ABSTRACT

MATTHEWS, JESSICA LOOCK. Sensitivity Analysis and Development of a Model that Quantifies the Effect of Soil Moisture and Plant Age on Leaf Conductance. (Under the direction of Ralph Smith.)

Global climate models are complex compilations of a number of inter-related submodels. Typical implementations include modules for atmospheric, oceanic, sea ice, and land interactions. Our focus is on the photosynthesis submodel of these larger global climate models, which is lacking either by it's omittance or it's oversimplification in many climate models.

A recent model by Niyogi et al. [48] links environmental conditions and plant physiological processes to closely coupled photosynthesis and leaf conductance rates. We start by carefully describing this model and performing sensitivity analysis to characterize the response of model state variables, namely photosynthetic rate, to changes in model parameter values. A novel data set produced from a unique water stress field studying four different levels of watering conditions with two different genotypes of soybeans over the course of two growing seasons is employed for model validation. This data set is applied to the model of Niyogi et al. to test model accuracy under varying soil moisture conditions. It is our conclusion that this model requires refinement to accurately predict total leaf conductance.

To address this, we present a new model to characterize total leaf conductance as a function of plant development and soil moisture conditions. This model is calibrated, via a two-step methodology of global and local optimization algorithms, with the data set describing conductance rates under varying water stress conditions. Estimation of confidence intervals for model parameters and predictions is investigated using asymptotic theory, Monte Carlo methods, and bootstrap methods. For this model, where the residuals exhibit heteroscedasticity, the confidence intervals estimated by the "wild" bootstrap method appear the most realistic of the methods investigated. This new model effectively quantifies the effect of soil moisture and plant age on leaf conductance.

Sensitivity Analysis and Development of a Model that Quantifies the Effect of Soil Moisture and Plant Age on Leaf Conductance

by
Jessica Loock Matthews

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Applied Mathematics

Raleigh, North Carolina

2010

APPROVED BY:

H. T. Banks	Edwin Fiscus
Mansoor Haider	Ralph Smith Chair of Advisory Committee

DEDICATION

To my loving and supportive husband, Quint.

BIOGRAPHY

The author was born in Carroll County, Maryland. After graduation from North Carroll High School in Hampstead, Maryland she attended two Pennsylvania colleges including Pennsylvania State University and York College where she explored such subjects as architecture and photography. As the confusion of teenage angst waned, she wandered to Raleigh, North Carolina due to the generosity of her older brother offering his hospitality. Here she gained focus and rediscovered her passion for studying mathematics. While in Raleigh she met a wonderful native southerner, Quint Kindley, whom she married in 2008.

Over the years the author has supported her fiercely independent nature with a variety of professions including (in chronological order): video store sales clerk, convenience store clerk, waitress, hair salon receptionist, machine shop worker, computer lab help desker, teacher, tutor, engineer at NASA Goddard Space Flight Institute, and biomathematician for the National Institute of Environmental Health Sciences. Following completion of this degree, Jessica will assume the position of Research Associate at the Cooperative Institute for Climate and Satellites in Asheville, NC.

ACKNOWLEDGEMENTS

Firstly, I would like to thank my advisor Dr. Ralph Smith for his direction, patience, and kindness throughout the years. Thanks also to my committee including Dr. H.T. Banks, Dr. Edwin Fiscus, and Dr. Mansoor Haider whom have all been a pleasure to work with during this entire process. I am especially grateful to Dr. Fiscus for generously providing the data that appears in this dissertation.

I want to thank all my colleagues at SRA International, Inc. who encouraged me to pursue my Ph.D. and were supportive throughout the process. Most especially I'd like to thank Marjo, Mike, Pat, Nicole, and Shawn. I will miss you all!

Without the support and encouragement of my family and friends, I would not be where I am today. To all of you, I am eternally grateful. Thanks to Mom and Miss Sue who were always a phone call away with calming words and inspirational guidance. Dad, I credit you for my resilient and independent spirit. Jason, thanks for reminding me that no matter what I do in life, I will always be your little sister. Courtney, thank you for keeping me afloat so many times over the years through your true friendship. Katie, I can always count on you to remind me of all the blessings I have. To my Kali, your loyalty and gentility have grounded me through the last 12 years. And finally to my sweet husband. Quint, you are my all, and I can't wait to start living the life we've dreamt of together.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction

Climate change is a current topic of great controversy in today's world. The arguments formed by every side stem from the problem that no one can predict the future. Although we can not tell with absolute certainty what the future holds for the earth, we do have the technology to produce mathematical models to predict outcomes and their associated prediction variabilities.

Global climate models are complex compilations of a number of inter-related submodels. Typical implementations include modules for atmospheric, oceanic, sea ice, and land interactions. Currently, a number of models exist but, since every model is a simplification of reality, improvements can always be made. Our focus is on the photosynthesis submodel of these larger global climate models. This component is lacking either by it's omittance or it's oversimplification in many climate models. In particular, we examine two models from the bank of models used by the Intergovernmental Panel on Climate Change (IPCC) to generate climate change predictions [55]. The model created by the National Oceanic and Atmospheric Administration (NOAA) omits entirely any photosynthesis component and the inclusion of stomatal conductance does not account for water-stressed conditions [45]. The model created by the National Center for Atmospheric Research (NCAR) includes a photosynthesis submodel based on the formulation presented by Collatz et al. [21], however there is no mechanism for change in stomatal conductance due to moisture conditions [53].

Much remains unknown about the particular causes and consequences of increasing ambient carbon dioxide levels in our atmosphere but, what is known is that the levels are increasing. Photosynthesis is the process by which carbon dioxide is converted to oxygen; therefore, the characterization of the photosynthetic process is critical in climate modeling. If a model omits a photosynthesis module entirely, prediction of carbon dioxide levels is not biologically relevant. Inclusion of the mechanism alone is not sufficient since it is critical that the submodel is accurate under varying environmental conditions. If photosynthetic capacity is overestimated,

the model will return lower than actual predictions of ambient carbon dioxide. Conversely, if photosynthetic capacity is underestimated, the carbon dioxide levels will be overpredicted. Potential impacts of climate change include periods of drought as well as the greater possibility of extreme precipitation events [44]. Because of the importance of correctly characterizing photosynthetic processes in the context of climate modeling, it is imperative that the response of photosynthesis to drought and extreme precipitation events be well understood and correctly modeled.

A number of models describing photosynthesis have been created over the years. Many [2, 21, 47, 60] are based on the so-called Ball-Berry model

$$g_s = m\frac{A_n h_s}{C_s} + b, \tag{1.1}$$

where g_s represents stomatal conductance (mol air m^{-2} s^{-1}), A_n represents the net photosynthesis rate (µmol CO_2 m^{-2} s^{-1}), h_s is the decimal relative humidity at the leaf surface (unitless), C_s is the CO_2 concentration at the leaf surface (µmol mol^{-1} air), and m and b (mol m^{-2} s^{-1}) are the species-specific slope and intercept terms, respectively.

The model presented in Niyogi et al. [48] uses (1.1) as a basis for linking plant physiological and ecological processes to stomatal conductance calculations. The model characterizes a number of biological mechanisms through the incorporation of dependencies on oxygen availability, temperature, species-specific maximum assimilation rate, soil moisture, wind, ambient pressure, leaf length, photosynthetically active radiation (PAR), and leaf scattering of PAR. Given the inclusion of these many mechanisms and model validation under idealized environmental conditions [48], we chose this model as a candidate for extension to accurately predict stomatal conductance and photosynthesis rates under water-stressed conditions in soybeans.

In Chapter 2, we carefully describe this model and perform sensitivity analysis to quantify the response of model state variables to changes in parameter values. In the words of Ball [5], "...understanding how the system functions can permit one to predict how the system might respond if portions of the system were altered; to consider whether or not the system is optimized with respect to specific criteria; and to develop rational strategies for improvement of plants or management practices." In Chapter 3, we consider a novel data set produced from a unique water stress field examining four different levels of watering conditions with two different genotypes of soybeans over the course of two growing seasons. This data set, supplied by Dr. Edwin Fiscus, Plant Physiologist for the Agricultural Research Service (ARS) in the United States Department of Agriculture (USDA), is applied to the model of Niyogi et al., analyzed in Chapter 2, to test model accuracy under varying soil moisture conditions. In Chapter 4, we develop a new model to characterize total leaf conductance as a function of plant development

and soil moisture conditions. This model is calibrated with the data set discussed in Chapter 3. Estimation of confidence intervals for model parameters and predictions is explored using asymptotic theory, Monte Carlo methods, and bootstrap methods. This new model effectively quantifies the effect of soil moisture and plant age on leaf conductance.

Chapter 2

Analysis of Existing Work

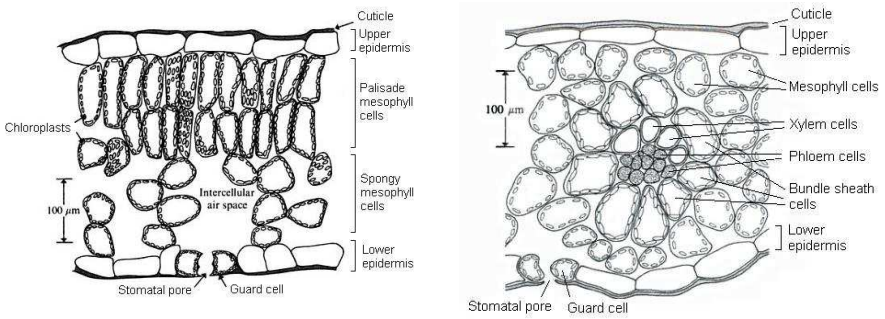
2.1 Background

The world of plants can be divided into three categories according to their photosynthetic processes: C3, C4, and CAM.

The vast majority of plants perform C3 photosynthesis, which is named C3 because atmospheric carbon dioxide (CO_2) is first incorporated into a three-carbon molecule in this photosynthetic process. This carbon fixation is catalyzed by the Rubisco enzyme. Rubisco has dual carboxylase and oxygenase activity, meaning that atmospheric oxygen (O_2) instead of CO_2 may be consumed during the alternative process of photorespiration. This results in less photosynthesis because the Rubsico enzyme is being utilized during the opposite reaction. However, C3 photosynthesis is more efficient than C4 and CAM under cool, moist, and normal light conditions because fewer enzymes are involved in the process. Common examples of C3 species include soybeans, rice, and cotton.

C4 plants are considered to be more advanced than their C3 counterparts. Bypassing the photorespiration pathway, C4 plants convert atmospheric CO_2 to a four-carbon compound in the mesophyll cells catalyzed by the phosphoenolpyruvate carboxylase (PEP-carboxylase) enzyme. This compound is then shuttled to the bundle sheath cells where CO_2 is delivered to continue processing along the conventional C3 pathway. This alternative delivery method enables C4 plants to photosynthesize faster than C3 plants in high light and high temperature conditions while avoiding photorespiratory losses. Figure 2.1 illustrates the anatomical differences between C3 and C4 plants [49]. C4 plants also exhibit better water use efficiency because their stomata do not need to remain open as long since their photosynthetic process is more efficient [10]. Corn, sugarcane, and crabgrass are a few of the thousands of identified C4 species.

Plant species performing Crassulacean acid metabolism (CAM) photosynthesis are mainly



(a) (b)

Figure 2.1: Schematic transverse section of an (a) C3 and (b) C4 leaf.

succulents and differ from C3 and C4 plants because their stomata open at night rather than during the day. Atmospheric CO_2 entering the leaf is converted to an acid and stored overnight until breaking down and releasing the CO_2 to Rubisco for photosynthesis during the day when abundant light energy is available. These plants thrive in arid conditions because their stomata open at night thus avoiding transpirational losses due to radiative heating. The small fraction of the plant kingdom which utilizes CAM photosynthesis includes many orchids, pineapple, and all varieties of cactus.

With the modification of a few equations and parameters the model in [48], analyzed subsequently, is applicable to both C3 and C4 plant species.

2.1.1 C3 Species

In plants, stomata are pores on the leaf and stem used for gas exchange. Oxygen produced by photosynthesis exits through the stomata which also restricts the outward flux of water (transpiration). Inward diffusion of air containing CO_2 and O_2 to sustain photosynthesis and respiration, respectively, also occurs through the stomata. Environmental variables affecting stomatal aperture include temperature, air humidity, wind speed, light intensity, water stress, and CO_2 concentration.

The model presented in Niyogi et al. [48] employs a stomatal model linking plant physiological and ecological processes to stomatal conductance calculations. In particular, they use the relative-humidity-based approach, introduced by Ball et al. [4] and expanded in Ball [5], which yields the relation

$$g_s = m\frac{A_n h_s P}{C_s} + b. \tag{2.1}$$

Here g_s represents stomatal conductance (mol air m^{-2} s^{-1}), A_n represents the net photosynthesis rate (mol CO_2 m^{-2} s^{-1}), h_s is the decimal relative humidity at the leaf surface (unitless), P is the surface pressure (Pa), C_s is the CO_2 partial pressure at the leaf surface (Pa), and m (unitless) and b (mol m^{-2} s^{-1}) are the species-specific slope and intercept terms, respectively. Note that we have included the surface pressure P, which was omitted from [48] but is required for unit balance. This so-called Ball-Berry model given in (2.1) is frequently a basis for coupled leaf photosynthesis and stomatal conductance models [2, 3, 21, 22, 47, 60]. The robustness of this model was validated in [5] with more than 500 data points from plants of 13 different species grown under varying light and nutrient conditions. It is noted that values for m and b are species specific, where m tends to be smaller in the C4 species than C3. Also, as A_n tends towards zero, as a result of low photon flux or low ambient CO_2 concentrations, the stomatal response is no longer linear meaning this model is not a good approximation of the kinetics. Therefore when applying this model, one must take notice of these environmental conditions [5].

Remark. To adhere to proper units in the calculation of g_s, the surface pressure term P is necessary. Although this term was omitted by [48], this corrected formulation is required when C_s is defined as partial pressure rather than concentration.

The net photosynthesis rate is defined as

$$A_n = A_g - R_d \tag{2.2}$$

where A_g is the gross carbon assimilation rate (mol CO_2 m^{-2} s^{-1}) and R_d is the rate of CO_2 evolution resulting from processes other than photorespiration which is known as the dark respiration rate (mol CO_2 m^{-2} s^{-1}). This equation seems to gain recognition with the work of Farquhar et al. [29] and is recurrent in the modeling literature [3, 21, 22, 47, 60, 63, 66].

Gross Carbon Assimilation Rate

The gross carbon assimilation rate is determined by three limiting factors

- Rubisco (w_c)
- Light (w_e)
- Plant capacity to utilize photosynthesis products (w_s)

via the coupled set of quadratic equations

$$\beta_1 w_p^2 - w_p(w_c + w_e) + w_e w_c = 0 \tag{2.3}$$

$$\beta_2 A_g^2 - A_g(w_p + w_s) + w_p w_s = 0 \tag{2.4}$$

where w_p is an interim variable representing the rate limitation imposed by Rubisco and light, and the β's are coupling coefficients.

Rubisco limitation: Rubisco, the common name for Ribulose-1,5-bisphosphate carboxylase oxygenase, is an enzyme used in the first step of the Calvin cycle, which is a series of biochemical reactions that take place in photosynthetic organisms to convert CO_2 into organic compounds for use by the organism. Following Collatz et al. [21], the first limitation to the C3 gross photosynthesis rate is quantified by the relation

$$w_c = V_m \frac{C_i - \Gamma^*}{C_i + K_c(1 + \frac{O_2}{K_o})} \tag{2.5}$$

where V_m represents the catalytic Rubisco capacity for the leaf (mol CO_2 m^{-2} s^{-1}), C_i is the partial pressure of CO_2 in intercellular spaces (Pa), Γ^* is the compensation point (Pa), K_c is the Michaelis-Menten constant for CO_2 (Pa), O_2 is the oxygen availability in leaf cells (Pa), and K_o is the oxygen inhibition constant (Pa). Similar formulations for the Rubisco limitation on photosynthesis in C3 plants are found in [29, 47, 60, 63].

Niyogi et al. [48] developed the unique representation

$$V_m = V_{max} f(T) f(w_2) 2.1^{Q_{10}} \tag{2.6}$$

which includes dependencies on both temperature and soil moisture. V_{max} defines the maximum Rubisco catalytic capacity (mol CO_2 m^{-2} s^{-1}), $f(T)$ is the temperature-dependence (unitless), and $f(w_2)$ is the soil moisture dependence (unitless). The unitless temperature-dependent term Q_{10} is defined as

$$Q_{10} = \frac{T_s - 298.0}{10} \tag{2.7}$$

where T_s defines the temperature at the leaf surface in degrees Kelvin. This response function defines the proportional increase of a parameter value for a 10 degree increase in temperature. Similar calculations to incorporate various parameter temperature-dependencies are found in [21, 22, 29, 33, 60, 67]. The portion 2.1^{Q10} of (2.6) is unique to this model. An additional temperature-dependency incorporated in the calculation of V_m, to mimic the whole plant response, is

$$f(T) = \frac{2^{Q_{10}} \left(1 + e^{0.3(S_4 - T_s)}\right)}{1 + e^{0.3(T_s - S_2)}} \tag{2.8}$$

where S_2 and S_4 are the high and low vegetation stress factors, respectively. The motivation for including this formulation of temperature dependency is to approximate the response of C4 plants to extreme temperatures [9]. A similar equation is utilized in Sellers et al. [60], with the exception that the coefficient in the denominator has a value of 0.2 instead of 0.3. The sensitivity of this coefficient value is explored in Section 2.3.24. Collatz et al. [22] also incorporates a similar temperature dependency except both parenthesis-enclosed terms of (2.8) occur in the denominator of their formulation.

The soil moisture-dependency of V_m is defined as

$$f(w_2) = \frac{w_2 - w_{wilt}}{w_{fc} - w_{wilt}} \tag{2.9}$$

where w_2 is the deep soil moisture content, w_{wilt} is the root soil moisture wilting value, and w_{fc} is the root level soil moisture field capacity value. The value of $f(w_2)$ is always between 0 and 1. This same regulation imposed by soil moisture stress appears throughout the literature, but is applied not to the catalytic Rubisco capacity, but to surface resistance [50], mesophyll conductance [16], and net carbon assimilation [2]. Sellers et al. [60] does incorporate a soil moisture stress factor to their calculation of V_m and, although it is dependent on soil type, it is significantly different than (2.9).

Before defining the CO_2 partial pressure in the intercellular spaces, we first examine the construction of the interaction between the leaf surface and the ambient environment controlled by the leaf boundary layer conductance g_b. The leaf boundary layer is a region of relatively still air adjacent to the surface of the leaf. Any gas which enters or exits the leaf must pass through this layer. The value of g_b influences the environmental conditions at the leaf surface which affects and is affected by stomatal conductance g_s. In general, as g_b decreases, the humidity of leaf surface air increases because more transpired water vapor is present in the boundary layer which causes g_s to increase [21]. The overall leaf boundary layer conductance is taken to be the maximum of forced and free convective conditions [47]. The conductance under forced conditions is defined as

$$g_{bfc} = cT_a^{0.56} \left[(T_a + 120)\frac{u}{dP}\right]^{0.5} \tag{2.10}$$

where c is a transfer coefficient (mol m^{-2} s^{-1}), T_a is the ambient temperature (K), u is the wind speed (m s^{-1}), and d is the leaf length scale (m) [47, 63]. This is considered forced due to the affect of wind, which is external to the leaf itself. The conductance under free conditions is

$$g_{bfr} = cT_s^{0.56}\left[\frac{T_s + 120}{P}\right]^{0.5}\left[\frac{|T_{vs} - T_{va}|}{d}\right]^{0.25} \tag{2.11}$$

where T_s is the surface leaf temperature (K), T_{vs} is the virtual surface leaf temperature (K), and T_{va} is the virtual ambient temperature (K) [47, 63].

In general, virtual temperatures are defined to be the temperature that dry air needs to have the same density as air containing moisture. Teten's formula

$$T_v = (1 + 0.61\omega)T \tag{2.12}$$

was used to compute T_{vs} and T_{va} where ω represents the mixing ratio (kg water vapor / kg dry air). Note that ω is related to the saturation mixing ratio, ω_s (kg water vapor / kg dry air), at saturation, by the relation

$$h_s = \frac{\omega}{\omega_s}. \tag{2.13}$$

Linear regression is performed to identify ω relating the values presented for ω_s in Table 2.1 to the values of h_s, T_s, and T_a used in the model simulation.

An alternative way to compute virtual temperatures takes pressure into account for the calculation (the previous method assumes pressure is 1000 millibars). First, find the saturation vapor pressure of water (e_s) in kiloPascal as defined in Monteith and Unsworth [46]:

$$e_s = 0.611e^{17.27\frac{T-273}{T-36}}. \tag{2.14}$$

Then find e from the relation $h_s = \frac{e}{e_s}$. Finally, we have

$$T_v = \frac{T}{1 - 0.378\frac{e}{P}}. \tag{2.15}$$

Table 2.1: Saturation mixing ratio (ω_s) values at varying temperatures and $P = 1000$ millibars.

Temperature (K)	ω_s (g/kg)
273.15	3.84
283.15	7.76
293.15	14.85
303.15	27.69
313.15	49.81
323.15	88.12

The carbon dioxide partial pressure at the leaf surface, C_s, is then calculated using the relation

$$C_s = C_a - \frac{A_n P}{g_b} \tag{2.16}$$

where C_a is the ambient carbon dioxide partial pressure (Pa). Therefore, the CO_2 partial pressure at the leaf surface is the partial pressure in the ambient air minus the CO_2 diffusing through the leaf boundary layer to be photosynthesized. Note that we have included P, which was omitted from [48], but is required for unit balance. This formulation for C_s also appears in [3, 63]. Nikolov et al. [47] and Sellers et al. [60] also use this formulation but suggest introducing an additional coefficient to A_n, $1.6^{\frac{2}{3}}$, to describe the effect of molecular diffusivity of H_2O and CO_2 under the influence of air flow in the leaf boundary layer.

Remark. To adhere to proper units in the calculation of C_s, the surface pressure term P is necessary. Although this term was omitted by [48], this corrected formulation is required when C_s is defined as partial pressure rather than concentration.

Finally, the intercellular CO_2 partial pressure, C_i is computed by

$$C_i = C_s - \frac{\eta A_n P}{g_s} \tag{2.17}$$

where η represents the ratio of molecular diffusivity of H_2O and CO_2 in still air through the stomatal pores. This version of C_i appears in [4, 22, 60]. It is important to notice by the formulations of C_s and C_i that the relationship $C_i < C_s < C_a$ will always hold.

Remark. The formula for C_i (2.17) is identical to that reported in [48]. However, they report the units as concentration whereas this formulation results in the calculation of partial pressure.

The compensation point

$$\Gamma^* = \frac{O_2}{2S} \tag{2.18}$$

is defined as the intercellular CO_2 partial pressure below which the leaf is unable to assimilate because of photorespiration. Here O_2 is the oxygen availability in leaf cells (Pa) and S is the unitless Rubisco specificity for CO_2 relative to O_2. This same formulation appears in [21, 60], both applied to modeling C3 photosynthesis. It is unclear why this dependency is also incorporated into modeling of C4 photosynthesis because the issue of Rubisco competition with oxygen is avoided.

The Rubsico specificity for carbon dioxide relative to oxygen is defined as

$$S = 2600 \cdot 0.57^{Q_{10}}. \tag{2.19}$$

The particular coefficient values in the expression for S are referenced to Collatz et al. [21]. Although no direct explanation is given for the derivation of these values, the work presented by Woodrow and Berry [67] have similar values derived from published data [14, 35]. The combination of (2.18) and (2.19) yield an exponential dependency of Γ^* on temperature while others formulate a polynomial dependency on temperature [14, 47].

The Michaelis-Menten constants of Rubisco for CO_2 and O_2 are defined as

$$K_c = 30 \cdot 2.1^{Q_{10}} \tag{2.20}$$

and

$$K_o = 30000 \cdot 1.2^{Q_{10}}, \tag{2.21}$$

respectively, having units in Pa. This Q_{10} formulation is also presented in [21, 60] and is likely motivated by the Arrhenius equations proposed to characterize these constants in [29, 39, 47].

Light limitation: The second factor limiting the rate of gross carbon assimilation is the amount of photosynthetically active radiation (PAR) absorbed by the leaf chlorophyll, which we describe as light limitation. Light limitation is quantified by the relation

$$w_e = PAR\epsilon(1 - w_\pi)\frac{C_i - \Gamma^*}{C_i + 2\Gamma^*} \tag{2.22}$$

where PAR is the photosynthetically active radiation (mol m^{-2} s^{-1}), ϵ is the quantum efficiency for CO_2 uptake (mol mol^{-1}), and w_π is the leaf-scattering coefficient for PAR (unitless). Similar formulations for the C3 species are presented in [16, 21, 32, 33, 60]. Several models present a different limiting factor in lieu of light limitation, namely limitation by the regeneration capacity of Rubisco which is controlled by electron transport [3, 26, 47, 63].

Plant capacity to utilize photosynthesis products: The third, and final, factor we are considering to limit the rate of gross carbon assimilation in C3 plants is the capacity of the plant to utilize the products of photosynthesis. In particular this is the rate of sucrose and starch (i.e., triose phosphate) synthesis. This limitation is expressed as

$$w_s = \frac{V_m}{2} \tag{2.23}$$

which is identical to the formulations in [21, 47, 60]. More complex alternative formulations are sometimes as used, an example of which is found in [26].

Dark Respiration Rate

To compute the net photosynthesis rate A_n, the dark respiration rate R_d is subtracted from the gross photosynthesis rate A_g as described in (2.2). In the literature, a variety of formulations for R_d are presented including temperature-based calculations [22, 29, 63] and linear relationships with maximum carboxylation velocity (V_m) [21, 47, 60]. In Niyogi et al. [48], the observations of van Heemst [64] motivate the expression

$$R_d = \frac{A_m}{9.0} \tag{2.24}$$

where A_m is the net assimilation rate (mol m^{-2} s^{-1}) limited by a deficit of CO_2.

The approach introduced by Goudriaan et al. [32], and further developed by Jacobs [33] and Calvet et al. [16], to compute A_m is applied to yield

$$A_m = A_{m,max}\left[1 - e^{-g_m \frac{C_i - \Gamma^*}{A_{m,max} P}}\right] \tag{2.25}$$

where $A_{m,max}$ is the maximum net assimilation (mol m^{-2} s^{-1}) and g_m is mesophyll conductance (m s^{-1}). Others consider $A_{m,max}$ as a variable dependent on temperature, but here it is treated as a species-dependent constant.

Remark. To adhere to proper units in the calculation of A_m, the surface pressure term P is necessary. Although this term was omitted by [48], this corrected formulation is required when C_i and Γ^* are defined as partial pressures rather than concentrations.

Mesophyll conductance, which describes the transport of CO_2 between the sub-stomatal cavity and the site of carboxylation, is quantified by the relation

$$g_m = g_{mp}\left[2^{Q_{10}} \frac{1 + e^{0.3(T_c - S_2)}}{1 + e^{0.3(S_4 - T_c)}} f(w_2)\right] \tag{2.26}$$

where g_{mp} is the species-specific maximum value of mesophyllic conductance (m/s). The dependence of g_m on soil moisture stress was first introduced in [16]. A critical difference between Niyogi et al.'s approach and that presented in Jacobs [33] is that the expression $1 + e^{0.3(T_c - S_2)}$ was in the denominator rather than the numerator. It is unclear why this change has been incorporated. Further, the formulation for g_m presented in [16, 33] is remarkably similar to the

expression for the temperature-dependent, substrate-saturated Rubisco capacity in [22]. Another notable difference in Jacobs' model is that the values of S_2 and S_4 are also species-specific when calculating g_m.

2.1.2 C4 Species

Model application to C4 species differs from that of C3 species in the three rate-limiting equations of gross photosynthetic rate as well as various parameter values. In particular the parameters which require species-dependent values are:

- Quantum efficiency for CO_2 uptake, ϵ
- Leaf-scattering coefficient for PAR, ω_π
- Maximum catalytic Rubisco capacity for leaf, V_{max}
- Maximum net assimilation rate, $A_{m,max}$
- Potential maximum value of mesophyllic conductance, g_{mp}
- Slope in stomatal conductance equation, m
- Intercept in stomatal conductance equation, b

The three factors limiting the gross photosynthetic rate in C4 species are: Rubisco (w_c), light (w_e), and PEP-Carboxylase (w_s).

Rubisco limitation: The photosynthetic rate of C4 plants, due to their more evolved carbon fixation mechanism, is limited by the Rubisco capacity but is not affected by utilization of Rubisco to catalyze photorespiration (which uses O_2 instead of CO_2 as in photosynthesis). Following Collatz et al. [22], the first limitation to the C4 gross photosynthesis rate is simply

$$w_c = V_m. \tag{2.27}$$

Light limitation: The second factor limiting the rate of gross carbon assimilation is the amount of photosynthetically active radiation (PAR) absorbed by the leaf chlorophyll, which we describe as light limitation. Light limitation is quantified by the relation

$$w_e = PAR\epsilon(1 - w_\pi) \tag{2.28}$$

where PAR is the photosynthetically active radiation (mol m^{-2} s^{-1}), ϵ is the quantum efficiency for CO_2 uptake (mol mol^{-1}), and w_π is the leaf-scattering coefficient for PAR (unitless). Similar formulations to account for the light limitation of A_g in C4 species are presented in [22, 60, 63]. As explained in [33], this formulation for w_e is identical to that for C3 plants, (2.22), substituting the CO_2 compensation point $\Gamma^* = 0$ because photorespiration is not an issue with C4 plants.

PEP-Carboxylase limitation The third, and final, factor we are considering to limit the rate of gross carbon assimilation is the PEP-Carboxylase capacity of C4 vegetation. PEP-Carboxylase, formally known as phosphoenolpyruvate carboxylase, is an enzyme that catalyzes carbon fixation in the mesophyll cells of C4 plants. This limitation is expressed as

$$w_s = 20000 V_m \frac{C_i}{P}. \tag{2.29}$$

Described instead as the CO_2-limited capacity for C4 photosynthesis, this formulation appears in [60]. A temperature-, CO_2-, and pressure-dependent representation which unlike (2.29) is notably independent of soil moisture, is described by [22].

2.2 Model Solution

Each simulation requires the parameters defined in Table 2.2. From these parameters we calculate the following terms for either the C3 or C4 species:

- temperature-dependency term
$$Q_{10} = \frac{T_s - 298.0}{10}, \tag{2.30}$$

- forced convection leaf boundary layer conductance
$$g_{bfc} = c T_a^{0.56} \left[(T_a + 120) \frac{u}{dP} \right]^{0.5}, \tag{2.31}$$

- free convective leaf boundary layer conductance
$$g_{bfr} = c T_s^{0.56} \left[\frac{T_s + 120}{P} \right]^{0.5} \left[\frac{|T_{vs} - T_{va}|}{d} \right]^{0.25}, \tag{2.32}$$

- leaf boundary layer conductance
$$g_b = \max(g_{bfc}, g_{bfr}), \tag{2.33}$$

Table 2.2: Model parameters from Niyogi et al. [48].

Parameter (units)	Physical representation	Value
T_s (K)	Surface temperature	297
T_c (K)	Canopy temperature	293
T_a (K)	Ambient temperature	295
P (Pa)	Surface pressure	1.01×10^5
u (m/s)	Wind speed	5
O_2 (Pa)	Oxygen availability in leaf cells	2.09×10^4
PAR (mol m^{-2} s^{-1})	Photosynthetically active radiation	1.38×10^{-3}
h_s (-)	Relative humidity at leaf surface	0.5
C_a (Pa)	Ambient carbon dioxide partial pressure	34
w_{wilt} (-)	Root level soil moisture wilting value	0.25
w_{fc} (-)	Field capacity value	0.3
w_2 (-)	Deep soil moisture content	0.27
ϵ (mol mol^{-1})	Quantum efficiency for CO_2 uptake	0.08 (C3) 0.05 (C4)
w_π (-)	Leaf-scattering coefficient for PAR	0.1 (C3) 0.2 (C4)
K_c (Pa)	Michaelis-Menten constant for CO_2	30 (C3)
K_o (Pa)	Michaelis-Menten constant for O_2	3. $\times 10^4$ (C3)
η (-)	Factor to account for different diffusivities of H_2O and CO_2 in the stomatal pores	1.6
S_2 (K)	High vegetation stress factor	310
S_4 (K)	Low vegetation stress factor	280
V_{max} (mol m^{-2} s^{-1})	Maximum catalytic Rubsico capacity for leaf	7.5 $\times 10^{-5}$ (C3) 3 $\times 10^{-5}$ (C4)
$A_{m,max}$ (mol m^{-2} s^{-1})	Maximum net assimilation rate	9.8 $\times 10^{-5}$ (C3) 7.48 $\times 10^{-5}$ (C4)
g_{mp} (m/s)	Potential maximum value of mesophyllic conductance	7. $\times 10^{-3}$ (C3) 17.5 $\times 10^{-3}$ (C4)
c (mol m^{-2} s^{-1})	Transfer coefficient	4.322×10^{-3}
d (m)	Leaf length scale	41. $\times 10^{-3}$
β_1 (-)	Coupling coefficient	0.8
β_2 (-)	Coupling coefficient	0.99
m (-)	Slope in stomatal conductance equation	9 (C3) 4 (C4)
b (mol m^{-2} s^{-1})	Intercept in stomatal conductance equation	0.01 (C3) 0.04 (C4)

- Rubisco specificity for CO_2 relative to O_2

$$S = 2600 \cdot 0.57^{Q_{10}}, \tag{2.34}$$

- CO_2 compensation point

$$\Gamma^* = \frac{O_2}{2S}, \tag{2.35}$$

- temperature dependency of V_m

$$f(T) = \frac{2^{Q_{10}} \left(1 + e^{0.3(S_4 - T_s)}\right)}{1 + e^{0.3(T_s - S_2)}}, \tag{2.36}$$

- soil moisture modulation

$$f(w_2) = \frac{w_2 - w_{wilt}}{w_{fc} - w_{wilt}}, \tag{2.37}$$

- maximum catalytic Rubisco capacity for the leaf

$$V_m = V_{max} f(T) f(w_2) 2.1^{Q10}, \tag{2.38}$$

- mesophyllic conductance

$$g_m = g_{mp} \left[2^{Q_{10}} \frac{1 + e^{0.3(T_c - S_2)}}{1 + e^{0.3(S_4 - T_c)}} f(w_2) \right]. \tag{2.39}$$

Then, we begin by using the equation defining carbon dioxide partial pressure in the intercellular spaces,

$$C_i = C_s - \frac{\eta A_n P}{g_s}, \tag{2.40}$$

and substituting the equation for stomatal conductance,

$$g_s = m \frac{A_n h_s P}{C_s} + b, \tag{2.41}$$

into (2.40) yielding

$$C_i = C_s - \frac{\eta A_n P}{m \frac{A_n h_s P}{C_s} + b}. \tag{2.42}$$

Further, inserting the equation for carbon dioxide at the leaf surface

$$C_s = C_a - \frac{A_n P}{g_b} \tag{2.43}$$

into (2.42) produces

$$C_i = C_a - \frac{A_n P}{g_b} - \frac{\eta A_n P}{m \frac{A_n h_s P}{C_a - \frac{A_n P}{g_b}} + b}. \tag{2.44}$$

Next we substitute the definition for maximum assimilation rate

$$A_m = A_{m,max} \left(1 - e^{-g_m \frac{C_i - \Gamma^*}{A_{m,max} P}} \right) \tag{2.45}$$

along with (2.44) into the definition for leaf respiration rate to yield

$$\begin{aligned} R_d &= \frac{A_m}{9} \\ &= \frac{A_{m,max}}{9} \left(1 - e^{-g_m \frac{C_i - \Gamma^*}{A_{m,max} P}} \right) \\ &= \frac{A_{m,max}}{9} \left(1 - e^{\frac{-g_m}{A_{m,max} P} \left(C_a - \frac{A_n P}{g_b} - \frac{\eta A_n P}{m \frac{A_n h_s P}{C_a - \frac{A_n P}{g_b}} + b} - \Gamma^* \right)} \right). \end{aligned} \tag{2.46}$$

2.2.1 C3 Species

For the C3 species, we additionally calculate:

- Michaelis-Menten constant of Rubisco for CO_2

$$K_c = 30 \cdot 2.1^{Q_{10}} \tag{2.47}$$

- Michaelis-Menten constant of Rubisco for O_2

$$K_o = 30000 \cdot 1.2^{Q_{10}}, \tag{2.48}$$

- capacity of C3 vegetation to utilize the photosynthesis products

$$w_s = \frac{V_m}{2}. \tag{2.49}$$

Substituting (2.44) into the definitions for Rubisco limitation of photosynthesis rate in C3 vegetation yields

$$\begin{aligned} w_c &= V_m \frac{C_i - \Gamma^*}{C_i + K_c(1 + \frac{O_2}{K_o})} \\ &= V_m \frac{C_a - \frac{A_n P}{g_b} - \frac{\eta A_n P}{m \frac{A_n h_s P}{C_a - \frac{A_n P}{g_b}} + b} - \Gamma^*}{C_a - \frac{A_n P}{g_b} - \frac{\eta A_n P}{m \frac{A_n h_s P}{C_a - \frac{A_n P}{g_b}} + b} + K_c(1 + \frac{O_2}{K_o})}, \end{aligned} \tag{2.50}$$

and light limitation of the photosynthesis rate in C3 vegetation yields

$$\begin{aligned} w_e &= PAR\epsilon(1 - w_\pi) \frac{C_i - \Gamma^*}{C_i + 2\Gamma^*} \\ &= PAR\epsilon(1 - w_\pi) \frac{C_a - \frac{A_n P}{g_b} - \frac{\eta A_n P}{m \frac{A_n h_s P}{C_a - \frac{A_n P}{g_b}} + b} - \Gamma^*}{C_a - \frac{A_n P}{g_b} - \frac{\eta A_n P}{m \frac{A_n h_s P}{C_a - \frac{A_n P}{g_b}} + b} + 2\Gamma^*}. \end{aligned} \tag{2.51}$$

Now we can define the minimum assimilation rate estimated between w_c and w_e, namely w_p. Here w_p is defined to be the smaller root of the quadratic equation

$$\beta_1 w_p^2 - w_p(w_c + w_e) + w_e w_c = 0. \tag{2.52}$$

Since we are assuming the limiting factors on gross photosynthesis, w_c, w_e, and w_s, as well as the coupling coefficient β_1 are all greater than or equal to 0, the smallest root of (2.52) is

$$w_p = \frac{(w_c + w_e) - \sqrt{(w_c + w_e)^2 - 4\beta_1 w_e w_c}}{2\beta_1}. \tag{2.53}$$

We then substitute (2.50) and (2.51) into (2.53), which produces an expression for w_p where A_n is the only unknown.

Recall that A_g is defined to be the smaller root of the quadratic equation

$$\beta_2 A_g^2 - A_g(w_p + w_s) + w_p w_s = 0. \tag{2.54}$$

Since we are assuming the coupling coefficient $\beta_2 > 0$, the smallest root of (2.54) is

$$A_g = \frac{(w_p + w_s) - \sqrt{(w_p + w_s)^2 - 4\beta_2 w_p w_s}}{2\beta_2}. \tag{2.55}$$

Substituting (2.53) and (2.49) into (2.55) produces an expression for A_g where A_n is the only unknown.

Finally, because the equation for gross photosynthetic rate is

$$A_g = A_n + R_d, \tag{2.56}$$

we can substitute (2.55) and (2.46) into (2.56) to form an algebraic expression for C3 vegetation where the only unknown remaining is A_n.

2.2.2 C4 Species

For the C4 species, we additionally calculate:

- Rubisco limitation of photosynthesis rate in C4 vegetation

$$w_c = V_m, \tag{2.57}$$

- light limitation of photosynthesis rate in C4 vegetation

$$w_e = PAR\epsilon(1 - w_\pi), \tag{2.58}$$

- minimum assimilation rate estimated between w_c and w_e

$$w_p = \frac{(w_c + w_e) - \sqrt{(w_c + w_e)^2 - 4\beta_1 w_e w_c}}{2\beta_1}. \tag{2.59}$$

Then, using (2.44), for C4 vegetation we may define the PEP-Carboxylase limitation on photosynthesis as

$$\begin{aligned} w_s &= 20000 V_m \frac{C_i}{P} \\ &= 20000 \frac{V_m}{P} \left(C_a - \frac{A_n P}{g_b} - \frac{\eta A_n P}{m \frac{A_n h_s P}{C_a - \frac{A_n P}{g_b}} + b} \right). \end{aligned} \tag{2.60}$$

Gross photosynthetic rate A_g, in terms of A_n, can be quantified with (2.55). Finally, substituting A_g and (2.46) into (2.56) forms an algebraic expression for C4 vegetation where the only unknown remaining is A_n.

2.2.3 Root Finding

Thus, the system of equations is combined into one algebraic expression

$$f(A_n) = 0 \tag{2.61}$$

for the net photosynthesis rate A_n. Due to the highly complex and nonlinear nature of this relation, analytic solution techniques are not feasible. Therefore we employ a numerical root finding algorithm.

Previous efforts to solve the system appear to lack numerical robustness and adequate measures to check for non-feasible state variables. We present an algorithm which is numerically stable and yields model predictions which are in the defined physical domain. In particular we impose the following constraints:

- $A_n \geq 0$,
- $C_s \geq 0$, and
- $C_i \geq \Gamma^*$.

We must implement these constraints because, although it is obvious that if any of these are violated the model is invalid (i.e., it does not make sense to consider negative photosynthetic rates or partial pressures and by definition of the compensation point Γ^*, C_i can not be lower than this minimum), mathematically it is possible to find roots of the system (2.61) which violate these constraints.

To identify the roots, a general grid search method is employed. We begin by defining possible A_n values

$$A_{n_i} = \frac{i}{10^7} \qquad \text{for } i\text{=}0,\ldots,1000. \tag{2.62}$$

Note that this definition for the possible domain for A_n inherently satisfies our first condition that $A_n \geq 0$. Then C_s is evaluated for each A_{n_i} using (2.16). The values A_{n_i} which yield $C_s < 0$ are removed from the possible A_n solution domain to enforce our second condition, $C_s \geq 0$. Then Γ^* is calculated and C_i is evaluated for each A_{n_i} using (2.17). The values A_{n_i} which yield $C_i < \Gamma^*$ are removed from the possible A_n solution domain to enforce our third constraint, $C_i \geq \Gamma^*$.

The function $f(A_n)$ is evaluated for the A_{n_i} values satisfying $C_s \geq 0$ and $C_i \geq \Gamma^*$. Since our goal is to locate the roots of $f(A_n)$, we now identify the intervals where the function evaluation crosses zero using Algorithm 1.

This algorithm should return points A_{n_i} where each point defines the starting value of the interval on the coarse grid for a root of $f(A_n)$. Now, as described in Algorithm 2, the function is evaluated from A_{n_i} incrementing by 10^{-10} until zero is crossed. Define $\hat{A}_{n_{i,1}}$ and $\hat{A}_{n_{i,2}}$ to be the values, differing only by 10^{-10}, which produce $f(\hat{A}_{n_{i,1}})$ and $f(\hat{A}_{n_{i,2}})$ on either side of zero. The ith root of $f(A_n)$ is

$$A_{n,i} = \arg \min_{(\hat{A}_{n_{i,1}}, \hat{A}_{n_{i,2}})} \left[|f(\hat{A}_{n_{i,1}})|, |f(\hat{A}_{n_{i,2}})| \right]. \tag{2.63}$$

Input: A function file 'fcn(A_n)' representing the algebraic expression of $f(A_n)$.
Input: A vector 'A_n' containing values in the domain of possible solutions to $f(A_n) = 0$.
Output: A vector 'crossInd' containing A_n values on the defined coarse grid immediately before a zero crossing.

fval = feval(fcn,A_n);
`/* changeInd is logical vector of length same as` A_n`, zeros where fcn(`A_n`(i))>=0 and ones where fcn(`A_n`(i))<0 */`
changeInd = fval < 0 ;
crossInd = [];
for *k = 2:length(changeInd)* **do**
  if *changeInd(k)* $\neq$ *changeInd(k-1)* **then**
    crossInd = [crossInd A_n(k-1)];
  end
end

Algorithm 1: Coarse grid search.

```
Input: A function file 'fcn(A_n)' representing the algebraic expression of f(A_n).
Input: A vector 'crossInd' containing A_n values on the defined coarse grid immediately
       before and after a zero crossing.
Output: A vector 'roots' containing A_n values within 10^-10 of the roots of 'fcn(A_n)'.
Output: A vector 'vals' containing the function evaluation associated with 'fcn(roots)'.
inc = 10^-10 ;
/* for each point defined on the coarse grid to be preceding a root,
   increment slowly to find the A_n-value closest to the actual root     */
for k=1:length(crossInd) do
    rootStart = crossInd(k) ;
    chgSgn = 1 ;                                    /* track when we cross 0 */
    valOld = feval(fcn,rootStart) ;
    rootStart = rootStart + inc ;
    while chgSgn do
        valNew = feval(fcn,rootStart) ;
        /* stop if we change signs, o.w.  keep going                         */
        if sign(valNew)*sign(valOld) < 0 then
            chgSgn = 0 ;
        else
            rootStart = rootStart + inc ;
        end
    end
    /* valNew and valOld are function values on either side of 0.  identify
       the one closest to 0 to be the root with tolerance < 10^-6.          */
    if abs(valNew) < abs(valOld) then
        roots(k) = rootStart ;
        vals(k) = valNew ;
    else
        roots(k) = rootStart - inc ;
        vals(k) = valOld ;
    end
end
```

Algorithm 2: Fine grid search.

2.3 Sensitivity Analysis

The key output of interest from this model is the photosynthesis rate A_n (mol m^{-2} s^{-1}). The sensitivity of A_n with respect to model parameters is studied by setting all parameters to those prescribed in Niyogi et al. [48] (see Table 2.2) and evaluating the impact on A_n as parameters are individually perturbed. In general, change in response was evaluated by perturbing parameters by 10% unless otherwise indicated. In most cases, we present sensitivity analysis for C3 and C4 jointly by computing the percent of change in A_n from the respective baseline values. Using the parameter values in Table 2.2, the C3 model predicts $A_n = 5.3791$ µmol m^{-2} s^{-1} and for the C4 model $A_n = 9.6809$ µmol m^{-2} s^{-1}.

2.3.1 Surface Temperature T_s

T_s is an input parameter in the calculation of several other parameters, namely the temperature dependence (Q_{10}), virtual surface temperature (T_{vs}), free convection leaf boundary layer conductance (g_{bfr}), and temperature impact on V_m ($f(T)$). None of these formulations restrict the range of T_s values, so the sensitivity analysis of T_s, presented in Figure 2.2, examines a large range of possible temperatures $T_s = [260, 320]$.

Clearly the calculation of A_n is extremely sensitive to the input value of surface temperature. In extreme heat, C3 photosynthesis rates tend towards zero, while as temperatures get cooler the model predicts photosynthesis to significantly increase. The local minimum and maximum present in Figure 2.2 are likely due to the change in sign around S_2 and S_4 within the exponential of $f(T)$; see (2.8). In general, model application to C3 species results in more extreme behavior as T_s varies compared to application to the C4 species. Notice that there is no feasible solution to the C4 model when $T_s > 300$ K. This temperature range is typical, and the inability to calculate a feasible solution within this range makes practical utilization of the model difficult at best.

2.3.2 Canopy Temperature T_c

T_c is used in the calculation of mesophyllic conductance (g_m), which yields no restrictions to the possible range of T_c. We examine the region $T_c = [265, 320]$ as shown in Figure 2.3.

Both the C3 and C4 models respond similarly to a change in T_c, with a slight monotonic decrease in A_n as the value of T_c increases and a pronounced decrease in A_n with extremely high canopy temperatures.

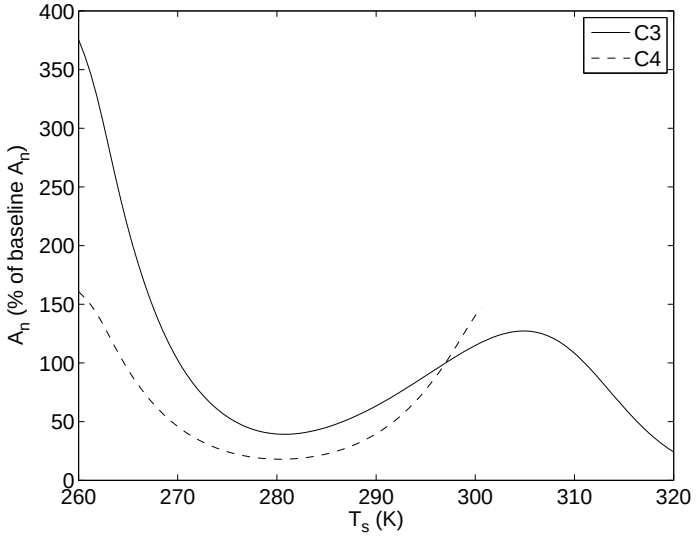


Figure 2.2: Response of A_n to surface temperature T_s.

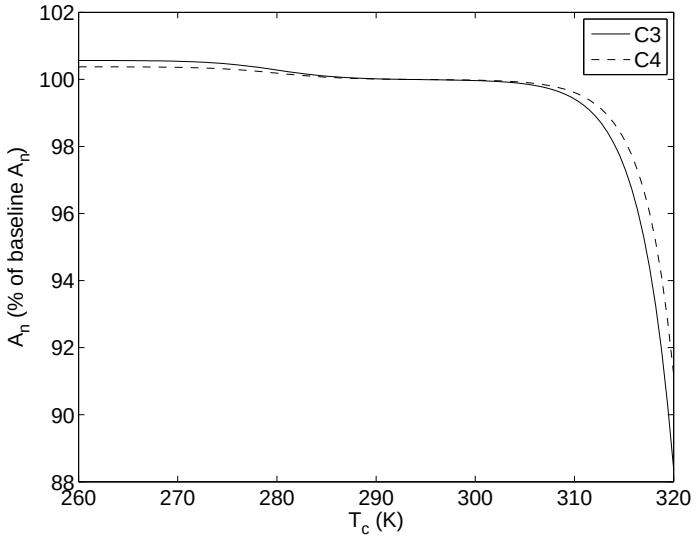


Figure 2.3: Response of A_n to canopy temperature T_c.

2.3.3 Ambient Temperature T_a

T_a is used in the calculation of the virtual ambient temperature T_{va} and the forced convection leaf boundary layer conductance g_{bfc}. Both of these variables impact the value of leaf boundary layer conductance g_b. As with the preceding other two temperature parameters, we examine the range $T_a = [260, 320]$ as shown in Figure 2.4.

Compared to the sensitivity exhibited to the other temperature parameters T_s and T_c, A_n is not as reactive to the value of T_a. Here A_n increases with increasing T_a. For C3, this is nearly a linear increase where for C4 it appears to be approaching an asymptote.

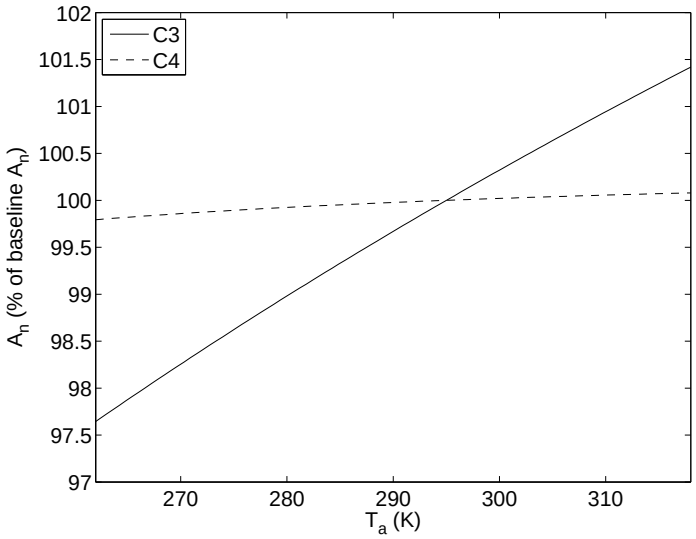


Figure 2.4: Response of A_n to ambient temperature T_a.

2.3.4 Surface Pressure P

Standard air pressure at sea level, approximately 1010 millibars or 101 kPa, indicates a typical value for P. Note that P is included in the calculation of the boundary layer conductance terms g_{bfc} and g_{bfr}, stomatal conductance g_s, net assimilation rate A_m, CO_2 partial pressure at the leaf surface C_s, and in intercellular spaces C_i for both models. Additionally, the PEP-Carboxylase limitation w_s for C4 species also depends on P.

As shown in Figure 2.5, application of the model to the C3 species shows a slight, linear decline in A_n as P increases. Because P is additionally involved in the limitation of the gross photosynthetic rate in the C4 species, the impact of changing P is slightly more important and nonlinear in the C4 as compared to the C3 species.

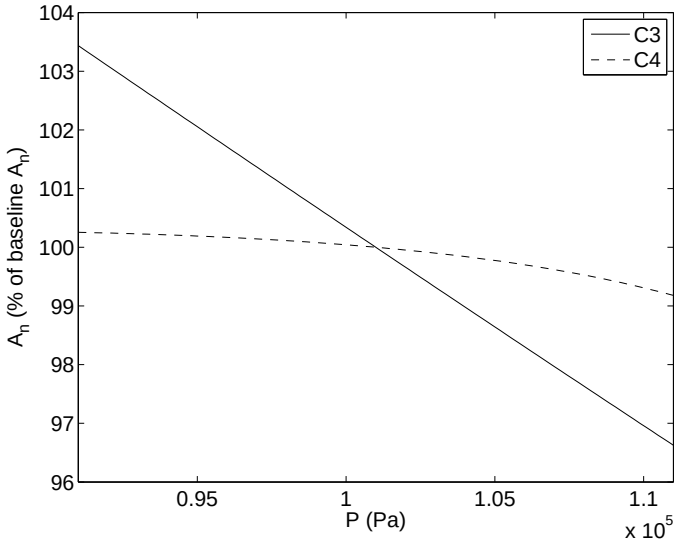


Figure 2.5: Response of A_n to surface pressure P.

2.3.5 Wind Speed u

Typical values for u can range from 0 - 10 m/s. The wind speed is only used in the calculation of g_{bfc} in the model. The leaf boundary layer conductance g_b is defined as the maximum of the free (g_{bfr}) and forced (g_{bfc}) conditions. Of all the parameters involved in the calculation of these two terms, the only instance where g_b is defined by g_{bfr} is when the value of u is less than 0.4 m/s. Therefore, when $u < 0.4$, g_b is defined by g_{bfr} which is independent of u; hence there is no change in A_n in this range of u.

Figure 2.6 shows that A_n exhibits strong sensitivity to the wind speed regardless of species. When free conditions define the leaf boundary layer conductance g_b, at $u < 0.4$ m/s, the value of A_n is reduced by more than 40% as compared to our baseline value of $u = 5$ m/s. As wind speed increases and forced conditions drive g_b, the net photosynthetic rate also increases. C3 species respond with continued increasing A_n through the maximum limit of the examined range (u = 10 m/s). However C4 species appear to reach a maximum value for A_n which changes little once wind speed reaches approximately 3 m/s. Also notice that there is no feasible solution to the C4 model when $u < 1$ m/s. This environmental condition is typical, and the inability to calculate a feasible solution within this range makes practical utilization of the model difficult at best.

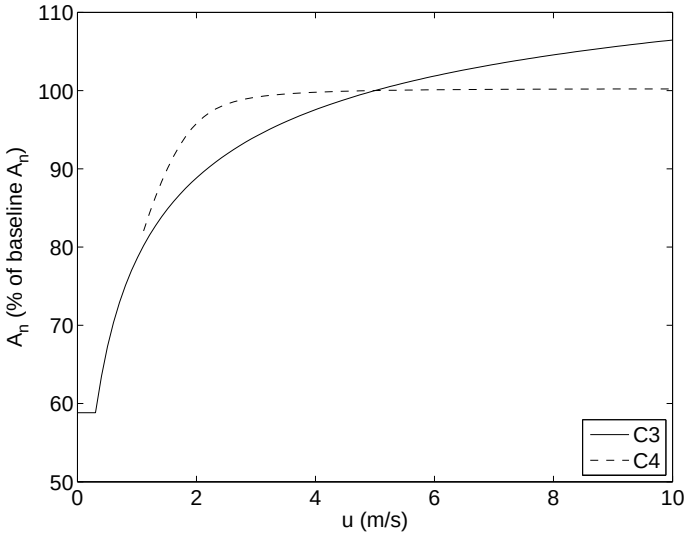


Figure 2.6: Response of A_n to wind speed u.

2.3.6 Oxygen Availability in Leaf Cells O_2

The utilized value for O_2 in several published models [21, 29, 47, 59] ranges from 20.9 to 21 kPa. Figure 2.7 illustrates the sensitivity of A_n to perturbing the baseline value of 2.09×10^4 Pa by 10%.

As expected, the C4 species exhibits nearly no change in A_n when varying this parameter. Recall that only the C3 species is impacted by photorespiration and therefore increased oxygen levels produce more oxygenase activity effectively inhibiting photosynthesis. This effect is clearly illustrated in Figure 2.7 where increasing levels of O_2 cause a reduction in A_n and reciprocally decreasing levels of O_2 allow A_n to increase.

2.3.7 Photosynthetically Active Radiation PAR

Photosynthetically active radiation is typically 45-65% net radiation depending on cloud cover [43]. On a non-cloudy day, net radiation is around 600 W/m^2, but on any given day can be between 400 and 1000 W/m^2. To convert from the energy units, W/m^2, to the quantum units employed in the model, mol m^{-2} s^{-1}, we use the relationship

$$E = \frac{hc}{\lambda} \tag{2.64}$$

where E = energy per photon (Joules), h = Planck's constant = 6.63×10^{-34} J $\cdot$ s, c = speed of light = 3×10^8 m/s, and λ = wavelength (m). Note also that 1 mol pho-

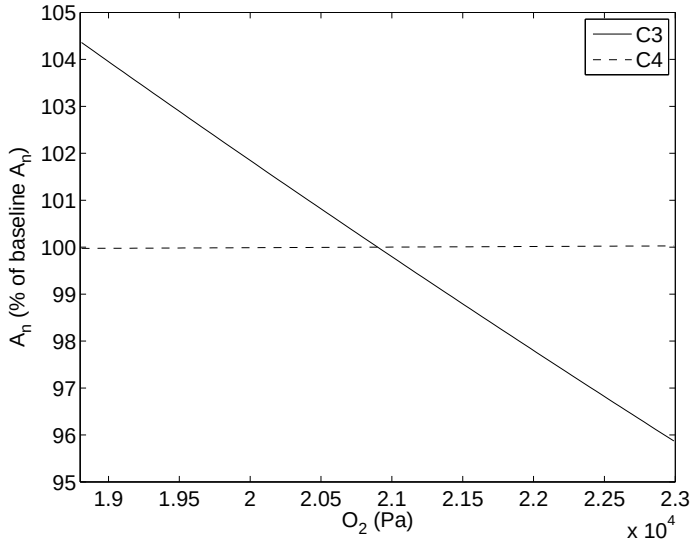


Figure 2.7: Response of A_n to oxygen availability in leaf cells O_2.

tons $= 6.02 \times 10^{23}$ photons (Avagadro's number). Assuming daylight has a wavelength of approximately 550 nm, which is midway in the range for PAR of 400-700 nm, then we have $\frac{1}{E} = \frac{\lambda}{hc} = \frac{550\times10^{-9}\text{m}}{(6.63\times10^{-34}\text{J}\cdot\text{s})(3\times10^{8}\text{m/s})} = 2.7652 \times 10^{18}$ photons per Joule. Therefore, in daylight, we have $\frac{2.7652\times10^{18}\text{photons/Joule}}{6.02\times10^{23}\text{photons/mol}} = 4.6 \times 10^{-6}$ mol/Joule. Since 1 Watt = 1 J/s, we can simply convert any value from W/m^2 to mol m^{-2} s^{-1} by multiplying by 4.6×10^{-6}. Thus we examine the range of PAR associated with $0.45 \cdot 400 = 180$ and $0.65 \cdot 1000 = 650$ W/m^2, or 8×10^{-4} to 3×10^{-3} mol m^{-2} s^{-1}.

PAR is used in the calculation of the light limiting factor w_e of A_g for both the C3 and C4 species, although the species require different formulations; see (2.22) and (2.28). Figure 2.8 illustrates that A_n is not remarkably sensitive to the amount of light. This implies that the remaining two limiting mechanisms, w_c and w_s, are more critical and that plants are not fully utilizing the light available.

2.3.8 Relative Humidity at Leaf Surface h_s

Normal atmospheric values for h_s fall between 0.2 and 0.99. We chose to examine the full range of possibility for h_s, from 0 to 1, as illustrated in Figure 2.9. Relative humidity is used in the model to relate stomatal conductance, net photosynthesis rate, and the CO_2 concentration at the leaf surface; see (2.1). Note that h_s is also used to calculate virtual temperatures via Teten's formula; see (2.13).

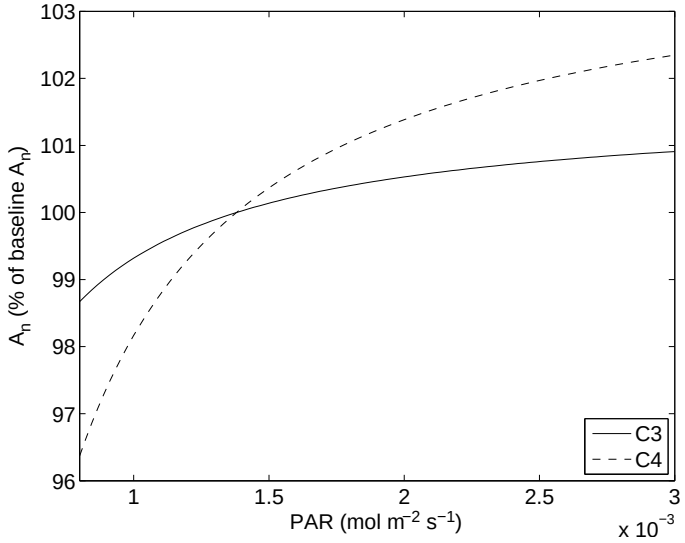


Figure 2.8: Response of A_n to photosynthetically active radiation PAR.

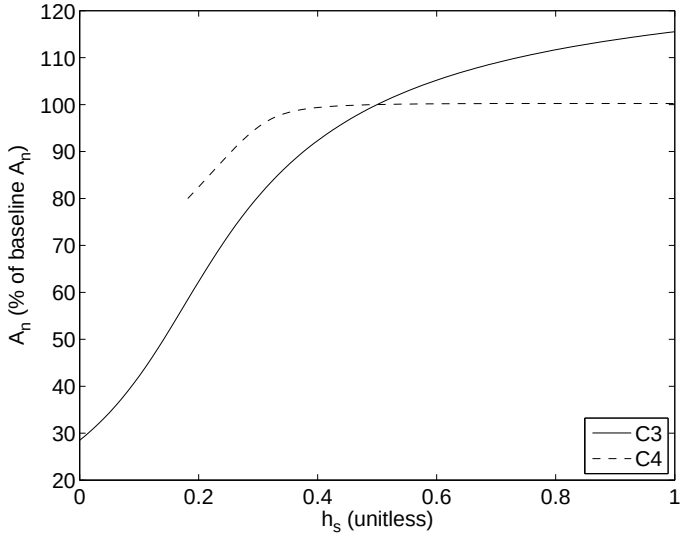


Figure 2.9: Response of A_n to relative humidity at leaf surface h_s.

The model exhibits a great deal of sensitivity to relative humidity. In particular, as levels approach 0 there is a significant decrease in A_n, ultimately resulting in approximately a 70% decrease for the C3 species. Although increasing relative humidity levels results in increasing A_n in both species, we again see the phenomenon of C4 species approaching a maximum possible value while the C3 species appears to continually increase A_n in response. For C4 there is no feasible solution to the C4 model when $h_s < 0.2$. This environmental condition can occur, and the inability to calculate a feasible solution within this range makes practical utilization of the model difficult at best.

2.3.9 Ambient CO_2 Partial Pressure C_a

The utilized value for C_a in several published models [2, 21, 37, 59, 66] ranges from 33 to 36 Pa. The sensitivity of photosynthetic rate to this parameter is especially of interest due to current concerns regarding rapidly increasing atmospheric CO_2 levels. Understanding how plants respond to these changing environmental conditions will aid scientists in modeling the phenomenon and possibly searching for ways to mitigate against negative climate impacts the increasing CO_2 levels may instigate.

The model uses C_a in the computation of carbon dioxide partial pressure at the leaf surface, C_s. As shown in Figure 2.10, increasing ambient CO_2 levels spurs an increase in A_n for both species. The response is more significant in the C3 species and it appears that continually increasing C_a will simply continually increase A_n. However, in the C4 species the model predicts an approaching maximum A_n regardless of the increase in C_a. If the model predictions are accurate, this implies that increasing the population of C3 species, rather than C4 species, worldwide may help to regulate environmental ambient CO_2 levels.

2.3.10 Root Level Soil Moisture Wilting Value w_{wilt}, Field Capacity Value w_{fc}, and Deep Soil Moisture Content w_2

Note that w_{wilt} is the soil moisture percentage where the plant permanently wilts while w_{fc} represents the maximum water percentage which a soil can hold against the forces of gravity. These parameters are dependent on soil type and may vary over time. Table 2.3 indicates values of w_{wilt}, w_{fc}, and w_{sat} (saturated moisture content) based on the soil classification of Clapp and Hornberger [20].

Several other published models consider the impact of soil moisture through the utilization of these soil-type based parameters. Calvet et al. [16] report values of w_{wilt} within the range of 0.15 - 0.34 m^3 water/m^3 soil and values for w_{fc} ranging from 0.32 - 0.43 m^3 water/m^3 soil [16, 17]. Others use a simple linear relationship to calculate w_{wilt} and w_{fc} values from w_{sat}

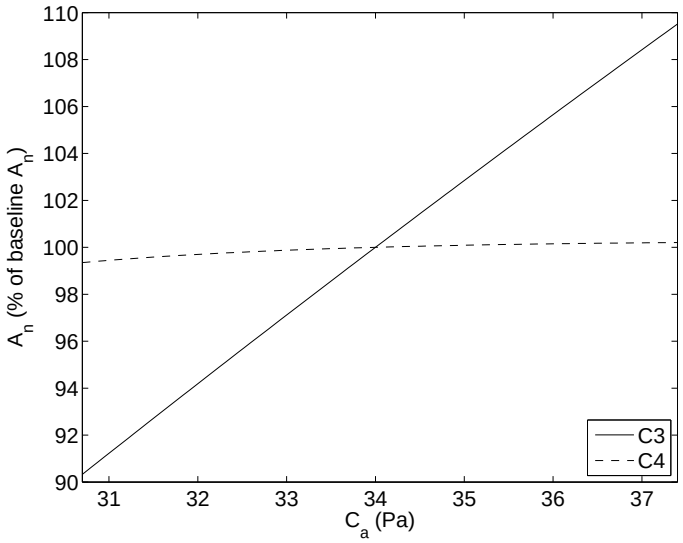


Figure 2.10: Response of A_n to ambient CO_2 partial pressure C_a.

values; particularly, they employ $w_{wilt} = 0.3 \cdot w_{sat}$ and $w_{fc} = 0.6 \cdot w_{sat}$ [51] or a relationship based on the clay content of soil [12].

The soil related parameter values of w_{wilt}, w_{fc}, and w_2 are only used during the calculation of $f(w_2)$ in this model by way of the relation

$$f(w_2) = \frac{w_2 - w_{wilt}}{w_{fc} - w_{wilt}}. \tag{2.65}$$

To satisfy $0 < f(w_2) < 1$, we require $w_{wilt} < w_2 < w_{fc}$. In general, as $f(w_2) \to 0$, $w_2 \to w_{wilt}$ and as $f(w_2) \to 1$, $w_2 \to w_{fc}$.

Note that $f(w_2)$ represents the soil moisture dependency for the catalytic Rubsico capacity for the leaf, V_m. In Figure 2.11, we illustrate the sensitivity of A_n to the value of $f(w_2)$. Clearly the model has strong sensitivity to the level of soil moisture. Under dry conditions, the model predicts A_n will approach 0. Dry conditions will influence the stomata to remain closed to minimize plant water loss, while at the same time minimizing the influx of CO_2, so this effect to A_n is realistic. Notice there is no feasible solution to the C4 model when $f(w_2) > 0.8$. The inability of the model to calculate a feasible solution for this range makes practical utilization of the model difficult at best.

Table 2.3: Soil moisture parameter values from Jacquemin and Noilhan [34]. Note that w_{fc} values are associated with a hydric conductivity of 0.1 mm/day and w_{wilt} values correspond to a moisture potential of -15 bar.

Soil type	w_{fc} (m^3/m^3)	w_{wilt} (m^3/m^3)	w_{sat} (m^3/m^3)
Sand	0.135	0.068	0.395
Loamy sand	0.150	0.075	0.410
Sandy loam	0.195	0.114	0.435
Silt loam	0.255	0.179	0.485
Loam	0.240	0.155	0.451
Sandy clay loam	0.255	0.175	0.420
Silty clay loam	0.322	0.218	0.477
Clay loam	0.325	0.250	0.476
Sandy clay	0.310	0.219	0.426
Silty clay	0.370	0.283	0.482
Clay	0.367	0.286	0.482

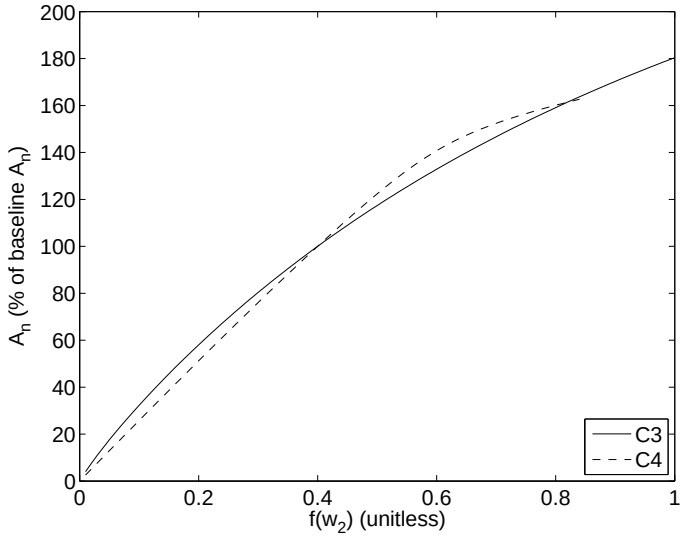


Figure 2.11: Response of A_n to $f(w_2)$.

2.3.11 Quantum Efficiency for Carbon Dioxide Uptake ϵ

The quantum efficiency ϵ is involved in the calculation of the light limited rate of assimilation (w_e) for both the C3 and C4 species, although each species requires different formulations; see (2.22) and (2.28). Collatz et al. [21, 22] present the cited C3 and C4 values, 0.08 mol/mol and 0.05 mol/mol respectively, which are also referenced in [61].

Figure 2.12 illustrates that A_n is not very sensitive to perturbations in the value for ϵ for either species. This again implies, as seen before in the analysis for PAR, that A_g is not limited by light and hence plants are not fully utilizing the potential of available light.

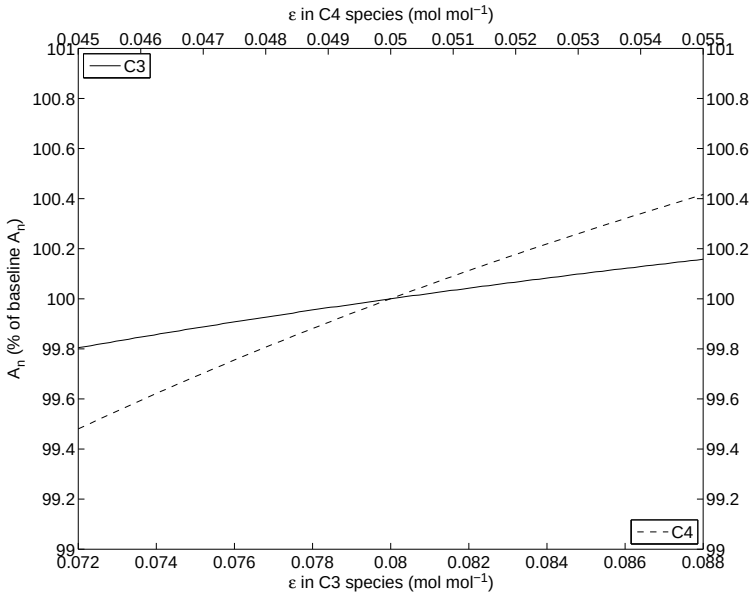


Figure 2.12: Response of A_n to quantum efficiency for carbon dioxide uptake ϵ.

2.3.12 Leaf-Scattering Coefficient for PAR w_π

The leaf-scattering coefficient for PAR, w_π, also is involved in the calculation of the light-limited rate of assimilation (w_e). Although the notation differs slightly, this concept seems to originate in the Collatz et al. models for plant photosynthesis where $w_\pi = 1 - a$ and a represents leaf absorptance to photosynthetically active quantum flux density. Using this relationship, they present $w_\pi = 0.14$ in C3 plants and 0.2 in C4 plants [21, 22]. The C4 value seems to be approximately related to the measurements of PAR absorptance recorded by Norman and Polley in C4 plants [52]. Values used by Niyogi et al. [48] are 0.1 for C3 plants and 0.2 for C4

plants.

The sensitivity of A_n to w_π, as shown in Figure 2.13, is fairly insignificant for both C3 and C4 species. Because w_π is only used in the calculation of w_e, as was similarly implied in the analysis for PAR and ϵ, the model suggests that plants of both species are not fully utilizing the available light, and A_n is being limited by other factors.

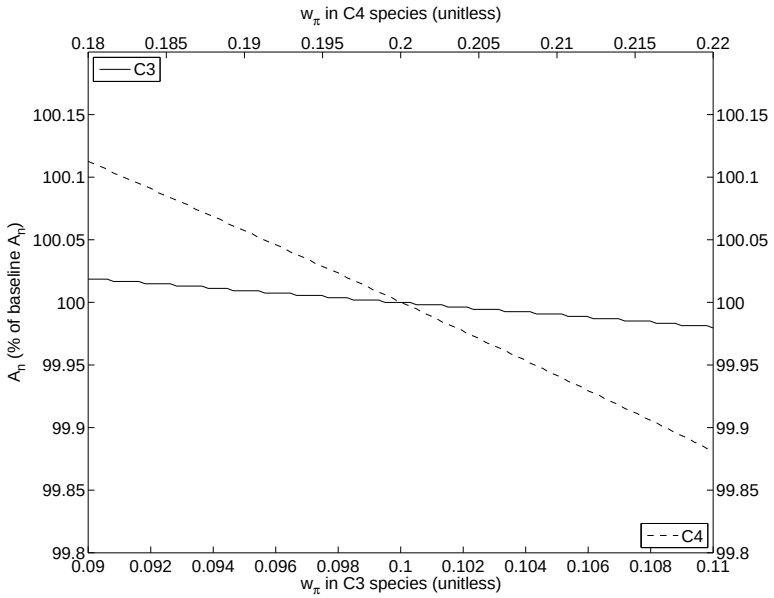


Figure 2.13: Response of A_n to leaf-scattering coefficient for PAR w_π.

2.3.13 Factor to Account for Different Diffusivities of H_2O and CO_2 in the Stomatal Pores η

Here η is the ratio of stomatal conductance to water vapor to that for CO_2 [5]. It is used in this model when applying Fick's law of diffusion to calculate the partial pressure of CO_2 in the intercellular spaces (C_i).

Figure 2.14 shows that as η increases, A_n decreases slightly in both C3 and C4 species. From (2.17) we see that increasing η will decrease the intercellular CO_2 partial pressure C_i which reduces that value of w_c in both species and w_e for the C3 species as shown in (2.5) and (2.22), respectively. The reduction of these limiting factors for A_g in turn decreases the value of A_n. Therefore the trends illustrated here are plausible.

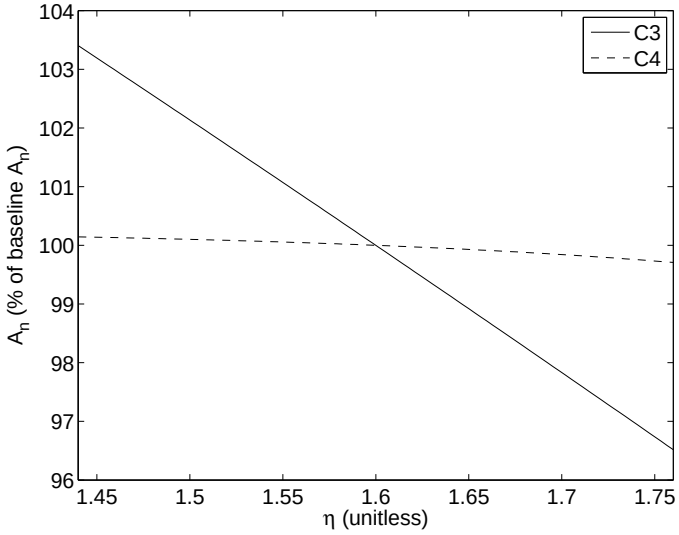
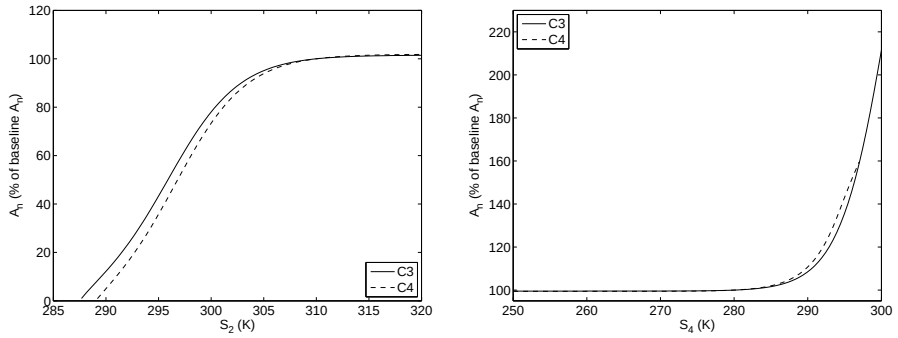


Figure 2.14: Response of A_n to the factor to account for different diffusivities of H_2O and CO_2 in the stomatal pores, η.

2.3.14 Vegetation Stress Factors, S_2 (high) and S_4 (low)

S_2 and S_4 are factors used to incorporate temperature-dependence into the calculations for catalytic Rubisco capacity for the leaf (V_m) and mesophyllic conductance (g_m). Sellers et al. [61] lists multiple values for each of these vegetation dependent parameters. Values for S_2 were within the range of 303-313 K for the evaluated samples and S_4 values ranged from 278 to 288 K.

Baseline values for both species, as noted in Table 2.2, are $S_2 = 310$ K and $S_4 = 280$ K. Figures 2.15(a) and (b) present the response of A_n to S_2 and S_4, respectively. As the value of S_2 decreases and approaches the value of S_4, A_n decreases in both species. Raising the S_2 value above it's baseline appears to have little effect on A_n in either species as an asymptotic maximum is present soon thereafter. Analogously, as the value of S_4 increases and approaches the value of S_2, A_n increases in both species. Reducing the S_4 value below it's baseline results in a nominal decrease in A_n which reaches steady state below approximately 275 K. These effects could be due to the expressions containing S_2 and S_4 being located on opposite sides of the divisor as indicated in (2.8) and (2.26).



(a) (b)

Figure 2.15: Response of A_n to vegetation stress factors (a) S_2 and (b) S_4.

2.3.15 Maximum Catalytic Rubisco Capacity for Leaf V_{max}

V_{max} is used in the calculation of the catalytic capacity for the leaf, V_m, which is used to find in both species the Rubisco limited assimilation rate (w_c), the C3 plant capacity to utilize photosynthesis products (w_s), and the C4 PEP-Carboxylase limited (w_s) assimilation rate. V_{max} is described as a physiological property of the leaf which is proportional to the Rubisco reserves of the leaf, and therefore it's nitrogen content [60]. This value differs for C3 and C4 vegetation and it appears the baseline values employed in Niyogi et al. [48] are approximated from the measurements of Wilson et al. [66] assuming C4 plants comprise the understory and C3 plants make up the overstory for the canopies under study. The measurements of V_{max} exhibit great variability caused by light environment, species type, and seasonal effects on leaves.

Figure 2.16 illustrates the impact on A_n when perturbing the species-specific V_{max} value by 10%. The top x-axis denotes the values in the C4 species and the bottom x-axis denotes the values in the C3 species. The general trend for both species is the same, that is, increasing V_{max} increases A_n. The sensitivity in C4 species is slightly more pronounced than in the C3 species. Notably A_n is fairly sensitive to this parameter indicating the importance of w_c and w_s as limiting factors to the assimilation rate.

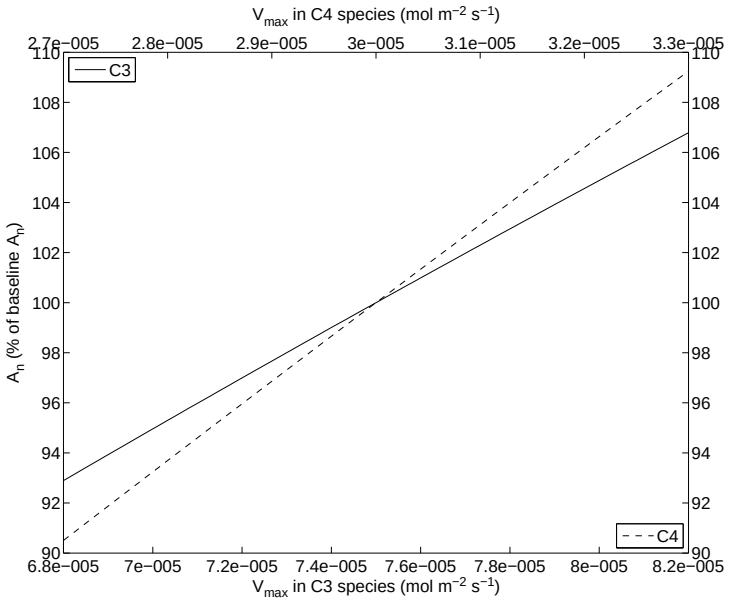


Figure 2.16: Response of A_n to maximum catalytic Rubisco capacity for leaf V_{max}.

2.3.16 Maximum Net Assimilation Rate $A_{m,max}$

$A_{m,max}$ is used to calculate the maximum net assimilation rate, A_m, to be used to find the dark respiration rate R_d. In Figure 2.17, we show the sensitivity of A_n when perturbing the value of $A_{m,max}$. Changing the value of this parameter has essentially no effect on the value of A_n for either species, possibly because these values for the maximum net assimilation rate are more than ten times larger than the value for the net assimilation rate A_n in either species.

It should be noted that others have computed $A_{m,max}$ to simply be a linear function of the maximum catalytic capacity of Rubisco, V_{max}. According to Jacobs [33], $A_{m,max}$ may be assumed to be proportional to V_{max}; that is, $A_{m,max} = k \cdot V_{max}$. For C4 species, $k = 1$ while for C3 species $k = 0.5$ [21, 22]. Therefore, this would imply we use a value for $A_{m,max} = V_{max} =$ 30 µmol m^{-2} s^{-1} for C4 and $A_{m,max} = 0.5 \cdot V_{max} = 37.5$ µmol m^{-2} s^{-1} for C3 plants. However, these $A_{m,max}$ values are still much higher than our A_n solutions, so modifying the formulation of $A_{m,max}$ to use this method has little affect on A_n.

Sensitivity analysis is also performed to determine the impact of a temperature-dependent formulation of $A_{m,max}$. Niyogi's model considers $A_{m,max}$ to be constant, where Jacobs includes the temperature dependency

$$A_{m,max} = \frac{x \cdot 2.0^{\frac{T_s - 298}{10}}}{(1 + e^{0.3(T_1 - T_s)})(1 + e^{0.3(T_s - T_2)})} \tag{2.66}$$

where $x = 1.7$ mg m^{-2} s^{-1}, $T_1 = 286$ K, and $T_2 = 311$ K for C4 plants and $x = 2.2$ mg m^{-2} s^{-1}, $T_1 = 281$ K, and $T_2 = 311$ K for C3 plants. Because the model requires $A_{m,max}$ to have molar units, we convert x to 38.6364 µmol m^{-2} s^{-1} for C4 plants and 50 µmol m^{-2} s^{-1} for C3 plants.

Figure 2.18 compares the response of A_n, when using the constant value of $A_{m,max}$, to the implementation of the temperature-dependent formulation in (2.66). Although the sensitivity shown in this figure seems substantial, the results are nearly identical to the model sensitivity to T_s (see Figure 2.2). Therefore, we may conclude that using the temperature-based formulation of $A_{m,max}$ as shown in (2.66) does not substantially affect A_n as opposed to considering it as a constant.

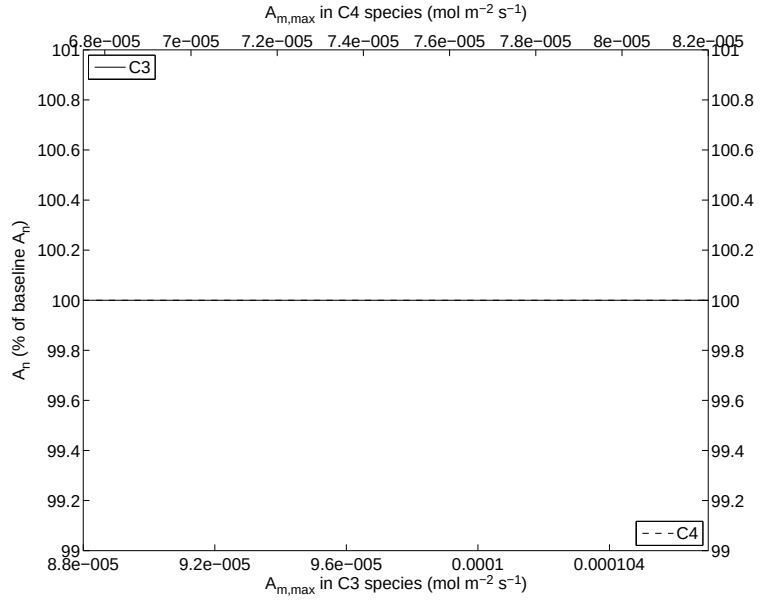


Figure 2.17: Response of A_n to maximum net assimilation rate $A_{m,max}$.

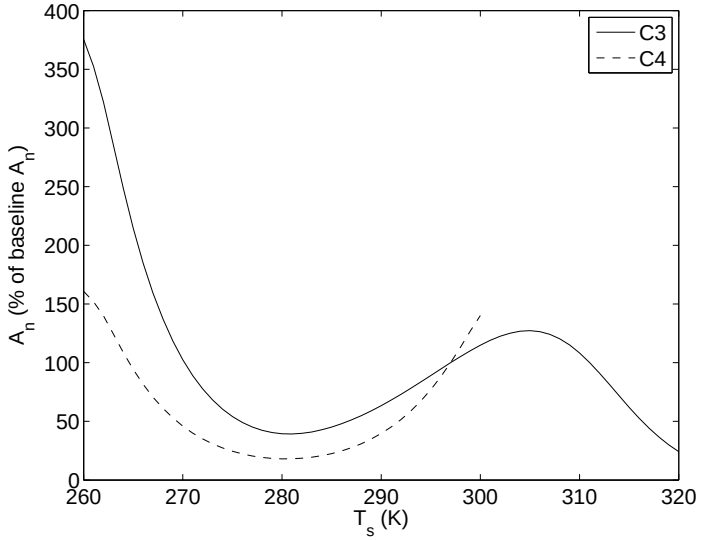


Figure 2.18: Response of A_n to a different $A_{m,max}$ formulation based on temperature.

2.3.17 Potential Maximum Value of Mesophyllic Conductance g_{mp}

In this model, g_{mp} is denoted as a potential maximum value to modulate the mesophyllic conductance, g_m. The value for g_{mp} is referenced to [33]; however, they do not use a value representing a maximum in their g_m formulation. The C3 species value of $g_{mp} = 7. \times 10^{-3}$ m/s and C4 species value of $g_{mp} = 17.5 \times 10^{-3}$ m/s seems to originate from the g_m values at $T_s =$ 298 K, which are in fact not the maximum g_m values possible in the Jacobs [33] model.

Figure 2.19 illustrates the impact on A_n when perturbing the species-specific g_{mp} value by 10%. The top x-axis denotes the values in the C4 species and the bottom x-axis denotes the values in the C3 species. The general trend for both species is the same, that is, changes in g_{mp} have little affect on A_n. This indicates that either the maximum net assimilation rate A_m is not sensitive to mesophyll conductance or that A_n is not very sensitive to changes in R_d.

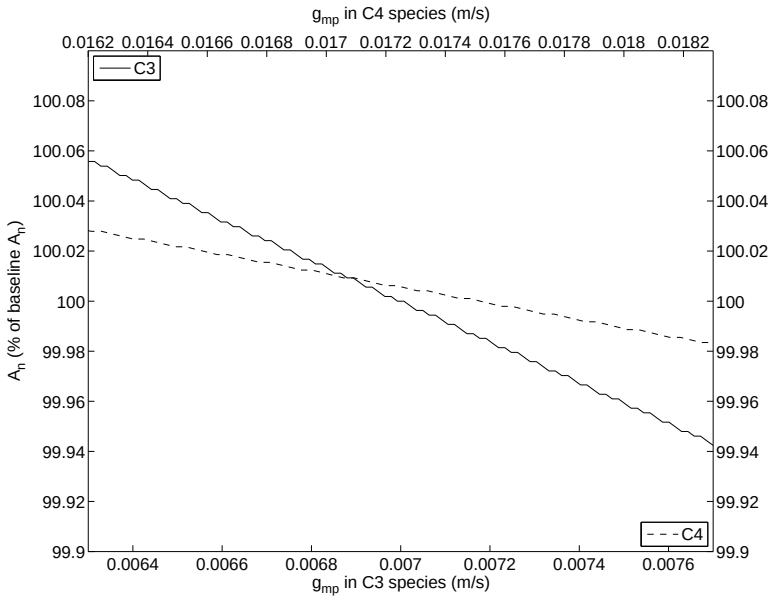


Figure 2.19: Response of A_n to potential maximum value of mesophyllic conductance g_{mp}.

2.3.18 Transfer Coefficient c

This transfer coefficient is used in the calculation of boundary layer conductance under free (g_{bfr}) and forced (g_{bfc}) convective conditions in the model. Niyogi's model considers the value of c to remain constant in the computations for g_{bfc} and g_{bfr}. However, Nikolov et al. [47] uses different values of c amongst these. Further, both Niyogi et al. and Nikolov et al. present unique values for c according to leaf type but not species type, meaning there are different values prescribed for broad leaf and coniferous shoots. Of the four values presented in Nikolov et al., the minimum is 0.8669×10^{-3} and the maximum is 4.322×10^{-3}.

Figure 2.20 presents the sensitivity of A_n to c where we use the same c in both the calculation for g_{bfc} and g_{bfr}, as presented in Niyogi et al. Here we examine the range of 10% below the minimum through 10% above the maximum indicated for c in Nikolov et al. For both species, there is a significant decrease in A_n as c decreases. In the model, decreasing c causes a decreased leaf boundary layer conductance g_b. This in turn decreases the intercellular CO_2 partial pressure C_i which increases R_d and decreases A_n. Notice there is no feasible solution to the C4 model when $c < 0.002$.

We summarize the impact of using different c values in the calculation of g_{bfc} and g_{bfr} in Table 2.4. As indicated in the sensitivity analysis for u, g_b is only defined by g_{bfr} when $u < 0.4$ m/s. Since we are using the baseline value of $u = 5$ m/s, g_{bfr} has no impact on the calculation of A_n. However, for future implementations of the model, where $u < 0.4$, using separate values for c could have a significant impact.

Table 2.4: Comparison of A_n values for different uses of c for C3 plants.

	c in g_{bfc}	c in g_{bfr}	A_n (µmol m^{-2} s^{-1})
Broad leaf example	4.322×10^{-3}	4.322×10^{-3}	5.3791
	4.322×10^{-3}	1.6361×10^{-3}	5.3791
Coniferous shoot example	1.2035×10^{-3}	1.2035×10^{-3}	3.3775
	1.2035×10^{-3}	0.8669×10^{-3}	3.3775

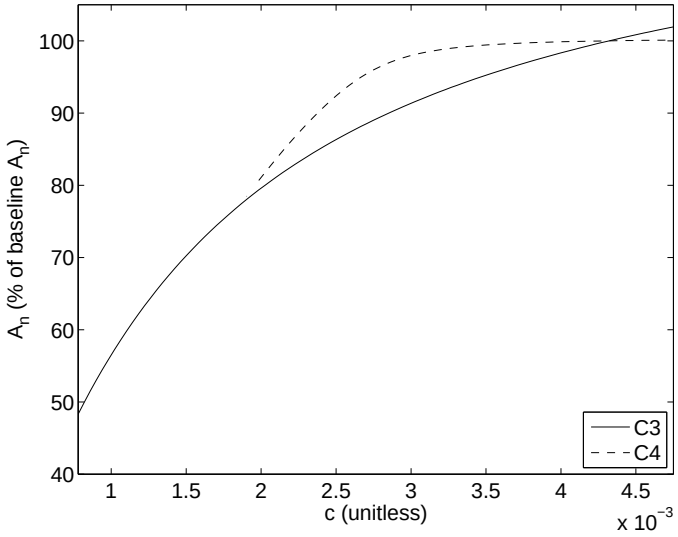


Figure 2.20: Response of A_n to transfer coefficient c.

2.3.19 Leaf Length Scale, d

Niyogi's model considers the value of d to remain constant in the computations for g_{bfc} and g_{bfr}. However, Nikolov et al. [47] used different values of d depending on if the plant is a broad leaf or coniferous species. For broad leaf species, in both instances d is defined to be leaf width. However, in the computation of g_{bfc} the needle diameter is used whereas in the computation of g_{bfr}, the shoot diameter is used. The reason for using the shoot diameter in free convection conditions is because in the absence of forced air flow, the shoot behaves as an intact object rather than a collection of needles [47]. Analogous to the analysis for c, since we are using the baseline value of $u = 5$ m/s > 0.4 m/s, g_{bfr} has no impact on the calculation of A_n. However, for future implementations of the model where $u < 0.4$ using separate values for d could have a significant impact.

Figure 2.21 shows the impact of leaf width from 1 mm to 10 cm on A_n. For both species, increasing leaf width results in a decreased A_n value. This is a result of d having the reciprocal effect that c has as explained in the previous section; see (2.10) and (2.11). The impact of reducing leaf size from the baseline value has a more significant impact on the C3 species.

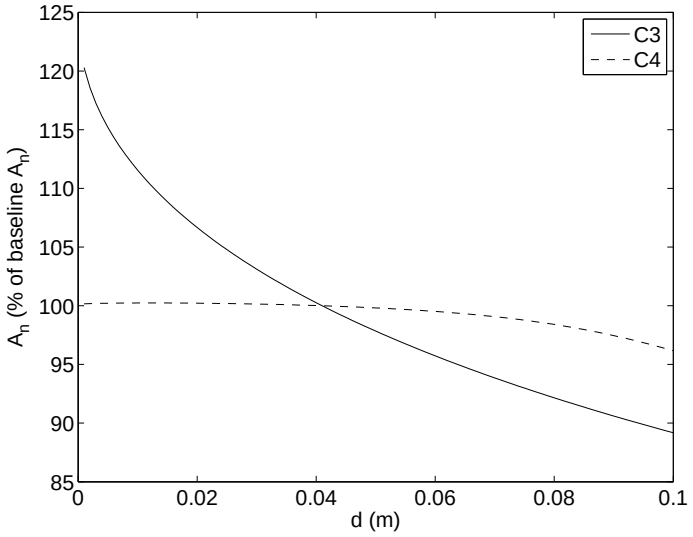


Figure 2.21: Response of A_n to leaf length scale d.

2.3.20 Coupling Coefficients, β_1 and β_2

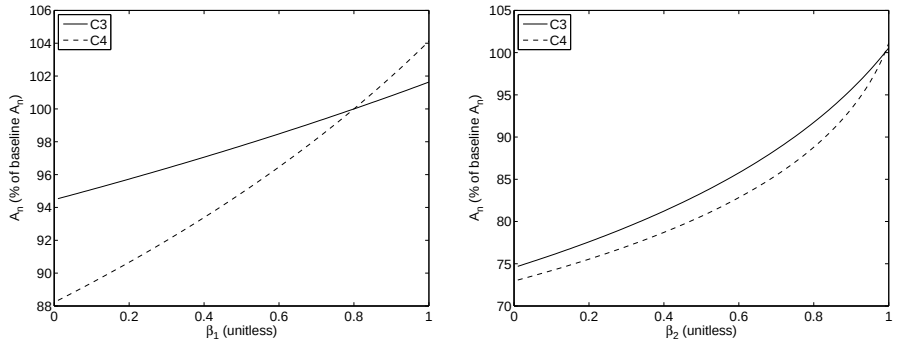
The coupling coefficients β_1 and β_2, whose values may range from 0 to 1, regulate the smoothness of transition between the three limiting factors which regulate photosynthesis. The limiting factors in C3 vegetation are (i) efficiency of Rubisco (w_c), (ii) light limitation (w_e), and (iii) plant capacity to utilize photosynthesis products (w_s). The limiting factors in C4 vegetation are (i) efficiency of Rubisco (w_c), (ii) light limitation (w_e), and (iii) PEP-Carboxylase limitation (w_s). The gross assimilation rate, A_g, is defined as the smallest root of the system

$$\begin{aligned} \beta_1 w_p^2 - w_p(w_c + w_e) + w_e w_c &= 0 \\ \beta_2 A_g^2 - A_g(w_p + w_s) + w_p w_s &= 0 \end{aligned} \tag{2.67}$$

where the coupling coefficients serve as smoothing terms.

Figures 2.22(a) and (b) illustrate the change in A_n as β_1 and β_2 vary from 0 to 1, respectively, and the other parameters are fixed to baseline values. Figures 2.23 and 2.24 present the change in A_n as both β_1 and β_2 vary for C3 and C4 plants, respectively. For both species, the value of A_n is much more sensitive to a change in β_2 than a perturbation of β_1. Given the system shown in (2.67), this implies that one of the terms, w_c or w_e, has little impact as a limiting factor.

Exploring this further, Table 2.5 presents the values for the limiting factors w_c, w_e, and w_s



(a) (b)

Figure 2.22: Response of A_n to coupling coefficients (a) β_1 and (b) β_2.

under baseline conditions. For both species, the model indicates that light is not a limiting factor for A_n as is evident by the w_e values being more than two-fold larger than the other two limiting factors. w_p is the interim variable representing the combination of w_c and w_e which is controlled by β_1. Clearly the value of w_p is dominated by the effect of w_c. A_g is then determined by w_p and w_s with curvature controlled by β_2. For both C3 and C4, w_p, as driven by w_c, is the smaller of these values implying that the gross photosynthetic rate is controlled by Rubisco limitation. This agrees with much of the sensitivity analysis we have seen thus far and suggests that if the model predictions are accurate, there is an opportunity to bioengineer the enzymatic capacities of plants to utilize the untapped potential of light energy.

Table 2.5: Comparison of values (µmol m^{-2} s^{-1}) for limiting factors and resultant assimilation rates.

	C3	C4
w_c	5.5788	10.252
w_e	55.300	55.2000
w_s	12.815	18.493
w_p	5.4592	9.8260
A_g	5.4195	9.7184
R_d	0.0405	0.0375
A_n	5.3791	9.6809

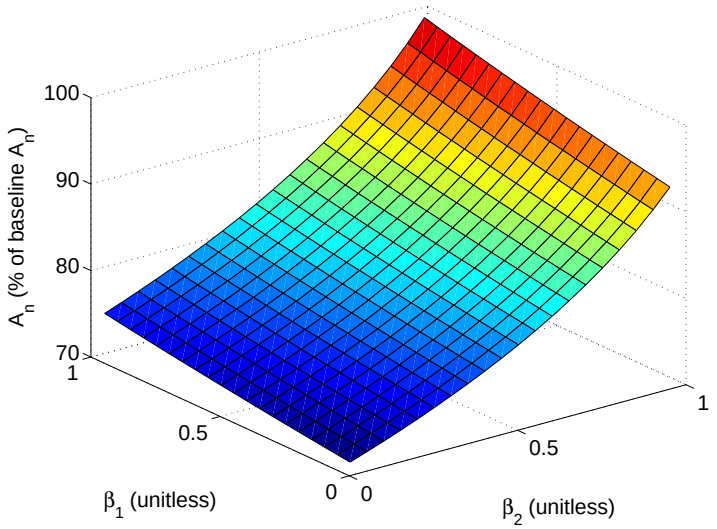


Figure 2.23: Response of A_n to coupled changes in β_1 and β_2 in C3 species.

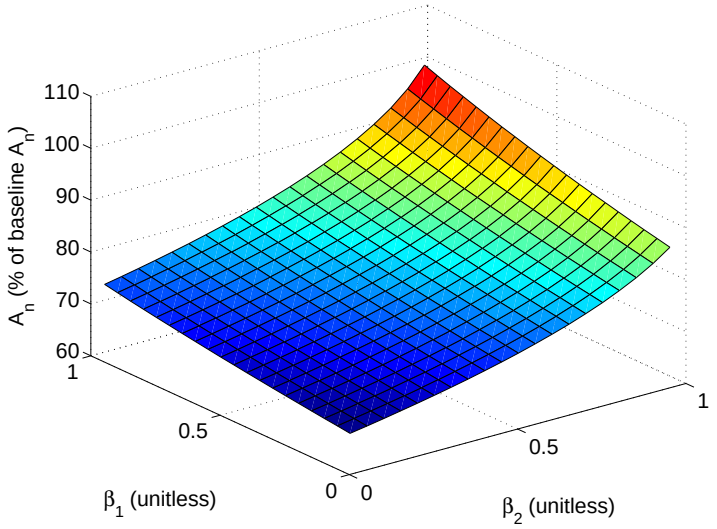


Figure 2.24: Response of A_n to coupled changes in β_1 and β_2 in C4 species.

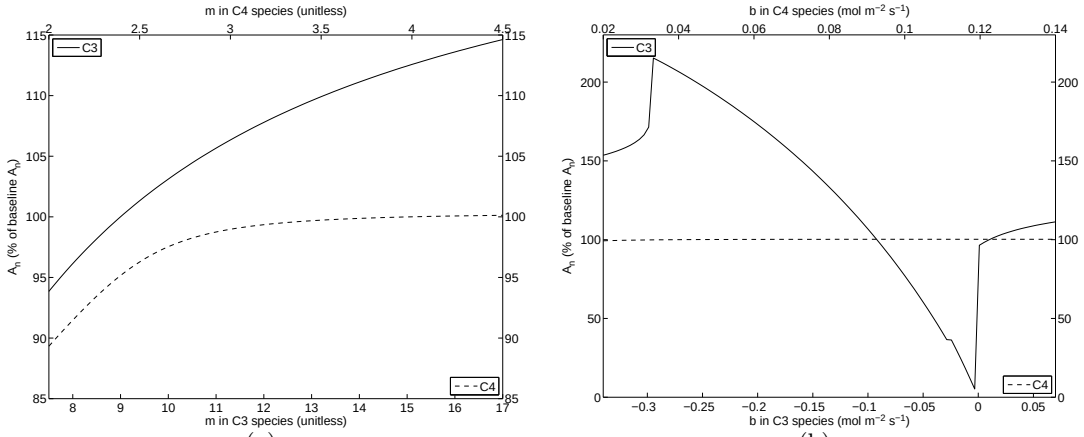
2.3.21 Slope m and Intercept b in Stomatal Conductance Equation

Several published models [2, 21, 60] cite the values of $m = 4$ and $b = 0.04$ for C4 species and $m = 9$ and $b = 0.01$ for C3 species, referencing the original groundbreaking work of Ball [5]. Other still use the methods of Ball to arrive at C4 value ranges of $m = 2.4 - 3.6$ and $b = 0.03 - 0.13$ mol m^{-2} s^{-1} and C3 value ranges of $m = 8.0 - 16.5$ and $b = -0.31 - 0.07$ mol m^{-2} s^{-1} [5, 22, 47, 52, 63].

Noted to be species-specific, m as used in the original model by Ball is described as the composite sensitivity of stomatal conductance to assimilation, CO_2 concentration, humidity, and temperature. It is apparent from the system of equations presented in Section 2.1 that in this model, m further incorporates sensitivity to wind speed, ambient pressure, plant species and leaf length scale. The free parameter b is the stomatal conductance that remains unaffected by the atmospheric environment or leaf biochemistry.

In Figure 2.25(a), we examine the ranges $m = 8 - 17$ for C3 species and $m = 2 - 4.5$ for C4 species. Both species exhibit a similar response to changes in m, in particular as the value of m increases, A_n increases as well. Figure 2.25(b) presents the sensitivity of A_n for the range of $b = -0.35 - 0.07$ in C3 plants and $b = 0.02 - 0.14$ in C4 plants. C4 plants appear to exhibit little sensitivity to the value of b when all other parameters are fixed to baseline values. The response to perturbations in b in the C3 species is interesting, in part because the values were permitted to be negative for a portion of the range examined.

Figures 2.26 and 2.27 illustrate the change in A_n as both m and b vary for C3 and C4 plants, respectively. The interesting trend in A_n in consistent across all values of m as b changes from negative to positive values in C3 plants as indicated by Figure 2.26. Note that Figure 2.25(b) shows little impact on A_n when b is varied using baseline values in C4 plants, and Figure 2.27 confirms this affect for most evaluated m values. The exception where more significant sensitivity of A_n in C4 plants exists is for small values of m and b.



(a) (b)

Figure 2.25: Response of A_n to (a) slope m and (b) intercept b in stomatal conductance equation.

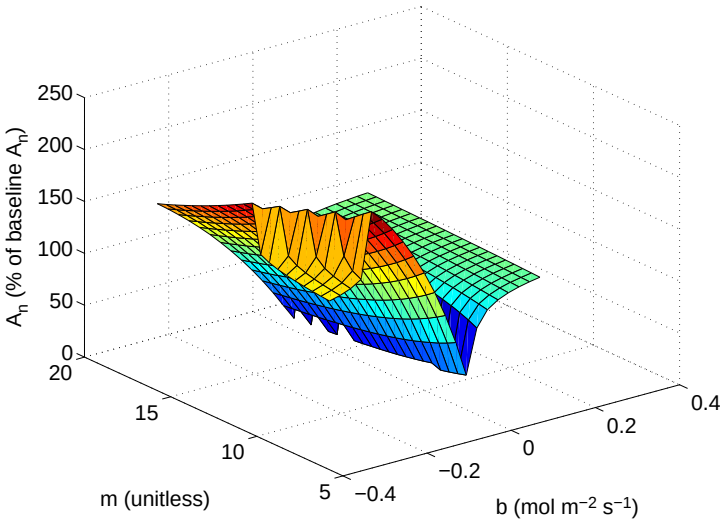


Figure 2.26: Response of A_n to m and b in C3 species.

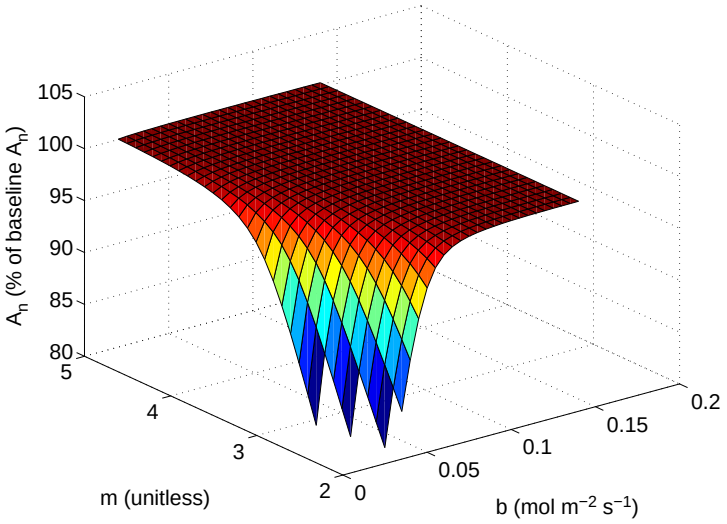


Figure 2.27: Response of A_n to m and b in C4 species.

2.3.22 Coefficient of w_s Equation

Recall that gross carbon assimilation rate A_g is computed as a result of the smoothed limit of three factors: w_c, w_e, and w_s. Although the Rubisco limitation w_c and light limitation w_e have different formulations between the C3 and C4 plant species, they in fact represent the same physical phenomenon. However, w_s represents different limitations based on plant species. For C3 plants, w_s represents the capacity of the plant to utilize the photosynthesis products while

in C4 plants, w_s quantifies the PEP-Carboxylase limitation on photosynthesis.

In C3 plants, the limitation by triose phosphate utilization, w_s, is simply defined as

$$w_s = \frac{V_m}{2}. \tag{2.68}$$

The coefficient of interest to determine sensitivity of A_n to is 1/2. In the C4 equation for PEP-Carboxylase limitation,

$$w_s = 2 \times 10^4 V_m \frac{C_i}{P}, \tag{2.69}$$

which is first presented in [60]; there is no detailed description for the origin of the coefficient 2×10^4.

Figure 2.28 presents the sensitivity of A_n when perturbing these coefficient values by 10%. Although both species have a slight increase in A_n as the coefficient of the w_s equation is increased, neither species is particularly sensitive to these coefficients. This is likely because w_s is not a limiting factor for either species as indicated in Table 2.5.

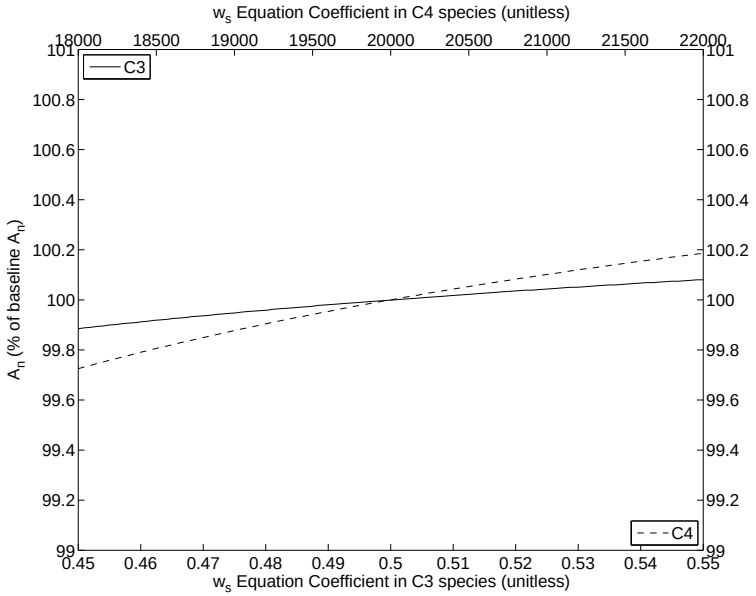


Figure 2.28: Response of A_n to coefficient of w_s equation.

2.3.23 Coefficient of R_d Equation

Along with other published models [16, 33], the definition of leaf respiration rate R_d is referenced to the original work of van Heemst [64]. van Heemst states that at normal temperatures, R_d satisfies the relation

$$R_d = \frac{A_m}{9}, \tag{2.70}$$

where A_m is the maximum net assimilation rate. Since no detailed evidence was presented to validate the choice of the coefficient 1/9, we examine the effect on A_n when perturbing the coefficient.

Figure 2.29 shows that neither species is incredibly sensitive to alteration of this value, although a slight decrease in A_n occurs as the coefficient increases. Given the definition of $A_n = A_g - R_d$, this impact is logical. Further, since Table 2.5 indicates that the magnitude of R_d is much smaller than A_g, the small change in A_n when changing the value of R_d is anticipated.

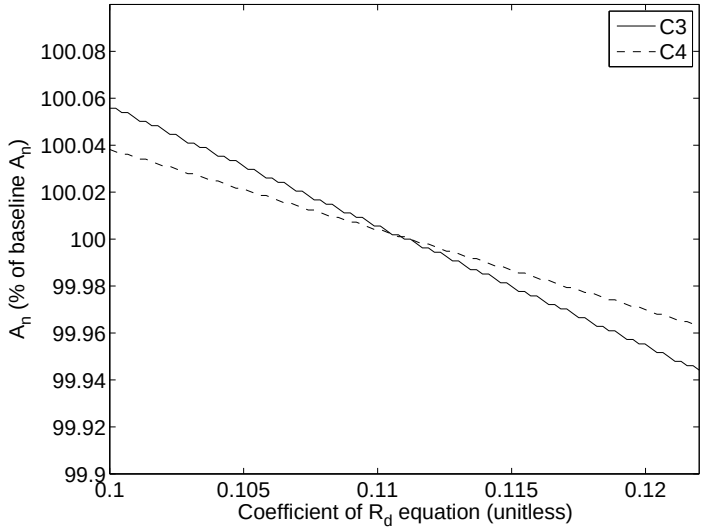
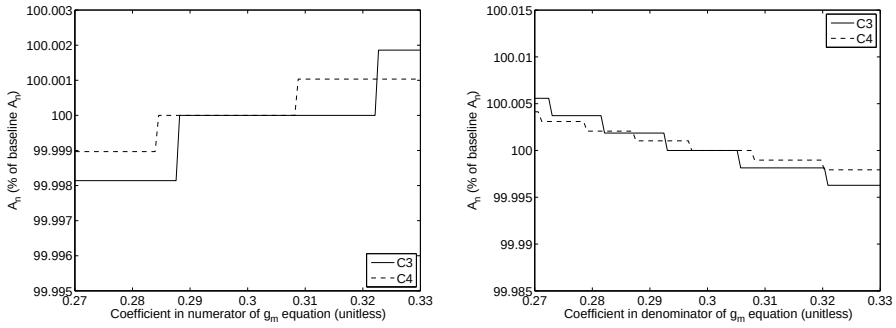
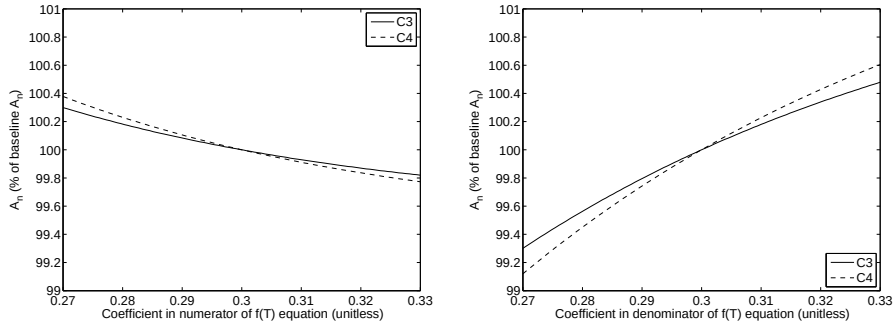


Figure 2.29: Response of A_n to coefficient of R_d equation.



(a) (b)

Figure 2.30: Response of A_n to coefficient in (a) numerator and (b) denominator of g_m equation.



(a) (b)

Figure 2.31: Response of A_n to coefficient in (a) numerator and (b) denominator of $f(T)$ equation.

2.3.24 Coefficients in Numerators and Denominators of g_m and $f(T)$ Equations

Although these coefficients (all with a value of 0.3) appear in multiple models [16, 33], no reference was found to indicate the origins of these values. Thus, Figures 2.30(a) and (b) illustrate the perturbation of the coefficients in the numerator and denominator of the equation for mesophyllic conductance g_m, respectively. Further, Figures 2.31(a) and (b) illustrate the perturbation of the coefficients in the numerator and denominator in the equation for temperature dependence for maximum catalytic Rubisco capacity, $f(T)$, respectively. Perturbing the coefficients by 10% had less than a 1% effect in both species for all occurances.

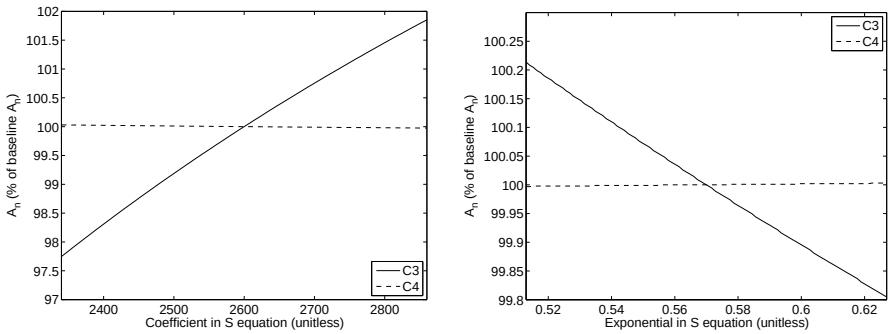
2.3.25 Constants in S Equation

The Rubisco specificity for CO_2 relative to O_2, S, is quantified by the relation

$$S = 2600 \cdot 0.57^{Q_{10}}. \tag{2.71}$$

S is then used to calculate the CO_2 compensation point Γ^* which is used to compute the maximum assimilation rate A_m in both plant species, as well as the limiting factors w_c and w_e in the C3 species. This exact formulation is presented in [21] and similar constant values are presented for this expression in [67]. However, no direct explanation was found to explain the determination of the constants 2600 and 0.57.

Therefore, Figures 2.32(a) and (b) illustrate the response of A_n to perturbation of these values by 10% in both species. As one would expect, the A_n predictions for the C4 species show little sensitivity to the calculation of S because O_2 does not compete for Rubisco activity in these plants. On the other hand, C3 plants are affected by the value of S, albeit to a small degree within this range of perturbation.



(a) (b)

Figure 2.32: Response of A_n to (a) constant and (b) exponential base of S equation.

2.3.26 Constants in K_c and K_o Equations

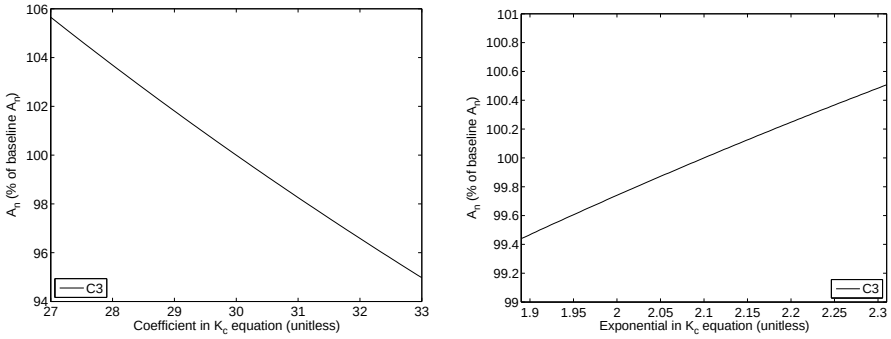
Rubisco limitation of A_g in C3 plants, w_c, incorporates the Michaelis-Menten constant for CO_2 (K_c) and the oxygen inhibition constant (K_o). These constants, both with units in Pa, are quantified by the relations

$$K_c = 30 \cdot 2.1^{Q_{10}} \tag{2.72}$$

and

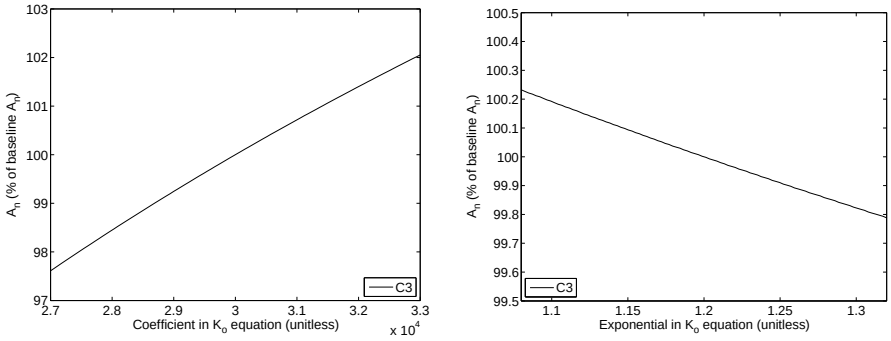
$$K_o = 30000 \cdot 1.2^{Q_{10}}. \tag{2.73}$$

Although several models utilize these equations to calculate K_c and K_o, little information was found to motivate the choice of constant values (30, 2.1, 30000, and 1.2). Therefore Figures 2.33 and 2.34 examine the affect on A_n when these constants vary by 10%. Because, as indicated in Table 2.5, w_c is the most limiting factor defining A_g with the baseline values, perturbations of these constants which affect w_c have an impact on A_n.



(a) (b)

Figure 2.33: Response of A_n to (a) constant and (b) exponential base of K_c equation.



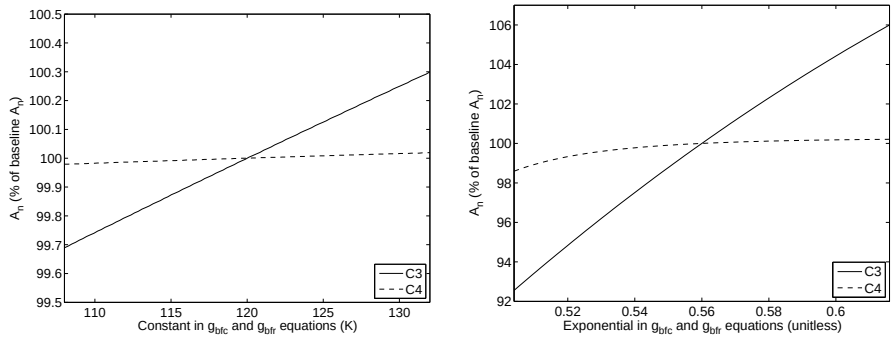
(a) (b)

Figure 2.34: Response of A_n to (a) constant and (b) exponential base of K_o equation.

2.3.27 Constants in g_{bfc} and g_{bfr} Equations

Leaf boundary layer conductance g_b in both C3 and C4 plants is defined as the smaller of conductance under forced conditions (g_{bfc}) and free conditions (g_{bfr}). Little explanation was found for the equations defining these terms; see (2.10) and (2.11). Sensitivity analysis was performed to examine the impact on A_n when the constants 120 K and 0.56 were perturbed in these equations.

Figure 2.35(a) shows that neither species predicts significant change to A_n when the constant 120 K is perturbed by 10%. Alternatively, as depicted in Figure 2.35(b), decreasing the exponential term from 0.56 even slightly causes a noticeable decrease in A_n in both species. As this value is increased from 0.56, the model predicts the value of A_n to approach a steady-state in C4 plants while for C3 plants it appears to increase without bound within this range of perturbation.



(a) (b)

Figure 2.35: Response of A_n to (a) constant and (b) exponent in g_{bfc} and g_{bfr} equations.

2.3.28 Comparison of Parameter Sensitivity

The previous subsections detail the model response in A_n to perturbation of individual parameters. To assess the comparative sensitivity of all these parameters, Figures 2.36 and 2.37 summarize the maximum percent change in A_n for the range examined of each parameter in C3 and C4 plants, respectively. Examining these figures allows us to assess to which parameters A_n is most sensitive for each species.

In C3 plants, the top five most sensitive parameters are: T_s, b, S_2, $f(w_2)$, and h_s. The top five most sensitive parameters for C4 plants are: S_2, $f(w_2)$, T_s, S_4, and β_2.

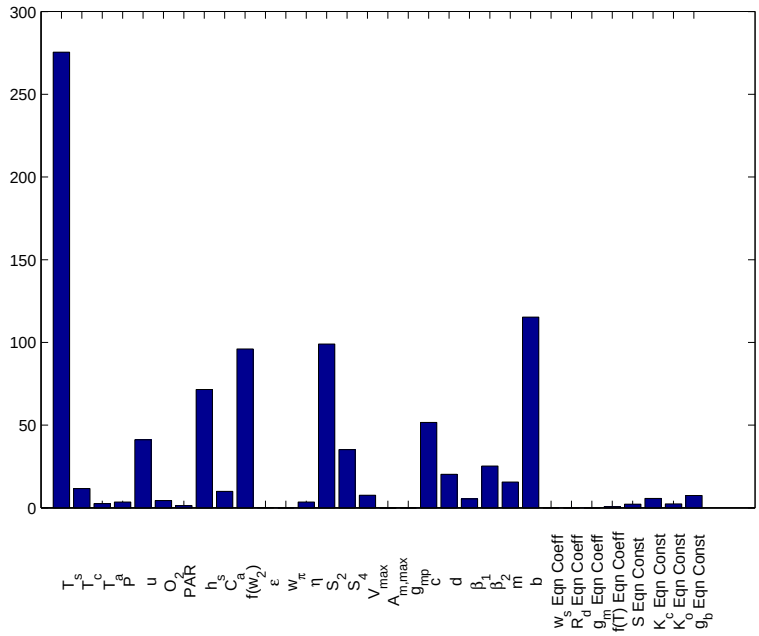


Figure 2.36: Maximum percent change in A_n over parameter ranges defined in previous subsections for C3 plants.

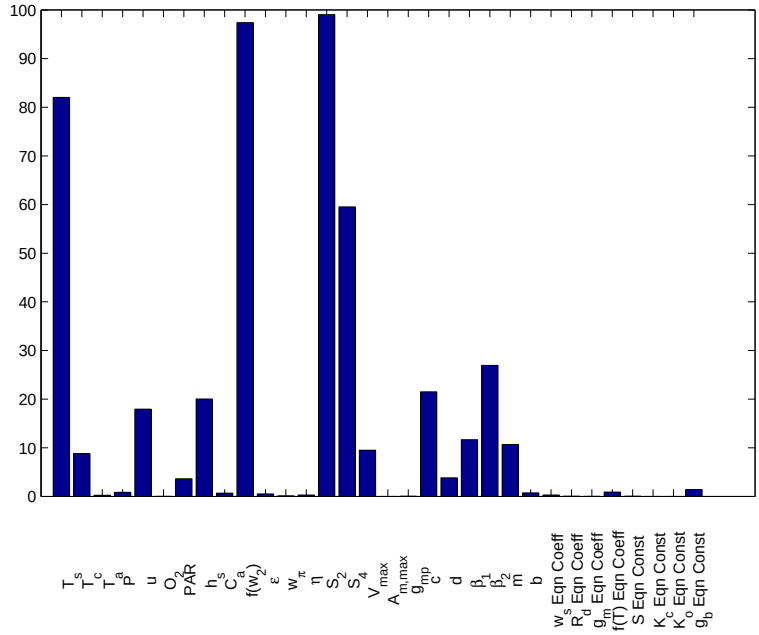


Figure 2.37: Maximum percent change in A_n over parameter ranges defined in previous subsections for C4 plants.

Chapter 3

Data Analysis and Niyogi Model Performance

Every mathematical model is built upon a framework of assumptions. Data is required to validate the formulation of any model. In Section 3.1 we present a new data set collected by Dr. Edwin Fiscus, Plant Physiologist for the Agricultural Research Service (ARS) in the United States Department of Agriculture (USDA). This conductance data along with the associated environmental conditions is used to validate the Niyogi et al. model in Section 3.2. Further, in Chapter 4, this conductance data is used to calibrate a new leaf conductance model based solely on plant age and soil moisture conditions.

3.1 Data

3.1.1 Description of Experimental Procedure

The controlled water stress plot developed by the USDA/ARS Plant Science Research Unit is located at the Inwood Road site in Raleigh, NC. The plot is approximately 23 meters long and 20 meters wide with rows oriented North to South. The overall slope of the field is approximately 0.5 meters from South to North and 0.2 meters from West to East. The field was plowed and furrowed before planting so that any precipitation could flow downhill to a drainage ditch at the northern end of the field. This field is comprised of four sections, three of which are covered by 6 millimeter Polyethylene sheeting to prevent penetration of precipitation to the soil underneath. A network of drip irrigation tubes was placed on the soil surface prior to covering with sheeting allowing for careful control of water conditions beneath the soil surface. The impact of differing water conditions on the stomatal conductance of soybean plants was studied for the 2008 and 2009 growing seasons. Significant events of the growing seasons are

summarized in Table 3.1.

The data under consideration is for two soybean genotypes: Haskell and N01. The Haskell genotype was registered in 1993 after being developed for disease resistance and high productivity [11]. This line is one of the typical lines grown in the southeastern United States. During conditions of increasing soil water deficit, the N01 genotype has been observed to exhibit slow wilting [58]. Therefore, these two contrasting genotypes were selected under the hypothesis that the N01 genotype would present more drought tolerance than the Haskell genotype.

The field is divided into four experimental studies for each soybean genotype: Dry, Medium, Wet, and Open. Figures 3.1 and 3.2 diagram the field layout for 2008 and 2009, respectively. Although tube and experiment locations remained the same for both study years, genotypic placement was randomized each year. In general, the "dry" experiment subjected the soil to drying cycles by withholding water to reach approximately 10% soil moisture. The "medium" experiment was irrigated with 0.75 inches of water per week. The "wet" experiment was irrigated with 3 inches of water per week in the 2008 growing season and 1.5 inches in the 2009 growing season. The open experiment did not have plastic sheeting covering the surface of the soil, so it received precipitation exposure as well as 3 inches of irrigation water treatment per week during 2008 and 1.5 inches in 2009. For all studies, well-watered conditions were applied in the beginning of the growing season to facilitate plant establishment.

Leaf conductance was measured with a model LI-1600M steady state porometer with a photosynthetically active radiation (PAR) sensor (LI-COR, Lincoln, NE). Additionally, the porometer collected simultaneous barometric pressure, relative humidity, PAR, leaf temperature, and cuvette temperature data. All measurements were made on the youngest, fully expanded leaf of the plant. Environmental data including barometric pressure, ambient CO_2 concentrations, relative humidity, PAR, temperature, and wind speed was measured by a weather station in close proximity to the controlled water stress plot.

3.1.2 Soil Moisture Data

Soil moisture measurements were taken by a HH2 Moisture Meter with a PR2 probe (Delta-T Devices, Cambridge, England). Soil moisture data, described as m^3 water/m^3 soil, was collected at three different depths: 20, 30, and 40 cm below the ground level surface. Eight

Table 3.1: Significant growing season events.

	Field planted	Irrigation stopped	Field harvested
2008	May 21	October 6	November 6
2009	May 22	October 21	November 9

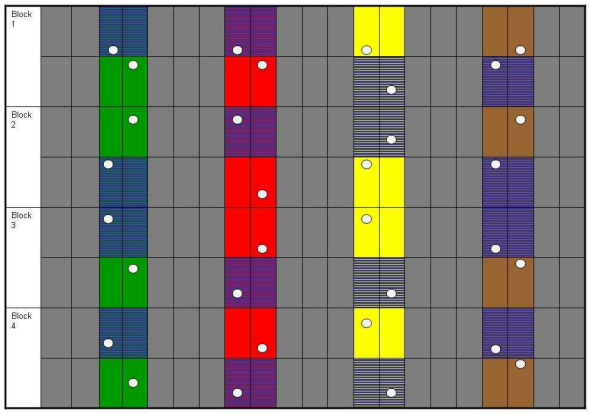


Figure 3.1: Field diagram for 2008 growing season. Color coding of experiments is as follows: Green (open), Red (wet), Yellow (medium), Brown (dry), Gray (border, unsampled). Horizontal hashed areas represent Haskell plant areas and open sections represent N01 areas. The white circles represent approximate locations of tubes where soil moisture measurements were collected.

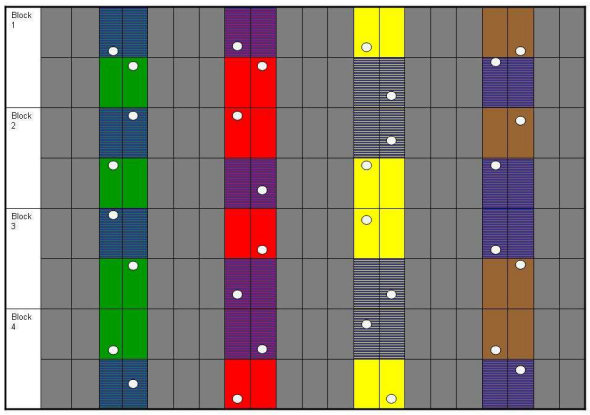


Figure 3.2: Field diagram for 2009 growing season. Color coding of experiments is as follows: Green (open), Red (wet), Yellow (medium), Brown (dry), Gray (border, unsampled). Horizontal hashed areas represent Haskell plant areas and open sections represent N01 areas. The white circles represent approximate locations of tubes where soil moisture measurements were collected.

locations within each of the four experimental study regions were selected for soil moisture measurements, as indicated by the white circle markings on Figures 3.1 and 3.2. Notice that there is an even distribution of moisture tubes between the two soybean genotypes. At each location and depth, three separate measurements were made by rotating the sensor by 120 degrees. Therefore, for each day, experiment, and depth, $8 \times 3 = 24$ measurements were made throughout the experimental plot. Readings were taken on 73 days from May 23 through December 30, 2008 and during 2009, readings were taken 78 times from January 14 through December 1.

The data presented for each date in Figures 3.3 - 3.6 represent the average of these 24 observations at each depth during the 2008 growing season for the "dry", "medium", "wet", and "open" treatment groups, respectively. In each of these figures, there is a line representing the average soil moisture value among the 20, 30, and 40 cm depths. Figure 3.7 summarizes the averages for each treatment group on one graph for comparative purposes. Figures 3.8 - 3.12 are analogous for the 2009 growing season.

In general, the soil moisture content increases with depth from the soil surface. Note that the "dry" and "medium" treatments began well-watered to establish the plants. From late June to early August, drought stress was imposed on these treatment groups. Then, well-watered conditions were maintained throughout the rest of the growing season.

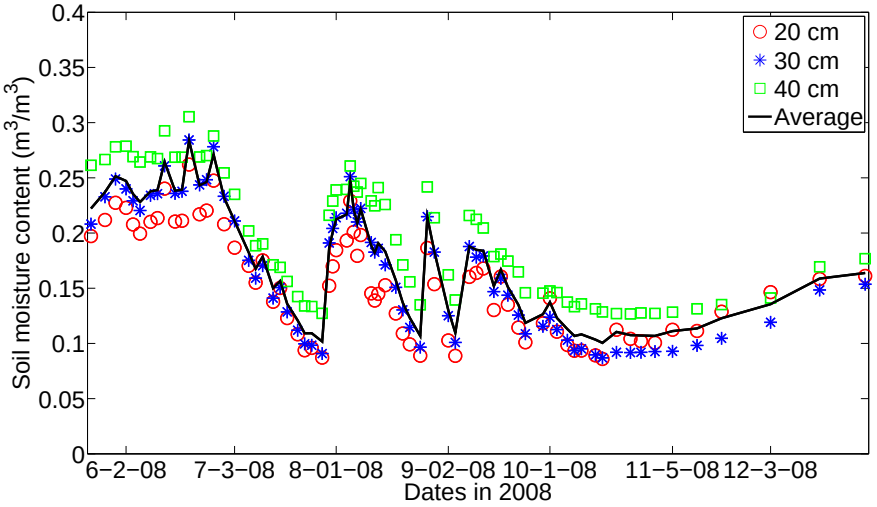


Figure 3.3: Soil moisture content for the "dry" experimental plot in 2008.

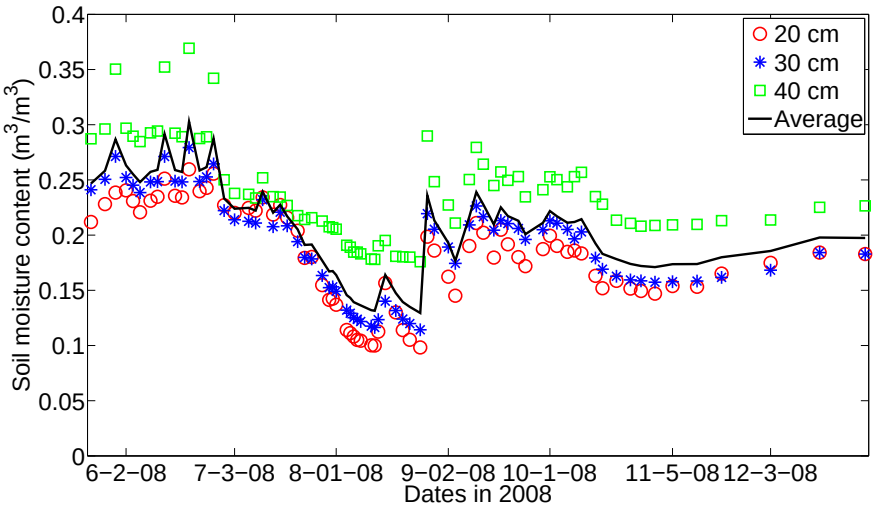


Figure 3.4: Soil moisture content for the "medium" experimental plot in 2008.

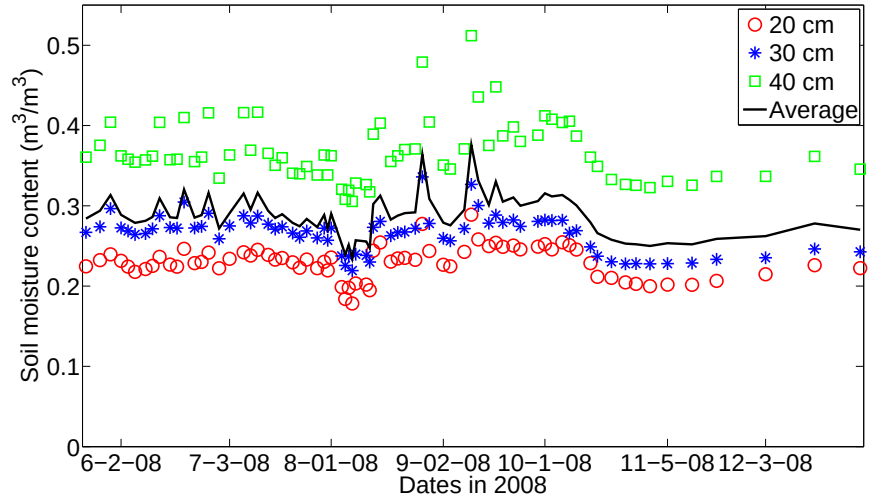


Figure 3.5: Soil moisture content for the "wet" experimental plot in 2008.

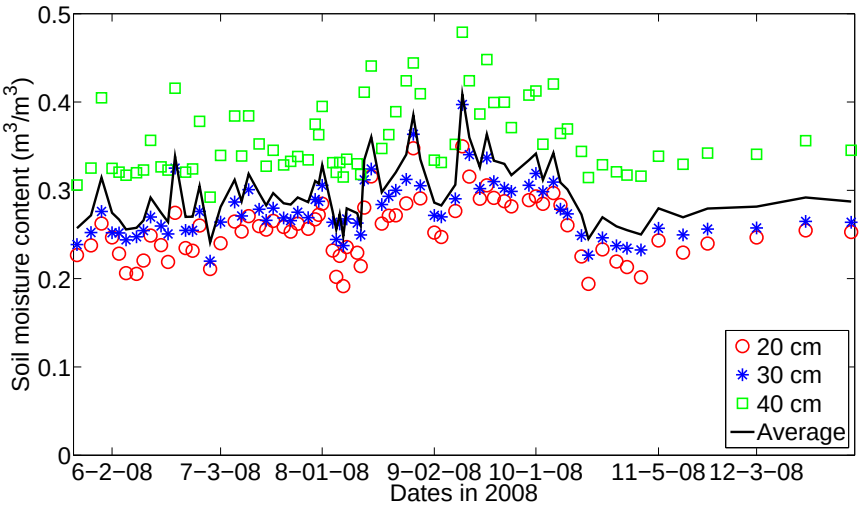


Figure 3.6: Soil moisture content for the "open" experimental plot in 2008.

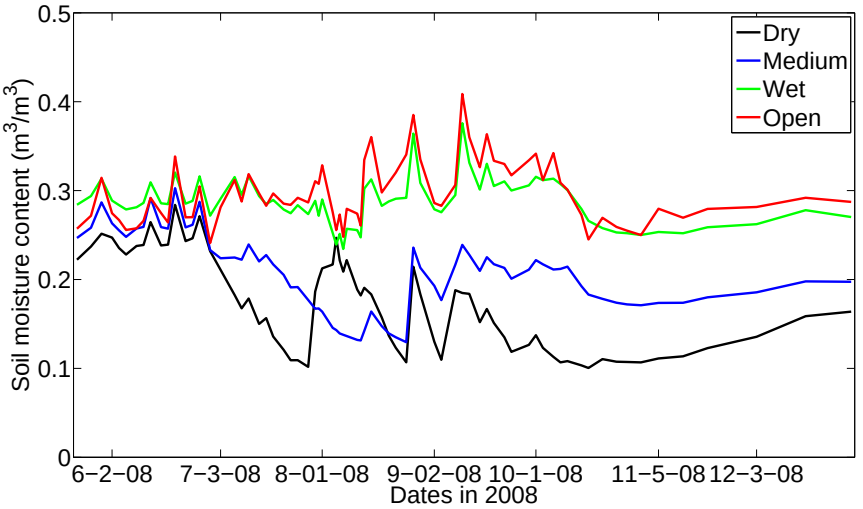


Figure 3.7: Soil moisture content averages for all experimental plots in 2008.

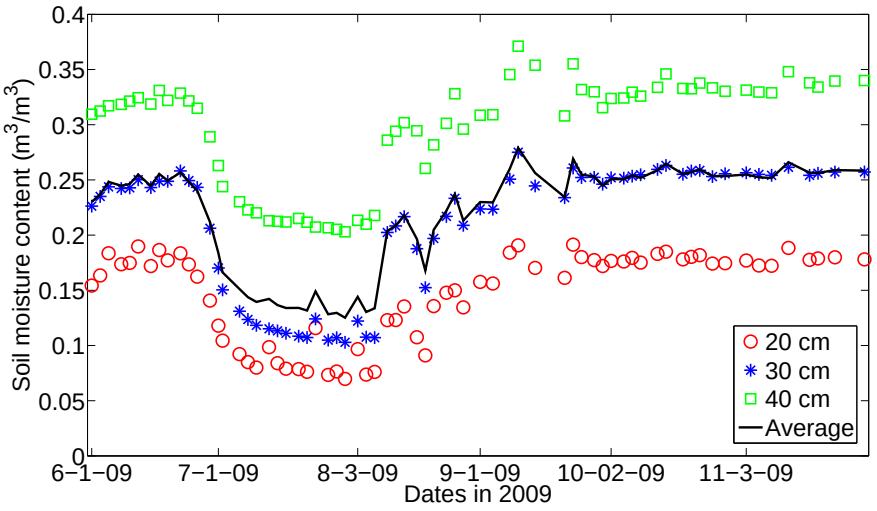


Figure 3.8: Soil moisture content for the "dry" experimental plot in 2009.

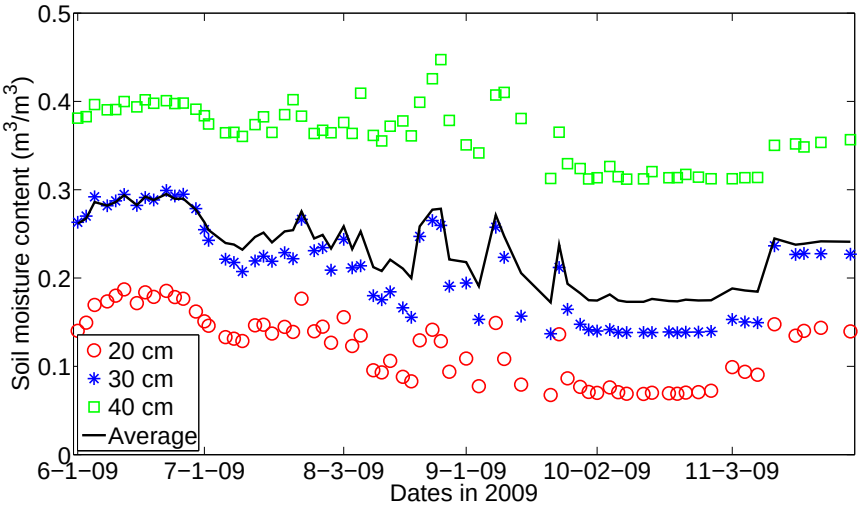


Figure 3.9: Soil moisture content for the "medium" experimental plot in 2009.

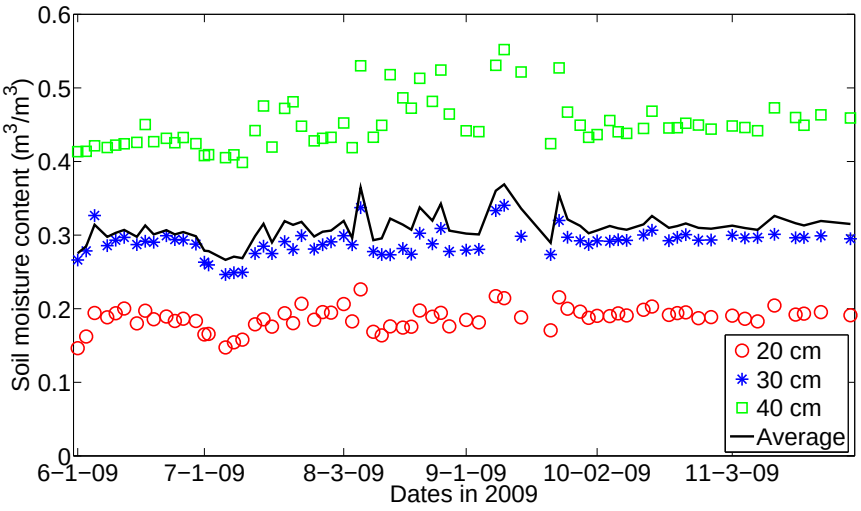


Figure 3.10: Soil moisture content for the "wet" experimental plot in 2009.

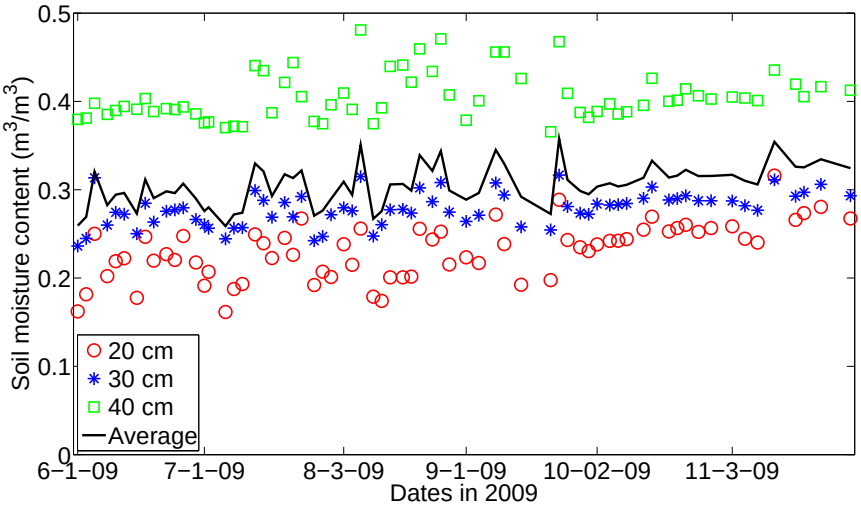


Figure 3.11: Soil moisture content for the "open" experimental plot in 2009.

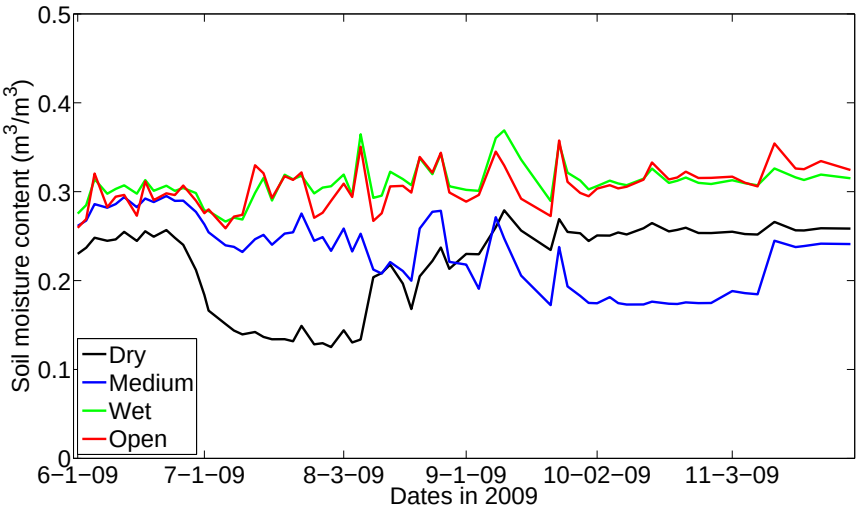


Figure 3.12: Soil moisture content averages for all experimental plots in 2009.

3.1.3 Barometric Pressure

Independent barometric pressure data is available for both the 2008 and 2009 growing seasons. The weather station records barometric pressure readings every few minutes with an Oakton Aneroid Barometer located approximately 1.8 meters above grade and 14 meters from the southeastern corner of the field. Additionally, the porometer requires barometric pressure as an initial manual input. The technician read this value at a barometer separate from the weather station and manually entered the value before taking conductance measurements. Therefore, comparisons between weather station values can be made with the porometer data readings to ensure reliability of the weather station measurements. Because the weather station provides more precise and incremental measurements, this data will be used as model inputs. However, comparing to the porometer technician's independent barometric pressure ensures reliability of the weather station measurements.

The porometer recorded the time of day when conductance measurements were calculated. The mean barometric pressure reading within this time period, usually between the hours of 11:00 am and 2:00 pm, from the weather station is used. Figure 3.13 summarizes the barometric pressure data from both the 2008 and 2009 growing seasons. The barometric pressure readings as recorded by the weather station are remarkably consistent with the independent barometer readings made by the porometer technician. Further note that the barometric pressure does not vary much over the course of either growing season.

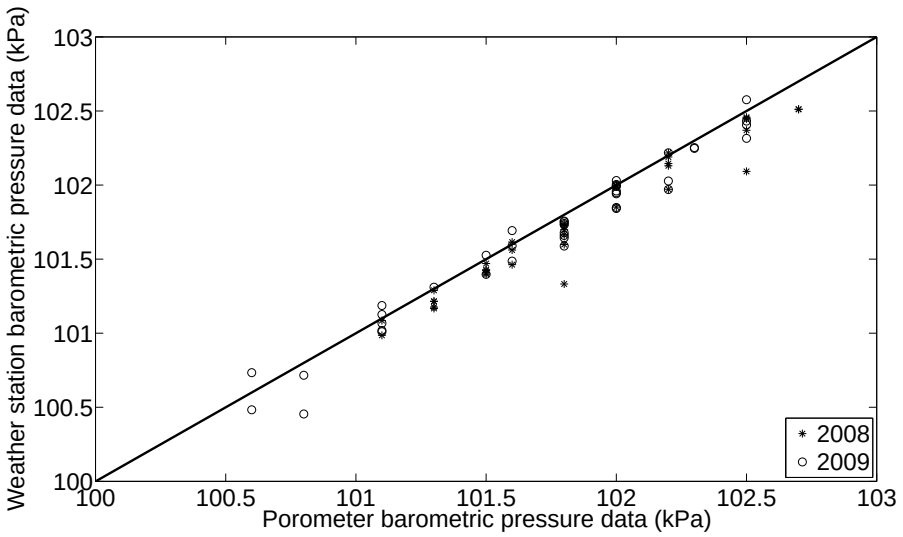


Figure 3.13: Barometric pressure data from 2008 and 2009 growing seasons.

3.1.4 Ambient CO_2 Levels

Ambient CO_2 levels were quantified by a model LI-6252 CO_2 analyzer (LI-COR, Lincoln, NE) with an intake approximately 14 meters from the southeast corner of the controlled water stress plot at a height of approximately 1 meter. Figures 3.14 and 3.15 illustrate the minimum, mean, and maximum ambient CO_2 level during the hours of 11:00 am through 2:00 pm, which is the general time period of porometer data collection, on each date. The minimums and maximums are shown by horizontal markings, while the mean is illustrated with the line connecting dates of the study. All ambient CO_2 concentrations for the experimental time period are between 340 and 380 ppm. However, it is interesting to note that ambient CO_2 levels can vary greatly not only amongst days but hours as well. For example, on July 1, 2009 the concentration varied from approximately 347 to 374 ppm just within a 3 hour time period.

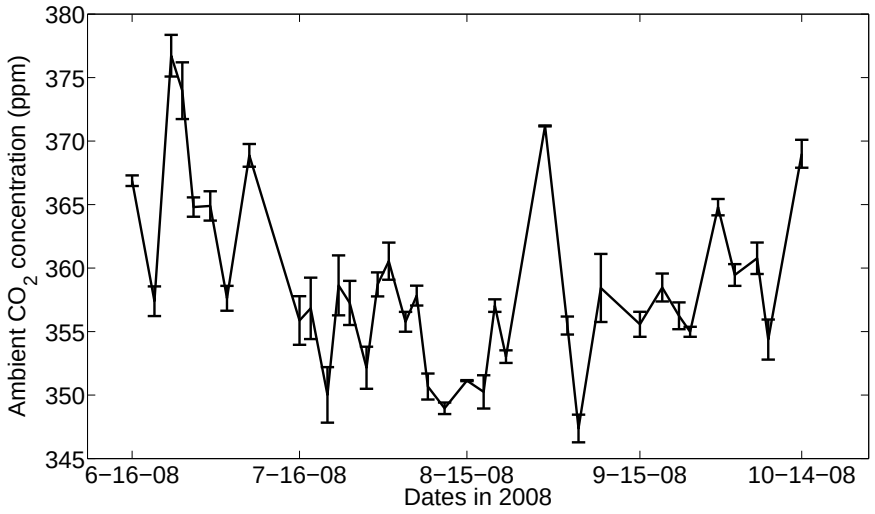


Figure 3.14: Minimum, mean, and maximum ambient carbon dioxide levels during the 2008 growing season.

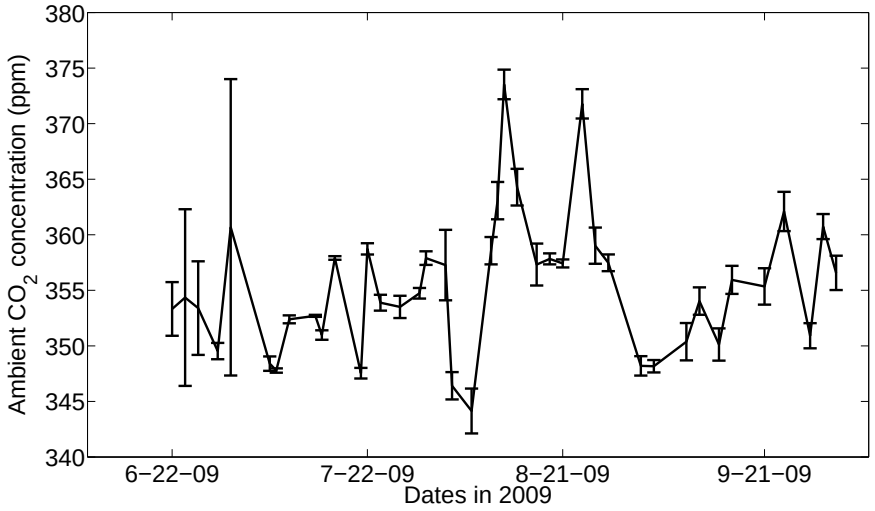


Figure 3.15: Minimum, mean, and maximum ambient carbon dioxide levels during the 2009 growing season.

3.1.5 Relative Humidity

Relative humidity data was measured both at the weather station and with the porometer at the precise time and location of leaf conductance measurements. A model 41382VC aspirated relative humidity/temperature probe (R.M. Young Company, Traverse City, MI) located approximately 10.7 meters from the southeast corner of the controlled water stress plot at a height of approximately 4.4 meters was used for relative humidity measurements from the weather station. The relative humidity data, as measured with the porometer, is used in Section 3.2 for model testing as it is more closely related to the leaf conductance measurements. However, to ensure reliability of the porometer relative humidity data values, it is compared with weather station data in Figure 3.16.

For the most part, relative humidity values, as recorded with the porometer, are consistent with the measurements taken at the weather station. On any given day the data points tend to vary by less than 20% which is not atypical given the affect that haze, spotty cloud cover, and wind can impose on relative humidity. However, note that the weather station has no reading less than 0.2, while the porometer does have a grouping of multiple dates in 2008 where the values were less than 0.2. All of these dates correspond to the first week of experimentation, and these values are considered to be unreliable given the unrealistically low value for relative humidity in this geographical region coupled with the probability that technicians were fairly inexperienced with the equipment during this time period.

3.1.6 Photosynthetically Active Radiation (PAR)

PAR was measured both at the weather station and with the porometer at the precise time and location of leaf conductance measurements. The PAR data measured with the porometer is plotted in Figures 3.17 and 3.18 and is used in Section 3.2 for model testing. Given the geometrical setup of the porometer instrument, PAR reading were made for both the upper and lower surfaces of the leaf. However, here we are only utilizing the incident PAR as received by the upper leaf surface.

Notice in Figures 3.17 and 3.18 that the values of PAR can significantly range over the course of even the same day. This variation can be due to a number of factors including changing cloud cover, haze, dust, and the angle of the sun.

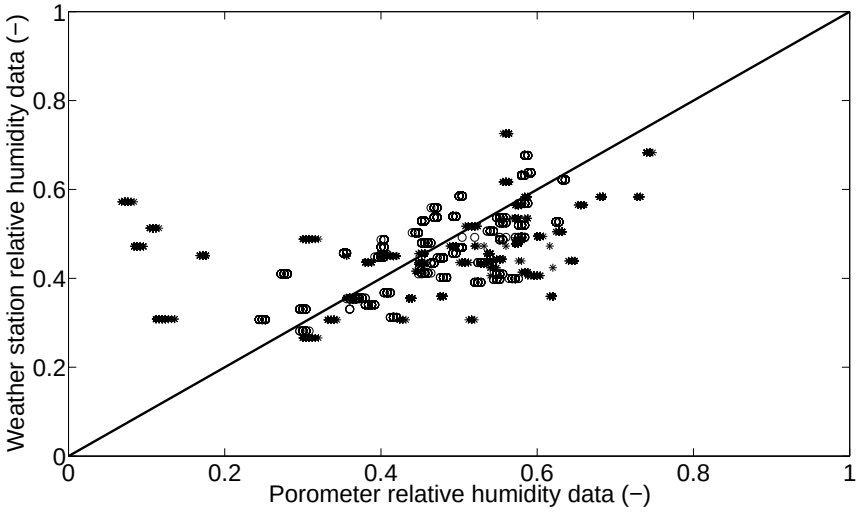


Figure 3.16: Relative humidity data. Stars are relative humidity values from the 2008 growing season and circles are data from 2009.

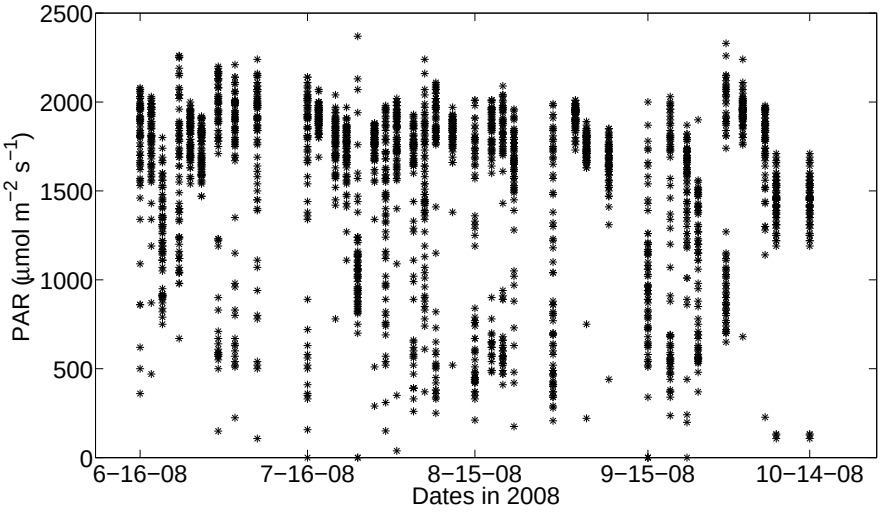


Figure 3.17: PAR data from 2008 growing season.

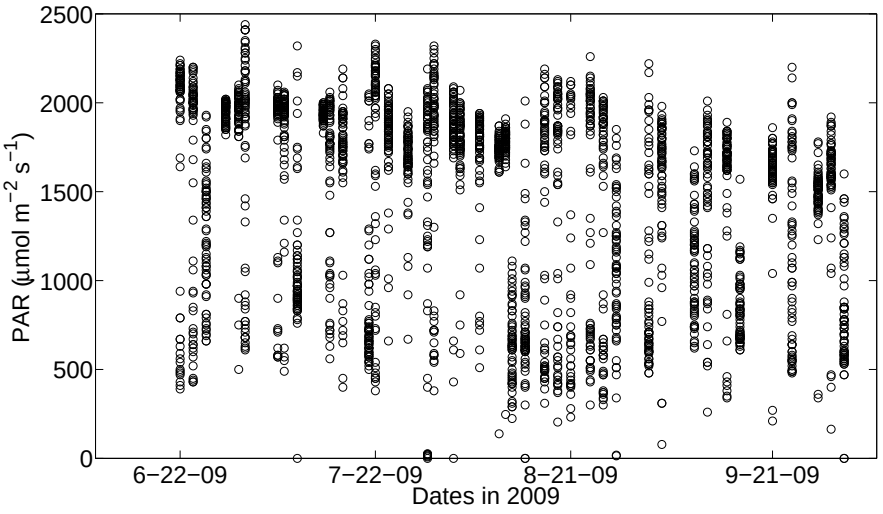


Figure 3.18: PAR data from 2009 growing season.

3.1.7 Temperature

The weather station collects temperature values with the same relative humidity/temperature probe used to collect relative humidity data at a height approximately 4.4 meters higher than the field surface. The porometer collects the leaf and cuvette temperatures during conductance calculation. The cuvette is the part of the porometer where conductance measures are made, and the cuvette temperature is considered as ambient temperature.

Figure 3.19 illustrates the relationship of cuvette and leaf temperature values which are measured simultaneously by the porometer. These values mostly differ by less than 5 degrees Kelvin. Notice that the data illustrates that there is no rule dictating which of these values is largest in magnitude.

Using the time period of conductance data collection (typically 11:00 am - 2:00 pm), we summarize the mean temperature value as recorded by the weather station in Figure 3.20. There is little difference between the temperature trends for these two growing seasons.

The model contains separate variables for three temperature values: canopy, ambient, and leaf surface temperature. Oftentimes the canopy temperature T_c is set equal to the ambient temperature T_a for model simulation. However, model testing in Section 3.2 uses the weather station temperature readings as T_c, the porometer cuvette temperatures as T_a, and the porometer leaf temperature readings as T_s.

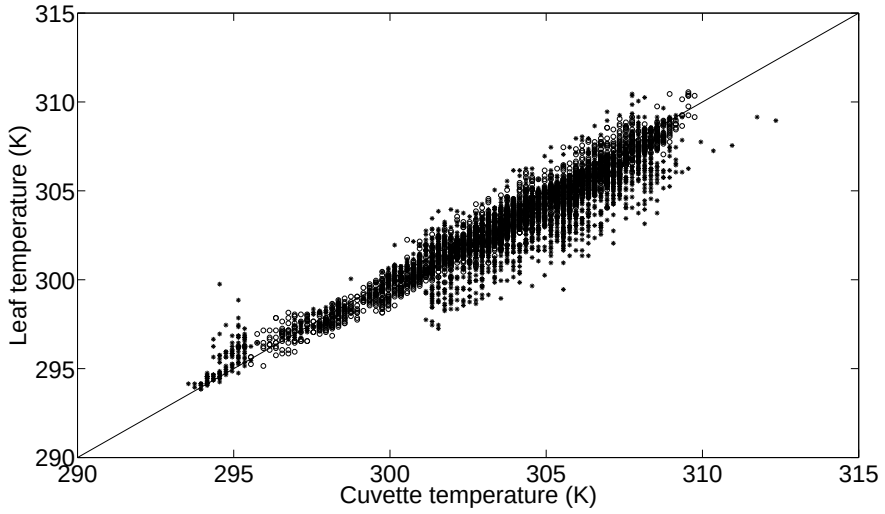


Figure 3.19: Cuvette temperature versus leaf temperature. Stars are values from the 2008 growing season, and circles are data from 2009.

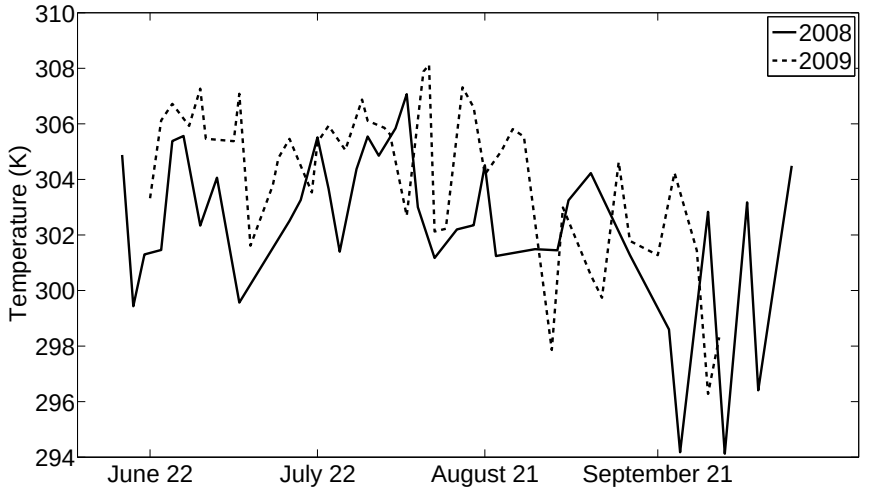


Figure 3.20: Weather station temperature readings during the 2008 and 2009 growing seasons.

3.1.8 Wind

Wind speed is a necessary input into the model, and this data was gathered from the weather station near the field. Wind direction and speed is measured with a model 12005 Gill MicroVane and 3-cup anemometer system (R.M. Young Company, Traverse City, MI) located at a height approximately 5.6 meters above grade and 10.7 meters from the southeast corner of the controlled water stress plot. The mean wind speed during the time period of conductance data collection was calculated for each day of the study. These mean wind speed values are used for model testing in Section 3.2 and are plotted in Figures 3.21 and 3.22 for the 2008 and 2009 growing seasons, respectively.

Note the existence of several issues with interpreting this data and it's relevance to the model:

- The weather station is approximately 5.6 meters in the air which is significantly higher than the plant height during any part of the growing season. Are winds at this height even relevant for the purpose of wind speed input to the model?

- Using the time period of conductance data collection (typically 11:00 am - 2:00 pm), we have found the mean wind speed value recorded during this time period. Wind speeds are recorded as instantaneous measurements every few minutes. However, we do not have an individual wind speed measurement for each leaf conductance data point.

- Does the clamp on the porometer itself make this part of the model useless anyway? Wind speed controls boundary layer conductance, but if the clamp is blocking the wind's interaction with the leaf surface, is wind even a relevant input given this type of conductance data?

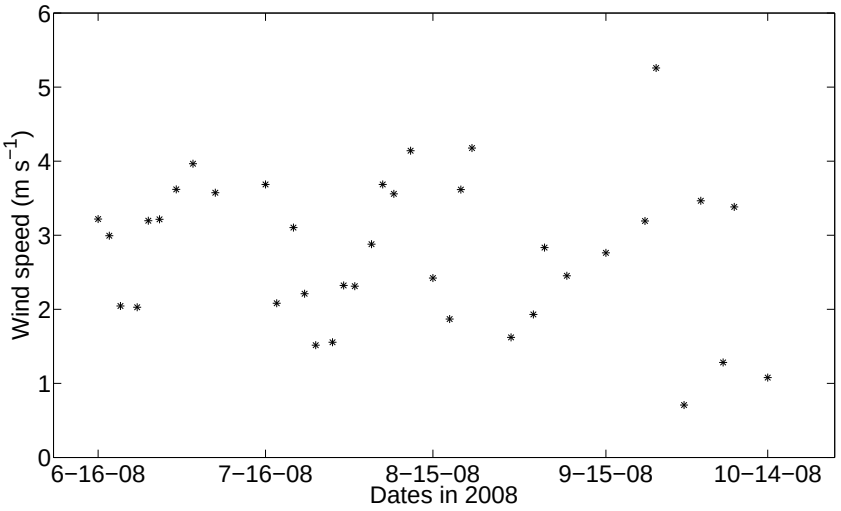


Figure 3.21: Wind speed data from 2008 growing season.

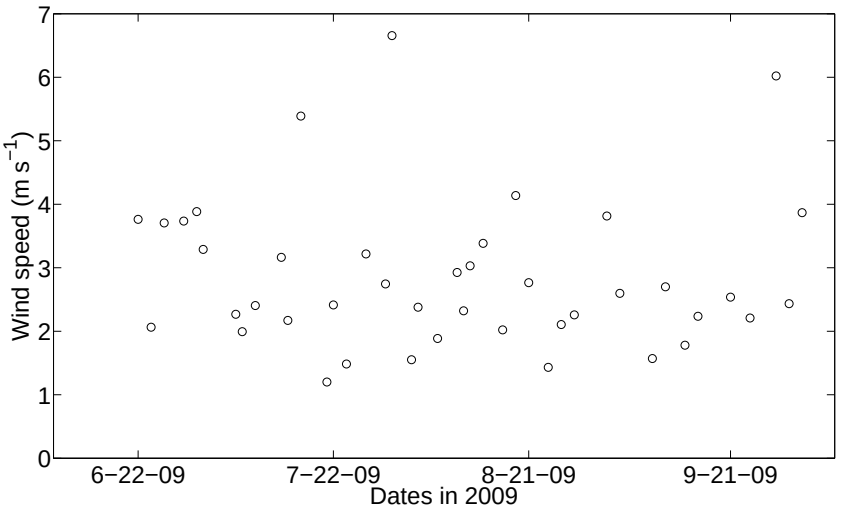


Figure 3.22: Wind speed data from 2009 growing season.

3.1.9 Leaf Conductance

Figures 3.23 through 3.26 summarize the total leaf conductance from the 2008 growing season for the "dry", "medium", "wet", and "open" experiments, respectively. On each figure, red markings illustrate data from the Haskell genotype and green markings data from the N01 genotype. Figure 3.27 summarizes the means for all experiments with the Haskell genotype in 2008 and Figure 3.28 summarizes the mean leaf conductance for all experiments with the N01 genotype in 2008. Both genotypes show decreased leaf conductance under conditions of soil moisture stress. Figures 3.29 through 3.34 are analogous plots for the 2009 growing season.

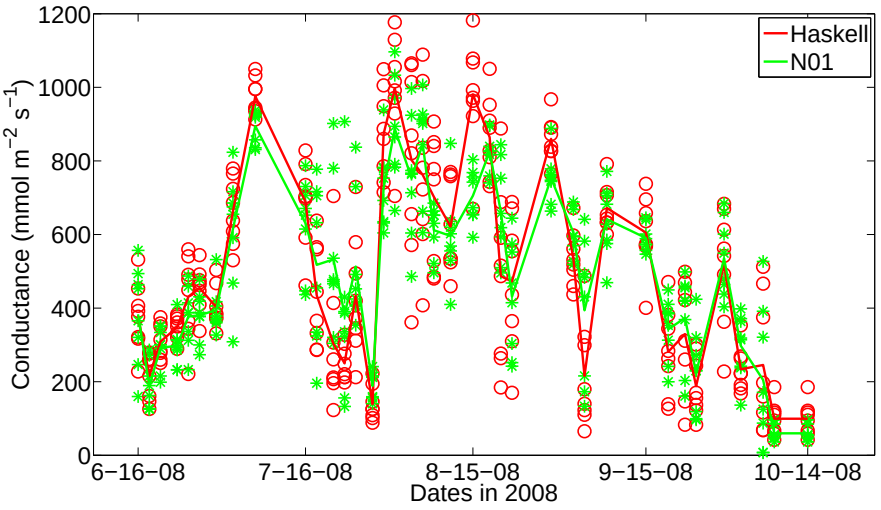


Figure 3.23: Conductance data from "dry" experimental plot in 2008. Red circles and lines indicate Haskell genotype, Green stars and lines indicate N01 genotype. Stars/circles are individual data points while lines represent means.

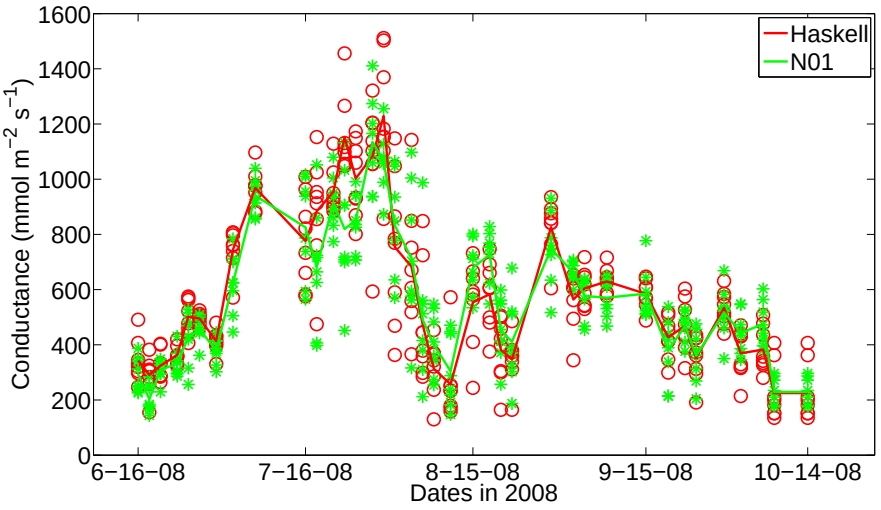


Figure 3.24: Conductance data from "medium" experimental plot in 2008. Red circles and lines indicate Haskell genotype, Green stars and lines indicate N01 genotype. Stars/circles are individual data points while lines represent means.

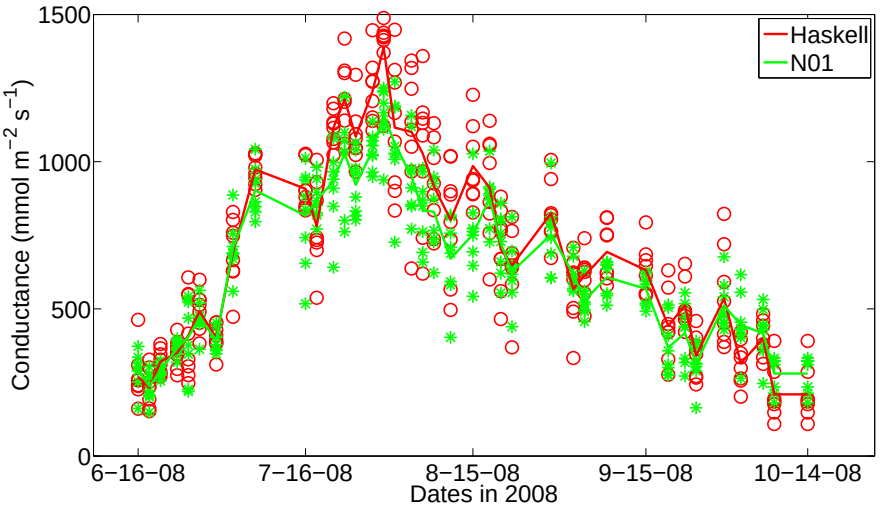


Figure 3.25: Conductance data from "wet" experimental plot in 2008. Red circles and lines indicate Haskell genotype, Green stars and lines indicate N01 genotype. Stars/circles are individual data points while lines represent means.

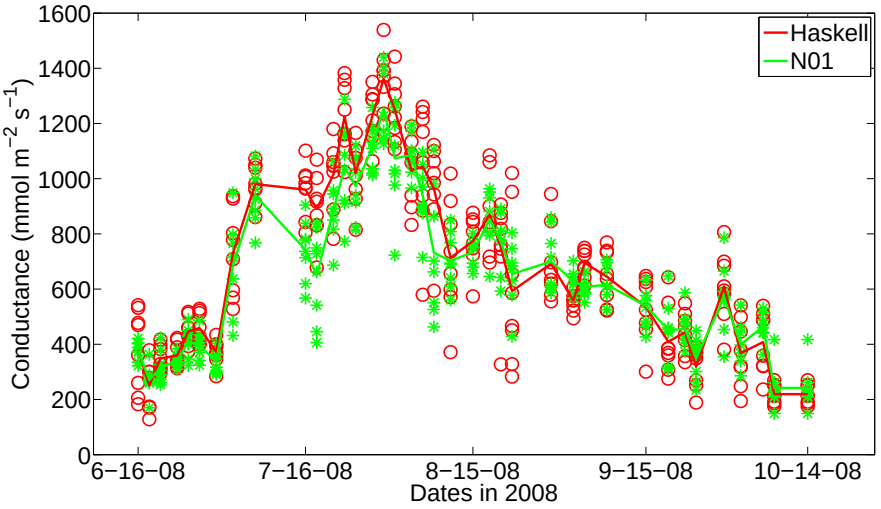


Figure 3.26: Conductance data from "open" experimental plot in 2008. Red circles and lines indicate Haskell genotype, Green stars and lines indicate N01 genotype. Stars/circles are individual data points while lines represent means.

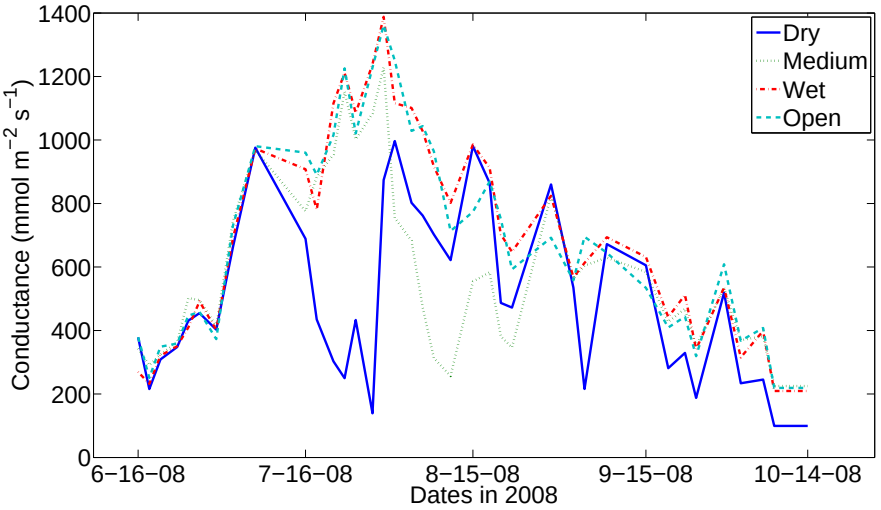


Figure 3.27: Conductance data averages for all experimental plots with Haskell genotype in 2008.

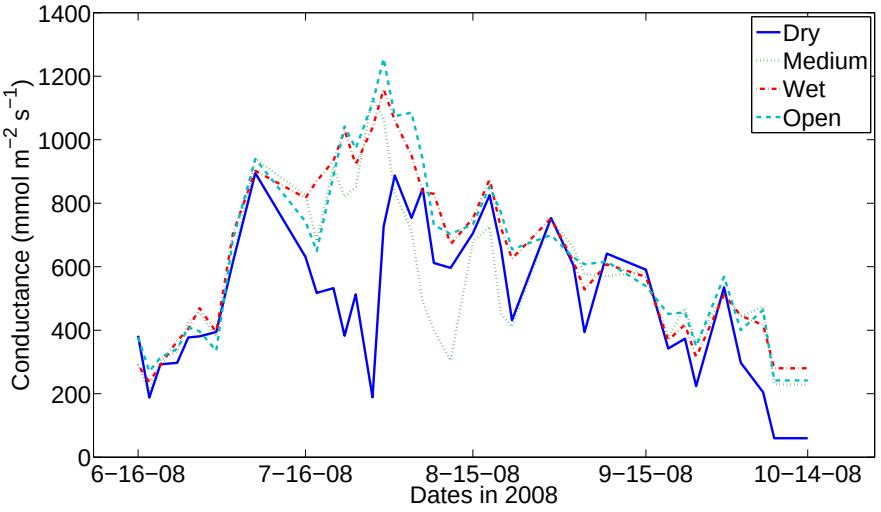


Figure 3.28: Conductance data averages for all experimental plots with N01 genotype in 2008.

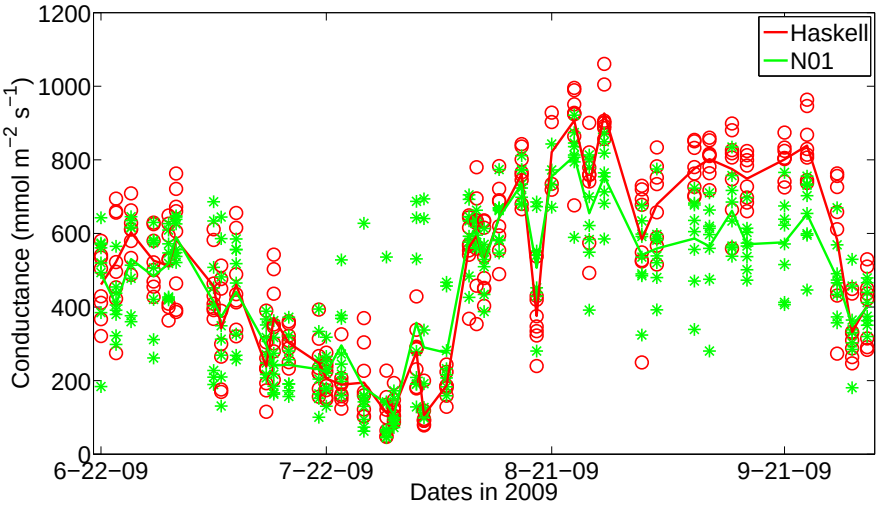


Figure 3.29: Conductance data from "dry" experimental plot in 2009. Red circles and lines indicate Haskell genotype, Green stars and lines indicate N01 genotype. Stars/circles are individual data points while lines represent means.

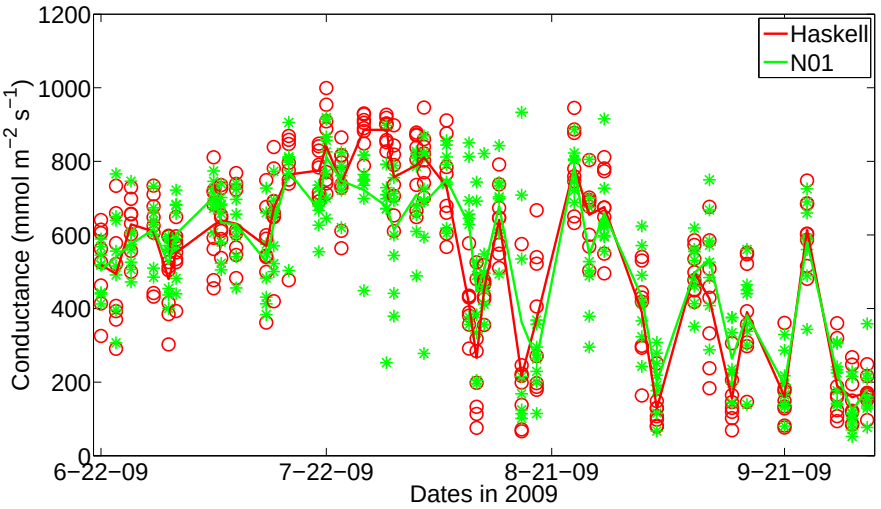


Figure 3.30: Conductance data from "medium" experimental plot in 2009. Red circles and lines indicate Haskell genotype, Green stars and lines indicate N01 genotype. Stars/circles are individual data points while lines represent means.

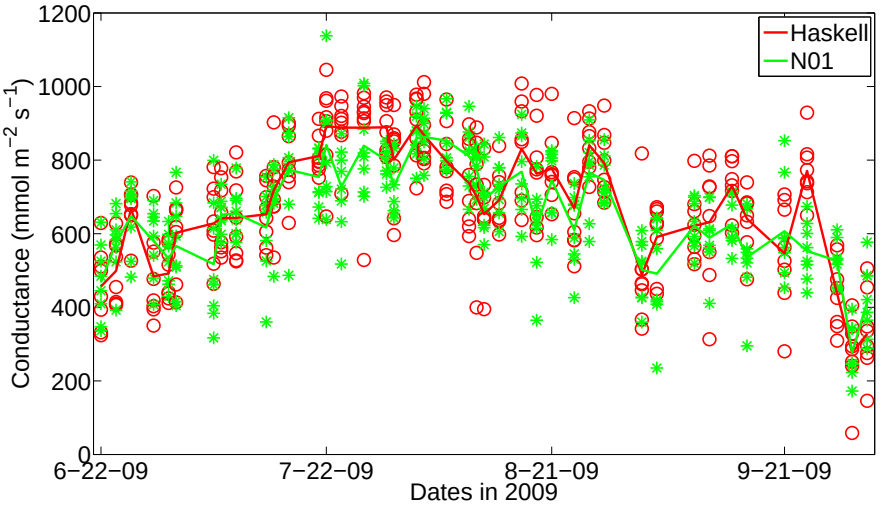


Figure 3.31: Conductance data from "wet" experimental plot in 2009. Red circles and lines indicate Haskell genotype, Green stars and lines indicate N01 genotype. Stars/circles are individual data points while lines represent means.

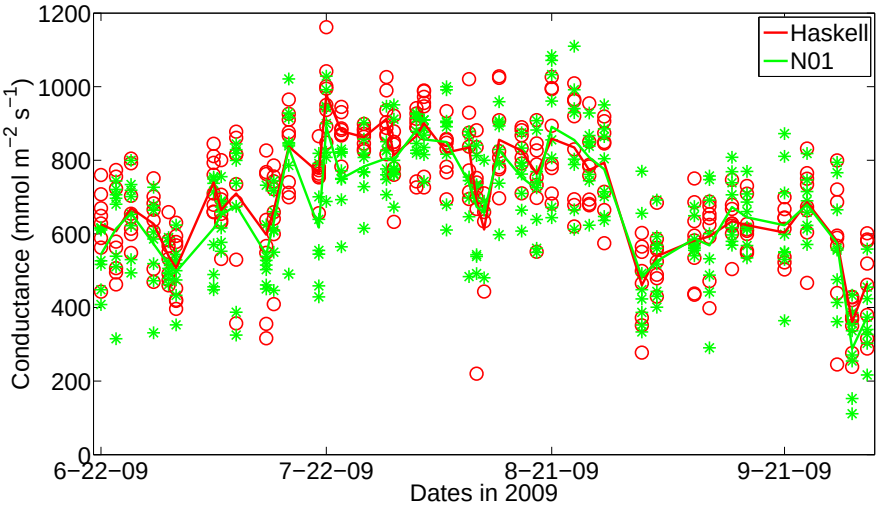


Figure 3.32: Conductance data from "open" experimental plot in 2009. Red circles and lines indicate Haskell genotype, Green stars and lines indicate N01 genotype. Stars/circles are individual data points while lines represent means.

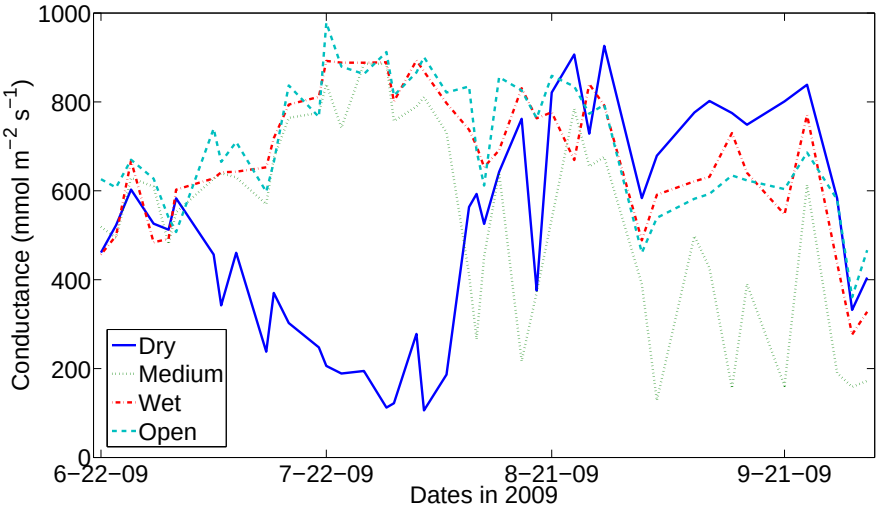


Figure 3.33: Conductance data averages for all experimental plots with Haskell genotype in 2009.

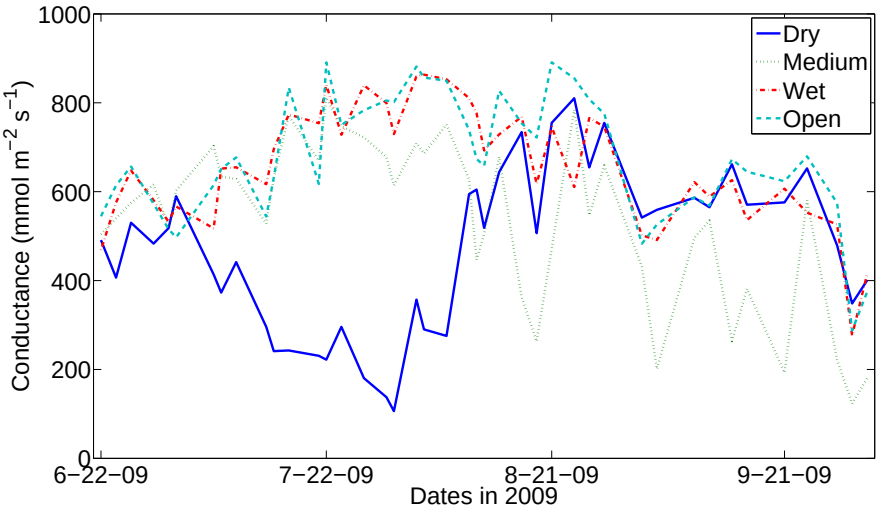


Figure 3.34: Conductance data averages for all experimental plots with N01 genotype in 2009.

3.2 Niyogi Model Performance

During the 2008 growing season, there were 38 days for which leaf conductance data was measured. On each day, 8 plants of each genotype were measured for each of the 4 studies, yielding a total of 32 conductance measurements. This would imply we could compare $38 \times 32 = 1216$ model predictions to data for each genotype. However, for some days we are lacking either CO_2, soil moisture, or environmental data which are required inputs for model simulation. In particular, as summarized in Table 3.2, 96 comparisons could not be made for each genotype due to lack of input data. For the 2009 growing season, there were 42 days for which leaf conductance data was gathered. This would imply a total of 1344 data points for 2009. However, due to rain events which resulted in incomplete data collection on several days, a total of 1325 data points were gathered for the Haskell genotype and 1327 for the N01. Moreover, 384 comparisons could not be made for each genotype because of a lack of input data. Additionally, several of the model simulations did not have a solution given the input data parameter values. These appear to happen in circumstances where the PAR reading was zero, soil conditions were extremely dry, or relative humidity was very low.

For every date soil moisture data was available in the 2008 growing season, the soil moisture content at 20, 30, and 40 cm was averaged. Therefore, for every date soil moisture data was collected, we had one soil moisture value for each study (dry, medium, wet, and open). These values were used as w_2 values in (2.37) for model simulation. The wilting value w_{wilt} was set to

Table 3.2: Conductance data point summary.

	2008		2009	
	Haskell	N01	Haskell	N01
Data used	1116	1116	935	937
No CO_2 data	32	32	32	32
No soil moisture data	32	32	352	352
No weather data	32	32	0	0
No model root	4	4	6	6
Total data employed	1216	1216	1325	1327

be the minimum of all averaged values. The field capacity was set to be the mean of the averaged values in the beginning of the growing season when all experiments were well-watered, but not over-watered as appears to be the case for later dates. In particular, the field capacity w_{fc} was set to be the mean of the average of the dates May 23 - June 27, 2008 and June 1 - June 26, 2009, respectively, where there were different field capacity values for each experimental study. Table 3.3 summarizes the values used for each respective grouping of simulations.

As described in Section 3.1.4, we have ambient carbon dioxide concentration values from a data recorder in close proximity to the soybean field. This data is recorded in ppm (or equivalently µmol CO_2/mol air); however the model requires input values in partial pressure with units of Pascal. To convert the readings from ppm to Pascals, Dalton's law

$$C_a(\text{Pa}) = C_a\left(\frac{\text{µmol } CO_2}{\text{mol air}}\right) \cdot P(\text{Pa}) \cdot \frac{1 \text{ mol } CO_2}{10^6 \text{µmol } CO_2} \tag{3.1}$$

of partial pressure is used where P is the barometric pressure. The mean ambient CO_2 value for each date is used for simulation.

The barometric pressure, canopy temperature, and wind values are gathered from the weather station data as described in Sections 3.1.3, 3.1.7, and 3.1.8, respectively. There is one of each of these values for each date. Relative humidity, leaf surface temperature, ambient

Table 3.3: Soil parameter values derived from data.

	2008		2009	
	w_{wilt}	w_{fc}	w_{wilt}	w_{fc}
Dry	0.1005	0.2456	0.1252	0.2463
Intermediate	0.1005	0.2662	0.1252	0.2845
Wet	0.1005	0.2933	0.1252	0.3005
Open	0.1005	0.2785	0.1252	0.2915

temperature, and PAR values are determined uniquely with each leaf conductance calculation by the porometer. These values are used for model simulation and are compared with the respective conductance data.

The model analyzed in Chapter 2 has 11 parameter values which may be tuned so simulation is appropriate for a given species. A literature review was conducted to identify parameter values specific to soybean. For seven of these 11 parameters, soybean values were found. General values appropriate to C3 plants were used for the remaining four parameters. These values, as well as their sources, are summarized in Table 3.4.

The conductance data, as collected by the porometer and detailed in Section 3.1.9, does not correct for leaf boundary layer conductance g_b. The data represents total leaf conductance (g_ℓ) which is a combination of the stomatal (g_s) and boundary layer (g_b) conductances. Recall that conductances are not additive, but resistances are. Further, conductance is merely the reciprocal of resistance. Hence we can calculate total leaf resistance

$$r_\ell = r_s + r_b = \frac{1}{g_s} + \frac{1}{g_b} = \frac{g_b + g_s}{g_b g_s} \tag{3.2}$$

thus implying total leaf conductance

$$g_\ell = \frac{1}{r_\ell} = \frac{g_b g_s}{g_b + g_s}. \tag{3.3}$$

Figures 3.35 through 3.38 illustrate the results of total leaf conductance predictions g_ℓ, using (3.3) with the Niyogi et al. [48] model predictions of g_b (2.33) and g_s (2.41), versus data for both genotypes of soybean and each growing season under study. As detailed in Table 3.2, Figures 3.35 and 3.36 illustrate 1116 comparisons and Figures 3.37 and 3.38 each have 938.

Table 3.4: Species-specific model parameter values used for simulation.

Parameter (units)	Value	Source
m (-)	10.6	Leakey et al. (2006)
b (mol m^{-2} s^{-1})	0.008	Leakey et al. (2006)
$A_{m,max}$ (mol m^{-2} s^{-1})	3.311×10^{-5}	Yu et al. (2004)
g_{mp} (m/s)	0.0035	Bunce (1985)
S_2 (K)	310	Sellers et al. (1996b) – for C3 species
S_4 (K)	280	Sellers et al. (1996b) – for C3 species
c (mol m^{-2} s^{-1})	4.322×10^{-3}	Nikolov et al. (1995) – for broad leaf C3 species
d (m)	0.1	Approximated
V_{max} (mol m^{-2} s^{-1})	9.115×10^{-5}	Vu et al. (2006), Parisi et al. (1998)
ϵ (mol mol^{-1})	0.0524	Boote et al. (1994), Ehlringer et al. (1977)
w_π (-)	0.18	Kasperbauer (1987)

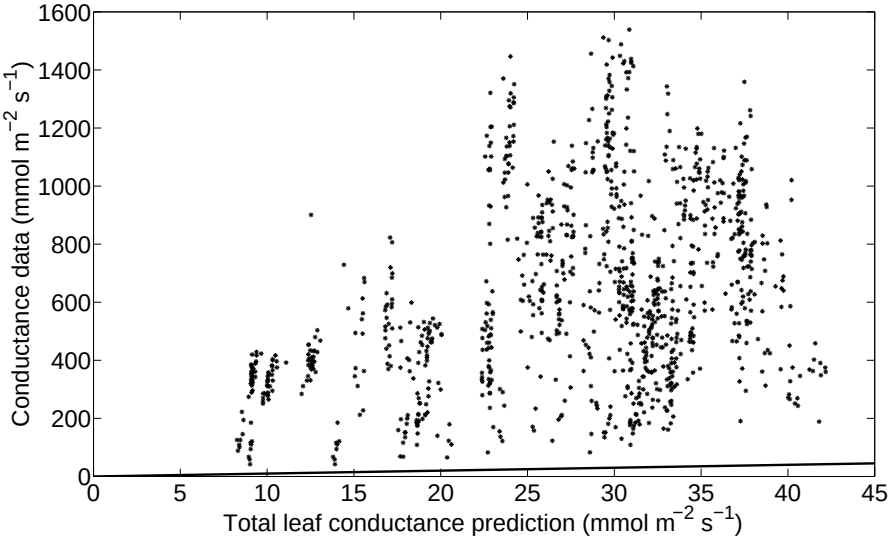


Figure 3.35: Conductance predictions g_ℓ versus data for all experimental plots with Haskell genotype in 2008. The 45° line is superimposed to indicate the scale of underprediction.

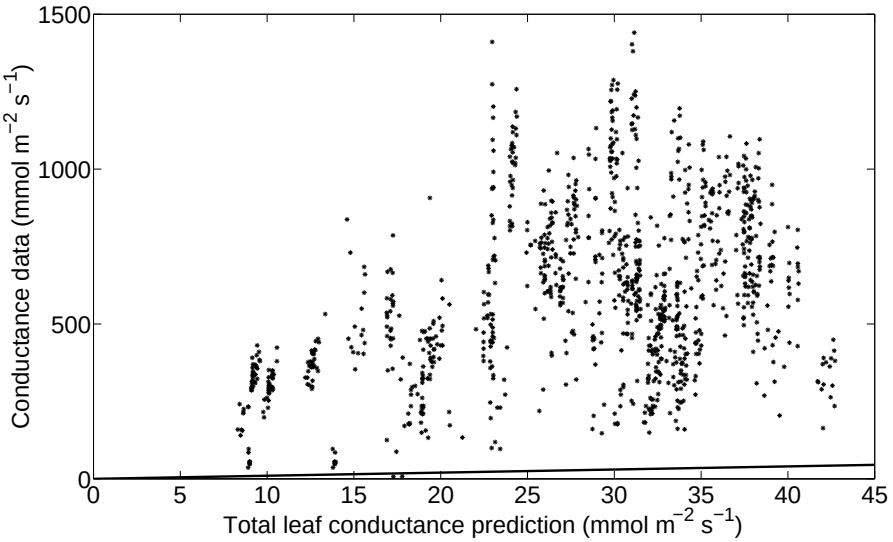


Figure 3.36: Conductance predictions g_ℓ versus data for all experimental plots with N01 genotype in 2008. The 45° line is superimposed to indicate the scale of underprediction.

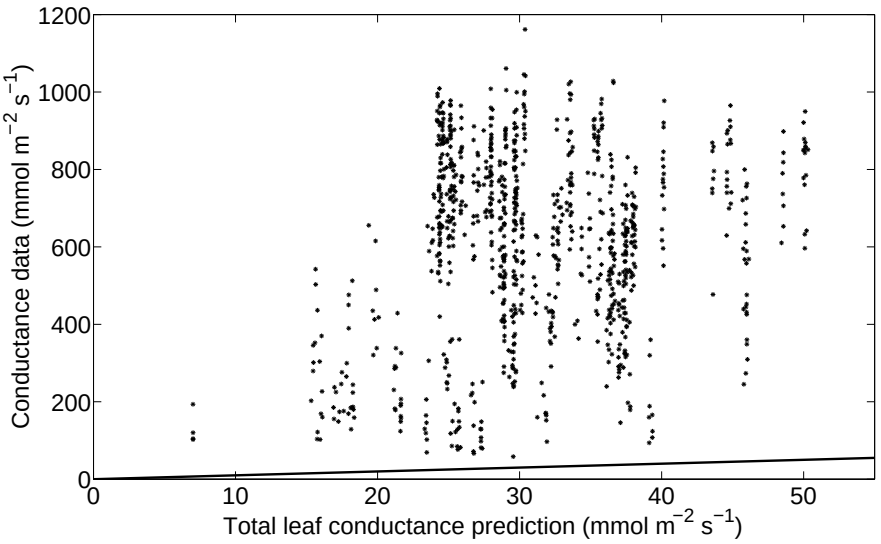


Figure 3.37: Conductance predictions g_ℓ versus data for all experimental plots with Haskell genotype in 2009. The 45° line is superimposed to indicate the scale of underprediction.

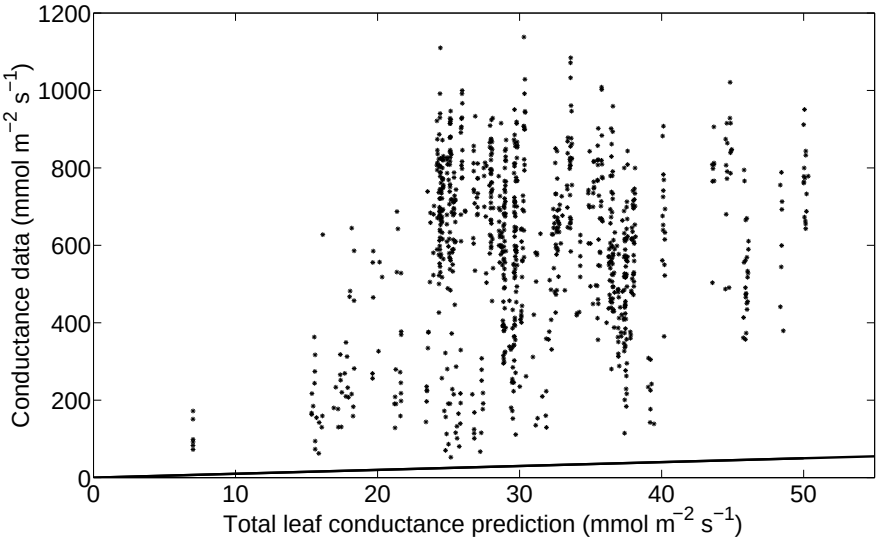


Figure 3.38: Conductance predictions g_ℓ versus data for all experimental plots with N01 genotype in 2009. The 45° line is superimposed to indicate the scale of underprediction.

As is illustrated in Figures 3.35 through 3.38, the model underpredicts the total leaf conductance as indicated by the location of all points above the 45° line imposed on the figures. Table 3.5 depicts the sum of squared error (SSE) values for each experiment, broken down by genotype and growing season. Define the input parameter vector p_k to contain the necessary environmental inputs, associated with the kth data point, for soil moisture w_2, wilting value w_{wilt}, field capacity w_{fc}, ambient CO_2 concentration C_a, barometric pressure P, wind speed u, relative humidity h_s, PAR, canopy temperature T_c, ambient temperature T_a, and leaf surface temperature T_s, that is,

$$p_k = [w_{2,k} \;\; w_{wilt,k} \;\; w_{fc,k} \;\; C_{a,k} \;\; P_k \;\; u_k \;\; h_{s,k} \;\; PAR_k \;\; T_{c,k} \;\; T_{a,k} \;\; T_{s,k}] \,. \tag{3.4}$$

Further, define the vector ϕ to be comprised of the eleven parameters as summarized in Table 3.4, that is,

$$\phi = [m \;\; b \;\; A_{m,max} \;\; g_{mp} \;\; S_2 \;\; S_4 \;\; c \;\; d \;\; V_{max} \;\; \epsilon \;\; w_\pi] \,. \tag{3.5}$$

Then we present the raw SSE computed as

$$\text{Raw SSE} = \sum_{k=1}^{n} [y_k - g_\ell(p_k; \phi)]^2 \tag{3.6}$$

where y_k is the data point for $k = 1, \ldots, n$ and $g_\ell(p_k; \phi)$ is the model prediction using the input environmental parameters p_k and the species-specific parameters ϕ. The normalized SSE is calculated as

$$\text{Normalized SSE} = \frac{\text{Raw SSE}}{n} \tag{3.7}$$

where n is the number of data points.

By examining the sum of squared error values as normalized by number of data points, the model fits the N01 genotype better than the Haskell genotype and the 2009 growing season better than the 2008 season for both genotypes. Table 3.6 further illustrates the SSE values

Table 3.5: Sum of squared errors.

	2008		2009	
	Haskell	N01	Haskell	N01
Raw SSE	5.3540×10^8	4.5684×10^8	3.6155×10^8	3.3720×10^8
No. data pts	1116	1116	935	937
Normalized SSE	4.7975×10^5	4.0936×10^5	3.8669×10^5	3.5987×10^5

as separated by experimental plot. The model predictions agree with data from the "dry" experiment better than the other scenarios.

As confirmed by both the visual comparison of model predictions to data as well as with the quantitative evidence presented by the large SSE values, this model requires refinement to adequately predict total leaf conductance. The conductance data show a strong relationship with plant development. During the early vegetative stages, conductance increases, and during senescence, the conductance declines. This model does not contain any dependence on plant development which likely inhibits its performance. The data also suggest a relationship between conductance and soil moisture status. Although the model does incorporate soil moisture status, it is only through the calculation of catalytic Rubisco capacity for the leaf, V_m. It is unclear if V_m is the correct or only state variable which is directly impacted by soil moisture. Further, the formulation of this dependence as $f(w_2)$ in (2.9) may not be the correct relation to characterize the relationship. Perhaps a formulation analogous to the temperature dependency $f(T)$ as shown in (2.8) would yield more accurate predictions. The model as presented in Niyogi et al. is a very detailed depiction of the coupled stomatal conductance and photosynthetic processes. Without validation data to confirm predictions of the many internal state variables (i.e., net photosynthesis rate A_n, CO_2 partial pressure at leaf surface C_s and in intercellular spaces C_i, etc), it is difficult to isolate the equations which require modification to improve model prediction agreement with data.

Table 3.6: Normalized sum of squared errors. Each entry is the SSE value divided by 10^5. The values in parenthesis are the number of data points used for that experiment.

	Dry	Medium	Wet	Open	Total
2008 Haskell	3.2271 (279)	4.1472 (278)	5.9821 (279)	5.8277 (280)	4.7975 (1116)
2008 N01	2.9541 (280)	3.7792 (279)	4.7466 (280)	4.9021 (277)	4.0936 (1116)
2009 Haskell	2.8207 (232)	3.1353 (232)	4.5041 (239)	4.9882 (232)	3.8669 (935)
2009 N01	2.3767 (234)	3.0583 (231)	4.2454 (239)	4.6984 (233)	3.5987 (937)

Chapter 4

Model Development and Calibration

4.1 Model Development

Whereas inadequacies in model predictions may be due to a number of factors, two critical phenomenon that are presently lacking are model dependence on plant development and soil moisture conditions. Therefore, we suggest that total leaf conductance (g_ℓ) can be modeled by a relation of the form

$$g_\ell(t, w_2; \theta) = f(t; \theta) \cdot h(w_2; \theta) \tag{4.1}$$

where $f(t;\theta)$ is a scaling function valued between 0 and 1 that depends on plant age (t) and $h(w_2;\theta)$ is a function describing conductance with units of mmol m^{-2} s^{-1} based on soil moisture fraction (w_2).

To model the total leaf conductance data using (4.1), we employ a polynomial

$$f(t;\theta) = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + \alpha_3 t^3 \tag{4.2}$$

to define the plant age dependency. Here t is the plant age in days and α_k for $k = 0, \ldots, 3$ are coefficients to be estimated through a least squares fit to data.

The conductance-dependence on soil moisture is defined as a Gompertz-like function

$$h(w_2;\theta) = \beta_1 e^{\beta_2 e^{\beta_3 w_2}} \tag{4.3}$$

where β_1 is the maximum allowable conductance, β_2 sets the displacement of the sigmoidal curve along the w_2 axis, and β_3 represents the extent of dependence of conductance on w_2 [56]. Both β_2 and β_3 are negative values. A more negative β_2 value shifts the curve towards higher w_2 values, indicating that smaller soil moisture values will have lower conductance rates. A

more negative β_3 value increases the slope of the sigmoid implying that the region of w_2 values transitioning from the lowest value of conductance to the highest becomes more narrow. The complete set of parameters is given by $\theta = [\alpha_0, \alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3]$.

4.2 Model Calibration

Parameter estimation was performed in a two-step process. In the first step, we employed a global search algorithm to obtain initial parameter values for a classical Nelder-Mead simplex method used in the second step. The data was collected at times t_{k_1}, $k_1 = 1, \ldots, n_1$, and soil moisture fractions w_{2,k_2}, $k_2 = 1, \ldots, n_2$, for a total of n measurements. Note that the total number of measurements for each case is summarized in Table 4.2. The cost function used throughout the optimization process is taken to be

$$\text{SSE} = \sum_{k_1=1}^{n_1} \sum_{k_2=1}^{n_2} [y_{k_1,k_2} - g_\ell(t_{k_1}, w_{2,k_2}; \theta)]^2 \tag{4.4}$$

$$= \sum_{k=1}^{n} [y_k - g_\ell(e_k; \theta)]^2 \tag{4.5}$$

where y_k for $k = 1, \ldots, n$ are the measured leaf conductances and $g_\ell(t_{k_1}, w_{2,k_2}; \theta) = g_\ell(e_k; \theta)$ are modeled values given by (4.1). Here e_k denotes the unique input associated with the measurement (t_{k_1}, w_{2,k_2}).

Because little information was known about feasible values for the unknown parameters, we chose to use a global search algorithm, capable of searching a large parameter space, for the first phase of optimization. In particular we used the Differential Evolution algorithm [62] defining the possible parameter space as indicated in Table 4.1. This algorithm begins with a

Table 4.1: Parameter space regions utilized with initial global search.

Parameter	Minimum value	Maximum value
α_0	-1000	1000
α_1	-1000	1000
α_2	-1000	1000
α_3	-1000	1000
β_1	0	10000
β_2	-100	0
β_3	-100	0

large population NP (here we used NP=1000) of parameter vectors chosen randomly to cover the entire parameter space. Essentially, this forms an $NP \times p$ matrix X where p is the length of each parameter vector (here p=7). The rows of this matrix are randomly perturbed to form three unique $NP \times p$ matrices X_1, X_2, and X_3. Then an F value, where $F \in [0, 2]$, is chosen to control the amplification of the differential variation

$$\hat{X} = X_3 + F \cdot (X_1 - X_2). \tag{4.6}$$

Our results reflect the choice of F=0.7. Equation 4.6 describes the mutation step of the algorithm which is then followed by the crossover step which increases the diversity of the perturbed vectors. In the crossover step, a CR value where $CR \in [0, 1]$ is chosen. Our results reflect the choice of CR=0.9. A random $NP \times p$ matrix Y containing values normally distributed over [0,1] is created. Finally, a new matrix $\tilde{X}$ is created, using the Y, $\hat{X}$, and X matrices, with components

$$\tilde{x}_{ij} = \begin{cases} \hat{x}_{ij} & \text{if } y_{ij} < CR \\ x_{ij} & \text{if } y_{ij} \geq CR \end{cases} \tag{4.7}$$

for $i = 1, \ldots, n$ and $j = 1, \ldots, p$. In general, a choice of a larger CR value creates a more perturbed population in the crossover step. At this point, the resultant cost values for the ith row of X are compared to the resultant cost values for the ith row of $\tilde{X}$ for $i = 1, \ldots, NP$. If the cost value for the parameter vector in $\tilde{X}$ is lower, the row of $\tilde{X}$ replaces that row of X. This completes the first generation. The entire process is continued for a user-specified number of generations (here 2000 generations were created). The parameter vector yielding the lowest cost in this final generation is returned as the globally minimizing parameter vector.

After completion of the initial global search for each case, the globally minimizing parameter vector was used as an initial guess for the Nelder-Mead simplex method which was subsequently employed to achieve local convergence.

Each of the four scenarios representing all combinations of growing season and genotype (that is, Haskell/2008, N01/2008, Haskell/2009, and N01/2009) were calibrated to the model described in (4.1) - (4.3). Table 4.2 summarizes the parameter estimates and resultant sum of squared error values. Figures 4.1 through 4.8 illustrate the model predictions as compared to data.

Note that as described in Table 3.2, not all available conductance data have associated soil moisture information. Figures 4.1 through 4.8 compare only conductance data having the associated soil moisture values to model predictions. Therefore, when comparing these figures to the raw data in Figures 3.23 - 3.34, one will notice the omission of some conductance

measurements. This effect is particularly noticeable in the 2009 results depicted in Figures 4.6 and 4.8 where there were 11 days worth of conductance measurements for which associated soil moisture values were not available.

Figure 4.9 depicts dependence on plant age of the modeled total leaf conductance. Each genotype and growing season is included for comparison. Note that as indicated in (4.2), the function $f(t)$ is valued between 0 and 1 for the entire plant age range. In all cases, the scaling imposed by plant age is lower in the early vegetative stage of growth, peaks at approximately 70 days of age, and declines again as the plant enters senescence. It is interesting to note the similarity of modeled age dependence between growing seasons rather than genotypes. This suggests that the model may be further improved by incorporating another mechanism such as temperature. Figure 4.10 illustrates the modeled maximum allowable conductance rate as a function of soil moisture fraction. Here the sigmoidal shape of (4.3) is evident for all genotypes and growing seasons. In general, as the soil moisture fraction increases, the potential conductance increases until it reaches a plateau. Again one may note the similarity of modeled soil moisture dependence between growing seasons rather than genotypes. However, in both the 2008 and 2009 growing seasons the N01 genotype has a smaller maximum allowable conductance rate than the Haskell genotype.

Table 4.2: Parameter estimates for each case.

	2008		2009	
Parameter	Haskell	N01	Haskell	N01
α_0	-1.5602×10^{0}	-1.4194×10^{0}	-1.6634×10^{-1}	-1.2527×10^{-1}
α_1	8.8321×10^{-2}	8.2211×10^{-2}	3.4967×10^{-2}	3.4066×10^{-2}
α_2	-9.4367×10^{-4}	-8.6564×10^{-4}	-3.0269×10^{-4}	-2.9766×10^{-4}
α_3	2.9187×10^{-6}	2.6421×10^{-6}	5.8304×10^{-7}	5.7756×10^{-7}
β_1	1.0173×10^{3}	9.0622×10^{2}	8.8439×10^{2}	8.5056×10^{2}
β_2	-5.4780×10^{1}	-3.1492×10^{1}	-1.3280×10^{1}	-1.0871×10^{1}
β_3	-3.3652×10^{1}	-3.2794×10^{1}	-1.7424×10^{1}	-1.6729×10^{1}
SSE	3.9567×10^{7}	3.0781×10^{7}	2.2497×10^{7}	1.9445×10^{7}
No. data pts	1184	1184	973	975

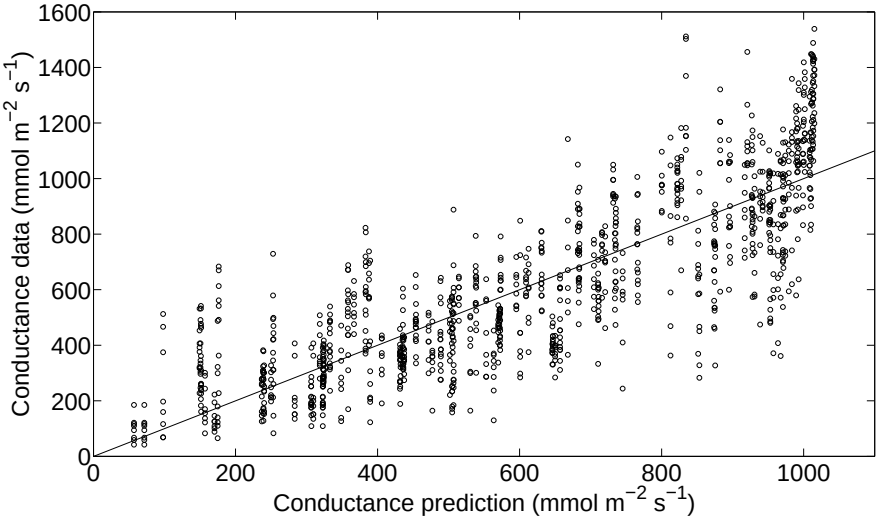


Figure 4.1: Conductance predictions versus data for all experimental plots with the Haskell genotype in 2008. The 45° line is superimposed to indicate scale.

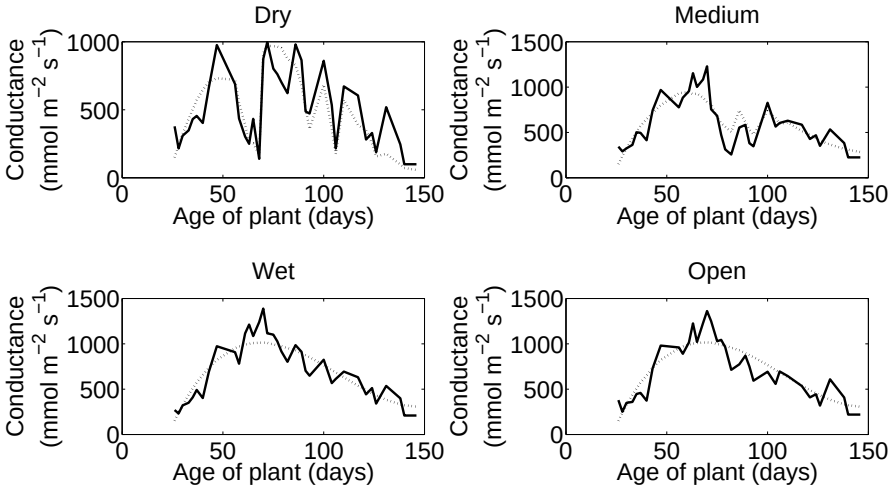


Figure 4.2: Conductance predictions (dotted lines) compared to mean data by date (solid lines) for the Haskell genotype in 2008.

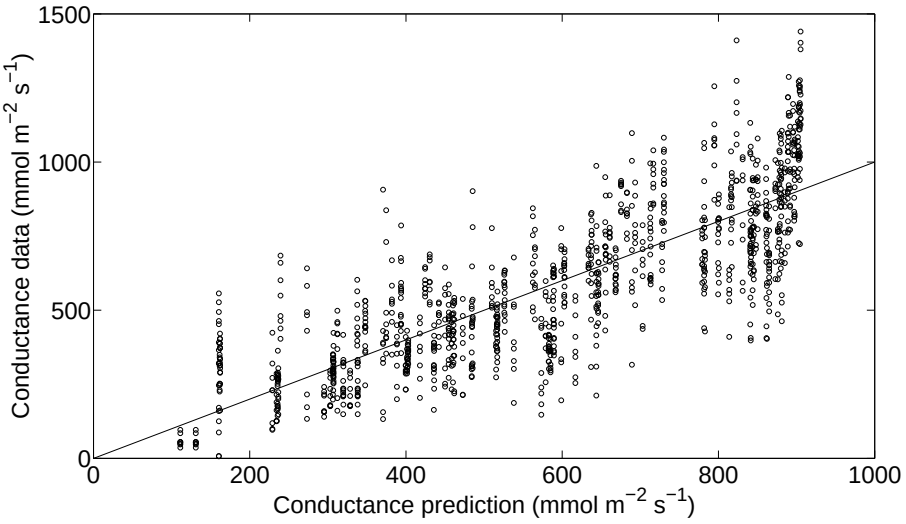


Figure 4.3: Conductance predictions versus data for all experimental plots with the N01 genotype in 2008. The 45° line is superimposed to indicate scale.

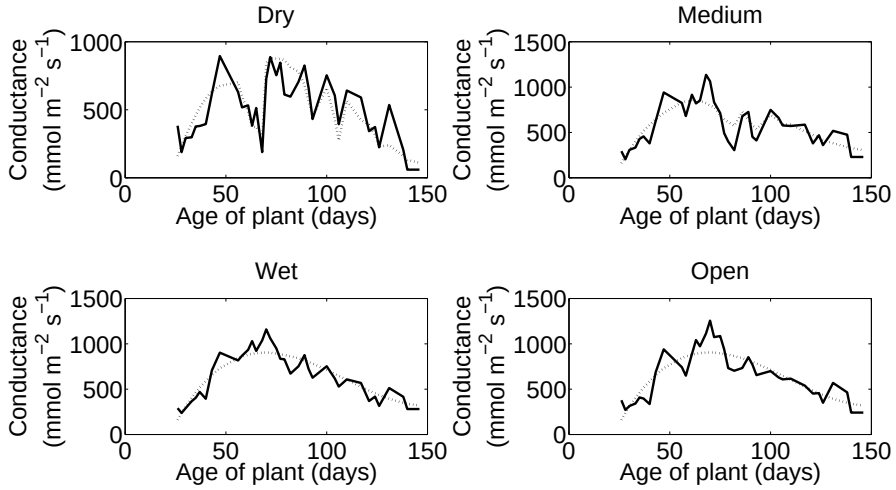


Figure 4.4: Conductance predictions (dotted lines) compared to mean data by date (solid lines) for the N01 genotype in 2008.

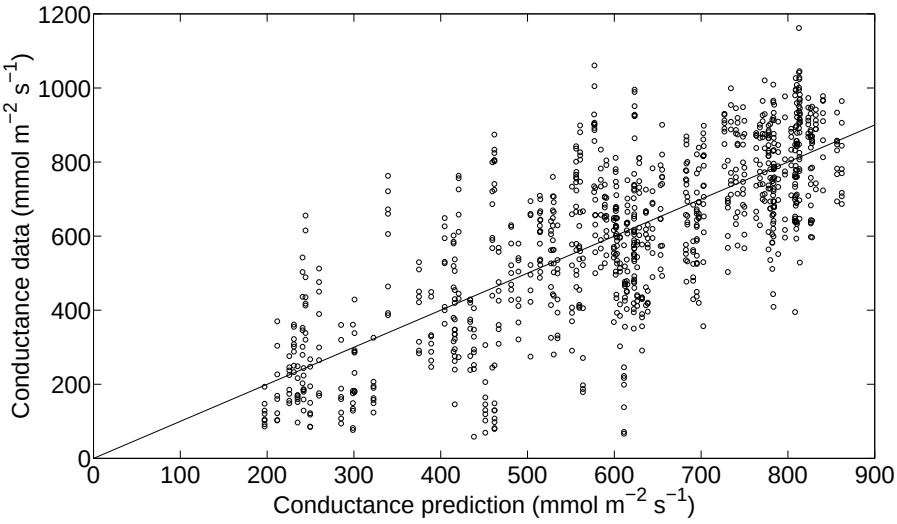


Figure 4.5: Conductance predictions versus data for all experimental plots with the Haskell genotype in 2009. The 45° line is superimposed to indicate scale.

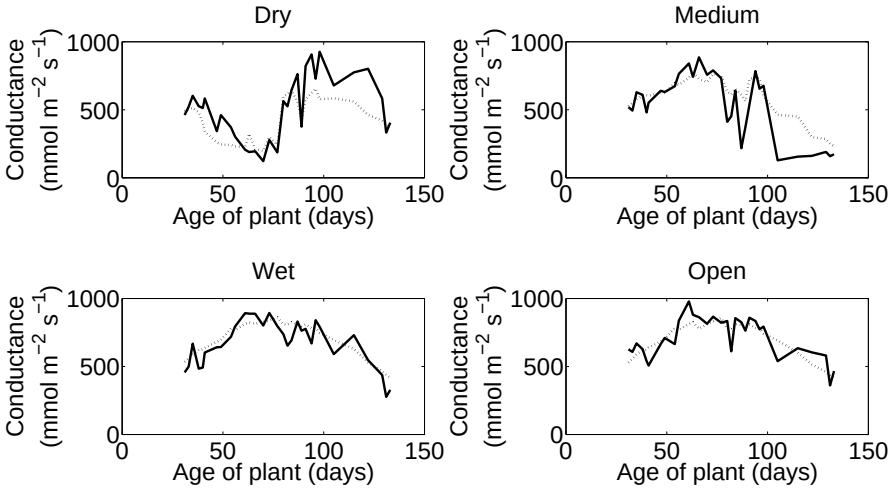


Figure 4.6: Conductance predictions (dotted lines) compared to mean data by date (solid lines) for the Haskell genotype in 2009.

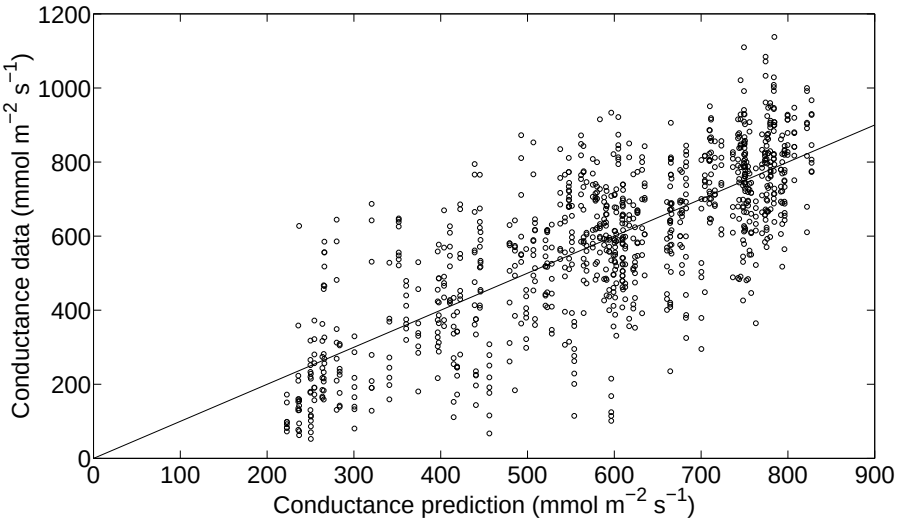


Figure 4.7: Conductance predictions versus data for all experimental plots with the N01 genotype in 2009. The 45° line is superimposed to indicate scale.

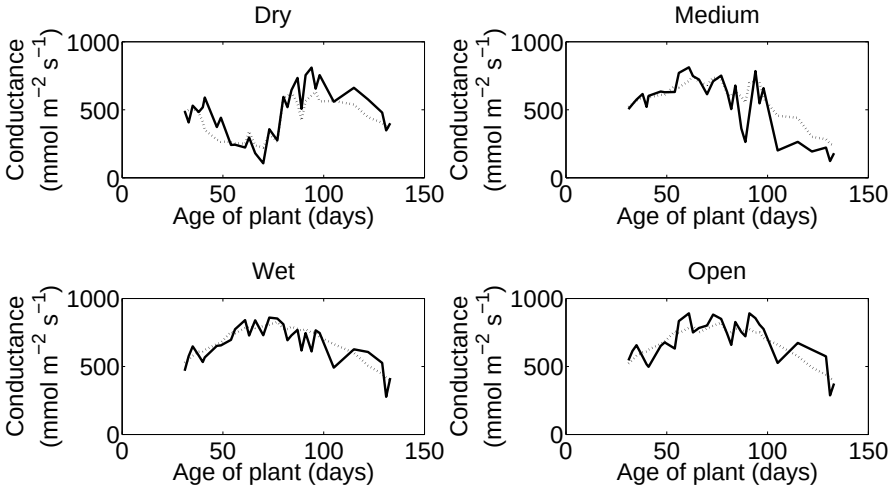


Figure 4.8: Conductance predictions (dotted lines) compared to mean data by date (solid lines) for the N01 genotype in 2009.

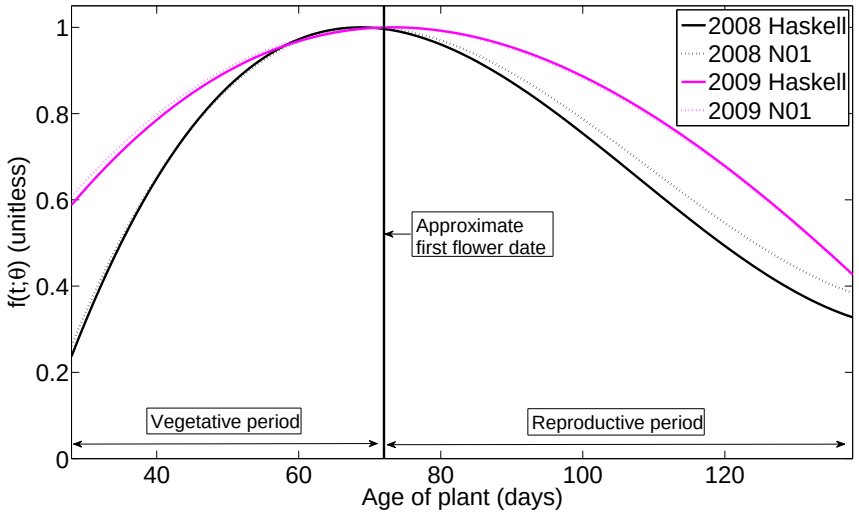


Figure 4.9: Model plant age dependence curves $f(t;\theta)$ as described by (4.2) and Table 4.2.

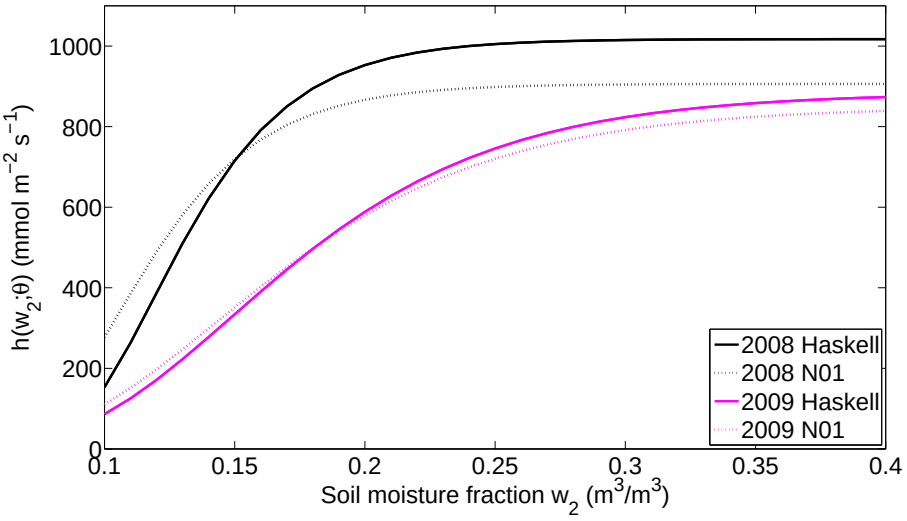


Figure 4.10: Model soil moisture dependence curves $h(w_2;\theta)$ as described by (4.3) and Table 4.2.

4.3 Confidence Intervals for Parameter Estimates

In Table 4.2 we present the parameter estimates for the seven unknown parameters of the model described by (4.1). These values are only point estimates and it is important to understand the accuracy and reliability of these parameter values. Confidence intervals for the parameter estimates provide an indication of this reliability. For example, if the derived confidence interval for a given parameter is large, this indicates to the modeler that the parameter estimate may be unreliable. Reciprocally, a small confidence interval indicates the parameter is likely well estimated.

We begin by applying traditional asymptotic theory to establish confidence intervals for the estimated parameters. However, as shown in Section 4.3.1, this method does not provide a computationally feasible solution. We then describe bootstrap methods in general and choose in particular the wild bootstrap, which is applicable to our case where the residuals exhibit heteroscedasticity, in Section 4.3.2.

4.3.1 Asymptotic Theory

We let n denote the number of data points, p the number of estimated parameters, and $\hat{\theta}$ the estimated parameter vector. Then, define the $n \times p$ sensitivity matrix $\chi(\hat{\theta})$ to have components

$$\chi_{ij}(\hat{\theta}) = \left.\frac{\partial g_\ell(e_i;\theta)}{\partial \theta_j}\right|_{\theta=\hat{\theta}} \tag{4.8}$$

for $i = 1, \ldots, n$ and $j = 1, \ldots, p$. The variance-covariance matrix V may then be estimated as

$$V = s^2 \left(\chi^T(\hat{\theta})\chi(\hat{\theta})\right)^{-1} \tag{4.9}$$

where the variance estimate is $s^2 = \frac{\text{SSE}}{n-p}$, with SSE defined in (4.5).

Oftentimes, a $1-\alpha$ joint inference region for estimated parameters within a nonlinear model can be adequately approximated via the inequality

$$\left(\theta - \hat{\theta}\right)^T V^{-1} \left(\theta - \hat{\theta}\right) \leq pF(p, n-p; \alpha) \tag{4.10}$$

where F represents the F distribution [8, 57]. The associated $1-\alpha$ confidence interval for an individual estimated parameter $\hat{\theta}_j$ is

$$\hat{\theta}_j \pm \sqrt{V_{jj}}\ \ t(n-p; \alpha/2) \tag{4.11}$$

where t represents the Student's T distribution.

Calculating (4.9) was the first step taken to approximate confidence intervals for the estimated parameters. Partial derivatives of the model, as described in (4.1), with respect to each of the parameters are

$$\chi_{i,1} = \frac{\partial g_\ell(e_i;\theta)}{\partial \alpha_0} = \beta_1 e^{\beta_2 e^{\beta_3 w_{2,i}}} \tag{4.12}$$

$$\chi_{i,2} = \frac{\partial g_\ell(e_i;\theta)}{\partial \alpha_1} = t_i \beta_1 e^{\beta_2 e^{\beta_3 w_{2,i}}} \tag{4.13}$$

$$\chi_{i,3} = \frac{\partial g_\ell(e_i;\theta)}{\partial \alpha_2} = t_i^2 \beta_1 e^{\beta_2 e^{\beta_3 w_{2,i}}} \tag{4.14}$$

$$\chi_{i,4} = \frac{\partial g_\ell(e_i;\theta)}{\partial \alpha_3} = t_i^3 \beta_1 e^{\beta_2 e^{\beta_3 w_{2,i}}} \tag{4.15}$$

$$\chi_{i,5} = \frac{\partial g_\ell(e_i;\theta)}{\partial \beta_1} = (\alpha_3 t_i^3 + \alpha_2 t_i^2 + \alpha_1 t_i + \alpha_0) e^{\beta_2 e^{\beta_3 w_{2,i}}} \tag{4.16}$$

$$\chi_{i,6} = \frac{\partial g_\ell(e_i;\theta)}{\partial \beta_2} = \beta_1 (\alpha_3 t_i^3 + \alpha_2 t_i^2 + \alpha_1 t_i + \alpha_0) e^{\beta_2 e^{\beta_3 w_{2,i}}} e^{\beta_3 w_{2,i}} \tag{4.17}$$

$$\chi_{i,7} = \frac{\partial g_\ell(e_i;\theta)}{\partial \beta_3} = \beta_1 \beta_2 w_{2,i} (\alpha_3 t_i^3 + \alpha_2 t_i^2 + \alpha_1 t_i + \alpha_0) e^{\beta_2 e^{\beta_3 w_{2,i}}} e^{\beta_3 w_{2,i}}. \tag{4.18}$$

These values yield the $n \times p$ sensitivity matrix χ

$$\chi = \begin{bmatrix} \chi_{1,1} & \chi_{1,2} & \chi_{1,3} & \chi_{1,4} & \chi_{1,5} & \chi_{1,6} & \chi_{1,7} \\ \chi_{2,1} & \chi_{2,2} & \chi_{2,3} & \chi_{2,4} & \chi_{2,5} & \chi_{2,6} & \chi_{2,7} \\ \vdots & & & & & & \vdots \\ \chi_{n,1} & \chi_{n,2} & \chi_{n,3} & \chi_{n,4} & \chi_{n,5} & \chi_{n,6} & \chi_{n,7} \end{bmatrix}. \tag{4.19}$$

This matrix χ was formed for the case of the Haskell genotype for the 2008 growing season. Using the parameter values indicated in Table 4.2 for this case we have

$$\chi^T\chi = \begin{bmatrix} 1\times10^{9} & 2\times10^{11} & 5\times10^{13} & 1\times10^{16} & 4\times10^{6} & 5\times10^{6} & -4\times10^{7} \\ 2\times10^{11} & 5\times10^{13} & 1\times10^{16} & 3\times10^{18} & 1\times10^{9} & 1\times10^{9} & -1\times10^{10} \\ 5\times10^{13} & 1\times10^{16} & 3\times10^{18} & 7\times10^{20} & 3\times10^{11} & 3\times10^{11} & -3\times10^{12} \\ 1\times10^{16} & 3\times10^{18} & 7\times10^{20} & 2\times10^{23} & 7\times10^{13} & 8\times10^{13} & -7\times10^{14} \\ 4\times10^{6} & 1\times10^{9} & 3\times10^{11} & 7\times10^{13} & 4\times10^{4} & 4\times10^{4} & -3\times10^{5} \\ 5\times10^{6} & 1\times10^{9} & 3\times10^{11} & 8\times10^{13} & 4\times10^{4} & 4\times10^{5} & -3\times10^{6} \\ -4\times10^{7} & -1\times10^{10} & -3\times10^{12} & -7\times10^{14} & -3\times10^{5} & -3\times10^{6} & 2\times10^{7} \end{bmatrix}.$$

Note that the order of magnitude of the entries in $\chi^T\chi$ varies from 10^4 to 10^{23}, which is a difference of 19 orders of magnitude. Based on the singular value decomposition of $\chi^T\chi$, Matlab computes that $\text{rank}(\chi^T\chi) = 3$. Since this matrix is not full rank, we can not compute $(\chi^T\chi)^{-1}$ as required by the linear approximation parameter confidence interval formula in (4.9). Therefore, asymptotic theory for this example does not provide a computationally feasible solution.

This problem of singularity of the $\chi^T\chi$ matrix is not uncommon for inverse problems with nonlinear models. A method to handle singularity of the Fisher Information Matrix $(\chi^T\chi)$ as a result of poor parameter identifiability and estimability has recently been introduced in [19]. This algorithm determines which parameter axes lie closest to the ill-conditioned directions of $\chi^T\chi$ and then implements a reduced-order estimation by fixing these associated parameter values at prior estimates. Because reasonable ranges for the estimated parameters in this application are unknown, we were unable to employ this method.

4.3.2 Bootstrap Method

As an alternative to the asymptotic theory to derive confidence intervals for the estimated parameters, we turn to a form of bootstrapping. This nonparametric approach to statistical inference substitutes intensive computation for more traditional asymptotic results and can be used to derive confidence intervals [30].

In particular, we apply the bootstrap to the residuals of our model. This method treats the model dependent variables (i.e., soil moisture, plant age) as fixed rather than random because they may be a product of experimental design.

Assume that the experimental data $(y_1, e_1), \ldots, (y_n, e_n)$ are from an underlying observation process

$$Y_k = g_\ell(e_k; \theta_0) + \mathcal{E}_k \tag{4.20}$$

for $k = 1, \ldots, n$ where the errors $\mathcal{E}_k$ are assumed to be independent and identically distributed

(iid) from a distribution F with mean 0 and constant variance, and θ_0 is the "true" parameter value.

First, the ordinary least squares estimate $\hat{\theta}$ is estimated by minimizing (4.5). Then, the residuals

$$u_k = y_k - g_\ell(e_k; \hat{\theta}) \tag{4.21}$$

for $k = 1, \ldots, n$ are computed. Bootstrapped data values $\bar{y}_k^m$ are formed by

$$\bar{y}_k^m = g_\ell(e_k; \hat{\theta}) + \bar{u}_k^m \tag{4.22}$$

where $\bar{u}_k^m$ is an element randomly selected with replacement from the residual vector u and $m = 1, \ldots, M$ where M is a sufficiently large number. Thus, each bootstrap error term $\bar{u}_k^m$ can take on n possible values, that is, the values in u, each with probability $1/n$. Finally, the ordinary least squares estimate $\hat{\theta}^m$ is determined by

$$\hat{\theta}^m = \arg\min_\theta \sum_{k=1}^{n} \left[\bar{y}_k^m - g_\ell(e_k; \theta)\right]^2 . \tag{4.23}$$

In the results presented here, we chose $M = 1000$. This technique of bootstrapping the residuals and re-estimating the parameter vector θ M times is referred to as residual bootstrapping [27, 30, 41].

However, the residual bootstrap as described is not valid if the error terms $\mathcal{E}_k$ are not iid because if the heteroscedasticity is of an unknown form, it cannot be emulated with the bootstrap distribution [24, 41]. Note that if the residuals are iid, a plot of residual values versus the associated model predictions will appear random [7, 18, 25]. As shown in Figure 4.11, the pattern in all cases does not appear random and therefore we conclude that the residuals are heteroscedastic.

Consider next the case where the experimental data $(y_1,\ e_1),\ \ldots,\ (y_n,\ e_n)$ are from an underlying observation process

$$Y_k = g_\ell(e_k; \theta_0)(1 + \mathcal{E}_k) \tag{4.24}$$

for $k = 1, \ldots, n$ where the $\mathcal{E}_k$ are iid from a distribution F with mean 0 and non-constant variance, and θ_0 is the "true" parameter value.

Several sophisticated variants of bootstrapping exist to address this case of heteroscedastic residuals. Among the many options are the generalized least squares (GLS) method and the wild bootstrap. GLS is well documented as a good theoretical alternative approach to parameter estimation in the presence of non-constant variance [6, 18, 23, 30, 31]. Briefly, in GLS, one

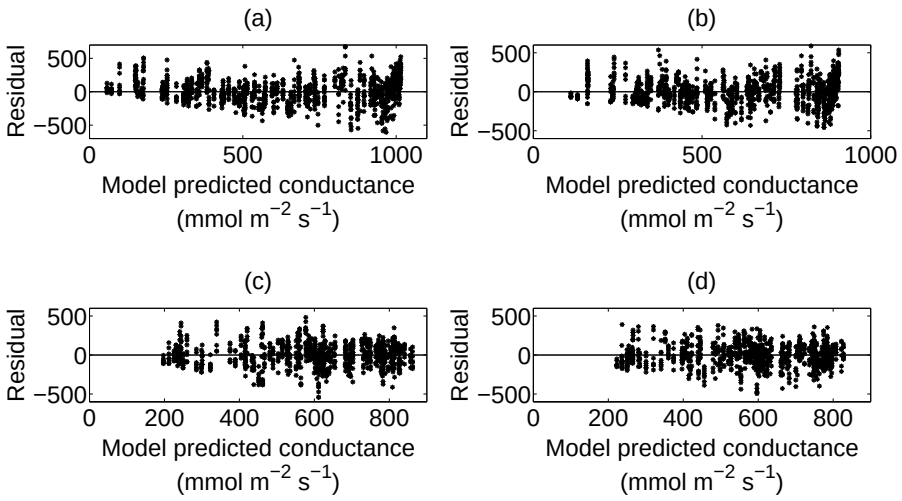


Figure 4.11: Conductance predictions versus residuals. (a) Haskell genotype in 2008, (b) N01 genotype in 2008, (c) Haskell genotype in 2009, and (d) N01 genotype in 2009.

searches for θ which minimizes

$$\text{SSE} = \sum_{k=1}^{n} \hat{w}_k \left[y_k - g_\ell(e_k; \theta) \right]^2 \tag{4.25}$$

where $\hat{w}_k$ are weights designed to give greater weight to observations with smaller variance. Unfortunately, in practice, the GLS method often does not live up to the theoretical advantages. Because the weights $\hat{w}_k$ are unknown and must be estimated, there is a possibility that the GLS regression may in fact "miss" the data entirely [57].

Hence, the method chosen to address the heteroscedastic residuals is the wild bootstrap method. This method was originally proposed by Wu and rigorously expanded by Liu as a general method for resampling residuals in the presence of error variance heteroscedasticity [40, 68]. The method is referred to as "wild" because n different distributions are estimated from only n observations. The procedure is identical to that described above for the residual bootstrap with the exception that (4.22) is replaced by

$$\bar{y}_k^m = g_\ell(e_k; \hat{\theta}) + h(u_k) v_k \tag{4.26}$$

where $h(u_k)$ is a transformation of the kth residual and v_k is a random variable from a distribution satisfying $\text{E}(v_k)=0$, $\text{E}(v_k^2)=1$, and $\text{E}(v_k^3)=1$. Clearly there are many options for how one

may choose to define $h(u_k)$ and the distribution of v_k. A simple choice for $h(u_k)$, and the one we chose, is the identity relation $h(u_k) = u_k$. A commonly employed two-point distribution for v_k is

$$v_k = \begin{cases} \frac{-(\sqrt{5}-1)}{2} & \text{with probability } p = (\sqrt{5}+1)/(2\sqrt{5}) \\ \frac{\sqrt{5}+1}{2} & \text{with probability } 1-p \end{cases} \tag{4.27}$$

[24, 41, 42]. Clearly this distribution satisfies the aforementioned requirements because

$$\begin{aligned} E(v_k) &= \frac{-(\sqrt{5}-1)}{2} \cdot \frac{\sqrt{5}+1}{2\sqrt{5}} + \frac{\sqrt{5}+1}{2} \cdot \left(1 - \frac{\sqrt{5}+1}{2\sqrt{5}}\right) \\ &= \frac{-5+1}{4\sqrt{5}} + \frac{10+2\sqrt{5}}{4\sqrt{5}} - \frac{5+2\sqrt{5}+1}{4\sqrt{5}} \\ &= \frac{-4+10+2\sqrt{5}-6-2\sqrt{5}}{4\sqrt{5}} \\ &= 0, \end{aligned}$$

$$\begin{aligned} E(v_k^2) &= \left(\frac{-(\sqrt{5}-1)}{2}\right)^2 \cdot \frac{\sqrt{5}+1}{2\sqrt{5}} + \left(\frac{\sqrt{5}+1}{2}\right)^2 \cdot \left(1 - \frac{\sqrt{5}+1}{2\sqrt{5}}\right) \\ &= \left(\frac{6-2\sqrt{5}}{4}\right)\left(\frac{\sqrt{5}+1}{2\sqrt{5}}\right) + \left(\frac{6+2\sqrt{5}}{4}\right)\left(1 - \frac{\sqrt{5}+1}{2\sqrt{5}}\right) \\ &= \frac{-4+4\sqrt{5}}{8\sqrt{5}} + \frac{20+12\sqrt{5}}{8\sqrt{5}} - \frac{16+8\sqrt{5}}{8\sqrt{5}} \\ &= \frac{8\sqrt{5}}{8\sqrt{5}} \\ &= 1, \end{aligned}$$

and

$$
\begin{aligned}
E(v_k^3) &= \left(\frac{-(\sqrt{5}-1)}{2}\right)^3 \cdot \frac{\sqrt{5}+1}{2\sqrt{5}} + \left(\frac{\sqrt{5}+1}{2}\right)^3 \cdot \left(1 - \frac{\sqrt{5}+1}{2\sqrt{5}}\right) \\
&= \left(\frac{16-8\sqrt{5}}{8}\right)\left(\frac{\sqrt{5}+1}{2\sqrt{5}}\right) + \left(\frac{16+8\sqrt{5}}{8}\right)\left(1 - \frac{\sqrt{5}+1}{2\sqrt{5}}\right) \\
&= \frac{-24+8\sqrt{5}}{16\sqrt{5}} + \frac{80+32\sqrt{5}}{2\sqrt{5}} - \frac{56+24\sqrt{5}}{16\sqrt{5}} \\
&= \frac{16\sqrt{5}}{16\sqrt{5}} \\
&= 1.
\end{aligned}
$$

Figures 4.12 - 4.15 illustrate the resultant histograms of parameter estimates from the wild bootstrap with M=1000 applied to the 2008 Haskell study, the 2008 N01 study, the 2009 Haskell study, and the 2009 N01 study, respectively. The vertical lines present in each subplot indicate the original $\hat{\theta}$ values.

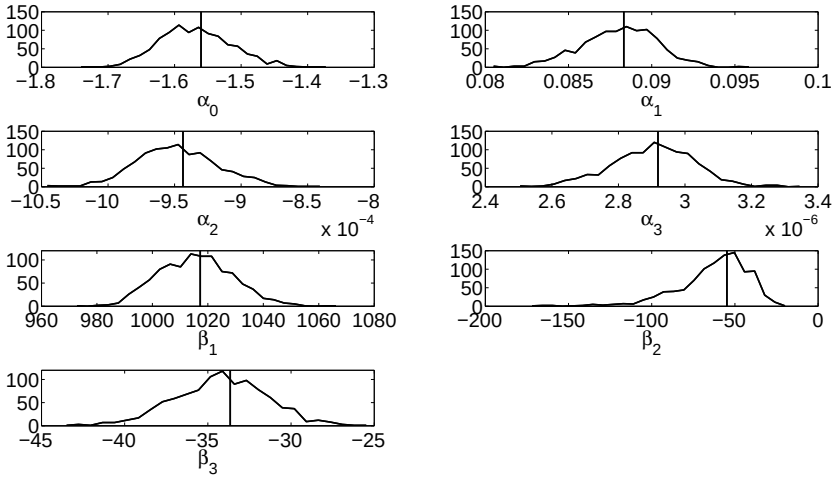


Figure 4.12: Histograms of bootstrapped parameter estimates of $\alpha_0 - \alpha_3$ and $\beta_1 - \beta_3$ for the Haskell genotype in 2008.

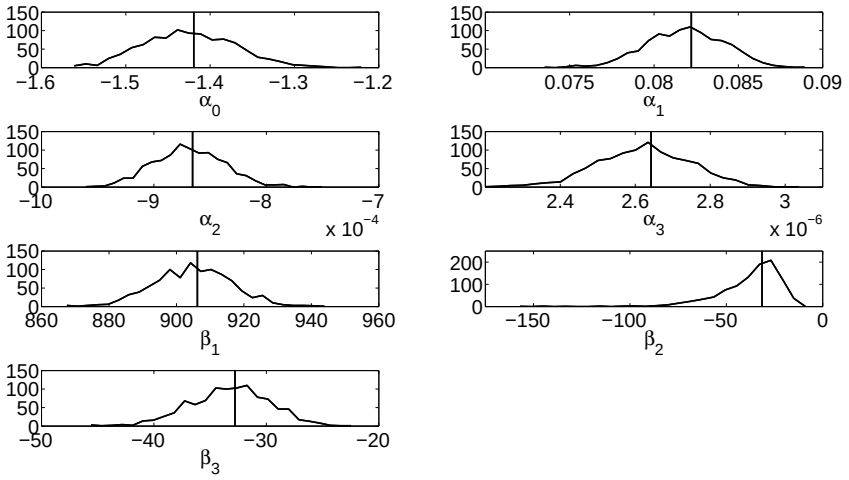


Figure 4.13: Histograms of bootstrapped parameter estimates of $\alpha_0 - \alpha_3$ and $\beta_1 - \beta_3$ for the N01 genotype in 2008.

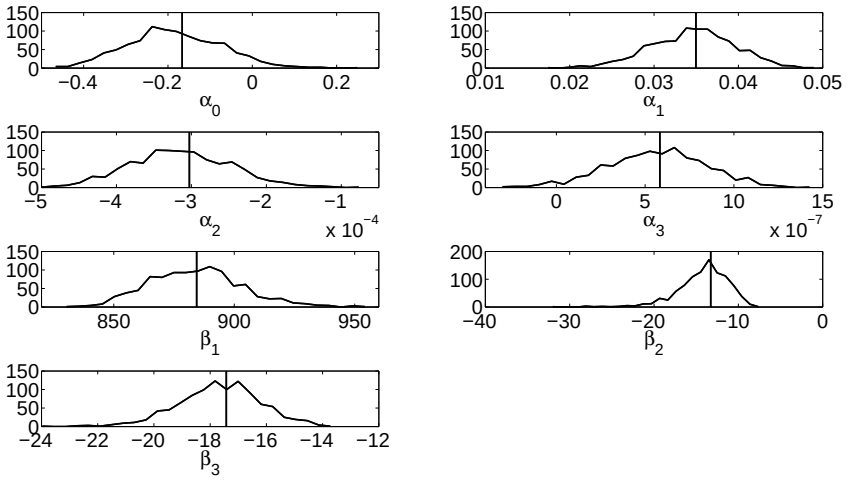


Figure 4.14: Histograms of bootstrapped parameter estimates of $\alpha_0 - \alpha_3$ and $\beta_1 - \beta_3$ for the Haskell genotype in 2009.

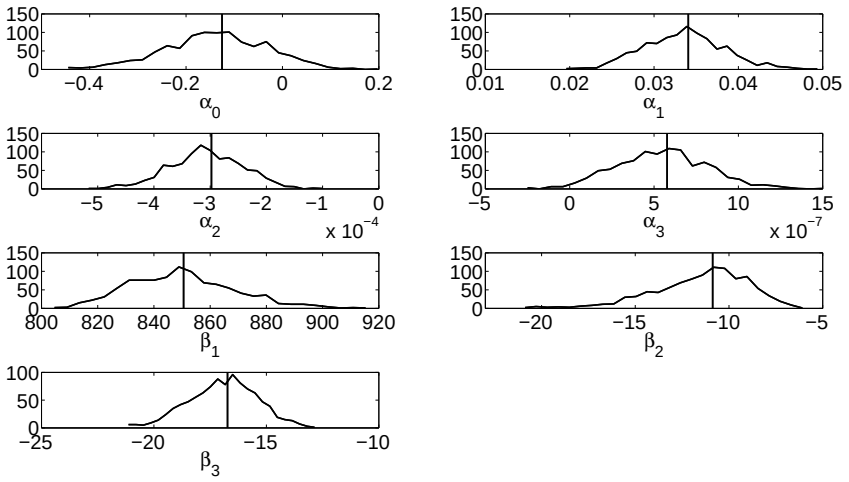


Figure 4.15: Histograms of bootstrapped parameter estimates of $\alpha_0 - \alpha_3$ and $\beta_1 - \beta_3$ for the N01 genotype in 2009.

From the bootstrap simulations, there are M estimates of θ as illustrated in the previous histograms. The percentiles of these histograms may be used to construct confidence intervals for the original estimate $\hat{\theta}$ [27]. Since in these studies $M = 1000$, the vector of M estimates is ordered and the 25th and 975th entry define the 95% confidence interval for each parameter. Table 4.3 summarizes the 95% confidence intervals for each parameter for each of the four studies.

The results of the bootstrap method may also be used to approximate the parameter covariance matrix using the relation

$$\mathrm{cov}(\hat{\theta}) = \frac{1}{M-1} \sum_{m=1}^{M} (\hat{\theta^m} - \hat{\theta}_{boot})^T (\hat{\theta^m} - \hat{\theta}_{boot}) \tag{4.28}$$

where

$$\hat{\theta}_{boot} = \frac{1}{M} \sum_{m=1}^{M} \hat{\theta}^m \tag{4.29}$$

is the mean of all M bootstrapped parameter vector estimates [18]. Using the definition given by (4.28) for the covariance matrix, the standard errors for each of each of the $j = 1, \ldots, p$ estimated parameters are [30]

$$\mathrm{SE}(\hat{\theta_j}) = \sqrt{\mathrm{cov}(\hat{\theta})_{jj}}. \tag{4.30}$$

In (4.11), the substitution of $\mathrm{SE}(\hat{\theta}_j)$ for $\sqrt{V_{jj}}$ yields our confidence intervals.

Recall that this equation defined confidence intervals for the estimated parameters as motivated by asymptotic theory. The bootstrap method was chosen because of the matrix singularity which rendered calculation of the covariance matrix via sensitivities based methods impossible. However, since the covariance matrix may be estimated from the bootstrap method, the confidence intervals as derived directly from the bootstrap method and as estimated with asymptotic theory may be compared. The 95% confidence intervals for the estimated parameters computed with the asymptotic theory method are summarized in Table 4.4.

Comparison of Tables 4.3 and 4.4 reveals that, in general, the confidence intervals derived from the asymptotic methods summarized in (4.11) and (4.30) are similar in size yet slightly positively shifted compared to those using the percentile method with the bootstrap simulation results. In practice, the confidence intervals from the percentile method with the bootstrap simulation results are likely more reliable. The asymptotic theory method inherently makes a number of assumptions and approximations mainly with the estimation of the covariance matrix and the assumption regarding the distribution choice of (4.11), while the percentile method bypasses these assumptions.

Table 4.3: 95% confidence intervals derived from the wild bootstrap method using the percentile method for each case.

	2008		2009	
Parameter	Haskell	N01	Haskell	N01
α_0	-1.6573×10^{0} -1.4441×10^{0}	-1.5187×10^{0} -1.3062×10^{0}	-3.7670×10^{-1} 4.7696×10^{-2}	-3.4876×10^{-1} 7.6688×10^{-2}
α_1	8.3599×10^{-2} 9.2630×10^{-2}	7.7591×10^{-2} 8.6427×10^{-2}	2.5110×10^{-2} 4.4253×10^{-2}	2.5257×10^{-2} 4.3850×10^{-2}
α_2	-1.0024×10^{-3} -8.8282×10^{-4}	-9.2411×10^{-4} -8.0216×10^{-4}	-4.3117×10^{-4} -1.6569×10^{-4}	-4.3493×10^{-4} -1.7769×10^{-4}
α_3	2.6616×10^{-6} 3.1532×10^{-6}	2.3810×10^{-6} 2.8748×10^{-6}	1.1779×10^{-8} 1.1333×10^{-6}	6.8845×10^{-8} 1.1425×10^{-6}
β_1	9.9378×10^{2} 1.0444×10^{3}	8.8432×10^{2} 9.2835×10^{2}	8.5196×10^{2} 9.2745×10^{2}	8.1909×10^{2} 8.9339×10^{2}
β_2	-1.1524×10^{2} -2.9177×10^{1}	-7.3537×10^{1} -1.3917×10^{1}	-2.0646×10^{1} -9.0897×10^{0}	-1.6948×10^{1} -7.1306×10^{0}
β_3	-3.9562×10^{1} -2.8460×10^{1}	-4.0057×10^{1} -2.6202×10^{1}	-2.0620×10^{1} -1.4632×10^{1}	-1.9915×10^{1} -1.3886×10^{1}

Table 4.4: 95% confidence intervals derived from (4.11) and (4.30) applying wild bootstrap simulations to estimate covariance matrices for each case.

	2008		2009	
Parameter	Haskell	N01	Haskell	N01
α_0	-1.6662×10^{0} -1.4542×10^{0}	-1.5274×10^{0} -1.3114×10^{0}	-3.8401×10^{-1} 5.1325×10^{-2}	-3.3513×10^{-1} 8.4589×10^{-2}
α_1	8.3842×10^{-2} 9.2801×10^{-2}	7.7622×10^{-2} 8.6800×10^{-2}	2.5264×10^{-2} 4.4669×10^{-2}	2.4693×10^{-2} 4.3439×10^{-2}
α_2	-1.0039×10^{-3} -8.8341×10^{-4}	-9.2707×10^{-4} -8.0420×10^{-4}	-4.3536×10^{-4} -1.7001×10^{-4}	-4.2627×10^{-4} -1.6906×10^{-4}
α_3	2.6749×10^{-6} 3.1626×10^{-6}	2.3959×10^{-6} 2.8883×10^{-6}	3.4382×10^{-8} 1.1317×10^{-6}	4.4807×10^{-8} 1.1103×10^{-6}
β_1	9.9117×10^{2} 1.0434×10^{3}	8.8349×10^{2} 9.2895×10^{2}	8.4674×10^{2} 9.2203×10^{2}	8.1372×10^{2} 8.8739×10^{2}
β_2	-9.7630×10^{1} -1.1930×10^{1}	-6.4482×10^{1} 1.4992×10^{0}	-1.9181×10^{1} -7.3786×10^{0}	-1.5873×10^{1} -5.8689×10^{0}
β_3	-3.9099×10^{1} -2.8206×10^{1}	-3.9769×10^{1} -2.5819×10^{1}	-2.0314×10^{1} -1.4534×10^{1}	-1.9698×10^{1} -1.3760×10^{1}

4.4 Confidence Intervals for Model Predictions

Figures 4.2, 4.4, 4.6, and 4.8 illustrate the model predictions as compared to mean data as a function of date. The individual conductance data points exhibit significant scatter, as is typical for any biological data set. Therefore, it is useful for a model to not only provide point estimates, but also the associated confidence intervals for the model predictions.

Three different methodologies are explored to determine confidence intervals for the model predictions. We begin in Section 4.4.1 by applying traditional asymptotic theory to establish confidence intervals for the model predictions. In Section 4.4.2, confidence intervals are estimated using Monte Carlo simulations. Finally, Section 4.4.3 details the bootstrap method to ascertain confidence intervals for model predictions.

4.4.1 Asymptotic Theory

Consider the gradient vector

$$\nabla g_\ell(t, w_2; \theta) = \begin{bmatrix} \frac{\partial g_\ell(t,w_2;\theta)}{\partial \alpha_0} \\ \frac{\partial g_\ell(t,w_2;\theta)}{\partial \alpha_1} \\ \frac{\partial g_\ell(t,w_2;\theta)}{\partial \alpha_2} \\ \frac{\partial g_\ell(t,w_2;\theta)}{\partial \alpha_3} \\ \frac{\partial g_\ell(t,w_2,\theta)}{\partial \beta_1} \\ \frac{\partial g_\ell(t,w_2,\theta)}{\partial \beta_2} \\ \frac{\partial g_\ell(t,w_2,\theta)}{\partial \beta_3} \end{bmatrix} \tag{4.31}$$

whose definitions are explicitly given in (4.12) through (4.18). The (1 - α) confidence interval for the conductance prediction can then be defined as

$$g_\ell(t, w_2; \hat{\theta}) \pm t(n-p; \alpha/2) \sqrt{\nabla^T g_\ell(t, w_2; \hat{\theta})\ \mathrm{cov}(\hat{\theta})\ \nabla g_\ell(t, w_2; \hat{\theta})} \tag{4.32}$$

which is commonly known as the delta method [30, 57]. The matrix cov($\hat{\theta}$) is calculated as indicated in (4.28). For the Haskell genotype in 2008,

$$\mathrm{cov}(\hat{\theta}) = \begin{bmatrix} 3 \cdot 10^{-3} & -1 \cdot 10^{-4} & 2 \cdot 10^{-6} & -6 \cdot 10^{-9} & -4 \cdot 10^{-1} & -7 \cdot 10^{-2} & -9 \cdot 10^{-3} \\ -1 \cdot 10^{-4} & 5 \cdot 10^{-6} & -7 \cdot 10^{-8} & 3 \cdot 10^{-10} & 1 \cdot 10^{-2} & 2 \cdot 10^{-3} & 1 \cdot 10^{-4} \\ 2 \cdot 10^{-6} & -7 \cdot 10^{-8} & 9 \cdot 10^{-10} & -4 \cdot 10^{-12} & -2 \cdot 10^{-4} & -6 \cdot 10^{-6} & 6 \cdot 10^{-7} \\ -6 \cdot 10^{-9} & 3 \cdot 10^{-10} & -4 \cdot 10^{-12} & 2 \cdot 10^{-14} & 5 \cdot 10^{-7} & -9 \cdot 10^{-9} & -8 \cdot 10^{-9} \\ -4 \cdot 10^{-1} & 1 \cdot 10^{-2} & -2 \cdot 10^{-4} & 5 \cdot 10^{-7} & 2 \cdot 10^{2} & 1 \cdot 10^{2} & 2 \cdot 10^{1} \\ -7 \cdot 10^{-2} & 2 \cdot 10^{-3} & -6 \cdot 10^{-6} & -9 \cdot 10^{-9} & 1 \cdot 10^{2} & 5 \cdot 10^{2} & 6 \cdot 10^{1} \\ -9 \cdot 10^{-3} & 1 \cdot 10^{-4} & 6 \cdot 10^{-7} & -8 \cdot 10^{-9} & 2 \cdot 10^{1} & 6 \cdot 10^{1} & 8 \cdot 10^{0} \end{bmatrix},$$

for the N01 genotype in 2008,

$$\text{cov}(\hat{\theta}) = \begin{bmatrix} 3 \cdot 10^{-3} & -1 \cdot 10^{-4} & 2 \cdot 10^{-6} & -6 \cdot 10^{-9} & -4 \cdot 10^{-1} & -8 \cdot 10^{-3} & -7 \cdot 10^{-3} \\ -1 \cdot 10^{-4} & 5 \cdot 10^{-6} & -7 \cdot 10^{-8} & 3 \cdot 10^{-10} & 1 \cdot 10^{-2} & -1 \cdot 10^{-3} & -8 \cdot 10^{-5} \\ 2 \cdot 10^{-6} & -7 \cdot 10^{-8} & 1 \cdot 10^{-9} & -4 \cdot 10^{-12} & -2 \cdot 10^{-4} & 3 \cdot 10^{-5} & 5 \cdot 10^{-6} \\ -6 \cdot 10^{-9} & 3 \cdot 10^{-10} & -4 \cdot 10^{-12} & 2 \cdot 10^{-14} & 7 \cdot 10^{-7} & -2 \cdot 10^{-7} & -3 \cdot 10^{-8} \\ -4 \cdot 10^{-1} & 1 \cdot 10^{-2} & -2 \cdot 10^{-4} & 7 \cdot 10^{-7} & 1 \cdot 10^{2} & 7 \cdot 10^{1} & 2 \cdot 10^{1} \\ -8 \cdot 10^{-3} & -1 \cdot 10^{-3} & 3 \cdot 10^{-5} & -2 \cdot 10^{-7} & 7 \cdot 10^{1} & 3 \cdot 10^{2} & 6 \cdot 10^{1} \\ -7 \cdot 10^{-3} & -8 \cdot 10^{-5} & 5 \cdot 10^{-6} & -3 \cdot 10^{-8} & 2 \cdot 10^{1} & 6 \cdot 10^{1} & 1 \cdot 10^{1} \end{bmatrix},$$

for the Haskell genotype in 2009,

$$\text{cov}(\hat{\theta}) = \begin{bmatrix} 1 \cdot 10^{-2} & -5 \cdot 10^{-4} & 7 \cdot 10^{-6} & -3 \cdot 10^{-8} & -8 \cdot 10^{-1} & -4 \cdot 10^{-2} & -3 \cdot 10^{-2} \\ -5 \cdot 10^{-4} & 2 \cdot 10^{-5} & -3 \cdot 10^{-7} & 1 \cdot 10^{-9} & 4 \cdot 10^{-2} & 3 \cdot 10^{-3} & 2 \cdot 10^{-3} \\ 7 \cdot 10^{-6} & -3 \cdot 10^{-7} & 5 \cdot 10^{-9} & -2 \cdot 10^{-11} & -5 \cdot 10^{-4} & -4 \cdot 10^{-5} & -3 \cdot 10^{-5} \\ -3 \cdot 10^{-8} & 1 \cdot 10^{-9} & -2 \cdot 10^{-11} & 8 \cdot 10^{-14} & 2 \cdot 10^{-6} & 2 \cdot 10^{-7} & 1 \cdot 10^{-7} \\ -8 \cdot 10^{-1} & 4 \cdot 10^{-2} & -5 \cdot 10^{-4} & 2 \cdot 10^{-6} & 4 \cdot 10^{2} & 4 \cdot 10^{1} & 2 \cdot 10^{1} \\ -4 \cdot 10^{-2} & 3 \cdot 10^{-3} & -4 \cdot 10^{-5} & 2 \cdot 10^{-7} & 4 \cdot 10^{1} & 9 \cdot 10^{0} & 4 \cdot 10^{0} \\ -3 \cdot 10^{-2} & 2 \cdot 10^{-3} & -3 \cdot 10^{-5} & 1 \cdot 10^{-7} & 2 \cdot 10^{1} & 4 \cdot 10^{0} & 2 \cdot 10^{0} \end{bmatrix},$$

and for the N01 genotype in 2009,

$$\text{cov}(\hat{\theta}) = \begin{bmatrix} 1 \cdot 10^{-2} & -5 \cdot 10^{-4} & 7 \cdot 10^{-6} & -3 \cdot 10^{-8} & -4 \cdot 10^{-1} & 9 \cdot 10^{-3} & 9 \cdot 10^{-4} \\ -5 \cdot 10^{-4} & 2 \cdot 10^{-5} & -3 \cdot 10^{-7} & 1 \cdot 10^{-9} & 2 \cdot 10^{-2} & -8 \cdot 10^{-5} & 2 \cdot 10^{-4} \\ 7 \cdot 10^{-6} & -3 \cdot 10^{-7} & 4 \cdot 10^{-9} & -2 \cdot 10^{-11} & -3 \cdot 10^{-4} & -2 \cdot 10^{-6} & -4 \cdot 10^{-6} \\ -3 \cdot 10^{-8} & 1 \cdot 10^{-9} & -2 \cdot 10^{-11} & 7 \cdot 10^{-14} & 1 \cdot 10^{-6} & 1 \cdot 10^{-8} & 2 \cdot 10^{-8} \\ -4 \cdot 10^{-1} & 2 \cdot 10^{-2} & -3 \cdot 10^{-4} & 1 \cdot 10^{-6} & 4 \cdot 10^{2} & 4 \cdot 10^{1} & 2 \cdot 10^{1} \\ 9 \cdot 10^{-3} & -8 \cdot 10^{-5} & -2 \cdot 10^{-6} & 1 \cdot 10^{-8} & 4 \cdot 10^{1} & 6 \cdot 10^{0} & 4 \cdot 10^{0} \\ 9 \cdot 10^{-4} & 2 \cdot 10^{-4} & -4 \cdot 10^{-6} & 2 \cdot 10^{-8} & 2 \cdot 10^{1} & 4 \cdot 10^{0} & 2 \cdot 10^{0} \end{bmatrix}.$$

Table 4.5 summarizes the percentages of data points covered by the estimated 95% confidence regions computed with (4.32) for each respective case. Figures 4.16-4.19 graphically depict each of the four studies for the Haskell genotype in 2008, the N01 genotype in 2008, the Haskell genotype in 2009, and the N01 genotype in 2009, respectively. Although the percentages of coverage indicated in the table seem very low, because of the scatter of the data as seen in Figures 3.23-3.34, the size of the confidence regions as illustrated in these figures is reasonable.

Table 4.5: Percentage of data points covered by estimated 95% confidence regions computed using (4.32).

Study year	Genotype	Dry	Medium	Wet	Open
2008	Haskell	13.85	13.18	11.15	9.12
2008	N01	10.81	13.85	14.19	13.18
2009	Haskell	13.53	13.33	10.88	14.11
2009	N01	11.89	12.50	14.58	15.57

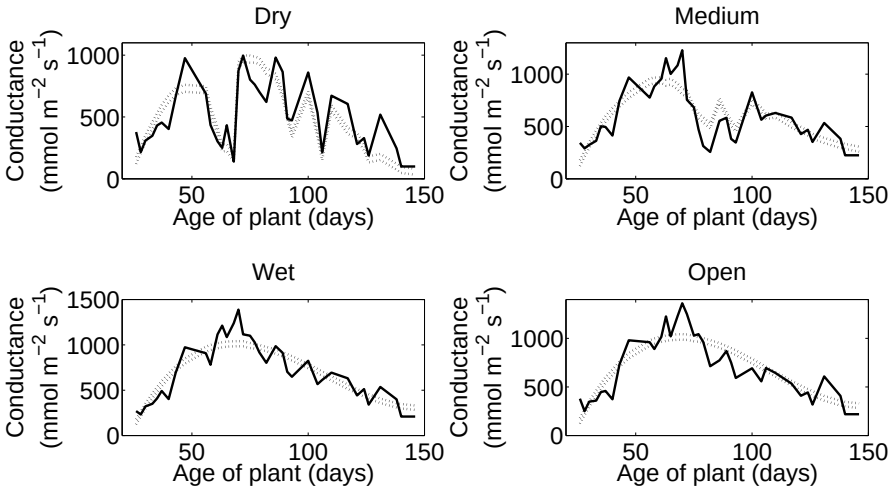


Figure 4.16: 95% confidence regions computed with (4.32) (dotted lines) compared to mean data by date (solid lines) for the Haskell genotype in 2008.

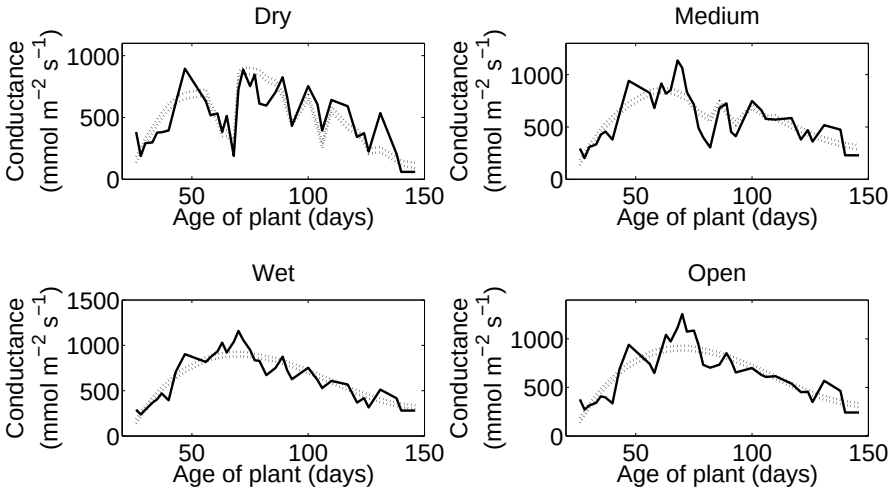


Figure 4.17: 95% confidence regions computed with (4.32) (dotted lines) compared to mean data by date (solid lines) for the N01 genotype in 2008.

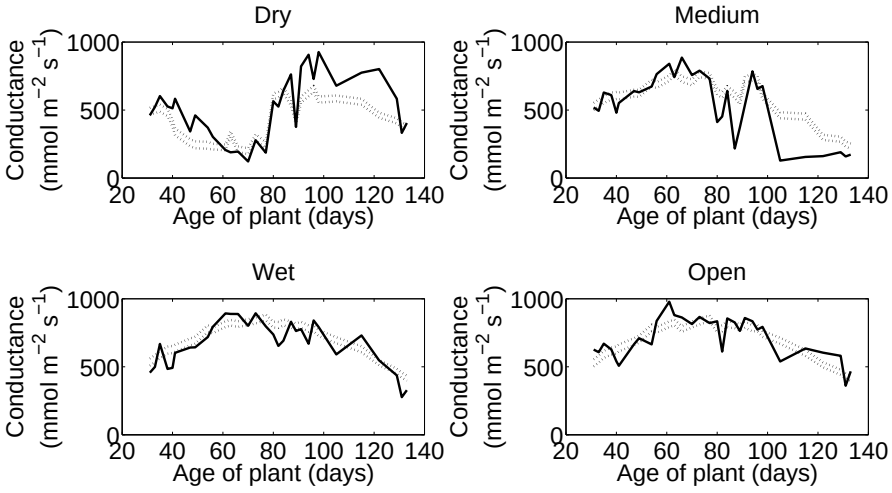


Figure 4.18: 95% confidence regions computed with (4.32) (dotted lines) compared to mean data by date (solid lines) for the Haskell genotype in 2009.

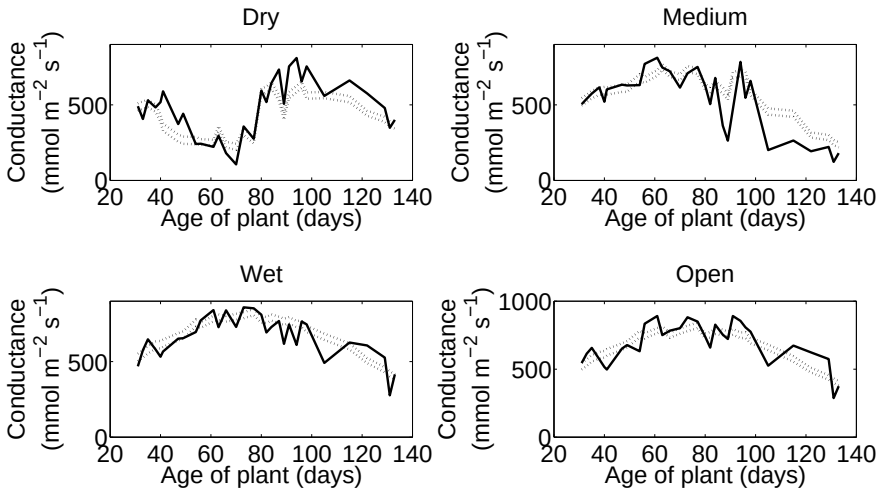


Figure 4.19: 95% confidence regions computed with (4.32) (dotted lines) compared to mean data by date (solid lines) for the N01 genotype in 2009.

4.4.2 Monte Carlo Method

A different method to estimate confidence intervals for model predictions involves Monte Carlo simulations. This method utilizes the estimated parameter vectors summarized in Table 4.2. Under the assumption that these estimated parameters originate from a normal distribution where the mean is the estimated parameter vector $\hat{\theta}$ and the covariance amongst parameters is quantified by the estimated covariance matrices $\text{cov}(\hat{\theta})$, 1000 parameter vectors $\hat{\theta}_m$ for $m = 1, \ldots, 1000$ are randomly chosen. Then, the associated model predictions $g_\ell(t, w_2; \hat{\theta}_m)$ are calculated. Hence, for each e_k, $k = 1, \ldots, n$, there are 1000 model predictions $g_\ell(e_k; \hat{\theta}_m)$. These 1000 predictions are sorted in ascending order and, using the percentile method to establish the 95% confidence region, the 25th and 975th predictions are selected as the lower and upper bounds, respectively.

Occasionally the randomly selected vectors were infeasible. In particular, we define $\beta_2 < 0$ and $\beta_3 < 0$. If by chance the random selection procedure produced a parameter vector which did not meet these criterion, a new vector was randomly selected until the criterion were met. Also, occasionally a randomly selected parameter vector would result in a negative conductance prediction g_ℓ, which is also infeasible. Upon this event, a new vector was randomly selected until a positive prediction g_ℓ was achieved.

Table 4.6 summarizes the percentages of data points covered by the estimated 95% confidence regions computed by this Monte Carlo method for each respective case. Figures 4.20 through 4.23 are analagous to Figures 4.16 - 4.19 of the previous section. Notice the 95% confidence regions predicted via this Monte Carlo method are larger than those derived via asymptotic theory, as indicated by the larger percentage of data points covered by the estimated confidence region. For example, the confidence region for the "dry" experiment for the Haskell genotype in 2006, as shown in Figure 4.20, indicates that at soil moisture of 0.12 m^3/m^3 and plant age of 61 days, the confidence region is approximately 330 mmol m^{-2} s^{-1} to 660 mmol m^{-2} s^{-1}. For a scientist, this information is fairly useless because of the large range. The failure of this method is likely due to the assumption of the normal distribution of the estimated parameters which the histograms in Figures 4.12 through 4.15 illustrate not to be true in many cases.

Remark. The reader should note that $\hat{\theta}_m$ defined for this Monte Carlo method is different from the $\hat{\theta}^m$ defined in (4.23) for the bootstrap method.

Table 4.6: Percentage of data points covered by estimated 95% confidence regions computed with Monte Carlo methods.

Study year	Genotype	Dry	Medium	Wet	Open
2008	Haskell	27.70	13.51	9.12	7.43
2008	N01	26.69	15.54	12.16	8.78
2009	Haskell	17.62	10.42	8.47	12.86
2009	N01	16.80	13.75	13.36	13.12

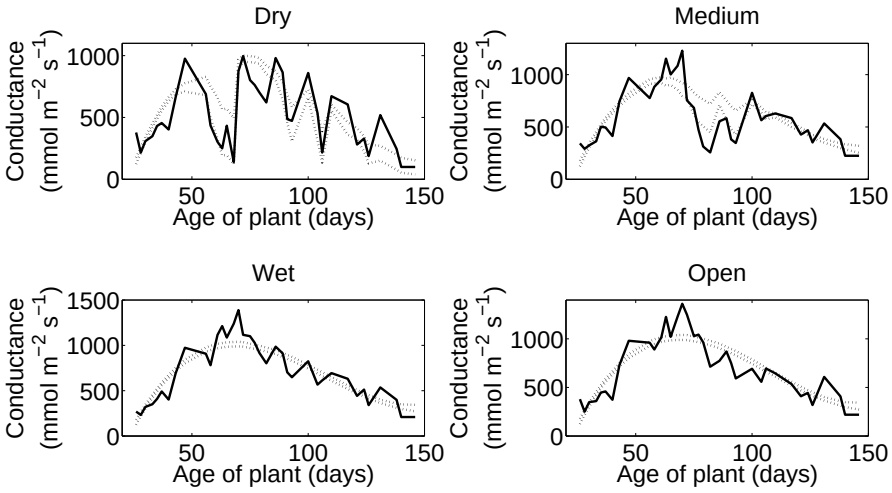


Figure 4.20: 95% confidence regions computed with Monte Carlo methods (dotted lines) compared to mean data by date (solid lines) for the Haskell genotype in 2008.

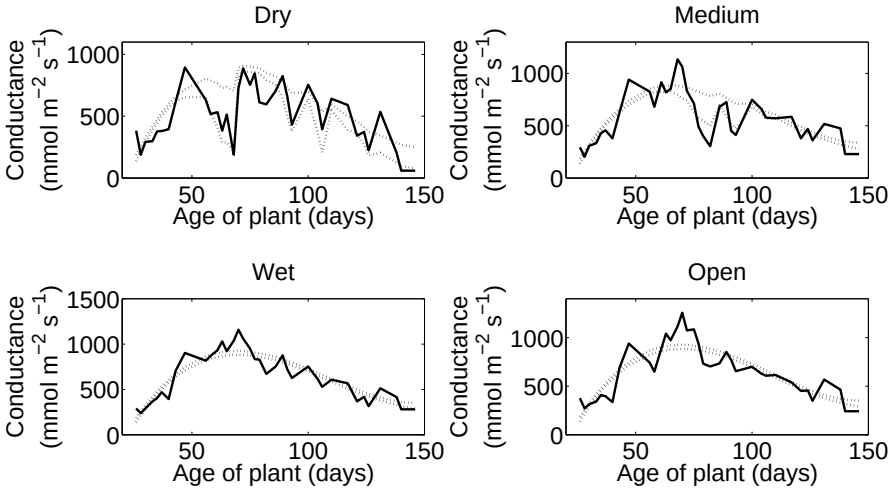


Figure 4.21: 95% confidence regions computed with Monte Carlo methods (dotted lines) compared to mean data by date (solid lines) for the N01 genotype in 2008.

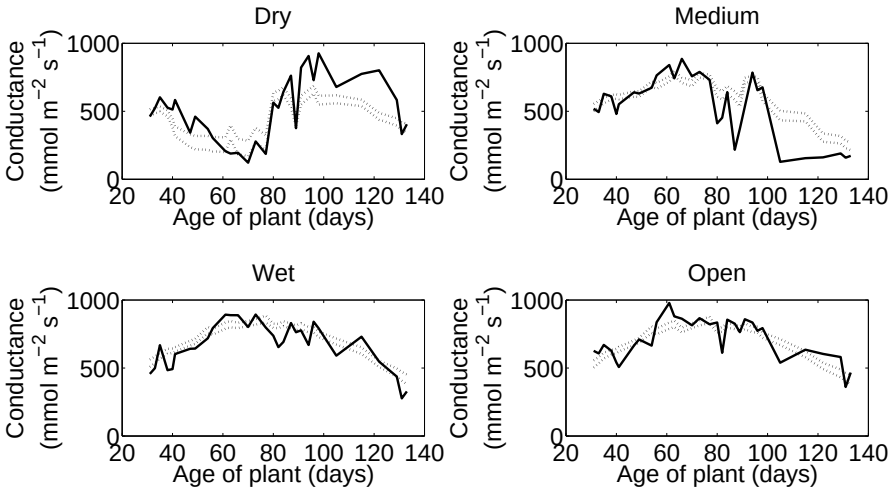


Figure 4.22: 95% confidence regions computed with Monte Carlo methods (dotted lines) compared to mean data by date (solid lines) for the Haskell genotype in 2009.

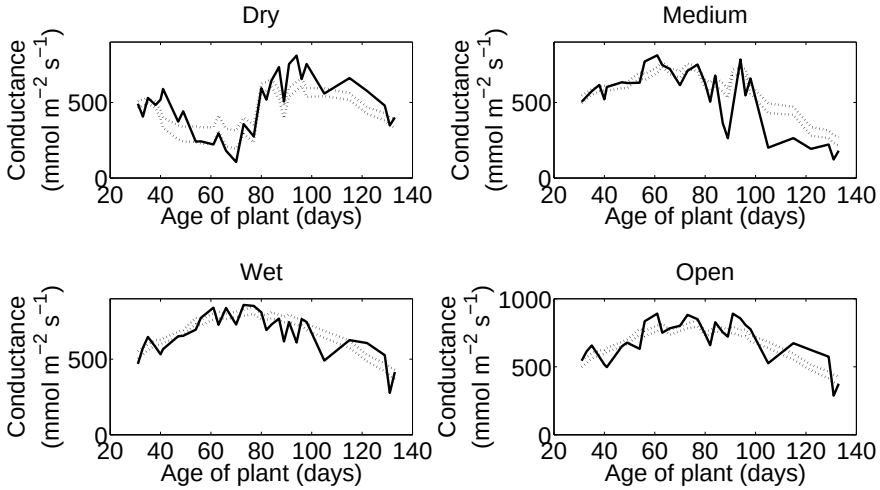


Figure 4.23: 95% confidence regions computed with Monte Carlo methods (dotted lines) compared to mean data by date (solid lines) for the N01 genotype in 2009.

4.4.3 Bootstrap Method

As described in (4.23), applying the wild bootstrap method yields the estimated parameter vectors $\hat{\theta}^m$ for $m = 1, \ldots, 1000$. Using these estimated parameter vectors, the associated model predictions $g_\ell(t, w_2; \hat{\theta}^m)$ can be calculated. Hence, for each e_k, $k = 1, \ldots, n$, there are 1000 model predictions $g_\ell(e_k; \hat{\theta}^m)$. These 1000 predictions are sorted in ascending order, and using the percentile method to establish the 95% confidence region for model predictions, the 25th and 975th predictions are selected as the lower and upper bounds, respectively.

Table 4.7 summarizes the percentages of data points covered by the estimated 95% confidence regions computed by the percentile method with bootstrap simulations. Figures 4.24 through 4.27 are analagous to Figures 4.16-4.19 and Figures 4.20-4.23 for the asymptotic and Monte Carlo methods, respectively. Although the percentages of coverage indicated in the table seem very low, because of the scatter of the data as seen in Figures 3.23 - 3.34, the size of the confidence regions as illustrated in Figures 4.24 through 4.27 is reasonable.

Although the range of these confidence regions is similar to those derived by asymptotic theory, recall that the delta method relies on the assumption for the distribution of model predictions; see (4.32). The confidence regions determined by the Monte Carlo method have the tendency to be overestimated due to violation of the normality assumption for estimated parameter distributions. Therefore, it is our opinion that of the three methods presented for estimating 95% confidence regions for model predictions, the bootstrap method is the most reliable. Coverage of data may be improved with continued model improvement as detailed in Chapter 5.

Table 4.7: Percentage of data points covered by estimated 95% confidence regions computed with percentile methods and bootstrap simulations.

Study year	Genotype	Dry	Medium	Wet	Open
2008	Haskell	13.51	9.80	9.46	7.43
2008	N01	12.16	11.15	11.49	8.78
2009	Haskell	13.93	9.58	7.66	11.62
2009	N01	14.34	11.67	11.34	11.48

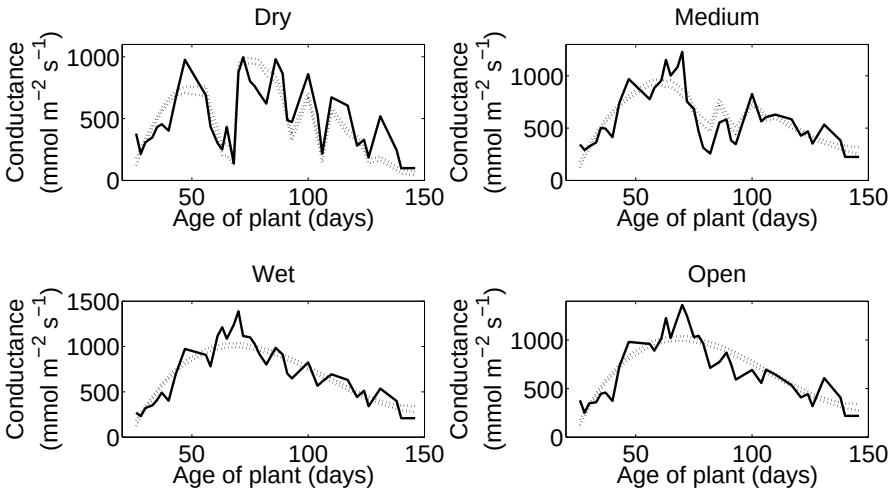


Figure 4.24: 95% confidence regions computed with percentile methods and bootstrap simulations (dotted lines) compared to mean data by date (solid lines) for the Haskell genotype in 2008.

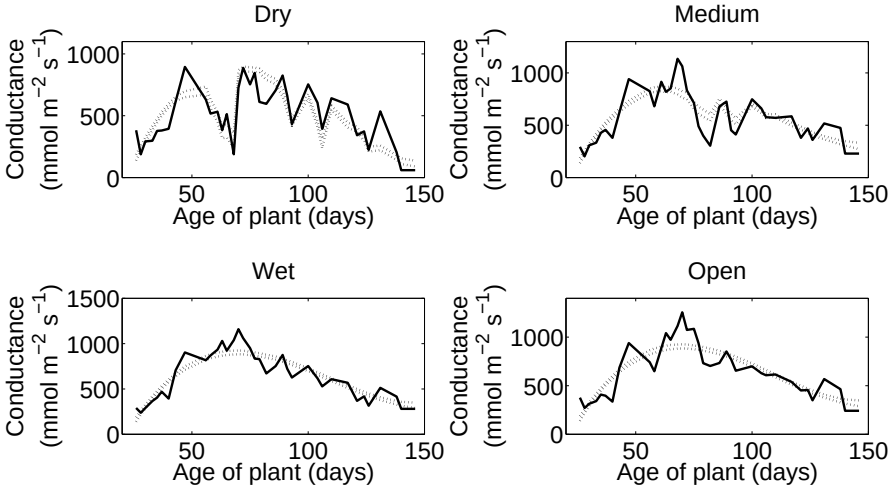


Figure 4.25: 95% confidence regions computed with percentile methods and bootstrap simulations (dotted lines) compared to mean data by date (solid lines) for the N01 genotype in 2008.

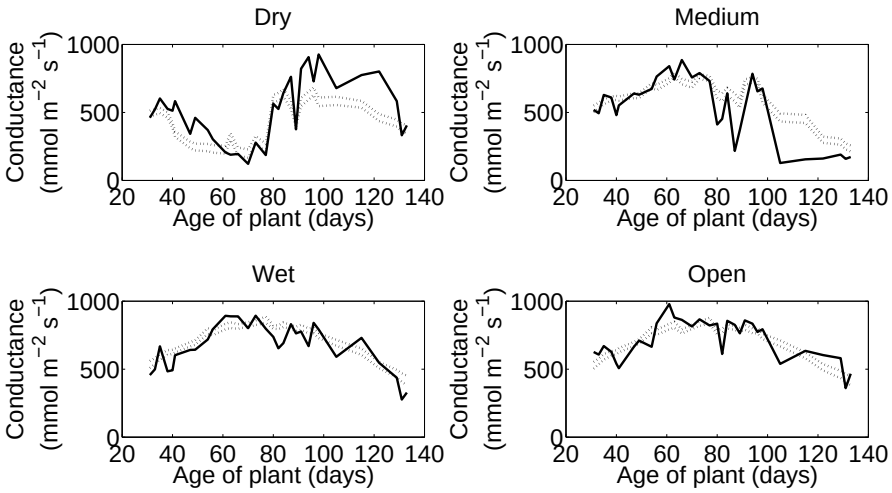


Figure 4.26: 95% confidence regions computed with percentile methods and bootstrap simulations (dotted lines) compared to mean data by date (solid lines) for the Haskell genotype in 2009.

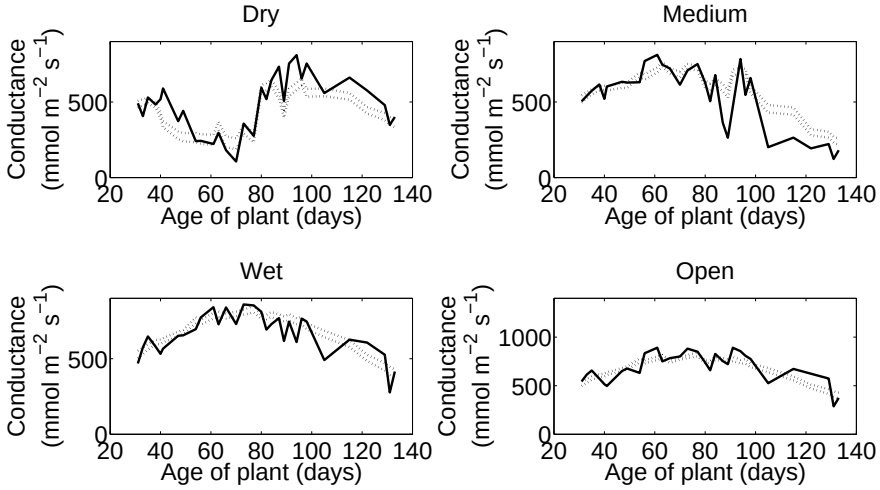


Figure 4.27: 95% confidence regions computed with percentile methods and bootstrap simulations (dotted lines) compared to mean data by date (solid lines) for the N01 genotype in 2009.

4.5 Comparison of Proposed Model to Niyogi's Model

As summarized in Chapter 3, the model presented by Niyogi et al. [48] did not predict the total leaf conductance well. In this chapter, a new model is introduced that characterizes the conductance response to plant development and soil moisture conditions. The differences between these two models are vast. Niyogi's model is a complicated system of equations, each describing a biological process involved in the coupled photosynthesis-conductance relationship. There are a number of model parameters and inputs required for simulation. The model presented here is very simple and includes only one equation incorporating soil moisture and plant age inputs. The comparison of the formulations of these two models exhibits a common dilemma faced by modelers. Many times the modeler must choose between a more complex highly-parameterized model depicting processes in a physical way as a opposed to a simple phenomenological formulation which is easy for scientists to implement and apply. By introducing the simple model we abide by Ockham's razor, the principle that entities must not be multiplied beyond necessity.

Visual comparison of Figures 3.35 - 3.38 depicting accuracy of Niyogi's model to Figures 4.1 - 4.8 implies that the current model is likely more representative of these data sets. For statistical rigor, Akaike's information criterion (AIC) is used to select the best model. The AIC for each model and case is calculated using the relation

$$AIC = 2 \cdot n_{par} + n_{data} \left(log_e \left(\frac{sse}{n_{data}} \right) \right) \tag{4.33}$$

where n_{par} is the number of estimated parameters, n_{data} is the number of data points used to calculate the SSE value, and SSE is calculated according to (4.5) [1]. These results are summarized in Table 4.8 where for every case the AIC test chooses the simpler model introduced in (4.1) over Niyogi's model.

Table 4.8: AIC model selection test results.

		2008		2009	
Model	Value	Haskell	N01	Haskell	N01
Niyogi	sse	5.3540×10^8	4.5684×10^8	3.6155×10^8	3.3720×10^8
	n_{data}	1116	1116	935	937
	n_{par}	0	0	0	0
	AIC	1.4598×10^4	1.4421×10^4	1.2029×10^4	1.1988×10^4
Matthews	sse	3.9567×10^7	3.0781×10^7	2.2497×10^7	1.9445×10^7
	n_{data}	1184	1184	973	975
	n_{par}	7	7	7	7
	AIC	1.2348×10^4	1.2050×10^4	9.7912×10^3	9.6671×10^3
Model choice		Matthews	Matthews	Matthews	Matthews

Chapter 5

Conclusions and Future Work

Potential impacts of climate change include periods of drought as well as the greater possibility of extreme precipitation events [44]. Because of the importance of correctly characterizing photosynthetic processes in the context of climate modeling, it is imperative that the response of stomatal conductance to drought and extreme precipitation events be well understood and applicably modeled.

In Chapter 2 an analysis of existing work was performed. Literature review and sensitivity analysis of a particular recent model [48] was presented. The chapter explicitly presents the parameters, equations, and solution algorithm for the model as applied to both C3 and C4 plants. Sensitivity analysis reveals that in C3 plants, the top five most sensitive parameters are: T_s, b, S_2, $f(w_2)$, and h_s. The top five most sensitive parameters for C4 plants are: S_2, $f(w_2)$, T_s, S_4, and β_2. This analysis raises concerns about the practicality of application of this model due to the inability to arrive at feasible solutions under a variety of typical environmental conditions.

Chapter 3 describes a novel data set containing conductance measurements of two genotypes of soybean under different water stress conditions. Environmental data collected simultaneously with conductance allowed testing of the model analyzed in Chapter 2 under varying soil moisture conditions. In general this model was found to underpredict conductance.

Under the hypothesis that the inadequacies of the model predictions for this data set are rooted in the absence or mischaracterization of two critical phenomenon (plant development and soil moisture conditions), a new model is presented. Model calibration and subsequent performance is summarized in Chapter 4. Statistical testing confirms that this model is superior in predicting conductance values under varying water stress conditions to that presented in [48]. Multiple methods, including the wild bootstrap, are explored to generate confidence intervals for the new model's parameters and predictions.

This work can be extended in several ways: model improvement, additional experimentation, and further exploration of statistical methods for confidence interval estimation.

Although the model presented herein describes the data set fairly well, improvements to a model are always possible. One suggestion to be explored in future work is the inclusion of an additional dependency on ambient temperature. Including the dependency on temperature may alleviate the difference between parameter values amongst growing seasons. As shown in Figures 4.9 and 4.10, the plant age and soil moisture dependencies were similar by growing season, when it would have been expected to see stronger similarities by genotype.

A suggestion for future experimentation is to collect coupled photosynthetic rate data along with conductance. The interdependency of these two processes is well documented, and it would be interesting to better understand their relationship under varying water stress conditions. This extended data set would allow for more complete testing of the model by Niyogi et al., as this model was only tested for conductance prediction. Identifying the relationship between conductance and photosynthetic rate under different watering scenarios may suggest a soil moisture dependency in the Ball-Berry formulation which could improve the predictive power of the Niyogi et al. model. Additionally, eliciting the relationship of conductance to photosynthesis would enable extension of the model presented herein to also predict photosynthetic rates.

As applied to the model presented in Chapter 4, bootstrap methods to estimate confidence intervals appear to be very promising. This field of research in the statistical community is still very active and novel methods are still being discovered. Further exploration of the many flavors of bootstrap methods available is warranted.

REFERENCES

[1] Akaike H. 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control* **19(6)**:716-723.

[2] Anderson MC, Norman JM, Meyers TP, Diak GR. 2000. An analytical model for estimating canopy transpiration and carbon assimilation fluxes based on canopy light-use efficiency. *Agricultural and Forest Meteorology* **101**:265-289.

[3] Baldocchi D. 1994. An analytical solution for coupled leaf photosynthesis and stomatal conductance models. *Tree Physiology* **14**:1069-1079.

[4] Ball J, Woodrow I, Berry J. 1987. A model predicting stomatal conductance and its contribution to the control of photosynthesis under different environmental conditions. In Progress in Photosynthesis Research, Vol IV, Martinus Nijhoff Pub., Dordrecht, 221-224.

[5] Ball JT. 1988. An analysis of stomatal conductance. PhD thesis, Stanford University, 89 pp.

[6] Banks HT, Holm K, Robbins D. 2009. Standard error computations for uncertainty quantification in inverse problems: asymptotic theory vs. bootstrapping. CRSC Technical Report TR09-13.

[7] Banks HT and Tran HT. 2009. Mathematical and experimental modeling of physical and biological processes. Boca Raton: CRC Press.

[8] Bates DM and Watts DG. 2007. Nonlinear regression analysis and its applications. New York: Wiley.

[9] Berry J and Bjorkman O. 1980. Photosynthetic response and adaptation to temperature in higher plants. *Annual Review of Plant Physiology* **31**:491-543.

[10] Bjorkman O and Berry J. 1973. High efficiency photosynthesis. *Scientific American* **229(4)**:80-93.

[11] Boerma HR, Hussey RS, Phillips DV, Wood ED, Finnerty SL. 1994. Registration of 'Haskell' soybean. *Crop Science* **34**:541.

[12] Boone A, Calvet J-C, Noilhan J. 1999. Inclusion of a third soil layer in a land surface scheme using the force-restore method. *Journal of Applied Meteorology* **38**:1611-1630.

[13] Boote KJ and Pickering NB. 1994. Modeling photosynthesis of row crop canopies. *HortScience* **29(12)**:1423-1434.

[14] Brooks A and Farquhar GD. 1985. Effect of temperature on the CO2/O2 specificity of ribulase-1,5-bisphosphate carboxylase/oxygenase and the rate of respiration in the light: Estimates from gas-exchange measurements on spinach. *Planta* **165**:397-406.

[15] Bunce JA. 1985. Effect of weather during leaf development on photosynthetic characteristics of soybean leaves. *Photosynthesis Research* **6**:215-220.

[16] Calvet J-C, Noilhan J, Roujean J-L, Bessemoulin P, Cabelguenne M, Olioso A, Wigneron J-P. 1998. An interactive vegetation SVAT model tested against data from six contrasting sites. *Agricultural and Forest Meteorology* **92**:73-95.

[17] Calvet J-C, Noilhan J, Bessemoulin. 1998. Retrieving the root-zone soil moisture from surface soil moisture or temperature estimates: a feasibility study based on field measurements. *Journal of Applied Meteorology* **37**:371-386.

[18] Carroll RJ and Ruppert D. 1988. Transformation and weighting in regression. New York: Chapman and Hall.

[19] Cintròn-Arias A, Banks HT, Capaldi A, Lloyd AL. 2009. A sensitivity matrix based methodology for inverse problem formulation. *Journal of Inverse and Ill-Posed Problems* **17**:545-564.

[20] Clapp RB and Hornberger GM. 1978. Empirical equations for some soil hydraulic properties. *Water Resources Research* **14**:601-604.

[21] Collatz GJ, Ball JT, Grivet C, Berry JA. 1991. Physiological and environmental regulation of stomatal conductance, photosynthesis and transpiration: a model that includes a laminar boundary layer. *Agricultural and Forest Meteorology* **54**:107-136.

[22] Collatz GJ, Ribas-Carbo M, Berry JA. 1992. Coupled photosynthesis-stomatal conductance model for leaves of C4 plants. *Australian Journal of Plant Physiology* **19**:519-538.

[23] Davidian M and Giltinan DM. 1995. Nonlinear models for repeated measurement data. New York: Chapman & Hall.

[24] Davidson R and Flachaire E. 2008. The wild bootstrap, tamed at last. *Journal of Econometrics* **146(1)**:162-169.

[25] Draper NR and Smith H. 1981. Applied regression analysis. New York: John Wiley & Sons.

[26] Dubois J-JB, Fiscus EL, Booker FL, Flowers MD, Reid CD. 2007. Optimizing the statistical estimation of the parameters of the Farquhar-van Caemmerer-Berry model of photosynthesis. *New Phytologist* **176**:402-414.

[27] Efron B and Tibshirani RJ. 1993. An introduction to the bootstrap. New York: Chapman & Hall.

[28] Ehleringer J and Bjorkman O. 1977. Quantum yields for CO_2 uptake in C_3 and C_4 plants. *Plant Physiology* **59**:86-90.

[29] Farquhar GD, von Caemmerer S, Berry JA. 1980. A biochemical model of photosynthetic CO2 assimilation in leaves of C3 species. *Planta* **149**:78-90.

[30] Fox J. 2008. Applied regression analysis and generalized linear models. Second edition. Los Angeles: Sage Publications, Inc.

[31] Gallant AR. 1987. Nonlinear statistical models. New York: John Wiley & Sons.

[32] Goudriaan J, van Laar H, van Keulen H, Lovwerse W. 1985. Photosynthesis, CO2 and plant production. In: Day W and Atkin RK (Eds), Wheat Growth and Modeling, NATO ASI Series, Series A, Vol. 86, Plenum Press, NY, 107-122.

[33] Jacobs C. 1994. Direct impact of atmospheric CO2 enrichment on regional transpiration. PhD thesis, Wageningen Agricultural University, The Netherlands, 179 pp.

[34] Jacquemin B and Noilhan J. 1990. Sensitivity study and validation of a land surface parameterization using the hapex-mobilhy data set. *Boundary-Layer Meteorology* **52**:93-134.

[35] Jordan DB and Ogren WL. 1984. The CO2/O2 specificity of ribulase 1,5-biphosphate carboxylase/oxygenase: Dependence on ribulosebisphosphate concentration, pH and temperature. *Planta* **161**:308-313.

[36] Kasperbauer MJ. 1987. Far-red light reflection from green leaves and effects on phytochrome-mediated assimilate partitioning under field conditions. *Plant Physiology* **85**:350-354.

[37] Kim J and Verma SB. 1991. Modeling canopy photosynthesis: scaling up from a leaf to canopy in a temperate grassland ecosystem. *Agricultural and Forest Meteorology* **57**:187-208.

[38] Leakey ADB, Bernacchi CJ, Ort DR, Long SL. 2006. Long-term growth of soybean at elevated [CO_2] does not cause acclimation of stomatal conductance under fully open-air conditions. *Plant, Cell and Environment* **29**:1794-1800.

[39] Leuning R. 1990. Modelling stomatal behaviour and photosynthesis of eucalyptus grandis. *Australian Journal of Plant Physiology* **17**:159-175.

[40] Liu RY. 1988. Bootstrap procedures under some non-i.i.d. models. *Annals of Statistics* **16(4)**:1696-1708.

[41] MacKinnon JG. 2006. Bootstrap methods in econometrics. *The Economic Record* **82**:S2-S18.

[42] Mammen E. 1993. Bootstrap and wild bootstrap for high dimensional linear models. *Annals of Statistics* **21(1)**:255-285.

[43] McCree KJ. 1966. A solarimeter for measuring photosynthetically active radiation. *Agricultural Meteorology* **3**:353-366.

[44] Meehl GA, Stocker TF, Collins WD, Friedlingstein P, Gaye AT, Gregory JM, Kitoh A, Knutti R, Murphy JM, Noda A, Raper SCB, Watterson IG, Weaver AJ, Zhao Z-C. 2007.

Global Climate Projections. In: *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* [Solomon S, Qin D, Manning M, Chen A, Marquis M, Averyt KB, Tignor M, Miller HL (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

[45] Milly PCD, Shmakin AB. 2002. Global modeling of land water and energy balances, Part I: The Land Dynamics (LaD) model. *Journal of Hydrometeorology* **3**:283-299.

[46] Monteith J and Unsworth M. 2008. Principles of environmental physics: 3rd edition. New York: Elsivier.

[47] Nikolov NT, Massman WJ, Schoettle AW. 1995. Coupling biochemical and biophysical processes at the leaf level: an equilibrium photosynthesis model for leaves of C3 plants. *Ecological Modelling* **80**:205-235.

[48] Niyogi D, Alapaty K, Raman S, Chen F. 2009. Development and evaluation of a coupled photosynthesis-based gas exchange evapotranspiration model (GEM) for mesoscale weather forecasting applications. *Journal of Applied Meteorology and Climatology* **48(2)**:349-368.

[49] Nobel PS. 2004 Physicochemical and environmental plant physiology: 3rd edition. Boston: Elsivier.

[50] Noilhan J and Planton S. 1989. A simple parameterization of land surface processes for meteorological models. *Monthly Weather Review* **117**:536-549.

[51] Noilhan J and Lacarrere P. 1995. GCM grid-scale evaporation from mesoscale modeling. *Journal of Climate* **8**:206-223.

[52] Norman JM and Polley W. 1989. Canopy photosynthesis. In Briggs WR (Ed.), *Photosynthesis* (pp. 227-241). New York: Liss.

[53] Oleson KW, et al. 2004. *Technical Description of the Community Land Model (CLM).* NCAR Technical Note NCAR/TN-461+STR, National Center for Atmospheric Research, Boulder, CO, 173 pp.

[54] Parisi AV, Wong JCF, Galea VJ. 1998. A study of the total ultraviolet exposure to all the leaves for small plant growth. *Journal of Photochemistry and Photobiology B:Biology* **45(1)**:36-42.

[55] Randall DA, Wood RA, Bony S, Colman R, Fichefet T, Fyfe J, Kattsow V, Pitman A, Shulka J, Srinivasan J, Stouffer RJ, Sumi A, Taylor KE. 2007. Climate Models and Their Evaluation. In: *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* [Solomon S, Qin D, Manning M, Chen A, Marquis M, Averyt KB, Tignor M, Miller HL (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

[56] Ratkowsky DA. 1983. Nonlinear regression modeling: A unified practical approach. New York: Marcel Dekker, Inc.

[57] Rawlings JO. 1988. Applied regression analysis: a research tool. Belmont, CA: Wadsworth & Brooks.

[58] Sadok W and Sinclair TR. 2009. Genetic variability of transpiration response to vapor pressure deficit among soybean cultivars. *Crop Science* **49**:955-960.

[59] Sellers PJ, Berry JA, Collatz GJ, Field CB, Hall FG. 1992. Canopy reflectance, photosynthesis, and transpiration. III. A reanalysis using improved leaf models and a new canopy integration scheme. *Remote Sensing of Environment* **42**:187-216.

[60] Sellers PJ, Randall DA, Collatz GJ, Berry JA, Field CB, Dazlich DA, Zhang C, Collelo GD, Bounoua L. 1996. A revised land surface parameterization (SiB2) for atmospheric GCMs. Part I: Model formulation. *Journal of Climate* **9**:676-705.

[61] Sellers PJ, Los SO, Tucker CJ, Justice CO, Dazlich DA, Collatz GJ, Randall DA. 1996b. A revised land surface parameterization (SiB2) for atmospheric GCMs. Part II: Generation of global fields of terrestrial biophysical parameters from satellite data. *Journal of Climate* **9**:706-737.

[62] Storn R and Price K. 1997. Differential evolution - a simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization* **11**:341-359.

[63] Su H-B, Paw U KT, Shaw RH. 1996. Development of a coupled leaf and canopy model for the simulation of plant-atmosphere interaction. *Journal of Applied Meteorology* **35**:733-748.

[64] van Heemst HDJ, 1986. Potential crop production. In van Keulan H, Wolf J (Eds.), *Modelling of agricultural production: weather, soil, and crops* (pp. 13-26). Wageningen, Netherlands: Pudoc.

[65] Vu JCV, Allen LH, Widodo W. 2006. Leaf photosynthesis and Rubisco activity and kinetics of soybean, peanut and rice grown under elevated atmospheric CO_2, supraoptimal air temperature and soil water deficit. *Current Topics in Plant Biology* **7**:27-41.

[66] Wilson KB, Baldocchi DD, Hanson PJ. 2000. Spatial and seasonal variability of photosynthetic parameters and their relationship to leaf nitrogen in a deciduous forest. *Tree Physiology* **20**:565-578.

[67] Woodrow IE and Berry JA. 1988. Enzymatic regulation of photosynthetic CO2 fixation in C3 plants. *Annual Review of Plant Physiology and Plant Molecular Biology* **39**:533-94.

[68] Wu CFJ. 1986. Jackknife, bootstrap and other resampling methods in regression analysis. *Annals of Statistics* **14(4)**:1261-1295.

[69] Yu G-R, Wang Q-F, Zhuang J. 2004. Modeling the water use efficiency of soybean and maize plants under environmental stresses: application of a synthetic model of photosynthesis-transpiration based on stomatal behavior. *Journal of Plant Physiology* **161**:303-318.