Quantitative Modelling of Root Growth and Carbon Allocation: Bridging Theory and Experiment

By

Clare Ziegler

A thesis submitted to
the University of Birmingham
for the degree of
DOCTOR OF PHILOSOPHY

Birmingham Institute of Forest Research
School of Biosciences
College of Life and Environmental Sciences
University of Birmingham
December 2020

Abstract

Plant root systems play a vital role in carbon sequestration, but the quantitative principles governing their growth and architecture remain poorly understood. The 'forward problem' of what root forms can arise from given models and parameters has been well studied through modelling and simulation, but comparatively little attention has been given to the 'inverse problem': what models and parameters can accurately simulate an experimentally observed root system? This thesis proposes the use of approximate Bayesian computation to infer mechanistic parameters governing root growth and architecture, allowing us to learn and quantify uncertainty in parameters and model structures using observed root architectures. In addition, large-scale field observations of fine roots, derived from belowground imaging and soil cores are combined with with image analysis, stochastic modelling, and statistical inference, to elucidate belowground root dynamics in a mature temperate deciduous forest under free-air carbon enrichment to 150ppm above ambient levels. Results show that elevated carbon dioxide leads to relatively faster root production. Also discussed and quantified is the large, but often neglected, uncertainties in such production measurements resulting from the experimental process, which may then be propagated in further work.

Acknowledgements

Firstly, I would like to thank my supervisors Iain and Rosemary for their endless help, support and patience. Thank you Iain for always having enthusiasm to spare when mine was running low, and for going above and beyond in everything from improving my research to hammering soil cores. Rosemary, thank you for giving me the confidence to do this in the first place, and for always being there with a sensible solution. In addition, I am grateful to Tom for your helpful advice during the monitoring process.

I could not have done this project without the Birmingham Institute of Forest Research, who funded this work with support from the University of Birmingham and the JABBS foundation. I am very grateful to have been given the opportunity to contribute to such important research. The diverse projects and the researchers involved have taught me a great deal, and I have enjoyed being a part of exciting outreach events and getting the opportunity to share my work. The Mill Haft welfare team are award-winning for a good reason, and I would like to thank them for all of the help with my site work. And thank you to all of the excellent BIFoR volunteers, who were instrumental in my data collection. Particularly the two Emilys, for being so consistent, enthusiastic and capable.

Thank you to all of my friends across the departments who have made this PhD enjoyable - and lent me an ear when it wasn't so much! The BIFoR students have brightened up many a site visit, meeting, outreach event, conference and always been a text message away if I needed help or advice. Thank you to the maths students for all of the laughs, and always making sure there was something fun going on. And thank you to all the people in Biosciences who were happy to answer my stupid questions, offer help when I looked lost in the lab, and grab a coffee when I needed a break.

I am endlessly grateful to my family for their support and interest in my pursuits, even going as far as volunteering with fieldwork. I would also like to thank Julia for all your help in the final stages, it made such a difference. Lastly, thank you to my wonderful husband who has been there for every day of this PhD. Thank you for always being there to support me, to cheer me up, to give me space when I needed to work, and to calm me down when I was worrying too much. I love you. And thank you to Chester the dog, who is the best thesis writing companion I could have asked for.

Contents

List of Figures **vi**

List of Tables **viii**

1 Introduction **1**
1.1 Carbon dioxide and climate change 1
1.2 The forestry carbon cycle 2
1.3 Challenges in observing root production 4
1.4 Root system architecture 5
1.5 Mathematical modelling for root studies 5
1.6 Hypothesis testing and likelihood 6
1.7 Thesis structure 6

2 Model selection and parameter estimation for root architecture models using likelihood-free inference **8**
2.1 Introduction 8
2.1.1 Mathematical models of plant root systems 9
2.1.2 Parameter inference using ABC MCMC 11
2.2 Methods 16
2.2.1 Plant growth 16
2.2.2 Model structure 16
2.2.3 ABC SMC implementation 17
2.2.4 Additional work 19
2.3 Results 20
2.3.1 An ABC SMC framework for inferring mechanistic parameters from root systems 20
2.3.2 Inferring parameters from a simulated root system 23
2.3.3 Inferring mechanistic parameters for other synthetic phenotypes and root simulation models 25
2.3.4 Inferring mechanistic parameters for wildtype *Arabidopsis thaliana* root systems 27
2.3.5 Model selection for root growth and branching mechanisms 29
2.3.6 Comparison between root structures 33
2.4 Discussion 34

3 Characterising root processes at the BIFoR FACE site **38**
3.1 Introduction . . . 38
3.1.1 Experimental observation of fine root dynamics under elevated carbon dioxide . . . 39
3.1.2 BIFoR and other FACE experiments . . . 40
3.1.3 Observing fine root growth using minirhizotrons . . . 42
3.1.4 Sampling root biomass using soil cores . . . 43
3.2 Methods . . . 45
3.2.1 Experimental setup at BIFoR FACE . . . 45
3.2.2 Minrhizotrons . . . 46
3.2.3 Soil cores . . . 48
3.2.4 Additional methods . . . 51
3.3 Results . . . 52
3.3.1 Root lengths and widths from minirhizotrons . . . 52
3.3.2 Root weights from soil cores . . . 53
3.3.3 Root biomass from minirhizotrons . . . 55
3.3.4 Root production . . . 55
3.3.5 Fold-change in fine root biomass . . . 60
3.4 Discussion . . . 60

4 Exploring data from BIFoR FACE using mathematical modelling **64**
4.1 Introduction . . . 64
4.1.1 Stochastic modelling and quantifying uncertainty . . . 65
4.1.2 NPP . . . 67
4.2 Methods . . . 68
4.2.1 BID model dynamics . . . 68
4.2.2 Statistical analysis and uncertainty quantification . . . 70
4.2.3 Maximum likelihood estimation for parameter values . . . 71
4.2.4 Bootstrapping on parameter estimation . . . 71
4.2.5 Calculating NPP from fine root production . . . 72
4.2.6 Calculating NPP using Caladis . . . 76
4.3 Results . . . 78
4.3.1 Stochastic modelling for fine root biomass . . . 78
4.3.2 NPP estimation and uncertainty . . . 79
4.4 Discussion . . . 83

5 Discussion **87**

Bibliography **91**

A Outreach and deliverables **106**
A.1 Conference presentations . . . 106
A.2 Awards . . . 106
A.3 Articles and videos . . . 107
A.4 Outreach . . . 107
A.5 Other deliverables . . . 107

B Published paper **108**
B.1 Introduction . . . 108
B.2 Results . . . 110
B.2.1 An ABC SMC framework for inferring mechanistic parameters from root systems . . . 110
B.2.2 Inferring parameters from a simulated root system . . . 111
B.2.3 Inferring mechanistic parameters for other synthetic phenotypes and root simulation models . . . 113
B.2.4 Inferring mechanistic parameters for wildtype *Arabidopsis thaliana* root systems . . . 113
B.2.5 Model selection for root growth and branching mechanisms . . . 115
B.2.6 Comparison between root structures . . . 117
B.3 Discussion . . . 119
B.4 Methods . . . 122
B.4.1 Plant growth . . . 122
B.4.2 Model structure . . . 122
B.4.3 ABC SMC implementation . . . 122
B.5 Competing Interests . . . 124
B.6 Acknowledgements . . . 124
B.7 Author Contributions . . . 124
B.8 Appendix . . . 125
B.9 Supplementary Information . . . 125

List of Figures

1.1 Forestry carbon cycle. 3
1.2 Relative root orders. 4

2.1 Validating root inference platform with synthetic *Arabidopsis* data 23
2.2 Posteriors on l_{max} for *Arabidopsis* wild-type seedlings 24
2.3 Posterior distributions on *Arabidopsis thaliana* and *Lupinus angustifolius* roots generated using RootBox. 26
2.4 Posterior distributions on b_{max} for *Arabidopsis thaliana* and *Lupinus angustifolius* roots generated using RootBox. 27
2.5 Example data and simulation output for the root inference framework . . . 28
2.6 Summary statistic comparison between data and parameterised model. . . 29
2.7 Posterior distributions on mechanistic parameters for *Arabidopsis* seedling roots. 30
2.8 Selection of competing mechanistic models for root growth using ABC SMC. 32
2.9 Distinguishing phenotypes with mechanistic inference. 35

3.1 Minirhizotron setup 44
3.2 Example of image analysis perfomed on minrhizotron images 48
3.3 Scanned roots from a soil core 50
3.4 Fine root observations under eCO_2 and control. 54
3.5 Root biomass per horizon sampled from soil cores 56
3.6 Root biomass per array sampled from soil cores 57
3.7 Imaging region of a minirhizotron for root production 58
3.8 Production across tubes in eCO_2 and control arrays 58
3.9 Root growth example 59
3.10 A comparison of the raw biomass per tube and the 'fold-change' transformation 61

4.1 Data collection and modelling schematic 65
4.2 Diagram of processes in a BID model 66
4.3 Model geometry for production calculations 73
4.4 Root biomass changes over time under eCO_2 and control conditions in the first year of sampling. 80
4.5 Root biomass changes over time under eCO_2 and control conditions in the second year of sampling. 81
4.6 Net primary productivity estimates and uncertainties. 83
4.7 NPP estimates from Caladis. 84

B.1 Validating root inference platform with synthetic *Arabidopsis* data 112
B.2 Posterior distributions on *Arabidopsis thaliana* and *Lupinus angustifolius* roots generated using RootBox. 114
B.3 Example data and simulation output for the root inference framework . . . 115
B.4 Summary statistic comparison between data and parameterised model. . . 116
B.5 Posterior distributions on mechanistic parameters for *Arabidopsis* seedling roots. 116
B.6 Selection of competing mechanistic models for root growth using ABC SMC. 118
B.7 Distinguishing phenotypes with mechanistic inference. 120
B.8 Posteriors on l_{max} for *Arabidopsis* wild-type seedlings 125
B.9 Posterior distributions on b_{max} for *Arabidopsis thaliana* and *Lupinus angustifolius* roots generated using RootBox. 125
B.10 Posteriors on l_{max} for *Arabidopsis* wild-type seedlings 126
B.11 Posterior distributions on b_{max} for *Arabidopsis thaliana* and *Lupinus angustifolius* roots generated using RootBox. 126

List of Tables

2.1 Model summary. 31

3.1 A selection of key forestry FACE experiments 42

List of abbreviations

- **ABC**: Approximate Bayesian Computation
- **BID**: Birth Immigration Death
- **BIFoR**: Birmingham Institute of Forest Research
- **BTC**: Bartz Technology Corporation
- **eCO_2**: elevated carbon dioxide
- **FACE**: Free Air Carbon Enrichment
- **GPP**: Gross Primary Production
- **LOESS**: Locally Estimated Scatterplot Smoothing
- **MCMC**: Markov Chain Monte Carlo
- **NPP**: Net primary production
- **ORNL**: Oak Ridge National Laboratory
- **PDE**: Partial Differential Equation
- **SMC**: Sequential Monte Carlo

Chapter 1

Introduction

1.1 Carbon dioxide and climate change

Carbon dioxide is a major greenhouse gas [Bolin and Doos, 1989], and increasing levels of carbon dioxide in the greenhouse layer are contributing to climate change [Hoyt, 1979]. While there is a great deal of focus on reducing human carbon emissions [Sawyer, 1972, Keeling, 1997, Fang et al., 2018], which are indeed a major contributing source [IPCC, 2013], carbon dioxide is also released by natural carbon sources [Lüthi et al., 2008] such as gas release from the oceans [Key et al., 2004] and decomposition of vegetation [Jobbágy and Jackson, 2000]. Conversely, terrestrial and oceanic ecosystems can act as carbon sinks, where carbon is stored for a period of time [Houghton, 2002]. Forests can provide an important carbon sink [Pan et al., 2011], as plants capture the carbon dioxide through photosynthesis before it reaches the greenhouse layer. Carbon may then pass into the soil [Bruce et al., 1999], be stored as plant matter [Min and Guangsheng, 2004], or be released through decomposition [Gougoulias et al., 2014] or respiration [Chambers et al., 2004]. An understanding of the carbon cycle within complex forestry ecosystems is therefore crucial to our understanding of the possible carbon uptake of the system [Grace, 2004]. This uptake may change as the ambient carbon dioxide levels increase [Gifford, 1992], and

forestry ecosystems across the globe may respond differently [Gower, 2003]. Therefore, we need to understand how the forestry carbon cycle will be affected by future ambient carbon dioxide levels in order to make accurate climate change predictions [Bradley and Pregitzer, 2007].

1.2 The forestry carbon cycle

There are many carbon stores, sinks and flows within a forestry ecosystem [Houghton, 1996], as shown in Figure 1.1. Carbon enters the system through photosynthesis; the amount of carbon fixed by all photosynthesisers within the system is known as the Gross Primary Productivity (GPP) [Gower, 2003]. Some of this carbon is then released through respiration, including aboveground respiration, root respiration, and soil respiration [Delucia et al., 2007], or stored as biomass within the system [Nabuurs et al., 2013]. Carbon may then enter the soil through biomass decomposition [Karhu et al., 2010], or root exhudates [Kumar et al., 2006]. Measuring net primary production (NPP) is crucial to our understanding of the carbon cycle [Fahey and Knapp, 2007]. In plants, NPP is the measurement of the energy stored as biomass [Fath, 2018] and represents the difference between energy gained through photosynthesis and that used for respiration [Waring, 2007]. Since NPP cannot be directly caculated [Clark et al., 2001a], it is often estimated using biomass production over a given time period [Clark et al., 2001b]. This contains many sources of error and may underestimate true NPP by up to 30% [Eviner, 2004]. In addition, belowground NPP is often neglected, sometimes assumed to be a proportion of aboveground NPP estimates [Medlyn et al., 2016, De Kauwe et al., 2016, Adame et al., 2017].

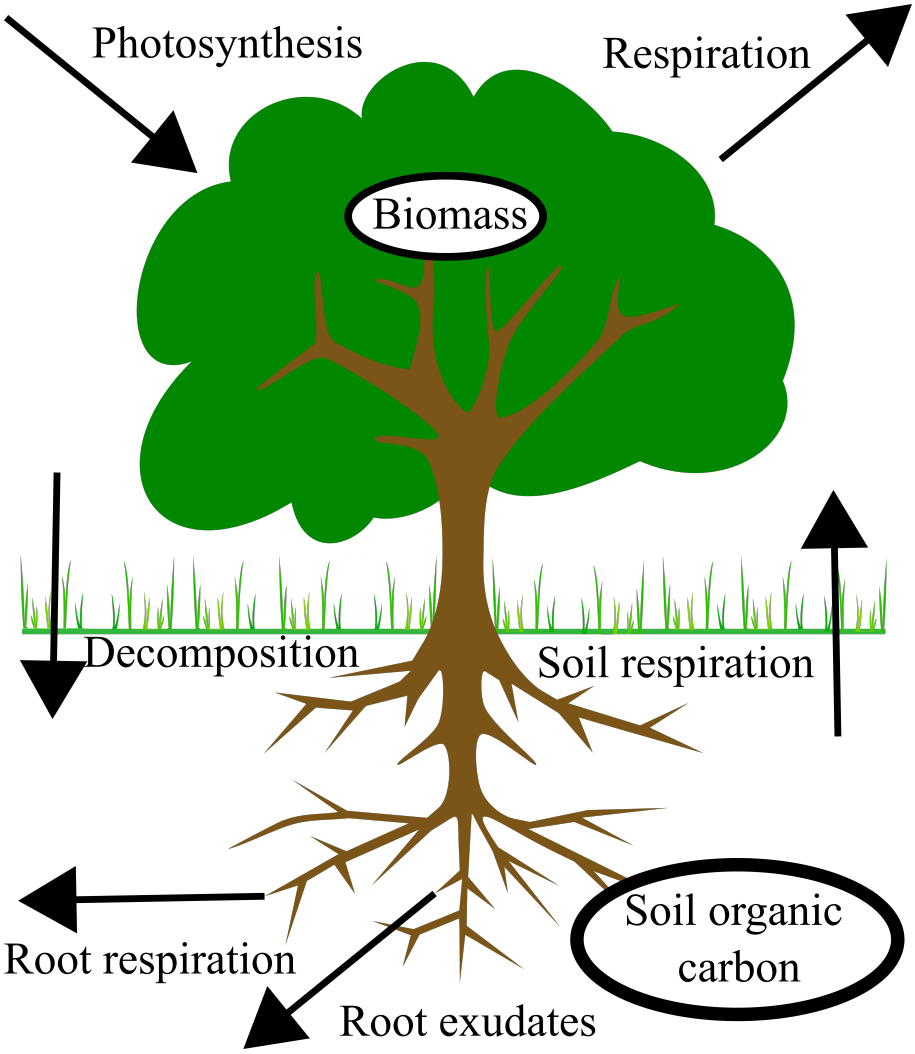


Figure 1.1: **The forestry carbon cycle.** A simplified representation of carbon flows within the forestry ecosystem. Arrows denote the movement of carbon and circles a carbon sink.

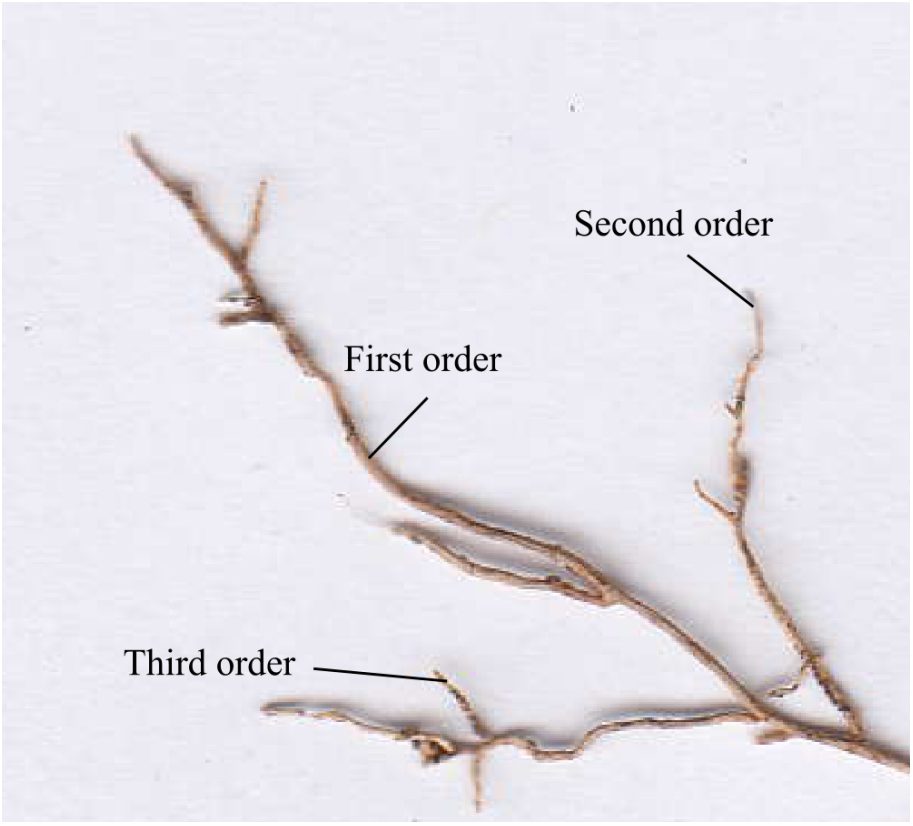


Figure 1.2: **Relative root orders.** A section of a bramble (*Rubus fruticosus*) root system with the relative root orders labelled. Note that since the whole root system is not imaged the labels denote relative root orders only.

1.3 Challenges in observing root production

While aboveground production is readily observed [Clark et al., 2001a], measurement of production belowground is far more challenging, particularly non-destructively [Majdi, 1996]. However, root systems are a crucial part of the forestry carbon cycle and should not be neglected, as they are estimated to be responsible for 5% by mass of the total global carbon budget [Jackson et al., 1997]. They stabilise plants [Yang et al., 2017], stabilise soils [Vannoppen et al., 2015], foster beneficial microbes [Reinhold-Hurek et al., 2015] and are the entry point for water and nutrients to the plant [Dotaniya and Meena, 2015]. The shape of a plant's root system is generated by a variety of physiological and signalling pathways within the plant, and understanding the generation of this system opens paths to its optimisation to maximise crop yield [Shahzad and Amtmann, 2017].

1.4 Root system architecture

Plant roots may follow a variety of structures depending on the phenotype [Morris et al., 2017]. Although a broad consensus has not been fully reached, there have been efforts to produce a general framework for root system taxonomy [Bodner et al., 2013, Liu et al., 2018]. Therefore, I will define some common terms that will be used within this work. The tap root is defined as the first root to emerge from the seed [Esau et al., 1965], and can clearly be seen in *Arabidopsis Thaliana*. Lateral roots are then defined as any root branching from another root [Péret et al., 2009]. When the whole root is observed, the order of a root branch can be taken as it's distance from the main root; a second order lateral is a branch from a first order lateral and so on. However, for roots sampled or observed in the field, the whole root system may not be visible and so we refer to relative root order, as seen in Figure 1.2.

1.5 Mathematical modelling for root studies

The interaction between roots (and plants) and carbon dioxide is a complex and multiscale process [Garnier et al., 2017, Verheijen et al., 2016, Bradley and Pregitzer, 2007, Jackson et al., 1997], compounded by the difficulties in making quantitative observations [Iversen et al., 2017]. In such problems, tools from mathematical modelling and statistics can increase the power and interpretability of underlying theory [Fowler et al., 1997]. Stochasticity is included in models across biology to capture the apparent randomness often observed in biological processes [Morgan, 2008]. For example, stochasticity is required to capture cell dynamics, cell populations and gene expression [Wilkinson, 2009]. Therefore, stochastic models are well suited to the noisy data obtained from experimental work at BIFoR FACE.

1.6 Hypothesis testing and likelihood

When addressing a research question, we typically have a null and an alternative hypothesis – for example, a null hypothesis may be that eCO_2 does not affect roots and the alternative is that eCO_2 increases root growth. Statistics gives us ways of comparing the support for these hypotheses given empirical observations. An understanding of statistics is crucial to accurate scientific reporting of research results [Ioannidis, 2005]. An commonly used statistic in hypothesis testing is the p-value, although the overreliance of this result comes with some controversy [Nuzzo, 2014, Ioannidis, 2018, Halsey et al., 2015]. The p-value, popularised by Sir Ronald Fisher in 1925 [Biau et al., 2010, Pearce, 1992], represents the probability that the observed effect would be as large or larger if the null hypothesis is true [Thiese et al., 2016]. Generally in biosciences a result is considered statistically significant if this probability is less than 0.05 [Dorey, 2010]. Fisher also first made the case for using likelihood to measure the fit of a statistical model to data [Fienberg and Kadane, 2001]. The likelihood function, $\mathcal{L}(\theta|x)$ measures the fit of a model to a given dataset, and is defined as the probability mass or density function of observed data x, viewed as a function of parameter set θ [Azzalini, 1996]. In this way, the likelihood function helps us identify the structures and parameterisations of hypothesised models that are most compatible with empirical observations. However for some models the likelihood function may be impossible or difficult to compute, necessitating the use of likelihood-free inference methods [Gutmann et al., 2018].

1.7 Thesis structure

The following chapters detail the use of mathematical and statistical modelling to gain greater insight into root growth, and to investigate the impact of elevated carbon dioxide in a temperate mature oak woodland. Chapter 2 contains work using likelihood-free inference to investigate root growth parameters. I detail the strength of this approach for

distinguishing phenotypic differences between root systems, extracting growth parameters from observed root systems, and in model selection. Chapter 3 contains experimental data collected from the BIFoR FACE experiment, and the calculation of NPP from this data. This work is then expanded upon in Chapter 4, where I use stochastic modelling and uncertainty quantification to gain greater insight into the data.

Chapter 2

Model selection and parameter estimation for root architecture models using likelihood-free inference

2.1 Introduction

In this chapter, I detail work using ABC SMC to investigate mathematical models of root systems. I propose a method of taking an existing root system and finding the growth and branching rates that contributed to its form. The application of the mathematical framework shows a clear separation of these rates for different root systems, and can also be used with simulated roots. The framework is applied to model selection, allowing us to compare and contrast how well existing root simulation models recreate root growth we observe in real plants. This work was published in *Journal of the Royal Society Interface*, and the full paper can be found in Appendix B.

2.1.1 Mathematical models of plant root systems

Many mathematical modelling techniques have been employed to represent root processes [Wu et al., 2005]. Source sink models such as the recent model produced by Feller *et al* [Feller et al., 2015] aim to capture resource allocation within a plant. Parts of the plant such as the shoot and root are described as compartments and include flow rate equations describing the movement of key nutrients within the plant. This approach provides a simple method of capturing nutrient movement but fails to take into account plant morphology or more complicated processes such as hormone pathways. This approach is however useful to consider flow from the shoot to the root system in a model of root growth and structure.

Plant growth models allow for investigation into the underlying processes without the difficulties of growing and imaging plant root systems [Fourcaud et al., 2008]. Functional-structural plant models are key to this approach, as they seek to produce a simulation of root system architectures, and can be useful in investigating root responses to various environments [Dunbabin et al., 2013]. Lungley (1973) produced one of the first computer simulated root systems, which generated two-dimensional outputs made up of ASCII characters [Lungley, 1973]. Root systems were simulated with prescribed rates of elongation and branching for primary and secondary roots, with a fixed inter-branch distance. Nutrient uptake from soil was characterised by looking at the percentage of total root length at different soil depths. A similar approach was taken by Gerwitz and Page (1974) who produced a model of nutrient uptake using a simple ODE fitted to data [Gerwitz and Page, 1974].

A constant growth rate was also used by Pages *et al* in 1989 to produce a 3D simulation of maize, but only for first order roots. Roots of second and third order have a negative exponential growth function dependent on the root age [Pages et al., 1989]. Growth direction is computed by resolving forces representing gravitropism, initial growth direction and resistive forces from the soil. A network-based approach was taken by Lynch *et al*

(1979), who modelled a root system as a network, with nodes as branches and inter-branch distances as edges. The model also included root radius, and volume chages along the growing root [Lynch et al., 1997]. A mass conservation law as used by Ozier-Lafontaine *et al* in their fractal analysis model, which defined the diameter of newly formed root branches according to a mass conservation law [Ozier-Lafontaine et al., 1999].

Previous modelling approaches were combined in the production of RootTyp in 2004 [Pagès et al., 2004], which allowed the simulation of root systems from a variety of plant species, in contrast to many earlier studies [Pages et al., 1989]. This has allowed the model to be adapted for use by other researchers in a range of studies [Collet et al., 2006, Garré et al., 2012]. Another key root architecture model is RootBox, [Leitner et al., 2010] which is designed to be combined with soil and water uptake models, along with allowing for the simulation of roots grown in containers of user-defined shape and dimensions to investigate root-soil interactions. This model has been recently improved with the advent of CRootBox, and there are plans to eventually extend the modelling approach to also consider above-ground plant growth [Schnepf et al., 2018]. An effective model is able to reproduce root systems of many different plant species, which necessitates the incorporation of root data collected in situ [Roose et al., 2016]. Due to advances in imaging techniques, models have begun to be informed by data collected through x-ray μct, where a 3D image is developed by stacking cross sectional x-rays, and MRI imaging [Metzner et al., 2015].

Root system data used to inform architecture models is collected from image analysis techniques from photos taken of lab-grown roots or from the aforementioned imaging techniques. Many pieces of software have been developed to analyse root images [Lobet et al., 2013], often highly specialised to each experimental approach [Lobet, 2017]. The Centre for Plant Integrated Biology at the University of Nottingham has taken this approach with RooTrak [Mairhofer et al., 2012], RootNav [Pound et al., 2013] and RooTrace [French et al., 2009]. A threshold algorithm is used to map the root system from user-identified start points, making the assumption that the roots are lighter than the background. This

semi-automated approach was also used in the production of SmartRoot [Lobet et al., 2011], but with the inclusion of a more manual approach, which can be used to improve accuracy if necessary. This makes the software particularly useful for less complex images. Key root system data can then be extracted for use in modelling.

2.1.2 Parameter inference using ABC MCMC

Approximate Bayesian Computation (ABC) provides a statistical framework for parameter inference with comparison to data, when a likelihood function is not easily computable [Didelot et al., 2011]. It is highly adaptable, but requires careful consideration of results in the context of how it has been applied [Bertorelle et al., 2010]. A simple ABC implementation is the rejection algorithm [Csilléry et al., 2010], in which a sample parameter value $\hat{\theta}$ is pulled from a predefined prior which describes the probability distribution of its value. The simulation is then run with $\hat{\theta}$, and a dataset $\hat{D}$ is produced as the model output. This dataset is compared to existing dataset D by some measure $\rho(\hat{D}, D)$ with a given tolerance ϵ so that if $\rho(\hat{D}, D) \leq \epsilon$, $\hat{\theta}$ is accepted and recorded or otherwise discarded. A new $\hat{\theta}$ is then pulled from the same prior distribution and the process repeated until a required number of hits have been reached. By recording the values of $\hat{\theta}$ that satisfy $\rho(\hat{D}, D) \leq \epsilon$ a posterior distribution can be produced that describes the likely values of parameter values θ to fit the dataset D [Toni et al., 2009]. This requires many decisions to be made during setup such as the choice of priors, the tolerance, the number of hits required and the choice of parameters. Each of these choices will impact the output of the model and therefore any evaluation of outputs should be done with careful consideration of these potential impacts [Didelot et al., 2011].

Markov Chain Monte Carlo (MCMC) is an ABC approach that converges more efficiently than the previously described rejection algorithm. $\hat{\theta}$ is pulled from the prior as before, but after a value has been accepted, a new $\hat{\theta}$ comes from perturbing the previously accepted value using a user defined perturbation kernel $K(\hat{\theta}, \theta)$. If the new value is

accepted it is perturbed for the next iteration, if not, the perturbation kernel is applied to the previously accepted value of $\hat{\theta}$ [Marjoram et al., 2003]. The perturbation kernel must be carefully chosen to facilitate exploration of the whole parameter space.

Both the rejection algorithm and MCMC methods can be applied to model selection, by introducing a random model selection before the generation of parameter values. This is difficult to implement with MCMC however, due to the dependence on previously accepted values. The rejection algorithm quickly becomes very inefficient when implemented for model selection, particularly for models with multiple parameters. SMC provides a more efficient ABC method and is well suited for model selection [Toni et al., 2009]. SMC works by considering T populations, with tolerances $\epsilon_1 < \epsilon_2 < ... < \epsilon_T$ for each population. The rejection algorithm is applied to produce the initial population of size N with tolerance ϵ_1. For the second population, values of $\hat{\theta}$ are taken from the N previously accepted values, weighted appropriately to compensate for the differing prior, perturbed by a perturbation kernel, and then used to generate $\hat{D}$ and accepted with tolerance ϵ_2. If a value is not accepted, a new value from the previous population is perturbed. This is repeated until N values of $\hat{\theta}$ have been accepted, and the process is then repeated until there are N accepted values in population T, which make up the posterior outputs. ABC SMC is a method that will be central to the work in this chapter.

While a great number of advances have been made in simulating root systems from a set of parameters, relatively little work has been done on the inverse problem: extracting the growth parameters from a given root system. This step is crucial to gain biological insight from these root models and gaining information about a plant from observing its root system. It is vital that we can validate root models and their application and can quantify uncertainty in the mechanisms and parameters. Advancing technologies are allowing observation of root systems in increasing detail, making it ever more important to bridge the gap between theory and observation.

A major challenge in solving the inverse problem is that of finding a likelihood for an

observed root system. This can be aided through the use of stochastic simulation and ABC techniques, which produce an approximation of the likelihood and remove the need for its explicit calculation. Another strength of these techniques lies in model selection and the inclusion of priors.

Root systems are essential to plants' structure and uptake of water and nutrients, and constitute more than 5% by mass of the total global carbon budget [Jackson et al., 1997]. They stabilise plants [Yang et al., 2017], stabilise soils [Vannoppen et al., 2015], foster beneficial microbes [Reinhold-Hurek et al., 2015] and are the entry point for water and nutrients to the plant [Dotaniya and Meena, 2015]. The shape of a plant's root system is generated by a variety of physiological and signalling pathways within the plant, and understanding the generation of this system opens paths to its optimisation to maximise crop yield [Shahzad and Amtmann, 2017].

Despite this importance, the mechanisms underlying root growth remain challenging to quantitatively understand [Dunbabin et al., 2013, Roose et al., 2016, Sievänen et al., 2014]. The complexity of root systems and their belowground nature poses observational challenges. Experimental techniques aiming to elucidate root architecture have historically included sketches of root systems and the use of hydroponics, then images of cleaned root systems. More recent advances have facilitated the imaging of plants *in situ* through the use of x-ray μ-Computed Tomography [Mairhofer et al., 2013], MRI scanning [Metzner et al., 2015, van Dusschoten et al., 2016], and transparent soil [Downie et al., 2012], which have been used to investigate root soil exploration and uptake of water and nutrients.

In parallel with this experimental elucidation, in an effort to understand how root systems grow, many physical and mathematical models of root growth and structure have been produced. These models solve the 'forward problem': given knowledge of the parameters governing growth processes in plants, they produce the details and dynamics of a likely simulated root system. The Lockhart equation described the elongation of a cell under turgor pressure [Lockhart, 1965], and has been widely adapted to describe the

growth of many plant organs, including roots [Dyson et al., 2014, Geitmann and Ortega, 2009]. Hackett and Rose produced the first root system model in the 1970s [Hackett and Rose, 1972] based on the growth and branching of barley roots, while Lungley [Lungley, 1973] produced a computational model which generated root systems represented using ASCII characters. Fitter [Fitter, 1987] introduced a topological model of root architecture where a root system was considered as a set of links. This idea was extended in the three dimensional model of Pagès [Pages et al., 1989] and Diggle, whose ROOTMAP model could be applied to a variety of plant species [Diggle, 1988], while Tatsumi *et al.* represented variation in root systems using fractal analysis [Tatsumi et al., 1989, Pagès et al., 2000] Lynch *et al.* modelled a root system as a network with nodes as branches and inter-branch distances as edges. The model also included root radius, and volume changes along the growing root [Lynch et al., 1997]. Advances in computation have led to a plethora of root system architecture models, which produce a three dimensional reproduction of a root system using a detailed parameter set [Dunbabin et al., 2013]. Previous modelling approaches were combined in the production of Root Typ in 2004 [Pagès et al., 2004]. This has allowed the model to be adapted for use by other researchers [Collet et al., 2006, Garré et al., 2012]. Another key root architecture model is RootBox, [Leitner et al., 2010] which is designed to be combined with soil and water uptake models along with allowing for the simulation of roots grown in containers of user-defined shape and dimensions. This model has been recently updated to produce CRootBox, and there are plans to eventually extend the modelling approach to also consider above-ground plant growth [Schnepf et al., 2018]. An effective model is able to reproduce root systems of many different plant species, which necessitates the incorporation of root data collected *in situ* [Roose et al., 2016].

While a great number of advances have been made in simulating root systems from a set of parameters, relatively little work has been done on the inverse problem: extracting the growth parameters from an observed root system. Model parameterisation is often limited by difficulties in root observation [Garré et al., 2012, Pagès and Pellerin, 1994],

however, this step is crucial to gain biological insight from these root models. It is vital that we can validate root models and the predictions they make, and quantify uncertainty in their mechanisms and parameters. A manual approach to the inverse problem, feeding specific measured parameters into generative models for root systems and assessing their ability to reproduce observations, has been used to gain biological insight and validate generative models [Chen et al., 2011]. However, without an automated approach, it remains challenging to explore the full ranges of parameters and mechanisms that could give rise to observed structures, and the likelihoods of each. Advancing technologies are allowing observation of root systems in increasing detail, making it ever more important to bridge the gap between theory and observation.

A major challenge in solving the inverse problem with traditional statistical methods is finding a likelihood function for an observed root system. Modern statistical approaches allow this challenge to be circumvented through the use of stochastic simulation and Approximate Bayesian Computation (ABC) techniques [Beaumont et al., 2002], which produce a computational approximation replacing the likelihood and remove the need for its explicit calculation. Another strength of these techniques lies in their natural capacity for model selection and the inclusion of prior knowledge about the system in an inference setting. Here, we report a novel pipeline by which ABC, embedded in an Sequential Monte Carlo (SMC) framework [Toni et al., 2009], can be used to learn the values of and uncertainty in generative, mechanistic parameters underlying root growth and architecture, and to compare different root architecture models. *Arabidopsis thaliana* (thale cress) is used in both computational and experimental investigation throughout as a model plant, but this process can readily be extended to any root system, as we also demonstrate with *Lupinus angustifolius* (narrowleaf lupin). We demonstrate how this framework can be used to identify generative parameters according to a given model, distinguish phenotypic differences, and evaluate the comparative effectiveness of models for root elongation and root branching processes, providing insight into the underlying mechanisms.

2.2 Methods

2.2.1 Plant growth

Arabidopsis Col-0 seeds were sterilised with three 3-minute washing steps in 50% domestic bleach and water rinses, then plated on $\frac{1}{2}$ MS agar in vertical plates. Plants were grown at constant 25°C on a 16h light / 8h dark cycle. Plates were photographed over a time course of 2, 5, 7, and 10 days to produce time-series images of the seedling growth. Summary statistics were extracted from the images using SmartRoot [Lobet et al., 2011], an imageJ plugin. The root systems were traced manually using thresholding, producing a skeleton over the original image. Summary statistics on root length and branch placement were then recorded from this skeleton.

2.2.2 Model structure

Root growth was simulated using a hybrid stochastic-deterministic algorithm. Primary root growth, by default, was assumed to follow a negative-exponential growth law:

$$l(t) = l_{max}\left(1 - e^{-gt/l_{max}}\right), \tag{2.1}$$

parameterised by a rate constant g and a scaling constant l_{max}. The alternative uniform growth model simply took the form $l(t) = gt$. Lateral roots grow according to the same growth law as the primary root, but with a multiplicative factor α applied to g so that for lateral roots $g_l = \alpha g$.

Branching was treated as a Poisson event with rate parameter b. The time until the next branching event is found using the Gillespie algorithm [Gillespie, 1977], and the length of existing branches is updated from the current time until the time of the branching event. The branching location was then determined by a specified branching

model, initially specified as a uniform probability distribution along the length of the primary root. In visualising structures, branching angle was always set to an angle of 45° from the growth direction, with equal change of being placed each side of the primary root, although these angles and positions play no role in the simulation. These steps are repeated until the time of the next branch exceeds the maximum simulation time, at which point the branch lengths are updated up to the maximum simulation time, and no branching event occurs. Once a branching event has occurred, the sidebranch grows according to the same growth law as the main root, scaled by parameter α; variability in lateral root length thus corresponds to variability in initial branching times and positions.

2.2.3 ABC SMC implementation

An ABC framework was implemented in Matlab. Model parameters were drawn from specified distributions and passed to the model as described in Model Structure above. Broadly, the simulated root systems are then compared to data, and the parameter values accepted if the simulation is sufficiently close to the data, with tolerance ϵ defined at the time of implementation. If the previous values were accepted, the parameter values are perturbed with a perturbation kernel K_t. If the previous values were not accepted, the parameter values were drawn from the priors as previously described. This process was repeated until 1000 hits were obtained at the specified tolerance.

We follow Ref. [Toni et al., 2009] in our ABC SMC implementation. For completeness, Algorithm 1 introduces a simple rejection sampling scheme under ABC. Algorithm 2 embeds this scheme in an SMC framework for parameter inference and model selection.

Algorithm 1. ABC rejection sampling for parameter inference.

1. Given N_p plant structures and $N_t(i)$ longitudinal observations for plant i, characterise the summary statistics $d_{ij} = \{B, L, \hat{l}\}$ from every plant i and observation j in

the dataset.

2. Draw a trial set of parameters θ^* from the prior distribution $\pi(\theta)$.
3. Simulate N_p instances of root growth, recording the state of structure i at each of the $N_t(i)$ time points corresponding to an experimental observation.
4. Compute ρ using Eqn. 2.4 above, to give the separation between each recorded structure and its simulated counterpart.
5. If $\rho < \epsilon$, where ϵ is a given tolerance, accept θ^* as a sample from the posterior.
6. If a termination condition is not met, return to 2.

Algorithm 2. ABC SMC for parameter inference and model selection.

1. Given N_p plant structures and $N_t(i)$ longitudinal observations for plant i, characterise the summary statistics $d_{ij} = \{B, L, \hat{l}\}$ from every plant i and observation j in the dataset.
2. Initialise tolerance vector $\mathbf{E}$ containing T elements. Set population indicator $t = 0$.
3. Set particle indicator $i = 1$.
4. Sample model indicator m^* from prior $\pi(m)$.
5. If $t = 0$, sample θ^{**} from $\pi(\theta(m^*))$. If $t > 0$, sample θ^* from the previous population $\{\theta(m^*)_{t-1}\}$ with weights $w(m^*)_{t-1}$, and set peturbation kernel $\theta^{**} \sim K_t(\theta|\theta^*)$.
6. If $\pi(\theta^{**}) = 0$, go to 4.
7. Simulate N_p instances of root growth using θ^{**}, recording the state of structure i at each of the $N_t(i)$ time points corresponding to an experimental observation.
8. Compute ρ using Eqn. 2.4 above.
9. If $\rho \geq \mathbf{E}[t]$, go to 4.
10. Set $m_t^{(i)} = m^*$ and add θ^{**} to the population $\{\theta(m^*)_t\}$. If $t = 0$, set weights $w_t^{(i)} = 0$, otherwise

$$w_t^{(i)} = \frac{\pi(\theta^{**})}{\sum_{j=1}^{N} w_{t-1}^{(j)} K_t(\theta_{t-1}^{(j)}, \theta^{**})}. \tag{2.2}$$

    If $i < N$, set $i = i + 1$, go to 4.
11. For every m, normalise the weights. If $t < T$, set $t = t + 1$, go to 3.

For a single model, the prior $\pi(m)$ associated with that model is unity and the choice of model indicator m^* plays no role in the inference process.

We used uniform priors over all model structures for $\pi(m)$ and uniform priors between 0-1.4 day^{-1} for g, 0-1.4 day^{-1} for b, 0-1 for α and 5-40cm for l_{max}. The perturbation kernel we used was $K_t \sim N(0, 0.1P)$, where P is the width of the uniform prior. The tolerance vector was $\mathbf{E} = \epsilon\{5, 3, 2, 1.5, 1\}$.

2.2.4 Additional work

In addition to work in this paper, other avenues were investigated and later abandoned, since they were unsuccessful or beyond the scope of this work. Initially, it was planned to create a new root simulation model to apply the ABC framework to. However, a great deal of excellent root models exist and are available to use [Schnepf et al., 2018, Pagès et al., 2004] and producing a model that was an improvement would be beyond the scope of this work. Therefore, a very simple model was used to clearly demonstrate the applicability of the ABC framework and for ease of model comparison.

Further to this, work was done in improving the root simulation output. Initially, the aim was to produce a realistic root output more in line with those obtained from models like CRootBox [Schnepf et al., 2018]. However, as discussed previously in this section, models exist and are available to use, and therefore improving the appearance of the roots would provide little scientific insight. Subsequently, the root output was greatly simplified to clearly demonstrate the summary statistics rather than prioritising the appearance of the root system.

The branching model comparison initially included a wider range of branching models. In addition to the uniform distribution for branching location included in this chapter, model selection was carried out on other distributions, such as using a combination of normal distributions with peaks spaced a distance, δ along the root. However, this work

was abandoned as it complicated the model selection and results, without gaining any particular mechanistic understanding. The model selection included is an illustration of how models can be compared using this framework, rather than making any claims about the best branching or growth model currently widely used in the literature.

2.3 Results

2.3.1 An ABC SMC framework for inferring mechanistic parameters from root systems

For generality, we begin by considering a highly simplified model for root growth (see Methods). Starting from an infinitesimal initial condition, a primary root elongates according to a growth law. Branches from this primary root occur stochastically according to a branching law. Branches elongate according to the primary root growth law multiplicatively scaled (allowing, for example, branches to grow at a slower rate than the primary root). Branching is for now restricted to first-order branches from the primary root, though nothing in our framework is dependent on this or any other structural choice.

This coarse-grained model was chosen to reflect the core behaviour shared at the intersection of several contemporary root models [Pagès et al., 2004, Leitner et al., 2010, Schnepf et al., 2018]. Its computational simplicity is an advantage but not a necessity for our inference framework; we later consider an alternative generative model to demonstrate the transferrability of our approach. The details of the model are described in Methods, but in this section we consider constructing the inference framework for a general mechanistic model, the parameters of which we denote θ. Our results mainly consider mechanistic models as we aim to capture the underlying processing contributing to the growth of a root system.

The platform proceeds by simulating outputs from this model with different trial parameterisations, using a distance function to compare these outputs to summary statistics of experimental observations, and iterating this process within a Bayesian framework to build up posterior distributions on model structures and parameters given the observed data.

To compare simulation to experiment, we focus on a mechanistically informative set of summary statistics. For a given root structure d, these are the number of branches B, length of the primary root L, and average length of the lateral roots $\hat{l}$. Within our scheme, these lengths are for convenience measured in cm, but different scalings of these features can be used to emphasise different aspects of root architecture in the simulation-data comparison. The distance function we use to compare two structures d_1, d_2 is similar to the Euclidean distance:

$$\rho(d_1, d_2) = \frac{1}{3}\left((B_1 - B_2)^2 + (L_1 - L_2)^2 + (\hat{l}_1 - \hat{l}_2)^2\right). \tag{2.3}$$

This was chosen to give equal weight to each component and as a straightforward distance measure.

A dataset $\mathcal{D}$ may consist of a set of structures d_{ij}, where i labels individual plants and j labels longitudinal observations. In this case, for each observed plant i, a model plant is simulated and its structure recorded at each of the times corresponding to the longitudinal observations. We will call these recorded structures d'_{ij} and are interested in the comparison between each recorded structure and its observed counterpart:

$$\rho(\mathcal{D}, \mathcal{D}') = \sum_i \sum_{j|i} \rho(d_{ij}, d'_{ij}). \tag{2.4}$$

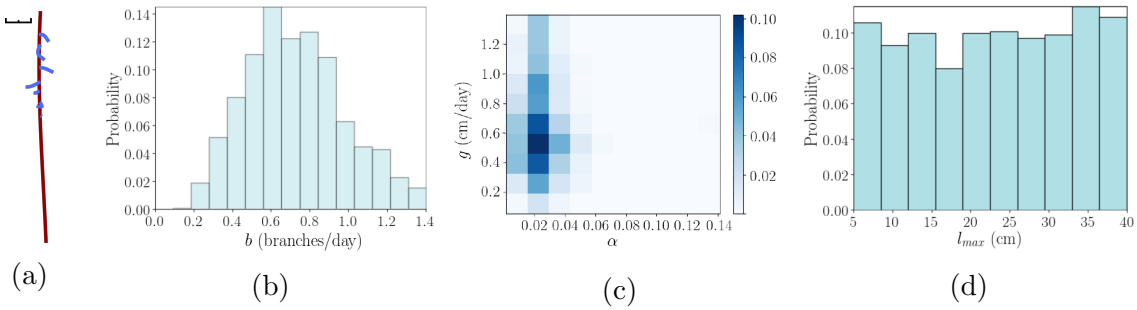
We deliberately choose this model and summary statistics to focus on the topological aspects of root architecture and ignore any specific physical embeddings (for example,

branching angles). This focus on topological degrees of freedom increases the generality of the approach, but features like branching angles and higher-order topological statistics can readily be included in the modelling and distance calculation if they reflect important degrees of freedom for the scientific question under consideration.

Eqn. 2.3 balances the ability to capture fine detail of the root system against the computational time required to obtain a reasonable number of samples from the posterior. Including more detail and/or degrees of freedom in the distance function will allow more detailed matching of observations, but will increase the sampling effort required to find regions of parameter space that match these criteria.

ABC involves accepting a trial set of parameters as a sample from the posterior distribution when $\rho(\mathcal{D}, \mathcal{D}') < \epsilon$, or in other words when the summary statistics of the structure emerging from simulation are 'close' to those arising from the experimental data. The posterior distribution on parameters θ built up from a set of samples taken in this way is $P(\theta|\mathcal{D}; \rho < \epsilon)$, which forms an increasingly good approximation to the true posterior $P(\theta|\mathcal{D})$ as ϵ is decreased [Csilléry et al., 2010].

For parametric inference within a fixed model, a simple rejection-sampling pipeline is then given by Algorithm 1 (see Methods). This approach would be sufficient to identify generative parameters from data, but rejection sampling is an inefficient paradigm, as any 'good' regions of parameter space are immediately forgotten when the next draw from the prior is made. To facilitate more efficient parametric inference as well as model selection, we use ABC embedded in a sequential Monte Carlo (SMC) framework as in Toni *et al.* [Toni et al., 2009]. ABC SMC first enforces only a relaxed fit to the data then sequentially uses the inferred parameter distributions as priors while enforcing a tighter fit to data. This sequential process is parameterised by a sequence of ϵ values describing the fit threshold required at each step in the sequence. Model selection can proceed by including a 'model index' parameter describing which model structure is to be used, applying a prior to this parameter (thus incorporating prior knowledge about which



(a) (b) (c) (d)

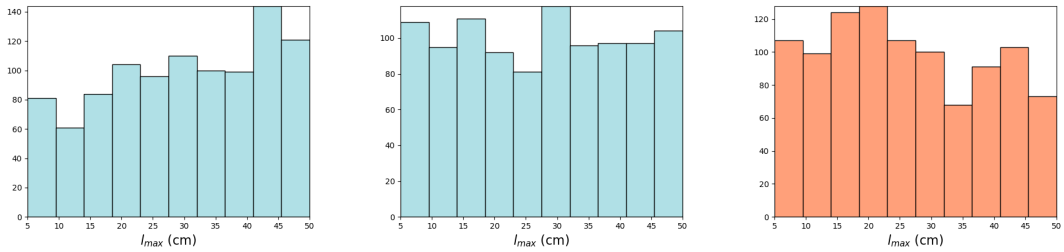
Figure 2.1: **Validating root inference platform with synthetic *Arabidopsis* data.** (a) Output from CRootBox simulation of *Arabidopsis* root growth; black scale bar is 1cm. (b-d) Output posteriors from an ABC SMC framework run on CRootBox output, with final ABC SMC tolerance $\epsilon = 0.5$ (see Methods). (b) Posterior distribution on branching rate b in the growth model. (c) Two-dimensional posterior on primary root growth rate g and lateral root growth scaling α. (d) Posterior distribution on l_{max}, the maximum length parameter in the negative-exponential growth model used.

model structures are more likely), then treating this index as a parameter to be inferred through SMC. Following Toni *et al.* [Toni et al., 2009], the coupled inference and model selection pipeline is then given by Algorithm 2 (see Methods).

2.3.2 Inferring parameters from a simulated root system

We first sought to test the applicability of our likelihood-free inference process on syntic root data, to confirm its ability to identify known generative parameters. To this end, the CRootBox root simulation model [Schnepf et al., 2018] was used to produce an example of an *Arabidopsis thaliana* root system. The governing parameters were mean growth rates of 0.49 cm day^{-1} for the primary root, 0.08 cm day^{-1} for the lateral roots, and an inter-lateral distance of 0.2cm, although inter-lateral distance is not an explicit parameter in our model (see next section). CRootBox adds an element of stochasticity to its generative parameters; in the default *Arabidopsis* case this corresponded to a coefficient of variation of 0.1 in the growth rates and 0.45 in the lateral spacing parameters. The simulation was run over 15 simulated days, yielding the structure in Fig. 2.1a).

To mirror the pipeline that will be used for experimental data, we analysed this simulation output with SmartRoot image analysis software [Lobet et al., 2011], obtaining the



(a) (b) (c)

Figure 2.2: **Posteriors on l_{max} for *Arabidopsis* wild-type seedlings** (a) from the initial pipeline described in section 2.3.4, and the wildtype (b) and *friendly* (c) comparison from section 2.3.6.

statistics of tap and lateral root length and placement. We then applied our ABC SMC framework to estimate posterior distributions on the mechanistic parameters of our simple growth model (see Methods). These parameters are g (primary root growth rate), l_{max} (primary root scaling constant), b (branching rate), and α (lateral root growth scaling).

As shown in Fig. 2.1b)-d), the growth rates and, notably, their variability are well captured in the resultant posteriors, with g inferred to lie around 0.55 ± 0.10 cm day^{-1}, compatible with the true growth rate parameterising the synthetic data. The branching rate parameter is more broadly spread, with a mean of 0.6 day^{-1} corresponding to the observed number of branches, and flexibility in the posterior reflecting the stochastic nature of this parameter's influence. α was inferred to lie around 0.021 ± 0.02, corresponding to a lateral growth rate around 0-0.04 cm day^{-1}; this is rather lower than the value used in the simulation, reflecting the rather limited lateral growth occurring in the specific simulated instance of the model. The posterior for l_{max} is close to recovering the prior which suggests that the model output is minimally dependent on the value of this parameter. We found this limited l_{max} dependence to generally be the case, and in subsequent sections will omit l_{max} from the posterior plots; all l_{max} posteriors, generally recovering priors, are plotted in Figure 2.2. This assessment of the relative importance of, and flexibility in, generative mechanistic parameters reflects a powerful aspect of this inverse modelling approach.

2.3.3 Inferring mechanistic parameters for other synthetic phenotypes and root simulation models

To test the wider applicability of our likelihood-free inference process, we next tested the ability to identfy known generative parameters when using a different, existing root simulation model, and for different plant species. RootBox [Leitner et al., 2010] was chosen for its wide application in the field. We embedded RootBox as the generative model in our inference framework, which was then applied as in section 2.3.2 to the previous synthetic *Arabidopsis thaliana* data and a simulated *Lupinus angustifolius* root system. The *Lupinus* simulation involved an initial growth rate of 1cm day^{-1} for the primary root, 0.2 cm day^{-1} for the laterals and an inter-root distance of 0.9cm, and proceeded for 15 simulation days.

RootBox employs a different branching protocol from our simple model above. Rather than allowing stochastic branching anywhere on the primary root, RootBox allows lateral branches to emerge at specified intervals along the primary structure. This interval d, and a value b_{max} governing the maximum number of allowed lateral branches, are parameters of the model and we therefore seek posterior distributions on these quantities as well as the other mechanistic parameters which directly map to those in our simple model.

Fig. 2.3 shows the resultant posteriors after applying our inference approach using RootBox as the core mechanistic model. Once more, the original generative parameters are well supported by the resulting posteriors, which also agree with the inferred values for primary and lateral growth rates using our simplified core model above (Fig. 2.1). The inter-lateral distances d, present in RootBox but not above, are also well recovered by the inference process. The maximum number of branches b_{max} is not tightly constrained

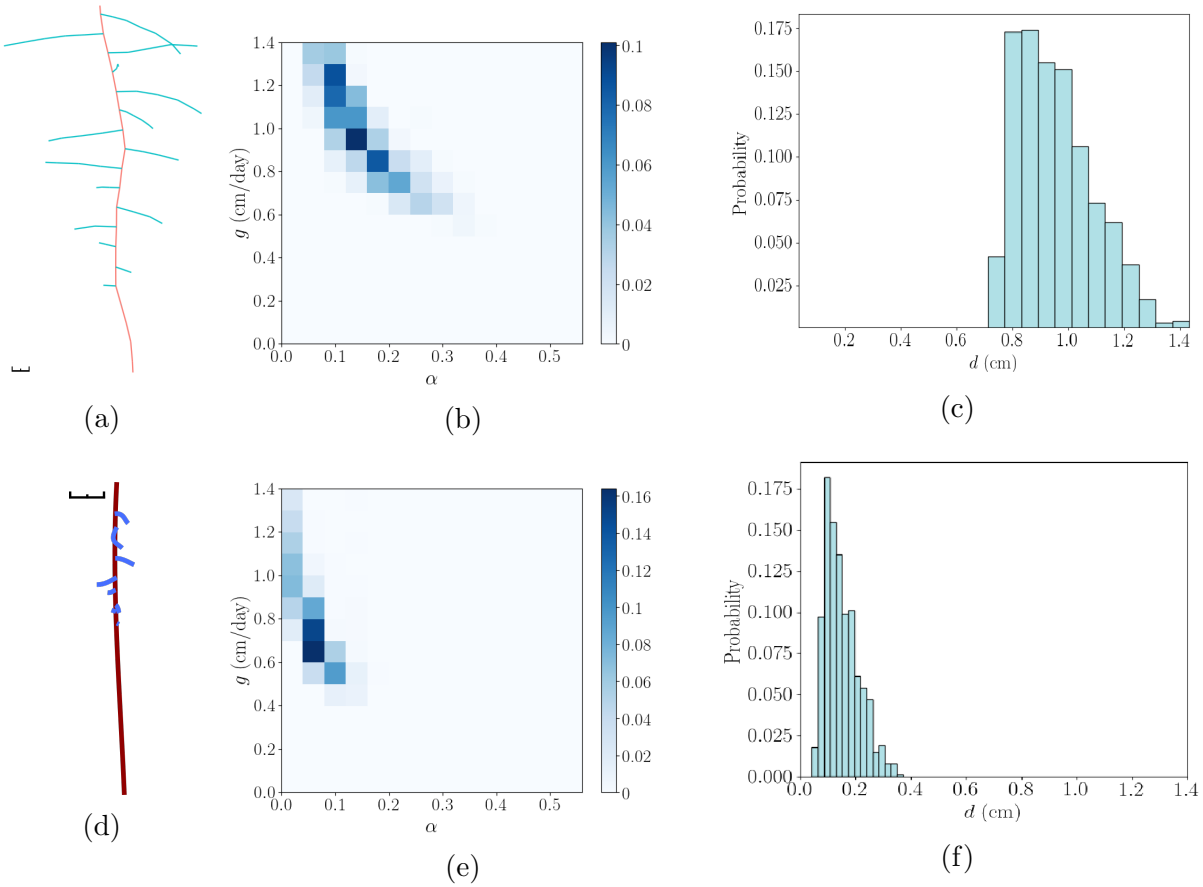
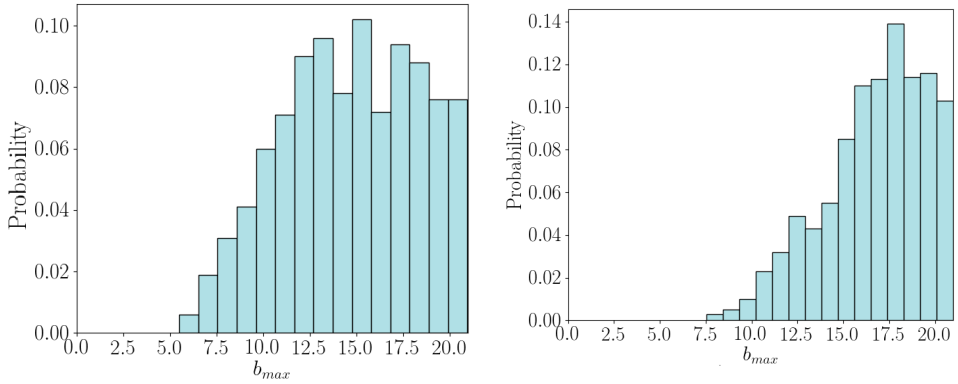


Figure 2.3: **Posterior distributions on *Arabidopsis thaliana* and *Lupinus angustifolius* roots generated using RootBox.** (a) Output from RootBox simulation of *Lupinus angustifolius* root growth; black scale bar is 1cm. (b-c) Output posteriors from an ABC SMC framework run on RootBox output of *Lupinus angustifolius*, with final ABC SMC tolerance $\epsilon = 0.4$ (see Methods). (b) Two-dimensional posterior on primary root growth rate g and lateral root growth scaling α. (c) Posterior distribution on branch separation d in the RootBox growth model. (d) Output from RootBox simulation of *Arabidopsis thaliana* root growth; black scale bar is 1cm. (e-f) Output posteriors from an ABC SMC framework run on RootBox output of *Arabidopsis thaliana*, with final ABC SMC tolerance $\epsilon = 0.4$ (see Methods). (e) Two-dimensional posterior on primary root growth rate g and lateral root growth scaling α. (f) Posterior distribution on branch separation d in the RootBox growth model.



(a) (b)

Figure 2.4: **Posterior distributions on b_{max} for a) *Arabidopsis thaliana* and b) *Lupinus angustifolius* roots generated using RootBox.**

by the synthetic data, as can be seen in Figure B.11).

2.3.4 Inferring mechanistic parameters for wildtype *Arabidopsis thaliana* root systems

To test the pipeline on experimental data, we grew *Arabidopsis* Col-0 plants on vertical $\frac{1}{2}$ MS agar plates (see Methods) and used a digital camera to capture their root system structure over several days. We used SmartRoot image analysis software [Lobet et al., 2011] to extract the lengths and placements of tap and lateral roots from these digital images at each sampled timepoint. An example of the digitised data is shown in Fig. 2.5.

We applied our ABC SMC framework to estimate the posterior distributions of the mechanistic parameters underlying the development of these root systems. The earlier populations of the SMC process gave a diverse range of simulated root structures; by the final population, the simulation outputs provide excellent visual matches to the observed experimental structures (Fig. 2.5) given the deliberate simplicity of the model. This intuitive snapshot matching is supported by the good agreement between the experimentally observed time series of summary statistics and those arising from simulation with the final posteriors (Fig. 2.6). Here, both the mean and the variability in the experimental

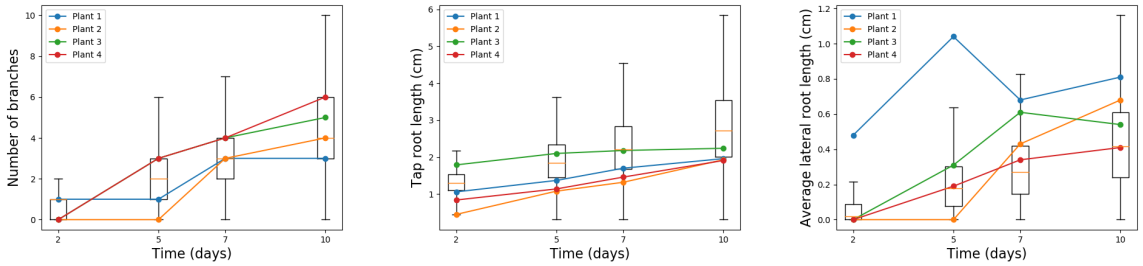
(a) (b) (c)

Figure 2.5: **Example data and simulation output for the root inference framework.** (a) *Arabidopsis* seedlings grown vertically on agar (see Methods) provide example root systems for the analysis pipeline. (b) Digitisation of the root systems using SmartRoot [Lobet et al., 2011] provide the quantitative data used in the inference process. (c) Example outputs from the stochastic growth model with parameters identified through the ABC SMC inference process. Black scale bar is 1cm.

statistics over time is captured by the distributions of simulated behaviour arising from the posteriors.

The posterior distributions themselves are shown in Fig. 2.7. The primary root growth rate g is reasonably well constrained, with a mean that intuitively falls around the total growth average. Notably, the posterior distribution on g is tighter than for the synthetic data example. This refinement reflects the strength of including time-course data in the inference platform. Observations of systems at different times provide more information on dynamic rate parameters, allowing better estimates than are available from single-instance observations alone.

The scaling of lateral growth rate α has a broader variance, reflecting the greater variability in average lateral root length observed in the data, and is correlated to some extent, as expected, with the value of g. The distribution of branching rate b is broader, reflecting a greater variability in the experimental observation of branch number over time, and also the stochastic nature of this process: as b reflects the mean rate of a Poisson process, the same branching structure can be achieved with a variety of different b values. The modal value of b matches the average branching rate observed in the data. Overall, therefore, the ABC SMC framework gives reliable and intuitive readouts linked



(a) (b) (c)

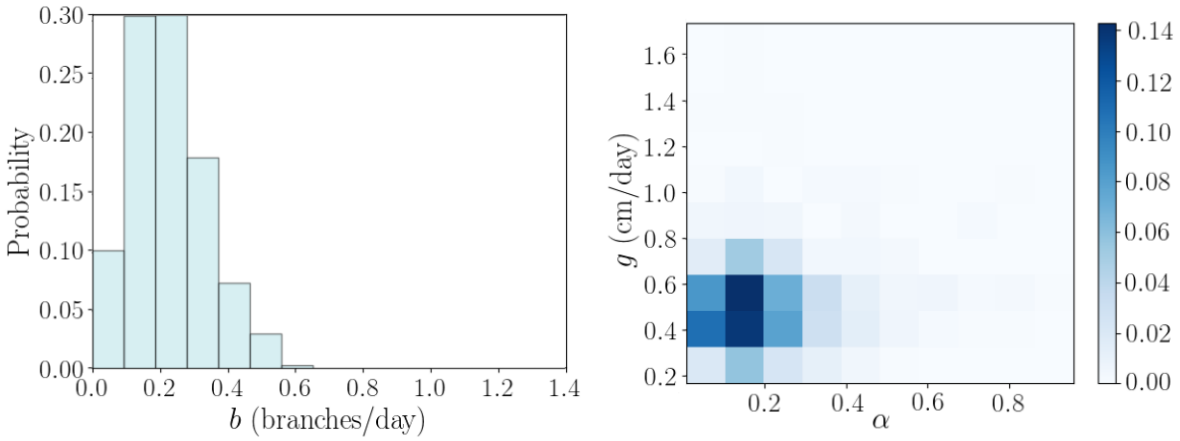
Figure 2.6: **Summary statistic comparison between data and parameterised model.** Individual line traces in each plot show time series of the summary statistic from observed *Arabidopsis* seedlings in Fig. 2.5; boxplots give the range of values arising from stochastic model simulation after parameter values have been learned. (a) Number of branch points; (b) primary root length; (c) average lateral root length.

to both the average observed behaviour and plant-to-plant variability in root structure.

2.3.5 Model selection for root growth and branching mechanisms

We next asked whether our approach could select between competing generative models, given time course data on the evolution of a root system. To this end, we considered a range of possible generative mechanisms for root growth and branching, as summarised in Table 2.1. We will employ uniform priors over competing models, reflecting the fact that before any observations are made, we have no belief that one mechanism is more likely than another. This prior belief can of course be arbitrarily changed within our Bayesian framework to reflect prior information. We then use our ABC SMC framework to identify the posterior support for each mechanistic model, given the observed data [Toni et al., 2009] (see Methods).

First, we consider different elongation laws for root growth. The first model involves root growth at a constant rate such that $l = gt$. This model is chosen for its simplicity, it is clear that it will not fully represent root growth; for example, the growth under this model is unbounded. This is compared with the second model, which involves a negative



(a) (b)

Figure 2.7: **Posteriors from our ABC SMC framework run on the data from figure 2.5 with final tolerance $\epsilon = 2.5$.** (a) Posterior distribution on branching rate b. (b) Two-dimensional posterior on primary root growth rate g and lateral root growth scaling α.

exponential growth law supported by [Schnepf et al., 2018] of the form

$$l(t) = l_{max}\left(1 - e^{-gt/l_{max}}\right), \tag{2.5}$$

where $l(t)$ is the length of the root at time t, parameterised by a rate constant g and scaling constant l_{max}. The posterior distribution over model index through the SMC process is shown in Fig. 2.8a)-b). The most permissive population (highest ϵ) shows less support for the exponential model to ensure model parsimony. As a better fit to the data is required, the support for the exponential model increases until it overcomes the lower weightings due to the additional parameter and is preferentially selected. Next, we explore a more nuanced mechanistic question underlying root architecture. We compared two models for branch placement positions. First, a uniform branching model, where the branching location was chosen at random anywhere along the primary root such that the branching location, x, is given by $x \sim U(0, l(t))$, where $l(t)$ is the length of the root at time t of the branching event occuring. Second, a minimally spaced model, which imposed a distance parameter δ around each existing branch where no further branching could occur. Branching was uniform as with the first model, but if a branch was attempted within this

distance no branch was implemented and the algorithm continues. Fig. 2.8b)-c) shows the model selection posteriors with decreasing tolerance, and the posterior on δ when the minimally spaced model was implemented. Here, the posteriors for the spacing model are lower for the more permissive populations, reflecting the increased model complexity – the extra parameter δ makes the model less parsimonious. The support for the model increases as a better fit to data is required, in the subsequent populations. By the final tolerance, the models have comparable support. Hence, the dataset suggests roughly equal support for both models despite their difference in complexity.

Model name	Model type	Description
Linear	Growth	Constant growth rate.
Exponential	Growth	Negative exponential growth rate as used in CRootBox [Schnepf et al., 2018], see equation 2.5.
No restriction	Branching	Branching can occur anywhere along the tap root.
Minimum spacing	Branching	No branching can occur within a distance, δ, of an existing branch, otherwise unrestricted as above.

Table 2.1: A summary of the growth and branching models used for model selection. Model selection posteriors are shown in Figure 2.8.

These simple experiments serve to illustrate the ability of ABC SMC to provide statistical support for competing mechanistic hypotheses (for example, linear versus negative-exponential root elongation laws). There is, however, nothing to prevent other targetted mechanistic questions being addressed using this framework (see Discussion).

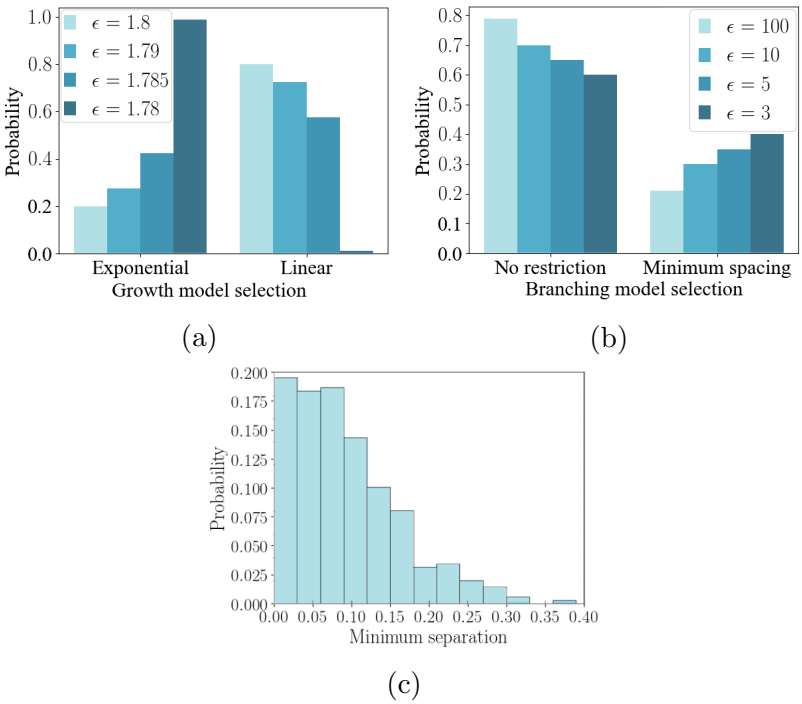


Figure 2.8: **ABC SMC allows selection of competing mechanistic models for root growth.** (a) Model selection posteriors from comparing a simple, constant-rate growth model to the negative exponential model used in CRootBox [Schnepf et al., 2018] for decreasing ABC tolerance ϵ. Tolerances are chosen to illustrate the convergence so values will depend on the models included. The negative exponential model acquires greater support as the tolerance decreases. (b) Model selection posteriors from comparing a uniform branching model to a model imposing a minimum separation from existing branches. The more parsimonious uniform model experiences higher support with relaxed tolerance but the minimum-separation model gains support with tighter tolerances (tolerances are higher for this model selection run, reflecting a tradeoff with the greater computational resource required). (c) Posterior for the minimum branch separation, δ, required when the minimum spacing model is implemented.

2.3.6 Comparison between root structures

Next, we asked whether the ABC SMC framework could distinguish between two phenotypes – those corresponding to wild-type *Arabidopsis* and the *friendly* mutant line. *FRIENDLY* is a mitochondrial fusion gene that when compromised has a range of bioenergetic effects which lead to reduced root growth [El Zawily et al., 2014].

Wildtype and *friendly* plants were grown under the same conditions as above (see Methods), and the inference pipeline was run as before, with exponential growth and uniform branching. The output posteriors in Figure 2.9 reflect the differing root systems shown in the tracings, with a clear separation in the parameter space between the two phenotypes. The branching rate b is fairly unconstrained as observed in section 2.3.4 due to inherent stochasticity in the branching mechanism. The values of g vary significantly between wildtype and *friendly*, as reflected in the tracings, with little change in the value of α. The distribution of g is substantially shifted towards lower values for the *friendly* plants, reflecting the known challenge to root growth resulting from this mutation. l_{max} shows a wide variability in both phenotypes, while representing clear differences consistent with the reduced root growth observed in the *friendly* mutant line. There is very little constraint in the value of l_{max}, suggesting little reliance on the value, although smaller values appear to be favoured for the *friendly* phenotype. This may be because we are comparing seedlings, and so the limiter l_{max} is not reached in either case.

Taken together, these results demonstrate that the physical parameters governing root architecture growth can be learned using this ABC SMC approach, and uncertainty in these learned outcomes quantified. The mechanistic model within our inference process both allows us to harness time-course data and dissects which parameters change (here, growth rate g) and which remain similar (here, lateral root scaling α) in different cases. This is a key strength of this method, and would be difficult to obtain through traditional parameter inference methods. The platform readily identifies the physically different

mechanisms underlying root architecture in a mutant line, and identifies accepted physical model structures for root growth.

2.4 Discussion

We have presented a framework for the inference of parameter values and mechanisms in root growth models when applied to an observed root system. While there has been much work undertaken producing plant root simulations from given parameters, our approach addressed the much less-studied 'inverse problem': that of finding generative parameter values and mechanisms that can reproduce a given root system. Knowledge about these mechanisms and parameters, and their ranges, flexibility, and relative importance, is necessary for an understanding of root growth processes such as growth and branching decisions, and how these may relate to biological processes within the plant. We hope that this highly general approach will allow for a more mechanistic understanding of root growth, and to quantify the efficacy of existing models.

We first used a very general growth model to (a) retain consistency with the 'core' of the maximum possible number of existing growth models, and (b) focus on parameters related to the growing plant and its phenotype, rather than the specifics of its physical embedding. We have demonstrated using RootBox [Leitner et al., 2010] that our approach can readily be adapted to other specific existing root models to allow the quantification of values of and uncertainty in generative parameters, furthering understanding of root system architecture. We also illustrated how alternative hypothesised mechanistic models can straightforwardly be compared, using SMC model selection. A strength of the Bayesian embedding here is that the most parsimonious model that is capable of explaining observations is naturally selected in the case of models with different numbers of parameters [Toni et al., 2009].

Advances in imaging techniques are allowing for greater insight into root system archi-

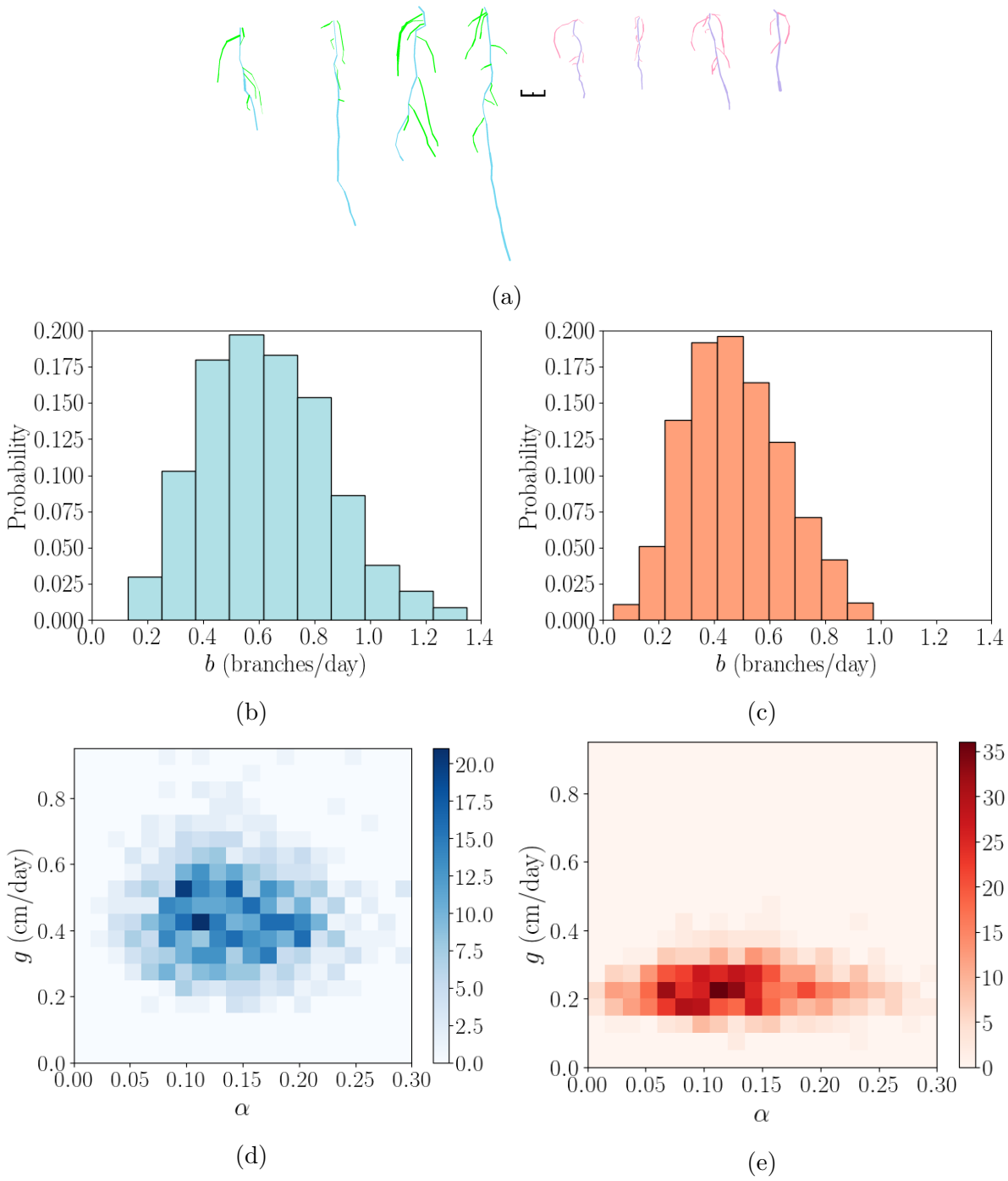


Figure 2.9: **Distinguishing phenotypes with mechanistic inference.** (a) (left) Wildtype and (right) *friendly Arabidopsis* seedlings grown in agar as described in section 2.2.1, demonstrating the root growth phenotype of *friendly*. The tracings were produced using SmartRoot software [Lobet et al., 2011] and the colours adjusted; black scale bar is 1cm. (b-c) Output posteriors for branching rate b show similar distributions for wildtype and *friendly*. (d-e) Two-dimensional posteriors on primary root growth rate g and lateral root growth scaling α demonstrate clear separation, reflecting the reduced root growth observed in *friendly*.

tecture [Roose et al., 2016], and specially designed image analysis software [Lobet et al., 2013] allows for increasingly efficient data collection from images. The combination of root models, advanced imaging techniques, image analysis software and an SMC framework could allow further advances in our understanding of root growth. We anticipate that, with the increasing developments in root imaging, this technique will find application in a growing variety of datasets, allowing for investigation of generative parameters for a wide variety of root system phenotypes. A natural future extension for this work would be to perform inference based directly on image data, rather than statistics of these data. This approach would require simulation of the imaging process as well as the generation of model root systems, for example, embedding the idealised root system in a simulated soil substrate and simulating the artefacts and noise involved in the imaging process. While (much) more computationally intensive, this approach would allow a more direct leveraging of phenotype data from experimental studies.

Notably, our approach allows inference based on time-course measurements of a developing root system, which increase the power and precision with which parameters and mechanisms can be identified. As demonstrated with our synthetic examples, this approach can readily be applied to single-instance observations, but also naturally leverages dynamic information to refine posterior distributions on physical rate parameters.

A stochastic modelling framework for root growth allows for a wide variety of possible outputs to be considered in the inference process, reflecting the variation between root systems in the real world. In this way, the modelling approach allows for investigation into the underlying mechanisms which are widely applicable, while avoiding a reliance on specificities and overfitting to a particular phenotype or growth environment. Predicting a branching event would require the consideration of processes such as genetics, cellular interactions, and organism-scale resource partitioning [Band et al., 2012], necessitating the development of a multiscale framework. As such multiscale approaches develop, we anticipate the use of likelihood-free inference to be further embraced to resolve inverse problems in parameter identification.

While the generality of our approach is appropriate for the scope of this study, greater specificity is required to gain a true understanding of plant processes. Care needs to be taken in the application of ABC techniques: choices must be made over elements such as the tolerance, priors, and summary statistics to achieve a balance between convergence rate and specificity of results. As specific choices for these values can be hard to interpret, simulation outputs must be verified to provide a reasonable match to genuine behaviour (as we have attempted throughout). We have worked with different models to explore the behaviour of our method under different generative assumptions. In Bayesian model selection, prior beliefs about models can strongly affect their support and interpretation must take this into account [Toni et al., 2009]. For example, when comparing the exponential and linear models, it should be noted that the models are going to function very similarly for early stages of root growth as represented here. However, we have aimed to demonstrate the strength of this approach when carefully applied and interpreted.

Overall, we have demonstrated a technique to allow for greater insight into model parameters for root systems, which could aid in increasing understanding of root growth mechanisms. The generalised approach allowed for investigation of the key aspects underlying root topology, while being highly adaptable for use with existing root architecture models.

Chapter 3

Characterising root processes at the BIFoR FACE site

3.1 Introduction

We now complement the theory in the previous chapter with a large-scale fieldwork analysis of root growth under elevated CO_2.This chapter details experimental work conducted at the BIFoR research site, and the results from this work. The experiment was set up with the null hypothesis that elevated carbon dioxide would have no effect on fine root growth in a temperate oak forest, and the alternative hypothesis that there will be an observable carbon fertilisation effect on fine root growth.

In order to quantify these effects four key metrics were observed; the root length, width, biomass and production. Here, the root length and width is the observed root measurements gained from analysis of images obtained from a minirhizotron or root scan (see methods). The biomass is obtained either by weighing roots extracted from soil cores, or by a conversion of observed root geometry to biomass from obtaining root density measurements. Production represents the amount of new root matter observed since the

last sampling. There is an important distiction here between biomass and production: biomass represents the roots observed, and production only the new roots that have grown since last observation. In this way, a root that has undergone no growth will be observed as a measure of biomass but not as production.

The experimental work conducted here represents the first FACE study of this scale of fine roots in a temperate oak forest, which will complement existing work on forestry systems around the world (see Table 3.1). As such, this chapter will detail a range of experimental methods for fine root observation followed by a more detailed introduction to the methods implemented as those most suited to the BIFoR FACE setup. This will be followed by the results from the experimental work performed.

3.1.1 Experimental observation of fine root dynamics under elevated carbon dioxide

Non-destructive observation of below-ground systems such as roots continues to be difficult. Although a powerful experiment, BIFoR offers a particular challenge due to its situation in a mature oak woodland, studying plants within a delicate preexisting ecosystem rather than a plantation as with many previous FACE experiments [Norton et al., 2008] [Hendrey et al., 1999]. Therefore, data collection methods need to be as non-destructive as possible. Minirhizotrons allow a small subset of a root system to be observed over time [Johnson et al., 2001], but provide little information regarding root biomass and by nature obstruct the natural structure of the root system [Taylor et al., 1990]. Soil cores are a more destructive sampling method and do not allow long-timescale observation, but allow for estimations of root biomass and turnover [Samson and Sinclair, 1994]. Both methods have been successfully employed in previous studies of fine root growth [Iversen, 2010].

Previous experiments allow for some predictions to be made about the effect of eCO_2 on root systems [Matamala et al., 2003], but are limited by their timespan, experimental

setup and type of plant studied. There is a general consensus that eCO_2 leads to an initial increase in carbon capture and growth in both above-ground [Norby et al., 2002] [Thilakarathne et al., 2015] and below-ground tissues [Allen et al., 2000]. This initial growth increase appears to be limited by nutrient availability [Johnson et al., 2004] [Norby et al., 2010] [Rogers et al., 1999], with many short-term experiments failing to capture this effect [Madhu and Hatfield, 2013] which may only be visible in experiments run over many years [Norby et al., 2010].

Increased growth under elevated carbon dioxide leads to a greater dependence on soil nutrient availability [Fernando et al., 2014], which is particularly important in agricultural studies due to decreased nutritional value in crops [Myers et al., 2014]. Nitrogen availability appears to be of particular importance in long-term growth response [Iversen, 2010], and may become a limiting factor in non-fertilised soils. Tight coupling between nitrogen levels and phosphate mobility in soils [Marklein and Houlton, 2012] may also lead to growth becoming phosphate limited [Edwards et al., 2006]. However, it is important to note that variations in plant species and growth environments may limit the application of these results to the UK temperate oak forest observed in this study [Lau et al., 2010].

It is clear that eCO_2 affects many key plant processes [Leakey et al., 2009], necessitating consideration of the entire plant when making root growth predictions. Lack of stomatal acclimation leads to greater water use efficiency [Ainsworth and Rogers, 2007], while photosynthetic activity is shown to increase in C3 plants [Leakey et al., 2009]. Any study of below-ground systems will therefore be greatly strengthened by consideration of results from other research conducted at BIFoR FACE.

3.1.2 BIFoR and other FACE experiments

One class of experiments aiming to elucidate the intersection between biosphere and carbon flow is Free Air Carbon Enrichment (FACE) experiments. FACE experiments aim

to understand the biosphere's response to elevated carbon dioxide, by artificially raising the level of carbon dioxide in a concentrated area, outside of a laboratory setting. Rather than enclosed growth chamber experiments, FACE experiments can be installed outside in a plantation or natural woodland allowing the study of plants *in situ*. FACE experiments have been performed around the world [Norby and Zak, 2011], on crops [Long et al., 2006], on forestry stands [Schlesinger et al., 2006, Norby et al., 2001], and on forests [Medlyn et al., 2016, Hart et al., 2019, Pepin and Körner, 2002].

The use of carbon dioxide for crop fertilisation has been established for 200 years as an agricultural practice, using the decomposition of biological matter [Allen Jr, 1992]. The 1960's saw the introduction of aerosolised CO_2 for this purpose as it became economically viable [Wittwer and Robb, 1964]. Following the progression of technology and increasing interest in climate change research, FACE experiments were performed in the 1990's on crops, forestry and grassland [Ainsworth and Long, 2005], with early tree-focused FACE experiments carried out on saplings grown in open-top chambers [Norby et al., 1999]. Duke FACE was the first FACE experiment of the type discussed in this chapter, beginning fumigation of multiple experimental plots in 1996 [McLeod and Long, 1999, Schlesinger et al., 2006].

Free Air Carbon Enrichment (FACE) experiments have been set up around the world to study plant responses to elevated carbon dioxide (eCO_2) [Nowak et al., 2004], as detailed in Table 3.1. These experiments provided valuable research in this field, but further experiments were required to cover the broad range of forestry types across the globe [Hickler et al., 2008]. This inspired the second generation of FACE experiments [Norby et al., 2016], including BIFoR and EucFACE which are currently operating [Medlyn et al., 2016, Hart et al., 2019]. In addition to increasing our understanding of the world's forests, BIFoR also provides an opportunity to study a varied forest ecosystem rather than a plantation.

Experiment name	Location	CO_2 elevation	Study area
BIFoR FACE	UK	Ambient + 150ppm	Native oak woodland
EucFACE	Australia	550ppm	Native eucalyptus forest
ORNL FACE	TN, USA	550ppm	Sweetgum plantation
Duke FACE	NC, USA	Ambient +200ppm	Loblolly pine stands
Aspen FACE	WI, USA	Ambient +200ppm	Native Aspen forest plantation

Table 3.1: A selection of key forestry FACE experiments, with their location, CO_2 levels and type of plants studied. Contents adapted from [Laboratory, 1243].

Experimental work in this thesis was conducted at the BIFoR FACE experimental site, which opened in 2016 in a temperate oak forest in Staffordshire [Hart et al., 2019]. The structure is very similar to the Australian EucFACE study, with 3 treatment and 3 control arrays based in native forest [Duursma et al., 2016]. The experimental setup consists of 6 roughly 30m diameter rings of towers which reach above the height of the oak canopy, and 3 similarly sized areas with minimal research infrastructure known as 'Ghost' arrays. The towers emit air which can be treated with additional CO_2, three of the rings are treated to raise the CO_2 level within the arrays to 150ppm above ambient levels, referred to as 'Treatment' arrays. The remaining three 'Control' arrays emit air with no additional CO_2 to act as a control. In this way, the experiment contains two levels of control; three Control arrays with the same setup as the treatment arrays without the additional CO_2, and three Ghost arrays with minimal infrastructure. The FACE site is host to a wide array of research looking at different areas of the forest [Hart et al., 2019].

3.1.3 Observing fine root growth using minirhizotrons

Minirhizotrons are a belowground observation system that allow for the observation of roots *in situ* over a longer timespan while minimising the belowground impact [Bates, 1937]. Clear tubes, often made of plastic [Withington et al., 2003], are installed in the

soil with care taken to prevent soil impaction [Phillips et al., 2000]. After a resettlement period of ideally one year [Johnson et al., 2001], roots growing around the tube can be monitored without further disturbance to the system. Following proper installation and resettlement, the area around the tube is assumed to be representative of bulk soil [Johnson et al., 2001], although this is a topic of some debate [Taylor et al., 1990, Samson and Sinclair, 1994, Majdi, 1996]. Roots are imaged through the use of a specialised periscopic camera system [Vamerali et al., 2012], which can be adapted for multi-spectral imaging [Wang et al., 1995, Pierret, 2008, Rahman et al., 2020]. Minirhizotron camera systems allow for high resolution imaging at small scales through both video and image capture [Johnson and Meyer, 1998], and the tracking of image location using an indexing handle [Ferguson and Smucker, 1989]. A minirhizotron and camera setup can be seen in Figure 3.1 Battery operated cameras are avaliable for use in remote systems [Upchurch and Ritchie, 1984]. Minirhizotron has been aided by advances in camera technology [Vamerali et al., 2012], while a great deal of image analysis softwares have been produced of variable efficacy [Lobet et al., 2013, Lobet, 2017]. However, the collection and analysis of images remains a time consuming and expensive endeavor, particularly in the image analysis as an effective automated solution does not appear to exist. Minirhizotrons can be used to estimate root productivity through the use of time-separated measurements of a growing root element [Norby et al., 2004, Tierney and Fahey, 2001, Majdi, 1996].

3.1.4 Sampling root biomass using soil cores

Sequential soil coring can be used to monitor root production [Makkonen and Helmisaari, 1999]. They continue to provide a far less expensive method of monitoring roots as no camera equipment is required, and are an effective choice when minirhizotrons are cost-prohibitive [Majdi et al., 2005] . Soil cores are a destructive method that allows samples of roots to be directly analysed [Borken et al., 2007]. Soil cores should provide better representation of the bulk soil due to fewer scaling assumptions [Rytter and Rytter,

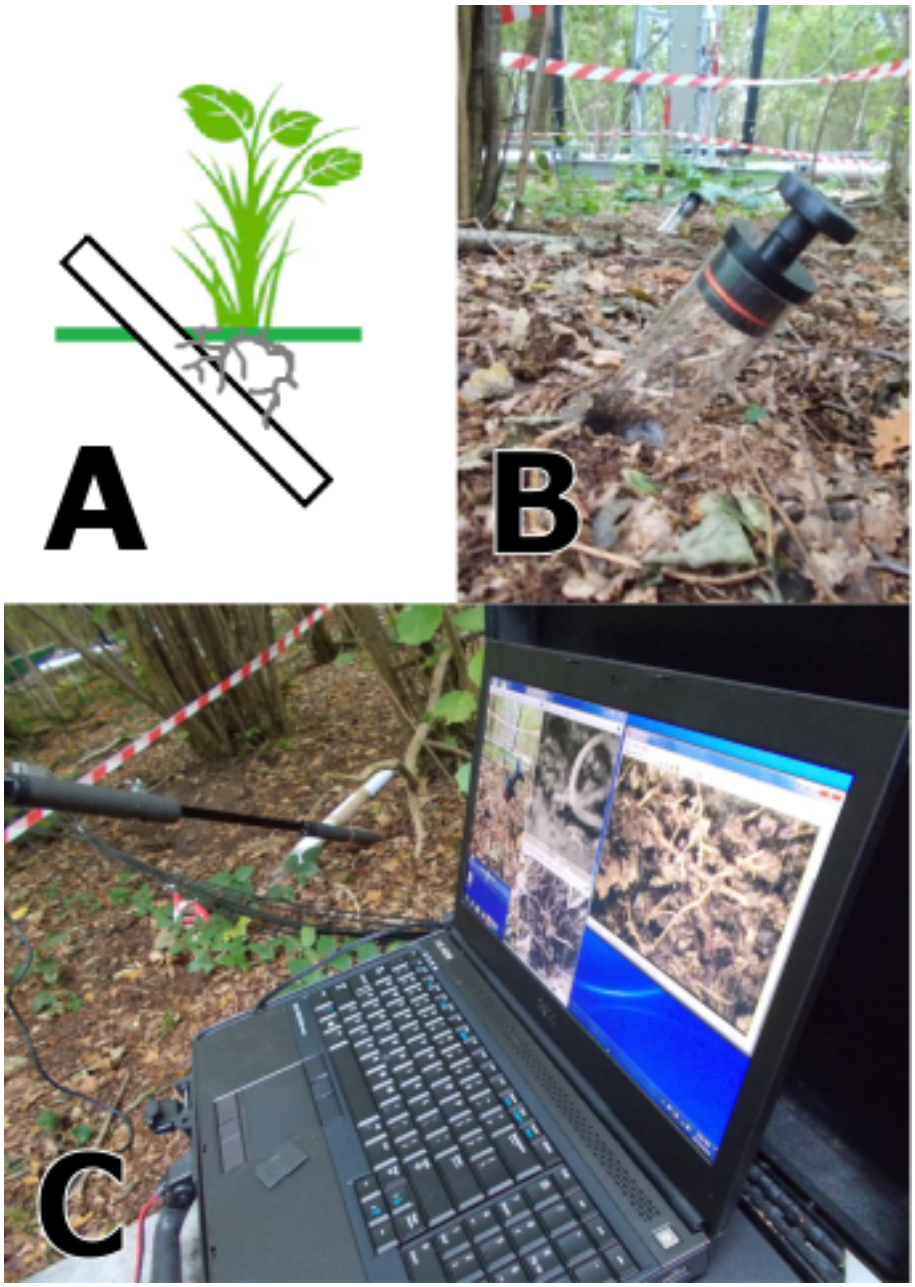


Figure 3.1: **The minirhizotron setup used for fine root observation.** (A) A diagram showing how roots can be observed using an installed tube. (B) A photo of a tube installed at the BIFoR FACE research site. (C) A photo of the camera setup used for imaging at the BIFoR research site. The system is a BTC I-CAP image capture system (Bartz Technology Corporation, Carpinteria, California, US), set up as for root imaging. Example images are visible on the laptop screen.

2012], but only provide a snapshot of the roots at the sampling site and do not allow for the monitoring of specific root segments over time. Due to the destructive nature of soil coring, it can be difficult to obtain enough samples to be confident they represent general trends without causing disturbance to a natural system [Taylor et al., 2013].

Ingrowth cores are a method that can be used to estimate fine root production and biomass [Makkonen and Helmisaari, 1999]. This involves removing a column of soil using a corer, then re-filling the empty column with soil containing no root matter. A core can then be taken at the same point to monitor how roots have grown into the area since the core was taken [Brunner et al., 2013]. Care should be taken to avoid mischaracterising compensatory root proliferation as natural growth [Pritchard et al., 2008]. In addition, ingrowth cores are particularly unreliable when performed in a species-rich ecosystem as there can be significant bias due to a variation in recovery rates between species [Addo-Danso et al., 2016]. For this reason, ingrowth cores are better suited to studies performed on a single species and less so to the complex ecosystem studied at BIFoR FACE. Recreating the soil environment within the core is difficult at sites with defined soil horizons, and also allows for ingrowth with no competition, which is a key element in a natural ecosystem [Lukac, 2012].

3.2 Methods

3.2.1 Experimental setup at BIFoR FACE

Our field observations were carried out at the Birmingham Institute of Forest Research (BIFoR) FACE facility [Hart et al., 2019] in Staffordshire, England. The facility has been built into a native mature deciduous woodland, dominated by oak (*Quercus robur*) interspersed with overstood hazel (*Corylus avellana*) coppice. Sycamore (*Acer pseudplatanus*), hawthorn (*Crataegus*) and holly (*Ilex aquifolium*) have self-seeded into gaps and, with the

hazel, form a distinct sub-canopy. The forest grows on centuries-old Orthic Luvisol soil with a mul-moder humus classification [Hart et al., 2019, Wrb, 2015]. The experimental design consists of three eCO_2 and three ambient-air control arrays (as in the Australian EucFACE study [Duursma et al., 2016]). The arrays are 30m diameter rings, with free-standing towers extending above the ~25m oak canopy.. Pipes attached to the towers emit treated air with increased CO_2, in order to raise the CO_2 levels within the eCO_2 arrays to 150 ppm above the ambient (calculated as the lowest CO_2 mixing ratio measured in the control arrays). The control arrays are identical, but the air released into the array contains no additional CO_2. Performance of the facility was excellent over the course of these experiments [Hart et al., 2019]. Air was emitted (treated or otherwise) during daylight hours from around April to October each year; treatment began with budburst in the spring and was stopped with leaf-fall in autumn.

3.2.2 Minrhizotrons

Minirhizotron installation

To facilitate time-series observations of fine roots *in situ*, four 50cm long minirhizotron tubes sealed at the bottom with a removable bung in the top, were installed in each array in winter 2016. Since the BIFoR arrays had been already paired and there were a small number of tubes being installed, the sites were chosen to keep the surrounding vegetation as consistent as possible rather than using a randomisation technique. A Van Walt 55mm corer was used to remove a cylinder of soil at a 45 degree angle. The minirhizotron tube was then inserted manually into the hole, with care taken to prevent damage to the exterior of the tube. Each set of four were installed in close proximity to each other within the array, and the area fenced with clear markers to minimise footfall. The tubes were then numbered 1-4 for ease of referencing. Following advice, above ground sections of the tubes were covered with black pipe insulation to prevent light from entering the tubes.

Four 50cm tubes were installed in each of the 6 control and treatment arrays, leading to 24 tubes in total. Data was also collected from two 2m tubes previously installed in Arrays 1 and 6 although this data was not used in investigate growth under elevated carbon dioxide, as both 2m tubes are installed in treatment arrays and therefore cannot be compared with control data. Difficulties during installation also led to poor soil contact with the tubes leading to poor quality images, particularly under wet conditions. A further 50cm tube was later installed at the entrance of Array 2 for demonstrational purposes. The tubes were left to settle until sampling began in April 2017.

Root imaging

A BTC ICAP image capture system (Bartz Technology Corporation, Carpinteria, California, US) was used to obtain the root images from the minirhizotrons, mounted on a Smucker manual indexing handle and paired with a BTC I-CAP image capture system. The bung was carefully removed from the end of the tube, with one hand stabilising the tube during removal, and the camera inserted until the outer cuff rested on the top of the tube. The camera was set to point 13 on the indexing handle, which corresponds to the bottom of the tube. The camera was rotated 360°, then moved up to point 12 on the indexing handle and rotated 360°. This process was repeated until points 1-13 on the handle were covered, where point 1 on the handle represented the top of the soil. In the first year roots were photographed and the depth, angle and tube number were recorded. The camera setup was improved for the second year of sampling, allowing for the capture of video. Therefore, in the second year videos were taken of the camera being moved 360° from point 13 up to point 1, covering the entire visible section of the tube. Screenshots were then taken for analysis. The images are then analysed using SmartRoot [Lobet et al., 2011] as described in section 3.2.2 and the amount of growth and branching is recorded through comparison with earlier images. Data is collected monthly from all tubes in all arrays.

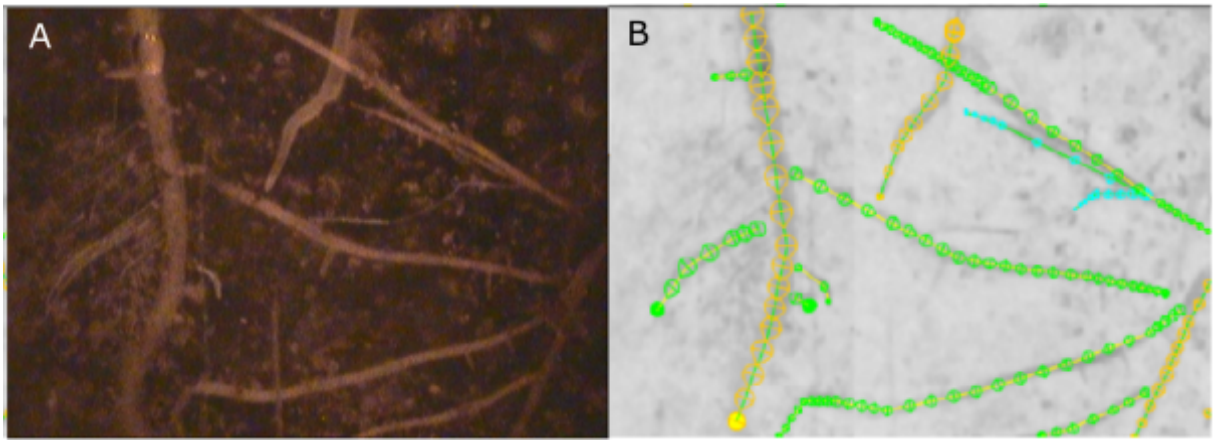


Figure 3.2: **Example of image analysis perfomed on minrhizotron images.** (A) A source image from minirhizotron sampling. (B) The same image following analysis with SmartRoot (see Methods). Each circle represents a root width measurement and the underlying skeleton the root length. Colours denote relative root order.

Image analysis

Image analysis was performed using SmartRoot [Lobet et al., 2011], an ImageJ plugin. ImageJ was installed as part of the Fiji package. Roots were manually identified using the software, and a skeleton was produced with periodic width measurements along the root (see figure 3.2). The length of each root segment was recorded along with an average width across each width measurement point. Each segment was recorded with a length and width measurement in cm, along with the array number, tube number, date of sampling, depth of sampling as taken from the Smucker manual indexing handle (see section 3.2.2), and angle of measurement when available. The image analysis was performed by myself so that any subjectivity in root identification would remain constant across treatment and control populations.

3.2.3 Soil cores

3 plots per array were designated for soil research to minimise impact across the site. For each sampling session soil cores of 3 x 2.5 cm were taken from these areas using a lined Van Walt 55mm corer, then the hole was filled in with sieved soil from outside of the

array and marked with a flag to prevent resampling of the same area. Each filled liner was sealed for transport back to the lab for analysis. The cores were then separated by horizon, and roots were removed by hand for further analysis as described in section 3.2.3

Root biomass from soil cores

Soil cores were taken periodically over the two year study period; in March 2017, March, July, November of 2018, and March 2019. 3 3 x 2.5cm cores were taken from each array at BIFoR using a lined Van Walt 55mm corer. The cores were separated by horizon. Roots were picked from the soil by hand and live and dead roots were separated. Live roots were identified by their elasticity, colour, and lack of evidence of decomposition. Similarly, dead roots would break upon bending, were often hollow inside and were darker in colour [Santantonio and Hermann, 1985]. The separation of live and dead roots was verified using Evans Blue vital stain on a random sample. Roots were then washed using a fine mesh sieve and left to dry before being weighed using a mg scale.

Root scans from soil cores

Roots were picked from soil, separated and washed as described in section 3.2.3. Roots were then blotted with paper towelling to remove excess water, and arranged on a scanner. Care was taken to separate root sections as much as possible, and roots were scanned with a background of graph paper for ease of scaling, to produce images as shown in Figure 3.3. Image analysis was then performed as described in section 3.2.2.

Calculating root density from soil cores

Root scans (see section 3.2.3 for method) were analysed using Smart Root (see section 3.2.2) to obtain the total root volume present in the scan. This number was used along with the weight of that sample (section 3.2.3) to calculate root density. This was repeated

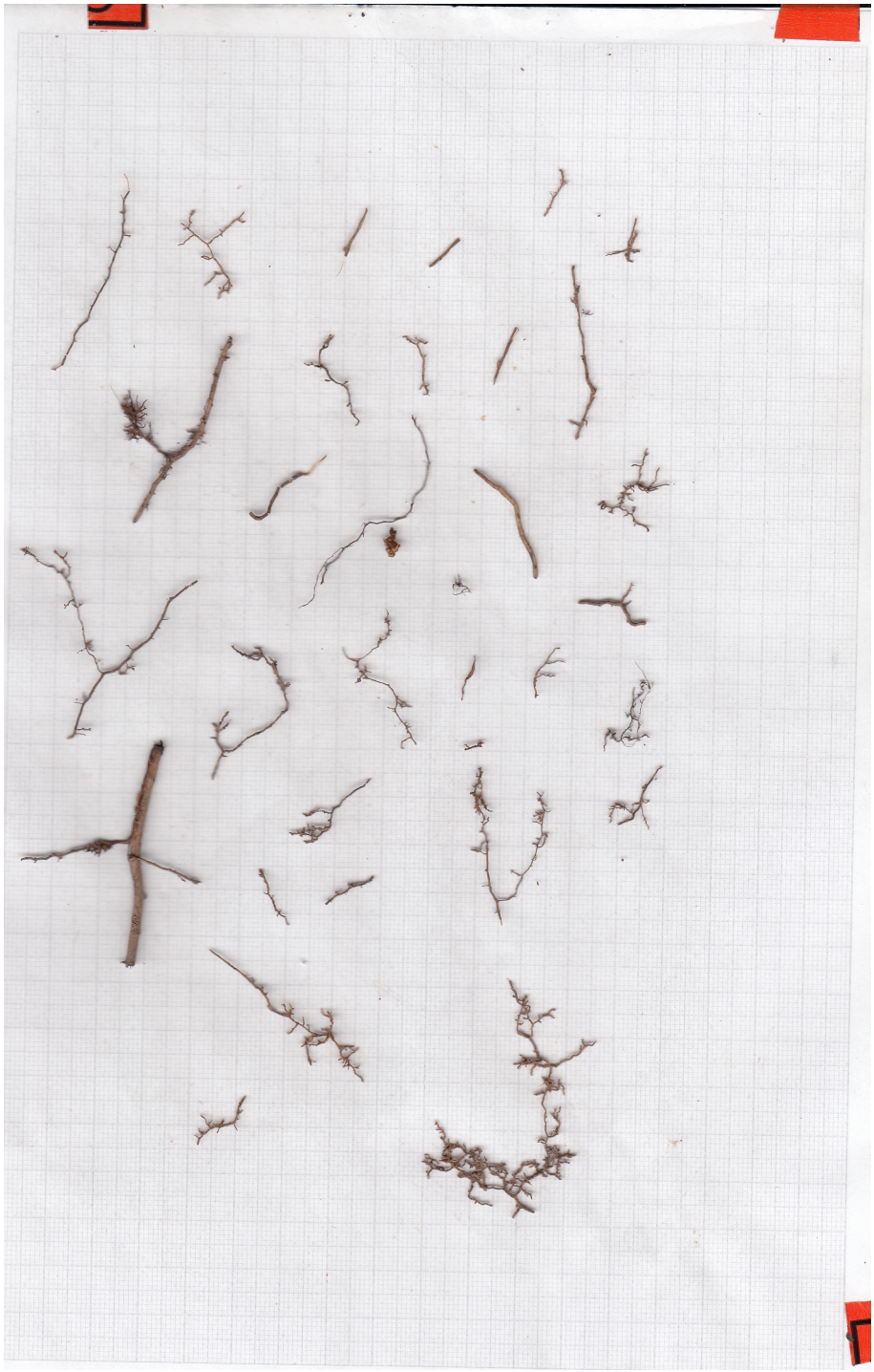


Figure 3.3: **A scanned image of roots from a soil core, as described in section 3.2.3**. Roots have been removed from soil, washed and arranged on graph paper. This image shows roots from the organic horizon of a soil core, which represents the top layer of soil.

for multiple scans, and the density was found to be consistent well within the accuracy of Smart Root. The average density was then used to scale from volume to biomass for root segments observed in minirhizotrons.

3.2.4 Additional methods

Fold-change in root biomass

The fold change in root biomass was calculated by first taking the total biomass per tube data shown in figure 3.10A. The 'earliest mean biomass' was then calculated. This was the average across the 4 tubes in a single array at the time of first sampling (11/04/2017). Therefore, each array had an earliest mean value by which to normalise all the future data. The fold change was found by taking the total biomass per tube data and dividing each point by the earliest mean measurement for the corresponding array. In this way, we obtain a fold change in biomass based on the amount of root matter at first sampling, which would have grown before the ambient carbon dioxide levels were raised in the treatment arrays.

Installation of root viewing windows

Root viewing windows were installed outside the experimental arrays at the BIFoR site as a test of their efficacy for fine root observation. Pits were dug with a shovel to around 30cm deep and 50cm wide, to match the dimensions of 3 perspex sheets to be installed, with one vertical face where the viewing windows were to be installed. This face was made as flat as possible to minimise space between the soil and the viewing window once *in situ*. The sheets were placed against the flat face of the pit, and left for the roots to recover. Canvas sheets were placed inside the pits and used to cover the viewing windows.

These pits were found to be unsuitable for data collection, since the space between the soil and the windows quickly filled with water, so the roots could not be seen. These windows would need to be installed under a permanent cover to be used successfully.

Whole root system excavation and scanning.

In order to observe root system architecture of a variety of plant species, whole plant root systems were carefully excavated outside of the experimental arrays at the BIFoR research site. Care was taken to preserve as much of the root system as possible. An array of plants were excavated to match those observed around the minirhizotrons installed in the arrays. Once removed, soil was carefully brushed off the roots, and roots were carefully teased apart as much as possible without damaging the structure. A scanner was used with a black background sheet to produce scanned images of the root systems, to aid in root identification in the images from the minirhizotrons.

3.3 Results

3.3.1 Root lengths and widths from minirhizotrons

Root growth around the minirhizotrons was monitored for a two year period beginning in April 2017 and ending in March 2019. Tubes were imaged monthly during treatment periods and alternate months when no treatment was taking place, as described in 3.2.1. Sampling was undertaken over a single day when possible, and as close together as logistically viable when the work could not be completed in a single day. Each sampling session including the imaging of the total viewable area inside each of the 24 tubes; four in each of the six arrays. Initially, images were taken with the help of BIFoR volunteers as two people were needed; one to operate the camera and one the laptop. Visits were coordinated with the volunteers so that they were fully supervised at all times and provided transport

when necessary. The volunteers operated the camera and assisted with equipment setup and removal. Later, damage to the video capture card necessitated an adjustment to the camera setup. During this process the hardware was adapted to allow video recording, which meant that a single person could operate the equipment and later take frames from the video for analysis. Analysis was performed using SmartRoot software [Lobet et al., 2011] on images containing a root, and root length, diameter and number of branches were recorded for each root segment. Over the two years nearly 18,000 measurements of length, diameter and branch numbers were recorded; one for each time a root segment was captured in an image.

Total root length per tube was taken at each time point, separated between tubes in arrays treated with eCO_2 (treatment), and those with ambient air (control), as shown in Figures 3.4A and 3.4C. Therefore, each sampling session contributes 12 datapoints in Figure 3.4 for each array type. The data shows some seasonal variation, with a great deal of variation between tubes. Initially, greater lengths were recorded in control tubes, but this gap is reduced in the second year.

Root widths were recorded, where each datapoint is the width of a single root segment recorded, with a LOESS fit to the data (3.4B). Any seasonal variation appears far less pronounced than in 3.4A 3.4C, though the widths do appear to increase in the second year of sampling, with a greater increase in the treatment array. Again, there is a great deal of variation across datapoints, with particularly wide segments appearing in the control arrays.

3.3.2 Root weights from soil cores

Root samples were analysed from soil cores as part of wider research at the BIFoR site. Cores were initially taken in March 2017, then quarterly from March 2018-2019 as described in section 3.2.3. Cores were separated visually by soil horizon; soil cores contained

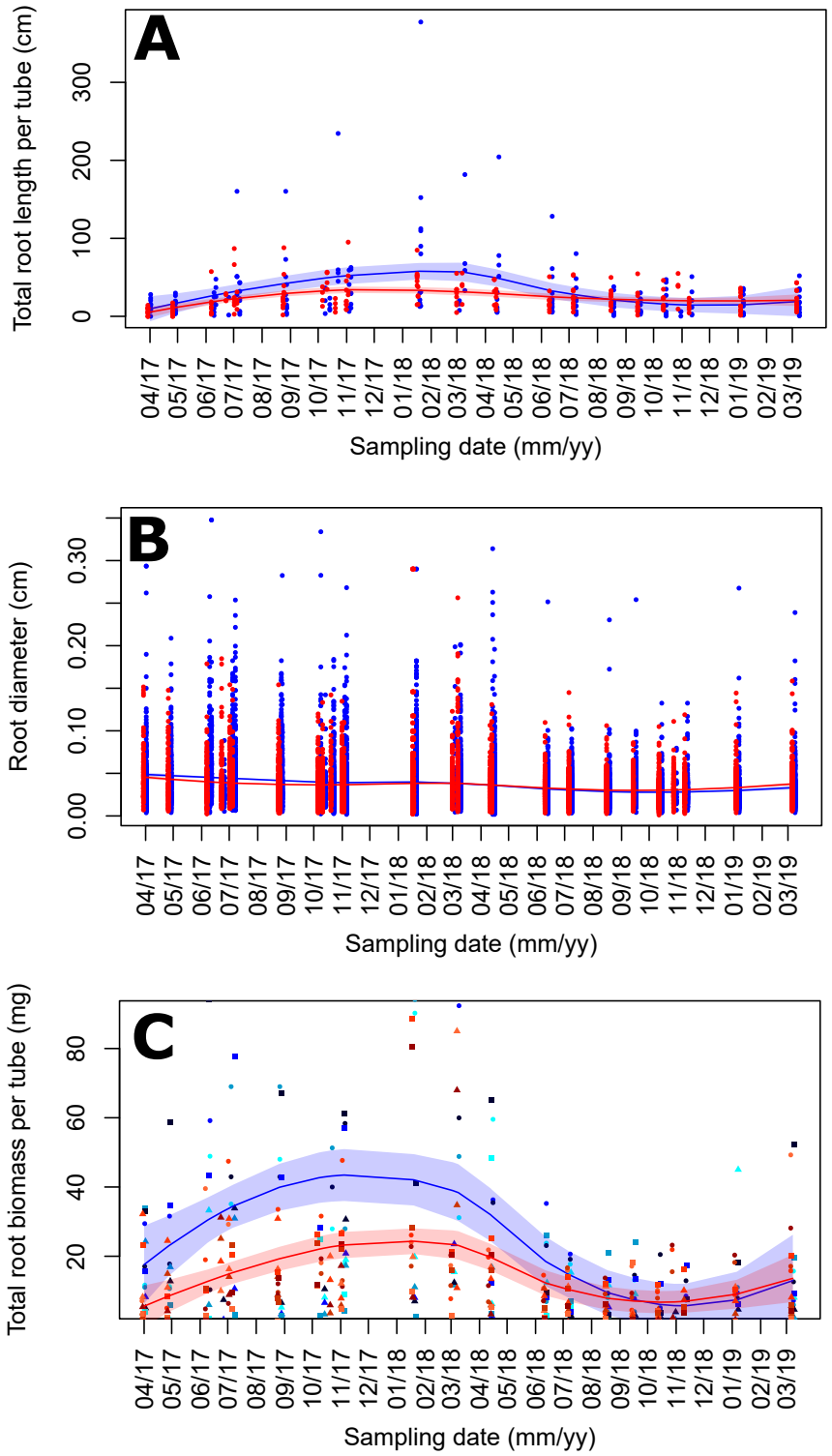


Figure 3.4: **Fine root observations in eCO_2 (red) and control (blue) experiments.** Each datapoint is the total root length in a single tube (A), the width of each segment observed (B), and the total biomass observed in a single tube (C). LOESS fits to the data are shown with 95% confidence intervals. To show the variation between tubes, (C) is plotted with different colours for each array, and a different plot marker for each tube within the array. As with the other plots, control arrays are shown in shades of blue and treatment in shades of red.

samples of the organic (O) horizon from the surface, topsoil (A) located below the organic layer, and the denser subsoil (B). Some cores did not contain any B horizon due to variations in horizon depth; if the O and A horizons were deep enough that the core did not reach the B horizon. Roots were separated from the soil by hand and processed as described in section 3.2.3

3.3.3 Root biomass from minirhizotrons

Root biomass was obtained from root data collected from minirhizotrons over the two year sampling period, as described in section 3.2.2, along with root density information using the method detailed in section 3.2.3. Total root biomass per tube was plotted, mirroring the root length plot (figure 3.4A). As seen in figure 3.4B, the biomass shows a more pronounced seasonal trend, mirroring that visible in figure 3.4A. The increase begins earlier in the first year of sampling than the second, beginning in spring of 2017, but not until late summer of 2018.

3.3.4 Root production

Root production data was collected by tracking singular roots growing along the top strip of a minirhizotron, in the region shown in Figure 3.7. Individual roots were identified and tracked between sampling periods (see figure 3.9), allowing for any growth to be recorded. The resulting production data is detailed in Figure 3.8. The production was greater in the first year, with a significant reduction between the first and second year of sampling. Both years saw greater production in treatment arrays than control, although this difference is lesser in the second year. These results generally follow the pattern shown in the length and biomass results (see Figures 3.4A and 3.4C).

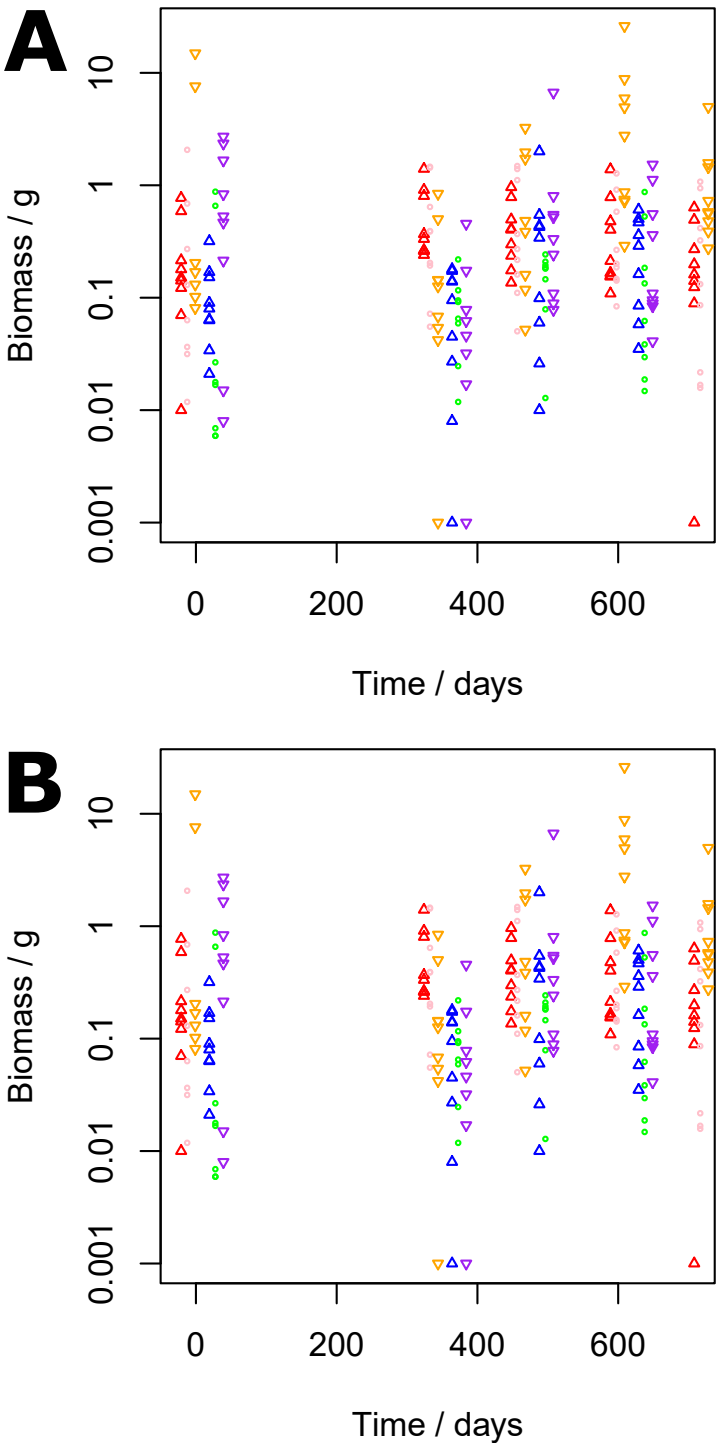


Figure 3.5: **Root biomass per horizon sampled from soil cores.** (A-B) (B) Biomass measurements separated by soil horizon for (A) control and (B) eCO_2 experiments. Horizons are O (up arrows), A (circles), B (down arrows); biomass is living (warm colours) and dead (cool colours). (C) Biomass summmed over all soil horizons for eCO_2 (red) and control (blue) experiments, classified by living, dead, and total. Each datapoint corresponds to a single array. LOESS fits with 95% confidence intervals are shown.

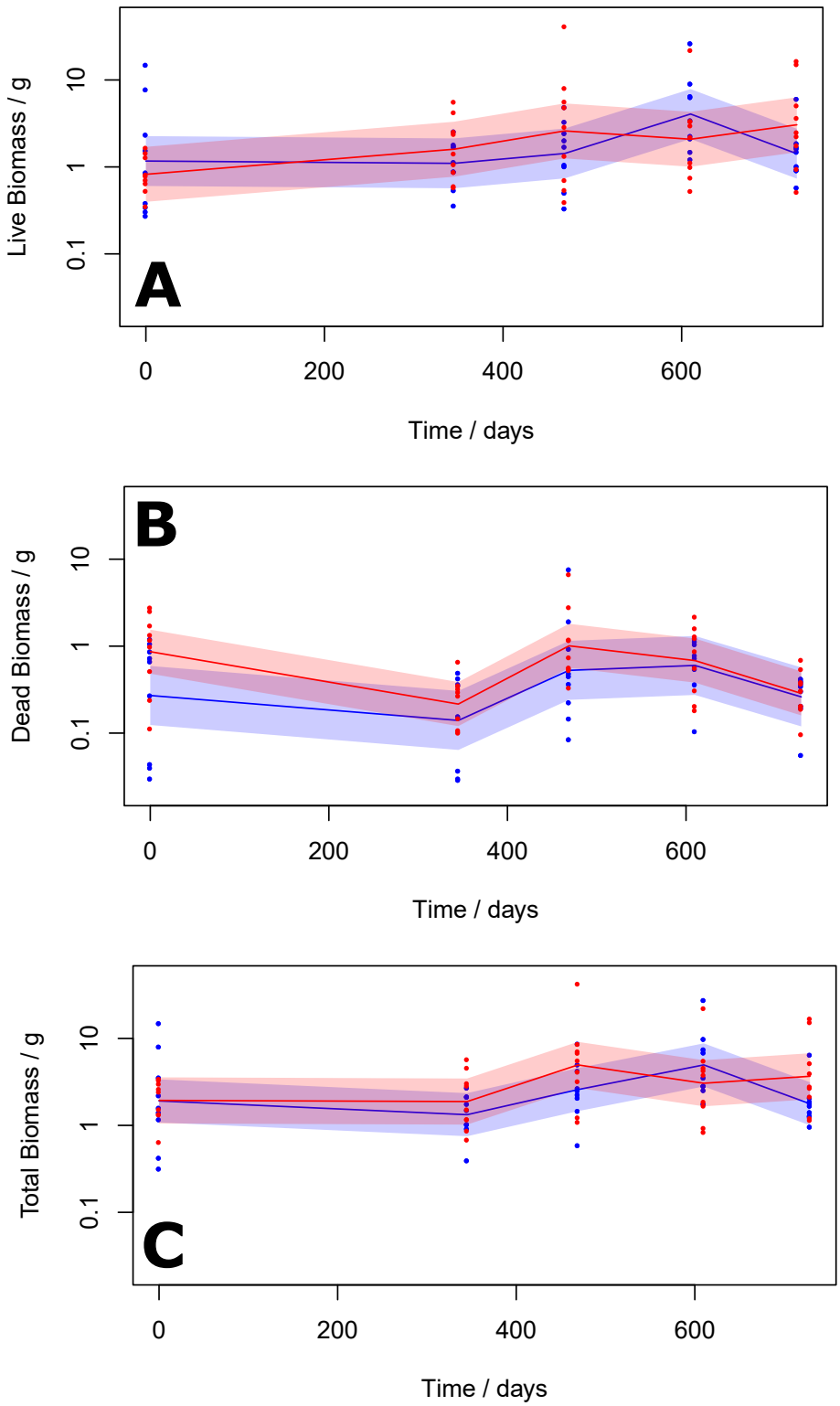


Figure 3.6: **Root biomass per array sampled from soil cores.** Biomass summmed over all soil horizons for eCO_2 (red) and control (blue) experiments, classified by (A) living, (B) dead, and (C) total. Each datapoint corresponds to a single array. LOESS fits with 95% confidence intervals are shown.

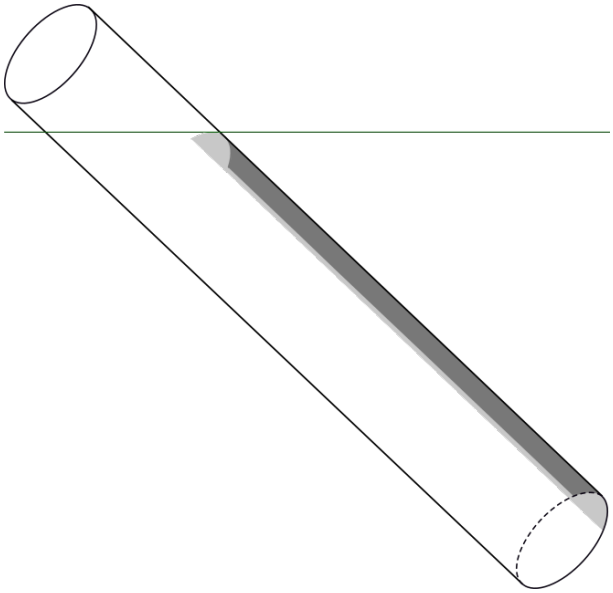


Figure 3.7: **The imaging region of a minirhizotron used for root production.** The tube is shown as installed in the ground, with the shaded area indicating the area in which root growth was monitored as detailed in section 3.3.4.

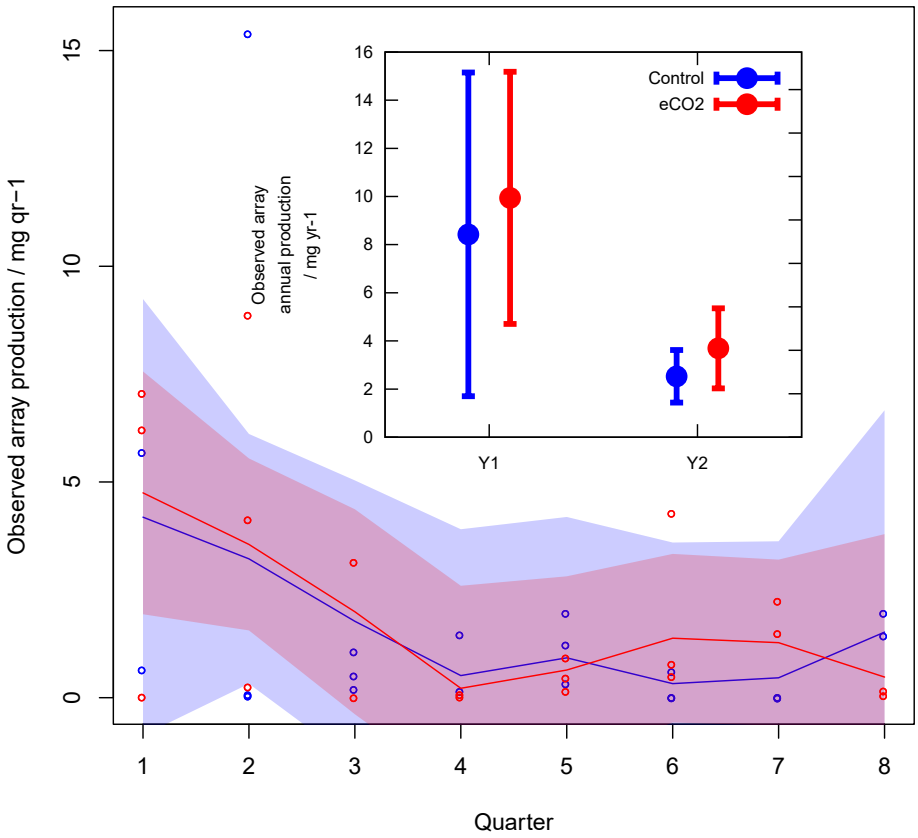


Figure 3.8: **The production across all of the tubes in all of the arrays, separated by eCO_2 (red), and control (blue).** The quarterly production is shown over the two year monitoring span with LOESS fits with 95% confidence intervals. The inset shows the total yearly production for each year. Error bars are standard error between tubes

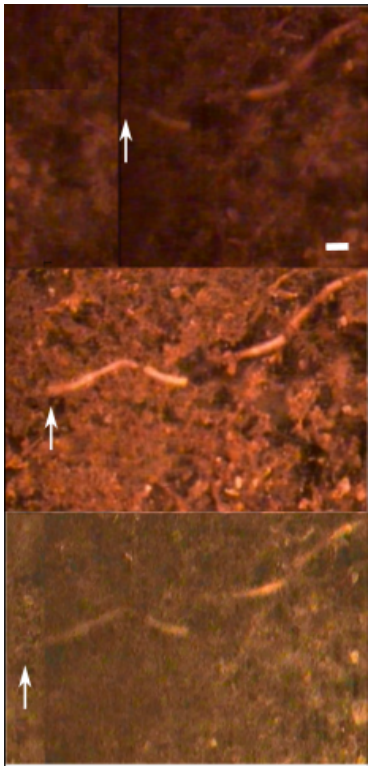


Figure 3.9: **Example of root growth used to calculate production.** Arrows identify the root tip and the horizontal scale bar represents 1mm. Images are taken month-by-month going from top to bottom.

3.3.5 Fold-change in fine root biomass

As discussed in Chapter 3, the raw data shows a great deal of variance and appears heavily dependent on sampling and environmental effects. Therefore, it would be greatly beneficial to find a method of separating any change in root dynamics due to elevated carbon dioxide from those caused by spatial or temporal variations in the natural ecosystem. In our biomass data, (see Chapter 3), the values observed appeared broadly similar for both treatment and control, but the rates of increase for spring-summer 2017 and late summer 2018 appeared much larger for the eCO_2 measurements compared to the control. To investigate this, a fold-change was calculated, so each datapoint showed the change in root biomass with respect to the biomass in the same tube at initial sampling (see section 3.2.4).

As can be readily seen from Figure 3.10, the increase in biomass for the treatment arrays appears substantially greater. LOSS fits applied to the data showed a maximum mean fold-change increase of 2.58 $\pm$ 0.20 for tubes in control arrays and 4.62$\pm$0.44 for those under elevated CO_2. This is also supported by Mann-Whitney tests performed on individual points for October 2017 and March 2018, showing values of 4.59 $\pm$ 0.58 for treatment and 2.49 $\pm$ 0.31 for control, where the errors represent standard error and p= 0.016. However, this approach suffers from an over reliance on the accuracy of the data from the first set of sampling, and the validity of this must be carefully considered before conclusions are drawn.

3.4 Discussion

The monitoring of natural root systems continues to be a challenging endeavor, both in finding accurate nondestructive methods and in collecting data that is representative of the true picture belowground [Hendricks et al., 2006]. Soil cores offer a relatively inexpensive and accessible method of quantifying root stocks, but are time consuming and are a

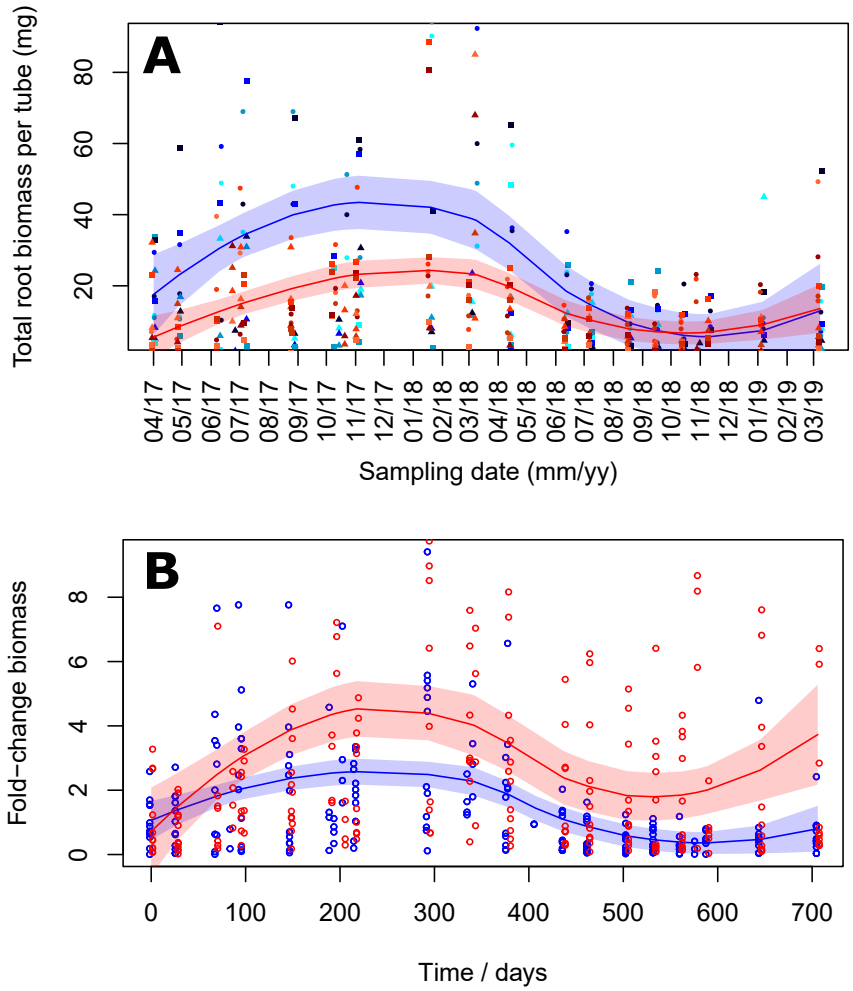


Figure 3.10: **A comparison of the raw biomass per tube and the 'fold-change' transformation.** (A) The biomass per tube data shown and detailed in Chapter 3. Fine root biomass observations in eCO2 (red) and control (blue) experiments. Each datapoint is an observation from a single minirhizotron site. LOESS fits to the data are shown with 95% confidence intervals (B) The same data having undergone the fold-change transformation described in section 3.2.4. Each datapoint is an observation from a single minirhizotron site, normalised by the initial biomass averaged across all sites in an individual array.

poor method for monitoring root growth over time. Ingrowth cores have been used in other studies [Maria do Rosário et al., 2000], but are a poor choice for a natural varied ecosystem such as that at BIFoR FACE since results are likely to be biased towards species with greater recovery rates, and are inappropriate for sites with high seasonal variability [Addo-Danso et al., 2016, Hertel and Leuschner, 2002]. Minirhizotrons have been the main method of data collection for this study, since they offer a more accurate estimation of productivity than soil cores [O'connell et al., 2003, Guo et al., 2004], however they present their own challenge to the researcher [Matamala et al., 2003, Norby and Jackson, 2000], and careful interpretation of the results is required [Yuan and Chen, 2012]. The use of both methods has provided a broader picture of root activity at BIFoR FACE.

Taken together, the raw results alone show little carbon fertilisation effect. Root length and biomass measurements per tube are greater in the control arrays than in treatment in the first year of sampling (see Figures 3.4A and 3.4C), with treatment results slightly larger than control by the second year. This result is also visible in root widths (Fig 3.4B), but the separation is very slight. However, this result is not echoed in the root production measurements (Figure 3.8), which shows greater production in the treatment arrays in both years of treatment, with a smaller difference in the second year. Figure 3.8 shows generally greater production in the treatment arrays in the second year, but a fair amount of variation month to month. Soil core results (Fig 3.6) suggest slightly greater root standing stock in treatment than control, but this is not consistent and well within error bars. There will also be an error present in the accumulation of production measurements, where the results will be dependent on sampling rate. This is not included in the results and would be useful to quantify, maybe by comparing results with different sampling rates.

A great deal of variation is shown in all of the results, reflecting heterogeneity across the site, with this dependence on sampling location visible in results from both minirhizotrons and soil cores. Heterogeneity is particularly visible in the soil core results (Fig 3.5A), where there is a great deal of variation in mass recorded in separate samples. For this reason,

results may have a large dependence on sampling location. However, as can be seen from Figure 3.4 C, the variability is present between sampling times and not consistent in each tube over time. Therefore, this variance appears to come from the sampling method and the sampling of a complex system rather than from just the tube locations. This effect is particularly concerning for productivity measurements, where a smaller region is used for sampling, and therefore there may be a greater impact on the results. While one of the strengths of the BIFoR face setup is its location in a natural ecosystem, spatial variation may be amplified by the number of competing species across the site. Since we do not identify species in any of our data, we are relying on our sampling methods to produce a broad picture of the overall ecosystem. The use of a natural ecosystem for this study also means that the results may be impacted by outside factors. The summer of 2018 (July - September) was particularly dry which may have affected root proliferation. The trees were also recovering from a springtime winter moth (*Operophtera brumata*) caterpillar infestation (17th May to 3rd June) [Hart, 2019]. Some increased growth in the first year could be attributed to a wounding response due to disturbance from minirhizotron installation [Pritchard et al., 2008]. All of these factors need to be considered when interpreting the results from the study.

The fold-change transformation in section 3.3.5 provided support for separation between the datasets with a Mann-Whitney test giving $p = 0.016$ in support of separate model fits for treatment and control data. This suggests that further modelling approaches could help to elucidate the effects of eCO_2 in the noisy data detailed in this chapter. Overall, it is difficult to see the real picture in these results due to the many uncontrolled factors affecting the data. Careful statistical analysis is required to account for this large variability in the results and obtain a clearer picture of the effect of eCO_2, as detailed in the following chapter.

Chapter 4

Exploring data from BIFoR FACE using mathematical modelling

4.1 Introduction

In the previous chapter, it was concluded that more statistical analysis was required to further elucidate the effect of eCO_2 on fine root growth. To this end, this chapter details modelling work performed on the data in Chapter 3. In addition, an estimation of Net Primary Production (NPP) is reported, along with careful error quantification. These methods provide further insight into the impact of eCO_2 on fine root growth at the BIFoR FACE site.

In section 3.3.5, it was shown how looking at the fold-change in root biomass showed a greater separation between treatment and control populations. However, this approach has a large dependence on the first set of measurements and can propagate error. A stochastic population model will mitigate this effect by including both initial and ongoing behaviour in the model fitting. This kind of model will account for all of the points, so the larger number of datapoints increases the strength of the model, while capturing the

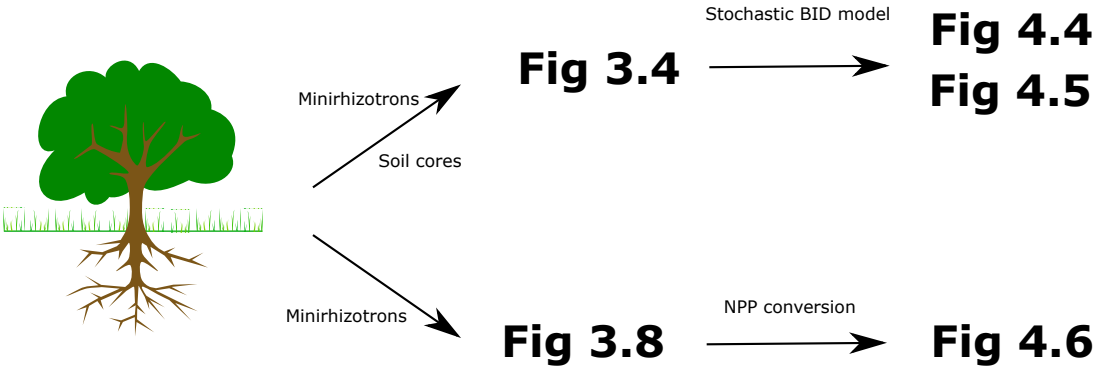


Figure 4.1: **Data collection and modelling schematic.** A schematic to show how the experimental work detailed in Chapter 3 is used in Chapter 4 and the modelling applied in each case.

distributional detail so the variance in the data is accounted for in addition to the mean trends. In comparison, ABC would be a poor choice for this kind of data, since it would be too computationally inefficient.

In addition to looking at the time-series measurements of root growth, a measurement of Net Primary Production (NPP) will allow for the inclusion of this work into larger Earth System Models which look at global carbon dynamics. This chapter details a novel approach to calculating NPP from the production data in section 3.3.4, along with a careful analysis of the uncertainty in this calculation. Overall, this chapter details how the application of stochastic modelling allows for greater insight into the kind of noisy data obtained from this type of study. In addition, we show the large uncertainty inherent in the calculation of NPP, and how this can be clearly quantified.

4.1.1 Stochastic modelling and quantifying uncertainty

Population models are used in a wide variety of biological applications [Novozhilov et al., 2006, Gani, 1984, Hastings and Palmer, 2003]. Stochastic modelling is particularly suited to the noisy systems often present in biology, and has been often used in cell modelling [Bressloff, 2014], and gene expression [Swain et al., 2002].

A birth-immigration-death process is an extension of the well-studied birth-death

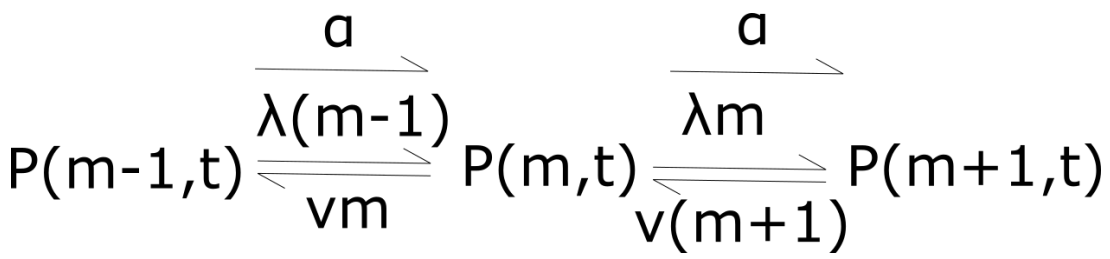


Figure 4.2: **Diagram of processes in BID model** $P(m,t)$ is the probability of the population being at size m at time t, λ is the birth rate, ν the death rate (both population dependent), and α the immigration rate.

stochastic process [Van Kampen, 1992], with the addition of the immigration term. Both models make the assumption of an infinite population, so there is no maximum limit from outside factors. Population members are modelled as particles that behave independently from one another; they can die or give birth at given rates. This model works at population level, and therefore the population birth and death rates will be dependent on the population size. The immigration term models an addition of new members at a given immigration rate. This is not dependent on the size of the population [Van Kampen, 1992].

This model is described by the master equation, which describes the evolution of the probabilities of the population being at a given state over time. To derive this model, we consider the probability of the population being at a given state m at time t. The master equation can then be derived by considering the probability to be at state $m-1$ at time t, and at $m+1$ at time t, see Figure 4.2. In this way, the model depends on a Markov assumption, where the changes in population are not influenced by what has happened prior to this, and it is assumed that the birth, immigration and death rates are time-independent [Van Kampen, 1992]. A full description of the model used in this chapter can be found in section 4.2.1.

A Birth - Immigration - Death (BID) stochastic model can be applied to root system data by considering a unit length of root as a member of the root 'population'. In this way, birth represents the growth and branching of existing roots, immigration a root hitting the minirhizotron and therefore being recorded, and death the decomposition of dead roots.

The birth-immigration-death (BID) model is well studied in mathematics [Van Kampen, 1992, Matis and Kiffe, 2000, Ekanayake and Allen, 2010], and has previously been applied to cell dynamics [Johnston and Jones, 2015]. The generating function can be readily written and solved, and moments can be calculated [Van Kampen, 1992, Johnston and Jones, 2015], making it suitable for modelling the noisy root data detailed in chapter 3.

4.1.2 NPP

Measuring net primary production (NPP) is crucial to our understanding of the carbon cycle [Fahey and Knapp, 2007]. In plants, NPP is the measurement of the energy stored as biomass [Fath, 2018] and represents the difference between energy gained through photosynthesis and that used for respiration [Waring, 2007]. Since NPP cannot be directly calculated [Clark et al., 2001a], it is often estimated using biomass production over a given time period [Clark et al., 2001b]. This contains many sources of error and may underestimate true NPP by up to 30% [Eviner, 2004]. In addition, belowground NPP is often neglected, sometimes assumed to be a proportion of aboveground NPP estimates [Medlyn et al., 2016, De Kauwe et al., 2016, Adame et al., 2017].

Previous FACE experiments estimate fine root NPP in a variety of ways. Aspen face did not include minirhizotrons in the study, and used estimates of fine root standing stock from soil cores in their total NPP calculation [King et al., 2005]. Duke FACE also did not include minirhizotrons in their study, calculating fine root turnover using estimates from soil cores [DeLucia et al., 1999]. ORNL FACE used minirhizotrons, and found an estimate of fine root NPP using a consideration of the volume of soil sampled from minirhizotrons [Norby et al., 2002]. The method from this study was expanded upon, using careful consideration of the geometry of the minirhizotron system to produce an estimate of NPP per unit area from fine root productivity observed in the minirhizotrons.

Variations in methodology make it very important to consider errors in the reporting of NPP values [Clark et al., 2001a]. In particular, fine root NPP from this study is

designed to be combined with other NPP figures from other experiments at BIFoR, to gain an understanding of the carbon cycle across the ecosystem. This figure could then be used in large scale carbon models [Flato, 2011]. It is readily seen how this could lead to a large propagation of error values. Therefore, clear consideration of the magnitude and sources of error is critical for reporting of NPP measurements.

4.2 Methods

4.2.1 BID model dynamics

A birth-immigration-death (BID) stochastic model was be applied to the time series root biomass data shown in Figure 3.4 C in Chapter 3. The model was conceptualised by considering a unit length of root as a member of the root 'population'. As in the main text, we take birth to represent the growth and branching of existing roots, immigration to represent a new root appearing in the minirhizotron viewing window and therefore entering the population, and death to represent the decomposition of dead roots. Taking m as the number of unit root lengths, λm the birth rate, νm the death rate and α the immigration rate, this model is described by the master equation

$$\frac{dP_m}{dt} = (\alpha + \lambda(m-1))P_{m-1} + \nu(m+1)P_{m+1} - (\alpha + \lambda m + \nu m)P_m, \tag{4.1}$$

for $P_m(t)$, the probability of a state with m unit root elements at time t. Initial conditions at $t = 0$, $\mathbb{E}(m(t=0)) = m_0$ and $\mathbb{V}(m(t=0)) = v_0$, are also parameters of the model.

BID statistics

We start with the BID master equation, describing the probability $P(m,t)$ of observing m length elements at time t under the influence of immigration α, birth λ and death ν:

$$\frac{dP(m,t)}{dt} = \alpha P(m-1,t)+\nu(m+1)P(m+1,t)+\lambda(m-1)P(m-1,t)-(\alpha+\nu m+\lambda m)P(m,t), \tag{4.2}$$

with initial condition enforcing that $m = m_0$ at $t = 0$:

$$P(m,0) = \delta_{m,m_0}. \tag{4.3}$$

Defining the generating function $G(z,t) = \sum_m z^m P(m,t)$, we obtain the following PDE from Eqn. 4.2

$$\frac{\partial G(z,t)}{\partial t} = \alpha(z-1)G(z,t) + (\nu(1-z) + \lambda(z^2-z))\frac{\partial G(z,t)}{\partial z}, \tag{4.4}$$

with initial condition

$$G(z,0) = z^{m_0}. \tag{4.5}$$

The solution is readily found through the method of characteristics [Johnston and Jones, 2015]:

$$G(z,t) = \left(\frac{\nu-\lambda}{\lambda e^{(\lambda-\nu)t}(z-1)-\lambda z+\nu}\right)^{\frac{\alpha}{\lambda}} \left(\frac{\nu e^{(\lambda-\nu)t}(z-1)-\lambda z+\nu}{\lambda e^{(\lambda-\nu)t}(z-1)-\lambda z+\nu}\right)^{m_0}. \tag{4.6}$$

We can obtain necessary information about $P(m,t)$ from the generating function (Eqn. 4.6). $P(m)$ is given by

$$P(m) = \left(\frac{1}{m!}\right)\left(\frac{\partial^m G}{\partial z^m}\right)\Bigg|_{z=0}, \tag{4.7}$$

the expected value by

$$\mathbb{E}(m) = \frac{\partial G}{\partial z}\Bigg|_{z=1}, \tag{4.8}$$

$$= e^{t(\lambda-\nu)}m_0 + \frac{e^{t(\lambda-\nu)\alpha}-1}{\lambda-\nu}, \tag{4.9}$$

and the variance by

$$\mathbb{V}(m) = \left(\frac{\partial^2 G}{\partial z^2} + \frac{\partial G}{\partial z} - \left(\frac{\partial G}{\partial z}^2\right)\right)\Bigg|_{z=1} \tag{4.10}$$

$$= \frac{\left(e^{t(\lambda-\nu)}-1\right)\left(e^{t(\lambda-\nu)}(\alpha\lambda + m_0(\lambda^2-\nu^2)) - \alpha\nu\right)}{(\lambda-\nu)^2}. \tag{4.11}$$

Note that for the purposes of this model we add an additional value v_0 to the variance. This linear noise contribution accounts for experimental variance due to noisy observations:

$$\mathbb{V}(m) = v_0 + \frac{\left(e^{t(\lambda-\nu)}-1\right)\left(e^{t(\lambda-\nu)}(\alpha\lambda + m_0(\lambda^2-\nu^2)) - \alpha\nu\right)}{(\lambda-\nu)^2}. \tag{4.12}$$

4.2.2 Statistical analysis and uncertainty quantification

Statistical analysis was performed in R [R Core Team, 2020] using custom scripts. LOESS fitting was performed using the default parameterisation of the *loess* command, specifically using a span α of 0.75 and a polynomial degree of 2.

Caladis [Johnston et al., 2014] was used for uncertainty propagation, specifically to track uncertainty through the calculation of NPP, as described in section 4.2.5. Caladis

is an online tool allowing for calculations using probability distributions [Johnston et al., 2014]. Each variable in a calculation is associated with a user-defined probability distribution reflecting uncertainty in that quantity, and when a calculation is performed the value of each variable is sampled from its distribution for use in the equation.

4.2.3 Maximum likelihood estimation for parameter values

Maximum likelihood estimation and bootstrapping for the BID model was performed in Mathematica [Inc.,]. The BID model admits a closed-form solution for an exact likelihood (which has been previously studied in stochastic biology [Johnston and Jones, 2015]), but for simplicity and because of the continuous nature of our root observations we employ a normal approximation. Hence, we set

$$\mathcal{L}(x, \tau) = \mathcal{N}(\mathbb{E}(m(t = \tau)), \sqrt{\mathbb{V}(m(t = \tau))}), \tag{4.13}$$

using expressions for $\mathbb{E}(m(t = \tau))$ (equation 4.8) and $\mathbb{V}(m(t = \tau))$ (equation 4.12), the mean and variance of root biomass observed at time τ.

The total log-likelihood was calculated for eCO_2, control and for both eCO_2 and control treated as one dataset. This came from the sum of the log-likelihood for each element in the corresponding dataset. These log-likelihood calculations were used to find parameter fits for the birth, immigration, death and the linear noise contribution (see section 4.2.1) using the NMaximize function in Mathematica [Inc.,] for numerical optimisation.

4.2.4 Bootstrapping on parameter estimation

Bootstrapping was performed on the maximum likelihood parameter estimation from section 4.2.3 using Mathematica [Inc.,]. 200 repetitions were performed, with samples the

same size as the original dataset. This was performed on the control, eCO_2 and combined dataset to produce histograms representing confidence intervals for the parameter fits for birth, immigration, death and linear noise contribution (see section 4.2.1).

4.2.5 Calculating NPP from fine root production

We now produce a novel scaling for the production data shown in Figure 3.8, to give an estimate for NPP.

The minirhizotron tube samples a proportion of the total volume of the soil column in which it is embedded:

$$V = \frac{V_s}{V_t} = \frac{l_I d(d+2r)}{2hr(r+d)\sin 2\phi}, \tag{4.14}$$

where h is the length of the viewing area, r is the radius of the minirhizotron tube, d is depth of viewing field, ϕ is the angle of the minirhizotron tube, and l_I is the viewing arc length. We further show that, given observed biomass production p through this sampling, the total NPP estimate is given by:

$$\text{NPP} = \frac{p}{VA}, \tag{4.15}$$

where A is the area on the surface covered by the viewing area, and V is calculated using equation 4.14.

Deriving an equation for NPP scaling

We model a rhizotron tube *in situ* at an angle ϕ with the horizontal soil surface. Root images are collected covering a segment with angle θ, and corresponding image width w,

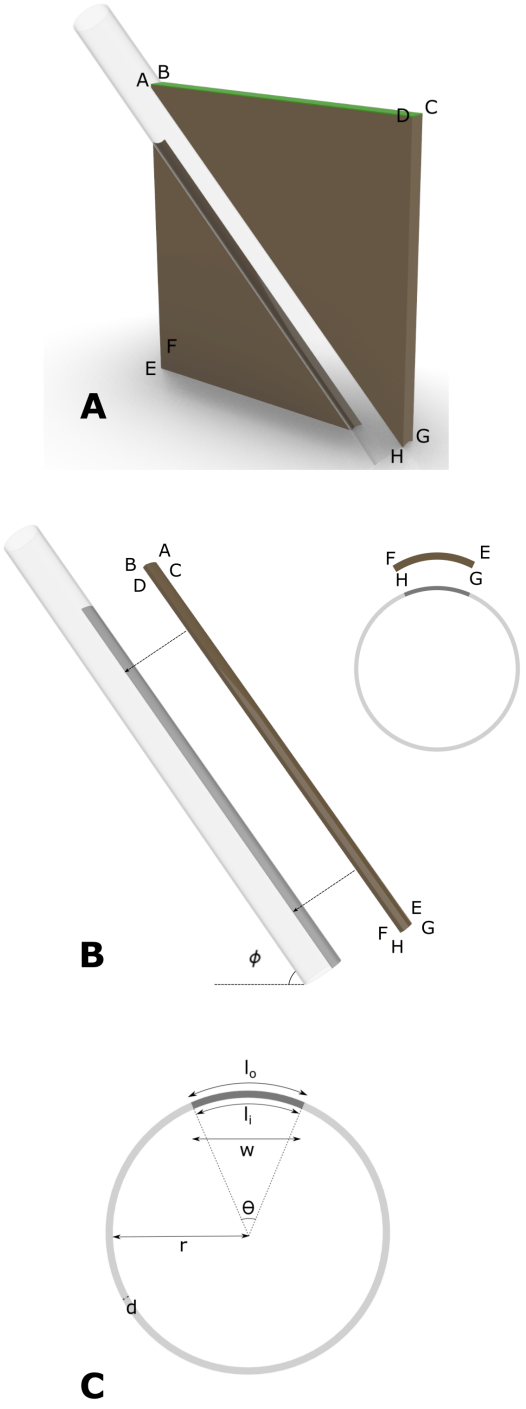


Figure 4.3: **Model geometry for production calculations.** (A) A minirhizotron tube in soil, with sampling volume V_s a subset of total volume V_t, which is volume ABCDEFGH coloured brown in the figure. The top surface is area A, shown as ABCD in green. (B) The sampling volume V_s shown here as volume ABCDEFGH away from the tube for clarity. The shaded area of the tube represents the region of the tube imaged for root production. Also shown is a view of the end of the tube, showing the shape of face EFGH of volume V_s, again shown away from the tube. (C) A cross section of the tube showing the geometry of the viewing area, shown as a shaded segment.

which allows a viewing window of width l_I with depth of field, d, as shown in Figure 4.3. We aim to produce a 2D projection, A, of the volume sampled by the minirhizotron, V_s, onto the surface for the estimation of root NPP as a production per unit area. To do this, we define the proportion of the total volume below this region (A) that has been sampled using the minirhizotron, and call this volume V_t (Figure 4.3A).

The viewing angle, θ (in radians) can be calculated from the viewing arc l_I and the tube radius r

$$\theta = \frac{l_I}{r}, \tag{4.16}$$

and therefore the outer arc l_O can be determined from the depth of field (Figure 4.3B):

$$l_O = \frac{(r+d)l_I}{r} \tag{4.17}$$

The sampling volume around the minirhizotron, V_s, Figure 4.3B. From the area of a circle sector, area ABO is $\frac{\theta}{2}(r+d)^2$ and area CDO is $\frac{\theta}{2}r^2$, so area $ABCD$ is $\frac{\theta}{2}((r+d)^2-r^2)$ and

$$V_s = \frac{\theta h}{2}((r+d)^2 - r^2), \tag{4.18}$$

$$= \frac{l_I h}{2r} d(d+2r). \tag{4.19}$$

The total volume V_t is delimited by the extreme points of the sampling volume $ABEFPQ$. From Figure 4.3B it is readily seen that lengths $AE = h\cos\phi$, $AQ = h\sin\phi$, $AB = (r+d)$, so

$$V_t = 2h^2(r+d)\sin\phi\cos\phi, \tag{4.20}$$

$$= h^2(r+d)\sin 2\phi. \tag{4.21}$$

Using equations 4.19 and 4.21, we obtain the proportion of the total volume sampled by the minirhizotron:

$$V = \frac{V_s}{V_t} = \frac{l_I d(d+2r)}{2hr(r+d)\sin 2\phi}. \tag{4.22}$$

Obtaining the sampled arc length l_I from the image width

The viewing arc, l_I, can be calculated from the image width, w, and the tube radius, r, using the following set of equations:

$$l_I = \arcsin\left(\frac{w}{h+\frac{w^2}{4h}}\right)\left(h+\frac{w^2}{4h}\right), \tag{4.23}$$

where

$$h = r\left(1-\frac{\cos\theta}{2}\right), \tag{4.24}$$

and

$$\theta = 2\sin^{-1}\left(\frac{w}{2r}\right). \tag{4.25}$$

Scaling to total fine root NPP

Let p be the fine root production. We can calculate total NPP as:

$$\text{NPP} = \frac{p}{VA}, \tag{4.26}$$

where A is the area at the surface that has been imaged (see Figure 4.3).

In this case A is given by

$$A = hw\cos\phi. \tag{4.27}$$

where h is the length and w the width of the viewing area and ϕ is the angle of the tube (see Figure 4.3).

4.2.6 Calculating NPP using Caladis

Caladis allows for calculations using probability distributions [Johnston et al., 2014]. Each variable is associated with a user-defined probability distribution, and when a calculation is performed the value of each variable is first pulled from its distribution for use in the equation. By repeating the calculation with a new set of variables each time, a histogram is produced showing the result of the calculation after a default of 20000 iterations.

Our NPP calculation uses the following equation (derived above):

$$\text{NPP} = \frac{2phr(r+d)\sin(2\phi)}{l_I d(d+2r)hw\cos(\phi)}. \tag{4.28}$$

Each variable in equation 4.28 has an associated probability distribution, defined below, chosen to reflect the confidence levels in the values given. A normal distribution is

used for measurements, or for normally distributed data while a uniform distribution was used when the values were all equally likely.

- $p \sim N(\mu_p, \sigma_p)$, where μ_p is the total root production (g yr^{-1}) and σ_p is the standard deviation of the production for each array, which comes from the data shown in Figure 3.8 (inset) in Chapter 3.

- $h \sim N(\mu_h, \sigma_h)$, where μ_h is the total viewing length along the rhizotron (m) (see Methods). Our calculation used $h \sim N(0.1755, 0.005)$ to allow for inacurracies in the measurement of the viewing windows.

- $r \sim N(\mu_r, \sigma_r)$, where μ_r is the the radius of the minirhizotron tube (m). Our calculation used $r \sim N(0.0275, 0.005)$ to allow for variations in tube manufacturing.

- $d \sim U(d_-, d_+)$, where d_- and d_+ represent the maximum and minimum depth of field values (m). These values were chosen to represent the spread of depth of field values in [Taylor et al., 2014], centred around the standard value used. A uniform distribution $d \sim U(0.0005, 0.0035)$ was used to reflect the uncertainty as explained in [Taylor et al., 2014], rather than this being a measurement.

- $\phi \sim N(\mu_\phi, \sigma_\phi)$, where μ_ϕ is the angle of the minirhizotron tube. Our calculation used $\phi \sim N(0.756, 0.196)$ (radians) to cover variations in tube installation.

- $w \sim N(\mu_w, \sigma_w)$, where μ_w is the width of the viewing area. Our calculation used $l_i \sim N(0.0192, 0.005)$ to capture inaccuracies in the measurement and calculation of this width.

- $l_i \sim N(\mu_{l_i}, \sigma_{l_i})$, where μ_{l_i} is the viewing arc length (m) (see Methods). Our calculation used $l_i \sim N(0.0189, 0.005)$ to capture inaccuracies in the measurement and calculation of this arc length.

The mean values given, and the midpoint of the depth of field distribution, were then used to calculate the NPP values quoted in section 4.3.2

4.3 Results

4.3.1 Stochastic modelling for fine root biomass

As discussed in section 3.3.5, the raw minirhizotron data shows a great deal of variation due to outside effects such as weather events, a caterpillar infestation and spread of vegetation across the site. This is readily seen in Chapter 3, and figure 3.10A. Therefore, it would be prudent to model the system in a way that allows for the influence of these outside processes that cannot be easily modelled deterministically. Stochastic modelling is suitable for a noisy system such as in this study. It aims to capture the distributional detail while allowing for variation within the dataset. A BID model is well suited to time series data, and is strengthened by the inclusion of multiple datapoints at each time. This consideration of the whole dataset represents a significant improvement upon the fold change model shown in chapter 3.

The mechanisms considered in this study are root growth, root decay, and new roots entering the area of observation. A model that readily maps to these three processes is a birth-immigration-death (BID) model, as described in section 4.2.1. In this model, we seek to find stochastic birth, immigration and death rates that correspond to the growth of roots, roots entering the viewing area, and decay of roots respectively; the se will be compared for treatment and control data.

The stochastic BID model described in section 4.2.1 was fitted to root biomass data shown in figure 3.4C by finding the parameters that maximise the likelihood function as detailed in section 4.2.3. A likelihood ratio test was performed, comparing the use of two models; the first model with separate datasets for control and eCO_2 data, with a second model containing the whole dataset. The first model will require the inclusion of 8 parameters: birth rate, immigration rate, death rate and linear noise contribution for both eCO_2 and control populations. In comparison, the second model only requires the fitting of 4 parameters: birth rate, immigration rate, death rate and linear noise for a

single population containing the whole dataset. Therefore, the null hypothesis would be that there is no separation between the control and eCO_2 datasets, and the alternative hypothesis that there is support for two distinct population. In this way, we look at the statistical evidence for separating the model, and compare whether separating the datasets provides improved model fit significant to offset the use of two sets of parameters. This showed that in the first 300 days there is a significant difference between treatment and control datasets ($p = 1.0 \times 10^{-11}$). This is supported by considering the maximum likelihood value of the difference between birth and death rates, which is positive (0.0045 day^{-1}) for treatment and negative (-0.0086 day^{-1}) for control. This separation can readily be seen in figure 4.4D which shows the result of boostrapping with the percentile method ($p = 0.041$). The data was separated by year since the expected seasonal variation would be poorly captured by this model type.

The same process applied to the second year data showed negative proliferation across both datasets, but with a less negative result in the fit for eCO_2. There is again a significant difference between the two datasets ($p = 1.3 \times 10^{-13}$). Both the separation and difference in root proliferation can readily be observed in figure 4.5, especially when compared to figure 4.4. Here the use of the BID stochastic model has shown a significant difference between the two datasets that was not easily seen from the raw data shown in Chapter 3.

4.3.2 NPP estimation and uncertainty

As discussed in Chapter 3, root productivity was estimated through continued observation of growing root segments along the top strip of minirhizotrons. In transforming this to a measurement of NPP, a method was sought that would also quantify the uncertainty in the measurement, both from the experimental setup and from assumptions used in the transformation.

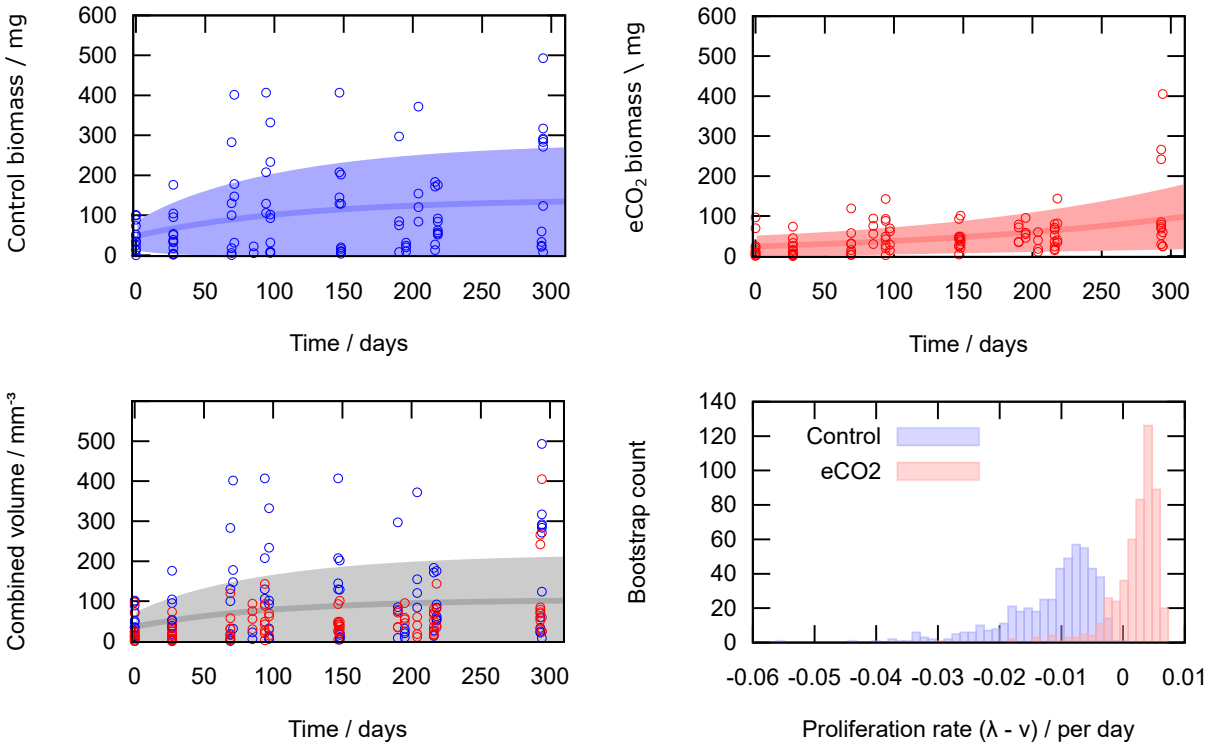


Figure 4.4: **Root biomass changes over time under eCO_2 and control conditions in the first year of sampling.** (A-C) The birth-immigration-death (BID) model described in the text, applied to (A) eCO_2, (B) control, and (C) combined biomass observations. The datapoints here are the biomass data for the first year of sampling as shown in figure 3.4C. Time axis gives days from 11/04/2017. The maximum likelihood BID parameterisation is found for each dataset, then the mean and standard deviation of the model for that parameterisation is plotted. A likelihood ratio test shows statistical support for the individual models (A)+(B) over the combined model (C) ($p = 1.3 \times 10^{-13}$). (D) Bootstrapped estimates for the difference between root elongation (birth, λ) and root decay (death, ν) parameters for eCO_2 and control data, as described in section 4.2.4. $\lambda - \nu$ is higher (with positive maximum likelihood estimate) for eCO_2, reflecting increasing root proliferation, and lower (with negative maximum likelihood estimate) for control, reflecting decreasing proliferation.

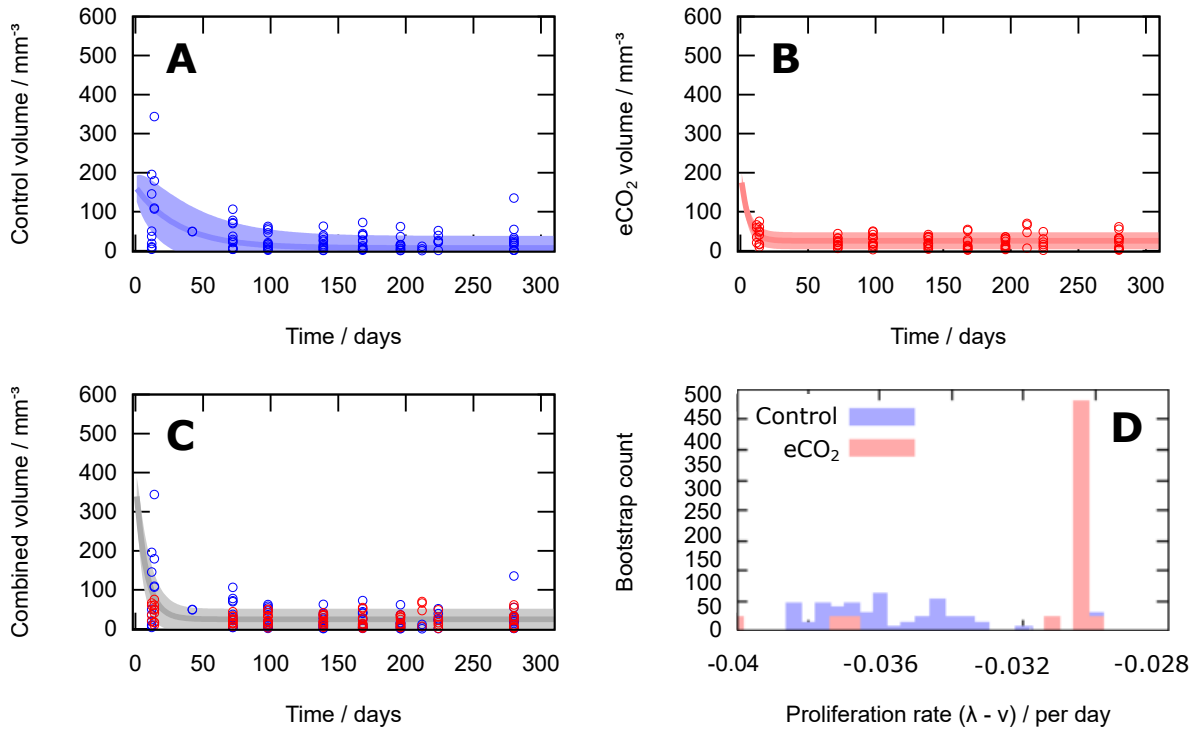


Figure 4.5: **Root biomass changes over time under eCO_2 and control conditions in the second year of sampling.** (A-C) The birth-immigration-death (BID) model described in the text, applied to (A) eCO_2, (B) control, and (C) combined biomass observations. The datapoints here are the biomass data for the second year of sampling as shown in figure 3.4C. Time axis gives days from 23/04/2018. The maximum likelihood BID parameterisation is found for each dataset, then the mean and standard deviation of the model for that parameterisation is plotted. A likelihood ratio test shows statistical support for the individual models (A)+(B) over the combined model (C) ($p = 1.3 \times 10^{-13}$). (D) Bootstrapped estimates for the difference between root elongation (birth, λ) and root decay (death, ν) parameters for eCO_2 and control data, as described in section 4.2.4. $\lambda - \nu$ is higher for eCO_2, although both have a negative maximum likelihood estimate, reflecting decreasing proliferation across the site.

An equation was derived for NPP from consideration of the geometry of the setup (see section 4.2.5 and figure 4.3):

$$\text{NPP} = \frac{2phr(r+d)\sin(2\phi)}{l_I d(d+2r)hw\cos(\phi)}. \tag{4.29}$$

Where p is the production observed through minirhizotron samples (see Methods), h is the length and w the width of the viewing area, r is the radius of the minirhizotron tube, d is depth of viewing field, ϕ is the angle of the minirhizotron tube and l_I is the viewing arc length. Briefly (full explanation in Methods), Eqn. 4.29 maps the productivity observed in the part-cylindrical viewing region of the minirhizotron tube to the corresponding surrounding volume of the soil column, and maps this volume to the corresponding 2D surface area for interpretation as a traditional NPP measurement. This equation was used with Caladis [Johnston et al., 2014], to quantify the uncertainty involved in each of the parameters that are used in the NPP calculation. In this way, the uncertainty involved in the measurement can be readily quantified.

As seen in Figure 4.6, the overall NPP was greater in under eCO_2 in both years of sampling, with mean values of 467 $\pm$ 373 g m^{-2} yr^{-1} control and 551 $\pm$ 290 g m^{-2} yr^{-1} for eCO_2 in year 1, and 140 $\pm$ 61 g m^{-2} yr^{-1} and 204 $\pm$ 92 g m^{-2} yr^{-1} in year 2. However, there is substantial uncertainty with the transformation to NPP, as visible in figure 4.7. The use of Caladis has facilitated the observation of this significant uncertainty.

The output from running Caladis on equation 4.28 with the probability distributions detailed in section 4.2.6 is shown in figure 4.7. The distributions show some larger NPP values in eCO_2 than control as in Figure 4.6, but the separation is relatively small. However, the spread of the histograms shows the massive uncertainty in the calculation of NPP that is not captured in a direct calculation as shown in figure 4.6. This very large uncertainty is illustrated by the impossibility of obtaining the values in figure 4.6 from the outputs in figure 4.7 because of the large spread in the results. As the results in

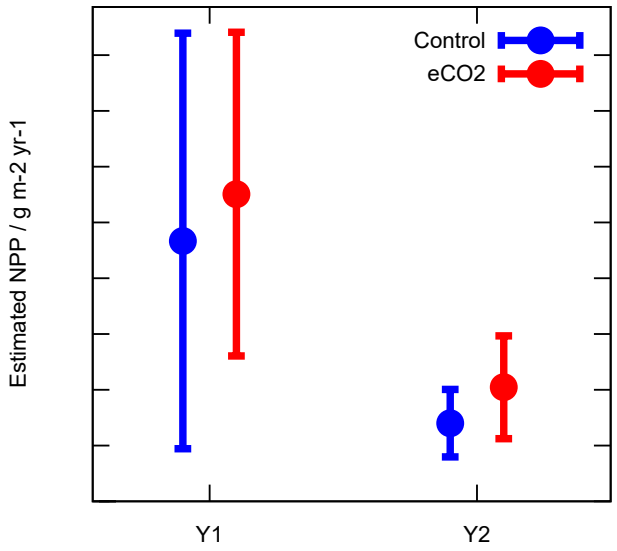


Figure 4.6: **Net primary productivity estimates and uncertainties.** Calculated NPP per quarter (traces) and per year (points) from root observations under eCO_2 (red) and ambient air (blue). Error bars are standard error between tubes.

section 4.3.1 show that there is some significant differences between eCO_2 and control, and the increase under eCO_2 is echoed in the NPP calculations (figure 4.6), this suggests that NPP may be a poor representation of root growth. The broad spread in the Caladis results (figure 4.7) show how much uncertainty there is in making the approximation from production to NPP.

4.4 Discussion

Two modelling approaches were applied to raw root data collected using minirhizotrons at the BIFoR FACE site, detailed in Chapter 3. Both approaches suggest an increase in root growth under eCO_2 with one of them showing statistically significant differences. The fitting of the stochastic birth-immigration-death (BID) model in section 4.3.1 was applied to root biomass data, and provided support for separation between the datasets with a Mann-Whitney tests giving $p = 1.0 \times 10^{-11}$ supporting separate model fits for treatment and control data. Although this model does not capture the apparent yearly seasonality in the data and is therefore a poor choice for forecasting, it does clearly show

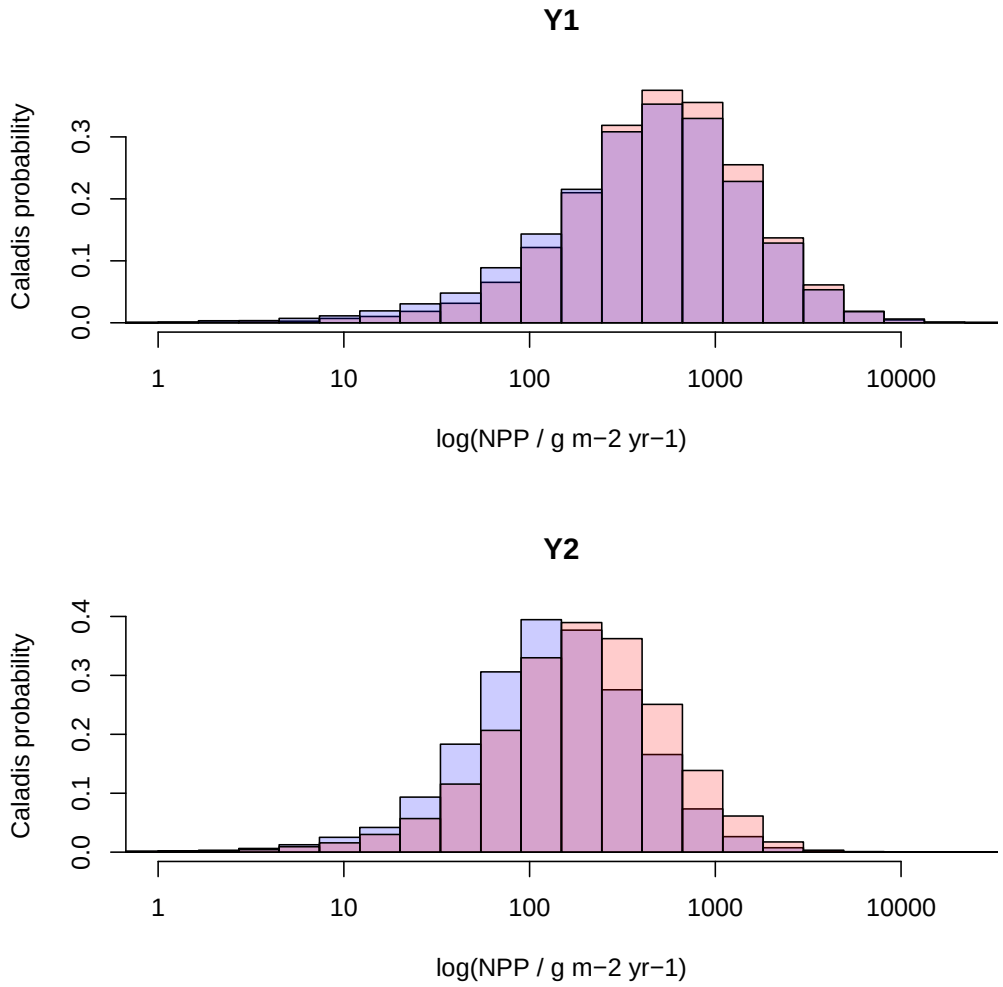


Figure 4.7: **NPP estimates from Caladis.** Estimates from Caladis [Johnston et al., 2014] of NPP in year 1 (top) and year 2 (bottom) of fine root productivity observations. Control data is shown in blue and eCO_2 in red. Distributions reflect the uncertainty in these estimates, derived by propagating uncertainty in each value involved in the calculation.

separation between the treatment and control datasets. In this way, a novel modelling approach has provided greater evidence for increased root biomass under eCO_2 that was not immediately clear from the initial data in Chapter 3. Initially, the raw results shown in figure 3.4 showed greater root biomass in the control array, although when only the root production was considered there was increased production under eCO_2 as shown in figure 3.8. This suggests that there was some difference that was difficult to extract due to the noisy system at BIFoR FACE, as discussed in Chapter 3. This was then supported by the transformation to fold change (figure 3.10). The novel application of a stochastic BID model to root data has elucidated this separation between eCO_2 and control populations.

The transformation to NPP applied to production data from Chapter 3 has clearly shown the great deal of 'hidden' uncertainty involved with a calculation of this type. Although there is an increase in NPP under eCO_2 for both years of sampling (see section 4.3.2), this is not statistically significant. Although NPP is an crucial figure for inclusion in wider research [Flato, 2011], it is important that errors in the calculation are clearly quantified for consideration when these values are included in larger models. The use of Caladis provided a valuable representation of the magnitude of this uncertainty. In summary, NPP may not be the best way of representing fine root growth in these kind of studies, due to the large amount of additional processing and measurements in what is already noisy data prone to variation. Scaling to NPP necessitates a consideration of depth of field, geometry of the minirhizotron and spacial area that are not present in other measurements and introduce additional error that is not quantified in the measurement. However, fine root NPP is a valuable measurement when considering the whole forestry ecosystem and the carbon cycle. Therefore, reporting of NPP values should be done carefully, with a greater consideration of sources of error than is currently present in the literature.

Overall, taking the three approaches together increases confidence that elevated ambient carbon dioxide stimulates root growth at BIFoR FACE. However, each approach comes with its own drawbacks and assumptions, and these must be carefully considered

when evaluating these results.

Chapter 5

Discussion

This project set out to investigate how plant root growth is impacted by elevated carbon dioxide (eCO_2). I began by looking at root growth in general, and investigated a method of gaining greater insight from existing root architecture models through the use of ABC SMC. To complement this theoretical analysis, I conducted a two year experimental study into the impact of eCO_2 on fine root growth in a temperate oak forest. Following the collection of these results, I used mathematical and statistical modelling techniques to gain further insight into the noisy data from a complex forest ecosystem, to extract the carbon fertilisation from other factors affecting the results. The central hypothesis was that eCO_2 would have a detectable fertilisation effect on root growth, and this work evidences that this is the case, with the modelling giving key insight that was not clearly evident in the initial results.

Chapter 2 details a novel framework for parameter inference and model selection that can be applied to existing root systems. This could be used to improve mechanistic understanding of root growth and in model selection for existing root architecture models. There are many excellent root system architecture models in the literature [Dunbabin et al., 2013], and this work provides a way to gain greater understanding from these root models, by showing the parameter space that could have produced an observed root

system. This was shown in Figure 2.3, where clear separation between the parameter spaces is visible for *Arabidopsis thaliana* and *Lupinus angustifolius* roots generated using RootBox [Leitner et al., 2010]. In addition, Figure 2.8 illustrates the strength of this approach for model selection, comparing a simple growth model to that used by CRootBox [Schnepf et al., 2018]. In this way, the applicability of this work to the existing literature is clearly shown. In addition, the ability to distinguish between phenotypes shown in Figure 2.9 could have many wider applications in future research. In summary, this readily applicable inference process allows for the consideration of the inverse problem which is poorly represented in current plant growth literature. In this way, quantitative and interpretable mechanistic insight can now readily be gained from root system architecture models.

While providing a significant contribution to root modelling literature [Postma and Black, 2020], the work detailed in Chapter 2 could not provide enough detail of root growth mechanisms to answer the central research question; the impact of eCO_2 on root growth. The two year fieldwork study conducted at BIFoR FACE was set up to answer this question. This study was designed and implemented incorporating established techniques in environmental science to elucidate belowground root dynamics, with careful application of each method in light of the difficulty in belowground sampling. This work will provide a key part in understanding carbon flows at BIFoR FACE. In turn, knowledge of the future carbon sink potential of temperate oak forests will improve the accuracy of current earth system models, and therefore help predict the future of our planet under climate change. Root data was collected through a variety of different methods; time series monitoring of the root lengths, widths, and calculated biomass visible within each minirhizotron at each sampling point (Figure 3.4), the root biomass from soil cores taken at the site (Figures 3.5 and 3.6) and the root production visible from monitoring root segments growing around the minirhizotron (Figure 3.8). Taken together, these results showed little evidence of increased root growth under eCO_2, indeed the root biomass monitoring suggested greater growth in the control arrays, the soil coring showed little discernible separation between

treatment and control, and the production measurement showed greater production under eCO_2 that was not statistically significant. These results will greatly benefit from more data, as BIFoR is intended to continue running well past the duration of this study. It is also possible that a greater fertilisation effect will be seen in later years as the forest continues to adapt. This is particularly likely for BIFoR, as it is situated in a mature oak forest rather than a fast-growing plantation. In addition, the resettlement period was shorter than would have been optimal, and that may have had an impact on the data observed, this will not be a problem in later years. Overall, a clear picture was not obtained from these fieldwork methods, and the results were conflicting. However, looking at the fold-change in root biomass (Figure 3.10 showed an increase under eCO_2, supporting the results from the production measurement. This evidenced the need for applying modelling techniques to these results to separate any fertilisation effect from eCO_2 from the noisy system at BIFoR FACE. In addition, a transformation from production to NPP was necessary to allow for a larger picture to be built of carbon flows within the whole forestry ecosystem.

The modelling in Chapter 4 provided greater evidence for a carbon fertilisation effect in the data. The application of a stochastic BID model to the biomass data provided statistical evidence of two separate populations (Section 4.3.1, Figure 4.4, Figure 4.5). This model was unable to capture the seasonality evident in the data, so is not able to make accurate forward predictions. However, it does provide evidence for carbon fertilisation that was not evident without the application of the model (Figure 3.4). This analysis shows the power of stochastic modelling for use in complex biological systems with a great deal of uncertainty, and further work would be valuable in adapting a stochastic modelling framework to capture the seasonality in the data. The application of stochastic modelling is a novel step for this type of root study, and presents a philosophy for greater understanding that could readily be applied to further research. The application of Caladis [Johnston et al., 2014] allowed for greater clarity on the inherent uncertainty in the NPP calculation. This has great importance when figures are incorporated into

large-scale models of the carbon cycle [MacDougall et al., 2017], and is something that should be considered when reporting figures from similar studies. The careful uncertainty quantification present in this work sets new standards for data collection and climate modelling in environmental science.

Overall, this project has made contributions to the fields of environmental science and mathematical biology, combining big data, fieldwork, modelling and statistics. The interdisciplinary approach has facilitated the use of statistical and modelling techniques in areas where they are rarely applied, and this has led to greater insights than a more traditional approach, particularly when quantifying uncertainty.

Bibliography

[Adame et al., 2017] Adame, M. F., Cherian, S., Reef, R., and Stewart-Koster, B. (2017). Mangrove root biomass and the uncertainty of belowground carbon estimations. *Forest ecology and management*, 403:52–60.

[Addo-Danso et al., 2016] Addo-Danso, S. D., Prescott, C. E., and Smith, A. R. (2016). Methods for estimating root biomass and production in forest and woodland ecosystem carbon studies: A review. *Forest Ecology and Management*, 359:332–351.

[Ainsworth and Long, 2005] Ainsworth, E. A. and Long, S. P. (2005). What have we learned from 15 years of free-air co_2 enrichment (face)? a meta-analytic review of the responses of photosynthesis, canopy properties and plant production to rising co_2. *New phytologist*, 165(2):351–372.

[Ainsworth and Rogers, 2007] Ainsworth, E. A. and Rogers, A. (2007). The response of photosynthesis and stomatal conductance to rising [co_2]: mechanisms and environmental interactions. *Plant, cell & environment*, 30(3):258–270.

[Allen et al., 2000] Allen, A., Andrews, J., Finzi, A., Matamala, R., Richter, D., and Schlesinger, W. (2000). Effects of free-air co_2 enrichment (face) on belowground processes in a pinus taeda forest. *Ecological Applications*, 10(2):437–448.

[Allen Jr, 1992] Allen Jr, L. H. (1992). Free-air co_2 enrichment field experiments: An historical overview. *Critical Reviews in Plant Sciences*, 11(2-3):121–134.

[Azzalini, 1996] Azzalini, A. (1996). *Statistical inference based on the likelihood*, volume 68. CRC Press.

[Band et al., 2012] Band, L. R., Fozard, J. A., Godin, C., Jensen, O. E., Pridmore, T., Bennett, M. J., and King, J. R. (2012). Multiscale systems analysis of root growth and development: modeling beyond the network and cellular scales. *The Plant Cell*, 24(10):3892–3906.

[Bates, 1937] Bates, G. (1937). A device for the observation of root growth in the soil. *Nature*, 139(3527):966.

[Beaumont et al., 2002] Beaumont, M. A., Zhang, W., and Balding, D. J. (2002). Approximate bayesian computation in population genetics. *Genetics*, 162(4):2025–2035.

[Bertorelle et al., 2010] Bertorelle, G., Benazzo, A., and Mona, S. (2010). Abc as a flexible framework to estimate demography over space and time: some cons, many pros. *Molecular ecology*, 19(13):2609–2625.

[Biau et al., 2010] Biau, D. J., Jolles, B. M., and Porcher, R. (2010). P value and the theory of hypothesis testing: an explanation for new researchers. *Clinical Orthopaedics and Related Research®*, 468(3):885–892.

[Bodner et al., 2013] Bodner, G., Leitner, D., Nakhforoosh, A., Sobotik, M., Moder, K., and Kaul, H.-P. (2013). A statistical approach to root system classification. *Frontiers in Plant Science*, 4:292.

[Bolin and Doos, 1989] Bolin, B. and Doos, B. R. (1989). Greenhouse effect.

[Borken et al., 2007] Borken, W., Kossmann, G., and Matzner, E. (2007). Biomass, morphology and nutrient contents of fine roots in four norway spruce stands. *Plant and Soil*, 292(1-2):79–93.

[Bradley and Pregitzer, 2007] Bradley, K. L. and Pregitzer, K. S. (2007). Ecosystem assembly and terrestrial carbon balance under elevated co_2. *Trends in Ecology & Evolution*, 22(10):538–547.

[Bressloff, 2014] Bressloff, P. C. (2014). *Stochastic processes in cell biology*, volume 41. Springer.

[Bruce et al., 1999] Bruce, J. P., Frome, M., Haites, E., Janzen, H., Lal, R., and Paustian, K. (1999). Carbon sequestration in soils. *Journal of soil and water conservation*, 54(1):382–389.

[Brunner et al., 2013] Brunner, I., Bakker, M. R., Björk, R. G., Hirano, Y., Lukac, M., Aranda, X., Børja, I., Eldhuset, T. D., Helmisaari, H.-S., Jourdan, C., et al. (2013). Fine-root turnover rates of european forests revisited: an analysis of data from sequential coring and ingrowth cores. *Plant and Soil*, 362(1):357–372.

[Chambers et al., 2004] Chambers, J. Q., Tribuzy, E. S., Toledo, L. C., Crispim, B. F., Higuchi, N., Santos, J. d., Araújo, A. C., Kruijt, B., Nobre, A. D., and Trumbore, S. E. (2004). Respiration from a tropical forest ecosystem: partitioning of sources and low carbon use efficiency. *Ecological Applications*, 14(sp4):72–88.

[Chen et al., 2011] Chen, Y., Dunbabin, V., Postma, J., Diggle, A., Palta, J., Lynch, J., Siddique, K., and Rengel, Z. (2011). Phenotypic variability and modelling of root structure of wild lupinus angustifolius genotypes. *Plant and Soil*, 348(1-2):345.

[Clark et al., 2001a] Clark, D. A., Brown, S., Kicklighter, D. W., Chambers, J. Q., Thomlinson, J. R., and Ni, J. (2001a). Measuring net primary production in forests: concepts and field methods. *Ecological applications*, 11(2):356–370.

[Clark et al., 2001b] Clark, D. A., Brown, S., Kicklighter, D. W., Chambers, J. Q., Thomlinson, J. R., Ni, J., and Holland, E. A. (2001b). Net primary production in tropical forests: an evaluation and synthesis of existing field data. *Ecological applications*, 11(2):371–384.

[Collet et al., 2006] Collet, C., Löf, M., and Pagès, L. (2006). Root system development of oak seedlings analysed using an architectural model. effects of competition with grass. *Plant and Soil*, 279(1):367–383.

[Csilléry et al., 2010] Csilléry, K., Blum, M. G., Gaggiotti, O. E., and François, O. (2010). Approximate bayesian computation (abc) in practice. *Trends in ecology & evolution*, 25(7):410–418.

[De Kauwe et al., 2016] De Kauwe, M. G., Keenan, T. F., Medlyn, B. E., Prentice, I. C., and Terrer, C. (2016). Satellite based estimates underestimate the effect of co_2 fertilization on net primary productivity. *Nature Climate Change*, 6(10):892–893.

[Delucia et al., 2007] Delucia, E. H., Drake, J. E., Thomas, R. B., and Gonzalez-Meler, M. (2007). Forest carbon use efficiency: is respiration a constant fraction of gross primary production? *Global Change Biology*, 13(6):1157–1167.

[DeLucia et al., 1999] DeLucia, E. H., Hamilton, J. G., Naidu, S. L., Thomas, R. B., Andrews, J. A., Finzi, A., Lavine, M., Matamala, R., Mohan, J. E., Hendrey, G. R., et al. (1999). Net primary production of a forest ecosystem with experimental co_2 enrichment. *Science*, 284(5417):1177–1179.

[Didelot et al., 2011] Didelot, X., Everitt, R. G., Johansen, A. M., Lawson, D. J., et al. (2011). Likelihood-free estimation of model evidence. *Bayesian analysis*, 6(1):49–76.

[Diggle, 1988] Diggle, A. (1988). ROOTMAP—a model in three-dimensional coordinates of the growth and structure of fibrous root systems. *Plant and Soil*, 105(2):169–178.

[Dorey, 2010] Dorey, F. (2010). In brief: The p value: What is it and what does it tell you?

[Dotaniya and Meena, 2015] Dotaniya, M. and Meena, V. (2015). Rhizosphere effect on nutrient availability in soil and its uptake by plants: a review. *Proceedings of the National Academy of Sciences, India Section B: Biological Sciences*, 85(1):1–12.

[Downie et al., 2012] Downie, H., Holden, N., Otten, W., Spiers, A. J., Valentine, T. A., and Dupuy, L. X. (2012). Transparent soil for imaging the rhizosphere. *PLoS One*, 7(9):e44276.

[Dunbabin et al., 2013] Dunbabin, V. M., Postma, J. A., Schnepf, A., Pagès, L., Javaux, M., Wu, L., Leitner, D., Chen, Y. L., Rengel, Z., and Diggle, A. J. (2013). Modelling root–soil interactions using three–dimensional models of root growth, architecture and function. *Plant and soil*, 372(1-2):93–124.

[Duursma et al., 2016] Duursma, R. A., Gimeno, T. E., Boer, M. M., Crous, K. Y., Tjoelker, M. G., and Ellsworth, D. S. (2016). Canopy leaf area of a mature evergreen eucalyptus woodland does not respond to elevated atmospheric [co_2] but tracks water availability. *Global change biology*, 22(4):1666–1676.

[Dyson et al., 2014] Dyson, R. J., Vizcay-Barrena, G., Band, L. R., Fernandes, A. N., French, A. P., Fozard, J. A., Hodgman, T. C., Kenobi, K., Pridmore, T. P., Stout, M.,

et al. (2014). Mechanical modelling quantifies the functional importance of outer tissue layers during root elongation and bending. *New Phytologist*, 202(4):1212–1222.

[Edwards et al., 2006] Edwards, E. J., McCaffery, S., and Evans, J. R. (2006). Phosphorus availability and elevated co_2 affect biological nitrogen fixation and nutrient fluxes in a clover-dominated sward. *New Phytologist*, 169(1):157–167.

[Ekanayake and Allen, 2010] Ekanayake, A. J. and Allen, L. J. (2010). Comparison of markov chain and stochastic differential equation population models under higher-order moment closure approximations. *Stochastic analysis and applications*, 28(6):907–927.

[El Zawily et al., 2014] El Zawily, A. M., Schwarzländer, M., Finkemeier, I., Johnston, I. G., Benamar, A., Cao, Y., Gissot, C., Meyer, A. J., Wilson, K., Datla, R., et al. (2014). Friendly regulates mitochondrial distribution, fusion, and quality control in arabidopsis. *Plant Physiology*, pages pp–114.

[Esau et al., 1965] Esau, K. et al. (1965). Plant anatomy. *Plant Anatomy.*, (2nd Edition).

[Eviner, 2004] Eviner, V. (2004). Biogeochemistry of terrestrial net. *Treatise on Geochemistry: Volume 9: Environmental Geochemistry*, 8:215.

[Fahey and Knapp, 2007] Fahey, T. J. and Knapp, A. K. (2007). *Principles and standards for measuring primary production*. Oxford University Press.

[Fang et al., 2018] Fang, J., Yu, G., Liu, L., Hu, S., and Chapin, F. S. (2018). Climate change, human impacts, and carbon sequestration in china. *Proceedings of the National Academy of Sciences*, 115(16):4015–4020.

[Fath, 2018] Fath, B. D. (2018). *Encyclopedia of ecology*. Elsevier.

[Feller et al., 2015] Feller, C., Favre, P., Janka, A., Zeeman, S. C., Gabriel, J.-P., and Reinhardt, D. (2015). Mathematical modeling of the dynamics of shoot-root interactions and resource partitioning in plant growth. *PloS one*, 10(7):e0127905.

[Ferguson and Smucker, 1989] Ferguson, J. and Smucker, A. (1989). Modifications of the minirhizotron video camera system for measuring spatial and temporal root dynamics. *Soil Science Society of America Journal*, 53(5):1601–1605.

[Fernando et al., 2014] Fernando, N., Panozzo, J., Tausz, M., Norton, R. M., Neumann, N., Fitzgerald, G. J., and Seneweera, S. (2014). Elevated co_2 alters grain quality of two bread wheat cultivars grown under different environmental conditions. *Agriculture, Ecosystems & Environment*, 185:24–33.

[Fienberg and Kadane, 2001] Fienberg, S. E. and Kadane, J. B. (2001). *Statistics: The Field*. Elsevier.

[Fitter, 1987] Fitter, A. (1987). An architectural approach to the comparative ecology of plant root systems. *New phytologist*, 106:61–77.

[Flato, 2011] Flato, G. M. (2011). Earth system models: an overview. *Wiley Interdisciplinary Reviews: Climate Change*, 2(6):783–800.

[Fourcaud et al., 2008] Fourcaud, T., Zhang, X., Stokes, A., Lambers, H., and Körner, C. (2008). Plant growth modelling and applications: the increasing importance of plant architecture in growth models. *Annals of Botany*, 101(8):1053–1063.

[Fowler et al., 1997] Fowler, A. C., Fowler, A. C., and Fowler, A. (1997). *Mathematical models in the applied sciences*, volume 17. Cambridge University Press.

[French et al., 2009] French, A., Ubeda-Tomás, S., Holman, T. J., Bennett, M. J., and Pridmore, T. (2009). High-throughput quantification of root growth using a novel image-analysis tool. *Plant physiology*, 150(4):1784–1795.

[Gani, 1984] Gani, J. (1984). Some population and epidemic models revisited. *Mathematical Medicine and Biology: A Journal of the IMA*, 1(3):277–287.

[Garnier et al., 2017] Garnier, E., Stahl, U., Laporte, M.-A., Kattge, J., Mougenot, I., Kühn, I., Laporte, B., Amiaud, B., Ahrestani, F. S., Bönisch, G., et al. (2017). Towards a thesaurus of plant characteristics: an ecological contribution. *Journal of Ecology*, 105(2):298–309.

[Garré et al., 2012] Garré, S., Laloy, E., Javaux, M., Vereecken, H., et al. (2012). Parameterizing a dynamic architectural model of the root system of spring barley from minirhizotron data. *Vadose Zone Journal*, 11(4).

[Geitmann and Ortega, 2009] Geitmann, A. and Ortega, J. K. (2009). Mechanics and modeling of plant cell growth. *Trends in plant science*, 14(9):467–478.

[Gerwitz and Page, 1974] Gerwitz, A. and Page, E. (1974). An empirical mathematical model to describe plant root systems. *Journal of Applied Ecology*, pages 773–781.

[Gifford, 1992] Gifford, R. (1992). Interaction of carbon dioxide with growth-limiting environmental factors in vegetation productivity: implications for the global carbon cycle. In *Advances in Bioclimatology 1*, pages 24–58. Springer.

[Gillespie, 1977] Gillespie, D. T. (1977). Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry*, 81(25):2340–2361.

[Gougoulias et al., 2014] Gougoulias, C., Clark, J. M., and Shaw, L. J. (2014). The role of soil microbes in the global carbon cycle: tracking the below-ground microbial processing of plant-derived carbon for manipulating carbon dynamics in agricultural systems. *Journal of the Science of Food and Agriculture*, 94(12):2362–2371.

[Gower, 2003] Gower, S. T. (2003). Patterns and mechanisms of the forest carbon cycle. *Annual Review of Environment and Resources*, 28(1):169–204.

[Grace, 2004] Grace, J. (2004). Understanding and managing the global carbon cycle. *Journal of Ecology*, 92(2):189–202.

[Guo et al., 2004] Guo, D. L., Mitchell, R. J., and Hendricks, J. J. (2004). Fine root branch orders respond differentially to carbon source-sink manipulations in a longleaf pine forest. *Oecologia*, 140(3):450–457.

[Gutmann et al., 2018] Gutmann, M. U., Dutta, R., Kaski, S., and Corander, J. (2018). Likelihood-free inference via classification. *Statistics and Computing*, 28(2):411–425.

[Hackett and Rose, 1972] Hackett, C. and Rose, D. (1972). A model of the extension and branching of a seminal root of barley, and its use in studying relations between root dimensions i. the model. *Australian Journal of Biological Sciences*, 25(4):669–680.

[Halsey et al., 2015] Halsey, L. G., Curran-Everett, D., Vowler, S. L., and Drummond, G. B. (2015). The fickle p value generates irreproducible results. *Nature methods*, 12(3):179–185.

[Hart, 2019] Hart, K. (2019). Personal communication.

[Hart et al., 2019] Hart, K. M., Curioni, G., Blaen, P., Harper, N. J., Miles, P., Lewin, K. F., Nagy, J., Bannister, E. J., Cai, X. M., Thomas, R. M., et al. (2019). Characteristics of free air carbon dioxide enrichment of a northern temperate mature forest. *Global change biology*.

[Hastings and Palmer, 2003] Hastings, A. and Palmer, M. A. (2003). A bright future for biologists and mathematicians? *Science*, 299(5615).

[Hendrey et al., 1999] Hendrey, G. R., Ellsworth, D. S., Lewin, K. F., and Nagy, J. (1999). A free-air enrichment system for exposing tall forest vegetation to elevated atmospheric co_2. *Global Change Biology*, 5(3):293–309.

[Hendricks et al., 2006] Hendricks, J. J., Hendrick, R. L., Wilson, C. A., Mitchell, R. J., Pecot, S. D., and Guo, D. (2006). Assessing the patterns and controls of fine root dynamics: an empirical test and methodological review. *Journal of Ecology*, 94(1):40–57.

[Hertel and Leuschner, 2002] Hertel, D. and Leuschner, C. (2002). A comparison of four different fine root production estimates with ecosystem carbon balance data in a fagus–quercus mixed forest. *Plant and soil*, 239(2):237–251.

[Hickler et al., 2008] Hickler, T., Smith, B., Prentice, I. C., Mjöfors, K., Miller, P., Arneth, A., and Sykes, M. T. (2008). Co_2 fertilization in temperate face experiments not representative of boreal and tropical forests. *Global Change Biology*, 14(7):1531–1542.

[Houghton, 1996] Houghton, R. (1996). Terrestrial sources and sinks of carbon inferred from terrestrial data. *Tellus B: Chemical and Physical Meteorology*, 48(4):420–432.

[Houghton, 2002] Houghton, R. (2002). Magnitude, distribution and causes of terrestrial carbon sinks and some implications for policy. *Climate Policy*, 2(1):71–88.

[Hoyt, 1979] Hoyt, D. V. (1979). An empirical determination of the heating of the earth by the carbon dioxide greenhouse effect. *Nature*, 282(5737):388–390.

[Inc.,] Inc., W. R. Mathematica, Version 12.0. Champaign, IL, 2019.

[Ioannidis, 2005] Ioannidis, J. P. (2005). Why most published research findings are false. *PLoS medicine*, 2(8):e124.

[Ioannidis, 2018] Ioannidis, J. P. (2018). The proposal to lower p value thresholds to 0.005. *Jama*, 319(14):1429–1430.

[IPCC, 2013] IPCC (2013). *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

[Iversen, 2010] Iversen, C. M. (2010). Digging deeper: fine-root responses to rising atmospheric co_2 concentration in forested ecosystems. *New Phytologist*, 186(2):346–357.

[Iversen et al., 2017] Iversen, C. M., McCormack, M. L., Powell, A. S., Blackwood, C. B., Freschet, G. T., Kattge, J., Roumet, C., Stover, D. B., Soudzilovskaia, N. A., Valverde-Barrantes, O. J., et al. (2017). A global fine-root ecology database to address below-ground challenges in plant ecology. *New Phytologist*, 215(1):15–26.

[Jackson et al., 1997] Jackson, R., Mooney, H., and Schulze, E.-D. (1997). A global budget for fine root biomass, surface area, and nutrient contents. *Proceedings of the National Academy of Sciences*, 94(14):7362–7366.

[Jobbágy and Jackson, 2000] Jobbágy, E. G. and Jackson, R. B. (2000). The vertical distribution of soil organic carbon and its relation to climate and vegetation. *Ecological applications*, 10(2):423–436.

[Johnson et al., 2004] Johnson, D., Cheng, W., Joslin, J., Norby, R., Edwards, N., and Todd, D. (2004). Effects of elevated co_2 on nutrient cycling in a sweetgum plantation. *Biogeochemistry*, 69(3):379–403.

[Johnson and Meyer, 1998] Johnson, M. and Meyer, P. (1998). Mechanical advancing handle that simplifies minirhizotron camera registration and image collection. *Journal of Environmental Quality*, 27(3):710–714.

[Johnson et al., 2001] Johnson, M., Tingey, D., Phillips, D., and Storm, M. (2001). Advancing fine root research with minirhizotrons. *Environmental and Experimental Botany*, 45(3):263–289.

[Johnston and Jones, 2015] Johnston, I. G. and Jones, N. S. (2015). Closed-form stochastic solutions for non-equilibrium dynamics and inheritance of cellular components over many cell divisions. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 471(2180):20150050.

[Johnston et al., 2014] Johnston, I. G., Rickett, B. C., and Jones, N. S. (2014). Explicit tracking of uncertainty increases the power of quantitative rule-of-thumb reasoning in cell biology. *Biophysical journal*, 107(11):2612–2617.

[Karhu et al., 2010] Karhu, K., Fritze, H., Hämäläinen, K., Vanhala, P., Jungner, H., Oinonen, M., Sonninen, E., Tuomi, M., Spetz, P., Kitunen, V., et al. (2010). Temperature sensitivity of soil carbon fractions in boreal forest soil. *Ecology*, 91(2):370–376.

[Keeling, 1997] Keeling, C. D. (1997). Climate change and carbon dioxide: An introduction. *Proceedings of the National Academy of Sciences*, 94(16):8273–8274.

[Key et al., 2004] Key, R. M., Kozyr, A., Sabine, C. L., Lee, K., Wanninkhof, R., Bullister, J. L., Feely, R. A., Millero, F. J., Mordy, C., and Peng, T.-H. (2004). A global ocean carbon climatology: Results from global data analysis project (glodap). *Global biogeochemical cycles*, 18(4).

[King et al., 2005] King, J. S., Kubiske, M. E., Pregitzer, K. S., Hendrey, G. R., McDonald, E. P., Giardina, C. P., Quinn, V. S., and Karnosky, D. F. (2005). Tropospheric o_3 compromises net primary production in young stands of trembling aspen, paper birch and sugar maple in response to elevated atmospheric co_2. *New Phytologist*, 168(3):623–636.

[Kumar et al., 2006] Kumar, R., Pandey, S., and Pandey, A. (2006). Plant roots and carbon sequestration. *Current Science*, pages 885–890.

[Laboratory, 1243] Laboratory, O. R. N. (19/05/17 12:43). Global list of face experiments. https://facedata.ornl.gov/global_face.html#UrbanaPlus.

[Lau et al., 2010] Lau, J., Shaw, R. G., Reich, P. B., and Tiffin, P. (2010). Species interactions in a changing environment: elevated co_2 alters the ecological and potential evolutionary consequences of competition.

[Leakey et al., 2009] Leakey, A. D., Ainsworth, E. A., Bernacchi, C. J., Rogers, A., Long, S. P., and Ort, D. R. (2009). Elevated co_2 effects on plant carbon, nitrogen, and water relations: six important lessons from face. *Journal of experimental botany*, 60(10):2859–2876.

[Leitner et al., 2010] Leitner, D., Klepsch, S., Bodner, G., and Schnepf, A. (2010). A dynamic root system growth model based on l-systems. *Plant and Soil*, 332(1-2):177–192.

[Liu et al., 2018] Liu, Y., Wang, G., Yu, K., Li, P., Xiao, L., and Liu, G. (2018). A new method to optimize root order classification based on the diameter interval of fine root. *Scientific reports*, 8(1):1–11.

[Lobet, 2017] Lobet, G. (2017). Image analysis in plant sciences: Publish then perish. *Trends in Plant Science*.

[Lobet et al., 2013] Lobet, G., Draye, X., and Périlleux, C. (2013). An online database for plant image analysis software tools. *Plant methods*, 9(1):38.

[Lobet et al., 2011] Lobet, G., Pagès, L., and Draye, X. (2011). A novel image-analysis toolbox enabling quantitative analysis of root system architecture. *Plant physiology*, 157(1):29–39.

[Lockhart, 1965] Lockhart, J. A. (1965). An analysis of irreversible plant cell elongation. *Journal of Theoretical Biology*, 8(2):264–275.

[Long et al., 2006] Long, S. P., Ainsworth, E. A., Leakey, A. D., Nösberger, J., and Ort, D. R. (2006). Food for thought: lower-than-expected crop yield stimulation with rising co_2 concentrations. *science*, 312(5782):1918–1921.

[Lukac, 2012] Lukac, M. (2012). Fine root turnover. In *Measuring roots*, pages 363–373. Springer.

[Lungley, 1973] Lungley, D. (1973). The growth of root systems—a numerical computer simulation model. *Plant and Soil*, 38(1):145–159.

[Lüthi et al., 2008] Lüthi, D., Le Floch, M., Bereiter, B., Blunier, T., Barnola, J.-M., Siegenthaler, U., Raynaud, D., Jouzel, J., Fischer, H., Kawamura, K., et al. (2008). High-resolution carbon dioxide concentration record 650,000–800,000 years before present. *Nature*, 453(7193):379–382.

[Lynch et al., 1997] Lynch, J. P., Nielsen, K. L., Davis, R. D., and Jablokow, A. G. (1997). Simroot: modelling and visualization of root systems. *Plant and Soil*, 188(1):139–151.

[MacDougall et al., 2017] MacDougall, A. H., Swart, N. C., and Knutti, R. (2017). The uncertainty in the transient climate response to cumulative co_2 emissions arising from the uncertainty in physical climate parameters. *Journal of Climate*, 30(2):813–827.

[Madhu and Hatfield, 2013] Madhu, M. and Hatfield, J. (2013). Dynamics of plant root growth under increased atmospheric carbon dioxide. *Agronomy Journal*, 105(3):657–669.

[Mairhofer et al., 2013] Mairhofer, S., Zappala, S., Tracy, S., Sturrock, C., Bennett, M. J., Mooney, S. J., and Pridmore, T. P. (2013). Recovering complete plant root system architectures from soil via x-ray μ-computed tomography. *Plant Methods*, 9(1):1.

[Mairhofer et al., 2012] Mairhofer, S., Zappala, S., Tracy, S. R., Sturrock, C., Bennett, M., Mooney, S. J., and Pridmore, T. (2012). Rootrak: automated recovery of three-dimensional plant root architecture in soil from x-ray microcomputed tomography images using visual tracking. *Plant physiology*, 158(2):561–569.

[Majdi, 1996] Majdi, H. (1996). Root sampling methods-applications and limitations of the minirhizotron technique. *Plant and Soil*, 185(2):255–258.

[Majdi et al., 2005] Majdi, H., Pregitzer, K., Moren, A.-S., Nylund, J.-E., and Ågren, G. I. (2005). Measuring fine root turnover in forest ecosystems. *Plant and soil*, 276(1-2):1–8.

[Makkonen and Helmisaari, 1999] Makkonen, K. and Helmisaari, H.-S. (1999). Assessing fine-root biomass and production in a scots pine stand–comparison of soil core and root ingrowth core methods. *Plant and Soil*, 210(1):43–50.

[Maria do Rosário et al., 2000] Maria do Rosário, G. O., Van Noordwijk, M., Gaze, S., Brouwer, G., Bona, S., Mosca, G., and Hairiah, K. (2000). Auger sampling, ingrowth cores and pinboard methods. In *Root methods*, pages 175–210. Springer.

[Marjoram et al., 2003] Marjoram, P., Molitor, J., Plagnol, V., and Tavaré, S. (2003). Markov chain monte carlo without likelihoods. *Proceedings of the National Academy of Sciences*, 100(26):15324–15328.

[Marklein and Houlton, 2012] Marklein, A. R. and Houlton, B. Z. (2012). Nitrogen inputs accelerate phosphorus cycling rates across a wide variety of terrestrial ecosystems. *New Phytologist*, 193(3):696–704.

[Matamala et al., 2003] Matamala, R., Gonzalez-Meler, M. A., Jastrow, J. D., Norby, R. J., and Schlesinger, W. H. (2003). Impacts of fine root turnover on forest npp and soil c sequestration potential. *Science*, 302(5649):1385–1387.

[Matis and Kiffe, 2000] Matis, J. H. and Kiffe, T. R. (2000). Nonlinear birth-death-migration models. In *Stochastic Population Models*, pages 161–171. Springer.

[McLeod and Long, 1999] McLeod, A. and Long, S. (1999). Free-air carbon dioxide enrichment (face) in global change research: a review. In *Advances in ecological research*, volume 28, pages 1–56. Elsevier.

[Medlyn et al., 2016] Medlyn, B. E., De Kauwe, M. G., Zaehle, S., Walker, A. P., Duursma, R. A., Luus, K., Mishurov, M., Pak, B., Smith, B., Wang, Y.-P., et al. (2016). Using models to guide field experiments: a priori predictions for the co_2 response of a nutrient-and water-limited native eucalypt woodland. *Global change biology*, 22(8):2834–2851.

[Metzner et al., 2015] Metzner, R., Eggert, A., van Dusschoten, D., Pflugfelder, D., Gerth, S., Schurr, U., Uhlmann, N., and Jahnke, S. (2015). Direct comparison of mri and x-ray ct technologies for 3d imaging of root systems in soil: potential and challenges for root trait quantification. *Plant methods*, 11(1):17.

[Min and Guangsheng, 2004] Min, Z. and Guangsheng, Z. (2004). Carbon storage of forest vegetation and its relationship with climatic factors. *Scientia Geographica Sinica*, 24(1):50–54.

[Morgan, 2008] Morgan, B. J. (2008). *Applied stochastic modelling*. CRC press.

[Morris et al., 2017] Morris, E. C., Griffiths, M., Golebiowska, A., Mairhofer, S., Burr-Hersey, J., Goh, T., Von Wangenheim, D., Atkinson, B., Sturrock, C. J., Lynch, J. P., et al. (2017). Shaping 3d root system architecture. *Current Biology*, 27(17):R919–R930.

[Myers et al., 2014] Myers, S. S., Zanobetti, A., Kloog, I., Huybers, P., Leakey, A. D., Bloom, A. J., Carlisle, E., Dietterich, L. H., Fitzgerald, G., Hasegawa, T., et al. (2014). Increasing co_2 threatens human nutrition. *Nature*, 510(7503):139–142.

[Nabuurs et al., 2013] Nabuurs, G.-J., Lindner, M., Verkerk, P. J., Gunia, K., Deda, P., Michalak, R., and Grassi, G. (2013). First signs of carbon sink saturation in european forest biomass. *Nature Climate Change*, 3(9):792–796.

[Norby et al., 2016] Norby, R. J., De Kauwe, M. G., Domingues, T. F., Duursma, R. A., Ellsworth, D. S., Goll, D. S., Lapola, D. M., Luus, K. A., MacKenzie, A. R., Medlyn, B. E., et al. (2016). Model–data synthesis for the next generation of forest free-air co_2 enrichment (face) experiments. *New Phytologist*, 209(1):17–28.

[Norby et al., 2002] Norby, R. J., Hanson, P. J., O'Neill, E. G., Tschaplinski, T. J., Weltzin, J. F., Hansen, R. A., Cheng, W., Wullschleger, S. D., Gunderson, C. A., Edwards, N. T., et al. (2002). Net primary productivity of a co_2-enriched deciduous forest and the implications for carbon storage. *Ecological Applications*, 12(5):1261–1266.

[Norby and Jackson, 2000] Norby, R. J. and Jackson, R. B. (2000). Root dynamics and global change: seeking an ecosystem perspective. *The New Phytologist*, 147(1):3–12.

[Norby et al., 2004] Norby, R. J., Ledford, J., Reilly, C. D., Miller, N. E., and O'Neill, E. G. (2004). Fine-root production dominates response of a deciduous forest to atmospheric co_2 enrichment. *Proceedings of the national Academy of Sciences*, 101(26):9689–9693.

[Norby et al., 2001] Norby, R. J., Todd, D. E., Fults, J., and Johnson, D. W. (2001). Allometric determination of tree growth in a co_2-enriched sweetgum stand. *New Phytologist*, 150(2):477–487.

[Norby et al., 2010] Norby, R. J., Warren, J. M., Iversen, C. M., Medlyn, B. E., and McMurtrie, R. E. (2010). Co_2 enhancement of forest productivity constrained by limited nitrogen availability. *Proceedings of the National Academy of Sciences*, 107(45):19368–19373.

[Norby et al., 1999] Norby, R. J., Wullschleger, S. D., Gunderson, C. A., Johnson, D. W., and Ceulemans, R. (1999). Tree responses to rising co_2 in field experiments: implications for the future forest. *Plant, Cell & Environment*, 22(6):683–714.

[Norby and Zak, 2011] Norby, R. J. and Zak, D. R. (2011). Ecological lessons from free-air co_2 enrichment (face) experiments. *Annual review of ecology, evolution, and systematics*, 42:181–203.

[Norton et al., 2008] Norton, R., Mollah, M., Fitzgerald, G., and McNeil, D. (2008). The australian grains free air carbon dioxide enrichment (agface) experiment–specifications and scope. In *Proceedings of the 14th Australian Agronomy Conference*.

[Novozhilov et al., 2006] Novozhilov, A. S., Karev, G. P., and Koonin, E. V. (2006). Biological applications of the theory of birth-and-death processes. *Briefings in bioinformatics*, 7(1):70–85.

[Nowak et al., 2004] Nowak, R. S., Ellsworth, D. S., and Smith, S. D. (2004). Functional responses of plants to elevated atmospheric co_2–do photosynthetic and productivity data from face experiments support early predictions? *New phytologist*, 162(2):253–280.

[Nuzzo, 2014] Nuzzo, R. (2014). Scientific method: statistical errors. *Nature News*, 506(7487):150.

[O'connell et al., 2003] O'connell, K. E., Gower, S. T., and Norman, J. M. (2003). Net ecosystem production of two contrasting boreal black spruce forest communities. *Ecosystems*, 6(3):248–260.

[Ozier-Lafontaine et al., 1999] Ozier-Lafontaine, H., Lecompte, F., and Sillon, J. F. (1999). Fractal analysis of the root architecture of gliricidia sepium for the spatial prediction of root branching, size and mass: model development and evaluation in agroforestry. *Plant and Soil*, 209(2):167–179.

[Pagès et al., 2000] Pagès, L., Asseng, S., Pellerin, S., and Diggle, A. (2000). Modelling root system growth and architecture. In *Root Methods*, pages 113–146. Springer.

[Pages et al., 1989] Pages, L., Jordan, M.-O., and Picard, D. (1989). A simulation model of the three-dimensional architecture of the maize root system. *Plant and Soil*, 119(1):147–154.

[Pagès and Pellerin, 1994] Pagès, L. and Pellerin, S. (1994). Evaluation of parameters describing the root system architecture of field grown maize plants (zea mays l.). *Plant and Soil*, 164(2):169–176.

[Pagès et al., 2004] Pagès, L., Vercambre, G., Drouet, J.-L., Lecompte, F., Collet, C., and Le Bot, J. (2004). Root typ: a generic model to depict and analyse the root system architecture. *Plant and Soil*, 258(1):103–119.

[Pan et al., 2011] Pan, Y., Birdsey, R. A., Fang, J., Houghton, R., Kauppi, P. E., Kurz, W. A., Phillips, O. L., Shvidenko, A., Lewis, S. L., Canadell, J. G., et al. (2011). A large and persistent carbon sink in the world's forests. *Science*, 333(6045):988–993.

[Pearce, 1992] Pearce, S. (1992). Introduction to fisher (1925) statistical methods for research workers. In *Breakthroughs in statistics*, pages 59–65. Springer.

[Pepin and Körner, 2002] Pepin, S. and Körner, C. (2002). Web-face: a new canopy free-air co_2 enrichment system for tall trees in mature forests. *Oecologia*, 133(1):1–9.

[Péret et al., 2009] Péret, B., Larrieu, A., and Bennett, M. J. (2009). Lateral root emergence: a difficult birth. *Journal of experimental botany*, 60(13):3637–3643.

[Phillips et al., 2000] Phillips, D. L., Johnson, M. G., Tingey, D. T., Biggart, C., Nowak, R. S., and Newsom, J. C. (2000). Minirhizotron installation in sandy, rocky soils with minimal soil disturbance. *Soil Science Society of America Journal*, 64(2):761–764.

[Pierret, 2008] Pierret, A. (2008). Multi-spectral imaging of rhizobox systems: new perspectives for the observation and discrimination of rhizosphere components. *Plant and Soil*, 310(1-2):263–268.

[Postma and Black, 2020] Postma, J. A. and Black, K. (2020). Ch02-advances in root architectural modeling during the last decade. Technical report, Pflanzenwissenschaften.

[Pound et al., 2013] Pound, M. P., French, A. P., Atkinson, J. A., Wells, D. M., Bennett, M. J., and Pridmore, T. (2013). Rootnav: navigating images of complex root architectures. *Plant Physiology*, 162(4):1802–1814.

[Pritchard et al., 2008] Pritchard, S. G., Strand, A. E., McCORMACK, M. L., Davis, M. A., Finzi, A. C., Jackson, R. B., Matamala, R., Rogers, H. H., and Oren, R. (2008). Fine root dynamics in a loblolly pine forest are influenced by free-air-co_2-enrichment: a six-year-minirhizotron study. *Global Change Biology*, 14(3):588–602.

[R Core Team, 2020] R Core Team (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

[Rahman et al., 2020] Rahman, G., Sohag, H., Chowdhury, R., Wahid, K. A., Dinh, A., Arcand, M., and Vail, S. (2020). Soilcam: A fully automated minirhizotron using multispectral imaging for root activity monitoring. *Sensors*, 20(3):787.

[Reinhold-Hurek et al., 2015] Reinhold-Hurek, B., Bünger, W., Burbano, C. S., Sabale, M., and Hurek, T. (2015). Roots shaping their microbiome: global hotspots for microbial activity. *Annual review of phytopathology*, 53:403–424.

[Rogers et al., 1999] Rogers, H. H., Runion, G. B., and Prior, A. (1999). Response of plants to elevated atmospheric co$_2$: Root growth, mineral. *Carbon dioxide and environmental stress*, page 215.

[Roose et al., 2016] Roose, T., Keyes, S., Daly, K., Carminati, A., Otten, W., Vetterlein, D., and Peth, S. (2016). Challenges in imaging and predictive modeling of rhizosphere processes. *Plant and Soil*, 407(1-2):9–38.

[Rytter and Rytter, 2012] Rytter, R.-M. and Rytter, L. (2012). Quantitative estimates of root densities at minirhizotrons differ from those in the bulk soil. *Plant and Soil*, 350(1-2):205–220.

[Samson and Sinclair, 1994] Samson, B. K. and Sinclair, T. R. (1994). Soil core and minirhizotron comparison for the determination of root length density. *Plant and Soil*, 161(2):225–232.

[Santantonio and Hermann, 1985] Santantonio, D. and Hermann, R. (1985). Standing crop, production, and turnover of fine roots on dry, moderate, and wet sites of mature douglas-fir in western oregon. In *Annales des Sciences Forestieres*, volume 42, pages 113–142. EDP Sciences.

[Sawyer, 1972] Sawyer, J. S. (1972). Man-made carbon dioxide and the "greenhouse" effect. *Nature*, 239(5366):23–26.

[Schlesinger et al., 2006] Schlesinger, W., Bernhardt, E., DeLucia, E., Ellsworth, D., Finzi, A., Hendrey, G., Hofmockel, K., Lichter, J., Matamala, R., Moore, D., et al. (2006). The duke forest face experiment: Co$_2$ enrichment of a loblolly pine forest. In *Managed Ecosystems and CO_2*, pages 197–212. Springer.

[Schnepf et al., 2018] Schnepf, A., Leitner, D., Landl, M., Lobet, G., Mai, T. H., Morandage, S., Sheng, C., Zörner, M., Vanderborght, J., and Vereecken, H. (2018). Crootbox: a structural–functional modelling framework for root systems. *Annals of botany*, 121(5):1033–1053.

[Shahzad and Amtmann, 2017] Shahzad, Z. and Amtmann, A. (2017). Food for thought: how nutrients regulate root system architecture. *Current opinion in plant biology*, 39:80–87.

[Sievänen et al., 2014] Sievänen, R., Godin, C., DeJong, T. M., and Nikinmaa, E. (2014). Functional–structural plant models: a growing paradigm for plant studies. *Annals of botany*, 114(4):599–603.

[Swain et al., 2002] Swain, P. S., Elowitz, M. B., and Siggia, E. D. (2002). Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proceedings of the National Academy of Sciences*, 99(20):12795–12800.

[Tatsumi et al., 1989] Tatsumi, J., Yamauchi, A., and Kono, Y. (1989). Fractal analysis of plant root systems. *Annals of Botany*, 64(5):499–503.

[Taylor et al., 2013] Taylor, B. N., Beidler, K. V., Cooper, E. R., Strand, A. E., and Pritchard, S. G. (2013). Sampling volume in root studies: the pitfalls of under-sampling exposed using accumulation curves. *Ecology letters*, 16(7):862–869.

[Taylor et al., 2014] Taylor, B. N., Beidler, K. V., Strand, A. E., and Pritchard, S. G. (2014). Improved scaling of minirhizotron data using an empirically-derived depth of field and correcting for the underestimation of root diameters. *Plant and Soil*, 374(1-2):941–948.

[Taylor et al., 1990] Taylor, H., Upchurch, D., and McMichael, B. (1990). Applications and limitations of rhizotrons and minirhizotrons for root studies. *Plant and Soil*, 129(1):29–35.

[Thiese et al., 2016] Thiese, M. S., Ronna, B., and Ott, U. (2016). P value interpretations and considerations. *Journal of thoracic disease*, 8(9):E928.

[Thilakarathne et al., 2015] Thilakarathne, C. L., Tausz-Posch, S., Cane, K., Norton, R. M., Fitzgerald, G. J., Tausz, M., and Seneweera, S. (2015). Intraspecific variation in leaf growth of wheat (triticum aestivum) under australian grain free air co_2 enrichment (agface): is it regulated through carbon and/or nitrogen supply? *Functional Plant Biology*, 42(3):299–308.

[Tierney and Fahey, 2001] Tierney, G. L. and Fahey, T. J. (2001). Evaluating minirhizotron estimates of fine root longevity and production in the forest floor of a temperate broadleaf forest. *Plant and Soil*, 229(2):167–176.

[Toni et al., 2009] Toni, T., Welch, D., Strelkowa, N., Ipsen, A., and Stumpf, M. P. (2009). Approximate bayesian computation scheme for parameter inference and model selection in dynamical systems. *Journal of the Royal Society Interface*, 6(31):187–202.

[Upchurch and Ritchie, 1984] Upchurch, D. and Ritchie, J. (1984). Battery-operated color video camera for root observations in mini-rhizotrons 1. *Agronomy journal*, 76(6):1015–1017.

[Vamerali et al., 2012] Vamerali, T., Bandiera, M., and Mosca, G. (2012). Minirhizotrons in modern root studies. In *Measuring roots*, pages 341–361. Springer.

[van Dusschoten et al., 2016] van Dusschoten, D., Metzner, R., Kochs, J., Postma, J. A., Pflugfelder, D., Bühler, J., Schurr, U., and Jahnke, S. (2016). Quantitative 3D analysis

of plant roots growing in soil using magnetic resonance imaging. *Plant physiology*, 170(3):1176–1188.

[Van Kampen, 1992] Van Kampen, N. G. (1992). *Stochastic processes in physics and chemistry*, volume 1. Elsevier.

[Vannoppen et al., 2015] Vannoppen, W., Vanmaercke, M., De Baets, S., and Poesen, J. (2015). A review of the mechanical effects of plant roots on concentrated flow erosion rates. *Earth-Science Reviews*, 150:666–678.

[Verheijen et al., 2016] Verheijen, L. M., Aerts, R., Bönisch, G., Kattge, J., and Van Bodegom, P. M. (2016). Variation in trait trade-offs allows differentiation among predefined plant functional types: implications for predictive ecology. *New Phytologist*, 209(2):563–575.

[Wang et al., 1995] Wang, Z., Burch, W. H., Mou, P., Jones, R. H., and Mitchell, R. J. (1995). Accuracy of visible and ultraviolet light for estimating live root proportions with minirhizotrons. *Ecology*, 76(7):2330–2334.

[Waring, 2007] Waring, R. (2007). Advances in eddy-flux analyses, remote sensing, and evidence of climate change. *Forest ecosystems: analysis at multiple scales*, pages 317–344.

[Wilkinson, 2009] Wilkinson, D. J. (2009). Stochastic modelling for quantitative description of heterogeneous biological systems. *Nature Reviews Genetics*, 10(2):122–133.

[Withington et al., 2003] Withington, J. M., Elkin, A. D., Bułaj, B., Olesiński, J., Tracy, K. N., Bouma, T. J., Oleksyn, J., Anderson, L. J., Modrzyński, J., Reich, P. B., et al. (2003). The impact of material used for minirhizotron tubes for root research. *New Phytologist*, 160(3):533–544.

[Wittwer and Robb, 1964] Wittwer, S. and Robb, W. (1964). Carbon dioxide enrichment of greenhouse atmospheres for food crop production. *Economic Botany*, 18(1):34–56.

[Wrb, 2015] Wrb, I. W. G. (2015). World reference base for soil resources 2014, update 2015: International soil classification system for naming soils and creating legends for soil maps.

[Wu et al., 2005] Wu, L., McGechan, M., Watson, C., and Baddeley, J. (2005). Developing existing plant root system architecture models to meet future agricultural challenges. *Advances in Agronomy*, 85:181–219.

[Yang et al., 2017] Yang, M., Défossez, P., Danjon, F., Dupont, S., and Fourcaud, T. (2017). Which root architectural elements contribute the best to anchorage of pinus species? insights from in silico experiments. *Plant and soil*, 411(1-2):275–291.

[Yuan and Chen, 2012] Yuan, Z. and Chen, H. Y. (2012). Indirect methods produce higher estimates of fine root production and turnover rates than direct methods. *PLoS One*, 7(11):e48989.

Appendix A

Outreach and deliverables

This section will outline deliverables and outreach events that did not form part of formal research.

A.1 Conference presentations

- The impact of elevated carbon dioxide on fine root growth in a temperate oak forest, *Time temperature and a transforming world*, SEB Seville 2019
- Likelihood-free inference reveals physical mechanisms and parameters governing root architecture, *In silico plants*, SEB Seville 2019
- Quantitative modelling of root growth and carbon allocation: bridging theory and experiment, *International Minirhizotron Working Group meeting*, Oak Ridge, Tennessee 2019
- Using stochastic modelling to investigate fine root growth under elevated CO_2, *Long-term ecological experiments in plant-soil ecosystems*, BES Buxton 2019
- Quantitative modelling of root growth and carbon allocation bridging theory and experiment, *BIFoR Annual Meeting*, UoB 2019

A.2 Awards

- Winner, UoB Research Poster Conference 2017
- Best talk, UoB BGRS Symposium 2019
- Runner up, UoB 3 Minute Thesis competition

A.3 Articles and videos

- Sampling in the Dark: Challenges in Fine-Root Research, *EOS*, 2019
- Elevated carbon dioxide: Getting to the root of the problem, *Botany One*, 2018
- Filmed Rapid Researcher segment for Old Joe alumni magazine, 2017

A.4 Outreach

- Gave short presentation of my research to HRH Prince Charles on his visit to the BIFoR research centre, 2018
- Appeared on BBC Midlands Today as part of the above visit, giving brief outline of my work
- Helped organise a Birmingham Pint of Science event, 2018
- Presented research at many BIFoR open days 2017 - 2019
- Exhibitor for BIFoR's stand at a Next Generation Innovators day at Malvern 3 Counties Showground, 2018

A.5 Other deliverables

- Supervised volunteers while root imaging at the BIFoR research site
- Managed larger groups of volunteers taking soil cores at BIFoR, and sorting roots in a lab at UoB.
- Participated in MSB-Net Multiscale Biology study groups in the UoB mathematics department in 2016 and 2017
- Teaching Assistant work in both Mathematics and Biosciences
- Worked in the Maths Support Centre in the university library

Appendix B

Published paper

Model selection and parameter estimation for root architecture models using likelihood-free inference

Clare Ziegler[1,2], Rosemary J. Dyson[2,3], Iain G. Johnston[1,2,4,*]

[1] School of Biosciences, University of Birmingham, UK

[2] Birmingham Institute of Forest Research, University of Birmingham, UK

[3] School of Mathematics, University of Birmingham, UK

[4] Alan Turing Institute, London, UK

[*] Correspondence to i.johnston.1@bham.ac.uk

Keywords: likelihood-free inference, root system architecture, root growth, approximate Bayesian computation

Abstract

Plant root systems play vital roles in the biosphere, environment, and agriculture, but the quantitative principles governing their growth and architecture remain poorly understood. The 'forward problem' of what root forms can arise from given models and parameters has been well studied through modelling and simulation, but comparatively little attention has been given to the 'inverse problem': what models and parameters are responsible for producing an experimentally observed root system? Here, we propose the use of approximate Bayesian computation (ABC) to infer mechanistic parameters governing root growth and architecture, allowing us to learn and quantify uncertainty in parameters and model structures using observed root architectures. We demonstrate the use of this platform on synthetic and experimental root data and show how it may be used to identify growth mechanisms and characterise growth parameters in different mutants. Our highly adaptable framework can be used to gain mechanistic insight into the generation of observed root system architectures.

B.1 Introduction

Root systems are essential to plants' structure and uptake of water and nutrients, and constitute more than 5% by mass of the total global carbon budget [Jackson et al., 1997]. They stabilise plants [Yang et al., 2017], stabilise soils [Vannoppen et al., 2015], foster beneficial microbes [Reinhold-Hurek et al., 2015] and are the entry point for water and nutrients to the plant [Dotaniya and Meena, 2015]. The shape of a plant's root system is generated by a variety of physiological and signalling pathways within the plant, and understanding the generation of this system opens paths to its optimisation to maximise crop yield [Shahzad and Amtmann, 2017].

Despite this importance, the mechanisms underlying root growth remain challenging to quantitatively understand [Dunbabin et al., 2013, Roose et al., 2016, Sievänen et al., 2014]. The complexity of root systems and their belowground nature poses observational challenges. Experimental techniques aiming to elucidate root architecture have historically included sketches of root systems and the use of hydroponics, then images of cleaned root systems. More recent advances have facilitated the imaging of plants *in situ* through the use of x-ray μ-Computed Tomography [Mairhofer et al., 2013], MRI scanning [Metzner et al., 2015, van Dusschoten et al., 2016], and transparent soil [Downie et al., 2012], which have been used to investigate root soil exploration and uptake of water and nutrients.

In parallel with this experimental elucidation, in an effort to understand how root systems grow, many physical and mathematical models of root growth and structure have been produced. These models solve the 'forward problem': given knowledge of the parameters governing growth processes in plants, they produce the details and dynamics of a likely simulated root system. The Lockhart equation described the elongation of a cell under turgor pressure [Lockhart, 1965], and has been widely adapted to describe the growth of many plant organs, including roots [Dyson et al., 2014, Geitmann and Ortega, 2009]. Hackett and Rose produced the first root system model in the 1970s [Hackett and Rose, 1972] based on the growth and branching of barley roots, while Lungley [Lungley, 1973] produced a computational model which generated root systems represented using ASCII characters. Fitter [Fitter, 1987] introduced a topological model of root architecture where a root system was considered as a set of links. This idea was extended in the three dimensional model of Pagès [Pages et al., 1989] and Diggle, whose ROOTMAP model could be applied to a variety of plant species [Diggle, 1988], while Tatsumi *et al.* represented variation in root systems using fractal analysis [Tatsumi et al., 1989, Pagès et al., 2000] Lynch *et al.* modelled a root system as a network with nodes as branches and inter-branch distances as edges. The model also included root radius, and volume changes along the growing root [Lynch et al., 1997]. Advances in computation have led to a plethora of root system architecture models, which produce a three dimensional reproduction of a root system using a detailed parameter set [Dunbabin et al., 2013]. Previous modelling approaches were combined in the production of Root Typ in 2004 [Pagès et al., 2004]. This has allowed the model to be adapted for use by other researchers [Collet et al., 2006, Garré et al., 2012]. Another key root architecture model is RootBox, [Leitner et al., 2010] which is designed to be combined with soil and water uptake models along with allowing for the simulation of roots grown in containers of user-defined shape

and dimensions. This model has been recently updated to produce CRootBox, and there are plans to eventually extend the modelling approach to also consider above-ground plant growth [Schnepf et al., 2018]. An effective model is able to reproduce root systems of many different plant species, which necessitates the incorporation of root data collected *in situ* [Roose et al., 2016].

While a great number of advances have been made in simulating root systems from a set of parameters, relatively little work has been done on the inverse problem: extracting the growth parameters from an observed root system. Model parameterisation is often limited by difficulties in root observation [Garré et al., 2012, Pagès and Pellerin, 1994], however, this step is crucial to gain biological insight from these root models. It is vital that we can validate root models and the predictions they make, and quantify uncertainty in their mechanisms and parameters. A manual approach to the inverse problem, feeding specific measured parameters into generative models for root systems and assessing their ability to reproduce observations, has been used to gain biological insight and validate generative models [Chen et al., 2011]. However, without an automated approach, it remains challenging to explore the full ranges of parameters and mechanisms that could give rise to observed structures, and the likelihoods of each. Advancing technologies are allowing observation of root systems in increasing detail, making it ever more important to bridge the gap between theory and observation.

A major challenge in solving the inverse problem with traditional statistical methods is finding a likelihood function for an observed root system. Modern statistical approaches allow this challenge to be circumvented through the use of stochastic simulation and Approximate Bayesian Computation (ABC) techniques [Beaumont et al., 2002], which produce a computational approximation replacing the likelihood and remove the need for its explicit calculation. Another strength of these techniques lies in their natural capacity for model selection and the inclusion of prior knowledge about the system in an inference setting. Here, we report a novel pipeline by which ABC, embedded in an Sequential Monte Carlo (SMC) framework [Toni et al., 2009], can be used to learn the values of and uncertainty in generative, mechanistic parameters underlying root growth and architecture, and to compare different root architecture models. *Arabidopsis thaliana* (thale cress) is used in both computational and experimental investigation throughout as a model plant, but this process can readily be extended to any root system, as we also demonstrate with *Lupinus angustifolius* (narrowleaf lupin). We demonstrate how this framework can be used to identify generative parameters according to a given model, distinguish phenotypic differences, and evaluate the comparative effectiveness of models for root elongation and root branching processes, providing insight into the underlying mechanisms.

B.2 Results

B.2.1 An ABC SMC framework for inferring mechanistic parameters from root systems

For generality, we begin by considering a highly simplified model for root growth (see Methods). Starting from an infinitesimal initial condition, a primary root elongates according to a growth law. Branches from this primary root occur stochastically according to a branching law. Branches elongate according to the primary root growth law multiplicatively scaled (allowing, for example, branches to grow at a slower rate than the primary root). Branching is for now restricted to first-order branches from the primary root, though nothing in our framework is dependent on this or any other structural choice.

This coarse-grained model was chosen to reflect the core behaviour shared at the intersection of several contemporary root models [Pagès et al., 2004, Leitner et al., 2010, Schnepf et al., 2018]. Its computational simplicity is an advantage but not a necessity for our inference framework; we later consider an alternative generative model to demonstrate the transferrability of our approach. The details of the model are described in Methods, but in this section we consider constructing the inference framework for a general mechanistic model, the parameters of which we denote θ.

The platform proceeds by simulating outputs from this model with different trial parameterisations, using a distance function to compare these outputs to summary statistics of experimental observations, and iterating this process within a Bayesian framework to build up posterior distributions on model structures and parameters given the observed data.

To compare simulation to experiment, we focus on a mechanistically informative set of summary statistics. For a given root structure d, these are the number of branches B, length of the primary root L, and average length of the lateral roots $\hat{l}$. Within our scheme, these lengths are for convenience measured in cm, but different scalings of these features can be used to emphasise different aspects of root architecture in the simulation-data comparison. The distance function we use to compare two structures d_1, d_2 is based on the Euclidean distance:

$$\rho(d_1, d_2) = \frac{1}{3}\left((B_1 - B_2)^2 + (L_1 - L_2)^2 + (\hat{l}_1 - \hat{l}_2)^2\right). \tag{B.1}$$

A dataset $\mathcal{D}$ may consist of a set of structures d_{ij}, where i labels individual plants and j labels longitudinal observations. In this case, for each observed plant i, a model plant is simulated and its structure recorded at each of the times corresponding to the longitudinal observations. We will call these recorded structures d'_{ij} and are interested in the comparison between each recorded structure and its observed counterpart:

$$\rho(\mathcal{D}, \mathcal{D}') = \sum_i \sum_{j|i} \rho(d_{ij}, d'_{ij}). \tag{B.2}$$

We deliberately choose this model and summary statistics to focus on the topological aspects of root architecture and ignore any specific physical embeddings (for example, branching angles). This focus on topological degrees of freedom increases the generality of the approach, but features like branching angles and higher-order topological statistics can readily be included in the modelling and distance calculation if they reflect important degrees of freedom for the scientific question under consideration.

Eqn. B.1 balances the ability to capture fine detail of the root system against the computational time required to obtain a reasonable number of samples from the posterior. Including more detail and/or degrees of freedom in the distance function will allow more detailed matching of observations, but will increase the sampling effort required to find regions of parameter space that match these criteria.

ABC involves accepting a trial set of parameters as a sample from the posterior distribution when $\rho(\mathcal{D}, \mathcal{D}') < \epsilon$, or in other words when the summary statistics of the structure emerging from simulation are 'close' to those arising from the experimental data. The posterior distribution on parameters θ built up from a set of samples taken in this way is $P(\theta|\mathcal{D}; \rho < \epsilon)$, which forms an increasingly good approximation to the true posterior $P(\theta|\mathcal{D})$ as ϵ is decreased [Csilléry et al., 2010].

For parametric inference within a fixed model, a simple rejection-sampling pipeline is then given by Algorithm 1 (see Methods). This approach would be sufficient to identify generative parameters from data, but rejection sampling is an inefficient paradigm, as any 'good' regions of parameter space are immediately forgotten when the next draw from the prior is made. To facilitate more efficient parametric inference as well as model selection, we use ABC embedded in a sequential Monte Carlo (SMC) framework as in Toni *et al.* [Toni et al., 2009]. ABC SMC first enforces only a relaxed fit to the data then sequentially uses the inferred parameter distributions as priors while enforcing a tighter fit to data. This sequential process is parameterised by a sequence of ϵ values describing the fit threshold required at each step in the sequence. Model selection can proceed by including a 'model index' parameter describing which model structure is to be used, applying a prior to this parameter (thus incorporating prior knowledge about which model structures are more likely), then treating this index as a parameter to be inferred through SMC. Following Toni *et al.* [Toni et al., 2009], the coupled inference and model selection pipeline is then given by Algorithm 2 (see Methods).

B.2.2 Inferring parameters from a simulated root system

We first sought to test the applicability of our likelihood-free inference process on synthetic root data, to confirm its ability to identify known generative parameters. To this end, the CRootBox root simulation model [Schnepf et al., 2018] was used to produce an example of an *Arabidopsis thaliana* root system. The governing parameters were mean growth rates of 0.49 cm day^{-1} for the primary root, 0.08 cm day^{-1} for the lateral roots, and an inter-lateral distance of 0.2cm, although inter-lateral distance is not an explicit parameter in our model (see next section). CRootBox adds an element of stochasticity to its generative parameters; in the default *Arabidopsis* case this corresponded to a coefficient of variation

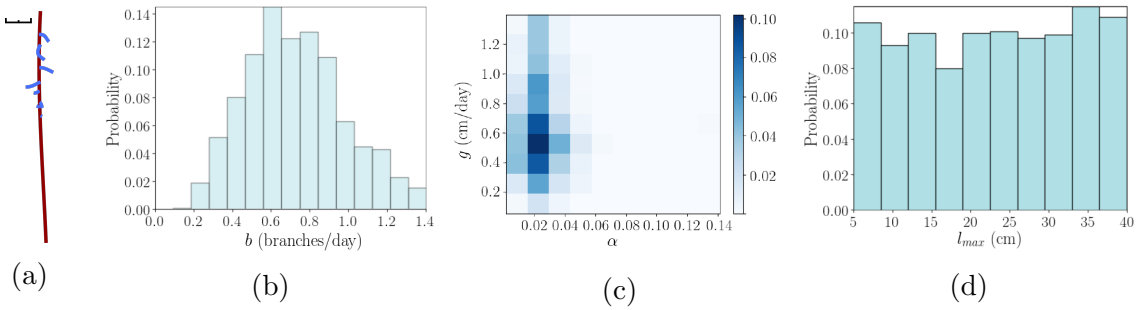


Figure B.1: **Validating root inference platform with synthetic *Arabidopsis* data**. (a) Output from CRootBox simulation of *Arabidopsis* root growth; black scale bar is 1cm. (b-d) Output posteriors from an ABC SMC framework run on CRootBox output, with final ABC SMC tolerance $\epsilon = 0.5$ (see Methods). (b) Posterior distribution on branching rate b in the growth model. (c) Two-dimensional posterior on primary root growth rate g and lateral root growth scaling α. (d) Posterior distribution on l_{max}, the maximum length parameter in the negative-exponential growth model used.

of 0.1 in the growth rates and 0.45 in the lateral spacing parameters. The simulation was run over 15 simulated days, yielding the structure in Fig. B.1a).

To mirror the pipeline that will be used for experimental data, we analysed this simulation output with SmartRoot image analysis software [Lobet et al., 2011], obtaining the statistics of tap and lateral root length and placement. We then applied our ABC SMC framework to estimate posterior distributions on the mechanistic parameters of our simple growth model (see Methods). These parameters are g (primary root growth rate), l_{max} (primary root scaling constant), b (branching rate), and α (lateral root growth scaling).

As shown in Fig. B.1b)-d), the growth rates and, notably, their variability are well captured in the resultant posteriors, with g inferred to lie around 0.55 ± 0.10 cm day^{-1}, compatible with the true growth rate parameterising the synthetic data. The branching rate parameter is more broadly spread, with a mean of 0.6 day^{-1} corresponding to the observed number of branches, and flexibility in the posterior reflecting the stochastic nature of this parameter's influence. α was inferred to lie around 0.021 ± 0.02, corresponding to a lateral growth rate around 0-0.02 cm day^{-1}; this is rather lower than the value used in the simulation, reflecting the rather limited lateral growth occurring in the specific simulated instance of the model. The posterior for l_{max} is close to recovering the prior which suggests that the model output is minimally dependent on the value of this parameter. We found this limited l_{max} dependence to generally be the case, and in subsequent sections will omit l_{max} from the posterior plots; all l_{max} posteriors, generally recovering priors, are plotted in Appendix B.8. This assessment of the relative importance of, and flexibility in, generative mechanistic parameters reflects a powerful aspect of this inverse modelling approach.

B.2.3 Inferring mechanistic parameters for other synthetic phenotypes and root simulation models

To test the wider applicability of our likelihood-free inference process, we next tested the ability to identify known generative parameters when using a different, existing root simulation model, and for different plant species. RootBox [Leitner et al., 2010] was chosen for its wide application in the field. We embedded RootBox as the generative model in our inference framework, which was then applied as in section B.2.2 to the previous synthetic *Arabidopsis thaliana* data and a simulated *Lupinus angustifolius* root system. The *Lupinus* simulation involved an initial growth rate of 1cm day^{-1} for the primary root, 0.2 cm day^{-1} for the laterals and an inter-root distance of 0.9cm, and proceeded for 15 simulation days.

RootBox employs a different branching protocol from our simple model above. Rather than allowing stochastic branching anywhere on the primary root, RootBox allows lateral branches to emerge at specified intervals along the primary structure. This interval d, and a value b_{max} governing the maximum number of allowed lateral branches, are parameters of the model and we therefore seek posterior distributions on these quantities as well as the other mechanistic parameters which directly map to those in our simple model.

Fig. B.2 shows the resultant posteriors after applying our inference approach using RootBox as the core mechanistic model. Once more, the original generative parameters are well supported by the resulting posteriors, which also agree with the inferred values for primary and lateral growth rates using our simplified core model above (Fig. B.1). The inter-lateral distances d, present in RootBox but not above, are also well recovered by the inference process. The maximum number of branches b_{max} is not tightly constrained by the synthetic data, and is therefore shown in the supplementary information (Fig. B.9).

B.2.4 Inferring mechanistic parameters for wildtype *Arabidopsis thaliana* root systems

To test the pipeline on experimental data, we grew *Arabidopsis* Col-0 plants on vertical $\frac{1}{2}$ MS agar plates (see Methods) and used a digital camera to capture their root system structure over several days. We used SmartRoot image analysis software [Lobet et al., 2011] to extract the lengths and placements of tap and lateral roots from these digital images at each sampled timepoint. An example of the digitised data is shown in Fig. B.3.

We applied our ABC SMC framework to estimate the posterior distributions of the mechanistic parameters underlying the development of these root systems. The earlier populations of the SMC process gave a diverse range of simulated root structures; by the final population, the simulation outputs provide excellent visual matches to the observed experimental structures (Fig. B.3) given the deliberate simplicity of the model. This intuitive snapshot matching is supported by the good agreement between the experimentally observed time series of summary statistics and those arising from simulation with the final posteriors (Fig. B.4). Here, both the mean and the variability in the experimental

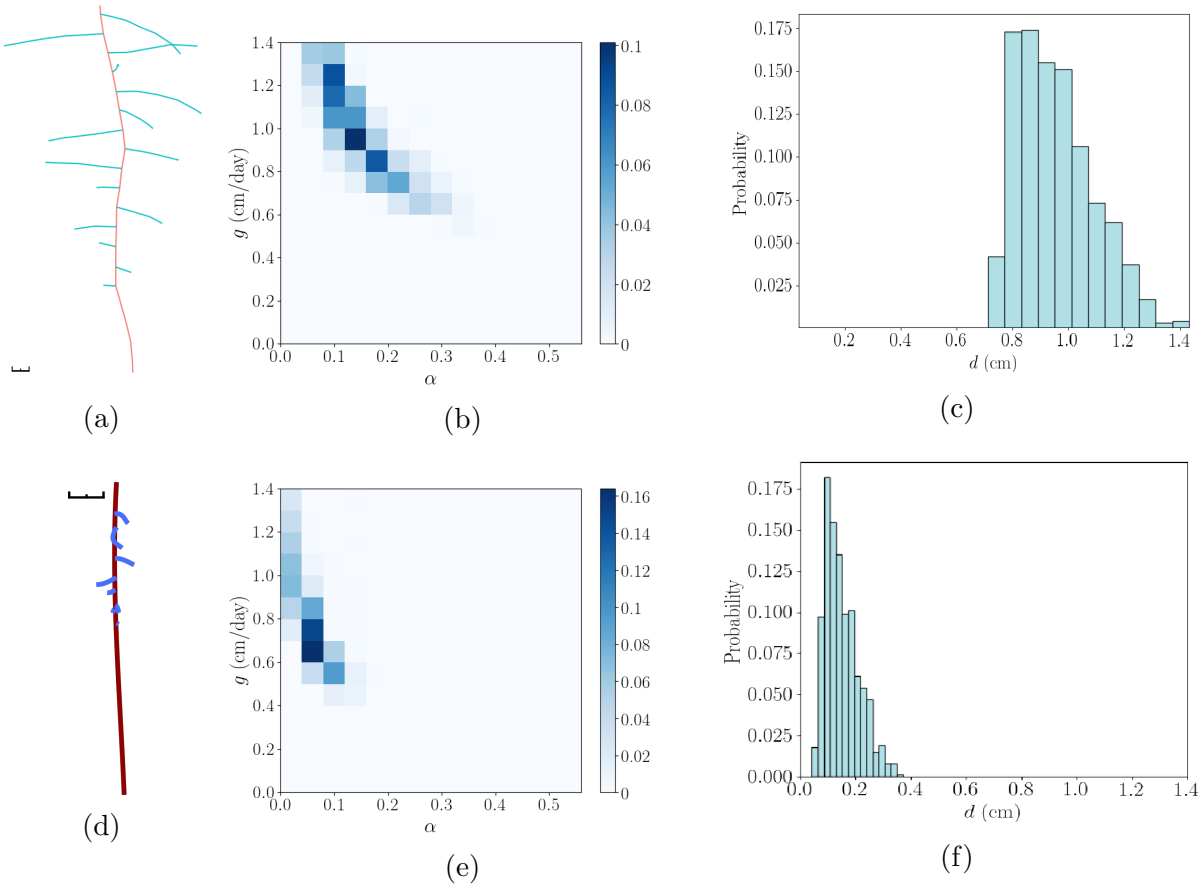


Figure B.2: **Posterior distributions on *Arabidopsis thaliana* and *Lupinus angustifolius* roots generated using RootBox.** (a) Output from RootBox simulation of *Lupinus angustifolius* root growth; black scale bar is 1cm. (b-c) Output posteriors from an ABC SMC framework run on RootBox output of *Lupinus angustifolius*, with final ABC SMC tolerance $\epsilon = 0.4$ (see Methods). (b) Two-dimensional posterior on primary root growth rate g and lateral root growth scaling α. (c) Posterior distribution on branch separation d in the RootBox growth model. (d) Output from RootBox simulation of *Arabidopsis thaliana* root growth; black scale bar is 1cm. (e-f) Output posteriors from an ABC SMC framework run on RootBox output of *Arabidopsis thaliana*, with final ABC SMC tolerance $\epsilon = 0.4$ (see Methods). (e) Two-dimensional posterior on primary root growth rate g and lateral root growth scaling α. (f) Posterior distribution on branch separation d in the RootBox growth model.

(a) (b) (c)

Figure B.3: **Example data and simulation output for the root inference framework**. (a) *Arabidopsis* seedlings grown vertically on agar (see Methods) provide example root systems for the analysis pipeline. (b) Digitisation of the root systems using SmartRoot [Lobet et al., 2011] provide the quantitative data used in the inference process. (c) Example outputs from the stochastic growth model with parameters identified through the ABC SMC inference process. Black scale bar is 1cm.

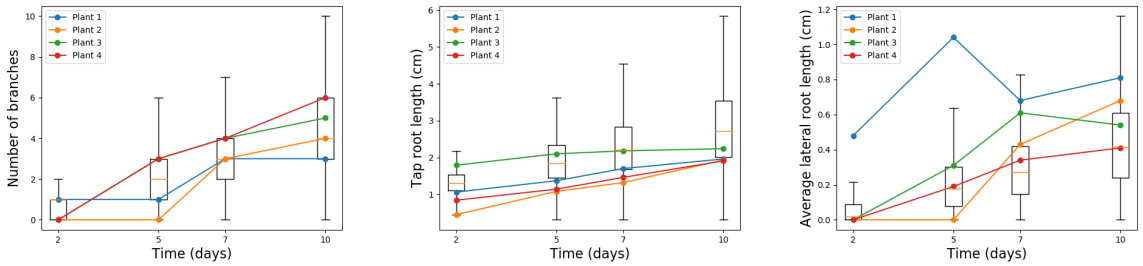
statistics over time is captured by the distributions of simulated behaviour arising from the posteriors.

The posterior distributions themselves are shown in Fig. B.5. The primary root growth rate g is reasonably well constrained, with a mean that intuitively falls around the total growth average. Notably, the posterior distribution on g is tighter than for the synthetic data example. This refinement reflects the strength of including time-course data in the inference platform. Observations of systems at different times provide more information on dynamic rate parameters, allowing better estimates than are available from single-instance observations alone.

The scaling of lateral growth rate α has a broader variance, reflecting the greater variability in average lateral root length observed in the data, and is correlated to some extent, as expected, with the value of g. The distribution of branching rate b is broader, reflecting a greater variability in the experimental observation of branch number over time, and also the stochastic nature of this process: as b reflects the mean rate of a Poisson process, the same branching structure can be achieved with a variety of different b values. The modal value of b matches the average branching rate observed in the data. Overall, therefore, the ABC SMC framework gives reliable and intuitive readouts linked to both the average observed behaviour and plant-to-plant variability in root structure.

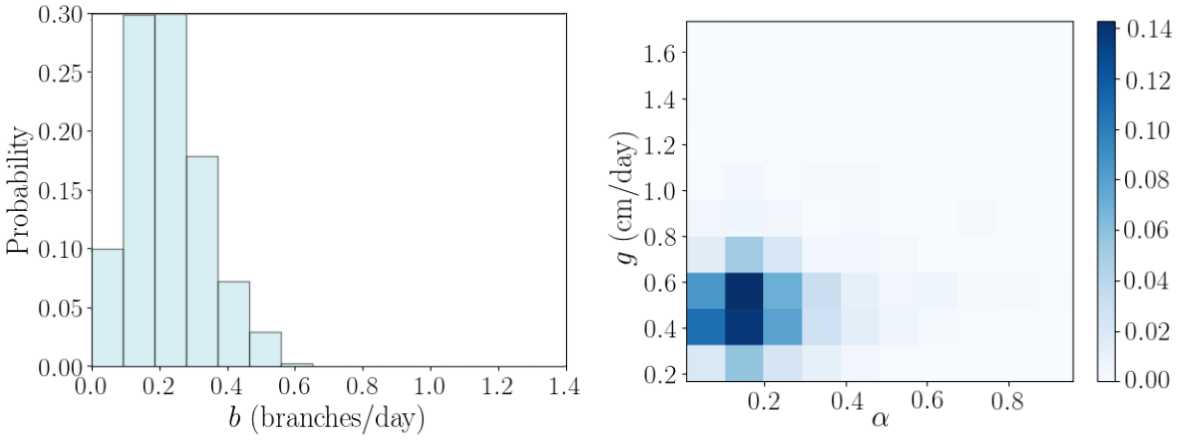
B.2.5 Model selection for root growth and branching mechanisms

We next asked whether our approach could select between competing generative models, given time course data on the evolution of a root system. To this end, we considered a range of possible generative mechanisms for root growth and branching. We will employ uniform priors over competing models, reflecting the fact that before any observations



(a) (b) (c)

Figure B.4: **Summary statistic comparison between data and parameterised model.** Individual line traces in each plot show time series of the summary statistic from observed *Arabidopsis* seedlings in Fig. B.3; boxplots give the range of values arising from stochastic model simulation after parameter values have been learned. (a) Number of branch points; (b) primary root length; (c) average lateral root length.



(a) (b)

Figure B.5: **Posterior distributions on mechanistic parameters for *Arabidopsis* seedling roots.** Posteriors from our ABC SMC framework run on the data from figure B.3 with final tolerance $\epsilon = 2.5$. (a) Posterior distribution on branching rate b. (b) Two-dimensional posterior on primary root growth rate g and lateral root growth scaling α.

are made, we have no belief that one mechanism is more likely than another. This prior belief can of course be arbitrarily changed within our Bayesian framework to reflect prior information. We then use our ABC SMC framework to identify the posterior support for each mechanistic model, given the observed data [Toni et al., 2009] (see Methods).

First, we consider different elongation laws for root growth. The first model involves root growth at a constant rate; the second involves a negative exponential growth law supported by [Schnepf et al., 2018] of the form

$$l(t) = l_{max}\left(1 - e^{-gt/l_{max}}\right), \tag{B.3}$$

where $l(t)$ is the length of the root at time t, parameterised by a rate constant g and scaling constant l_{max}. The posterior distribution over model index through the SMC process is shown in Fig. B.6a)-b). The most permissive population (highest ϵ) shows less support for the exponential model to ensure model parsimony. As a better fit to the data is required, the support for the exponential model increases until it overcomes the lower weightings due to the additional parameter and is preferentially selected.

Next, we explore a more nuanced mechanistic question underlying root architecture. We compared two models for branch placement positions. First, a uniform branching model, where the branching location was chosen at random anywhere along the primary root. Second, a minimally spaced model, which imposed a distance parameter δ around each existing branch where no further branching could occur. If a branch was attempted within this distance no branch was implemented and the algorithm continues. Fig. B.6b)-c) shows the model selection posteriors with decreasing tolerance, and the posterior on δ when the minimally spaced model was implemented. Here, the posteriors for the spacing model are lower for the more permissive populations, reflecting the increased model complexity – the extra parameter δ makes the model less parsimonious. The support for the model increases as a better fit to data is required, in the subsequent populations. By the final tolerance, the models have comparable support. Hence, the dataset suggests roughly equal support for both models despite their difference in complexity.

These simple experiments serve to illustrate the ability of ABC SMC to provide statistical support for competing mechanistic hypotheses (for example, linear versus negative-exponential root elongation laws). There is, however, nothing to prevent other targetted mechanistic questions being addressed using this framework (see Discussion).

B.2.6 Comparison between root structures

Next, we asked whether the ABC SMC framework could distinguish between two phenotypes – those corresponding to wild-type *Arabidopsis* and the *friendly* mutant line. *FRIENDLY* is a mitochondrial fusion gene that when compromised has a range of bioenergetic effects which lead to reduced root growth [El Zawily et al., 2014].

Wildtype and *friendly* plants were grown under the same conditions as above (see Methods), and the inference pipeline was run as before, with exponential growth and

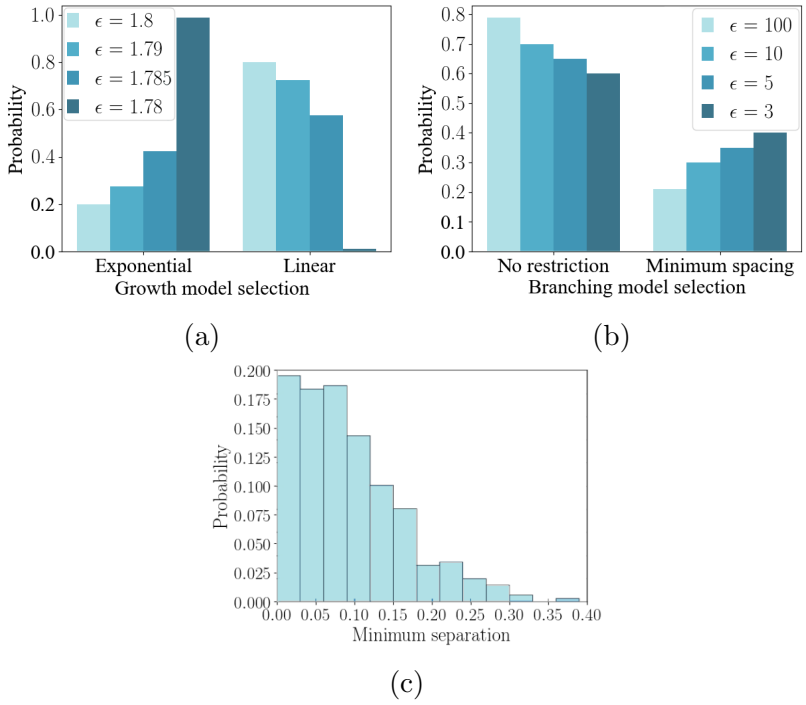


Figure B.6: **ABC SMC allows selection of competing mechanistic models for root growth.** (a) Model selection posteriors from comparing a simple, constant-rate growth model to the negative exponential model used in CRootBox [Schnepf et al., 2018] for decreasing ABC tolerance ϵ. The negative exponential model acquires greater support as the tolerance decreases. (b) Model selection posteriors from comparing a uniform branching model to a model imposing a minimum separation from existing branches. The more parsimonious uniform model experiences higher support with relaxed tolerance but the minimum-separation model gains support with tighter tolerances (tolerances are higher for this model selection run, reflecting a tradeoff with the greater computational resource required). (c) Posterior for the minimum branch separation, δ, required when the minimum spacing model is implemented.

uniform branching. The output posteriors in Figure B.7 reflect the differing root systems shown in the tracings, with a clear separation in the parameter space between the two phenotypes. The branching rate b is fairly unconstrained as observed in section B.2.4 due to inherent stochasticity in the branching mechanism. The values of g vary significantly between wildtype and *friendly*, as reflected in the tracings, with little change in the value of α. The distribution of g is substantially shifted towards lower values for the *friendly* plants, reflecting the known challenge to root growth resulting from this mutation. l_{max} shows a wide variability in both phenotypes, while representing clear differences consistent with the reduced root growth observed in the *friendly* mutant line. There is very little constraint in the value of l_{max}, suggesting little reliance on the value, although smaller values appear to be favoured for the *friendly* phenotype.

Taken together, these results demonstrate that the physical parameters governing root architecture growth can be learned using this ABC SMC approach, and uncertainty in these learned outcomes quantified. The mechanistic model within our inference process both allows us to harness time-course data and dissects which parameters change (here, growth rate g) and which remain similar (here, lateral root scaling α) in different cases. This is a key strength of this method, and would be difficult to obtain through traditional parameter inference methods. The platform readily identifies the physical different mechanisms underlying root architecture in a mutant line, and identifies accepted physical model structures for root growth.

B.3 Discussion

We have presented a framework for the inference of parameter values and mechanisms in root growth models when applied to an observed root system. While there has been much work undertaken producing plant root simulations from given parameters, our approach addressed the much less-studied 'inverse problem': that of finding generative parameter values and mechanisms that can reproduce a given root system. Knowledge about these mechanisms and parameters, and their ranges, flexibility, and relative importance, is necessary for an understanding of root growth processes such as growth and branching decisions, and how these may relate to biological processes within the plant. We hope that this highly general approach will allow for a more mechanistic understanding of root growth, and to quantify the efficacy of existing models.

We first used a very general growth model to (a) retain consistency with the 'core' of the maximum possible number of existing growth models, and (b) focus on parameters related to the growing plant and its phenotype, rather than the specifics of its physical embedding. We have demonstrated using RootBox [Leitner et al., 2010] that our approach can readily be adapted to other specific existing root models to allow the quantification of values of and uncertainty in generative parameters, furthering understanding of root system architecture. We also illustrated how alternative hypothesised mechanistic models can straightforwardly be compared, using SMC model selection. A strength of the Bayesian embedding here is that the most parsimonious model that is capable of explaining observations is naturally selected in the case of models with different numbers of parameters [Toni et al., 2009].

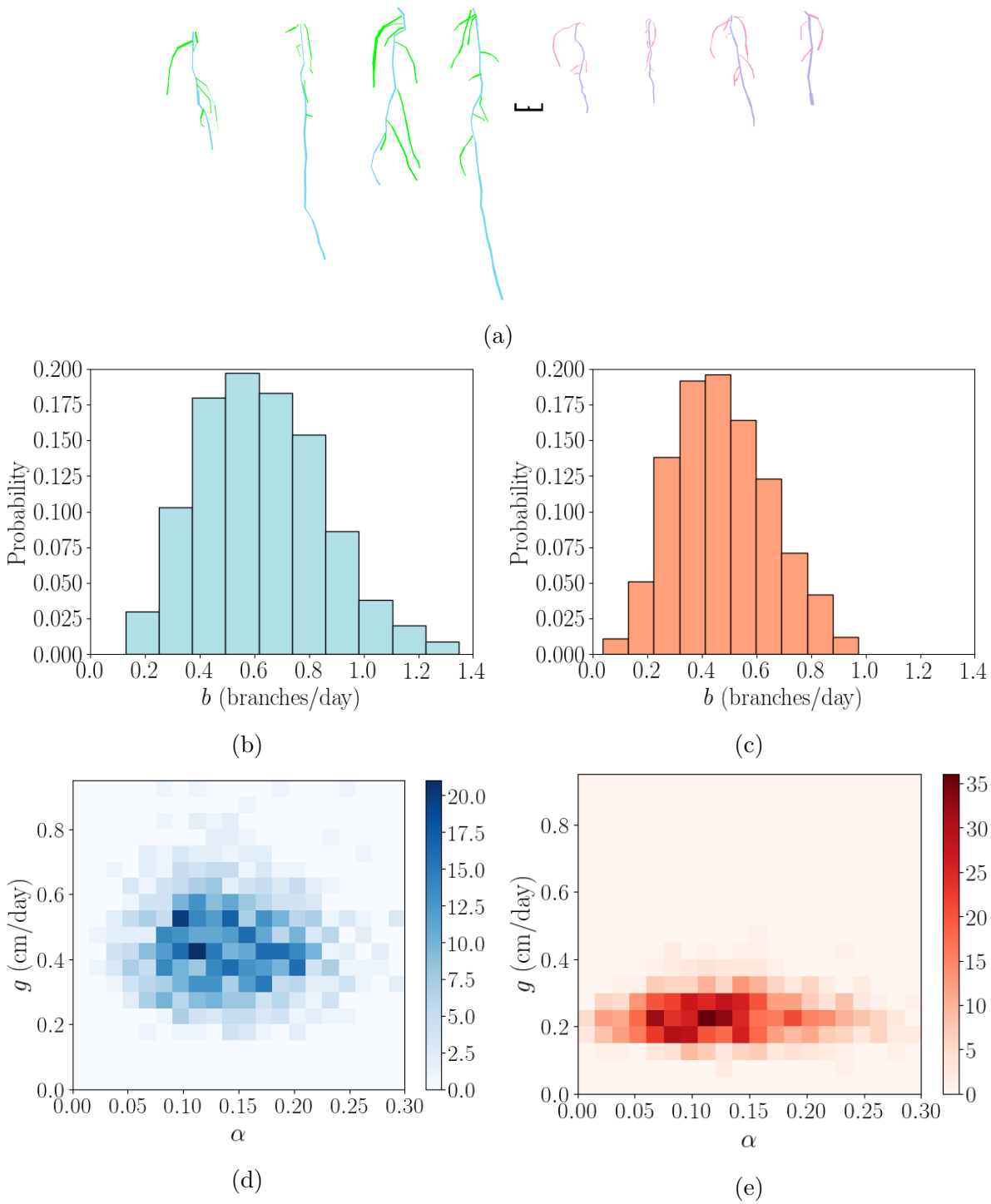


Figure B.7: **Distinguishing phenotypes with mechanistic inference.** (a) (left) Wild-type and (right) *friendly Arabidopsis* seedlings grown in agar as described in section B.4.1, demonstrating the root growth phenotype of *friendly*. The tracings were produced using SmartRoot software [Lobet et al., 2011] and the colours adjusted; black scale bar is 1cm. (b-c) Output posteriors for branching rate b show similar distributions for wildtype and *friendly*. (d-e) Two-dimensional posteriors on primary root growth rate g and lateral root growth scaling α demonstrate clear separation, reflecting the reduced root growth observed in *friendly*.

Advances in imaging techniques are allowing for greater insight into root system architecture [Roose et al., 2016], and specially designed image analysis software [Lobet et al., 2013] allows for increasingly efficient data collection from images. The combination of root models, advanced imaging techniques, image analysis software and an SMC framework could allow further advances in our understanding of root growth. We anticipate that, with the increasing developments in root imaging, this technique will find application in a growing variety of datasets, allowing for investigation of generative parameters for a wide variety of root system phenotypes. A natural future extension for this work would be to perform inference based directly on image data, rather than statistics of these data. This approach would require simulation of the imaging process as well as the generation of model root systems, for example, embedding the idealised root system in a simulated soil substrate and simulating the artefacts and noise involved in the imaging process. While (much) more computationally intensive, this approach would allow a more direct leveraging of phenotype data from experimental studies.

Notably, our approach allows inference based on time-course measurements of a developing root system, which increase the power and precision with which parameters and mechanisms can be identified. As demonstrated with our synthetic examples, this approach can readily be applied to single-instance observations, but also naturally leverages dynamic information to refine posterior distributions on physical rate parameters.

A stochastic modelling framework for root growth allows for a wide variety of possible outputs to be considered in the inference process, reflecting the variation between root systems in the real world. In this way, the modelling approach allows for investigation into the underlying mechanisms which are widely applicable, while avoiding a reliance on specificities and overfitting to a particular phenotype or growth environment. Predicting a branching event would require the consideration of processes such as genetics, cellular interactions, and organism-scale resource partitioning [Band et al., 2012], necessitating the development of a multiscale framework. As such multiscale approaches develop, we anticipate the use of likelihood-free inference to be further embraced to resolve inverse problems in parameter identification.

While the generality of our approach is appropriate for the scope of this study, greater specificity is required to gain a true understanding of plant processes. Care needs to be taken in the application of ABC techniques: choices must be made over elements such as the tolerance, priors, and summary statistics to achieve a balance between convergence rate and specificity of results. As specific choices for these values can be hard to interpret, simulation outputs must be verified to provide a reasonable match to genuine behaviour (as we have attempted throughout). We have worked with different models to explore the behaviour of our method under different generative assumptions. In Bayesian model selection, prior beliefs about models can strongly affect their support and interpretation must take this into account [Toni et al., 2009].However, we have aimed to demonstrate the strength of this approach when carefully applied and interpreted.

Overall, we have demonstrated a technique to allow for greater insight into model parameters for root systems, which could aid in increasing understanding of root growth mechanisms. The generalised approach allowed for investigation of the key aspects underlying root topology, while being highly adaptable for use with existing root architecture models.

B.4 Methods

B.4.1 Plant growth

Arabidopsis Col-0 seeds were sterilised with three 3-minute washing steps in 50% domestic bleach and water rinses, then plated on $\frac{1}{2}$ MS agar in vertical plates. Plants were grown at constant 25°C on a 16h light / 8h dark cycle. Plates were photographed over a time course of 2, 5, 7, and 10 days to produce time-series images of the seedling growth. Summary statistics were extracted from the images using SmartRoot [Lobet et al., 2011], an imageJ plugin. The root systems were traced manually using thresholding, producing a skeleton over the original image. Summary statistics on root length and branch placement were then recorded from this skeleton.

B.4.2 Model structure

Root growth was simulated using a hybrid stochastic-deterministic algorithm. Primary root growth, by default, was assumed to follow a negative-exponential growth law:

$$l(t) = l_{max}\left(1 - e^{-gt/l_{max}}\right), \tag{B.4}$$

parameterised by a rate constant g and a scaling constant l_{max}. The alternative uniform growth model simply took the form $l(t) = gt$. Lateral roots grow according to the same growth law as the primary root, but with a multiplicative factor α applied to g so that for lateral roots $g_l = \alpha g$.

Branching was treated as a Poisson event with rate parameter b. The time until the next branching event is found using the Gillespie algorithm [Gillespie, 1977], and the length of existing branches is updated from the current time until the time of the branching event. The branching location was then determined by a specified branching model, initially specified as a uniform probability distribution along the length of the primary root. In visualising structures, branching angle was always set to an angle of 45° from the growth direction, with equal change of being placed each side of the primary root, although these angles and positions play no role in the simulation. These steps are repeated until the time of the next branch exceeds the maximum simulation time, at which point the branch lengths are updated up to the maximum simulation time, and no branching event occurs. Once a branching event has occurred, the sidebranch grows according to the same growth law as the main root, scaled by parameter α; variability in lateral root length thus corresponds to variability in initial branching times and positions.

B.4.3 ABC SMC implementation

An ABC framework was implemented in Matlab. Model parameters were drawn from specified distributions and passed to the model as described in Model Structure above.

Broadly, the simulated root systems are then compared to data, and the parameter values accepted if the simulation is sufficiently close to the data, with tolerance ϵ defined at the time of implementation. If the previous values were accepted, the parameter values are perturbed with a perturbation kernel K_t. If the previous values were not accepted, the parameter values were drawn from the priors as previously described. This process was repeated until 1000 hits were obtained at the specified tolerance.

We follow Ref. [Toni et al., 2009] in our ABC SMC implementation. For completeness, Algorithm 1 introduces a simple rejection sampling scheme under ABC. Algorithm 2 embeds this scheme in an SMC framework for parameter inference and model selection.

Algorithm 1. ABC rejection sampling for parameter inference.

1. Given N_p plant structures and $N_t(i)$ longitudinal observations for plant i, characterise the summary statistics $d_{ij} = \{B, L, \hat{l}\}$ from every plant i and observation j in the dataset.
2. Draw a trial set of parameters θ^* from the prior distribution $\pi(\theta)$.
3. Simulate N_p instances of root growth, recording the state of structure i at each of the $N_t(i)$ time points corresponding to an experimental observation.
4. Compute ρ using Eqn. B.2 above, to give the separation between each recorded structure and its simulated counterpart.
5. If $\rho < \epsilon$, where ϵ is a given tolerance, accept θ^* as a sample from the posterior.
6. If a termination condition is not met, return to 2.

Algorithm 2. ABC SMC for parameter inference and model selection.

1. Given N_p plant structures and $N_t(i)$ longitudinal observations for plant i, characterise the summary statistics $d_{ij} = \{B, L, \hat{l}\}$ from every plant i and observation j in the dataset.
2. Initialise tolerance vector $\mathbf{E}$ containing T elements. Set population indicator $t = 0$.
3. Set particle indicator $i = 1$.
4. Sample model indicator m^* from prior $\pi(m)$.
5. If $t = 0$, sample θ^{**} from $\pi(\theta(m^*))$. If $t > 0$, sample θ^* from the previous population $\{\theta(m^*)_{t-1}\}$ with weights $w(m^*)_{t-1}$, and set $\theta^{**} \sim K_t(\theta|\theta^*)$.
6. If $\pi(\theta^{**}) = 0$, go to 4.
7. Simulate N_p instances of root growth using θ^{**}, recording the state of structure i at each of the $N_t(i)$ time points corresponding to an experimental observation.
8. Compute ρ using Eqn. B.2 above.
9. If $\rho \geq \mathbf{E}[t]$, go to 4.
10. Set $m_t^{(i)} = m^*$ and add θ^{**} to the population $\{\theta(m^*)_t\}$. If $t = 0$, set weights $w_t^{(i)} = 0$, otherwise
$$w_t^{(i)} = \frac{\pi(\theta^{**})}{\sum_{j=1}^{N} w_{t-1}^{(j)} K_t(\theta_{t-1}^{(j)}, \theta^{**})}. \tag{B.5}$$
If $i < N$, set $i = i + 1$, go to 4.
11. For every m, normalise the weights. If $t < T$, set $t = t + 1$, go to 3.

For a single model, the prior $\pi(m)$ associated with that model is unity and the choice of model indicator m^* plays no role in the inference process.

We used uniform priors over all model structures for $\pi(m)$ and uniform priors between 0-1.4 day^{-1} for g, 0-1.4 day^{-1} for b, 0-1 for α and 5-40cm for l_{max}. The perturbation kernel we used was $K_t \sim N(0, 0.1P)$, where P is the width of the uniform prior. The tolerance vector was $\mathbf{E} = \epsilon\{5, 3, 2, 1.5, 1\}$.

B.5 Competing Interests

The authors declare that no competing interests exist.

B.6 Acknowledgements

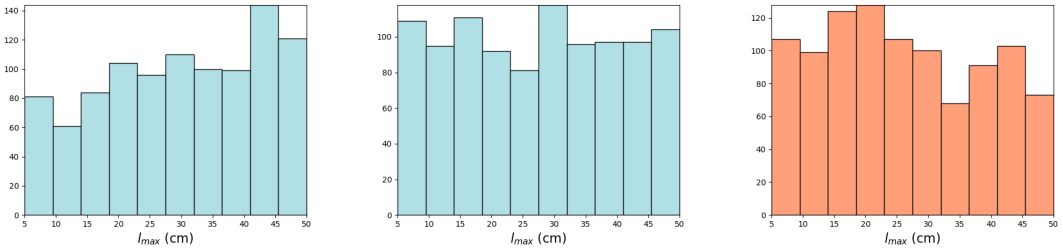
All authors acknowledge support from the Birmingham Institute of Forest Research. IGJ acknowledges support from a Birmingham Fellowship for the University of Birmingham and Turing Fellowship from the Alan Turing Institute.

B.7 Author Contributions

CZ developed the code, performed the plant growth experiments, and drafted the manuscript. IGJ conceived the project. RJD and IGJ supervised the project. All authors wrote the final manuscript.

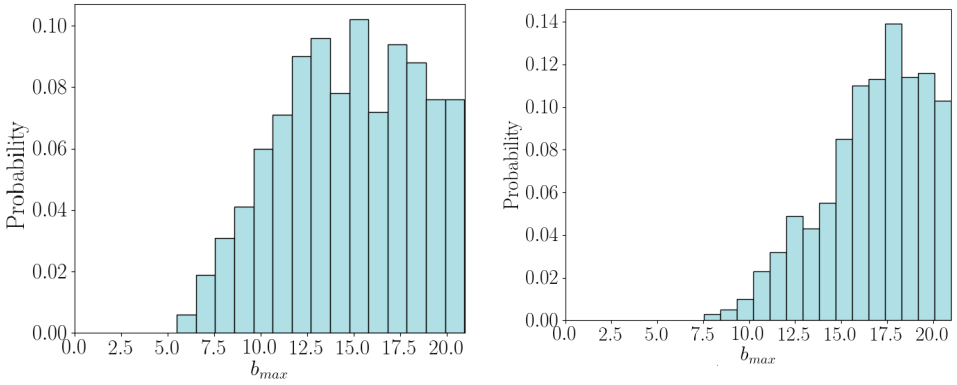
B.8 Appendix

Figures B.8 and B.9 show posteriors on l_{max} and b_{max} omitted from the main text.



(a) (b) (c)

Figure B.8: Posteriors on l_{max} for *Arabidopsis* wild-type seedlings (a) from the initial pipeline described in section B.2.4, and the wildtype (b) and *friendly* (c) comparison from section B.2.6.

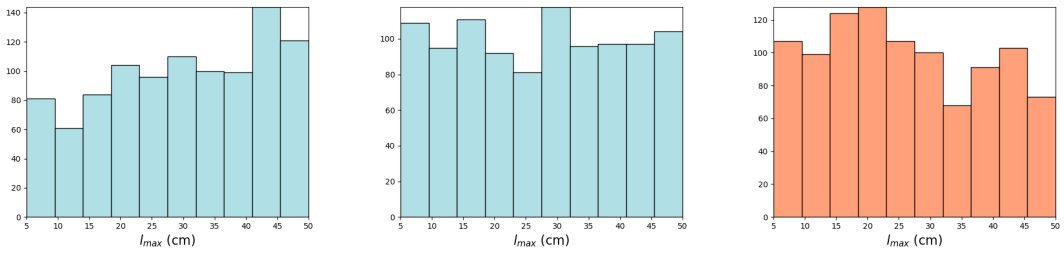


(a) (b)

Figure B.9: Posterior distributions on b_{max} for a) *Arabidopsis thaliana* and b) *Lupinus angustifolius* roots generated using RootBox.

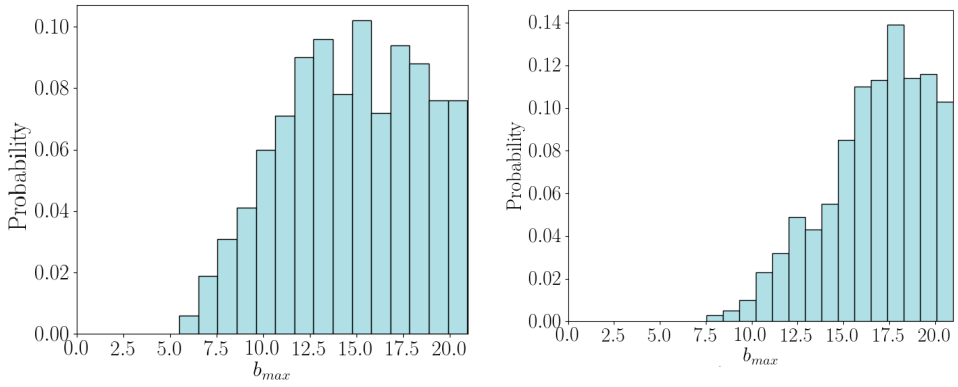
B.9 Supplementary Information

Figures B.10 and B.11 show posteriors on l_{max} and b_{max} omitted from the main text.



(a) (b) (c)

Figure B.10: Posteriors on l_{max} for *Arabidopsis* wild-type seedlings (a) from the initial pipeline described in section 2.4, and the wildtype (b) and *friendly* (c) comparison from section 2.6.



(a) (b)

Figure B.11: Posterior distributions on b_{max} for a) *Arabidopsis thaliana* and b) *Lupinus angustifolius* roots generated using RootBox.