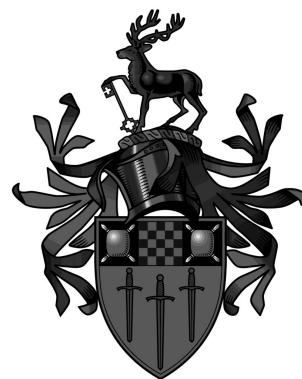


A Qualitative Analysis of the Data Assimilation Linked Ecosystem Carbon Model, DALEC

By Ann Mary Chuter

Submitted for the Degree of
Doctor of Philosophy



University of Surrey
Faculty of Engineering and Physical Sciences
Guildford, Surrey, GU2 7XH, U.K.

5 February 2013

©Ann Mary Chuter 2012

Abstract

Due to changes in our climate and environment, partly caused by human behaviour, it is becoming increasingly important to understand the processes involved in Earth systems, such as the carbon cycle. There are many models that attempt to describe the dynamical behaviour of carbon stocks and stores, but, despite their complexity in attempting to describe all crucial processes, significant uncertainties remain. Our aim is to look at the qualitative behaviour of one of the simplest carbon cycle models, the Data Assimilation Linked Ecosystem Carbon (DALEC) model, and consider in detail the processes involved. DALEC is a simple vegetation model of processes involved in the carbon cycle of forests. Our analysis shows that the dynamics of both evergreen and deciduous forests in DALEC are dependent on a few key parameters and it is possible to find a tipping point at which stable sustainable behaviour of a forest gives way to widespread mortality. We also study and simplify the Gross Primary Production (GPP), a complex photosynthesis function at the heart of DALEC, and create annual maps of the five carbon pools involved, using the simplified and averaged GPP. These results are then used to examine the effects of moisture shocks on the tipping point, as well as to examine parameter sensitivity from both a numerical and analytical point of view. The net ecosystem exchange (NEE) is an important measure of whether a forest is a CO₂ sink or source and also serves as a means to find parameters in the model. During the process of examining parameter sensitivity we simplify the Net Ecosystem Exchange (NEE) equation in various ways, depending on which time period is considered. We find that we can identify the sensitivity of NEE to the parameters and furthermore, we find that the NEE becomes more sensitive to some parameters over time and less to others. Furthermore we find that during a certain important time period NEE can be expressed in terms of the annual mean GPP. These results give important insights into what affects the NEE.

Contents

1	Introduction	15
1.1	Overview	15
1.2	Aims and Objectives	17
1.3	Structure of the Thesis	19
2	Literature Review	22
2.1	Introduction	22
2.2	The Carbon Cycle	23
2.3	Dynamic Global Vegetation Models	25
2.3.1	The Dynamic Global Phytogeography Model (DOLY)	27
2.3.2	The Biosphere Energy-Transfer Hydrology Scheme (BETHY)	29
2.3.3	Data Assimilation Linked Ecosystem Carbon, DALEC	33
2.3.4	Comparison of the Models	37

2.4	Data Assimilation	38
2.4.1	What is Data Assimilation?	40
2.4.2	4DVAR Data Assimilation	42
2.4.3	The Cost Function	42
2.4.4	The 4DVAR Algorithm	45
2.5	Conclusion	48
3	Models, Data and Assumptions	49
3.1	DALEC Evergreen	49
3.2	DALEC Deciduous	51
3.3	The Aggregated Canopy Model, ACM	53
3.4	NEE	54
3.5	Model Assumptions and Data	54
4	Tipping Points	56
4.1	Introduction	56
4.2	Linear Maps	57
4.2.1	Variation 1	57
4.2.2	Variation 2	59
4.3	Behaviour of the Foliar Carbon Pool	61

4.3.1	Finding Fixed Points	61
4.3.2	Varying the Parameters to Find Tipping Points	65
4.3.3	Smoothing the Drivers	70
4.3.4	Improving Smoothing of the Irradiance	74
4.3.5	Results of Smoothing on the Line of Tipping Points	79
4.4	Behaviour of the Other Carbon Pools	84
4.4.1	Fixed Points for Root Carbon, C_r	86
4.4.2	Fixed Points for Woody Carbon, C_w	87
4.4.3	Fixed Points for Litter Carbon \mathbf{C}_{lit}	88
4.4.4	Fixed Points for Soil and Organic Matter Carbon, \mathbf{C}_{som}	90
4.4.5	Periods of the Carbon Pools' Fixed Points	92
4.5	Preliminary Results for DALEC Deciduous	92
4.6	Conclusion	95
5	Gross Primary Production and \mathbf{C}_f Revisited	99
5.1	Introduction	99
5.2	Simplifying and Averaging the GPP	100
5.3	An Annual Map for \mathbf{C}_f	112
5.3.1	Approximation 1	113

5.3.2	Approximation 2	119
5.4	Linearising Around a Fixed Point to Simplify the Annual Map for C_f Further	124
5.5	Conclusion	129
6	Experimenting with Shocks	131
6.1	Introduction	131
6.2	The Experiment	132
6.2.1	What Happens When ψ Changes?	136
6.3	Conclusion	139
7	Constraining of Parameters	140
7.1	Introduction	140
7.2	Numerical Sensitivity Analysis for C_f and NEE	141
7.2.1	Table 7.1	142
7.2.2	Table 7.2	146
7.3	Analytical Sensitivity Analysis for NEE	148
7.3.1	Contributions of Carbon Pools	148
7.3.2	Contributions to NEE in More Detail	154
7.3.3	Conclusion	194

8 Conclusion	198
8.1 Overview of the Problem and Results	198
8.2 Impact of Results and Future Work	205
8.3 Conferences, Seminars, Workshops, Publications and Other Activities . .	207
8.3.1 Conferences, Seminars and Workshops	207
8.3.2 Publications	208
8.3.3 Other Activities	208

Acknowledgements

I gratefully acknowledge support from NERC via a studentship funded under the auspices of the National Centre for Earth Observation. I am also grateful to Mathew Williams, University of Edinburgh and Timothy Hill, University of Edinburgh for helpful discussions regarding this work.

I would like to thank my supervisors, Dr. Anne C. Skeldon and Dr. Philip J. Aston, who have consistently been supportive and approachable and have taught me so much. Prof. Ian Roulstone, my co-supervisor, has been supportive throughout and helped me to find the job I will start upon finishing my PhD.

Last but certainly not least I would like to thank all my family, but in particular my husband, Frank, my two sons, David and Daniël and my two daughters, Jasmin and Nina, for supporting me and believing in me.

**I dedicate this thesis to my
husband, Frank.**

List of Figures

2.1	<i>The carbon cycle. (Source: PhysicalGeography.net).</i>	24
2.2	<i>The increase in the concentration of CO₂ in the atmosphere from 1976 - 2010 (in parts per million), with projected values for 2013.</i>	25
2.3	<i>An overview of the JULES land surface model, using TROLL as its GVDM (source: jules.jchmr.org/model-description).</i>	27
2.4	<i>The DOLY model.</i>	28
2.5	<i>Structure of the BETHY model.</i>	31
2.6	<i>Modules of JSBACH, (Source: Max Planck Institute for Meteorology).</i>	38
2.7	<i>Structure of the SDGVM, (Source: Centre for Terrestrial Carbon Dynamics).</i>	39
2.8	<i>A graph showing the different data assimilation approaches as a function of time. (Source: ECMWF website).</i>	41
2.9	<i>Schematic representation of variational methods [15]: 1st adaption and 2nd adaption are trajectories created from newly estimated initial conditions, found by the data assimilation algorithm..</i>	43

2.10	<i>The analysis state, \mathbf{x}_a, (\mathbf{x}_0 in our cost function (2.1)), is positioned between the background state, \mathbf{x}_b, and first observation (obs) [40].</i>	43
3.1	<i>A schematic overview of the carbon cycle of an evergreen needle leaf forest according to DALEC EV.</i>	50
3.2	<i>A schematic overview of the carbon cycle of a deciduous forest according to DALEC DE. Parameters p_{12} and p_{13} are not shown, but are used to determine when m_{tl} and m_{tf} are switched on and off.</i>	52
4.1	<i>Results for varying parameter p_5.</i>	62
4.2	<i>Path of fixed points for the annual map for $p_{23} = 0.1392$. The top branch is stable, the bottom branch is unstable and there is also a stable fixed point branch at $C_f = 0$. On the vertical axis the value of C_f on day 1 of each year (December 21st) is shown.</i>	65
4.3	<i>Tipping points divide parameter plane into two different regions.</i>	67
4.4	<i>The line of tipping points for the Oregon forest.</i>	68
4.5	<i>Fold bifurcations for various values of C_a.</i>	69
4.6	<i>Smooth and variable drivers over a three year period.</i>	71
4.7	<i>Plot of C_f over three years, comparing C_a.</i>	72
4.8	<i>Comparison of smoothed and non-smoothed drivers.</i>	73
4.9	<i>Plot of the cost function, using the irradiance as a Fourier series.</i>	76
4.10	<i>Contour plot for irradiance expressed as a Fourier series, showing a number of local minima.</i>	76

4.11	<i>The result of data assimilation using the optimal Fourier series irradiance.</i>	77
4.12	<i>Plot of the cost function.</i>	80
4.13	<i>Contour plot of the cost function, showing a minimum.</i>	80
4.14	<i>The result of data assimilation using the optimal quadratic irradiance.</i>	81
4.15	<i>The irradiance smoothed by a Fourier series and a quadratic function.</i>	81
4.16	<i>Error smoothed drivers for all the pools.</i>	82
4.17	<i>C_f over three years shown using original drivers and smoothed drivers. Although the fit is very good for three years, the curves would grow a little further apart if shown on a longer timescale.</i>	83
4.18	<i>Results of smoothed drivers for tipping point.</i>	83
4.19	<i>Results of the effect of smoothed drivers for line of tipping points.</i>	84
4.20	<i>Behaviour of C_f, C_r and C_{lit} over 20 years.</i>	85
4.21	<i>Annual maps of the pools.</i>	85
4.22	<i>Sketch of the phase plane for the fixed points (C_f, C_r), showing two stable fixed points and one saddle fixed point.</i>	87
4.23	<i>C_w crawling to a fixed point over a very long period of time.</i>	89
4.24	<i>Pools showing one period of their non-trivial stable fixed point.</i>	91
4.25	<i>The foliar and labile pools plotted over one year.</i>	94
4.26	<i>Limit point for parameter p_{14} on a path of fixed points of the annual map for C_f for DALEC DE.</i>	94

4.27 Limit point for parameter p_{14} on a path of fixed points of the annual map for C_{lab} for DALEC DE.	95
5.1 The coefficient of $\tilde{C}_i(t)$ in equation (5.8), $\frac{-(C_a + q - p(t))}{280}$, in terms of C_f and time.	102
5.2 The coefficient in equation (5.8), $\frac{C_a q - p(t) a_3}{280^2}$, in terms of C_f and time. .	103
5.3 The coefficient of $C_{ai}(t)$ in equation (5.12), $\frac{-(C_a - q + p(t))}{92}$, in terms of C_f and time.	104
5.4 The constant coefficient in equation (5.12), $\frac{p(t)(C_a - a_3)}{92^2}$, in terms of C_f and time.	105
5.5 The term $\frac{4p(C_a - a_3)}{(C_a - q + p)^2}$ (in terms of C_f and t) is small relative to 1 in equation (5.20).	106
5.6 $C_a - C_i$ original and simplified.	107
5.7 GPP original and GPP using simplified $C_a - C_i$	108
5.8 C_f plotted using the original GPP and the simplified GPP as in equation (5.31).	111
5.9 Original and simplified GPP as in equation (5.31).	111
5.10 C_f plotted using simplified GPP as in equation (5.31) and a fitting factor of approximate value 1.085.	112
5.11 Comparisons of annual maps for C_f	115

5.12 An annual map of the original GPP and of the simplified GPP as in equation (5.31). Both annual maps were created by taking the mean yearly GPP from the daily calculation of GPP.	116
5.13 Fold bifurcation found using original GPP and simplified GPP.	120
5.14 Line of limit points found using the original GPP and the averaged, simplified GPP.	120
5.15 Annual map of C_f plotted using the original GPP, created from yearly mean C_f calculated by the daily map. Also the annual map of C_f using Approximation 1 and the annual map of C_f using Approximation 2.	123
5.16 The limit point calculated using the original GPP (continuous line) and Approximation 2 for the averaged GPP (dashed line), with fitting factor of 1.139.	123
5.17 The line of limit points calculated using the original GPP (continuous line) and Approximation 2 for the averaged GPP (dashed line), with fitting factor of 1.139.	124
5.18 Plot showing that linearisation of the annual map for C_f works well, even from far away from the fixed point.	126
5.19 Plot of the normal form for a limit point, equation (5.90); values for C_1 and C_2 are both 1.	127
5.20 Plots for equations (5.95) and (5.101).	129
6.1 Varying ψ using the original model (continuous lines) and using the simplified model (dotted lines).	134

6.2	<i>Path of limit points for ψ and p_5.</i>	135
6.3	<i>Path of limit points for p_5 and C_f.</i>	135
6.4	<i>The ψ story.</i>	137
6.5	<i>Here the critical value for C_f when ψ changes from 2 to 0.5 is shown as well as the unstable fixed point for C_f.</i>	138
7.1	<i>The five carbon pools scaled from fixed points.</i>	151
7.2	<i>Zoomed plot of scaled \tilde{C}_{som} and \tilde{C}_w.</i>	151
7.3	<i>The five scaled carbon pools, starting from initial conditions.</i>	152
7.4	<i>The different contributions to NEE.</i>	154
7.5	<i>Growth of C_w and C_{som} between 0 and 30 years.</i>	156
7.6	<i>Growth of C_f, C_r and C_{lit} between 0 and 30 years.</i>	157
7.7	<i>A visual representation of equation (7.28).</i>	160
7.8	<i>Contributions to the annual NEE as per equation (7.45).</i>	166
7.9	<i>The NEE expressed in three different ways between 0 and 30 years.</i>	180

7.10	<i>The parameter estimation for DALEC evergreen with synthetic data [13]. The vertical line in each of the panels represents the true value of the parameter. The circles are the values of the parameters as estimated by the different data assimilation algorithms (indicated on the y-axis) and the brackets around the circles indicate the size of the 90% confidence intervals. The upper and lower bounds of each parameter are shown on the x-axes. For all parameters beginning with T the x-axes are log-scaled.</i>	182
7.11	<i>Parameter estimation metrics using nine different algorithms based on syn- thetic data for evergreen forest [13]. Lower values indicate better results.</i>	184
7.12	<i>The annual map for C_{som} plotted from the original daily map and as an annual map, no extra fitting factor.</i>	186
7.13	<i>The NEE expressed in two different ways between 30 and 1,000 years.</i>	187
7.14	<i>The annual C_w and C_{som} during Phase II; the growth of C_w is much steeper than the growth of C_{som} in the first 150 years or so.</i>	188
7.15	<i>The annual NEE calculated from the original daily maps, using yearly av- erages for NEE, and annual NEE in terms of mean GPP.</i>	189
7.16	<i>With observations too early in the life of the forest, (7.92) underestimates the correct values for the unknowns, C_f^* and $f'(C_f^*)$.</i>	192
7.17	<i>With observations from approximately half way through phase I, (7.92) pro- duces the approximate correct values for C_f^* and $f'(C_f^*)$.</i>	192
7.18	<i>The growth of C_{som} and C_w on a log scale.</i>	193

Chapter 1

Introduction

1.1 Overview

Since the Industrial Revolution the amount of carbon dioxide in the atmosphere has risen exponentially as a result of human activities, such as the burning of fossil fuels and deforestation [28]. Carbon dioxide is a greenhouse gas which absorbs and emits infrared radiation and traps the heat in the atmosphere. It affects the Earth's climate and the functioning of its ecosystems, such as the terrestrial ecosystem. In return, ecosystems can also affect the climate through biophysical and biochemical processes. It is known that a considerable amount of atmospheric CO₂ is taken up by vegetation. However, with increasing CO₂ levels and the consequent global warming, terrestrial vegetation, which formerly acted as a carbon sink, may not be able to absorb as much CO₂ as before or may even contribute to more atmospheric CO₂ by dying and decomposing [12]. This is a strong motivation to understand in depth the Earth's systems, such as the carbon cycle: the exchange of carbon between the biosphere, pedosphere, geosphere, hydrosphere and the atmosphere.

Dynamic Global Vegetation Models (DGVMs) are models which simulate the key physiological processes of the terrestrial carbon cycle through both empirical and process based methods, based on climate data, and are typically quite complex. DGVMs are used not only for determining carbon stocks, but also to simulate the effects of future climate change on natural vegetation.

Nowadays there exist many quite complex and sophisticated DGVMs used in analyses of the carbon cycle. However, the underlying dynamics of the carbon cycle is not well understood, not least because the values of many of the parameters that appear in the models are not known with confidence. Uncertainties in observations and parameters can influence the outcome of a simulation of a complex system, such as a GCM, as a consequence of changes in the dynamics of the core model.

Various projects have been designed to use techniques from data assimilation in an attempt to constrain the values of the parameters by ‘confronting the models with data’ [54, 17]. The majority of modern data assimilation schemes use the forward model as part of the algorithm and it is vital to understand the nature of the system dynamics if one is to employ the data assimilation methods with any confidence. For example, variational data assimilation techniques are based on assumptions of quasi-linearity and the use of Gaussian statistics. The presence of multiple timescales, either in the frequency of observations or in the dynamics of the forward model, is a feature that presents challenges to data assimilation schemes. It is therefore important to understand whether a problem such as the carbon cycle is amenable to scrutiny using such methods and, in particular, how the intrinsic separation of timescales, from days (for foliar carbon) to millennia (for soil organic matter), presents fundamental challenges to data assimilation schemes.

Better understanding of the behaviour of carbon cycle models could help to reduce some of the uncertainty. In order to begin studying the issues that arise in more complex systems we first study a relatively simple process model which is typical of that which

underpins many GCMs. This is a strategy not dissimilar to that found in other areas of environmental prediction, for example the Lorenz 1963 equations [27], which were used to study thermal convection in a very simple setting and which brought to light fundamental properties of dynamical systems.

1.2 Aims and Objectives

The aim of this thesis is to try to understand in depth the qualitative dynamical behaviour of the Data Assimilation Linked Ecosystem model (DALEC) [51]. DALEC is a dynamic vegetation model which simulates the carbon cycle of forests. It was especially created for data assimilation and specifically designed for calibration and testing against eddy flux data. DALEC is a simple box model of carbon pools connected via fluxes running at a daily time-step, driven by daily climate data such as temperature, irradiance and atmospheric carbon. There are many similarities between DALEC and other models but also some differences. One similarity is that DALEC's photosynthesis model is based on the Farquhar photosynthesis model, which is used by other DGVMs, such as Hybrid 3.0 [14], BETHY [23], ED2 [31] and DOLY [53]. The main difference between DALEC and many other DGVMs is that it is an extremely simplified model, yet it aims to capture the essential dynamics of the carbon cycle.

Different versions of DALEC have been used on several occasions, for example to show how data assimilation techniques can improve estimations of carbon fluxes and provide insights into ecosystem carbon exchange [51] and to analyse the magnitude and contribution of parameter and driver uncertainty to confidence intervals on estimates of carbon fluxes [47]. Other research by Hill *et al.* [17] has concentrated on the use of CO₂ flux time series for parameter and carbon stock estimation in carbon cycle research and Quaife *et al.* [37] used DALEC to assimilate canopy reflectance data into the DALEC model with an Ensemble Kalman filter. DALEC is also used in the ABACUS project, which is a

linked programme of plant and soil process studies, isotope analyses, flux measurements, micro-meteorology, process modelling, and aircraft and satellite observations designed to improve predictions of the response of the arctic terrestrial biosphere to global change.

We formulate the process model for DALEC as a dynamical system, iterative in time, and we study the model using dynamical systems methods, employing both analytical and numerical approaches.

The work done for this thesis was motivated by a masters thesis by Ilett [20], but there are some differences in approach. Ilett focused on two aspects of the model; firstly, the dynamics of the model for a three year period, considering how the model predictions varied with different initial conditions and different parameter values; and secondly, considering how the model behaved if all time dependence in the drivers, such as the temperature and the irradiance, was removed. Our analysis sets Ilett's work in a broader context, explaining why different initial conditions lead to different behaviour and showing how the generic behaviour of the DALEC model is controlled by various parameters, resulting in an interesting tipping point structure. By studying and simplifying the complex photosynthesis function and using this to create annual maps from the carbon pools' daily maps, we are able to recreate identical qualitative tipping point behaviour without the use of complicated software tools, such as continuation software. We also experiment with moisture shocks and examine what the DALEC model itself can tell us about parameter sensitivity based on observations of net ecosystem exchange (NEE) and leaf area index (LAI) from both an analytical and numerical point of view, whilst using the results of the simplifications to express the NEE in various ways to clarify the sensitivity. LAI, which is defined as the one sided green leaf area per unit ground area in broadleaf canopies or as the projected needle leaf area per unit ground area in needle canopies (definition by www.uni-giessen.de/~gh1461/plapada/lai/lai.html), is a measurement that is dependent on the foliar carbon, C_f and can be established both in the field and from satellite data. NEE is a measurement of how much carbon is entering and leaving the ecosystem and is

measured using flux towers. NEE plays an important role in establishing carbon stocks and stores globally.

1.3 Structure of the Thesis

Chapter 2, the literature review, discusses the importance of understanding the Earth's systems, such as the carbon cycle and looks at how this could be achieved in the form of mathematical models, which simulate the key physiological processes of such systems. We look at ways in which carbon stocks and stores are being determined and compare three vegetation models, one of which is the Data Assimilation Linked Ecosystem Carbon model, DALEC, the subject of this thesis. We also look at data assimilation, which is an important way of fitting models and observations and we discuss where it comes from and how it is being used, and discuss in greater detail one particular type of data assimilation, four dimensional variational data assimilation, which is used in Chapters 4 and 7 of this thesis.

Chapter 3 sets out and explains the structure of the DALEC evergreen version, DALEC EV and the DALEC deciduous version, DALEC DE, as well as the Aggregated Canopy Model, ACM, which is the photosynthesis model derived from the Soil-Plant Atmosphere model, SPA. The photosynthesis function is called the Gross Primary Production or GPP. This section also discusses the data which is used throughout the thesis and the assumptions which have been made regarding the models and the data.

In Chapter 4 we note that the foliar carbon pool, C_f , in the DALEC EV model decouples from the other four carbon pools. Hence the forest dynamics are driven by and depend on C_f and its parameters. By making the climate drivers periodic, we show that the long-term behaviour of the foliar carbon pool is a periodic oscillation of period one year. We find fixed points for the annual foliar carbon cycle, which then allows us to trace

out the behaviour of these periodic states as a function of one of the parameters using continuation software. Doing this we discover a fold bifurcation, also called tipping point, at which sustainable behaviour of the forest gives way to widespread mortality. By varying a second parameter we find a line of tipping points and this shows for which values of the parameters a forest can be expected to exhibit sustainable behaviour and for which values the forest will certainly die without a chance of coming back to life.

The drivers for the DALEC model are highly variant and we study the effect of smoothing them on our tipping point results. Furthermore we study the behaviour of the other four carbon pools and find their fixed points and we show that the deciduous version of the DALEC model exhibits the same type of behaviour as the evergreen version.

The aim of Chapter 5 is to write the GPP in such a way that it becomes clear how the GPP depends on the foliar carbon, C_f . The equation for GPP is rather complicated and contains a square root term in the equation for the C_i , which is the CO₂ concentration at the site of carboxylation. By simplifying this square root term the GPP becomes less complicated. We then look at different methods of averaging the GPP over an annual cycle, either by averaging all the time dependent functions or by averaging all the time dependent functions except for the irradiance. One of these simplified GPP equations can then be used to create an annual map for C_f and allows us to produce the same limit point graphs that we found in Chapter 4 without the use of continuation software.

Chapter 6 uses the annual map created in Chapter 5 and shows that it has become extremely simple to study what effect varying the moisture parameter has on the tipping point structure. This short chapter is intended to give an example of how the DALEC model can be used to study shocks, such as extended dry periods or extreme temperatures. By varying the shock parameter in the model we show how the tipping point moves, which explains why in some cases a forest in the DALEC model could die after a prolonged dry spell.

It is possible to obtain information regarding the constraining of the DALEC parameters through data assimilation, as was done in the REFLEX project in 2007 [13], a project to compare the strengths and weaknesses of various data assimilation methods for estimating parameters and predicting carbon fluxes using DALEC, based on three years of driver data. However, an examination of the model equations should also give useful information on which parameters are likely to be constrainable and which are not. This is explored in Chapter 7. The constrainability of parameters is very much dependent on the data that is fed into the model, for example Leaf Area Index, LAI and Net Ecosystem Exchange, NEE. NEE is both input and output for any GDVM as well as for any GCM, which, similar to DALEC, can employ data assimilation methods to produce better predictions short term and long term. Our aim in this section is to try to find out from the model equations alone which parameters can be expected to be reasonably constrained from NEE and LAI. A numerical sensitivity analysis for both C_f and NEE gives us information about the parameter sensitivity of a specific location whereas an analytical sensitivity analysis allows us to make a more general statement about the parameter sensitivity of any location. To this end we examine the different contributions of the pools to each other and the NEE and expand and simplify the NEE equation, using some results found in earlier chapters, such as the annual map for the foliar carbon, in various different ways, depending on the time period we are considering, as we find that the sensitivity of the NEE to parameters changes over time. We draw comparisons between our results and the results from the REFLEX study.

Finally, in Chapter 8 we look back over what has been achieved and what more could be done in the future.

Chapter 2

Literature Review

2.1 Introduction

It is widely known nowadays that our climate is changing and that this is caused to some extent by human activities. Section 2.2 discusses the importance of understanding the carbon cycle and looks at how this could be achieved in the form of mathematical models, which simulate the key physiological processes of this cycle. We look at ways in which carbon stocks and stores are determined and compare three vegetation models in Section 2.3, one of which is DALEC, the subject of this thesis. Data assimilation is an important way of fitting models and observations and Section 2.4 discusses what it is, where it comes from and how it is being used. We go into more detail about one particular method of data assimilation, which is used in Chapters 4 and 7 of this thesis.

2.2 The Carbon Cycle

Due to changes in climate and environment, which are partly caused by human behaviour, such as fossil fuel burning, large scale deforestation and industrial processes, it has become important to understand processes on Earth, like the carbon cycle. Natural questions that arise are: What happens during the carbon cycle process, which is crucial to life on Earth? How can we monitor the carbon cycle to be able to predict any changes to our environment and how can we influence it in such a way that we can preserve and improve our environment?

Carbon is vital to life on Earth, it is present in all organic matter, alive and dead. It is also stored in soils, fossil fuels, like oil, and in sedimentary rock deposits such as chalk. We find it in the atmosphere in the form of carbon dioxide, CO_2 , and, as this is slightly soluble, in the oceans, where it either remains dissolved in the water or is stored in marine organisms, such as shells and body parts of clams. On land the atmosphere provides ecosystems with CO_2 . Plants, for example, use CO_2 , together with the sun's radiation and water, to produce complex carbohydrates, a process called photosynthesis. A byproduct of photosynthesis is oxygen, O_2 , which is breathed in by humans, animals and also plants. During photosynthesis the carbon is stored in plant leaves, stems and roots and when the plant dies decay will take over: partially rotted organic matter is absorbed by the soil where it eventually decomposes into water and CO_2 , which is released back into the atmosphere, with the help of micro-organisms. Plants also pass on their carbon by providing food to humans and animals, which in turn release carbon back into the atmosphere in the shape of CO_2 through respiration. See Fig. 2.1 for a graphic overview of the carbon cycle.

Scientists cannot balance the storage and flow of carbon in the global carbon cycle. Findings show that there is less carbon in the atmosphere than expected, based on the existing knowledge about flows for carbon in from and out of the atmosphere. Although oceans

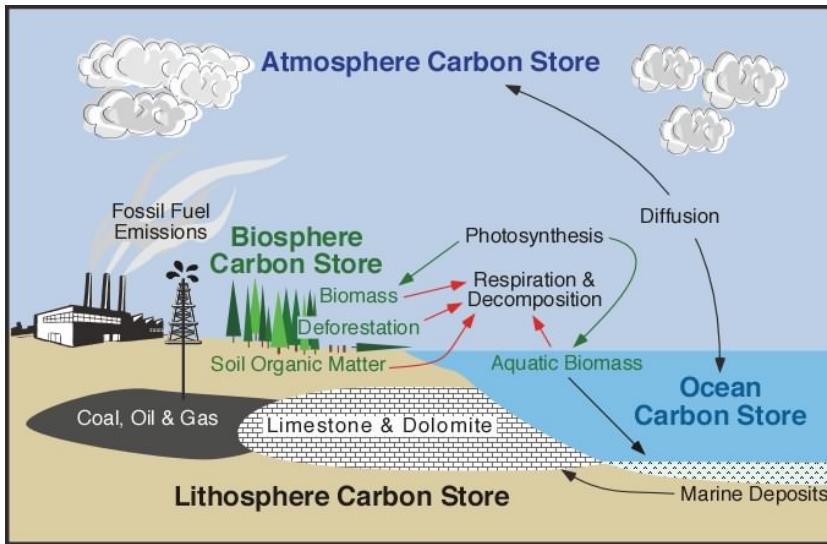


Figure 2.1: *The carbon cycle.* (Source: PhysicalGeography.net).

act as carbon sinks, the increase in the amount of carbon taken up by the oceans is not enough to explain this difference. One hypothesis is that net primary production is increased by the growing concentration of CO_2 in the atmosphere, which increases the rate at which carbon is taken from the atmosphere. For example, it is thought that carbon stocks in plants and soil have increased during the fossil fuel era [44], due to stimulation of photosynthesis by CO_2 increases [53]. Another hypothesis is related to global warming. The increasing amount of CO_2 (which is a greenhouse gas) in the atmosphere raises the global temperature, which could stimulate plant growth, which again would increase the rate with which plants take carbon from the atmosphere for photosynthesis. Despite this, it is clear from measurements of CO_2 taken over the years, that there is more carbon entering the atmosphere than leaving it [35].

Since the Industrial Revolution, which started around 1750, there has been a rise of more than 30% in the concentration of CO_2 in the atmosphere, currently increasing by about 2ppm (parts per million) [39] per year. The amount of CO_2 in the atmosphere before 1750, given by IPCC 2001 [18] on page 185, is 280 ± 10 parts per million (ppm). This value comes from measurements of CO_2 taken from old air trapped in bubbles in layers of

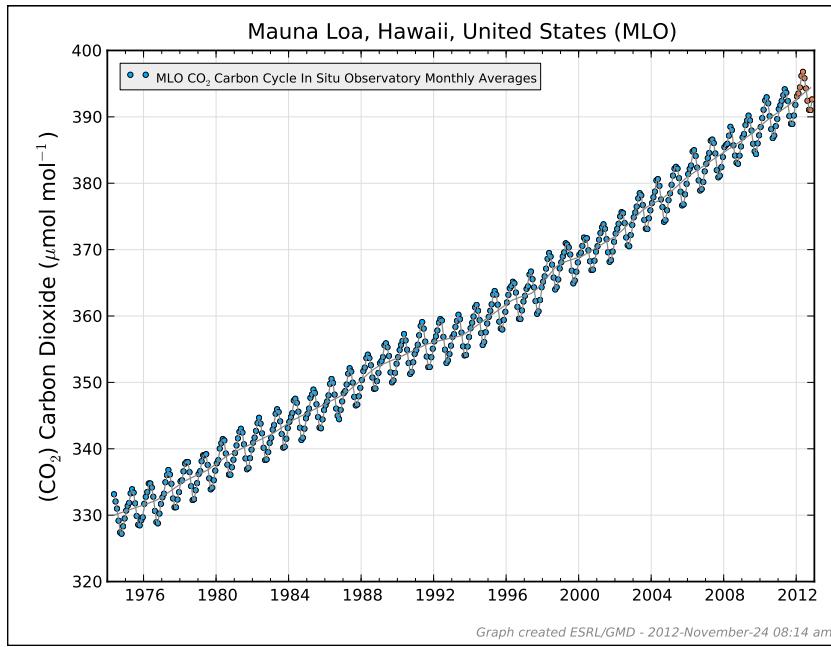


Figure 2.2: This graph shows the increase in the concentration of CO_2 in the atmosphere from 1976 (in parts per million) with projected values for 2013, measured via air intakes at the top of towers in Mauna Loa, Hawaii. (Source: Earth System Research Laboratory (ESRL), Global Monitoring Division).

ice that have built up over over hundreds of years in Antarctica. Measurements taken via air intakes at the top of towers in Manua Loa, Hawai average 390.1 ppm for 2011. Fig. 2.2 shows the increase of CO_2 in the atmosphere between 1976 and 2012.

2.3 Dynamic Global Vegetation Models

There are broadly speaking two ways of determining carbon stocks and budgets, either by empirical methods or by process-based models. The empirical approach measures changes in carbon stocks over time or measures fluxes of carbon. The advantage of this approach is that the measurements are usually cheap and simple. The disadvantages are that there are often gaps in the data and that some carbon pools are difficult to measure (for example carbon stored in soil and fine roots). Also, measurements are generally

confined to limited areas. Through improvement of old methods and introducing new technologies, like for example eddy flux towers, it has become easier to measure CO₂ and water vapour exchange between the land and the atmosphere, but there are still many gaps in the data and filling in the gaps can cause uncertainty and introduce bias [53].

Process-based models simulate the key processes involved in the carbon cycle. The advantages of these models are that predictions can be made about future behaviour and the models can be used over larger areas of land. The disadvantages of these models are that models may not be accurate and that it is difficult to get correct values for some parameters in the models. By tuning parameters incorrectly it is possible to get the right answer, but for the wrong reasons [48] or the wrong answer.

Dynamic Global Vegetation Models (DGVMs) are process-based models used to simulate the terrestrial carbon cycle. Nowadays there exist many quite complex DGVMs, such as DOLY [53], TRIFFID [7], LPJ-DGVM [45], BETHY [22], SEIB-DGVM [41] and ED2 [31]. These models have detailed processes and often include a large number of plant types. For example, LPJ-DGVM allows for ten different plant types, whereas TRIFFID uses five, and SEIB even simulates the local interaction between individual trees. Photosynthesis, evapotranspiration and soil water dynamics are modelled, using monthly, daily, hourly or half-hourly meteorological data and considering different layers of soil. Although many models are able to run independently, they can also be incorporated within general circulation models (GCMs), which are used extensively for weather forecasting, understanding the global climate and climate change. For example, TRIFFID is the DGVM for the community land surface model JULES [1], see Fig. 2.3 and has also been coupled to HADCM3 [7]. In a study by Schaphoff, LPJ-DGVM was coupled to five different GCMs to compare land biosphere carbon uptake between the GCMs under five different simulations of climate change driven by a common emission scenario of CO₂ increase [42]. In Sections 2.3.2, 2.3.1 and 2.3.3 we will look at three different vegetation models: BETHY, DOLY and DALEC in more detail and try to draw some comparisons

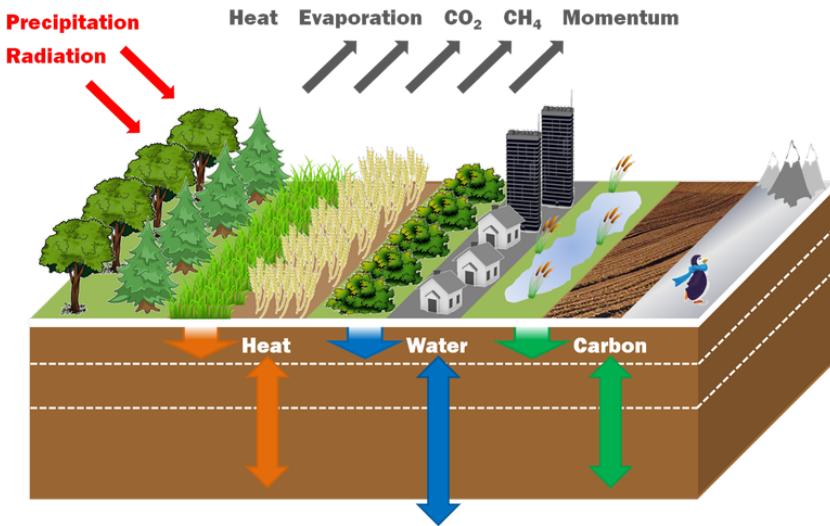


Figure 2.3: An overview of the JULES land surface model, using TROLL as its GVDM (source: jules.jchmr.org/model-description).

between the different models.

2.3.1 The Dynamic Global Phytogeography Model (DOLY)

DOLY, created by Woodward *et al.*, University of Sheffield [53], was developed to simulate aspects of changes in phytogeography (phytogeography is the study of the geographic distribution of plants), with the aim of analysing vegetation responses to changes in the global environment, such as atmospheric CO₂ increase and climatic change. DOLY works on the assumption that global relationships between ecosystem variables and climate can be explained by basic plant physiological models, rather than vegetation maps, demonstrated by Bonan [2]. This makes it possible to generalise global models more. The photosynthesis model used is the one by Farquhar *et. al* [11], which is suitable for global vegetation analysis as it simulates the influence of biochemistry on photosynthesis for a large variety of C3 plants (woody, roundleafed plants) and environmental conditions. It is also possible to extend the model for C4 plants (drier, hotter land plants, such as grasses, sedges, grains) by using the photosynthesis model by Collatz *et al.* [6].

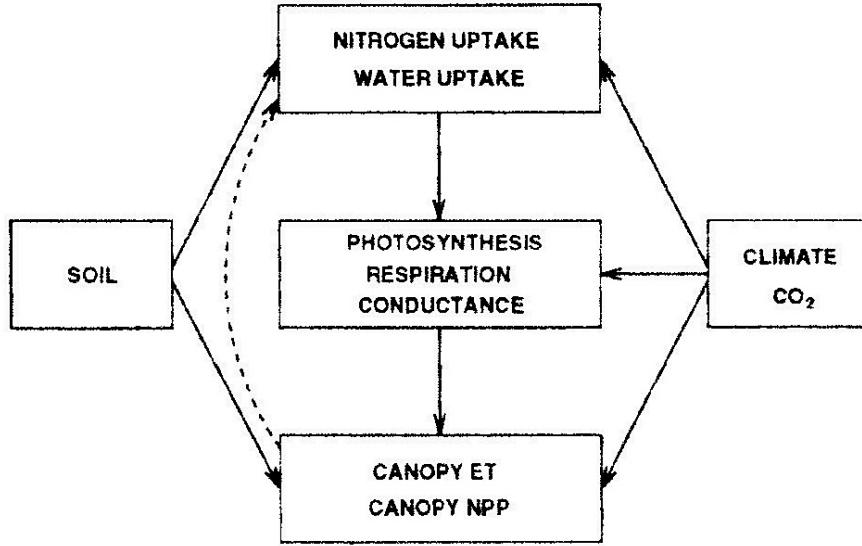


Figure 2.4: *Leaf and canopy gas exchange is predicted based on information on soil nutrient status, water holding capacity, climate and CO_2 concentration. The next step is the prediction of leaf photosynthesis, dark respiration (a form of respiration in plants where carbon dioxide is released without the aid of (sun)light) and stomatal conductance. Finally, these responses are run through a year to predict canopy evapotranspiration (ET) and net primary productivity (NPP). The dashed arrow indicates feedback from canopy processes to plant uptake of water and leaf gas exchange [53].*

The DOLY model is split up into independent modules and is used to simulate the global distributions of leaf area index and annual Net Primary Productivity, NPP, which is the gross primary production GPP (the amount of carbon fixed by photosynthesis) minus plant respiration. Observations of soil carbon and nitrogen are used to constrain the model. Inputs to the model are data on soil nutrient status, water holding capacity, climate and CO_2 concentration. Outputs of the model are canopy evapotranspiration and NPP. Gas exchange is dependent on stomatal conductance (the speed at which water evaporates from pores in a plant), which depends on temperature and soil moisture. Fig. 2.4 shows an overview of the model.

DOLY was tested using climate data assembled by Leemans and Cramer [25] and soil carbon and nitrogen data assembled by Zinke and co-workers [55]. With these data the

model could be evaluated on a yearly basis for each $0.5^\circ \times 0.5^\circ$ grid on Earth using climate information calculated from long-term monthly means of temperature, rainfall and relative humidity. Solar radiation was calculated by standard methods [19]. The time-step for the model is one day, the influence of the day and night cycle was approximated, temperature and relative humidity data were interpolated and rainfall was assumed to be uniformly distributed across the days of the month.

Although the results of the experiments with DOLY are reasonable, indicating that it can provide useful understanding of global vegetation, the creators of DOLY recognised in 1995 that the model only treated the most fundamental environmental factors and processes that determine variables such as leaf area index or primary productivity and that it was not constrained by observations of such variables. Also, some important phenomena, such as fire disturbances, were not considered. The results indicate that the role of climate is very important in determining distribution of vegetation type and function, but that monthly mean climate data are inadequate for the use of detailed global process-based model studies. They also suggest that it is possible to explain the dependence of processes and parameters involved in primary production in terms of functional relationships, rather than in terms of characteristics and parameters of vegetation types. Finally, to quote from [53]: “*Our tests and other studies indicate that the representations in our model of biochemical processes and of the dependence of stomatal conductance on assimilation, temperature and soil moisture are satisfactory*”.

2.3.2 The Biosphere Energy-Transfer Hydrology Scheme (BETHY)

Intercomparison studies, such as the one by Cramer *et al.* 1999 [8] revealed large discrepancies in the results of the various global terrestrial models and give cause for concern. In 2001 Knorr and Heimann published a paper addressing problems with accurate modelling [22]. The paper questioned whether the discrepancies were caused by modelling strate-

gies or by uncertainties in model parameters. A new process-based vegetation model was introduced, called BETHY [23]. BETHY was created to explore different solutions to the problem of calculating net primary production (NPP). Questions the researchers were trying to answer were: (1) What is the most likely global distribution of NPP? (2) Which uncertainties arise from modeling strategies, input data and parameterisation of processes? (3) Is it possible to use field measurements of NPP to reduce these uncertainties?

Knorr *et al.* 2001 believe it is essential to include the link between energy balance and photosynthesis for comprehensive error determination. DOLY includes most of the energy balance and photosynthesis required, but this was not described in the publication by Woodward *et al.* [53]. Other models were deemed to be too complex or the calculations too time consuming, therefore a new scheme was presented, using a different stomatal conductance model. BETHY is flexible, so as to explore various solutions to the NPP calculation problem, allowing two different photosynthesis models to be used as well as a choice of three different vegetation maps. The vegetation micro-climate is simulated on the basis of realistic climate maps, rather than by using mechanistic details (explicitly calculating energy balance and photosynthesis), which are expensive in terms of computer usage. Fig. 2.5 shows the structure of the BETHY model.

BETHY consists of four modules: Energy and Water Balance, Photosynthesis, Phenology and Carbon Balance.

Gridded input data include monthly means of temperature and its diurnal range, precipitation, number of wet days, solar radiation and maps of soil type, soil brightness and type of vegetation cover. Grids are $1^\circ \times 1^\circ = 100 \times 100\text{km}^2$ and exist for the whole planet.

Non-gridded input data include various vegetation-specific parameters (ranging from parameters for tropical evergreen forests to parameters for grass and tundra vegetation, rice, cotton, etc.) and the mean atmospheric CO₂ concentration.

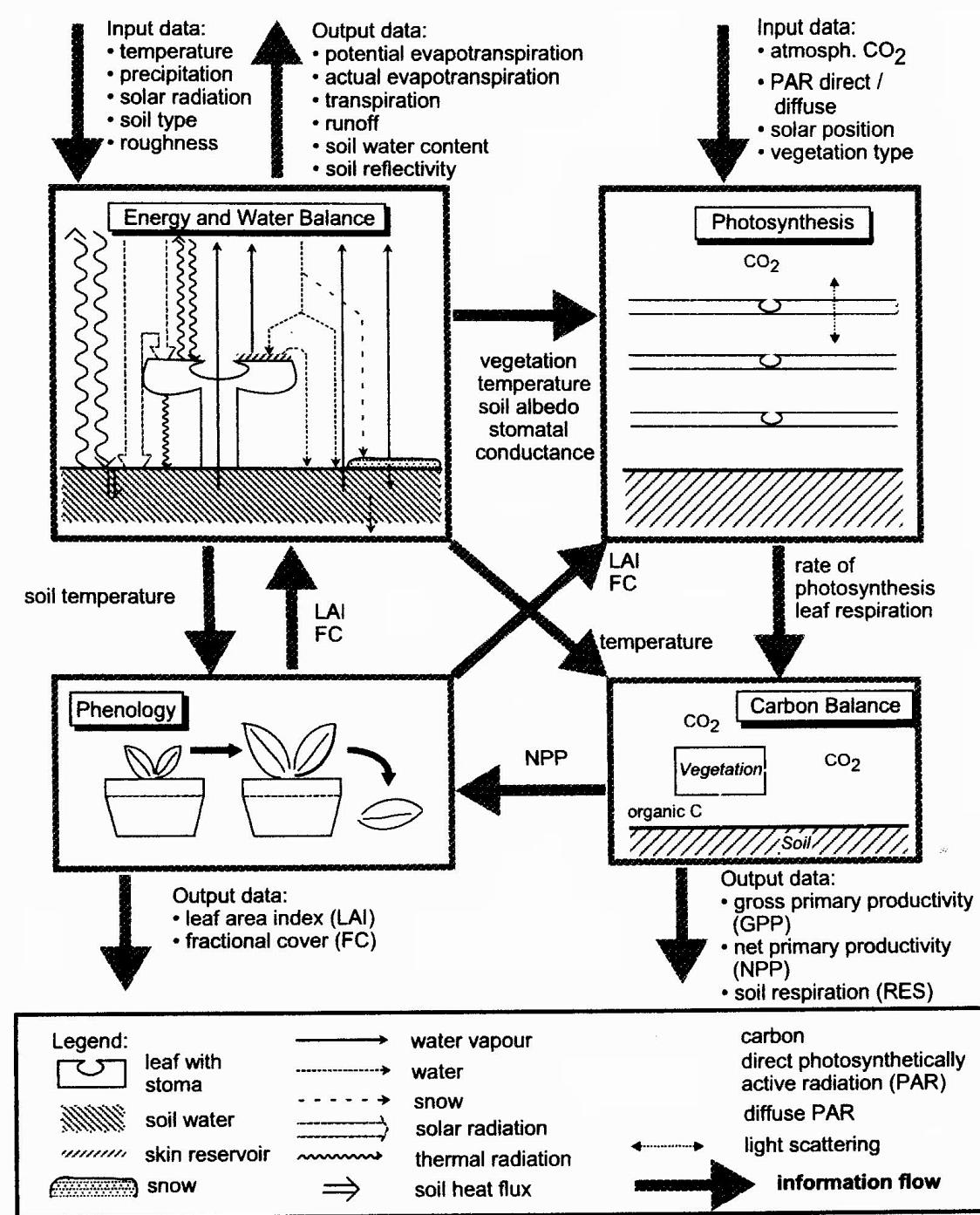


Figure 2.5: Structure of the BETHY model showing input and output as well as flow of information between the four model components and the various fluxes of water, light and carbon [23].

The model outputs include monthly fields of actual and potential evapotranspiration (the sum of evaporation and plant transpiration from the Earth's land surface to the atmosphere, where evaporation is the change of water into gas and transpiration is the water loss from plants through their stomata), transpiration alone, runoff, soil water content and soil reflectance from the energy and water balance module, LAI, fractional cover from the phenology module, GPP, NPP and soil respiration from the carbon balance module.

The model calculates water, energy and carbon balance separately for each grid point. The time step is one day for the water and carbon balance and one hour for the energy balance and photosynthesis.

During the testing experiment 24 different variants of BETHY were tested and compared. Some variants included a change in input data sets or model parts, other variants included a change in certain parameter values. Climate data, including precipitation, near ground daily mean temperature and amplitude were taken from Leemans and Cramer [25]. Other data, including number of wet days, came from station data by Müller [33] and incident critical forcing of photosynthesis (photosynthetic active radiation or PAR) was derived from satellite data from the International Satellite Cloud Climatology Project of 1987 [36]. The soil map data was used from Dunne and Willmot [10], which specifies 128 different types of soil for grids of $0.5^\circ \times 0.5^\circ$; each type gave fixed ranges of percentages of sand, silt and clay. Soil albedo came from a global map by Wilson and Henderson-Sellers [52] of light, medium and dark soils. Three methods of determining vegetation types were: (1) collect data from traditional studies, (2) collect data from time series analysis of the NDVI (normalized difference vegetation index) derived from satellite measurements and (3) assemble data from mean climate. Comparisons were also made with field measurements.

The results demonstrated that large uncertainties still exist for process-based modelling

of global vegetation activity. The paper stated that the discrepancies between different models can probably be explained by existing uncertainties in model parameters, in particular photosynthetic capacity, plant respiration and root depth. It seemed that one of the most important tasks at hand was to improve understanding of plant respiration. One of the other conclusions was that at the time, direct measurements of NPP were probably not suitable for checking global process-based models and that the most serious problem seemed to be the accurate measurement of underground carbon stocks as this is generally underestimated.

2.3.3 Data Assimilation Linked Ecosystem Carbon, DALEC

As we have seen so far, there are two main approaches to determine carbon land stocks: (1) measuring carbon stocks over time or measuring carbon fluxes directly and (2) using process-based models which simulate key processes to approximate carbon stocks. There are problems with both approaches in terms of data gaps, bias and uncertainty for the first approach and problems with creating subjective models using parameters that may be difficult to define, for the second approach. Data assimilation is a technique that combines observable data with a process-based model to produce better estimates of carbon dynamics and to improve the model where necessary. More about data assimilation will be discussed in Section 2.4.

A process-based model especially created for data assimilation and specifically designed for calibration and testing against eddy flux data is DALEC [51]. This model is the main focus of this thesis and there are several versions: DALEC Evergreen, DALEC Deciduous and DALEC Water. The simplest version is DALEC Evergreen. DALEC evergreen involves five carbon pools, namely foliage (C_f), woody stems and coarse roots (C_w), fine roots (C_r), fresh leaf and fine root litter (C_{lit}) and soil organic matter and coarse woody debris (C_{som}), see Fig. 3.1 in Chapter 3.

Inputs to the model (called drivers) are minimum and maximum daily temperatures, atmospheric CO₂ concentration and solar irradiance. The outputs give information on the carbon stocks in the different pools and net ecosystem exchange (NEE) and NPP can be calculated by taking away the appropriate respiration from the GPP ($NPP = GPP - R_a$, $NEE = GPP - R_a - R_h$), where R_a is autotrophic respiration: the carbon that is returned to the atmosphere as CO₂ during plant metabolism and R_h is heterotrophic respiration: respiration of CO₂ during the process of decomposition of organic matter in the soil by soil decomposer organisms (definitions by

<http://www.nrel.colostate.edu/projects/irc/public/Documents/Science/Glossary.htm>).

GPP, calculated by the Aggregated Canopy Model (ACM), is a pseudo-pool, a daily accumulation of photosynthate, which determines the amount of carbon allocated to each of the carbon pools. The GPP is used up completely every day. GPP is a function of the foliar carbon, as foliage produces the photosynthate. Therefore the foliar carbon pool is crucial in determining the amount of carbon the other pools receive and determines the fate of the forest.

Aggregated Canopy Model, ACM

The ACM by Williams *et al.* [50], is a daily step model that estimates GPP as a function of LAI, foliar nitrogen, total daily irradiance, maximum and minimum daily temperature, day length, atmospheric CO₂ concentration, soil-plant water potential and total soil-plant hydraulic resistance. Using either ground based measurements or satellite measurements of leaf carbon mass per area, estimates of the LAI for the ACM are determined from the C_f pool.

The ACM has been derived from the SPA model (Soil-Plant-Atmosphere) by Williams *et al.*, [49]. The SPA model is a process-based multi-layer (fine-scale) model, simulating

ecosystem photosynthesis and water balance at a 30 minute time step and for multiple canopy and soil layers. It uses the Farquhar model of leaf-level photosynthesis and the Penman-Monteith equation to predict evapotranspiration [21]. What makes this model unique is the treatment of the stomatal opening, coupling water flows from the soil to the atmosphere with carbon fixation [48]. The rate at which water can be supplied to the canopy, which is constrained by hydraulic capacity, determines and limits transpiration and affects the stomata openings. The SPA model was developed and parametrised for a mixed deciduous oak and maple wood in central Massachusetts, USA. Given several simplifications and assumptions, it adequately predicts CO₂ uptake by the canopy and transpiration for this forest and the creators felt confident they could “*employ the model to simulate seasonal patterns of C and water exchanges, given the availability of relevant parameters*” [49].

The equations for the ACM were derived from SPA using cumulative or average values of the most sensitive driving variables. Then the ACM was calibrated so that the estimates of the GPP were similar to the estimates of GPP of the SPA model across a wide range of these driving variables. The time step is daily, rather than 30 minutes for the fine-scale model.

The ACM is a broad-leaf model (coarse scale), aggregated model. Aggregated models have the effects of nonlinearities built into the model parameters and so avoid the need for spatial data at leaf level. The parameters though, need to be estimated directly from the coarse scale, ie. the canopy, as fine-scale data are not suitable for these estimations [49]. The time step for the ACM is one day. Daily drivers are average daily temperature, daily temperature range, atmospheric CO₂ concentration, day of year and solar irradiance. The other variables in the ACM function, for example soil-plant water potential and foliar nitrogen, can be set to appropriate values for the evergreen vegetation in question. latitude also plays a big role, as it influences day length and therefore the amount of photosynthate produced.

The ACM has been tested using measurements of NEE from an eddy covariance system at Harvard Forest, Massachusetts, USA. The eddy covariance technique determines the exchange rate of CO₂ between the atmosphere and a plant canopy by measuring the covariance between fluctuations in vertical wind velocity and CO₂ mixing ratio. Eddy covariance data is collected from flux towers. The ACM has also been tested on wet sedge tundra in the Arctic and on a range of forest types across Oregon Transect Ecosystem Research (OTTER) sites in Oregon, USA. In all cases the ACM fitted the data well. The creators state that the most important drivers for the ACM are irradiance and leaf area index. Soil moisture stress, foliar nitrogen and ambient CO₂ are in second place. “*We can be confident that we have included the key driving variables in our aggregated model necessary to make predictions of ecosystem GPP over a wide range of sites and conditions. the aggregated model should provide an effective tool for developing regional estimates of C uptake and should be easily incorporated into regional models of C cycling*” [50].

The creators of DALEC, Williams *et al.*, recognised that parameter definition is of the greatest importance and so they initiated a project called REFLEX [13]. The aims of the project were to use different data assimilation techniques and models to estimate carbon model parameters and to predict carbon fluxes. It also aimed to address errors and biases which are introduced when predicting fluxes and using measured data from flux towers. Synthetic data was provided for NEE and LAI, generated from DALEC with added noise, and observed NEE and LAI data from European flux towers. Some results of this project [13] indicated that some parameters linked to GPP and respiration were quite well characterised, whilst other parameters related to allocation and turnover of fine root and wood pools were not so well constrained. It was found to be important to add constraints to the model to help to reduce uncertainties for those model parameters which were not supported by eddy covariance data. This means adding data on the wood, soil and fine root carbon pools.

2.3.4 Comparison of the Models

It is not easy to compare the three process-based carbon models described in this literature review, based on the papers read. The main difference seems to be that DOLY does not use vegetation specific data, such as vegetation maps, but relies on basic plant physiological models and the Farquhar model for photosynthesis for C3 plants [11], although an extension for C4 plants allows a different photosynthesis model by Collatz [6] to be used. DOLY is constrained by observations of soil carbon and nitrogen. BETHY is quite a detailed and complex model, using different vegetation maps and able to run two different photosynthesis models. DALEC is vegetation specific; there are separate versions, such as for evergreen and deciduous vegetation. It varies within these versions by changing latitude and parameters. BETHY and DOLY are both modular, whilst DALEC has a more integrated approach.

Both DALEC (through the REFLEX project) and BETHY found that more data on underground carbon stocks needed to be established. BETHY, DOLY and DALEC all realise that climate is important in estimating carbon stocks. BETHY has been used in data assimilation projects by the Carbon Cycle Data Assimilation System (CCDAS), which was built around BETHY and coupled to atmospheric transport models [24]; it has also been incorporated into a land surface model, Jena Scheme for Biosphere-Atmosphere Coupling (JSBACH) at the Max Planck Institute for Meteorology in Hamburg, Germany, see Fig. 2.6. DOLY has been incorporated into the Sheffield Global Dynamical Vegetation Model (SGDVM), see Fig. 2.7. SGDVM and SPA/DALEC are in use within the Centre for Terrestrial Carbon Dynamics [38]. DALEC was also in use with Abacus, WP7: Modelling and synthesis (Leader: Williams and Harding, Blythe, Disney, Moncrieff).

BETHY and DOLY both used climate data assembled by Leemans and Cramer [25] but used different other data to test the models. It would be interesting to see how the three models would compare when each was given the exact same data, appropriate to their

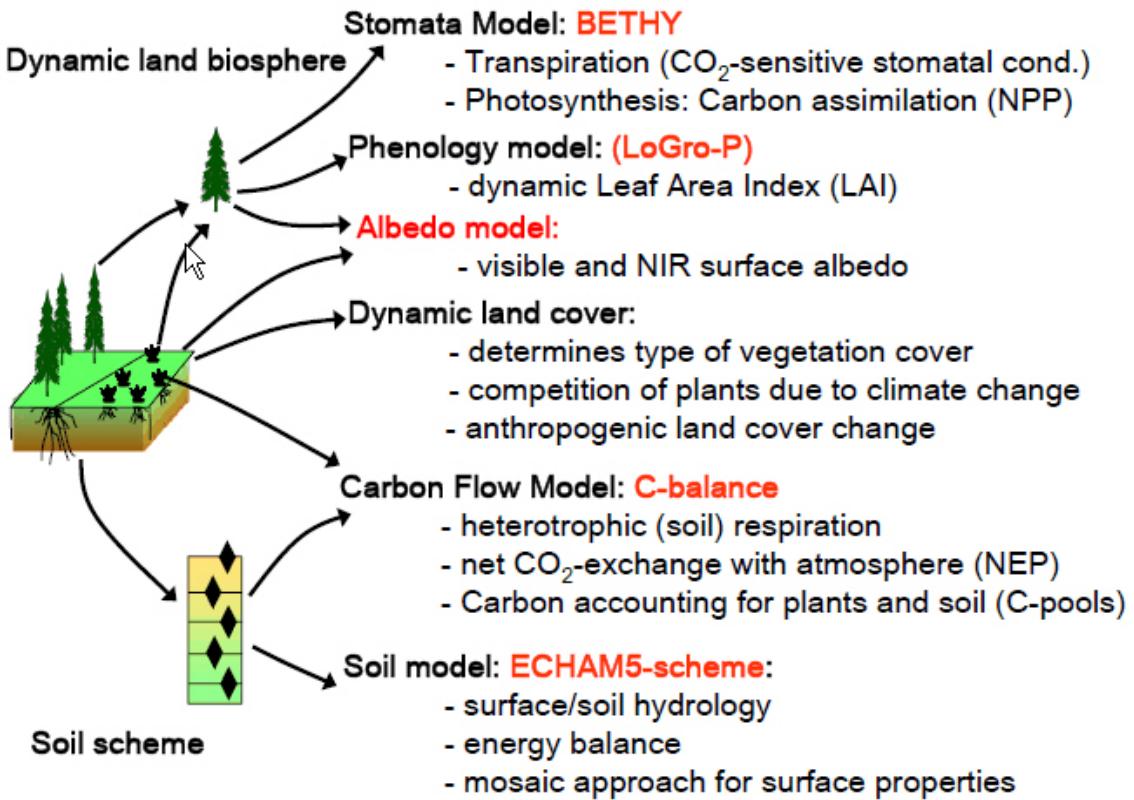


Figure 2.6: *Modules of JSBACH*, (Source: Max Planck Institute for Meteorology).

requirements.

2.4 Data Assimilation

DALEC, which is the subject of this thesis, is a model, like many other models, that can be used in conjunction with data assimilation. This section describes in short what data assimilation is and where it comes from. Although in this thesis we do not use 4DVAR with the goal of finding state variables or parameters, we do employ some four-dimensional data variation (4DVAR) techniques when we try to fit functions to data. Therefore we also describe in Section 2.4.2 the 4DVAR technique.

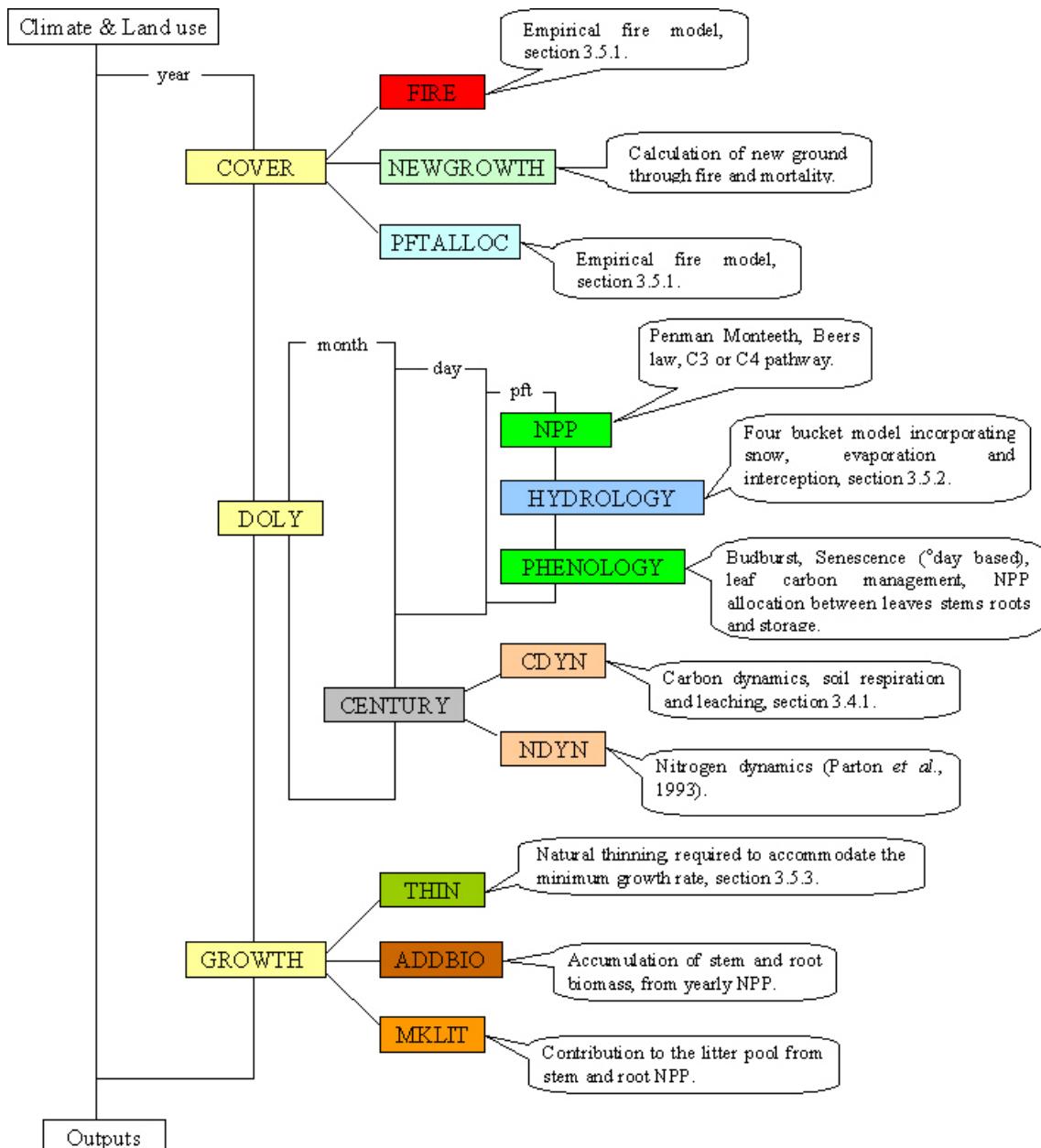


Figure 3: Loop structure of the SDGVM.

Figure 2.7: Structure of the SDGVM, (Source: Centre for Terrestrial Carbon Dynamics).

2.4.1 What is Data Assimilation?

Recently it has become possible and more and more important to model the system dynamics of environmental, biological, physical or economic systems in a mathematical and numerical manner. If the initial values of these mathematical models are known it is in theory possible to predict the behaviour of the system. In reality, the models often contain errors and if the initial values are not precise (especially in chaotic systems), the predictions can be incorrect. To improve the quality of the models, observations over time can be incorporated in the mathematical equations so that new estimates can be derived. Data assimilation is a technique that does precisely this and can be applied to all general models, whether they be simple linear, deterministic, continuous ordinary differential equation models or sophisticated nonlinear stochastic partial-differential continuous or discrete models. Data assimilation strives to derive as accurate estimates of the current and future states as possible, whilst also taking into account the level of uncertainty in the estimations [34].

There are two approaches to data assimilation, real-time (sequential) or retrospective (non-sequential). Real-time assimilation takes into account observations made in the past up until the time of analysis. Retrospective data assimilation can also take future observations into account (for example when data are re-analysed). The observations can be processed in small batches (intermittent) or they can be considered over a long time (continuous), see Fig. 2.8. An example of a retrospective data assimilation algorithm is the Kalman smoother, an example of real time data analysis algorithm is four-dimensional variational data assimilation or 4DVAR.

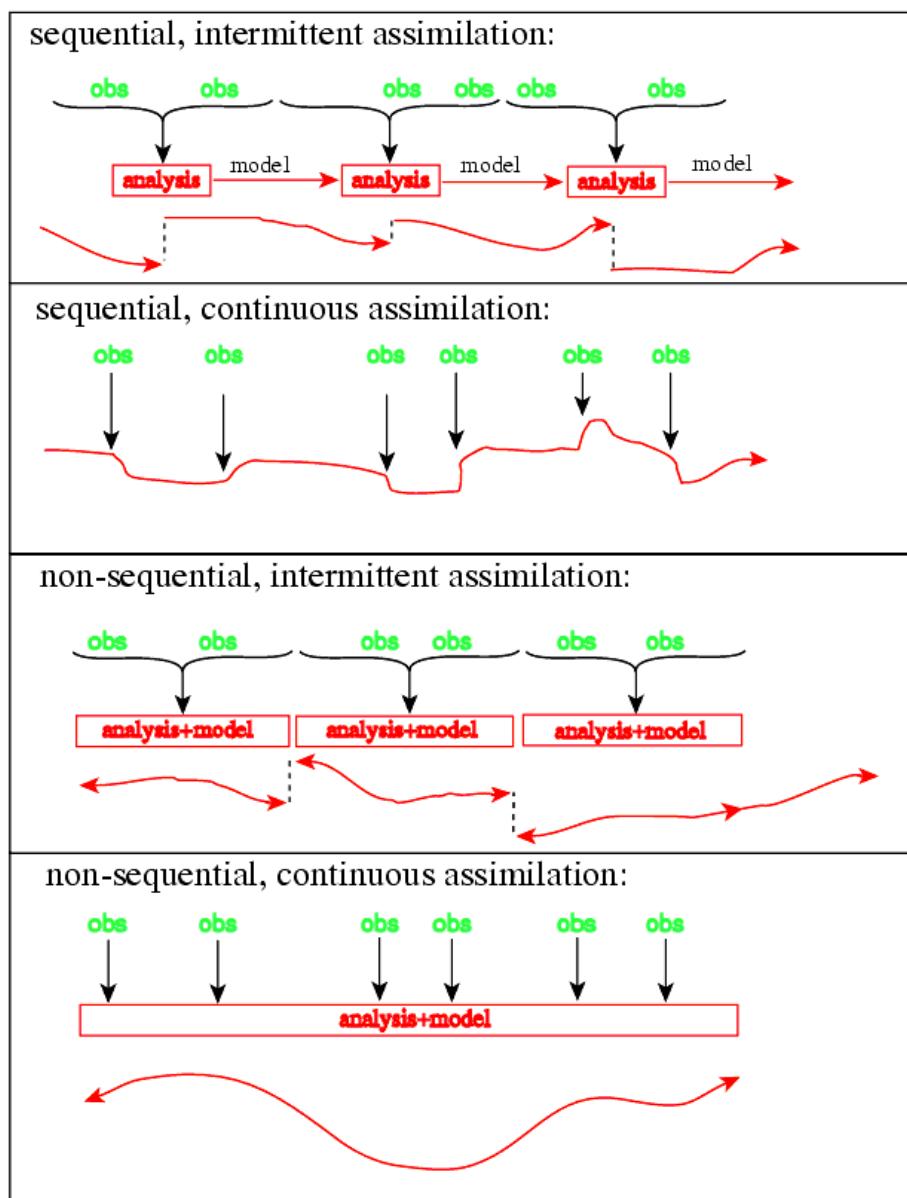


Figure 2.8: A graph showing the different data assimilation approaches as a function of time. (Source: ECMWF website).

2.4.2 4DVAR Data Assimilation

Four-dimensional variational data assimilation takes into account both the space dimension of observations and the time window that they are taken from. The aim is to try and fit a model trajectory which is as close to the observations, measured over a particular assimilation window (time window), as possible. This produces an optimal solution, or analysis state. To achieve the optimal solution a cost function is introduced which describes the fit between the analysed trajectory and the observations. This trajectory is based on an initial estimate as to the whereabouts of the initial conditions of the observational trajectory, see Fig. 2.9. The cost function is often minimised by computing the gradient with respect to these estimated initial conditions (also called control variables), using an iterative algorithm. During this algorithm a new estimate for the initial conditions is found and a new trajectory is created from these new values, which is an improvement on the last estimate and fits the observations better. One feature of 4DVAR is a so-called background term. This is an additional term that is included in the cost function and contains information about the initial condition for the assimilation window obtained from a previous forecast.

2.4.3 The Cost Function

The cost function, J is, in its simplest form, just a least squares fit. For 4DVAR data assimilation the cost function is split into a background term, J^b , which shows the fit between the model state and the background, and an observation term, J^o , which shows the fit between model state and the observations, see Fig. 2.10:

$$J = J^b + J^o$$

Consider a discrete case of 4DVAR, although everything discussed here also applies to

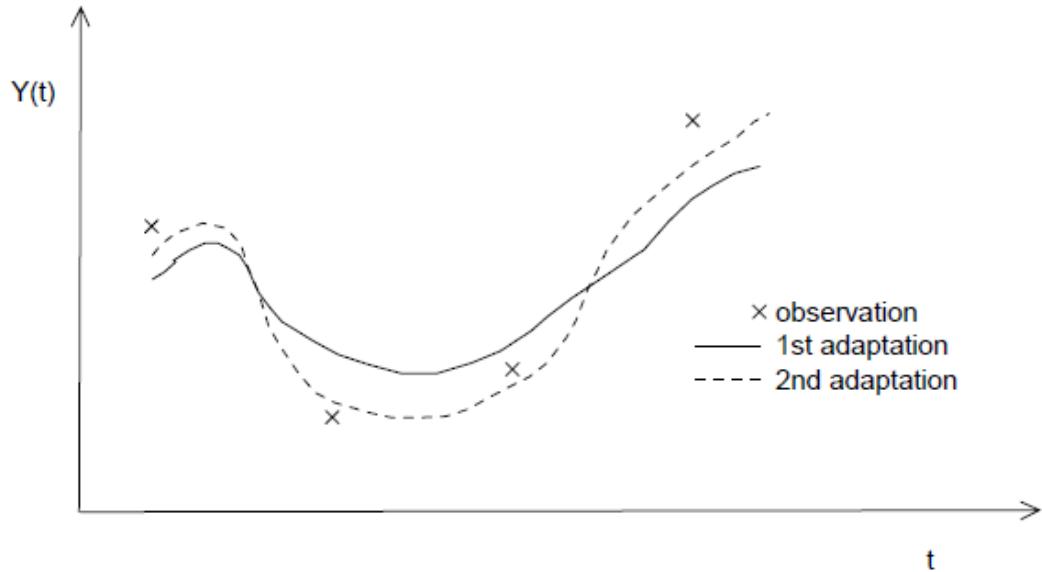


Figure 2.9: *Schematic representation of variational methods [15]: 1st adaption and 2nd adaption are trajectories created from newly estimated initial conditions, found by the data assimilation algorithm..*

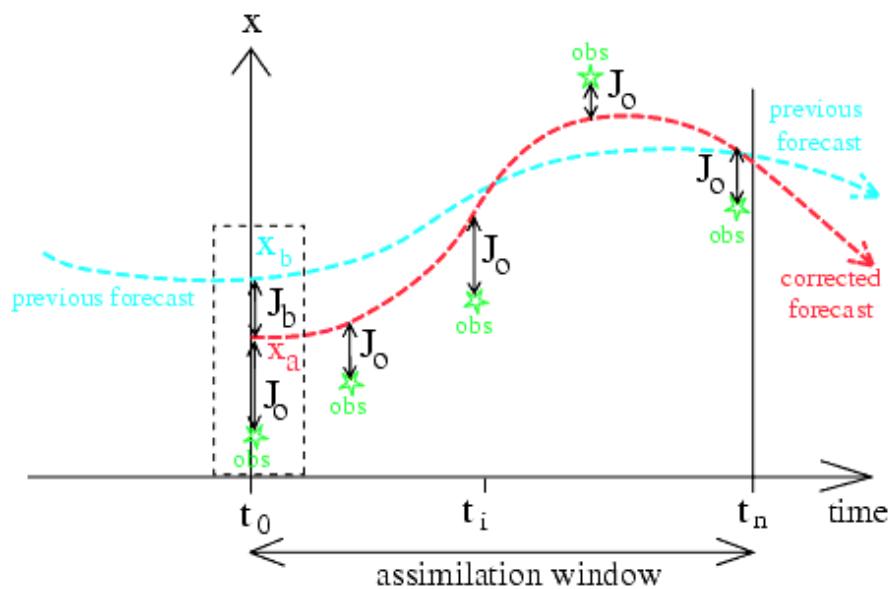


Figure 2.10: *The analysis state, x_a , (x_0 in our cost function (2.1)), is positioned between the background state, x_b , and first observation (obs) [40].*

the continuous case (which is not widely used in real models, where the adjoint of the discretised model must be used [4] and is not described in this thesis).

Equations for the terms J^b and J^o are:

$$J^b(\mathbf{x}_0) = \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}_b)^T \mathbf{B}^{-1}(\mathbf{x}_0 - \mathbf{x}_b)$$

and

$$J^o(\mathbf{X}) = \frac{1}{2} \sum_{i=0}^{N-1} (\mathbf{o}_i - \mathbf{H}(\mathbf{x}_i))^T \mathbf{R}^{-1} (\mathbf{o}_i - \mathbf{H}(\mathbf{x}_i))$$

where $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \dots, \mathbf{x}_N\}$,

And therefore the cost function J is defined as follows:

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}_b)^T \mathbf{B}^{-1}(\mathbf{x}_0 - \mathbf{x}_b) + \frac{1}{2} \sum_{i=0}^{N-1} (\mathbf{o}_i - \mathbf{H}(\mathbf{x}_i))^T \mathbf{R}^{-1} (\mathbf{o}_i - \mathbf{H}(\mathbf{x}_i)). \quad (2.1)$$

subject to the constraint:

$$\mathbf{x}_{i+1} = \mathbf{M}(\mathbf{x}_i), \quad i = 0, \dots, N-1.$$

The constraint is strong which means that the model errors are assumed to be small in comparison with the other sources of errors and therefore they can be neglected (assumption of perfect model).

$\mathbf{M} : \mathbb{R}^n \rightarrow \mathbb{R}$. \mathbf{M} is the model containing the mathematical equations describing the system we are interested in, propagating it forward in time. For example \mathbf{M} could be represented by the DALEC model.

$$J^b : \mathbb{R}^n \rightarrow \mathbb{R}.$$

$$J^o : \mathbb{R}^{n \times N} \rightarrow \mathbb{R}.$$

$\mathbf{x}_i \in \mathbb{R}^n$ is the state vector at time i .

$\mathbf{x}_0 \in \mathbb{R}^n$ is the value of the state vector at time 0.

$\mathbf{x}_b \in \mathbb{R}^n$ is the background vector.

$\mathbf{B} \in \mathbb{R}^{n \times n}$ is the background error covariance matrix, containing information on the background.

$\mathbf{o}_i \in \mathbb{R}^n$ is the vector of observations at time i .

\mathbf{H} is the observation operator. Data that has been produced by the forward model during an estimation may not be of the same type as the observations we are considering. The observation operator adjusts the estimated data in such a way that it is compatible with the observations.

$\mathbf{R} \in \mathbb{R}^{n \times n}$ is the observation error covariance matrix.

In this thesis the background term, J^b , is ignored and \mathbf{R} is set to the identity matrix, \mathbf{I} ; we assume the observations are perfect. Data produced by the forward model is of the correct type, so that $H(\mathbf{x}_i) = \mathbf{x}_i$.

2.4.4 The 4DVAR Algorithm

The discrete case for the 4DVAR data assimilation problem is described as follows:

Minimise the cost function $J(\mathbf{X})$, where $\mathbf{X} = (\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_N)$, defined over an assimilation window $[t_0, t_N]$, $0 \leq t_0 < t_N$, subject to the constraint $\mathbf{x}_{i+1} = \mathbf{M}(\mathbf{x}_i)$, where $\mathbf{x}_i \in \mathbb{R}^n$ is a state vector, \mathbf{M} is the mathematical model representing an environmental, biological, physical or economic system. $J : \mathbb{R}^{n(N+1)} \rightarrow \mathbb{R}$, $\mathbf{M} : \mathbb{R}^n \rightarrow \mathbb{R}^n$. All functions are assumed to be sufficiently smooth.

The constraint, $\mathbf{x}_{i+1} = \mathbf{M}(\mathbf{x}_i)$, is strong, which means model error is neglected.

The best way to minimise the cost function, subject to the constraint, is to introduce

Lagrange multipliers, $\boldsymbol{\lambda}_i \in \mathbb{R}^n$, $i = 1...N$:

$$G(\mathbf{X}, \boldsymbol{\Lambda}) = J(\mathbf{X}) + \sum_{i=0}^{N-1} \boldsymbol{\lambda}_{i+1} \cdot (\mathbf{x}_{i+1} - \mathbf{M}(\mathbf{x}_i)), \quad (2.2)$$

where $\boldsymbol{\Lambda}_i = (\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2, \dots, \boldsymbol{\lambda}_N) \in \mathbb{R}^{n \times N}$.

The cost function describes the fit between the previous trajectory and the observations. The previous trajectory is based on an initial estimate of the initial conditions for the observational trajectory. As stated above it is required to minimise the cost function with respect to the initial conditions.

The gradient of G with respect to \mathbf{x}_i is:

$$\nabla_{\mathbf{x}_i} G = \nabla_{\mathbf{x}_i} J + \boldsymbol{\lambda}_i - \mathbf{M}_x^T(\mathbf{x}_i) \boldsymbol{\lambda}_{i+1}, \quad i = 1...N-1. \quad (2.3)$$

Requiring $\nabla_{\mathbf{x}_i} G = 0$ we can now see that:

$$\boldsymbol{\lambda}_i = \mathbf{M}_x^T(\mathbf{x}_i) \boldsymbol{\lambda}_{i+1} - \nabla_{\mathbf{x}_i} J, \quad i = 1...N-1. \quad (2.4)$$

and by analogy:

$$\boldsymbol{\lambda}_0 = \mathbf{M}_x^T(\mathbf{x}_0) \boldsymbol{\lambda}_1 - \nabla_{\mathbf{x}_0} J. \quad (2.5)$$

We refer to $\boldsymbol{\lambda}_i$ in Equation (2.4) as the adjoint variables. $\mathbf{M}_x^T(\mathbf{x}_i)$ is the adjoint model and $\nabla_{\mathbf{x}_i} J$ is called the adjoint forcing or the innovation vector. Also from (2.2), we note that $\nabla_{\mathbf{x}_0} G = -\boldsymbol{\lambda}_0$.

The gradient of G with respect to \mathbf{x}_0 is:

$$\nabla_{\mathbf{x}_0} G = \nabla_{\mathbf{x}_0} J - \mathbf{M}_x^T(\mathbf{x}_0) \boldsymbol{\lambda}_1. \quad (2.6)$$

We set the adjoint variables at time N to zero:

$$\boldsymbol{\lambda}_N = 0. \quad (2.7)$$

The gradient of G with respect to $\boldsymbol{\lambda}_{i+1}$ is:

$$\begin{aligned}\nabla_{\boldsymbol{\lambda}_{i+1}} G &= \mathbf{x}_{i+1} - \mathbf{M}(\mathbf{x}_i), \quad i = 0 \dots N-1 \\ \nabla_{\boldsymbol{\lambda}_{i+1}} G &= 0 \implies \mathbf{x}_{i+1} = \mathbf{M}(\mathbf{x}_i), \quad i = 0 \dots N-1.\end{aligned}$$

This is the *forward model* and is exactly the equation for the constraint. The forward model is used to calculate \mathbf{x}_i as it is propagating itself forward in time from the initial conditions.

The objective is to find the gradient of G with respect to \mathbf{x}_0 , so that we can find the minimum of the cost function and run the forward model from the new initial values found.

This is done according to the instructions below:

1. Calculate all \mathbf{x}_i using the forward model $\mathbf{x}_{i+1} = \mathbf{M}(\mathbf{x}_i)$, $i = 0, \dots, N-1$, from a estimate of the control vector (the initial values, \mathbf{x}_0) to ensure that $\nabla_{\boldsymbol{\lambda}_{i+1}} G = 0$, $i = 0, \dots, N-1$
2. The adjoint vectors, $\boldsymbol{\lambda}_i$ are chosen such that $\nabla_{\mathbf{x}_i} G = 0$ for $i = 0, \dots, N-1$:

$$\boldsymbol{\lambda}_N = 0,$$

$$\boldsymbol{\lambda}_i = \mathbf{M}_{\mathbf{x}}^T(\mathbf{x}_i) \boldsymbol{\lambda}_{i+1} - \nabla_{\mathbf{x}_i} J, \quad i = N-1, \dots, 0$$

until we have found $\boldsymbol{\lambda}_0$, see equation (2.5). The vectors $\boldsymbol{\lambda}_i$ are calculated backwards using the already calculated \mathbf{x}_i .

3. Find the gradient of G with respect to \mathbf{x}_0 , using equation (2.2), which leads to:

$$\nabla_{\mathbf{x}_0} G = -\boldsymbol{\lambda}_0.$$

4. To minimise G a minimisation algorithm is used, for example Newton's method, method of steepest descent, etc. [43].

Once G is minimised, we have found the initial values of the trajectory that fits the observations best.

2.5 Conclusion

In this section we have considered why it is necessary to understand the Earth's systems such as the carbon cycle and have looked at what techniques, processes and models are available. We have compared three dynamic vegetation models: DOLY, BETHY and DALEC, which are all in use currently and were created around the same time. As many models are used in conjunction with data assimilation, and DALEC in particular, we have discussed what data assimilation is, where it comes from and how it is used. Four-dimensional data variational assimilation was described in more detail, as this technique is used in this thesis.

Chapter 3

Models, Data and Assumptions

This chapter describes and explains the model equations for the DALEC evergreen and deciduous version, DALEC EV and DALEC DE, as well as the Aggregated Canopy Model, which is the photosynthesis model in DALEC. We also describe the assumptions we have made and discuss the data used to run our analyses.

3.1 DALEC Evergreen

DALEC EV [51] describes the time evolution of five carbon pools, namely foliage (C_f), fine roots (C_r), woody stems and coarse roots (C_w), fresh leaf and fine root litter (C_{lit}) and soil organic matter and coarse woody debris (C_{som}) and one pseudo pool, Gross Primary Production (GPP). Fig. 3.1 shows the carbon cycle of an evergreen forest according to DALEC. The pools are connected by arrows, which represent allocation and production fluxes. Sunlight, water and CO₂ are needed for photosynthesis which takes place in the needles. The photosynthate is stored in the GPP, which distributes the available carbon to the C_f , C_w and C_r pools. During photosynthesis carbon is also respired into the atmosphere by the tree; a process called autotrophic respiration (R_a). Decomposition of

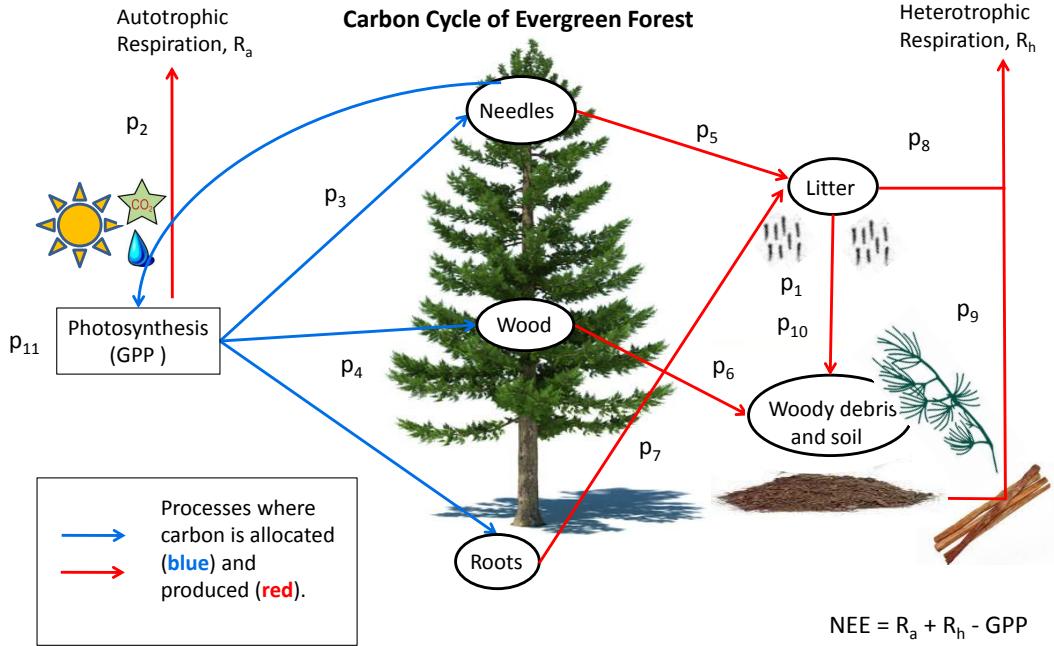


Figure 3.1: A schematic overview of the carbon cycle of an evergreen needle leaf forest according to DALEC EV.

litter in and on the soil results in heterotrophic respiration (R_h).

The DALEC EV model equations [13] written as a discrete dynamical system, take the form of daily maps, specifically:

$$C_f(t+1) = (1 - p_5)C_f(t) + p_3(1 - p_2)GPP(C_f(t), t), \quad (3.1)$$

$$C_r(t+1) = (1 - p_7)C_r(t) + p_4(1 - p_3)(1 - p_2)GPP(C_f(t), t), \quad (3.2)$$

$$C_w(t+1) = (1 - p_6)C_w(t) + (1 - p_4)(1 - p_3)(1 - p_2)GPP(C_f(t), t), \quad (3.3)$$

$$C_{lit}(t+1) = (1 - (p_8 + p_1)T(t))C_{lit}(t) + p_5C_f(t) + p_7C_r(t), \quad (3.4)$$

$$C_{som}(t+1) = (1 - p_9T(t))C_{som}(t) + p_6C_w(t) + p_1T(t)C_{lit}(t), \quad (3.5)$$

where t is time in days and $T(t) = 0.5 \exp(p_{10}T_m(t))$ is the temperature sensitive rate parameter, which is important for decomposition. $T_m(t)$ is the mean daily air temperature.

For a description of the parameters, p_i , $i = 1, \dots, 10$ see Appendix C. Typically, we have used the following initial conditions for the five pools: $C_f(0) = 150 \text{ gC m}^{-2}$, $C_r(0) = 160 \text{ gC m}^{-2}$, $C_{lit}(0) = 60 \text{ gC m}^{-2}$, $C_w(0) = 9200 \text{ gC m}^{-2}$ and $C_{som}(0) = 11000 \text{ gC m}^{-2}$. These values of the initial conditions for the wood and soil and organic carbon pools are based on suggestions made by the REgional FLux Estimation eXperiment (REFLEX) [13] for the forest data we used. The initial conditions for the foliar pool, the fine roots and litter pool are estimations based on LAI available.

3.2 DALEC Deciduous

DALEC Deciduous, DALEC DE, is another version of the DALEC model, see Fig. 3.2. The main difference between DALEC EV and DALEC DE is that DALEC DE has an extra labile pool, C_{lab} , which is used in the spring to kick-start the growth of leaves. The labile pool depletes over the spring and then builds up again in the autumn, when the leaves are dropping.

As for DALEC EV, the DALEC DE model equations [13] take the form of daily maps, specifically:

$$\begin{aligned} C_f(t+1) &= (1 - p_5(p_{14} + (1 - p_{14})T(t))m_{tf})C_f(t) \\ &\quad + (\min(p_{17} - C_f(t), p_3(1 - p_2)GPP(C_f(t), t))) \\ &\quad + p_{15}(1 - p_{16})C_{lab}(t)T(t)m_{tl}, \end{aligned} \quad (3.6)$$

$$C_{lab}(t+1) = (1 - p_{15}T(t)m_{tl})C_{lab}(t) + p_5(1 - p_{14})(1 - p_{16})C_f(t)T(t)m_{tf}, \quad (3.7)$$

$$\begin{aligned} C_r(t+1) &= (1 - p_7)C_r(t) + p_4((1 - p_2)GPP(C_f(t), t)) \\ &\quad - m_{tl}(\min(p_{17} - C_f(t), p_3(1 - p_2)GPP(C_f(t), t))), \end{aligned} \quad (3.8)$$

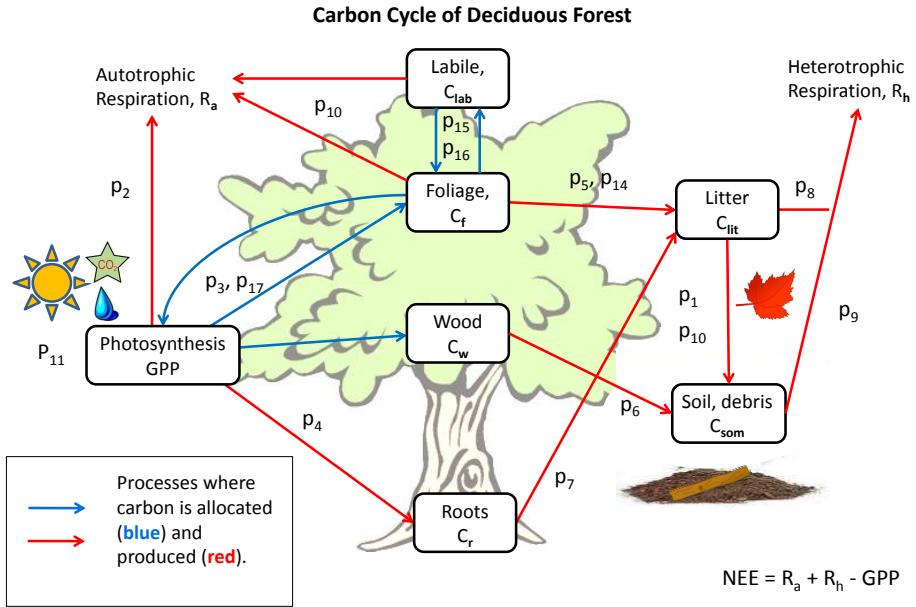


Figure 3.2: A schematic overview of the carbon cycle of a deciduous forest according to DALEC DE. Parameters p_{12} and p_{13} are not shown, but are used to determine when m_{tl} and m_{tf} are switched on and off.

$$\begin{aligned} C_w(t+1) &= (1-p_6)C_w(t) + (1-p_2)(1-p_4)GPP(C_f(t), t) \\ &\quad (1-p_4)\min(p_{17} - C_f(i), p_3(1-p_2)GPP(C_f(t), t))m_{tl}, \end{aligned} \quad (3.9)$$

$$C_{lit}(t+1) = (1-(p_1+p_8)T(t))C_{lit}(t) + p_5p_{14}m_{tf}C_f(t) + p_7C_r(t), \quad (3.10)$$

$$C_{som}(t+1) = (1-p_9T(t))C_{som}(t) + p_6C_w(t) + p_1T(t)C_{lit}(t), \quad (3.11)$$

where t is time in days and $T(t) = 0.5 \exp(p_{10}T_m(t))$ is the temperature sensitive rate parameter. $T_m(t)$ is the mean daily air temperature. There are 17 parameters in DALEC DE; for a description see Appendix C. The variable m_{tf} switches the turnover of foliar carbon on (in the autumn, so that the trees drop leaves) or off (during the summer, when the trees do not drop leaves and in the spring, when the leaves are growing) and m_{tl} switches the turnover of labile carbon on (used to kick-start the foliar carbon in the spring) or off (in the summer, when it has been depleted).

We have used the following initial conditions for the five pools: $C_f(0) = 0 \text{ g C m}^{-2}$, $C_{lab}(0) = 100 \text{ g C m}^{-2}$, $C_r(0) = 5 \text{ g C m}^{-2}$, $C_{lit}(0) = 5 \text{ g C m}^{-2}$, $C_w(0) = 5 \text{ g C m}^{-2}$ and $C_{som}(0) = 9900 \text{ g C m}^{-2}$. The GPP is the same as for the DALEC EV model, as given in equation (3.12). The parameter values and the values of the initial conditions for the carbon pools are based on research by Hill *et. al* [17]. We used the same climate data that we used for DALEC EV [13] in its unsmoothed version.

We see that in the DALEC DE model, the foliar and labile pools are coupled but decouple from the other four equations.

3.3 The Aggregated Canopy Model, ACM

The GPP represents a daily accumulation of carbon produced by photosynthesis and is the same for DALEC EV and DALEC DE. It takes the following form:

$$GPP(C_f(t), t) = \frac{E_0(t)I(t)g_c(t)(C_a(t) - C_i(t))}{E_0(t)I(t) + g_c(t)(C_a(t) - C_i(t))}(a_2s(t) + a_5). \quad (3.12)$$

The functions in the GPP are collectively called the Aggregated Canopy Model (ACM). A description of the various functions in the GPP can be found in Appendix A.

In order to compute the GPP, input of daily drivers of maximum and minimum temperature (which produce the temperature range), atmospheric CO₂ concentration, solar irradiance and day of the year are needed. The other parameters in the ACM, for example soil-plant water potential and foliar nitrogen, can be set to appropriate fixed values for the vegetation in question. Latitude also plays an important role, as it influences day length and therefore the amount of photosynthate produced, but again this is another fixed parameter.

In the DALEC model the GPP depends on the foliar carbon, C_f , via two functions, namely

$E_0(t)$ (canopy level quantum yield) and $C_i(t)$ (CO_2 concentration at site of carboxylation). This is because foliage produces the photosynthate, which none of the other carbon pools do. Therefore the foliar carbon pool is crucial in determining the amount of carbon the other pools receive and we will show that it is this pool that determines the fate of the forest, see Section 4.3.2.

3.4 NEE

From the model it is also possible to evaluate other quantities. These include the net ecosystem exchange (NEE), which is the net carbon uptake or loss by the ecosystem, and net primary production (NPP), which is the net carbon flux from the atmosphere into green plants. NEE and NPP can be calculated as follows:

$$\text{NEE}(t) = \text{GPP}(C_f(t), t) - R_a(t) - R_h(t), \quad (3.13)$$

$$\text{NPP}(t) = \text{GPP}(C_f(t), t) - R_a(t), \quad (3.14)$$

where

$$R_a(t) = p_2 \text{GPP}(C_f(t), t) \quad (3.15)$$

and

$$R_h(t) = R_{h_1}(t) + R_{h_2}(t) = p_8 T(t) C_{lit}(t) + p_9 T(t) C_{som}(t) \quad (3.16)$$

3.5 Model Assumptions and Data

In order to run the DALEC model, estimates of the parameters and values for the driver data are needed. For the DALEC EV version we have used realistic data both for parameters and climate data from a Netherlands flux site, Loobos, which was provided for the REFLEX project [13]. The purpose of this project was to compare the strengths and

weaknesses of various data assimilation methods for estimating parameters and predicting carbon fluxes using DALEC, based on three years of driver data. However, in our analysis of DALEC EV the behaviour of the model is studied over a longer period of time, so driver data was created by averaging the three years of available data (including the atmospheric carbon, C_a), creating average daily data. This produces data which varies on a daily basis in a realistic manner, but is the same each year. The averaging introduces some smoothing but this has little qualitative impact.

The time scale has been shifted ten days back from the 1st January, so that $t = 0$ occurs on the 21st December, the shortest day. This was done in order to ensure that the daylength function is even around zero (which made the analysis less complex). The last ten days of data were moved up to the beginning of the file to provide data for the first ten days. For simplicity, we ignore leap years, so that every year has 365 days.

For DALEC DE we use the same (unsmoothed) climate data as for DALEC EV, but the parameter values and the values of the initial conditions for the carbon pools are based on research by Hill *et. al* [17]. The combination of the climate and parameter data is however realistic. Although we are using periodic drivers in this analysis, in reality the drivers are not periodic. However, we can assume that drivers are periodic with a small perturbation, which effectively corresponds to noise in the system, which would not change the overall structure, so therefore we expect that our results will remain valid. Similarly, if the data showed an upward or downward trend we would expect to find that the fixed point changes every year, effectively becoming a noisy fixed point.

In the original Fortran code for DALEC EV, provided by REFLEX [13], the LAI, which appears in the GPP in the C_i function and the E_0 function, is set to $\max(0.1, C_f/lma)$, which prevents it from going below 0.1. This is done in order to prevent C_f from ever reaching zero, which represents the death of the forest. We remove this artificial constraint and therefore write $LAI = C_f/lma$.

Chapter 4

Tipping Points

4.1 Introduction

We start this chapter with a description of how to create annual maps from two different linear daily maps. This technique will be used later on in Section 4.4 to create annual maps for four of the carbon pools.

In Section 4.3 we find that the C_f pool in the DALEC EV model decouples from the other four carbon pools and conclude that hence the forest dynamics are driven by and depend on C_f and its parameters, p_5 , the combination $p_3(1 - p_2)$ and the p_{11} parameter in the GPP. The long-term behaviour of the foliar carbon pool is a periodic oscillation of period one year. In Section 4.3.1 we first find fixed points for the annual foliar carbon cycle, which then allows us in Section 4.3.2 to trace out the behaviour of this periodic state as a function of one of the parameters. Doing this we discover a tipping point at which sustainable behaviour of the forest gives way to widespread mortality. By varying a second parameter we find a line of tipping points and this shows for which values of the parameters a forest can be expected to exhibit sustainable behaviour and for which

values the forest will certainly die without a chance of coming back to life. An interesting feature is that the parameter values from the data set we have used for the majority of this thesis lie very close to the line of limit points. Out of interest we use a second data set to see if we get a similar result.

The drivers for the DALEC model are highly variant (meaning they do not follow a smooth line) and in Section 4.3.3 we study the effect of smoothing them on our tipping point results.

In Section 4.4 we study the behaviour of the other four carbon pools and find their fixed points. And lastly, before our conclusion in Section 4.6 we show in Section 4.5 that the deciduous version of the DALEC model exhibits the same type of behaviour as the evergreen version.

4.2 Linear Maps

In this section we describe how we can create annual maps out of two different daily linear maps. We will use this technique later on in Section 4.4.

4.2.1 Variation 1

Suppose we have a linear daily iteration of the form:

$$x_{t+1} = ax_t + b_t,$$

where b_t has period 365. Then we can define an annual map which has a time-step of 1 year.

To do this, we note that:

$$\begin{aligned}
 x_1 &= ax_0 + b_0, \\
 x_2 &= ax_1 + b_1 \\
 &= a(ax_0 + b_0) + b_1 \\
 &= a^2x_0 + ab_0 + b_1 \\
 x_3 &= ax_2 + b_2 \\
 &= a(a^2x_0 + ab_0 + b_1) + b_2 \\
 &= a^3x_0 + a^2b_0 + ab_1 + b_2 \\
 &\vdots
 \end{aligned}$$

This leads to:

$$x_t = a^t x_0 + \sum_{i=0}^{t-1} a^i b_{t-i-1}$$

Since the term b_i has period 365, then we can define the annual map:

$$x_{365} = a^{365} x_0 + \sum_{i=0}^{364} a^i b_{364-i} \quad (4.1)$$

We use superscript to denote the year, i.e. $x^{(n)}$ is the value of x in year n . Thus, the annual map is given by:

$$x^{(n+1)} = a^{365} x^{(n)} + \sum_{i=0}^{364} a^i b_{364-i} \quad (4.2)$$

Now, to find a fixed point of the annual map, we require $x^{(n+1)} = x^{(n)} = x_0$ and so we

require that:

$$\begin{aligned} x_0 &= a^{365}x_0 + \sum_{i=0}^{364} a^i b_{364-i} \\ x_0 &= \frac{\sum_{i=0}^{364} a^i b_{364-i}}{1 - a^{365}} \end{aligned}$$

The fixed point is stable if the absolute value of the derivative of the annual map at the fixed point is strictly less than 1 and unstable if it is strictly greater than 1. As b_t is not dependent on x , the derivative for equation (4.2) with respect to $x^{(n)}$ is:

$$\frac{dx^{(n+1)}}{dx^{(n)}} = a^{365}. \quad (4.3)$$

Therefore the fixed point of the annual map (4.2) is stable if $|a^{365}| < 1$ and clearly this is equivalent to $|a| < 1$.

4.2.2 Variation 2

Now consider a linear daily iteration of the form:

$$x_{t+1} = a_t x_t + b_t, \quad t = 0, 1, 2, 3, \dots$$

where a_t and b_t have period 365.

In this case we see that:

$$\begin{aligned} x_1 &= a_0 x_0 + b_0, \\ x_2 &= a_1 x_1 + b_1 \\ &= a_1(a_0 x_0 + b_0) + b_1 \end{aligned}$$

$$\begin{aligned}
 &= a_0 a_1 x_0 + a_1 b_0 + b_1 \\
 x_3 &= a_2 x_2 + b_2 \\
 &= a_2(a_0 a_1 x_0 + a_1 b_0 + b_1) + b_2 \\
 &= a_0 a_1 a_2 x_0 + a_1 a_2 b_0 + a_2 b_1 + b_2 \\
 &\vdots
 \end{aligned}$$

This leads to:

$$x_t = \left(\prod_{i=0}^{t-1} a_i \right) x_0 + \sum_{i=0}^{t-1} \left(b_i \prod_{j=i+1}^{t-1} a_j \right)$$

where we use the convention that the product term has the value 1 if the range is invalid.

Since the terms a_t and b_t have period 365, then we can define the annual map:

$$x^{(n+1)} = \left(\prod_{i=0}^{364} a_i \right) x^{(n)} + \sum_{i=0}^{364} \left(b_i \prod_{j=i+1}^{364} a_j \right) \quad (4.4)$$

where a superscript has been used to denote the year.

To find the fixed points of the annual map, we require that:

$$\begin{aligned}
 x_0 &= \left(\prod_{i=0}^{364} a_i \right) x_0 + \sum_{i=0}^{364} \left(b_i \prod_{j=i+1}^{364} a_j \right) \\
 x_0 &= \frac{\sum_{i=0}^{364} \left(b_i \prod_{j=i+1}^{364} a_j \right)}{1 - \prod_{i=0}^{364} a_i}
 \end{aligned} \quad (4.5)$$

As b_t is not dependent on x , the derivative of the annual map, equation (4.4), with respect to $x^{(n)}$ is $\prod_{i=0}^{364} a_i$ and so the fixed point of the annual map is stable if $\left| \prod_{i=0}^{364} a_i \right| < 1$.

4.3 Behaviour of the Foliar Carbon Pool

4.3.1 Finding Fixed Points

The foliar carbon pool is crucial in determining the amount of carbon the other pools receive. In equations (3.1) - (3.5), this is seen from the fact that the foliar carbon pool, C_f , decouples from the other carbon pools. Therefore the dynamics of the forest is driven by and depends on C_f and on the parameters in this pool, namely p_2 (fraction of GPP respired), p_3 (fraction of NPP allocated to foliage), p_5 (daily turnover fraction of foliage) and p_{11} (nitrogen use efficiency parameter in the ACM). For the purpose of our analysis, we only consider parameters p_2 , p_3 and p_5 as they are the parameters that appear in the C_f equation. Parameters p_2 and p_3 only appear in combination with each other as $p_3(1 - p_2)$ in equation (3.1); we regard them in our analysis as one parameter. Thus we define $p_{23} = p_3(1 - p_2)$. This parameter combination represents the fraction of GPP allocated to the foliar carbon every day. For different parameter regimes there is different behaviour of C_f .

Fig. 4.1 shows two graphs of C_f , resulting from running the model forward over a number of years, for two different values of p_5 . The solid line ($p_5 = 0.0024$) shows the foliar carbon growing to a stable periodic state with a period of one year. During this annual cycle there is more foliar carbon during the summer than during the winter, as photosynthesis is increased in the summer. In contrast, the dashed line ($p_5 = 0.0035$) shows the foliar carbon steadily decreasing on average over a period of 8 to 9 years and then asymptotting to zero, which represents the forest dying. This raises the question: “How does the change in behaviour take place and for which parameter values?”

The transient behaviour of C_f that leads to a stable periodic state, as shown in Fig. 4.1, suggests that the DALEC EV model has a periodic solution with a period of one year for certain values of the parameters. These periodic solutions can be found either by running

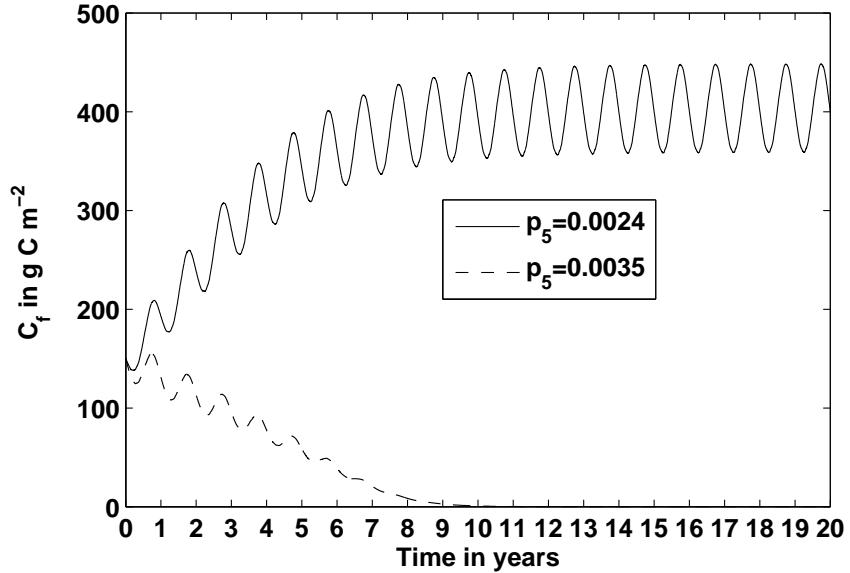


Figure 4.1: The top line shows that for a value of $p_5 = 0.0024$, C_f stabilises on a periodic state. The bottom line shows C_f decreasing and going to zero for $p_5 = 0.0035$, showing several years of decline resulting in mortality. The values of the other parameters are $p_{23} = 0.1392$ and $p_{11} = 7.4$. Day 1 for each year is December 21st, the shortest day of the year at a latitude of 52° .

the DALEC EV model forward in time, as was done in order to produce Fig. 4.1, or by recognising that, for a periodic solution, the amount of carbon on day t in year n is the same as on day t in year $n + 1$:

$$C_f(t) = C_f(t + 365). \quad (4.6)$$

Periodic solutions can therefore be found as fixed points of an annual map, which is obtained by iterating the daily map 365 times.

We cannot analytically find the fixed points for $C_f(t)$, but we can find them numerically, using the Matlab function ‘fzero’, which results in three fixed points for the foliar carbon: $C_f = 0$, $C_f = 86.7858$ and $C_f = 242.2016$ for values $p_2 = 0.52$, $p_3 = 0.29$ and $p_5 = 0.0028$. These are the parameter values from the Loobos forest data set [13].

We cannot analytically prove stability for fixed points $C_f = 86.7858$ and $C_f = 242.2016$.

By using finite differencing to find the derivative of the fixed points we found that $C_f = 86.7858$ is unstable with a derivative of magnitude 1.2962 and $C_f = 242.2016$ is stable with a derivative of magnitude 0.7650.

We can prove stability for $C_f = 0$ though.

Lemma 4.3.1. *The fixed point $C_f = 0$ is linearly stable*

Proof: As mentioned before, proof of stability of a fixed point lies in evaluating the Jacobian at the fixed point.

Let us express the equation for the daily C_f as:

$$C_f(t+1) = F(C_f(t)).$$

and its derivative is:

$$\frac{dC_f(t+1)}{dC_f(t)} = F'(C_f(t)).$$

Then the annual map can be expressed as:

$$C_f(365) = F^{365}(C_f(0))$$

and therefore

$$C_f^{(n+1)} = F^{365}(C_f^{(n)}).$$

The derivative for the annual map is:

$$\frac{dC_f^{(n+1)}}{dC_f^{(n)}} = \prod_{i=365n}^{365(n+1)-1} F'(C_f(i)).$$

The derivative of equation (3.1), the daily dynamical foliar carbon equation, is:

$$F'(C_f(t)) = \frac{dC_f(t+1)}{dC_f(t)} = 1 - p_5 + \frac{dGPP(C_f(t), t)}{dC_f(t)}. \quad (4.7)$$

We will now prove that $\frac{dGPP(C_f(t), t)}{dC_f(t)}|_{C_f(t)=0} = 0$, which will simplify finding the derivative of the annual map at $C_f = 0$. Recall equation (3.12), the equation for $GPP(C_f(t), t)$. For legibility we will leave the $(C_f(t), t)$ out for this proof. Using the chain rule, the derivative of GPP with respect to C_f is:

$$\frac{dGPP}{dC_f} = \frac{\partial GPP}{\partial E_0} \frac{dE_0}{dC_f} + \frac{\partial GPP}{\partial C_i} \frac{dC_i}{dC_f}. \quad (4.8)$$

Let's start with dE_0/dC_f . Recall equation (8.4) from Appendix A, the equation for $E_0(t)$. To simplify matters, we will let $\beta = 1/lma$. The derivative of this equation with respect to C_f is:

$$\frac{dE_0}{dC_f} = \frac{2a_7\beta^2C_f}{\beta^2C_f^2 + a_9} - \frac{2a_7\beta^4C_f^3}{(\beta^2C_f^2 + a_9)^2}.$$

It is easy to see that when $C_f = 0$, both numerators will be zero and the denominators will be non-zero and therefore $dE_0/dC_f = 0$. Also, $\partial GPP/\partial C_i = 0$ for $C_f = 0$. With both $dE_0/dC_f = 0$ and $\partial GPP/\partial C_i = 0$, it follows that $dGPP/dC_f = 0$ when $C_f = 0$. So for $C_f(t) = 0$, $F(C_f(t)) = 0$ and $F'(C_f(t)) = 1 - p_5$ and therefore $dC_f(365)/dC_f(0) = (1 - p_5)^{365}$.

Therefore we end up with a derivative of the annual map of $(1 - p_5)^{365}$ evaluated at $C_f = 0$, which has magnitude smaller than 1, provided that $0 < p_5 < 1$, and, as $p_5 = 0.0028$, this means that the fixed point $C_f = 0$ is stable. Being the fraction of the foliar carbon that goes out into litter, p_5 cannot physically go beyond 1 and so this fixed point will always be stable. \square

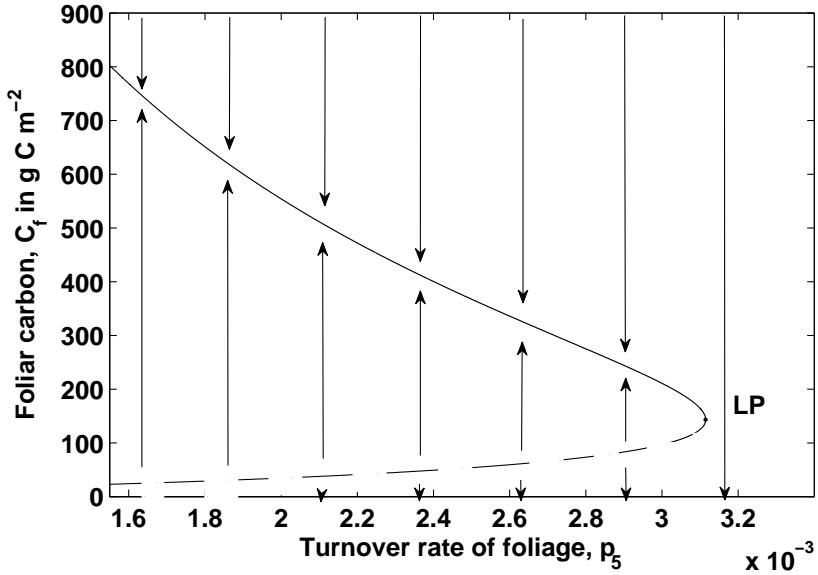


Figure 4.2: *Path of fixed points for the annual map for $p_{23} = 0.1392$. The top branch is stable, the bottom branch is unstable and there is also a stable fixed point branch at $C_f = 0$. On the vertical axis the value of C_f on day 1 of each year (December 21st) is shown.*

4.3.2 Varying the Parameters to Find Tipping Points

The benefits of finding the periodic states by finding fixed points of the annual map are that it is then possible to trace out the behaviour of the periodic state as a function of the parameters, by varying either one or more parameters, and that it is possible to follow both stable and unstable solutions. Although unstable solutions are not seen in simulations, they play a significant role as they allow one to map out regions of qualitatively different behaviour in phase space. Using CL_Matcont_for_maps [9] results in Fig. 4.2, which shows how the periodic state changes when parameter p_5 is varied. It has two branches, representing two paths of fixed points, which are connected by a limit point or fold bifurcation.

This limit point represents a tipping point, where the stability of a solution is suddenly lost as p_5 increases and the system makes a sudden transformation to completely different

behaviour. The top branch of the curve in Fig. 4.2 is stable and the bottom branch is unstable. The role of the unstable solutions in dividing up phase space is illustrated in the following example.

If a simulation is carried out with the initial value of $C_f(0) = C_{f_0}$, at a point anywhere on the graph between the top and bottom branch, such as at $p_5 = 0.0024$ and $C_{f_0} = 140$, then C_f will be attracted to the top branch and will grow to a stable periodic state. If the initial condition is at a point below the bottom branch, for example at $p_5 = 0.0024$ and $C_{f_0} = 20$, it will be repelled from the bottom branch and attracted to the fixed point $C_f = 0$. If p_5 is chosen to the right of the limit point, for example $p_5 = 0.0035$ then the iterates will also be attracted to $C_f = 0$ for all initial values of C_f . For a forest this means that for an initial value of $140 \text{ g } C \text{ m}^{-2}$ of foliar carbon and a daily turnover rate of foliage, $p_5 = 0.0024$, it will grow and find an equilibrium, see Fig. 4.1, but for an initial value of $20 \text{ g } C \text{ m}^{-2}$ of foliar carbon and $p_5 = 0.0024$, it would die over a period of time. If $p_5 = 0.0035$ the forest would eventually die for any initial value of the foliar carbon, see Fig. 4.1. Once the forest has died, it would not be possible for it to re-grow (according to this model), as the line $C_f = 0$ is an attractor.

In Fig. 4.2 the limit point of C_f is shown for a particular value of p_{23} . However, different values of p_{23} would all produce different limit points. Using CL_Matcont_for_maps and varying p_{23} as well as parameter p_5 creates a line of limit or tipping points, as for each value of the parameter p_{23} there is a curve, such as the one in Fig. 4.2.

If p_{23} changes, the curve will move and the limit point will move with it. In this way one can imagine a whole series of new curves with new limit points. In the three-dimensional space (C_f, p_5, p_{23}) , there will be a two-dimensional surface of fixed points and a one-dimensional line of limit points. By finding the limit points of all the possible combinations it becomes possible to define for which values of the parameters a forest has a chance of growing to a periodic state and living, depending on the initial value of C_f , and for which

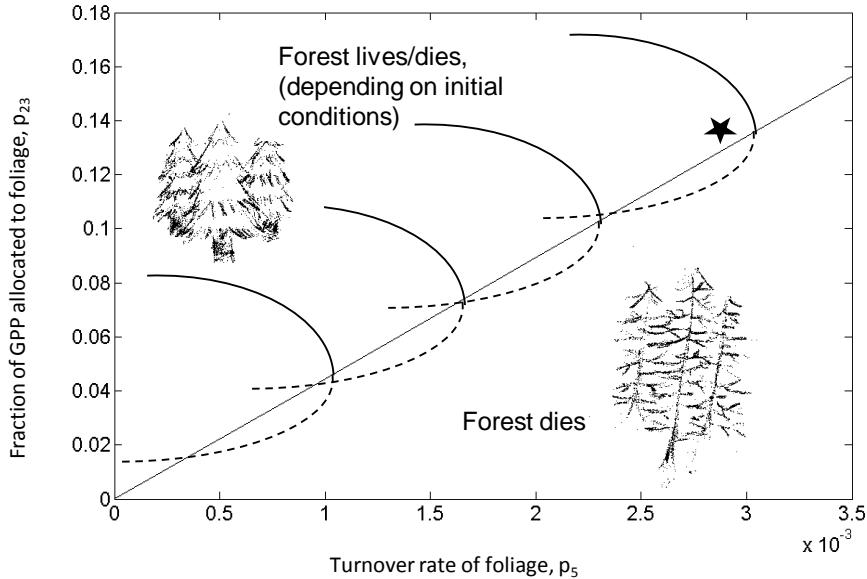


Figure 4.3: *Schematic diagram showing how the limit points (tipping points) divide the parameter plane into two different regions. Note the star near the line of tipping points, which represents the parameter values used in the REFLEX project [13] (drawings of trees by Nina Chuter).*

values the forest is doomed to widespread mortality. Fig. 4.3 shows the line of tipping points along with a schematic series of curves illustrating how the line of tipping points varies.

As C_f is central to the distribution of carbon to the different carbon pools, it is possible to conclude that for values of p_{23} and p_5 , situated on the left hand side of the line of limit points, the evergreen forest either grows to a stable periodic state and lives or it will eventually die, depending on the initial condition for C_f . For values of the parameters on the right hand side of the line of limit points the conditions are such that an evergreen forest will eventually die, whatever the initial value of C_f . This answers the question on how the behaviour of the foliar carbon pool changes for different values of the parameters, p_2 , p_3 and p_5 . It is interesting to note that the parameter values from the REFLEX project result in a point near the tipping point line, see Fig. 4.3.

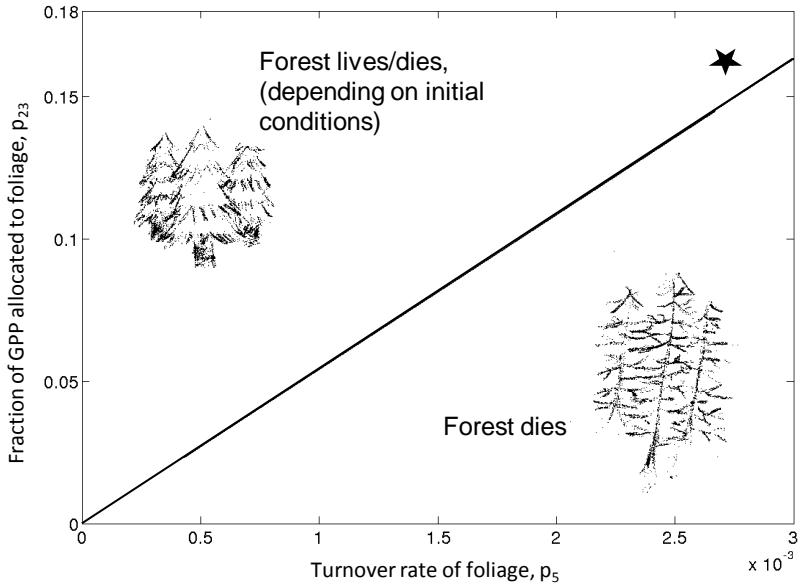


Figure 4.4: *The line of tipping points for the Oregon forest. The star represents the parameter values. For this forest $p_5 = 0.0027$ and $p_3(1 - p_2) = 0.1659$.*

We used climate and parameter data from another, younger, forest to see if the parameter values would also result in a point near the tipping point line. The forest is a young Ponderosa pine site located in Metolius, a Research Natural Area in the eastern Cascades, near Sisters, Oregon, USA [50], which, similarly to the Loobos forest used in the Reflex project, enjoys a temperate climate. Nearly all parameter values for this forest are different, including the a -parameters in the ACM, see Appendix D. The star in Fig. 4.4 is again near the line of tipping points, although perhaps not as close as for the Reflex data.

The DALEC model is used to simulate the carbon cycle of an evergreen forest and, although in our analysis we have assumed that the drivers are periodic on an annual timescale, in reality the atmospheric carbon is increasing year by year, partly due to human activities such as the burning of fossil fuel [5]. The effect of atmospheric carbon is included in the model of photosynthesis, see equation (8.2) in Appendix A. Fig. 4.5 shows the effect that raising the atmospheric carbon, C_a has on the curve in Fig. 4.2. For fixed p_{23} , as C_a increases, relatively speaking, the value of p_5 increases significantly.

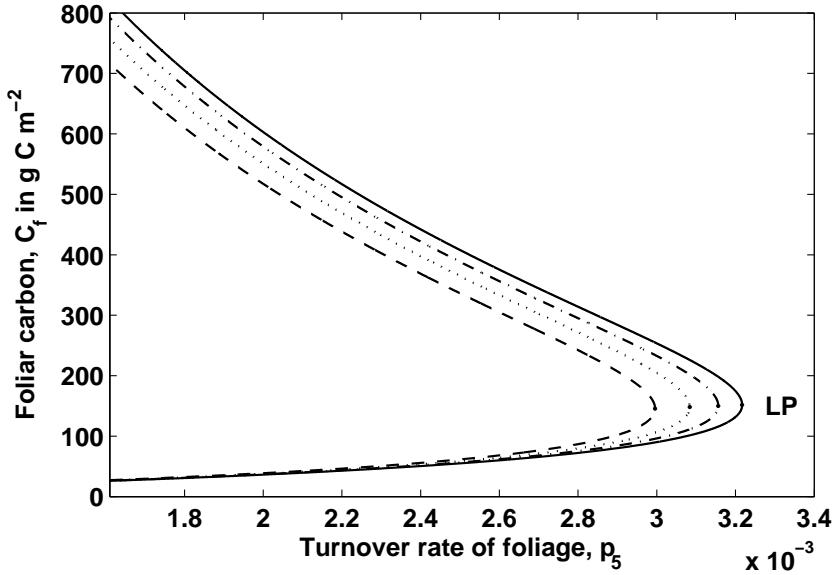


Figure 4.5: Fold bifurcations for $C_a = 380$ (dashed), $C_a = 430$ (dot), $C_a = 480$ (dash-dot) and $C_a = 530$ (solid).

The value of C_f at the limit points also increases slightly. For a given value of p_5 , with more CO₂ used in photosynthesis, the forest grows towards a higher periodic solution and the region on the graph where a forest can live is bigger. Such a response is in agreement with McMahon *et al.* [30], who identify increased atmospheric carbon as a possible factor in increased growth of trees through carbon fertilisation. However, growth is also dependent on other resources. Research by McCarthy *et al.* [29] found that during a six year experiment at a Free Air CO₂ Enrichment (FACE) site in Duke Forest, USA, pines receiving elevated CO₂ had on average about 17 percent more needles than untreated pines, mainly depending on the amount of nitrogen present in the soil.

Interestingly, the increase of mean daily temperature by a constant amount of 2.4° C has no significant effect on the curve. Several studies suggest that this is the approximate amount the temperature will rise by the end of the 21st century even if the atmospheric carbon level is kept at its current value [16]. This lack of effect could be due to the fact that the data used is from a forest in a moderate climate. There are only 23 days in the

data where the temperature during the day is slightly below zero. Also, in DALEC EV no provision is made for when the temperature goes below zero, which is when photosynthesis slows down or stops [3]. In DALEC EV photosynthesis takes place whatever the temperature. One might anticipate more temperature dependence for deciduous forests, as a temperature increase here could result in an increase in the length of the growing season.

4.3.3 Smoothing the Drivers

The drivers for temperature and irradiance used are highly variable, see Fig. 4.6. We now consider the effect that smoothing the drivers has on the results as, for mathematical analysis purposes, it is more desirable to work with simple functions rather than raw data.

Atmospheric CO₂ Concentration, C_a

This driver has only three different values over three years (377, 380 and 383 ppm respectively). We have chosen the average of these values: $C_a = 380$ ppm. This does not make a lot of difference to the foliar carbon graph, see Fig. 4.7. Of course, this is completely unrealistic, as we know that the amount of CO₂ in the atmosphere is rising steadily (with an annual cycle). For the purpose of our analysis though, it will suffice, as the percentage change per year is relatively small (less than 1%).

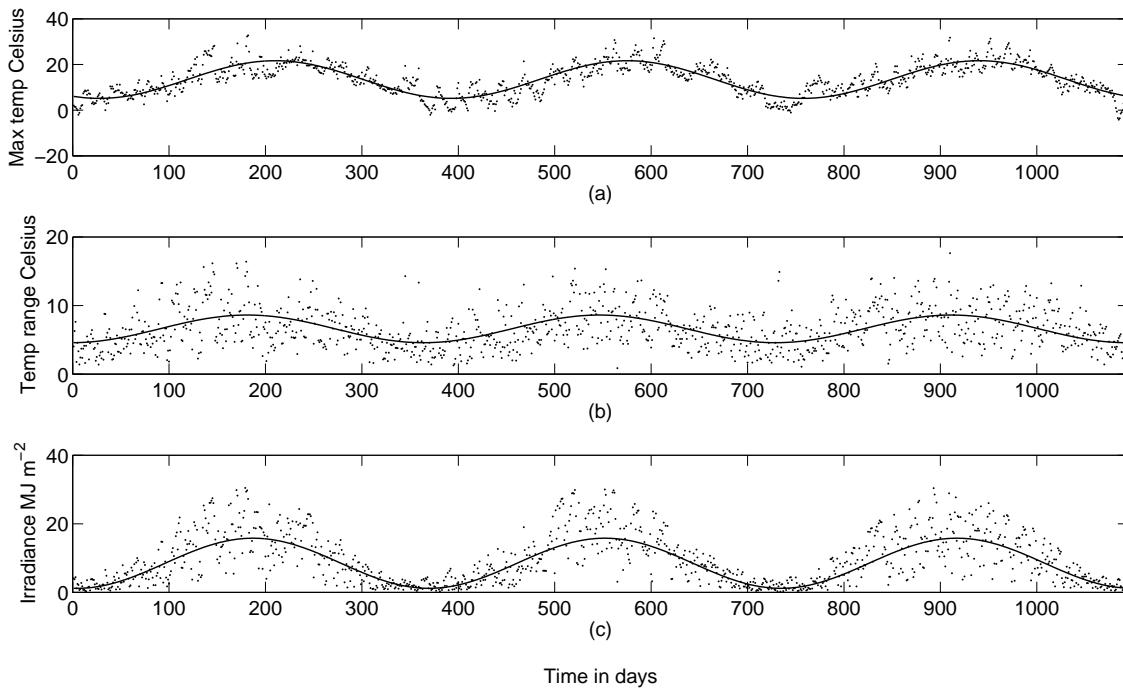


Figure 4.6: Smoothed and original drivers over a three year period, data from Loobos forest in the Netherlands [13]: (a) the maximum temperature, (b) the temperature range and (c) the irradiance. Note that the maximum temperature rarely drops below zero.

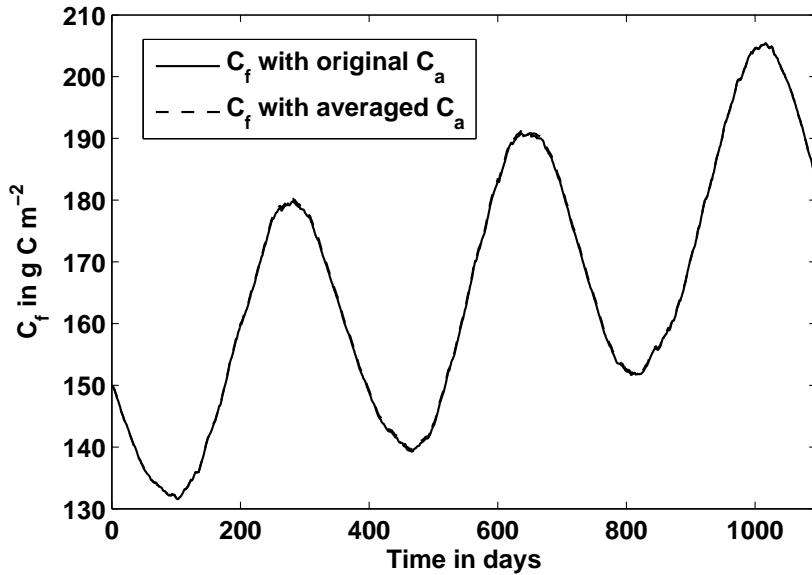


Figure 4.7: Plot of the foliar carbon over three years, using all the highly variable drivers, comparing C_a varied over three years with C_a set to constant 380.

The Maximum Temperature and Temperature Range

The maximum temperature and the temperature range, which are both expressed in °C, can be smoothed using a Fourier series (created with a fast Fourier transform (FFT)):

$$\begin{aligned}
 T_{max} &= 13.4070 - 7.3239 \cos(2\pi t/365) - 3.7831 \sin(2\pi t/365) \text{ } ^\circ \text{C} \\
 &= 13.4070 - 8.2433 \cos(2\pi(t - 27.698)/365) \text{ } ^\circ \text{C} \\
 T_{range} &= 6.5933 - 2.0156 \cos(2\pi t/365) + 0.04256 \sin(2\pi t/365) \text{ } ^\circ \text{C} \\
 &= 6.5933 - 2.0161 \cos(2\pi(t + 1.226)/365) \text{ } ^\circ \text{C}
 \end{aligned}$$

It was decided to ignore the remaining cosine and sine terms in the Fourier series, as adding more terms did not make a significant difference to the smoothed shape of the graphs. Fig. 4.6 shows the highly variable maximum temperature and temperature range.

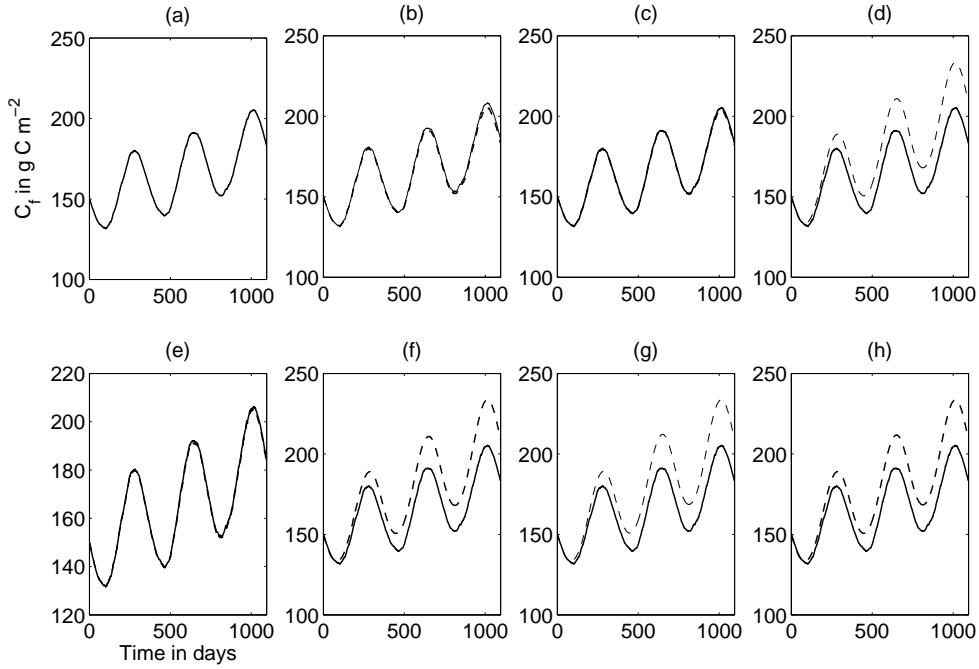


Figure 4.8: (a) Non-smoothed C_f , (b) C_f with smoothed temperature range, (c) C_f with smoothed maximum temperature, (d) C_f with smoothed irradiance, (e) C_f with smoothed temperature range and maximum temperature, (f) C_f with smoothed temperature range and irradiance, (g) C_f with smoothed maximum temperature and irradiance, (h) C_f with smoothed irradiance, maximum temperature and temperature range. All plots show a comparison with the original foliar carbon constructed from all the highly variable drivers (continuous line) with the foliar carbon constructed from the smoothed drivers (dashed line).

Irradiance

The irradiance driver is highly variable. The data can be smoothed as a Fourier series, using FFT, as shown in Fig. 4.6, which is as follows:

$$\begin{aligned} I(t) &= 9.5598 - 8.3958 \cos(2\pi t/365) - 0.2060 \sin(2\pi t/365) \text{ } m^{-2} \text{ } d^{-1} \\ &= 9.5598 - 8.3983 \cos(2\pi(t - 1.4251)/365) \text{ } m^{-2} \text{ } d^{-1} \end{aligned}$$

Results of Smoothing on C_f

Now that we have smoothed irradiance, maximum temperature and temperature range, it is interesting to see the effect of the smoothing on C_f . Fig. 4.8 shows graphs of the foliar carbon, which were created using different combinations of smoothed and highly variable drivers in the GPP. As can be seen in Fig. 4.8 (b), (c) and (e), smoothing the maximum temperature and the temperature range has little impact on the predictions of the model. However this is not true for the irradiance, see Fig. 4.8 (d), (f), (g) and (h).

4.3.4 Improving Smoothing of the Irradiance

To see if we can improve on the agreement between the smoothed model and the original model we approximate the irradiance as a Fourier series as:

$$I(t) = a + b \cos(2\pi t/365) + c \sin(2\pi t/365).$$

and use 4DVAR, see Section 2.4.2, to find values of a , b , and c , so that the trajectory of the foliar carbon matches the corresponding trajectory based on the highly variable drivers as closely as possible.

The cost function for this particular problem is:

$$J(C_{ft}) = \frac{1}{2} \sum_{t=0}^{1095} (C_{ft} - \mathbf{M}(C_{ft}))^2,$$

subject to the constraint:

$$C_{ft+1} = \mathbf{M}(C_{ft}), \quad t = 0, \dots, 1095.$$

\mathbf{M} is the model equation, in this case the equation for the foliar carbon pool, equation

(1). The constraint is strong, so the model is assumed to be perfect.

The auxiliary function, G , as mentioned in Section 2.4.2, looks as follows:

$$G(C_{f_t}, a, b, c, \lambda) = J(C_{f_t}) + \sum_{i=0}^{1095} \lambda_{t+1} \cdot (C_{f_{t+1}} - \mathbf{M}(C_{f_t})),$$

It is required to minimise the cost function with respect to the Fourier series coefficients a , b and c . As described in Section 2.4.2, to find the derivative of the auxiliary function G , with respect to the parameters, we need to apply:

$$\nabla_{par} G = - \sum_{t=0}^{1095} \lambda_{t+1}^T \mathbf{M}_{par}(C_{f_t}, par),$$

where par is the vector of Fourier coefficients a , b and c .

So we need to calculate λ_i in order to find the derivative.

Data Assimilation for the Irradiance

Matlab has a function, called ‘fminunc’, a minimisation routine for an unconstrained problem, which was used as the minimisation algorithm in the data assimilation programs which were written to find the derivative. ‘fminunc’ can be used because the functional G is now an unconstrained problem. However, ‘fminunc’ tends to have problems finding the global minimum, according to the Matlab documentation.

A plot of the cost function, Fig. 4.9, showed that there is a number of local minima. A contour plot, Fig. 4.10, shows these minima in a different way.

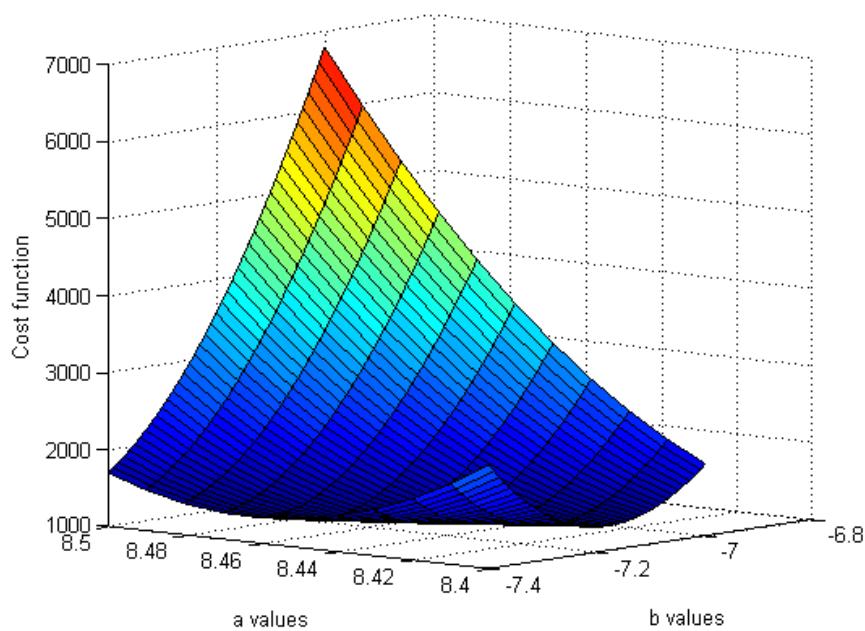


Figure 4.9: Plot of the cost function, using the irradiance as a Fourier series.

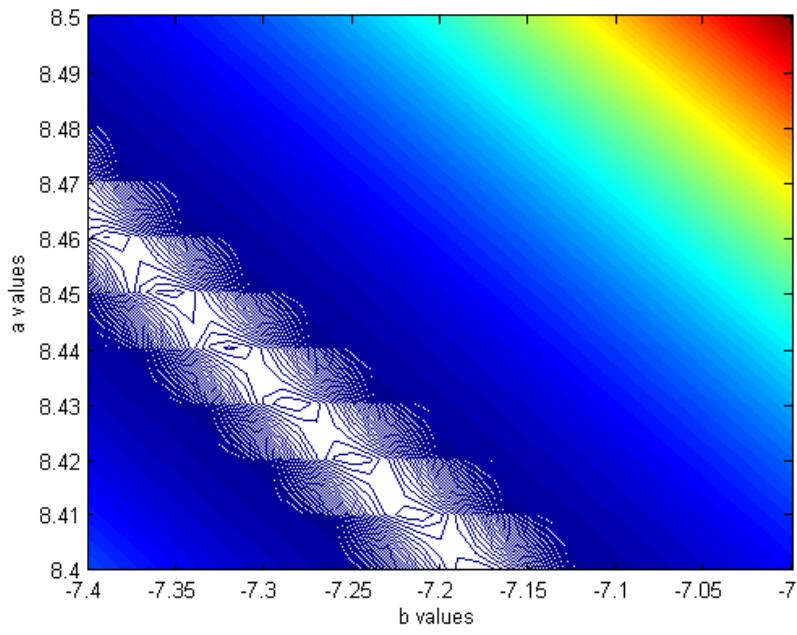


Figure 4.10: Contour plot for irradiance expressed as a Fourier series, showing a number of local minima.

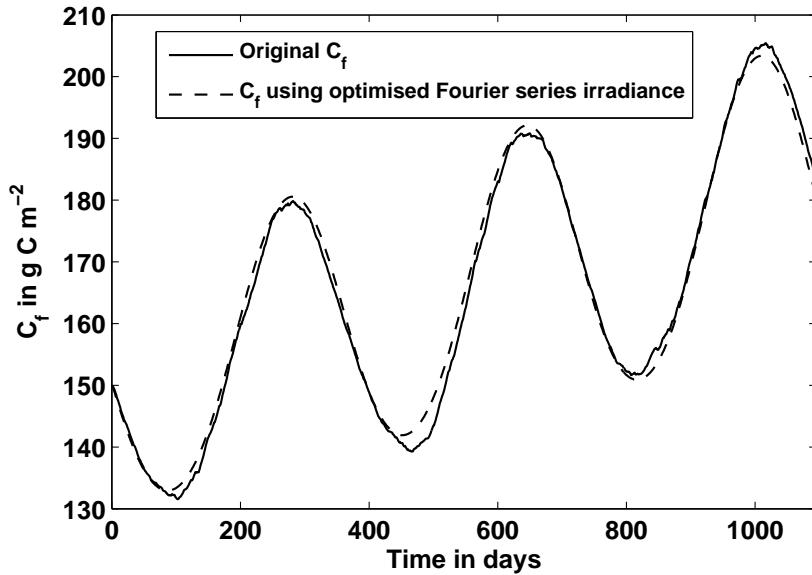


Figure 4.11: The result of data assimilation using the optimal Fourier series irradiance.

After much trial and error, values for the Fourier series were found:

$$I(t) = 8.4774 - 7.2819 \cos(2\pi t/365) - 0.5789 \sin(2\pi t/365)$$

$$I(t) = 8.4774 - 7.3 \cos(2\pi(t-2)/365)$$

The value of the cost function is 752.324. Although we would like the cost function to be as close to zero as possible, due to the fact that there is more than one local minimum and because of the limitations in using ‘fminunc’, we accept the result based on Fig. 4.11, which shows the graph of C_f plotted over three years as original and using the optimal Fourier series with values as in (4.9).

Using Piecewise Quadratic Approximation for the Irradiance

As well as describing the irradiance by a Fourier Series, we attempted to describe it by using a quadratic polynomial, created for one year periodically extended, (by mapping it onto itself twice). The reason for this is that the Fourier series does not reach the higher peaks shown in the highly variable plot of the irradiance (Fig. 4.6). A piecewise polynomial will address this problem.

The quadratic function for the irradiance, with the coefficients a , the shift α and c as parameters can be expressed as:

$$I(t) = a(t + \alpha)^2 + c.$$

The time shift α was included in order to account for the fact that the minimum irradiance does not occur at $t = 0$. Initial values were found through the basic fitting tool in Matlab, which fits a polynomial to the data in a least squares sense. Thus we found an initial quadratic equation for the irradiance:

$$I(t) = 0.00059(t + 1.9)^2 + 2.8.$$

Fig. 4.14 shows that this is not a very good fit for the C_f function. So therefore data assimilation was applied to find the optimum values for the coefficients a , c and α .

A plot of the cost function and a contour plot (see Figs 4.12 and 4.13) show there is a minimum.

Again, after much trial and error, using the same technique as for the Fourier series in the last section, to find the appropriate values for the coefficients, the resulting function

for the irradiance was found:

$$I(t) = 0.00059(t - 5)^2 + 2.05.$$

The graph of the foliar carbon, produced for these values of a , c and α , for the quadratic approximation of the radiance is shown in Fig. 4.14. We can see that for C_f , this quadratic approximation of the irradiance is not much of an improvement on the Fourier series approximation in Fig. 4.11, even though now the higher peaks in Fig. 4.6 are included. We have chosen the Fourier series approximation in the rest of this thesis, as it is easier to use.

Fig. 4.15 shows a comparison between the irradiance as highly variable data, smoothed by the Fourier series approximation and smoothed by the quadratic approximation.

Although we have only used the foliar carbon equation to find the smoothed drivers, Fig. 4.16 shows the error between the highly variable drivers and smoothed drivers for the other carbon pools over three years. We can see that the error slowly grows over the three year period for the foliar pool, the root pool, the soil and organic matter pool and the woody pool, but is on average not too far away from zero. The error for the litter pool does not seem to grow over the 3 year period, the error for the woody pool is the largest and the error for the soil organic matter pool is smallest. The error for the foliar and root pools seem to be similar in value.

4.3.5 Results of Smoothing on the Line of Tipping Points

In Fig. 4.17 the graph for C_f is obtained by running the model forward over three years and drawn with the original drivers and the smoothed drivers. This shows that the effect of the smoothing on C_f is minimal and certainly does not change the qualitative behaviour.

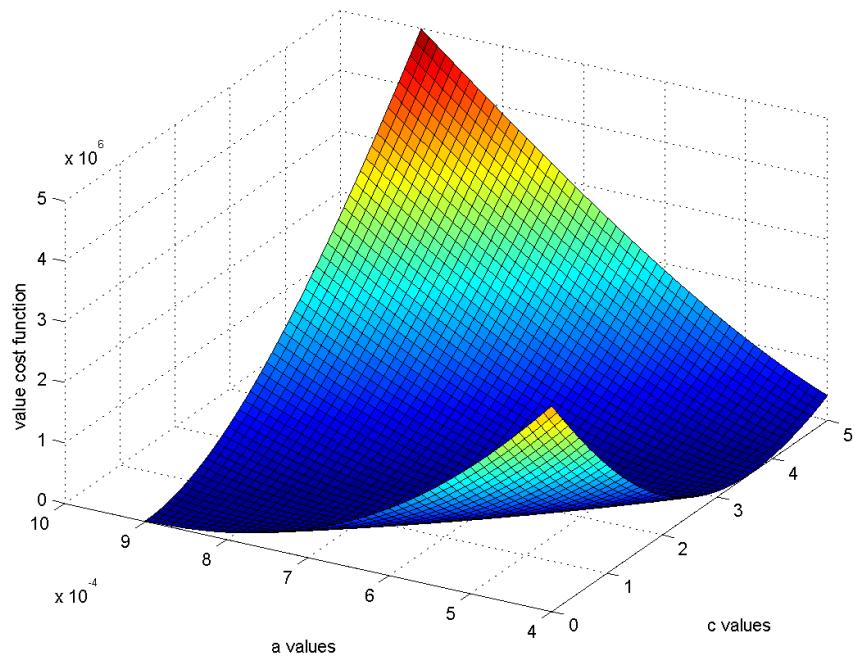


Figure 4.12: Plot of the cost function.

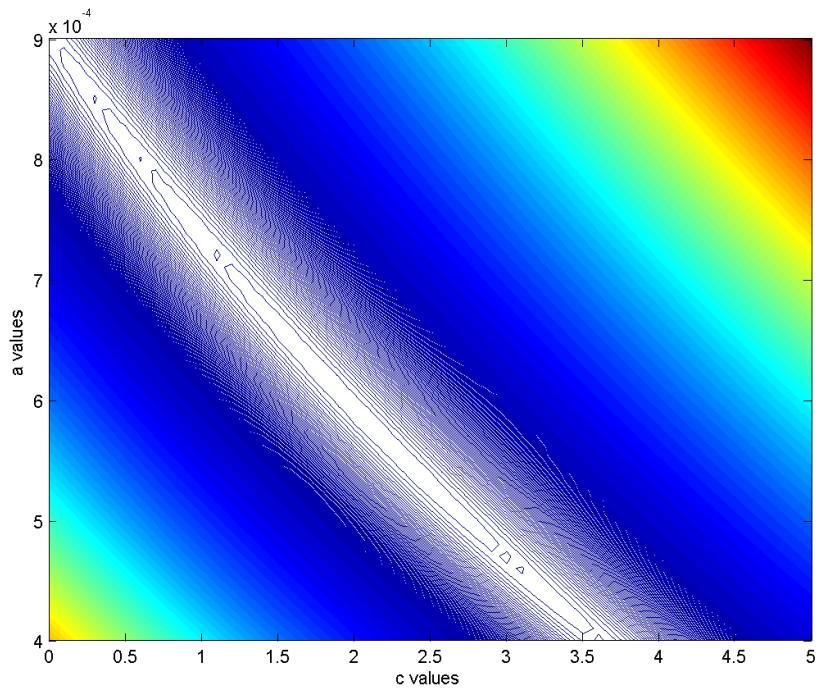


Figure 4.13: Contour plot of the cost function, showing a minimum.

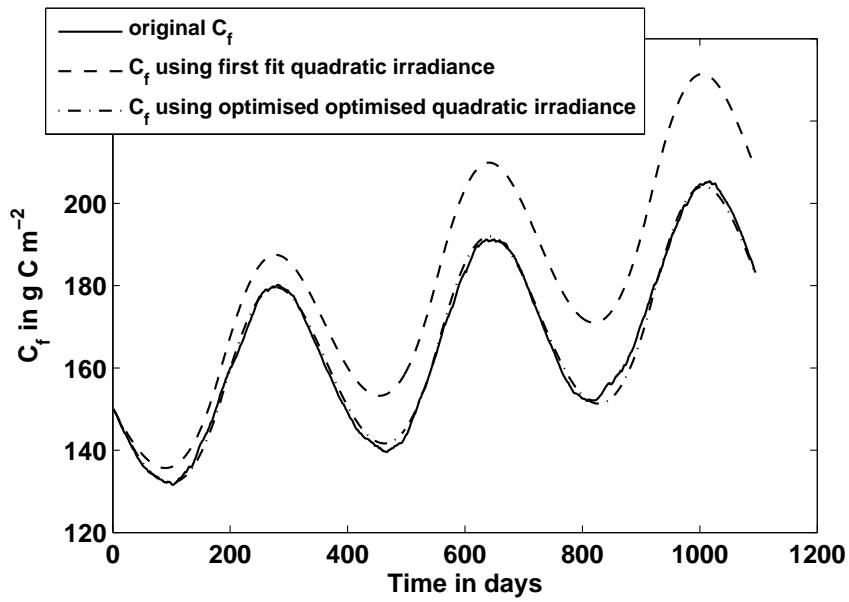


Figure 4.14: *The result of data assimilation using the optimal quadratic irradiance.*

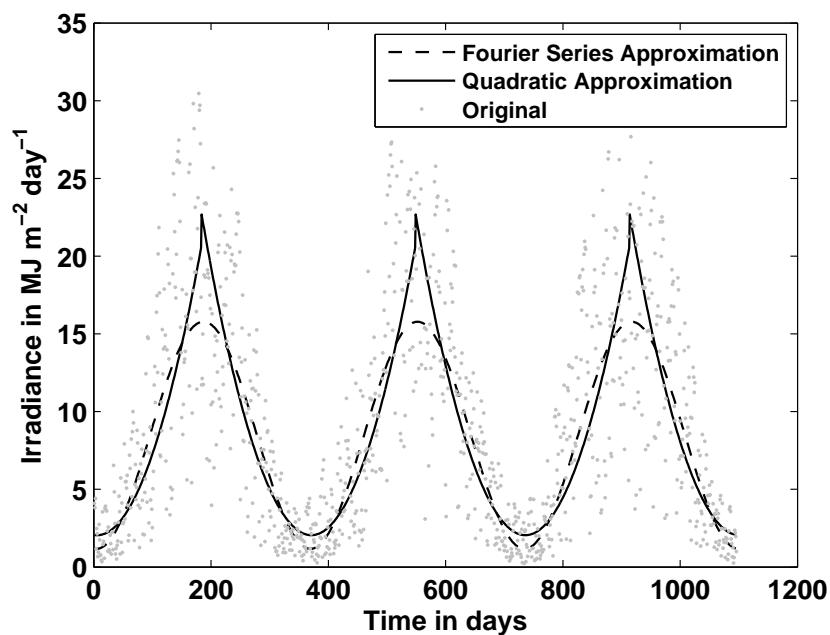


Figure 4.15: *The irradiance smoothed by a Fourier series and a quadratic function.*

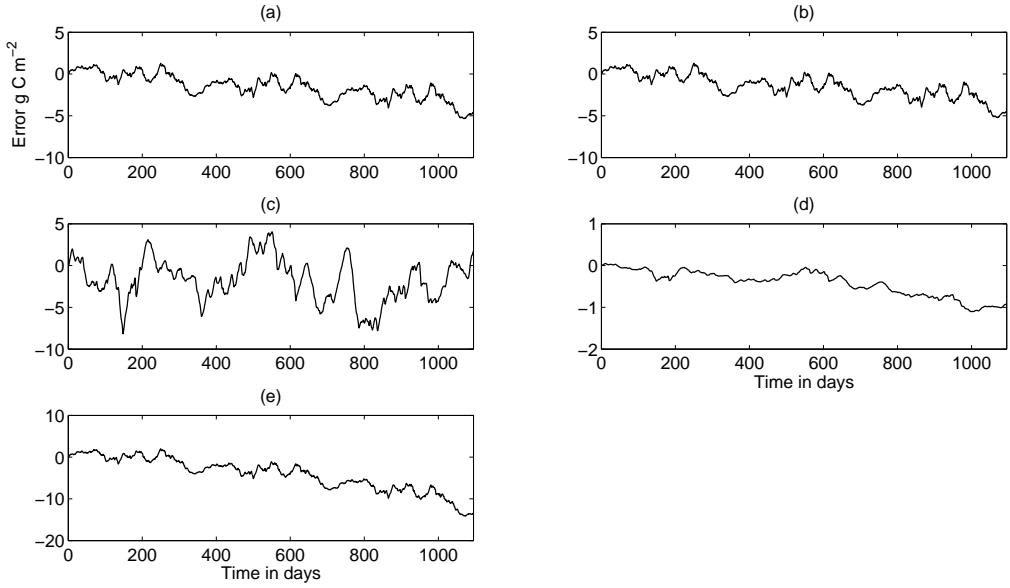


Figure 4.16: The plots shown above show the error between using the original drivers and the smoothed drivers (only the Fourier series for the irradiance was used) for (a) the foliar carbon pool, (b) the root carbon pool, (c) the litter carbon pool, (d) the soil and organic matter carbon pool and (e) the wood pool. All the errors are in g C m^{-2} , so therefore the error for the wood pool (which is a large pool) is very small and not proportionally larger than the rest.

The bifurcation graph and the line of limit points graph produced using the smoothed drivers also show qualitatively very similar behaviour, see Figs 4.18 and 4.19. The only difference is a small change in the path of fixed points for the foliar carbon and in the position of the limit point. The use of smoothed drivers greatly simplifies the analysis of the model as expressing data in terms of functions means it is not necessary to use separate data files. It can also simplify data collection as only three parameters are required, namely the mean, amplitude and phase.

According to Medvige *et al.* [32], natural variability in drivers is important and ecosystems respond to it. His research shows that by taking away high-frequency variability decadal Net Ecosystem Production (NEP), GPP and total respiration is significantly enhanced. However, the ACM only takes into account daily measures of climate data and therefore

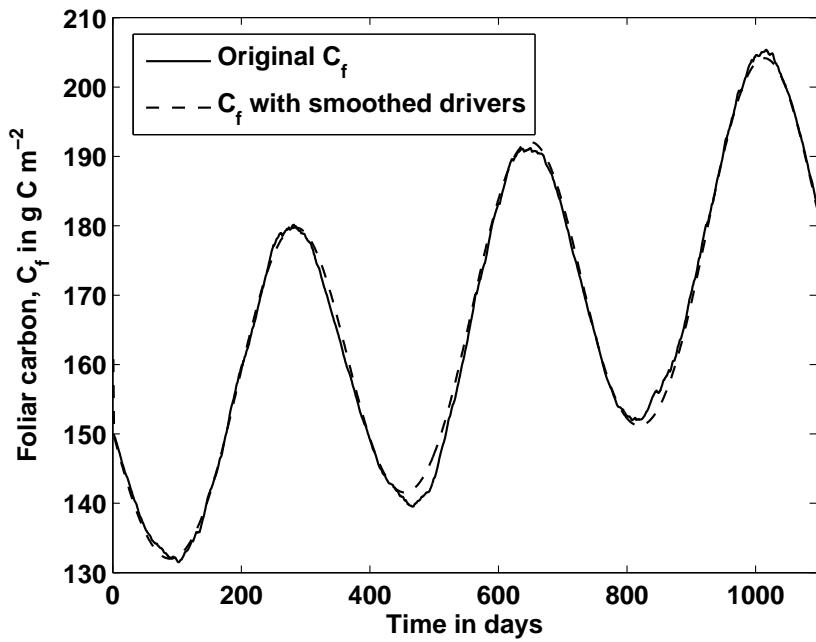


Figure 4.17: C_f over three years shown using original drivers and smoothed drivers. Although the fit is very good for three years, the curves would grow a little further apart if shown on a longer timescale.

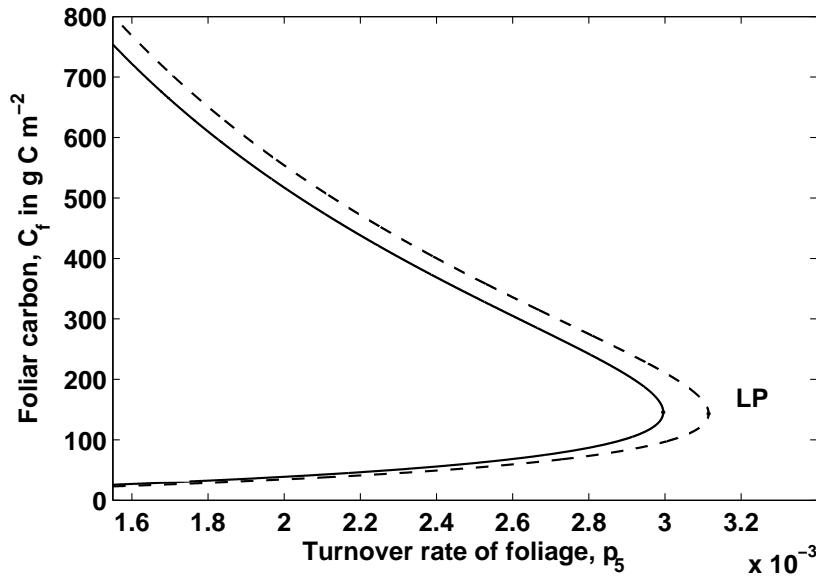


Figure 4.18: Results of the effect of using smoothed drivers on the tipping point structure. The solid curve is the result of using smoothed drivers, the dashed curve is the result of using the original drivers.

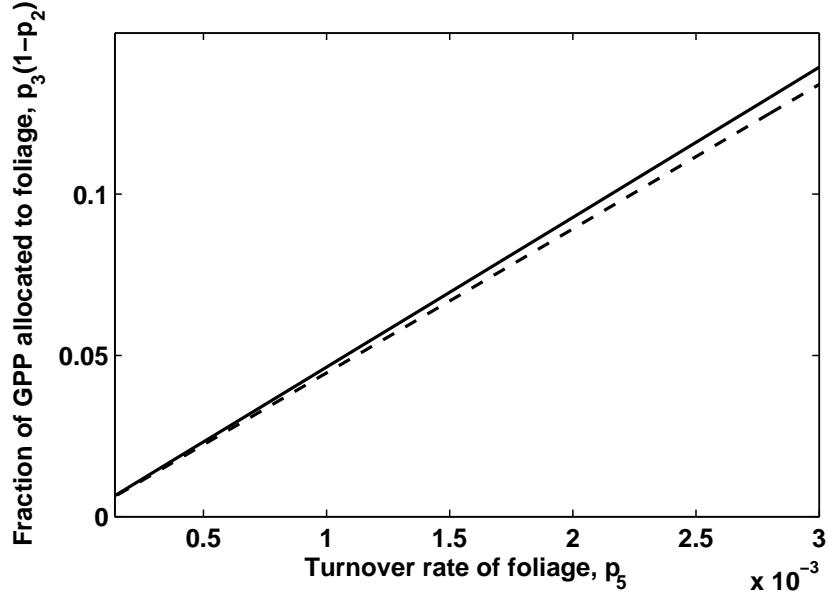


Figure 4.19: *Results of the effect of smoothed drivers for line of tipping points. The continuous line of limit points is the result of using smoothed drivers, the dashed line is the result of using highly variable drivers.*

the high frequency variability has already been taken out.

4.4 Behaviour of the Other Carbon Pools

For each of the fixed points of the annual map for the foliar carbon pool we can find corresponding fixed points for the other four carbon pools. For the fixed point $C_f = 0$ it can be shown that all other pools also have a fixed point at zero. The roots carbon pool, C_r , and the litter carbon pool, C_{lit} , reach their fixed points relatively quickly, on the same time scale as the foliar carbon pool, see Fig. 4.20. However, the soil and organic matter carbon pool, C_{som} , and the woody carbon pool, C_w take a very long time (thousands of years) to reach their fixed points, due to the very small values of the parameters p_6 and p_9 . The wood pool, C_w , grows faster than the soil and organic matter pool, as is shown in Fig. 4.21. Due to the very slow growth of C_{som} , we can regard the initial value of C_{som} as a pseudo-fixed point on a decadal scale.

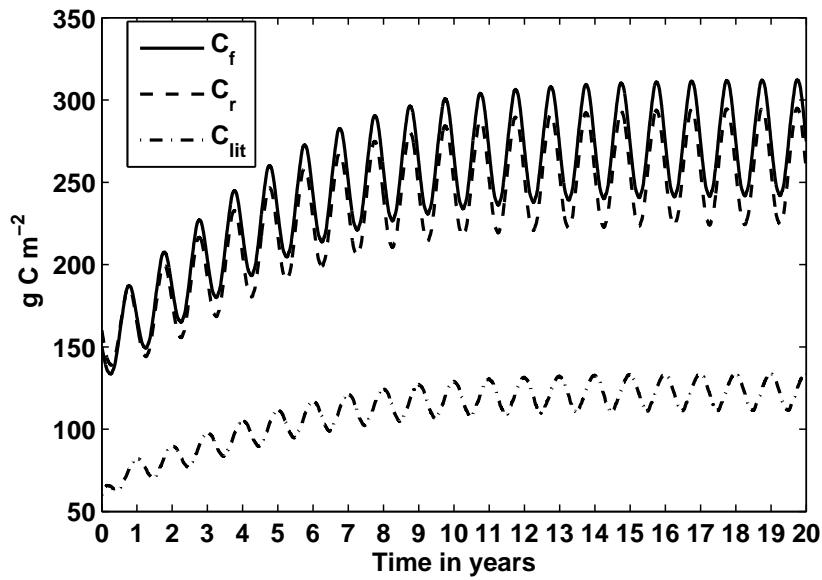


Figure 4.20: Behaviour of C_f , C_r and C_{lit} over 20 years; all three of these carbon pools reach a periodic state within this time.

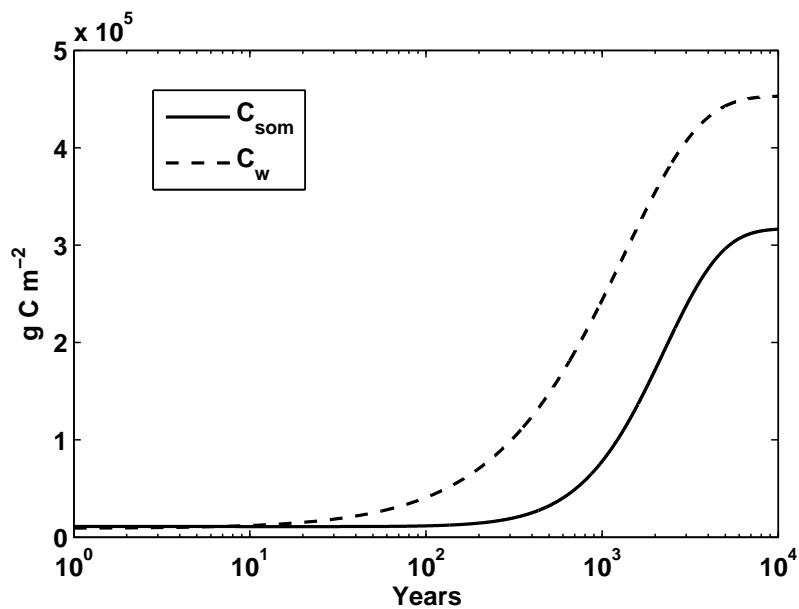


Figure 4.21: Only one point per year plotted, removing the periodic behaviour: C_{som} and C_w do not reach their fixed points until after 10,000 years (note the log scale).

4.4.1 Fixed Points for Root Carbon, C_r

To find fixed points for the root carbon pool, we apply the technique in Section 4.2.1 to find an annual map. However, the C_r pool is dependent on $GPP(C_f(t))$ and therefore its fixed points are dependent on the fixed points for C_f .

The carbon pool for roots is propagated forward in time using the equation:

$$C_r(t+1) = (1 - p_7)C_r(t) + p_4(1 - p_2)(1 - p_3)GPP(C_f(t), t).$$

The annual map for C_r :

$$\begin{aligned} C_r^{(n+1)} &= (1 - p_7)^{365}C_r^{(n)} + p_4(1 - p_2)(1 - p_3) \sum_{i=365n}^{365(n+1)-1} (1 - p_7)^i \\ &\quad \times GPP(C_f(365(n+1) - 1 - i), 365(n+1) - 1 - i). \end{aligned} \quad (4.9)$$

The equation to find fixed points for C_r , assuming that C_f is now periodic, is:

$$C_{r0} = \frac{p_4(1 - p_2)(1 - p_3) \sum_{i=0}^{364} (1 - p_7)^i GPP(C_f(364 - i), 364 - i)}{1 - (1 - p_7)^{365}}.$$

For $C_f = 0$, the fixed point for the root carbon, C_r , is also zero, as $GPP(0, t) = 0$. The other fixed points are $C_r = 225.7 \text{ g C m}^{-2}$ and $C_r = 80.9 \text{ g C m}^{-2}$, for $p_2 = 0.52$, $p_3 = 0.29$, $p_4 = 0.41$ and $p_7 = 0.003$.

The derivative for the annual map of C_r with respect to $C_r^{(n)}$ is $(1 - p_7)^{365}$ and therefore any fixed point of C_r is stable if $|1 - p_7| < 1$, which is equivalent to $0 < p_7 < 2$. Physically, for the same reasons as for the foliar carbon, p_7 cannot be any larger than 1. So therefore, although C_f has two stable fixed points and one unstable fixed point, the fixed points for C_r are all stable. This implies that the unstable fixed point for C_f is a saddle point in the

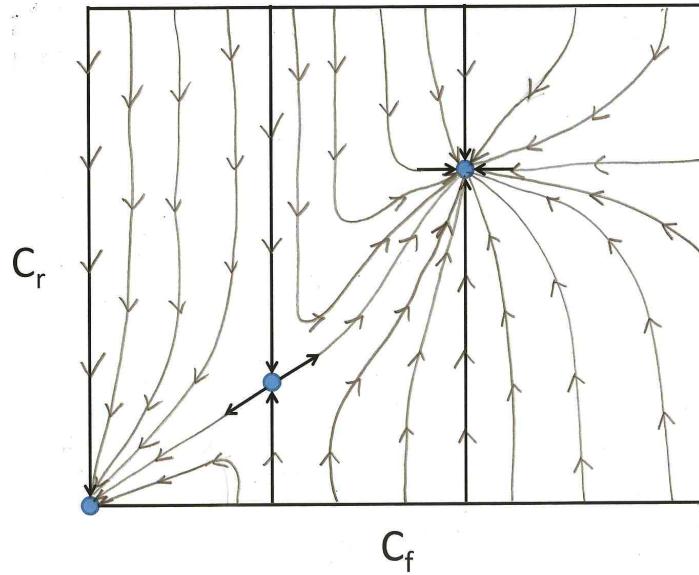


Figure 4.22: Sketch of the phase plane for the fixed points (C_f, C_r) , showing two stable fixed points and one saddle fixed point.

two-dimensional (C_f, C_r) plane. The phase plane in Fig. 4.22 illustrates this fact. The stable manifold acts as the boundary of the two basins of attraction for the two stable fixed points as also shown in Fig. 4.22.

4.4.2 Fixed Points for Woody Carbon, C_w

The fixed points for the woody carbon pool are found in the same manner as the fixed points for the root carbon pool. As the C_w pool is dependent on $GPP(C_f(t), t)$, its fixed points are dependent on the fixed points of C_f .

The carbon pool for wood is propagated forward in time using the equation:

$$C_w(t+1) = (1 - p_6)C_w(t) + (1 - p_4)(1 - p_2)(1 - p_3)GPP(C_f(t), t).$$

The annual map for C_w is similar to that of C_r :

$$\begin{aligned} C_w^{(n+1)} &= (1 - p_6)^{365} C_w^{(n)} + (1 - p_4)(1 - p_2)(1 - p_3) \\ &\quad \times \sum_{i=365n}^{365(n+1)-1} (1 - p_6)^i GPP(C_f(365(n+1) - 1 - i), 365(n+1) - 1 - i). \end{aligned} \quad (4.10)$$

The equation to find a fixed point for the annual map of C_w , assuming that C_f is now periodic, is:

$$C_{w0} = \frac{(1 - p_4)(1 - p_2)(1 - p_3) \sum_{i=0}^{364} (1 - p_6)^i GPP(C_f(364 - i), 364 - i)}{1 - (1 - p_6)^{365}},$$

For $C_f = 0$, the fixed point for the woody carbon, C_w , is zero, as $GPP(0, t) = 0$. The other fixed points are $C_w = 474526.2 \text{ g C m}^{-2}$ and $C_w = 168443.4 \text{ g C m}^{-2}$, for $p_2 = 0.52$, $p_3 = 0.29$, $p_4 = 0.41$ and $p_6 = 0.00000206$, see Fig. 4.23.

The derivative for the annual map of C_w is $(1 - p_6)^{365}$ for any fixed point. Therefore C_w is stable if $|1 - p_6| < 1$, which is equivalent to $0 < p_6 < 2$. Physically, for the same reasons as for the foliar carbon, p_6 cannot be any larger than 1. So all three fixed points of C_w are stable. As the fixed points of C_w depend on the fixed points of C_f we will take note of the fact that in the two-dimensional (C_f, C_w) plane, the unstable fixed point for C_f is a saddle point.

4.4.3 Fixed Points for Litter Carbon C_{lit}

Finding the fixed points for this pool requires a slightly different solution, as the coefficient of C_{lit} is time dependent. We now use the technique in Section 4.2.2.

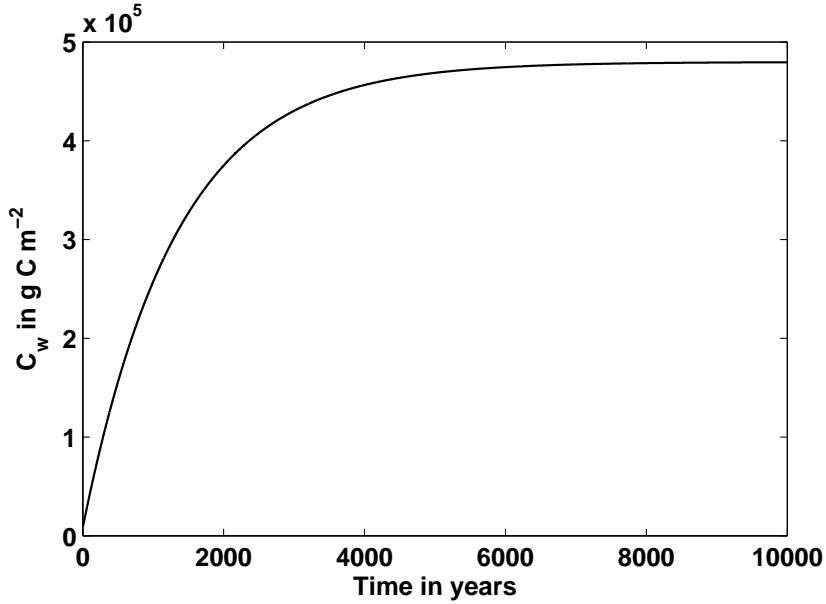


Figure 4.23: C_w crawling to a fixed point over a very long period of time. This is because p_6 is very small.

The carbon pool for litter is propagated forward in time using this equation:

$$C_{lit}(t+1) = (1 - (p_1 + p_8)T(t))C_{lit}(t) + p_5C_f(t) + p_7C_r(t).$$

According to equation (4.4) the annual map for C_{lit} can be written as:

$$C_{lit}^{(n+1)} = \prod_{i=0}^{364} (1 - (p_1 + p_8)T(i)) C_{lit}^{(n)} + \sum_{i=365n}^{365(n+1)-1} [(p_5C_f(i) + p_7C_r(i)) \quad (4.11)]$$

$$\times \prod_{j=i+1}^{365(n+1)-1} (1 - (p_1 + p_8)T(j)). \quad (4.12)$$

According to equation (4.5) the equation to find a fixed point for the annual map of C_{lit} , assuming that C_f and C_r are now periodic, is:

$$C_{lit_0} = \frac{\sum_{i=0}^{364} [(p_5C_f(i) + p_7C_r(i)) \prod_{j=i+1}^{364} (1 - (p_1 + p_8)T(j))]}{1 - \prod_{i=0}^{364} (1 - (p_1 + p_8)T(i))}.$$

For $C_f = 0$ the fixed point for C_{lit} is also zero, as $C_r = 0$ for $C_f = 0$ and $\prod_{i=0}^{364} (1 - (p_1 + p_8))T(i) < 1$ for values of $p_1 = 0.00000441$ and $p_8 = 0.02$. The other fixed points are $C_{lit} = 80.3 \text{ g C m}^{-2}$ and $C_{lit} = 28.8 \text{ g C m}^{-2}$ for $p_5 = 0.0028$, $p_7 = 0.003$ and $T(t) = 0.5 \exp(p_{10}T_m(t))$. $T_m(t)$ is the mean daily temperature.

From equation (4.4) it follows that, for all fixed points of C_{lit} , the derivative is $\prod_{i=0}^{364} (1 - (p_1 + p_8))T(i)$. Stability depends on the value of this derivative: $\prod_{i=0}^{364} (1 - (p_1 + p_8))T(i) = 3.5578 \times 10^{-4}$. As this is smaller than 1, this means that all fixed points are stable. As before, for C_r and C_w , the fixed points of C_{lit} are dependent on those for C_f and in this case, on C_r .

4.4.4 Fixed Points for Soil and Organic Matter Carbon, C_{som}

The equation for the soil and organic matter carbon pool is of a similar construction to the litter carbon pool equation, again using the technique in Section 4.2.2:

$$C_{som}(t+1) = (1 - p_9T(t))C_{som}(t) + p_6C_w(t) + p_1T(t)C_{lit}(t).$$

The annual map for C_{som} can be written as, :

$$\begin{aligned} C_{som}^{(n+1)} &= \prod_{i=0}^{364} (1 - p_9T(i)) C_{som}^{(n)} \\ &+ \sum_{i=365n}^{365(n+1)-1} \left[(p_6C_w(i) + p_1T(i)C_{lit}(i)) \prod_{j=i+1}^{365(n+1)-1} (1 - p_9T(j)) \right]. \end{aligned} \quad (4.13)$$

The equation to find a fixed point for the annual map of C_{som} , assuming that C_w and C_{lit} are now periodic, is:

$$C_{som_0} = \frac{\sum_{i=0}^{364} \left[(p_6C_w(i) + p_1T(i)C_{lit}(i)) \prod_{j=i+1}^{364} (1 - p_9T(j)) \right]}{1 - \prod_{i=0}^{364} (1 - p_9T(i))},$$

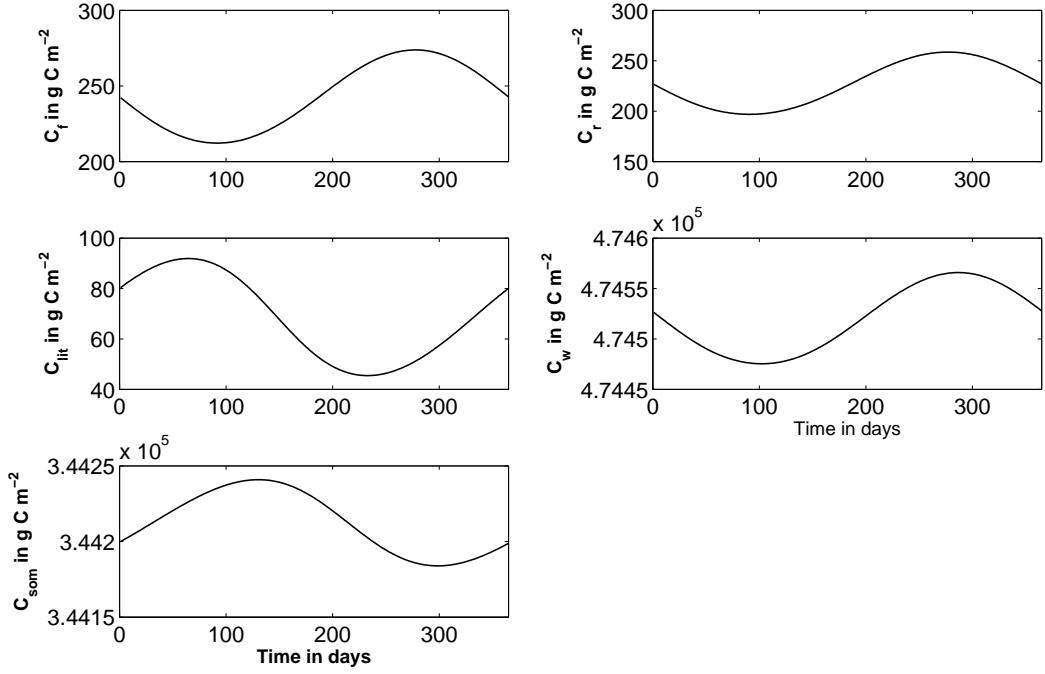


Figure 4.24: Pools showing one period of their non-trivial stable fixed point.

For $C_f = 0$, the fixed point for C_{som} is also zero, as $C_w = 0$ and $C_{lit} = 0$ for $C_f = 0$ and $\prod_{i=0}^{364} (1 - p_9 T(i)) < 1$ for a value of $p_9 = 0.00000265$. The other fixed points are $C_{som} = 344200 \text{ g C m}^{-2}$ and $C_{som} = 122190 \text{ g C m}^{-2}$ for $p_1 = 0.00000441$, $p_6 = 0.00000206$, and $T(t) = 0.5 \exp(p_{10} T_m(t))$. $T_m(t)$ is the mean daily temperature.

For all fixed points of C_{som} , the derivative is $\prod_{i=0}^{364} (1 - p_9 T(i))$. So the magnitude of the derivative for any fixed point depends on $T(i)$. Stability depends on $\prod_{i=0}^{364} (1 - p_9 T(i)) = 0.9992$. As this is smaller than 1, this means that all fixed points are stable. This is possible as the fixed points for C_{som} depend on the fixed points for C_w and C_{lit} .

4.4.5 Periods of the Carbon Pools' Fixed Points

Fig. 4.24 shows one period of each carbon pool's non-trivial stable fixed point. The foliar carbon pool is the only pool which has an unstable fixed point. The lowest point of C_f occurs around the end of March, which is the start of the spring for this latitude and its highest point occurs at the end of the summer. The other pools' fixed points are always stable for a given cycle of C_f , as long as their p -parameters stay within their physical range. So for the full map of all the five pools, there are two stable fixed points and one saddle point with a one-dimensional unstable manifold and a four-dimensional stable manifold.

We can see that, if we assume day 1 is December 21st, the litter carbon pool has its lowest value around day 230, which is in July and its highest value around day 60, which is near March. In contrast, the root carbon pool experiences its lowest value around April and its highest value around October, which is similar to the lowest and highest values for the woody carbon pool. The soil and organic matter carbon pool shows its lowest value near November and its highest value near May. This makes sense, as the litter pool will decompose and its carbon will disappear into the soil a bit later.

4.5 Preliminary Results for DALEC Deciduous

In this section we find that the model structure of DALEC DE is similar to that of DALEC EV. As for DALEC EV the model equations [13] take the form of daily maps, see Chapter 3.

The main difference between the two models is that DALEC DE has an extra labile pool, C_{lab} , which is used in the spring to kick-start the growth of leaves. It depletes in the spring and then builds up again in the autumn, when the leaves are dropping. Fig. 4.25 shows

the foliar and labile pools plotted over one year. In the spring, approximately between day 100 and day 150, foliar turnover is switched off and labile turnover is switched on, allowing leaves to start growing again. This will depend on temperature and the time of year. In the summer, both the labile pool and the foliar pool remain unchanged for some time (approximately between day 200 and 250), which is when both the foliar turnover and the labile turnover are switched off. In the autumn, foliar turnover is switched on, when the minimum temperature goes below the value of p_{13} after more than 200 days of the year have passed. The trees drop their leaves and photosynthesis comes to a halt. The foliar carbon pool is depleted and carbon is stored in the labile pool until spring.

We see that in this model, the foliar and labile pools are coupled but decouple from the other four equations. Even though the dynamics of the forest now depends on two carbon pools, we found a bifurcation structure which is similar to that of DALEC EV. However, this time, as it is important that the foliar carbon depletes to zero in the autumn and labile carbon also depletes to zero in the spring, we found that parameter p_{14} , which is the fraction of leaf loss to litter, is a more realistic parameter to use for finding a tipping point, as varying this parameter retained the depletion of both pools. Varying p_5 would stop the foliar carbon pool going to zero in the autumn. So when we keep all the other parameters fixed and vary p_{14} , we find a similar tipping point bifurcation for both the foliar carbon pool C_f and the labile carbon pool, C_{lab} , see Figs 4.26 and 4.27. It is interesting to note that the top branches of both graphs go up much steeper than the top branch in Fig. 4.2. Consequently a small change in p_{14} causes a large change in the value of the fixed point.

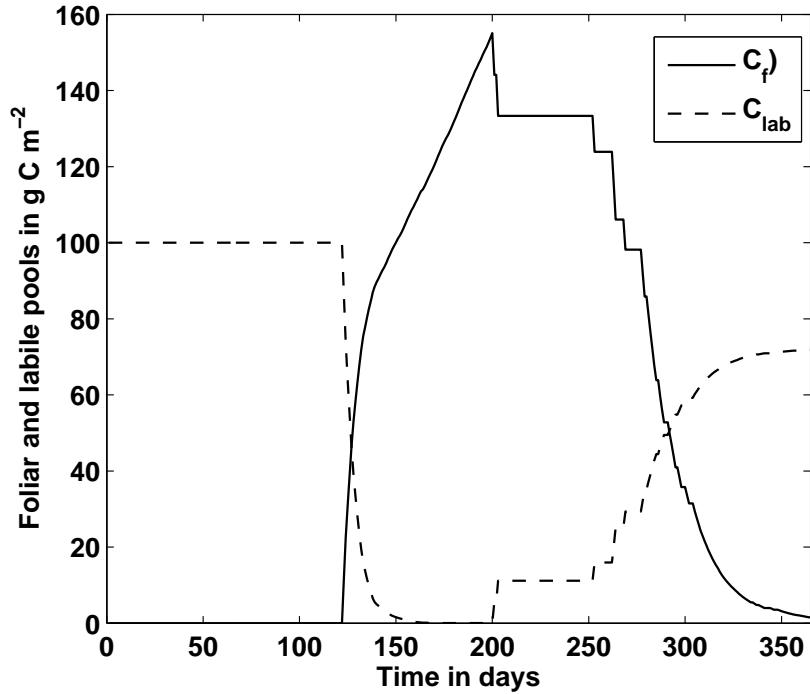


Figure 4.25: The foliar and labile pools plotted over one year.

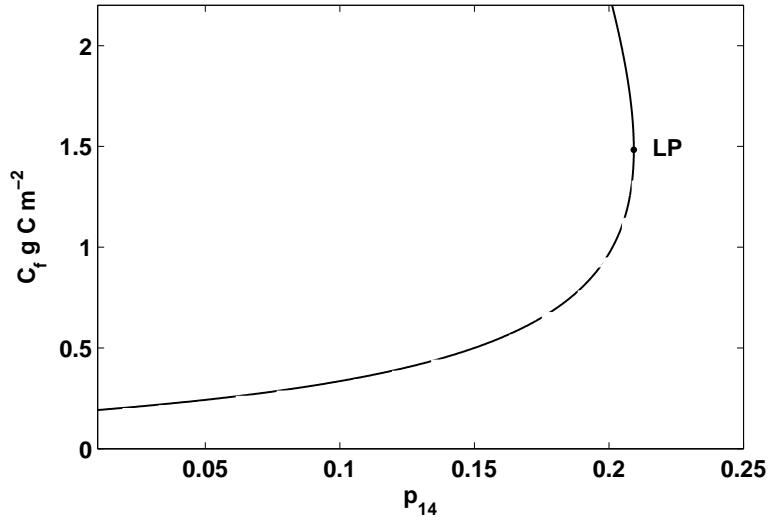


Figure 4.26: Limit point for parameter p_{14} on a path of fixed points of the annual map for C_f for DALEC DE. The top branch is stable, the bottom branch is unstable and there is also a fixed point branch at $C_f = 0$. The fixed points are sampled on day 1 of each year.

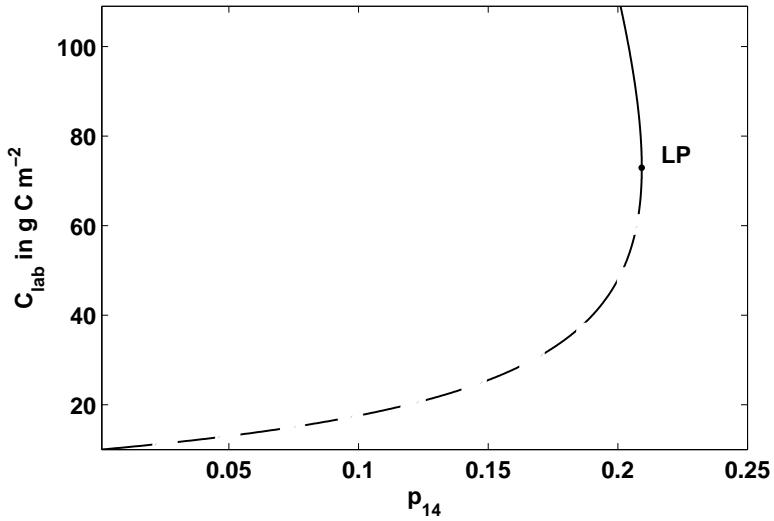


Figure 4.27: Limit point for parameter p_{14} on a path of fixed points of the annual map for C_{lab} for DALEC DE. The top branch is stable, the bottom branch is unstable and there is also a fixed point branch at $C_{lab} = 0$. The fixed points are sampled on day 1 of each year.

4.6 Conclusion

In this chapter we found that the dynamics of an evergreen forest, according to the DALEC EV model, fundamentally depend on the behaviour of the foliar carbon pool and the parameters involved in this pool, $p_3(1 - p_2)$ (fraction of GPP allocated to foliage) and p_5 (daily turnover rate of foliage). The dynamics of the foliar carbon pool show a tipping point, which is dependent on the value of these parameters and explains how, for certain values of the parameters, according to the model, a forest is expected to die out, without any chance of coming back to life. Likewise, for other values and depending on the initial conditions of the foliar pool, a forest grows to an annual cycle and lives. For example, Fig. 4.2 shows the tipping point for the daily turnover rate of foliage, p_5 , with a value of approximately 0.00315. Beyond this value and for a fixed value of p_{23} , the trees lose too much carbon through the loss of their needles to survive. Since leaves produce photosynthate, this means that the trees cannot feed themselves and they will die eventually. The three parameters are dependent on each other and there needs to be

a balance between them for the forest to thrive. Suppose, in Fig. 4.3, that the forest is at a point in the graph where $p_{23} = 0.04$ and $p_5 = 0.002$. This is a point in the area of the graph where a forest is expected to show widespread mortality. The reason for this is that the fraction of GPP allocated to foliage is not in balance with the daily turnover rate of the foliage. For a value of $p_{23} = 0.04$, the daily foliar turnover rate, p_5 , would need to be smaller than 0.001 for a forest to have a chance to stabilise and live, depending on the initial conditions of the foliar carbon. We could look at it from yet a different perspective: if $p_3(1 - p_2)$ is very small then this will almost always result in the forest dying, since very little of the GPP is then allocated to the foliage. Clearly, these conditions should be avoided in order for the forest to live.

A fixed point analysis of the other four pools showed that, as they are dependent on C_f , the foliar pool, which has two stable fixed points and one unstable fixed point, all have stable fixed points. This implies that in combination with the fixed points of the other pools, the unstable fixed point for C_f is a saddle point. Due to the nature of the unstable saddle fixed point we will only in reality observe the fixed points which are dependent on the stable fixed points of C_f . The root pool, C_r , and the litter pool, C_{lit} , each reach their fixed point fairly soon, similar to the C_f pool, but due to the very small value of parameters p_6 and p_9 in the wood pool, C_w , and soil and organic matter pool, C_{som} , these two pools take much longer to reach their fixed points.

We carried out a preliminary study of DALEC deciduous, which is a slightly more complex model due to leaf loss in the autumn and the requirement of an extra carbon pool and found that it exhibits a similar tipping point structure to the evergreen model. In this case, we see C_f and C_{lab} decouple from the other pools. For DALEC DE we chose parameter p_{14} , the fraction of leaf loss transferred to litter, as a more realistic parameter to find the tipping point, instead of parameter p_5 , the daily turnover rate of foliage. It would be interesting to examine the roles that the other parameters in the C_f and the C_{lab} equation, the combination $p_3(1 - p_2)$, p_5 , p_{15} , p_{16} and p_{17} , play.

It is interesting to note that, using the Reflex project data for the Loobos site in the Netherlands and the parameter values given to us by the University of Edinburgh in the DALEC EV version, the parameter values are near the tipping point. When we used climate and parameter data from a different pine forest, in Oregon, USA, we found that those parameter values were also near the line of tipping points, although not as close as the parameters from the Reflex data. We speculate that this may be the most energy efficient situation for a tree to be in. This brings up an interesting question: if this would be a common occurrence, would the values of the parameters change and move close to the limit point again when the area of possible growth and stabilisation on the left hand side of the line of limit points increases, for example through an increase in atmospheric CO₂? We could speculate that in that case the tree may want to decrease its fraction of GPP allocated to foliage ($p_3(1 - p_2)$) and concentrate on putting more carbon into the wood, allowing the tree to grow, thereby moving its position on the graph closer to a limit point again.

We have seen that it may not be necessary in DALEC, for a forest that is living in a steady climate, to use daily driver data, but that it is sufficient to represent the drivers by a smooth cosine wave, where all that is needed is to measure the amplitude, mean and phase of the annual cycle. We do realise that smoothing the drivers would not always be sensible, certainly not when the model is used to study certain shock phenomena. Also, we do acknowledge that smoothing the drivers takes away the natural variability, to which ecosystems respond. According to Medvigy *et al.* [32], removal of high-frequency variability significantly enhances decadal NEE, GPP and total respiration. Interestingly, according to their research, solar radiation has a strong effect, whereas temperature variability only has a minor impact. During our smoothing exercise, we also found that the model seems to be more sensitive to solar radiation, which is why we had to resort to data assimilation methods to make the function fit, than to the temperature data. In DALEC, using the smoothed drivers did not make much difference to the results, perhaps as the

high frequency variability was already taken out, the data being daily averages rather than hourly observations.

We have been using periodic drivers in this analysis, but in reality the drivers are not periodic. However, if one assumes that drivers are periodic with a small perturbation, which effectively corresponds to noise in the system, then this would not change the overall structure and therefore our results would still remain valid.

Chapter 5

Gross Primary Production and C_f Revisited

5.1 Introduction

In this chapter we write the Gross Primary Production (GPP) in such a way that it becomes clear how it depends on C_f . The equation for GPP is complicated and contains a square root term in the equation for the C_i , which is the CO₂ concentration at the site of carboxylation. By simplifying this square root term the GPP becomes less complicated. We then look at different methods of averaging the GPP over an annual cycle, either by averaging all the time dependent functions or by averaging all the time dependent functions bar the irradiance. We select one of these simplified GPP equations to create an annual map for C_f and this allows us to produce the same limit point graph and line of limit points graph that we found in Chapter 4, but this time without the use of continuation software.

5.2 Simplifying and Averaging the GPP

The GPP looks is defined as follows:

$$GPP(t) = \frac{E_0(t)I(t)g_c(t)(C_a - C_i(t))}{E_0(t)I(t) + g_c(t)(C_a - C_i(t))}(a_2s(t) + a_5), \quad (5.1)$$

where only $E_0(t)$ and $C_i(t)$ are functions of $C_f(t)$. These and the other functions in the GPP are described in Appendix A.

We concentrate on $(C_a - C_i(t))$ in the equation for the GPP to see if we can simplify it. This would make it much easier to express the GPP in terms of C_f . The equation for $C_i(t)$, the CO₂ concentration at site of carboxylation, is given by:

$$C_i(t) = \frac{1}{2} \left[C_a + q - p(t) + \sqrt{(C_a + q - p(t))^2 - 4(C_aq - p(t)a_3)} \right], \quad (5.2)$$

where:

$$\begin{aligned} C_a &= 380 \text{ (atmospheric CO}_2\text{ concentration, constant averaged over three years);} \\ q &= a_3 - a_4 = -204.64527, \\ a_3 &= 4.22273, \\ a_4 &= 208.868, \\ p(t) &= \text{photosynthate or } \gamma(t)C_f(t), \\ \text{where } \gamma(t) &= \frac{p_{11}N \exp(a_8 T_{max}(t))}{lma g_c(t)}. \end{aligned}$$

This comes from the paper by Williams *et al.* 1997, where the Aggregated Canopy Model, ACM, a daily time step model for photosynthesis (which is GPP in DALEC) is derived from the Soil-Plant-Atmosphere model, SPA, a half-hourly step model for photosynthesis [50].

Williams *et al.* expresses the whole-canopy N -limited (where N is foliar nitrate) photo-

synthetic capacity, $p_N(t)$ as:

$$p_N(t) = a_1 N L(t) \exp(a_2 T(t)), \quad (5.3)$$

where $a_1 = 2.95$ and $a_2 = 0.018$ are calibrated parameters, N with a value between $[1.0, 4.0]$ is the foliar nitrate concentration, $L(t)$ with a value between $[0.5, 8.0]$ is LAI $= C_f(t)/lma$ and $T(t)$, with values between $[7, 30]$ is the mean daily temperature.

The photosynthetic capacity corrected for CO₂ limitation, $p_C(t)$, is modelled as:

$$p_C(t) = \frac{p_N(t)(C_i(t) - \theta)}{k + (C_i(t) - \theta)}, \quad (5.4)$$

where $\theta = 32.6$ is the CO₂ compensation point in micromoles per mole, and $k = 576.7$, a half-saturation constant in micromoles per mole.

The rate of diffusion of CO₂ to the state of C fixation, $p_D(t)$, is modelled as:

$$p_D(t) = g_c(t)(C_a - C_i(t)), \quad (5.5)$$

where Williams states that C_a can have a value between $[250, 750]$ ppm and $g_c(t)$ is expressed as:

$$g_c(t) = \frac{-\psi_d \exp(b_1 T(t))}{b_2 H + D_T}, \quad (5.6)$$

where ψ_d , with a value between $[-4.5, -0.5]$ MPa, is the soil moisture index, $b_1 = -0.029$ and $b_2 = 0.315$ are temperature coefficients, H , with a value between $[0.5, 5.0]$ meters is canopy height and D_T , with values between $[1, 12]$ °C, is the temperature range.

In order to find an equation for $C_i(t)$, assuming a steady state, $p_C(t)$ is set equal to $p_D(t)$. This gives the quadratic equation for $C_i(t)$:

$$C_i(t)^2 - (C_a + q - p(t))C_i(t) + C_a q - p(t)a_3 = 0, \quad (5.7)$$

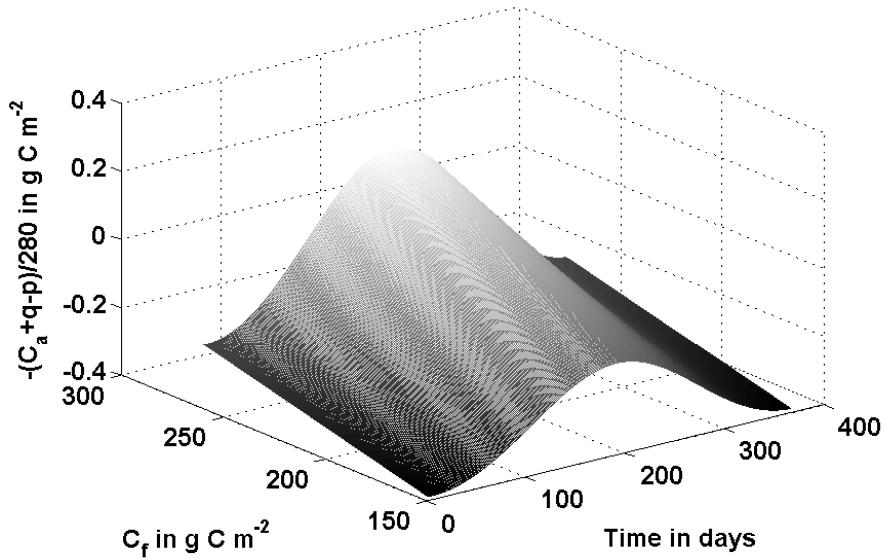


Figure 5.1: The coefficient of $\tilde{C}_i(t)$ in equation (5.8), $\frac{-(C_a + q - p(t))}{280}$, in terms of C_f and time.

where $q = \theta - k$, $a_3 = \theta$ and $p(t) = p_N(t)/g_c(t)$.

Our aim is to find an approximate expression for $C_i(t)$ that does not involve a square root. First we consider if any of the terms in the quadratic equation (5.2) are dominant. We rescale $C_i(t)$ by defining $\tilde{C}_i(t) = \frac{C_i(t)}{280}$ (where 280 is roughly the mean of $C_i(t)$ at C_f 's fixed point), which makes the $\tilde{C}_i^k(t)$ for $k = 0, 1, 2$ close to 1 and therefore moves the scaling into the coefficients. Dividing equation 5.7 through by 280^2 gives:

$$\tilde{C}_i(t)^2 - \frac{(C_a + q - p(t))}{280} \tilde{C}_i(t) + \frac{(C_a q - p(t)a_3)}{280^2} = 0. \quad (5.8)$$

Figs 5.1 and 5.2 shows plots of the coefficients against $C_f(t)$ and t . We see that the orders of the contributions of the $\tilde{C}_i^k(t)$ are too close; the first coefficient is of order 10^{-1} and the second of order 10^0 . This means no term is dominant enough and therefore it will not be straightforward to simplify equation (5.2).

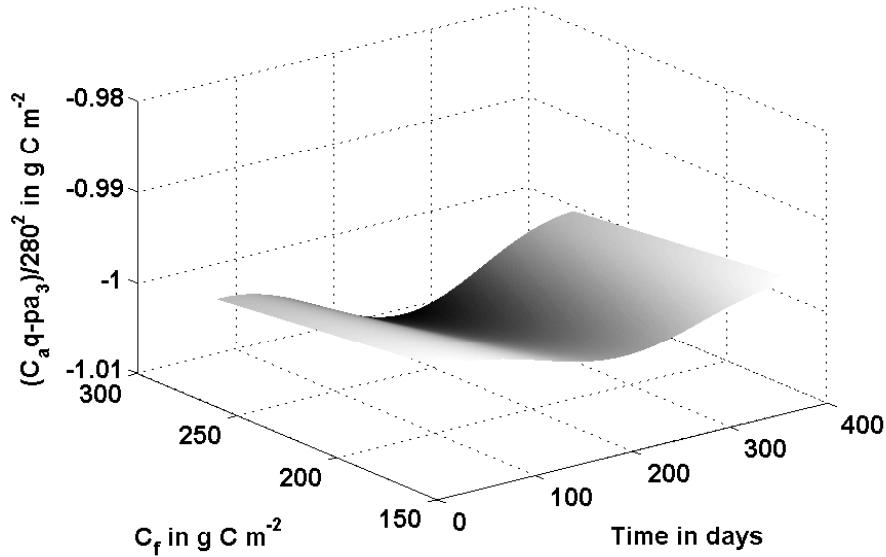


Figure 5.2: The coefficient in equation (5.8), $\frac{C_a q - p(t) a_3}{280^2}$, in terms of C_f and time.

Our next attempt at simplifying equation (5.1) is to consider the expression $C_a - C_i(t)$ in equation 5.1 and to see if we can express the term under the square root in equation (5.2) by a binomial approximation. For simplicity we will refer to $C_i(t)$ as C_i and to $p(t)$ as p .

First we write the quadratic in equation (5.7) in terms of $C_i - C_a$:

$$(C_i - C_a + C_a)^2 - (C_a + q - p)(C_i - C_a + C_a) + C_a q - a_3 p = 0, \quad (5.9)$$

$$(C_i - C_a)^2 + (C_a - q + p)(C_i - C_a) + p(C_a - a_3) = 0, \quad (5.10)$$

or equivalently:

$$(C_a - C_i)^2 - (C_a - q + p)(C_a - C_i) + p(C_a - a_3) = 0. \quad (5.11)$$

Now to rescale $(C_a - C_i)$ we define $C_{ai} = \frac{(C_a - C_i)}{92}$, as 92 is the approximate value of

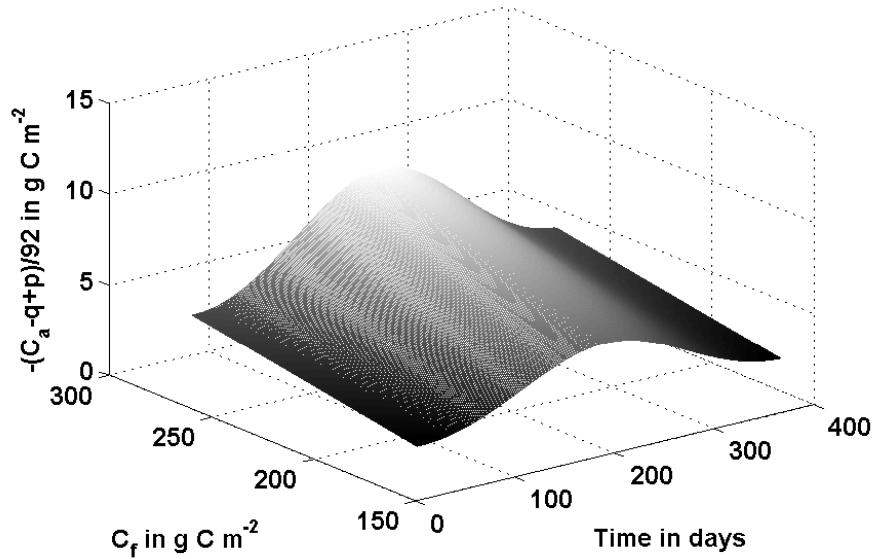


Figure 5.3: The coefficient of $C_{ai}(t)$ in equation (5.12), $\frac{-(C_a - q + p(t))}{92}$, in terms of C_f and time.

the mean of $(C_a - C_i)$ at the fixed point of C_f .

$$C_{ai}^2 - \frac{(C_a - q + p)}{92} C_{ai} + \frac{p(C_a - a_3)}{92^2} = 0. \quad (5.12)$$

Figs 5.3 and 5.4 show the results of the rescaling. We can see that both the coefficient for C_{ai} and the constant term are larger than 1, but of opposite signs and balancing. We could neglect the C_{ai}^2 term in equation (5.11) and write:

$$-(C_a - q + p)C_{ai} + p(C_a - a_3) = 0, \quad (5.13)$$

which means that:

$$C_a - C_i = \frac{p(C_a - a_3)}{C_a - q + p} \quad (5.14)$$

Now an alternative but more rigorous approach to simplifying the $C_a - C_i(t)$ term is

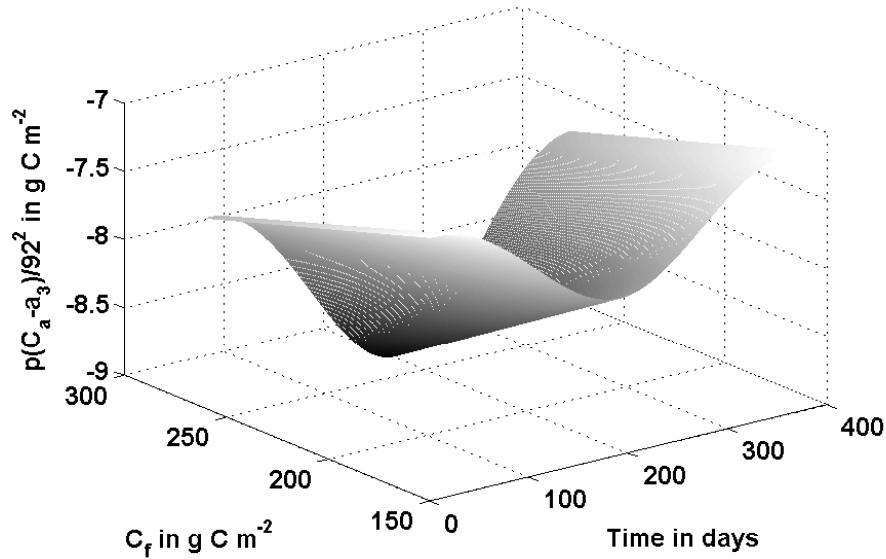


Figure 5.4: The constant coefficient in equation (5.12), $\frac{p(t)(C_a - a_3)}{92^2}$, in terms of C_f and time.

described. From equation (5.11) we can see that:

$$C_a - C_i = \frac{1}{2} \left[(C_a - q + p) \pm \sqrt{(C_a - q + p)^2 - 4p(C_a - a_3)} \right] \quad (5.15)$$

and so we need to know which sign to take for the square root term. Subtracting C_i given by equation (5.2) from C_a gives:

$$C_a - C_i = C_a - \frac{1}{2} \left[C_a + q - p + \sqrt{(C_a + q - p)^2 - 4(C_a q - p a_3)} \right], \quad (5.16)$$

$$= \frac{1}{2} \left[C_a - q + p - \sqrt{(C_a + q - p)^2 - 4(C_a q - p a_3)} \right]. \quad (5.17)$$

Now:

$$(C_a + q - p)^2 = (C_a - q + p)^2 + 4C_a(q - p).$$

So,

$$C_a - C_i = \frac{1}{2} \left[C_a - q + p - \sqrt{(C_a - q + p)^2 - 4p(C_a - a_3)} \right] \quad (5.18)$$

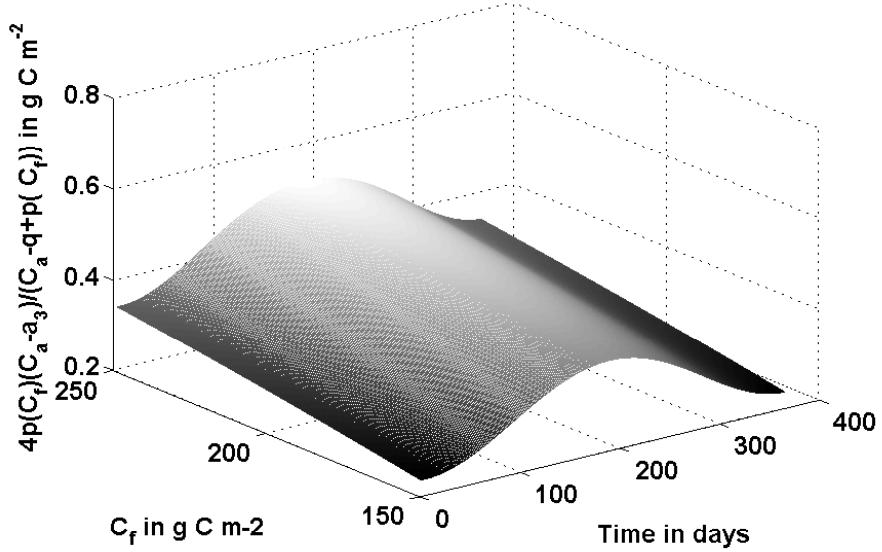


Figure 5.5: The term $\frac{4p(C_a - a_3)}{(C_a - q + p)^2}$ (in terms of C_f and t) is small relative to 1 in equation (5.20).

and comparing this with (5.15), we see that we require the negative square root.

To simplify $(C_a - C_i)$, we note that:

$$C_a - C_i = \frac{1}{2} \left[C_a - q + p - |C_a - q + p| \sqrt{1 - \frac{4p(C_a - a_3)}{(C_a - q + p)^2}} \right]. \quad (5.19)$$

From Fig. 5.5 we can see that the term $\frac{4p(C_a - a_3)}{(C_a - q + p)^2}$ is small relative to 1. Hence:

$$\left(1 - \frac{4p(C_a - a_3)}{(C_a - q + p)^2}\right)^{\frac{1}{2}} = 1 - \frac{2p(C_a - a_3)}{(C_a - q + p)^2} + O(2). \quad (5.20)$$

It must be true that $C_a - q + p > 0$, since $q < 0$ and $p(t) > 0$ for all t and so we obtain:

$$\begin{aligned} C_a - C_i &\approx \frac{1}{2}[C_a - q + p - (C_a - q + p)\left(1 - \left(\frac{2p(C_a - a_3)}{(C_a - q + p)^2}\right)\right)] \\ &= \frac{p(C_a - a_3)}{C_a - q + p} \end{aligned} \quad (5.21)$$

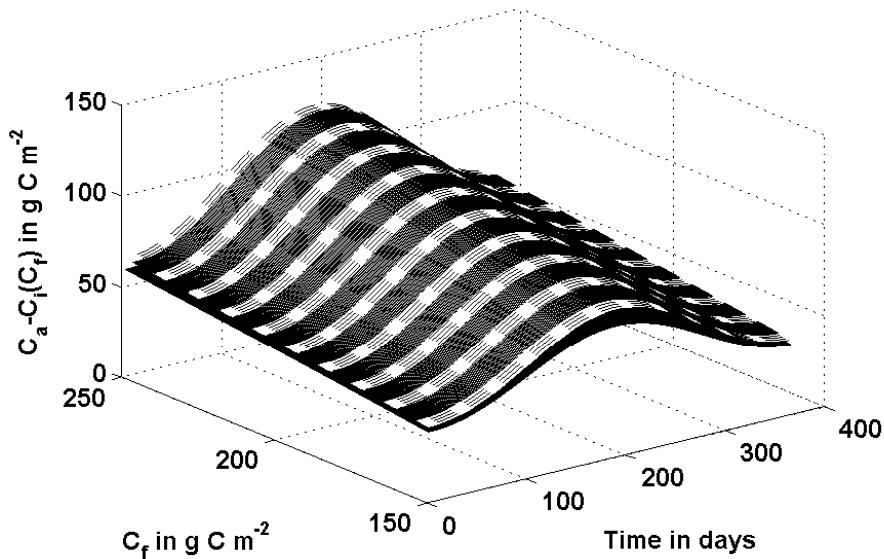


Figure 5.6: The dashed surface represents $C_a - C_i$ as it is originally, (5.2) and the black surface represents when $C_a - C_i$ is simplified as in equation (5.21), both in terms of C_f and time.

Numerically, we find that, using a mean value for $p(t) = 104.26$ for the year with initial condition $C_f = 150$:

$$\left(1 - \frac{4p(C_a - a_3)}{(C_a - q + p)^2}\right)^{\frac{1}{2}} \approx 0.82 \quad (5.22)$$

and

$$1 - \left(\frac{2p(C_a - a_3)}{(C_a - q + p)^2}\right) \approx 0.83, \quad (5.23)$$

which is a difference of around 2%. At the fixed point for C_f these values would be a bit smaller, between 0.76 and 0.79, as the average for $p(t)$ goes up. Equation (5.22) confirms that $\left(\frac{4p(C_a - a_3)}{(C_a - q + p)^2}\right)$ is small relative to 1 on average.

Equations (5.14) and (5.21) are the same. To explain this we need to look at the general equation of a quadratic function. Suppose we have:

$$ax^2 + bx + c = 0 \quad (5.24)$$

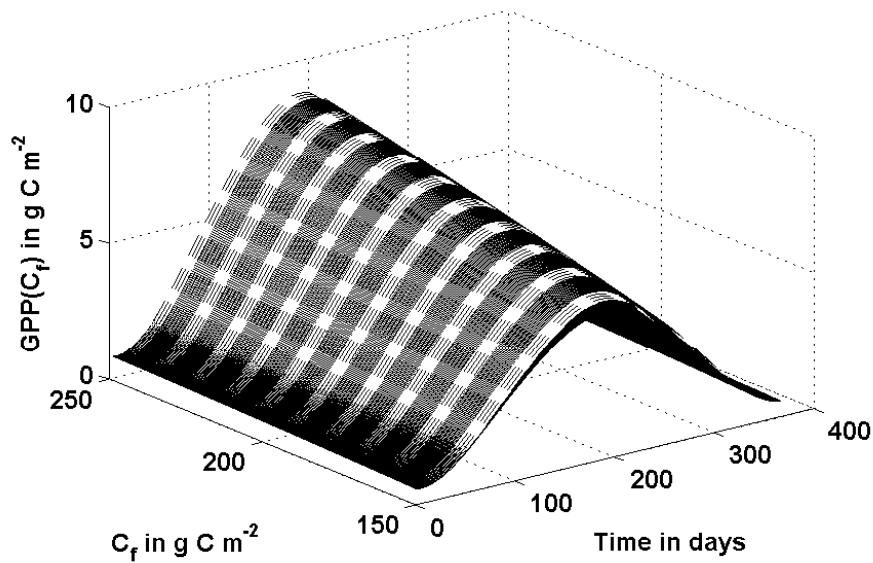


Figure 5.7: The dashed surface represents GPP plotted using equation (5.1) and the black surface represents GPP using the simplified $C_a - C_i$ in equation (5.21), both in terms of C_f and time.

To achieve equation (5.14), we assume that both $b \gg a$ and $c \gg a$. Therefore:

$$-bx + c \approx 0 \quad (5.25)$$

and therefore

$$x \approx \frac{c}{b}. \quad (5.26)$$

For the method used to achieve equation (5.21) we use that:

$$x = \frac{b \pm \sqrt{b^2 - 4ac}}{2a} \quad (5.27)$$

and choose the negative root.

Therefore we now can say that:

$$x = \frac{b}{2a} - \frac{|b|}{2a} \left\{ 1 - \frac{4ac}{b^2} \right\}^{\frac{1}{\alpha}} \quad (5.28)$$

If $\frac{4ac}{b^2} \ll 1$, then:

$$x \approx \frac{b}{2a} - \frac{|b|}{2a} + \frac{1}{2} \frac{|b|}{2a} \frac{4ac}{b^2} \quad (5.29)$$

Since $b > 0$, this becomes:

$$x \approx \frac{c}{b}, \quad (5.30)$$

as before.

Fig. 5.7 shows the GPP as it is originally and using the simplified $C_a - C_i$. There is not very much difference between the two versions and we can now use this simplification for our next purpose, namely creating a relatively simple annual map for \mathbf{C}_f , which creates a simple mechanism for experimenting with shocks to the model.

Substituting the expression for $C_a - C_i$ given in equation (5.21) into equation (5.1) and then expanding and simplifying gives:

$$GPP(C_f(t), t) = \frac{a(t)C_f(t)^2}{b(t) + c(t)C_f(t) + d(t)C_f(t)^2}, \quad (5.31)$$

where:

$$\begin{aligned}
 a(t) &= a_7 I(t) \gamma(t) g_c(t) (a_3 - C_a) (a_2 s(t) + a_5), \\
 b(t) &= \gamma(t) g_c(t) a_9 lma^2 (a_3 - C_a), \\
 c(t) &= -a_7 I(t) (C_a - q), \\
 d(t) &= -\gamma(t) (a_7 I(t) + g_c(t) (C_a - a_3)),
 \end{aligned}$$

(all expressed in g C m⁻²)

Fig. 5.8 shows C_f over a number of years, in its exact form and with the approximated GPP; we can see that the fixed point for C_f changes when we use the approximation for the GPP. The fixed point is a lot lower than when using the exact GPP, as the approximated GPP is lower than the exact GPP, see Fig. 5.9, which shows the GPP over a number of years. The reason why the approximated GPP seems to fit better to the bottom of the original GPP is that at certain times of the year, particularly in the summer, $\frac{4p(C_a - a_3)}{(C_a - q + p)^2}$ from equation (5.22) is not much smaller than 1. Although this does not change the analytical behaviour of the model, we can solve this by using a constant fitting factor, f_0 , in the equation for C_f , which is close to 1 and varies with the type of forest used:

$$C_f(t+1) = (1 - p_5)C_f(t) + f_0 p_3 (1 - p_2) GPP(C_f(t), t) \quad (5.32)$$

The fitting factor, f_0 , which has been optimised by data assimilation (by fitting it as a parameter to C_f), has a value of around 1.085. Fig. 5.10 shows a plot for C_f with the fitting factor included. The reason we need a fitting factor close to 1 is that once the GPP increases by 8%, this has a compound effect, as every time we calculate the GPP, we again multiply the already increased GPP by the fitting factor. It would therefore in particular benefit the summer values, which are much higher than the winter values.

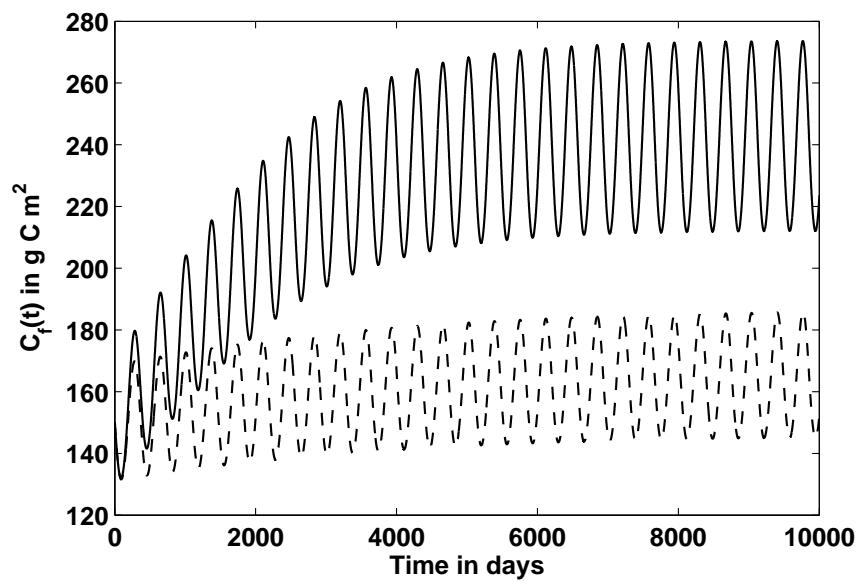


Figure 5.8: C_f plotted using the original GPP and the simplified GPP as in equation (5.31).

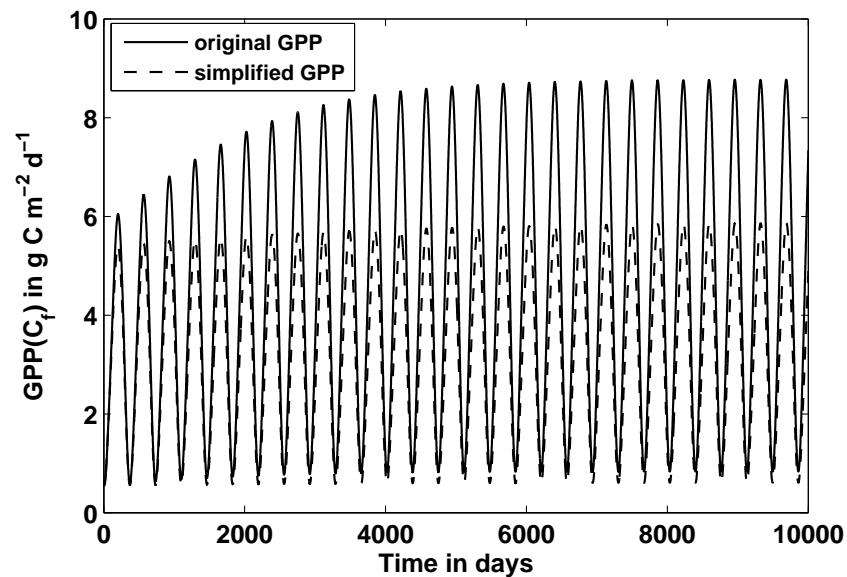


Figure 5.9: Original and simplified GPP as in equation (5.31).

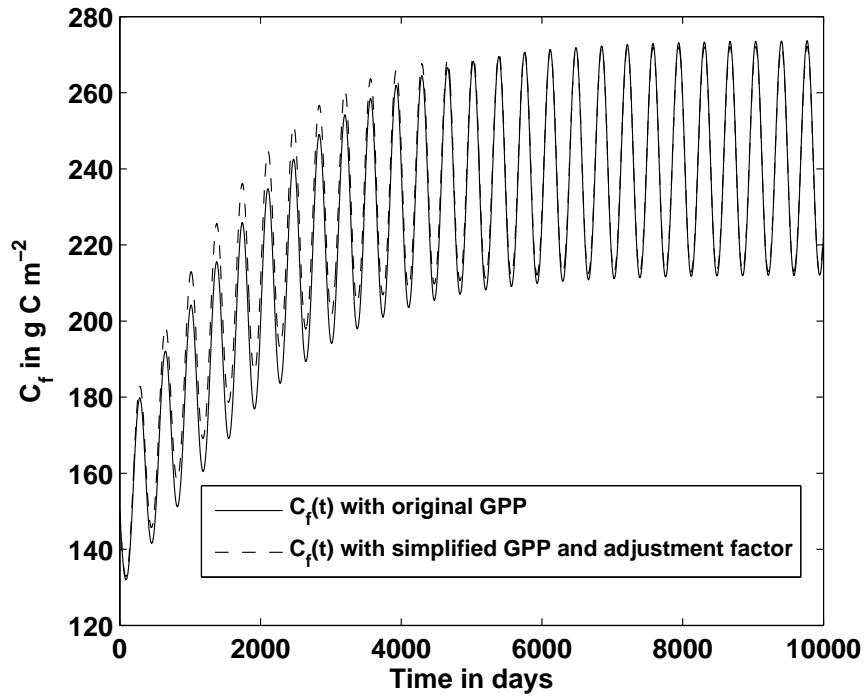


Figure 5.10: C_f plotted using simplified GPP as in equation (5.31) and a fitting factor of approximate value 1.085.

5.3 An Annual Map for C_f

The next step is to create an annual map for C_f using the simplified GPP, given in equation (5.31), which can then be used as a mechanism for experimenting with shocks to the model. Iterating the daily map 365 times to give an annual map is not possible analytically due to the nonlinear GPP function. Thus we replace the GPP by a constant mean during an annual cycle, assuming that $C_f(i) = C_f(0)$ (which is a constant) in the GPP.

The daily map during the first year then becomes:

$$C_f(i+1) = (1 - p_5)C_f(i) + p_3(1 - p_2)\overline{GPP}(C_f(0)). \quad (5.33)$$

Iterating this map 365 times gives:

$$C_f(365) = (1 - p_5)^{365} + p_3(1 - p_2)\overline{GPP}(C_f(0)) \sum_{i=0}^{364} (1 - p_5)^i, \quad (5.34)$$

using the technique described in Section 4.3.1. The value of C_f used in the GPP is updated at the beginning of each annual cycle, so therefore the annual map for C_f is:

$$C_f^{(n+1)} = (1 - p_5)^{365} C_f^{(n)} + \left(\frac{(1 - (1 - p_5)^{365})}{p_5} \right) p_3(1 - p_2)\overline{GPP}(C_f^{(n)}), \quad (5.35)$$

where n is the year. The value of $C_f^{(n)}$ is the mean value of the foliar pool over one year and is updated at the beginning of each year.

An approximation for the mean of the GPP can be found in a number of ways. We discuss three different approximations.

5.3.1 Approximation 1

The first approximation we consider is to approximate $GPP(C_f(t), t)$ by taking the form (5.31) with all the time dependent functions that form $a(t)$, $b(t)$, $c(t)$ and $d(t)$ taken as their average value over one year. So, for example, $I(t) \approx \bar{I} = 8.4774$.

In this case our annual map becomes:

$$C_f^{(n+1)} = (1 - p_5)^{365} C_f^{(n)} + \left(\frac{(1 - (1 - p_5)^{365})}{p_5} \right) p_3(1 - p_2) \frac{\bar{a} C_f^{(n)2}}{\bar{b} + \bar{c} C_f^{(n)} + \bar{d} C_f^{(n)2}}, \quad (5.36)$$

either where, option 1:

$$\bar{a} = a_7 \bar{I} \bar{g}_c \bar{\gamma} (a_3 - C_a) (a_2 \bar{s} + a_5) \quad (5.37)$$

$$\bar{b} = \bar{\gamma} \bar{g}_c a_9 l m a^2 (a_3 - C_a) \quad (5.38)$$

$$\bar{c} = -a_7 \bar{I} (C_a - q) \quad (5.39)$$

$$\bar{d} = -\bar{\gamma} (a_7 \bar{I} + \bar{g}_c (C_a - a_3)), \quad (5.40)$$

(all expressed in g C m⁻²)

or, as an alternative approach, option 2, to average each of the coefficients, in which case we have:

$$\bar{a} = \overline{a_7 I g_c \gamma (a_3 - C_a) (a_2 \bar{s} + a_5)} \quad (5.41)$$

$$\bar{b} = \overline{\gamma g_c a_9 l m a^2 (a_3 - C_a)} \quad (5.42)$$

$$\bar{c} = \overline{-a_7 I (C_a - q)} \quad (5.43)$$

$$\bar{d} = \overline{-\gamma (a_7 I + g_c (C_a - a_3))} \quad (5.44)$$

(all expressed in g C m⁻²)

The annual map gives the average value of C_f for every year. To compare this annual map with the original C_f , we have created a plot of all the yearly averages of the daily C_f , which is in Fig. 5.11, using option 1. To make the annual map fit better, we introduce a fitting factor, f_1 , which multiplies $\overline{GPP}(C_f(0))$, so the annual map becomes:

$$C_f^{(n+1)} = (1 - p_5)^{365} C_f^{(n)} + f_1 \left(\frac{(1 - (1 - p_5)^{365})}{p_5} \right) p_3 (1 - p_2) \frac{\bar{a} C_f^{(n)2}}{\bar{b} + \bar{c} C_f^{(n)} + \bar{d} C_f^{(n)2}}. \quad (5.45)$$

The value of f_1 is determined by matching the fixed point for the annual map with that of the annually averaged daily map. The fitting factor is close to 1, but is different for

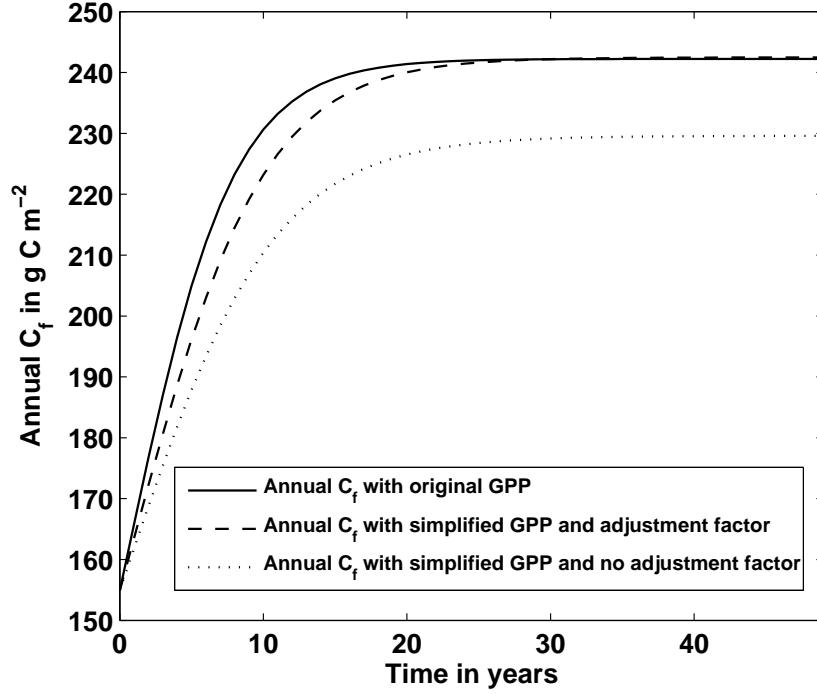


Figure 5.11: An annual map of C_f using the simplified GPP as in equation (5.31) with and without a fitting factor, versus the annual map using the original GPP. For both plots annual C_f was created from the daily map, taking the mean C_f per year.

each forest. For the Loobos forest $f_1 \approx 1.016$.

Fig. 5.12 shows the annual map for the original GPP and created using the simplified GPP as in equation (5.31). This shows clearly why we need a fitting factor for the annual map for C_f .

We now examine to what extent this annual map for C_f (5.45) with approximations for \overline{GPP} gives a good approximation for the line of limit points.

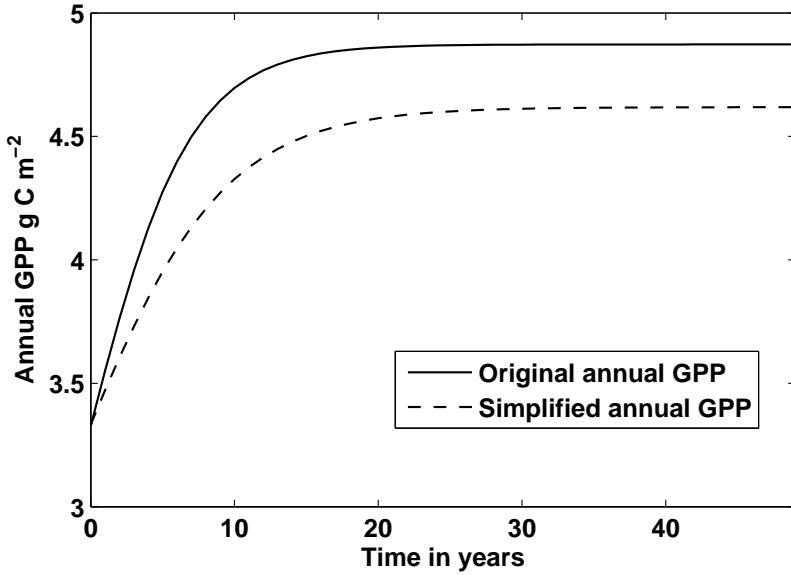


Figure 5.12: An annual map of the original GPP and of the simplified GPP as in equation (5.31). Both annual maps were created by taking the mean yearly GPP from the daily calculation of GPP.

Fixed point solutions of the annual map (5.45) satisfy $C_f^{(n+1)} = C_f^{(n)} \equiv C_f$. So we get:

$$C_f = (1 - p_5)^{365} C_f + f_1 \left(\frac{(1 - (1 - p_5)^{365})}{p_5} \right) p_3 (1 - p_2) \frac{\bar{a} C_f^2}{\bar{b} + \bar{c} C_f + \bar{d} C_f^2}, \quad (5.46)$$

$$0 = (1 - p_5)^{365} C_f - C_f + f_1 \left(\frac{(1 - (1 - p_5)^{365})}{p_5} \right) p_3 (1 - p_2) \frac{\bar{a} C_f^2}{\bar{b} + \bar{c} C_f + \bar{d} C_f^2}. \quad (5.47)$$

This equation has one solution $C_f = 0$, which is the stable trivial fixed point that also exists for the original map. To find the non trivial fixed points, we divide both sides by C_f :

$$(1 - p_5)^{365} - 1 + f_1 \left(\frac{(1 - (1 - p_5)^{365})}{p_5} \right) p_3 (1 - p_2) \frac{\bar{a} C_f}{\bar{b} + \bar{c} C_f + \bar{d} C_f^2} = 0. \quad (5.48)$$

Dividing by $1 - (1 - p_5)^{365}$ gives:

$$-1 + \frac{f_1 p_3 (1 - p_2) \bar{a} C_f}{p_5 (\bar{b} + \bar{c} C_f + \bar{d} C_f^2)} = 0. \quad (5.49)$$

Therefore:

$$p_5 \bar{b} + (p_5 \bar{c} - f_1 p_3 (1 - p_2) \bar{a}) C_f + p_5 \bar{d} C_f^2 = 0, \quad (5.50)$$

which is a quadratic equation in C_f .

The solutions for this quadratic are:

$$C_f = \frac{-p_5 \bar{c} + f_1 p_3 (1 - p_2) \bar{a} \pm \sqrt{(p_5 \bar{c} - f_1 p_3 (1 - p_2) \bar{a})^2 - 4p_5^2 \bar{b} \bar{d}}}{2p_5 \bar{d}}. \quad (5.51)$$

Varying $p_3(1 - p_2)$ or p_5 will give us the stable and unstable branches of fixed points that we found before with the continuation software using the original GPP, see Fig. 5.13.

The discriminant, $(p_5 \bar{c} - f_1 p_3 (1 - p_2) \bar{a})^2 - 4p_5^2 \bar{b} \bar{d}$, set to zero will give us the line of limit points, equivalent to the line that was found before with the continuation software using the original GPP, see Fig: 5.14.

We can either express $p_3(1 - p_2)$ as a function of p_5 or the other way around. Here we have chosen the first option:

$$p_3(1 - p_2) = \frac{p_5(\bar{c} \pm 2\sqrt{\bar{b}\bar{d}})}{f_1 \bar{a}}, \quad (5.52)$$

which is a straight line through the origin in the $(p_3(1 - p_2), p_5)$ plane.

The negative solution is the correct one. Using option 1, equations (5.37)-(5.40), of

averaged values for the GPP gives:

$$\bar{a} = -1.75 \times 10^3, \quad (5.53)$$

$$\bar{b} = -3.14 \times 10^6, \quad (5.54)$$

$$\bar{c} = -3.57 \times 10^4, \quad (5.55)$$

$$\bar{d} = -1.65 \times 10^2, \quad (5.56)$$

(all expressed in g C m⁻²)

$$f_1 = 1.016, \quad (5.57)$$

The line of limit points can now be expressed as:

$$p_3(1 - p_2) = 45.562p_5 \quad (5.58)$$

or, according to the second option for the averaged values of a , b , c and d , equations (5.41)-(5.44) give:

$$\bar{a} = -1.73 \times 10^3, \quad (5.59)$$

$$\bar{b} = -2.99 \times 10^6, \quad (5.60)$$

$$\bar{c} = -3.57 \times 10^4, \quad (5.61)$$

$$\bar{d} = -1.65 \times 10^2. \quad (5.62)$$

(all expressed in g C m⁻²)

$$f_1 = 1.023, \quad (5.63)$$

The line of limit points can be expressed as:

$$p_3(1 - p_2) = 44.623p_5. \quad (5.64)$$

We can see from equations (5.58) and (5.64) that it does not make much of a difference whether we average each time dependent function in $a(t)$, $b(t)$, $c(t)$ and $d(t)$ in equation (5.31) separately or whether we average all of the term $a(t)$, $b(t)$, $c(t)$ and $d(t)$ at once. From now on we use only Option 1 of the averaged values of the GPP: equations (5.53) - (5.57).

Fig. 5.13 shows the fold bifurcation found using the original GPP and the simplified GPP with fitting factor and without. This plot shows that the fitting factor does not change the position of the limit point with respect to C_f .

Fig. 5.14 shows the original line of limit points we found for the daily C_f and the result of plotting equation (5.58). The line of limit points computed using the first approximation compares well with the original model.

5.3.2 Approximation 2

We know from Section 4.3.3 that the model is particularly sensitive to the irradiance. In the second approximation we see if including time dependence only in the irradiance term in the calculation of the GPP is sufficient. We express the irradiance as $I(t) = \alpha + \beta \cos\left(\frac{2\pi t}{365}\right)$. Then the GPP becomes:

$$GPP(C_f, t) = \frac{\tilde{a}(\alpha + \beta \cos(2\pi t/365)) C_f^2}{\tilde{b} + \tilde{c}(\alpha + \beta \cos(2\pi t/365)) C_f + (\tilde{d}(\alpha + \beta \cos(2\pi t/365)) + \tilde{e}) C_f^2}, \quad (5.65)$$

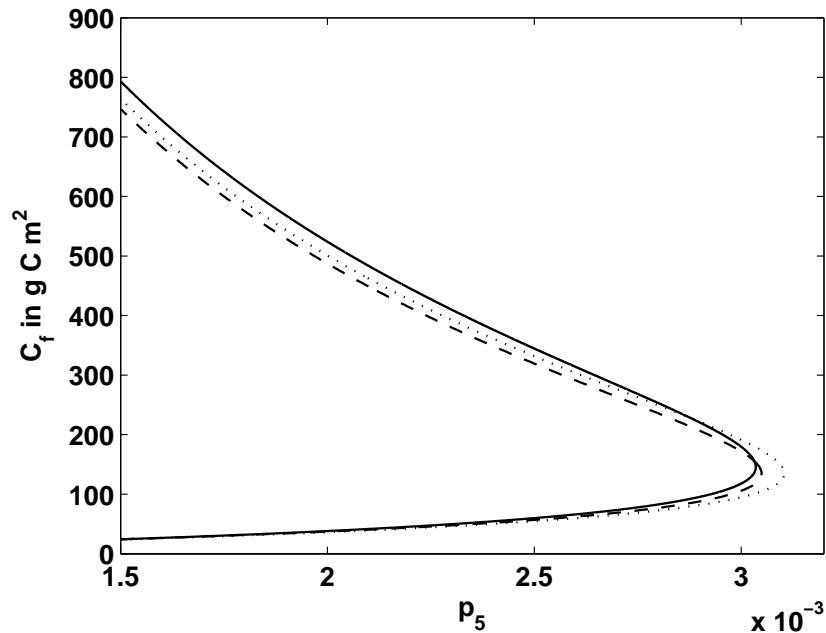


Figure 5.13: Fold bifurcation found using original GPP (continuous line) and simplified GPP (dashed line), varying p_5 , using equations (5.37)-(5.40) of the averaged values for GPP, with a fitting factor of 1.016. The dotted line is the fold bifurcation found using the simplified GPP and no fitting factor

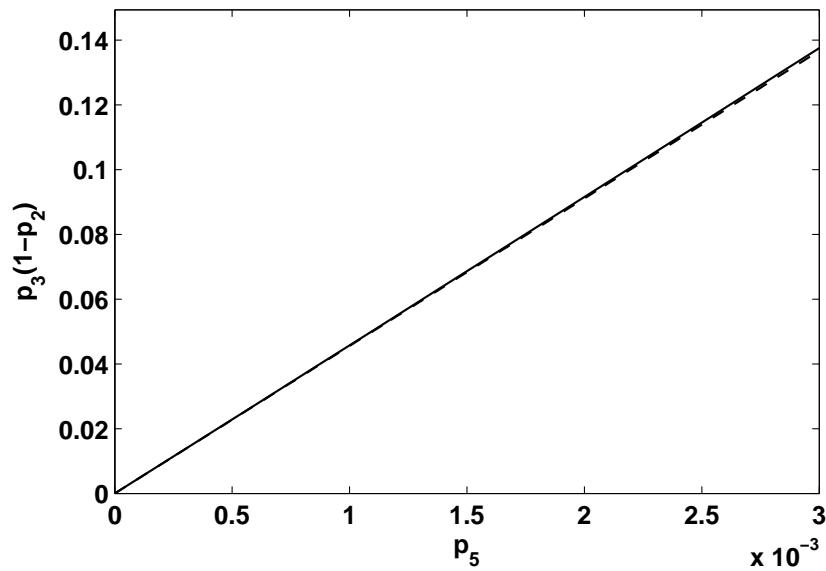


Figure 5.14: Line of limit points found using the original GPP (continuous line) and the averaged simplified GPP (dashed line), varying both p_5 and $p_3(1 - p_2)$, using equations (5.37)-(5.40) of the averaged values for GPP, with a fitting factor of 1.016.

where:

$$\tilde{a} = a_7 \bar{g}_c \bar{\gamma} (a_3 - C_a) (a_2 \bar{s} + a_5) \quad (5.66)$$

$$\tilde{b} = \bar{\gamma} \bar{g}_c a_9 lma^2 (a_3 - C_a) \quad (5.67)$$

$$\tilde{c} = -a_7 (C_a + q) \quad (5.68)$$

$$\tilde{d} = -a_7 \bar{\gamma} \quad (5.69)$$

$$\tilde{e} = \bar{\gamma} \bar{g}_c (a_3 - C_a) \quad (5.70)$$

(all expressed in g C m⁻²)

$$\alpha = 8.4774 \quad (5.71)$$

$$\beta = -7.281 \quad (5.72)$$

When we integrate the GPP with respect to t , assuming that C_f is constant, from $t = 0$ to 365, we get the following equation for the average GPP over one year:

$$\begin{aligned} \overline{GPP}(C_f) = & \tilde{a} C_f \left(\pm (\tilde{b} + \tilde{d} \tilde{e} C_f^2) \right. \\ & + \sqrt{(\tilde{b} + \tilde{c}(\alpha + \beta) C_f + (\tilde{d}(\alpha + \beta) + \tilde{e}) C_f^2)(\tilde{b} + \tilde{c}(\alpha - \beta) C_f + (\tilde{d}(\alpha - \beta) + \tilde{e}) C_f^2)} \Big) \\ & /(\tilde{c} + \tilde{d} C_f) \sqrt{(\tilde{b} + \tilde{c}(\alpha + \beta) C_f + (\tilde{d}(\alpha + \beta) + \tilde{e}) C_f^2)(\tilde{b} + \tilde{c}(\alpha - \beta) C_f + (\tilde{d}(\alpha - \beta) + \tilde{e}) C_f^2)}. \end{aligned} \quad (5.73)$$

The sign of the term $\tilde{b} + \tilde{d} \tilde{e} C_f^2$ is determined by comparing each solution with an annual map for GPP created from its daily map. Using this approximation for the average GPP in the annual map (5.45), we note that we still have $C_f = 0$ as a fixed point.

It is not possible to solve the equation for the annual C_f analytically for C_f for this

version of the GPP, but when we substitute:

$$\tilde{a} = -206.75 \quad (5.74)$$

$$\tilde{b} = -3.14 \times 10^6 \quad (5.75)$$

$$\tilde{c} = -4205.30 \quad (5.76)$$

$$\tilde{d} = -4.84 \quad (5.77)$$

$$\tilde{e} = -123.63 \quad (5.78)$$

(all expressed in g C m⁻²)

$$\alpha = 8.4774, \quad (5.79)$$

$$\beta = -7.2819, \quad (5.80)$$

and introduce a fitting factor, $f = 1.139$, as we did before, our plot is very similar to that of the annual C_f using Approximation 1 of the simplified GPP with $\alpha = 8.4774$ as the mean of the irradiance, see Fig. 5.15.

Although we cannot solve the equation for the annual C_f analytically for C_f , we can solve it numerically for p_5 and plot for various values of C_f , which gives us the bifurcation figure, Fig. 5.16. We can calculate a line of limit points by fixing a value for $p_3(1 - p_2)$ and let Matlab calculate simultaneously $f(C_f) - C_f = 0$ and $f'(C_f) = 0$, see Fig. 5.17, where $f(C_f)$ denotes a function in C_f . As we can see, as for Approximation 1, the fit is very good.

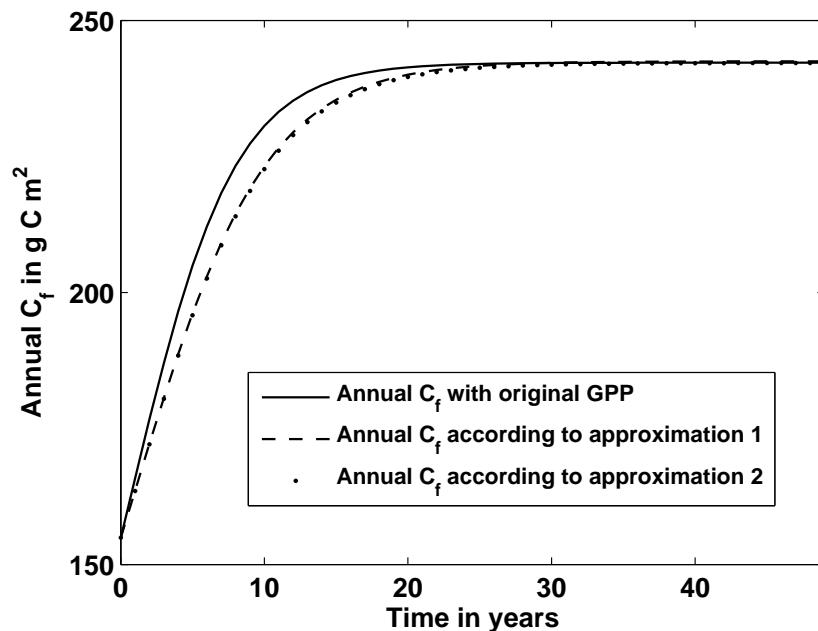


Figure 5.15: Annual map of C_f plotted using the original GPP, the annual map of C_f using Approximation 1 and the annual map of C_f using Approximation 2. A fitting factor was used with both variants of the simplified GPP to model the fixed point of the iterations.

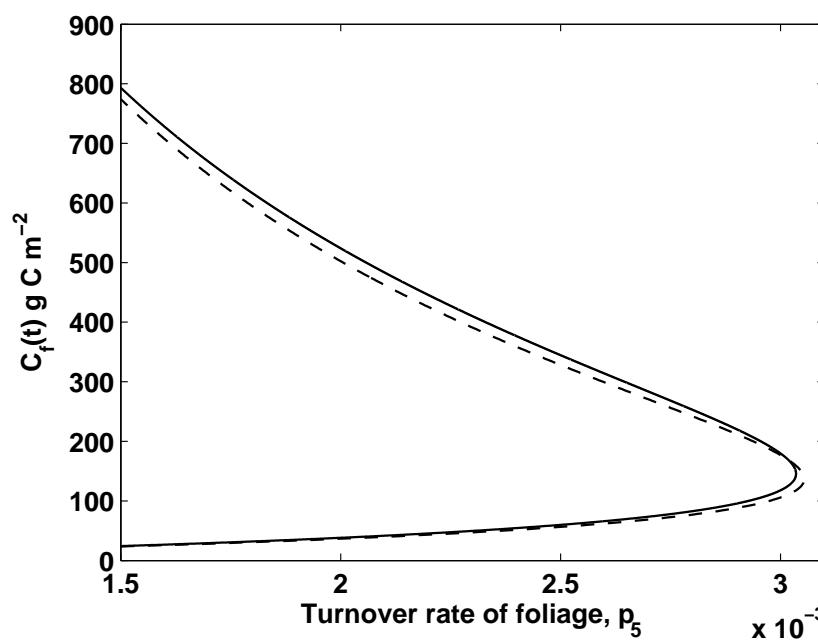


Figure 5.16: The limit point calculated using the original GPP (continuous line) and Approximation 2 for the averaged GPP (dashed line), with fitting factor of 1.139.

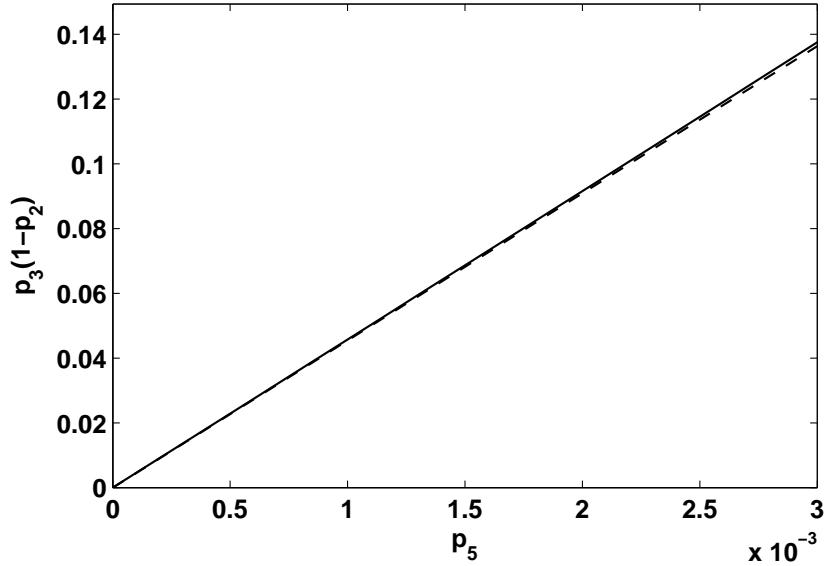


Figure 5.17: The line of limit points calculated using the original GPP (continuous line) and Approximation 2 for the averaged GPP (dashed line), with fitting factor of 1.139.

5.4 Linearising Around a Fixed Point to Simplify the Annual Map for C_f Further

By linearising around the fixed point we can simplify the annual map for C_f even more and create a linear map.

We express the equation for the annual map of C_f (5.45) as:

$$C_f^{(n+1)} = f(C_f^{(n)}). \quad (5.81)$$

Let the fixed point for this map be C_f^* and define:

$$C_f^{(n)} = C_f^* + \delta C_f^{(n)}. \quad (5.82)$$

Substituting this into (5.81) and expanding, retaining only linear terms in $f(C_f^{(n)})$ gives

the linear map:

$$\delta C_f^{(n+1)} = f'(C_f^*) \delta C_f^{(n)}. \quad (5.83)$$

Iterating this map n times starting with $\delta C_f^{(0)}$ we obtain:

$$\delta C_f^{(n)} = f'(C_f^*)^n \delta C_f^{(0)}. \quad (5.84)$$

For the annual map given by equation (5.45), with the fitting factor, we see that :

$$f(C_f^*) = (1 - p_5)^{365} C_f^* + \frac{f_1(1 - (1 - p_5)^{365})p_3(1 - p_2)\bar{a}C_f^{*2}}{p_5(\bar{b} + \bar{c}C_f^* + \bar{d}C_f^{*2})} \quad (5.85)$$

and so:

$$\begin{aligned} f'(C_f^*) &= (1 - p_5)^{365} + \frac{2f_1(1 - (1 - p_5)^{365})p_3(1 - p_2)\bar{a}C_f^*}{p_5(\bar{b} + \bar{c}C_f^* + \bar{d}C_f^{*2})} \\ &\quad - \frac{f_1(1 - (1 - p_5)^{365})p_3(1 - p_2)\bar{a}C_f^{*2}(\bar{c} + 2\bar{d}C_f^*)}{p_5(\bar{b} + \bar{c}C_f^* + \bar{d}C_f^{*2})^2}. \end{aligned} \quad (5.86)$$

For values of p_2 , p_3 , p_5 , a , b , c , d and f as used before and a value of $C_f^* = 242.23$, $f'(C_f^*) = 0.81$. Consequently:

$$C_f^{(n)} = 242.23 + 0.81^n \delta C_f^{(0)}, \quad (5.87)$$

where

$$\delta C_f^{(0)} = C_f^{(0)} - C_f^* = 242.23 - 154.92 = -87.32. \quad (5.88)$$

We can find evidence for why this works so well (see Fig. 5.18) by looking at the following ODE (which can be discretised and made into a map if necessary). The simplest form of

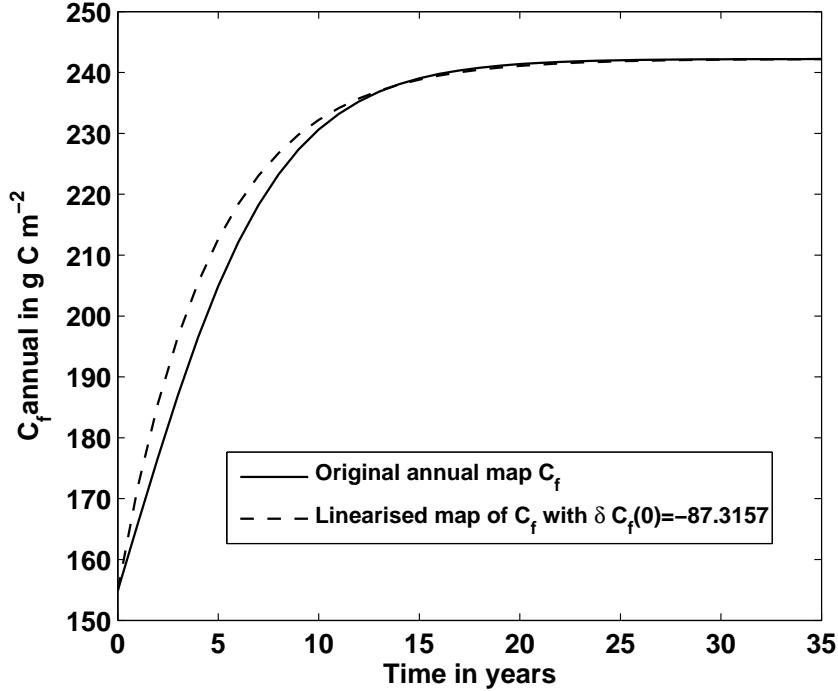


Figure 5.18: Plot showing that linearisation of the annual map for C_f works well, even from far away from the fixed point.

an ODE with a limit point structure similar to that of $C_f(t)$ is:

$$\frac{dx}{dt} = f(x) \quad (5.89)$$

$$f(x) = c_1\lambda - c_2x^2, \quad (5.90)$$

where $c_1, c_2 \in \mathbf{R}$ and λ is the bifurcation parameter.

The fixed points of this equation are:

$$x^* = \pm \sqrt{\frac{c_1}{c_2}\lambda} \text{ for } \frac{c_1\lambda}{c_2} > 0. \quad (5.91)$$

We can solve the ODE for x by separation of variables:

$$\int \frac{1}{c_1\lambda - c_2x^2} dx = \int dt \quad (5.92)$$

$$\frac{\operatorname{arctanh}\left(\frac{c_2x}{\sqrt{c_1c_2\lambda}}\right)}{\sqrt{c_1c_2\lambda}} = t + A \quad (5.93)$$

$$x(t) = \frac{\tanh((A+t)\sqrt{c_1c_2\lambda})\sqrt{c_1c_2\lambda}}{c_2}. \quad (5.94)$$

If we require $x(0) = 0$, so that the limit point sits on $(0, 0)$, see Fig. 5.19, then $A = 0$, so we get:

$$x(t) = \frac{\tanh(t\sqrt{c_1c_2\lambda})\sqrt{c_1c_2\lambda}}{c_2}. \quad (5.95)$$

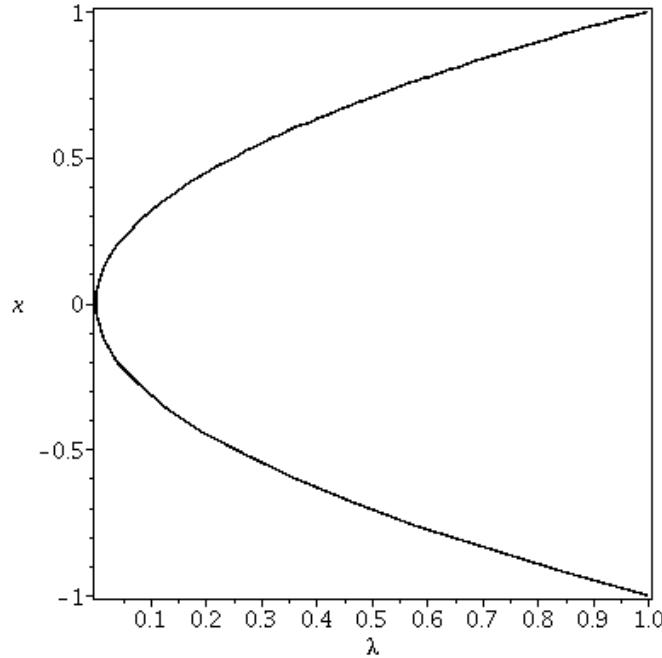


Figure 5.19: Plot of the normal form for a limit point, equation (5.90); values for C_1 and C_2 are both 1.

Now we linearise around the fixed point, x^* , using a Taylor expansion:

$$x(t) = x^* + \delta x(t). \quad (5.96)$$

Retaining only the linear terms in δx , we obtain:

$$\frac{d(\delta x)}{dt} = f_x(x^*)\delta x. \quad (5.97)$$

Using $f(x)$ defined by equation (5.90) gives:

$$f_x(x^*) = -2c_2\sqrt{\frac{c_1}{c_2}}\lambda = -2\sqrt{c_1 c_2 \lambda}, \quad (5.98)$$

$$\therefore \frac{d(\delta x)}{dt} = -2\sqrt{c_1 c_2 \lambda}\delta x. \quad (5.99)$$

We can solve equation (5.99) by separation of variables as before, which gives:

$$x(t) = x^* + Ae^{-2\sqrt{c_1 c_2 \lambda}t} \quad (5.100)$$

If we require $x(0) = 0$ then $A = -x^*$, so we get:

$$x(t) = x^* - x^*e^{-2\sqrt{c_1 c_2 \lambda}t} \quad (5.101)$$

Fig. 5.20 shows the plots for equations (5.95) and (5.101). We can see that when we vary the initial condition, the value for A changes and the fit for the linearised equation improves slightly.

If we compare the first quadrant of the plots in Fig. 5.20 with our annual map for C_f we can see the similarities. We have now simplified the annual map for C_f , which contained

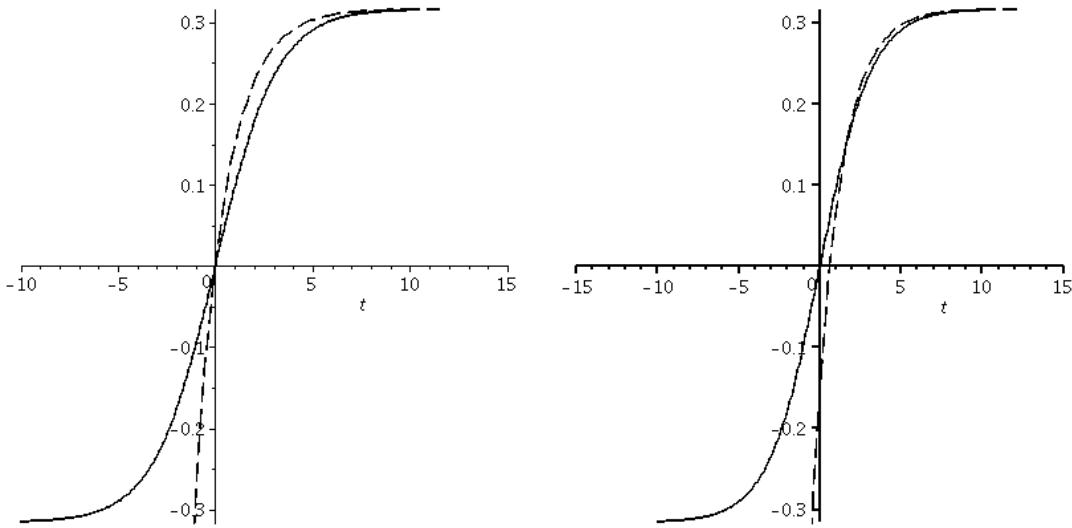


Figure 5.20: On the left is the plot for equations (5.95) (dashed line) and (5.101) (solid line). On the right is the plot for equations (5.95) (dashed line) and (5.101) (solid line), but changing the initial condition, so that A has a different value.

a complicated nonlinear function for the GPP, and can express C_f as a simple linear map.

5.5 Conclusion

In this chapter we have greatly simplified the GPP equation by applying the binomial theorem to the square root term in the C_i equation of the GPP. We then employed different techniques to simplify the GPP even more by averaging its time dependent functions, either by averaging all of them or by leaving the time dependence in the irradiance, but averaging the rest of the time dependent functions. Approximation 1 provided the simplest way of averaging GPP, by averaging all time dependent functions, including the irradiance. Approximation 2, where all the time dependent functions except for the irradiance were averaged, did not achieve a better result than Approximation 1, but was more complicated. Having found the average GPP allowed us to create an annual map for C_f which shows evolution of the foliar carbon pool over the years without the annual fluctuations. Using Approximation 1, which is the easiest approximation to work with, it

was possible to create the bifurcation graph and the graph with the line of limit points, shown in Section 4.3.2, without the use of continuation software. The results for these simplifications and approximations are qualitatively as well as quantitatively very good, especially when we look at Fig. 5.14, where we can see that with a very small fitting factor (1.016, which induces a change of approximately 1.6% in the annual GPP in the annual map for C_f), we can fit the approximated line of limit points almost exactly onto the original line of limit points.

The fitting factors we used are close to 1 and would be different for every forest. The values we have used here are appropriate for the Loobos data. Qualitatively, if the fitting factors were not there, the behaviour of the simplified annual maps would still be similar to that of the original map for C_f . Fig. 5.13 shows that the fitting factor does not change the position of the limit point with respect to C_f , although it makes a significant difference ($\approx 5\%$) to the annual values of C_f in Fig. 5.11. It is expected that fitting factors will be close to 1 and work in general, as they are applied to an approximation of the carbon foliar pool and the GPP of the forests itself.

Section 5.4 showed that we can simplify the C_f even more and express it as a linear map. We explained that this works because of the limit point structure in the annual map of C_f , which can be compared with the normal form of a limit point bifurcation. Both maps show qualitatively similar behaviour. As the graph of the linear map in Fig. 5.18 approximately matches the original annual map of C_f , we are able to calculate an approximate annual map of C_f over time. Note though, the linear map only describes how solutions converge to the fixed point; it is not possible, for example, to recover the limit point structure.

Chapter 6

Experimenting with Shocks

6.1 Introduction

The material in this section is meant as a flavour of what is possible in terms of experimenting with shocks to a model such as DALEC EV. We are using the simplified annual map for C_f from Chapter 5, which uses periodic (averaged) drivers. We do recognise that in reality drivers are not quite periodic and in the case of shocks annual drivers may show a marked difference between one year and another. However, it is also possible to experiment with shocks to the model using continuation software and non-periodic drivers. Using the simplified model is merely a quick and simple way to show how a shock, such as a change in moisture, can affect the position of the tipping point.

In this chapter we look at the effects of varying the maximum soil-leaf water potential difference, ψ , which is seen as a constant in the model, on the tipping point. In DALEC the value of this parameter is set between -0.5 MPa, reflecting a highly drought-stressed soil and -4.0 MPa, reflecting a wet soil with drought-resistant vegetation. In the calculations $|\psi|$ is used.

There is an extension to DALEC available, called DALEC Water [46], in which the maximum soil-leaf water potential difference is not a constant, but depends on a range of conditions in the soil layers, effectively creating a ψ driver. However, we decided to vary the ψ constant in the DALEC EV model, which is sufficient for our purposes.

6.2 The Experiment

We define $\tilde{\psi} = |\psi|^{a_{10}}$ for notational purposes only. As the three different variants of the annual map for C_f are very similar we choose the simplest one to work with, see equation (5.45), and, written in terms of $\tilde{\psi}$, the equation for the annual map of C_f is as follows:

$$C_f^{(n+1)} = (1-p_5)^{365} C_f^{(n)} + \frac{(1 - (1 - p_5)^{365})}{p_5} p_3(1-p_2)f_1 \frac{\hat{a}C_f^{(n)2}}{\hat{b} + \hat{c}C_f^{(n)} + (\frac{\hat{d}_1}{\tilde{\psi}} + \hat{d}_2)C_f^{(n)2}}, \quad (6.1)$$

where

$$\hat{a} = a_7 \bar{I} p_{11} N \exp(a_8 \bar{T})(a_3 - C_a)(a_2 \bar{s} + a_5) = 1.66 \times 10^2 \quad (6.2)$$

$$\hat{b} = p_{11} N \exp(a_8 \bar{T}_m)(a_9 lma^2(a_3 - C_a) = -2.09 \times 10^6 \quad (6.3)$$

$$\hat{c} = -a_7 \bar{I}(C_a - q) = -3.57 \times 10^4 \quad (6.4)$$

$$\hat{d}_1 = -p_{11} N \exp(a_8 \bar{T}_m)(0.5 \bar{T}_r + a_6 R_{tot})(a_7 \bar{I}) = -69.98 \quad (6.5)$$

$$\hat{d}_2 = -p_{11} N \exp(a_8 \bar{T}_m)(C_a - a_3) = -1.17 \times 10^2 \quad (6.6)$$

(all expressed in g C m⁻²)

where \bar{T}_m is the mean maximum temperature and \bar{T}_r is the mean temperature range in °C. These values have been calculated using the values for a , b , c and d in Approximation 1 in Section 5.3.1.

To calculate the non-zero fixed point solutions for the map, we can write the above

equation as a quadratic in C_f :

$$p_5 \hat{b} + (p_5 \hat{c} - p_3(1 - p_2)f_1 \hat{a})C_f + p_5 \left(\frac{\hat{d}_1}{\tilde{\psi}} + \hat{d}_2 \right) C_f^2 = 0, \quad (6.7)$$

and find solutions:

$$C_f = \frac{-p_5 \hat{c} + p_3(1 - p_2)f_1 \hat{a} \pm \sqrt{(p_5 \hat{c} - p_3(1 - p_2)f_1 \hat{a})^2 - 4p_5^2 \hat{b} \left(\frac{\hat{d}_1}{\tilde{\psi}} + \hat{d}_2 \right)}}{2p_5 \left(\frac{\hat{d}_1}{\tilde{\psi}} + \hat{d}_2 \right)}. \quad (6.8)$$

The solutions vary with the value we choose for ψ . In Fig. 6.1 we can see effect of varying ψ . The dotted lines are the result of varying ψ using the simplified model, the continuous lines in the same colour show the result of varying ψ using the original model. As before, the bottom half of the curve represents the unstable solutions and the top half of the curve represents the stable solutions. We can see that the results of varying ψ lie very close to each other.

The black line on the graph shows the limiting curve as $\psi \rightarrow \infty$. Multiplying equation (6.7) by $\tilde{\psi}$ and letting $|\psi| \rightarrow 0$, shows that the only solution in this case is $C_f = 0$ when $p_5 \neq 0$ or C_f can take any value when $p_5 = 0$. We can see from Fig. 6.1 that the latter is the limiting case when $|\psi| \rightarrow 0$.

As in Section 5.3.1, where we found a path of limit points for variables p_5 and p_{23} from the discriminant of equation (5.51), we can solve the discriminant in equation (6.8) for $\tilde{\psi}$ and plot a path of limit points for p_5 and $|\psi|$, as shown in Fig. 6.2. Similarly, Fig. 6.3 shows the path of limit points for C_f and p_5 .

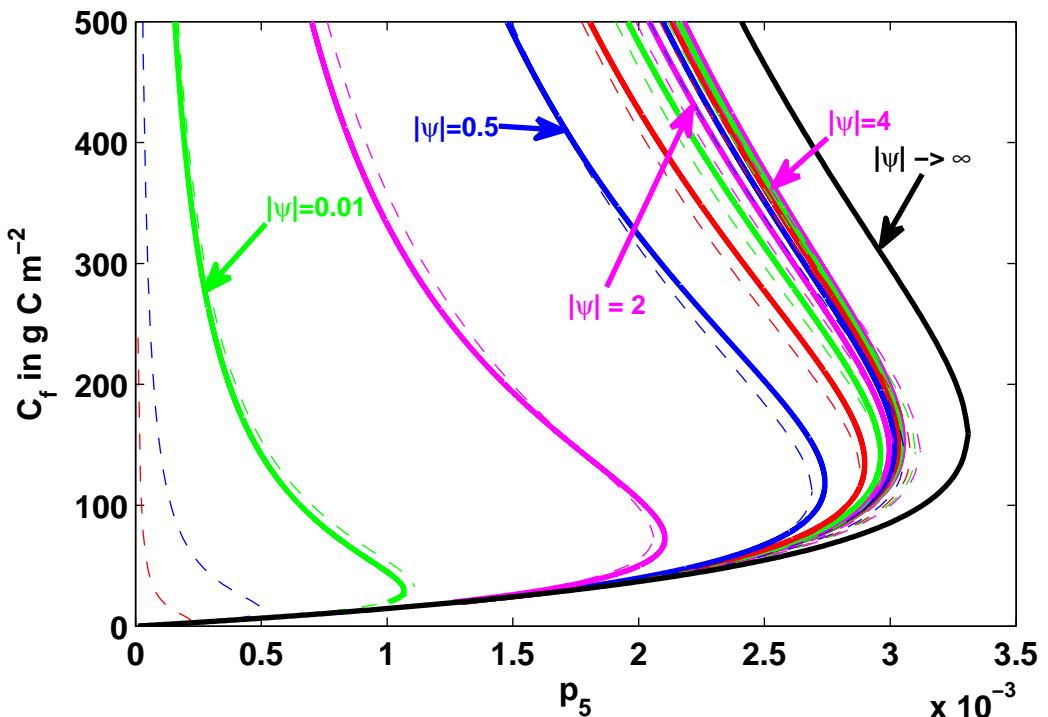


Figure 6.1: Varying ψ using the original model (continuous lines) and using the simplified model (dotted lines). The top half of the curve represent the stable solutions, whereas the bottom half of the curve represents the unstable solutions. The black curve shows the limiting curve as $|\psi| \rightarrow \infty$. Very wet soil would not directly be a problem for forest survival as trees are limited in how much water they can take up.

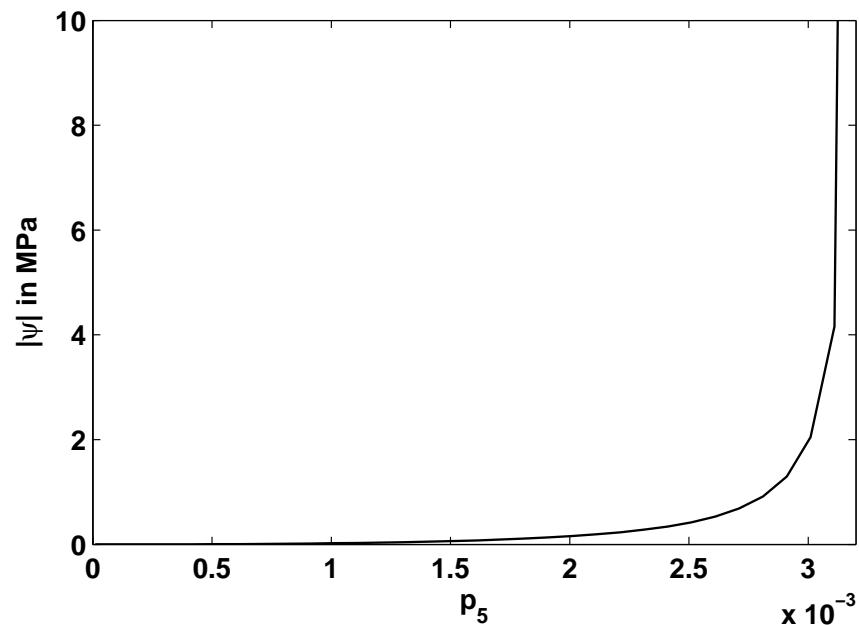


Figure 6.2: Path of limit points for $|\psi|$ and p_5 .

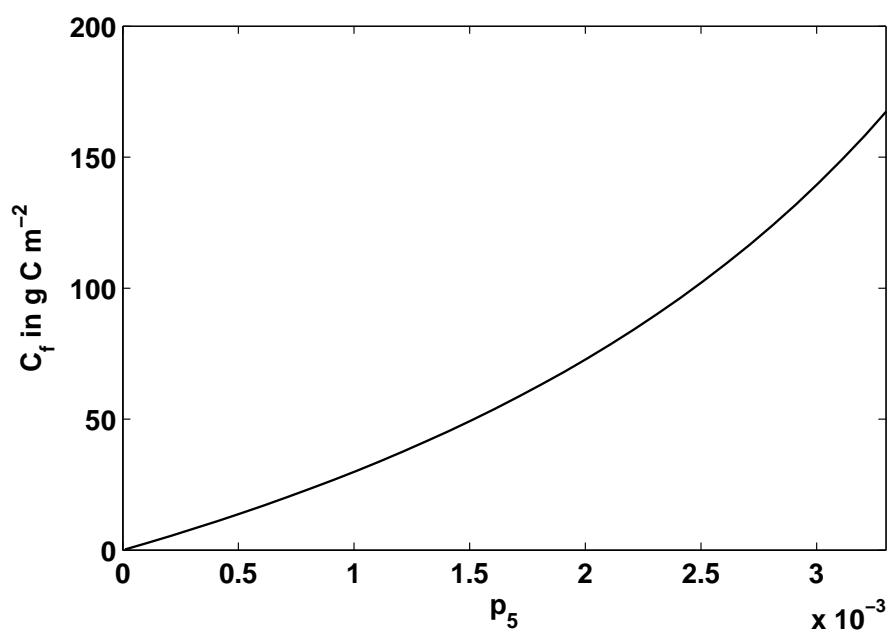


Figure 6.3: Path of limit points for p_5 and C_f .

Solving the discriminant equal to zero in equation (6.8) for $\tilde{\psi}$ gives:

$$\tilde{\psi} = \frac{4p_5^2 \hat{b} \hat{d}_1}{(p_5 \hat{c} - p_3(1-p_2)f_1 \hat{a})^2 - 4p_5^2}. \quad (6.9)$$

To find the line of limit points for C_f and p_5 , we substitute the values for $\tilde{\psi}$ and p_5 into equation (6.8), which will give us the appropriate value for C_f .

Corresponding values of C_f are:

$$C_f = \frac{-p_5 \hat{c} + p_3(1-p_2)f_1 \hat{a}}{2p_5(\frac{\hat{d}_1}{\tilde{\psi}} + \hat{d}_2)} \quad (6.10)$$

The asymptote in Fig. 6.2 occurs when the denominator of equation (6.9) is zero.

6.2.1 What Happens When ψ Changes?

Looking at Fig. 6.1 we can imagine what happens when ψ changes in the forest, for example when there is a drought.

Consider the case when $p_5 = 0.0028$, as it has been throughout this report and $|\psi| = 2$. In Fig. 6.4 we suppose that the forest is the dot on the continuous line representing $|\psi| = 2$ (this is where the forest has its fixed point). When ψ changes from $|\psi| = 2$ to $|\psi| = 0.5$, represented by the dashed line, the dot starts dropping downwards, as the curve for $|\psi| = 2$ disappears; there are no fixed points in that region anymore. It is possible for ψ to recover and if that happens fast enough the forest can find another stable fixed point before it has dropped down too far. We can also see in Fig. 6.4 that the value to which ψ needs to recover to save the forest is $|\psi| = 0.65821$; we call this the ‘critical value’ of ψ for the value of $p_5 = 0.0028$.

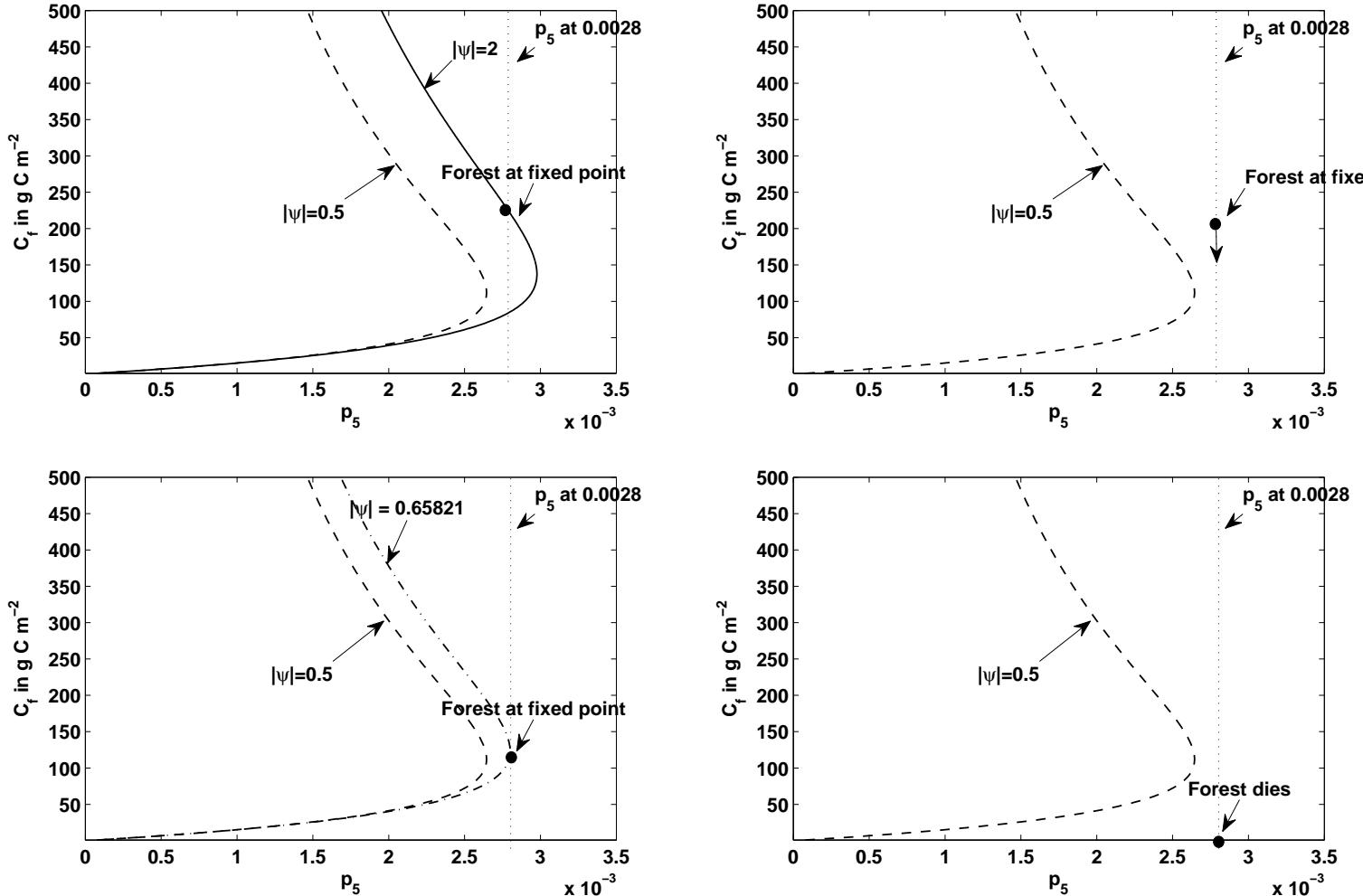


Figure 6.4: On the top left hand side the forest is on a fixed point of the continuous curve where $|\psi| = 2$. The value of $|\psi|$ changes to 0.5. In the plot on the top right hand side we can see the continuous curve is gone and the forest's fixed point changes, represented by the dot falling down. On the bottom left hand side: if ψ recovers enough to at least the critical value for $p_5 = 0.0028$, the forest will be able to settle into a new stable fixed point. On the bottom right hand side: if ψ does not recover or recovers too late, the dot will keep falling down towards the stable fixed point $C_f = 0$ and the forest will die.

However, if C_f drops too far before ψ recovers, then the forest will die. Fig. 6.5 shows, for $p_5 = 0.0028$ and for $|\psi| = 2$ changing to $|\psi| = 0.5$, how long it takes for C_f to reach the critical value from which it cannot recover, if ψ does not recover. We can calculate the critical value of ψ for any p_5 by solving the discriminant of the solutions to the annual map for C_f , equation (6.9), equal to zero for ψ , whilst substituting the appropriate value of p_5 . For each value of p_5 , there will be a critical value of ψ which can be calculated in this way.

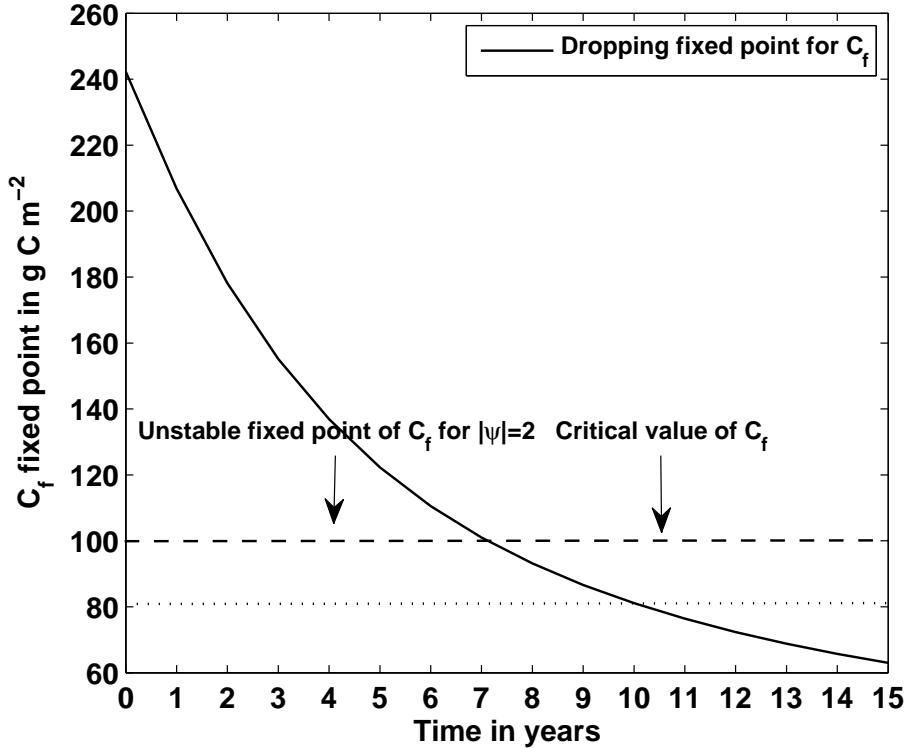


Figure 6.5: Here the critical value for C_f when $|\psi|$ changes from 2 to 0.5 is shown as well as the unstable fixed point for C_f .

6.3 Conclusion

We have used the simplified annual map for C_f to create a simple mechanism for experimenting with soil moisture shocks, varying the parameter for moisture, which is fixed as a constant in the DALEC EV model. We have assumed in this experiment that the value of p_5 , which is the daily turnover of foliage, stays constant, as well as all other parameters. This may not be realistic, but it is simplest case, illustrating the effect of shocks. We found that in the DALEC EV model the tipping point moves when the soil moisture parameter value changes. If the soil moisture parameter becomes smaller, representing drier conditions, and does not recover in time, a forest is not able to adjust and dies. We note however, that if parameter p_5 also changes and for example becomes larger (more needles are dropped daily) when ψ becomes smaller, the forest may have a decreased chance to find a new stable, sustainable fixed point, as the change will occur in two directions.

We showed that it will take several years for a forest to reach the critical point from which it cannot recover after the moisture parameter changes from $|\psi| = 2$ to $|\psi| = 0.5$ for the parameter values we are working with, and therefore the forest in this model is quite robust with respect to drought, meaning that there is plenty of chance for the moisture parameter to recover. For a moisture parameter change in the other direction, the model indicates that the forest will not die, as it will remain between the unstable and stable lines of the limit point curve and therefore will find a new stable fixed point.

Chapter 7

Constraining of Parameters

7.1 Introduction

Data assimilation can be used to attempt to constrain parameters, which was done in the REFLEX project [13], but an examination of the model equations should also give useful information on which parameters are likely to be constrainable and which are not. The constrainability of parameters is very much dependent on the data that is fed into the model. The REFLEX participants were given leaf area index (LAI), which is directly proportional to C_f ($LAI = C_f/lma$) and net ecosystem exchange (NEE), as well as climate and location data and initial conditions for C_{som} and C_w . LAI is a measurement that can be established both in the field and from satellite data. NEE is measured using flux towers and plays an important role in establishing carbon stocks globally. It is both input and output for any global dynamic vegetation model (GDVM) as well as for any general circulation model (GCM), which, similar to DALEC, can employ data assimilation methods to produce better predictions short term and long term.

Our aim in this chapter is to try to find out from the model equations alone which

parameters can be expected to be reasonably constrained from NEE and LAI. In earlier chapters, since LAI is directly proportional to C_f , we have already discovered that we cannot expect to constrain any parameters other than p_2 , p_3 , p_5 and p_{11} from LAI alone. A numerical sensitivity analysis should confirm this. The numerical sensitivity analyses for both C_f and NEE are performed in Section 7.2 and will give us information about the parameter sensitivity of a specific location (in this case the Loobos which we have used throughout this project) whereas the analytical sensitivity analysis of NEE, performed in Section 7.3, will allow us to make a more general statement about the parameter sensitivity of NEE at any location. To this end, in Section 7.3.1 we will examine the different contributions of the pools to each other and the NEE and in Section 7.3.2 we expand and simplify the NEE equation in various ways, depending on the time period we are considering. In Section 7.3.2 we draw comparisons between our results and the results from the REFLEX study and we also find that the length of time for which the model runs has an impact on the constrainability of parameters. An interesting conclusion is that, by combining work done in Chapter 5 and the simplification of the annual NEE in Section 7.3.2, we find a very simple carbon model that could be used to estimate NEE during a certain time period.

It must be noted that some of the results in this chapter, in particular numerical results, are dependent on the initial conditions that we used for our calculations.

7.2 Numerical Sensitivity Analysis for C_f and NEE

In this section we perform a numerical sensitivity analysis, measuring sensitivity of the periodic solution for C_f and NEE to parameter changes and also measuring sensitivity of C_f and total NEE over one year to parameter changes for fixed initial conditions. For this purpose we use a central difference approximation and define the sensitivity of C_f and NEE to parameter p_i as:

$$\frac{\partial C_f}{\partial p_i} = \frac{C_f(p_i + p_i h) - C_f(p_i - p_i h)}{2p_i h} \quad (7.1)$$

and

$$\frac{\partial NEE}{\partial p_i} = \frac{NEE(p_i + p_i h) - NEE(p_i - p_i h)}{2p_i h}, \quad (7.2)$$

respectively where $h = 1 \times 10^{-4}$. The results are displayed in Tables 7.1 and 7.2. These values give an indication of the change in C_f and NEE that would result from a given small absolute change in p_i .

The absolute sensitivity of C_f and NEE with respect to a percentage change in the parameter, p_i , is calculated as follows:

$$\frac{\partial C_f}{\partial p_i} \times \frac{p_i}{100} \quad (7.3)$$

and

$$\frac{\partial NEE}{\partial p_i} \times \frac{p_i}{100}. \quad (7.4)$$

The absolute sensitivity of C_f and NEE are shown in Tables 7.1 and 7.2. These values give an indication of the change in C_f that would result from a given small percentage change in p_i . We have not just calculated the sensitivity of NEE to the p -parameters, but also to other parameters, such as lma and C_a . The parameters a , b and c are the parameters used in the Fourier approximation of the irradiance.

7.2.1 Table 7.1

This table shows the change in the fixed point for C_f after each parameter (p_i) changes by 1%. To determine sensitivity of NEE to the parameters we find the fixed points for all the other carbon pools after the parameter change and then sum NEE over one year.

As we can see, C_f is very sensitive to parameters p_2 , p_3 , p_5 and p_{11} . This shows clearly the decoupling of the foliar carbon pool from the other four carbon pools. The foliar carbon pool is also quite sensitive to N (foliar nitrate). When there is more nitrate available, foliar carbon increases. The sensitivity of C_f to lma is about -6.7, which makes sense, as $LAI = C_f/lma$. The larger the lma , the smaller the LAI, which will have an effect on the GPP by causing the foliar carbon pool to decrease. The parameter a refers to the mean of the Fourier approximation for the irradiance. Earlier we concluded that the foliar carbon pool is sensitive to irradiance, which is why we needed data assimilation to approximate the irradiance by a Fourier approximation.

A positive value for the sensitivity means that when the parameter is increased, the fixed point solution for C_f increases and when the parameter is decreased the fixed point solution for C_f decreases. If the values are negative however, it means that when the parameter increases the fixed point decreases and vice-versa. Therefore, an increase of parameter p_2 , which is the fraction of GPP respired by autotrophic respiration, causes a decrease in the foliar carbon pool. This makes sense, because if more carbon has been respired by the tree back into the atmosphere there is less left in the foliar carbon pool. However parameter p_3 , the fraction of NPP allocated to foliage, which shows a positive sensitivity, causes an increase in the foliar carbon when it is increased and a decrease when it is decreased. This also makes sense, because if more of the NPP is taken up, there is more carbon in the foliage.

It is clear from Table 7.1 that the foliar carbon pool decouples from the other four pools, as C_f is only sensitive to its own parameters plus p_{11} which is part $p(t)$ (8.3), needed to calculate GPP. Recall the smoothing of the temperature and irradiance data in Chapter 4.3.2. We found that it was straightforward to approximate the temperature data by a Fourier series, showing that C_f is not very sensitive to temperature fluctuation, but that C_f is quite sensitive to irradiance, which is why we applied data assimilation to C_f with regard to irradiance. In Table 7.1 we can see this confirmed, C_f is a lot less sensitive

p_i	C_f at fixed points		NEE at fixed points. NEE totalled over year		
	p_i values	$\frac{\partial C_f}{\partial p_i}$	$\frac{\partial C_f}{\partial p_i} \frac{p_i}{100}$	$\frac{\partial NEE}{\partial p_i}$	$\frac{\partial NEE}{\partial p_i} \frac{p_i}{100}$
P_1	$4.41E-06$	0.00	0.00	-0.26	$-1.15E-08$
P_2	0.52	-1899.20	-9.88	$-1.17E-05$	$-6.08E-08$
P_3	0.29	3143.50	9.12	$1.31E-05$	$3.79E-08$
P_4	0.41	0.00	0	$1.30E-06$	$5.32E-09$
P_5	0.0028	$-3.27E+05$	-9.17	-0.002	$-4.95E-08$
P_6	$2.06E-06$	0.00	0.00	3.28	$6.75E-08$
P_7	0.003	0.00	0.00	$-4.97E-05$	$-1.49E-09$
P_8	0.02	0.00	0.00	$-3.12E-05$	$-6.25E-09$
P_9	$2.65E-06$	0.00	0.00	0.64	$1.71E-08$
P_{10}	0.0693	0.00	0.00	$-3.26E-05$	$-2.26E-08$
P_{11}	7.40	55.85	4.13	$5.95E-08$	$4.40E-09$
$ \psi $	2.00	29.34	0.59	$1.66E-06$	$3.32E-08$
R_{tot}	1.00	-6.66	-0.07	$-4.76E-06$	$-4.76E-08$
N	4.00	103.33	4.13	$1.10E-07$	$4.40E-09$
lat	0.9071	155.99	1.42	$-8.81E-06$	$-8.00E-08$
lma	110.00	-6.09	-6.69	$4.54E-08$	$4.99E-08$
Ca	380.00	0.68	2.58	$4.85E-09$	$1.84E-08$
a	8.4774	66.39	5.63	$1.58E-07$	$1.34E-08$
b	-7.2819	19.22	-1.39	$1.27E-07$	$9.27E-09$
c	-0.5789	-2.01	0.01	$-5.51E-06$	$-3.19E-08$
\bar{T}_{max}	13.407	4.59	0.62	$-1.01E-07$	$-1.36E-08$
\bar{T}_{range}	6.59	-4.21	-0.28	$3.47E-07$	$2.29E-08$

Table 7.1: *Sensitivity analysis results for C_f and NEE where every parameter change results in different fixed points for the five carbon pools. The initial fixed points for the different pools to 2 decimal places are: $C_f = 242.20$, $C_r = 226.55$, $C_w = 475598.93$, $C_{lit} = 80.35$ and $C_{som} = 344203.18$. Units for C_f and NEE are g C mr.*

to \bar{T}_{max} (which is the mean of the Fourier series approximation for the maximum temperature) and \bar{T}_{range} (the mean of the Fourier series approximation for the temperature range) than to the mean irradiance (parameter a).

The results show that there is no sensitivity of the NEE to changes in the parameters.

We now show that this is because the NEE is always zero at a fixed point of the carbon pools.

Lemma 7.2.1. *The NEE summed over one year at a fixed point of the carbon pools is zero.*

Proof. We define the total carbon in the model, C_{tot} , by:

$$C_{tot}(t) = C_f(t) + C_r(t) + C_w(t) + C_{lit}(t) + C_{som}(t).$$

Adding the five equations for the carbon pools, equations (3.1) until (3.5), together, gives:

$$C_{tot}(t+1) = C_{tot}(t) - NEE(t),$$

where

$$NEE(t) = p_8 T(t) C_{lit}(t) + p_9 T(t) C_{som}(t) - (p_2 - 1) GPP(C_f(t), t) \quad (7.5)$$

So

$$C_{tot}(t+1) - C_{tot}(t) = -NEE(t). \quad (7.6)$$

Summing the left hand side of equation (7.6) over one year gives:

$$\begin{aligned} \sum_{t=0}^{364} (C_{tot}(t+1) - C_{tot}(t)) &= C_{tot}(365) - C_{tot}(364) + C_{tot}(364) - C_{tot}(363) + \\ &\quad \dots + C_{tot}(1) - C_{tot}(0), \\ &= C_{tot}(365) - C_{tot}(0), \end{aligned}$$

hence

$$C_{tot}(365) - C_{tot}(0) = - \sum_{t=0}^{364} NEE(t). \quad (7.7)$$

If all pools are periodic, then $C_{tot}(365) = C_{tot}(0)$ and so $\sum_{t=0}^{364} NEE(t) = 0$. \square

For $\sum_{t=0}^{364} NEE(t) > 0$, the forest is a source of carbon and if $\sum_{t=0}^{364} NEE(t) < 0$, the forest is a sink. For a growing forest, depending on the initial conditions, each of the carbon pools will generally increase over a period of a year, in which case $C_{tot}(365) > C_{tot}(0)$ and so

p_i	C_f from fixed initial condition At day 365		NEE from fixed initial conditions NEE totalled over year		
	p_i values	$\frac{\partial C_f}{\partial p_i}$	$\frac{\partial C_f}{\partial p_i} \frac{p_i}{100}$	$\frac{\partial NEE}{\partial p_i}$	$\frac{\partial NEE}{\partial p_i} \frac{p_i}{100}$
P_1	$4.41E-06$	0.00	0.00	-14663.00	$-6.47E-04$
P_2	0.52	-352.14	-1.83	1584.80	8.24
P_3	0.29	582.86	1.69	-845.48	-2.45
P_4	0.41	0.00	0.00	117.01	0.48
P_5	0.0028	-56863.00	-1.59	$1.22E+05$	3.43
P_6	$2.06E-06$	0.00	0.00	1876.20	$3.87E-05$
P_7	0.003	0.00	0.00	31778.00	0.95
P_8	0.02	0.00	0.00	2137.30	0.43
P_9	$2.65E-06$	0.00	0.00	$4.31E+06$	0.11
P_{10}	0.0693	0.00	0.00	496.82	0.34
P_{11}	7.40	11.27	0.83	-50.85	-3.76
$ \psi $	2.00	3.26	0.07	-14.62	-0.29
R_{tot}	1.00	-0.74	-0.01	3.32	0.03
N	4.00	20.85	0.83	-94.07	-3.76
lat	0.9071	29.09	0.26	-135.14	-1.23
lma	110.00	-1.48	-1.63	6.71	7.38
C_a	380.00	0.11	0.43	-0.51	-1.93
a	8.4774	12.21	1.04	-54.48	-4.62
b	-7.2819	3.57	-0.26	-15.40	1.12
c	-0.5789	0.18	-0.001	-4.10	0.02
\bar{T}_{max}	13.407	0.93	0.12	-0.43	-0.06
\bar{T}_{range}	6.59	-0.05	-0.004	0.21	0.01

Table 7.2: C_f and NEE sensitivity for fixed initial conditions for all pools: $C_f(0) = 150$, $C_r(0) = 160$, $C_w(0) = 9200$, $C_{som}(0) = 11000$ and $C_{lit}(0) = 60$. Total NEE after one year is -235.48. Units for C_f and NEE are $g C m^2$.

$\sum_{t=0}^{364} NEE(t) < 0$. Thus, as expected, a growing forest is a carbon sink. However, a newly planted forest could also be a source for a period of time, as the carbon in the soil may be at a higher level than fits the growing forest and therefore will initially decrease.

7.2.2 Table 7.2

Table 7.2 shows the sensitivity of C_f at day 365 and NEE summed over one year with respect to the parameters when measured from initial conditions. The results are much lower for this table than for Table 7.1, although for C_f the same parameters as for Table

7.1 have large sensitivity values, namely p_2 , p_3 , p_5 , p_{11} , N , lma and a . The reason for the different outcomes is because we are measuring different things. Table 7.2 measures sensitivity over 1 year of a growing forest, whereas Table 7.1 measures the change in the fixed points, the forest being in an equilibrium.

The NEE is most sensitive to p_2 and lma . The coefficient for p_2 is positive: when p_2 increases, the fraction of GPP respired increases, which causes the NEE to increase as there is more carbon coming into the atmosphere. As said earlier, the lma is closely related to C_f : the larger lma the smaller the LAI, which means less carbon taken up by any of the pools, which in turn means an increase in NEE. It is interesting to note that NEE seems least sensitive to p_6 , the daily turnover rate of wood. There are three very small parameters in DALEC evergreen, p_1 , p_6 and p_9 ; the sensitivity for NEE to p_1 is very small too, of order 10^{-4} , but NEE's sensitivity to p_9 is of the order 10^{-1} , which is a lot higher than the sensitivity to p_1 and p_6 .

The sensitivity coefficients for C_f and NEE always have opposite signs (when non-zero). Carbon either goes into the tree or into the atmosphere and so they will be negatively correlated.

Looking at C_a , the atmospheric carbon concentration, we can see that as C_a increases, NEE decreases, meaning that if there is more atmospheric carbon, the forest acts as a stronger sink or weaker source. NEE is not very sensitive to the mean of the maximum temperature or the mean of the temperature range, confirming our earlier findings, which we discussed under the section for Table 7.1. NEE is sensitive to the a and b parameters, which are the mean and the amplitude of the cosine wave for the Fourier approximation of the irradiance.

7.3 Analytical Sensitivity Analysis for NEE

In the following sections we aim to get a clearer idea of the different contributions of the carbon pools to each other and to the NEE. To this end we rescale the pools in Section 7.3.1, both from an initial conditions perspective and their fixed points. We then expand and simplify the NEE in Section 7.3.2 in three different ways to find out what role the carbon pools and the different parameters play. This then gives further insight into the sensitivity of the NEE to the parameters.

7.3.1 Contributions of Carbon Pools

Before we expand and simplify the NEE equation, in order to analytically examine the parameter sensitivity, we first get a clearer idea of the size of the different contributions of the carbon pools to each other and the NEE. As the initial conditions and fixed points for the different pools vary widely, which makes it difficult to see the size of the different contributions, we rescale the pools, both from a initial conditions perspective and their fixed points. By rescaling, each pool fluctuates around 1 and the contributions from the different pools are now noticeable by looking at the coefficients only. The GPP stays of $\mathcal{O}(1)$.

Recall from Section 7.2 the equation for NEE:

$$NEE(t) = p_8 C_{lit}(t)T(t) + p_9 C_{som}(t)T(t) - (1 - p_2)GPP(C_f(t), t) \quad (7.8)$$

Remember, the equations for the five carbon pools are:

$$\begin{aligned}
 C_f(t+1) &= (1-p_5)C_f(t) + p_3(1-p_2)GPP(C_f(t), t), \\
 C_r(t+1) &= (1-p_7)C_r(t) + p_4(1-p_2)(1-p_3)GPP(C_f(t), t), \\
 C_w(t+1) &= (1-p_6)C_w(t) + (1-p_4)(1-p_2)(1-p_3)GPP(C_f(t), t), \\
 C_{lit}(t+1) &= (1-(p_1+p_8)T(t))C_{lit}(t) + p_5C_f(t) + p_7C_r(t), \\
 C_{som}(t+1) &= (1-p_9T(t))C_{som}(t) + p_6C_w(t) + p_1T(t)C_{lit}(t).
 \end{aligned}$$

To rescale a function $x(t)$ we define $\tilde{x}(t) = \frac{x(t)}{x^*}$, where x^* can be either an initial condition or a fixed point. We can apply this to the five carbon pools:

$$\tilde{C}_f(t) = \frac{C_f(t)}{C_f^*}, \quad (7.9)$$

$$\tilde{C}_r(t) = \frac{C_r(t)}{C_r^*}, \quad (7.10)$$

$$\tilde{C}_w(t) = \frac{C_w(t)}{C_w^*}, \quad (7.11)$$

$$\tilde{C}_{lit}(t) = \frac{C_{lit}(t)}{C_{lit}^*}, \quad (7.12)$$

$$\tilde{C}_{som}(t) = \frac{C_{som}(t)}{C_{som}^*}. \quad (7.13)$$

This results in the following five equations:

$$\tilde{C}_f(t+1) = (1-p_5)\tilde{C}_f(t) + \frac{p_3(1-p_2)}{C_f^*}GPP\left(\tilde{C}_f(t)C_f^*, t\right), \quad (7.14)$$

$$\tilde{C}_r(t+1) = (1-p_7)\tilde{C}_r(t) + \frac{p_4(1-p_2)(1-p_3)}{C_r^*}GPP\left(\tilde{C}_f(t)C_f^*, t\right), \quad (7.15)$$

$$\tilde{C}_w(t+1) = (1-p_6)\tilde{C}_w(t) + \frac{(1-p_2)(1-p_3)(1-p_4)}{C_w^*}GPP\left(\tilde{C}_f(t)C_f^*, t\right), \quad (7.16)$$

$$\tilde{C}_{lit}(t+1) = (1-(p_1+p_8)T(t))\tilde{C}_{lit}(t) + \frac{p_5C_f^*}{C_{lit}^*}\tilde{C}_f(t) + \frac{p_7C_r^*}{C_{lit}^*}\tilde{C}_r(t), \quad (7.17)$$

$$\tilde{C}_{som}(t+1) = (1-p_9T(t))\tilde{C}_{som}(t) + \frac{p_6C_w^*}{C_{som}^*}\tilde{C}_w(t) + \frac{p_1C_{lit}^*}{C_{som}^*}T(t)\tilde{C}_{lit}(t). \quad (7.18)$$

Equation (7.8), now becomes:

$$NEE(t) = p_8 C_{lit}^* T(t) \tilde{C}_{lit}(t) + p_9 C_{som}^* T(t) \tilde{C}_{som}(t) - (1 - p_2) GPP(\tilde{C}_f(t) C_f^*, t). \quad (7.19)$$

We consider two different scalings, one using the initial condition, $C_f^* = 150$ and one with $C_f^* = 242.2016$ which is the value of the fixed point. Similarly for the other carbon pools. Figs 7.1 and 7.2 show the scaled pools for their fixed points, which are $C_r = 225.7127$, $C_{lit} = 80.3470$, $C_w = 474526.211$ and $C_{som} = 344200$. It seems as if \tilde{C}_{som} and \tilde{C}_w are constants in Fig. 7.1, but they do vary slightly, as can be seen in Fig. 7.2.

In fact, the peak-to-peak amplitudes for the five carbon pools for the fixed point version are:

$$\begin{aligned}\tilde{C}_{f_A} &= 0.2544, \\ \tilde{C}_{r_A} &= 0.2758, \\ \tilde{C}_{w_A} &= 0.0002, \\ \tilde{C}_{lit_A} &= 0.5780, \\ \tilde{C}_{som_A} &= 0.0002\end{aligned}$$

The coefficients of the scaled pools (\tilde{C}_f , \tilde{C}_r , etc.) in these equations are of particular interest. Table 7.3 shows the values of these coefficients for the fixed point version and the initial condition version of the scaled pools and NEE.

Fig. 7.3 shows the scaled pools when they are starting from their initial conditions, which are $C_f(0) = 150$, $C_r(0) = 160$, $C_w(0) = 9200$, $C_{lit}(0) = 60$ and $C_{som}(0) = 11000$.

From Table 7.3 we can understand why C_{som} and C_w are almost constant in Figs 7.1 and 7.3. For the $\tilde{C}_{som}(t+1)$ equation, the coefficient for $\tilde{C}_{som}(t)$, $(1 - p_9 T(t))$, is nearly

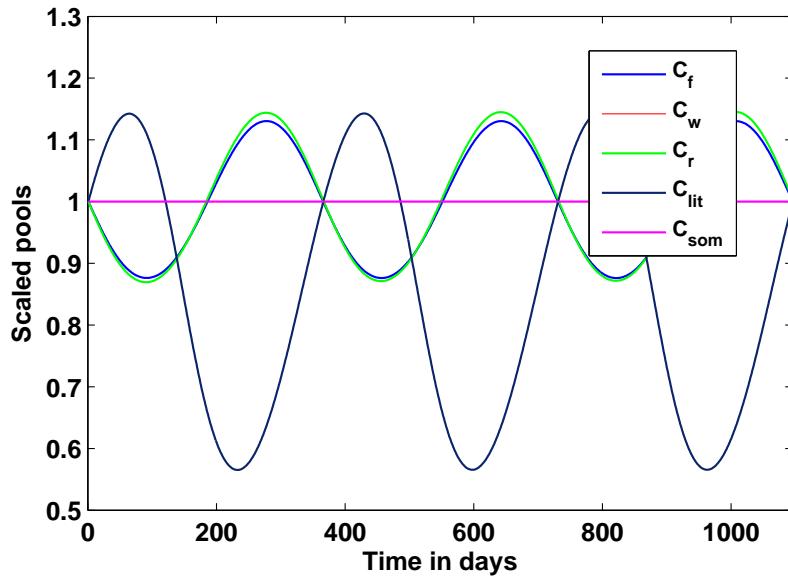


Figure 7.1: *The five carbon pools scaled from fixed points. We cannot see the difference between \tilde{C}_{som} and \tilde{C}_w here.*

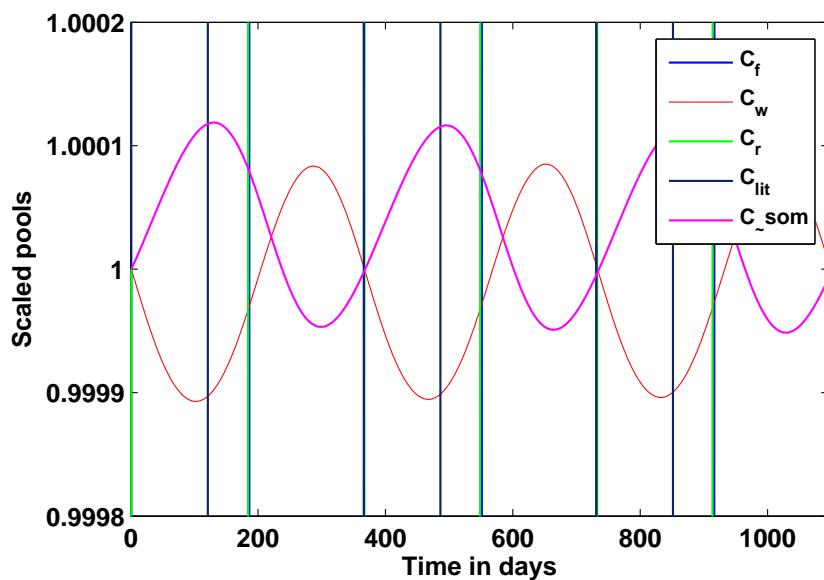


Figure 7.2: *The five carbon pools scaled from fixed points. Here we can see C_{som} and C_w up close. Their amplitudes are very small.*

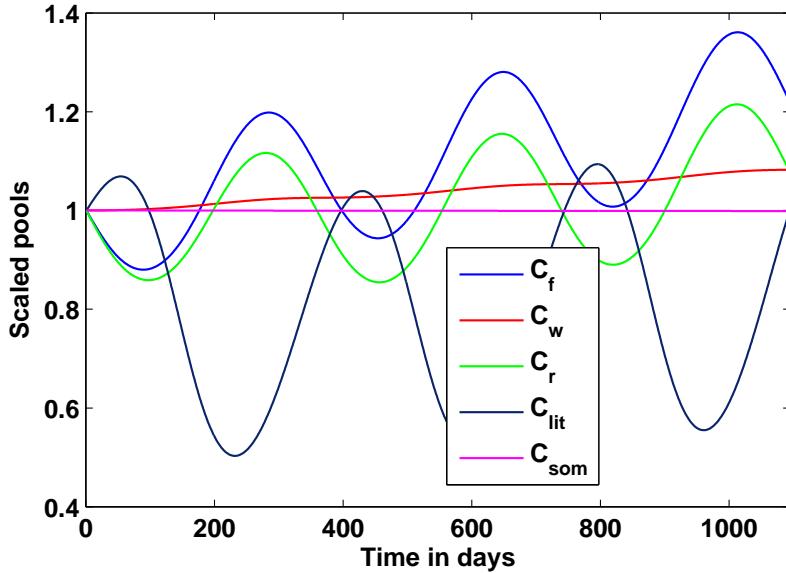


Figure 7.3: The five scaled carbon pools, starting from initial conditions.

1, as p_9 is very small. The coefficient of \tilde{C}_w , $\frac{p_6 C_w^*}{C_{som}^*}$, is of order 10^{-6} for both the fixed point version and the initial condition version and is very small. Therefore we can say that there is not a large contribution of \tilde{C}_w to this equation. Similarly, the coefficient for \tilde{C}_{lit} , $\frac{p_1 C_{lit}^*}{C_{som}^*}$, is very small, of order 10^{-9} for the fixed point version and of order 10^{-8} for the initial condition version and therefore the contribution of \tilde{C}_{lit} to this equation is negligible. All this together causes \tilde{C}_{som} to be almost constant in the medium term and to grow very slowly towards its fixed point. Also, for the $\tilde{C}_w(t+1)$ equation, the coefficient for $\tilde{C}_w(t)$, $(1-p_6)$, is nearly 1, as p_6 is very small. The coefficient of the GPP, $\frac{(1-p_2)(1-p_3)(1-p_4)}{C_w^*}$, is of order 10^{-7} for the fixed point version and 10^{-5} for the initial condition version, causing C_w to grow very slowly towards its fixed point (although not as slowly as C_{som}) and to be regarded as almost constant.

We also look at the size of the contributions of C_{som} and C_{lit} to the NEE equation, which are given by $p_8 C_{lit}^*$ and $p_9 C_{som}^*$. The coefficient for \tilde{C}_{lit} , $p_8 C_{lit}^*$, is significant for both fixed point and initial condition version and so is the contribution of the coefficient for \tilde{C}_{som} , $p_9 C_{som}^*$, even though parameter p_9 is very small.

	Term	From fixed point	From initial condition
In C_f equation	$\frac{p_3(1-p_2)}{C_f^*}$	$5.75E-04$	$9.28E-04$
	$(1-p_5)$	0.9972	0.9972
In C_r equation	$\frac{p_4(1-p_2)(1-p_3)}{C_r^*}$	$6.19E-04$	$8.73E-04$
	$(1-p_7)$	0.9970	0.9970
In C_w equation	$\frac{(1-p_2)(1-p_3)(1-p_4)}{C_w^*}$	$4.24E-07$	$2.19E-05$
	$(1-p_6)$	0.9999	0.9999
In C_{lit} equation	$\frac{p_5 C_f^*}{C_{lit}^*}$	0.0085	0.0070
	$\frac{p_7 C_r^*}{C_{lit}^*}$	0.0085	0.0080
	$(1 - (p_1 + p_8)(\bar{T}))$	0.9785	0.9785
In C_{som} equation	$\frac{p_6 C_w^*}{C_{som}^*}$	$2.85E-06$	$1.72E-08$
	$\frac{p_1 C_{lit}^*}{C_{som}^*}$	$1.03E-09$	$2.41E-08$
	$(1 - p_9(\bar{T}))$	0.9999	0.9999
In NEE equation	$p_8 C_{lit}^*$	1.61	1.20
	$p_9 C_{som}^*$	0.91	0.03

Table 7.3: Values of the coefficients for fixed point version and initial condition version of the scaled pools and NEE. $T(t)$ has been averaged over one year.

The coefficients of the scaled pools are all of more or less the same order for the initial condition version as for the fixed point version apart from $\frac{(1-p_2)(1-p_3)(1-p_4)}{C_w^*}$, $\frac{p_1 C_{lit}}{C_{som}^*}$ and $p_9 C_{som}^*$. The value of $\frac{(1-p_2)(1-p_3)(1-p_4)}{C_w^*}$ for the fixed point version is lower because C_w takes a much higher value at the fixed point than at the initial conditions, so the numerator is over a much larger denominator. Similarly, for $\frac{p_1 C_{lit}}{C_{som}^*}$ and $p_9 C_{som}^*$, the contribution of \tilde{C}_{som} to the NEE is higher for the fixed point version, because at the

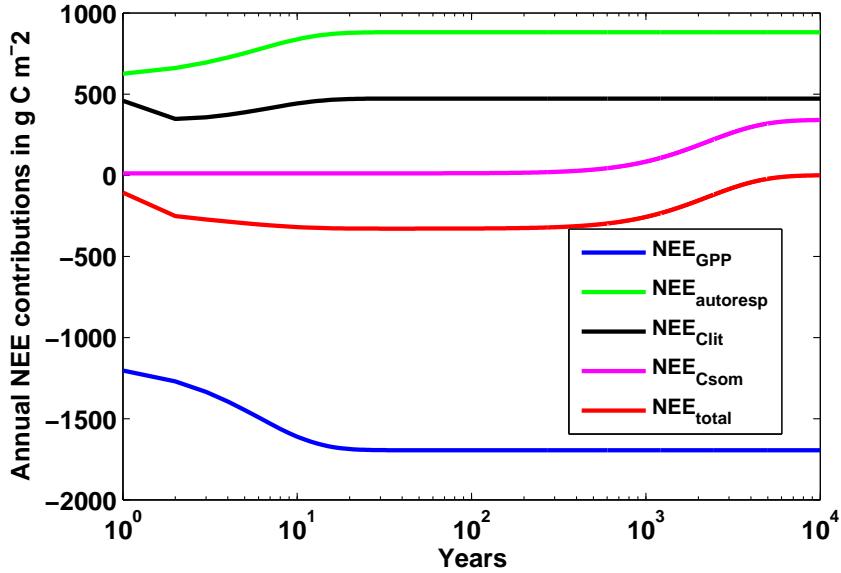


Figure 7.4: The different contributions to NEE.

fixed point C_{som} takes a much higher value. C_f , C_r and C_{lit} have fixed points of the same order as their initial conditions, in contrast to C_w and C_{som} , which have fixed points of order 10^2 and order 10 respectively higher than their initial conditions.

7.3.2 Contributions to NEE in More Detail

By scaling the pools and looking at the coefficients of the scaled pools in the model equations, we have gained a better numerical insight into how the pools contribute to each other and the NEE and why C_w and C_{som} take much longer to reach their fixed points than the other three pools. We now take a closer look at the contributions to the NEE.

As seen in equation (7.8), there are four contributions to the NEE, namely $p_8C_{lit}(t)T(t)$ and $p_9C_{som}(t)T(t)$, which together make up the heterotrophic respiration, $p_2GPP(C_f(t), t)$, the autotrophic respiration and $-GPP(C_f(t), t)$, the amount of carbon created by photosynthesis. Three of these represent sources of carbon terms, namely $p_8C_{lit}(t)T(t)$,

$p_9 C_{som}(t) T(t)$ and $p_2 GPP(C_f(t), t)$ and one a sink term, $-GPP(C_f(t), t)$. Fig. 7.4 shows these contributions, represented by $NEE_{C_{lit}}$, $NEE_{C_{som}}$, $NEE_{autoresp}$ and NEE_{GPP} respectively, along with the total NEE, NEE_{total} as calculated from initial conditions for the pools mentioned in Section 3.1. It can be seen that the C_{som} contribution seems to be constant for a long time, until about 1,000 years and then climbs, whilst the other three contributions settle fairly quickly, within about 30 years. The total NEE stays negative for a long time until, at around 1,000 years, it starts climbing and reaches zero at around 10,000 years.

It must be noted that Fig. 7.4 and consequent figures may look different if other initial conditions for the pools are used, in particular during the first tens of years. For example, in an area where a new forest is planted on the site of a capped old forest, the C_{som} may have a large amount of C left over from the old forest and may need to decrease whilst the young forest is growing, which means the forest would be a source rather than a sink.

In this subsection we find that NEE is more or less sensitive to certain parameters at particular periods in time, which we call

1. Phase I: the period from the initial conditions until C_f , C_r and C_{lit} have reached their fixed points;
2. Phase II: The period between when C_f , C_r and C_{lit} have reached their fixed points and approximately 1000 years;
 - Phase IIa: The period between when C_f , C_r and C_{lit} have reached their fixed points and roughly 150 years;
3. Phase III: The period after 1000 years.

We also seek to find explicit expressions for the sensitivities of the NEE to different parameters. This is valuable because the numerical sensitivity analysis only holds for

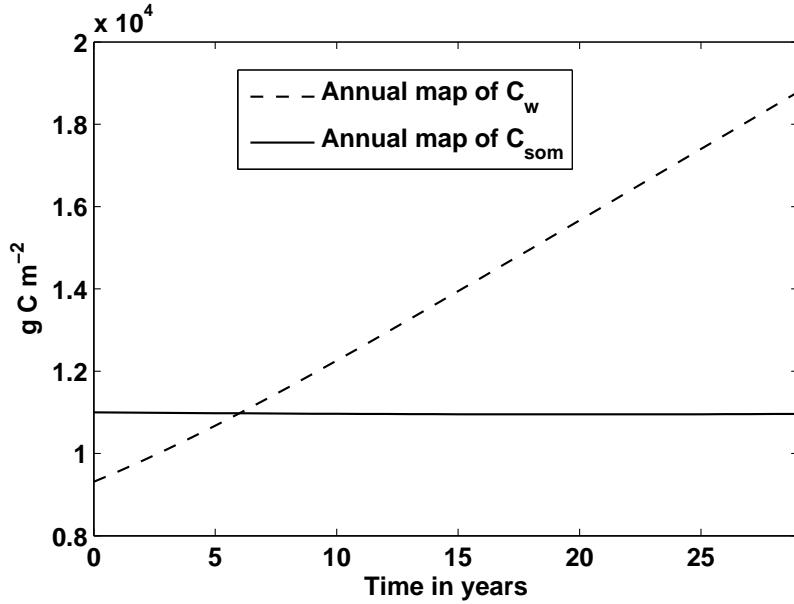


Figure 7.5: *Growth of C_w and C_{som} between 0 and 30 years.*

one particular forest with one set of climate data, whereas explicit expressions are more general.

NEE in Phase I

In this section we regard the NEE in the period from the initial conditions until C_f , C_r and C_{lit} have reached their fixed points, which is at around 30 years for the initial conditions we have used so far in this thesis. We then compare our results with the REFLEX results, as they represent the first three years from the initial conditions. From Figs 7.5 and 7.6 we can see that C_f , C_r , C_{lit} and C_w are all growing and it seems that C_{som} is constant. This has consequences for the parameter sensitivity of NEE with respect to parameter p_6 , as the equation for C_{som} is dependent on the C_w equation, which includes parameter p_6 . We consider two ways of writing the NEE: (1) as expanded from the daily maps and (2) as the sum of all the carbon pools to find out which method gives us the most information on the sensitivity of the NEE to the parameters.

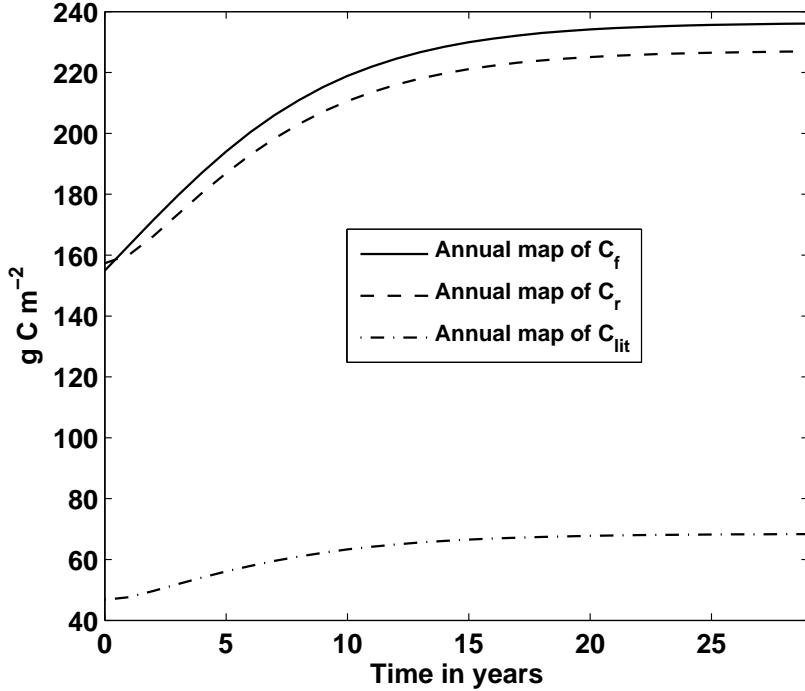


Figure 7.6: Growth of C_f , C_r and C_{lit} between 0 and 30 years.

NEE Expanded From the Daily Map

Recall equation (7.5):

$$NEE(t) = p_8 T(t) C_{lit}(t) + p_9 T(t) C_{som}(t) - (p_2 - 1) GPP(C_f(t), t).$$

Treating C_{som} as a constant, we can write equation (7.5) as:

$$NEE(t) = p_8 T(t) C_{lit}(t) + p_9 C_{som}(0) T(t) - (p_2 - 1) GPP(C_f(t), t). \quad (7.20)$$

Consider $p_8 T(t) C_{lit}(t)$ in equation (7.20). Recall that the equation for $C_{lit}(t+1)$ is:

$$C_{lit}(t+1) = (1 - (p_1 + p_8)T(t))C_{lit}(t) + p_5 C_f(t) + p_7 C_r(t) \quad (7.21)$$

and that given a map of the form, recall Section 4.2:

$$x_{i+1} = a_i x_i + b_i$$

then

$$x_t = \prod_{i=0}^{t-1} a_i x_0 + \sum_{i=0}^{t-1} b_i \prod_{j=i+1}^{t-1} a_j.$$

Since, also $p_1 \ll p_8$, we can write:

$$C_{lit}(t) \approx \prod_{i=0}^{t-1} (1 - p_8 T(i)) C_{lit}(0) + \sum_{i=0}^{t-1} (p_5 C_f(i) + p_7 C_r(i)) \prod_{j=i+1}^{t-1} (1 - p_8 T(j)). \quad (7.22)$$

Multiplying through by $p_8 T(t)$ then gives:

$$p_8 T(t) C_{lit}(t) = p_8 T(t) \prod_{i=0}^{t-1} (1 - p_8 T(i)) C_{lit}(0) + p_8 T(t) \sum_{i=0}^{t-1} (p_5 C_f(i) + p_7 C_r(i)) \prod_{j=i+1}^{t-1} (1 - p_8 T(j)). \quad (7.23)$$

Now consider from equation (7.23), the term $p_8 T(t) \prod_{i=0}^{t-1} (1 - p_8 T(i))$:

$$\begin{aligned} p_8 T(t) \prod_{i=0}^{t-1} (1 - p_8 T(i)) &= (-1 + p_8 T(t) + 1) \prod_{i=0}^{t-1} (1 - p_8 T(i)), \\ &= \prod_{i=0}^{t-1} (1 - p_8 T(i)) - \prod_{i=0}^t (1 - p_8 T(i)). \end{aligned}$$

We define:

$$A_{i,t} = \prod_{j=i}^t (1 - p_8 T(j)). \quad (7.24)$$

Then:

$$p_8 T(t) C_{lit}(t) = (A_{0,t-1} - A_{0,t}) C_{lit}(0) + \sum_{i=0}^{t-1} (p_5 C_f(i) + p_7 C_r(i)) (A_{i+1,t-1} - A_{i+1,t}) \quad (7.25)$$

Consider from equation (7.25) the $C_r(i)$ term. Recall that the equation for $C_r(t+1)$ is:

$$C_r(t+1) = (1 - p_7) C_r(t) + p_4(1 - p_2)(1 - p_3) GPP(C_f(t), t).$$

Also recall that given a map of the form:

$$x_{i+1} = ax_i + b_i,$$

then:

$$x_t = a^t x_0 + \sum_{i=0}^{t-1} b_{t-i-1} a^i.$$

Therefore we can write:

$$C_r(i) = (1 - p_7)^i C_r(0) + p_4(1 - p_2)(1 - p_3) \sum_{j=0}^{i-1} GPP(C_f(i-j-1))(1 - p_7)^j, \quad (7.26)$$

and therefore:

$$\begin{aligned} p_7 \sum_{i=0}^{t-1} C_r(i) (A_{i+1,t-1} - A_{i+1,t}) &= p_7 C_r(0) \sum_{i=0}^{t-1} (1 - p_7)^i (A_{i+1,t-1} - A_{i+1,t}) \\ &+ p_4 p_7 (1 - p_2)(1 - p_3) \sum_{i=1}^{t-1} \sum_{j=0}^{i-1} GPP(C_f(i-j-1))(1 - p_7)^j (A_{i+1,t-1} - A_{i+1,t}). \end{aligned} \quad (7.27)$$

Note that in the double sum we have changed the starting value from $i = 0$ to $i = 1$, since the second sum is not defined for $i = 0$ (and is assumed to be zero there).

We can manipulate the double sum in equation (7.27) according to the next lemma:

Figure 7.7: A visual representation of equation (7.28).

Lemma 7.3.1.

$$\sum_{i=1}^n \sum_{j=0}^{i-1} \alpha_i \beta_j \gamma_{i-j-1} = \sum_{m=0}^{n-1} \gamma_m \sum_{j=0}^{n-m-1} \alpha_{m+j+1} \beta_j. \quad (7.28)$$

for $n = 1, \dots$

Proof. Fig. 7.7 shows what happens to α , β and γ during the summing of $\sum_{i=1}^n \sum_{j=0}^{i-1} \alpha_i \beta_j \gamma_{i-j-1}$:

We could sum over the diagonals of the table in Fig. 7.7 by defining $m = i - j - 1$, which results in:

$$\sum_{m=0}^{n-1} \gamma_m \sum_{j=0}^{n-m-1} \alpha_{m+j+1} \beta_j. \quad (7.29)$$

□

Corollary 7.3.2.

$$\begin{aligned} & \sum_{i=0}^{t-1} \sum_{j=0}^{i-1} GPP(C_f(i-j-1)) (1-p_7)^j (A_{i+1,t-1} - A_{i+1,t}) \\ &= \sum_{m=0}^{t-2} GPP(C_f(m)) \sum_{j=0}^{t-2-m} (1-p_7)^j (A_{m+j+2,t-1} - A_{m+j+2,t}) \end{aligned} \quad (7.30)$$

Therefore:

$$\begin{aligned}
 p_7 \sum_{i=0}^{t-1} C_r(i)(A_{i+1,t-1} - A_{i+1,t}) &= p_7 C_r(0) \sum_{i=0}^{t-1} (1-p_7)^i (A_{i+1,t-1} - A_{i+1,t}) + p_4 p_7 \\
 &\times (p_2 - 1)(1-p_3) \sum_{m=0}^{t-2} GPP(C_f(m)) \sum_{j=0}^{t-2-m} (A_{m+j+2,t-1} - A_{m+j+2,t}) (1-p_7)^j \quad (7.31)
 \end{aligned}$$

and we can substitute equation (7.25) to give:

$$\begin{aligned}
 p_8 T(t) C_{lit}(t) &= (A_{0,t-1} - A_{0,t}) C_{lit}(0) + p_5 \sum_{i=0}^{t-1} C_f(i)(A_{i+1,t-1} - A_{i+1,t}) \\
 &+ p_7 C_r(0) \sum_{i=0}^{t-1} (1-p_7)^i (A_{i+1,t-1} - A_{i+1,t}) + p_4 p_7 (p_2 - 1)(1-p_3) \\
 &\times \sum_{m=0}^{t-2} GPP(C_f(m)) \sum_{j=0}^{t-2-m} (1-p_7)^j (A_{m+j+2,t-1} - A_{m+j+2,t}). \quad (7.32)
 \end{aligned}$$

So now we can write equation (7.20) as:

$$\begin{aligned}
 NEE(t) &= (A_{0,t-1} - A_{0,t}) C_{lit}(0) + p_5 \sum_{i=0}^{t-1} C_f(i)(A_{i+1,t-1} - A_{i+1,t}) \\
 &+ p_7 C_r(0) \sum_{i=0}^{t-1} (1-p_7)^i (A_{i+1,t-1} - A_{i+1,t}) \\
 &+ p_4 p_7 (p_2 - 1)(1-p_3) \sum_{m=0}^{t-2} GPP(C_f(m)) \sum_{j=0}^{t-2-m} (1-p_7)^j (A_{m+j+2,t-1} - A_{m+j+2,t}) \\
 &+ p_9 T(t) C_{som}(0) - (p_2 - 1) GPP(C_f(t)). \quad (7.33)
 \end{aligned}$$

We can write the total NEE accumulated in one year (for $N = 364$) as:

$$\begin{aligned}
 \sum_{t=0}^N NEE(t) &= \sum_{t=0}^N \left\{ (A_{0,t-1} - A_{0,t})C_{lit}(0) + p_5 \sum_{i=0}^{t-1} C_f(i)(A_{i+1,t-1} - A_{i+1,t}) \right. \\
 &\quad + p_7 C_r(0) \sum_{i=0}^{t-1} (1 - p_7)^i (A_{i+1,t-1} - A_{i+1,t}) \\
 &\quad + p_4 p_7 (p_2 - 1)(1 - p_3) \sum_{m=0}^{t-2} GPP(C_f(m)) \\
 &\quad \times \left. \sum_{j=0}^{t-m-2} (1 - p_7)^j (A_{m+j+2,t-1} - A_{m+j+2,t}) \right\} \\
 &\quad + p_9 C_{som}(0) \sum_{t=0}^N T(t) - (p_2 - 1) \sum_{t=0}^N GPP(C_f(t)). \tag{7.34}
 \end{aligned}$$

This can be simplified further. Consider from equation (7.34) the following part:

$$\sum_{t=0}^N \sum_{m=0}^{t-2} GPP(C_f(m)) \sum_{j=0}^{t-m-2} (1 - p_7)^j (A_{m+j+2,t-1} - A_{m+j+2,t}). \tag{7.35}$$

We would like to write this in the style of $(p_2 - 1) \sum_{t=0}^N GPP(C_f(t))$ in equation (7.34) so that we can combine the two and simplify. Therefore we manipulate equation (7.35) as follows:

$$\begin{aligned}
 &\sum_{t=0}^N \sum_{m=0}^{t-2} GPP(C_f(m)) \sum_{j=0}^{t-m-2} (1 - p_7)^j (A_{m+j+2,t-1} - A_{m+j+2,t}) = \\
 &\sum_{m=0}^{N-2} GPP(C_f(m)) \sum_{t=m+2}^N \sum_{j=0}^{t-m-2} (1 - p_7)^j (A_{m+j+2,t-1} - A_{m+j+2,t}) \tag{7.36}
 \end{aligned}$$

and write $m = t$ and $t = t'$:

$$\begin{aligned} & \sum_{m=0}^{N-2} GPP(C_f(m)) \sum_{t=m+2}^N \sum_{j=0}^{t-m-2} (1-p_7)^j (A_{m+j+2,t-1} - A_{m+j+2,t}) = \\ & \sum_{t=0}^{N-2} GPP(C_f(t)) \sum_{t'=t+2}^N \sum_{j=0}^{t'-t-2} (1-p_7)^j (A_{t+j+2,t'-1} - A_{t+j+2,t'}). \end{aligned} \quad (7.37)$$

Now consider $\sum_{t=0}^N (A_{0,t-1} - A_{0,t})$ from equation (7.34). We can write :

$$\sum_{t=0}^N (A_{0,t-1} - A_{0,t}) = A_{0,1} - A_{0,N} = 1 - A_{0,N}, \quad (7.38)$$

using the convention that a product has the value 1 if the upper index is smaller than the lower index.

Equation (7.34) can then be written as follows:

$$\begin{aligned} \sum_{t=0}^N NEE(t) &= (1 - A_{0,N}) C_{lit}(0) + \sum_{t=0}^N \left\{ p_5 \sum_{i=0}^{t-1} C_f(i) (A_{i+1,t-1} - A_{i+1,t}) \right. \\ &\quad \left. + p_7 C_r(0) \sum_{i=0}^{t-1} (1-p_7)^i (A_{i+1,t-1} - A_{i+1,t}) \right\} - (p_2 - 1) \sum_{t=0}^{N-2} GPP(C_f(t)) \\ &\quad \times \left\{ 1 - p_4 p_7 (1-p_3) \sum_{t'=t+2}^N \sum_{j=0}^{t'-t-2} (1-p_7)^j (A_{t+j+2,t'-1} - A_{t+j+2,t'}) \right\} \\ &\quad - (p_2 - 1) (GPP(C_f(N-1)) + GPP(C_f(N))) + p_9 C_{som}(0) \sum_{t=0}^N T(t). \end{aligned} \quad (7.39)$$

We can manipulate the double sums in this equation in a similar way to Lemma 7.3.1:

$$\sum_{t=0}^N \sum_{i=0}^{t-1} = \sum_{i=0}^{N-1} \sum_{t=i+1}^N \quad (7.40)$$

and

$$\sum_{t'=t+2}^N \sum_{j=0}^{t'-t-2} = \sum_{j=0}^{N-t-2} \sum_{t'=j+t+2}^N , \quad (7.41)$$

which means that:

$$\begin{aligned} \sum_{t=0}^N p_5 \sum_{i=0}^{t-1} C_f(i) (A_{i+1,t-1} - A_{i+1,t}) &= p_5 \sum_{i=0}^{N-1} C_f(i) \sum_{t=i+1}^N (A_{i+1,t-1} - A_{i+1,t}) \\ &= p_5 \sum_{i=0}^{N-1} C_f(i) (1 - A_{i+1,N}) \end{aligned} \quad (7.42)$$

and

$$\begin{aligned} \sum_{t=0}^N p_7 C_r(0) \sum_{i=0}^{t-1} (1 - p_7)^i (A_{i+1,t-1} - A_{i+1,t}) &= p_7 C_r(0) \sum_{i=0}^{N-1} \sum_{t=i+1}^N (1 - p_7)^i (A_{i+1,t-1} - A_{i+1,t}) \\ &= p_7 C_r(0) \sum_{i=0}^{N-1} (1 - p_7)^i \sum_{t=i+1}^N (A_{i+1,t-1} - A_{i+1,t}) \\ &= p_7 C_r(0) \sum_{i=0}^{N-1} (1 - p_7)^i (1 - A_{i+1,N}) \end{aligned} \quad (7.43)$$

and

$$\begin{aligned} \sum_{t'=t+2}^N \sum_{j=0}^{t'-t-2} (1 - p_7)^j (A_{t+j+2,t'-1} - A_{t+j+2,t'}) &= \sum_{j=0}^{N-t-2} (1 - p_7)^j \sum_{t'=j+t+2}^N (A_{j+t+2,t'-1} - A_{j+t+2,t'}) \\ &= \sum_{j=0}^{N-t-2} (1 - p_7)^j (1 - A_{j+t+2,N}) \end{aligned} \quad (7.44)$$

Therefore we can write an equation for the sum of the NEE over one year, which we will

now call the annual NEE, as:

$$\begin{aligned}
 NEE_{ann}^{(n+1)} = & (1 - A_{0,N}) C_{lit}^{(n)} + p_5 \sum_{t=365n}^{\tilde{N}-1} C_f(t)(1 - A_{l+1,N}) \\
 & + p_7 C_r^{(n)} \sum_{t=0}^{N-1} (1 - p_7)^t (1 - A_{t+1,N}) \\
 & - (p_2 - 1) \sum_{t=365n}^{\tilde{N}-2} GPP(C_f(t)) \left\{ 1 - p_4 p_7 (1 - p_3) \right. \\
 & \times \left. \sum_{j=0}^{N-l-2} (1 - p_7)^j (1 - A_{j+l+2,N}) \right\} - (p_2 - 1) (GPP(C_f(\tilde{N} - 1)) \\
 & + GPP(C_f(\tilde{N}))) + p_9 C_{som}(0) \sum_{t=0}^N T(t),
 \end{aligned} \tag{7.45}$$

where $n \geq 0$ represents the year, $N = 364$, $\tilde{N} = 365(n + 1) - 1$, $l = t - 365n$ and:

$$A_{i,t} = \prod_{j=i}^t (1 - p_8 T(j)) \tag{7.46}$$

We assumed that $T(t)$ is periodic with period 365 and hence the totals that include the drivers do not involve the year.

To make sure that the approximations of neglecting p_1 and regarding C_{som} as a constant, thereby effectively ignoring p_6 , are reasonable, we have run some tests. They show us how much the equation (7.45) for the NEE is different from the calculations for the original formula, equation (7.8). The exact value of the first year's NEE for initial conditions of the carbon pools as $C_f(0) = 150$, $C_r(0) = 160$, $C_w(0) = 9200$, $C_{som}(0) = 11000$ and $C_{lit}(0) = 60$, is -235.48 . For equation (7.45) we get a value of -236.16 for the first year's NEE. That is approximately 0.3% difference. Therefore we can say that equation (7.45) is a good approximation for year 1.

From equation (7.45) we can see which parts of the NEE equation contribute positively or negatively, being sources or sinks respectively:

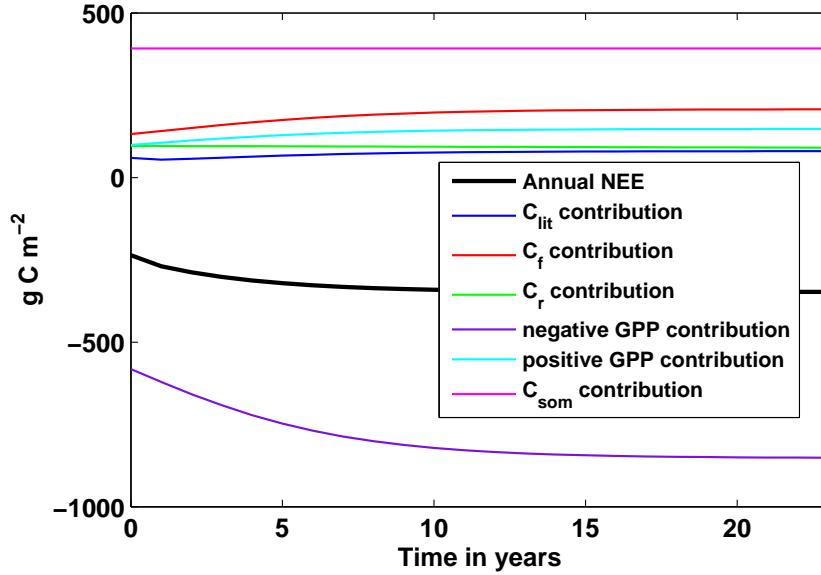


Figure 7.8: Contributions to the annual NEE as per equation (7.45).

Negative contributions to NEE:

$$(p_2 - 1) \sum_{t=365n}^{\tilde{N}-2} GPP(C_f(t)) \quad (7.47)$$

Positive contributions to NEE:

$$(1 - A_{0,N})C_{lit}^{(n)} \quad (7.48)$$

$$p_5 \sum_{t=365n}^{\tilde{N}-1} C_f(t)(1 - A_{l+1,N}) \quad (7.49)$$

$$p_7 C_r^{(n)} \sum_{t=0}^{N-1} (1 - p_7)^t (1 - A_{t+1,\tilde{N}}) \quad (7.50)$$

$$p_4 p_7 (p_2 - 1)(1 - p_3) \sum_{t=365n}^{\tilde{N}-2} GPP(C_f(t)) \left\{ \sum_{j=0}^{N-l-2} (1 - p_7)^j (1 - A_{j+l+2,N}) \right\} \quad (7.51)$$

$$p_9 C_{som}(0) \sum_{t=0}^N T(t). \quad (7.52)$$

Fig. 7.8 shows the negative and positive contributions to NEE. We can see that it does not take long for the NEE to become more or less constant. The negative contribution of $(p_2 - 1) \sum_{t=365n}^{\tilde{N}} GPP(C_f(t))$ is the largest, followed by the positive contribution of C_{som} .

We now consider the dependence on NEE on each of the parameters. From equation (7.45), we can see that the NEE is linear in p_4 and p_9 and nonlinear in all the other parameters, apart from p_1 and p_6 as we neglected those. Remember, p_{10} is part of T , the temperature coefficient.

To determine the sensitivity of the NEE to the various parameters, we differentiate NEE with respect to the parameters:

$$\begin{aligned} \frac{\partial NEE_{ann}^{(n+1)}}{\partial p_2} &= (1 - A_{0,N}) \frac{\partial C_{lit}^{(n)}}{\partial p_2} + p_5 \sum_{t=365n}^{\tilde{N}-1} \frac{\partial C_f(t)}{\partial p_2} (1 - A_{l+1,N}) + p_7 \frac{\partial C_r^{(n)}}{\partial p_2} \sum_{t=0}^{N-1} (1 - p_7)^t (1 - A_{t+1,N}) \\ &\quad + \sum_{t=365n}^{\tilde{N}-2} \left(GPP(C_f(t)) - (p_2 - 1) \frac{\partial GPP(C_f(t))}{\partial p_2} \right) \left\{ 1 - p_4 p_7 (1 - p_3) \sum_{j=0}^{N-l-2} (1 - p_7)^j (1 - A_{j+l+2,N}) \right\} \\ &\quad + GPP(C_f(\tilde{N} - 1)) + GPP(C_f(\tilde{N})) - (p_2 - 1) \frac{\partial GPP(C_f(\tilde{N} - 1))}{\partial p_2} + \frac{\partial GPP(C_f(\tilde{N}))}{\partial p_2} \end{aligned} \quad (7.53)$$

$$\begin{aligned} \frac{\partial NEE_{ann}^{(n+1)}}{\partial p_3} &= (1 - A_{0,N}) \frac{\partial C_{lit}^{(n)}}{\partial p_3} + p_5 \sum_{t=365n}^{\tilde{N}-1} \frac{\partial C_f(t)}{\partial p_3} (1 - A_{l+1,N}) + p_7 \frac{\partial C_r^{(n)}}{\partial p_3} \sum_{t=0}^{N-1} (1 - p_7)^t (1 - A_{t+1,N}) \\ &\quad - (p_2 - 1) \sum_{t=365n}^{\tilde{N}-2} GPP(C_f(t)) \left\{ p_4 p_7 \sum_{j=0}^{N-l-2} (1 - p_7)^j (1 - A_{j+l+2,N}) \right\} \\ &\quad - (p_2 - 1) \sum_{t=365n}^{\tilde{N}-2} \frac{\partial GPP(C_f(t))}{\partial p_3} \left\{ 1 - p_4 p_7 (1 - p_3) \sum_{j=0}^{N-l-2} (1 - p_7)^j (1 - A_{j+l+2,N}) \right\} \\ &\quad - (p_2 - 1) \left(\frac{\partial GPP(C_f(\tilde{N} - 1))}{\partial p_3} + \frac{\partial GPP(C_f(\tilde{N}))}{\partial p_3} \right) \end{aligned} \quad (7.54)$$

$$\begin{aligned} \frac{\partial NEE_{ann}^{(n+1)}}{\partial p_4} &= (1 - A_{0,N}) \frac{\partial C_{lit}^{(n)}}{\partial p_4} + p_7 \frac{\partial C_r^{(n)}}{\partial p_4} \sum_{t=0}^{N-1} (1 - p_7)^t (1 - A_{t+1,N}) \\ &\quad + (p_2 - 1) \sum_{t=365n}^{\tilde{N}-2} GPP(C_f(t)) \left\{ p_7 (1 - p_3) \sum_{j=0}^{N-l-2} (1 - p_7)^j (1 - A_{j+l+2,N}) \right\} \end{aligned} \quad (7.55)$$

$$\begin{aligned}
\frac{\partial NEE_{ann}^{(n+1)}}{\partial p_5} = & (1 - A_{0,N}) \frac{\partial C_{lit}^{(n)}}{\partial p_5} + \sum_{t=365n}^{\tilde{N}-1} (1 - A_{l+1,N}) \left[C_f(t) + p_5 \frac{\partial C_f(t)}{\partial p_5} \right] \\
& + p_7 \frac{\partial C_r^{(n)}}{\partial p_5} \sum_{t=0}^{N-1} (1 - p_7)^t (1 - A_{t+1,N}) - (p_2 - 1) \left\{ \sum_{t=365n}^{\tilde{N}-2} \frac{GPP(C_f(t))}{\partial p_5} \right. \\
& \times \left[1 - p_4 p_7 (1 - p_3) \sum_{j=0}^{N-l-2} (1 - p_7)^j (1 - A_{j+l+2,N}) \right] + \frac{\partial GPP(C_f(\tilde{N}-1))}{\partial p_5} + \frac{\partial GPP(C_f(\tilde{N}))}{\partial p_5} \left. \right\} \quad (7.56)
\end{aligned}$$

$$\begin{aligned}
\frac{\partial NEE_{ann}^{(n+1)}}{\partial p_7} = & (1 - A_{0,N}) \frac{\partial C_{lit}^{(n)}}{\partial p_7} + p_7 \frac{\partial C_r^{(n)}}{\partial p_7} \sum_{t=0}^{N-1} (1 - p_7)^t (1 - A_{t+1,N}) \\
& + C_r^{(n)} \sum_{t=0}^{N-1} ((1 - p_7)^t - p_7 t (1 - p_7)^{t-1}) (1 - A_{t+1,N}) \\
& + p_4 (p_2 - 1) (1 - p_3) \sum_{t=365n}^{\tilde{N}-2} GPP(C_f(t)) \sum_{j=0}^{N-l-2} ((1 - p_7)^j - p_7 j (1 - p_7)^{j-1}) (1 - A_{j+l+2,N}) \quad (7.57)
\end{aligned}$$

$$\begin{aligned}
\frac{\partial NEE_{ann}^{(n+1)}}{\partial p_8} = & (1 - A_{0,N}) \frac{\partial C_{lit}^{(n)}}{\partial p_8} - \frac{\partial A_{0,N}}{\partial p_8} C_{lit}^{(n)} - p_5 \sum_{t=365n}^{\tilde{N}-1} C_f(t) \frac{\partial A_{l+1,N}}{\partial p_8} \\
& - p_7 C_r^{(n)} \sum_{t=0}^{N-1} (1 - p_7)^t \frac{\partial A_{t+1,N}}{\partial p_8} - (p_2 - 1) \sum_{t=365n}^{\tilde{N}-2} GPP(C_f(t)) \\
& \times \left\{ p_4 p_7 (1 - p_3) \sum_{j=0}^{N-l-2} (1 - p_7)^j \frac{\partial A_{j+l+2,N}}{\partial p_8} \right\} \quad (7.58)
\end{aligned}$$

$$\frac{\partial NEE_{ann}^{(n+1)}}{\partial p_9} = C_{som}(0) \sum_{t=0}^N T(t) \quad (7.59)$$

$$\frac{\partial NEE_{ann}^{(n+1)}}{\partial p_{10}} = (1 - A_{0,N}) \frac{\partial C_{lit}^{(n)}}{\partial p_{10}} - \frac{\partial A_{0,N}}{\partial p_{10}} C_{lit}^{(n)} - p_5 \sum_{t=365n}^{\tilde{N}-1} C_f(t) \frac{\partial A_{l+1,N}}{\partial p_{10}}$$

$$-p_7 C_r^{(n)} \sum_{t=0}^{N-1} (1 - p_7)^t \frac{\partial A_{t+1,N}}{\partial p_{10}} - (p_2 - 1) \sum_{t=365n}^{\tilde{N}-2} GPP(C_f(t)) \\ \times \left\{ p_4 p_7 (1 - p_3) \sum_{j=0}^{N-l-2} (1 - p_7)^j \frac{\partial A_{j+l+2,N}}{\partial p_{10}} \right\} + p_9 C_{som}(0) \sum_{t=0}^N \frac{\partial T(t)}{\partial p_{10}} \quad (7.60)$$

$$\frac{\partial NEE_{ann}^{(n+1)}}{\partial p_{11}} = (1 - A_{0,N}) \frac{\partial C_{lit}^{(n)}}{\partial p_{11}} + p_5 \sum_{t=365n}^{\tilde{N}-1} \frac{\partial C_f(t)}{\partial p_{11}} (1 - A_{l+1,N}) \\ + p_7 \frac{\partial C_r^{(n)}}{\partial p_{11}} \sum_{t=0}^{N-1} (1 - p_7)^t (1 - A_{t+1,N}) - (p_2 - 1) \left\{ \sum_{t=365n}^{\tilde{N}-2} \frac{GPP(C_f(t))}{\partial p_{11}} \right. \\ \times \left\{ 1 - p_4 p_7 (1 - p_3) \sum_{j=0}^{N-l-2} (1 - p_7)^j (1 - A_{j+l+2,N}) \right\} \\ \left. + \frac{\partial GPP(C_f(\tilde{N}-1))}{\partial p_{11}} + \frac{\partial GPP(C_f(\tilde{N}))}{\partial p_{11}} \right\} \quad (7.61)$$

The only parameters for which the formulae are relatively straightforward are p_4 , p_7 , p_9 and p_{11} . Parameters p_8 and p_{10} are embedded in $A_{i,t}$ and therefore it is difficult to write down useful derivatives. The derivatives for parameters p_2 , p_3 and p_5 are complicated by the fact that they include recurrence relations in the various pools, which is illustrated by the following example for parameter p_2 :

$$\begin{aligned} C_f(t+1) &= (1 - p_5)C_f(t) + p_3(p_2 - 1)GPP(C_f(t), t) \\ \therefore \frac{\partial C_f(t+1)}{\partial p_2} &= (1 - p_5)\frac{\partial C_f(t)}{\partial p_2} - p_3GPP(C_f(t)) \\ &\quad + p_3(p_2 - 1)\frac{\partial GPP}{\partial C_f(t)}\frac{\partial C_f(t)}{\partial p_2} \end{aligned} \tag{7.62}$$

As we do not have an analytical solution for $C_f(t)$, it is impossible to construct a useful analytical derivative of NEE with respect to parameters p_2 , p_3 and p_5 .

There are no derivatives for parameters p_1 and p_6 as they have been neglected.

NEE as a Sum of the Carbon Pools

The last section showed us that by expanding and simplifying the original NEE we have found relatively simple expressions for the derivatives of p_4 , p_7 , p_9 and p_{11} , but much more complicated derivatives for p_2 , p_3 , p_5 , p_8 and p_{10} and therefore it is still very difficult to see exactly what role those parameters play. Therefore in this section we write the NEE as the total sum of all the carbon pools and find out if this makes it any easier.

We can express the total NEE produced in year $n + 1$, where $n \geq 0$, as follows:

$$\begin{aligned} NEE_{ann}^{(n+1)} = & - \left(C_f^{(n+1)} - C_f^{(n)} + C_r^{(n+1)} - C_r^{(n)} + C_w^{(n+1)} - C_w^{(n)} \right. \\ & \left. + C_{lit}^{(n+1)} - C_{lit}^{(n)} + C_{som}^{(n+1)} - C_{som}^{(n)} \right). \end{aligned} \quad (7.63)$$

As before we regard C_{som} as a constant. Therefore $NEE_{ann}^{(n+1)}$ becomes:

$$\begin{aligned} NEE_{ann}^{(n+1)} = & - \left(C_f^{(n+1)} - C_f^{(n)} + C_r^{(n+1)} - C_r^{(n)} + C_w^{(n+1)} - C_w^{(n)} \right. \\ & \left. + C_{lit}^{(n+1)} - C_{lit}^{(n)} \right). \end{aligned} \quad (7.64)$$

From Section 5.3 we know we can express the annual map for C_f as follows:

$$C_f^{(n+1)} = (1 - p_5)^{365} C_f^{(n)} + \frac{(1 - (1 - p_5)^{365})}{p_5} p_3 (1 - p_2) f_1 \overline{GPP}(C_f^{(n)}), \quad (7.65)$$

using the averaged GPP and $f_1 = 1.016$ (for the Loobos data).

In the same way we can create an annual map for C_w :

$$C_w^{(n+1)} = (1 - p_6)^{365} C_w^{(n)} + f_2 \frac{1 - (1 - p_6)^{365}}{p_6} (1 - p_2)(1 - p_3)(1 - p_4) f_2 \overline{GPP}(C_f^{(n)}). \quad (7.66)$$

using the averaged GPP and $f_2 = 1.05$.

As p_6 is of order 10^{-6} , we can manipulate $C_w^{(n+1)} - C_w^{(n)}$ as follows:

$$\begin{aligned} C_w^{(n+1)} - C_w^{(n)} &= ((1 - p_6)^{365} - 1) C_w^{(n)} \\ &\quad + \frac{1 - (1 - p_6)^{365}}{p_6} (1 - p_2)(1 - p_3)(1 - p_4) \overline{GPP}(C_f^{(n)}) \\ &= -365p_6 C_w^{(n)} + 365(1 - p_2)(1 - p_3)(1 - p_4) f_2 \overline{GPP}(C_f^{(n)}). \end{aligned} \quad (7.67)$$

The annual map for C_r can be calculated in a similar way to those of C_f and C_w and can be expressed as:

$$\begin{aligned} C_r^{(n+1)} &= (1 - p_7)^{365} C_r^{(n)} + \frac{(1 - (1 - p_7)^{365})}{p_7} \\ &\quad \times p_4(1 - p_2)(1 - p_3)f_3\overline{GPP}(C_f^{(n)}), \end{aligned} \quad (7.68)$$

where $f_3 = 1.042$, found by data assimilation to achieve a best fit to the annual map created from the mean values of C_r calculated by the daily map over one year.

It is slightly more complicated to create an annual map for C_{lit} as we have to consider C_f and C_r in this equation. Recall the equation for the original C_{lit} :

$$C_{lit}(t+1) = (1 - (p_8 + p_1)T(t))C_{lit}(t) + p_5C_f(t) + p_7C_r(t). \quad (7.69)$$

As $p_1 \ll p_8$, we neglect p_1 in the following calculations. In addition we average $T(t)$, so that equation (7.69) becomes:

$$C_{lit}(t+1) = (1 - p_8\bar{T})C_{lit}(t) + p_5C_f(t) + p_7C_r(t). \quad (7.70)$$

According to the technique described in Section 4.4.1 the annual map for C_{lit} can be written as follows:

$$C_{lit}^{(n+1)} = (1 - p_8\bar{T})^{365}C_{lit}^{(n)} + \sum_{i=0}^{364} (1 - p_8\bar{T})^{364-i} [p_5C_f(i) + p_7C_r(i)]. \quad (7.71)$$

We know that:

$$C_f(i) = (1 - p_5)^i C_{f0} + \frac{1 - (1 - p_5)^i}{p_5} p_3(1 - p_2)f_1\overline{GPP}(C_f^{(0)}) \quad (7.72)$$

and

$$C_r(i) = (1 - p_7)^i C_{r0} + \frac{1 - (1 - p_7)^i}{p_7} p_4(1 - p_2)(1 - p_3)f_3 \overline{GPP}(C_f^{(0)}). \quad (7.73)$$

Replacing C_{f0} and C_{r0} with $C_f^{(n)}$ and $C_r^{(n)}$, we get:

$$\begin{aligned} C_{lit}^{(n+1)} = & (1 - p_8 \bar{T})^{365} C_{lit}^{(n)} + f_4 \left\{ p_5 C_f^{(n)} \sum_{i=0}^{364} (1 - p_8 \bar{T})^{364-i} (1 - p_5)^i \right. \\ & + p_3(1 - p_2)f_1 \overline{GPP}(C_f^{(n)}) \sum_{i=0}^{364} (1 - p_8 \bar{T})^{364-i} (1 - (1 - p_5)^i) \\ & + p_7 C_r^{(n)} \sum_{i=0}^{364} (1 - p_8 \bar{T})^{364-i} (1 - p_7)^i \\ & + p_4(1 - p_2)(1 - p_3)f_3 \overline{GPP}(C_f^{(n)}) \\ & \left. \times \sum_{i=0}^{364} (1 - p_8 \bar{T})^{364-i} (1 - (1 - p_7)^i) \right\}, \end{aligned} \quad (7.74)$$

where $f_4 = 1.08$ and has been found by using data assimilation, to achieve an optimal fit to the annual map created from the mean values of C_{lit} calculated by the daily map over one year.

Now we can write the annual NEE as follows:

$$\begin{aligned}
 NEE_{ann}^{(n+1)} = & - \left(\left\{ \left((1 - p_5)^{365} - 1 \right) + f_4 p_5 \sum_{i=0}^{364} (1 - p_8 \bar{T})^{364-i} (1 - p_5)^i \right\} C_f^{(n)} \right. \\
 & + \left\{ \left((1 - p_7)^{365} - 1 \right) + f_4 p_7 \sum_{i=0}^{364} (1 - p_8 \bar{T})^{364-i} (1 - p_7)^i \right\} C_r^{(n)} \\
 & + \left((1 - p_8 \bar{T})^{365} - 1 \right) C_{lit}^{(n)} - 365 p_6 C_w^{(n)} \\
 & + (1 - p_2) \overline{GPP}(C_f^{(n)}) \left\{ \frac{(1 - (1 - p_5)^{365})}{p_5} p_3 f_1 \right. \\
 & + \frac{(1 - (1 - p_7)^{365})}{p_7} p_4 (1 - p_3) f_3 + 365 (1 - p_4) (1 - p_3) f_2 \\
 & + p_3 f_1 f_4 \sum_{i=0}^{364} (1 - p_8 \bar{T})^{364-i} (1 - (1 - p_5)^i) \\
 & \left. \left. + p_4 (1 - p_3) f_3 f_4 \sum_{i=0}^{364} (1 - p_8 \bar{T})^{364-i} (1 - (1 - p_7)^i) \right\} \right), \quad (7.75)
 \end{aligned}$$

where $n \geq 0$ represents the year.

To make sure that the approximations of neglecting p_1 and regarding C_{com} as a constant, thereby effectively ignoring p_9 , are reasonable, we have run some tests. They show us how much the equation (7.75) for the NEE is different from the calculations for the original formula, equation (7.8). The exact value of the first year's NEE for initial conditions of the carbon pools as $C_f(0) = 150$, $C_r(0) = 160$, $C_w(0) = 9200$, $C_{som}(0) = 11000$ and $C_{lit}(0) = 60$, is -235.48 . For equation (7.75) we get a value of -243.89 for the first year's NEE. That is approximately 3.5% difference. Therefore we can say that equation (7.75) is a reasonable approximation for year 1.

Differentiating the NEE with respect to its parameters gives the following result:

$$\begin{aligned} \frac{\partial NEE_{ann}^{(n+1)}}{\partial p_2} = & - \left\{ ((1-p_5)^{365} - 1) + f_4 p_5 \sum_{i=0}^{364} (1-p_8 \bar{T})^{364-i} (1-p_5)^i \right\} \frac{\partial C_f^{(n)}}{\partial p_2} - \left\{ ((1-p_7)^{365} - 1) + f_4 p_7 \sum_{i=0}^{364} (1-p_8 \bar{T})^{364-i} \right. \\ & \times (1-p_7)^i \left. \right\} \frac{\partial C_r^{(n)}}{\partial p_2} - ((1-p_8 \bar{T})^{365} - 1) \frac{\partial C_{lit}^{(n)}}{\partial p_2} + 365 p_6 \frac{\partial C_w^{(n)}}{\partial p_2} - \left((1-p_2) \frac{\partial \overline{GPP}(C_f^{(n)})}{\partial p_2} - \overline{GPP}(C_f^{(n)}) \right) \\ & \times \left\{ \frac{(1-(1-p_5)^{365})}{p_5} p_3 f_1 + \frac{(1-(1-p_7)^{365})}{p_7} p_4 (1-p_3) f_3 + 365 (1-p_4) (1-p_3) f_2 \right. \\ & + p_3 f_1 f_4 \sum_{i=0}^{364} (1-p_8 \bar{T})^{364-i} (1-(1-p_5)^i) + p_4 (1-p_3) f_3 f_4 \sum_{i=0}^{364} (1-p_8 \bar{T})^{364-i} (1-(1-p_7)^i) \left. \right\} \end{aligned} \quad (7.76)$$

$$\begin{aligned} \frac{\partial NEE_{ann}^{(n+1)}}{\partial p_3} = & - \left\{ ((1-p_5)^{365} - 1) + f_4 p_5 \sum_{i=0}^{364} (1-p_8 \bar{T})^{364-i} (1-p_5)^i \right\} \frac{\partial C_f^{(n)}}{\partial p_3} - \left\{ ((1-p_7)^{365} - 1) + f_4 p_7 \sum_{i=0}^{364} (1-p_8 \bar{T})^{364-i} \right. \\ & \times (1-p_7)^i \left. \right\} \frac{\partial C_r^{(n)}}{\partial p_3} - ((1-p_8 \bar{T})^{365} - 1) \frac{\partial C_{lit}^{(n)}}{\partial p_3} + 365 p_6 \frac{\partial C_w^{(n)}}{\partial p_3} - (1-p_2) \overline{GPP}(C_f^{(n)}) \left\{ \frac{(1-(1-p_5)^{365})}{p_5} \right. \\ & \times f_1 - \frac{(1-(1-p_7)^{365})}{p_7} p_4 f_3 - 365 (1-p_4) f_2 + f_1 f_4 \sum_{i=0}^{364} (1-p_8 \bar{T})^{364-i} (1-(1-p_5)^i) - p_4 f_3 f_4 \sum_{i=0}^{364} (1-p_8 \bar{T})^{364-i} \\ & \times (1-(1-p_7)^i) \left. \right\} - (1-p_2) \frac{\partial \overline{GPP}(C_f^{(n)})}{\partial p_3} \left\{ \frac{(1-(1-p_5)^{365})}{p_5} p_3 f_1 + \frac{(1-(1-p_7)^{365})}{p_7} p_4 (1-p_3) f_3 + 365 (1-p_4) \right. \\ & \times (1-p_3) f_2 + p_3 f_1 f_4 \sum_{i=0}^{364} (1-p_8 \bar{T})^{364-i} (1-(1-p_5)^i) + p_4 (1-p_3) f_3 f_4 \sum_{i=0}^{364} (1-p_8 \bar{T})^{364-i} (1-(1-p_7)^i) \left. \right\} \end{aligned} \quad (7.77)$$

$$\begin{aligned} \frac{\partial NEE_{ann}^{(n+1)}}{\partial p_4} = & - \left\{ ((1-p_7)^{365} - 1) + f_4 p_7 \sum_{i=0}^{364} (1-p_8 \bar{T})^{364-i} (1-p_7)^i \right\} \frac{\partial C_r^{(n)}}{\partial p_4} - ((1-p_8 \bar{T})^{365} - 1) \frac{\partial C_{lit}^{(n)}}{\partial p_4} \\ & + 365 p_6 \frac{\partial C_w^{(n)}}{\partial p_4} - (1-p_2) \overline{GPP}(C_f^{(n)}) \left\{ \frac{(1-(1-p_7)^{365})}{p_7} (1-p_3)f_3 - 365(1-p_3)f_2 \right. \\ & \left. + (1-p_3)f_3 f_4 \sum_{i=0}^{364} (1-p_8 \bar{T})^{364-i} (1-(1-p_7)^i) \right\} \end{aligned} \quad (7.78)$$

$$\begin{aligned} \frac{\partial NEE_{ann}^{(n+1)}}{\partial p_5} = & - \left\{ ((1-p_5)^{365} - 1) + f_4 p_5 \sum_{i=0}^{364} (1-p_8 \bar{T})^{364-i} (1-p_5)^i \right\} \frac{\partial C_f^{(n)}}{\partial p_5} - \left\{ -365(1-p_5)^{364} \right. \\ & \left. + f_4 \sum_{i=0}^{364} (1-p_8 \bar{T})^{364-i} ((1-p_5)^i - p_5 i (1-p_5)^{i-1}) \right\} C_f^{(n)} \\ & - \left\{ ((1-p_7)^{365} - 1) + f_4 p_7 \sum_{i=0}^{364} (1-p_8 \bar{T})^{364-i} (1-p_7)^i \right\} \frac{\partial C_r^{(n)}}{\partial p_5} \\ & - ((1-p_8 \bar{T})^{365} - 1) \frac{\partial C_{lit}^{(n)}}{\partial p_5} + 365 p_6 \frac{\partial C_w^{(n)}}{\partial p_5} - (1-p_2) \overline{GPP}(C_f^{(n)}) \left\{ \left(365(1-p_5)^{364} \right. \right. \\ & \left. \left. - \frac{(1-(1-p_5)^{365})}{p_5} \right) \frac{p_3 f_1}{p_5} + p_3 f_1 f_4 \sum_{i=0}^{364} (1-p_8 \bar{T})^{364-i} (i(1-p_5)^{i-1}) \right\} - (1-p_2) \frac{\partial \overline{GPP}(C_f^{(n)})}{\partial p_5} \\ & \times \left\{ \frac{(1-(1-p_5)^{365})}{p_5} p_3 f_1 + \frac{(1-(1-p_7)^{365})}{p_7} p_4 (1-p_3) f_3 + 365(1-p_4)(1-p_3) f_2 \right. \\ & \left. + p_3 f_1 f_4 \sum_{i=0}^{364} (1-p_8 \bar{T})^{364-i} (1-(1-p_5)^i) + p_4 (1-p_3) f_3 f_4 \sum_{i=0}^{364} (1-p_8 \bar{T})^{364-i} (1-(1-p_7)^i) \right\} \end{aligned} \quad (7.79)$$

$$\frac{\partial NEE_{ann}^{(n+1)}}{\partial p_6} = 365 \left(C_w^{(n)} + p_6 \frac{\partial C_w^{(n)}}{\partial p_6} \right) \quad (7.80)$$

$$\begin{aligned} \frac{\partial NEE_{ann}^{(n+1)}}{\partial p_7} = & - \left(\left\{ -365(1-p_7)^{364} + f_4 \sum_{i=0}^{364} (1-p_8\bar{T})^{364-i} ((1-p_7)^i - p_7 i (1-p_7)^{i-1}) \right\} C_r^{(n)} \right. \\ & -365p_6 \frac{\partial C_w^{(n)}}{\partial p_7} + (1-p_2) \overline{GPP}(C_f^{(n)}) \left\{ \left(-365(1-p_7)^{364} - \frac{(1-(1-p_7)^{365})}{p_7} \right) \frac{p_4(1-p_3)f_3}{p_7} \right. \\ & \left. \left. + p_4(1-p_3)f_3f_4 \sum_{i=0}^{364} (1-p_8\bar{T})^{364-i} (i(1-p_7)^{i-1}) \right\} \right) \end{aligned} \quad (7.81)$$

$$\begin{aligned} \frac{\partial NEE_{ann}^{(n+1)}}{\partial p_8} = & - \left(-f_4p_5 C_f^{(n)} \sum_{i=0}^{364} (364-i)(1-p_8\bar{T})^{363-i} (1-p_5)^i \bar{T} - f_4p_7 C_r^{(n)} \sum_{i=0}^{364} (364-i)(1-p_8\bar{T})^{363-i} (1-p_7)^i \bar{T} \right. \\ & -365(1-p_8\bar{T})^{364} \bar{T} C_{lit}^{(n)} + ((1-p_8\bar{T})^{365} - 1) \frac{\partial C_{lit}^{(n)}}{\partial p_8} - (1-p_2) \overline{GPP}(C_f^{(n)}) \left\{ p_3f_1f_4 \sum_{i=0}^{364} (364-i) \right. \\ & \times (1-p_8\bar{T})^{363-i} \bar{T} (1-(1-p_5)^i) + p_4(1-p_3)f_3f_4 \sum_{i=0}^{364} (364-i)(1-p_8\bar{T})^{363-i} \bar{T} (1-(1-p_7)^i) \left. \right\} \right) \end{aligned} \quad (7.82)$$

178

$$\begin{aligned} \frac{\partial NEE_{ann}^{(n+1)}}{\partial p_{11}} = & - \left\{ ((1-p_5)^{365} - 1) + f_4p_5 \sum_{i=0}^{364} (1-p_8\bar{T})^{364-i} (1-p_5)^i \right\} \frac{\partial C_f^{(n)}}{\partial p_{11}} - \left\{ ((1-p_7)^{365} - 1) \right. \\ & + f_4p_7 \sum_{i=0}^{364} (1-p_8\bar{T})^{364-i} (1-p_7)^i \left\} \frac{\partial C_r^{(n)}}{\partial p_{11}} - ((1-p_8\bar{T})^{365} - 1) \frac{\partial C_{lit}^{(n)}}{\partial p_{11}} + 365p_6 \frac{\partial C_w^{(n)}}{\partial p_{11}} \\ & - (1-p_2) \frac{\partial \overline{GPP}(C_f^{(n)})}{\partial p_{11}} \left\{ \frac{(1-(1-p_5)^{365})}{p_5} p_3f_1 + \frac{(1-(1-p_7)^{365})}{p_7} p_4(1-p_3)f_3 + 365(1-p_4)(1-p_3)f_2 \right. \\ & \left. + p_3f_1f_4 \sum_{i=0}^{364} (1-p_8\bar{T})^{364-i} (1-(1-p_5)^i) + p_4(1-p_3)f_3f_4 \sum_{i=0}^{364} (1-p_8\bar{T})^{364-i} (1-(1-p_7)^i) \right\} \end{aligned} \quad (7.83)$$

There are no derivatives for parameters p_1 and p_9 as they have been neglected. The derivative for parameter p_{10} is very complicated due to the fact that it is embedded in $(1 - p_8 \bar{T})^{364-i}$ and has therefore been omitted. The only parameters for which the derivatives are straightforward are p_4 , p_6 , p_7 , p_8 and p_{11} . Therefore, as we now have a sensible analytical derivative for p_8 (although it is by no means simple), we have improved on the expansion of the original NEE equation. We found a derivative for p_6 , but no derivative for p_9 .

We have made several approximations to come to this result:

- $GPP(C_f(t), t)$ is averaged over one year
- $T(t)$ is averaged over one year
- C_w is approximated as in equation (7.66)
- p_1 is ignored
- C_{som} is regarded as constant, thereby disappearing from the equation and thus we effectively ignore p_9
- fitting factors are used.

Fig. 7.9 shows the fit between the different approximations to the annual NEE from the initial conditions for the pools as mentioned in Chapter 3.1. We cannot distinguish between the annual NEE calculated from the original equation (7.8) and the annual NEE calculated from expanding the original NEE, even though we have neglected p_1 and regarded C_{som} as a constant. The annual NEE calculated by summing all the carbon pools over one year is a reasonable fit, although it overestimates the annual NEE in the first 17 years and underestimates it in the last 13 years.

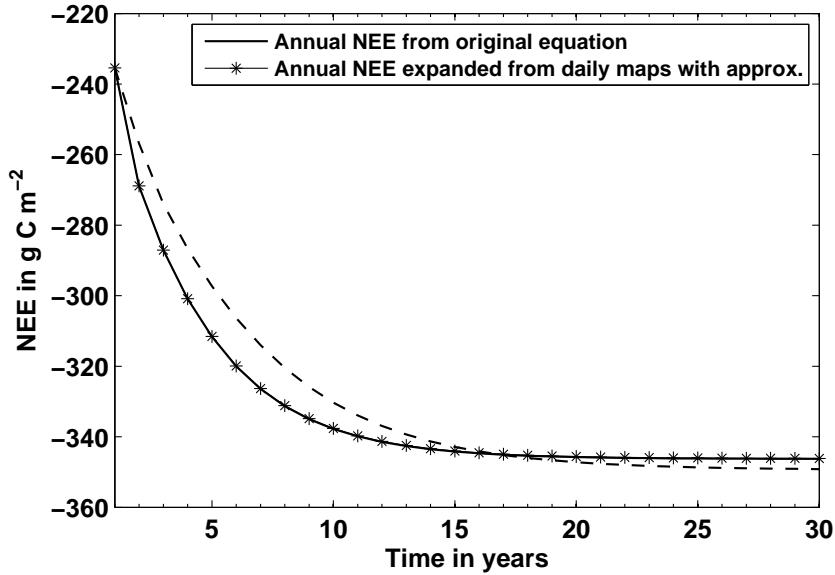


Figure 7.9: *The NEE expressed in three different ways between 0 and 30 years.*

Comparison with REFLEX Results

It is interesting to compare our findings with the findings of the paper by Fox *et al.* [13] regarding the REFLEX project, particularly as one of the REFLEX project's objectives was to try to constrain parameters from different data assimilation techniques and we are claiming that the model itself should give information about the sensitivity of the observables (LAI and NEE) to parameters.

Fig. 7.10 shows the REFLEX parameter estimation results for DALEC evergreen synthetic data (data not from measurements). This particular table is the most appropriate one to look at because it represents the results of using data assimilation to predict parameter values based on synthetic data generated by DALEC using the same forest and climate data that we have used (from the Loobos forest). The vertical line in each of the panels in Fig. 7.10 represents the true value of the parameter. The circles in the panels represent the values of the parameters as estimated by the different data assimilation algorithms (indicated on the *y*-axis) and the brackets around the circles indicate the size of the 90%

confidence intervals. The upper and lower bounds of each parameter are shown on the x -axes. For all parameters beginning with T the x -axes are log-scaled. We can see from Fig. 7.10 that the parameters which are best estimated are p_2 (Fg), p_3 (Fnf), p_4 (Fnrr), p_5 (Tf), p_8 (Tl) and p_{10} (Et). The parameters which are worst estimated are p_1 (Td), p_6 (Tw) and p_9 (Ts). The other parameters p_7 (Tr) and p_{11} (Pr) are somewhere in the middle of the other two groups.

If we compare this with our sensitivity analysis and also take into account our equations for the sum of NEE over a year, equations (7.45) and (7.75), we can make some sense out of why some parameters can be estimated better than others. For example, in our numerical sensitivity analysis (see Table 7.2), we have shown that NEE is most sensitive to parameters p_2 (fraction of GPP respired), p_3 fraction of NPP allocated to foliage), p_5 (daily turnover rate of foliage) and p_{11} (N use efficiency parameter) and least sensitive to parameters p_1 decomposition rate and p_6 (daily turnover rate of wood). This agrees with the REFLEX results. Parameters p_2 , p_3 and p_5 not only appear in equations (7.45) and (7.75), but are also the main parameters for C_f and GPP. Parameter p_{10} (Parameter in exponential term of temperature dependent rate) appears to be well constrained by most DA algorithms. It also appears many times in the NEE equation. Parameter p_1 does not appear at all in our equations (7.75) and (7.45). The reason for this is that parameter p_1 appears in the equation for the C_{lit} pool in a sum with p_8 , (mineralisation rate of litter) which is of a much higher order than p_1 . Parameter p_6 has been neglected in equation (7.45) as it is part of the C_{som} pool, which we ignore, as it is constant. Parameter p_9 (mineralisation rate of C_{som}) is the coefficient of a constant of $\mathcal{O}(10^4)$ and so if p_9 is small (as it is for the Loobos forest, $\mathcal{O}(10^{-6})$) and changes by a small amount, the resulting term is still only $\mathcal{O}(10^{-2})$. In equation (7.75) p_9 does not appear at all as this is the NEE expressed as the total sum of all carbon pools, which does not have C_{som} in it, as it is regarded as a constant.

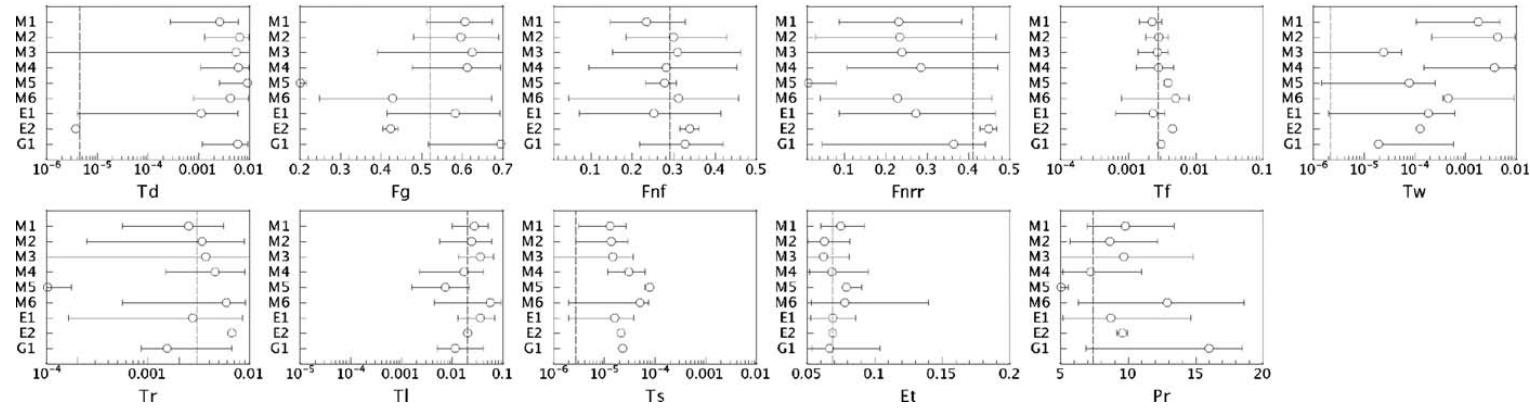


Figure 7.10: The parameter estimation for DALEC evergreen with synthetic data [13]. The vertical line in each of the panels represents the true value of the parameter. The circles are the values of the parameters as estimated by the different data assimilation algorithms (indicated on the y-axis) and the brackets around the circles indicate the size of the 90% confidence intervals. The upper and lower bounds of each parameter are shown on the x-axes. For all parameters beginning with T the x-axes are log-scaled.

This goes some way to explaining why NEE is not very sensitive to parameters p_1 , p_6 and p_9 and why they cannot be constrained very well. The parameters for which the sensitivity analysis produced mediocre results do play a role in equations (7.75) and (7.45), especially p_7 (daily turnover rate of roots) and p_8 , which appear in these equations several times.

REFLEX also used a set of metrics to describe the outcomes of the parameter estimation. For each parameter of the synthetic data three relative distance metrics were computed: d_1 , d_2 and d_3 .

d_1 : d_1 describes consistency among algorithms:

$$\sigma(m_1, \dots, m_9)/(P_{max} - P_{min}), \quad (7.84)$$

where σ is the standard deviation, m_i is algorithm i 's best estimate of the parameter, P_{max} is the pre-specified upper limit and P_{min} is the pre-specified lower limit of the parameter.

d_2 : d_2 tests to what degree the posterior estimate of the parameter is an improvement on the prior:

$$\mu(CI_1, \dots, CI_9)/(P_{max} - P_{min}), \quad (7.85)$$

where μ is a mean and CI_i is the width of the parameter's confidence interval for algorithm i .

d_3 : d_3 tests for consistency with the true value of the parameter:

$$|t - \mu(m_1, \dots, m_9)|/(P_{max} - P_{min}), \quad (7.86)$$

where t is the true value of the parameter.

We can see from the table in Fig. 7.11 that the parameters for which the value was best estimated (the lowest d_3 metric) are p_2 , p_3 , p_5 , p_7 , p_8 and p_{10} . The parameters that were best constrained (have the narrowest confidence intervals) are p_5 , p_8 , p_9 and p_{10} (the

Evergreen: EV-SYN			
Param	d_1	d_2	d_3
T_d	0.26	0.36	0.75
F_g	0.30	0.41	0.02
F_{nf}	0.07	0.49	0.00
F_{nrr}	0.24	0.65	0.31
T_f	0.06	0.20	0.03
T_w	0.22	0.40	0.69
T_r	0.27	0.52	0.03
T_l	0.07	0.22	0.03
T_s	0.05	0.16	0.21
E_t	0.04	0.24	0.00
P_r	0.21	0.47	0.15

Figure 7.11: Parameter estimation metrics using nine different algorithms based on synthetic data for evergreen forest [13]. Lower values indicate better results.

lowest d_2 metric) and most consistency amongst the algorithms was found for p_3 , p_5 , p_8 , p_9 and p_{10} .

It is interesting to notice that, according to the REFLEX project [13], p_9 (mineralisation rate of C_{som}) is regarded as well constrained with narrow confidence intervals. However, the d_3 metric indicates that consistency with the true value of p_9 was not very impressive and according to our results, NEE is not very sensitive to p_9 . On the other hand, according to the REFLEX paper, p_2 is poorly constrained, with wide confidence intervals, but the value of p_2 was found consistently.

One problem regarding the above paragraph is that it is not clear which starting values for the parameters were used by the nine groups. If the model is not very sensitive to a particular parameter, but the initial choice for this parameter is (unknowingly) very close to the truth, it may be possible to get a good result for this parameter for the wrong reasons. On the other hand, even though there may be consistency in algorithms or the confidence intervals are narrow, if the true value of a parameter is not found, what value do metrics d_1 and d_2 have? Our opinion is that there are weaknesses in the

interpretation of whether a parameter is well constrained and estimated or not and that it would probably have been appropriate to give all the participants the same starting value for the parameters. The intervals from which the participants could estimate the parameter values were very large [13].

NEE in Phase II

Phase II covers the period from when C_f , C_r and C_{lit} have reached their fixed points, which, with the data that we have been using is approximately 30 years until about 1000 years. This means that we could expand the NEE again from its original equation, but now, instead of using a constant value of C_{som} , including the daily map for the C_{som} . This will become very complicated, so therefore we express the NEE again as the total sum of all the carbon pools, but now C_f , C_r and C_{lit} are regarded as constants and C_{som} is not:

$$NEE_{ann}^{(n+1)} = - \left(C_w^{(n+1)} - C_w^{(n)} + C_{som}^{(n+1)} - C_{som}^{(n)} \right). \quad (7.87)$$

for $n \geq 0$.

Recall the equation for $C_{som}(t)$:

$$C_{som}(t+1) = (1 - p_9 T(t)) C_{som}(t) + p_6 C_w(t) + p_1 T(t) C_{lit}(t) \quad (7.88)$$

We regard C_{lit} as a constant at this time (as it has reached its fixed point) and as this contribution is very small, of $\mathcal{O}(10^{-4})$, we ignore it. Therefore we can write the annual

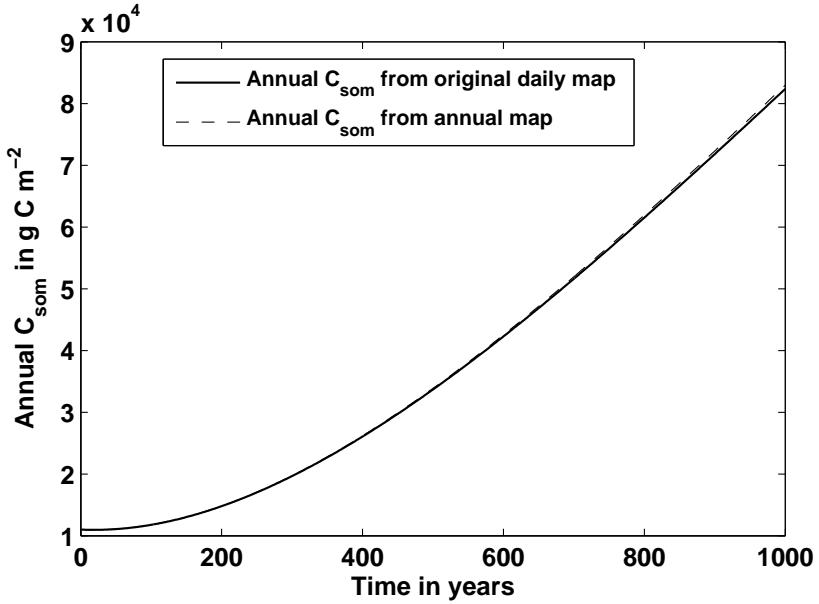


Figure 7.12: The annual map for C_{som} plotted from the original daily map and as an annual map, no extra fitting factor.

map for C_{som} as follows:

$$\begin{aligned}
 C_{som}^{(n+1)} = & (1 - p_9 \bar{T}) C_{som}^{(n)} + \left[p_6 (1 - p_6)^{365} C_w^{(n)} \right. \\
 & \left. - 365 p_6 (1 - p_2) (1 - p_3) (1 - p_4) f_2 \overline{GPP}(C_f^{(n)}) \right] \\
 & \times \sum_{i=0}^{364} (1 - p_9 \bar{T})^{364-i}
 \end{aligned} \tag{7.89}$$

From Fig. 7.12 we can see that we do not need an extra fitting factor in this case, as the fit is very good, even though we have ignored the contribution of $p_1 \bar{T} C_{lit}$. In fact, this approximation is valid from year 0 onwards.

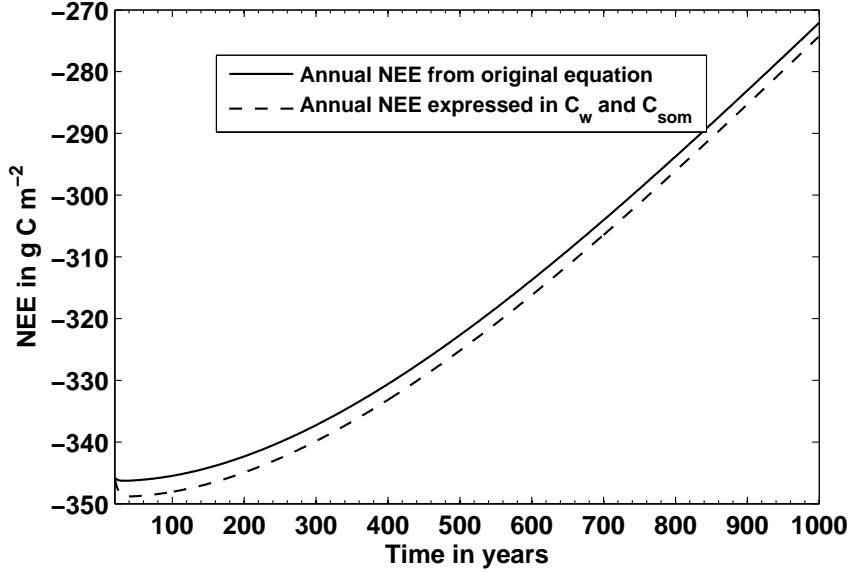


Figure 7.13: The NEE expressed in two different ways between 30 and 1,000 years.

Therefore the annual NEE can be written as:

$$\begin{aligned}
 NEE_{ann}^{(n+1)} = & - \left\{ \left((1 - p_9 \bar{T})^{365} - 1 \right) C_{som}^{(n)} - (1 - p_2)(1 - p_3)(1 - p_4)f_2 \right. \\
 & \times GPP(C_f^{(n)}) \left[\left(365 \left(p_6 \sum_{i=0}^{364} (1 - p_9 \bar{T})^{364-i} \right) + 1 \right) \right] \\
 & \left. + \left(p_6(1 - p_6)^{365} C_w^{(n)} \right) \sum_{i=0}^{364} (1 - p_9 \bar{T})^{364-i} + 365p_6C_w^{(n)} \right\}, \quad (7.90)
 \end{aligned}$$

which is much simpler than the previous formula.

Fig. 7.13 shows the annual NEE as calculated from its original formula and the annual NEE in terms of C_w and C_{som} only. We can see that equation (7.90) underestimates the original NEE by less than 1%.

We could again write down the derivatives for NEE with respect to the parameters, but we would not gain very much, due to the recurrence relation between the parameters and the pools.

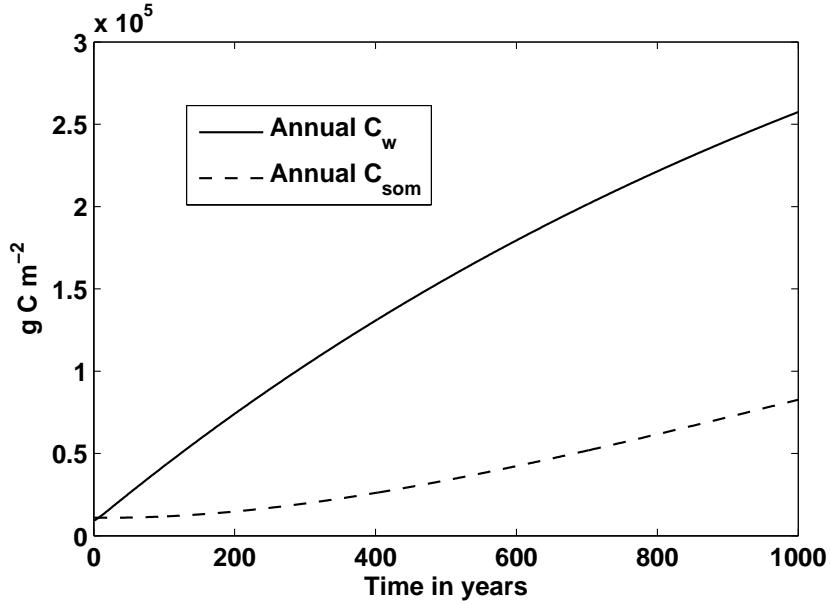


Figure 7.14: The annual C_w and C_{som} during Phase II; the growth of C_w is much steeper than the growth of C_{som} in the first 150 years or so.

We can however look at the contributions of the different terms to the NEE. To this end we consider that approximately in the first 150 years of Phase II, C_w grows much faster than C_{som} , see Fig. 7.14. After that the difference between their growth rates becomes less. Fig. 7.15 shows that the NEE expressed in terms of GPP only: $-365(1 - p_2)(1 - p_3)(1 - p_4)\overline{GPP}(C_f^{(n)})$, which is a term from $C_w^{(n+1)} - C_w^{(n)}$, accounts for the most of the NEE in the first approximately 150 years of Phase II; we call this period Phase IIa. Therefore we can write the annual NEE in Phase IIa as follows:

$$NEE_{ann}^{(n+1)} = -365f_2(1 - p_2)(1 - p_3)(1 - p_4)\overline{GPP}(C_f^{(n)}). \quad (7.91)$$

This is interesting, as the difference between the annual NEE calculated from the original equation and the annual NEE calculated as per equation (7.91) is less than 1% (overestimating) and therefore equation (7.91) could be used as a rough estimation of NEE during Phase IIa. This is a great, but nevertheless valid simplification of the NEE, which depends only on p_2 , p_3 , p_4 , p_5 and p_{11} .

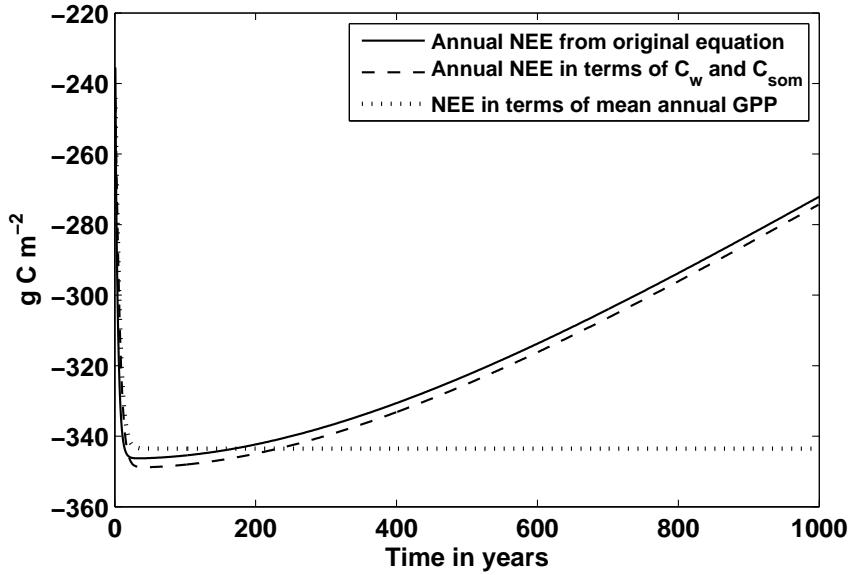


Figure 7.15: The annual NEE calculated from the original daily maps, using yearly averages for NEE, and annual NEE in terms of mean GPP.

A numerical sensitivity calculation for equations (7.90) and (7.91) performed at year 1 and at year 100 compared with the numerical sensitivity calculation of the original NEE equation at year 1 and year 100 in Table 7.4 shows that at year 1 equations (7.90) and (7.91) are not suitable: the values for year 1 using the original equation for NEE and year 1 using equations (7.90) and (7.91) are completely different. However at year 100 we can see that these inequalities have smoothed out considerably. When we consider the original NEE and equation (7.90) we can see that the largest increase in sensitivity between years 1 and 100 is for p_6 , which is 0.019 at year 100 as opposed to 0.00 at year 1.

Hill *et al.* [17] showed that increasing the length of the observation time series has a positive effect on the constraining of parameters. Perhaps the above results in Table 7.4 give an indication as to why this may be so. Increasing the time span increases the sensitivity of the NEE to certain parameters, which may make it easier to retrieve them. Interestingly, the sensitivity of the original NEE equation to parameter p_9 did not increase at all. Therefore it may be difficult to retrieve this parameter. On the other hand, the

	NEE original		NEE in terms of C_w and C_{som}		NEE in terms of GPP	
	1 year	100 years	1 year	100 years	1 year	100 years
p₂	8.24	14.67	4.04	14.67	0.40	14.69
p₃	-2.45	-8.52	-0.29	-8.52	-0.29	-8.50
p₄	0.48	2.48	1.70	2.47	1.70	2.48
p₅	3.43	10.02	1.22	10.02	1.22	10.03
p₆	0.00	0.019	0.00	0.019	0.00	0.00
p₇	0.95	0.00	0.00	0.00	0.00	0.00
p₈	0.43	0.00	0.00	0.00	0.00	0.00
p₉	0.11	0.11	0.11	0.11	0.00	0.00
p₁₀	0.34	0.09	0.09	0.09	0.00	0.00
p₁₁	-3.76	-6.15	-1.84	-6.15	-1.85	-6.16

Table 7.4: Numerical sensitivity as calculated by $\frac{\partial NEE_{ann}^{(n+1)}}{\partial p_i} \frac{p_i}{100}$ for equations (7.8) and (7.90) at year 1 and year 100. The calculations are done on the total of NEE accumulated in year 1 and in year 100.

sensitivity to NEE for p_7 , p_8 and p_{10} seems to decrease, indicating that increasing the time period of observation may not improve the possibility of constraining these parameters.

A Very Simple Model to Determine NEE

An interesting observation is that, for the type of forest we have used (a sink forest) we can combine some simplified equations and form a very simple carbon model, which can be used in Phase I to predict future NEE in Phase IIa. Phase IIa starts when the fast pools have reached their fixed point, which in turn depends on the initial conditions of the fast pools.

Recall from Section 5.4 that the equation for the linear map of C_f is:

$$C_f^{(n)} = C_f^* + f'(C_f^*)^n (C_f^{(0)} - C_f^*) \quad (7.92)$$

The known quantities are $C_f^{(0)}$ and $C_f^{(n)}$, which are observations measured at one particular day of each year. It is possible to calculate C_f^* and $f'(C_f^*)$ if we have two or more observations. Once we know the fixed point for C_f , we can calculate the average annual GPP, which can then be substituted in the simplified NEE equation.

With the use of a simple 4DVAR data assimilation algorithm, fitting the linear model to the annual C_f by estimating the unknowns, we found that this equation works dependent on the amount of data and the age of the forest (therefore also depending on the initial conditions of the fast pools). Annual observations of C_f early on in the life of a forest will not produce the correct output, see Fig. 7.16. This figure shows the annual C_f as calculated from equation (5.35) using the Loobos data (in the plot this is called 'Original annual C_f '). The five stars show five observations taken from that annual map of C_f . The DA algorithm uses those five observations and (7.92) to find an optimal fit. The dashed line shows this optimal fit. We can see that it greatly overestimates the fixed point. However, if the age of the forest is known and it is deemed to be approximately half way or longer through Phase I, this equation will produce approximately correct unknowns with as few as three observations of the annual C_f , see Fig. 7.17.

So, for an established sink forest, a very simple model to approximate future NEE consists of two equations:

$$C_f^{(n)} = C_f^* + f'(C_f^*)^n(C_f^{(0)} - C_f^*) \quad (7.93)$$

$$NEE_{ann}^{(n+1)} = -365(1 - p_4)(1 - p_3)(1 - p_2)\overline{GPP}(C_f^{(n)}). \quad (7.94)$$

For a source forest we would need to use a more complete form of NEE, in terms of both C_{som} and C_w .

Although this model will probably not be suitable for ecologists, it may be valid to use in cases where one is interested in how much CO₂ is taken up or expelled by a forest for

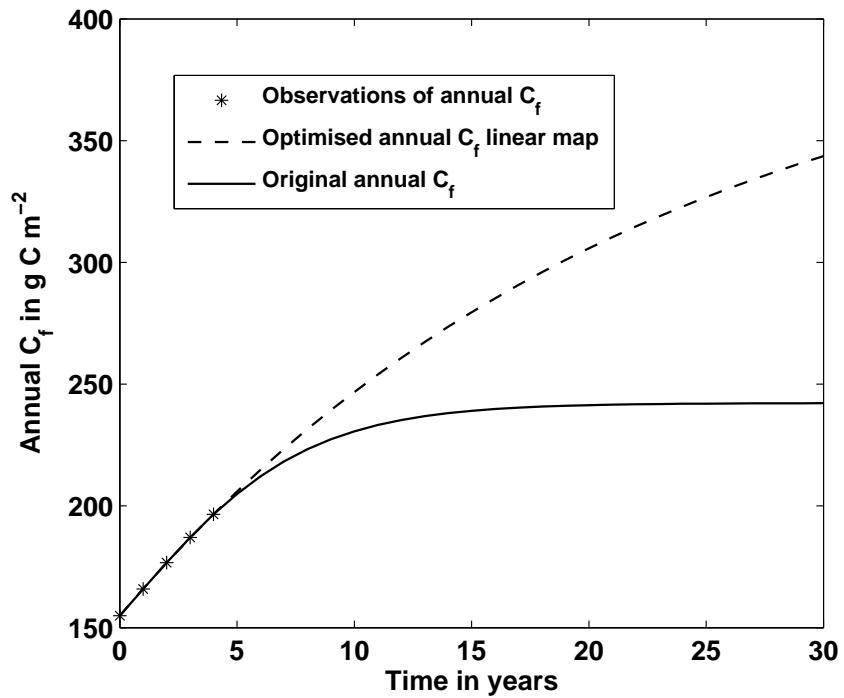


Figure 7.16: *With observations too early in the life of the forest, (7.92) underestimates the correct values for the unknowns, C_f^* and $f'(C_f^*)$.*

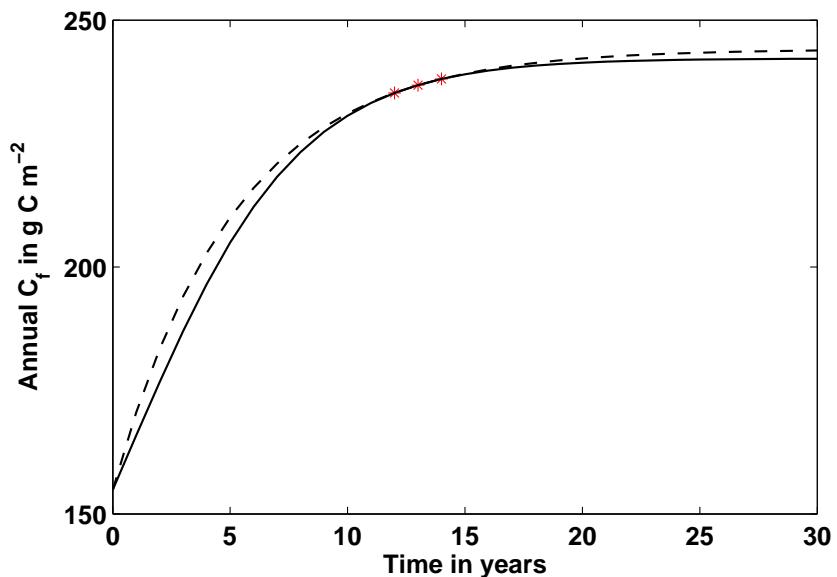


Figure 7.17: *With observations from approximately half way through phase I, (7.92) produces the approximate correct values for C_f^* and $f'(C_f^*)$.*

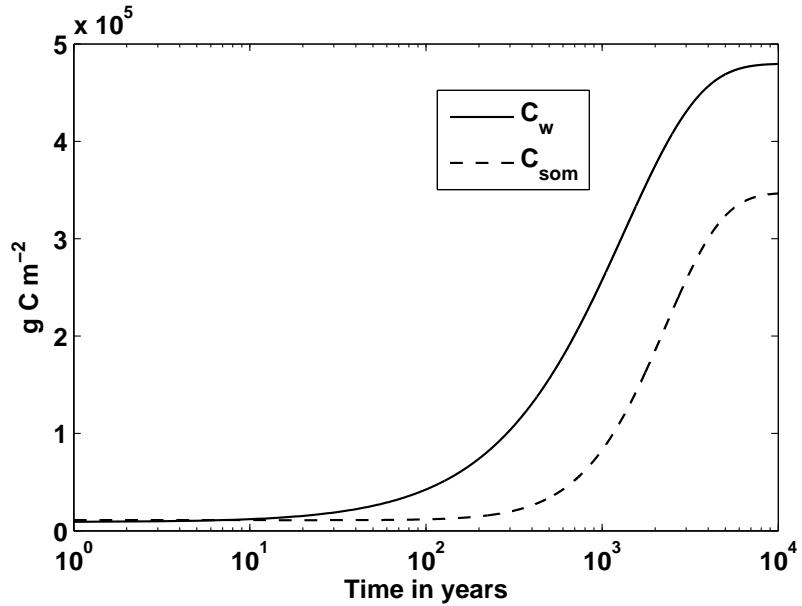


Figure 7.18: The growth of C_{som} and C_w on a log scale.

climate modelling purposes.

NEE in Phase III

In phase III, which is the period after 1,000 years until NEE reaches its fixed point, C_w starts growing much faster, which in turn causes C_{som} also to grow much faster, see Fig. 7.18, which shows the growth of both pools on a log scale.

If we compare Fig. 7.18 with Fig. 7.4, we can see that this correlates very well; between 1,000 and 10,000 years NEE starts to grow due to increased growth of C_w and C_{som} and finally reaches its fixed point when both C_w and C_{som} have reached their fixed points at around 10,000 years.

During this phase we need to express the annual NEE in terms of both C_w and C_{som} and cannot simplify this any further. Clearly this range is beyond any practical interest.

7.3.3 Conclusion

In this chapter we have performed a numerical sensitivity analysis for NEE and C_f , which showed that there is more or less sensitivity to all p -parameters, apart from p_1 and p_6 . The insensitivity to parameters p_1 and p_6 could be related to the fact that they are very small, $\mathcal{O}(10^{-6})$. A percentage change in such values cannot have much effect. As the initial conditions and fixed points of the five carbon pools widely differ we rescaled the pools, so that we could get a clear idea of the contributions of the carbon pools to each other and to the NEE.

We then examined what happens to the NEE over time, which we divided into three phases:

1. Phase I: the period between 0 and 30 years (when C_f , C_r and C_{lit} approximately reach their fixed points for initial conditions and parameters for the Loobos site). All the pools are growing, but C_{som} seems almost constant.
2. Phase II: the period between 30 years and 1,000 years, when C_f , C_r and C_{lit} have reached their fixed points, but C_w and C_{som} are still growing, although C_w is growing faster than C_{som} .
 - Phase IIa: the period between 30 and 150 years, when C_w is growing much faster than C_{som} .
3. Phase III: the period between 1,000 and 10,000 years, when C_w is starting to grow much faster, which causes C_{som} also to grow faster.

In Phase I we tried to find an analytical expression for the sensitivity of NEE to the various parameters. Therefore we expanded and simplified the NEE from its daily map, neglecting p_1 , as p_1 was used in a sum with p_8 , where $p_1 \ll p_8$ and regarding C_{som} as a constant, effectively neglecting p_6 . The resulting expression for the sum of the NEE

over one year (which we call the annual NEE), equation (7.45) is linear in p_4 , nonlinear in all the other parameters and we could determine which parts of the NEE contribute positively or negatively. The role of some parameters became clearer but the analytical sensitivity of NEE to the parameters could still not be determined fully as it was difficult or impossible to write down an analytical derivative of the NEE with respect to some of the parameters. The only parameters for which the formulae for the derivatives of NEE with respect to the parameters were relatively straightforward are p_4 , p_7 , p_9 and p_{11} . Parameters p_8 and p_{10} are embedded in $A_{i,t}$ and therefore it was difficult to write down their derivatives analytically. The derivatives for p_2 , p_3 and p_5 were complicated by the fact that they include a recurrence relation in C_f . Therefore our next step was to write the annual NEE as the sum of all the carbon pools, expressed as annual maps. This notation introduced a lot of approximation and averaging, but was found to be valid and had the advantage that it now included C_w . However, an analytical expression for the sensitivity of the NEE with respect to the parameters was still not very good. Although we now managed to obtain an analytical derivative for parameter p_8 , the derivatives were still complicated and some had recurrence relations in the equations for the pools. Even though we gained a relatively straightforward expression for the derivative of NEE with respect to parameter p_6 , we lost the one for parameter p_9 , as, due to the assumption that C_{som} is constant during this phase, it was not included in this expression for the annual NEE.

When we compared our results with the REFLEX results, which are results from Phase I, we found that they largely agreed, but we stated that there are weaknesses in the REFLEX interpretation of whether a parameter is well constrained or not, as it was not clear which starting values for the parameters were used by the REFLEX participants. We noted that it would have been appropriate to give all participants the same starting values for the parameters.

In Phase II C_f , C_r and C_{lit} have reached their fixed points, so the only influence on NEE

are C_w and C_{som} and therefore we could now write the annual NEE in terms of the total of the C_w and C_{som} pools (as annual maps). Doing so we discovered that, due to the faster growth rate of C_w during the period between approximately 30 and 150 years, which we call Phase IIa, the main term determining the NEE is:

$$NEE_{ann}^{(n+1)} = -365(1 - p_4)(1 - p_3)(1 - p_2)\overline{GPP}(C_f^{(n)}), \quad (7.95)$$

which means that for this time period we can greatly reduce the complexity of the expression for NEE.

We performed a numerical sensitivity analysis of the NEE for year 1 and for year 100, using the equations we found for NEE in terms of C_w and C_{som} and in terms of the mean annual GPP only and discovered that the sensitivity of the NEE to some parameters increases over time, whilst for others it decreases. It also confirmed that our simplifications of the NEE are valid and correct for this period in time.

From a simplification of the annual C_f in Section 5.4 and the simplification of NEE, which can in Phase IIa be written in terms of the annual mean GPP only, we found a simpler model than DALEC EV to estimate and predict future NEE, which could be useful for climate modelling purposes.

In Phase III we find that both C_w and C_{som} are growing much faster, reaching their fixed points at around year 10,000, which is the same time at which the NEE reaches its fixed point.

Although the research in this chapter did not result in achieving our initial objectives, namely to be able to write the sensitivity of the NEE with respect to all the parameters in a useful analytical form, which was hampered by the fact that there are recurrence relationships between the parameters and the pools, important discoveries are that:

1. In Phase II the annual NEE can be expressed in terms of C_w and C_{som} only;
2. In Phase IIa the annual NEE can be expressed in terms of the annual mean GPP only;
3. The sensitivity of NEE to some parameters involved increases over time, which may be of benefit in constraining parameters (sensitivity to other parameters decreases);
4. We can define an even simpler model of two equations and six parameters to estimate the annual NEE in Phase IIa.

According to Lemma 7.2.1 $NEE = 0$ at a fixed point of the carbon pools. However, we have seen that, for the parameter values used for the Loobos, according to the DALEC model, it takes approximately 10,000 years to achieve the final fixed point due to the slow growth of C_w and C_{som} . Thus, even when C_f , C_r and C_{lit} have reached their fixed points, the forest used in our calculations (Loobos) is still a sink for a long time.

To establish whether a forest is at a total equilibrium, measurements of all the pools would need to be taken over a large amount of years. It can be assumed that a forest is at a pseudo steady state when the measurements of the fast scale pools are equal over a number of years. The timescale for a forest to reach absolute steady state is dependent on the initial conditions from all the pools. From the copious literature available it is not clear how long it could take the slow pools, and in particular the soil pool, to reach their equilibria in reality as there are many opposing statements and opinions.

Chapter 8

Conclusion

8.1 Overview of the Problem and Results

The continual increase in computational power allows ever-more detailed studies of Earth systems, such as the carbon cycle. Many complex models exist nowadays which are analysing and predicting carbon stocks and stores, but the underlying dynamics of the carbon cycle itself is not well understood. One important aspect of this is that the values of many of the parameters that appear in the models are not known with confidence.

Various projects have been designed to use techniques from data assimilation in an attempt to constrain the values of the parameters. Most modern data assimilation schemes use the forward model as part of the algorithm, and it is of vital importance to understand the nature of the system dynamics to be able to employ the data assimilation methods with any confidence. Challenges are presented to data assimilation schemes due to the presence of multiple timescales, either in the frequency of observations or in the dynamics of the forward model. It is therefore important to understand whether a problem such as the carbon cycle is amenable to scrutiny using such methods and, in particular, how

the intrinsic separation of timescales, from days (for foliar carbon) to millennia (for soil organic matter), presents fundamental challenges to data assimilation schemes.

In this thesis we have focussed on DALEC, a very simple model of the terrestrial carbon cycle. In Chapter 4 we have shown how the generic behaviour of the DALEC EV model in the short run is controlled by the foliar carbon pool parameters p_2 (fraction of GPP respired), p_3 (fraction of NEE allocated to foliage), p_5 (daily turnover rate of foliage) and p_{11} (N use efficiency parameter used in ACM) and that these parameters, which reflect the circumstances surrounding the forest, play a crucial role in the survival of the forest. The dynamics of the foliar pool show a tipping point, which is dependent on the value of these parameters and explains how, for certain values of the parameters, according to the model, a forest is expected to die out, without any chance of coming back to life. Likewise, for other values of the parameters, and depending on the initial values of the foliar pool, a forest grows to an annual cycle and lives. This happens because in DALEC EV the forest's dynamics depend on the foliar carbon pool, therefore if there is no photosynthesis and the foliar turnover is too high (which could be caused for example by a severe long lasting frost or drought), the forest dies out. In reality a forest has other resources, such as seeds, that may resurrect or replenish a forest.

An interesting feature of the tipping point structure, which can be seen from the plot showing the line of tipping points, Fig. 4.3 in Section 4.3, is that it becomes possible to see how close model parameter values bring a forest to a tipping point. In two cases we found that the forests in question were quite near to a tipping point. This brings up two questions: (1) is this a common occurrence? and if so, (2) would the values of parameters change when circumstances and conditions change and move the forest close to the tipping point again when the area of possible growth and stabilisation on the left hand side of the line of tipping points increases, for example through an increase of CO₂?

Tipping points are a topic of interest in climate studies and are well known mechanisms

which could lead to sudden climate change. Although many subsystems may be involved in climate tipping, many tipping points seem to be caused by one particular sub-system. Lenton *et al.* [26] listed nine tipping elements considered to be possible causes of climate tipping, such as the a collapse of the Greenland ice sheet and the loss of Arctic summer sea-ice. These subsystems slowly evolve over time until the tipping point is reached. In the DALEC model reaching the tipping point could lead to the forest's death. From the literature on DALEC [48, 51, 50, 49, 13] it is not clear whether the parameters which cause the bifurcation structure in the evergreen or the deciduous version are fixed once set or change over time, so therefore the bifurcation structure in the DALE C model caused by parameters p_5 (daily turnover rate of foliage) in the evergreen version and p_{14} (fraction of leaf loss transferred to litter) in the deciduous version could either mean a slow change over time in parameter values which eventually causes a tipping point or, if the parameter values do not change over time, it gives us information on which parameter values are sensible to choose if we are modelling a sustainable forest.

A fixed point analysis of the pools showed that in the DALEC EV model, although the foliar, root and litter pools reach their fixed point fairly soon (tens of years), the wood and soil and organic matter pools take much longer (ten thousands of years), due to the very small value of parameters p_6 (daily turnover of wood) and p_9 (daily mineralisation rate of SOM/CWD). Therefore, although the fast pools determine if a forest survives or lives in the short run, the wood and soil pools determine in the long run whether a forest is a sink or a source. In DALEC the values of parameters p_6 and p_9 have an effect on how long it takes for a forest to reach the point where it is no longer a sink or a source.

We have also shown in Chapter 4 that in DALEC it may not be necessary to use highly variable climate drivers for forests that live in a steady climate, but that it is sufficient to represent the drivers by a smooth cosine wave, where all that is needed is to measure the amplitude, mean and phase of an annual cycle. We acknowledge that according to Medvighy [32] smoothing the drivers takes away the natural variability to which ecosystems

respond. However, smoothing drivers in DALEC did not make much of a difference, most likely due to the fact that the highest variability has already been taken out by providing the model with daily climate data, rather than with data collected on a shorter time scale. We did find however, that the model seems to be more sensitive to solar radiation than to temperature, which according to Medvigy has the greatest impact.

Having used periodic climate data in our research, we acknowledge that in reality climate data will not be periodic but data could be seen as periodic data with noise, producing noisy fixed points. However, we expect that this will not make a difference to the qualitative behaviour we have found.

A short preliminary study of the DALEC DE model, an example of a slightly more complex model due to leaf loss in the autumn and the requirement of an extra pool, showed a similar tipping point structure to the DALEC EV model. In contrast to DALEC EV, DALEC DE involves two equations which decouple from the other pools and therefore control the dynamics of the forest, as well as more parameters, requiring closer scrutiny.

Our next focus was on the GPP, a complex function representing the photosynthesis process, which we studied and simplified in Chapter 5. Our aim was to lay bare the dependence of the GPP on C_f . By simplifying certain elements in the GPP we managed to express it in a much simpler form, namely a fraction with a quadratic equation in C_f , the foliar pool, in both the numerator and the denominator, clearly showing the non-linearity of this function. Another aim was to simplify the process of finding a limit point structure without the help of continuation software. To this end we expressed the foliar pool equation in terms of an annual map rather than a daily map, using a mean annual GPP. We tried out different ways of finding an expression for the annual mean GPP and found that the easiest way was to express the coefficients of C_f^n in the numerator and denominator of the simplified GPP in terms of their means, averaging all the time dependent sub-functions and drivers of the GPP. Expressing the foliar pool as an annual

pool with a mean annual GPP function in C_f allowed us to find a tipping point and consequently a line of tipping points. The same qualitative behaviour was exhibited as when we used the daily maps and the continuation software.

Now that we had a very simple mechanism in place, we could show in Chapter 6 that by varying certain constants in the model, such as the soil moisture constant, ψ , we could show how a shock, such as drought, would influence the position of the tipping point of a forest. DALEC EV and DALEC DE do not have a hydrological model and assume that forests are in a constant moisture state. In DALEC Water [46] there is a hydrological model attached, which feeds the soil moisture component into the forward model as a daily driver. Therefore we decided to experiment with the soil moisture constant in the DALEC EV model. Results showed that the tipping point changes drastically when the moisture constant varies and that, in the case of drought, if a forest does not recover in time, it would die, as it would not be able to adjust itself to a new sustainable fixed point, although this could take several years. This part of our research was meant to be an illustration of what can be done with the model and how shocks affect the tipping point. Using the simplifications of equations for pools and GPP created in earlier chapters made this an easy task, but we acknowledge that in reality it may not be possible to use smoothed climate drivers or annual means. However, in this case continuation software would be most suitable. Although it requires more complexity and takes more time, it does not require drivers to be smooth or functions to be averaged.

All our previous research culminated in Chapter 7 in which we considered a numerical and analytical sensitivity analysis of the NEE. The motivation for this was that a data assimilation project, REFLEX [13], was in part set up to attempt to constrain parameters. Apart from climate data the only observational data for this project was data for leaf area index, LAI, and net ecosystem exchange, NEE. Our aim was to find out from the model equations, which parameters could be constrained from this observational data alone. It is clear that from LAI alone, which is directly proportional to C_f , only the parameters

that play a role in the C_f equation can be constrained. However, the NEE equation also only directly includes a small subset of all parameters. By expanding the NEE using the daily maps of the various pools involved, it is possible to lay bare the role of any parameter involved.

Research by Hill *et al.* [17] shows that it may be possible to constrain more parameters by using longer time series of CO₂ flux. This inspired us to consider the NEE at different periods in time. We already knew that the foliar, roots and litter pools take a relatively short time to reach their fixed points, so after about 30 years (depending on the initial conditions and the values of the parameters involved), they will not have any influence on NEE anymore. A plot of the NEE and its components showed that for the initial conditions used, although those three pools reach their fixed points relatively soon, the NEE remains negative and almost constant for a very long time, due to the wood and soil and organic matter pools not having reached their fixed points. Interestingly, it is not until around 1,000 years that the NEE starts to increase to reach its fixed point, $NEE = 0$, at around 10,000 years.

Therefore we split the 10,000 year time period into three phases. Phase I starts from chosen initial conditions up until the foliar, roots and litter pools have reached their fixed points, at around 30 years. Phase II covers the period between 30 and 1000 years, with a sub phase, Phase IIa, between 30 and 150 years and Phase III starts from around 1,000 years up until 10,000 years. We found that in Phase I the NEE needs to be expressed in terms of all carbon pools, apart from the soil and organic matter pool (which seems to be almost constant). This is the time where we can compare our results with the REFLEX results, as the participants were only required to apply their chosen data assimilation scheme over the first three years from initial conditions. We found that our numerical and analytical results largely agreed with the REFLEX results. It was not possible to write down a useful analytical expression for the NEE with respect to its sensitivity to all parameters due to recurrence relationships between the parameters and the pools. We

were not sure that the REFLEX results were entirely valid, as we did not know what the starting values for the parameters in the REFLEX project chosen by the various participants were and it would be certainly possible to choose a value that lies close to the true value of a parameter. The result may show that the parameter is constrainable, but this may be an accidental positive result. In our opinion it would have been appropriate to give all participants equal starting values for the parameters, rather than requiring them to choose values from indicated intervals.

In Phase II the NEE largely depends on the wood and soil and organic matter pools, but the wood pool is growing much faster than the soil and organic matter pool. Therefore most of the influence on NEE comes from the wood pool and for a short but significant period of time, in Phase IIa, for a sink forest, such as the one we have been using, the NEE can be expressed in terms of mean annual GPP only, involving only parameters p_2 (fraction of GPP respired), p_3 (fraction of NPP allocated to foliage), p_4 , (fraction of NPP2 allocated to roots) p_5 (daily turnover rate of foliage) and p_{11} (Nitrogen use efficiency parameter). A numerical sensitivity analysis of this equation confirms that the NEE is mainly sensitive to those five parameters during this phase. It is interesting to note that the NEE becomes more sensitive to certain parameters over a longer period of time and less sensitive to others. Sensitivity to parameters p_7 (daily turnover rate of roots) and p_8 (daily mineralisation rate of litter) even drops to zero. This may explain why Hill *et al.* [17] found that it becomes possible to constrain parameters better by using longer time series. The rest of Phase II is still dominated by the wood pool, but less so as plots show that the growth rate of the soil and organic matter pool slowly starts to increase until in Phase III the compound growth causes both pools to grow much faster. We could have divided the period between 30 and 10,000 years in a different way, where Phase IIa would be Phase 2 and the remaining time Phase III, but we felt it important to distinguish between the periods where the growth rates of the pools were distinctly different.

Combining the results from Phase IIa and results in Section 5.4, where we found that we

could simplify the annual map for the foliar carbon even more and express it as a linear map (thereby losing the tipping point structure, due to losing non-linearity) we can now, for a sink forest such as Loobos, formulate an extremely simplified model to determine and predict NEE in Phase IIa, for which only a small number of annual observations of LAI are necessary, depending on the age of the forest. This may be applicable to climate research, where all one is interested in is the NEE predicted over a longer period of time.

It is interesting to note that in Chapter 2 the creators of DOLY [53] and BETHY [23] both concluded that it was important to obtain more information on underground carbon stores. Our analysis of the DALEC model with respect to the NEE equation confirms this, as the NEE equation (7.8) depends mainly on the soil and organic matter pool, once the other pools have reached their fixed points, which for many forests may already be the case at the time when they are studied. It is also important to have more information on the wood pool and the root pool, as the sensitivity of NEE with respect to these pools is low and information on them would allow DA schemes to be able to better constrain the parameters involved, namely p_6 (daily turnover rate of wood) for the wood pool, p_7 (daily turnover rate of roots) for the root pool and p_9 (daily mineralisation rate of SOM/CWD) for the soil and organic pool. Also, perhaps by having more information on the litter pool, parameter p_1 (daily decomposition rate) could be constrained better, as this is also a very small parameter. In summary, to be able to constrain all parameters, observational information on all the pools is necessary.

8.2 Impact of Results and Future Work

The research into the tipping point structure could well be of benefit to users of the DALEC model in deciding on sensible values for some parameters, as it is very clear from the plots which values are and are not suitable in certain circumstances. A short preliminary test using the appropriate forest data would show whether parameter values

are viable or not.

A study of how close forests are to their tipping points may provide insight into how forests survive and adapt. In order to be able to study deciduous forests as well as evergreen forests with respect to their proximity to the fixed point, the research into the DALEC DE model would have to be extended to understand in detail what the roles of the other parameters are in determining the tipping point structure. In DALEC DE two pools decouple from the others as opposed to only one pool in DALEC EV and DALEC DE has more parameters.

Also, although many models are more complex than the DALEC models, one should expect a similar bifurcation structure to occur in these more complex models. Common sense tells us that if, for example, the turnover rate of leaves is too high and does not improve, a tree or even a forest may not be able to recover. The ACM model of photosynthesis, which comprises of the GPP, is based on the SPA model, which uses the (nonlinear) Farquhar model of photosynthesis. Photosynthesis models, such as the Farquhar model, sit at the heart of all process-based models and it is the nonlinearity of the photosynthesis model that underlies the structure that is seen in DALEC EV and DALEC. Further research could give insight into whether this is the case.

Another area of interest is the effect of shocks on the DALEC model. We have only done a simple straightforward study of moisture shocks on the DALEC EV model, varying the soil moisture constant. In order to get a more thorough idea of the effect of moisture shocks it would be important to consider more complicated models, such as the DALEC Water model [46], which creates a moisture driver. Using DALEC Water, effects such as precipitation could be taken into account. Other shocks, involving temperature, irradiance or perhaps disturbances such as fire, could also be studied, taking into account the fact that in case of temperature shocks, in the DALEC EV or DE models photosynthesis takes place whatever the temperature, even during frosty days. This is perhaps something that

could be easily remedied by adjusting the DALEC code slightly.

Finally, our research into NEE has shown why some parameters can be constrained better than others and has perhaps thrown light on research by Hill *et al.* in terms of why longer time series of CO₂ fluxes produce better constrainability of parameters and also perhaps on why one data assimilation scheme works better than the other. A scheme that puts less weight on earlier observations may not be able to constrain parameters as well as a scheme that puts weight on both early and later observations. More research, using various DA schemes would be able to give more insight.

8.3 Conferences, Seminars, Workshops, Publications and Other Activities

8.3.1 Conferences, Seminars and Workshops

- Attendance: NCEO Junior Researchers Conference, 2010
- Attendance: Young Researchers in Maths Conference, 2010
- Poster: University of Surrey PGR Conference, 2010
- Poster: NCEO Annual Conference, 2010
- Presentation: MILES University of Surrey, 2011
- Workshop: Climate Change University of Surrey, 2011
- Presentation: SIAM in Snowbird, Utah, USA, 2011
- Presentation: DA Symposium, Durham, 2011
- Presentation: NCEO Annual Conference 2011
- Presentation: University of Surrey (PGRs), 2012

- Presentation: University of Reading, 2012

8.3.2 Publications

- Paper on the tipping point structure soon to be submitted to Ecological Modelling
- Paper on NEE to be submitted in due course

8.3.3 Other Activities

During my time as PhD student I have spent some time tutoring students of all levels, from primary school to university level. This inspired me to study for the Graduate Teaching Certificate at the University of Surrey, which consisted of four modules in one year. I have passed all modules. I also withdrew from the PhD programme for three months to teach A-level students at Chichester College as an Associate Lecturer.

Bibliography

- [1] M.J. Best *et al.* The joint UK land environment simulator (JULES), model description – Part 1: Energy and water fluxes. *Geoscientific Model Development*, 4:677–699, 2011.
- [2] G.B. Bonan. Physiological derivation of the observed relationship between net primary production and mean annual air temperature. *Tellus*, 45B(5):397–408, 1993.
- [3] P.F. Bördeau. Seasonal variations of the photosynthetic efficiency of evergreen conifers. *Ecology*, 40:63–67, 1959.
- [4] F. Bouttier and P. Courtier. Data assimilation concepts and methods. www.ecmwf.int/newsevents/training/rcourse_notes/DATA_ASSIMILATION/ASSIM_CONCEPTS/Assim_concepts11.html, March 1999.
- [5] J.G. Canadell *et al.* Contributions to accelerating atmospheric CO₂ growth from economic activity, carbon intensity and efficiency of natural sinks. *Proceedings of the National Academy of Sciences*, 104:18866–18870, 2007.
- [6] G.J. Collatz, M. Ribas-Carbo, and J.A. Berry. Coupled photosynthesis-stomatal conductance model for leaves of C4 plants. *Australian Journal of Plant Physiology*, 19:519 – 538, 1992.
- [7] P.M. Cox. Description of the TRIFFID dynamic global vegetation model. Technical Report 24, Hadley Centre, 2001.

- [8] W. Cramer *et al.* Comparing global models of terrestrial net primary productivity (npp): overview and key results. *Global Change Biology*, 5:1–18, 1999.
- [9] A. Dhooge, W. Govaerts, and Yu. A. Kuznetsov. Matcont: A Matlab package for numerical bifurcation analysis of odes. *ACM Transactions on Mathematical Software (TOMS) TOMS Homepage table of contents archive*, 29:141–164, June 2003.
- [10] K.A. Dunne. Global distribution of plant-extractable water capacity of soil. *International Journal of Climatology*, 16:841–859, 1996.
- [11] G.D. Farquhar and S. von Caemmerer. A biochemical model of photosynthetic CO₂ assimilation in leaves of C3 species. *Planta*, 149:78–90, 1980.
- [12] J.A. Foley *et al.* Green surprise? How terrestrial ecosystems could affect Earth’s climate. *Frontiers in Ecology and the Environment*, 1:38–44, 2003.
- [13] A. Fox *et al.* The REFLEX project: Comparing different algorithms and implementations for the inversion of a terrestrial ecosystem model against eddy covariance data. *Agricultural and Forest Meteorology*, 149:1597–1615, 2009.
- [14] A.D. Friend *et al.* A process-based, terrestrial biosphere model of ecosystem dynamics (Hybrid v3.0). *Ecological Modelling*, 95:248–287, 1997.
- [15] R. Giering and T. Kaminsky. Recipes for adjoint code construction. *ACM. Transactions of Mathematical Software*, 4:437–474, 1998.
- [16] C. Greene, B. Monger, and M. Huntley. Geo-engineering: The inescapable truth of getting to 350. *Solutions*, 1:57–66, 2010.
- [17] T. Hill, E. Ryan, and M. Williams. The use of CO₂ flux time series for parameter and carbon stock estimation in carbon cycle research. *Global Change Biology*, 18:179–193, 2012.
- [18] J.T. Houghton *et al.* *IPCC2001. Climate Change 2001: The Scientific Basis*. Cambridge University Press, 2001.

- [19] R.D. Hungerford *et al.* A mountain microclimate simulation model. *Res. Pap. INT-414, Intermt. Res. Sta., For. Serv. U.S. Dep. Agric., Ogden, Utah*, 1989.
- [20] T. Ilett. An analysis of a discrete-time nonlinear dynamical system. Master's thesis, University of York, 2008.
- [21] H.G. Jones. *Plants and microclimate: a quantitative approach to environmental plant physiology*. Cambridge University Press, second edition, 1992.
- [22] W. Knorr. Annual and interannual CO₂ exchanges of the terrestrial biosphere: process-based simulations and uncertainties. *Global Ecology and Biogeography*, 9:225–252, 2000.
- [23] W. Knorr and M. Heimann. Uncertainties in global terrestrial biosphere modeling, Part I: A comprehensive sensitivity analysis with a new photosynthesis and energy balance scheme. *Global Biogeochemical Cycles*, 15:207–225, March 2001.
- [24] E. Koffi *et al.* Sensitivity of climate data assimilation to transport models and observational networks. *Proceedings of 8th Carbon Dioxide Conference at Jena, 2009*.
- [25] R. Leemans and W.P. Cramer. The IIASA database for mean monthly values of temperature, precipitation and cloudiness on a global terrestrial grid. *Int. Inst. Appl. Syst. Anal., Laxenburg, Austria*, 1991.
- [26] T.M. Lenton *et al.* Tipping elements in the earth's climate system. *Proceedings of the National Academy of Sciences of the United States of America*, 105:1786–1793, 2008.
- [27] E.N. Lorenz. Deterministic nonperiodic flow. *Journal of the Atmospheric Sciences*, 20:130–141, 1963.
- [28] D.J.C. MacKay. *Sustainable Energy, without the hot air*. UIT Cambridge Ltd., 2009.

- [29] H. McCarthy *et al.* Temporal dynamics and spatial variability in the enhancement of canopy leaf area under elevated atmospheric CO₂. *Global Change Biology*, 13:2479–2497, 2007.
- [30] S.M. McMahon, G.G. Parker, and D.R. Millera. Evidence for a recent increase in forest growth. *Proceedings of the National Academy of Sciences*, 107:3611–3615, 2010.
- [31] D. Medvigy *et al.* Mechanistic scaling of ecosystem function and dynamics in space and time: Ecosystem Demography model version 2. *Journal of Geophysical Research*, 114:GO1002, doi: 10.1029/2008JG000812, 2009.
- [32] D. Medvigy *et al.* Responses of terrestrial ecosystems and carbon budgets to current and future environmental variability. *Proceedings of the National Academy of Sciences*, 107:8275–8280, 2010.
- [33] M.J. Müller. Selected climatic data for a global set of standard stations for vegetation science. *Dr. Junk, Norwell Mass.*, 1982.
- [34] N.K. Nichols. *Data assimilation aims and basic concepts (from the book “Data assimilation for the earth systems”)*. Kluwer Academic, 2003.
- [35] M. Pidwirny and J. Guldedge. Carbon cycle:
http://www.eoearth.org/article/carbon_cycle. *Encyclopedia of Earth*, May 2010.
- [36] R.T. Pinker and I. Laszlo. Global distribution of photosynthetically active radiation as observed from satellites. *American Meteorological Society*, 5(56-65), 1992.
- [37] T. Quaife *et al.* Assimilating canopy reflectance data into an ecosystem model with an ensemble kalman filter. *Remote Sensing of Environment*, 112:1347–1364, 2008.
- [38] S. Quegan and J. Grace. Estimating biogenic carbon fluxes over the UK. *On behalf of the Centre for Terrestrial Carbon Dynamics*, May 2006.

- [39] D. Reay *et al.* Carbon dioxide: http://www.eoearth.org/article/carbon_dioxide. *Encyclopedia of Earth*, September 2006.
- [40] A.C. Rudd. *The effect of nonlinearity on the variational assimilation of satellite observations using a simple column model*. PhD thesis, 2009.
- [41] H. Sato, A. Itoh, and T. Kohyama. SEIB-DGVM: A new dynamic global vegetation model using a spatially explicit individual-based approach. *Ecological Modelling*, 200:279–307, 2007.
- [42] S. Schaphoff *et al.* Terrestrial biosphere carbon storage under alternative climate projections. *Climatic Change*, 74:97–122, 2006.
- [43] J.R. Shewchuk. An introduction to the conjugate gradient method without the agonizing pain, edition 1, 1994.
- [44] U. Siegenthaler and J.L. Sarmiento. Atmospheric carbon dioxide and the ocean. *Nature*, 365, September 1993.
- [45] S. Sitch *et al.* Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model. *Global Change Biology*, 9:161–185, 2003.
- [46] L. Spadavecchia, M. Williams, and B.E. Law. Uncertainty in predictions of forest carbon dynamics: separating driver error from model error. *Ecological Applications*, 21:1506–1522, 2011.
- [47] L. Spadavecchia *et al.* Uncertainty in prediction of forest carbon dynamics: Separating driver error from model error. *Ecological Applications*, 21(5):1506–1522, 2005.
- [48] M. Williams, B.J. Bond, and M.G. Ryan. Evaluating different soil and plant hydraulic constraints on tree function using a model and sap flow data from ponderosa pine. *Plant, Cell and Environment*, 24:679–690, 2001.

- [49] M. Williams *et al.* Modelling the soil-plant-atmosphere continuum in a *Quercus*-*Acer* stand at Harvest Forest: the regulation of stomatal conductance by light, nitrogen and soil/plant hydraulic properties. *Plant, Cell and Environment*, 19:911–927, 1996.
- [50] M. Williams *et al.* Predicting gross primary productivity in terrestrial ecosystems. *Ecological Applications*, 7(3):882–894, 1997.
- [51] M. Williams *et al.* An improved analysis of forest carbon dynamics using data assimilation. *Global Change Biology*, 11:89–105, 2005.
- [52] M.F. Wilson and A. Henderson-Sellers. A global archive of land cover and soils data for use in general circulation climate models. *Journal of Climatology*, 5:119–143, 1985.
- [53] F.I. Woodward, T.M. Smith, and W.R. Emmanuel. A global land primary productivity and phytogeography model. *Global Biogeochemical Cycles*, 9:471–490, 1995.
- [54] T. Ziehn *et al.* Improving the predictability of global CO₂ assimilation rates under climate change. *Geophysical Research Letters*, 38:L10404 doi:10.1029/2011GL047182, 2011.
- [55] P.J. Zinke *et al.* Worldwide organic soil carbon and nitrogen data. *Rep. ORNL/TM-8857, Oak Ridge Nati. Lab. Tenn.*, 1984.

Appendix A

This appendix contains the functions in the GPP.

Canopy Conductance, $g_c(t)$

$$g_c(t) = \frac{|\psi|^{a_{10}}}{0.5T_{range}(t) + a_6R_{tot}}, \quad (8.1)$$

where:

$$\begin{aligned} |\psi| &= 2 \text{ (constant),} \\ R_{tot} &= 1 \text{ (constant).} \end{aligned}$$

CO₂ Concentration at Site of Carboxylation, C_i(t)

$$C_i(t) = \frac{1}{2} \left[C_a + q - p(t) + \sqrt{(C_a + q - p(t))^2 - 4(C_aq - p(t)a_3)} \right], \quad (8.2)$$

where:

$$\begin{aligned} C_a &= 380 \text{ (constant averaged over three years);} \\ q &= a_3 - a_4 = -204.64527 \end{aligned}$$

Photosynthate, p(t)

$$p(t) = \frac{p_{11}NL(t)}{g_c(t)} \exp(a_8T_{max}(t)), \quad (8.3)$$

where:

$$\begin{aligned} N &= 4 \text{ (EV) and } 2.7 \text{ (DE) (constant),} \\ L(t) &= C_f(t)/lma, \\ lma &= 110 \text{ (EV), } 60 \text{ (DE) (constant).} \end{aligned}$$

Canopy Level Quantum Yield, $E_0(t)$

$$E_0(t) = \frac{a_7 L(t)^2}{L(t)^2 + a_9}, \quad (8.4)$$

where:

$$L(t) = C_f(t)/lma,$$

Solar Declination

$$\delta(t) = -23.4 \cos(2\pi t/365)(\pi/180) = -0.408 \cos(2\pi t/365). \quad (8.5)$$

Day Length

$$s(t) = 24 \cos^{-1}(-\tan(lat) \tan(\delta(t))) / \pi, \quad (8.6)$$

where:

$$lat = 52^\circ \text{ (EV, constant)}, 42.5^\circ \text{ (DE, constant)}$$

Appendix B

This appendix contains a list of symbols.

Symbol Description

ψ	Max soil-leaf water potential difference (MPa)
T_r	Daily temperature range ($^{\circ}\text{C}$)
R_{tot}	Total plant-soil hydraulic resistance (MPa $\text{m}^2\text{s mmol}^{-1}$)
N	Foliar nitrogen (g N m^{-2} leaf area)
lma	Leaf mass per area (g C m^{-2} leaf area)
L	Leaf area index ($\text{m}^2 \text{ m}^{-2}$) (LAI)
T_{max}	Maximum daily temperature ($^{\circ}\text{C}$)
T_m	Mean daily air temperature ($^{\circ}\text{C}$)
C_a	Atmospheric CO ₂ concentration ($\mu\text{mol mol}^{-1}$)
δ	Solar declination (radians)
D	Day of the year
NPP	Net Primary Production (GPP - R_a)
NPP ₂	NPP after allocation to foliage
lat	Site latitude ($^{\circ}$)
I	Irradiance ($\text{MJ}^{-1} \text{ m}^{-2} \text{ day}^{-1}$)
R_a	Autotrophic respiration
R_h	Heterotrophic respiration, $R_{h_1} + R_{h_2}$ (g C m^{-2})
R_{h_1} :	$p_8 C_{lit} T(t)$
R_{h_2} :	$p_9 C_{som} T(t)$

Appendix C

This appendix contains the parameter values used in our research for DALEC DE and DALEC EV.

Parameters to be calibrated:

Only used in DALEC EV

	Description	Value in our analysis
p_1	Daily decomposition rate	0.0000044100
p_2	Fraction of GPP respired	0.52
p_3	Fraction of NPP allocated to foliage	0.29
p_4	Fraction of NPP remainining after allocation to foliage, allocated to roots	0.41
p_5	Daily turnover rate of foliage	0.0028
p_6	Daily turnover rate of wood	0.00000206
p_7	Daily turnover rate of roots	0.003
p_8	Daily mineralisation rate of litter	0.02
p_9	Daily mineralisation rate of soil and organic matter	0.00000265
p_{10}	Parameter in exponential term of temperature dependent parameter	0.0693
p_{11}	Nitrogen use efficiency parameter in ACM	7.4

Only used in DALEC DE

	Description	Value in our analysis
p_1	Daily decomposition rate	0.00001
p_2	Fraction of GPP respired	0.45
p_3	Fraction of NPP allocated to foliage	0.4
p_4	Fraction of NPP remainining after allocation to foliage, allocated to roots	0.4
p_5	Daily turnover rate of foliage	0.06
p_6	Daily turnover rate of wood	0.00007
p_7	Daily turnover rate of roots	0.008
p_8	Daily mineralisation rate of litter	0.03
p_9	Daily mineralisation rate of soil and organic matter	0.00003
p_{10}	Parameter in exponential term of temperature dependent parameter	0.073
p_{11}	Nitrogen use efficiency parameter in ACM	14
p_{12}	GDD value causing leaf out (GDD is the growing degree day factor)	240
p_{13}	Minimum daily temperature causing leaf fall	9
p_{14}	Fraction of leaf loss transferred to litter	0.48
p_{15}	Daily turnover rate of labile carbon	0.09
p_{16}	Fraction of labile transfers respired	0.15
p_{17}	Maximum C_f value (g C m^{-2})	300

Parameters viewed as constants, all in ACM:

Optimised Value

a_2 0.0156

a_3 4.22273

a_4 208.868

a_5 0.0453

a_6 0.3783

a_7 7.1929

a_8 0.0111

a_9 2.1001

a_{10} 0.7897

Appendix D

This appendix contains the parameter and other constant values for the Oregon forest.

Parameters viewed as constants, all in ACM:

	Optimised Value
a_2	0.0142
a_3	0.980
a_4	217.9
a_5	0.155
a_6	2.653
a_7	4.309
a_8	0.060
a_9	1.062
a_{10}	0.0006

Parameters to be calibrated:

	Description	Value in our analysis
p_1	Daily decomposition rate	0.0000044100
p_2	Fraction of GPP respired	0.473267
p_3	Fraction of NPP allocated to foliage	0.314951
p_4	Fraction of NPP remainining after allocation to foliage, allocated to roots	0.434401
p_5	Daily turnover rate of foliage	0.00266518
p_6	Daily turnover rate of wood	0.00000206
p_7	Daily turnover rate of roots	0.00248
p_8	Daily mineralisation rate of litter	0.028
p_9	Daily mineralisation rate of soil and organic matter	0.00000265
p_{10}	Parameter in exponential term of temperature dependent parameter	0.0693
p_{11}	Nitrogen use efficiency parameter in ACM	2.155

Other:

	Value in our analysis
$ \psi $	0.8502
R_{tot}	1
N	2.7
lat	52°
lma	111
C_a	357 (averaged over three years of data)

Declaration of Originality

This thesis and the work to which it refers are the results of my own efforts. Any ideas, data, images or text resulting from the work of others (whether published or unpublished) are fully identified as such within the work and attributed to their originator in the text, bibliography or footnotes. This thesis has not been submitted in whole or in part for any other academic degree or professional publication. I agree that the University has the right to submit my work to the plagiarism detection service TurnitinUK for originality checks. Whether or not drafts have been so-assessed, the University reserves the right to require an electronic version of the final document (as submitted) for assessment as above.

Signed: