Universität Kassel
Fachbereich Ökologische Agrarwissenschaften
Fachgebiet Agrartechnik

Combining crop growth models with the Precision Agriculture concept of yield gap analysis to evaluate yield limiting and reducing factors

By: Emir Memic

Witzenhausen 2022

Combining crop growth models with the Precision Agriculture concept of yield gap analysis to evaluate yield limiting and reducing factors

Universität Kassel
Fachbereich Ökologische Agrarwissenschaften
Fachgebiet Agrartechnik
Prof. Dr. sc. agr. Oliver Hensel

Dissertation zur Erlangung des akademischen Grades Doktor der
Agrarwissenschaften (Dr. agr.)

von
M.Sc. Emir Memic
geboren in Sarajevo, Bosnien und Herzegowina

Witzenhausen 2022

Tag der mündlichen Prüfung: 20.12.2022

Declaration

"I herewith give assurance that I completed this dissertation independently without prohibited assistance of third parties or aids other than those identified in this dissertation. All passages that are drawn from published or unpublished writings, either word-for-word or in paraphrase, have been clearly identified as such. Third parties were not involved in the drafting of the content of this dissertation; most specifically, I did not employ the assistance of a dissertation advisor. No part of this thesis has been used in another doctoral or tenure process."

Erklärung

„Hiermit versichere ich, dass ich die vorliegende Dissertation selbständig, ohne unerlaubte Hilfe Dritter angefertigt und andere als die in der Dissertation angegebenen Hilfsmittel nicht benutzt habe. Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten oder unveröffentlichten Schriften entnommen sind, habe ich als solche kenntlich gemacht. Dritte waren an der inhaltlichen Erstellung der Dissertation nicht beteiligt; insbesondere habe ich nicht die Hilfe eines kommerziellen Promotionsberaters in Anspruch genommen. Kein Teil dieser Arbeit ist in einem anderen Promotions- oder Habilitationsverfahren durch mich verwendet worden."

21.01.2023, Emir Memic

Acknowledgments

I am very grateful to Prof. Simone Graeff for constant support and Prof. Oliver Hensel and Prof. Hans W. Griepentrog. In spite all the time and effort I have invested in the thesis and thesis-related work, finishing the thesis would not be possible without their support and willingness to back me up when needed. I am grateful to the co-authors of my publications for finding time to support me and help me whenever they had time.

I am grateful to my parents for genuine support through this entire process.

The support of many friends and colleagues was very important the whole time. Without naming each of them, I keep them in my mind and am grateful for having them in my life. They know who I mean by this :)

Table of Contents

Parts of this thesis have already been published as follows:

1) Memic, E.; Graeff, S.; Claupein, W.; Batchelor, W. D. (2019): GIS-based spatial nitrogen management model for maize: short- and long-term marginal net return maximising nitrogen application rates. *Precision Agriculture* 20 (2), S. 295–312. DOI: 10.1007/s11119-018-9603-4.
   Chapter 4

2) Memic, E.; Graeff-Hönninger, S.; Hensel, O.; Batchelor, W.D. Extending the CSM-CERES-Beet Model to Simulate Impact of Observed Leaf Disease Damage on Sugar Beet Yield. *Agronomy* 2020, *10*, 1930. https://doi.org/10.3390/agronomy10121930 **Chapter 5**

3) Memic, E.; Graeff, S.; Boote, K. J.; Hensel, O.; Hoogenboom, G. (2021): Cultivar Coefficient Estimator for the Cropping System Model Based on Time-Series Data: A Case Study for Soybean. *Transactions of the ASABE* 64 (4), S. 1391–1402. DOI: 10.13031/trans.14432
   Chapter 6

List of Figures

List of Tables

List of Abbreviations

AI	Artificial Intelligence
Apollo	**A**pplication of **P**recisi**O**n Agricu**L**ture for Fie**L**d Management **O**ptimization
APSIM	Agricultural Production Systems sIMulator
Avg-nRMSE	Average-normalised root mean square error
C	Carbon
CERES	Crop Environment Resource Synthesis
CSM	Cropping System Models
DAP	Days after planting
DGPS	Differential Global Positioning System
DISLA	Photosynthesis reduction due to leaf disease
DM	Dry matter
DSS	Decision support system
DSSAT	Decision Support System for Agrotechnology Transfer
d-statistics	Index of agreement
FSPM	Functional-structural plant models
GENCALC	Genotype Coefficient Calculator
GIS	Geographic Information System
GLUE	Generalised likelihood uncertainty estimation
GPS	Global Positioning System
IBSNAT	International Benchmark Sites Network for Agrotechnology Transfer
K	Potassium
LAD	Leaf area damage
LAI	Leaf area index
LDPF	Leaf disease pressure index
LFWT	Leaf weight dry matter
N	Nitrogen
NPM	Nitrogen Prescription Model
nRMSE	Normalised root mean square error
P	Phosphorus
PA	Precision Agriculture

PLA	Plant leaf area
QGIS	Quantum Geographic Information System
RMSE	Root mean square error
RTK-GPS	Real Time Kinematics Global Positioning System
RUE	Radiation use efficiency
SENLA	Growth-related senescence
SCS	Soil Conservation Service
SOM	Soil organic matter
SRW	Storage root weight
SY	Sugar yield
TSE	Time-Series cultivar coefficient Estimator
USAID	United States Agency for International Development
WLIDOT	Daily disease damage

1   General introduction

The agricultural sector and herein the agricultural production is considered as a major backbone of every society, being related directly to the fulfilment of one of the basic human needs – food (Gillespie and van den Bold, 2017). Without reliable and sustainable food production it is not possible to imagine a "functional" society. Since the industrial revolution and beginnings of the technological advancement used in agricultural production, societies were changing and evolving accordingly. First world industrial countries development and better living conditions for members of those societies was greatly dependent on early investment in conventional farming machinery (Awokuse and Xie, 2015). Increase in the agricultural production productivity due to major investments in improving conventional farming equipment resulted in greater product output with less labour force required (Binswanger, 1986; Pingali, 2007). Technological advancement in general relies greatly on enormous investment, which is not always available in the agricultural sector. Only after specific technologies are present in industry long enough, thus the initial costs are low, an overflow of these technologies into the agricultural sector has been observed throughout history (Lowenberg-DeBoer, 1999). The reason for this is that agricultural production depends on relatively low profit margins with high production risks. Towards the end of the previous century and beginning of this, high investment in industrial technology resulted in higher commercialisation of specific technologies resulting in lower initial investment costs that found its way into the agricultural sector and agricultural production (Lowenberg-DeBoer, 1999). Nevertheless, for already highly efficient agricultural production systems in developed countries one of the main goals, next to the technological advancement of agricultural production methods and means, is to minimize yield gaps and environmental degradation. The projected population increase and requirement of additional agricultural area for crop cultivation and renewable energy production, the competition for additional land is forcing agricultural producers to invest in higher productivity per land area. In addition, because of projected climate change effects, the importance of a more sustainable and less volatile rainfed agricultural production has to be considered. The productivity per land area is expected to be affected on already established and relatively productive agricultural production areas in developed countries.

In agricultural production potential yield can be defined as yield of a specific variety cultivated under non-limiting and non-reducing conditions. Simplified, yield limiting factors include factors such as water and nutrients (abiotic) and reducing factors such as pests and diseases (biotic). There are other factors influencing potential yield which are not "manageable" for the purpose of attaining higher yield on a large-scale production, such as solar radiation and temperatures. The factors: nutrients and plant disease, which were selected for this study, are manageable to a certain extent by means of nutrient application rates or disease protective measures, that farmers can devise based on crop needs and economic feasibilities. Other factors influencing potential yield and consequently yield gap such as selection of suitable variety for specific region (based on required growing degree days and photoperiod), planting date, crop maturity and harvest date fall into crop management practices, which are manageable and can have a major influence on potential yield and yield gap.

The yield gap, depending on the definition of yield, can be analysed from different perspectives such as genetic potential (breeding), environmental factors (different growing environments), management practices affected by various market and socio-political

imperatives (Guilpart et al., 2017; Snyder et al., 2017) etc. The definition of the yield gap, in the context of the yield potential or actual yield in a farmers' field and that attained in experimental research stations can cause a confusion in yield gap analytical methods (Lobell et al., 2009), making it difficult to address the problem in the most adequate way. Throughout history yield gaps in general were analysed predominantly in the context of the food security of the rising population (Guilpart et al., 2017) as important broad scale parameter indicating social welfare. In technical terms based on the production means and resources within specific cropping systems (individual crops on farm level) yield gaps can be considered as narrow scale parameter. The term global (broad – region, state etc.) yield gap is used for describing or evaluating a multi-crop-based production output as a form of socio-economic balance that integrates crop-specific yield gap from one targeted cropping system (individual crop) on farm production level (narrow – farm level crop specific economic balance).

Cropland management, sustainable food production and supply chains as an element of social welfare gained entirely new dimensions in the second half of the previous century, due to raising awareness about complexity of the social, food production and climate interactions, backed up with huge amounts of information collected. The need for decision support systems was recognised and gave a rise to various cropland management decision support platforms (Ewert et al., 2015; Jones et al., 2017; Muller and Martre, 2019). Due to the complexity of the task and challenges that came about, early decision support in form of crop models was very specific (Ewert et al., 2015; Jones et al., 2017). They were mostly targeting specific tasks such as better understanding of crop physiology based on the observed in-season dynamics and interactions of the plant with specific environmental factors for specific fields with respect to specific management practices (Ewert et al., 2015). With the emergence of various "specialised" (specific crop under specific management practice) crop modelling solutions, new opportunities of linking bio-physiological development of the plant based on in-season dynamics to the economic viability of specific crops with regard to the specific crop management practices were created. The combination of the bio-physiological and economic dimension of crop management backed up with advancement of information technologies (higher computation power of used devices) paved the way to more complex decision support platforms capable of integrating a full spectrum of the factors affecting crop management. Later on, crop models were integrated into computational platforms usable for analysing global crop management trends with respect to future climate changes (Muller and Martre, 2019). Since field and greenhouse experiments are costly and time consuming, crop growth modelling environments offer unique chances to investigate crop growth and yield constraining factors in order to get a better insight into the factors contributing the most to the size of the yield gap from "crop" perspective. Since agricultural production is an economic entrepreneur biophysical yield maximisation of crops has to be further extended with economic analysis. All available and used resources in agricultural production: water, soil fertility maintenance, soil degradation reduction, management practices etc. are genuinely even more difficult to address on the level of smallholder farms where investment is limited and cannot rely on expensive field trials or advanced technology. Decision support tools and qualitative and quantative analysis tools could provide a form of "cheaper" technical support to enhance the overall resource use efficiency. Various technological and manual methods providing huge amounts of data about crop growth and efficiency of the agricultural resources used have to be interpreted in a meaningful way. Information about in-season biomass accumulation, soil texture and water, and indirectly N uptake and spread of disease in the field can enable more precise (temporally and spatially) crop management within decision

support systems capable integrating all of this information. Due to complex interactions of physiological, physical and economic factors involved, the need for decision support systems enabling informed management is greater than ever. Optimal N management (without over- or under-applying) in agriculture has a direct effect on farm level revenues, soil degradation and environmental pollution. Already existing decision support systems even though advanced still need to be improved. The studies conducted in this thesis aimed in investigating specific aspects of yield limiting and reducing factors with emphasis of potential improvement and disclosure of limitations.

There are still very important gaps in informed management of the fertiliser and chemicals in agricultural production under irrigated and rainfed conditions with respect to the economics of farm level crop management. Within the studies conducted in this thesis, analysis priority was given to the cropping system-based yield gap (one crop analysis) affected on a site-specific scale by spatial and temporal variabilities, biotic and abiotic stresses within the field of specific cropping systems. It has to be pointed out that a site-specific yield gap within designated cropping systems may have positive or negative effects on the broad yield gap in the context of multi-crop based cropping system interactions (Guilpart et al., 2017). The reason for ambiguity that arises when determining what is the most "reasonable" starting point for analysing the yield gap, either from cropping system (narrow scale) or global multi-crop-based yield gap (broad scale) is found in the specific "trade-off" resulting from objectives engraved in the yield maximisation concept itself. Yield maximisation, which seems to be a straight forward objective, connects three extremely "fluid" interconnected objectives: <u>agronomical (bio-physiological)</u>, <u>economic</u> and <u>socio-economic</u> yield maximisation. Some of these three underlying dimensions of the cropping system-based yield gap closure were analysed within the current thesis in the context of the optimisation of the nutrient use efficiency on a site-specific level (nitrogen) and field-specific biotic stress impact evaluation (leaf disease).

1.1   Research question – aims and objectives

Based on the existing theoretical knowledge of yield gaps derived from a socio-economic perspective of agricultural sectors across countries and from a "crop" perspective, general theoretical agreements can be found in literature as to how to solve it (Sumberg, 2012; van Oort et al., 2017). The problem with alleged yield gap closure "solutions" is that they are rarely practical solutions, but merely broad theoretical claims not derived directly from yield gap analysis (Sumberg, 2012). In the most cases the theoretical agreement in the essence cannot offer practical solutions and enable practical policies without proper quantitate analysis of causes and potential solutions. The broader the scope of yield gap analysis (socio-economic, regional or multi-cropping) the less focussed can practical solutions be (Sumberg, 2012; Van Ittersum et al., 2013). With more detailed and more focussed yield gap analysis such as one crop or crop rotations on farm level, the easier is to empirically quantify factors causing the yield gap (difference between potential and actual yield) as indicated in studies such as van Oort et al. (2017), Grassini et al. (2015), Van Rees et al. (2014). Based on a more specific analysis from a "crop" perspective, yield gap analysis can offer focussed policy and management recommendations empirically supported, with unfortunately limited applicability on the entire agricultural sector level (Sumberg, 2012).

The underlying objectives of the scientific studies conducted within this thesis was to quantify in-field heterogeneity with respect to the yield limiting and reducing factors. In the analysis of yield gaps, the priority was given first to the empirical understanding of the given yield gap

with respect to spatial and temporal variabilities based on empirically demonstrable bio-physiological factors causing this yield gap (such as site-specific soil profile characterisation). Crop models have enabled a more detailed scientific analysis of bio-physiological factors causing the yield gap with respect to spatial and temporal variabilities affecting measured yield variability in the field. Furthermore, bio-physiological net contributors were analysed economically based on a simplified marginal net return concept, in order to give a better insight into two major aspects of the crop production: bio-physiological yield constraints and economic implications when it comes to management recommendations.

With the focus on the nutrient management on commercial farm production, where commonly uniform fertiliser application rates are applied, the difference between potential yield and actual yield can be directly related to the site-specific nutrient availability and uptake conditioned by different soil properties and water and nutrient availability. If hypothetically all other non-controllable weather factors would be assumed static, then differences between potential yield and actual yield in the field would be the result of the in-field heterogeneity and as such this difference could be minimised by site-specific fertiliser management based on the site-specific soil profile requirements. Site-specific management of fertiliser is not focussed explicitly on the reduction of applied fertiliser in some of the in-field zones compared to the uniform fertiliser application. More or less, it is based on the plant requirement with respect to the soil profile characterisation and thus the indicated or calculated appropriate amount even if this leads to higher amounts than a previous uniform application rate would have indicated. In the same fashion as with site-specific management of fertiliser application rates chemicals applied for controlling pest and leaf disease in the field should be considered, without complying with the general myth that site-specific fertiliser and chemical management will inevitably lead to reduction of fertiliser and chemicals in the field.

Only after bio-physiological factors were quantified and the theoretical framework is supported by data, socio-economic imperatives could be included into the scientific analysis, before any practical policy or management recommendations could be devised. Economic and social constraints are external factors dictating policy and management recommendations, as crop production is an economic entrepreneur. Having said that, it has to be pointed out that bio-physiologically maximised yield with all yield increasing and protective measures does not paint a real picture of profitable agricultural production without revenues and costs included in production management. Only if additional revenue from additional yield (gain through yield increasing and protection measures) is covering the costs of yield increasing and protective measures based on marginal net return analysis, the yield gap closure is a meaningful entrepreneur.

The thesis is assembled of three research articles already published. All three articles are a form of independent research based on crop growth models, conducted with the overall objective to contribute to the improvement of existing agricultural practices. Crop growth models were used to target specifically applied N rates as yield limiting factor (attainable yield) and leaf disease as yield reducing factor (actual yield).

The thesis objectives were defined based on the farm "crop perspective" with respect to major factors influencing the yield gap. Potential yield can be defined based on the importance of factors contributing the most to the yield gap (hierarchical significance): crop growth defining factors (e.g. plant genetics, temperatures, solar radiation), yield limiting factors (e.g. water availability and nutrients) and yield reducing factors (e.g. diseases, pests, pollutants).

Conducted studies included field trials relying on rainfed conditions. Thus water availability was implicitly investigated through its effect on nutrient transfer and leaf disease favourable weather conditions. Nutrient (mis)management can lead to sub-optimal growing conditions and can increase the yield gap to a large extent before biotic factors (disease, pests) might further reduce yield. In this regard a nitrogen optimisation study was considered as a very important starting point of the thesis. Overall nutrient management can be considered as a form of yield "increasing" measures with respect to underlying economics. If crop growth and nutrient management can be considered to be in optimum the next major yield influencing factors are diseases, pests and pollutants. Especially for sugar beet leaf diseases play a major role in reducing yield and extractable sugar. This has a direct influence on the economic rentability (rate of return ROR) of the farm crop management. In addition, specific mathematical aspects of decision support tools (crop model) were analysed in order to demonstrate the potential and limitations of the specific methodological approaches used.

Thus, the overall objectives of the thesis addressed within the conducted three studies were:

1) Crop model-based analysis of yield gaps in maize production based on economic optimization of nitrogen as an example for a yield limiting factor.

2) Crop model-based analysis of leaf disease impact on sugar beet yield as an example for a yield reducing factor.

3) Analysis of specific mathematical abstractions used in crop models. Influence of predicted mathematical constants defining plant growth and in-season biomass accumulation rates on potential yield and yield gap.

2 State of the art

2.1 Yield gap overview

Within the overall context of this thesis, the potential crop yield and associated yield gaps were considered in the framework of crop growth simulating platforms and yield limiting and reducing factors that can be captured with these platforms. Crop production can be separated into three different conceptual frameworks: <u>potential</u>, <u>attainable</u> and <u>actual</u> yield, based on the defining factors (Figure 2-1). <u>Potential yield</u> defining factors are: CO_2, solar radiation, day length crop characteristics, physiology, phenology and canopy architecture, the traits specific for site and cultivar grown (Lövenstein et al., 1992) .The <u>attainable yield</u> refers to the attained yield based on the yield limiting factors such as water and nutrients (nitrogen, phosphorus etc.) (Lövenstein et al., 1992). The <u>actual yield</u> is defined based on the yield reducing factors such as weeds, pests, leaf diseases and various pollutants (Lövenstein et al., 1992; Sumberg, 2012) in addition to already mentioned yield limiting factors.
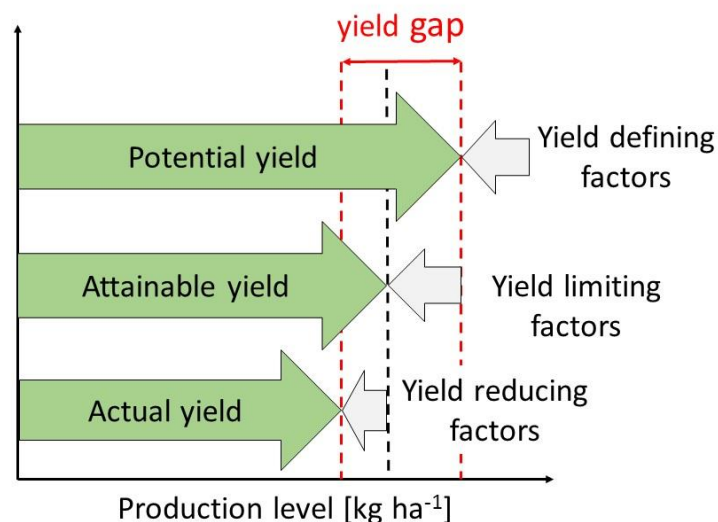
Figure 2-1 Production frameworks: potential, attainable and actual yield. Own illustration based on World Food Production: Biophysical Factors of Agricultural Production (1992) (own illustration).

The importance of separating yield situations into these three conceptual frameworks lies in the underlying means to address problems within each conceptual yield framework. Potential yield is defined by selected location and cultivar. The attainable yield is defined by yield limiting factors and <u>yield increase measures</u> that can be implemented. Finally, the actual yield is defined according to the <u>yield reducing factors</u> and <u>crop protection measures</u>. Different crop simulation models were developed with different goals and underlying structural concepts (process-oriented models, mechanistic etc.) in order to capture crop×environment×management processes with respect to climate (Ewert et al., 2015), plant development and C (carbon) and N (nitrogen) balances. Maximum attainable yield, as defined, is the highest yield at a specific location resulting from specific genetic×environment×managment interactions in agricultural production. Maximum attainable yield can be achieved in controlled environments under fully controlled conditions

within different experiments and is a great indicator of specific aspects of the agricultural production that impair its accomplishment. Large scaled fields, due to specific environmental factors such as heterogeneity of the soil and non-regular water and nutrient supply, commonly do not deliver maximum potential yield, but actual yield averages that vary from year to year.

2.2   Overview of agricultural systems

Many farmers and agricultural scientists were aware of yield variability throughout history, but lacked tools, methods and incentives to cope with these problems and factors causing them (Braga et al., 2015; Jones et al., 2017). New technologies have enabled a better insight into the scale of given site-specific variability based on observed spatial and temporal variabilities (Sishodia et al., 2020), giving rise to a "new" production approach in agriculture. The Precision Agriculture (PA) concept gained entirely new dimensions due to the industrial technological advancement that found its way into the agricultural sector (Lowenberg-DeBoer, 1999) (Sishodia et al., 2020).With support of various smart farming technologies, management of agricultural inputs on a site-specific level with respect to all occurring spatial and temporal variabilities became more feasible.

Management of site-specific inputs can be separated into three conceptual frameworks: meta-analysis, empirical models and process-based crop growth models. These three different approaches were widely used in different forms for addressing excessive use of N fertiliser for different crops in recent decades (Parent et al., 2017). The main differences among these approaches is the complexity of the underlying N-related decisions based on a different amount of data used in the process. Crop growth models are very complex and require relatively detailed information about crop genetics, environment and management practices when compared to other approaches, which made them not readily usable in practice in the past. With technological advancement of sensors used in agriculture in the context of PA deeper insight has been gained on the presence of variabilities by e.g. yield monitoring sensors or other remote-sensing technologies. Various sensors have been developed for investigating causes of yield variations capturing dynamics of plant physical transfers (water, nitrogen, etc.) (Gopala et al., 1999) and created potential for better economic and environmental management of production inputs in agriculture (Thorp et al., 2006). Economic and environmental management in agricultural production roughly consist of better management of yield limiting and reducing factors. With remote- and near-sensing technologies a better insight can be gained in physical transfers occurring during plant growth (Flowers et al., 2003) as well as in soil and atmosphere enabling farmers to manage fertiliser and chemicals more economically with regard to the environment.

The simplest way of defining spatial and temporal variability can be done with an example of fertiliser used in agricultural production, where application of the fertiliser is conducted on site-specific units that have relatively homogenous site-specific properties (spatial variability management) at the time when the plant is ready for uptake in order to accumulate biomass and protein based on the available water in the soil (temporal variability management) (Maestrini and Basso, 2018). In this case, spatial unit delineation predominantly depends on the soil-related properties, and available water in the soil with respect to timing of precipitation/irrigation (temporal variability caused by weather conditions changing over time) (Paz et al., 2001). In-field spatial and temporal variability of weeds, pests and leaf diseases are a form of yield reducing factors that are extremely difficult to quantify and delineate (Rasche and Taylor, 2017), compared to the soil properties-related variabilities. The

spatial and temporal scales of these factors are more difficult to quantify and treat even if they exhibit a patchy in-field behaviour that might be predictable to a certain extent, in the context of favourable weather conditions and plant density related micro climate changes.

The easiest and the least costly approach of measuring in-field variability in the field is through observed yield variability as a final output of entire plant in-season development. In-field yield variability observed on a single location in one season (one year) is an indicator of spatial yield variability, while yield variability over multiple seasons on the same location is an indicator of temporal variability, as it relates effects of changing weather conditions over time (Maestrini and Basso, 2018). Even though the crop growing season (one year) is subjected to specific temporal factors (that vary in time), when considered in the context of observed end-of-season yield (point-based indicator), the end-of-season yield still can be used as an important indicator of spatial variability. Based on the conceptualisation of spatial and temporal variability with point measures of (end-of-season) yield on site-specific units, where yield is consistently higher (lower) compared to the field average yield, N recommendations (less/more apply) are relatively "simple" depending on management criteria. The problem occurs in the cases when measured end-of-season yield varies from year to year "inconsistently" (on one location) preventing a more generic management approach as the causes of year-to-year fluctuations are not easy to determine and to address (Maestrini and Basso, 2018). Whether end-of-season data can be used as an indicator of spatial or temporal variability and to what extent, will be addressed later on in the thesis with examples.

Current and future trends in utilising smart farming technologies aim to establish farms based on semi- or fully-autonomous platforms relying on complimentary technologies serving the same purpose. Originally segmented efforts of using specific technologies such as Global Positioning Systems (GPS), Geographic Information Systems (GIS), and computation machines with more power for analysing larger amounts of agronomic data when unified under decision support platforms can lead to the establishment of futuristic Smart Farms. On one side automation of the vehicles used in agriculture evaluated through positioning accuracy from GPS, through Differential GPS (DGPS) to Real Time Kinematics GPS (RTK-GPS) is a great example of how beneficial technological progress can be for implementing agricultural tasks. The improvement of the positioning accuracy interpreted by the implementation of a specific agricultural production task can be summarised as the use of less accurate GPS signals for observation and data collection about crop development in the field on larger scales (yield monitoring), to more accurate vehicle guidance for reducing overlaps during the fertiliser or chemical applications, and finally to very accurate positioning where positioning information can be used for mechanical weeding (inter- and intra-row) (Griepentrog et al., 2006). Remote- (satellite-based) and near-sensing (airborne-based and sensors mounted on vehicle) images used for yield and soil mapping show great potential in providing insight in in-season crop development non-destructively. In addition to the data collection with satellite and airborne machines real-time sensors mountable on tractors and field scouting autonomous robots are also commercially available (e.g. Yara, Claas Crop Sensor Isaria, Topcon CropSpec etc.).

These sensors are measuring N status in the plant canopy with respect to chlorophyll in relative terms based on the spectral reflectance. Even though the measurements are in relative terms they are a great indicator of the spatial variability in the field. Whether sensors readings (in absolute or relative terms) that indicate in-field variability can be used in real-time for N fertilisation is open for discussion and interpretation, but when those in-field variabilities are

coupled with soil maps, historical yield maps and decision support based on bio-physiological plant knowledge, they offer a great opportunity for further N application rates optimisation on a site-specific level. Real-time sensor measurements and N application are based on one very simple deterministic approach. In many cases, the underlying assumption is: "If the plant did not take-up N (based on sensor readings), N is not available (or it is not there for various reasons)", which can result in the decision of additional N application. This can result in over-fertilisation in the cases where N is still in the soil, but for some reasons not available to the plant. In-season plant development depends on various biotic and abiotic factors that limit and reduce plant growth potential over time. Measurements of N status in the canopy in relative terms does not give enough insight into underlying dynamics of plant development with respect to genetics, management and environment. Similar deterministic approaches can be found in sensor solutions used for measuring soil texture (e.g. EM-38) (Heil and Schmidhalter, 2017) based on soil electrical conductivity. Sensors measuring electrical conductivity indeed can be used for delineating soil spatial variability, but should be complimented with ground-truth data before using. Due to the large amount of data that can be collected with various sensor technologies the need for computational mechanisms capable of making "sense" in terms of crop production management is more needed than ever. Various forms of machine learning, neural networks and artificial intelligence (AI) are being employed for analysis of sensor-based data for helping "informed" agricultural production management decisions due to the potential of "effortless" data analysis based on generic mathematical patterns used in written underlying algorithms. The AI based approaches depend on large amounts of data required for training and evaluation of the corresponding algorithms, which might be attainable non-destructively via remote sensing technologies (Jung et al., 2021). The problem with using the tremendous amount of non-destructively collected data is often a lack of ground-truth data that should accompany it. Without ground-truth data it is often not possible to make bio-physiologically informed crop management decisions.

With all available technologies and agronomic knowledge, a very important distinction has to be made when it comes to the methodological approach of implementing site-specific agricultural input management with respect to the technological solutions. Originally two main approaches were devised: sensor-based (real-time) and map-based.

The map-based approach is a form of pre-sampling where the data is subsequently processed and used in the process of informed decision making. This approach can utilise various information from different sources such as in-field sensors, destructive sample collection, historical yield data etc., aiming in generating field maps for producing field scale or site-specific maps providing the base for informed decision making in agricultural inputs management (e.g. variable N application rates). The map-based approach enables inclusion of data that is not necessarily easy to collect about soil properties without proper ground-truth sampling. Site-specific field maps are very useful when it comes to specific data that is less variable over multiple years and have a major influence on management of agricultural production inputs. One of the most important indicators affecting yield variability that is constant over time is soil texture, which has direct influence on the site-specific soil water holding capacity and N uptake throughout the season, depending on precipitation/irrigation seasonal patterns. Based on site-specific soil maps, variable N application can be controlled by applying the amount of N per site-specific unit according to the agronomic recommendations and seasonal precipitation patterns.

With technological advancement in sensors applicable for the estimation of N status in canopy based on chlorophyll, colour and light reflectance, a form of real-time sensing is possible with real-time variable N application controllers. With this approach, the sensors are measuring plant N status on-the-go with the sensor mounted on the front part of the vehicle and send signals to the N spreader at the back of the vehicle in order to adjust the N application rates based on the canopy N status (Figure 2-2).
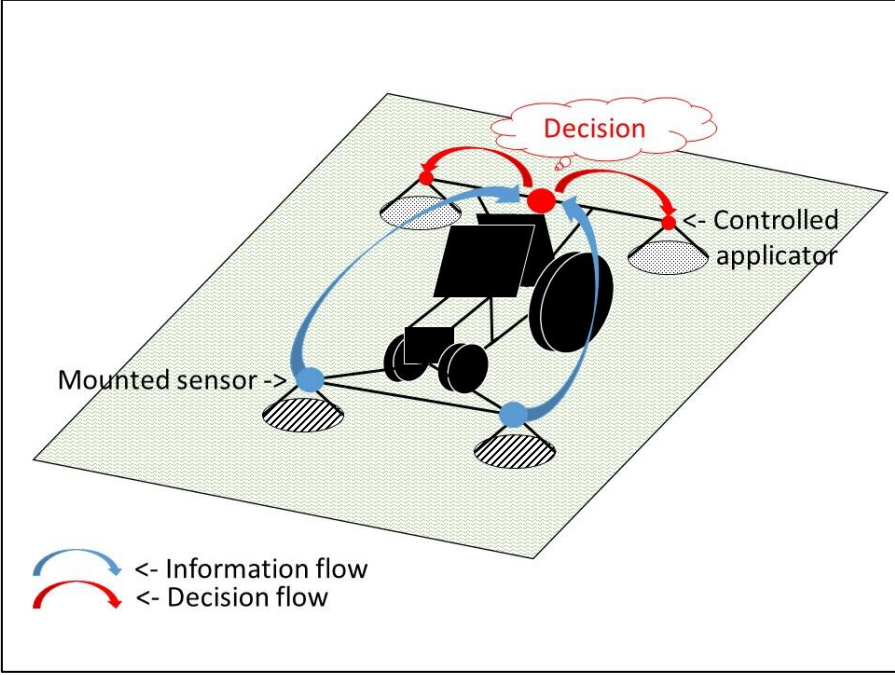
Figure 2-2 Sensor-based approach: real-time canopy sensing with application controller (own illustration).

A combination of the map- and sensor-based approaches is a methodological approach promising to utilise all benefits of both map- and sensor-based approaches, while eliminating the disadvantages (Figure 2-3). The biggest disadvantage of the map-based approach is the post-processing element and decision making that takes place subsequently. This implies the use of the post-processed data later in the season or in next years without proper ability to account for the field factors that change over time forcing it to an adjustment that is always one step behind. The most important problem of the sensor-based (real-time) approach is that specific plant growth aspects cannot be captured and lead to "informed" decisions based on one in-season measurement in real-time, because the in-season plant growth and biomass accumulation is a dynamical process resulting from various soil, weather and nutrient interactions over time. Crop growth models seem to be promising tools due to advanced computational power of new machines and their ability to integrate map-based approaches benefits into field management that can be updated and adjusted based on on-the-go sensor readings.
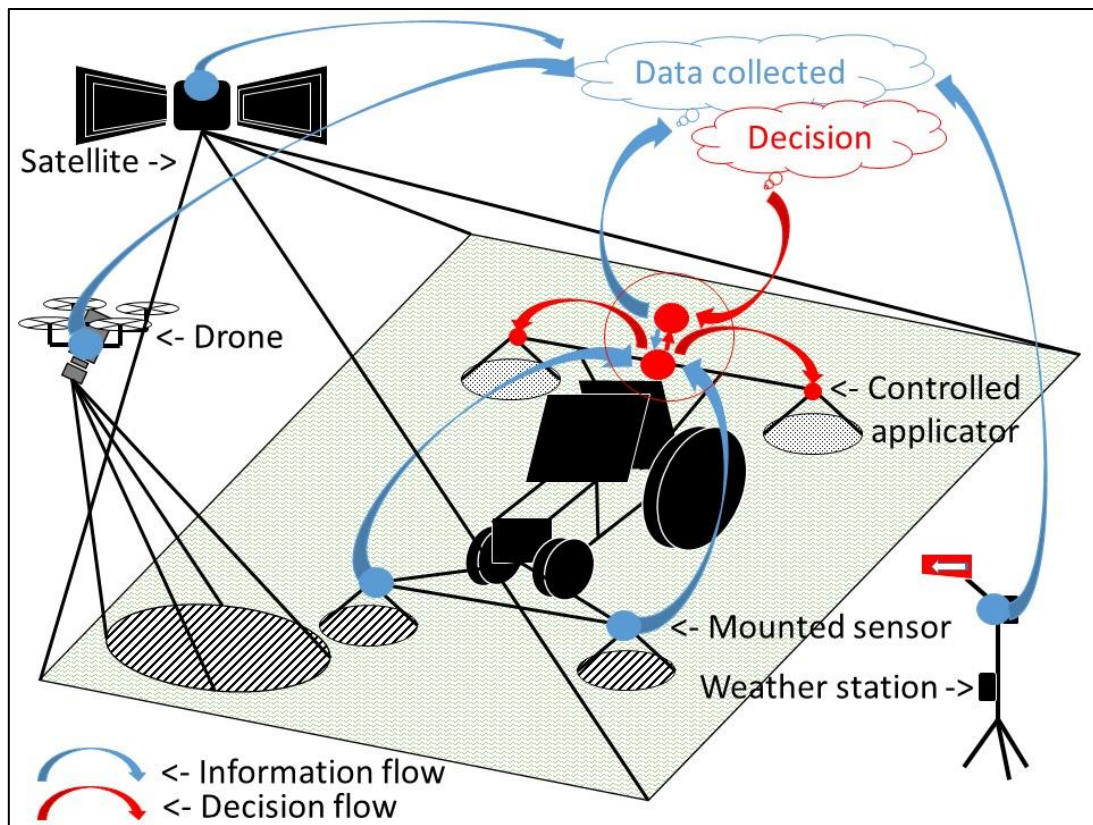
Figure 2-3 Combination of the map- and sensor-based approaches (own illustration).

Among other technological advancements originally developed for other purposes such as military, GPS and satellite-based remote sensing can only reach their true potential in agricultural production if they are used in combination with agronomic knowledge (inclusion of all biotic and abiotic factors in crop management decisions). There have been various efforts in establishing decision support platforms capable of using all sorts of data relevant for a more efficient crop production such as those based on fuzzy logic theory (Papadopoulos et al., 2011; Heiß et al., 2021). Crop growth models are one of those specialised cases that show great potential. In order to unify the large amount of sensor based collected data with plant growth dynamics and to spatially delineate management guidelines for the farmer, process-oriented crop models have been developed in the previous century (Jones et al., 2017; Muller and Martre, 2019). With further development of sensor technologies real-time management of agricultural production inputs with regard to spatial and temporal variability with accurate GPS navigation of the agricultural vehicles has become more feasible. The agricultural production management can then be based on crop growth models as real-time decision support tools for various sensors in the framework of PA. Further, crop models can be used for evaluating the trade-off between economic benefits and environmental consequences of applying various agricultural production inputs in interaction with sensor-based data. Various process-oriented crop growth models exist with underlying cropping system models for simulating in-season plant development and yield prediction based on in-season plant×environment×genetic dynamics. As mentioned previously various economic, social, technological and environmental factors played a decisive role in the rise of crop simulation models in order to enable the understanding and interactions of very complex factors (Jones et al., 2017). Underlying imperatives behind the rise of agricultural systems modelling were summarised nicely by Jones et al. (2017) with a historical timeline based on prime motivators

behind agricultural system crop modelling. Based on the brief history overview conducted by Jones et al. (2017) important lessons can be learned. With regard to specific environmental, economic and social factors interactions and long-term consequences, a "catch-up" game has been played with decision support systems capable of capturing complexity beyond individual human mind and understanding. Compared to the period 40-50 years ago todays technological advancement is more fluid with lower investment costs enabling scientists, researches and the private sector involved to shorten "catch-up" time gaps between mentioned factors supported by high machine computational power in various forms of cropping system models, machine learning, artificial intelligence etc. Based on the experience of pioneers of decision support systems from previous centuries more robust decision support systems can be achieved through open-source harmonised data sharing efforts of different market stake holders with easier access to know–how of different disciplines in order to enable inclusion that can offer benefits to all stakeholders.

At the beginnings of various crop model developing efforts, different crop simulation models were developed by different scientific communities without much interactions or feedback (Jones et al., 2017; Muller and Martre, 2019), resulting in a great variety of crop models that were interconnected or based on the same or similar theoretical framework with minor/major variations in basics, as illustrated in Muller and Martre (2019). In the last decade there have been few initiatives of creating global crop modelling networks in order to compare CSM models and enable improvement, such as the International Crop Modelling Symposium (iCROPM) (Hoogenboom et al., 2020) or the Agricultural Intercomparison and Improvement Project (AgMIP) (Rosenzweig et al., 2013; Müller et al., 2019).

All crop models can be separated into two groups: cropping system models (CSM) and functional-structural plant models (FSPM). The CSM models were mainly simulating crop development on canopy level where either the entire plant is simulated as a whole with an underlying radiation use efficiency approach (such as DSSAT-CERES) or plant development is simulated at leaf-level with its underlying energy balance mechanism (such as DSSAT-CROPGRO). On the other hand, FSPM were developed to consider the 3-dimensional morphology and architecture of the plants. Both modelling approaches have advantages and disadvantages and can be used for addressing different aspects of plant growth and development depending on the research or practical objective. The CSM takes into account the entire plant cycle with respect to the environment and genetics, while the FSPM are/were mostly used for targeting specific processes with limited scope in regard to environment and genetics (Muller and Martre, 2019). Due to the large number of research communities and lack of communication among them, a huge number of crop growth models has been published throughout the last 50-60 years (Muller and Martre, 2019). In many cases crop model development, extension to additional crops, theoretical updates in the crop model source code was not consistently updated and reported in the corresponding documentation. Muller and Martre (2019) grouped models based on the publication chronology and scientific communities such as those in Netherland, France, US and the rest.

2.3   The DSSAT crop modelling

DSSAT was developed as a form of decision-making tool for different environments for reducing the costs of conducting field experiments and for saving time in decision making. Among the variety of available crop growth models (CSM) (Müller et al., 2019; Ewert et al., 2015; Jones et al., 2017; Muller and Martre, 2019) DSSAT was selected for conducting analysis

of yield limiting and reducing factors in this thesis, because it can be used for various economically "relevant" crops and enables knowledge transfer across multiple locations and climates, while integrating local information about crop production. The DSSAT crop growth models have been used for different purposes such as for studying and optimising various agricultural production practices including fertiliser optimisation, pest management, smart farming etc. (Jones et al., 2003). Within the DSSAT-CSM sensitivity tools, yield limiting and yield reducing factors can be investigated for better understanding of the yield gap and underlying reasons for its occurrence and management scope. Estimation of yield gap in the model is based on the potential maximum yield being limited by water and/or N stress (abiotic), and reduced by pest and leaf diseases (biotic) factors (Madani et al., 2018). The origins of DSSAT go back to the early 1970s initiative of the United States Agency for International Development (USAID) to support systemic analysis of agricultural production systems for addressing food supply in developing countries. Early initiatives in analysing agricultural production systems had more focus on specific aspects of the production such as impact of soil type. These approaches led to the conclusion that focussing on one specific aspect of the production cannot lead to an overall improvement in agricultural production. Finally, this insight resulted in new initiatives that can capture various aspects of the production such as genetics, management and environment. Early work of different scientific communities was then unified under the International Benchmark Sites Network for Agrotechnology Transfer (IBSNAT), which resulted in a more generic Minimum Data Set approach with standardised inputs for the use of crop simulation models.

Crop growth models are mathematical algorithms simulating crop growth and yield as a function of the entire crop environment including physical transfers occurring in the soil, weather conditions, and crop management practices (Jones et al., 2003; Baey et al., 2014). Most of the available crop growth models are developed to simulate plant growth or yield on a field scale neglecting entirely or averaging within fields present spatial and temporal variabilities (Braga et al., 2015). The DSSAT crop models have been successfully used for simulating crop growth and yield on site-specific level (Paz et al., 2001; Batchelor et al., 2004). DSSAT was initially developed and revised in a way to integrate all information about farming practice with respect to soil conditions, weather and to enable easier transfer of plant knowledge from one location to another (Jones et al., 2003). The model requirement consists of a minimum of four different daily weather input variables (solar radiation, precipitation, minimum and maximum temperature), crop management practices (sowing date, plant population rate and fertiliser type and amount etc.), soil parameters (% of silt, clay, sand, rooting depth, etc.), and crop cultivar characteristics (genetic coefficients). The DSSAT model is the sum of various mathematical functions governing phenological traits of different crops in different conditions. In order to simulate specific crop growth, the model requires genetic information of the grown cultivar as input. Genetic information enables the users to differentiate cultivar specific phenological traits for different crops by externally adjusting specific model processes internally defined in the code. The process of identification of externally adjustable genetic coefficients based on the physiological traits of plant is defined as calibration.

One of the earliest and the most substantial attempts of using the DSSAT crop model for precision farming was the APOLLO program for site-specific soil and nitrogen optimisation in wheat and maize (Batchelor et al., 2004). Models that were originally developed to work on larger scales (hectare basis), are not necessarily able to capture sub-field (site-specific) environmental dynamics. Site-specific division of the field (site-specific units delineation) itself

can be conducted based on the economic benefits (Link et al., 2006a; Koch et al. 2004), and not solely based on the bio-physical potential of the plants. Yield maximisation based on bio-physiological full growth potential and marginal net return maximising yield are not the same and as such have to be addressed accordingly. The real potential of site-specific nitrogen applications can be implemented only if variability exists and if the given variability is not too random to the point where it would make more sense to apply fertiliser uniformly. This means that field scale variability can be delineated on site-specific field level, where site-specific units are relatively homogeneous and can accordingly be treated uniformly.

In this thesis the DSSAT crop growth models: CSM-CERES-Maize and CSM-CERES-Beet are simulating genetic×environment×management dynamic interactions on a daily basis with respect to daily weather data. Commonly done static statistical analysis are not able to capture the dynamic interactions of various factors affecting crop growth and plant stresses. Daily simulation, from planting to harvest, provide the chance for investigating dynamics behind the ability of a crop to fix CO_2 as a function of energy required over time. The correlation between solar radiation and biomass growth with respect to phenological phases of the plant based on temperature response have been in the focus of early modelling efforts. Both used crop growth models are based on the radiation use efficiency (RUE) approach according to the Beer's low (Ritchie et al., 1984; Jones and Kiniry 1986). With this simple mechanism daily plant growth is a product of intercepted solar radiation based on the RUE approach. Even though the underlying mechanism of RUE is very simplistic, it proved to be very robust and was the base for many crop growth simulation models (Ritchie et al., 1984; Jones and Kiniry 1986).

2.3.1   DSSAT water balance
It is impossible to speak about fertiliser efficiency, especially N, without considering soil water balance with respect to seasonal rainfall patterns and soil properties. The original version of the DSSAT soil water balance subroutine was developed for CERES-Wheat (Ritchie and Otter 1995). The generic version of this model was later integrated and used in all DSSAT crop models, since the DSSAT 3.5 version. Broadly speaking this one-dimensional soil water balance uses a water infiltration amount (precipitation and/or irrigation), water drainage, and unsaturated water flow, soil evaporation in combination with water taken up by roots for calculating soil water content on a daily basis with respect to soil lower limit, soil upper limit and soil saturation rate. Soil layer based daily water content is based on the tipping-bucket approach for calculating water drainage above soil upper limit. More specifically daily soil water content is calculated with the Eq. 3.

$$\Delta S = P + I - EP - ES - R - D \tag{3}$$

where ΔS is the change in soil water content based on the water infiltration amount calculated from precipitation (P) + irrigation (I) as added water to the soil profile, and transpiration (EP) + soil evaporation (ES) + surface runoff (R) + drainage from soil profile (D) as water removed from the soil profile.

2.3.2   DSSAT N-balance
The crop model nitrogen balance is simulated with the soil carbon and nitrogen balance sub-module (Soil Organic Matter - SOM). The original SOM sub-module was developed and described in Godwin and Jones (1991) and Goodwin and Singh (1998). It was integrated in the current DSSAT-Cropping System Module (CSM), with minor structural modifications. Currently it provides the template for simulating crop growth of more than 40 crops in the DSSAT-CSM. Soil N dynamics are integrated through two soil organic matter modules and an

inorganic N module. The inorganic N module is used for integrating inorganic N additions into the system for accumulating additional plant biomass with respect to the soil conditions and N transformation from one type to another. The N balance includes available N from decomposing organic matter and can be included within crop model input files in a form of residues and expected rates and amounts of decomposition with respect to environmental factors such as soil temperature, soil moisture, soil pH etc. Soil crop model N balance includes: plant N uptake, immobilisation, leaching, N gas losses, denitrification and nitrification. Simulated gas emissions are based on the DayCent model (Del Grosso et al., 2001) and include N_2O, NO and CO_2 resulting from organic matter decomposition, denitrification and nitrification. The soil organic dynamics can be calculated with two different methods in DSSAT: CERES-based module (Godwin and Singh, 1998) and Century-based module (Gijsman et al., 2002), while the Century-based module can compute the balance based on three soil organic pools. Initialisation of three separate soil organic matter pools is more complex and relies on additional input data. The N balance and within in it, the N transport through soil layers is directly dependent on the water flux values calculated in the soil water sub-module.

2.3.3 DSSAT leaf disease modelling

In the DSSAT-CSM the damage caused by pest and leaf diseases (yield reducing factors) can only be introduced as an observation on specific dates and the effects on yield can be simulated through different methods. The DSSAT-CSM does not have simulation subroutines for simulating pest or leaf disease spread rate. A different approach was recently adapted in the Agricultural Production Systems sIMulator (APSIM). APSIM as a crop modelling platform does not include direct pest and disease sub-modules, but uses an approach that enables the integration of the external pest damage simulation models (APSIM-DYMEX link) and simulates their impact on crop development indirectly (Donatelli et al., 2017).

With the CERES family models, few major pest damage types are available for simulating the effects on leaf, stem, root, seed weight, leaf area index and daily carbon assimilation, among others (Batchelor et al., 1993; Jones et al., 2003; Boote et al., 2010; Dodds et al., 2019). The information about pest damage is passed into the model through time-series file (file T) and can be entered as date specific input. Among four different types of damage one at the time can be selected in the PEST file in order to affect photosynthetically active biomass (to reduce it) through specific state variables, and by doing so affect daily carbon assimilation rates and ability of the plant to produce additional biomass. For the simulation of pest or/and leaf diseases, damage observations have to be provided as external input.

3 Materials and methods

3.1 Crop model-based analysis of yield limiting factors – N case study

Within the current thesis the yield limiting factor N was investigated in a study conducted with the DSSAT-CERES-Maize crop growth model. The study aimed for the optimisation of N usage based on short- and long-term marginal net return while maximising variable N application rates on site-specific level. For maize (*Zea mays L.*) production, N is crucial for maximising biomass accumulation. In most cases, N-based fertilisers contain ammonium (NH_4^+) and highly dissolvable nitrate (NO_3^-) that leaches into the groundwater causing pollution (Martínez-Dalmau et al., 2021). One of the potential solutions is a timely application of inorganic N at the periods when the plant is ready for uptake, depending on water availability, in the manner of site-specific N application (right amount at right time in right place). Due to political and social imperatives for plant-based renewable energy sources, backed up with the German Renewable Energy Act, cereal production, with maize among them, increased in Germany over the last two decades (Appel et al., 2016; Theuerl et al., 2019). Further, larger maize cultivation areas increased the pressure for more sustainable and efficient production. As a result, additional N demand was created leading to environmental pollution and affecting the sustainability principles requested by EU regulations.

This study aimed at developing a QGIS based software solution that can be used with a crop growth model for evaluating (recommending) site-specific variable N application rates based on a simplified marginal net return concept and observed in-field yield variability. Within this study the DSSAT-CERES-Maize based site-specific nitrogen prescription optimisation program was developed. The program was designed for estimating site-specific variable N application optimums based on marginal net return over long periods of weather data for different planting populations of maize. The developed Nitrogen Prescription Model (NPM) estimated yield and variable N levels at harvest for defined site-specific units. Based on the estimated yield and N levels required for attaining that yield, marginal net return was calculated. For the approach the modified DSSAT-CERES-Maize (v 3.7) was used with site-specific soil profiles estimated with GeoSim (Thorp et al., 2006; Thorp et al., 2008; Link et al., 2013). Two different fields were used for evaluating the NPM program: McGarvey field (US) and Riech field (Germany).

Objectives-summary:

1) Development of an open source software plug-in for evaluation of economic consequences of N recommendations for maize.
2) Evaluation of different plant population densities on maize yield with respect to site-specific soil and nutrient requirements.
3) Evaluation of the developed software plug-in with two maize datasets, from Germany and US.

3.2 Crop model-based analysis of yield reducing factors – leaf disease case study

In the context of yield reducing factors, a study was conducted within the current thesis on sugar beet (*Beta vulgaris L.*) focusing on Cercospora leaf spot disease. The *Cercospora beticolla* leaf disease pathogen causing Cercospora leaf spot disease has a major effect on harvested sugar beet yield in south-west Germany (Wolf and Verreet, 2002). Sugar beet yield reduction (harvested sugar beet dry matter) directly affects extractable sugar as shown in Shane and

Teng (1992). Other relevant diseases observed in south-west Germany are caused by *Ramularia beticola*, *Uromyces betae* and *Phoma betae* and are commonly treated with the same fungicide while treating *Cercospora beticolla*.

This study describes the integration of a leaf disease subroutine into an already existing sugar beet crop growth model. Approximately 20% of sugar produced in the world is extracted from sugar beet (Eurostat, 2020). 50% of sugar beet used for sugar production in the world is cultivated in EU (Eurostat, 2020). When considering the impact of foliar disease on crop growth, temporal and spatial in-field variability of the disease have to be accounted for. Temporal variability management depicts disease favourable weather conditions throughout the growing season, while spatial variability management depicts spatial distribution of spores in the field. In this study an existing sugar beet model was used for simulating yield losses based on the mesured Cercospora leaf spot ratings. In order to develop functional decision support systems for quantifying leaf disease impact on sugar, yield loss estimates are essential. The objective of this study was the development of a decision support tool that can use data from sensors (in future) able to quantify leaf disease damage on sugar beet canopy.

Objectives-summary:

1) Development of a crop model sub-routine for simulating impact of measured Cercospora leaf spot disease ratings on sugar beet canopy that subsequently affects harvested yield and extracted sugar by using the CSM-CERES-Beet model.
2) Evaluation of the crop model leaf disease sub-routine with three years of observed data from south-west Germany.
3) Evaluation of sugar yield and sugar yield losses due to occurring Cercospora disease indirectly through sugar beet canopy losses and its impact on the measured storage root dry matter quantities.

3.3   Crop model parameters accuracy influencing potential yield and yield gap - case study

Since crop models are a mathematical representation of plant development, specific aspects of plant growth are to a certain extent mathematical abstraction in a form of mathematical constants. Plant phenological development and above- and below-ground biomass accumulation is defined directly through crop species and cultivar coefficients (Hogenboom et al., 2011; Hogenboom et al., 2019a). In the DSSAT model two tools are available for "semi-automatic" cultivar coefficient estimation: Genotype Coefficient Calculator (GENCALC) (Hunt et al., 1993) and generalised likelihood uncertainty estimation (GLUE) (He et al. 2010). Both tools use end-of-season field observations for estimating cultivar coefficients. Most of the crop model output variables do not have constant (linear) growth rates throughout the growing period and because of that one end-of-season observation cannot provide proper insight into the in-season growth rates, based on which cultivar coefficients are estimated. Due to the importance of the in-season dynamics on simulation of the in-season biomass accumulations in the context of the used N and observed leaf disease it is very important to use cultivar coefficients representative of those in-season dynamics. In order to achieve the inclusion of the in-season plant growth dynamics and biomass partitioning among different plant organs, the use of time-series data for the cultivar coefficient estimation process was investigated within an additional study of this thesis.

In this study a newly developed external python plug-in for the DSSAT model was described and tested with a new underlying mechanism of using time-series data for estimating cultivar coefficients for all crops available in the DSSAT shell (more than 40 crops). The time-series estimator plug-in was made available in GitHub repository with user guidelines, and is already being used by few other crop modellers (https://github.com/memicemir/TSE). The Time-Series cultivar coefficient Estimator (TSE) was developed and designed to use multiple in-season observations on biomass accumulation throughout the season of multiple target variables (such as LAI, leaf weight, grain weight etc.) simultaneously. The objective was the development of a program with an intuitive user interface that enables users to automatically calibrate crop genetics.

Objectives-summary:

1) Development of an error minimisation methodology for estimation of phenology- and growth-related cultivar coefficients for the DSSAT CSM.
2) Evaluation of an error minimisation methodology including single and multiple locations and seasons based on experimental in-season observations with respect to multiple-treatment based observations for establishing more robust cultivar coefficients capable reflecting plant growth in multiple locations and seasons.

4 GIS-based spatial nitrogen management model for maize: short- and long-term marginal net return maximising nitrogen application rates

Abstract

Crop growth models including CERES-Maize and CROPGRO-Soybean have been used in the past to evaluate causes of spatial yield variability and to evaluate economic consequences of variable rate prescriptions. In this work, a nitrogen prescription program has been developed that simulates the consequences of different nitrogen prescriptions using the DSSAT crop growth models. The objective of this paper is to describe a site-specific nitrogen prescription and economic optimizer program developed for computing spatially optimum N rates over long periods of weather and plant population for maize (*Zea mays L.*) using the CERES-Maize model. The application of the model was demonstrated on a field in Germany and another one in the USA to evaluate the concept across different environmental conditions. The user can determine the short- and the long-term optimal spatial nitrogen prescription based on crop price and nitrogen cost. The program simulated short-term optimum N applications that averaged 9% (McGarvey field, USA) and 48% (Riech field, Germany) lower than the uniform rates actually applied in the fields. The program indicated different site-specific N management options for low and high yielding fields under the assumed prices for maize and N. The implementation of a site-specific plant population management was investigated. A site-specific-optimization of plant population showed a higher profitability in the heterogeneous field in Germany. Hard pan depth, hard pan factor, root distribution factor and the percentage of available soil water across the heterogeneous field were useful indicators in predicting the magnitude of site-specific plant population benefits over uniform rates.

Keywords: CERES-Maize, nitrogen management, plant population, nitrogen balance, marginal net return

4.1 Introduction

Precision agriculture is a revolutionary technology for crop production around the world. The importance of decision support tools complimenting available sensor technologies for economic and environmental risk assessment of farming practices is increasing. The DSSAT crop growth models are designed to simulate the crop growth and development processes from sowing to harvest by incorporating all information of farming practices, soil properties, crop genetics and weather (Jones et al., 2003). The DSSAT models can be used to evaluate farming practices at different locations by incorporating site-specific information on soil properties, management and weather.

Today, producers can measure spatial yields, obtain aerial images of crop biomass, gather information such as soil water content and spatial N levels, and use this information to manage

their crops precisely at a small spatial scale. However, producers have difficulties in interpreting the vast amount of data available and turning that data into production decisions.

Nitrogen (N) is critical for crop production, but over-application of N can reduce profits and cause environmental degradation. In 2014, nearly 2 Mt of N was used for agricultural production in Germany, while 12.5 Mt were used in the USA (FAO 2014). Over-application of N is common in these countries and around the world and there is a great need to reduce NO_3–N amounts left in the soil after harvest to prevent leaching into the ground water. Optimizing N timing and spatial application rate to better match crop needs can lead to a reduction in N losses to the environment.

Crop growth models have been used in the past to estimate long term optimum spatial N rates for maize (*Zea mays L.*). Paz et al. (1999) developed a technique to calibrate spatial soil properties for the CERES-Maize v3.7 model for individual grids within a field to minimize the error between simulated and measured yields over multiple years. In this approach, they divided fields into smaller grids and collected spatial yield data over several seasons for each grid. They developed crop model input files and weather files for each year and grid. Then they coupled the CERES-Maize model to an optimizer that used a simulated annealing algorithm to estimate optimum soil parameters for each grid that minimized the error between simulated and measured yield, over multiple seasons for each grid. Using this technique, they were able to explain over 80% of spatial yield variability in maize and soybean fields in Iowa based on variable soil properties. This technique was integrated into a software system called Apollo (**A**pplication of **P**recisi**O**n Agricu**L**ture for FieL**d** Management **O**ptimization) which was developed in Visual Basic for implementation on a personal computer (Batchelor et al., 2004). The Apollo software was extended to evaluate optimum N rates and plant population for each grid. The software was designed to use calibrated soil properties for each grid to simulate different combinations of N application timing and rate, and plant populations over many seasons of historical weather data in order to determine the N rate and population that maximized the long-term marginal net return for assumed values of N and seed cost and yield price (Batchelor et al., 2004). The software was written in Visual Basic for Windows XP and is no longer operational on recent Microsoft Windows platforms due to substantial changes made to Visual Basic to integrate it with the Microsoft.NET platform supported in recent releases of Windows.

Other researchers have implemented these techniques for soybean (Paz et al., 2001), wheat (Link et al., 2008) and maize (Miao et al., 2006), but required extensive training from the Apollo model developers. Using process-oriented crop growth models to evaluate spatial yields and prescriptions at small spatial scales is complex, because the crop models require numerous inputs and specialized software such as Apollo to assist with spatial model calibration and prescription development. Thus, these techniques have not been widely adopted, and a comprehensive software package does not exist to use crop growth models for precision agriculture decisions (Link et al., 2006a).

Thorp and Bronson (2013) developed an open source model optimization software package called GeoSim, which is distributed as a plug-in to the open source QGIS geographic information software (QGIS Development Team, 2009). The purpose of GeoSim is to allow users to calibrate parameters for any environmental or crop model using an optimizer based on the simulated annealing optimization technique. This software offers a modern open source replacement to the calibration procedures developed in the Apollo software. Using QGIS and

GeoSim, users can develop a map of a field, divide the field into management units, set up crop modelling input files for each grid, and calibrate soil properties to minimize the error between simulated and measured spatial yields over multiple seasons of weather and yields. GeoSim was written in Python, which is an open source language and is available as a plug-in for QGIS. It can be downloaded at http://www.qgis.org/. While GeoSim provides an excellent platform for model calibration, it does not contain software to develop and evaluate N management prescriptions.

The objectives of this work were to 1) develop a prototype open source software package to evaluate economic consequences of N management prescriptions for maize, 2) evaluate effects of different plant population rates on maize yield based on the site-specific concept, and 3) test the software for two maize datasets in Germany and the US. The long-term goal is to distribute this software as an additional plug-in to the QGIS software to be used in conjunction with GeoSim. This pair of plug-ins to QGIS will provide users with the tools needed to calibrate crop models to simulate historical spatial yield variability and to develop optimum N management prescriptions using long-term historical or future climate change weather records.

4.2   Materials and Methods

4.2.1   Field Experiments Description
McGarvey field

The 20.25 ha McGarvey field is located near Perry, Iowa, USA (41.93080°N, 94.07254°W). Spatial maize yield data were collected every even-numbered year from 1994-2002, as maize was planted with soybean in a crop rotation. The field was divided into 100 grids 0.2025 ha in size. Weather data were measured at a weather station directly at the site. In 1994 and 1996, a uniform N rate of 207 kg N ha^{-1} and 40.8 kg ha^{-1} phosphorus (P) and potassium (K) was applied just before planting. In 1998, 2000 and 2002, 224 kg N ha^{-1} and 121 kg ha^{-1} of P and K were applied each season. A more detailed description of the soil information and the related crop management practices can be found in Thorp at al. (2006).

Riech field

The Riech field is 10 ha in size and located at Ihinger Hof, Agricultural Research Station, University of Hohenheim, Germany (48.666°N, 8.967°W). Maize was planted in 2006, 2007 and 2008 following standard farmer's management practices. Weather data were taken from a local weather station at the research station. The field was divided into 80 grids (0.125 ha) for this analysis. Yield was measured each season using a yield monitor implemented on a combine harvester. Soil information was available for crop model input file development based on the publication of Link at al. (2013). Model inputs were developed for each grid and year. Initial NO$_3$-N was measured for each grid (Fig. 5), prior to sowing. The farmer's practice was to apply 160 kg N ha^{-1} as KAS (26 % N) as a uniform rate.

4.2.2   Model Development Methodology
In this project, the GeoSim Nitrogen Prescription Model (GeoSim NPM) was developed as a stand-alone Python program to simulate optimum N prescriptions for maize. The program uses optimum soil parameters calibrated using GeoSim (Thorp and Bronson, 2013) to run different combinations of N rates and application dates using a user specified number of historical (or future) years of weather data. The program generates yield and N levels in the

soil at harvest for user specified grids and weather years. The economic optimizer component of GeoSim NPM allows the user to enter the selling price for maize, the cost of N, and the cost of leaving N in the field to account for policies such as the current German compensation payment, which incentivises producers to limit N left in the field. It then computes the marginal net return (MNR) for a range of N rates for user specified historical weather years. The seasonal MNRs are then used to compute the N prescription that maximizes the long-term MNR over the user selected seasons of weather data.

Figure 4-1 shows a block diagram of the system. Grey boxes represent computational parts of GeoSim and GeoSim NPM, while white boxes represent passed or computed (simulated) parameters or additional information necessary to compute final result. The modified version of CERES-Maize (v 3.7) used in Apollo allows the optimization of up to 10 soil-related input parameters for each grid in the field, including SCS (*Soil Conservation Service*) runoff curve number, drainage rate, effective tile drainage rate, saturated hydraulic conductivity of deep impermeable layer, hard pan factor, depth of hard pan, root distribution reduction factor, N mineralization factor, soil fertility factor and adjustment of soil water availability (Thorp at al., 2008). The setup of soil-related crop model input parameters was based on the given field-specific soil properties (Thorp et al., 2006; Link et al. 2013). The model uses a self-annealing algorithm to optimize the given soil-related input parameters for each grid to the obtained yield. Site-specific soil parameter optimization resulting in a small gap between observed and simulated yield can theoretically be achieved, when more soil-related input parameters are used, but overfitting of the soil profile is unlikely to give a good fit in the final validation process (Thorp et al. 2008). Combinations of these 10 parameters are optimized by GeoSim (Fig. 1a) and passed to GeoSim NPM using a text file (Fig. 1b).

In GeoSim NPM (Fig. 1c), the user specifies the historical weather seasons to be simulated, as well as the date of N applications, range and increments within the range of N levels to simulate yield and N left in the field at the end of each season in order to compute the optimum N rate for each grid or management zone over long term seasons of weather. The user can also define different plant population densities (population rate) (Fig. 1c). Plant population density has an important role in maize growth due to the interplant competition (Tetio-Kagho and Gardner, 1988). According to Duncan (1958), plant population density increase led to individual plant yield reduction while increasing maize yield per area unit. Plant population densities that maximise yield and economic return often vary from 3 to 9 plants per square meter, due to infield site-specific variabilities (Olson and Sanders, 1988). The CERES-Maize model is then run for all combinations of N rates/plant population densities and seasons of weather data for each grid using the parameters calibrated for each grid by GeoSim and stores this information in a database for future analysis.

The GeoSim NPM software also requires the user to specify economic information including N price and value of yield in order to simulate marginal net return (MNR) for different combinations of N rates. Marginal net return is computed with Eq. 1.

$$MNR = Yield * Price - NRate * NCost + Compensation\ Payment \qquad (1)$$

where, *MNR* is the marginal net return (\$ ha^{-1}), *Yield* is simulated crop yield (kg ha^{-1}), *Price* is crop price (\$ kg^{-1}), *NRate* is the N application rate (kg N ha^{-1}), *NCost* is nitrogen cost (\$ kg^{-1}), and *Compensation Payment* is the value or penalty for leaving N in the field at harvest (\$ kg N^{-1} ha^{-1}). Once the database is computed for all combinations of N rates and seasons of weather,

GeoSim NPM searches the database to determine the N rate that maximizes the average MNR computed over all seasons.

Marginal net return for simulating different combinations of N rates and population densities simultaneously is computed with Eq. 3. Seed costs are calculated as plant population times cost per seed (Eq. 3).

$$SeedCost = Population\ rate\ (ha) * Single\ Seed\ Price \qquad (2)$$

$$MNR = Yield * Price - NRate * NCost - SeedCost + Compensation\ Payment \qquad (3)$$

Eq. 3 is used for MNR calculations only if the *Population Rate* option is activated in GeoSim NPM (Fig. 1c), in order to see the effect of different population densities on MNR. Once GeoSim NPM computes the MNR for combinations of N rate, population and years of weather, the optimum prescription can be determined for each grid by searching for the N rate and population that maximizes the average MNR over all seasons.
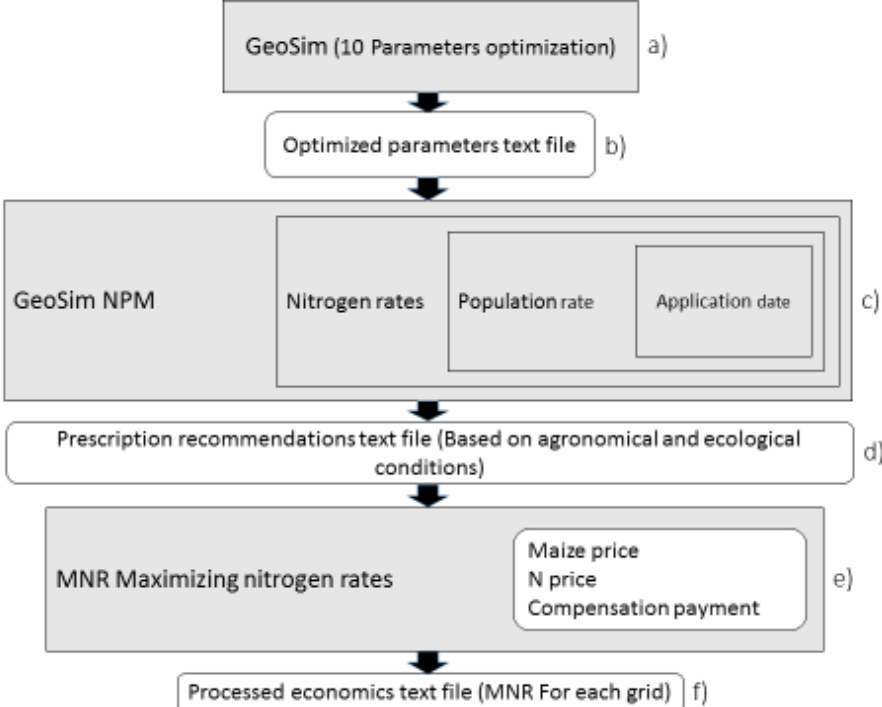
Figure 4-1 Flow diagram of the optimization and simulation process in GeoSim NPM (Memic et al., 2019).

4.3 Results

4.3.1 McGarvey field

The process of site-specific soil parameters optimization based on GeoSim indicated that the three soil parameters, effective tile drain spacing, saturated hydraulic conductivity of the lower impermeable layer and the percentage of available soil water in each soil layer were the major soil factors that described spatial yield variability. The calibration of these three soil parameters minimized the error between simulated and observed maize yields in each of the 100 grids over five seasons, while the impact on yield of all other available soil parameters in the model could be neglected. Crop rotation effects of nitrogen fixing soybean as a previous crop where accounted for in the initial conditions of the model. The calibration results of the simulated and observed maize yields for the 100 grids and five years are shown in Figure 4-2.

The R^2 between simulated and observed yield over all grids and years was 0.94, which is consistent with results reported by Thorp et al. (2006), who used the Apollo model to conduct a similar calibration for this dataset. The calibrated soil properties explained 94% of given spatial yield variability over the 100 grids and 5 seasons.
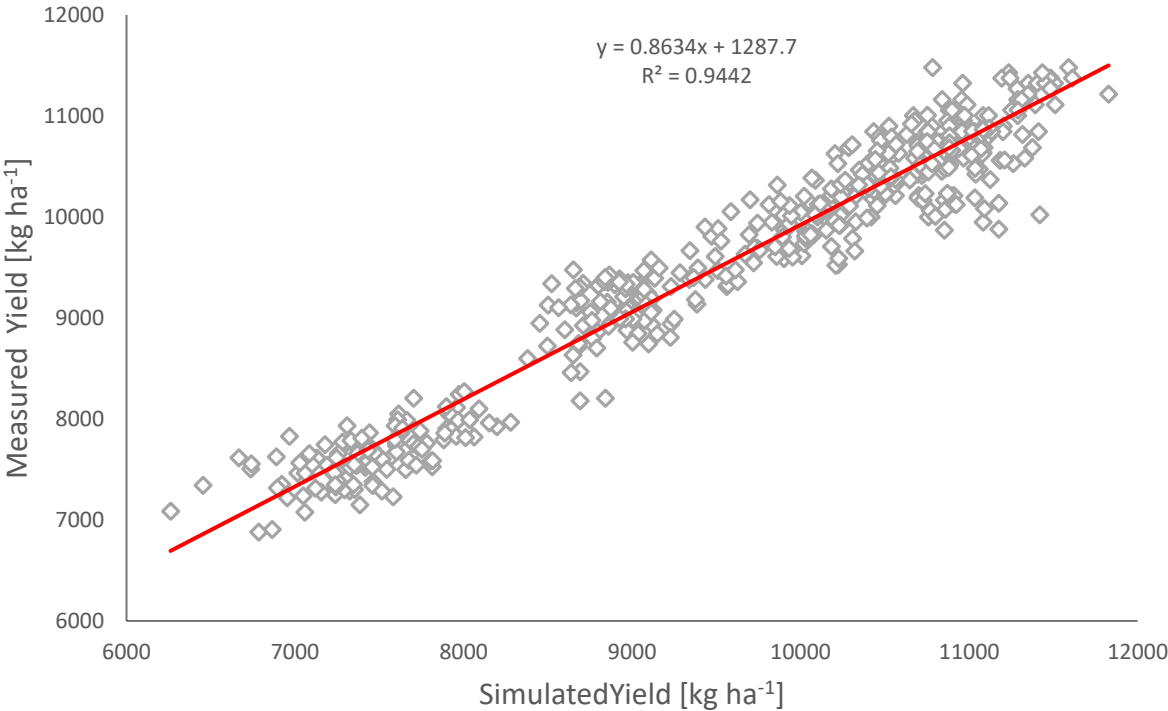
Figure 4-2 Relationship between simulated and measured maize yield [kg ha^{-1}] for the McGarvey field, Perry, Iowa using the following three soil parameters: the optimum effective tile drain spacing, the saturated hydraulic conductivity and the percentage of available soil water (n=500) (Memic et al., 2019).

GeoSim NPM was then used to compute MNR for different combinations of N rates for these five seasons (1994, 1996, 1998, 2000 and 2002) using measured weather data from the site. The price of maize and N fertiliser was assumed to be 0.13 \$ kg^{-1} and 0.5 \$ kg^{-1}, respectively. The N compensation payment was set to 0 \$ kg^{-1} ha^{-1} since there is no compensation payment for N management in the US. The GeoSim NPM simulation of N rates were defined in a range between 40 and 240 kg N ha^{-1} with an increase in increments of 10 kg N ha^{-1}. The N rate associated with the highest MNR for each grid was selected as the optimum N rate for each season. Table 4-1 shows the field level computed MNR for the producer's practice. Different grids had different simulated optimum N rates that maximized MNR for each year. Table 4-1 shows the simulated optimum N rates and MNR averaged over all grids to compare to the producer's practices at the field level. Following the optimum N rates simulated by GeoSim NPM for each season, the producer would have obtained a 5% increase in MNR and 9% reduction in the applied amount of N compared to his current practice. Table 4-1 shows averaged values of NO$_3$-N and NH$_4$-N left after harvest each year over 100 grids based on the GeoSim NPM model.

Figure 4-3 shows the simulated optimum N rate for each grid. The years 1998 and 2002 had a low variability in simulated optimum N rate, with most grids having an optimum N rate of 201-230 kg N ha^{-1}. The year 2000 had lower optimum N levels, which corresponded with lower simulated yields (Table 4-1) due to unfavourable weather conditions. Lower simulated yield potential led to lower simulated optimum N rates due to the relative differences in yield value and cost of N. However, years 1994, 1998 and 2000 had higher simulated optimum N rates, ranging from 111-230 kg N ha^{-1}.

Table 4-1 Calculated marginal net return (MNR) for measured yields and simulated MNR, measured grain yield kg ha^{-1}, simulated optimum N rate averaged over all grids for each growing season, N leached and left in the soil after harvest averaged over 100 grids for each season (Memic et al., 2019).

Year	[kg ha^{-1}] M Yield	[kg ha^{-1}] S Yield	[$ ha^{-1}] M MNR	[$ ha^{-1}] S MNR	[kg ha^{-1}] Applied N	Simulated N	Avg. N Leached	Avg. NO$_3$-N Left	Avg. NH$_4$-N Left
1994	10788	11021	1299	1331	207	202	0.2	4.1	6.3
1996	9076	8887	1076	1061	207	189	13.1	3.1	6.3
1998	9862	10172	1170	1210	224	225	5.8	3.2	6.3
2000	7598	8768	876	1064	224	152	0.4	3.1	6.3
2002	10431	10661	1244	1274	224	225	8.4	3.2	6.3
Change	4 %		5 %		-9 %				

M – Measured, S – Simulated, MNR – Marginal Net Return, N – Nitrogen, Avg. – Averaged

Table 4-2 shows the simulated N kg ha^{-1} rates grouped in representative application ranges and number of corresponding covered grids. The geospatial spread of N groups across the field for every simulated year is shown in Figure 4-3. Because of missing data in the years 1996, 1998 and 2002, no N application rates could be simulated for six grids out of 500. Grid maps were generated in QGIS and exported as images in QGIS Print Composer.

Table 4-2 Number of the grids with simulated optimum N ranges over five growing seasons (Memic et al., 2019).

Year	N Simulation Ranges [kg ha^{-1}]				Grids
	111-140	141-170	171-200	201-230	
1994	0	3	50	47	100
1996	1	23	49	25	98
1998	1	1	3	94	99
2000	20	77	3	0	100
2002	0	0	5	92	97
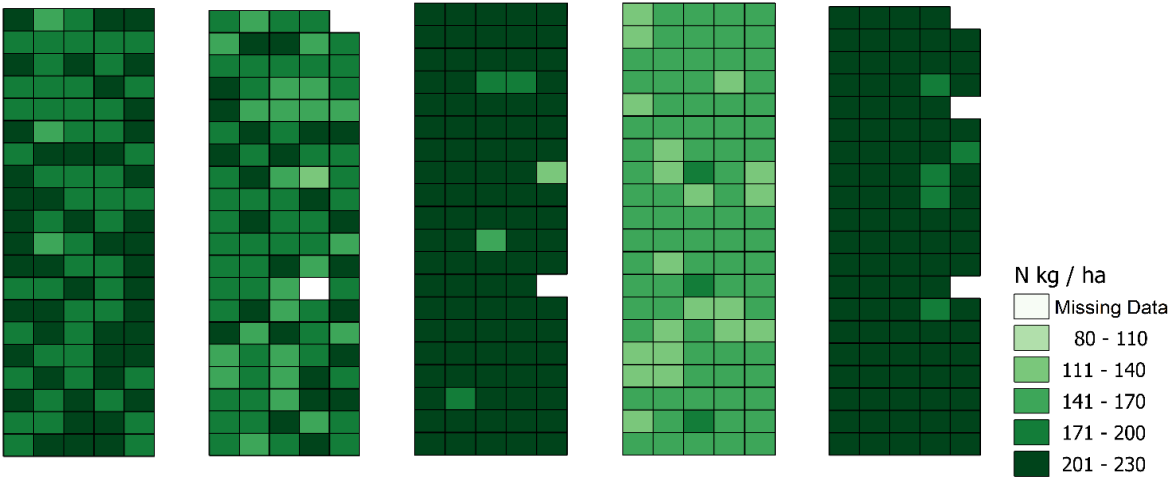
Figure 4-3 Maps of simulated N application rates that maximized MNR in each growing season (from left: 1994, 1996, 1998, 2000 and 2002) using GeoSim NPM (Memic et al., 2019).

Additional simulations were conducted to simulate both optimum N rate and plant population densities that maximized simulated MNR for each of the five seasons. Two simulation scenarios were compared: N optimisation (*N Only*) and N with population densities optimisation (*N and Pop*). The N simulation scenarios were 40-240 kg N ha^{-1} with an increase in increments of 10 kg N ha^{-1}. Population densities were simulated in a range between 3-9 plants per square meter with one plant increments. Price of a single seed was assumed to be 0.0022 $. Table 4-3 shows the results of the simulation scenarios, in which *N Only MNR* and *N and Pop MNR* included seed costs. According to the simulation results, a site-specific population density increase of 8 % per m² resulting in a plant density of 7-9 plants per m² instead of 7-8 would require an additional 2 % N kg ha^{-1} increase to approximately maintain grain yields at the same level. In retrospect, site-specific plant population densities optimisation in combination with N optimisation would result in higher MNR.

Table 4-3 Simulation scenarios of N levels (*Only N*), and combination of N rates and plant population (*N and Pop*) as conducted for each simulation scenario with corresponding MNR, N-rates, and plant population (number of the plants per square meter) (seed costs are included in MNR calculations of N simulation scenarios (*N Only*) and plant population simulation scenarios (*N and Pop*)) (Memic et al., 2019).

	Simulated [kg ha^{-1}]		[$ ha^{-1}]		Simulated [kg ha^{-1}]		Measured [m^{-2}]	Simulated [m^{-2}]
Simulation Scenarios	N Only	N and Pop	N Only	N and Pop	N Only	N and Pop	N Only	N and Pop
Year	Yield	Yield	MNR	MNR	N	N	Plant Pop	Plant Pop
1994	11021	11782	1148	1234	202	199	8.3	9
1996	8887	8881	878	887	189	193	8.3	7
1998	10172	11183	1025	1153	225	222	8.4	9
2000	8768	9630	901	967	152	173	7.4	9
2002	10661	11070	1111	1131	225	227	7.4	9
Change	6 %		6 %		2 %			8 %

N Only – Nitrogen optimisation only, N and Pop – Nitrogen and plant population optimisation combined, N – Nitrogen, Plant Pop – Plant population

Farmer's profit maximising N rates over 18 years of weather data (1966-1972, 1975-1982 and 1985-1987) were simulated with GeoSim NPM and are shown in Table 4-4. In these simulation scenarios, field-specific optimised soil parameters, farmer's practice (*Uniform N Rates*) and current prices were fixed. In GeoSim NPM, only daily temperature, precipitation and solar radiation varied on a daily basis over 18 years. The long-term N optimum showed 6% lower N application rates, averaged over five years.

The profit maximising simulated N rates in a long-term, especially over long periods of weather variability, did not convey optimum N rates for any specific grid or year. In a long-term with over 18 years of weather, simulated N rates would maximise the difference between the farmer's income and the costs for N applications with current maize and N price.

Table 4-4 Calculated MNR for simulated yield, and simulated N kg ha^{-1} compared to the applied uniform N kg ha^{-1} against 18 years of weather data (seed costs are not included in the calculations of simulated MNR in a long-term simulation scenario) (Memic et al., 2019).

Year	Simulated Yield [kg ha^{-1}]	MNR [$ ha^{-1}]	Uniform N [kg ha^{-1}]	Simulated N [kg ha^{-1}]
1994	9590	1150	207	192
1996	9892	1187	207	198
1998	10053	1202	224	209
2000	10111	1209	224	210
2002	10168	1217	224	210
Change				-6 %

MNR – Marginal Net Return, N - Nitrogen

4.3.2 Riech field

In the simulation scenario for the Riech field, four soil parameters provided the best fit between observed and simulated yield values over three seasons and were chosen as the best soil-related input parameters combination of the field-specific soil properties. The soil-related crop model input parameters indicating a major impact on yield were: hard pan depth, hard pan factor, root distribution factor and percentage of available soil water. The results of the calibrations are shown in Figure 4-4. The R^2 between simulated and measured yields across all grids and seasons was 0.75. Thus, 75 % of the spatial yield variability in the field across all grids and seasons were explained by the four soil parameters.
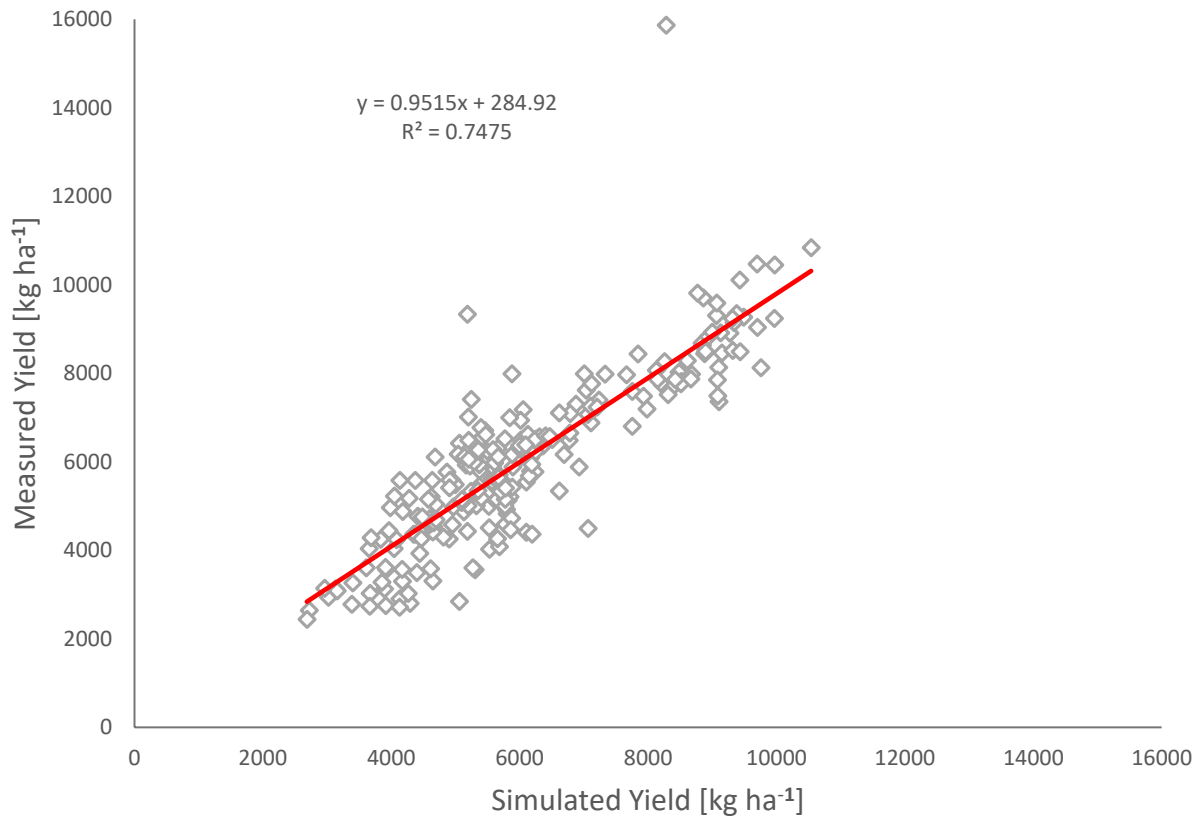
Figure 4-4 Relationship between simulated and measured maize grain yield [kg ha^{-1}] calibrated using four soil parameters for the Riech field, Ihinger Hof, Germany (n=240) (Memic et al., 2019).

GeoSim NPM was then used to compute MNR for different combinations of N rates for these three seasons (2006, 2007, and 2008). The price of maize and N fertiliser was assumed to be 0.13 \$ kg^{-1} and 0.5 \$ kg^{-1}, respectively. First, an N optimisation scenario was conducted without compensation payment (0 \$ $^{-1}$ ha^{-1}). GeoSim NPM simulation N rates were set in a range between 40 and 180 kg N ha^{-1} with an increase in increments of 10 kg N ha^{-1}. Simulation results suggested a 48 % lower N rate, resulting in 11 % higher MNR over three growing seasons compared to the farmer's actual practice (Table 4-5).

Riech field had substantially lower yields kg ha^{-1} when compared to the McGarvey field. Due to this, costs of N have a higher impact on MNR levels. Lower N rates in the simulations are also a result of the already existing high NO$_3$-N levels in the soil (Fig. 5), which were considered during the soil parameter optimisation in the GeoSim.

Table 4-5 Simulated MNR, grain yield kg ha^{-1} and simulated N kg ha^{-1} compared with calculated MNR for measured yields, measured yield grain kg ha^{-1} and applied N kg ha^{-1}, over three growing seasons (seed costs are not included in calculations of M or S MNR and compensation payment was not activated, N leached and left in the soil after harvest averaged over 80 grids for each season (Memic et al., 2019).

Year	[kg ha^{-1}]		[$ ha^{-1}]		[kg ha^{-1}]				
	M Yield	S Yield	M MNR	S MNR	Applied N	Simulated N	Avg. N Leached	Avg. NO$_3$-N Left	Avg. NH$_4$-N Left
2006	5369	6167	617	766	160	71	0.0	3.9	7.2
2007	7016	7294	832	903	160	90	0.0	3.4	6.1
2008	5715	5459	662	664	160	91	0.0	3.7	6.4
Change	5 %		11 %		- 48 %				

M – Measured yield, S – Simulated yield, MNR – Marginal Net Return, N – Nitrogen, Avg. – Averaged

Additional simulations including the compensation payment (165 $ ha^{-1}) were conducted for the Riech field. GeoSim NPM compensation payment option is based on the NO$_3$-N and NH$_4$-N left in the upper soil layer at harvest (0-0.90 m). The N compensation threshold was set to less than 45 kg of NO$_3$–N plus NH$_4$-N kg ha^{-1} left in the soil after harvest. Due to the simulated low amounts of NO$_3$–N and NH$_4$-N kg ha^{-1} left in the field at harvest, the compensation payment did not affect N rate optimisation. All simulation output values were the same as in the basic N rate optimisation scenario, as shown in Table 4-5.

Over three growing seasons, the applied N rates are grouped in representative application amounts kg ha^{-1} (Table 4-6). They are graphically shown in Figure 4-5. Orange boxes in Figure 4-5 indicate the amount of NO$_3$-N in the soil for each grid prior to sowing, based on the soil samplings. Grid maps were generated in QGIS and exported as images in QGIS Print Composer.

Table 4-6 Number of the field grids with specific N amount ranges over three growing seasons (Memic et al., 2019).

Year	N Simulation Ranges [kg ha^{-1}]							Grids
	40-70	71-85	86-100	101-115	116-130	131-145	146-160	
2006	54	16	8	2	0	0	0	80
2007	13	15	42	7	3	0	0	80
2008	19	22	20	5	9	0	5	80
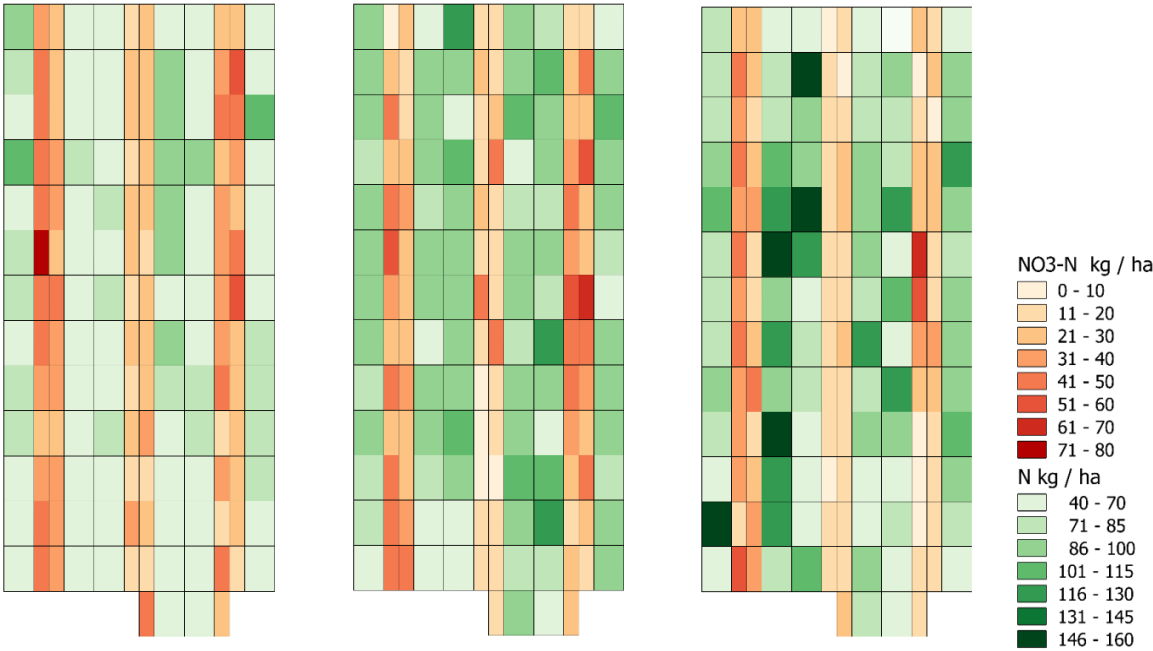
Figure 4-5 Maps of simulated N application rates that maximized MNR in each growing season (2006, 2007 and 2008) for the Riech field (green squares) using GeoSim NPM. Orange boxes indicate the NO3-N kg ha^{-1} levels before sowing in the soil as mean of the corresponding grids. Amount of NO$_3$-N kg ha^{-1} for each grid was considered in the computation of necessary N application rates and thus in the overall N balance of the vegetation period of maize (Memic et al., 2019).

Additional simulations were conducted to test the influence of population densities on MNR levels and optimum N rates. GeoSim NPM simulation N rates were set in a range between 40 and 180 kg N ha^{-1} with increments of 10 kg N ha^{-1}. Plant population densities were set in a range between 5 and 10 plants per square meter with one plant increase for each run. GeoSim NPM nitrogen and plant population optimisation included seed costs. Price of one single seed was assumed to be 0.0022 $. Over three growing seasons, simulated MNR was 3 % higher when plant population (*N and Pop*) was considered, and both nitrogen and population densities were optimised simultaneously (Table 4-7). With relatively low yield kg ha^{-1} as is the case of Riech field, profit maximisation would be achieved by reducing the amount of seeds by 15 % and N by 8 % (Table 4-7).

Table 4-7 Simulation scenarios of N levels (*Only N*), and the combination of N rates and plant population (*N and Pop*) as conducted for each simulation scenario with corresponding MNR, N-rates, plant population (number of the plants per square meter), (seed costs are included in MNR calculations of N simulation scenarios (*N Only*) and plant population simulation scenarios (*N and Pop*)) (Memic et al., 2019).

Simulation Scenarios	Simulated [kg ha^{-1}]		[$ ha^{-1}]		Simulated [kg ha^{-1}]		Measured [m^{-2}]	Simulated [m^{-2}]
	N Only	N and Pop	N Only	N and Pop	N Only	N and Pop	N Only	N and Pop
Year	Yield	Yield	MNR	MNR	N	N	Plant Pop	Plant Pop
2006	6167	6194	586	600	71	71	8.2	7
2007	7294	7348	701	712	90	86	9.2	9
2008	5459	5186	479	503	91	74	8.4	6
Change		-1 %		3 %		-8 %		-15 %

N Only – Nitrogen optimisation only, N and Pop – Nitrogen and plant population optimisation combined, Plant Pop – Plant population

Farmer's profit maximising N rates over 11 years of weather data (1992-2002) were simulated with GeoSim NPM and are shown in Table 4-8. In these simulation scenarios, field specific optimised soil parameters, farmer's practice (*Uniform N rates*), current prices and observed soil values were fixed. Long-term optimum N showed 45 % lower N rates, averaged over three years.

The profit maximising simulated N rates in a long-term did not convey optimum N rates for any specific grid or year. In a long-term with over 11 years of weather, simulated N rates would maximise the difference between the farmer's income and the costs for N applications with current maize and N price.

Table 4-8 Calculated MNR for simulated yield, and simulated N kg ha⁻¹ compared to the applied uniform N kg ha⁻¹ against 11 years of weather data (seed costs are not included in the calculations of simulated MNR in a long-term simulation scenario) (Memic et al., 2019).

Year	Simulated Yield [kg ha^{-1}]	MNR [\$ ha^{-1}]	Uniform N [kg ha^{-1}]	Simulated N [kg ha^{-1}]
2006	6107	750	160	89
2007	6251	769	160	89
2008	5487	671	160	85
Change				-45 %

MNR – Marginal Net Return, N – Nitrogen 43.4 Discussion

According to the results of the simulations, it can be concluded that low yielding fields (kg ha⁻¹) can maximize their profit with lower N amounts applied, depending on the N price. If the gap between price of maize kg⁻¹ and price of N kg⁻¹ is high, profit can be maximised with lower amounts of applied N, because yield increases in low yielding fields are not high enough to cover N costs. With GeoSim NPM, it would be possible to investigate which price gap (the grain yield price kg⁻¹ and N price kg⁻¹) and grain yield range (maximum and minimum yield kg⁻¹ ha out of all defined grids in the field) would be a good indicator for a farmer to adjust an already existing uniform rate, if not ready to switch to variable N application.

Results of the model soil properties calibration for fields in Germany and the USA explained 75% and 94% of historical spatial yield variability. This indicates that the adjustments of soil parameters accounted for a significant amount of the spatial and temporal yield variability across the field.

The lower R^2 between simulated and observed yield in the field in Germany can be explained by a higher variability within the field, wherefore a higher level of insecurity is associated with the simulated N prescriptions for the field in Germany.

For the McGarvey field, the simulated average yield across the field over five years was 9902 kg⁻¹ ha with a standard deviation of 524 kg⁻¹ ha. For the Riech field, the simulated average grain yield over three years and all grids was 6307 kg ha⁻¹ with a standard deviation of 1445 kg ha⁻¹. The higher standard deviation of the Riech field may be a result of the given higher spatial variably within the field in comparison to the McGarvey field as well as the availability of only three years of yield data in comparison to five years

The GeoSim NPM was used to compute the optimum N rates that maximized marginal net return. Results indicated that N rates could be reduced in both fields compared with current producer practices. In the McGarvey field, N rates could be 9 % lower and in Riech field N rates could be 48 % lower without profit loss.

As can be seen from the results of short (McGarvey 9 % and Riech 48 % lower N rates) and long-term (McGarvey 6 % and Riech 45 % lower N rates) N optimisation, site-specific N application has a short-term management potential, based on averaging N rates across longer periods (long-term). The profit maximising N rates over long periods of weather data could result in an over- application in low yielding years and N deficiencies in high yielding years. However, in the long-term, the farmer's profit would be maximised. In order to quantify a potential increase in uncertainty associated with predicted weather data for the rest of the growing season, a more detailed analysis of weather data is needed.

The impact of year due to changing weather conditions was obvious regarding the different grain yield amounts over a few seasons on the same field, assuming that all other inputs and practices were not changed. Additional analysis could be done to test GeoSim NPM N application timing (temporal variability) options and to see if different N application timings would have more influence on grain yield with more efficient use of N. According to the difference between short- and long-term differences lower site-specific variability in the field leads to more uniform N rate prescriptions.

It has to be noted that the results regarding plant population and N are all model-based and were not validated with field samplings or experiments. However, they can serve as a good indicator of farm cost management strategies of N and plant population site-specifically in the context of fields indicating a high spatial heterogeneity (Riech) versus fields with a lower heterogeneity (McGarvey). If the causes of the spatial variability are soil-related, they cannot be changed easily or without substantial costs involved. In that case, the farmer can consider the option of reducing existing costs involved in the production, rather than trying to further increase the yield.

The objective of the site-specific N and plant population simulations was to highlight the impact of crop production related costs on decision making in precision farming. Crop models can play a major role in trying to minimize the uncertainty associated with certain management actions as they integrate and consider multiple factors for the decision-making process. However, the results of model calibration were affected by the soil parameters used for calibration. Overall the ideal combination of soil parameters used for the calibration process seems to be determined by the underlying factors leading to spatial yield variability. Further research is needed to determine a suitable approach for the assessment of soil parameters in model calibration that captures enough information to represent spatial yield variability and temporal stability at a scale appropriate to finally optimize crop management and reduce yield gaps. The aim of using the model is of course to find a fertilization and sowing strategy better than the one used so far by the farmer and, in this way, increase the profit of the farmer. But while the model results seem to generate good profit characteristics, it has to be considered that the model responses to N application and population changes, N leaching and N left in the soil after harvest need to be validated. Based on a qualitative analysis of the results, it can be concluded that the model is doing well and could be applied in practice as a management decision support tool to achieve good profit performance of the farm. Further validation experiments would be an asset as model predictions are highly dependent on calibrated soil and yield parameters.

4.4   Conclusion

The potential of crop models as decision support for variety broad range of field scouting and sampling sensor technologies is evident. However, the collected data has to be linked with proper decision support tools to reach the full potential and gain further insights in to existing complexity and interactions between different parameters influencing crop growth and thus final management.

In this project, an open source software package has been developed that can be used in conjunction with the GeoSim open source software and QGIS to allow users to calibrate the CERES-Maize model to simulate historical spatial yield variability (GeoSim) and evaluate the economic consequences of variable rate N and plant population prescriptions (GeoSim NPM). While GeoSim NPM is currently operated as a stand-alone program, future work will focus on making this an open source plug-in for QGIS, which can be installed with the GeoSim plug-in.

5 Extending the CSM-CERES-Beet Model to Simulate Impact of Observed Leaf Disease Damage on Sugar Beet Yield

Abstract

A CSM-CERES-Beet pest damage routine was modified to simulate the impact of Cercospora leaf spot disease effects on sugar beet yield. Foliar disease effects on sugar beet growth and yield were incorporated as daily damage to leaf area and photosynthesis, which was linked to daily crop growth and biomass accumulation. An experiment was conducted in Southwest Germany (2016–2018) with different levels of disease infection. Data collected included time-series leaf area index, top weight, storage root weight and Cercospora leaf spot disease progress. The model was calibrated using statistical and visual fit for one treatment and evaluated for eight treatments over three years. Model performance of the calibration treatment for all three variables resulted in R^2 values higher than 0.82 and d-statistics higher than 0.94. Evaluation treatments for all three observation groups resulted in high R^2 and d-statistics with few exceptions mainly caused by weather extremes. Root mean square error values for calibration and evaluation treatments were satisfactory. Model statistics indicate that the approach can be used as a suitable decision support system to simulate the impact of observed Cercospora leaf spot damage on accumulated above-ground biomass and storage root yield on a plot/site-specific scale.

Keywords: Cercospora leaf spot in sugar beet; *Cercospora beticola*; CSM-CERES-Beet; decision support system

5.1 Introduction

The EU is the largest producer of sugar beet (*Beta vulgaris L.*) in the world with approximately 50% of global production (Eurostat, 2020). Approximately 20% of global sugar is produced from sugar beet (Eurostat, 2020). The sugar beet industry plays a very important role in the EU rural and agricultural economy and as such requires studies for increasing competitiveness of the sugar beet crop. Due to the abolishment of the production quotas, the EU farmers have increased sugar beet production. In 2017, cultivated area used for sugar beet production increased by 17.2% compared to 2016 (Eurostat, 2020). The EU-28 production quantities of sugar beet in 2017 was 27.3% higher than in 2016 (Eurostat, 2020). In 2018, sugar beet sown area was 1.2% lower when compared to the previous year with harvested sugar beet being 16.5% lower (Eurostat, 2020). The sudden drop in harvested amount of sugar beet was very likely caused by recorded drought conditions and not quota abolishment related price volatility (Eurostat, 2020). Based on the general economic theory with market defined prices, more volatility is to be expected in sugar beet pricing, which will affect production quantities in the EU. The increase in sugar beet production in 2017 led to a fall in prices by an average of 5.4% in real terms when compared to the prices from 2016 (Eurostat, 2020) and because of this, improved management of production resources might help to mitigate the impact on sugar beet production profitability.

Cercospora beticola (Sacc.) is a leading leaf pathogen affecting sugar beets in Germany (Wolf and Verreet, 2002). Economic consequences of Cercospora leaf spot in sugar beet are evident and quantifiable in the context of storage root yield and extractable sugar losses as reported by Shane and Teng (1992). Shane and Teng (1992) reported that sugar loss due to reductions in storage root and sugar concentration had more impact on dollar return than sugar loss to molasses due to impurities. The Cercospora leaf spot (caused by *Cercospora beticola*) has a significant influence on sugar beet yield, causing up to 30% yield losses (Wolf and Verreet, 2002). Economically less significant sugar beet leaf diseases in Germany are caused by *Ramularia beticola*, *Uromyces betae*, and *Phoma betae*, which are normally not treated with specific fungicides, as some appear later in the growing season and/or are slow to develop (Wolf and Verreet, 2002).

Crop growth models were developed as a tool to help researchers and farmers understand how genetics, environment and management impact daily crop growth and yield. They have been used to identify crop management practices (Jones et al., 2003), and the impact of climate change on yield (Olesen et al., 2011). However, most crop growth models do not simulate the impact of pest damage on crop growth and yield (Röll et al., 2019; Batchelor et al. 2020). One of the earliest efforts to simulate Cercospora leaf spot effects on yield was undertaken by Bourgeois (1989), based on the peanut crop growth (PNUTGRO) model (Boote et al., 1985). Similar efforts for evaluating disease effects on photosynthesis and yield estimates were performed by Nokes and Young (1991) and Batchelor et al. (1993).

The CERES-Beet (Crop Environment Resource Synthesis – Beet) model developed by Leviel (2000) simulates growth and development processes of sugar beet. It has been tested using data from France, Romania (Leviel et al., 2003), and North Dakota, USA (Anar et al., 2015). Anar et al. (2015) modified the CERES-Beet model of Leviel (2003) and incorporated it as the CSM-CERES-Beet (Cropping System Model) model in the Decision Support System for Agrotechnology Transfer 4.6 (DSSAT4.6). A model comparison of five sugar beet crop models showed that the CERES model provided overall good simulations of plant growth and yield, based on the evaluation criteria that consisted of: relative root mean square error, model efficiency coefficient and yield prediction error (Baey et al., 2014). The CSM-CERES-Beet model was improved and successfully tested with additional data from North Dakota, USA (2016–2018) and made available within GitHub DSSAT 4.7 release (Hoogenboom et al., 2019a; Hoogenboom et al., 2019b; Anar et al., 2019). The CERES-Beet and the CSM-CERES-Beet models are derivatives of the CERES-Maize model (Jones et al., 1986). The CERES-Maize model is deterministic and simulates different phenology events of the crop, including growth rates and biomass partitioning among crop organs (roots, stem, leaves and kernels) on a daily basis (López-Cedrón et al., 2008). The model requires a minimum of four different daily weather input variables (solar radiation, minimum temperature, maximum temperature, and precipitation), crop management practices (sowing date, plant population density and fertiliser amounts) and crop cultivar characteristics (genetic coefficients). The CSM-CERES-Beet considers sugar beet as an annual crop for beet production purposes and classifies the phenology into five events: sowing, germination, emergence, vegetative phase, and harvest. The CSM-CERES-Beet model did not include leaf disease damage.

The objectives of this study were to: (1) to develop a method to simulate the impact of observed Cercospora leaf spot disease on sugar beet yield and sugar content using the CSM-CERES-Beet model, (2) to evaluate the leaf disease model with three years of observed data from Southwest

Germany, (3) to evaluate sugar yield based on the measured storage root dry matter (DM) quantities in defined plots.

This is an extended version of the conference paper published and presented at the 12th European Precision Agriculture Conference in Montpellier, France 2019 as preliminary work under the title "Extending the CERES-Beet model to simulate leaf disease in sugar beet", Memic et al. (2019b).

5.2   Materials and Methods

5.2.1   Field Experiment Description and Data Collection

In 2016–2018, field experiments were conducted at three different fields located at Ihinger Hof, Agricultural Research Station of Hohenheim University (30 km from Stuttgart, latitude: 48.666, longitude: 8.967, elevation: approximately 490 m). Weather data (solar radiation, rainfall, minimum and maximum temperature) for model simulations were collected from a local station 200–600 m away from the fields. According to the World Reference Base (FAO, 2006), experimental soils can be characterized as vertic Luvisol and vertic Cambisol. Organic carbon was assumed to be on average 3.8%, based on historic site measurements recorded at the research station over multiple years (unpublished).

In 2017, after sugar beet had been planted, temperatures dropped below 0 °C (Figure 5-1). Sugar beet in the field managed to recover, but in the model, low temperatures had a significant influence on the simulation of the crop emergence and early growth as can be seen in the results section.
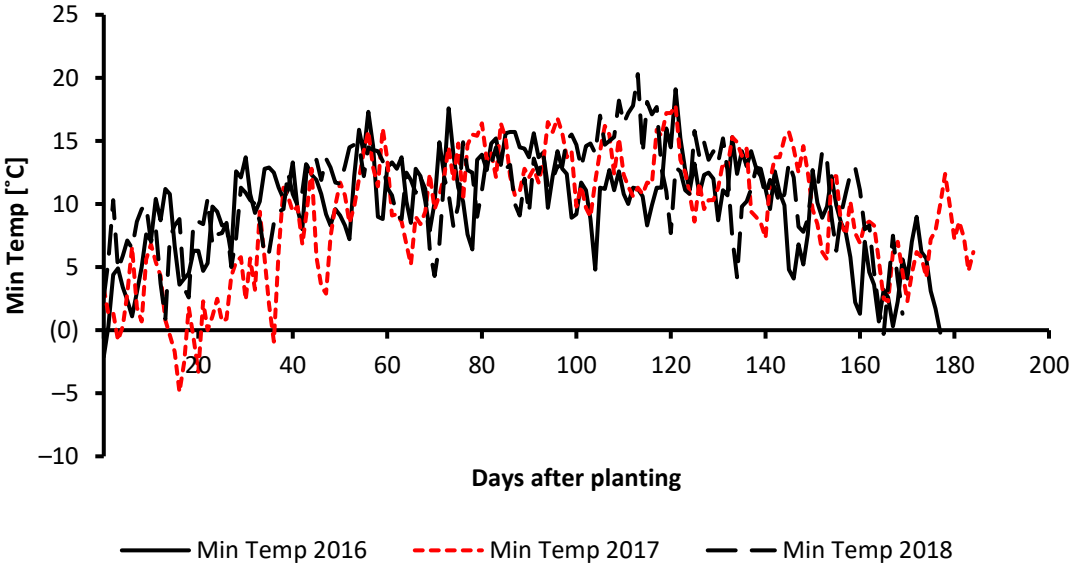
Figure 5-1 Daily minimum temperatures (Min Temp) (°C) observed at the weather station near to the sugar beet field experiment (2016–2018) (Memic et al., 2020).

In 2018, drought was recorded in the region as can be seen from seasonal cumulative rain curves shown in Figure 5-2, and the impact on simulated sugar beet growth was not entirely captured by the model.
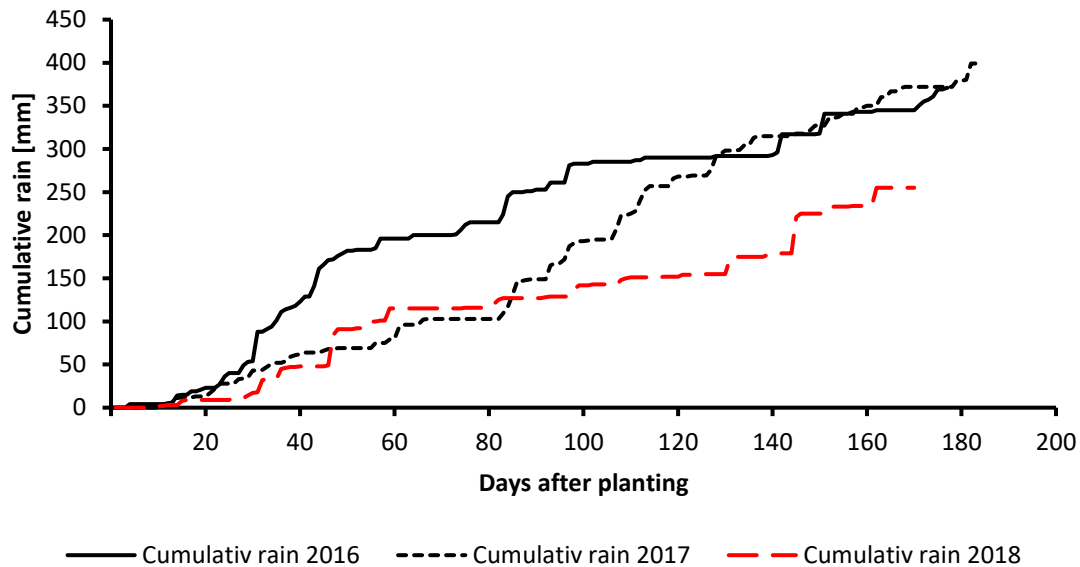
Figure 5-2 Cumulative rain (mm) observed at the weather station near to the sugar beet field experiment (2016–2018) (Memic et al., 2020).

The sugar beet cultivar BTS940 (Betaseed GmbH, 60,325 Frankfurt am Main, Germany), a Rhizomania tolerant cultivar moderately susceptible to *Cercospora beticola*, was planted in all experiments. In 2016, sugar beet was sown on 29 April 2016 (120th day of the year) and harvested 177 days after planting. In 2017, sugar beet was sown on 4 April 2017 (94th day) and harvested 184 days after planting. In 2018, sowing took place on 18 April 2018 (108th day) and sugar beet was harvested 169 days after planting. Furthermore, 107.000 seeds ha^{-1} were planted in 2 cm depth with Khun Maxima precision seed drill, with an inter-row spacing of 50 cm and intra-row spacing of 19 cm.

In 2016, five different fungicide levels, consisting of 0, 25, 50, 75 and 100% of the recommended rates, were applied. The fungicide Spyrale (Syngenta Agro, Basel, Switzerland) was applied twice (according to the application recommendation), first at the beginning of August and second three weeks later. For the 100% fungicide treatment, 1 L of Spyrale was solved in 350 L water for application per hectare. Treatments 75, 50, and 25% consisted of 0.75, 0.5, and 0.25 L, respectively. Fungicide levels 0, 50 and 100% were investigated in order to estimate level of disease leaf area within the plots (defined as plot-specific units). Respected repetitions for investigated treatments were averaged in order to get robust model evaluation values across different treatments. The plots were fully randomised with three replications in 2016. Observed values used for the simulations were the mean values of the three repetitions. Total plot size was 576 m² (24 × 24 m) in 2016. Plots were evenly divided into sampling and harvesting areas.

In 2017 and 2018, the treatment consisted of different amounts of *Cercospora beticola* inoculum per plot named as 0% inoculum (no inoculum) and 100% inoculum (with inoculum), with 4 replications in both years. Inoculum was collected at the field in 2016 and 2017. The number of *Cercospora beticola* spores was analysed in the laboratory. In both years, 1 g m^{-2} of inoculum was applied with a rate of 5.35 × 10⁶ spores g^{-1} in the 100% inoculum treatments with Massey Ferguson (90 horsepower) at 4.5 km h^{-1} with 12 m wide sprayer. The sugar beets in the plots were first wetted with 400 L ha^{-1} and then an inoculum semolina mixture was spread. In 2017, plot size was 96 m² (12 × 8 m) and in 2018, 192 m² (24 × 8 m). Emergence rate was 7.5 beets m^{-2}

in 2016, 9 beets m^{-2} in 2017 and 8 beets m^{-2} in 2018. In 2016, within a week after sowing, 130 kg N ha^{-1} was applied on the field as calcium ammonium nitrate (CAN, 27% N). In 2017, 150 kg ha^{-1} and in 2018 140 kg ha^{-1} was applied as CAN (27% N) within a week after sowing.

During the growing seasons, leaf area index (LAI) was measured every two weeks non-destructively using a LAI 2000 (LICOR Inc., Lincoln, NE, USA), by taking one reference measurement above the canopy and four measurements within the canopy. Top weight (leaves and petiole separately) and storage root weight were sampled in two to four-week intervals during the growing period (2016–2018). For each storage root, sampling date, sugar yield was measured as percent of storage root dry matter. Each sampled beet (storage root) was cut in half, where one half was used for determining sugar content and the other half for nitrogen analysis. The sugar content was analysed according to the polarisation method (ICUMSA 2003).

Leaf disease ratings were conducted for Cercospora leaf spot, after canopy closure (approximately 90% of leaves from one row were touching those in neighbouring rows), starting at the end of June in each year. Minor incidents of *Ramularia beticola* and *Pseudomonas syringae* pathogens were observed in inspected plots. The damage caused by these two pathogens was minor, when compared to *Cercospora beticola*. Every 2–3 weeks, 10 middle leaves (long-term leaves) were inspected from 10 plants in 2016 (three plot repetitions), 4 plants in 2017 (four plot repetitions) and from 5 plants in 2018 (four plot repetitions) per plot, mostly as part of destructive sampling. Based on the diseased leaf area, plot-specific leaf area disease progress (%) was recorded (Table 5-1) and used as input for the model. Model calibration was conducted on the 2016 100% fungicide treatment. The remaining data of 2016 were used in addition to the treatments of the years 2017 and 2018 for model evaluation.

Table 5-1 Observed Cercospora leaf spot, as leaf area disease progress (%), on indicated days after planting (DAP) for three different fungicide treatments: 0%, 50% and 100% in which 0, 0.5, and 1 L ha^{-1} of Spyrale fungicide was applied, respectively, and 0% and 1 g m^{-2} of inoculum was applied, respectively (Memic et al., 2020).

Year	DAP	Cercospora Leaf Spot Leaf Area Disease Progress (%)		
2016		0% fungicide	50% fungicide	100% fungicide
	63	0	0	0
	83	3	3	1
	103	13	20	20
	125	22	21	24
	138	47	25	22
	152	48	33	33
	177	48	33	33
2017		0% inoculum	100% inoculum	
	106	0	0	
	127	18	46	
	140	56	74	
	169	92	100	
	184	92	100	
2018		0% inoculum	100% inoculum	
	106	0	0	
	119	1	7	
	126	46	52	
	133	49	57	
	147	75	79	
	154	100	100	
	169	100	100	

5.2.2 Leaf Disease Damage Coupling Points

The CSM-CERES-Beet model (Leviel, 2000; Anar et al., 2019) inherited pest coupling points for simulation of many types of pest damage. Currently, there are four approaches for applying foliar damage through pest coupling points that reduce daily state variables or growth rate processes: (1) daily absolute damage rate, (2) percent observed damage (measured by comparison of different treatments), (3) daily percent damage rate, and (4) daily absolute damage rate with preference and competition (Batchelor et al., 1993). The existing pest damage module structure in the DSSAT (Figure 5-3) was used for simulating Cercospora leaf spot impact on plant growth and yield by means of "daily percent damage rate (no. 3)". This method was selected for introducing the damage caused by Cercospora leaf spot disease because it showed the required flexibility and reliability using the collected data. In combination with linear interpolation (between two in-season observations), the chosen approach provided acceptable results based on the model evaluation criteria described in the Results section.

Figure 5-3 shows the simplified modular structure with a minimal input data approach hypothetically required for evaluation of leaf disease impact on crop yield. The crop model simulates in-season crop growth and yield, which is then reduced by yield limiting factors such as leaf disease through the reduction in cumulative leaf area in the pest module (Pest.for) and damage calculated in the vegetation damage sub-module (VEGDEM.for) (Figure 5-3). In-

season above- and below-ground accumulation rates are defined in the crop model genetic coefficients input file (Figure 5-3) and through genetic coefficients and above- and below-ground biomass in-season growth ratios, a direct connection is established between leaf disease damage (in this case, Cercospora leaf spot) and storage root (yield) simulated in the model. With this approach, Cercospora leaf spot damage does not have an instant effect on reduction in the storage root in the CSM-CERES-Beet model, but rather limits further storage root growth due to reduced photosynthesis rates and the radiation use efficiency approach implemented in the model. The disease damage passed into the model earlier had a more devastating effect on yield when compared to the damage introduced later (high percentages close to harvest time).
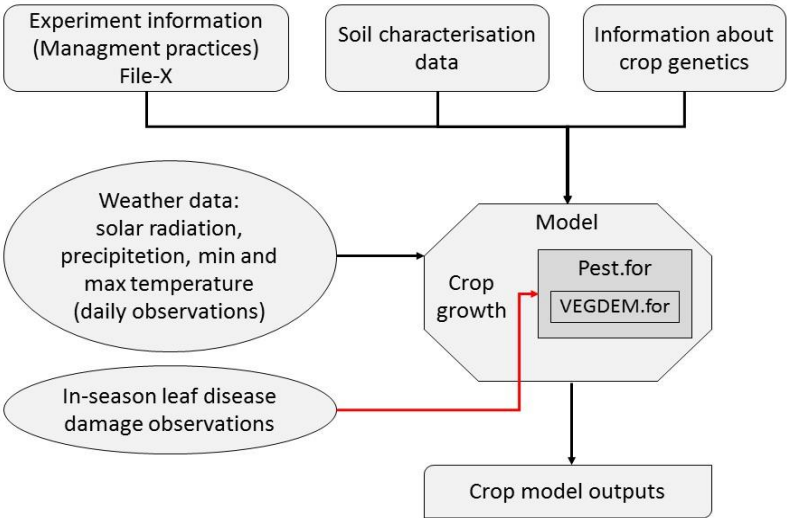
Figure 5-3 Crop model inputs: experimental file, soil characterisation file, genotype file and daily weather observations in the model with pest modular structure and vegetation damage coupling point (Memic et al., 2020).

The foliar disease Cercospora leaf spot was integrated into the model based on Equations (1) and (2), where x_{it} is the modified cumulative leaf area state variable after applying daily damage D_i on cumulative leaf area state variable X_i on day t.

$$x_{it} = X_{it} - D_{it} \tag{1}$$

The "percent of the cumulative leaf area" damage method was defined and selected in the crop model pest file, as shown in Table 5-2. Based on the selected method, Cercospora leaf spot damage is computed on a daily basis. The daily disease damage (WLIDOT) is subtracted from simulated plant leaf area (PLA), Equation (2). SENLA in Equation (2) refers to growth related senescence, and LFWT is simulated leaf weight dry matter based on the plant population (PLTPOP).

$$\text{PLA} = \text{PLA} - WLIDOT * \frac{PLA - SENLA}{LFWT * PLTPOP} \tag{2}$$

For application of leaf damage caused by *Cercospora beticola* leaf pathogen, the model variables "LAI" and "cumulative leaf area" were selected as the primary coupling points. LAI is calculated from specific leaf area. Based on the calculations in the model and the direct relationship between leaf area and LAI, applied damage will affect both cumulative leaf area and LAI, as LAI damage is subsequently deduced based on cumulative leaf area damage. Leaf

mass damage proportionally affects leaf N concentration, photosynthesis and subsequently storage root dry mater. The daily crop growth reduction rate computation is shown in Equation (3), where CARBO is daily biomass production and DISLA is photosynthesis reduction due to leaf disease.

$$\text{CARBO} = \text{CARBO} * \left[1 - \frac{DISLA}{PLA * PLTPOP} \right] \tag{3}$$

Manual estimates of Cercospora leaf spot leaf disease area progress (%) (or sensor-based leaf disease area progress (%)) can be passed into the model as percent damage. Percent damage is calculated in Equation (4):

$$D_{it} = \frac{R_{it}}{100} X_{it} \tag{4}$$

where R_{it} is the observed percent damage applied to the coupling point on day t. The CSM-CERES-Beet model simulates individual leaf development (Leviel, 2000), and leaf area is computed based on leaf number and leaf weight. Leaf disease severity is measured in the field and computed as an average per plant. The model can simulate disease effects by entering observed leaf area disease progress levels in an input file, which is used to simulate daily damage on leaf area and subsequently, light interception and daily photosynthesis rate. This method computes cumulative leaf area damage rates between two disease observation points from scouting and uses a linear interpolation (for simplicity) to convert time-series scouting observations into daily damage. Disease progress interpolation between field observation dates is computed in the model. The leaf disease damage method with damage rates is defined in the pest definition file for DSSAT sugar beet model, BSCER047.PST (Table 5-2). A more detailed description of the disease related model structure can be found in the recent publication on leaf disease damage application in Cropsim-CERES-Wheat (Röll et al., 2019; Batchelor et al., 2020) and in Batchelor et al. (Batchelor et al., 1993).

Table 5-2 Information and format of the pest definition file for the CSM-CERES-Beet (Cropping System Model – Crop Environment Resource Synthesis – Beet) model, used for defining coupling point and damage method in the crop model programming code (BSCER047.PST) (Memic et al., 2020).

No.	PID	Method Name	DM	CP	Coeff.
1	PCLA	Observed % defoliation	3	LAD	2.0

PID—leaf disease damage identifier. DM—damage characterisation method: 1–4 (3—daily percent damage rates). CP—coupling point identifier in the model. Coeff. —damage application rates. PCLA—percent cumulative leaf area.

Observed leaf disease percentages are passed into the model through the time-series file, called the T-File (treatment file), which contains observed damage on each field scouting day. The leaf area damage (LAD) method uses time series observations from the T-File to calculate percent leaf mass, leaf N and leaf area damage based on the LAD formulation in the code. The function uses a coefficient value of 2.0 to double the impact of necrotic leaf area on daily photosynthesis. The adapted damage approach in the model was based on the principle of defoliation by insects. Only physically missing leaf parts were reported in the model as observed damage affecting the overall photosynthetical activity of the plant. For Cercospora leaf spot, the diseased leaf is still present and absorbs light, but its photosynthetic activity is reduced. Jones et al. (1955) reported that removal of the foliage (by 25%) mechanically at the

eight-leaf stage did not have an appreciable effect on storage root weight and sugar content because leaves were able to re-grow and regain photosynthetic activity. Mechanically removed leaves do not block the sunlight for remaining leaves.

5.3   Results

The genetic coefficients of the model were calibrated using the 2016 100% fungicide treatment (the treatment, described in materials and methods section, in which 1 L of Spyrale was applied per hectare), based on given regulations in Germany that limited the amount of fungicide to be applied. The 100% fungicide application treatment was used to calibrate genetic coefficients, even though there were some disease incidents in this treatment. For calibration, a two-step approach was used. First, the genetic coefficients were calibrated assuming no disease was present, which helped us to understand the magnitude of each genetic coefficient required for the model to fit measured growth data. Next, the observed disease levels were incorporated into the model and the genetic coefficients were re-adjusted to attain the optimum calibration combination.

Genetic coefficients (Table 5-3) were manually adjusted to obtain the best visual and statistical fit, based on the crop model evaluation criteria used in this study, between simulated and observed values for LAI, top weight, and storage root. During the calibration process, each coefficient range was kept within minimum and maximum values recommended for the model.

Table 5-3 Sugar beet cultivar specific genetic coefficients for CSM-CERES-Beet model (BTS940) (Memic et al., 2020).

	Definition	Units	BTS940
P1	Growing Degree Days from the seedling emergence to the end of juvenile phase (juvenile group of leaves, depending on the cultivar up to 15–20 leaves)	°C-d	760.0
P2	Photo period sensitivity	hr^{-1}	0.0
P5	Thermal time from leaf growth to physiological maturity	°C-d	700.0
G2	Leaf expansion rate during leaf growth stage	cm^2 cm^{-2} d^{-1}	420.0
G3	Maximum root growth rate	gm^{-2} d^{-1}	27.5
PHINT	Phyllochron interval, the interval in thermal time between successive leaf tip appearances	°C-d	43.0

For the evaluation, two treatments from 2016 (0 and 50% fungicide treatments) and treatments from 2017 (0% and 100% inoculum level) and 2018 (0 and 100% inoculum level) were used. The calibrated model was used to simulate LAI, top weight, and storage root dry weight.

For statistical evaluation (as model performance evaluation criteria) of the simulated results, the coefficient of determination, the root mean square error (RMSE) and the d-statistics (index of agreement) were used. RMSE was used to estimate the deviation between measured (x_i) and simulated (y_i) values in the same unit (absolute measure of fit) Equation (5).

$$RMSE = \sqrt{\sum (y_i - x_i)^2 / n} \qquad (5)$$

Model performance was evaluated with the index of agreement (unitless measure), because it is more sensitive to larger deviations than smaller, due to the calculation of the difference

between simulated and observed as squared values by Equation (6), as described in Yang et al. (2014).

$$d = 1 - \sum (y_i - x_i)^2 / \sum (|y_i - \bar{x}| + |x_i - \bar{y}|)^2 \qquad (6)$$

Variation range of the index of agreement is 0.0–1.0, and values closer to 1 indicate a better fit.

5.3.1   Calibration Results

Calibration results are shown as time series graphs (Figures 5-4 and 5-5). As Cercospora leaf spot disease was introduced as damage on cumulative leaf area per plot, top and storage root weight were considered as important indicators of the overall model performance. In Figure 4, LAI (a) and top weight (b) are shown with observed values, and storage root yield in Figure 5. The model gave reasonably good estimations, based on the model evaluation criteria with high R^2 and d-stat. and relatively low RMSE of observed data, as can be seen from figures and statistics in Table 5-4. For LAI, top weight, and storage root weight (DM) R^2 were >0.82 and d-statistics > 0.94 (Table 5-4). Even though RMSE for storage root was 1696 kg ha^{-1}, it is not an indicator of bad model performance due to the existence of two large deviations from the observed time-series trend (Figure 5-5, Table 5-4). The dip in the simulation curve of top weight (Figure 5-4b) on the 140th day after planting was not caused by Cercospora leaf spot disease damage integration as can be seen from the disease-free curve ("no-dis" sugar beet growth simulated with current genetics without disease ratings being included in simulation process—(red) dotted line), but a result of a structural issues within the model.

Detailed views of the measured data and the effects of disease levels on the calibration treatment (100% fungicide in 2016) are shown in Figure 5-6. Cercospora leaf spot disease impact on sugar beet growth was demonstrated with manually measured data with corresponding impacts on top weight (primary y axis) and storage root (secondary y axis) (Figure 5-6).

(a)   (b)
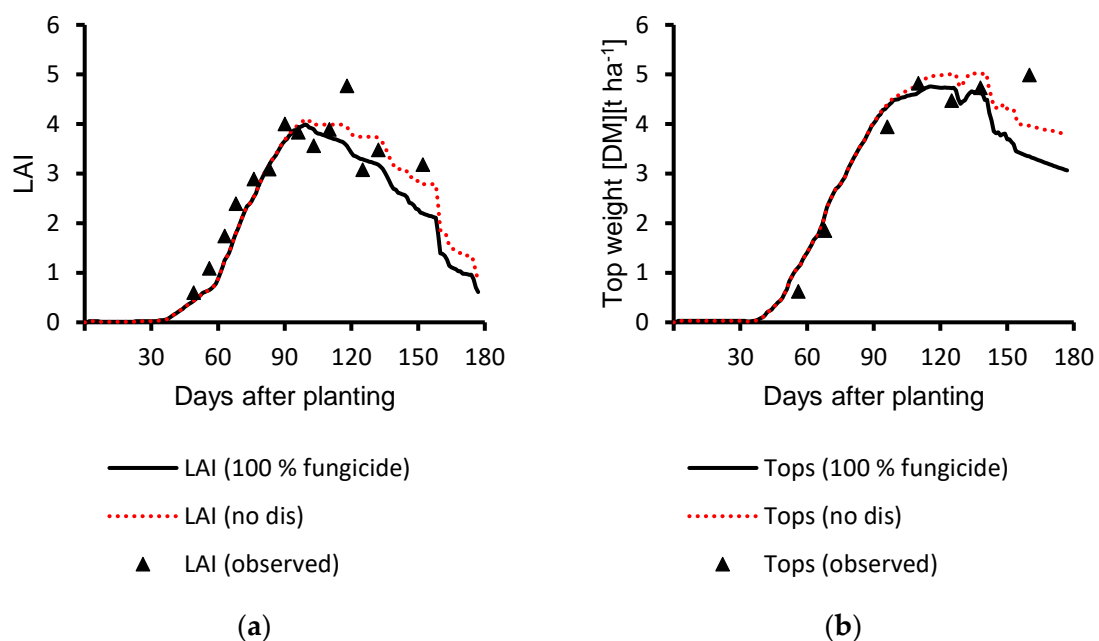
Figure 5-4 Simulated and observed values (2016: 100% fungicide) with Cercospora leaf spot ratings included in the simulation results: (a) leaf area index (LAI); (b) top weight (DM t ha^{-1}) for the calibration treatment with sugar beet growth simulated with current genetics and

without Cercospora leaf spot disease ratings being included in the simulation process as "no dis" treatment (Memic et al., 2020).
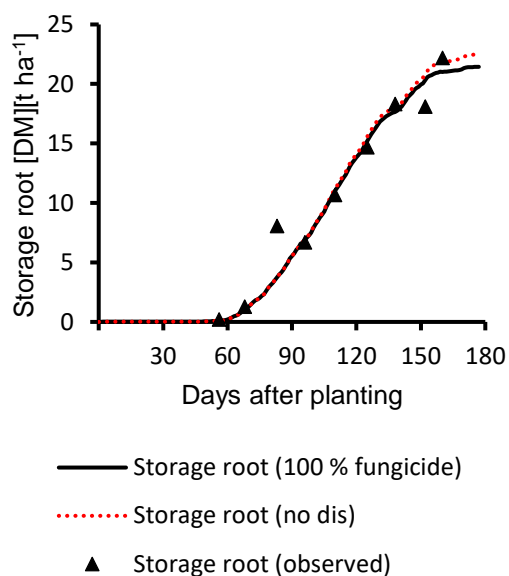
Figure 5-5 Simulated and observed values (2016: 100% fungicide) with Cercospora leaf spot ratings included in the simulation results: storage root weight (DM t ha^{-1}) for the calibration treatment with sugar beet growth simulated with current genetics and without Cercospora leaf spot disease ratings being included in the simulation process as "no dis" treatment.

Table 5-4 Detailed statistics for simulated and observed values of LAI (m^2 m^{-2}), top weight (DM) kg ha^{-1}, and storage root weight (DM) kg ha^{-1} for the calibration treatment (100% fungicide in 2016) (Memic et al., 2020).

Year	Variable	Treatment	R²	RMSE	d Stat.	Total Obs.
	LAI	100% fungicide	0.87	0.52	0.95	14
2016	Top weight	100% fungicide	0.82	686	0.94	7
	Storage root	100% fungicide	0.95	1696	0.99	9

Total Obs.—number of in-season observations used

64
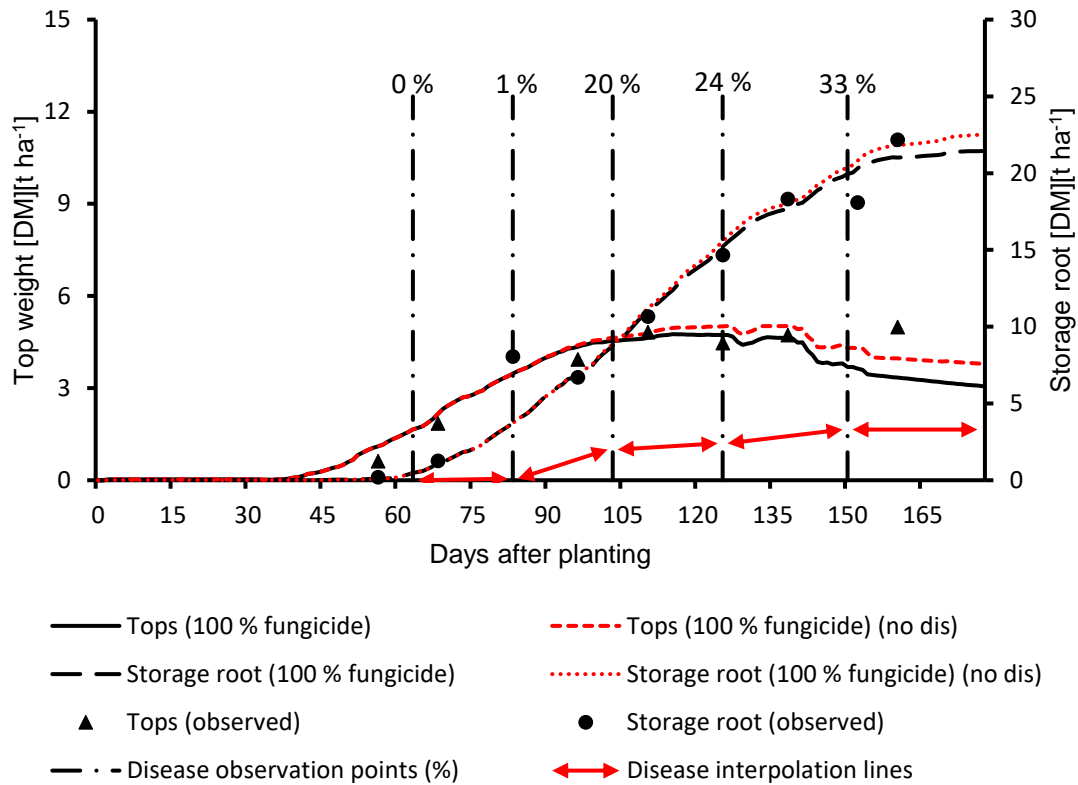
Figure 5-6 Simulated and observed values with Cercospora leaf spot ratings included in the simulation results: (y0 axis) top weight; (y1 axis) storage root weight and observed Cercospora leaf spot disease ratings on specific dates, and sugar beet growth simulated with current genetics without disease ratings being included in the simulation process as "no dis" treatment (100% fungicide in 2016) of top and storage root weight (Memic et al., 2020).

5.3.2 Evaluation Results

In 2016, the 0 and 50% fungicide treatments were available for evaluating model performance. For each evaluation treatment, 14 observations were available for LAI, seven for top weight and nine for storage root. Simulated LAI, top weight and storage root curves showed fewer fluctuations when compared among each other than observations on the same sampling dates across different treatments (Figures 5-7 and 5-8). Simulated LAI was underestimated compared to observed values (Figure 5-7a). Overall statistics (Table 5-5) and visual fit (Figure 5-7a) were adequate for rough estimates, with d-statistics >0.91. Top weight (Table 5-5, Figure 5-7b) had the same problem as in calibration treatments after 140th day (the dip in simulation curve) but had a d statistic >0.92. Visual model fit for storage root (Figure 5-8) was good. Storage root R^2 and d-statistics (Table 5-5) were very good with exception of the RMSE.

In 2017, the model simulated top weight and storage root well with the exception of the LAI. LAI was only partially satisfying (Figure 5-9a, Table 5-6) due to over-estimation of observed values with R^2 being 0.54, 0.81 and d-statistics 0.63, 0.83 for 0% and 100% inoculum treatment, respectively (Table 5-6). Top weight (Figure 5-9b) and storage root (Figure 5-10) R^2 and d-statistics were >0.96 with exception of top weight d-statistics for 0% inoculum treatment being 0.74 (Table 5-6).

—— LAI (0 % fungicide)	
– – – LAI (50 % fungicide)	
·········· LAI (no dis)	
▲ LAI (0 %) (observed)	
● LAI (50 %) (observed)	

—— Tops (0 % fungicide)	
– – – Tops (50 % fungicide)	
·········· Tops (no dis)	
▲ Tops (0 %) (observed)	
● Tops (50 %) (observed)	

(**a**) (**b**)
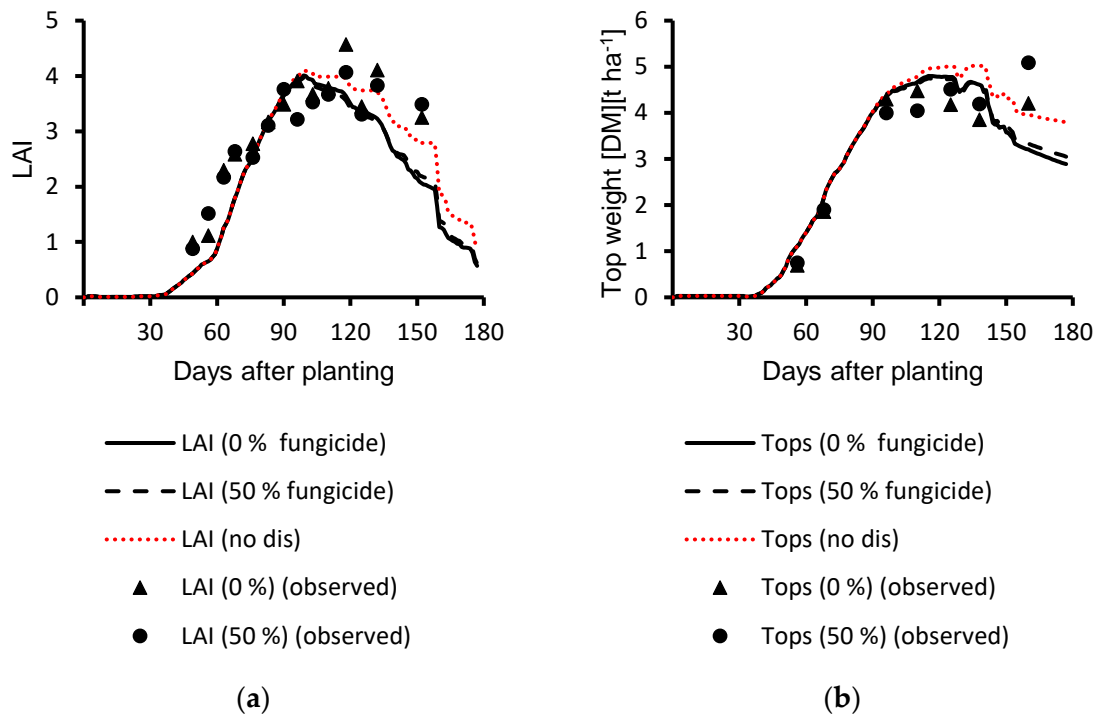
Figure 5-7 Simulated and observed values (2016) with Cercospora leaf spot ratings included in the simulation results: (a) LAI; (b) top weight (DM t ha^{-1}) for the evaluation treatments with sugar beet growth simulated with current genetics and without Cercospora leaf spot disease ratings being included in the simulation process as "no dis" treatment (Memic et al., 2020).

—— Storage root (0 % fungicide)
– – – Storage root (50 % fungicide)
·········· Storage root (no dis)
▲ Storage root (0 %) (observed)
● Storage root (50 %) (observed)
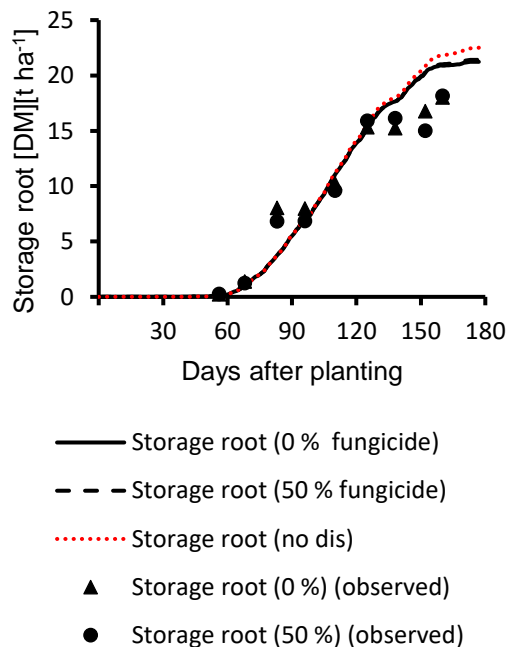
Figure 5-8 Simulated and observed values (2016): storage root weight (DM t ha^{-1}) for the evaluation treatments with (0% fungicide, 50% fungicide) and without ('no dis") Cercospora leaf spot ratings included in the simulation (Memic et al., 2020).

Table 5-5 Detailed statistics for simulated and observed values of LAI (m² m⁻²), top weight (DM) kg ha⁻¹ and storage root weight (DM) kg ha⁻¹ for the evaluation treatments 0% and 50% fungicide in 2016 (Memic et al., 2020).

Year	Variable	Treatment	R^2	RMSE (kg ha⁻¹)	d Stat.	Total Obs.
2016	LAI	0% fungicide	0.85	0.63	0.92	14
	LAI	50% fungicide	0.80	0.63	0.91	14
	Top weight	0% fungicide	0.85	565	0.95	7
	Top weight	50% fungicide	0.74	751	0.92	7
	Storage root	0% fungicide	0.94	2270	0.97	9
	Storage root	50% fungicide	0.94	2362	0.97	9

Total Obs.—number of in-season observations used

(**a**)   (**b**)
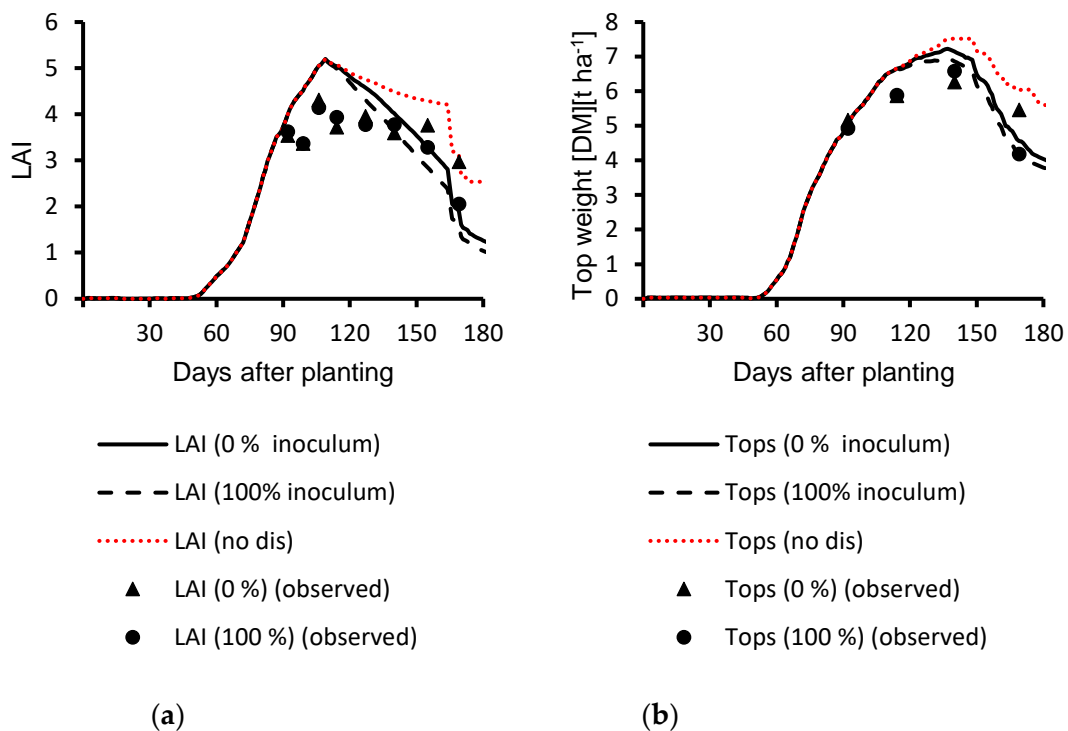
Figure 5-9 Simulated and observed values (2017): (a) LAI; (b) top weight (DM t ha⁻¹) for the evaluation treatment with (0% inoculum, 100% inoculum) and without ("no dis") simulated Cercospora leaf spot disease damage included in the simulation (Memic et al., 2020).

Table 5-6 Detailed statistics for simulated and observed values of LAI (m^2 m^{-2}), top weight (DM) kg ha^{-1}, and storage root weight (DM) kg ha^{-1} for the evaluation treatments 0% and 100% inoculum in 2017 (Memic et al., 2020).

Year	Variable	Treatment	R^2	RMSE (kg ha^{-1})	d Stat.	Total Obs.
2017	LAI	0% inoculum	0.54	0.85	0.63	8
	LAI	100% inoculum	0.81	0.72	0.83	8
	Top weight	0% inoculum	0.80	747	0.74	4
	Top weight	100% inoculum	0.96	402	0.96	4
	Storage root	0% inoculum	0.93	2399	0.98	4
	Storage root	100% inoculum	0.87	3124	0.96	4

Total Obs.—number of in-season observations used
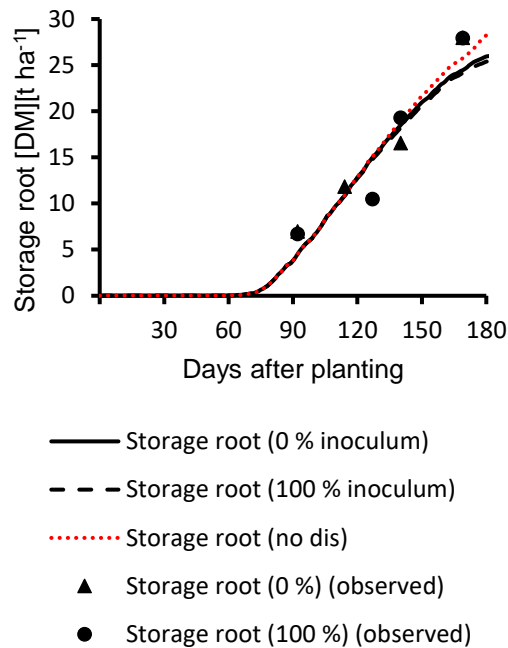
Figure 5-10 Simulated and observed values (2017): storage root weight (DM t ha^{-1}) for the evaluation treatment with (0% inoculum, 100% inoculum) and without ("no dis") simulated Cercospora leaf spot disease damage ratings included in the simulation (Memic et al., 2020).

Evaluation results for 2018 are shown in Figures 5-11 and 5-12 with corresponding statistics in Table 5-7. In 2018, a drought period occurred in the region. The model did not entirely capture drought effects causing minor over- and under-estimations of observed values. Still, the visual fit (Figures 5-11 and 5-12) indicated a satisfactory performance supported with reasonably good statistics (Table 5-7), based on the model evaluation criteria. R^2 of 0.76 (0% inoculum treatment) and 0.70 (100% inoculum treatment) were due to the over-estimation of LAI. LAI curve trend was simulated well, as can be seen in Figure 5-11a and LAI d-statistics 0.87 and 0.84 for 0% inoculum and 100% inoculum treatment, respectively. The same over-estimation of the observed values occurred for top weight dry matter (Figure 5-11b) with slightly better R^2 and d-statistics than for LAI (Table 5-7). Storage root dry matter simulation results (Figure 5-12) were partially satisfying with under-estimation close to harvest time. For both treatments, storage root dry matter R^2 and d-statistics were >0.94 (Table 5-7).

Table 5-7 Detailed statistics for simulated and observed values of LAI (m² m⁻²), top weight (DM) kg ha⁻¹, and storage root weight (DM) kg ha⁻¹ for the evaluation treatments 0% and 100% inoculum in 2018 (Memic et al., 2020).

Year	Variable	Treatment	R^2	RMSE (kg ha⁻¹)	d Stat.	Total Obs.
2018	LAI	0% inoculum	0.76	0.57	0.87	9
	LAI	100% inoculum	0.70	0.64	0.84	9
	Top weight	0% inoculum	0.92	812	0.78	6
	Top weight	100% inoculum	0.81	953	0.71	6
	Storage root	0% inoculum	0.97	3045	0.94	6
	Storage root	100% inoculum	0.99	3486	0.93	6

Total Obs.—number of in-season observations used
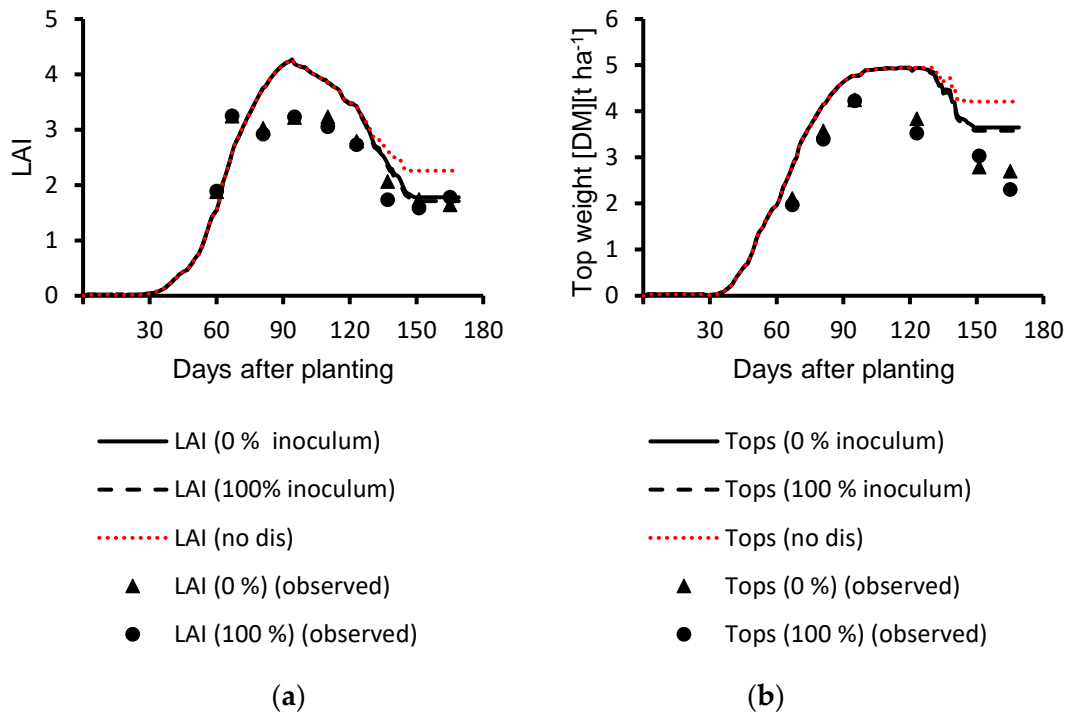
Figure 5-11 Simulated and observed values with Cercospora leaf spot ratings included in simulation results: (a) LAI; (b) top weight (DM t ha⁻¹) for the evaluation treatment with sugar beet growth simulated with current genetics and without Cercospora leaf spot disease ratings being included in the simulation process as "no dis" treatment (Memic et al., 2020).
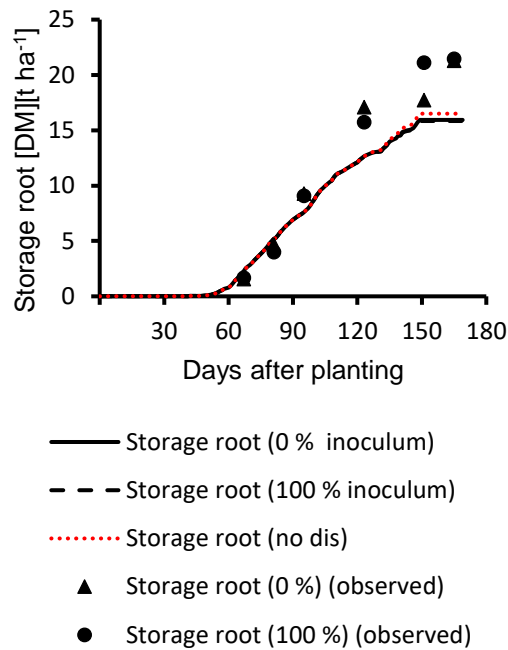
Figure 5-12 Simulated and observed values with Cercospora leaf spot ratings included in simulation results: storage root weight (DM t ha^{-1}) for the evaluation treatment with sugar beet growth simulated with current genetics and without Cercospora leaf spot disease ratings being included in the simulation process as "no dis" treatment (Memic et al., 2020).

5.3.3   Model-Based Yield Losses Evaluation Results

Sugar content was analysed for every sampling date, as described in the Methodology section. Sugar content was measured as percent of dry matter within weeks before harvest (three samples each in 1–2 weeks interval) and the average across all treatments (2016–2018) was 68% as shown in Table 5-8.

Table 5-8 Observed storage root weight (SRW) (DM) kg ha^{-1}, sugar yield (SY) kg ha^{-1} and sugar content as (%) of (DM) (Sc (%)) for three different fungicide treatments: 0%, 50% and 100% in which 0, 0.5, and 1 L ha^{-1} of Spyrale fungicide was applied, respectively, and 0% and 100% inoculum treatments in which 0 and 1 g m^{-2} of inoculum was applied, respectively (Memic et al., 2020).

	DAY	SRW kg ha^{-1}	SY kg ha^{-1}	Sc (%)	SRW kg ha^{-1}	SY kg ha^{-1}	Sc (%)	SRW kg ha^{-1}	SY kg ha^{-1}	Sc (%)
2016		0% fungicide			50% fungicide			100% fungicide		
	138	15,241	10,171	67	16,153	10,625	66	18,319	12,438	68
	152	16,768	11,343	68	15,005	9906	66	18,091	12,366	68
	160	18,024	12,072	67	18,180	12,231	67	22,184	14,911	67
Avg.				67			66			68
2017		0% inoculum			100% inoculum					
	114	11,820	8233	70	10,461	7109	68			
	140	16,538	10,970	66	19,270	13,384	69			
	169	27,956	20,166	72	27,915	19,944	71			
Avg.				69			69			
2018		0% inoculum			100% inoculum					
	119	17,115	11,839	69	15,736	10,937	70			
	147	17,748	12,842	72	21,111	13,941	66			
	161	21,308	14,890	70	21,462	15,533	72			
Avg.				70			69			

In retrospect, sugar yield was quantified as percent of dry matter per experimental plot (Table 5-8) and used for quantifying sugar yield in simulated dry matter. To simulate sugar yield losses (SY loss, based on the measured sugar content percentage shown in Table 5-8), the model-based storage root dry matter quantities (SRW) were evaluated by comparing the no disease treatment ("no dis"—Cercospora leaf spot disease damage ratings not included in the simulation process) and disease treatments ("dis"—with Cercospora leaf spot disease damage ratings included) based on Equation (7) with results shown in Table 5-9.

$$SRW\ loss\ [DM] = SRW'no\ dis'[DM] - SRW'dis'[DM] \tag{7}$$

Table 5-9 Simulated storage root (DM) kg ha⁻¹: disease free storage root ("no dis"—Cercospora leaf spot disease ratings not included in the model input files), storage root simulated quantity with Cercospora leaf spot disease ratings included in the model input files ("dis") and storage root weight (SRW) losses (Memic et al., 2020).

		Simulated Storage Root (DM) (kg ha⁻¹)						
	DAY	'No Dis'	'Dis'	SRW Loss	'Dis'	SRW Loss	'Dis'	SRW Loss
2016				0% fungicide		50% fungicide		100% fungicide
	63	478	478	0	478	0	478	0
	83	3738	3728	10	3728	10	3735	3
	103	8867	8776	91	8741	126	8759	108
	125	15,565	15,269	296	15,202	363	15,207	358
	138	17,983	17,636	347	17,645	338	17,655	328
	152	20,802	20,221	581	20,291	511	20,307	495
	177	22,506	21,232	1274	21,424	1082	21,440	1066
2017				0% inoculum		100% inoculum		
	106	2742	2742	0	2742	0		
	127	9792	9787	5	9779	13		
	140	12,577	12,529	48	12,456	121		
	169	21,108	20,481	627	20,145	963		
	184	28,518	26,068	2450	25,489	3029		
2018				0% inoculum		100% inoculum		
	106	10,971	10,971	0	10,971	0		
	119	12,654	12,651	3	12,632	22		
	126	13,171	13,082	89	13,047	124		
	133	14,418	14,187	231	14,133	285		
	147	16,506	15,939	567	15,841	665		
	154	16,506	15,939	567	15,841	665		
	169	16,506	15,939	567	15,841	665		

Sugar yield loss (SY loss kg ha⁻¹) was then computed with Equation (8), by a sugar content (Sc) of 68% (Table 8) based on the storage root losses (SRW loss) shown in Table 5-9.

$$SY\ loss = SRW\ loss\ [DM] * Sc\ [\%] \tag{8}$$

The results of simulated sugar yield loss (SY loss) based on the simulated storage root weight loss (SRW loss) are shown in Table 5-10. In 2016, three fungicide application rates (0%, 50% and 100%) resulted in different observed leaf disease percentages. Higher applied fungicide amount resulted in lower disease damage and lower storage root loss and consequently lower sugar yield loss as shown in Table 5-10. In 2017 and 2018, *Cercospora beticola* inoculum was mechanically spread as two different treatments (0% and 100%) in order to cause additional leaf disease damage for investigating the impact on above-ground biomass and storage root accumulation rates. Applied inoculum resulted in higher observed leaf disease percentages that correlated with higher storage root losses (Table 5-10). For modelling purposes, in-field leaf disease was observed for designated plots without investigating direct relationships between observed leaf disease percentages and quantities of applied fungicide and inoculum.

Table 5-10 Simulated storage root weight (DM) losses (SRW) (kg ha⁻¹) and corresponding sugar yield losses (SY) (kg ha⁻¹) at harvest as days after planting (DAY) (Memic et al., 2020).

Year	DAY	Treatments	SRW Loss (DM) (kg ha⁻¹)	SY Loss (kg ha⁻¹)
2016	177	0% fungicide	1274	866
		50% fungicide	1082	735
		100% fungicide	1066	725
2017	184	0% inoculum	2450	1666
		100% inoculum	3029	2060
2018	169	0% inoculum	567	386
		100% inoculum	665	452

5.4   Discussion

In 2017, sugar beet was planted at the beginning of April in a period with lower temperatures (Figure 5-1). The model simulated the overall LAI pattern quite well. However, likely due to the lower temperatures around emergence, the model did not simulate LAI growth correctly 20 to 30 days after planting.

The Cercospora leaf spot leaf area disease progress (%) was capturing plot-based disease status very well until leaf disease patches in the plots were observed (within days after leaf area disease progress (%) = 100). Above that level, leaf area disease progress (%) did not reflect the severity of the spread. During the study, Cercospora leaf spot patches were recorded on plot level. A further method needs to be developed that compliments the approach (leaf area disease progress (%)), to enable the user to add information on disease severity based on the observed Cercospora leaf spot disease patches. This additional method will help to integrate Cercospora leaf spot disease severity information from the point where leaf area disease progress (%) method of the middle 10 leaves (long-term leaves) lose explanatory power, closer to harvest time.

When interpreting the results, many factors have to be included, such as the length of the growing period, timing of Cercospora leaf spot occurrence and leaf area disease percentages.

For example, lower Cercospora leaf spot disease damage introduced earlier in the CSM-CERES-Beet model led to higher reduction in above-ground biomass compared to higher leaf disease damage introduced in the model closer to the harvest time. Sugar beet has the ability to produce new leaves throughout its entire vegetative stage (in its first year). Depending on the soil and weather conditions, it can "replace" lost leaves. Currently, the model does not account for this, as it is accumulating dry matter on a daily basis.

The occurrence of Cercospora leaf spot disease depends on specific weather factors, as optimum daily temperatures are 20 to 25 °C (Wolf and Verreet, 2002). Cercospora leaf spot can occur at lower temperatures and a broad range of humidity (Wolf and Verreet, 2002) and changing microclimates within a plant stand. It was observed in the experimental field that Cercospora leaf spot often exhibited a patchy distribution later in the growing season, close to harvest.

Reliable and timely assessments of Cercospora leaf spot occurrence and spread are the basis for planning targeted plant protection activities in the field. Visual plant disease estimations by extension officers is one way to collect these data, or leaf disease spread simulation models such as that developed by Rossi and Battailani. Rossi and Battailani (1991) used the CERCOPRI model to quantify the effects of Cercospora leaf spot on sugar yield. Rossberg et al. (2000) modified CERCOPRI in order to simulate early Cercospora leaf spot epidemic's impact on sugar beet growth and yield and to evaluate the impacts of fungicide applications. Their work resulted in the development of CERCBET 1, which was further improved by Racca et al. (2004). Other significant Cercospora leaf spot forecasting models are: the leaf spot model for sugar beet (Windels et al., 1998), the integrated pest management system in Germany (Wolf and Verreet, 2002), and the integrated surveillance of leaf disease in sugar beet (Mittler et al., 2004). All of these model development efforts were conducted in order to moderate application of chemicals based on the environmental conditions combined with field scouting reports.

If crop growth models can be coupled in the future with suitable sensor technologies (Bock et al., 2020) or models capable of predicting leaf disease occurrence based on the leaf disease favouring weather conditions, their potential as decision support tools is enormous. Hyperspectral imaging can be used for analysis of Cercospora leaf spot as shown in Leucker et al. (2017). More importantly, various aspects of the crop growth and leaf disease dynamics and interactions can be investigated in detail. Using a crop model, impact of soil profile (e.g., soil texture, soil water holding capacity, soil organic matter etc.) and daily weather data (temperature minimum and maximum, precipitation, and solar radiation) on overall crop growth can be investigated in more detail. With the ability to estimate growth limiting factors and leaf disease effects, a detailed economic analysis can be conducted based on the detailed field information included in the crop model analysis. With further development and improvement, the CSM-CERES-Beet might be used as a decision support system, coupled with sensors capable of quantifying Cercospora leaf spot diseases in sugar beet. Overall, further model developments are needed as leaf disease severity information is used for the evaluation of sugar beet dry matter losses per defined plot. Nevertheless, three years of observed data for this specific cultivar are not enough for determining Cercospora leaf spot damage. There is a need to look at more than one cultivar and in a greater diversity of fields and environmental conditions to further improve the models.

5.5   Conclusions

Field experiments were conducted over three years to develop and test the Cercospora leaf spot disease subroutines for simulating the damage caused by Cercospora leaf spot disease in sugar beet with CSM-CERES-Beet. Values for Cercospora leaf spot leaf area disease progress (%) were converted into leaf disease damage rates (internally in the model) and applied to the selected disease coupling point. Introducing leaf disease impact played a very important role in simulating storage root yield and sugar content during the later sugar beet growing period and led to an overall better fit between observed and simulated values when compared to the results where disease damage was not reported or included. The approach can serve as a suitable decision support system to simulate the impact of observed Cercospora leaf spot damage on accumulated above-ground biomass and storage root yield on a plot/site-specific scale.

6 Cultivar Coefficient Estimator for the Cropping System Model Based on Time-Series Data - A Case Study for Soybean

Abstract

The Decision Support System for Agrotechnology Transfer is one of the most popular crop modelling software solutions for predicting crop growth and yield while capturing the effect of management practices and interactions between the crop and environment. An accurate estimation of the crop cultivar specific coefficients that govern in-season growth and development is critical for correct yield estimates. The manual cultivar coefficient estimation process is time consuming and results in user-dependent, subjective optimums that are difficult to reproduce. Typically, end-of-season observations (point based) are used for estimating dynamic in-season biomass accumulation rates. The objective of this study was the development of a tool capable of using multiple in-season observations for estimating coefficients that define in-season growth and partitioning. Using the time-series cultivar coefficient estimator, coefficients were estimated based on multiple in-season observations for leaf area index, shoot, leaf, and grain weights. The cultivar coefficients were estimated from single- and multiple-treatment (seasons/locations) in-season observations. This was done for two cultivars for six management × environment combinations. Estimated multiple-treatment-based cultivar coefficients were evaluated with an independent data set and compared to DSSAT standard (manual) and the cultivar coefficients estimated with the GLUE tool. The average normalised root mean squared error for grain weight, leaf area index, shoot weight and leaf weight was 26% lower for one cultivar and about the same for the other when compared to the DSSAT standard. Since GLUE uses an end-of-season point-based cultivar coefficient estimation approach, the grain weight over time was under-estimated in earlier phases and more accurate towards harvest. The times-series estimator estimated cultivar coefficients, based on 346 in-season observations across multiple target variables and six experiments, reflected more accurately in-season growth and grain weight without compromising final grain weight predictions.

Keywords. DSSAT; CROPGRO-Soybean; genetic coefficients; normalized root mean square error minimization; time-series observations

6.1 Introduction

A wide range of crop models have been developed for various purposes, such as yield prediction, agricultural production inputs management evaluation, and assessment of long-term impacts of agricultural management practices on soil and environment degradation (Boote et al., 2010; Ewert et al., 2015; Rötter et al., 2015; Tsuji et al., 1998). In general, these models are capable of predicting crop growth and quantifying yield limiting factors (Hoogenboom et al., 2019a; Thorp et al., 2010), while capturing effects of crop management (fertilizer, sowing date, sowing density, etc.) and interactions between crops and environment

(soil, weather etc.) (Jones et al., 2003). The Decision Support System for Agrotechnology Transfer (DSSAT) software represents a conceptual and practical solution for capturing many important factors affecting agricultural production of more than 40 crops (Hoogenboom et al., 2019a). Within the DSSAT, the Cropping System Model (CSM)-CROPGRO-legume model (Boote et al., 1998) simulates crop growth and development from planting to harvest on a daily basis (carbon and nitrogen balances) throughout the vegetative and reproductive phases with different biomass and yield accumulation rates.

The CROPGRO model simulates canopy photosynthesis on an hourly basis based on leaf-level photosynthesis parameters and simulates complete plant C and N balance including N effects on photosynthesis, biomass accumulation and grain growth (Boote et al., 1998). Model outputs are crop growth variables, soil water and crop N balance on a daily basis. Within the CROPGRO model, various legumes such as soybean (*Glycine max* L. Merr.), peanut (*Arachis hypogaea* L.), dry bean (*Phaseolus vulgaris* L.), cowpea (*Vigna unguiculata* L.), faba bean (*Vicia faba* L.), velvet bean (*Mucuna pruriens* L.), and chickpea (*Cicer aretinum* L.) (Boote et al., 2009; 2021) can be simulated. The conceptual design of defining specific genetic traits through input files (species, ecotype and cultivar input files) enables the use of generic algorithms for simulating crop growth and development of multiple crops without modifying crop development and growth subroutines individually. The cultivar input files contain information for differentiating cultivars within a crop species and contain parameters important for influencing photoperiodic response, photothermal durations of specific growth phases, leaf appearance rate, seed fill duration and composition (Hoogenboom et al., 2011; Hoogenboom et al., 2019a). The CROPGRO model cultivar file contains 18 cultivar coefficients. These cultivar coefficients are externally defined and setup in a way to enable users to modify internally defined crop growth processes such as defining the influence of specific cultivar parameters on crop growth and development rates. In general, the overall model application relies considerably on the estimation of cultivar coefficients and the reliability and accuracy of their evaluation (Seidel et al., 2018). In many cases, comprehensive experimental data sets for model evaluation are rare (Hoogenboom et al., 2012; White et al., 2013; Boote et al., 2015; Kersebaum et al., 2015).

Currently, two different tools are available within DSSAT for optimization of the cultivar coefficients: genotype coefficient calculator (GENCALC) (Hunt et al., 1993) and generalized likelihood uncertainty estimation (GLUE) (He et al., 2010; Jones et al., 2011; Boote, 2019). Both tools have been included in DSSAT Version 4.5 (Hoogenboom et al., 2011) and later DSSAT versions (Hoogenboom et al., 2019b). Both GENCALC and GLUE use *end-of-season field observations* (point based) introduced into the model through "FileA" for estimation of the cultivar coefficients. FileA contains data that are collected only once in a season such as yield and yield components, or are summarized during the season such as anthesis and maturity dates and maximum leaf area index (LAI). Within GENCALC, the selection of optimal cultivar coefficients is based on Root Mean Square Error (RMSE) (Jones et al., 1998). The cultivar coefficients combination with the smallest difference between observed and model simulated values is taken as an optimum value. Within GLUE, a Monte Carlo distribution sampling method based on a Bayesian estimation approach with Gaussian likelihood function is used for optimizing cultivar coefficients (He et al., 2010; Jones et al., 2011). Similar to GENCALC, only end-of-season observations can be introduced for multiple target variables into GLUE through "FileA" for error minimization between model-simulated outputs and observed data. The approach optimizes all phenology-related coefficients at one time, then all growth-related

cultivar coefficients, with the strongly recommended possibility to consider multiple-treatments of the same cultivar (He et al., 2010; Jones et al., 2011; Gao et al., 2020).

However, **GENCALC** and **GLUE** *cannot handle inputs of time-series observations* (Buddhaboon et al., 2018). Therefore, the only option for a user to use time-series observations for optimization of coefficients is to conduct manually. Calibration of models against time-series data traditionally attempts to reduce total model prediction error by changing parameters so that simulations match or closely resemble observations. The goal of this study was to develop a *cultivar coefficient estimator* tool based on both *time-series observations* and end-of-season observations. It is expected that time-series experimental observations (*multiple in-season observations*) will enable a more accurate estimation of cultivar coefficients, because the time-series observations are the outcome of cultivar coefficient effects on dynamically-varying growth rates. By extending the list of the target variables to total shoot weight, leaf area index (LAI), leaf weight, and other observations measured during the growing season, additional aspects of the crop model performance throughout the season can be evaluated and used for estimating cultivar coefficients in addition to the onset of flowering and crop physiological maturity.

The specific objectives of the study were: 1) to develop a method for the estimation of phenology- and growth-related cultivar coefficients for the CSM model of DSSAT, and 2) to evaluate an error minimization method that includes single and multiple seasons/treatments of experimental in-season observations with an emphasis on the multiple-treatment-based cultivar coefficient estimates for establishing more robust cultivar coefficient values representative of multiple locations and seasons.

6.2   Materials and Methods

6.2.1   Experimental data

Experimental data for soybean conducted in different locations in the USA were selected for testing the new tool. The experimental data were initially collected for evaluating the impact of irrigation management and weather on the performance of soybean. The experiment conducted in Ohio in 1990 was meant to be irrigated but never required irrigation. The experimental files including management practices, weather, soil characteristics, and observations of crop growth and yield are found within the official DSSAT Version 4.7 (Hoogenboom et al., 2019b). Four selected irrigated treatments (Table 6-1) of soybean cultivar Bragg were grown on an Arredondo fine sand soil with planting dates in June. Detailed information of environmental conditions (seasonal rainfall, irrigation amounts and mean temperatures) are shown in Table 6-1 with the corresponding row spacing, plant populations, observed yield and biomass (Wilkerson et al., 1983; Boote et al., 1997). The soybean cultivar Williams (four selected treatments) was grown in Ohio in 1988 (Irrigated) and 1990 and in Iowa in 1988 and 1990 (Table 6-1). Detailed environmental conditions and measured data are shown in Table 6-1, with corresponding row spacing and plant populations.

Table 6-1 Cultivar and related crop management information for the experiments used for model calibration and evaluation (Wilkerson et al., 1983; Boote et al., 1997) (From Memic et al., (2021). Used with permission.).

Location	Treatment	Planting date (dd/mm)	Row spacing (cm)	Plant population (plants m^{-2})	Seasonal rainfall (mm)	Total irrigation (mm)	Mean temp. (°C)	Seed yield (kg ha^{-1})	Biomass at maturity (kg ha^{-1})
Bragg									
Gainesville (1976)	Irrigated	05/5	30	12.9	776	75	26.12	3439	6848
Gainesville (1978)	Irrigated	15/06	91	29.9	534	196	27.12	3041	6068
Gainesville (1979)	Irrigated	19/06	91	47.0	695	113	26.95	2891	5781
Gainesville (1984)	Irrigated	12/6	76	31.1	469	382	26.34	3723	6689
Williams									
Ohio (1988)	Irrigated	01/05	19	31.1	370	557	20.61	3976	8090
Ohio (1990)	Rainfed	30/4	19	41.2	581	-	19.24	3149	7113
Iowa (1988)	Rainfed	11/05	70	27.2	381	-	23.26	3222	6617
Iowa (1990)	Rainfed	08/5	70	19.2	776	-	21.50	3168	6674

Seasonal rainfall, mean temp. - Crop growing season, from planting to harvest, Ohio 1990, Iowa 1990 - Irrigation was intended but never required, Iowa 1988 - Rainfed with significant water stress

6.2.2 Time-Series cultivar coefficient Estimator

The **T**ime-**S**eries observations-based cultivar coefficient **E**stimator (**TSE**) was developed as a potential DSSAT plug-in with generic algorithm, written in python with an intuitive interface. The TSE requires functional experimental input files in DSSAT4.7 such as complete crop management, weather data, soil surface and profile data, and observations stored in a time-series file (FileT) and a summary file (FileA). It is assumed that functioning default genetics files are provided for species, ecotype and cultivar.

The program setup and optimization were designed in three steps. During the first step (Figure 6-1; Step 1) the experiment treatment(s) (or multiple treatments) that will be used for computation of the cultivar coefficients is (are) selected by a user. At this stage the user can select optimization of the phenology- and growth-related cultivar coefficients separately. Phenology-related coefficients in CROPGRO-Soybean are optimized during the first round. The optimization of the phenological events such as the onset of flowering and physiological maturity is not based on time-series observations and, as such, is easier to optimize by minimizing the difference between simulated and observed day of onset of flowering or physiological maturity. The optimization of the growth-related cultivar coefficients is conducted by using all available time-series observations throughout the season and by minimizing the normalized root mean square error between simulated and observed values of total crop dry weight, grain dry weight, leaf dry weight, and leaf area index. Based on the selected treatment, all available in-season observations are read in and saved in a temporary file for a later comparison with simulation outputs. In this step (Figure 6-1; Step 1.) initial ranges of selected cultivar coefficients and associated incremental steps are defined by the user. The allowed range of cultivar coefficients is defined by the user in a program input file, based on literature knowledge and previously determined cultivars within the desired cultivar group. The initial coefficient ranges and incremental steps are read by the program, but enable the user to modify the ranges and increment steps if deemed necessary during the cultivar coefficient estimation procedure. During the second step, the crop model execution is completed (Figure 6-1; Step 2). The model is executed for each given coefficient combination and simulations were conducted for as many cultivar coefficient combinations as defined during coefficient preparation in Step 1, Figure 6-1. After each model run, the simulated output values are extracted from the model time-series simulation outputs and saved in a TSE output file containing both simulated and observed outputs. During the final step, the TSE output file is analyzed and the coefficient combination giving the smallest average nRMSEs is selected as the optimum (Figure 6-1; Step 3.).
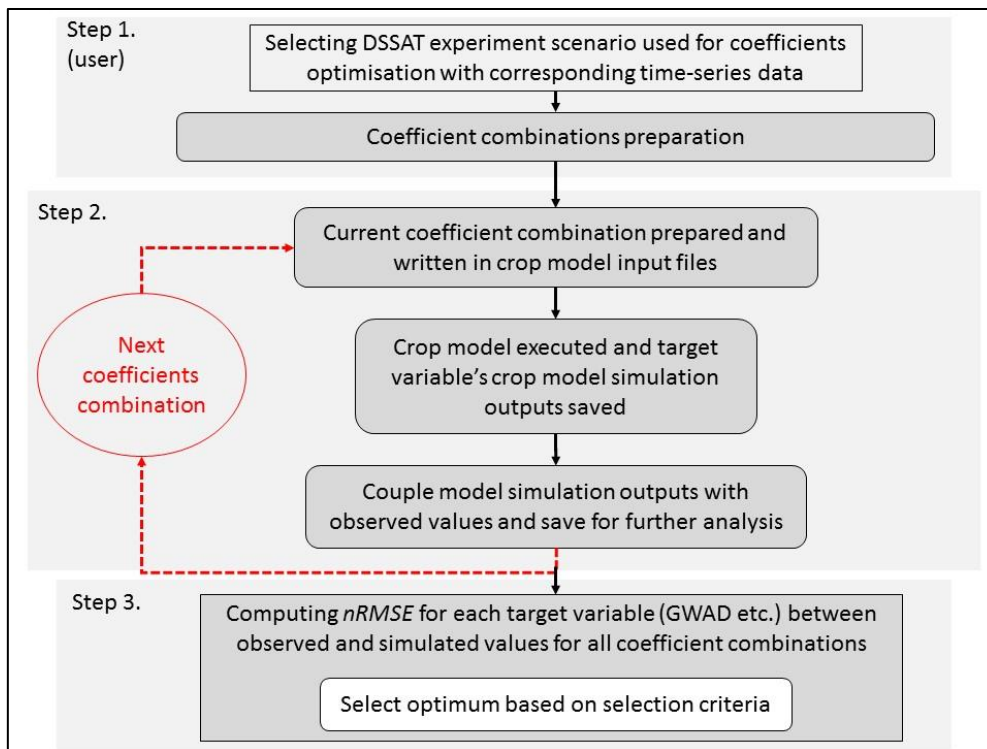
Figure 6-1 Flow chart showing the overall approach used for model calibration (Röll et al., 2020) (From Memic et al., (2021). Used with permission.).

The TSE generic algorithm, when executed and depending on selected DSSAT version, reads all available crop models listed in SIMULATION.CDE file located in the native DSSAT directory (Figure 6-2b). After selecting the desired model from the list, the TSE algorithm collects information required for running the model and modifying the cultivar files from the DSSAT configuration file (DSSATPRO) located in the DSSAT native directory (Figure 6-2b).

a   b
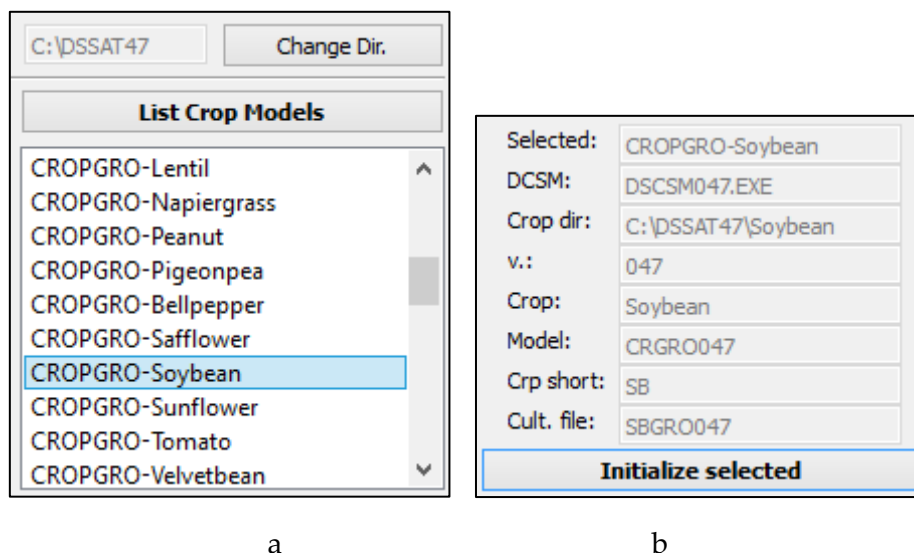
Figure 6-2 The Time-Series observations-based cultivar coefficient Estimator (TSE) interface section of a) the list of crop models and b) corresponding model specifications of the DSSATPRO file (From Memic et al., (2021). Used with permission.).

Based on the selected model and model-specific properties shown in Figure 6-2b the TSE algorithm lists the names of all available cultivars in the cultivar file. For the example shown in Figure 6-2b, the TSE algorithm locates the soybean cultivar file (SBGRO047.CUL) and lists all cultivars available in the file for selection in an additional list widget window (Figure 6-3a). After the cultivar is selected from the list, the TSE algorithm locates all experiment files available in the crop directory and lists for selection only those experiment files (Figure 6-3b) that contain the selected cultivar name (cultivar Bragg selected in Figure 6-3a).

a   b
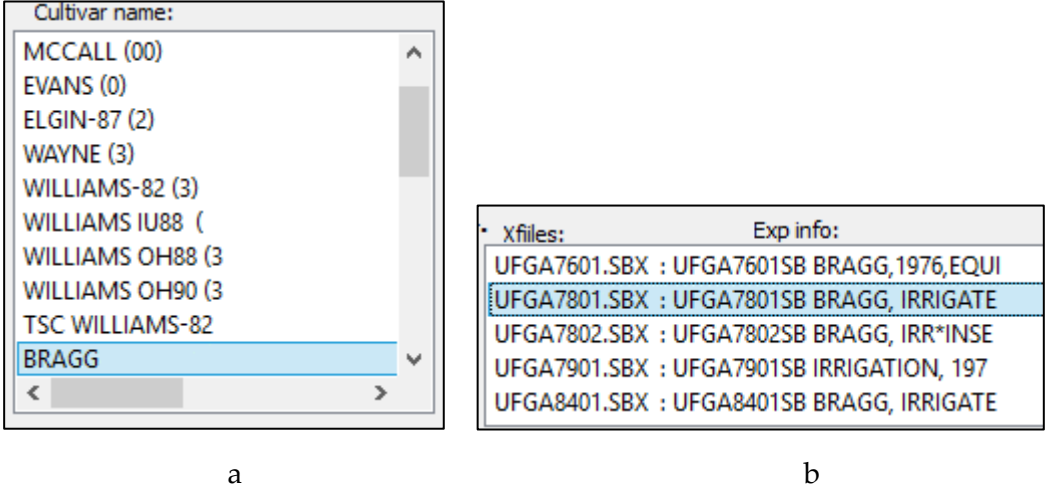
Figure 6-3 The Time-Series observations-based cultivar coefficient Estimator (TSE) interface section of a) the cultivar list and b) the corresponding experiment files list (From Memic et al., (2021). Used with permission.).

Based on the selected crop model experiment file (Figure 6-3b) the TSE generic algorithm locates this specific experiment file with related in-season observations and lists target variable names in the list widget window for selection and initialization (Figure 6-4a). At the same time in a different list widget window, all cultivar coefficients available in the cultivar file are listed for optimization selection (Figure 4b).

a   b
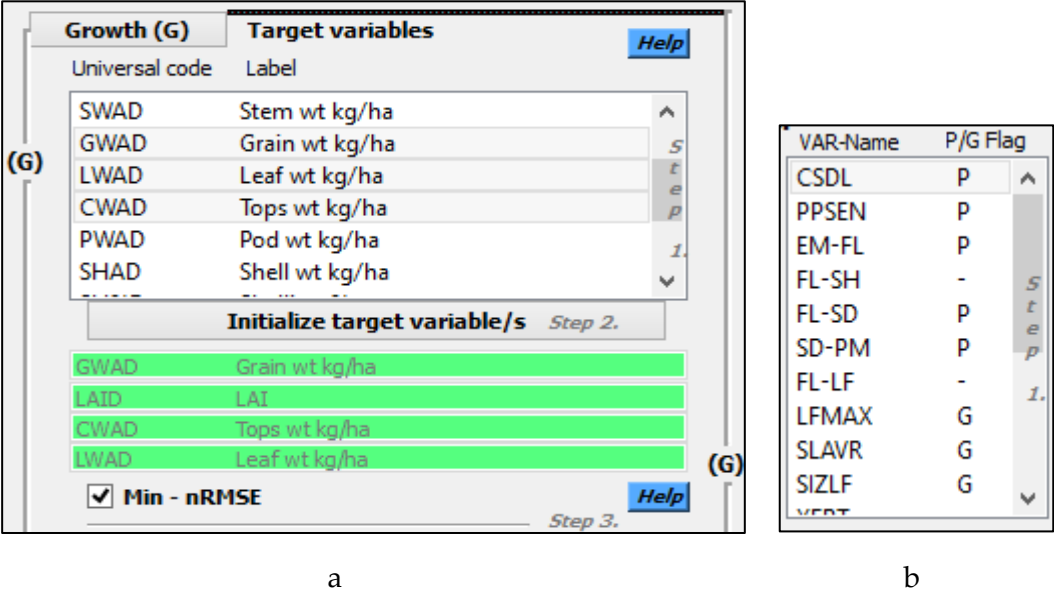
Figure 6-4 The Time-Series observations-based cultivar coefficient Estimator (TSE) interface section with a) an example for selecting growth-related target variables and b) the available cultivar coefficients (From Memic et al., (2021). Used with permission.).

The TSE algorithm dynamically locates the cultivar header line in the cultivar file and lists possible genetic coefficients in the list widget window (Figure 6-4b). One of the options in the program is to enable access only to those cultivar coefficients with predefined **P**henology/**G**rowth related flags (=P or G), with a separate coefficient flags file. The functionality of the TSE algorithm is not dependent on the mentioned flags, but they can be used to help users if they are uncertain whether a coefficient is meant for optimization of a given phenological event or growth-related variable. The generic TSE algorithm was written in a way to enable cultivar coefficient estimation of all the crop models available in the DSSAT interface including CROPGRO and CERES-style models. A more detailed description of the TSE algorithm functionality is available in the user guidelines available in the GitHub repository (https://github.com/memicemir/TSE).

Within the CROPGRO-Soybean cultivar file, there are a total of 18 coefficients available, used for determining plant development and growth aspects of different cultivar groups (Boote et al., 2003) and are listed in list widget shown in Figure 6-4b. The program was used for estimating only eleven coefficients of those listed in Table 6-2. The selected cultivar coefficients are those that differ the most within cultivar groups and have a substantial effect on the simulated outputs and were designated with Phenology or Growth flags visible in Figure 6-4b. The coefficients in Table 6-2 are separated into those that are used for determining the timing of cultivar-specific phenological events such as onset of first flowering day, first pod day, first seed day and physiological maturity day (Figure 6-4b, P flagged) and those that are related to biomass and yield accumulation rates during specific growth phases (G flagged). The distinction of the coefficients into these two groups has an important theoretical and practical background. Crop phenology depends on day length and temperature and, thus, is mostly independent from growth, but on the other hand, crop growth (biomass and yield accumulation rates) are affected by all environmental and management factors as well as phenological development (Jones et al., 2011).

Table 6-2 List of coefficients targeted for optimization and their definitions in the cultivar file for the CROPGRO-Soybean model (Boote et al., 2003) (From Memic et al., (2021). Used with permission.).

Coeff.	Definitions	Units
CSDL (P)	Critical Short-Day Length below which reproductive development progresses with no day length effect (for short day plants)	hour
PPSEN (P)	Slope of the relative response of development to photoperiod with time (positive for short day plants)	1/hour
EM-FL (P)	Time between plant emergence and flower appearance (R1)	photothermal days
Fl-SH (P)	Time between first flower and first pod (R3)	photothermal days
FL-SD (P)	Time between first flower and first seed (R5)	photothermal days
SD-PM (P)	Time between first seed (R5) and physiological maturity (R7)	photothermal days
LFMAX (G)	Maximum leaf photosynthesis rate at 30 C, 350 vpm CO_2, and high light	mg CO_2/m^2-s
SLAVR (G)	Specific leaf area of cultivar under standard growth conditions	cm^2/g
SIZLF (G)	Maximum size of full leaf (three leaflets)	cm$_2$
WTPSD (G)	Maximum weight per seed	g
SFDUR (G)	Seed filling duration for pod cohort at standard growth conditions	photothermal days

P - Phenology-related coefficient, G - Growth-related coefficient

6.2.3 Error minimization

Single-treatment based normalized RMSE

For quantifying the variation between simulated (S_i) and observed (O_i) values, the statistical method of nRMSE (Eq. 2) was used. The nRMSE, is RMSE (Eq. 1) normalized by mean of all available in-season observations ($\overline{O}$), for each observed crop variable in Eq. 2.

$$RMSE = \left[\frac{1}{n}\sum_{i=1}^{n}(S_i - O_i)^2\right]^{0.5} \quad (1)$$

$$nRMSE = \frac{RMSE}{\overline{O}} \quad\quad (2)$$

The nRMSE is a simplified selection criterion that is applicable across multiple target variables with different scales, and because it is normalized, it can be averaged over multiple target variables. The coefficients (n *coefficient combinations*) are estimated across multiple target variables, with the aim of a low average nRMSE over all variables (Eq. 3).

$$nRMSE_{avg(n)} = (nRMSE_{gwad} + nRMSE_{lai} + nRMSE_{cwad} + nRMSE_{lwad})/4 \quad (3)$$

The selection of the coefficient combination with the lowest nRMSEs averaged across all target variables proves to be a good solution as illustrated for the 1978 Gainesville experiment (Table 6-3, AVG-nRMSE=0.12).

Table 6-3 The nRMSE - simplified example of varying one of the cultivar parameters affecting growth (G) related target variables, i.e., grain weight, LAI, shoot weight, and leaf weight, for the 1978 Gainesville experiment for Bragg and optimum selection based on the lowest average nRMSE (AVG-nRMSE) over multiple target variables (From Memic et al., (2021). Used with permission.).

Coefficient	nRMSE (unitless)				
LFMAX	Grain weight	LAI	Shoot weight	Leaf weight	AVG-nRMSE
0.8	0.208	0.22	0.185	0.203	0.204
0.912	0.131	0.153	0.119	0.146	0.137
1.024	0.082	0.145	0.109	0.144	**0.12**
1.136	0.078	0.171	0.137	0.173	0.14
1.248	0.109	0.209	0.178	0.211	0.177

LFMAX is defined in Table 2, AVG-nRMSE-average of normalized RMSE over four target variables (grain weight, leaf area index, shoot weight, leaf weight)

Multiple-treatment based goodness of fit criteria

The new TSE tool allows a user to estimate cultivar coefficients over multiple treatments. The optimization of cultivar coefficients for a single season and location is extended further for optimization over multiple seasons and locations with emphasis on nRMSE values. An example of optimizing LFMAX (as defined in Table 6-2) based on the observations from three different treatments is shown in Table 6-4. Calculated average nRMSE (AVG-nRMSE) over four target variables (as shown in Table 6-3) are extended and used for localizing the optimum over multiple target variables and multiple treatments (Table 6-4). If the number of the observations is higher for one treatment compared to other treatments, this procedure limits the over-representation of single treatment effects on the multi-treatment based average nRMSE.

For each individual treatment, the same value for minimum (0.85), maximum (1.25) and increment step (0.1) for the LFMAX cultivar coefficient values were passed into the cultivar file and simulations were conducted (Table 6-4). For each cultivar coefficient (LFMAX) value, the crop model simulations were conducted and the average nRMSE was calculated for the four target variables (grain weight, LAI, shoot weight and leaf weight). Single treatment "optimums" are available (Table 6-4), but they were ignored for the calculation of the multi-treatment based "optimum" (Table 6-4). For each coefficient value, a multi-treatment-based average of a single treatment normalized RMSE (AVG-nRMSE) averages was calculated. Based on the lowest multiple-treatment nRMSE average [(TRT$_1$+TRT$_2$+TRT$_3$)/3] value, the multiple-treatment based cultivar coefficient optimum with LFMAX=0.95 was selected (Table 6-4) for having the smallest AVG-nRMSE that equals 0.214 (Table 6-4). According to the AVG-nRMSE values shown in Table 6-4, the "optimum" coefficient range for LFMAX falls between 0.95 and 1.05, since these two LFMAX values had the lowest average nRMSEs of 0.214 and 0.218, respectively, when compared to the AVG-nRMSEs of other LFMAX coefficient such as 0.85 or 1.15. The step increment could be reduced, e.g., to 0.01 to determine whether a lower average nRMSE could be obtained with a value for LFMAX between 0.95 and 1.05.

Table 6-4 The effect of the variation of the LFMAX cultivar coefficient from 0.85 to 1.25 with an increment of 0.1 while localizing single- and multiple-treatment optimums based on average nRMSE (AVG-nRMSE) for three experiments conducted in Gainesville, FL (From Memic et al., (2021). Used with permission.).

Single-treatment "optimums"				Multi-treatment based "optimum"				
Year	TRT	LFMAX	AVG nRMSE	Year	TRT	LFMAX	AVG nRMSE	Multiple TRT average
1978	1	0.85	0.17	1978	1	0.85	0.17	
1978	1	0.95	0.128	1979	2	0.85	0.119	
1978	1	1.05	0.122	1984	3	0.85	0.408	0.232
1978	1	1.15	0.142	1978	1	0.95	0.128	
1978	1	1.25	0.175	1979	2	0.95	0.14	
1979	2	0.85	0.119	1984	3	0.95	0.374	0.214
1979	2	0.95	0.14	1978	1	1.05	0.122	
1979	2	1.05	0.177	1979	2	1.05	0.177	
1979	2	1.15	0.21	1984	3	1.05	0.355	0.218
1979	2	1.25	0.239	1978	1	1.15	0.142	
1984	3	0.85	0.408	1979	2	1.15	0.21	
1984	3	0.95	0.374	1984	3	1.15	0.346	0.233
1984	3	1.05	0.355	1978	1	1.25	0.175	
1984	3	1.15	0.346	1979	2	1.25	0.239	
1984	3	1.25	0.345	1984	3	1.25	0.345	0.253

TRT-treatment, LFMAX-defined in Table 2, AVG-nRMSE-average of normalized RMSE over four target variables (grain weight, leaf area index, shoot weight, leaf weight).

Single- and multiple-treatment error minimization study cases

The TSE normalized RMSE minimization method was tested as a single-treatment for the estimation of the cultivar coefficients for two soybean cultivars, i.e., Bragg, and Williams, for a total of six treatments. For Bragg and Williams, three treatments were used for each cultivate to estimate the multiple-treatment based cultivar coefficients. Multiple-treatment based cultivar coefficient "optimums" of Bragg and Williams were evaluated with data sets that were not included in the calibration process (Gainesville 1976 and Iowa 1990). The 1976 Gainesville experiment was selected for evaluation, because observed values for onset of flowering day, first pod, first seed and physiological maturity were not available and, therefore, any form of phenological events optimization was not possible. Both cultivars had been tested in the past for multiple years and for different environments and demonstrated the ability of the model to accurately simulate growth, flowering and physiological maturity, and grain yield for each cultivar for different environmental conditions (Boote et al., 1997). The cultivar coefficients obtained with the TSE tool were compared with those cultivar coefficients that were distributed with the official DSSAT Version 4.7 release (Hoogenboom et al., 2019b), which were shown as *DSSAT standard* Bragg and Williams in this study. The values for the cultivar coefficients that are distributed with the official DSSAT Version 4.7 release had been manually estimated by DSSAT developers based on seven treatments for Bragg and eight treatments for Williams cultivar to obtain the best performance over multiple environments and seasons per cultivar. All treatments were simulated with soil and plant water and nitrogen balance "on", and the leaf level hedgerow photosynthesis method selected in the corresponding experimental files. For the 1988 Iowa experiment, drought was observed with expected stress on the observed target variables. Ideally stress-free observations should be used for optimization of cultivar coefficients, but in this study one stressed treatment was used in order to check the cultivar coefficient values estimated directly from stressed in-season observations. During the first round, phenology-related cultivar coefficients were optimized with respect to the observed anthesis date, first pod date, first seed date, and physiological maturity. During the second round, the growth-related cultivar coefficients were estimated based on four target variables, i.e., grain weight, leaf area index, shoot weight and leaf weight, with in-season observations.

Because GLUE is an existing cultivar coefficient optimizer that is distributed with the DSSAT Software, it was important to compare the developed TSE approach (time-series) to the GLUE approach (end-of-season). In order to enable a direct comparison of the *DSSAT standard*, the TSE approach and the GLUE approach, the GLUE tool was used for estimating multiple-treatment based cultivar coefficients with the same sets of three. For estimation of the Bragg cultivar coefficients with multiple-treatments with GLUE, the Gainesville 1978, Gainesville 1979 and Gainesville 1984 treatments were used, and for estimating the cultivar coefficients for Williams, the Ohio 1988, Ohio 1990 and Iowa 1988 were used. The number of simulations for the GLUE coefficient estimation process was 6000 for optimizing phenology-related and 6000 for growth-related target variables.

6.3   Results

6.3.1   Goodness of fit – Single- and Multiple-treatment

The phenology- and growth-related statistics resulting from the TSE single-treatment (S-T) based Bragg and Williams cultivar coefficients are shown in Tables 6-5 and 6-6. The agreement

of simulated with observed onset of flowering day, first pod day, first seed day, and physiological maturity day is shown as the total phenological event error (Table 6-5). Phenological events simulated values of S-T are shown in Table 6-5 with "Error" line (Error=simulated-observed), indicating the difference between simulated and observed phenological events. The total error over all simulated phenological events with *TSE* for the Gainesville 1978 (S-T) experiment was 5 days, and was calculated as the sum of absolute errors of each phenological event in order to prevent error compensations in the process of summation [|(+2)|+|(-1)|+|(0)|+|(-2)|=|5|]. The phenology-related TSE coefficients optimization resulted in an almost perfect agreement between simulated and observed phenological events for the 1979 and 1984 (S-T) experiments conducted in Gainesville with a total error of 0 and 1 day, respectively (Table 6-5). For the 1988 Iowa (S-T) experiment the TSE optimization resulted in only 2 days total error, and for the Ohio 1988 in 4 days total error over four phenological events (Table 6-5). For the 1990 Ohio (S-T) experiment, a perfect agreement between simulated and observed three phenological events was achieved, with the onset of flowering reported as missing.

Multiple-treatment (M-T) *TSE* based Bragg cultivar coefficients were used for simulating phenological events and four target output traits for the 1978, 1979, and 1984 experiments conducted in Gainesville. The corresponding statistics for phenology are shown in Table 6-5 and for growth in Table 6-6. The total sum of absolute error over all M-T simulated phenological events with *TSE* for the 1978 Gainesville experiment was 3 days, for the 1979 Gainesville experiment it was 6 days and for the 1984 Gainesville experiment it was 2 days, therefore resulting in an almost perfect agreement between simulated and observed data. Multiple-treatment based coefficients of Williams were used for simulating phenological development (Table 6-5) and four target traits (Table 6-6) for the 1988 and 1990 Iowa experiments and the 1988 Ohio experiment. The total summed absolute error over all simulated phenological events with *TSE* for the 1988 Ohio experiment was 6 days, for the 1990 Ohio experiment 5 days and for the 1988 Iowa experiment 9 days as shown in Table 6-5.

Table 6-5 Measures of agreement between simulated and observed phenological events as days after planting (DAP) of single-treatment (S-T) and multiple-treatment (M-T) based phenology-related TSE cultivar coefficient estimates for the cultivars Bragg and Williams evaluated over three locations each (From Memic et al., (2021). Used with permission.).

-	-	Bragg						Williams					
-	-	Gainesville 1978		Gainesville 1979		Gainesville 1984		Ohio 1988		Ohio 1990		Iowa 1988	
-	-	DAP	Error	DAP	Error	DAP	Error	DAP	Error	DAP	Error	DAP	Error
S-T	Anthesis	47	(+2)	43	(0)	48	(+1)	71	(0)	76	(ND)	51	(0)
S-T	First pod	66	(-1)	60	(0)	66	(0)	87	(-4)	96	(0)	70	(-1)
S-T	First seed	77	(0)	69	(-)	78	(0)	100	(0)	101	(0)	83	(+1)
S-T	Maturity	114	(-2)	119	(0)	121	(0)	145	(0)	150	(0)	129	(0)
S-T	|Total error|		|5|		|0|		|1|		|4|		|0|		|2|
M-T	Anthesis	45	(0)	44	(+1)	45	(-2)	74	(+3)	73	(ND)	55	(+4)
M-T	First pod	65	(-2)	64	(+4)	66	(0)	91	(0)	92	(-4)	74	(+3)
M-T	First seed	77	(0)	75	(ND)	78	(0)	98	(-2)	101	(0)	83	(+1)
M-T	Maturity	117	(+1)	118	(-1)	122	(0)	146	(+1)	149	(-1)	128	(-1)
M-T	|Total error| [b]		|3|		|6|		|2|		|6|		|5|		|9|

Error = (simulated – observed), -/+ = underestimated/overestimated, |Total error| - Total absolute error, ND - No data

For estimating S-T growth-related cultivar coefficients, a total of 346 in-season observations were used over six experiments reflecting in-season growth of four target variables for each experiment. With the exception of the grain weight in-season observations, all target variables had on average greater than 15 observations per season (Table 6-6). The *TSE* lowest AVG-nRMSE of Bragg and Williams cultivars for various coefficient combinations over four target variables, e.g., grain weight, LAI, shoot weight and leaf weight, as single-treatment (S-T) based cultivar coefficient estimation was between 0.101 and 0.201, indicating very good results (Table 6-6). The RMSE between simulated and observed grain weight ranged between 140 and 251 kg ha^{-1}, which was very low and an indicator of extremely good performance, since the RMSE error maintains target variable unit and was calculated over multiple in-season observations. The RMSE measure of agreement of simulated and observed for all experiment data was extremely good, as can be seen in Table 6-6, with the exceptions of the shoot weight and LAI for Ohio 1988.

The *TSE* lowest AVG-nRMSE for the cultivars Bragg and Williams for various growth coefficient combinations over four target variables ranged between 0.14 and 0.234, (Table 6-6). The RMSE error of simulated and observed grain weight was between 140 and 320 kg ha^{-1} over six treatments, for multiple in-season observations and indicated very good performance.

Table 6-6 Measures of agreement between simulated and observed growth variables for single-treatment (S-T) and multiple-treatment (M-T) based growth-related TSE cultivar coefficient estimates for the cultivars Bragg and Williams evaluated over three locations each with multiple in-season observations (From Memic et al., (2021). Used with permission.).

-	-	Bragg						Williams					
-	-	Gainesville 1978		Gainesville 1979		Gainesville 1984		Ohio 1988		Ohio 1990		Iowa 1988	
-	-	RMSE	No. Obs	RMSE	No. Obs	RMSE	No. Obs	RMSE	No. Obs	RMSE	No. Obs	RMSE	No. Obs
S-T	Grain weight	140	7	166	8	235	7	236	17	166	18	251	10
S-T	LAI	0.41	15	0.54	17	0.39	17	0.77	17	0.96	18	0.39	9
S-T	Shoot weight	486	15	414	17	305	17	826	17	491	18	548	9
S-T	Leaf weight	143	15	179	17	99	17	305	17	253	18	207	9
M-T	Grain weight	228	7	248	8	140	7	320	17	152	18	183	10
M-T	LAI	0.47	15	0.58	17	0.51	17	0.83	17	1.02	18	0.36	9
M-T	Shoot weight	511	15	550	17	597	17	1025	17	678	18	497	9
M-T	Leaf weight	168	15	190	17	180	17	364	17	276	18	233	9
		nRMSE		nRMSE		nRMSE		nRMSE		nRMSE		nRMSE	
S-T	Grain weight	0.07	7	0.086	8	0.114	7	0.194	17	0.163	18	0.319	10
S-T	LAI	0.145	15	0.192	17	0.13	17	0.208	17	0.303	18	0.149	9
S-T	Shoot weight	0.108	15	0.102	17	0.066	17	0.14	17	0.104	18	0.13	9
S-T	Leaf weight	0.141	15	0.186	17	0.097	17	0.229	17	0.236	18	0.187	9
S-T	AVG-nRMSE	0.116		0.141		0.101		0.193		0.201		0.196	
M-T	Grain weight	0.115	7	0.129	8	0.068	7	0.264	17	0.149	18	0.232	10
M-T	LAI	0.165	15	0.204	17	0.168	17	0.225	17	0.32	18	0.134	9
M-T	Shoot weight	0.114	15	0.135	17	0.129	17	0.173	17	0.144	18	0.118	9
M-T	Leaf weight	0.166	15	0.198	17	0.175	17	0.273	17	0.257	18	0.211	9
M-T	AVG-nRMSE	0.14		0.166		0.135		0.234		0.217		0.174	

No. Obs. - Number of in-season observations used in cultivar coefficient estimation process, RMSE (kg ha⁻¹): grain weight, shoot weight and leaf weight

6.3.2   Comparison of the DSSAT Standard (manual), GLUE and TSE – Independent evaluation

The TSE and GLUE multiple-treatment based estimation for the cultivar coefficients for Bragg was conducted using the 1978, 1979, and 1984 Gainesville in-season observations (TSE only) for determining robust values for the cultivar coefficients (Table 6-7). For the cultivar Williams, three treatments including the 1988 and 1990 Ohio experiments and the 1988 Iowa experiment were used simultaneously for determining TSE and GLUE multiple-treatment based cultivar coefficients (Table 6-7). Multiple-treatment based TSE phenology- and growth-related cultivar coefficient values of Bragg and Williams were more or less similar to the *DSSAT standard* values for the cultivars Bragg and Williams (Table 6-7), with minor exceptions for the cultivar Williams for coefficients CSDL and PPSEN (phenology) and SLAVR (growth). Based on the multiple-treatment based cultivar coefficient values and comparison with the *DSSAT standard* Bragg and Williams values (Table 6-7), it can be concluded that *TSE* was able to determine "good values" for the cultivar coefficients that are comparable to those that had been solved manually by the DSSAT development team using seven or eight experimental data sets for each cultivar.

Table 6-7 Multiple-treatment based TSE and GLUE phenology- and growth-related cultivar coefficient values for Bragg and Williams for three locations each and compared with the DSSAT standard values (From Memic et al., (2021). Used with permission.).

Cultivar	Bragg			Williams		
Coefficients	DSSAT standard	GLUE: Gainesville 1978, Gainesville 1979, Gainesville 1984	TSE: Gainesville 1978 Gainesville 1979 Gainesville 1984	DSSAT standard	GLUE: Ohio 1988, Ohio 1990, Iowa 1988	TSE: Ohio 1988 Ohio 1990 Iowa 1988
CSDL (P)	12.33	12.36	12.56	13.40	12.86	12.64
PPSEN (P)	0.320	0.365	0.389	0.285	0.225	0.182
EM-FL (P)	19.5	17.44	18.86	19.0	16.03	19.19
FL-SH (P)	10.0	10.0	10.11	8.3	8.3	8.48
FL-SD (P)	15.2	19.21	17.29	14.2	19.22	12.60
SD-PM (P)	37.6	33.72	38.85	32.2	27.85	34.92
LFMAX (G)	1.00	1.000	0.995	0.99	1.171	1.013
SLAVR (G)	355.0	302.6	356.0	385.0	377.9	425.3
SIZLF (G)	170.0	187.1	162.0	180.0	228.5	166.1
WTPSD (G)	0.17	0.170	0.219	0.18	0.189	0.199
SFDUR (G)	25.0	17.00	23.0	26.0	18.80	26.77

DSSAT standard: 7 treatments for Bragg and 8 treatments for Williams, P - Phenology-related coefficients defined in Table 2, G - Growth-related coefficients defined in Table 2

The TSE multiple-treatment based cultivar coefficients that were estimated for Bragg and Williams (Table 6-7) were used for simulating four phenological events and four target traits for one independent data set for each cultivar. The multiple-treatment based TSE cultivar coefficient estimates resulted in simulations that were comparable to those of the *DSSAT standard* Bragg and Williams coefficients, thus indicating potential of using the cultivar coefficient estimator for estimating generic phenology-related coefficients. The 1990 Iowa experiment was used for evaluating the TSE-based Williams cultivar coefficients; in this experiment the first pod day and first seed day were not observed. Physiological maturity simulated with TSE Williams cultivar coefficients (Table 6-8) was more accurate when compared to the *DSSAT standard* values for Williams, based on the observed date.

The TSE AVG-nRMSE for Bragg for various coefficient combinations for the four target traits, e.g., grain weight, LAI, shoot weight and leaf weight, was 26% lower for the 1976 Gainesville experiment and for Williams it was about the same for the 1990 Iowa experiment, when compared to the *DSSAT standard* Bragg and Williams cultivar coefficient model outputs (Table 6-8). For the Gainesville simulation results with the TSE multiple-treatment coefficient 'optimums' (Table 6-7), the values for RMSE and nRMSE were lower when compared to the *DSSAT standard* Bragg coefficient simulation results for all four target traits (Table 6-8). For the Williams cultivar simulated for Iowa, the values for the four target traits were more or less the same compared to the *DSSAT standard* Williams coefficients (Table 6-8), with minor improvements in simulation of grain weight and worsening in LAI.

Table 6-8 Measures of agreement between simulated and observed phenological events as days after planting (DAP) and growth variables for DSSAT standard, GLUE and TSE cultivar coefficients when evaluated for Bragg and Williams with a treatment not included in the cultivar coefficient estimation process (From Memic et al., (2021). Used with permission.).

-	Bragg (Gainesville 1976)				Williams (Iowa 1990)			
	DSSAT Standard	GLUE	TSE	Used Obs.	DSSAT Standard	GLUE	TSE	Used Obs.
Anthesis DAP	50	49	49	ND	62	57	61	66
First pod DAP	80	86	83	ND	81	79	80	ND
First seed DAP	92	106	99	ND	94	101	89	ND
Maturity DAP	143	148	147	ND	131	135	135	139
	RMSE:				RMSE:			
Grain weight	582	254	408	7	174	269	153	9
LAI	1.39	1.03	1.11	20	0.24	0.62	0.33	9
Shoot weight	335	621	288	20	454	1112	410	9
Leaf weight	300	251	192	20	158	408	160	9
	nRMSE:				nRMSE:			
Grain weight	0.368	0.161	0.258	7	0.214	0.331	0.188	9
LAI	0.422	0.312	0.337	20	0.081	0.21	0.112	9
Shoot weight	0.075	0.138	0.064	20	0.106	0.261	0.096	9
Leaf weight	0.246	0.206	0.158	20	0.161	0.414	0.162	9
AVG-nRMSE	0.277	0.204	0.204		0.14	0.304	0.14	

Used Obs. - Number of in-season observations used for evaluation, ND - No data, RMSE (kg ha^{-1}): grain weight, shoot weight and leaf weight

The GLUE multiple-treatment based Bragg and Williams cultivar coefficients (Table 6-7) were evaluated with one independent experimental data set for Bragg (Gainesville 1976) and Williams (Iowa 1990), similar to the TSE approach. The resulting statistics (Table 6-8) show some important variation for the first seed and first pod dates, while for the onset of flowering and crop maturity the results were within an acceptable range for all three optimization approaches (Table 6-8).

Based on the resulting RMSE errors (kg ha^{-1}) for the selected four target variables, all three optimization approaches were satisfactory, with TSE resulting in lower overall errors. With the TSE approach, additional biomass variables were targeted in the optimization process such as shoot and leaf weight. The time-series graphs of grain weight and shoot weight for Bragg cultivar (Gainesville 1976) and Williams cultivar (Iowa 1990), are shown in Figure 6-5 and 6-6, respectively. Since GLUE is using an end-of-season point-based cultivar coefficient estimation approach, the grain mass over time was under-estimated during the earlier phases and more accurate towards harvest maturity (Figure 6-5a and 6-6a). A higher number of optimization combinations in GLUE resulted in more accurate grain yield simulations, at the cost of other target variable accuracy, as can be seen from shoot weight simulation outputs in Figure 6-5b and 6-6b and Table 6-8. For *DSSAT standard*, developers were aware of the grain yield and shoot weight time-series observations throughout the season and tried to balance resulting over- and under-estimation of multiple target variables throughout the season. The TSE tool is similarly using all in-season observations of multiple target variables in the estimation process along with final end-of-season variables, and thus strikes a balance during the process of estimating cultivar coefficients throughout the growing season.

a   b
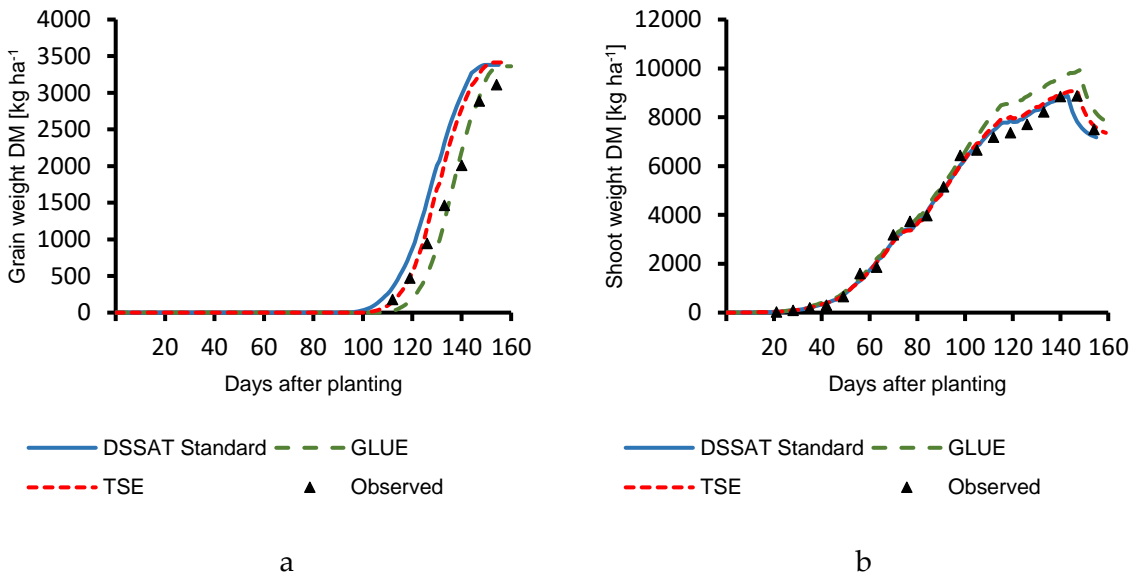
Figure 6-5 Time-series graph showing simulation output results of Bragg cultivar with observed for a) grain weight and b) shoot weight, using three different optimization approaches: DSSAT standard, GLUE and TSE. Observed corresponds to Gainesville 1976 experiment data which was not used in cultivar coefficient estimation process (From Memic et al., (2021). Used with permission.).
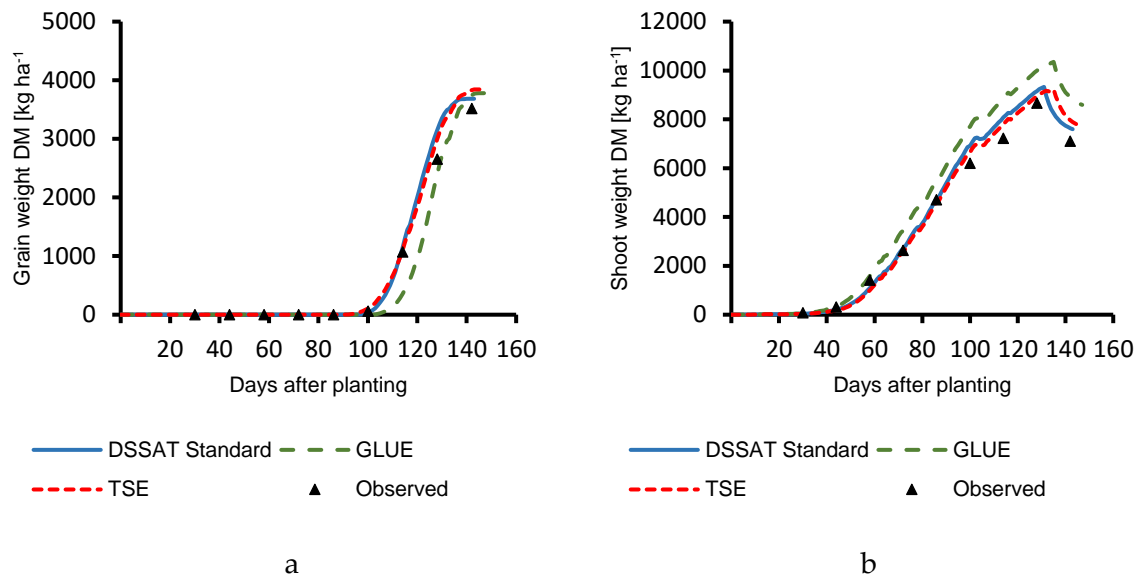
Figure 6-6 Time-series graph showing simulation output results of Williams cultivar with observed for a) grain and b) shoot weight, from three different optimization approaches: DSSAT standard, GLUE and TSE. Observed corresponds to Iowa 1990 experiment data which was not used in cultivar coefficient estimation process (From Memic et al., (2021). Used with permission.).

6.4   Discussion

Single-treatment based cultivar coefficient estimation is based on observations from one treatment of a field experiment (single season and location). With one year of in-season field observations, cultivar coefficients can be adjusted to provide a very good agreement between simulated outputs and field observations. Ideally, the estimated cultivar coefficients depicting phenological plant development and growth-related values should be applicable without any modifications across multiple locations and seasons. However, determining cultivar coefficient values based on only one specific treatment and environment may result in field- and season-specific biases being integrated into the cultivar coefficients, thus, causing underperformance of the model for different seasons for the same location or for other locations. Ideally, optimal seasonal and environmental conditions (stress-free) for multiple seasons and locations are required for estimating robust representative cultivar coefficients. Since ideal conditions are rarely met in field experiments, cultivar coefficient estimation using multiple-treatments consisting of multiple seasons and environments is a *must* in order to minimize the influence of season- and environment-based biases on cultivar specific coefficients.

Most of the output variables, such as grain yield, shoot weight, LAI, and leaf weight simulated with CROPGRO do not have constant (linear) in-season growth rates. Hence, one observation at the end of the season is not enough for evaluation of the results. Even if a single observation at the end of the growing season fits perfectly simulated, it will not guarantee realistic insight in the dynamic growth rates during the season. The influence of in-season growth rates of various variables can have an enormous influence on the above- and below-ground biomass accumulation rates, and on ratios within above-ground biomass, for example stem to leaf to grain ratios. However, these ratios will have a direct influence on the photosynthetic activity of the plant and consequently affect simulated yield at the end of the season.

Within this study a TSE tool for the estimation of cultivar coefficients was developed. The TSE tool provides accurate simulation of above-ground biomass and plant components based on commonly collected multiple in-season samples of four targeted crop variables, with emphasis on the grain yield. For estimating growth-related cultivar coefficients, a total 346 in-season observations of four growth target variables for six treatments (experiments) were used with TSE.

The multiple-treatment based TSE coefficient "optimums" shown in Table 6-7, resulted in low AVG-nRMSE for Gainesville 1978 (0.14), Gainesville 1979 (0.166) and Gainesville 1984 (1.35) (Table 6-6). As expected, the robustness of the cultivar coefficients that were solved over more than one treatment/location resulted in less accurate single treatment-based statistics when compared to single treatment-based cultivar coefficients of Bragg for the 1978, 1979, and 1984 experiments conducted in Gainesville with an AVG-nRMSE of 0.116, 0.141, and 0.101, respectively (Table 6-6). As single treatment optimization allows over-fitting of coefficients, obtained results may not apply as broadly to other years and locations. The Williams multiple-treatment coefficient "optimums" resulted in low AVG-nRMSE of 0.234, 0.217 and 0.174 for the growth variables of the 1988 and 1990 Ohio experiments and the 1988 Iowa experiment, respectively (Table 6-6). For comparison, the single-treatment based cultivar coefficients of Williams for the 1988 and 1990 Ohio experiments and the 1988 Iowa experiment had AVG-nRMSE values of 0.193, 0.201 and 0.196 respectively (Table 6-6). Multiple-treatment based cultivar coefficient 'optimums' evaluated with data not used in the coefficient estimation process showed great potential, especially when phenology related 'optimums' were evaluated for Gainesville 1976. This, along with excellent agreement for simulated and observed four phenological events, allows the conclusion that the developed TSE performed very well in solving for the genetic coefficients.

Although use of any of the three compared cultivar coefficient optimization approaches (*DSSAT standard* – manual, GLUE and TSE) can be seen as a form of success based on the goals of specific studies, there seems to be a very important advantage of using multiple in-season observations across multiple target variables for estimating cultivar coefficients that correctly reflect growth of different plant components simulated throughout entire season. The TSE tool allows users to select which target variables to prioritize in the cultivar coefficient optimization approach, depending on the collected samples. For the users that target as accurate as possible yield estimates at the end of season, the GLUE approach will very likely provide quite accurate statistics. For the users that are investigating specific management practices with the crop model where it is very important to simulate as accurately as possible specific plant component growth throughout season (including stem to leaf to grain ratios), in order to accurately evaluate the impact of those management practices, the TSE approach is recommended.

Most importantly, the TSE method of cultivar coefficients optimization based on mathematically formed selection thresholds will eliminate user-dependent bias from the process. With removal of the user bias from the cultivar coefficient estimation process, multi-model approaches can be implemented for specific groups of models within DSSAT: three wheat models (CERES, N-Wheat and Cropsim) and two maize models (CERES-Maize and CSM-IXIM), as described in Röll et al. (2020).

6.5 Conclusions

The TSE program is designed to work with in-season field observations, by minimizing differences between simulated and observed values. Single-treatment-based nRMSE error minimization showed the ability of the TSE tool to search and find cultivar coefficients. Based on the results of the multiple-treatment-based cultivar coefficient estimation, the model performed very well and provided robust cultivar coefficients. The program is written in a way to enable optimization of cultivar coefficients of all available crop models in the DSSAT shell. So far it has been tested with CERES-Maize and CERES-Wheat with satisfactory results (unpublished). It should work for many of the CROPGRO crops, especially the grain legumes, because they share common cultivar coefficients and definitions. Future work will test the program directly with CROPGRO and CERES models with a focus on using time-series plus end-of-season observations rather than only end-of-season based observations for deriving crop specific cultivar coefficients, which is a limitation of the GENCALC and GLUE methods presently in DSSAT.

The program is available in the GitHub repository (https://github.com/memicemir/TSE). It can be used without any given warranty or usage restrictions. Feedback regarding program performance or suggestions for improvement is welcome.

7 General discussion

The objective of the thesis was to investigate yield limiting (nitrogen) and reducing factors (Cercospora leaf spot disease) of specific cropping systems using the DSSAT crop growth models. Input management in agriculture can still be improved with respect to observable and quantifiable in-field variabilities with various remote- and near-sensor systems unified within crop growth decision support platforms. Better understanding of in-field variabilities is expected to result in improved management of nutrients and chemicals used for increasing yield and crop protection.

The cropping system platforms can capture genetic×environment×management interactions with respect to spatial and temporal in-field variabilities. This is of extreme importance when it comes to management of agricultural inputs used in production and use of the remote- or near-sensor measurements, because nutrient and pest/disease management can impair interpretation of sensor-based measurements based on which certain decisions are made. In the case of nutrient management sensor-based readings can help to the extent to which specific bio-physiological traits of plants are considered. Sensor readings that rely on the visible spectrum for quantifying N status are useful but plant development stage, pest and leaf disease status have to be considered too, in order to lead to informed management of agricultural inputs. Every field in itself is a unique environment and sensor-based readings can help improving management of agricultural inputs based on local field status with respect to the soil properties and weather affecting in-field microclimates favouring pest and leaf disease development.

Thus, the true potential of input management depends on integration of the site-specific knowledge about environment and weather with respect to management practices complimented with various sensor systems when combined within a decision support platform.

7.1 Site-specific marginal net return maximising N application rates

Among other very important macro nutrients required for plant growth, N is the one that has the most influence on yield and is required in highest quantities in production of cereal crops (Fowler et al., 2013). It has relatively low costs and very low use efficiency based on the fertilized amounts and quantities taken by the plant (Martínez-Dalmau et al., 2021) and as such causes major problems in environmental pollution and underlying economics of profitable agricultural inputs management (Alotaibi et al., 2018; Wang et al., 2020). When it comes to the N based fertiliser (organic and inorganic) used in crop production in EU, the balance of inputs and outputs is in favour of inputs, which means that fertiliser used in production is not entirely utilised by the crop but lost to the environment (Sutton et al., 2011).

Based on the studies conducted by Thorp and Link with APPOLO and the NPM approach demonstrated in chapter 4, yield maximisation without detailed analysis of underlying economics of N application (N price) and benefits (yield increase), N use efficiency (yield increase in kg ha^{-1} per N kg ha^{-1} applied) cannot be properly managed. As shown in the study conducted in chapter 4, there is a major difference between two seemingly intertwined concepts: bio-physiological yield maximisation and marginal net return maximising yield, especially in the context of low-yielding (Germany, Riech) and high-yielding (USA, McGarvay) fields. The objective was to investigate the potential of variable N application rates with respect to maize and N prices, and effects of the plant population densities on yield and

marginal net return. The results indicated that for the low-yielding fields better management of N application has more influence on the marginal net return than additional N that cannot produce additional yield to cover or justify additional costs in the production (chapter 4). A similar trend was observed with plant population rates.

A similar study was conducted focussing on the economically optimal N amounts in maize Northeast China with two different soil types. The study evaluated the effects of weather and management practices on variable N application rates (Wang et al., 2020). The authors tested maize production with different planting densities and different N application rates focusing on economic optimums. Wang et al. (2020) had similar conclusions related to the soil type as Link et al. (2013). In a sandy soil, variable N application showed larger variability in economic optimum N with respect to the weather conditions, when compared to a black soil (Wang et al., 2020). Within their study, based on the two different soil types they concluded that a soil type-specific N application rate might lead to improvement of nitrogen use efficiency and consequently might positively affect marginal net return (Wang et al., 2020) in soils that are more prone to N loss. In a study conducted at a location in Canada over 12 years (Alotaibi et al., 2018) N application rates led to optimized corn yields when the calculation was based on the soil texture.

In general, knowledge on fertiliser applications are a result of in nature static statistical analysis. These statistics were used to cope with a problem that is truly dynamic and a result of specific in-season dynamics of the plant nutrient uptake, soil water balance (in mostly heterogeneous soils) and weather related factors influencing the efficiency of the used fertiliser (Martínez-Dalmau et al., 2021). To generate this knowledge, different amounts of fertiliser were applied in different plots and the effects of those different rates on final crop yield were analysed in order to derive "optimum" fertiliser application amounts, maximising yield or marginal net return. These analyses were conducted with various regression models that were not able to account for or give any insight into temporal variabilities observed within season that have a major effect on the fertiliser utilisation or losses. With crop growth models a huge potential for sensitivity analysis of the given spatial and temporal variability exists and can help to reveal the underlying complexity, finally leading to more accurate fertiliser applications.

As more sensors became available in the agricultural sector over the last 20-30 years, the potential to quantify yield variability within field increased. Hence, the solution for timely site-specific N applications seemed to be reachable, solvable and implementable in practice with relatively low costs (Sishodia et al., 2020). The first problem faced, when considering site-specific N application, consists of defining the size of the site-specific management sub-units (Link et al., 2006a). Since the goal of site-specific N application is often the minimisation of the given in-field yield variability, the most direct solution seemed to be to delineate site-specific management units based on the observed yield variability within a field. The yield variability is relatively easy quantifiable with new sensor technologies, and is a good starting point in understanding yield limiting factors causing this variability. Yield limiting factors can roughly be split in two groups: long-term soil properties and yield response to the specific agricultural production inputs (Maestrini and Basso, 2018). The yield limiting factors causing low- and high-yielding zones in the field potentially can be considered as those that vary from year to year (previous crop residues in the field) and long-term factors such as soil type and texture (Maestrini and Basso, 2018). If yield variability (high/low yielding zones) is stable over multiple years (Maestrini and Basso, 2018) it offers the possibility to be used for long-term site-

specific management zone delineation. If yield variability within a field varies from year to year it has to be treated differently.

7.2 Decision support system for evaluating the impact of observed leaf disease damage on sugar beet yield

Coupling of the crop growth model state variables (about plant canopy) with suitable sensor readings (about plant canopy) can potentially lead to decision support tools for managing leaf disease outbursts in the field (Thorp et al., 2010; Röll et al., 2019; Batchelor et al., 2020). Various aspects of crop growth (such as plant organ partitioning) and leaf disease effects on different plant organs can be investigated in detail (Batchelor et al., 1993). With quantification of the yield reducing factors and leaf disease impact, extensive economic analysis can be conducted. With CSM-CERES-Beet and model based genetic specifications of the cultivar used in the experiment, sugar beet dry matter losses and consequently extractable sugar losses were simulated and evaluated with observed data. Simulation of the leaf disease effects in the context of sugar beet development is complex. Leaf disease such as Cercospora leaf spot disease and occurring necrosis spots cannot be entirely excluded in their impact on biomass data used for evaluating disease effects through tops weight dry matter losses per defined field unit. Different problems can occur in the process of detecting and quantifying various pathogens causing leaf diseases (Hillnhütter et al., 2011; Zhang et al., 2014) due to mainly overlapping of the leaves and spread from plant bottom to top (Cao et al., 2015).

Leaf disease impact on sugar beet canopy development was demonstrated with manually measured data with corresponding impacts on tops weight (y0 axis) and storage root (y1axis) (chapter 5, Figure 4-6). Detailed demonstration of the measured data and respected effects on 100 % fungicide treatment (2016) are shown in chapter 5, Figure 4-6. Linearly interpolated daily damage rates are shown in chapter 5, Figure 4-6 as straight (red) lines with arrows between observed disease points (0%, 1%, 20%, 24% and 33%). If detailed information on leaf disease progress would be available, a non-linear interpolation could be introduced in the future. However, the current concept based on manually collected data is a good starting point for developing decision support tools able to integrate sensor-based disease ratings.

Figure 7-1 shows how the concept for a DSS can be combined with sensor-based data. For quantifying leaf disease effects on tops weight two scenarios have to be formed: a baseline and a disease scenario. In a first step the crop model is executed without having disease ratings included in the simulation process (baseline scenario), assuming e.g. no disease is present in the field. In the second step the model is executed with disease ratings included in the simulation process (disease scenario). The difference between the baseline and the disease scenario enables the quantification of leaf disease effects on tops weight, and consequently on storage root yield. The storage root yield to tops weight ratio is defined externally through cultivar coefficients and the underlying dry matter partitioning assumptions among different plant organs.

At the time the disease is observed in the field (Figure 7-1, point 1) the model can simulate only a baseline (simulating tops weight without disease information). At this point the user has simulated 'potential' tops weight, observed disease and observed tops weight (observed tops weight affected by disease already). By introducing diseases into the model and rerunning the model, the user can quantify effects of the disease on tops weight by subtracting the disease scenario from the baseline (tops weight loss, Figure 7-1, phase II). Tops weight observed at the moment of measuring the disease in the field is point 3 in Figure 7-1. In order

to introduce disease effects on tops weight on a daily basis (Figure 7-1, point 2-3) a starting point has to be selected (no disease in the field) to enable the use of the proposed interpolation method that enables introduction of the disease and disease damage on a daily basis (Figure 7-1, point 2-3).

In the next phase (III) (Figure 7-1, point 4), hypothetical projections, of future potential disease development under the given disease pressure and weather conditions can be used for hypothetical yield loss estimates from the moment of the observation (Figure 7-1, point 3-4).
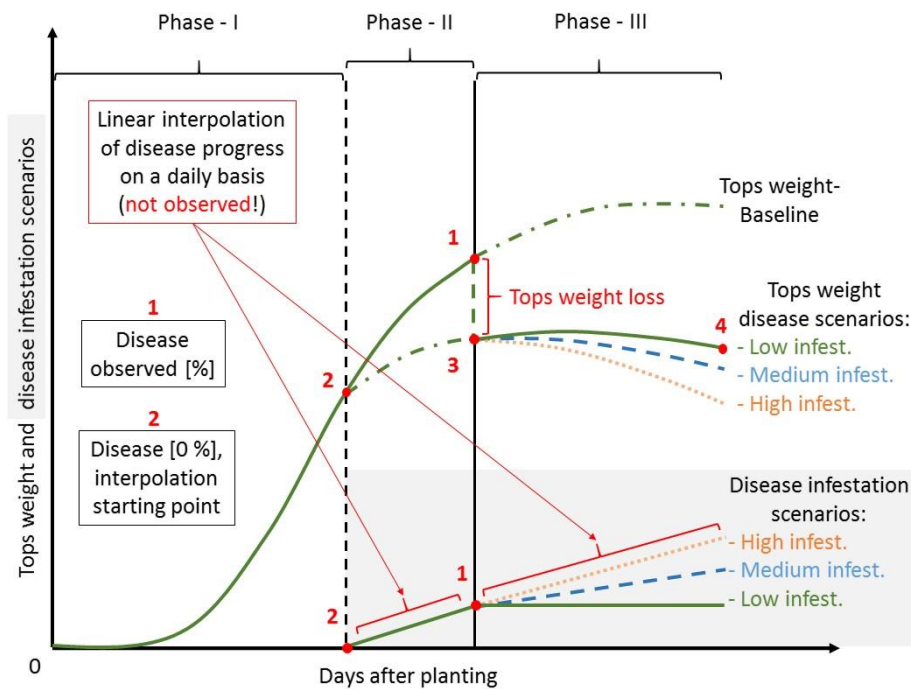
Figure 7-1 Own illustration, baseline and disease scenario example for tops weight – tops weight loss and creation of hypothetical disease treatments (own illustration).

7.3   Impact of cultivar coefficients on simulated yield

The cultivar coefficient estimation process is extremely complex and should be conducted with extreme care. As representative cultivar coefficients have to be estimated, the data being used in the process of deriving these coefficients has to be representative of plant behaviour in phenology and growth (Hunt et al., 1993; Guillaume et al., 2011; Wallach et al., 2011). As already discussed to a certain extent in chapter 6 the objective of the crop modelling study should decide on the needed approach in the cultivar coefficient estimation process: use of end-of-season grain yield or use of multiple in-season observations of multiple target variables (e.g. LAI, tops weight etc.). If the user has only few phenology-related observations (onset of flowering, physiological maturity etc.) and grain yield, the choice is simple based on the availability of the data. If the user possesses in-season observations of multiple target variables they should be used as they are an important indicator of the crop in-season growth dynamics. In a study on the comparison of cultivar coefficient estimation procedures conducted by Guillaume et al. (2011) the authors found that the use of in-season observations of multiple target variables resulted in worsening of end-of-season estimates of grain yield and grain N concentration, compared to the approach, which was based on estimating cultivar coefficients

only based on end-of-season grain yield. Guillaume et al. (2011) concluded to a certain extent that if the goal of a crop modelling study is the overall accuracy improvement of end-of-season variables, the use of information about in-season dynamics is not "useful". Wheather the overall crop model performace (entier seasonal crop growth) is to be subjected to end-of-season mathematical fit of grain yield and grain N concentration, as in the case of Guillaume et al. (2011), is up to the user. Guillaume et al. (2011) also indicated that the use of in-season information for estimating cultivar coefficients can improve the overall crop model performance (more accurate depiction of plant growth from planting to harvest) and indicate what parts of the model under-perform, based on in-season dynamics. However, the user should be aware of the fact that cultivar coefficients estimated with only end-of-season observations might not perform well in other environments, as they are not able to replicate plant-specific in-seasonal behavior that depends on certain environmental apsects, and susequently will give wrong end-of-season yield estimates, or "correct", for wrong reasons.

As explained in chapter 6, the importance of using more than just end-of-season crop observations should be clearer. A larger amount of available data documenting the within-season crop×environment×management practice interactions, could lead to more representative (robust) predictions. The predicted yield and management practice based on some of the within-season plant behaviour will enable better insight into the within-season dynamics that have a major influence on the yield and above-ground biomass predictions. The influence of the cultivar coefficient "accuracy" on N balance with the respect to the predicated above-ground biomass and yield is shown in Table 7-1. For this example, an independent data set shown in chapter 6 was used. This data set was used for direct comparison of the three different cultivar coefficient estimation approaches: DSSAT Standard (manual), GLUE and TSE (chapter 6, Figure 6-5). Different cultivar coefficient estimates resulted in different shoot weight and yield estimates and corresponding N balances as can be seen in Table 7-1. As it can be seen from this example (Table 7-1) the N balance has to be interpreted based on the intensity of N in the system. In this example production is not relying on the additional inorganic N fertilizer. The plant growth and accumulated biomass has specific N requirements that have to be met in order to produce final grain yield. Different cultivar coefficient values will affect the amount of simulated biomass and plant N uptake from soil (Table 7-1). Based on the plant N uptake (in kg ha^{-1}) DSSAT Standard and TSE indicated lower rates when compared to GLUE. If additional inorganic N was given, these different cultivar coefficients would result in different N recommendations and N use efficiencies. Since DSSAT Standard and TSE were relying on the in-season biomass accumulations rates for deriving cultivar coefficients, it is very likely that they represent better insight in the plant N uptake requirements and thus in the total N balance.

Table 7-1 Detailed soil N-balance with respect to simulated biomass with cultivar coefficients estimated with three different approaches (tools).

	DSSAT Standard		GLUE		TSE	
	Initial	Final	Initial	Final	Initial	Final
Unit	kg N ha^{-1}					
Soil NO$_3$	22.27	54.76	22.27	46.06	22.27	57.11
Soil NH$_4$	19.31	3.60	19.31	3.61	19.31	3.47
Soil Urea	0.00	0.00	0.00	0.00	0.00	0.00
Fertilizer N	0.00		0.00		0.00	
Mineralized N	178.50		186.90		183.14	
Leached NO$_3$		14.72		14.14		14.76
N Denitrified		0.25		0.32		0.29
N Uptake from soil		141.86		159.46		144.20
Ammonia volatilization		0.00		0.00		0.00
N Immobilized		4.90		4.90		4.90
Total N Balance	220.09	220.09	228.49	228.49	224.72	224.72

Initial – soil N at start of the simulation, Final – soil N at harvest

However, a critical issue in the cultivar coefficient estimation is the initial cultivar coefficient range setup as discussed in chapter 6 with respect to other tools and literature. If the user is defining for a specific cultivar coefficient a lower and upper range limit, within which sensitivity analysis is conducted in order to select the best value based on the AVG-nRMSE, as is the case for TSE, these lower and upper range limits will have a direct influence on the cultivar coefficient values in cases where estimated values are getting closer to the lower or upper range limit. If the lower and upper range limit is set based on the physiological traits of the plant, they should not be crossed. For an example, if the cultivar coefficient defining leaf size for a specific crop is defined based on in-field observed minimum and maximum leaf size, the simulation of abnormal leaf sizes in order to get better statistical fit of predicted end-of-season yield, is definitely not recommended. The cultivar coefficient estimation process consists of varying the values for each cultivar coefficient and comparing a statistical fit of simulated outputs with field observations in order to determine the coefficient combination providing the best agreement between simulated and observed values. Various cultivar coefficients have potentially wide ranges (minimum and maximum values difference) with many in-between values that depend on the increment step size (Inc). The so called *Exhaustive gridding - coefficient variation* (Table 7-2) (Röll et al., 2020), can be used to systematically investigate coefficient ranges in search for coefficient values that provide the best statistical fit. For example, P5 coefficient value for minimum 100, maximum 900 and increment step 3.3 can be passed into the cultivar file and the model will be executed for each coefficient value. In this example as shown in Table 7-2 it can be seen that for a coefficient range from 100 to 900 with increment steps of 3.3 a total of 243 coefficient variations are executed for the P5 coefficient with Exhaustive gridding method. In order to overcome time losses in the process of cultivar coefficients estimation based on the statistical fit (lowest nRMSE), a range reduction method was implemented (Table 7-2). With this range reduction method four global phases are conducted in the process of estimating cultivar coefficients with smallest average nRMSE. Greater increment steps are used in the first phase for each given coefficient range with P5 coefficient having minimum 100, maximum 900 and increment step 200. Based on the lowest

nRMSE, the value for each coefficient is selected, i.e. 300. In the second phase, the new coefficient ranges with a narrower increment step are executed with P5 having a minimum of 180, maximum 420 and increment steps of 60. Based on the lowest nRMSE, the new coefficient "optimum" is selected, i.e. P5=300. In the third phase new ranges for each coefficient are defined with P5 having a minimum of 270, a maximum of 330 and increment steps of 15. In the final phase the P5 minimum 263.2 and the maximum 276.7 with increment steps of 3.3 are passed into the cultivar file and based on the lowest nRMSE the value P5=263.2 is selected, in the process of determining local minimum (Duan et al., 1992). Based on the range reduction approach, 48% fewer combinations are executed when compared to exhaustive gridding coefficient variations. The range reduction method as described in Table 7-2 is expected to retain a systematic optimum localisation approach (achievable with exhaustive gridding variation), and is expected to provide more realistic values for the coefficients when compared to the random generation of cultivar coefficients for allowed ranges. The range reduction is flexible programmed and will work with different scaled coefficients.

Table 7-2 Exhaustive gridding and range reduction methods example with P5 coefficient.

Exhaustive gridding		Range reduction method			
P5		Phase 1	Phase 2	Phase 3	Phase 4
Min	100	100	180	270	263.2
Max	900	900	420	330	267.7
Inc. step	**+3.3**	**+200**	**+60**	**+15**	**+3.3**
	100	100	180	270	263.2 →263.2
	103.3	300 →300	240	285	266.5
	…	500	300 →300	300	269.9
	→**263.2**	700	360	315	273.3
	…	900	420	330 →270	276.7
	900				
No. comb.	243	5	5	5	5
Total	**243**	**20**			

108

The range reduction method was implemented in addition to the method described in chapter 6 and is already available in the software solution shared on GitHub, with examples described in the user guidelines.

7.4 Crop model-based analysis of yield limiting and reducing factors: potential and limitations

It was discovered within the study (see chapter 4) that the crop-model based tool can be used for analysing common N management practices. It was found that soil properties and weather play a major role in correctly capturing and interpreting in-season N dynamics (Thorp et al., 2006). Different studies by Link at al. were conducted for south-west Germany (Link et al., 2006a; Link et al., 2006b; Link et al., 2008) indicating benefits of using, at the time APOLLO, or similar model-based approaches for optimising N fertilisation management. The study conducted with APOLLO indicated that model-based site-specific N optimisation has potential and the focus should be kept on soil category and water holding capacity in the context of N management (Link et al., 2013). The N management optimisation should be conducted with respect to positive social externalities (reduction of the ground water pollution) and economic benefits (Link et al., 2006b). It has to be pointed out that a major restriction of the APPOLO and NPM approaches is the underlying CERES-Maize v3.7 used for optimising soil profile-related parameters, which is no longer updated. The soil profile optimisation in the soil and atmosphere model (of the DSSAT v3.7) were not carried over further in following versions of the DSSAT model. Currently we are developing a new approach based on the method and concept of the developed tool in chapter 4 for optimising soil profile-related parameters based on the generic soil profile setup available in DSSAT 4.7 and future versions.

When it comes to yield reduction due to leaf disease damage, regardless of how good in-field disease estimates are at a certain point, the future weather forecasts that plays a major role in disease development and spread is an important obstacle. Currently, the biggest challenge in developing a functional decision support tool is the improvement of the hypothetical disease projections in the context of leaf disease favourable weather conditions. Leaf disease management relies on the predictability of weather conditions favouring leaf disease development, and as such is in the focus of developing hypothetical disease progress curves. Spatial disease management and spatial spread of spores in the field, could only be considered implicitly within the observed leaf disease damage ratings in this thesis, as disease submodules capable of simulating the spread or accumulation of spores based on leaf disease favourable conditions over time are not available in DSSAT. In the DSSAT 4.7 there are two tools that can be used for analyses and prediction of weather conditions: Weather Analogue and Weatherman. The Weather Analogue and WeatherMan tools are tested, but these tools are not suitable for precision agricultural applications without proper sensitivity analysis. Overall it is quite difficult to obtain realistic estimates of future weather. But these two tools can at least enable the user to create different weather scenarios based on historical weather data (Weather Analogue) and stochastic estimates (Weatherman) for sensitivity analysis of leaf disease progress. In addition, a basic concept of a Leaf Disease Pressure Factor (LDPF) is being developed. This pressure factor will be used in the DSS for determining the slope of the hypothetical yield losses. LDPF values consist of leaf disease favouring weather factors

(minimalistic approach consist of: humidity, minimum and maximum temperatures) and a cultivar susceptibility factor to the corresponding disease.

It has to be pointed out that cropping systems are not a perfect representation of the natural plant growth as they are mathematical abstractions of extremely complex systems. To which degree the bio-physiological development of a plant is accurately simulated depends in many cases on the data used for determining plant phenology and biomass accumulation as a function of the entire environment such as solar radiation, temperatures, water, soil properties etc. As demonstrated in chapter 6, certain mathematical abstractions such as cultivar coefficients can play a major role in "accurate" prediction of the plant phenology and growth. The users have to be fully aware of that fact that estimated cultivar coefficients based on specific in-season biomass observations and cultivar coefficient range limits are derived based on these two underlying conditions. If there is user-bias in the sampled data or unrealistic coefficient range limits, the derived cultivar coefficient estimates will integrate those biases in the cultivar coefficient values, which will make them less representative over different seasons and locations.

7.5 Remote- and near-sensing data integration into the crop model – conceptual solution based on the method and software solution described in chapter 6

Optical sensing technologies are used for evaluation of the plant N status (Antille et al., 2018). These techniques are based on the assumption that plant tissue compounds such as chlorophyll are an important indicator of N status and as such can be used as an indicator of the plant N uptake. The optical sensing is based on common principles of: multispectral reflectance, multispectral transmittance, chlorophyll fluorescence and hyperspectral reflectance (Antille et al., 2018). For macronutrients used in agricultural production, such as N, relatively good correlation has been found between plant canopy N status and accumulated plant biomass (Van Maarschalkerweerd and Husted, 2015). In situations where chlorophyll synthesis is not impaired by other factors indeed plant nutrient status can be assessed to a certain extent with indirect measurement of organic compounds with optical sensors (Tremblay et al., 2010) and can result in informed N application decision. The problem occurs in situations when available soil N is high but chlorophyll synthesis is impaired by other soil-related properties or leaf damage. In addition to the already mentioned "questionable" assumption that plant N status is indeed a valuable indicator of available soil N, depending on strength of correlation of plant N status and available soil N, measurement of N status in different crop growth phases and corresponding growth phase observable biomass plays a major role in measurement accuracy. The interpretation of the optical sensor-based N status readings can be improved to a certain extent if fully fertilised in-field reference spots are used for adjustment, under some soil homogeneity and pest/leaf disease free assumptions.

The crop model platforms offer a unique opportunity to include non-destructive information about in-season crop development with respect to time and spatial dimension collected with drone or satellites (Seidl et al., 2004; Sishodia et al., 2020) or any other sensor capable measuring in-season N-status and biomass progress. In-season biomass accumulation progress and LAI data (as time-series) can be used based on site-specific grid delineation with the tools developed within chapter 6 for improving specific inputs based on the measurements on a weekly basis throughout the growing season. Different studies showed that satellite images based indices such as NDVI can provide usable information about in-season biomass progress (Haboudane et al., 2004; Xie et al., 2016). The complex dynamic interaction of the

yield variability underlying factors cannot be solved only through intensive yield and/or soil monitoring sensors. A linkage between the available data collected by sensors and crop growth models, will help to put the data in an overall needed context of plant-soil-weather relations.

In order to show a conceptual design for integrating sensor-based in-season measurements into the crop model a simplified flow diagram is shown in Figure 7-2. In order to setup crop mode inputs (Figure 7-2, step 1.) already mentioned input data are required: experiment data, soil characterisation, genetics and measured daily solar radiation, temperature (min and max) and precipitation. Based on the inputs setup, the crop growth model simulates various aspects of plant growth on a daily basis as shown in Figure 7-2 step 2 for grain weight (kg ha^{-1}). The simulated grain weight is a form of post-processing analysis in which all required data was collected and used for simulation of grain weight based on the daily weather parameters. Ideally in-season or end-of-season (harvested yield) measurements of grain weight (in this example) are used to evaluate the accuracy of estimated yield. Based on the simplified analysis shown in Figure 7-2 through steps 1-3 sensitivity analysis can be conducted in order to check if the production management strategies can be improved by varying either planting date, fertilisation amounts, planting density, harvesting date etc. Commonly crop growth models have been used in the past for this form of post-production sensitivity analysis.

Because of the available sensors that can measure specific aspects of plant in-season growth, the post-production concept can be improved and adapted in a form of a real-time decision support system. If in the specific case a field-related crop model setup is ready, in-season measurements (indicators) of biomass progress can be included in the model. This will help to evaluate model performance based on the measurement with respect to the measured daily weather parameters to evaluate the efficiency of for example N application rate, especially if a split N application strategy is used. Sensor based data about in-season biomass progress can be included into the system to replace commonly destructive biomass samples included in the system shown in Figure 7-2, step 2. For real-time setup other state variables can be used instead of grain weight as shown in Figure 7-2. Crop model state variables such as LAI, above-ground biomass are important indicators of crop in-season development and efficiency of fertilisation strategy based on the weather and soil dynamics. Since crop growth models simulate various aspects of the crop above-ground biomass accumulation and N content on a daily basis, sensor-based measurement of either biomass or N-status can be used to evaluate the simulation results with respects to soil and weather dynamics. Based on the model simulation of the in-season N balance it can be determined whether additional N should be applied or not. It has to be pointed out that the real-time crop model application is dependent on weather forecasts. Without weather forecasts future projections cannot be formed, as in the case of leaf disease, since the crop model cannot be run without weather-related parameters. Different weather scenarios can be used, based on historical data or stochastic estimates, to conduct sensitivity analysis based on which management practice can be adjusted.

Either sensors mounted on agricultural vehicles or satellite/airborne can be used for deriving in-season biomass (canopy) information as shown in Figure 7-2, step 4. Theoretically specific genetic aspects of crop model inputs can be re-calibrated in order to improve simulated crop-genotype-environment dynamics (simplified ratios of leaf to stem etc.), or soil properties that vary in short term, based on the observed biomass. The idea is to re-calibrate specific crop model inputs and re-run the model to better reflect in-season growth with respect to measured weather and measured above-ground biomass. The method and the tool described in chapter 6 offer a perfect opportunity for these tasks as it can optimise crop model input parameters

based on multiple in-season observations of multiple target variables (different plant organs such as stem, leaf etc.). The time-series estimation tool (TSE) was used in chapter 6 for estimating cultivar coefficients, but it can be used also for estimating soil-related properties such as site-specific root distribution factor, soil organic matter etc. as they all influence in our example the mentioned N balance. Beside re-calibration other approaches are available for integrating in-season measurements in the model, such as force fit and update of specific crop model state variables. Force fit and update of state variable such as LAI do not aid in improving and resolving underlying causes for inaccurate estimates of LAI in this example. Force fit and update relays on over-writing crop model estimates independent of in-season genetic×environment×management dynamics, and as such are not recommended.
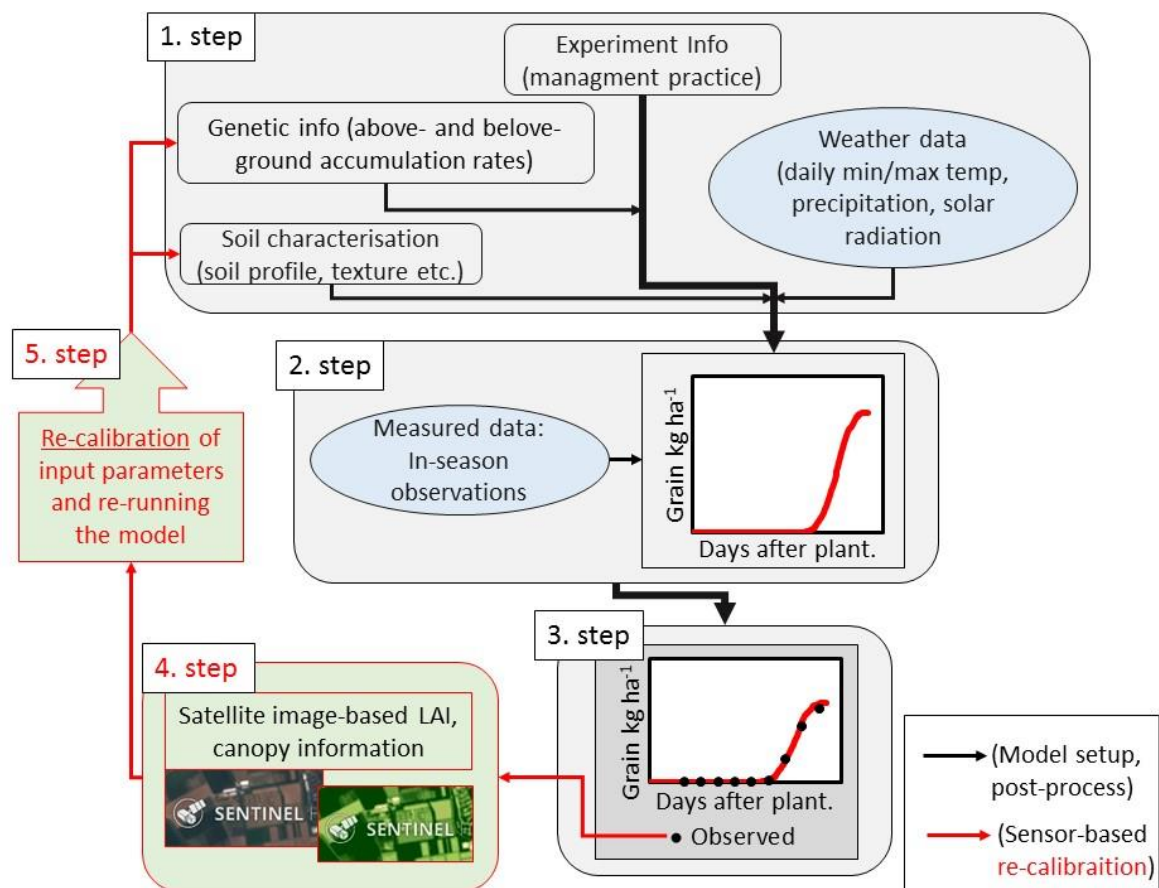
Figure 7-2 Crop model-based post processing analysis (step 1-3) with conceptual design of integrating in-season sensor-based in-season observations (step 4-5) based on re-calibration approach (own illustration).

Figure 7-3 shows a more detailed example of the re-calibration process for winter wheat with respect to LAI, total soil N and N uptake. Because winter wheat is commonly fertilised three times (split N application) in order to maximise N uptake with regard to biomass and protein accumulation, a range of different more meaningful sensitivity analysis can be conducted. In Figure 7-3 the example is based on LAI, N uptake and total N available in the soil. For this example, LAI (leaf area per area unit) was used as an indicator of the in-season plant development based on which N management is evaluated. Total soil N in this case is used as an indicator of the available N in the soil for plant uptake in various forms (ammonium, nitrate, urea etc.), and plant N uptake for connecting soil N with biomass accumulation indirectly based on LAI. The example for winter wheat shown in Figure 7-3 is a simplified

form of a sensitivity analysis based on two different N application rates with three split N applications each, where full lines for LAI, total soil N and N uptake represent the treatment with a total of 240 kg N ha^{-1} (100+100+40). The doted lines represent the treatment with 80 kg N ha^{-1} (30+30+20). Split N application rate dates are the same for both N application rates. Based on Figure 7-3 it can be seen that different N application rates will result in different LAI. The re-calibration concept can be implemented in a way when measured in-season LAI is higher/lower than the one simulated in the crop model, based on soil N and N uptake. Later N application rates can be adjusted in order to prevent over/under fertilisation. The soil-plant-atmosphere dynamics shown in Figure 7-3 are over-simplified in order to demonstrate potential practical application.
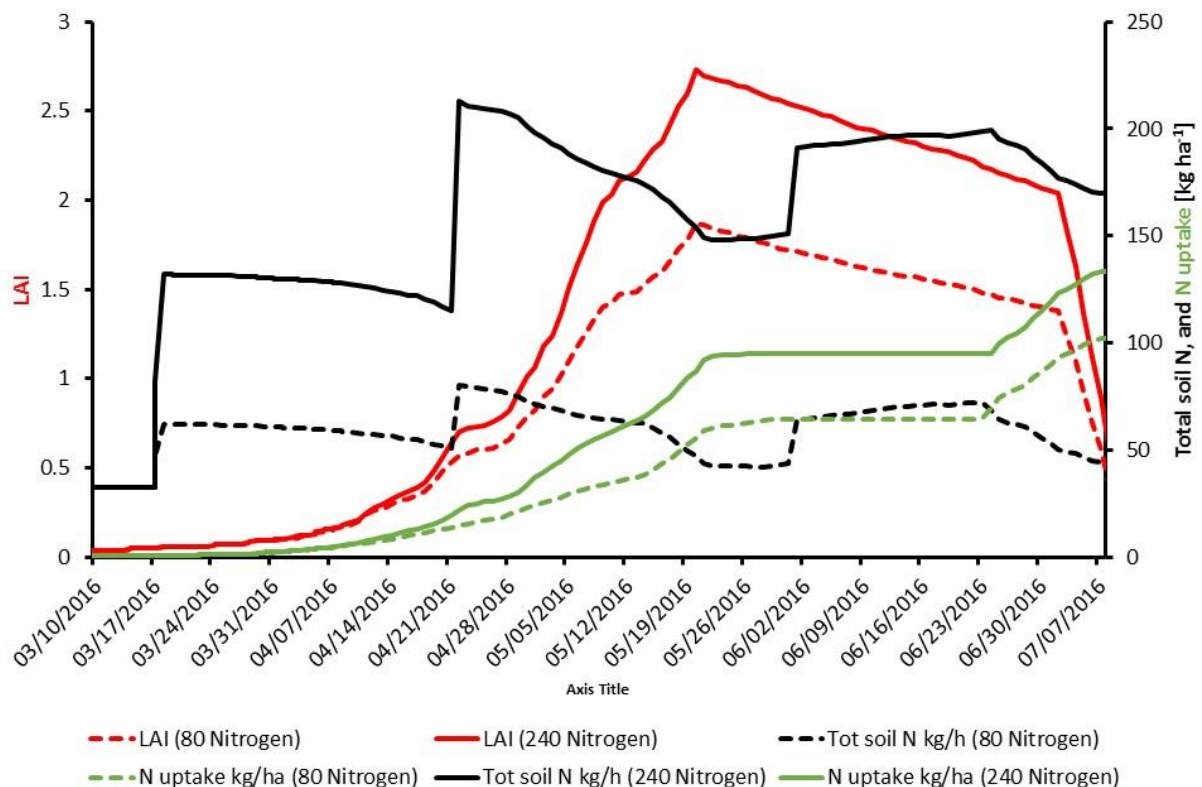
Figure 7-3 The in-season interaction of LAI, total soil N and N uptake (kg ha^{-1}) in winter wheat for two different split N application rates: 240 (100+100+40) and 80 (30+30+20) kg ha^{-1}.

7.6   Critical review of the thesis

The greatest question that might come to one's mind at this point in time: "*why are we not already there, and what is the problem when it comes to practical implementation of scientifically meaningful concepts, since farm digitalisation started basically with the first form of electric devices and most of the sensor technology and methodology usable in agricultural production is already available in some form for 20-30 years?*".

Within this thesis crop model-based analysis was conducted in order to investigate the potential yield and consequently factors causing yield gap. As already mentioned while reviewing the yield gap concept, the main difficulty is the hierarchy of three important

dimensions of agricultural production and yield maximisation: bio-physiological, economic and socioeconomic. As any other economic activity agricultural production is subjected to economically feasible profit making with respect to positive and negative externalities for society and environment. Different legal regulations are in place to limit the use of specific nutrients and chemicals in agricultural production for balancing profitable management of agricultural producers and pollution caused by them. Because of this, the problems faced by agricultural producers are even more complex than only uncertainties that rise from crop biological and environmental constraints affecting yield maximisation.

When it comes to crop growth model based potential yield and yield gap analysis, or in similar sense meta-analysis or statistical analysis, the greatest challenge is the input data used for evaluation of specific agricultural production frameworks used for future projections. Different approaches require a different amount of input data resulting in different complexity levels. Too detailed analysis of specific processes may cause over-specialised solutions that might not be successful when applied on a larger scale. It is of great importance to strike the balance between input data requirement and practical benefits of different analysis approaches. Even though crop growth models require a relatively large quantity of input data for depicting bio-physiological and physical processes relevant for simulating plant growth, it has to be pointed out that commonly collected data for these purposes mostly reflect above-ground biomass growth and not below-ground biomass accumulation and growth. Physical processes occurring in the soil with respect to soil water dynamics and nutrient transfer play major roles in defining above-ground biomass growth. Data collection of below-ground plant traits is time consuming and expensive and as such neglected to certain extent. On top of all, data collection is always impaired implicitly by inaccuracies due to specific data collection biases based on different levels of expertise used in the process. Even though great research has been conducted so far, certain physiological and physical processes of plant-environment interaction are not clearly defined and as such cannot be correctly integrated into crop growth models. Commonly experimental designs used for crop model development consist of calibration data and evaluation. Rarely the models are validated through entirely independent data for validating phenology- and growth-related plant behaviour in different climates, continents or weather extremes (under stress). Especially in weather extremes crop growth models do not perform very well as they are mostly calibrated and evaluated under water and nutrient stress-free data. For crop growth models adjusted through stress-free conditions to perform well in weather extreme conditions, the great importance is to be found in soil and biomass dynamics, which are not easy to depict due to complexities that arise in relation to soil texture and water holding capacity (soil water lower limit, soil water upper limit and soil water saturation rate).

The selection criteria of crop case studies (maize, sugar beet and soybean) used in this thesis can raise various practical questions but at the same time gives insight in already spoken input data requirements. The main technical and practical background of the case studies was discussed throughout the general discussion with respect to relevant scientific studies. The main underlying imperative behind using maize and sugar beet crops lays in economic and environmental positive and negative externalities and their implications on further improvement of cropping systems with respect to profitable agricultural production in studied regions. In order to demonstrate the importance of time-series in-season observations used for deriving cultivar coefficients for crop growth models, a huge amount of data is required.

Ideally data sets should be collected and representative for the simulated regions in a more consistent way in future studies.

The current conventional farming approach consisting of uniform agricultural input management has to be further extended to site-specific management, where possible. Site-specific field management is not a form of solution where one approach fits all problems. Sub-field zones designated into homogeneously manageable site-specific units raises various difficulties in implementation. As shown in chapter 4, based on the field and crop model setup specifically for maize, meaningful site-specific sizes can be derived based on additional ground truth data with respect to economics related to grain and N prices. Based on the study presented in chapter 4, the Precision Agriculture concept is no longer based only on agronomic optimums. The given practices can only be further improved by an underlying economic analysis. The study shown in chapter 4 is a simplified example based on simple marginal net return analysis with grain and N prices. Even though the study revealed some very interesting aspects of site-specific management of high- and low-yielding fields and their effects on marginal net return maximising N application rates, it did not include various additional costs involved in crop management and as such is still a scientific and not a practical recommendation.

The study of chapter 5, where leaf disease effects on sugar beet storage root dry matter loss and extractable sugar yield were investigated, has to be treated in the same sense. It is more of a scientific study rather than a practical recommendation. With all technological advancement and sensors available there are still uncertainties related to leaf disease quantification that prevent practical application. Indeed, a crop model with a leaf disease subroutine can to a certain extent give practical insight into potential yield losses, but leaf disease spread and progress under leaf disease favourable weather conditions are major obstacles in a more practical approach. Both N and leaf disease management are impaired by future weather forecasts. Without weather forecasts these kinds of daily-based crop model analysis cannot be conducted. Already mentioned weather extrapolation methods can be used for sensitivity analysis as long as users are aware of extrapolation method limitations.

As shown in chapter 6 crop growth models and specific approaches used in their underlying programming solutions and practical approaches still leave room for improvement. Overall the results of all chapters together revealed great potential for crop models and future precision agriculture applications but also important limitations that have to be kept in mind and further investigated.

7.7 Outlook and future prospects

The awareness of the spatial and temporal variability in crop production is known and spatial and temporal agricultural inputs management was advanced over the last 20-30 years within conceptual frameworks such as Precision Agriculture. Throughout history conventional crop management defined as applying uniformly nutrients, pesticides, herbicides etc. on field level has been a success story in economic terms, relatively speaking in developed countries. Due to already described social and technological imperatives this story can be made an even greater success within the Precision Agriculture conceptual framework of managing inputs site-specifically, thus contributing to environmental protection and higher productivity.

The first task in tackling the problem caused by such variabilities have been identified, by means of various ground-truth sampling methods and technologies. This segment of the entire

approach is of crucial importance. After entirety of all influencing factors have been observed, the understanding of the dynamics affecting the growth as the next step has been done to a certain extent in various forms of decision support systems in order to be able to predict crop growth.

Modelling the life and life cycle of complex living organisms such as plants brings with it the complexity that is not necessarily easy to capture through mathematical abstractions. The crop growth models offer to a certain extent opportunity to replicate complex dynamics of interacting crop growth, environment and crop management, and help in understanding how the things were done so far and how "should" they be done in the future. Simulating "life" by measuring the daily weather, the soil-related dynamics with respect to the crop species, ecotype and cultivar traits based on the measurable phenotypical behaviour of the plants, and integrating the knowledge into existing ecosystems of the crop management with various machinery and technology through human commitment is not an easy task. The potential of faster collection of large amounts of data with sensors relevant for evaluating plant growth and yield prediction even though evident is yet to be practically implemented. Transforming large amounts of digital raw data about plant biomass, soil water etc. into agronomical meaningful plant growth and development indicators for managing crops and finally integrating them into various decision support mechanisms, is a great challenge. Striking the balance between large amounts of data and a "meaningful" amount of data in the context of optimal crop management with respect to economics involved is difficult and requires involvement of different segments of farm management and technologies.

Based on the studies conducted within this thesis and research work conducted by international scientists and precision agriculture enthusiast farmers, it can be concluded that there is a great potential in moving from field-scale to site-specific scale. The approaches described in chapter 4, 5 and 6 still contain to a certain degree mathematical abstraction that might scare away potential users. Additional experiments with greater focus on easily attainable in-field measurements by technological solutions are still required to make the approach more robust and easier marketable to potential end users. Chapter 6 pointed out that there is still room for improving currently used crop growth models, either through adjustment of certain approaches used in models or through externally attainable sensor data and maybe in the future also through a combination of crop growth models with machine learning and neural networks. Crop growth models then would contain agronomic knowledge and machine learning advanced technological approaches.

Further, the described remote- and near-sensing data integration offers a theoretical framework for uniting agronomic knowledge and technological advancement of various non-destructive data collections and should be pursued in future. Sub-chapter 2.2 described the potential combination of near- and remote-sensing technology integrated into the agricultural machinery and outlined to what extent crop production system can be made autonomous.

Based on the current trends and publications agriculture seems to move away from traditional destructive data collection in field experiments towards the use of non-destructive sensing technologies. A similar trend can be observed with crop modelling and machine learning based decision support systems. Higher priority is commonly given to machine learning, which does not necessarily include plant physiology and physical transfer occurring in the soil. At this point it is too difficult to claim what would be the more meaningful approach, and common trends observable indirectly through various publications and the amount of

publications related to a specific topic in agricultural journals will very likely influence the development direction of future decision support systems.

It is possible from time to time to get lost in the scientific and technological futuristic wishes, but one must not be forgotten: *sustainably attained higher yields in an environmentally friendly manner with respect to economics involved in farm and crop management will always be a compromise between an eco-friendly society and profit maximising farmers*.

8   Summary

The agricultural sector is considered as one of the main elements of a "functional" society and its welfare for providing means to fulfil the most basic human need for food. In the past, an improved overall welfare of developed countries was directly related to early investment in conventional farming machinery parallel to the industrial development. With the expected future population increase and climate change, farm yield gaps are expected to increase and become more volatile. The awareness of negative externalities of agricultural production and yield maximisation led to many regulations e.g. in the EU aiming to control agricultural inputs to prevent ground water and environmental pollution. The problems arising from maintaining the balance between socio-economic and environmental aspects of agricultural production while maximising yield are complex and cannot be solved with one single solution.

For the studies conducted in this dissertation a simplified yield definition was used with a major focus on yield limiting (nutrient) and yield reducing factors (leaf disease). There are still major knowledge gaps in fully understanding dynamics of potential yield (attainable under fully controlled conditions on e.g. research stations or in controlled environments) and actual yield attained by farmers due to spatial and temporal heterogeneity occurring in the field on large scale crop production. As more insight into soil- and weather-related dynamics was gained with various technological solutions available for agricultural practices, a more detailed investigation of the variabilities affecting yield came about. These technologies include various remote- and near-sensing technologies specialised for measuring various aspects of crop growth and field spatial and temporal variabilities. The measurement range of remote- and near-sensing technologies is able to capture soil×crop×weather dynamics by direct (e.g. N status measuring sensors) and indirect measurement (e.g. measuring plant chlorophyll, multi-spectral reflectance etc.) of soil parameters, plant biomass accumulation (for predicting yield) and weather parameters affecting in-season crop growth. Decision support tools such as process oriented crop growth models (DSSAT-Decision Support System for Agrotechnlogy Transfer) have the potential to capture in-season dynamics and rely on feedback of various sensor data for more accurate depiction of plant growth and yield prediction. Some of the major problems related to crop growth models are the underlying mathematical abstractions of plant phenological development and required calibration of plant genetics. In the future, sensor-based N status can be used for checking the model prediction of plant N uptake and for adjusting further in-season N prescriptions on a site-specific level. Various sensors have also potential to be used for evaluating leaf disease status in the field for predicating yield losses.

The overall objectives of dissertation were:
1) Crop model-based analysis of yield gaps in maize production based on economic optimization of nitrogen as an example for a yield limiting factor.
2) Crop model-based analysis of leaf disease impact on sugar beet yield as an example for a yield reducing factor.
3) Analysis of specific mathematical abstractions used in crop models and their accuracy level influencing the predicted potential yield and yield gap.

The case study on the yield limiting factor nitrogen was based on eight years of maize trials from two different locations using the DSSAT-CERES-Maize crop growth model. Short- and long-term marginal net return maximising N application rates were investigated on a site-specific level. Within this study a QGIS based software solution was developed to complement the existing crop growth model and enable site-specific variable N and plant density

optimisation based on price of grain and N fertiliser over a longer period of weather data for two different fields (Germany and US). The N application recommendations were on average 9% lower for the McGarvey field (US, 5 years) and 48% lower for the Riech field (Germany, 3 years) when compared to the uniform N application rate commonly applied by the farmer. The study indicated that for low yielding heterogeneous fields, such as the Riech field, variable N application is crucial for-profit maximisation. Even though a simplified marginal net return analysis was conducted, it was apparent that additionally applied N in Riech did not increase grain yields in quantities to cover additional N costs.

In a second case study Cercospora leaf spot disease in sugar beet was investigated as an example for a yield reducing factor. The study was based on field trial in Southwest Germany by using the DSSAT-CERES-Sugarbeet model. Data on Cercospora leaf spot disease was collected throughout the growing season starting end of June each year. Cercospora leaf spot disease has a major influence on harvested sugar beet yield and consequently the amount of extractable sugar. The study consisted of modifying and adapting an existing pest subroutine in DSSAT enabling the model to simulate sugar beet yield losses based on the observed Cercospora leaf spot disease ratings from the field. Simulation of the leaf disease effects on yield is an extremely complex task due to spatial (spread of spores in the field) and temporal variability (leaf disease favouring weather conditions) of factors that are not easy to capture. Based on the data collected in Southwest Germany in the three-year field experiment, the model was able to simulate sugar beet yield losses based on the observed Cercospora leaf spot disease ratings, and consequently sugar yield losses required for economic evaluation. Model performance of the calibration treatment for LAI, tops weight and storage root beet yield resulted in R^2 higher than 0.82 and d-statistics higher than 0.94. Evaluation treatments resulted in high R^2 and d-statistics, with few exceptions mostly caused by drought.

Crop growth models are a mathematical representation of the real physiological and physical transfers occurring during plant development with plant behaviour derived from field experiments and correlations based on statistical evaluations. Plant phenological development and in-season above-and below-ground biomass accumulation are defined through crop cultivar coefficients, which have to be calibrated based on the field experiment measured data. Commonly cultivar coefficients are estimated manually by trial and error method. Within the third case study of this thesis, a cultivar coefficient estimation tool (TSE, external Python plug-in) was developed enabling users to use all in-season measured data for estimating cultivar coefficients. Cultivar coefficients estimated for two different cultivar types were based on 346 in-season observations over multiple target variables (eg. LAI, tops weight, grain weight etc.) of six experiments from different locations in the US. Multiple-treatment based TSE cultivar coefficient estimates were evaluated with independent data set and compared to the Standard DSSAT coefficients manually estimated by trial and error done by model developers. The averaged normalised root mean squared error for LAI, shoot weight, leaf weight and grain weight of TSE coefficient estimates was 26% lower for one cultivar and about the same for the other cultivar when compared to the Standard DSSAT coefficients-based crop model predictions. Correct simulation of the phenological events and above- and below-ground biomass in-season accumulation rates is crucial for accurate evaluation of specific management practices such as N application or leaf disease damage effects on yield.

All studies conducted within this dissertation, even though offering insight into some very important aspects of the yield maximisation are still scientific recommendations and not practical recommendations and solutions. Based on the studies and results obtained, the

potential seems obvious especially when considered in the context of new technologies to compliment already existing agronomic knowledge that can be unified within a crop growth model-based decision support platform.

9   Zusammenfassung

Der Agrarsektor gilt als eines der Hauptelemente einer "funktionierenden" Gesellschaft und ihres Wohlstands, da er Mittel zur Deckung des menschlichen Grundbedürfnisses nach Nahrung bereitstellt. In der Vergangenheit stand die Verbesserung des allgemeinen Wohlstands in den Industrieländern in direktem Zusammenhang mit dem Zeitpunkt der Investitionen in konventionelle landwirtschaftliche Maschinen parallel zur industriellen Entwicklung. Angesichts des erwarteten Bevölkerungswachstums und des Klimawandels wird erwartet, dass die Ertragsunterschiede in der Landwirtschaft in Zukunft weiter zunehmen und stärker schwanken werden. Das Bewusstsein für die negativen externen Effekte der landwirtschaftlichen Produktion und der Ertragsmaximierung führte zu zahlreichen Verordnungen, z. B. in der EU, die darauf abzielen, den landwirtschaftlichen Input zu kontrollieren, um Grundwasser- und Umweltverschmutzung zu vermeiden. Die Probleme, die sich aus der Aufrechterhaltung des Gleichgewichts zwischen sozioökonomischen und ökologischen Aspekten der landwirtschaftlichen Produktion bei gleichzeitiger Ertragsmaximierung ergeben, sind komplex und lassen sich nicht mit einer einzigen Lösung lösen. Für die in dieser Dissertation durchgeführten Studien wurde eine vereinfachte Ertragsdefinition verwendet, wobei der Schwerpunkt auf ertragsbegrenzenden Faktoren (Nährstoffe) und ertragsmindernden Faktoren (Blattkrankheiten) lag. Es gibt immer noch große Wissenslücken im Hinblick auf das vollständige Verständnis der Dynamik des potenziellen Ertrags (der unter vollständig kontrollierten Bedingungen, z. B. auf Forschungsstationen oder in kontrollierten Umgebungen, erzielt werden kann) und des tatsächlichen Ertrags, der von den Landwirten erzielt wird. Mit einem zunehmenden Einblick in die boden- und wetterbedingte Dynamik und den verschiedenen technologischen Lösungen, die für die landwirtschaftliche Praxis zur Verfügung stehen, wurde eine detailliertere Untersuchung der ertragsbeeinflussenden Schwankungen möglich. Zu diesen Technologien gehören verschiedene Fern- und Naherkundungstechnologien, die auf die Messung verschiedener Aspekte des Pflanzenwachstums und der räumlichen und zeitlichen Variabilität des Feldes spezialisiert sind. Das Messspektrum der Fern- und Naherkundungstechnologien ist in der Lage, die Dynamik von Boden, Pflanze und Wetter durch direkte (z. B. Sensoren zur Messung des N-Status) und indirekte Messungen (z. B. Messung des Chlorophylls der Pflanzen, multispektrale Reflexion usw.) von Bodenparametern, der pflanzlichen Biomasseakkumulation (zur Vorhersage des Ertrags) und von Wetterparametern, die sich auf das Pflanzenwachstum während der Saison auswirken, zu erfassen. Entscheidungshilfen wie prozessorientierte Pflanzenwachstumsmodelle (DSSAT - Decision Support System for Agrotechnlogy Transfer) haben das Potenzial, die Dynamik während der Saison zu erfassen, und stützen sich auf die Rückkopplung verschiedener Sensordaten für eine genauere Darstellung des Pflanzenwachstums und der Ertragsprognose. Einige der Hauptprobleme im Zusammenhang mit Pflanzenwachstumsmodellen sind die zugrundeliegenden mathematischen Abstraktionen der phänologischen Entwicklung von Pflanzen und die erforderliche Kalibrierung der Pflanzengenetik. In Zukunft soll der sensorgestützt erfasste N-Status zur Überprüfung der Modellvorhersage der pflanzlichen N-Aufnahme und zur Anpassung weiterer saisonaler N-Applikationen auf teilflächenspezifischer Ebene verwendet werden. Verschiedene Sensoren haben auch das

Potenzial, zur Bewertung des Blattkrankheitsstatus im Feld eingesetzt zu werden, um Ertragsverluste vorherzusagen.

Die allgemeinen Ziele der Dissertation waren:

1) Modellbasierte Analyse von Ertragsunterschieden im Maisanbau auf der Grundlage der wirtschaftlichen Optimierung von Stickstoff als Beispiel für einen ertragsbegrenzenden Faktor.
2) Modellgestützte Analyse der Auswirkungen von Blattkrankheiten auf den Zuckerrübenertrag als Beispiel für einen ertragsmindernden Faktor.
3) Analyse spezifischer mathematischer Abstraktionen, die in Pflanzenwachstumsmodellen verwendet werden, und ihres Genauigkeitsgrades, der den vorhergesagten potenziellen Ertrag und die Ertragsunterschiede beeinflusst.

Die Studie über den ertragsbegrenzenden Faktor Stickstoff basiert auf achtjährigen Maisversuchen an zwei verschiedenen Standorten unter Verwendung des DSSAT-CERES-Mais-Wachstumsmodells. Auf teilflächenspezifischer Ebene wurden kurz- und langfristige, den Grenzertrag maximierende N-Ausbringungsraten untersucht. Im Rahmen dieser Studie wurde eine QGIS-basierte Softwarelösung entwickelt, um das bestehende Pflanzenwachstumsmodell zu ergänzen und eine teilflächenspezifische variable N- und Pflanzendichte-Optimierung auf der Grundlage der Preise für Getreide und N-Dünger über einen längeren Zeitraum mit Wetterdaten für zwei verschiedene Felder (Deutschland und USA) zu ermöglichen. Die Empfehlungen für die N-Düngung lagen im Durchschnitt um 9 % für das McGarvey-Feld (USA, 5 Jahre) und um 48 % für das Riech-Feld (Deutschland, 3 Jahre) niedriger als die einheitliche N-Düngung, die der Landwirt üblicherweise vornimmt. Die Studie zeigte, dass bei ertragsschwachen, heterogenen Feldern, wie dem Beispiel des Riech-Feldes, eine variable N-Ausbringung für die Gewinnmaximierung entscheidend ist. Obwohl eine vereinfachte marginale Nettoertragsanalyse durchgeführt wurde, zeigte sich, dass der zusätzlich ausgebrachte Stickstoff auf der Fläche Riech die Kornerträge nicht in dem Maße erhöhte, das die zusätzlichen Stickstoffkosten gedeckt werden konnten.

In einer zweiten Studie wurde die Cercospora-Blattfleckenkrankheit bei Zuckerrüben als Beispiel für einen ertragsmindernden Faktor untersucht. Die Studie basierte auf einem Feldversuch in Südwestdeutschland unter Verwendung des DSSAT-CERES-Zuckerrübenmodells. Die Daten zur Cercospora-Blattfleckenkrankheit wurden während der gesamten Vegetationsperiode ab Ende Juni eines jeden Jahres gesammelt. Die Cercospora-Blattfleckenkrankheit hat einen großen Einfluss auf den Ertrag der geernteten Zuckerrüben und folglich auf die Menge des extrahierbaren Zuckers. Im Rahmen der Studie wurde ein bestehendes Pest-Programm in DSSAT modifiziert und angepasst, um das Modell in die Lage zu versetzen, Ertragsverluste bei Zuckerrüben auf der Grundlage der beobachteten Cercospora-Blattfleckenkrankheitswerte aus dem Feld zu simulieren. Die Simulation der Auswirkungen der Blattfleckenkrankheit auf den Ertrag ist aufgrund der räumlichen (Ausbreitung der Sporen auf dem Feld) und zeitlichen Variabilität (die Blattfleckenkrankheit begünstigende Witterungsbedingungen) von Faktoren, die nicht einfach zu erfassen sind, eine äußerst komplexe Aufgabe. Auf der Grundlage der im Rahmen des dreijährigen Feldversuchs in Südwestdeutschland gesammelten Daten konnte das Modell auf der Grundlage der beobachteten Cercospora-Blattfleckenkrankheiten Ertragseinbußen bei Zuckerrüben und folglich auch die für die wirtschaftliche Bewertung erforderlichen Zuckerertragsverluste simulieren. Die Modellleistung der Kalibrierungsbehandlung für LAI, Biomasse und

Rübenertrag ergab ein R² von über 0.82 und eine d-Statistik von über 0.94. Die Auswertungsbehandlungen ergaben hohe R²- und d-Statistiken, mit wenigen Ausnahmen, die hauptsächlich durch Trockenheit verursacht wurden.

Pflanzenwachstumsmodelle sind eine mathematische Darstellung der realen physiologischen und physikalischen Zusammenhänge, die während der Pflanzenentwicklung auftreten, wobei die Reaktion der Pflanze aus Feldversuchen und Korrelationen auf der Grundlage statistischer Auswertungen abgeleitet wird. Die phänologische Entwicklung der Pflanzen und die saisonale Akkumulation von ober- und unterirdischer Biomasse werden durch genetische Koeffizienten definiert, die anhand von Messdaten aus Feldversuchen kalibriert werden müssen. Üblicherweise werden die Koeffizienten sortenspezifisch manuell nach der Trial-and-Error-Methode geschätzt. Im Rahmen der dritten Studie dieser Arbeit wurde ein Tool zur Schätzung der genetischen Koeffizienten (TSE, externes Python-Plug-in) entwickelt, das es den Nutzern ermöglicht, alle während der Saison gemessenen Daten zur Schätzung der Koeffizienten zu verwenden. Die im Rahmen dieser Studie geschätzten genetischen Koeffizienten für zwei verschiedene Sorten basieren auf 346 Beobachtungen während der Saison für mehrere Zielvariablen (z. B. LAI, Gewicht der Spitzen, Korngewicht usw.) aus sechs Versuchen an verschiedenen Standorten in den USA. Die auf Mehrfachbehandlungen basierenden TSE- Koeffizientenschätzungen wurden mit unabhängigen Datensätzen bewertet und mit den Standard-DSSAT-Koeffizienten verglichen, die von den Modellentwicklern manuell durch Trial-and-Error-Methode ursprünglich geschätzt wurden. Der RMSE (root mean squared error) für LAI, Stängel-, Blatt-und Korngewicht der TSE-Schätzungen war für eine Sorte um 26 % niedriger und für die andere Sorte ungefähr gleich, wenn man sie mit den auf den Standard-DSSAT-Koeffizienten basierenden Vorhersagen des Modells vergleicht. Die korrekte Simulation der phänologischen Ereignisse und der Akkumulationsraten der ober- und unterirdischen Biomasse während der Saison ist entscheidend für die genaue Bewertung und Durchführung spezifischer Bewirtschaftungsmaßnahmen, wie z. B. die Applikation von Stickstoff oder die Auswirkungen von Blattkrankheiten auf den Ertrag.

Alle im Rahmen dieser Dissertation durchgeführten Studien bieten zwar Einblicke in einige sehr wichtige Aspekte der Ertragsmaximierung, sind aber dennoch wissenschaftliche Empfehlungen und keine praktischen Empfehlungen und Lösungen. Auf der Grundlage der Studien und der erzielten Ergebnisse scheint das Potenzial dieser Ansätze allerings offensichtlich, insbesondere wenn man es im Zusammenhang mit neuen Technologien betrachtet, die das bereits vorhandene agronomische Wissen ergänzen und in einer auf einem Pflanzenwachstumsmodell basierenden Plattform zur Entscheidungsunterstützung vereint werden können.

References

Alotaibi, K.D., Cambouris, A.N., St. Luce, M., Ziadi, N., Tremblay, N., **2018**. Economic optimum nitrogen fertilizer rate and residual soil nitrate as influenced by soil texture in corn production. *Agron. J.,* 110, 2233–2242. https://doi.org/10.2134/agronj2017.10.0583

Anar, M.J., Lin, Z., Teboh, J., Ostlie, M., **2015**. Modeling energy beets using the Decision Support System for Agrotechnology Transfer. *Trans. ASABE*, RRV15-048, doi:10.13031/RRV15-048.

Anar, M.J., Lin, Z., Hoogenboom, G., Shelia, V., Batchelor, W.D., Teboh, J.M., Ostlie, M., Schatz, B.G., Khan, M., **2019**. Modeling Growth, Development and Yield of Sugarbeet using DSSAT. *Agric. Syst.*, *169*, 58–70.

Antille, L.D., Lobsey, R.C., McCarthy, L. C., Thomasson, J. A., **2018**. A review of the state of the art in agricultural automation. Part IV: Sensor-based nitrogen management technologies. *ASABE Meet.,* Pap. 1–14.

Appel, F., Ostermeyer-Wiethaup, A., Balmann, A., **2016**. Effects of the German Renewable Energy Act on structural change in agriculture – The case of biogas. *Util. Policy,* 41, 172–182. https://doi.org/10.1016/j.jup.2016.02.013

Awokuse, T.O., Xie, R., **2015**. Does agriculture really matter for economic growth in developing countries? *Can. J. Agric. Econ.,* 63, 77–99. https://doi.org/10.1111/cjag.12038

Baey, C., Didier, A., Lemaire, S., Maupas, F., Cournède, P., **2014**. Parametrization of five classical plant growth models applied to sugar beet and comparison of their predictive capacity on root yield and total biomass. *Ecol. Modell.,* 290, 11–20. https://doi.org/10.1016/j.ecolmodel.2013.11.003

Batchelor, W.D., Jones, J.W., Boote, K.J., Pinnschmidt, H.O., **1993**. Extending the use of crop models to study pest damage. Trans. - *Am. Soc. Agric. Eng. Gen.,* Ed. 36, 551–558. https://doi.org/10.13031/2013.28372

Batchelor, W. D., Paz, J. O., Thorp, K.R., **2004**. Development and evaluation of a decision support system for precision agriculture, in: Mulla, D.J. (Ed.), Proceedings of the 7th International Conference on Precision Agriculture and Other Precision Resources Management. Minneapolis, MN, USA, pp. 2005–2009.

Batchelor, W.D., Suresh, L.M., Zhen, X., Beyene, Y., Wilson, M., Kruseman, G., Prasanna, B., **2020**. Simulation of maize lethal necrosis (MLN) damage using the CERES-Maize model. *Agron.,* 10, 1–14. https://doi.org/10.3390/agronomy10050710

Binswanger, H., **1986**. Agricultural mechanization: A comparative historical perspective. World Bank *Res. Obs.,* 1, 27–56. https://doi.org/10.1093/wbro/1.1.27

Bock, C.H., Barbedo, J.G.A., Del Ponte, E.M., Chiang, K.S., Bohnenkamp, D., Mahlein, A.K., **2020**. From visual estimates to fully automated sensor-based measurements of plant disease severity: Status and challenges for improving accuracy. *Phytopathol. Res.*, *2*, doi:10.1186/s42483-020-00049-8.

Boote, K.J., Jones, J.W., Mishoe, J.W., Wilkerson, G.G. Modelling growth and yield of groundnut. In *Agrometeorology of Groundnut Proceedings of the International Symposium, ICRISAT Sahelian Centre*, *Niamey, Niger, 21–26 August 1985*; Sivakumar, M.V.K., Virmani, S.M., Eds.; ICRISAT: Patancheru, India, 1986; pp. 243–254.

Boote K.J., Jones J.W., Hoogenboom G., Wilkerson G.G., **1997**. Evaluation of the CROPGRO-Soybean model over a wide range of experiments. In Kropff M.J. et al. (Eds.), *Applications of Systems Approaches at the Field Level* (Vol. 6, pp. 113-133). Springer, Dordrecht. https://doi.org/10.1007/978-94-017-0754-1_8

Boote K.J., Jones J.W., Hoogenboom G., Pickering N.B., **1998**. The CROPGRO model for grain legumes. In Tsuji G.Y., Hoogenboom G., & Thornton P.K. (Eds.), *Understanding Options for Agricultural Production* (Vol. 7, pp. 99-128). Springer, Dordrecht. https://doi.org/10.1007/978-94-017-3624-4_6

Boote, K.J., Jones, J.W., Batchelor, W.D., Nafziger, E.D., Myers, O., **2003**. Genetic coefficients in the CROPGRO–Soybean model: Links to field performance and genomics. *Agron. J., 95*, 32–51. https://doi.org/10.2134/agronj2003.0032

Boote, K.J., Hoogenboom, G., Jones, J.W., Ingram, K.T., **2009**. Modeling N-Fixation and Its Relationship to N Uptake in the CROPGRO Model. In Ma, L., Ahuja, L., & Bruulsema, T. (Eds.), *Quantifying and Understanding Plant Nitrogen Uptake for Systems Modeling* (1st. ed.), Boca Raton, Florida: Taylor & Francis Group LLC. https://doi.org/10.1201/9781420052978

Boote, K.J., Jones, J.W., Hoogenboom, G., White, J.W., **2010**. The role of Crop systems simulation in agriculture and environment. *Int. J. Agric. Environ. Inf. Syst.,* 1, 41–54. https://doi.org/10.4018/jaeis.2010101303

Boote, K.J., Porter, C., Jones, J.W., Thorburn, P.J., Kersebaum, K.C., Hoogenboom, G., White, J.W., Hatfield, J.L., **2015**. Sentinel Site Data for Crop Model Improvement—Definition and Characterization. In Hatfield, J.L., Fleisher, D. (Eds.), *Improving Crop modeling Tools to Assess Climate Change Effects on Crop Response. Advances in Agricultural Systems Modeling* (Vol. 7, pp. 125-158). Madison, Wisconsin: American Society of Agronomy, Crop Science Society of America, and Soil Science Society of America. https://doi.org/10.2134/advagricsystmodel7.2014.0019

Boote, K. (Ed.)., **2019**. *Advances in crop modelling for a sustainable agriculture* (1st ed.). London, England: Burleigh Dodds Science Publishing. https://doi.org/10.1201/9780429266591

Boote, K.J. Jones, J.W., Hoogenboom, G., **2021**. Incorporating realistic trait physiology intro crop growth models to support genetic improvement. *isP., 3*(1). https://doi:10.1093/insilicoplants/diab002

Bourgeois, G. Interrelationships between Late Leafspot Disease and Florunner Peanut: A Modelling Approach. Ph.D. Thesis, University of Florida, Gainesville, FL, USA, **1989**.

Braga, R.P., Jones, J.W., Basso, B., **2015**. Weather Induced Variability in Site-Specific Management Profitability: A Case Study. Proc. 4th Int. Conf. Precis. Agric., 1853–1863. https://doi.org/10.2134/1999.precisionagproc4.c89b

Buddhaboon, C., Jintrawet, A., Hoogenboom, G., **2018**. Methodology to estimate rice genetic coefficients for the CSM-CERES-Rice model using GENCALC and GLUE genetic coefficient estimators. *J. Agric. Sci., 156*(4), 482-492. https://doi.org/10.1017/S0021859618000527

Cao, X., Luo, Y., Zhou, Y., Fan, J., Xu, X., West, J.S., Duan, X., Cheng, D., **2015**. Detection of powdery mildew in two winter wheat plant densities and prediction of grain yield using canopy hyperspectral reflectance. *PLoS One,* 10, 1–14. https://doi.org/10.1371/journal.pone.0121462

Del Grosso, S.J., Parton, W.J., Mosier, A.R., Hartman, M.D., Brenner, J., Ojima, D.S., Schimel, D.S., **2001**. Simulated interaction of carbon dynamics and nitrogen trace gas fluxes using the DAYCENT model. In: Schaffer, M., Ma, L., Hansen, S. (Eds.), Modeling Carbon and Nitrogen Dynamics for Soil Management. CRC Press, Boca Raton, Florida, pp. 303e332.

Dodds, B., In, S., Science, A., **2019**. Advances in crop modelling for a sustainable agriculture. *Adv. Crop Model. a Sustain. Agric*., https://doi.org/10.1201/9780429266591

Donatelli, M., Magarey, R.D., Bregaglio, S., Willocquet, L., Whish, J.P.M., Savary, S., **2017**. Modelling the impacts of pests and diseases on agricultural systems. *Agric. Syst.,* 155, 213–224. https://doi.org/10.1016/j.agsy.2017.01.019

Duan, Q., Sorooshian, S., Gupta, V., **1992**. *SAC-SMA,* 28, 1015–1031.

Duncan, W. G., **1958**. The relationship between corn population and yield. *Agron. J.*, 50, 82-84.

Eurostat, Agricultural Production—Crops. Available online: https://ec.europa.eu/eurostat/statistics-explained/index.php/Agricultural_production_-_crops#cite_note-3 (**accessed on 5 October 2020**).

Ewert, F., Rötter, R.P., Bindi, M., Webber, H., Trnka, M., Kersebaum, K.C., Olesen, J.E., van Ittersum, M.K., Janssen, S., Rivington, M., Semenov, M.A., Wallach, D., Porter, J.R., Stewart, D., Verhagen, J., Gaiser, T., Palosuo, T., Tao, F., Nendel, C., Roggero, P.P., Bartošová, L., Asseng, S., **2015**. Crop modelling for integrated assessment of risk to food production from climate change. Environ. *Model. Softw.,* 72, 287–303. https://doi.org/10.1016/j.envsoft.2014.12.003

FAOSTAT (2014). Statistical database of the Food and Agriculture Organization (FAO) of the United Nations, http://www.fao.org/faostat/en/#country/. **Last accessed August 2018**.

Fowler, D., Coyle, M., Skiba, U., Sutton, M.A., Cape, J.N., Reis, S., Sheppard, L.J., Jenkins, A., Grizzetti, B., Galloway, J.N., Vitousek, P., Leach, A., Bouwman, A.F., Butterbach-Bahl, K., Dentener, F.,

Stevenson, D., Amann, M., Voss, M., **2013**. The global nitrogen cycle in the Twentyfirst century. *Philos. Trans. R. Soc. B Biol. Sci.*, 368. https://doi.org/10.1098/rstb.2013.0164

Gao, Y., Wallach, D., Liu, B., Dingkuhn, M., Boote, K.J., Singh, U., Asseng, S., Kahveci, T., He, J., Zhang, R., Confalonieri, R., Hoogenboom, G., **2020**. Comparison of three calibration methods for modeling rice phenology. *Agric. For. Meteorol.*, *280* (107785). https://doi.org/10.1016/j.agrformet.2019.107785

Gijsman, A.J., Hoogenboom, G., Parton, W.J., Kerridge, P.C., **2002**. Modifying DSSAT for low-input agricultural systems, using a soil organic matter*/residue module from CENTURY. *Agron. J.*, 94 (3), 462/474.

Gillespie, S., van den Bold, M., **2017**. Agriculture, Food Systems, and Nutrition: Meeting the Challenge. Glob. Challenges 1, 1600002. https://doi.org/10.1002/gch2.201600002

Godwin, D.C., Jones, C.A., **1991**. Nitrogen dynamics in soil/ plant systems. In: Hanks, J., Ritchie, J.T. (Eds.), Modeling Plant and Soil Systems (Agronomy monograph no. 31). ASA, CSSA, and SSSA, Madison, WI, pp. 287/321.

Godwin, D.C., Singh, U., **1998**. Nitrogen balance and crop response to nitrogen in upland and lowland cropping systems. In: Tsuji, G.Y., Hoogenboom, G., Thornton, P.K. (Eds.), Understanding Options for Agricultural Production. System Approaches for Sustainable Agricultural Development. Kluwer Academic Publishers, Dordrecht, Netherlands, pp. 55/77.

Gopala Pillai, S., Tian, L., Bullock, D., **1999**. Yield Mapping with Digital Aerial Color Infrared (CIR) Images. SAE Transactions, 108, 303–316. http://www.jstor.org/stable/44723053

Grassini, P., Torrion, J.A., Yang, H.S., Rees, J., Andersen, D., Cassman, K.G., Specht, J.E., **2015**. Soybean yield gaps and water productivity in the western U.S. Corn Belt. *F. Crop. Res.*, 179, 150–163. https://doi.org/10.1016/j.fcr.2015.04.015

Griepentrog, H.W., Nørremark, M., Nielsen, J., **2006**. Autonomous intra-row rotor weeding based on GPS, in: CIGR World Congress Agricultural Engineering for a Better World.

Guillaume, S., Bergez, J.E., Wallach, D., Justes, E., **2011.** Methodological comparison of calibration procedures for durum wheat parameters in the STICS model. *Eur. J. Agron.,* 35, 115–126. https://doi.org/10.1016/j.eja.2011.05.003

Guilpart, N., Grassini, P., Sadras, V.O., Timsina, J., Cassman, K.G., **2017**. Estimating yield gaps at the cropping system level. *F. Crop. Res.*, 206, 21–32. https://doi.org/10.1016/j.fcr.2017.02.008

Haboudane, D., Miller, J.R., Pattey, E., Zarco-Tejada, P.J., Strachan, I.B., **2004**. Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture. *Remote Sens. Environ.,* 90, 337–352. https://doi.org/10.1016/j.rse.2003.12.013

He, J., Jones, J.W., Graham, W.D., Duke, M.D., **2010**. Influence of likelihood function choice for estimating crop model parameters using the generalized likelihood uncertainty estimation methods. *Agric. Syst., 103*, 256–264. https://doi.org/10.1016/j.agsy.2010.01.006

Heil, K., Schmidhalter, U., **2017**. The application of EM38: Determination of soil parameters, selection of soil sampling points and use in agriculture and archaeology. *Sensors* (Switzerland) 17. https://doi.org/10.3390/s17112540

Heiß, A., Paraforos, D.S., Sharipov, G.M., Griepentrog, H.W., **2021**. Modeling and simulation of a multi-parametric fuzzy expert system for variable rate nitrogen application. *Comput. Electron. Agric.,* 182. https://doi.org/10.1016/j.compag.2021.106008

Hillnhütter, C., Mahlein, A.K., Sikora, R.A., Oerke, E.C., **2011**. Remote sensing to detect plant stress induced by Heterodera schachtii and Rhizoctonia solani in sugar beet fields. *F. Crop. Res.,* 122, 70–77. https://doi.org/10.1016/j.fcr.2011.02.007

Hoogenboom, G., Jones, J.W., Wilkens, P.W., Porter, C.H., Boote, K.J., Hunt, L.A., Singh, U., Lizaso, J.L., White, J.W., Uryasev, O., Royce, F.S., Ogoshi, R., Gijsman, A.J., Tsuji, G.Y., **2011**. Decision Support System for Agrotechnology Transfer Version 4.5. (www.DSSAT.net). Honolulu, HI, USA: University of Hawaii.

Hoogenboom, G., Jones, J.W., Traore, P.C.S., Boote, K.J., **2012**. Experiments and data for model evaluation and application. In Kihara, J., Fatondji, D., Jones, J.W., Hoogenboom, G., Tabo, R., &

Bationo, A. (Eds), *Improving Soil Fertility Recommendations in Africa using the Decision Support Systems for Agrotechnology Transfers (DSSAT)*(pp. 9-18.). Springer, Dordrecht, the Netherlands.

Hoogenboom, G., Porter, C.H., Boote, K.J., Shelia, V., Wilkens, P.W., Singh, U., White, J.W., Asseng, S., Lizaso, J.I., Moreno, L.P., Pavan, W., Ogoshi, R., Hunt, L.A., Tsuji, G.Y., Jones, J.W., **2019a**. The DSSAT crop modeling ecosystem. In Boote, K.J. (Ed.), *Advances in Crop Modeling for a Sustainable Agriculture* (1st ed., pp. 173-216). Burleigh Dodds Science Publishing, Cambridge, United Kingdom. http://dx.doi.org/10.19103/AS.2019.0061.10

Hoogenboom, G., Porter, C.H., Shelia, V., Boote, K.J., Singh, U., White, J.W., Hunt, L.A., Ogoshi, R., Lizaso, J.I., Koo, J., Asseng, S., Singels, A., Moreno, L.P., Jones, J.W., **2019b**. Decision Support System for Agrotechnology Transfer (DSSAT) Version 4.7 (www.DSSAT.net). DSSAT Foundation, Gainesville, Florida, USA.

Hoogenboom, G., Justes, E., Pradal, C., Launay, M., Asseng, S., Ewert, F., Martre, P., **2020**. ICROPM 2020: Crop Modeling for the Future. *J. Agric. Sci.,* 158, 791–793. https://doi.org/10.1017/S0021859621000538

Hunt, L.A., Pararajasingham, S., Jones, J.W., Hoogenboom, G., Imamura, D.T., Ogoshi, R.M., **1993**. GENCALC: Software to facilitate the use of crop models for analyzing field experiments. *Agron. J.,* 85, 1090–1094. https://doi.org/10.2134/agronj1993.00021962008500050025x

ICUMSA: Methods Book. Method GS6-3: The determination of the polarisation of sugar beet by the macerator or cold aqueous digestion method using aluminium sulphate as clarifying agent-official. In *Internation-al Commission for Uniform Methods of Sugar Analysis*; Publisher: Berlin, Germany, **2003**.

Jones, F.G.W., Dunning, R.A., Hamphries, K.P., **1955**. The effect of defoliation and loss of stand upon yield of sugar beets. *Ann. Appl. Biol.*, *43*, 63–70.

Jones, C.A., Kiniry, J.R., **1986**. *CERES-Maize: A Simulation Model of Maize Growth and Development*; Texas A&M University Press: College Station, TX, USA, p. 194.

Jones, J.W., Tsuji, G.Y., Hoogenboom, G., Hunt, L.A., Thornton, P.K., Wilkens, P.W., Imamura, D.T., Bowen, W.T., Singh, U., **1998**. Decision support system for agrotechnology transfer: DSSAT v3. In Tsuji G.Y., Hoogenboom G., & Thornton P.K. (Eds.), *Understanding Options for Agricultural Production. Systems Approaches for Sustainable Agricultural Development* (Vol. 7, pp. 157-177). Springer, Dordrecht. https://doi.org/10.1007/978-94-017-3624-4_8

Jones, J.W., Hoogenboom, G., Porter, C.H., Boote, K.J., Batchelor, W.D., Hunt, L.A., Wilkens, P.W., Singh, U., Gijsman, A.J., Ritchie, J.T., **2003**. The DSSAT cropping system model, *Europ. J. of Agron.* https://doi.org/10.1016/S1161-0301(02)00107-7

Jones, J.W., He, J., Boote, K.J., Wilkens, P., Porter, C.H., Hu, Z., **2011**. Estimating DSSAT Cropping System Cultivar-Specific Parameters Using Bayesian Techniques. In Ahuja, L.R., & Ma, L. (Eds.), *Methods for Introducing System Models into Agricultural Research* (Vol. 2, pp. 365-393). https://doi.org/10.2134/advagricsystmodel2.c13

Jones, J.W., Antle, J.M., Basso, B., Boote, K.J., Conant, R.T., Foster, I., Godfray, H.C.J., Herrero, M., Howitt, R.E., Janssen, S., Keating, B.A., Munoz-Carpena, R., Porter, C.H., Rosenzweig, C., Wheeler, T.R., **2017**. Brief history of agricultural systems modeling. *Agric. Syst.,* 155, 240–254. https://doi.org/10.1016/j.agsy.2016.05.014

Jung, J., Maeda, M., Chang, A., Bhandari, M., Ashapure, A., Landivar-Bowles, J., **2021**. The potential of remote sensing and artificial intelligence as tools to improve the resilience of agriculture production systems. Curr. Opin. Biotechnol. 70, 15–22. https://doi.org/10.1016/j.copbio.2020.09.003

Kersebaum K.C., Boote, K.J., Jorgenson, J.S., Nendel, C., Bindi, M., Fruehauf, C., Gaiser, T., Hoogenboom, G., Kollas, C., Olesen, J.E., Rotter, R.P., Ruget, F., Thorburn, P.J., Trnka, M., Wegehenkel, M., **2015**. Analysis and classification of data sets for calibration and validation of agro-ecosystem models. *Environ. Model. Softw., 72*, 402-417. https://doi.org/10.1016/j.envsoft.2015.05.009

Koch, B., Khosla, R., Frasier, W.M., Westfall, D.G., Inman, D., **2004**. Economic feasibility of variablerate nitrogen application utilizing site-specific management zones. *Agron. J.,* 96, 1572–

1580

Leviel, B. Evaluation of Risks and Monitoring of Nitrogen Fluxes at the Crop Level on the Romanian and Bulgarian Plain. Application to Maize, Wheat, Rapeseed and Sugar-Beet. Ph.D. Thesis, Institut National Polytechnique de Toulouse, Toulouse, France, **2000**.

Leviel, B., Crivineanu, C., Gabrielle, B. CERES-Beet, a model for the production and environmental impact of sugar beet. In Proceedings of the Joint Colloquium on Sugar Beet Growing and Modelling, Lille, France, **12 September 2003**; pp. 143–152.

Leucker, M., Wahabzada, M., Kersting, K., Peter, M., Beyer, W., Steiner, U., Mahlein, A.K., Oerke, E.C., **2017**. Hyperspectral imaging reveals the effect of sugar beet quantitative trait loci on Cercospora leaf spot resistance *Funct. Plant Biol.*, *44*, 1–9.

Link, J., Graeff, S., Batchelor, W., Claupein, W., **2006a**. Spatial variability and temporal stability of corn (Zea mays L.) grain yields - Relevance of grid size. *Arch. Agron. Soil Sci.*, 52, 427–439. https://doi.org/10.1080/03650340600775487

Link, J., Graeff, S., Batchelor, W.D., Claupein, W., **2006b**. Evaluating the economic and environmental impact of environmental compensation payment policy under uniform and variable-rate nitrogen management. *Agric. Syst.*, 91, 135–153. https://doi.org/10.1016/j.agsy.2006.02.003

Link, J., Batchelor, W.D., Graeff, S., Claupein, W., **2008**. Evaluation of current and model-based site-specific nitrogen applications on wheat (Triticum aestivum L.) yield and environmental quality. *Precis. Agric.*, 9, 251–267.https://doi.org/10.1007/s11119-008-9068-y

Link, J., Graeff, S., Claupein, W., **2013**. Comparison of uniform control and site-specific model-based nitrogen prescription in terms of grain yield, nitrogen use efficiency and economic aspects in a heterogeneous corn field. *J. fur Kult.*, 65, 236–247. https://doi.org/10.5073/JFK.2013.06.03

Lobell, D.B., Cassman, K.G., Field, C.B., **2009**. Crop yield gaps: Their importance, magnitudes, and causes. *Annu. Rev. Environ. Resour.*, 34, 179–204. https://doi.org/10.1146/annurev.environ.041008.093740

Lowenberg-DeBoer, J., **1999**. Risk Management Potential of Precision Farming Technologies. *J. Agric. Appl. Econ.*, 31, 275–285. https://doi.org/10.1017/s1074070800008555

López-Cedrón, X.F., Boote, J.K., Piñeiro, J., Sau, F., **2008**. Improving the CERES-Maize Model Ability to Simulate Water Deficit Impact on Maize Production and Yield Components. *Agron. J.*, *100*, 296–307.

Lövenstein, H.M., Rabbinage, R., Keulen, H. van, **1992**. World Food Production: Biophysical Factors in Agricultural Production.

Madani, B., Shekari, A.M., Imahori, Y., **2018**. Physiological responses to stress, Postharvest Physiology and Biochemistry of Fruits and Vegetables. *Elsevier Inc.* https://doi.org/10.1016/B978-0-12-813278-4.00020-8

Maestrini, B., Basso, B., **2018**. Drivers of within-field spatial and temporal variability of crop yield across the US Midwest. *Sci. Rep.*, 8, 1–9. https://doi.org/10.1038/s41598-018-32779-3

Martínez-Dalmau, J., Berbel, J., Ordóñez-Fernández, R., **2021**. Nitrogen fertilization. A review of the risks associated with the inefficiency of its use and policy responses. Sustain. 13, 1–15. https://doi.org/10.3390/su13105625

Memic, E., Graeff, S., Claupein, W., Batchelor, W. D., **2019**. GIS-based spatial nitrogen management model for maize: short- and long-term marginal net return maximising nitrogen application rates. *Precision Agriculture,* 20 (2), S. 295–312. DOI: 10.1007/s11119-018-9603-4.

Memic, E., Graeff, S., Batchelor, W. Extending the CERES-Beet model to simulate leaf disease in sugar beet. In *Precision Agriculture '19 Proceedings of the European Conference on Precision Agriculture, Montpellier, France, 8 July 2019b*; John, V.S., Ed.; Wageningen Academic Publishers: Wageningen, The Netherlands, 2019; doi:10.3920/978-90-8686-888-9.

Memic, E., Graeff-Hönninger, S., Hensel, O., Batchelor, W.D., **2020**. Extending the CSM-CERES-Beet Model to Simulate Impact of Observed Leaf Disease Damage on Sugar Beet Yield. *Agron.*, *10*, 1930. https://doi.org/10.3390/agronomy10121930

Memic, E., Graeff, S., Boote, K. J., Hensel, O., Hoogenboom, G., **2021**. Cultivar Coefficient Estimator

for the Cropping System Model Based on Time-Series Data: A Case Study for Soybean. *Transactions of the ASABE* 64 (4), S. 1391–1402. DOI: 10.13031/trans.14432

Miao, Y., Mulla, D. J., Batchelor, W. D., Paz, J. O., Robert, P. C., **2006**. Evaluating management zone optimal N rates with a crop growth model. *Agron. J.,* 98(3), 545-553.

Mittler, S., Petersen, J., Jörg, E., Racca, P. Integrierte Bekämpfung von Blattkrankheiten bei Zuckerrüben (Integrated treatment of leaf disease in sugar beet). In Proceedings of the 67th IIRB Congress, Brüssel, Belgium, **11-12 February 2004**.

Muller, B., Martre, P., **2019**. Plant and crop simulation models: Powerful tools to link physiology, genetics, and phenomics. *J. Exp. Bot.,* 70, 2339–2344. https://doi.org/10.1093/jxb/erz175

Müller, C., Elliott, J., Kelly, D., Arneth, A., Balkovic, J., Ciais, P., Deryng, D., Folberth, C., Hoek, S., Izaurralde, R.C., Jones, C.D., Khabarov, N., Lawrence, P., Liu, W., Olin, S., Pugh, T.A.M., Reddy, A., Rosenzweig, C., Ruane, A.C., Sakurai, G., Schmid, E., Skalsky, R., Wang, X., de Wit, A., Yang, H., **2019**. The Global Gridded Crop Model Intercomparison phase 1 simulation dataset. Sci. Data 6, 1–22. https://doi.org/10.1038/s41597-019-0023-8

Nokes, S.E., Young, J.H., **1991**. Simulation of the temporal spread of leafspot and the effect on peanut growth. *Trans. ASAE*, *34*, 653–662.

Olesen, J.E., Trnka, M., Kersebaum, K.C., Skjelvåg, A.O., Seguin, B., Peltonen-Sainio, P., Rossi, F., Kozyra, J., Micale, F., **2011**. Impacts and adaptation of European crop production systems to climate change. *Eur. J. Agron.*, 34, 96–112.

Olson, R. A., Sanders, D. H., **1988**. Maize production. In G. F. Spraggue & J. W. Dudley (Eds.), *Corn and corn improvement* (pp. 639-686). Madison, WI, USA: American Society of Agronomy.

Papadopoulos, A., Kalivas, D., Hatzichristos, T., **2011**. Decision support system for nitrogen fertilization using fuzzy theory. *Comput. Electron. Agric.,* 78, 130–139. https://doi.org/10.1016/j.compag.2011.06.007

Parent, S.É., Leblanc, M.A., Parent, A.C., Coulibali, Z., Parent, L.E., **2017**. Site-specific multilevel modeling of potato response to nitrogen fertilization. *Front. Environ. Sci.,* 5, 1–18. https://doi.org/10.3389/fenvs.2017.00081

Paz, J. O., Batchelor, W. D., Colvin, T. S., Logsdon, S. D., Kaspar, T. C., Karlen, et al., **1999**. Model-based techniques to determine variable rate nitrogen for corn. *Agric. Sys.,* 60, 69-75.

Paz, J.O., Batchelor, W.D., Tylka, G.L., Hartzler, R.G., **2001**. A modelling approach to quantifyy the effects of spatial soybean yield limiting factors. *Trans. - Am. Soc. Agric. Eng.,* 44, 1329–1334.

Pingali, P., **2007**. Chapter 54 Agricultural Mechanization: Adoption Patterns and Economic Impact. *Handb. Agric. Econ.,* 3, 2779–2805. https://doi.org/10.1016/S1574-0072(06)03054-4

QGIS Development Team (2009). QGIS Geographic Information System. Open Source Geospatial Foundation. https://plugins.qgis.org/plugins/geospatialsimulation/. **Last accessed August 2018**.

Racca, P., Jörg, E., Mittler, S., Petersen, J. CERCBET 3—A forecaster for epidemic development of *Cercospora beticola*. In Proceedings of the 67th IIRB Congress, Brüssel, Belgium, **11-12 February 2004**.

Rasche, L., Taylor, R.A.J., **2017**. A generic pest submodel for use in integrated assessment models. Trans. ASABE 60, 147–158. https://doi.org/10.13031/trans.11931

Ritchie, J.T., Godwin, D.C., Otter, S., **1984**. CERES-Wheat: A user oriented wheat yield model. Preliminary documentation. AGRISTARS Publication No. YM-U3-04442-JSC-18892, East Lansing, Michigan, p. 252.

Ritchie, J.T., Otter, S., **1985**. Description and performance of CERES-Wheat: a user-oriented wheat yield model. In: ARS Wheat Yield Project. ARS-38. Natl Tech Info Serv, Springf ield, Missouri, pp. 159/175.

Rossi, V., Battilani, P. CERCOPRI: A forecasting model for primary infections of Cercospora leaf spot of sugar beet. *Bull. OEPP,* **1991**, *21*, 527–531.

Rossberg, D., Racca, P., Joerg, E., Kleinhenz, B. Erste Erfahrung mit dem Modell CERCBET 1 (First application of CERCBEET 1 model). *Nachr Dtsch Pflanzenschutzd,* **2000**, *52*, 153–159.

Rosenzweig, C., Jones, J.W., Hatfield, J.L., Ruane, A.C., Boote, K.J., Thorburn, P., Antle, J.M., Nelson,

G.C., Porter, C., Janssen, S., Asseng, S., Basso, B., Ewert, F., Wallach, D., Baigorria, G., Winter, J.M., **2013**. The Agricultural Model Intercomparison and Improvement Project (AgMIP): Protocols and pilot studies. *Agric. For. Meteorol.,* 170, 166–182. https://doi.org/10.1016/j.agrformet.2012.09.011

Röll, G., Batchelor, W.D., Castro, A.C., Simón, M.R., Graeff-Hönninger, S., **2019**. Development and evaluation of a leaf disease damage extension in Cropsim-CERES wheat. *Agron.,* 9, 1–17. https://doi.org/10.3390/agronomy9030120

Röll, G., Memic, E., Graeff-Hönninger, S., **2020**. Implementation of an automatic time-series calibration method for the DSSAT wheat models to enhance multi-model approaches. *Agron. J.,* 112, 3891–3912. https://doi.org/10.1002/agj2.20328

Rötter, R.P., Tao, F., Höhn, J.G., Palosuo, T., **2015**. Use of crop simulation modelling to aid ideotype design of future cereal cultivars. *J. Exp. Botany., 66* (12), 3463–3476. https://doi.org/10.1093/jxb/erv098

Seidl, M.S., Batchelor, W.D., Paz, J.O., **2004**. Integrating remotely sensed images with a soybean model to improve spatial yield simulations. *Transactions of the ASAE.,* 47(6): 2081-2090. doi:10.13031/2013.17793

Seidel, S.J., Palosuo, T., Thorburn, P., Wallach, D., **2018**. Towards improved calibration of crop models—where are we now and where should we go? *Eur. J. Agron., 94*, 25–35. https://doi.org/10.1016/j.eja.2018.01.006

Shane, W. W., Teng, P.S., **1992**. Impact of Cercospora leaf spot on root weight, sugar yield, and purity of Beta vulgaris. *Plant Dis.,* 76, 812–820.

Sishodia, R.P., Ray, R.L., Singh, S.K., **2020**. Applications of remote sensing in precision agriculture: A review. *Remote Sens.,* 12, 1–31. https://doi.org/10.3390/rs12193136

Snyder, K.A., Miththapala, S., Sommer, R., Braslow, J., **2017**. THE YIELD GAP: CLOSING the GAP by WIDENING the APPROACH. *Exp. Agric.,* 53, 445–459. https://doi.org/10.1017/S0014479716000508

Sumberg, J., **2012**. Mind the (yield) gap (s). *Food Secur.* ,4, 509–518. https://doi.org/10.1007/s12571-012-0213-0

Sutton, M., Oenema, O., Erisman, J. et al., **2011**. Too much of a good thing. *Nature,* 472, 159–161. https://doi.org/10.1038/472159a

Tetio-Kagho, F., Gardner, F. P., **1988**. Response of maize to plant population density: II. Reproductive development, yield and yield adjustment. *Agron. J.*, 80, 935-940.

Theuerl, S., Herrmann, C., Heiermann, M., Grundmann, P., Landwehr, N., Kreidenweis, U., Prochnow, A., **2019**. The future agricultural biogas plant in Germany: A vision, Energies. https://doi.org/10.3390/en12030396

Thorp, K.R., Batchelor, W.D., Paz, J.O., Steward, B.L., Caragea, P.C., **2006**. Methodology to link production and environmental risks of precision nitrogen management strategies in corn. *Agric. Syst.,* 89, 272–298. https://doi.org/10.1016/j.agsy.2005.09.005

Thorp, K.R., DeJonge, K.C., Kaleita, A.L., Batchelor, W.D., Paz, J.O., **2008**. Methodology for the use of DSSAT models for precision agriculture decision support. Comput. *Electron. Agric.,* 64, 276–285. https://doi.org/10.1016/j.compag.2008.05.022

Thorp, K.R., Hunsaker, D.J., French, A.N., **2010**. Assimilating leaf area index estimates from remote sensing into the simulation of cropping systems model. *Transactions of the ASABE.*, 53(1), 251–262.

Thorp, K. R., Bronson, K., **2013**. A model-independent open-source geospatial tool for managing point-based environmental model simulations at multiple spatial locations. *Environ. Mod. & Soft.,* 50, 25-36.

Tremblay, N., Wang, Z., Belec, C., **2010**. Performance of dualex in spring wheat for crop nitrogen status assessment, yield prediction and estimation of soil nitrate content. *J. Plant Nutr.,* 33, 57–70. https://doi.org/10.1080/01904160903391081

Tsuji G.Y., **1998**. Network management and information dissemination for agrotechnology transfer. In

Tsuji G.Y., Hoogenboom G., & Thornton P.K. (Eds.), *Understanding Options for Agricultural Production.* (Vol. 7). Springer, Dordrecht. https://doi.org/10.1007/978-94-017-3624-4_18

Van Ittersum, M.K., Cassman, K.G., Grassini, P., Wolf, J., Tittonell, P., Hochman, Z., **2013**. Yield gap analysis with local to global relevance-A review. *F. Crop. Res.,* 143, 4–17. https://doi.org/10.1016/j.fcr.2012.09.009

Van Maarschalkerweerd, M., Husted, S., **2015**. Recent developments in fast spectroscopy for plant mineral analysis. *Front. Plant Sci.,* 6, 1–14. https://doi.org/10.3389/fpls.2015.00169

van Oort, P.A.J., Saito, K., Dieng, I., Grassini, P., Cassman, K.G., van Ittersum, M.K., **2017**. Can yield gap analysis be used to inform R&D prioritisation? *Glob. Food Sec.,* 12, 109–118. https://doi.org/10.1016/j.gfs.2016.09.005

Van Rees, H., McClelland, T., Hochman, Z., Carberry, P., Hunt, J., Huth, N., Holzworth, D., **2014**. Leading farmers in South East Australia have closed the exploitable wheat yield gap: Prospects for further improvement. *F. Crop. Res.,* 164, 1–11. https://doi.org/10.1016/j.fcr.2014.04.018

Wallach, D., Buis, S., Lecharpentier, P., Bourges, J., Clastre, P., Launay, M., Bergez, J.E., Guerif, M., Soudais, J., Justes, E., **2011**. A package of parameter estimation methods and implementation for the STICS crop-soil model. *Environ. Model. Softw.,* 26, 386–394. https://doi.org/10.1016/j.envsoft.2010.09.004

Wang, X., Miao, Y., Dong, R., Chen, Z., Kusnierek, K., Mi, G., Mulla, D.J., **2020**. Economic optimal nitrogen rate variability of maize in response to soil and weather conditions: Implications for site-specific nitrogen management. Agronomy 10. https://doi.org/10.3390/agronomy10091237

White, J.W., Hunt, L.A, Boote, K.J., Jones, J.W., Koo, J., Kim, S., Porter, C.H., Wilkens, P.W., Hoogenboom, G., **2013**. Integrated description of agricultural field experiments and production: The ICASA Version 2.0 Data Standards. *Comput. Electron. Agric.*, 96(1), 1-12.

Wilkerson, G.G., Jones, J.W., Boote, K.J., Ingram, K.T., Mishoe, J.W., **1983**. Modeling soybean growth for crop management. *Trans. of ASAE,* 26, 63-73. https://doi.org/10.13031/2013.33877

Windels, C.E., Lamey, H.A., Hilde, D., Widner, J., Knudsen, T.A., **1998**. Cercospora leaf spot model for sugar beet: In practice by an industry. *Plant Dis.*, 82, 716–726.

Wolf, P.F.J., Verreet, J.A., **2002**. An integrated pest management system in Germany for the control of fungal leaf diseases in sugar beet. *Plant Dis.,* 86, 336–344. https://doi.org/10.1094/PDIS.2002.86.4.336

FAO, **2006**. World Reference Base for Soil Resources. *World Soil Resources Reports No. 103*; Rome, Italy.

Xie, Q., Huang, W., Zhang, B., Chen, P., Song, X., Pascucci, S., Pignatti, S., Laneve, G., Dong, Y., **2016**. Estimating Winter Wheat Leaf Area Index from Ground and Hyperspectral Observations Using Vegetation Indices. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.,* 9, 771–780. https://doi.org/10.1109/JSTARS.2015.2489718

Yang, M., Yang, J.Y., Liu, S., Hoogenboom, G., **2014**. An evaluation of the statistical methods for testing the performance of crop models with observed data. *Agric. Syst.*, 127, 81–89.

Zhang, J., yuan, L., Pu, R., Loraamm, R.W., Yang, G., Wang, J., **2014**. Comparison between wavelet spectral features and conventional spectral features in detecting yellow rust for winter wheat. *Comput. Electron. Agric.,* 100, 79–87. https://doi.org/10.1016/j.compag.2013.11.001