	5-Jul	4-Aug	5-Sep	7-Oct
Cape Canso	10-May	8-Jun	5-Jul	4-Aug	3-Sep	6-Oct
Louisbourg	11-May	11-Jun	10-Jul	5-Aug	8-Sep	8-Oct
St Peters	13-May	10-Jun	8-Jul	5-Aug	4-Sep	8-Oct

## Appendix B: Additional sampling procedures

### *B.1. Chlorophyll-a point-sample data*

#### B.1.1. Obtaining the water sample

Point sample measures of the chlorophyll-*a* (chl-*a*) concentration at each site were taken during each monthly site visit during the 2014 field season. Three water samples were taken during each visit from a depth of approximately 1.5 m using a Van Dorn Bottle water sampler. Each sample was about 500 ml in volume. The three water bottles used to house to 500 ml water samples were rinsed out with the first portion of the water sample from the Van Dorn Bottle to ensure there was no dilution or contamination from previous samples. This was unlikely regardless, as the water bottles were cleaned between all sites to minimize the potential for unwanted species introductions. Once the water sample was stored in the bottle, they were wrapped in aluminium foil, labelled, and placed in a cooler box. These steps are to prevent light getting into the water sample and to reduce heat exposure, which could degrade the chl-*a* in the sample.

#### B.2.1. Water sample filtration, chl-*a* extraction, and sample preservation.

Filtration of the water sample to remove and store the chl-*a* was conducted as soon as possible, always within four hours of the water sample being taken. A clean graduated cylinder was rinsed with each water sample prior to 100 ml of the sample being measured. A hand-pump filtration system was used to filter the water. A 25 mm glass fibre filter (GFF) was placed in its holder in the filtration system. The filtration beaker was rinsed with a portion of the original water sample before being locked into place above the filtration paper,

creating an air-tight seal. The GFF holder and the filtration beaker fit on top of a container that receives the filtrate. The 100 ml water sample was poured into the beaker and the hand-pump is squeezed until sufficient pressure is generated to slowly filter the water sample through the GFF. Once the water had fully passed through the paper, the beaker is unlocked and the GFF is removed using a forceps and placed into a scintillation vial with 10 ml of 90 % acetone solution. Each scintillation vial was clearly labelled, covered with aluminium foil to eliminate any irradiance and then placed in a cooler box until it could be stored in a freezer for long-term storage.

### B.3.1. Fluorometry

Fluorometric analysis was carried out on the chl-*a* extractions using a Turner Design Model 10-005 R Fluorometer at the Bedford Institute of Oceanography (BIO), Nova Scotia. Samples were transported to BIO in a cooler and allowed to sit at room temperature for an hour prior to analysis. A control ‘blank’ is run first, consisting of a cuvette with filled with 90% acetone only. For sample measurement, I began by rinsing the cuvette with 90% acetone twice. After shaking the scintillation vial to ensure the solution is homogenous, I rinsed the cuvette with the sample twice before filling the cuvette. I took an initial reading from the fluorometer output display, adjusting the sensitivity switches as required (these allow for very low and high quantities of chl-*a* to be detected using the same machine) before adding two drops of HCl and taking a second reading. I calculated the chl-*a* measurements (mg/l) according to the latest fluorometer calibration equation and the readings taken from the machine.

### B.4.1. Results

**Table B.1: Chl-*a* measurement results from fluorometry analysis presented for each site and date of measurement.**

Site code	Date	Mean chl- <i>a</i> (mg/l)	Standard error	No. Samples
CC	5/3/14	0.95	0.02	3
CC	6/7/14	1.58	0.01	3
CC	7/4/14	2.99	0.60	3
CC	8/1/14	0.96	0.48	3
CC	9/6/14	2.63	0.01	3
CC	10/6/14	1.85	0.35	3
DW	5/7/14	0.54	0.02	3
DW	6/10/14	0.74	0.03	3
DW	7/7/14	2.54	0.13	3
DW	8/5/14	5.42	0.55	3
DW	9/11/14	1.34	NA	1

DW	10/9/14	0.69	NA	1
<b>Table B.1. cont.</b>				
FP	5/3/14	0.74	0.06	3
FP	6/7/14	1.58	0.06	3
FP	7/4/14	0.37	0.09	3
FP	8/1/14	1.75	0.33	3
FP	9/6/14	1.92	0.23	3
FP	10/6/14	0.71	0.06	3
IP	5/2/14	0.84	0.35	4
IP	6/6/14	0.31	0.12	3
IP	7/3/14	0.41	0.02	3
IP	7/31/14	0.70	0.04	3
IP	9/5/14	0.43	0.02	3
IP	10/6/14	0.91	0.07	3
LR	5/7/14	0.25	0.10	3
LR	6/10/14	0.65	0.07	3
LR	7/7/14	1.98	0.76	3
LR	8/5/14	1.00	0.04	3
LR	9/11/14	17.91	NA	1
LR	10/8/14	0.34	NA	1
PB	5/9/14	0.56	0.07	3
PB	5/16/14	0.36	0.03	3
PB	6/1/14	0.88	0.09	3
PB	6/12/14	0.79	0.07	3
PB	7/8/14	1.19	0.27	3
PB	8/4/14	1.38	0.37	3
PB	9/9/14	0.78	0.02	3
PB	10/9/14	0.95	NA	1
PG	5/8/14	0.36	0.02	3
PG	5/16/14	0.27	0.02	3
PG	6/11/14	1.67	0.78	3
PG	7/6/14	3.60	2.46	3
PG	8/3/14	4.26	0.83	3
PG	9/10/14	0.86	0.03	3
PG	10/12/14	2.17	NA	1
SA	5/7/14	2.18	0.16	3
SA	6/10/14	0.49	0.01	3
SA	7/7/14	1.21	0.05	3
SA	8/5/14	0.89	0.04	3
SA	9/11/14	2.32	NA	1
SA	10/8/14	2.84	NA	1
SH	5/1/14	1.40	0.13	3
SH	6/5/14	1.14	0.05	3
SH	7/3/14	1.78	0.03	3
SH	7/31/14	1.11	0.07	3
SH	9/5/14	0.99	0.08	3

SH	10/6/14	1.74	0.28	2
<b>Table B.1. cont.</b>				
VC	5/29/14	0.46	0.11	3
VC	6/11/14	0.79	0.13	3
VC	7/6/14	1.66	0.19	3
VC	8/3/14	2.19	0.10	3
VC	9/10/14	1.42	0.02	3
VC	10/9/14	2.36	NA	1
WP	5/4/14	1.94	0.23	3
WP	6/6/14	1.72	0.09	3
WP	7/4/14	2.19	0.08	3
WP	8/1/14	1.73	0.14	3
WP	9/6/14	2.53	0.07	3
WP	10/7/14	2.91	0.57	3
YB	5/3/14	1.54	0.26	2
YB	6/6/14	1.42	0.10	3
YB	7/4/14	4.20	0.13	3
YB	8/2/14	2.12	0.17	3
YB	9/7/14	3.83	0.04	3
YB	10/7/14	3.00	0.37	3
ZZ	6/25/14	0.59	0.02	3
ZZ	7/8/14	0.52	0.03	3
ZZ	8/4/14	0.85	0.04	3
ZZ	9/9/14	1.61	0.04	3
ZZ	10/11/14	0.40	NA	1

## B.2. *Ciona intestinalis* sample collection for genetic analysis

During the final round of site visits in October of both years, *C. intestinalis* samples were collected from all sites where it was present. At sites where *C. intestinalis* was in low abundance, all individuals of a length greater than 1 cm were collected. The specimens were carefully removed from the settlement plates ensuring the tunics of the animals were not torn. Gently, any excess water that it was possible to squeeze out of the samples was removed. This was to minimize dilution of the 100% ethanol that the samples were placed in for storage. The samples in ethanol were kept in a cooler until they could be stored in a freezer upon return to the laboratory. At sites with high *C. intestinalis* abundance, roughly 30 individuals greater than 1 cm in length were collected. Where possible, an even number of individuals were removed from multiple settlement plates with no specific selection preference for larger individuals, except for the greater than 1 cm criterion. The 2014 samples were sent to Dr Aibin Zhan, a collaborator at the Research Center for Eco-Environmental

Science at the Chinese Academy of Sciences in Beijing, China. Genetic sequencing has been carried out on the 2014 samples following the methods in Zhan et al. (2010). The 2015 samples are awaiting transfer to China.

## **Appendix C: Multivariate community analysis**

### *C.1. Introduction*

While the main focus of this study was *C. intestinalis* growth and has been modelled in isolation using the NLME model approach, there are other members of the biofouling community, native and introduced, that share (some of) the space. The of the most prominent species co-existing on the settlement plates with *C. intestinalis* were two other invasive tunicate species, the colonial species *Botryllus schlosseri* and *Botrylloides violaceus*. In addition to invasive tunicates, the invasive *Membranipora membranacea* was also relatively common, if not abundant. The remainder of the community was made up of other bryozoan species, hydroid species, *Mytilus edulis*, *Balanus* sp., and occasionally the odd small sea anemone, tubeworm. There were also times when a single unidentifiable species would occur on the plates. In 2014, at one site, VC, *Styela clava* was present, but it was not there again in 2015. In 2015, although not included in the following analysis, a molgulid species was present in high abundance at Louisbourg.

To conduct analysis that test the degree of similarity between the biological communities at each site during each measurement period, PRIMER v7 and PERMANOVA+ for PRIMER were used (Clarke and Gorley, 2015). This program also allows for the incorporation of the abiotic data to explain how much of the variation in the biological community is due to abiotic variables, and which ones are the most important. It is a program that relies predominantly on non-metric and permutation based forms of analysis and as such makes very little assumption about the form of the response or independent data. The permutation is carried out by randomly reordering the sample names the specified amount of time, while a test statistic is calculate for each random permutation. This creates a frequency distribution of the test statistics, which can then be used to determine the importance of the test statistic calculated according to the observed data. This is advantageous as there is no desire for the biological data to be normally distributed. The end result is a robust and widely applicable range of analyses for data exploration and testing. The basic premise behind the program is to create dis/similarity matrices between samples so that the similarity, or lack thereof, between

biological communities can be used for a range of ordinations and multivariate analyses. Then abiotic data can be used to interpret the relationships

### *C.2. Methods*

The biological data used for all of the analyses is the species percentage cover results by sample (settlement plate within measurement period within site). I only included the data from the core sites - minus absent, as although this is a community analysis, I still want to explain how the abiotic data influences the biological community at sites where *C. intestinalis* was present. For each year, the abiotic data used were the same abiotic variable metrics from the preselection process in the NLME model development (section 2.7.2.). For the same reasons this was carried out before the NLME model, this is to reduce collinearity and eliminate redundant variables that would lead to overly complicated models. In 2014, the four abiotic variable metrics used were mean temperature, mean salinity, minimum pH, and minimum daily variance in acceleration (Table 3.6). In 2015, the four abiotic variable metrics used were mean temperature, maximum salinity, maximum pH, and maximum daily variance in acceleration (Table 3.9).

The sparsity of species presence at the beginning of an observational study of growth over time such as this can cause some trouble PRIMER. The lack of species presence for two samples means that when they are compared using a Bray-Curtis similarity matrix the coefficient between the two samples is undefined, as joint absences are ignored. Or if there are only one or two species and low abundances of these species, the resultant similarities calculated in the matrix can vary wildly from 0 – 100 % similar based on very little difference. If the only difference between two plates is a single occurrence of a barnacle, the Bray-Curtis similarity matrix would define the two as 0% similar, despite the difference in reality being very marginal. The best way to account for both undefined coefficients and highly irregular similarity measures at the beginning of the observational study, a zero-adjustment can be made to the Bray-Curtis equation quite simply (Clarke and Gorley, 2015). I used a modified similarity measure, adding a single dummy variable to the biological dataset before the Bray-Curtis similarity matrices were calculated to account for these aforementioned issues.

The biological data and abiotic data were imported into PRIMER v7 as separate datasheets. No data transformation was implemented on the biological data. A resemblance matrix was created for the biological data using a Bray-Curtis similarity measure between each sample,

with a single dummy variable added at this point. This results in a matrix with a value of similarity between every single sample.

A draftsman plot was used to visually assess relationships between the abiotic metrics, as a follow up to the preselection of least correlated metrics. The abiotic metrics did not require any transformation in any cases. A duplicate copy of the abiotic dataset was created, and then the ‘missing’ tool in PRIMER v7 was used on the duplicate to create an abiotic dataset without missing values. The ‘missing’ tool using an expectation-maximisation (EM) algorithm, a standard method for continuous environmental variables, was used to replace missing values (Clarke and Gorley, 2015). This was required for the DistLM procedure, as it cannot be performed with missing data are present in the abiotic dataset. The BEST procedure has no such requirement, but I performed BEST analysis using both the dataset with missing values and the dataset without missing values to investigate the difference the ‘missing’ procedure had. This allowed me to inspect the differences in the BEST results to judge how much of an impact using the ‘missing’ tool had and the implications for the DistLM results. On both datasets a normalization procedure was carried out to normalize the scale of the four different abiotic variable metrics. The final step here can be to create a resemblance matrix, much like the one created for the biological data, except a Euclidean distance measure is used to calculate the similarity between samples for environmental data rather than the Bray-Curtis. The results are not presented here, but comparing the biological and abiotic resemblance matrices is a basic exploratory analysis to inspect the degree that the matrices relate to one another. The following procedures described go a step further and interpret the proportion of importance that each abiotic metric has in explaining the biological patterns of similarity.

#### C.2.1. BEST

I used the BEST procedure in PRIMER v7 to match the biological patterns with the best explanatory abiotic metrics. There are two methods that can be implemented within BEST. The BVSTEP method is a stepwise approach of testing abiotic metrics and is useful if there are a lot of metrics. I used the BIOENV method as I was only including four abiotic metrics and this method tests all combinations of the metrics. The patterns in the biological resemblance matrices are compared to the patterns in the corresponding resemblance matrix of abiotic data, in this case either the normalized data with or without missing values. I used the Kendall’s tau correlation coefficient method to assess the pattern similarities, as it is

appropriate for time-series data and also tends to provide a more conservative coefficient compared to Spearman or Pearson. The number of permutation tests was set to 999. BEST allows for within-level factors to be assigned. I used the measurement period as a within-level factor. This within-level factor allows the procedure to only compare samples within the same level of the factor, not with samples of other levels. This was a way of accounting the effect of time on the pattern of growth in the biological communities, allowing for the method to explain the amount of variation in the biological data within a given measurement period with the corresponding abiotic data for that measurement period. This does not take into account the covariance and correlation that is likely to occur within and between measurement periods.

### C.2.2. Distance-based linear modelling (DistLM)

DistLM is a routine used to analyse and model relationship between a multivariate data cloud (a biological resemblance matrix and one or more abiotic predictor variables) using method called distance-based redundancy analysis (dbRDA) (Clarke and Gorley, 2015). This is similar to a multiple regression analysis and unlike the BEST procedure, the DistLM provides a modelled measure of the explained variance in these data, opposed to a simple correlation result from BEST. Like the BEST procedure, the biological resemblance matrix is used, while the normalized abiotic data that has had the missing data replaced via the EM algorithm is used as the predictor variable data. There are a number of different methods that can be used for the predictor variable selection. I used the ‘Best’ method, where all of the abiotic metric combinations are tested. AICc was used as the selection criterion for ranking the models while 999 permutations were used. A dbRDA plot was also created to visualise the biological data’s ordination in relation to the fitted models using the predictor abiotic metrics. DistLM does not have the capacity to incorporate a within-level factor.

## C.3. Results

### C.3.1. Cumulative biological data - 2014

#### C.3.1.1. BEST 2014

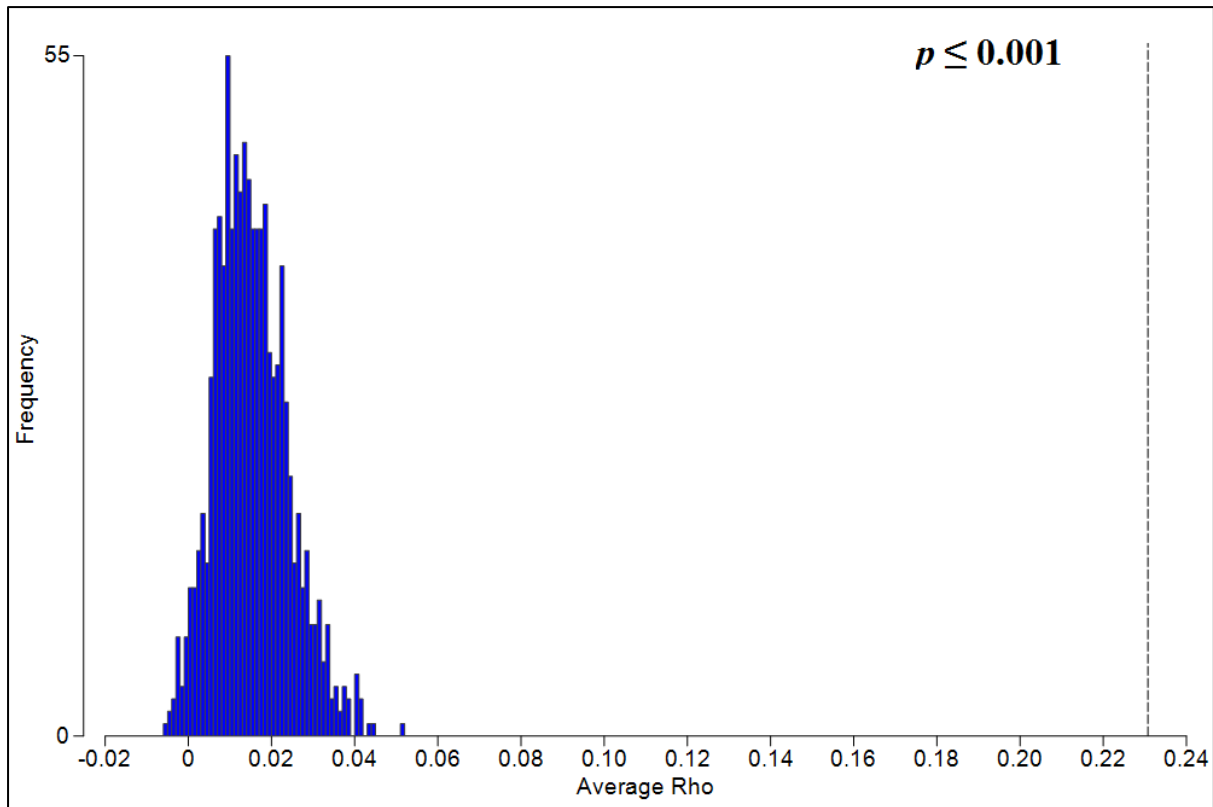
**Table C.1: Best abiotic variable combinations at explaining the cumulative biological community data from 2014, with measurement period as a within-level factor in the BIOENV procedure. Dataset contains missing abiotic values.**

No. abiotic metrics	Best metric combination	Kendall’s tau ( $\tau$ )
---------------------	-------------------------	--------------------------



1	Mean salinity	0.23
3	Mean sal + Min pH + Mean temp	0.21
2	Mean salinity + Min pH	0.2
2	Mean salinity + Mean temperature	0.19
2	Mean salinity + Min var accel	0.17
4	Mean sal + Min pH + Mean temp + Min var accel	0.16
3	Mean sal + Min pH + Min var accel	0.16
3	Mean sal + Mean temp + Min var accel	0.16
1	Min pH	0.15
2	Min pH + Mean temp	0.15

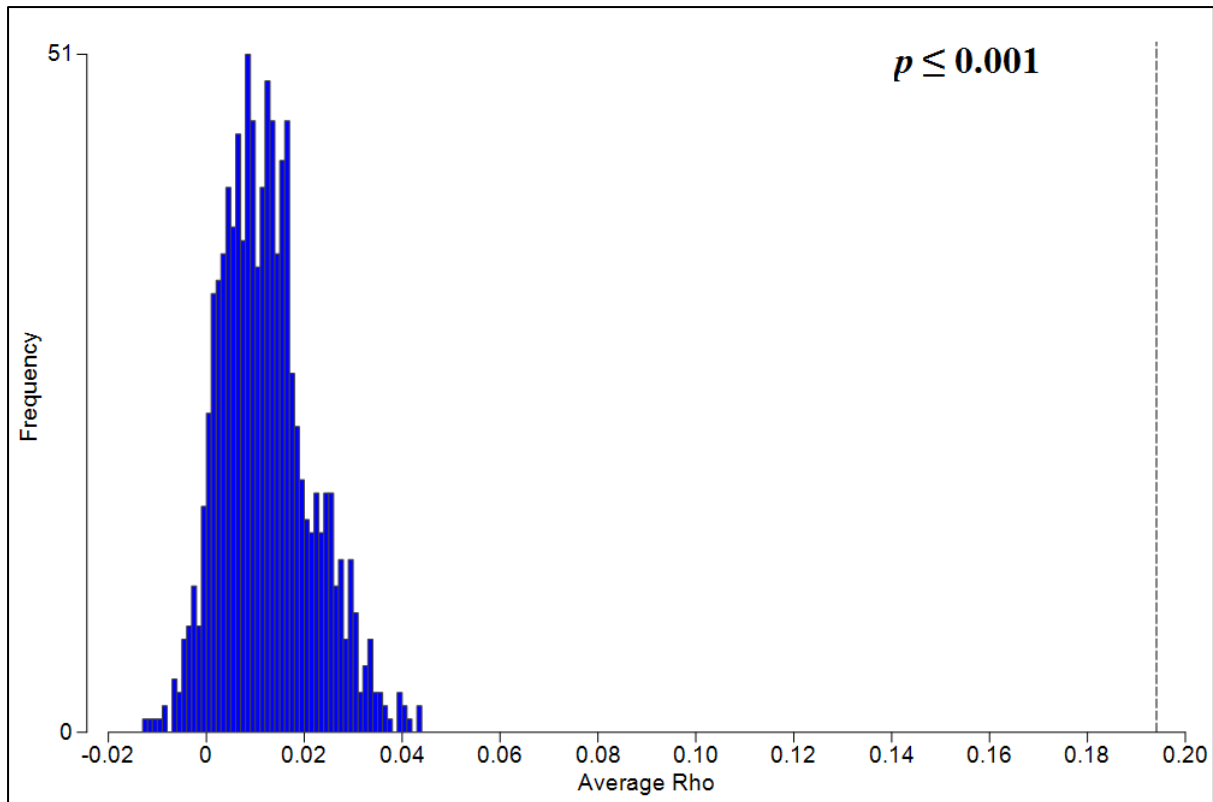
The BIO-ENV procedure indicates that the patterns in the 2014 biological community, within each measurement period, are best correlated to mean salinity ( $\tau = 0.23$ , Table C.1,  $p \leq 0.001$ , Fig. C.1). The correlation coefficients were similar for combinations that included mean temperature and minimum pH with mean salinity (Table C.2). When missing data was replaced, the abiotic metrics with the highest correlation with the pattern in the biological community, within each measurement period, are mean temperature and mean salinity ( $\tau = 0.19$ , Table C.2,  $p \leq 0.001$ , Fig. C.2). A variety of combinations including minimum pH, mean temperature and mean salinity all produce very similar correlation coefficients (Table C2.).



**Figure C.1: The distribution of rho values from the 999 permutation tests carried out for the 2014 BIOENV procedure using the dataset with missing values. The rho value (dotted vertical line) of the best predictor (Mean salinity, Table C.1) variable and its significance are displayed.**

**Table C.2: Best abiotic variables at explaining the cumulative biological community data from 2014 with measurement period as a factor within the BIOENV procedure. Dataset contains no missing abiotic values (replacement via EM algorithm).**

No. of abiotic metrics	Best metric combination	Kendall's tau
2	Mean sal + Mean temp	0.19
3	Mean sal + Mean temp + Min pH	0.19
1	Mean sal	0.19
2	Mean sal + Min pH	0.19
4	Mean sal + Mean temp + Min pH + Min var accel	0.17
3	Mean sal + Min pH + Min var accel	0.16
2	Mean sal + Min var accel	0.16
3	Mean sal + Mean temp + Min var accel	0.16
2	Mean temp + Min pH	0.12
1	Min pH	0.11



**Figure C.2:** The distribution of rho values from the 999 permutation tests carried out for the 2014 BIOENV procedure using the dataset with no missing values (missing data replaced using EM algorithm). The rho value (dotted vertical line) of the best predictor variables (Mean salinity + Mean temp, Table C.2) and its significance are displayed.

*C.3.1.2. DistLM 2014*

**Table C.3:** Marginal test results from the DistLM analysis in 2014.

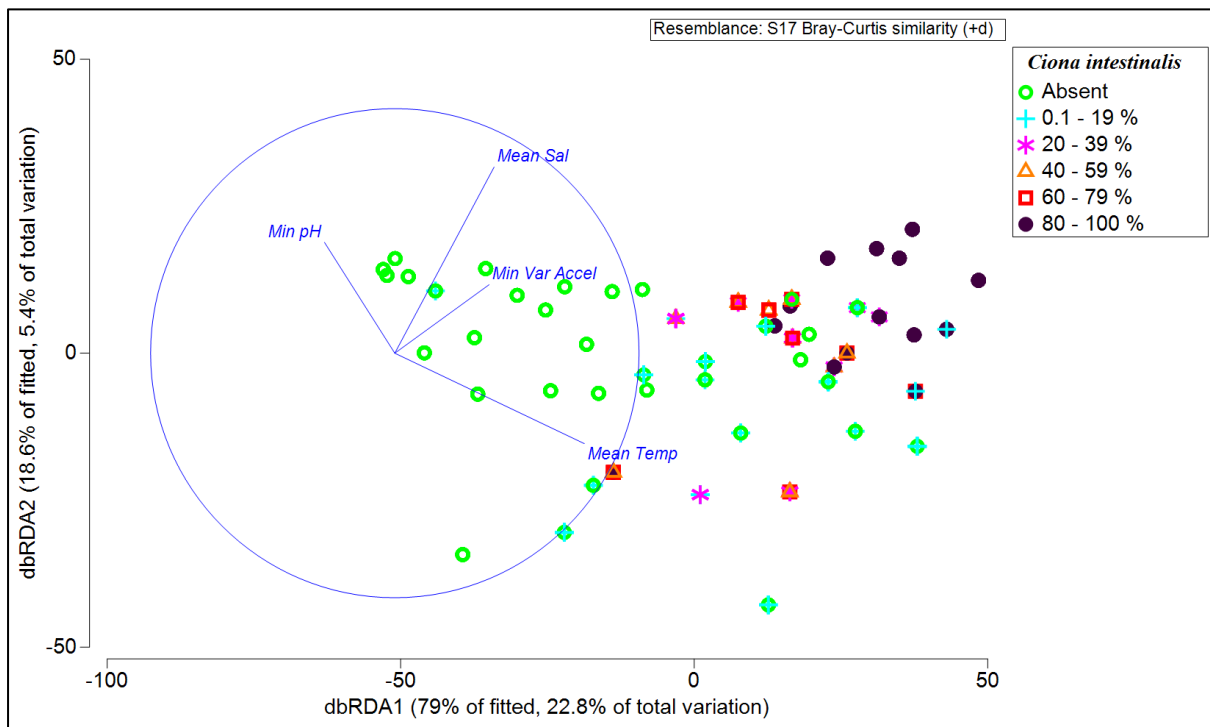
Variable	Sum of squares	Pseudo-F	P-value	Proportion of variation explained
Min pH	1.3 E+05	39.5	0.001	0.073353
Min Var Accel	1.17 E+05	35.358	0.001	0.066168
Mean Sal	1.33 E+05	40.354	0.001	0.074819
Mean Temp	3.33 E+05	115.55	0.001	0.18803

The marginal test results (the relationship between response variable and individual explanatory variables in isolation) from the DistLM analysis carried out on the 2014 data show that mean temperature explains the most variation in the biological community variation by far, at ~ 19 % (Table C.3). The other three metrics all explain ~ 7 % of the variation in the biological community variation, when considered in isolation. All four metrics are significant ( $P \leq 0.001$ , Table C.3). The BEST method of testing which combination of abiotic metrics performed the best within the DistLM suggests that including all four of the abiotic metrics in the final model provides the best model fit (Table C.4). This

model explains 29 % of the variation in the biological community, while a series of models containing either two or three of the abiotic metrics provide similar measures of explained variance, but they ranked lower according to AICc comparison (Table C.4).

**Table C.4: Best models in 2014 from the DistLM procedure according to AICc ranking.**

AICc	R <sup>2</sup>	No. abiotic metrics	Selections
3933.2	0.29	4	Mean temp + Mean sal + Min var accel + Min pH
3935.8	0.28	3	Mean temp + Mean sal + Min var accel
3942.3	0.27	3	Mean temp + Mean sal + Min pH
3949.2	0.26	2	Mean temp + Mean sal
3976.2	0.22	3	Mean temp + Min var accel + Min pH
3977.9	0.22	2	Mean temp + Min var accel
3980.3	0.22	3	Mean sal + Min var accel + Min pH
3990.1	0.2	2	Mean temp + Min pH
3993.2	0.19	1	Mean temp
4014.7	0.16	2	Min var accel + Min pH



**Figure C.3: Ordination of the 2014 cumulative biological data according to distance-based redundancy analysis using the abiotic data as predictor variables. A vector**

overlay represents the distance and proportion of the abiotic predictor effects from the best model (Table C.3).

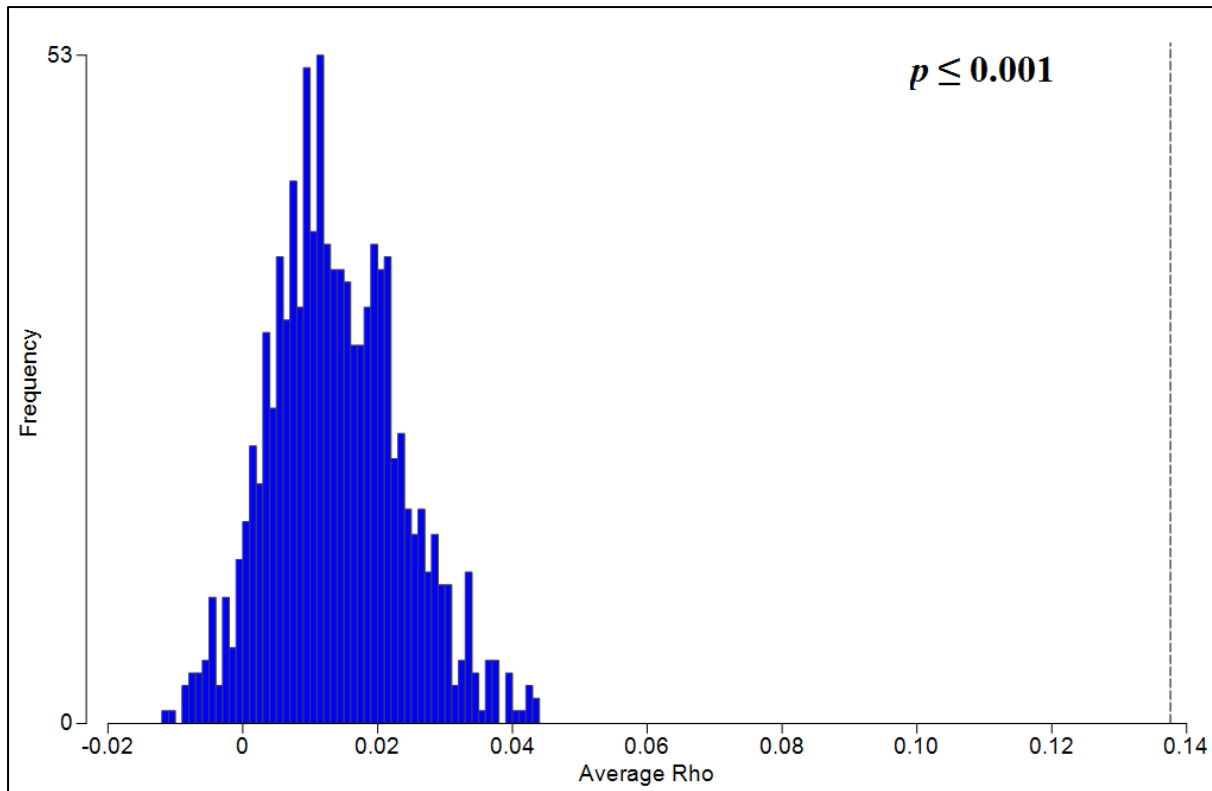
### C.3.2. Cumulative biological data - 2015

#### C.3.2.1. BEST 2015

The BEST procedure indicates that the patterns in the 2015 biological community, within each measurement period, are best correlated to mean temperature ( $\tau = 0.14$ , Table C.5,  $p \leq 0.001$ , Fig. C.4). There was a noticeable drop-off in the correlation coefficients for the next best combinations that included maximum daily variance in acceleration, maximum salinity, and maximum pH (Table C.5).

**Table C.5: Best abiotic variable combinations at explaining the cumulative biological community data from 2015, with measurement period as a within-level factor in the BIOENV procedure.**

No. of abiotic metrics	Best metric combination	Kendall's tau
1	Mean temp	0.14
2	Mean temp + Max var accel	0.09
2	Mean temp + Max sal	0.08
2	Mean temp + Max pH	0.08
3	Mean temp + Max var accel + Max sal	0.07
4	Mean temp + Max var accel + Max sal + Max pH	0.07
1	Max sal	0.07
3	Mean temp + Max var accel + Max pH	0.07
3	Mean temp + Max sal + Max pH	0.06
2	Max var accel + Max sal	0.05



**Figure C.4:** The distribution of rho values from the 999 permutation tests carried out for the 2015 BIOENV procedure using the dataset. The rho value (dotted vertical line) of the best predictor variable (Mean temp, Table C.5) and its significance are displayed.

*C.3.2.2. DistLM 2015*

**Table C.6:** Marginal test results from the DistLM analysis in 2015.

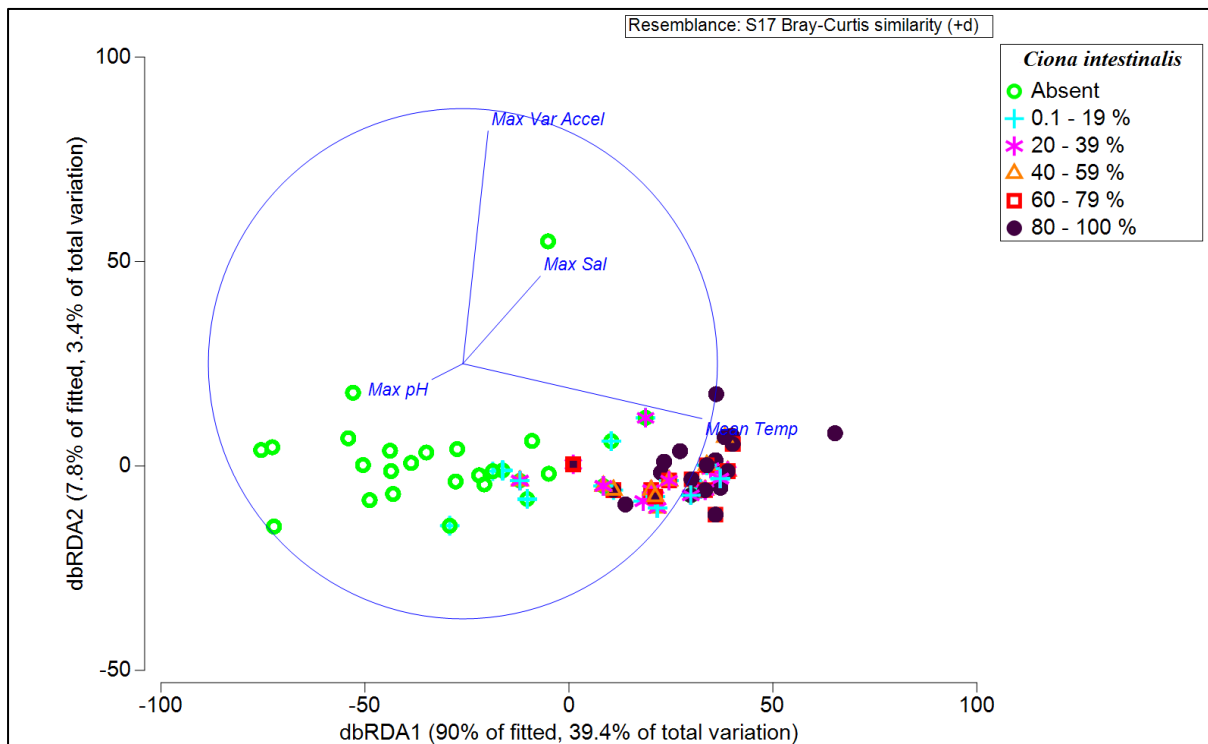
Variable	Sum of squares	Pseudo-F	<i>P</i> -value	Proportion of variation explained
Max Var Accel	6.32 E+04	20.505	0.001	0.026271
Max salinity	1.66 E+05	56.257	0.001	0.06892
Max pH	1.59 E+05	53.87	0.001	0.06619
Mean Temp	8.95 E+05	450.14	0.001	0.37197

The marginal test results from the DistLM analysis carried out on the 2015 data show that mean temperature explains the most variation in the biological community variation by far, at ~ 37 % (Table C.6). Maximum salinity and maximum pH explain ~ 7 % of the variation in the biological community variation, while maximum daily variance in acceleration explains ~ 3 %. All four metrics are significant ( $P \leq 0.001$ , Table C.6). The BEST method used to test which combination of abiotic metrics performed the best within the DistLM suggests that including all four of the abiotic metrics in the final model provides the best model fit (Table C.7). This top model explains 44 % of the variation in the biological community, while the ordination of the biological community can be seen in Fig. C.5. A series of models containing

either two or three of the abiotic metrics provide similar measures of explained variance, but they ranked much lower according to AICc comparison (Table C.7).

**Table C.7: Best models in 2015 from the DistLM procedure according to AICc ranking.**

AICc	R <sup>2</sup>	No. of abiotic metrics	Selections
5711.7	0.44	4	Mean temp + Max var accel + Max sal + Max pH
5729.3	0.42	3	Mean temp + Max var accel + Max sal
5744.4	0.41	3	Mean temp + Max var accel + Max pH
5757.6	0.4	3	Mean temp + Max sal + Max pH
5759.7	0.4	2	Mean temp + Max var accel
5774.3	0.39	2	Mean temp + Max sal
5774.6	0.39	2	Mean temp + Max pH
5789.3	0.37	1	Mean temp
6015.3	0.16	3	Max var accel + Max sal + Max pH
6049.7	0.12	2	Max var accel + Max sal



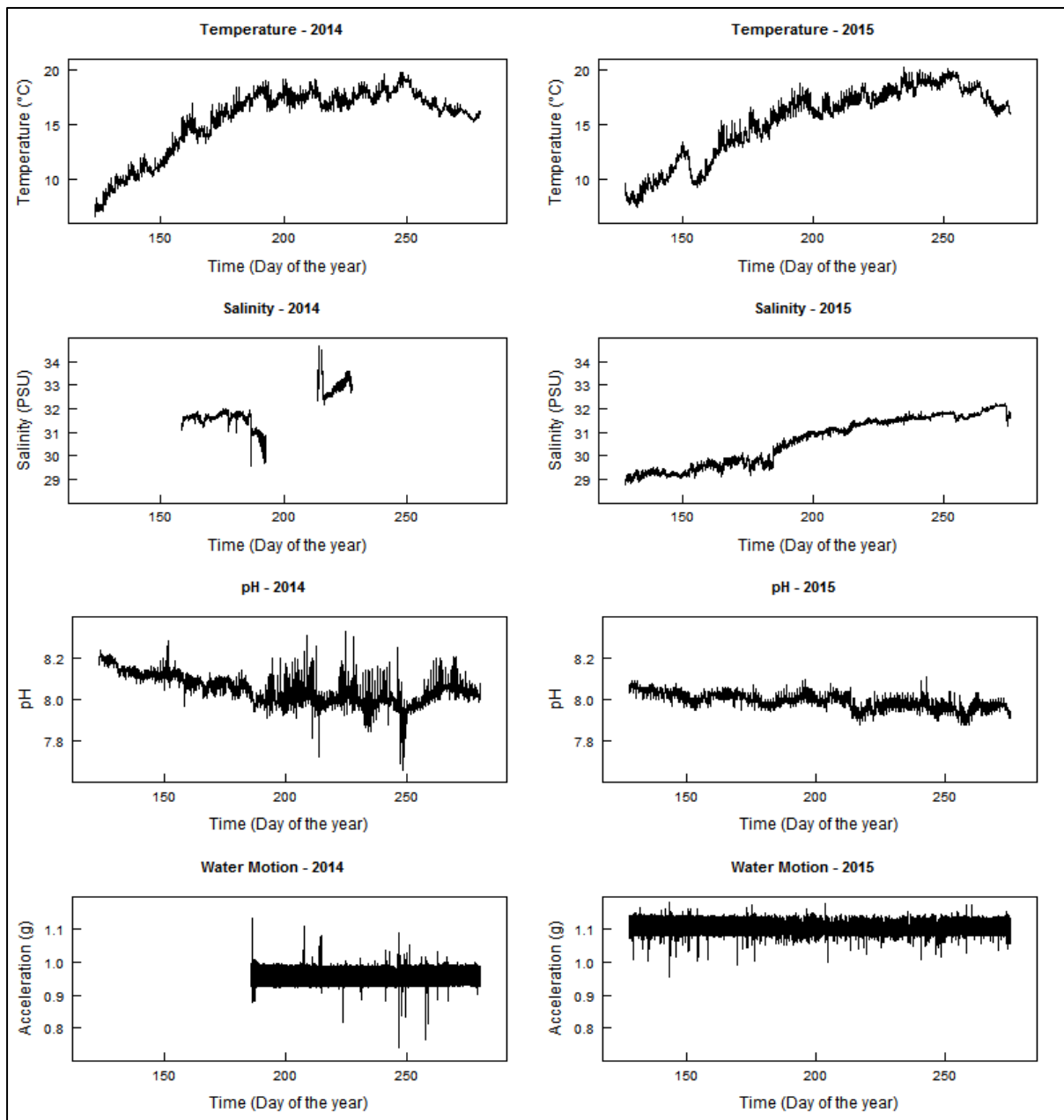
**Figure C.5: Ordination of the 2015 cumulative biological data according to distance-based redundancy analysis using the abiotic data as predictor variables. A vector overlay represents the distance and proportion of the abiotic predictor effects from the best model (Table C.5).**

#### *C.4. Discussion*

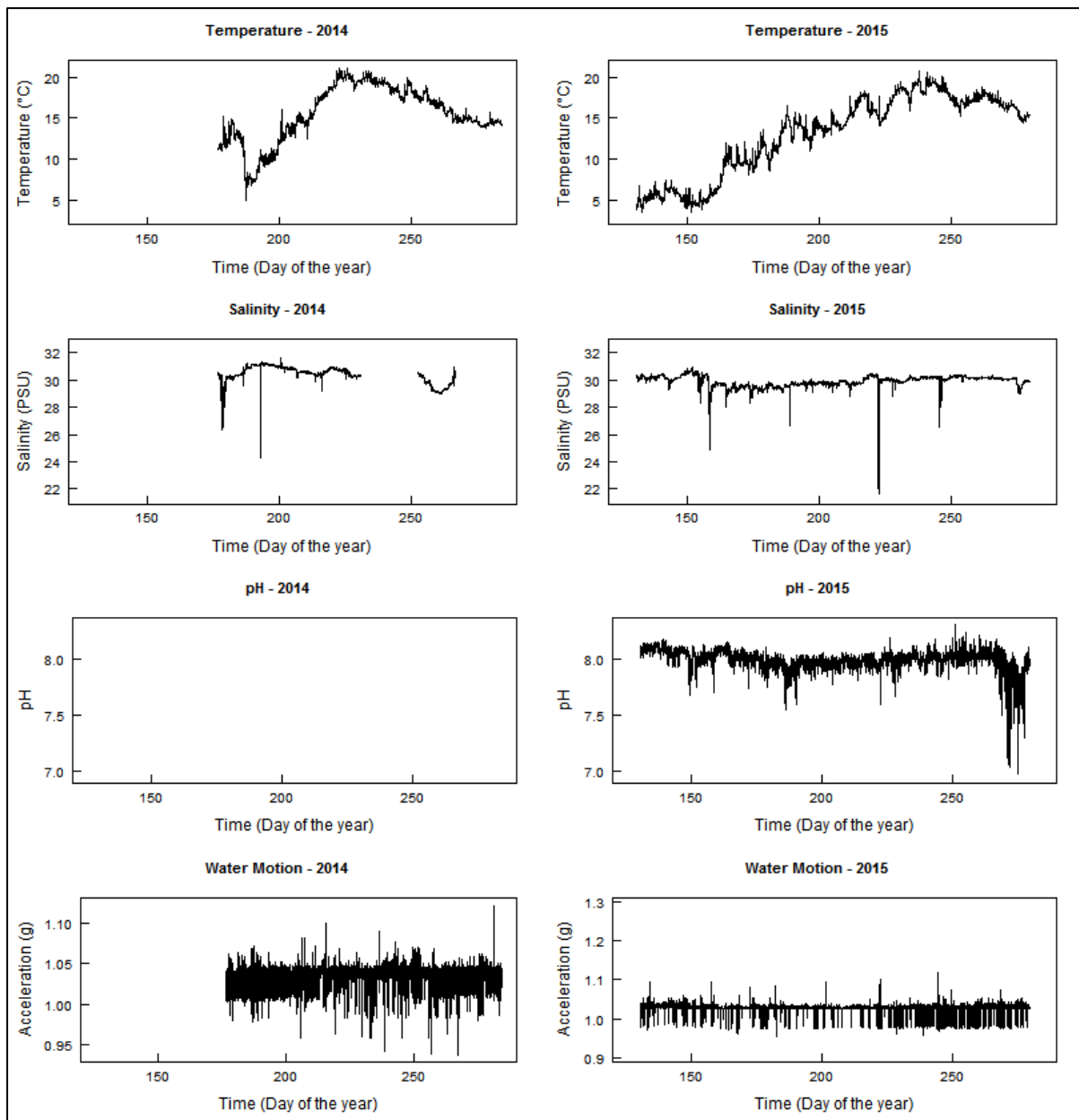
The results from the analyses conducted in PRIMER v7 and PERMANOVA+ describing the similarity in the patterns displayed in the biological community data and in the abiotic data, as well as which abiotic variables best explain these patterns, are very similar to the results obtained through the NLME modelling of the cumulative *C. intestinalis* abundance data. The single abiotic metric with the strongest relationship with the variation in the biological community in 2014 was mean salinity (Table C.1). In 2015, mean temperature had the strongest association with the variability in the biological community (Table C.5). In a similar fashion to the NLME modelling, the best predictors used to model the variation within the biological community data were composed of the metric from all four abiotic variables (Table C.7). In 2014, the best predictive model also included all metrics from the four abiotic metrics, but it should be noted that this was using an abiotic dataset that had missing data replaced.



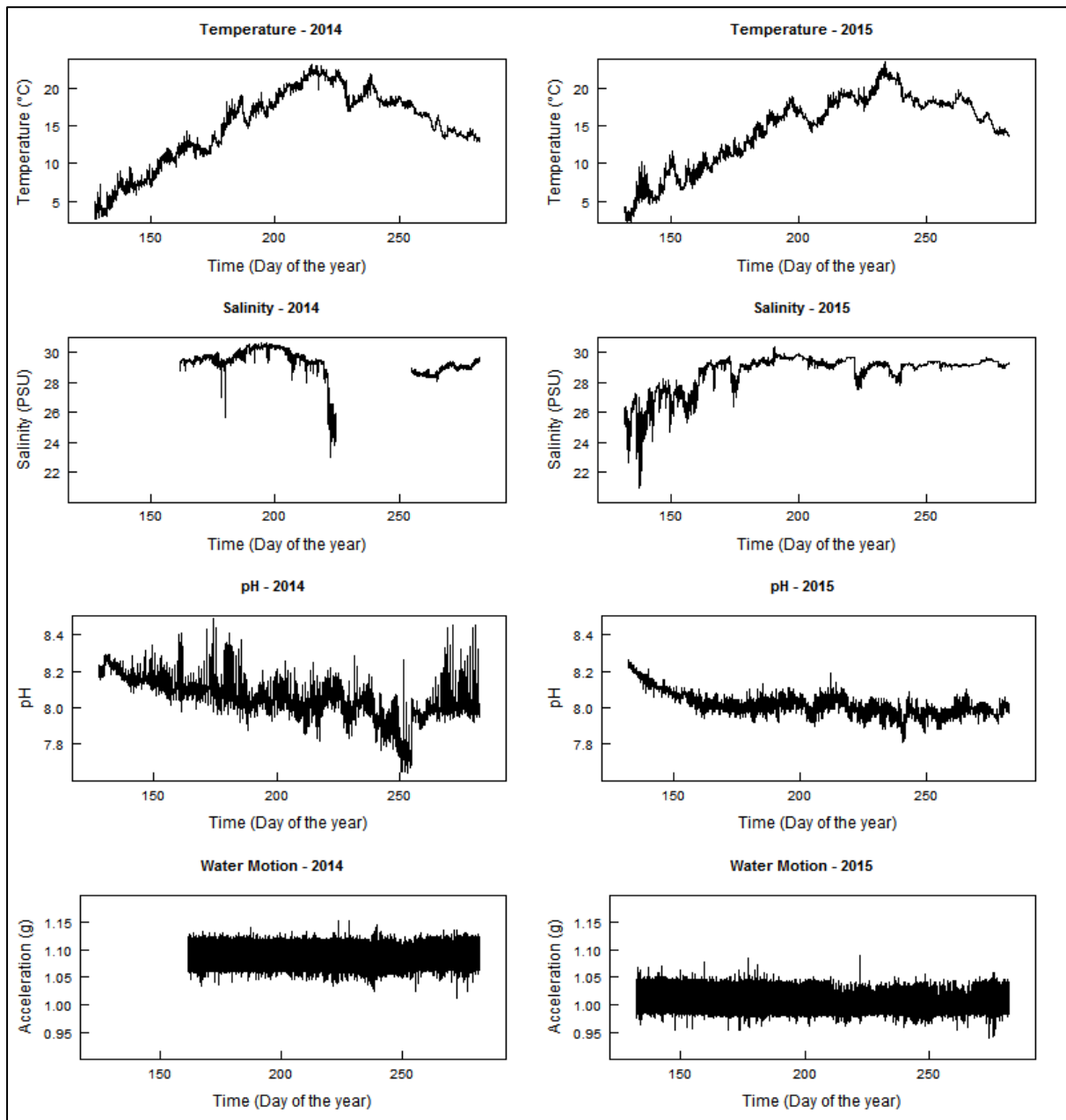
## Appendix D: Raw abiotic data plots



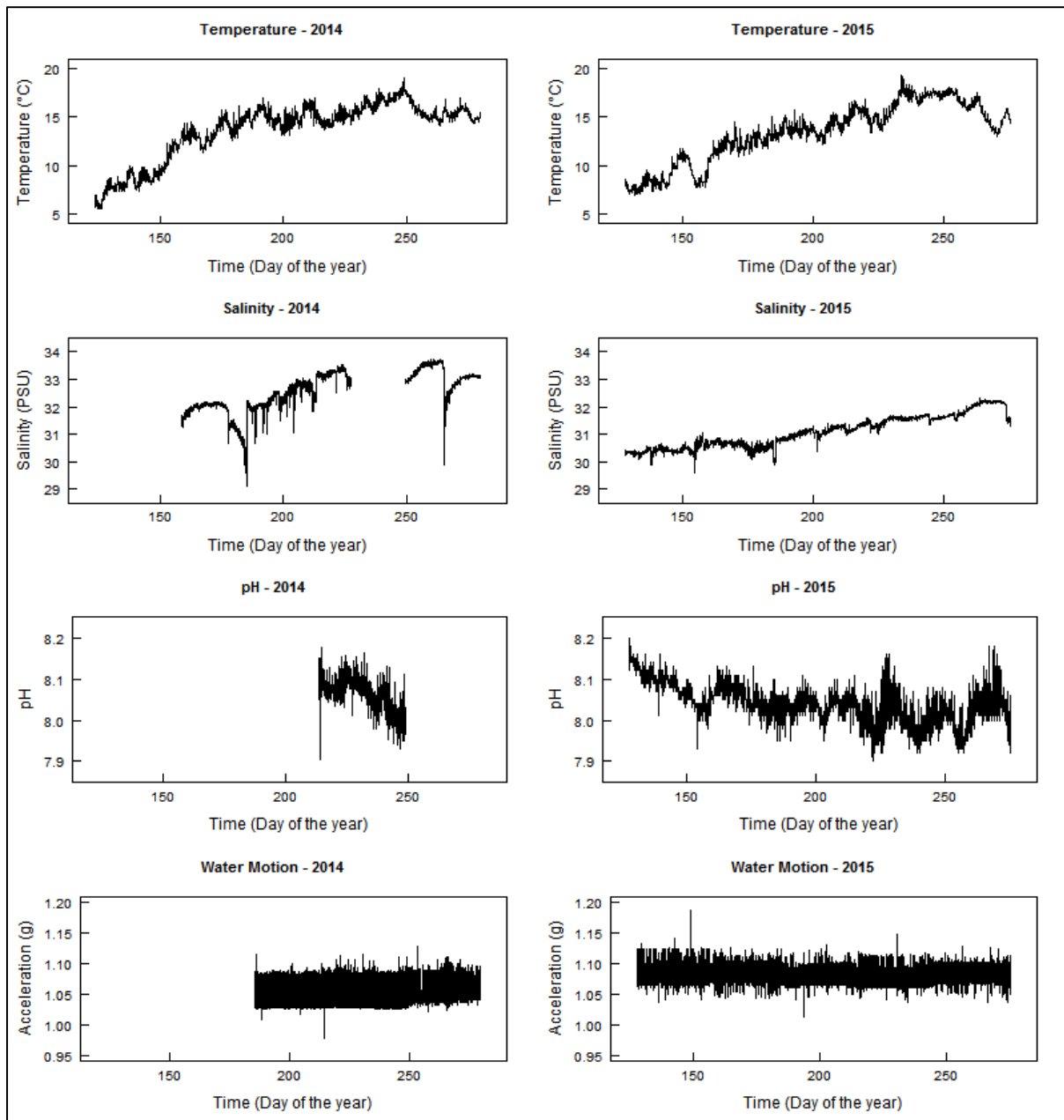
**Figure D.1: 5-minute interval data for each of the four abiotic variables measured in 2014 and in 2015 at Camp Cove (CC). The 2014 plots are in the first column, while 2015 plots are in the second. From the top row to the bottom, the abiotic variables are as follows: temperature; salinity; pH; and water motion.**



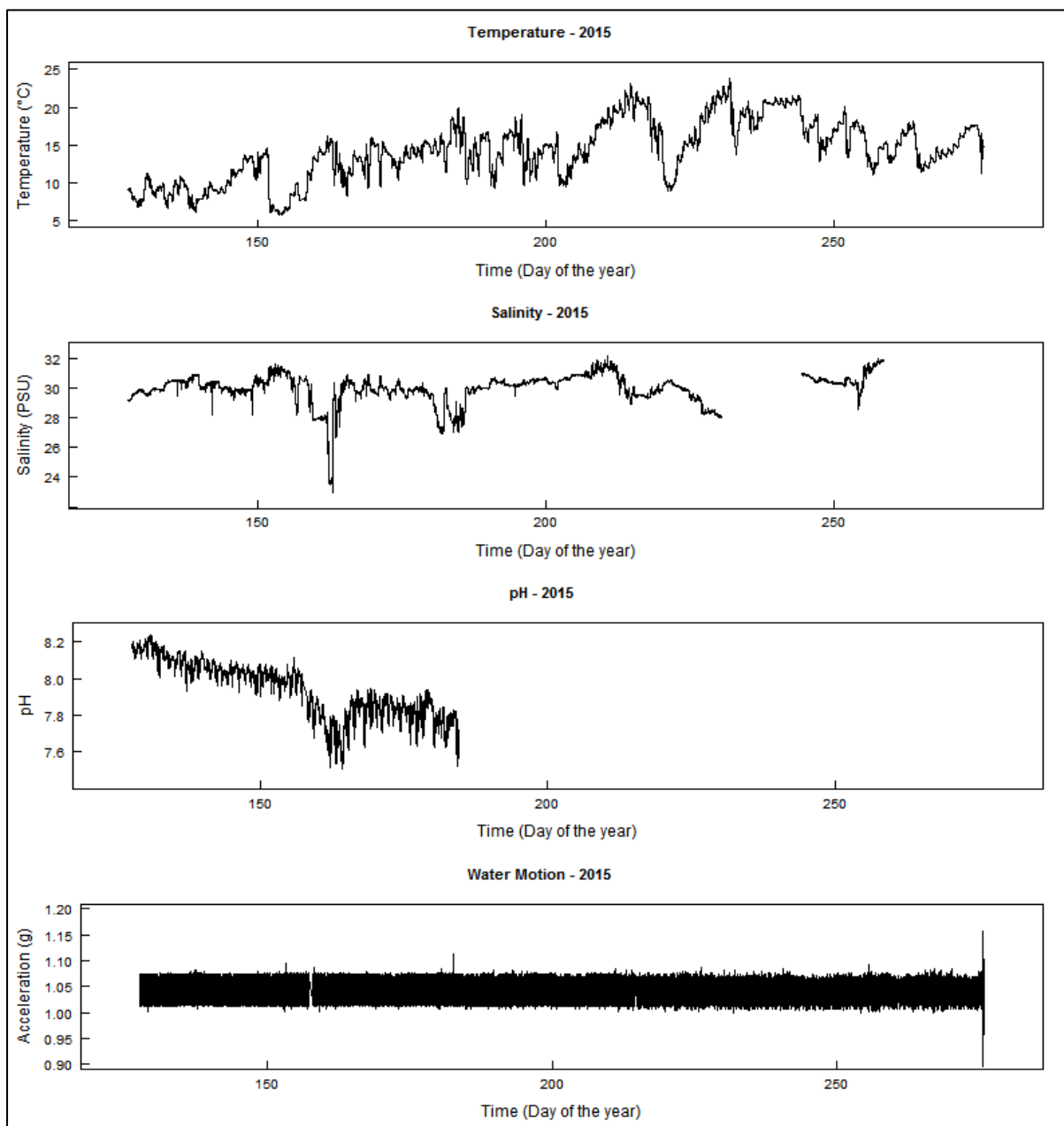
**Figure D.2: 5-minute interval data for each of the four abiotic variables measured in 2014 and in 2015 at Cape Canso (ZZ). The 2014 plots are in the first column, while 2015 plots are in the second. From the top row to the bottom, the abiotic variables are as follows: temperature; salinity; pH (2014 missing); and water motion.**



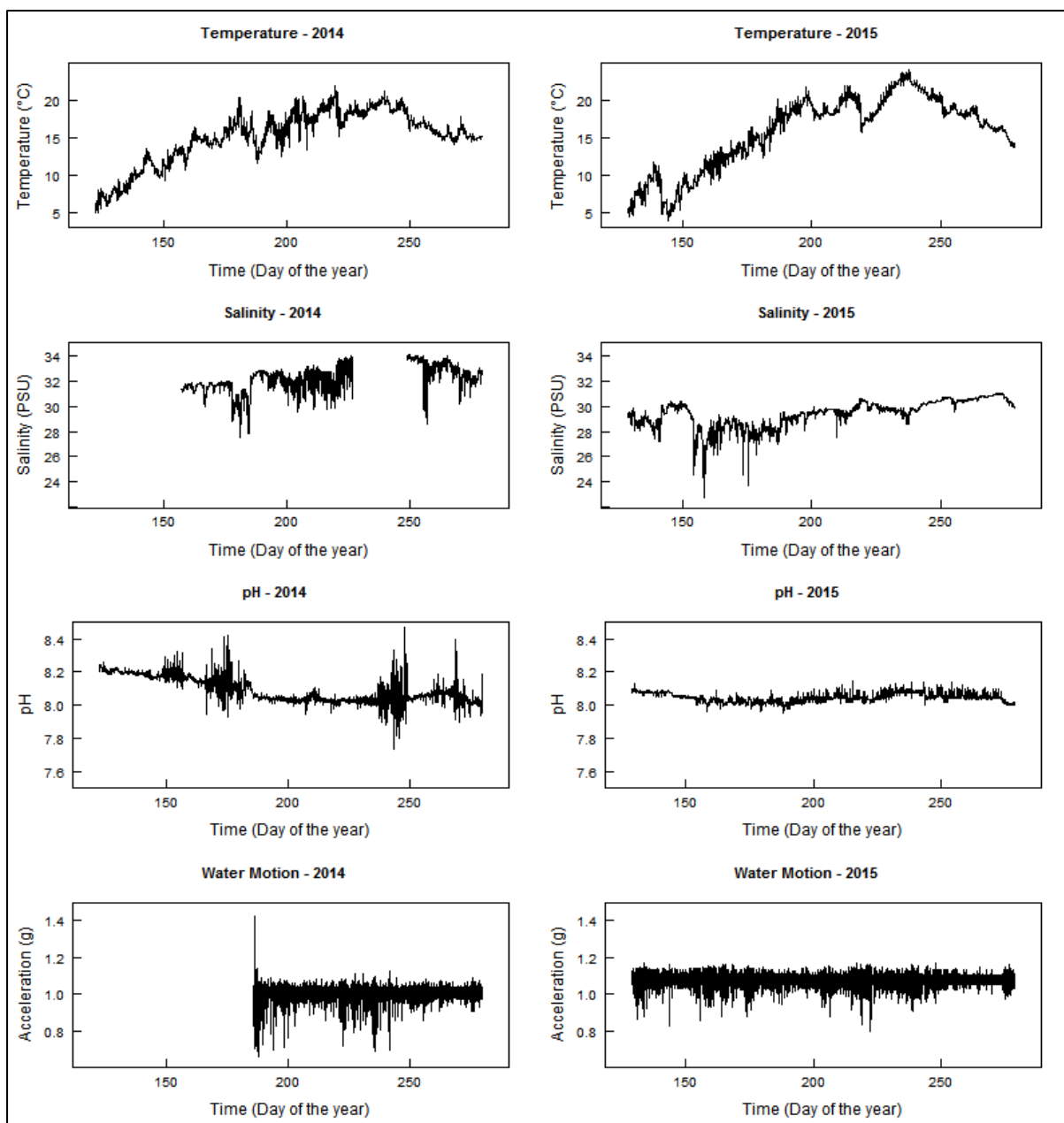
**Figure D.3: 5-minute interval data for each of the four abiotic variables measured in 2014 and in 2015 at Dingwall (DW). The 2014 plots are in the first column, while 2015 plots are in the second. From the top row to the bottom, the abiotic variables are as follows: temperature; salinity; pH; and water motion.**



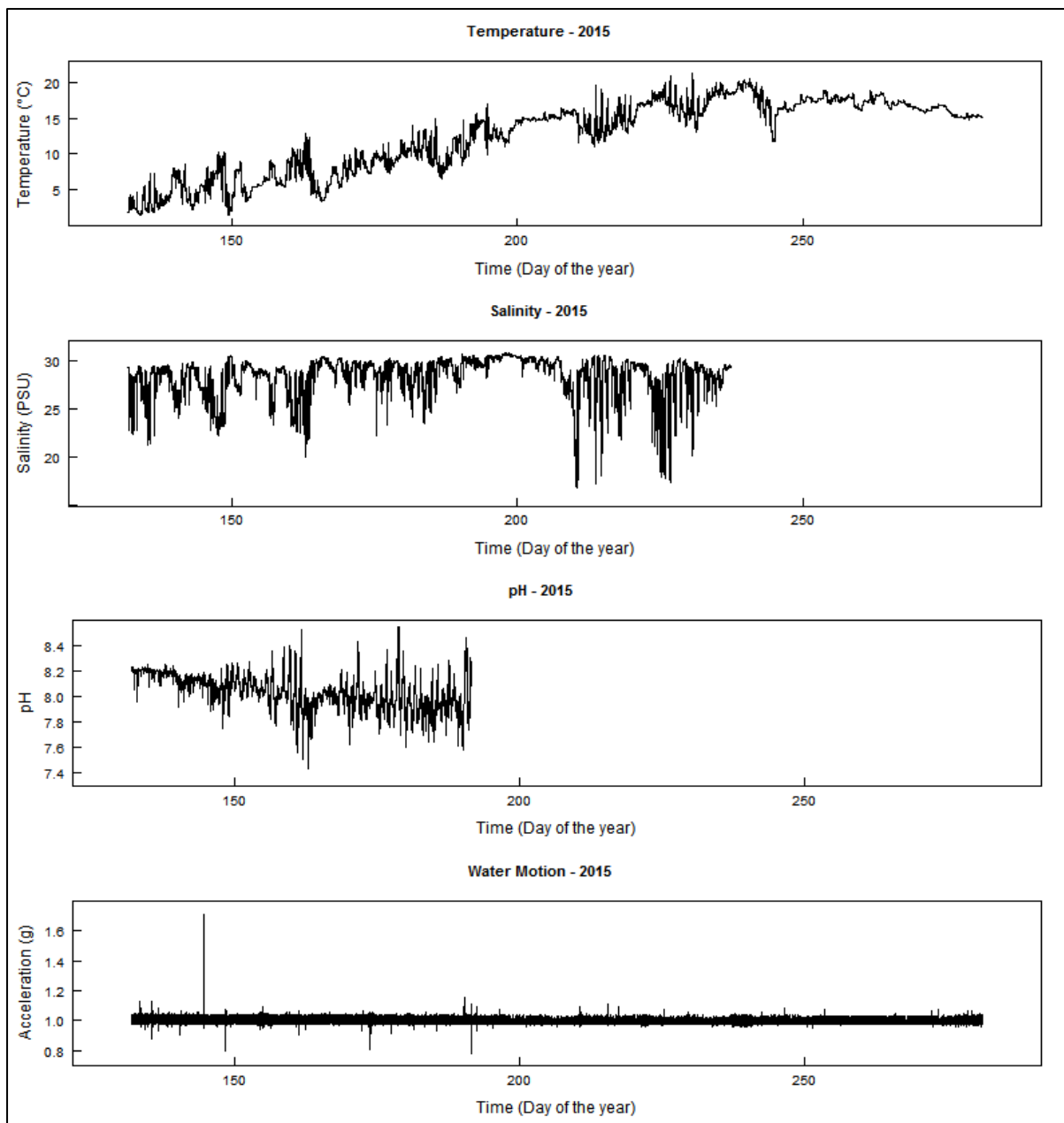
**Figure D.4: 5-minute interval data for each of the four abiotic variables measured in 2014 and in 2015 at Fall's Point (FP). The 2014 plots are in the first column, while 2015 plots are in the second. From the top row to the bottom, the abiotic variables are as follows: temperature; salinity; pH; and water motion.**



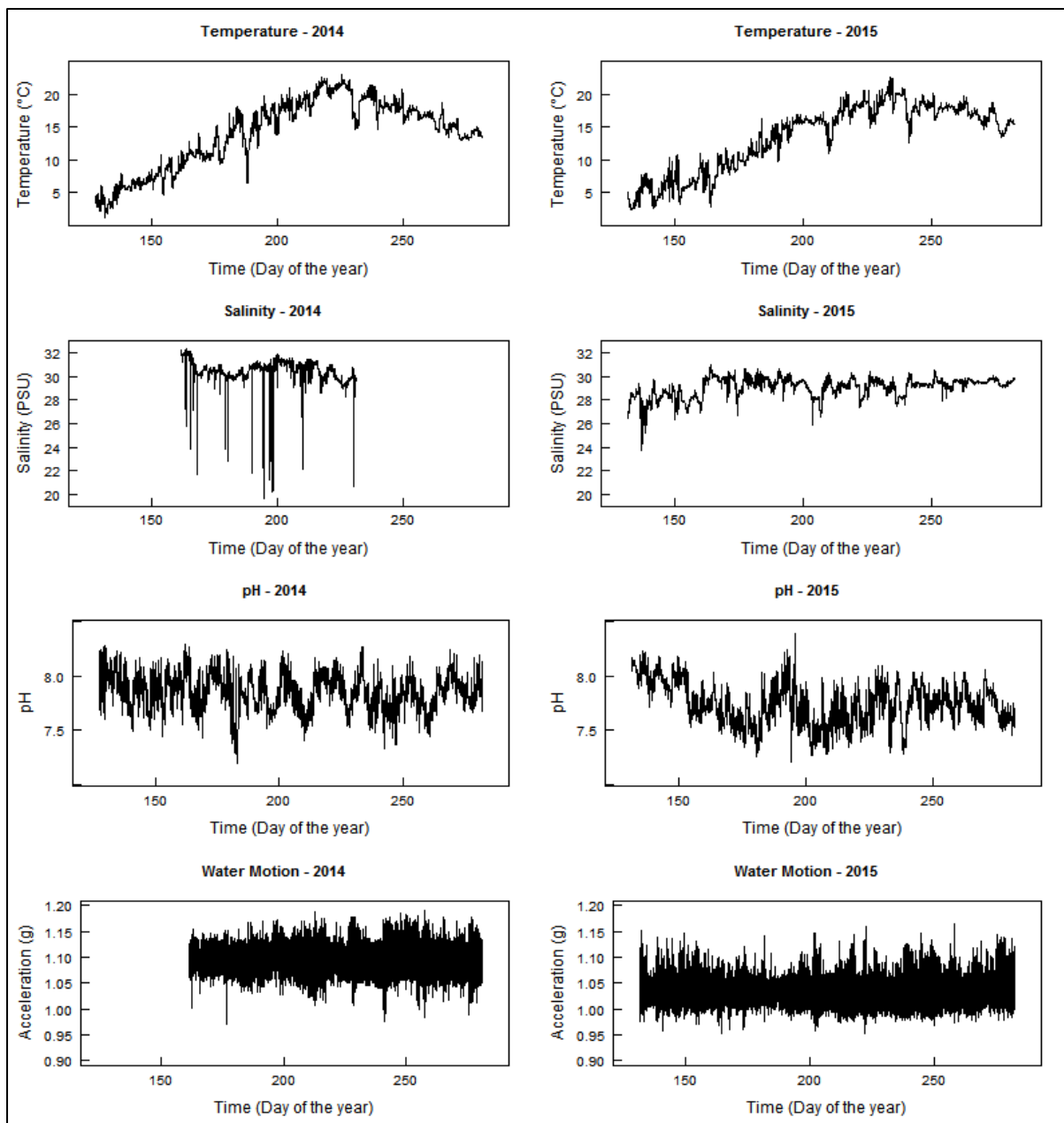
**Figure D.5: 5-minute interval data for each of the four abiotic variables measured in 2015 at Ingomar (IN). From the top row to the bottom, the abiotic variables are as follows: temperature; salinity; pH; and water motion.**



**Figure D.6: 5-minute interval data for each of the four abiotic variables measured in 2014 and in 2015 at Indian Point (IP). The 2014 plots are in the first column, while 2015 plots are in the second. From the top row to the bottom, the abiotic variables are as follows: temperature; salinity; pH; and water motion.**

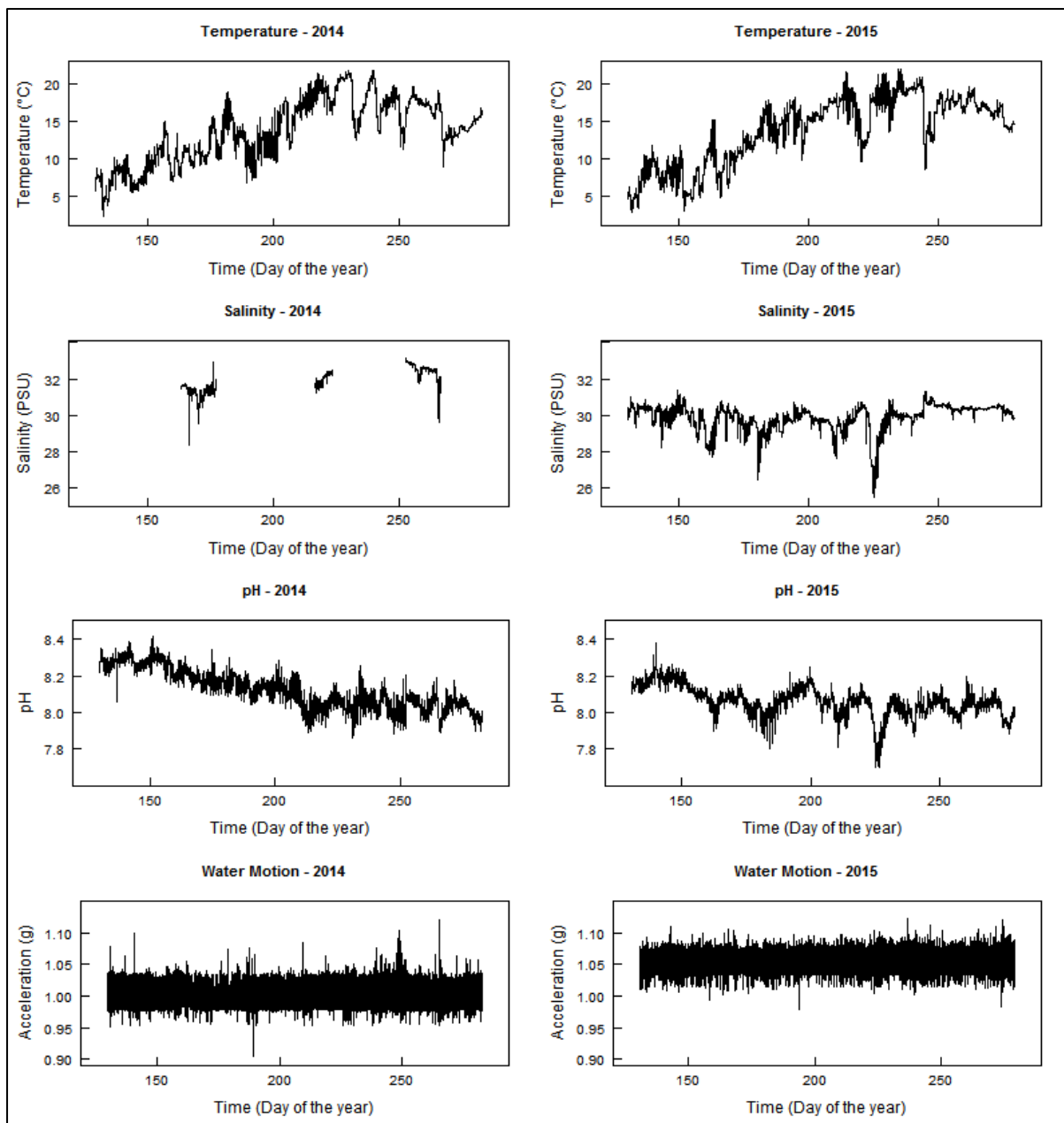


**Figure D.7: 5-minute interval data for each of the four abiotic variables measured in 2015 at Louisbourg (LB). From the top row to the bottom, the abiotic variables are as follows: temperature; salinity; pH; and water motion.**

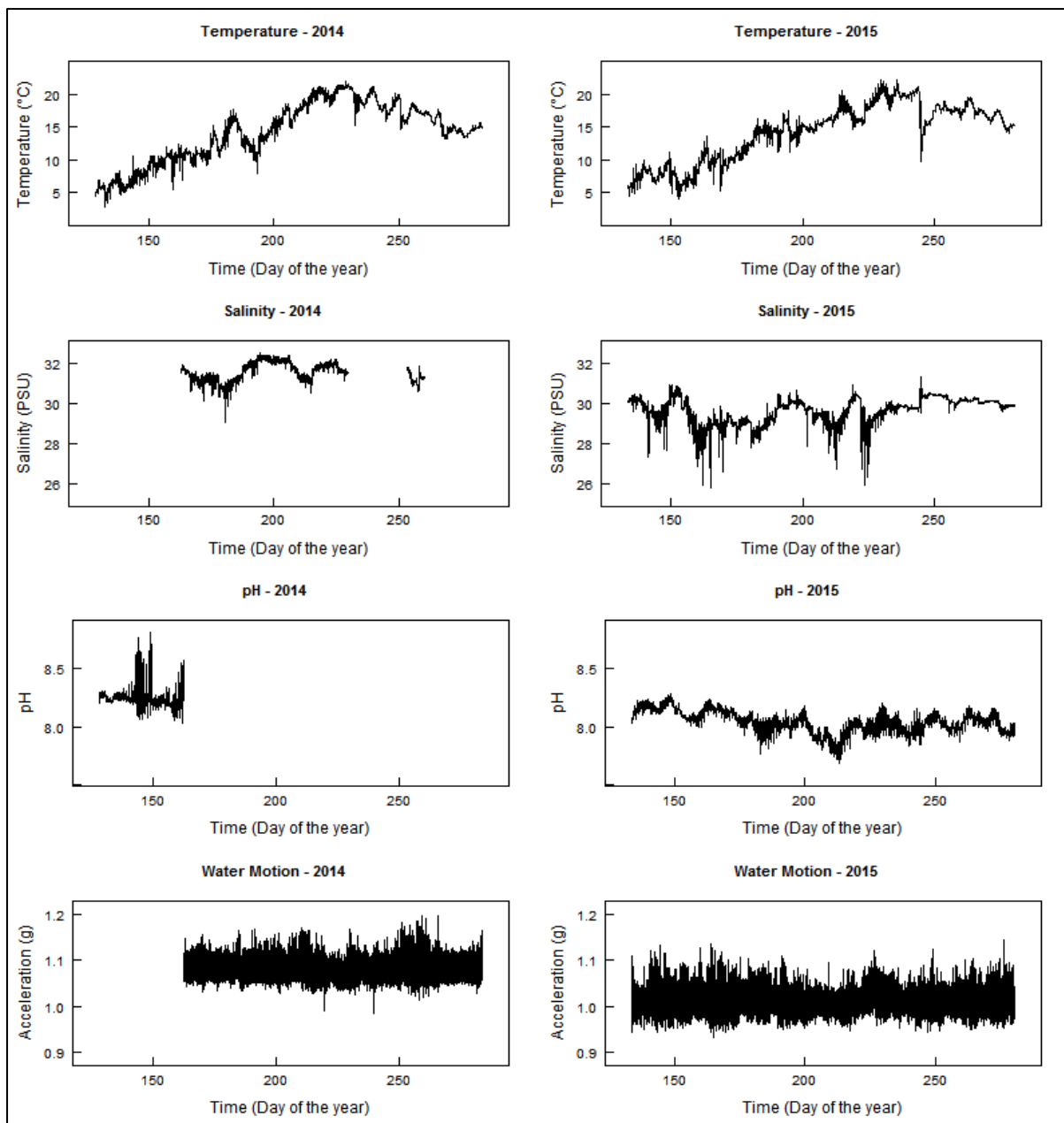


**Figure D.8: 5-minute interval data for each of the four abiotic variables measured in 2014 and in 2015 at Little River (LR). The 2014 plots are in the first column, while 2015 plots are in the second. From the top row to the bottom, the abiotic variables are as follows: temperature; salinity; pH; and water motion.**

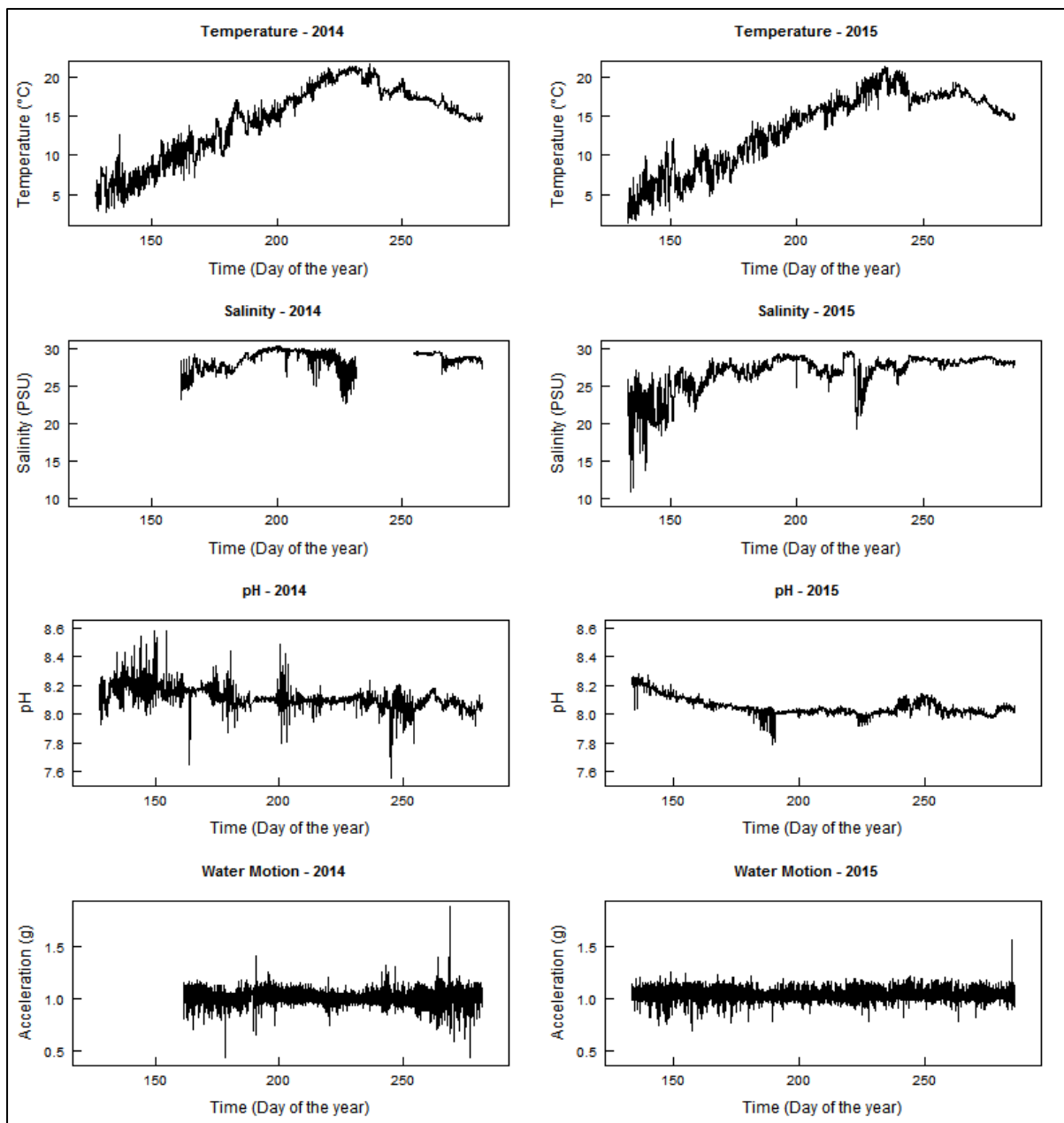




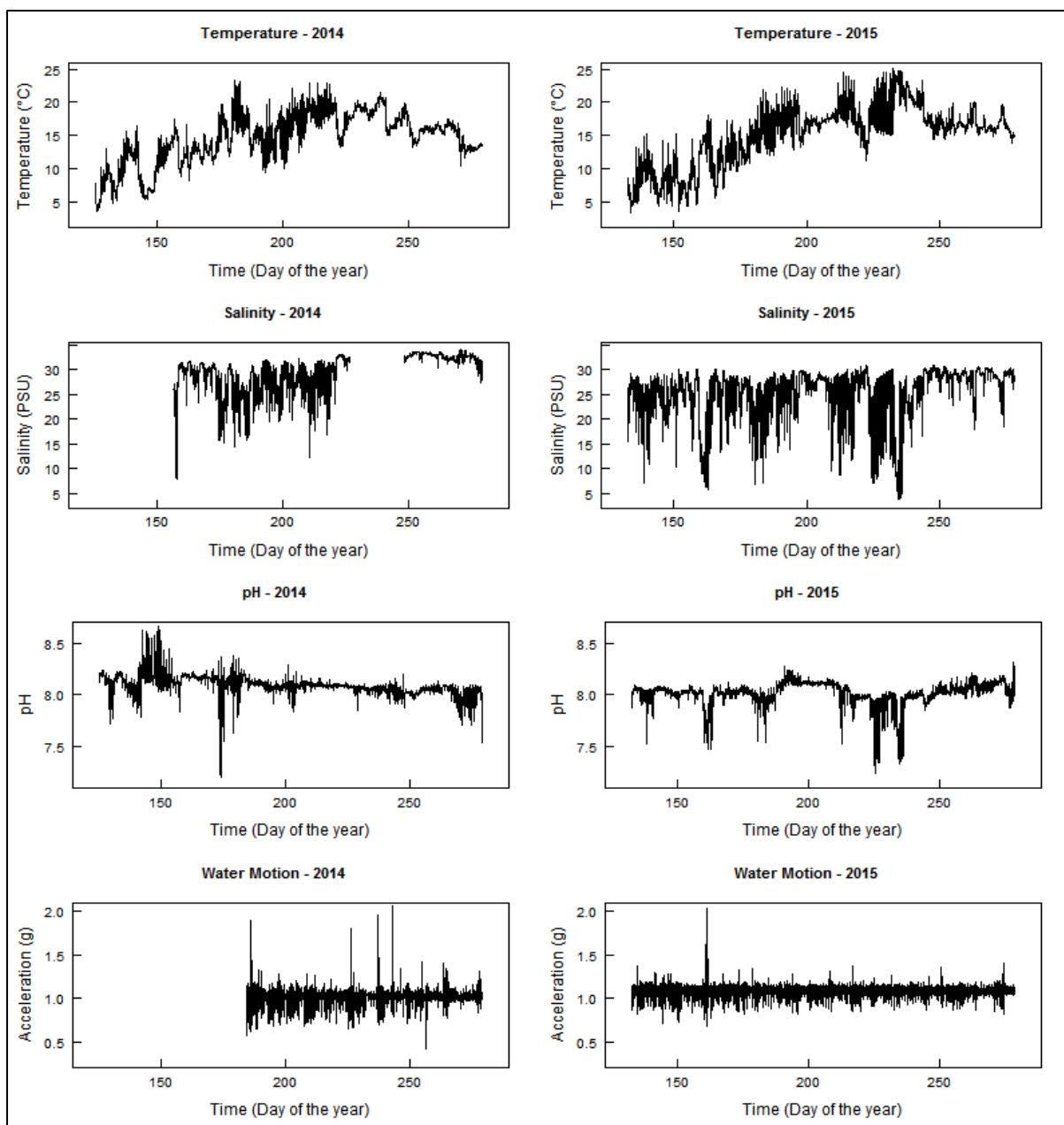
**Figure D.9: 5-minute interval data for each of the four abiotic variables measured in 2014 and in 2015 at Port Bickerton (PB). The 2014 plots are in the first column, while 2015 plots are in the second. From the top row to the bottom, the abiotic variables are as follows: temperature; salinity; pH; and water motion.**



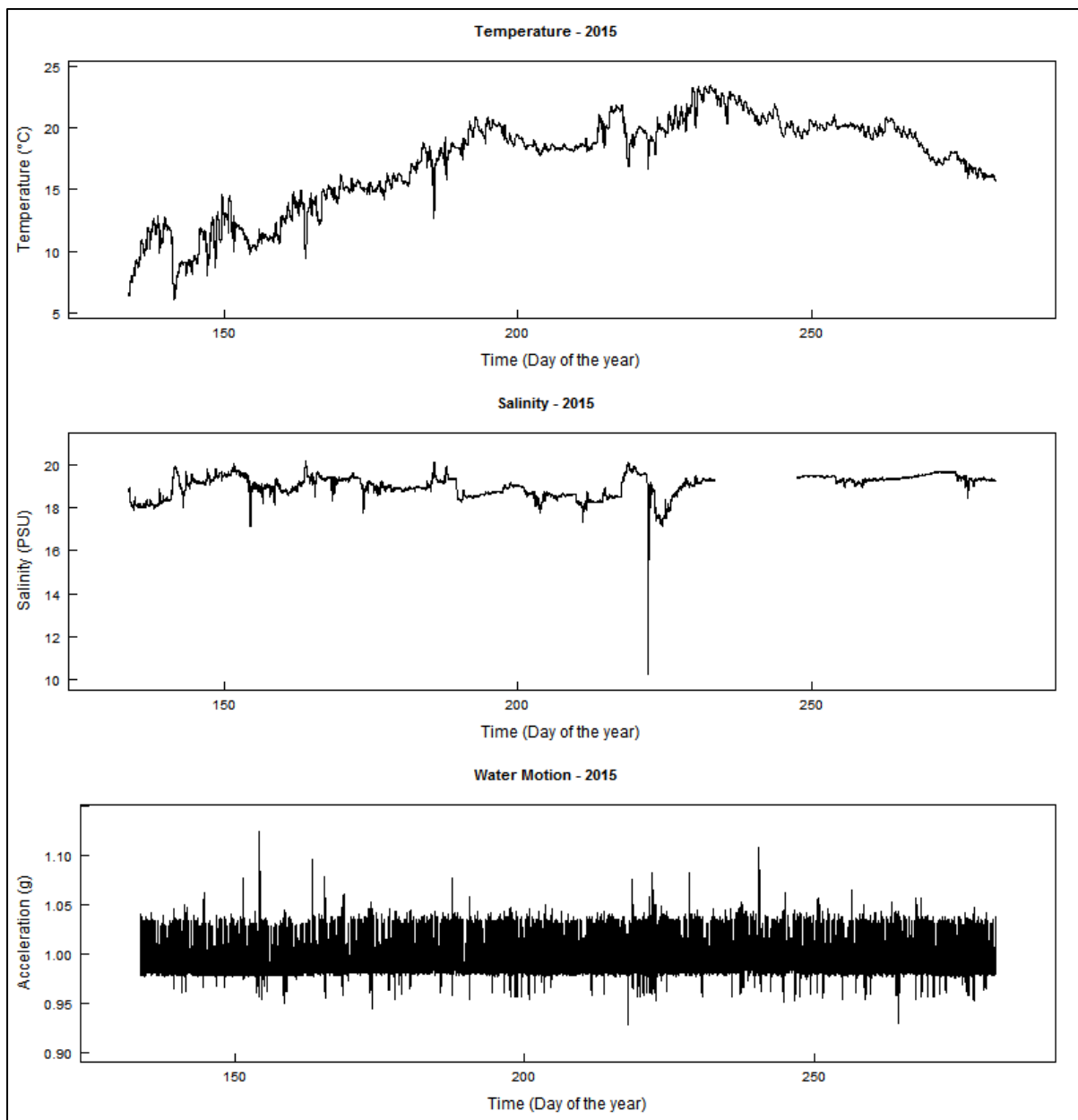
**Figure D.10: 5-minute interval data for each of the four abiotic variables measured in 2014 and in 2015 at Petit-de-Grat (PG). The 2014 plots are in the first column, while 2015 plots are in the second. From the top row to the bottom, the abiotic variables are as follows: temperature; salinity; pH; and water motion.**



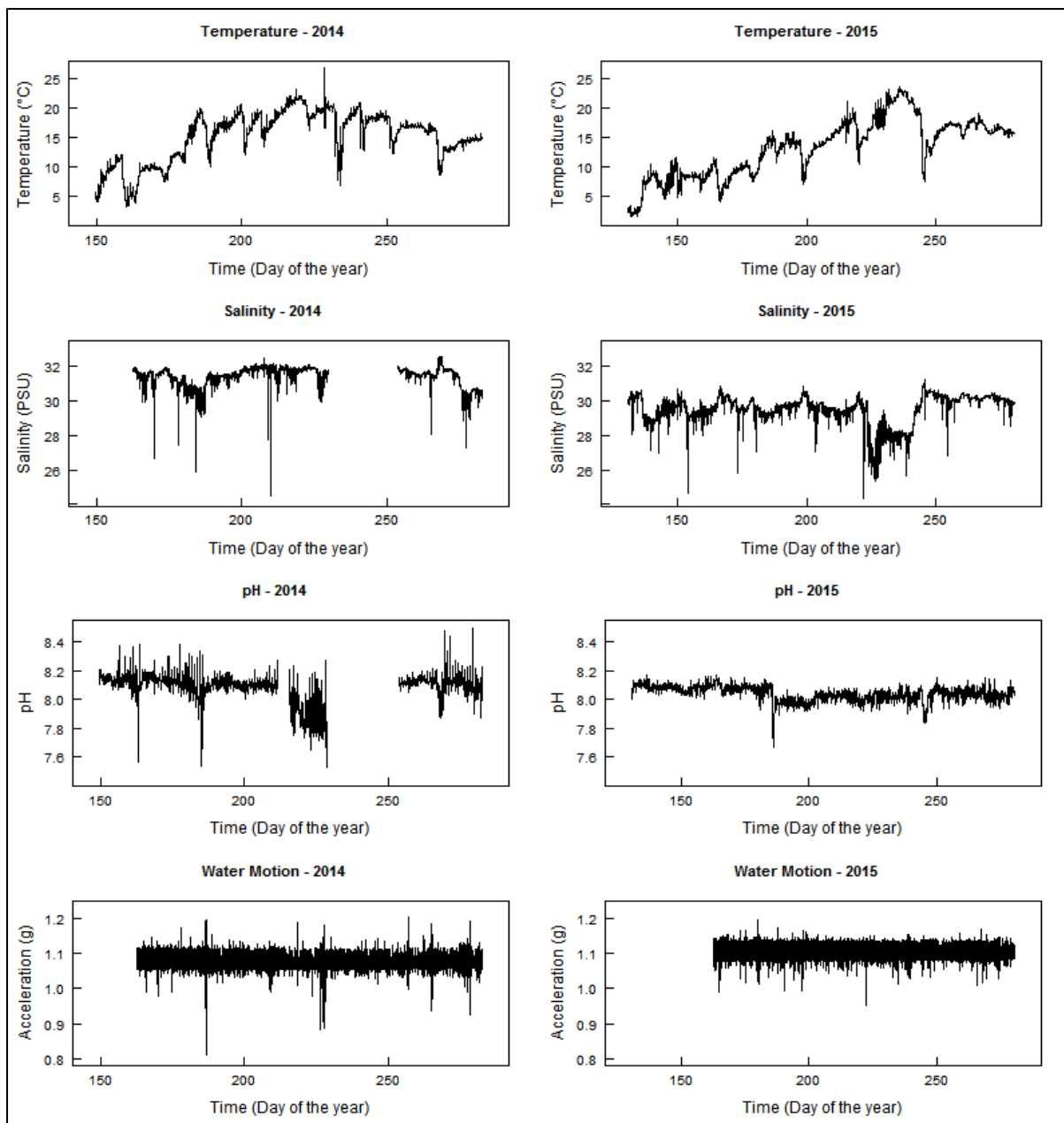
**Figure D.11: 5-minute interval data for each of the four abiotic variables measured in 2014 and in 2015 at St. Ann's Bay (SA). The 2014 plots are in the first column, while 2015 plots are in the second. From the top row to the bottom, the abiotic variables are as follows: temperature; salinity; pH; and water motion.**



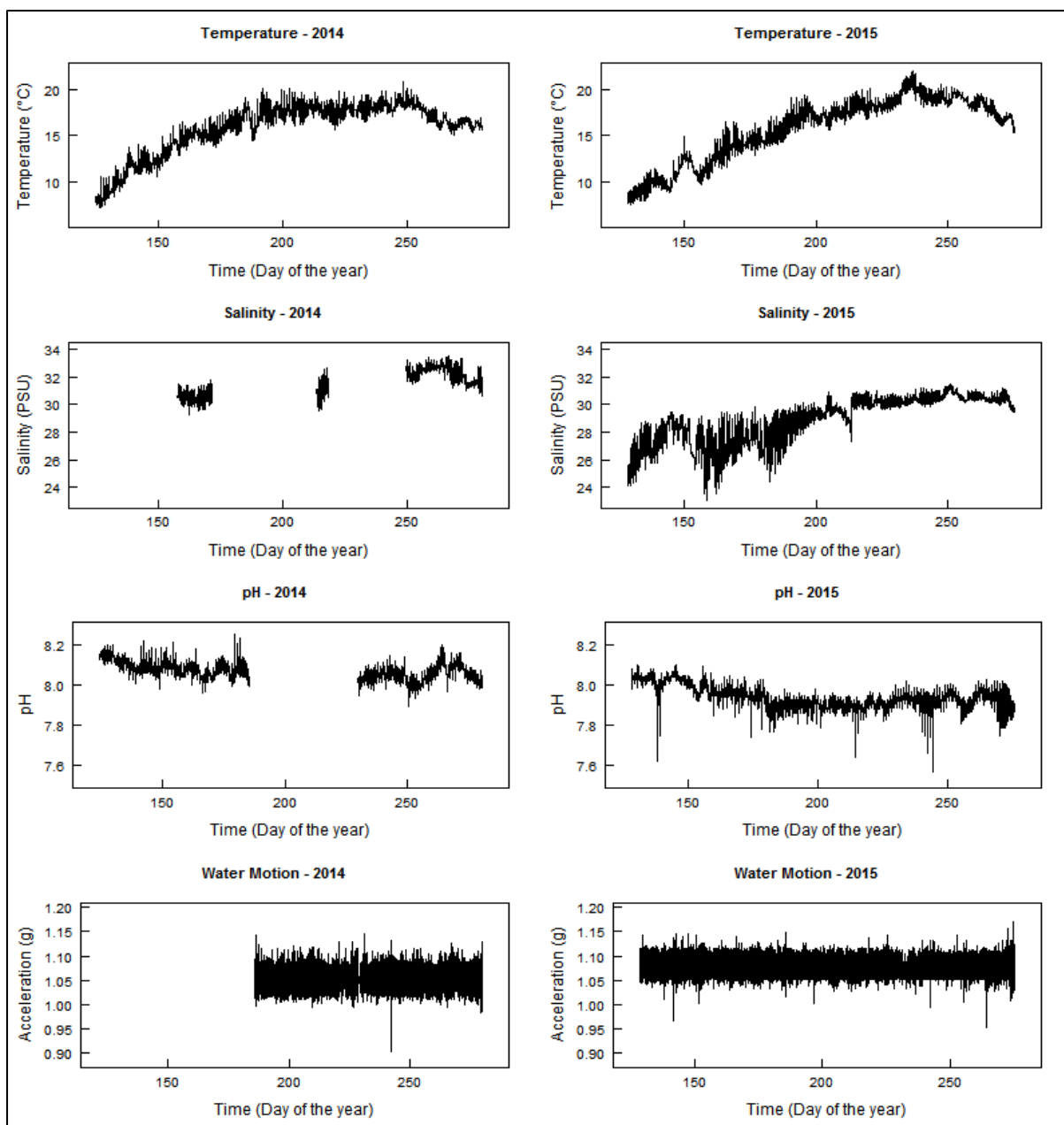
**Figure D.12: 5-minute interval data for each of the four abiotic variables measured in 2014 and in 2015 at Ship Harbour (SH). The 2014 plots are in the first column, while 2015 plots are in the second. From the top row to the bottom, the abiotic variables are as follows: temperature; salinity; pH; and water motion.**



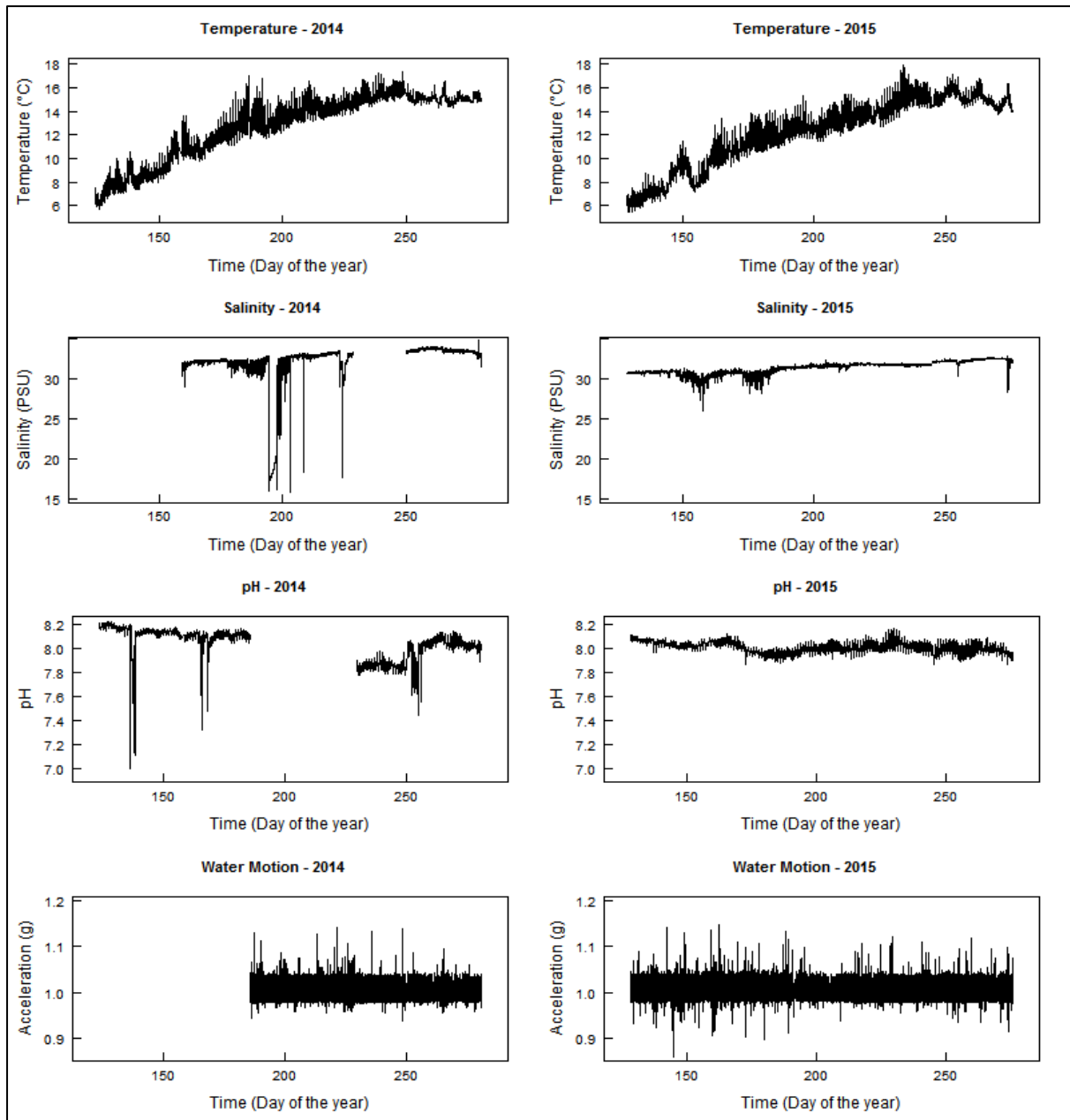
**Figure D.13: 5-minute interval data for each of the three abiotic variables measured in 2015 at St Peters (SP). From the top row to the bottom, the abiotic variables are as follows: temperature; salinity; and water motion.**



**Figure D.14: 5-minute interval data for each of the four abiotic variables measured in 2014 and in 2015 at Venus Cove (VC). The 2014 plots are in the first column, while 2015 plots are in the second. From the top row to the bottom, the abiotic variables are as follows: temperature; salinity; pH; and water motion.**



**Figure D.15: 5-minute interval data for each of the four abiotic variables measured in 2014 and in 2015 at Wedgeport (WP). The 2014 plots are in the first column, while 2015 plots are in the second. From the top row to the bottom, the abiotic variables are as follows: temperature; salinity; pH; and water motion.**



**Figure D.16: 5-minute interval data for each of the four abiotic variables measured in 2014 and in 2015 at Yarmouth Bar (YB). The 2014 plots are in the first column, while 2015 plots are in the second. From the top row to the bottom, the abiotic variables are as follows: temperature; salinity; pH; and water motion.**

## Appendix E: 2014 and 2015 combined NLME model

### *E.1. Preselection of abiotic variables*

The four metrics remaining after the stepwise removal process was carried out on the pairwise-complete observations of the 2014 and the 2015 combined abiotic data are mean temperature, minimum salinity, maximum pH, and minimum variance in acceleration (Table



E.1). There were a number of very highly correlated ( $r > 0.9$ ) metrics, seen in Table E.1. All of the very highly correlated metrics were found between correlations within the same abiotic variable, e.g. mean and maximum temperature ( $r = 0.94$ ).

**Table E.1: Stepwise removal of a single abiotic variable metric from the most highly correlated pair using the combined 2014 and 2015 abiotic data. The pair of metrics most correlated at each step is highlighted. The light grey shaded metric was kept (Keep) while the metric shaded with dark grey was removed (Rem.) from the process. There was a selection rule to always have one metric from each variable remaining. Only pairwise complete observations were used in the abiotic metric correlations.**

Variable	Metric	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7	Step 8*	Step 9
Temp.	Mean	Keep			Keep					
	Minimum				Rem.					
	Maximum	Rem.								
Salinity	Mean		Rem.							
	Minimum		Keep				Keep	Keep	Keep	Keep
	Maximum						Rem.			
pH	Mean					Rem.				
	Minimum									Rem.
	Maximum					Keep				
Water Motion Accel.	Variance			Rem.						
	Minimum							Rem.		
	Maximum			Keep					Rem.	
Pearson's r		0.94	0.91	0.91	0.9	0.88	0.58	-0.55	-0.53	0.47

As the highly collinear metrics were reduced, some moderate correlations were found between metrics from two different abiotic variables, e.g. Step 7 (Table E.1): minimum salinity and minimum variance in acceleration ( $r = -0.55$ ). If the selection rule requiring one metric from each abiotic variable be present in the final four was not in place, all correlations of  $r > 0.5$  would have been removed. In one case, in the last removal step neither of the most correlated pair could be removed, as they were the last representative for their respective abiotic variables. The second most correlated pair was used instead (Step 9, Table E.1). I was unable to achieve a full top-down model that converged with the four final metrics from Table E.1, so instead of mean temperature, I used minimum temperature, as this resulted in

model convergence. I used both mean and minimum temperature in the bottom-up model fitting (Table E.2).

### *E.2. Bottom-up and top-down candidate sets*

The ‘best’ bottom-up model (BBU) contains minimum temperature and maximum daily variance in acceleration in the Days50 and Lag parameters, while it has mean temperature, minimum salinity and maximum variance in acceleration in the MaxPop parameter (Table E.2). Its Akaike weight ( $w_i$ ) is 0.54 out of the total ( $N = 32$ ) set of compared models (candidate set) for the BU approach (Table E.2). Only seven models in the BU candidate set had a  $\Delta_i$  of  $< 10$ . The second best model had a  $\Delta_i$  of  $\sim 2$  compared to the BBU (Table E.2). The evidence ratio between these top two models is 2.8, meaning that the top model is almost three times as likely to be the best model of the candidate set, based on these data.

The TD candidate set appears to have provided a much better model than those fit by the BU approach (Table E.2 & E.3), with a large  $\Delta_i \sim 20$  between the BBU and ‘best’ top-down model (BTD). While not in the region of  $> 0.9$  suggested by (Burnham and Anderson, 2003) as describing a clearly strongest model, the  $w_i$  of 0.59 is the largest of any ‘best’ model from any candidate set of any year. There is only one other model that is within the  $\Delta_i$  of 2 (Table E.3). The abiotic metrics contained within the model parameters are like the global model (GM), except for the exclusion of minimum salinity from MaxPop and Lag. The second BTD model, with a  $w_i$  of 0.36 and an evidence ratio compared to the BTD of 1.6, is almost the GM, only excluding maximum pH from Lag.

There is less uncertainty in the best model selection in both BU and TD candidate sets for the combined yearly data when compared to either 2014 or 2015 in isolation. This is exemplified by the top two models in the TD combined candidate set having a cumulative  $w_i$  of 0.95, while the top three models ( $\Delta_i$  of  $< 4$ ) in the BU candidate set have a cumulative  $w_i$  of 0.85. The models do have a relatively high number of parameters, represented by all of the abiotic variables. This suggests that each abiotic variable contributes valuable information to the likelihood of the best models, without over parametrizing.

**Table E.2: The seven best models (total no. models = 32) from the bottom-up model approach (combined), ranked by lowest AICc value. Each model parameter (MaxPop, D50, and Lag) could have all four of the abiotic metrics in it in any given model. The abiotic metrics in the table have been assigned shorthand names as follows: Mean temp. = mean temperature; Temp = minimum temperature; Sal = minimum salinity; pH = maximum pH; and Accel = maximum variance in acceleration. K = no. model parameters.**

Model rank	K	AICc	$\Delta_i$	Model Likelihood	Model $w_i$	Cumulative $w_i$	Maximum population (MaxPop)				Days50 (D50)				Lag phase (Lag)			
1	13	7764.23	0	1	0.54	0.54	Mean Temp.	Sal	Accel		Temp		Accel		Temp		Accel	
2	14	7766.29	2.06	0.36	0.19	0.73	Mean Temp.	Sal	Accel		Temp		Accel	pH	Temp		Accel	
3	15	7767.26	3.04	0.22	0.12	0.85	Mean Temp.	Sal	Accel		Temp		Accel	pH				
4	17	7768.5	4.28	0.12	0.06	0.91	Mean Temp.	Sal	Accel	pH	Temp		Accel	pH	Temp	Sal	Accel	pH
5	16	7769.2	4.97	0.08	0.04	0.96	Mean Temp.	Sal	Accel	pH	Temp		Accel	pH	Temp	Sal	Accel	
6	12	7770.36	6.13	0.05	0.03	0.98	Mean Temp.	Sal	Accel				Accel		Temp		Accel	
7	14	7773.23	9	0.01	0.01	0.99		Sal	Accel	pH	Temp		Accel	pH	Temp		Accel	

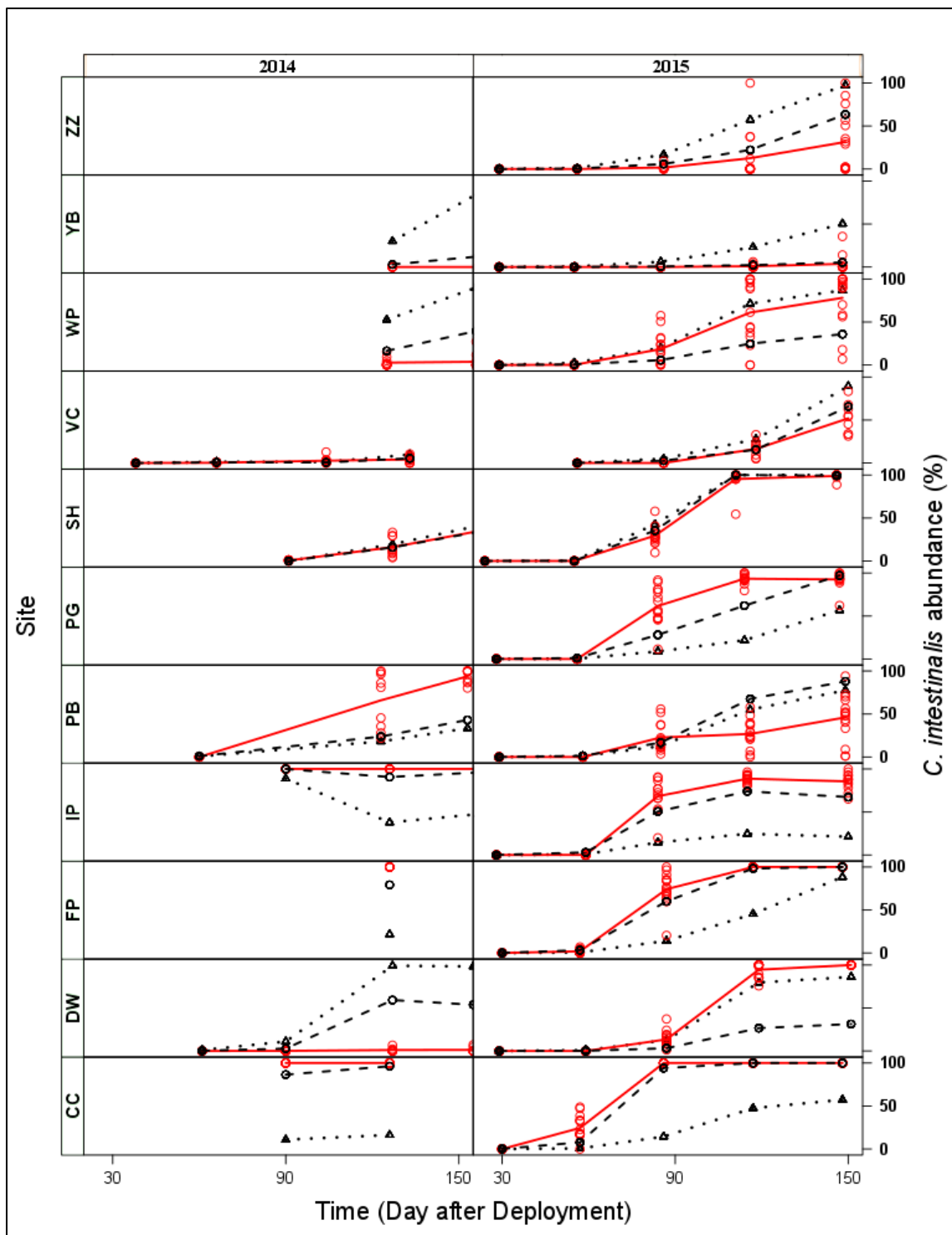
**Table E.3: The six best models (total no. models = 29) from the top-down model approach (combined), ranked by lowest AICc value. Each model parameter (MaxPop, D50, and Lag) could have all four of the abiotic metrics in it in any given model. The abiotic metrics in the table have been assigned shorthand names as follows: Temp = minimum temperature; Sal = minimum salinity; pH = maximum pH; and Accel = maximum variance in acceleration. K = no. model parameters.**

Model rank	K	AICc	$\Delta_i$	Model Likelihood	Model $w_i$	Cumulative $w_i$	Maximum population (MaxPop)				Days 50 (D50)				Lag phase (Lag)			
							Temp	Sal	Accel	pH	Temp	Sal	Accel	pH	Temp	Sal	Accel	pH
1	16	7745.00	0.00	1.00	0.59	0.59	Temp		Accel	pH	Temp	Sal	Accel	pH	Temp		Accel	pH
2	17	7746.01	1.01	0.60	0.36	0.95	Temp	Sal	Accel	pH	Temp	Sal	Accel	pH	Temp	Sal	Accel	
3	16	7751.37	6.37	0.04	0.02	0.97	Temp	Sal		pH	Temp	Sal	Accel	pH	Temp	Sal		pH
4	17	7753.25	8.26	0.02	0.01	0.98	Temp	Sal	Accel	pH	Temp	Sal	Accel	pH		Sal	Accel	pH
5	17	7754.48	9.49	0.01	0.01	0.99	Temp	Sal	Accel	pH	Temp	Sal	Accel	pH	Temp		Accel	pH
6	16	7754.54	9.54	0.01	0.01	0.99	Temp		Accel	pH	Temp		Accel	pH	Temp	Sal	Accel	pH

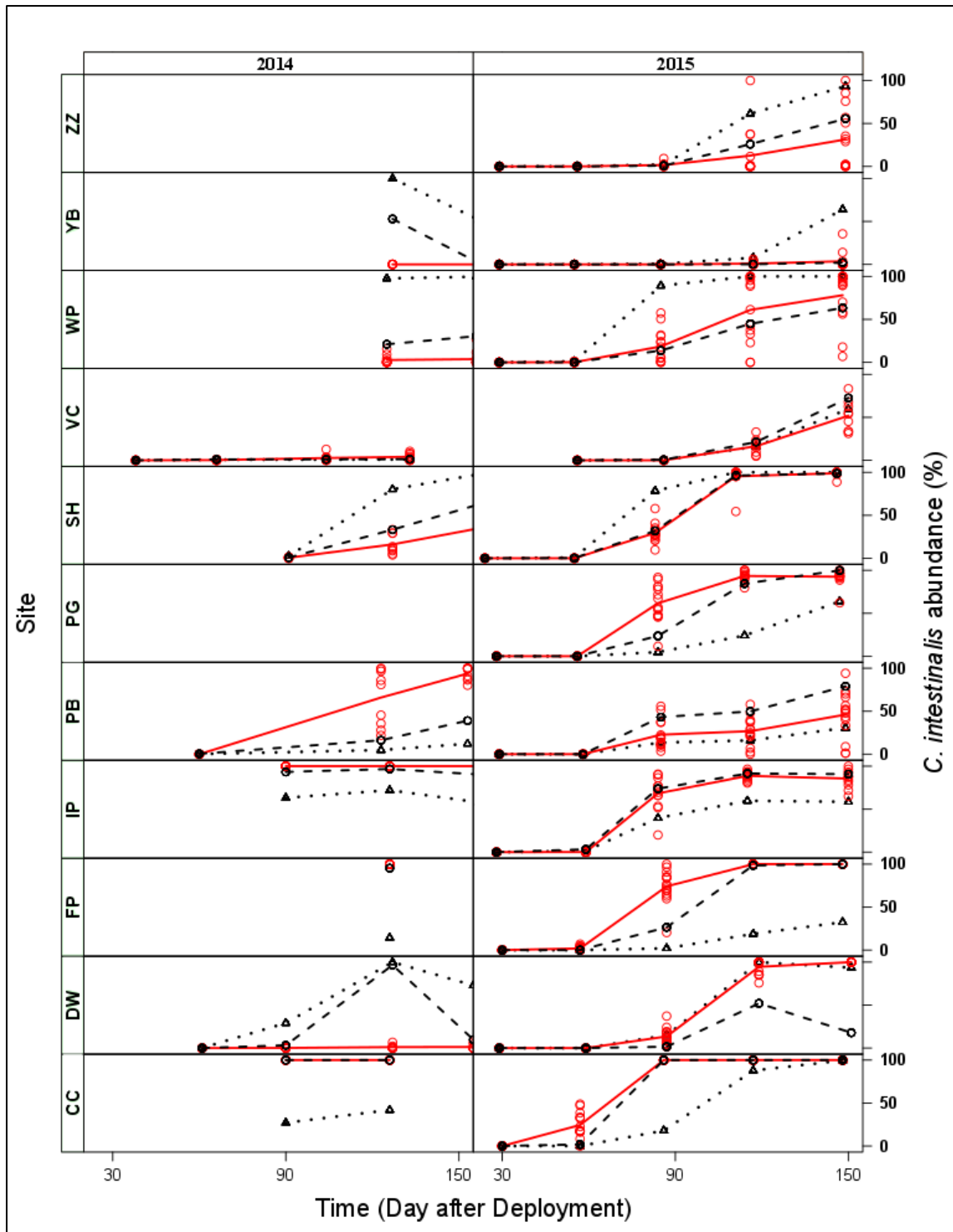
### E.3. Observed values vs. fitted values

The scale of missing data in 2014 can be seen clearly in the left panel of Fig. E.1. Due to pH, salinity, water flow and temperature being included in the development of the bottom-up candidate set, any measurement period with missing data had to be removed. The full-fit of the best bottom-up model (BBU) deviates from the observed average *C. intestinalis* growth, to a greater or lesser extent depending on the site and year (Fig. E.1). As is consistent throughout all of the models, inter-site variation of species-wide effects contribution to the fit is evident (Fig. E.1). There is also intra-site variation in species-wide effects between the two years. Two sites, VC and SH are fit very well by the species-wide effects in both years, if slightly less so for VC in 2015. DW is a prime example of a poor full-fit, due to the vastly different patterns of growth at DW in the two years. WP is similarly a poor fit as it was another site to have very different observed values of *C. intestinalis* between the years. SH is a very good fit in both years, despite having markedly higher growth in 2015. This could be an indication that the change in *C. intestinalis* growth at DW and WP between the years is largely due to factors not accounted for in this project, but the difference in growth at SH between the years is largely due to *C. intestinalis* responding consistently to variable abiotic conditions. ZZ, PG, and FP (except one measurement period of data) have no data to work with in 2014. The rate of growth and maximum population size at the end of the monitoring period is over predicted at ZZ by the species-wide effect, while rate of growth is under predicted at PG for 2015. The initial rate of growth is poorly fit for FP in 2015, but it then increases so that the final maximum population fit is quite accurate. PG on the other hand has a much lower final maximum population fit by the species-wide effects than was observed. YB, IP, and PB are quite poorly fit by the species-wide effects, while CC is very poorly fit by the species-wide effects.

The combined years best top-down model (BTD) full-fit deviates from the observed average *C. intestinalis* growth, to a greater or lesser extent depending on the given site and year (Fig. E.2). The species-wide effects contribute more so at some sites like VC and IP, while site-level effects clearly improve the fit at sites such as CC and WP. The species-wide effects fit worsen at some site when compared to the BBU, such as SH, while other improve, like IP.



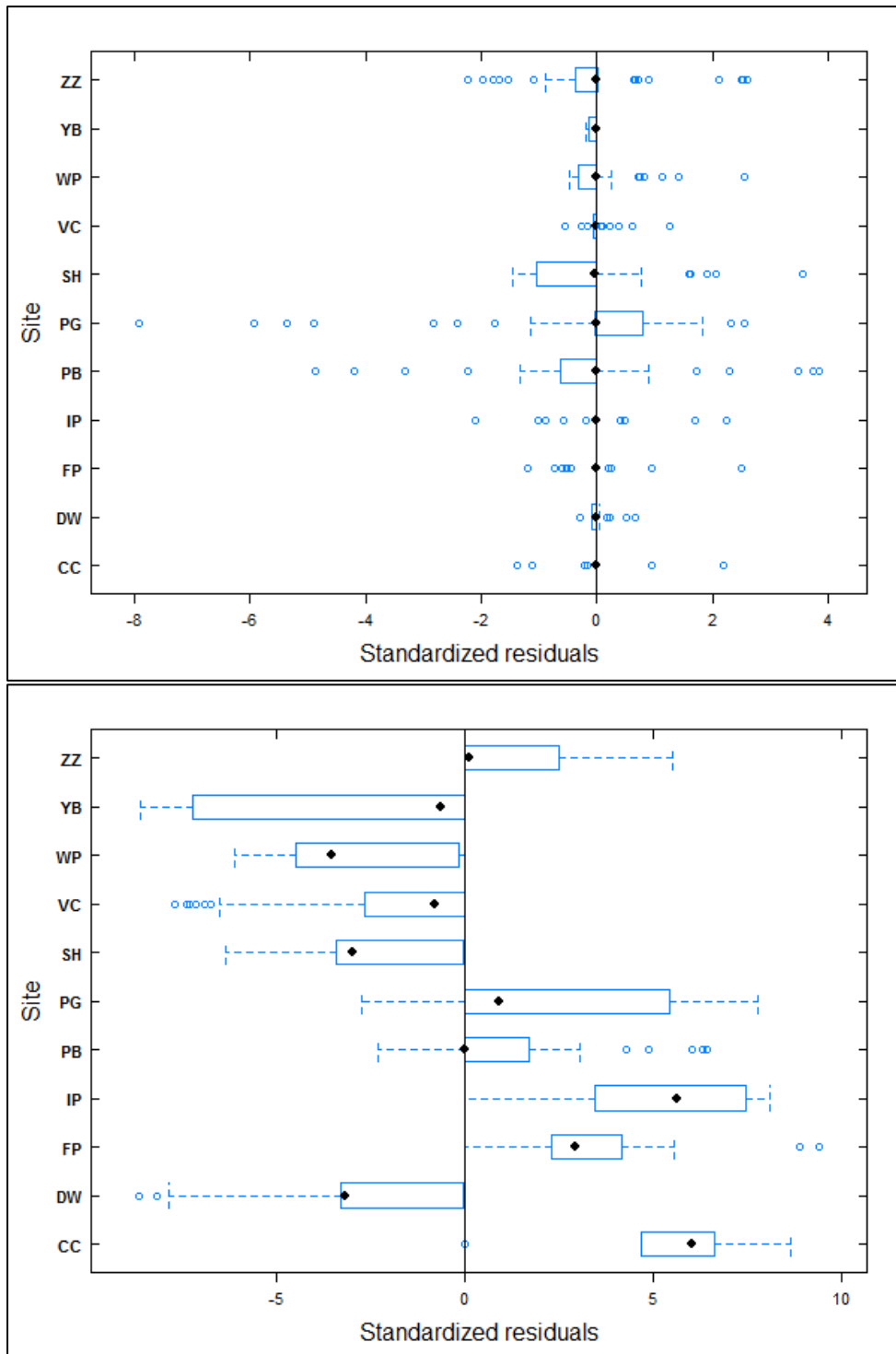
**Figure E.1: Red line and circle point markers represent the observed values. Dashed black line and circle markers represent the fitted values. The black dotted line and triangle markers represent the species-wide effect only fitted values.**



**Figure E.2:** Red line and circle point markers represent the observed values. Dashed black line and circle markers represent the fitted values. The black dotted line and triangle markers represent the species-wide effect only fitted values.

## Appendix F: Residual plots

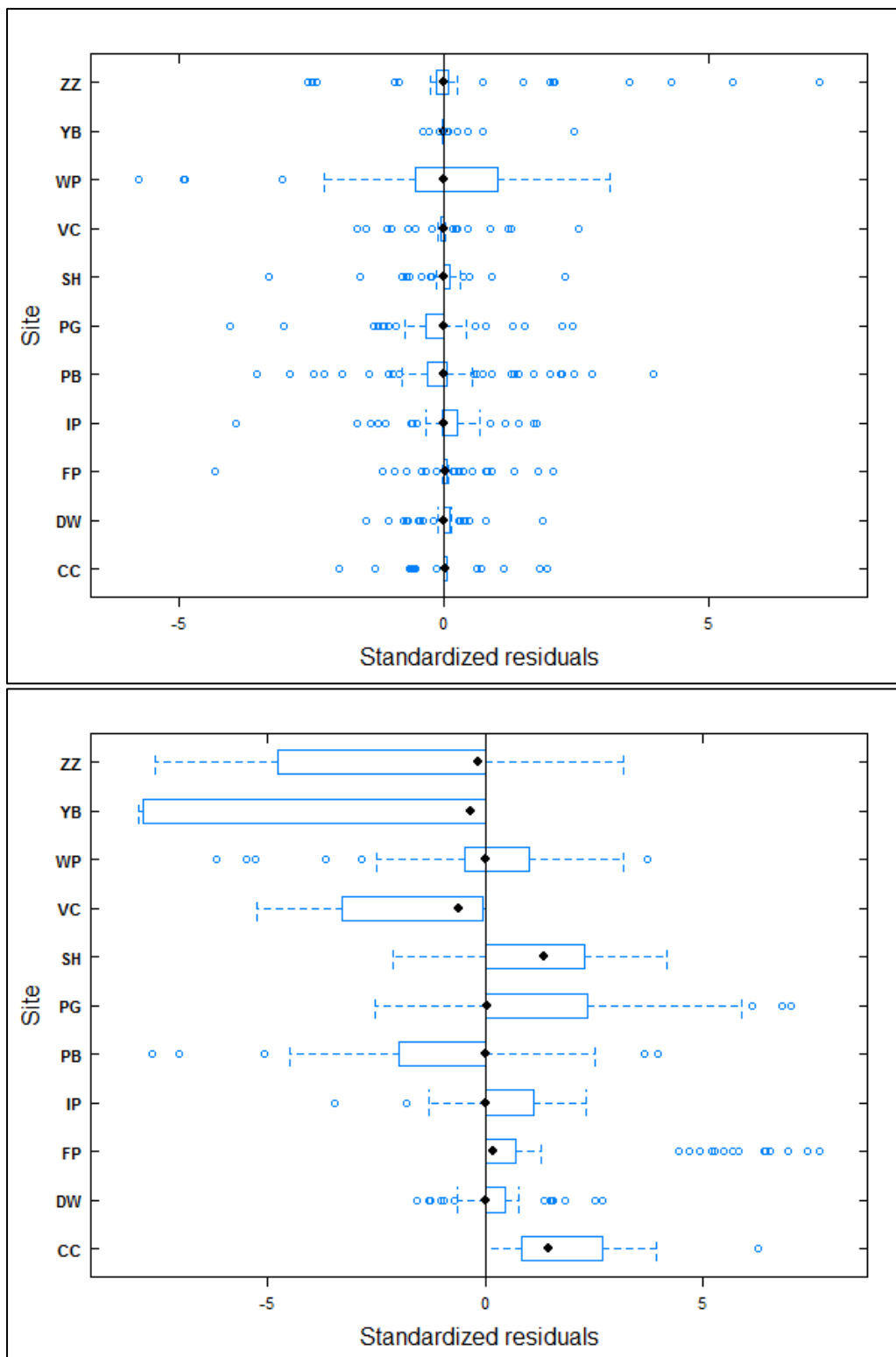
### F.1. Best temperature-only model (BTO) 2014



**Figure F.1: BTO 2014 residual boxplots by site, ordered alphabetically. The full-model residuals are displayed in the upper plot, while the species-wide effects residuals in isolation are displayed in the lower plot.**

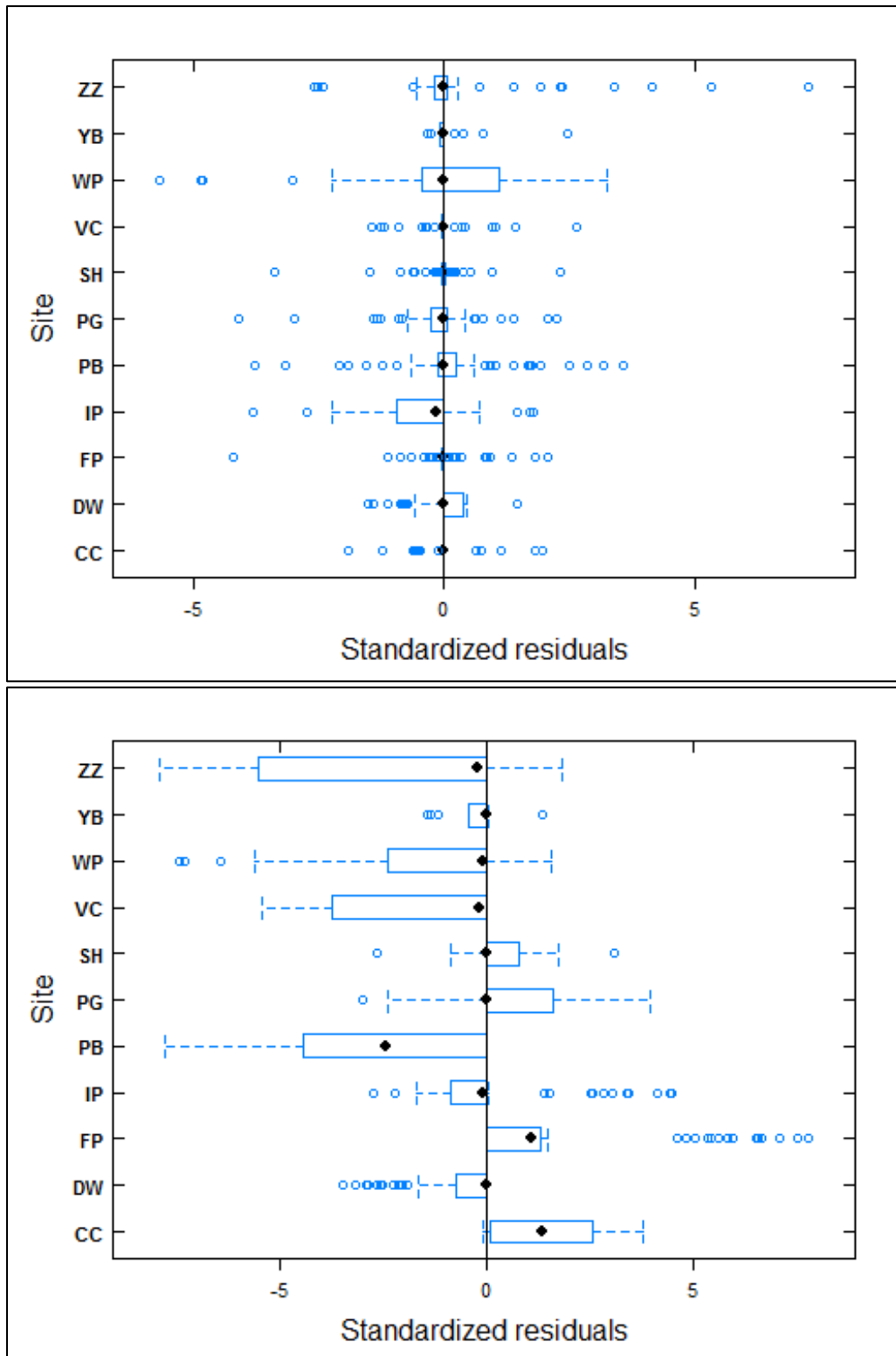
*F.2. Best bottom-up model (BBU) 2015*





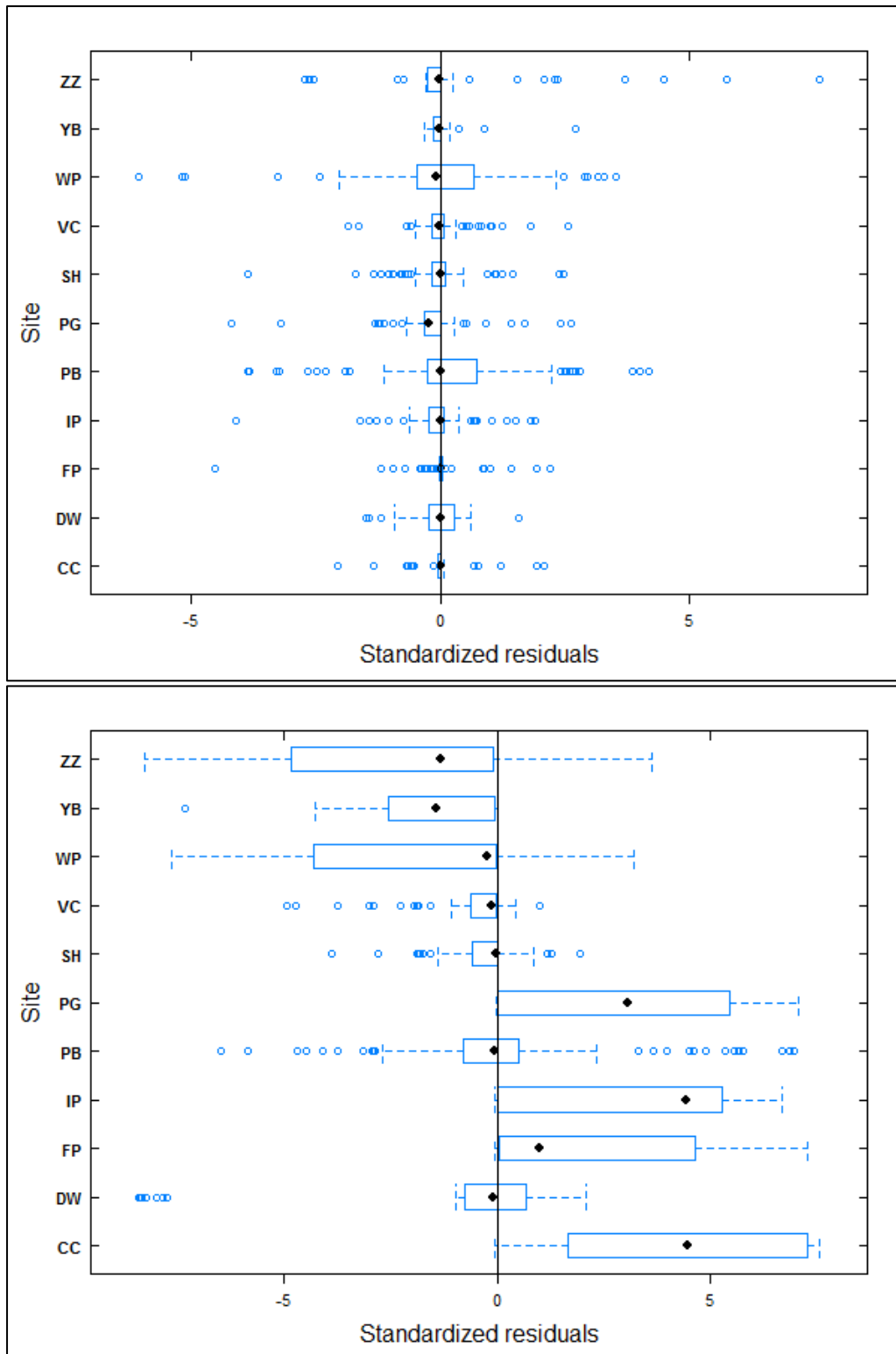
**Figure F.2: BBU 2015 residual boxplots by site, ordered alphabetically. The full-model residuals are displayed in the upper plot, while the species-wide effects residuals in isolation are displayed in the lower plot.**

*F.3. Best temperature-only model (BTO) 2015*



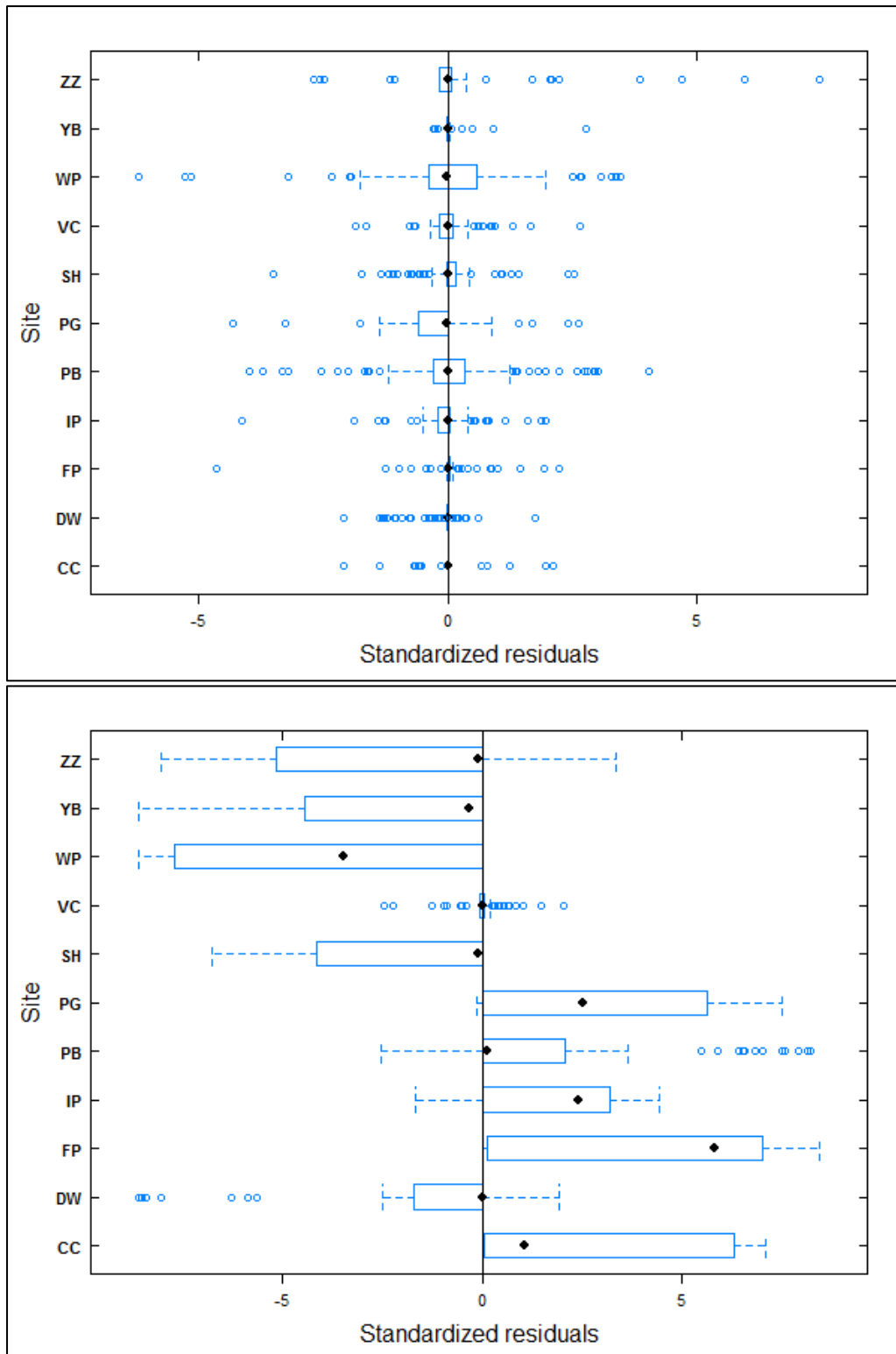
**Figure F.3: BTO 2015 residual boxplots by site, ordered alphabetically. The full-model residuals are displayed in the upper plot, while the species-wide effects residuals in isolation are displayed in the lower plot.**

*F.4. Best bottom-up model (BBU) combined*



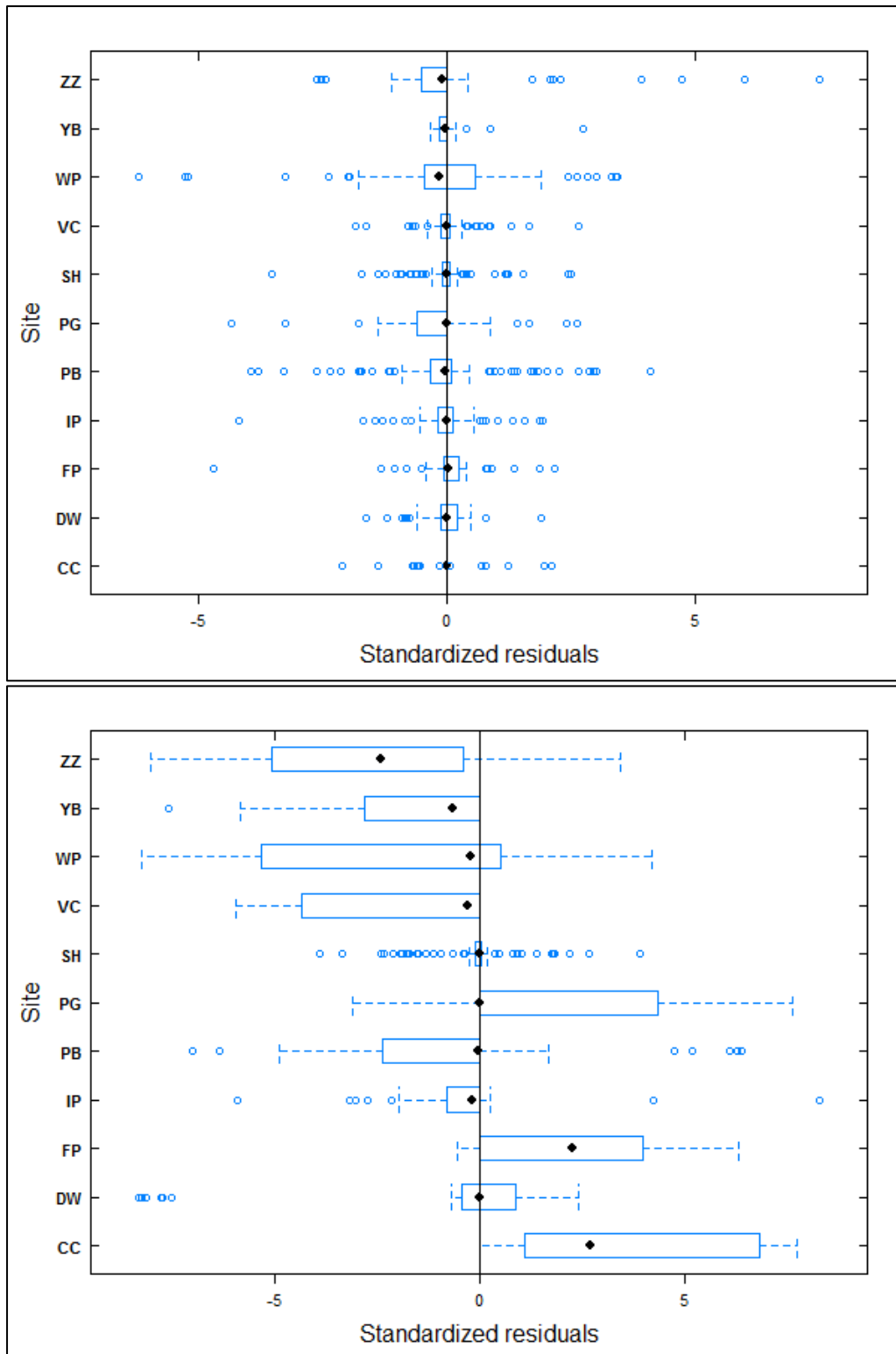
**Figure F.4: BBU combined residual boxplots by site, ordered alphabetically. The full-model residuals are displayed in the upper plot, while the species-wide effects residuals in isolation are displayed in the lower plot.**

*F.5. Best top-down model (BTD) combined*



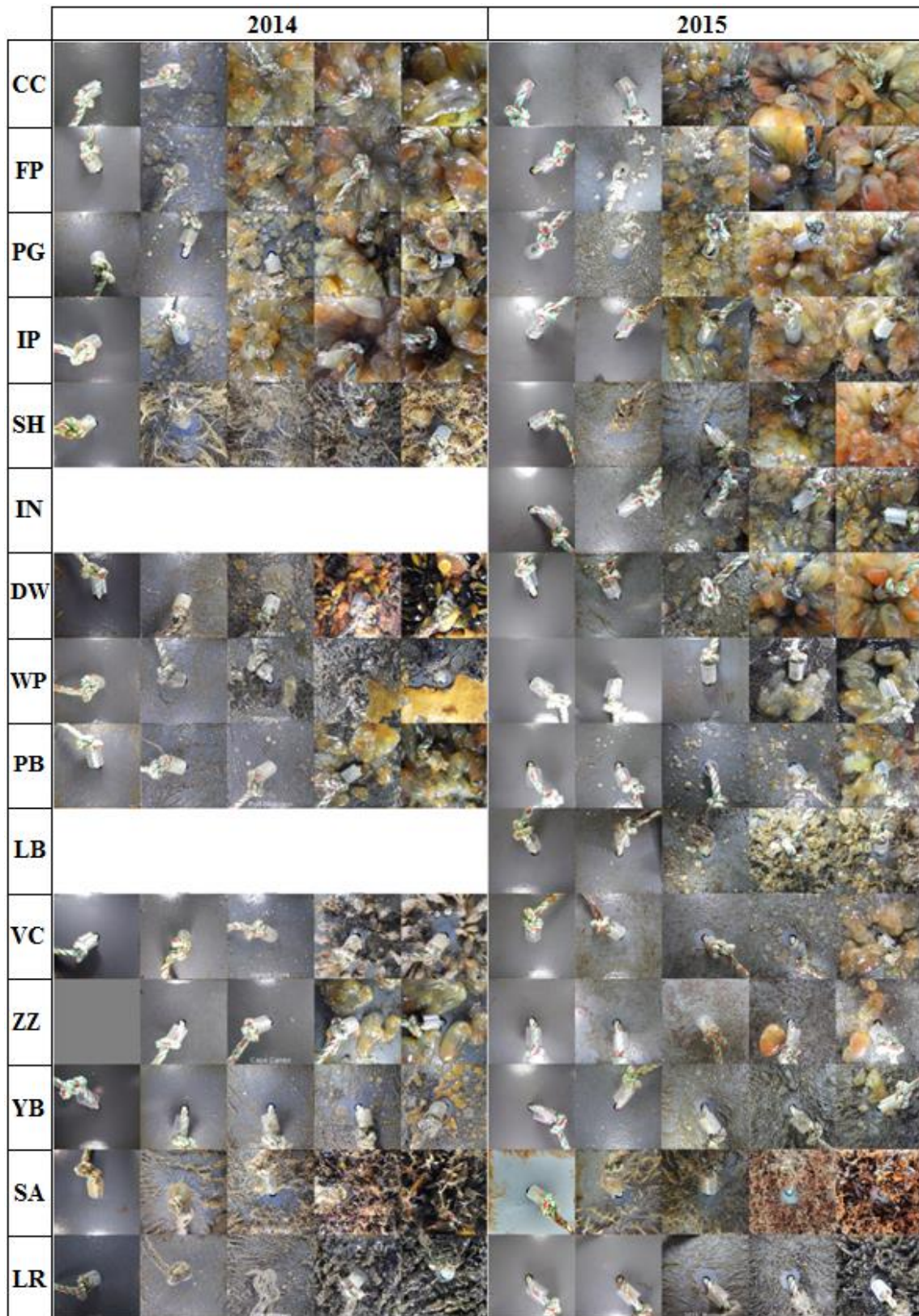
**Figure F.5: BTD combined residual boxplots by site, ordered alphabetically. The full-model residuals are displayed in the upper plot, while the species-wide effects residuals in isolation are displayed in the lower plot.**

*F.6. Second best top-down model (SBTD) combined*



**Figure F.6: SBTD combined residual boxplots by site, ordered alphabetically. The full-model residuals are displayed in the upper plot, while the species-wide effects residuals in isolation are displayed in the lower plot.**

**Appendix G: Settlement collector images**



**Figure G.1: Representative settlement plate images for each measurement period in both years for the expanded sites, showing the different biofouling communities present at that point in time. Ranked from highest 2015 mean *C. intestinalis* abundance.**

## Appendix H: ImageJ macro and R codes

### H.1. ImageJ

#### H.1.1. Image scaling

Open ImageJ > Process > Batch > Macro...

Choose Input Folder with all images that you desire the process to be implemented on.

Choose Output, a new folder where the resultant files will appear.

Insert code below into the large empty section:

```
run("Scale...", "x=- y=- width=3000 height=3000 interpolation=Bicubic average create title=create");
```

Press 'Process'.

#### H.1.2. Custom grid

Open ImageJ > Plugins > Macros > Record...

Make sure 'Record:' is set to 'Macro'. It should be as a default.

Create a custom name in the 'Name:' section, i.e. "Custom\_Grid.ijm".

Insert the code below (in red text) into the recorder:

```
nLines = 9;
color = "green";
if (nImages==0)
  newImage("Untitled", "8-bit black", 512, 512, 1);
run("Remove Overlay");
width = getWidth;
height = getHeight;
tileWidth = width/(nLines+1);
tileHeight = tileWidth;
xoff=tileWidth;
while (xoff<width) { // draw vertical lines
  makeLine(xoff, 0, xoff, height);
  run("Add Selection...", "stroke="+color);
  xoff += tileWidth;
}
yoff=tileHeight;
while (yoff<height) { // draw horizontal lines
  makeLine(0, yoff, width, yoff);
```

```

run("Add Selection...", "stroke="+color);
yoff += tileHeight;
}
run("Select None");

```

Press 'Create'

A new panel will appear. You can either run the macro from there, or save the macro to a specified folder and then install it as a permanent feature by following these steps.

Open ImageJ > Plugins > Macros > Install...

Choose the macro you saved in your specified folder.

Now when you open 'Macros' in ImageJ, the custom macro is an option to select.

## H.2. R Code

### H.2.1. Residual boxplots

#### H.2.1.1. Full-model residuals

```

plot(ExampleModel.nlme, SiteCode~resid(., type='pearson'),
     abline=0, aspect = 0.7,
     xlab = list(
       "Standardized residuals", cex = 1.3),
     ylab = list(
       "Site", cex = 1.3),
     scales = list(
       y=list(
         cex=0.9, font=2, tck =c(-1,-1)),
       x=list(
         cex=0.9, tck =c(1,-1), font=1)))

```

#### H.2.1.2. Species-wide effects residuals

```

plot(ExampleModel.nlme, SiteCode~resid(., type='pearson', level=0),
     abline=0, aspect = 0.7,
     xlab = list(
       "Standardized residuals", cex = 1.3),
     ylab = list(
       "Site", cex = 1.3),
     scales = list(
       y=list(
         cex=0.9, font=2, tck =c(-1,-1)),
       x=list(
         cex=0.9, tck =c(1,-1), font=1)))

```



## H.2.2. Observed values vs. fitted values

```
# Create sub-dataset that matches that used in the given model

datasub <- ExampleData

# requires datasub to be the appropriate subset used for the ExampleModel.nlme

# Create copy of datasub to avoid altering original

datasubcopy <- datasub

# Add new column "OF" (Observed/Fitted) and fill in with "observed" values corresponding to
#the observed "PercentCiona" values

datasubcopy$OF <- "observed"

# Assign the species-wide effects fitted values a name

fixfittedvals <- fitted(ExampleModel.nlme, level = 0)

# Assign a new datasubcopy that will be for the fitted values from ExampleModel.nlme

datasubwFR <- datasubcopy

# Replace the observed "PercentCiona" values with ExampleModel.nlme species-wide effects
# fitted values

datasubwFR$PercentCiona <- fixfittedvals

# Add new column "FR" (Fixed/Random) and fill in with "Fixed" values corresponding to the
Fixed-effect fitted "PercentCiona" values

datasubwFR$FR <- "Fixed"

# Assign a new datasubcopy that will be for the site-level effects fitted values from
# ExampleModel.nlme

datasubwRF <- datasubcopy

# Assign the fixed-effect fitted values a name

Ranfittedvals <- fitted(ExampleModel.nlme, level = 1)

# Replace the observed "PercentCiona" values ExampleModel.nlme species-wide effects #
fitted values

datasubwRF$PercentCiona <- Ranfittedvals
```

```
# Add new column "FR" (Fixed/Random) and fill in with "Random" values corresponding to
the Random-effect fitted "PercentCiona" values
```

```
datasubwRF$FR <- "Random"
```

```
# Merge the two versions of datasubcopy that have the fitted and observed values.
```

```
# Add new column "OF" (Observed/Fitted) and fill in with "observed" values corresponding to
the observed "PercentCiona" values
```

```
datasubcopy$FR <- "observed"
```

```
datasubcomb <- rbindlist(list(datasubwFR,datasubwRF, datasubcopy))
```

```
#####
# Plot the species-wide effects and site-level effects fitted values on the same plot #
#####
```

```
xypplot(PercentCiona~Daysafterdeploy |factor(SiteCode, levels(reorder(SiteCode,
MeanPercentCiona2015))),
  groups = FR,
  layout = c(1,11),
  xlab = list(
    "Time (Day after deployment)", cex = 1.35),
  ylab = list(
    "Site", cex = 1.35),
  ylab.right = list(
    label=substitute(paste(italic('C. intestinalis'), " relative abundance (%))), cex = 1.35),
  data = datasubcomb,
  type = c("a","p"), pch = c(2,1,1), cex = c(.6,.9, .8), col = c("black", "red", "black"), lty =
c("dotted","solid","dashed"), lwd = c(2.75, 2, 2),
  scales = list(
    y=list(
      at=seq(0,100,50),labels = c("0", "50", "100"), cex=0.85, font=2, tck =c(0,1),
alternating = c(2, 0)),
    x=list(
      at=seq(30,150,60), cex=.85, tck =c(1,-1), alternating = c(1, 1), font=1)),
  xlim = c(20, 155),
  strip = FALSE,
  strip.left = strip.custom(horizontal = FALSE, bg = 'white'),
  par.strip.text=list(cex=1, font=2))
```

### H.2.3. Confidence and prediction intervals

```
# 10-repeated 10-fold cross-validation
```

```
runs = 10
```

```
nfold = 10
```

```
results = c()
```

```
for (run in 1:runs)
```

```
{
```

```

split = sample(rep(1:nfold, length = n), n) # Split n in 10 sets, n = length of dataset
resample = lapply(1:nfold, function(x,spl) list(cal=which(spl!=x), val=which(spl==x)),
spl=split) # Create values for 10 cal/val sets.

```

```

# Here we loop through all cal/val combinations
for(fold in 1:nfold)

```

```

{
# Create the cal/val data sets for the current 'fold'
cal = Data_Edit[resample[[fold]]$cal,]
val = Data_Edit[resample[[fold]]$val,]

```

```

# Fit a model with the current calibration data set
model = ExamplModel.nlmf

```

```

# Calculate MEF and put it in results
observations = val$PercentCiona
predictions = predict(model, newdata=val)

```

```

MAE = mae(observations, predictions)
#RMSE = rmse(observations, predictions)
#MEF = MEF(predictions, observations)
#cor = cor(predictions, observations, method = "spearman")
results = c(results, MAE) # This attaches 'MAE' to any previous content of 'results'
#results = c(results, RMSE) # This attaches 'RMSE' to any previous content of 'results'
#results = c(results, MEF) # This attaches 'MEF' to any previous content of 'results'
#results = c(results, cor) # This attaches 'cor' to any previous content of 'results'

```

```

}
}

```

```

Cross_Val_Results <- mean(results)

```