UNIVERSITY OF TECHNOLOGY SYDNEY
School of Mathematical and Physical Sciences/Climate Change Cluster

Data assimilation techniques for aquatic ecological processes

by

Bojana Manojlovic

A Thesis Submitted
in Fulfillment of the
Requirements for the Degree

Doctor of Philosophy

Sydney, Australia

2021

Certificate of Original Authorship

I, Bojana Manojlovic declare that this thesis, is submitted in fulfilment of the requirements for the award of Doctor of Philosophy, in the Climate Change Cluster and School of Mathematical and Physical Sciences at the University of Technology Sydney.

This thesis is wholly my own work unless otherwise reference or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

This document has not been submitted for qualifications at any other academic institution.

This research is supported by the Australian Government Research Training Program.

Production Note:
Signature removed prior to publication.

April 2021

Acknowledgements

Thank you to my supervisors, Stephen Woodcock, Emlyn Jones, Mark Baird and Peter Ralph, and collaborator Christian Evenhuis, for your time, effort and guidance. Thank you to all data collaborators Tim Smith, Stacey Trevathan-Tackett, Peter Wood, Environmental Protection Agency Victoria, and the Bureau of Meteorology. This PhD was supported by University of Technology Sydney Climate Change Cluster (C3), Commonwealth Scientific and Industrial Research Organisation (CSIRO), the Industry Doctoral Training Centre (IDTC) and the Coastal Carbon Cluster.

Bojana Manojlovic

Sydney Australia, 2021.

Preface

To avoid repetition, the data assimilation technique used in this thesis is presented in Chapter 1 and each application (Chapter 2-4) that follows has been prepared in pseudo publication format with corresponding Introduction, Methods, Results and Discussion sections. Post examination each application chapter (Chapter 2-4) will be formatted as a manuscript for submission to a scientific peer-reviewed journal.

ABSTRACT

Data assimilation techniques for aquatic ecological processes

by

Bojana Manojlovic

Data assimilation applications for three aquatic ecological processes are explored through the Particle Marginal Metropolis Hastings technique. Datasets collected from five seagrass meadow sites at Port Phillip Bay, Australia are assimilated with a seagrass growth model for state and parameter estimation and prediction. Predictions are validated and assessed against their ability to accurately predict biomass and sea-floor cover for ongoing monitoring of seagrass meadows. The second application proposes a more flexible alternative to the conventional modelling of seagrass detrital decay using a time-varying decay rate. The proposed approach is tested against an in-situ field study and a laboratory experiment investigating the effects of temperature and nutrients on seagrass in Fagans Bay, Australia. The third application evaluates how combining high density sensor data with a data assimilating microalgae model informs on underlying processes occurring within environmental photo-bioreactors for the production of bio-fuels.

Contents

Certificate ii

Acknowledgments iii

Preface iv

Abstract v

List of Figures x

Abbreviation xxviii

1 Introduction 1

1.1 Data assimilation 1

1.2 State-space modelling in a general Bayesian hierarchical framework . . 7

1.3 Particle Marginal Metropolis-Hastings algorithm for obtaining the posterior distribution 9

1.3.1 Diagnostics: Convergence, mixing, and acceptance rates of the PMMH 12

1.3.2 Implementation in LibBi 14

2 Data assimilation of Port Phillip Bay seagrass meadows for state and parameter estimation and prediction 15

2.1 Introduction 16

2.2 Methods 19

2.2.1 Process Model: The seagrass growth model 19

2.2.2 Formulating the deterministic process model to incorporate stochasticity . . . 23
2.2.3 Parameter Model: Setting priors and initial conditions . . . 27
2.2.4 Data Model: Port Phillip Bay data description and collection methods . . . 32
2.3 Results . . . 39
2.3.1 Prior to posterior results for 2012 . . . 40
2.3.2 Posterior results for 2012 with a prediction for 2013 . . . 61
2.3.3 Prior to posterior results for 2013 . . . 68
2.3.4 Comparing the predicted 2013 results with the posterior 2013 results . . . 84
2.3.5 Posterior results across 2012-2013 . . . 103
2.4 Discussion . . . 130

3 Modelling seagrass detrital decay using a time-varying decay rate 134

3.1 Introduction . . . 135
3.2 Methods . . . 137
3.2.1 Random walk modelling approach for all phases of decay . . . 137
3.2.2 Set-up of seagrass decomposition field study investigating the impact of three different site conditions . . . 138
3.2.3 Experimental design of laboratory study investigating the effects of temperature and nutrient addition on time-varying decay rates . . . 138
3.3 Results . . . 141

3.3.1 Proof of concept results for the random walk model compared to multi-phase exponential decay of a seagrass decomposition field study 141

3.3.2 Data assimilated results of time-varying decay rates for an additional laboratory dataset investigating the effects of temperature and nutrient addition 146

3.4 Discussion 149

4 A data assimilating state-space model for algal growth under controlled conditions within a photo-bioreactor 154

4.1 Introduction 155

4.2 Methods 157

4.2.1 Data Model: Photo-bioreactor setup, experimental design and data collection methods 157

4.2.2 Data model: Data treatment, distributions and measurement error 159

4.2.3 Process model: Carbon chemistry 161

4.2.4 Process model: Gas transfer equilibrium concentrations for O_2 and CO_2 167

4.2.5 Process model: Flux into cells by photosynthesis and respiration 169

4.2.6 Process model: Dilution 169

4.2.7 Process model summary and parameter model 171

4.2.8 Design and setup of data assimilation model for a twin experiment and experimental data 172

4.3 Results 174

4.3.1 Posterior results of twin experiment where photosynthesis and respiration are constant through time 174

4.3.2 Posterior results of twin experiment where photosynthesis and respiration are changing through time 180

4.3.3 Posterior results with experimental data where photosynthesis, respiration and respiratory quotients are changing through time 187

4.3.4 Posterior results with experimental data where photosynthesis, respiration and respiratory quotients are changing through time and an offset on O_2 is introduced . . . 196

4.3.5 Posterior results with experimental data where photosynthesis, respiration and respiratory quotients are changing through time, O_2 has an offset and the O_2 observations were thinned further 206

4.4 Discussion 212

5 Concluding remarks 216

Appendices 221

A LibBi model code and MCMC diagnostics for Port Phillip Bay seagrass meadows 221

B LibBi model code and MCMC diagnostics for modelling detrital decay using a time-varying decay rate 239

C LibBi model code and MCMC diagnostics for microalgae growth within a photo-bioreactor 248

References 264

List of Figures

2.1 Equilibrium fraction of below-ground biomass (f_b) distribution for global seagrass and only *Zostera nigricaulis* used to set the prior. . . 28

2.2 The relationship between above-ground biomass and proportion cover with realisations of Equation 2.9 at Ω values 0.001, 0.003, 0.005, 0.01 and 0.1 (left). Histogram of Ω values calculated from above-ground biomass and proportion cover observations at all Port Phillip Bay sites (right). 29

2.3 Map of five seagrass sampling sites (seagrass), nearest BOM solar stations (sun), and EPA water quality sites (blue lines) in Port Phillip Bay, Victoria. Aerial image sourced from the Bureau of Meteorology. 33

2.4 Box-plots of above-ground biomass (g DW m^{-2}), proportion cover, and canopy height (cm) observations for each site through time (n=9 for each time point). The mean is indicated by a blue circle, and potential outliers by red crosses. Time is representative of months, starting from October 2011 (O) ranging to January 2014 (J). 35

2.5 Photosynthetically active radiation at water's surface, E_{wc} (mol photon m^{-2} d^{-1}), of the 5 BOM sites representative of the seagrass sampling sites. 37

2.6 Light attenuation coefficient k_d (m^{-1}) calculated from a surface bottle sample. 38

2.7 Above-ground biomass (g DW m^{-2}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Altona in 2012. 41

2.8 Above-ground biomass (g DW m^{-2}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Blairgowrie in 2012. 41

2.9 Above-ground biomass (g DW m^{-2}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Point Henry in 2012. 42

2.10 Above-ground biomass (g DW m^{-2}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Point Richards in 2012. 42

2.11 Above-ground biomass (g DW m^{-2}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Swan Bay South in 2012. 43

2.12 Proportion cover fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Altona in 2012. 43

2.13 Proportion cover fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Blairgowrie in 2012. 44

2.14 Proportion cover fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Point Henry in 2012. 44

2.15 Proportion cover fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Point Richards in 2012. 45

2.16 Proportion cover fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Swan Bay South in 2012. 45

2.17 Canopy height (m) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Altona in 2012. 46

2.18 Canopy height (m) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Blairgowrie in 2012. 46

2.19 Canopy height (m) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Point Henry in 2012. 47

2.20 Canopy height (m) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Point Richards in 2012. 47

2.21 Canopy height (m) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Swan Bay South in 2012. 48

2.22 Light attenuation coefficient (m^{-1}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Altona in 2012. 48

2.23 Light attenuation coefficient (m^{-1}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Blairgowrie in 2012. 49

2.24 Light attenuation coefficient (m^{-1}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Point Henry in 2012. 49

2.25 Light attenuation coefficient (m^{-1}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Point Richards in 2012. 50

2.26 Light attenuation coefficient (m^{-1}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Swan Bay South in 2012. 50

2.27 Long-term means of the first order autoregressive parameter prior and posterior distributions for each Port Phillip Bay site in 2012. . . 51

2.28 Parameter prior and posterior distributions for each Port Phillip Bay site in 2012. 52

2.29 Above-ground biomass posterior medians (solid blue line) and 95% credible intervals (shaded blue) for the first year of observations, and prediction medians (solid orange line) and 95% credible intervals (shaded orange) for the second year of observations. Observations (n=9 at each sampling time-point) are shown as box-plots with potential outliers (+). 62

2.30 Proportion cover posterior medians (solid blue line) and 95% credible intervals (shaded blue) for the first year of observations, and prediction medians (solid orange line) and 95% credible intervals (shaded orange) for the second year of observations. Observations (n=9 at each sampling time-point) are shown as box-plots with potential outliers (+). 63

2.31 Canopy height posterior medians (solid blue line) and 95% credible intervals (shaded blue) for the first year of observations, and prediction medians (solid orange line) and 95% credible intervals (shaded orange) for the second year of observations. Observations (n=9 at each sampling time-point) are shown as box-plots with potential outliers (+). 64

2.32 Light attenuation coefficient posterior medians (solid blue line) and 95% credible intervals (shaded blue) for the first year of observations, and prediction medians (solid orange line) and 95% credible intervals (shaded orange) for the second year of observations. Observations (n=1 at each sampling time-point) are shown as solid black circles. 65

2.33 Below-ground biomass posterior medians (solid blue line) and 95% credible intervals (shaded blue) for the first year of observations, and prediction medians (solid orange line) and 95% credible intervals (shaded orange) for the second year of observations. Observations (n=9 at each sampling time-point) are shown as box-plots with potential outliers (+). 66

2.34 Above-ground biomass (g DW m^{-2}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Altona in 2013. 69

2.35 Above-ground biomass (g DW m^{-2}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Blairgowrie in 2013. 69

2.36 Above-ground biomass (g DW m^{-2}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Point Henry in 2013. 70

2.37 Above-ground biomass (g DW m^{-2}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Point Richards in 2013. 70

2.38 Above-ground biomass (g DW m^{-2}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Swan Bay South in 2013. 71

2.39 Proportion cover fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Altona in 2013. 71

2.40 Proportion cover fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Blairgowrie in 2013. 72

2.41 Proportion cover fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Point Henry in 2013. 72

2.42 Proportion cover fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Point Richards in 2013. 73

2.43 Proportion cover fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Swan Bay South in 2013. 73

2.44 Canopy height (m) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Altona in 2013. 74

2.45 Canopy height (m) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Blairgowrie in 2013. 74

2.46 Canopy height (m) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Point Henry in 2013. 75

2.47 Canopy height (m) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Point Richards in 2013. 75

2.48 Canopy height (m) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Swan Bay South in 2013. 76

2.49 Light attenuation coefficient (m^{-1}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Altona in 2013. 76

2.50 Light attenuation coefficient (m^{-1}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Blairgowrie in 2013. 77

2.51 Light attenuation coefficient (m^{-1}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Point Henry in 2013. 77

2.52 Light attenuation coefficient (m^{-1}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Point Richards in 2013. 78

2.53 Light attenuation coefficient (m^{-1}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Swan Bay South in 2013. 78

2.54 Long-term means of the first order autoregressive parameter prior and posterior distributions for each Port Phillip Bay site in 2013. . . 79

2.55 Parameter prior and posterior distributions for each Port Phillip Bay site in 2013. 80

2.56 Above-ground biomass predicted medians (solid orange line) and 95% credible intervals (shaded orange), posterior medians (solid blue line) and 95% credible intervals (shaded blue) for the second year of observations. Observations (n=9 at each sampling time-point) are shown as box-plots with potential outliers (+). 85

2.57 Proportion cover predicted medians (solid orange line) and 95% credible intervals (shaded orange), posterior medians (solid blue line) and 95% credible intervals (shaded blue) for the second year of observations. Observations (n=9 at each sampling time-point) are shown as box-plots with potential outliers (+). 86

2.58 Canopy height predicted medians (solid orange line) and 95% credible intervals (shaded orange), posterior medians (solid blue line) and 95% credible intervals (shaded blue) for the second year of observations. Observations (n=9 at each sampling time-point) are shown as box-plots with potential outliers (+). 87

2.59 Light attenuation coefficient predicted medians (solid orange line) and 95% credible intervals (shaded orange), posterior medians (solid blue line) and 95% credible intervals (shaded blue) for the second year of observations. Observations (n=1 at each sampling time-point) are shown as solid black circles. 88

2.60 Canopy height long-term mean ($\mu_{\mathrm{h_c}}$) predicted vs posterior parameter kernel density estimates for the five Port Phillip Bay sites. 93

2.61 Light attenuation coefficient long-term mean (μ_{k_d}) predicted vs posterior parameter kernel density estimates for the five Port Phillip Bay sites. 93

2.62 Effective absorbance long-term mean (μ_{A}) predicted vs posterior parameter kernel density estimates for the five Port Phillip Bay sites. 94

2.63 Leaf mortality rate long-term mean (μ_{ζ_A}) predicted vs posterior parameter kernel density estimates for the five Port Phillip Bay sites. 94

2.64 Root/rhizome mortality rate long-term mean (μ_{ζ_B}) predicted vs posterior parameter kernel density estimates for the five Port Phillip Bay sites. 95

2.65 Translocation rate long-term mean (μ_{τ}) predicted vs posterior parameter kernel density estimates for the five Port Phillip Bay sites. 95

2.66 Compensation scalar irradiance long-term mean ($\mu_{\mathrm{E}_{comp}}$) predicted vs posterior parameter kernel density estimates for the five Port Phillip Bay sites. 96

2.67 Dry weight specific area of seagrass (Ω) predicted vs posterior parameter kernel density estimates for the five Port Phillip Bay sites. 96

2.68 Equilibrium fraction of biomass below-ground (f_b) predicted vs posterior parameter kernel density estimates for the five Port Phillip Bay sites. 97

2.69 Water column height (h_{wc}) predicted vs posterior parameter kernel density estimates for the five Port Phillip Bay sites. 97

2.70 Radiation use efficiency (r_{ue}) predicted vs posterior parameter kernel density estimates for the five Port Phillip Bay sites. 98

2.71 Seagrass diversity factor (SDF) predicted vs posterior parameter kernel density estimates for the five Port Phillip Bay sites. 98

2.72 Above-ground biomass (g DW m^{-2}) posterior medians (solid blue line) and 95% credible intervals (shaded blue) for the full set of observations. Observations (n=9 at each sampling time-point) are shown as box-plots with potential outliers (+). 104

2.73 Proportion cover posterior medians (solid blue line) and 95% credible intervals (shaded blue) for the full set of observations. Observations (n=9 at each sampling time-point) are shown as box-plots with potential outliers (+). 105

2.74 Canopy height (m) posterior medians (solid blue line) and 95% credible intervals (shaded blue) for the full set of observations. Observations (n=9 at each sampling time-point) are shown as box-plots with potential outliers (+). 106

2.75 Light attenuation coefficient (m^{-1}) posterior medians (solid blue line) and 95% credible intervals (shaded blue) for the full set of observations. Observations (n=1 at each sampling time-point) are shown as solid black circles. 107

2.76 Below-ground biomass (g DW m^{-2}) posterior medians (solid blue line) and 95% credible intervals (shaded blue) for the full 2012-2013 period. 108

2.77 Above-ground biomass (g DW m^{-2}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Altona in 2012-2013. 109

2.78 Above-ground biomass (g DW m^{-2}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Blairgowrie in 2012-2013. 109

2.79 Above-ground biomass (g DW m^{-2}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Point Henry in 2012-2013. 110

2.80 Above-ground biomass (g DW m^{-2}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Point Richards in 2012-2013. 110

2.81 Above-ground biomass (g DW m^{-2}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Swan Bay South in 2012-2013. 111

2.82 Proportion cover fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Altona in 2012-2013. 111

2.83 Proportion cover fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Blairgowrie in 2012-2013. 112

2.84 Proportion cover fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Point Henry in 2012-2013. 112

2.85 Proportion cover fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Point Richards in 2012-2013. . . . 113

2.86 Proportion cover fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Swan Bay South in 2012-2013. . . 113

2.87 Canopy height (m) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Altona in 2012-2013. 114

2.88 Canopy height (m) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Blairgowrie in 2012-2013. 114

2.89 Canopy height (m) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Point Henry in 2012-2013. 115

2.90 Canopy height (m) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Point Richards in 2012-2013. . . . 115

2.91 Canopy height (m) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Swan Bay South in 2012-2013. . . 116

2.92 Light attenuation coefficient (m^{-1}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Altona in 2012-2013. 116

2.93 Light attenuation coefficient (m^{-1}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Blairgowrie in 2012-2013. 117

2.94 Light attenuation coefficient (m^{-1}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Point Henry in 2012-2013. 117

2.95 Light attenuation coefficient (m^{-1}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Point Richards in 2012-2013. 118

2.96 Light attenuation coefficient (m^{-1}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Swan Bay South in 2012-2013. 118

2.97 Long-term means of the first order autoregressive parameter prior and posterior distributions for each Port Phillip Bay site in 2012-2013. 119

2.98 Parameter prior and posterior distributions for each Port Phillip Bay site in 2012-2013. 120

3.1 Experimental set-up: incubation containers mimicking natural decay conditions. 139

3.2 Posterior medians (solid blue line) and 95% credible intervals (shaded) of proportion of weight remaining detritus (W), with observations (black) at each site for leaf (left) and root/rhizome (right) tissue. 143

3.3 Posterior medians (solid blue line) and 95% credible intervals (shaded) of decay rate K (d^{-1}) at each site for leaf (left) and root/rhizome (right) tissue. 143

3.4 Prior (blue) and posterior (orange) distributions for parameter σ_r at each site for leaf (left) and root/rhizome (right) tissue. 144

3.5 Proportion of weight remaining (W) over the length of the experiment (80 days), for each temperature, nutrient and tissue combination. 147

3.6 Decay rates, K (d^{-1}), through time for each temperature, nutrient and tissue combination. 147

3.7 Stochasticity parameter, σ_r, for each temperature, nutrient and tissue combination. 148

3.8 Proportion of weight remaining (W) when the original and alternative fresh weight to dry weight conversion was applied to each treatment group. 152

4.1 Full O_2 (μM L^{-1}) and pH ($\log_{10}$(-mol/L H+)) datasets (black) with thinned O_2 and pH observations (yellow) over the four day experiment. 160

4.2 CO2SYS 161

4.3 Iterative (3rd iteration) vs exact solution for carbon chemistry CO_2 (μM L^{-1}), HCO_3 (μM L^{-1}), CO_3 (μM L^{-1}), pH ($\log_{10}$(-mol/L H+)) and state variables O_2 (μM L^{-1}), DIC (μM L^{-1}), and TA (μM L^{-1}). 166

4.4 Twin experiment posterior medians (solid blue line), 95% credible intervals (shaded blue), and synthetic observations (black) for O_2 and pH over 5 days where photosynthesis and respiration were constant through time. 176

4.5 Twin experiment posterior medians (solid blue line), 95% credible intervals (shaded blue), and synthetic observations (black) for DIC, TA and C_p over 5 days where photosynthesis and respiration were constant through time. 177

4.6 Twin experiment posterior medians (solid blue line) and 95% credible intervals (shaded blue) for HCO_3, CO_2 and CO_3 over 5 days where photosynthesis and respiration were constant through time. . . 177

4.7 Model parameter priors (orange), posteriors (purple) and true values (green) for the twin experiment where photosynthesis and respiration were constant through time. 178

4.8 Parameter trace-plots for the twin experiment where photosynthesis and respiration were constant through time. 179

4.9 Posterior medians (solid blue line), 95% credible intervals (shaded blue), and synthetic observations (black) of O_2 and pH for the twin experiment where photosynthesis and respiration were changing through time. X-axis is time in days. 182

4.10 Posterior medians (solid blue line), 95% credible intervals (shaded blue), and synthetic observations (black) of DIC, TA and C_p for the twin experiment where photosynthesis and respiration were changing through time. X-axis is time in days. 183

4.11 Posterior medians (solid blue line) and 95% credible intervals (shaded blue) of HCO_3, CO_2 and CO_3 for the twin experiment where photosynthesis and respiration were changing through time. X-axis is time in days. 184

4.12 Model parameter priors (orange), posteriors (purple) and true values (green) for the twin experiment where photosynthesis and respiration were changing through time. 185

4.13 Photosynthesis (P_1) and respiration (R_1) random walk posterior medians (solid blue), 95% credible intervals (shaded blue), and true values (green) for the twin experiment in the left panel (x-axis is time in days). σ_{rP} and σ_{rR} priors (orange) and posteriors (purple) in the centre, and traces on the right. 186

4.14 Parameter trace-plots for the twin experiment where photosynthesis and respiration were changing through time. 187

4.15 Posterior medians (solid blue line), 95% credible intervals (shaded blue), and experimental observations (black) for O_2 and pH when photosynthesis, respiration and respiratory quotients were changing through time. X-axis is time in days. 189

4.16 Posterior medians (solid blue line), 95% credible intervals (shaded blue), and experimental observations (black) for DIC, TA and C_p when photosynthesis, respiration and respiratory quotients were changing through time. X-axis is time in days. 190

4.17 Posterior medians (solid blue line) and 95% credible intervals (shaded blue) for HCO_3, CO_2 and CO_3 when photosynthesis, respiration and respiratory quotients were changing through time. X-axis is time in days. 191

4.18 Photosynthesis (P_1) and respiration (R_1) random walk posterior medians (solid blue line) and 95% credible intervals (shaded blue) after assimilating experimental data when photosynthesis, respiration and respiratory quotients were changing through time. X-axis is time in days. 192

4.19 Respiratory quotients (RQ_d and RQ_n) posterior medians (solid blue line) and 95% credible intervals (shaded blue) after assimilating experimental data when photosynthesis, respiration and respiratory quotients were changing through time. X-axis is time in days. 193

4.20 Model parameter priors (orange), and posteriors (purple) after assimilating experimental data when photosynthesis, respiration and respiratory quotients were changing through time. 194

4.21 Parameter trace-plots after assimilating experimental data when photosynthesis, respiration and respiratory quotients were changing through time. 195

4.22 Posterior medians (solid blue line), 95% credible intervals (shaded blue), and experimental observations (black) for O_2 and pH when photosynthesis, respiration and respiratory quotients were changing through time and an O_2 offset was introduced. X-axis is time in days. 199

4.23 Posterior medians (solid blue line), 95% credible intervals (shaded blue), and experimental observations (black) for DIC, TA and C_p when photosynthesis, respiration and respiratory quotients were changing through time and an O_2 offset was introduced. X-axis is time in days. 200

4.24 Posterior medians (solid blue line) and 95% credible intervals (shaded blue) for HCO_3, CO_2 and CO_3 after assimilating experimental data when photosynthesis, respiration and respiratory quotients were changing through time and an O_2 offset was introduced. X-axis is time in days. 201

4.25 Photosynthesis (P_1) and respiration (R_1) random walk posterior medians (solid blue line) and 95% credible intervals (shaded blue) after assimilating experimental data when photosynthesis, respiration and respiratory quotients were changing through time and an O_2 offset was introduced. X-axis is time in days. 202

4.26 Respiratory quotients (RQ_d and RQ_n) posterior medians (solid blue line) and 95% credible intervals (shaded blue) after assimilating experimental data when photosynthesis, respiration and respiratory quotients were changing through time and an O_2 offset was introduced. X-axis is time in days. 203

4.27 Model parameter priors (orange) and posteriors (purple) after assimilating experimental data when photosynthesis, respiration and respiratory quotients were changing through time and an O_2 offset was introduced. 204

4.28 Parameter trace-plots after assimilating experimental data when photosynthesis, respiration and respiratory quotients were changing through time and an O_2 offset was introduced. 205

4.29 Full O_2 (μM L^{-1}) and pH ($\log_{10}$(-mol/L H+)) datasets (black) with further thinned O_2 and pH observations (yellow). X-axis is time in days. 207

4.30 Posterior medians (solid blue line), 95% credible intervals (shaded blue), and further thinned experimental observations (black) for O_2 and pH. X-axis is time in days. 207

4.31 Posterior medians (solid blue line), 95% credible intervals (shaded blue), and further thinned experimental observations (black) for DIC, TA and C_p. X-axis is time in days. 208

4.32 Posterior medians (solid blue line) and 95% credible intervals (shaded blue) for HCO_3, CO_2 and CO_3 after assimilating further thinned experimental data. X-axis is time in days. 209

4.33 Photosynthesis (P_1) and respiration (R_1) random walk posterior medians (solid blue line) and 95% credible intervals (shaded blue) after assimilating further thinned experimental data. X-axis is time in days. 210

4.34 Respiratory quotients (RQ_d and RQ_n) posterior medians (solid blue line) and 95% credible intervals (shaded blue) after assimilating further thinned experimental data. X-axis is time in days. 211

4.35 Model parameter priors (orange) and posteriors (purple) after assimilating further thinned experimental data. 211

4.36 Parameter trace-plots after assimilating further thinned experimental data. 212

A.1 Directed acyclic graph (DAG) of LibBi model file seagrass.bi 231

A.2 2012 log-likelihoods across samples for each site after initial 10,000 samples were discarded as burn-in. 233

A.3 2012 parameter traces across samples for each site after burn-in was discarded. 234

A.4 2013 log-likelihoods across samples for each site after initial 10,000 samples were discarded as burn-in. 235

A.5 2013 parameter traces across samples for each site after burn-in was discarded. 236

A.6 2012-2013 log-likelihoods across samples for each site after burn-in was discarded. 237

A.7 2012-2013 parameter traces across samples for each site after burn-in was discarded. 238

B.1 Directed acyclic graph of the LibBi model file decay.bi 242

B.2 Laboratory study random walk (r) posterior through time for each temperature, nutrient and tissue combination. 244

B.3 Laboratory study stochasticity parameter (σ_r) trace across samples, for each treatment group after burn-in was discarded. 245

B.4 Laboratory study log-likelihoods across samples, for each treatment group after burn-in was discarded. 245

B.5 Field study random walk (r) posterior through time for each site and tissue type (left: leaf, right: root/rhizome). 246

B.6 Field study stochasticity parameter (σ_r) trace across samples for each site and tissue type (left: leaf, right: root/rhizome) after burn-in was discarded. 247

B.7 Field study log-likelihoods across samples for each site and tissue type (left: leaf, right: root/rhizome) after burn-in was discarded. . . . 247

C.1 Directed Acyclic Graph of the LibBi model file micro_iterative.bi . . . 260

C.2 Log-likelihood for the twin experiment where photosynthesis and respiration were constant through time. 261

C.3 Log-likelihood for the twin experiment where photosynthesis and respiration were changing through time. 261

C.4 Log-likelihood for the model with experimental data when photosynthesis, respiration and respiratory quotients were changing through time. 262

C.5 Log-likelihood for the model with experimental data when photosynthesis, respiration and respiratory quotients were changing through time and an O_2 offset was introduced. 262

C.6 Log-likelihood for the model with further thinned O_2 and pH experimental data. 263

Abbreviations

BHM: Bayesian Hierarchicial Model

BOM: Bureau of Meteorology

DA: Data Assimilation

EMS: Environmental Modelling Suite

EPA: Environmental Protection Agency

GP-GPU: General Purpose computing on Graphics Processing Unit

GPU: Graphics Processing Unit

IQR: Interquartile Range

KDE: Kernel Density Estimate

LibBi: Library for Bayesian Inference

MCMC: Markov Chain Monte Carlo

MH: Metroplois Hastings

ODE: Ordinary Differential Equation

OM: Organic Matter

PF: Particle Filter

PMMH: Particle Marginal Metropolis Hastings

PPB: Port Phillip Bay

RMSE: Root Mean Squared Error

SMC: Sequential Monte Carlo

SSM: State-Space Model

Chapter 1

Introduction

1.1 Data assimilation

Conventional ocean modelling consists of solving the model equations as accurately as possible, and then comparing the results with observations. While encouraging levels of quantitative agreement have been obtained, as a rule there is significant quantitative disagreement owing to many sources of error: model formulation, model inputs, computation and the data themselves. Computational errors aside, the errors made both in formulating the model and in specifying its inputs usually exceed the errors in the data. Thus it is unsatisfactory to have a model solution which is uninfluenced by the data. Bennett (Inverse Methods in Physical Oceanography,1st edn. Cambridge University Press, New York, p. 112, 1992)

Data assimilation quantitatively combines models with a diverse set of observations to provide a more consistent view of the physical and biological state of the ocean. It is essentially a problem in statistical estimation i.e. combining dynamical models and data to provide state or parameter estimates [43]. The application of data assimilation separates into two types of problems: i) Parameter estimation and ii) State estimation. The two approaches reflect different ways of fusing the model with the observations. Both approaches use data to constrain the evolution of the state variables, but provide different information. In parameter estimation, the model parameters are modified to fit the constraints, while in state estimation the state variables are modified to fit the observations [68].

Inverse methods are typically used for estimation of parameters (a list of studies which have used data assimilation methods to estimate model parameters of ecosystem models is provided in [43]), where models are considered deterministic functions of the parameters, a cost function is posed that measures the discrepancy between the model and data, and optimisation procedures are used to minimise it [34]. Consequently, the data are applied in a series of activities prior to integration of the assimilation model to obtain the best set of parameter values to match the observations. The parameters are then inserted into the model and integrated forward in time just like a free-run model. A common feature of biogeochemical (BGC) models is the large number of model parameters that must be specified to simulate the BGC model. Using observations to set these parameters can sometimes be applied, but for many parameters no direct observation estimate is available. The attraction of data assimilation is that it provides a means to generate a set of model parameters that reflects the observations, determines the values of the poorly known parameters, and provides insight into which parameters are constrained by the model [58]. A drawback, that should be kept in mind, of the parameter estimation approach is that it is possible to estimate unrealistic parameter values if important processes have been excluded from the model formulation.

Sequential data assimilation methods, widely used for state estimation, are recursive algorithms concerned with estimation of the ecological state as the system evolves through time [15]. In sequential data assimilation, there are no activities occurring prior to the integration of the model, instead, the model is integrated forward in time until data are available. The model results are modified by the data, typically using statistical procedures. The model is re-initialised and integrated forward in time to the next data assimilation event. These methods are used mostly for state and flux estimation, and only rarely for parameter optimisation. The goal is simply to provide the best state estimate by driving model outputs toward the

data through constant confrontation with data [43].

The use of data assimilation methods has led to considerable improvements in the predictive ability of numerical weather prediction (NWP) [8], and ocean forecasting [10, 74]. The aim of this project is to improve estimates and predictions of aquatic ecological processes, which can be achieved through marine BGC modelling. There are a number of critical differences between NWP, ocean forecasting and BGC modelling. Firstly, the structure and governing equations of marine BGC models are built on complex biological and physiological behaviours that are parametrised by simple functional forms, unlike NWP and ocean forecasting. Secondly, measurements of state variables and rate processes are difficult and often do not observe exactly what is modelled. For example, temperature measurements carry tiny analytical errors, while comparatively, measurements of in-situ Chlorophyll-a fluorescence which are loosely related to phytoplankton biomass, commonly carry errors of up to 50%. One of the biggest challenges for marine prediction is the identification and development of appropriate data assimilation methods to integrate the variety of data sources with marine ecosystem models. Observations of marine ecological state variables come from a variety of sources and sensors e.g. satellites, water samples, moored and profiling instruments. They are characterised by being sparse, noisy and non-Gaussian in their distribution [25].

Bayesian Framework

There are a number of drivers for the fusion of models and observations; the need for more objective and quantitative measures or metrics to evaluate model predictions against observations, the desire for more accurate predictions, and increasing demand for formal assessments of the uncertainty in model predictions (these are frequently used to inform decision making and risk management).

Errors enter into an integrated model-data system from at least three sources.

First, there are errors in the process of making observations, which typically provide a fragmented view of the underlying reality. All data come bundled with error. In particular, along with the obvious errors associated with measuring, manipulating, and archiving, there are other errors, such as discrete spatial and temporal sampling of naturally continuous systems. Consequently, there are always scales of variability that are unresolvable and that will further affect the observations [23]. One consequence is that we do not know the exact state of the system when we initialise dynamic models. Second, process models make simplifying assumptions and approximations. A model is a simplification of some well-chosen aspects of a system of interest, as it is impossible to completely capture all aspects in a model. Many ecological and Earth system models are dynamic models, predicting the evolution of system trajectories over time, and model errors are typically stochastic, leading to divergence of simulated trajectories over time [23]. Errors may also arise from the notion that observations collected may not be the same as the quantities modelled (e.g. modelling spatial means but having only point estimates as observations), these are known as errors of representativeness or change in support. Finally, process models incorporate a number of parameters whose values are uncertain and can include stochastic forcing as another source of uncertainty. Therefore, not only is the reality of what we are trying to model uncertain, our attempts to explain this are uncertain, and our measurements of our uncertain system are uncertain. Just as the physical and biological sciences have the notions of mass balance and energy balance, Cressie and Wikle [23] state that statistical science has a notion of variability balance. The total variability is modelled with variability due to measurement, variability due to using an uncertain model of how the world works, and variability due to uncertainty on parameters that control the measurement and model variabilities [23]. Although real-world systems may in principle be partially deterministic, our information is incomplete at each of the stages of observation, summarisation, and

inference, and consequently uncertainty interferes with our understanding. Therefore, by the time the inference stage is reached, the lack of certainty will influence how much knowledge we can gain from the data.

Hierarchical statistical modelling represents a way to express uncertainties through well defined levels of conditional probabilities. Following Berliner's [12] terminology: At the top level is the data model, which expresses the distribution of the data given a hidden process. This hidden process can be thought of as the "true process", uncorrupted by any measurement of it. At the level directly underneath the data model is the process model, which models scientific uncertainty in the hidden ("true") process through a probability distribution of the phenomenon of interest. It is quite possible that the process model is itself made up of sub-models whose uncertainties are also expressed at sub-levels through conditional probabilities. The result is a hierarchical model (HM). The components of a HM are conditional probability distributions that, when multiplied together, yield the joint probability distribution of all quantities in the model. The quantities of interest could be as simple as random variables and as complicated as space-time stochastic processes of random sets. All the conditional probability distributions specified in the HM typically depend on unknown parameters.

A HM is a Bayesian Hierarchical Model (BHM), if a lower level (underneath the data model and the process model) is established by specifying the joint probability distribution of all the unknown parameters. This probability model at the lowest level, which we define as the parameter model, completes the sequence: data model (top level) followed by process model (second level) followed by parameter model (bottom level). An alternative approach to specifying a parameter model is to assume that the parameters are fixed and to estimate them using the data, they are then substituted into the data model and the process model as if they were known. The result is an Empirical Hierarchical Model. It is possible to put prior distribu-

tions on some parameters and to estimate others (the term "prior distribution" is synonymous with "parameter model").

The last two decades have seen increasing advocacy of Bayesian approaches to data assimilation [13, 22, 104]. Bayesian methods typically yield posterior distributions for the inferred state and parameters. These can be particularly useful in applied contexts, where users may be interested in the probability distribution of performance measures derived from model predictions. A key attraction of the Bayesian approach is its ability to formally incorporate prior information about models and parameters. Given that the rationale for using mechanistic, process-based models is that they build on prior scientific knowledge about the structure and function of system components, it makes sense to use methods that allow this knowledge to be formally represented in model-data comparisons [82]. It is possible to use the Bayesian formulation, while disregarding prior information, through uninformative priors or empirical Bayes methods. In these cases, Bayesian methods can generally be shown to be equivalent to classical methods [92].

Within the broader Bayesian tradition, BHM offers a particularly attractive framework for the integration of mechanistic process models and observations. It provides a consistent, formal probabilistic framework combining error or uncertainty in model parameters, model state, model processes and observations [13, 22, 96]. This framework encourages the modeller to think carefully and systematically about the approximations and assumptions involved in process model formulation, about the observation process and the relationship between model state variables and observations, and about the relationship between model parameters and independent prior knowledge. Berliner [13] describes the deep integration of mechanistic and statistical modelling as "physical-statistical" modelling. The last two decades have seen an increase in research into Bayesian applications of environmental sciences, including but not limited to population dynamics and dispersal [6, 18, 21, 65, 96],

plant ecology and terrestrial surface fluxes [4, 84, 100], ocean circulation and climate [13, 14], and marine ecosystem and biogeochemistry [29, 30, 37, 58, 54, 68, 69, 76].

1.2 State-space modelling in a general Bayesian hierarchical framework

State-space models (SSM) are a special class of Bayesian Hierarchical Model (BHM), which explicitly define three hierarchical levels of uncertainty. The first level expresses the uncertainty coming from measurement error in observations, this is referred to as the data model. The second level of the hierarchy expresses the uncertainty of modelling scientific processes that are simplified or perhaps not completely understood, expressed as the process model. The third level of the hierarchy expresses uncertainty in parameters, known as the parameter model often referred to as the prior distribution. The general Bayesian Hierarchical modelling framework defines the joint probability distribution $[\boldsymbol{Y}, \boldsymbol{X}, \boldsymbol{\theta}]$ of observations $\boldsymbol{Y}$, processes $\boldsymbol{X}$, and parameters $\boldsymbol{\theta}$, as the likelihood times the prior,

$$[\boldsymbol{Y}, \boldsymbol{X}, \boldsymbol{\theta}] = [\boldsymbol{Y} \mid \boldsymbol{X}, \boldsymbol{\theta}][\boldsymbol{X}, \boldsymbol{\theta}] \tag{1.1}$$

where $[\boldsymbol{Y}]$ denotes the probability distribution of $\boldsymbol{Y}$, and $[\boldsymbol{X} \mid \boldsymbol{\theta}]$ denotes the conditional probability of $\boldsymbol{X}$ given $\boldsymbol{\theta}$. Using the laws of conditional probability, the joint distribution can be rewritten as,

$$[\boldsymbol{Y}, \boldsymbol{X}, \boldsymbol{\theta}] = [\boldsymbol{Y} \mid \boldsymbol{X}, \boldsymbol{\theta}][\boldsymbol{X} \mid \boldsymbol{\theta}][\boldsymbol{\theta}] \tag{1.2}$$

which represent the three levels of the hierarchy: the data model $[\boldsymbol{Y} \mid \boldsymbol{X}, \boldsymbol{\theta}]$, process model $[\boldsymbol{X} \mid \boldsymbol{\theta}]$, and parameter model $[\boldsymbol{\theta}]$.

We are interested in applying the BHM to modelling a dynamic system that has been observed through time. If we impose the Markov property on the state process $\boldsymbol{X}_t$, and at each time point t the observations $\boldsymbol{Y}_t$ depend only on the parameters $\boldsymbol{\theta}$

and state $\boldsymbol{X}_t$, our model becomes a special case of the BHM known as the state-space model. Therefore, the next state is only dependent on the current state and $\boldsymbol{X}$ is a Markov-model described by $[\boldsymbol{X}_t \mid \boldsymbol{X}_{t-1}, \boldsymbol{\theta}]$ for t=1,...,T. Observations at time t, $\boldsymbol{Y}_t$ are conditional only on the state $\boldsymbol{X}_t$ and assumed to be independent of observations at other times. These conditional independences can be factored into Equation 1.2 by splitting the prior into a parameter density, an initial state density, and a product of transition densities, while the likelihood becomes a product of observation densities. The joint probability distribution thus takes the following form,

$$[\boldsymbol{Y}_{1:T}, \boldsymbol{X}_{0:T}, \boldsymbol{\theta}] = [\boldsymbol{\theta}][\boldsymbol{X}_0 \mid \boldsymbol{\theta}](\prod_{t=1}^{T}[\boldsymbol{X}_t \mid \boldsymbol{X}_{t-1}, \boldsymbol{\theta}])(\prod_{t=1}^{T}[\boldsymbol{Y}_t \mid \boldsymbol{X}_t, \boldsymbol{\theta}]) \tag{1.3}$$

The main aim of Bayesian inference is to make inference on the process state and parameters given a set of observations. This is known as the posterior,

$$[\boldsymbol{X}_{0:T}, \boldsymbol{\theta} \mid \boldsymbol{Y}_{1:T}] = [\boldsymbol{\theta} \mid \boldsymbol{Y}_{1:T}][\boldsymbol{X}_{0:T} \mid \boldsymbol{\theta}, \boldsymbol{Y}_{1:T}] \tag{1.4}$$

The posterior is split into two parts, the first $[\boldsymbol{\theta} \mid \boldsymbol{Y}_{1:T}]$, parameter estimation, and the second $[\boldsymbol{X}_{0:T} \mid \boldsymbol{\theta}, \boldsymbol{Y}_{1:T}]$, state estimation conditioned on a particular set of parameters drawn from the first. An important distinction here is that the BHM framework is more comprehensive than the standard state-space model solutions which can be thought of as 'partially-Bayesian' in the sense that certain parameters or hyper-parameters are fixed for convenience or otherwise estimated in the absence of prior information. Complex models often lead to integrals that cannot be solved analytically making the posterior in Bayesian inference generally analytically intractable, and alternatively either an approximate closed form is fit or Monte Carlo sampling performed. Modern computing has enabled advanced research into these alternatives such as variational techniques, set of Kalman-based filters, and Markov Chain Monte Carlo (MCMC) methods.

In the next section, I will describe an algorithm from the MCMC family called

the Particle Marginal Metropolis Hastings (PMMH) [1] that solves the state and parameter inference problem for state-space models. It is a sampling based technique chosen for its accuracy, ability to handle high dimensional problems and generalisability to any state-space model. While it is computationally expensive, it was chosen over variational techniques which are fast, approximate, optimisation based algorithms, and Kalman filters that compute likelihoods directly, because there is no limitation to the types of models it can handle. While state estimation techniques like the Kalman filter work well, they are limited to particular types of SSM that assume linear and Gaussian transition and observation densities, and a Gaussian initial state model. A Sequential Monte Carlo (SMC) method, the particle filter [28, 39, 60], is much more generalisable and can be used for any SSM. Sequential Monte Carlo methods are alternative simulation-based algorithms for solving analytically intractable integrals. In these methods, a (partially) continuous probability distribution is approximated by a discrete distribution made of weighted draws termed particles. From one iteration of the algorithm to the next, particles are updated to approximate one distribution after another by changing the particle's location on the support of the distribution and their weights. The particle filter relies on importance sampling and has various techniques available, the original bootstrap particle filter algorithm is described in pseudo-code in Algorithm 1 and discussed in the next section along with the Metropolis-Hastings algorithm (Algorithm 2).

1.3 Particle Marginal Metropolis-Hastings algorithm for obtaining the posterior distribution

The Particle Marginal Metropolis-Hastings algorithm sequentially employs a particle filter for state estimation and the Metropolis-Hastings algorithm for parameter estimation. The particle filter estimates the marginal likelihood $[\boldsymbol{Y} \mid \boldsymbol{\theta}]$ which is then passed to the Metropolis-Hastings algorithm to draw samples of the parame-

ters from $[\boldsymbol{\theta} \mid \boldsymbol{Y}]$. The particle filter is then used again to sample from the state $[\boldsymbol{X} \mid \boldsymbol{\theta}, \boldsymbol{Y}]$.

State estimation is the sampling of $[\boldsymbol{X} \mid \boldsymbol{\theta}, \boldsymbol{Y}]$ conditioned on a particular set of parameters $\Theta = \theta$. For that given θ, the particle filter initialises particles by drawing P_x number of random samples, $\boldsymbol{x}^j(t_0)$ from $p(\boldsymbol{X}(t_0)|\boldsymbol{\theta})$, for j = 1, . . . , P_x, and weighting each uniformly with $w^j(t_0) = \frac{1}{P_x}$. It moves sequentially through observation times through a series of propagation, weighting and re-sampling steps (Algorithm 1). In the propagation step each particle is advanced to the next observation time, then weighted with the likelihood of the new observation. The re-sampling step is a key ingredient of the particle filter which restores the particle supply to equal weights by re-sampling particles with replacement, where the probability of each particle being drawn is proportional to its weight w^j. Particles with relatively small weights tend to be discarded, while particles with relatively large weights tend to be replicated several times. It is this process that determines the ancestor indices for the next time propagation. The particle filter is used to compute an unbiased estimate of the likelihood of $\boldsymbol{\theta}$, marginalised over $\boldsymbol{X}(t_{0:T})$ and to produce a sample by tracing a single particle back through its ancestry $\hat{\boldsymbol{x}}'(t_{i-1}) = \boldsymbol{x}^j(t_{i-1})$ (last step in Algorithm 1).

The likelihood estimate and state sample from an initial parameter set $\boldsymbol{\theta}$ are fed into the Metropolis-Hastings algorithm (Algorithm 2). A parameter jump is proposed $\boldsymbol{\theta}'$ by sampling from a proposal distribution that relies on the current parameter value $q(\boldsymbol{\theta}'|\boldsymbol{\theta})$. The likelihood l' and state sample $\boldsymbol{x}'(t_{0:T})$ are estimated for the proposed parameter. Then the acceptance ratio $\frac{l'p(\boldsymbol{\theta}')q(\boldsymbol{\theta}|\boldsymbol{\theta}')}{lp(\boldsymbol{\theta})q(\boldsymbol{\theta}'|\boldsymbol{\theta})}$ is calculated and if it is greater than or equal to a randomly drawn value from a uniform distribution between 0 and 1, the proposed parameter is accepted otherwise it is rejected and the parameter remains at its current value.

Algorithm 1 Particle Filter Algorithm $(\boldsymbol{\theta}, t_{r:s}) \rightarrow (\hat{l}, \hat{\boldsymbol{x}}'(t_{0:s}))$

if $r = 0$

    for each j $\in \{1, ..., P_x\}$

        Initialise particle j: $\boldsymbol{x}^j(t_0) \sim p(\boldsymbol{X}_0|\boldsymbol{\theta})$

        Initialise weight j: $w^j(t_0) = \frac{1}{P_x}$

for $i = r + 1, ..., s$

    for each j $\in \{1, ..., P_x\}$.

        Ancestor for particle j: $a^j(t_i) \sim multinomial(w(t_{i-1}))$

        Propagate particle j: $\boldsymbol{x}^j(t_i) \sim p(\boldsymbol{X}(t_i)|\boldsymbol{x}^{a^j(t_i)}(t_{i-1}), \boldsymbol{\theta})$

        Weight particle j: $w^j(t_i) = p(\boldsymbol{y}(t_i)|\boldsymbol{x}^j(t_i), \boldsymbol{\theta})$

Marginal likelihood: $\hat{l} = \prod_{i=1}^{s} (\frac{1}{P_x} \sum_{j=1}^{P_x} w^j(t_i))$

$j \sim multinomial(\boldsymbol{w}(t_s))$

$\hat{\boldsymbol{x}}'(t_s) = \boldsymbol{x}^j(t_s)$

for $i = s, ..., 1$

    $j = a^j(t_i)$

    $\hat{\boldsymbol{x}}'(t_{i-1}) = \boldsymbol{x}^j(t_{i-1})$

return $(\hat{l}, \hat{\boldsymbol{x}}'(t_{0:s}))$

Algorithm 2 Marginal Metropolis Hastings Algorithm $(\boldsymbol{\theta}, \boldsymbol{x}(t_{0:T}), l) \rightarrow (\boldsymbol{\theta}, \boldsymbol{x}(t_{0:T}), l)$

1: Initialise by setting $\boldsymbol{\theta}$ arbitrarily and run Algorithm 1 to obtain a likelihood estimate l and state sample $\boldsymbol{x}(t_{0:T})$
2: Propose a parameter jump: $\boldsymbol{\theta}' \sim q(\boldsymbol{\theta}'|\boldsymbol{\theta})$
3: Using the proposed parameter $\boldsymbol{\theta}'$ estimate the likelihood l' and state sample $\boldsymbol{x}'(t_{0:T})$ from Algorithm 1
4: Sample α uniformly between 0 and 1: $\alpha \sim U(0, 1)$
5: **if** $\alpha \leq \frac{l'p(\boldsymbol{\theta}')q(\boldsymbol{\theta}|\boldsymbol{\theta}')}{lp(\boldsymbol{\theta})q(\boldsymbol{\theta}'|\boldsymbol{\theta})}$
6: accept $\boldsymbol{\theta}'$ **return** $(\boldsymbol{\theta}', \boldsymbol{x}'(t_{0:T}), l')$
7: **else**
8: reject $\boldsymbol{\theta}'$ **return** $(\boldsymbol{\theta}, \boldsymbol{x}(t_{0:T}), l)$

1.3.1 Diagnostics: Convergence, mixing, and acceptance rates of the PMMH

PMMH theory states that if we run the PMMH sampler long enough, the samples we obtain will be samples from the joint posterior distribution (target/stationary distribution) and not dependent on the starting point. This raises some questions, one is how many samples do we need to run the Markov chain for it to explore the posterior distribution adequately. Another question is how to tell if the chain is mixing well. Mixing of the chain will determine how quickly we obtain useful results.

Trace-plots, marginal densities, log-likelihoods and correlation plots are all used as diagnostic tools to evaluate the convergence and mixing of MCMC algorithms. Trace-plots show the values a parameter took during the runtime of a chain and indicate how well the parameter space is being explored. Choice of proposal densities are crucial to the rapid convergence of the Metropolis-Hastings algorithm. The most important issue is how to scale the proposal (how to choose the proposal standard

deviation). If the proposal standard deviation is too small then the chain will move too slowly, if it is too large then the proposals will usually be rejected and the chain will not move enough. In either case, it will then take longer to come to equilibrium.

Marginal density plots are histograms of the values in the trace-plot, i.e. the distribution of the values of the parameter in the chain (you can also plot a smoothed version of the histogram as the marginal density plot). Marginal densities are an average over the values a parameter takes with all other parameters "marginalised", i.e. other parameters having any values according to their posterior probabilities. Often, marginal densities are treated as the main output of a Bayesian analysis by reporting their mean and standard deviation or median and interquartile range (IQR), but this practice shouldn't be looked at alone without further analysis and sometimes information on joint distributions may be informative. The reason is that marginal densities hide correlations between parameters, and if there are correlations, parameter uncertainties appear to be much greater in the marginals than they actually are.

The variance of the estimated likelihood ($\hat{l}$) will affect the convergence speed of the algorithm. Specifically, if the variance is very large, then the PMMH algorithm tends to get stuck for many consecutive iterations, not accepting any proposed moves. The reason for this is that large variance in the estimate is often related to a large skewness. The acceptance probability will tend to be small and the sampler can get stuck for many iterations. For the method to work satisfactorily, we therefore need to ensure that the variance in the likelihood ($\hat{l}$) is not overly large, which in turn implies that we need to use a sufficiently large number of particles in the underlying particle filter. What "sufficiently" means is problem dependent.

1.3.2 Implementation in LibBi

Library for Bayesian inference (LibBi) [72] is the software used to implement the PMMH for each SSM in the coming chapters. LibBi is its own modelling language, written in Perl with a parser and compiler and consisting of a C++ template library. LibBi is used for state-space modelling and Bayesian inference on high-performance computer hardware, intentionally designed for parallel computing.

The reason for choosing LibBi over other current Bayesian Inference software such as BUGS, JAGS or Stan is that it is specialised for state-space models rather than more general Bayesian hierarchical models. Its inference methods focus on SMC methods, and include Particle Markov Chain Monte Carlo (PMCMC), PMMH, Kalman filters and additional parameter optimisation routines, making it the most efficient and appropriate choice for our applications.

LibBi was run on CSIRO's high performance computing (HPC) cluster for each SSM application in the coming chapters. Pre (creation of forcing and observation files) and post processing (visualising and analysing results) was completed in Matlab. Python and R were also used in this thesis for targeted mathematical and statistical analyses respectively.

Chapter 2

Data assimilation of Port Phillip Bay seagrass meadows for state and parameter estimation and prediction

Abstract

In recent years, there has been increasing research interest into the modelling and monitoring of seagrass meadows, owing to the global role they play as significant carbon sinks, their ability to improve water quality and provide habitats for coastal fisheries. Of particular interest are methods examining seagrass dynamics, causes of decline and the capability to accurately predict seagrass loss for environmental management practices. Fusing spatially and temporally sparse data due to destructive sampling techniques with seagrass models that encapsulate key ecosystem relationships is the focus of this chapter. Data assimilation is used to draw ecologically informative state and parameter estimates from five sites in Port Phillip Bay, Victoria for the years 2012 and 2013. A prediction and further evaluation of seagrass biomass and proportion cover states is shown to be successful at three of the five sites. A quantitative meta-analysis of parameter values currently in the literature is used for the construction of informative priors and complements existing work towards evidence based priors and parameters for aquatic ecosystem modelling through a parameter library.

2.1 Introduction

Seagrass are marine flowering plants found in coastal waters that support important marine ecosystems. Their ecological importance not only lies in their ability to capture and bury carbon in the sediments below them [85, 86], but also through their ability to improve the surrounding water quality [75], and habitat provided for coastal fisheries [52]. Seagrass meadows have been found to be one of the most productive plant communities on Earth [70]; however 30% have been destroyed globally [94]. This causes concerns for the potential lost benthic communities that inhabit seagrass meadows, diminished coastal water quality, and the disturbance of stored carbon being released back into the atmosphere. Seagrass is a major benthic habitat within Port Phillip Bay (38° 09′ S, 144° 52′ E), a large semi-enclosed bay located in southern Victoria, Australia. *Zostera nigricaulis* (formerly *Heterozostera tasmanica*; [61, 53]) is the single dominant sub-tidal, meadow-forming seagrass in Port Phillip Bay. *Zostera muelleri*, *Halophila australis* and *Amphibolis antartica* are also present in smaller quantities [16]. While *Zostera nigricaulis* is commonly found in the west and south of the bay, its presence fluctuates over time and has recently declined [5].

The need to model and monitor seagrass meadows has attracted attention in recent years [75, 94] resulting in studies examining seagrass dynamics and the causes of their decline. Of particular interest are methods for predicting seagrass loss, since these processes have numerous negative impacts, and restoration efforts have generally proven to be difficult. While particular seagrass species have been extensively studied, there are still many whose response to environmental perturbations is uncertain. Literature shows that seagrasses exhibit species-specific physiological responses, proving a general quantitative description of a generic seagrass response difficult, as current mathematical models describing seagrass dynamics are sensitive to changes in parameters and initial conditions. There are a number of mechanis-

tic seagrass process models currently in the literature with differing foci, and on different scales including *Zostera capensis* Setchell responses to different freshwater related and river mouth-breaching scenarios [98], stochastic population model for *Zostera muelleri* to identify environmental indicators for Victoria's coastal environments [9], a bio-optical model of irradiance distribution and photosynthesis in seagrass canopies [103], a biogeochemical seagrass growth model driven by light and nutrient limitation [3], and a Bayesian belief network to quantify the relationship between seagrass and environmental drivers in the Great Barrier Reef [40].

Having a rigorous quantitative description of the state of a seagrass meadow that changes over time, requires a number of challenges to be overcome. Firstly, seagrass observations are sparse in space and time, so they alone cannot provide a complete representation of the state of a meadow. Furthermore, the sampling techniques used to obtain observations need to be considered carefully. Non-destructive sampling techniques such as presence/absence, percent cover and instrument probes are useful for monitoring changes in seagrass meadow health through time. While destructive sampling techniques such as sediment coring, elemental analyses, isotope analyses, are able to provide more precise data, they are damaging to the meadow and cannot be resolved on a dense temporal scale without disturbing and permanently damaging the meadow. Furthermore, parameters inferred from observational or experimental data are limited to the particular species and a narrow set of environmental conditions in question.

Data assimilation is a particularly attractive framework for combining seagrass models and observations as it readily addresses a number of the current problems faced during seagrass monitoring and subsequent modelling. The framework is based on a mechanistic model driving future estimates, and every time an observation is collected the joint state and parameter estimates are updated while conditioned on a defined level of uncertainty. The data assimilation framework easily deals with

missing data and unbalanced statistical experimental designs, while providing a platform to combine a number of different sources of data, which could be sampled at different times and on different scales. Another attractive feature of data assimilation is the ability to quantify uncertainty at various levels: during data collection, in model formulation, and in assigning parameters. After fusing the model and data, the final results are in the form of probability distributions; either as posterior estimates or posterior predictions, delivering solutions of highest probability density and uncertainty (credibility) surrounding those values. State-space models are particularly useful for risk-informed decision making as they provide rigorous probabilistic treatment of uncertainties in parameters, states and observations. There are current applications of data assimilation in marine biogeochemistry [76, 54, 29, 30], but none have been developed in seagrass biogeochemistry.

This chapter layout is as follows; firstly, a current seagrass growth model is described with a particular relationship between above-ground biomass and proportion cover. It attempts to capture and represent the main biogeochemical processes driving seagrass growth. Next, a literature search is conducted on the parameters that drive the process model to explore their values for distinct species and different environmental conditions. This sets up the parameter space for data assimilation of a state-space model. Then, a dataset collected to investigate different seagrass dynamics within a bay in south-east Australia over a two year period is explored. Finally, inference is drawn across the year 2012 and 2013 for the 5 sites in Port Phillip Bay, followed by a discussion comparing results from each year and site within the bay. The data assimilation procedure is applied to one site at a time. The year 2012 is also used as a training set to make predictions across all sites in 2013, which are then evaluated against the validation data and posteriors for 2013. Lastly, posterior inference is drawn across the full dataset (2012-2013) and compared to that of the individual years.

2.2 Methods

In this section we will formulate the problem as a state-space model (a specific class of Bayesian Hierarchical Model), which we introduced in Section 1. Firstly we will describe the process model, followed by the parameter model and lastly the data model.

2.2.1 Process Model: The seagrass growth model

The seagrass model chosen for this study is an adaptation of that used in the biogeochemical model developed by the Commonwealth Scientific and Industrial Research Organisation (CSIRO) for the Environmental Modelling Suite (EMS) (for full details see [3]). While EMS is a 60 variable ecosystem model, only those pools affecting seagrass biomass were included in this model formulation resulting in 5 state variables; above-ground biomass, below-ground biomass, detritus, dissolved inorganic carbon and oxygen. The detrital pool, dissolved inorganic carbon and oxygen pools were included in the model as variables solely to conserve mass and act as states where the seagrass processes of mortality, photosynthesis and respiration are accumulating. While dissolved inorganic carbon and oxygen are not explicitly coupled to a biogeochemical model for the purpose of this thesis, they could be coupled in a future iteration if needed. Consequently, these state variables have not been discussed in the results or discussion sections.

The seagrass sub-model within the full biogeochemical model evolves over timescales of months to years, whereas many of the other biogeochemical processes occur over hours to days. Therefore the state estimation of the seagrass sub-model is quasi-independent to the other biogeochemical processes, making it reasonable to look at the seagrass model in "quasi-isolation". A degree of caution should be used though when results are interpreted as feedbacks between the other "fast" biogeochemical process impact the "slow" seagrass process and these fast processes have not been

considered here.

There are two significant differences between the seagrass model in EMS and the one described in the following sections. The first is the model currency is biomass units of grams dry weight per metre squared (g DW m^{-2}) instead of grams Nitrogen per metre squared (g N m^{-2}) to match the biomass field observations that are collected in g DW m^{-2}. Next, EMS models seagrass growth as a minimum function describing either light, nitrogen or phosphorus limitation. To avoid over parametrisation and many unmeasurable free parameters, nutrient limitation was not included in the current model formulation, thus making the assumption that *Zostera nigricaulis* growth in Port Phillip Bay is only light limited. A recent study [51] found there was only one site in Port Phillip Bay that had a pronounced response to in-situ nutrient enrichment (indicating potential nutrient limitation) during the dates our observations were collected. The implications of the results on this site in particular, and this assumption in general, are discussed in Section 2.4.

Growth of above-ground seagrass biomass

To model seagrass growth, we start by describing the amount of light available at the top of the seagrass canopy. The amount of photosynthetically active radiation (E_c) that reaches the top of the seagrass canopy follows the Beer-Lambert equation [59] that describes light attenuation through the water column as an exponential decay governed by a vertical attenuation coefficient (k_d) that indicates water clarity, the light available at the water's surface (E_{wc}), and the distance the light has to travel from the top of the water column to the top of the seagrass canopy (h_{wc} - h_c),

$$E_c = E_{wc} e^{-k_d(h_{wc}-h_c)} \tag{2.1}$$

The amount of light that is used for photosynthesis (k_E) (mol photon m^{-2} d^{-1}) is a function of available light at the top of the seagrass canopy (E_c), proportion cover of seagrass (P_c), and effective leaf absorbance A (A= $A_L$sinB, where A_L =

leaf absorbance, and sinB= sine of the canopy bending angle),

$$\mathrm{k}_E = \mathrm{E}_c(1 - e^{-\mathrm{A}\Omega \mathrm{S}_A}) \tag{2.2}$$

Respiration (k_{resp}) (mol photon m^{-2} d^{-1}) is a function of compensation scalar irradiance (E_{comp}), proportion cover (P_c), and effective leaf absorbance (A),

$$\mathrm{k}_{resp} = \mathrm{E}_{comp}\mathrm{A}\Omega \mathrm{S}_A \tag{2.3}$$

Growth (g_A) (g m^{-2} d^{-1}) is then the difference between photosynthesis and respiration, unless respiration exceeds photosynthesis, then growth is 0 (i.e. we cannot have negative growth), and a conversion factor from mol photon to g DW known as radiation use efficiency (r_{ue}),

$$\mathrm{g}_A = \max(0, \frac{\mathrm{k}_E - \mathrm{k}_{resp}}{\mathrm{r}_{ue}}) \tag{2.4}$$

It is important to highlight here that this is a simplified formulation of photosynthesis where fast processes such as light-driven photosynthesis (occurring minutes to hours) are averaged over to allow for the use of daily rates (g_A and r_{ue}). To ensure a realistic growth rate is estimated, we limit the seagrass to not grow faster than a maximum growth rate (g_A^{max}),

$$\mathrm{g}_A = \min[\mathrm{g}_A^{max}\mathrm{S}_A, \max(0, \frac{\mathrm{k}_E - \mathrm{k}_{resp}}{\mathrm{r}_{ue}})] \tag{2.5}$$

Total above-ground seagrass growth (g m^{-2} d^{-1}) is then included in the ode as,

$$\frac{d\mathrm{S}_A}{dt} = \mathrm{g}_A - translocation - mortality \tag{2.6}$$

The mass conserving dissolved inorganic carbon (DIC) pool (g C m^{-3}) reflects photosynthesis,

$$\frac{d\mathrm{DIC}}{dt} = -\frac{550}{30}\frac{12}{14}0.0192\frac{\mathrm{g}_A}{\mathrm{h}_{wc}} \tag{2.7}$$

while the mass conserving oxygen (O_2) pool (g O m^{-3}) reflects respiration,

$$\frac{d\mathrm{O}_2}{dt} = \frac{716}{30}\frac{32}{14}0.0192\frac{\mathrm{g}_A}{\mathrm{h}_{wc}} \tag{2.8}$$

where 0.0192 [32] is the conversion from g DW to g N, the ratio of C:N:P is the Atkinson ratio 550:30:1 [2] for benthic plants, and the equation for organic matter formation gives the stoichiometric constants:
14 g N mol N^{-1}; 12 g C mol C^{-1}; 31 g P mol P^{-1}; 32 g O mol O_2^{-1}

Relationship between proportion cover and above-ground biomass

The following relationship is proposed [3] between proportion cover (P_c) and above-ground seagrass biomass (S_A), governed by the dry weight specific area of seagrass (Ω),

$$P_c = 1 - e^{-\Omega S_A} \tag{2.9}$$

where $\Omega\ S_A$ represents the leaf area index (LAI).

Translocation of biomass between seagrass tissue

Translocation of biomass can go in either direction; biomass can flow from above-ground to below-ground, or from below-ground to above-ground depending on the S_A:S_B ratio. Translocation (Υ) (g m^{-2} d^{-1}) is modelled using an equilibrium fraction of biomass below-ground (f_b) that *Zostera nigricaulis* prefers to maintain, the current state of above-ground (S_A) and below-ground (S_B) biomass, and the translocation rate (τ),

$$\Upsilon = (f_b(S_A + S_B) - S_B)\tau \tag{2.10}$$

An important note here is that increasing below-ground biomass only occurs through translocation. Updating our differential equations we now have,

$$\frac{dS_A}{dt} = g_A - \Upsilon - mortality \tag{2.11}$$

$$\frac{dS_B}{dt} = \Upsilon - mortality \tag{2.12}$$

Mortality and detritus

Seagrass leaves die and leave the above-ground biomass pool at a leaf mortality rate (ζ_A) (d^{-1}),

$$\frac{d\mathrm{S}_A}{dt} = \mathrm{g}_A - \Upsilon - \zeta_A \mathrm{S}_A \tag{2.13}$$

Similarly, seagrass roots/rhizomes die and leave the below-ground biomass pool with a root/rhizome mortality rate (ζ_B) (d^{-1}),

$$\frac{d\mathrm{S}_B}{dt} = \Upsilon - \zeta_B \mathrm{S}_B \tag{2.14}$$

All of the dead above-ground and dead below-ground biomass forms the detritus pool (D) in the sediment, taking into account porosity (ϕ) and thickness of the top sediment layer (h_{sed}),

$$\frac{d\mathrm{D}}{dt} = \frac{\zeta_A \mathrm{S}_A + \zeta_B \mathrm{S}_B}{\mathrm{h}_{sed}\phi} \tag{2.15}$$

A summary of the final process model is in Table 2.1.

2.2.2 Formulating the deterministic process model to incorporate stochasticity

To capture the effects of approximations and errors in the process representation described in section 2.2.1, we take our deterministic process model (Table 2.1) and introduce stochasticity through an autoregressive process. Stochasticity could also have been introduced through additive error, either as a continuous Weiner process added to the ordinary differential equations or as a Gaussian error term at each time step in a discretised application. As this approach violates conservation of mass, we instead adopt the approach defined in [76] where a stochastic version of a deterministic marine biogeochemical model for Nutrient-Phytoplankton-Zooplankton-Detritus (NPZD) was formulated. Similarly to the phytoplankton and zooplankton communities, we introduce stochasticity by replacing select constant parameters in the deterministic model with stochastic processes that represent within community

variance or unresolved model error. This seems more biologically intuitive than independently drawing community properties from an underlying distribution at each time step, as it allows for some level of persistence in seagrass community composition through time by being treated as the outcome of a first-order autoregressive (AR(1)) stochastic process. Each stochastic process B is defined as,

$$\mathrm{B}(t+\Delta t) = \mathrm{B}(t).(1-\Delta t/\mathrm{St}) + \mathrm{r}_B(t)\Delta t/\mathrm{St}, \text{ for } |1-\Delta t/\mathrm{St}| < 1 \tag{2.16}$$

where Δt is the discrete time step, St is the characteristic time of the autoregressive process, and $\{\mathrm{r}_B(t)\}$ represents a sequence of independent and identically distributed random variables with distribution $[\mathrm{r}_B]$. A detailed derivation can be found in Appendix B of the supplementary material of [76].

The parameters we choose to be represented as autoregressive processes are translocation rate (τ), leaf mortality rate (ζ_A), root/rhizome mortality rate (ζ_B), total compensation scalar irradiance (E_{comp}), effective absorbance (A), canopy height (h_c), and light attenuation coefficient (k_d). This moves them from the parameter space to the state space, but keeps the means of the autoregressive processes as unknown parameters to be estimated and prior distributions to be assigned. Prior distributions are also assigned to initial conditions of all state variables, and it is important to note that the introduction of stochasticity to parameters can cause bias if the initial conditions for the state variables are not included as part of the estimation. One important distinction to note here is that the growth rate is treated as a static parameter with a dependence on effective absorbance, which is treated as a dynamic parameter. This alternative to representing photosynthesis from the usual approach of photosynthesis irradiance curves (which requires fine-time resolution) has considerable computational advantages that allow PMMH to be feasible for this application. The parameters are interpretable as daily averages and a rate of change, respectively, and the estimates are also ecologically informative.

Seagrass process model equations

$$\mathrm{E}_c = \mathrm{E}_{wc} e^{-\mathrm{k}_d(\mathrm{h}_{wc}-\mathrm{h}_c)} \tag{2.17}$$

$$\mathrm{k}_E = \mathrm{E}_c(1 - e^{-\mathrm{A}\Omega \mathrm{S}_A}) \tag{2.18}$$

$$\mathrm{k}_{resp} = \mathrm{E}_{comp}\mathrm{A}\Omega\mathrm{S}_A \tag{2.19}$$

$$\mathrm{g}_A = \min[\mathrm{g}_A^{max}\mathrm{S}_A, \max(0, \frac{\mathrm{k}_E - \mathrm{k}_{resp}}{\mathrm{r}_{ue}})] \tag{2.20}$$

$$\Upsilon = (\mathrm{f}_b(\mathrm{S}_A + \mathrm{S}_B) - \mathrm{S}_B)\tau \tag{2.21}$$

$$\frac{d\mathrm{S}_A}{dt} = \mathrm{g}_A - \zeta_A \mathrm{S}_A - \Upsilon \tag{2.22}$$

$$\frac{d\mathrm{S}_B}{dt} = -\zeta_B \mathrm{S}_B + \Upsilon \tag{2.23}$$

$$\frac{d\mathrm{D}}{dt} = \frac{\zeta_A \mathrm{S}_A + \zeta_B \mathrm{S}_B}{\mathrm{h}_{sed}\phi} \tag{2.24}$$

$$\frac{d\mathrm{DIC}}{dt} = -\frac{550}{30}\frac{12}{14}0.0192\frac{\mathrm{g}_A}{\mathrm{h}_{wc}} \tag{2.25}$$

$$\frac{d\mathrm{O}_2}{dt} = \frac{716}{30}\frac{32}{14}0.0192\frac{\mathrm{g}_A}{\mathrm{h}_{wc}} \tag{2.26}$$

Table 2.1 : Table of seagrass process model equations.

Symbol	**Description**	**Units**
S_A	Above-ground seagrass biomass	g DW m^{-2}
S_B	Below-ground seagrass biomass	g DW m^{-2}
D	Detritus in the sediment	g DW m^{-3}
DIC	Water column DIC concentration	g C m^{-3}
O_2	Water column oxygen concentration	g O m^{-3}
P_c	Proportion cover	-
Υ	Translocation rate	g DW m^{-2} d^{-1}
ζ_A	Leaf mortality rate	d^{-1}
ζ_B	Root/rhizome mortality rate	d^{-1}
E_{wc}	Total photosynthetically active radiation at water's surface	mol photon m^{-2} d^{-1}
E_c	Solar irradiance at the top of the seagrass canopy	mol photon m^{-2} d^{-1}
E_{comp}	Total compensation scalar irradiance	mol photon m^{-2} d^{-1}
g_A	Above-ground growth	g DW m^{-2} d^{-1}
g_A^{max}	Maximum above-ground growth rate	d^{-1}
f_b	Equilibrium fraction of biomass below-ground	-
Ω	Dry weight specific area of seagrass	(g DW m^{-2})$^{-1}$
τ	Translocation rate	d^{-1}
A	Effective Absorbance	-
ϕ	Sediment porosity	-
h_{sed}	Top sediment layer thickness	m
r_{ue}	Radiation use efficiency	g DW (mol photon)$^{-1}$
h_{wc}	Water column height	m
h_c	Seagrass canopy height	m
k_d	Light attenuation coefficient	m^{-1}
St	Time-scale of community composition changes	d
SDF	Seagrass diversity factor	-

Table 2.2 : Table of seagrass model variables and parameters.

2.2.3 Parameter Model: Setting priors and initial conditions

In data-scarce situations, informative priors become increasingly important to guide the model when there are no observations available. Seagrass datasets are notoriously sparse, with weeks, months and sometimes seasons between consecutive sampling time-points. Setting priors for biogeochemical model parameters can be difficult (see [82] for a detailed justification on evidence-based priors for aquatic ecosystem models) as often parameters take small values (close to but not below zero) and occasionally can become large. This reasoning drives the choice to set most of the parameter priors for our seagrass process model as log normal distributions (Table 2.3), to allow for long right tails with low probabilities. A literature review was conducted, and for those parameters where estimates could be found from previous studies, priors were set by deriving prior means and coefficients of variation from the distributions of parameter values reported from those studies. Where available, *Zostera nigricaulis* data was used, otherwise a global seagrass prior for that parameter was set by obtaining a range of values for a range of seagrass species under varying environmental conditions. Of the global priors, those parameters that clustered around species-specific values [33] (such as fraction of below-ground biomass), and had substantial sample sizes were identified and priors were made more informative by taking only those values of species similar to *Zostera nigricaulis*.

A global prior for equilibrium fraction of biomass below-ground (f_b) ranged from 0.28 to 0.89 covering species *Zostera capricorni, Zostera marina, Zostera noltii, Zostera nigricaulis, Thalassia hemprichii, Thalassia testudinum, Halodule uninervis, Halodule wrightii, Halophila decipiens, Halophila johnsonii, Halophila ovalis, Halophila stipulacea, Cymodocea nodosa, Cymodocea rotundatam, Cymodocea serrulata, Enhalus acoroides, Phyllospadix scouleri, Phyllospadix torreyi, Posidonia angustifolia, Posidonia oceanica, Syringodium filiforme,* and *Syringodium isoetifolium* [33, 48, 87]. The local prior based only on *Zostera nigricaulis* clustered values in

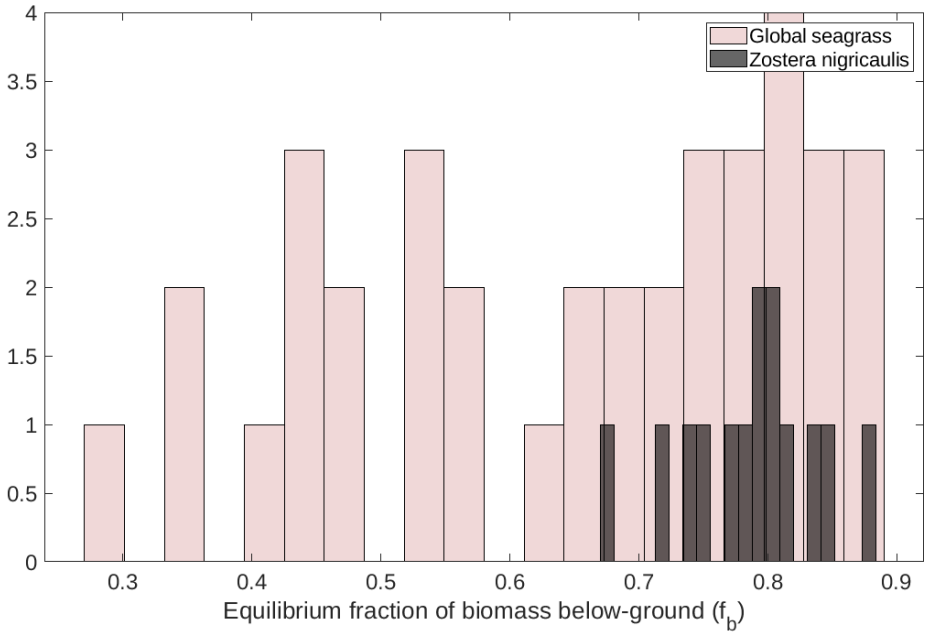

Figure 2.1 : Equilibrium fraction of below-ground biomass (f_b) distribution for global seagrass and only *Zostera nigricaulis* used to set the prior.

the high end of f_b spanning from 0.67 to 0.88 (Fig. 2.1). We fit an informative species-specific prior to the distribution of f_b as a normal distribution with mean 0.7721 and standard deviation 0.0273 (Table 2.3).

Above-ground biomass was plotted against proportion cover across all sites, including replicates and time-points to gauge where the boundaries of the relationship described in Equation 2.9 were. Five realisations of Equation 2.9 with Ω values ranging from 0.001 to 0.1 were laid over the observations (Figure 2.2 (left)) as indicative. A distribution of Ω values (Figure 2.2 (right)) was built by rearranging Equation 2.9 to $\Omega = -\frac{\ln(1-P_c)}{S_A}$ and substituting all P_c and S_A observation pairs to then fit a log normal prior distribution (Table 2.3).

Radiation use efficiency (r_{ue}) as the conversion efficiency of absorbed radiation into dry weight biomass [41] is known to be both species-specific and dependent upon environmental conditions [89]. Only two seagrass studies measuring r_{ue} were

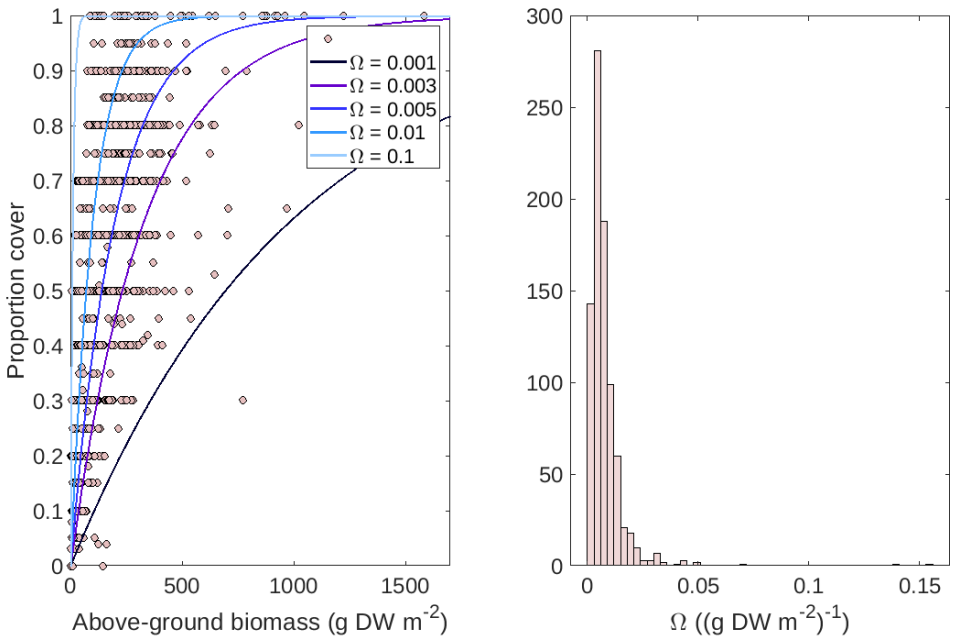

Figure 2.2 : The relationship between above-ground biomass and proportion cover with realisations of Equation 2.9 at Ω values 0.001, 0.003, 0.005, 0.01 and 0.1 (left). Histogram of Ω values calculated from above-ground biomass and proportion cover observations at all Port Phillip Bay sites (right).

found, one for *Posidonia oceanica* comparing effects of different depths and different light regimes [64] and the other *Amphibolis griffithii* under winter and summer light climates to describe energy flows through a canopy [19]. The prior distribution for effective absorbance (A) was based on a combination of bending angle [3] and leaf absorbance [77]. Compensation scalar irradiance (E_{comp}) was based on [20], and the translocation rate was loosely based on [55]. Leaf and root/rhizome mortality rates were essentially not measured in previous studies, so the prior means follow [20] and were assigned large standard deviations to reflect the uncertainty in these values. Water column height (h_{wc}) was not measured while sampling consequently the prior distribution was set based on data made available from an unpublished study conducted by Deakin University (Tim Smith, personal communication).

The seagrass diversity factor was assigned a prior distribution of $\text{Log}\mathcal{N}(\ln(0.2), 0.35)$ to represent the variance scaling of marine community compositions and reflect one dominant species at each site, instead of multiple species. This was illustrated in the equilibrium fraction of biomass below-ground (f_b) prior (Fig. 2.1) where the variance of *Zostera nigricaulis* was approximately 0.2 of the global prior. This was in line with the SDF prior in the NPZD study [76], where in marine community compositions a single species had approximately 20% of the community variance.

The light attenuation coefficient (k_d) prior was estimated from EPA measurements and varied across sites, $\text{Log}\mathcal{N}(\ln(0.2), 0.6)$ at Altona and Blairgowrie, $\text{Log}\mathcal{N}(\ln(0.3), 0.6)$ at Point Henry and Point Richards, and $\text{Log}\mathcal{N}(\ln(0.1), 0.6)$ at Swan Bay South. The following parameters were assigned constant and known, $g_A^{max} = 0.1\,\text{d}^{-1}$, St = 10 d, $h_{sed} = 0.15\,\text{m}$, and $\phi = 0.5$.

Initial conditions for above-ground biomass, below-ground biomass and proportion cover were different at each site and for each year (appendix Table A.1) and were constructed from the observations collected [87]. Initial conditions for translocation

Parameter	**Prior**	**Source**	**Species**
μ_τ	Log$\mathcal{N}$(ln(0.001), 0.95)	[55]	*Zostera marina*
μ_{ζ_A}	Log$\mathcal{N}$(ln(0.005), 1)	[20]	*Zostera capricorni*
μ_{ζ_B}	Log$\mathcal{N}$(ln(0.0001, 1)	[20]	*Zostera capricorni*
$\mu_{\mathrm{E}_{comp}}$	Log$\mathcal{N}$(ln(3), 0.5)	[20]	*Zostera capricorni*
μ_{A}	Log$\mathcal{N}$(ln(0.15), 0.5)	[3, 77]	*Zostera muelleri*
μ_{h_c}	Log$\mathcal{N}$(ln(0.4), 0.5)	[87]	*Zostera nigricaulis*
μ_{k_d}	Log$\mathcal{N}$(ln(0.1-0.3), 0.6)	EPA data	Water quality
Ω	Log$\mathcal{N}$(ln(0.005), 1)	[87]	*Zostera nigricaulis*
h_{wc}	Log$\mathcal{N}$(ln(0.5), 0.2)	(Tim Smith, personal communication)	Water depth
f_b	$\mathcal{N}$(0.7721, 0.0273)	[33, 48, 87]	*Zostera nigricaulis*
SDF	Log$\mathcal{N}$(ln(0.2), 0.35)		
r_{ue}	Log$\mathcal{N}$(ln(0.4), 0.5)	[19, 64]	*Amphibolis griffithii*, *Posidonia oceanica*

Table 2.3 : Summary of the prior distributions for parameters and the source of their values.

rate (τ), leaf mortality rate (ζ_A), root/rhizome mortality rate (ζ_B), total compensation scalar irradiance (E_{comp}), effective absorbance (A), canopy height (h_c), were the same across site and year and were sampled from the prior distributions assigned in Table 2.3. Initial conditions for k_d varied by site, but were consistently sampled from the prior distributions assigned.

2.2.4 Data Model: Port Phillip Bay data description and collection methods

All seagrass observations were obtained through collaboration with Tim Smith, Deakin University [87], water quality observations in correspondence with Environmental Protection Agency (EPA) and light measurements from the Bureau of Meteorology (BOM) website.

Study site

Covering an area of approximately 1,950 km^2 with a coastline spanning 325 km in length, Poty Phillip Bay is densely populated with more than 4 million people living around its coastline. Most of the population resides in Melbourne to the north of the bay which also sees freshwater inputs from the Yarra River. The western part of the bay is home to the Western Treatment Plant (WTP) that discharges treated sewage causing influxes of nitrogen and toxins into those parts of the bay. Approximately half of Port Phillip Bay is less than 8 m deep, with parts of the bay reaching a maximum depth of 24 m.

Seagrass biomass, percent cover, canopy height

Five sites in Port Phillip Bay were sampled monthly in the flowering season and every three months between from October 2011 to January 2014 [87]. At each site, three replicate quadrats were randomly thrown within three permanent 10 m $\times$ 10 m plots, that were 50-100 m apart and sampled. Plots were located at the

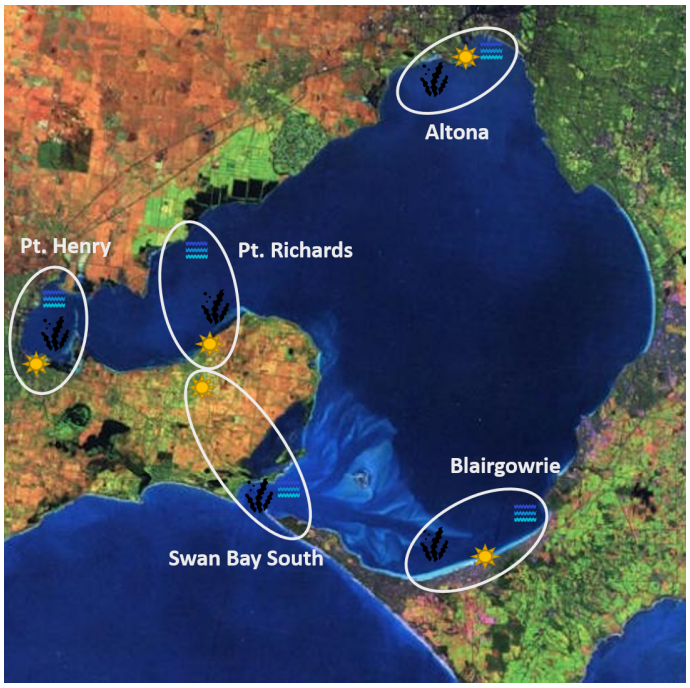

Figure 2.3 : Map of five seagrass sampling sites (seagrass), nearest BOM solar stations (sun), and EPA water quality sites (blue lines) in Port Phillip Bay, Victoria. Aerial image sourced from the Bureau of Meteorology.

interior of patches, and gaps in seagrass cover avoided to ensure consistency within sites. Samples consisted of a 0.25 m × 0.25 m quadrat where canopy height (the highest point of the canopy in m) was measured and all seagrass cut at the sediment surface and collected. Above-ground samples were frozen prior to being processed in the laboratory for biomass, measured in grams of dry weight. Additionally, percent cover (the percentage of the seafloor covered by seagrass when viewed from above) measurements were collected on 0.5 m × 0.5 m quadrats overlapping those 0.25 m × 0.25 m quadrats used previously. Samples from Point Henry in January 2013 were not collected due to poor weather. The five sites and their locations were Altona (37°87 S, 144°84 E), Blairgowrie (38°36 S, 144°79 E), Point Henry (38°14 S, 144°42 E), Point Richards (38°11 S, 144°63 E) and Swan Bay South (38°27 S, 144°65 E).

To visualise the central tendency and spread of above-ground biomass, proportion cover and canopy height observations through time and across site, box-plots for each time-point were plotted (Fig. 2.4). Above-ground biomass observations on average follow a seasonal trend with peaks in the October to January period. The variability appears consistently low, except for October to January 2011 at Blairgowrie where there is twice as much variability in October, and January, three times as much in December and four times as much in November than is captured otherwise. With the exception of October to January 2011 at Blairgowrie, the data-points are roughly symmetric and the mean and median lie at similar values. Proportion cover observations display high variability because while the absolute range for proportion cover is 0 (no seagrass) to 1 (the whole sea-floor is covered by seagrass), there are times when almost the entire range is covered; December 2011, January and April 2012 at Altona, January to September 2012 at Point Richards, January July 2012 at Swan Bay South. Means are deviating from the medians and there are skews in both the positive and negative direction. Proportion cover observations are known to be highly variable because of the subjective nature of their data col-

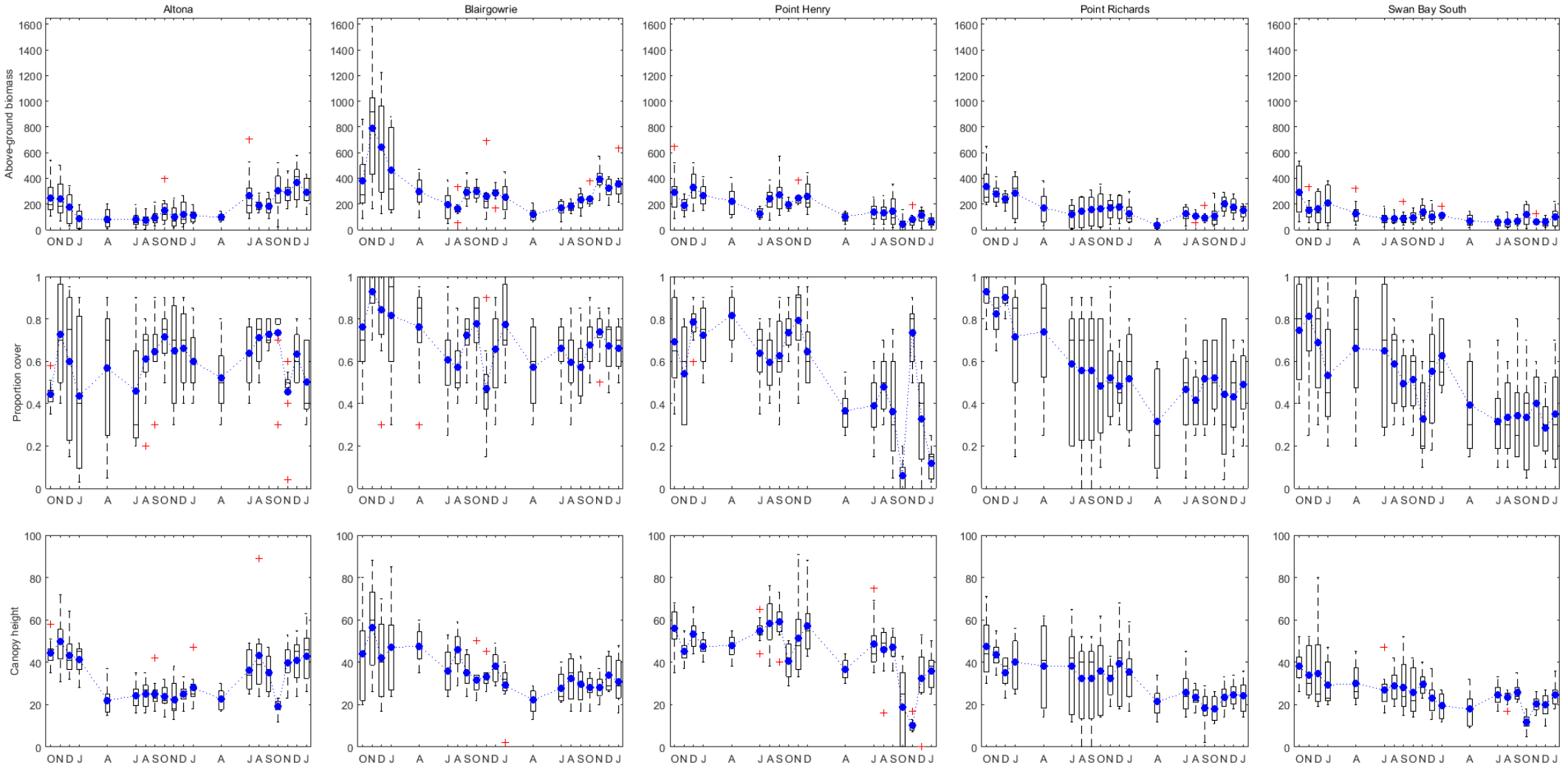

Figure 2.4 : Box-plots of above-ground biomass (g DW m^{-2}), proportion cover, and canopy height (cm) observations for each site through time (n=9 for each time point). The mean is indicated by a blue circle, and potential outliers by red crosses. Time is representative of months, starting from October 2011 (O) ranging to January 2014 (J).

lection. There is a benchmark guide for recording proportion cover, but this varies between species, density and data collector. Means across plots for each site at each time-point were used as the observations assimilated (blue in Figure 2.4).

Solar data

Daily surface irradiance was determined from daily global solar exposure (GSE) measurements made between October 2011 and January 2014. These observations were obtained from the Australian Government Bureau of Meterology (BOM) website (www.bom.gov.au/climate/data/) and are defined as the total solar energy for a day falling on a horizontal surface, measured from midnight to midnight. To obtain a representative GSE measurement, the BOM stations closest to the seagrass biomass data sites (Fig. 2.3) were identified for each site and assumed to represent the daily surface light dose at each seagrass sampling site. The following BOM stations were used: Altona (City Offices) (station number: 87131, 37°86 S, 144°83 E) for Altona, South Channel Island (station number: 86344, 38°31 S, 144°80 E) for Blairgowrie, Breakwater (Geelong Racecourse) (station number: 87184, 38°17 S, 144°38 E) for Point Henry, Portarlington VIC (station number: 87053, 38°12 S, 144°64 E) for Point Richards, and Drysdale (Brimdale) (station number: 87114, 38°17 S, 144°59 E) for Swan Bay South. The BOM GSE measurements (MJ m^{-2} d^{-1}) were converted to surface dose of photosynthetically active radiation, E_{wc} (mol photon m^{-2} d^{-1}), using the conversion factor 0.5639 MJ mol^{-1} [45]. The photosynthetically active radiation at the water's surface is similar across sites, ranging from 1 to 20 (mol photon m^{-2} d^{-1}) with similar seasonal magnitudes (Fig. 2.5). Photosynthetically active radiation is used as the model forcing.

Water quality data

The Environmental Protection Agency (EPA) Victoria monitors the water quality of a number of fixed sites in Port Phillip Bay. They have been monitoring

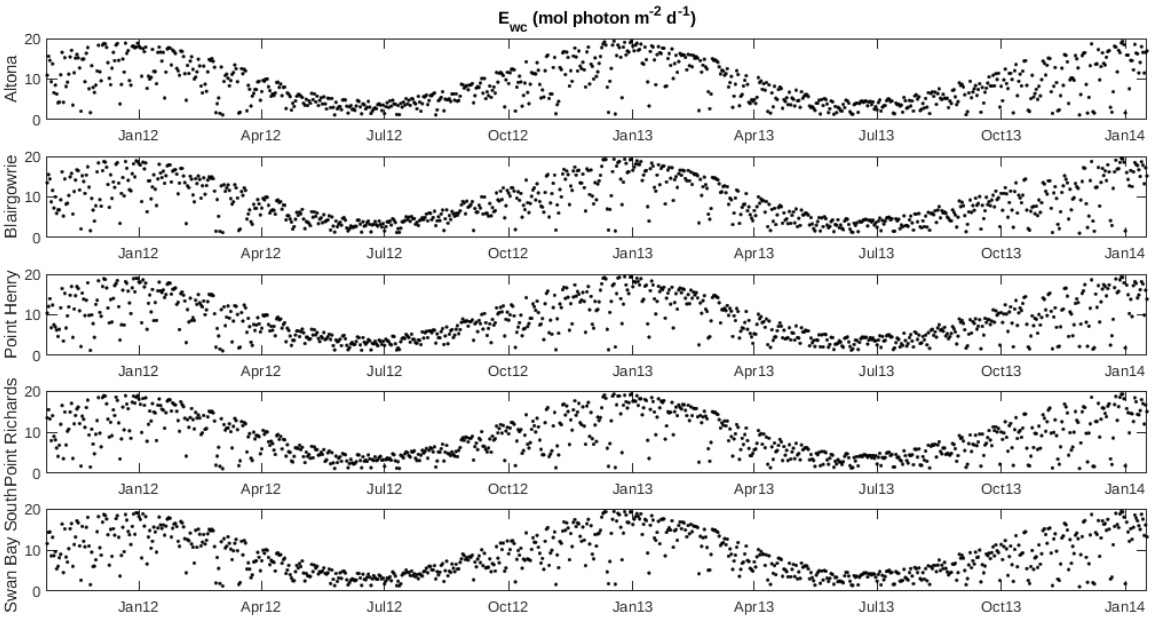

Figure 2.5 : Photosynthetically active radiation at water's surface, E_{wc} (mol photon m^{-2} d^{-1}), of the 5 BOM sites representative of the seagrass sampling sites.

these sites since 1984 to understand the long-term trends of environmental conditions in Victoria's marine environments. Monthly samples of nutrients (nitrogen, phosphorus and silicate), water clarity (total suspended solids and turbidity), dissolved oxygen, ammonia, salinity, chlorophyll-a, heavy metals, water temperature and photosynthetically active radiation are collected and analysed. The EPA monitoring site closest to each seagrass biomass data site (Figure 2.3) was identified for measurements of light attenuation. The water quality sites closest to the seagrass sampling sites were Hobsons Bay (site number: 1991, 37°8702 S, 144°9338 E) to Altona, Dromana (site number: 1282, 38°3034 S, 144°9913 E) to Blairgowrie, Corio Bay (site number: 1911, 38°1010 S, 144°3987 E) to Point Henry, Long Reef (site number: 369, 38°0293 S, 144°5928 E) to Point Richards, and Popes Eye (site number: 2100, 38°2750 S, 144°6989 E) to Swan Bay South. There was not enough EPA nutrient data to use for assimilation, as each of these sites only contained 1-3

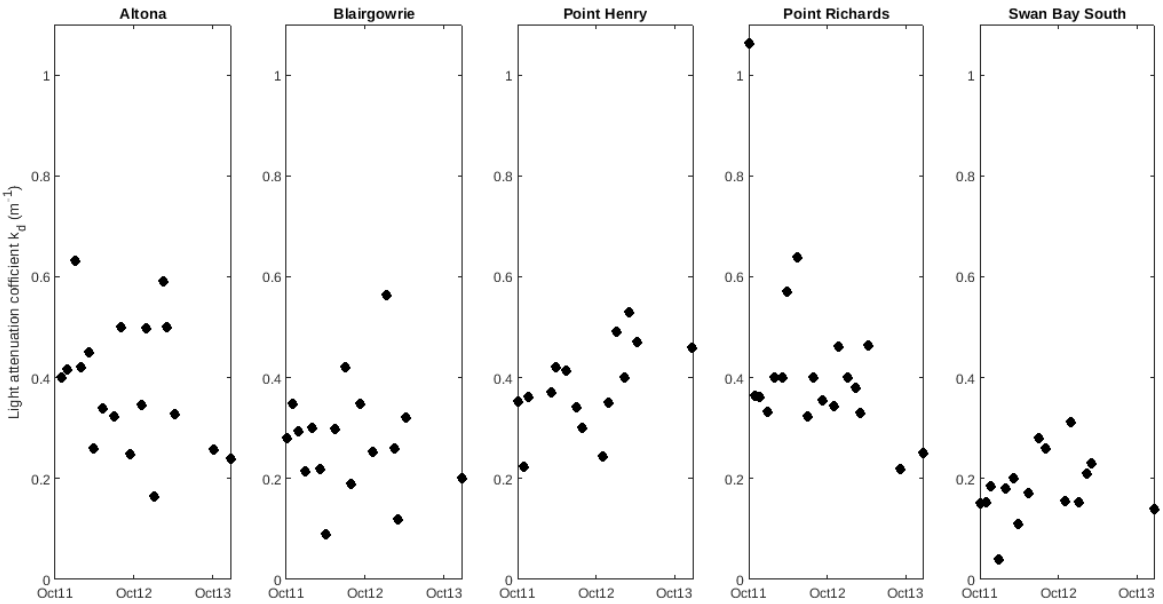

Figure 2.6 : Light attenuation coefficient k_d (m^{-1}) calculated from a surface bottle sample.

data-points across 2012-2013.

Data model

In the data model, above-ground biomass and proportion cover observations were set as log normal distributions of their states with 30% multiplicative error, $S_A$obs $\sim$ Log$\mathcal{N}$(ln(S_A),0.3), $P_C$obs $\sim$ Log$\mathcal{N}$(ln(P_c),0.3). While canopy height and light attenuation coefficient were set as log normal distributions of their states with 50% multiplicative error, $k_d$obs $\sim$ Log$\mathcal{N}$(ln(k_d),0.5), $h_c$obs $\sim$ Log$\mathcal{N}$(ln(h_c),0.5). The light attenuation coefficient was assigned a larger error because the EPA sampling sites were not the same as the seagrass sites, but close by (Figure 2.3) and there was a sample size of 1 at each time-point. Canopy height had a large observation error because it was specified as an autoregressive and therefore would need to fluctuate around a mean over time instead of following increases or decreases in observations.

Table 2.4 : PMMH acceptance rates for 2012, 2013, and 2012-2013 posteriors.

Site	**2012**	**2013**	**2012-2013**
Altona	20.48%	25.80%	11.59%**
Blairgowrie	13.64%	20.09%	9.51%
Point Henry	25.92%	28.48%	15.29%**
Point Richards	18.50%	24.18%	18.10%
Swan Bay South	20.80%	29.35%	12.67%

2.3 Results

The PMMH algorithm was run using LibBi [72] for all five sites, executed one site at a time. Firstly it was used to sample the prior and posterior for 2012, then an ensemble prediction was generated for 2013 by doing a free-run of the dynamical model. The PMMH algorithm was then used to sample from the prior and posterior for 2013 for comparison against the 2013 prediction. Lastly, the algorithm was run to sample from the prior and posterior across 2012-2013. Runge-Kutta 4(3) was the ODE solver used (to solve equations 2.17-2.26) and PMMH proposal distributions were assigned according to each parameter's prior distribution (Table 2.3) with proposal standard deviation calculated as 10% of the prior standard deviation. After tuning the PMMH algorithm for convergence, each MCMC chain contained 50,000 samples with the initial 10,000 discarded as burn-in, and each set of posterior results achieved sufficient MCMC mixing (Figure A.3, Figure A.5, Figure A.7) and likelihood convergence (Figure A.2, Figure A.4, Figure A.6) using 1,024 particles with the exception of 2012-2013 Altona and Point Henry that saw larger variance in the likelihood and consequently the number of particles was increased to 4,096. Acceptance rates ranged from 9.51% to 29.35% (Table 2.4), which are in agreement with acceptance rates for models with equivalent numbers of parameters.

This section illustrates the results of the state variables above-ground biomass (S_A), below-ground biomass (S_B), proportion cover (P_c) and the first order autoregressive processes (translocation rate (τ), leaf mortality rate (ζ_A), root/rhizome mortality rate (ζ_B), total compensation scalar irradiance (E_{comp}), effective absorbance (A), canopy height (h_c), and light attenuation coefficient (k_d)). Prior and posterior states are illustrated through heat-maps to identify areas of high probability density before and after data assimilation. Separately, observations are displayed as box-plots and posteriors and predictions are plotted as medians and 95% credible intervals to show the values of highest probability density and spread through time when displayed with the data.

The parameters Ω, f_b, h_{wc}, r_{ue}, SDF, and all the long-term means of the autoregressive processes (translocation rate (μ_τ), leaf mortality rate (μ_{ζ_A}), root/rhizome mortality rate (μ_{ζ_B}), total compensation scalar irradiance ($\mu_{E_{comp}}$), effective absorbance (μ_A), canopy height (μ_{h_c}), and light attenuation coefficient (μ_{k_d})) are plotted either as histograms or kernel density estimates (kde). Parameter priors and posteriors are plotted as overlaid histograms to visualise what updates the observations had on the priors. Kernel density estimates are used to compare posteriors across sites and to discover where the differences in parameters lie between the 2013 predicted and posterior results. Due to the skewed nature of our parameter priors and posteriors, the median is chosen as the measure of central tendency, and the interquartile range (IQR) as the measure of spread.

2.3.1 Prior to posterior results for 2012

Above-ground biomass fraction of density heat-maps illustrated tightening of posterior density after assimilating observations at each site. There was 10 times the posterior density than the sampled prior at those points (Figure 2.7, Figure 2.8, Figure 2.9, Figure 2.10, Figure 2.11). Altona, Blairgowrie, Point Richards and

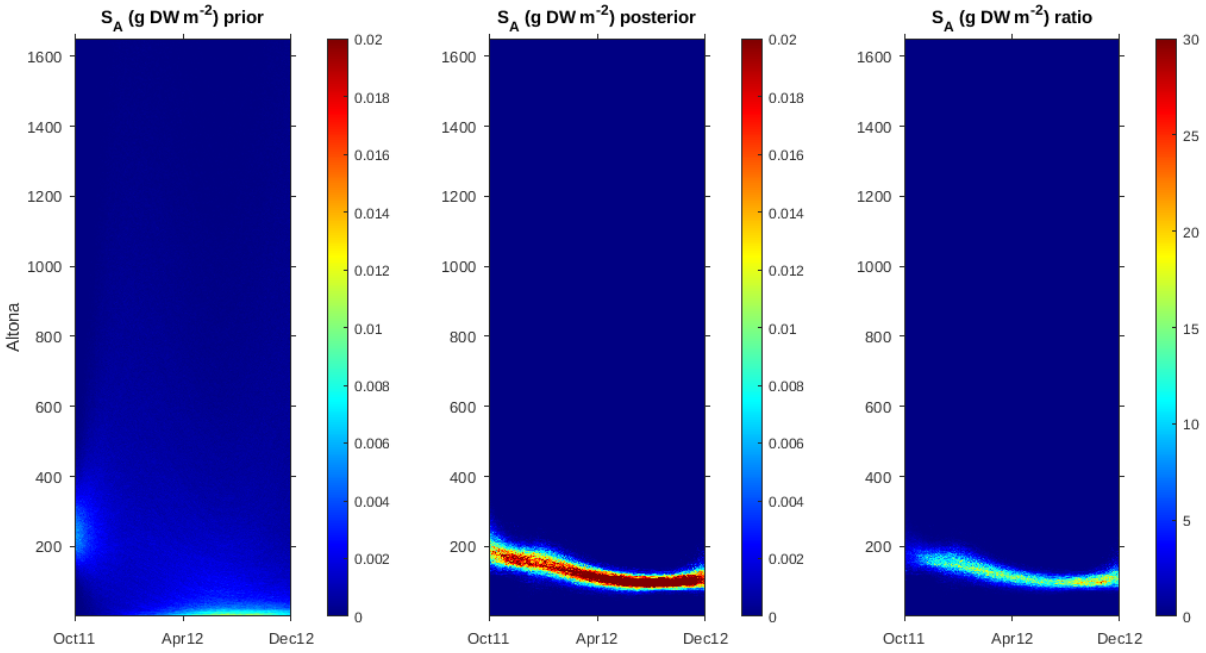

Figure 2.7 : Above-ground biomass (g DW m^{-2}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Altona in 2012.

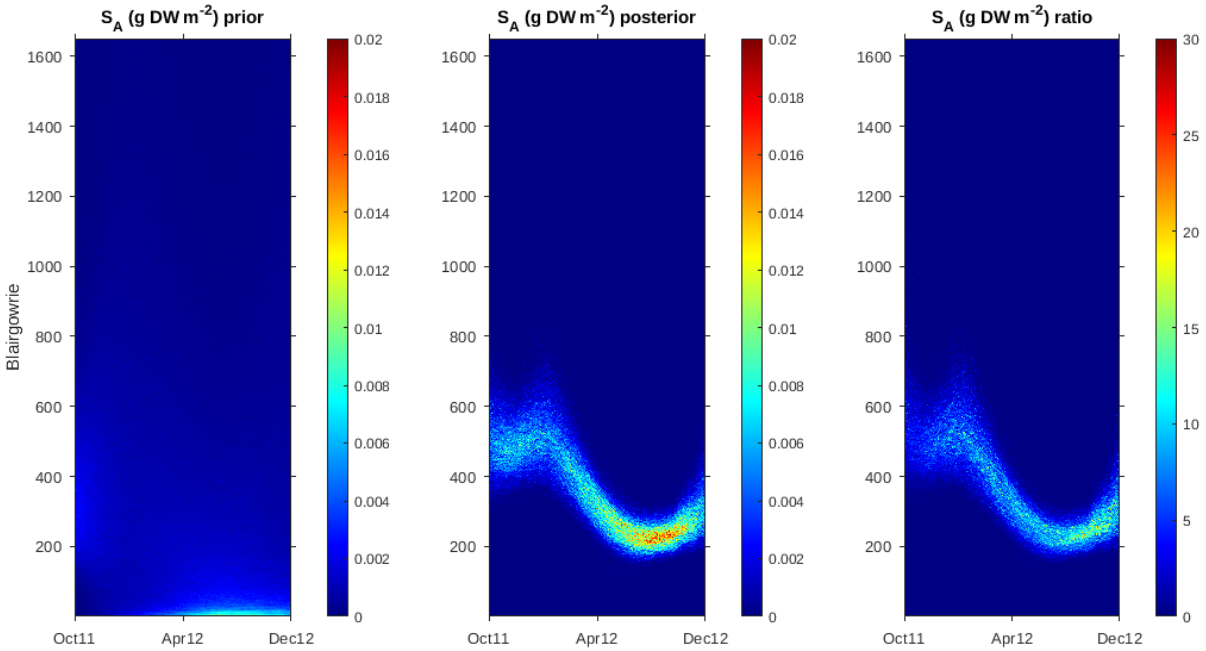

Figure 2.8 : Above-ground biomass (g DW m^{-2}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Blairgowrie in 2012.

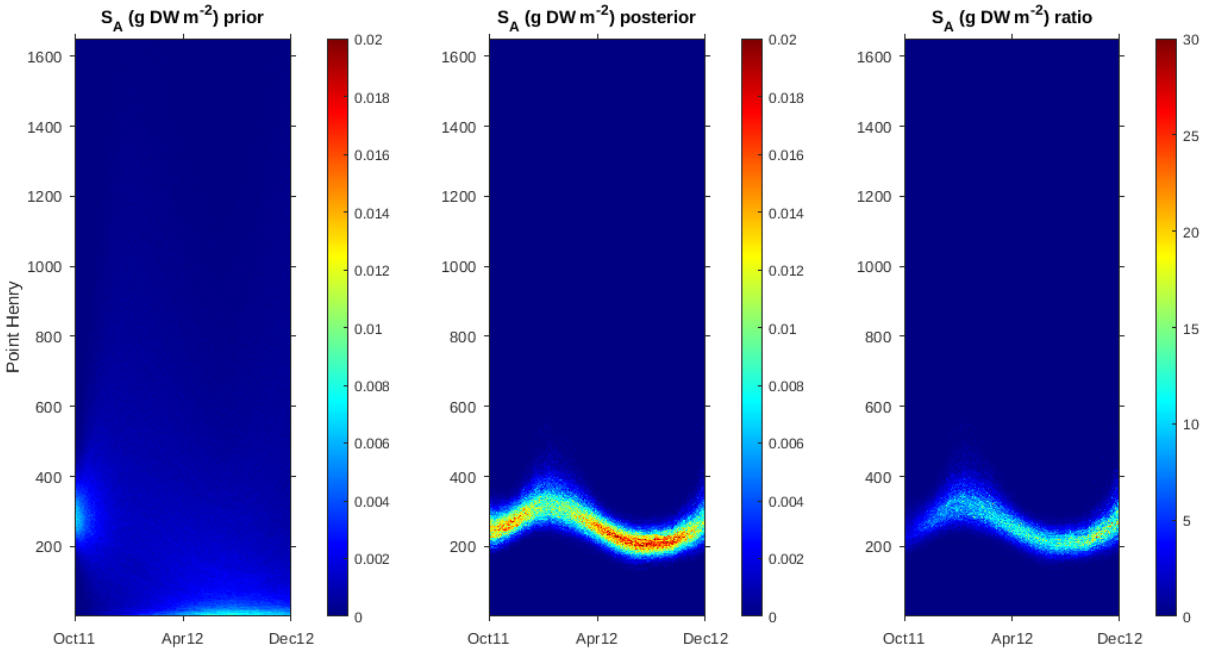

Figure 2.9 : Above-ground biomass (g DW m^{-2}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Point Henry in 2012.

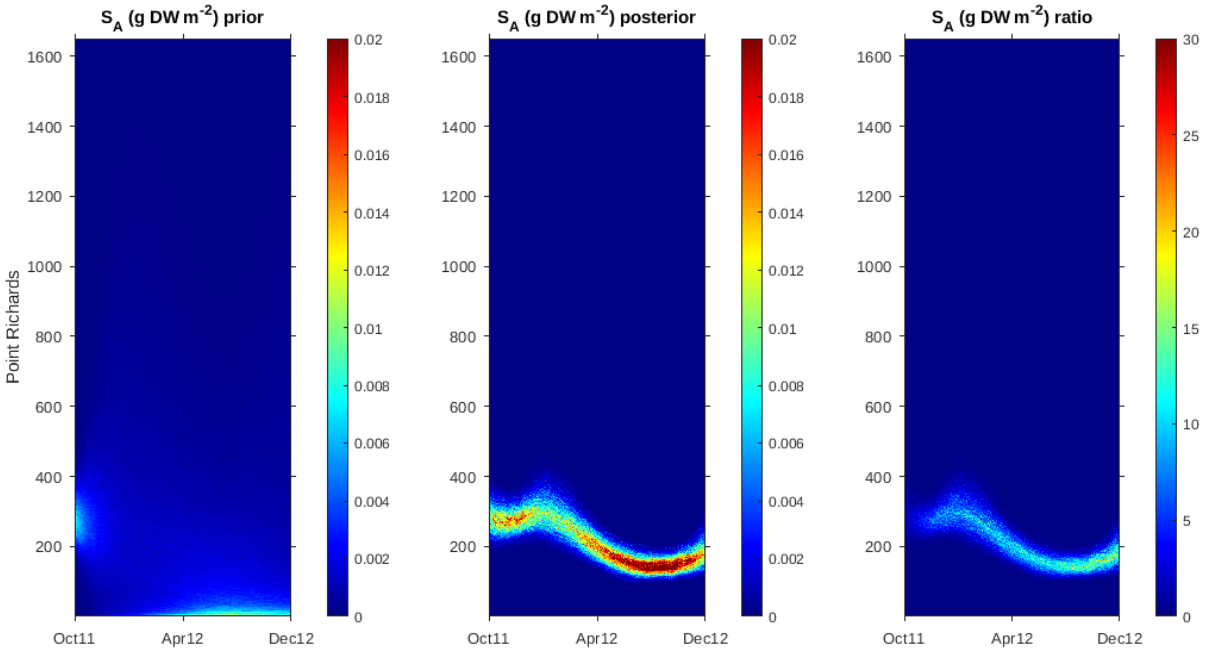

Figure 2.10 : Above-ground biomass (g DW m^{-2}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Point Richards in 2012.

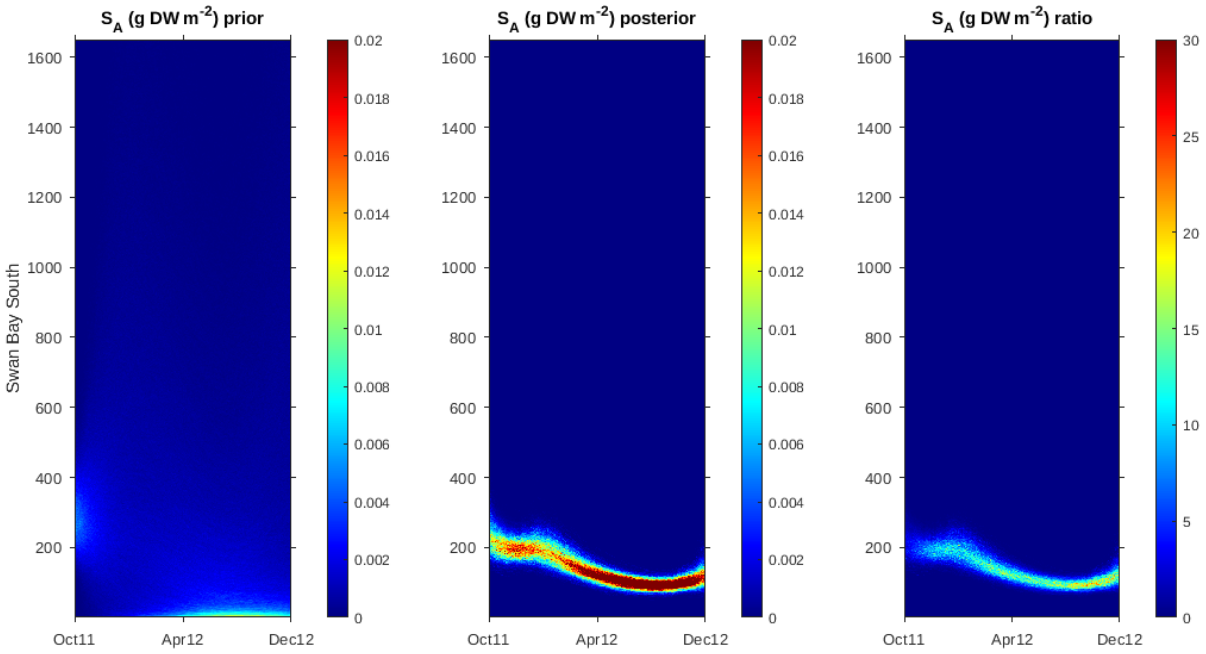

Figure 2.11 : Above-ground biomass (g DW m^{-2}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Swan Bay South in 2012.

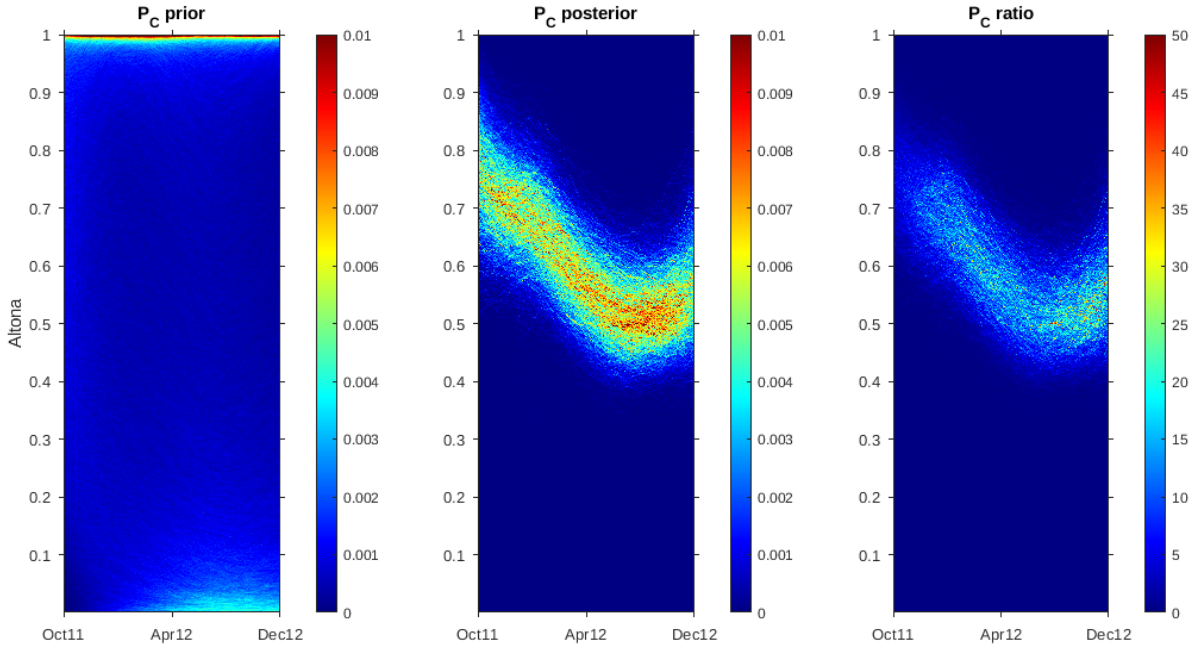

Figure 2.12 : Proportion cover fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Altona in 2012.

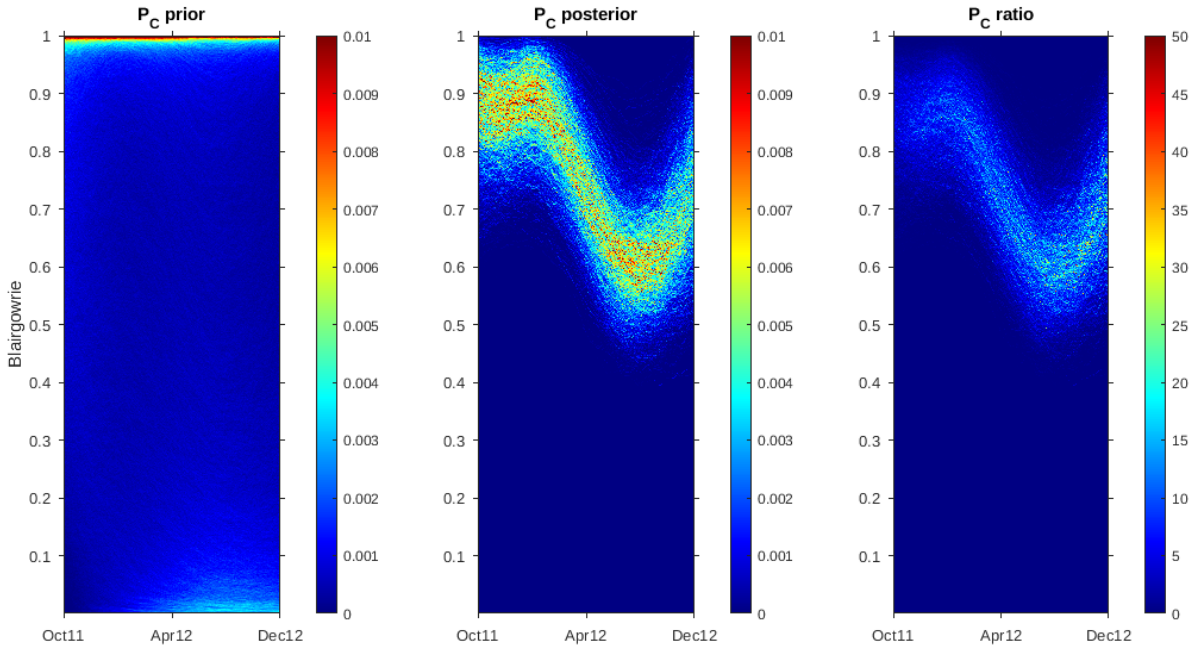

Figure 2.13 : Proportion cover fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Blairgowrie in 2012.

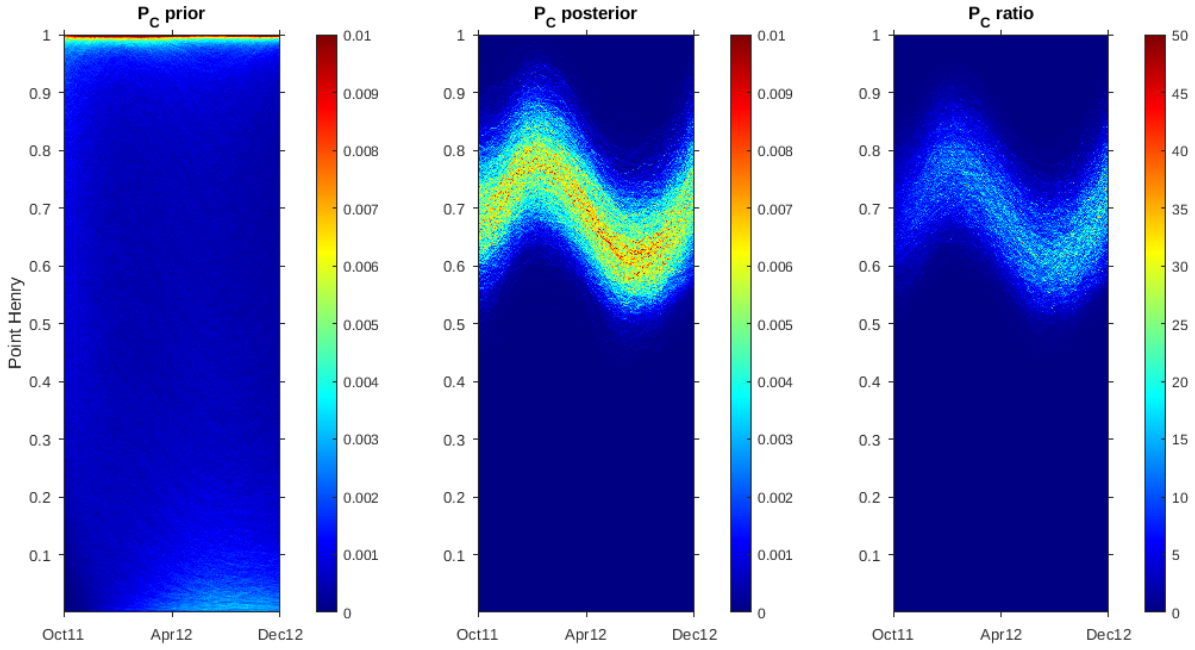

Figure 2.14 : Proportion cover fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Point Henry in 2012.

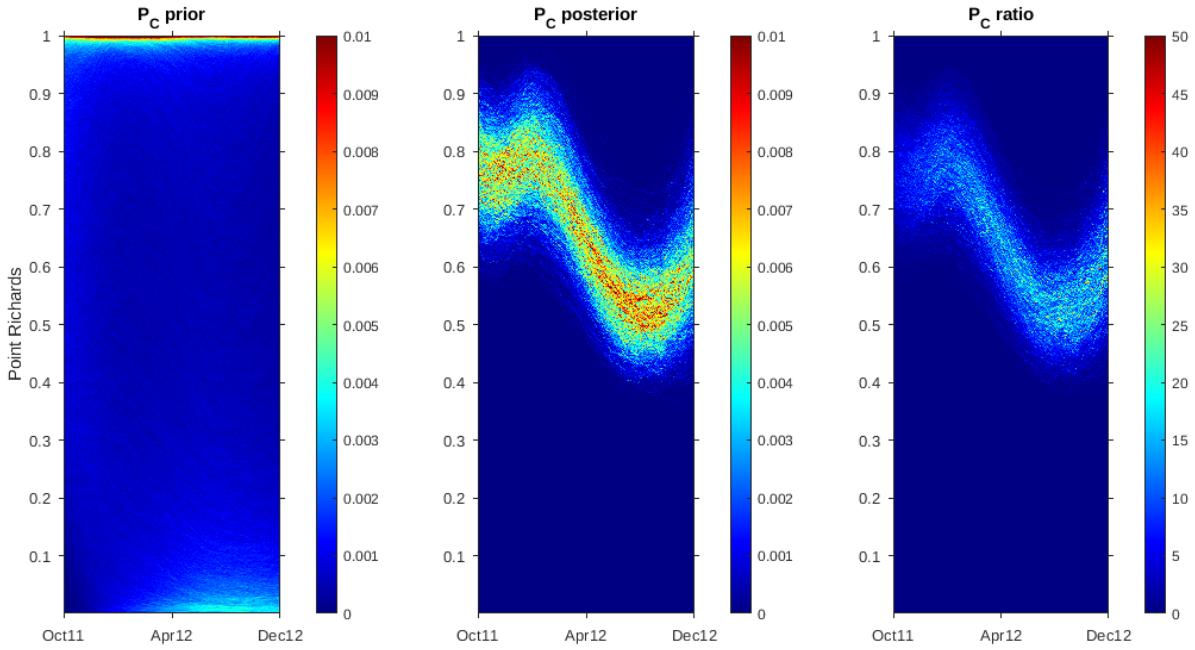

Figure 2.15 : Proportion cover fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Point Richards in 2012.

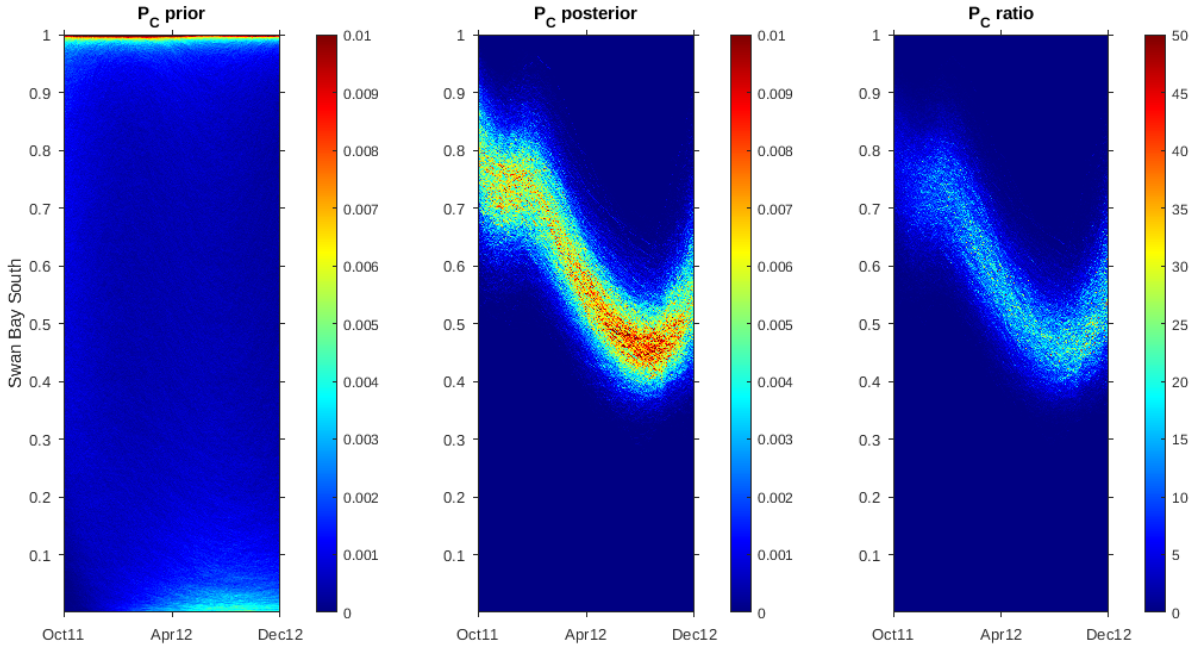

Figure 2.16 : Proportion cover fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Swan Bay South in 2012.

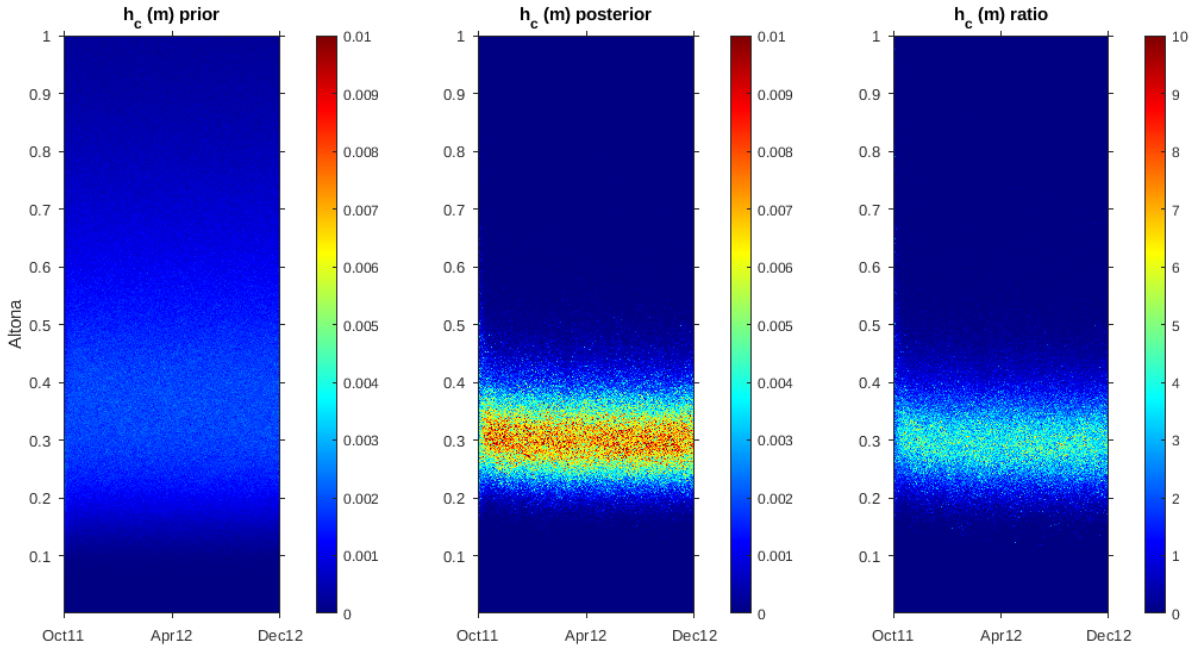

Figure 2.17 : Canopy height (m) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Altona in 2012.

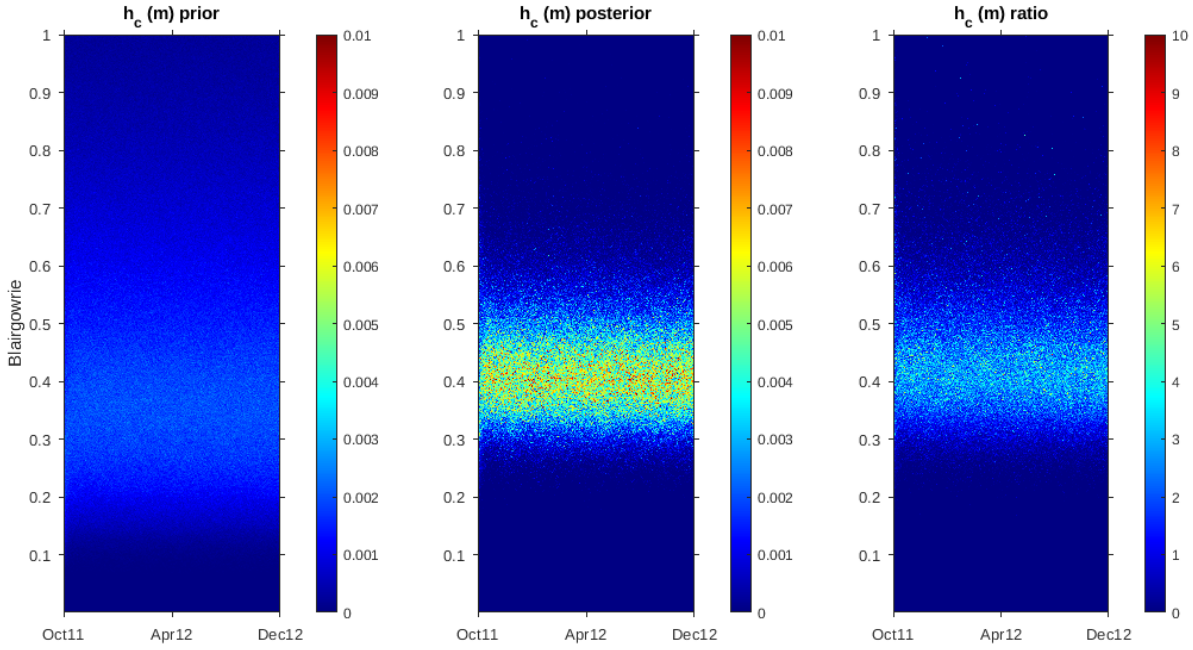

Figure 2.18 : Canopy height (m) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Blairgowrie in 2012.

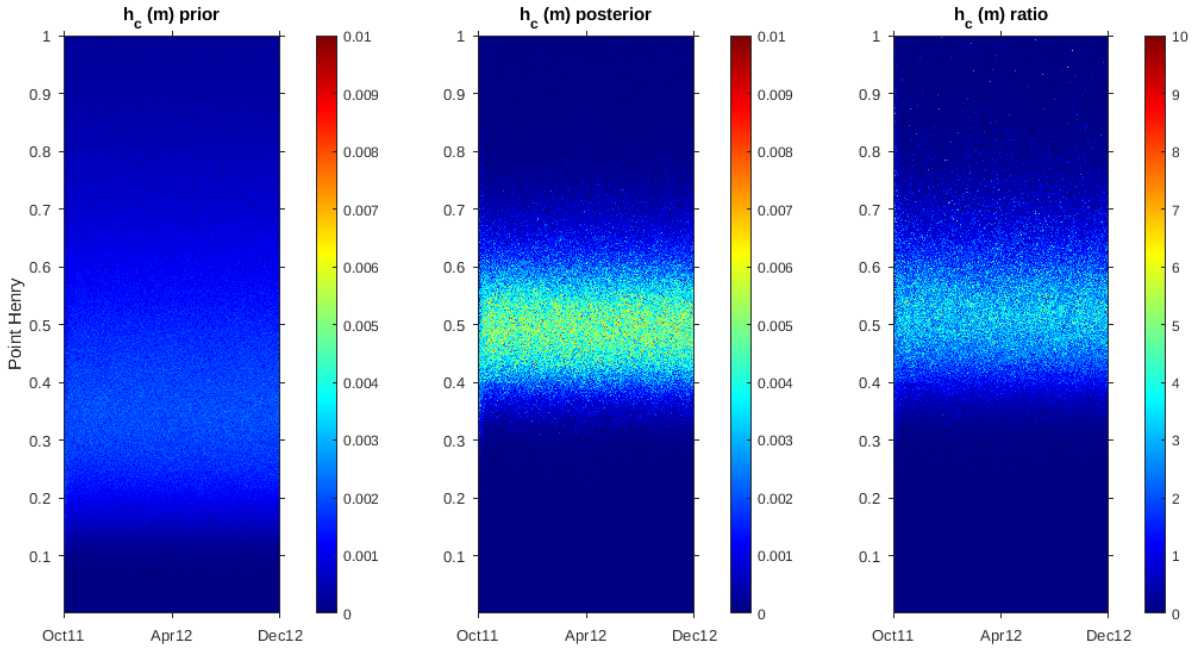

Figure 2.19 : Canopy height (m) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Point Henry in 2012.

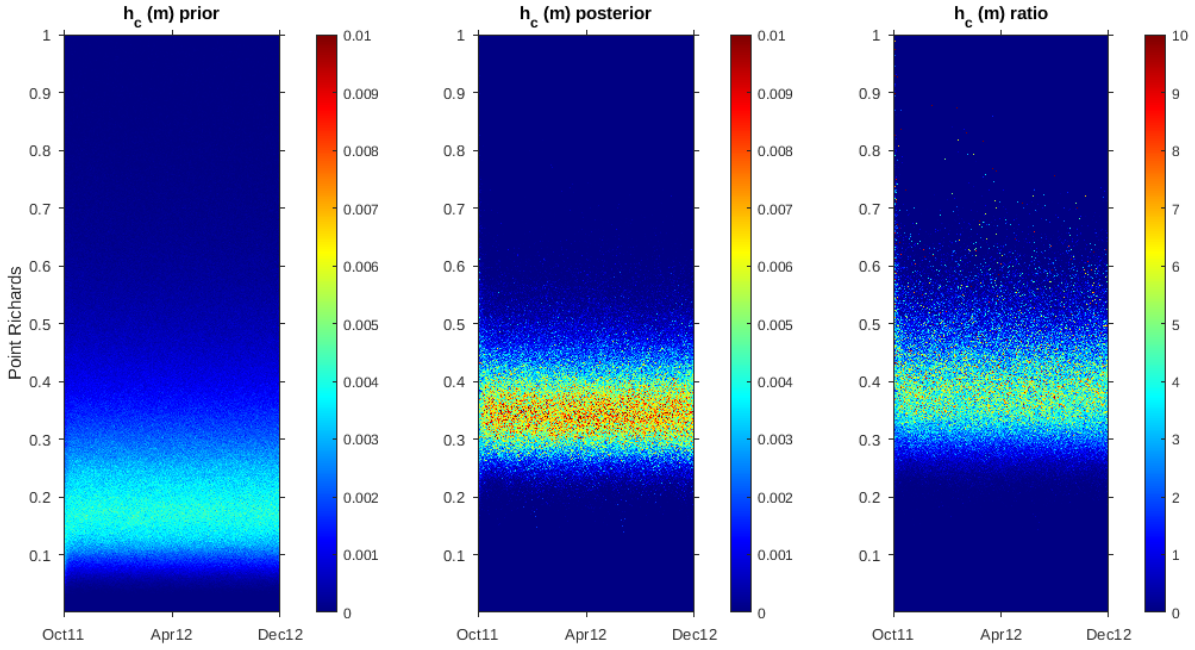

Figure 2.20 : Canopy height (m) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Point Richards in 2012.

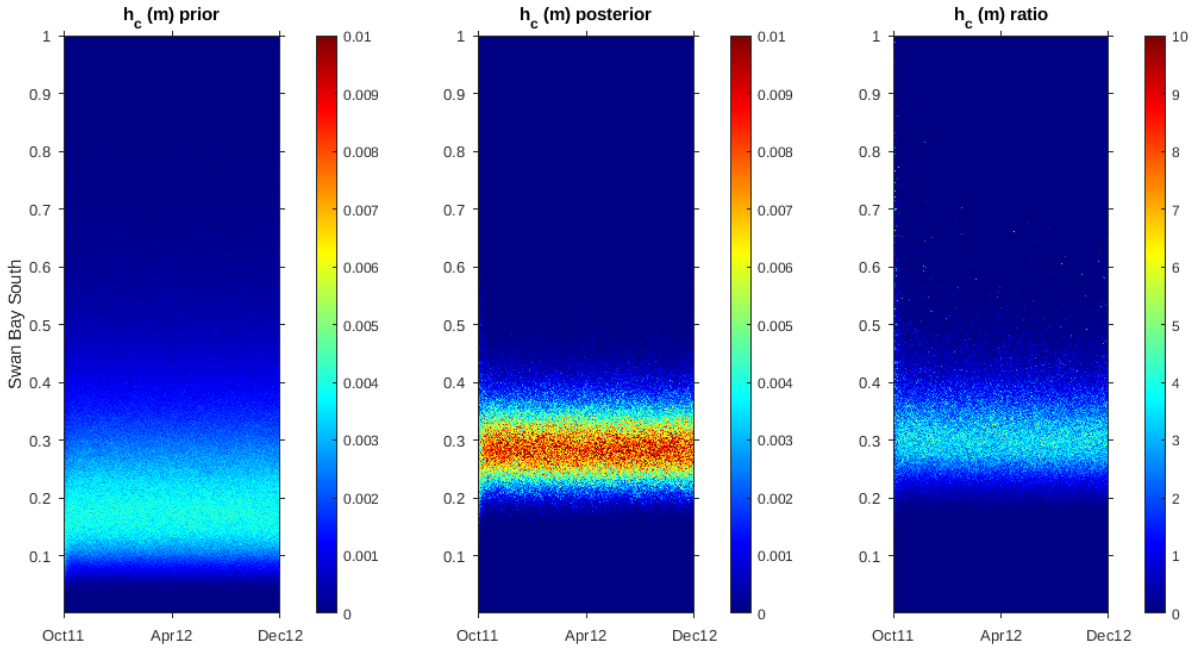

Figure 2.21 : Canopy height (m) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Swan Bay South in 2012.

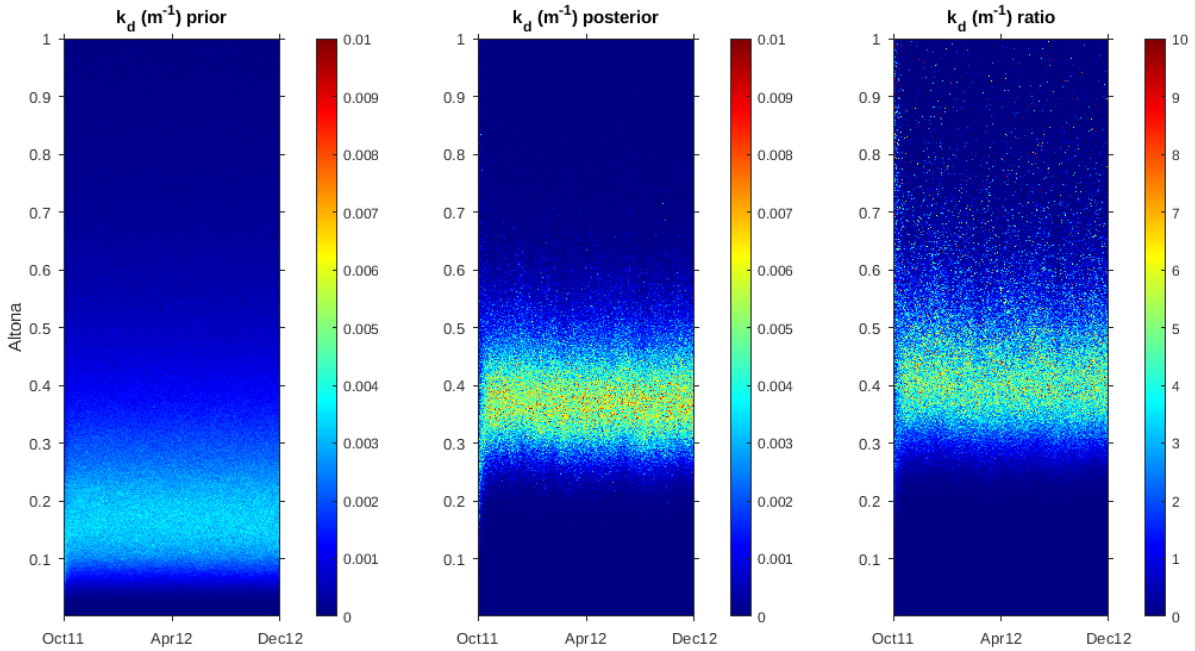

Figure 2.22 : Light attenuation coefficient (m^{-1}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Altona in 2012.

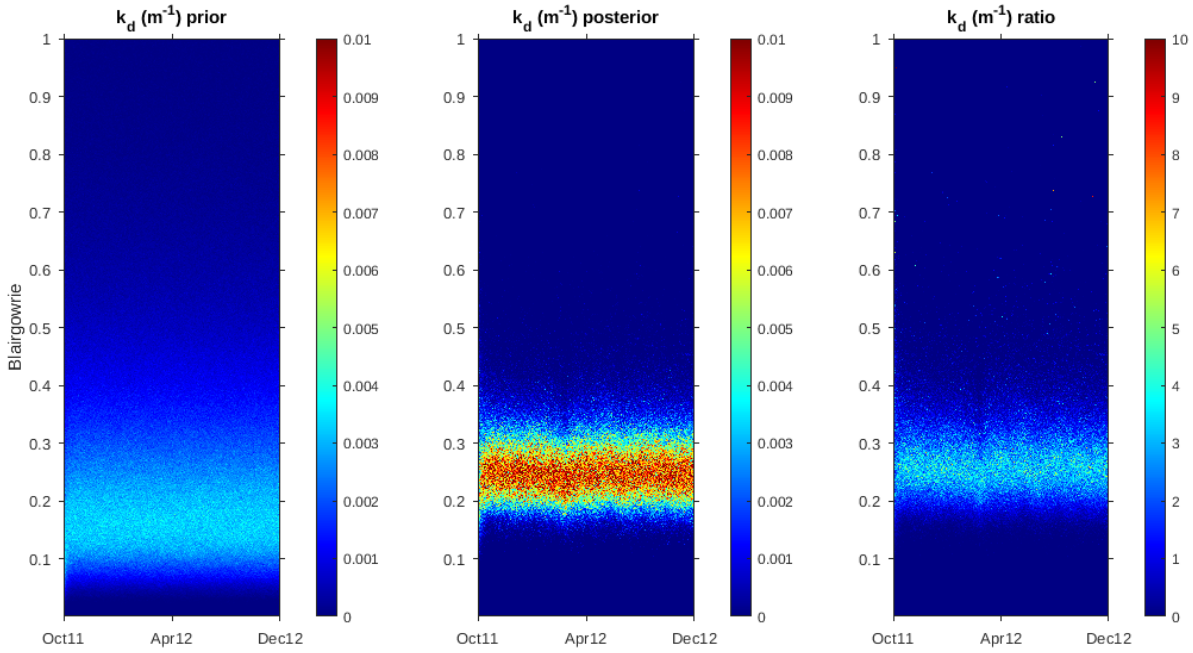

Figure 2.23 : Light attenuation coefficient (m^{-1}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Blairgowrie in 2012.

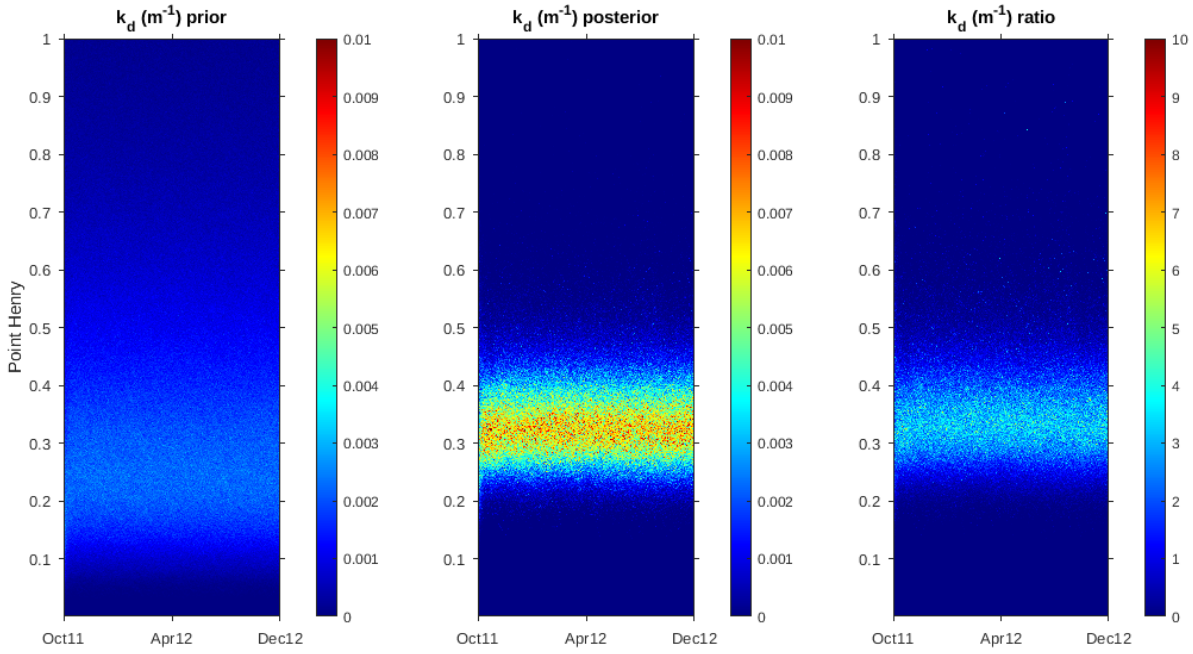

Figure 2.24 : Light attenuation coefficient (m^{-1}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Point Henry in 2012.

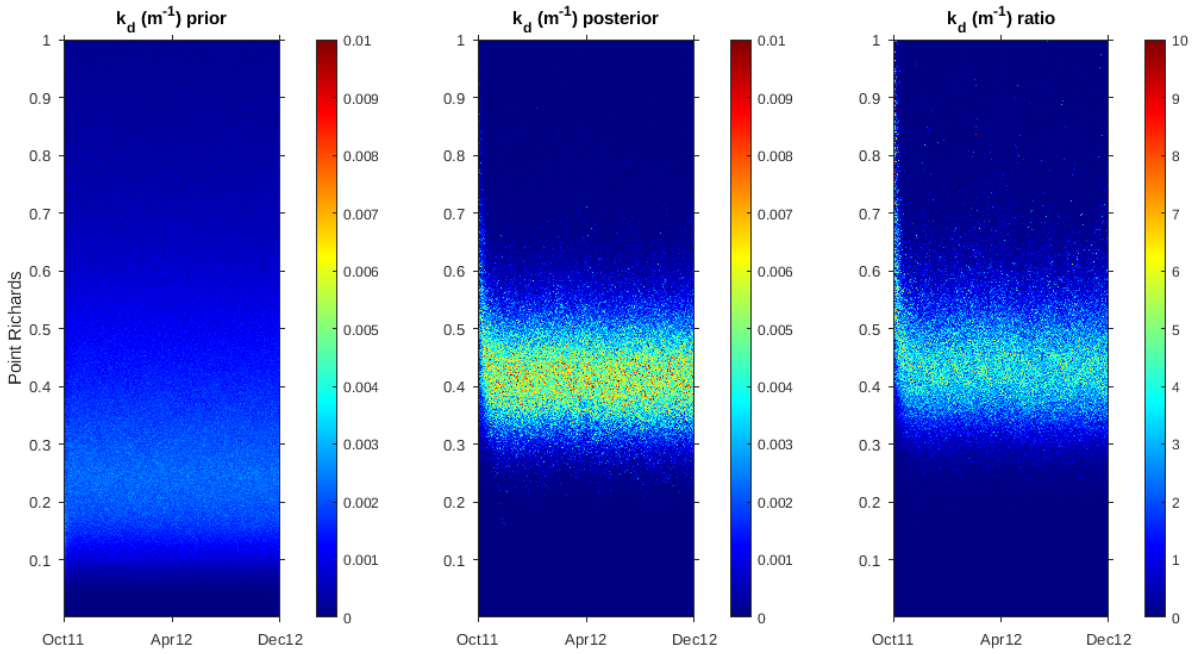

Figure 2.25 : Light attenuation coefficient (m^{-1}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Point Richards in 2012.

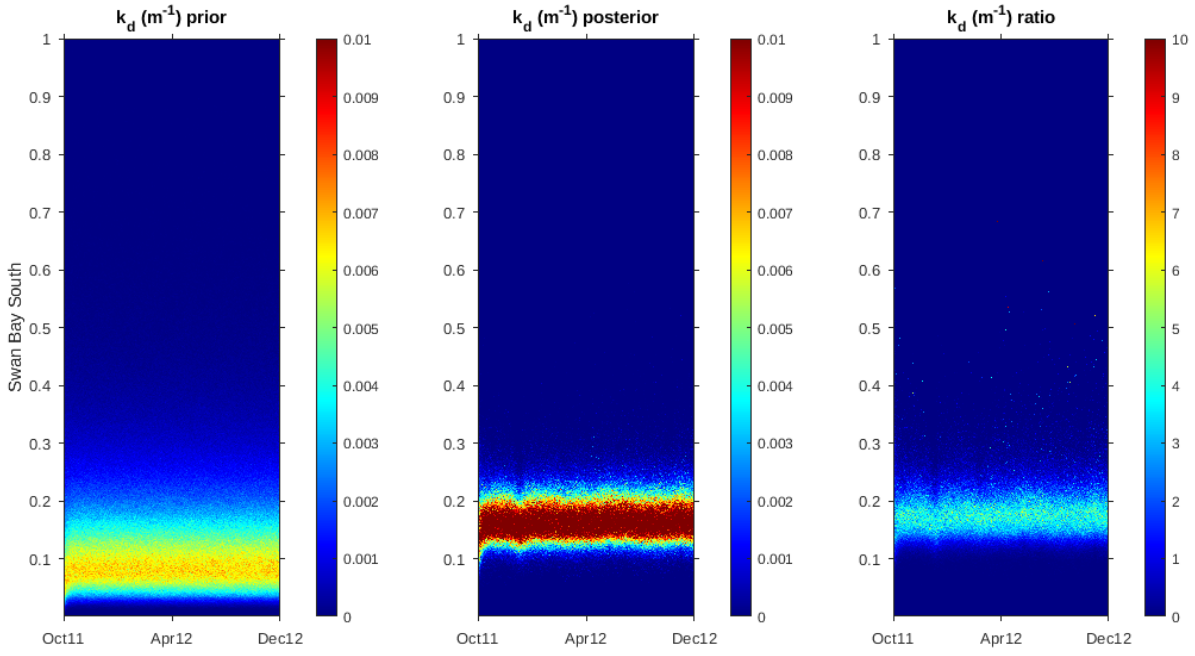

Figure 2.26 : Light attenuation coefficient (m^{-1}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Swan Bay South in 2012.

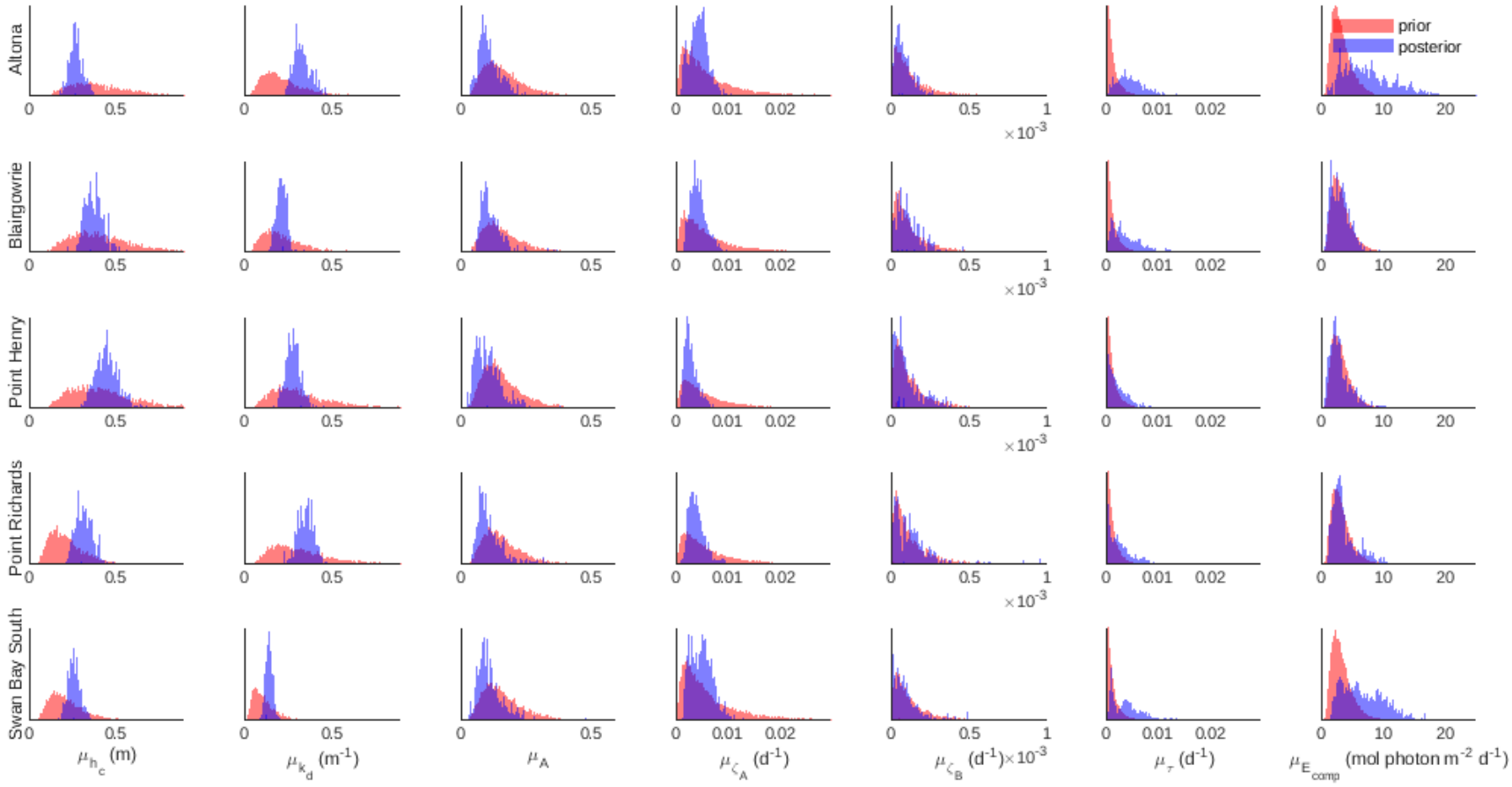

Figure 2.27 : Long-term means of the first order autoregressive parameter prior and posterior distributions for each Port Phillip Bay site in 2012.

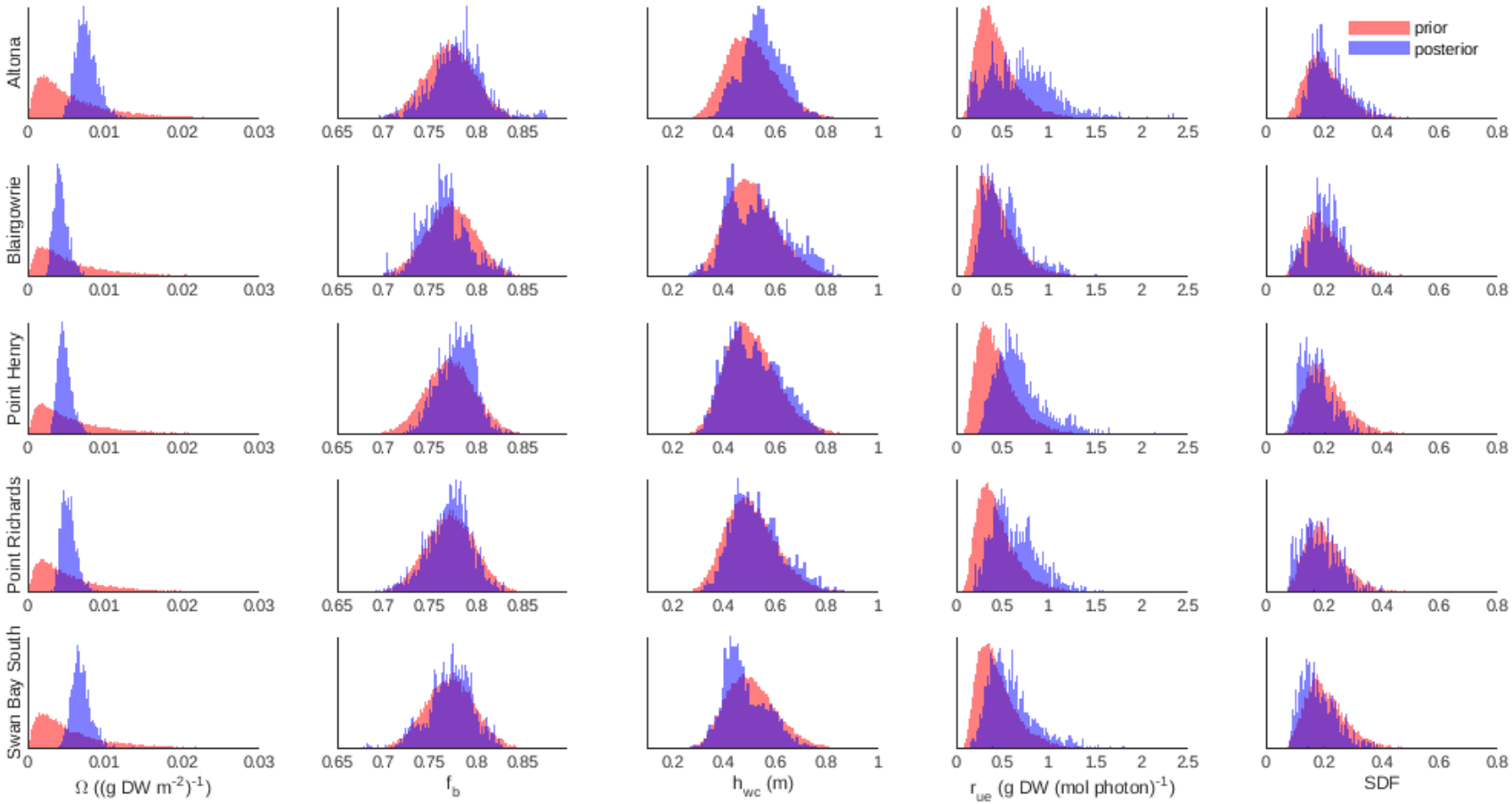

Figure 2.28 : Parameter prior and posterior distributions for each Port Phillip Bay site in 2012.

Swan Bay South posteriors all trend downwards over 2012. Blairgowrie had the largest initial posterior biomass, and largest spread of posterior density of all the sites during October 2011-February 2012 (Figure 2.8). Point Henry was the only site not trending downwards over 2012, with a seasonal trend seeing the biomass return to the same starting values of October 2011 (Figure 2.9).

Proportion cover fraction of density heat-maps showed the prior sampled almost the entire space between 0 and 1, with higher densities at either complete cover during all of 2012 or declining to zero during the second half of 2012. The heat-maps illustrated tightening of posterior density after assimilating observations at each site. There was 20 times the posterior density than the sampled prior at those points (Figure 2.12, Figure 2.13, Figure 2.14, Figure 2.15, Figure 2.16). Altona, Blairgowrie, Point Richards and Swan Bay South posteriors all trend downwards over 2012. Blairgowrie had the largest initial proportion cover density ranging from 0.8 to 1 (Figure 2.13). Point Henry initialised at lower values than the other four sites and exhibited a seasonal trend seeing it return to the same starting values of October 2011, 0.6-0.8 (Figure 2.14).

Canopy height fraction of density heat-maps illustrated tightening of posterior density after assimilating observations at each site. There was 4-5 times the posterior density than the sampled prior at those points (Figure 2.17, Figure 2.18, Figure 2.19, Figure 2.20, Figure 2.21). At each site the mean canopy height posterior IQR decreased by 2.5-5 times from the prior IQR (Table 2.5). The posterior medians at Altona, Blairgowrie and Point Henry all shifted from the prior median 0.4000 m. Altona shifted downwards to 0.2731 m, Blairgowrie had a more subtle shift downwards to 0.3707 m, while Point Henry shifted upwards to 0.4463 m. Point Richards and Swan Bay South posterior medians (0.3137 m and 0.2579 m respectively) shifted upwards form their prior median of 0.2000 m. Point Henry had the tallest canopy height in 2012, followed by Blairgowrie and Point Richards, with Altona and Swan

Bay South having the shortest canopy heights (Figure 2.27).

Light attenuation coefficient fraction of density heat-maps illustrated tightening of posterior density after assimilating observations at each site. There was 4-5 times the posterior density than the sampled prior at those points (Figure 2.22, Figure 2.23, Figure 2.24, Figure 2.25, Figure 2.26). Altona (Figure 2.22) and Point Richards (Figure 2.25) had the largest spread of posterior density of all the sites, while Swan Bay South (Figure 2.26) was the most tightly constrained. The mean light attenuation coefficient posterior had smaller spread than the prior, with posterior IQR one quarter of the prior IQR (Figure 2.27). Altona and Blairgowrie saw shifts in posterior medians to $0.3238\,\mathrm{m}^{-1}$ and $0.2149\,\mathrm{m}^{-1}$ from a prior median of $0.2000\,\mathrm{m}^{-1}$. From a prior median of $0.3000\,\mathrm{m}^{-1}$, Point Henry's posterior median shifted slightly to $0.2804\,\mathrm{m}^{-1}$, while Point Richards shifted further to $0.3545\,\mathrm{m}^{-1}$. Swan Bay South prior median $0.1000\,\mathrm{m}^{-1}$ shifted to $0.1409\,\mathrm{m}^{-1}$ which was the lowest light attenuation coefficient followed by Blairgowrie, Point Henry, Altona, with Point Richards as the highest site for 2012.

Effective absorbance, while an unobserved parameter, displayed shifts in the posterior from the prior indicating that the observations had an effect on constraining this parameter. The same prior was assigned to all sites with median 0.1500 and IQR 0.1000 and approximately the same amount of learning occurred at each site, with posterior medians 0.0963-0.1066 and IQR 0.0507-0.0609.

Leaf mortality rate posteriors are more concentrated from the priors at each site. All posterior IQR decreased by 2-4 times from the prior IQR. Swan Bay South had the highest leaf mortality rate posterior median $0.4939\times10^{-2}\,\mathrm{d}^{-1}$ and was the only site that remained relatively unchanged from the prior median $0.5000\times10^{-2}\,\mathrm{d}^{-1}$. Altona followed with a leaf mortality rate posterior median of $0.4412\times10^{-2}\,\mathrm{d}^{-1}$, then Blairgowrie ($0.4136\times10^{-2}\,\mathrm{d}^{-1}$), Point Richards ($0.3787\times10^{-2}\,\mathrm{d}^{-1}$) and Point

Henry ($0.2695\times10^{-2}\,d^{-1}$) was the lowest. There were no significant shifts from root/rhizome mortality rate priors to posteriors at all sites.

Translocation rate posterior medians all shifted upwards from the prior medians, as well as increased posterior IQR at each site. Altona had the highest posterior median ($0.4599\times10^{-2}\,d^{-1}$), followed by Swan Bay South ($0.3647\times10^{-2}\,d^{-1}$), Blairgowrie ($0.3220\times10^{-2}\,d^{-1}$), Point Richards ($0.2408\times10^{-2}\,d^{-1}$) and Point Henry ($0.1923\times10^{-2}\,d^{-1}$).

Compensation scalar irradiance priors and posteriors remained relatively unchanged at Blairgowrie, Point Henry and Point Richards. Altona and Swan Bay South posterior central tendency shifted from a prior of 3.0000 mol photon m^{-2} d^{-1} to 7.1438 mol photon m^{-2} d^{-1} at Altona and 6.2655 mol photon m^{-2} d^{-1} at Swan Bay South with both sites doubling their variability.

Dry weight specific area of seagrass was constrained well by the observations shown in the concentrated posterior distributions with lower variability (Figure 2.28). Altona had the highest posterior median (0.7494×10^{-2} (g DW m^{-2})$^{-1}$), followed by Swan Bay South (0.6739×10^{-2} (g DW m^{-2})$^{-1}$), Point Richards (0.5218×10^{-2} (g DW m^{-2})$^{-1}$), Point Henry (0.4660×10^{-2} (g DW m^{-2})$^{-1}$) and Blairgowrie (0.4164×10^{-2} (g DW m^{-2})$^{-1}$) had the lowest. All sites have posterior IQR 0.1175×10^{-2} - 0.1883×10^{-2} (g DW m^{-2})$^{-1}$ down from prior IQR 0.7300×10^{-2} (g DW m^{-2})$^{-1}$ (Table 2.5). Equilibrium fraction of biomass below-ground posteriors remain unchanged from the priors at all sites. Water column height at Blairgowrie, Point Henry and Point Richards remained unchanged from prior to posterior. The posterior median at Altona shifted upwards to 0.5453 m, while the posterior median at Swan Bay South shifted downwards to 0.4626 m. Radiation use efficiency posterior distributions all shifted upwards from the prior distribution (median 0.4000 g DW (mol photon)$^{-1}$, IQR 0.2700 g DW (mol photon)$^{-1}$). Altona had the largest shift in posterior me-

dian (0.6999 g DW (mol photon)$^{-1}$), but also twice the IQR (0.5124 g DW (mol photon)$^{-1}$). All other sites had a similar posterior variability to the prior. Point Henry had the next largest shift in posterior median to 0.6393 g DW (mol photon)$^{-1}$, followed by Point Richards (0.5892 g DW (mol photon)$^{-1}$), Swan Bay South (0.5892 g DW (mol photon)$^{-1}$), and Blairgowrie had the smallest (0.4864 g DW (mol photon)$^{-1}$). Seagrass diversity factor saw minimal change in spread from prior to posterior with small shifts in central tendency priors to posteriors.

Parameter	**Site**	**Prior 2012**	**Posterior 2012**
μ_{h_c}	Altona	0.4015 (0.2768)	0.2731 (0.0532)
	Blairgowrie	0.3988 (0.2745)	0.3707 (0.0724)
	Point Henry	0.4000 (0.2721)	0.4463 (0.0837)
	Point Richards	0.2003 (0.1368)	0.3137 (0.0637)
	Swan Bay South	0.2006 (0.1382)	0.2579 (0.0511)
μ_{k_d}	Altona	0.1995 (0.1664)	0.3238 (0.0640)
	Blairgowrie	0.1992 (0.1652)	0.2149 (0.0417)
	Point Henry	0.2985 (0.2470)	0.2804 (0.0597)
	Point Richards	0.2989 (0.2494)	0.3545 (0.0652)
	Swan Bay South	0.1010 (0.0831)	0.1409 (0.0282)
μ_{A}	Altona	0.1503 (0.1028)	0.1036 (0.0563)
	Blairgowrie	0.1501 (0.1026)	0.1066 (0.0507)
	Point Henry	0.1484 (0.1016)	0.0963 (0.0609)
	Point Richards	0.1505 (0.1022)	0.0992 (0.0571)
	Swan Bay South	0.1501 (0.1029)	0.0999 (0.0592)
μ_{ζ_A}	Altona	0.4974×10^{-2} (0.7228×10^{-2})	0.4412×10^{-2} (0.2333×10^{-2})

	Blairgowrie	0.5024×10^{-2} (0.7195×10^{-2})	0.4136×10^{-2} (0.1964×10^{-2})
	Point Henry	0.4972×10^{-2} (0.7205×10^{-2})	0.2695×10^{-2} (0.1673×10^{-2})
	Point Richards	0.5065×10^{-2} (0.7374×10^{-2})	0.3787×10^{-2} (0.2057×10^{-2})
	Swan Bay South	0.5006×10^{-2} (0.7183×10^{-2})	0.4939×10^{-2} (0.3165×10^{-2})
μ_{ζ_B}	Altona	0.0999×10^{-3} (0.1474×10^{-3})	0.0760×10^{-3} (0.0732×10^{-3})
	Blairgowrie	0.0996×10^{-3} (0.1451×10^{-3})	0.0992×10^{-3} (0.0950×10^{-3})
	Point Henry	0.0993×10^{-3} (0.1428×10^{-3})	0.0823×10^{-3} (0.1110×10^{-3})
	Point Richards	0.0991×10^{-3} (0.1456×10^{-3})	0.1097×10^{-3} (0.1556×10^{-3})
	Swan Bay South	0.1006×10^{-3} (0.1472×10^{-3})	0.0820×10^{-3} (0.0920×10^{-3})
μ_{τ}	Altona	0.1007×10^{-2} (0.1364×10^{-2})	0.4599×10^{-2} (0.4133×10^{-2})
	Blairgowrie	0.0992×10^{-2} (0.1359×10^{-2})	0.3220×10^{-2} (0.3785×10^{-2})
	Point Henry	0.0998×10^{-2} (0.1368×10^{-2})	0.1923×10^{-2} (0.2587×10^{-2})
	Point Richards	0.0995×10^{-2} (0.1359×10^{-2})	0.2408×10^{-2} (0.3201×10^{-2})
	Swan Bay South	0.0986×10^{-2} (0.1350×10^{-2})	0.3647×10^{-2} (0.4411×10^{-2})
$\mu_{\mathrm{E}_{comp}}$	Altona	2.9766 (2.0473)	7.1438 (5.9472)
	Blairgowrie	2.9932 (2.0709)	2.9868 (2.1179)
	Point Henry	3.0106 (2.0631)	2.6726 (2.0854)

	Point Richards	2.9942 (2.0466)	3.2284 (2.3697)
	Swan Bay South	2.9951 (2.0602)	6.2655 (5.1140)
Ω	Altona	0.5013×10^{-2} (0.7233×10^{-2})	0.7494×10^{-2} (0.1883×10^{-2})
	Blairgowrie	0.5026×10^{-2} (0.7359×10^{-2})	0.4164×10^{-2} (0.1175×10^{-2})
	Point Henry	0.4986×10^{-2} (0.7253×10^{-2})	0.4660×10^{-2} (0.1192×10^{-2})
	Point Richards	0.4991×10^{-2} (0.7268×10^{-2})	0.5218×10^{-2} (0.1320×10^{-2})
	Swan Bay South	0.4983×10^{-2} (0.7302×10^{-2})	0.6739×10^{-2} (0.1613×10^{-2})
f_b	Altona	0.7719 (0.0369)	0.7812 (0.0343)
	Blairgowrie	0.7725 (0.0368)	0.7648 (0.0323)
	Point Henry	0.7722 (0.0369)	0.7805 (0.0285)
	Point Richards	0.7723 (0.0370)	0.7729 (0.0298)
	Swan Bay South	0.7722 (0.0368)	0.7733 (0.0321)
h_{wc}	Altona	0.5008 (0.1342)	0.5453 (0.1086)
	Blairgowrie	0.5011 (0.1357)	0.5277 (0.1767)
	Point Henry	0.4999 (0.1365)	0.5039 (0.1614)
	Point Richards	0.5002 (0.1351)	0.5171 (0.1355)
	Swan Bay South	0.5006 (0.1359)	0.4626 (0.1218)

r_{ue}	Altona	0.4003 (0.2721)	0.6999 (0.5124)
	Blairgowrie	0.3997 (0.2738)	0.4864 (0.2838)
	Point Henry	0.4005 (0.2763)	0.6393 (0.3303)
	Point Richards	0.3994 (0.2729)	0.5892 (0.3419)
	Swan Bay South	0.4017 (0.2770)	0.5462 (0.3172)
SDF	Altona	0.2003 (0.0951)	0.2138 (0.0959)
	Blairgowrie	0.2001 (0.0956)	0.2028 (0.0754)
	Point Henry	0.1999 (0.0945)	0.1665 (0.0737)
	Point Richards	0.1995 (0.0950)	0.1804 (0.0891)
	Swan Bay South	0.1990 (0.0950)	0.1720 (0.0903)

Table 2.5 : Parameter median and interquartile range (IQR) at each site for 2012 prior and posterior.

2.3.2 Posterior results for 2012 with a prediction for 2013

The above-ground biomass posterior estimates for 2012 assimilated the trend of the observations well (Figure 2.29). Blairgowrie has the highest initial above-ground biomass in October 2011 followed by Point Richards, Point Henry, Swan Bay South then Altona. All sites exhibited seasonality with above-ground biomass peaks in the summer months. Point Henry was the only site that did not exhibit a downward trend across 2012, the remaining four sites all exhibited seasonality with a downward trend by December 2012. Blairgowrie and Point Henry are predicted to have the highest above-ground biomass for 2013 followed by Point Richards, Swan Bay South and Altona. Above-ground biomass predictions for Blairgowrie, Point Richards and Swan Bay South line up with the validation data for 2013, while the Altona prediction under-predicts the observations and the Point Henry prediction over-predicts them (Figure 2.29). The 2013 prediction for Altona continues a downward trend in above-ground biomass from the posterior estimate for 2012. This prediction misses the validation data completely as it shows a clear return to biomass levels at the start of October 2011. The combination of posterior parameter estimates for 2012 and forcings used in the model provided no indication of an upturn in 2013. Potentially we have not captured all forcings present at Altona and further environmental data should be obtained and analysed to understand whether there were environmental changes driving the difference in the two years. Similarly, the same investigation should be undertaken at Point Henry, particularly as the proportion cover predictions exhibit the same unsuccessful 2013 predictions. Similarly to above-ground biomass, proportion cover predicts the 2013 observations well at Blairgowrie, Point Richards and Swan Bay South while under-predicting at Altona. Point Henry has quite volatile 2013 observations which span from 0 to 90% cover and while the prediction overlaps some of the observations it also over-predicts many of them (Figure 2.30). Proportion cover posterior estimates at each site align reasonably well with

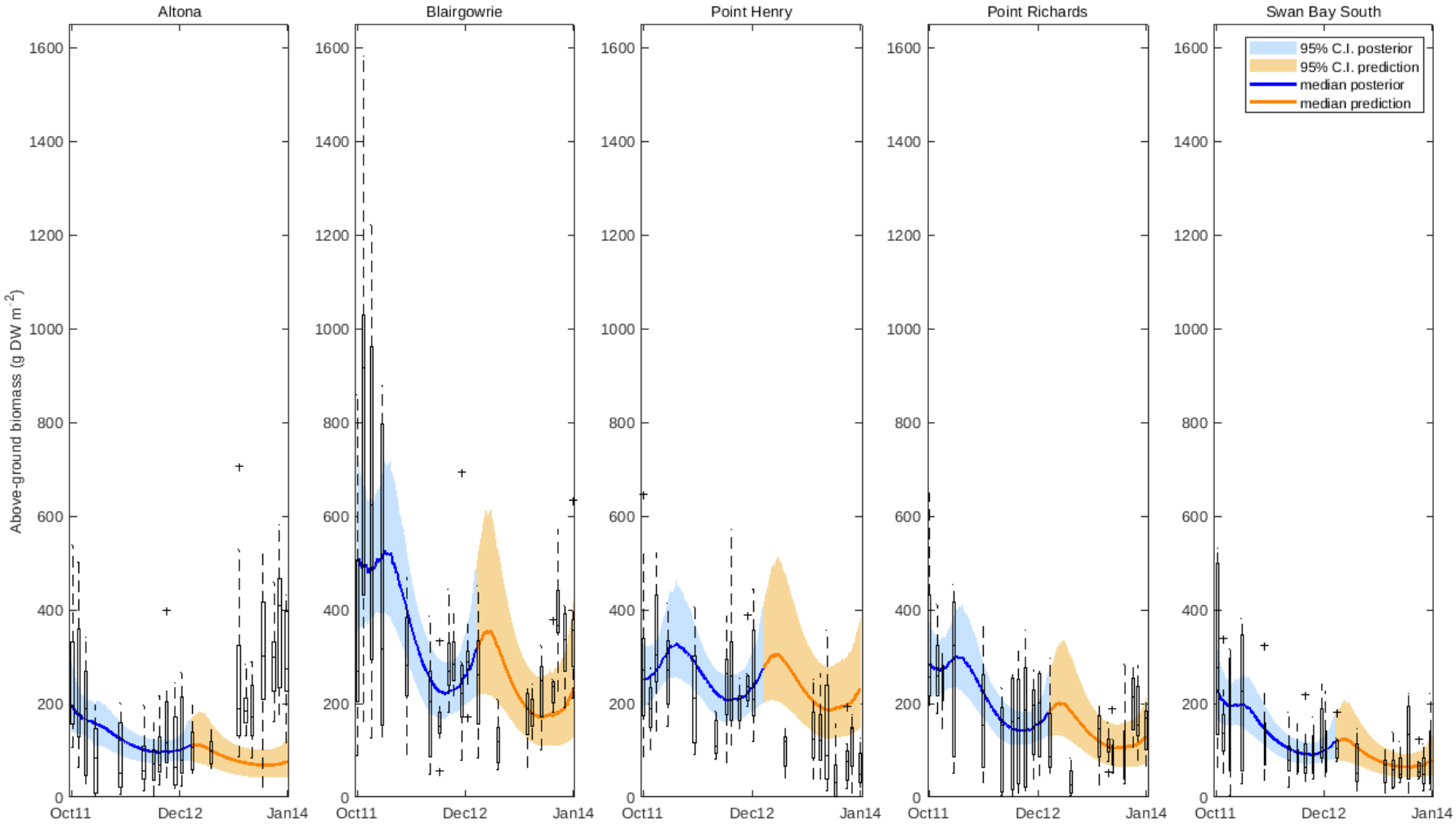

Figure 2.29 : Above-ground biomass posterior medians (solid blue line) and 95% credible intervals (shaded blue) for the first year of observations, and prediction medians (solid orange line) and 95% credible intervals (shaded orange) for the second year of observations. Observations (n=9 at each sampling time-point) are shown as box-plots with potential outliers (+).

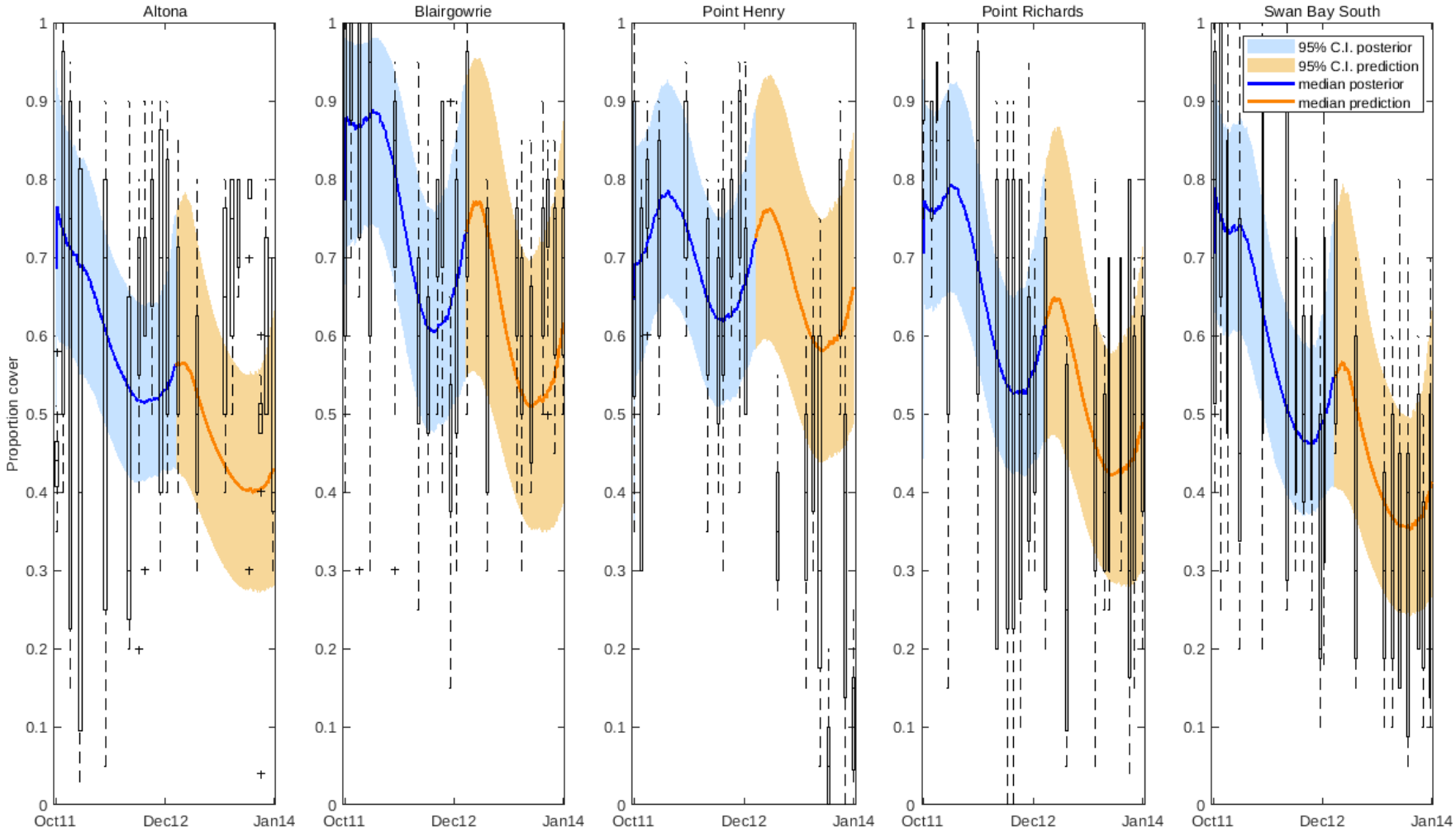

Figure 2.30 : Proportion cover posterior medians (solid blue line) and 95% credible intervals (shaded blue) for the first year of observations, and prediction medians (solid orange line) and 95% credible intervals (shaded orange) for the second year of observations. Observations (n=9 at each sampling time-point) are shown as box-plots with potential outliers (+).

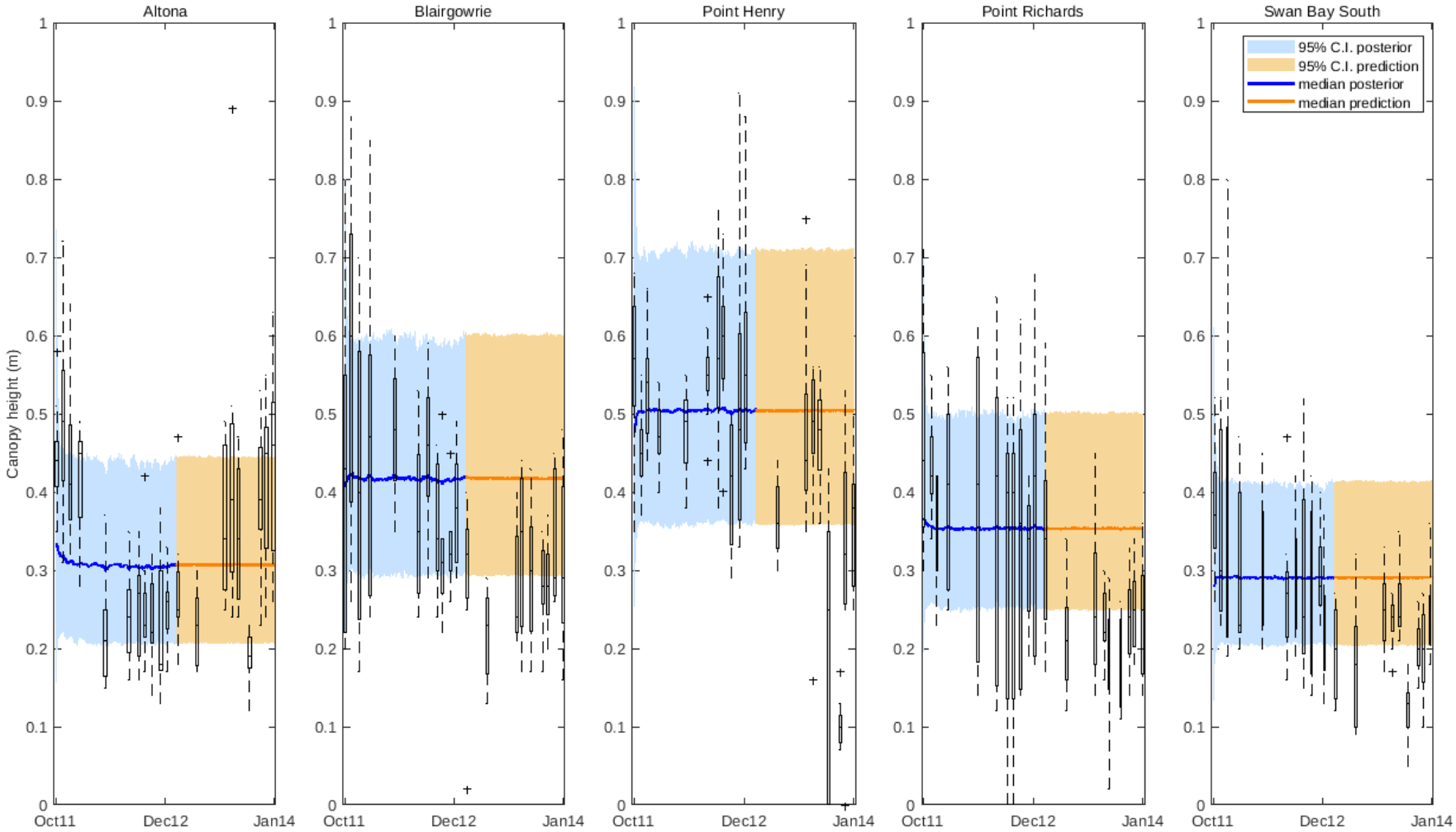

Figure 2.31 : Canopy height posterior medians (solid blue line) and 95% credible intervals (shaded blue) for the first year of observations, and prediction medians (solid orange line) and 95% credible intervals (shaded orange) for the second year of observations. Observations (n=9 at each sampling time-point) are shown as box-plots with potential outliers (+).

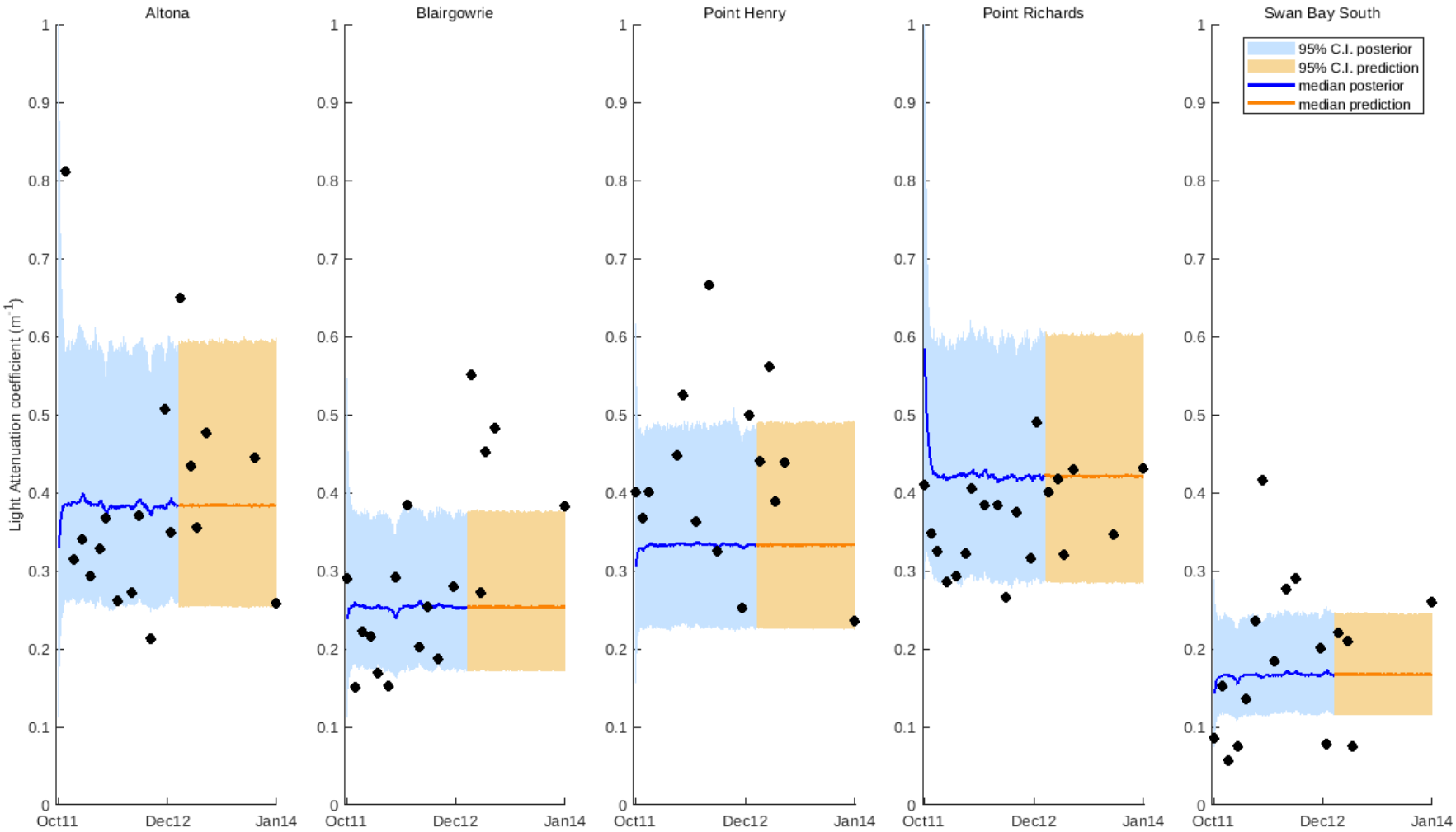

Figure 2.32 : Light attenuation coefficient posterior medians (solid blue line) and 95% credible intervals (shaded blue) for the first year of observations, and prediction medians (solid orange line) and 95% credible intervals (shaded orange) for the second year of observations. Observations (n=1 at each sampling time-point) are shown as solid black circles.

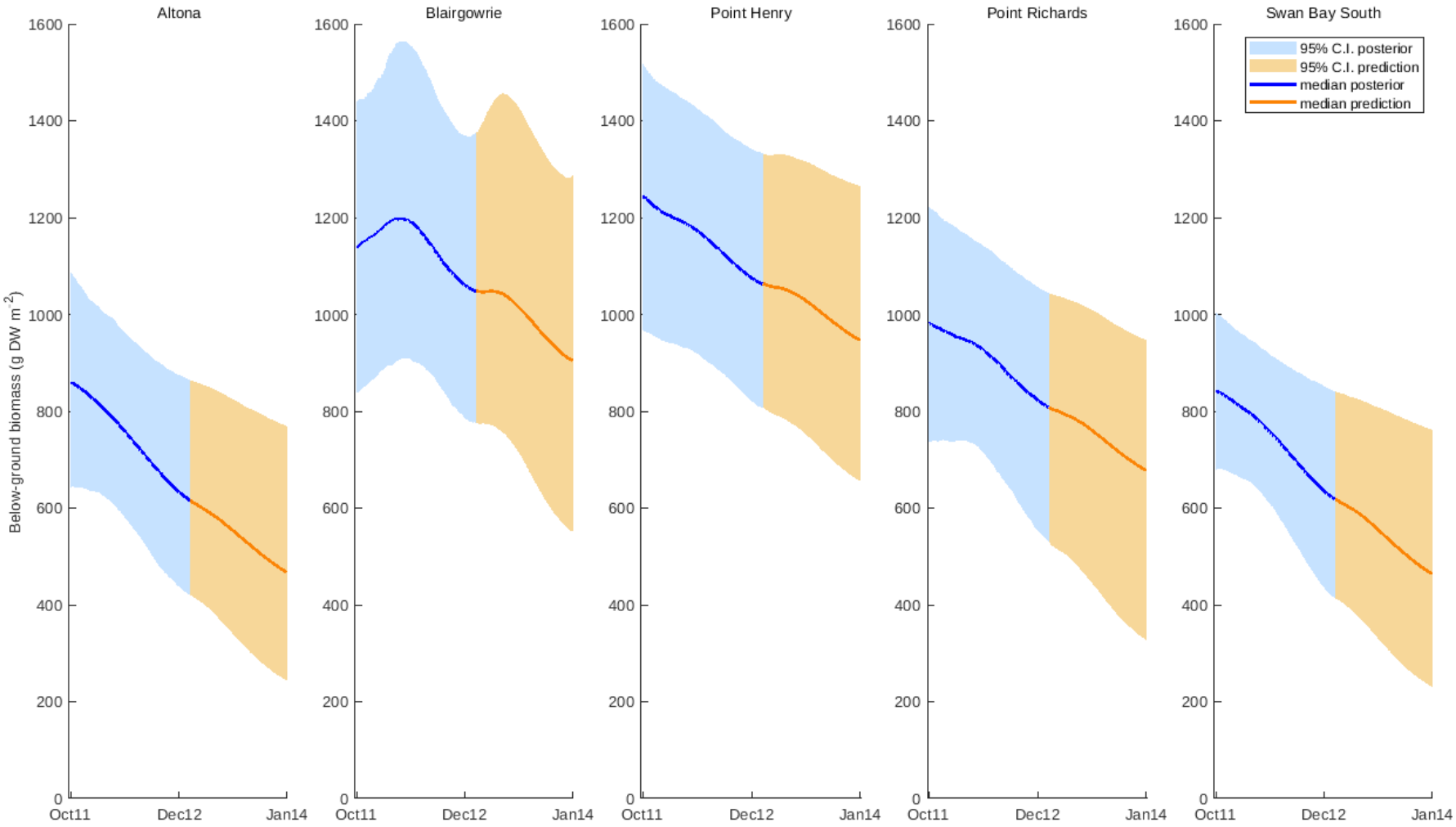

Figure 2.33 : Below-ground biomass posterior medians (solid blue line) and 95% credible intervals (shaded blue) for the first year of observations, and prediction medians (solid orange line) and 95% credible intervals (shaded orange) for the second year of observations. Observations (n=9 at each sampling time-point) are shown as box-plots with potential outliers (+).

the observations for 2012, given the large variability at some time-points (Figure 2.30). Blairgowrie has the largest initial proportion cover followed by Swan Bay South, Point Richards, Altona then Point Henry. Seasonality in the form of proportion cover peaks during summer is present at all sites with Altona, Blairgowrie, Point Richards and Swan Bay South trending downwards during 2012. The predicted trend for 2013 at these four sites is a continued drop in proportion cover, while at Point Henry the predicted trend remains stable and unchanged for 2013.

Canopy height during 2012 was tallest at Point Henry, followed by Blairgowrie and Point Richards, with Altona and Swan Bay South having the shortest seagrass canopies. The first order autoregressive process representing the canopy height posterior during 2012 on average fits the observations well at Blairgowrie, Point Henry, Point Richards and Swan Bay South. The observations at Altona during 2012 look to be grouped into two stages, an initial taller height from October 2011 to January 2012 centred around 0.4-0.5 m, and a secondary lower height from April 2012 centred around 0.2000 m. The highest area of posterior probability is 0.3000 m throughout 2012, meaning the first order autoregressive is capturing the long-term stationary central tendency but not necessarily capturing the observed quantities of canopy height. Blairgowrie, Pont Henry, Point Richards and Swan Bay South all over-predicted the 2013 observations while Altona produced a more aligned 2013 prediction to the observations (Figure 2.31). An important highlight here is that canopy height model estimates displayed almost no variability through time while there was substantial variability in the data. The same is apparent for the light attenuation coefficient and may suggest the autoregressive process parameters need to be adjusted in future iterations to allow for more movement over time. Point Richards and Altona have the highest light attenuation coefficient posteriors across 2012, followed by Point Henry, Blairgowrie and then Swan Bay South with the lowest light attenuation. Altona, Blairgowrie and Swan Bay South posteriors capture the

observations in 2012 well. Point Richards posterior is centred on the higher end of the observations, but includes most of them in the 95% credible interval. Point Henry is centred on the lower end of the observations for 2012. Light attenuation coefficient predicts unseen 2013 observations well for Altona, Point Henry, Point Richards, and Swan Bay South, but under-predicts for Blairgowrie (Figure 2.32).

Point Henry and Blairgowrie have the highest initial below-ground biomass posteriors in October 2011, followed by Point Richards, then Altona and Swan Bay South with the lowest of the five sites. Below-ground biomass posteriors for Altona, Point Henry, Point Richards and Swan Bay South are all decreasing over 2012, Blairgowrie is the only site that sees an initial increase before decreasing for the remainder of 2012 (Figure 2.33). All predictions for 2013 see further decreases in below-ground biomass.

2.3.3 Prior to posterior results for 2013

Heat-maps of fraction of above-ground biomass density illustrated tightening of posterior density after assimilating observations at each site. There was 10-20 times the posterior density than the sampled prior at those points (Figure 2.34, Figure 2.35, Figure 2.36, Figure 2.37, Figure 2.38). Altona (Figure 2.34) and Blairgowrie (Figure 2.35) displayed an increasing posterior trend over 2013, Point Henry decreased (Figure 2.36), Point Richards and Swan Bay South remained stable (Figure 2.37, Figure 2.38).

Heat-maps of fraction of proportion cover density showed the prior sampled almost the entire space between 0 and 1 with higher densities at either complete cover during all of 2013 or declining to zero from April 2013. The heat-maps illustrated tightening of posterior density after assimilating observations at each site. There was 20-25 times the posterior density than the sampled prior at those points (Figure 2.39, Figure 2.40, Figure 2.41, Figure 2.42, Figure 2.43). Point Henry trended downwards

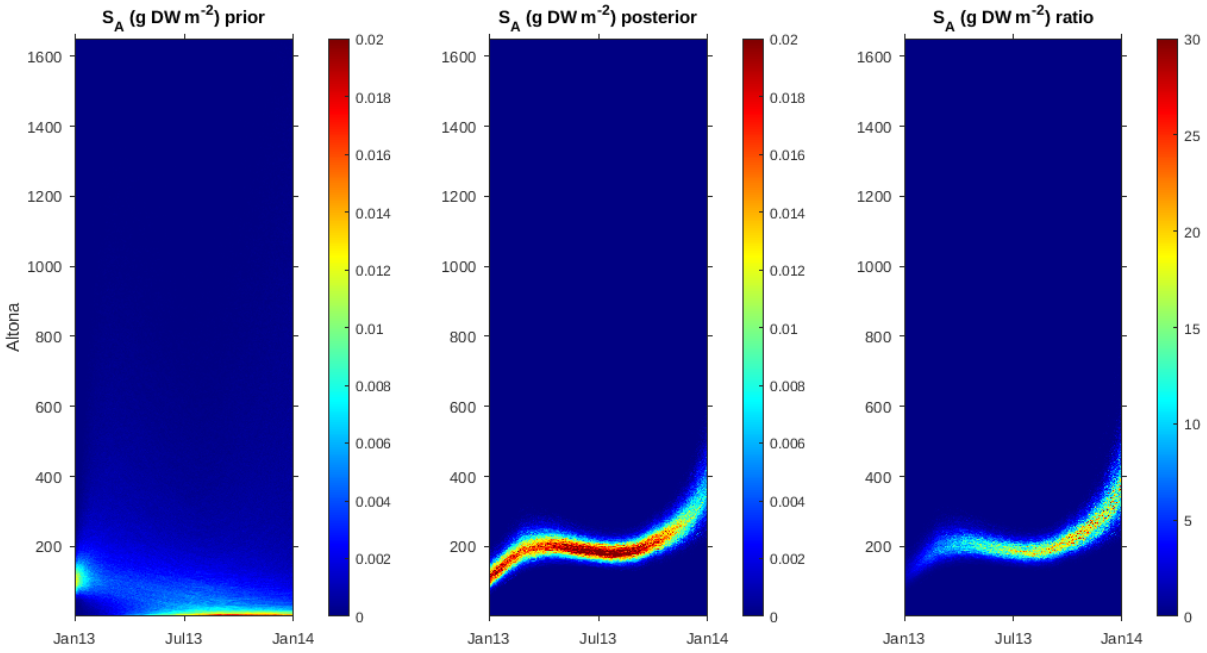

Figure 2.34 : Above-ground biomass (g DW m^{-2}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Altona in 2013.

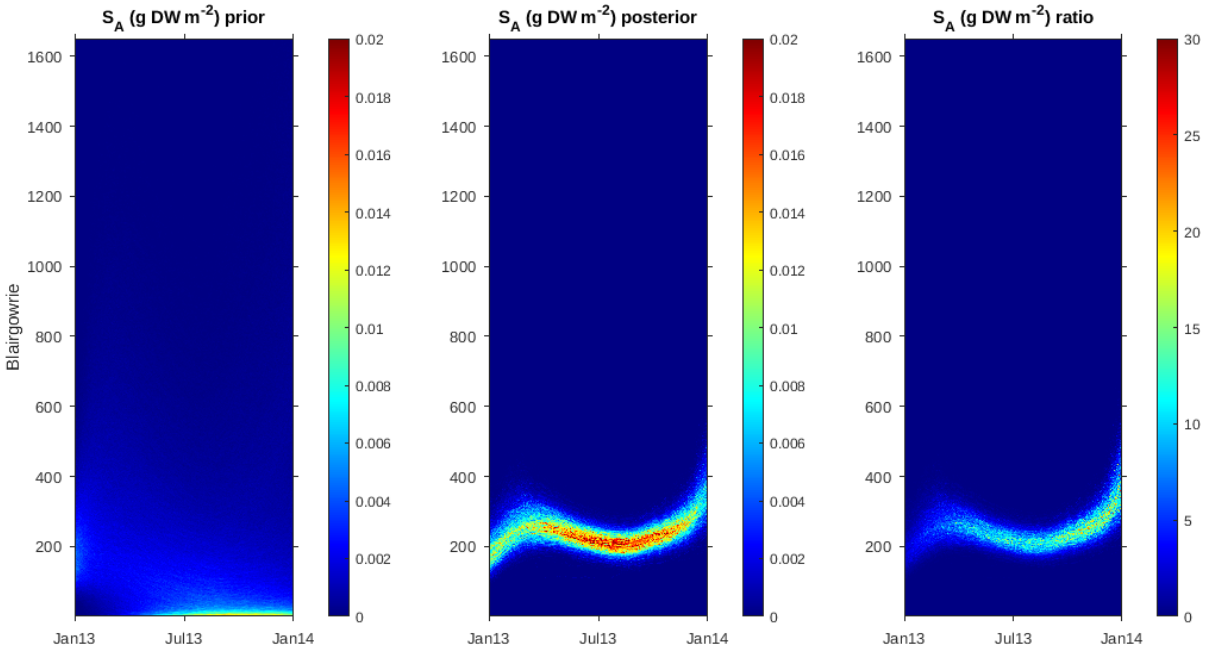

Figure 2.35 : Above-ground biomass (g DW m^{-2}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Blairgowrie in 2013.

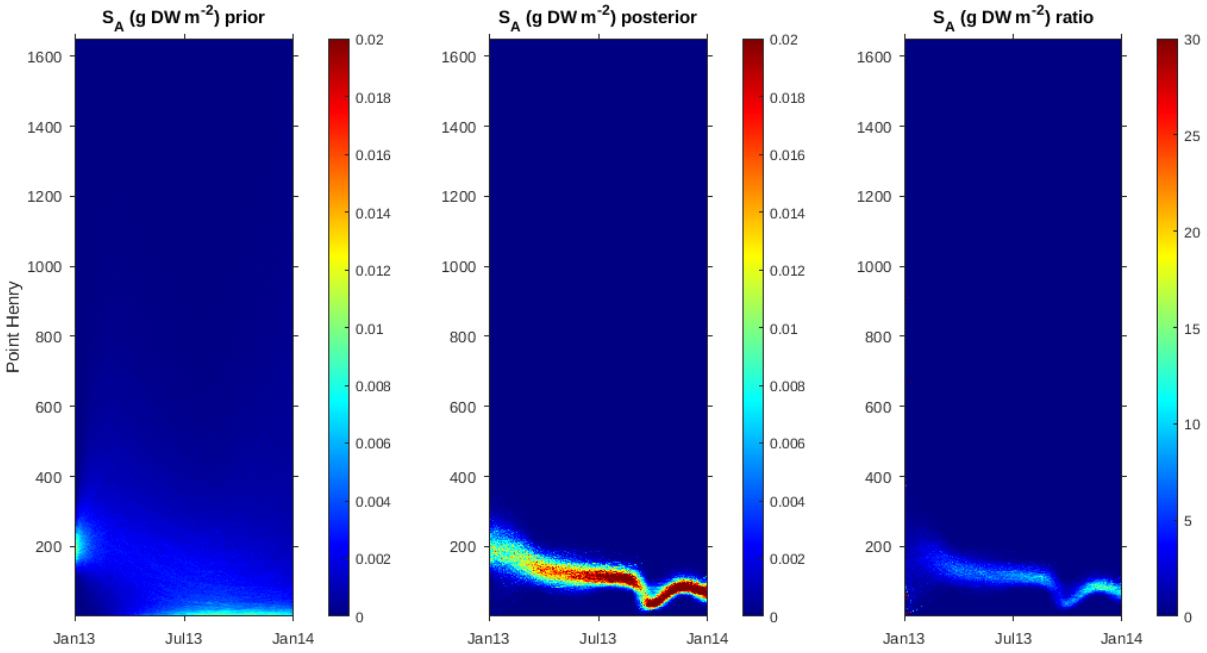

Figure 2.36 : Above-ground biomass (g DW m^{-2}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Point Henry in 2013.

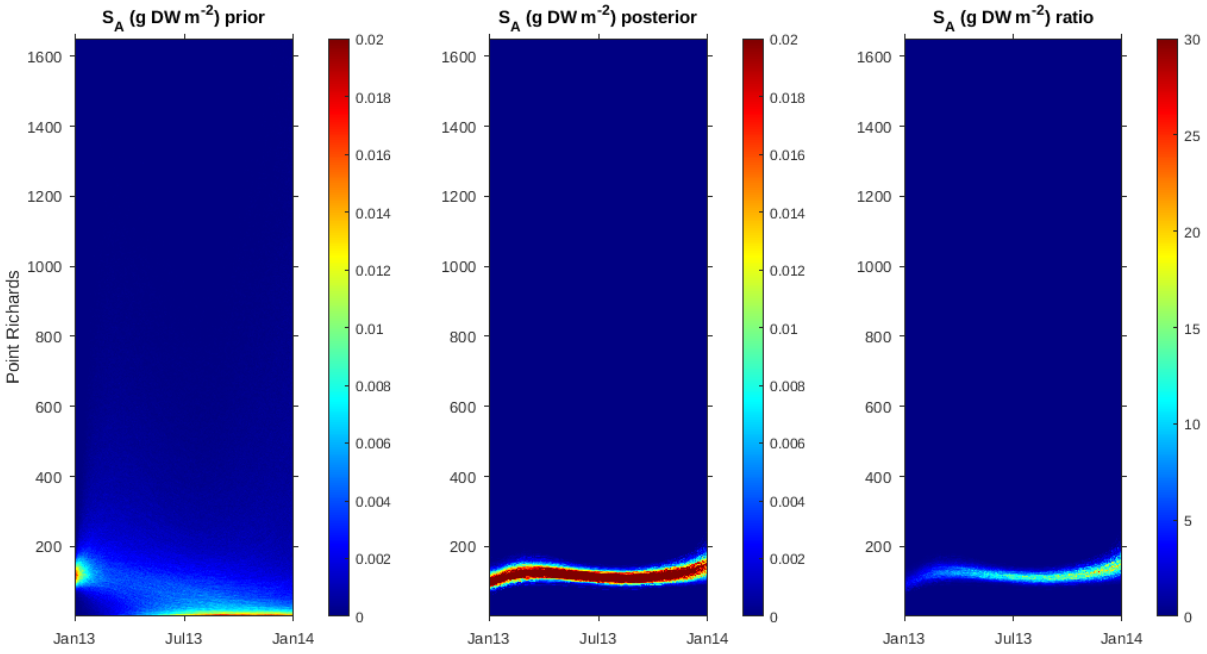

Figure 2.37 : Above-ground biomass (g DW m^{-2}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Point Richards in 2013.

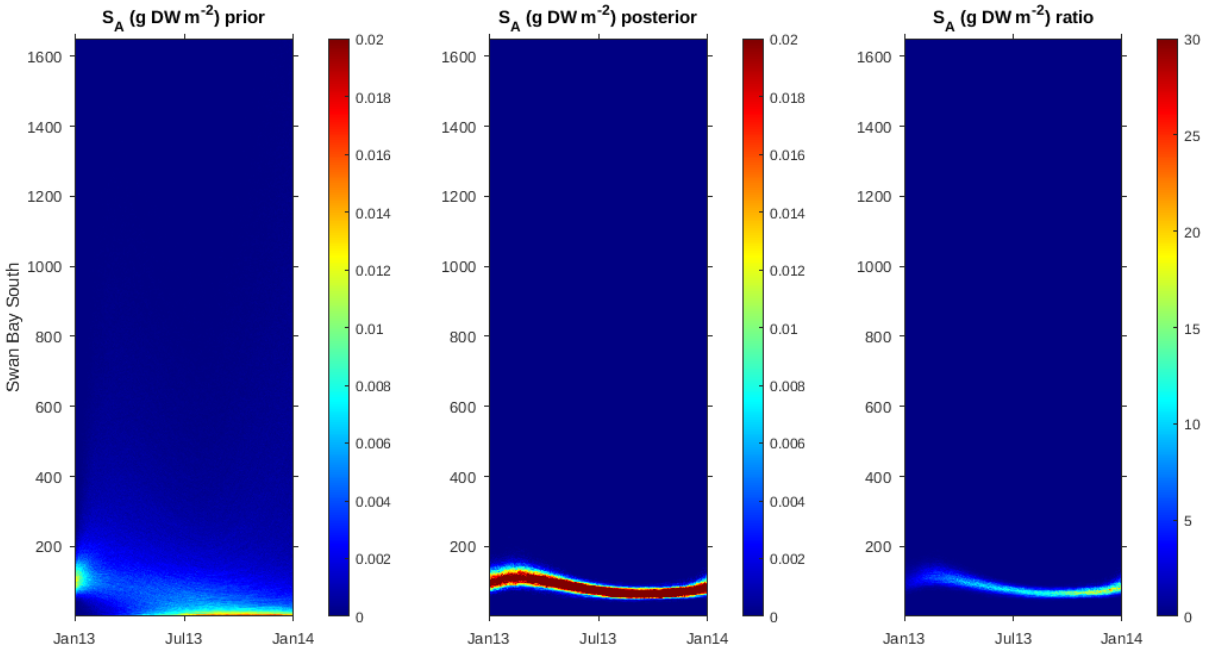

Figure 2.38 : Above-ground biomass (g DW m^{-2}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Swan Bay South in 2013.

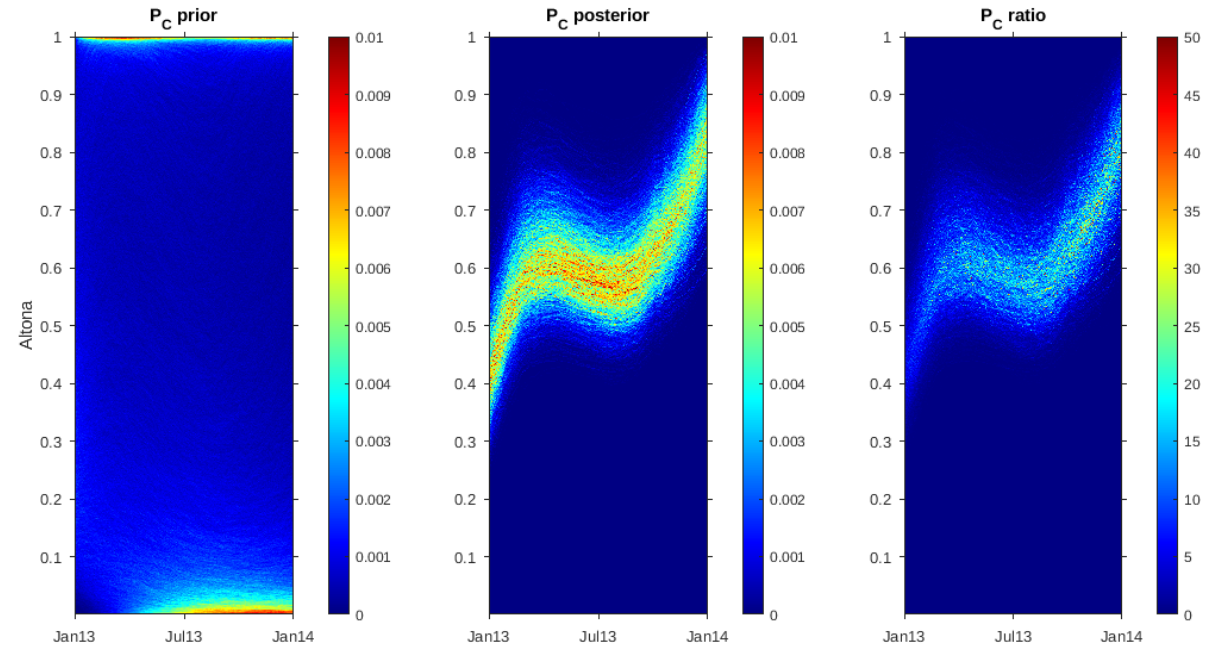

Figure 2.39 : Proportion cover fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Altona in 2013.

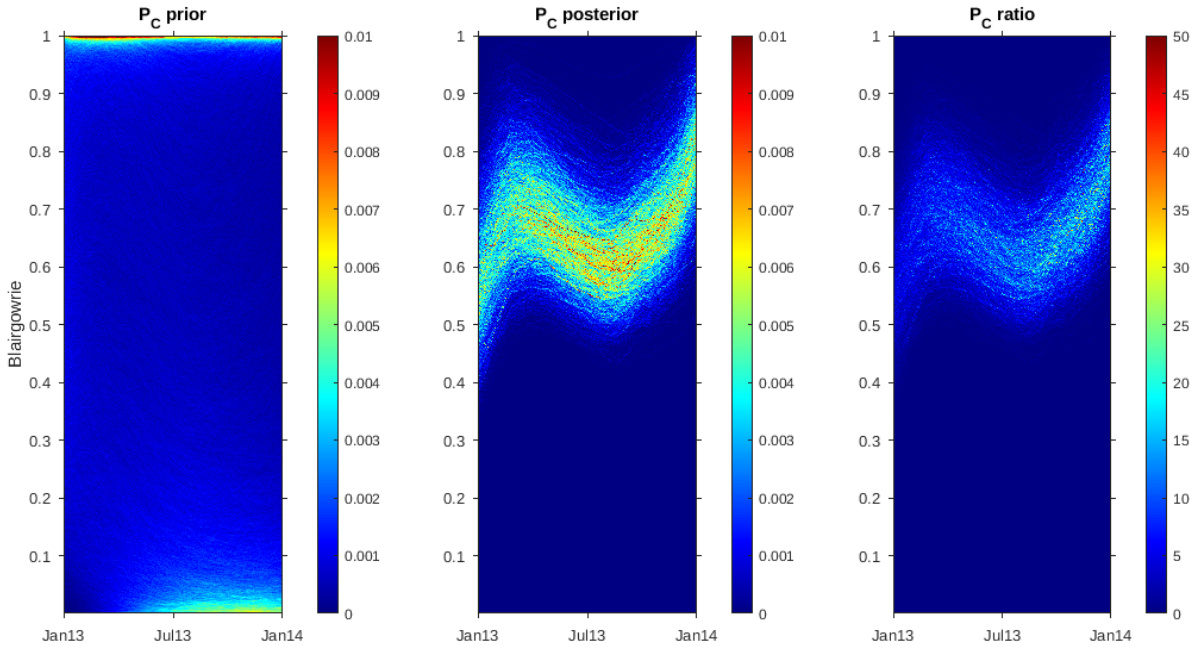

Figure 2.40 : Proportion cover fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Blairgowrie in 2013.

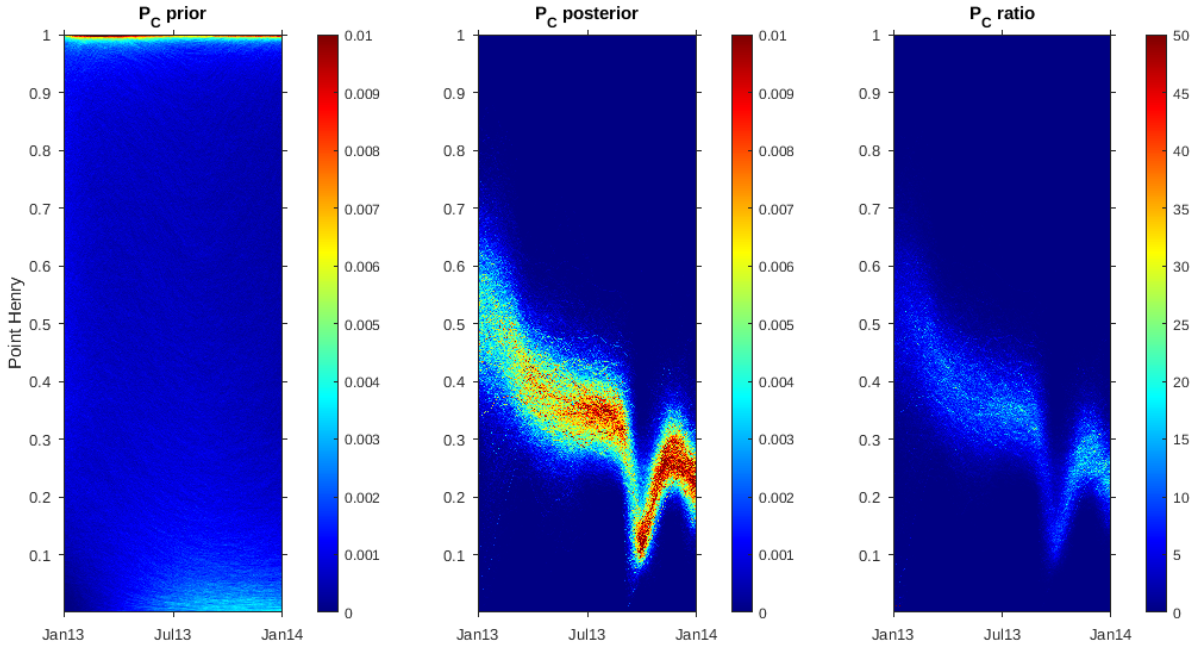

Figure 2.41 : Proportion cover fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Point Henry in 2013.

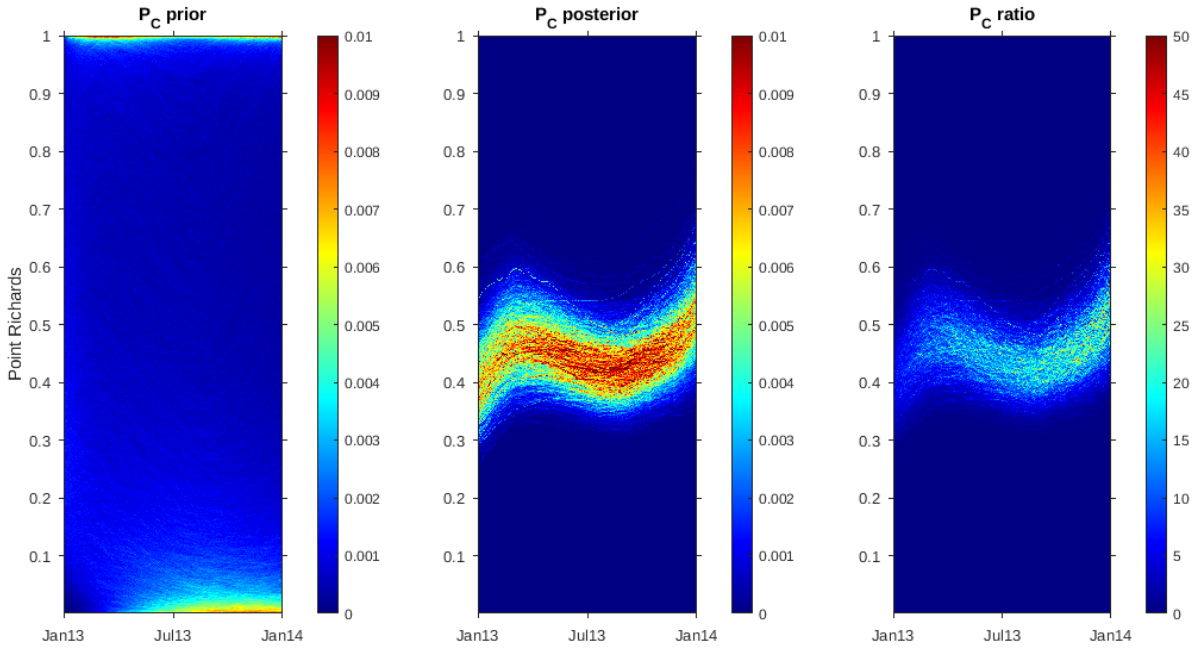

Figure 2.42 : Proportion cover fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Point Richards in 2013.

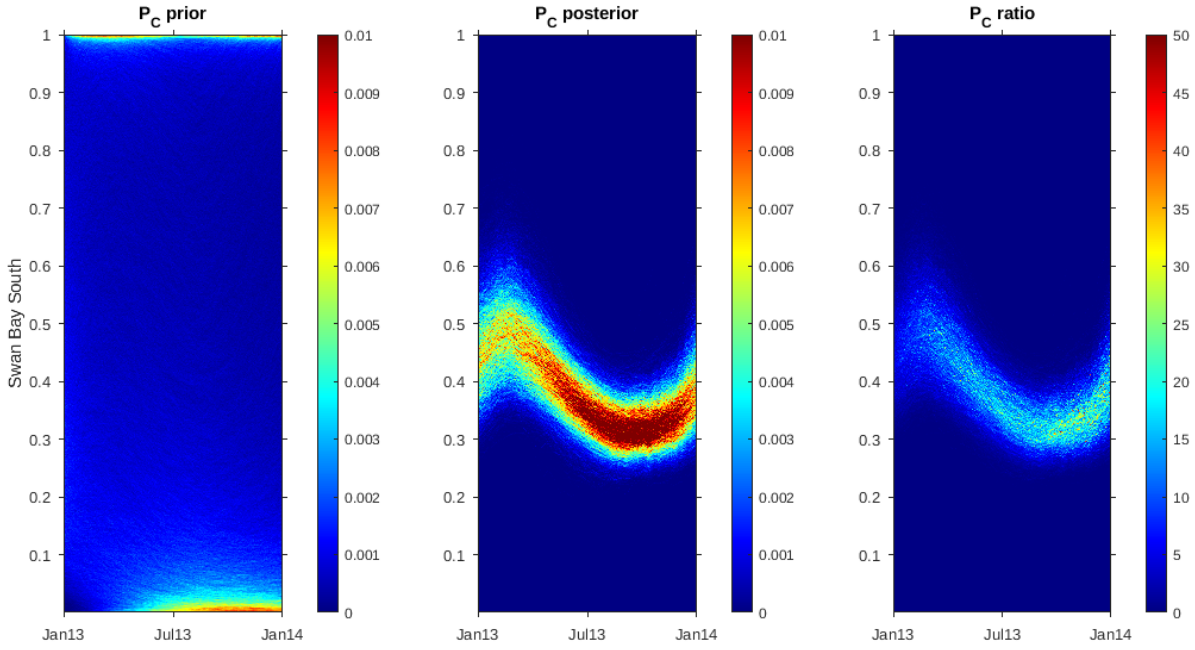

Figure 2.43 : Proportion cover fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Swan Bay South in 2013.

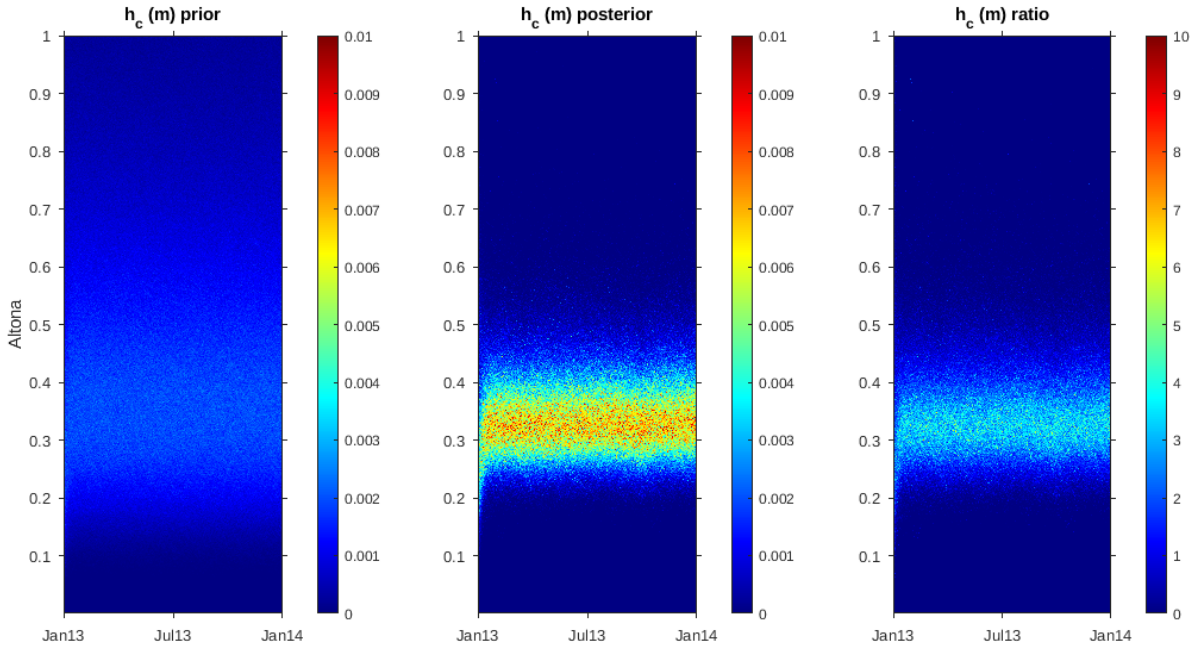

Figure 2.44 : Canopy height (m) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Altona in 2013.

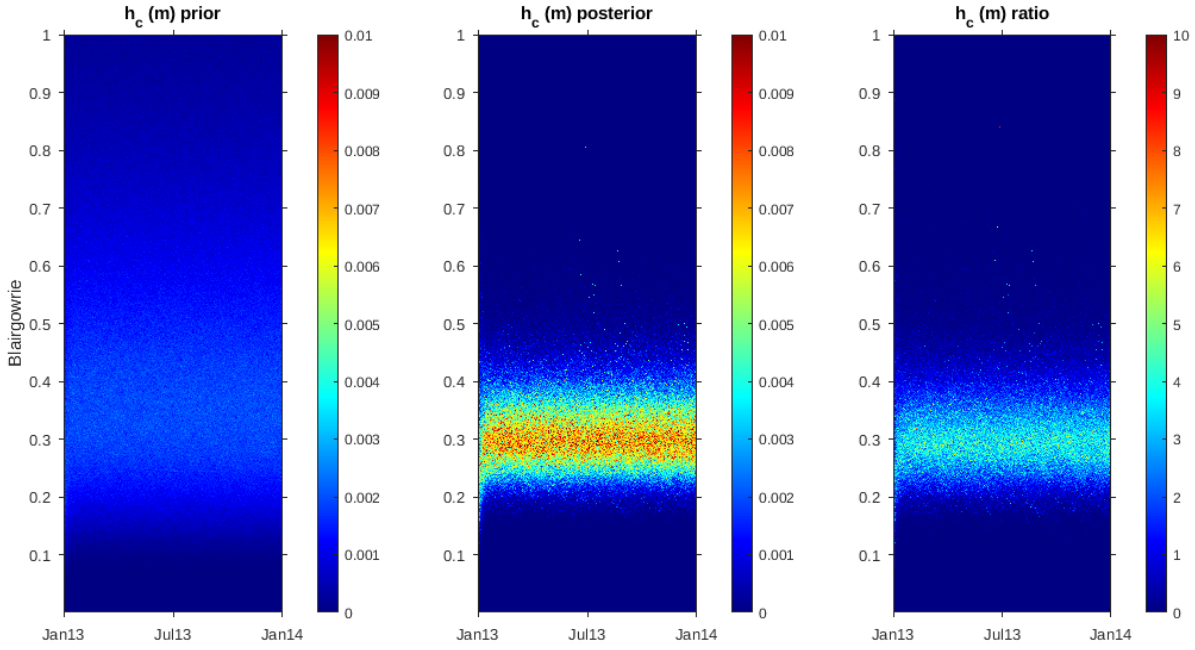

Figure 2.45 : Canopy height (m) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Blairgowrie in 2013.

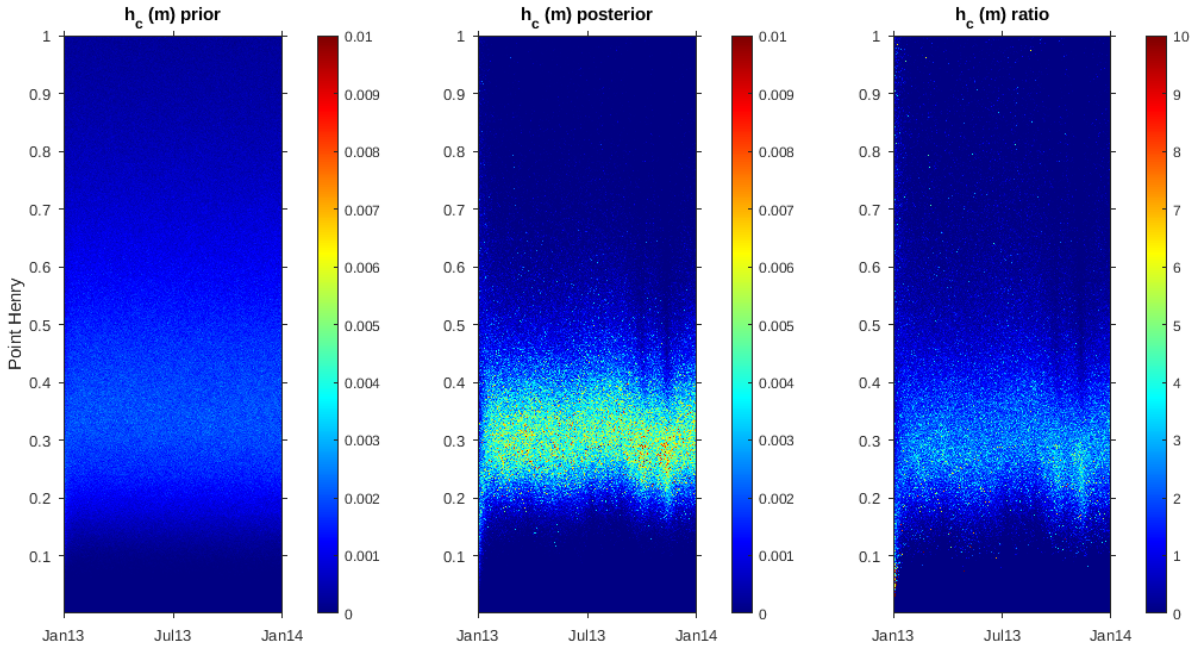

Figure 2.46 : Canopy height (m) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Point Henry in 2013.

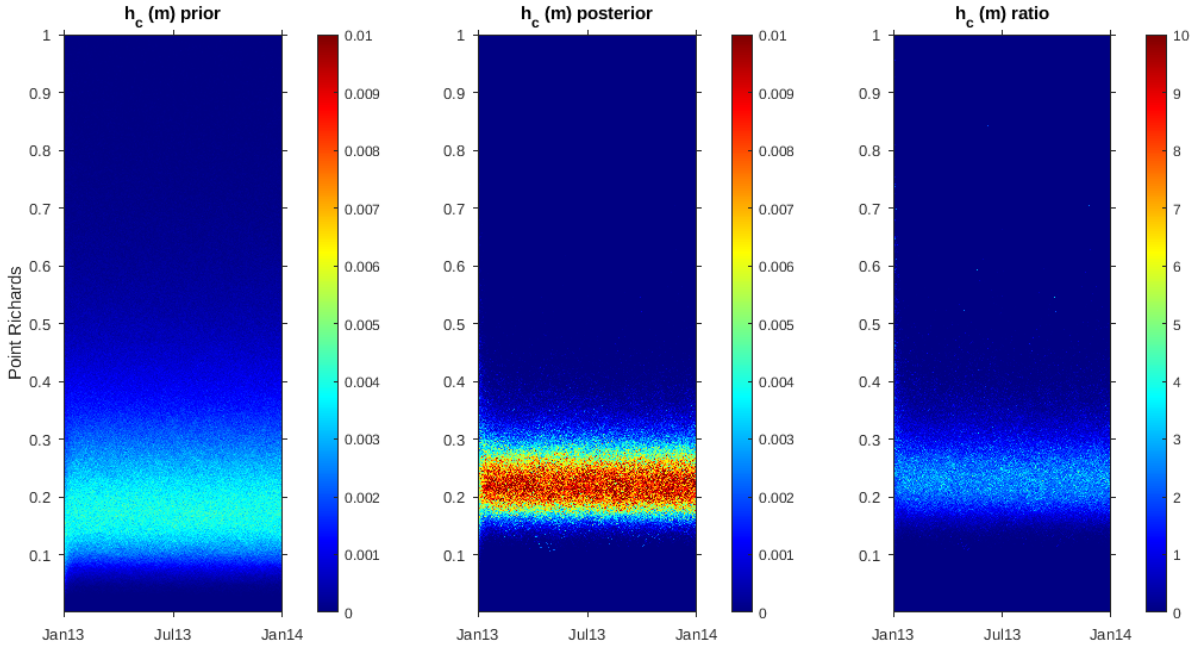

Figure 2.47 : Canopy height (m) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Point Richards in 2013.

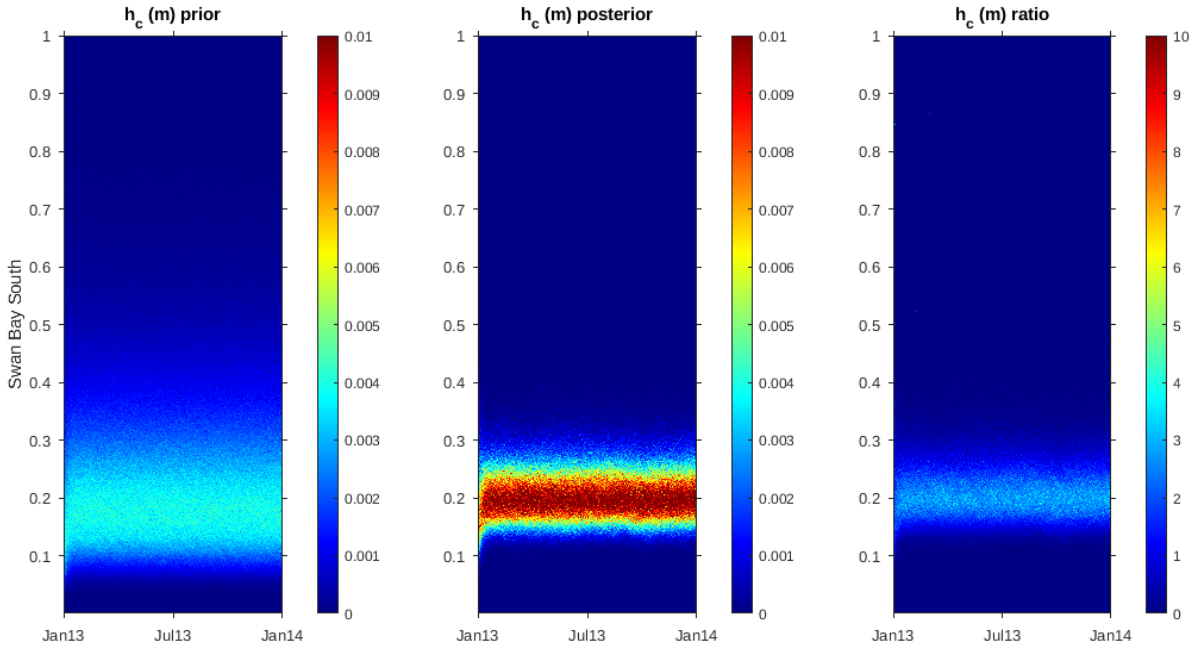

Figure 2.48 : Canopy height (m) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Swan Bay South in 2013.

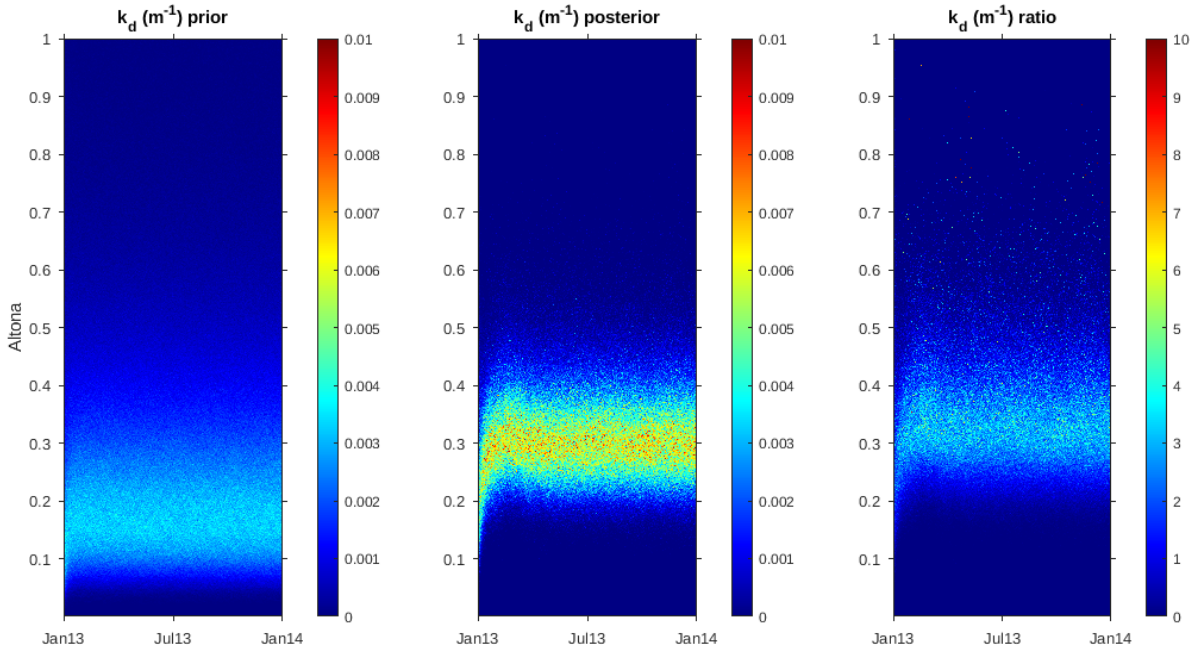

Figure 2.49 : Light attenuation coefficient (m^{-1}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Altona in 2013.

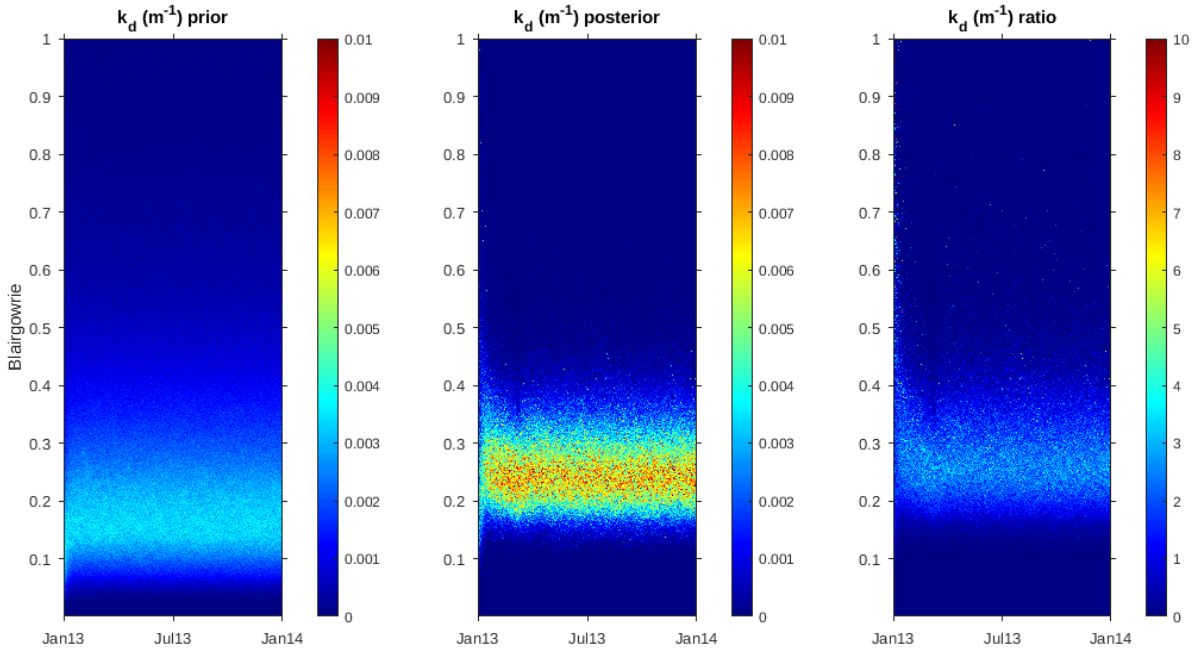

Figure 2.50 : Light attenuation coefficient (m^{-1}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Blairgowrie in 2013.

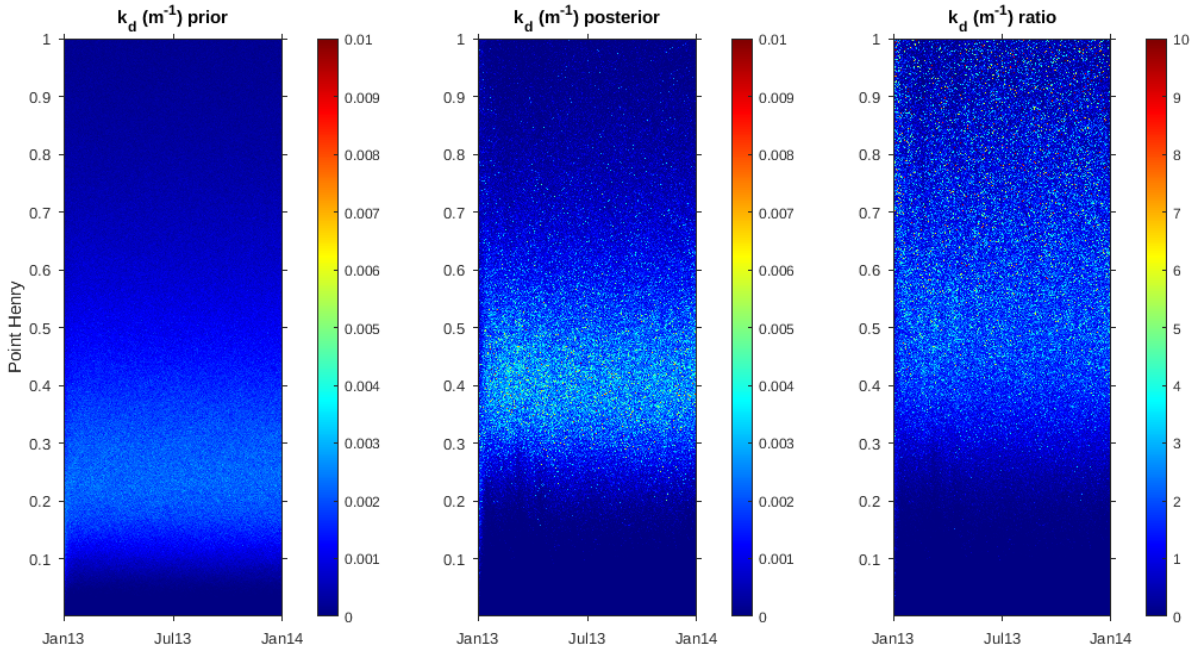

Figure 2.51 : Light attenuation coefficient (m^{-1}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Point Henry in 2013.

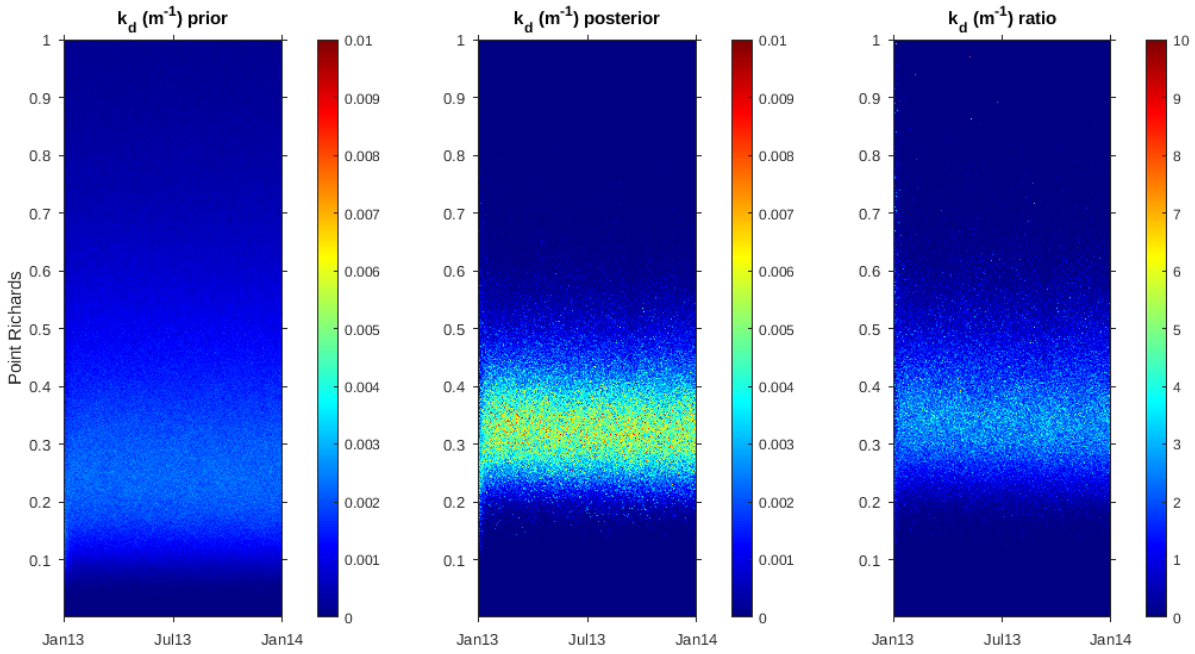

Figure 2.52 : Light attenuation coefficient (m^{-1}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Point Richards in 2013.

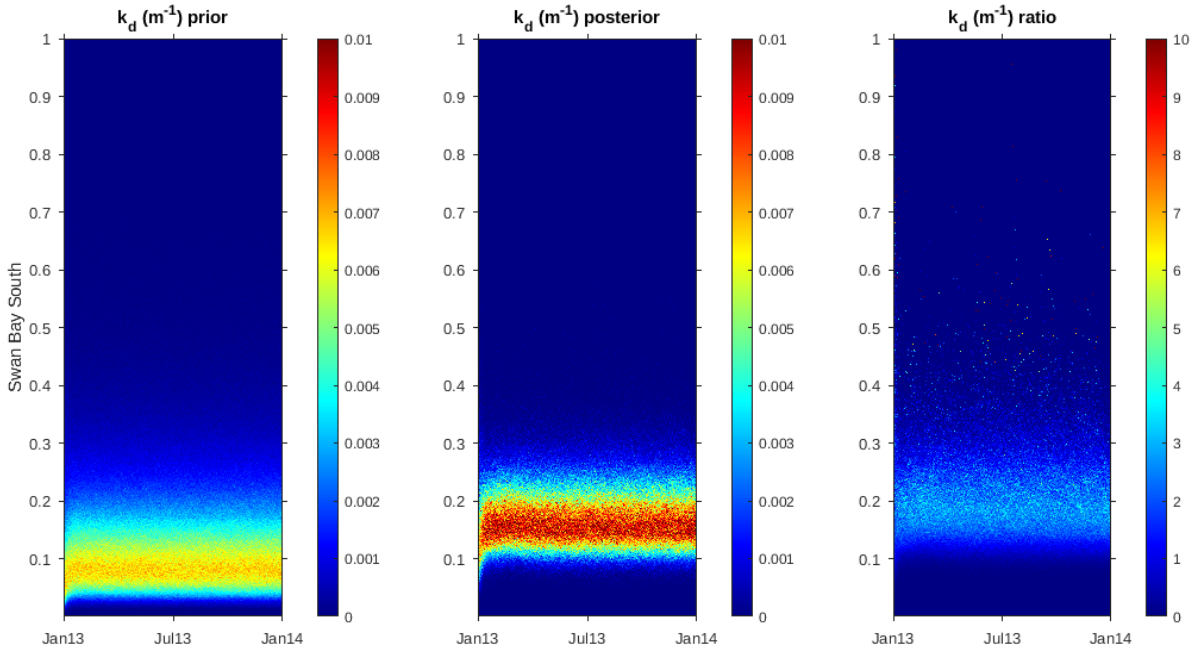

Figure 2.53 : Light attenuation coefficient (m^{-1}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Swan Bay South in 2013.

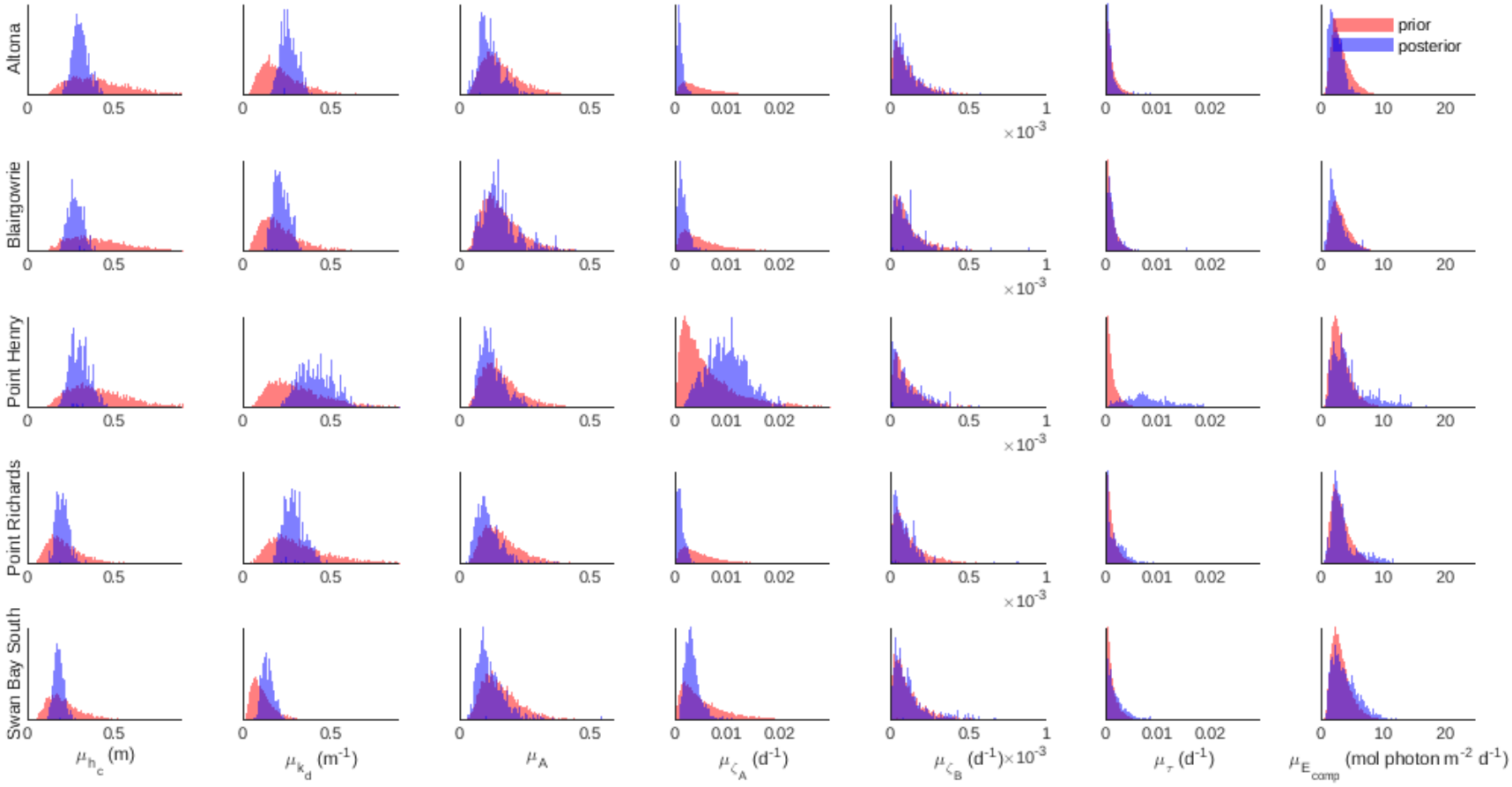

Figure 2.54 : Long-term means of the first order autoregressive parameter prior and posterior distributions for each Port Phillip Bay site in 2013.

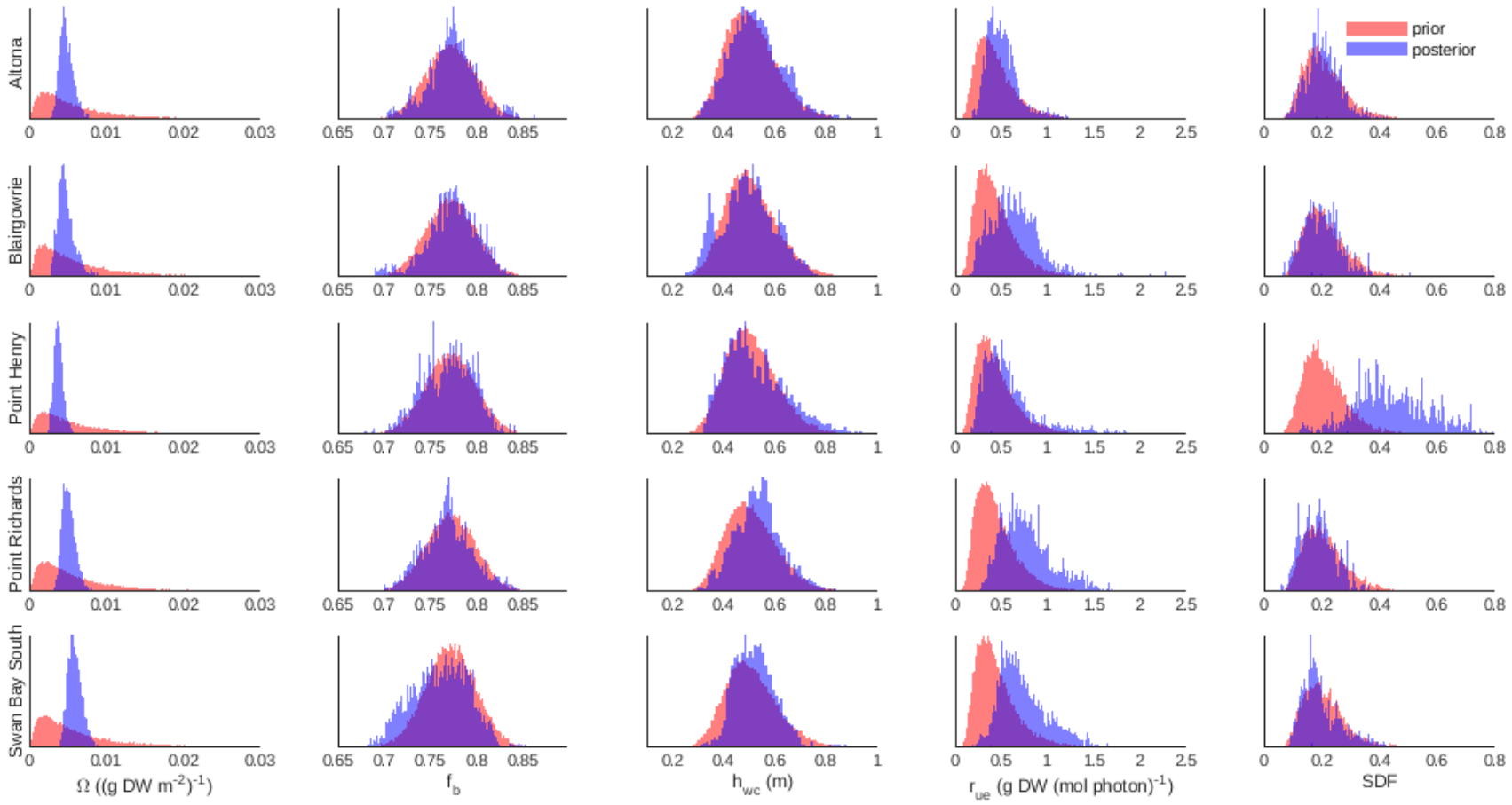

Figure 2.55 : Parameter prior and posterior distributions for each Port Phillip Bay site in 2013.

over 2013, Altona and Blairgowrie trended upwards, and Point Richards and Swan Bay South remained stable.

Heat-maps of fraction of canopy height density illustrated tightening of posterior density after assimilating observations at each site. There was 3-4 times the posterior density than the sampled prior at those points (Figure 2.44, Figure 2.45, Figure 2.46, Figure 2.47, Figure 2.48). Point Henry showed the largest spread in canopy height posterior over 2013 (Figure 2.46). At each site the long-term mean canopy height posterior IQR decreased to approximately a third of the prior IQR (Table 2.6). The posterior medians at Altona, Blairgowrie and Point Henry all shifted downwards from the prior median of 0.4000 m. Altona shifted downwards to 0.3011 m, Blairgowrie to 0.2741 m, and Point Henry to 0.2969 m. Point Richards posterior median 0.2023 m remained unchanged from the prior median of 0.2000 m. Swan Bay South posterior median of 0.1808 m shifted subtly downwards from its prior median of 0.2000 m. Altona and Point Henry had the tallest canopy height in 2013, followed by Blairgowrie, Point Richards, and Swan Bay South had the shortest canopy height (Figure 2.54).

Heat-maps of fraction of light attenuation coefficient density illustrated tightening of posterior density after assimilating observations at each site. There was 4-5 times the posterior density than the sampled prior at those points (Figure 2.49, Figure 2.50, Figure 2.51, Figure 2.52, Figure 2.53). Point Henry (Figure 2.51) had the largest spread of posterior density of all the sites. Point Henry had a posterior IQR half of the prior IQR, and was the site with the largest spread in long-term mean light attenuation coefficient over 2013. The long-term mean light attenuation coefficient posterior had smaller spread than the prior, with posterior IQR one third of the prior IQR for sites Altona, Blairgowrie, Point Richards and Swan Bay South (Figure 2.54). Altona and Blairgowrie saw shifts in posterior medians to $0.2608\,\mathrm{m}^{-1}$ and $0.2151\,\mathrm{m}^{-1}$ from a prior median of $0.2000\,\mathrm{m}^{-1}$. From a prior median

of $0.3000\,\mathrm{m}^{-1}$, Point Richard's posterior median shifted slightly to $0.2865\,\mathrm{m}^{-1}$, while Point Henry shifted further to $0.3990\,\mathrm{m}^{-1}$. Swan Bay South prior median $0.1000\,\mathrm{m}^{-1}$ shifted to $0.1417\,\mathrm{m}^{-1}$ which was the lowest light attenuation coefficient followed by Blairgowrie, Altona, Point Richards, with Point Henry as the highest site for 2013.

The same prior was assigned to effective absorbance at all sites with median of 0.1500 and IQR of 0.1000 and approximately the same amount of learning occurred at Altona, Point Henry, Point Richards and Swan Bay South, with posterior medians 0.1007-0.1177 and IQR 0.0603-0.0674. A subtle shift at Blairgowrie to a posterior median of 0.1408 and an equal amount of reduction in spread with an IQR 0.0768 (Table 2.6).

All sites moved away from the leaf mortality rate prior median of $0.5000\times10^{-2}\,\mathrm{d}^{-1}$ and IQR $0.7200\times10^{-2}\,\mathrm{d}^{-1}$. Point Henry had the highest leaf mortality rate posterior median of $0.9705\times10^{-2}\,\mathrm{d}^{-1}$ and the largest IQR $0.5293\times10^{-2}\,\mathrm{d}^{-1}$, followed by Swan Bay South with a leaf mortality rate posterior median of $0.3185\times10^{-2}\,\mathrm{d}^{-1}$ and IQR $0.1875\times10^{-2}\,\mathrm{d}^{-1}$. Altona and Point Richards had the lowest medians with $0.0996\times10^{-2}\,\mathrm{d}^{-1}$ and $0.1074\times10^{-2}\,\mathrm{d}^{-1}$, and also the smallest IQR ($0.0719\times10^{-2}\,\mathrm{d}^{-1}$ and $0.0927\times10^{-2}\,\mathrm{d}^{-1}$). Blairgowrie had a posterior median of $0.1455\times10^{-2}\,\mathrm{d}^{-1}$ and IQR $0.1207\times10^{-2}\,\mathrm{d}^{-1}$. There were no significant shifts from root/rhizome mortality rate priors to posteriors at all sites (Table 2.6).

Altona had the smallest translocation rate posterior median $0.0821\times10^{-2}\,\mathrm{d}^{-1}$ and smallest IQR $0.1076\times10^{-2}\,\mathrm{d}^{-1}$ of all the sites. Blairgowrie, Point Richards and Swan Bay South performed similarly across 2013 with posterior medians of 0.1324-$0.1721\times10^{-2}\,\mathrm{d}^{-1}$ and IQR 0.2001-$0.2758\times10^{-2}\,\mathrm{d}^{-1}$. Point Henry had the largest translocation rate posterior median of $0.8133\times10^{-2}\,\mathrm{d}^{-1}$, the largest shift upwards from prior median of $0.1000\times10^{-2}\,\mathrm{d}^{-1}$, and the largest IQR $0.7048\times10^{-2}\,\mathrm{d}^{-1}$, 5 times larger than the prior IQR $0.1400\times10^{-2}\,\mathrm{d}^{-1}$.

Compensation scalar irradiance priors and posteriors remained relatively unchanged at Point Richards. Altona and Blairgowrie posterior central tendency shifted from a prior of 3 mol photon m^{-2} d^{-1} to 2.3197 mol photon m^{-2} d^{-1} at Altona and 2.3721 mol photon m^{-2} d^{-1} at Blairgowrie with both sites decreasing their variability. Point Henry and Swan Bay South posterior central tendency shifted to 3.7530 mol photon m^{-2} d^{-1} at Point Henry and 3.6127 mol photon m^{-2} d^{-1} at Swan Bay South with both sites increasing in variability.

Dry weight specific area of seagrass was constrained well by the observations shown in the concentrated posterior distributions with lower variability (Figure 2.55). Swan Bay South had the highest posterior median (0.5828×10^{-2} (g DW $m^{-2})^{-1}$), followed by Point Richards (0.5018×10^{-2} (g DW $m^{-2})^{-1}$), Altona (0.4709×10^{-2} (g DW $m^{-2})^{-1}$), Blairgowrie (0.4593×10^{-2} (g DW $m^{-2})^{-1}$) and Point Henry (0.3725×10^{-2} (g DW $m^{-2})^{-1}$) had the lowest. All sites had posterior IQR 0.0827×10^{-2} - 0.1335×10^{-2} (g DW $m^{-2})^{-1}$ down from prior IQR 0.7300×10^{-2} (g DW $m^{-2})^{-1}$ (Table 2.6).

Equilibrium fraction of biomass below-ground posteriors remained unchanged from the prior at all sites. Water column height at Altona, Blairgowrie, Point Henry and Swan Bay South remained unchanged from prior to posterior. The posterior median at Point Richards displayed a subtle shift upwards to 0.5349 m.

Radiation use efficiency posterior distributions all shifted upwards from the prior distribution (median 0.4000 g DW (mol photon)$^{-1}$, IQR 0.2700 g DW (mol photon)$^{-1}$). Point Richards had the largest shift in posterior median (0.7376 g DW (mol photon)$^{-1}$), but also an increase in IQR (0.3652 g DW (mol photon)$^{-1}$). Swan Bay South had the next largest shift in posterior median to 0.6869 g DW (mol photon)$^{-1}$, followed by Blairgowrie (0.6292 g DW (mol photon)$^{-1}$), Point Henry (0.5136 g DW (mol photon)$^{-1}$), and Altona had the smallest (0.4856 g DW (mol photon)$^{-1}$). Altona had the smallest IQR (0.2133 g DW (mol photon)$^{-1}$), with Blair-

gowrie, Point Henry and Swan Bay South all between 0.3242 g DW (mol photon)$^{-1}$-0.3553 g DW (mol photon)$^{-1}$.

Seagrass diversity factor saw minimal change in central tendency and spread from prior to posterior at Altona, Blairgowrie, Point Richards and Swan Bay South. Point Henry posterior median doubled to 0.4250 from a prior median of 0.2002, and saw an increase in posterior IQR to 0.1659 from a prior of 0.0956 (Table 2.6).

2.3.4 Comparing the predicted 2013 results with the posterior 2013 results

The above-ground biomass posteriors captured the observations well across all five sites, with notable differences for Altona and Point Henry from their predictions (Figure 2.56). The Altona posterior had an upward shift from the 2013 prediction, while the Point Henry posterior had a downward shift from the prediction. Proportion cover posteriors performed well across Blairgowrie, Point Richards and Swan Bay South. Blairgowrie and Point Richards proportion cover posteriors started lower than predicted and ended higher than predicted. Swan Bay South's posterior state consistently estimated slightly lower than predicted. Point Henry overall estimated the proportion cover observations well apart from November 2013, with a large downward shift in values from the predicted estimates. Proportion cover posterior at Altona shifted upwards from the prediction, to capture many of the observations in the 95% credible intervals (Figure 2.57).

Canopy height posteriors performed well at all sites with most observations falling within the 95% credible intervals (Figure 2.58). They improved on the predictions at each site, with respect to capturing the central tendency and spread of the observations. Altona, Blairgowrie and Point Henry on average have higher canopy height estimates than Point Richards and Swan Bay South. Altona was the only site to have an increased canopy height long-term mean (Figure 2.60) from predicted

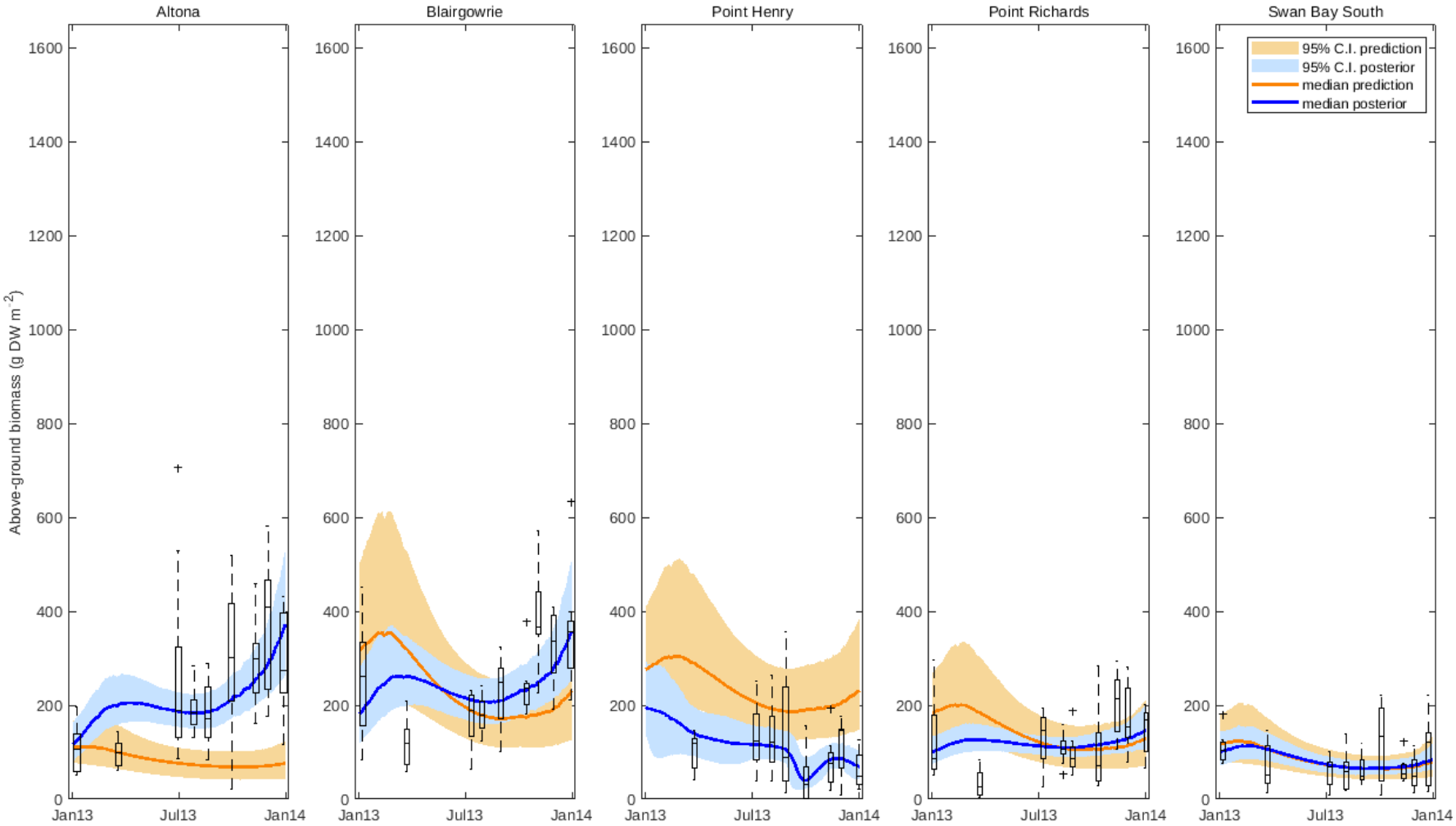

Figure 2.56 : Above-ground biomass predicted medians (solid orange line) and 95% credible intervals (shaded orange), posterior medians (solid blue line) and 95% credible intervals (shaded blue) for the second year of observations. Observations (n=9 at each sampling time-point) are shown as box-plots with potential outliers (+).

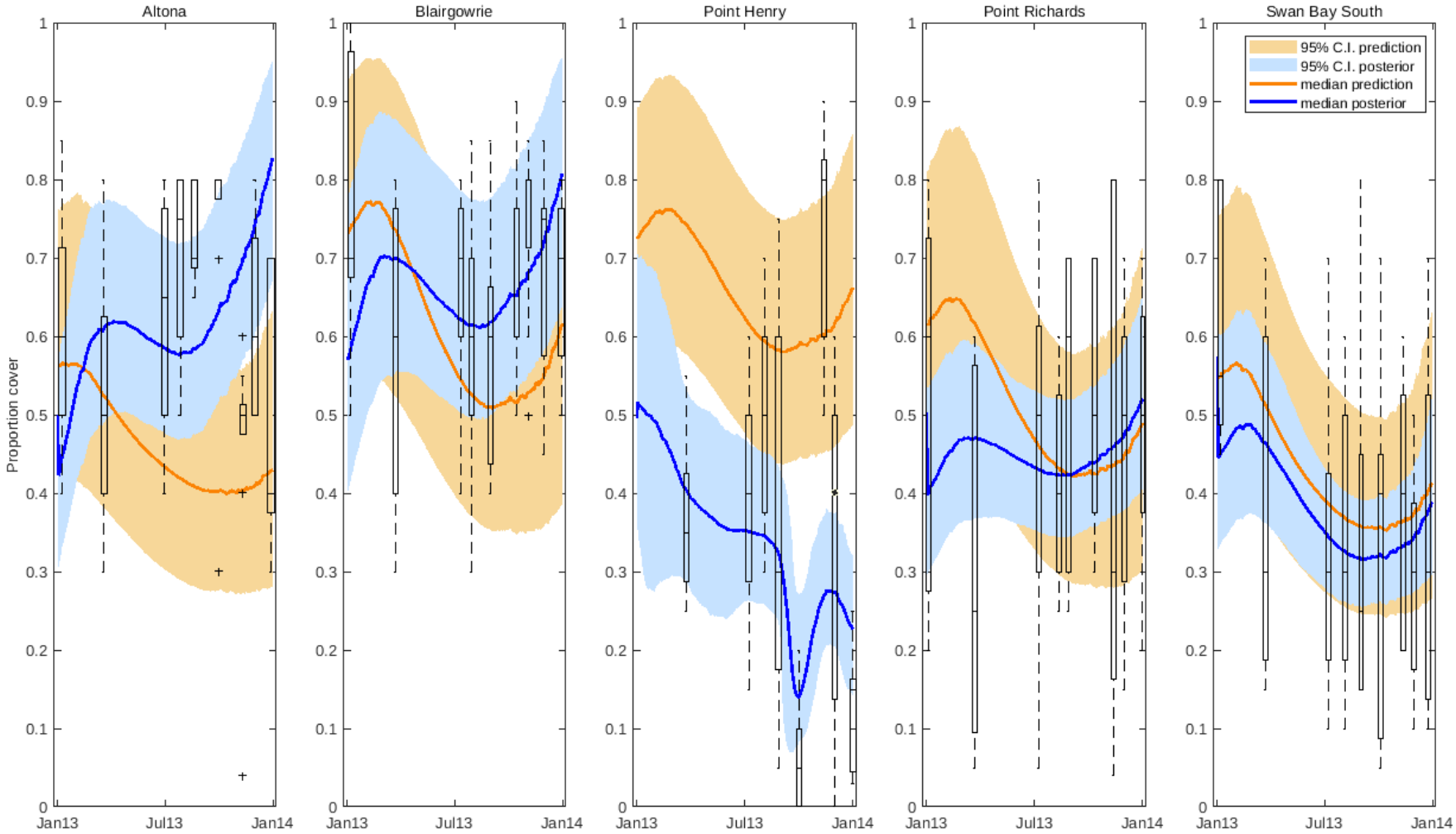

Figure 2.57 : Proportion cover predicted medians (solid orange line) and 95% credible intervals (shaded orange), posterior medians (solid blue line) and 95% credible intervals (shaded blue) for the second year of observations. Observations (n=9 at each sampling time-point) are shown as box-plots with potential outliers (+).

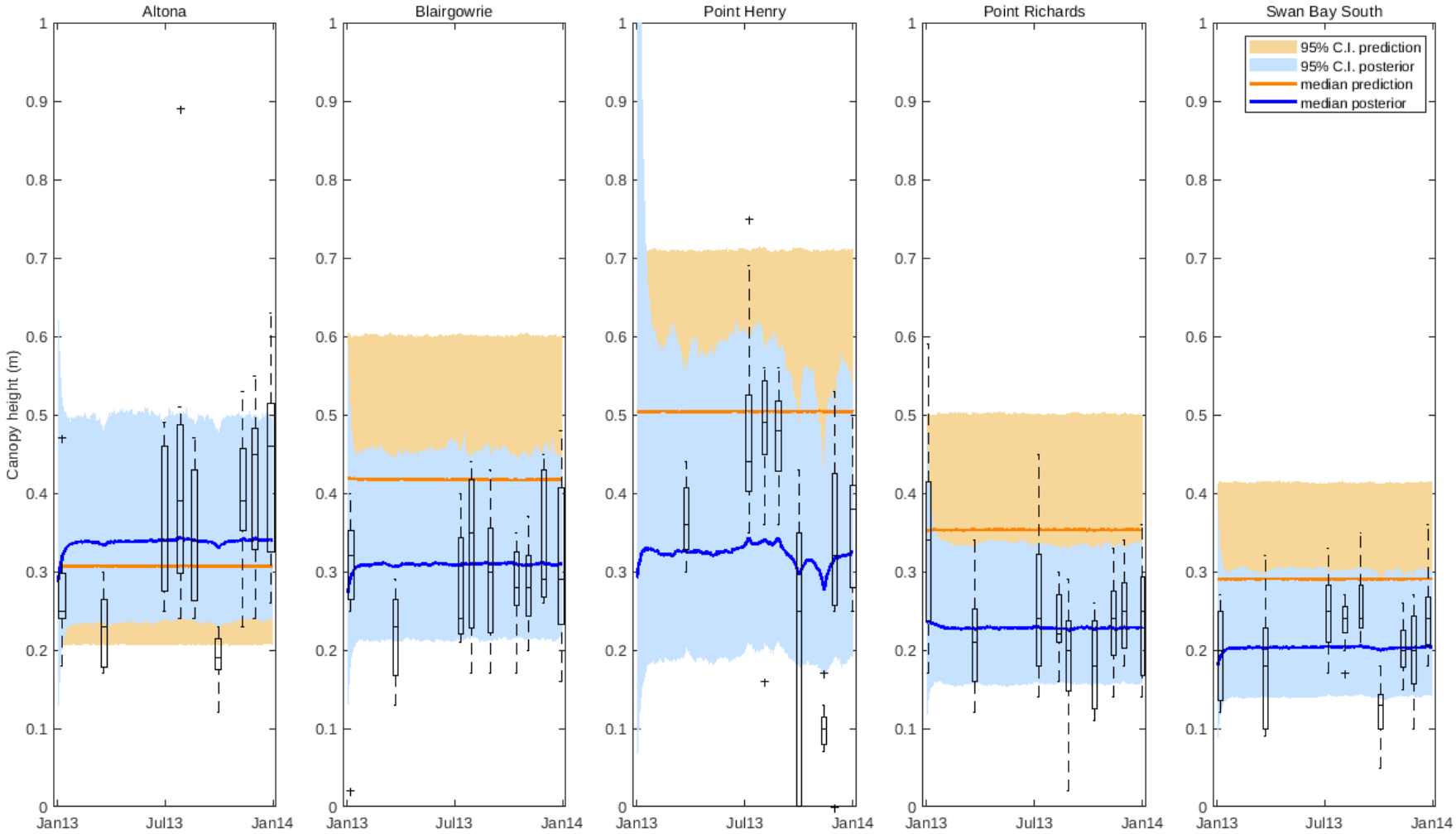

Figure 2.58 : Canopy height predicted medians (solid orange line) and 95% credible intervals (shaded orange), posterior medians (solid blue line) and 95% credible intervals (shaded blue) for the second year of observations. Observations (n=9 at each sampling time-point) are shown as box-plots with potential outliers (+).

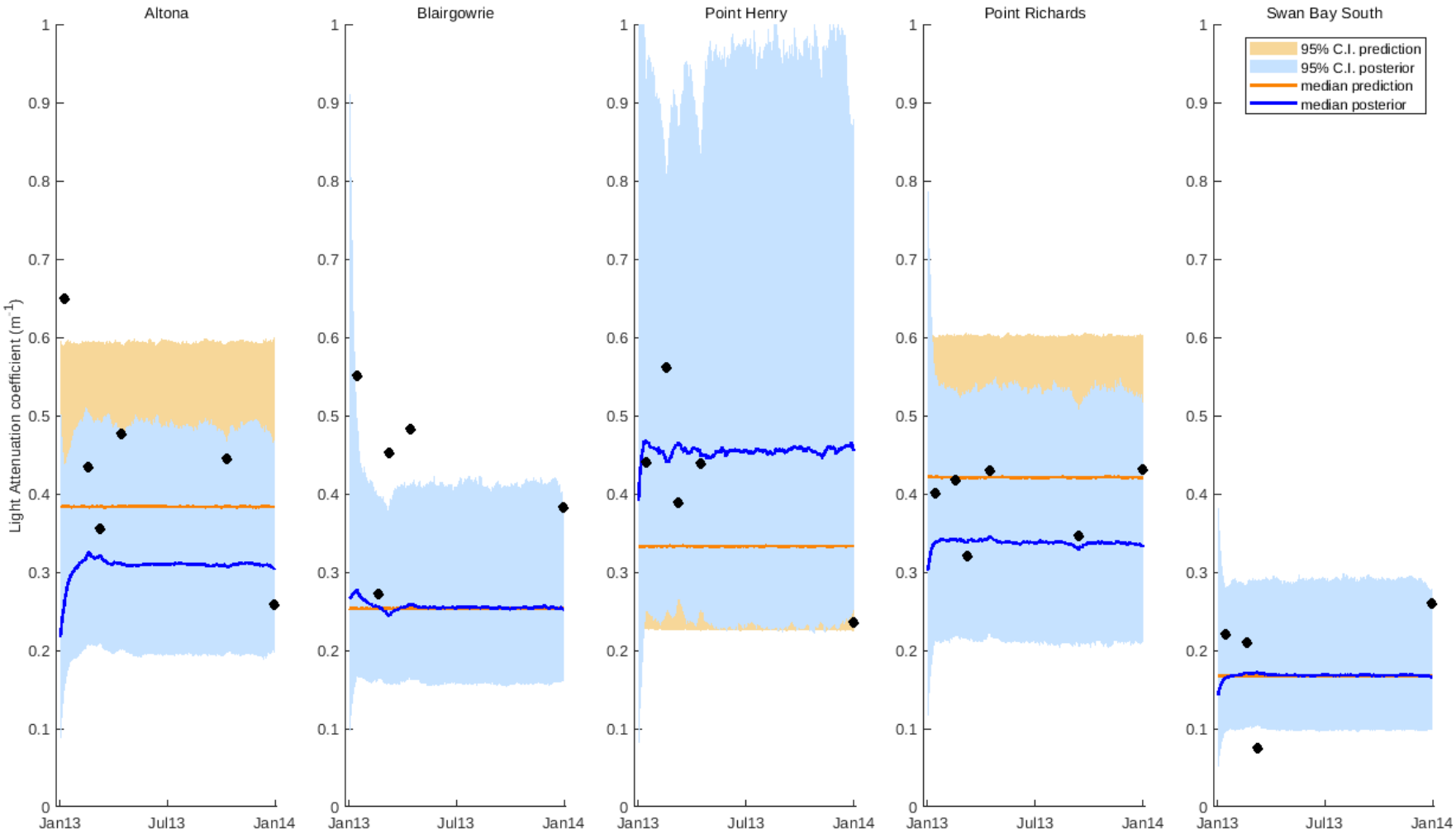

Figure 2.59 : Light attenuation coefficient predicted medians (solid orange line) and 95% credible intervals (shaded orange), posterior medians (solid blue line) and 95% credible intervals (shaded blue) for the second year of observations. Observations (n=1 at each sampling time-point) are shown as solid black circles.

0.2731 m (0.0532 m) to posterior 0.3011 m (0.0616 m). Altona had the highest posterior canopy height, along with Blairgowrie 0.2741 m (0.0619 m) and Point Henry 0.2969 m (0.0782 m) which also decreased from predicted 0.3707 m (0.0724 m) and 0.4463 m (0.0837 m), respectively. Point Richards and Swan Bay South posteriors grouped lower at 0.2023 m (0.0488 m) and 0.1808 m (0.0403 m) (Table 2.6).

Light attenuation coefficient posterior medians at Point Henry, Point Richards, and Swan Bay South captured the central tendency for 2013 well with the 95% credible intervals encompassing most observations. Altona and Blairgowrie posterior median estimates pulled towards the lower end of the observations, and while Blairgowrie remained unchanged from the prediction, the Altona posterior state has shifted downwards from the 2013 prediction and further from most of the observations (Figure 2.59). The light attenuation coefficient long-term mean posterior for Blairgowrie 0.2151 m^{-1} (0.0634 m^{-1}) and Swan Bay South 0.1417 m^{-1} (0.0471 m^{-1}) remained unchanged from the predicted 0.2149 m^{-1} (0.0417 m^{-1}) and 0.1409 m^{-1} (0.0282 m^{-1}) (Figure 2.61). Point Henry saw an increase in light attenuation coefficient from the predicted 0.2804 m^{-1} (0.0597 m^{-1}) to the posterior 0.3990 m^{-1} (0.1376 m^{-1}). Altona and Point Richards both saw a shift downwards from predicted 0.3238 m^{-1} (0.0640 m^{-1}) and 0.3545 m^{-1} (0.0652 m^{-1}) to posterior 0.2608 m^{-1} (0.0645 m^{-1}) and 0.2865 m^{-1} (0.0817 m^{-1}) (Table 2.6).

Effective absorbance long-term mean predictions were similar across site (Figure 2.62) with medians 0.0963-0.1066 and IQRs 0.0507-0.0609. Blairgowrie showed a shift from predicted 0.1066 (0.0507) to posterior 0.1408 (0.0768). The remaining four sites posteriors had no noticeable shifts from their predictions (Table 2.6).

Leaf mortality rate long-term mean posteriors for Altona 0.0996×10^{-2} d^{-1} (0.0719×10^{-2} d^{-1}), Blairgowrie 0.1455×10^{-2} d^{-1} (0.1207×10^{-2} d^{-1}) and Point Richards 0.1074×10^{-2} d^{-1} (0.0927×10^{-2} d^{-1}), all decreased from their predicted values 0.3787-0.4412×10^{-2} d^{-1}

(0.1964-0.2333$\times10^{-2}$ d^{-1}). Swan Bay South predicted central tendency and spread 0.4939$\times10^{-2}$ d^{-1} (0.3165$\times10^{-2}$ d^{-1}) that shifted downwards to a posterior estimate of 0.3185$\times10^{-2}$ d^{-1} (0.1875$\times10^{-2}$ d^{-1}). Point Henry had the lowest predicted leaf mortality rate of 0.2695$\times10^{-2}$ d^{-1} (0.1673$\times10^{-2}$ d^{-1}), but had the highest and largest spread of posterior distribution with median of 0.9705$\times10^{-2}$ d^{-1} and IQR 0.5293$\times10^{-2}$ d^{-1} (Figure 2.63).

Root/rhizome mortality rate long-term means remained unchanged from prediction to posterior and across site (Figure 2.64) distributed with medians of 0.0865-0.0961$\times10^{-3}$ d^{-1} and IQR 0.1051-0.1419$\times10^{-3}$ d^{-1}.

Point Henry had the lowest predicted translocation rate of 0.1923$\times10^{-2}$ d^{-1} (0.2587$\times10^{-2}$ d^{-1}) and the highest posterior translocation rate of 0.8133$\times10^{-2}$ d^{-1} (0.7048$\times10^{-2}$ d^{-1}) (Figure 2.65). Altona was predicted to have the highest translocation rate at 0.4599$\times10^{-2}$ d^{-1} (0.4133$\times10^{-2}$ d^{-1}), but had the lowest posterior translocation rate of 0.0821$\times10^{-2}$ d^{-1} (0.1076$\times10^{-2}$ d^{-1}). Blairgowrie, Point Richards and Swan Bay South had similar posterior estimates, all shifted to lower values than predicted (Table 2.6).

Altona and Swan Bay South were predicted to have the highest compensation scalar irradiance values and largest spread of 7.1438 mol photon m^{-2} d^{-1} (5.9472 mol photon m^{-2} d^{-1}) and 6.2655 mol photon m^{-2} d^{-1} (5.1140 mol photon m^{-2} d^{-1}), both posteriors notably shifted to lower values and smaller variability 2.3197 mol photon m^{-2} d^{-1} (1.4803 mol photon m^{-2} d^{-1}) and 3.6127 mol photon m^{-2} d^{-1} (2.9533 mol photon m^{-2} d^{-1}). Blairgowrie and Point Richards also shifted from predictions of 2.9868 mol photon m^{-2} d^{-1} (2.1179 mol photon m^{-2} d^{-1}) and 3.2284 mol photon m^{-2} d^{-1} (2.3697 mol photon m^{-2} d^{-1}) to slightly lower posteriors of 2.3721 mol photon m^{-2} d^{-1} (1.7478 mol photon m^{-2} d^{-1}) and 3.0427 mol photon m^{-2} d^{-1} (2.2337 mol photon m^{-2} d^{-1}). Point Henry had the only increase in central tendency, but also

in spread from a predicted 2.6726 mol photon m^{-2} d^{-1} (2.0854 mol photon m^{-2} d^{-1}) to 3.7530 mol photon m^{-2} d^{-1} (3.5701 mol photon m^{-2} d^{-1}) (Figure 2.66).

Altona predicted the highest dry weight specific area of seagrass (Figure 2.67) with median of 0.7494×10^{-2} (g DW $m^{-2})^{-1}$ and IQR (0.1883×10^{-2} (g DW $m^{-2})^{-1}$), followed by Swan Bay South 0.6739×10^{-2} (g DW $m^{-2})^{-1}$ (0.1613×10^{-2} (g DW $m^{-2})^{-1}$), Point Richards 0.5218×10^{-2} (g DW $m^{-2})^{-1}$ (0.1320×10^{-2} (g DW $m^{-2})^{-1}$), Point Henry 0.4660×10^{-2} (g DW $m^{-2})^{-1}$ (0.1192×10^{-2} (g DW $m^{-2})^{-1}$) and Blairgowrie with the lowest 0.4164×10^{-2} (g DW $m^{-2})^{-1}$ (0.1175×10^{-2} (g DW $m^{-2})^{-1}$). All posteriors experienced a shift downwards except for Blairgowrie which increased to 0.4593×10^{-2} (g DW $m^{-2})^{-1}$ (0.1335×10^{-2} (g DW $m^{-2})^{-1}$). Swan Bay South posterior had the largest dry weight specific area of seagrass with 0.5828×10^{-2} (g DW $m^{-2})^{-1}$ (0.1238×10^{-2} (g DW $m^{-2})^{-1}$), followed by Point Richards 0.5018×10^{-2} (g DW $m^{-2})^{-1}$ (0.1167×10^{-2} (g DW $m^{-2})^{-1}$), then Altona and Blairgowrie with 0.4709×10^{-2} (g DW $m^{-2})^{-1}$ (0.1226×10^{-2} (g DW $m^{-2})^{-1}$) and 0.4593×10^{-2} (g DW $m^{-2})^{-1}$ (0.1335 (g DW $m^{-2})^{-1}$), respectively.

All sites predicted similar values of equilibrium fraction of biomass below-ground (Figure 2.68) with medians ranging from 0.7648 to 0.7812 and IQR ranging from 0.0285 to 0.0343. There was no noticeable difference between the predictions and posteriors that had consistent medians of 0.7625-0.7741 and marginally larger IQRs 0.0314-0.0459 (Table 2.6). Swan Bay South predicted the lowest water column height (Figure 2.69) with 0.4626 m (0.1218 m) and Altona predicted the highest with 0.5453 m (0.1086 m). Point Henry posterior 0.5028 m (0.1669 m), remained relatively unchanged from the predicted 0.5039 m (0.1614 m). Altona and Blairgowrie posteriors estimated slightly lower than predicted, while Point Richards and Swan Bay South posteriors estimated slightly higher than predicted (Table 2.6).

Altona and Point Henry were predicted to have the highest measures of radi-

ation use efficiency, 0.6999 g DW (mol photon)$^{-1}$ (0.5124 g DW (mol photon)$^{-1}$) and 0.6393 g DW (mol photon)$^{-1}$ (0.3303 g DW (mol photon)$^{-1}$), with Blairgowrie, Point Richards and Swan Bay South predicting lower values of 0.4864 g DW (mol photon)$^{-1}$ (0.2838 g DW (mol photon)$^{-1}$), 0.5892 g DW (mol photon)$^{-1}$ (0.3419 g DW (mol photon)$^{-1}$) and 0.5462 g DW (mol photon)$^{-1}$ (0.3172 g DW (mol photon)$^{-1}$). In the posterior distributions, these two groups traded places with Altona and Point Henry at the lower end with medians and IQRs of 0.4856 g DW (mol photon)$^{-1}$ (0.2133 g DW (mol photon)$^{-1}$) and 0.5136 g DW (mol photon)$^{-1}$ (0.3242 g DW (mol photon)$^{-1}$), respectively and Blairgowrie, Point Richards and Swan Bay South on the higher end with 0.6292 g DW (mol photon)$^{-1}$ (0.3553 g DW (mol photon)$^{-1}$), 0.7376 g DW (mol photon)$^{-1}$ (0.3652 g DW (mol photon)$^{-1}$), and 0.6869 g DW (mol photon)$^{-1}$ (0.3473 g DW (mol photon)$^{-1}$) (Figure 2.70, Table 2.6). The only noticeable shift from predicted to posterior values of seagrass diversity factor was at Point Henry (Figure 2.71) where the median and IQR shifted from 0.1665 (0.0737) to 0.4250 (0.1659) (Table 2.6).

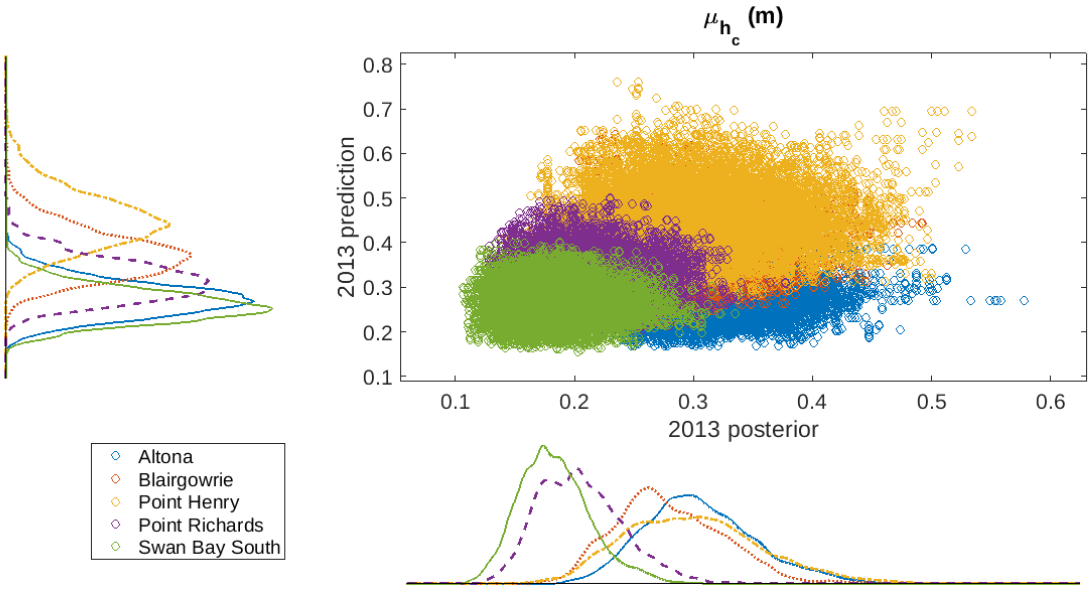

Figure 2.60 : Canopy height long-term mean (μ_{h_c}) predicted vs posterior parameter kernel density estimates for the five Port Phillip Bay sites.

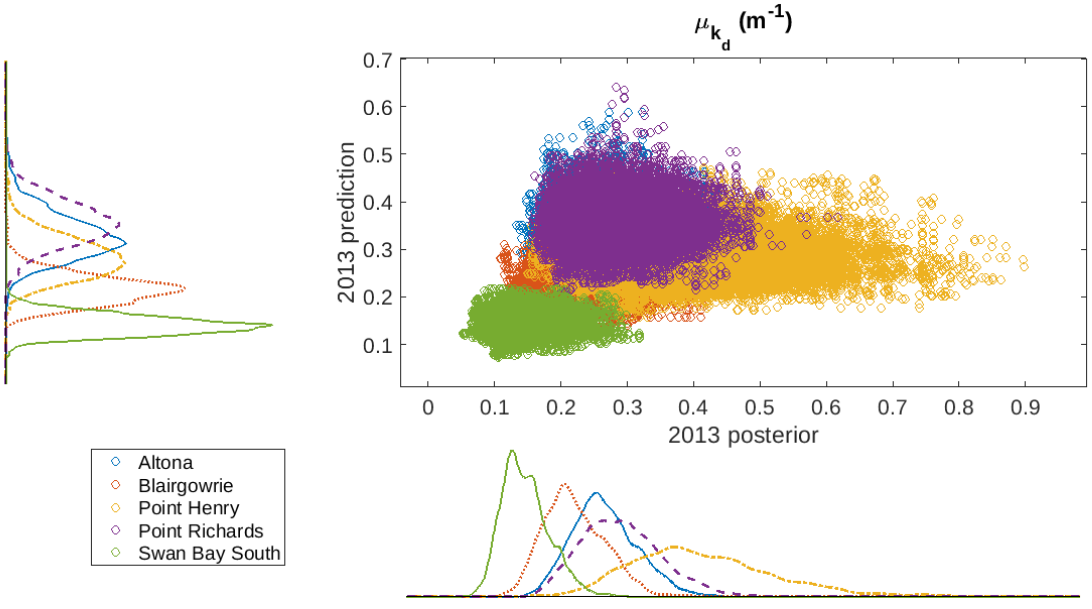

Figure 2.61 : Light attenuation coefficient long-term mean (μ_{k_d}) predicted vs posterior parameter kernel density estimates for the five Port Phillip Bay sites.

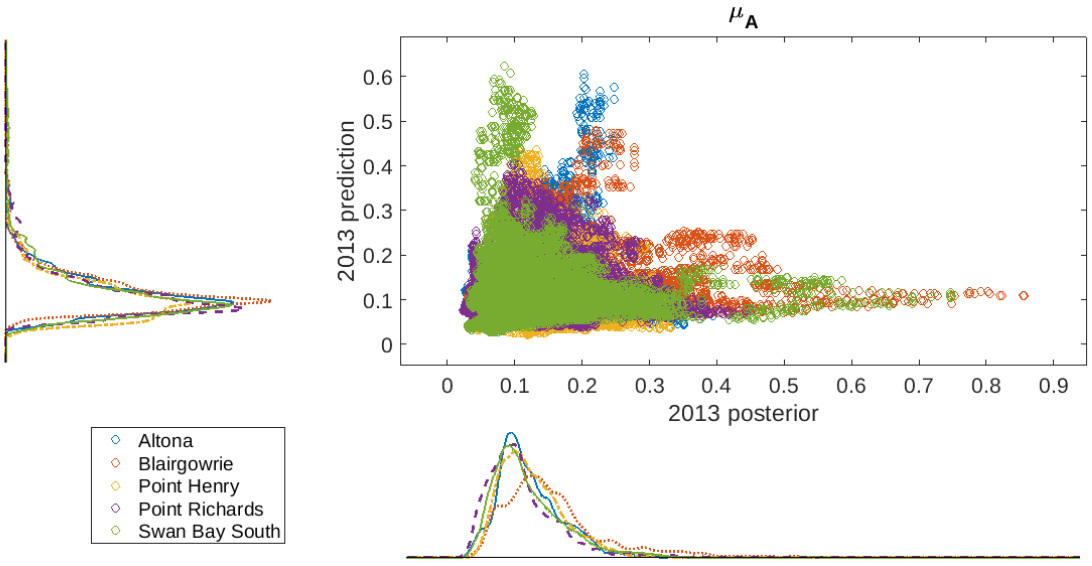

Figure 2.62 : Effective absorbance long-term mean (μ_A) predicted vs posterior parameter kernel density estimates for the five Port Phillip Bay sites.

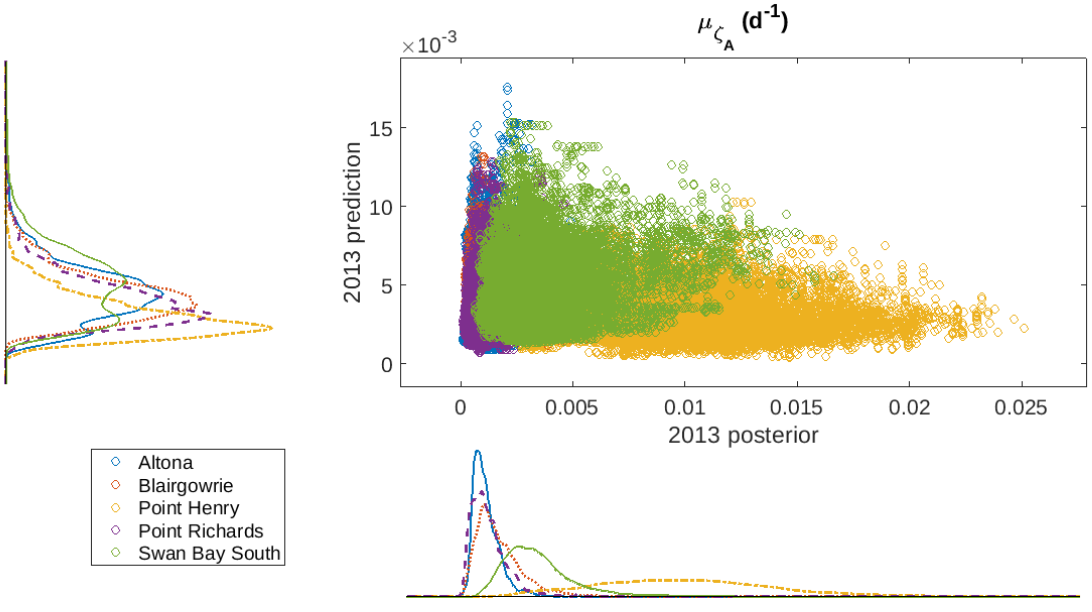

Figure 2.63 : Leaf mortality rate long-term mean (μ_{ζ_A}) predicted vs posterior parameter kernel density estimates for the five Port Phillip Bay sites.

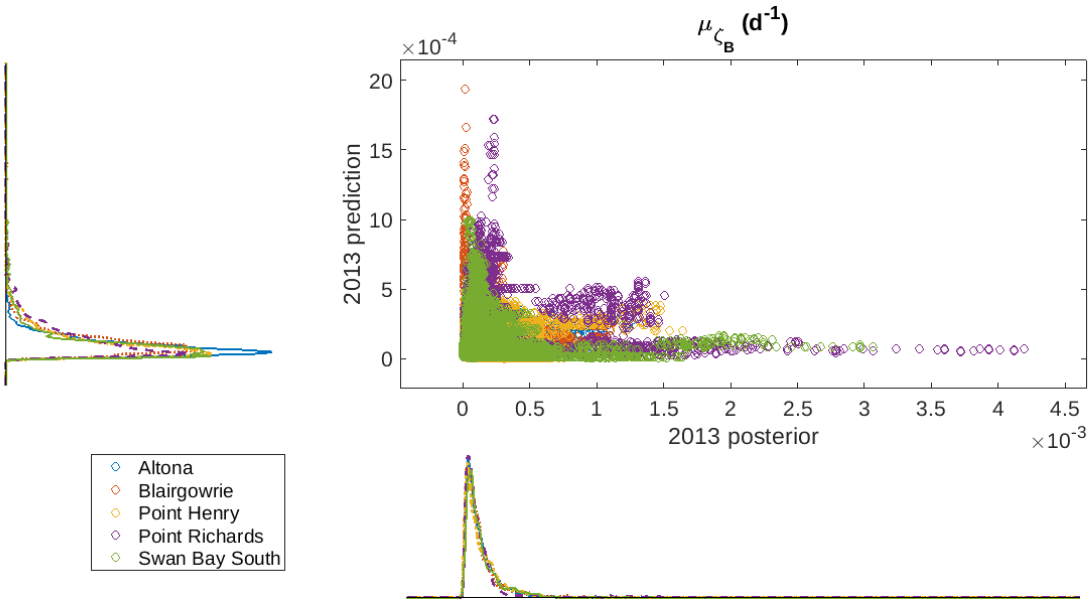

Figure 2.64 : Root/rhizome mortality rate long-term mean (μ_{ζ_B}) predicted vs posterior parameter kernel density estimates for the five Port Phillip Bay sites.

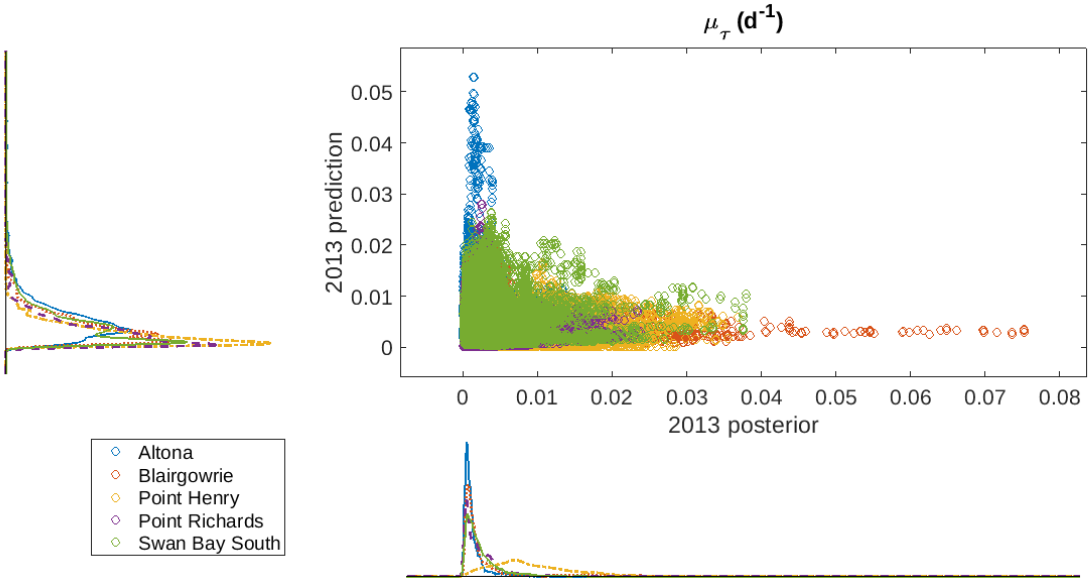

Figure 2.65 : Translocation rate long-term mean (μ_τ) predicted vs posterior parameter kernel density estimates for the five Port Phillip Bay sites.

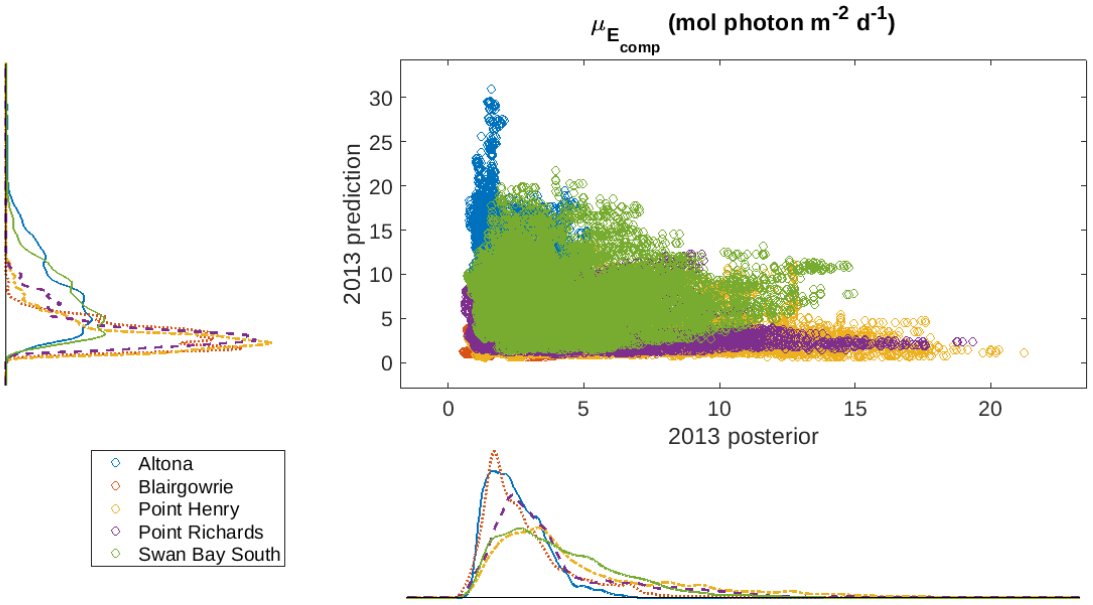

Figure 2.66 : Compensation scalar irradiance long-term mean ($\mu_{\mathrm{E}_{comp}}$) predicted vs posterior parameter kernel density estimates for the five Port Phillip Bay sites.

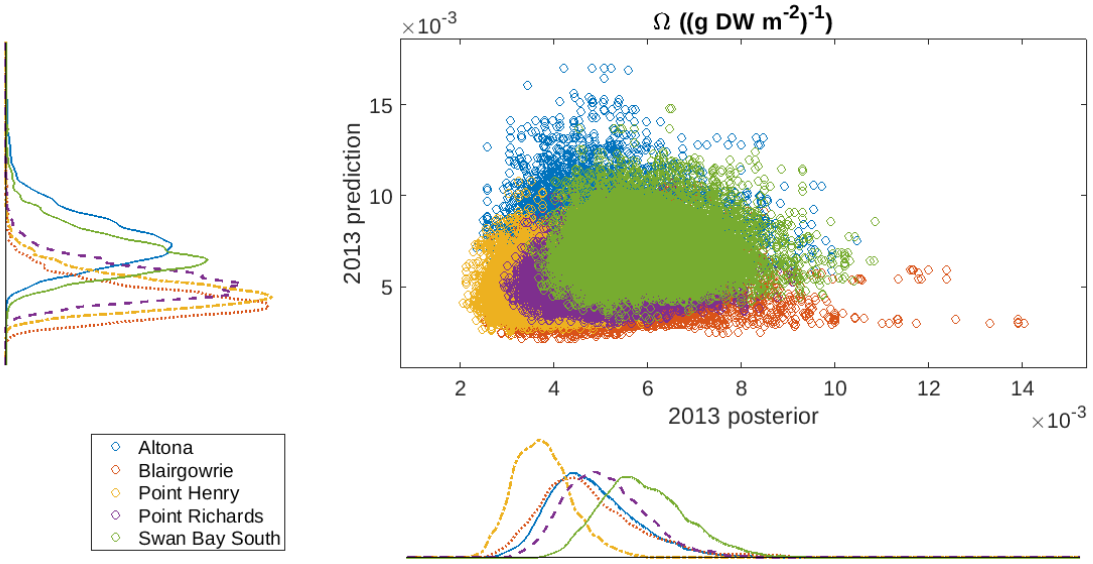

Figure 2.67 : Dry weight specific area of seagrass (Ω) predicted vs posterior parameter kernel density estimates for the five Port Phillip Bay sites.

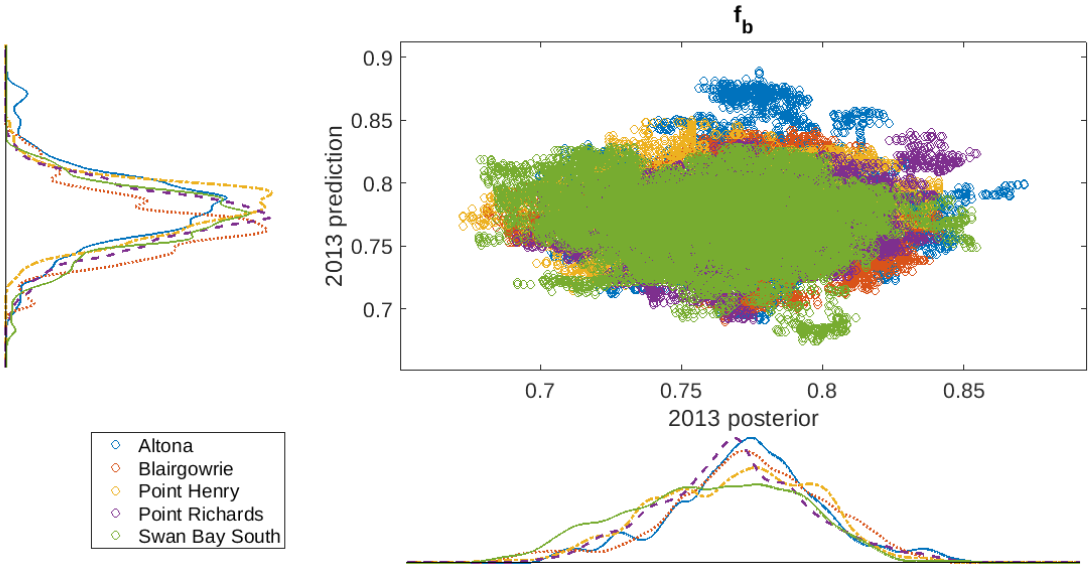

Figure 2.68 : Equilibrium fraction of biomass below-ground (f_b) predicted vs posterior parameter kernel density estimates for the five Port Phillip Bay sites.

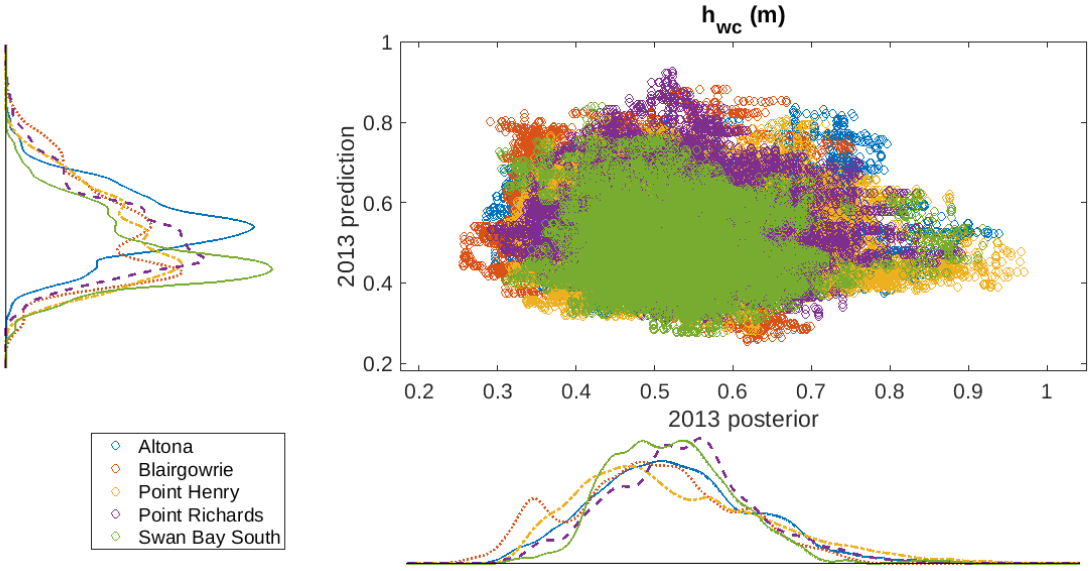

Figure 2.69 : Water column height (h_{wc}) predicted vs posterior parameter kernel density estimates for the five Port Phillip Bay sites.

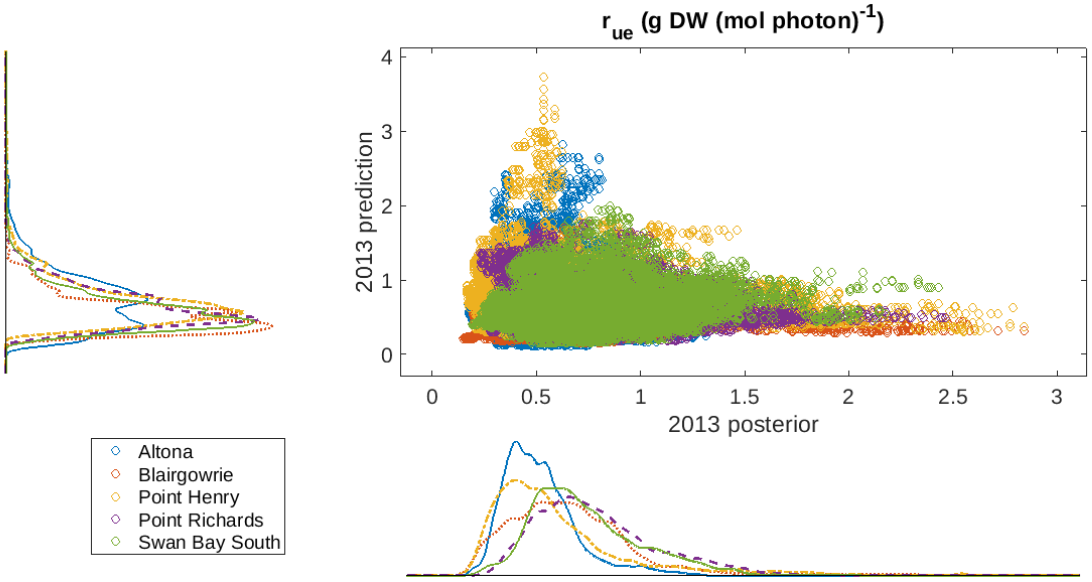

Figure 2.70 : Radiation use efficiency (r_{ue}) predicted vs posterior parameter kernel density estimates for the five Port Phillip Bay sites.

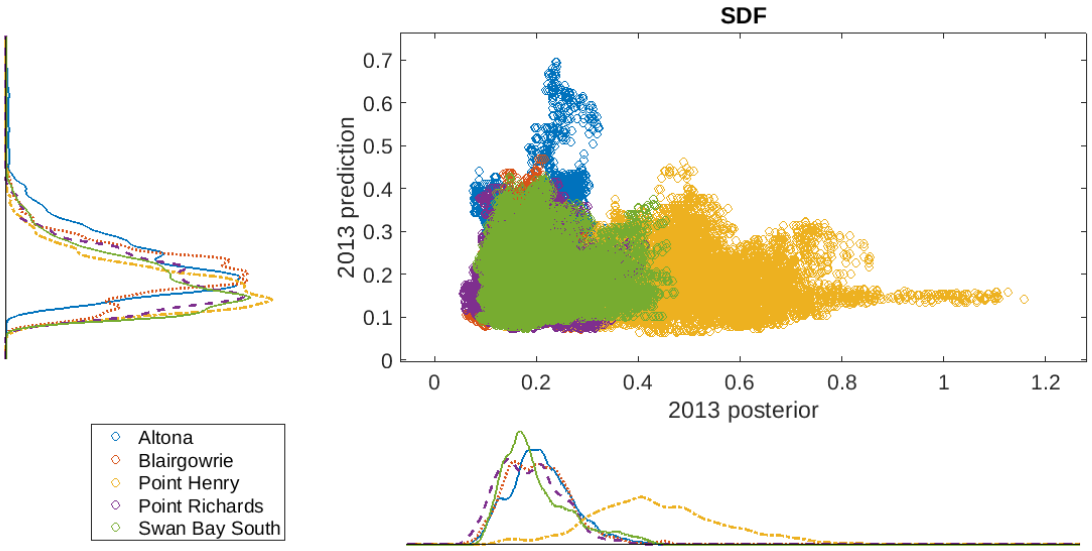

Figure 2.71 : Seagrass diversity factor (SDF) predicted vs posterior parameter kernel density estimates for the five Port Phillip Bay sites.

Parameter	Site	Prior 2013	Prediction 2013	Posterior 2013
μ_{h_c}	Altona	0.4002 (0.2754)	0.2731 (0.0532)	0.3011 (0.0616)
	Blairgowrie	0.3982 (0.2759)	0.3707 (0.0724)	0.2741 (0.0619)
	Point Henry	0.4007 (0.2760)	0.4463 (0.0837)	0.2969 (0.0782)
	Point Richards	0.1999 (0.1360)	0.3137 (0.0637)	0.2023 (0.0488)
	Swan Bay South	0.2003 (0.1377)	0.2579 (0.0511)	0.1808 (0.0403)
μ_{k_d}	Altona	0.2009 (0.1662)	0.3238 (0.0640)	0.2608 (0.0645)
	Blairgowrie	0.1999 (0.1653)	0.2149 (0.0417)	0.2151 (0.0634)
	Point Henry	0.3000 (0.2481)	0.2804 (0.0597)	0.3990 (0.1376)
	Point Richards	0.2991 (0.2489)	0.3545 (0.0652)	0.2865 (0.0817)
	Swan Bay South	0.1003 (0.0832)	0.1409 (0.0282)	0.1417 (0.0471)
μ_{A}	Altona	0.1500 (0.1024)	0.1036 (0.0563)	0.1116 (0.0603)
	Blairgowrie	0.1509 (0.1028)	0.1066 (0.0507)	0.1408 (0.0768)
	Point Henry	0.1496 (0.1026)	0.0963 (0.0609)	0.1177 (0.0622)
	Point Richards	0.1507 (0.1029)	0.0992 (0.0571)	0.1007 (0.0603)
	Swan Bay South	0.1493 (0.1009)	0.0999 (0.0592)	0.1073 (0.0674)
μ_{ζ_A}	Altona	0.5024×10^{-2} (0.7292×10^{-2})	0.4412×10^{-2} (0.2333×10^{-2})	0.0996×10^{-2} (0.0719×10^{-2})

	Blairgowrie	0.5036×10^{-2} (0.7245×10^{-2})	0.4136×10^{-2} (0.1964×10^{-2})	0.1455×10^{-2} (0.1207×10^{-2})
	Point Henry	0.5028×10^{-2} (0.7237×10^{-2})	0.2695×10^{-2} (0.1673×10^{-2})	0.9705×10^{-2} (0.5293×10^{-2})
	Point Richards	0.4973×10^{-2} (0.7207×10^{-2})	0.3787×10^{-2} (0.2057×10^{-2})	0.1074×10^{-2} (0.0927×10^{-2})
	Swan Bay South	0.5001×10^{-2} (0.7229×10^{-2})	0.4939×10^{-2} (0.3165×10^{-2})	0.3185×10^{-2} (0.1875×10^{-2})
μ_{ζ_B}	Altona	0.1011×10^{-3} (0.1462×10^{-3})	0.0760×10^{-3} (0.0732×10^{-3})	0.0940×10^{-3} (0.1273×10^{-3})
	Blairgowrie	0.1000×10^{-3} (0.1435×10^{-3})	0.0992×10^{-3} (0.0950×10^{-3})	0.0915×10^{-3} (0.1151×10^{-3})
	Point Henry	0.0996×10^{-3} (0.1457×10^{-3})	0.0823×10^{-3} (0.1110×10^{-3})	0.0925×10^{-3} (0.1419×10^{-3})
	Point Richards	0.0996×10^{-3} (0.1438×10^{-3})	0.1087×10^{-3} (0.1556×10^{-3})	0.0865×10^{-3} (0.1051×10^{-3})
	Swan Bay South	0.0992×10^{-3} (0.1438×10^{-3})	0.0820×10^{-3} (0.0920×10^{-3})	0.0961×10^{-3} (0.1340×10^{-3})
μ_τ	Altona	0.1002×10^{-2} (0.1371×10^{-2})	0.4599×10^{-2} (0.4133×10^{-2})	0.0821×10^{-2} (0.1076×10^{-2})
	Blairgowrie	0.0995×10^{-2} (0.1366×10^{-2})	0.3220×10^{-2} (0.3785×10^{-2})	0.1324×10^{-2} (0.2001×10^{-2})
	Point Henry	0.0999×10^{-2} (0.1378×10^{-2})	0.1923×10^{-2} (0.2587×10^{-2})	0.8133×10^{-2} (0.7048×10^{-2})
	Point Richards	0.1006×10^{-2} (0.1382×10^{-2})	0.2408×10^{-2} (0.3201×10^{-2})	0.1696×10^{-2} (0.2758×10^{-2})
	Swan Bay South	0.0996×10^{-2} (0.1379×10^{-2})	0.3647×10^{-2} (0.4411×10^{-2})	0.1721×10^{-2} (0.2624×10^{-2})
$\mu_{\mathrm{E}_{comp}}$	Altona	2.9922 (2.0758)	7.1438 (5.9472)	2.3197 (1.4803)
	Blairgowrie	3.0120 (2.0741)	2.9868 (2.1179)	2.3721 (1.7478)
	Point Henry	2.9900 (2.0649)	2.6726 (2.0854)	3.7530 (3.5701)

	Point Richards	2.9958 (2.0671)	3.2284 (2.3697)	3.0427 (2.2337)
	Swan Bay South	2.9888 (2.0500)	6.2655 (5.1140)	3.6127 (2.9533)
Ω	Altona	0.5040×10^{-2} (0.7361×10^{-2})	0.7494×10^{-2} (0.1883×10^{-2})	0.4709×10^{-2} (0.1226×10^{-2})
	Blairgowrie	0.4993×10^{-2} (0.7222×10^{-2})	0.4164×10^{-2} (0.1175×10^{-2})	0.4593×10^{-2} (0.1335×10^{-2})
	Point Henry	0.4976×10^{-2} (0.7261×10^{-2})	0.4660×10^{-2} (0.1192×10^{-2})	0.3725×10^{-2} (0.0827×10^{-2})
	Point Richards	0.5064×10^{-2} (0.7257×10^{-2})	0.5218×10^{-2} (0.1320×10^{-2})	0.5018×10^{-2} (0.1167×10^{-2})
	Swan Bay South	0.5006×10^{-2} (0.7191×10^{-2})	0.6739×10^{-2} (0.1613×10^{-2})	0.5828×10^{-2} (0.1238×10^{-2})
f_b	Altona	0.7720 (0.0371)	0.7812 (0.0343)	0.7741 (0.0314)
	Blairgowrie	0.7721 (0.0366)	0.7648 (0.0323)	0.7739 (0.0347)
	Point Henry	0.7723 (0.0369)	0.7805 (0.0285)	0.7712 (0.0419)
	Point Richards	0.7724 (0.0368)	0.7729 (0.0298)	0.7685 (0.0346)
	Swan Bay South	0.7723 (0.0370)	0.7733 (0.0321)	0.7625 (0.0459)
h_{wc}	Altona	0.4989 (0.1344)	0.5453 (0.1086)	0.5191 (0.1353)
	Blairgowrie	0.4993 (0.1346)	0.5277 (0.1767)	0.4941 (0.1398)
	Point Henry	0.4997 (0.1350)	0.5039 (0.1614)	0.5028 (0.1669)
	Point Richards	0.4998 (0.1362)	0.5171 (0.1355)	0.5349 (0.1165)
	Swan Bay South	0.4991 (0.1347)	0.4626 (0.1218)	0.5178 (0.1089)

r_{ue}	Altona	0.4016 (0.2740)	0.6999 (0.5124)	0.4856 (0.2133)
	Blairgowrie	0.4000 (0.2731)	0.4864 (0.2838)	0.6292 (0.3553)
	Point Henry	0.3983 (0.2761)	0.6393 (0.3303)	0.5136 (0.3242)
	Point Richards	0.4010 (0.2748)	0.5892 (0.3419)	0.7376 (0.3652)
	Swan Bay South	0.4004 (0.2753)	0.5462 (0.3172)	0.6869 (0.3473)
SDF	Altona	0.1998 (0.0957)	0.2138 (0.0959)	0.2048 (0.0778)
	Blairgowrie	0.2003 (0.0954)	0.2028 (0.0754)	0.2000 (0.0858)
	Point Henry	0.2002 (0.0956)	0.1665 (0.0737)	0.4250 (0.1659)
	Point Richards	0.2002 (0.0954)	0.1804 (0.0891)	0.1835 (0.0885)
	Swan Bay South	0.1990 (0.0952)	0.1720 (0.0903)	0.1824 (0.0861)

Table 2.6 : Parameter median and interquartile range (IQR) at each site for 2013 prior, prediction and posterior.

2.3.5 Posterior results across 2012-2013

The above-ground biomass posterior estimates for 2012-2013 assimilated the trend of the observations well (Figure 2.72). Blairgowrie has the highest initial above-ground biomass in October 2011, followed by Point Richards and Point Henry, Swan Bay South then Altona. All sites exhibited seasonality with above-ground biomass peaks in the summer months. Altona was the only site that displayed an upward trend across 2012-2013, the remaining four sites all exhibited seasonality with a downward trend by January 2014. Above-ground biomass fraction of density heat-maps illustrated tightening of posterior density after assimilating observations at each site. There was 15-25 times the posterior density than the sampled prior at those points (Figure 2.77, Figure 2.78, Figure 2.79, Figure 2.80, Figure 2.81). Blairgowrie, Point Henry, Point Richards and Swan Bay South posteriors all trend downwards over 2012-2013. Altona was the only site with above-ground biomass trending upwards over 2012-2013 (Figure 2.77).

Proportion cover posterior estimates at each site aligned reasonably well with the observations for 2012-2013, given the large variability at some time-points (Figure 2.73). Blairgowrie has the largest initial proportion cover followed by Point Richards, Swan Bay South, Point Henry then Altona. Seasonality in the form of proportion cover peaks during summer were present at all sites with Blairgowrie, Point Henry, Point Richards and Swan Bay South trending downwards during 2012-2013. Similarly to above-ground biomass, proportion cover at Altona is trending upwards during 2012-2013. Proportion cover fraction of density heat-maps showed the prior sampled almost the entire space between 0 and 1 with higher densities at either complete cover during all of 2012-2013 or declining to zero during 2013. The heat-maps illustrated tightening of posterior density after assimilating observations at each site. There was 20-30 times the posterior density than the sampled prior at those points (Figure 2.82, Figure 2.83, Figure 2.84, Figure 2.85, Figure 2.86). Blair-

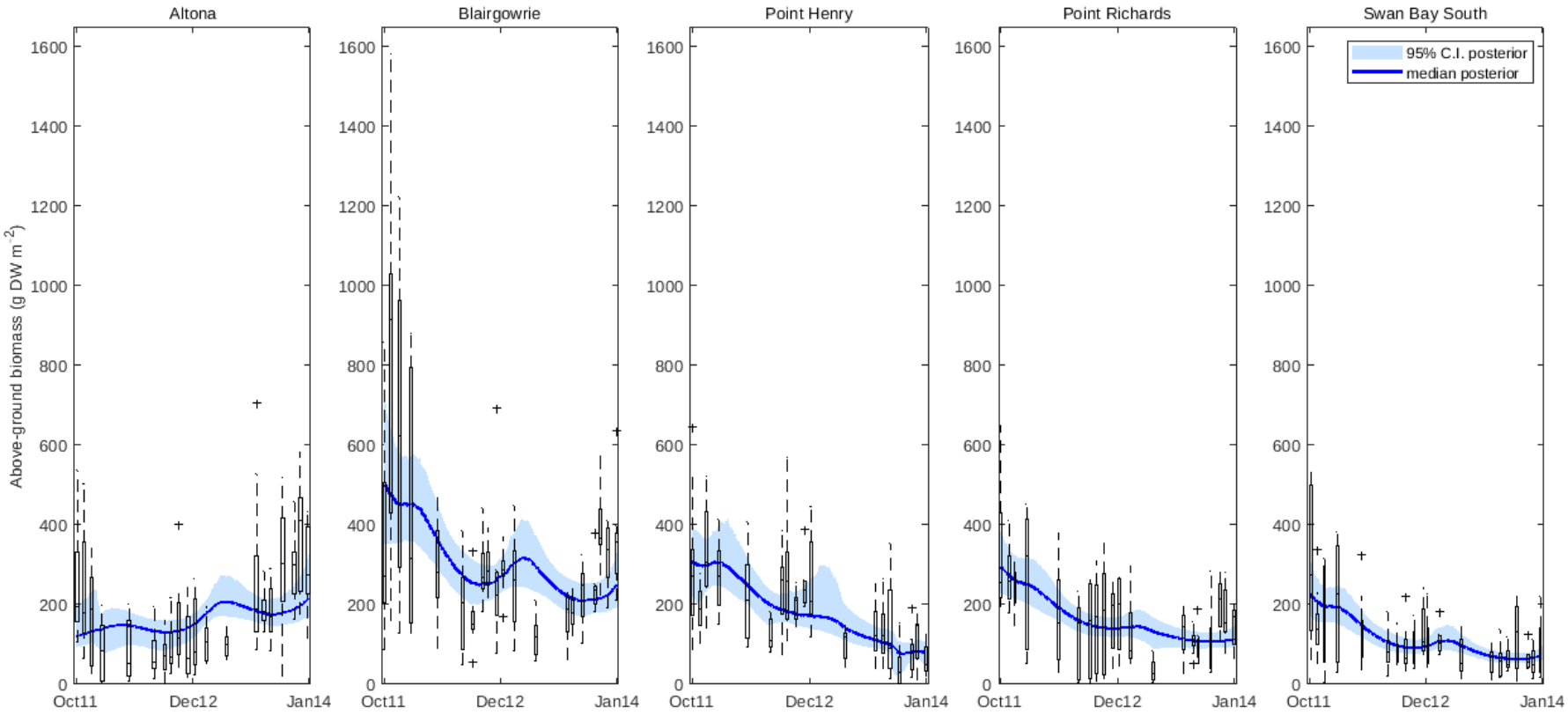

Figure 2.72 : Above-ground biomass (g DW m^{-2}) posterior medians (solid blue line) and 95% credible intervals (shaded blue) for the full set of observations. Observations (n=9 at each sampling time-point) are shown as box-plots with potential outliers (+).

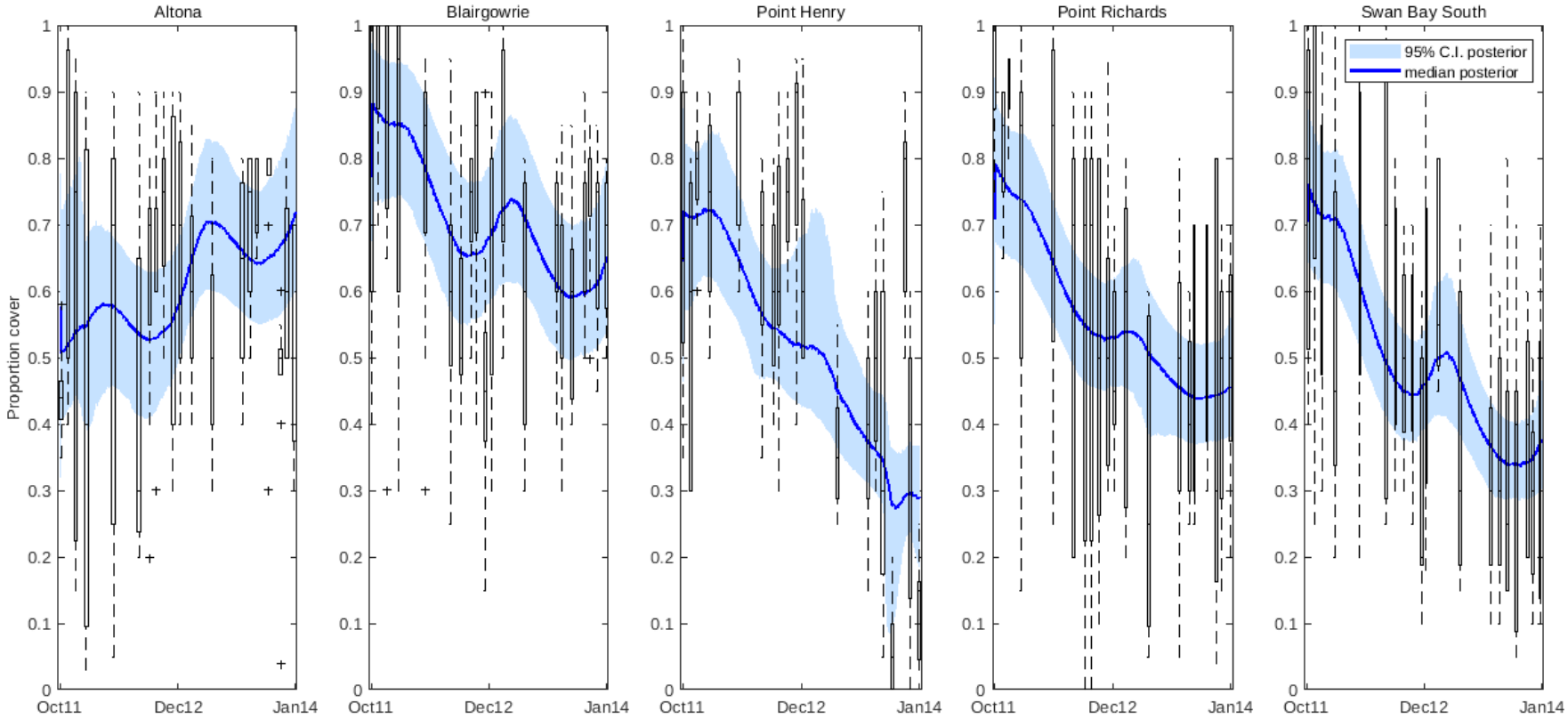

Figure 2.73 : Proportion cover posterior medians (solid blue line) and 95% credible intervals (shaded blue) for the full set of observations. Observations (n=9 at each sampling time-point) are shown as box-plots with potential outliers (+).

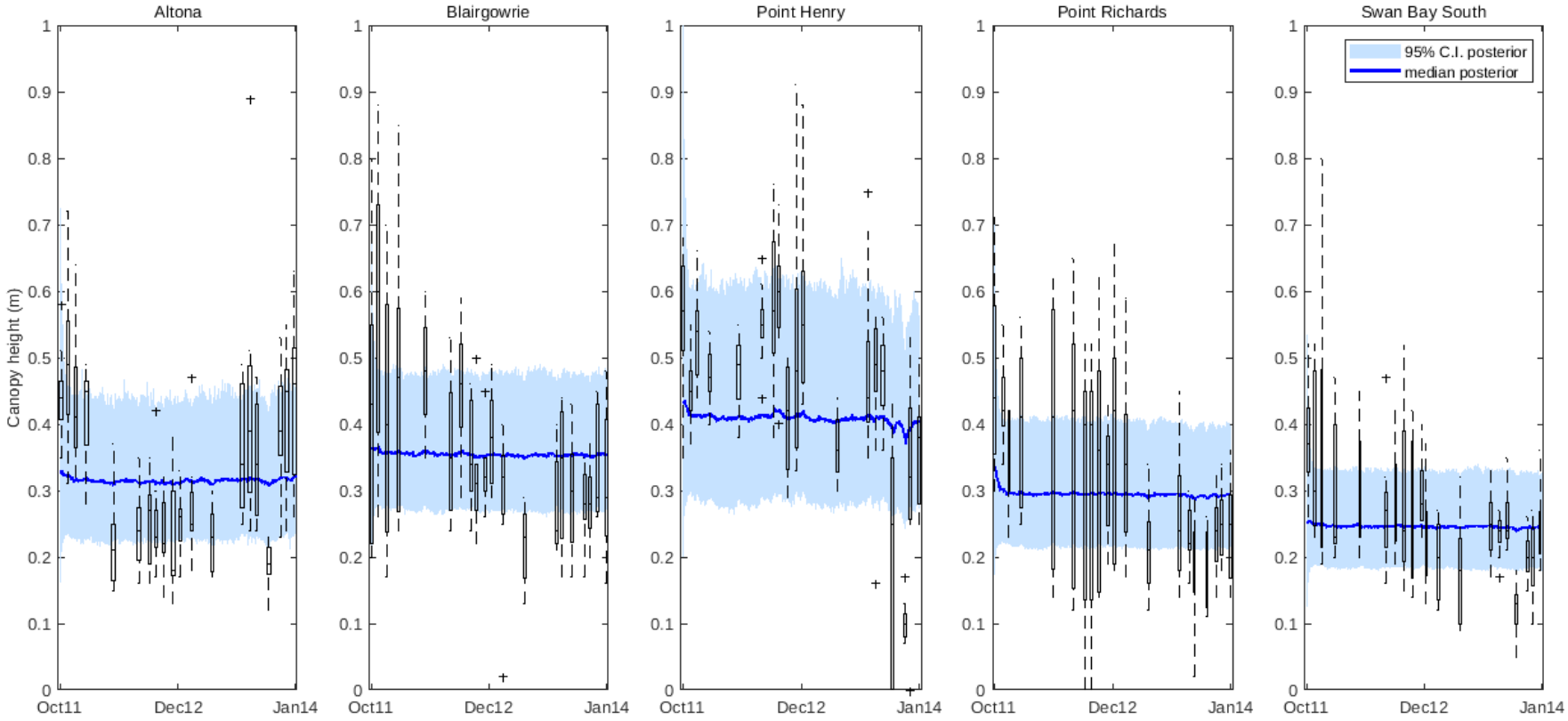

Figure 2.74 : Canopy height (m) posterior medians (solid blue line) and 95% credible intervals (shaded blue) for the full set of observations. Observations (n=9 at each sampling time-point) are shown as box-plots with potential outliers (+).

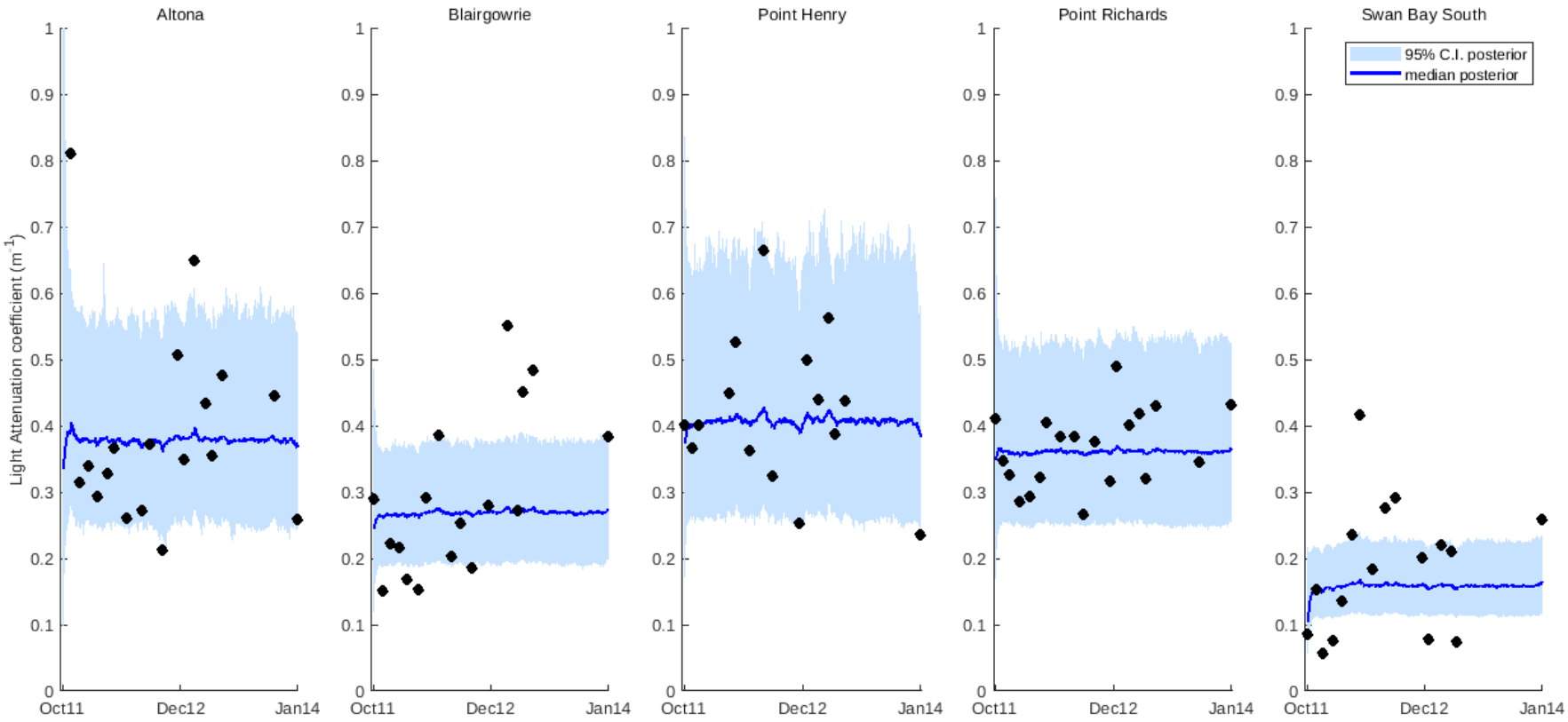

Figure 2.75 : Light attenuation coefficient (m^{-1}) posterior medians (solid blue line) and 95% credible intervals (shaded blue) for the full set of observations. Observations (n=1 at each sampling time-point) are shown as solid black circles.

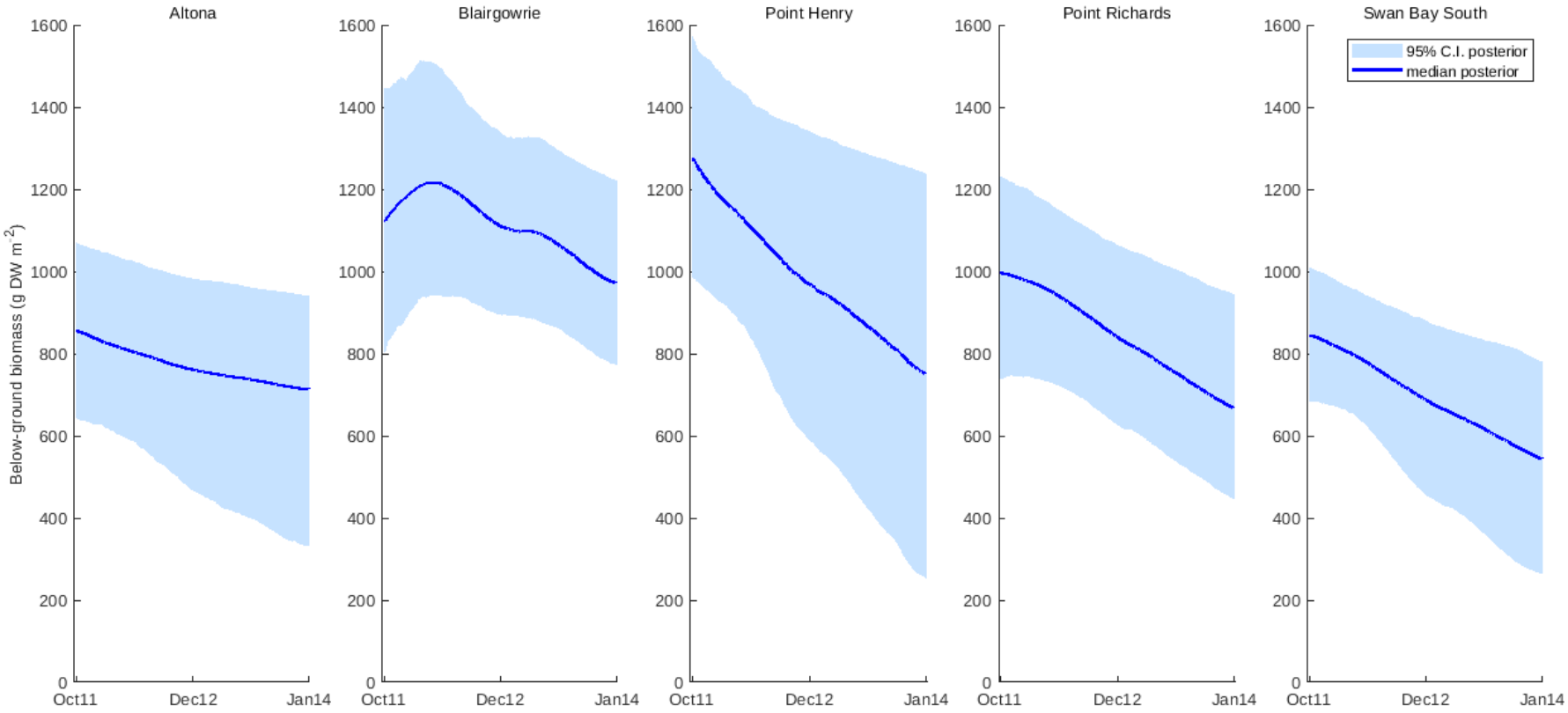

Figure 2.76 : Below-ground biomass (g DW m^{-2}) posterior medians (solid blue line) and 95% credible intervals (shaded blue) for the full 2012-2013 period.

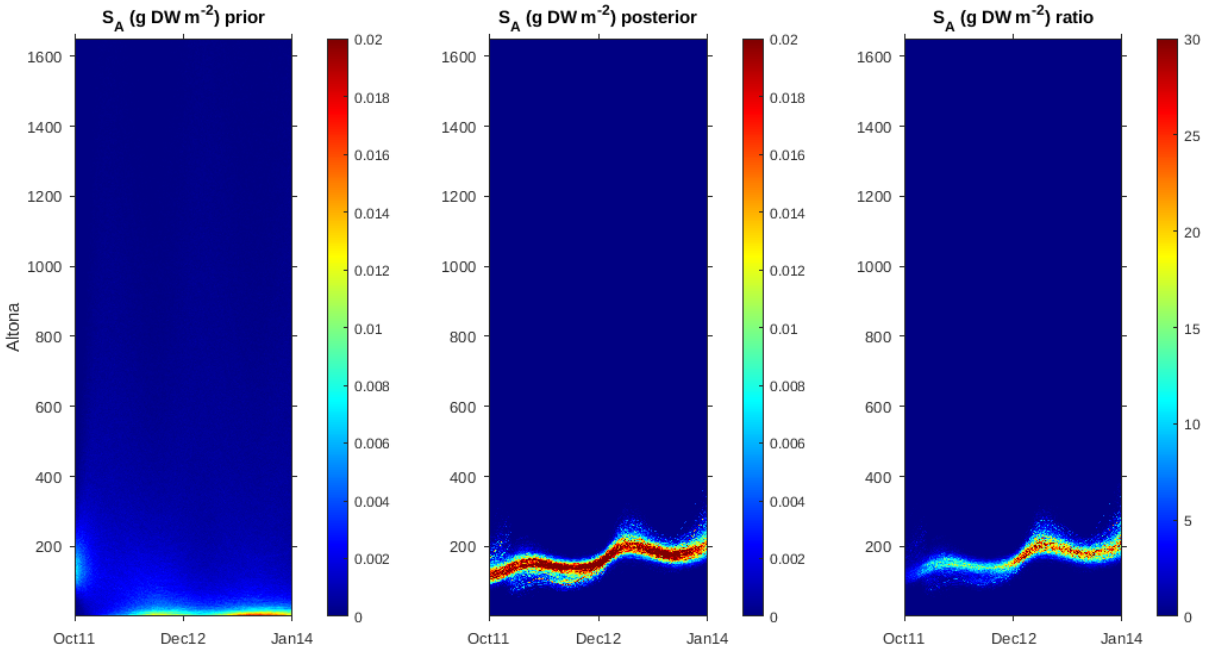

Figure 2.77 : Above-ground biomass (g DW m^{-2}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Altona in 2012-2013.

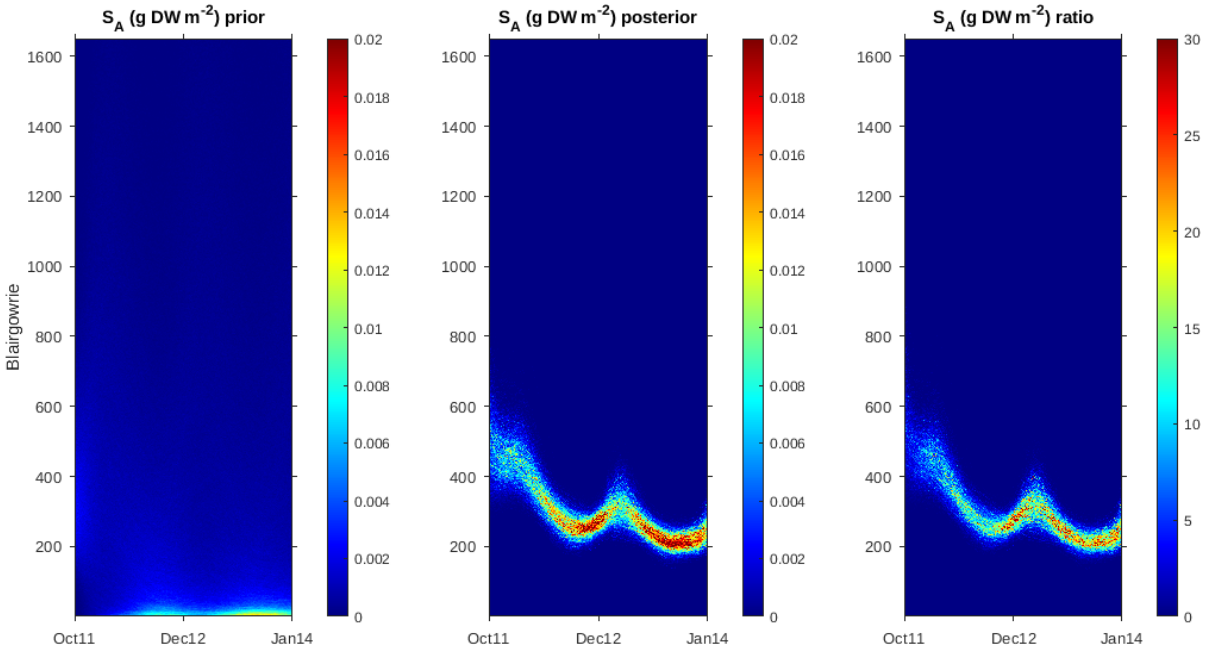

Figure 2.78 : Above-ground biomass (g DW m^{-2}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Blairgowrie in 2012-2013.

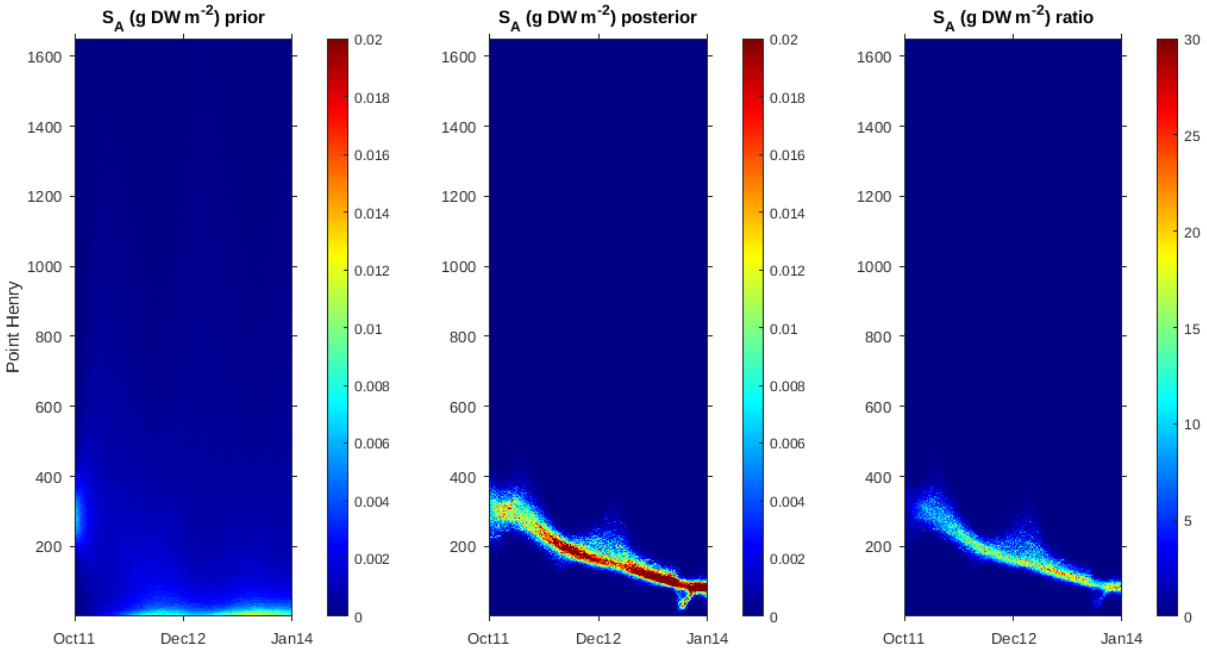

Figure 2.79 : Above-ground biomass (g DW m^{-2}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Point Henry in 2012-2013.

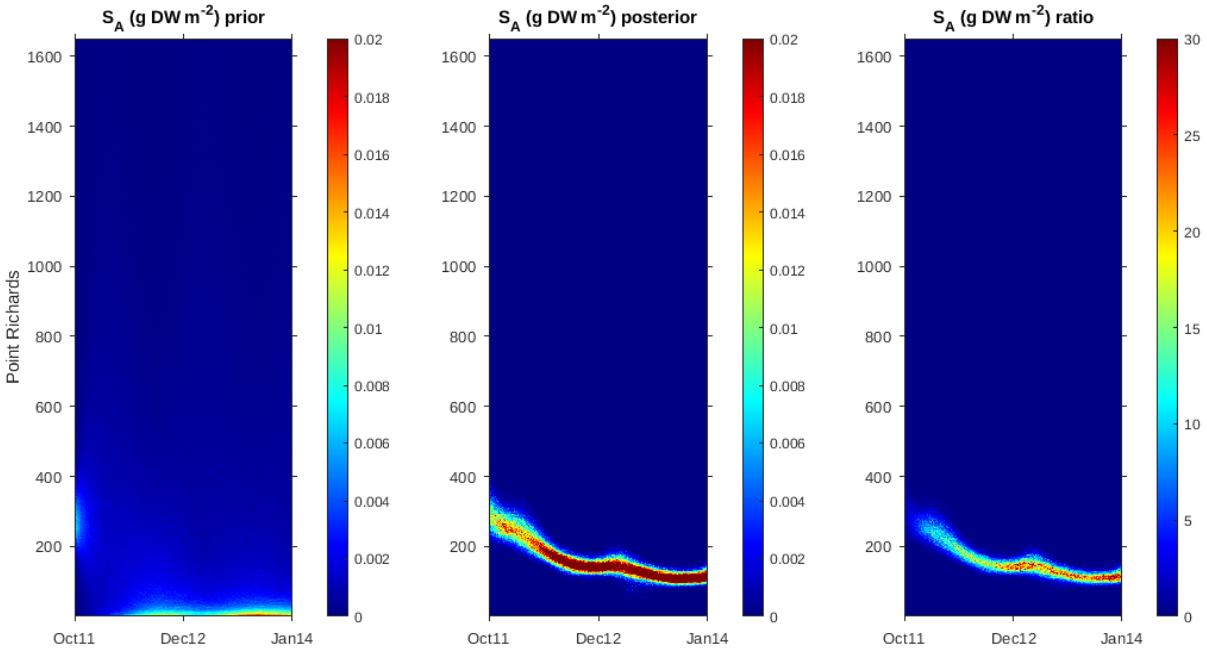

Figure 2.80 : Above-ground biomass (g DW m^{-2}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Point Richards in 2012-2013.

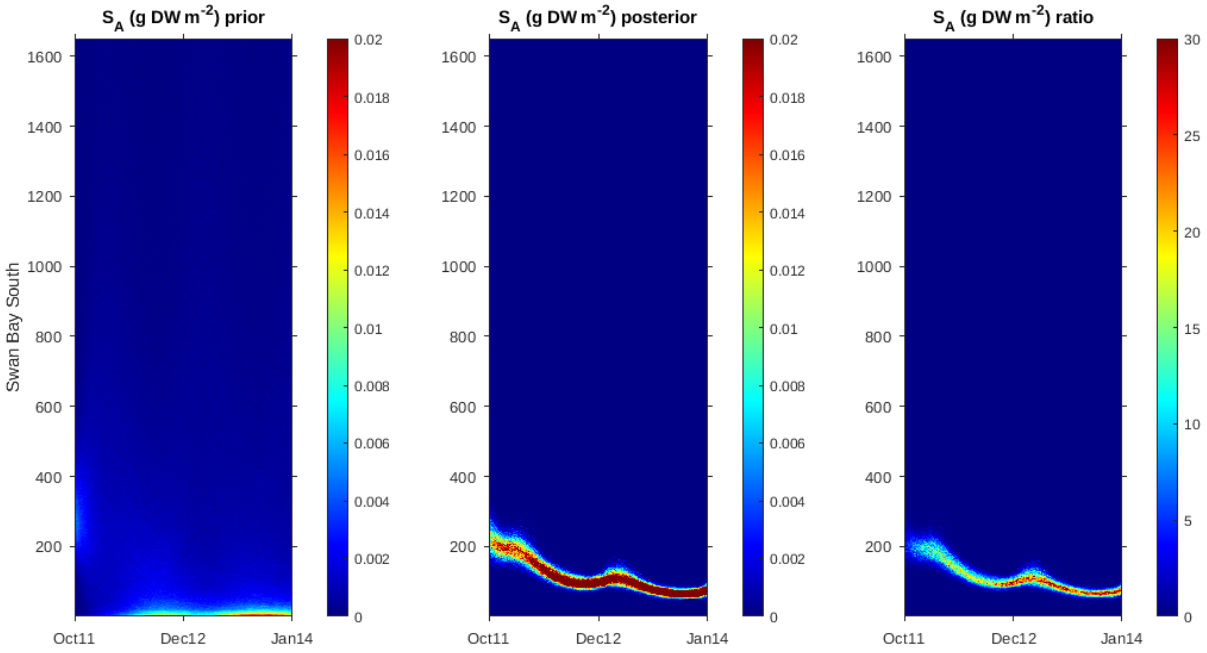

Figure 2.81 : Above-ground biomass (g DW m^{-2}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Swan Bay South in 2012-2013.

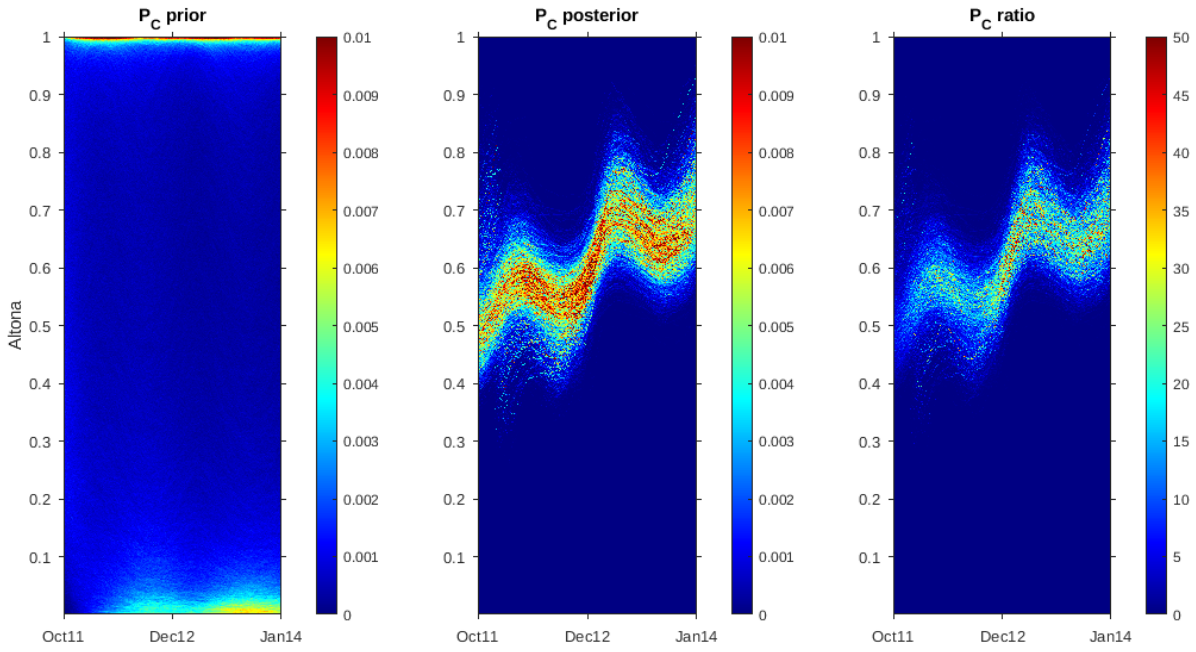

Figure 2.82 : Proportion cover fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Altona in 2012-2013.

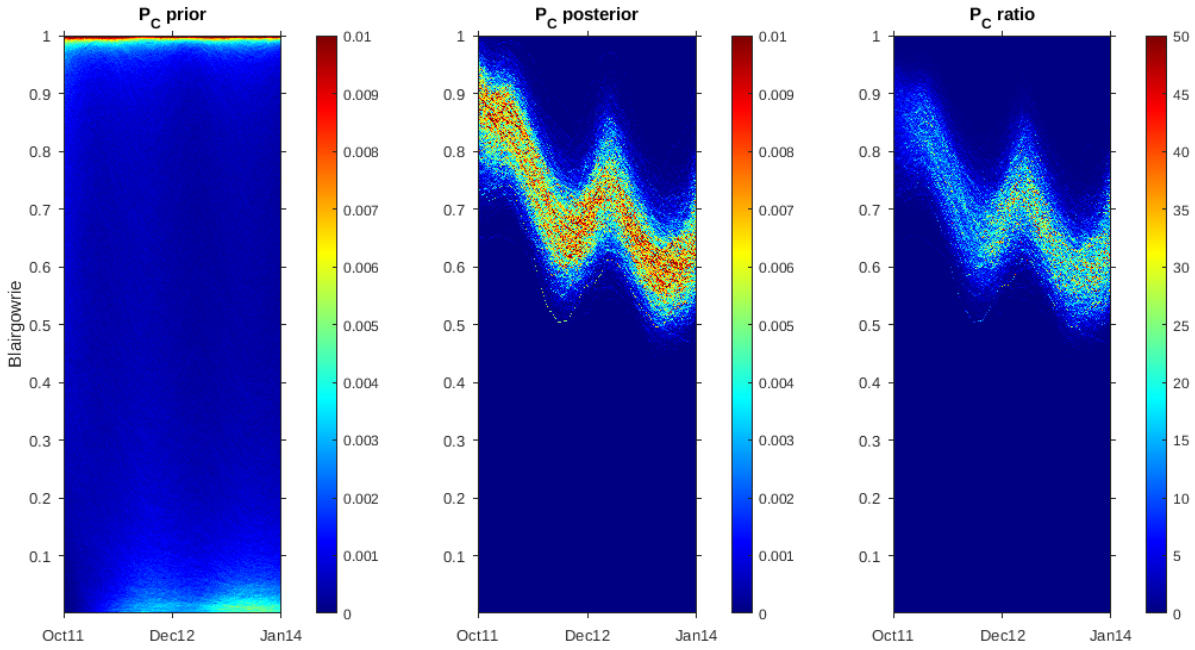

Figure 2.83 : Proportion cover fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Blairgowrie in 2012-2013.

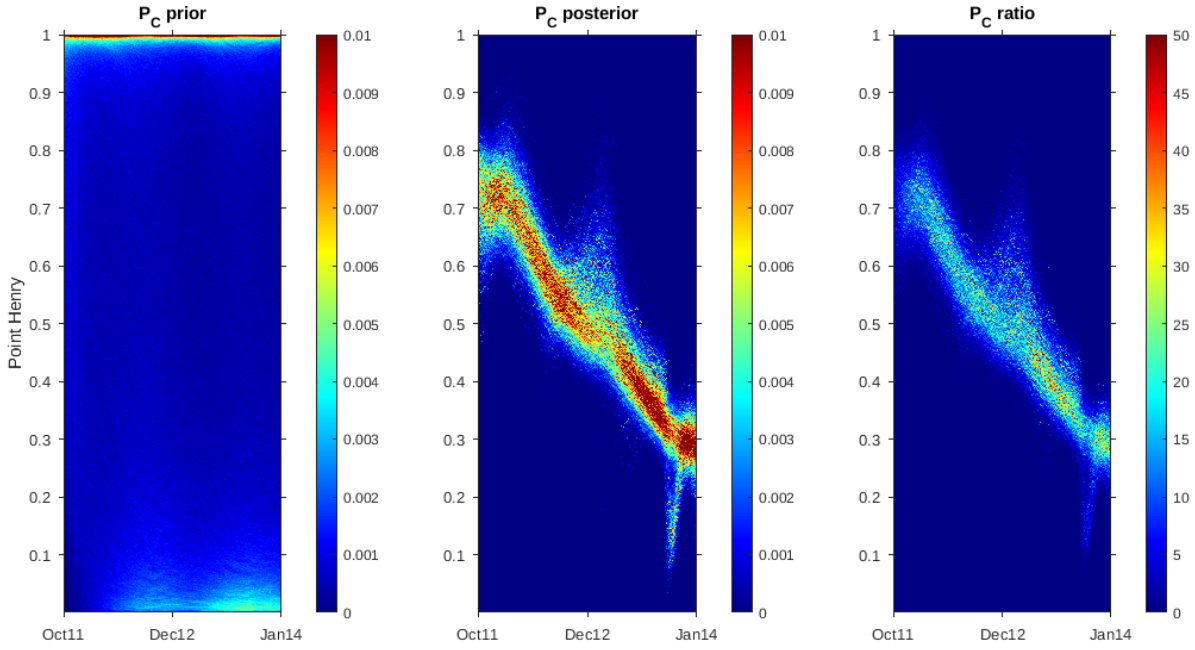

Figure 2.84 : Proportion cover fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Point Henry in 2012-2013.

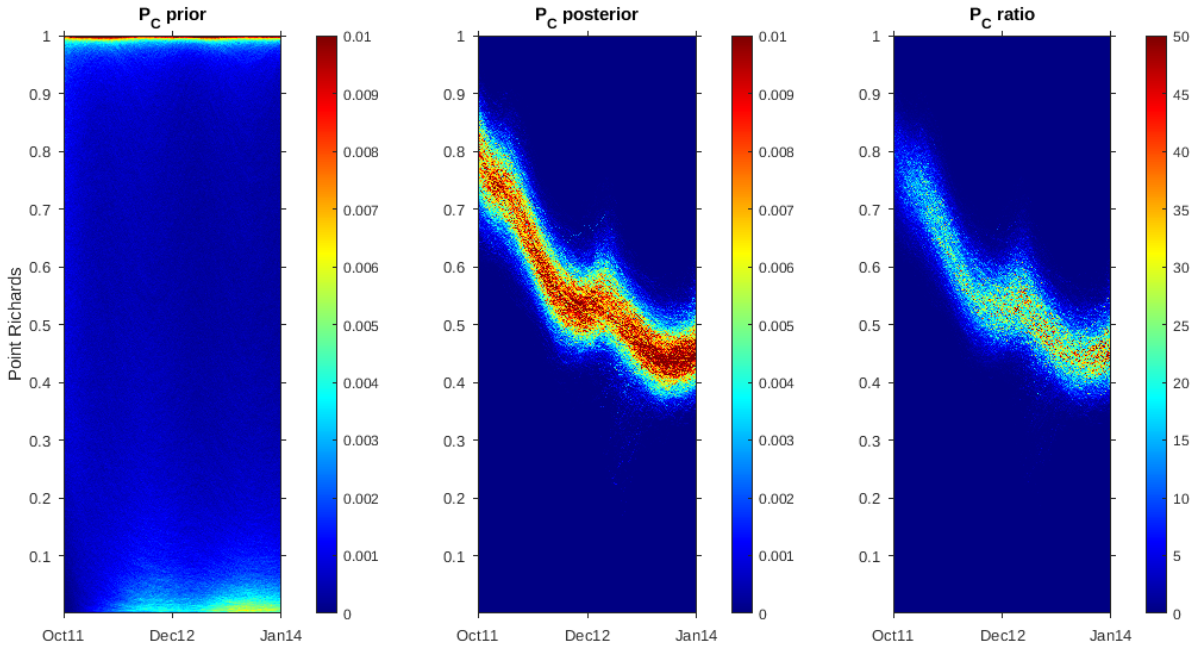

Figure 2.85 : Proportion cover fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Point Richards in 2012-2013.

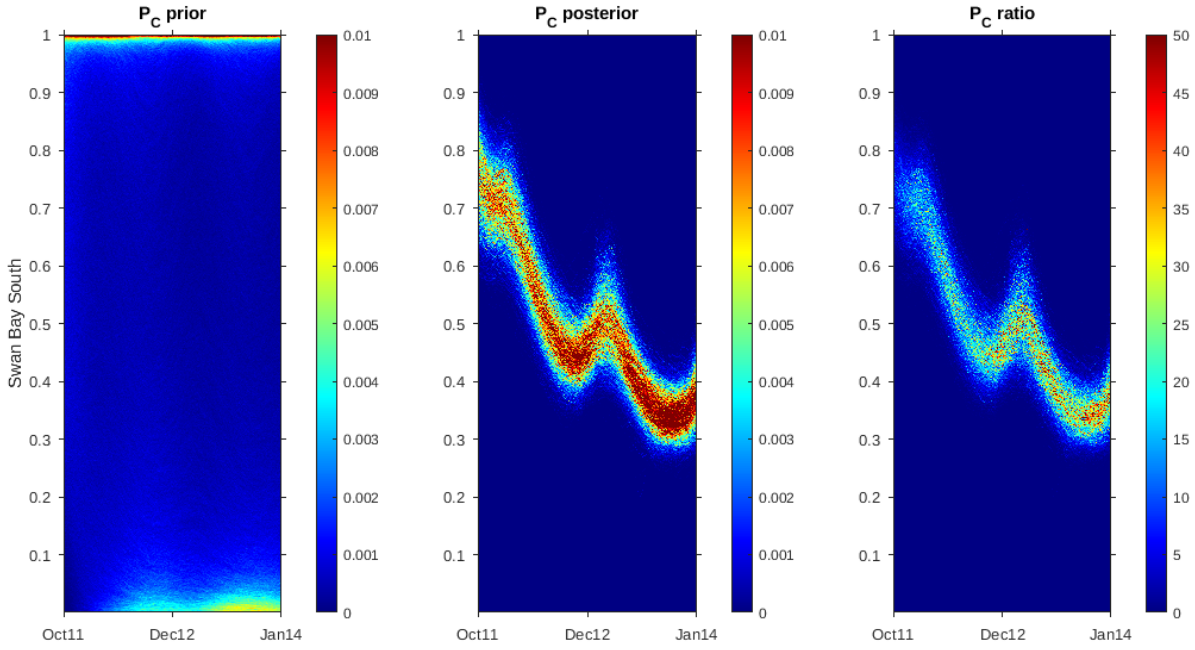

Figure 2.86 : Proportion cover fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Swan Bay South in 2012-2013.

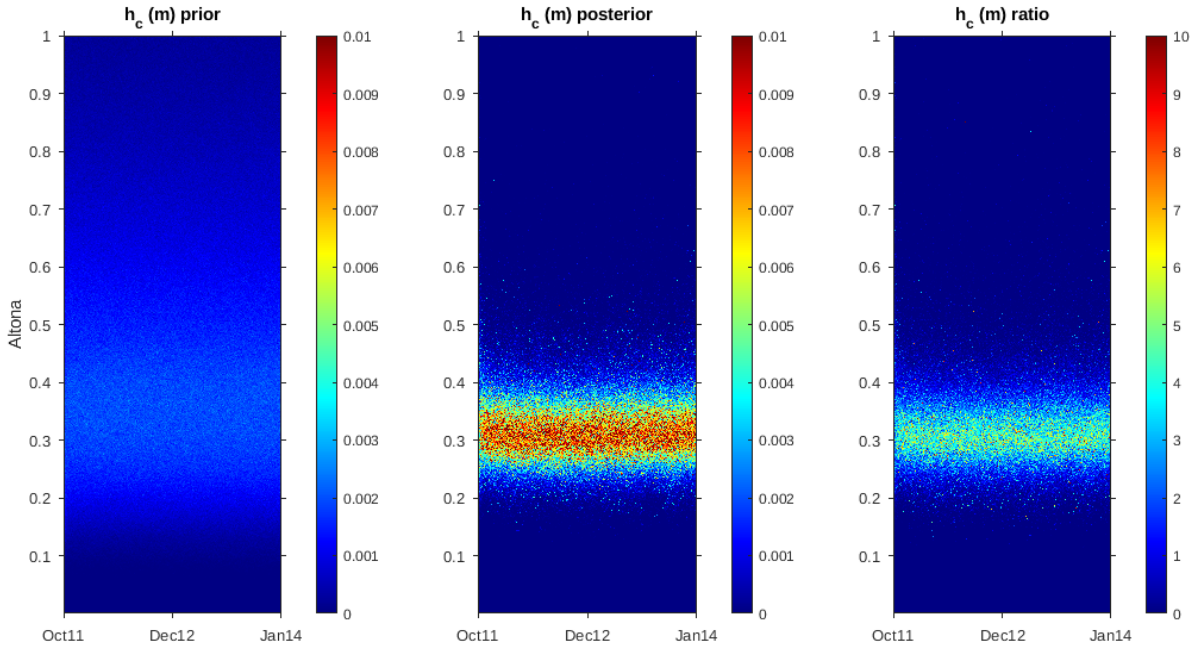

Figure 2.87 : Canopy height (m) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Altona in 2012-2013.

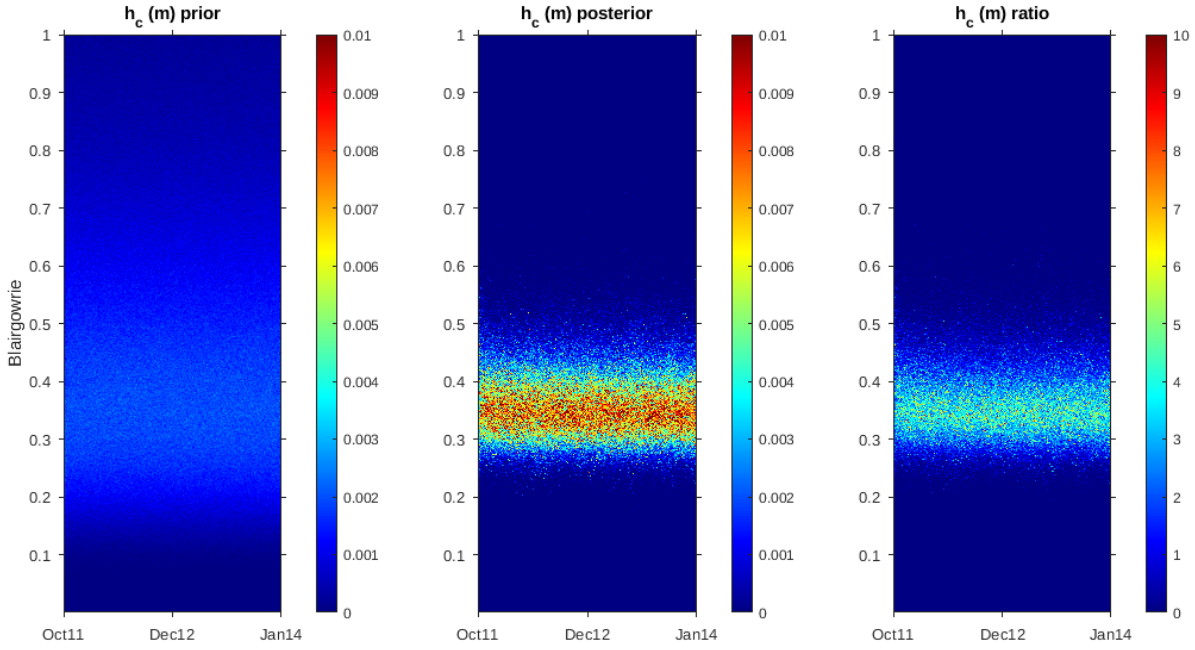

Figure 2.88 : Canopy height (m) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Blairgowrie in 2012-2013.

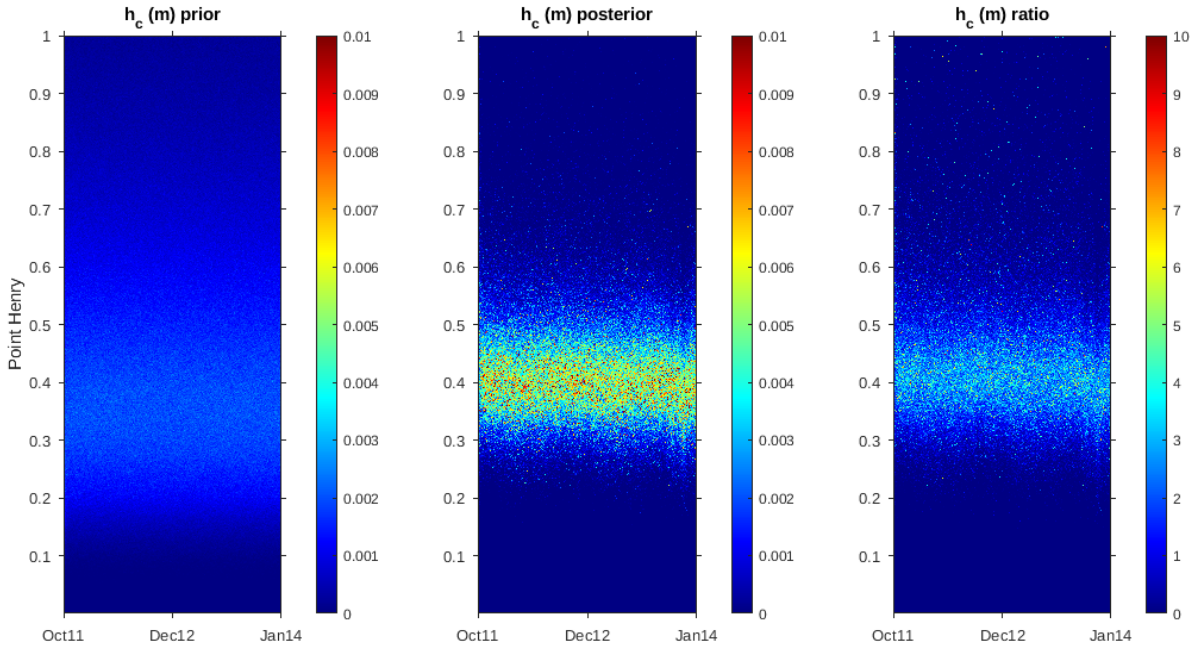

Figure 2.89 : Canopy height (m) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Point Henry in 2012-2013.

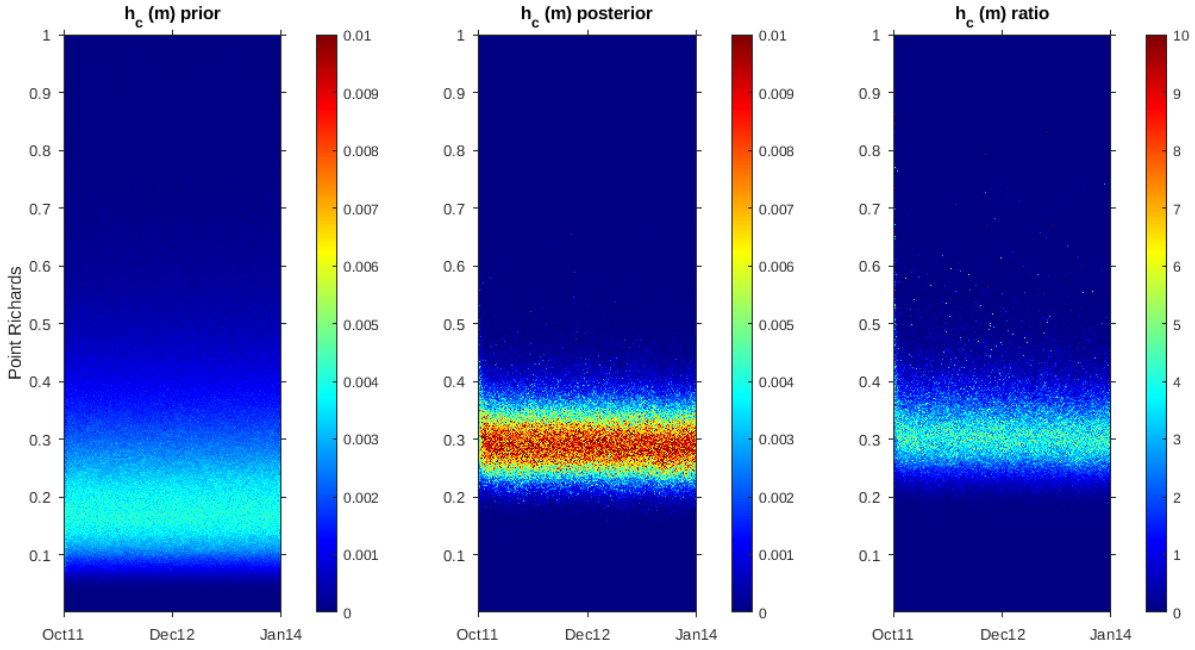

Figure 2.90 : Canopy height (m) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Point Richards in 2012-2013.

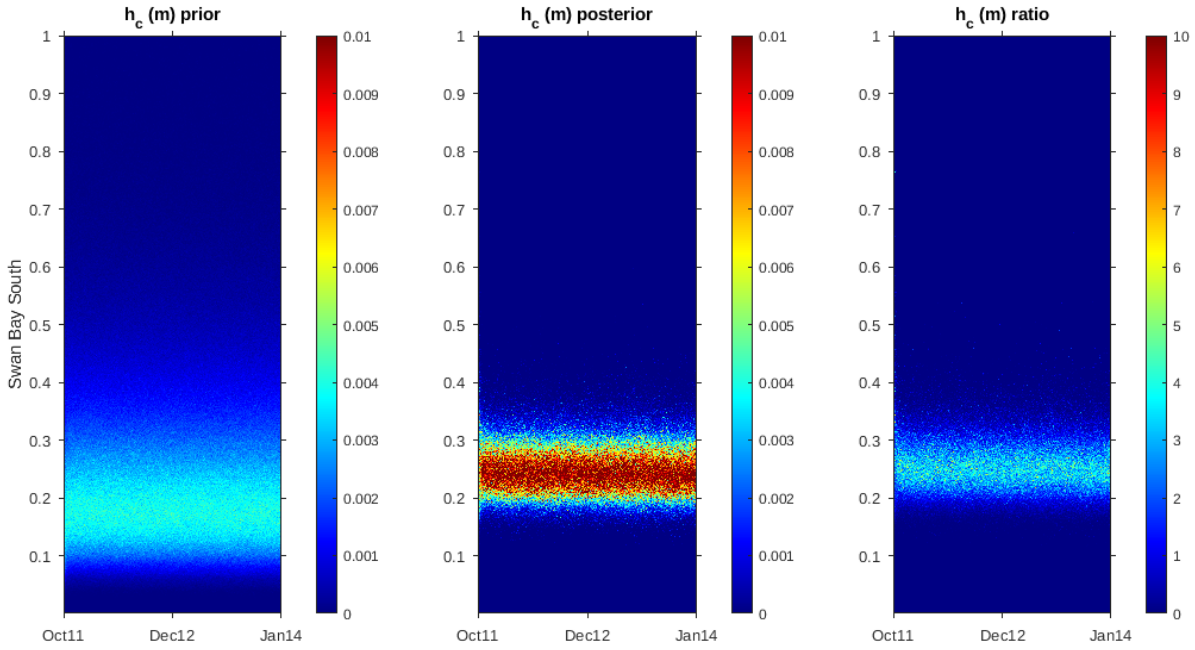

Figure 2.91 : Canopy height (m) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Swan Bay South in 2012-2013.

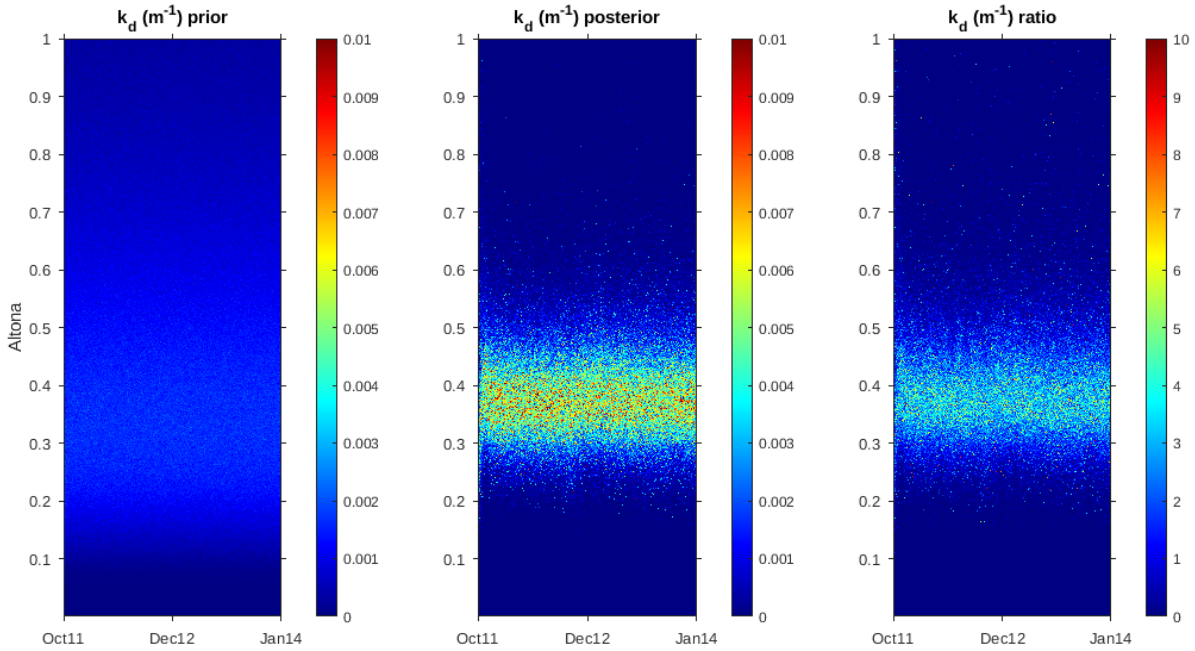

Figure 2.92 : Light attenuation coefficient (m^{-1}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Altona in 2012-2013.

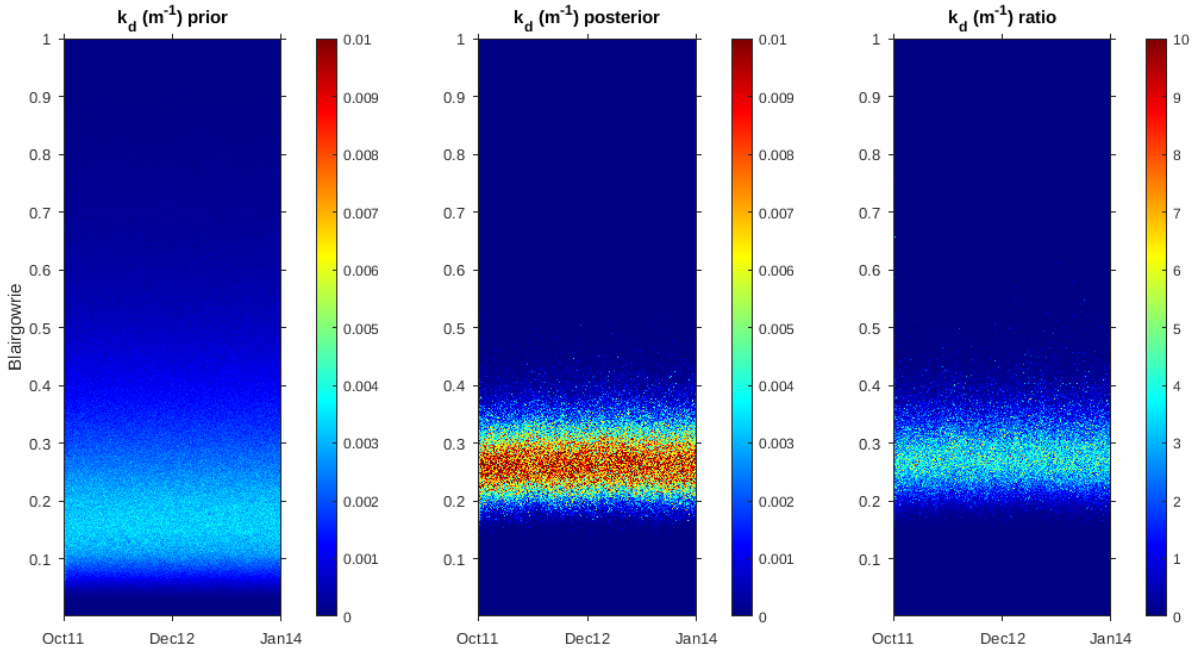

Figure 2.93 : Light attenuation coefficient (m^{-1}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Blairgowrie in 2012-2013.

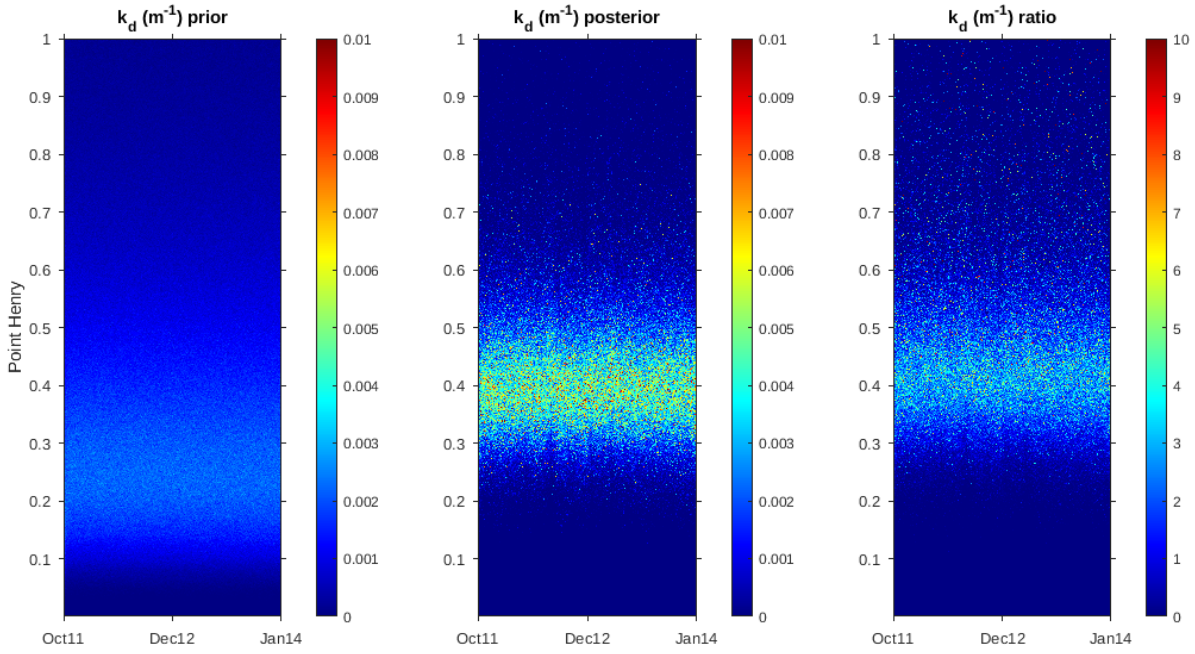

Figure 2.94 : Light attenuation coefficient (m^{-1}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Point Henry in 2012-2013.

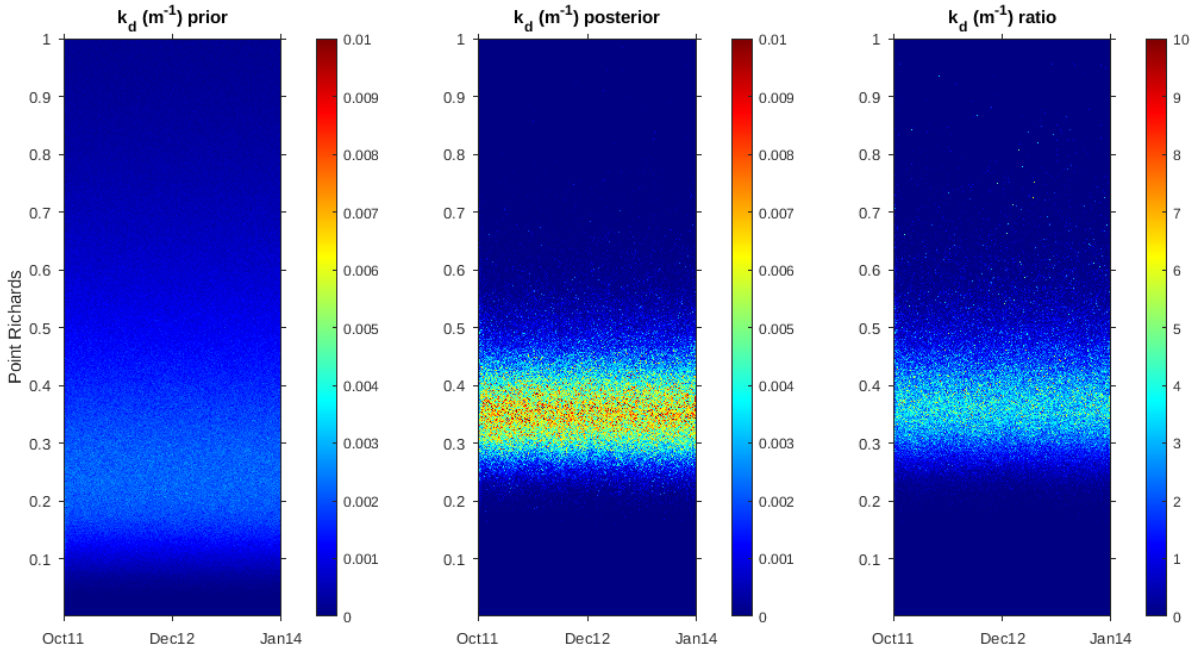

Figure 2.95 : Light attenuation coefficient (m^{-1}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Point Richards in 2012-2013.

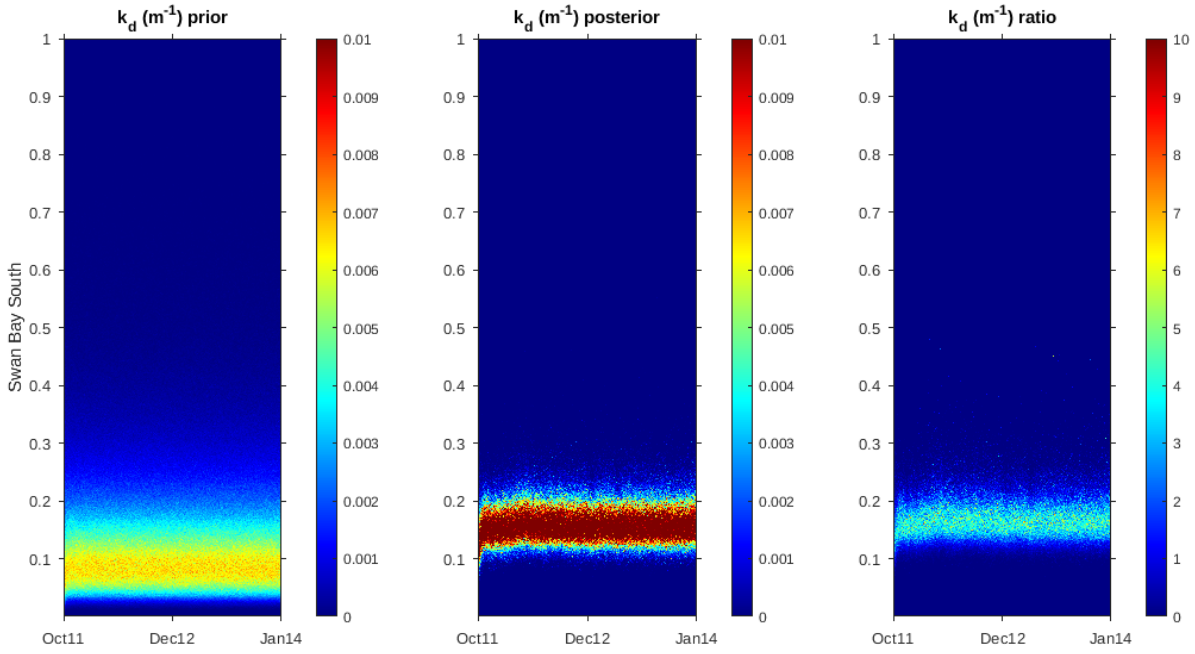

Figure 2.96 : Light attenuation coefficient (m^{-1}) fraction of density heatmap of priors, posteriors and their ratio (posterior/prior) for Swan Bay South in 2012-2013.

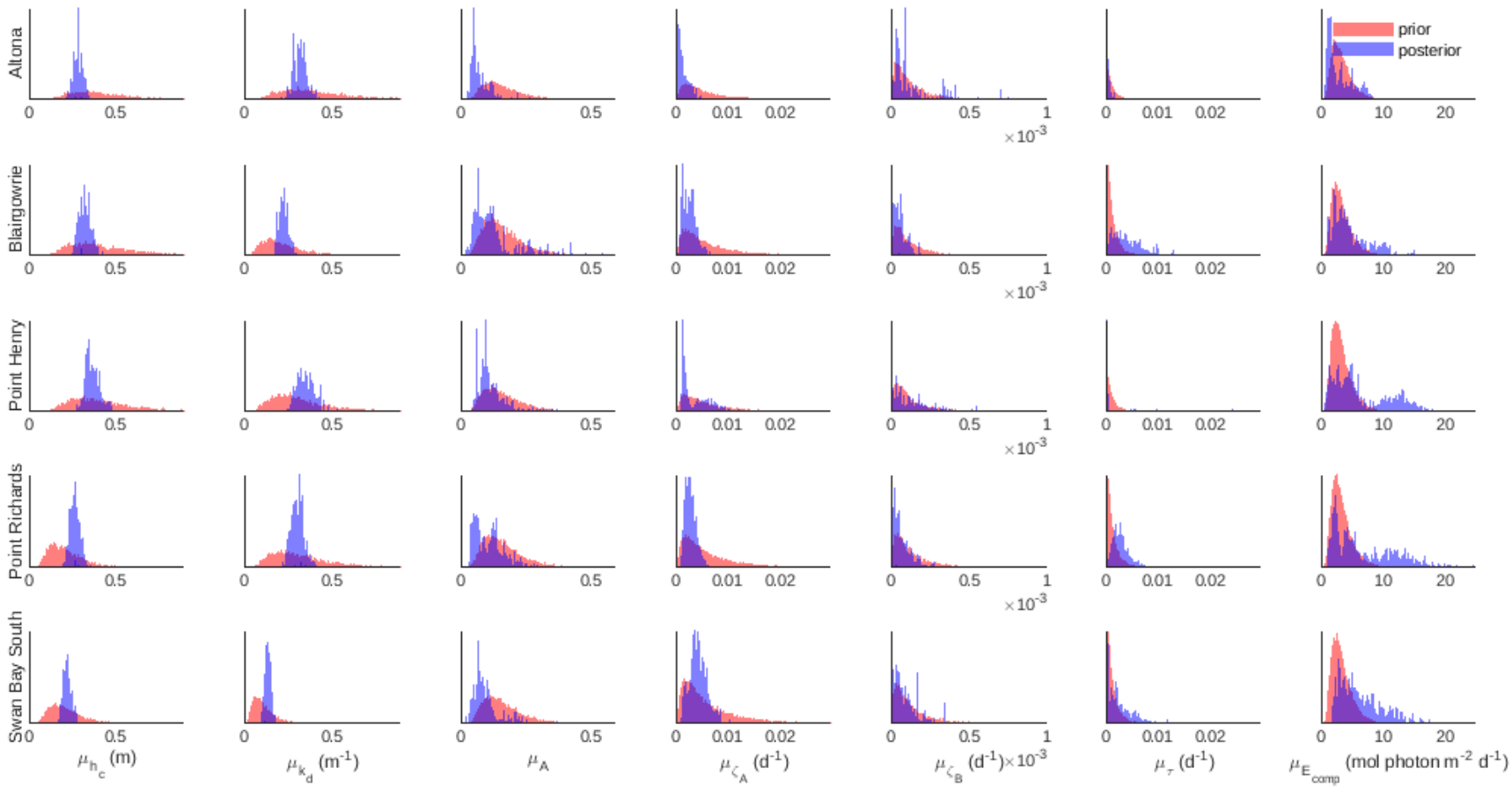

Figure 2.97 : Long-term means of the first order autoregressive parameter prior and posterior distributions for each Port Phillip Bay site in 2012-2013.

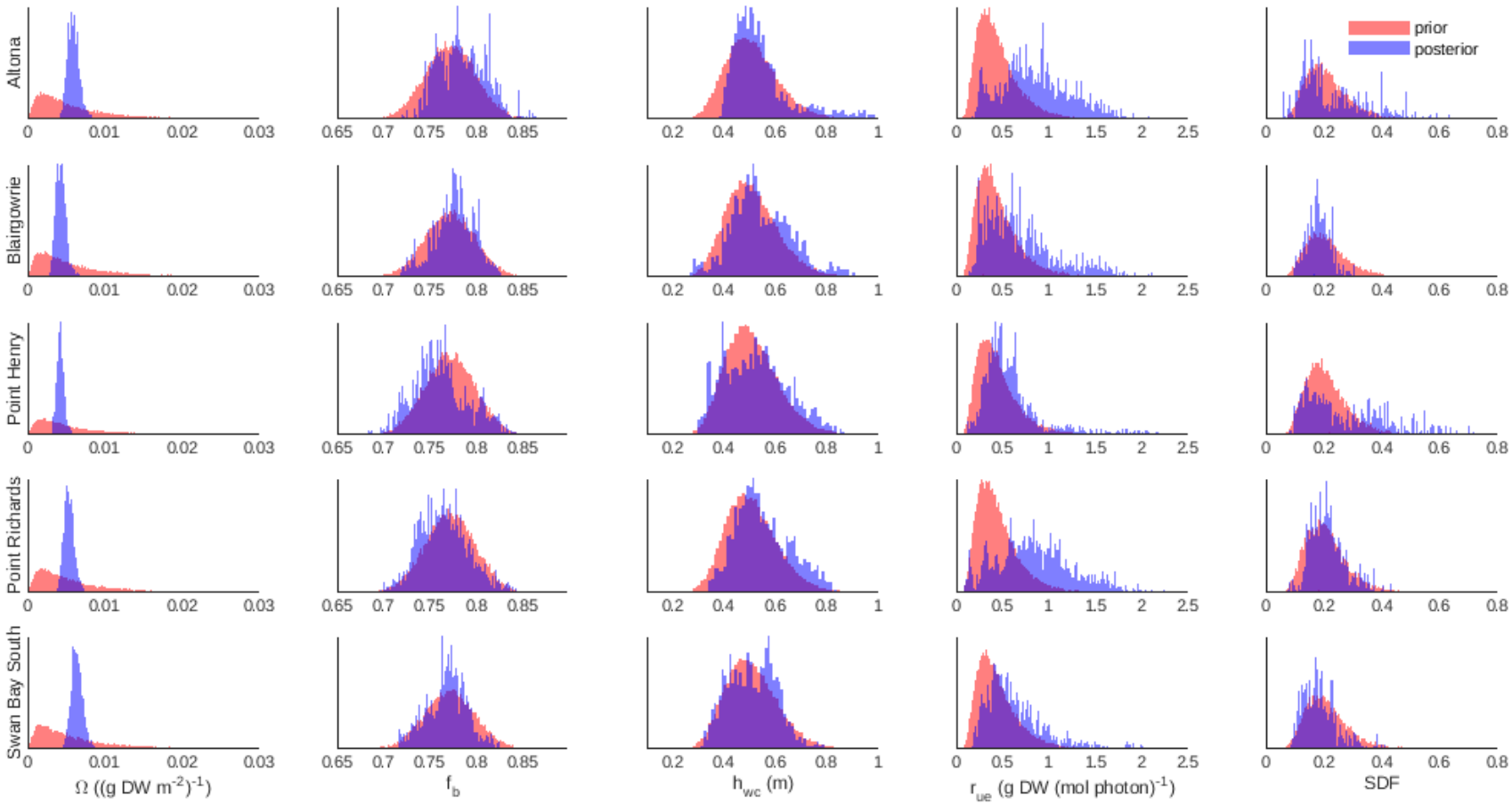

Figure 2.98 : Parameter prior and posterior distributions for each Port Phillip Bay site in 2012-2013.

gowrie, Point Henry, Point Richards and Swan Bay South posteriors all trended downwards over 2012-2013. Altona was the only site with proportion cover trending upwards over 2012-2013 (Figure 2.82).

Canopy height during 2012-2013 was tallest at Point Henry, followed by Blairgowrie, Point Richards and Altona, with Swan Bay South having the shortest seagrass canopies. The first order autoregressive process representing the canopy height posterior during 2012-2013 on average fit the observations well at Blairgowrie, Point Henry, Point Richards and Swan Bay South; potentially not capturing a decreasing trend overall at these sites. The observations at Altona during 2012-2013 look to be grouped into three stages, an initial taller height from October 2011 to January 2012 centred around 0.4000-0.5000 m, then a secondary lower height from April 2012 centred around 0.2000 m, with a third phase returning to the taller height during the second half of 2013. The highest area of posterior probability is 0.3000 m throughout 2012-2013, meaning the first order autoregressive is capturing the long-term stationary central tendency, but not necessarily capturing the changes in observed quantities of canopy height (Figure 2.74). Canopy height fraction of density heatmaps illustrated tightening of posterior density after assimilating observations at each site. There was 4-5 times the posterior density than the sampled prior at those points (Figure 2.87, Figure 2.88, Figure 2.89, Figure 2.90, Figure 2.91). Point Henry had the largest spread of canopy height posterior over 2012-2013 (Figure 2.89). At each site the long-term mean canopy height posterior IQR decreased to 0.0346-0.0580 m from the prior IQR 0.1380-0.2758 m (Table 2.7). The posterior medians at Altona, Blairgowrie and Point Henry all shifted downwards from the prior median 0.4000 m. Altona shifted to 0.2816 m, Blairgowrie shifted to 0.3145 m, and Point Henry shifted to 0.3641 m. Point Richards and Swan Bay South posterior medians (0.2630 m and 0.2181 m respectively) shifted upwards from their prior median of 0.2000 m. Point Henry had the tallest canopy height during 2012-2013, followed

by Blairgowrie, Altona, Point Richards, and Swan Bay South having the shortest canopy heights (Figure 2.97).

Point Henry and Altona had the highest light attenuation coefficient posteriors across 2012-2013, followed by Point Richards, Blairgowrie and then Swan Bay South with the lowest light attenuation (Figure 2.75). Altona, Point Henry and Point Richards posteriors captured the observations in 2012-2013 well. Light attenuation observations at Blairgowrie appear to be increasing over 2012-2013 which was not well captured by a first order autoregressive (Figure 2.75). Light attenuation coefficient fraction of density heat-maps illustrated tightening of posterior density after assimilating observations at each site. There was 4-5 times the posterior density than the sampled prior at those points (Figure 2.92, Figure 2.93, Figure 2.94, Figure 2.95, Figure 2.96). Altona (Figure 2.92), Point Henry (Figure 2.94) and Point Richards (Figure 2.95) had the largest spread of posterior density of all the sites, while Blairgowrie (Figure 2.93) and Swan Bay South (Figure 2.96) were the most tightly constrained. The long-term mean light attenuation coefficient posterior had smaller spread than the prior, with posterior IQR approximately one quarter of each sites prior IQR (Figure 2.97) with Swan Bay South having the smallest spread of $0.0231\,\mathrm{m}^{-1}$. Altona saw a shift in posterior median to $0.3205\,\mathrm{m}^{-1}$ from a prior median of $0.4000\,\mathrm{m}^{-1}$. Point Richards remained unchanged, while Point Henry shifted to a posterior median of $0.3476\,\mathrm{m}^{-1}$ from a prior of $0.3000\,\mathrm{m}^{-1}$. Blairgowrie shifted subtly upwards to $0.2258\,\mathrm{m}^{-1}$ from a prior of $0.2000\,\mathrm{m}^{-1}$. Swan Bay South shifted to $0.1344\,\mathrm{m}^{-1}$ from prior median $0.1000\,\mathrm{m}^{-1}$. Swan Bay South had the lowest central tendency and spread of light attenuation coefficient while Point Henry had the highest central tendency and spread of all the sites over 2012-2013.

Point Henry and Blairgowrie have the highest initial below-ground biomass posteriors in October 2011, followed by Point Richards, then Altona and Swan Bay South with the lowest of the five sites. Below-ground biomass posteriors for Altona,

Point Henry, Point Richards and Swan Bay South are all decreasing over 2012-2013, Blairgowrie is the only site that sees an initial increase during 2012 before decreasing for the remainder of 2012 and 2013 (Figure 2.76).

Effective absorbance, while unobserved, displayed shifts in the posterior from the prior indicating that the observations had an effect on constraining this parameter. The same prior was assigned to all sites with a median of 0.1500 and IQR 0.1000. Blairgowrie, Point Henry and Point Richards had the highest posterior medians 0.1014-0.1093 and largest IQR 0.0606-0.0843. Swan Bay South (posterior median of 0.0825 and IQR 0.0486) and Altona (posterior median of 0.0644 and IQR 0.0496) had the lowest effective absorbance during 2012-2013.

Leaf mortality rate posteriors were more concentrated from the prior IQR $0.7300\times10^{-2}\,\mathrm{d}^{-1}$ at each site. Swan Bay South had the highest leaf mortality rate posterior median of $0.4108\times10^{-2}\,\mathrm{d}^{-1}$ and was the site closest to the prior median of $0.5000\times10^{-2}\,\mathrm{d}^{-1}$. Point Richards followed with a leaf mortality rate posterior median of $0.2732\times10^{-2}\,\mathrm{d}^{-1}$ and IQR $0.1500\times10^{-2}\,\mathrm{d}^{-1}$, then Point Henry (median $0.2564\times10^{-2}\,\mathrm{d}^{-1}$ and IQR $0.4616\times10^{-2}\,\mathrm{d}^{-1}$), Blairgowrie (median $0.2551\times10^{-2}\,\mathrm{d}^{-1}$ and IQR $0.1774\times10^{-2}\,\mathrm{d}^{-1}$) and Altona (median $0.1082\times10^{-2}\,\mathrm{d}^{-1}$ with IQR $0.1468\times10^{-2}\,\mathrm{d}^{-1}$) was the lowest.

There were minimal shifts from root/rhizome mortality rate priors (median $0.1000\times10^{-3}\,\mathrm{d}^{-1}$ and IQR $0.1500\times10^{-3}\,\mathrm{d}^{-1}$) to posteriors at Altona and Point Henry. Blairgowrie, Point Richards and Swan Bay South saw slight shifts in root/rhizome mortality rate posterior medians of $0.0591\text{-}0.0642\times10^{-3}\,\mathrm{d}^{-1}$ and subtle learnings in IQR $0.0657\text{-}0.0840\times10^{-3}\,\mathrm{d}^{-1}$.

Altona was the only site where the translocation rate posterior median $0.0353\times10^{-2}\,\mathrm{d}^{-1}$ shifted downwards and IQR $0.0385\times10^{-2}\,\mathrm{d}^{-1}$ tightened from the prior median $0.1000\times10^{-2}\,\mathrm{d}^{-1}$ and IQR $0.1400\times10^{-2}\,\mathrm{d}^{-1}$ during 2012-2013. Point

Henry had a small central tendency but the largest spread of any of the sites (IQR $0.8742\times10^{-2}\,d^{-1}$), 6 times the spread of the prior distribution. Blairgowrie, Point Richards and Swan Bay South all had increases in posterior medians 0.2302-$0.3939\times10^{-2}\,d^{-1}$, but also widely spread IQR 0.2164-$0.4094\times10^{-2}\,d^{-1}$.

The compensation scalar irradiance posterior variability was larger at every site than the prescribed prior (IQR 2.0000 mol photon m^{-2} d^{-1}). Point Henry, Point Richards and Swan Bay South had the highest central tendency, but also the largest variability (IQR 5.4215-8.2368 mol photon m^{-2} d^{-1}) across 2012-2013.

Dry weight specific area of seagrass was constrained well by the observations shown in the concentrated posterior distributions with lower variability (Figure 2.98). Swan Bay South had the highest posterior median (0.6347×10^{-2} (g DW $m^{-2})^{-1}$), followed by Altona (0.5849×10^{-2} (g DW $m^{-2})^{-1}$), Point Richards (0.5322×10^{-2} (g DW $m^{-2})^{-1}$), Blairgowrie (0.4243×10^{-2} (g DW $m^{-2})^{-1}$) and Point Henry (0.4202×10^{-2} (g DW $m^{-2})^{-1}$) had the lowest. All sites had posterior IQR 0.0631-0.1076×10^{-2} (g DW $m^{-2})^{-1}$ down from prior IQR 0.7300×10^{-2} (g DW $m^{-2})^{-1}$ (Table 2.7).

Equilibrium fraction of biomass below-ground posteriors remained relatively unchanged from the priors at all sites. Water column height at all sites remained relatively unchanged from prior (median 0.5000 m and IQR 0.1300 m) to posterior (medians 0.5080-0.5341 m and IQR 0.1121-0.1839 m).

Radiation use efficiency posterior distributions all shifted upwards from the prior distribution (median 0.4000 g DW (mol photon)$^{-1}$, IQR 0.2700 g DW (mol photon)$^{-1}$). Altona and Point Richards had the largest shifts in posterior median (0.8110 g DW (mol photon)$^{-1}$), but also twice the IQR (0.4986 g DW (mol photon)$^{-1}$ and 0.5343 g DW (mol photon)$^{-1}$). Blairgowrie had the next largest posterior median radiation use efficiency (0.6153 g DW (mol photon)$^{-1}$) for 2012-2013 and twice

the variability of the prior (IQR 0.4745 g DW (mol photon)$^{-1}$). Followed by Swan Bay South (0.5518 g DW (mol photon)$^{-1}$) and Point Henry (0.5276 g DW (mol photon)$^{-1}$) with the smallest radiation use efficiency of all the sites.

Seagrass diversity factor saw minimal change in central tendency and spread from prior to posterior at Blairgowrie, Point Richards and Swan Bay South. Altona saw an increase in variability and Point Henry saw an increase in central tendency (median 0.2771) and spread (IQR 0.2057) across 2012-2013 (Table 2.7).

Parameter	**Site**	**Prior 2012-2013**	**Posterior 2012-2013**
μ_{h_c}	Altona	0.3990 (0.2758)	0.2816 (0.0416)
	Blairgowrie	0.3991 (0.2722)	0.3145 (0.0491)
	Point Henry	0.3987 (0.2728)	0.3641 (0.0580)
	Point Richards	0.1999 (0.1383)	0.2630 (0.0399)
	Swan Bay South	0.2001 (0.1380)	0.2181 (0.0346)
μ_{k_d}	Altona	0.3971 (0.3301)	0.3205 (0.0566)
	Blairgowrie	0.1999 (0.1681)	0.2258 (0.0402)
	Point Henry	0.2992 (0.2493)	0.3476 (0.0679)
	Point Richards	0.2999 (0.2499)	0.3058 (0.0525)
	Swan Bay South	0.1002 (0.0833)	0.1344 (0.0231)
μ_{A}	Altona	0.1498 (0.1031)	0.0644 (0.0496)
	Blairgowrie	0.1502 (0.1047)	0.1032 (0.0785)
	Point Henry	0.1500 (0.1038)	0.1014 (0.0606)
	Point Richards	0.1511 (0.1029)	0.1093 (0.0843)
	Swan Bay South	0.1498 (0.1024)	0.0825 (0.0486)
μ_{ζ_A}	Altona	0.4965×10^{-2} (0.7233×10^{-2})	0.1082×10^{-2} (0.1468×10^{-2})

	Blairgowrie	0.5059×10^{-2} (0.7325×10^{-2})	0.2551×10^{-2} (0.1774×10^{-2})
	Point Henry	0.5018×10^{-2} (0.7289×10^{-2})	0.2564×10^{-2} (0.4616×10^{-2})
	Point Richards	0.5034×10^{-2} (0.7288×10^{-2})	0.2732×10^{-2} (0.1500×10^{-2})
	Swan Bay South	0.4999×10^{-2} (0.7303×10^{-2})	0.4108×10^{-2} (0.2327×10^{-2})
μ_{ζ_B}	Altona	0.1004×10^{-3} (0.1453×10^{-3})	0.0835×10^{-3} (0.1453×10^{-3})
	Blairgowrie	0.1005×10^{-3} (0.1444×10^{-3})	0.0602×10^{-3} (0.0657×10^{-3})
	Point Henry	0.0996×10^{-3} (0.1460×10^{-3})	0.1112×10^{-3} (0.1623×10^{-3})
	Point Richards	0.1015×10^{-3} (0.1481×10^{-3})	0.0591×10^{-3} (0.0747×10^{-3})
	Swan Bay South	0.1000×10^{-3} (0.1445×10^{-3})	0.0642×10^{-3} (0.0840×10^{-3})
μ_{τ}	Altona	0.1002×10^{-2} (0.1378×10^{-2})	0.0353×10^{-2} (0.0385×10^{-2})
	Blairgowrie	0.1008×10^{-2} (0.1373×10^{-2})	0.3939×10^{-2} (0.4094×10^{-2})
	Point Henry	0.0999×10^{-2} (0.1372×10^{-2})	0.0442×10^{-2} (0.8742×10^{-2})
	Point Richards	0.1001×10^{-2} (0.1369×10^{-2})	0.2689×10^{-2} (0.2164×10^{-2})
	Swan Bay South	0.0997×10^{-2} (0.1354×10^{-2})	0.2302×10^{-2} (0.3289×10^{-2})
$\mu_{\mathrm{E}_{comp}}$	Altona	3.0170 (2.0808)	2.2430 (3.1040)
	Blairgowrie	3.0123 (2.0711)	3.7584 (3.8344)
	Point Henry	3.0035 (2.0846)	5.2266 (7.9099)

	Point Richards	3.0074 (2.0609)	5.4194 (8.2368)
	Swan Bay South	3.0013 (2.0746)	5.6833 (5.4215)
Ω	Altona	0.4975×10^{-2} (0.7193×10^{-2})	0.5849×10^{-2} (0.1076×10^{-2})
	Blairgowrie	0.5044×10^{-2} (0.7358×10^{-2})	0.4243×10^{-2} (0.0915×10^{-2})
	Point Henry	0.4987×10^{-2} (0.7283×10^{-2})	0.4202×10^{-2} (0.0631×10^{-2})
	Point Richards	0.4942×10^{-2} (0.7178×10^{-2})	0.5322×10^{-2} (0.0884×10^{-2})
	Swan Bay South	0.5046×10^{-2} (0.7311×10^{-2})	0.6347×10^{-2} (0.0985×10^{-2})
f_b	Altona	0.7722 (0.0366)	0.7837 (0.0403)
	Blairgowrie	0.7722 (0.0368)	0.7777 (0.0288)
	Point Henry	0.7721 (0.0368)	0.7590 (0.0373)
	Point Richards	0.7720 (0.0370)	0.7653 (0.0359)
	Swan Bay South	0.7716 (0.0371)	0.7701 (0.0296)
h_{wc}	Altona	0.5005 (0.1349)	0.5165 (0.1121)
	Blairgowrie	0.5008 (0.1355)	0.5282 (0.1586)
	Point Henry	0.4997 (0.1358)	0.5200 (0.1839)
	Point Richards	0.5008 (0.1359)	0.5341 (0.1491)
	Swan Bay South	0.5004 (0.1363)	0.5080 (0.1438)

r_{ue}	Altona	0.4005 (0.2742)	0.8110 (0.4986)
	Blairgowrie	0.3988 (0.2767)	0.6153 (0.4745)
	Point Henry	0.4018 (0.2752)	0.5276 (0.2933)
	Point Richards	0.4026 (0.2753)	0.8110 (0.5343)
	Swan Bay South	0.3998 (0.2751)	0.5518 (0.3692)
SDF	Altona	0.2004 (0.0946)	0.1993 (0.1565)
	Blairgowrie	0.2009 (0.0952)	0.1801 (0.0497)
	Point Henry	0.1997 (0.0954)	0.2771 (0.2057)
	Point Richards	0.2006 (0.0943)	0.2073 (0.0850)
	Swan Bay South	0.2004 (0.0950)	0.1762 (0.0784)

Table 2.7 : Parameter median and interquartile range (IQR) at each site for 2012-2013 prior and posterior.

2.4 Discussion

After assimilating data from 2012, we were able to accurately predict seagrass above-ground biomass and proportion cover states for 2013 at three out of five sites; Blairgowrie, Point Richards and Swan Bay South (Figure 2.56, Figure 2.57), with prediction medians and 95% credible intervals capturing the trend of the unseen 2013 observations. From the data collected and posteriors generated during 2012, there was no indication of a decline in biomass at Point Henry during 2013. Point Henry was the only site to have stable above-ground biomass (Figure 2.9) and proportion cover (Figure 2.14) posteriors during 2012 while the other four sites showed decreasing trends (Figure 2.7 - 2.11, Figure 2.12 - 2.16). Thus the prediction for 2013 at Point Henry over-predicted the actual observations, as the model training dataset (2012) did not display any kind of loss or decline to indicate the future 2013 behaviour. All four of the other sites declined during 2012 (Figure 2.29, Figure 2.30) which resulted in three sites (Blairgowrie, Point Richards and Swan Bay South), where a similar pattern of behaviour was illustrated during 2013, resulting in a well performing model with accurate predictions. Of the three sites with successful predictions for 2013, all three exhibited above-ground biomass and proportion cover seasonality with a downward trend across 2012 and 2013. Point Richards seagrass meadows were found to have the highest effective absorbance and highest radiation use efficiency of the three sites across the 2 year period. The highest translocation rate of the three sites was estimated at Blairgowrie. Swan Bay South seagrass meadows had the highest leaf mortality rate and highest dry weight specific area of seagrass.

Point Henry posteriors estimated a higher leaf mortality, translocation, and respiration than predicted in 2013 (Section 2.3.4). Increases in each of these processes caused a larger loss to above-ground biomass and proportion cover during 2013. In

addition to these, the seagrass diversity factor was higher indicating the scale of the variance on the autoregressive processes was much higher in the posterior than the prediction. This is indicative of an unquantified process in the model formulation or missing forcing at Point Henry, which could be nutrient limitation, changes to epiphyte loads or various environmental events causing substantial seagrass meadow loss. The parameters for Point Henry should be interpreted with caution as they could be overcompensating for an un-modelled process or missing forcing. For future work, collecting and assimilating fundamental environmental observations is key to confirming this. The different behaviour of seagrass at Point Henry suggests there are differing local environmental factors driving seagrass growth at the different sites in Port Phillip Bay, like epiphyte blooms occurring when conditions are favourable at local scales [87].

Variations in seagrass growth can be linked to environmental factors such as temperature and rainfall, where reduced rainfall has the potential to cause nitrogen limitation, particularly following periods of drought (porosity is another alternative) [50]. Blairgowrie, Point Richards and Swan Bay South results indicate that they lend themselves well to a light limited model which coincides with current literature for Point Richards and Swan Bay South, with Swan Bay South having the highest sediment pore water nutrient levels in the bay [50]. This contradicts literature for Blairgowrie however, that reports Blairgowrie having the lowest nutrient levels and being the only site that exhibited a pronounced response to nutrient enrichment from a recent in-situ experiment [50, 51].

At Altona, in contrast to Point Henry, the model under-predicted the seagrass states in 2013 after assimilating 2012 data. Altona posteriors estimated a lower leaf mortality, translocation, and respiration than predicted in 2013 (Section 2.3.4). Each decrease in these processes allowed more above-ground biomass growth and proportion cover during 2013. Altona is exposed to the two largest nutrient and

turbidity inputs into the bay, the Yarra River and Western Treatment Plant [62]. It has high nutrient content, fine sediment grain size, medium sediment organic matter content, and high wave height [87]. While it is possible the exposure to high nutrient inputs caused an irregular growth in 2013, it is also equally possible that one year of data is not enough to capture seasonal dynamics at this site. Had there been access to a longer time series of observations, 2+ years for example, to train the model and then produce a prediction, this would likely improve predictive results. This may also be an indication that one year is too large a time frame to attempt to predict and perhaps 3-6 months is a more appropriate prediction window for seagrass state. A future work direction would be to optimise how many months/years of data is needed to make an accurate 3 or 6 monthly prediction of above-ground seagrass biomass and proportion cover.

It is difficult to draw conclusions in high dimensional environments where a complex model is coupled with many unobserved parameters, and a sparse dataset. While a longer time-series would be ideal, due to the destructive nature of sampling techniques, this is generally not feasible. This makes key relationships like the one between proportion cover and above-ground biomass increasingly important. In the future, given this well defined quantitative relationship has been validated to hold true, non destructive datasets of proportion cover could be used to infer above-ground biomass without the need for destructive sampling of seagrass meadows. However, the variability of proportion cover observations exhibited within plots was high using this collection method (Figure 2.4). These could be combined with automated image processing techniques like those used in [5] that historically mapped seagrass cover in Port Phillip Bay between 1939 and 2011, to tease out some of the within plot variability and scale up the proportion cover observations.

Zostera nigricaulis above-ground biomass posteriors were tightly constrained estimates at all sites during all years and in agreement with values historically recorded

in Port Phillip Bay [17], but greater than reported for Western Australian populations [66].

A key achievement of this chapter is the quantitative meta-analysis of parameter values currently in the literature for the construction of informative priors in Section 2.2.3. This complements the work by Barbara Robson towards evidence based priors and parameters for aquatic ecosystem modelling through a parameter library [82].

Some areas of refinement as potential future work could further investigate the translocation rate relationship to state variables as it is a key parameter in drawing long term trends in above-ground biomass. It is the key driver of the above-ground to below-ground biomass equilibrium, and a translocation upwards could be indicating a deficit in above-ground biomass growth. We note that AR(1) parameters like light attenuation and canopy height do not display much variation around the mean (Figure 2.75). A potential extension of this thesis could be to treat the AR1 parameter St as subject to estimation as a means of exploring parameter dynamics over time. Alternatively a different functional form to AR(1) could also be explored.

Chapter 3

Modelling seagrass detrital decay using a time-varying decay rate

Abstract

In recent years, there has been increasing research interest into seagrass meadows, owing to the global role they play as significant carbon sinks. It is now believed that blue carbon ecosystems (seagrasses, mangroves and saltmarshes) match the sequestration capability of terrestrial systems. When these plants die and are broken down, this has a sizeable impact upon the cycling of carbon and nutrients in marine environments. Seagrass detrital decay has conventionally been modelled as multi-component exponential decay. Such models represent the different stages of decay due to bacterial interaction with detritus. Often, however, the size of observed datasets is small, which leads to the over-parameterisation of multi-component exponential decay models. This chapter introduces an alternate method of modelling seagrass decay, which incorporates stochasticity in a single exponential decay model through a random walk. The model presented here is shown to be more parsimonious than those currently used and it has the ability to capture simple or more complex decay dynamics without increasing the parameter space, allowing flexibility in use and biological justification.

3.1 Introduction

Seagrass ecosystems are known to sequester and store atmospheric carbon effectively. Some carbon is stored within the plant tissues, in particular the roots and rhizomes, while most of the carbon stock is stored within the sediments of seagrass ecosystems [35]. Seagrasses are responsible for contributing 50% [57] to 88% [80] of seagrass sediment carbon. Undisturbed, healthy seagrass ecosystems have been shown to store sedimentary carbon for centennial or millennial timescales [35]. A key component to our understanding of the carbon cycle is the fate of the organic material in the sediment as these plants die and their detritus is broken down or remineralised [49, 97].

The rate of detrital decay is known to change through time. Historically, this has typically been described by a multiple-component exponential decay model [36, 90]. This model describes the remaining detritus by the weighted sum of two, or sometimes three, exponential decay curves. The biological justification for this has been that this captures multiple phases of chemical and biological decay. The first of these is a rapid initial phase where soluble materials are quickly leached from seagrass detritus. During the second phase, bacteria are actively utilising relatively more of the recalcitrant plant tissue at an intermediate rate of decay. A third phase is sometimes included in the model [38] to capture an additional slower rate due to the microbial breakdown of a specific form of carbon, lignin.

The resulting models are of the form $\mathrm{W(t)}=\alpha_1 e^{-k_1 t}+\alpha_2 e^{-k_2 t}+\alpha_3 e^{-k_3 t}$ where W(t) describes the remaining detritus at time $\mathrm{t}\geq 0$, k_1, k_2, $\mathrm{k}_3 \geq 0$ are decay constants and $\alpha_j (j \in 1, 2, 3)$ are the proportions of initially-available detritus which decays at rate k_j. This has six free parameters. When the model is simplified to a two-stage decay model (four parameters), this simply collapses by setting $\alpha_3 = 0$.

Although biological justification for such models exist [11, 91], they are not

wholly satisfactory. Most problematic is that this model effectively segregates the pool of decomposing biomass into discrete subsections, each of which decays at a constant rate. Although the proportions are fitted by the model, they are never directly measured, so it remains difficult to know whether their values are accurate, or whether they simply operate as free parameters, which can be adjusted to compensate for poorer fits for other mechanistic parameters.

Mathematically, such models are even more problematic. Because of the destructive nature of data collection, seagrass biomass used in decay studies has to be removed from its habitat and destroyed, therefore experiments of this kind have been limited to small datasets, which are often sparse in time. Consequently, the models used to fit these small datasets have been over-parameterised, often with almost as many free parameters as observed data-points [36]. Furthermore, as many field experiments acknowledge that the decay rate is greatest immediately after their decay begins, studies have tended to gather measurements more frequently initially and more sparsely in time after the initial rapid decay has passed. When fitted with a simple least squares approach, such models are prone to having later data-points with unduly high leverage.

The purpose of this chapter is to estimate the decay rate by proposing an alternative approach to modelling seagrass detrital decay that is more mathematically and biologically flexible. In the following sections, we fit the proposed model to observations from a field study of seagrass decomposition and compare its performance to the conventional multiple-component exponential approach. Then we apply the new time-varying decay rate model to an additional dataset: a laboratory seagrass decomposition incubation experiment studying the effect of temperature and nutrients on decay rates. The data assimilation procedure is applied separately to each field study site and experimental set-up. Finally, we examine the data collection methodologies and discuss the affect experimental design has on the resulting

observations and conclusions that can be drawn from them.

3.2 Methods

3.2.1 Random walk modelling approach for all phases of decay

In addition to fitting the traditional multi-phase exponential decays, here we propose an alternative approach to modelling all phases of seagrass decay. We do this by modifying the original exponential decay model to include a stochastic representation of a microbial process. Rather than relying on a weighted sum of different exponential decays, we describe the rate of detrital decay by a single first order differential equation

$$\frac{dW}{dt} = -kW$$

in time where k is the decay rate. We, however, allow this rate to change through time. To capture the changing nature of this decay rate through time, we replace the simple deterministic model with a random walk stochastic process by setting the decay rate at any time to be a function of the decay rate at the previous time-step. We do this by taking k, previously a constant biogeochemical parameter, and replacing it by K(t). Here, we take K(t) to be such that

$$K(t + \Delta t) = K(t) + r$$

where $r \sim \mathcal{N}(0, \sigma_r)$, and Δt is the length of discrete time-step. For the purpose of the Bayesian analysis here, σ_r is treated as a parameter to be inferred and a prior distribution is prescribed. In the interest of informing the model during time-steps when no observations are available we set a quasi-informative prior on σ_r to be a normal distribution with mean 0.1 and variance 0.01. This was chosen to allow sufficient flexibility for the observations to influence the posterior while the assumption of Gaussian noise essentially imposes an element of smoothness, disallowing large jumps per time-step. To ensure the decay rate did not fluctuate

between positive and negative numbers during the random walk, the solution was computed in log space and transformed back before solving.

Additionally, we assume $\Delta t = 1$ day. We undertook some sensitivity analyses across a range of plausible values (Δt ranging from around 0.5 to 2) and found that the resulting fits showed low sensitivity to changes in time-step.

3.2.2 Set-up of seagrass decomposition field study investigating the impact of three different site conditions

All data collection methods for this chapter were part of a series of experiments examining dynamics of refractory carbon in seagrass meadows (Stacey Trevathan-Tackett 2016 UTS PhD [91]).

Datasets were gathered from three natural field sites in the Brisbane Waters Estuary, Central Coast, NSW, Australia [91]. Two sites contained a *Zostera muelleri* Irmisch ex Ascherson seagrass meadow at Saratoga (33.4732S, 151.3357E), and Fagans Bay (33.4306S, 151.3211E), were classified as Pristine and Impacted, respectively. A third site in Fagans Bay was chosen as a reference bare site, void of seagrass (33.4347S, 151.3228E). Litter bags containing *Zostera muelleri* leaf (wrack) or fresh rhizome and root were deployed in Autumn 2013. Subsequently, seagrass detritus was sampled after 14, 42, 98, 168, 389, 519 and 729 days from litter bags that were buried either 2 cm below the sediment (rhizome and root tissue) or anchored on top of the sediment (leaf tissue) mimicking natural decomposition conditions.

3.2.3 Experimental design of laboratory study investigating the effects of temperature and nutrient addition on time-varying decay rates

An additional dataset was generated during a laboratory-based experiment to investigate the effects of temperature and nutrient addition on seagrass decomposition [91]. The motivation for a temperature treatment was to examine how detritus-

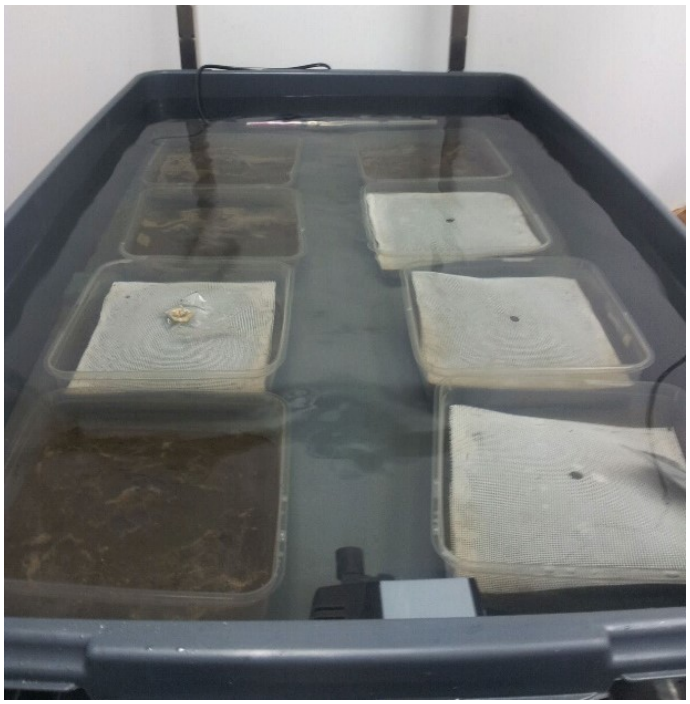

Figure 3.1 : Experimental set-up: incubation containers mimicking natural decay conditions.

bacteria interactions and resulting decay rates respond to predicted rising seawater temperatures. Two temperature treatments were chosen based on current temperature data for Fagans Bay (NSW Australia); the chosen data collection site. An ambient temperature of 23°C was the annual mean, and a high temperature of 30°C was approximately the average maximum during summer [91].

To assess the impact of eutrophication on seagrass decomposition, two nutrient treatments were chosen. A control level and an inorganic nutrient addition (+Nutrient) level. The sediments collected from Fagans Bay had high organic matter content (OM>9%) in the top 10 cm [91]. The following section will describe how the sampled detritus and sediment were used in the laboratory to try and replicate field conditions.

A laboratory seagrass decomposition incubation study was conducted with sea-

grass tissue and sediments collected from a *Zostera muelleri* Irmisch ex Ascherson meadow in Fagans Bay. Two types of seagrass tissue were collected, leaf and root/rhizome material. To mimic natural decomposition conditions, incubation containers (225 cm^2) filled with 3.5-4.0 cm of sediment had leaf litterbags deployed on top of the sediment and root/rhizome litterbags buried under 1 cm of sediment (Fig. 3.1). 1 g/L of controlled-release fertiliser pellets were added to the sediments assigned the +Nutrient treatment. The temperature loggers recorded 23.5 ± 2°C for the ambient temperature treatment, and 30 ± 1°C for the high temperature treatment controlled with aquatic heaters.

Leaf material was collected as fresh wrack (still green) along the shoreline, while root/rhizome material was collected as live plants. After being cleaned of sediments and attached fauna, all seagrass material was stored frozen prior to the start of the experiment to prevent microbial degradation. This seagrass biomass is what we refer to as fresh seagrass and subsequently its weight in grams as fresh weight (g FW).

Whole leaves were separated into groups of approximately 78 g FW, while 2-4 cm segmented roots and rhizomes were separated into groups of approximately 30 g FW and each group was packed into a nylon mesh litterbag. The initial fresh weight measurement was recorded for 160 litter bags respectively to accommodate an experimental design of two tissue types (leaf, root/rhizome) x four treatments (23°C and control, 30°C and control, 23°C and +Nutrient, 30°C and +Nutrient) x five time-points (2, 7, 14, 28, 84 days) x four replicates.

The initial time-point (day 0) was treated separately to the rest of the data and a conversion factor used to calculate the corresponding dry weight of each litterbag's initial fresh weight. This was achieved by separating additional groups of tissue (n = 3 for leaf material and n = 1 for root/rhizome material) where no treatments were applied, and weighing them before and after drying to get a fresh weight to dry

weight conversion. All initial fresh weight biomass measurements were multiplied by their respective tissue conversion factors, 0.2 for leaf and 0.153 for root/rhizome.

At every sampled time-point, four litterbags from each temperature, nutrient and tissue combination were removed from their incubators, weighed as fresh weight, then oven dried at 60°C and weighed post-drying to get their dry weight (g DW) biomass. To get the proportion of weight remaining (W), each decayed litterbag dry weight biomass was divided by the litterbag's initial converted dry weight at day 0, then averaged over the four replicates.

3.3 Results

3.3.1 Proof of concept results for the random walk model compared to multi-phase exponential decay of a seagrass decomposition field study

For the proposed random walk model, we implemented the Particle Marginal Metropolis Hastings (PMMH) algorithm using LibBi [72]. LibBi is a state-space modelling and Bayesian inference software designed for use on high-performance computer hardware, and used in a number of environmental modelling studies [47, 76]. Each posterior run had 50,000 samples where the first 10,000 were discarded as burn-in. 1024 particles were propagated and the proposal distribution variances were set to 10% of each parameter prior variance. Acceptance rates ranged between 41-59% (Table B.2) with observation errors of 10%. Posterior medians and 95% credible intervals of remaining detritus were well constrained by the observations for each site and tissue type (Figure 3.2). The decay rate posterior medians and 95% credible intervals varied through time, across site and tissue type (Figure 3.3). While the prior and posterior distributions (Figure 3.4), and posterior quantiles (Table 3.1) of parameter σ_r illustrated that we suitably captured the stochasticity for these datasets. If we had misspecified the prior on σ_r, we would have seen a posterior

distributed at much higher values of σ_r, indicating that a larger stochasticity was needed.

For the existing multi-phase exponential decay model, this was fitted using a simple least squares approach. Both fits were then evaluated by examining the resulting R-squared values (Table 3.2). In general, it should not be assumed that a higher R-squared value is indicative of a "better" model as simply adding additional parameters can increase the R-squared even without any meaningful improvement in the fit. Indeed, in this case, the R-squared for a three phase exponential decay model would necessarily be no lower than that for a two phase, even if there truly were only two phases of decay. Here, however, we do not require a model selection criterion which penalises the goodness of fit for the inclusion of additional degrees of freedom. For all six datasets examined here, the proposed random walk model produced higher R-squared values (Table 3.2). The model is not only more parsimonious, it also better captures the observed variability in the remaining detrital pools.

All site and tissue combinations except for leaf tissue at the impacted site exhibited an initial increase in decay rate over time followed by a slow decrease towards zero, forming a visible peak in decay rate (Figure 3.3). The leaf tissue decay rate at the impacted site was the only site to exhibit two peaks, an initial peak at 120 days with a second peak at 500 days. The leaf tissue decay rate at the pristine site peaked at 150 days compared to the bare site peak later at 250 days. There is no visible difference of the root/rhizome tissue decay rates between sites, with each displaying a peak at approximately 150 days.

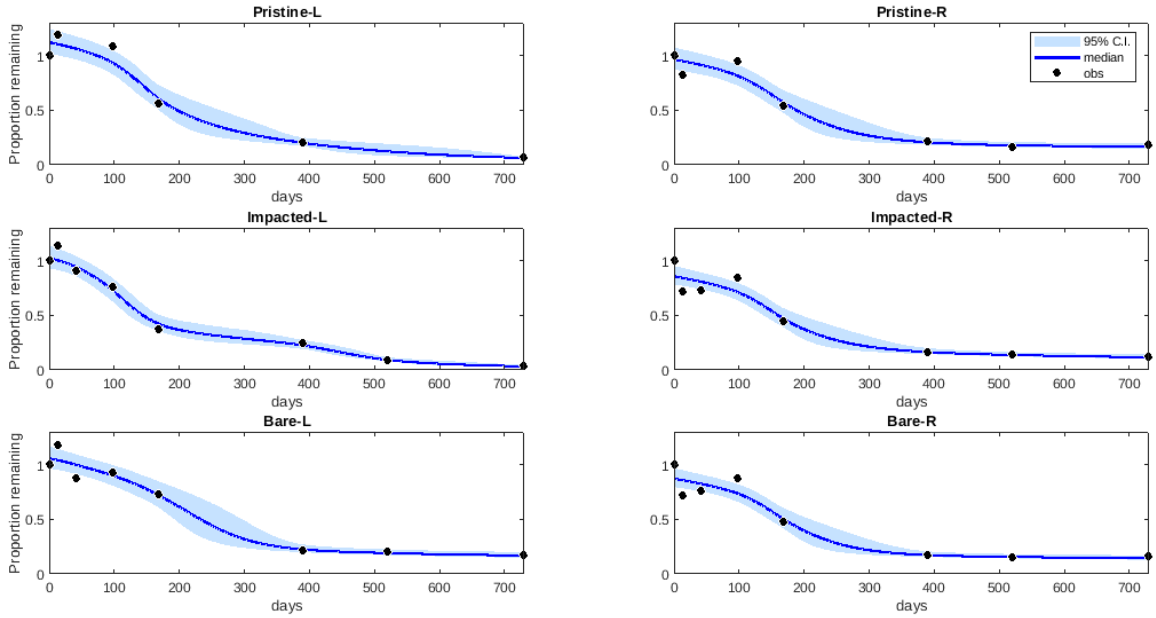

Figure 3.2 : Posterior medians (solid blue line) and 95% credible intervals (shaded) of proportion of weight remaining detritus (W), with observations (black) at each site for leaf (left) and root/rhizome (right) tissue.

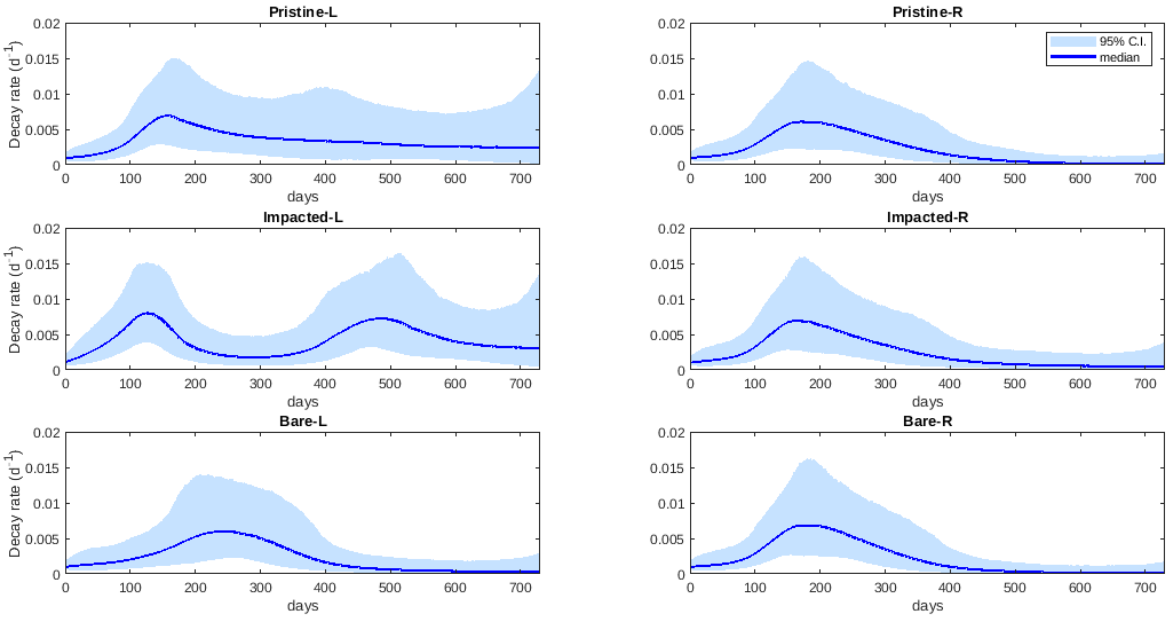

Figure 3.3 : Posterior medians (solid blue line) and 95% credible intervals (shaded) of decay rate K (d^{-1}) at each site for leaf (left) and root/rhizome (right) tissue.

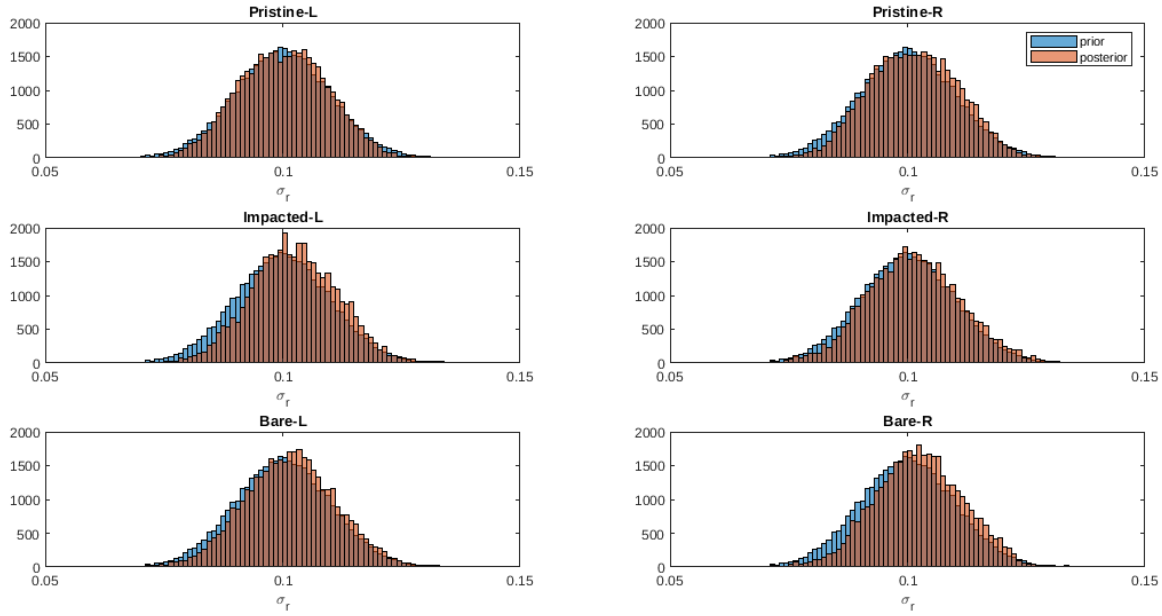

Figure 3.4 : Prior (blue) and posterior (orange) distributions for parameter σ_r at each site for leaf (left) and root/rhizome (right) tissue.

Table 3.1 : Posterior quantiles for parameter σ_r across tissue type and site.

Parameter	**Tissue**	**Site**	**2.5%**	**25%**	**50%**	**75%**	**97.5%**
σ_r	Leaf	Pristine	0.079	0.094	0.101	0.109	0.121
		Impacted	0.085	0.095	0.101	0.108	0.121
		Bare	0.081	0.094	0.102	0.110	0.121
	Root/	Pristine	0.084	0.094	0.100	0.107	0.119
	rhizomes	Impacted	0.082	0.096	0.103	0.109	0.120
		Bare	0.086	0.097	0.104	0.109	0.121

Table 3.2 : Coefficients and R-squared values for the existing multi-phase exponential decay and R squared values for the random walk model.

	Leaf			**Root/rhizomes**		
	Pristine	**Impacted**	**Bare**	**Pristine**	**Impacted**	**Bare**
α_1	0.000	0.060	0.000	0.193	0.067	0.131
α_2	0.387	0.512	0.495	0.000	0.471	0.038
α_3	0.613	0.428	0.505	0.807	0.462	0.830
k_1	0.081	0.004	0.000	0.003	0.000	0.990
k_2	0.003	0.004	0.003	0.003	0.005	0.003
k_3	0.003	0.004	0.003	0.003	0.005	0.003
R^2	0.840	0.948	0.909	0.906	0.879	0.891
R^2 random walk model	0.953	0.980	0.966	0.959	0.925	0.924

3.3.2 Data assimilated results of time-varying decay rates for an additional laboratory dataset investigating the effects of temperature and nutrient addition

Applying the random walk model described in Section 3.2.1 to the additional laboratory dataset, we implemented the Particle Marginal Metropolis Hastings (PMMH) algorithm using LibBi [72]. Each posterior run had 50,000 samples, where the first 10,000 were discarded as burn-in. 1024 particles were propagated and the proposal distribution variances were set to 50% of each parameter prior variance, thus achieving good MCMC mixing (Fig. B.3). The prior on σ_r was assigned a normal distribution with mean 0.1 and variance 0.01 for leaf tissue. A larger stochasticity prior was assigned to the root/rhizome tissue, a normal distribution with mean 0.7 and variance 0.1, to capture the rapid mass loss over the first two days (Table B.1).

Similarly to Section 3.3.1, we set $\Delta t = 1$ day. Acceptance rates ranged between 27-39% with observation errors of 10% (Table B.1). Posterior medians and 95% credible intervals of remaining detritus were well constrained by the observations for each temperature, nutrient and tissue combination (Figure 3.5). The decay rate posterior medians and 95% credible intervals varied through time, across site and between tissue type (Figure 3.6). While the prior and posterior distributions (Figure 3.7), and posterior quantiles (Table 3.3) of parameter σ_r illustrate that we suitably captured the stochasticity for these datasets.

Leaf tissue results with no additional nutrients, show an increasing decay rate through time at high temperature compared to the lower and approximately constant rate of decay at ambient temperature (Figure 3.6). Additional nutrient treatments on leaf tissue results show no noticeable differences between ambient and high temperatures. No differences in root/rhizome tissue results can be seen between temperature or nutrient treatments. The largest difference in decay rates

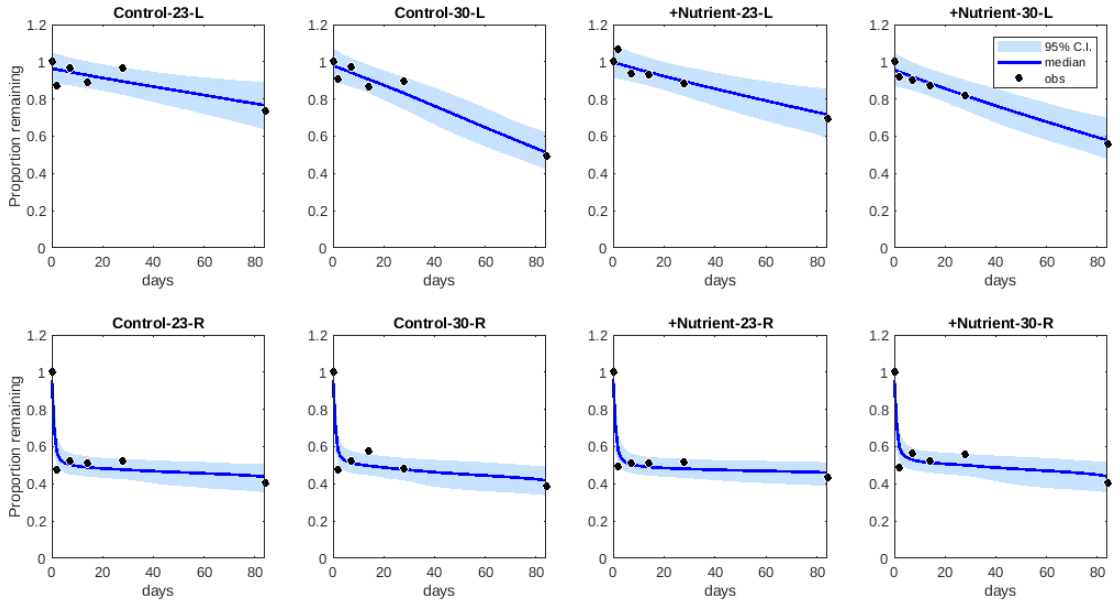

Figure 3.5 : Proportion of weight remaining (W) over the length of the experiment (80 days), for each temperature, nutrient and tissue combination.

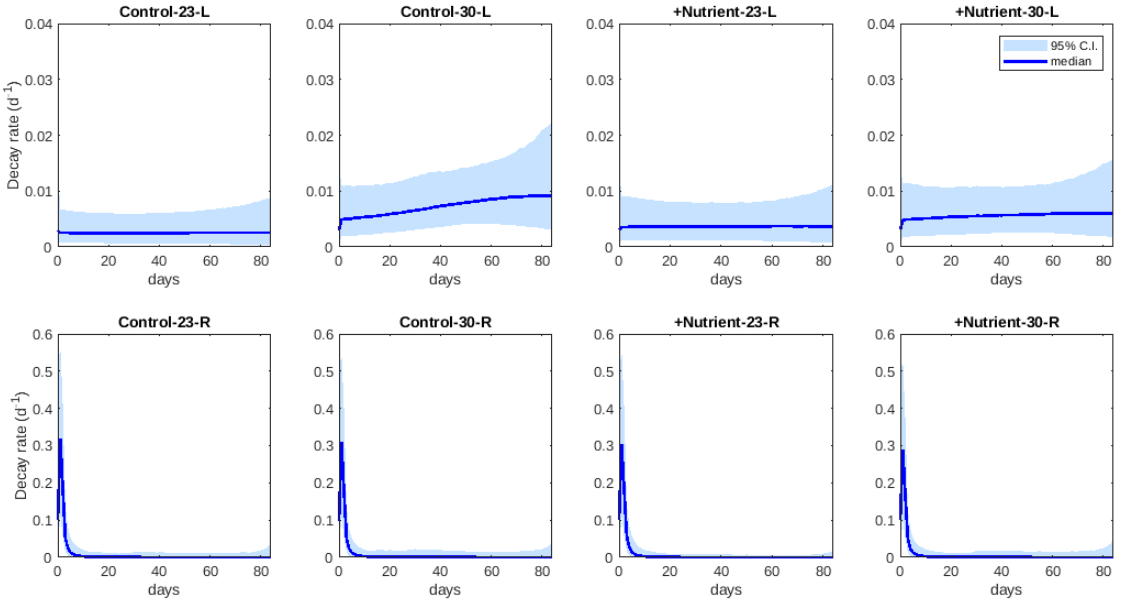

Figure 3.6 : Decay rates, K (d^{-1}), through time for each temperature, nutrient and tissue combination.

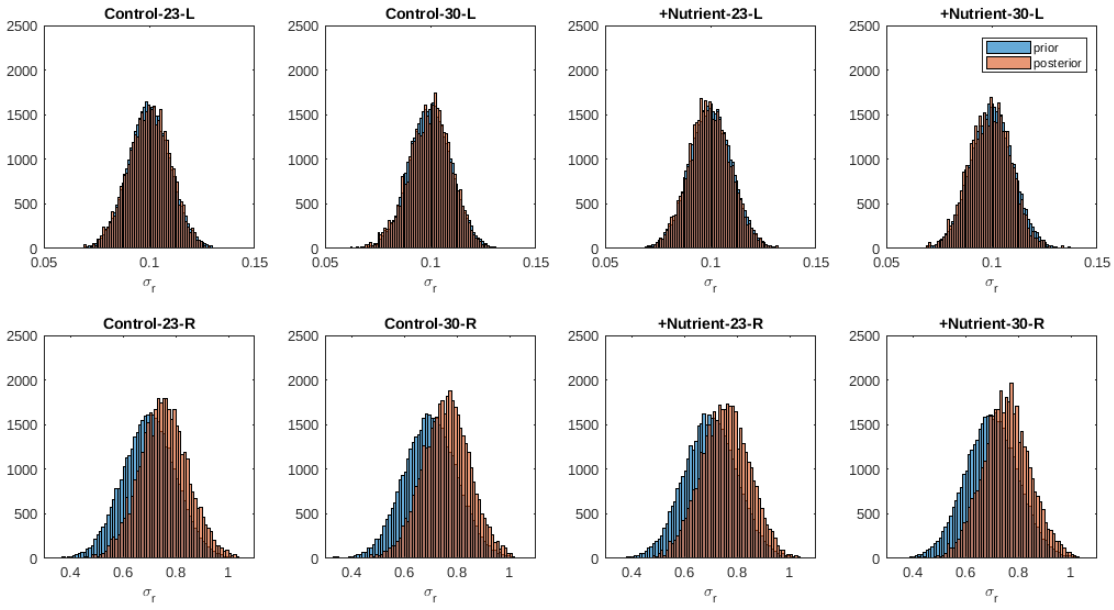

Figure 3.7 : Stochasticity parameter, σ_r, for each temperature, nutrient and tissue combination.

is visible between tissue types. Figure 3.5 and 3.6 demonstrate the random walk model's flexibility in capturing two distinct behaviours of decomposition. Leaf tissue that remains roughly constant over the course of the decomposition experiment, while root/rhizome decay spikes rapidly upwards in the first two days and then flattens close to zero for the remainder of the experiment. Both responses are readily handled by the model.

Table 3.3 : Posterior quantiles for parameter σ_r across tissue type, treatment and temperature.

Tissue	**Treatment**	**Temperature**	**2.5%**	**25%**	**50%**	**75%**	**97.5%**
Leaf	Control	23°C	0.079	0.093	0.100	0.107	0.119
		30°C	0.079	0.093	0.100	0.107	0.119
	+Nutrient	23°C	0.081	0.093	0.100	0.106	0.119
		30°C	0.080	0.092	0.099	0.106	0.119
Root/	Control	23°C	0.575	0.693	0.753	0.815	0.937
rhizomes		30°C	0.581	0.700	0.761	0.820	0.937
	+Nutrient	23°C	0.573	0.692	0.755	0.817	0.932
		30°C	0.580	0.695	0.754	0.814	0.936

3.4 Discussion

Overall, the time-varying decay rate model demonstrated the ability to capture 3 distinct decay rate profiles through time. Soft peaks with slow increase and decreases before and after in the field study (Figure 3.3). Approximately constant and linearly increasing decay rates in the experiment leaf tissue, and a sharp spike and rapid decrease in decay rates that then flat line close to zero in the experiment root/rhizome tissue (Figure 3.6). This demonstrates the flexibility of the framework to capture diverse and complex decay dynamics which would have otherwise been lost in the traditional multi-component model.

The field study saw a decay rate peak earliest at the impacted site with leaf tissue, 30 days later for leaf tissue at the pristine site and all root/rhizome tissue, then 100 days after that for leaf tissue at the bare site. All site and tissue combinations except for leaf tissue at the impacted site exhibited an initial increase in decay rate

over time followed by a slow decrease towards zero, forming a visible peak in decay rate (Figure 3.3). The leaf tissue decay rate at the impacted site was the only site to exhibit two peaks, an initial peak at 120 days with a second peak at 500 days. The leaf tissue decay rate at the pristine site peaked at 150 days compared to the bare site peak later at 250 days. There is no visible difference of the root/rhizome tissue decay rates between sites, with each displaying a peak at approximately 150 days.

The aim of the laboratory experiment was to discern whether temperature, tissue or nutrient addition have an impact on the detrital decay of *Zostera muelleri* in Fagans Bay. With warming sea temperature predictions, it would be expected that decomposing seagrass would spend longer periods of time under elevated temperature conditions. Leaf tissue results with no additional nutrients, show an increasing decay rate through time at high temperature compared to the lower and approximately constant rate of decay at ambient temperature. Additional nutrient treatments on leaf tissue results show no noticeable differences between ambient and high temperatures. No differences in root/rhizome tissue results can be seen between temperature or nutrient treatments. Indeed, the largest difference in decay rates is visible between the tissue types (Figure 3.6). However, it is difficult to draw this conclusion for a number of reasons.

Potentially the quick initial leeching stage of decay was missed in the leaf tissue data, because wrack was collected instead of live leaf material. Leaf tissue was collected as green wrack from the shoreline indicating it was already removed from the plant for 1-2 days before collection. The decay rate spike for root/rhizomes in the first two days corresponds to halving the proportion of biomass remaining by day 2 of the experiment. Data collection occurred on day 0 and day 2, and without finer scale observations, we are unable to pinpoint when or how the decay behaves in this early stage, only that by day 2 half the seagrass detritus has decomposed.

When wrack versus live plant material is considered, this affects which time-points in the degradation process we can consider to be comparable. For example, the root/rhizome tissue after day 2 may correspond to the start of the intermediate recalcitrant phase of decay after a quick leeching phase, whereas leaf tissue day 0 may already be in the recalcitrant phase as the wrack's age is unknown.

The initial time-point (day 0) was treated separately to the other time-points and did not receive any temperature or nutrient treatment. The root/rhizome initial fresh weight to dry weight (FW:DW) conversion factor was determined by a sample size of n=1. Consequently, the FW:DW conversion factor impacts the remainder of the dataset by shifting it entirely upwards when FW:DW is smaller, alternatively shifting values downwards when FW:DW is larger (Figure 3.8). While the field study employed the same data collection methodology as the laboratory study, we do not see the same differentiation in results; however, there are differences in the initial FW:DW conversions. In the field study, the FW:DW conversions were very similar when comparing the initial time-point (leaf: 0.147 ± 0.015; root/rhizome: 0.169 ± 0.018) to the rest of the time-points (leaf: 0.143 ± 0.022, root/rhizome: 0.161 ± 0.030), while the laboratory study showed notably different results between the initial time-point (leaf: 0.153 ± 0.056; root/rhizome: 0.201) and the other time-points (leaf: 0.137 ± 0.018, root/rhizome: 0.146 ± 0.024). Had the conversions been more like those in the post collection data, we would have seen data-points like the magenta in Figure 3.8.

While part of the uncertainty in interpreting the decay rates between leaf and root/rhizome tissue lies in the nature of their data collection, it is important to consider the practical implications of this sampling technique. For live plant material (roots/rhizomes) this is a destructive sampling technique and consequently there are limitations (collection permits) to how much can be collected by disturbing a meadow. In contrast, wrack is already shed leaf material that has washed up on

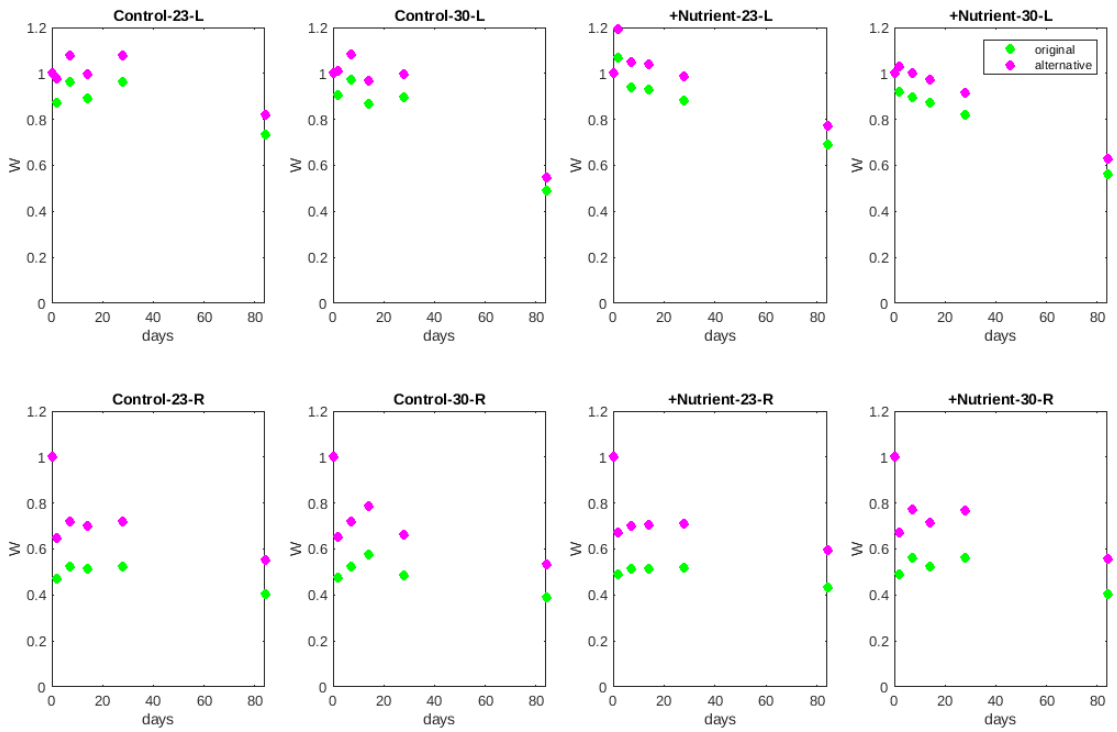

Figure 3.8 : Proportion of weight remaining (W) when the original and alternative fresh weight to dry weight conversion was applied to each treatment group.

shore and is not included in the same quantity collection restrictions as the live material.

The work presented here should be regarded as a proof-of-concept that, in the presence of data, the complex decay dynamics of seagrasses can be captured without reliance on over-parametrised models or excessive numbers of free parameters. Mathematically, the alternative we present is more parsimonious. It relies on fitting one free parameter (σ_r), not six, and is relatively insensitive to time-step (Δt). Additionally, we would also argue that this model has greater biological justification and does not rely on explanations involving community fractions which are unmeasurable.

This framework is sufficiently flexible to capture more complex decay dynamics as the parameter space would not grow as building additional phases into a multi-stage decay model does. Also, where additional explanatory variables were available, the fits of this model could easily be studied in reference to these. For example, when this model has identified periods of more rapid decay, these might be identified as co-occurring when spikes in temperature or in rainfall happen. While we have not been able to do so in this work, largely because of the relative sparsity of the dataset available to us, this remains an exciting avenue for future work and one which this modelling approach would readily support without necessitating the inclusion of further parameters.

Chapter 4

A data assimilating state-space model for algal growth under controlled conditions within a photo-bioreactor

Abstract

Due to the devastating impacts of fossil fuels on the environment, clean energy and renewable sources of biofuels are two major areas of interest. Microalgae are seen as excellent candidates for sustainable biofuel production and high biomass cultivation can be used to potentially mitigate carbon dioxide and other greenhouse gases as microalgae photosynthesise and grow. The key to scaling microalgae production is in the successful fusion of high temporal frequency non-destructive data, such as sensors, and a microalgae model. This chapter used data assimilation to fully exploit the use of non-destructive high temporal resolution data and a well defined model to produce ecologically informative estimates of algal growth by way of photosynthesis and respiration. A key achievement of this chapter was successfully and accurately resolving CO2SYS in LibBi to calculate carbon chemistry and accurately depicting the carbon state of the system. This methodology allows LibBi and other Bayesian inference software, designed for parallel processing and simplicity limitations because of their use of GPU, to expand in use to more complex models and equations.

4.1 Introduction

Due to the devastating impacts of fossil fuels on the environment, there is great interest around the world in clean energy, particularly renewable sources of biofuels. Microalgae are great candidates for sustainable production of biofuels and associated bioproducts. In addition to converting microalgal biomass to fuel, high biomass cultivation can be used to potentially mitigate carbon dioxide and other greenhouse gases as microalgae photosynthesise and grow.

The key goal of biofuels production is the optimisation of biomass productivity in large-scale microalgal culturing systems such as open ponds or closed photo-bioreactors. Carbon and light availability are two of the most common limiting factors of biomass productivity [78]. Australia offers natural advantages for algal cultivation, including extensive areas of non-arable land, unrestricted access to sea-water and plenty of natural light. The primary goal is to maximise the production of raw biomass (i.e. fast growth rates) coupled with high lipid production to produce biofuels in a sustainable and profitable way.

Microalgae has long been viewed as a potential platform for bioengineering. During the 1970's and 2000's much of the research focussed on biofuel production, while recently the potential for pharmaceuticals and other high-value chemicals has been more actively explored. Whether the desired product is a primary metabolite (eg lipids in biofuel production) or secondary metabolite (eg astaxanthin as high-value dye) it is important to be able to quantify processes such as photosynthesis and respiration that take place in the cells.

At the broadest scale, the growth of microalgae can be expressed as the following chemical reaction in which carbon (C) is taken from the atmosphere as carbon dioxide (CO_2), along with aqueous nutrients nitrogen (N) and phosphorus (P), to

produce biomass

$$106\,CO_2(g) + 122\,H_2O + 16\,NO_3{}^-(aq) + PO_4{}^{3-}(aq) + 19\,H^+ \xrightarrow{[10^3 - 10^4]\,\gamma} C_{106}\,H_{263}\,O_{110}\,N_{16}\,P_1 + 138\,O_2\,(g)$$

This equation captures a number of features of interest. The total biomass is determined by the concentrations of N and P. A typical growth media for microalgae is Guillard's Marine Enriched Seawater (F/2), which has nitrogen concentration of 880 μM, phosphorus concentration of 36 μM, and total alkalinity (TA) of 2300 μM. Taking the limiting nutrient to be nitrogen, the maximum biomass of algae is 200 mg L^{-1}, of which approximately 70 mg L^{-1} is carbon. This is equivalent to 0.12 L of pure CO_2 at standard temperature and pressure (25 °C at 100 kPa), which approximates to 325 L of air for each litre of media. Alkalinity is a central concept in seawater chemistry that allows one to calculate how much carbon can be dissolved in solution, and what form it takes i.e. dissolved inorganic carbon (DIC) exists as CO_2, $HCO_3{}^-$ and $CO_3{}^{2-}$. Alkalinity measures the charge imbalance between strongly and weakly dissociating ions in solution. As the charged nutrients (N and P) are removed from the solution during growth the alkalinity increases and more carbon is able to be dissolved in the media.

The key to scaling microalgae production is in the successful fusion of a microalgae model and high temporal frequency non-destructive data, such as sensors which are able to collect information without the need to remove or destroy parts of the system being measured. Data assimilation is a prime candidate for this type of model-data fusion technique to fully exploit the use of non-destructive high temporal resolution measurements and produce ecologically informative estimates of algal growth by way of photosynthesis and respiration.

The key goals of this chapter are to,

- Use data assimilation to fuse a microalgae model with both high temporal frequency observations and infrequently sampled datasets to produce informative estimates of algal growth by way of photosynthesis and respiration.
- Fully exploit the use of non-destructive high temporal resolution measurements.
- Successfully and accurately resolve CO2SYS in LibBi to calculate carbon chemistry and accurately depict the carbon state of the system.

4.2 Methods

The work in this chapter was completed in collaboration with Christian Evenhuis (UTS) and extends on a recently published empirical process model to predict microalgal carbon fixation rates in photobioreactors, with an application for *Chlorella vulgaris* [88].

4.2.1 Data Model: Photo-bioreactor setup, experimental design and data collection methods

All data collection methods for this chapter were part of a series of experiments examining microalgal responses to photobiorector treatments (Peter Wood 2019 UTS PhD).

Microalgal culture *Nannochloropsis oceanica* (Droop) Green (strain CS-179) obtained from the Australian National Algae Culture Collection was cultured in 200 mL conical flasks; maintained in an incubator (Labec Pty Ltd) at 20°C, under an irradiance of 50 μmol photon m^{-2} s^{-1} of cool-white fluorescent light at a 12 hour light/12 hour dark cycle. Stock cultures were grown in F/2 saltwater medium [46] and diluted 5 days prior to the start of experiments to ensure that *N. oceanica* was in the

exponential growth phase and not nutrient deprived. F/2 was sparged prior to stock culture dilutions to maximise carbon and oxygen content.

N. oceanica was cultured in four, 500 mL environmental photo-bioreactors (ePBRs, Phenometrics Inc) with a 10% v/v inoculation of stock culture. Top-side illumination over a path length of 25 cm was provided by a cool-white light LED, whilst temperature was maintained at 27°C using a Peltier heater-cooler connected to a water jacket. In-built thermocouples, calibrated against external temperature sensors attached to the Firesting module (TeX4; PyroScience GmbH), measured every 5 minutes were used to control the Peltier heater-cooler jacket through a feedback loop to an accuracy of ± 0.2°C. pH was also measured in 5 min intervals by in-built pH electrodes (Van London Inc); controlled by periodic CO_2 (5%) injections using valves in the ePBRs. pH was 3-point calibrated using pH buffer solutions at pH 4.00 ± 0.02, pH 7.00 ± 0.02 and pH 10.00 ± 0.02. ePBR mixing was controlled by magnetic stirring bars at 110 rpm. All four ePBRs were aerated with filtered/humidified air through a 1.2 mm needle valve (Terumo Co).

A period of 2 days was allowed for *N. occulata* to acclimate to the ePBRs at an irradiance of 500 μmol photon m^{-2} s^{-1} and a temperature of 27°C. Following this acclimation period, the ePBR was set to the experimental condition of 2,000 μmol photon m^{-2} s^{-1} for another 2 days and a 12 hour light/12 hour dark cycle with a temperature of 27°C. ePBRs were maintained at an optical density (OD) of 0.4 using manual dilutions, creating a semi-batch culturing system. Dilutions occurred once per day (one hour before the commencement of the light cycle), using aerated F/2 media. The experiment was conducted over a period of 4 days, samples were extracted post and prior dilution, as well as half way through the light cycle. 50 mL was extracted to examine total alkalinity (TA) and dissolved inorganic carbon concentration. Dissolved oxygen (O_2) was measured using a 3 mm robust optical probe (OXROB10-OI; PyroScience GmbH) attached to a FireStingO2

logger (PyroScience GmbH). DO measurements were taken every 60 seconds and temperature-corrected using a temperature extension module (TeX4; PyroScience GmbH). DO was two-point calibrated using air-saturated seawater (100% saturation) and sodium sulfate-saturated water (0% saturation). At 2 hour intervals, a solenoid valve (SMC Pneumatics Pty. Ltd.) was used to stop aeration for 10 minutes to allow for observations of net photosynthesis (visible as the spikes in O_2 data in Figure 4.1).

Alkalinity and DIC was measured twice a day closely following the Standard Operating Procedures (SOP) outline in [27]. Approximately 30 mL of the *N. oceanica* media was titrated against 0.1 M hydrochloric acid on an auto-titrator (800 Dosino; Metrohm AG) [SOP3b Open-cell tritration]. The total alkalinity and DIC were calculated from the output of the auto-titrator (volume of HCl delivered, pH) by calculating the the pH as a function of the volume of the acid delivered (see SOP3a Annexe 1).

4.2.2 Data model: Data treatment, distributions and measurement error

Gas valve, temperature, light (normalised to 0/1) and dilution rates were used to force the model. Dissolved oxygen, pH, dissolved inorganic carbon and total alkalinity observations for 4 days post acclimation were assimilated. While pH observations were calibrated and corrected, O_2 observations were not despite having visibly experienced some sensor drift during the experiment.

The data model assigned log normally distributed observation errors for each instrument; $O_{2_{obs}} \sim \text{Log}\mathcal{N}(\log(O_2), \sigma_{O_2})$, $pH_{obs} \sim \text{Log}\mathcal{N}(\log(pH), \sigma_{pH})$, $DIC_{obs} \sim \text{Log}\mathcal{N}(\log(DIC), \sigma_{DIC})$, $TA_{obs} \sim \text{Log}\mathcal{N}(\log(TA), \sigma_{DIC})$, where the standard deviations (σ_{O_2}, σ_{pH}, σ_{DIC}) were unknown parameters to be estimated as part of the assimilating model. Dissolved inorganic carbon and total alkalinity measurements

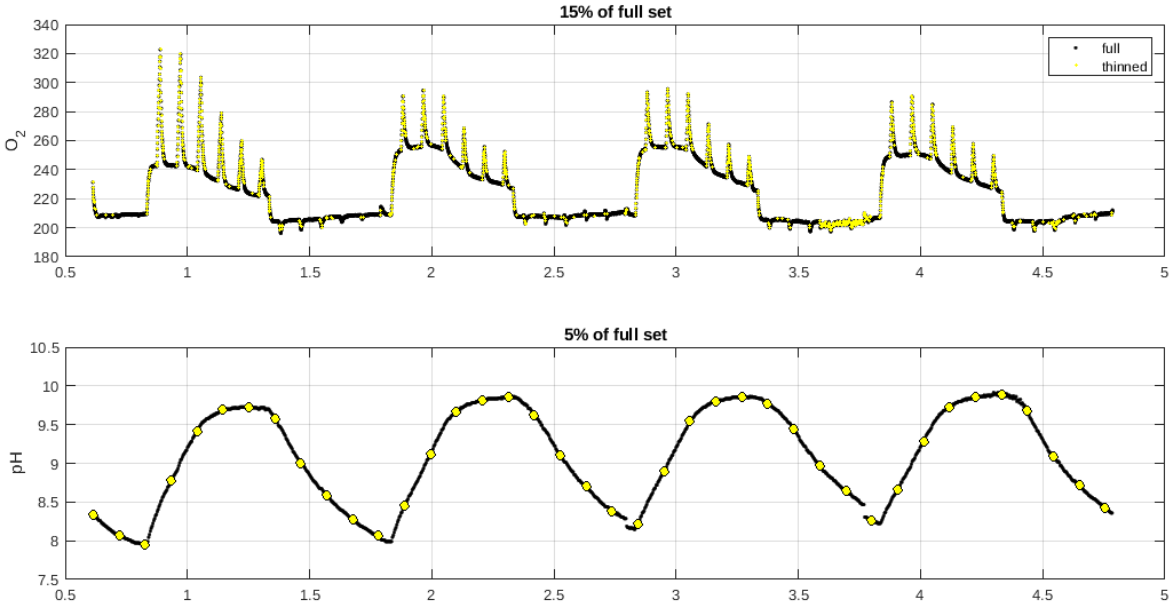

Figure 4.1 : Full O_2 (μM L^{-1}) and pH ($\log_{10}$(-mol/L H+)) datasets (black) with thinned O_2 and pH observations (yellow) over the four day experiment.

were obtained from the same instrument, thus the error is shared between these states.

Due to different instruments used in data collection, O_2 observations were collected most frequently (6,008 data-points), followed by pH observations (1,179 data-points), with DIC and TA collected intermittently (11 data-points). Initial posterior runs with these observation densities resulted in the MCMC chain unable to mix properly due to the high frequency of observations. In an attempt to improve MCMC mixing, the denser observation sets were thinned/sampled down. The pH observations were thinned to approximately 5% of the full pH dataset by taking every 30^{th} observation. The O_2 observations were thinned to approximately 15% of the original O_2 dataset by only taking those consecutive observations that had at least 1 μM L^{-1} difference (Figure 4.1).

4.2.3 Process model: Carbon chemistry

When carbon dioxide dissolves in water it undergoes a series of reactions which results in three species that exist in equilibrium, CO_2, HCO_3^- and CO_3^{2-}, accompanied by the release of hydrogen ions. CO2SYS is a program developed for calculating and returning a detailed state of the carbonate system of oceanographic water samples in seawater and freshwater [63]. This chemical system has been the focus of extensive research as it plays a central role in understanding the effect of anthropogenic carbon emissions and is generally referred to as "Ocean Acidification". The dependence of these equilibrium constants on temperature and salinity has been experimentally determined, functional forms have been fitted and integrated into CO2SYS. A thorough explanation of CO2SYS, and seawater carbon chemistry in general, can be found in Zeebe and Wolf-Gladrow [102]. A concise summary of the equations can be found in SOP 3a, Annexe 1 of [27].

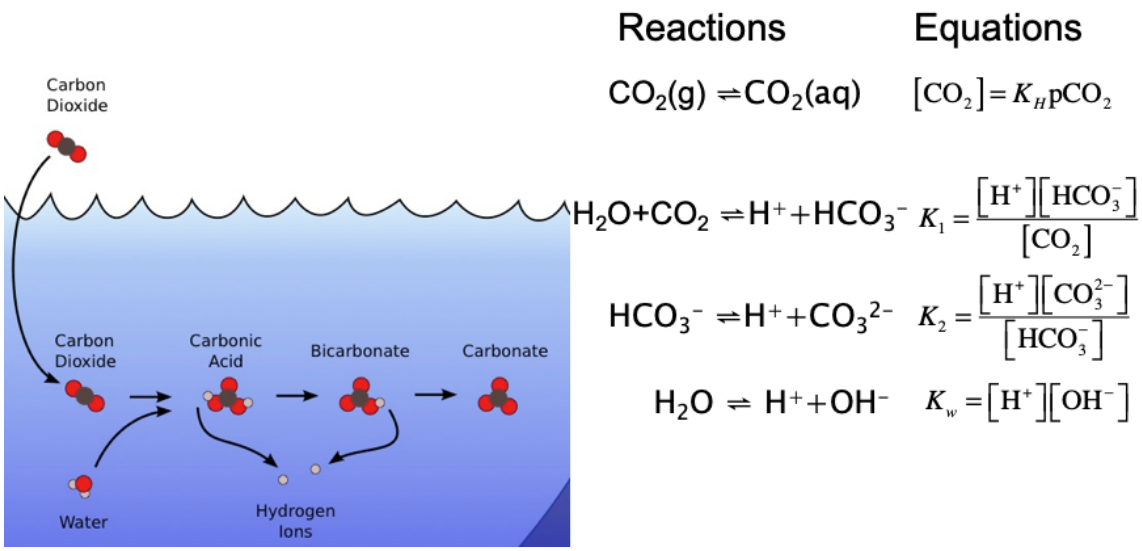

Figure 4.2 : CO2SYS

CO2SYS is recognised as the de facto standard for carbon system calculations and is widely used in oceanographic and marine biology. There are four measurable

carbonate system parameters (total alkalinity, total inorganic CO_2, pH, and either fugacity fCO_2 or partial pressure of CO_2). Given two parameters, the other two can be calculated for a set of input conditions (temperature and pressure). In the model, the state variables TA and DIC are given and pH and pCO_2 are calculated.

Ideally CO2SYS [63] would have been used to calculate the carbon chemistry of the photo-bioreactor. However, it is not possible to directly implement CO2SYS in LibBi due the constraints imposed by the GP-GPU computing model. CO2SYS iteratively solves the set of chemical equilibrium equations which requires a for/while loop and conditional statements like if/else which are not available in LibBI. A number of approaches were explored to circumvent this restriction:

1. Creating functional approximations of the solutions from CO2SYS.

2. Recasting the steady state equilibrium as a set of coupled ODEs.

3. Manually 'unrolling' the while loop as fixed number of iterations.

In addition to being arduous, it was not possible to achieve the accuracy needed via functional approximation. In particular, ensuring that the invariants of the system (such as non-negativity and that the sum of the carbon species equals the dissolved inorganic carbon) is not trivial. The set of ODEs that describe the chemical reactions are notoriously stiff [101] which resulted in a prohibitively small step size. Coding a fixed number of iterations also presented issues because when the number of iterations was too low the solution was inaccurate, while too many iterations was inefficient and may be limited by memory available to the process on the GPU.

CO2SYS was successfully incorporated into LibBi by a combination of the first and third approaches. The functional approximation from approach 1 was used to precondition the solution and 3 iterations were explicitly coded (Eq. 4.4 - 4.12) of the Newton-Raphson method for finding approximations to roots of real valued

functions. The Newton-Raphson method is an iterative process $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$ considering a function $f(x_n)$, its derivative $f'(x_n)$ and an initial starting value x_0. The approximate root x_{n+1} converges to the exact solution very quickly if a close initial starting value is picked. To ensure the quick convergence of the Newton-Raphson method, an approximating equation for pH_0 (Eq. 4.3) was obtained by fitting a stepwise regression with interactions to a range of simulated CO2SYS input parameters (temperature: 20-30°C, salinity: 30-40 ppt, DIC: 200-2500 μM L^{-1}, and alkalinity: 1500-3000 μM L^{-1}). A range of initial conditions and parameter values were tested, and each converged with root mean squared error (RMSE) < 0.01 across pH, HCO_3^-, CO_2, and CO_3, DIC, O_2, TA by the 3rd iteration (Figure 4.3 and Table 4.1).

CO2SYS constants and iterative solution for pH, HCO_3^-, CO_2, and CO_3

This section details the exact CO2SYS solution implemented into LibBi, with choice of constants, approximating equations and Newtown-Raphson iterations to ultimately calculate the carbon chemistry of the photo-bioreactor.

Total Sulfur is defined:

$$
\begin{aligned}
TS &= \frac{0.14}{96.062} * \frac{S}{1.8065} \\
IS &= 19.924 \frac{S}{(1000 - 1.005S)} \\
KS_{int} &= -\frac{4276.1}{T_K} + 141.328 - 23.093 log(T_K) + (-\frac{13856.0}{T_K} + 324.57 \\
&\quad - 47.986 log(T_K))\sqrt{IS} + (\frac{35474}{T_K} - 771.54 + 114.723 log(T_K))IS \\
&\quad - \frac{2698}{T_K} IS^{1.5} + \frac{1776}{T_K} IS^2 \\
KS &= (1 - 0.001005S) e^{(KS_{int})}
\end{aligned}
$$

Fluorine is defined:

$$TF = \frac{\frac{0.000067S}{18.9984}}{1.80655}$$
$$KF = e^{(-(-\frac{874.0}{T_K} - 0.111\sqrt{S} + 9.68))}$$
$$SWS_{2_T} = \frac{(1 + \frac{TS}{KS})}{(1 + \frac{TS}{KS} + \frac{TF}{KF})}$$
$$Free_{2_T} = 1 + \frac{TS}{KS}$$

H_2O dissociation:

$$KW = e^{(148.9802 - \frac{13847.26}{T_K} - 23.6521 log(T_K) + (\frac{118.67}{T_K} - 5.977 + 1.0495 log(T_K))\sqrt{S} - 0.01615S)}$$

Boron:

$$KB = exp(\frac{(-8966.90 - 2890.53\sqrt{S} - 77.942S + 1.728S\sqrt{S} - 0.0996S^2)}{T_K}$$
$$+ 148.0248 + 137.1942\sqrt{S} + 1.62142S$$
$$- (24.4344 + 25.085\sqrt{S} + 0.2474S) log(T_K) + 0.053105 T_K \sqrt{S})$$
$$TB = 0.0004326 \frac{S}{35}$$

Choice of carbonate dissociation constants K_1 and K_2 were Mehrbach [71] (refit by Dickson and Millero [26]) with $1.23K_1$ and $0.53K_2$ measured experiment specific adjustments:

$$K_1 = 10^{(-(\frac{3633.86}{T_K} - 61.2172 + 9.6777 log(T_K) - 0.011555S + 0.0001152S^2))} * 1.23 \quad (4.1)$$
$$K_2 = 10^{(-(\frac{471.8}{T_K} + 25.9290 - 3.16967 log(T_K) - 0.01781S + 0.0001122S^2))} * 0.53 \quad (4.2)$$

Approximating equation for the starting value of pH:

$$pH_0 = 12.26 - 0.0030605DIC - 0.043752T - 0.013625S + 0.00011315TA$$
$$+ 1.3463e-5DIC * T + 5.2215e-7DIC * TA \quad (4.3)$$

Newton-Raphson iterations:

$$h_n = 10^{-pH_n} \tag{4.4}$$

$$h_{n_{free}} = \frac{h_n}{Free_{2_T}} \tag{4.5}$$

$$\begin{aligned} f_n = (DIC * 1e-6 * \frac{K_1 h_n + 2K_1K_2}{h_n^2 + K_1 h_n + K_1K_2} \\ - h_{n_{free}} + \frac{KW}{h_n} - TA * 1e-6 + \frac{TB}{1 + \frac{h_n}{KB}}) * 1e6 \end{aligned} \tag{4.6}$$

$$\begin{aligned} df_n = (DIC * 1e-6 * \frac{K_1 + 2K_1K_2}{h_n^2 + K_1 h_n + K_1K_2} \\ - DIC * 1e-6 * \frac{(K_1 h_n + 2K_1K_2)}{(h_n^2 + K_1 h_n + K_1K_2)^2}(2h_n + K_1) \\ - TB\frac{1}{(1 + \frac{h_n}{KB})^2}/KB \\ - \frac{KW}{h_n^2} - \frac{1}{Free_{2_T}}) * 1e6 * (-log(10) * 10^{-pH}) \end{aligned} \tag{4.7}$$

$$pH_{n+1} = pH_n - \frac{f_n}{df_n} \tag{4.8}$$

$$H_{n+1} = 10^{-pH_{n+1}} \tag{4.9}$$

$$CO_{2_{n+1}} = \frac{H_{n+1}^2 DIC}{H_{n+1}^2 + K_1 H_{n+1} + K_1K_2} \tag{4.10}$$

$$HCO_{3_{n+1}} = \frac{H_{n+1} K_1 DIC}{H_{n+1}^2 + K_1 H_{n+1} + K_1K_2} \tag{4.11}$$

$$CO_{3_{n+1}} = \frac{K_1K_2 DIC}{H_{n+1}^2 + K_1 H_{n+1} + K_1K_2} \tag{4.12}$$

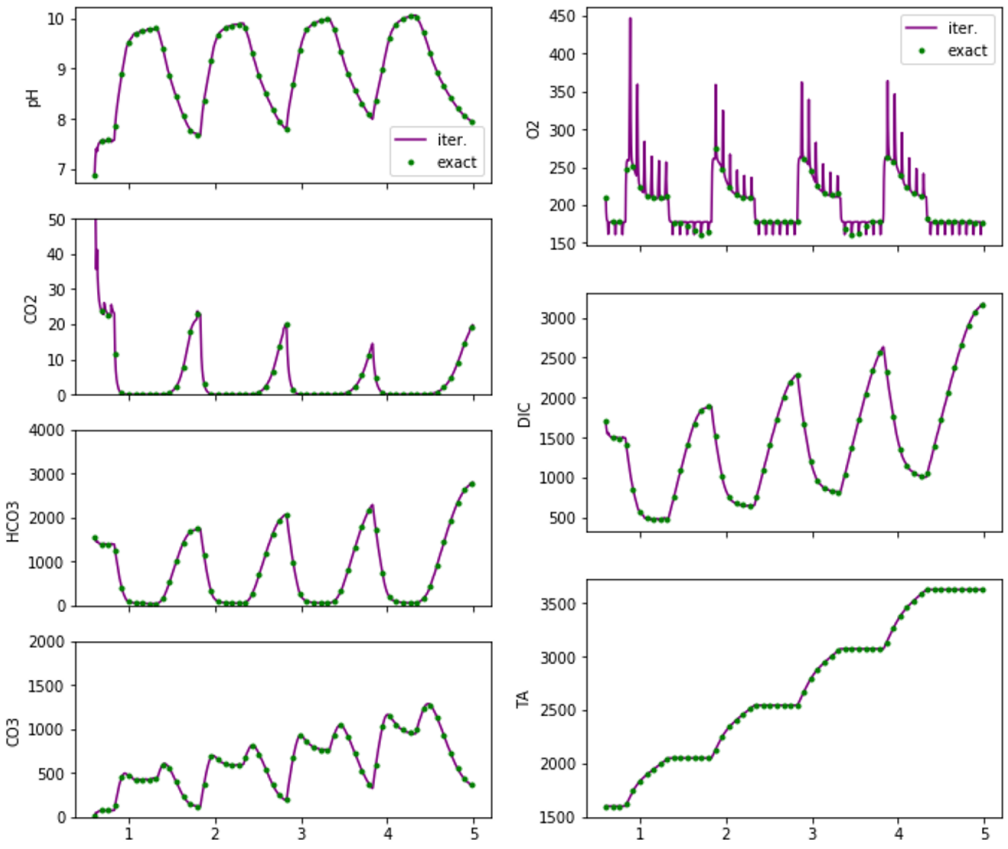

Figure 4.3 : Iterative (3rd iteration) vs exact solution for carbon chemistry CO_2 (μM L^{-1}), HCO_3 (μM L^{-1}), CO_3 (μM L^{-1}), pH ($\log_{10}$(-mol/L H+)) and state variables O_2 (μM L^{-1}), DIC (μM L^{-1}), and TA (μM L^{-1}).

Variable	Iter. 1	Iter. 2	Iter. 3	Iter. 4	Iter. 5
pH	0.036092734	0.002355758	1.41E-05	6.93E-06	6.93E-06
CO_2	2.109401968	0.145719349	0.001222812	0.000866728	0.000866727
HCO_3	19.81869214	1.21021115	0.008016765	0.001025002	0.001025139
CO_3	20.89660704	1.307061652	0.00867642	0.001102278	0.001102434
DIC	16.78775711	0.958511825	0.005229318	0.002305054	0.002333411
O_2	0.308389964	0.016044284	4.18E-05	6.89E-05	7.59E-05
TA	2.607767674	0.160897272	0.000688102	0.001257725	0.001218981

Table 4.1 : RMSE for 5 iterations of the Newton-Raphson carbon chemistry iterative solution.

4.2.4 Process model: Gas transfer equilibrium concentrations for O_2 and CO_2

The equilibrium concentration for CO_2 solubility in water, CO_{2H} (μM L^{-1}), is calculated using Henry's law,

$$CO_{2H} = K0_{CO2} * fCO2 * 1.0220 * 1e6 \tag{4.13}$$

where fCO2 (atm) is the fugacity or approximately the partial pressure of CO_2, 1.0220 is the density of seawater (kg L^{-1}) at salinity 34 ppt and temperature 27°C [79, 42]. $K0_{CO2}$ (mol/kg_{soln}/atm) is the solubility of gas in seawater and is calculated from the fitted van't Hoff equation and the logarithmic Setchenow salinity dependence [95],

$$K0_{CO2} = e^{(-60.2409+93.4517\frac{100}{T_K}+23.3585*log(\frac{T_K}{100})+S(0.023517-0.023656\frac{T_K}{100}+0.0047036(\frac{T_K}{100})^2))} \tag{4.14}$$

where T_K is the temperature (K) and S is salinity (ppt). Similarly the equilibrium concentration for O_2 solubility in water O_{2H} is calculated using Henry's law,

$$O_{2H} = K0_{O2} * fO2 * 1.0220 * 1e-6 \tag{4.15}$$

where fO2 (atm) is the fugacity or approximately the partial pressure of O_2, 1.0220 is the density of seawater (kg L^{-1}) at salinity 34 ppt and temperature 27°C [79, 42], and $K0_{O2}$ (mol/kg_{soln}/atm) is the solubility of oxygen in seawater with an adjusted salinity dependence [7],

$$K0_{O2} = \frac{e^{(-1282.8704+\frac{36619.96}{T_K}+223.1396log(T_K)-0.354707T_K+S(5.957e-3-\frac{3.7353}{T_K})+3.68e-6S^2)}}{0.2094e-6} \tag{4.16}$$

where T_K is the temperature (K) and S is salinity (ppt). The equilibrium concentrations for O_2 and CO_2 are modelled together with the gas turning on and off during the experiment, as

$$Q^{air}kLa^{air}_{O_2}(O_{2H} - O_2) \tag{4.17}$$

$$Q^{air}kLa^{air}_{CO_2}(CO_{2H} - CO_2) \tag{4.18}$$

where Q^{air} is the gas state (1= on, 0= off), $kLa^{air}_{O_2}$ and $kLa^{air}_{CO_2}$ are the mass transfer coefficients for air (day^{-1}), and 0.893 is the ratio between measured O_2 and CO_2 mass transfer constants [44].

4.2.5 Process model: Flux into cells by photosynthesis and respiration

The carbon flux into cells is introduced through net photosynthesis,

$$\frac{dDIC}{dt} = -(PI\frac{HCO_3^-}{K_m + HCO_3^-} - R) + \text{gas transfer} + \text{dilution} \tag{4.19}$$

$$\frac{dO_2}{dt} = \frac{1}{(RQ_dI + RQ_n(1-I))}(PI\frac{HCO_3^-}{K_m + HCO_3^-} - R) + \text{gas transfer} + \text{dilution} \tag{4.20}$$

$$\frac{dTA}{dt} = R_R(PI\frac{HCO_3^-}{K_m + HCO_3^-}) + \text{dilution} \tag{4.21}$$

$$\frac{dC_p}{dt} = (PI\frac{HCO_3^-}{K_m + HCO_3^-} - R) + \text{dilution} \tag{4.22}$$

where the parameters common to all states are P, the maximum photosynthesis rate (μM L^{-1} $hour^{-1}$), I the light indicator (0/1), and the Michaelis-Menton term $\frac{HCO_3^-}{K_m+HCO_3^-}$ representing the photosynthetically active carbon used for photosynthesis. This can be CO_2, HCO_3^-, or a combination of both, $CO_2 + HCO_3^-$ if the microalgae are using both carbon dioxide and bicarbonate for photosynthesis. K_m is the carbon restriction term (μM L^{-1}).

The respiration rate R (μM L^{-1} $hour^{-1}$), is present in the net photosynthesis calculation for dissolved inorganic carbon (DIC), oxygen (O_2), and carbon concentration in the form of cells (C_p). Oxygen also accounts for the day (RQ_d) and night respiratory quotients (RQ_n), the ratio of CO_2 produced and O_2 consumed by a cell. Total alkalinity (TA) only increases due to photosynthesis, while accounting for the Redfield ratio (R_R).

4.2.6 Process model: Dilution

To maintain a semi-batch culturing system, manual dilutions that occurred once per day over the period of the experiment were extracted to measure total alkalinity

and dissolved inorganic carbon (described in more detail in Section 4.2.1). This also keeps the algae in the growth stage of the biomass curve, resetting it every day so that it never reaches the logit plateau. In the ODEs this affects every state variable, dissolved organic carbon (DIC), oxygen (O_2), total alkalinity (TA) and amount of carbon in the form of cells (C_p),

$$\frac{dDIC}{dt} = \text{flux into cells} + \text{gas transfer} + \frac{Q^M}{V}(DIC^M - DIC) \quad (4.23)$$

$$\frac{dO_2}{dt} = \text{flux into cells} + \text{gas transfer} + \frac{Q^M}{V}(O_2^M - O_2) \quad (4.24)$$

$$\frac{dTA}{dt} = \text{flux into cells} + \frac{Q^M}{V}(TA^M - TA) \quad (4.25)$$

$$\frac{dC_p}{dt} = \text{flux into cells} + \frac{Q^M}{V}(- C_p) \quad (4.26)$$

where Q^M is the measured dilution rate (mL day^{-1}) used to force the model, V is the volume of the photo-bioreactor fixed at 500 mL, and DIC^M, O_2^M, and TA^M are the respective concentrations of the media. The media concentrations were calculated using CO2SYS (at temperature = 27°C and salinity = 34 ppt) and set to be fixed throughout the experiment as DIC^M= 1,724.20 μM L^{-1}, O_2^M = 226.65 μM L^{-1}, and TA^M = 1,797.90 μM L^{-1}.

4.2.7 Process model summary and parameter model

A summary of the ODEs that make up the process model described in previous sections:

$$
\begin{array}{lrlll}
\text{Rate} & & \text{flux into cells} & \text{gas transfer} & \text{dilution} \\
\frac{dDIC}{dt} = & -(PI\frac{HCO_3^-}{K_m + HCO_3^-} - R) & +\hat{Q}^{air} kLa^{air}_{CO_2}(CO_2^{air} - CO_2) & +\frac{Q^M}{V}(DIC^M - DIC) \\
\frac{dO_2}{dt} = & \frac{(PI\frac{HCO_3^-}{K_m+HCO_3^-} - R)}{(RQ_d I + RQ_n(1-I))} & +\hat{Q}^{air} kLa^{air}_{O_2}(\ O_2^{air} - \ O_2) & +\frac{Q^M}{V}(\ O_2^M - \quad O_2) \\
\frac{dTA}{dt} = & R_R(PI\frac{HCO_3^-}{K_m + HCO_3^-} \quad) & & +\frac{Q^M}{V}(\ TA^M - \ TA) \\
\frac{dC_p}{dt} = & (PI\frac{HCO_3^-}{K_m + HCO_3^-} - R) & & +\frac{Q^M}{V}(\quad - \ C_p)
\end{array}
\tag{4.27}
$$

	Symbol	Description	Prior / Value	Unit
Initial conditions	DIC^0	Dissolved inorganic carbon	$\text{Log}\mathcal{N}(\log(1300), 0.2)$	μM L^{-1}
	O_2^0	Oxygen	$\text{Log}\mathcal{N}(\log(225), 0.2)$	μM L^{-1}
	TA^0	Total alkalinity	$\text{Log}\mathcal{N}(\log(1750), 0.1)$	μM L^{-1}
	C_p^0	Carbon in the form of cells	$\text{Log}\mathcal{N}(\log(300), 0.2)$	μM L^{-1}
	pH^0	-	$\text{Log}\mathcal{N}(\log(8.5), 0.2)$	$\log_{10}$(-mol/L H+)
	${CO_2}^0$	Carbon dioxide	$\text{Log}\mathcal{N}(\log(5), 0.4)$	μM L^{-1}
	HCO_3^{-0}	Bicarbonate	$\text{Log}\mathcal{N}(\log(1500), 0.3)$	μM L^{-1}
	CO_3^{2-0}	Carbonate	$\text{Log}\mathcal{N}(\log(100), 0.4)$	μM L^{-1}
Flux into cells	P	Maximum photosynthesis rate	*	μM L^{-1} $hour^{-1}$
	R	Respiration rate	*	μM L^{-1} $hour^{-1}$
	K_m	Carbon restriction	*	μM L^{-1}
	RQ_d	Daytime respiratory quotient	*	-
	RQ_n	Night respiratory quotient	*	-
	R_R	Redfield ratio	*	-

	I	Light indicator	forcing (0/1)	-
Gas transfer terms	$\hat{Q}^{air}$	Indicator for flow in air line	forcing (0/1)	-
	$x^{air}_{CO_2}$	Mole fraction of CO_2 atmosphere	400	ppm
	CO_{2H}	Equilibrium CO_2 concentration	Eq. 4.13	μM L^{-1}
	CO_2^{air}	Sat CO_2 conc with atmosphere	$x^{air}_{CO_2}$ CO_{2H}	
	$kLa^{air}_{CO_2}$	Mass transfer coefficient for CO_2	$0.893kLa^{air}_{O_2}$	day^{-1}
	$x^{air}_{O_2}$	Mole fraction of O_2 atmosphere	0.2094	atm
	O_{2H}	Equilibrium O_2 concentration	Eq. 4.15	μM L^{-1}
	O_2^{air}	Sat O_2 conc with atmosphere	$x^{air}_{O_2}$ O_{2H}	
	τ	half-life of $kLa^{air}_{O_2}$	range(2-20)	min^{-1}
	$kLa^{air}_{O_2}$	Mass transfer coefficient for O_2	$\ln(2) * 24 * 60/\tau$	day^{-1}
Dilution terms	Q^M	Dilution rate	forcing	mL day^{-1}
	V	Volume of the reactor	500	mL
	DIC^M	Media dissolved inorganic carbon	1724.20	μM L^{-1}
	O_2^M	Media oxygen concentration	226.65	μM L^{-1}
	TA^M	Media total alkalinity	1797.90	μM L^{-1}

Table 4.2 : State variable, parameter and forcing definitions with units, and their assignments: either fixed values, priors on initial condition or priors on parameters (*) defined later in respective sections.

4.2.8 Design and setup of data assimilation model for a twin experiment and experimental data

Firstly, a twin experiment was set up with simulated observations, followed by assimilation of the experimental data described in Section 4.2.1 and 4.2.2. In the twin experiment, the process model described in Section 4.2.7 was run with a fixed set of parameters (P = 200, R = 30, $kLA^{air}_{O_2}$ = 200, K_m = 150, RQ_d = 0.85, RQ_n

= 0.95, R_R = 0.075) and initial conditions (O_2^0 = 225, DIC^0 = 1250, TA^0 = 1750) to provide a "true" simulated set of results. Normally distributed noise was added to the sub-sampled model generated synthetic dataset with standard deviations of $2\,\mu$M L^{-1} for O_2, 0.01 for pH and $50\,\mu$M L^{-1} for DIC and TA, to simulate random observation error comparable to measured versus CO2SYS derived estimates. This synthetic dataset was then assimilated into the model to generate posterior distributions for states and parameters where,

1. All parameters were treated as constant through time (Section 4.3.1).
2. Photosynthesis and respiration were treated as random walks through time (Section 4.3.2).

The success of the twin experiment was judged by the degree of agreement between the posterior results and the true state and parameter values. These results were followed by assimilation of experimental data with several model and parameter formulations (Section 4.3.3 to 4.3.5),

1. Photosynthesis, respiration and respiratory quotients are random walks through time.
2. Including an O_2 offset.
3. Thinning out the experimental observations further.

Each result section includes the specific treatment of parameters, priors and proposal distributions for clarity. Distributions and their notations are the uniform ($\mathcal{U}$), normal ($\mathcal{N}$), log normal (Log$\mathcal{N}$), and truncated normal (Trun$\mathcal{N}$) which imposes an upper and lower bound on the normal distribution.

4.3 Results

State posteriors are visualised by plotting the median and shading 95% credible intervals, while parameter priors and posteriors are displayed by histograms. Interval estimates (25%, 75%) and (5%, 95%) of parameter posterior distributions are also quoted. Where the synthetic data were assimilated the true parameter values are indicated by solid green lines.

4.3.1 Posterior results of twin experiment where photosynthesis and respiration are constant through time

Parameters P, R, kLA_{O_2}, K_m, R_R, RQ_d, and RQ_n were all treated as constant through time but unknown with prior distributions and proposal distributions defined in Table 4.3. The other model parameters were defined earlier in Table 4.2.7. The data model assigned log normally distributed observation errors for each instrument (Section 4.2.2) with observation error standard deviation values cited in Table 4.3.

After running the PMMH for 50,000 samples (10,000 discarded as burn-in) with 1,024 particles and tuning to an acceptance rate of 28.79%, the true values of parameters R, $kLA^{air}_{O_2}$, R_R, and RQ_d were recovered to the 25^{th} and 75^{th} percentile (Table 4.4, Figure 4.7). While the true values of K_m and RQ_n did not lie within the 25^{th} and 75^{th} quantile, they were captured by the 5^{th} and 95^{th} percentiles (Table 4.4, Figure 4.7). P was the only parameter whose true value lied on the cusp of the 5^{th} percentile. Parameters P, R, $kLA^{air}_{O_2}$, and R_R mixed well, while parameters K_m, RQ_d, and RQ_n did not mix as well, some autocorrelation between samples was present (Figure 4.8).

Observed state variables O_2, DIC, TA, and observed pH posteriors were in excellent agreement with observations, with all observations fitting within the 95% cred-

Table 4.3 : Twin experiment parameters, priors and proposal distributions where photosynthesis and respiration were constant through time (* indicates the parameter was held fixed).

Parameter	**Prior**	**Proposal**
P	$\text{Log}\mathcal{N}(\log(250.0), 0.8)$	$\text{Log}\mathcal{N}(\log(\text{P}), 0.08)$
R	$\text{Log}\mathcal{N}(\log(20.0), 0.8)$	$\text{Log}\mathcal{N}(\log(\text{R}), 0.08)$
kLA_{O2}	$\text{Log}\mathcal{N}(\log(200.0), 0.3)$	$\text{Log}\mathcal{N}(\log(\text{kLA}_{O2}), 0.03)$
K_m	$\text{Log}\mathcal{N}(\log(200.0), 0.6)$	$\text{Log}\mathcal{N}(\log(\text{K}_m), 0.06)$
R_R	$\mathcal{U}(0, 0.2)$	$\text{Trun}\mathcal{N}(\text{R}_R, 0.01, \text{lower} = 0, \text{upper} = 0.2)$
RQ_d	$\mathcal{U}(0.6, 1)$	$\text{Trun}\mathcal{N}(\text{RQ}_d, 0.005, \text{lower} = 0.6, \text{upper} = 1.0)$
RQ_n	$\mathcal{U}(0.6, 1)$	$\text{Trun}\mathcal{N}(\text{RQ}_n, 0.005, \text{lower} = 0.6, \text{upper} = 1.0)$
σ_{O_2}	0.3	*
σ_{pH}	0.3	*
σ_{DIC}	0.3	*

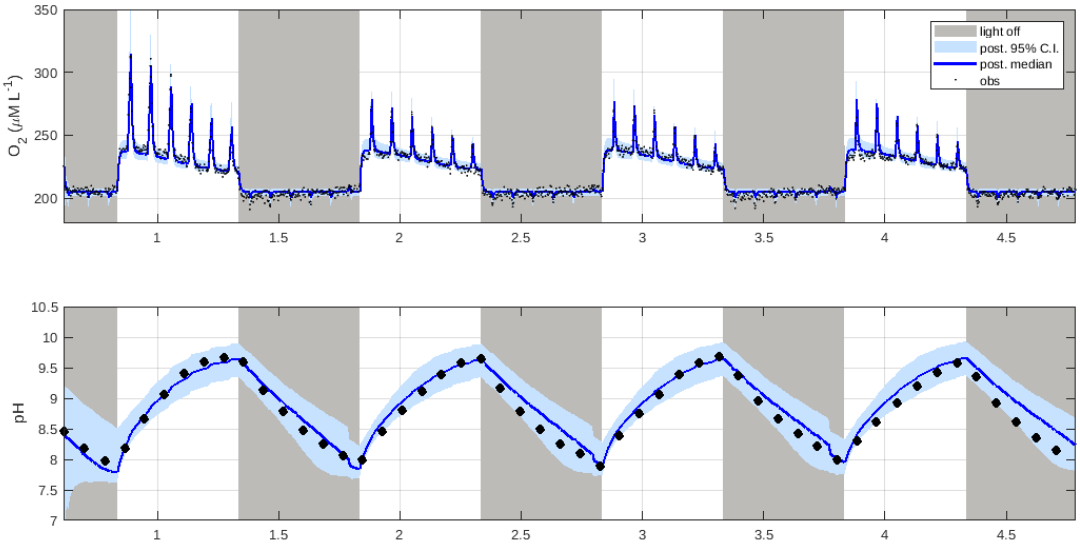

Figure 4.4 : Twin experiment posterior medians (solid blue line), 95% credible intervals (shaded blue), and synthetic observations (black) for O_2 and pH over 5 days where photosynthesis and respiration were constant through time.

ible interval posteriors (Figure 4.4, Figure 4.5). Unobserved state variable posterior C_p appeared to be constrained throughout the experiment (Figure 4.5). Similarly, the carbon chemistry variables HCO_3, CO_2, and CO_3 also had tightly constrained posteriors (Figure 4.6).

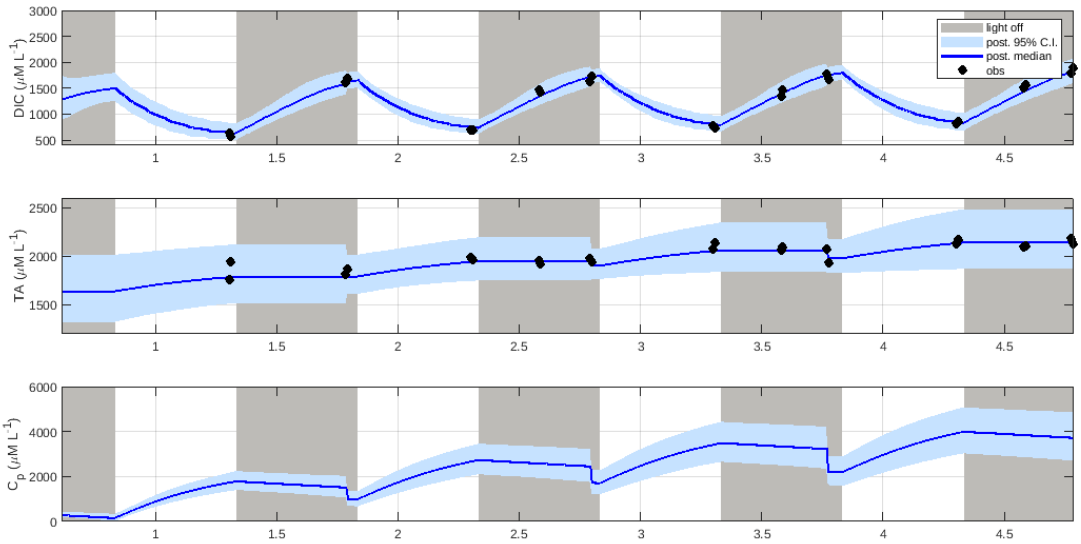

Figure 4.5 : Twin experiment posterior medians (solid blue line), 95% credible intervals (shaded blue), and synthetic observations (black) for DIC, TA and C_p over 5 days where photosynthesis and respiration were constant through time.

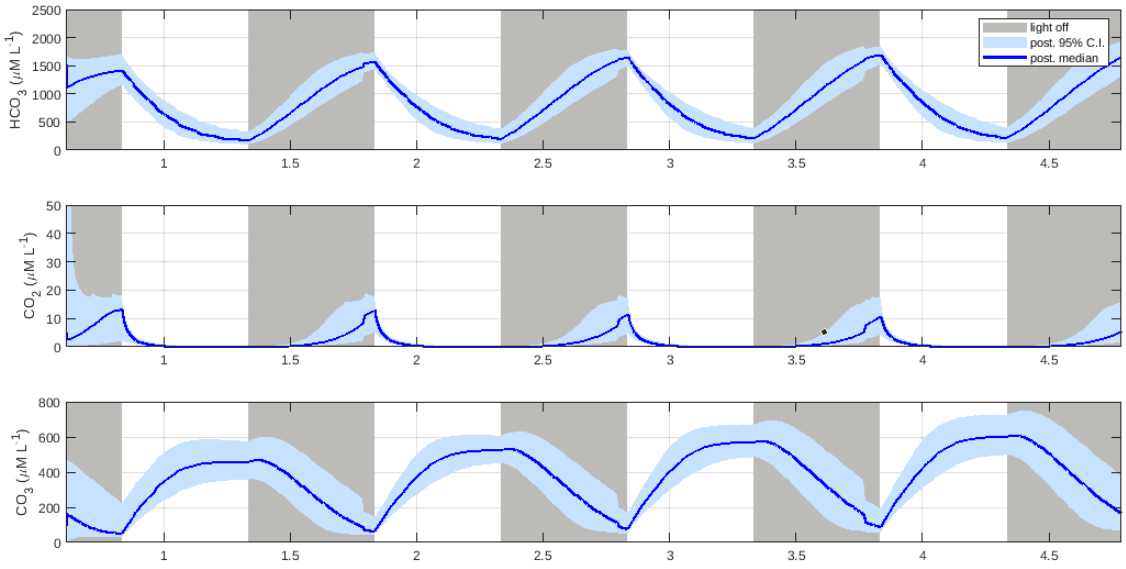

Figure 4.6 : Twin experiment posterior medians (solid blue line) and 95% credible intervals (shaded blue) for HCO_3, CO_2 and CO_3 over 5 days where photosynthesis and respiration were constant through time.

Table 4.4 : True values for parameters used to create synthetic observations, and posterior (25%, 75%), (5%, 95%) quantiles after assimilation for the twin experiment where photosynthesis and respiration were constant through time.

Parameter	**Quantiles (25%, 75%)**	**Quantiles (5%, 95%)**	**True value**
P	(237.2167, 302.0565)	(200.8971, 365.1751)	200
R	(17.5370, 32.3501)	(10.6537, 45.3096)	30
$kLA_{O_2}^{air}$	(177.1383, 222.9641)	(148.2838, 267.5934)	200
K_m	(182.6643, 362.6714)	(118.4125, 643.8518)	150
R_R	(0.0504, 0.1041)	(0.0196, 0.1458)	0.075
RQ_d	(0.7529, 0.8718)	(0.6894, 0.9666)	0.85
RQ_n	(0.7469, 0.9338)	(0.6400, 0.9920)	0.95

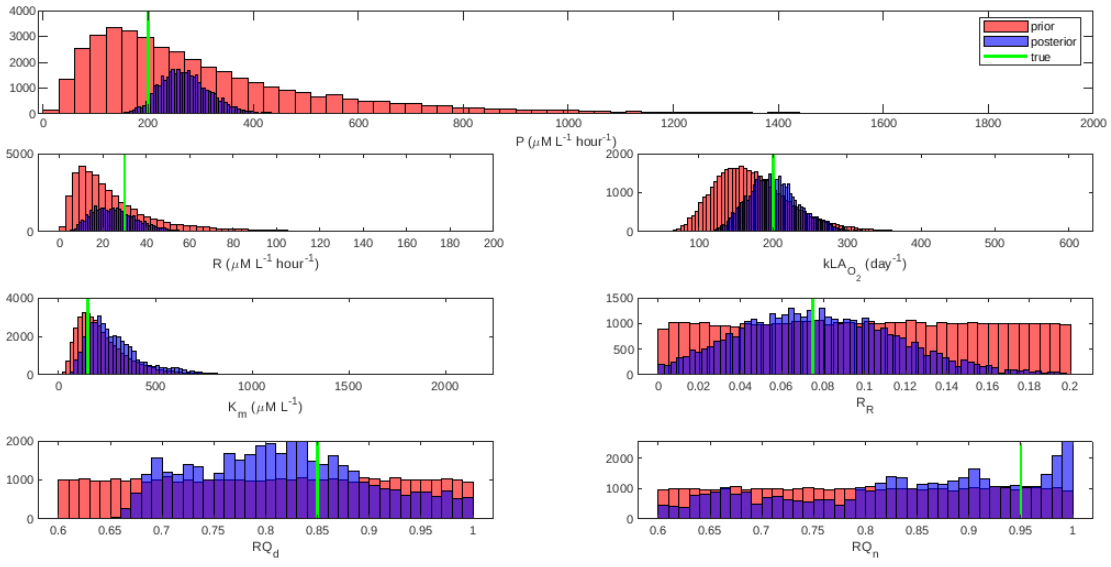

Figure 4.7 : Model parameter priors (orange), posteriors (purple) and true values (green) for the twin experiment where photosynthesis and respiration were constant through time.

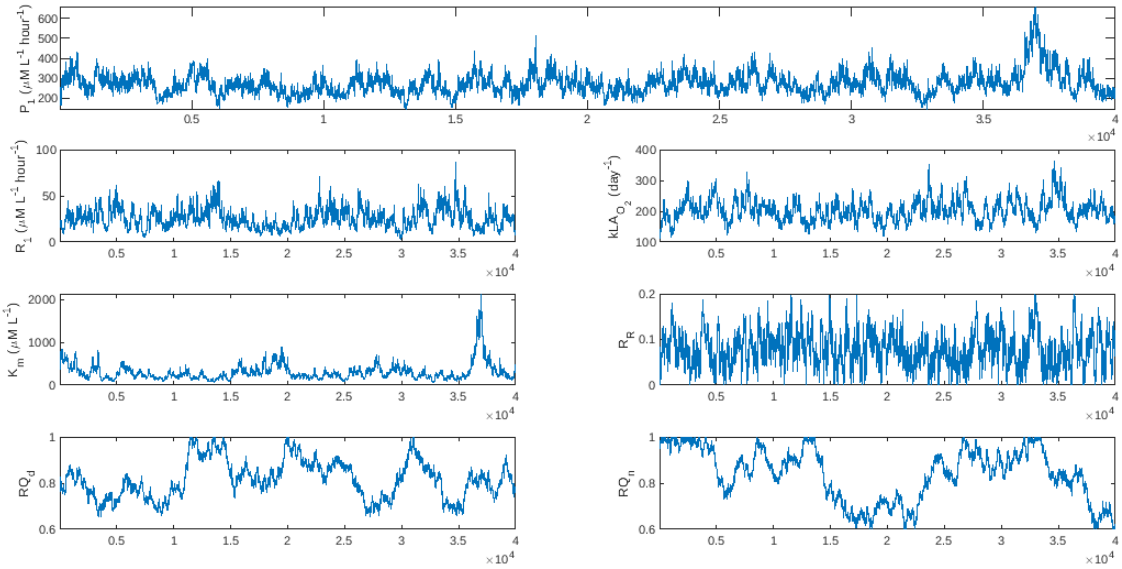

Figure 4.8 : Parameter trace-plots for the twin experiment where photosynthesis and respiration were constant through time.

4.3.2 Posterior results of twin experiment where photosynthesis and respiration are changing through time

Parameters kLA_{O_2}, K_m, R_R, RQ_d, and RQ_n were all treated as constant through time but unknown with prior distributions and proposal distributions defined in Table 4.5. Photosynthesis (P_1) and respiration (R_1) were both modelled as random walks, by taking P and R, previously constant parameters, and replacing them by P_1(t) and R_1(t). Here, we took P_1(t) and R_1(t) to be such that

$$P_1(t + \Delta t) = P(t) + r_P$$

$$R_1(t + \Delta t) = R(t) + r_R$$

where $r_P \sim \mathcal{N}(0, \sigma_{r_P})$, $r_R \sim \mathcal{N}(0, \sigma_{r_R})$, and Δt is the length of discrete time-step. For the purpose of the Bayesian analysis here, σ_{r_P} and σ_{r_R} were treated as parameters to be inferred. The other model parameters were defined earlier in Table 4.2.7. The data model assigned log normally distributed observation errors for each instrument (Section 4.2.2) with observation error standard deviation values cited in Table 4.5.

After running the PMMH for 50,000 samples (10,000 discarded as burn-in) with 1,024 particles and tuning to get an acceptance rate of 20.23%, the true values of parameters $kLA^{air}_{O_2}$, R_R, and RQ_d were recovered to the 25$^{\text{th}}$ and 75$^{\text{th}}$ percentile (Table 4.6). While the true values of K_m and RQ_n did not lie within the 25$^{\text{th}}$ and 75$^{\text{th}}$ quantile, they were captured by the 5$^{\text{th}}$ and 95$^{\text{th}}$ percentiles (Table 4.6). Parameters σ_{rP}, σ_{rR}, $kLA^{air}_{O_2}$, and R_R mixed well, while parameters K_m, RQ_d, and RQ_n did not mix as well, some autocorrelation between samples was present (Figure 4.14). The true values of P_1 (200) and R_1 (30) lay within the posterior 95% credible intervals (Figure 4.13). Observed state variables O_2, DIC, TA, and observed pH posteriors were in excellent agreement with observations, with all observations fitting within the 95% credible interval posteriors (Figure 4.9, Figure 4.10).

Table 4.5 : Parameters, priors and proposal distributions for the twin experiment where photosynthesis and respiration were changing through time (* indicates the parameter was held fixed).

Parameter	**Prior**	**Proposal**
kLA_{O2}	$\mathrm{Log}\mathcal{N}(\log(200.0), 0.3)$	$\mathrm{Log}\mathcal{N}(\log(\mathrm{kLA}_{O2}), 0.03)$
K_m	$\mathrm{Log}\mathcal{N}(\log(200.0), 0.6)$	$\mathrm{Log}\mathcal{N}(\log(\mathrm{K}_m), 0.06)$
R_R	$\mathcal{U}(0, 0.2)$	$\mathrm{Trun}\mathcal{N}(\mathrm{R}_R, 0.01, \mathrm{lower} = 0, \mathrm{upper} = 0.2)$
RQ_d	$\mathcal{U}(0.6, 1)$	$\mathrm{Trun}\mathcal{N}(\mathrm{RQ}_d, 0.005, \mathrm{lower} = 0.6, \mathrm{upper} = 1.0)$
RQ_n	$\mathcal{U}(0.6, 1)$	$\mathrm{Trun}\mathcal{N}(\mathrm{RQ}_n, 0.005, \mathrm{lower} = 0.6, \mathrm{upper} = 1.0)$
σ_{r_P}	$\mathcal{N}(0.01, 0.001)$	$\mathcal{N}(\sigma_{r_P}, 0.0001)$
σ_{r_R}	$\mathcal{N}(0.01, 0.001)$	$\mathcal{N}(\sigma_{r_R}, 0.0001)$
σ_{O_2}	0.3	*
σ_{pH}	0.3	*
σ_{DIC}	0.3	*

Table 4.6 : True values for parameters used to create synthetic observations, and posterior (25%, 75%), (5%, 95%) quantiles after assimilation for the twin experiment where photosynthesis and respiration were changing through time.

Parameter	**Quantiles (25%, 75%)**	**Quantiles (5%, 95%)**	**True value**
$\mathrm{kLA}_{O_2}^{air}$	(170.9854, 216.0578)	(145.6899, 253.4652)	200
K_m	(187.3586, 386.4103)	(93.9670, 641.1432)	150
R_R	(0.0521, 0.1075)	(0.0192, 0.1544)	0.075
RQ_d	(0.7906, 0.8965)	(0.6833, 0.9711)	0.85
RQ_n	(0.6654, 0.8334)	(0.6164, 0.9739)	0.95

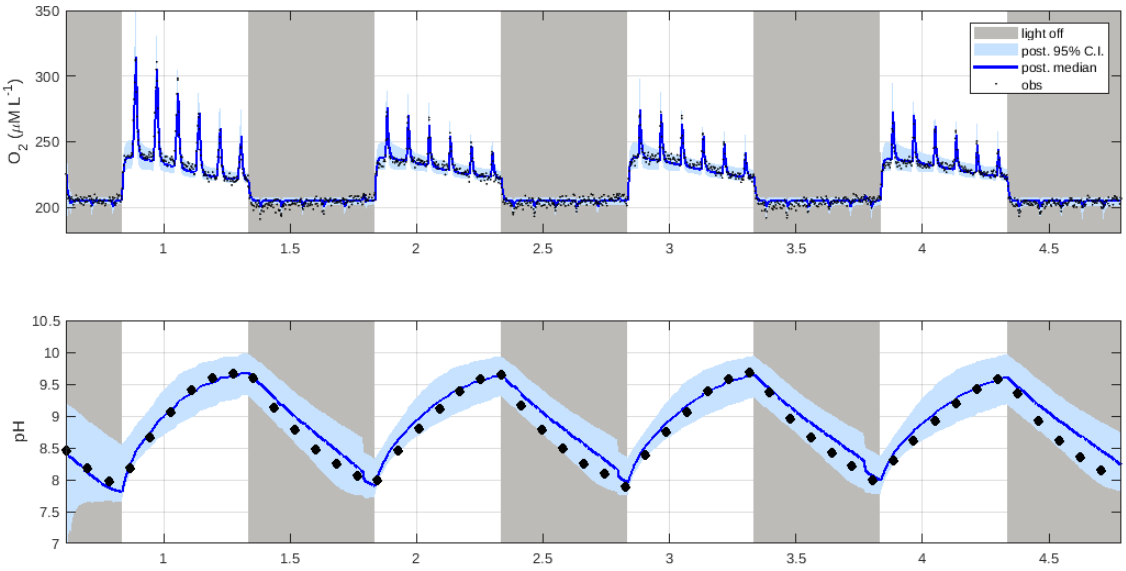

Figure 4.9 : Posterior medians (solid blue line), 95% credible intervals (shaded blue), and synthetic observations (black) of O_2 and pH for the twin experiment where photosynthesis and respiration were changing through time. X-axis is time in days.

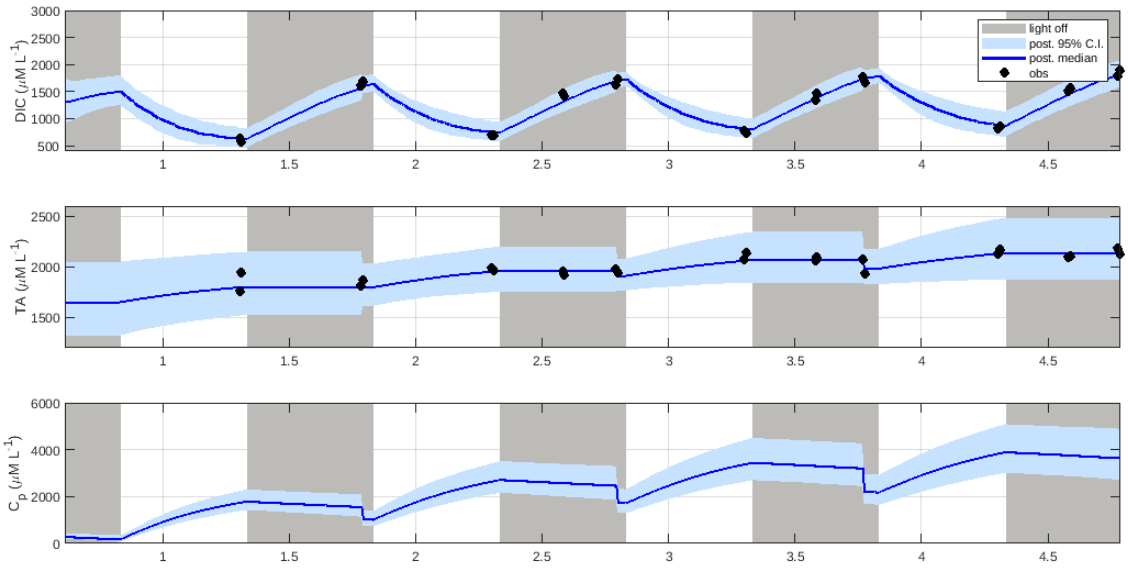

Figure 4.10 : Posterior medians (solid blue line), 95% credible intervals (shaded blue), and synthetic observations (black) of DIC, TA and C_p for the twin experiment where photosynthesis and respiration were changing through time. X-axis is time in days.

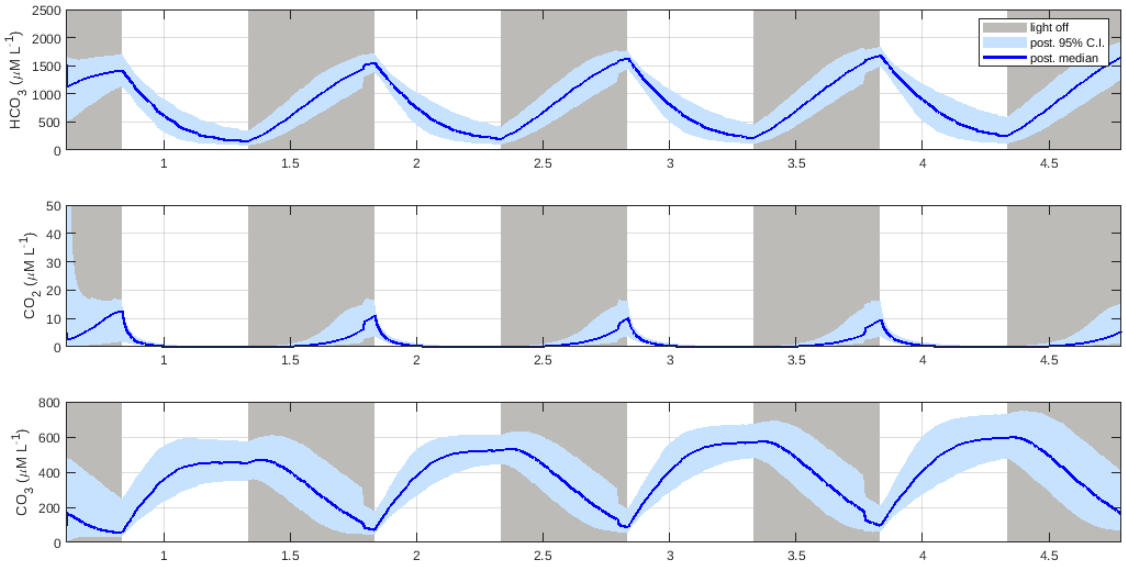

Figure 4.11 : Posterior medians (solid blue line) and 95% credible intervals (shaded blue) of HCO_3, CO_2 and CO_3 for the twin experiment where photosynthesis and respiration were changing through time. X-axis is time in days.

Unobserved state variable posterior C_p appeared to be constrained throughout the experiment (Figure 4.10). Similarly, the carbon chemistry variables HCO_3, CO_2, and CO_3 also had tightly constrained posteriors (Figure 4.11).

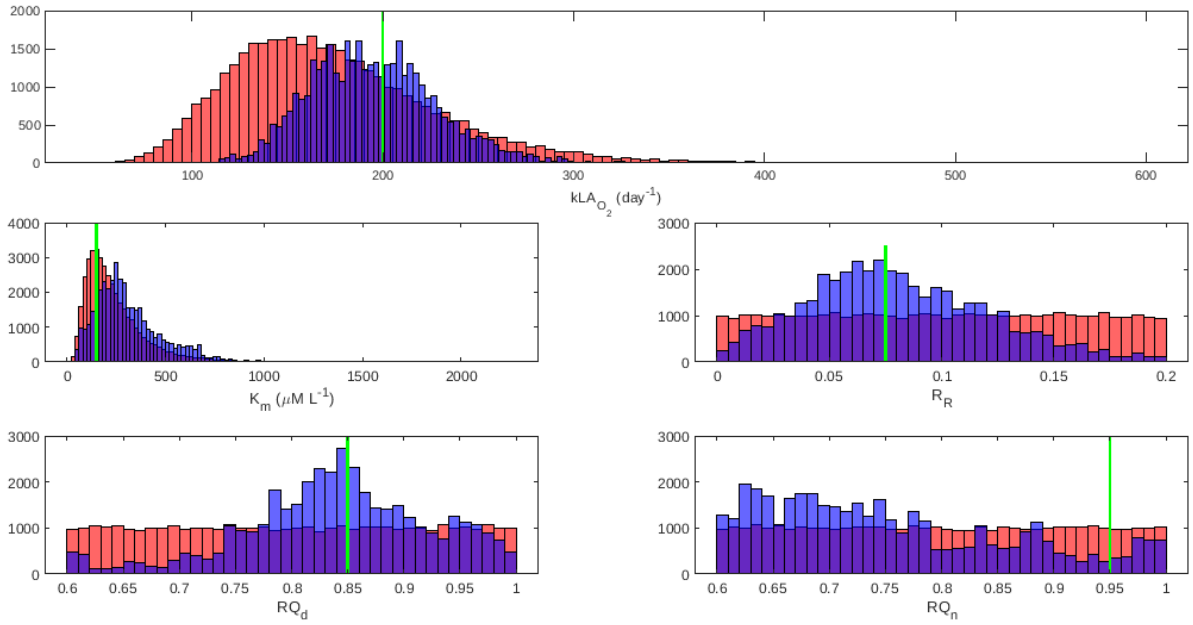

Figure 4.12 : Model parameter priors (orange), posteriors (purple) and true values (green) for the twin experiment where photosynthesis and respiration were changing through time.

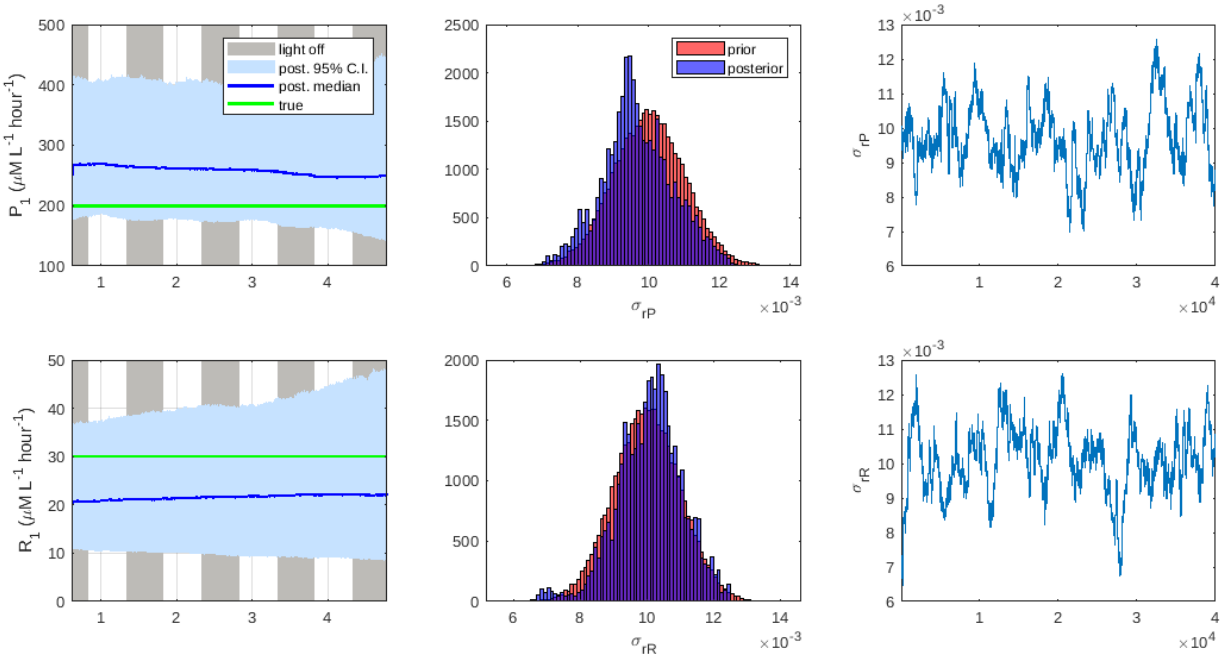

Figure 4.13 : Photosynthesis (P_1) and respiration (R_1) random walk posterior medians (solid blue), 95% credible intervals (shaded blue), and true values (green) for the twin experiment in the left panel (x-axis is time in days). σ_{rP} and σ_{rR} priors (orange) and posteriors (purple) in the centre, and traces on the right.

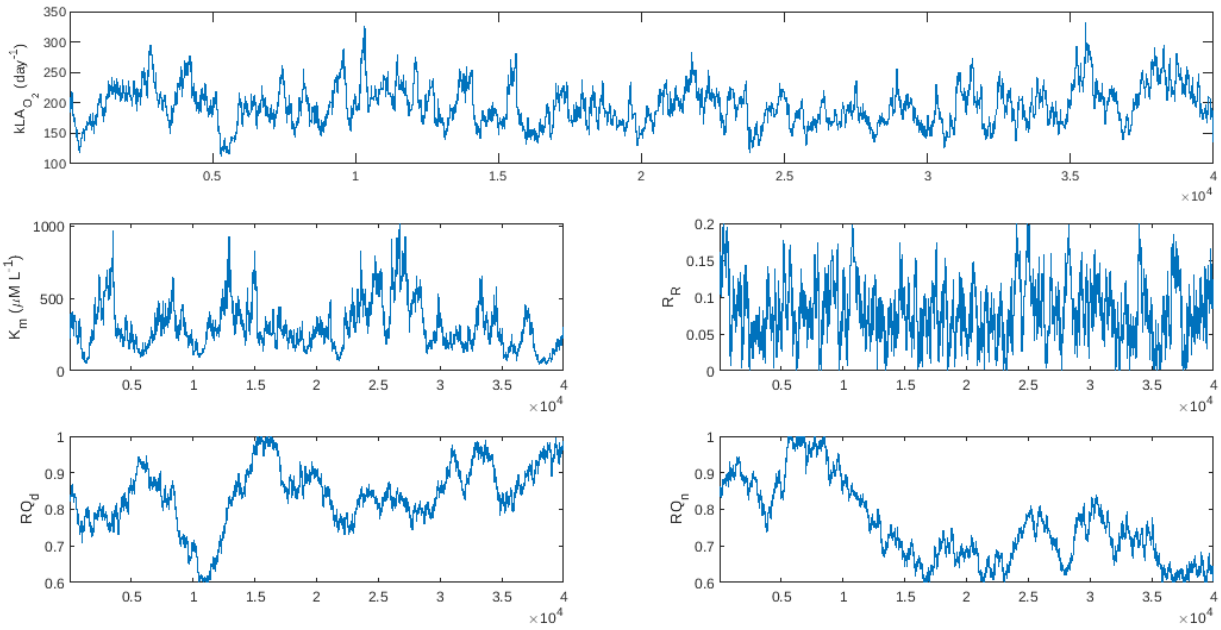

Figure 4.14 : Parameter trace-plots for the twin experiment where photosynthesis and respiration were changing through time.

4.3.3 Posterior results with experimental data where photosynthesis, respiration and respiratory quotients are changing through time

Photosynthesis (P_1) and respiration (R_1) were both modelled as random walks as described in Section 4.3.2. Similarly, the respiratory quotients RQ_d and RQ_n were also treated as random walks, changing through time, with Wiener processes as standard deviations. Parameters kLA_{O_2}, K_m, R_R, σ_{rP}, σ_{rR} were all treated as constant through time but unknown with prior distributions and proposal distributions defined in Table 4.7. The data model assigned log normally distributed observation errors for each instrument (Section 4.2.2) with observation error standard deviation values cited in Table 4.7.

Table 4.7 : Parameters, priors and proposal distributions with experimental data when photosynthesis, respiration and respiratory quotients were changing through time (* indicates the parameter was held fixed).

Parameter	**Prior**	**Proposal**
kLA_{O2}	$\text{Log}\mathcal{N}(\log(200.0), 0.3)$	$\text{Log}\mathcal{N}(\log(\text{kLA}_{O2}), 0.03)$
K_m	$\text{Log}\mathcal{N}(\log(200.0), 0.6)$	$\text{Log}\mathcal{N}(\log(\text{K}_m), 0.06)$
R_R	$\mathcal{U}(0, 0.2)$	$\text{Trun}\mathcal{N}(\text{R}_R, 0.005, \text{lower} = 0, \text{upper} = 0.2)$
σ_{r_P}	$\mathcal{N}(0.02, 0.002)$	$\mathcal{N}(\sigma_{r_P}, 0.0002)$
σ_{r_R}	$\mathcal{N}(0.01, 0.001)$	$\mathcal{N}(\sigma_{r_R}, 0.0001)$
σ_{O_2}	0.3	*
σ_{pH}	0.3	*
σ_{DIC}	0.5	*

The PMMH was run for 40,000 samples with 1,024 particles, where 10,000 samples were discarded as burn-in (40,000 samples was the maximum number of samples that could complete in the maximum time allocation on the HPC cluster). The PMMH was tuned to achieve a final acceptance rate of 20.44%.

Observed state variables DIC and TA posteriors performed well with all observations lying within the 95% credible intervals (Figure 4.16). O_2 posteriors tracked the observations well while the light was on, with all observations falling inside tightly constrained 95% credible intervals. During times when there was no light, most days the posteriors would fit the observations well to start and then potentially there was a sensor drift causing increasing observations that the model was not accounting for (Figure 4.15). pH captured most observations within the 95% credible intervals except for 2 points on the 1st day while the light was off (Figure 4.15). Unobserved

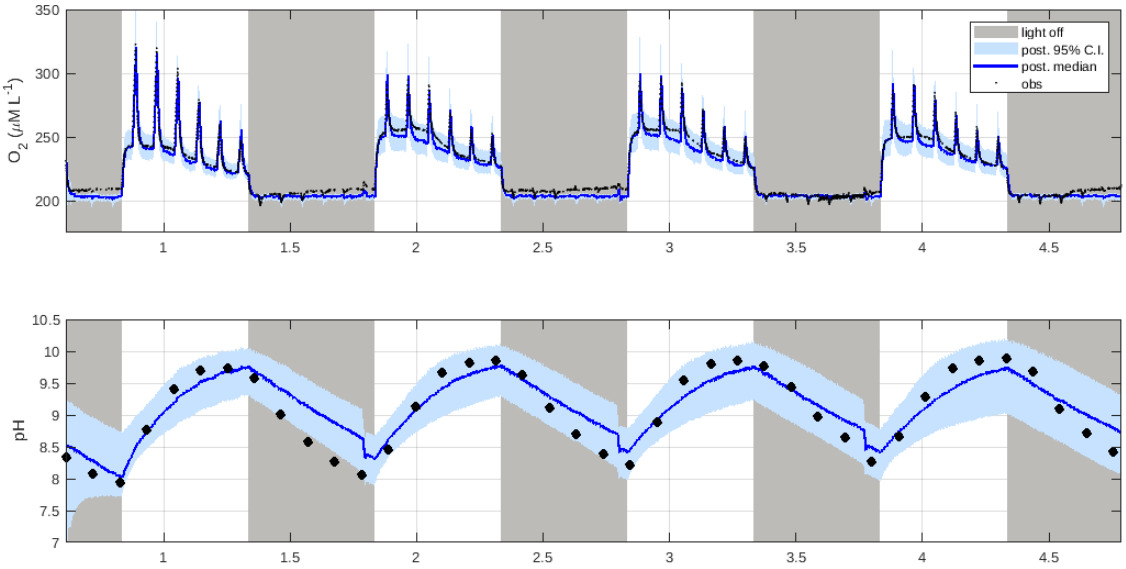

Figure 4.15 : Posterior medians (solid blue line), 95% credible intervals (shaded blue), and experimental observations (black) for O_2 and pH when photosynthesis, respiration and respiratory quotients were changing through time. X-axis is time in days.

Table 4.8 : Posterior (25%, 75%), (5%, 95%) quantiles for parameters after assimilating experimental data when photosynthesis, respiration and respiratory quotients were changing through time.

Parameter	**Quantiles (25%, 75%)**	**Quantiles (5%, 95%)**
$\mathrm{kLA}_{O_2}^{air}$	(139.5979, 170.8102)	(120.9171, 204.5474)
K_m	(168.1931, 378.8001)	(104.2598, 599.6881)
R_R	(0.0815, 0.1512)	(0.0284, 0.1844)
σ_{r_P}	(0.0178, 0.0202)	(0.0159, 0.0222)
σ_{r_R}	(0.0094, 0.0113)	(0.0081, 0.0121)

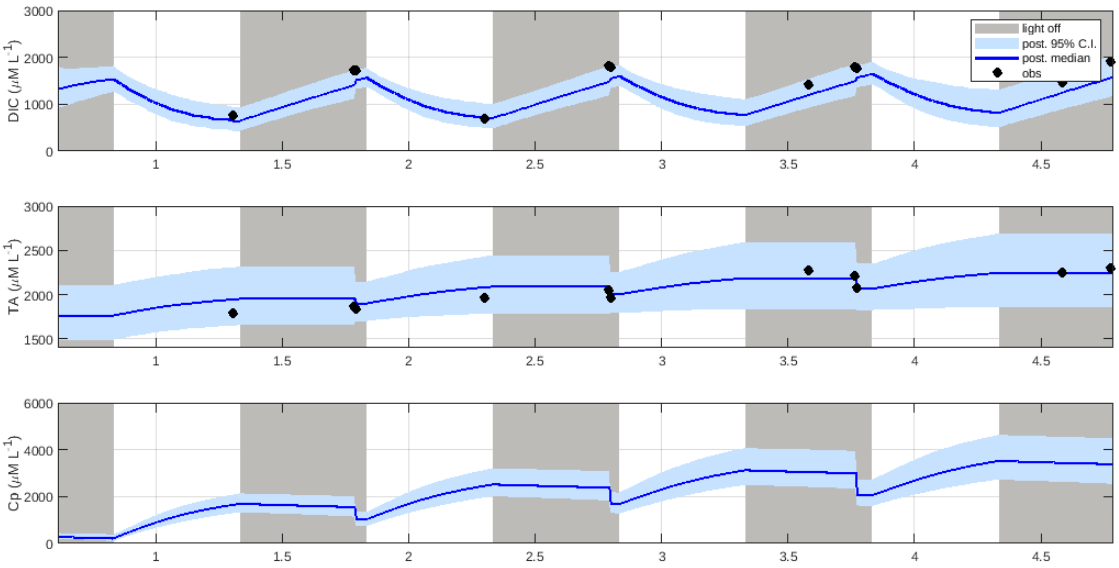

Figure 4.16 : Posterior medians (solid blue line), 95% credible intervals (shaded blue), and experimental observations (black) for DIC, TA and C_p when photosynthesis, respiration and respiratory quotients were changing through time. X-axis is time in days.

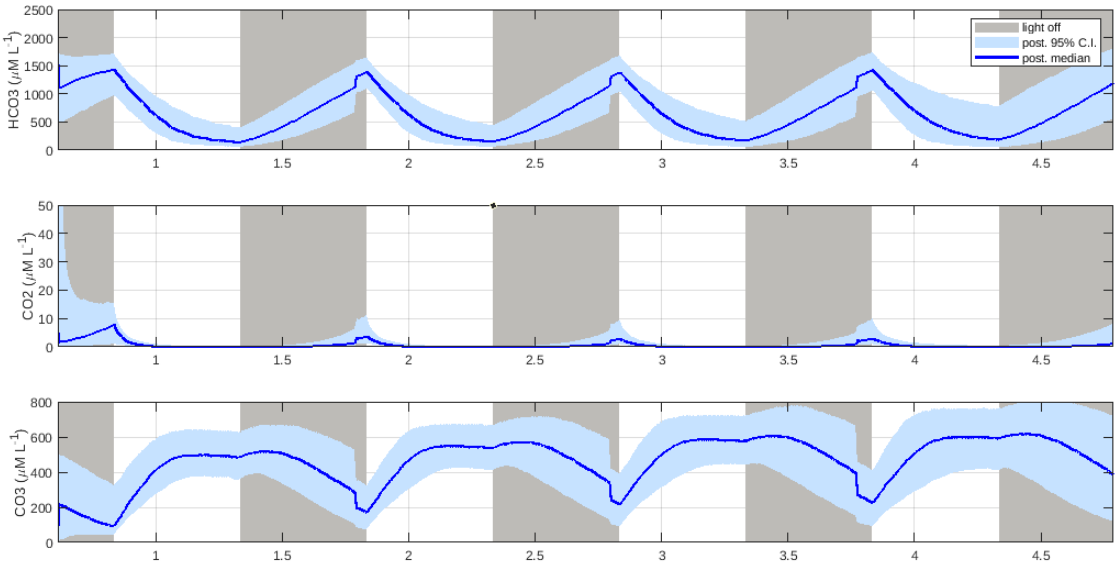

Figure 4.17 : Posterior medians (solid blue line) and 95% credible intervals (shaded blue) for HCO_3, CO_2 and CO_3 when photosynthesis, respiration and respiratory quotients were changing through time. X-axis is time in days.

state variable C_p posterior increased during each day, flattened during the nights and dropped suddenly with every dilution event (Figure 4.16).

The photosynthesis rate (P_1) centred around 240 μM L^{-1} hour^{-1} with small oscillations during daylight. The 95% credible intervals lay approximately between 150 μM L^{-1} hour^{-1} and 400 μM L^{-1} hour^{-1} with tighter intervals at the start of each day (Figure 4.18). The respiration rate (R_1) remained centred around 12 μM L^{-1} hour^{-1} with 95% credible intervals between 4 μM L^{-1} hour^{-1} and 30 μM L^{-1} hour^{-1} through the course of the experiment (Figure 4.18).

The daytime respiratory quotient RQ_d centred around 0.75 with a 95% credible interval of 0.6-1 during the first day of the experiment. The centre dropped to 0.6 during the 2nd, 3rd and 4th days, with 95% credible intervals of 0.4-0.8 (Figure 4.19). The night-time respiratory quotient RQ_n was initially centred around 0.85 and slowly

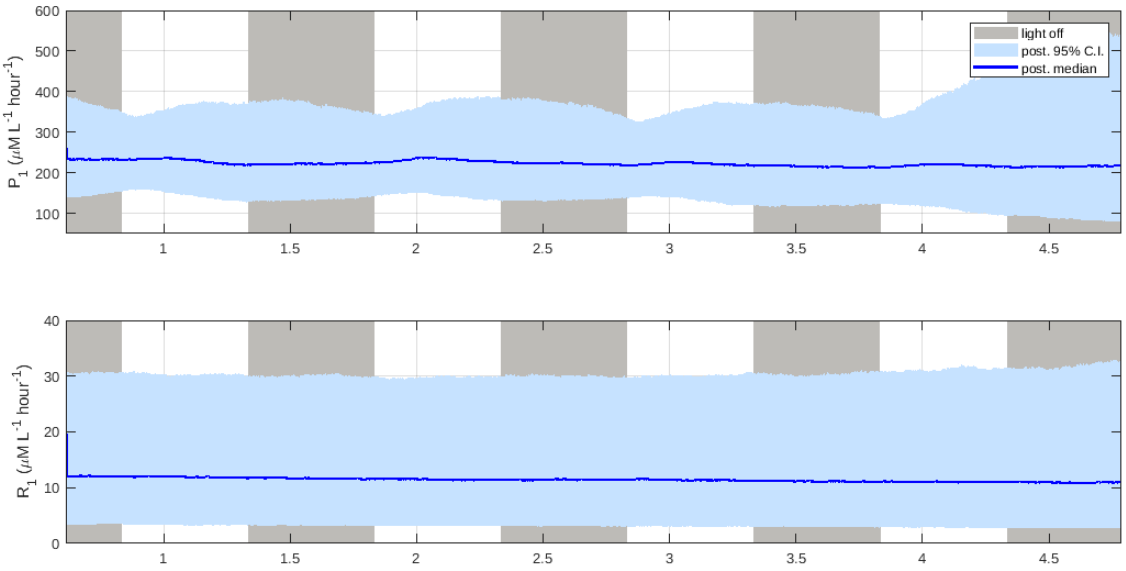

Figure 4.18 : Photosynthesis (P_1) and respiration (R_1) random walk posterior medians (solid blue line) and 95% credible intervals (shaded blue) after assimilating experimental data when photosynthesis, respiration and respiratory quotients were changing through time. X-axis is time in days.

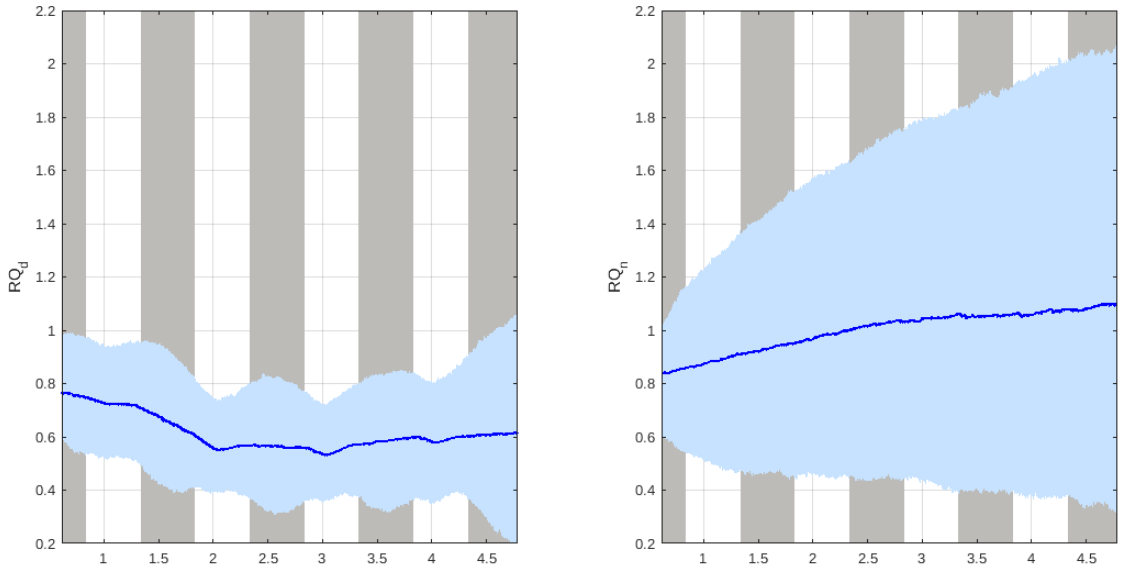

Figure 4.19 : Respiratory quotients (RQ_d and RQ_n) posterior medians (solid blue line) and 95% credible intervals (shaded blue) after assimilating experimental data when photosynthesis, respiration and respiratory quotients were changing through time. X-axis is time in days.

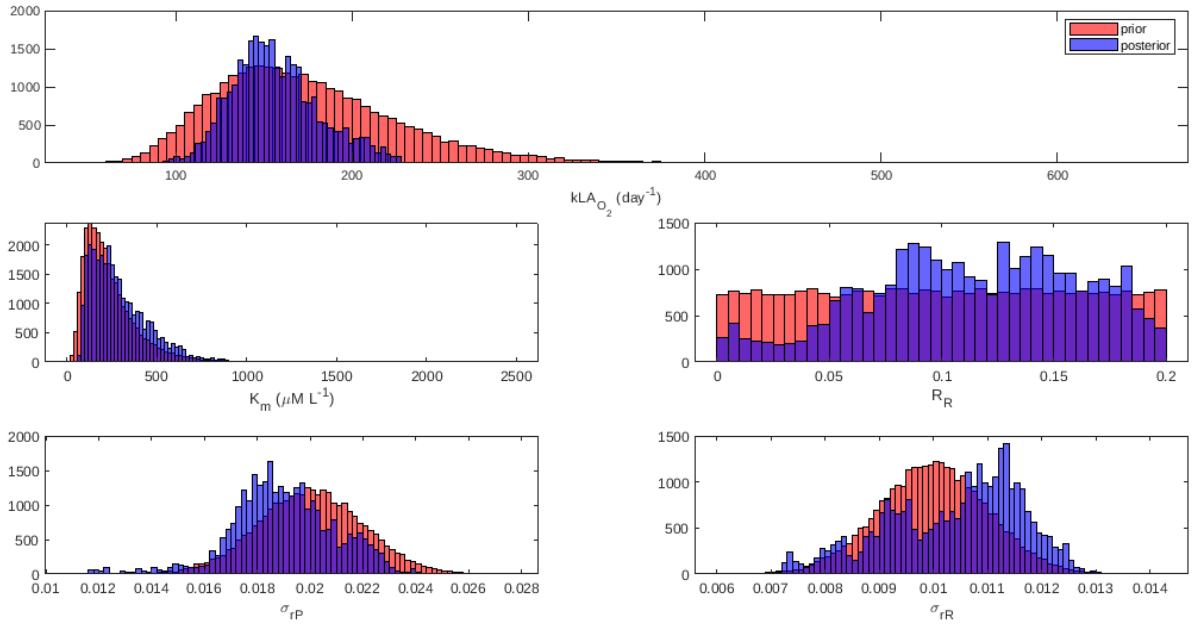

Figure 4.20 : Model parameter priors (orange), and posteriors (purple) after assimilating experimental data when photosynthesis, respiration and respiratory quotients were changing through time.

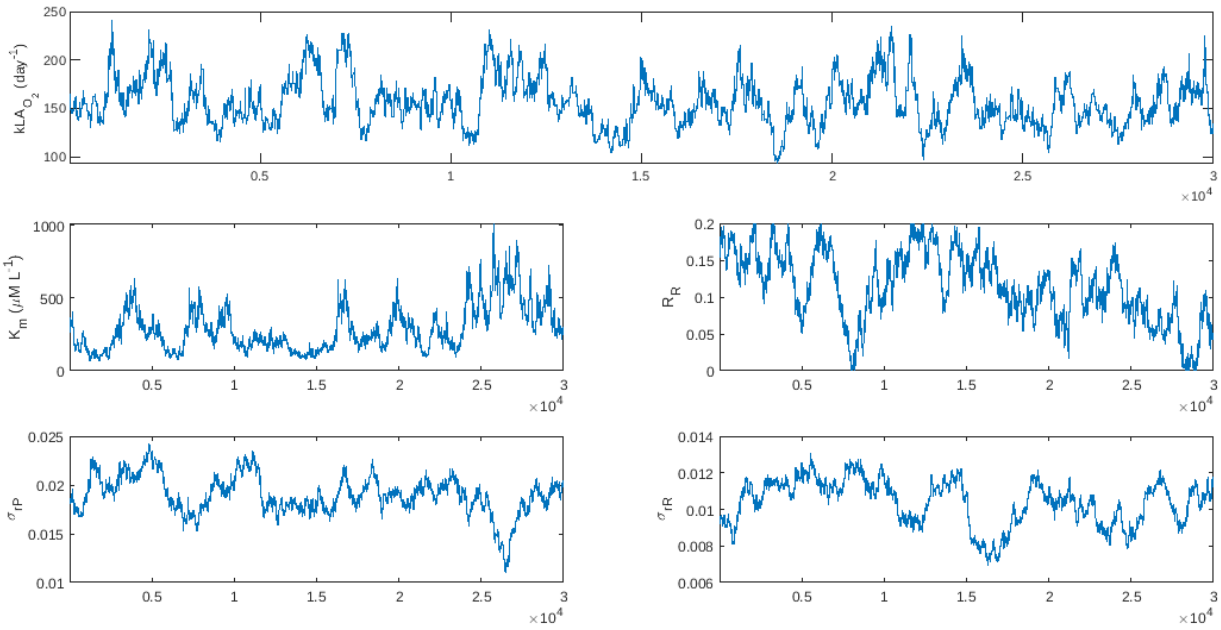

Figure 4.21 : Parameter trace-plots after assimilating experimental data when photosynthesis, respiration and respiratory quotients were changing through time.

rose to 1.1 by the end of the experiment. The 95% credible intervals started with 0.6-1 and increased its span steadily to 0.4-2 by the end of the experiment (Figure 4.19).

Parameter $\text{kLA}_{O_2}^{air}$ saw a more constrained posterior distribution with 25% and 75% quantiles (139.5979, 170.8102) (Table 4.8, Figure 4.20) compared to the assigned prior (Table 4.7). K_m had little change between prior and posterior distribution (Figure 4.20). While the R_R posterior distribution spanned the whole prior range (0, 0.2), there was higher density concentrated around the 25% and 75% quantiles (0.0815, 0.1512) (Figure 4.20, Table 4.8). Parameters σ_{r_P} and σ_{r_R} showed no large shifts from prior to posterior. Parameter trace-plots showed $\text{kLA}_{O_2}^{air}$, K_m, σ_{rP}, σ_{rR} and R_R that mixing needed to be improved, some autocorrelation between samples was present (Figure 4.21).

4.3.4 Posterior results with experimental data where photosynthesis, respiration and respiratory quotients are changing through time and an offset on O_2 is introduced

Photosynthesis (P_1) and respiration (R_1) were both modelled as random walks as described in Section 4.3.2, and similarly respiratory quotients RQ_d and RQ_n were also modelled as random walks (Section 4.3.3). An offset parameter ($offset_{O_2}$) was added to the O_2 ODE to see if any of the instrument mis-calibration could be quantified. Parameters kLA_{O_2}, K_m, R_R, σ_{rP}, σ_{rR}, and $offset_{O_2}$ were all treated as constant through time but unknown with prior distributions and proposal distributions defined in Table 4.9. The data model assigned log normally distributed observation errors for each instrument (Section 4.2.2) with tighter observation error standard deviation values cited in Table 4.9.

Table 4.9 : Parameters, priors and proposal distributions with experimental data when photosynthesis, respiration and respiratory quotients were changing through time and an O_2 offset was introduced (* indicates the parameter was held fixed).

Parameter	**Prior**	**Proposal**
kLA_{O2}	$\text{Log}\mathcal{N}(\log(200.0), 0.3)$	$\text{Log}\mathcal{N}(\log(\text{kLA}_{O2}), 0.021)$
K_m	$\text{Log}\mathcal{N}(\log(200.0), 0.6)$	$\text{Log}\mathcal{N}(\log(\text{K}_m), 0.042)$
R_R	$\mathcal{U}(0, 0.2)$	$\text{Trun}\mathcal{N}(\text{R}_R, 0.0035, \text{lower} = 0, \text{upper} = 0.2)$
σ_{r_P}	$\mathcal{N}(0.02, 0.002)$	$\mathcal{N}(\sigma_{r_P}, 0.00014)$
σ_{r_R}	$\mathcal{N}(0.01, 0.001)$	$\mathcal{N}(\sigma_{r_R}, 0.00007)$
offset_{O_2}	$\mathcal{N}(0, 5.0)$	$\mathcal{N}(\text{offset}_{O_2}, 0.5)$
σ_{O_2}	0.1	*
σ_{pH}	0.1	*
σ_{DIC}	0.2	*

Running 40,000 samples with 1,024 particles generally proved to generate fatal acceptance rates, so the PMMH was run for 20,000 samples with 2,048 particles, where 10,000 samples were discarded as burn-in. 20,000 samples was the maximum that could complete with 2,048 particles in the maximum time allocation on the HPC cluster.

Observed state variables DIC and TA posteriors perform well with all observations lying within the 95% credible intervals (Figure 4.23). O_2 posteriors tracked the observations excellently while the light was on, with all observations falling inside tightly constrained 95% credible intervals. During times when there was no light, each day the posteriors would fit the observations well to start and then potentially there was a sensor drift causing increasing observations that the model was not accounting for (Figure 4.22). pH captured most observations within the 95% credible intervals (Figure 4.22). Unobserved state variable C_p posterior increased during each day, flattened during the nights and dropped suddenly with every dilution event (Figure 4.23).

The photosynthesis rate (P_1) was centred between 200 and 250 μM L^{-1} $hour^{-1}$ oscillating subtly with daylight peaks. The 95% credible intervals lay approximately between 150 μM L^{-1} $hour^{-1}$ and 400 μM L^{-1} $hour^{-1}$ with tighter intervals at the start of each day (Figure 4.25). The respiration rate (R_1) remained centred around 9 μM L^{-1} $hour^{-1}$ with 95% credible intervals between 4 μM L^{-1} $hour^{-1}$ and 20 μM L^{-1} $hour^{-1}$ through the course of the experiment (Figure 4.25).

The daytime respiratory quotient RQ_d was centred around 0.7 with a 95% credible interval of 0.6-0.9 during the first day of the experiment. The centre dropped to 0.5 during the 2^{nd}, 3^{rd}, and 4^{th} days with tighter 95% credible intervals of 0.4-0.6 (Figure 4.26). The night-time respiratory quotient RQ_n was initially centred around 0.9 and slowly rose to 1.1 by the second day and continued around 1.1 for

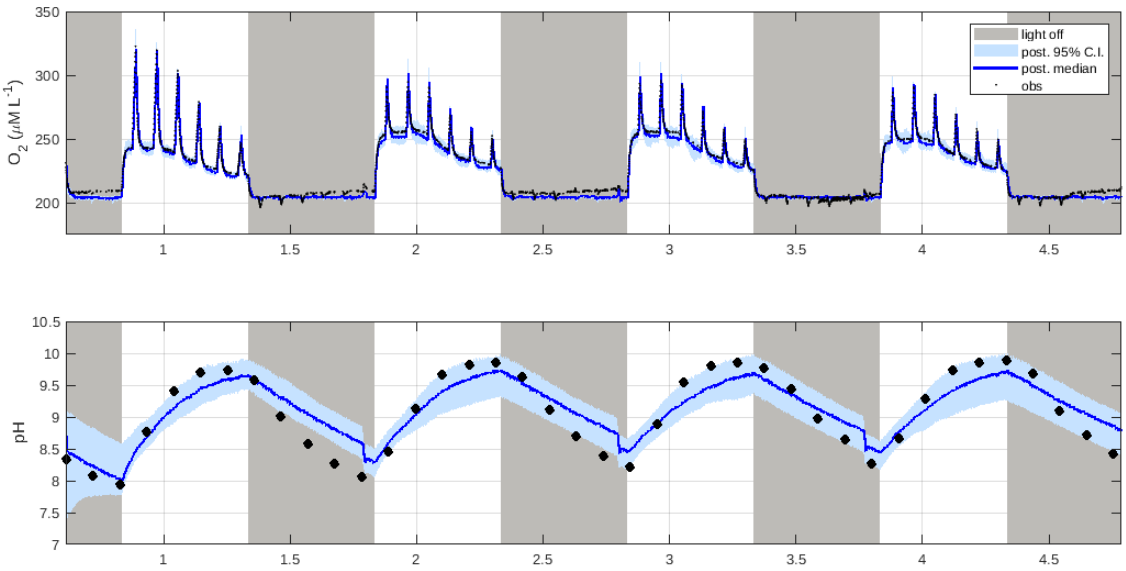

Figure 4.22 : Posterior medians (solid blue line), 95% credible intervals (shaded blue), and experimental observations (black) for O_2 and pH when photosynthesis, respiration and respiratory quotients were changing through time and an O_2 offset was introduced. X-axis is time in days.

Table 4.10 : Posterior (25%, 75%), (5%, 95%) quantiles for parameters after assimilating experimental data when photosynthesis, respiration and respiratory quotients were changing through time and an O_2 offset was introduced.

Parameter	**Quantiles (25%, 75%)**	**Quantiles (5%, 95%)**
$kLA_{O_2}^{air}$	(147.0307, 161.6729)	(139.0878, 171.0318)
K_m	(264.5106, 373.5021)	(157.8467, 485.1311)
R_R	(0.1538, 0.1701)	(0.1403, 0.1963)
σ_{r_P}	(0.0180, 0.0190)	(0.0176, 0.0194)
σ_{r_R}	(0.0097, 0.0100)	(0.0096, 0.0102)
$offset_{O_2}$	(0.3097, 2.4730)	(-1.2734, 3.5217)

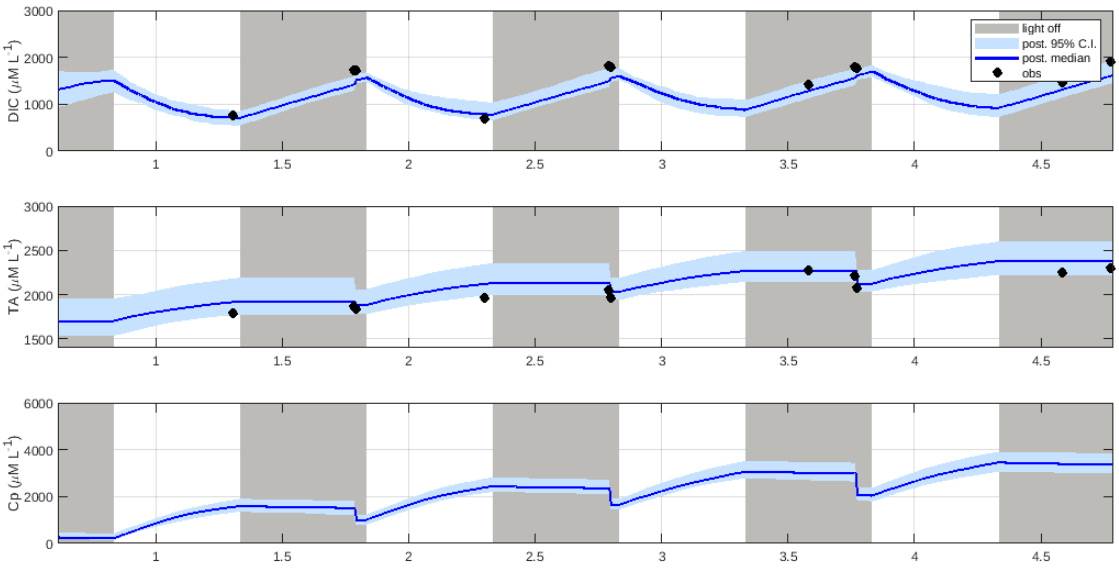

Figure 4.23 : Posterior medians (solid blue line), 95% credible intervals (shaded blue), and experimental observations (black) for DIC, TA and C_p when photosynthesis, respiration and respiratory quotients were changing through time and an O_2 offset was introduced. X-axis is time in days.

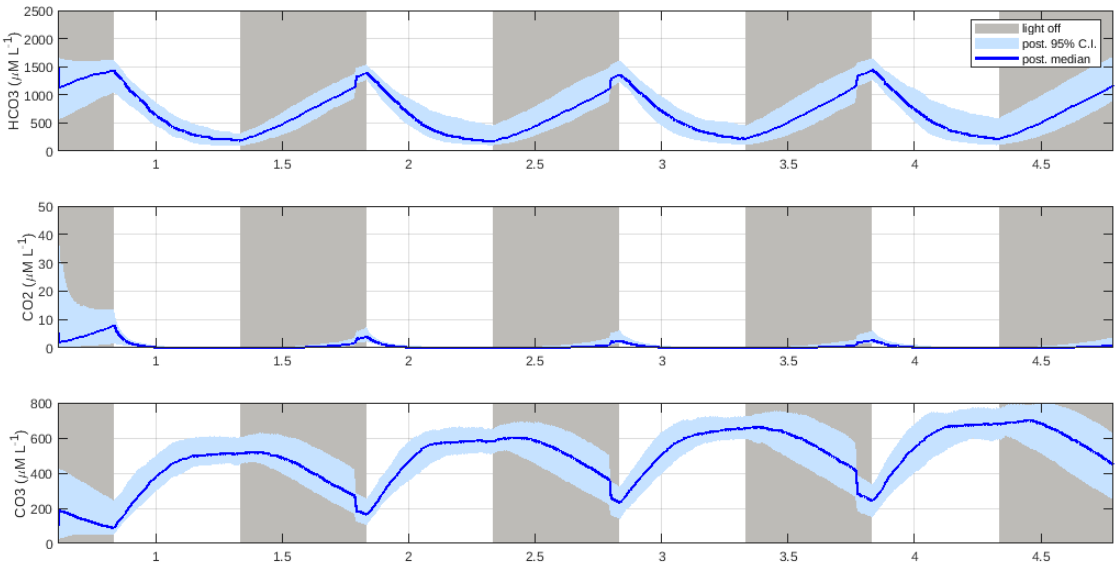

Figure 4.24 : Posterior medians (solid blue line) and 95% credible intervals (shaded blue) for HCO_3, CO_2 and CO_3 after assimilating experimental data when photosynthesis, respiration and respiratory quotients were changing through time and an O_2 offset was introduced. X-axis is time in days.

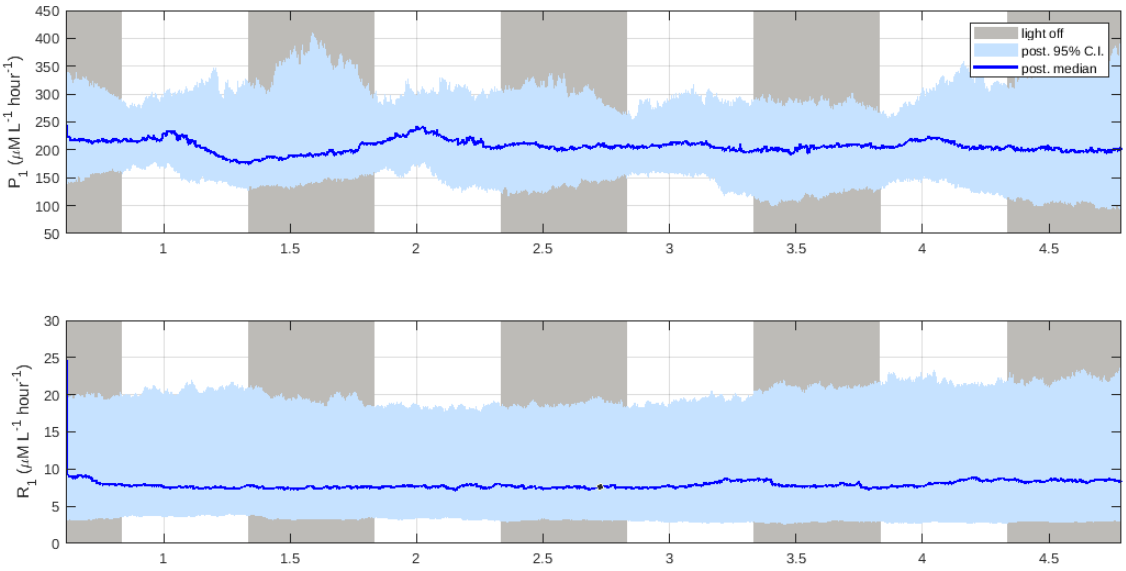

Figure 4.25 : Photosynthesis (P_1) and respiration (R_1) random walk posterior medians (solid blue line) and 95% credible intervals (shaded blue) after assimilating experimental data when photosynthesis, respiration and respiratory quotients were changing through time and an O_2 offset was introduced. X-axis is time in days.

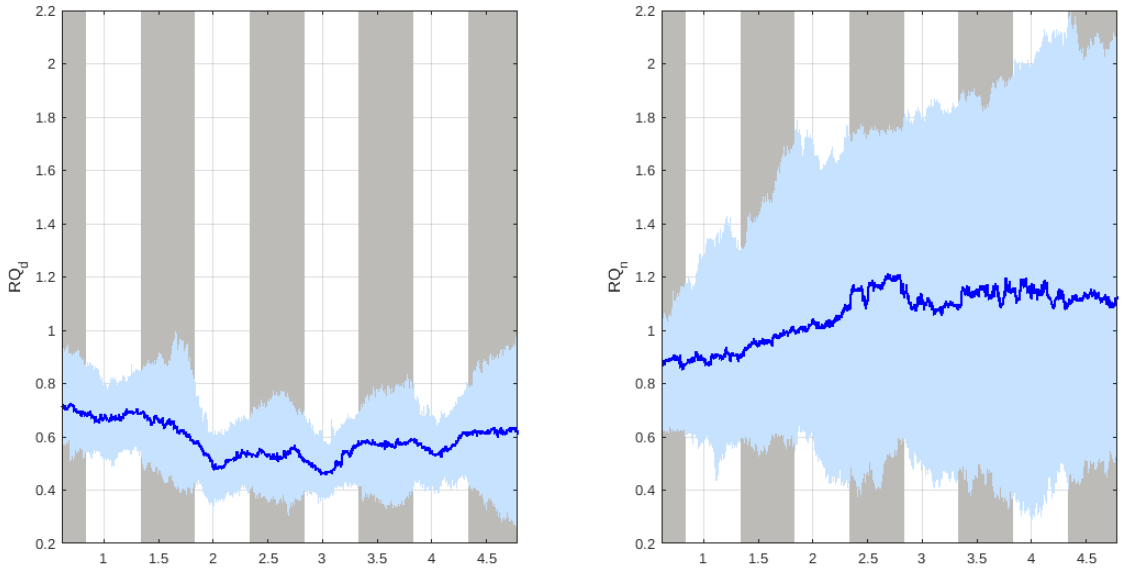

Figure 4.26 : Respiratory quotients (RQ_d and RQ_n) posterior medians (solid blue line) and 95% credible intervals (shaded blue) after assimilating experimental data when photosynthesis, respiration and respiratory quotients were changing through time and an O_2 offset was introduced. X-axis is time in days.

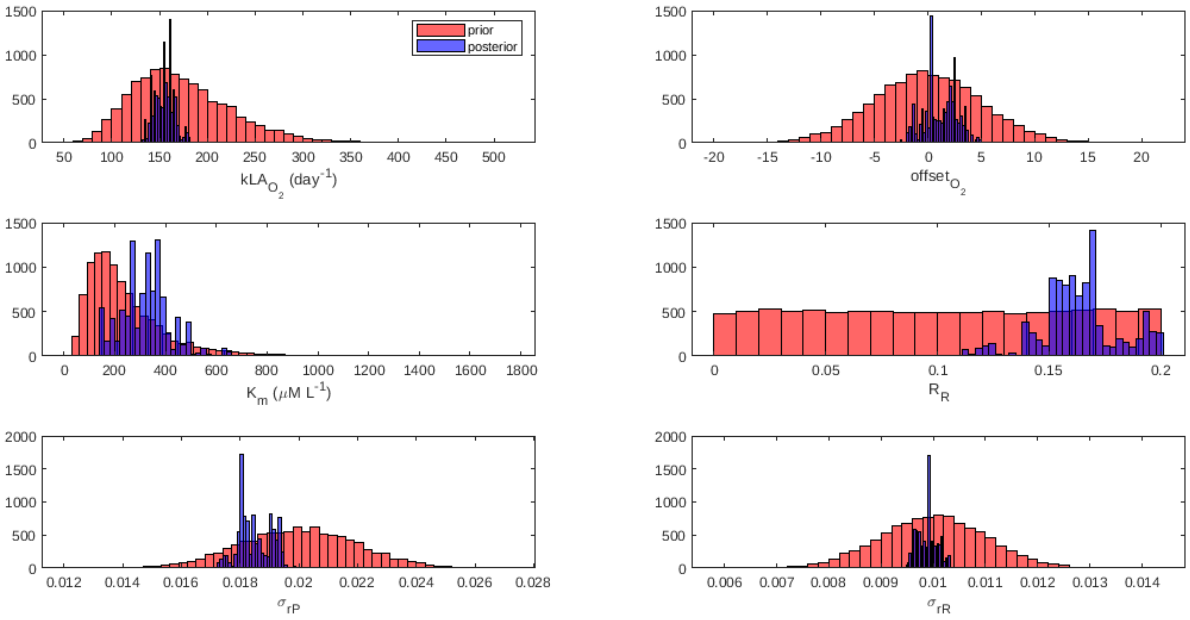

Figure 4.27 : Model parameter priors (orange) and posteriors (purple) after assimilating experimental data when photosynthesis, respiration and respiratory quotients were changing through time and an O_2 offset was introduced.

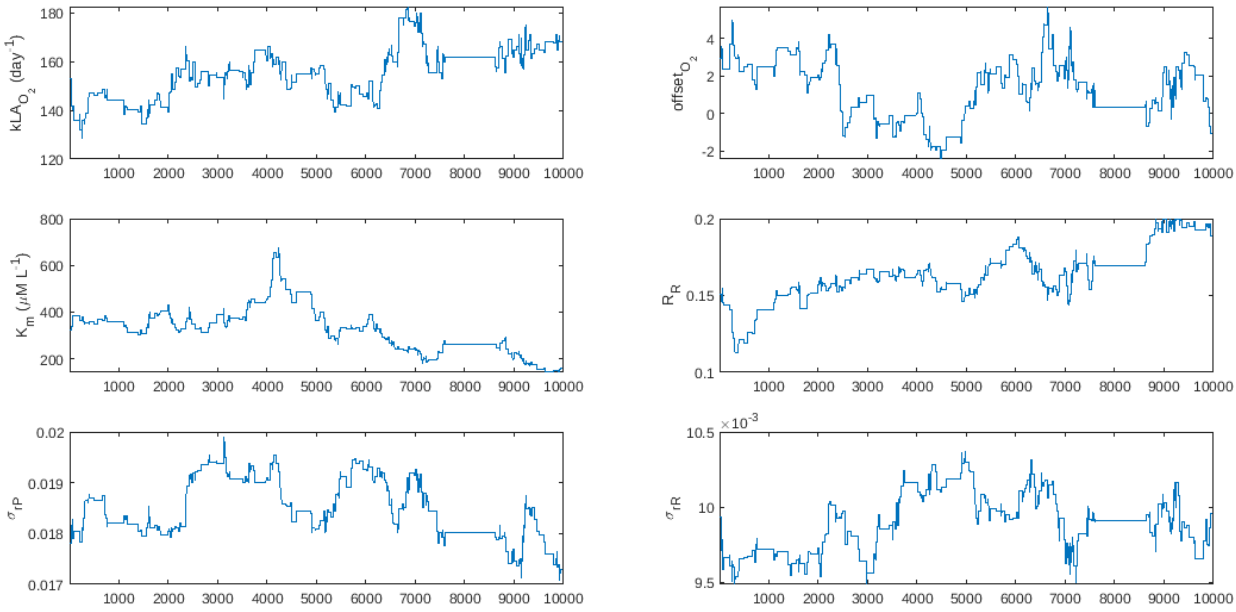

Figure 4.28 : Parameter trace-plots after assimilating experimental data when photosynthesis, respiration and respiratory quotients were changing through time and an O_2 offset was introduced.

the remainder of the experiment. The 95% credible intervals started with 0.6-1 and increased its span steadily to 0.5-2 by the end of the experiment (Figure 4.26).

Each parameter posterior showed learnings from the prior distributions with much tighter distributions in comparison (Figure 4.27). Parameter $\text{kLA}_{O_2}^{air}$ had a very tightly constrained posterior distribution with 25% and 75% quantiles (147.0307, 161.6729) (Table 4.10, Figure 4.27). K_m tightened its posterior distribution with 25% and 75% quantiles (264.5106, 373.5021) (Figure 4.27). The R_R posterior distribution had high density concentrations around the 25% and 75% quantiles (0.1538, 0.1701) (Figure 4.27, Table 4.10). Parameters σ_{r_P} and σ_{r_R} saw more constrained distributions from prior to posterior (Figure 4.27). The offset parameter offset_{O_2} showed a much tighter posterior distribution compared to prior (Figure 4.27), but spanned across 0 for the 5% and 95% quantiles (-1.2734, 3.5217).

Overall, mixing of the PMMH was poor with an acceptance rate of 3.57% and slowly mixing parameter traces with some correlation between samples that were unable to be improved by increasing particles or tweaking proposal distributions (Figure 4.28). Improving mixing is the focus of the next section (Section 4.3.5).

4.3.5 Posterior results with experimental data where photosynthesis, respiration and respiratory quotients are changing through time, O_2 has an offset and the O_2 observations were thinned further

This section aimed to test whether thinning out the O_2 observations further (Figure 4.29) helped improve the poorly mixed chain of Section 4.3.4. All priors and model specifications remained the same as Section 4.3.4, with the only difference being the further reduced O_2 dataset to 5% of the full observation set.

Similarly to Section 4.3.4, 20,000 samples were run (10,000 discarded as burn-in) with 2,048 particles. The acceptance rate improved to 13.61% with less dense observations.

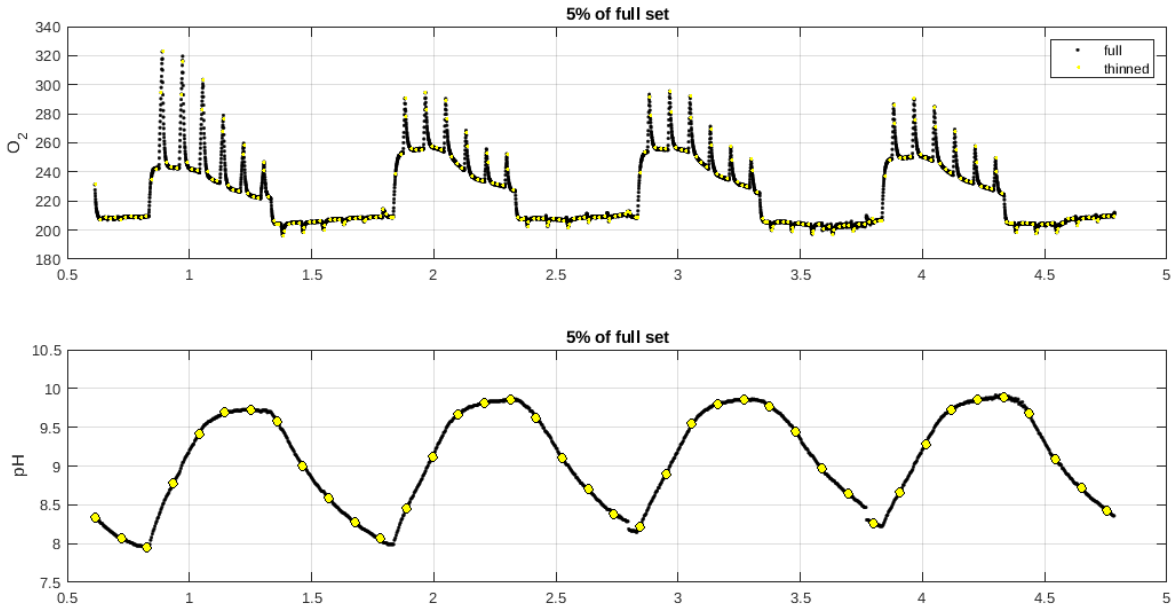

Figure 4.29 : Full O_2 (μM L^{-1}) and pH ($\log_{10}$(-mol/L H+)) datasets (black) with further thinned O_2 and pH observations (yellow). X-axis is time in days.

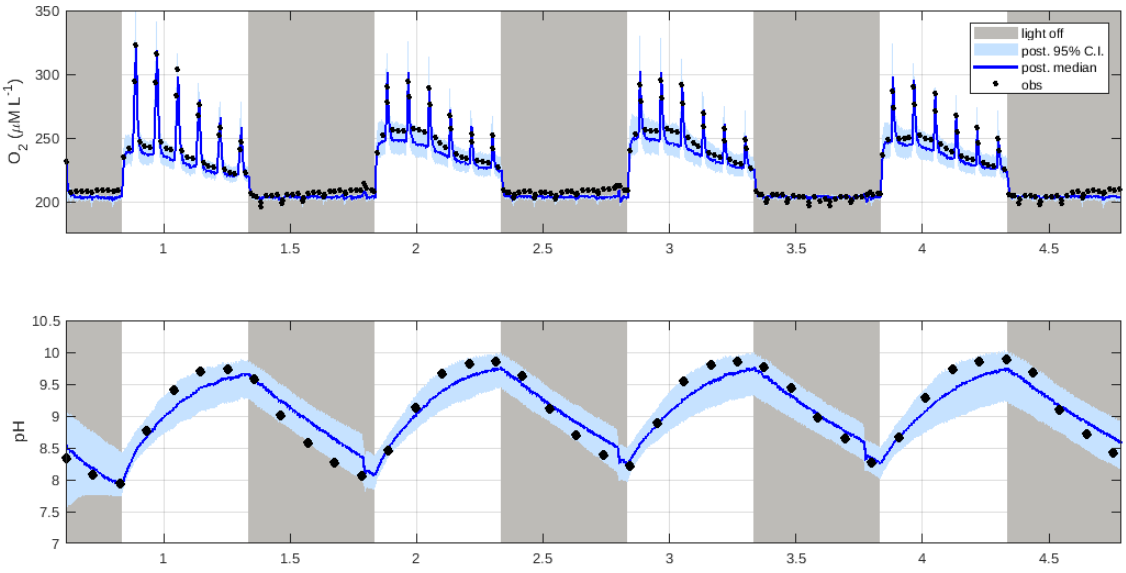

Figure 4.30 : Posterior medians (solid blue line), 95% credible intervals (shaded blue), and further thinned experimental observations (black) for O_2 and pH. X-axis is time in days.

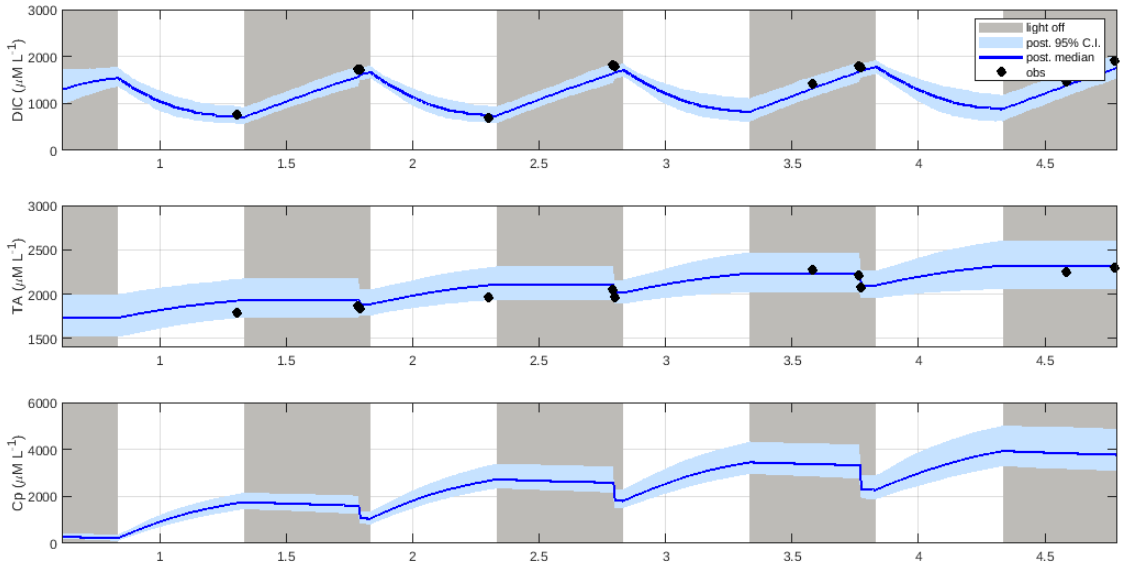

Figure 4.31 : Posterior medians (solid blue line), 95% credible intervals (shaded blue), and further thinned experimental observations (black) for DIC, TA and C_p. X-axis is time in days.

Table 4.11 : Posterior (25%, 75%), (5%, 95%) quantiles for parameters after assimilating further thinned experimental data.

Parameter	**Quantiles (25%, 75%)**	**Quantiles (5%, 95%)**
$kLA_{O_2}^{air}$	(172.4048, 202.2482)	(153.4338, 227.8861)
K_m	(224.2554, 352.8140)	(183.9269, 506.7944)
R_R	(0.0939, 0.1527)	(0.0639, 0.1684)
σ_{r_P}	(0.0190, 0.0220)	(0.0176, 0.0229)
σ_{r_R}	(0.0091, 0.0096)	(0.0084, 0.0099)
$offset_{O_2}$	(-3.6927, 7.7906)	(-7.4953, 10.7980)

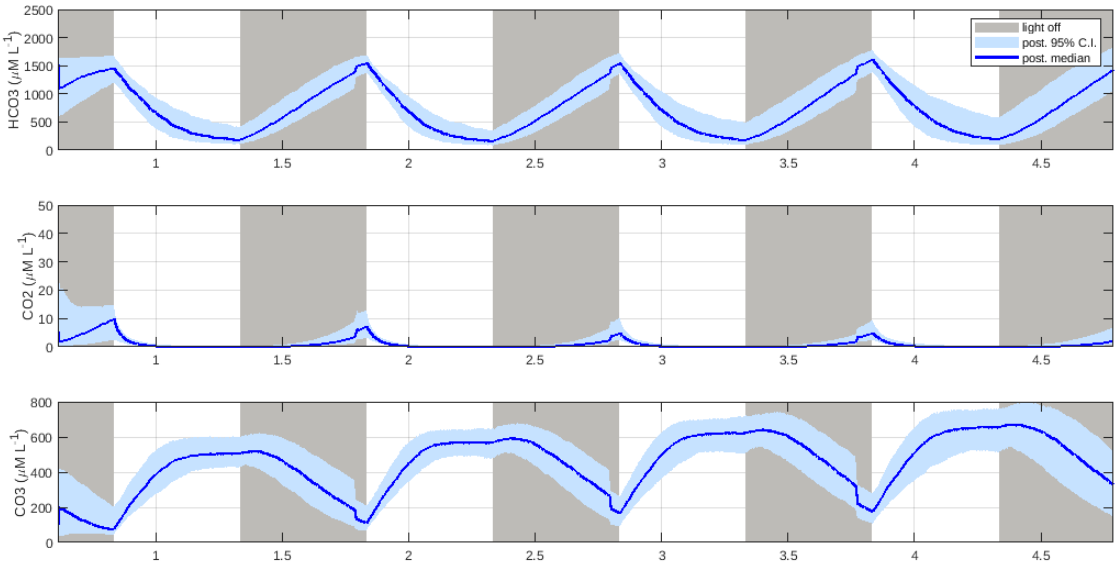

Figure 4.32 : Posterior medians (solid blue line) and 95% credible intervals (shaded blue) for HCO_3, CO_2 and CO_3 after assimilating further thinned experimental data. X-axis is time in days.

There was an improvement in pH posterior with the 95% credible interval now capturing all observations (Figure 4.30). Visually there does not appear to be much difference in O_2 posterior with both runs capturing the same observations within their 95% credible intervals. The credible intervals are wider during the day for the further thinned observations when compared to Figure 4.22. There was not much difference between posteriors for DIC, TA, and C_p (Figure 4.31) apart from the C_p 95% credible interval being wider when there are less O_2 observations. Similarly, the carbon chemistry behaved similarly in both sections (Figure 4.32).

P_1 behaved similarly across sections, with more smoothness through the median and slightly wider credible intervals when less O_2 observations were assimilated. R_1 notably was centred around 15 μM L^{-1} $hour^{-1}$ with 95% credible intervals 5-30 μM L^{-1} $hour^{-1}$ (Figure 4.33) compared to 9 μM L^{-1} $hour^{-1}$ with 95% credible intervals

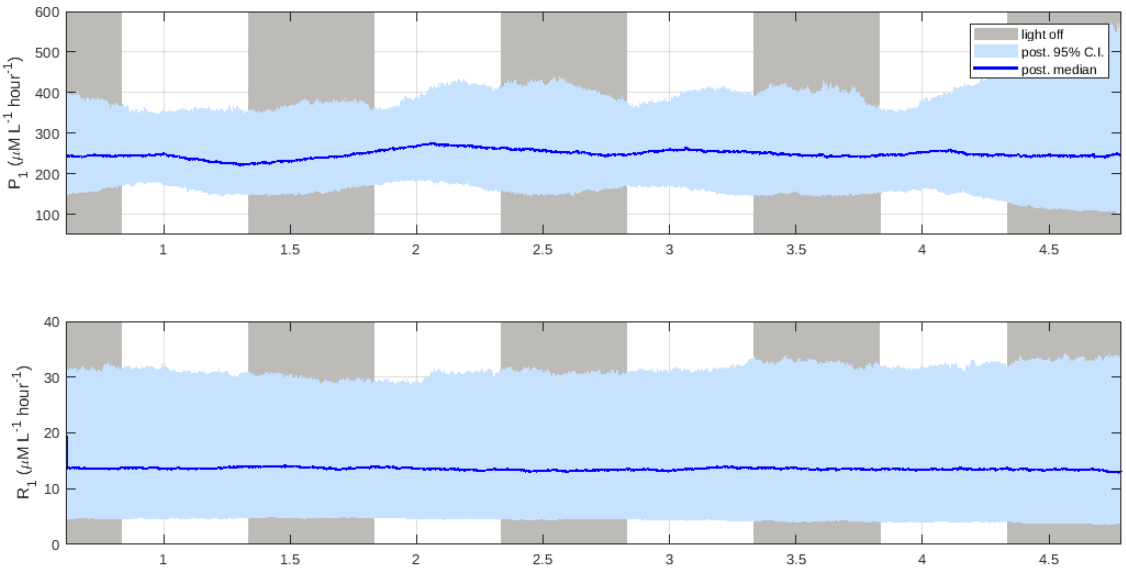

Figure 4.33 : Photosynthesis (P_1) and respiration (R_1) random walk posterior medians (solid blue line) and 95% credible intervals (shaded blue) after assimilating further thinned experimental data. X-axis is time in days.

between $4\,\mu$M L^{-1} $hour^{-1}$ and $20\,\mu$M L^{-1} $hour^{-1}$ (Figure 4.25). Respiratory quotients had smoother medians throughout the experiment with less O_2 data (Figure 4.34), otherwise not much change was observed.

$kLA_{O_2}^{air}$ posterior shifted to higher values compared to Section 4.3.4. $offset_{O_2}$ had a wider posterior than that of the previous section, but still centred around approximately the same value. K_m posterior shifted to lower values while R_R displayed a wider posterior distribution than the previous section as well as being centred lower. Posteriors for σ_{rP} and σ_{rR} were both wider when there were less O_2 observations to be assimilated (Figure 4.35). Parameter traces improved their mixing (Figure 4.36) from Section 4.3.4, but still displayed correlation between samples.

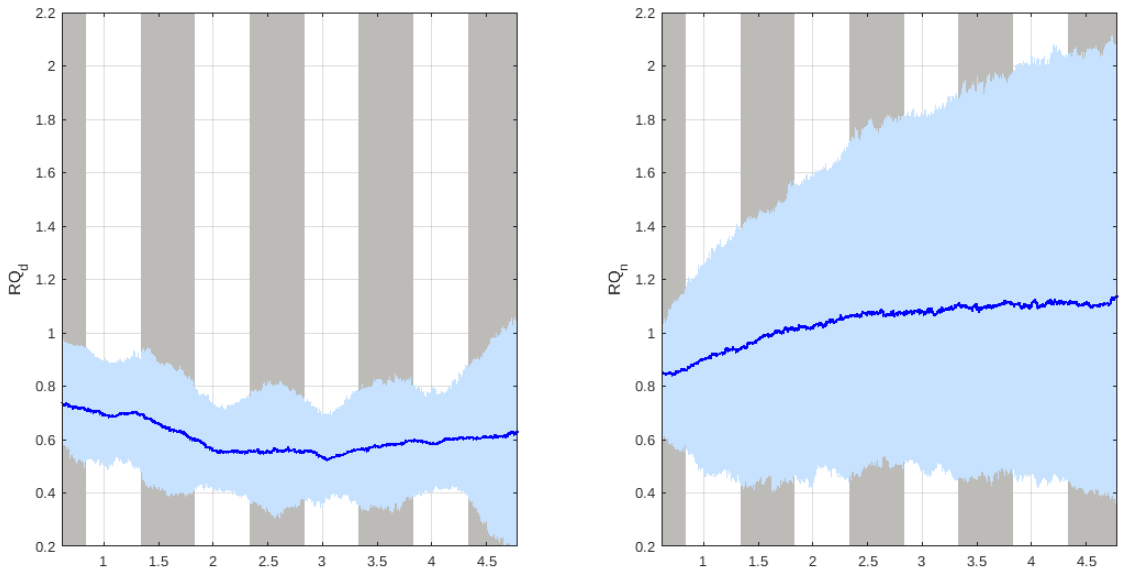

Figure 4.34 : Respiratory quotients (RQ_d and RQ_n) posterior medians (solid blue line) and 95% credible intervals (shaded blue) after assimilating further thinned experimental data. X-axis is time in days.

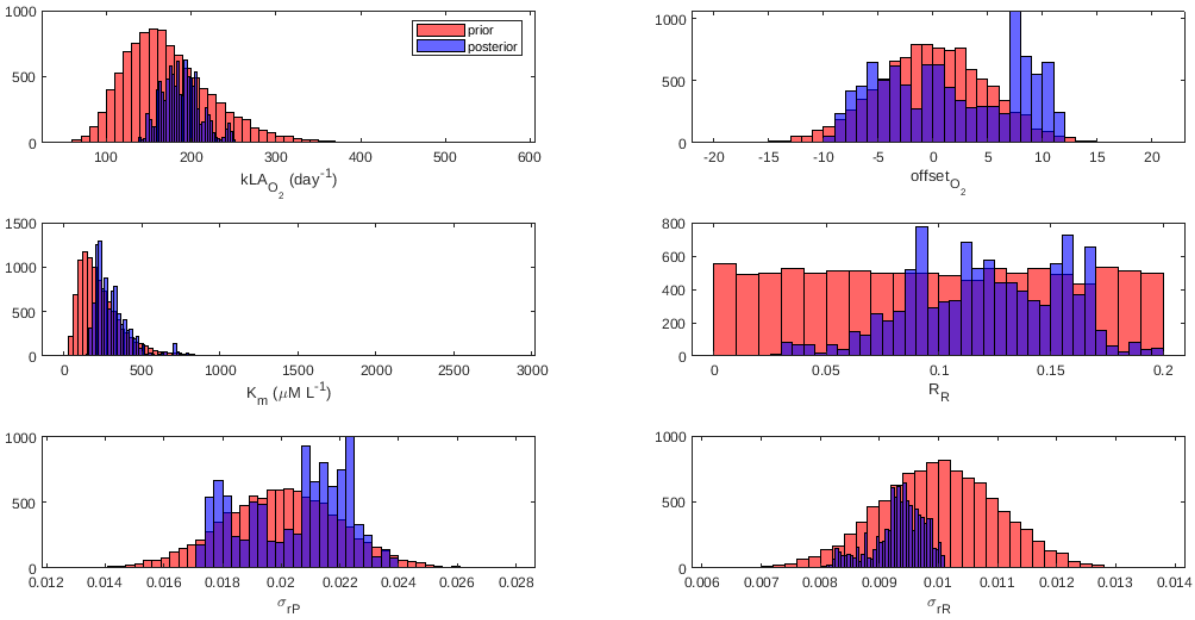

Figure 4.35 : Model parameter priors (orange) and posteriors (purple) after assimilating further thinned experimental data.

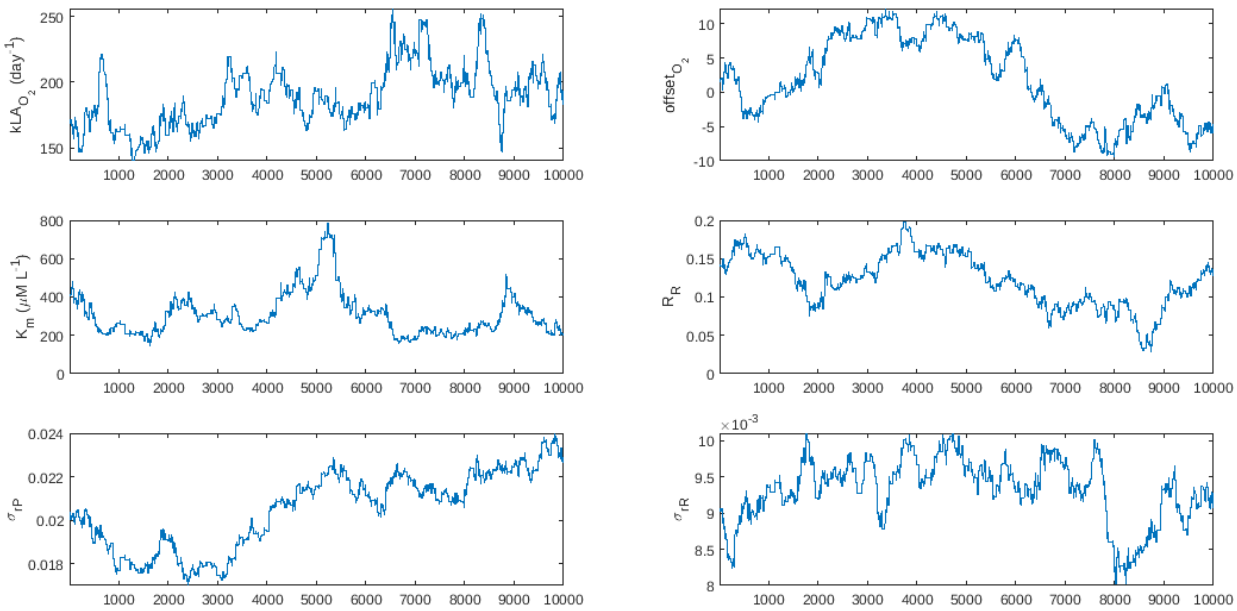

Figure 4.36 : Parameter trace-plots after assimilating further thinned experimental data.

4.4 Discussion

Non-destructive measurements allowed us to use sensors to obtain high temporal frequency measurements instead of destructive sampling techniques. Consequently, we obtained informative estimates from sensor measurements that combined with a data assimilating model informed on the main processes of interest without the expense of destructive sampling and limited data-points. This provided a significant advantage in the attempt to improve the MCMC mixing and computational run-time of the PMMH as it allowed for thinning of observations while still capturing the underlying signal. An interesting area of future work could be to calculate the critical number of data points needed before diminishing returns from higher sampling rates and whether these points need to be localised and dependent on the underlying process (e.g. photosynthesis, gas valve).

While the pH measurements were calibrated and corrected, it was more difficult to assess the bias from O_2 instrument errors as it was completely unmeasured. Where the noise on instrument errors is easily measurable, the bias is not. This was the motivation for introducing an offset on the O_2 measurements in Section 4.3.4 and included when testing a further thinned dataset in Section 4.3.5. During early assimilation runs, it was visible that model states were sensitive to small perturbations in parameters, which makes addressing instrument bias even more essential.

During the model formulation, the respiratory quotients RQ_d and RQ_n were assigned random bounded walks between biologically feasible values 0.6 and 1 to restrict their wandering into biologically unexplainable spaces. These runs did not result in successfully running posteriors as RQ_n was pushing against the top bound and RQ_d was pushing against the bottom bound with nearly no proposed jumps being accepted. It was decided to allow them to randomly walk unbounded to see which spaces they would enter and if it improved model performance. This resulted in RQ_n spanning into values larger than 1, while RQ_d centred around 0.6 for 3 out of the 4 experiment days and had credible intervals spanning down to 0.4. One possible explanation for the respiratory quotient posteriors going into values outside those that biology would allow (larger than 1 or less than 0.6) was that, in reality, there were more processes taking place than those in the model description. The respiratory quotient could have been compensating for these un-modelled processes, for example the parameter RQ_n can compensate for a drop or increase in the input O_2 gas line. In the model, O_2 concentration is treated as fixed when in reality this is a further un-modelled process that could have diurnal changes, Henry's law shifts, or physical lab components such as increases because there are fewer people there during the night than during the day allowing it higher values. For future work, as much of these components can and should be measured to adjust the observational

datasets rather than increasing the complexity of the process model and not having informative measurements to constrain those extra processes. Another possible reason for unrealistic respiratory quotient posterior values could be highly correlated model parameters, which is a regular result of non-linear ecosystem models [56, 67, 99, 93]. This can cause problems for our data assimilation technique. Sensitivity analyses and degrees of correlation between model parameters are important techniques that assist in helping define a Bayesian hierarchical model and have been undertaken in previous parameter estimation data assimilating literature [30, 37, 67]. For example, an error-covariance matrix from a simulated annealing optimisation routine, gave both the sensitivity of the model to each model parameter and the correlation coefficients between all parameter pairs [67]. An alternative sensitivity analysis was used to estimate the degrees of correlation between parameters [37], while a third used variations in log-likelihood to define sensitivity [30]. Employing a similar technique to these studies would be an informative avenue for future work.

There were many attempts made to improve the mixing of the MCMC chains in each results section, such as tuning the proposal distribution step size, increasing the number of particles, and further thinning out very dense datasets (O_2). The final runs presented here are the best possible achieved within computational resources available. There was still autocorrelation present between samples which could have been addressed by running the MCMC chain for longer, say 100,000 to 150,000 to deal with the autocorrelated samples. This was not presently possible because of the computationally expensive nature of the PMMH, with every increase in particles or samples requiring an increase in computational resources. Since the same poor mixing issues are present in the twin experiments, the issue isn't likely model misspecification (similar problems persist with real and synthetic data). An alternative area of investigation for resolving the poor mixing could be the issue of parameter identifiability. When parameters are able to trade off against one another,

then mixing will be compromised unless the offending parameters are individually tightly constrained by the priors.

One of the major achievements of this chapter was being able to successfully and accurately resolve CO2SYS in LibBi. This had two major impacts, the first was depicting the carbon state of the system accurately as CO2SYS was sensitive to producing unfeasible carbon values like negative HCO_3 for example, due to a small discrepancies when attempting to find a viable approximation. The sensitivity was present on different carbon concentration scales and must be accurate as we had large values for HCO_3 (in thousands) and tiny values for CO_2 (less than 10), while pH is on a log scale meaning that errors in approximation for one are more consequential than the other. The flow-on effects are present in the fluxes because the carbon state of the system is what drives them. The second major impact is that successfully implementing an accurate representation of CO2SYS within LibBi constraints which are designed for parallel processing, and simplicity limitations because of their use of GPU, opens up LibBi to more complex models and equations. These limitations are the same across other Bayesian inference software like Stan and the approach used in this chapter can be applied there also.

Chapter 5

Concluding remarks

This thesis illustrates data assimilation applications for three aquatic ecosystem processes. Data assimilation of Port Phillip Bay seagrass meadows for state and parameter estimation and prediction in Chapter 2 demonstrated the benefit of a quantitative meta-analysis of parameter values currently found in literature for the construction of informative priors. This approach proved particularly effective despite a complex process model and small dataset. It demonstrated the ability of the light limiting seagrass growth model from EMS to accurately predict above-ground biomass and proportion cover at three out of five sites in Port Phillip Bay for the year 2013 when observations from 2012 had been assimilated. Of the three sites with successful predictions for 2013, all three exhibited above-ground biomass and proportion cover seasonality with a downward trend across 2012 and 2013. Point Richards seagrass meadows were found to have the highest effective absorbance and highest radiation use efficiency of the three sites across the 2 year period. The highest translocation rate of the three sites was estimated at Blairgowrie. Swan Bay South seagrass meadows had the highest leaf mortality rate and highest dry weight specific area of seagrass. Assimilating additional environmental forcing data such as temperature, rainfall, and nutrient loads is a key extension to this work and will provide insight into why the model predictions at Altona and Point Henry were unsuccessful. Accurate models and predictions are of particular significance to examining seagrass dynamics, determining causes of decline and informing environmental management practices with flow on effects to quantifying carbon sequestration, improvements to water quality and providing habitats for coastal fisheries.

Chapter 3 presented an alternative approach to modelling seagrass decay rates than currently adopted as general practice. Without assuming a functional form for decay it allowed a more flexible formulation without segregating the pool of decomposing biomass into discrete subsections, each decaying at a constant rate. As a proof of concept, it was applied to two datasets; one in-situ study and one laboratory experiment of temperature and nutrient effects on *Zostera muelleri* detrital decay. The results illustrated that in the presence of data, various complex decay dynamics were captured with the time-varying model that were not captured with the conventional multi component decay. Specifically, the time-varying decay rate model demonstrated the ability to capture 3 distinct decay rate profiles through time. Soft peaks with slow increase and decreases before and after in the field study. Approximately constant and linearly increasing decay rates in the experiment leaf tissue, and a sharp spike and rapid decrease in decay rates that then flat line close to zero in the experiment root/rhizome tissue. This demonstrated the flexibility of the framework to capture diverse and complex decay dynamics which would have otherwise been lost in the traditional multi-component model. The field study saw the impacted site with leaf tissue exhibited the earliest decay rate peak, at least 1 month prior to the other site and tissue combinations, and also the only site and tissue combination to have a second decay rate peak. The laboratory experiment showed an increasing decay rate through time at high temperature in leaf tissue with no additional nutrients compared to the lower and approximately constant rate of decay at ambient temperature. Additional nutrient treatments on leaf tissue results showed no noticeable differences between ambient and high temperatures. No differences in root/rhizome tissue results were seen between temperature or nutrient treatments. Indeed, the largest difference in decay rates was visible between tissue types.

Chapter 4 demonstrated how combining high temporal frequency sensor data

with a data assimilating model informed on underlying processes without the expense of destructive sampling and limited data-points. The posterior state estimates were in excellent agreement with the high frequency observations, in turn producing insightful parameter posteriors. Photosynthesis and respiration, the key processes underpinning algal growth were estimated over time and centered around 240 μM L^{-1} $hour^{-1}$ and 12 μM L^{-1} $hour^{-1}$ with 95% credible intervals between 150 μM L^{-1} $hour^{-1}$ and 400 μM L^{-1} $hour^{-1}$ and 4 μM L^{-1} $hour^{-1}$ and 30 μM L^{-1} $hour^{-1}$ respectively. Having an accurate model is key to the successful scaling of microalgae production as a candidate in the industry of renewable sources of biofuels. Their high biomass cultivation can also be used to mitigate carbon dioxide and other greenhouse gases as microalgae photosynthesise and grow. Another key achievement of this chapter was accurately resolving CO2SYS within LibBi using a hybrid functional approximation and three iterations of the Newton-Raphson method resulted in an accurate depiction of the carbon state of the system. It also opens up LibBi, and other GPU software, to more complex models and equations using the same approach.

The computational effort to successfully carry out inference on Bayesian state-space models for aquatic ecological processes is significant even with programs like LibBi that are specifically designed to carry out these tasks. Debugging probabilistic programs remains challenging and errors can arise from ignoring dependencies and correlation between variables, poorly chosen parameters and incorrect statistical models [73]. Non-linearity in ecosystem models regularly results in highly correlated model parameters [67]. This can cause problems for our data assimilation technique. Sensitivity analyses and degrees of correlation between model parameters are important techniques that assist in helping define a Bayesian hierarchical model and have been undertaken in previous parameter estimation data assimilating literature [30, 37, 67]. For example, an error-covariance matrix from a simulated annealing

optimisation routine, gave both the sensitivity of the model to each model parameter and the correlation coefficients between all parameter pairs [67]. An alternative sensitivity analysis was used to estimate the degrees of correlation between parameters [37], while a third used variations in log-likelihood to define sensitivity [30]. Employing a similar technique to these studies would be an informative avenue for future work.

Tuning proposal distributions of the Metropolis-Hastings algorithm are required to reach convergence of MCMC chains. Determining which proposal is best for a particular target distribution is both important and difficult for high dimensions, and usually approached in an ad-hoc manner involving trial and error. The biogeochemical seagrass model (Chapter 2), and the microalgae rate transfer model (Chapter 4) mixing results were not able to be improved further using this approach. There are now techniques that estimate optimal proposal scalings and adaptive algorithms to attempt to find good proposals automatically called adaptive MCMC which automatically "learn" better parameter values "on the fly" (while an algorithm runs) [83]. Adaptive MCMC algorithms have been used to successfully speed up convergence on high-dimensional test problems [81]. Taking into consideration how much time and effort ad-hoc tuning of proposal distribution scalings took during this project, trying an adaptive MCMC algorithm would be the potential direction for future work. Like we saw in Chapter 4, MCMC can often be slow especially if the chain moves slowly, with many iterations and needing to compute or estimate the likelihood at every iteration. Hamiltonian Monte Carlo [31] is another alternative MCMC algorithm that quickly explores high dimensional parameter spaces at the cost of a large number of gradient evaluations. Sub-sampling MCMC is an even faster alternative that estimates the likelihood from a subsample in each MCMC iteration [24]. An additional dimension for consideration is the rapidly evolving field of artificial intelligence and machine learning, which may provide algorithms for use

in constraining these types of models.

Improving initial conditions is another way of improving mixing of the MCMC chain. Apart from doing this by completing some runs and manually adjusting and trialling different initial conditions, there are optimisation schemas that can be run, but are not currently set up in LibBi to sensibly work with a particle filter. Running the chain for longer, say 100,000 or 150,000 samples would also have improved mixing and autocorrelation but we were limited by computational constraints for consecutive jobs on the HPC. Running the chain for longer, then thinning out the samples is another way of dealing with autocorrelation, similarly, this would have also required more computational effort than is currently available.

In conclusion, data assimilation is an effective methodology for quantitatively combining the knowledge gained from previous studies with aquatic ecological models and data.

Appendix A

LibBi model code and MCMC diagnostics for Port Phillip Bay seagrass meadows

LibBi model file: seagrass.bi

```
model seagrass{

const h_sed= 0.15

const St= 10.0

const phi= 0.5


input E_wc


param SDF

param r_ue

param f_b

param h_wc

param omega


state S_A

state S_B

state P_C

state D

state E_c

state k_I

state k_resp
```

```
state light

state DIC

state O_2

state z1

state g_SA

state g

state trans

state tau

state zeta_A

state zeta_B

state E_comp

state A

state k_d

state h_c


noise r_tau, r_zeta_A, r_zeta_B, r_E_comp, r_A, r_k_d, r_h_c


param mu_tau, mu_zeta_A, mu_zeta_B, mu_E_comp, mu_A, mu_k_d, mu_h_c


const theta_tau = 0.95

const theta_zeta_A = 1.0

const theta_zeta_B = 1.0

const theta_E_comp = 0.5

const theta_A = 0.5

const theta_k_d = 0.6

const theta_h_c = 0.5
```

```
inline sigma_tau = sqrt(log(1.0 + (2.0*St - 1.0)*SDF*SDF*(exp(
   theta_tau*theta_tau) - 1.0)))
inline sigma_zeta_A = sqrt(log(1.0 + (2.0*St - 1.0)*SDF*SDF*(exp(
   theta_zeta_A*theta_zeta_A) - 1.0)))
inline sigma_zeta_B = sqrt(log(1.0 + (2.0*St - 1.0)*SDF*SDF*(exp(
   theta_zeta_B*theta_zeta_B) - 1.0)))
inline sigma_E_comp = sqrt(log(1.0 + (2.0*St - 1.0)*SDF*SDF*(exp(
   theta_E_comp*theta_E_comp) - 1.0)))
inline sigma_A = sqrt(log(1.0 + (2.0*St - 1.0)*SDF*SDF*(exp(theta_A*
   theta_A) - 1.0)))
inline sigma_k_d = sqrt(log(1.0 + (2.0*St - 1.0)*SDF*SDF*(exp(
   theta_k_d*theta_k_d) - 1.0)))
inline sigma_h_c = sqrt(log(1.0 + (2.0*St - 1.0)*SDF*SDF*(exp(
   theta_h_c*theta_h_c) - 1.0)))

inline gamma_tau = log(mu_tau) + pow(theta_tau, 2.0)/2.0 - pow(
   sigma_tau, 2.0)/2.0
inline gamma_zeta_A = log(mu_zeta_A) + pow(theta_zeta_A, 2.0)/2.0 -
   pow(sigma_zeta_A, 2.0)/2.0
inline gamma_zeta_B = log(mu_zeta_B) + pow(theta_zeta_B, 2.0)/2.0 -
   pow(sigma_zeta_B, 2.0)/2.0
inline gamma_E_comp = log(mu_E_comp) + pow(theta_E_comp, 2.0)/2.0 -
   pow(sigma_E_comp, 2.0)/2.0
inline gamma_A = log(mu_A) + pow(theta_A, 2.0)/2.0 - pow(sigma_A,
   2.0)/2.0
inline gamma_k_d = log(mu_k_d) + pow(theta_k_d, 2.0)/2.0 - pow(
   sigma_k_d, 2.0)/2.0
```

```
inline gamma_h_c = log(mu_h_c) + pow(theta_h_c, 2.0)/2.0 - pow(
   sigma_h_c, 2.0)/2.0

obs S_A_obs
obs P_C_obs
obs k_d_obs
obs h_c_obs


sub parameter {
SDF ~ log_normal(log(0.2), 0.35)
r_ue ~ log_normal(log(0.4), 0.5)
f_b ~ normal(0.7721, 0.0273)
h_wc ~ log_normal(log(0.5), 0.2)
omega ~ log_normal(log(0.005), 1.0)

mu_tau ~ log_normal(log(0.001), theta_tau)
mu_zeta_A ~ log_normal(log(0.005), theta_zeta_A)
mu_zeta_B ~ log_normal(log(0.0001), theta_zeta_B)
mu_E_comp ~ log_normal(log(3.0), theta_E_comp)
mu_A ~ log_normal(log(0.15), theta_A)
mu_k_d ~ log_normal(log(0.2), theta_k_d)
mu_h_c ~ log_normal(log(0.4), theta_h_c)
}

const prop_std = 0.1;
```

```
sub proposal_parameter {
SDF ~ log_normal(log(SDF), 0.35*prop_std)
r_ue ~ log_normal(log(r_ue), 0.5*prop_std)
f_b ~ normal(f_b, 0.0273*prop_std)
h_wc ~ log_normal(log(h_wc), 0.2*prop_std)
omega ~ log_normal(log(omega), 1.0*prop_std)

mu_tau ~ log_normal(log(mu_tau), theta_tau*prop_std)
mu_zeta_A ~ log_normal(log(mu_zeta_A), theta_zeta_A*prop_std)
mu_zeta_B ~ log_normal(log(mu_zeta_B), theta_zeta_B*prop_std)
mu_E_comp ~ log_normal(log(mu_E_comp), theta_E_comp*prop_std)
mu_A ~ log_normal(log(mu_A), theta_A*prop_std)
mu_k_d ~ log_normal(log(mu_k_d), theta_k_d*prop_std)
mu_h_c ~ log_normal(log(mu_h_c), theta_h_c*prop_std)
}

sub initial {
tau ~ log_normal(log(mu_tau), sigma_tau)
zeta_A ~ log_normal(log(mu_zeta_A), sigma_zeta_A)
zeta_B ~ log_normal(log(mu_zeta_B), sigma_zeta_B)
E_comp ~ log_normal(log(mu_E_comp), sigma_E_comp)
A ~ log_normal(log(mu_A), sigma_A)
k_d ~ log_normal(log(mu_k_d), sigma_k_d)
h_c ~ log_normal(log(mu_h_c), sigma_h_c)

z1 ~ log_normal(log(0.3), 0.1)
E_c ~ log_normal(log(0.7), 0.1)
```

```
k_resp ~ log_normal(log(0.3), 0.1)
k_I ~ log_normal(log(0.3), 0.1)
g_SA ~ log_normal(log(0.3), 0.1)
g ~ log_normal(log(0.3), 0.1)
trans ~ log_normal(log(0.3), 0.1)

S_A ~ log_normal(log(246.45), 0.3)
S_B ~ normal(858.19, 108.14)
P_C ~ normal(0.5, 0.1)
}

sub transition(delta = 1.0) {
z1 <- max(0.0, h_wc - h_c)
E_c <- E_wc*exp(-k_d*z1)
k_resp <- E_comp*A*omega*S_A
k_I <- E_c*(1.0 - exp(-A*omega*S_A))
g_SA <- max(0.0, (k_I - k_resp)/r_ue)
g <- min(0.1*S_A, g_SA)
trans <- (f_b*(S_A + S_B) - S_B)*tau

r_tau ~ log_normal(gamma_tau, sigma_tau)
r_zeta_A ~ log_normal(gamma_zeta_A, sigma_zeta_A)
r_zeta_B ~ log_normal(gamma_zeta_B, sigma_zeta_B)
r_E_comp ~ log_normal(gamma_E_comp, sigma_E_comp)
r_A ~ log_normal(gamma_A, sigma_A)
r_k_d ~ log_normal(gamma_k_d, sigma_k_d)
r_h_c ~ log_normal(gamma_h_c, sigma_h_c)
```

```
tau <- tau*(1.0 - 1.0/St) + r_tau/St
zeta_A <- zeta_A*(1.0 - 1.0/St) + r_zeta_A/St
zeta_B <- zeta_B*(1.0 - 1.0/St) + r_zeta_B/St
E_comp <- E_comp*(1.0 - 1.0/St) + r_E_comp/St
A <- A*(1.0 - 1.0/St) + r_A/St
k_d <- k_d*(1.0 - 1.0/St) + r_k_d/St
h_c <- h_c*(1.0 - 1.0/St) + r_h_c/St

ode(h = 0.01, atoler = 1.0e-6, rtoler = 1.0e-9, alg = 'RK4(3)'){
dS_A/dt = g_SA - zeta_A*S_A - trans
dS_B/dt = -zeta_B*S_B + trans

dD/dt   = (zeta_A*S_A + zeta_B*S_B)/(h_sed*phi)

dDIC/dt = -550.0/30.0*12.0/14.0*0.0192*(g_SA*S_A)/h_wc
dO_2/dt = 716.0/30.0*32.0/14.0*0.0192*(g_SA*S_A)/h_wc
}

P_C <- 1.0-exp(-omega*S_A)

}


sub observation {
S_A_obs~log_normal(log(S_A),0.3)
```

```
P_C_obs~log_normal(log(P_C),0.3)
k_d_obs~log_normal(log(k_d),0.5)
h_c_obs~log_normal(log(h_c),0.5)

}
}
```

LibBi 2012 prior sampling file:

```
--target prior
--model-file seagrass.bi
--nsamples 50000
--start-time 734801.0
--end-time 735251.0
--noutputs 450
--input-file data/input_E_PAR_PPB1.nc
--output-file results/prior_1stYear.nc
```

LibBi 2012 posterior sampling file:

```
--target posterior
--model-file seagrass.bi
--input-file data/input_E_PAR_PPB1.nc
--obs-file data/all_obsPPB1_1stYear.nc
--nsamples 50000
--nparticles 1024
--start-time 734801.0
--end-time 735251.0
--noutputs 450
--output-file results/posterior_1stYear.nc
```

```
--with-transform-initial-to-param
```

LibBi prediction sampling file:

```
--target prediction
--model-file seagrass.bi
--input-file data/input_E_PAR_PPB1.nc
--start-time 735251.0
--end-time 735616.0
--noutputs 365
--nsamples 50000
--init-file results/posterior_1stYear.nc
--output-file results/prediction.nc
```

LibBi 2013 prior sampling file:

```
--target prior
--model-file seagrass.bi
--nsamples 50000
--start-time 735251.0
--end-time 735616.0
--noutputs 365
--input-file data/input_E_PAR_PPB1.nc
--output-file results/prior_2ndYear.nc
```

LibBi 2013 posterior sampling file:

```
--target posterior
--model-file seagrass.bi
--input-file data/input_E_PAR_PPB1.nc
--obs-file data/all_obsPPB1_1stYear.nc
```

```
--nsamples 50000

--nparticles 1024

--start-time 735251.0

--end-time 735616.0

--noutputs 365

--output-file results/posterior_2ndyear.nc

--with-transform-initial-to-param
```

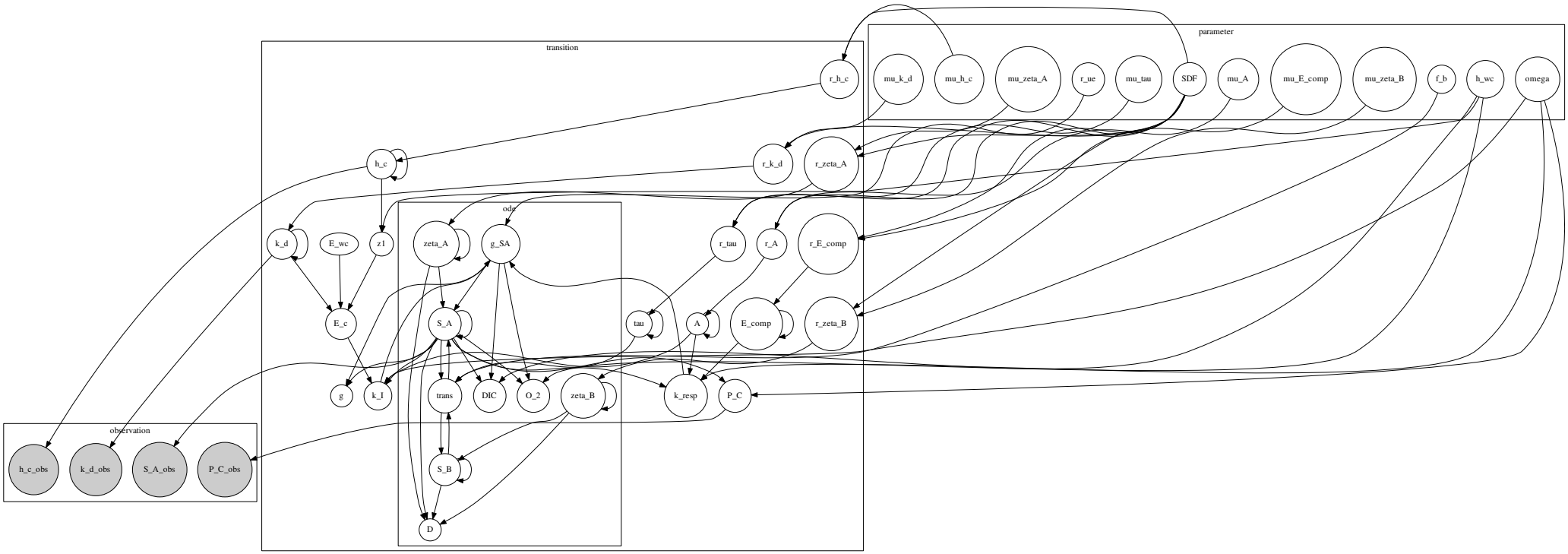

Figure A.1 : Directed acyclic graph (DAG) of LibBi model file seagrass.bi

Table A.1 : Initial condition priors for above-ground biomass, below-ground biomass and proportion cover at each site for each year (2012 and 2012-2013 were assigned the same initial conditions).

Variable	**Site**	**Initial condition 2012**	**Initial condition 2013**
S_A	Altona	Log$\mathcal{N}$(ln(246.45), 0.3)	Log$\mathcal{N}$(ln(110), 0.3)
	Blairgowrie	Log$\mathcal{N}$(ln(377.69), 0.5)	Log$\mathcal{N}$(ln(200), 0.5)
	Point Henry	Log$\mathcal{N}$(ln(291.11), 0.2)	Log$\mathcal{N}$(ln(200), 0.2)
	Point Richards	Log$\mathcal{N}$(ln(275.06), 0.2)	Log$\mathcal{N}$(ln(120), 0.2)
	Swan Bay South	Log$\mathcal{N}$(ln(290.29), 0.3)	Log$\mathcal{N}$(ln(110), 0.2)
S_B	Altona	$\mathcal{N}$(858.19, 108.14)	$\mathcal{N}$(400.0, 60.0)
	Blairgowrie	$\mathcal{N}$(1121.81, 159.69)	$\mathcal{N}$(600.0, 150.0)
	Point Henry	$\mathcal{N}$(1244.58, 142.65)	$\mathcal{N}$(1000.0, 140.0)
	Point Richards	$\mathcal{N}$(985.56, 123.27)	$\mathcal{N}$(450.0, 50.0)
	Swan Bay South	$\mathcal{N}$(840.56, 82.59)	$\mathcal{N}$(400.0, 50.0)
P_c	Altona	Trun$\mathcal{N}$(0.75, 0.3, 0.0, 1.0)	Trun$\mathcal{N}$(0.5, 0.3, 0.0, 1.0)
	Blairgowrie	Trun$\mathcal{N}$(0.85, 0.3, 0.0, 1.0)	Trun$\mathcal{N}$(0.6, 0.3, 0.0, 1.0)
	Point Henry	Trun$\mathcal{N}$(0.7, 0.3, 0.0, 1.0)	Trun$\mathcal{N}$(0.5, 0.3, 0.0, 1.0)
	Point Richards	Trun$\mathcal{N}$(0.8, 0.3, 0.0, 1.0)	Trun$\mathcal{N}$(0.5, 0.3, 0.0, 1.0)
	Swan Bay South	Trun$\mathcal{N}$(0.8, 0.3, 0.0, 1.0)	Trun$\mathcal{N}$(0.6, 0.3, 0.0, 1.0)

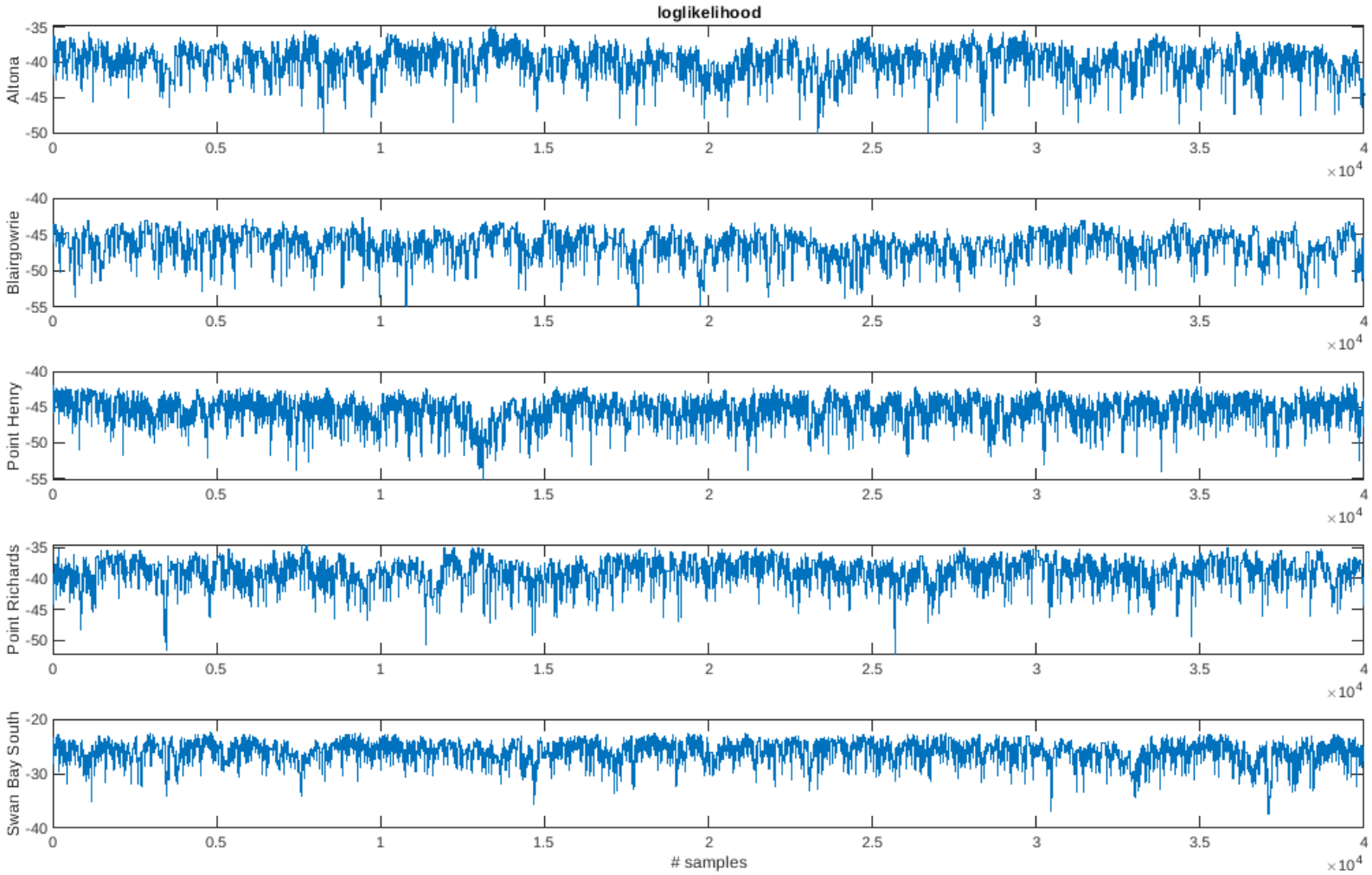

Figure A.2 : 2012 log-likelihoods across samples for each site after initial 10,000 samples were discarded as burn-in.

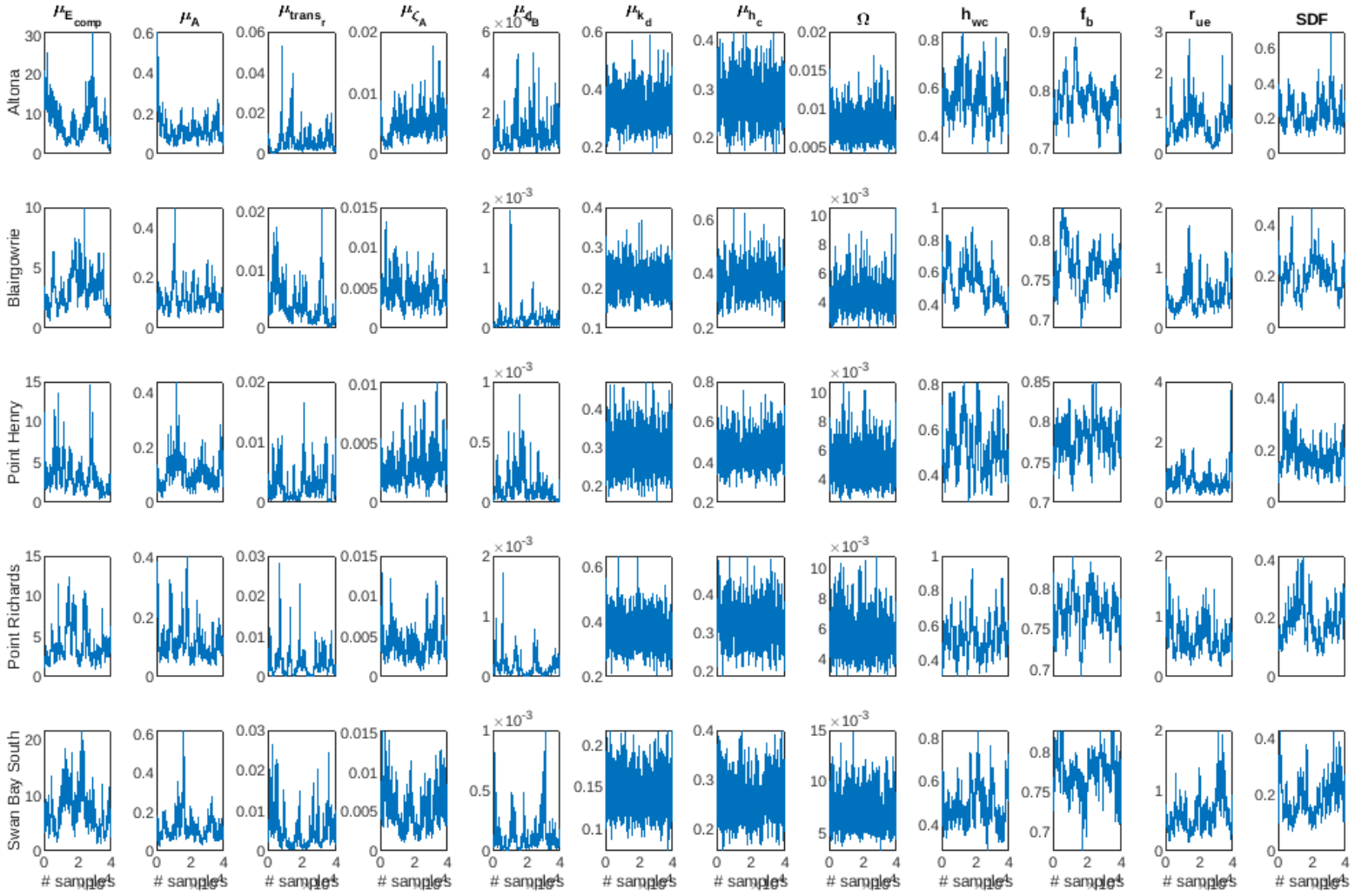

Figure A.3 : 2012 parameter traces across samples for each site after burn-in was discarded.

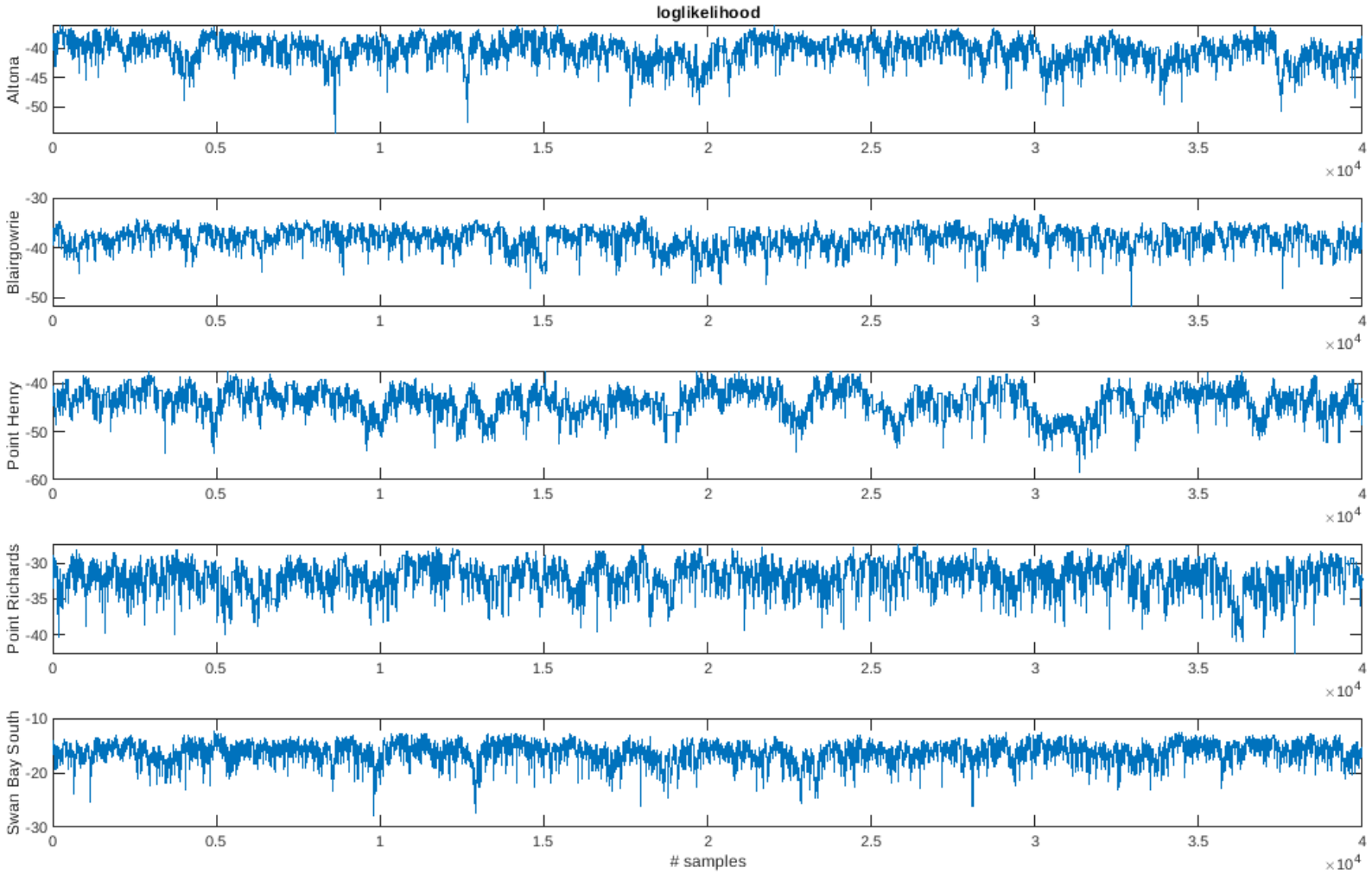

Figure A.4 : 2013 log-likelihoods across samples for each site after initial 10,000 samples were discarded as burn-in.

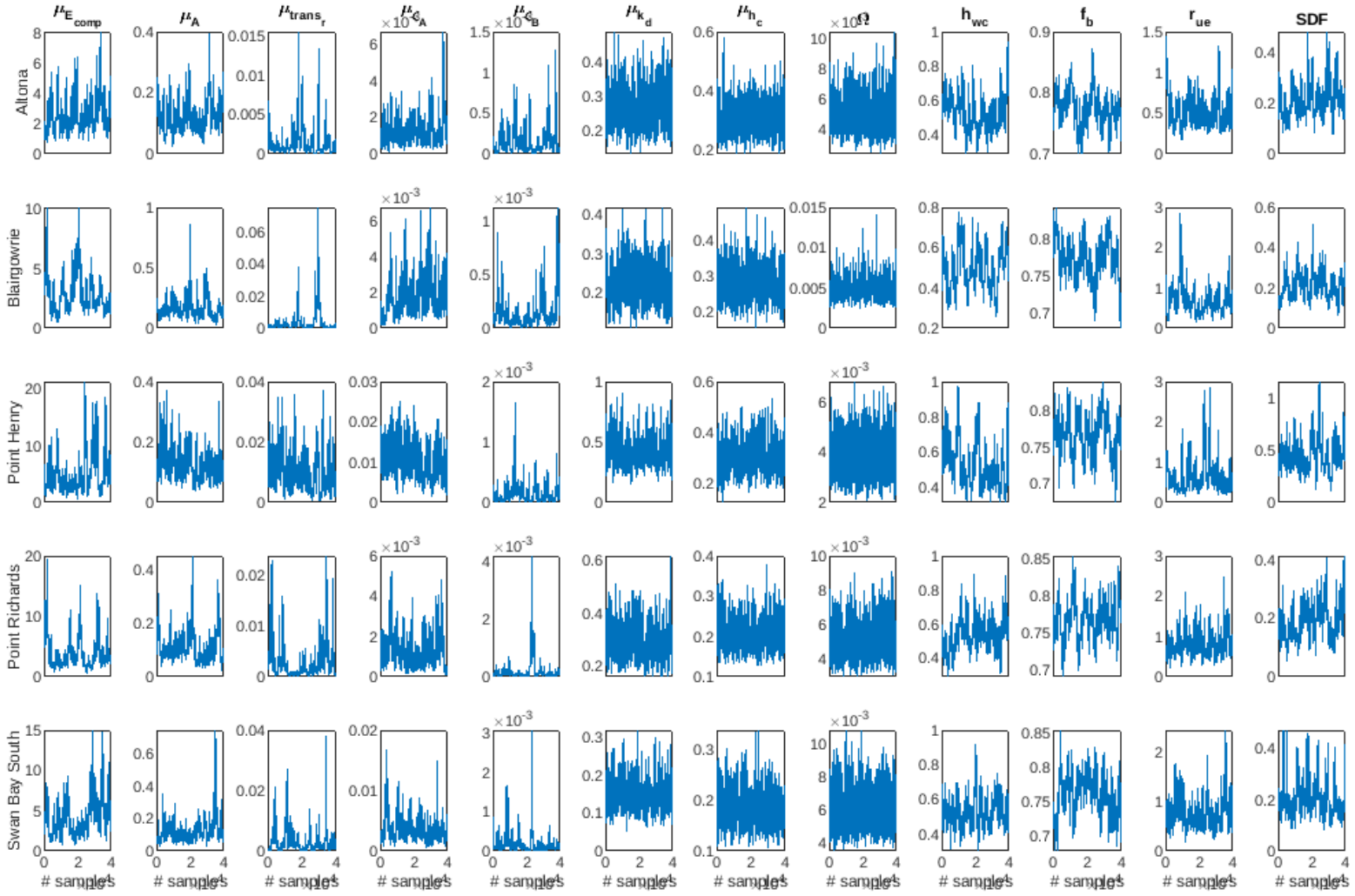

Figure A.5 : 2013 parameter traces across samples for each site after burn-in was discarded.

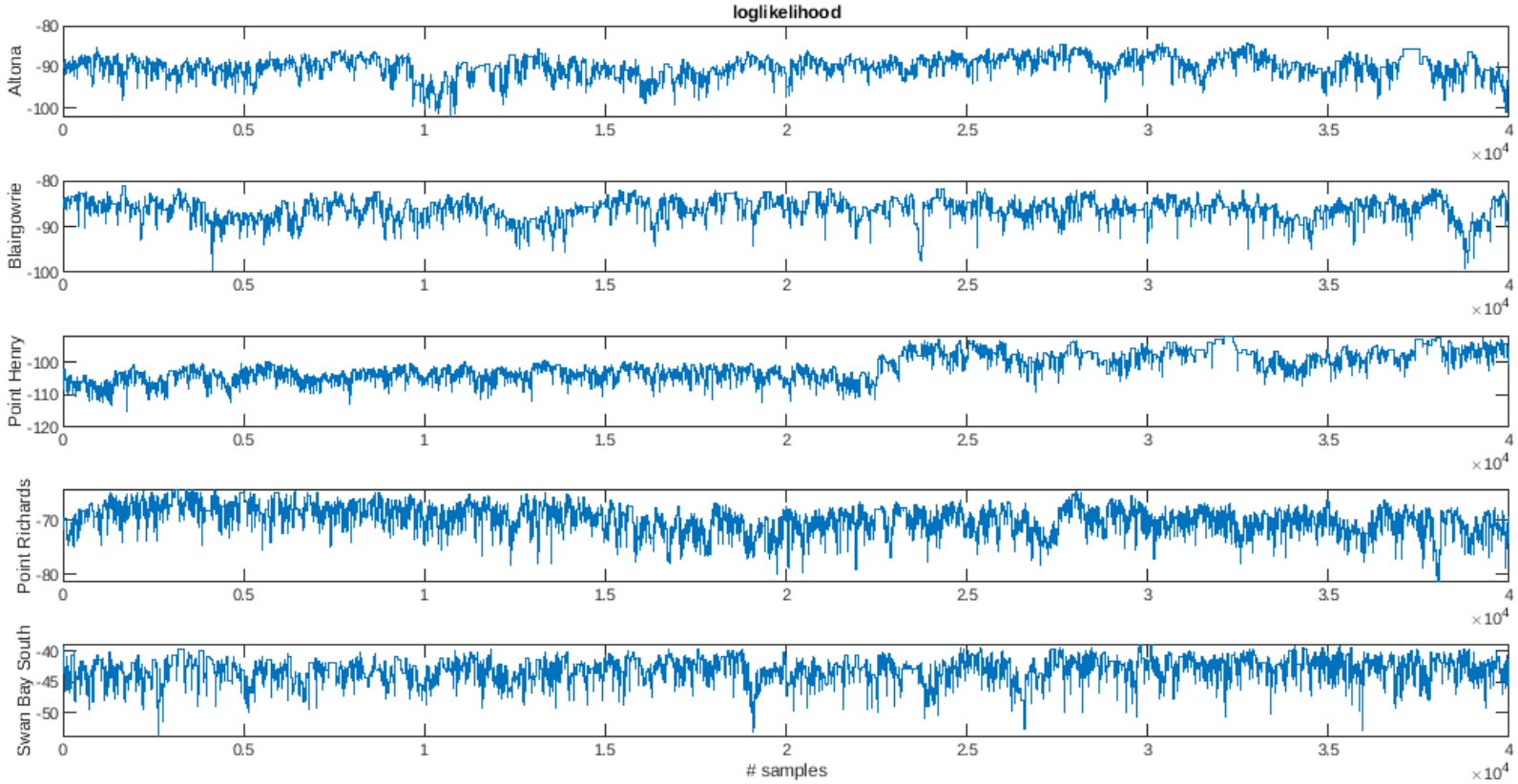

Figure A.6 : 2012-2013 log-likelihoods across samples for each site after burn-in was discarded.

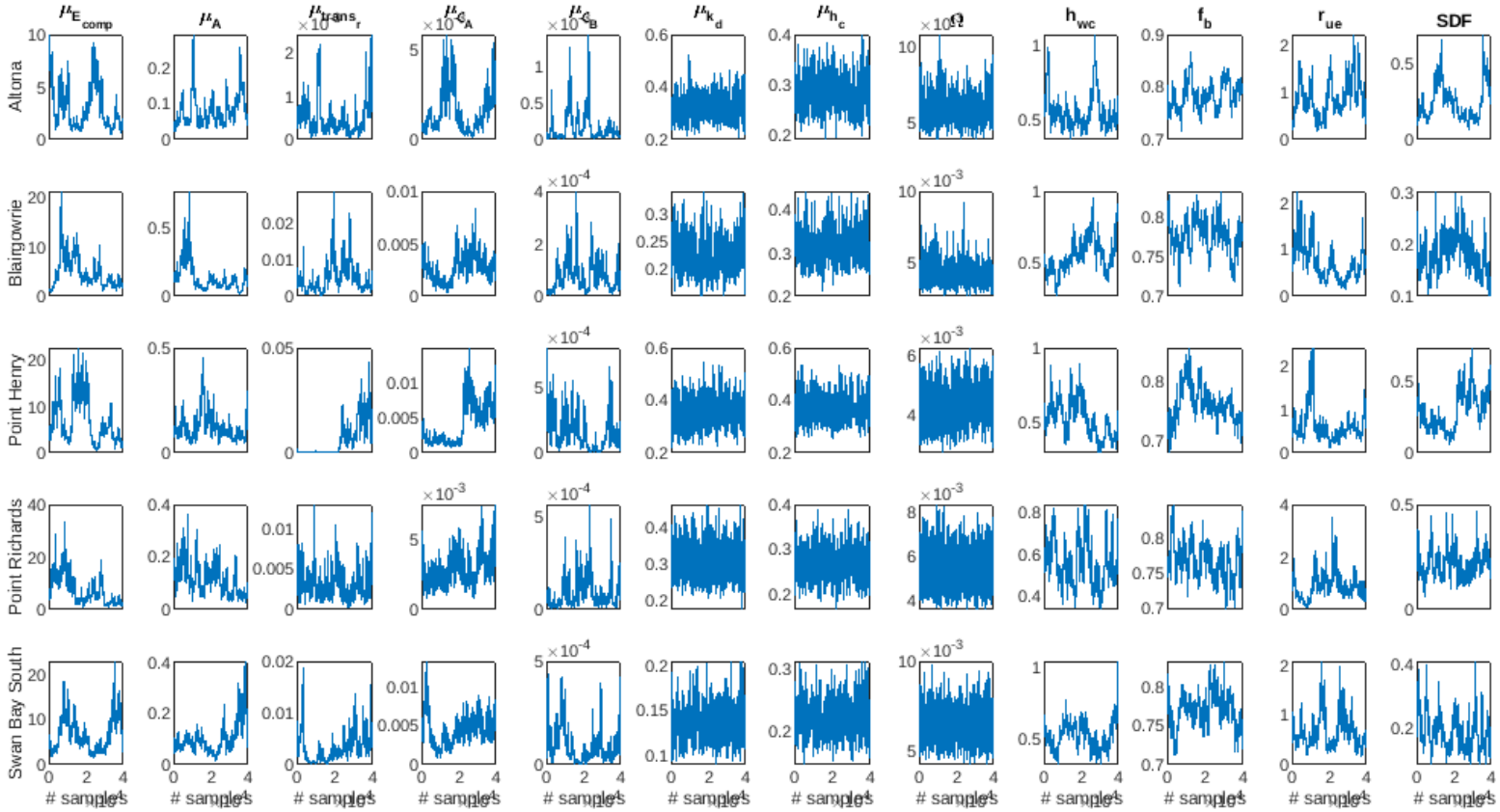

Figure A.7 : 2012-2013 parameter traces across samples for each site after burn-in was discarded.

Appendix B

LibBi model code and MCMC diagnostics for modelling detrital decay using a time-varying decay rate

LibBi laboratory study model file: decay.bi

```
model decay {


state W, k, K

noise r

param sigma_r

obs W_obs


sub parameter {

sigma_r ~ normal(0.1, 0.01)

}


const prop_std = 0.5;

sub proposal_parameter {

sigma_r ~ normal(sigma_r, 0.01*prop_std)

}


sub initial {

k ~ normal(log(0.003), 0.8)

K ~ log_normal(log(0.003), 0.8)
```

```
W ~ normal(1.0, 0.1)
}

sub transition(delta = 1.0) {
r ~ normal(0.0, sigma_r)
k <- k + r
K <- exp(k)

ode(h = 0.01, atoler = 1.0e-6, rtoler = 1.0e-9, alg = 'RK4(3)'){
dW/dt = -K*W
}
}

sub observation {
W_obs~log_normal(log(W),0.1)
}

}
```

LibBi laboratory study prior sampling file: prior.conf

```
--target prior
--model-file decay.bi
--nsamples 50000
--start-time 0.0
--end-time 84.0
--noutputs 84
--output-file results/prior.nc
```

LibBi laboratory study posterior sampling file: posterior.conf

```
--target posterior

--model-file decay.bi

--obs-file data/obs_no_23_lW.nc

--nsamples 50000

--nparticles 1024

--start-time 0.0

--end-time 84.0

--noutputs 84

--output-file results/posterior.nc

--with-transform-initial-to-param
```

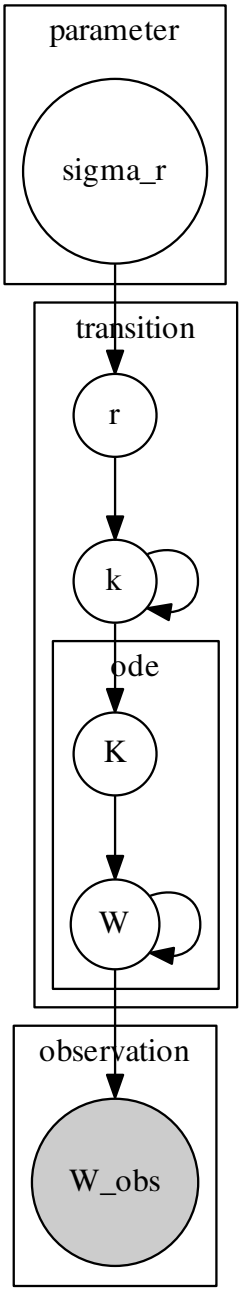

Figure B.1 : Directed acyclic graph of the LibBi model file decay.bi

For the laboratory study, the parameter priors that remained constant across all tissue types, treatments, and temperatures were the initial conditions of proportion of weight remaining ($W \sim \mathcal{N}(1,\, 0.1)$). Those that varied were parameter priors for σ_r, and initial condition priors for decay rate k (Table B.1). The PMMH proposal distribution was set to 50% jumps of the original distribution standard deviation $\sigma_r \sim \mathcal{N}(\sigma_r,\, 0.5{\times}0.01)$ for leaf tissue and $\sigma_r \sim \mathcal{N}(\sigma_r,\, 0.5{\times}0.1)$ for root/rhizome tissue.

For the field study, the parameter priors that remained constant across all tissue types and sites were the stochasticity parameter ($\sigma_r \sim \mathcal{N}(0.1,\, 0.01)$) and initial conditions of decay rate ($k \sim \text{Log}\mathcal{N}(\ln(0.001),\, 0.3)$). Those that varied were initial condition priors for the proportion of weight remaining W (Table B.2). The PMMH proposal distribution was set to 20% jumps of the original distribution standard deviation $\sigma_r \sim \mathcal{N}(\sigma_r,\, 0.2{\times}0.01)$.

Table B.1 : Parameter priors for the stochasticity parameter (σ_r), initial condition priors for decay rate (k), and PMMH acceptance rates of each tissue type, treatment and temperature combination of the laboratory study.

Tissue	**Treatment**	**Temp.**	$\boldsymbol{\sigma_r}$ **prior**	$\boldsymbol{k}$ **initial condition**	**Accpt. rate**
Leaf	Control	23°C	$\mathcal{N}(0.1, 0.01)$	$\text{Log}\mathcal{N}(\ln(0.003), 0.8)$	0.3436
		30°C	$\mathcal{N}(0.1, 0.01)$	$\text{Log}\mathcal{N}(\ln(0.003), 0.8)$	0.2829
	+Nutrient	23°C	$\mathcal{N}(0.1, 0.01)$	$\text{Log}\mathcal{N}(\ln(0.003), 0.8)$	0.3695
		30°C	$\mathcal{N}(0.1, 0.01)$	$\text{Log}\mathcal{N}(\ln(0.003), 0.8)$	0.2733
Root/	Control	23°C	$\mathcal{N}(0.7, 0.1)$	$\text{Log}\mathcal{N}(\ln(0.1), 0.8)$	0.3645
rhizomes		30°C	$\mathcal{N}(0.7, 0.1)$	$\text{Log}\mathcal{N}(\ln(0.1), 0.8)$	0.3674
	+Nutrient	23°C	$\mathcal{N}(0.7, 0.1)$	$\text{Log}\mathcal{N}(\ln(0.1), 0.8)$	0.3945
		30°C	$\mathcal{N}(0.7, 0.1)$	$\text{Log}\mathcal{N}(\ln(0.1), 0.8)$	0.3780

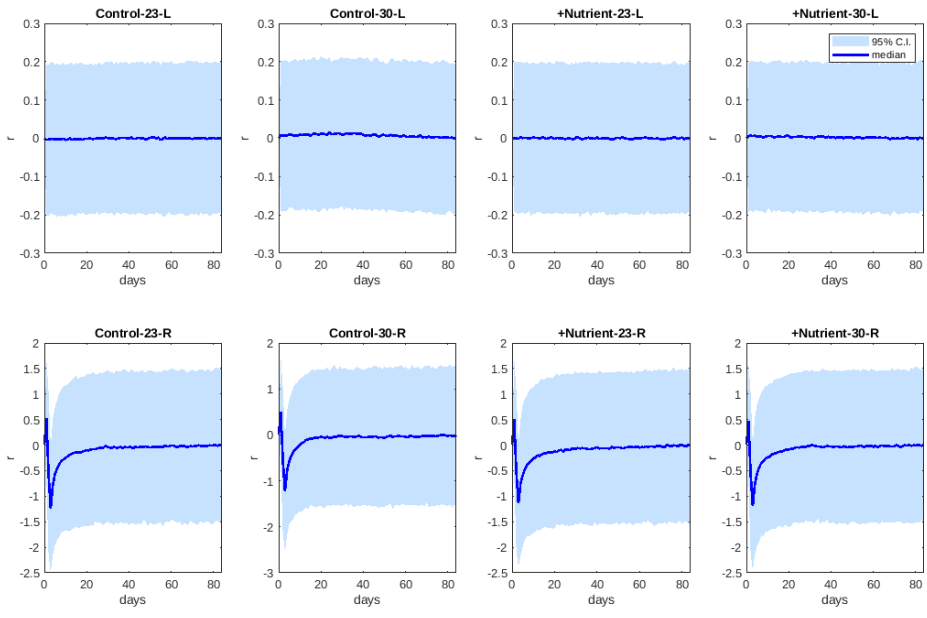

Figure B.2 : Laboratory study random walk (r) posterior through time for each temperature, nutrient and tissue combination.

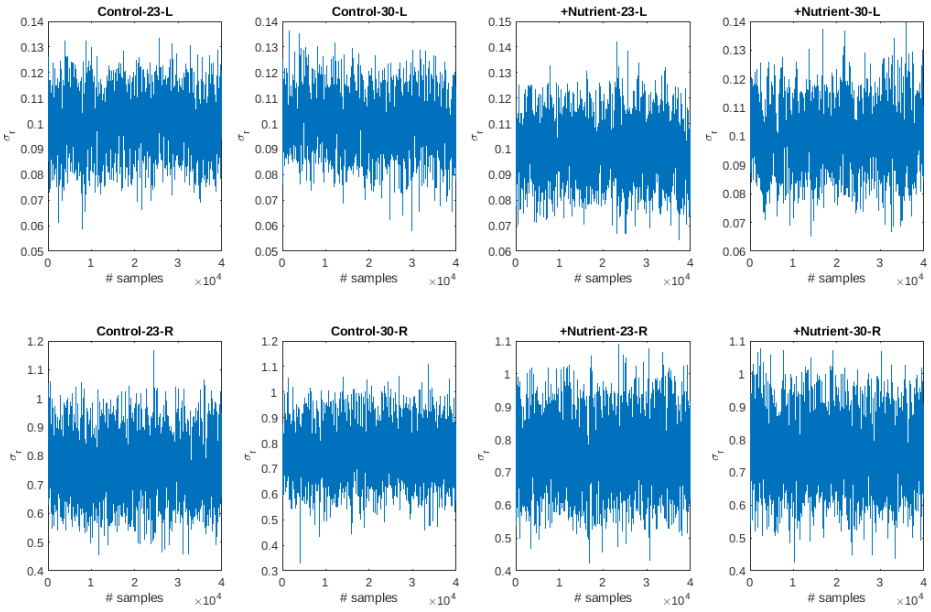

Figure B.3 : Laboratory study stochasticity parameter (σ_r) trace across samples, for each treatment group after burn-in was discarded.

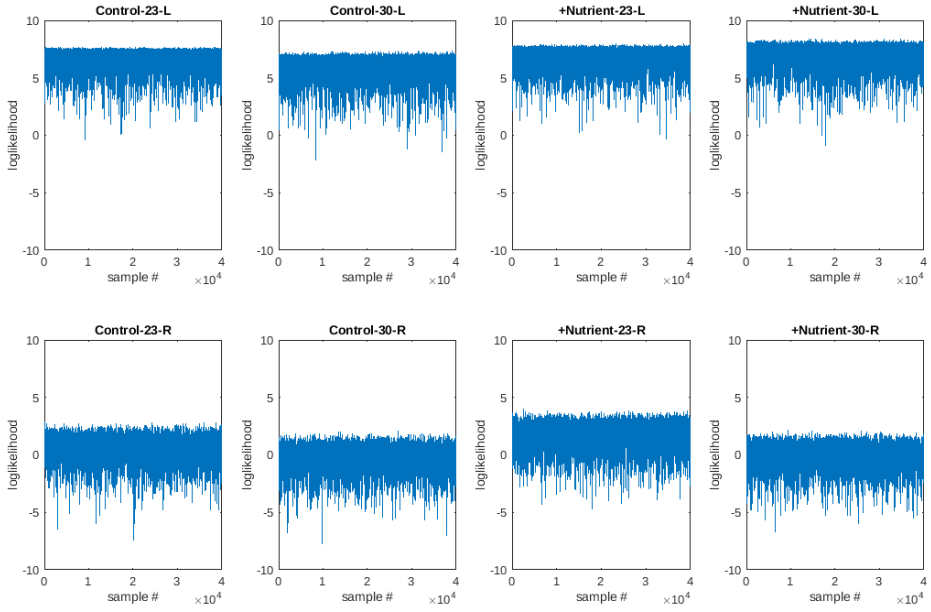

Figure B.4 : Laboratory study log-likelihoods across samples, for each treatment group after burn-in was discarded.

Table B.2 : Field study initial condition priors for proportion of weight remaining (W), and PMMH acceptance rates of each tissue type and site combination.

Tissue	**Site**	***W* initial condition**	**Acceptance rate**
Leaf	Pristine	$\mathcal{N}(1.1, 0.1)$	0.5883
	Impacted	$\mathcal{N}(1, 0.1)$	0.4233
	Bare	$\mathcal{N}(1.1, 0.1)$	0.4738
Root/ rhizomes	Pristine	$\mathcal{N}(1.1, 0.1)$	0.5227
	Impacted	$\mathcal{N}(0.9, 0.1)$	0.4183
	Bare	$\mathcal{N}(0.9, 0.1)$	0.4517

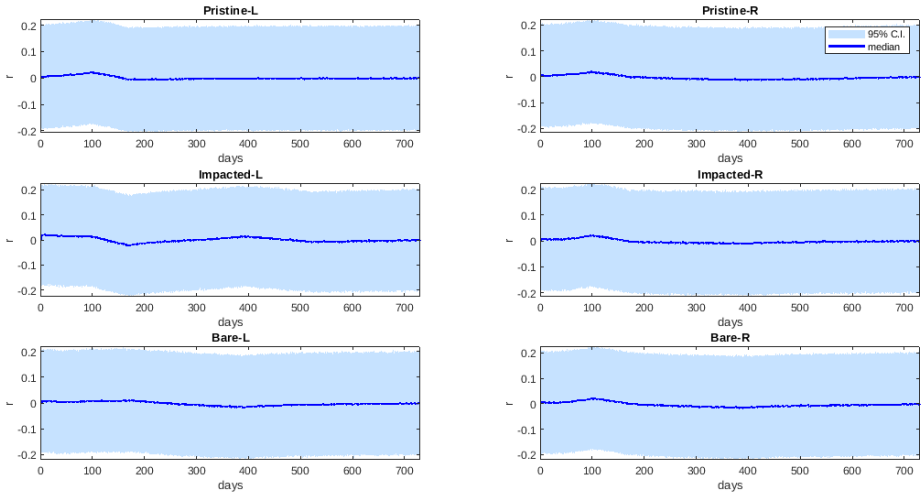

Figure B.5 : Field study random walk (r) posterior through time for each site and tissue type (left: leaf, right: root/rhizome).

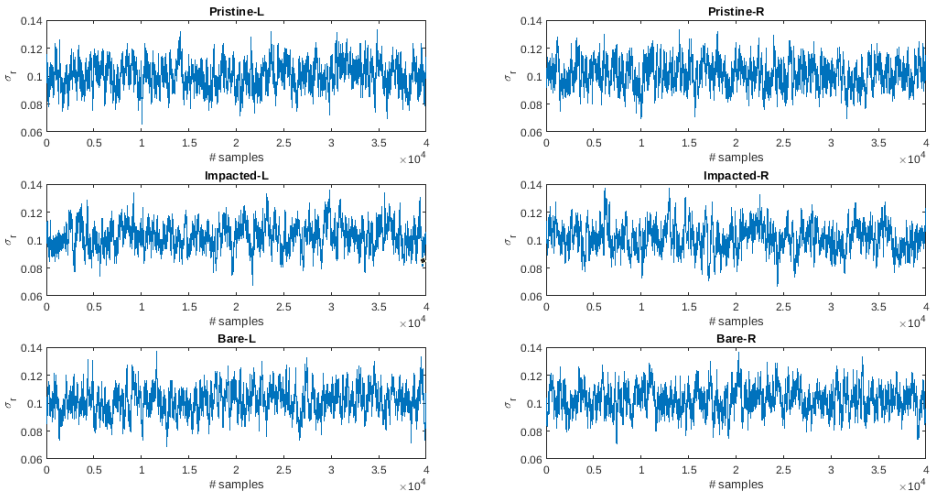

Figure B.6 : Field study stochasticity parameter (σ_r) trace across samples for each site and tissue type (left: leaf, right: root/rhizome) after burn-in was discarded.

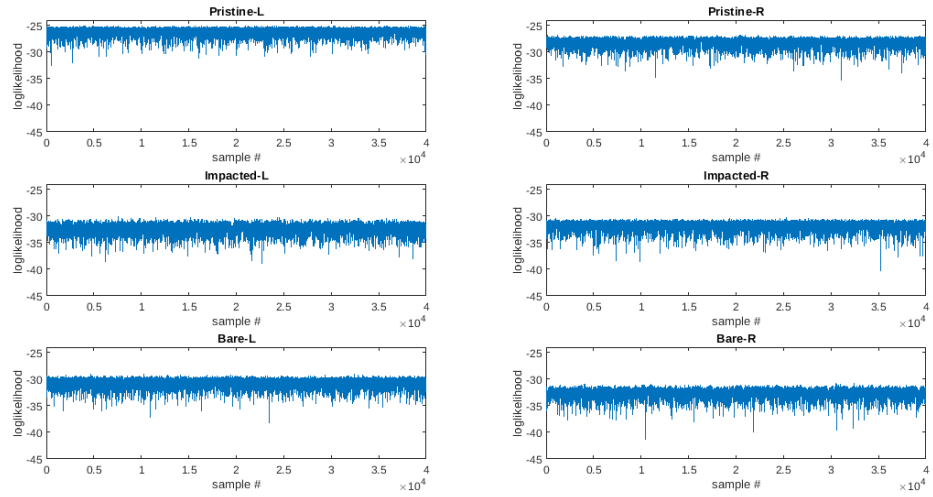

Figure B.7 : Field study log-likelihoods across samples for each site and tissue type (left: leaf, right: root/rhizome) after burn-in was discarded.

Appendix C

LibBi model code and MCMC diagnostics for microalgae growth within a photo-bioreactor

LibBi model file: micro_iterative.bi

```
model micro_iterative {


const FO2       = 0.2094

const FCO2      = 397e-6

const S         = 34.0

const V         = 500.0

const DIC_M     = 1724.20

const O_2_M     = 226.65

const alk_M     = 1797.90

const tau       = 6.0

const kLAO2_m   = log(2.0)*24.0*60.0/tau


param kLAO2

param Km

param RR

param RQ_d

param RQ_n

param sigma_O_2

param sigma_pH

param sigma_DIC
```

```
param offset_O_2


input I                 // light intensity
input T                 // temperature (C)
input gas               // gas on/off
input dil               // dilution rate


state DIC // state variables
state O_2
state pH
state Cp
state mich_ment
state O2H_pr
state CO2H_pr
state R
state R1
state P
state P1
state alk
state CO2
state HCO3
state CO3
state O_2H
state CO2H
state h_3
state h_free_3
```

```
noise r_R
noise r_P

/* random walk parameters */
param sigma_r_R
param sigma_r_P

obs O2_obs
obs pH_obs
obs DIC_obs
obs alk_obs

sub parameter {/* prior distribution over parameters */
Km      ~ log_normal(log(100.0), 0.5)
kLAO2   ~ log_normal(log(kLAO2_m), 0.3)
RR      ~ uniform(0.0001, 0.2)
RQ_d    ~ uniform(0.66, 1.0)
RQ_n    ~ uniform(0.66, 1.0)

sigma_O_2 ~ log_normal(log(0.03), 0.5)
sigma_pH  ~ log_normal(log(0.03), 0.5)
sigma_DIC ~ log_normal(log(0.03), 0.5)

offset_O_2 ~ normal(0, 2.0)

sigma_r_R       ~ normal(0.01, 0.001)
sigma_r_P       ~ normal(0.05, 0.01)
```

```
}

const prop_std = 0.1;
sub proposal_parameter {
Km      ~ log_normal(log(Km), 0.5*prop_std)
kLAO2   ~ log_normal(log(kLAO2), 0.3*prop_std)
RR      ~ truncated_normal(RR, 0.2*prop_std, lower = 0.0001, upper =
   0.2)
RQ_d    ~ truncated_normal(RQ_d, 0.2*prop_std, lower = 0.66, upper =
   1.0)
RQ_n    ~ truncated_normal(RQ_n, 0.2*prop_std, lower = 0.66, upper =
   1.0)


sigma_O_2       ~ log_normal(log(sigma_O_2), 0.5*prop_std)
sigma_pH        ~ log_normal(log(sigma_pH), 0.5*prop_std)
sigma_DIC       ~ log_normal(log(sigma_DIC), 0.5*prop_std)

offset_O_2      ~ normal(offset_O_2, 2.0*prop_std)

sigma_r_R       ~ normal(sigma_r_R, 0.001*prop_std)
sigma_r_P       ~ normal(sigma_r_P, 0.01*prop_std)
}

sub initial {/* prior distribution over initial conditions, given
   parameters */
R       ~ normal(log(20.0), 0.4)
```

```
R1      ~ log_normal(log(20.0), 0.4)
P       ~ normal(log(200.0), 0.4)
P1      ~ log_normal(log(200.0), 0.4)

Cp      ~ log_normal(log(300.0), 0.2)
alk     ~ log_normal(log(1750.0), 0.1)
DIC     ~ log_normal(log(1300.0), 0.2)
O_2     ~ log_normal(log(225.0), 0.2)
pH      ~ log_normal(log(8.5), 0.2)
CO2     ~ log_normal(log(3.0), 0.4)
HCO3    ~ log_normal(log(1000.0), 0.3)
CO3     ~ log_normal(log(300.0), 0.4)
O_2H    ~ log_normal(log(200.0), 0.2)
CO2H    ~ log_normal(log(10.0), 0.2)
}


sub transition(delta = 0.0021) {

/* processes */

inline TK     = T + 273.15              // temp in kelvin
inline K0_CO2 = exp(-60.2409 + 93.4517*(100.0/TK) + 23.3585*log(TK
   /100.0)+ S*(0.023517 - 0.023656*(TK/100) + 0.0047036*(TK/100.0)*(
   TK/100.0)))
CO2H          <- K0_CO2*FCO2*1.0220*1e6
```

```
inline K0_O2  =  (exp(-1282.8704 + 36619.96/TK + 223.1396*log(TK)
   -0.354707*TK + S*(5.957e-3 -3.7353/TK) + 3.68e-6*S*S))/(0.2094e
   -06)
O_2H          <- K0_O2*FO2*1.0220*1e-6

inline PAC    = HCO3
inline mm     = PAC/(Km + PAC)

// CO2SYS iterative solution
// set up all the constants

inline logTK  = log(TK)
inline S2     = S*S
inline sqrtS  = sqrt(S)

// total sulphur

inline TS     = (0.14/96.062)*(S/1.80655)
inline IS     = 19.924*S/(1000.0 - 1.005*S)

inline KS_int = -4276.1/TK + 141.328 - 23.093*logTK + (-13856.0/TK +
   324.57  -  47.986*logTK)*sqrt(IS) + ( 35474.0/TK - 771.54  +
   114.723*logTK)*IS - 2698.0/TK*IS**1.5 + 1776.0/TK*IS**2
inline KS     = exp(KS_int)*(1 - 0.001005*S)

// Fluorine
```

```
inline TF       = 0.000067*S/18.9984/1.80655
inline KF       = exp(-(-874.0/TK - 0.111*sqrtS + 9.68))
inline SWS_2_T  = (1.0 + TS/KS)/(1.0 + TS/KS + TF/KF)
inline Free_2_T = 1.0 + TS/KS

// H2O dissoc

inline KW = exp(148.9802 - 13847.26/TK  - 23.6521*logTK + (118.67/TK
   - 5.977 + 1.0495*logTK)*sqrtS - 0.01615*S)

// Boron

inline KB = exp((-8966.90 - 2890.53*sqrtS - 77.942*S + 1.728*S*sqrtS
   - 0.0996*S2)/TK + 148.0248 + 137.1942*sqrtS + 1.62142*S -
   (24.4344 + 25.085*sqrtS + 0.2474*S)*logTK + 0.053105*sqrtS*TK)
inline TB = 0.0004326*S/35.0

// Carbon eq constants

inline K1 = 10**(-(3633.86/TK - 61.2172 + 9.6777 *logTK - 0.011555*S
   + 0.0001152*S**2))*1.23    //1.23 experiment specific and
   measured
inline K2 = 10**(-( 471.8/TK + 25.9290 - 3.16967*logTK - 0.01781*S +
   0.0001122*S**2))*0.53      //0.53 experiment specific and
   measured

// end all the constants
```

```
// intial guess at the pH (using the approximating equation)

inline pH_init = 12.26 -0.0030605*DIC -0.043752*T -0.013625*S+
   0.00011315*alk + 1.3463e-05*DIC*T + 5.2215e-07*DIC*alk

// iteration 1

inline h_1      = 10.0**(-pH_init)
inline h_free_1 = h_1/Free_2_T
inline f0_1     = (DIC*1e-6*(K1*h_1 + 2.0*K1*K2)/(h_1*h_1 + K1*h_1 +
   K1*K2) - h_free_1 + KW/h_1 - alk*1e-6 + TB/(1.0 + h_1/KB))*1e6
inline df0_1    = (DIC*1e-6*(K1 + 2.0*K1*K2)/(h_1**2.0 + K1*h_1 + K1*
   K2) - DIC*1e-6*(K1*h_1 + 2.0*K1*K2)/(h_1**2.0 + K1*h_1 + K1*K2)
   **2.0*(2.0*h_1 + K1) - TB*1.0/(1.0 + h_1/KB)**2.0/KB - KW/h_1
   **2.0 - 1.0/Free_2_T)*1e6*(-log(10.0)*10.0**(-pH_init))
inline pH_1     = pH_init - f0_1/df0_1

// iteration 2

inline h_2      = 10.0**(-pH_1)
inline h_free_2 = h_2/Free_2_T
inline f0_2     = (DIC*1e-6*(K1*h_2 + 2.0*K1*K2)/(h_2*h_2 + K1*h_2 +
   K1*K2) - h_free_2 + KW/h_2 - alk*1e-6 + TB/(1.0 + h_2/KB))*1e6
inline df0_2    = (DIC*1e-6*(K1 + 2.0*K1*K2)/(h_2**2.0 + K1*h_2 + K1*
   K2) - DIC*1e-6*(K1*h_2 + 2.0*K1*K2)/(h_2**2.0 + K1*h_2 + K1*K2)
   **2.0*(2.0*h_2 + K1) - TB*1.0/(1.0 + h_2/KB)**2.0/KB - KW/h_2
```

```
   **2.0 - 1.0/Free_2_T)*1e6*(-log(10.0)*10.0**(-pH_1))
inline pH_2     = pH_1 - f0_2/df0_2

// iteration 3

h_3             <- 10.0**(-pH_2)
h_free_3        <- h_3/Free_2_T
inline f0_3     = (DIC*1e-6*(K1*h_3 + 2.0*K1*K2)/(h_3*h_3 + K1*h_3 +
   K1*K2) - h_free_3 + KW/h_3 - alk*1e-6 + TB/(1.0 + h_3/KB))*1e6
inline df0_3    = (DIC*1e-6*(K1 + 2.0*K1*K2)/(h_3**2.0 + K1*h_3 + K1*
   K2) - DIC*1e-6*(K1*h_3 + 2.0*K1*K2)/(h_3**2.0 + K1*h_3 + K1*K2)
   **2.0*(2.0*h_3 + K1) - TB*1.0/(1.0 + h_3/KB)**2.0/KB - KW/h_3
   **2.0 - 1.0/Free_2_T)*1e6*(-log(10.0)*10.0**(-pH_2))
pH              <- pH_2 - f0_3/df0_3


// calculate the final concentrations

inline H                = 10.0**(-pH)
inline H2               = H*H
inline denom            = (H2 + K1*H + K1*K2)
CO2                     <- DIC*H2/denom
HCO3                    <- DIC*H*K1/denom
CO3                     <- DIC*K1*K2/denom

// end CO2SYS iterative solution
```

```
/* R and P as random walks */

r_R     ~ normal(0.0, sigma_r_R)
R       <- R + r_R
R1      <- exp(R)

r_P     ~ normal(0.0, sigma_r_P)
P       <- P + r_P
P1      <- exp(P)

ode(h = 0.1, atoler = 1.0e-6, rtoler = 1.0e-6, alg = 'RK4(3)'){
dDIC/dt  =  -P1*24.0*I*mm + R1*24.0 + gas*0.893*kLAO2*(CO2H - CO2) +
   dil/V*(DIC_M - DIC)
dO_2/dt  =  (P1*24.0*I*mm - R1*24.0)/(RQ_d*I + RQ_n*(1.0-I)) + gas*
   kLAO2*(O_2H - O_2) + dil/V*(O_2_M - O_2)     + offset_O_2

dalk/dt  =   RR*P1*24.0*I*mm + dil/V*(alk_M - alk)
dCp/dt   =   (P1*24.0*I*mm - R1*24.0) + dil/V*(Cp)

}

mich_ment <- mm
O2H_pr    <- O_2H
CO2H_pr   <- CO2H

}
```

```
sub observation {

O2_obs  ~ log_normal(log(O_2), sigma_O_2)
pH_obs  ~ log_normal(log(pH),  sigma_pH)
DIC_obs ~ log_normal(log(DIC), sigma_DIC)
alk_obs ~ log_normal(log(alk), sigma_DIC)
}
}
```

LibBi prior sampling file: prior.conf

```
--target prior
--model-file micro_iterative.bi
--nsamples 500
--start-time 0.61304
--end-time 4.7866
--noutputs 6049
--input-file data/input_all_2018_normalised.nc
--output-file results/prior_micro_iterative.nc
```

LibBi posterior sampling file: posterior.conf

```
--target posterior
--model-file micro_iterative.bi
--input-file data/input_all_2018_normalised.nc
--obs-file data/obs_all_2018.nc
--nsamples 500
--nparticles 1024
```

```
--start-time 0.61304

--end-time 4.7866

--noutputs 6049

--output-file results/posterior_micro_iterative.nc

--with-transform-initial-to-param
```

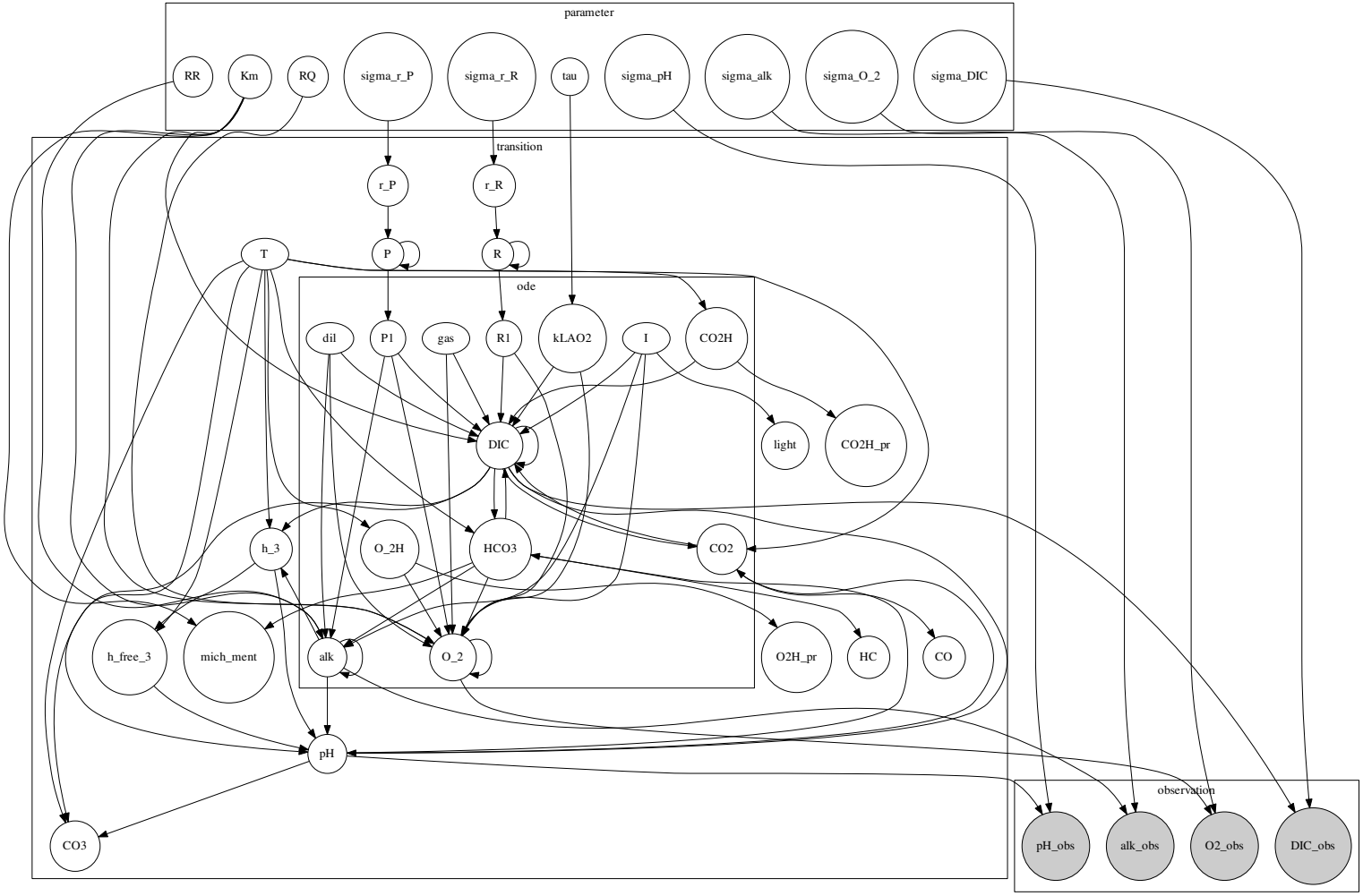

Figure C.1 : Directed Acyclic Graph of the LibBi model file micro_iterative.bi

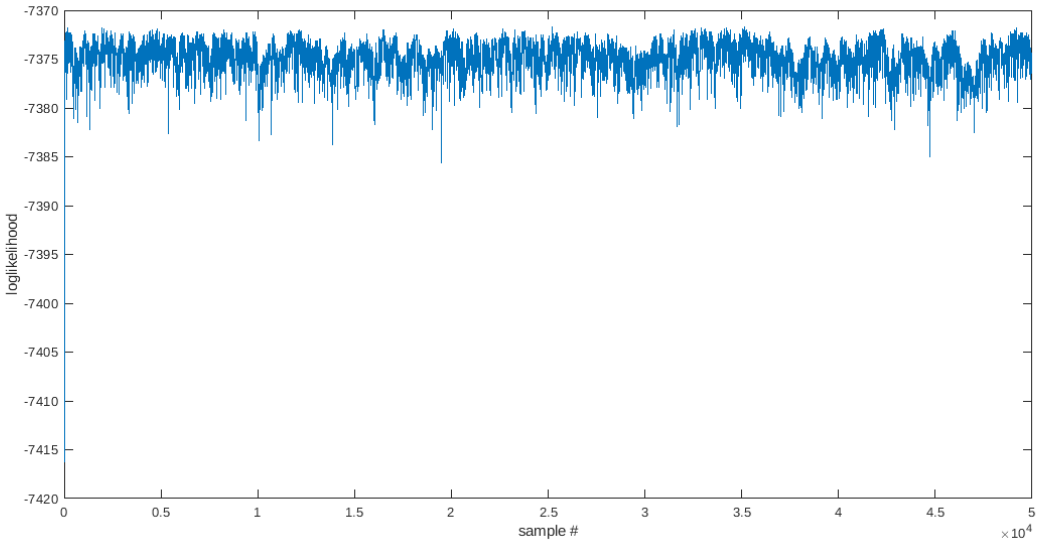

Figure C.2 : Log-likelihood for the twin experiment where photosynthesis and respiration were constant through time.

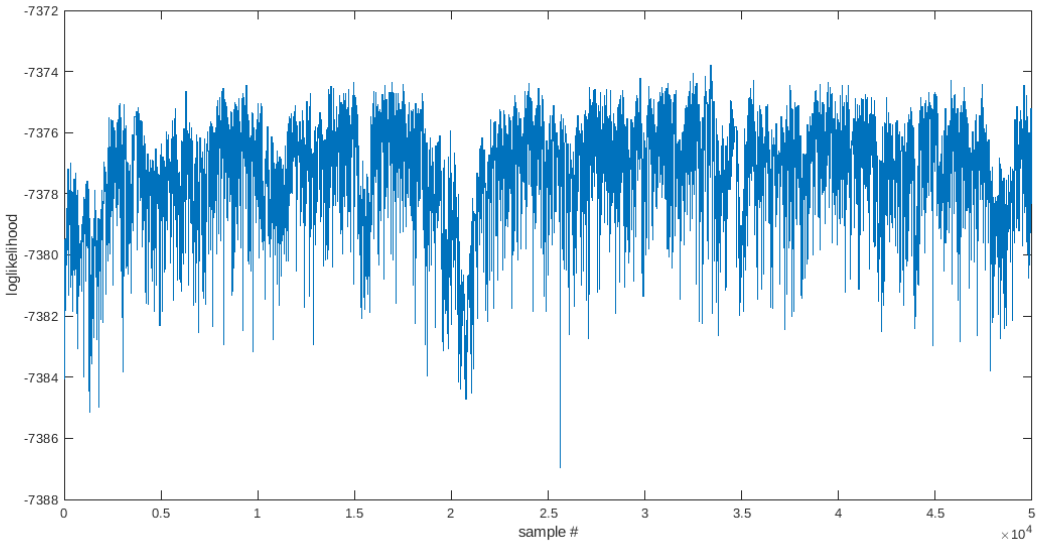

Figure C.3 : Log-likelihood for the twin experiment where photosynthesis and respiration were changing through time.

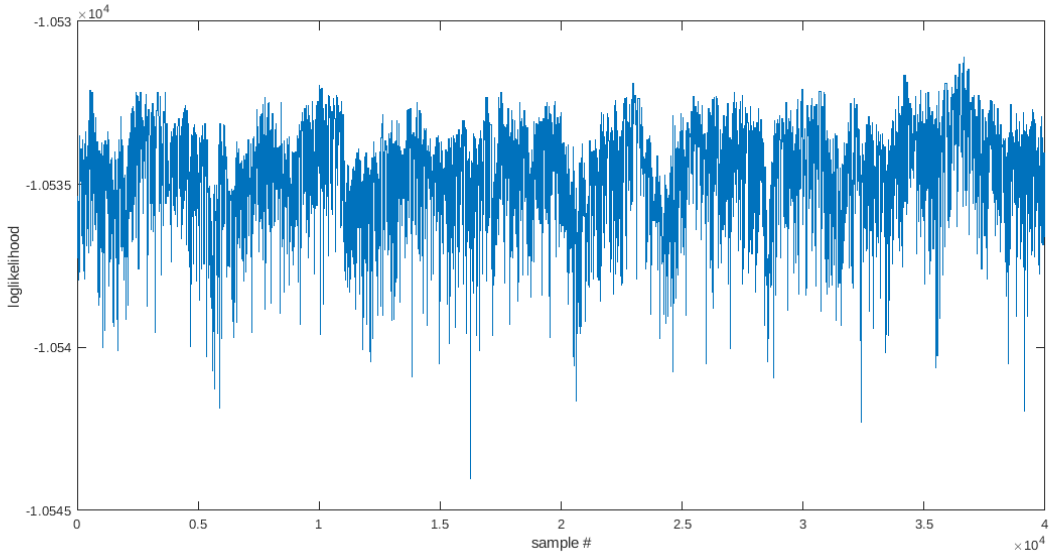

Figure C.4 : Log-likelihood for the model with experimental data when photosynthesis, respiration and respiratory quotients were changing through time.

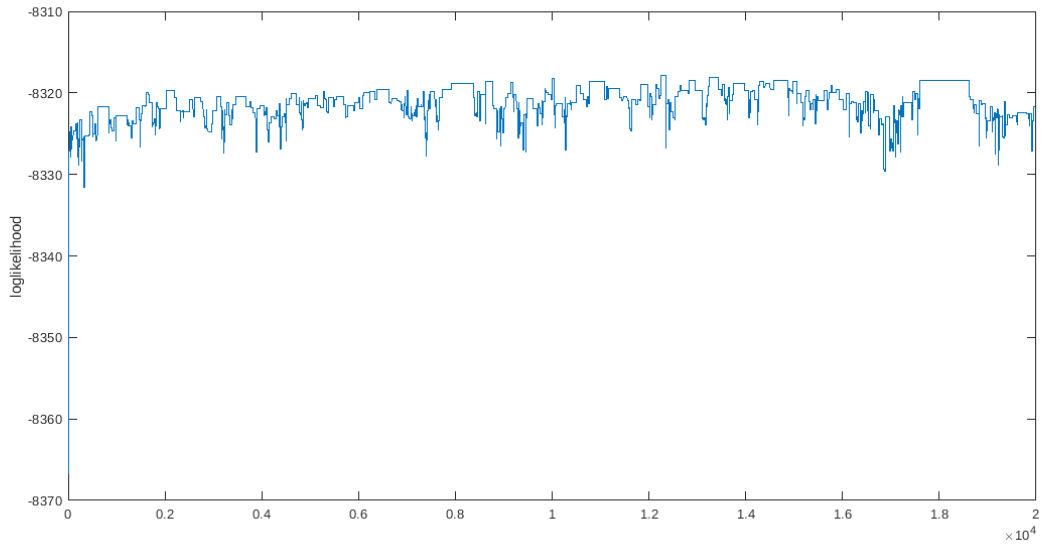

Figure C.5 : Log-likelihood for the model with experimental data when photosynthesis, respiration and respiratory quotients were changing through time and an O_2 offset was introduced.

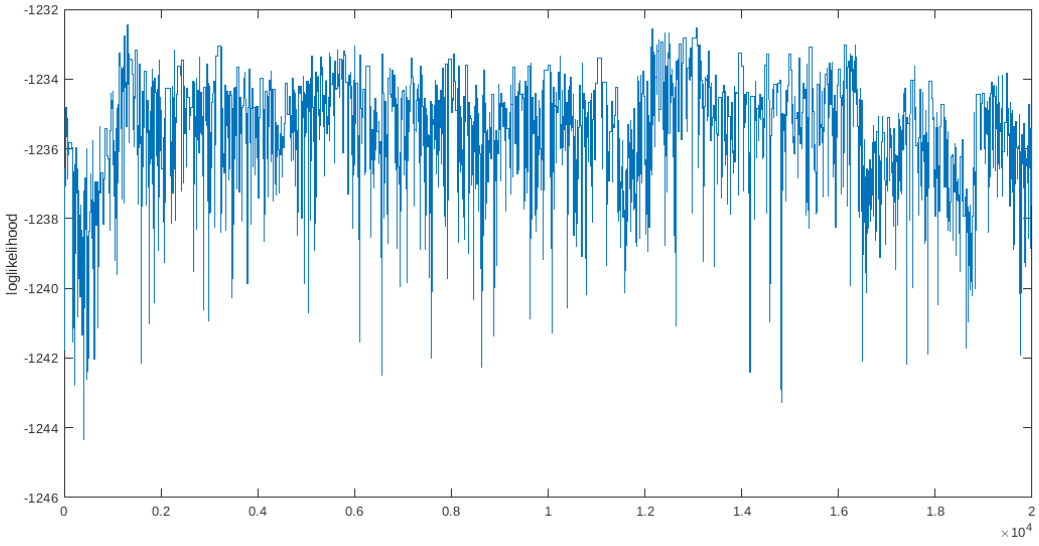

Figure C.6 : Log-likelihood for the model with further thinned O_2 and pH experimental data.

References

[1] C. Andrieu, A. Doucet, and R. Holenstein, "Particle Markov chain Monte Carlo methods," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 72, no. 3, pp. 269–342, 2010.

[2] M. Atkinson and S. Smith, "C: N: P ratios of benthic marine plants," *Limnology and Oceanography*, vol. 28, no. 3, pp. 568–574, 1983.

[3] M. E. Baird, M. P. Adams, R. C. Babcock, K. Oubelkheir, M. Mongin, K. A. Wild-Allen, J. Skerratt, B. J. Robson, K. Petrou, P. J. Ralph *et al.*, "A biophysical representation of seagrass growth for application in a complex shallow-water biogeochemical model," *Ecological Modelling*, vol. 325, pp. 13–27, 2016.

[4] D. F. Baker, S. C. Doney, and D. S. Schimel, "Variational data assimilation for atmospheric CO2," *Tellus B: Chemical and Physical Meteorology*, vol. 58, no. 5, pp. 359–365, 2006.

[5] D. Ball, M. Soto-Berelov, and P. Young, "Historical seagrass mapping in Port Phillip Bay, Australia," *Journal of Coastal Conservation*, vol. 18, no. 3, pp. 257–272, 2014.

[6] J. J. Barber and A. E. Gelfand, "Hierarchical spatial modeling for estimation of population size," *Environmental and Ecological Statistics*, vol. 14, no. 3, pp. 193–205, 2007.

[7] R. Battino, T. R. Rettich, and T. Tominaga, "The solubility of oxygen and

ozone in liquids," *Journal of physical and chemical reference data*, vol. 12, no. 2, pp. 163–178, 1983.

[8] P. Bauer, A. Thorpe, and G. Brunet, "The quiet revolution of numerical weather prediction," *Nature*, vol. 525, no. 7567, p. 47, 2015.

[9] A. R. Bearlin, M. A. Burgman, and H. M. Regan, "A stochastic model for seagrass (Zostera muelleri) in Port Phillip Bay, Victoria, Australia," *Ecological Modelling*, vol. 118, no. 2, pp. 131–148, 1999.

[10] M. J. Bell, R. M. Forbes, and A. Hines, "Assessment of the FOAM global data assimilation system for real-time operational ocean forecasting," *Journal of Marine Systems*, vol. 25, no. 1, pp. 1–22, 2000.

[11] R. Benner, M. L. Fogel, and E. K. Sprague, "Diagenesis of belowground biomass of Spartina alterniflora in salt-marsh sediments," *Limnology and Oceanography*, vol. 36, no. 7, pp. 1358–1374, 1991.

[12] L. M. Berliner, "Hierarchical Bayesian time series models," in *Maximum entropy and Bayesian methods*. Springer, 1996, pp. 15–22.

[13] ——, "Physical-statistical modeling in geophysics," *Journal of Geophysical Research: Atmospheres*, vol. 108, no. D24, 2003.

[14] L. M. Berliner, C. K. Wikle, and N. Cressie, "Long-lead prediction of Pacific SSTs via Bayesian dynamic modeling," *Journal of climate*, vol. 13, no. 22, pp. 3953–3968, 2000.

[15] L. Bertino, G. Evensen, and H. Wackernagel, "Sequential data assimilation techniques in oceanography," *International Statistical Review*, vol. 71, no. 2, pp. 223–241, 2003.

[16] S. Blake and D. Ball, *Seagrass Mapping of Port Phillip Bay.* Marine and Freshwater Resources Institute, 2001.

[17] D. A. Bulthuis and W. J. Woelkerling, "Seasonal variation in standing crop, density and leaf growth rate of the seagrass, Heterozostera tasmanica, in Western Port and Port Phillip Bay, Victoria, Australia," *Aquatic Botany*, vol. 16, no. 2, pp. 111–136, 1983.

[18] C. Calder, M. Lavine, P. Müller, and J. S. Clark, "Incorporating multiple sources of stochasticity into dynamic population models," *Ecology*, vol. 84, no. 6, pp. 1395–1402, 2003.

[19] T. Carruthers and D. Walker, "Light climate and energy flow in the seagrass canopy of Amphibolis griffithii (JM Black) den Hartog," *Oecologia*, vol. 109, no. 3, pp. 335–341, 1997.

[20] K. Chartrand, P. Ralph, K. Petrou, M. Rasheed *et al.*, "Development of a light-based seagrass management approach for the Gladstone western basin dredging program," *Fisheries Queensland, Cairns*, p. 126, 2012.

[21] J. S. Clark and O. N. Bjørnstad, "Population time series: process variability, observation errors, missing values, lags, and hidden states," *Ecology*, vol. 85, no. 11, pp. 3140–3150, 2004.

[22] N. Cressie, C. A. Calder, J. S. Clark, J. M. V. Hoef, and C. K. Wikle, "Accounting for uncertainty in ecological analysis: the strengths and limitations of hierarchical statistical modeling," *Ecological Applications*, vol. 19, no. 3, pp. 553–570, 2009.

[23] N. Cressie and C. K. Wikle, *Statistics for spatio-temporal data.* John Wiley & Sons, 2015.

[24] K.-D. Dang, M. Quiroz, R. Kohn, M.-N. Tran, and M. Villani, "Hamiltonian Monte Carlo with energy conserving subsampling," *arXiv preprint arXiv:1708.00955*, 2017.

[25] T. D. Dickey, "Emerging ocean observations for interdisciplinary data assimilation systems," *Journal of Marine Systems*, vol. 40, pp. 5–48, 2003.

[26] A. Dickson and F. Millero, "A comparison of the equilibrium constants for the dissociation of carbonic acid in seawater media," *Deep Sea Research Part A. Oceanographic Research Papers*, vol. 34, no. 10, pp. 1733–1743, 1987.

[27] A. G. Dickson, C. L. Sabine, and J. R. Christian, *Guide to best practices for ocean CO2 measurements.* North Pacific Marine Science Organization, 2007, no. PICES Special Publication 3.

[28] A. Doucet, N. De Freitas, and N. Gordon, "An introduction to sequential Monte Carlo methods," in *Sequential Monte Carlo methods in practice*. Springer, 2001, pp. 3–14.

[29] M. Dowd, "A sequential Monte Carlo approach for marine ecological prediction," *Environmetrics*, vol. 17, no. 5, pp. 435–455, 2006.

[30] ——, "Estimating parameters for a stochastic dynamic marine ecological system," *Environmetrics*, vol. 22, no. 4, pp. 501–515, 2011.

[31] S. Duane, A. D. Kennedy, B. J. Pendleton, and D. Roweth, "Hybrid monte carlo," *Physics letters B*, vol. 195, no. 2, pp. 216–222, 1987.

[32] C. M. Duarte, "Seagrass nutrient content." *Marine ecology progress series. Oldendorf*, vol. 6, no. 2, pp. 201–207, 1990.

[33] C. M. Duarte and C. L. Chiscano, "Seagrass biomass and production: a reassessment," *Aquatic botany*, vol. 65, no. 1, pp. 159–174, 1999.

[34] G. T. Evans, "Defining misfit between biogeochemical models and data sets," *Journal of marine systems*, vol. 40, pp. 49–54, 2003.

[35] J. W. Fourqurean, C. M. Duarte, H. Kennedy, N. Marbà, M. Holmer, M. A. Mateo, E. T. Apostolaki, G. A. Kendrick, D. Krause-Jensen, K. J. McGlathery *et al.*, "Seagrass ecosystems as a globally significant carbon stock," *Nature geoscience*, vol. 5, no. 7, p. 505, 2012.

[36] J. W. Fourqurean and J. E. Schrlau, "Changes in nutrient content and stable isotope ratios of C and N during decomposition of seagrasses and mangrove leaves along a nutrient availability gradient in Florida Bay, USA," *Chemistry and Ecology*, vol. 19, no. 5, pp. 373–390, 2003.

[37] M. A. Friedrichs, "A data assimilative marine ecosystem model of the central equatorial Pacific: Numerical twin experiments," *Journal of marine research*, vol. 59, no. 6, pp. 859–894, 2001.

[38] G. L. Godshalk and R. G. Wetzel, "Decomposition of aquatic angiosperms. III. Zostera marina L. and a conceptual model of decomposition," *Aquatic Botany*, vol. 5, pp. 329–354, 1978.

[39] N. J. Gordon, D. J. Salmond, and A. F. Smith, "Novel approach to nonlinear/non-Gaussian Bayesian state estimation," in *IEE Proceedings F (Radar and Signal Processing)*, vol. 140, no. 2. IET, 1993, pp. 107–113.

[40] A. Grech and R. Coles, "An ecosystem-scale predictive model of coastal seagrass distribution," *Aquatic Conservation: marine and freshwater ecosystems*, vol. 20, no. 4, pp. 437–444, 2010.

[41] C. Green, "Genotypic differences in the growth of Triticum aestivum in relation to absorbed solar radiation," *Field Crops Research*, vol. 19, no. 4, pp. 285–295, 1989.

[42] A. E. Greensberg, L. Clesceri, A. D. Eaton, and M. Franson, "Standard methods for the examination of water and wastewater," *American Public Health Asociation, Whashington, DC*, 1992.

[43] W. W. Gregg, M. A. Friedrichs, A. R. Robinson, K. A. Rose, R. Schlitzer, K. R. Thompson, and S. C. Doney, "Skill assessment in ocean biological data assimilation," *Journal of Marine Systems*, vol. 76, no. 1-2, pp. 16–33, 2009.

[44] E. M. Grima, J. S. Pérez, F. Í. Garc Camacho, and A. R. Medina, "Gas-liquid transfer of atmospheric CO2 in microalgal cultures," *Journal of Chemical Technology & Biotechnology*, vol. 56, no. 4, pp. 329–337, 1993.

[45] A. R. Grinham, "Downstream effects of land use on shallow-water benthic microalgal communities in Moreton Bay, Australia and Marovo Lagoon, Solomon Islands," 2007.

[46] R. R. Guillard and J. H. Ryther, "Studies of marine planktonic diatoms: I. Cyclotella nana Hustedt, and Detonula confervacea (Cleve) Gran." *Canadian journal of microbiology*, vol. 8, no. 2, pp. 229–239, 1962.

[47] S. Hadley, E. Jones, C. Johnson, K. Wild-Allen, and C. Macleod, "A Bayesian inference approach to account for multiple sources of uncertainty in a macroalgae based integrated multi-trophic aquaculture model," *Environmental Modelling & Software*, vol. 78, pp. 120–133, 2016.

[48] J. W. Hansen, J. W. Udy, C. J. Perry, W. C. Dennison, and B. A. Lomstein, "Effect of the seagrass Zostera capricorni on sediment microbial processes," *Marine Ecology Progress Series*, vol. 199, pp. 83–96, 2000.

[49] P. G. Harrison, "Detrital processing in seagrass systems: a review of factors affecting decay rates, remineralization and detritivory," *Aquatic Botany*, vol. 35, no. 3-4, pp. 263–288, 1989.

[50] A. Hirst, A. Longmore, D. Ball, P. Cook, and G. Jenkins, "Linking nitrogen sources utilised by seagrass in a temperate marine embayment to patterns of seagrass change during drought," *Marine Ecology Progress Series*, vol. 549, pp. 79–88, 2016.

[51] A. J. Hirst and G. P. Jenkins, "Experimental test of N-limitation for Zostera nigricaulis seagrass at three sites reliant upon very different sources of N," *Journal of experimental marine biology and ecology*, vol. 486, pp. 204–213, 2017.

[52] E. L. Jackson, A. A. Rowden, M. J. Attrill, S. J. Bossey, and M. B. Jones, "The importance of seagrass beds as a habitat for fishery species," *Oceanography and marine biology*, vol. 39, pp. 269–304, 2001.

[53] S. Jacobs and D. Les, "New combinations in Zostera (Zosteraceae)," *Telopea*, vol. 12, no. 3, pp. 419–423, 2009.

[54] E. Jones, J. Parslow, and L. Murray, "A Bayesian approach to state and parameter estimation in a Phytoplankton-Zooplankton model," *Australian Meteorological and Oceanographic Journal*, vol. 59, no. SP, pp. 7–16, 2010.

[55] J. E. Kaldy, C. A. Brown, and C. P. Andersen, "In situ 13C tracer experiments elucidate carbon translocation rates and allocation patterns in eelgrass Zostera marina," *Marine Ecology Progress Series*, vol. 487, pp. 27–39, 2013.

[56] D. E. Kaufman, M. A. Friedrichs, J. C. Hemmings, and W. O. Smith Jr, "Assimilating bio-optical glider data during a phytoplankton bloom in the southern ross sea," *Biogeosciences*, vol. 15, no. 1, pp. 73–90, 2018.

[57] H. Kennedy, J. Beggins, C. M. Duarte, J. W. Fourqurean, M. Holmer, N. Marbà, and J. J. Middelburg, "Seagrass sediments as a global carbon sink:

Isotopic constraints," *Global Biogeochemical Cycles*, vol. 24, no. 4, 2010.

[58] M. Kidston, R. Matear, and M. E. Baird, "Parameter optimisation of a marine ecosystem model at two contrasting stations in the Sub-Antarctic Zone," *Deep Sea Research Part II: Topical Studies in Oceanography*, vol. 58, no. 21-22, pp. 2301–2315, 2011.

[59] J. T. Kirk, "Effects of suspensoids (turbidity) on penetration of solar radiation in aquatic ecosystems," *Hydrobiologia*, vol. 125, no. 1, pp. 195–208, 1985.

[60] G. Kitagawa, "Monte Carlo filtering and smoothing method for non-Gaussian nonlinear state space model," *Inst. Statist. Math. Res. Memo.*, 1993.

[61] J. Kuo, "A revision of the genus Heterozostera (Zosteraceae)," *Aquatic Botany*, vol. 81, no. 2, pp. 97–140, 2005.

[62] R. S. Lee, K. P. Black, C. Bosserel, and D. Greer, "Present and future prolonged drought impacts on a large temperate embayment: Port Phillip Bay, Australia," *Ocean Dynamics*, vol. 62, no. 6, pp. 907–922, 2012.

[63] E. Lewis, D. Wallace, and L. J. Allison, "Program developed for CO {sub 2} system calculations," Brookhaven National Lab., Dept. of Applied Science, Upton, NY (United States ..., Tech. Rep., 1998.

[64] M. Libes, "Productivity-irradiance relationship of Posidonia oceanica and its epiphytes," *Aquatic Botany*, vol. 26, pp. 285–306, 1986.

[65] W. A. Link, E. Cam, J. D. Nichols, and E. G. Cooch, "Of BUGS and birds: Markov chain Monte Carlo for hierarchical modeling in wildlife research," *The Journal of wildlife management*, pp. 277–291, 2002.

[66] N. Marbà and D. I. Walker, "Growth, flowering, and population dynamics of temperate Western Australian seagrasses," *Marine Ecology Progress Series*, vol. 184, pp. 105–118, 1999.

[67] R. J. Matear, "Parameter optimization and analysis of ecosystem models using simulated annealing: a case study at Station P," *Journal of Marine Research*, vol. 53, no. 4, pp. 571–607, 1995.

[68] R. J. Matear and E. Jones, "Marine biogeochemical modelling and data assimilation," in *Operational Oceanography in the 21st Century*. Springer, 2011, pp. 295–317.

[69] J. P. Mattern, K. Fennel, and M. Dowd, "Estimating time-dependent parameters for a biological ocean model using an emulator approach," *Journal of Marine Systems*, vol. 96, pp. 32–47, 2012.

[70] E. Mcleod, G. L. Chmura, S. Bouillon, R. Salm, M. Björk, C. M. Duarte, C. E. Lovelock, W. H. Schlesinger, and B. R. Silliman, "A blueprint for blue carbon: toward an improved understanding of the role of vegetated coastal habitats in sequestering CO2," *Frontiers in Ecology and the Environment*, vol. 9, no. 10, pp. 552–560, 2011.

[71] C. Mehrbach, C. Culberson, J. Hawley, and R. Pytkowicx, "Measurement of the apparent dissociation constants of carbonic acid in seawater at atmospheric pressure1," *Limnology and Oceanography*, vol. 18, no. 6, pp. 897–907, 1973.

[72] L. M. Murray, "Bayesian state-space modelling on high-performance hardware using LibBi," *arXiv preprint arXiv:1306.3277*, 2013.

[73] C. Nandi, D. Grossman, A. Sampson, T. Mytkowicz, and K. S. McKinley, "Debugging probabilistic programs," in *Proceedings of the 1st ACM*

SIGPLAN International Workshop on Machine Learning and Programming Languages. ACM, 2017, pp. 18–26.

[74] P. R. Oke, G. B. Brassington, D. A. Griffin, and A. Schiller, "The Bluelink ocean data assimilation system (BODAS)," *Ocean Modelling*, vol. 21, no. 1-2, pp. 46–70, 2008.

[75] R. J. Orth, T. J. Carruthers, W. C. Dennison, C. M. Duarte, J. W. Fourqurean, K. L. Heck Jr, A. R. Hughes, G. A. Kendrick, W. J. Kenworthy, S. Olyarnik *et al.*, "A global crisis for seagrass ecosystems," *Bioscience*, vol. 56, no. 12, pp. 987–996, 2006.

[76] J. Parslow, N. Cressie, E. P. Campbell, E. Jones, and L. Murray, "Bayesian learning and predictability in a stochastic nonlinear dynamical model," *Ecological applications*, vol. 23, no. 4, pp. 679–698, 2013.

[77] K. Petrou, I. Jimenez-Denness, K. Chartrand, C. McCormack, M. Rasheed, and P. Ralph, "Seasonal heterogeneity in the photophysiological response to air exposure in two tropical intertidal seagrass species," *Marine Ecology Progress Series*, vol. 482, pp. 93–106, 2013.

[78] C. Posten, "Design principles of photo-bioreactors for cultivation of microalgae," *Engineering in Life Sciences*, vol. 9, no. 3, pp. 165–177, 2009.

[79] N. Ramsing and J. Gundersen, "Seawater and gases," *Limnol. Oceanogr*, vol. 37, pp. 1307–1312, 2011.

[80] R. Reef, T. B. Atwood, J. Samper-Villarreal, M. F. Adame, E. M. Sampayo, and C. E. Lovelock, "Using eDNA to determine the source of organic carbon in seagrass meadows," *Limnology and Oceanography*, vol. 62, no. 3, pp. 1254–1265, 2017.

[81] G. O. Roberts and J. S. Rosenthal, "Examples of adaptive MCMC," *Journal of Computational and Graphical Statistics*, vol. 18, no. 2, pp. 349–367, 2009.

[82] B. J. Robson, G. B. Arhonditsis, M. E. Baird, J. Brebion, K. F. Edwards, L. Geoffroy, M.-P. Hébert, V. van Dongen-Vogels, E. M. Jones, C. Kruk *et al.*, "Towards evidence-based parameter values and priors for aquatic ecosystem modelling," *Environmental modelling & software*, vol. 100, pp. 74–81, 2018.

[83] J. S. Rosenthal *et al.*, "Optimal proposal distributions and adaptive MCMC," *Handbook of Markov Chain Monte Carlo*, vol. 4, no. 10.1201, 2011.

[84] W. J. Sacks, D. S. Schimel, R. K. Monson, and B. H. Braswell, "Model-data synthesis of diurnal and seasonal CO2 fluxes at Niwot Ridge, Colorado," *Global Change Biology*, vol. 12, no. 2, pp. 240–259, 2006.

[85] T. P. Scoffin, "The trapping and binding of subtidal carbonate sediments by marine vegetation in Bimini Lagoon, Bahamas," *Journal of Sedimentary Research*, vol. 40, no. 1, 1970.

[86] O. Serrano, M. Mateo, P. Renom, and R. Julià, "Characterization of soils beneath a Posidonia oceanica meadow," *Geoderma*, vol. 185, pp. 26–36, 2012.

[87] T. M. Smith, P. H. York, P. I. Macreadie, M. J. Keough, D. J. Ross, and C. D. Sherman, "Spatial variation in reproductive effort of a southern Australian seagrass," *Marine environmental research*, vol. 120, pp. 214–224, 2016.

[88] B. Tamburic, C. R. Evenhuis, J. R. Crosswell, and P. J. Ralph, "An empirical process model to predict microalgal carbon fixation rates in photobioreactors," *Algal research*, vol. 31, pp. 334–346, 2018.

[89] N. Trápani, A. J. Hall, V. O. Sadras, and F. Vilella, "Ontogenetic changes in radiation use efficiency of sunflower (Helianthus annuus L.) crops," *Field Crops Research*, vol. 29, no. 4, pp. 301–316, 1992.

[90] S. M. Trevathan-Tackett, J. R. Seymour, D. A. Nielsen, P. I. Macreadie, T. C. Jeffries, J. Sanderman, J. Baldock, J. M. Howes, A. D. Steven, and P. J. Ralph, "Sediment anoxia limits microbial-driven seagrass carbon remineralization under warming conditions," *FEMS microbiology ecology*, vol. 93, no. 6, 2017.

[91] S. M. Trevathan-Tackett, "Dynamics of refractory carbon in seagrass meadows," Ph.D. dissertation, 2016.

[92] J. M. Ver Hoef, "Parametric empirical Bayes methods for ecological applications," *Ecological Applications*, vol. 6, no. 4, pp. 1047–1055, 1996.

[93] B. A. Ward, M. A. Friedrichs, T. R. Anderson, and A. Oschlies, "Parameter optimisation techniques and the problem of underdetermination in marine biogeochemical models," *Journal of Marine Systems*, vol. 81, no. 1-2, pp. 34–43, 2010.

[94] M. Waycott, C. M. Duarte, T. J. Carruthers, R. J. Orth, W. C. Dennison, S. Olyarnik, A. Calladine, J. W. Fourqurean, K. L. Heck, A. R. Hughes *et al.*, "Accelerating loss of seagrasses across the globe threatens coastal ecosystems," *Proceedings of the National Academy of Sciences*, vol. 106, no. 30, pp. 12 377–12 381, 2009.

[95] R. Weiss, "Carbon dioxide in water and seawater: the solubility of a non-ideal gas," *Marine chemistry*, vol. 2, no. 3, pp. 203–215, 1974.

[96] C. K. Wikle, "Hierarchical Bayesian models for predicting the spread of ecological processes," *Ecology*, vol. 84, no. 6, pp. 1382–1394, 2003.

[97] J. Wilson, I. Valiela, and T. Swain, "Carbohydrate dynamics during decay of litter of Spartina alterniflora," *Marine Biology*, vol. 92, no. 2, pp. 277–284, 1986.

[98] J. Wortmann, J. W. Hearne, and J. B. Adams, "A mathematical model of an estuarine seagrass," *Ecological Modelling*, vol. 98, no. 2-3, pp. 137–149, 1997.

[99] Y. Xiao and M. A. Friedrichs, "Using biogeochemical data assimilation to assess the relative skill of multiple ecosystem models in the mid-atlantic bight: effects of increasing the complexity of the planktonic food web," *Biogeosciences*, vol. 11, no. 11, pp. 3015–3030, 2014.

[100] T. Xu, L. White, D. Hui, and Y. Luo, "Probabilistic inversion of a terrestrial ecosystem model: Analysis of uncertainty in parameter estimation and model prediction," *Global Biogeochemical Cycles*, vol. 20, no. 2, 2006.

[101] R. E. Zeebe, D. Wolf-Gladrow, and H. Jansen, "On the time required to establish chemical and isotopic equilibrium in the carbon dioxide system in seawater," *Marine Chemistry*, vol. 65, no. 3-4, pp. 135–153, 1999.

[102] R. E. Zeebe and D. Wolf-Gladrow, *CO2 in seawater: equilibrium, kinetics, isotopes*. Gulf Professional Publishing, 2001, no. 65.

[103] R. C. Zimmerman, "A biooptical model of irradiance distribution and photosynthesis in seagrass canopies," *Limnology and oceanography*, vol. 48, no. 1, 2003.

[104] J. Zobitz, A. Desai, D. Moore, and M. Chadwick, "A primer for data assimilation with ecological models using Markov Chain Monte Carlo (MCMC)," *Oecologia*, vol. 167, no. 3, p. 599, 2011.