THE UNIVERSITY *of* EDINBURGH

A Multi-Scale Whole-Plant Model of Arabidopsis Growth to Flowering

Yin Hoon Chew

Thesis submitted for the degree of

Doctor of Philosophy

Institute of Structural and Molecular Biology

School of Biological Sciences

University of Edinburgh

Scotland

United Kingdom

May 2013

Contents

Abstract

This study employed a systems biology approach to determine how the environment impacts plant growth, combining models in a modular fashion that integrates information from the molecular level up to the ecosystem level. This provides a means to understand the combined effects of environmental signalling on cellular function and metabolic flux, hence on carbon partitioning and ultimately plant form. It also allows us to predict how daily or seasonal changes in the weather, combined for the first time with gene network dynamics, alter plant morphology.

In this study, theoretical and experimental approaches were combined, using Arabidopsis as the studied species. The multi-scale model incorporates the following, existing sub-models: a phenology model that can predict the flowering time of plants grown in the field, a gene circuit of the circadian clock network that regulates flowering through the photoperiod pathway, a process-based model describing carbon assimilation and resource partitioning, and a functional-structural module that determines shoot structure for light interception and root growth. First, the phenology model was examined on its ability to predict the flowering time of field plantings at different sites and seasons in light of the specific meteorological conditions that pertained. This analysis suggested that the synchrony of temperature and light cycles is important in promoting floral initiation. New features were incorporated into the phenology model that improved its predictive accuracy across seasons. Using both lab and field data, this study has revealed an important seasonal effect of night temperatures on flowering time. Further model adjustments to describe phytochrome (*phy*) mutants supported the findings and implicated phyB in the temporal gating of temperature-induced flowering. The improved phenology model was next linked to the clock gene circuit model. Simulation of clock mutants with different free-running periods highlighted the complex mechanism associated with daylength responses for the induction of flowering. Finally, the carbon assimilation and functional-structural growth modules were integrated to form the multi-component, whole-plant model. The integrated model was successfully validated with experimental data from a few genotypes grown in the laboratory.

In conclusion, the model has the ability to predict the flowering time, leaf biomass and ecosystem exchange of plants grown under conditions of varying light intensity, temperature, CO_2 level and photoperiod, though extensions of some model components to incorporate more biological details would be relevant. Nevertheless, this meso-scale model creates obvious application routes from molecular and cellular biology to crop improvement and biosphere management. It could provide a framework for whole-organism modelling to help address global issues such as food security and the energy crisis.

Declarations

This Thesis contains original work based on the following published paper:

- Chew, Y. H., Wilczek, A. M., Williams, M., Welch, S. M., Schmitt, J., & Halliday, K. J. (2012). An augmented Arabidopsis phenology model reveals seasonal temperature control of flowering time. *New Phytologist, 194*(3), 654 - 665.

I hereby declare that this Thesis is my own work except where explicitly stated. No part of this Thesis has been previously submitted for a professional qualification or a degree at the University of Edinburgh, or any other university.

Yin Hoon Chew
Edinburgh, May 2013

Acknowledgements

First, I would like to extend my deepest gratitude to my principal supervisor, Dr. Karen Halliday, and my co-supervisors, Professor Dr. Mathew Williams and Professor Dr. Andrew Millar, for their supervision and useful advice throughout my PhD study. Without their endless support, this would not have been possible.

I would like to thank Dr. Bénédicte Wenden, who had patiently guided me during my first year of study, Dr. Kelly Stewart and Mr. Gavin Steel, for their technical assistance and practical guidance which had prevented me from making mistakes in the lab, and Dr. Eve-Marie Josse, for making me feel at home right from the beginning. To Åke Henrik Johansson and Joseph Rustat Hemsted, both of you had been great companies in the student office. To Canan Külahoglu, your brief presence in the lab will always be missed. Not forgotten are Dr. Julia Foreman, Dr. Tomasz Zielinski, Dr. Jayne Griffiths, Dr. Gabriela Toledo Ortiz, Robert Smith and other past and present members of the Halliday and Millar Labs for all their help in getting me to understand plant biology. Special thanks to Kelly, Gavin and Henrik, who conducted a few gas exchange measurements for me while I was away on a conference.

I am also grateful to Professor Dr. Amity Wilczek, Professor Dr. Stephen Welch and Professor Dr. Johanna Schmidt who had generously shared their data and provided honest comments on my manuscript.

This study was made possible with the scholarship awarded by The Darwin Trust of Edinburgh. It was also supported by the Biotechnology and Biological Sciences Research Council (BBSRC) through the ROBuST grant. Thus, I wish to extend my greatest gratitude to both funding bodies.

Last but not least, I would like to express my deepest appreciation for my parents, family and friends for their encouragement and support.

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Background of study

A systems approach to model biological events is becoming an important tool to understand how plants respond to changes in the environment, which is central to crop improvements and biosphere management. As crop yields are predicted to decline in light of climate change (Estrella, Sparks, & Menzel, 2007; Luterbacher et al., 2007; Tao, Yokozawa, Xu, Hayashi, & Zhang, 2006), renewed efforts would be required that bring together studies addressing the interaction between plants and the environment at all levels of organisation.

On the smaller scale, genetic studies at the molecular level have revealed specific genes associated with the sensing of environmental factors. These 'sensors' activate a suite of downstream pathways to regulate metabolic reactions at the cellular level, which in turn control growth at the tissue and organ level, as well as development at the whole-plant level (Lucas, Laplaze, & Bennett, 2011). All these interactions collectively enable individual plants to survive and cope with extremely variable conditions.

In recent years, our understanding of plant systems at each level has improved significantly with the development of new analytical tools, both in wet and dry laboratories. As the vast amount of experimental data being generated is becoming difficult to handle manually, the emergence of bioinformatics has facilitated the analysis of complicated networks, which helps to identify novel components and biological pathways. Mathematical modelling of these pathways in turn drives new hypotheses and experimental designs (Di Ventura, Lemerle, Michalodimitrakis, & Serrano, 2006).

Plant models have been in existence since the eighteenth century. Early phenology models studied the effects of observable environmental factors such as temperature and daylength on developmental events at the whole-plant level (Robertson, 1968). Since then different types of models have surfaced, ranging from spatial to temporal and simple statistical relationships to mechanistic models (Lucas et al., 2011). The complexity of these models depends on the level of study, desired outcome and the amount of data available. Most of the classical models were limited to mainly one level of organisation with little consideration of the interaction or feedback between levels. Contemporary models now make use of our growing knowledge by incorporating genetic and metabolic information either explicitly or through indirect representation (Salazar et al., 2009; Thomas, 1973; White & Hoogenboom, 1996; Wilczek et al., 2009). The interaction between plant architecture and function has also received attention with the development of functional-structural models (Mathieu, Cournede, Barthelemy, & De Reffye, 2008; Sievanen et al., 2000; Vos et al., 2010). The systems approach has now triggered the emergence of cross-scaled models which could provide improved predictive power while offering better capacity to understand system behaviour (Lucas et al., 2011). With this in mind, the main aim of this study was to develop a multi-scale whole-plant model that integrates information from the molecular level up to the ecosystem level, using *Arabidopsis thaliana* as the studied species.

1.2 Overview of study

Light and temperature are among the primary environmental factors regulating plant growth and development. Therefore, these factors have large impact on crop yield. Studies on the model species *Arabidopsis thaliana* have improved our understanding on the regulatory system in plants in response to these factors.

In Arabidopsis, photoreceptors such as phytochromes and cryptochromes act as sensors to provide light-signalling input for the regulations (Sullivan & Deng, 2003). They interact with the endogenous circadian oscillator (Yanovsky & Kay, 2002), which synchronises plant physiological and metabolic rhythms with the diurnal and seasonal timing (Mas & Yanovsky, 2009). Some photoreceptors have been found to be temperature-sensitive (Blazquez, Ahn, & Weigel, 2003; Halliday, Salter, Thingnaes, & Whitelam, 2003). In the development to flowering, temperature effects can occur in two ways, i.e. warm ambient temperature promotes floral initiation (Balasubramanian, Sureshkumar, Lempe, & Weigel, 2006; Kumar et al., 2012), while long period of cold temperature accelerates flowering in a process known as vernalisation (Amasino, 2010). These plant responses ensure that the transition from vegetative to reproductive growth match seasonal changes, which can increase survival rate. Biomass accumulation can also be affected following this developmental switch due to changes in resource allocation. In addition, temperature influences growth-related processes such as respiration (ap Rees et al., 1988) and photosynthesis (Berry & Bjorkman, 1980), which is dependent on light intensity, CO_2 level and plant architecture.

As there is a large concentrate of data available on Arabidopsis, this model species was therefore selected for the current study, where a multi-scale model was developed to describe Arabidopsis growth and development to flowering. A combination of theoretical and experimental approaches was used and the interaction between external and internal regulations was considered. Various models that describe the plant system at different levels of organisation were combined in a modular fashion and tested in physiological experiments (Fig. 1.1). The integrated

model incorporates four existing sub-models. These include a phenology model that can predict the flowering time of plants grown in the field (Wilczek et al., 2009), a gene circuit of the circadian clock network that regulates flowering through the photoperiod pathway (Salazar et al., 2009), a process-based model describing carbon assimilation and resource partitioning (Rasse & Tocquin, 2006), and a functional-structural module (Christophe et al., 2008) that determines shoot structure for light interception and root growth. This multi-scale model estimates plant yield from germination to flowering given environmental inputs such as daylength, temperature, light intensity and CO_2 level (Fig 1.1). Such a model could provide a framework for whole-organism modelling to help address global issues such as food security and the energy crisis.

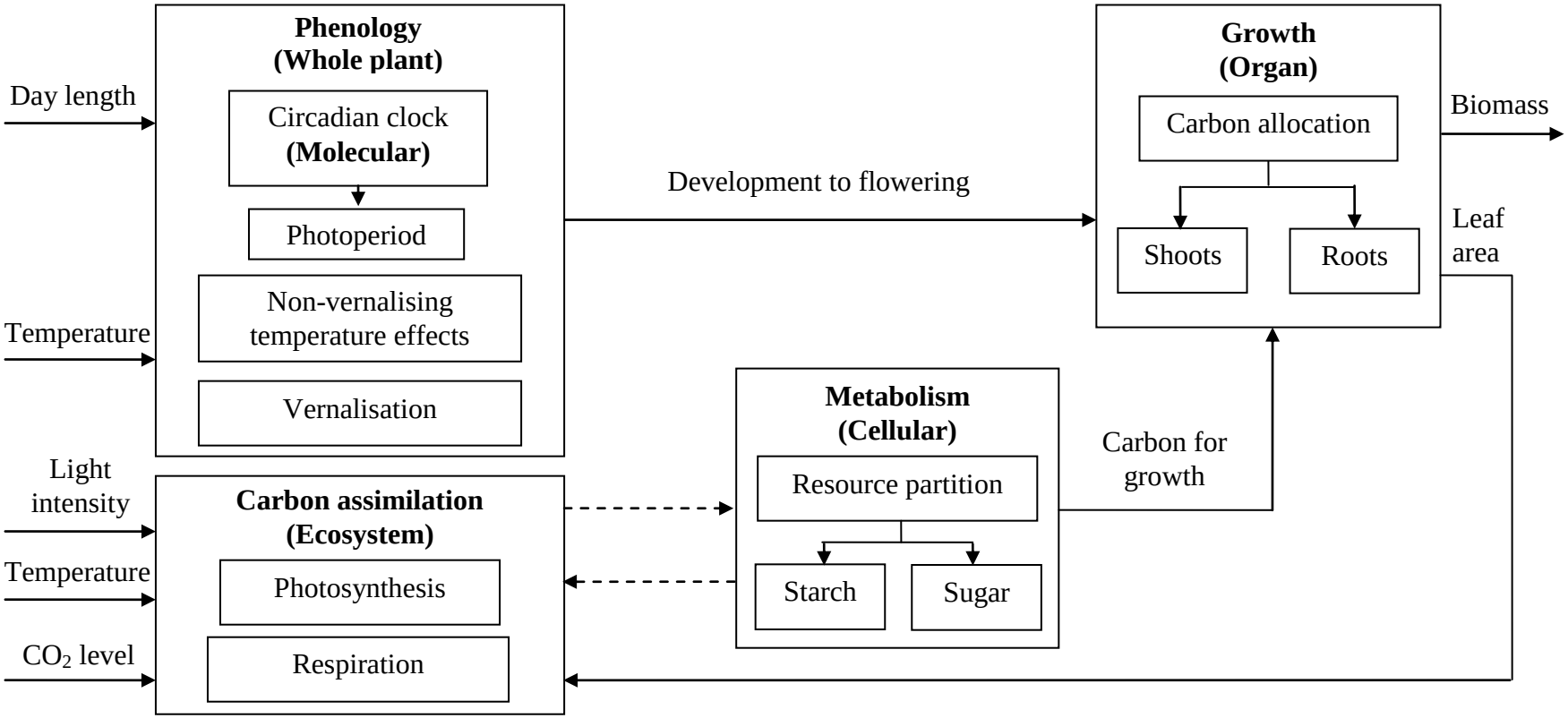


Figure 1.1: A schematic diagram showing the components of the multi-scale integrated model. Texts in brackets indicate the levels of organisation involved in the components. The dashed arrows illustrate the feedforward and feedback of information between the carbon assimilation and metabolism components.

1.2.1 Chapter layout

Chapter 2 presents a literature review of the flowering studies in Arabidopsis and the advancement in multi-scale plant modelling, including its challenges and current limitations. The research methodology adopted in the current study is described in Chapter 3, followed by the results and discussions in Chapter 4, which is divided into three sub-chapters. First, a detailed analysis of the Wilczek et al. phenology model alongside meteorological data and field measurements is presented (Chapter 4: Part I). This analysis was conducted to re-assess the contributions of temperature and light period to promoting floral induction, and the results led to the development of an improved version of the model. Following this, the role of the circadian clock and its importance in plant life cycle were investigated by combining meteorological data with gene network dynamics (Part II). The integrated multi-scale model is presented in Part III, where the model was validated and tested using experimental data of single plants and small populations to explore model potential at a broader scope. Chapter 5 presents a general discussion highlighting model limitations, future extensions and the implications of study. Finally, Chapter 6 concludes the study with a brief summary of the findings.

Chapter 2

Literature Review

2.1 Genetic and environmental control of flowering in Arabidopsis

Arabidopsis thaliana is a rosette plant. During its vegetative phase, rosette leaves are formed at the shoot apical meristem with minimal internode elongation. Upon transition to the reproductive phase, leaf development is suppressed and floral primordia are initiated at the shoot apical meristem (Poethig, 2003). A morphogenetic unit comprising a cauline leaf, an axillary flowering shoot and an internode is produced (Pouteau & Albertini, 2009) to denote the beginning of internode elongation and floral transition. This intermediate region between rosette and true inflorescence is known as the 'cauline leaf zone' and the event is called bolting. The flowering of Arabidopsis has been studied extensively in the last two decades. Two indicators are usually used to define 'flowering time'. The first indicator, also known as flowering time or days to bolting (DTB), measures the time taken from germination to flowering initiation. The second indicator records the number of leaves produced before floral initiation. It has been found that flowering time often correlates to total leaf number at bolting (TLN) (Steynen, Bolokoski, & Schultz, 2001).

A number of major genetic pathways that regulate flowering time in Arabidopsis have been identified (Fig. 2.1). Photoperiod is one of the important inductive signals for flowering in Arabidopsis. As a facultative long-day plant, its flowering is accelerated under long days and the ability to perceive daylength is due to a clock-controlled mechanism through the photoperiod pathway (Imaizumi, Schultz, Harmon, Ho, & Kay, 2005; Salazar et al., 2009; Song, Smith, To, Millar, & Imaizumi, 2012; Suarez-Lopez et al., 2001). The circadian clock component *GIGANTEA* (*GI*) regulates the daily cycle of *CONSTANS* (*CO*) transcription in the

photoperiod pathway, which in turn activates the transcription of floral promoter *FLOWERING LOCUS T* (*FT*) (David, Armbruster, Tama, & Putterill, 2006; Fowler et al., 1999; Imaizumi & Kay, 2006; Park et al., 1999; Tiwari et al., 2010; Yu et al., 2008). In day/night cycles, the timing of the peak *CO* mRNA relative to the day length ensures that CO promotes flowering only in long-day (LD) photoperiods that exceed a critical day length (Fig. 2.2) (Suarez-Lopez et al., 2001; Yanovsky & Kay, 2002). Under LDs *CO* mRNA peaks before the end of the day as a result of both direct interaction with clock-controlled and light-activated FLAVIN-BINDING, KELCH-REPEAT, F-BOX (FKF1) (Song et al., 2012) and the indirect activity of FKF1 with GI, which degrades CYCLING DOF FACTOR 1 (CDF1) and related repressors of *CO* (Fornara et al., 2009; Imaizumi et al., 2005; Yanovsky & Kay, 2002). As this peak occurs during daylight, CO protein is stabilised by the action of phytochrome A (phyA) and cryptochrome 2 (cry2) and flowering is thus promoted through CO regulation of *FT* (Valverde et al., 2004). In short days (SD), as *CO* mRNA peaks during the night, *CO* transcript levels are relatively low during the day. Under these conditions, the CO protein does not accumulate and *FT* mRNA levels remain low (Corbesier et al., 2007b; Valverde et al., 2004; Yanovsky & Kay, 2002). Arabidopsis does eventually flower in SDs, as the gibberellin (GA) pathway takes over to promote floral induction (Dorca-Fornell et al., 2011; Mai, Wang, & Yang, 2011; Moon et al., 2003; Mutasa-Gottgens & Hedden, 2009; Wilson, Heckman, & Somerville, 1992).

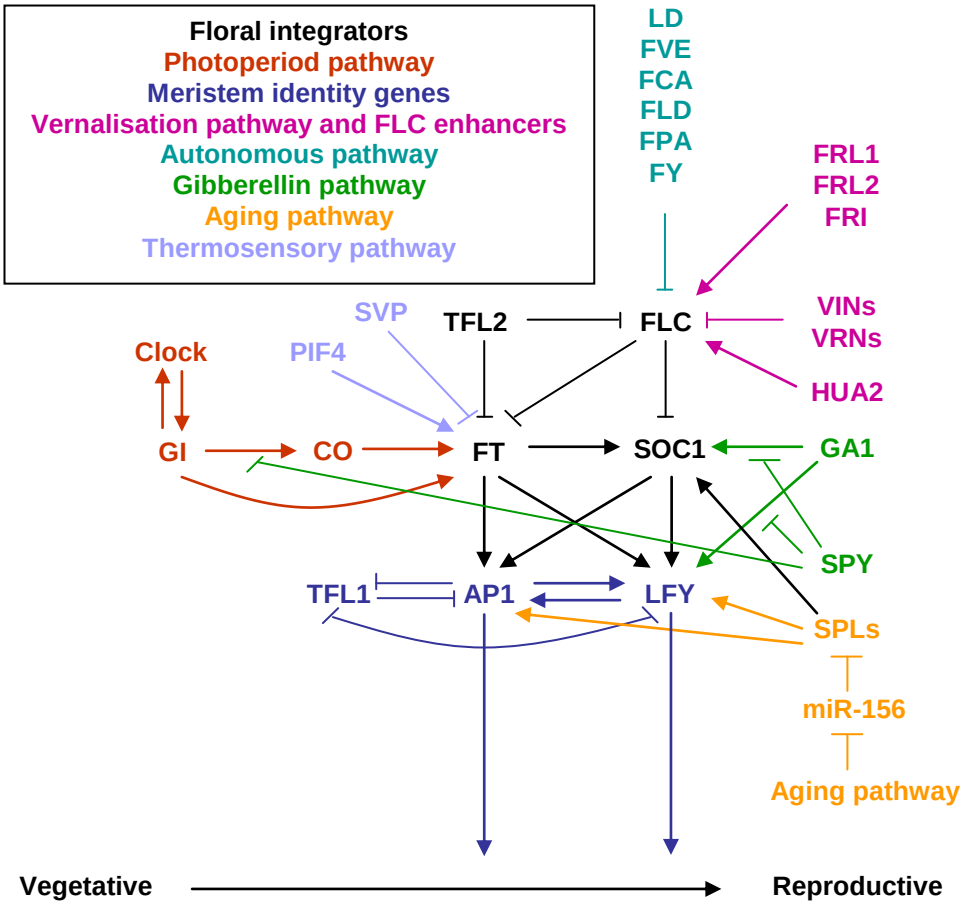


Figure 2.1: The major genetic pathways that regulate flowering time in Arabidopsis. This diagram was reproduced from Wilczek et al. (2009), with the addition of the aging and thermosensory pathways.

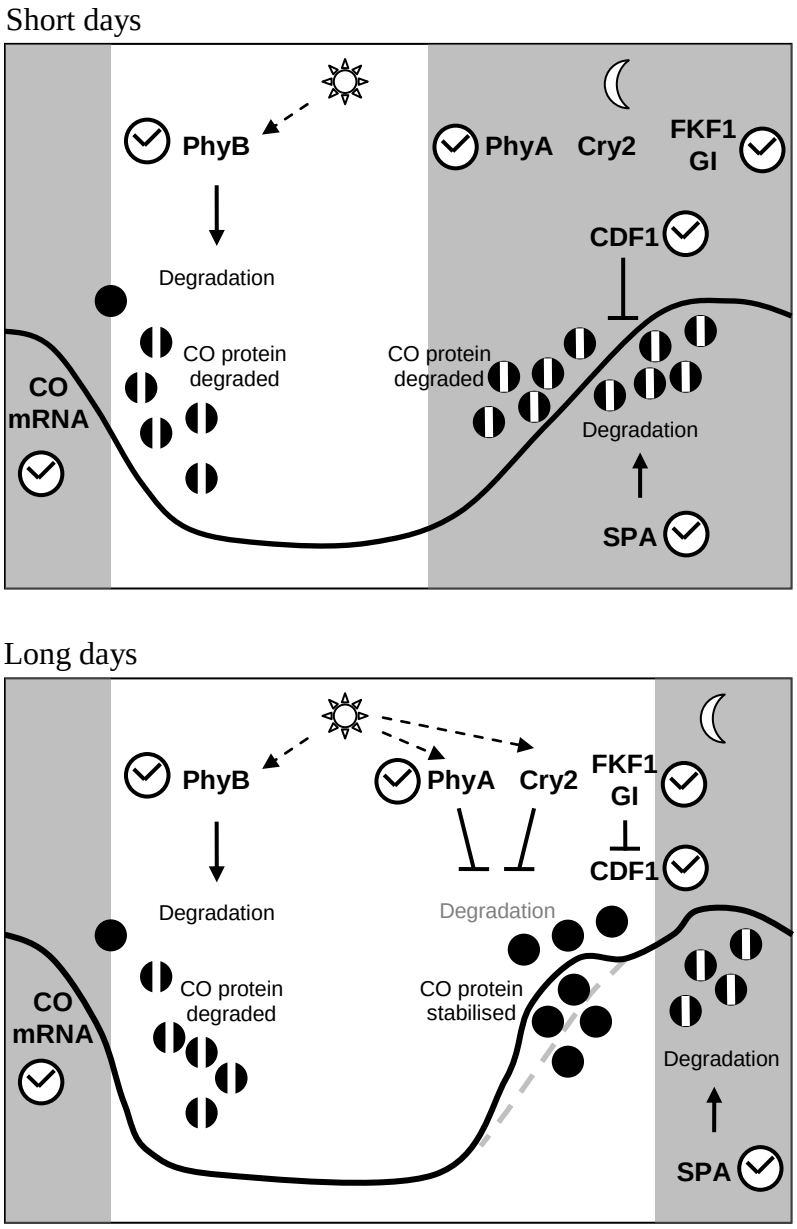


Figure 2.2: Regulation of *CO* gene at both the mRNA (curves) and protein (spheres) levels. In short days, CO protein level remains low at day time due to the degradation (as indicated by the split spheres) promoted by activated PhyB. In long days, CO protein is stabilised near the end of the day as a result of the combined action of light-activated PhyA and Cry2, and the antagonistic action of clock-controlled FKF1 and GI on repressor CDF1. This diagram was reproduced from Lagercrantz (2009). Keys: Clock sign – component(s) regulated by the clock; Full arrow – positive activation; Dashed arrow – light activation; Bar-headed line – inhibition. **Note:** Not all known components and/or activities in the regulation of *CO* are included in the diagram.

Flowering can also be accelerated by exposure to periods of prolonged cold in a process known as vernalisation. This is a strategy adopted by many species that overwinter to ensure that flowering occurs in more favourable spring conditions (Amasino, 2010). A central regulator of the vernalisation pathway is *FLOWERING LOCUS C* (*FLC*) that represses flowering (Michaels & Amasino, 1999) through *FT* and *SUPPRESSOR OF OVEREXPRESSION OF CO 1* (*SOC1*) inhibition (Hepworth, Valverde, Ravenscroft, Mouradov, & Coupland, 2002). *FLC* is gradually inactivated through epigenetic silencing when plants are exposed to a sustained period of cold. This time-dependent process relieves *FLC* suppression of flowering (Bastow et al., 2004; Gendall, Levy, Wilson, & Dean, 2001; Greb et al., 2007; Heo & Sung, 2011; Sung & Amasino, 2004; Swiezewski, Liu, Magusin, & Dean, 2009). In contrast to vernalisation however, short periods of cold can delay flowering (Kim et al., 2004) while warm temperature accelerates flowering (Balasubramanian et al., 2006; Blazquez et al., 2003). A number of genes such as *TERMINAL FLOWER 1* (*TFL1*), *SHORT VEGETATIVE PHASE* (*SVP*), *FVE* and *FCA* from the classical autonomous pathway, and some micro-RNAs have been reported to mediate these effects of ambient growth temperature on flowering (Blazquez et al., 2003; Hanano & Goto, 2011; Kim et al., 2004; Kim, Ahn, Chiou, & Ahn, 2011; Lee et al., 2010; Lee et al., 2007; Strasser, Alvarez, Califano, & Cerdan, 2009). A more recent study suggested a thermosensory pathway mediated through direct activation of *FT* by PHYTOCHROME-INTERACTING FACTOR 4 (PIF4) (Kumar et al., 2012).

Besides photoperiod and temperature, environmental light signals are known to impose strong regulation on flowering time. High vegetative shade or high light intensity can lead to a dramatic acceleration in flowering time (Franklin & Quail, 2010; King, Hisamatsu, Goldschmidt, & Blundell, 2008). These effects are largely mediated through the phytochrome photoreceptors that are accurate sensors of light quality and fluence rates. Two members of this photoreceptor family, phyA and phyB have been shown to operate, at least in part, by regulating the stability of CO protein (Valverde et al., 2004). PhyA promotes flowering by boosting CO accumulation towards the end of a long day, while phyB suppresses flowering by preventing CO accumulation in the morning (Fig. 2.2). *phyA* and *phyB* mutants have been shown to have altered flowering responses to different day lengths compared to wild-type (Giakountis et al., 2010). It has also been shown that the effects of photoreceptors on flowering are temperature sensitive (Blazquez et al., 2003; Halliday et al., 2003), illustrating that the photoreceptor pathways may provide a link between the temperature and photoperiod pathways in the regulation of flowering. This link was further examined in the current study using a modelling approach (Section 4.1).

2.2 Experimental systems to study flowering

Most of the genetic pathways outlined in the previous section were identified in the laboratory where specific environmental factors can be isolated and tightly controlled. However, to understand the genetics of flowering variation, natural conditions are more desirable. Weinig et al. (2002) initiated the first Arabidopsis field study of Quantitative Trait Loci (QTL) for bolting time, and they identified a large number of QTLs with major effects on bolting time in the field environments that were not detectable in controlled conditions. A number of studies have attempted to simulate natural seasonal variation in the lab by programming growth chambers based on the day length, temperature, light quality and intensity, as well as relative humidity from native geographic locations (Li, Huang, Bergelson, Nordborg, & Borevitz, 2010; Li, Roycewicz, Smith, & Borevitz, 2006). By using such artificial natural conditions, these studies have managed to isolate the standard local conditions to understand local adaptation while eliminating any effects of random factors that only corresponded to particular years. A more common alternative in studies opting for natural conditions is the greenhouse, but a comparison study involving 20 000 plants of 184 worldwide natural accessions found a poor correlation in flowering time variation between data from field experiment and those previously obtained under greenhouse conditions, possibly due to the effects of greenhouse-specific factors (Brachi et al., 2010). This highlighted the need to study adaptive variation in flowering time using real field conditions.

Recently, a large-scale co-ordinated field study was conducted at various sites and seasons, covering a range of genotypes that included mutants impaired in specific response to the environment (Wilczek et al., 2009). The data collected from this field study was used to parameterise the Wilczek et al. phenology model. Following parameterisation with field data, the model was tested on independent lab data and managed to predict the flowering time of these validation data. Such combinatory use of lab and field data is therefore useful not only for testing model applicability, but also to validate its ability to capture the effect of an isolated environmental factor, e.g. night temperatures, as shown in the the current study (Section 4.1.4).

2.3 Approaches to modelling plant biology

In the context of systems biology, modelling can be generally divided into two: 1) static graphical representations of the relationships between elements in a network and; 2) spatio-temporal dynamic models. The following sub-sections only covers dynamic modelling, as it is the approach used in the current study.

2.3.1 Types of models

The types of models used in plant modelling are very diverse, ranging from simple statistical relationships to models with comprehensive details. Statistical plant models mainly consider the overall relationship between independent variables (environmental factors such as temperature, humidity, soil moisture and rainfall) and dependent variables (growth, yields and plant distribution) (Prasad et al., 2006; Thuiller, Araújo, & Lavorel, 2003; Yee & Mitchell, 1991). The parameters in these models do not usually represent any biological or physical functions. On the other hand, process-based mechanistic models attempt to describe the relevant physiological processes that channel the environmental effects to the growth outcome. For example, mass balance principles are used in CROPGRO (Hoogenboom et al., 1994) and CERES-Maize (Jones & Kiniry, 1986) to simulate daily crop growth in response to carbon, nitrogen and water availability. These two groups of models are widely used in ecological studies, and a comparison between both model types has shown that they can complement each other to offset uncertainties, but statistical models do not take phenotypic plasticity and local adaptation into account (Morin & Thuiller, 2009). In agronomy, process-based models have gained considerable attention with the advent of modern computers (Vos et al., 2010). However, classical process-based models considered plant canopies as one big leaf that does not distinguish between layers or individual leaves (Allen, Pereira, Raes, & Smith, 1998; Jones & Kiniry, 1986), and this concept clearly could not describe any environmental and genetic control of morphogenesis in plants (Dingkuhn, Luquet, Quilot, & de Reffye, 2005). This called for models that consider

plant architecture, which were pioneered by Honda (1971) and Lindenmayer (1968a, 1968b). Early 3D plant models were mainly descriptive and static, and they were built to determine light interception or for landscaping purposes (Chelle & Andrieu, 1998; Dauzat, 1994; Dauzat & Eroy, 1997; Jaeger & De Reffye, 1992). Functional-structural plant modelling (Godin & Sinoquet, 2005; Vos, Marcelis, & Evers, 2007; Yan, Kang, de Reffye, & Dingkuhn, 2004), which combined process-based and architectural models, later emerged to capture the dynamic feedbacks between plant structure (shaped by individual organs) and plant function as a whole (Vos et al., 2010).

The types of models presented so far mainly describe biological events at the ecosystem, population, whole-organism and organ levels. Recent advances in molecular biology and systems biology have facilitated the modelling of events at scales as small as the genetic level. Gene regulatory network (GRN) models are used in representing the dynamics of genes, mRNA or proteins that are often entangled in complex interactions. Early GRN models of plant development described the dynamics of floral organ specification and root hair development (Mendoza & Alvarez-Buylla, 1998; Mendoza & Alvarez-Buylla, 2000). Others include gene network model of plant development to flowering (Welch et al., 2005) and the circadian clock network (Locke et al., 2005; Locke et al., 2006).

2.3.2 Model formalism

Models can also be classified based on the formalism used. For example, architectural models adopting the Lindenmayer-Systems use a formal language comprising a set of rules for rewriting a string of symbols (Prusinkiewicz & Hanan, 1990; Prusinkiewicz, Karwowski, & Lane, 2007). Each symbol represents a plant structure such as an internode or a leaf, which together forms a repetitive growth unit known as the phytomer (Vos et al., 2010). GreenLab is another type of architectural models and it also adopts the concept of phytomer. In GreenLab, new phytomers appear at the beginning of each growth cycle to form new branches, therefore each

phytomer can be described by its order relative to the main axis (physiological age) and the time elapsed since its emergence (chronological age) (Cournede et al., 2006; Yan et al., 2004). Both the Lindenmayer-Systems and GreenLab formalisms have been very useful in simulating the topological structures in plants. Process-based models, on the other hand, do not have any fixed form. The mathematical representations used can range from a simple proportionality to mass-action or more complex kinetic models, depending on the process to be described and how much mechanistic details are known (Allen et al., 1998; Farquhar, Caemmerer, & Berry, 1980; Rasse & Tocquin, 2006). However, environment x genotype interactions are actually governed by a complex series of molecular reactions and feedback loops, and most crop models employ simple empirical relationships to represent the interactions. Yin and Struik (2010) highlighted the need to reduce the empiricism of these models by either adding mechanistic details or introducing alternative mathematical functions to describe the biological processes. It was suggested that crop models should be revised to incorporate the vast information from functional genomics and biochemistry, which have seen rapid progress with the emergence of systems biology. Approaches for modelling biological events at the cellular and molecular levels include ordinary differential equations (ODE) (Jonsson, Heisler, Shapiro, Meyerowitz, & Mjolsness, 2006; Locke et al., 2005), Boolean network modelling (Espinosa-soto, Padilla-Longoria, & Alvarez-Buylla, 2004; Schlatter et al., 2009) and flux balance analysis (Grafahrend-Belau, Schreiber, Koschutzki, & Junker, 2009; Shastri & Morgan, 2005). On the other hand, systems biology should also draw on the existing development in crop modelling (Yin and Struik, 2010).

In general, plant models can be deterministic or stochastic. Deterministic models compute the same output given the same input variables, while stochastic models contain some probabilistic functions causing each simulation output to be different. Stochasticity has been applied to different types of models. For instance, the GreenLab functional-structural model comes in both deterministic and stochastic versions (Cournede et al., 2006). In the latter, stochasticity is introduced in the branching either by adding a simple randomising function in phytomer appearance and bud death, or through feedback from biomass acquisition (Cournede et al., 2006;

Kang, Cournede, de Reffye, Auclair, & Hu, 2008). At the cellular and molecular levels, stochastic ODE and Boolean models are used to account for any noise due to small numbers of molecules or other sources of uncertainties (Shmulevich, Dougherty, Kim, & Zhang, 2002; Shmulevich, Dougherty, & Mang, 2002; Tian & Burrage, 2006). Recent studies on fluctuations in biochemical activities have garnered new perspectives on noise-induced phenomena, suggesting that they are not random behaviours but could enhance sensitivity and perform certain biological functions (Qian, 2012). Stochastic simulations of these events could reveal such non-trivial behaviours and therefore provide additional information on biological systems (Twycross, Band, Bennett, King, & Krasnogor, 2010).

2.4 Multi-scale plant modelling

Multi-scale modelling is not a distinct type of approach; it only indicates that multiple spatial and/or temporal scales are considered, and it can take any types and forms outlined in Section 2.3. As the current study attempted to link independent models of different scales, a review of the current progress in multi-scale modelling and problems asscociated with it is therefore presented as a separate section here.

Multi-scale modelling is not something new; it has been widely used in physics, chemistry and engineering (Flemisch et al., 2011; Gates, Odegard, Frankland, & Clancy, 2005; Raghavan, Bai, & Ghosh, 2004). In biology, this field has emerged in biomedical sciences such as cardiology (e.g. heart models) and oncology (e.g. cell cycle and tissue dynamics models) with the aims of enhancing systemic understanding and improving therapeutic effects (Qu, Garfinkel, Weiss, & Nivala, 2011; Schnell, Grima, & Maini, 2007; Southern et al., 2008; Stolarska, Kim, & Othmer, 2009). For plant systems, multi-scale models have been applied over a wide range of spatial and temporal scales, though these models mainly cluster around opposite ends of the scales. On one end, ecological models are developed to understand the interaction between hierarchical scales from individual plant species to the whole vegetation (Koniak & Noy-Meir, 2009; van Wijk, 2007), and these models can describe events from one season to a few decades. On the other end, there are multi-scale models describing cellular and tissue dynamics in the fruits, vascular system, shoot apical meristem and roots (Feugier, Mochizuki, & Iwasa, 2005; Ho, Verboven, Verlinden, & Nicolai, 2010; Jonsson et al., 2006; Swarup et al., 2005), which are systems with a relatively shorter time scale. Studies bridging these clusters have surfaced and one notable attempt was a wheat growth-and-development model (Evers et al., 2010), which linked together a 3D architectural model that determines organ-level light interception and photosynthesis to whole-plant assimilate partitioning for a small plot of wheat plants. Attempts have also been made to incorporate genotype-phenotype relationships into crop models using information from QTL analyses (Letort, Mahe, Cournede, De Reffye, & Courtois, 2008; Xu, Henke, Zhu, Kurth, & Buck-Sorlin, 2011; Yin, Chasalow, Dourleijn,

Stam, & Kropff, 2000). Model parameters were linked to specific traits and modified either based on measured values extracted from the literature or through model optimisation. By incorporating such links, the random errors associated with the measured crop model input parameters were partly removed by the QTL analysis statistics, thus improving model prediction of yield differences among different inbred lines (Yin, Kropff, Goudriaan, & Stam, 2000; Yin, Stam, Kropff, & Schapendonk, 2003). These studies therefore demonstrated how physiological modelling and genetic mapping can be combined to resolve environment x genotype interactions to support plant breeding efforts.

2.5 A systems biology approach in the current study

Plants display plasticity by altering development and metabolism, allowing them to adapt to fluctuating environment. To facilitate predictive biology and crop improvement, there is a need to fully understand plant physiology at all levels of organisation, including interlevel interactions. This can be achieved using the systems biology approach, which attempts to understand the behaviour and functional relationships between all the elements in the whole system or organism (Ideker, Galitski & Hood, 2001). To do this, it is necessary to develop whole-organism models comprising sub-models that describe biological events at different levels of organisation. There already exist mathematical models describing various plant processes and Arabidopsis has the concentrate of data sets required to support and test such models. However, these data may be incompatible if there is excessive variation in growth across labs. The current study aimed to explore the feasibility of capitalising on the vast existing plant models by integrating four models from different domains and scales in a modular fashion, and validating the integrated model using independent data.

Each of the four models utilised in this study adopts a different modelling approach. The models include: 1) a GRN model of the circadian clock in the form of ODEs (Salazar et al., 2009); 2) a phenology (crop) model containing empirical functions and scaling factors (Wilczek et al., 2009); 3) a processed-based model of carbon assimilation and resource partitioning (Rasse & Tocquin, 2006); 4) a functional-structural model based on GreenLab (Christophe et al., 2008). As illustrated in the approaches outlined in Section 2.3, different types of models are built for different purposes. It is therefore not desirable to adopt any one type of model to describe a whole organism, particularly for complex multi-scale systems like plants. A recently published bacteria whole-cell model showed a good example of how different types of models could be separately developed, parameterised and then combined using a modular approach (Karr et al., 2012). Thus, the current study adopted a similar approach to develop a multi-scale whole-plant growth model for Arabidopsis.

For model validation, as the integrated model comprises existing models of different types, it is difficult to generalise the type of data that is suitable. Usually, each type of models utilises a specific type of experimental data for both model training and validation. For instance, ecological models are mostly trained and validated using on-site measurement of gas exchange, precipitation, soil composition etc., and/or remote satellite imaging or monitoring system. Functional-structural models, on the other hand, require data of plant architecture such as branching patterns, plant geometry or spatial orientation, and shoot morphology, e.g. the size and shape of leaves. Many 3D digitising devices have been developed for these purposes (Sinoquet & Rivet, 1997; Watanabe et al., 2005). For validation of growth and resource partitioning models, some of the common data used include biomass of whole plants and individual organs, carbon composition and nitrogen content. At the cellular and molecular levels, ODE modelling utilises time-series data of model components, while stochastic models most often require high-resolution data down to the copy number of molecular components in a single cell. These examples are only generalisations; however, a trend can be seen where the type of data correlates with the spatio-temporal scale of the model and also depends on the objective(s) of the model development. For this study, one possibility would be to generate various data to validate all sub-models of different scales. Alternatively, since most model variables are correlated, validation can be conducted on a few selected output variables. The latter was adopted here, where the multi-scale model was validated using biomass data of individual leaves and gas exchange measurement of small plant populations taken at different time points in the vegetative stage.

This study also aimed to bridge between crop modelling and plant systems biology, which not only links models from both domains but also reduces empiricisms in doing so. For example, an empirical function in the phenology (crop) model (Wilczek et al., 2009) was linked to a clock (gene network) model for Arabidopsis (Salazar et al., 2009). The GreenLab growth model (Christophe et al., 2008) that links plant architecture to light interception was also incorporated, where its simple light use efficiency function was replaced by a (biochemical) model describing leaf-

level photosynthesis and resource partitioning (Rasse & Tocquin, 2006). A full description of the multi-scale model is presented in Chapter 4 (Part III).

As mentioned earlier, one of the many advantages of a systems approach is the ability to define an organisation both as a whole and as smaller constituents that make up the whole system. Therefore, in addition to developing a multi-scale plant model, this study also aimed to achieve the following objectives:

1. To enhance our understanding on how external signals, i.e. light and temperature, regulate flowering time;
2. To gain insight into the circadian clock regulation of flowering under natural environments;
3. To test the applicability of the integrated plant model at different scales and levels of organisation.

These objectives build on the current understanding of the genetic pathways controlling flowering time in Arabidopsis (Section 2.1), using the experimental systems described in Section 2.2.

Chapter 3

Methodology

This study adopted a combination of experimental and theoretical approaches. Experimental data from both the field and the laboratory were used in the study. Field measurement and meteorological data were kindly provided by collaborators Professor Dr. Amity Wilczek, Professor Dr. Stephen Welch and Professor Dr. Johanna Schmidt. Data from controlled chamber experiments were either generated as part of this study or were unpublished data of other members in the laboratory. This chapter only described the procedures for experimental work conducted as part of the current study.

3.1 Computational methods and statistical analysis

Modelling work was carried out in MATLAB (Mathworks, Cambridge, UK). Of the four existing sub-models used, only the clock model (Salazar et al., 2009) was readily available in the MATLAB m-file format. Therefore, existing submodels that were not available in Matlab codes had to be re-written in m-files and compared to the original implementation to ensure minimum numerical differences. First, the Wilczek et al. phenology model, which was originally in the Excel Spreadsheet format, was re-coded. The carbon-assimilation-and-metabolic model was originally published in FORTRAN, but was later rebuilt by Simulistics in Simile, which is an object-based modelling platform with a Graphic User Interface. The Simile version of this model has been deposited in the model repository PlaSMo (Plant Systems-biology Modelling). Using the Simile version as a reference, this model was next re-coded. As for the functional-structural model used in this study, it was based on the concept of the GreenLab model (Christophe et al., 2008). Even though the GreenLab model exists in the MATLAB format, it was developed in a structure aimed at model

parameterisation and is not suitable for integration with other models. Therefore, a new functional-structural model was developed in MATLAB to facilitate the integration work in the current study, while re-using the concept and parameter values of the GreenLab model (Christophe et al., 2008).

Model analysis such as parameter estimation and statistical work were also conducted in MATLAB. For model re-parameterisation of the Wilczek et al. phenology model (Chapter 4: Part I), the global optimisation tool (active-set algorithm) was used. Analysis of covariance (ANCOVA) was conducted using the statistics toolbox and interactive tool (Chapter 4: Part II). Ordinary differential equations (ODEs) in the circadian clock model were solved using the MATLAB ode15s variable-step solver.

3.2 Plant materials and growth conditions

Arabidopsis thaliana from the Landsberg *erecta* (Ler) and Feira (Fei) (kindly provided by Dr. Ronan Sulpice from Max Planck Institute for Molecular Plant Physiology, Potsdam-Golm, Germany) accessions were used in this study. All seeds were first sown on half strength Murashige and Skoog (MS) solution and stratified in darkness at 4 ^{0}C for 5 days before being exposed to white light at the desired photoperiod and temperature. Four-day-old seedlings were then transferred to trays of 6 x 4 pots containing Levington seed and modular compost (plus sand). Light was provided by Polylux XL_R 835 (GE Lighting) triphosphor white fluorescent tubes, 130 μmol m^{-2} s^{-1}.

3.3 Gas exchange measurement

An EGM-4 Environmental Gas Monitor for CO2 (PP Systems, US) was used for CO_2 flux measurement. A 50 cm x 35 cm x 20 cm perspex chamber was fabricated and connected to the EGM-4 with two butyl tubes for closed-loop measurement (Fig. 3.1a). An airtight seal was created between shelf top and chamber by lining adjoining surfaces with neoprene and pressing the chamber tightly to the shelf during measurement. Each individual measurement was taken by placing a tray of plants inside the chamber for approximately 60 seconds, during which the EGM-4 recorded CO_2 concentration (μmol mol^{-1}) every 4.6 seconds. The air in the chamber was mixed using two small fans. CO_2 flux was calculated from the slope of CO_2 concentration plotted against time as follows:

$$F_c = \frac{dC}{dt} \times \rho V \,. \tag{3.1}$$

$$\rho = \frac{P}{RT}, \tag{3.2}$$

where

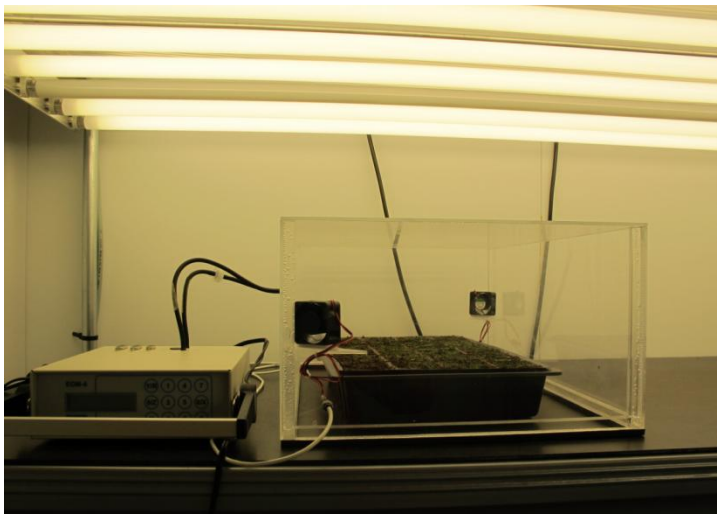
F_c	=	net CO_2 flux (μmol s^{-1})
ρ	=	air density (mol m^{-3})
V	=	air volume in the chamber (m^3)
dC/dt	=	slope of CO_2 concentration against time (umol mol^{-1} s^{-1})
P	=	pressure (Pa)
R	=	ideal gas constant, 8.314 (m^3 Pa mol^{-1} K^{-1})
T	=	temperature (K)

CO_2 flux of 4-week-old plants was measured every few days until flower buds first became visible to the naked eye. As a control, gas exchange from a tray of soil without plants was also measured. Hourly CO_2 concentration at leaf level was also monitored by connecting the EGM-4 to a computer for data logging. The average hourly CO_2 level logged over the weekend, when there was no human interference, was used as input to the model (Fig. 3.1b).

3.4 Leaf number and biomass assay

The number of total leaves (including the cotyledons) was recorded every 3-4 days from germination until the flower buds appeared. Any leaves exceeding 1 mm^2 in size (estimated with the naked eye) were considered in the total leaf count. Biomass assay was also conducted at different time points throughout the vegetative stage. Plants were selected randomly from each accession and rosettes were dissected for the measurement of individual leaf biomass. For dry biomass assay, dissected plants were dried in the oven at 80 ^{0}C for 7 days. For the assay at the final time point, only plants with visible flower buds were selected to ensure they were all at the same developmental stage. Area analysis was conducted using Image J (Schneider, Rasband & Eliceiri, 2012). Each image was first processed with colour thresholding to isolate the green region and convert it into the binary format. The area was then determined using the Analyze Particles tool.

(a)



(b)

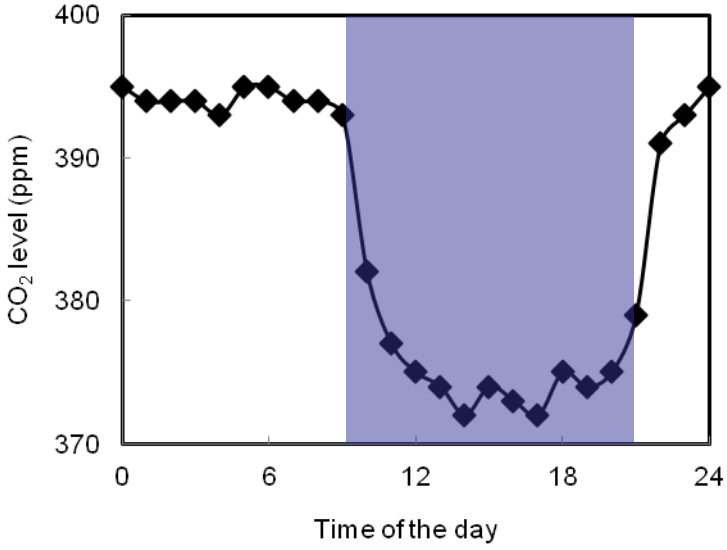


Figure 3.1: a) Experimental setup for gas exchange measurement; b) Hourly CO_2 concentration at leaf level in the growth room.

Chapter 4

Results and Discussions

PART I: Phenology Model

This section is now a published work (Chew et al., 2012). All the analyses conducted in this section were the original work of this thesis. Field data used were either from a published work (Wilczek et al., 2009) or kindly provided by Professor Dr. Amity Wilczek, Professor Dr. Stephen Welch and Professor Dr. Johanna Schmidt. The field experiments were supported by the NSF Frontiers in Integrative Biological Research program grant EF-0425759 and an Alexander von Humboldt Research Award to Professor Schmidt.

Summary: In this section, the behaviour of an earlier published phenology model (Wilczek et al., 2009) was analysed together with meteorological data. Results revealed a seasonal temperature control of flowering time in Arabidopsis. An improved version of the model was also developed and extended to describe photoreceptor mutants, which brought to light a potential mechanism that explains the temperature-dependent role of phyB.

4.1.1 Description of the Arabidopsis phenology model

The basic concept for phenology models was conceived in the eighteenth century by Réaumur, who suggested that transitional events such as flowering occurred when a critical level of accrued daily temperature was exceeded (as reviewed in Robertson (1968)). The notion of accumulated temperature, referred to as ‘degree-days’ or thermal time, suggested that developmental rate was a linear function of temperature

(Monteith, 1981). It was later discovered that the accumulated temperature threshold required to trigger flowering changed with day length. This observation led to the development of a number of improved models that incorporated both temperature and photoperiod data (Robertson, 1968; Weir, Bragg, Porter, & Rayner, 1984; Williams, 1974). Early photothermal models offered good predictive power even though they were based solely on field observations, recorded dates and meteorological data. The early phenology models, however, did not consider genetic variability or the underlying biological pathways. A number of contemporary models exploit our growing knowledge of the molecular pathways that interpret environmental cues. In these models the incorporation of genetic and molecular information improved their predictive power as well as offering improved capability to decipher network behaviour (Messina, Jones, Boote, & Vallejos, 2006; Salazar et al., 2009; Wenden et al., 2009; White & Hoogenboom, 1996). This class of model is more applicable to crop forecasting and molecular-assisted breeding programmes.

The Arabidopsis phenology model (Wilczek et al., 2009) used in this study was developed based on the classical crop modelling concepts, which integrate photoperiod and thermal time. This model also incorporated the impact of vernalisation, a feature that is considered in many phenology models for temperate species (Chuine, 2000; Harrington, Gould, & St Clair, 2010; Wang & Engel, 1998). In addition, this model is contemporary as individual scaling factors were used to link model parameters to the activities of specific genes, particularly in the photoperiod and vernalisation pathways, enabling it to consider genetic variability.

There are three main components in the model: photoperiod, thermal time, and vernalisation (Fig. 4.1.1). It calculates a modified photothermal unit (MPTU), which is a product of the main components, on an hourly basis.

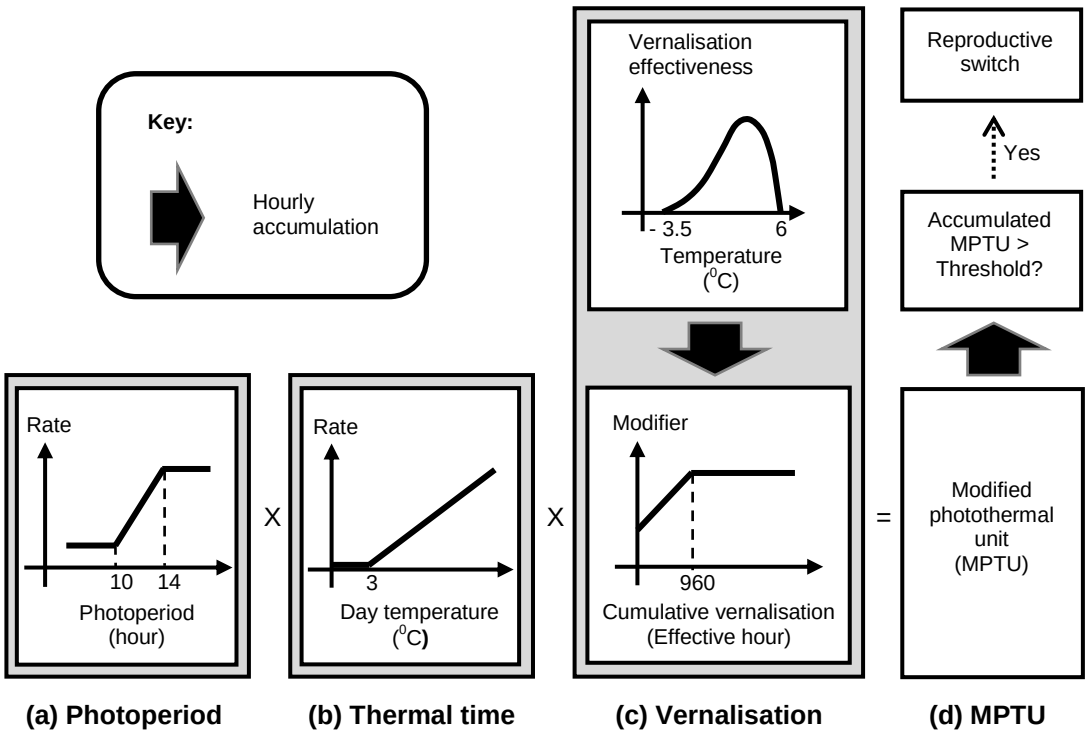


Figure 4.1.1: The three components of the Wilczek et al. photothermal model, with each consisting of a piece-wise linear transition function. (a) Photoperiod component, which is divided into three sections by two critical day lengths; (b) Thermal time component, which considers only daytime hourly temperatures above 3 ^{0}C; (c) Vernalisation component, which has two sub-components. The modifier represents the extent of vernalisation and depends on the accumulated vernalisation effectiveness at different temperatures; (d) The Modified Photothermal Unit (MPTU) is a product of the three components. MPTU is accumulated every hour until a threshold value is reached to indicate the reproductive switch.

The photoperiod component consists of a piece-wise linear transition function, with three sections divided by two critical day lengths. At day length (*dl*) below the critical short day length (CSDL), rate to bolting is at its minimum. As photoperiod lengthens, the rate increases linearly until the critical long day length (CLDL) is reached, where there is no further increase. The effect of photoperiod on rate to bolting at each time-point *t* is therefore:

$$Photoperiod(t) = \begin{cases} D_{SD}, & dl(t) \le CSDL; \\ D_{LD}, & dl(t) \ge CLDL; \\ D_{SD} + \dfrac{(dl(t) - CSDL)(D_{LD} - D_{SD})}{CLDL - CSDL}, & \text{otherwise,} \end{cases} \quad (4.1.1)$$

where D_{SD} and D_{LD} are the minimum and maximum rates respectively. This photoperiodic response has been observed experimentally (Pouteau & Albertini, 2009; Pouteau et al., 2008; Pouteau, Ferret, & Lefebvre, 2006; Wilczek et al., 2009) and could be reproduced theoretically using the clock gene circuit model (Salazar et al., 2009). Genotypes which are insensitive to photoperiod induction maintain a constant minimum rate.

The second component calculates the thermal time in degree-hour, as follows:

$$Thermal(t) = \begin{cases} (T(t) - T_b) \times P(t_{24}), & T(t) \ge T_b; \\ 0, & \text{otherwise.} \end{cases} \quad (4.1.2)$$

$$P(t_{24}) = \begin{cases} 1, & \text{sunrise} \le t_{24} \le \text{sunset;} \\ 0, & \text{otherwise,} \end{cases} \quad (4.1.3)$$

where t_{24} is the time of the day expressed in the 24-hour format. Only day temperatures above a base value, T_b, of 3 ^{0}C (Granier et al., 2002) are considered in the model. The *P* function is included in the equation to filter out night temperatures.

The effect of vernalisation is computed in the model by using first a beta function to represent vernalisation effectiveness (v_e) at different temperatures within a vernalising-temperature range:

$$v_e(t) = \exp(\kappa)(T(t) - T_{V\min})^{\omega}(T_{V\max} - T(t))^{\xi}. \quad (4.1.4)$$

κ, ω and ξ are parameters in the beta function. The minimum ($T_{V\min}$) and maximum ($T_{V\max}$) vernalising-temperatures were fixed at -3.5 and 6 ^{0}C and the three beta function parameters were determined based on statistical analysis and fitting to laboratory and field data (Napp-Zinn, 1957; Wilczek et al., 2009).

Cumulative vernalisation hours (V_h) up to and including the hour that ends at t is computed using a summation function to determine the period of time exposed to the effective vernalising temperatures, as follows:

$$V_h(t) = \sum_{t=0}^{t} (v_e(t) \times \Delta t). \quad (4.1.5)$$

Δt is the time step used in the model, which is set to one hour. The extent of vernalisation, *Vern*(t), is then determined using the following:

$$Vern(t) = \begin{cases} F_b + \dfrac{V_h(t) \times (1 - F_b)}{V_{sat}}, & V_h(t) \le V_{sat}; \\ 1, & \text{otherwise.} \end{cases} \quad (4.1.6)$$

F_b is a parameter representing baseline *FLC* repression. As the period of cold exposure increases, *FLC* becomes more repressed until a point when it is permanently inactivated. This is known as the saturation point, V_{sat}, which has a value of 960 hours or 40 days (Lee & Amasino, 1995; Napp-Zinn, 1957). For vernalisation insensitive genotypes, *Vern*(t) remains at the F_b value regardless of the period of exposure to vernalising temperatures.

Every hour, MPTU is calculated using equation 4.1.7 and accumulates until a threshold T_h is reached:

$$MPTU(t) = Photoperiod(t) \times Thermal(t) \times Vern(t). \quad (4.1.7)$$

$$CumMPTU(t) = \sum_{t=0}^{CumMPTU(t)=T_h} MPTU(t). \quad (4.1.8)$$

The threshold feature in the model is similar to that used in classical crop models. It is also consistent with molecular switch-like traits such as those shown by *FT*, *LEAFY* (*LFY*) and *APETALA1* (*AP1*) (Corbesier et al., 2007a; Jack, 2004; Sablowski, 2007).

The model was parameterised in Wilczek et al. (2009) using field data from seven genotypes: two in the Landsberg *erecta* (Ler) background, i.e. wild type and *co-2*; and five in the Columbia (Col) background, i.e. wild type, *gi-2*, Col-*FRI-Sf2*, *vin3-1* (in Col-*FRI*-Sf2) and *fve-3*. Field experiments were carried out at five sites across Europe, i.e. Valencia, Oulu, Cologne, Halle and Norwich in the spring, summer and/or autumn (Fig. 4.1.2a) (Wilczek et al., 2009). The sites cover a range of latitudes and local climates, such as Mediterranean in Valencia, oceanic in Norwich, and subarctic in Oulu. Cologne and Halle are located at approximately the same latitude as Norwich but experience a continental climate, so cohorts at these sites would experience comparable day lengths but different local weather at similar times of the year. There were a total of nine plantings, including a repeat of Norwich Summer in 2007. Planting was timed to coincide with the observed natural germination flush in the wild population except in Oulu, where natural germination occurs in September.

The optimisation strategy used in this study was similar to that in Wilczek et al. (2009), which was to minimise the coefficient of variation for the set of line x planting MPTU totals. The active-set algorithm in MATLAB was used. Only D_{SD} and F_b were optimised, and constrained to values between 0 and 1 as they were

scaling factors. D_{LD} was therefore set to 1 to represent maximal scaling. *CLDL* and *CSDL* were set based on experimental data of flowering time response to photoperiod (Wilczek et al., 2009). The threshold value T_h would take the mean of planting MPTU totals calculated from the optimised parameter values. This optimisation strategy was adopted as it enabled parameter values to be driven by environmental inputs (Stephen Welch, Personal communication 2011). Parameters for Ler and Col were estimated separately to ensure switch-gene isogenicity.

4.1.2 Analysis of model behaviour

The attributes of the Wilczek et al. model was first investigated by analysing model performance alongside meteorological data. According to Fig. 4.1.2b, the model displayed a good fit for cohorts in the Summer and Spring plantings. For Autumn cohorts, however, there are three distinctive regions of fit: 1) the Valencia Autumn cohort scattered along the diagonal line, indicating a good match between observed and predicted values (e); 2) flowering times of Halle and Cologne Autumn cohorts were mostly overestimated (f and g); 3) flowering times of late flowering genotypes (*gi-2*, Col-*FRI-Sf2*, *vin3-1* and *fve-3*) from Norwich Autumn were considerably underestimated (h). Therefore, the model was further analysed by breaking the data into subsets based on the three groups of goodness of fit (GoF), then matched with the respective meteorological data (Fig. 4.1.2c-h). Comparison of GoF was based on the root mean squared error (RMSE) expressed in days, or RMSE normalised by the mean value of the observations which forms the coefficient of variation, also known as CV(RMSE) expressed as a percentage. Larger values of the RMSE or CV(RMSE) indicate more substantial relative differences between model predictions and field data.

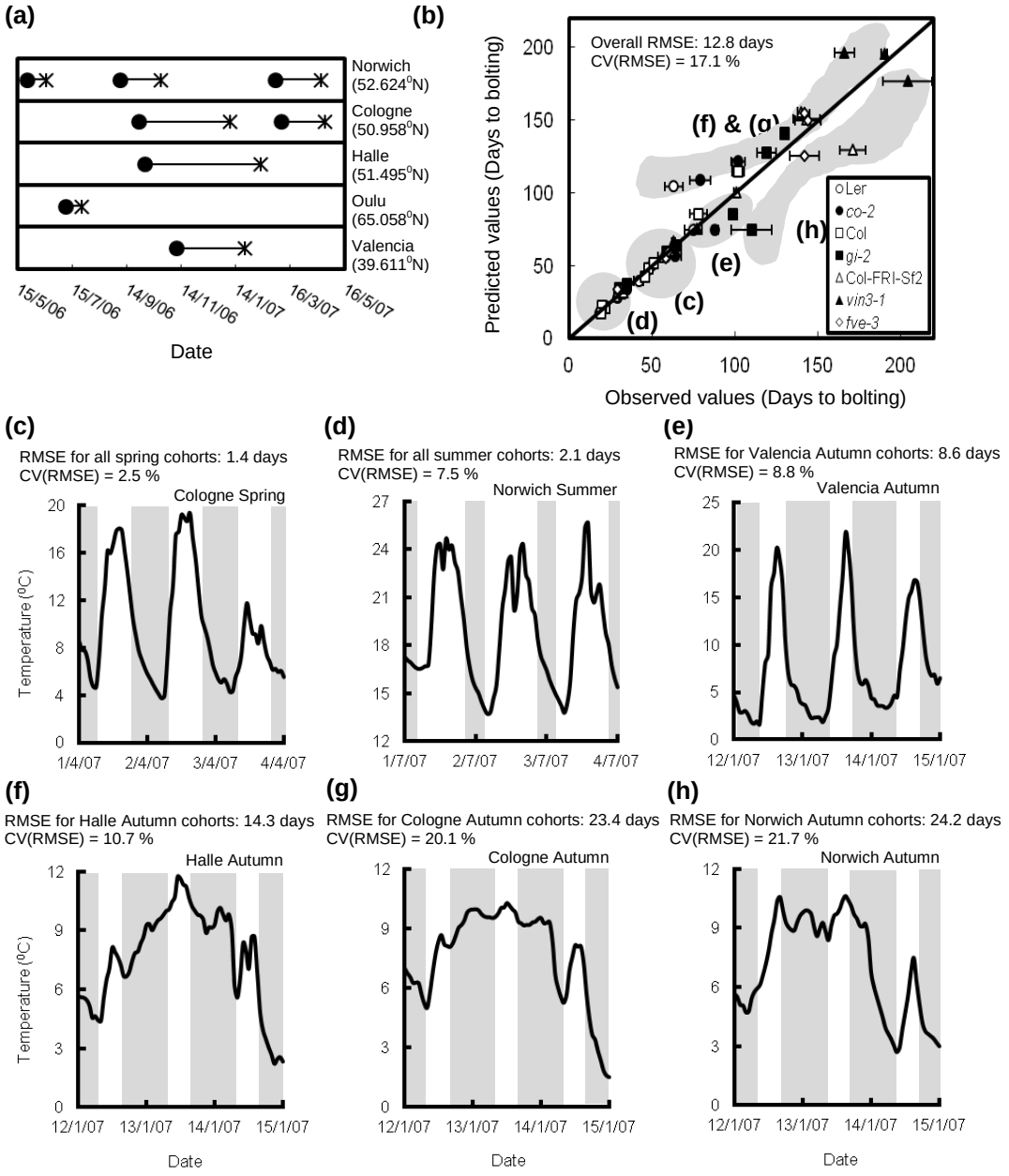


Figure 4.1.2: Analysis of model behaviour and meteorological data. (a) The timing of sowing (●) and Col bolting (*) in eight experimental plantings, which were timed to coincide with the germination of local natural populations (except Oulu, where natural germination occurs in September). A ninth planting, Norwich Summer 2007, took place within one week from the Julian dates of the 2006 Summer cohort. The latitude for each site is shown in brackets. This diagram was reproduced from Wilczek et al. (2009); (b) Predicted versus observed bolting times, with the diagonal line representing perfect fit. Error bars represent one standard error. The goodness of fit (GoF) showed distinctive regions as highlighted. Letters in parentheses link each region to the illustrations of associated meteorological data of a few representative days shown in (c) to (h). Model GoF (root mean squared error or RMSE shown in days and as a percentage) are shown at the top of each meteorological plot.

Fig. 4.1.2c-h shows the seasonal variation in daily temperature time series for a few representative days. Extended through time, this variation leads to distinctly different patterns of day vs. night thermal time accumulation (Fig. 4.1.3). It was found that RMSE was low for cohorts that experienced daily temperature cycles (Fig. 4.1.2c-e). The model could reliably match the flowering time data from Summer, Spring and Valencia Autumn cohorts, where plants typically experienced cooler night and warmer day-time temperatures. However, an incremental rise in RMSE was seen for the Autumn cohorts at Halle, followed by Cologne and Norwich, where the deviations between model predictions and field data were the largest (Fig. 4.1.2f-h). At these sites, the temperature rhythm was less predictable, with occasional peaks at night. These results suggested that the model could match the flowering time data when plants had experienced strong phase synchrony between light and temperature cycles, but was less accurate when this was not the case.

A different behaviour was observed for the Norwich Autumn cohort, where plants fell into two discreet groups, i.e. rapid-cyclers and winter-annuals. As the planting of this cohort began earlier compared to other sites to coincide with its natural germination flush (Fig. 4.1.2a), plants experienced long hours of day light (11 to 13 hours) and relatively warm daytime temperatures during early development. This early induction triggered plants to bolt in late autumn without much vernalisation. Model prediction was good for Ler, *co-2* and Col genotypes from this cohort which displayed the rapid life cycle. However, the model underestimated late flowering genotypes, i.e. *gi-2*, Col-*FRI-Sf2*, *vin3-1* and *fve-3*, from Norwich Autumn. These genotypes did not flower until the following spring, and experienced periods of unsynchronised light-temperature cycles.

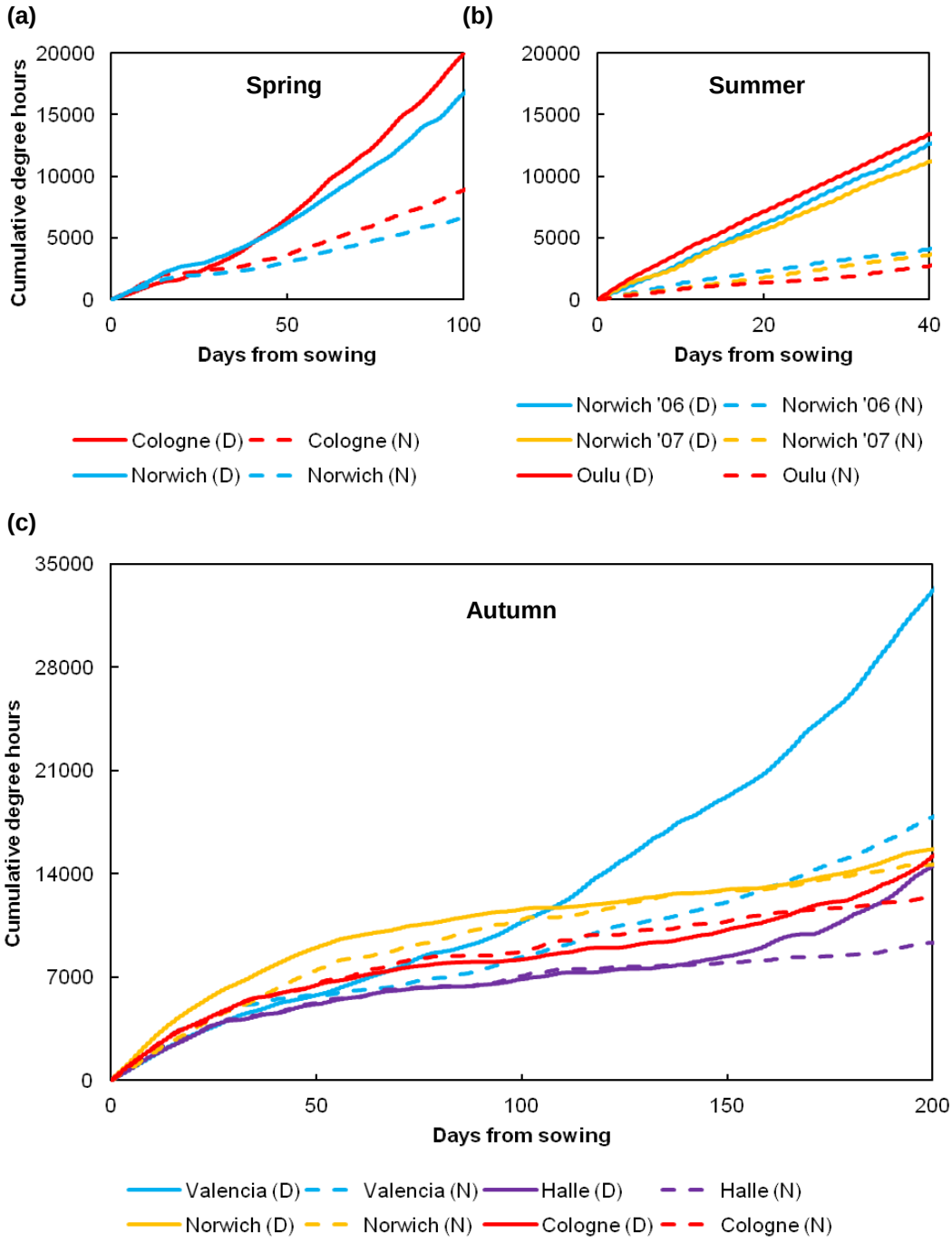


Figure 4.1.3: Day-cumulative (solid lines) and night-cumulative (dashed lines) thermal time (temperature above 3 ^{0}C) at different sites and seasons. In the spring (a) and summer (b), accrued thermal time during the day was always higher than during the night on all sites. In the autumn (c), daytime accumulation rate was also higher in Valencia due to its diurnal temperature pattern (Fig. 4.1.2e). On other sites however, the occurrence of non-vernalising promotory temperatures during the night was comparable to that during the day. Cologne in particular experienced a long period during which thermal time accumulated at a higher rate in the night than in the day. As the season proceeded towards late winter and early spring, the relative contribution of day and night thermal time resumed to the spring/summer profile.

4.1.3 New model variants

The synchrony of thermal and photo-cycles was not explicitly included in the Wilczek et al. photothermal model. However, the lowest cost function was achieved during model optimisation when the "thermal time" component (see Fig. 4.1.1) considered only daytime temperatures (Wilczek et al., 2009). The model was therefore developed to include a filter *P* (equation 4.1.3) that captured the effect of temperature during the day, and night temperatures were ignored in the accumulation of degree-days (Fig. 4.1.4a). This simple function generated higher MPTUs for genotypes exposed to highly synchronised light-temperature cycles as those experienced in the Summer, Spring and Valencia Autumn. On the other hand, meteorological data for Cologne and Halle Autumn illustrates that the temperature profile was more inconsistent with extended periods of instability and occasional rises in night temperature relative to day (Fig. 4.1.2). As night temperatures can affect floral initiation (Moe, 1990; Thingnaes, Torre, Ernstsen, & Moe, 2003; Yin, Kropff, & Goudriaan, 1996), the model may be less accurate at predicting flowering time of autumn/winter cohorts because warm temperatures for the accumulation of degree-hours occurred frequently during the night but were disregarded.

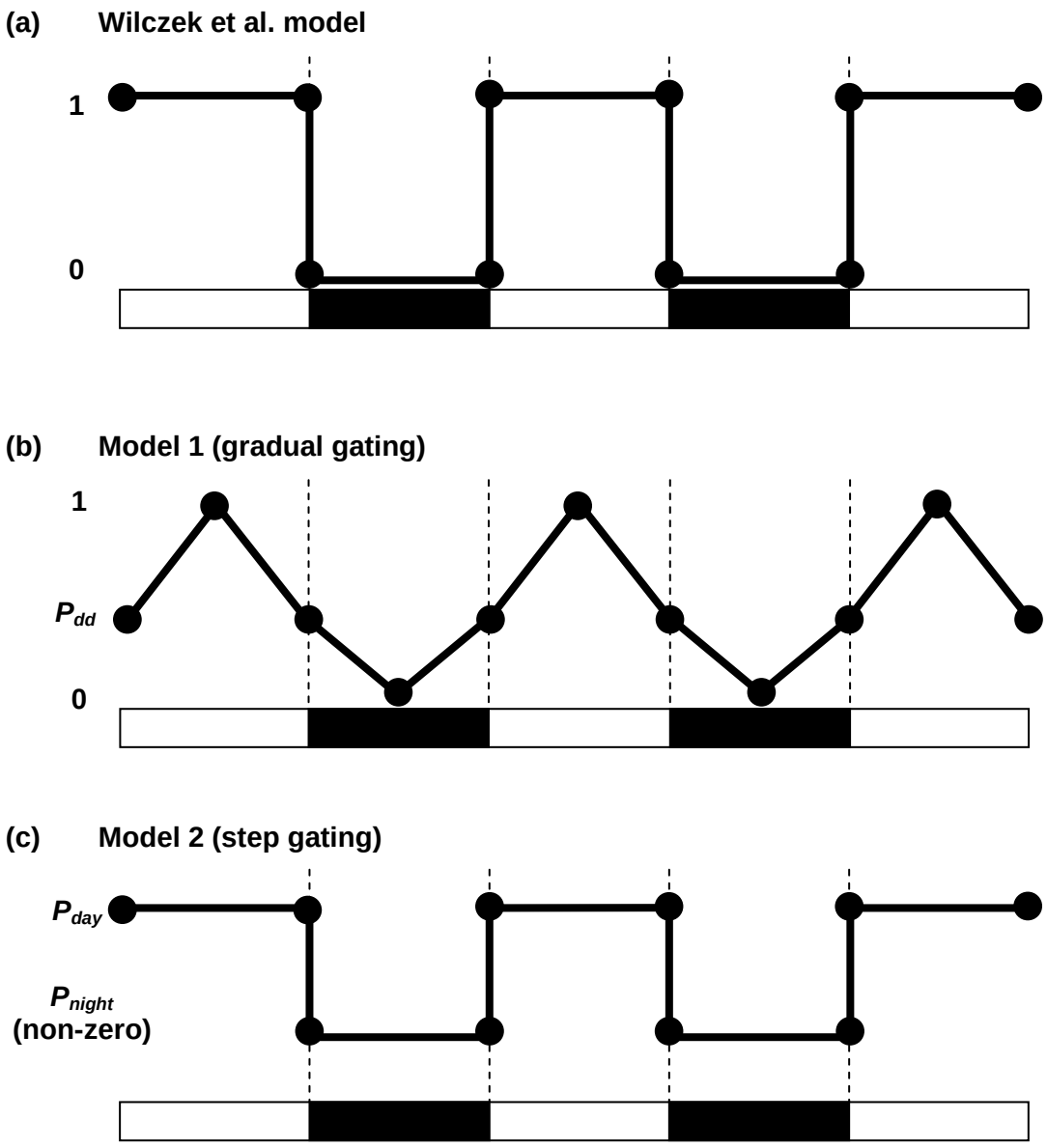


Figure 4.1.4: Filter functions *P* that account for the differential effects of day and night temperatures. The black and white bars represent light-dark cycles. (a) In the Wilczek et al. model, only day temperatures are considered in the thermal time component by multiplying by a factor of 1 at day-time (P_{day}) and 0 at night-time (P_{night}), thus forming a square waveform; (b) In Model 1 (gradual gating), a 'triangle' waveform is used with maximum factor (1) at mid-day and minimum factor (0) at mid-night. A constant factor P_{dd} is locked to dawn and dusk; (c) In Model 2 (step gating), a square waveform with a non-zero night factor (P_{night}) is used.

In order to improve model performance, two new models which incorporated night temperatures in the accrued degree-hours were developed. Variations of the filter function *P* that considered night temperatures were introduced in these models (Fig. 4.1.4). In the Wilczek et al. model, day temperatures were effectively taken into account by multiplying by a factor of 1 (P_{day}), while night temperatures were disregarded using a factor of 0 (P_{night}), thus forming a square waveform (equation 4.1.3 and Fig. 4.1.4a). In the first model variant (Model 1), a 'triangle' waveform function was used instead, where 1 and 0 were fixed to the middle of day and night respectively (Fig. 4.1.4b). Such a function allows both day and night time temperatures to be considered in the accumulation of thermal time units that promote flowering, with the highest effect at mid-day and the lowest effect in the middle of the night. This function could serve as a proxy for circadian gating of the temperature response that is known to occur in Arabidopsis and other species (Espinoza et al., 2010; Fowler, Cook, & Thomashow, 2005; Rikin, Dillwith, & Bergman, 1993). Thus, Model 1 accounts for a gradual gating of temperature effects. A constant factor (P_{dd}) was locked to sunrise and sunset to allow tracking of dawn and dusk (Edwards et al., 2010). In the second model variant (Model 2), a step gating function was introduced by setting (*a priori*) P_{day} as 1 as in the Wilczek et al. model, however a non-zero night temperature factor (P_{night}) applied universally across all plantings was estimated through model-fitting of field data (Fig. 4.1.4c). Both model variants contained each an additional parameter to be optimised, i.e. P_{dd} for Model 1 and P_{night} for Model 2.

The new models achieved comparable GoF of Spring, Summer and Valencia Autumn cohort data when compared to the Wilczek et al. model (with changes in RMSE less than 0.5 day). In addition, both new models improved the fit for Autumn data. For the Cologne Autumn cohort, the RMSE was reduced from 23.4 days (20.1 %) in the Wilczek et al. model to 16.1 days (13.8 %) in Model 1 and 16.4 days (14.0 %) in Model 2 (Fig. 4.1.2 and Fig. 4.1.5). There was also improvement for the Halle Autumn cohort, where RMSE decreased from 14.3 days (10.7 %) in the Wilczek et al. model to 8.8 days (6.6 %) in Model 1 and 10.2 days (7.7 %) in Model 2. According to Fig. 4.1.3, thermal time accumulated at a faster rate during the day time

compared to the night at the Spring, Summer and Valencia Autumn sites, which was in agreement with the observed daily temperature cycles (Fig. 4.1.2). On the contrary, thermal time accumulation during the day and night time was comparable at the Halle and Cologne sites. This seasonal difference in thermal time accumulation rate was due to both reduced day-night amplitudes as well as longer nocturnal durations in the autumn. These collective results indicate that the inclusion of night temperature effects in the thermal time component could be important for determining flowering time of Autumn cohorts.

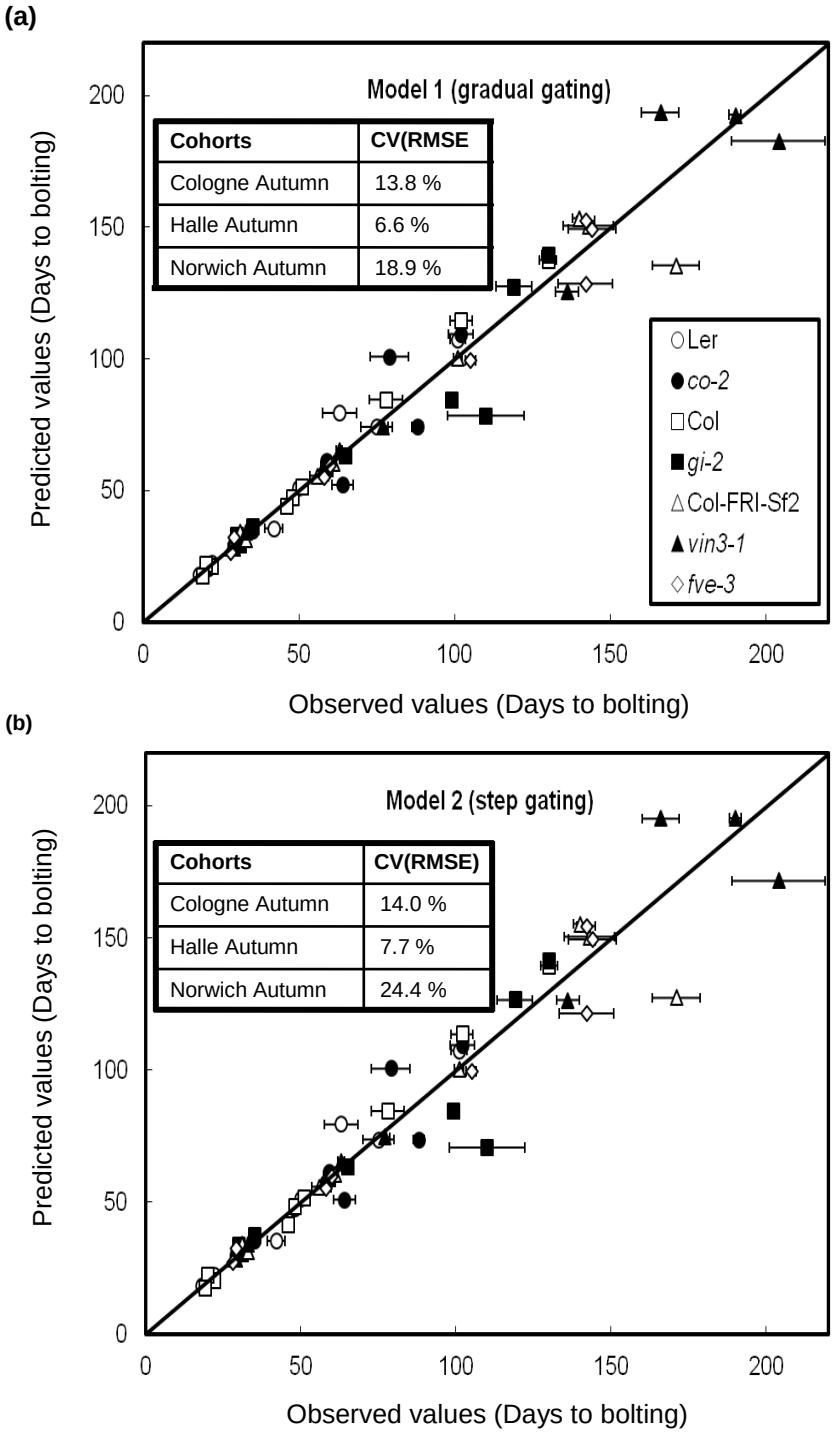


Figure 4.1.5: Bolting times as observed experimentally and predicted using Model 1 (a) and Model 2 (b). The diagonal lines represent perfect fit. Error bars represent one standard error.

Interestingly, the new model variants could still not describe the late flowering *gi-2*, Col-*FRI-Sf2*, *vin3-1* and *fve-3* in Norwich Autumn. This put forward the possibility that the relatively poor performance of the Wilczek et al. model for Autumn cohorts was only a bias as the optimiser tried to fit and divide between divergent groups, i.e. the four 'outlier' genotypes vs. the others. If that were true, removing the outliers should improve the fits of both the Wilczek et al. model as well as the new model variants. Therefore the Wilczek et al. model and both the new models were re-parameterised without the *gi-2*, Col-*FRI-Sf2*, *vin3-1* and *fve-3* data from Norwich Autumn. As can be seen in Table 4.1.1, there was not much improvement in the Wilczek et al. model even without the 'outlier' data, but both the new models which incorporated night temperature improved significantly. These results supported the inclusion of night temperature in the models to accurately describe Autumn cohort data. As the proportion of temperature data considered by the Wilczek et al. model reduced with decreasing day length, model improvements may simply arise from extending the period during which temperature was considered in the autumn and winter. However, this possibility has been explored previously and a degrading fit was found with the incremental inclusion of temperature hours after dusk (Wilczek et al., 2009). Alternatively, the improved fitting in the new model variants may reflect seasonal differences in the effectiveness of day and night time temperatures in controlling flowering time.

Table 4.1.1: Comparison between photothermal model variants

Model	No. of Estimated Parameters	Dataset	RMSE	AICc	
				Col	Ler
Wilczek et al. model	Ler: 2 (D_{SD}, F_b)	+	12.8 days	102	46
	Col: 4 (D_{SD}, F_{bfri}, F_{bFRI}, F_{bfve})	-	9.9 days	76	
Model 1	Ler: 3 (D_{SD}, F_b)	+	10.0 days	100	41
	Col: 5 (D_{SD}, F_{bfri}, F_{bFRI}, F_{bfve}, P_{dd})	-	7.0 days	74	
Model 2	Ler: 3 (D_{SD}, F_b)	+	11.8 days	107	41
	Col: 5 (D_{SD}, F_{bfri}, F_{bFRI}, F_{bfve}, P_{night})	-	7.1 days	75	

Note: + Using all field data
- Excluding *gi-2*, Col-FRI-Sf2, *vin3-1* and *fve-3* from Norwich Autumn

Lower RMSE indicates a better goodness of fit while lower AICc values indicate the more strongly supported models. As the optimisation was conducted separately for Ler and Col with different number of estimated parameters, the AICc value for each model was calculated separately for each line. AICc can only be compared when the number of datasets used are similar, so different comparisons were performed for model with (+) and without (-) the late-flowering genotypes from Norwich Autumn.

Both the new model variants displayed improved fit, which was statistically expected with an increase in the number of parameters but this could cause overfitting. Therefore, the second-order Akaike Information Criterion (AICc), which compares model performance but penalises for increase in the number of parameters or model complexity, was used to assist model selection (Table 4.1.1). Models with lower AICc values indicate that they have stronger statistical support in terms of the balance between model accuracy and model complexity. In general, both the new model variants displayed lower AICc values compared to the original model in all cases except one, indicating that they have strong statistical backing. Model 1 displayed the best improved fit and lowest AICc values compared to Model 2. Nevertheless, due to the additional flexibility offered by the *P* function in Model 2 (see section 4.1.5), this step-gating model variant was selected for subsequent study. The parameter values for Model 2 are listed in Appendices A1 and A2.

4.1.4 Seasonal effects of night temperature

Results from the previous sections have shown that the inclusion of night temperature markedly improved the ability of the models to describe the field data, particularly for the Autumn cohorts (Fig. 4.1.6). This model adjustment was also applicable to published data generated in controlled conditions where plants were grown under laboratory conditions that combined day/night temperatures of 12, 17, 22 and 27 ^{0}C (Thingnaes et al., 2003). The Wilczek et al. model predicted the same flowering time for plants grown under the same day temperature, regardless of night temperatures (Fig. 4.1.7). In contrast, for each set of day temperatures, step-gating Model 2 predicted a decreasing trend in flowering time as night temperature increased. This improved the overall GoF considerably, reducing the RMSE from 12.6 days (30.2%) in the Wilczek et al. model to 7.7 days (18.5%) in the new model. Improvement was especially significant when plants were grown under low day temperature (12 ^{0}C) in combination with various night temperatures, conditions that had naturally occurred in Cologne and Halle during the autumn (Fig. 4.1.2f-g and Fig. 4.1.3c). These results indicate that in comparison to the earlier Wilczek et al. model, the new model can more accurately simulate flowering time of plants subject to cool days vs. warm nights, while retaining ability to describe plants grown under warm days vs. cool nights.

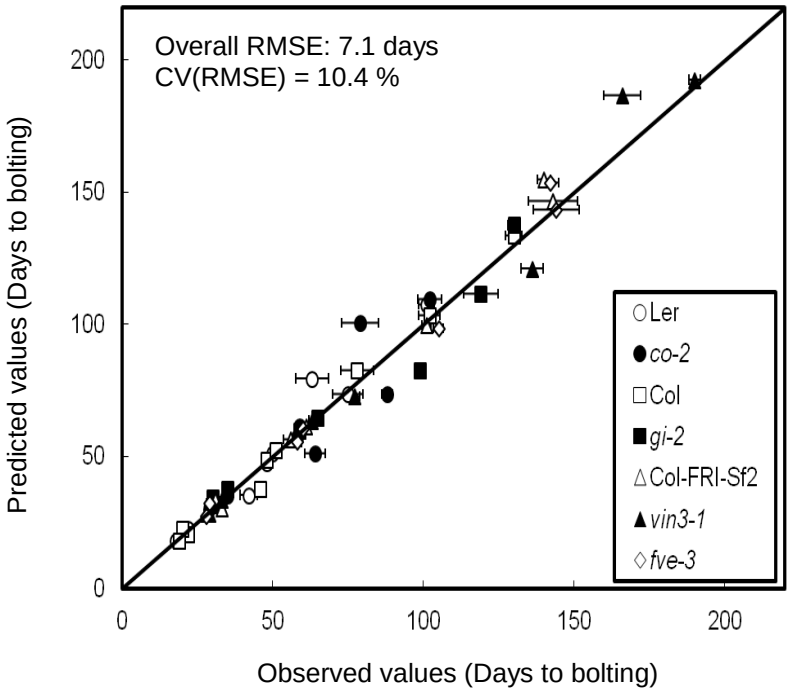


Figure 4.1.6: Bolting times as observed in the field and predicted using step-gating Model 2, with the diagonal line representing perfect fit. Late-flowering genotypes *gi-2*, Col-*FRI-Sf2*, *vin3-1* and *fve-3* from Norwich Autumn were excluded during parameter estimation. Therefore, these data points are not included in the figure.

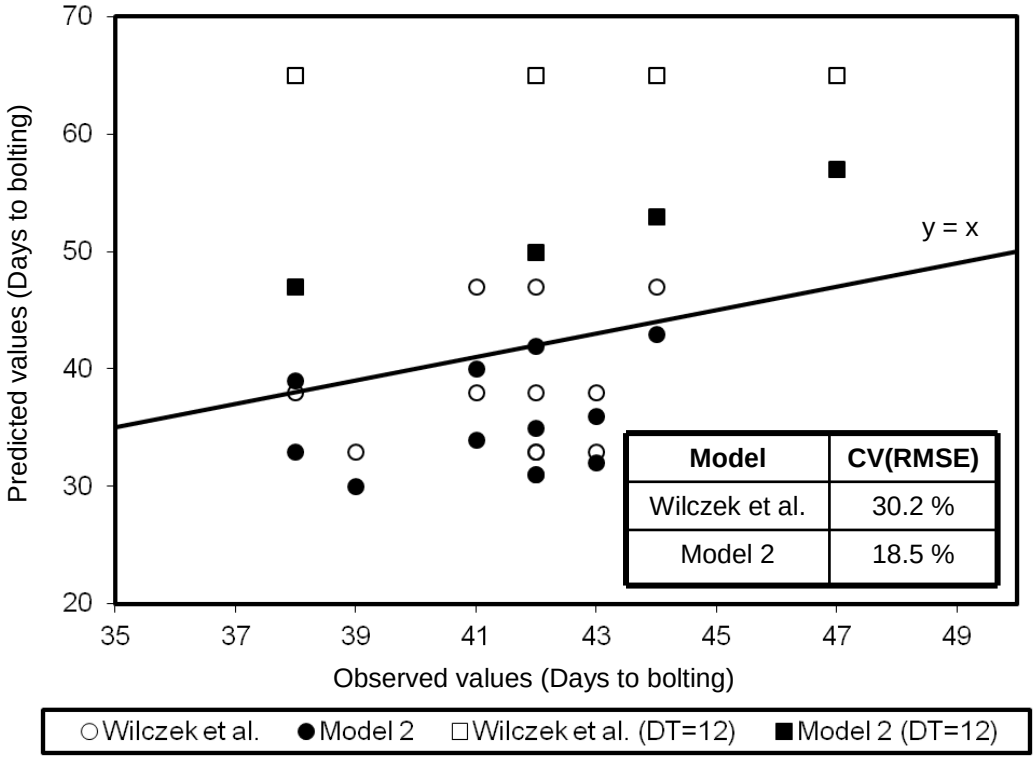


Figure 4.1.7: Validation of Model 2 on its ability to describe the effect of day/night temperatures on flowering time using lab data. The goodness of fit (GoF) of the Wilczek et al. model (open markers) and step-gating Model 2 (closed markers) are indicated by the CV(RMSE). Model predictions were generated without re-optimisation to fit these lab data. The observed experimental data were taken from Thingnaes et al. (2003). In the experiments, plants were grown in controlled chambers under 16 different combinations of day and night temperatures (12, 17, 22, 27 ^{0}C). Square markers indicate the data of plants grown under a day temperature of 12 ^{0}C in combination with various night temperatures.

4.1.5 Model extension to *phyA* and *phyB* mutants

Photoreceptors phyA and phyB play important roles in the control of flowering time (Halliday, Koornneef, & Whitelam, 1994; Halliday & Whitelam, 2003; Valverde et al., 2004). Modelling of genotypes lacking these receptors may therefore increase our understanding on light and temperature regulation of flowering time.

As phyA and phyB are involved in the control of CO stability in the photoperiod pathway (Valverde et al., 2004), the published leaf number data of *phyA-201* and *phyB-1* mutants at different photoperiods (Giakountis et al., 2010) was compared to that of the Ler wild-type (Fig. 4.1.8a). These mutant alleles were also included in the field study (Wilczek et al., 2009). Leaf number data was used for phenotypic comparison as this indicator has been shown to be tightly coupled to bolting time within a wide photoperiod window (Koornneef, Hanhart, & Vanderveen, 1991; Pouteau, et al., 2006). Leaf number data are also more widely available in the literature. Model was extended for *phyA* and *phyB* mutants by modifying the parameters in the photoperiod component (*CSDL*, *CLDL*, D_{SD} and D_{LD} in equation 4.1.1) in Model 2 based on the proportional differences between the mutants and wild type (Fig. 4.1.8a). While a proportional rate informed by leaf number data may not be quantitatively accurate, the data displayed a qualitative photoperiod response that supports the role of these mutants in the photoperiod pathway (see Fig. 4.1.8a and the next paragraph). This qualitative response also concurs with early flowering time phenotype that has been reported for *phyB* mutants in both long and short days (Cerdan & Chory, 2003; Mockler, Guo, Yang, Duong, & Lin, 1999).

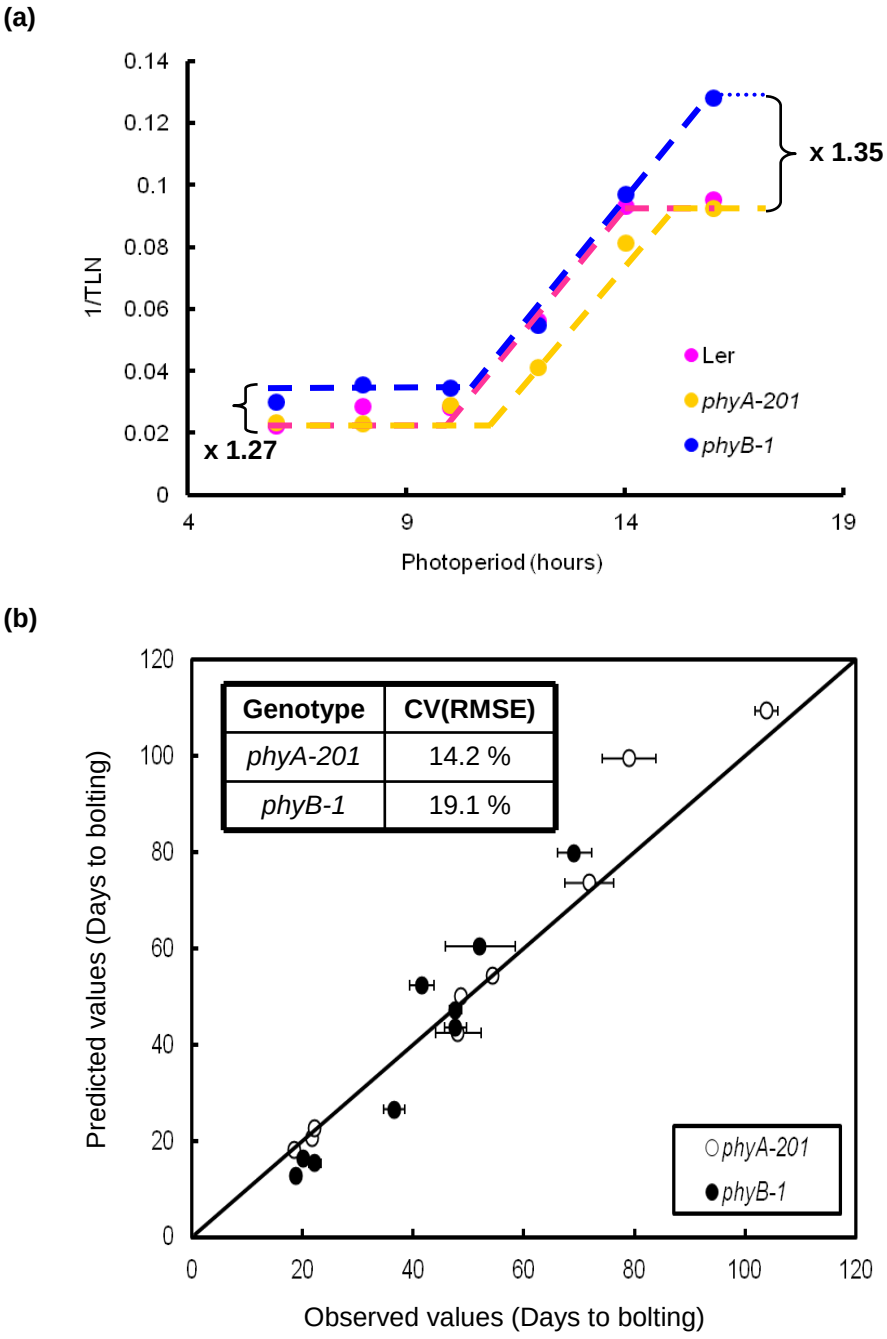


Figure 4.1.8: Model adjustment for *phyA-201* and *phyB-1* mutants. (a) Rate to bolting as estimated by the reciprocal of Total Leaf Number (TLN). The TLN data were taken from Giakountis et al. (2010), where experiments were conducted over a range of photoperiod from 6 to 16 hours. Curly brackets indicate the proportional differences between wild type and mutant. For *phyB-1* mutant, it was assumed that a maximum rate would be achieved at 16 hours and above (blue dotted line); (b) Predicted versus observed bolting times of *phyA-201* and *phyB-1* mutants using Model 2 after modification in the photoperiod component. Error bars represent one standard error.

According to Fig. 4.1.8a, the rate to bolting for *phyA-201* during long days (with photoperiod above 10 hours) was lower, correlating with a loss of phyA activity in stabilising CO protein (Fig. 2.2) (Valverde et al., 2004). However, the maximum rate was not altered, as there are other layers of CO regulation by FKF1 and GI (Salazar et al., 2009). For *phyB-1*, the rate was higher in general, following the role of activated phyB in promoting the degradation of CO protein in the morning (Valverde et al., 2004). In the model adjustment, it was assumed that *phyB-1* rate to bolting would achieve its maximum at photoperiods of 16 hours or above. Adjusted parameters are listed in Appendix A. Fig. 4.1.8b shows that the modified model could predict the bolting times of *phyA-201* mutant grown in the same field plantings in Wilczek et al. (2009) (Appendix A3), with a RMSE of 7.4 days (14.2 %). The RMSE for *phyB-1* was 7.6 days (19.1 %) but deviations were uneven, with all Spring/Summer cohorts underestimated while Autumn cohorts were overestimated.

Published data have shown that phyB has a temperature-dependent role in flowering in addition to its function in the photoperiod pathway (Halliday et al., 2003; Halliday & Whitelam, 2003; Valverde et al., 2004). Earlier results in this study suggested that the phase relationship between light and temperature cycles was important. Therefore analysis of the *P* function in the thermal-gating model (Fig. 4.1.4) might reveal information regarding the dual role of phyB in light and temperature signalling. Using the new photoperiod parameters for *phyB-1* as described above (Fig. 4.1.8a), both the day and night factors (P_{day} and P_{night}) in step gating Model 2 (Fig. 4.1.4c) were re-parameterised to fit the *phyB-1* field data while holding all other parameters to the values estimated earlier for wild-type. Intriguingly, an optimal fitting was achieved when P_{day} and P_{night} were at values of 0.5959 and 0.6856, respectively. This unexpected result suggested that temperature gating was almost abolished in phyB-deficient plants. To test this, the model for *phyB-1* was re-parameterised by constraining P_{day} and P_{night} to be equal and constant (Fig. 4.1.9b), in other words abolishing the gating effect. The modified model, with a constant 'gating' factor of 0.6279, showed a marked improvement with a RMSE of 4.9 days (12.5 %) (Fig. 4.1.9a) compared to 7.6 days (19.1 %) in Fig. 4.1.8b. Nevertheless, this improvement could be a mathematical artefact to compensate for the changes made in the

photoperiod component. However, using Model 2 alone without any modification, i.e. the model for wild-type Ler, resulted in a RMSE of 19.8 days (50.0 %) suggesting that modification(s) was indeed required to describe *phyB-1* field data. To further investigate this, the estimation of P_{day} and P_{night} was repeated but without altering the photoperiod parameters. In this case, *phyB-1* mutants experienced fully accelerated rate on days of intermediate (14 hr) and not just very long (16 hr) photoperiods. The optimised values for P_{day} and P_{night} were 0.8291 and 0.6992, respectively, which again displayed a reduced gating effect. Comparison of AICc values for different modification schemes (Table 4.1.2) supported the notion of constant gating for *phyB-1*. In addition, the model with double modifications that embodied a constant gating showed the lowest set of RMSE and AICc values. These results suggested that in order to describe the *phyB-1* mutant field data, both the photoperiod and thermal-gating modifications were required.

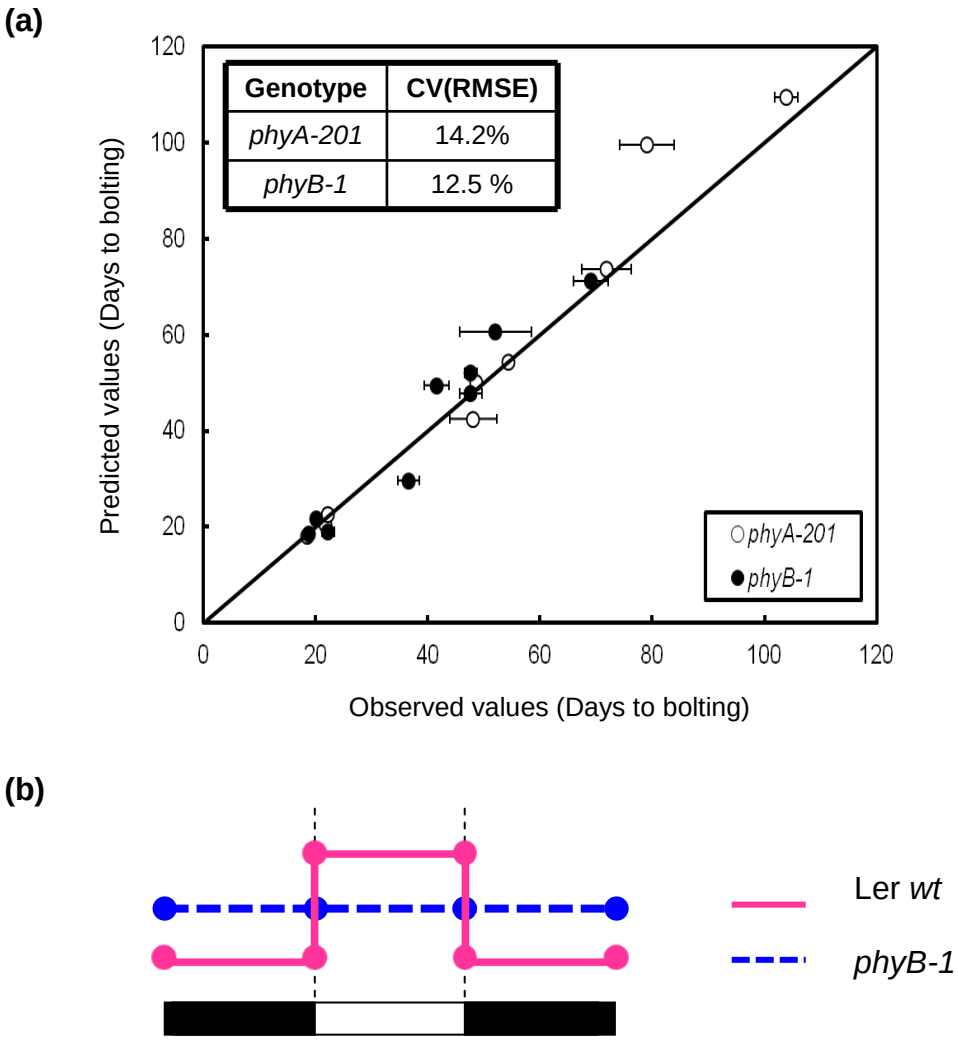


Figure 4.1.9: Dual role of phyB in light and temperature signalling: (a) Predicted versus observed bolting times of *phyA-201* and *phyB-1* mutants using Model 2 (step gating). The photoperiod component was modified according to Fig. 4.1.8a for both mutants. For *phyB-1* a constant gating of 0.6279 was adopted. The observed values are field data from the same plantings in Wilczek et al. (2009). The diagonal line represents perfect fit; (b) Proposed thermal-gating mechanisms of wild-type plant (Ler *wt*) and *phyB-1* mutant.

Table 4.1.2: Comparison between model adjustments for *phyA* and *phyB* mutants

Modification in Model 2 for *phyB-1*	**No. of Estimated Parameters**	**RMSE**	**AICc**
Photoperiod and Gating	2 (P_{day}, P_{night})	4.7 days	18
Photoperiod and Constant Gating	1 ($P_{day=night}$)	4.9 days	15
Gating only	2 (P_{day}, P_{night})	6.1 days	20
Constant gating only	1 ($P_{day=night}$)	5.5 days	16

4.1.6 Discussion

Modelling is a useful tool not only for predictive study, but analysis of model behaviour could extract additional information out of complicated data. The variation in model performance (Fig. 4.1.2) and improvement of model GoF (Fig. 4.1.6 and Fig. 4.1.7) in the current study suggested a seasonal effect of night temperature on flowering time. Meteorological data indicated that Spring/Summer cohorts typically experienced warmer days than nights and photoperiods of 10 to 20 hours (Fig. 4.1.3; Wilczek et al. (2009)). The original Wilczek et al. photothermal model, which considered vernalising temperatures during both day and night but only daytime temperatures in the accumulated thermal time, was able to accurately predict the flowering time of Spring/Summer cohorts, but performed less well for the majority of Autumn cohorts (Fig. 4.1.2b). This suggests that, at least under long photoperiods, the dramatic rise (of up to 20^0C) in near-ground daytime temperature is a strong determinant of flowering time, while reduced night temperatures have little impact so they can be ignored in the model. Incorporating night temperature in the model either through gradual gating (Model 1) or step gating (Model 2) improved the fit for the Autumn data without causing changes for the Spring/Summer data. These results comply with the dramatic switch in diurnal temperature pattern in the autumn where the occurrence of non-vernalising promotory temperatures during the night was comparable to that during the day (Fig. 4.1.3). Model validation using flowering data of plants subjected to a range of day/night temperatures in the laboratory (Thingnaes et al., 2003) confirmed the importance of including night temperature effects (Fig. 4.1.7), particularly when day temperatures are cooler.

Increased overall accuracy in the thermal-gating models was due to the improved fit to Autumn data, where plants had been subject to cooler but still inductive temperatures, with less predictable daily fluctuations. The GA pathway is known to regulate flowering time across photoperiods, but it has a predominant role in SDs (Mutasa-Gottgens & Hedden, 2009). As day length shortens in the autumn, there is a switch from the CO-photoperiod to GA floral regulatory pathway (Moon et al., 2003; Wilson et al., 1992). The increased responsiveness to night temperature under SDs

may reflect the changeover to GA signalling that is known to have differential sensitivity to day and night temperatures (Arana, Marin-de la Rosa, Maloof, Blazquez, & Alabadi, 2011; Stavang et al., 2009; Stavang, Junttila, Moe, & Olsen, 2007). Modelling approach in the current study highlights the importance of this change in responsiveness to non-vernalising promotory temperatures from daytime in LDs to day/night in SDs.

Neither the original nor the improved models could fit the Norwich Autumn *gi-2*, Col-*FRI-Sf2*, *vin3-1* and *fve-3* data as these genotypes were unusually late flowering at this location. As the Norwich planting occurred earlier in contrast to Cologne and Halle, plants in Norwich were not exposed to vernalisation for a considerable time post-germination (Appendix A4). Instead plants were exposed to frequent drops in daily temperature, conditions that have been shown to induce C-REPEAT BINDING FACTORS (CBFs) and boost FLC expression levels (Seo et al., 2009). This period of intermittent cold could have caused an increase in FLC levels and thus increased the vernalisation requirement for these late flowering genotypes. Additionally, there appears to be a feedback loop where FLC also enhances CBF expression through inhibition of SOC1, a negative regulator of CBFs (Seo et al., 2009). This causes a further delay of flowering. Including the crosstalk between CBFs and FLC in future models therefore may improve their accuracy.

Previous work in the literature has implicated the photoreceptors phyA and phyB in flowering time regulation, operating through both CO-dependent and independent pathways (Cerdan & Chory, 2003; Ishikawa, Kiba, & Chua, 2006; Suarez-Lopez et al., 2001; Valverde et al., 2004). Model with modified photoperiod parameters in the current study could predict *phyA-201* bolting time of the field cohorts with reasonable accuracy. This result suggests that under field conditions phyA may operate largely through the photoperiod pathway. On the contrary, modification of photoperiod parameters alone was insufficient to match the *phyB-1* bolting data. Remarkably, an accurate fit to data was in fact only achieved when the temperature gating was removed in addition to the photoperiod adjustments. The implication here is that wild type phyB is not only required for photoperiod perception, it also

mediates the impact of day and night temperature on flowering time. This potential role for phyB in temperature gating is in agreement with the reported temperature-dependent phenotype of *phyB* in flowering time (Blazquez et al., 2003; Halliday et al., 2003; Mazzella, Bertero, & Casal, 2000). This proposition also offers a mechanistic explanation for the paradoxical observation that phyB-overexpressors and *phyB* loss-of-function mutants are both early flowering (Bagnall et al., 1995; Reed, Nagpal, Poole, Furuya, & Chory, 1993). If *phyB*-null mutants lose the ability to sense day temperature effectively compared to the wild type (Fig. 4.1.9b), phyB-overexpressors would be predicted to have increased sensitivity to day temperature and this could lead to a more rapid accumulation of MPTUs at day time and thus early flowering. Certainly, co-ordination between light signalling, temperature perception, the circadian clock regulation and the photoperiod pathway enables tracking of seasons and ensures plants flower under favourable conditions (Franklin, 2009).

PART II: Linking the Clock to Phenology

Summary: This section describes how a clock gene circuit model was connected to the phenology model. Model 2 (step gating) from the previous section was used. The combined model allowed the study of gene circuit dynamics in response to changes in the natural environments. Simulation of clock mutants with different free-running periods also highlighted the complex mechanism associated with daylength responses in the induction of flowering.

4.2.1 Photoperiodic and clock control of flowering in Arabidopsis

It has been outlined earlier in Chapter 2 how the circadian clock regulates flowering time through a cascade of activities at the molecular level (Fig. 2.2). The combined antagonistic and protagonistic actions of clock-regulated GI, FKF1 and CDF1 (Fornara et al., 2009; Imaizumi et al., 2005; Song et al., 2012; Yanovsky & Kay, 2002) collectively result in *CO* peaking before the end of LDs, thus promoting *FT* expression. In SDs, as *CO* mRNA peaks during the night and CO protein is unstable at day time, *FT* mRNA levels remain low (Corbesier et al., 2007b; Valverde et al., 2004; Yanovsky & Kay, 2002). Such regulatory mechanism of *FT* provides a means for plants to time their life cycle according to the season. Here, a gene circuit model describing the regulation of *FT* by the clock (Salazar et al., 2009) was integrated into the Wilczek et al. phenology model, thus associating molecular dynamics to their effects at the phenological level.

4.2.2 Linking the clock gene network to the Arabidopsis phenology model

The clock gene circuit model (Locke et al., 2006; Locke et al., 2005) consists of a system of ODEs representing the dynamics of core genes in the clock network that are inter-regulated in feedback loops. Previous work has combined the clock model to the photoperiodic regulation of flowering time by modelling the activation of *FT* through *CO* (Fig. 4.2.1a) (Salazar et al., 2009). In that study, a few model variants were explored by considering different mechanisms for the activation of *CO* and *FT* with or without FKF1. These model variants have managed to reproduce the different *FT* mRNA abundance profiles during LD and SD. A recent study (Song et al., 2012) has also incorporated into the model the CDF1 component and its interaction with FKF1, which are responsible for producing the first CO peak at the end of long days (Fornara et al., 2009; Imaizumi et al., 2005). However, these models used experimental data of CDF1 and FKF1 as model inputs and thus could only describe flowering regulation at the few photoperiods where data were available. Future models that include these regulators as part of the modelled components would be required for studies under the natural range of photoperiod that varies daily. Therefore the current study utilised the simplest form of the Salazar et al. model (Model 3 that does not include FKF1). This simple model may not fully represent the exact biphasic dynamics of CO (Fig. 2.2) or the phase relations of all clock components; however it could provide good qualitative output of flowering regulation relative to light period (Fig. 4.2.1) (Salazar et al., 2009), which was the aim of this study. All the equations and parameter values for this model are displayed in Appendix B.

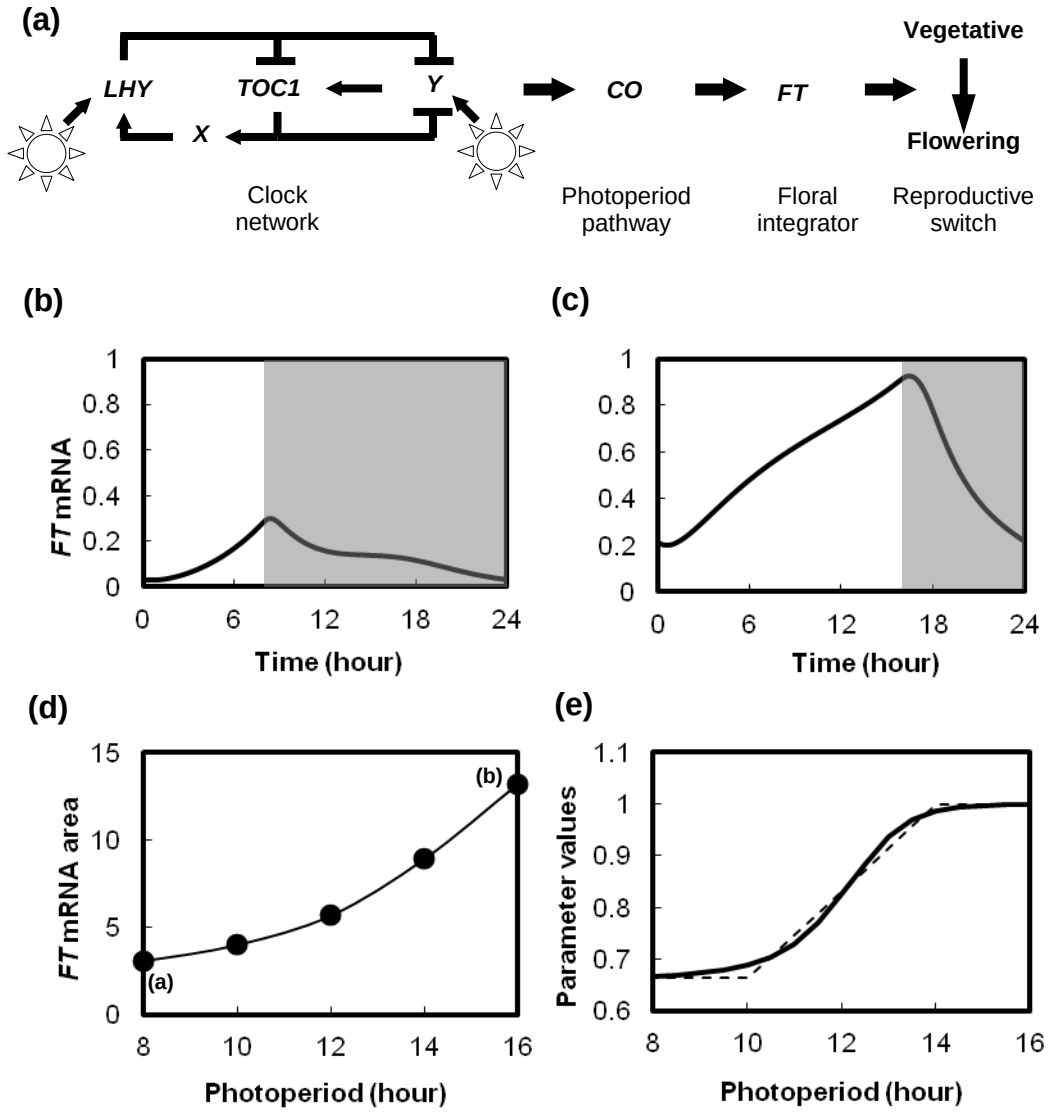


Figure 4.2.1: Calibrating the clock gene circuit model (Salazar et al., 2009) to the photoperiod component of the Arabidopsis phenology model (Wilczek et al., 2009). (a) Schematic diagram of the clock-photoperiod circuit model (Model 3 in Salazar et al. (2009)). In this model, *X* and *Y* are hypothetical components required to describe experimental data, and *GI* was found to be a strong candidate for *Y* (Locke et al., 2006). The time series of *FT* mRNA simulated by the clock model for: (b) 8-hour photoperiod and; (c) 16-hour photoperiod; (d) A plot of the integral (area under the curve) of simulated *FT* mRNA time series at different photoperiods. Letters in brackets link to the associated time series. These integrals were translated into the photoperiod component using the sigmoid function in equation 4.2.1; (e) Comparison between the piece-wise linear transition function in Wilczek et al. (2009) (dashed line) and the calibrated photoperiod component (full line). The parameter values of this component are constrained to values between 0 and 1 as they are used as scaling factors.

To link both models, the piece-wise linear transition function of the photoperiod component in the phenology model was replaced by the Salazar et al. clock-photoperiod circuit using the following calibration equation (equation 4.2.1). A sigmoid function as described in Salazar et al. (2009) was used, where developmental rate reached its limits at both photoperiod extremes after certain critical day lengths (Welch, Dong, Roe, & Das, 2005; Welch, Roe, et al., 2005; Wilczek et al., 2009). This function relates the integral of *FT* mRNA over 24 hours produced by the clock model for each photoperiod, to the associated scaling factor in the photoperiod component (equation 4.1.1), which is bounded by a minimum value D_{SD} and a maximum value of 1 (D_{LD}).

$$Photoperiod = a + b\left[\frac{c^n}{c^n + (FTarea)^n}\right]. \qquad (4.2.1)$$

All the constant values are listed in Table 4.2.1.

Table 4.2.1: Parameter values for the sigmoid function that calibrates the integral of simulated *FT* mRNA to the photoperiod component in the phenology model for both the Wilczek et al. and Model 2 (step gating) variant.

Parameter	Wilczek et al. model		Model 2	
	Ler	Col	Ler	Col
a	1	1	1	1
b	-0.3374	-0.3985	-0.3635	-0.3985
c	5.733241	5.733241	5.733241	5.733241
n	7	7	7	7

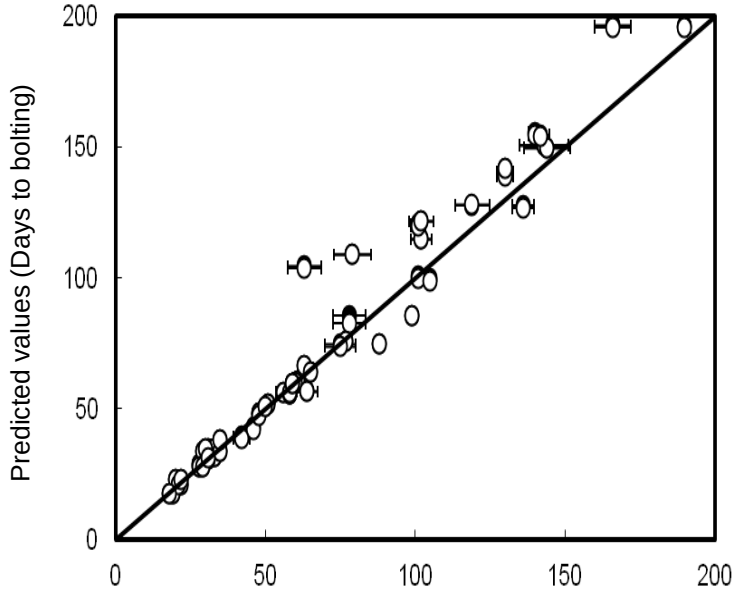
Theses values were optimised using Excel Solver by constraining the maximum and minimum values of the sigmoid function to equal D_{LD} and D_{SD} (equation 4.1.1), respectively. *n* was also constrained to take an integer value.

All the parameter values in equations 4.2.1 were determined by entraining the clock model to different constant photoperiods. During each model entrainment, a unique solution was reached for every photoperiod where the concentration level of each component achieved a stable cycle of 24-hour period (Figs. 4.2.1a and 4.2.1b). In the natural environment, day length changes every day, and the concentration level of each component is carried forward to the following day. Therefore, for simulation of flowering time using the combined model, only the first day was entrained. The concentration levels of all components at the end of each day were then used as the initial values in the ODE simulation for the next day.

One feature of the Wilczek et al. phenology model (Chapter 4: Part I) is that the photoperiod component for every hour on the same day shares the same value, because day length is determined by the sunrise and sunset times of the associated day. This indirectly assigns the notion that plants know in advance the time of sunrise and sunset before the instant, which should not be the case biologically speaking. However, as the phenology model was parameterised with this inherent feature (Wilczek et al., 2009), the formality was retained in the calibrated model. Therefore the photoperiod component for every hour of the same day was calculated by summing up the concentration of *FT* mRNA from hour 0 to hour 24, and then calibrated using equation 4.2.1.

Both the models calibrated to the Wilczek et al. phenology model and step-gating Model 2 could reproduce all the predicted values from the original non-calibrated models as shown in Fig. 4.2.2. In the subsequent study, only Model 2 was utilised.

(a)



(b)

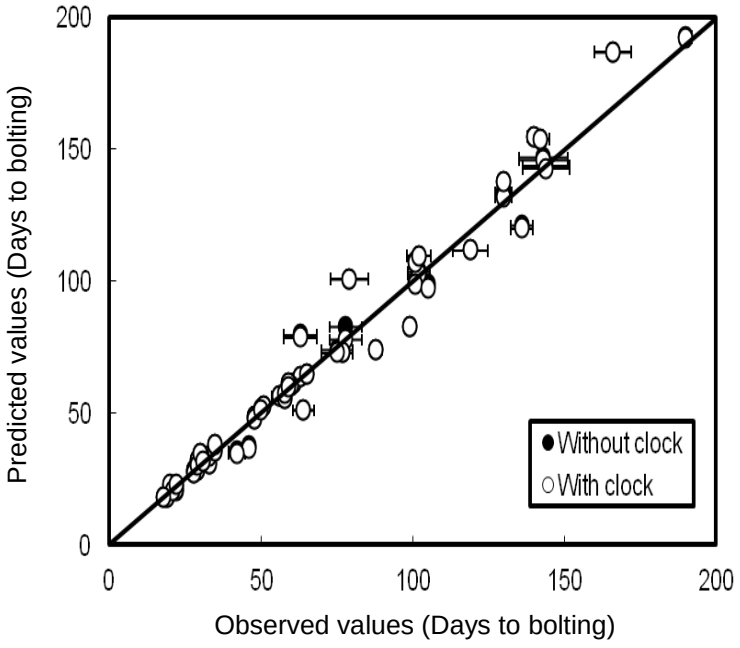


Figure 4.2.2: Days to bolting of different genotypes as predicted by the clock-linked (open circles) and non-linked (closed circles) models for: a) Wilczek et al. version; b) Model 2 (step gating) version of the photothermal model. The genotypes shown were those published in Wilczek et al. (2009). Closed circles are not really seen because they overlap with the open circles, indicating that the clock-linked models could successfully reproduce the predictions of the original non-linked models.

As shown in the previous section (Chapter 4: Part I), each genotype behaved differently depending on the growing season and the model has managed to capture this behaviour by changing the genotype-associated parameter(s). Therefore, linking the clock gene circuit to the phenology model could enable the prediction of flowering time behaviour caused by mutations of the clock components. This would increase our understanding of clock functions and its implications on flowering time, as described next.

4.2.3 Simulation of clock period mutants

To generate clock mutants with different free-running periods, the expression rate of the clock component *LHY* (and a partially redundant gene *CCA1*) was modified in the model. Short-period mutants were generated by simulating a single and double mutation of *LHY*/*CCA1* through reduction of *LHY* translation rate ($v6$ in equation B2.3) to 0.5 and 0.001 of its original value, respectively. A long-period mutant was simulated by increasing *LHY* translation rate ($v6$) by a factor of 10. The free-running period of these mutants under constant light, τ_{LL}, was measured as the time interval between successive expression peaks of *TIMING OF CAB 1* (*TOC1*), a clock component upstream of *CO* in the model. The values are as shown in Table 4.2.2.

Table 4.2.2: Modification and scaling factors used to generate clock period mutants

Genotypes	Modification factor of *LHY* translation rate	Scaling factor (*SF*) for TOC1 protein level	Free-running period under constant light, τ_{LL}
Wild type	-	-	25
lhy/*cca1* single mutant	0.5	0.79	23
lhy/*cca1* double mutant	0.001	0.34	18.5
Long-period mutant	10	1.44	27

As the simulated expression level of *TOC1* was affected by the parameter modifications (Fig. 4.2.3a), a scaling factor (*SF*) was used to re-adjust TOC1 protein peak level using wild type as the standard (equation 4.2.2, from equation B2.14). This was done to ensure that changes in the expression level of the photoperiod components, *CO* and *FT*, were caused only by phase shift of the clock components.

$$\frac{d[\mathrm{CO}]}{dt} = vCOm \cdot SF \cdot [\mathrm{TOC1}_n] - \frac{(1-\Theta_{light}) \cdot vCOp \cdot [\mathrm{CO}]}{kCOp + [\mathrm{CO}]}, \qquad (4.2.2)$$

where *vCOm* is the rate constant of *CO* (or *TOC1*) mRNA translation, *vCOp* is the maximum rate of light-dependent CO protein degradation, and *kCOp* is the Michaelis constant of CO protein degradation.

The period mutants were next simulated using as entrainment input the naturally changing photoperiod in various latitudes (Wilczek et al., 2009). This enabled the study of clock gene dynamics in real seasonally changing daylength, performed for the first time in our clock gene circuit models. The flowering time of period mutants simulated in different seasons were also generated using the newly calibrated phenology model (Section 4.2.2). Additionally, simulation of plants without any functional clock was conducted where the modelled plants were not sensitive to changing photoperiods. In these control studies, flowering time was generated by setting the photoperiod parameter of the phenology model always at the lowest (as if under constant SD), medium (12 hours' daylight) or highest (LD) photoperiod induction.

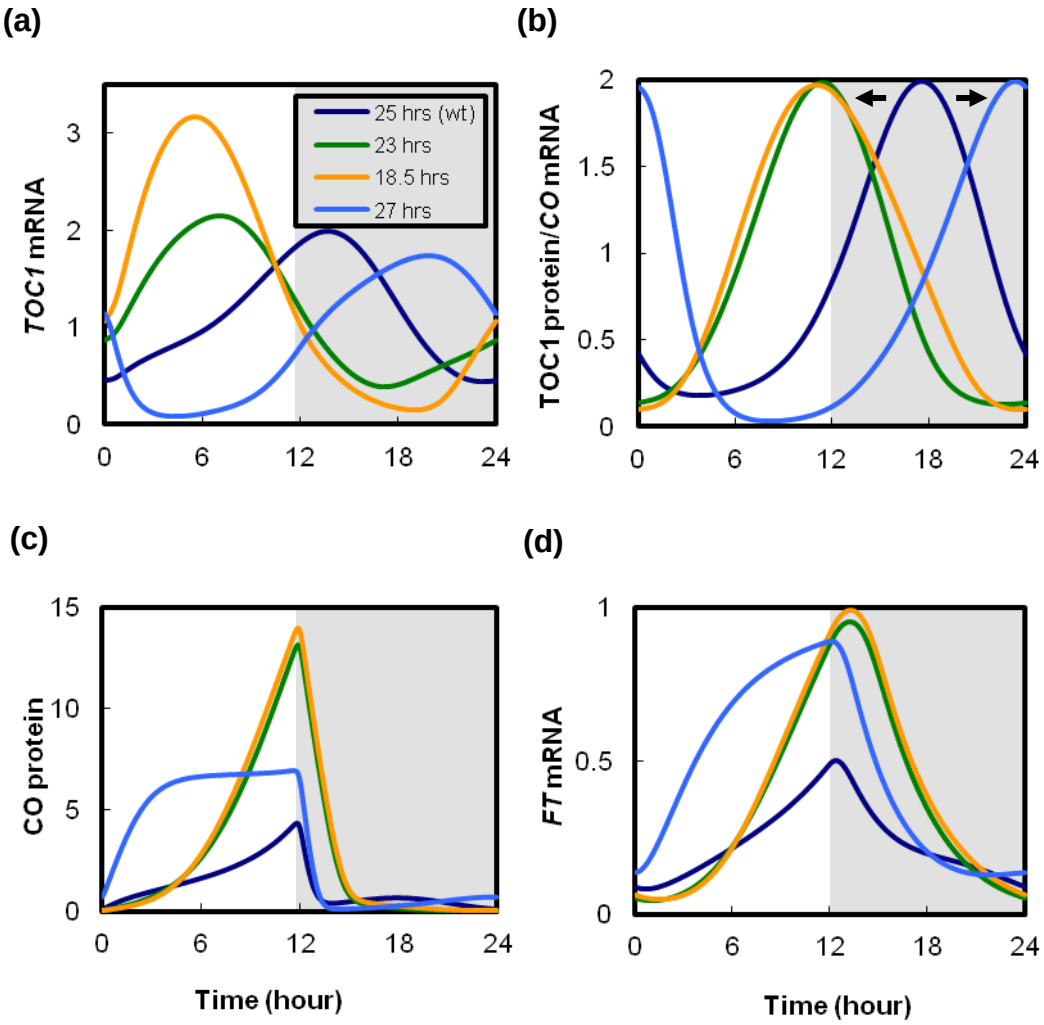


Figure 4.2.3: Phase relationship of simulated *TOC1* mRNA (a), TOC1 protein/*CO* mRNA (b), CO protein (c) and *FT* mRNA (d) time series between clocks with different free-running periods, entrained under a 12:12 light/dark cycle. The peak levels of TOC1 protein for the period mutants have been re-scaled during simulation, therefore the different levels of CO protein (c) and *FT* mRNA (d) shown were only due to the phase shift in the period mutants relative to the wild type as indicated by the arrows in panel (b).

4.2.4 Simulation of flowering time for period mutants in different seasons

The calibrated phenology model predicted that all the period mutants would flower much earlier than the wild type if planted in the autumn (Fig. 4.2.4). The long-period mutant simulation (τ_{LL} = 27 hrs) deviated from wild type the most, flowering at the same time as the control study with constant maximal photoperiod induction. The 18.5-hour-period and 23-hour-period mutants were also predicted to flower earlier than the wild type, though slightly later than the long-period mutant. All genotypes with different free-running periods were expected to flower at the same time in the Spring/Summer.

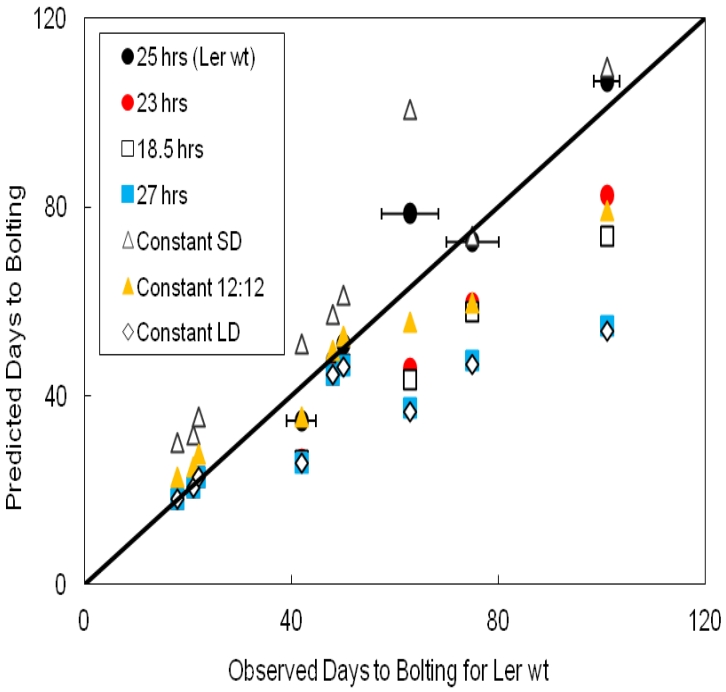


Figure 4.2.4: Predicted days to bolting of Ler wild type, period mutants (with the legend showing the free-running period) and controls (which are simulated plants without a functional clock). The abscissa represent the observed days to bolting of wild type plants in different plantings in the field (Wilczek et al., 2009), while the ordinate shows the values predicted by the calibrated model for each simulated genotype. The three controls were generated by setting the photoperiod component constantly at high (long-day induction), medium (12-hour photoperiod) or low (short-day) value.

To determine the basis for the early flowering observed here in all the mutants, the dynamics of *CO* and *FT* in the period mutants were compared (Fig. 4.2.3b-d). Phase shifts in *CO* expression were observed in the period mutants (Fig. 4.2.3b), as has been reported previously (Mizoguchi et al., 2005) and these changed the relative abundance of CO protein during the photoperiod (Fig. 4.2.3c). In short-period mutants, a large part of the expression peak was advanced into the light period. On the other hand, the phase of *CO* mRNA in the long-period mutant was delayed such that the level at dawn was still high (Fig. 4.2.3b), which led to a high abundance of CO protein that remained throughout the day (Fig. 4.2.3c). Consequently a sharp increase in *FT* expression was simulated early in the morning from ZT1 (or 1 hour after dawn) onwards in the long-period mutant (Fig. 4.2.3d). Consistent with this, analysis at different photoperiods illustrated that a large percentage of the daily abundance of *CO* mRNA was expressed during the day in long-period mutant under most conditions, while the percentage in short-period mutants increased sharply from very low photoperiod onwards compared to the wild type (Fig. 4.2.5a). Concurrently, the photoperiodic window for floral induction shifted dramatically in the period mutants (Fig. 4.2.5b). For short-period mutants, their critical day lengths were 2 hours earlier than wild type but the photoperiodic response in long-period mutant was 6 hours in advance. Thus, the simulated long-period mutant started being responsive to photoperiod at a critical short day length of 4 hours and reached its maximum at around 8 hours, a condition commonly adopted as SD in most studies. These again explained the early-flowering phenotype of the simulated mutants, particularly for the autumn cohort.

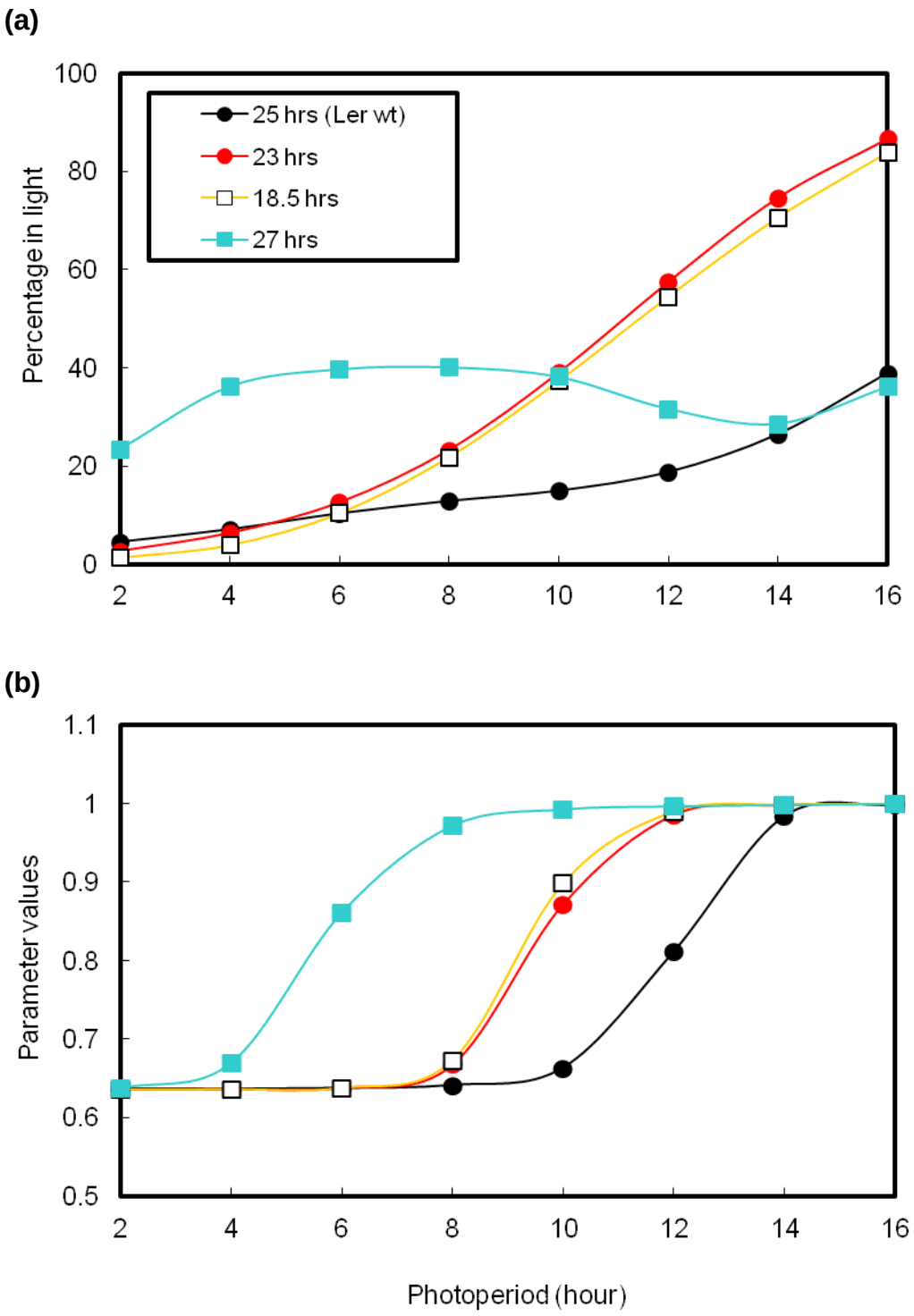


Figure 4.2.5: (a) The percentage of *CO* mRNA in the light relative to its total expression for each simulated period mutant at different photoperiods; (b) Parameter values of the photoperiod component at different photoperiods for each simulated genotype. The parameter values were computed from equation 4.2.1.

4.2.5 Gene dynamics in naturally occurring photoperiods

Molecular mechanisms in response to the environment are usually discovered based on laboratory studies that implement discriminating and often unnatural conditions. Therefore, comprehensive ecological observations are essential for the collection of meteorological, physiological and genomic data to reveal the adaptive strategies of plants in their natural habitats.

This study has linked for the first time the clock gene circuit to meteorological conditions at various latitudes and in different seasons, and this could facilitate the study of gene dynamics in realistically varying conditions. For example, the dynamics of *FT* mRNA across the season between sites at different latitudes could be compared (Fig. 4.2.6 and Appendix B4). For wild type plants, sensitivities of *FT* expression were higher at long photoperiods (sharp peaks against flat troughs). This increased sensitivity to LD in the simulation corresponds with Arabidopsis as a long-day plant. All simulated mutants displayed sharp troughs, indicating higher sensitivity relative to wild type in short days. Simulation for long-period mutant displayed steeper peaks compared to short-period mutants, suggesting it has a higher sensitivity at long photoperiods. Noisy fluctuations could be seen in the beginning of the simulation for Norwich and Oulu Summer plantings (Appendix B4), when the initial photoperiods were high. This was due to the unstable entrainment of the clock model at photoperiods above 16 hours, which is a characteristic of this model version. Nevertheless, this would not affect the simulation of flowering time as the developmental rate saturates at these photoperiods (Fig. 4.2.5).

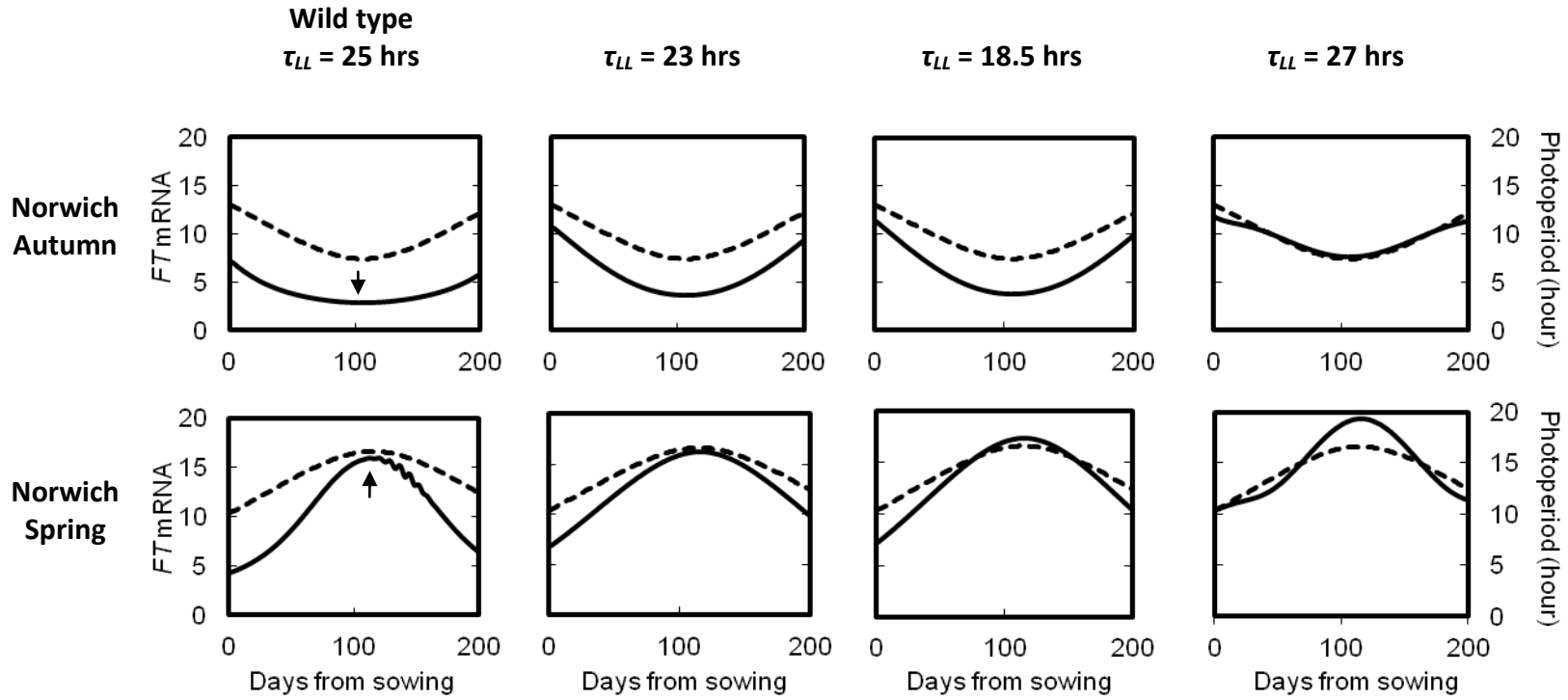


Figure 4.2.6: Integral (over 24 hours) of simulated *FT* mRNA level (full line) for wild type and period mutants in the Norwich Autumn and Spring plantings. The plantings were the same as those published in Wilczek et al. (2009). The simulated free-running period of each genotype under constant light, τ_{LL} (after entrainment in 12:12 light/dark cycle), is shown at the top of the column. Dashed lines indicate the naturally occurring photoperiod. Arrows point to the peak and the trough of expression at high and low photoperiods, respectively.

Another behaviour that can be observed from this analysis is that photoperiodic induction does not seem to depend on the direction or rate of change in day length in wild type plants. Fig. 4.2.7 shows that the photoperiod induced *FT* mRNA abundance was symmetrical, indicating that it was not affected by whether day length was increasing or decreasing. Sensitivity analysis of *FT* expression was also conducted to determine the effect of rate of change in day length. Here, *FT*-expression sensitivity was defined as the unit change in the integral of simulated *FT* mRNA for each unit of change in daily photoperiod, which is dependent on latitude. *FT*-expression sensitivity was then compared between sites of various latitudes (varying up to 20^0), which experience different rates of photoperiod change (Fig. 4.2.8). There was very little variation between sites, especially in low photoperiods. The scatter of data points at high photoperiods for some sites was a numerical artefact; the photoperiods (the denominator) were recorded to the nearest 0.6 minute, thus any small noise in the record was significantly amplified in the calculated sensitivity. An analysis of covariance (ANCOVA) was also conducted to investigate any variation between sites. First, the hypothesis that all the sites had different correlations was tested (Table 4.2.3a). However, no significant difference was found for the interaction between site and photoperiod (with a *p*-value of 0.1708, which was very much larger than the alpha value of 0.05), indicating that the slope of the correlations at different sites was similar. Therefore, the intercept was tested next, setting the correlations to be parallel (Table 4.2.3b). Again, no significant difference was found between the three sites (*p*-value = 0.2117). The conclusion here was that the clock had the same sensitivity regardless of sites or the rate of photoperiod change. When all the data were pooled together, a strong dependence of *FT*-sensitivity on photoperiod was found (*p*-value = 9.7833e-151; R^2 = 0.9517). This was consistent with the expression analysis in Fig. 4.2.6 showing an increased sensitivity at high photoperiods. These results collectively suggested that photoperiod induction of flowering in Arabidopsis is dependent mainly on the absolute photoperiod.

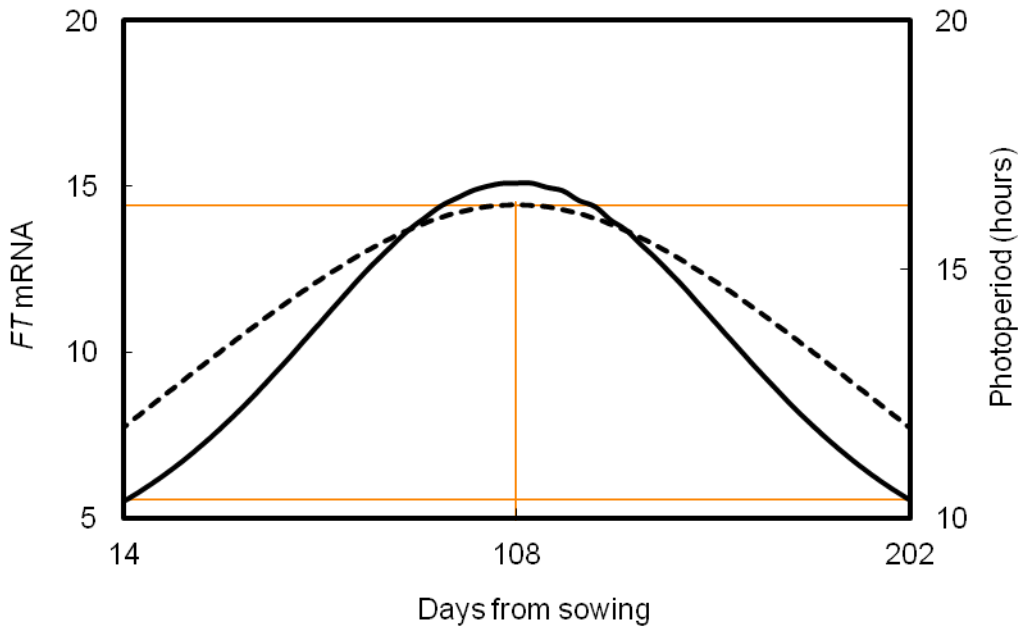


Figure 4.2.7: Integral (over 24 hours) of simulated *FT* mRNA level (full line) for wild type in naturally changing photoperiod (dashed line). The example shown here is data taken from Cologne in the spring. Brown lines indicate the symmetry in *FT* abundance on both sides of the maximum day length.

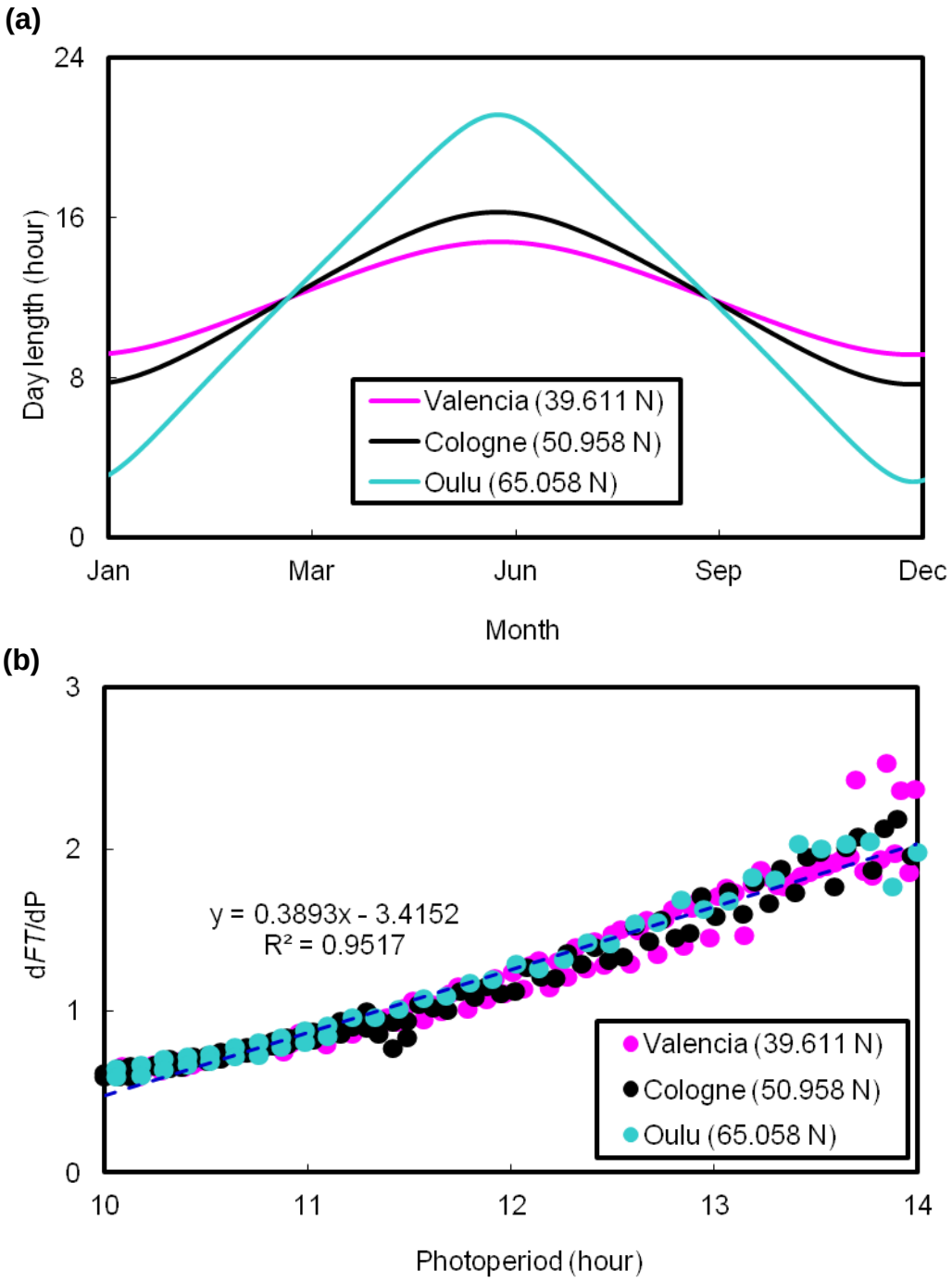


Figure 4.2.8: Sensitivity analysis of *FT* expression simulated in wild type plants at different sites. (a) Day length variation in Valencia, Cologne and Oulu; (b) Graph showing the unit change in the integral of *FT* mRNA per unit daily change in photoperiod at different day lengths for the three sites. The latitude of each site is shown in brackets in the legend. Only a photoperiod range between 10 and 14 hours (the critical day lengths) were considered as the effect of photoperiod on flowering plateaus at both extremes. The dashed line was the linear regression of all the data regardless of sites. The equation and regression coefficient are also displayed on the graph.

Table 4.2.3: Results from the analysis of covariance (ANCOVA) at alpha = 0.05 comparing *FT*-expression sensitivity at different sites

a)

Source	d.f.	Sum Sq	Mean Sq	F	Prob > F
Site	2	0.0346	0.0173	1.57	0.2094
Photoperiod	1	48.5317	48.5317	4413	0
Site*Photoperiod	2	0.0392	0.0196	1.78	**0.1708**
Error	222	2.4414	0.011		

b)

Source	d.f.	Sum Sq	Mean Sq	F	Prob > F
Site	2	0.0346	0.0173	1.56	**0.2117**
Photoperiod	1	48.5317	48.5317	4382.42	0
Error	224	2.4806	0.0111		

c)

Source	d.f.	Sum Sq	Mean Sq	F	Prob > F
Photoperiod	1	49.5927	49.5927	4456.01	**9.78329e-151**
Error	226	2.5152	0.0111		

4.2.6 Discussion

Day length is one of the important environmental signals for plants to sense daily and seasonal changes (Franklin & Whitelam, 2004). Characteristics of the circadian oscillators enable the system to provide a good mechanism for sensing day length in the regulation of growth and development (Pittendrigh & Daan, 1976b). Without a clock (such as in the controls in Fig. 4.2.4), plants may flower either too early or extremely late regardless of the day length or season. Such a strategy would not be ideal for optimal plant growth and survival as abiotic stresses vary with seasons.

In general, a circadian clock is the most stable under a 24-hour entrainment cycle, and allows plants to time their development according to the environment (Pittendrigh & Daan, 1976a, 1976b). However, natural variation in the free-running period of the clock (ranging from ~22 to 28 hours) has been reported and studies suggested that this variation was necessary to enhance plant fitness over a wide geographical range (Lou et al., 2011; Michael et al., 2003). The modelling work in this study has allowed us to understand the effects of clock variation, in a varying environment. When the external light/dark cycles differed from the plants' internal period, the simulated timing of gene expression was shifted relative to dawn and dusk. Accordingly, flowering time was also predicted to alter. Results in the current study were in agreement with published data which showed that *LHY*/*CCA1* loss-of-function mutants had shorter period and flowered at the same time as the wild type in LD but were early flowering in SD (Mizoguchi et al., 2002). Fig. 4.2.4 also concurs with their results showing that a double mutation has a more severe flowering time phenotype compared to a single mutation. On the other hand, a semi-dominant loss of function mutation in ZEITLUPE (ZTL), which caused an increase in free-running period, has been reported to delay flowering in LD (Somers, Schultz, Milnamow, & Kay, 2000). In contrast, a dosage-dependent delayed in flowering was also observed in ZTL overexpressors under LD, though it was suggested to be caused by the direct ZTL regulation of *CO*/*FT* expression independent of the clock (Somers, Kim, & Geng, 2004). A double mutation of PSEUDO-RESPONSE REGULATOR 7 and 9

(PRR7/PRR9) that exhibited a long period also flowered late in LD (Nakamichi, Kita, Ito, Yamashino, & Mizuno, 2005). Thus, the flowering behaviour of long-period mutants simulated in the current study contradicted most of those reported in the literature. In the model, early flowering was predicted for long-period mutants due to a delayed phase in the simulated *CO* mRNA causing a high level at dawn (Fig. 4.2.3b), and this has been observed in wild type plants subjected to light/dark cycles shorter than 24 hours (Roden, Song, Jackson, Morris, & Carre, 2002). The disagreement in flowering behaviour therefore suggested that the long-period mutant, which has a delayed phase (Fig. 4.2.3b), may not be just a mirror image of the short-period mutants, which have an advanced phase. Different molecular mechanisms may be involved in the opposite cases, as has been reported previously in hamsters (Schwartz, Tavakoli-Nezhad, Lambert, Weaver, & de la Iglesia, 2011), or more complex mechanisms may be involved in day length perception to increase flexibility of floral induction in different environments (Fornara et al., 2009; Roden et al., 2002; Song et al., 2012). Phase shift may also cause a difference in their coincidences with interacting factors. For example, studies using *CO*-overexpressors have shown that CO is highly unstable during the day (Fornara et al., 2009; Song et al., 2012) due to morning degradation of CO protein by phyB (Valverde et al., 2004) but this has not been included explicitly in the Salazar et al. model. Inclusion of phyB interaction in the photoperiod gene circuit model would be required to generate a more realistic CO level early in the morning, particularly to better match the effects of delayed phase observed in the literature.

Interestingly, some studies have proposed that the rate of change in photoperiod or the change in sunset/sunrise time could be more important as seasonal cues especially for plants in the tropical regions, given the negligible variation in day length near the Equator (Borchert et al., 2005; Clerget et al., 2004). On the other hand, studies on temperate species have reached contrasting conclusions. Strong correlation was observed between the rate of photoperiod change and leaf appearance rate for wheat and barley in southern England and Scotland (Baker, Gallagher, & Monteith, 1980; Ellis & Russell, 1984; Kirby, Appleyard, & Fellowes, 1982), but no significant effect was found on the developmental and leaf emergence rate in both

species in southern Australia (Kernich, Slafer, & Halloran, 1995; Slafer, Connor, & Halloran, 1994; Slafer, Halloran, & Connor, 1994). The modelling results for Arabidopsis in this study seem to be consistent with the latter, where the photoperiod pathway is only sensitive to absolute photoperiod in the control of flowering time (Fig. 4.2.7 and 4.2.8). This is also in agreement with Salazar et al. (2009), which showed that after entrainment to 8-hour photoperiod, the integral of simulated *FT* increased linearly upon transfer to higher photoperiod up to a drastic change of 8 hours in light period. However, as light intensity and light quality may also affect photoperiod sensitivity but are not currently considered in the model, future studies that incorporate these factors and comparisons with different plant species may increase our understanding on the diverse photoperiodic regulation of plant developmental event.

PART III: Arabidopsis Multi-Scale Model

Summary: The Arabidopsis multi-scale model is presented in this section. This integrated model incorporates the following, existing sub-models: a phenology model for the prediction of flowering time from Part I, a gene circuit of the circadian clock network from Part II, a process-based model describing carbon assimilation and resource partitioning, and a functional-structural module that determines shoot structure for light interception and root growth. All models are simplified, to limit the total complexity. The latter two models are first presented briefly, followed by a description of how they were combined in a modular fashion to form the multi-scale model. Validation using experimental data demonstrated a synergy of all model components. The results suggested that the multi-scale model can provide application routes from molecular and cellular biology to crop improvement and biosphere management.

4.3.1 Carbon assimilation and metabolism model

This module (Fig. 4.3.1) consists of two mechanistic sub-models, i.e. carbon assimilation and sugar-starch metabolism (Rasse & Tocquin, 2006). In the carbon assimilation sub-model, leaf photosynthesis is determined using the Farquhar et al. (1980) equations that consider CO_2 level, light intensity and temperature. This classical biochemical model considers the two rate-limiting factors in photosynthesis: the carboxylase activity of RuBisCO and electron transport (Farquhar, von Caemmerer, & Berry, 2001). An important parameter in the former is the maximum rate of carboxylation (V_{cmax}), which can be determined by conducting a series of gas exchange measurements at different CO_2 levels under saturating irradiance. On the other hand, electron transport is light dependent with a potential rate of J_{max} that can also be determined experimentally. In the Rasse and Tocquin (2006) study, only V_{cmax} was measured, while J_{max} was set to 2.1 times V_{cmax}, the same ratio as calculated in Farquhar et al. (1980).

In the sugar-starch partitioning sub-model, a fixed proportion of the photoassimilate calculated from the assimilation sub-model is turned into transitory starch at a baseline rate as suggested previously (Dewar, Medlyn, & McMurtrie, 1998; Sun, Okita, & Edwards, 1999), while the rest is converted into soluble sugars which are used for growth and respiration. Two types of respiration are considered in the model: maintenance respiration that depends on sugar concentration (Ogren, 2000; Rasse & Tocquin, 2006), and growth respiration which is a fixed fraction of the growth demand. The remaining carbon available for growth is then allocated to leaves and roots based on the root-to-shoot allocation ratio, *RS*. As growth demand is limited to a maximum rate, any excess photoassimilate is converted into starch through an overflow mechanism as suggested by Eichelmann and Laisk (1994), and Stitt (1996). At night time when there is no photosynthesis, starch is broken down into sugar to sustain growth. The model adopted a linear breakdown rate where a constant percentage of the end-of-day starch is consumed by the end of the night regardless of the night length (Fig 4.3.2). Such a behaviour has been observed in Arabidopsis (Gibon et al., 2004; Gibon et al., 2009) and was found to be circadian regulated (Graf, Schlereth, Stitt, & Smith, 2010; Lu, Gehan, & Sharkey, 2005). Translocation from the organs, i.e. leaves and roots, is also allowed to maintain a minimum sugar level. This element was created in the Rasse and Tocquin model to reflect the observations in an Arabidopsis starchless mutant in which sugar was consumed rapidly in the dark period but never reached zero, while genes involved in the breakdown of non-starch energy sources were up-regulated at the end of the night (Caspar, Huber, & Somerville, 1985; Thimm et al., 2004). Model equations and parameter values are listed in Appendix C. The parameter values were either extracted from the literature or determined through experimental measurements in their study (Rasse & Tocquin, 2006).

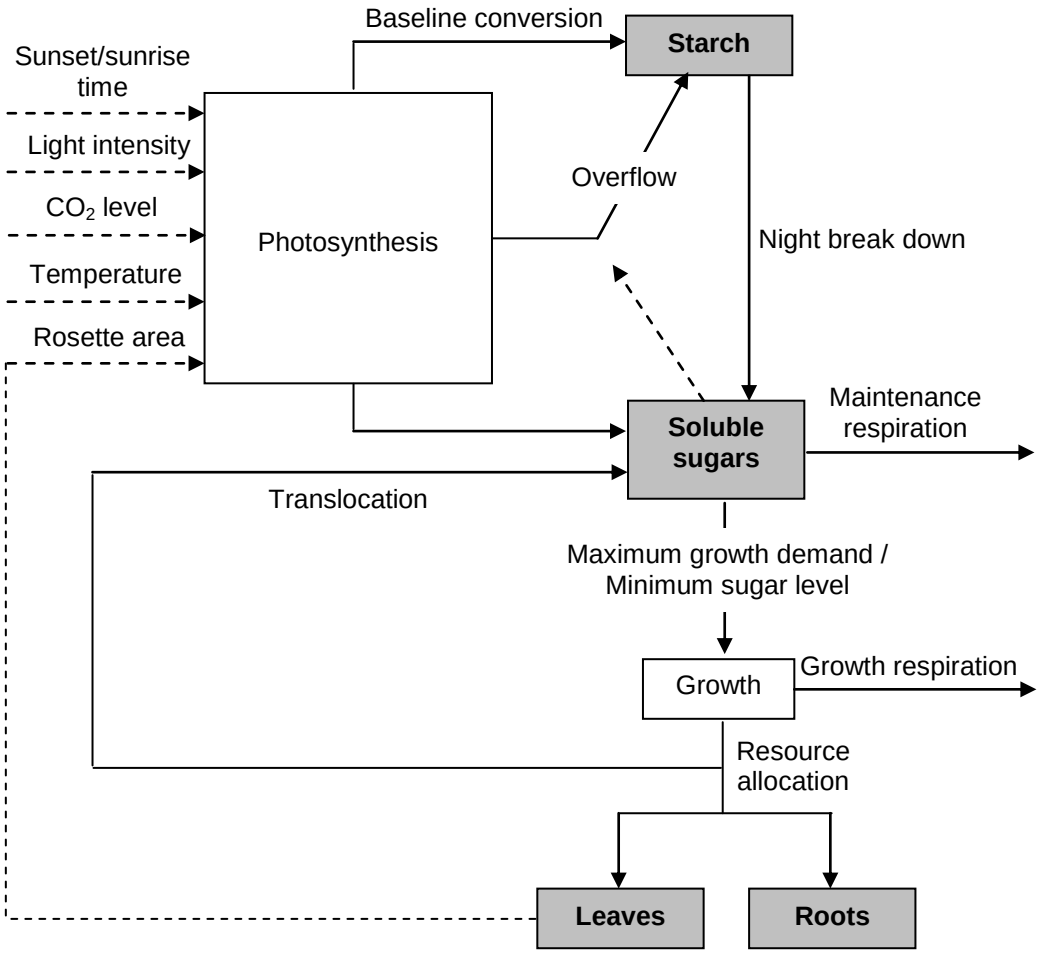


Figure 4.3.1: Schematic diagram of the carbon assimilation and sugar-starch metabolism model for Arabidopsis (Rasse & Tocquin, 2006). Carbon is transported (solid arrows) in and/or out of four different pools (grey boxes) through various processes. Dashed arrows indicate information or feedback input.

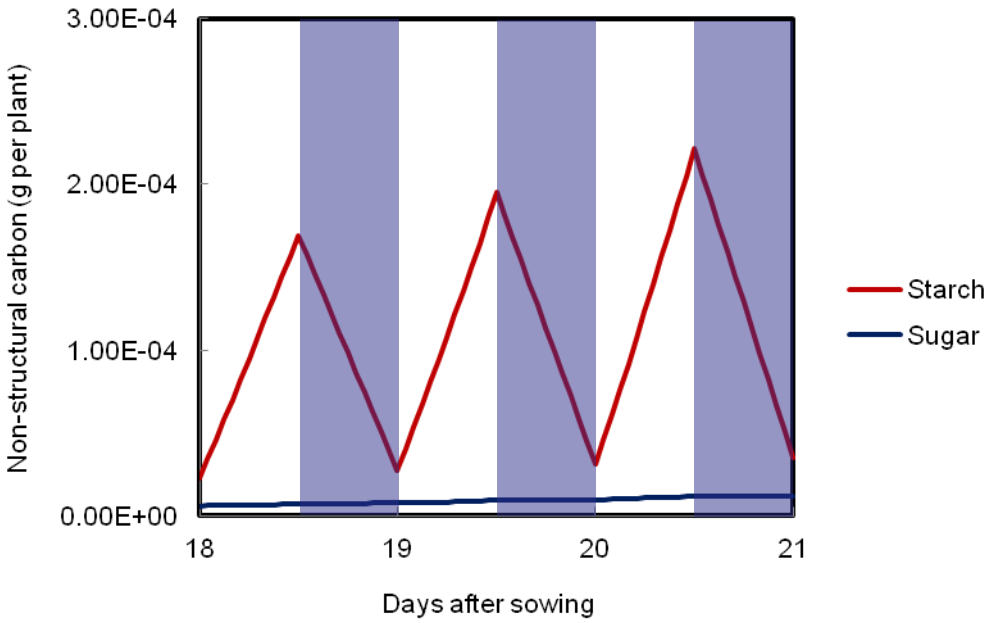


Figure 4.3.2: Simulated time series of non-structural carbon (NSC). Two pools of NSC are considered in the model, i.e. sugar and starch. Starch is accumulated at day time and broken down at a linear rate during the night, while sugar content is maintained relatively low and stable. The time series shown was simulated using a 12:12 light/dark cycle input.

4.3.2 Functional-structural plant growth model

The GreenLab functional-structural plant growth model for Arabidopsis (Christophe et al., 2008) was utilised here (Fig. 4.3.3). The general GreenLab model adopts the concept of the elementary growth unit known as the phytomer or metamer. For an Arabidopsis rosette, each phytomer consists of a leaf with negligible internode (stem) elongation. Therefore, the GreenLab model was adapted to the simpler non-branching structure of Arabidopsis rosette, and physiological age was redefined as the time since organ emergence to reflect the physiological transition in leaves from sink to source as they mature. Plant age indicates the time since sowing (t) while the physiological age of each leaf (phytomer) is the thermal time elapsed since its appearance (n). A growth cycle (numbered in j) is the period between the appearances of two successive phytomers (phyllochron).

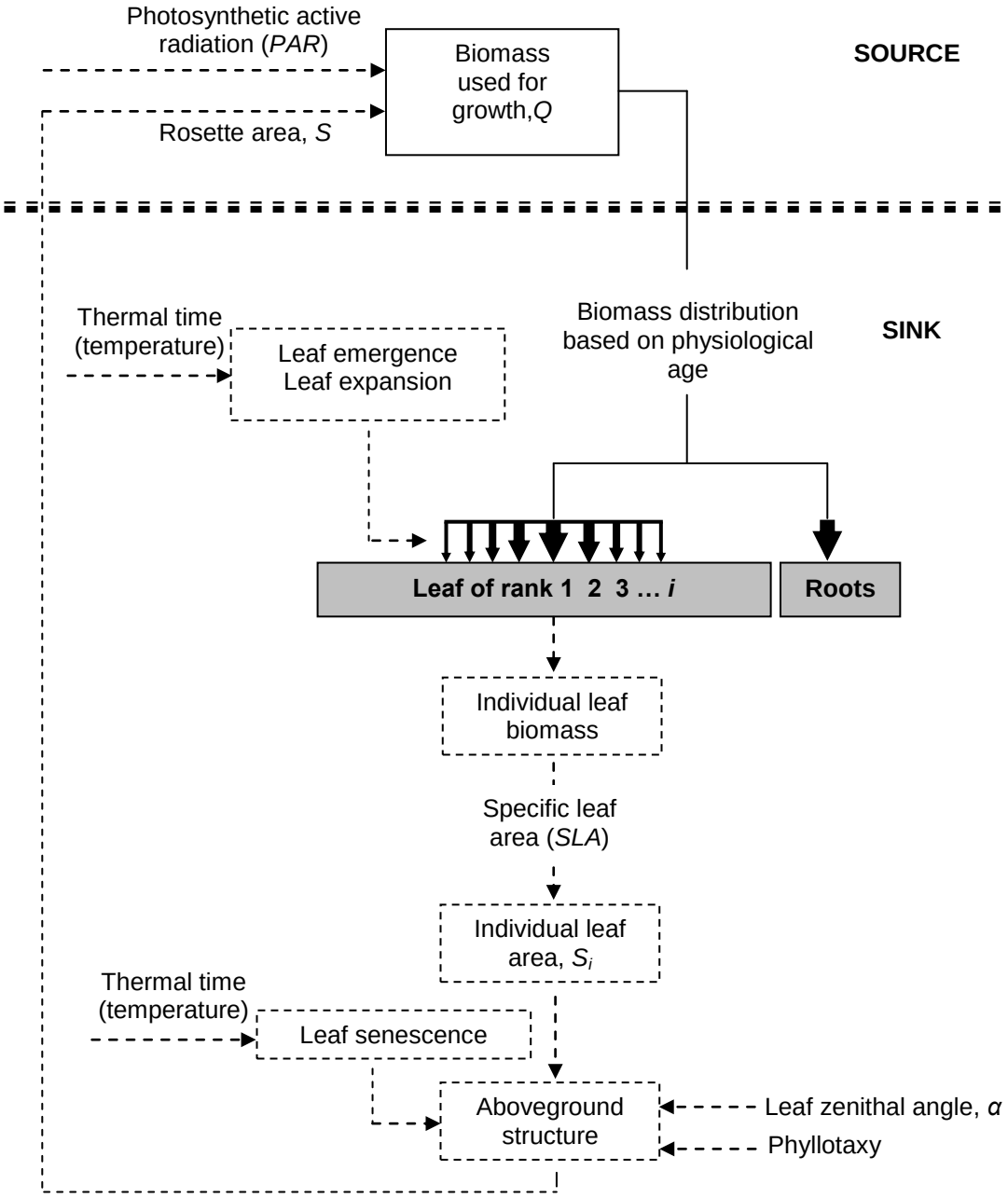


Figure 4.3.3: Schematic diagram of the Arabidopsis functional-structural plant model (Christophe et al., 2008). Biomass *Q* is distributed (solid arrows) to different organs (grey boxes) based on their physiological age. The above-ground structure, which determines the area for light interception (*S*), is also considered in the model. Dashed arrows indicate information or feedback input.

The first growth cycle starts from the time of sowing until plant emergence, which is the stage when the cotyledons (the first phytomer) are fully opened (Boyes et al., 2001). Upon plant emergence, the juvenile stage follows, where one phytomer appears approximately every 30.3 degree days (^{0}Cd) (Christophe et al., 2008). This is followed by the adult stage with one phytomer appearing every 11.9^{0}Cd. Therefore, the number of phytomers (y) that exist at each thermal time since plant emergence (TT) is given by:

$$y(j) = \begin{cases} 0.033TT(j) + 1, & TT(j) \le 355^0\text{Cd}; \\ 0.084TT(j) - 16.75, & \text{otherwise.} \end{cases} \quad (4.3.1)$$

As the first phytomer consists of two leaves, the number of existing leaves at any time point, L, is therefore:

$$L(j) = y(j) + 1. \quad (4.3.2)$$

Besides organogenesis, growth processes are also computed in each growth cycle and they can be generally divided into two groups: the source and the sink (Fig. 4.3.3). For the first growth cycle, seed biomass is assumed to be the only source (Christophe et al., 2008).

From the second cycle onwards, biomass production Q (source) is determined by a simple function of radiation use efficiency (RUE), rosette area at the end of the previous cycle for light interception (S) and the photosynthetic active radiation (PAR) as in the following:

$$Q(j) = RUE(j) \times S(j-1) \times PAR(j). \quad (4.3.3)$$

Only the vertically projected rosette area is considered in the model, therefore S depends on the phyllotaxy and zenithal angle (from the horizontal surface) of each leaf. In Arabidopsis, both cotyledons appear opposite each other during plant emergence. The third leaf is initiated at an angle close to 90^0 from one of the

cotyledons and the fourth leaf is almost 180^0 from the third leaf. For the fifth leaf onwards, leaves appear with a spiral phyllotaxy (Medford, Behringer, Callos, & Feldmann, 1992) of angle varying from 137.5^0 (Chenu et al., 2005) to 138.2^0 (Mundermann, Erasmus, Lane, Coen, & Prusinkiewicz, 2005). This range of angle is approximated by a phyllotaxy of 5/13 (138.46^0), where after 5 rotations or 13 leaves, the fourteenth leaf appears on the same orientation as the first leaf. Therefore, when the rosette leaf number exceeds 15 (as the first three leaves do not follow the spiral phyllotaxy), only the 13 largest functional leaves are considered for light interception due to self-shading:

$$S(j-1) = \underset{k\in\{1,\ldots,L_j\}}{\mathrm{Max}}\left(\sum_{i=k}^{k+12} S_i(j-1)\times\cos\alpha_i\right), \tag{4.3.4}$$

where S_i is the area of leaf of rank i. The zenithal angle of each leaf, α_i, changes from 70^0 at leaf emergence to 10^0 at the end of leaf expansion (Chenu et al., 2005), which is calculated using:

$$\alpha_i = \begin{cases} 10, & 1 \le i \le i_{\max}; \\ 10 + 60\cdot\dfrac{i - i_{\max}}{i_{\text{current}} - i_{\max}}, & i_{\max} < i \le i_{\text{current}}, \end{cases} \tag{4.3.5}$$

where $i_{\max}$ is the rank of the largest leaf while i_{current} is the newest leaf in the growth cycle.

The biomass produced in each cycle j is next distributed to two types of sink organ, i.e. leaves (l) and roots (r). Each leaf is considered as an individual organ while the roots are regarded as a single entity. In GreenLab, the incremental growth of each organ depends on the trophic competition among all the organs existing at each time point. This is decided based on the demand of each organ (d_{organ}), which is a function of sink strength (P_{organ}) and sink variation (f_{organ}) as follows:

$$d_{organ}(n) = P_{organ}\times f_{organ}(n). \tag{4.3.6}$$

Sink strength is an organ-type specific parameter while sink variation is a beta-law (Yin, Goudriaan, Lantinga, Vos, & Spiertz, 2003) or bell-shaped curve that depends on the thermal time since organ emergence (or physiological age, *n*). This curve is normalised to its maximal value *M* as shown in the following:

$$f_{organ}(n) = \frac{1}{M}\left(\frac{n+0.5}{T_{organ}}\right)^{a_{organ}-1}\left(1-\frac{n+0.5}{T_{organ}}\right)^{b_{organ}-1}, \quad (4.3.7)$$

where a_{organ} and b_{organ} are parameters for the organ-specific beta law and T_{organ} is the duration of organ expansion. The duration of leaf expansion was set to a constant value while the duration of root system expansion was set to 1.3 fold the flowering time, based on experimental observations (Christophe et al., 2008).

The total sink demand (*D*) of each growth cycle is the sum of all the existing individual leaf (as illustrated in Fig. 4.3.4) and root demands as follows:

$$D(j) = \sum_{i=1}^{L_j} d_i(n) + d_r(n). \quad (4.3.8)$$

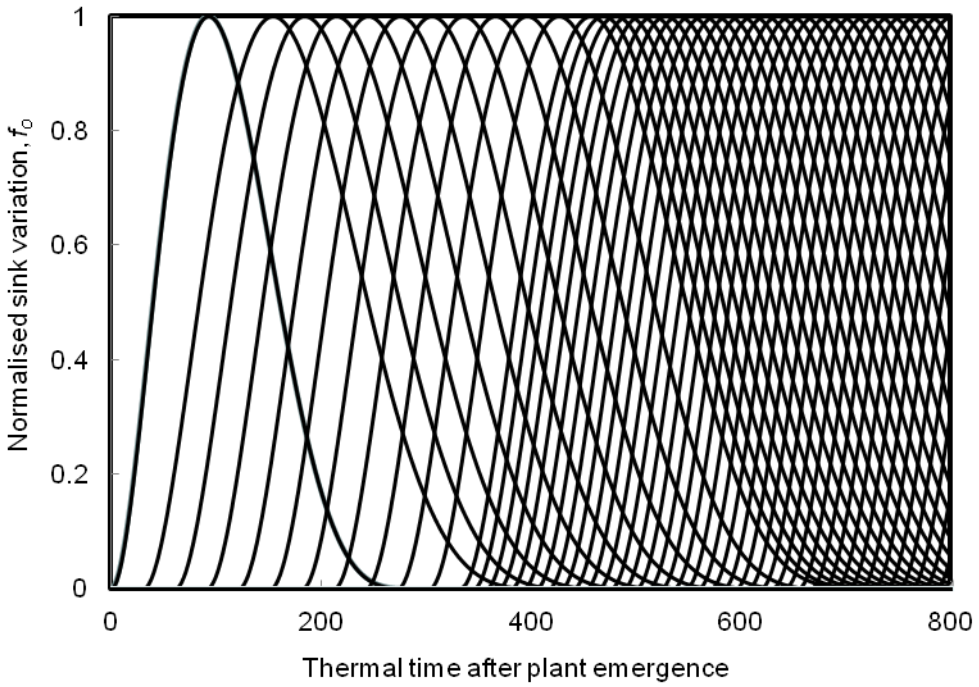


Figure 4.3.4: The sink variation of existing phytomer(s) at each thermal time after plant emergence. The first curve at the left represents the sink variation of the cotyledons. The curves are stacked closer to each other at later stage as phytomers appear at a higher rate during the adult phase.

The increase in biomass for each organ is therefore:

$$\Delta q_{organ}(j,n) = \frac{d_{organ}(n)}{D(j)} \times Q(j)\,. \tag{4.3.9}$$

The dry weight of each organ (m_{organ}) at the end of every cycle can then be calculated by:

$$m_{organ}(j) = m_{organ}(j-1) + \Delta q_{organ}(j)\,. \tag{4.3.10}$$

Next, the dry weight of each leaf (in g) is converted to leaf area (in m^2) for use in the next cycle using the specific leaf area (*SLA* in $m^2\ g^{-1}$) that changes with thermal time since plant emergence, *TT* (Christophe et al., 2008):

$$SLA(j) = 0.144 \exp(-0.002 TT(j)). \quad (4.3.11)$$

In the model, leaves are divided into three functional groups: (i) young leaves that function as both source and sink (from leaf emergence until the end of leaf expansion; (ii) mature leaves that function only as source (from the end of expansion till the end of its lifespan, T_s); (iii) senesced leaves that do not have any function. Only leaves in the first two groups are considered for light interception. All parameter values are listed in Table 4.3.1.

Table 4.3.1: Parameter values for the functional-structural plant growth model

Parameter description	**Symbol**	**Value and unit**
Measured parameters in Christophe et al. (2008):		
Seed input	$Q(1)$	1.6 x 10^{-5} g
Duration of leaf expansion	T_l	300 ^{0}Cd (cotyledons)
		400 ^{0}Cd (true leaves)
Leaf lifespan after full expansion	T_s	160 ^{0}Cd
Optimised parameters in Christophe et al. (2008):		
Sink strengths:		
Leaf	P_l	1
Root system	P_r	2.64
Sink variation parameters:		
Leaf	a_l	3.07
	b_l	5.59
Root	a_r	13.03
	b_r	9.58

The GreenLab growth model runs in steps of growth cycles. As this model is not linked to any phenology and assimilation models, input data from experimental measurement such as emergence and flowering time are required, while the *RUE* of each growth cycle has to be determined through model optimisation as demonstrated in Christophe et al. (2008). In the current study, this model was linked to a phenology component (Part I) where flowering time could be simulated based on environmental conditions. Biomass for growth (*Q*) would be computed using the photosynthesis and metabolic model (Section 4.3.1), as described next.

4.3.3 Issues in linking models of different scales

Two main issues required considerations in linking models of different scales in this study. They were:

(i) Time step

The phenology model runs on an hourly basis while the clock model has a small time step within seconds that is changed constantly by the ODE solver depending on the model dynamics. For the process-based metabolic model, a short time step of 6s was utilised in the original work (Rasse & Tocquin, 2006) whereas the growth model adopted a thermal-time step of one growth cycle (Christophe et al., 2008);

(ii) Biomass unit

The metabolic model only considers carbon mass while the functional-structural model was optimised using dry biomass, not all of which is carbon.

To standardise the time unit for all the modules, an hourly step as used in the phenology model was adopted except for the clock module which maintained its

ODE time step. In the previous sub-chapter (Part II), it has been presented how the clock model was integrated into the phenology model. Daily sunrise and sunset times were sent to the clock model for the computation of the area under the curve over 24 hours of *FT* mRNA simulated by equation B2.15. This daily information was then calibrated to a value in the phenology photoperiod component (equation 4.2.1), which was shared and used on an hourly basis in equation 4.1.7 until a new set of sunrise/sunset times was available the following day. Both the metabolic and growth models were also modified to run every one hour. Therefore, equations 4.3.1 to 4.3.11 were re-annotated by replacing growth cycle *j* with plant age *t* in hour. The thermal time unit (in degree days) for every hour can thus be calculated as:

$$Thermaltime(t) = \frac{T(t) - T_b}{24}, \tag{4.3.12}$$

where the numerator is the temperature above a base value of T_b = 3 ^{0}C (as used previously in equation 4.1.2, Part I: Phenology Model).

The different biomass units used in the metabolic and growth models were maintained to ensure compatibility with measurement and optimisation carried out in previous studies (Christophe et al., 2008; Rasse & Tocquin, 2006). However, unit conversion factors based on carbon content were introduced in the current work so that biomass information could be sent backwards and forwards between the two models. For simplicity, an average value of 0.3398 g carbon per g leaf dry mass was used based on the reported total leaf carbon content per dry weight minus sugar and starch contents (Gorsuch, Pandey, & Atkin, 2010a, 2010b). Root carbon content was fixed at 0.35 g per g dry weight (Kumar, Udawatta, & Anderson, 2010; Prendergast-Miller & Sohi, 2010).

4.3.4 Model initialisation

The GreenLab growth model requires as input data the thermal time after plant emergence, and any events or processes before this critical point are not considered. In the current study, the growth model remains inactive until plant emergence, TT_0. It is marked as the time point corresponding to 110 ^{0}Cd, which was the accumulated thermal time when the cotyledons were fully opened (Stage 1) in Boyes et al. (2001). The thermal time post-plant emergence, *TT*, is therefore given by:

$$TT(t) = Thermaltime(t) - TT_0 . \tag{4.3.13}$$

As soon as TT_0 is reached, both the metabolic and growth models are triggered. Initial plant biomass is assumed to be the same as seed input (Table 4.3.1) (Christophe et al., 2008), which is distributed to the sink organs, i.e. cotyledons, roots and hypocotyl, based on a constant fraction. This fraction is determined by the sink demand function (equations 4.3.8 and 4.3.9), assuming that the period of root expansion begins on the day of sowing and the above-ground sink strength is similar to that of rosette leaves ($P_{above\text{-}ground}$ = 1) and at its maximum ($f_{above\text{-}ground}$ = 1). Above-ground biomass is shared between the cotyledons and hypocotyl by setting the initial cotyledon area as 1 mm^2 and the specific cotyledon area as the *SLA* at plant emergence. These give:

$$\text{Hypocotyl biomass} = \text{Seed biomass} - \text{Root biomass} - 2 \times \text{Cotyledon area} \times SLA_{TT=0}. \tag{4.3.14}$$

The biomass of hypocotyl is set as a constant thereafter and the secondary growth of stem is assumed to be negligible. Therefore, any subsequent allocation to the shoot is distributed to leaves only.

Using the calculated biomass at emergence, the initial shoot and root carbon contents can next be generated with the biomass unit conversion factors. Initial sugar carbon and starch carbon for use in the metabolic model are also determined as:

Sugar_carbon(emergence) = 2 x Cotyledon area x Sugar content per unit area.

(4.3.15)

Starch_carbon(emergence) = Initial starch:sugar ratio x Sugar_carbon(emergence).

(4.3.16)

These initial ratio and specific sugar content are the same as those used during the initialisation of the metabolic model in Rasse and Tocquin (2006).

4.3.5 Connecting growth and metabolic models

In order to connect the process-based metabolic model (Section 4.3.1) and the functional-structural plant growth module (Section 4.3.2), a few connection points had to be identified and redundant equations were replaced or abolished. One redundant equation is 4.3.3 from the growth model (Section 4.3.2) which determines the biomass production *Q* (source). This component was replaced with the carbon assimilation and metabolic model in the current study (Fig. 4.3.5). Both models are executed simultaneously with information passed back and forth in between.

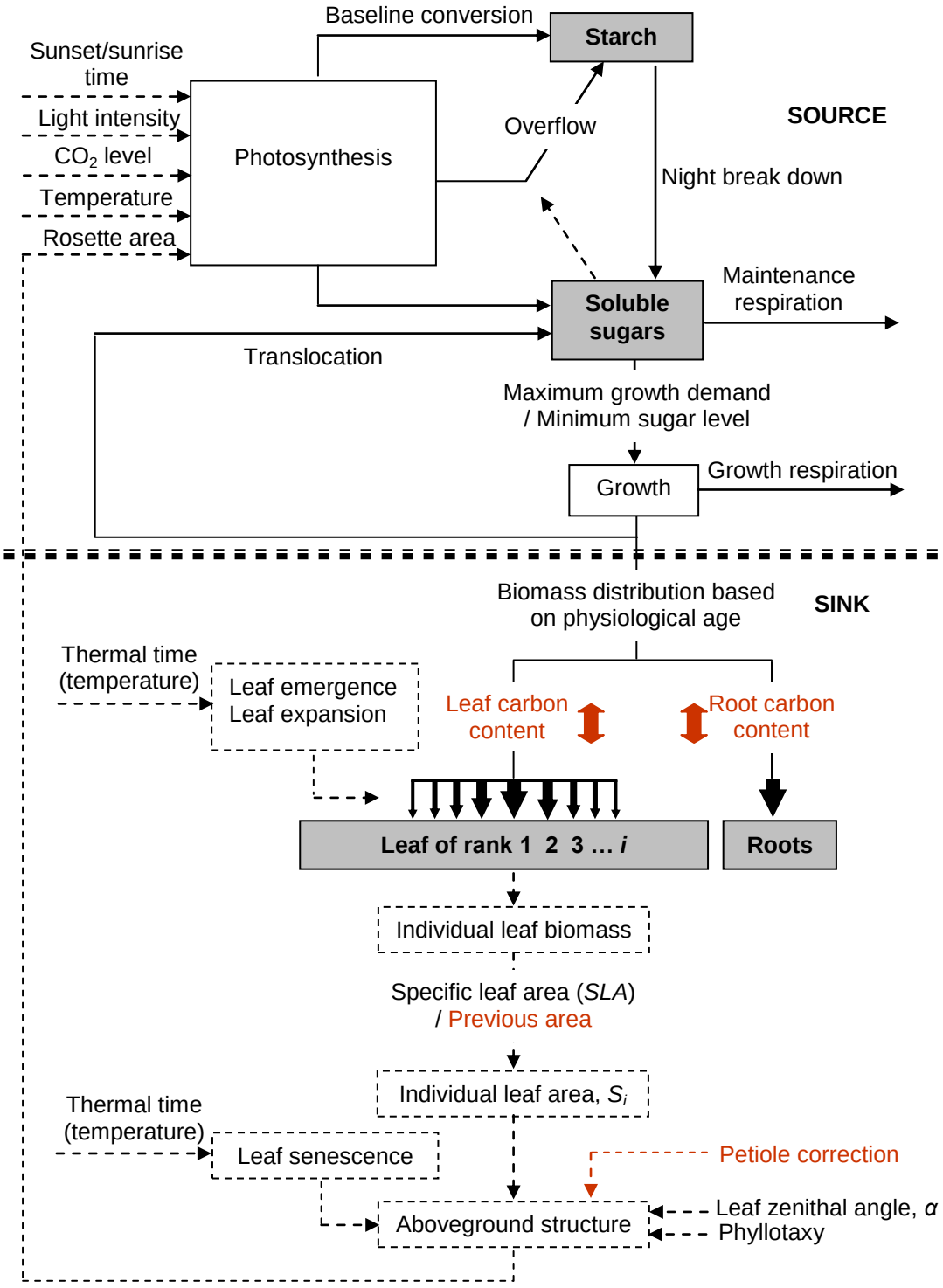


Figure 4.3.5: Schematic diagram illustrating the combined growth and metabolic model. It is a combination of Figs. 4.3.1 and 4.3.3. Texts in red highlight the new features that were introduced in the current work. The red double-sided arrows indicate the connection points where biomass information between the two models is sent back and forth.

To begin simulation, the number of leaves (equations 4.3.1 and 4.3.2) are first calculated so the dry biomass sink demand for each of the organs can be computed (equation 4.3.6) and converted to carbon sink demand using the unit conversion factors. These values are then used to determine the carbon root-to-shoot allocation ratio (*RS*) by summing all the leaf demands as follows:

$$RS(t) = \frac{d_r(n)}{\sum_{i=1}^{L_t} d_i(n)} \times \frac{\text{Root carbon content}}{\text{Leaf carbon content}}, \tag{4.3.17}$$

where t is the plant age (or hours after sowing), n is the physiological age of each organ (thermal time after organ emergence) and L_t is the number of existing leaves at time t. The area for light interception (equations 4.3.4 and 4.3.5) is also calculated but a simple *crowding* correction was introduced in the current study to account for petiole elongation which reduced self-shading among the leaves when they exceeded 15 in number. Equation 4.3.4 therefore was replaced by:

$$S(t-1) = crowding(t) \cdot \max_{k \in \{1,\ldots,L_j\}} \left(\sum_{i=k}^{k+12} S_i(t-1) \times \cos \alpha_i \right). \tag{4.3.18}$$

$$crowding(t) = \begin{cases} 1, & L_t \le 15; \\ 1 + \dfrac{pet - 1}{t_{end} - t_{L=15}} \left(t - t_{L=15}\right), & \text{otherwise}, \end{cases} \tag{4.3.19}$$

where *pet* is the maximum petiole correction factor, which would be determined from experimental data (see Table 4.3.4).

Information on rosette area is passed to the carbon assimilation model (together with environmental input such as light intensity, temperature and CO_2 level) for determination of photosynthesis, respiration and carbon flow. The growth pool in this model (Fig. 4.3.1) is then allocated to the roots and shoots using the *RS* ratio in equation 4.3.17, which replaced the *RS* ratio (as a function of rosette area, equation C1.29) that was originally used (Rasse & Tocquin, 2006). Once the carbon has been

allocated, contents in the roots and shoots are re-converted to dry mass unit. Distribution of the incremental shoot dry mass to different leaves is later determined using a modified version of equation 4.3.9 that was adjusted to represent trophic competition among the leaves, as shown in the following:

$$\Delta q_i(t,n) = \frac{d_i(n)}{\sum_{i=1}^{L_t} d_i(n)} \times \text{Total increment in shoot dry mass}(t). \quad (4.3.20)$$

The new dry biomass for each leaf can thus be calculated using equation 4.3.10 and the individual leaf area determined from the specific leaf area, *SLA* (equation 4.3.11). As *SLA* decreases with thermal time, a constraint is applied to ensure that leaf area does not shrink in later time points, as follows:

$$S_i(t) = \text{Max}\{SLA(t) \times m_i(t), S_i(t-1)\}. \quad (4.3.21)$$

All the parameter values required in the combined model are listed in Table 4.3.2.

Table 4.3.2: Parameter values for the combined plant growth and metabolic model.

Parameter description	**Symbol**	**Value and unit**
Thermal time at plant emergence	TT_0	110 ^{0}Cd
Initial area of cotyledons	S_1(emergence) S_2(emergence)	1 mm^2 1 mm^2
Sugar content per unit area		0.1 g sugar-C m^{-2}
Initial starch:sugar ratio		2
Leaf carbon content		0.3398 g C per g leaf dry mass
Root carbon content		0.35 g C per g root dry mass

4.3.6 Integrated multi-scale model

The phenology model (Part I and Part II) only predicts the flowering time of Arabidopsis without giving any information on plant size and biomass. On the other hand, the combined growth-and-metabolic model in Section 4.3.5 has unlimited run-time as long as input data are provided, unless execution is terminated. In other words, the simulation process could run incessantly, though simulated rosette growth may eventually end when the carbon budget runs out due to self-shading, leaf senescence, and/or suboptimal growth conditions. The current study therefore connected both models together, with the phenology component providing the time of flowering to the growth-and-metabolic component so that rosette growth can be terminated accordingly. Four input data, i.e. temperature, light intensity, CO_2 level and sunrise/sunset time, are necessary to run the integrated model. The former three input data are read hourly, while sunrise/sunset times are read by the clock model once at the start of each calendar day (Fig. 4.3.6). The phenology component starts execution instantaneously without any conditions, whereas the growth-and-metabolic compartment has a conditional feature. The latter is only activated when the accumulated thermal time exceeds the threshold for seedling emergence. Once activated, this compartment begins its implementation using initial values from model initialisation. The circuit runs continuously until flowering time, as informed by the phenology component, is reached. At this point, the whole integrated model stops execution. Outputs from the model include time series data of sugar and starch contents, and the biomass and size of the whole plant and individual organs.

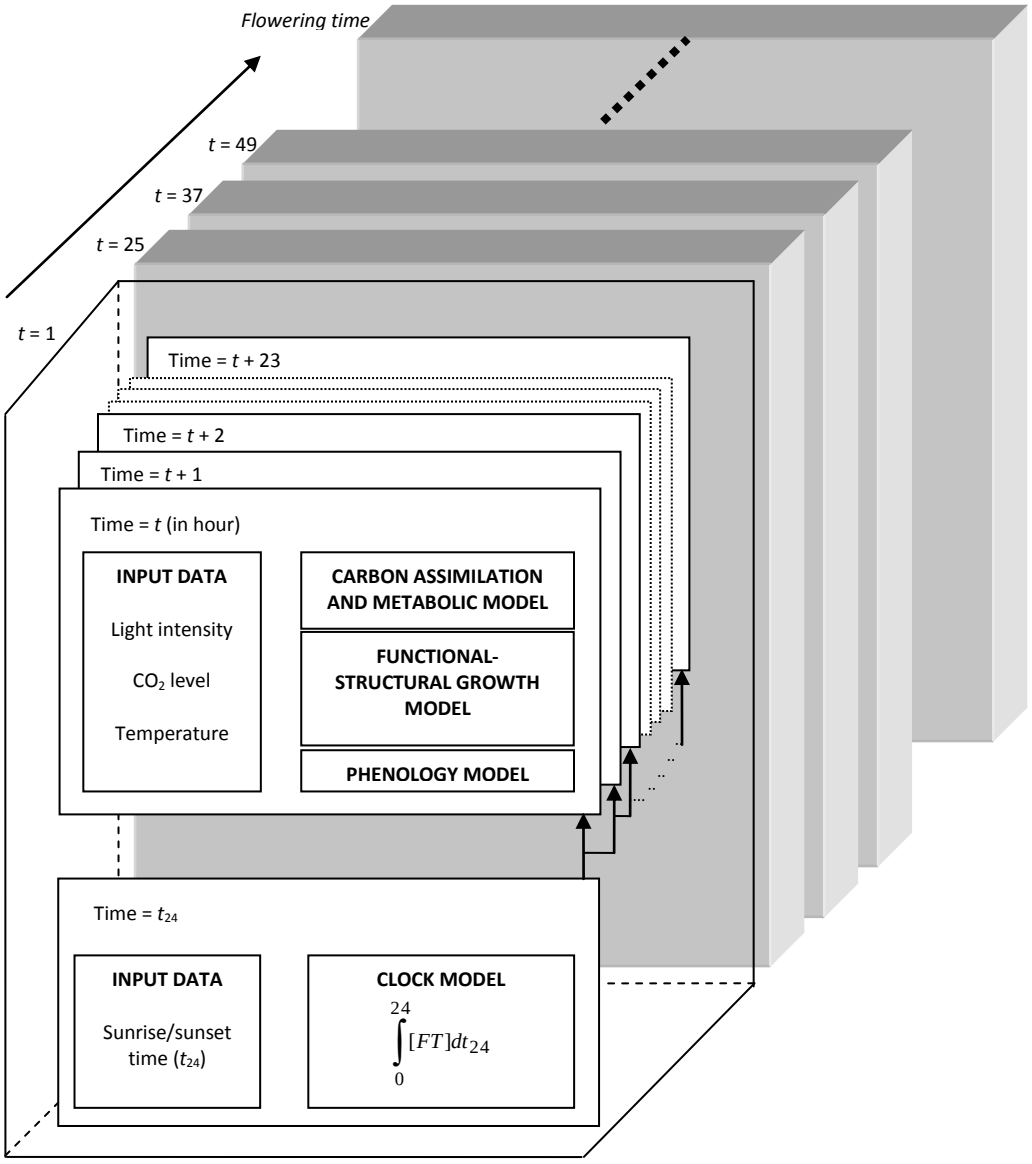


Figure 4.3.6: Time-stepping scheme of the integrated multi-scale model. Integral of *FT* mRNA is simulated by the clock model using input data of sunrise and sunset times (in the 24-hour format), which are the same for the same calendar day. Output from the clock model is therefore used repeatedly for each one-hour time step at which the other models are running, until the next calendar day. This is repeated until flowering time.

4.3.7 Model validation

The original growth (Christophe et al., 2008), assimilation and resource partitioning (Rasse & Tocquin, 2006) models were developed and optimised using experimental data of the Col ecotype. Therefore, the model was first validated using Col data from an experiment conducted earlier in the lab (B. Wenden and C. I. Garcia, unpublished data).

Two modifications were made in the model for the Col data (Table 4.3.4). As the plants bolted slightly later than the flowering time predicted by the phenology model, flowering threshold was adjusted to extend the simulation of the plants' exponential growth until the last biomass assay. Another modification was the J_{max}:V_{cmax} ratio (or R_{JV} in equation C1.8). The reported values of this ratio in the literatures suggested a photoperiod-dependent tendency (Table 4.3.3), with an increasing ratio as photoperiod shortened. The ratio used in Rasse and Tocquin (2006) was 2.1, where the model was tested on plants grown under an 8-hour photoperiod. As the Col data were collected from plants grown under a 12-hour photoperiod, the current study adopted the J_{max}:V_{cmax} ratio measured under the same photoperiod, which corresponded to 1.7 (Flexas et al., 2007). The value for *pet* was calculated from the last biomass assay by taking the ratio between measured rosette area and the total measured area of the 13 largest leaves corrected with zenithal angle (equation 4.3.4).

Table 4.3.3: The measured ratio of J_{max}:V_{cmax} at different photoperiods from the literatures

Photoperiod (Light:Dark)	$\boldsymbol{J_{max}:V_{cmax}}$	**Source**
8:16	2.04 – 2.26	(Pons, 2012)
12:12	1.7	(Flexas et al., 2007)
14:10	1.36	(Bunce, 2008)

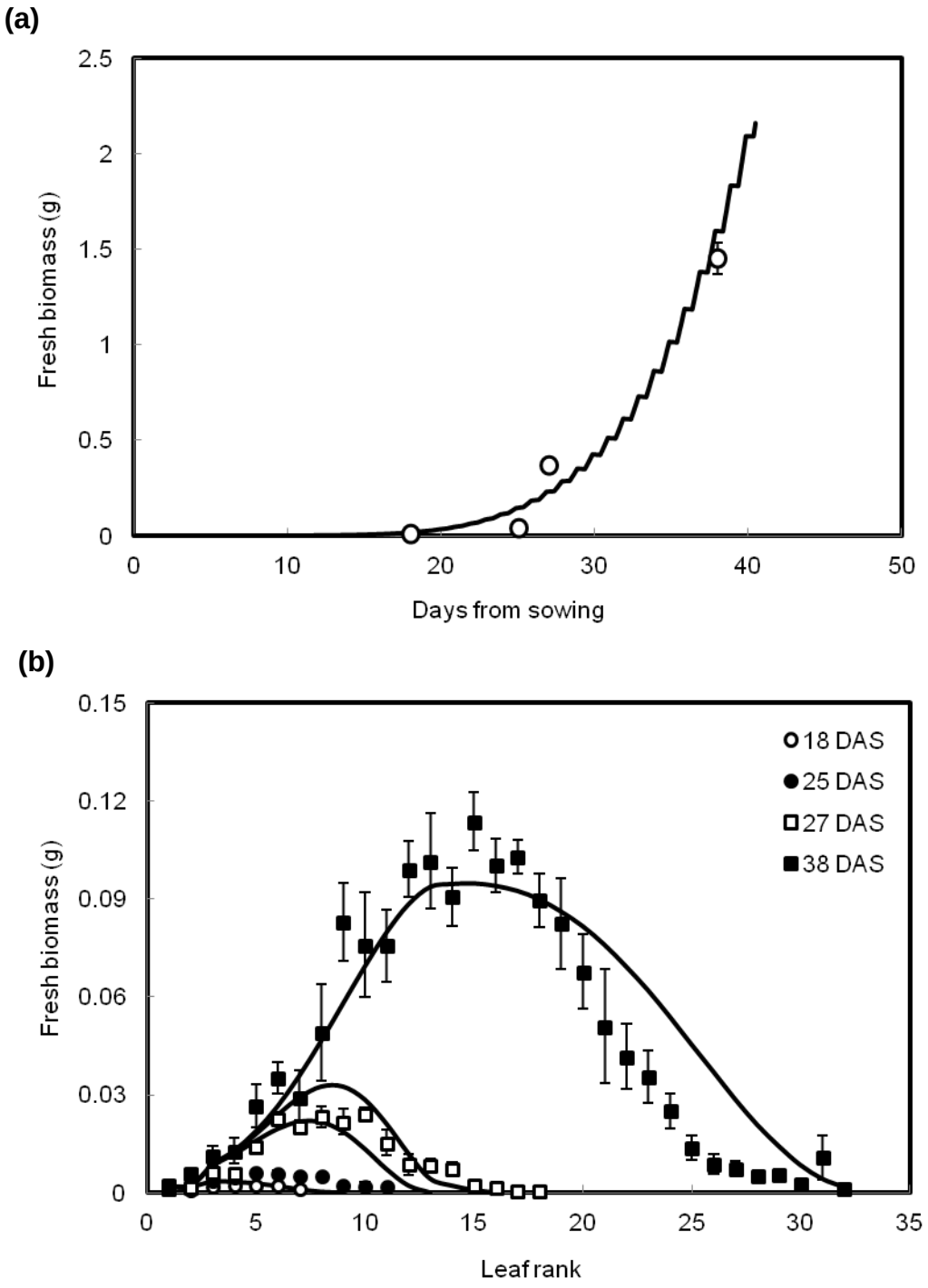


Figure 4.3.7: Model simulation (solid line) and experimental data (symbol) of Col: a) above-ground biomass and; (b) individual leaf biomass at different time points. Leaves were ranked according to the order of appearance. Growth conditions: ~21.3 ^{0}C; 12:12 light/dark cycle; light intensity = 110 μmol m^{-2} s^{-1}; CO_2 level as measured in the growth room (Fig 3.1 in Chapter 3). Error bars shown were the standard errors of 5 plants.

In general, the integrated multi-scale model managed to capture the dynamics of rosette growth in Col (Fig. 4.3.7). A diurnal growth rhythm can be seen in the model simulation. The biomass of each individual leaf during the early and late stages was also well described by the model. There was however an overestimation of the biomass measured on 25 days after sowing (DAS) and for young leaves of high ranking on 38 DAS. Model performance, as measured by the normalised root mean squared error, is summarised in Table 4.3.5.

As the integrated growth component is a combination of modules independently parameterised with different experimental data (Christophe et al., 2008; Rasse & Tocquin, 2006), model consistency was also examined at the connection points where certain elements have been replaced. In this case, the root-to-shoot allocation ratio (*RS*) simulated by the integrated model (equation 4.3.17) was compared (Fig. 4.3.8) to the original ratio calibrated from experimental measurement in Rasse and Tocquin (2006) (equation C1.29). In the Rasse and Tocquin model, *RS* has a minimum threshold of 0.12 when the rosette area is 1.22 cm^2, which was the earliest available measurement to calibrate the model in that study. No comparison was therefore feasible before this stage. In contrast, the integrated model was initialised at plant emergence when the total cotyledon area was 0.02 cm^2 and the calculated *RS* was 2.85e-06. As leaf surface increased, the *RS* values simulated by both models also increased, with the new model producing a lower value until the area reached ~15 cm^2, then exceeded the calibrated ratio (by about 8.8 % at the maximum point) before decreasing to values below the extrapolated plateau. Altogether, the comparison between both models displayed some differences at different stages; however the new model could generally re-produce the calibrated ratio to a consistent magnitude, suggesting sufficient compatibility within the integrated modules.

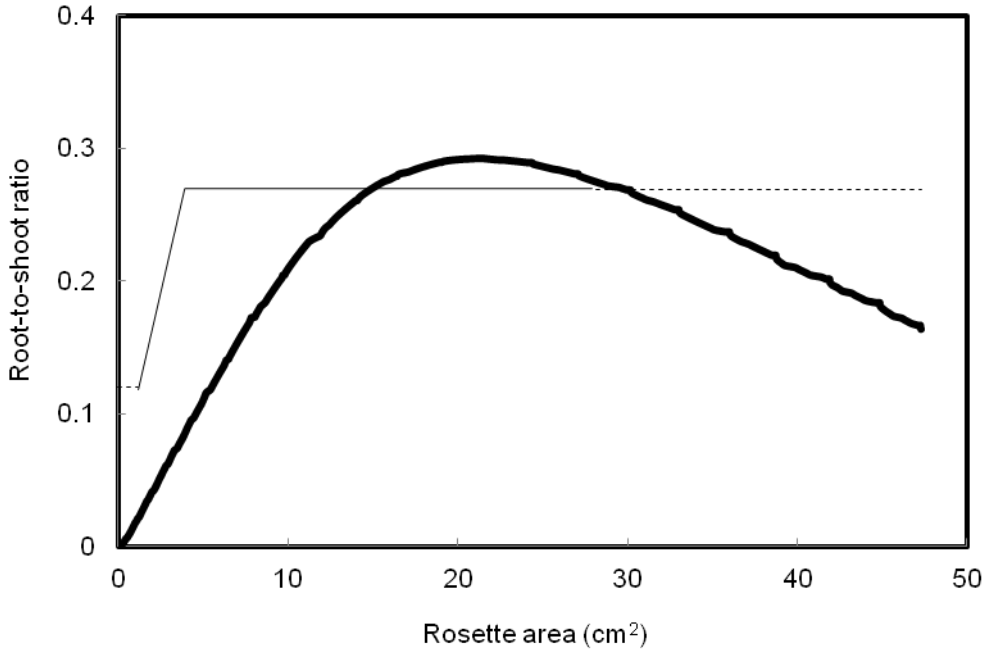


Figure 4.3.8: Comparison between Rasse and Tocquin (2006) (thin line) and the current study (thick line) for modelled root-to-shoot allocation ratio (*RS*) as a function of rosette area. In the Rasse and Tocquin model, the function was derived numerically by assuming a minimum threshold and a maximum plateau, based on experimental measurement of plants from a size of 1.22 cm^2 up to ~28 cm^2. Ratios at rosette area outside the measured regions were extrapolated (dashed lines). For the integrated model, simulation began at plant emergence (0.02 cm^2) until flowering time.

4.3.8 Model applicability to Ler and Fei

One application of plant (crop) modelling is the prediction of growth and development of different plant varieties under an ever-changing environment. This could facilitate the selection for varieties with desired traits or breeding of super-hardy crops. The multi-scale model was therefore tested on its ability to predict the growth of Arabidopsis accessions with different traits by only modifying model parameters associated with the specific traits, without re-optimisation to fit experimental data. Two accessions were selected for this purpose, as follows:

1. Ler, which is another one of the most popular Arabidopsis accessions used in molecular and genetic studies (Anderson & Mulligan, 1992). Due to a mutation in *ERECTA* (*ER*), Ler plants have different leaf morphology, i.e. round leaves and short petioles (Bowman, 1993; Redei, 1992; Torii et al., 1996) (inset of Fig. 4.3.10a).

2. Fei, a relatively less studied ecotype, which has been reported to have a higher leaf appearance rate (Mendez-Vigo, de Andres, Ramiro, Martinez-Zapater, & Alonso-Blanco, 2010) (Fig. 4.3.9a).

Parameter modifications for the two accessions are summarised in Table 4.3.4. Similar adjustments as applied to Col, i.e. flowering threshold, petiole correction factor and J_{max}:V_{cmax} ratio, were also applied to Ler and Fei. In addition, parameters associated with leaf appearance were altered based on experimental data. Consistent with its characteristic as a fast-growing ecotype, Fei was found to reach plant emergence early. This was incorporated by modifying the thermal time at plant emergence to that observed for Fei, which was enough for the model to re-produce its leaf appearance rate (Fig. 4.3.9b). Ler on the other hand emerged at the expected time, but its leaf appearance rate at the adult stage was slightly lower than the rate derived from Col data that was used in the GreenLab model (Christophe et al., 2008). Therefore, the leaf appearance rate for Ler was modified in the model to reflect this observation (Table 4.3.4).

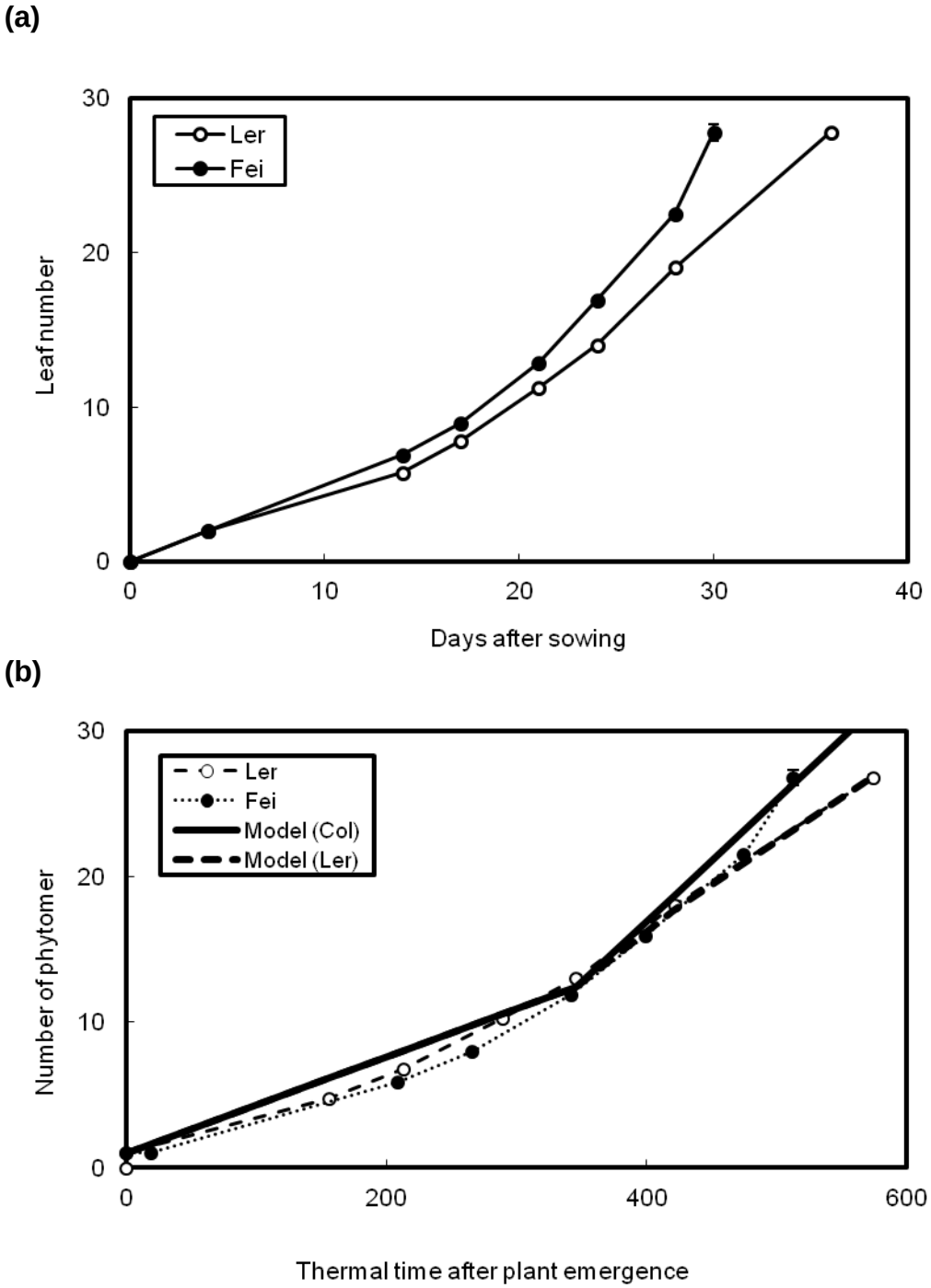


Figure 4.3.9: Vegetative shoot growth of Ler and Fei from germination until the appearance of flower buds. (a) Leaf number plotted against days after sowing; (b) Phytomer number plotted as a function of thermal time after plant emergence. Also shown are phytomer numbers as modelled using equation 4.3.1 derived from Col data in the GreenLab model (Christophe et al., 2008) and the modified rate for Ler (Table 4.3.4). Error bars showing the standard error of 24 plants were very small and therefore hidden by the data markers.

The multi-scale model could fit the biomass data of both Ler and Fei at different time points to a good accuracy (Figs. 4.3.10 and 4.3.11). For Ler, the model was able to predict all data of the total above-ground biomass except for a slight underestimation at 29 DAS. The model also displayed good fit of the individual leaf biomass particularly at the early stages. However, the biomass of early leaves (rank 5 – 10) was well matched at 23 DAS (Fig. 4.3.10b) but underestimated at 29 DAS, when the model predicted leaves 6 – 8 to be fully expanded. The measured biomass of these leaves showed that they continued to grow, as observed at 37 DAS (arrow in Fig. 4.3.10b). This suggested that the duration of leaf expansion used in the model was too low, causing the model to direct all carbon sources to newer sink organs which were still expanding. Simulation results for Fei were in general similar though biomass at the last two time points were overestimated (Fig. 4.3.11a), and these overestimations were mainly distributed among young leaves (Fig. 4.3.11b). Model goodness-of-fit for both accessions are shown in Table 4.3.5.

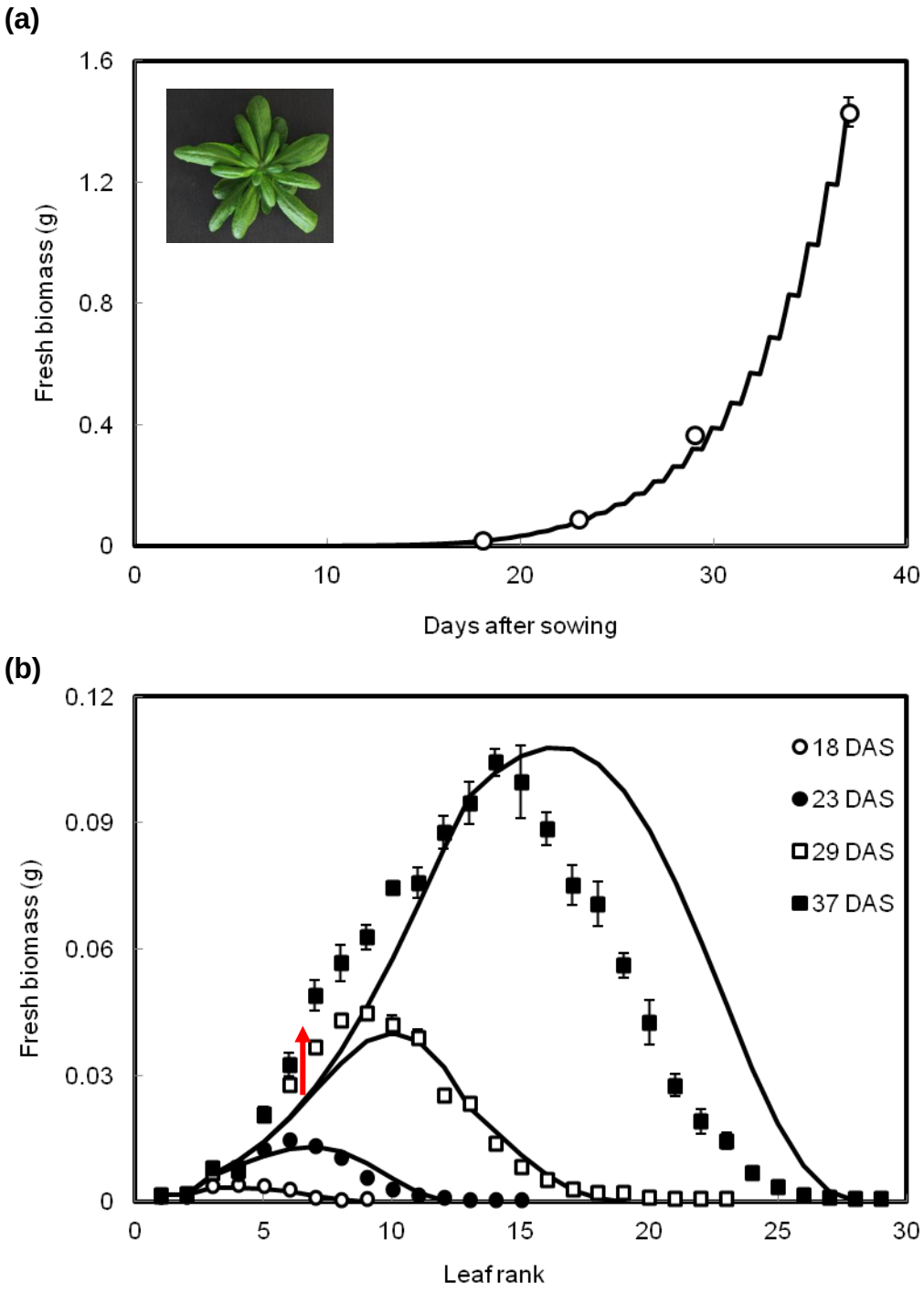


Figure 4.3.10: Model simulation (solid line) and experimental data (symbol) of above-ground (a) and individual leaf biomass (b) at different time points for Ler. Leaves were ranked according to the order of appearance. Inset: The rosette structure of Ler. Growth conditions: 22 ^{0}C; 12:12 light/dark cycle; light intensity = 130 µmol m^{-2} s^{-1}; CO_2 level as measured in the growth room (Fig 3.1 in Chapter 3). Error bars shown were the standard errors of: *n* = 10 plants for total above-ground biomass; *n* = 5 plants for individual leaf biomass. The arrow indicates that early leaves continued to grow, in contrast to model prediction.

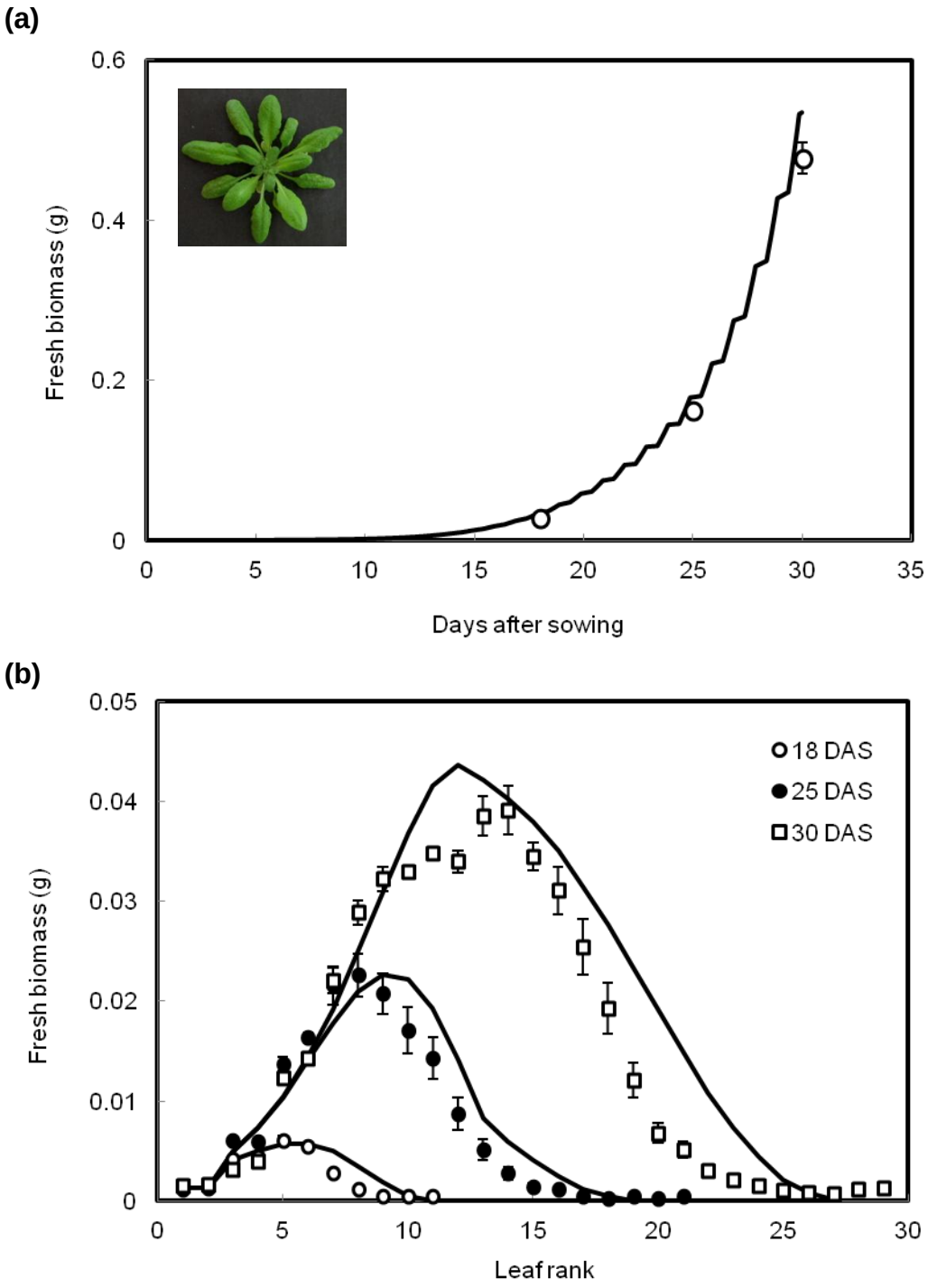


Figure 4.3.11: Model simulation (solid line) and experimental data (symbol) of above-ground (a) and individual leaf biomass (b) at different time points for Fei. Leaves were ranked according to the order of appearance. Inset: The rosette structure of Fei. Growth conditions: 22 ^{0}C; 12:12 light/dark cycle; light intensity = 130 μmol m^{-2} s^{-1}; CO_2 level as measured in the growth room (Fig 3.1 in Chapter 3). Error bars shown were the standard errors of: *n* = 10 plants for total above-ground biomass; *n* = 5 plants for individual leaf biomass.

Table 4.3.4: Modifications of parameters in the multi-scale model for different genotypes

Parameter description	Symbol and equation	Col	Ler	Fei	Source
Phenology model:					
Flowering threshold	T_h (4.1.8)	3955 [3212]	3560 [2907]	2886* [2907]	This study
Growth model:					
Thermal time at plant emergence	TT_0 (4.3.13)	[110]	[110]	57 [110]	This study
Leaf appearance rate in the adult phase ($TT(t) > 355\ ^0$Cd)	(4.3.1)	[$y(t) = 0.084TT(t) – 16.75$]	$y(t) = 0.06TT(t) – 7.6$	[$y(t) = 0.084TT(t) – 16.75$]	This study
Maximum petiole correction factor	*pet* (4.3.19)	1.1	1.06	1.1	This study
Assimilation model:					
Ratio of maximum electron transport to maximum carboxylation capacity	J_{max}:V_{cmax} or R_{JV} (C1.8)	1.7 [2.1]	1.7 [2.1]	1.7 [2.1]	(Flexas et al., 2007)

Notes: Numbers and equations in square brackets (where applicable) indicate the original parameter values before modifications. If no new values are shown, no modifications have been employed.
*This value was optimised by assuming that all other parameters in the phenology component adopted the values of Ler's

Table 4.3.5: Goodness of fit of the multi-scale model as indicated by root mean squared error (RMSE) normalised to the mean of observations, also known as the coefficient of variation of RMSE (CV(RMSE))

Model output	**Col**	**Ler**	**Fei**
Fresh biomass (FW):			
Total shoot	0.2394	0.0493	0.1552
Individual leaves	0.3925 (18 DAS) 3.0095 (25 DAS) 0.4907 (27 DAS) 0.3316 (38 DAS)	0.2486 (18 DAS) 0.6150 (23 DAS) 0.2812 (29 DAS) 0.5388 (37 DAS)	0.4184 (18 DAS) 0.3493 (25 DAS) 0.3475 (30 DAS)
Dry biomass (DW):			
Total shoot	0.1947	0.1338	0.1758
Individual leaves	0.4772 (18 DAS) 1.6707 (25 DAS) 0.3869 (27 DAS) 0.2742 (38 DAS)	0.5353 (18 DAS) 0.3808 (23 DAS) 0.4240 (29 DAS) 0.4615 (37 DAS)	0.7781 (18 DAS) 0.3612 (25 DAS) 0.2681 (30 DAS)
Leaf number	0.1482	0.0945	0.0620

Note: In the simulation of fresh and dry biomass, water content measured in the experiments was used in the model. The measured water content for each accession was as follows: Col (92.0 %); Ler (90.7 %); Fei (88.2 %).

4.3.9 Model scale-up: Net ecosystem exchange (NEE)

Net ecosystem exchange (NEE) is defined as the net CO_2 flux from the ecosystem to the atmosphere, and it is one of the main components of carbon (C) cycling that plays a central role in climate change studies (Chapin et al., 2006). In this section the integrated model, which describes the growth of a single plant, was scaled up to a small population to examine its applicability at the higher level of organisation. Assuming no interaction such as competition for light between each plant in the population, model output in terms of (negative) net carbon uptake was multiplied by the number of plants on a tray and compared to the NEE measured from the whole tray (Figs. 4.3.12 and 4.3.13). Two types of measured values were used for comparison: measured values corrected with the control, which comprised a tray of soil without plants; and non-corrected measured values. The control measurement was included to account for heterotrophic respiration or gas-exchange activities of other autotrophs on the soil surface. All the control measurements had negative NEE (Appendix C4), indicating a dominance of autotrophic activities on the soil from moss and photosynthetic microbes. For Ler, simulated NEE values mostly fell between the corrected and non-corrected measured values. Simulation results were closer to the measured values corrected with control when the plants were still small. As the plants grew larger, modelled NEE values shifted towards the non-corrected measurements and eventually deviated from both measured values. This trend was expected since the larger plants covered more soil surface and this suppressed other gas-exchange activities on the soil. The bigger rosettes also started shading one another (Fig. 4.3.12c) but this was not considered in the simulation, thus explaining the overestimation of NEE for older plants. For Fei, only three measurements were available as the plants flowered earlier. Simulation results for Fei did not follow the same trend as Ler. Modelled values overlapped with the non-corrected values, and there was a slight negative deviation at the last time point even though plant coverage of soil surface and inter-shading were relatively low compared to Ler (Fig. 4.3.13c). This pointed to a general overestimation of NEE by the model at all time points. An overestimation has also been observed earlier of biomass in Fei (Fig. 4.3.11a), offering a possible explanation to the higher NEE simulated by the model.

To test if the qualitative interdependence between biomass and NEE could account for the quantitative overestimation, the model was adjusted to reduce its biomass output. This was done by changing a parameter that only has an indirect effect on biomass but no direct connection to daytime NEE in the model. Starch turnover at night (ST_c in equation C1.15) was selected as an example for this purpose. Graded changes in ST_c and therefore biomass led to changes in simulated NEE (Fig. 4.3.14). In the worst case scenario (70% reduction in ST_c) where biomass was underestimated, there was also an underestimation in the NEE simulation, as indicated by the overlapping of modelled and corrected values at the first time point even though the plants already covered quite a large area of soil surface that should have led to less negative control (Fig. 4.3.13b). On the other hand, the simulation that corrected the biomass fit, i.e. 85% reduction in ST_c, translated into NEE values between the corrected and non-corrected measurements, which were the expected trend. However, as the biomass and NEE are interdependent, increasing or decreasing one naturally affects the other and *vice versa*. Reducing photosynthesis itself (thus NEE) could also lead to a decrease in biomass (Appendix C5). This suggested that there could be various ways of improving data fit for Fei, but the relation between variables in the integrated model is still realistic despite all its multi-scale components.

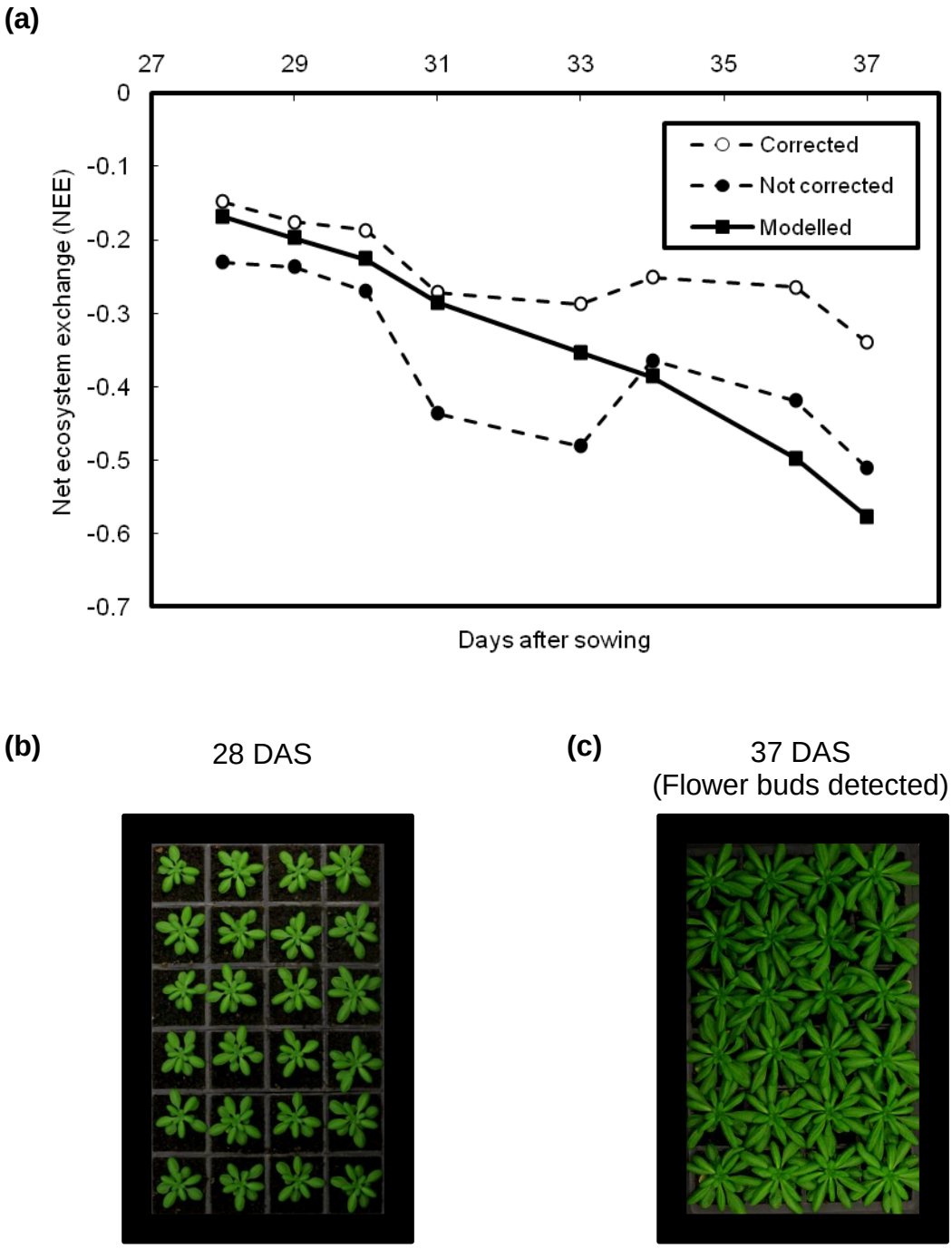


Figure 4.3.12: Modelled and measured net ecosystem exchange (NEE) for a small population of Ler (a). Measurements were either corrected (by subtraction) or not corrected with the control, which was a tray of soil without plants (see Appendix C3 for calculation details). Images of plant population at two different time points are also illustrated (b and c). **Note:** NEE has a negative sign here by convention as it is defined by atmospheric scientists as a C input to the atmosphere (Chapin et al., 2006).

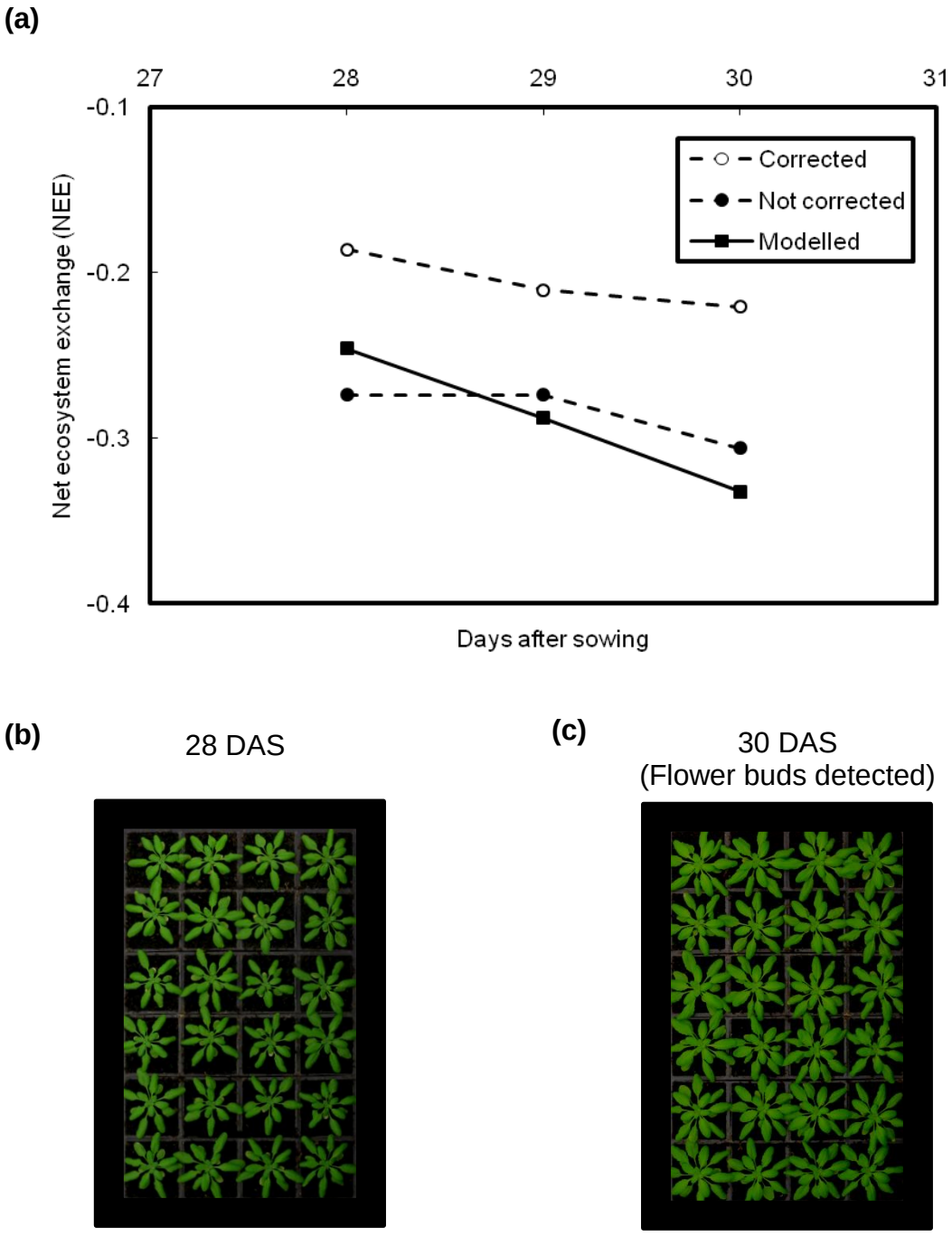


Figure 4.3.13: Modelled and measured net ecosystem exchange (NEE) for a small population of Fei (a). Images of plant population at two different time points are also illustrated (b and c).

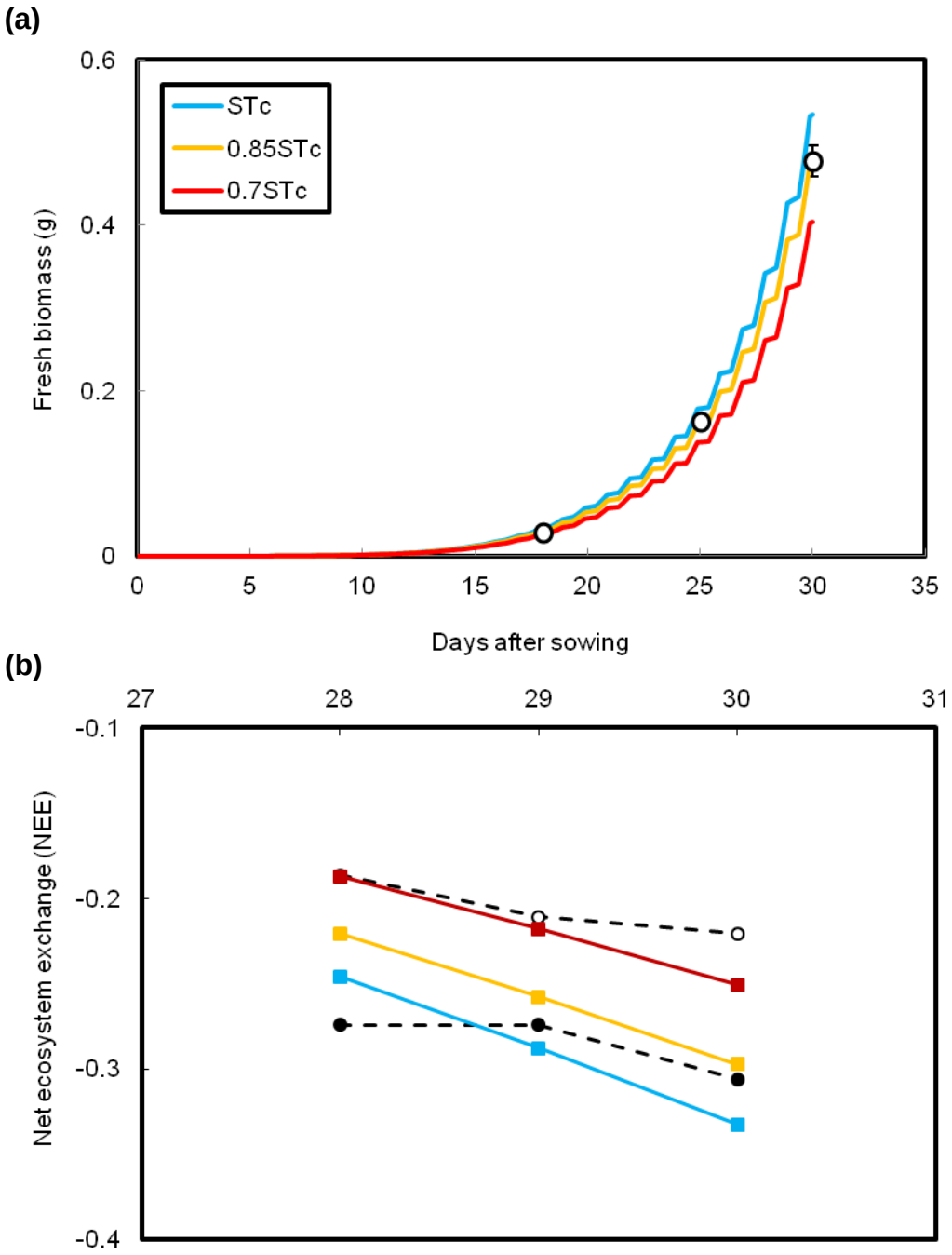


Figure 4.3.14: Interdependence between simulated biomass (a) and net ecosystem exchange (NEE) (b). ST_c was reduced progressively in each simulation run. The percentage of reduction is shown in the legend. Dashed lines indicate the corrected and non-corrected measured values as in Fig. 4.3.13a.

4.3.10 Discussion

One of the challenges in linking existing models is their compatibility upon integration. Each existing model is usually parameterised using different data sets, and there may be inconsistency among different scales and model types. The Arabidopsis multi-scale model developed in this study has demonstrated that combining plant models in a modular fashion is feasible given observed experimental variation across laboratories.

Validation with different accessions, i.e. Col, Ler and Fei, produced results with fairly good accuracy (Figs. 4.3.7, 4.3.10 and 4.3.11). There were some small discrepancies between modelled and measured biomass for Col in the early stage, which could be due to a different sowing method that affected early development. In that experiment, seeds were sown directly onto soil and thinned later, while in the current study seeds were first sown on agar and seedlings were later transplanted to soil (Chapter 3). Nevertheless, the coefficient of variation of RMSE (CV(RMSE)) for total Col shoot biomass was less than 25% (Table 4.3.5). The model also showed surprisingly good results when tested with Ler and Fei by adjusting only four genotype-specific parameters based on experimental data. Model performed the best in describing Ler data (CV(RMSE)) for total shoot biomass ~ 5%), followed by Fei (~15%). This demonstrated the robustness of the model framework. Inspection of the result (Fig. 4.3.10b) suggested that model performance might be further improved such as by modifying leaf expansion duration for Ler. However, it was the aim of the validation test to have minimal modification unless supported by experimental data. As leaf expansion duration was not measured specifically in the experiment, this parameter was not adjusted in the model.

The model also predicted realistic NEE values for both Ler and Fei, showing that it has potential for the application of scaling up to a small population. There was an overestimation of NEE for Fei which correlated with model overestimation of biomass. Reducing starch turnover (ST_c) or carbon assimilation rate improved model prediction of both variables (Fig. 4.3.14 and Appendix C5). Recently, a study has

found a duplication-and-loss event in the RuBisCO small subunit (RBCS) gene family in Fei (Schwarte & Tiedemann, 2011), though this does not seem to have affected its RuBisCO protein content (5.567 mg/g FW) relative to Col (5.426 mg/g FW) and Ler (5.114 mg/g FW) (Supplemental Dataset 1 in Sulpice et al. (2010)). Starch content at the end of the day was found to be slightly lower in Fei (35 μmol/g FW) compared to Col (40 μmol/g FW) and Ler (39 μmol/g FW) in that same dataset of 129 Arabidopsis accessions (mean = 40 μmol/g FW, standard error = 0.44 μmol/g FW) (Sulpice et al., 2010). However, attempt at reducing baseline starch synthesis in the model (equation C1.13) did not generate any change in the simulation as it was compensated by the overflow mechanism in the model (equation C1.37). Further information and detailed analyses would be required to determine how Fei compares to the others for additional model adjustment(s). Nevertheless, the results altogether still demonstrated that the integrated multi-scale model could simulate carbon cycling, usually described using simple big-leaf models in ecosystem studies, to a good accuracy.

Phenotyping assay of Fei to parameterise the model has generated interesting observations that are usually not analysed. Leaf production rate of Fei was found to be faster than Ler as has been reported earlier (Mendez-Vigo et al., 2010), but when plotted against thermal time after plant emergence, the difference was reduced and Fei could be described using the Col rate (Fig. 4.3.9). The timing of plant emergence is seldom recorded, but it could be an important dimension that affects observed leaf production rate and leaf number at flowering. Future studies may therefore benefit from comparative analyses of plant emergence across genotypes and treatments. In terms of flowering time, flower buds appeared slightly earlier in Fei than in Ler under 12:12 light/dark cycle in this study (Fig. 4.3.9). This contradicted published data where Fei was shown to flower later than Ler in both LD (16-hour photoperiod) and SD (8-hour photoperiod) conditions (Mendez-Vigo et al., 2010). One possibility was that Fei already achieved its maximum developmental rate at 12-hour photoperiod, as suggested by the comparable leaf number at flowering here (27.8) and that published (~31) under LD (Mendez-Vigo et al., 2010). Indeed, a large variation has been found in photoperiod response among different accessions, with

some deviating from the typical critical daylengths (Giakountis et al., 2010). There is currently very limited data on this accession in the literature, thus the photoperiodic window for Fei is not clear. Unfortunately, it is not feasible to refer to data of accessions collected from similar latitude as no strong correlation was found between photoperiod response and latitude (Giakountis et al., 2010). As both the photoperiod pathway and starch degradation rate are regulated by the circadian clock (Graf et al., 2010; Lu et al., 2005), it is tempting to hypothesise that Fei has a lower critical daylength and that the same regulatory mechanism is causing starch turnover to decrease, thus supporting the ST_c adjustment exercise. This is however only a speculation, and until more data are available for Fei and which support the assumption, it will remain one.

4.3.11 Model synergy

Models with increased complexity often come with improved fit but reduced flexibility. In the case of integrating existing models of biological processes either by direct linking or component replacement, the increased complexity offers more biological details but does not necessarily improve fit, while flexibility comes in the form of species specificity down to the molecular level.

In this study, the multi-scale whole-plant model not only displayed reasonably good fit to validation data of different accessions, but it could also provide extra information and synergy compared to the individual model components. For example, the integrated model could now simulate growth from plant emergence until flowering, displaying a synergy between the Rasse and Tocquin carbon assimilation model and the Wilczek et al. phenology model. Whereas the assimilation model only considers the rosette as one big leaf, including the GreenLab functional-structural growth model (Christophe et al., 2008) gave information on the growth of individual organs and how each organ contributes to the whole plant structure. In return, model integration diminishes dependency of the GreenLab model

on optimisation for radiation use efficiency, which is a lumped parameter of photosynthesis, sugar-starch partitioning, respiration and daily allowable growth rate.

In addition to model synergy, the multi-scale model could simulate the diurnal growth rhythm for the whole vegetative stage, which interestingly illustrated a sink-limited phase in young plants, as indicated by the spikes in sugar content during the night (Fig. 4.3.15). Since the model has not been validated with any metabolite data, this might be the result of limited biological details in the model, but it also points out that the source-sink relation throughout plant life cycle could be very dynamic, as has been demonstrated in some plants (Lewis, Wang, Griffin, & Tissue, 2002) and crop cultivars (Egli, 1999; Pavel & Dejong, 1993). Studies have shown that starch degradation rate can adjust very quickly to night length to avoid immediate carbon starvation (Gibon et al., 2004; Gibon et al., 2009; Graf et al., 2010; Lu et al., 2005), and this rate adjustment is included in the model from Rasse and Tocquin (2006). By the same logic, carbon utilisation could be optimised in the long run if starch turnover rate is adjusted according to sink status to avoid carbon wastage. If the sink demand is low, transitory starch would be better saved for future use instead of being converted to non-needed sugar (Fig. 4.3.15a), which could in turn lead to high maintenance respiration (Ogren, 2000). Alternatively, photosynthesis rate can be reduced to lower the overall costs (Krapp, Hofmann, Schafer, & Stitt, 1993). These feedback relations between the source and the sink are not currently described in the model. A cost-benefit analysis of the whole plant system throughout its life cycle may provide a better understanding of the need for any tight coupling or feedback mechanisms between different pathways. Such analysis might also explain the diel and/or circadian rhythm of growth that has been reported in leaves, hypocotyls and roots (Dowson-Day & Millar, 1999; Nozue et al., 2007; Poire et al., 2010; Wiese, Christ, Virnich, Schurr, & Walter, 2007; Yazdanbakhsh, Sulpice, Graf, Stitt, & Fisahn, 2011). These emergent properties can only be analysed using whole-plant models that describe biological processes to a certain level of detail.

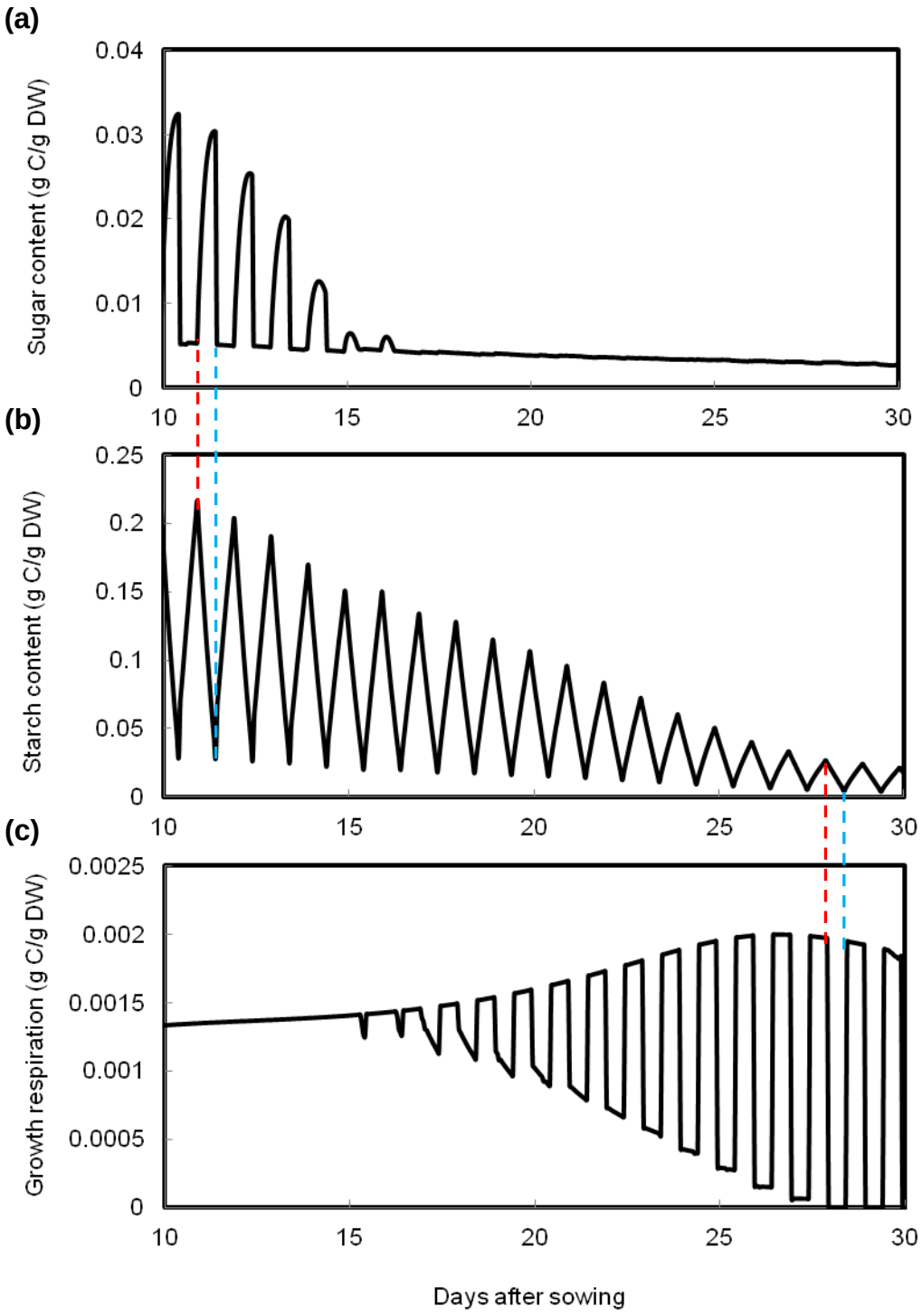


Figure 4.3.15: Simulated time series during the vegetative stage for Ler using the same input conditions as in Fig. 4.3.10. (a) Sugar content per total shoot dry mass; (b) Starch content per total shoot dry mass; (c) Growth respiration per total shoot dry mass. Dashed lines indicate end-of-day (red) and end-of night (blue) in the time series. A change from sink-limited (indicated by high sugar content at night and a constantly high growth respiration) to source-limited growth (indicated by low sugar content and very low growth respiration at night) was observed.

Chapter 5

General Discussion

Results presented in the previous chapter have demonstrated how the integrated model can provide a framework that already matched biomass and gas exchange data from the lab to good accuracies, and upon further extension, may facilitate analyses for improving our understanding of the whole-plant system. In this chapter, model limitations and future extensions are presented, followed by the implications of study.

5.1 Model limitations and future extensions

At present, the Arabidopsis multi-scale model is necessarily simplistic. Complex biological pathways, biochemical reactions and physiological responses are only represented by simple equations due to incomplete understanding of the system, technical limitations in data collection and model availabilities.

The most detailed module in the framework is currently the clock network (Appendix B), where the early two-loop model (Appendix B1) (Locke et al., 2005) is utilised. Many new clock components have since been added to the clock gene circuit (Pokhilko et al., 2012; Pokhilko et al., 2010), which improves model behaviour under a wider range of light conditions and more genotypes can be described. Once the *FT*-activation component is connected to the new model versions and re-parameterised in the future, the two-loop model can be replaced in a straightforward manner by re-calibrating integral of *FT* to the phenology photoperiod component (Section 4.2.2). Comprehensive analyses of flowering data, i.e. flowering time and leaf number, and *FT* expression levels across a range of photoperiods would improve such calibration method in the future. A technical limitation of the

phenology model, however, is that it uses the times of sunrise and sunset in hourly calculations throughout the day (Section 4.2.2). While the clock may allow plants to anticipate dawn and dusk, the relevant molecular events are dynamically modulated by ongoing environmental stimuli, which may either cause immediate responses or be stored as "memory" that affects later processes (Tafforeau, Verdus, Norris, Ripoll, & Thellier, 2006; Verdus, Ripoll, Norris, & Thellier, 2012). An immediate response in this case would mean that photoperiodic induction of flowering changes depending on the hourly level of *FT*. A memory concept may be more applicable, where the total amount (or integral) of *FT* accumulated over the previous 24 hours from the hour at hand is considered. This is consistent with the observation that a single long day is insufficient for photoperiodic induction in young Arabidopsis, suggesting that information accumulates over several days (Corbesier, Gadisseur, Silvestre, Jacqmard, & Bernier, 1996). It could also embody the lag time required to transport FT protein from the leaves to the meristem (Corbesier et al., 2007a). Implementing this change in the model would not affect the simulation for constant photoperiod conditions; however, under varying photoperiods the predicted flowering time may be different. This dynamic (memory) response may therefore increase model precision, and more importantly, analysis of the differences will allow us to identify potential challenges in the direct linking of gene dynamics to scaling factors that are widely used in crop phenology models.

Another module with more kinetic details in the integrated model is the carbon assimilation component. In this component, the Farquhar et al. (1980) model utilises a combination of mass-action kinetics, the Michaelis-Menten model, and specially-derived functions to consider the first two of the main stages in the Calvin-Benson cycle, i.e. carboxylation, reduction and regeneration. Arnold and Nikoloski (2011) highlighted that this model is overall the best in terms of accuracy, robustness and sensitivity in describing carbon fixation processes, but it would not be suitable for metabolic engineering applications compared to some of its counterparts, which described sub-processes down to the level of single reactions (Giersch, Lammel, & Farquhar, 1990; Laisk, Eichelmann, & Oja, 2006; Poolman, Fell, & Thomas, 2000). The necessity of adding more kinetic details therefore depends on the objectives to

be achieved. In any case, linking model parameters such as carboxylation and electron transport capacities, i.e. V_{cmax} and J_{max}, to protein or molecular contents would be useful. Studies have shown that these two parameters are dependent on temperature and light intensity (Bunce, 2008; Pons, 2012), and comparison of their measured values seems to indicate a dependency on photoperiod (Table 4.3.3). This suggested a possible regulation by either the clock, carbon status, or both. Many key metabolic pathways are regulated by the clock (Harmer et al., 2000), one of which is starch degradation rate (Graf et al., 2010). In the model, this regulation is simplified to a linear rate adjusted automatically to night length (C1.15). Linking the clock module to automate this rate (and other clock-regulated metabolic reactions) would require knowledge of the responsible clock component(s) and its downstream mechanism. These, and other feedback machineries, are the subject of ongoing investigations (Pyl et al., 2012; Stitt & Zeeman, 2012; Timm et al., 2012). Most components of the Rasse and Tocquin model are parameterised with data from optimal growth conditions, and under these conditions regulatory effects may well not be obvious. Consequently, the model's capacity to describe carbon assimilation and partitioning under excessive conditions is still very limited. A previous study applying a range of constant photoperiods down to 3 hours revealed a close coordinated adjustment of starch turnover, protein content and relative growth rate to avoid carbon starvation (Gibon et al., 2009). This study is one example of how artificial conditions that are not natural in the field could sometimes reveal new testable hypotheses. Therefore, identification and inclusion of the underlying metabolic and growth regulations in the multi-scale model could enable simulation of plant behaviour under extreme conditions such as constant light and skeleton photoperiods (light pulses), which may reveal additional emergent properties.

The multi-scale model also included a module of biomass allocation to the shoots and the roots based on organ sink strength and expansion duration (equations 4.3.6 and 4.3.7). Currently, the model assumes the root expansion period to be a function of flowering time (and indirectly temperature and photoperiod) founded on only one set of experimental observation (Christophe et al., 2008). This expansion period has significant effect on the root-to-shoot allocation ratio (*RS*) since it determines root sink variation (equation 4.3.7). The integrated model has not been validated with any root data, and examination of module compatibility in *RS* illustrated some differences between models during the early and later stages (Fig. 4.3.8). The Rasse and Tocquin model employs a minimum *RS* threshold for seedlings, whereas the multi-scale model determines the ratio based on organ sink variation that depends on temperature (roots and leaves) and photoperiod (roots). It is not clear if a minimum ratio exists since the Rasse and Tocquin model was not calibrated with data from young seedlings, but roots appear first during germination supporting a high *RS* assumption. On the other hand, root and shoot development can be very plastic depending on nutrient availability (Roycewicz & Malamy, 2012), and they are also coordinated by light (Salisbury, Hall, Grierson, & Halliday, 2007; Sassi et al., 2012). At the moment, germination events and the early seedling stage are not described in the multi-scale model, and most initial values such as sugar/starch level and cotyledon size at plant emergence are assumed to be constant regardless of earlier events (Section 4.3.4). Studies on seedlings are already well established, with large amount of data available on hypocotyl and cotyledon growth for various light and temperature conditions. Incorporating a module that describes these events in the future would allow impact study of early events on plant life cycle. As the plants age, a general rule for herbaceous plants is that *RS* decreases (Bourdot, Saville, & Field, 1984; Bray, 1963; Ryle, Arnott, & Powell, 1981; Troughton, 1956), and this trend was also shown in the simulation for older plants in the integrated model (Fig. 4.3.8). A periodic harvest of roots from germination until senescence of plants grown under different conditions would be required to establish any relationship between root growth and flowering time and/or other variables, though such data are still limited due to experimental difficulties.

Besides root growth, a correlation between leaf expansion and photoperiod has been reported (Cookson, Chenu, & Granier, 2007). In short days, leaf expansion rate was found to decrease while the expansion duration increased. These two variables are very plastic in dicotyledonous leaves, and their opposing responses to environmental changes may sometimes offer partial or complete compensation such that the final leaf area is not affected (Granier & Tardieu, 2009). At present, leaf expansion duration in the model is controlled only by temperature (thermal time) without any effect from photoperiod. The model also assumes that all leaves that are still expanding function as carbon sink, such that the ontogenic switch from metabolic to hydraulic control in leaf expansion (Pantin, Simonneau, Rolland, Dauzat, & Muller, 2011) is not considered. Nor does the model consider any environmental regulation of petiole elongation and circadian control of leaf movement. A comprehensive analysis of published data and additional experiments would enhance our understanding and provide correlations that improve the environmental-functional-structural relationship in the model.

The model is of a deterministic nature at the moment. Stochasticity occurs at all levels of organisation, but the non-symmetrical branching in trees provides strong evidence that organogenesis is highly variable and plastic to the environment. Data of leaf appearance rate observed under long and short day conditions for Arabidopsis (Mendez-Vigo et al., 2010) seemed to indicate an effect of photoperiod, possibly due to a feedback mechanism of carbon availability as has been observed on plants grown under different sucrose concentrations (Roldan, Gomez-Mena, Ruiz-Garcia, Salinas, & Martinez-Zapater, 1999). Studies on the shoot apical meristem have revealed that leaf formation and phyllotaxis are regulated by auxin transport and signalling (Benkova et al., 2003; Leyser, 2006; Reinhardt et al., 2003). There are already models, both deterministic and stochastic, that describe these regulations (Jonsson et al., 2006; Twycross et al., 2010). Therefore, integrating these modules in the future not only connects the current work to hormone signalling and developmental biology, but it could also introduce some stochasticity into the model to explain variability in plant growth to better represent a population.

Most of the model extensions discussed so far would require data generated under controlled conditions in the lab. Eventually, it would also be important to test the model on plants grown in the field as part of the validation process. Analysis of any discrepancies that arise could help identify model components that may require further extension. Only with such iterative process can the model be truly exploited to its full potential, both in knowledge discovery and to help address global issues.

5.2 Implications of study

The Arabidopsis multi-scale whole-plant model developed here represents a new class of models that embody the concept of crop systems biology (Yin & Struik, 2010), and it can provide a framework to bridge between the reductionism and holistic approaches. The model is a good example of how information from functional genomics can be applied at the whole-plant level. Not only are gene functions incorporated by simple modification of parameter values, but gene dynamics at the transcriptional and translational levels are linked to model parameters. Analysis of the model has revealed emergent properties associated with competition for carbon sources within the plants at the organ and metabolic levels. It is also possible to incorporate more biological mechanisms in the future due to the modular structure of the model. These properties are in line with the attributes of crop systems biology outlined in Yin and Struik (2010).

Global concerns regarding the impact of climate change have recently generated various research activities to identify alternatives that improve plant yield and productivity. Ongoing efforts to improve photosynthetic capacity include manipulating the function of enzymes such as RuBisCO, and exploring the potential engineering of photosynthetic components in C4 plants into C3 plants (Leakey & Lau, 2012). Another group of studies focuses on enhancing starch storage by manipulating starch biosynthesis and remobilisation to improve growth at night or during carbon shortage (Geigenberger, 2011). These elements are also considered in the Arabidopsis multi-scale model, either with kinetic details or as simple functions

where more details can be incorporated. As a model species, the genome of Arabidopsis has been fully sequenced (Kaul et al., 2000), followed by crop species such as rice (Goff et al., 2002; Matsumoto et al., 2005; Yu et al., 2002), maize (Schnable et al., 2009) and soybean (Schmutz et al., 2010). There is still a long way towards developing a plant model that considers the full genome, as has been successfully demonstrated in the breakthrough bacteria whole-cell model (Karr et al., 2012). Nevertheless, the current study is part of the initial steps, and coupled with the knowledge-swapping efforts between Arabidopsis and crops (Chew & Halliday, 2011), could pave the way for a full-genome plant model and facilitate synthetic biology (Collins, 2012).

Chapter 6

Conclusions

In conclusion, this study demonstrated how a multi-scale whole-plant model was developed for Arabidopsis using a modular approach. The model was successfully validated with biomass data from individual plants and gas exchange measurements for a small population. Modifying model parameters associated with measured traits enabled the model to describe various genotypes. As each module is independent and can be analysed separately, examination of the phenology component and meteorological data has revealed a seasonal effect of night temperatures on flowering time. Further analysis suggested the involvement of photoreceptor phyB in the gating of temperature effects on flowering. This study has also linked a simple phenology model to gene dynamics, an initiative that could bridge the gap between classical crop modelling, which have a long-standing history, and modern science. Future work that addresses model limitations would improve its capability and provide application routes from fundamental biology to crop improvement and biosphere management.

References

Allen, R. G., Pereira, L. S., Raes, D., & Smith, M. (1998). *Crop evapotranspiration. Guidelines for computing crop water requirements.* Paper presented at the FAO Irrigation and Drainage ROME.

Amasino, R. (2010). Seasonal and developmental timing of flowering. *Plant Journal, 61*(6), 1001-1013.

Anderson, M., & Mulligan, B. (1992). *Arabidopsis* mutant collection. In C. Koncz, N. H. Chua & J. Schell (Eds.), *Methods in Arabidopsis Research* (pp. 419 - 437). Singapore: World Scientific.

ap Rees, T., Burrell, M. M., Entwistle, T. G., Hammond, J. B. W., Kirk, D., & Kruger, N. J. (1988). Effects of Low-Temperature on the Respiratory Metabolism of Carbohydrates by Plants. *Plants and Temperature, 42*, 377-393.

Arana, M. V., Marin-de la Rosa, N., Maloof, J. N., Blazquez, M. A., & Alabadi, D. (2011). Circadian oscillation of gibberellin signaling in Arabidopsis. *Proc Natl Acad Sci U S A, 108*(22), 9292-9297.

Arnold, A., & Nikoloski, Z. (2011). A quantitative comparison of Calvin-Benson cycle models. *Trends in Plant Science, 16*(12), 676-683.

Bagnall, D. J., King, R. W., Whitelam, G. C., Boylan, M. T., Wagner, D., & Quail, P. H. (1995). Flowering Responses to Altered Expression of Phytochrome in Mutants and Transgenic Lines of Arabidopsis-Thaliana (L) Heynh. *Plant Physiology, 108*(4), 1495-1503.

Baker, C. K., Gallagher, J. N., & Monteith, J. L. (1980). Daylength Change and Leaf Appearance in Winter-Wheat. *Plant Cell and Environment, 3*(4), 285-287.

Balasubramanian, S., Sureshkumar, S., Lempe, J., & Weigel, D. (2006). Potent induction of Arabidopsis thaliana flowering by elevated growth temperature. *Plos Genetics, 2*(7), 980-989.

Bastow, R., Mylne, J. S., Lister, C., Lippman, Z., Martienssen, R. A., & Dean, C. (2004). Vernalization requires epigenetic silencing of FLC by histone methylation. *Nature, 427*(6970), 164-167.

Benkova, E., Michniewicz, M., Sauer, M., Teichmann, T., Seifertova, D., Jurgens, G., et al. (2003). Local, efflux-dependent auxin gradients as a common module for plant organ formation. *Cell, 115*(5), 591-602.

Berry, J., & Bjorkman, O. (1980). Photosynthetic Response and Adaptation to Temperature in Higher-Plants. *Annual Review of Plant Physiology and Plant Molecular Biology, 31*, 491-543.

Blazquez, M. A., Ahn, J. H., & Weigel, D. (2003). A thermosensory pathway controlling flowering time in Arabidopsis thaliana. *Nature Genetics, 33*(2), 168-171.

Borchert, R., Renner, S. S., Calle, Z., Navarrete, D., Tye, A., Gautier, L., et al. (2005). Photoperiodic induction of synchronous flowering near the Equator. *Nature, 433*(7026), 627-629.

Bourdot, G. W., Saville, D. J., & Field, R. J. (1984). The Response of Achillea-Millefolium L (Yarrow) to Shading. *New Phytologist, 97*(4), 653-663.

Bowman, J. (1993). *Arabidopsis: An Atlas of Morphology and Development*. New York: Springer-Verlag.

Boyes, D. C., Zayed, A. M., Ascenzi, R., McCaskill, A. J., Hoffman, N. E., Davis, K. R., et al. (2001). Growth stage-based phenotypic analysis of arabidopsis: A model for high throughput functional genomics in plants. *Plant Cell, 13*(7), 1499-1510.

Brachi, B., Faure, N., Horton, M., Flahauw, E., Vazquez, A., Nordborg, M., et al. (2010). Linkage and Association Mapping of Arabidopsis thaliana Flowering Time in Nature. *Plos Genetics, 6*(5).

Bray, J. R. (1963). Root production and the estimation of net productivity. *Canadian Journal of Botany, 41*, 65 - 72.

Bunce, J. A. (2008). Acclimation of photosynthesis to temperature in Arabidopsis thaliana and Brassica oleracea. *Photosynthetica, 46*(4), 517-524.

Caspar, T., Huber, S. C., & Somerville, C. (1985). Alterations in Growth, Photosynthesis, and Respiration in a Starchless Mutant of Arabidopsis-Thaliana (L) Deficient in Chloroplast Phosphoglucomutase Activity. *Plant Physiology, 79*(1), 11-17.

Cerdan, P. D., & Chory, J. (2003). Regulation of flowering time by light quality. *Nature, 423*(6942), 881-885.

Chapin, F. S., Woodwell, G. M., Randerson, J. T., Rastetter, E. B., Lovett, G. M., Baldocchi, D. D., et al. (2006). Reconciling carbon-cycle concepts, terminology, and methods. *Ecosystems, 9*(7), 1041-1050.

Chelle, M., & Andrieu, B. (1998). The nested radiosity model for the distribution of light within plant canopies. *Ecological Modelling, 111*(1), 75-91.

Chenu, K., Franck, N., Dauzat, J., Barczi, J. F., Rey, H., & Lecoeur, J. (2005). Integrated responses of rosette organogenesis, morphogenesis and architecture to reduced incident light in Arabidopsis thaliana results in higher efficiency of light interception. *Functional Plant Biology, 32*(12), 1123-1134.

Chew, Y. H., & Halliday, K. J. (2011). A stress-free walk from Arabidopsis to crops. *Current Opinion in Biotechnology, 22*(2), 281-286.

Chew, Y. H., Wilczek, A. M., Williams, M., Welch, S. M., Schmitt, J., & Halliday, K. J. (2012). An augmented Arabidopsis phenology model reveals seasonal temperature control of flowering time. *New Phytologist, 194*(3), 654 - 665.

Christophe, A., Letort, V., Hummel, I., Cournede, P. H., de Reffye, P., & Lecoeur, J. (2008). A model-based analysis of the dynamics of carbon balance at the whole-plant level in Arabidopsis thaliana. *Functional Plant Biology, 35*(11), 1147-1162.

Chuine, I. (2000). A unified model for budburst of trees. *Journal of Theoretical Biology, 207*(3), 337-347.

Clerget, B., Dingkuhn, M., Chantereau, J., Hemberger, J., Louarn, G., & Vaksmann, M. (2004). Does panicle initiation in tropical sorghum depend on day-to-day change in photoperiod? *Field Crops Research, 88*(1), 21-37.

Collins, J. (2012). SYNTHETIC BIOLOGY Bits and pieces come to life. *Nature, 483*(7387), S8-S10.

Cookson, S. J., Chenu, K., & Granier, C. (2007). Day length affects the dynamics of leaf expansion and cellular development in Arabidopsis thaliana partially through floral transition timing. *Annals of Botany, 99*(4), 703-711.

Corbesier, L., Gadisseur, I., Silvestre, G., Jacqmard, A., & Bernier, G. (1996). Design in Arabidopsis thaliana of a synchronous system of floral induction by one long day. *Plant Journal, 9*(6), 947-952.

Corbesier, L., Vincent, C., Jang, S. H., Fornara, F., Fan, Q. Z., Searle, I., et al. (2007a). FT protein movement contributes to long-distance signaling in floral induction of Arabidopsis. *Science, 316*(5827), 1030 - 1033.

Corbesier, L., Vincent, C., Jang, S. H., Fornara, F., Fan, Q. Z., Searle, I., et al. (2007b). FT protein movement contributes to long-distance signaling in floral induction of Arabidopsis. *Science, 316*(5827), 1030-1033.

Cournede, P. H., Kang, M. Z., Mathieu, A., Barczi, J. F., Yan, H. P., Hu, B. G., et al. (2006). Structural factorization of plants to compute their functional and architectural growth. *Simulation, 82*(7), 427 - 438.

Dauzat, J. (1994). Radiative-Transfer Simulation on Computer-Models of Elaeis-Guineensis. *Oleagineux, 49*(3), 81-90.

Dauzat, J., & Eroy, M. N. (1997). Simulating light regime and intercrop yields in coconut based farming systems. *European Journal of Agronomy, 7*(1-3), 63-74.

David, K. M., Armbruster, U., Tama, N., & Putterill, J. (2006). Arabidopsis GIGANTEA protein is post-transcriptionally regulated by light and dark. *Febs Letters, 580*(5), 1193-1197.

Dewar, R. C., Medlyn, B. E., & McMurtrie, R. E. (1998). A mechanistic analysis of light and carbon use efficiencies. *Plant Cell and Environment, 21*(6), 573-588.

Di Ventura, B., Lemerle, C., Michalodimitrakis, K., & Serrano, L. (2006). From in vivo to in silico biology and back. *Nature, 443*(7111), 527-533.

Dingkuhn, M., Luquet, D., Quilot, B., & de Reffye, P. (2005). Environmental and genetic control of morphogenesis in crops: towards models simulating phenotypic plasticity. *Australian Journal of Agricultural Research, 56*(11), 1289-1302.

Dorca-Fornell, C., Gregis, V., Grandi, V., Coupland, G., Colombo, L., & Kater, M. M. (2011). The Arabidopsis SOC1-like genes AGL42, AGL71 and AGL72 promote flowering in the shoot apical and axillary meristems. *Plant Journal, 67*(6), 1006-1017.

Dowson-Day, M. J., & Millar, A. J. (1999). Circadian dysfunction causes aberrant hypocotyl elongation patterns in Arabidopsis. *Plant Journal, 17*(1), 63-71.

Edwards, K. D., Akman, O. E., Knox, K., Lumsden, P. J., Thomson, A. W., Brown, P. E., et al. (2010). Quantitative analysis of regulatory flexibility under changing environmental conditions. *Molecular Systems Biology, 6*, -.

Egli, D. B. (1999). Variation in leaf starch and sink limitations during seed filling in soybean. *Crop Science, 39*(5), 1361-1368.

Eichelmann, H., & Laisk, A. (1994). CO_2 Uptake and Electron-Transport Rates in Wild-Type and a Starchless Mutant of Nicotiana-Sylvestris - the Role and Regulation of Starch Synthesis at Saturating Co2 Concentrations. *Plant Physiology, 106*(2), 679-687.

Ellis, R. P., & Russell, G. (1984). Plant Development and Grain-Yield in Spring and Winter Barley. *Journal of Agricultural Science, 102*(Feb), 85-95.

Espinosa-soto, C., Padilla-Longoria, P., & Alvarez-Buylla, E. R. (2004). A gene regulatory network model for cell-fate determination during Arabidopsis thalianal flower development that is robust and recovers experimental gene expression profiles. *Plant Cell, 16*(11), 2923-2939.

Espinoza, C., Degenkolbe, T., Caldana, C., Zuther, E., Leisse, A., Willmitzer, L., et al. (2010). Interaction with Diurnal and Circadian Regulation Results in Dynamic Metabolic and Transcriptional Changes during Cold Acclimation in Arabidopsis. *Plos One, 5*(11).

Estrella, N., Sparks, T. H., & Menzel, A. (2007). Trends and temperature response in the phenology of crops in Germany. *Global Change Biology, 13*(8), 1737-1747.

Evers, J. B., Vos, J., Yin, X., Romero, P., van der Putten, P. E. L., & Struik, P. C. (2010). Simulation of wheat growth and development based on organ-level photosynthesis and assimilate allocation. *Journal of Experimental Botany, 61*(8), 2203-2216.

Farquhar, G. D., Caemmerer, S. V., & Berry, J. A. (1980). A Biochemical-Model of Photosynthetic Co2 Assimilation in Leaves of C-3 Species. *Planta, 149*(1), 78-90.

Farquhar, G. D., von Caemmerer, S., & Berry, J. A. (2001). Models of photosynthesis. *Plant Physiology, 125*(1), 42-45.

Feugier, F. G., Mochizuki, A., & Iwasa, Y. (2005). Self-organization of the vascular system in plant leaves: Inter-dependent dynamics of auxin flux and carrier proteins. *Journal of Theoretical Biology, 236*(4), 366-375.

Flemisch, B., Darcis, M., Erbertseder, K., Faigle, B., Lauser, A., Mosthaf, K., et al. (2011). DuMu(x): DUNE for multi-{phase, component, scale, physics, ...} flow and transport in porous media. *Advances in Water Resources, 34*(9), 1102-1112.

Flexas, J., Ortuno, M. F., Ribas-Carbo, M., Diaz-Espejo, A., Florez-Sarasa, I. D., & Medrano, H. (2007). Mesophyll conductance to CO2 in Arabidopsis thaliana. *New Phytologist, 175*(3), 501-511.

Fornara, F., Panigrahi, K. C. S., Gissot, L., Sauerbrunn, N., Ruhl, M., Jarillo, J. A., et al. (2009). Arabidopsis DOF Transcription Factors Act Redundantly to Reduce CONSTANS Expression and Are Essential for a Photoperiodic Flowering Response. *Developmental Cell, 17*(1), 75-86.

Fowler, S., Lee, K., Onouchi, H., Samach, A., Richardson, K., Coupland, G., et al. (1999). GIGANTEA: a circadian clock-controlled gene that regulates photoperiodic flowering in Arabidopsis and encodes a protein with several possible membrane-spanning domains. *Embo Journal, 18*(17), 4679-4688.

Fowler, S. G., Cook, D., & Thomashow, M. E. (2005). Low temperature induction of Arabidopsis CBF1, 2, and 3 is gated by the circadian clock. *Plant Physiology, 137*(3), 961-968.

Franklin, K. A. (2009). Light and temperature signal crosstalk in plant development. *Current Opinion in Plant Biology, 12*(1), 63-68.

Franklin, K. A., & Quail, P. H. (2010). Phytochrome functions in Arabidopsis development. *Journal of Experimental Botany, 61*(1), 11-24.

Franklin, K. A., & Whitelam, G. C. (2004). Light signals, phytochromes and cross-talk with other environmental cues. *Journal of Experimental Botany, 55*(395), 271-276.

Gates, T. S., Odegard, G. M., Frankland, S. J. V., & Clancy, T. C. (2005). Computational materials: Multi-scale modeling and simulation of nanostructured materials. *Composites Science and Technology, 65*(15-16), 2416-2434.

Geigenberger, P. (2011). Regulation of Starch Biosynthesis in Response to a Fluctuating Environment. *Plant Physiology, 155*(4), 1566-1577.

Gendall, A. R., Levy, Y. Y., Wilson, A., & Dean, C. (2001). The VERNALIZATION 2 gene mediates the epigenetic regulation of vernalization in Arabidopsis. *Cell, 107*(4), 525-535.

Giakountis, A., Cremer, F., Sim, S., Reymond, M., Schmitt, J., & Coupland, G. (2010). Distinct Patterns of Genetic Variation Alter Flowering Responses of Arabidopsis Accessions to Different Daylengths. *Plant Physiology, 152*(1), 177-191.

Gibon, Y., Blasing, O. E., Palacios-Rojas, N., Pankovic, D., Hendriks, J. H. M., Fisahn, J., et al. (2004). Adjustment of diurnal starch turnover to short days: depletion of sugar during the night leads to a temporary inhibition of carbohydrate utilization, accumulation of sugars and post-translational activation of ADP-glucose pyrophosphorylase in the following light period. *Plant Journal, 39*(6), 847-862.

Gibon, Y., Pyl, E. T., Sulpice, R., Lunn, J. E., Hohne, M., Gunther, M., et al. (2009). Adjustment of growth, starch turnover, protein content and central metabolism to a decrease of the carbon supply when Arabidopsis is grown in very short photoperiods. *Plant Cell and Environment, 32*(7), 859-874.

Giersch, C., Lammel, D., & Farquhar, G. (1990). Control Analysis of Photosynthetic Co2 Fixation. *Photosynthesis Research, 24*(2), 151-165.

Godin, C., & Sinoquet, H. (2005). Functional-structural plant modelling. *New Phytologist, 166*(3), 705-708.

Goff, S. A., Ricke, D., Lan, T. H., Presting, G., Wang, R. L., Dunn, M., et al. (2002). A draft sequence of the rice genome (Oryza sativa L. ssp japonica). *Science, 296*(5565), 92-100.

Gorsuch, P. A., Pandey, S., & Atkin, O. K. (2010a). Temporal heterogeneity of cold acclimation phenotypes in Arabidopsis leaves. *Plant Cell and Environment, 33*(2), 244-258.

Gorsuch, P. A., Pandey, S., & Atkin, O. K. (2010b). Thermal de-acclimation: how permanent are leaf phenotypes when cold-acclimated plants experience warming? *Plant Cell and Environment, 33*(7), 1124-1137.

Graf, A., Schlereth, A., Stitt, M., & Smith, A. M. (2010). Circadian control of carbohydrate availability for growth in Arabidopsis plants at night. *Proceedings of the National Academy of Sciences of the United States of America, 107*(20), 9458-9463.

Grafahrend-Belau, E., Schreiber, F., Koschutzki, D., & Junker, B. H. (2009). Flux Balance Analysis of Barley Seeds: A Computational Approach to Study Systemic Properties of Central Metabolism. *Plant Physiology, 149*(1), 585-598.

Granier, C., Massonnet, C., Turc, O., Muller, B., Chenu, K., & Tardieu, F. (2002). Individual leaf development in Arabidopsis thaliana: a stable thermal-time-based programme. *Annals of Botany, 89*(5), 595-604.

Granier, C., & Tardieu, F. (2009). Multi-scale phenotyping of leaf expansion in response to environmental changes: the whole is more than the sum of parts. *Plant Cell and Environment, 32*(9), 1175-1184.

Greb, T., Mylne, J. S., Crevillen, P., Geraldo, N., An, H. L., Gendall, A. R., et al. (2007). The PHD finger protein VRN5 functions in the epigenetic silencing of Arabidopsis FLC. *Current Biology, 17*(1), 73-78.

Halliday, K. J., Koornneef, M., & Whitelam, G. C. (1994). Phytochrome B and at Least One Other Phytochrome Mediate the Accelerated Flowering Response of Arabidopsis-Thaliana L to Low Red/Far-Red Ratio. *Plant Physiology, 104*(4), 1311-1315.

Halliday, K. J., Salter, M. G., Thingnaes, E., & Whitelam, G. C. (2003). Phytochrome control of flowering is temperature sensitive and correlates with expression of the floral integrator FT. *Plant Journal, 33*(5), 875-885.

Halliday, K. J., & Whitelam, G. C. (2003). Changes in photoperiod or temperature alter the functional relationships between phytochromes and reveal roles for phyD and phyE. *Plant Physiology, 131*(4), 1913-1920.

Hanano, S., & Goto, K. (2011). Arabidopsis TERMINAL FLOWER1 Is Involved in the Regulation of Flowering Time and Inflorescence Development through Transcriptional Repression. *Plant Cell, 23*(9), 3172-3184.

Harmer, S. L., Hogenesch, L. B., Straume, M., Chang, H. S., Han, B., Zhu, T., et al. (2000). Orchestrated transcription of key pathways in Arabidopsis by the circadian clock. *Science, 290*(5499), 2110-2113.

Harrington, C. A., Gould, P. J., & St Clair, J. B. (2010). Modeling the effects of winter environment on dormancy release of Douglas-fir. *Forest Ecology and Management, 259*(4), 798-808.

Heo, J. B., & Sung, S. (2011). Vernalization-Mediated Epigenetic Silencing by a Long Intronic Noncoding RNA. *Science, 331*(6013), 76-79.

Hepworth, S. R., Valverde, F., Ravenscroft, D., Mouradov, A., & Coupland, G. (2002). Antagonistic regulation of flowering-time gene SOC1 by CONSTANS and FLC via separate promoter motifs. *Embo Journal, 21*(16), 4327-4337.

Ho, Q. T., Verboven, P., Verlinden, B. E., & Nicolai, B. M. (2010). A model for gas transport in pear fruit at multiple scales. *Journal of Experimental Botany, 61*(8), 2071-2081.

Honda, H. (1971). Description of Form of Trees by Parameters of Tree-Like Body - Effects of Branching Angle and Branch Length on Shape of Tree-Like Body. *Journal of Theoretical Biology, 31*(2), 331-&.

Hoogenboom, G.J., Jones, J.W., Wilkens, P.W., Batchelor, W.D., Bowen, W.T., Hunt, L.A., Pickering, N., Singh, U., Godwin, D.C., Baer, B., Boote, K.J., Ritchie, J.T., & White, J.W. (1994). Crop models. In: Tsuji, G.Y., Uehara, G., Balas, S. (eds.), DSSAT v3, vol. 2 /2. University of Hawaii, Honolulu, Hawaii.

Ideker, T., Galitski, T. & Hood, L. (2001). A new approach to decoding life: Systems biology. *Annual Review of Genomics and Human Genetics, 2*, 343-372.

Imaizumi, T., & Kay, S. A. (2006). Photoperiodic control of flowering: not only by coincidence. *Trends in Plant Science, 11*(11), 550-558.

Imaizumi, T., Schultz, T. F., Harmon, F. G., Ho, L. A., & Kay, S. A. (2005). FKF1F-BOX protein mediates cyclic degradation of a repressor of CONSTANS in Arabidopsis. *Science, 309*(5732), 293-297.

Ishikawa, M., Kiba, T., & Chua, N. H. (2006). The Arabidopsis SPA1 gene is required for circadian clock function and photoperiodic flowering. *Plant Journal, 46*(5), 736-746.

Jack, T. (2004). Molecular and genetic mechanisms of floral control. *Plant Cell, 16*, S1-S17.

Jaeger, M., & De Reffye, P. (1992). Basic Concepts of Computer-Simulation of Plant-Growth. *Journal of Biosciences, 17*(3), 275-291.

Jones, C. A., & Kiniry, J. R. (1986). *CERES-Maize: A simulation model of maize growth and development*. Texas: Texas A&M University Press.

Jonsson, H., Heisler, M. G., Shapiro, B. E., Meyerowitz, E. M., & Mjolsness, E. (2006). An auxin-driven polarized transport model for phyllotaxis. *Proceedings of the National Academy of Sciences of the United States of America, 103*(5), 1633-1638.

Kang, M. Z., Cournede, P. H., de Reffye, P., Auclair, D., & Hu, B. G. (2008). Analytical study of a stochastic plant growth model: Application to the GreenLab model. *Mathematics and Computers in Simulation, 78*(1), 57-75.

Karr, J. R., Sanghvi, J. C., Macklin, D. N., Gutschow, M. V., Jacobs, J. M., Bolival, B., et al. (2012). A Whole-Cell Computational Model Predicts Phenotype from Genotype. *Cell, 150*(2), 389-401.

Kaul, S., Koo, H. L., Jenkins, J., Rizzo, M., Rooney, T., Tallon, L. J., et al. (2000). Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. *Nature, 408*(6814), 796-815.

Kernich, G. C., Slafer, G. A., & Halloran, G. M. (1995). Barley Development as Affected by Rate of Change of Photoperiod. *Journal of Agricultural Science, 124*, 379-388.

Kim, H. J., Hyun, Y., Park, J. Y., Park, M. J., Park, M. K., Kim, M. D., et al. (2004). A genetic link between cold responses and flowering time through FVE in Arabidopsis thaliana. *Nature Genetics, 36*(2), 167-171.

Kim, W., Ahn, H. J., Chiou, T. J., & Ahn, J. H. (2011). The role of the miR399-PHO2 module in the regulation of flowering time in response to different ambient temperatures in Arabidopsis thaliana. *Molecules and Cells, 32*(1), 83-88.

King, R. W., Hisamatsu, T., Goldschmidt, E. E., & Blundell, C. (2008). The nature of floral signals in Arabidopsis. I. Photosynthesis and a far-red photoresponse independently regulate flowering by increasing expression of FLOWERING LOCUS T (FT). *Journal of Experimental Botany, 59*(14), 3811-3820.

Kirby, E. J. M., Appleyard, M., & Fellowes, G. (1982). Effect of Sowing Date on the Temperature Response of Leaf Emergence and Leaf Size in Barley. *Plant Cell and Environment, 5*(6), 477-484.

Koniak, G., & Noy-Meir, I. (2009). A hierarchical, multi-scale, management-responsive model of Mediterranean vegetation dynamics. *Ecological Modelling, 220*(8), 1148-1158.

Koornneef, M., Hanhart, C. J., & Vanderveen, J. H. (1991). A Genetic and Physiological Analysis of Late Flowering Mutants in Arabidopsis-Thaliana. *Molecular & General Genetics, 229*(1), 57-66.

Krapp, A., Hofmann, B., Schafer, C., & Stitt, M. (1993). Regulation of the Expression of Rbcs and Other Photosynthetic Genes by Carbohydrates - a

Mechanism for the Sink Regulation of Photosynthesis. *Plant Journal, 3*(6), 817-828.

Kumar, S., Udawatta, R. P., & Anderson, S. H. (2010). Root length density and carbon content of agroforestry and grass buffers under grazed pasture systems in a Hapludalf. *Agroforestry Systems, 80*(1), 85-96.

Kumar, S. V., Lucyshyn, D., Jaeger, K. E., Alos, E., Alvey, E., Harberd, N. P., et al. (2012). Transcription factor PIF4 controls the thermosensory activation of flowering. *Nature, 484*(7393), 242-U127.

Lagercrantz, U. (2009). At the end of the day: a common molecular mechanism for photoperiod responses in plants? *Journal of Experimental Botany, 60*(9), 2501-2515.

Laisk, A., Eichelmann, H., & Oja, V. (2006). C-3 photosynthesis in silico. *Photosynthesis Research, 90*(1), 45-66.

Leakey, A. D. B., & Lau, J. A. (2012). Evolutionary context for understanding and manipulating plant responses to past, present and future atmospheric [CO2]. *Philosophical Transactions of the Royal Society B-Biological Sciences, 367*(1588), 613-629.

Lee, H., Yoo, S. J., Lee, J. H., Kim, W., Yoo, S. K., Fitzgerald, H., et al. (2010). Genetic framework for flowering-time regulation by ambient temperature-responsive miRNAs in Arabidopsis. *Nucleic Acids Research, 38*(9), 3081-3093.

Lee, I., & Amasino, R. (1995). Effect of Vernalization, Photoperiod, and Light Quality on the Flowering Phenotype of Arabidopsis Plants Containing the FRIGIDA Gene. *Plant Physiology, 108*(1), 157 - 162.

Lee, J. H., Yoo, S. J., Park, S. H., Hwang, I., Lee, J. S., & Ahn, J. H. (2007). Role of SVP in the control of flowering time by ambient temperature in Arabidopsis. *Genes & Development, 21*(4), 397-402.

Letort, V., Mahe, P., Cournede, P. H., De Reffye, P., & Courtois, B. (2008). Quantitative genetics and functional-structural plant growth models: Simulation of quantitative trait loci detection for model parameters and application to potential yield optimization. *Annals of Botany, 101*(8), 1243-1254.

Lewis, J. D., Wang, X. Z., Griffin, K. L., & Tissue, D. T. (2002). Effects of age and ontogeny on photosynthetic responses of a determinate annual plant to elevated CO2 concentrations. *Plant Cell and Environment, 25*(3), 359-368.

Leyser, O. (2006). Dynamic integration of auxin transport and signalling. *Current Biology, 16*(11), R424-R433.

Li, Y., Huang, Y., Bergelson, J., Nordborg, M., & Borevitz, J. O. (2010). Association mapping of local climate-sensitive quantitative trait loci in Arabidopsis thaliana. *Proceedings of the National Academy of Sciences of the United States of America, 107*(49), 21199-21204.

Li, Y., Roycewicz, P., Smith, E., & Borevitz, J. O. (2006). Genetics of Local Adaptation in the Laboratory: Flowering Time Quantitative Trait Loci under Geographic and Seasonal Conditions in Arabidopsis. *Plos One, 1*(2).

Lindenmayer, A. (1968a). Mathematical Models for Cellular Interactions in Development .2. Simple and Branching Filaments with 2-Sided Inputs. *Journal of Theoretical Biology, 18*(3), 300-&.

Lindenmayer, A. (1968b). Mathematical Models for Cellular Interactions in Development .I. Filaments with 1-Sided Inputs. *Journal of Theoretical Biology, 18*(3), 280-&.

Locke, J. C. W., Kozma-Bognar, L., Gould, P. D., Feher, B., Kevei, E., Nagy, F., et al. (2006). Experimental validation of a predicted feedback loop in the multi-oscillator clock of Arabidopsis thaliana. *Molecular Systems Biology, 2*.

Locke, J. C. W., Southern, M. M., Kozma-Bognar, L., Hibberd, V., Brown, P. E., Turner, M. S., et al. (2005). Extension of a genetic network model by iterative experimentation and mathematical analysis. *Molecular Systems Biology, 1*.

Lou, P., Xie, Q., Xu, X., Edwards, C. E., Brock, M. T., Weinig, C., et al. (2011). Genetic architecture of the circadian clock and flowering time in Brassica rapa. *Theoretical and Applied Genetics, 123*(3), 397-409.

Lu, Y., Gehan, J. P., & Sharkey, T. D. (2005). Daylength and circadian effects on starch degradation and maltose metabolism. *Plant Physiology, 138*(4), 2280-2291.

Lucas, M., Laplaze, L., & Bennett, M. J. (2011). Plant systems biology: network matters. *Plant Cell and Environment, 34*(4), 535-553.

Luterbacher, J., Liniger, M. A., Menzel, A., Estrella, N., Della-Marta, P. M., Pfister, C., et al. (2007). Exceptional European warmth of autumn 2006 and winter 2007: Historical context, the underlying dynamics, and its phenological impacts. *Geophysical Research Letters, 34*(12), -.

Mai, Y. X., Wang, L., & Yang, H. Q. (2011). A Gain-of-Function Mutation in IAA7/AXR2 Confers Late Flowering under Short-day Light in Arabidopsis. *Journal of Integrative Plant Biology, 53*(6), 480-492.

Mas, P., & Yanovsky, M. J. (2009). Time for circadian rhythms: plants get synchronized. *Current Opinion in Plant Biology, 12*(5), 574-579.

Mathieu, A., Cournede, P. H., Barthelemy, D., & De Reffye, P. (2008). Rhythms and alternating patterns in plants as emergent properties of a model of interaction between development and functioning. *Annals of Botany, 101*(8), 1233-1242.

Matsumoto, T., Wu, J. Z., Kanamori, H., Katayose, Y., Fujisawa, M., Namiki, N., et al. (2005). The map-based sequence of the rice genome. *Nature, 436*(7052), 793-800.

Mazzella, M. A., Bertero, D., & Casal, J. J. (2000). Temperature-dependent internode elongation in vegetative plants of Arabidopsis thaliana lacking phytochrome B and cryptochrome 1. *Planta, 210*(3), 497-501.

Medford, J. I., Behringer, F. J., Callos, J. D., & Feldmann, K. A. (1992). Normal and Abnormal-Development in the Arabidopsis Vegetative Shoot Apex. *Plant Cell, 4*(6), 631-643.

Mendez-Vigo, B., de Andres, M. T., Ramiro, M., Martinez-Zapater, J. M., & Alonso-Blanco, C. (2010). Temporal analysis of natural variation for the rate of leaf production and its relationship with flowering initiation in Arabidopsis thaliana. *Journal of Experimental Botany, 61*(6), 1611-1623.

Mendoza, L., & Alvarez-Buylla, E. R. (1998). Dynamics of the genetic regulatory network for *Arabidopsis thaliana* flower morphogenesis. *Journal of Theoretical Biology, 193*, 307 – 319.

Mendoza, L., & Alvarez-Buylla, E. R. (2000). Genetic regulation of root hair development in *Arabidopsis thaliana*: a network model. *Journal of Theoretical Biology, 204*, 311 – 326.

Messina, C. D., Jones, J. W., Boote, K. J., & Vallejos, C. E. (2006). A gene-based model to simulate soybean development and yield responses to environment. *Crop Science, 46*(1), 456-466.

Michael, T. P., Salome, P. A., Yu, H. J., Spencer, T. R., Sharp, E. L., McPeek, M. A., et al. (2003). Enhanced fitness conferred by naturally occurring variation in the circadian clock. *Science, 302*(5647), 1049-1053.

Michaels, S. D., & Amasino, R. M. (1999). FLOWERING LOCUS C encodes a novel MADS domain protein that acts as a repressor of flowering. *Plant Cell, 11*(5), 949-956.

Mizoguchi, T., Wheatley, K., Hanzawa, Y., Wright, L., Mizoguchi, M., Song, H. R., et al. (2002). LHY and CCA1 are partially redundant genes required to maintain circadian rhythms in Arabidopsis. *Developmental Cell, 2*(5), 629-641.

Mizoguchi, T., Wright, L., Fujiwara, S., Cremer, F., Lee, K., Onouchi, H., et al. (2005). Distinct roles of GIGANTEA in promoting flowering and regulating circadian rhythms in Arabidopsis. *Plant Cell, 17*(8), 2255-2270.

Mockler, T. C., Guo, H. W., Yang, H. Y., Duong, H., & Lin, C. T. (1999). Antagonistic actions of Arabidopsis cryptochromes and phytochrome B in the regulation of floral induction. *Development, 126*(10), 2073-2082.

Moe, R. (1990). Effect of Day and Night Temperature Alternations and of Plant-Growth Regulators on Stem Elongation and Flowering of the Long-Day Plant Campanula-Isophylla Moretti. *Scientia Horticulturae, 43*(3-4), 291-305.

Monteith, J. L. (1981). Climatic Variation and the Growth of Crops. *Quarterly Journal of the Royal Meteorological Society, 107*(454), 749-774.

Moon, J., Suh, S. S., Lee, H., Choi, K. R., Hong, C. B., Paek, N. C., et al. (2003). The SOC1 MADS-box gene integrates vernalization and gibberellin signals for flowering in Arabidopsis. *Plant Journal, 35*(5), 613-623.

Morin, X., & Thuiller, W. (2009). Comparing niche- and process-based models to reduce prediction uncertainty in species range shifts under climate change. *Ecology, 90*(5), 1301-1313.

Mundermann, L., Erasmus, Y., Lane, B., Coen, E., & Prusinkiewicz, P. (2005). Quantitative modeling of Arabidopsis development. *Plant Physiology, 139*(2), 960-968.

Mutasa-Gottgens, E., & Hedden, P. (2009). Gibberellin as a factor in floral regulatory networks. *Journal of Experimental Botany, 60*(7), 1979-1989.

Nakamichi, N., Kita, M., Ito, S., Yamashino, T., & Mizuno, T. (2005). PSEUDO-RESPONSE REGULATORS, PRR9, PRR7 and PRR5, together play essential roles close to the circadian clock of Arabidopsis thaliana. *Plant and Cell Physiology, 46*(5), 686-698.

Napp-Zinn, K. (1957). Untersuchungen iiber das Vernalisationsverhalten einer winterannuellen Rasse von Arabidopsis thaliana. *Planta, 50*, 177 - 210.

Nozue, K., Covington, M. F., Duek, P. D., Lorrain, S., Fankhauser, C., Harmer, S. L., et al. (2007). Rhythmic growth explained by coincidence between internal and external cues. *Nature, 448*(7151), 358-U311.

Ogren, E. (2000). Maintenance respiration correlates with sugar but not nitrogen concentration in dormant plants. *Physiologia Plantarum, 108*(3), 295-299.

Pantin, F., Simonneau, T., Rolland, G., Dauzat, M., & Muller, B. (2011). Control of Leaf Expansion: A Developmental Switch from Metabolics to Hydraulics. *Plant Physiology, 156*(2), 803-815.

Park, D. H., Somers, D. E., Kim, Y. S., Choy, Y. H., Lim, H. K., Soh, M. S., et al. (1999). Control of circadian rhythms and photoperiodic flowering by the Arabidopsis GIGANTEA gene. *Science, 285*(5433), 1579-1582.

Pavel, E. W., & Dejong, T. M. (1993). Source-Limited and Sink-Limited Growth Periods of Developing Peach Fruits Indicated by Relative Growth-Rate Analysis. *Journal of the American Society for Horticultural Science, 118*(6), 820-824.

Pittendrigh, C. S., & Daan, S. (1976a). Functional-Analysis of Circadian Pacemakers in Nocturnal Rodents .4. Entrainment - Pacemaker as Clock. *Journal of Comparative Physiology, 106*(3), 291-331.

Pittendrigh, C. S., & Daan, S. (1976b). Functional-Analysis of Circadian Pacemakers in Nocturnal Rodents .5. Pacemaker Structure - Clock for All Seasons. *Journal of Comparative Physiology, 106*(3), 333-355.

Poethig, R. S. (2003). Phase change and the regulation of developmental timing in plants. *Science, 301*(5631), 334-336.

Poire, R., Wiese-Klinkenberg, A., Parent, B., Mielewczik, M., Schurr, U., Tardieu, F., et al. (2010). Diel time-courses of leaf growth in monocot and dicot species: endogenous rhythms and temperature effects. *Journal of Experimental Botany, 61*(6), 1751-1759.

Pokhilko, A., Fernandez, A. P., Edwards, K. D., Southern, M. M., Halliday, K. J., & Millar, A. J. (2012). The clock gene circuit in Arabidopsis includes a repressilator with additional feedback loops. *Molecular Systems Biology, 8*.

Pokhilko, A., Hodge, S. K., Stratford, K., Knox, K., Edwards, K. D., Thomson, A. W., et al. (2010). Data assimilation constrains new connections and components in a complex, eukaryotic circadian clock model. *Molecular Systems Biology, 6*.

Pons, T. L. (2012). Interaction of temperature and irradiance effects on photosynthetic acclimation in two accessions of Arabidopsis thaliana. *Photosynthesis Research*, 1-13.

Poolman, M. G., Fell, D. A., & Thomas, S. (2000). Modelling photosynthesis and its control. *Journal of Experimental Botany, 51*, 319-328.

Pouteau, S., & Albertini, C. (2009). The significance of bolting and floral transitions as indicators of reproductive phase change in Arabidopsis. *Journal of Experimental Botany, 60*(12), 3367-3377.

Pouteau, S., Carre, I., Gaudin, V., Ferret, V., Lefebvre, D., & Wilson, M. (2008). Diversification of Photoperiodic Response Patterns in a Collection of Early-Flowering Mutants of Arabidopsis. *Plant Physiology, 148*(3), 1465-1473.

Pouteau, S., Ferret, V., & Lefebvre, D. (2006). Comparison of environmental and mutational variation in flowering time in Arabidopsis. *Journal of Experimental Botany, 57*(15), 4099-4109.

Prasad, A. K., Chai, L., Singh, R. P., & Kafatos, M. (2006). Crop yield estimation model for Iowa using remote sensing and surface parameters. *International Journal of Applied Earth Observation and Geoinformation, 8*(1), 26 – 33.

Prendergast-Miller, M., & Sohi, S. P. (2010). *Investigating biochar impacts on plant roots and root carbon*. Paper presented at the Soil Organic Matter Conference, Cote d'Azur, France.

Prusinkiewicz, P., & Hanan, J. S. (1990). Visualization of botanical structures and processes using paramateric L-systems. In D. Thanlmann (Ed.), *Scientific visualization and graphics simulation* (pp. 183 - 201). New York: John Wiley and Sons.

Prusinkiewicz, P., Karwowski, R., & Lane, B. (2007). Modelling architecture of crop plants using L-systems. In J. Vos, L. F. M. Marcelis, P. H. B. De Visser, P. C. Struik & J. B. Evers (Eds.), *Functional-structural plant modelling in crop production* (pp. 27 - 42). Dordrecht: Springer.

Pyl, E. T., Piques, M., Ivakov, A., Schulze, W., Ishihara, H., Stitt, M., et al. (2012). Metabolism and Growth in Arabidopsis Depend on the Daytime Temperature but Are Temperature-Compensated against Cool Nights. *Plant Cell, 24*(6), 2443-2469.

Qian, H. (2012). Cooperativity in Cellular Biochemical Processes: Noise-Enhanced Sensitivity, Fluctuating Enzyme, Bistability with Nonlinear Feedback, and Other Mechanisms for Sigmoidal Responses. *Annual Review of Biophysics, Vol 41, 41*, 179-204.

Qu, Z. L., Garfinkel, A., Weiss, J. N., & Nivala, M. (2011). Multi-scale modeling in biology: How to bridge the gaps between scales? *Progress in Biophysics & Molecular Biology, 107*(1), 21-31.

Raghavan, P., Bai, H., & Ghosh, S. (2004). Multi-scale model for damage analysis in fiber-reinforced composites with debonding. *Materials Processing and Design: Modeling, Simulation and Applications, Pts 1 and 2, 712*, 1911-1916.

Rasse, D. P., & Tocquin, P. (2006). Leaf carbohydrate controls over Arabidopsis growth and response to elevated CO2: an experimentally based model. *New Phytologist, 172*(3), 500-513.

Redei, J. P. (1992). A note on Columbia wild type and Landsberg *erecta*. In C. Koncz, N. H. Chua & J. Schell (Eds.), *Methods in Arabidopsis Research* (pp. 3). Singapore: World Scientific.

Reed, J. W., Nagpal, P., Poole, D. S., Furuya, M., & Chory, J. (1993). Mutations in the Gene for the Red Far-Red Light Receptor Phytochrome-B Alter Cell Elongation and Physiological-Responses Throughout Arabidopsis Development. *Plant Cell, 5*(2), 147-157.

Reinhardt, D., Pesce, E. R., Stieger, P., Mandel, T., Baltensperger, K., Bennett, M., et al. (2003). Regulation of phyllotaxis by polar auxin transport. *Nature, 426*(6964), 255-260.

Rikin, A., Dillwith, J. W., & Bergman, D. K. (1993). Correlation between the Circadian-Rhythm of Resistance to Extreme Temperatures and Changes in Fatty-Acid Composition in Cotton Seedlings. *Plant Physiology, 101*(1), 31-36.

Robertson, G. W. (1968). A Biometeorological Time Scale for a Cereal Crop Involving Day and Night Temperatures and Photoperiod. *International Journal of Biometeorology, 12*(3), 191-223.

Roden, L. C., Song, H. R., Jackson, S., Morris, K., & Carre, I. A. (2002). Floral responses to photoperiod are correlated with the timing of rhythmic expression relative to dawn and dusk in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America, 99*(20), 13313-13318.

Roldan, M., Gomez-Mena, C., Ruiz-Garcia, L., Salinas, J., & Martinez-Zapater, J. M. (1999). Sucrose availability on the aerial part of the plant promotes morphogenesis and flowering of Arabidopsis in the dark. *Plant Journal, 20*(5), 581-590.

Roycewicz, P., & Malamy, J. E. (2012). Dissecting the effects of nitrate, sucrose and osmotic potential on Arabidopsis root and shoot system growth in laboratory assays. *Philosophical Transactions of the Royal Society B-Biological Sciences, 367*(1595), 1489-1500.

Ryle, G. J. A., Arnott, R. A., & Powell, C. E. (1981). Distribution of Dry-Weight between Root and Shoot in White Clover Dependent on N-2 Fixation or Utilizing Abundant Nitrate Nitrogen. *Plant and Soil, 60*(1), 29-39.

Sablowski, R. (2007). Flowering and determinacy in Arabidopsis. *Journal of Experimental Botany, 58*(5), 899-907.

Salazar, J. D., Saithong, T., Brown, P. E., Foreman, J., Locke, J. C. W., Halliday, K. J., et al. (2009). Prediction of Photoperiodic Regulators from Quantitative Gene Circuit Models. *Cell, 139*(6), 1170-1179.

Salisbury, F. J., Hall, A., Grierson, C. S., & Halliday, K. J. (2007). Phytochrome coordinates Arabidopsis shoot and root development. *Plant Journal, 50*(3), 429-438.

Sassi, M., Lu, Y. F., Zhang, Y. H., Wang, J., Dhonukshe, P., Blilou, I., et al. (2012). COP1 mediates the coordination of root and shoot growth by light through modulation of PIN1-and PIN2-dependent auxin transport in Arabidopsis. *Development, 139*(18), 3402-3412.

Schlatter, R., Schmich, K., Vizcarra, I. A., Scheurich, P., Sauter, T., Borner, C., et al. (2009). ON/OFF and Beyond - A Boolean Model of Apoptosis. *Plos Computational Biology, 5*(12).

Schmutz, J., Cannon, S. B., Schlueter, J., Ma, J. X., Mitros, T., Nelson, W., et al. (2010). Genome sequence of the palaeopolyploid soybean. *Nature, 463*(7278), 178-183.

Schnable, P. S., Ware, D., Fulton, R. S., Stein, J. C., Wei, F. S., Pasternak, S., et al. (2009). The B73 Maize Genome: Complexity, Diversity, and Dynamics. *Science, 326*(5956), 1112-1115.

Schneider, C.A., Rasband, W.S., & Eliceiri, K.W. (2012).NIH Image to ImageJ: 25 years of image analysis. *Nature Methods*, 9, 671-675.

Schnell, S., Grima, R., & Maini, P. K. (2007). Multiscale modeling in biology - New insights into cancer illustrate how mathematical tools are enhancing the understanding of life from the smallest scale to the grandest. *American Scientist, 95*(2), 134-142.

Schwarte, S., & Tiedemann, R. (2011). A Gene Duplication/Loss Event in the Ribulose-1,5-Bisphosphate-Carboxylase/Oxygenase (Rubisco) Small Subunit Gene Family among Accessions of Arabidopsis thaliana. *Molecular Biology and Evolution, 28*(6), 1861-1876.

Schwartz, W. J., Tavakoli-Nezhad, M., Lambert, C. M., Weaver, D. R., & de la Iglesia, H. O. (2011). Distinct patterns of Period gene expression in the suprachiasmatic nucleus underlie circadian clock photoentrainment by advances or delays. *Proceedings of the National Academy of Sciences of the United States of America, 108*(41), 17219-17224.

Seo, E., Lee, H., Jeon, J., Park, H., Kim, J., Noh, Y. S., et al. (2009). Crosstalk between Cold Response and Flowering in Arabidopsis Is Mediated through the Flowering-Time Gene SOC1 and Its Upstream Negative Regulator FLC. *Plant Cell, 21*(10), 3185-3197.

Shastri, A. A., & Morgan, J. A. (2005). Flux balance analysis of photoautotrophic metabolism. *Biotechnology Progress, 21*(6), 1617-1626.

Shmulevich, I., Dougherty, E. R., Kim, S., & Zhang, W. (2002). Probabilistic Boolean networks: a rule-based uncertainty model for gene regulatory networks. *Bioinformatics, 18*(2), 261-274.

Shmulevich, I., Dougherty, E. R., & Mang, W. (2002). From Boolean to probabilistic Boolean networks as models of genetic regulatory networks. *Proceedings of the Ieee, 90*(11), 1778-1792.

Sievanen, R., Nikinmaa, E., Nygren, P., Ozier-Lafontaine, H., Perttunen, J., & Hakula, H. (2000). Components of functional-structural tree models. *Annals of Forest Science, 57*(5-6), 399-412.

Sinoquet, H., & Rivet, P. (1997). Measurement and visualization of the architecture of an adult tree based on a three-dimensional digitising device. *Trees-Structure and Function, 11*(5), 265-270.

Slafer, G. A., Connor, D. J., & Halloran, G. M. (1994). Rate of Leaf Appearance and Final Number of Leaves in Wheat - Effects of Duration and Rate of Change of Photoperiod. *Annals of Botany, 74*(5), 427-436.

Slafer, G. A., Halloran, G. M., & Connor, D. J. (1994). Development Rate in Wheat as Affected by Duration and Rate of Change of Photoperiod. *Annals of Botany, 73*(6), 671-677.

Somers, D. E., Kim, W. Y., & Geng, R. S. (2004). The F-box protein ZEITLUPE confers dosage-dependent control on the circadian clock, photomorphogenesis, and flowering time. *Plant Cell, 16*(3), 769-782.

Somers, D. E., Schultz, T. F., Milnamow, M., & Kay, S. A. (2000). ZEITLUPE encodes a novel clock-associated PAS protein from Arabidopsis. *Cell, 101*(3), 319-329.

Song, Y. H., Smith, R. W., To, B. J., Millar, A. J., & Imaizumi, T. (2012). FKF1 conveys timing information for CONSTANS stabilization in photoperiodic flowering. *Science, 336*, 1045 - 1049.

Southern, J., Pitt-Francis, J., Whiteley, J., Stokeley, D., Kobashi, H., Nobes, R., et al. (2008). Multi-scale computational modelling in biology and physiology. *Progress in Biophysics & Molecular Biology, 96*(1-3), 60-89.

Stavang, J. A., Gallego-Bartolome, J., Gomez, M. D., Yoshida, S., Asami, T., Olsen, J. E., et al. (2009). Hormonal regulation of temperature-induced growth in Arabidopsis. *Plant Journal, 60*(4), 589-601.

Stavang, J. A., Junttila, O., Moe, R., & Olsen, J. E. (2007). Differential temperature regulation of GA metabolism in light and darkness in pea. *Journal of Experimental Botany, 58*(11), 3061-3069.

Steynen, Q. J., Bolokoski, D. A., & Schultz, E. A. (2001). Alteration in flowering time causes accelerated or decelerated progression through Arabidopsis vegetative phases. *Canadian Journal of Botany-Revue Canadienne De Botanique*, *79*(6), 657-665.

Stitt, M. (1996). Metabolic regulation of photosynthesis. In N. R. Baker (Ed.), *Photosynthesis and the Environment* (pp. 635 - 640). Dordrecht, The Netherlands: Kluwer Academic.

Stitt, M., & Zeeman, S. C. (2012). Starch turnover: pathways, regulation and role in growth. *Current Opinion in Plant Biology, 15*(3), 282-292.

Stolarska, M. A., Kim, Y., & Othmer, H. G. (2009). Multi-scale models of cell and tissue dynamics. *Philosophical Transactions of the Royal Society a-Mathematical Physical and Engineering Sciences, 367*(1902), 3525-3553.

Strasser, B., Alvarez, M. J., Califano, A., & Cerdan, P. D. (2009). A complementary role for ELF3 and TFL1 in the regulation of flowering time by ambient temperature. *Plant Journal, 58*(4), 629-640.

Suarez-Lopez, P., Wheatley, K., Robson, F., Onouchi, H., Valverde, F., & Coupland, G. (2001). CONSTANS mediates between the circadian clock and the control of flowering in Arabidopsis. *Nature*, *410*(6832), 1116-1120.

Sullivan, J. A., & Deng, X. W. (2003). From seed to seed: the role of photoreceptors in Arabidopsis development. *Developmental Biology, 260*(2), 289-297.

Sulpice, R., Trenkamp, S., Steinfath, M., Usadel, B., Gibon, Y., Witucka-Wall, H., et al. (2010). Network Analysis of Enzyme Activities and Metabolite Levels and Their Relationship to Biomass in a Large Panel of Arabidopsis Accessions. *Plant Cell, 22*(8), 2872-2893.

Sun, J. D., Okita, T. W., & Edwards, G. E. (1999). Modification of carbon partitioning, photosynthetic capacity, and O-2 sensitivity in arabidopsis plants with low ADP-glucose pyrophosphorylase activity. *Plant Physiology, 119*(1), 267-276.

Sung, S. B., & Amasino, R. M. (2004). Vernalization in Arabidopsis thaliana is mediated by the PHD finger protein VIN3. *Nature, 427*(6970), 159-164.

Swarup, R., Kramer, E. M., Perry, P., Knox, K., Leyser, H. M. O., Haseloff, J., et al. (2005). Root gravitropism requires lateral root cap and epidermal cells for transport and response to a mobile auxin signal. *Nature Cell Biology, 7*(11), 1057-1065.

Swiezewski, S., Liu, F. Q., Magusin, A., & Dean, C. (2009). Cold-induced silencing by long antisense transcripts of an Arabidopsis Polycomb target. *Nature, 462*(7274), 799-U122.

Tafforeau, M., Verdus, M. C., Norris, V., Ripoll, C., & Thellier, M. (2006). Memory Processes in the Response of Plants to Environmental Signals. *Plant Signaling and Behavior, 1*(1), 9 - 14.

Tao, F. L., Yokozawa, M., Xu, Y. L., Hayashi, Y., & Zhang, Z. (2006). Climate changes and trends in phenology and yields of field crops in China, 1981-2000. *Agricultural and Forest Meteorology, 138*(1-4), 82-92.

Thimm, O., Blasing, O., Gibon, Y., Nagel, A., Meyer, S., Kruger, P., et al. (2004). MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. *Plant Journal, 37*(6), 914-939.

Thingnaes, E., Torre, S., Ernstsen, A., & Moe, R. (2003). Day and night temperature responses in Arabidopsis: Effects on gibberellin and auxin content, cell size, morphology and flowering time. *Annals of Botany, 92*(4), 601-612.

Thomas, R. (1973). Boolean Formalization of Genetic-Control Circuits. *Journal of Theoretical Biology, 42*(3), 563-585.

Thuiller, W., Araújo, M., & Lavorel, S. (2003). Generalized model vs. Classification tree analysis: Predicting spatial distributions of plant species at different scales. *Journal of Vegetation Science*, 14, 669 – 680.

Tian, T. H., & Burrage, K. (2006). Stochastic models for regulatory networks of the genetic toggle switch. *Proceedings of the National Academy of Sciences of the United States of America, 103*(22), 8372-8377.

Timm, S., Mielewczik, M., Florian, A., Frankenbach, S., Dreissen, A., Hocken, N., et al. (2012). High-to-Low CO2 Acclimation Reveals Plasticity of the Photorespiratory Pathway and Indicates Regulatory Links to Cellular Metabolism of Arabidopsis. *Plos One, 7*(8), e42809.

Tiwari, S. B., Shen, Y., Chang, H. C., Hou, Y. L., Harris, A., Ma, S. F., et al. (2010). The flowering time regulator CONSTANS is recruited to the FLOWERING LOCUS T promoter via a unique cis-element. *New Phytologist, 187*(1), 57-66.

Torii, K. U., Mitsukawa, N., Oosumi, T., Matsuura, Y., Yokoyama, R., Whittier, R. F., et al. (1996). The arabidopsis ERECTA gene encodes a putative receptor protein kinase with extracellular leucine-rich repeats. *Plant Cell, 8*(4), 735-746.

Troughton, A. (1956). Studies on the growth of young grass plants with special reference to the relatioship between the shoot and root systems. *Journal of the British Grassland Society, 6*, 56 - 65.

Twycross, J., Band, L. R., Bennett, M. J., King, J. R., & Krasnogor, N. (2010). Stochastic and deterministic multiscale models for systems biology: an auxin-transport case study. *Bmc Systems Biology, 4*.

Valverde, F., Mouradov, A., Soppe, W., Ravenscroft, D., Samach, A., & Coupland, G. (2004). Photoreceptor regulation of CONSTANS protein in photoperiodic flowering. *Science, 303*(5660), 1003-1006.

van Wijk, M. T. (2007). Predicting ecosystem functioning from plant traits: Results from a multi-scale ecophysiological modeling approach. *Ecological Modelling, 203*(3-4), 453-463.

Verdus, M. C., Ripoll, C., Norris, V., & Thellier, M. (2012). The Role of Calcium in the Recall of Stored Morphogenetic Information by Plants. *Acta Biotheoretica, 60*(1-2), 83-97.

Vos, J., Evers, J. B., Buck-Sorlin, G. H., Andrieu, B., Chelle, M., & de Visser, P. H. B. (2010). Functional-structural plant modelling: a new versatile tool in crop science. *Journal of Experimental Botany, 61*(8), 2101-2115.

Vos, J., Marcelis, L. F. M., & Evers, J. B. (2007). Functional-structural plant modelling in crop production: adding a dimension. In J. Vos, L. F. M. Marcelis, P. H. B. de Visser, P. C. Struik & J. B. Evers (Eds.), *Functional-structural plant modelling in crop production* (pp. 1 - 12). Dordrecht: Springer.

Wang, E. L., & Engel, T. (1998). Simulation of phenological development of wheat crops. *Agricultural Systems, 58*(1), 1-24.

Watanabe, T., Hanan, J. S., Room, P. M., Hasegawa, T., Nakagawa, H., & Takahashi, W. (2005). Rice morphogenesis and plant architecture: Measurement, specification and the reconstruction of structural development by 3D architectural modelling. *Annals of Botany*, *95*(7), 1131-1143.

Weinig, C., Ungerer, M. C., Dorn, L. A., Kane, N. C., Toyonaga, Y., Halldorsdottir, S. S., et al. (2002). Novel loci control variation in reproductive timing in Arabidopsis thaliana in natural environments. *Genetics, 162*(4), 1875-1884.

Weir, A. H., Bragg, P. L., Porter, J. R., & Rayner, J. H. (1984). A Winter-Wheat Crop Simulation-Model without Water or Nutrient Limitations. *Journal of Agricultural Science, 102*(Apr), 371-382.

Welch, S. M., Dong, Z. S., Roe, J. L., & Das, S. (2005). Flowering time control: gene network modelling and the link to quantitative genetics. *Australian Journal of Agricultural Research, 56*(9), 919-936.

Welch, S. M., Roe, J. L., Das, S., Dong, Z., He, R., & Kirkham, M. B. (2005). Merging genomic control networks and soil-plant-atmosphere-continuum models. *Agricultural Systems, 86*(3), 243-274.

Wenden, B., Dun, E. A., Hanan, J., Andrieu, B., Weller, J. L., Beveridge, C. A., et al. (2009). Computational analysis of flowering in pea (Pisum sativum). *New Phytologist, 184*(1), 153-167.

White, J. W., & Hoogenboom, G. (1996). Simulating effects of genes for physiological traits in a process-oriented crop model. *Agronomy Journal, 88*(3), 416-422.

Wiese, A., Christ, M. M., Virnich, O., Schurr, U., & Walter, A. (2007). Spatio-temporal leaf growth patterns of Arabidopsis thaliana and evidence for sugar control of the diel leaf growth cycle. *New Phytologist, 174*(4), 752-761.

Wilczek, A. M., Roe, J. L., Knapp, M. C., Cooper, M. D., Lopez-Gallego, C., Martin, L. J., et al. (2009). Effects of Genetic Perturbation on Seasonal Life History Plasticity. *Science, 323*(5916), 930-934.

Williams, G. D. (1974). Deriving a Biophotothermal Time Scale for Barley. *International Journal of Biometeorology, 18*(1), 57-69.

Wilson, R. N., Heckman, J. W., & Somerville, C. R. (1992). Gibberellin Is Required for Flowering in Arabidopsis-Thaliana under Short Days. *Plant Physiology, 100*(1), 403-408.

Xu, L. F., Henke, M., Zhu, J., Kurth, W., & Buck-Sorlin, G. (2011). A functional-structural model of rice linking quantitative genetic information with morphological development and physiological processes. *Annals of Botany, 107*(5), 817-828.

Yan, H. P., Kang, M. Z., de Reffye, P., & Dingkuhn, M. (2004). A dynamic, architectural plant model simulating resource-dependent growth. *Annals of Botany, 93*(5), 591-602.

Yanovsky, M. J., & Kay, S. A. (2002). Molecular basis of seasonal time measurement in Arabidopsis. *Nature, 419*(6904), 308-312.

Yazdanbakhsh, N., Sulpice, R., Graf, A., Stitt, M., & Fisahn, J. (2011). Circadian control of root elongation and C partitioning in Arabidopsis thaliana. *Plant Cell and Environment, 34*(6), 877-894.

Yee, T. W., & Mitchell, N. D. (1991). Generalized additive models in plant ecology. *Journal of Vegetation Science*, 2, 587 – 602.

Yin, X. Y., Chasalow, S. D., Dourleijn, C. J., Stam, P., & Kropff, M. J. (2000). Coupling estimated effects of QTLs for physiological traits to a crop growth model: predicting yield variation among recombinant inbred lines in barley. *Heredity, 85*(6), 539-549.

Yin, X. Y., Goudriaan, J., Lantinga, E. A., Vos, J., & Spiertz, H. J. (2003). A flexible sigmoid function of determinate growth. *Annals of Botany, 91*(3), 361-371.

Yin, X. Y., Kropff, M. J., & Goudriaan, J. (1996). Differential effects of day and night temperature on development to flowering in rice. *Annals of Botany, 77*(3), 203-213.

Yin, X. Y., Kropff, M. J., Goudriaan, J., & Stam, P. (2000). A model analysis of yield differences among recombinant inbred lines in barley. *Agronomy Journal, 92*(1), 114-120.

Yin, X. Y., Stam, P., Kropff, M. J., & Schapendonk, A. H. C. M. (2003). Crop modeling, QTL mapping, and their complementary role in plant breeding. *Agronomy Journal, 95*(1), 90-98.

Yin, X. Y., & Struik, P. C. (2010). Modelling the crop: from system dynamics to systems biology. *Journal of Experimental Botany, 61*(8), 2171-2183.

Yu, J., Hu, S. N., Wang, J., Wong, G. K. S., Li, S. G., Liu, B., et al. (2002). A draft sequence of the rice genome (Oryza sativa L. ssp indica). *Science, 296*(5565), 79-92.

Yu, J. W., Rubio, V., Lee, N. Y., Bai, S. L., Lee, S. Y., Kim, S. S., et al. (2008). COP1 and ELF3 Control Circadian Function and Photoperiodic Flowering by Regulating GI Stability. *Molecular Cell, 32*(5), 617-630.

Appendices

Appendix A

A1. Parameter values of the Wilczek et al. model and Model 2 (step gating) for the Ler lines

Parameter	**Wilczek et al. Model[1]**	**Model 2 (step gating)**
Threshold MPTU	2392	2907
Non-vernalising development		
T_b	3	
CSDL	10	
CLDL	14	
D_{LD}	1	
D_{SD}	0.6626	0.6365
P_{day}	1	1 (set *a priori*)
P_{night}	0	0.2118
Vernalisation submodel parameters		
T_{Vmin}	-3.5 ^{0}C	
T_{Vmax}	6 ^{0}C	
V_{sat}	960	
F_b	0.3756	0.4259
κ	-5.17	
ω	2.23	
ξ	1.00	

Adjustments in Model 2 for *phyA-201* and *phyB-1* mutants (in the Ler background):

Parameter	***phyA-201***	***phyB-1***
CSDL	11	10.5
CLDL	15	16
D_{LD}	No change	1.350
D_{SD}	No change	0.8084
P_{day}	No change	0.6279
P_{night}	No change	0.6279

[1] For consistency, the parameter values for the Wilczek et al. model listed here were re-optimised using the global optimisation tool in MATLAB (Mathworks, Cambridge, UK) in contrast to the Excel Solver used in Wilczek et al. (2009).

A2. Parameter values of the Wilczek et al. model and Model 2 (step gating) for the Col lines

Parameter		**Wilczek et al. Model[1]**	**Model 2 (step gating)[2]**
Threshold MPTU		2604	3212
Non-vernalising development			
T_b		3	
CSDL		10	
CLDL		14	
D_{LD}		1	
D_{SD}		0.6015	0.6015
P_{day}		1	1 (set *a priori*)
P_{night}		0	0.1782
Vernalisation submodel parameters			
T_{Vmin}		-3.5 ^{0}C	
T_{Vmax}		6 ^{0}C	
V_{sat}		960	
F_b	Col fri	0.4075	0.4743
	Col-FRI-Sf2	0.2635	0.3176
	fve	0.2722	0.3307
κ		-5.17	
ω		2.23	
ξ		1.00	

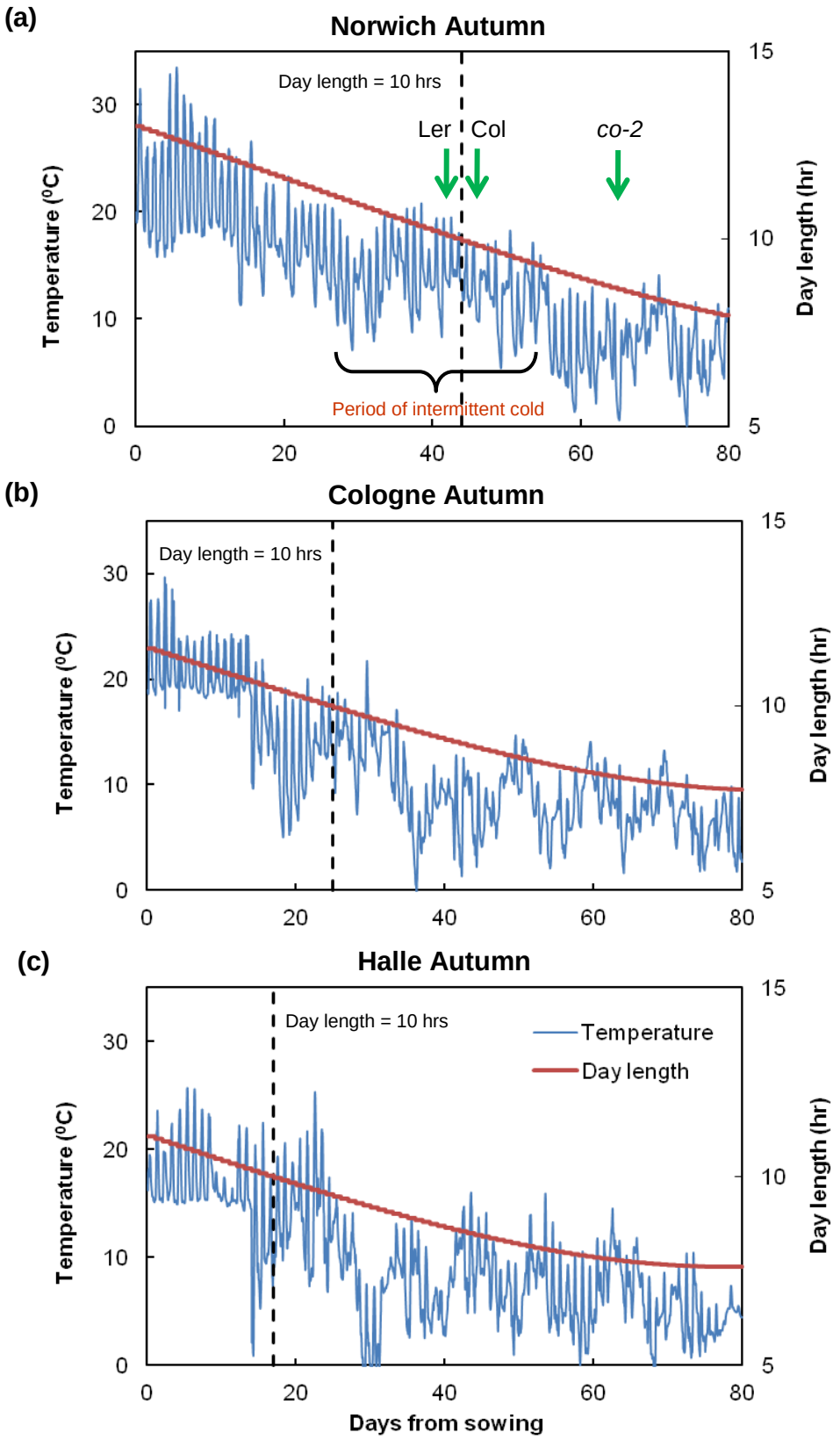
1 For consistency, the parameter values for the Wilczek et al. model listed here were re-optimised using the global optimisation tool in MATLAB (Mathworks, Cambridge, UK) in contrast to the Excel Solver used in Wilczek et al. (2009).

2The parameter values listed here were estimated without the four late-flowering genotypes (*gi-2*, Col-FRI-Sf2, *vin3-1* and *fve-3*) from Norwich Autumn.

A3. Field data of Ler *wt*, *phyA-201* and *phyB-1* mutants

Plantings	Ler *wt*			*phyA-201*			*phyB-1*		
	Mean DTB	Standard error	Number of plants	Mean DTB	Standard error	Number of plants	Mean DTB	Standard error	Number of plants
Norwich Summer ‘06	21	0.40	5	22	0.70	8	22	1.11	4
Oulu Summer	18	0.44	24	19	0.43	23	19	0.55	20
Norwich Autumn	42	2.83	11	48	4.16	12	37	1.92	11
Cologne Autumn	63	5.47	10	79	4.87	13	42	2.20	12
Halle Autumn	101	2.46	17	104	2.09	17	69	3.08	16
Valencia Autumn	75	5.05	14	72	4.42	13	52	6.32	10
Norwich Spring	50	0.32	14	54	0.60	12	48	1.10	14
Cologne Spring	48	0.18	9	49	0.21	6	48	2.00	7
Norwich Summer ‘07	22	0.74	14	22	0.83	11	20	0.36	10

A4. Leaf-level temperature and day length experienced by the plants at different sites in the autumn.



In these experiments, plants were kept in the greenhouse for 2 weeks before being transplanted to the field. The black dashed lines represent the critical point when day length decreased below 10 hours (CSDL). The green arrows indicate the bolting times of Ler, *co-2* and Col at Norwich Autumn.

Appendix B

B1. The ODEs of the clock-photoperiod circuit model (Model 3 in Salazar et al. (2009)).

$$\frac{d[\mathrm{PIF3}]}{dt} = 0.5\times\left(1-\Theta_{light}\right)-[\mathrm{PIF3}]\times\Theta_{light}-\frac{1.2\times[\mathrm{PIF3}]}{1.2+[\mathrm{PIF3}]}. \tag{B2.1}$$

$$\frac{d[LHY]}{dt} = \Theta_{light}\times v1\times[\mathrm{PIF3}]+\frac{v2\times[\mathrm{X}_c]^{mu}}{v3+[\mathrm{X}_c]^{mu}}-\frac{v4\times[LHY]}{v5+[LHY]}. \tag{B2.2}$$

$$\frac{d[\mathrm{LHY}_c]}{dt} = v6\times[LHY]-v7\times[\mathrm{LHY}_c]+v8\times[\mathrm{LHY}_n]-\frac{v9\times[\mathrm{LHY}_c]}{v10+[\mathrm{LHY}_c]}. \tag{B2.3}$$

$$\frac{d[\mathrm{LHY}_n]}{dt} = v7\times[\mathrm{LHY}_c]-v8\times[\mathrm{LHY}_n]-\frac{v11\times[\mathrm{LHY}_n]}{v12+[\mathrm{LHY}_n]}. \tag{B2.4}$$

$$\frac{d[TOC1]}{dt} = \frac{v13\times[\mathrm{Y}_n]^{nu}}{v14+[\mathrm{Y}_n]^{nu}}\times\frac{v15}{v16+[Y_n]^{nu}}-\frac{v17\times[TOC1]}{v18+[TOC1]}. \tag{B2.5}$$

$$\frac{d[\mathrm{TOC1}_c]}{dt} = v19\times[TOC1]-v20\times[\mathrm{TOC1}_c]+v21\times[\mathrm{TOC1}_n] - \frac{\left(v22\times\left(1-\Theta_{light}\right)+v23\right)\times[\mathrm{TOC1}_c]}{v24+[\mathrm{TOC1}_c]}. \tag{B2.6}$$

$$\frac{d[\mathrm{TOC1}_n]}{dt} = v20\times[\mathrm{TOC1}_c]-v21\times[\mathrm{TOC1}_n]-\frac{\left(v25\times\left(1-\Theta_{light}\right)+v26\right)\times[\mathrm{TOC1}_n]}{v27+[\mathrm{TOC1}_n]}. \tag{B2.7}$$

$$\frac{d[X]}{dt} = \frac{v28\times[\mathrm{TOC1}_n]^{xi}}{v29+[\mathrm{TOC1}_n]^{xi}}-\frac{v30\times[X]}{v31+[X]}. \tag{B2.8}$$

$$\frac{d[\mathrm{X}_c]}{dt} = v32 \times [X] - v33 \times [\mathrm{X}_c] + v34 \times [\mathrm{X}_n] - \frac{v35 \times [\mathrm{X}_c]}{v36 + [\mathrm{X}_c]}. \quad \text{(B2.9)}$$

$$\frac{d[\mathrm{X}_n]}{dt} = v33 \times [\mathrm{X}_c] - v34 \times [\mathrm{X}_n] - \frac{v37 \times [\mathrm{X}_n]}{v38 + [\mathrm{X}_n]}. \quad \text{(B2.10)}$$

$$\frac{d[Y]}{dt} = \left(\Theta_{light} \times v53 \times [\mathrm{PIF3}] + \frac{\Theta_{light} \times v39 + v40}{v41 + [\mathrm{TOC1}_n]^{ytoc}} \right) \times \frac{v51}{[\mathrm{LHY}_n]^{ylhy} + v52} - \frac{v42 \times [Y]}{v43 + [Y]}. \quad \text{(B2.11)}$$

$$\frac{d[\mathrm{Y}_c]}{dt} = v44 \times [Y] - v45 \times [\mathrm{Y}_c] + v46 \times [\mathrm{Y}_n] - \frac{v47 \times [\mathrm{Y}_c]}{v48 + [\mathrm{Y}_c]}. \quad \text{(B2.12)}$$

$$\frac{d[\mathrm{Y}_n]}{dt} = v45 \times [\mathrm{Y}_c] - v46 \times [\mathrm{Y}_n] - \frac{v49 \times [\mathrm{Y}_n]}{v50 + [\mathrm{Y}_n]}. \quad \text{(B2.13)}$$

Assuming that *CO* mRNA is the same as TOC1 nuclear protein, and *FT* mRNA is activated by CO protein:

$$\frac{d[\mathrm{CO}]}{dt} = vCOm \times [\mathrm{TOC1}_n] - \frac{(1 - \Theta_{light}) \times vCOp \times [\mathrm{CO}]}{kCOp + [\mathrm{CO}]}. \quad \text{(B2.14)}$$

$$\frac{d[FT]}{dt} = Bco + \frac{VCO \times [\mathrm{CO}]}{KCO + [\mathrm{CO}]} - \frac{vFT \times [FT]}{kFT + [FT]}. \quad \text{(B2.15)}$$

$$\Theta_{light} = \frac{1}{4} \times (1 + \tanh(6 \times (t - sunrise))) \times (1 - \tanh(6 \times (t - sunset))). \quad \text{(B2.16)}$$

The concentration (indicated by [ ]) of each component is written in italics to represent the corresponding mRNA, or in non-italics to represent the protein form. The subscripts *c* (cytoplasm) and *n* (nucleus) indicate the corresponding compartment.

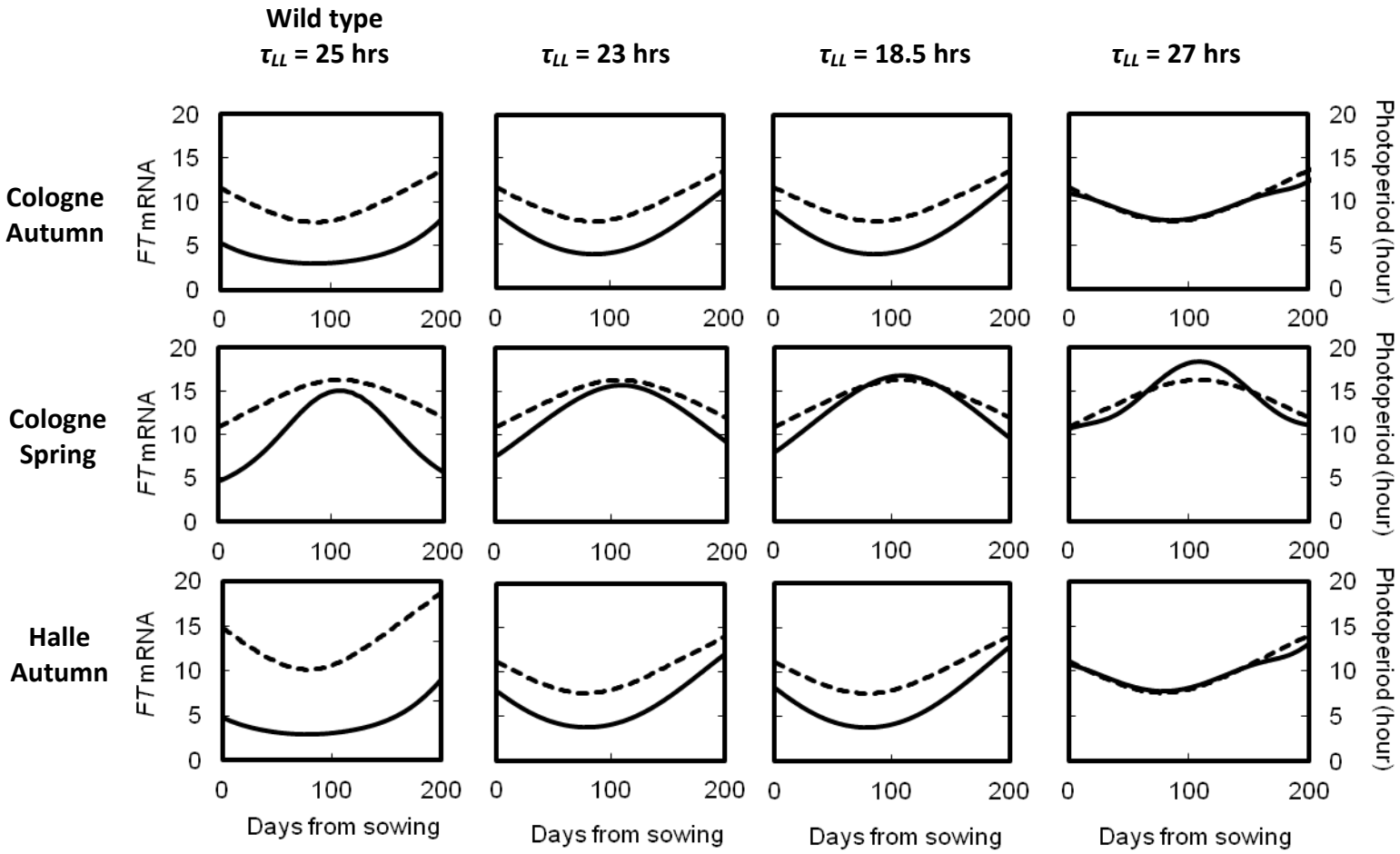
B2. Parameter values of the Salazar et al. model

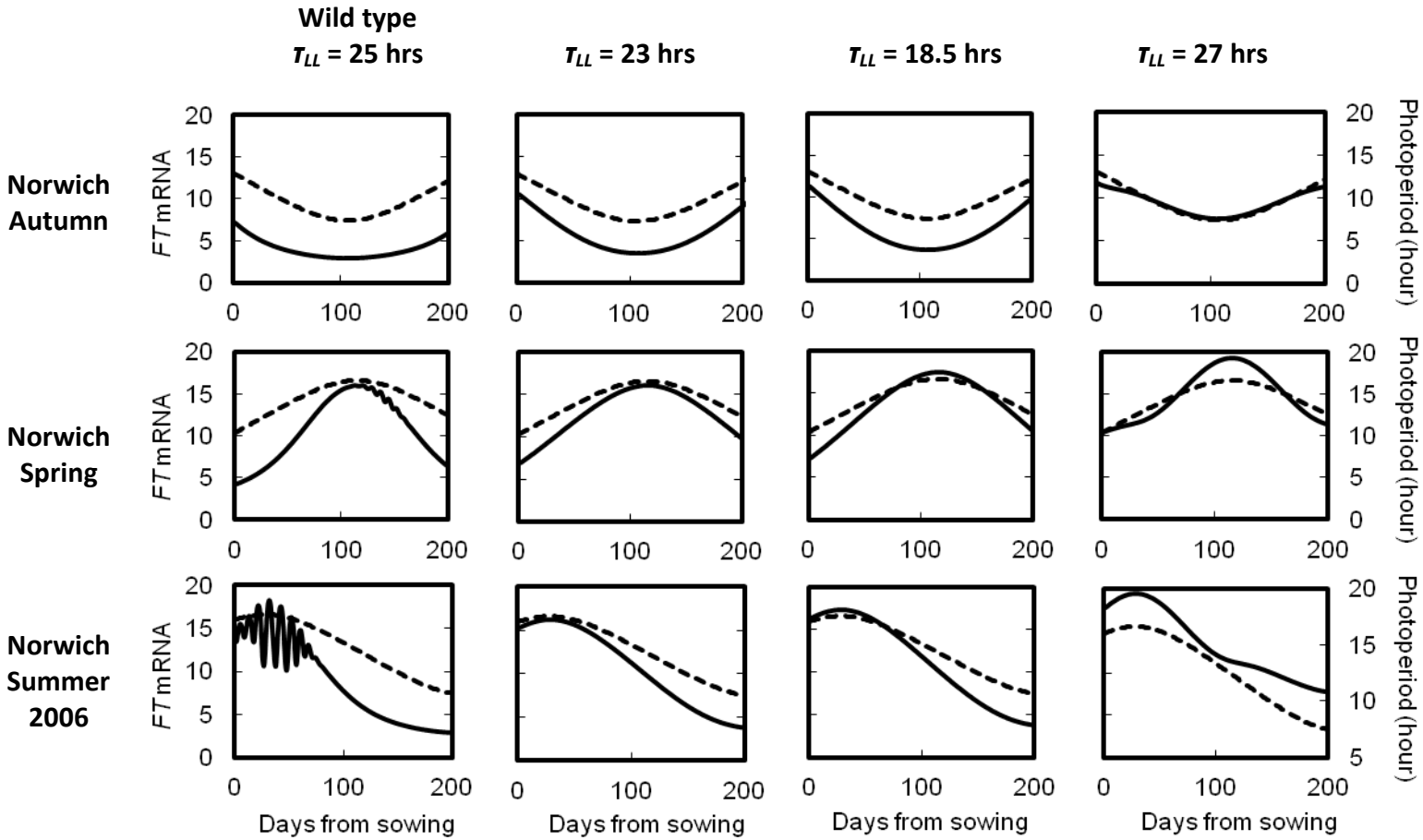
Symbol	Parameter value	Unit	Parameter description
$v1$	2.4514	h^{-1}	Coupling constant of light activation of *LHY* transcription
$v2$	5.1694	nM h^{-1}	Maximum rate of light-independent *LHY* transcription
$v3$	0.6473	nM	Constant of activation by protein X
$v4$	1.5283	nM h^{-1}	Maximum rate of *LHY* mRNA degradation
$v5$	1.8170	nM	Michaelis constant of *LHY* mRNA degradation
$v6$	0.8295	h^{-1}	Rate constant of *LHY* mRNA translation
$v7$	16.8363	h^{-1}	Rate constant of LHY transport into the nucleus
$v8$	0.1687	h^{-1}	Rate constant of LHY transport out of the nucleus
$v9$	20.4400	nM h^{-1}	Maximum rate of cytoplasmic LHY degradation
$v10$	1.5644	nM	Michaelis constant of cytoplasmic LHY degradation
$v11$	3.6888	nM h^{-1}	Maximum rate of nuclear LHY degradation
$v12$	1.2765	nM	Michaelis constant of nuclear LHY degradation
$v13$	1.3956	nM h^{-1}	Maximum activation by protein Y
$v14$	0.0338	nM	Constant of activation by protein Y
$v15$	0.5539	nM^{-1}	Maximum rate of *TOC*1 transcription
$v16$	0.2569	nM	Constant of repression by LHY
$v17$	3.8231	nM h^{-1}	Maximum rate of *TOC*1 mRNA degradation
$v18$	2.5734	nM	Michaelis constant of *TOC*1 mRNA degradation
$v19$	4.3240	h^{-1}	Rate constant of *TOC*1 mRNA translation
$v20$	0.3166	h^{-1}	Rate constant of TOC1 movement into the nucleus
$v21$	2.1509	h^{-1}	Rate constant of TOC1 movement out of the nucleus
$v22$	0.0013	nM h^{-1}	Maximum rate of light-dependent cytoplasmic TOC1 degradation

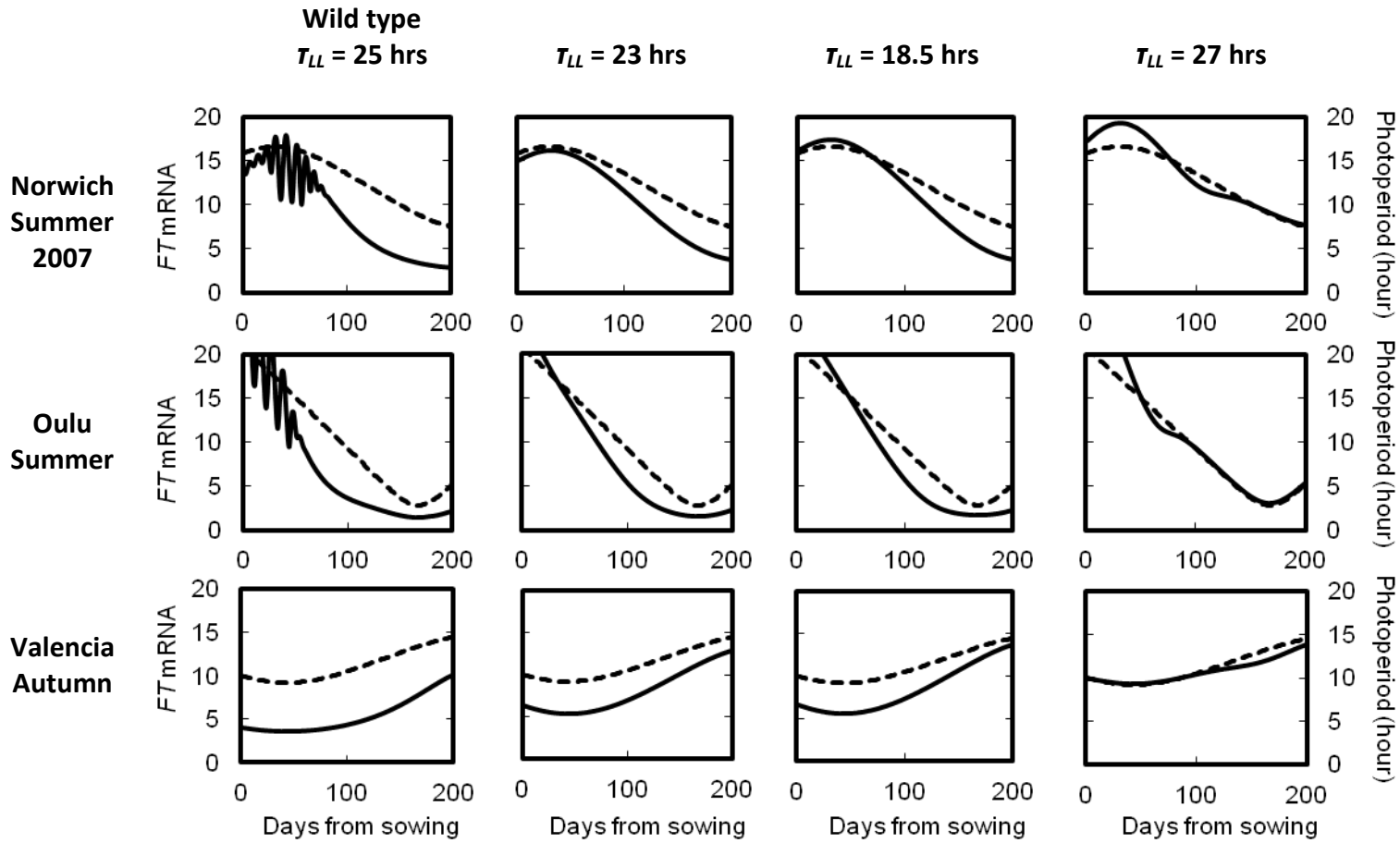
Symbol	Parameter value	Unit	Parameter description
*v*23	3.1741	nM h^{-1}	Maximum rate of light-independent cytoplasmic TOC1 degradation
*v*24	2.7454	nM	Michaelis constant of cytoplasmic TOC1 degradation
*v*25	0.0492	nM h^{-1}	Maximum rate of light-dependent nuclear TOC1 degradation
*v*26	4.0424	nM h^{-1}	Maximum rate of light-independent nuclear TOC1 degradation
*v*27	0.4033	nM	Michaelis constant of nuclear TOC1 degradation
*v*28	0.2431	nM h^{-1}	Maximum transcription rate of protein X
*v*29	0.4099	nM	Constant of activation by TOC1 protein
*v*30	10.1132	nM h^{-1}	Maximum rate of degradation of *X* mRNA
*v*31	6.5585	nM	Michaelis constant of *X* mRNA degradation
*v*32	2.1470	h^{-1}	Rate constant of *X* mRNA translation
*v*33	1.0352	h^{-1}	Rate constant of protein X movement into the nucleus
*v*34	3.3017	h^{-1}	Rate constant of protein X movement out of the nucleus
*v*35	0.2179	nM h^{-1}	Maximum rate of degradation of cytoplasmic protein X
*v*36	0.6632	nM	Michaelis constant of cytoplasmic protein X degradation
*v*37	3.3442	nM h^{-1}	Maximum rate of degradation of nuclear protein X
*v*38	17.111	nM	Michaelis constant of nuclear protein X degradation
*v*39	1.8272	nM h^{-1}	Light-dependent component of Y transcription
*v*40	3.5159	nM h^{-1}	Light-independent component of Y transcription
*v*41	1.8056	nM	Constant of repression by TOC1
*v*42	4.2970	nM h^{-1}	Maximum rate of degradation of *Y* mRNA
*v*43	1.7303	nM	Michaelis constant of *Y* mRNA degradation

Symbol	Parameter value	Unit	Parameter description
*v*44	0.2485	h^{-1}	Rate constant of *Y* mRNA translation
*v*45	2.2123	h^{-1}	Rate constant of protein Y movement into the nucleus
*v*46	0.2002	h^{-1}	Rate constant of protein Y movement out of the nucleus
*v*47	0.1347	nM h^{-1}	Maximum rate of degradation of cytoplasmic protein Y
*v*48	1.8258	nM	Michaelis constant of cytoplasmic protein Y degradation
*v*49	0.6114	nM h^{-1}	Maximum rate of degradation of nuclear protein Y
*v*50	1.8066	nM	Michaelis constant of nuclear protein Y degradation
*v*51	0.0051	nM h^{-1}	Maximum rate of Y transcription
*v*52	0.0604	nM	Constant of repression by LHY
*v*53	28.3562	h^{-1}	Coupling constant of light activation of *Y* mRNA transcription
nu	1.0258		Hill coefficient of activation by protein Y
mu	3.3064		Hill coefficient of activation by protein X
xi	1.4422		Hill coefficient of activation by TOC1
ylhy	1.0237		Hill coefficient of repression by LHY
ytoc	3.6064		Hill coefficient of repression by TOC1
vCOm	1.1452	h^{-1}	Rate constant of *CO* (or *TOC*1) mRNA translation
vCOp	9.2242	nM h^{-1}	Maximum rate of light-dependent CO protein degradation
kCOp	2.0976	nM	Michaelis constant of CO protein degradation
VCO	0.5800	nM h^{-1}	Maximum rate of *FT* activation by CO
KCO	7.3533	nM	Michaelis constant of *FT* activation by CO
vFT	1.8974	nM h^{-1}	Maximum rate of *FT* mRNA degradation
kFT	5.3925	nM	Michaelis constant of *FT* mRNA degradation
Bco	0	nM h^{-1}	Basal rate of *FT* transcription

B3. Integral of simulated *FT* mRNA level across the season




Wild type
τ_{LL} = 25 hrs
τ_{LL} = 23 hrs
τ_{LL} = 18.5 hrs
τ_{LL} = 27 hrs
Norwich Autumn
Norwich Spring
Norwich Summer 2006
FT mRNA
Photoperiod (hour)
Days from sowing
0
5
10
15
20
100
200


Wild type
τ_{LL} = 25 hrs
τ_{LL} = 23 hrs
τ_{LL} = 18.5 hrs
τ_{LL} = 27 hrs
Norwich Summer 2007
Oulu Summer
Valencia Autumn
FT mRNA
Photoperiod (hour)
Days from sowing
0
5
10
15
20
100
200

Appendix C

C1. Equations of the carbon assimilation and partitioning model (Rasse & Tocquin, 2006)

Carbon assimilation:

The following equations represent the biochemical model of photosynthetic CO_2 assimilation in C_3 plants developed by Farquhar et al. (1980). In general, this model considers two rate limiting factors in the Calvin-Benson cycle: (1) RuBisCO carboxylation; and (2) electron transport. For more information on the model, please refer to Farquhar et al. (1980).

RuBisCO activity:

The parameter values of kinetic constants at leaf temperature, T_{leaf}, are first determined using values at the standard temperature 25^0 C as follows:

$$K_c(T_{leaf}) = K_c(25^0\text{C}) \times \exp\left(\frac{H_c \times (T_{leaf} - 25)}{298 \times R \times (T_{leaf} + 273)}\right). \quad \text{(C1.1)}$$

$$K_{o_2}(T_{leaf}) = K_{o_2}(25^0\text{C}) \times \exp\left(\frac{H_{o_2} \times (T_{leaf} - 25)}{298 \times R \times (T_{leaf} + 273)}\right). \quad \text{(C1.2)}$$

$$V_{c_{\max}}(T_{leaf}) = V_{c_{\max}}(25^0\text{C}) \times \exp\left(\frac{H_V \times (T_{leaf} - 25)}{298 \times R \times (T_{leaf} + 273)}\right). \quad \text{(C1.3)}$$

Rate of assimilation limited by RuBisCO, A_c, is then calculated as:

$$A_c = V_{c_{\max}}(T_{leaf}) \times \frac{[\mathrm{CO_2}]_i - \Gamma^*}{[\mathrm{CO_2}]_i + K_c(T_{leaf}) \times \left(1 + \frac{[\mathrm{O_2}]}{K_{o_2}(T_{leaf})}\right)}, \quad \text{(C1.4)}$$

where $[O_2]$ is the O_2 partial pressure having a constant value of 20500 Pa, while $[CO_2]_i$ is the intercellular CO_2 partial pressure given by

$$[\mathrm{CO_2}]_i = p1 \times [\mathrm{CO_2}]. \quad \text{(C1.5)}$$

The CO_2 compensation point in the absence of mitochondrial respiration, Γ^*, is determined as

$$\Gamma^* = p2 + p3 \times (T_{leaf} - 25) + p4 \times (T_{leaf} - 25)^2. \quad \text{(C1.6)}$$

Electron transport:

The potential rate of electron transport at leaf temperature, J_{max} (T_{leaf}), is determined using the following equation:

$$J_{\max}(T_{leaf}) = J_{\max}(25^0\mathrm{C}) \times \frac{\exp\left(\frac{H_J \times (T_{leaf} - 25)}{298 \times R \times (T_{leaf} + 273)}\right) \times \left(1 + \exp\left(\frac{298 \times p5 - p6}{298 \times R}\right)\right)}{1 + \exp\left(\frac{(T_{leaf} + 273) \times p5 - p6}{(T_{leaf} + 273) \times R}\right)}. \quad \text{(C1.7)}$$

$J_{\max}(25^0\,\mathrm{C})$ can be measured experimentally, but in Rasse and Tocquin (2006) and in the current study, this parameter is estimated as

$$J_{\max}(25^0\mathrm{C}) = R_{JV} \times V_{c_{\max}}(25^0\mathrm{C}). \quad \text{(C1.8)}$$

The rate of electron transport, *J*, depends on the irradiance *PAR* (μmol photon m^{-2} s^{-1}) and is calculated by solving the following quadratic function:

$$p7 \times J^2 - \left(\frac{PAR \times (1 - f_{spec})}{2} + J_{\max}(T_{leaf}) \right) \times J + \frac{PAR \times (1 - f_{spec})}{2} \times J_{\max}(T_{leaf}) = 0 . \quad \text{(C1.9)}$$

Rate of assimilation limited by electron transport, A_j, is determined as:

$$A_j = \frac{J \times ([CO_2]_i - \Gamma *)}{4 \times ([CO_2]_i - 2 \times \Gamma *)} . \quad \text{(C1.10)}$$

Net rate of carbon assimilation:

The net rate of leaf photosynthesis per unit area, A_{net} (μmol CO_2 m^{-2} s^{-1}), is given by the minimum of the two rate limiting factors described earlier. Therefore:

$$A_{net} = \mathrm{Min}\{A_c, A_j\}. \quad \text{(C1.11)}$$

In the case where no solution can be found for equation C1.9, A_{net} takes the value of A_c.

The total amount of carbon assimilated per plant (in g C) in each hour, which is the time step used in this study, is calculated using conversion factors as follows:

Carbon assimilation (t) = 12 x 10^{-6} x 3600 x A_{net} (t) x Rosette area (t-1). (C1.12)

The projected rosette area at the end of the previous time point is used as the total area available for light interception.

Carbon partitioning:

The following equations represent the mechanistic allocation and growth model for Arabidopsis developed by Rasse and Tocquin (2006). This model considers four carbon pools (Fig. 4.3.1): (1) Starch carbon; (2) Sugar carbon; (3) Leaf carbon; and (4) Root carbon. Please refer to Rasse and Tocquin (2006) for more information on model assumptions and calibrated parameters.

Starch-sugar partitioning:

At day time, a fixed baseline portion of carbon assimilates is stored as transitory starch and the rest is turned into sugar. Therefore:

Starch synthesis (t) = ST_{br} x Carbon assimilation (t) . (C1.13)

Partition to sugar (t) = Carbon assimilation (t) – Starch synthesis (t). (C1.14)

At night time when there is no photosynthesis, a fixed portion of the starch accumulated at the end of the day is degraded at a linear rate to form sugar, as follows:

$$\text{Starch degradation } (t) = \frac{ST_c \times \text{End} - \text{of} - \text{day starch}}{\text{Night length}}, \quad \text{(C1.15)}$$

where night length is measured in hour.

Maintenance respiration:

Carbon is constantly respired for plant maintenance, and leaf maintenance respiration per unit area, R_l, at 20^0 C was found to have the following linear correlation with sugar content (Rasse & Tocquin, 2006):

$$R_l(20^0\text{C}) = p8 \times \text{Sugar content per unit rosette area } (t\text{-}1) + p9. \quad \text{(C1.16)}$$

The amount of carbon respired for maintenance in each corresponding time step is estimated from the sugar content at the end of the previous time step. If sugar content is 0, R_l is set to 0.

The above-ground and below-ground maintenance respirations at leaf temperature are computed as:

$$R_l(T_{leaf}) = R_l(20^0\text{C}) \times \exp\left(\frac{H_r \times (T_{leaf} - 20)}{293 \times R \times (T_{leaf} + 273)}\right). \quad \text{(C1.17)}$$

Total leaf maintenance respiration:

$$R_{above}(t) = R_l(T_{leaf}) \text{ x Rosette area } (t\text{-}1). \quad \text{(C1.18)}$$

Total root maintenance respiration:

$$R_{below}(t) = R_{above}(t) \times \frac{\text{Root carbon } (t-1)}{\text{Leaf carbon } (t-1)}. \quad \text{(C1.19)}$$

Carbon available for growth:

The transient amount of carbon available for growth, Q_{trans}, is the excess from the sugar pool after balancing all the carbon fluxes in the corresponding time step as follows:

$$Q_{trans}(t) = \text{Sugar carbon } (t\text{-}1) + S(\text{t}) - R_{above}(t) - R_{below}(t), \quad \text{(C1.20)}$$

where

$$S(t) = \begin{cases} \text{Partition to sugar}(t), & sunrise < t \le sunset; \\ \text{Starch degradation}(t), & \text{otherwise.} \end{cases} \quad \text{(C1.21)}$$

However, a minimum sugar content of $SSU_{\min}$ (g sugar-C m^{-2}) has to be maintained. Therefore carbon availability, C_{avail}, is calculated as:

$$C_{avail}(t) = Q_{trans}(t) - SSU_{\min} \text{ x Rosette area } (t\text{-}1). \quad \text{(C1.22)}$$

If C_{avail} has a negative value, no carbon is available for growth in that time step. Instead, carbon is translocated from the leaf and root carbon pools to the sugar pool to maintain its minimum level. As such:

Component	**Value used**		
	$\boldsymbol{C_{avail}(t) \ge 0}$	$\boldsymbol{C_{avail}(t) < 0}$	
Carbon available for growth, $Q_c(t)$	$C_{avail}(t)$	0	(C1.23)
Translocation from leaf carbon, $TransL(t)$	0	$-C_{avail}(t) \times \frac{\text{Leaf carbon } (t\text{-}1)}{\text{Total}\,(t\text{-}1)}$	(C1.24)
Translocation from root carbon, $TransR(t)$	0	$-C_{avail}(t) \times \frac{\text{Root carbon } (t\text{-}1)}{\text{Total}\,(t\text{-}1)}$	(C1.25)

where

$$\text{Total } (t\text{-}1) = \text{Leaf carbon } (t\text{-}1) + \text{Root carbon } (t\text{-}1). \quad \text{(C1.26)}$$

Organ growth demand:

Leaf (or rosette) growth demand per time step, d_L (in g C per hour), is limited by a maximum daily relative growth rate, GR_{max} (g g^{-1} d^{-1}), as shown in the following:

$$d_L\,(t) = GR_{\max} \times \text{Leaf carbon } (t\text{-}1) \times \frac{1}{24}. \qquad \text{(C1.27)}$$

Note: In the Rasse and Tocquin model, the whole rosette is treated as one big leaf without consideration of each individual leaf.

A fraction of the total growth demand is used for growth respiration, therefore the required leaf growth respiration, d_{RL}, is:

$$d_{RL}\,(t) = \frac{d_L(t) \times \alpha}{1-\alpha}. \qquad \text{(C1.28)}$$

Root growth demand, d_R, is calculated from the leaf growth demand using the root-to-shoot allocation ratio, *RS*, as in the following:

$$RS(t) = \begin{cases} 0.12, & \text{Rosette area } (t-1) < 1.22e-4; \\ 0.0496 + 555 \times \text{Rosette area } (t-1), & 1.22e\text{-}4 < \text{Rosette area } (t-1) < 3.96e\text{-}4; \\ 0.27, & \text{otherwise.} \end{cases} \qquad \text{(C1.29)}$$

$$d_R\,(t) = RS(t) \times d_L\,(t). \qquad \text{(C1.30)}$$

Note: In the current study, *RS* in equation C1.29 has been replaced by the ratio of root sink demand to shoot sink demand (equation 4.3.17).

Similarly, the required root growth respiration, d_{RR}, can be calculated as:

$$d_{RR}(t) = \frac{d_R(t) \times \alpha}{1-\alpha}. \quad \text{(C1.31)}$$

Therefore, the total growth demand for carbon, D_c, is given by:

$$D_c(t) = d_L(t) + d_{RL}(t) + d_R(t) + d_{RR}(t). \quad \text{(C1.32)}$$

Allocation:

The actual growth depends on whether the total growth demand is met by the carbon available for growth, as shown below:

Growth component	**Value used**		
	Full demand is met $D_c(t) \le Q_c(t)$	**Full demand is not met** $D_c(t) > Q_c(t)$	
Leaf growth, $L(t)$	$d_L(t)$	$\frac{d_L(t)}{D_c(t)} \times Q_c(t)$	(C1.33)
Leaf growth respiration, $R_{ml}(t)$	$d_{RL}(t)$	$\frac{d_{RL}(t)}{D_c(t)} \times Q_c(t)$	(C1.34)
Root growth, $R(t)$	$d_R(t)$	$\frac{d_R(t)}{D_c(t)} \times Q_c(t)$	(C1.35)
Root growth respiration, $R_{mr}(t)$	$d_{RR}(t)$	$\frac{d_{RR}(t)}{D_c(t)} \times Q_c(t)$	(C1.36)

In the case where the full demand is met, any excess carbon is transferred to the starch pool in the model to represent starch production through an overflow mechanism, O_{sta}. This overflow starch production is constrained to the light period.

$$O_{sta}(t) = \begin{cases} Q_c(t) - D_c(t), & sunrise < t \le sunset; \\ 0, & \text{otherwise.} \end{cases} \quad \text{(C1.37)}$$

Amount of carbon in each pool:

Leaf carbon (t) = Leaf carbon (t-1) + $L(t)$ – $TransL(t)$. (C1.38)

Root carbon (t) = Root carbon (t-1) + $R(t)$ – $TransR(t)$. (C1.39)

Starch carbon (t) = Starch carbon (t-1) + Starch synthesis (t) –
Starch degradation (t) + $O_{sta}(t)$. (C1.40)

Sugar carbon (t) = Sugar carbon (t-1) + $S(t) - R_{above}(t) - R_{below}(t) - R_{ml}(t) - R_{mr}(t) - O_{sta}(t) - L(t) - R(t) + TransL(t) + TransR(t)$. (C1.41)

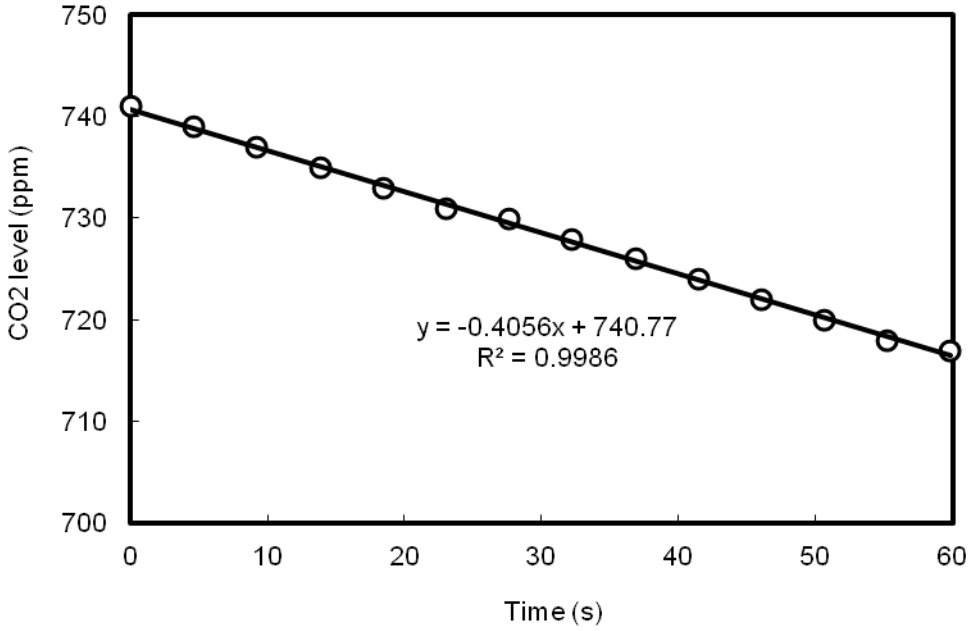
C2. Parameter values of the carbon assimilation and partitioning model (Rasse & Tocquin, 2006)

Symbol	Parameter value	Unit	Parameter description
R	8.314	J K^{-1} mol^{-1}	Gas constant
$K_c(25^0\,C)$	40.4	Pa	Michaelis constant of carboxylation at 25^0 C
H_c	59400	J mol^{-1}	Activation energy for carboxylation constant
$K_{o_2}(25^0\,C)$	24800	Pa	Michaelis constant of oxygenation at 25^0 C
H_{o_2}	36000	J mol^{-1}	Activation energy for oxygenation constant
$V_{c_{max}}(25^0\,C)$	29.6875	μmol m^{-2} s^{-1}	Maximum carboxylation rate at 25^0 C
H_V	64800	J mol^{-1}	Activation energy for maximum carboxylation rate
$[O_2]$	20500	Pa	O_2 partial pressure
H_J	37000	J mol^{-1}	Activation energy for maximum electron transport
R_{JV}	2.1 1.7		(Rasse & Tocquin, 2006) (for an 8-h photoperiod simulation) This study (for 12-h photoperiod
f_{spec}	0.15		Spectral correction factor due to absorbance of irradiance by tissues other than the chloroplast lamella
ST_{br}	0.125		Baseline starch production coefficient
ST_c	0.84		Proportion of night-time starch breakdown
H_r	66400	J mol^{-1}	Activation energy for leaf respiration

Symbol	Parameter value	Unit	Parameter description
SSU_{min}	0.05	g sugar-C m^{-2}	Minimum sugar content in the leaves
GR_{max}	0.408	g g^{-1} d^{-1}	Maximum relative growth rate
α	0.195		Growth respiration coefficient
$p1$	0.7		Ratio of intercellular to ambient CO_2 level
$p2$	3.69	Pa	Quadratic constant of the CO_2 compensation point
$p3$	0.188	Pa K^{-1}	Quadratic constant of the CO_2 compensation point
$p4$	0.0036	Pa K^{-2}	Quadratic constant of the CO_2 compensation point
$p5$	710	J K^{-1} mol^{-1}	Electron transport temperature response parameter
$p6$	220000	J mol^{-1}	Curvature parameter of J_{max}
$p7$	0.7		Curvature of electron transport response to irradiance
$p8$	0.085		Slope of the linear regression between sugar content and maintenance respiration
$p9$	0.016	g C m^{-2}	Slope of the linear regression between sugar content and maintenance respiration

C3. Modelled and measured values of net ecosystem exchange (NEE)

Due to human presence in the growth room during gas exchange measurement, the CO_2 level inside the perplex chamber was higher than usual (Fig. 3.1 in Chapter 3 and figure below). To ensure a valid comparison with measured values, simulated NEE were generated separately by using the CO_2 level measured in the chamber as model input but only at the specific time points as described below.



Measured NEE:

The graph shown above was the measurement for Ler on 37 DAS. Net ecosystem exchange (NEE) was determined from the slope of regression using equation 3.1 (Chapter 3). Measurements were corrected with the control by the following equation:

$$\text{Corrected} = [\text{Non-corrected}] - \frac{[CO_2]_{plants}}{[CO_2]_{control}} \times [\text{Control}]. \qquad \text{(C1.42)}$$

As there was a difference in CO_2 level during the measurement of the plants and the control, the control term in the equation above was scaled to the CO_2 level of the measurement with plants, assuming that the CO_2 levels were at the linear region of the *A*-C_i curve (the plot of assimilation *vs*. CO_2 level) for the autotrophs.

Modelled NEE:

First, the model was executed using as inputs the usual growth room conditions. To simulate the NEE at the specific time points when gas exchange was measured, model outputs such as rosette area and sugar content at those time points were extracted. The model was then executed for one time step using the extracted data together with the measured CO_2 level. In the example shown above, the measured CO_2 level was 740.77 ppm at time zero. Other growth conditions were maintained. Modelled NEE per plant (g C h^{-1}) was calculated as follows:

$$\text{NEE per plant} = -(\text{Carbon assimilation} - R_{above} - R_{below} - R_{ml} - R_{mr}). \quad \text{(C1.43)}$$

The NEE per tray in μmol m^{-2} s^{-1} was estimated by:

$$\text{NEE per tray} = \text{NEE per plant} \times 24 \text{ plants per tray} \times \frac{1}{3600} \times \frac{1}{12} \times 10^6. \quad \text{(C1.44)}$$

Note: NEE has a negative value here as it represents carbon loss from the atmosphere (surroundings) as measured.

C4. Experimental data from the gas exchange measurements of control and plants during the vegetative phase

DAS	Control			Ler			Fei		
	Intercept (CO_2 level)	Slope (Gas exchange rate)	R^2	Intercept (CO_2 level)	Slope (Gas exchange rate)	R^2	Intercept (CO_2 level)	Slope (Gas exchange rate)	R^2
28	686.94	-0.0721	0.9506	623.54	-0.1830	0.9894	664.00	-0.2174	1
29	736.11	-0.0540	0.8018	662.63	-0.1882	0.9918	688.00	-0.2174	1
30	723.26	-0.0755	0.5276	633.63	-0.2145	0.9787	647.40	-0.2427	0.9721
31	1227.50	-0.1524	0.8951	1050.9	-0.3459	0.9839			
33	743.66	-0.1414	0.9882	806.2	-0.3817	0.9960			
34	696.89	-0.0965	0.8623	647.00	-0.2891	0.9459			
36	729.51	-0.1366	0.9621	658.03	-0.3330	0.9190			
37	712.40	-0.1304	0.9808	740.77	-0.4056	0.9986			

C5. Simulation of biomass and net ecosystem exchange (NEE) at 98 % of original carbon assimilation (C1.12)