



# To what extent does phosphorus limit agricultural production at the global scale?: A process modelling based approach

Marko Kvakic

## ► To cite this version:

Marko Kvakic. To what extent does phosphorus limit agricultural production at the global scale?: A process modelling based approach. Environmental Engineering. Université de Bordeaux, 2019. English. NNT : 2019BORD0191 . tel-02365204

**HAL Id: tel-02365204**

**<https://theses.hal.science/tel-02365204>**

Submitted on 15 Nov 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



THÈSE PRÉSENTE POUR OBTENIR LE GRADE DE:  
DOCTEUR DE L'UNIVERSITÉ DE BORDEAUX

ÉCOLE DOCTORALE: Sciences et environnement  
SPÉCIALITÉ: Biogéochimie et écosystèmes  
Par Marko KVAKIĆ:

**Dans quelle mesure le phosphore limite-t-il la production agricole à l'échelle mondiale? Une approche de modélisation basée sur les processus.**

ou en Anglais:

**To what extent does phosphorus limit agricultural production at the global scale? A process modelling based approach.**

SOUS LA DIRECTION DE:

Bruno RINGEVAL, Chargé de Recherche, ISPA-INRA

Philippe CIAIS, Directeur de recherche, LSCE-CEA

Sylvain PELLERIN, Directeur de recherche, ISPA-INRA

Soutenue le 22/10/2019 devant la jury composé de:

Mr MAKOWSKI David	Directeur de Recherche	INRA (FR)	Rapporteur, Président du jury
Mme COMTE Myriam	Professeur	Sorbonne Université (FR)	Rapporteur
Mme BONDEAU Alberte	Chargé de recherche	CNRS (FR)	Examinatrice
Mr FRANKLIN Oskar	Research Scholar	IIASA (AT)	Examineur
Mr RINGEVAL Bruno	Chargé de recherche	INRA (FR)	Directeur
Mr CIAIS Philippe	Directeur de Recherche	CEA (FR)	Co-directeur
Mr PELLERIN Sylvain	Directeur de Recherche	INRA (FR)	Invité

*“Take things as they are. Punch when you have to punch.  
Kick when you have to kick”.*  
*Bruce Lee*



*“A lot of people give up just before they’re about to  
make it. You know you never know when that next  
obstacle is going to be the last one”.*  
*Chuck Norris*

## Resumé

Le rôle du phosphore (P) en agriculture est indéniable: le P est un nutriment essentiel dont tous les êtres vivants ont besoin pour fonctionner, et est donc nécessaire pour maintenir les rendements agricoles à l'échelle globale dans les années à venir. Une grande partie du P utilisé pour fertiliser les cultures se présente sous forme d'engrais chimique et provient de mines de roches phosphatées. Cette ressource finie est gérée de manière non-optimale: dans certains endroits du Monde, le P est utilisé de manière excessive et peut nuire à l'environnement, alors qu'à d'autres endroits, le P apporté est insuffisant et conduit à des baisses de rendement importantes. Cette hétérogénéité, combinée à des problématiques d'accès à la ressource, qui dépend également de facteurs économiques et politiques, conduit à de sérieuses questions sur les impacts potentiels du P sur la sécurité alimentaire mondiale.

Des études récentes se sont penchées sur les principaux facteurs limitant les rendements agricoles dans le Monde, mais présentent des difficultés à séparer la contribution de ces différents facteurs, et en particulier du P. Dans un premier temps, j'ai combiné des simulations de la distribution du P dans les sols agricoles et des simulations de croissance des céréales dans des conditions idéales (i.e. non limitantes en eau, azote, etc.), tout en prenant en compte, de manière fine, les mécanismes de transfert du P entre le sol et la plante. J'ai montré que le P pourrait contribuer de manière significative à une baisse de rendement par rapport au rendement potentiel de 22, 55 et 26 % en blé d'hiver, maïs et riz. Cette diminution n'est que partiellement impactée quand les apports actuels de P par fertilisants chimiques sont considérés et ceci s'explique principalement par l'historique du bilan en P des sols (qui a contribué à fortement augmenter les stocks de P des sols). Cependant, la non prise en compte de certains processus, à savoir ceux liés aux ajustements des plantes dans des conditions limitantes en P, ont pu fortement biaiser ces estimations.

Pour mieux représenter ces processus d'ajustements, j'ai ensuite développé un modèle d'allocation du carbone (C) et du P basé sur des principes d'optimisation d'utilisation des ressources au sein de la plante. Le modèle est capable de simuler la réponse de la plante à une limitation en P: augmentation du ratio racines / biomasse aérienne, diminution de la biomasse totale et de la concentration en P. Le modèle a été testé dans un gradient de disponibilité en P à différentes échelles (plante en hydroponie et au champ) et reproduit raisonnablement le comportement des plantes. Malgré des hypothèses simplistes qui ne permettent pas de capturer la nature exacte de l'allocation, le modèle présenté peut être introduit dans un modèle de végétation plus physique, permettant l'étude de la limitation en P de manière plus générique.

Le couplage du modèle d'allocation idéalisé à un modèle de végétation physique a été réalisé en utilisant ORCHIDEE, un modèle de végétation dynamique utilisé pour étudier les interactions végétation-climat. Les paramétrisations de processus fondamentaux au sein d'ORCHIDEE (assimilation, etc.) ont été utilisées pour piloter le modèle d'allocation en fonction de la disponibilité en C et en P, et les simulations ont été comparées à deux jeux d'observations sur maïs irrigué. Les résultats ont montré le potentiel de la combinaison de ces deux modèles pour simuler le fonctionnement des cultures dans différents environnements. Le modèle ainsi obtenu pourra être utilisé pour mieux quantifier, à l'échelle mondiale, la contribution du P à la baisse de rendement des cultures par rapport à leur potentiel.

*Mots-clés: global, rendement, phosphore, carbone, modèle de végétation, ORCHIDEE*

## Abstract

The global role of phosphorus (P) in agriculture is undeniable: P is an essential nutrient required by all living beings to function, and thus necessary for sustaining yields worldwide in the time to come. In global agriculture, most of the P used to grow crops comes in form of chemical fertilizer which is mined from existing soil deposits. This in itself would not be an issue, was it not for the way we globally (mis)manage this potentially finite resource. While some places use P to the point of harming the environment, others do not have enough to sustain their yields and feed themselves. Combined with uncertainties of equitable P supply in the future which depend on economical and political factors as well, serious questions arise on the potential impacts of P on global food security.

Recent studies have looked into the main drivers of yield worldwide, but have difficulties separating P' contribution, as they lack the information to do so due to their empirical nature. As an initial step, we combined simulated global information on agricultural soil P and cereal growth in ideal conditions, while accounting for mechanisms of soil-plant P transfer more faithfully. We have found that P could significantly contribute to existing global production gaps with an average yield gap of 22, 55 and 26 % in winter wheat, maize and rice; lowering only slightly with today's P fertilizer use. This is mainly to be due to the global P management history or the net soil P balance up to date. But the idealized nature of the employed models ignored other processes, namely plant adjustment in P limited environments, which have a significant potential to change our diagnostic estimates.

To better represent plant adjustment, we have then developed an carbon (C) & P allocation model based on optimal functioning principles. The idealized model is capable of simulating primary plant response to a P limited environment: root-shoot ratio change, biomass and P concentration decrease. It was compared to plant growth across a P availability gradient at different scales (hydroponic to field) and has been found to reasonably predict observed plant behaviour. In spite of its simplistic assumptions which do not capture the exact nature of P flow within a plant, the idealized model could be introduced into a more physical vegetation one to allow the study of P limitation in a generic growing environment.

The coupling of our idealized allocation model to a physical vegetation one was performed using ORCHIDEE, a dynamic vegetation model used to study global vegetation-climate interaction. Its parameterizations of fundamental plant processes were used to drive our model as function of C and P availability, and compared to two irrigated maize observation datasets. The results have shown the potential of their combination to simulate crops in different growing environments, which is to be used on a global scale and finally help us better understand contribution of P to crop productivity globally.

*Keywords: global, yield, phosphorus, carbon, vegetation model, ORCHIDEE*

# Acknowledgements

First of all I would like to thank my sister and my mother, Iva and Ljiljanka, for their hard work and unreserved support over all these years which made the person I am today. Not only have they acted as role models and pillars of my life, but have enabled the pursuit of my wishes and dreams such as one can rarely find; all of my achievements up to this point can rightly be attributed to them.

I would like to thank my principal supervisor, Bruno Ringeval, for all of the support and advice he has bestowed upon me during my work. Most of all, I appreciate the freedom he has given me to pursue my interests and thank him for the extraordinary patience he has shown in my attempts to navigate through this personal and professional trial. I would also like to thank my other two supervisors, Philippe Ciais and Sylvain Pellerin, who made this PhD a possibility. Especially for providing additional financial support initially, which has enabled me to extend my studies and bring this work to a more satisfying end. Finally, I thank all of my supervisors for showing their faith in my work, which I hope will find its rightful place.

A special thank you goes to my BFF, Fotis, without whom I can't imagine my life in Bordeaux in any other possible way. For all of the good times and the bad ones, his jokes and good nature is what made the most of my Bordeaux experience the way it is. Another special thank you goes to my cool and sexy room-mates, Pietro and Rémi, who share my every-day cleaning burden with heroic demeanour but, even more, who share all of their life's experiences and advice so I can better understand my own. A special thank you also goes to Morgane, who is showing me even better things to come and just for being awesome as she is already.

Last but not least, I would like to thank all of the special people who I've crossed during my time in Bordeaux. Hilary, for giving me one of the most important lessons in my life. Mariellen, for sharing her experiences which helped me understand myself even more. Tristan, for his philosophical musings (which I never really understood) and stories of debauchery (which I understood even less). Annette, for sharing good food and almost making me break my neck. Bofang, Mohammed, Fredw, Florent and Andreas, for their company which made every day at work much more bearable. Tania, for her awesome thanksgiving turkey. Nico, Laurent and Mark, for sharing their scientific wisdom and letting me dabble in entomology. Jean-Yves, for being an awesome office mate. Sylvie, for

all of the nice Monday cake. Romain, Pierre and Dorian for letting me be a rockstar for a small bit. And all the others who I have crossed paths with at one time or the other; thank you all for writing this chapter of my life.

# Abbreviations

## Elements:

Al	Aluminium
C	Carbon
Fe	Iron
H	Hydrogen
K	Potassium
N	Nitrogen
P	Phosphorus

## Compounds:

CaCo <sub>3</sub>	Calcium carbonate
CO <sub>2</sub>	Carbon dioxide
H <sub>2</sub> O	Water
PO <sub>4</sub> <sup>3-</sup>	Phosphate, orthophosphate

## Acronyms:

DGVM	Dynamic global vegetation model
GGCM	Global gridded crop model
LAI	Leaf area index
RSR	Root-shoot ratio

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Global agriculture and the role of P . . . . .	1
1.2	The P cycle: from soil to plant . . . . .	3
1.3	Effects of P limitation: from plant to field . . . . .	4
1.4	P limitation in croplands: from field to the globe . . . . .	7
1.5	Motivation, objectives, questions . . . . .	13
1.6	Outline . . . . .	14
<b>2</b>	<b>Quantifying the limitation to world cereal production due to soil phosphorus status</b>	<b>15</b>
2.1	Introduction . . . . .	17
2.2	Methods . . . . .	18
2.2.1	Overview . . . . .	18
2.2.2	Crop P demand . . . . .	19
2.2.3	Soil P supply . . . . .	22
2.2.4	P limitation and yield gap due to P limitation . . . . .	24
2.2.5	Analysis of uncertainty, commonality analysis and global averages . . . . .	25
2.3	Results . . . . .	26
2.3.1	Potential root P uptake . . . . .	26
2.3.2	Crop P demand . . . . .	28
2.3.3	P limitation and yield gap . . . . .	30
2.4	Discussion . . . . .	32
<b>3</b>	<b>Carbon and phosphorus allocation in annual plants: an optimal resource use approach</b>	<b>36</b>
3.1	Introduction . . . . .	38
3.2	Methods . . . . .	40
3.2.1	Modelling framework . . . . .	40
3.2.2	Calibration : observations and protocol . . . . .	43
3.2.3	Sensitivity analysis of the modeled response . . . . .	45
3.3	Results . . . . .	46

3.3.1	Modeled plant response . . . . .	46
3.3.2	Sobol sensitivity analysis . . . . .	47
3.3.3	Comparison with field data . . . . .	49
3.3.4	Comparison with hydroponic studies . . . . .	51
3.4	Discussion . . . . .	52
<b>4</b>	<b>Optimal allocation of carbon and phosphorus in irrigated maize using ORCHIDEE</b>	<b>57</b>
4.1	Introduction . . . . .	59
4.2	Methods . . . . .	60
4.2.1	ORCHIDEE description . . . . .	60
4.2.2	Optimal functioning implementation . . . . .	62
4.2.3	Calibration with observations and ORCHIDEE comparison . . . . .	66
4.3	Results . . . . .	67
4.3.1	Comparison with Nebraska . . . . .	67
4.3.2	Comparison with Tartas . . . . .	68
4.3.3	Comparison between Nebraska and Tartas . . . . .	69
4.3.4	Comparison with original ORCHIDEE . . . . .	69
4.4	Discussion . . . . .	70
<b>5</b>	<b>Discussion</b>	<b>75</b>
5.1	Results summary and their place in current research . . . . .	75
5.2	Plant adjustment effects on global P limitation . . . . .	76
5.3	Why optimal functioning? . . . . .	78
5.4	Complexity vs. transparency . . . . .	80
5.5	Global P limitation using simpler models . . . . .	82
5.6	Global productivity, but also feasibility and impact . . . . .	85
<b>A</b>	<b>Bibliography</b>	<b>87</b>
<b>B</b>	<b>Supplementary information</b>	<b>106</b>
B.1	Chapter 2 supplementary information . . . . .	107
B.1.1	Introduction . . . . .	107
B.1.2	QUEFTS method . . . . .	107
B.1.3	Potential root P uptake . . . . .	109
B.1.4	P in soil solution . . . . .	110
B.1.5	Supplementary Tables . . . . .	112
B.1.6	Supplementary Figures . . . . .	114
B.2	Chapter 3 supplementary information . . . . .	121
B.2.1	Supplementary Tables . . . . .	121
B.2.2	Supplementary Figures . . . . .	123
B.3	Chapter 4 supplementary information . . . . .	126

B.3.1	Supplementary Tables . . . . .	126
B.3.2	Supplementary Figures . . . . .	130
B.4	Chapter 5 supplementary information . . . . .	135
B.4.1	Supplementary Figures . . . . .	135

# Chapter 1

## Introduction

### 1.1 Global agriculture and the role of P

Since the advent of modern agriculture, increasing crop yields has been the main objective of agricultural effort worldwide. This comes as no surprise as the current world population stands around six to seven billion, and is projected to reach nine in the next 30 years (Godfray et al., 2010). The expected rise in the world's population implies an increase in global food demand, estimated to be about 75% to 100% higher than today (Royal Society of London, 2009). One of the main ways to achieve this goal apart from fundamental changes to our food consumption habits (Kearney John, 2010) and systems of production in place (Tilman, 1998) is to intensify our crop production even further (Foley et al., 2011).

The push for global intensification has its roots in the "Green Revolution", a period of about last 50 years when crop production tripled with only a 30% increase in cultivated land area (Pingali, 2012). The increase was made possible through the adoption of improved crop varieties that ensured higher and more stable yields (Khush, 2001). Of these, the three most important ones are maize, rice and wheat as they comprise about 40% of total food intake globally (FAOSTAT, 2019). With our increasing ability to grow crops, fertilizer use (among other resources like water [H<sub>2</sub>O] and pesticides) has naturally followed suit (Fig. 1.1).

Fertilizer P comes in organic or chemical form. The organic one stems from production and recycling of organic waste (animal or human), whereas chemical one is mined in form of rock containing high concentrations of the ion phosphate (PO<sub>4</sub><sup>3-</sup>). Organic P sources are starting to take hold in the global agricultural P supply with at least 6-8 Mt P year<sup>-1</sup> applied to fields globally (Smil, 2000; Liu et al., 2008). But due to the limited practicality of organic fertilizer P use, chemical P is still the mainstay for the majority of the world crop production (or around 10 Mt P year<sup>-1</sup>, Smil, 2000). The exploitation of earth's rock P reserves has begun to accelerate exponentially during the 19th century and has recently grown to be about 13-16 Mt P year<sup>-1</sup>, driven by the increase in global consumption that reached its peak at 16.5 Mt P year<sup>-1</sup> in 1988 (Smil, 2000). From it, as was put into context

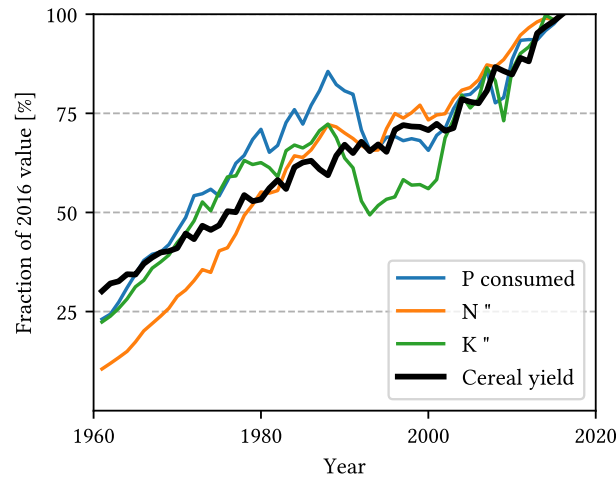


Figure 1.1: Total world cereal production and fertilizer consumption from 1960 to 2016 (FAOSTAT, 2019). Values are given relative to the ones at year 2016 when cereal production was 2.91 Gt grain, P consumption was 48.6 Mt P, N consumption was 110.2 Mt N and K consumption was 38.7 Mt K.

by Smil (2000), the Earth's agricultural soils received about 550 Mt P in between 1850 and 2000 which is equivalent to almost 10% of all world's arable soil's total P content.

The main issue with chemical P fertilizer is the finite nature of its availability in the long term. This is due to the depletion of existing, high quality reserves which makes further P exploration and mining economically unfeasible (Cordell, Drangert, and White, 2009). Conservative estimates by Cordell, Drangert, and White (2009) show that peak P production should happen around 2033, although this result seems to be contested recently (Fixen and Johnston, 2012; Kauwenbergh, M. Stewart, and Mikkelsen, 2013). One certain thing is that P addition in global croplands is highly unbalanced and, despite a positive global input balance, almost 30% of total arable land (MacDonald et al., 2011) experiences a P deficit (or more P removed through harvest than added with fertilizer). To give a typical example: European countries apply on average  $25 \text{ kgP ha}^{-1}$  to their fields whereas African ones apply only  $3 \text{ kgP ha}^{-1}$  (Liu et al., 2008); a difference of almost tenfold...

The global issue of P becomes even more pressing once social and environmental factors come into play (Obersteiner et al., 2013; J. Elser and E. Bennett, 2011). Despite the mentioned global P input surplus, only about 20% makes it to the final product of which (sadly enough) almost half is spoiled or wasted, with a big part running off into  $\text{H}_2\text{O}$  bodies and contributing to environmental pollution via eutrophication (J. Elser and E. Bennett, 2011). But the careless use of earth's P reserves will not be able to continue for long, as P reserves are concentrated in a handful of countries which (combined with others' uneven ability to procure it) makes international diplomatic tension due to P access a real possibility (Obersteiner et al., 2013). In spite of a grim forecast though, viable but habit changing solutions exist in form of technological improvements, reasonable P use, recycling and plant/animal genetics (Childers et al., 2011) in order to lessen our unreasonable

dependence on new P inputs.

## 1.2 The P cycle: from soil to plant

P has no natural gaseous forms or little atmospheric input (P deposition via aerosols; R. Wang et al. 2015) that would facilitate its transfer and uptake by an ecosystem unlike other major nutrients like C and nitrogen (N). The supply of P is thus strongly determined by existing soil reserves that are gradually released and taken up by the biome. From there, P cycles among various organic and inorganic pools of which the only one available for direct uptake is the  $\text{PO}_4^{3-}$  ion, determining P availability and ultimately the productivity of a given ecosystem. Since the process of P release happens on geological timescales, natural ecosystems have evolved to recycle existing P very efficiently (P. Vitousek, 1982)

The most well known framework for explaining P fate in natural soils is the model of Walker and Syers (1976). P is primarily derived from weathering of parent soil material of which the most common form is apatite, a calcium (Ca)  $\text{PO}_4^{3-}$  mineral that accounts for 95% of all P in the Earth's crust (Smil, 2000). Through the process of weathering,  $\text{PO}_4^{3-}$  is released and takes on forms of different availability that can be quantified using various soil extraction techniques (Bray and Kurtz, 1945; Nelson, 1953; Olsen, 1954; Hedley, J. W. B. Stewart, and Chauhan, 1982). Total soil P can be divided in the simplest manner into three fractions : non-occluded inorganic, occluded inorganic and organic P (Fig. 1.2). Non-occluded inorganic fraction stands for one that is sorbed to various soil components like clays, metal oxides (iron and aluminium [Fe and Al]) and calcium carbonate ( $\text{CaCO}_3$ ; Gérard 2016). This fraction can be exchanged with the soil solution at varying time-scales and finally taken up by the plant. Occluded inorganic fraction stands for plant inaccessible forms, which are due to physical occlusion within pores of Fe and Al oxides. Finally, organic fraction is the P bound in organic forms (phytates and nucleic acids) that can be mineralized via phosphatase excretion and subsequently recycled by plant roots or soil micro-organisms. Since all forms of P are subject to leaching that removes P from a given system and unless there is a significant input from atmospheric particle deposition (R. Wang et al., 2015), all natural systems should tend towards a 'terminal steady state' (Walker and Syers, 1976) where P becomes progressively unavailable to plants and the whole ecosystem P limited (P. M. Vitousek et al., 2010).

The main principle of soil-plant P exchange is the diffusion of  $\text{PO}_4^{3-}$  bound to different soil constituents due to its highly reactive nature (Morel et al., 2000). Concentration of P in soil solution depends on the sorption-desorption equilibrium, determined by physical and chemical soil properties like pH, temperature or electrolyte concentration (Barrow, 1983). As plants take up P, its concentration will decrease in the vicinity of the root and establish a gradient, in turn driving the diffusion of P towards the root surface (Barber, 1995). Since concentration of P in soil  $\text{H}_2\text{O}$  is too low to provide adequate quantities via advection (0.1

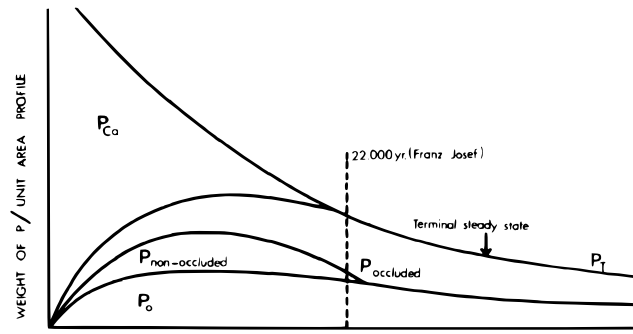


Figure 1.2: Graphical representation of the Walker and Syers (1976) model (reproduced from the original article).  $P_{Ca}$  stands for primary P source apatite.  $P_{occluded}$  and  $P_{non-occluded}$  stand for inorganic P forms that are respectively accessible or non-accessible to plants.  $P_T$  stands for total soil P. The vertical dashed line represents observed P fractions at one of the analyzed chronosequences in the original article.

-  $1 \text{ mgP L}^{-1}$ , Achat et al., 2016) diffusion is the key determinant of P available for plant uptake (Barber, 1995).

The mechanism of P uptake by roots is an important key for P transfer, and is most concisely described with an uptake kinetic specific to an individual species (Barber, 1995). These kinetics try to capture the fact that roots consist (roughly) of an outer permeable layer where absorption of  $\text{H}_2\text{O}$  and nutrients happens, and an inner part tasked with their transport into the rest of the plant, with the two separated by an ion barrier called the Casperian strip (Fig. 1.3). Even though root systems vary greatly among species and environments, a common point between them is root branching which serves to increase the root surface and subsequently the absorption of resources (Barber, 1995). Taking into account all of the physical aspects of root nutrient and  $\text{H}_2\text{O}$  transfer as well as the influence of root architecture has been subject of a long-standing inquiry. From analytical treatments of nutrient depletion in root vicinity (Barber, 1995; Tinker and Nye, 2000; Roose and Kirk, 2009) to architectural models of root growth and branching (Dupuy, Gregory, and Bengough, 2010; Pagès, 2019), as well as numerical solutions of their combined effect on root development and P distribution in the soil column (De Willigen and van Noordwijk, 1994; Comte, 2018) these studies demonstrate the underlying complexity of P transfer from soil to the plant. Most importantly, upscaling of these processes to the plant and agroecosystem scale is necessary if the complex nature of plant nutrition is to be studied (Hinsinger et al., 2011).

### 1.3 Effects of P limitation: from plant to field

P is an essential element to all living things and as such, its reduced availability can strongly impact growth and manifest limitation symptoms (Fig. 1.4). In crops these include stunted growth, reduced tillering, lower leaf and grain number, delay in reproductive growth and most notoriously the appearance of reddish or purple leaves (V. D. Fageria, 2001). This is because P is a fundamental building block of nucleic acids, membrane lipids, energy

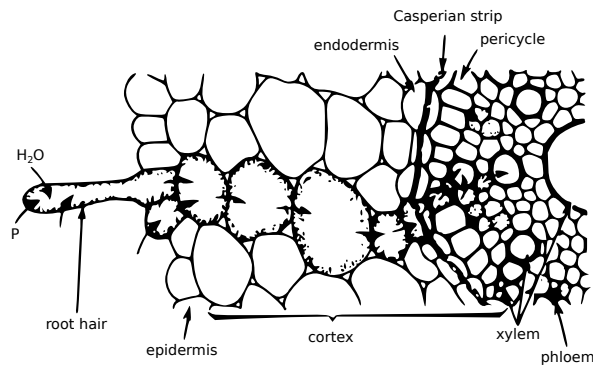


Figure 1.3: Root cross section of cells and tissues involved in ion absorption (adapted from Barber 1995, page 50).

metabolites and intermediates in the photosynthetic C cycle in plants (Plaxton and Lambers, 2015). Once the supply of P is limited, plants naturally respond to combat the reduced P availability through root morphological changes (to facilitate top soil foraging and increase root length), increased excretion of organic acids/enzymes (to mobilize existing soil organic P reserves), and internal adjustments to modify the metabolism's P use efficiency (Plaxton and Lambers, 2015; Hinsinger et al., 2009). Additional strategies include formation of cluster roots and association with mycorrhizae (Hinsinger et al., 2009) whose function is to increase the soil volume explored by roots.



Figure 1.4: Growth of corn plants at the  $0 \text{ mg P kg}^{-1}$  soil (left),  $50 \text{ mg P kg}^{-1}$  (middle), and  $175 \text{ mg P kg}^{-1}$  (right) grown on a Brazilian Oxisol (reproduced from V. D. Fageria 2001, page 98)

In an agricultural system, the primary concern regarding P availability is the decrease in harvested grain yield and quality (or concentration). Since its main mode of function is the removal of P via grain and silage, farmers will rely mostly on additional inputs to maintain the steady P loss. As P availability depends strongly on soil properties and existing P reserves, the applied fertilizer will be more or less efficient (or yield gained per nutrient applied; Valkama et al., 2009; V. D. Fageria, 2001). In general, soils with a longer history of P addition (and thus higher soil water P concentration) will respond less

to additional P inputs (Valkama et al., 2009; V. D. Fageria, 2001). Furthermore, soils with finer composition and higher organic matter content have a greater P buffering capacity, enabling the supply of P to plants over a longer period of time (Valkama et al., 2009). For the influence of pH, different forms of the  $\text{PO}_4^{3-}$  ion will combine to form insoluble compounds (with Fe and Al in acidic, and Ca in alkaline soils) and determine P recovery from applied fertilizer, which can go down even to 10 - 20% in such soil as Brazilian Oxisols (V. D. Fageria, 2001).

Apart from applying fertilizer, farmers employ different management practices to increase soil P availability and the effectiveness of applied P (V. D. Fageria, 2001). Liming is a common practice in acidic soils, where soil pH is increased through the addition of Ca and magnesium rich materials which can (by forming carbon dioxide [ $\text{CO}_2$ ] and  $\text{H}_2\text{O}$  with H ions) reduce the P immobilization capacity of Fe and Al compounds (V. D. Fageria, 2001). The type of P fertilizer can also play an important role, with water soluble ones being more suited to high pH (>6.0) soils as rock  $\text{PO}_4^{3-}$  requires an adequate supply of H ions to break it down (V. D. Fageria, 2001). Exact timing and placement can significantly improve effectiveness of applied P. Banding is a practice of incorporating P fertilizer in the vicinity of the planted crop (just before sowing), which tries to avoid unnecessary contact with soil as opposed to broadcasting where fertilizer is thrown over a much larger surface (V. D. Fageria, 2001). Other practices (to name a couple) include optimal application of P with respect to other major nutrients (primarily N and potassium [K]) to avoid losses due to co-limitation by either one, as well as the use of P efficient genotypes which produce more biomass per P taken up, or have better P acquisition capabilities via better root exploration (V. D. Fageria, 2001).

Accounting for all of the different mechanisms that modify soil-plant P transfer and how they impact cropland productivity is a daunting task, most often performed in field trials to separate exact contribution of each sought factor over one or many growing seasons (Amanullah and Inamullah, 2016; Batten, Khan, and Cullis, 1984; Dai et al., 2013; Dobermann et al., 1998; N. K. Fageria and Oliveira, 2014; Gallet et al., 2003; G. P. D. Jones, Blair, and Jessop, 1989; Sahrawat et al., 1995; J. Shen et al., 2011; Takahashi and Anwar, 2007; Tang et al., 2008; Tonitto and Ricker-Gilbert, 2016; Valkama et al., 2009; van der Velde et al., 2013; Wissuwa and Ae, 2001; L. Wu et al., 2015). In these, properties of the soil and the harvested plant, as well as nutrient addition are tracked to construct yield curves (Steenbjerg, 1951) for the purpose of optimizing a management practice in a specific growing environment (van Keulen, 1982). Even though many variations exist on how yield is related to a certain property (plant, soil or other; eg. Janssen et al. 1990, Bai et al. 2013, Morel et al. 2014) they all boil down to fitting empirical functions in order to explain plant response (Wit 1953; Fig. 1.5)

To overcome the limited descriptive extent of any kind of empirical treatment, crop simulations have been developed to better account for the underlying forces and their

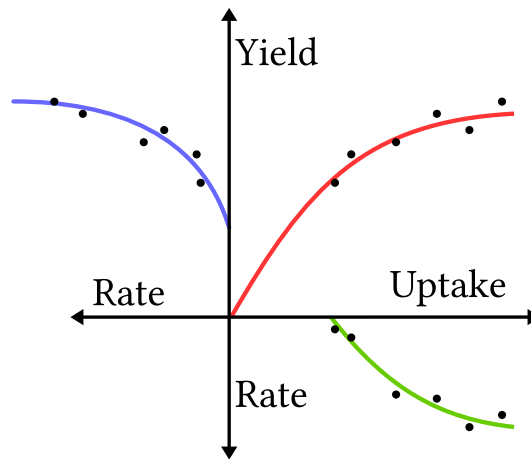


Figure 1.5: A depiction of the well-known “three-quadrant” method pioneered by Wit (1953). The plots describe a hypothetical high input crop system (seen by the blue curve intersect) through rates of nutrient input ( $\text{kg ha}^{-1}$ ), grain yield ( $\text{t ha}^{-1}$ ) and grain nutrient uptake ( $\text{kg ha}^{-1}$ ). Black dots would represent observed data. Nutrient input rate and grain uptake could be either N, P or K.

consequences on final crop production, as well the surrounding environment (eg. Brisson et al. 2009; Williams et al. 1990; James W Jones et al. 2003; Keating et al. 2003; Van Diepen et al. 1989). Crop simulations describe a dynamic system where state variables (like leaf area index [LAI], biomass, soil water, soil nutrient pools, etc.) are driven by external ones (like radiation, temperature, precipitation, irrigation, nutrient addition, etc.) and are integrated to reproduce a growing plant, most often on a daily time step (Wallach, Makowski, and J. W. Jones, 2006). Since our knowledge of the individual processes guiding crop development is substantial (leaf assimilation, transpiration, respiration, root nutrient uptake, etc.), their combination should ideally provide a reasonable framework for studying crop response in hypothetical growing conditions (Wallach, Makowski, and J. W. Jones, 2006). Furthermore, the uncertainty and sensitivity of the simulated result can easily be studied and used to perform risk analyses (Wallach, Makowski, and J. W. Jones, 2006). The advantages of crop simulations in studying P limitation should thus be clear, due to the intricacies of soil-plant P cycling and their possible variation in different growing conditions; not to mention the future uncertainty of a changing climate (Asseng et al., 2013).

## 1.4 P limitation in croplands: from field to the globe

Field trials (and their simulated counterparts) offer the most precise and trustworthy estimates of systemic effects a certain practice might have in a field of crops, but they also depend on the very specific conditions they were obtained in, making extrapolation to larger scales a questionable affair (van Ittersum et al., 2013). To this end, two methodologies have surfaced as a way to understand limits to crop yield on the regional and global scale. The first one is based on a statistical description of yield response, whereas the

second one tries to achieve the same through crop simulation (van Ittersum et al., 2013).

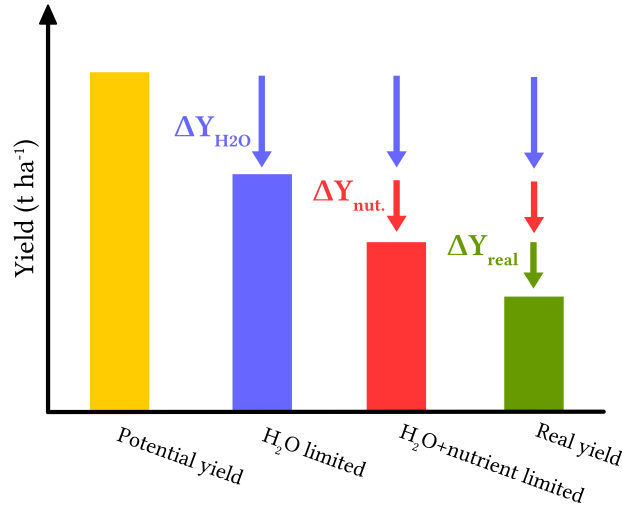


Figure 1.6: The yield gap concept (adapted from van Ittersum and Rabbinge 1997, page 6). Bars denote yields which experience losses or yield gaps ( $\Delta Y$ ) due to sub-optimal water ( $\Delta Y_{H_2O}$ ) and nutrient ( $\Delta Y_{nut.}$ ) availability, or additional presence of reducing factors like pests or disease observed in the field ( $\Delta Y_{real}$ )

To understand the effects of a certain practice and quantify its impact on crop yield, it is useful to look at its potential to increase productivity relative to the case of it being absent; commonly termed a “yield gap” as in van Ittersum and Rabbinge (1997). A yield gap is defined as the difference between the potential, limited and real yield (Fig. 1.6). Potential yield is determined by growth-defining factors such as climate and plant genetic makeup, and should be the maximum achievable one when all negative yield effects are removed. Limited yield is the one experiencing losses due to a sub-optimal resource supply, which are  $H_2O$  and nutrients. Real yield is the one where all negative effects are taken into account (resource limitation, pests, disease, etc.) and should correspond to what is observed in the field. These terms serve to identify the most important crop, soil and management factors limiting yields in a specific growing environment, and can serve as a base for further economic and policy analysis (van Ittersum et al., 2013).

The yield gap concept has often been adopted by the mentioned statistical methods, which rely on spatial and temporal datasets of major crop yield drivers: climate, irrigation and nutrient management (van Ittersum et al., 2013). These are used to form robust empirical relationships to observed yields, enabling the quantification of the main response (Fig. 1.7) as well as the uncertainty that goes with it (D. Lobell and C. Field, 2007). For example, spatially aggregated climate trends (of growing season temperature and precipitation) can explain at least 30% of year-to-year change in the world’s six most widely grown crops (D. Lobell and C. Field, 2007). Similarly, historical temporal variation in a region’s climate can explain up to 40% (as a global average) of the historical variability in maize, rice, wheat and soybean yield (Ray et al., 2015). Similar links can be found for management practice impacts, as the history of  $H_2O$  and nutrient addition can explain a

big portion of today's global crop production, pointing to their significance in combatting existing production gaps (Lassaletta et al., 2014; Mueller et al., 2012; Sinclair and Rufty, 2012). The main strength of these approaches is their ability to aggregate all additional effects tied to the main input driver, which would otherwise require an elaborate experiment design. But seeing as they are empirical in nature, data availability and the reduced description of the underlying system are the main barriers to their extension and, most importantly, extrapolation outside of the range of the environments inquired (temporal, spatial or otherwise; D. Lobell and Burke 2010). This gap between insufficient information and the need for knowledge on the potential impacts is the one crop simulations try to fill in.

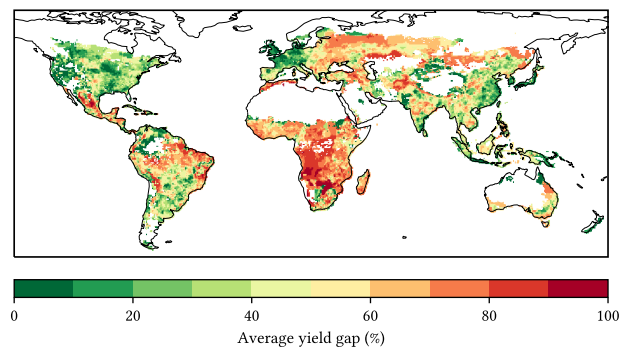


Figure 1.7: The average yield gaps for maize, wheat and rice combined (adapted from Mueller et al. 2012). The yield gap was calculated using the potential yield, which is in turn estimated statistically via climate binning techniques (van Wart et al., 2013)

Crop simulations from regional to global scale have been made too many too count, but with one recent inter-comparison project standing out called The Agricultural Model Intercomparison and Improvement Project (AgMIP; Rosenzweig et al. 2013; Elliott et al. 2015). AgMIP is an effort to better understand feedbacks of climate change on global crop production while improving our crop simulation knowledge, with the final goal of estimating regional vulnerabilities, producing targeted risk assessment and adaptation strategies for future cropland systems (Rosenzweig et al., 2013). At the core of its global crop yield study, AgMIP puts a dozen of well known global gridded crop models (GGCMs) head to head in predefined future management and climate scenarios, and looks at the implications for the future global yield (Elliott et al., 2015). Their results are unequivocal that, in the short term, yield increases could be observed but with long term gains turning into losses due to mainly temperature increase (Fig. 1.8, Rosenzweig et al. 2014; Asseng et al. 2015). But as the AgMIP collaborators readily admit, significant uncertainties exist due to the underlying parametrization (mostly regarding CO<sub>2</sub>, H<sub>2</sub>O and nutrient effects) whose results diverge quite a lot among different GGCMs (Asseng et al., 2013; Rosenzweig et al., 2014; C. Müller et al., 2017) and which seems to happen irrespective of the amount of parametrizing information thrown at them (Bassu et al., 2014). These problems seemingly originate from

out-of-date concepts still serving as a base for a lot of modern crop simulations, that were originally conceived as engineering tools rather than biological models (Passioura, 1996) and which urgently need to integrate newer and more recent findings in plant physiology (Rötter et al., 2011; Franklin et al., 2012; I. Prentice, 2013).

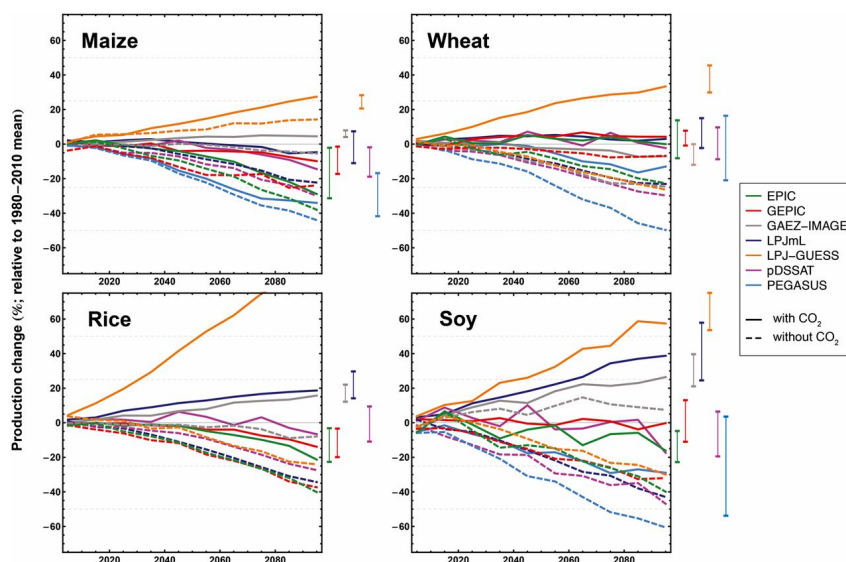


Figure 1.8: Figure reproduced from Rosenzweig et al. (2014). The plot shows relative change (%) in RCP8.5 decadal mean production for each GGCM (based on current agricultural lands and irrigation distribution) from ensemble median for all GCM combinations with (solid) and without (dashed) CO<sub>2</sub> effects for maize, wheat, rice and soybean; bars show range of all GCM combinations with CO<sub>2</sub> effects. GEPIC, GAEZ-IMAGE, and LPJ-GUESS (GGCMs participating in AgMIP) only contributed one GCM without CO<sub>2</sub> effects

To come back to the issue of P limitation effects on global yield, estimates have been partially provided using statistical treatment of observed global yield by relating it to irrigation intensity and chemical fertilizer (NPK) input (Mueller et al., 2012). Apart from critically missing pieces like soil nutrient availability and losses, application of organic fertilizer and the inability to distinguish rainfed from irrigated yield in the utilised census data (among others; Mueller et al. 2012), the true difficulty lies in the co-linearity of the input drivers which makes it difficult to separate their true contribution (best seen in Fig. 1.1). This should make it clear that, for the purpose of P (as mentioned in the last section) crop simulation should ideally be a better candidate. To get to this stage though, a better understanding of how historic use of P is needed, as well as how this might impact agricultural soils worldwide.

The issue of P impact on agricultural soils on regional to global scale has been subject of current research, where the temporal and spatial variation of the P input/output balance or the various soil P fractions have been studied (S. Z. Sattari et al. 2012; Bouwman et al. 2013; Ringeval et al. 2017; J. Zhang et al. 2017). Much of this work had been sparked by MacDonald et al. (2011) who looked into recent worldwide agricultural P use balance (year 2000) and found that, in spite of a global P surplus (fertilizer P input greater than P

removal via crop harvest) almost 30% of total cropland area experiences a P deficit (Fig. 1.9). Subsequently, S. Z. Sattari et al. (2012) relied on national and regional census data to distinguish the temporal evolution of soil P inputs and outputs, and how these might affect soil P fractions via a simple soil P cycling model. Bouwman et al. (2013) allowed for the spatial extension of the same, by providing global maps of agricultural soil P balance during the 20th century. These results have been used to drive further modelling studies like Ringeval et al. (2017) and J. Zhang et al. (2017) who estimated the global spatio-temporal evolution of various soil P fractions using soil P cycling models, thus providing a suitable base for investigating global crop P limitation by statistical or crop simulation approaches.

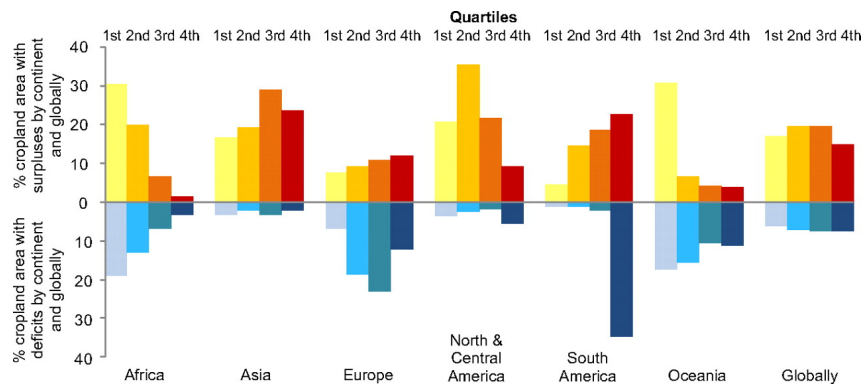


Figure 1.9: Distributions of P surpluses and deficits by quartiles shown as percent of total cropland area in each continent and as percent of global cropland area. Reproduced from MacDonald et al. (2011).

In crop simulations, the plant-soil nutrient cycle is typically described using a soil transformation and a plant uptake module (Brisson et al., 2009; C. A. Jones et al., 1984) which interact to reproduce the long term effects of nutrient availability on crop yield. While the form of the soil nutrient module is quite similar to ones mentioned before (S. Z. Sattari et al., 2012; Ringeval et al., 2017; J. Zhang et al., 2017), the plant one relies on an empirical description of nutrient requirement at different development stages (Lemaire, 1997; Stockle and Debaeke, 1997). To reconcile the two, stress functions are computed on the basis of soil nutrient supply and plant nutrient demand while taking into account internal or external factors (like plant nutrient status or soil quality; respectively) and applied to different time-integrated quantities (like LAI or biomass change) to limit plant growth (Brisson et al., 2009; C. A. Jones et al., 1984). The most astonishing thing is that an actual mechanism for root growth and nutrient uptake is seldom used, or at least one which provides direct feedback on plant growth (E. Wang and Smith, 2004). Root mass is instead diagnosed using prescribed allocation patterns (similar to plant nutrient demand) and subsequently modified according to soil properties to reproduce observed root profiles (E. Wang and Smith, 2004). This (quite heavily parametrized) approach stems from the mentioned field-scale models and encounters significant difficulties once applied to regional and global scales, where it has been shown that soil-type induced yield variability can outweigh even the inter-annual one (Folberth et al., 2016). This points not only to

the need of more soil information (Folberth et al., 2016) but yet again to the urgent need of more up-to-date concepts in simulating crop yields worldwide (Passioura, 1996; Rötter et al., 2011; Franklin et al., 2012; I. Prentice, 2013); especially for the purpose simulating P limitation.

A good candidate for studying P limitation in a more general fashion are dynamic vegetation models (DGVMs) which simulate plant growth as a product of fundamental physiological processes (Ball, Woodrow, and Berry, 1987; Farquhar, von Caemmerer, and Berry, 2001). These are combined with models of surface H<sub>2</sub>O, energy and C exchange (Krinner et al., 2005) to describe the complete atmosphere-biosphere coupling and study global change impacts, mostly in natural eco-systems (C. Peng, 2000). In DGVMs, the importance of modelling P limitation has duly been recognized (Reed, X. Yang, and Thornton, 2015) with research on the subject steadily moving forward (Y. P. Wang, Houlton, and C. B. Field, 2007; Q. Zhang et al., 2011; Goll et al., 2012; X. Yang et al., 2014; Goll et al., 2017; Sun et al., 2017). Even though these models look at the complete atmosphere-biosphere coupling, they recently try to answer the same fundamental question: how will P ultimately dictate a system's productivity and how will this play out in the future? (Reed, X. Yang, and Thornton, 2015). Thus, their generic nature should allow for the inclusion of an agricultural soil P model like the one of Ringeval et al. (2017) which, combined with a root P uptake model like the one of De Willigen and van Noordwijk (1994), could enable the study of P limitation in crops on a global scale; in the past, present and the future.

## 1.5 Motivation, objectives, questions

Taking into account the previously highlighted issues and knowledge gaps, coupled with the potential of P to impact global agriculture, the basic driving question is **how much does P contribute to world crop production limits?** But to arrive to this answer, we need to establish **the necessary minimum needed to explain plant response to a changing P environment**. This could be done by formulating a system that describes soil-plant P transfer more faithfully, in which a plant responds to and evolves as function of P availability. Only if the previous item has been acceptably addressed, can we **implement models of crop response to P limitation**. Once this model has been verified against observed crop growth in different P availability conditions, we can attempt to **estimate current effects and make projections of P management impact in crops worldwide**.

To go about this, we should first have a tentative look at current global P limitation using existing soil P and crop yield information, and identify potential solutions to improve the representation of the P cycle and ultimately its limitation. Since models of soil P cycling and root P uptake already contain a sound conceptual basis, we should focus on a way to couple the two by relying less on an empirical description (common in today's crop simulations) and more on current physiological concepts of plant nutrient use. This should result in a hypothetical plant model, which should be developed and tested on observed growth in varying P conditions. The hypothetical model should then be integrated within a more physically based one to allow the study of real crop systems. Finally, global limitation due to P management should be investigated.

All of the above summarized, the questions are following:

- How much can current soil P availability limit today's cereal production?
- How can we model the impacts of P availability on plant growth?
- How can we enable the study of P limitation in crops using the previous model?
- What are the consequences of historical and future P management scenarios on global cereal yield?

## 1.6 Outline

The following dissertation will consider the previously mentioned questions through the following structure:

- Chapter 2 diagnoses global P limitation in cereals using existing simulated information on soil P (Ringeval et al., 2017) and crop yield (X. Wu et al., 2016), while combining them through a mechanistic description of the underlying P transfer processes (De Willigen and van Noordwijk, 1994). Results and their uncertainty are evaluated, while trying to identify key components which need to be taken into account when combining the two models (soil and crop one).
- Chapter 3 deals with applying the physiological concepts of optimal functioning (Dewar et al., 2009) to annual plant growth, resulting in a conceptual allocation model that is evaluated against observed maize field data and a set of hydroponic studies. The aim here is to provide a parsimonious and transparent basis for studying C & P interaction in plants, which is to be coupled to a more physically driven vegetation model to allow the study of P limitation in a generic growth environment.
- Chapter 4 deals with implementing the previously developed allocation model to a physically driven one (Krinner et al., 2005) and evaluates it against observed field maize trials in two temperate sites. The results, current model short-comings and future opportunities are discussed for the application of optimal functioning principles in DGVMs, and their use in studying P limitation in crops worldwide.
- Chapter 5 wraps up the findings of the work so far, discusses their implications (particularly with respect to the initial findings) and provides some thoughts on how to approach and understand crop P limitation effects globally.

## Chapter 2

# Quantifying the limitation to world cereal production due to soil phosphorus status

Marko Kvakić, Sylvain Pellerin, Philippe Ciais, David L. Achat, Laurent Augusto, Pascal Denoroy, James S. Gerber, Daniel Goll, Alain Mollier, Nathaniel D. Mueller, Xuhui Wang, Bruno Ringeval

Published in *Global Biogeochemical Cycles* 31:1 (2018)

## Abstract

Phosphorus (P) is an essential element for plant growth. Low P availability in soils is likely to limit crop yields in many parts of the world, but this effect has never been quantitatively assessed at the global scale by process-based models. Here we attempt to estimate P limitation in 3 major cereals worldwide for the year 2000 by combining information on soil P distribution in croplands and a generic crop model, while accounting for the nature of soil-plant P transport. As a global average, the diffusion limited soil P supply meets the crop's P demand corresponding to the climatic yield potential. This is due to the large amount of legacy soil P in highly fertilized areas. However, when focusing on the spatial distribution of P supply vs. demand, we found strong limitation in regions like North and South America, Africa and Eastern Europe. Integrated over grid-cells where P supply exactly meets or is lower than potential P demand, the global yield gap due to P limitation is estimated to be 22, 55 and 26 % in wheat, maize and rice. Assuming that a fraction (20%) of the annual P applied in fertilizers is directly available to the plant, the global P yield gap lowers by only 5 – 10 % underlying the importance of the existing soil P supply in sustaining crop yields. The study offers a base for exploring P limitation in crops worldwide, but with certain limitations remaining. These could be better accounted for by describing the agricultural P cycle with a fully coupled and mechanistic soil-crop model.

## 2.1 Introduction

Phosphorus (P) is recognised as one of the foremost concerns for global crop production (Cordell, Drangert, and White, 2009) because as a macro-element, it is indispensable for plant function and thus growth. P is a relatively immobile element compared to carbon (C) and nitrogen (N) as it reacts with soil constituents that buffer its availability to plants, and through time becomes unavailable altogether (Filippelli, 2002; Sample, Soper, and Racz, 1980). In agriculture, natural levels of P in most soils are not adequate to sustain crop production in the long-term, so additional inputs in form of chemical fertilizer or manure are needed. Depending on the access to fertilizer (chemical or manure) P management in agricultural systems shows a wide variety of outcomes (P. M. Vitousek et al., 2009) ranging from soil nutrient depletion and degradation on one extreme, to fertilizer over-consumption and pollution of water resources on the other.

These issues have recently been highlighted in studies (MacDonald et al., 2011; Mueller et al., 2012) showing the imbalance of nutrient flows in croplands globally. Looking at the soil's P budget, 30% of the worldwide cropland area experiences a P deficit (MacDonald et al., 2011) despite a global positive P balance. Even more, it has been estimated that around 70% of underachieving areas could resolve their yield losses by solely focusing on nutrient limitation and not irrigation infrastructure (Mueller et al., 2012). These studies give a strong case for P mismanagement on the global scale, but how this ultimately affects yields requires further attention.

Nutrient limitation in crops is most commonly estimated in field trials (Dai et al., 2013; Dobermann et al., 1998; Gallet et al., 2003; Shen et al., 2004; Takahashi and Anwar, 2007; Tonitto and Ricker-Gilbert, 2016; Valkama et al., 2009; Wissuwa and Ae, 2001), where the crop yield is recorded and related to soil nutrient availability or the amount of applied fertilizer. It is often shown that P addition can significantly increase yields, especially through its interaction with N (V. D. Fageria, 2001). These studies give a precise measure of nutrient requirement, but the variety of experimental conditions makes a general estimate of P limitation hard to obtain, especially when dealing with the issue globally. More importantly, yield increases due to P are highly dependent on the properties of the soil underneath. Soil texture and organic content play a significant role (Tonitto and Ricker-Gilbert, 2016; Valkama et al., 2009) by changing the P buffering capacity of the soil, whereas soil metal or carbonate content (J. Shen et al., 2011; Takahashi and Anwar, 2007) determines the effectiveness of applied fertilizer through formation of less available P compounds. P availability, which is intimately tied with the soil, is thus an important factor in sustaining crop yields.

Nutrient limitation manifests most importantly as a loss of yield, termed a yield gap (van Ittersum and Rabbinge, 1997). To estimate it on a global level, a set of methodologies are usually employed as summarised by van Ittersum et al. (2013). They are broadly categorised into statistical and crop-modelling approaches. Statistical approaches rely on yield

census data of varying size, which are then combined with global data (e.g. climate, fertilizer inputs, irrigation area). Based on climate binning techniques, they divide the world into contingent areas of similar growth environment, looking at main drivers of yield variation within. On the other hand, global gridded crop models (GGCMs) simulate growth and yield using relatively generic parametrizations and calibrated parameters to mimic the effects of genetics (e.g. maximum leaf area index and harvest index). These models can be run globally on a grid, forced by daily weather and management-related data (eg. planting date, sowing density, irrigation, annual fertilizer application). Only a few GGCMs participating in inter comparison projects have an explicit representation of nutrient cycling in soils and plants (Rosenzweig et al., 2014) and their parametrization remains rather coarse, with no representation of key process such as the limitation of root uptake by soil P diffusion. Moreover, the use of crop models to estimate limitation due to soil P globally has been limited by a lack of information on the distribution of P in agricultural soils.

Recent attempts (Ringeval et al., 2017; J. Zhang et al., 2017) have simulated the evolution of P in agricultural soils, by combining spatially explicit datasets on soil P inputs/outputs and a soil P dynamics model. Such data can be combined with GGCMs output of potential yield, to assess the yield limitation due to insufficient P supply compared to the potential crop P demand. Here we attempt to diagnose P limitation due to soil P status in three major cereals in the world (wheat, maize and rice) for the year 2000, by combining the outputs of a generic GGCM known as ORCHIDEE-CROP (X. Wu et al., 2016) and an agricultural soil P model (Ringeval et al., 2017), while accounting for the mechanism of soil to plant P transfer by diffusion (De Willigen and van Noordwijk, 1994).

## 2.2 Methods

### 2.2.1 Overview

Quantifying P limitation on yield is attempted using a supply vs. demand approach (Fig. 2.1). Crop P demand is the P needed to achieve a certain level of yield, which is in our case the potential one or one that is limited only by climate (temperature, light) and genetic make-up. On the other hand, the soil P supply is derived from the soil P availability by accounting for the diffusive nature of P transport at the soil-root boundary.

To estimate crop P demand we used ORCHIDEE-CROP (X. Wu et al., 2016, red box in Fig. 2.1), a GGCM that simulates plant growth in irrigated conditions without any other stress (nutrients, pests or weeds) at 0.5° resolution over the globe. From the modelled crop carbon pools, we derive spatially explicit estimates of P demand using three different approaches described in the following sections. To estimate soil P supply we used an agricultural soil P dynamics model (Ringeval et al., 2017), which gives the amount of soil P available for plant uptake (blue in Fig. 2.1). Because the soil P supply is determined both by soil P availability and the plant's ability to take it up, a model of root uptake (green box

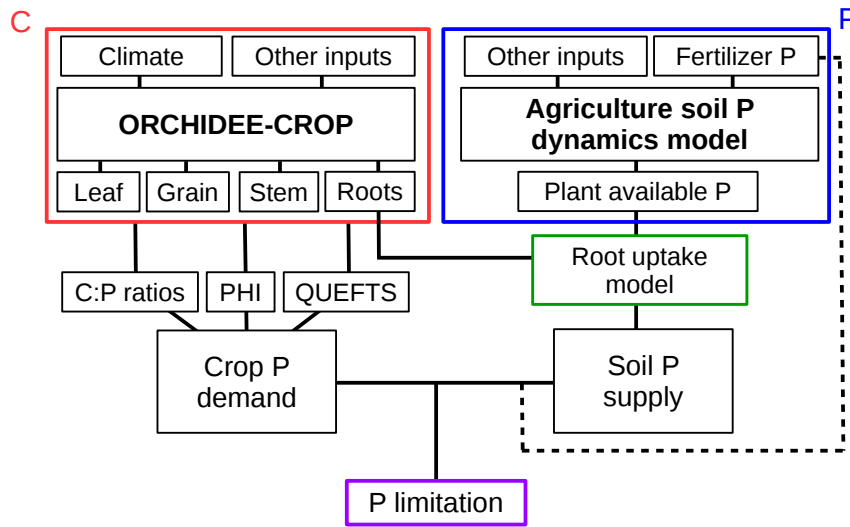


Figure 2.1: Information flow diagram describing the methodology of this study. The red and blue rectangles delineate the main sources of data, and correspond to carbon and phosphorus respectively. The box in green delineates the root uptake model, which acts as an interface for P in soil that is available for plant uptake. These three (colored) boxes form the basis for estimating potential crop P demand and soil P supply, and finally P limitation. An additional sensitivity test takes into account fertilizer P (chemical + manure) in addition to soil P supply (dotted line)

in Fig. 2.1) was employed to compute potential P uptake by plants. By comparing crop P demand to soil P supply, the limitation to crop production due to soil P status can be quantified. Additionally, a scenario with an improved soil P supply is carried out, which assumes that a part of the fertilizer P inputs (chemical and manure) could be incorporated into crop biomass (dotted line in Fig. 2.1).

### 2.2.2 Crop P demand

P demand is diagnosed from the different plant C pools simulated by ORCHIDEE-CROP: leaf, grain, stem and root (the model has also a pool for carbohydrate reserves, which is assumed to contain no P). The simulations used here were performed following the ISIMIP2a protocol (Warszawski et al., 2014) assuming no stress occurs (water, nutrient, weed, pests). To estimate P demand we used 3 different methods: an empirical relationship between yield and uptake (QUEFTS), P harvest index (PHI) and stoichiometric ratios (C:P). These methods provide us information on the amount of P needed to achieve a certain biomass or yield at maturity. When searching for methods to infer crop P demand, we have tried to pick ones whose data encompasses a wide range of growing conditions (genetic variety, geographic location, etc.) to obtain a general and a robust estimate. Additionally, the use of multiple methods allows us to give a range of possible values and account for their inherent bias. Details of the crop model and estimation of P demand are detailed in the following sub-sections.

## ORCHIDEE-CROP

Crop biomass pools are simulated by ORCHIDEE-CROP (X. Wu et al., 2016), a version of the ORCHIDEE dynamic global vegetation model (Krinner et al., 2005) modified to simulate phenology, temporal evolution of LAI and C allocation (including grain filling that determines yield) of crops by integrating some parameterizations of the generic crop model STICS (Brisson et al., 2009). The model has been calibrated and evaluated for maize and winter wheat in temperate areas (X. Wu et al., 2016), and extended to include rice (X. Wang et al., 2017).

The simulated crop growth is divided into seven development stages (e.g. emergence, vegetative growth, grain filling, etc.). The timing and duration of each stage is calculated based on development units, which describe the physiological requirements of crops. These development units are calculated as growing degree days weighted by limiting functions to account for photoperiodism and vernalization. LAI is the primary driver of canopy photosynthesis, and is described by a logistic curve depending on the plant phenology stage. LAI is used in conjunction with the leaf-scale models for C3 and C4 plants used in ORCHIDEE (Krinner et al., 2005) to determine gross primary productivity (GPP). Carbon assimilates from GPP are distributed daily to leaves, stem, grain, roots and reserves. A fraction of GPP and biomass is used for maintenance and growth plant respiration, to calculate net primary productivity (NPP). The priority of C allocation changes in the model with the plant development stage, and reflects different resource use strategies. ORCHIDEE-CROP also has water and nutrient modules, whose purpose is to regulate plant growth. Water uptake is modelled as a function of plant transpiration and root water availability using a detailed and realistic multilayer soil hydrological model (M. Guimberteau et al., 2014). The nutrient component is highly idealized: it considers only N, whose soil-plant budget is not explicitly simulated. Instead, leaf-scale photosynthetic capacity is increased with N fertilizer addition through an empirical relationship, with no feedbacks to C allocation.

The crop simulations used here follow the protocol set out by “The Global Gridded Crop Model Intercomparison” (GGCMI) project (Elliott et al., 2015). It specifies crop species, their management and a set of inputs (gridded climate and crop management data) to insure comparability among different models. The most important crops are maize, rice and wheat as they comprise 43% of total global food intake (Elliott et al., 2015). The output used to calculate crop biomass pools in this study were the mean monthly values, which were simulated with a daily time-step at a  $0.5^\circ$  resolution globally with irrigated conditions in the year 2000. Irrigated conditions were simulated by setting the soil moisture to field capacity everywhere. N was assumed non-limiting by setting a high N fertilizer application rate (150 kg N/ha) everywhere.

## QUEFTS, PHI and C:P ratios

QUEFTS (Janssen et al., 1990) is an empirical model used to estimate yield response according to fertilizer inputs and soil fertility. It was developed to provide advice on optimal nutrient application in tropical soils in a simple and robust way, and has been extensively used to characterize cereal production worldwide (S. Sattari et al., 2014). It consists of four different steps, which define the supply and demand of the three nutrients mostly widely used in agriculture (N, P and K) to give a final yield estimate. The model sections of interest in this study are the optimal nutrient supply combinations and the uptake-yield curves (details in section B.1.2). The latter one provides the yield response depending on the potential yield as function of the plant uptake of P. In our calculation, the yield potential is defined as the ORCHIDEE-CROP simulated yield (irrigated and high N application everywhere). We calculated the crop P demand as the P uptake which enables the crop to achieve 95% of the yield potential. Hereafter, this crop P demand is called  $P_{demand}^{QUEFTS}$  (kgP ha<sup>-1</sup> yr<sup>-1</sup>).

The PHI method is based on a combination of the grain yield simulated by ORCHIDEE-CROP ( $Y_{grain}$ ) and parameters derived from literature corresponding to 1) the ratio of P contained in grain vs. P in the aboveground biomass (known as P harvest index or  $PHI$ ) and 2) P concentrations in plant organs. With this method, the total P demand ( $P_{demand}^{PHI}$ , kgP ha<sup>-1</sup> yr<sup>-1</sup>) is computed by distinguishing the P demand in shoot ( $P_{shoot}$ , kgP ha<sup>-1</sup> yr<sup>-1</sup>) and root ( $P_{root}$ , kgP ha<sup>-1</sup> yr<sup>-1</sup>) as follows:

$$P_{demand}^{PHI} = P_{shoot} + P_{root} = \frac{P_{\%,grain} \cdot Y_{grain}}{PHI} + P_{\%,root} \cdot Y_{root} \quad (2.1)$$

Where  $P_{\%,grain}$  and  $P_{\%,root}$  are grain and root P concentrations expressed as the amount of P per gram of plant biomass (gP g<sup>-1</sup> biomass), respectively.  $PHI$  (kgP grain kgP<sup>-1</sup> shoot) is the P harvest index. We have included the root biomass to avoid underestimating the P demand, even though this method should rely on grain only. Crop specific values for P concentrations are taken from literature and correspond to values at plant maturity (references in SI Table B.1). Both PHI and HI were taken from a study (Duivenbooden, 1992) compiling fertilizer trial data on nutrient uptake and yield response, coming from a multitude of climatic and socio-economic environments. Values were derived from field experiments in stressed conditions, focusing on the linear part of the uptake-yield curve when nutrient use efficiency is maximal. Consequently, the PHI method gives us an estimate of the minimum amount of P required to achieve a certain grain yield.

The C:P ratios method relies on the stoichiometry of crop tissue (grain, stem, leaf and roots) at maturity based on data found in literature, in combination with the magnitude of each crop C pool given by ORCHIDEE-CROP at harvest. This method has been used for projecting future P demand (Peñuelas et al., 2013) in natural ecosystems, with models that do not describe the P cycle and its constraints on primary productivity. The P demand for

the whole plant was calculated as follows:

$$P_{demand}^{C:P} = \sum_i \frac{P_{\%,i}}{C_{\%,i}} \cdot C_i \quad (2.2)$$

Where  $P_{\%,i}$  and  $C_{\%,i}$  are the P and C concentration of a biomass pool (g P,C g<sup>-1</sup> biomass), respectively.  $C_i$  is the amount of carbon in that pool (gC m<sup>-2</sup>) given by ORCHIDEE-CROP. The different C pools considered here are  $C_{leaf}$ ,  $C_{grain}$ ,  $C_{stem}$  and  $C_{root}$ . Concentration values were obtained from literature (SI Table B.2) for crops, corresponding to plant maturity. If a nutrient factorial design existed, values were taken from the most nutrient limited one (usually the control experiment) to avoid bias in concentrations due to luxury consumption.

### 2.2.3 Soil P supply

The soil P supply is determined by two factors: the amount of plant available P in the soil and the plant ability to take it up. The first factor was provided by a modelling approach (Ringeval et al., 2017) simulating the yearly evolution of soil P status during the 20th century. The second factor, or the ability of a crop to take up P, is determined by the plant root system and the nature of P transport in soil solution. Uptake of P depends almost exclusively on diffusion (Barber, 1980), as concentrations are often too low to provide P by mass flow (or advection). A model of root uptake describing the diffusive flux of P in the root vicinity was thus employed (De Willigen and van Noordwijk, 1994).

### Agricultural soil P

The global distribution of plant available P was given by a modelling approach (Ringeval et al., 2017) that combined global datasets describing the drivers of P cycle and a soil P dynamics model. The soil P dynamics model is based on the Hedley fractionation concept (Hedley, J. W. B. Stewart, and Chauhan, 1982), where P is distributed among various pools of different chemical extractability. These pools interact on different time-scales, and correspond to flows of P between mobile forms at a biological level (mineralization, immobilization, uptake) and highly stable or ‘occluded’ forms of P which are unavailable to plants (Cross and Schlesinger, 1995; Walker and Syers, 1976). From a global vegetation or crop modelling perspective, this representation has emerged favourably in attempts to include nutrient limitation (Goll et al., 2017; Y. P. Wang, Law, and Pak, 2010; X. Yang et al., 2014) on plant development.

The soil P dynamics model combines multiple sources of spatially explicit data which are deemed main drivers of the P cycle: soil biogeochemical background, soil inputs/outputs resulting from farm practices (chem. fertilizer, manure, crop uptake and residues), land use and land cover change, soil temperature and humidity, losses through erosion, atmospheric deposition and soil-order dependent buffering capacity (Ringeval et al., 2017). The

soil P model calculates bulk transformations in the first 30 cm with a yearly time step from 1900 to 2005, separating crop and pasture landuse at 0.5° resolution globally. The different P pools are (from the most to the least mobile): labile inorganic, labile/stabile organic, secondary mineral, apatite and occluded P. The only pool used to calculate the supply of P in our approach is the labile inorganic pool ( $P_{ILAB}$ ), consisting of water soluble, resin and bicarbonate inorganic Hedley fractions. In the model,  $P_{ILAB}$  is used as proxy of the plant available P. It depends strongly on farm soil inputs/outputs and is assumed to be in equilibrium with P bound on secondary minerals ( $P_{SEC}$ ) at a yearly time scale. Here, we assume no interaction between  $P_{ILAB}$  and  $P_{SEC}$  during the growing season. Thus,  $P_{ILAB}$  represents plant available P at the beginning of the growing season, before fertilizer is applied and a new equilibrium with  $P_{SEC}$  is formed.

We used gridded  $P_{ILAB}$  estimates for the year 2000 (Fig. B.3) which consist of 30 different simulations according to the uncertainty analysis of the original study (Ringeval et al., 2017). In the analysis, all of the previously mentioned drivers varied (except land use change) within a predefined range to estimate how each driver's uncertainty affects the spatial distribution of  $P_{ILAB}$ .

### Potential root P uptake

The potential root uptake model (De Willigen and van Noordwijk, 1994) describes roots as a system composed of parallel and vertically placed root cylinders, which are uniformly distributed in the soil column and take up nutrients at the same constant rate. Every single root has a cylinder of influence whose radius depends on total root density. By assuming uniform water flow across this cylinder of influence, the authors (De Willigen and van Noordwijk, 1994) found an analytical solution describing solute transport in a homogeneous soil medium. We use a special case of this solution, when concentration at root surface reaches zero and uptake of P is the same as the rate at which it diffuses there. The root uptake model enabled us to calculate the potential root P uptake ( $P_{uptake}$ , kgP ha<sup>-1</sup>) determined by soil P availability and the crop root system:

$$P_{uptake} = \sum_{i=1}^{12} \pi \cdot z \cdot L_{rv,i} \cdot D \cdot \frac{\rho^2 - 1}{G(\rho)} \cdot C_P \quad (2.3)$$

Where  $\Delta z$  is soil depth (cm),  $L_{rv}$  is monthly root length density (cm cm<sup>-3</sup>),  $D$  is the coefficient of P diffusion (cm<sup>2</sup> day<sup>-1</sup>),  $C_P$  is the mean concentration of orthophosphate ions in soil solution in the top 30 cm (mgP L<sup>-1</sup>),  $\rho$  is a dimensionless ratio of soil cylinder to root radius and  $G(\rho)$  a dimensionless geometric function related to uptake by diffusion only.  $L_{rv}$  was calculated from ORCHIDEE-CROP monthly root biomass values with specific root length (SRL), which was held constant. SRL varies during the vegetative stage, but this effect was deemed negligible compared to seasonal changes in root biomass and the root-shoot ratio. SRL as well as root diameter values were taken at maturity stage.  $D$

was computed by a ‘constant slope impedance factor’ model (Olesen et al., 2001) to mimic decreased diffusivity in porous media. This effect is due to the amount of water filled pore space, which determines the diffusion path length. Further details on estimation of  $\rho$ ,  $G(\rho)$ ,  $L_{rv}$  and  $D$  can be found in the SI section B.1.3.

$C_P$  (mgP L<sup>-1</sup>) was derived from  $P_{ILAB}$  (mgP kg<sup>-1</sup> soil) thanks to an empirical relationship fitted from a database (Achat et al., 2016) compiling isotopic dilution measurements (providing orthophosphate concentration in soil solution) and Hedley measurements (providing  $P_{ILAB}$ ) in various ecosystems (forests, croplands and grasslands). It was calculated using a Freundlich-type equation, which describes the soil solution equilibrium:

$$P_{ILAB} = a \cdot C_P^b \quad (2.4)$$

Where  $a$  and  $b$  varied as a function of soil-order (details in SI section B.1.4) to roughly account for the sensitivity of sorption/desorption to soil properties like pH, oxide concentration, etc. In our approach, we assumed that  $C_P$  is constant during the growing season (given the fact that  $P_{ILAB}$  is computed at a yearly time-step) and thus not modified by the here computed potential plant uptake (Eq. 2.3). A map of bulk density (Hengl et al., 2014) was used to convert  $P_{ILAB}$  from kgP ha<sup>-1</sup> to mgP kg<sup>-1</sup> soil as needed in Eq. 2.4.

#### 2.2.4 P limitation and yield gap due to P limitation

We compare the crop P demand and the soil P supply by computing the following ratio, hereafter called  $R$ :

$$R^i = \frac{P_{supply}}{P_{demand}^i} \quad (2.5)$$

Where  $P_{demand}^i$  is the crop P demand (kgP ha<sup>-1</sup>) and the superscript  $i$  refers to QUEFTS, PHI or C:P ratio methods (Section 2.2.2).  $P_{supply}$  (kgP ha<sup>-1</sup>) is calculated in the following manner:

$$P_{supply} = \begin{cases} P_{uptake} & (I) \\ P_{uptake} + 0.2 \cdot \frac{P_{fert}}{CI} & (II) \end{cases} \quad (2.6)$$

Case I consists of the soil P supply, calculated via the root P uptake model ( $P_{uptake}$ , kgP ha<sup>-1</sup>) (Eq. 2.3) and serves as a baseline of P supply. Case II includes farm P inputs ( $P_{fert}$ , kgP ha<sup>-1</sup>) along with soil P supply, to include the effects of today’s P management practice.  $P_{fert}$  is the labile fraction of annual fertilizer P input (chemical and manure) used as one of the drivers of the soil P dynamics model (Ringeval et al., 2017). We assume that only 20% of applied fertilizer P can be recovered in the same season (Balemi and Negisho, 2012), which seems to be a generous estimate according to the literature.  $CI$  (yr<sup>-1</sup>) is the number of crop harvests per year.  $CI$  was estimated from a dataset on observed cropping areas around the year 2000 (Portmann, Siebert, and Döll, 2010), which is defined as the ratio

of annually cropped area to the cropland extent area excluding fallow land (Eq 2, section 3.3 in Portmann, Siebert, and Döll, 2010). The reason for using the CI parameter is to avoid a high bias in the annual P application rate in multi cropping areas. Case I represents P limitation due to either low soil P legacy or a biophysical limit to P diffusion, whereas case II gives an estimate on how much the current fertilizer application alleviates this limitation. Case II could be considered as a lower limit to crop P limitation where only 20% of the applied fertilizer is incorporated into plant biomass, bypassing the diffusion+root uptake mechanism.

We calculate the yield gap (van Ittersum and Rabbinge, 1997) due to P limitation ( $Y_{gap,P}$  in %) by assuming a linear relationship between the  $Y_{gap,P}$  and the demand:supply ratio (Eq. 2.5) as follows :

$$Y_{gap,P}^i = \max(0, 1 - R^i) \quad (2.7)$$

We restrict this computation to the case where the P supply is smaller or equal to the P demand by assuming that plants cannot take up more P than needed during the growing season. P limitation (and its yield gap) was computed using the ORCHIDEE-CROP simulations performed in irrigated and high N fertilizer conditions that represent the true yield potential (i.e. limited only by climate and genetic make up). We compared our estimates to a globally gridded dataset (Mueller et al., 2012) on the observed yield gaps. In this dataset, the world grid points are divided into climate bins, defined by different pairs of growing degree days and amount of yearly precipitation. Within a climate bin, the potential yield is defined as the area-weighted 95th percentile of the observed yield within the same climate bin. The yield gap then is the difference of the mean and the potential yield within the same bin. These yield gaps were expressed as fractions of their yield potential to make them comparable with our estimates.

### 2.2.5 Analysis of uncertainty, commonality analysis and global averages

A Monte-Carlo (MC) method was used to estimate uncertainty in the P limitation estimates. Namely, the parameters were varied around their mean in 1000 replicates according to their standard error (SE) provided by the source literature. If SE was not provided, we assumed it as a fraction of the mean as either 5% or 20%. In PHI (2.1) and C:P ratios (2.2) methods, parameters/variables whose uncertainty was considered are: nutrient concentration of the plant organs, PHI and the HI parameters. In the QUEFTS method, we varied the parameters related to optimal nutrient uptake. In Eq. 2.3 and 2.4, an uncertainty was considered for:  $P_{ILAB}$ , Freundlich equation parameters, constant-slope impedance diffusion parameters, SRL and root diameter. In Eq. 2.6  $P_{fert}$  uncertainty was considered.

A commonality analysis (CA, Reichwein Zientek and Thompson, 2006) was used to explain factors contributing to the spatial and the gridpoint variability of the root uptake relation (Eq. 2.3). To determine the contribution to spatial variability, twelve monthly val-

ues at every grid point were averaged in time, their median from 1000 MC replicates taken, log transformed and then z-scored around the global mean to form a multiple linear model of the dependent variable  $P_{uptake}$  with predictors  $L_{rv}$ ,  $D$ ,  $C_P$  and  $P_{ILAB}$ . Similarly, to estimate the contribution of the predictors at every grid point, the variables were averaged in time, log transformed and z-scored around the mean of the 1000 MC replicates. The CA protocol determines variance belonging to all possible combinations ( $15=2^4-1$ ) of predictor variables. It then successively assigns unique or common (due to collinearity) variance to each of the predictors, by combining the  $R^2$  values of each multiple linear model combination. In the spatial variability CA, 1000 bootstrap samples were taken to obtain the 95% uncertainty interval around the global median CA coefficients. The bootstrap procedure was not performed for the grid point uncertainty estimates.

In the following, global averages of P supply, demand and limitation are calculated by weighting the grid-cell values with the total observed crop area (Portmann, Siebert, and Döll, 2010), and 95% uncertainty interval is given (in brackets) to the median value of the 1000 MC replicates. A similar technique was applied when presenting histograms of these quantities in each grid-cell, where the contribution of every grid-cell depends on the fraction of crop area within it. When presenting histograms, uncertainty is shown as the 68% interval around the median value of 1000 MC replicates for every grid-cell. The 68% interval would correspond to 1 standard deviation, if the 1000 MC values at each cell were normally distributed.

We masked and filtered some grid-cells before performing our calculation. For each crop, we started from grid-cells where the observed crop area is non-null (Portmann, Siebert, and Döll, 2010). We then kept grid-cells considered in the ORCHIDEE-CROP model for which a value of  $P_{ILAB}$  is provided by the soil P dynamics model. These criteria reduce the considered crop area to 78%, 71% and 63% of the observed crop area for wheat, maize and rice respectively. Additionally, we have removed grid-cells for which ORCHIDEE-CROP simulated no grain, as well as grid-cells where the crop model simulates plants that remain in the vegetative phase during the whole growing season. The no grain condition especially concerns wheat, since we simulated winter wheat only. In the model, wheat has a vernalization requirement and can not grow properly in low latitudes. Once these were removed, the final crop area covers 35%, 67% and 61% of the observed area (Portmann, Siebert, and Döll, 2010) for wheat, maize and rice respectively.

## 2.3 Results

### 2.3.1 Potential root P uptake

Global averages of potential root P uptake are 20.2 (15.5 – 26.3), 20.9 ( 15.3 – 31.9) and 13.6 (10.5 – 16.8) kgP ha<sup>-1</sup> respectively for wheat, maize and rice. The ranking of potential uptake between different crops (maize  $\approx$  wheat  $\geq$  rice) is explained by the root biomass

potential of each crop, maize having the highest and rice the lowest. Wheat might be additionally biased, as our calculation concerns only 35% of the global observed wheat area which is concentrated in areas with high soil P levels (Northern Hemisphere). Fig. 2.2 shows the spatial variability of potential root P uptake, which does not vary much among crop species. The top 25% values consistently show up in Western Europe, South-East Asia and Central America corresponding to high production intensity areas. Interestingly, North America and Brazil show widespread low potential P uptake. This is explained by low levels of  $P_{ILAB}$  (SI Fig. B.3) and high Mollisol and Oxisol coverage, with high P sorption and hence low  $C_P$  (SI Eq. 9). Other areas that show low potential P uptake such as Africa, Eastern Europe, Argentina and Paraguay are characterized by low  $P_{ILAB}$  levels, likely due to negative soil P input/output balance (MacDonald et al., 2011).

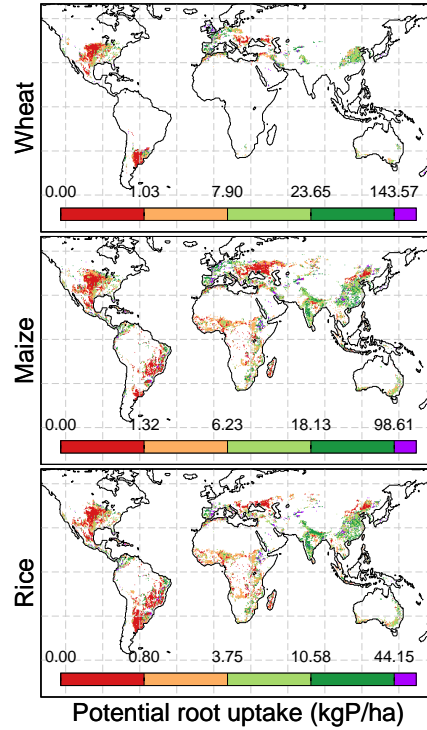


Figure 2.2: Potential root P uptake during the growing season ( $P_{uptake}$ ,  $\text{kgP ha}^{-1}$ ), determined as the median from 1000 MC replicates. The colorbar depicts quantiles (0-25-75, 90). Rows are wheat, maize and rice.

The spatial distribution of potential uptake closely follows that of  $P_{ILAB}$  (SI Fig. B.3). At least 80% of the spatial distribution of the potential root P uptake can be explained by the amount of  $P_{ILAB}$  (Fig. 2.3, top). In this figure, the commonality coefficient CC12 describes common variance to variables  $P_{ILAB}$  and  $C_P$ , whereas CC1 ( $P_{ILAB}$ ) and CC2 ( $C_P$ ) describe the additional variance not shared with the other variable. Since  $C_P$  is a function of  $P_{ILAB}$ , CC 2 shows the additional variability from the fitted Freundlich equation (Eq. 2.4), whereas CC 1 has no new information. Other variables are much less important than  $C_P$  or  $P_{ILAB}$  in determining the spatial distribution of potential uptake. This comes as no surprise, as high soil water content (consequence of the irrigated conditions)

presents no obstacles to P diffusion. Another point to note is the lack of influence of root length density ( $L_{rv}$ ) on the spatial distribution of rice P uptake (Figure 3, top, CC4). This is explained by the fact that the root system of rice achieves its biomass potential in most of the considered grid-cells, thus exhibiting no climate variability in our crop simulations.

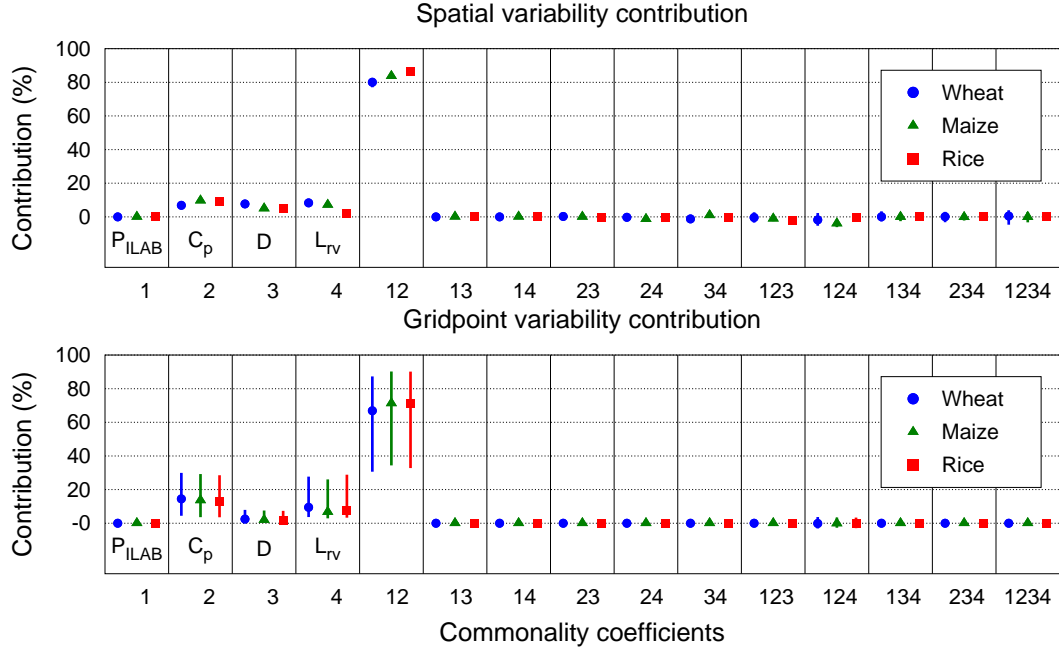


Figure 2.3: CA of potential P root uptake (Eq. 2.3). The plots show the contribution of each of the drivers ( $P_{ILAB}$ ,  $C_p$ ,  $D$ ,  $L_{rv}$ ) to the spatial variability (top) and variability at each grid point (bottom) of the calculated root P uptake. Values 1 to 4 show the unique contribution not shared with others. Other combinations (12 to 1234) show variance common to all variables within that combination. Different markers and colors correspond to different crop species (see legend). In the top plot, errorbars show the 95% uncertainty interval around the global median value (obtained with 1000 bootstrap samples). In the bottom plot, errorbars show the 68% quantile interval around the global median value.  $P_{ILAB}$  is soil inorganic labile P,  $C_p$  is P in soil solution,  $D$  is the P diffusion coefficient and  $L_{rv}$  is the root length density

Uncertainty at each grid point level is mainly due to the amount of plant available P or  $P_{ILAB}$  (Fig. 3, bottom) and is higher than variability due to other variables ( $C_p$ ,  $D$ ,  $L_{rv}$ ) in 70 – 80 % of the world grid points (not shown). The uncertainty shifts from  $P_{ILAB}$  to other variables in areas of low soil P levels such as North America, South America and Eastern Europe (Supp. Fig. 3).

### 2.3.2 Crop P demand

The values of the global average of crop P demand are given in Table 2.1. We found that the demand is larger for maize than for the two others crops (wheat and rice). In general, we found a relative consistency between the three methods except for rice, where the C:P ratios method leads to higher P demand compared to the PHI and QUEFTS methods. This could be explained by the fact that QUEFTS and PHI parameters are obtained in stressed environments where nutrient use efficiency is highest, while the C:P ratio parameters are

	Wheat	Maize	Rice
QUEFTS	6.6 (4.9-10.1)	14.6 (10.4-22.1)	5.5 (4.5-6.3)
PHI	7.8 (7.8-7.9)	19.0 (18.7-19.3)	6.8 (6.7-7.0)
C:P ratios	7.1 (4.6-9.9)	21.4 (18.3-24.7)	13.1 (12.2-14.0)

Table 2.1: Global average of P demand ( $\text{kgP ha}^{-1} \text{ yr}^{-1}$ ) calculated with different methods (QUEFTS, PHI and C:P ratios). Values in brackets are the 95% quantile interval around the median value from 1000 MC replicates.

obtained in conditions of luxury consumption (see sections Methods and Discussion).

There were some inconsistencies between the methods, especially for the C:P method. The deviation between C:P and the two other methods was larger for rice and could be explained by the difference in allocation between the crops. In ORCHIDEE-CROP simulations, grain receives about 35% of NPP in case of rice against 65 and 50% for wheat and maize, respectively. This tends to decrease further the contribution of grain in the demand estimate by the C:P method for rice. This is amplified by inter-species differences in organ P concentrations (SI Fig. B.2), which was mentioned earlier in the previous paragraph and is further presented in the discussion.

Comparing global average values of demand (Tab. 2.1) and potential root uptake (Section 2.3.1), we observed that the demand was not larger than the soil P supply. A different picture emerges though, if looking at the range of these values across croplands worldwide. If histograms of P supply and demand are plotted side by side (Fig. 2.4), they represent the case of most optimal redistribution of the two. In this case, only 15 – 40 % of wheat and maize area would achieve its potential yield using soil P only and 25 – 65 % in case of rice, which points to the inefficiency of P distribution in agricultural soils worldwide.

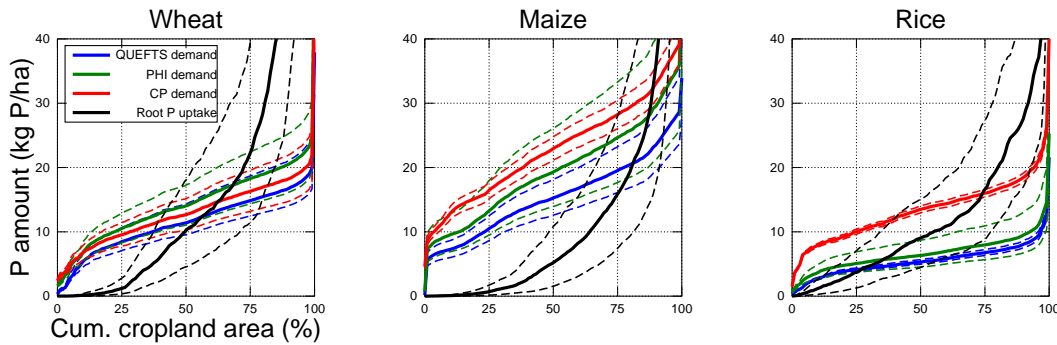


Figure 2.4: Cumulative histograms of crop P demand and potential root P uptake globally [ $\text{kg P ha}^{-1}$ ]. Demand is drawn with color, while potential root uptake is drawn in black. Each color represents a different way of calculating P demand (blue for QUEFTS, green for PHI and red for C:P ratios). Dashed lines correspond to the 68% quantile interval around the median of 1000 MC replicates.

Lastly, the difference in the shapes of the demand and potential uptake curves (Fig. 2.4) can be explained by their respective source of spatial variability. Since the crop P demand is calculated from the ORCHIDEE-CROP biomass pools in irrigated conditions,

the climate determines the yield potential and subsequently the distribution shape. For potential root P uptake, we have shown that 80% of the spatial variability is explained by the  $P_{ILAB}$  distribution (Section 2.3.1), which in turn depends mostly on the historical farm input/output balance and soil biogeochemical background (Ringeval et al., 2017).

### 2.3.3 P limitation and yield gap

The global average of the supply:demand ratio showed that the world soil P supply by itself is largely adequate in satisfying crop P demand. The three demand methods give an average of 160 (81 – 473) % in wheat, 132 (55 – 263) % in maize and 205 (69 – 544) % in rice. By adding P from fertilizer the global average increases to 182 (98 – 514) % in wheat, 155 (72 – 299) % in maize and 278 (96 – 624) % in rice. The differences among species are due to their geographic location as mentioned in section 2.3.1. Even though the addition of fertilizer increases the amount of available P, it is hard to distinguish the increase due to high uncertainty of the initial  $P_{ILAB}$  data.

A clearer picture of global P limitation is presented when looking at the spatial distribution of the supply:demand ratio (Fig. 2.5, focusing on the QUEFTS method; see SI Fig. B.5 for the PHI and C:P based ones). The soil supply only sustains the demand of P for potential yields in certain parts of the world like South-East Asia, Western Europe, Central America, Ethiopia and Oceania. By contrast, large parts of Central Asia, Africa, North and South America experience severe limitation that needs to be amended with additional inputs to increase yields. These areas have a known low soil P supply, due to high sorption capacity (Mollisols in North America, Oxisols in Brazil) and historically low to negative P balance (Ringeval et al., 2017) (Central Asia, Africa). When including fertilizer P from the current year in the supply, the limitation is somewhat relieved except in the western part of North America, Africa, Central Asia and southern South American states (Argentina, Paraguay and Uruguay).

The histogram of the yield gap is plotted in Fig. 2.6 (top row). We found that on average 52 – 57 % of wheat area, 69 – 79 % of maize area and 36 – 63 % of rice area shows a yield gap due to P limitation in some form. Differences between the three crops are mainly explained by their geographic location. Maize is the most limited crop because its potential biomass is higher in ORCHIDEE-crop, leading to an higher demand (Section 2.3.2) but also because it is grown all over the world, including in areas of low soil P supply (Central Asia and Africa). Wheat limitations are in between those from maize and rice because wheat is grown mostly in northern latitudes, where soil P supply is relatively higher. Rice shows the lowest P limitation because its production is highly concentrated in South-East Asia, an area with high levels of soil P supply (Fig. 2.2).

The overall uncertainty is large and comparable across species (except for rice) and across methods. Because the potential P uptake is common to all crops/methods, this points to the uptake as the main determinant of the yield gap uncertainty. Since the un-

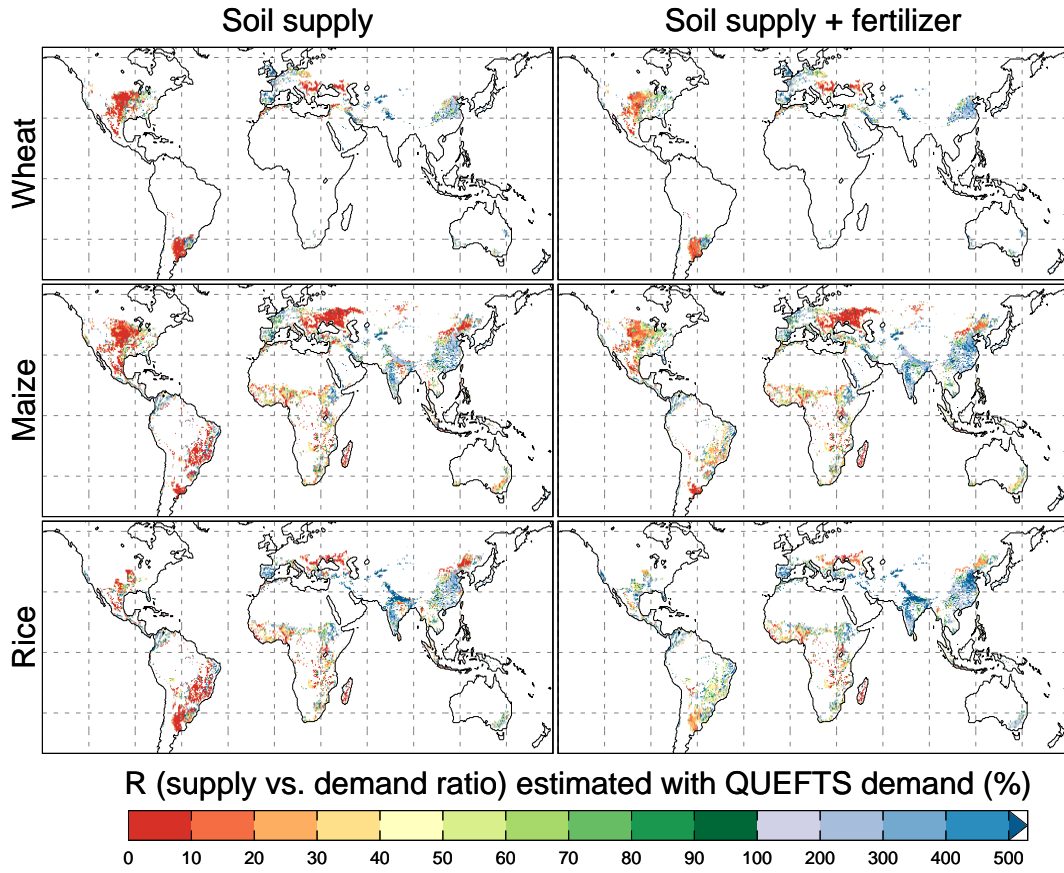


Figure 2.5: Spatial distribution of the  $R$  ratio (supply:demand; %) calculated using the QUEFTS method to estimate the demand. The ratio is provided when only the soil  $P$  supply is considered (case I in Eq.6 ; left column) as well as when both soil  $P$  supply and fertilizer  $P$  are considered (case II in Eq.6; right column). Rows are different species (top to bottom: wheat, maize, rice). SI Fig. B.5 show ratios for PHI and C:P methods.

certainty in the potential root  $P$  uptake is largely explained by  $P_{ILAB}$  (Fig. 2.3, bottom), this uncertainty can be traced back to the uncertainty of the main drivers in the  $P_{ILAB}$  pool, which are the history of farm inputs/outputs and the soil biogeochemical background (Ringeval et al., 2017). For rice, there is a notable difference between the three limitation methods (Fig. 2.6, top right) which is due to differences between the methods to estimate the demand (as explained in Section 2.3.2). Adding  $P$  from fertilizer to the potential root  $P$  uptake, the total area experiencing  $P$  limitation decreases to 43 – 46 % for wheat, 65 – 72 % for maize and 23 – 50 % for rice (dashed lines in Fig. 2.6, bottom row). But as mentioned earlier, the decrease in limitation is hardly noticeable due to variability of the initial  $P_{ILAB}$  data.

Based on our different estimates, the global average of the  $P$  yield gap is 22 (18–28) % in wheat, 55 (47–66) % in maize and 26 (18–46) % in rice. Amended with fertilizer  $P$ , the yield gap drops to 17 (14 – 21) % in wheat, 46 (36 – 55) % in maize and 15 (10 – 32) % in rice. Comparing our  $P$  yield gap estimates to the ones observed (Mueller et al., 2012) (black lines in Fig. 2.6, bottom), we found that  $P$  limitation explains only a portion of the

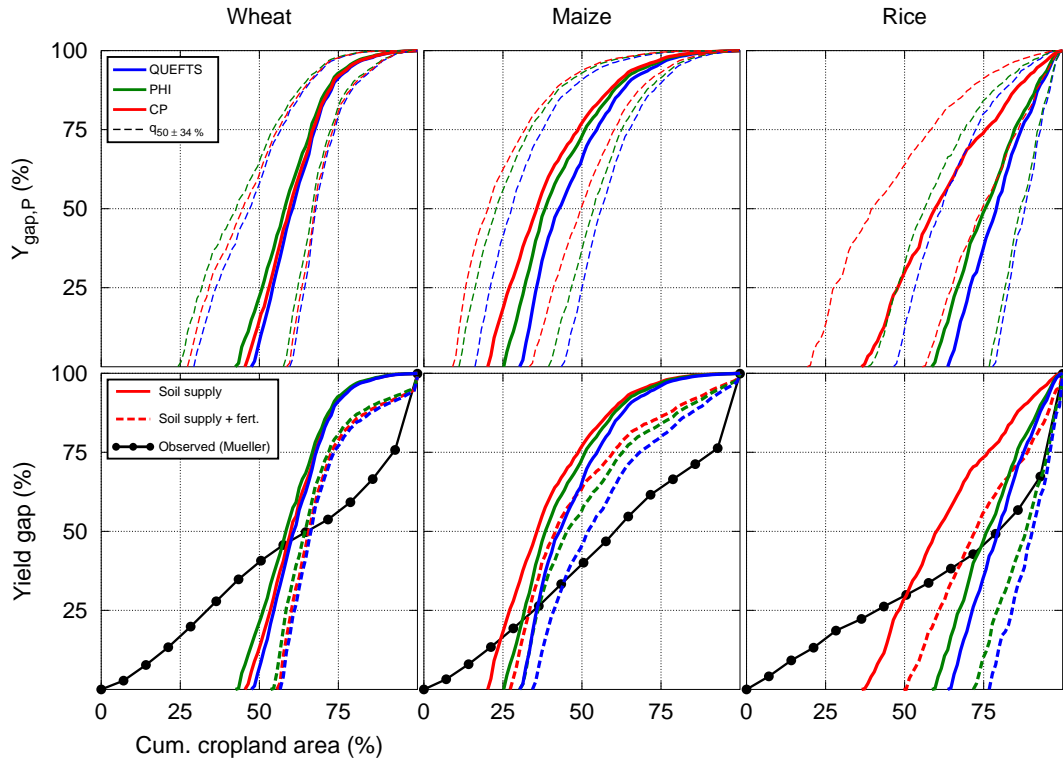


Figure 2.6: Cumulative histogram of the P yield gap in irrigated simulations (%). Upper figures show the global P limitation (due to potential root uptake P) calculated with different P demand methods. Full lines are the median of the 1000 MC replicates, with 68% quantile range around it in dashed lines. Lower figures show the limitation due to root uptake P (full colored line) amended with fertilizer P (dashed colored lines). Additionally, the yield gap from Mueller et al., 2012 (black line with circles) is plotted for comparison.

observed yield gap. The difference could be due to us considering only P as the limiting factor, whereas the observed yield gaps includes all of the macro-nutrients (NPK) as well as contributions from irrigation, weed, pests and disease.

## 2.4 Discussion

We have investigated how soil P supply can lead to P limitation of crop yields, by combining data on the current global distribution of agricultural soil P and an idealized crop simulation, while accounting for the basic nature of plant-soil transfer of P. Whereas previous studies estimate the global nutrient yield gap through statistical approaches (Mueller et al., 2012), our diagnostic approach aims to represent more explicitly key variables and processes involved in the P limitation, namely the cropland soil P, soil P diffusion and root uptake. We showed that P is a key factor of yield gap at the global scale with estimated global yield gap due to P of 22, 55 and 26 % for wheat, maize and rice, respectively. Although the global average of the potential soil P supply is larger than P demand, due to the amount of soil legacy P in highly fertilized areas, we found that a significant portion of cropland area has soil P levels too low to sustain the climate potential yield. These results

corroborate the heterogeneity of today's global P management practice and its potential impacts on food production (Mueller et al., 2012; MacDonald et al., 2011) which need to be dealt with a comprehensive set of global measures like: more equitable access to fertilizer P, P redistribution and reducing P waste through technological improvements (Obersteiner et al., 2013; Tilman et al., 2011)

Soil P availability plays a more important role than current fertilizer application, according to our results. A sensitivity test, where we assumed that 20% of the annual applied P in fertilizer is directly incorporated into crop biomass showed that P limitation can be relieved only slightly, with a global yield gap reduction of 5 – 10 %. This small alleviation suggests, as mentioned in the previous paragraph, that the spatial distribution of current fertilizer application is relatively non-optimal. Surprisingly, we found significant P limitation in intense production areas (e.g. North America, Brazil) even in our fertilizer sensitivity test. Even though these areas largely consist of high-fixing P soils (Roy et al., 2016), we suspect the P limitation stems from the nature of the models used. One explanation is the simplicity of the root uptake model, which does not include some key processes that increase P acquisition (Hinsinger et al., 2011) like root branching/architecture, exudates/phosphatase and mycorrhizae association. Another reason for the severe P limitation in high production areas could be the bulk nature of the soil P model, which homogeneously mixes the fertilizer P in the first 30 cm. In reality, plant nutrient uptake can be increased (G. W. Randall and Hoeft, 1988) through a practice called banding, whereby fertilizer is placed close to the roots to increase its efficiency, especially in soils with low P levels and high fixing capacity (G. W. Randall and Hoeft, 1988). This shows that simple bulk flow calculations might not be sufficient in diagnosing P limitation, and that a more process oriented description of P exchange is needed.

Global yield gaps have so far been studied (Licker et al., 2010; Mueller et al., 2012; Neumann et al., 2010) in a spatial manner using globally gridded yield census data (Monfreda, Ramankutty, and Foley, 2008) and relating it to crop management information. These provide a robust picture of current limitation in world's crop production, and can be a valuable knowledge source for targeted regional management. But the strength of these approaches, which is in dealing with aggregated effects of agricultural practices, is also a limit if one wants to take a closer look at each limiting factor. Irrigation and nutrient application are strongly related to yield variation globally (Mueller et al., 2012), but separating their contribution remains a problem, since these factors exhibit a high degree of correlation. A more process-based approach would enable us to isolate the effects of P limitation from others (like water and N). However, the use of a generic crop model gives estimates that are partly disconnected from reality, and makes the comparison with the observed yield gap difficult. The parametrization of crop models against rich data sources for local climate and soil conditions, for example the GYGA project (van Ittersum et al., 2013), would allow both distinction of the limiting factors and connection to a local context.

Our estimates contain considerable uncertainty, which is mainly due to the amount of plant available P in soil ( $P_{ILAB}$ ) variable. In the source of the  $P_{ILAB}$  information (Ringeval et al., 2017), the uncertainty stems from the soil’s biogeochemical history which is poorly constrained by the original data (X. Yang et al., 2013). Apart from this, there are notable differences across the three P demand methods (Tab. 2.1 and Figure 2.4). In PHI and QUEFTS, the estimates are based on traditionally measured quantities of aboveground nutrient off take, and provide comparable values as they incorporate data from a large set of growing environments (Duivenbooden, 1992; S. Sattari et al., 2014). Measurement of C:P ratios in plant organs (references given in SI Table B.2) does not seem to be a common practice in agriculture, given by the fact that we are using literature from year 1934 for maize (Latshaw and Miu, 1934). For rice, the data comes from an experiment in highly fertilized soil (average application of 240 kgN ha<sup>-1</sup>, 120 kgP<sub>2</sub>O<sub>5</sub> ha<sup>-1</sup> and 120 kgK<sub>2</sub>O ha<sup>-1</sup> annually; Ye et al., 2014) and leads to a notable P demand bias compared to the other two, possibly due to luxury consumption even in the nutrient omission plot. The scarcity of C:P ratio data in agriculture might serve as a call for more experimental inquiry, as ecosystem models move towards a description of the plant’s carbon cycle which include the effects of nutrient availability (Goll et al., 2017; Y. P. Wang, Law, and Pak, 2010; X. Yang et al., 2014).

This study is a step towards a process-based assessment of the P limitation in agricultural ecosystems. However, to arrive to these estimates, a number of broad assumptions and constraints had to be considered: missing representation of the plant response to P limitation, rudimentary description of P in soil solution, and linear reduction of yield according to the supply:demand ratio. Firstly, P limitation affects C allocation during plant development (A. Mollier and S. Pellerin, 1999), redirecting assimilates to roots instead of leaves to increase the plant’s P foraging ability. The consequence is a lower shoot biomass which demands less P, but also relatively greater roots with higher capacity for P uptake. Secondly, our description of the soil solution at an annual time-step is simple (Eq. 2.4), with a relation that describes the long term equilibrium between soil solution and sorbed P without any seasonal change; this limitation is evident in our treatment of fertilizer P (Eq. 2.6). When added to  $P_{ILAB}$  prior to the computing the  $C_P - P_{ILAB}$  equilibrium, the additional P has no effects on the simulated potential uptake (not shown). This points to the importance of transient and short-term effects in fertilizer P sorption (Barrow, 1983), a process that is highly dependent on soil composition and chemical properties. Finally, we calculated the yield gap as a linear decrease according to the supply:demand ratio, which is not an unreasonable simplification. Studies looking at the effects of nutrient deficiency show that (Amanullah and Inamullah, 2016; Batten, Khan, and Cullis, 1984; N. K. Fageria and Oliveira, 2014) the harvest index stays relatively stable at low P additions, with greater variability due to genetics and climate. However, this might not be true in extreme situations (Sahrawat et al., 1995), where allocation to grain might be disrupted to insure functioning of other plant parts.

The above-mentioned issues could be tackled by coupling the soil P dynamics (Ringeval et al., 2017) and the ORCHIDEE-CROP model (X. Wu et al., 2016). For example, a dynamic representation of P stress on a daily basis would allow us to simulate the interaction of P with other factors, notably water and N. Furthermore, physiological responses to P stress could be simulated such as a decrease in leaf area index (Plénet et al., 2000) or change in the shoot:root ratio (A. Mollier and S. Pellerin, 1999), which would influence the final grain yield and more faithfully describe yield loss. Exchange of P between the soil and crop through crop uptake/residues, fertilizer application and their incorporation into the soil, would allow us to model long-term effects of P management on crop yields. We surmise that, by properly accounting for the flow of P between the soil and plant, more accurate estimates and a better understanding of P limitation in agricultural ecosystems is possible, which are needed in the face of the growing phosphorus issue worldwide (Cordell, Drangert, and White, 2009; Obersteiner et al., 2013; Peñuelas et al., 2013).

## Chapter 3

# Carbon and phosphorus allocation in annual plants: an optimal resource use approach

Marko Kvakić, George Tzagkarakis, Sylvain Pellerin, Philippe Ciais, Daniel Goll, Alain Mollier, Bruno Ringeval

To be submitted to *Frontiers in plant science*

## Abstract

Phosphorus (P) is the second most important nutrient after nitrogen (N) and can greatly diminish plant productivity if P supply is not adequate. Plants respond to soil P availability by adjusting root biomass to maintain uptake and productivity due to P use. In spite of our vast knowledge on P effects on plant growth, how to functionally model enhanced root biomass allocation in low P environments is not fully explored.

We develop a dynamic plant model based on the principle of optimal carbon (C) and P allocation to investigate growth and functional response to contrasting levels of soil P availability. By describing plant growth as a balance of growth and respiration processes, we optimize C and P allocation in order to maximize leaf productivity and drive plant response. We compare our model to a field trial and a set of hydroponic experiments which describe plant response at varying P availabilities.

The model is able to reproduce long-term plant functional response to different P levels like change in root-shoot ratio (RSR), total biomass and organ P concentration. But it is not capable of fully describing the time evolution of organ P uptake and cycling within the plant. Most notable is the underestimation of organ P uptake during the vegetative growth stage which is due to the model's leaf productivity formalism.

In spite of the model's parsimonious nature, which optimizes for and predicts whole plant response through leaf productivity alone, the optimal growth hypothesis can provide a reasonable framework for modelling plant response to environmental change that can be used in more physically driven vegetation models.

### 3.1 Introduction

Plants respond to low P supply by growing more roots (Marschner, Kirkby, and Cakmak, 1996) as they are responsible for its uptake. Depressing the shoot C allocation and preferential partitioning of assimilates to roots is a well documented phenomenon at low P availability (Fredeen, Rao, and Terry, 1989; Rychter and D. D. Randall, 1994). The importance of nutrient control on C allocation between shoots and roots is even more clear under magnesium (Mg) and potassium (K) limitation (Cakmak, Hengeler, and Marschner, 1994) since they have a direct role in assimilate transport from leaves and can disrupt the plants' mechanism to counter reduced P availability. Once P supply is limited, plants exhibit lower leaf surface (Fredeen, Rao, and Terry, 1989) and reduced photosynthetic capacity (Fredeen et al., 1990) all of which contribute to reduction of productivity and total plant biomass. Another important point is the stoichiometric flexibility of plant tissue (J. J. Elser et al., 2010) as a response to environmental change; especially to change in nutrient availability. Since plant size and nutrient concentration are strongly connected due to the underlying machinery that drives growth (G. I. Ågren, 2008) representing stoichiometry change is a crucial step when coupling the C and nutrient cycles.

One of the main hypotheses explaining plant response to nutrient, CO<sub>2</sub> and water availability is that plants make optimal use of resources in order to maximize growth (Bloom, Chapin, and Mooney, 1985). Resources are acquired by the plant and distributed to its organs, all of which serve a different function (eg. leaves for C assimilation, roots for nutrient uptake, stems for structural support). Investing into an organ will increase its capacity to perform a certain function, but will necessarily incur a cost as resources are devoted to its formation and upkeep. The plant should thus invest into an organ and maximize the organ's efficiency, which is the amount of functional gain per amount of resource invested. If we assume individual plants grow as much as they can in order to survive, there should be an optimal way to distribute resources in order to do so. Applying this principle to the problem of nutrient limitation, a plant should grow roots (which take up nutrients) in such a way that uptake efficiency is highest (or the least amount of roots needed to satisfy the growth requirements). Consequently, at different nutrient availabilities the plant will exert more or less effort in growing roots and thus change its root-shoot balance.

The question of how to model C allocation due to nutrient limitation is a long standing one (Ågren and Wikström, 1993; John H. M. Thornley, 1995; Franklin et al., 2012). Optimal resource distribution as an approach would fall in between functional balance and teleonomic ones. Functional balance states that allocation of C towards an organ will be driven by its function (C assimilation in leaves, P uptake in roots) whereas teleonomic states that allocation is guided with a specific goal in mind, which is maximizing productivity and fitness. Optimal function approaches have shown promise in explaining plant response to a changing environment (McMurtrie et al., 2008; Mäkelä, Valentine, and Helmisaari, 2008;

Franklin et al., 2009; Dewar et al., 2009). By describing plant growth as a balance of growth and respiration processes, the concept of optimizing (or maximising) productivity provides a consistent framework for describing plant response to nutrient availability (Dewar et al., 2009). Furthermore, some avenues to improve this approach are proposed (Dewar et al., 2009) such as: extension to different time-scales, inclusion of multiple resources (energy, water, nutrients, etc.) and developing practical methods to be included in dynamic global vegetation models (DGVM).

Previously mentioned optimal function approaches (Dewar et al., 2009) deal with steady state vegetation growth. Even though this is a robust representation of ecosystem response, a dynamical representation of the underlying processes (via a DGVM) is warranted as vegetation growth is inherently seasonal. One of the main goals of DGVMs (Krinner et al., 2005) is to bridge the gap between the fast (order of 1 hour) hydrologic and biophysical processes of water and energy exchange on one hand and slow (order of 1 year and more) vegetation dynamics like fire, sapling establishment, light competition, tree mortality on the other. This is achieved with the use of physical parametrizations of C exchange through photosynthesis (Farquhar, von Caemmerer, and Berry, 2001; Collatz, Ribas-Carbo, and Berry, 1992), stomatal conductance (Ball, Woodrow, and Berry, 1987) and respiration models (Parton, J. W. B. Stewart, and Cole, 1988; Ruimy, Dedieu, and Saugier, 1996) which drive vegetation growth and its interaction with the biosphere on a daily time step.

The optimal approaches mentioned so far (Dewar et al., 2009) have focused on nitrogen (N) only. This is understandable as N is the principal nutrient required for plant structure and metabolism functioning. Here we focus on P because its effects on plant productivity can extend to the eco-system level, where it has been shown (James J. Elser et al., 2007) that P has a similar potential to N across terrestrial biomes. Even more, P has the highest potential of all major nutrients (Barber, 1995) to affect plant investment into roots (and subsequently whole plant response) due to the physical mechanism of soil P transfer via diffusion. Representing the effects of P limitation is thus as important as N if plant response to a changing environment is to be investigated.

To this end, we propose a model which optimizes plant growth according to P availability as a balance of leaf productivity, root P uptake and organ respiration on a daily time step. The goal is a consistent description of plant response like change in root-shoot ratio, stoichiometry and total biomass. We present an optimization tool (Dantzig and Thapa, 1997) which allows for dynamic (day-by-day) modelling and could be extended to include other growth limiting resources, as well as implemented within a more physically based vegetation model (Krinner et al., 2005)

## 3.2 Methods

### 3.2.1 Modelling framework

We assume C and P are the main currency for biomass growth. Leaves assimilate C and roots take up P on a daily basis. Although C is the main resource for biomass growth, P is needed to sustain the plant's metabolism. Assuming that individual plants try to grow as much as they can, a balance between the amount of leaf and root should be established according to P availability. The problem can be translated into an optimal resource use one: given a limited supply of C and P, how can they be distributed to grow the plant as much as possible? To solve this problem we employ linear programming (Dantzig and Thapa, 1997) a method which calculates the best possible outcome given a problem stated by a system of linear equations. Linear programming is a well established theory used in many areas of operational research (Dantzig and Thapa, 1997) where performance of a system is maximised (like profit or units produced) given a limited amount of resources (like construction material, capital, labour or time).

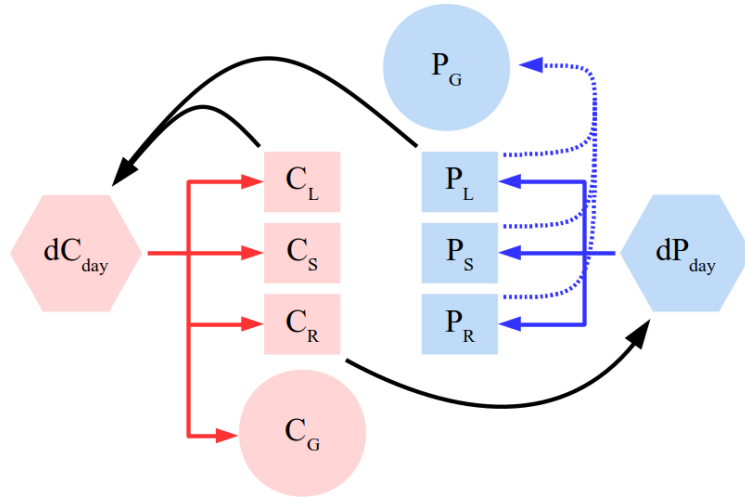


Figure 3.1: Simplified schema of the plant model. Full arrows depict C (red) and P (blue) flux directions. Black color depicts feedbacks of different pools on gross productivity ( $dC_{day}$ ) and root P uptake ( $dP_{day}$ ). Different letters correspond to leaf (L), root (R), stem (S) and grain (G). Dashed blue line represents grain P remobilisation. The model additionally contains C respiratory fluxes and allometric constraints, which were not depicted here for clarity sake and are detailed in the Methods section.

The plant is modeled as a system of linear mass-balance equations (model schema in Fig. 3.1) describing assimilation, root uptake and respiration. The plant consists of several organs: roots, leaves, stem and grain. Leaves take up C and roots take up P on a daily basis. We allocate C and P to grow plant organs as much as possible while respecting constraints in form of allometric relationships. For root P uptake, the model does not describe external physical mechanisms that determine P supply like soil diffusion. Instead, P availability is determined by the root P uptake rate which is simply the amount of P taken up per unit of root biomass per time. Stated as a linear program, the maximisation objective at each

time-step is following:

$$\text{Maximize } dC_{day} \quad (3.1)$$

Where  $dC_{day}$  is daily gross productivity, which is the amount of new C that can be allocated to different pools.  $dC_{day}$  is proportional to leaf biomass ( $C_L^*$ ) via the assimilation rate ( $k_{CL}$ ) and is weighted by the leaf's light ( $LUE$ ) and phosphorus ( $PUE$ ) use efficiency:

$$dC_{day} = k_{CL} \cdot dt \cdot C_L^* \cdot LUE \cdot PUE \quad (3.2)$$

$$LUE = \frac{C_{L,max}}{C_L^* + C_{L,max}} \quad (3.3)$$

$$PUE = \frac{P_L^*}{P_L^* + C_L^* \cdot \rho_{max}} \quad (3.4)$$

$LUE$  represents saturated assimilation via leaf shading (J. H. M. Thornley, 1976) whereas  $PUE$  represents saturated photosynthetic capacity due to leaf P concentration (Fredeen et al., 1990).  $C_{L,max}$  and  $\rho_{max}$  are the leaf dry weight and P concentration at which  $LUE$  and  $PUE$  reach half of their maximum efficiency, respectively. Equations 3.2 - 3.4 are a standard hypothesis in optimal function approaches (Dewar et al., 2009) and are the main driver of plant response to P availability.  $C_L^*$  and  $P_L^*$  are future leaf C and P pools, respectively. Other pools which exist, but are not photosynthetically active are roots (R), stems (S) and grain (G). Future pool refers to updating the current one at every time-step. For C, updated pools are a balance (Eq.3.5) between C allocation (Eq. 3.6) and respiration (Eq.3.7):

$$C_i^* = C_i + dC_i - rC_i \quad (3.5)$$

$$dC_{day} = \sum_{i=L,R,S,G} dC_i \quad (3.6)$$

$$rC_i = \lambda_{C,i} \cdot C_i \cdot dt \quad (3.7)$$

The variables here are  $C_i^*$  (future C pool),  $dC_i$  (amount of C allocated) and  $rC_i$  (respired C).  $C_i$  (current C pool) is a state variable and  $\lambda_{C,i}$  (respiration rate) is time-invariant. Index  $i$  denotes different organs. In the case of P, future pools are a balance (Eq.3.8) of root P uptake (Eq.3.9), P allocation (Eq.3.10) and P remobilisation (Eqs.3.11 and 3.12):

$$P_i^* = P_i + dP_i - mP_i \quad (3.8)$$

$$dP_{day} = k_{CR} \cdot C_R^* \quad (3.9)$$

$$dP_{day} = \sum_{i=L,R,S} dP_i \quad (3.10)$$

$$mP_i = k_{mP_i} \cdot P_i \cdot dC_G; \quad i = L, R, S \quad (3.11)$$

$$\sum_{i=L,R,S,G} mP_i = 0 \quad (3.12)$$

The variables are:  $P_i^*$  (future P pool),  $dP_i$  (P increment) and  $mP_i$  (remobilised P).  $P_i$  (current P pool) is a state variable while  $k_{CR}$  (root P uptake rate) and  $k_{mP_i}$  (remobilised P fraction) are time-invariant. We assume no saturating effect for root uptake (Eq. 3.9) to reduce model complexity, assuming it can be due to external physical limitations to P supply (eg. soil diffusion) which are not modeled here. P remobilisation flux (Eq. 3.11) is proportional to the amount of C going towards the grain ( $dC_G$ ) and the fraction of the remobilisable P ( $k_{mP_i} \cdot P_i$ ). Allometric constraints are applied on stem and grain C filling (Eq. 3.13, 3.14) as well as on P filling of photosynthetically non-active tissue (Eq.3.15):

$$f_{C_S} \cdot dC_L = dC_S \quad (3.13)$$

$$f_{C_G} \cdot rC_S = dC_G \quad (3.14)$$

$$f_{P_i} \cdot dP_L = dP_i; \quad i = R, S \quad (3.15)$$

$f_{C_S}$  and  $f_{C_G}$  are the fraction of C going towards stem and grain, and are tied to the amount of C allocated to leaf ( $dC_L$ ) and stem respired C ( $rC_S$ ) respectively.  $f_{P_i}$  is the fraction of P going towards the photosynthetically non-active tissue (root and stem) and is tied to the amount of leaf allocated P ( $dP_L$ ). Equation 3.13 follows the principle of the pipe-theory model (Shinozaki et al., 1964). For grain (Eq. 3.14) we rely on the concept of Iwasa and Roughgarden (1984) where grain filling is triggered once a plant reaches its maturity. In our approach, we model grain filling as a continuous process where the grain C flow reaches its peak when the vegetative part stops growing (or stem respiration is highest). Equation 3.15 is a necessary assumption if we want to fill the photosynthetically non-active pools with P, since the plant does not confer any benefit from doing so. Details on the various variables and parameters are given in Table 3.1.

Table 3.1: Model variables and parameters. Units correspond to values that would be expected in an agricultural field trial. For the hydroponic studies, the units change according to the scale of the experiment to  $\text{gDW pot}^{-1}$  for biomass and  $\text{mgP pot}^{-1}$  for  $P_i$ .  $i$  denotes organs which are leaves (L), roots (R), stem (S) and grain (G). DW stands for dry weight

Variable	Description	Unit
$dC_{day}$	Daily gross productivity	$\text{tDW ha}^{-1}$
$dP_{day}$	Daily root P uptake	$\text{kgP ha}^{-1}$
$C_i$	Current C pool	$\text{tDW ha}^{-1}$
$P_i$	Current P pool	$\text{kgP ha}^{-1}$
$C_i^*$	Updated C pool	$\text{tDW ha}^{-1}$
$P_i^*$	Updated P pool	$\text{kgP ha}^{-1}$
$dC_i$	Daily C allocation	$\text{tDW ha}^{-1}$
$rC_i$	Daily C respiration	$\text{tDW ha}^{-1}$
$dP_i$	Daily P allocation	$\text{kgP ha}^{-1}$
$mP_i$	Daily grain remobilised P	$\text{kgP ha}^{-1}$
Constant		
$dt$	Time step	day
$k_{CL}$	Assimilation rate	$\text{tDW tDW}^{-1} \text{ leaf day}^{-1}$
$k_{CR}$	Root P uptake rate	$\text{kgP tDW}^{-1} \text{ root day}^{-1}$
$C_{L,max}$	LUE half-saturation point	$\text{tDW leaf ha}^{-1}$
$\rho_{L,max}$	PUE half-saturation point	$\text{kgP tDW}^{-1} \text{ leaf}$
$f_{C,i=S,G}$	C allocation fraction	$\text{tDW tDW}^{-1} \text{ leaf}$
$\lambda_{C,i=L,R,S,G}$	C respiration rate	$\text{day}^{-1}$
$f_{P,i=R,S}$	P allocation fraction	$\text{kgP kgP}^{-1} \text{ leaf}$
$k_{mP,i=R,L,S}$	Grain P remobilisation fraction	$\text{tDW}^{-1} \text{ grain day}^{-1}$

### 3.2.2 Calibration : observations and protocol

We use the following datasets to calibrate our model with observations : a field trial and two hydroponic studies. The field trial contains information on maize (*Zea mays*) shoot biomass and organ P uptake during a whole growing season in 1996 (with an interval of 7 to 20 days) and comes from a long-term experiment in Tartas, France (Pl  net et al., 2000) where maize response to different P input levels was recorded (3 levels, 4 replicates). The hydroponic studies describe early vegetative response (within 3 to 4 weeks) of shoot biomass, root biomass and shoot P uptake across five different P levels at experiment end. They contain 7 temperate annual pasture species (Asher and Loneragan, 1967) and 5 cereal/legume ones (N. K. Fageria and Baligar, 1989). The pasture species are subterranean clover (*Trifolium subterraneum*), barrel medic (*Medicago tribuloides*), blue lupin (*Lupinus digitatus*), smooth flatweed (*Hypochoeris glabra*), erodium (*Erodium botrys*), capeweed (*Cryptostemma calendula*), silver grass (*Vulpia myuros*) and brome grass (*Bromus rigidus*). The cereal/legume species are alfalfa (*Medicago sativa*), red clover (*Trifolium pratense*), common bean (*Phaseolus vulgaris*), rice (*Oryza sativa*) and wheat (*Triticum aes-*

*tivum*).

When calibrating, we use all of the available information to constrain the model parameters (time and P level wise). This entails 13 time points across 12 different P experiments (3 P levels x 4 blocks) in the field trial and only 1 time point across 5 different P experiments for each hydroponic study. During calibration we assume the root P uptake rate  $k_{CR}$  to be the only parameter changing among the different P experiments as it reflects P availability. All of the other parameters are kept constant across the different P experiments. In this manner the obtained values should be specific to the investigated species and the environment they were grown in, where plant response depends on P availability alone. Since we are concerned with P only, the calibrated values should correspond to growth that is equally limited by other major abiotic factors (like water, light or N) across all different P experiments. For the hydroponic studies we assume a plant consisting of only leaves and roots without allocation of P to roots. This is because aboveground biomass is grouped into shoots in both hydroponic studies (Asher and Loneragan, 1967; N. K. Fageria and Baligar, 1989) and no root P concentration is given in N. K. Fageria and Baligar (1989).

All of the parameters were kept constant during integration and calibrated in order to minimize the prediction error until the incremental parameter change falls below 1%. To start the calibration process we provide initial guess values based on maize, and assume these are the same when calibrating other species. To calculate the assimilation rate  $k_{CL}$  we assume leaf specific leaf area (SLA) to be around  $150 \text{ cm}^2 \text{ g DW}^{-1}$  based on our observation dataset, and refer to Kim et al. (2006) for the net assimilation rate  $A_{max} = 50 \mu \text{ mol CO}_2 \text{ m}^{-2} \text{ s}^{-1}$ . This results in  $k_{CL}$  of approximately  $1.5 \text{ gDW gDW}^{-1} \text{ leaf day}^{-1}$  if we assume dry biomass consists of 50% C. To calculate the root P uptake rate we use data from André et al. (1978) where average final root dry weight is  $18 \text{ gDW plant}^{-1}$  and total P taken up  $1.71 \text{ gP plant}^{-1}$  over a period of 100 days, which results in  $k_{CR}$  of approximately  $1 \text{ mgP gDW}^{-1} \text{ root day}^{-1}$ . LUE half-saturation leaf biomass  $C_{L,max}$  was estimated from Lindquist et al. (2005) where the average Beer-Lambert extinction rate  $k$  was measured at  $0.67 \text{ LAI}^{-1}$ . If we assume half of the light intensity is lost at same LAI ( $\text{LAI}_{LUE=1/2} = \frac{\log(2)}{k} = C_{L,max} \cdot \text{SLA}$ ) it gives a  $C_{L,max}$  of around  $1.5 \text{ tDW leaf ha}^{-1}$  with the previously mentioned SLA. PUE half-saturation leaf concentration  $\rho_{L,max}$  was estimated from Jacob and Lawlor (1991) where the majority of P limitation happens in the  $1.8 - 7.2 \text{ mmolP m}^{-2}$  range. This translates to  $0.8 - 3.2 \text{ mgP gDW}^{-1}$  with previously mentioned SLA so we assume a  $\rho_{L,max}$  value of  $1 \text{ mgP gDW}^{-1}$ . For plant respiration we rely on information on a whole plant basis (André et al., 1978; Vries, Witlage, and Kremer, 1979; Vries, 1972) which aggregates contributions due to growth and maintenance respiration, since we do not distinguish them neither. It can be found that maize respiration rates fall in the  $0.2 - 0.3 \text{ day}^{-1}$  range in optimal growing conditions so we assume :  $0.30 \text{ day}^{-1}$  for roots,  $0.10 \text{ day}^{-1}$  for leaves,  $0.03 \text{ day}^{-1}$  for the stem and  $0.01 \text{ day}^{-1}$  for the grain. Since concentration and final mass are a product of growth and respiration and are not easily

described by C:P ratios, we manually adjust the parameters related to allocation of C and P towards stem and root as well as grain P remobilisation, until the modeled plant somewhat resembled maize growth in our observation dataset. Afterwards we rely on the calibration procedure to pinpoint the parameter values.

Table 3.2: Parameter initial guess values. For the hydroponic studies, the units change according to the scale of the experiment to  $\text{gDW pot}^{-1}$  for biomass and  $\text{mgP pot}^{-1}$  for P. The asterix symbol (\*) signifies the parameter was not used during calibration of hydroponic studies. DW stands for dry weight.

Variable	Value	Unit	Description	Reference
$dt$	0.1	day	Time step	-
$k_{CL}$	1.5	$\text{tDW tDW}^{-1} \text{ leaf day}^{-1}$	Assimilation rate	Kim et al., 2006
$k_{CR}$	1.0	$\text{kg P tDW}^{-1} \text{ root day}^{-1}$	P uptake rate	André et al., 1978
$C_{L,max}$	1.5	$\text{tDW leaf}$	LUE half-saturation biomass	Lindquist et al., 2005
$\rho_{L,max}$	1.0	$\text{kgP tDW}^{-1} \text{ leaf}$	PUE half-saturation concentration	Jacob and Lawlor, 1991
$\lambda_{CR}$	0.30	$\text{day}^{-1}$	Root respiration rate	André et al., 1978 Vries, Witlage, and Kremer, 1979 Vries, 1972
$\lambda_{CL}$	0.10	$\text{day}^{-1}$	Leaf respiration rate	"
$\lambda_{CS}$	0.03*	$\text{day}^{-1}$	Stem respiration rate	"
$\lambda_{CG}$	0.01*	$\text{day}^{-1}$	Grain respiration rate	"
$f_{CS}$	0.5*	$\text{tDW stem tDW}^{-1} \text{ leaf}$	Stem C filling fraction	Manually adjusted
$f_{CG}$	1.0*	$\text{tDW grain tDW}^{-1} \text{ leaf}$	Grain C filling fraction	Manually adjusted
$f_{PR}$	0.1*	$\text{kgP root kgP}^{-1} \text{ leaf}$	Root P filling fraction	Manually adjusted
$f_{PS}$	0.8*	$\text{kgP stem kgP}^{-1} \text{ leaf}$	Stem P filling fraction	Manually adjusted
$k_{mP_i}$	0.05*	$\text{tDW}^{-1} \text{ grain day}^{-1}$	P remobilisation coefficient	Manually adjusted
$C_{L,0}$	0.1	$\text{tDW leaf ha}^{-1}$	Initial leaf biomass	Manually adjusted
$P_{L,0}$	0.1	$\text{kgP leaf ha}^{-1}$	Initial leaf P	Manually adjusted

Table 3.2 provides initial guess values and references for additional clarity. To calibrate the parameters we used Scipy's optimize package (E. Jones, Oliphant, and Peterson, 2001). To solve the linear programming problem we linearise equations 3.2 - 3.4 and use the package CVXOPT (Andersen, Dahl, and Vandenberghe, 2018). We integrate the model 160 days for the field trials and 30 days for the hydroponic studies. The calibrated values are compared to our initial guess estimates in Table 3.3.

As we do not describe the soil-plant mechanisms leading to P limitation, we connect the modeled root P uptake rate ( $k_{CR}$ ) to the observed soil solution concentration ( $C_P$ ) after the calibration. This is done using a Michaelis-Menten kinetic (3.16) and results are provided in the Supplementary Information.

$$k_{CR} = a_{C_P} \cdot \frac{C_P}{b_{C_P} + C_P} \quad (3.16)$$

### 3.2.3 Sensitivity analysis of the modeled response

To provide a sense of each parameter's impact on the modeled outputs we perform Sobol's method (Sobol, 1993) which is a global and model independent variance based sensitivity

analysis (Nossent, Elsen, and Bauwens, 2011). Here, the output model variance ( $V$ ) is decomposed into contributions due to each parameter input ( $V_i$ ) and their interactions with others ( $V_{ij}$ ) which allows the calculation of Sobol indices:

$$S_i = \frac{V_i}{V}; \quad S_{ij} = \frac{V_{ij}}{V}; \quad S_{Ti} = S_i + \sum_{i \neq j} S_{ij} + \dots \quad (3.17)$$

These are ratios of partial to total variance due to the parameter's main effect  $S_i$  (or first order index), its interactions  $S_{ij}$  (or second order index) or the sum of all of them together  $S_{Ti}$  (or total index). For a detailed derivation of these indices please refer to Sobol (1993) and Nossent, Elsen, and Bauwens (2011). In our case the modeled outputs were total plant biomass, total plant concentration and the RSR at simulation end. Sobol sensitivity analysis was performed with the SALib package (Herman and Usher, 2017) in the 10% - 200% range of the initial guess values (Table 3.2) using 170 000 samples.

### 3.3 Results

#### 3.3.1 Modeled plant response

The model is able to reproduce the typical growth pattern (J. H. M. Thornley, 1976) in annual plants : an early exponential growth transitioning into steady state at maturity, when grain filling takes place and P remobilisation occurs (Fig. 3.2, A and B). The evolution of organ concentration is related to the dynamic of leaf LUE and PUE (Eqs. 3.3 and 3.4) which is reflected in the evolution of the root-shoot ratio (or RSR, Fig. 3.2, C). There are two distinct phases separated by the maximum growth point around 30 days. The plant initially starts to grow P poor leaves, as the biomass gain due to LUE exceeds the PUE one. As LUE decreases with self-shading, roots are progressively grown more to provide P and sustain growth. After the maximum growth point around 30 days, the plant decreases its RSR as the already accumulated leaf P can support additional growth. We provide the dynamic of changing leaf LUE and PUE in Supplementary Fig. B.8 for more clarity.

Towards maturity, the RSR never falls to zero due to grain P remobilization from leaf. The leaf P remobilisation induces a decrease in PUE and is compensated by maintaining roots, which provide additional P to prevent it from decreasing further. This has the effect of increasing total plant concentration due to the additional P flux into the whole plant. Roots exhibit high P concentration towards maturity (Fig. 3.2, D) because accumulation of P in root is driven by the amount of P going towards the leaf (Eq. 3.15) and not the root C balance. This means that, since the plant accumulates enough P during its lifetime to support leaf PUE, the decrease in root growth after the maximum growth point will increase its concentration rapidly (Fig. 3.2, D). The "jagged" nature of C allocation (best seen in the RSR, Fig. 3.2 C) is due to the linear approximation of daily productivity (Eq. 3.2) and the optimization algorithm, which chooses to instantaneously ( $dt = 0.1$  day) grow

either leaf or root.

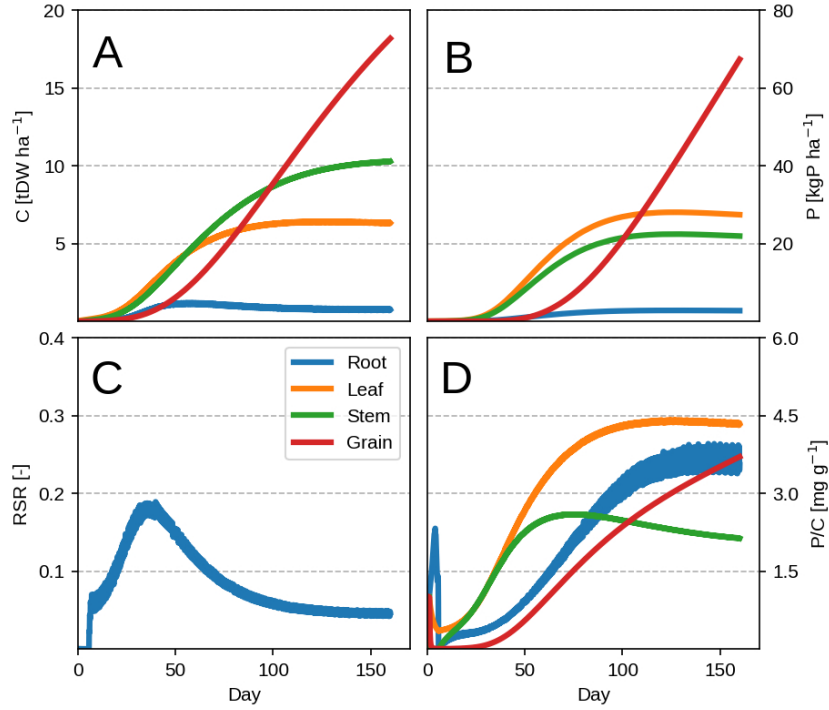


Figure 3.2: Modeled plant response with initial guess values (Table 3.2). Plots show the time evolution at high P availability ( $k_{CR} = 1.0 \text{ kgP tDW}^{-1} \text{ root ha}^{-1} \text{ day}^{-1}$ ). Upper row depicts C pools (left) and P pools (right). Lower row depicts the RSR (left) and organ concentration (right). Colors depict different plant organs.

Lowering the P availability (or the root P uptake rate  $k_{CR}$ ) has the effect of decreasing total plant biomass as more C is devoted to the roots (seen by the increasing RSR in Fig. 3.3, C) and a decrease in leaf PUE. Across different  $k_{CR}$  values, the RSR might seem lower than what is usually observed (0.2 - 0.8 in Amos and Walters 2006) as  $C_R$  represents active uptake root tissue. This could be improved by adding a separate passive root C pool with a lower respiration rate that would bring down total root concentration (Fig. 3.3, D)

### 3.3.2 Sobol sensitivity analysis

Looking at parameter influence on final plant biomass (Fig. 3.4, A, red) we can see assimilation ( $k_{CL}$ ) and self-shading ( $C_{L,max}$ ) have the most direct effect, as they determine the amount of C available for growth. This fact can be also be seen in the individual interaction terms (Fig. 3.4, B) where  $k_{CL} \times C_{L,max}$  stands out from the rest. Other contributions come mostly from interactions by organ respiration rates ( $\lambda_{Ci}$ ), C allocation fractions ( $f_{Ci}$ ), root P uptake rate ( $k_{CR}$ ) and leaf P demand ( $\rho_{L,max}$ ). For  $\lambda_{Ci}$  and  $f_{Ci}$ , allocation of C to various organs determines the amount of lost C due to respiration; a fixed ratio for stem and grain, and a variable one for root (depending on  $k_{CR}$  and  $\rho_{L,max}$ )

Looking at parameter influence on final plant concentration (Fig. 3.4, A, blue) assimi-

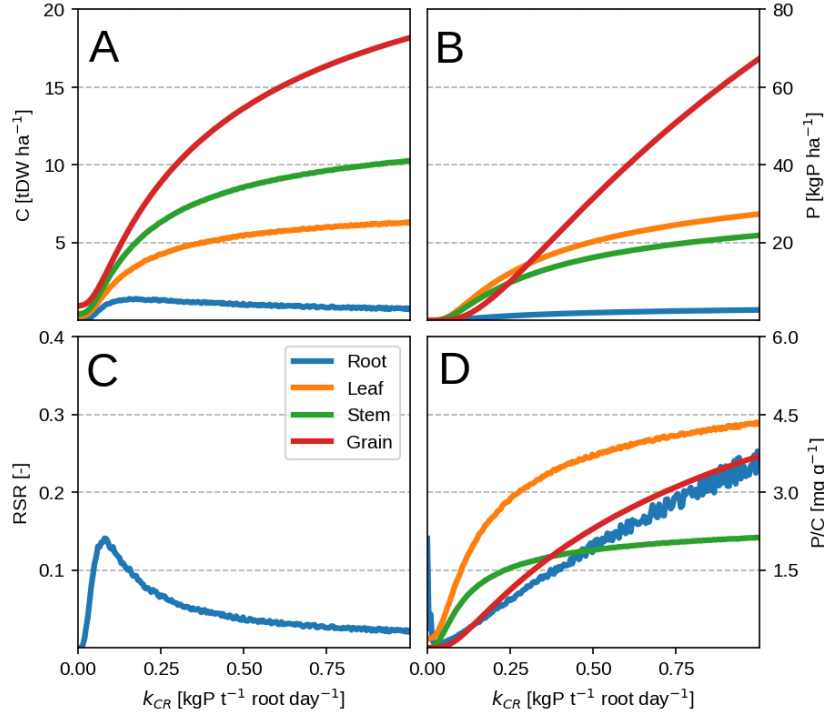


Figure 3.3: Modeled plant response with initial guess values (Table 3.2). Plots show modeled values at season end as function of root P uptake rate. Upper row depicts C pools (left) and P pools (right). Lower row depicts the RSR (left) and organ concentration (right). Colors depict different plant organs.

tion ( $k_{CL}$ ) and leaf respiration ( $\lambda_{CL}$ ) exert the most direct effect. Increasing  $k_{CL}$  increases total plant concentration as more C is available for root growth, which in turn increases the flux of P towards leaf and the whole plant. Increasing  $\lambda_{CL}$  brings down total plant concentration on the other hand, since a higher respiration rate both raises leaf P concentration and lowers the amount of C available for root growth, in turn lowering the P flux to the leaf and ultimately the whole plant. The other parameters responsible for final plant concentration are root P uptake ( $k_{CR}$ ), leaf P demand ( $\rho_{L,max}$ ), C allocation fractions ( $f_{Ci}$ ), root respiration rate ( $\lambda_{CR}$ ) and leaf remobilisation rate ( $k_{mPL}$ ). For  $k_{CR}$  and  $\rho_{L,max}$  this is not surprising, as increasing P availability and leaf P demand increases the overall flux of P into the whole plant. For  $f_{Ci}$ , C allocation to non-root organs lowers the P flux into the plant as well concentration due to enhanced growth. For  $\lambda_{CR}$ , it determines the amount of P gained per C lost to root respiration. For  $k_{mPL}$ , it is due to leaf PUE upkeep as explained in the previous sub-section.

Parameter influence on final RSR (Fig. 3.4, A, green) is similar to one obtained for plant concentration (Fig. 3.4, A, blue) with more contribution due to parameter interaction. Also, a much higher contribution of root respiration rate ( $\lambda_{CR}$ ) can be seen, as it determines the net root growth rate and subsequently the plant RSR.

Parameters which seemingly do not affect neither of the three modeled quantities (Fig. 3.4, A) are the amount of P allocated to roots ( $f_{PR}$ ), grain respiration rate ( $\lambda_{CG}$ ), P remobil-

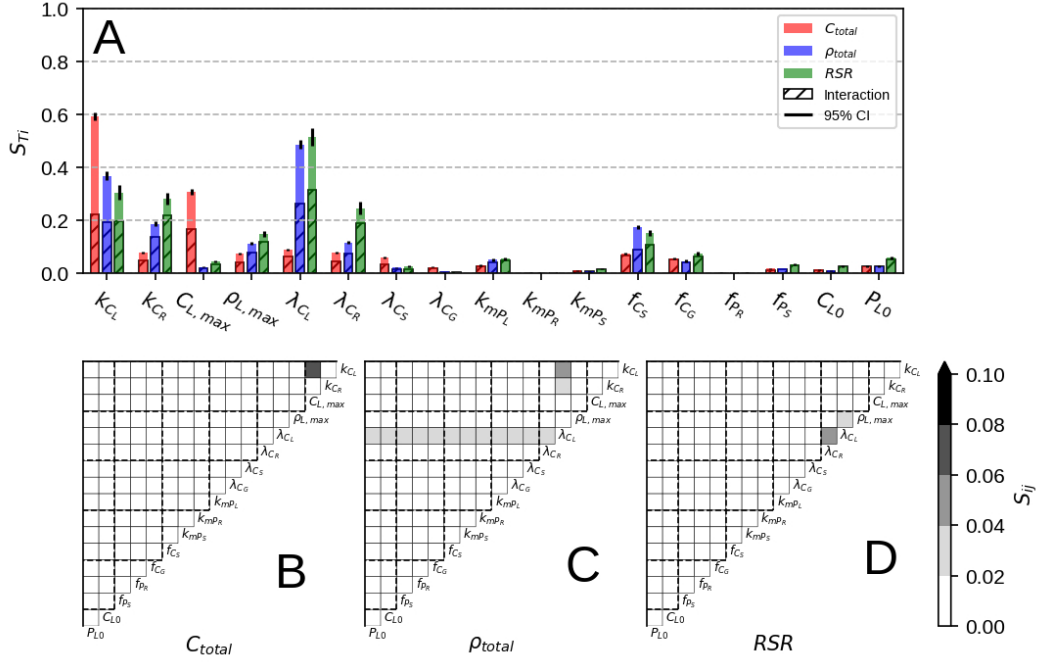


Figure 3.4: Results of Sobol sensitivity analysis on  $C_{total}$ ,  $\rho_{total}$  and  $RSR$  at simulation end. Upper row shows the total contribution (coloured) and sum of all interactions (hatched) for each model parameter. Lower row shows individual interaction contribution between each parameter. Red dots denote values that are significantly higher than zero. Parameter description can be found in Table 3.2.

isation rates ( $k_{mP_i}$ ) and initial leaf conditions ( $C_{L,0}$  and  $P_{L,0}$ ). For  $f_{PR}$ , it is due to roots' low biomass and low P allocation priority. For  $\lambda_{CG}$ , it is due to the grain respiration value affecting the whole plant net C balance only slightly. For  $k_{mPR}$  and  $k_{mPS}$ , it is due to non-leaf remobilisation only redistributing P among different organs.  $C_{L,0}$  and  $P_{L,0}$  only affect the initial adjustment stage (before the max. growth point) after which biomass and P pools are a product of C and P availability.

### 3.3.3 Comparison with field data

The model is able to reproduce very well the evolution of C pools during a growing season (Fig. 3.5, A and C) with discrepancies probably stemming from seasonal temperature effects which are not present in our model (via respiration and photosynthesis). When it comes to the observed P pools, the model is not able to reproduce the evolution of tissue P accumulation during the vegetative stage (Fig. 3.5, B and D) due to the previously mentioned  $LUE \times PUE$  dynamic. In the observations, it seems most of the plant P is taken up during this period and remobilised by the grain at a later stage.

In spite of the model's inability to predict the timing of vegetative P uptake, it manages to predict well the final plant response at different P levels (Fig. 3.6, B). The consistency of the response can be also be seen when comparing predicted root uptake rates and measured soil P availability (Supplementary Fig. B.9). A part of the mismatch comes from the fact

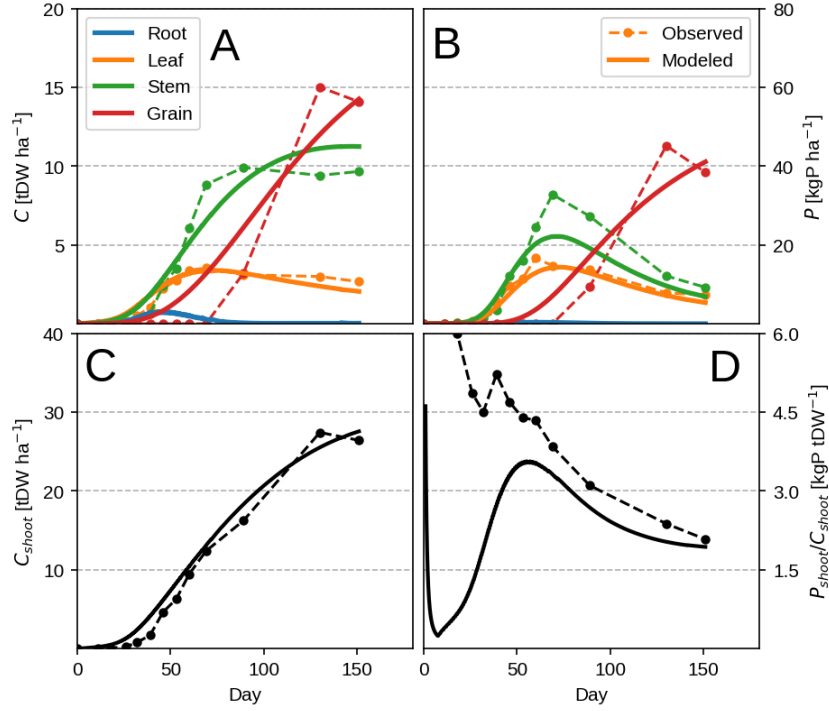


Figure 3.5: Modeled vs. observed time-series for the highest P input field maize experiment (block 4). Upper row depicts root, leaf, stem and grain C pools (left) and P pools (right). Lower row depicts total plant C (left) and total plant P concentration (right). Individual organs are depicted with different colors, where as total plant quantities are depicted with black.

that we use the whole season to calibrate our model. The other mismatch comes at high P availability when luxury uptake is observed (Fig. 3.6, C) but is not reproducible by our method as the plant grows in the most frugal way possible. Additionally, concentration of vegetative organs starts to decrease with higher P availability (Fig. 3.6, C) due to excessive grain P remobilisation. This might be due to observations falling mostly in the luxury uptake range, and could potentially be better constrained by having more P limited data.

The differences in calibrated and initial guess values (Table 3.3) are mainly due to differences in the cultivar and the growth environment. The assimilation rate ( $k_{CL}$ ) should depend on cultivar as well as temperature and the amount of light, where as LUE half-saturation point ( $C_{L,max}$ ) depends on planting density. Even though rates of root P uptake ( $k_{CR}$ ) and respiration ( $\lambda_{CR}$ ) are set during calibration, they might not contain reliable information since they are not constrained directly.  $\lambda_{CR}$  deviates much more than  $k_{CR}$  from the initial guess during calibration because its influence on the model is lower (Fig. 3.4, A). But these items should not pose a serious problem as overall influence of  $k_{CR}$  and  $\lambda_{CR}$  on the model is moderate (Fig. 3.4, A) and the simulated root mass is not grossly misrepresented (Figs. 3.5 and 3.6). Leaf respiration ( $\lambda_{CL}$ ) is much higher than the initial guess, most probably due to senescence or additional temperature dependent mechanisms which we do not account for. The same goes for stem and grain respiration rates ( $\lambda_{CS}$

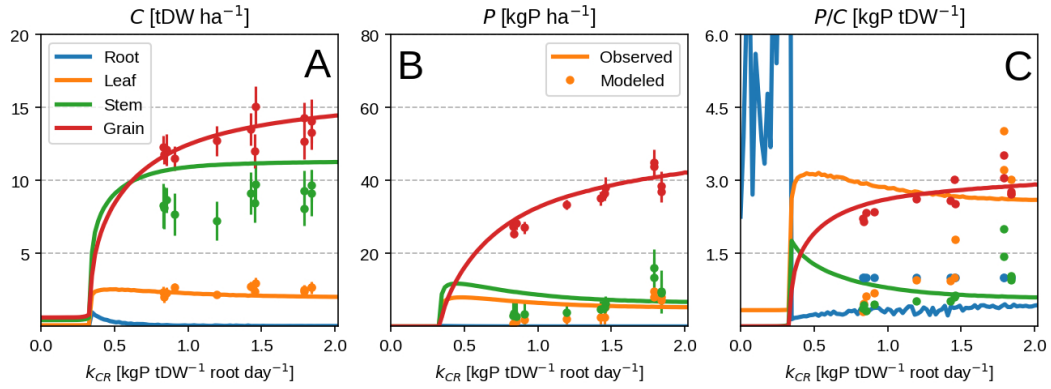


Figure 3.6: Modeled vs. observed values at field maize experiment season end as function of root P uptake rate. Leaf P remobilisation is turned on in the top row and turned off in the bottom one. From left to right the columns show C pool, P pool and concentration for leaf, stem and grain. Different organs are depicted with different colors. Lines are modeled values and markers are observed ones.

and  $\lambda_{CG}$ ) although  $\lambda_{CS}$  turns out to be lower in the end. Initial leaf biomass and P values ( $C_{L0}$  and  $P_{L0}$ ) show a high degree of uncertainty as they do not impact the overall model behaviour, but rather the initial adjustment period until the plant reaches the maximum growth point.

### 3.3.4 Comparison with hydroponic studies

The general form of plant response to P availability is reproduced well : an increase in shoot biomass, shoot fraction and concentration (Fig. 3.7) with increasing P in the nutrient solution (Supplementary Fig. B.10). Some discrepancies remain due to toxicity effects and luxury uptake, which are present in observations but are not reproducible by the model. They can be seen at high shoot P concentration (Fig. 3.7, A and B) when shoot biomass remains the same or starts to decline, contrary to the model which in principle has no limit to growth. P toxicity effects are usually linked to interactions with zinc (Zn, Loneragan et al., 1979) where Zn transport from root to shoot is inhibited and Zn deficiency is induced which could be represented in our model by a form of P control on Zn uptake. Luxury uptake of P luxury can not be reproduced as mentioned previously, because the plant is grown in the most P efficient way possible.

The calibrated values are different from the initial guess values as expected (Table 3.3) due to species diversity and differences in the growing environment as mentioned before. The values obtained are quite similar to ones from the maize field trial (Table 3.3) with the exception of root respiration rate ( $\lambda_{CR}$ ) which is two to three times higher (0.3 - 0.6  $\text{day}^{-1}$  for hydroponics vs. 0.14  $\text{day}^{-1}$  for field maize). This might be due to the experiment nature that looks at initial vegetative growth only, where younger roots expend much more C. Additionally, all of the calibrated parameters show a high degree of uncertainty which is due to a relatively small number of observations (information at experiment end only).

Table 3.3: Calibrated values for both the field trial and the hydroponic studies. Values for hydroponic studies are given as an average over all species. All parameters were kept the same across different P levels except  $k_{CR}$  which reflects P availability (minimum and maximum provided here). Parameter description can be found in Tables 3.1 and 3.2. Refer to Supplementary Tables B.4 and B.5 for per-species hydroponic values and a complete set of  $k_{CR}$  values.

Parameter	Initial guess	Calibrated values $\pm$ std. error	
		Field trial	Hydroponic studies
$k_{CL}$	1.5	$1.71 \pm 0.11$	$1.65 \pm 0.93$
$k_{CR} \text{ min}$	1.0	$0.83 \pm 0.09$	$0.09 \pm 0.69$
$k_{CR} \text{ max}$	1.0	$1.84 \pm 0.20$	$1.52 \pm 0.83$
$C_{L,max}$	1.5	$0.90 \pm 0.09$	$0.64 \pm 0.38$
$\rho_{L,max}$	1.0	$1.09 \pm 0.09$	$1.56 \pm 0.59$
$\lambda_{CR}$	0.30	$0.14 \pm 0.03$	$0.41 \pm 0.57$
$\lambda_{CL}$	0.10	$0.16 \pm 0.02$	$0.11 \pm 0.24$
$\lambda_{CS}$	0.03	$0.01 \pm 0.01$	-
$\lambda_{CG}$	0.01	$0.02 \pm 0.01$	-
$f_{CS}$	0.5	$0.40 \pm 0.05$	-
$f_{CG}$	1.0	$2.36 \pm 0.21$	-
$f_{PS}$	0.8	$1.56 \pm 0.07$	-
$f_{PR}$	0.1	$0.04 \pm 0.09$	-
$k_{mPL}$	0.05	$0.06 \pm 0.01$	-
$k_{mPR}$	0.05	$0.25 \pm 0.12$	-
$k_{mPS}$	0.05	$0.07 \pm 0.01$	-
$C_{L,0}$	0.1	$0.01 \pm 0.03$	$0.09 \pm 0.47$
$P_{L,0}$	0.1	$0.03 \pm 0.02$	$0.06 \pm 0.60$

### 3.4 Discussion

According to the model results, the effects of P limitation on plants can largely be described as a compromise between root and leaf growth combined with changing efficiency of leaf P use. In situations of low P supply, plants devote relatively more C to roots at the expense of shoot to sustain plant P demand. This sets the stage for P limitation as lower leaf biomass translates into a lower biomass potential. Increasing the P availability will increase the plant's ability to grow due to higher leaf concentration, but the efficiency of this gain decreases as more roots are needed to sustain additional growth.

The model itself does not attempt to include all of the mechanisms that modulate C and P availability which could render it into a fully-fledged vegetation one ( Krinner et al., 2005; Mäkelä, Valentine, and Helmisaari, 2008; McMurtrie et al., 2008; Franklin et al., 2009). Instead, it tries to describe plant response by relying on the functional role of different plant organs (which are acquisition of C in leaves and P in roots). These are simply

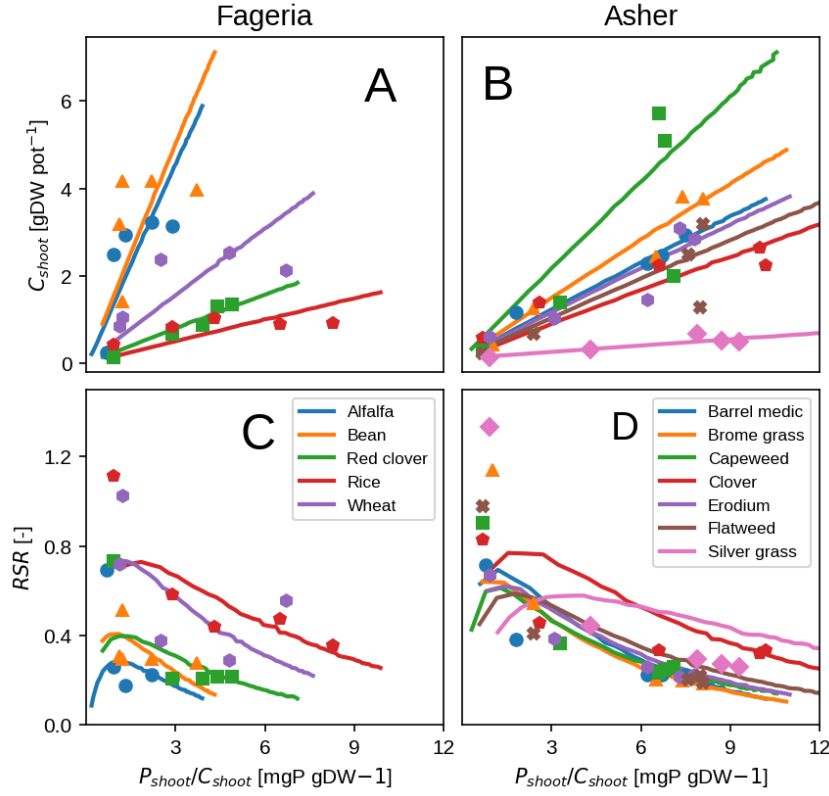


Figure 3.7: Modeled vs. observed shoot biomass (top row) and root-shoot ratio (bottom row) versus shoot P concentration. Results show the response at at hydroponic experiment end. Left column depicts grass species (Asher and Loneragan, 1967). Right column depicts cereals and legumes (N. K. Fageria and Baligar, 1989).

expressed as resource uptake rates (the leaf assimilation rate  $k_{CL}$  and root P uptake rate  $k_{CR}$ ) to facilitate comprehension, whereas in reality they are strongly modified by the plant's environment. Leaf assimilation depends on temperature, incoming radiation and CO<sub>2</sub> partial pressure (Berry and Bjorkman, 1980) as well as water availability (Hsiao, 1973). Root P uptake depends on the properties and the amount of P in the underlying soil, as well as the physiological and physico-chemical limits to P transport (Barber, 1995). Inclusion of these processes via known mechanistic models like Farquhar's photosynthesis (Farquhar, von Caemmerer, and Berry, 2001) or Barber's P diffusion one (Barber, 1995) would enable this approach to be extended to a multitude of growing environments. But the generalization and the additional level of detail would not necessarily improve the model's ability to reproduce plant behaviour which, in essence, only tries to describe how to allocate C and P between root and leaf.

Additionally, most of the model response rests on the shoulders of leaves (Fig. 3.4, A) which is not surprising as the whole concept of optimal functioning revolves around maximizing leaf productivity. This is a serious drawback of the approach because plants possess a multitude of architectural and development strategies to cope with reduced P availability (Niu et al., 2012; Plaxton and Lambers, 2015). Architectural ones (Niu et al., 2012) include

increased top-soil foraging, lateral root growth promotion, root length density increase and cluster root formation. The first two could be modeled using a 2-dimensional root model (Heinen, Alain Mollier, and Willigen, 2003) but require a vertical description of soil P fractions instead of a bulk soil P quantity. The latter two (dealing with root length density) could be attempted by implementing different root orders with varying C respiration and loss rates. Development strategies (Plaxton and Lambers, 2015) include root exudation of acids and enzymes, mycorrhizal association, lowering the metabolic cost of photosynthesis and increasing P remobilisation efficiency/seed P content among others. But these require a much more complex description of soil P chemistry, as well as a full accounting of C costs and benefits for each of the underlying processes. Nevertheless, we foresee no fundamental issues with their implementation which we avoid here to facilitate comprehension of the model and its results.

Our approach is similar to previous optimal function ones (McMurtrie et al., 2008; Mäkelä, Valentine, and Helmisaari, 2008; Franklin et al., 2009) where plant productivity is modified by leaf nutrient concentration and allocation among different organs optimized in order to reach maximum growth. This assumption provides a reasonable starting point for describing plant physiological and functional response, but overlooks the fact that leaves are the only ones impacted by the PUE assumption (Eq. 3.4). According to this hypothesis (which is driven by leaf concentration alone) there is no benefit from investing P into non-assimilating organs. To overcome this, allocation of P is linked via stoichiometry constraints (Eq. 3 in Franklin et al., 2009) or directly to the leaf concentration (Eq. 8 in Mäkelä, Valentine, and Helmisaari, 2008). We follow the second approach, where filling of non-assimilating organs is driven by the leaf P flux (Eq. 3.15). An answer to this issue might be to increase model complexity like in Thornley's transport-resistance approach (John H. M. Thornley, 1995) where a plant is described as a network of nutrient and C exchanging organs, whose individual growth is determined by the labile (or exchanged) substrate concentration. But the problem using this approach is the reliance on parameters which are seldom measured (G. Ågren and Franklin, 2003) like rates of nutrient productivity, substrate utilization rates and transport resistances.

One notable difference between our model and some of the mentioned studies (Mäkelä, Valentine, and Helmisaari, 2008; Franklin et al., 2009) is the lack of maintenance respiration due to P. Maintenance respiration is defined as a C cost deducted from gross primary productivity (GPP) to support the nutrient's metabolic activity. For N, this cost is related to protein turnover which requires energy for replacement and repair (Ryan et al., 1996) where as for P the cost could be related to formation of nucleic acid and triose phosphate (Rowland et al., 2017). We chose not to implement this process primarily because the available information is confined to forest eco-systems in tropical and sub-tropical area (Meir, Grace, and Miranda, 2001; Rowland et al., 2017) which are not studied here. On a conceptual level, this assumption is not necessary to elicit plant response to varying nutrient

availability. The decreasing P efficiency (via the PUE kinetic, Eq. 3.4) is a sufficient condition, where additional C gains are more and more expensive due to increasing requirement for root growth. Introducing additional C costs due to P maintenance respiration should modify only the final C balance at different P levels, but separating P contributions to maintenance respiration is difficult at this point due to dearth of data as mentioned.

In spite of some short-comings, optimal function approaches provide a sensible description of plant development that relies less on empirical constraints (like potential leaf-area curves, thermal sum driven phenology and prescribed allocation patterns) which are often employed in eco-physiological models (Sönke Zaehle and Dalmonech, 2011; Rosenzweig et al., 2014). Even though seasonal dynamics of P allocation are not captured well, it might not be a critical issue as final distribution of P among plant organs should have an effect on long-term P cycling (P. M. Vitousek et al., 2010) and can be captured well by final plant concentration. What is more important is the representation of C phenology, as it has a more direct connection to the plant's energy and water balance. This paves way for connection to the plant's metabolism that is strongly moderated by temperature and water availability (Berry and Bjorkman, 1980; Atkin and Tjoelker, 2003; Hsiao, 1973).

For temperature, this should be pretty straightforward by moderating rates of assimilation ( $k_{CL}$ , Table 3.1) and respiration ( $\lambda_{Ci}$ , Table 3.1) via known mechanistic or empirical relationships (Berry and Bjorkman, 1980; Atkin and Tjoelker, 2003). For water, an increased level of complexity is warranted. The main mechanism of plant response to water stress is the regulation of stomata closure to maintain internal water potential (Hsiao, 1973). Even though a multitude of physiological symptoms emerge in water stressed conditions (Hsiao, 1973) the most pertinent one is the loss in photosynthetic capacity due to a decrease in leaf  $\text{CO}_2$  flux with stomata closure. This presents two avenues to modify plant productivity, either by increasing root water uptake or decreasing leaf transpiration. Increasing root uptake will increase C costs due to root growth and maintenance, where as decreasing leaf transpiration will decrease C gains through lower assimilation capacity as mentioned before. One way to approach this problem is to prioritize water by prescribing stomate behaviour at different water levels, and then optimize C allocation to maximize nutrient driven productivity like in McMurtrie et al., 2008. But taking into account both of these strategies requires additional links via the assimilation rate  $k_{CL}$  and the plant's water balance, that can not be captured by a simple resource use kinetic typically utilised in optimal function approaches (Dewar et al., 2009).

So far we tacitly assume P limitation can be treated in the same way as N. Even though this is not justified physiologically, the observed functional response of root biomass change (Poorter et al., 2012) can serve as an argument for it. From a leaf productivity perspective the primary role of N is to provide the enzymes that catalyze photosynthesis where as for P, it is the ribosomes that allow their formation (G. I. Ågren, 2008). Decreasing availability of both N and P should have an effect on leaf photosynthetic efficiency and drive plant re-

sponse. This is what our PUE kinetic (Eq. 3.4) tries to achieve similar to the other N based studies (Dewar et al., 2009). Furthermore, their interaction could be attempted by weighing gross productivity (Eq. 3.2) with individual resource use kinetics ones like our PUE one (Eq. 3.4). It is very likely though that caveats encountered here would be even more pronounced with multiple limiting factors, since plant response is driven by leaf alone.

Our model could be integrated within a dynamic vegetation model (Krinner et al., 2005) where growth can be driven by a more physical description of underlying plant processes. In our approach, assimilation and root uptake are directly connected to leaf and root biomass. These could be transformed to leaf area index and root length density, making them amenable to explicit photosynthesis (Farquhar, von Caemmerer, and Berry, 2001) and soil-root diffusion (Barber, 1995) parametrizations. Furthermore, connection to the underlying soil can be done via a soil P model (Y. P. Wang, Houlton, and C. B. Field, 2007; Ringeval et al., 2017). Such coupling could help us investigate effects of P limitation in eco-systems (Peñuelas et al., 2013) while accounting for plant adjustment and investigate long term effects of P cycling (Goll et al., 2012; Ringeval et al., 2017).

## Chapter 4

# Optimal allocation of carbon and phosphorus in irrigated maize using ORCHIDEE

Marko Kvakić, Philippe Ciais, Daniel Goll, Sylvain Pellerin, Bruno Ringeval

To be submitted to *Geoscientific Model Development*

## Abstract

The knowledge on how global crop production will respond to a changing environment is direly needed, as research points to an uncertain future, putting into question our capacity to adapt to the coming times. Impacts of climate and management are unquestionable to this date, but how their interaction will steer global production requires a more comprehensive look. To this end, we describe a crop model based on optimal functioning (OF) principles, which we combine with a generic vegetation one (ORCHIDEE) and apply it to the issue of phosphorus (P) limitation, for which it has already been shown a great potential to contribute to existing global production gaps.

OF framework tries to model plant response based on growth resource availability, which in our case are carbon (C) and P. Combining ORCHIDEEs physical parametrizations with OF allows for a dynamic feedback of plant growth to a changing environment, which we compare to two irrigated monoculture maize systems (Nebraska,USA and Tartas,France). Additionally, we qualitatively compare the performance of the OF version to the original one over a climate gradient.

OF is largely capable of capturing the evolution of different C pools in a growing season, but not so for P. This is due to the parsimonious nature of its underlying hypotheses, which center on leaf productivity alone. Additional improvements need to be made, most notably the inclusion of water (H<sub>2</sub>O) effects and a complete soil P cycling module. But in spite of these processes missing, a consistent framework is provided for simulating hypothetical future scenarios which could potentially be extended to the global scale.

## 4.1 Introduction

Quantification of global crop productivity is of great importance, especially in light of recent global change, as agricultural systems all over are exposed to a multitude of factors determining food security for the near future (Foley et al., 2011; Wheeler and Braun, 2013). Today's limits to the majority of world crop production can be explained in large part by climate trends and variability (D. Lobell and C. Field, 2007; Ray et al., 2015), as well as employed management practices (Mueller et al., 2012; Sinclair and Rufty, 2012; Lassaletta et al., 2014) all of which points to the climate-management interaction as the main driver of crop production to this date. But assessment of their combined historical effects and how these will play out in the future is needed, as research points to "a looming and growing agricultural crisis" (Ray et al., 2013).

The previously mentioned estimates are based on statistical approaches, which work by establishing strong empirical links with main drivers of yield response (climate, irrigation and nutrient addition) and permit a robust accounting of their implicit means of action (D. Lobell and C. Field, 2007). But the simplified nature of treating aggregate effects on crop yield might overlook other mechanisms that could relieve limitation (like plant adjustment or existing nutrient reserves) and whose information on a global scale is not easily obtainable, restricting the use of such relationships to the range of environments and periods inquired. A useful counterpart to statistical approaches are global gridded crop models (GGCMs) which include parametrization of plant-soil processes and management effects, allowing a more integrated description of crop growth that can be extended to hypothetical scenarios (Deryng et al., 2014; Rosenzweig et al., 2014).

An example highlighting the need for a more elaborate description of limitation in crops is phosphorus (P), a major plant nutrient (N. K. Fageria, 2009). Distribution of P in global cropland is highly unbalanced (MacDonald et al., 2011) and its availability to crops strongly determined by soil properties and history of P input (Ringeval et al., 2017) due to the nature of soil-plant P exchange (Barber, 1995). The issue of P (as with any other growth limiting resource) beckons for a deeper look into crop limitation mechanisms and how a plant reacts to them, if we are to have a clearer picture of crop production today and in the future.

In the case of nutrient limitation, or specifically P in ours, the primary plant response is reallocation of carbon (C) towards roots to change its root-shoot ratio (RSR) and counteract the decreased P supply (Marschner, Kirkby, and Cakmak, 1996). Since P serves an important role in the plant's metabolism (Plaxton and Lambers, 2015) lowering the P supply will necessarily impact growth and ultimately crop yield. One of the ways to model plant adjustment to nutrient limitation is the optimal functioning (OF) approach (Dewar et al., 2009) where a plant is described as a set of organs serving a certain function (assimilation of C in leaves, uptake of P in roots). The activity of each organ is guided with one goal in mind: maximising plant productivity as a proxy for individual fitness. Even though

a simplistic view of plant behaviour (John H. M. Thornley, 1995; Franklin et al., 2012), OF approach provides a parsimonious and transparent way of modelling plant adaptation to nutrient availability and environmental change (Dewar et al., 2009).

Accounting for climate effects on plant growth is possible using well established mechanistic relationships of processes like photosynthesis (Farquhar, von Caemmerer, and Berry, 2001), transpiration (Ball, Woodrow, and Berry, 1987) and respiration (Ruimy, Dedieu, and Saugier, 1996). These are an integral part of dynamic global vegetation models (DGVMs, C. Peng, 2000) which try to capture the surface exchange of C, water (H<sub>2</sub>O) and energy for the purpose of vegetation-climate feedback modelling. To this end we rely on the model ORCHIDEE (Krinner et al., 2005) and combine it with OF principles to model plant growth as function of C and P availability. By employing a generic description of the underlying mechanisms we hope to enable the study of C x P interaction in a multitude of environments, which is to be used on a global scale.

## 4.2 Methods

### 4.2.1 ORCHIDEE description

The employed DGVM is an improved version of ORCHIDEE (Matthieu Guimberteau et al., 2018) which consists of two main parts. The first one deals with short-term processes (30 minutes) of hydrologic/thermal exchange and photosynthesis. The second one deals with long-term processes (1 day to 1 year) of C respiration, allocation, plant phenology and plant community dynamics. The improved version (Matthieu Guimberteau et al., 2018) focuses on the parametrization of high-latitude processes like: permafrost physics, snow melting/insulation and soil organic C feedback on soil physical properties. Most of these processes are turned off for this work (as we model irrigated maize in a temperate climate) so we shall focus on ones that are directly coupled to the proposed OF approach. For an exhaustive description of other parametrizations please consult the mentioned references: Krinner et al. (2005) for the original and Matthieu Guimberteau et al. (2018) for the improved one.

ORCHIDEEs main paradigm is the use of plant functional types (PFTs) of which there are 13 (bare soil, 8 types of forests, 2 types of grasses and 2 types of crops). These PFTs occupy a certain fraction of every simulated grid point and determine parameter values for all underlying parametrizations. As we simulate irrigated maize monoculture, we prescribe the entire area as maize PFT and do not change it. During simulation, we rely on ORCHIDEE parametrizations to drive plant growth and optimally allocate C in the OF version. For this we need several ORCHIDEE calculated quantities: the flux of C via assimilation and respiration, and the flux of P via root P uptake.

In ORCHIDEE, C<sub>4</sub> species assimilation rate ( $A$ ,  $\mu\text{molCO}_2 \text{ m}^{-2} \text{ s}^{-1}$ ) is calculated using photosynthesis/stomatal models (Collatz, Ribas-Carbo, and Berry, 1992; Ball, Woodrow,

and Berry, 1987) and leaf area index (LAI) to calculate gross primary production (GPP). This is done every 30 minutes to account for changes in H<sub>2</sub>O availability and temperature, and summed up each day to produce daily GPP (gC m<sup>-2</sup> day<sup>-1</sup>) as available C for plant respiration and growth. LAI is connected to leaf biomass via constant specific leaf area (SLA) and a Beer-Lambert extinction law is used to calculate the effective assimilation rate of an expanding canopy:

$$I(x) = I_0 \cdot e^{-k \cdot x} \quad (4.1)$$

$$GPP \propto \int_0^{LAI} A(I) dx \quad (4.2)$$

$A$  is the leaf photosynthetic rate ( $\mu\text{molCO}_2 \text{ m}^{-2} \text{ s}^{-1}$ ),  $LAI$  the leaf area index (-),  $k$  the light extinction rate ( $LAI^{-1}$ ),  $I_0$  the incident radiation at canopy top ( $\text{W m}^{-2}$ ) and  $x$  is the cumulative canopy LAI (-). Effects of H<sub>2</sub>O, nitrogen (N) and leaf age can additionally be taken into account, but are disabled here since we assume the plant not to be limited by neither.

In ORCHIDEE, total plant respiration is calculated as the sum of maintenance and growth respiration. Maintenance respiration rates ( $\lambda_i$ ) are calculated as a linear function of temperature (Ruimy, Dedieu, and Saugier, 1996):

$$\lambda_{Ci} = \lambda_0 \cdot (1 + \lambda_{Ti} \cdot T); \quad i = R, L, S, G, X \quad (4.3)$$

$\lambda_{Ci}$  is the organ respiration rate ( $\text{day}^{-1}$ ),  $\lambda_0$  is the respiration rate at zero degree Celsius ( $\text{day}^{-1}$ ),  $\lambda_{Ti}$  is the change in respiration rate with annual temperature ( $\text{day}^{-1} \text{ } ^\circ\text{C}^{-1}$ ) and  $T$  is temperature in Celsius. Index  $i$  denotes organs : roots ( $R$ ), leaves ( $L$ ), stem ( $S$ ), grain ( $G$ ) and reserve ( $X$ ). Maintenance respiration can use up to 80% of available assimilates, after which plant biomass is spent (or removed) to satisfy maintenance costs. The remaining assimilates are used to grow plant organs, from which growth respiration is deducted using a fixed ratio (28%). For OF purposes, we turn growth respiration off and model total plant respiration using only the maintenance calculated rates (Eq. 4.3) as growth respiration is calculated during allocation (described below).

C allocation in the original ORCHIDEE is performed using prescribed allocation fractions as function of light, H<sub>2</sub>O and N availability. Since considerable changes were introduced in ORCHIDEE (see Eqs. A29 - A41 in Krinner et al. 2005), we present the allocation scheme in it's original form for clarity sake (Friedlingstein et al., 1999):

$$f_R = 3 \cdot r_0 \cdot \frac{L}{L + 2 \cdot \min(W, N)} \quad (4.4)$$

$$f_S = 3 \cdot s_0 \cdot \frac{\min(W, N)}{2 \cdot L + \min(W, N)} \quad (4.5)$$

$$f_L = 1 - f_R - f_S \quad (4.6)$$

$f_R$ ,  $f_S$  and  $f_L$  are allocation fractions for roots, stems and leaves.  $L$ ,  $W$  and  $N$  are scalars (0.1 - 1.0) representing light, H<sub>2</sub>O and N availability.  $r_0$  and  $s_0$  are root and stem allocation fractions in unlimited conditions (or when  $L$ ,  $W$  and  $N$  are equal to one). We circumvent this step in our version, as C and P are allocated using OF principles.

Nutrient availability in ORCHIDEE is prescribed as an empirical function of soil humidity and temperature for N only. Since we are interested in P and no explicit description of plant-soil nutrient cycling exists, we prescribe P availability as a constant root uptake rate while N (like H<sub>2</sub>O) is considered to be not limiting.

ORCHIDEE describes rules for growth initiation based on meteorological criteria like temperature sums or moisture thresholds, but we use prescribed initiation days to represent planting dates (as we simulate monoculture maize)

#### 4.2.2 Optimal functioning implementation

We provide a schematic depicting how the previously mentioned processes are coupled to the optimization process (Fig. 4.1). Within ORCHIDEE, respiration rates ( $\lambda_i$ ) and GPP ( $GPP_{day}$ ) are calculated using plant biomass ( $C_{plant}$ ) and LAI ( $LAI$ ) on a daily basis. These values are passed to the optimization routine to allocate C and P. Since there is no feedback of plant P ( $P_{plant}$ ) on ORCHIDEE parametrizations,  $P_{plant}$  is considered only within the optimization routine. Following optimal allocation,  $C_{plant}$ ,  $P_{plant}$  and  $LAI$  are updated before the next time-step, when the whole process repeats.  $\lambda_i$  are passed directly, whereas  $GPP_{day}$  includes leaf shading effects (Eqs. 4.1 and 4.2) which are removed before optimization is performed.

In the optimization routine, we describe the plant as a set of organs (root, leaf, stem, grain and reserve) where leaves and roots acquire respectively C and P. These are allocated in order to maximize leaf productivity ( $dC_{day}$ ) limited by self-shading ( $LUE$ ) and leaf P status ( $PUE$ ) in the form of a Liebig's law of minimum:

$$Maximize(dC_{day}) \quad (4.7)$$

$$dC_{day} \leq k_{CL} \cdot C_{L,max} \cdot LUE \cdot dt \quad (4.8)$$

$$dC_{day} \leq k_{CL} \cdot C_{L,max} \cdot PUE \cdot dt \quad (4.9)$$

$$LUE = \frac{C_L^*}{C_L^* + C_{L,max}} \quad (4.10)$$

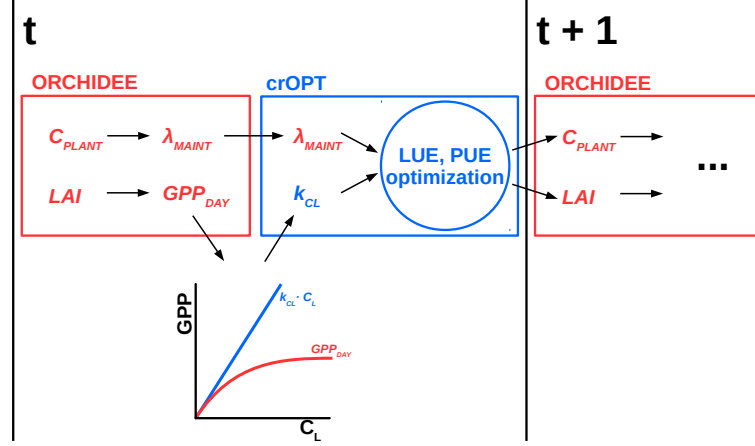


Figure 4.1: Schematic of optimal allocation within ORCHIDEE. Arrows show feedbacks between ORCHIDEE (red) and the optimization routine (blue) from one time-step to the other.

$$PUE = \frac{P_L^*}{P_L^* + C_L^* \cdot \rho_{L,max}} \quad (4.11)$$

$dC_{day}$  is the daily GPP ( $\text{gC m}^{-2}$ ) and  $k_{CL}$  the potential leaf assimilation rate ( $\text{gC gC}^{-1} \text{day}^{-1}$ ).  $C_{L,max}$  and  $\rho_{L,max}$  are leaf biomass ( $\text{gC m}^{-2}$ ) and P concentration ( $\text{mgP gC}^{-1}$ ) at which maximum leaf productivity is halved.  $C_L^*$  and  $P_L^*$  are updated leaf biomass ( $\text{gC m}^{-2}$ ) and P ( $\text{mgP m}^{-2}$ ).  $dt$  is the optimization time-step (1 day).  $k_{CL}$  is calculated by dividing ORCHIDEEs GPP with current leaf biomass ( $C_L$ ) after shading effects have been removed (Eq. 4.1 and 4.2).

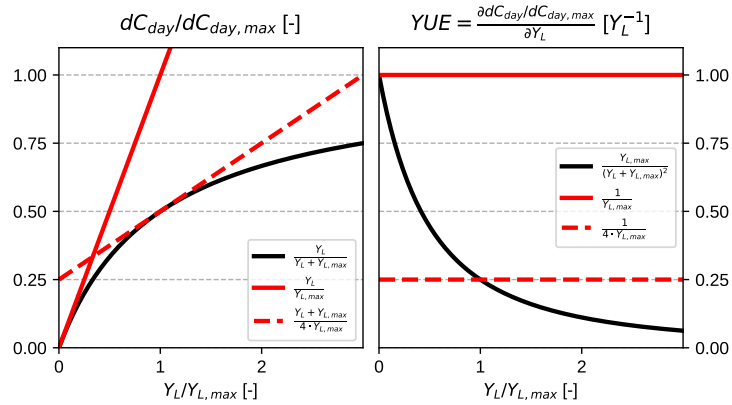


Figure 4.2: Relative leaf productivity (left) and  $Y_L$  use efficiency (YUE, right) where  $Y_L$  is a generic leaf-dependent element. Dark line represents the form used in our work (Eqs. 4.10 and 4.11). Potential productivity rate is depicted with a full red line and a dashed one when quartered.  $dC_{day,max}$  is maximum leaf productivity in fully unlimited conditions equal to  $k_{CL} \cdot C_{L,max}$

The mechanism of *LUE* and *PUE* interaction is depicted in Fig. 4.2 where leaf productivity is drawn as a function of a generic leaf-dependent element ( $Y_L$ ) which in our case are  $C_L$  and  $\rho_L$ . Low  $Y_L$  availability equates to low productivity but also to highest efficiency of  $Y_L$  use (*YUE*) when the production rate is highest. Leaf productivity increases with  $Y_L$  (but with decreasing *YUE*) the limit of which converges to  $k_{CL} \cdot C_{L,max} \cdot dt$ ; the maximum leaf productivity in fully unlimited conditions. The rest of the optimization describes the flow of C and P in the plant using the following set of equations. For C these are:

$$C_i^* = C_i + dC_i - rC_i \quad (4.12)$$

$$dC_{day} = \sum_{i=L,R,S,G,X} dC_i \quad (4.13)$$

$$rC_i = \lambda_{C_i} \cdot C_i \cdot dt \quad (4.14)$$

$C_i^*$  is the updated C pool as a balance of allocated C ( $dC_i$ ) and respired C ( $rC_i$ ).  $C_i$  is the current C pool and  $\lambda_{C_i}$  is the respiration rate. Index  $i$  denotes different organs. For P, the equations are following:

$$P_i^* = P_i + dP_i - mP_i \quad (4.15)$$

$$dP_{day} = k_{CR} \cdot C_R^* \cdot dt \quad (4.16)$$

$$dP_{day} = \sum_{i=L,R,S,X} dP_i \quad (4.17)$$

$$mP_i = k_{mP_i} \cdot P_i \cdot dt; \quad i = L, R, S \quad (4.18)$$

$$\sum_{i=L,R,S,G} mP_i = 0 \quad (4.19)$$

$P_i^*$  is the updated P pool of each organ  $i$ , which a balance allocated P ( $dP_i$ ) taken up by the roots ( $dP_{day}$ ), and remobilised P ( $mP_i$ ).  $P_i$  is the current P pool,  $k_{CR}$  is the root P uptake rate and  $k_{mP_i}$  is the P remobilisation rate. Index  $i$  denotes different organs. We assume no saturating effect for root uptake (Eq. 4.16) to reduce model complexity, assuming it can be due to external physical limitations to P supply (eg. soil diffusion) which are not modeled here. Allometric constraints are applied on stem and grain C filling, as

well as on P filling of non-assimilating tissue:

$$f_{C_S} \cdot dC_L = dC_S \quad (4.20)$$

$$f_{C_G} \cdot rC_S = dC_G \quad (4.21)$$

$$f_{P_i} \cdot dP_L = dP_i; \quad i = R, S \quad (4.22)$$

$f_{C_S}$  and  $f_{C_G}$  are the fraction of C going towards stem and grain, and are tied to the amount of C allocated to leaf ( $dC_L$ ) and stem respired C ( $rC_S$ ) respectively.  $f_{P_i}$  is the fraction of P going towards the photosynthetically non-active tissue (root and stem) and is tied to the amount of leaf allocated P ( $dP_L$ ). Equation 4.20 follows the principle of the pipe-theory model (Shinozaki et al., 1964). For grain (Eq. 4.21) we rely on the concept of Iwasa and Roughgarden (1984) where grain filling is triggered once a plant reaches its maturity. In our approach, we model grain filling as a continuous process where the grain C flow reaches its peak when the vegetative part stops growing (or stem respiration is highest). Equation 4.22 is a necessary assumption if we want to fill the photosynthetically non-active pools with P, since the plant does not confer any benefit from doing so. Details on the various variables and parameters are given in Table 4.1.

Table 4.1: Optimal allocation variables and constants.  $i$  denotes different organs : leaves (L), roots (R), stem (S), grain (G) and reserve(X).

Variable	Description	Unit
$k_{CL}$	Potential assimilation rate	gC gC <sup>-1</sup> leaf day <sup>-1</sup>
$dC_{day}$	Daily gross productivity	gC m <sup>-2</sup>
$dP_{day}$	Daily root P uptake	mgP m <sup>-2</sup>
$C_i$	Current C pool	gC m <sup>-2</sup>
$P_i$	Current P pool	mgP m <sup>-2</sup>
$C_i^*$	Updated C pool	gC m <sup>-2</sup>
$P_i^*$	Updated P pool	mgP m <sup>-2</sup>
$dC_i$	Daily C allocation	gC m <sup>-2</sup>
$rC_i$	Daily C respiration	gC m <sup>-2</sup>
$dP_i$	Daily P allocation	mgP m <sup>-2</sup>
$mP_i$	Daily P remobilised	mgP m <sup>-2</sup>
Constant		
$dt$	Time step	day
$k_{CR}$	Root P uptake rate	mgP gC <sup>-1</sup> root day <sup>-1</sup>
$C_{L,max}$	LUE half-saturation point	gC leaf m <sup>-2</sup>
$\rho_{L,max}$	PUE half-saturation point	mgP gC <sup>-1</sup> leaf
$f_{C,i=S,G}$	C allocation fraction	gC gC <sup>-1</sup> leaf
$\lambda_{C,i=L,R,S,G}$	C respiration rate	day <sup>-1</sup>
$f_{P,i=R,S}$	P allocation fraction	mgP mgP <sup>-1</sup> leaf
$k_{mP,i=R,L,S}$	P remobilisation fraction	gC <sup>-1</sup> grain

### 4.2.3 Calibration with observations and ORCHIDEE comparison

We use two different datasets to calibrate our model, both of which describe growth of irrigated monoculture maize in a temperate environment (with information given every 7 to 14 days). The first one comes from an eddy-covariance site in Nebraska, United States (B. Peng et al., 2018) and the other from a long-term fertilization experiment in Tartas, France (Plénet et al., 2000). The Nebraska dataset contains information on aboveground biomass allocation, LAI and C exchange (GPP and ecosystem respiration) for a single site in the period 2003-2012, with management conditions assumed to be constant. The Tartas dataset contains information on LAI, aboveground biomass and P allocation for 3 different sites in the year 1996, with each site having a different P management history (3 P levels). Due to the nature of information availability in each dataset, Nebraska should better constrain the C portion of the model whereas Tartas should better constrain the P one. The aboveground organs are leaves, stems and grain. No below ground plant information is given in neither dataset, although Tartas contains information on soil solution P concentration ( $C_P$ ). To extract plant respiration ( $R_{plant}$ ) from the Nebraska dataset, we model ecosystem respiration ( $R_{eco}$ ) as a linear function of observed GPP ( $GPP$ ) and air temperature ( $T$ ), while subtracting the soil terms ( $a_{soil}$  and  $b_{soil}$ ):

$$R_{eco} = R_{soil} + R_{plant} = (a_{soil} + b_{soil} \cdot T) + (b_{GPP} \cdot GPP + \varepsilon) \quad (4.23)$$

We provide a figure (SI Fig. B.11) to show the effect of removing soil respiration using Eq. 4.23.

When calibrating, we use all of the available information to constrain the model parameters. This entails 115 time points for Nebraska and 13 time points across 12 different P experiments in Tartas (3 P levels x 4 blocks). For Tartas, we assume the root P uptake rate ( $k_{CR}$ ) to be the only parameter changing among the different P experiments as it reflects P availability. All of the other parameters are kept constant. In this manner the obtained values should be specific to the maize variety, where plant response depends on seasonal change in meteorological conditions and P availability in case of Tartas. The final calibrated values should correspond to growth that is equally and weakly limited by other major abiotic factors like  $H_2O$ , light or N across experiments for each dataset (different years in Nebraska and P levels in Tartas).

To solve our system we linearise Equations 4.8 - 4.11 and solve it daily using the linear programming package GLPK (Makhorin, 2001). To calibrate parameter values we minimize the mean squared error using the Levenberg-Marquardt algorithm (Moré, 1978) implemented within Scipy's optimize package (E. Jones, Oliphant, and Peterson, 2001). The parameter list and the initial guess values are given for Nebraska and Tartas in SI Tables B.6 and B.7, whereas final calibrated ones are given in SI Tables B.8 and B.9.

We compare the original and the OF versions along a climate gradient (depicted in SI

Fig. B.14) using Nebraska calibrated parameters. To make the two versions comparable we remove growth respiration, the internal LAI cap (equal to 2.5), as well as implicit N and age effects from the original one. In this manner, the only differences are due to the *LUE* shape (exponential vs. a Michaelis-Menten one) and the C allocation scheme (prescribed vs. optimal).

Additional comparison of the modeled result (using Nebraska calibrated parameters) is made to observations using Mueller et al. dataset (2012) which contains globally gridded statistical estimates of the maize yield potential. We qualitatively compare zonal means to assess yield climate response (due to rainfall, temperature and light) since zonal means should remove horizontal management variability (in irrigation intensity and nutrient availability). The Mueller et al (2012) maize yield potential is depicted in SI Fig. B.16.

## 4.3 Results

### 4.3.1 Comparison with Nebraska

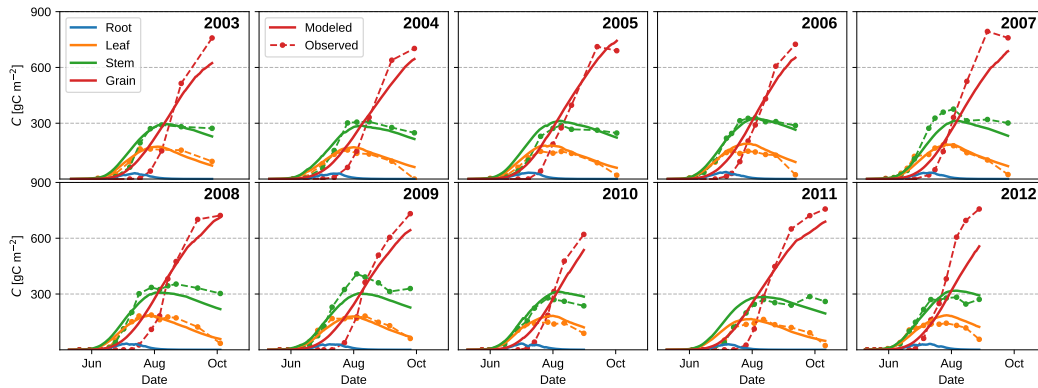


Figure 4.3: Nebraska observed vs modeled C pools.

Looking at modelled C pools (Fig. 4.3) the model seems to capture the overall behaviour, even though discrepancies remain in form of biases and offsets in time. This can best be seen with the grain C pool, which is significantly underestimated at season end due to either model misrepresentation or constant planting dates. Grain filling (Eq. 4.21) is prescribed in a continuous fashion and simulates a grain C pool that exists from the beginning, whereas reproductive growth starts at a later stage in the observations. For the planting dates, a better agreement could be achieved by having different planting days each year, as modeled bias follows closely the observed one (SI Fig. B.12)

Looking at modelled C exchange (Fig. 4.4) the model seems to capture the overall behaviour again but with slight biases as mentioned previously. These could either be due to variable planting dates or the way respiration is modeled. In our model, we have kept the temperature response of the respiration rate linear (Eq. 4.3) where as it might be more

suitable to have it exponential (Atkin and Tjoelker, 2003) as respiration rates during the final growth stages seem to be over-estimated.

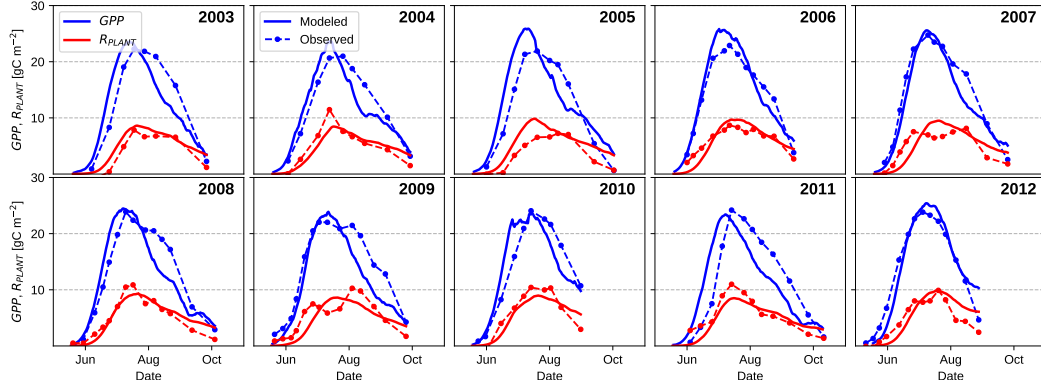


Figure 4.4: Nebraska observed vs modeled C exchange.  $R_{plant}$  is not a true observation, but is modeled using Eq. 4.23

### 4.3.2 Comparison with Tartas

In Tartas, the model gives satisfactory results for C but less so for P (Fig. 4.5). For P, most notable is the underestimation of plant P concentration due to the  $LUE \times PUE$  interaction which (during initial plant growth) prefers to grow bigger leaves instead of supplying them with more P. Nevertheless the total plant P concentration settles to a value which is closer to the observed one by the time the plant is harvested (Fig. 4.5, lower right).

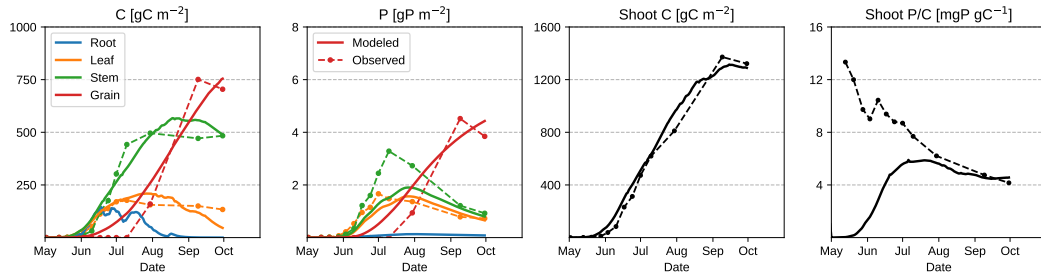


Figure 4.5: Tartas observed vs. modelled time-series for block 4 in the highest P level experiment.

The model is able to predict final plant response to P availability (Fig. 4.6) although biases exist due to the mentioned  $LUE \times PUE$  interaction. Also, most of the observed data falls in the luxury uptake range, which does not permit the model to be constrained in truly P limited conditions. This can best be seen by underestimated P concentration in vegetative organs at high P availability (SI Fig. B.13). The RSR is much lower than observed (around 0.2 in Amos and Walters, 2006) because we model actively absorbing roots. As the modeled plant leaf P can support additional growth after the maximum growth point, the plant stops supplying root C and they quickly die off (Fig. 4.5, left-most panel).

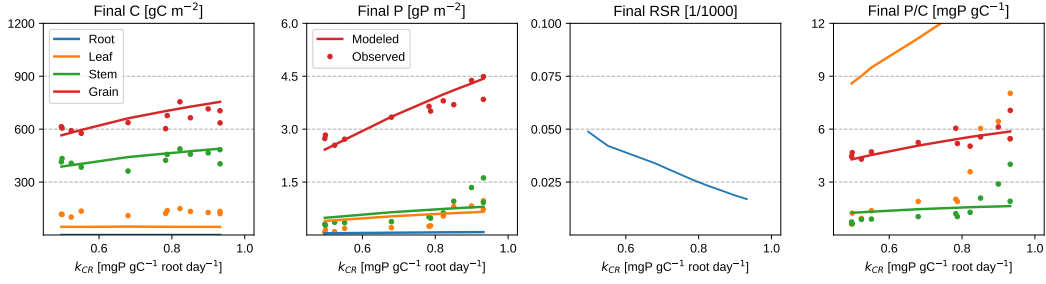


Figure 4.6: Tartas response to P availability as function of root P uptake rate  $k_{CR}$ .

### 4.3.3 Comparison between Nebraska and Tartas

Looking at parameter values after calibration (Table 4.2) most of them are surprisingly close, with differences stemming from the employed maize variety and the nature of the data used (containing information on either plant P uptake or NEE, apart from above-ground biomass). In Tartas, parameters related to the NEE ( $V_{C,max}$ ,  $\lambda_{T,i}$  and  $C_{L,max}$ ) acquire higher values but with organ growth remaining similar (Fig. 4.5, upper left) since the increase in C assimilation is compensated by increased respiration. If we were to have information on Tartas NEE, these values would be better constrained and probably decrease as in Nebraska. In Nebraska, a similar effect can be found for certain P parameters ( $f_{Pi}$  and  $k_{mPi}$ ) which acquire much lower values, since no information exists to directly constrain them.

### 4.3.4 Comparison with original ORCHIDEE

Even though the growing season  $GPP$  and  $R_{plant}$  averages are similar across the two versions, their cumulative balance amounts to a significantly different plant biomass (Fig. 4.7, top row). The reason for this can be either the way  $LUE$  is parametrized (exponential vs. Michaelis-Menten) or how C is allocated (prescribed vs. optimal). It is very likely that the allocation scheme is the culprit though, since the average  $GPP$  across latitudes deviates much more than  $R_{plant}$  (Fig. 4.7, top row). This means the average productivity is the same for either the exponential or Michaelis-Menten  $LUE$  function.

The reason for the  $R_{plant}$  differences stems from the way C is allocated in the original version, which might not be the most effective in the sense of plant productivity. This can clearly be seen in  $C_{plant}$  and  $C_{leaf}$  differences where (in spite of a larger final  $C_{leaf}$ ) the original version produces a smaller plant (Fig. 4.7, two right columns). This might be improved by calibrating the different terms in the Friedlingstein et al. (1999) allocation scheme which was not done here (Eqs. 4.4 - 4.6).

Due to  $H_2O$  stress removal in both of the versions, plant biomass could further increase towards the equator whereas observed potential maize yield shows a sharp decline below  $40^\circ$  latitude (SI Fig. B.16). Similarly at higher latitudes, simulated maize yield decreases

Table 4.2: Comparison of Nebraska and Tartas calibrated values. Only minimum and maximum  $k_{CR}$  values are given for Tartas. Refer to SI Tables B.8 and B.9 for the full parameter sets and their error estimates.

Variable	Description	Units	Initial guess	Nebraska	Tartas
$V_{C,max}$	Max. carboxylation rate	$\mu\text{mol C m}^{-2} \text{s}^{-1}$	60	49.3	94.2
$SLA$	Specific leaf area	$\text{m}^2 \text{gC}^{-1}$	0.03	0.031	0.032
$\lambda_0$	Respiration rate at 0 C°	$\text{day}^{-1}$	0.03	0.026	0.041
$\lambda_{TL}$	Leaf respiration temperature slope	C° <sup>-1</sup>	0.02	0.063	0.034
$\lambda_{TR}$	Root "	C° <sup>-1</sup>	0.02	0.014	0.083
$\lambda_{TS}$	Stem "	C° <sup>-1</sup>	0.005	0.005	0.005
$\lambda_{TG}$	Grain "	C° <sup>-1</sup>	0.0001	0.0002	0.0006
$\lambda_{TX}$	Reserve "	C° <sup>-1</sup>	0.0001	0.0002	0.0010
$k_{CR}$	Root P uptake rate	$\text{mgP gC}^{-1} \text{root day}^{-1}$	0.10	0.99	0.50 - 0.93
$C_{L,max}$	LUE half-saturation biomass	$\text{gC m}^{-2}$	60	52.1	98.9
$\rho_{L,max}$	PUE half-saturation concentration	$\text{mgP gC}^{-1}$	1	3.28	3.54
$f_{CG}$	Grain C filling fraction	-	2	1.16	1.95
$f_{CS}$	Stem C "	-	1	0.15	0.80
$f_{PR}$	Root P "	-	0.1	0.15	0.07
$f_{PS}$	Stem P "	-	0.1	0.89	1.21
$k_{mPL}$	Leaf P remobilisation rate	-	0.01	0.002	0.016
$k_{mPR}$	Root P "	-	0.01	0.002	0.009
$k_{mPS}$	Stem P "	-	0.01	0.019	0.016
$t_{ini}$	Planting date	day of year	120	132	121

much more than the observed one (SI Fig. B.16) which could be due to parameters being suited for temperate climates only.

## 4.4 Discussion

Simulating plant growth using DGVM parametrization of photosynthesis and respiration (Krinner et al., 2005; Matthieu Guimberteau et al., 2018) while applying the OF principle to allocate C and P (Dewar et al., 2009), a generic framework for simulating P limitation in crops is provided. Here, we calibrate our model using two different irrigated maize field trials (in Nebraska and Tartas) and show the DGVM-OF combination to be largely capable of capturing climate and nutrient management effects in irrigated monoculture systems. But in spite the model’s ability, plenty of room is left for its improvement.

Firstly, partial data availability prevents us from properly constraining the model. In Nebraska, information on organ P accumulation is missing whereas in Tartas, it is the plant respiration one (through NEE). This results in significantly different model parameter values (Table 4.2) which might also not be realistic. Two options present itself to counteract this: either by constraining the parameter search space, or initiating values for one experiment by the other one after it has been calibrated. Both of these approaches are not truly desirable, as they predetermine model performance in a self-fulfilling manner. The sole

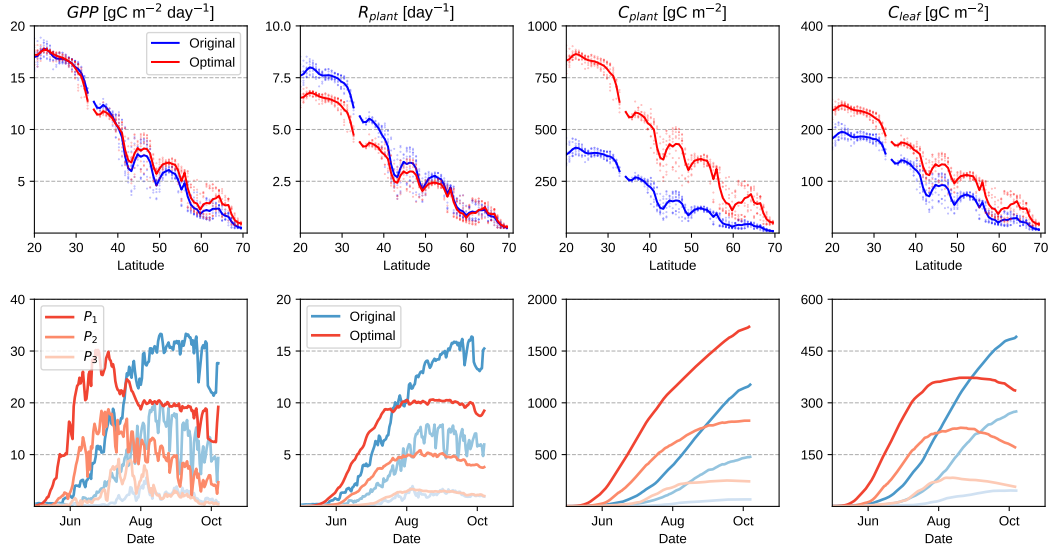


Figure 4.7: Mean latitude response of time averages (upper row) and the time series across 3 different locations (lower row) of original vs. optimally allocated ORCHIDEE.

reason why (for example) Tartas  $V_{c,max}$  and  $\lambda_i$  parameters acquire much higher values is because the model's net C balance is the main determinant of plant biomass (Chapter 3). Since there is no direct information to constrain neither plant respiration nor leaf assimilation in Tartas, the calibration software naively finds the most error minimizing parameter combination which can lie well outside the range of normally observed values. We surmise the same issue would happen with the original ORCHIDEE version, but which is yet to be seen as the calibration protocol was not followed in its case (detailed more in the 4th paragraph below this one).

The arbitrary nature of calibrated parameter values would seem concerning, as the point of a DGVM is to explain growth in a universal manner, and not rely on an experiment-by-experiment basis. But as highlighted before, data availability is the true constraint. Even though plenty of crop growth references exist, we have found only B. Peng et al. (2018) to provide information necessary for DGVM calibration (GPP and NEE). A similar conclusion goes for the P uptake dataset (Plénet et al., 2000) which was communicated internally. The need for appropriate and quality data is unquestionable (van Ittersum et al., 2013) and only when one has access to it, can the results truly be extrapolated to the global scale.

Most obvious model shortcoming is the lack of a soil P module which we do not include, as its implementation requires a separate model on its own (Barber, 1995; Ringeval et al., 2017; X. Yang et al., 2014). Instead, we model P availability with a constant root uptake rate which is a sufficient requirement to explain basic plant response (Iwasa and Roughgarden, 1984). To make our model more mechanistic (or to account for soil properties and its P status) we could use a saturating function of soil P concentration (provided in the Tartas

dataset). But the addition of a few soil related parameters would not fundamentally change how our model behaves, since plants should respond to what they can acquire, irrespective of the composition or the chemical profile of the soil underneath. Additionally, long term plant-soil P cycling demands a higher degree of soil model complexity such as had been done for natural vegetation in ORCHIDEE (Goll et al., 2017), but which was not used as a starting point for this work.

Even though versions of ORCHIDEE including crop LAI phenology (X. Wu et al., 2016) and the P cycle (Goll et al., 2017) exist, we chose not to build on top of them as the underlying assumptions do not fit OF principles. In the crop LAI version (X. Wu et al., 2016) a typical crop-modelling approach is employed (Wallach, Makowski, and J. W. Jones, 2006) where a potential LAI curve is prescribed as a function of a temperature sum (or growing degree days). Even though it makes for a robust empirical description of LAI growth, the plant's net C balance and growth are calculated independently from it because LAI is a completely separate entity. In the OF approach, leaves and LAI (via SLA) are grown as a response to plant needs and resource availability. For the P cycle version (Goll et al., 2017) the main issue is the different conceptual approach to allocation, where C is distributed in order to maintain an optimal leaf concentration (Eq 18, Text S1, S. Zaehle and Friend, 2010) instead of it being a product of plant's internal limits to productivity; or what OF tries to achieve when maximising leaf productivity as a function of leaf P.

Comparison between the original and the OF version was done, with only differences due to the *LUE* function (exponential vs. Michaelis-Menten) and C allocation scheme (Friedlingstein et al., 1999). To make the two even more comparable, the *LUE* term (Eq. 4.10) should be modeled with a saturating exponential in the optimal version, and the various allocation constants calibrated in the original one (Eqs. 4.4 - 4.6) apart from the assimilation and respiration ones. We surmise that (in the original version) the plant C balance should be captured much better, but not the P one. Firstly, there are no allocation rules to distribute P among different organs, although this can be dealt with fixed P allocation fractions like we do here. But more importantly, the root allocation fraction in Eq. 4.4 will result in P uptake regardless of plant needs. How this will affect plant P concentration and productivity is yet to be seen.

Biases and time offsets in the modeled result stem from model misrepresentation. For example, the grain filling mechanism (Eq. 4.21) prescribes a grain C pool that exists since growth initiation, whereas a more appropriate way would be to trigger it during the season like in Iwasa and Roughgarden (1984). For this purpose, linear programming (Dantzig and Thapa, 1997) could serve as a suitable method since it allows optimization of discrete values in order to simulate decisions (1 for yes and 0 for no). But how to implement this strategy consistently within the OF framework requires more attention.

Another model misrepresentation comes in form of underestimating plant P concentration during initial growth (Fig. 4.5). We found this to be due to the *PUE* term which

depends on leaf concentration (Eq. 4.11). If the generic  $Y_L$  productivity element (Fig. 4.2) in the  $PUE$  term is described as the total leaf P amount or  $P_L$  ( $\text{mgP m}^{-2}$ ) instead of leaf P concentration ( $\rho_L$ ), leaf P concentration settles to a constant value which does not change in alternative growing conditions and could be closer to the observed ones (SI Fig. B.15). The main issue with this approach is the apparent disconnect of leaf C and P (from a physiological standpoint) as both contribute to leaf productivity separately. But this just shows that OF principle does not exactly describe the specific internal mechanisms of plant adjustment. Instead, it tries to describe the general principle of matter flow and organization (C and P) that insures maximum plant productivity.

Introducing variable planting dates would decrease model error in Nebraska, but the discrete nature of the planting date variable (expressed in days) introduces difficulties for the calibrating software (E. Jones, Oliphant, and Peterson, 2001). Since days are integers, the modelled error stays the same once the searching step falls below 1 day. When this happens, other variables (continuous) are adjusted towards a lower error value and usually converge to a combination that is not the most optimal one. This issue becomes compounded once 10 different planting days are introduced. But in spite of a constant planting date across all years, the model still provides reasonable results (Fig. 4.4)

Self shading effects were modeled via the  $LUE$  function (Eq. 4.10) which mimics the decreasing efficiency of light capture in an expanding canopy. A more mechanistic way of describing this phenomenon would be to simulate vertical discretization via multiple stem and leaf C pools, with different layers capturing less light the further they are away from the canopy top. This would not only simulate the Beer-Lambert law more faithfully (Eq. 4.1), but would also allow the optimization principle to extend to the vertical organization of C as leaves and stem would be grown to capture C in the most efficient manner. Planting density effects could also be modeled via changes to the light attenuation coefficient, although the exact details on how to go about it (apart from an empirical relation) need to be investigated.

One of the most important mechanisms that is overlooked is the impact of  $\text{H}_2\text{O}$  availability. We avoid this by turning off  $\text{H}_2\text{O}$  limitation effects due to the potential complexity of the underlying regulation processes (Hsiao, 1973) that are not easily extendible to the OF principle in current ORCHIDEE form. The way this is done is by solving a set of equations that couple leaf photosynthesis (Farquhar, von Caemmerer, and Berry, 2001; Collatz, Ribas-Carbo, and Berry, 1992) and stomata opening (Ball, Woodrow, and Berry, 1987) while being forced with an empirical stress factor based on soil  $\text{H}_2\text{O}$  availability (Eq. A11 in Krinner et al., 2005). To explain in more detail, current soil  $\text{H}_2\text{O}$  availability determines the  $\text{H}_2\text{O}$  limited  $V_{C,max}$  and the stomatal conductance slope. The  $\text{H}_2\text{O}$  limited  $V_{C,max}$  is then used to solve the photosynthesis-stomata equations to calculate leaf assimilation after which C is assimilated,  $\text{H}_2\text{O}$  transpired and soil  $\text{H}_2\text{O}$  updated. This makes for an elegant way to couple  $\text{H}_2\text{O}$  transpiration and C assimilation, but does not truly provide

feedback for root growth as it is prescribed separately (Eqs. 4.4 - 4.6). Though, the lack of a dynamic root feedback seems to be a common thread in most of today's DGVMs (Farrior et al., 2015).

To make the plant H<sub>2</sub>O response in line with OF principles, the costs and benefits of different actions need to be taken into account. To model this, internal plant H<sub>2</sub>O status needs to be tracked, to which leaf assimilation can be tied. This presents two options to affect plant productivity: either by increasing plant H<sub>2</sub>O potential through root growth or maintaining it through stomata closure. The first response would be triggered in increased CO<sub>2</sub> environments, where the fertilization effect supplies extra carbon for root growth, increasing the internal H<sub>2</sub>O potential via enhanced H<sub>2</sub>O uptake. The second response would be triggered in abundant H<sub>2</sub>O conditions when stomata closure reduces plant H<sub>2</sub>O intake, and where the assimilated C needs not be wasted on root growth. Both of these mechanisms serve to maintain productivity in conditions where either of the two resources are abundant or lacking, and make for a more consistent description of plant response to H<sub>2</sub>O availability. This approach should also be in line with the work of I. C. Prentice et al. (2014) who model stomatal response as a product of minimizing carboxylation and transpiration costs in leaves, and are able to predict leaf traits in differing climates.

Furthermore, effects of root architecture could be introduced from a functional standpoint, as different root orders have specific tasks in resource acquisition (McCormack et al., 2015). Explained in the most general manner (based on Fig. 4 in McCormack et al., 2015), low root orders are tasked with resource absorption and have high maintenance costs, whereas higher root orders are tasked with resource transport and have progressively lower costs. ORCHIDEE describes the temperature and H<sub>2</sub>O soil profile across 11 layers in the improved version (Matthieu Guimberteau et al., 2018), which permits the introduction of root functional role effects across it. Multiple root orders for each soil layer could be simulated, with differing transport and absorptive capacities while connecting them vertically through allometric ratios (similar to the stem-leaf ratio employed in Eq. 4.20). This could be possible for H<sub>2</sub>O but not for P, due to the bulk nature of currently implemented soil P models (X. Yang et al., 2013; Ringeval et al., 2017; Goll et al., 2017).

To conclude, OF provides a simple and a robust way to integrate different mechanisms of plant adjustment to its growing environment (Dewar et al., 2009). By combining it with a DGVM that contains mechanistic parametrization of main plant growth drivers (Krinner et al., 2005), a consistent framework is provided for simulating historical and future scenarios of global crop management. This can be applied to the issue of P in global cropland (MacDonald et al., 2011; Ringeval et al., 2017) and its contribution to limits in world crop production investigated.

## Chapter 5

# Discussion

### 5.1 Results summary and their place in current research

In this work we have tried to tackle the issue of P availability on worldwide crop production, with special focus on three staple cereals: maize, wheat and rice. While we might have fallen short of our initial goals (Chapter 1) of which the crown one was separating P' contribution to global crop production today and in the future, we have provided an initial glimpse into the matter (Chapter 2) and followed up with necessary simulation tools to study the issue in a more complete and fundamental manner (Chapters 3 and 4).

Our initial work diagnosed today's P limitation in cereals worldwide, where existing simulated estimates of soil available P (Ringeval et al., 2017) were combined with potential crop growth (X. Wu et al., 2016) through a more realistic description of soil-plant P transfer (De Willigen and van Noordwijk, 1994). These results showed, albeit in a constrained manner, the significant potential of existing soil P reserves to limit global cereal production; in line with (MacDonald et al., 2011; Mueller et al., 2012) and as an extension of research up to that point (Bouwman et al., 2013; Ringeval et al., 2017). The obtained results contained considerable uncertainty due to poorly constrained initial data sources (Ringeval et al. 2017; and similar to Folberth et al. 2016) but more importantly, key issues in form of process misrepresentation were identified that need to be taken care of when coupling the two models (soil and crop one). Since models of soil P dynamics and root uptake already contained a sound conceptual basis, we've decided to focus on plant adjustment to P limitation as it should have an impact on crop development and ultimately grain yield.

The coupling of soil P availability and plant growth was done by developing an ideal C & P allocation model that shies away from standard (and highly parametrized) ways in crop simulation (Stockle and Debaeke, 1997), and instead relies on more modern and fundamental concepts in plant physiology (Franklin et al., 2012); more specifically the optimal functioning one (Dewar et al., 2009). Using optimal functioning, plant response at different P availabilities has been reproduced, where organ growth and subsequently

yield are a product of internal limits to productivity rather than a simple supply/demand balance (as in Chapter 3). In spite of the allocation model's parsimonious nature, which can not explain the exact nature of C & P flow during a growing cycle, it provides a much simpler and more transparent way of simulating crop P limitation.

The C & P allocation model was then implemented into the dynamic vegetation model ORCHIDEE (Krinner et al., 2005) containing parametrizations of fundamental plant growth processes (Farquhar, von Caemmerer, and Berry, 2001; Ball, Woodrow, and Berry, 1987; Ruimy, Dedieu, and Saugier, 1996), which allows its extension to any kind of growing environment (in principle). The model was evaluated at two different field sites, and has shown potential in reproducing observed growth in spite of key pieces missing; most notably a soil P module. Ways of improving the new vegetation model, as well as the extension of the optimal function principle (Dewar et al., 2009) to other plant-soil-atmosphere processes is proposed (concerning reproductive timing, canopy effects, root functional roles and water use regulation) which could go towards a more biologically and physically consistent description of crop growth. This model would be a useful addition to the already existing attempts (Y. P. Wang, Houlton, and C. B. Field, 2007; Y. P. Wang, Law, and Pak, 2010; Goll et al., 2012; X. Yang et al., 2014; Goll et al., 2017) at deciphering P' role in future eco-systems (Reed, X. Yang, and Thornton, 2015) but one which is focused on crops, as the future of P management and its impact on global food production remain largely uncertain (Cordell, Drangert, and White, 2009; J. Elser and E. Bennett, 2011; MacDonald et al., 2011; Peñuelas et al., 2013; Obersteiner et al., 2013; Makowski et al., 2014).

## 5.2 Plant adjustment effects on global P limitation

To show the potential effect of plant adjustment (Chapter 3) on our global P limitation estimates (Chapter 2), we performed a simulation using the ideal OF allocation model and mapped the responses of either a linear or quadratic yield model to it (Fig. 5.1). The linear model corresponds to the C:P ratio and PHI estimates, while the quadratic model corresponds to the QUEFTS ones (Chapter 3). The OF response curve was obtained using the initial guess calibration values (Table 3.2) and simulating over a range of  $k_{CR}$  values (similar to Fig. 3.3, panel A and B). The biomass potential was chosen as total plant biomass at  $k_{CR} = 1$ , which is far away from P limited range when RSR starts to decrease (Fig. 3.3). The OF yield response curve is then simply scaled across all  $k_{CR}$  values (or P availabilities) with respect to total plant C and P at  $k_{CR} = 1$  (Fig. 5.1).

Taking a much higher  $k_{CR}$  value (eg. ten times higher) puts more weight on more P limited growth (SI Fig. B.17), but changes the OF curve shape only slightly. Furthermore, the conceptual model does not have limits to biomass production as there are no limits to  $k_{CR}$  increase. This can be seen by the increasing OF curve after the  $P_{supply}/P_{demand} = 1$  point in Fig. 5.1. In reality, P diffusion kicks in as a physical limit to P transport from the

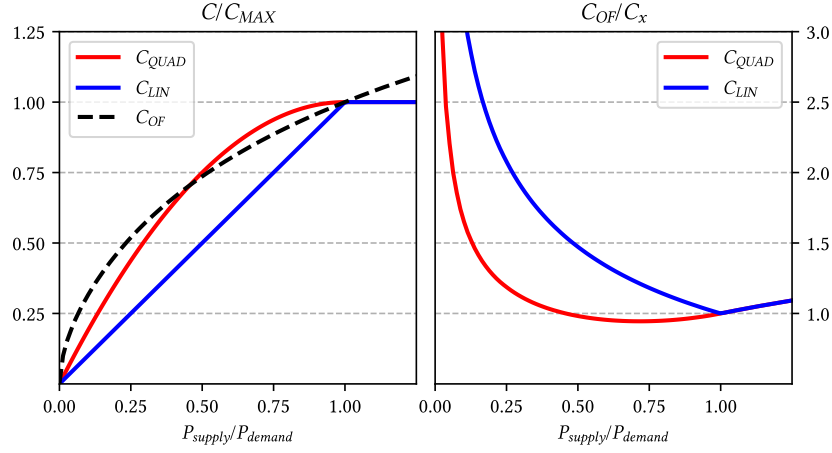


Figure 5.1: Yield limitation curves for a linear, quadratic and OF adjustment model as function of  $P$  supply and demand. Global  $P$  limitation estimates from Chapter 2 follow a similar pattern as the blue ( $C:P$  ratios and  $\Phi$ ) and red (QUEFTS) curves. Black curve describes a response curve following OF allocation as in Chapter 3. Right plot shows the underestimation of the linear or quadratic model relative to the OF one.

soil; as was done in Chapter 2 when calculating the potential root  $P$  uptake. If we were to implement a similar mechanism in the hypothetical model, the OF curve would asymptotically approach the intersection curve instead of reaching it at  $P_{supply}/P_{demand} = 1$  (Fig. 5.1). Similarly, we chose the initial guess values to construct the OF curve solely for simplicity sake. Introducing additional complexity for these rule-of-thumb estimates (which would require significant parametrization and sensitivity testing) should not be necessary since the initial  $P$  limitation estimates already contain significant uncertainty.

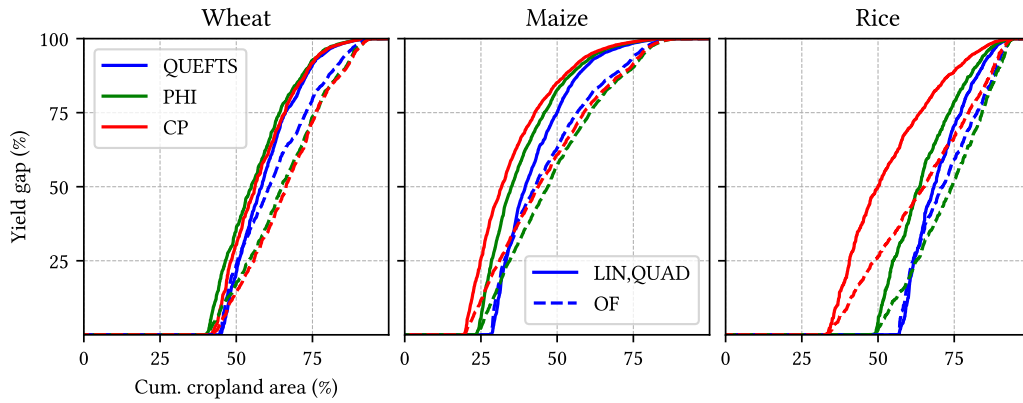


Figure 5.2: Cumulative histogram of the  $P$  yield gap (%) in irrigated simulations as function of potential soil  $P$  supply only. This figure is similar to the bottom row in Fig. 2.6 except for the dashed lines, which (here) represent adjustment due to OF. Lines represent the 1000 MC sample median.

Looking at the cumulative histogram of the  $P$  yield gap (Fig. 5.2), a slight relief can be noticed depending on the way yield limitation was initially modeled. The relief increases in areas with higher yield gaps, consistent with the way the linear or square model over-

estimate limitation relative to the OF one (Fig. 5.1, right). Once these contributions are integrated to form a global average, OF plant adjustment shows a high potential to change the initial estimates (Tab. 5.1). Even though it is not possible to attribute significant differences, OF adjusted maize yield gap lowers about the same when compared to direct fertilizer uptake (or 20% of today’s applied fertilizer; a generous estimate as explained in Chapter 2). Although, the similarity of these two effects is probably due to the global maize cropland extent which encompasses more low P input areas (eg. Africa). In rice and wheat, OF adjustment has a similar or lower effect on reducing the yield gap as cultivated area co-occurs with high P addition spots (SI Fig. B.3).

	Wheat	Maize	Rice
Soil P supply	22 (18-28) %	55 (47-66) %	26 (18-46) %
" + P fertilizer	17 (14-21) %	46 (36-55) %	15 (10-32) %
" + OF adjusted	19 (15-26) %	47 (42-56) %	21 (15-35) %

Table 5.1: Global averages of the P yield gap (%) as function of soil P supply, soil P supply amended with P fertilizer, and soil P supply with OF plant adjustment. Values are medians with 95% interval in the bracket.

Even though the provided global estimates are not methodologically robust, they point out the need for a conceptually deeper description rather than increased model complexity, which often requires additional layers of observable information. In spite of their simplistic nature though, mass balance approaches like these serve as a good motivational base for further research as had been done for N by Hungate et al. (2003) or P by Peñuelas et al. (2013); as was done here.

### 5.3 Why optimal functioning?

The need for a deeper conceptual accounting was invoked, but one which begs a very straightforward question: how to actually go about it? Two options should present themselves here, corresponding to either a bottom-up or or a top-down approach (Franklin et al., 2012). The bottom-up approach tries to describe all the minutiae of the processes involved, and attempts to describe a system’s response from their collective interaction. This approach is typically used in crop simulations and DGVMs, where the individual processes (eg. root uptake, stomate opening, nutrient requirement) are precisely tuned and let loose on the simulated plant during integration. On the other hand, the top-down approach tries to look for an over-arching principle that dictates how the individual elements come together and make the whole system function. Even though a top-down approach might contain a very detailed process description, it is the presence of foresight (or purpose) to plant functioning that sets it apart. For the bottom-up approach, more or less sophisticated rules can be formulated to orchestrate the individual parts’ action, but the fact remains

the simulated plant naively sticks to them in all circumstances. Instead, by looking for a principle that guides plant behaviour in a reasonable way (evolutionarily, physiologically, community-wise or other) a more autonomous system can be created, and which responds to whatever environment it finds itself in.

For the bottom-up approach, a typical example can be shown with STICS (Brisson et al., 2009), a crop simulation model containing much of the same apparatus employed by the other AgMIP model candidates:

$$DELTA I_L(I) = DELTA I_{dev}(I) \cdot DELTA I_T(I) \cdot DELTA I_{dens}(I) \cdot DELTA I_{stress}(I) \quad (5.1)$$

$$DELTA I_{dev}(I) = \begin{cases} \frac{DLAIMAXBRUT_P}{1 + \exp(PENTLAIMAX \cdot (VLAIMAX_P - ULAI(I)))} & \text{if } ULAI(I) < UDLAIMAX_P \\ DELTA I_{max} \cdot \left(1 - \frac{ULAI(I) - UDLAIMAX_P}{3 - UDLAIMAX_P}\right)^2 & \text{if } ULAI(I) \geq UDLAIMAX_P \end{cases} \quad (5.2)$$

$$DELTA I_T(I) = \begin{cases} 0.0 & \text{if } TCULT(I) < TCMIN_P \\ \frac{TCULT(I) - TCMIN_P}{TCMAX_P - TCMIN_P} & \text{if } TCMIN_P < TCULT(I) \leq TCMAX_P \\ \frac{TCMAX_P - TCXSTOP_P}{TCMAX_P - TCMIN_P} \cdot (TCULT(I) - TCXSTOP_P) & \text{if } TCMAX_P < TCULT(I) \leq TCXSTOP_P \end{cases} \quad (5.3)$$

$$DELTA I_{dens}(I) = \begin{cases} DENSITE(I) \cdot \frac{DENSITE(I)^{ADENS_P}}{BDENS_P} & \text{if } LAI(I) \geq LAICOMP_P \text{ and } DENSITE(I) \geq BDENS_P \\ DENSITE(I) & \text{if } LAI(I) < LAICOMP_P \text{ and } DENSITE(I) < BDENS_P \end{cases} \quad (5.4)$$

$$DELTA I_{stress}(I) = \min(TURFAC(I), INNLA I(I)) \cdot EXOLA I(I) \quad (5.5)$$

The equations above prescribe LAI growth (only) as function of temperature, phenology, plant density and stress. The meaning of different parts need not be understood, as their purpose is to illustrate the underlying complexity one often needs to employ to make the different parts work together. But if looked at closer, the equations describe an empirical growth curve (Eq. 5.2) which progressively gets squashed down with each additional stress factor (Eqs. 5.3 - 5.5). These equations can be constrained by well defined datasets and show remarkable skill at reproducing observed yields (C. Müller et al., 2017), but once they are taken outside of the calibration range, they encounter understandable difficulties which are not easily overcome with more data (Asseng et al., 2013; Rosenzweig et al., 2014; Bassu et al., 2014; Folberth et al., 2016; C. Müller et al., 2017). This problem was best summarized by Passioura (1996) who wrote:

*Crop simulation models ... are typically flawed by being based on untestable guesses about the processes that control growth. They may, however, provide useful self-education for their developers.*

To overcome this seemingly arbitrary way of describing growth, where a plant consists of highly interconnected but truly unrelated parts which together try to smother it out of existence (Eq. 5.1 - 5.5), a more sensible manner would be to make a plant somehow “decide” what to do by using the tools it already has. This is what optimal functioning tries to do when maximizing productivity (as a proxy of individual fitness) and subjecting the whole of plant functioning to its achievement (Dewar et al., 2009). In spite of its reasonable basic assumptions, valid questions arise on its theoretical foundation as mentioned

in Dewar et al. (2009): Why should plants optimize productivity, or what should they optimize? Do these optimization goals change in different environments? Or even more, why should a plant optimize anything at all? Of all of the problems optimal functioning faces, the most difficult one is not taking into account competitive interaction between individuals, as it brings into question its feasibility in an evolutionary sense (maximising individual fitness is not necessarily the best survival strategy; Schieving and Poorter 1999; Anten 2005; Anten and During 2011). But in spite of these (still) theoretical considerations, the fact remains it explains plant response as an emergent phenomenon, rather than its prescription through a complex set of scheduling rules (Dewar et al., 2009).

Optimal functioning itself is part of a greater push towards implementing more general plant ecology concepts in vegetation modelling, for the purpose of predicting global ecosystem response (Franklin et al., 2012). The main issue being the over-whelming presence of still very basic concepts behind long-term vegetation change, which obviously need to be improved on when studying climate change impacts (Purves and Pacala, 2008; Ise et al., 2010; Rötter et al., 2011). This is especially true since these models provide the basis for policy implementation on a global scale (see United Nations 2019 for further details). But not to get too ahead of ourselves (or conversely too deep into the vegetation modelling rabbit hole), the main reason for the choice of optimal functioning was its simplicity, which makes prediction and analysis of nutrient impacts on crop growth much more transparent and conceptually satisfying (I. Prentice, 2013; Stocker et al., 2016).

## 5.4 Complexity vs. transparency

Even though a clear methodology path was cut out from the start, where P limitation effects on a global scale were to be studied using soil P cycling and crop growth simulations, the question begs itself: was it really necessary to go into this much mechanistic detail? A good example to illustrate the point is Chapter 2 where, in spite of a better mechanistic representation, the additional sources of uncertainty produce estimates of questionable use. This point goes hand in hand with previous remarks, where increasing model complexity necessitates more and more information. To make our point, we will focus on the root uptake parametrization from Chapter 2 (De Willigen and van Noordwijk, 1994) as it allows easier back-tracing of added uncertainties as opposed to the used soil P dynamics and crop growth simulation models (Ringeval et al., 2017; X. Wu et al., 2016).

The employed root uptake model (De Willigen and van Noordwijk, 1994) is an analytical steady state solution, where nutrient mass transport equation by diffusion and advection is solved, by assuming a homogeneously distributed root mass in form of vertically placed cylinders. The full derivation of the analytical solution can be found in De Willigen and van Noordwijk (1994), so here we will focus on the final utilized solution, where the rate of nutrient uptake is equal to the diffusion transport limit within the root cylinder

termed "zero-sink" uptake:

$$P_{uptake} = \sum_{i=1}^{12} \pi \cdot \Delta z \cdot L_{rv,i} \cdot D \cdot \frac{\rho^2 - 1}{G(\rho)} \cdot C_P \quad (5.6)$$

Additionally, we will repeat all of the different sources used to parametrize each of these terms (explained in detail in SI section B.1.3) to paint a better picture of the additional complexity employed: Olesen et al. (2001), Hengl et al. (2014), Achat et al. (2016), Duivenbooden, Wit, and Keulen (1995), Hocking (1994), Latshaw and Miu (1934), Ye et al. (2014), Nakhforoosh et al. (2014), Li et al. (2016), and Biscani, Izzo, and Yam (2010). Again, this is for root uptake only... Not surprisingly then, do the final estimates show such high spread (Table 5.1) as they also include uncertainty due to the soil P and the crop simulation models (Ringeval et al., 2017; X. Wu et al., 2016). If looking more closely at Eq. 5.6 the main drivers of P uptake are: root mass (through root length density  $L_{rv}$ ), soil water content (through the diffusion coefficient  $D$ ) and the amount of plant available P in soil (through the empirical relation for soil water P concentration  $C_P$ ). This boils down to the following equation:

$$P_{uptake}(C_R, P_{lab}, \theta) = k \cdot C_R \cdot P_{lab} \cdot \max(\theta / \theta_0, 0) \quad (5.7)$$

Where  $C_R$  is root biomass (tDW ha<sup>-1</sup> or gC m<sup>-2</sup>) and  $P_{lab}$  is plant available P in soil (kgP ha<sup>-1</sup> or mgP m<sup>-2</sup>), while  $\theta$  and  $\theta_0$  are soil water content and its diffusion cut-off point (-).  $k$  is a species and soil specific conversion factor that computes  $P_{uptake}$  (kgP ha<sup>-1</sup> or mgP m<sup>-2</sup>) as a measure of crop-soil specific P uptake efficiency, and is not entirely different from the bibliography determined coefficients. With this in mind, one can simply ask: but how do we generalize this kind of model to the global scale? An honest answer would be that you can not. And even more, you probably should not; at least not in this methodological form. Jumping from the field scale to the global one by relying on observed data only, not matter how precise it might be, requires a bit of scientific faith. This was best illustrated by one of the Referees for the Chapter 2 submitted paper who responded:

*This model-based examination ... depends heavily on the basic P geochemical data ... plus a whole set of P cycling and plant utilization equations. I agree with the former quantification issues, but believe that the authors did what they could... And for the latter modeling-heavy aspects, particularly those dealing with plant utilization, I admit to being largely clueless*

But on the other hand, using a reduced model form (like the one in Eq. 5.7) should allow us to decipher robust observation datasets like FAOSTAT (2019) in a more understandable fashion. Instead of adding more processes which require more parametrizing information and obscure the path to the final result, maybe it would be prudent to just parametrize a much simpler model using large scale data? Because, when looking at these transfer

relationships on a regional to global scale, the variety of soil conditions and crop species can not in any way be reproduced with these very precise, but still very specific estimates. Instead, we should use the observations to constrain a kind of a macro property which, even though it should contain significant uncertainty, would more robustly reflect the process of soil-plant P transfer at these scales. This process simplification should also be preferred, not because simplicity equates to more profound theoretical value (Webb, 1996), but since it makes any subsequent analysis more reproducible and falsifiable (Hoffmann, Minkin, and B. K. Carpenter, 1997; Baker, 2007).

## 5.5 Global P limitation using simpler models

One of the more appropriate approaches to estimating P limitation on a regional to global scale using dynamic simulation is the one of S. Z. Sattari et al. (2012). The principal aim of this work was estimating the amount of P stored in agricultural soils, as a consequence of P management during the late 20th century. This was attempted by describing a very simple soil cycling model consisting of two P pools (Fig. 5.3) that are mainly driven with census data on fertilizer P application, crop yield and harvested area (FAOSTAT, 2019) with predefined assumptions regarding P runoff/erosion, soil P weathering, P deposition and organic P use (S. Z. Sattari et al., 2012). Very robust estimates of soil P accumulation were obtained using this approach, but more importantly, the hysteretic nature of P buildup and subsequent release were reproduced (Fig. 4 in S. Z. Sattari et al. 2012), normally seen by the sudden increase in P application efficiency during the 1990s (Fig. 1.1).

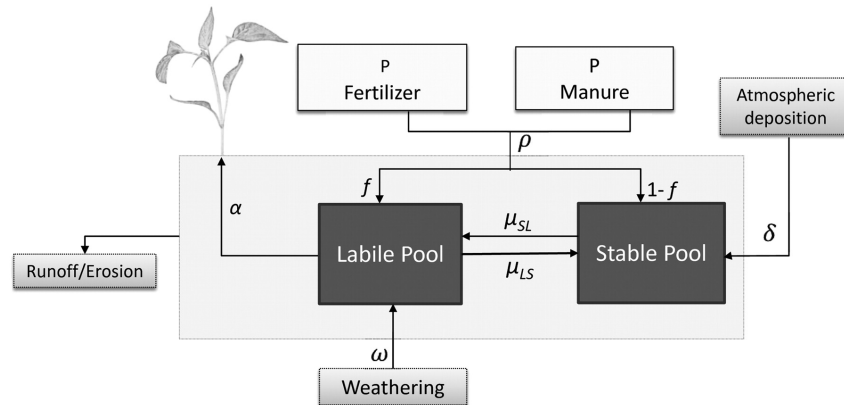


Figure 5.3: The soil P cycling model used in S. Z. Sattari et al. (2012). This scheme was adapted from Figure 1 in the original paper, where details on each of the terms can also be found.

To extend this model to the study of yield limitation, an additional crop yield biomass pool could be added and connected to the labile P one using a saturating function (similar to J. Zhang et al. 2017) instead of a simple removal coefficient (as was done via  $\alpha$  in Fig. 5.3). In this manner, more mechanistic representation of P transfer and use would implicitly be modeled (like root uptake via P diffusion and changing plant P use efficiency) to allow a

more faithful description of yield limitation (like in Fig. 5.1). Once a basic model form is established, parameter inference methods should preferably be utilised (Geyer, 1992) instead of relying on literature parameter values and errors (a common practice in mass balance studies) since the reduced model necessarily bundles certain processes together. But more importantly, parameter estimates and their hypothetical distributions would be obtained, in order to estimate future trends and uncertainties in P fertilizer use and impact on yields; the main thing we are interested in when studying P limitation on a global scale.

Modelling complex biogeochemical and climate-atmosphere phenomena using reduced-form models is an established practice in the field of climate modelling (Claussen et al., 2002; Ricciuto, Davis, and Keller, 2008; Meinshausen, Raper, and Wigley, 2011; Rogelj, Meinshausen, and Knutti, 2012). The main motivation behind is the discrepancy in the simulated output of different general circulation models (or GCMs; Ricciuto, Davis, and Keller 2008) which, due to their often profound complexity and a very high computational cost, makes sensitivity or uncertainty analysis of their results impractical or even unfeasible (Claussen et al., 2002). Reduced models try to step in by describing the system (eg. global climate or C cycle) with a limited amount of spatially aggregated state variables (like surface vegetation/soil C pools, atmospheric temperature, CO<sub>2</sub> concentration) which together interact through purposefully (more than less) parametrized interactions (Claussen et al., 2002). All of this is done with the aim of reproducing systemic feedbacks in a more intelligible and computationally faster fashion (Gasser et al., 2017), as well as provide robust uncertainty estimates for the design of economically efficient future C management strategies (Ricciuto, Davis, and Keller, 2008; Rogelj, Meinshausen, and Knutti, 2012).

Apart from the previously mentioned “simple model” approach, statistical treatment is still the most straight-forward way to understanding aggregated effects of agricultural practice on global yields (Wallach, Makowski, and J. W. Jones, 2006; D. B. Lobell, Cassman, and Christopher B. Field, 2009; Mueller et al., 2012; Lassaletta et al., 2014). But due to the variety of physical and biological processes involved, not to mention their interaction with other major growth limiting resources (like H<sub>2</sub>O or N), the soil-plant P cycle should still remain a tough problem to solve. A way around this issue could be to employ empirical methods like machine-learning (ML) that can handle inherent complexities of ecological systems (time lags, historical legacy effects, non-linearities, etc.) that traditional statistics have a hard time of doing (Olden, Lawler, and Poff, 2008). Explained in the most simple manner, ML makes predictions based on already existing patterns in observations, without actually prescribing explicit functional relations to do so (Alpaydin, 2010). Instead, ML relies on sheer data quantity to approximate the underlying processes (as well as their internal links) to produce a much more simple model; a process conveniently called “data mining” or an analogy to the act of processing huge quantities of material to produce minute amounts of precious ore (Alpaydin, 2010). Even though many different aspects of

ML methodology exist (Alpaydin, 2010), the most appropriate one for studying P limitation would be “supervised learning” or ML regression, which works by “training” an ML model on a known subset of data and using its predictions to extrapolate further (Alpaydin, 2010).

To apply this approach to the issue of P limitation, we would obviously need to compile a large database of nutrient limitation experiments similar to Valkama et al. (2009), van der Velde et al. (2014) or Tonitto and Ricker-Gilbert (2016). The input variables should be ones that have a direct impact on P availability (soil quality & nutrient level, fertilizer applied, crop species, irrigation intensity, temperature, precipitation, etc.), while the output ones should simply be grain yield and quality (or nutrient concentration). Once this model is “trained” on a subset of the initial dataset, its capacity to explain the rest should be tested as well as the uncertainty along it. Once the ML model is verified, global maps of the input factors could be used (Monfreda, Ramankutty, and Foley, 2008; Portmann, Siebert, and Döll, 2010; X. Yang et al., 2013; Bouwman et al., 2013; Hengl et al., 2014) and extrapolate the ML model results over a wide geographical area. This point is one of the more challenging ones for ML (as well as for statistical approaches) as global information on some of the input factors are either not available (eg. irrigation intensity) or based on database extrapolation techniques (basically, all of the datasets in the previous sentence). Another foreseeable issue is initial data consolidation, as various (more or less) standard measuring techniques are employed. This is best pictured with the numerous amount of ways to measure soil P availability: Bray and Kurtz (1945), Nelson (1953), Olsen (1954), Hedley, J. W. B. Stewart, and Chauhan (1982), and Morel et al. (2000).

ML regression techniques have already proven useful in extrapolating point observation sites (or often very rich sources of information) to wider spatial extents by using remote sensing products. To cite a few, Mascaro et al. (2014) map tropical forest soil C stocks in North Peru with sparse field-scale data by using Landsat products (Woodcock et al., 2008). F. Yang et al. (2007) map FLUXNET GPP estimates (Baldocchi et al., 2001) across the whole North American continent using MODIS products (Justice et al., 1998). Stocker et al. (2018) uses similar sources (MODIS and FLUXNET) but quantifies light use efficiency across the whole globe (or the ratio of GPP to light intensity). Both MODIS and Landsat contain crop distribution products, which could be (for example) combined with FLUXNET cropland sites (of which there are about a dozen; Fluxnet Network 2019). Even though these eddy-covariance sites focus much more on surface C and H<sub>2</sub>O exchange unlike agronomic ones, which usually provide better info on soil-plant nutrient status and yield (needed for estimating P limitation), the existing ML methodology should be transferable if a suitable compilation of the two is made. At least, ML could be more desirable as it seems to traverse spatial scales with much more grace (Mascaro et al., 2014; F. Yang et al., 2007; Stocker et al., 2016).

## 5.6 Global productivity, but also feasibility and impact

Throughout most of this work (if not all of it) the focus was put on managing productivity, with two obvious factors left out: economic feasibility and environmental impact. These two issues should form the basis of any kind of scenario analysis, not only because they are the basic motivating points of our work, but also because they represent true limits to implementing any kind of P management policy (E. M. Bennett, S. R. Carpenter, and Caraco, 2001; Cordell, Drangert, and White, 2009; P. M. Vitousek et al., 2009; Peñuelas et al., 2013; Obersteiner et al., 2013).

In Chapter 1 we mention that majority of today's fertilizer P comes from a potentially finite source, which is mainly due to the ever increasing costs of exploring lower quality reserves in the future (Cordell, Drangert, and White, 2009; Fixen and Johnston, 2012; Kauwenbergh, M. Stewart, and Mikkelsen, 2013). What is also quite concerning, is that P availability to any country depends on its current economic value, determined by the world's commodity market (de Ridder, 2012; Heckenmüller, Narita, and Klepper, 2014). It is not surprising that the founding work of most of today's global P impact research (Cordell, Drangert, and White, 2009) coincides with the 2008 spike in P fertiliser price (Khabarov and Obersteiner, 2017). Specific actors and circumstances can be identified and analyzed (Khabarov and Obersteiner, 2017), but the fact remains that world production of P is controlled by a handful of countries (or a global cartel; Heckenmüller, Narita, and Klepper 2014), and naturally prompts serious policy discussion on regional and global scale (de Ridder, 2012; Weber et al., 2014; Rosemarin and Ekane, 2016).

Similar to how economical concerns have reignited our interest in the future of global P, environmental impacts have been the main driver of global P impact research in the past; namely eutrophication of H<sub>2</sub>O bodies (Daniel, Sharpley, and Lemunyon, 1998; Correll, 1998; Smil, 2000). Today, the discussion on P' environmental footprint revolves mainly around food quality (as fertilizer P contains significant amounts of cadmium, a toxic trace element; de Ridder 2012) and especially the future of sustainable P use. While many quite similar works exist on this topic, the majority agree this is to be done by reducing existing losses in the P supply chain, changing our dietary habits and lessening our dependence on additional P inputs through recycling (Childers et al., 2011; Neset and Cordell, 2012; Chowdhury et al., 2017).

Taking all of these daunting issues into account and connecting them to the work presented here, a logical step would be to implement our work into so called Integrated Assessment Models (IAMs; Antle and Capalbo 2001; Weyant 2017; Ackerman et al. 2009). IAMs attempt to quantify and reproduce complex natural and human systems, as well as their interaction (eg. climate change and global agriculture), by integrating different process-based models of their underlying drivers (be they physical, biological, economical, political or otherwise; Antle and Capalbo 2001). Their final goal is to simulate behaviour of these systems outside the observed data range, in ways that are consistent with estab-

lished scientific understanding (Antle and Capalbo, 2001). This translates most obviously into simulating the effects of global strategies like: lowering greenhouse gas emissions, climate change adaptation and geo-engineering (Weyant, 2017). Most importantly though, they are used to evaluate the cost and benefit of various climate policy options to devise the most optimal ones for the coming future (Ackerman et al., 2009). These models (like any other ones) do not come without criticism. It seems the biggest part of it focuses on the underlying economical hypotheses, which de-value near-term impacts, downplay uncertainty of expected damage and over-estimate mitigation costs (Ackerman et al., 2009; Pindyck, 2013).

But not to linger on the IAM criticism (primarily because it surpasses anything touched upon in this work) and come back to the issue of P limitation, we would like to address the potential structural IAM issues through an unfortunately specific example; GLOBIOM (Ermolieva et al., 2015) or more specifically its crop module EPIC (Williams et al., 1990; Balkovič et al., 2013; Folberth et al., 2016). Like the previously mentioned crop simulation model STICS (Brisson et al., 2009), EPIC contains a similar empirical backbone (Williams et al., 1990) which makes its use in future projections rather questionable as mentioned (Passioura, 1996; Rötter et al., 2011; Folberth et al., 2016; C. Müller et al., 2017). Using much simpler models of yield formation, the ad-hoc parametrization of EPIC simulations (Christoph Müller et al., 2019) can be uncovered in form of regional fitting coefficients which exist to improve the model's predictive performance (Bruno Ringeval, personal communication). While this technique should not be immediately dismissed as fitting observations can help us better understand the current system state (eg. cultivar diversity) the need for such an intervention puts the entire underlying methodology into question (Rötter et al., 2011; Passioura, 1996; I. Prentice, 2013); especially if it forms the base of any kind of future impact study (Stolbovoy, Montanarella, and Panagos, 2007; Sylvain Pellerin et al., 2019).

While the previous criticism (or in fact, most of it up to this point) might sound quite harsh or borderline megalomaniac, its purpose is not to single out any single methodology over the other as the end solution to all problems. Instead, the purpose is to point out issues which can and should be dealt with possibly simpler, and definitely much more transparent approaches.

## Appendix A

## Bibliography

- Achat, David Ludovick et al. (Jan. 7, 2016). "Soil Properties Controlling Inorganic Phosphorus Availability: General Results from a National Forest Network and a Global Compilation of the Literature". In: *Biogeochemistry* 127.2-3, pp. 255–272. ISSN: 0168-2563, 1573-515X.
- Ackerman, Frank et al. (Aug. 2009). "Limitations of Integrated Assessment Models of Climate Change". In: *Climatic Change* 95.3-4, pp. 297–315. ISSN: 0165-0009, 1573-1480.
- Ågren, Göran I. (Dec. 2008). "Stoichiometry and Nutrition of Plant Growth in Natural Communities". In: *Annual Review of Ecology, Evolution, and Systematics* 39.1, pp. 153–170. ISSN: 1543-592X, 1545-2069.
- Ågren, Goran I and J Fredrik Wikstrom (1993). "MODELLING CARBON ALLOCATION—A REVIEW". In: *New Zealand Journal of Forestry Science*, p. 11.
- Ågren, Goran and Oskar Franklin (Dec. 2003). "Root : Shoot Ratios, Optimization and Nitrogen Productivity". In: *Annals of Botany* 92.6, pp. 795–800. ISSN: 0305-7364.
- Alpaydin, Ethem (2010). *Introduction to Machine Learning*. 2nd ed. Adaptive Computation and Machine Learning. OCLC: ocn317698631. Cambridge, Mass: MIT Press. 537 pp.
- Amanullah and Inamullah (Mar. 2016). "Dry Matter Partitioning and Harvest Index Differ in Rice Genotypes with Variable Rates of Phosphorus and Zinc Nutrition". In: *Rice Science* 23.2, pp. 78–87. ISSN: 16726308.
- Amos, B. and D. T. Walters (2006). "Maize Root Biomass and Net Rhizodeposited Carbon". In: *Soil Science Society of America Journal* 70.5, p. 1489. ISSN: 1435-0661.
- Andersen, M.S., J. Dahl, and L. Vandenberghe (2018). *CVXOPT: A Python Package for Convex Optimization (Version 1.2)*. URL: <https://cvxopt.org/>.
- André, M. et al. (1978). "Daily Patterns under the Life Cycle of a Maize Crop". In: *Physiologia Plantarum* 44.3, pp. 197–204. ISSN: 1399-3054.
- Anten, Niels P. R. (Feb. 1, 2005). "Optimal Photosynthetic Characteristics of Individual Plants in Vegetation Stands and Implications for Species Coexistence". In: *Annals of Botany* 95.3, pp. 495–506. ISSN: 0305-7364.
- Anten, Niels P. R. and Heinjo J. During (Oct. 2011). "Is Analysing the Nitrogen Use at the Plant Canopy Level a Matter of Choosing the Right Optimization Criterion?" In: *Oecologia* 167.2, pp. 293–303. ISSN: 0029-8549, 1432-1939.
- Antle, John M. and Susan M. Capalbo (May 1, 2001). "Econometric-Process Models for Integrated Assessment of Agricultural Production Systems". In: *American Journal of Agricultural Economics* 83.2, pp. 389–401. ISSN: 0002-9092.
- Asher, C. J. and J. F. Loneragan (Apr. 1967). "RESPONSE OF PLANTS TO PHOSPHATE CONCENTRATION IN SOLUTION CULTURE: I. GROWTH AND PHOSPHORUS CONTENT". In: *Soil Science* 103.4, p. 225. ISSN: 0038-075X.
- Asseng, S. et al. (Sept. 2013). "Uncertainty in Simulating Wheat Yields under Climate Change". In: *Nature Climate Change* 3.9, pp. 827–832. ISSN: 1758-678X, 1758-6798.

- Asseng, S. et al. (Feb. 2015). "Rising Temperatures Reduce Global Wheat Production". In: *Nature Climate Change* 5.2, pp. 143–147. ISSN: 1758-678X, 1758-6798.
- Atkin, Owen K. and Mark G. Tjoelker (July 1, 2003). "Thermal Acclimation and the Dynamic Response of Plant Respiration to Temperature". In: *Trends in Plant Science* 8.7, pp. 343–351. ISSN: 1360-1385.
- Bai, Zhaohai et al. (Nov. 2013). "The Critical Soil P Levels for Crop Yield, Soil Fertility and Environmental Safety in Different Soil Types". In: *Plant and Soil* 372.1-2, pp. 27–37. ISSN: 0032-079X, 1573-5036.
- Baker, Alan (2007). "Occam's Razor in Science: A Case Study from Biogeography". In: *Biology and Philosophy* 22.2, pp. 193–215. ISSN: 0169-3867.
- Baldocchi, Dennis et al. (2001). "FLUXNET: A New Tool to Study the Temporal and Spatial Variability of Ecosystem-Scale Carbon Dioxide, Water Vapor, and Energy Flux Densities". In: *Bulletin of the American Meteorological Society* 82.11, pp. 2415–2434. ISSN: 0003-0007.
- Balemi, T. and K. Negisho (Sept. 2012). "Management of Soil Phosphorus and Plant Adaptation Mechanisms to Phosphorus Stress for Sustainable Crop Production: A Review". In: *Journal of soil science and plant nutrition* 12.3, pp. 547–562. ISSN: 0718-9516.
- Balkovič, Juraj et al. (Sept. 2013). "Pan-European Crop Modelling with EPIC: Implementation, up-Scaling and Regional Crop Yield Validation". In: *Agricultural Systems* 120, pp. 61–75. ISSN: 0308-521X.
- Ball, J. Timothy, Ian E. Woodrow, and Joseph A. Berry (1987). "A Model Predicting Stomatal Conductance and Its Contribution to the Control of Photosynthesis under Different Environmental Conditions". In: *Progress in Photosynthesis Research: Volume 4 Proceedings of the VIIth International Congress on Photosynthesis Providence, Rhode Island, USA, August 10–15, 1986*. Ed. by J. Biggins. Dordrecht: Springer Netherlands, pp. 221–224.
- Barber, Stanley A. (1980). "Soil-Plant Interactions in the Phosphorus Nutrition of Plants". In: *The role of phosphorus in agriculture (theroleofphosph)*, pp. 591–615.
- (Apr. 3, 1995). *Soil Nutrient Bioavailability: A Mechanistic Approach*. John Wiley & Sons. 434 pp.
- Barrow, N. J. (Dec. 1, 1983). "A Mechanistic Model for Describing the Sorption and Desorption of Phosphate by Soil". In: *Journal of Soil Science* 34.4, pp. 733–750. ISSN: 1365-2389.
- Bassu, Simona et al. (2014). "How Do Various Maize Crop Models Vary in Their Responses to Climate Change Factors?" In: *Global Change Biology* 20.7, pp. 2301–2320. ISSN: 1365-2486.
- Batten, G. D., M. A. Khan, and B. R. Cullis (1984). "Yield Responses by Modern Wheat Genotypes to Phosphate Fertilizer and Their Implications for Breeding". In: *Euphytica* 33.1, pp. 81–89.
- Bennett, Elena M., Stephen R. Carpenter, and Nina F. Caraco (Jan. 3, 2001). "Human Impact on Erodable Phosphorus and Eutrophication: A Global Perspective Increasing Accu-

- mulation of Phosphorus in Soil Threatens Rivers, Lakes, and Coastal Oceans with Eutrophication". In: *BioScience* 51.3, pp. 227–234. ISSN: 0006-3568, 1525-3244.
- Berry, Joseph A. and O Bjorkman (1980). "Photosynthetic Response and Adaptation to Temperature in Higher Plants". In: *Annual Review of Plant Physiology* 31.1, pp. 491–543.
- Biscani, Francesco, Dario Izzo, and Chit Hong Yam (Apr. 21, 2010). "A Global Optimisation Toolbox for Massively Parallel Engineering Optimisation". In:
- Biscarini, Filippo et al. (May 26, 2016). "Genome-Wide Association Study for Traits Related to Plant and Grain Morphology, and Root Architecture in Temperate Rice Accessions". In: *PLOS ONE* 11.5, e0155425. ISSN: 1932-6203.
- Bloom, A J, F S Chapin, and H A Mooney (1985). "Resource Limitation in Plants-An Economic Analogy". In: *Annual Review of Ecology and Systematics* 16.1, pp. 363–392.
- Bouwman, Lex et al. (Dec. 24, 2013). "Exploring Global Changes in Nitrogen and Phosphorus Cycles in Agriculture Induced by Livestock Production over the 1900–2050 Period". In: *Proceedings of the National Academy of Sciences* 110.52, pp. 20882–20887. ISSN: 0027-8424, 1091-6490.
- Bray, Roger H and LT Kurtz (1945). "Determination of Total, Organic, and Available Forms of Phosphorus in Soils". In: *Soil science* 59.1, pp. 39–46. ISSN: 0038-075X.
- Brisson, N. et al. (2009). *Conceptual Basis, Formalisations and Parameterization of the Stics Crop Model*. Collection Update Sciences & Technologies. Quae.
- Cakmak, Ismail, Christine Hengeler, and Horst Marschner (1994). "Partitioning of Shoot and Root Dry Matter and Carbohydrates in Bean Plants Suffering from Phosphorus, Potassium and Magnesium Deficiency". In: *Journal of Experimental Botany* 45.9, pp. 1245–1250.
- Childers, Daniel L. et al. (Feb. 1, 2011). "Sustainability Challenges of Phosphorus and Food: Solutions from Closing the Human Phosphorus Cycle". In: *BioScience* 61.2, pp. 117–124. ISSN: 0006-3568.
- Chowdhury, Rubel Biswas et al. (2017). "Key Sustainability Challenges for the Global Phosphorus Resource, Their Implications for Global Food Security, and Options for Mitigation". In: *Journal of Cleaner Production* 140, pp. 945–963. ISSN: 0959-6526.
- Claussen, Martin et al. (2002). "Earth System Models of Intermediate Complexity: Closing the Gap in the Spectrum of Climate System Models". In: *Climate dynamics* 18.7, pp. 579–586. ISSN: 0930-7575.
- Collatz, G. J., M. Ribas-Carbo, and Joseph A. Berry (1992). "Coupled Photosynthesis-Stomatal Conductance Model for Leaves of C4 Plants". In: *Functional Plant Biology* 19.5, pp. 519–538. ISSN: 1445-4416.
- Comte, Myriam (2018). "Analysis and Simulation of a Model of Phosphorus Uptake by Plant Roots". In: *Current Research in Nonlinear Analysis: In Honor of Haim Brezis and Louis*

- Nirenberg. Ed. by Themistocles M. Rassias. Springer Optimization and Its Applications. Cham: Springer International Publishing, pp. 85–97.
- Cordell, Dana, Jan-Olof Drangert, and Stuart White (May 2009). “The Story of Phosphorus: Global Food Security and Food for Thought”. In: *Global Environmental Change* 19.2, pp. 292–305. ISSN: 09593780.
- Correll, David L (1998). “The Role of Phosphorus in the Eutrophication of Receiving Waters: A Review”. In: *Journal of environmental quality* 27.2, pp. 261–266. ISSN: 0047-2425.
- Cross, Anne Fernald and William H. Schlesinger (Jan. 1995). “A Literature Review and Evaluation of the Hedley Fractionation: Applications to the Biogeochemical Cycle of Soil Phosphorus in Natural Ecosystems”. In: *Geoderma* 64.3–4, pp. 197–214. ISSN: 0016-7061.
- Dai, Xiaoqin et al. (Dec. 12, 2013). “Variation in Yield Gap Induced by Nitrogen, Phosphorus and Potassium Fertilizer in North China Plain”. In: *PLoS ONE* 8.12. ISSN: 1932-6203.
- Daniel, TC, AN Sharpley, and JL Lemunyon (1998). “Agricultural Phosphorus and Eutrophication: A Symposium Overview”. In: *Journal of environmental quality* 27.2, pp. 251–257. ISSN: 0047-2425.
- Dantzig, George B. and Mukund Narain Thapa (1997). *Linear Programming*. Springer Series in Operations Research. New York: Springer. 2 pp.
- De Willigen, Peter and Meine van Noordwijk (1994). “MASS FLOW AND DIFFUSION OF NUTRIENTS TO A ROOT WITH CONSTANT OR ZERO-SINK UPTAKE I. CONSTANT UPTAKE.” In: *Soil Science* 157.3, pp. 162–170.
- De Ridder, M. (2012). *Risks and Opportunities in the Global Phosphate Rock Market : Robust Strategies in Times of Uncertainty*. Rapport / Centre for Strategic Studies;No. 17 | 12 | 12. Den Haag: The Hague Centre for Strategic Studies.
- Deryng, Delphine et al. (Mar. 2014). “Global Crop Yield Response to Extreme Heat Stress under Multiple Climate Change Futures”. In: *Environmental Research Letters* 9.3, p. 034011. ISSN: 1748-9326.
- Dewar, Roderick C. et al. (Feb. 1, 2009). “Optimal Function Explains Forest Responses to Global Change”. In: *BioScience* 59.2, pp. 127–139. ISSN: 0006-3568.
- Dobermann, A. et al. (1998). “Management of Phosphorus, Potassium, and Sulfur in Intensive, Irrigated Lowland Rice”. In: *Field Crops Research* 56.1-2, pp. 113–138.
- Duivenbooden, N. van (1992). *Sustainability in Terms of Nutrient Elements with Special Reference to West - Africa*. Wageningen: CABO-DLO.
- Duivenbooden, N. van, C. T. de Wit, and H. van Keulen (Feb. 1995). “Nitrogen, Phosphorus and Potassium Relations in Five Major Cereals Reviewed in Respect to Fertilizer Recommendations Using Simulation Modelling”. In: *Fertilizer research* 44.1, pp. 37–49. ISSN: 0167-1731, 1573-0867.

- Dupuy, L., P. J. Gregory, and A. G. Bengough (May 1, 2010). "Root Growth Models: Towards a New Generation of Continuous Approaches". In: *Journal of Experimental Botany* 61.8, pp. 2131–2143. ISSN: 0022-0957, 1460-2431.
- Elliott, J. et al. (Feb. 11, 2015). "The Global Gridded Crop Model Intercomparison: Data and Modeling Protocols for Phase 1 (v1.0)". In: *Geoscientific Model Development* 8.2, pp. 261–277. ISSN: 1991-9603.
- Elser, J. J. et al. (May 2010). "Biological Stoichiometry of Plant Production: Metabolism, Scaling and Ecological Response to Global Change: Tansley Review". In: *New Phytologist* 186.3, pp. 593–608. ISSN: 0028646X.
- Elser, James J. et al. (2007). "Global Analysis of Nitrogen and Phosphorus Limitation of Primary Producers in Freshwater, Marine and Terrestrial Ecosystems". In: *Ecology letters* 10.12, pp. 1135–1142.
- Elser, James and Elena Bennett (Oct. 6, 2011). "Phosphorus Cycle: A Broken Biogeochemical Cycle". In: *Nature* 478.7367, pp. 29–31. ISSN: 0028-0836.
- Ermolieva, T Yu et al. (2015). "Systems Analysis of Robust Strategic Decisions to Plan Secure Food, Energy, and Water Provision Based on the Stochastic GLOBIOM Model". In: *Cybernetics and Systems Analysis* 51.1, pp. 125–133. ISSN: 1060-0396.
- Fageria, N. K. (2009). *The Use of Nutrients in Crop Plants*. OCLC: 181142705. Boca Raton: CRC Press. 430 pp.
- Fageria, N. K. and V. C. Baligar (Sept. 1989). "Response of Legumes and Cereals to Phosphorus in Solution Culture". In: *Journal of Plant Nutrition* 12.9, pp. 1005–1019. ISSN: 0190-4167, 1532-4087.
- Fageria, N. K. and J. P. Oliveira (Aug. 24, 2014). "Nitrogen, Phosphorus and Potassium Interactions in Upland Rice". In: *Journal of Plant Nutrition* 37.10, pp. 1586–1600. ISSN: 0190-4167, 1532-4087.
- Fageria, V. D. (July 31, 2001). "NUTRIENT INTERACTIONS IN CROP PLANTS". In: *Journal of Plant Nutrition* 24.8, pp. 1269–1290. ISSN: 0190-4167, 1532-4087.
- FAOSTAT (2019). *FAOSTAT Database*. URL: <http://www.fao.org/faostat/en/#home> (visited on 02/28/2019).
- Farquhar, Graham D., Susanne von Caemmerer, and Joseph A. Berry (2001). "Models of Photosynthesis". In: *Plant Physiology* 125.1, pp. 42–45.
- Farrior, Caroline E. et al. (June 9, 2015). "Decreased Water Limitation under Elevated CO<sub>2</sub> Amplifies Potential for Forest Carbon Sinks". In: *Proceedings of the National Academy of Sciences* 112.23, pp. 7213–7218. ISSN: 0027-8424, 1091-6490.
- Filippelli, G. M. (Jan. 1, 2002). "The Global Phosphorus Cycle". In: *Reviews in Mineralogy and Geochemistry* 48.1, pp. 391–425. ISSN: 1529-6466.
- Fixen, Paul E. and Adrian M. Johnston (2012). "World Fertilizer Nutrient Reserves: A View to the Future". In: *Journal of the Science of Food and Agriculture* 92.5, pp. 1001–1005. ISSN: 1097-0010.

- Fluxnet Network, The (2019). *List of FLUXNET 2015 Sites*. URL: <https://fluxnet.fluxdata.org/sites/site-list-and-pages/> (visited on 08/18/2019).
- Folberth, Christian et al. (June 21, 2016). “Uncertainty in Soil Data Can Outweigh Climate Impact Signals in Global Crop Yield Simulations”. In: *Nature Communications* 7, p. 11872. ISSN: 2041-1723.
- Foley, Jonathan A. et al. (Oct. 12, 2011). “Solutions for a Cultivated Planet”. In: *Nature* 478.7369, pp. 337–342. ISSN: 0028-0836, 1476-4687.
- Franklin, Oskar et al. (Jan. 2009). “Forest Fine-Root Production and Nitrogen Use under Elevated CO<sub>2</sub> : Contrasting Responses in Evergreen and Deciduous Trees Explained by a Common Principle”. In: *Global Change Biology* 15.1, pp. 132–144. ISSN: 13541013, 13652486.
- Franklin, Oskar et al. (June 1, 2012). “Modeling Carbon Allocation in Trees: A Search for Principles”. In: *Tree Physiology* 32.6, pp. 648–666. ISSN: 0829-318X.
- Fredeen, Arthur L., I. Madhusudana Rao, and Norman Terry (Jan. 1, 1989). “Influence of Phosphorus Nutrition on Growth and Carbon Partitioning in Glycine Max”. In: *Plant Physiology* 89.1, pp. 225–230. ISSN: 0032-0889, 1532-2548.
- Fredeen, Arthur L. et al. (June 1, 1990). “Effects of Phosphorus Nutrition on Photosynthesis in Glycine Max (L.) Merr.” In: *Planta* 181.3, pp. 399–405. ISSN: 1432-2048.
- Friedlingstein, P. et al. (Oct. 1, 1999). “Toward an Allocation Scheme for Global Terrestrial Carbon Models”. In: *Global Change Biology* 5.7, pp. 755–770. ISSN: 1365-2486.
- Gallet, Anne et al. (2003). “Effect of Phosphate Fertilization on Crop Yield and Soil Phosphorus Status”. In: *Journal of Plant Nutrition and Soil Science* 166.5, pp. 568–578.
- Gasser, Thomas et al. (Jan. 20, 2017). “The Compact Earth System Model OSCAR v2.2: Description and First Results”. In: *Geoscientific Model Development* 10.1, pp. 271–319. ISSN: 1991-959X.
- Gérard, Frédéric (Jan. 15, 2016). “Clay Minerals, Iron/Aluminum Oxides, and Their Contribution to Phosphate Sorption in Soils — A Myth Revisited”. In: *Geoderma* 262, pp. 213–226. ISSN: 0016-7061.
- Geyer, Charles J (1992). “Practical Markov Chain Monte Carlo”. In: *Statistical science*, pp. 473–483. ISSN: 0883-4237.
- Godfray, H. Charles J. et al. (Feb. 12, 2010). “Food Security: The Challenge of Feeding 9 Billion People”. In: *Science* 327.5967, pp. 812–818. ISSN: 0036-8075, 1095-9203.
- Goll, D. S. et al. (Sept. 6, 2012). “Nutrient Limitation Reduces Land Carbon Uptake in Simulations with a Model of Combined Carbon, Nitrogen and Phosphorus Cycling”. In: *Biogeosciences* 9.9, pp. 3547–3569. ISSN: 1726-4189.
- Goll, D. S. et al. (Oct. 12, 2017). “A Representation of the Phosphorus Cycle for ORCHIDEE (Revision 4520)”. In: *Geosci. Model Dev.* 10.10, pp. 3745–3770. ISSN: 1991-9603.

- Guimberteau, M. et al. (June 6, 2014). “Testing Conceptual and Physically Based Soil Hydrology Schemes against Observations for the Amazon Basin”. In: *Geosci. Model Dev.* 7.3, pp. 1115–1136. ISSN: 1991-9603.
- Guimberteau, Matthieu et al. (Jan. 15, 2018). “ORCHIDEE-MICT (v8.4.1), a Land Surface Model for the High Latitudes: Model Description and Validation”. In: *Geoscientific Model Development* 11.1, pp. 121–163. ISSN: 1991-959X.
- Heckenmüller, Markus, Daiju Narita, and Gernot Klepper (2014). *Global Availability of Phosphorus and Its Implications for Global Food Supply: An Economic Overview*. Working Paper 1897. Kiel Working Paper.
- Hedley, M. J., J. W. B. Stewart, and B. S. Chauhan (1982). “Changes in Inorganic and Organic Soil Phosphorus Fractions Induced by Cultivation Practices and by Laboratory Incubations”. In: *Soil Science Society of America Journal* 46.5, pp. 970–976. ISSN: 0361-5995.
- Heinen, Marius, Alain Mollier, and Peter De Willigen (May 1, 2003). “Growth of a Root System Described as Diffusion. II. Numerical Model and Application”. In: *Plant and Soil* 252.2, pp. 251–265. ISSN: 0032-079X, 1573-5036.
- Hengl, Tomislav et al. (Aug. 29, 2014). “SoilGrids1km — Global Soil Information Based on Automated Mapping”. In: *PLOS ONE* 9.8, e105992. ISSN: 1932-6203.
- Herman, Jon and Will Usher (Jan. 10, 2017). *SALib: An Open-Source Python Library for Sensitivity Analysis*. URL: <http://joss.theoj.org> (visited on 05/16/2019).
- Hinsinger, Philippe et al. (Aug. 2009). “Rhizosphere: Biophysics, Biogeochemistry and Ecological Relevance”. In: *Plant and Soil* 321.1-2, pp. 117–152. ISSN: 0032-079X, 1573-5036.
- Hinsinger, Philippe et al. (July 14, 2011). “Acquisition of Phosphorus and Other Poorly Mobile Nutrients by Roots. Where Do Plant Nutrition Models Fail?” In: *Plant and Soil* 348.1-2, pp. 29–61. ISSN: 0032-079X, 1573-5036.
- Hocking, P. J. (July 1994). “Dry-matter Production, Mineral Nutrient Concentrations, and Nutrient Distribution and Redistribution in Irrigated Spring Wheat”. In: *Journal of Plant Nutrition* 17.8, pp. 1289–1308. ISSN: 0190-4167, 1532-4087.
- Hoffmann, Roald, Vladimir I Minkin, and Barry K Carpenter (1997). “Ockham’s Razor and Chemistry”. In: *International Journal for the Philosophy of Chemistry* 3, pp. 3–28.
- Hsiao, Theodore C. (June 1, 1973). “Plant Responses to Water Stress”. In: *Annual Review of Plant Physiology* 24.1, pp. 519–570. ISSN: 0066-4294.
- Hungate, Bruce A. et al. (Nov. 28, 2003). “Nitrogen and Climate Change”. In: *Science* 302.5650, pp. 1512–1513. ISSN: 0036-8075, 1095-9203.
- Ise, Takeshi et al. (2010). “Comparison of Modeling Approaches for Carbon Partitioning: Impact on Estimates of Global Net Primary Production and Equilibrium Biomass of Woody Vegetation from MODIS GPP”. In: *Journal of Geophysical Research: Biogeosciences* 115.G4. ISSN: 2156-2202.

- Iwasa, Yoh and Jonathan Roughgarden (Feb. 1, 1984). "Shoot/Root Balance of Plants: Optimal Growth of a System with Many Vegetative Organs". In: *Theoretical Population Biology* 25.1, pp. 78–105. ISSN: 0040-5809.
- Jacob, J. and D. W. Lawlor (Jan. 8, 1991). "Stomatal and Mesophyll Limitations of Photosynthesis in Phosphate Deficient Sunflower, Maize and Wheat Plants". In: *Journal of Experimental Botany* 42.8, pp. 1003–1011. ISSN: 0022-0957, 1460-2431.
- Janssen, B. H. et al. (1990). "A System for Quantitative Evaluation of the Fertility of Tropical Soils (QUEFTS)". In: *Geoderma* 46.4, pp. 299–318.
- Jones, C. A. et al. (1984). "A Simplified Soil and Plant Phosphorus Model: I. Documentation". In: *Soil Science Society of America Journal* 48.4, pp. 800–805.
- Jones, Eric, Travis Oliphant, and Pearu Peterson (2001). *SciPy: Open Source Scientific Tools for Python*. URL: <http://www.scipy.org/>.
- Jones, G. P. D., G. J. Blair, and R. S. Jessop (1989). "Phosphorus Efficiency in Wheat—a Useful Selection Criterion?" In: *Field Crops Research* 21.3-4, pp. 257–264.
- Jones, James W et al. (2003). "The DSSAT Cropping System Model". In: *European journal of agronomy* 18.3-4, pp. 235–265. ISSN: 1161-0301.
- Justice, Christopher O et al. (1998). "The Moderate Resolution Imaging Spectroradiometer (MODIS): Land Remote Sensing for Global Change Research". In: *IEEE transactions on geoscience and remote sensing* 36.4, pp. 1228–1249. ISSN: 0196-2892.
- Kauwenbergh, Steven J Van, Mike Stewart, and Robert Mikkelsen (2013). "World Reserves of Phosphate Rock... a Dynamic and Unfolding Story". In: 97.3, p. 3.
- Kearney John (Sept. 27, 2010). "Food Consumption Trends and Drivers". In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 365.1554, pp. 2793–2807.
- Keating, Brian A et al. (2003). "An Overview of APSIM, a Model Designed for Farming Systems Simulation". In: *European journal of agronomy* 18.3-4, pp. 267–288. ISSN: 1161-0301.
- Khabarov, Nikolay and Michael Obersteiner (June 14, 2017). "Global Phosphorus Fertilizer Market and National Policies: A Case Study Revisiting the 2008 Price Peak". In: *Frontiers in Nutrition* 4. ISSN: 2296-861X.
- Khush, Gurdev S. (Oct. 2001). "Green Revolution: The Way Forward". In: *Nature Reviews Genetics* 2.10, pp. 815–822. ISSN: 1471-0064.
- Kim, Soo-Hyung et al. (2006). "Canopy Photosynthesis, Evapotranspiration, Leaf Nitrogen, and Transcription Profiles of Maize in Response to CO<sub>2</sub> Enrichment". In: *Global Change Biology* 12.3, pp. 588–600. ISSN: 1365-2486.
- Krinner, G. et al. (2005). "A Dynamic Global Vegetation Model for Studies of the Coupled Atmosphere-Biosphere System: DVGM FOR COUPLED CLIMATE STUDIES". In: *Global Biogeochemical Cycles* 19.1. ISSN: 08866236.

- Lassaletta, Luis et al. (Oct. 2014). "50 Year Trends in Nitrogen Use Efficiency of World Cropping Systems: The Relationship between Yield and Nitrogen Input to Cropland". In: *Environmental Research Letters* 9.10, p. 105011. ISSN: 1748-9326.
- Latshaw, W. L. and E. C. Miu (1934). "ELEMENTAL COMPOSITION OF THE CORN PLANT<sup>1</sup>". In: *Journal of Agricultural Research* 27.11.
- Lemaire, Gilles, ed. (1997). *Diagnosis of the Nitrogen Status in Crops*. Berlin, Heidelberg: Springer Berlin Heidelberg.
- Li, Haixiao et al. (May 19, 2016). "The Long-Term Effects of Tillage Practice and Phosphorus Fertilization on the Distribution and Morphology of Corn Root". In: *Plant and Soil*. ISSN: 0032-079X, 1573-5036.
- Licker, Rachel et al. (Nov. 1, 2010). "Mind the Gap: How Do Climate and Agricultural Management Explain the 'Yield Gap' of Croplands around the World?" In: *Global Ecology and Biogeography* 19.6, pp. 769–782. ISSN: 1466-8238.
- Lindquist, John L. et al. (Jan. 1, 2005). "Maize Radiation Use Efficiency under Optimal Growth Conditions". In: *Agronomy Journal* 97.1, pp. 72–78. ISSN: 1435-0645.
- Liu, Yi et al. (Apr. 1, 2008). "Global Phosphorus Flows and Environmental Impacts from a Consumption Perspective". In: *Journal of Industrial Ecology* 12.2, pp. 229–247. ISSN: 1530-9290.
- Lobell, D.B. and M.B. Burke (Oct. 15, 2010). "On the Use of Statistical Models to Predict Crop Yield Responses to Climate Change". In: *Agricultural and Forest Meteorology* 150.11, pp. 1443–1452. ISSN: 0168-1923.
- Lobell, D.B. and C.B. Field (Mar. 2007). "Global Scale Climate–Crop Yield Relationships and the Impacts of Recent Warming". In: *Environmental Research Letters* 2.1, p. 014002. ISSN: 1748-9326.
- Lobell, David B., Kenneth G. Cassman, and Christopher B. Field (Nov. 2009). "Crop Yield Gaps: Their Importance, Magnitudes, and Causes". In: *Annual Review of Environment and Resources* 34.1, pp. 179–204. ISSN: 1543-5938, 1545-2050.
- Loneragan, J. F. et al. (1979). "Phosphorus Toxicity as a Factor in Zinc-Phosphorus Interactions in Plants <sup>1</sup>". In: *Soil Science Society of America Journal* 43.5, pp. 966–972. ISSN: 0361-5995.
- MacDonald, G. K. et al. (Feb. 15, 2011). "Agronomic Phosphorus Imbalances across the World's Croplands". In: *Proceedings of the National Academy of Sciences* 108.7, pp. 3086–3091. ISSN: 0027-8424, 1091-6490.
- Mäkelä, Annikki, Harry T. Valentine, and Heljä-Sisko Helmisaari (Oct. 2008). "Optimal Co-Allocation of Carbon and Nitrogen in a Forest Stand at Steady State". In: *New Phytologist* 180.1, pp. 114–123. ISSN: 0028646X, 14698137.
- Makhorin, Andrew (2001). "Gnu Linear Programming Kit (Glpk)". In: *Department for Applied Informatics, Moscow Aviation Institute, Moscow, Russia*.

- Makowski, D. et al. (Apr. 1, 2014). "Global Agronomy, a New Field of Research. A Review". In: *Agronomy for Sustainable Development* 34.2, pp. 293–307. ISSN: 1773-0155.
- Marschner, Horst, Ernest Kirkby, and Ismail Cakmak (Aug. 1, 1996). "Effect of Mineral Nutritional Status on Shoot-Root Partitioning of Photoassimilates and Cycling of Mineral Nutrients". In: *Journal of Experimental Botany* 47 (Special\_Issue), pp. 1255–1263. ISSN: 0022-0957, 1460-2431.
- Mascaro, Joseph et al. (Jan. 28, 2014). "A Tale of Two "Forests": Random Forest Machine Learning Aids Tropical Forest Carbon Mapping". In: *PLoS ONE* 9.1. Ed. by Ben Bond-Lamberty, e85993. ISSN: 1932-6203.
- McCormack, M. Luke et al. (2015). "Redefining Fine Roots Improves Understanding of Below-Ground Contributions to Terrestrial Biosphere Processes". In: *New Phytologist* 207.3, pp. 505–518. ISSN: 1469-8137.
- McMurtrie, Ross E. et al. (2008). "Why Is Plant-Growth Response to Elevated CO<sub>2</sub> Amplified When Water Is Limiting, but Reduced When Nitrogen Is Limiting? A Growth-Optimisation Hypothesis". In: *Functional Plant Biology* 35.6, p. 521. ISSN: 1445-4408.
- Meinshausen, Malte, Sarah CB Raper, and Tom ML Wigley (2011). "Emulating Coupled Atmosphere-Ocean and Carbon Cycle Models with a Simpler Model, MAGICC6–Part 1: Model Description and Calibration". In: *Atmospheric Chemistry and Physics* 11.4, pp. 1417–1456. ISSN: 1680-7316.
- Meir, P., J. Grace, and A. C. Miranda (June 2001). "Leaf Respiration in Two Tropical Rainforests: Constraints on Physiology by Phosphorus, Nitrogen and Temperature". In: *Functional Ecology* 15.3, pp. 378–387. ISSN: 0269-8463, 1365-2435.
- Mollier, A. and S. Pellerin (1999). "Maize Root System Growth and Development as Influenced by Phosphorus Deficiency". In: *Journal of Experimental Botany* 50.333, pp. 487–497.
- Monfreda, Chad, Navin Ramankutty, and Jonathan A. Foley (Mar. 1, 2008). "Farming the Planet: 2. Geographic Distribution of Crop Areas, Yields, Physiological Types, and Net Primary Production in the Year 2000". In: *Global Biogeochemical Cycles* 22.1, GB1022. ISSN: 1944-9224.
- Moré, Jorge J. (1978). "The Levenberg-Marquardt Algorithm: Implementation and Theory". In: *Numerical Analysis*. Springer, pp. 105–116.
- Morel, Christian et al. (2000). "Transfer of Phosphate Ions between Soil and Solution: Perspectives in Soil Testing". In: *Journal of Environmental Quality* 29.1, pp. 50–59. ISSN: 0047-2425.
- Morel, Christian et al. (Aug. 2014). "Modeling of Phosphorus Dynamics in Contrasting Agroecosystems Using Long-Term Field Experiments". In: *Canadian Journal of Soil Science* 94.3, pp. 377–387. ISSN: 0008-4271, 1918-1841.
- Mueller, Nathaniel D. et al. (Oct. 2012). "Closing Yield Gaps through Nutrient and Water Management". In: *Nature* 490.7419, pp. 254–257. ISSN: 0028-0836, 1476-4687.

- Müller, C. et al. (Apr. 4, 2017). “Global Gridded Crop Model Evaluation: Benchmarking, Skills, Deficiencies and Implications”. In: *Geosci. Model Dev.* 10.4, pp. 1403–1422. ISSN: 1991-9603.
- Müller, Christoph et al. (May 8, 2019). “The Global Gridded Crop Model Intercomparison Phase 1 Simulation Dataset”. In: *Scientific Data* 6.1, pp. 1–22. ISSN: 2052-4463.
- Nakhforoosh, Alireza et al. (July 2014). “Wheat Root Diversity and Root Functional Characterization”. In: *Plant and Soil* 380.1-2, pp. 211–229. ISSN: 0032-079X, 1573-5036.
- Nelson, WL (1953). “The Development, Evaluation, and Use of Soil Tests for Phosphorus Availability”. In: *Agronomy* 4, pp. 153–188.
- Neset, Tina-Simone S and Dana Cordell (2012). “Global Phosphorus Scarcity: Identifying Synergies for a Sustainable Future”. In: *Journal of the Science of Food and Agriculture* 92.1, pp. 2–6. ISSN: 0022-5142.
- Neumann, Kathleen et al. (June 2010). “The Yield Gap of Global Grain Production: A Spatial Analysis”. In: *Agricultural Systems* 103.5, pp. 316–326. ISSN: 0308-521X.
- Niu, Yao Fang et al. (Dec. 23, 2012). “Responses of Root Architecture Development to Low Phosphorus Availability: A Review”. In: *Annals of Botany*, mcs285. ISSN: 0305-7364, 1095-8290.
- Nossent, Jiri, Pieter Elsen, and Willy Bauwens (Dec. 2011). “Sobol’ Sensitivity Analysis of a Complex Environmental Model”. In: *Environmental Modelling & Software* 26.12, pp. 1515–1525. ISSN: 13648152.
- Obersteiner, Michael et al. (2013). “The Phosphorus Trilemma”. In: *Nature Geoscience* 6.11, p. 897.
- Olden, Julian D., Joshua J. Lawler, and N. LeRoy Poff (June 1, 2008). “Machine Learning Methods Without Tears: A Primer for Ecologists”. In: *The Quarterly Review of Biology* 83.2, pp. 171–193. ISSN: 0033-5770.
- Olesen, T. et al. (2001). “CONSTANT SLOPE IMPEDANCE FACTOR MODEL FOR PREDICTING THE SOLUTE DIFFUSION COEFFICIENT IN UNSATURATED SOIL”. In: *Soil Science* 166.2. ISSN: 0038-075X.
- Olsen, Sterling R (1954). *Estimation of Available Phosphorus in Soils by Extraction with Sodium Bicarbonate*. United States Department Of Agriculture; Washington.
- Pagès, Loïc (2019). “Analysis and Modeling of the Variations of Root Branching Density Within Individual Plants and Among Species”. In: *Frontiers in Plant Science* 10.
- Parton, W. J., J. W. B. Stewart, and C. V. Cole (Feb. 1, 1988). “Dynamics of C, N, P and S in Grassland Soils: A Model”. In: *Biogeochemistry* 5.1, pp. 109–131. ISSN: 1573-515X.
- Passioura, John B. (1996). “Simulation Models: Science, Snake Oil, Education, or Engineering?” In: *Agronomy Journal* 88.5, pp. 690–694.
- Pellerin, Sylvain et al. (2019). “A Model-Based Assessment of the Soil C Storage Potential at the National Scale: A Case Study from France”. In: International Conference Food Security Climate Change: 4 per Mille Initiative New Tangible Global Challenges for

- the Soil. Book of abstracts: International Conference Food Security Climate Change: 4 per mille initiative new tangible global challenges for the soil. public: INRA, p. 49.
- Peng, Bin et al. (Mar. 15, 2018). "Improving Maize Growth Processes in the Community Land Model: Implementation and Evaluation". In: *Agricultural and Forest Meteorology* 250-251, pp. 64–89. ISSN: 0168-1923.
- Peng, Changhui (2000). "From Static Biogeographical Model to Dynamic Global Vegetation Model: A Global Perspective on Modelling Vegetation Dynamics". In: *Ecological modelling* 135.1, pp. 33–54. ISSN: 0304-3800.
- Peñuelas, Josep et al. (Dec. 17, 2013). "Human-Induced Nitrogen–Phosphorus Imbalances Alter Natural and Managed Ecosystems across the Globe". In: *Nature Communications* 4. ISSN: 2041-1723.
- Pindyck, Robert S (2013). "Climate Change Policy: What Do the Models Tell Us?" In: *Journal of Economic Literature* 51.3, pp. 860–72. ISSN: 0022-0515.
- Pingali, P. L. (July 31, 2012). "Green Revolution: Impacts, Limits, and the Path Ahead". In: *Proceedings of the National Academy of Sciences* 109.31, pp. 12302–12308. ISSN: 0027-8424, 1091-6490.
- Plaxton, William C. and Hans Lambers, eds. (2015). *Phosphorus Metabolism in Plants*. Annual Plant Reviews 48. OCLC: 898924009. Chichester: Wiley-Blackwell. 449 pp.
- Plénet, D. et al. (2000). "Growth Analysis of Maize Field Crops under Phosphorus Deficiency". In: *Plant and Soil* 223.1-2, pp. 119–132.
- Poorter, Hendrik et al. (Jan. 1, 2012). "Biomass Allocation to Leaves, Stems and Roots: Meta-Analyses of Interspecific Variation and Environmental Control". In: *New Phytologist* 193.1, pp. 30–50. ISSN: 1469-8137.
- Portmann, Felix T., Stefan Siebert, and Petra Döll (Mar. 1, 2010). "MIRCA2000—Global Monthly Irrigated and Rainfed Crop Areas around the Year 2000: A New High-Resolution Data Set for Agricultural and Hydrological Modeling". In: *Global Biogeochemical Cycles* 24.1, GB1011. ISSN: 1944-9224.
- Prentice, I. (2013). "Ecosystem Science for a Changing World - Grantham Discussion Paper 4". In: Imperial College Press.
- Prentice, I. Colin et al. (Jan. 1, 2014). "Balancing the Costs of Carbon Gain and Water Transport: Testing a New Theoretical Framework for Plant Functional Ecology". In: *Ecology Letters* 17.1, pp. 82–91. ISSN: 1461-023X.
- Purves, Drew and Stephen Pacala (June 13, 2008). "Predictive Models of Forest Dynamics". In: *Science* 320.5882, pp. 1452–1453. ISSN: 0036-8075, 1095-9203.
- Randall, G. W. and R. G. Hoeft (1988). "Placement Methods for Improved Efficiency of P and K Fertilizers: A Review". In: *Journal of Production Agriculture* 1.1, pp. 70–79.
- Ray, Deepak K. et al. (June 19, 2013). "Yield Trends Are Insufficient to Double Global Crop Production by 2050". In: *PLOS ONE* 8.6, e66428. ISSN: 1932-6203.

- Ray, Deepak K. et al. (Jan. 22, 2015). "Climate Variation Explains a Third of Global Crop Yield Variability". In: *Nature Communications* 6, p. 5989. ISSN: 2041-1723.
- Reed, S. C., X. Yang, and P. E. Thornton (Oct. 2015). "Incorporating Phosphorus Cycling into Global Modeling Efforts: A Worthwhile, Tractable Endeavor". In: *New Phytologist* 208.2, pp. 324–329. ISSN: 0028646X.
- Reichwein Zientek, Linda and Bruce Thompson (July 1, 2006). "Commonality Analysis: Partitioning Variance to Facilitate Better Understanding of Data". In: *Journal of Early Intervention* 28.4, pp. 299–307. ISSN: 1053-8151.
- Ricciuto, Daniel M., Kenneth J. Davis, and Klaus Keller (June 1, 2008). "A Bayesian Calibration of a Simple Carbon Cycle Model: The Role of Observations in Estimating and Reducing Uncertainty". In: *Global Biogeochemical Cycles* 22.2. ISSN: 0886-6236.
- Ringeval, Bruno et al. (2017). "Phosphorus in Agricultural Soils: Drivers of Its Distribution at the Global Scale". In: *Global change biology* 23.8, pp. 3418–3432. ISSN: 1354-1013.
- Rogelj, Joeri, Malte Meinshausen, and Reto Knutti (Apr. 2012). "Global Warming under Old and New Scenarios Using IPCC Climate Sensitivity Range Estimates". In: *Nature Climate Change* 2.4, pp. 248–253. ISSN: 1758-6798.
- Roose, T. and G. J. D. Kirk (Mar. 1, 2009). "The Solution of Convection–Diffusion Equations for Solute Transport to Plant Roots". In: *Plant and Soil* 316.1-2, pp. 257–264. ISSN: 0032-079X, 1573-5036.
- Rosemarin, Arno and Nelson Ekane (2016). "The Governance Gap Surrounding Phosphorus". In: *Nutrient cycling in agroecosystems* 104.3, pp. 265–279. ISSN: 1385-1314.
- Rosenzweig, C. et al. (Mar. 15, 2013). "The Agricultural Model Intercomparison and Improvement Project (AgMIP): Protocols and Pilot Studies". In: *Agricultural and Forest Meteorology. Agricultural Prediction Using Climate Model Ensembles* 170, pp. 166–182. ISSN: 0168-1923.
- Rosenzweig, C. et al. (Mar. 4, 2014). "Assessing Agricultural Risks of Climate Change in the 21st Century in a Global Gridded Crop Model Intercomparison". In: *Proceedings of the National Academy of Sciences* 111.9, pp. 3268–3273. ISSN: 0027-8424, 1091-6490.
- Rötter, Reimund P. et al. (July 2011). "Crop–Climate Models Need an Overhaul". In: *Nature Climate Change* 1.4, pp. 175–177. ISSN: 1758-678X, 1758-6798.
- Rowland, Lucy et al. (May 1, 2017). "Scaling Leaf Respiration with Nitrogen and Phosphorus in Tropical Forests across Two Continents". In: *New Phytologist* 214.3, pp. 1064–1077. ISSN: 1469-8137.
- Roy, Eric D. et al. (Apr. 18, 2016). "The Phosphorus Cost of Agricultural Intensification in the Tropics". In: *Nature Plants* 2.5, p. 16043. ISSN: 2055-0278.
- Royal Society of London (2009). *Reaping the Benefits: Science and the Sustainable Intensification of Global Agriculture*. London: The Royal Society.

- Ruimy, Anne, Gérard Dedieu, and Bernard Saugier (1996). "TURC: A Diagnostic Model of Continental Gross Primary Productivity and Net Primary Productivity". In: *Global Biogeochemical Cycles* 10.2, pp. 269–285. ISSN: 1944-9224.
- Ryan, Michael G. et al. (Mar. 1, 1996). "Foliage, Fine-Root, Woody-Tissue and Stand Respiration in *Pinus Radiata* in Relation to Nitrogen Status". In: *Tree Physiology* 16.3, pp. 333–343. ISSN: 0829-318X.
- Rychter, Anna M. and Douglas D. Randall (July 1, 1994). "The Effect of Phosphate Deficiency on Carbohydrate Metabolism in Bean Roots". In: *Physiologia Plantarum* 91.3, pp. 383–388. ISSN: 1399-3054.
- Sahrawat, K. L. et al. (Jan. 1, 1995). "Response of Sorghum to Fertilizer Phosphorus and Its Residual Value in a Vertisol". In: *Fertilizer research* 41.1, pp. 41–47. ISSN: 0167-1731, 1573-0867.
- Sample, E. C., R. J. Soper, and G. J. Racz (1980). "Reactions of Phosphate Fertilizers in Soils". In: *The role of phosphorus in agriculture* (theroleofphosph), pp. 263–310.
- Sattari, S. Z. et al. (Apr. 17, 2012). "Residual Soil Phosphorus as the Missing Piece in the Global Phosphorus Crisis Puzzle". In: *Proceedings of the National Academy of Sciences* 109.16, pp. 6348–6353. ISSN: 0027-8424, 1091-6490.
- Sattari, S.Z. et al. (Feb. 2014). "Crop Yield Response to Soil Fertility and N, P, K Inputs in Different Environments: Testing and Improving the QUEFTS Model". In: *Field Crops Research* 157, pp. 35–46. ISSN: 03784290.
- Schieving, Feike and Hendrik Poorter (1999). "Carbon Gain in a Multispecies Canopy: The Role of Specific Leaf Area and Photosynthetic Nitrogen-Use Efficiency in the Tragedy of the Commons". In: *New Phytologist* 143.1, pp. 201–211. ISSN: 1469-8137.
- Shen, J. et al. (July 1, 2011). "Phosphorus Dynamics: From Soil to Plant". In: *PLANT PHYSIOLOGY* 156.3, pp. 997–1005. ISSN: 0032-0889, 1532-2548.
- Shen, J et al. (Mar. 10, 2004). "Crop Yields, Soil Fertility and Phosphorus Fractions in Response to Long-Term Fertilization under the Rice Monoculture System on a Calcareous Soil". In: *Field Crops Research* 86.2–3, pp. 225–238. ISSN: 0378-4290.
- Shinozaki, Kichiro et al. (1964). "A Quantitative Analysis of Plant Form;the Pipe Model Theory,1." In: *JAPANESE JOURNAL OF ECOLOGY* 14.3, pp. 97–105. ISSN: 00215007.
- Sinclair, Thomas R. and Thomas W. Rufty (Dec. 1, 2012). "Nitrogen and Water Resources Commonly Limit Crop Yield Increases, Not Necessarily Plant Genetics". In: *Global Food Security* 1.2, pp. 94–98. ISSN: 2211-9124.
- Smil, Vaclav (2000). "Phosphorus in the Environment: Natural Flows and Human Interferences". In: *Annual review of energy and the environment* 25.1, pp. 53–88.
- Sobol, Ilya M (1993). "Sensitivity Estimates for Nonlinear Mathematical Models". In: *Mathematical modelling and computational experiments* 1.4, pp. 407–414.
- Steenbjerg, F. (Apr. 1951). "Yield Curves and Chemical Plant Analyses". In: *Plant and Soil* 3.2, pp. 97–109. ISSN: 0032-079X, 1573-5036.

- Stocker, Benjamin D. et al. (June 2016). “Terrestrial Nitrogen Cycling in Earth System Models Revisited”. In: *New Phytologist* 210.4, pp. 1165–1168. ISSN: 0028646X.
- Stocker, Benjamin D. et al. (June 1, 2018). “Quantifying Soil Moisture Impacts on Light Use Efficiency across Biomes”. In: *New Phytologist* 218.4, pp. 1430–1449. ISSN: 0028-646X.
- Stockle, CO and P Debaeke (1997). “Modeling Crop Nitrogen Requirements: A Critical Analysis”. In: *European Journal of Agronomy* 7.1-3, pp. 161–169. ISSN: 1161-0301.
- Stolbovoy, Vladimir, Luca Montanarella, and Panagiotis Panagos (2007). *Carbon Sink Enhancement in Soils of Europe: Data, Modeling, Verification*. EUR - Scientific and Technical Research Reports. OPOCE.
- Sun, Yan et al. (July 1, 2017). “Diagnosing Phosphorus Limitations in Natural Terrestrial Ecosystems in Carbon Cycle Models”. In: *Earth’s Future* 5.7, pp. 730–749. ISSN: 2328-4277.
- Takahashi, Shigeru and Muhuddin R. Anwar (Mar. 5, 2007). “Wheat Grain Yield, Phosphorus Uptake and Soil Phosphorus Fraction after 23 Years of Annual Fertilizer Application to an Andosol”. In: *Field Crops Research* 101.2, pp. 160–171. ISSN: 0378-4290.
- Tang, Xu et al. (Sept. 23, 2008). “Phosphorus Efficiency in Long-Term (15 Years) Wheat–Maize Cropping Systems with Various Soil and Climate Conditions”. In: *Field Crops Research* 108.3, pp. 231–237. ISSN: 0378-4290.
- Thornley, J. H. M. (1976). *Mathematical Models in Plant Physiology: A Quantitative Approach to Problems in Plant and Crop Physiology*. Academic Press. 344 pp.
- Thornley, John H. M. (1995). “Shoot Root Allocation with Respect to C, N and P: An Investigation and Comparison of Resistance and Teleonomic Models”. In: *Annals of Botany* 75.4, pp. 391–405. ISSN: 0305-7364.
- Tilman, David (Nov. 1998). “The Greening of the Green Revolution”. In: *Nature* 396.6708, pp. 211–212. ISSN: 1476-4687.
- Tilman, David et al. (Dec. 13, 2011). “Global Food Demand and the Sustainable Intensification of Agriculture”. In: *Proceedings of the National Academy of Sciences* 108.50, pp. 20260–20264. ISSN: 0027-8424, 1091-6490.
- Tinker, P. B. and P. H. Nye (Mar. 2, 2000). *Solute Movement in the Rhizosphere*. Oxford University Press. 465 pp.
- Tonitto, Christina and Jacob E. Ricker-Gilbert (Feb. 10, 2016). “Nutrient Management in African Sorghum Cropping Systems: Applying Meta-Analysis to Assess Yield and Profitability”. In: *Agronomy for Sustainable Development* 36.1, pp. 1–19. ISSN: 1774-0746, 1773-0155.
- United Nations, The (Aug. 11, 2019). *Intergovernmental Panel on Climate Change*. In: *Wikipedia*. Page Version ID: 910298706. The United Nations.
- Valkama, Elena et al. (Apr. 2009). “Phosphorus Fertilization: A Meta-Analysis of 80 Years of Research in Finland”. In: *Agriculture, Ecosystems & Environment* 130.3–4, pp. 75–85. ISSN: 0167-8809.

- Van der Velde, Marijn et al. (Apr. 2, 2013). "Affordable Nutrient Solutions for Improved Food Security as Evidenced by Crop Trials". In: *PLoS ONE* 8.4. Ed. by Luis Herrera-Estrella, e60075. ISSN: 1932-6203.
- Van der Velde, Marijn et al. (Apr. 2014). "African Crop Yield Reductions Due to Increasingly Unbalanced Nitrogen and Phosphorus Consumption". In: *Global Change Biology* 20.4, pp. 1278–1288. ISSN: 13541013.
- Van Diepen, CA van et al. (1989). "WOFOST: A Simulation Model of Crop Production". In: *Soil use and management* 5.1, pp. 16–24. ISSN: 0266-0032.
- Van Ittersum, M. K. and R. Rabbinge (June 1, 1997). "Concepts in Production Ecology for Analysis and Quantification of Agricultural Input-Output Combinations". In: *Field Crops Research* 52.3, pp. 197–208. ISSN: 0378-4290.
- Van Ittersum, M. K. et al. (Mar. 1, 2013). "Yield Gap Analysis with Local to Global Relevance—A Review". In: *Field Crops Research*. Crop Yield Gap Analysis – Rationale, Methods and Applications 143, pp. 4–17. ISSN: 0378-4290.
- Van Keulen, H. (Sept. 1982). "Graphical Analysis of Annual Crop Response to Fertiliser Application". In: *Agricultural Systems* 9.2, pp. 113–126. ISSN: 0308521X.
- Van Wart, Justin et al. (Mar. 1, 2013). "Use of Agro-Climatic Zones to Upscale Simulated Crop Yield Potential". In: *Field Crops Research*. Crop Yield Gap Analysis – Rationale, Methods and Applications 143, pp. 44–55. ISSN: 0378-4290.
- Vitousek, Peter (Apr. 1982). "Nutrient Cycling and Nutrient Use Efficiency". In: *The American Naturalist* 119.4, pp. 553–572. ISSN: 0003-0147, 1537-5323.
- Vitousek, Peter M. et al. (2009). "Nutrient Imbalances in Agricultural Development". In: *Science* 324.5934, p. 1519.
- Vitousek, Peter M. et al. (2010). "Terrestrial Phosphorus Limitation: Mechanisms, Implications, and Nitrogen-Phosphorus Interactions". In: *Ecological applications* 20.1, pp. 5–15.
- Vries, F. W. T. Penning De (1972). "Respiration and Growth". In: *Crop Processes in Controlled Environments*. 2. Academic Press, pp. 327–347.
- Vries, F. W. T. Penning De, J. M. Witlage, and D. Kremer (Nov. 1979). "Rates of Respiration and of Increase in Structural Dry Matter in Young Wheat, Ryegrass and Maize Plants in Relation to Temperature, to Water Stress and to Their Sugar Content". In: *Annals of Botany* 44.5, pp. 595–609. ISSN: 1095-8290, 0305-7364.
- Walker, T. W. and J. K. Syers (Jan. 1, 1976). "The Fate of Phosphorus during Pedogenesis". In: *Geoderma* 15.1, pp. 1–19. ISSN: 0016-7061.
- Wallach, D., D. Makowski, and J. W. Jones, eds. (2006). *Working with Dynamic Crop Models: Evaluation, Analysis, Parameterization, and Applications*. 1st ed. Amsterdam ; Boston: Elsevier. 447 pp.

- Wang, E. and C. J. Smith (2004). "Modelling the Growth and Water Uptake Function of Plant Root Systems: A Review". In: *Australian Journal of Agricultural Research* 55.5, p. 501. ISSN: 0004-9409.
- Wang, Rong et al. (Jan. 2015). "Significant Contribution of Combustion-Related Emissions to the Atmospheric Phosphorus Budget". In: *Nature Geoscience* 8.1, pp. 48–54. ISSN: 1752-0894.
- Wang, Y. P., B. Z. Houlton, and C. B. Field (Mar. 1, 2007). "A Model of Biogeochemical Cycles of Carbon, Nitrogen, and Phosphorus Including Symbiotic Nitrogen Fixation and Phosphatase Production". In: *Global Biogeochemical Cycles* 21.1, GB1018. ISSN: 1944-9224.
- Wang, Y. P., R. M. Law, and B. Pak (July 23, 2010). "A Global Model of Carbon, Nitrogen and Phosphorus Cycles for the Terrestrial Biosphere". In: *Biogeosciences* 7.7, pp. 2261–2282. ISSN: 1726-4189.
- Warszawski, Lila et al. (Mar. 4, 2014). "The Inter-Sectoral Impact Model Intercomparison Project (ISI-MIP): Project Framework". In: *Proceedings of the National Academy of Sciences* 111.9, pp. 3228–3232. ISSN: 0027-8424, 1091-6490.
- Webb, Geoffrey I (1996). "Further Experimental Evidence against the Utility of Occam's Razor". In: *Journal of Artificial Intelligence Research* 4, pp. 397–417. ISSN: 1076-9757.
- Weber, Olaf et al. (2014). "Trade and Finance as Cross-Cutting Issues in the Global Phosphate and Fertilizer Market". In: *Sustainable Phosphorus Management*. Springer, pp. 275–299.
- Weyant, John (Jan. 1, 2017). "Some Contributions of Integrated Assessment Models of Global Climate Change". In: *Review of Environmental Economics and Policy* 11.1, pp. 115–137. ISSN: 1750-6816.
- Wheeler, Tim and Joachim von Braun (Aug. 2, 2013). "Climate Change Impacts on Global Food Security". In: *Science* 341.6145, pp. 508–513. ISSN: 0036-8075, 1095-9203.
- Williams, J. R. et al. (Sept. 29, 1990). "The Erosion Productivity Impact Calculator (EPIC) Model: A Case History". In: *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 329.1255, pp. 421–428.
- Wissuwa, M. and N. Ae (Feb. 1, 2001). "Genotypic Variation for Tolerance to Phosphorus Deficiency in Rice and the Potential for Its Exploitation in Rice Improvement". In: *Plant Breeding* 120.1, pp. 43–48. ISSN: 1439-0523.
- Wit, C. T. de (1953). *A Physical Theory on Placement of Fertilizers*. Landbouwk. 7 pp.
- Woodcock, Curtis E et al. (2008). "Free Access to Landsat Imagery". In: *Science* 320.5879, pp. 1011–1011. ISSN: 0036-8075.
- Wu, Liangquan et al. (Aug. 15, 2015). "Change in Phosphorus Requirement with Increasing Grain Yield for Chinese Maize Production". In: *Field Crops Research* 180, pp. 216–220. ISSN: 0378-4290.

- Wu, X. et al. (Mar. 1, 2016). “ORCHIDEE-CROP (v0), a New Process-Based Agro-Land Surface Model: Model Description and Evaluation over Europe”. In: *Geosci. Model Dev.* 9.2, pp. 857–873. ISSN: 1991-9603.
- Yang, Feihua et al. (Sept. 2007). “Developing a Continental-Scale Measure of Gross Primary Production by Combining MODIS and AmeriFlux Data through Support Vector Machine Approach”. In: *Remote Sensing of Environment* 110.1, pp. 109–122. ISSN: 00344257.
- Yang, X. et al. (Apr. 16, 2013). “The Distribution of Soil Phosphorus for Global Biogeochemical Modeling”. In: *Biogeosciences* 10.4, pp. 2525–2537. ISSN: 1726-4189.
- Yang, X. et al. (Mar. 28, 2014). “The Role of Phosphorus Dynamics in Tropical Forests – a Modeling Study Using CLM-CNP”. In: *Biogeosciences* 11.6, pp. 1667–1681. ISSN: 1726-4189.
- Ye, Yushi et al. (July 3, 2014). “Carbon, Nitrogen and Phosphorus Accumulation and Partitioning, and C:N:P Stoichiometry in Late-Season Rice under Different Water and Nitrogen Managements”. In: *PLOS ONE* 9.7, e101776. ISSN: 1932-6203.
- Zaehle, S. and A. D. Friend (2010). “Carbon and Nitrogen Cycle Dynamics in the O-CN Land Surface Model: 1. Model Description, Site-Scale Evaluation, and Sensitivity to Parameter Estimates”. In: *Global Biogeochemical Cycles* 24.1. ISSN: 1944-9224.
- Zaehle, Sönke and Daniela Dalmonech (Oct. 2011). “Carbon–Nitrogen Interactions on Land at Global Scales: Current Understanding in Modelling Climate Biosphere Feedbacks”. In: *Current Opinion in Environmental Sustainability* 3.5, pp. 311–320. ISSN: 18773435.
- Zhang, J. et al. (Apr. 20, 2017). “Spatiotemporal Dynamics of Soil Phosphorus and Crop Uptake in Global Cropland during the 20th Century”. In: *Biogeosciences* 14.8, pp. 2055–2068. ISSN: 1726-4189.
- Zhang, Q. et al. (2011). “Limitations of Nitrogen and Phosphorous on the Terrestrial Carbon Uptake in the 20th Century”. In: *Geophysical Research Letters* 38.22. ISSN: 0094-8276.

## **Appendix B**

### **Supplementary information**

## B.1 Chapter 2 supplementary information

### B.1.1 Introduction

The supplementary information gives further detail on the following methods, which are mentioned in the main article:

- Details of QUEFTS nutrient uptake vs. grain yield computation
- Details of potential root uptake computation
- Details of soil solution P vs. labile P computation

Additionally, figures and tables are provided which are referred to in the SI and the main text, but were not included in the chapter.

### B.1.2 QUEFTS method

QUEFTS (S. Sattari et al., 2014; Janssen et al., 1990) is an empirical approach to estimate yield response, according to nutrient supply (NPK) and crop demand. The method is roughly divided into an empirical and a ‘theoretical’ section, which calculate the nutrient supply and crop demand to form the final yield estimate. It consists of 4 steps:

1. Computation of the soil available supply of nutrients ( $\text{kgNPK ha}^{-1}$ ) via empirical relations
2. Building of nutrient supply vs. uptake curves, taking into account nutrient interaction
3. Building of yield vs. uptake curves, estimated from previous step
4. Estimating the ultimate yield, calculated from the full combination of yield-uptake curves

Step 1 was not used in our study, since we estimate the potential P supply ourselves. Furthermore, the empirical relations describing the soil nutrient supply have a limited range of validity, and are mostly suited for tropical soils. In our study, we focus on steps 2 to 4 which allow us to estimate the amount of P needed to achieve a certain yield. Step 2 entails the construction of parabolic supply vs. uptake curves, determined by observed nutrient content of the plant aboveground part. The parameters of interest are the yield at maximum nutrient accumulation ( $a$ ) and maximum nutrient dilution ( $d$ ). These correspond to the “maximum” and “minimum” amount of nutrient in aboveground plant parts at grain harvest (determined at 2.5 and 97.5 percentile level, S. Sattari et al. 2014, Table 5) for

the three cereals, which were obtained in field trials spanning a wide variety of soils and climates. The calculation of supply vs. uptake using the parameters  $a$  and  $d$  is following:

$$U_{ij} = \begin{cases} S_i & (1) \text{ for } S_i < S_j \cdot a_j/d_i \\ S_j \cdot d_j/a_i & (2) \text{ for } S_i > S_j \cdot (2 \cdot d_j/a_i - a_j/d_i) \\ S_i - \frac{1}{4} \cdot \frac{(S_i - S_j \cdot a_j/d_i)^2}{S_j \cdot (d_j/a_i - a_j/d_i)} & (3) \text{ for } S_1 > S_i > S_2 \end{cases} \quad (\text{B.1})$$

$a_i$  is the maximum nutrient accumulation (kg grain kg<sup>-1</sup> NPK),  $d_i$  is the maximum nutrient dilution (kgNPK kg<sup>-1</sup> shoot),  $U_{ij}$  is the uptake (kg ha<sup>-1</sup>) and  $S_{i,j}$  the supply (kg ha<sup>-1</sup>) of the nutrient  $i$  and  $j$ , which can be N, P or K. Optimal nutrient uptake triples ( $U_N$ ,  $U_P$  and  $U_K$ ) are determined from these supply vs uptake curves, which are needed for the following steps 3 and 4. In determining the uptake triples it is assumed that uptake efficiency ( $PhE$ ) is highest:

$$U_i = \min(\varepsilon_{ij} \cdot U_{ij}, \varepsilon_{ik} \cdot U_{ik})$$

$$\text{Maximize}(PhE = \sum_i U_i/S_i) \quad (\text{B.2})$$

$\varepsilon_{ij,ik}$  is the Levi-Civita symbol. This works by defining a range of  $S_i$  values (eg. 0-200 kg N, 0-80 kgP, 0-200 kg K) with a predefined number of steps (eg. 100). For each of the  $100^3$   $S_i$  combinations, six  $U_{ij}$  uptake values are determined (according to Eq. B.1). We determine the optimal  $U_N$ ,  $U_K$  and  $U_P$  by starting from a certain  $S_P$  value and chose the one  $S_N$  and  $S_K$  combination (from  $100^2$ ) where  $PhE$  is the highest (Eq. B.2). With optimal uptake triples from the previous step, we can advance to steps 3 and 4. These are closely related, where the final yield is given depending on the amount of nutrient taken up and the crop yield potential. Step 3 first calculates yields at maximum nutrient accumulation  $Y_i^a$  and dilution  $Y_i^d$ , which are the minimum and maximum achievable yield at a certain level of optimum nutrient uptake  $U_i$ :

$$Y_i^a = a_i \cdot U_i$$

$$Y_i^d = d_i \cdot U_i \quad (\text{B.3})$$

These estimates form the basis of the step 4 for determining the ultimate yield  $Y_U$ . Here, nutrient interaction is assumed by calculating the maximum achievable yield, considering the most limiting nutrient  $Y_{min}$  and the water limited, climate potential yield  $Y_{max}$  :

$$\begin{aligned}
Y_{i,min} &= \varepsilon_{ij} \cdot \min(Y_j^a, Y_j^d, Y_{max}) \\
Y_{ij} &= Y_j^a + \frac{(Y_{min} - Y_j^a) \cdot (U_i - Y_j^a/d_i)}{Y_{min}/a_i - Y_j^a/d_j} \cdot \left[ 2 - \frac{(Y_{min} - Y_j^a) \cdot (U_i - Y_j^a/d_i)}{(Y_{min}/a_i - Y_j^a/d_j)} \right] \\
Y_U &= \frac{1}{6} \cdot \sum_i \varepsilon_{ij} \cdot Y_{ij}
\end{aligned} \tag{B.4}$$

$Y_{max}$  is a predefined parameter, which reflects the fact that there are genetic and climatic limits to crop production in any environment. In calculating P demand, we have first determined the optimal uptake triples using step 2 (Eqs. B.1 and B.2) and values provided by S. Sattari et al. (2014) (Table 5 in the original paper). The range of different  $a$  and  $d$  values allowed us to estimate their uncertainty, which was dealt with a Monte Carlo approach as mentioned in the main part of this article. With the optimal uptake triples, we then calculated maximum and minimum achievable yields (Eq. B.3) and the ultimate yield (Eq. B.4). The ultimate yield depends additionally on the potential yield  $Y_{max}$ , which was taken at each grid point from the ORCHIDEE-CROP simulations of crop yield. Finally, P demand was calculated as uptake which satisfies 95% of the calculated ultimate yield  $Y_U$ .

### B.1.3 Potential root P uptake

The conceptual details of the ‘zero-sink’ uptake model (De Willigen and van Noordwijk, 1994) are given in the main article (Section 2.2.3). Here we give details about the computation of each parameter. To recap, the ‘zero-sink’ uptake relation (Eq 2.3 in the main text) is following:

$$P_{uptake} = \sum_{i=1}^{12} \pi \cdot \Delta z \cdot L_{rv,i} \cdot D \cdot \frac{\rho^2 - 1}{G(\rho)} \cdot C_P \tag{B.5}$$

$P_{uptake}$  is the cumulative root uptake during one year ( $\text{kgPha}^{-1} \text{ year}^{-1}$ ),  $\Delta z$  is soil height (30 cm),  $L_{rv}$  is monthly root length density ( $\text{cm cm}^{-3}$ ),  $D$  is the coefficient of P diffusion ( $\text{cm}^2 \text{ day}^{-1}$ ),  $C_P$  is the mean concentration of  $\text{PO}_4^{-3}$  ions in the top 30 cm ( $\text{mgP L}^{-1}$ ),  $\rho$  ratio of soil cylinder to root radius (-) and  $G(\rho)$  a dimensionless geometric function (-) related to uptake by diffusion only.  $L_{rv,i}$  is the monthly root length calculated by multiplying the ORCHIDEE-CROP root biomass with the specific root length (SRL) parameter, which was taken at the maturity stage. Values and their standard errors are given in Table B.3.  $D$  is the coefficient of P diffusion given by the ‘constant slope impedance factor’ model (Olesen et al., 2001). It mimics the decreased solute diffusivity in unsaturated soils, while taking into account soil texture and bulk density. The decrease in diffusivity  $f$  is

called impedance (-) and is the ratio of diffusivity in soil and water:

$$f = \frac{D_s}{D_0} = 1.1 \cdot (\theta - \theta_{th}) \quad (B.6)$$

$D_s$  is the P diffusivity in soil ( $\text{m}^2 \text{s}^{-1}$ ),  $D_0$  is the P diffusivity in water ( $\text{m}^2 \text{s}^{-1}$ ),  $\theta$  is the soil water content ( $\text{m}^3 \text{m}^{-3}$ ) and  $\theta_{th}$  is the diffusivity cut-off point ( $\text{m}^3 \text{m}^{-3}$ ).  $\theta$  is the modeled soil water content, and comes from the ORCHIDEE hydrology module called SECHIBA (Krinner et al., 2005). The impedance decreases with  $\theta$ , when there is less water filled pore space and the diffusivity path (or tortuosity) increases. The soil water content cut-off point  $\theta_{th}$  is a parameter which depends on soil texture and bulk density:

$$\begin{aligned} \theta_{th} = & + 0.81 \cdot CF - 0.90 \cdot CF^2 \\ & - 0.60 \cdot \rho_b + 0.22 \cdot \rho_b^2 \\ & - 0.07 \cdot SF + 0.42 + RMSE_{\theta_{th}} \end{aligned} \quad (B.7)$$

$CF$  is the clay fraction (-),  $SF$  is the silt fraction (-) and  $\rho_b$  is the bulk density ( $\text{g cm}^{-3}$ ). Global soil texture and bulk density information was read from the Soilgrids1km maps (Hengl et al., 2014). Additional error to  $\theta_{th}$  was added according to it's RMSE value ( $= 0.034 \text{ m}^3 \text{m}^{-3}$ ) when performing Monte-Carlo.  $\rho$  is the normalized radius (-) and  $G(\rho)$  is the geometry function (-) which arise in deriving the 'zero-sink' solution in the original model (De Willigen and van Noordwijk, 1994). They are defined as following:

$$\begin{aligned} \rho &= \frac{1}{R_0 \cdot \sqrt{\pi \cdot L_{rv}}} \\ G(\rho) &= \frac{1}{2} \cdot \left( \frac{1 - 3 \cdot \rho^2}{4} + \frac{\rho^4 \cdot \log(\rho)}{\rho^2 - 1} \right) \end{aligned} \quad (B.8)$$

$R_0$  is the root radius (mm) and  $L_{rv}$  is the root length density ( $\text{cm cm}^{-3}$ ). Value and their standard errors are given in Table B.3. The procedure for calculation of P in soil solution ( $C_P$ ) is explained in the following section. For Monte Carlo estimation of uncertainty, all of the previous parameters were varied around their mean value with their standard error. If the standard error wasn't provided, we assumed a CV of 5 to 20% depending on the parameter in question (details in Tables B.1 - B.3)

#### B.1.4 P in soil solution

The soil P dynamic model (Ringeval et al., 2017) does not represent P in soil solution ( $C_P$ ), but only inorganic labile P ( $P_{ILAB}$ ) as the plant available P.  $C_P$  is needed for the potential root uptake relation (Eq. B.5). To bridge this gap we had to use a separate database (Achat et al., 2016) to infer the soil solution P from the labile P pool. The dataset contained results

from Hedley fractionation studies in diverse ecosystems (forests, grasslands and croplands) covering 379 data points. The depths of soil samples covered the profile up to a depth of 165 cm and were classified according to the USDA system. Samples deeper than 40 cm were discarded (around 10%) since the soil P model covers only the first 30 cm. All of the data points had measurements of either soil solution ( $C_P$ , mg dm<sup>-3</sup> of water), resin, inorganic bicarbonate P or total inorganic labile P. Inorganic labile ( $P_{ILAB}$ , mg kg<sup>-1</sup> of soil) was calculated as a sum of the resin and the inorganic bicarbonate fractions if provided.  $C_P$  was taken as it was measured. These values were log transformed and a Freundlich type relation was fitted:

$$\log(P_{ILAB}) = a \cdot \log(C_P) + b \quad (\text{B.9})$$

When fitting the parameters, we assumed an identical slope  $a$  across all soil orders and a soil order dependent  $b$  (Fig. B.1). The fitted parameters are  $a_{GLOBAL} = 4.28145 \pm 0.12071$  ( $p < 0.001$ ) and  $b_{GLOBAL} = 0.62882 \pm 0.05787$  ( $p < 0.001$ ). The soil-type dependent bias for the  $b$  parameter ( $\Delta b_{SOILTYPE} = b_{SOILTYPE} - b_{GLOBAL}$ ) was found significant only for Oxisols and Mollisols. The values are  $\Delta b_{Mollisol} = 0.941379 \pm 0.327989$  ( $p < 0.001$ ) and  $\Delta b_{Oxisol} = 1.573335 \pm 0.310246$  ( $p < 0.05$ ).

### B.1.5 Supplementary Tables

*Supplementary Table B.1: Parameters used to calculate P demand ( $\text{kgP ha}^{-1}$ ) with the PHI method (Section 2.2.2, Eq. 2.1). Mean values are shown with their standard error. If a standard error wasn't provided in the source material, a CV of 20% was assumed. Grain information was obtained from Duivenbooden, Wit, and Keulen (1995). Root information for wheat, maize and rice was obtained respectively from Hocking (1994), Latshaw and Miu (1934) and Ye et al. (2014).*

Description	Parameter	Unit	Wheat	Maize	Rice
Grain P concentration at maturity	$P_{\%,grain}$	$\text{g kg}^{-1}$	$3.70 \pm 0.80$	$2.90 \pm 0.80$	$2.00 \pm 0.80$
Amount of P in grain vs the shoot	$PHI$	-	$0.78 \pm 0.07$	$0.67 \pm 0.13$	$0.61 \pm 0.13$
Harvest index	$HI$	-	$0.42 \pm 0.12$	$0.44 \pm 0.08$	$0.41 \pm 0.08$
Root P concentration at maturity	$P_{\%,root}$	$\text{g kg}^{-1}$	1.01	1.2	$1.31 \pm 0.21$

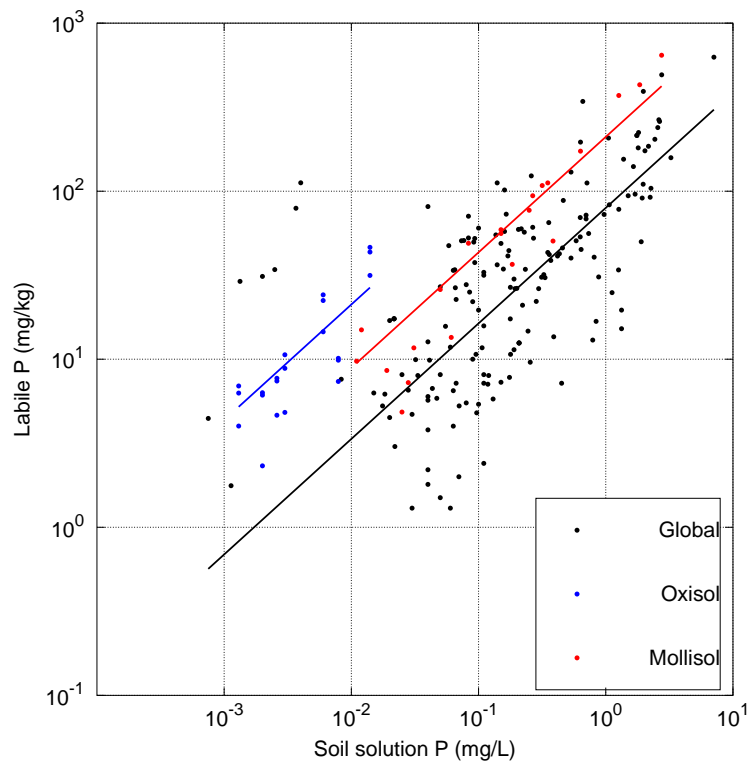
*Supplementary Table B.2: Parameters used to calculate P demand ( $\text{kgP ha}^{-1}$ ) with the C:P ratio method (Section 2.2.2, Eq. 2.2). Mean values are shown with their standard error. If a standard error wasn't provided in the source material, a CV of 20% was assumed for P concentration and 5% for C concentrations. As the reference paper for wheat didn't contain C concentrations, we've assumed C concentration of 45%. Wheat, maize and rice information was obtained respectively from Hocking (1994), Latshaw and Miu (1934) and Ye et al. (2014).*

Description	Parameter	Unit	Wheat	Maize	Rice
Grain P concentration	$P_{\%,grain}$	$\text{g kg}^{-1}$	3.78	$3.42 \pm 0.34$	$3.58 \pm 0.15$
Leaf P "	$P_{\%,leaf}$	$\text{g kg}^{-1}$	0.69	$2.08 \pm 0.34$	$1.64 \pm 0.14$
Stem P "	$P_{\%,stem}$	$\text{g kg}^{-1}$	0.54	$0.89 \pm 0.22$	$1.64 \pm 0.14$
Root P "	$P_{\%,root}$	$\text{g kg}^{-1}$	1.01	$1.20 \pm 0.15$	$1.31 \pm 0.21$
Grain C "	$C_{\%,grain}$	%	45.00	$44.72 \pm 0.50$	$40.37 \pm 0.17$
Leaf C "	$C_{\%,leaf}$	%	45.00	$41.27 \pm 0.58$	$38.04 \pm 0.27$
Stem C "	$C_{\%,stem}$	%	45.00	$44.51 \pm 1.05$	$38.04 \pm 0.27$
Root C "	$C_{\%,root}$	%	45.00	42.31	$38.04 \pm 1.68$

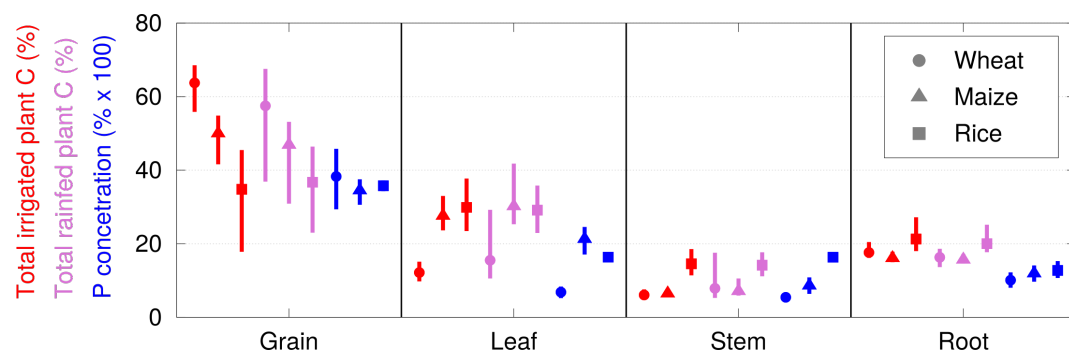
Supplementary Table B.3: Values used to determine  $L_{rv}$  ( $\text{cm cm}^{-3}$ ),  $\rho$  and  $G_\rho$  in the 'zero-sink' uptake model (Eq. B.5). When assessing uncertainty, a CV of 20% was assumed for each of the variables. Wheat, maize and rice information was obtained respectively from Nakhforoosh et al. (2014), Li et al. (2016) and Biscarini et al. (2016).

Description	Parameter	Unit	Wheat	Maize	Rice
Specific root length	$SRL$	$\text{m g}^{-1} \text{ root}$	129	74	146
Root diameter	$d_{root}$	mm	0.42	0.28	0.23

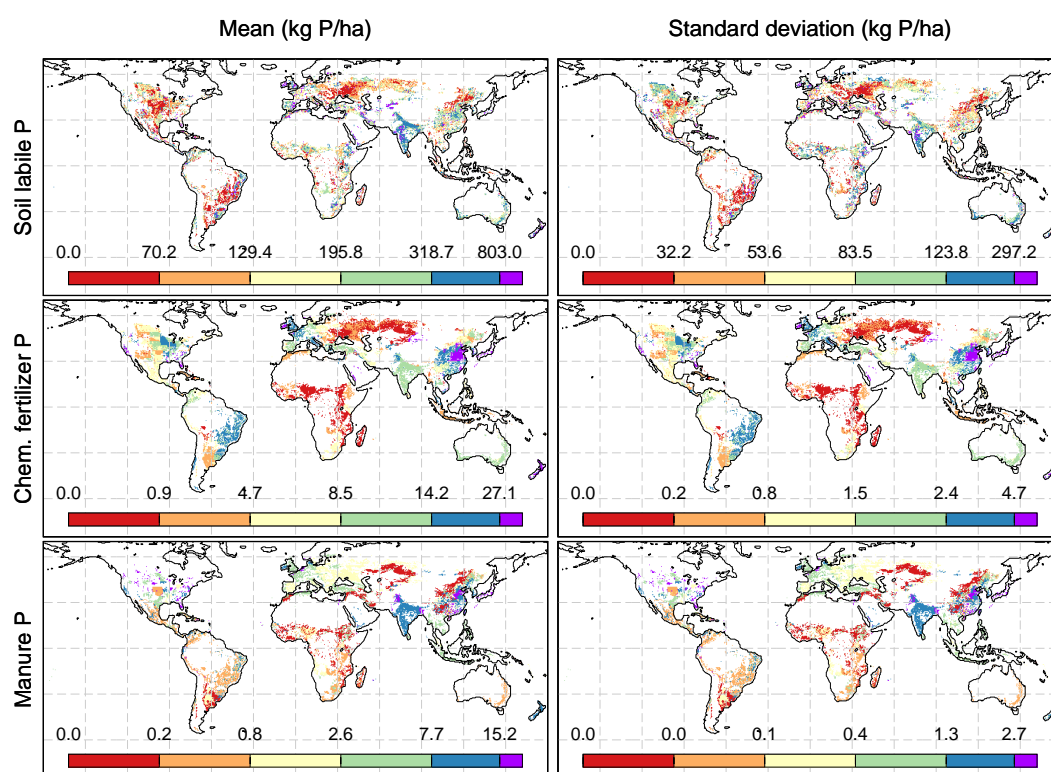
### B.1.6 Supplementary Figures



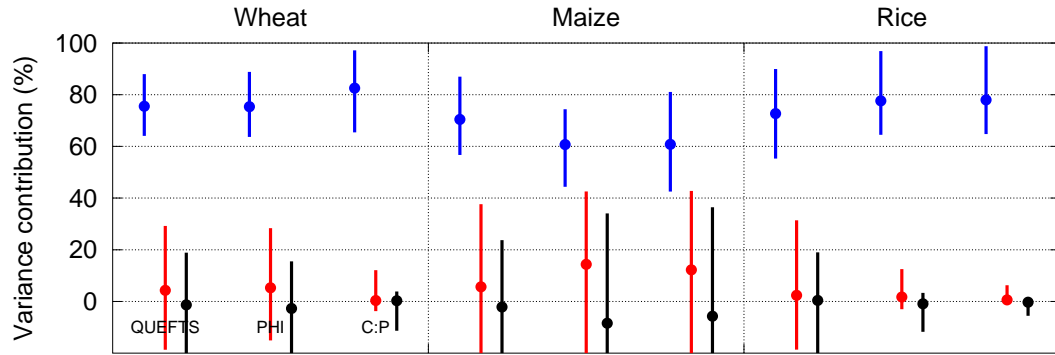
Supplementary Figure B.1: Soil solution ( $C_P$ ,  $\text{mg dm}^{-3}$ ) vs. inorganic labile P ( $P_{ILAB}$ ,  $\text{mg kg}^{-1}$ ) data used to obtain the Freundlich type relationship (Eq. B.9). Points are measurements and solid lines are the modelled values. Different colors represent the values for Mollisols (red), Oxisols (blue) and all other USDA soil types (dark).



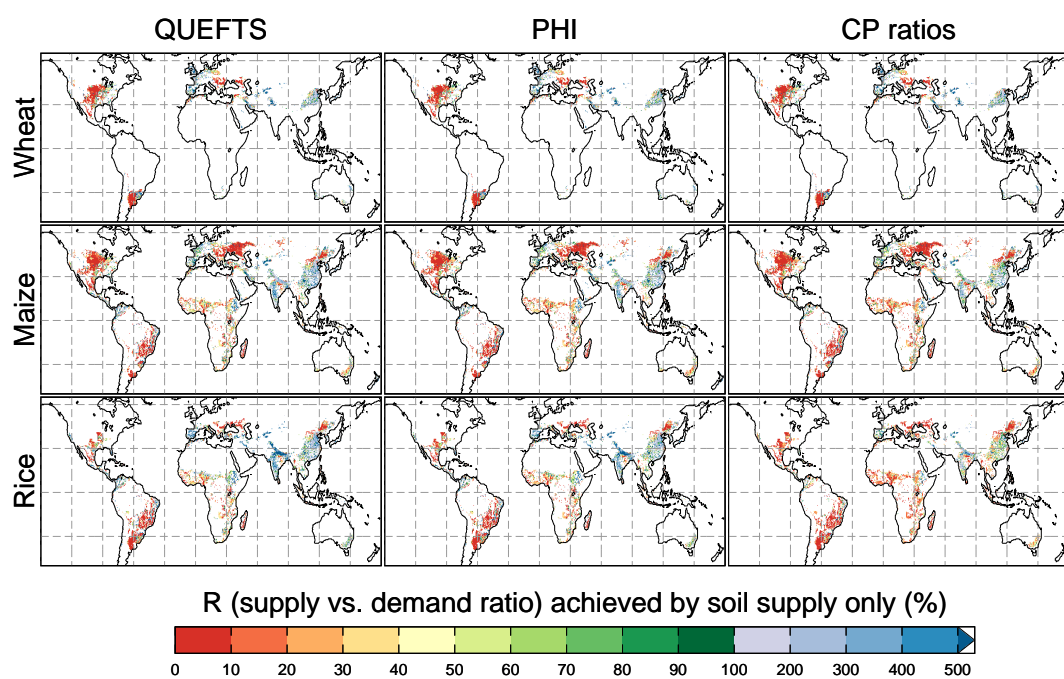
Supplementary Figure B.2: Global values of total plant C fraction at maturity in irrigated (red) and rainfed simulations (violet), as well as organ P concentrations used in the C:P ratio method (blue). Markers shows the global median values with a 68% quantile interval around it. Different markers correspond to different species. P concentration values are given in Table B.2.



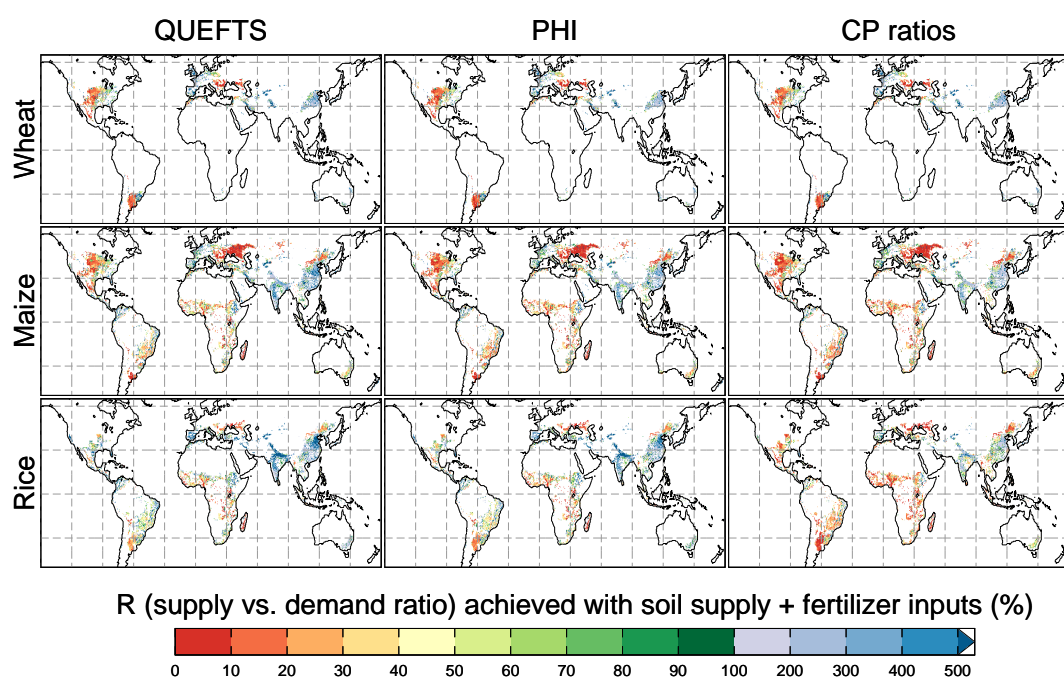
Supplementary Figure B.3: Maps of soil labile, chemical fertilizer and manure P ( $\text{kgP ha}^{-1}$ ) for the year 2000 (top to bottom). Maps shows the mean value (left column) and it's standard deviation (right column) coming from 30 different simulations as part of the soil P model's sensitivity test 7, where the main drivers randomly varied in a predefined interval. Colorbars depict global quantiles (0-20-80, 95)



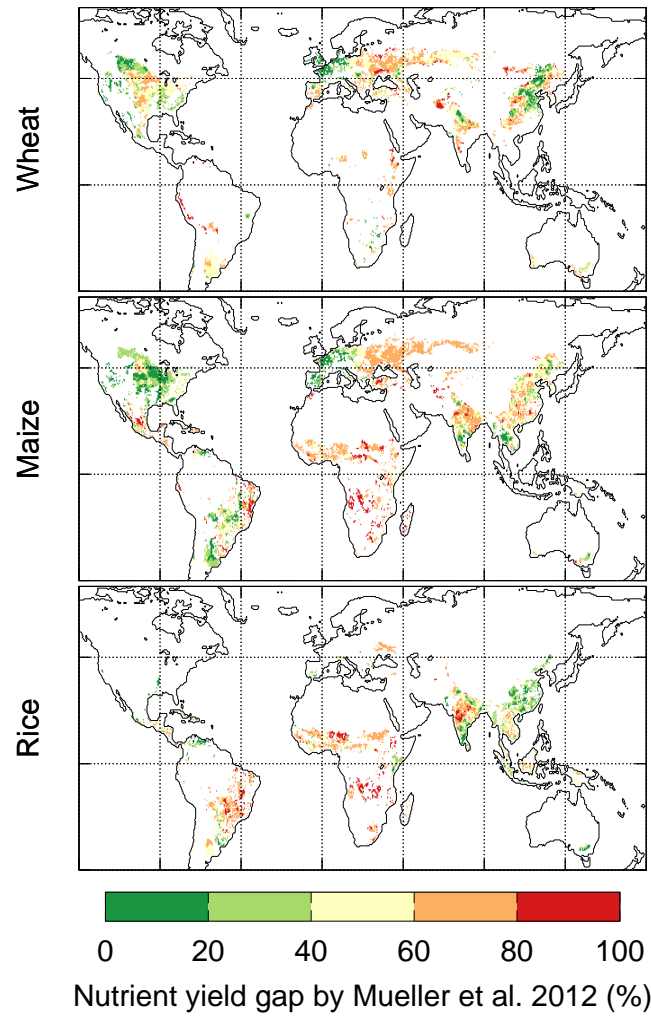
Supplementary Figure B.4: Commonality analysis of the potential P root uptake (Eq. B.5). The plots show the contribution of each of the drivers ( $P_{ILAB}$ ,  $C_P$ ,  $D$ ,  $L_{tv}$ ) to the spatial variability (top) and variability at each grid point (bottom) of the calculated root P uptake. Values 1 to 4 show the unique contribution not shared with others. Other combinations (12 to 1234) show variance common to all variables within that combination. Different markers and colors correspond to different crop species (see legend). In the top plot, errorbars show the 95% uncertainty interval of the global median (obtained with 1000 bootstrap samples). In the bottom plot, errorbars show the 68% quantile interval around the global median value.



Supplementary Figure B.5: P limitation as a ratio of supply vs. demand (%), calculated using only the potential soil P supply. Rows are different species (top to bottom: wheat, maize, rice). Grid points without any grain were not plotted.



Supplementary Figure B.6: P limitation as a ratio of supply vs. demand (%), calculated using the potential soil P supply + annual fertilizer P (chemical and manure). Rows are different species (top to bottom: wheat, maize, rice). Grid points without any grain were not plotted.



Supplementary Figure B.7: Nutrient yield gap (%) according to Mueller et al. (2012). The yield gap was calculated as the ratio of mean observed 10 and the attainable yield. Attainable yield was calculated by dividing the world into climate bins (from 100 to 400) and calculating the 95th percentile yield within the same climate bin.

## B.2 Chapter 3 supplementary information

### B.2.1 Supplementary Tables

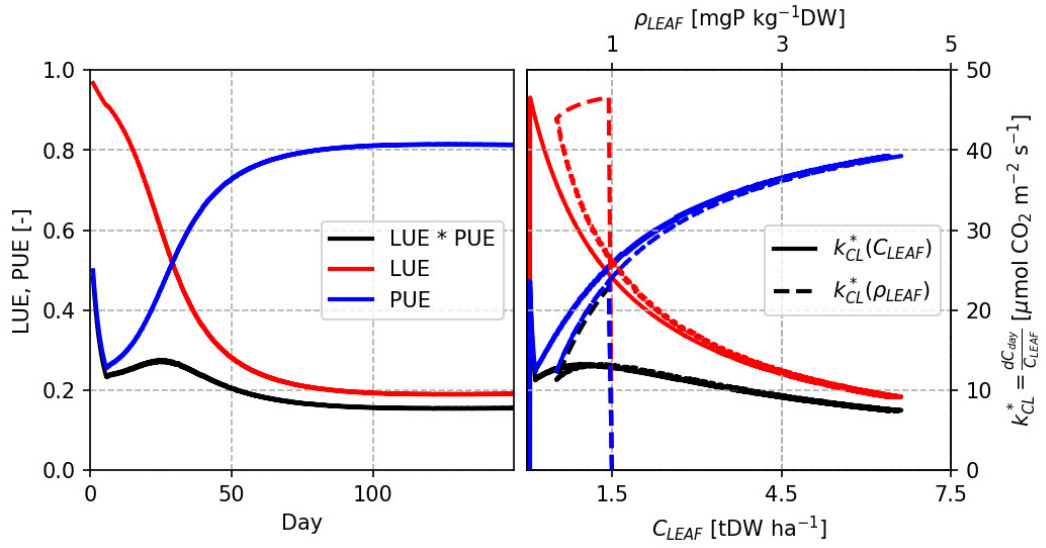
*Supplementary Table B.4: Calibrated values for the maize field trial. Parameter  $k_{CR}$  is given for each of the 12  $P$  experiments (3  $P$  levels  $\times$  4 blocks) and all others kept the same. Parameter description can be found in Tables 3.1 and 3.2.*

Parameter	Initial guess	Calibrated values $\pm$ std. error
$k_{CL}$	1.5	$1.7137 \pm 0.1106$
$k_{CR,1}$	1.0	$0.9101 \pm 0.1017$
$k_{CR,2}$	1.0	$0.8302 \pm 0.0922$
$k_{CR,3}$	1.0	$0.8545 \pm 0.0925$
$k_{CR,4}$	1.0	$0.8376 \pm 0.0969$
$k_{CR,5}$	1.0	$1.4577 \pm 0.1673$
$k_{CR,6}$	1.0	$1.4526 \pm 0.1529$
$k_{CR,7}$	1.0	$1.1963 \pm 0.1111$
$k_{CR,8}$	1.0	$1.4287 \pm 0.1573$
$k_{CR,9}$	1.0	$1.7878 \pm 0.2064$
$k_{CR,10}$	1.0	$1.8376 \pm 0.1784$
$k_{CR,11}$	1.0	$1.7906 \pm 0.1966$
$k_{CR,12}$	1.0	$1.8377 \pm 0.1859$
$C_{L,max}$	1.5	$0.9042 \pm 0.0862$
$\rho_{L,max}$	1.0	$1.0927 \pm 0.0904$
$\lambda_{CR}$	0.30	$0.1362 \pm 0.0291$
$\lambda_{CL}$	0.10	$0.1608 \pm 0.0141$
$\lambda_{CS}$	0.03	$0.0117 \pm 0.0020$
$\lambda_{CG}$	0.01	$0.0153 \pm 0.0048$
$f_{CS}$	0.5	$0.4036 \pm 0.0495$
$f_{CG}$	1.0	$2.3523 \pm 0.2097$
$f_{PS}$	0.8	$1.5956 \pm 0.0660$
$f_{PR}$	0.1	$0.0399 \pm 0.0850$
$k_{mPL}$	0.05	$0.0584 \pm 0.0081$
$k_{mPR}$	0.05	$0.2541 \pm 0.1193$
$k_{mPS}$	0.05	$0.0686 \pm 0.0096$
$C_{L,0}$	0.1	$0.0064 \pm 0.0250$
$P_{L,0}$	0.1	$0.0298 \pm 0.0180$

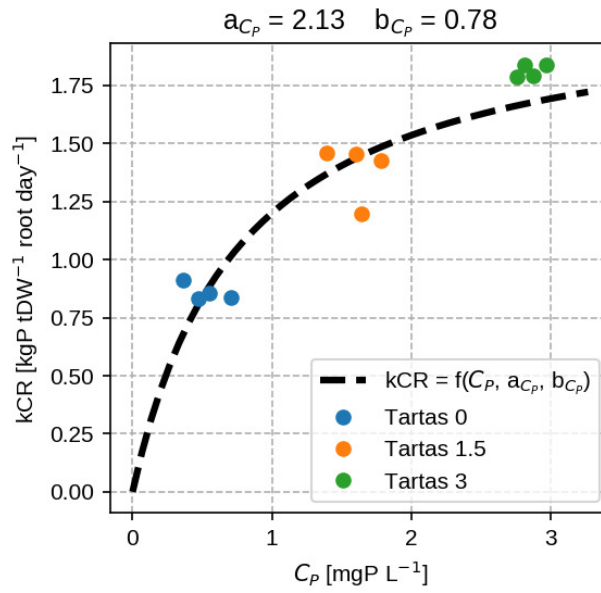
Supplementary Table B.5: Calibrated parameter values for the hydroponic studies. Parameter  $k_{CR}$  is given for each of the five  $P$  experiments and all others kept the same for each species (Asher and Loneragan, 1967; N. K. Fageria and Baligar, 1989). Parameter description can be found in Tables 3.1 and 3.2.

Variable	$k_{CL}$	$C_{L,max}$	$\rho_{L,max}$	$\lambda_{CR}$	$\lambda_{CL}$	$C_{L,0}$
Initial guess	1.5	1.0	1.5	0.3	0.1	0.1
Barrel medic	$1.2217 \pm 1.2479$	$1.4816 \pm 0.9336$	$0.8967 \pm 0.2541$	$0.3261 \pm 0.4065$	$0.1195 \pm 0.1727$	$0.0441 \pm 0.2109$
Brome grass	$1.8221 \pm 0.5271$	$0.4289 \pm 0.0644$	$1.2571 \pm 0.2453$	$0.4145 \pm 0.1946$	$0.1187 \pm 0.0651$	$0.1211 \pm 0.4818$
Capeweed	$1.2570 \pm 0.5783$	$1.6330 \pm 1.1550$	$0.9148 \pm 0.4691$	$0.2489 \pm 1.2535$	$0.1224 \pm 0.4597$	$0.0524 \pm 0.8952$
Clover	$1.7286 \pm 0.3377$	$0.4744 \pm 0.1035$	$1.1299 \pm 0.2715$	$0.4721 \pm 0.1357$	$0.0847 \pm 0.0482$	$0.1395 \pm 0.1960$
Erodium	$2.0614 \pm 2.4349$	$0.7016 \pm 1.0327$	$1.1258 \pm 1.3077$	$0.6225 \pm 1.1166$	$0.0720 \pm 0.3201$	$0.0956 \pm 0.5452$
Flatweed	$2.0698 \pm 0.6796$	$0.4879 \pm 0.1347$	$2.0429 \pm 0.4540$	$0.4136 \pm 0.2458$	$0.1366 \pm 0.1897$	$0.1208 \pm 0.2399$
Silver grass	$1.7724 \pm 0.7744$	$0.4667 \pm 0.2487$	$1.4596 \pm 1.0746$	$0.4734 \pm 0.7083$	$0.1048 \pm 0.2218$	$0.1309 \pm 0.4042$
Alfalfa	$1.7676 \pm 1.0372$	$0.4844 \pm 0.1900$	$1.7205 \pm 0.4770$	$0.4672 \pm 0.5006$	$0.1132 \pm 0.2824$	$0.1419 \pm 0.4264$
Bean	$1.4107 \pm 0.4049$	$0.3595 \pm 0.0642$	$1.3110 \pm 0.2805$	$0.5136 \pm 0.3943$	$0.1099 \pm 0.1339$	$0.1300 \pm 0.3123$
Red clover	$1.5970 \pm 1.4740$	$0.3258 \pm 0.2304$	$1.9161 \pm 0.7067$	$0.3124 \pm 0.3857$	$0.1165 \pm 0.1431$	$0.0873 \pm 0.5166$
Rice	$1.4473 \pm 0.9212$	$0.1676 \pm 0.0903$	$3.7253 \pm 1.1436$	$0.3411 \pm 1.1679$	$0.1157 \pm 0.6166$	$0.0858 \pm 0.8162$
Wheat	$1.7524 \pm 0.7460$	$0.6345 \pm 0.2700$	$1.2003 \pm 0.4162$	$0.3365 \pm 0.3818$	$0.1091 \pm 0.2223$	$0.0447 \pm 0.5497$
Variable	$P_{L,0}$	$k_{CR,1}$	$k_{CR,2}$	$k_{CR,3}$	$k_{CR,4}$	$k_{CR,5}$
Initial guess	0.1	1.0	1.0	1.0	1.0	1.0
Barrel medic	$0.0214 \pm 0.5136$	$0.1588 \pm 3.3886$	$0.6731 \pm 2.0856$	$0.8325 \pm 1.6676$	$1.1213 \pm 1.0650$	$1.2884 \pm 0.7138$
Brome grass	$0.0723 \pm 0.2866$	$0.0360 \pm 0.1446$	$0.2408 \pm 0.2552$	$1.1367 \pm 0.3647$	$1.3340 \pm 0.4073$	$1.8140 \pm 0.3765$
Capeweed	$0.0609 \pm 0.3320$	$0.2978 \pm 0.4454$	$0.4815 \pm 0.5419$	$0.5965 \pm 0.6870$	$0.8538 \pm 0.8076$	$1.1816 \pm 0.9310$
Clover	$0.1203 \pm 0.2558$	$0.0348 \pm 0.0862$	$0.2690 \pm 0.1442$	$1.1347 \pm 0.3132$	$2.2824 \pm 0.6717$	$2.0929 \pm 0.5477$
Erodium	$0.0415 \pm 1.6314$	$0.0215 \pm 1.1265$	$0.3436 \pm 2.0639$	$0.7719 \pm 2.3206$	$1.6700 \pm 2.8556$	$1.8104 \pm 3.0153$
Flatweed	$0.0961 \pm 0.4203$	$0.0782 \pm 0.1879$	$0.3508 \pm 0.1543$	$0.9953 \pm 0.1750$	$1.6565 \pm 0.4316$	$1.4294 \pm 0.2642$
Silver grass	$0.1066 \pm 1.3494$	$0.0809 \pm 0.5959$	$0.3281 \pm 0.6141$	$0.7678 \pm 0.4249$	$1.8235 \pm 0.6355$	$1.7826 \pm 0.8745$
Alfalfa	$0.0713 \pm 0.8038$	$0.0259 \pm 0.4894$	$0.2313 \pm 0.4821$	$0.9014 \pm 0.3704$	$1.5848 \pm 0.3303$	$2.1503 \pm 0.6900$
Bean	$0.0464 \pm 0.1495$	$0.0630 \pm 0.1667$	$0.5332 \pm 0.2332$	$0.8130 \pm 0.2835$	$1.2538 \pm 0.3362$	$1.4391 \pm 0.3720$
Red clover	$0.0719 \pm 0.7136$	$0.1330 \pm 0.5427$	$0.4411 \pm 0.5719$	$0.6859 \pm 0.4559$	$0.8600 \pm 0.5091$	$1.0676 \pm 0.4154$
Rice	$0.0468 \pm 0.4628$	$0.0548 \pm 0.8097$	$0.4648 \pm 0.5638$	$1.4896 \pm 0.9318$	$1.2552 \pm 0.7613$	$1.2643 \pm 0.8733$
Wheat	$0.0440 \pm 0.3336$	$0.1627 \pm 0.4041$	$0.2031 \pm 0.4443$	$0.5298 \pm 0.6615$	$0.8987 \pm 1.0080$	$1.0202 \pm 0.9067$

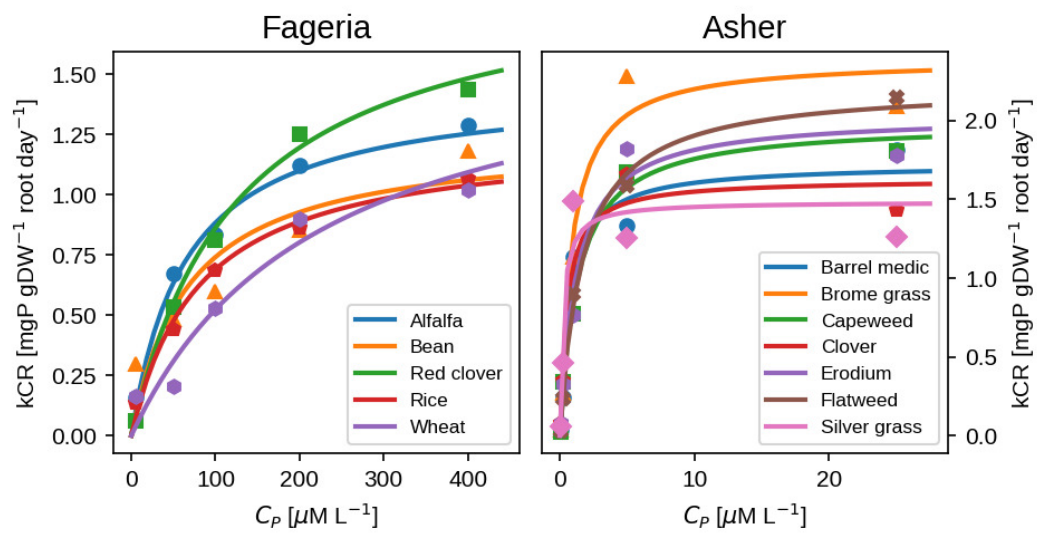
## B.2.2 Supplementary Figures



Supplementary Figure B.8: Time evolution of leaf LUE and PUE (left) and the effective assimilation rate ( $k_{CL}^*$ , right) as function of leaf biomass and concentration. Root uptake rate  $k_{CR}$  is set to  $1.0 \text{ kgP tDW}^{-1} \text{ root ha}^{-1} \text{ day}^{-1}$ .



Supplementary Figure B.9: Predicted field trial root P uptake ( $k_{CR}$ ) vs. measured soil solution P concentration. Different markers and colors depict different levels of P addition. Dark dashed line depicts a Michaelis-Menten kinetic (similar to Eq. 3.16) which best describes the relationship between the two.



Supplementary Figure B.10: Predicted hydroponic studies root P uptake ( $k_{CR}$ ) vs. measured solution P concentration. Different markers and colors depict different levels of P addition. Dark dashed line depicts a Michaelis-Menten kinetic (similar to Eq. 3.16) which best describes the relationship between the two.

## B.3 Chapter 4 supplementary information

### B.3.1 Supplementary Tables

Supplementary Table B.6: Nebraska initial guess

Name	Initial guess	Variable	Description	Unit
VCMAX25	60	$V_{C,max}$	Max. carboxylation rate	$\mu\text{mol C m}^{-2} \text{ s}^{-1}$
SLA	0.03	$SLA$	Specific leaf area	$\text{m}^2 \text{ gC}^{-1}$
MAINT_RESP_SLOPE_C	0.03	$\lambda_0$	Respiration rate at 0 C°	$\text{day}^{-1}$
CM_ZERO_LEAF	0.02	$\lambda_{TL}$	Leaf respiration temperature slope	$\text{C}^\circ^{-1}$
CM_ZERO_ROOT	0.02	$\lambda_{TR}$	Root respiration temperature slope	$\text{C}^\circ^{-1}$
CM_ZERO_SAPABOVE	0.005	$\lambda_{TS}$	Stem respiration temperature slope	$\text{C}^\circ^{-1}$
CM_ZERO_FRUIT	0.0001	$\lambda_{TG}$	Grain respiration temperature slope	$\text{C}^\circ^{-1}$
CM_ZERO_RESERVE	0.0001	$\lambda_{TX}$	Reserve respiration temperature slope	$\text{C}^\circ^{-1}$
CROPT_kCR	1	$k_{CR}$	Root P uptake rate	$\text{mgP gC}^{-1} \text{ root day}^{-1}$
CROPT_CLmax	60	$C_{L,max}$	LUE half-saturation biomass	$\text{gC m}^{-2}$
CROPT_PLmax	1	$\rho_{L,max}$	PUE half-saturation concentration	$\text{mgP gC}^{-1}$
CROPT_frCG	2	$f_{CG}$	Grain C filling fraction	-
CROPT_frCS	1	$f_{CS}$	Stem C filling fraction	-
CROPT_frPR	0.1	$f_{PR}$	Root P filling fraction	-
CROPT_frPS	0.1	$f_{PS}$	Stem P filling fraction	-
CROPT_kmPL	0.01	$k_{mPL}$	Leaf P remobilisation rate	-
CROPT_kmPR	0.01	$k_{mPR}$	Root P remobilisation rate	-
CROPT_kmPS	0.01	$k_{mPS}$	Stem P remobilisation rate	-
CROPT_tini	120	$t_{ini}$	Planting date	day of year

Supplementary Table B.7: Tartas initial guess

Name	Initial guess	Variable	Description	Unit
VCMAX25	60	$V_{C,max}$	Max. carboxylation rate	$\mu\text{mol C m}^{-2} \text{ s}^{-1}$
SLA	0.03	$SLA$	Specific leaf area	$\text{m}^2 \text{ gC}^{-1}$
MAINT_RESP_SLOPE_C	0.03	$\lambda_0$	Respiration rate at 0 C°	$\text{day}^{-1}$
CM_ZERO_LEAF	0.02	$\lambda_{TL}$	Leaf respiration temperature slope	$\text{C}^\circ^{-1}$
CM_ZERO_ROOT	0.02	$\lambda_{TR}$	Root respiration temperature slope	$\text{C}^\circ^{-1}$
CM_ZERO_SAPABOVE	0.005	$\lambda_{TS}$	Stem respiration temperature slope	$\text{C}^\circ^{-1}$
CM_ZERO_FRUIT	0.0001	$\lambda_{TG}$	Grain respiration temperature slope	$\text{C}^\circ^{-1}$
CM_ZERO_RESERVE	0.0001	$\lambda_{TX}$	Reserve respiration temperature slope	$\text{C}^\circ^{-1}$
CROPT_kCR_00	0.10	$k_{CR,1}$	P1, B1 root P uptake rate	$\text{mgP gC}^{-1} \text{ root day}^{-1}$
CROPT_kCR_01	0.10	$k_{CR,2}$	P1, B2 "	"
CROPT_kCR_02	0.10	$k_{CR,3}$	P1, B3 "	"
CROPT_kCR_03	0.10	$k_{CR,4}$	P1, B4 "	"
CROPT_kCR_04	0.33	$k_{CR,5}$	P2, B1 "	"
CROPT_kCR_05	0.33	$k_{CR,6}$	P2, B2 "	"
CROPT_kCR_06	0.33	$k_{CR,7}$	P2, B3 "	"
CROPT_kCR_07	0.33	$k_{CR,8}$	P2, B4 "	"
CROPT_kCR_08	1.00	$k_{CR,9}$	P3, B1 "	"
CROPT_kCR_09	1.00	$k_{CR,10}$	P3, B2 "	"
CROPT_kCR_10	1.00	$k_{CR,11}$	P3, B3 "	"
CROPT_kCR_11	1.00	$k_{CR,12}$	P3, B4 "	"
CROPT_CLmax	60	$C_{L,max}$	LUE half-saturation biomass	$\text{gC m}^{-2}$
CROPT_PLmax	1	$\rho_{L,max}$	PUE half-saturation concentration	$\text{mgP gC}^{-1}$
CROPT_frCG	2	$f_{CG}$	Grain C filling fraction	-
CROPT_frCS	1	$f_{CS}$	Stem C filling fraction	-
CROPT_frPR	0.1	$f_{PR}$	Root P filling fraction	-
CROPT_frPS	0.1	$f_{PS}$	Stem P filling fraction	-
CROPT_kmPL	0.01	$k_{mPL}$	Leaf P remobilisation rate	-
CROPT_kmPR	0.01	$k_{mPR}$	Root P remobilisation rate	-
CROPT_kmPS	0.01	$k_{mPS}$	Stem P remobilisation rate	-
CROPT_tini	120	$t_{ini}$	Planting date	day of year

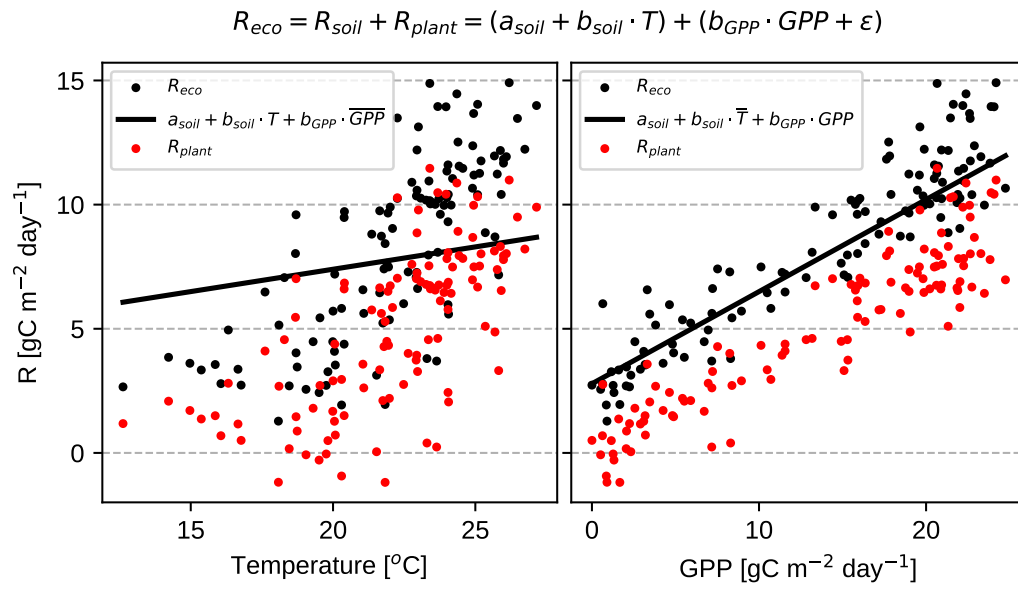
*Supplementary Table B.8: Nebraska calibrated values*

Name	Initial guess	Calibrated values ( $\pm$ std. error)
VCMAX25	60	$49.345800 \pm 0.667849$
SLA	0.03	$0.030902 \pm 0.000105$
MAINT_RESP_SLOPE_C	0.03	$0.026173 \pm 0.003043$
CM_ZERO_LEAF	0.02	$0.063503 \pm 0.015571$
CM_ZERO_ROOT	0.02	$0.014829 \pm 0.000655$
CM_ZERO_SAPABOVE	0.005	$0.005452 \pm 0.000210$
CM_ZERO_FRUIT	0.0001	$0.000213 \pm 0.000431$
CM_ZERO_RESERVE	0.0001	$0.000203 \pm \text{NaN}$
CROPT_kCR	0.10	$0.988825 \pm 0.035788$
CROPT_CLmax	60	$52.099800 \pm 0.647469$
CROPT_PLmax	1	$3.282030 \pm 0.298035$
CROPT_frCG	2	$1.157330 \pm 0.007158$
CROPT_frCS	1	$0.148733 \pm 0.084369$
CROPT_frPR	0.1	$0.148733 \pm 0.084369$
CROPT_frPS	0.1	$0.888210 \pm 0.208662$
CROPT_kmPL	0.01	$0.002189 \pm 0.000626$
CROPT_kmPR	0.01	$0.002443 \pm 0.000000$
CROPT_kmPS	0.01	$0.018644 \pm 0.000000$
CROPT_tini	120	$132.036000 \pm 1.232960$

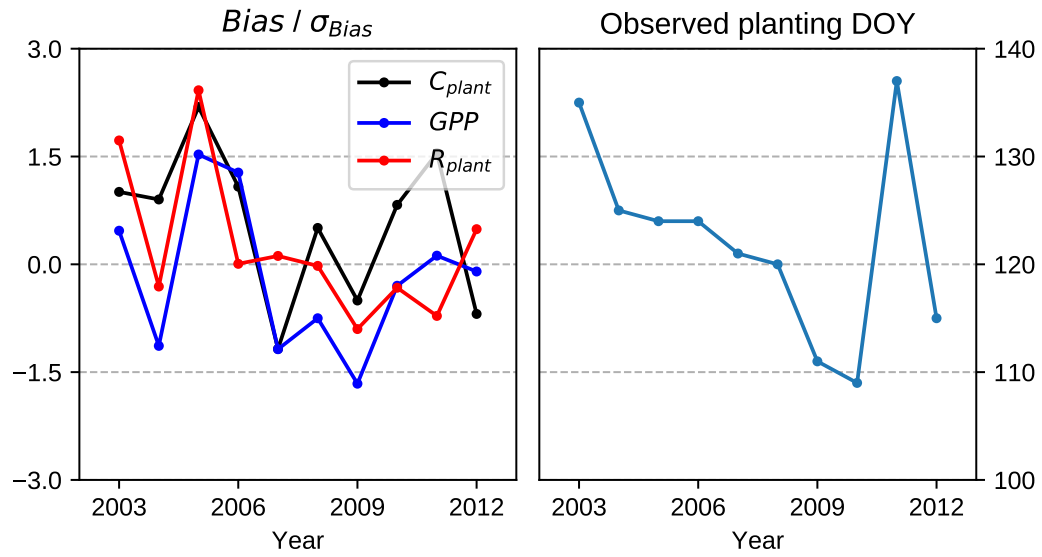
Supplementary Table B.9: Tartas calibrated values

Name	Initial guess	Calibrated values ( $\pm$ std. error)
VCMAX25	60	94.160000 $\pm$ 4.195620
SLA	0.03	0.032080 $\pm$ 0.000867
MAINT_RESP_SLOPE_C	0.03	0.041130 $\pm$ 0.007687
CM_ZERO_LEAF	0.02	0.034130 $\pm$ 0.003509
CM_ZERO_ROOT	0.02	0.082810 $\pm$ 0.012167
CM_ZERO_SAPABOVE	0.005	0.004911 $\pm$ 0.000931
CM_ZERO_FRUIT	0.0001	0.000605 $\pm$ 0.000991
CM_ZERO_RESERVE	0.0001	0.001000 $\pm$ NaN
CROPT_kCR_00	0.10	0.551200 $\pm$ 0.065082
CROPT_kCR_01	0.10	0.496700 $\pm$ 0.041635
CROPT_kCR_02	0.10	0.498500 $\pm$ 0.031745
CROPT_kCR_03	0.10	0.523700 $\pm$ 0.041312
CROPT_kCR_04	0.33	0.822100 $\pm$ 0.067137
CROPT_kCR_05	0.33	0.782800 $\pm$ 0.121015
CROPT_kCR_06	0.33	0.679600 $\pm$ 0.066054
CROPT_kCR_07	0.33	0.787200 $\pm$ 0.069152
CROPT_kCR_08	1.00	0.899600 $\pm$ 0.074646
CROPT_kCR_09	1.00	0.850800 $\pm$ 0.073362
CROPT_kCR_10	1.00	0.932300 $\pm$ 0.054426
CROPT_kCR_11	1.00	0.932000 $\pm$ 0.053944
CROPT_CLmax	60	98.910000 $\pm$ 4.519330
CROPT_PLmax	1	3.544000 $\pm$ 0.222627
CROPT_frCG	2	1.954000 $\pm$ 0.382329
CROPT_frCS	1	0.808400 $\pm$ 0.054284
CROPT_frPR	0.1	0.070440 $\pm$ 0.264154
CROPT_frPS	0.1	1.211000 $\pm$ 0.080423
CROPT_kmPL	0.01	0.016120 $\pm$ 0.002262
CROPT_kmPR	0.01	0.009299 $\pm$ 0.078614
CROPT_kmPS	0.01	0.015970 $\pm$ 0.001807
CROPT_tini	120	121.60000 $\pm$ 1.850340

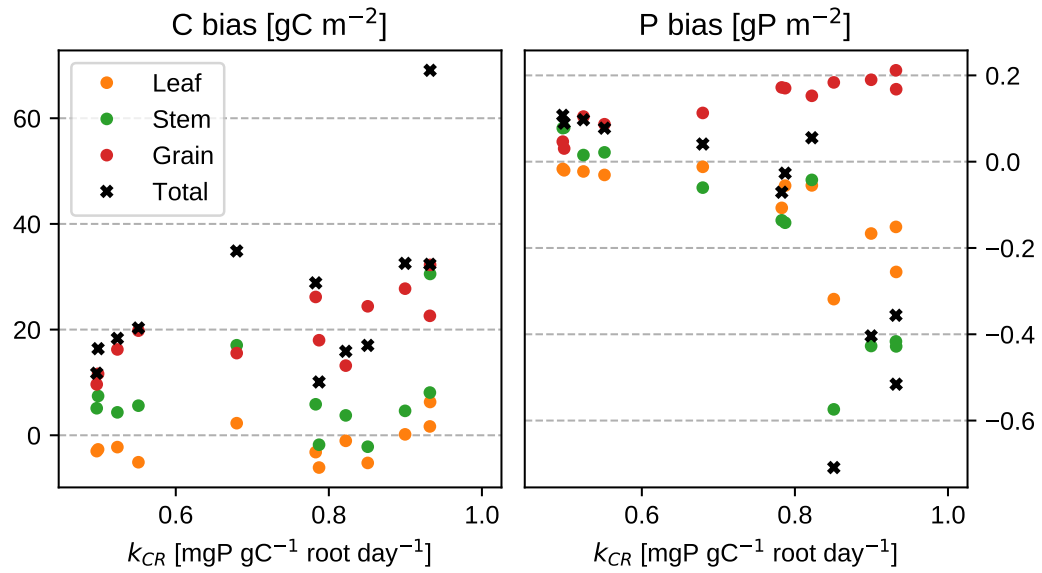
### B.3.2 Supplementary Figures



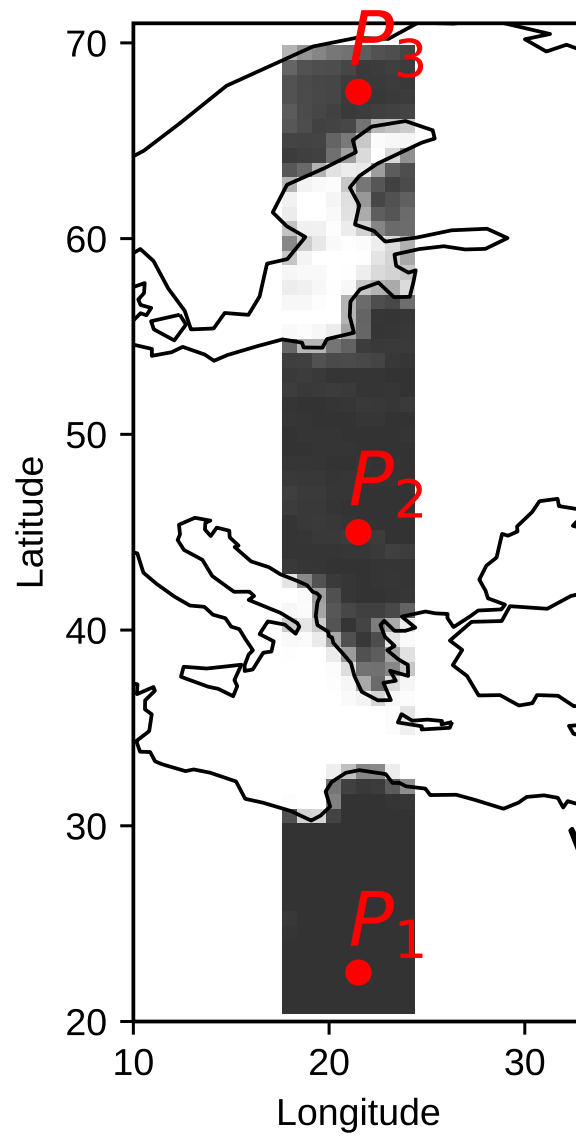
Supplementary Figure B.11: Nebraska GPP and eco-system respiration ( $R_{eco}$ ) using Eq. 4.23



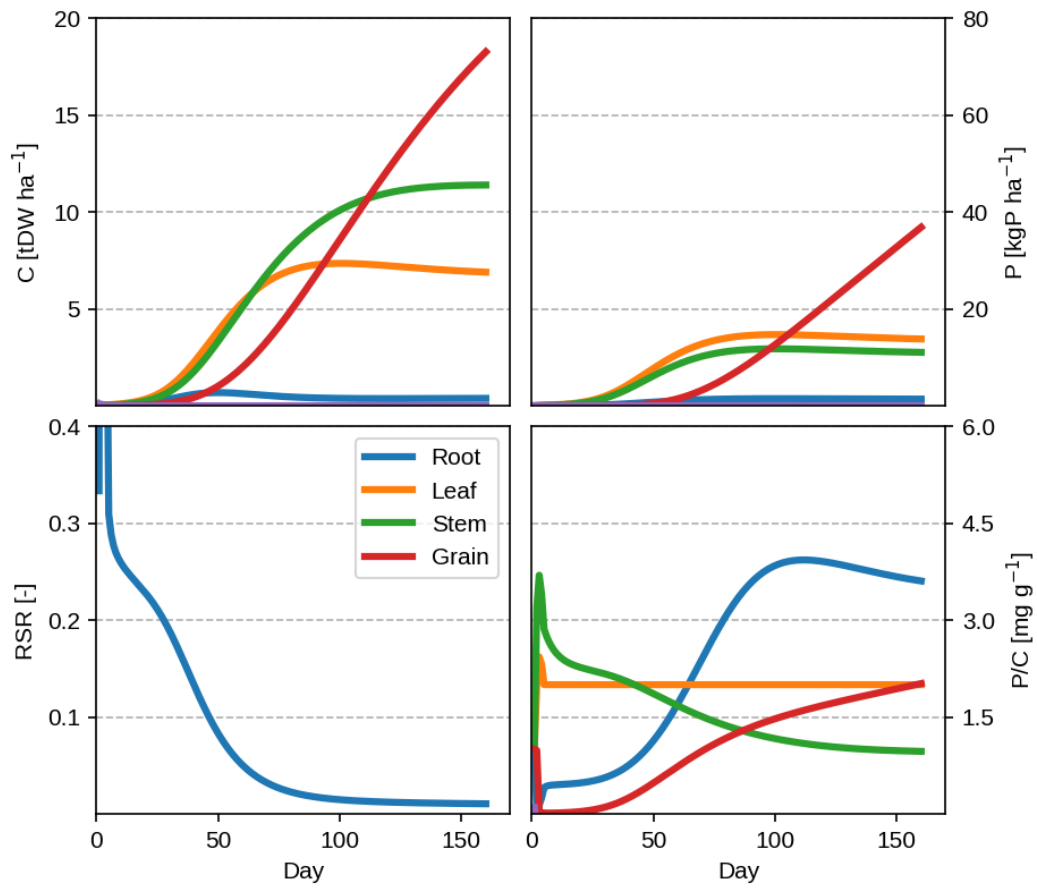
Supplementary Figure B.12: Nebraska modeled bias (left) and the observed planting dates (right)



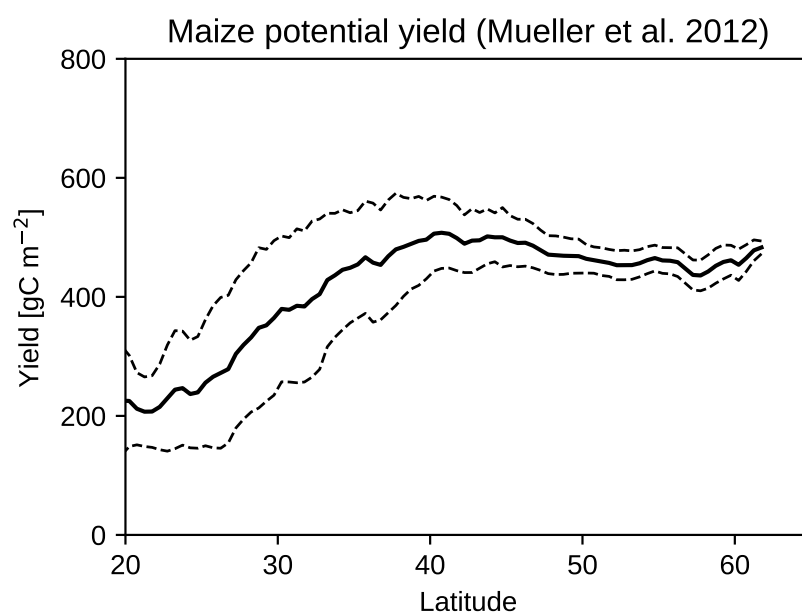
Supplementary Figure B.13: Tartas modeled bias as function of root P uptake rate ( $k_{CR}$ )



*Supplementary Figure B.14: Plot of the domain over which the original and the optimal functioning ORCHIDEE version are compared. Red line denotes points at which the two versions' time series were compared.*



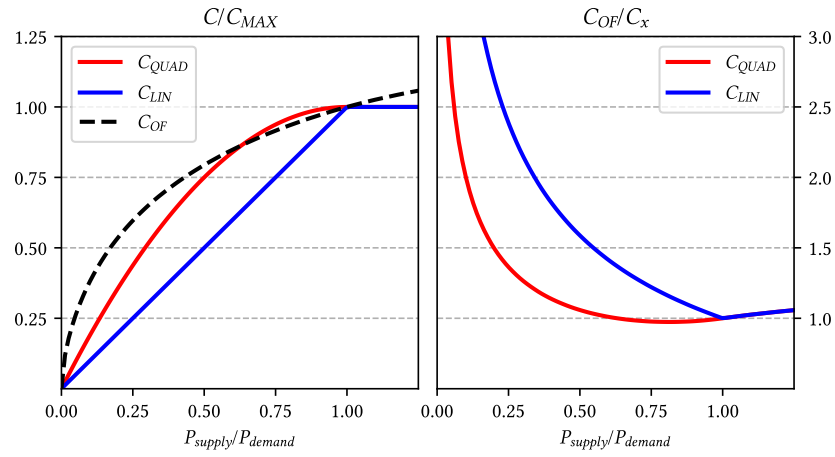
Supplementary Figure B.15: Conceptual model result with the PUE term (Eq. 4.11) modeled with total leaf P ( $P_{L,max}$ , mgP m<sup>-2</sup>) instead of leaf P concentration ( $\rho_{L,max}$ , mgP gC<sup>-1</sup>)



Supplementary Figure B.16: Maize potential yield from Mueller et al. (2012). Zonal mean values are shown with the standard deviation.

## B.4 Chapter 5 supplementary information

### B.4.1 Supplementary Figures



Supplementary Figure B.17: Yield limitation curves for a linear, quadratic and OF adjustment model as function of  $P$  supply and demand similar to Fig. 5.1.