

---

Parameterisation and Evaluation of the Crop Growth Model SSM-iCrop for  
Winter Wheat Grown in Eastern Austria

---

By

Dipl.-Ing. Wolfgang Fuchs BSc

A dissertation in partial fulfilment of the requirements for the academic degree

Doktor der Bodenkultur

Doctor rerum naturalium technicarum

(Dr.nat.techn.)

Universität für Bodenkultur Wien

University of Natural Resources and Life Sciences, Vienna

Department of Crop Sciences, Institute of Agronomy

Study Program: Doctoral studies of Natural Resources and Life Sciences

Supervisor: Assoc. Prof. Dr. Ahmad M. Manschadi



February 2021



1. Contents	
2. Declaration of Authorship.....	VII
3. Acknowledgements.....	VIII
4. Abstract .....	IX
5. Introduction.....	1
5.1. Winter Wheat.....	1
5.1.1. Wheat Production in Eastern Austria.....	1
5.1.2. Morphology and Phenology of the Wheat Crop .....	1
5.1.3. Agronomic Aspects of Wheat Production.....	3
5.2. Crop Growth Models.....	4
5.2.1. Brief History of Crop Modelling.....	5
5.2.2. Model Parameterisation and Calibration .....	6
5.2.3. Model Evaluation.....	7
5.2.4. SSM-iCrop – a Process-based Crop Growth Model.....	7
5.2.5. Aspects of Wheat Simulation.....	8
5.3. Research Gap Filled by this Thesis .....	13
6. Research Question and Hypotheses.....	15
7. Materials and Methods .....	15
7.1. Field Experiments.....	15
7.2. Long Term Wheat Data Set .....	21
7.3. The “iCrop” Crop Growth Model.....	21
7.3.1. Overview .....	21
7.3.2. Phenology.....	22
7.3.3. Leaf Development.....	24
7.3.4. Dry Mass Production and Partitioning.....	25
7.3.5. Nitrogen Uptake and Allocation.....	27
7.3.6. Grain Yield .....	29

7.3.7.	Soil Water .....	29
7.3.8.	Soil Nitrogen .....	31
7.4.	Parameter Calculation.....	32
7.4.1.	Weather Data Input.....	33
7.4.2.	Soil Parameters .....	33
7.4.3.	Crop/Cultivar Parameters .....	34
7.5.	Statistics.....	39
8.	Results .....	41
8.1.	Field Experiments.....	41
8.1.1.	Weather .....	41
8.1.2.	Sampling Dates .....	42
8.1.3.	Observed Soil Processes .....	42
8.1.4.	Phenology.....	45
8.1.5.	Leaf Canopy Development .....	48
8.1.6.	Yield and Dry Mass.....	50
8.1.7.	Crop Nitrogen .....	53
8.1.8.	Yield Components .....	54
8.2.	iCrop Source Code Updates.....	55
8.3.	iCrop Setup and Parameters .....	55
8.3.1.	Management .....	56
8.3.2.	Environment .....	56
8.3.3.	Emergence.....	57
8.3.4.	Vernalisation and Photoperiod .....	57
8.3.5.	Development Phases .....	57
8.3.6.	Phyllochron.....	58
8.3.7.	Leaf Area .....	59
8.3.8.	Radiation Use Efficiency.....	60
8.3.9.	Leaf:Stem Partitioning.....	61

8.3.10.	Yield.....	62
8.3.11.	Root Growth.....	64
8.3.12.	Stem Nitrogen.....	64
8.3.13.	Leaf Nitrogen .....	65
8.3.14.	Grain Nitrogen.....	67
8.3.15.	Nitrogen Uptake Rate.....	68
8.3.16.	Soil Water .....	69
8.3.17.	Soil Nitrogen .....	71
8.4.	iCrop Simulation Results .....	72
8.4.1.	Phenology.....	72
8.4.2.	Leaf Canopy Development .....	73
8.4.3.	Dry Mass Accumulation and Partitioning.....	75
8.4.4.	Crop Nitrogen Uptake and Partitioning.....	80
8.4.5.	Soil Water .....	85
8.4.6.	Soil Mineral Nitrogen .....	88
8.5.	iCrop Evaluation.....	89
8.5.1.	Long Term Data.....	89
8.5.2.	Evaluation Simulation Setup.....	90
8.5.3.	Phenology.....	90
8.5.4.	Yield.....	91
9.	Discussion .....	94
9.1.	Field Experiments.....	94
9.1.1.	Weather .....	94
9.1.2.	Soil Water .....	95
9.1.3.	Soil Nitrogen .....	95
9.1.4.	Phenology.....	97
9.1.5.	Leaf Development.....	98
9.1.6.	Dry mass.....	99

9.1.7.	Crop Nitrogen .....	99
9.1.8.	Yield.....	100
9.2.	Calculation of Model Parameters and Initials.....	100
9.2.1.	Simulation Initialisation.....	100
9.2.2.	Climate Parameter: CO <sub>2</sub> .....	101
9.2.3.	Soil Parameters .....	101
9.2.4.	Effect of Rooting Parameters on Yield.....	102
9.2.5.	Effect of Soil Runoff on Crop Dry Mass .....	103
9.2.6.	Phenology Parameters.....	104
9.2.7.	Leaf Development: Phyllochron, LAI.....	105
9.2.8.	Dry Mass: Radiation Use Efficiency .....	106
9.2.9.	Crop Nitrogen Parameters .....	106
9.2.10.	Yield Parameters.....	108
9.2.11.	Effect of Sampling Date on Parameter Calculation.....	108
9.2.12.	Overall Cultivar Differences.....	109
9.3.	Crop Model Simulations of the Field Experiments.....	109
9.3.1.	Soil Water .....	110
9.3.2.	Phenology.....	111
9.3.3.	Leaf Development.....	112
9.3.4.	Dry Mass .....	114
9.3.5.	Effect of Development-dependent SNCS.....	116
9.3.6.	Crop and Soil Nitrogen.....	116
9.3.7.	Yield.....	119
9.4.	Crop Model Evaluation .....	120
9.5.	Application of the iCrop Model.....	122
10.	Conclusions .....	124
11.	References .....	127

## 2. Declaration of Authorship

I hereby declare that this thesis and the work presented in it is entirely my own. Where I have consulted the work of others, this is always clearly stated.

Ich erkläre eidesstattlich, dass ich die Arbeit selbständig angefertigt habe. Es wurden keine anderen als die angegebenen Hilfsmittel benutzt. Die aus fremden Quellen direkt oder indirekt übernommenen Formulierungen und Gedanken sind als solche kenntlich gemacht. Diese schriftliche Arbeit wurde noch an keiner Stelle vorgelegt.

Vienna, February 2021

Dipl.-Ing. Wolfgang Fuchs BSc

### 3. Acknowledgements

I would like to thank my supervisor, Dr. Ahmad Manschadi, and supervisory team, Dr. Hans-Peter Kaul and Dr. Josef Eitzinger, as well as Dr. Afshin Soltani for their helpful critique and skilled advice. My gratitude also goes to my colleagues Craig Jackson MA, Tamara Kolodziej, Dipl.-Ing. Lukas Koppensteiner, Dipl.-Ing. Marion Kreuzhuber, Sarah Mühlbacher, Marlene Palka MSc, Dipl.-Ing. Stefan Ryall, and Dipl.-Ing. Martin Schneider, who were excellent technical and mental support.

My work was supported by two research projects and a scholarship: I would like to express my thanks to the COMBIRISK (COMBIned weather related RISK assessment monitor for tailoring climate change adaptation in Austrian crop production, KRI5AC8K12614) project, funded by ACRP (Austrian Climate Research Program), the Farm/IT (ICT for Decision Making in Farming) project, funded by the Austrian Research Promotion Agency (FFG), and the NFB (NÖ Forschung und Bildung, Federal State of Lower Austria) for the “Science Call: Dissertationen” scholarship.

Finally, I want to thank my parents, Reinhard and Hedwig, my brother Markus, my partner Alina, and all of my friends, especially Philipp and Stefan. You not only helped me finish this thesis but have been my best friends at all times. And for that, I am truly grateful. Thank you!

## 4. Abstract

Eastern Austria represents Austria's most important cereal production region, with winter wheat being the most widely grown arable crop achieving highest grain qualities. Farmers' management practices often rely on experience without considering difficult to access seasonal information, such as current soil water or crop nitrogen (N) status. Crop growth models are software tools which allow simulation of crop growth and development in response to environmental, soil, management, and cultivar-specific genetic factors. After comprehensive parameterisation accounting for region and specific cultivars, the models can deliver a range of highly relevant soil and crop information to support farmers' management decisions. So far, only a few modelling studies have addressed winter wheat in Eastern Austria, and many models either lack documentation, require numerous input parameters, or do not allow free-of-charge commercial application. Additionally, few models have found their way into user-friendly software-implementations which support farmers' and stakeholders' decisions. The aim of this PhD thesis was the detailed cultivar-specific parameterisation and evaluation of the well-documented, relatively simple, and freely available SSM-iCrop (iCrop) model for several modern winter wheat cultivars grown in multi-environment trials in Eastern Austria.

Two field experiments (2017/18 and 2018/19) were conducted in Tulln (Eastern Austria) using four local winter wheat cultivars (Arnold, Aurelius, Bernstein, Emilio) and four N-fertilisation levels (N0 to N3; 0 to 210 kg N ha<sup>-1</sup> in 70 kg N ha<sup>-1</sup> steps). Phenology was scored regularly, and crop and soil samples were collected and analysed on multiple dates. The collected data set was used for parameterisation, while an independent long-term data set from Eastern Austria was used for evaluation of the iCrop model regarding anthesis timing and yield.

While several calculated parameters did not show cultivar differences, including all crop N related parameters, cultivar-specific parameters were still found, exclusively for phenology, leaf area index (LAI), and dry mass (DM) partitioning between leaves and stems. iCrop simulated cultivar-specific phenology for the two field experiments (anthesis error: max. 5 days) as well as for the long-term evaluation data (RMSE: 2.9 days) with high accuracy. LAI simulations of the field experiments were moderately accurate (anthesis RMSE: 0.8). Using a biphasic stem N parameter, iCrop captured N-specific leaf and stem N content dynamics correctly. Observed differences in post-anthesis leaf traits (LAI, N content) indicated different stay-green behaviours of the cultivars. However, these were not captured by iCrop due to missing soil data and parameters (e.g. root exploration factor). Soil water simulations were partially excellent, at the same time they appeared to be sensitive to small changes in soil parameters of single layers. Soil mineral N content was systematically overestimated, indicated by total crop N overestimation at the

end of the season. Average leaf, stem, and grain N estimates were good, while cultivar-differences for grain N were poorly simulated. This was at least partially caused by biased N-dilution due to overestimation of crop total DM (percentage error at harvest: +13.7%) and yield (+6%). Simulated leaf:stem DM partitioning was poor, but the relative cultivar differences were represented well nevertheless. The model evaluation with long-term data was inconclusive for yield due to lack of soil data.

This study shows that the iCrop model can be used for successful cultivar-specific winter wheat simulation in Eastern Austria, thus enabling further scientific investigation and practical application. However, identified weaknesses of the model and options to improve these are given in the study. The parameterised iCrop model has been implemented in two research projects, COMBIRISK and Farm/IT, the latter directly delivering scientific advances to farmers and stakeholders in agriculture via modern software tools.

## 5. Introduction

### 5.1. Winter Wheat

#### 5.1.1. Wheat Production in Eastern Austria

Lower Austria is the largest of Austria's nine federated states, covering significant parts of the agricultural production region in Eastern Austria. Lower Austria's total area covers more than 1.9 Mha of which over 1.6 Mha were cultivated by farms and agricultural businesses in 2016. The largest land uses were forests with 43% closely followed by arable lands with 42% (over 682.000 ha). Winter wheat was mostly grown arable crop by area, covering 23% (over 157.000 ha) of arable lands, followed by grain maize (incl. corn-cob mix) with 11% (75.000 ha) and spring barley with 6% (39.000 ha) (Grüner Bericht Niederösterreich, 2018). Lower Austria accommodated over 61% of Austria's total winter wheat area (ca. 255.000 ha) and 57% (705.000 t) of the country's wheat yield in 2018. 77% of Austria's wheat production were classified as quality wheat (i.e. min. 14% grain protein), with best qualities found in Eastern Austria (Grüner Bericht, 2019).

#### 5.1.2. Morphology and Phenology of the Wheat Crop

Perry and Belford (2000) have neatly described the evolution and biology of the wheat crop, as subsumed in the following.

For over 10.000 years, cereals have evolved alongside humans, starting in the Middle East. The evolution of cereals from their low-yielding wild ancestors towards modern high yielding cultivars was a key to the development of humanity's society. All modern cereals are annual grasses, including wheat (*Triticum*), maize (*Zea*), barley (*Hordeum*), oat (*Avena*), rye (*Secale*), as well as the artificially made triticale (*Triticosecale*) (Perry and Belford, 2000).

While there are worldwide more than 30 species of wheat and over 40.000 cultivars, only three species are commercially relevant: club wheat (*Triticum compactum*), a hexaploid and usually soft-grained wheat used for cake flour, durum wheat (*Triticum durum* cv. *durum*), tetraploid hard wheat well known for its use in pasta, and the most popular common or bread wheat (*Triticum aestivum*) (Perry and Belford, 2000).

The main development phases of wheat (and cereals in general) are (i) germination and seedling establishment, (ii) initiation and development of leaves, (iii) tillering, (iv) root growth, (v) ear formation and growth, (vi) stem elongation, and (vii) flowering and grain growth (Perry and Belford, 2000).

Leaf development is initiated from primordia at the shoot apex. About three primordia are already present at germination, more being produced sequentially and on alternating sides of the apex during plant growth. The last developed leaf before the emergence of the ear is called flag leaf. All wheat leaves are relatively long and narrow, separated into a basal part encircling the stem called leaf sheath and the upper extension serving as main region for photosynthesis named leaf blade (Figure 1) (Perry and Belford, 2000).

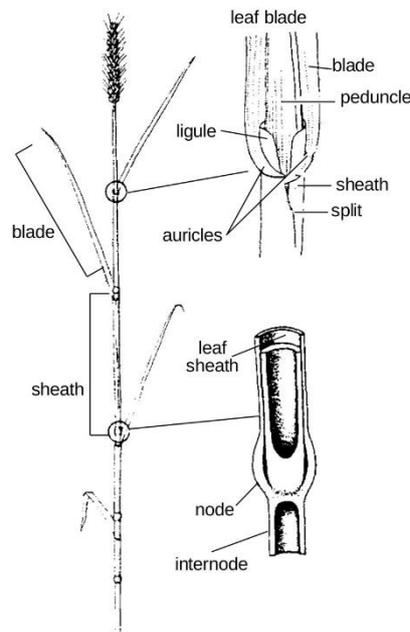


Figure 1 Morphology of the wheat plant (figure from "The wheat book", Perry and Belford, 2004).

The leaf appearance rate (phyllochron) of a wheat main stem is defined as the time a single leaf needs to fully emerge. It is strongly correlated with thermal time and much less correlated with chronological time (Gallagher, 1979). While chronological time refers to the count of hours or days, the thermal time is the sum of daily average temperatures above a base threshold over a specific period expressed in degree days ( $^{\circ}\text{C d}$ ).

Tillers are basal side shoots that originate from buds in the leaf axils of the main stem. Since they are virtually identical to the main stem, they have the potential to produce ears and grains. Leaves on tillers appear at  $90^{\circ}$  relative to the main stem leaves. Tillers are usually referred to by the number of the leaf axil on the main stem from which they appear (e.g. tiller one appearing from main stem leaf one). Primary tillers appear from the main stem, while secondary tillers appear from leaves on primary tillers. Tertiary tillers may appear but are rare. Sometimes tillers appear from the coleoptile (a modified leaf structure that protects and encloses the shoot during emergence) named coleoptile tillers (Perry and Belford, 2000).

The wheat stem consists of several nodes and hollow internodes and is wrapped in the sheaths of the surrounding leaves. The sheaths enhance the stem's mechanical strength, thereby reducing lodging. The nodes build regions where other structures (leaves, roots, tillers, spikelets) join the stem. The tissue between adjacent nodes is called internode. The young wheat plant appears to grow leaves from a single point. However, the leaves actually appear from 8-14 nodes which are closely stacked above each other, separated by very short (<1 mm) internodes. During the phenological phase named stem elongation, internodes begin to grow, thereby forming the characteristic tall wheat stem of the mature plant. A significant proportion of the stem's dry mass at ear emergence constitutes of carbohydrates and nutrients which can be translocated to grains during grain filling. Thereby, the stem fulfils both a mechanical/structural as well as a storage function (Perry and Belford, 2000).

The wheat ear (inflorescence) is a compound spike that consists of a rachis with spikelets attached on each side. The spikelets contain up to ten florets (flowers) of which only two to four usually develop into seeds (grains, kernels). Typical wheat ears develop between 30 to 50 grains (Perry and Belford, 2000).

#### 5.1.3. Agronomic Aspects of Wheat Production

Common agricultural practices for growing winter wheat in Eastern Austria are, for instance, sowing in mid-October, the use of recommended N-fertilisation rates, and herbicide application. Irrigation of wheat is mostly uncommon. Management strategies usually aim for high yields and possibly high grain protein content, depending on the target use of the harvest product such as fodder cereal (lower grain protein) or quality wheat (higher grain protein) for milling and baking.

Official references exist for N fertilisation application amounts and dates. The AGES (Austrian Federal Agency for Health and Food Safety) (Baumgarten et al., 2017) recommends winter wheat N fertilisation for average yield expectations (5 t ha<sup>-1</sup>, BMLRT, 2021) of 110-130 kg N ha<sup>-1</sup> split among three dates: (i) start of the vegetation period in spring (BBCH varies; ca. 21-24), (ii) beginning of stem elongation (BBCH 30), and (iii) prior to beginning of heading (BBCH 51). Apparently, all three of these dates are highly dependent on the wheat crop's phenological development. While the timing of the first N application (i) depends mainly on wheat base temperature above which plant physiological processes start (commonly assumed 0 °C; e.g. Porter and Gawith, 1999) which is considered mostly cultivar-independent, initiation of stem elongation (ii) and heading (iii) may vary significantly between cultivars and even more between seasons. Cultivar differences originate from genetically set durations of phenological phases and their reaction to environmental influence such as day length (photoperiod) and cold temperatures

(vernalisation). Seasonal differences result from environmental effects, mainly regarding year-to-year variation of temperature and precipitation. In practice, farmers have to observe their fields to determine the occurrence of critical phenological stages such as beginning of stem elongation to find the optimum date for management. Other seasonal variables such as weather forecasts and soil water status are usually not considered when N application is planned. Reasons for the exclusion of these informations may be their inaccessibility (mid-range skilful weather forecasts) or difficulty and expensiveness of their determination (soil sampling and determination of soil water content).

Crop growth models offer the possibility to estimate some of these data. Soil water content, for instance, can be estimated from past weather data over several months. Online software solutions can automatically integrate historical weather data as well as short and long-term weather forecasts into crop model simulations (Manschadi et al., 2019, 2020b). Farmers may adapt their fertilisation and/or irrigation strategy specifically to the current season based on the crop model simulation results (Boote, 2020; Van Evert, 2020). Overall, crop growth models are tools that have the potential to support informed decision making in agriculture (Sinclair and Seligman, 1996; Van Evert, 2020) by integrating numerous relevant data which can, due to their vast amount, hardly be taken into account manually.

## 5.2. Crop Growth Models

Modelling can be defined as the transformation of scientific understanding into a mathematical structure. Due to the infinite complexity of the real world, models necessarily include simplifications and logical assumptions. Otherwise, when allowing for too many details in the model, it becomes either vastly complex or even insolvable (Eck et al., 2011). Crop growth models, as a sub-type of models, calculate a crop's development and growth as influenced by a wide range of conditions. Essential components are energy input (intercepted solar radiation,  $\text{MJ m}^{-2} \text{d}^{-1}$ ) and output (crop growth rate,  $\text{g d}^{-1}$ ), time (growing period), and production of crop dry mass as a result of energy input, a PAR (photosynthetically active radiation) factor, and radiation use efficiency (RUE,  $\text{g MJ}^{-1}$ ). Beyond these variables, models usually include responses to temperature, day length, soil water and nutrient status (often nitrogen), precipitation, irrigation, fertilisation, and other management options such as sowing density. Specific model can also include effects of pests, diseases, and weeds (Soltani and Sinclair, 2012 cited in Fuchs, 2016).

In a previous work (Fuchs, 2016) I have summarised literature on the history of crop growth modelling (5.2.1) and the processes of parameterising and calibrating (5.2.2) crop models. In this PhD thesis, I have condensed this information in this section (5.2) and its subsections, adding supplementary information and references where appropriate.

### 5.2.1. Brief History of Crop Modelling

Crop modelling dates back to the 1970s (Passioura, 1996; Sinclair and Seligman, 1996) and its predecessors even to the 1960s when C.T. de Wit published a key report titled “Photosynthesis of leaf canopies” (Bouman et al., 1996; de Wit, 1965; Soltani et al., 2013). 25 years ago, Sinclair and Seligman (1996) described the development of crop models in the style of a human’s life, “from infancy to maturity”. Although a quarter century ago, their description is still up to date and, therefore, recapped in the following (with a few updates such as “smart farming”).

The first models were rather simple, but the idea of quantifying crop growth and development motivated researchers to quickly improve them. New findings in crop physiology as well as the prospect to enable prediction of yields, reduction of experiments, and evaluation of new genetic material lead to rapidly increasing numbers of parameters and model complexity. Awareness grew that this could not go on forever. The task of trying to model the complexity of a crop while avoiding an oversaturation of details led to extensive reductionism. Scientists tried to reduce as much as possible into basic physical, chemical, and physiological processes. The adverse effect was an increase in complexity without improvements of the predictions. The idea of developing universal models, applicable in each and every situation, was given up. However, the importance of model evaluation became clear. While models can, technically, not be validated due to the lack of a single, falsifiable hypothesis, the act of trying to evaluate a model by testing its performance in different situations gives model users at least an idea of where a model has proven useful and where not. Nowadays, crop models are being used as powerful aids in research, teaching, and smart farming. In models, knowledge of a crop is organized logically, where the correctness of assumptions can be verified or falsified. Simple and easy-to-understand models serve to teach students interactions such as the factors influencing crop production under different circumstances. In research, crop models are used to create sound concepts reflecting current knowledge. Moreover, they are tools for analysing experimental results by investigating the causes of differences. In early applications of smart farming, crop models have been used for pest management in cotton and wheat already in the 1980s (Sinclair and Seligman, 1996 cited in Fuchs, 2016).

Crop models can be classified into mechanistic (process-based) and empirical models. While process-based models try to imitate real world processes (on a physical and physiological basis), empirical models use mathematical functions chosen to best fit observations without consideration of the underlying processes (Monteith, 1996). However, process-based models also become empirical (i.e. purely mathematical) at a lower organizational layer (Sinclair and Seligman, 1996; Soltani and Sinclair, 2012).

Soltani and Sinclair (2015) compared four wheat models with different complexity, including SSM-iCrop (simple) and APSIM (complex). They found that the simpler models were generally more robust than the complex ones. Obviously, simpler models are advantageous as they are easier to parameterise and investigate. Therefore, it is not surprising that Passioura (1996) noted very early that models should be kept as simple as possible and demand only little input data.

#### 5.2.2. Model Parameterisation and Calibration

The input parameters of a mechanistic crop growth model represent important crop/cultivar characteristics necessary to model bio-physiological processes. Parameters are (mathematical) variables whose values are generally constant within a simulation, such as a crop's transpiration efficiency. Using observed data from literature and/or experimentation, parameter values are calculated for a specific genotype x environment combination. This procedure is referred to as "parameterisation". In contrast, the term "calibration" means randomly changing parameter values, selecting those fitting best to expectations such as observed yield (Sinclair and Seligman, 1996; Soltani and Sinclair, 2012). Some models need parameters which are not related to a measurable size and can only be calibrated by using the model itself. Since calibration is a purely "trial and error" method that ignores scientific knowledge on crop physiological processes and their interactions with the environment it should be avoided where possible (Monteith, 1996; Soltani and Sinclair, 2015, 2012). Therefore, calibration of a single parameter in relation to a measured size is acceptable, but calibration of the whole model for a specific output (e.g. to fit observed yields) is not. While some authors do not distinguish between calibration and parameterisation, in this thesis, the above-mentioned definition will be used.

There is no standardised approach on how to parameterise and calibrate a crop model. While there is a general need for high quality data sets to enable sound parameterisation (Rötter et al., 2011), the process itself is tackled differently. In a survey among crop modellers, Seidel et al. (2018) found that parameters are estimated mostly by trial and error in a multiple stage process, usually started with phenology parameters. The term "trial and error" would suggest a pure calibration exercise, in terms of the above-mentioned description. However, it can probably be safely assumed that most scientific modellers will base their trial and error approach on previously measured and calculated values, only adjusting them to optimise results. Apparently, the difference between the terms "parameterisation" and "calibration" is not a sharp line. However, in their survey, Seidel et al. (2018) found that among crop modellers the time spent to parameterise a model ranged from only one day to as much as five years (median 25 days) indicating the complexity and challenge of the parameterisation exercise.

### 5.2.3. Model Evaluation

In the context of crop growth modelling, the term “evaluation”, also called testing or validation (Boote et al., 1996), is the process of gaining confidence in the model itself or a specific application of the model, such as a newly derived set of parameters defining a previously untested cultivar. Frequently evaluated crop model outputs are the timing of emergence, anthesis, and maturity, leaf area index at anthesis, total crop biomass (dry mass) at anthesis and maturity, and grain yield. Depending on the subject of interest, other simulation results may be evaluated in detail, such as soil water content or grain nitrogen concentration.

Commonly applied methods to evaluate model results include qualitative assessments as well as quantitative tests. Concerning qualitative assessments, visual comparisons of simulated time series with observed points are frequently used in literature from early modelling (e.g. Monteith, 1965) until today (e.g. Asseng et al., 2019; Ebrahimi et al., 2016; Manschadi et al., 2020b; Webber et al., 2017). Field experiments carried out for crop modelling are usually designed with at least three replications to facilitate statistical testing. However, as exact mathematical models, crop models mostly lack any randomness, thereby making replications of the same simulation a useless task, as results would be identical. However, if several different environments are simulated, quantitative tests (i.e. statistical tests) can be used to assess model performance (e.g. Hunt and Boote, 1998 cited in Cao et al., 2012).

A quantitative method often used to describe crop model results is the goodness-of-fit (e.g. regression analysis simulated vs. observed). Sometimes the same dataset is used for parameterisation and evaluation, which leads to over-optimistic results (Efron 1983 cited in Seidel et al. 2018). Thus, it is better to have independent parameterisation and evaluation datasets (Seidel et al., 2018). In general, test statistics can be classified into two groups: those based on correlation, such as the correlation coefficient  $R^2$ , and those based on deviation, such as the root mean squared error (RMSE).

### 5.2.4. SSM-iCrop – a Process-based Crop Growth Model

Given the range of available crop models nowadays, the question “why make another model?” arises. The developers of SSM-iCrop (Soltani et al., 2013; Soltani and Sinclair, 2012) argued that previous models often (i) lack of documentation, (ii) are very complex, (iii) restrict access to model code, (iv) lack intuitive user interfaces, and (v) have parameters with unclear meaning. Regarding (v), one of the key goals of the developers was to exclusively use measurable parameters in SSM-iCrop. Other authors such as Seidel et al. (2018) supported the critique concerning the lack of documentation (i) and problems with high complexity (ii) among many models and

even claimed these issues to be of major importance during the process of model parameterisation.

The development of SSM-iCrop (SSM: Simple Simulation Models; hereafter referred to as “iCrop”) dates back more than 30 years when the first model version for soybean was described by Sinclair (1986). Since then, the model has been applied to a range of important field crops such as wheat (Sinclair and Amir, 1992; Soltani et al., 2013), maize (Sinclair and Muchow, 1995), barley (Wahbi and Sinclair, 2005), sorghum (Sinclair et al., 1997), peanut (Hammer et al., 1995), and chickpea (Soltani and Sinclair, 2011). In a study by Soltani and Sinclair (2015) the model’s performance was tested against three widely used and recognised crop models and iCrop was found to be transparent and robust. A simplified version of the model for large area simulations (SSM-iCrop2) was parameterised and tested for more than 30 crop species and has been used in large scale studies in different continents for various crop species (Sinclair et al., 2020; Soltani et al., 2020).

The version of iCrop presented and used in this PhD thesis was based on the original SSM-iCrop model as described in the book by Soltani and Sinclair (2012) and in the journal article by Soltani et al. (2013).

#### 5.2.5. Aspects of Wheat Simulation

In the following, some of the aspects of crop growth models simulating winter wheat are presented and different approaches are shown. A detailed description of the iCrop model itself is provided in the Materials and Methods chapter.

Soil is an essential part of a crop model, incorporating several crucial processes (Figure 2). While some soil-related variables are often completely ignored in some crop models (e.g. root biomass) others are usually included with varying detail (e.g. soil water content) depending on the variables of interest in the specific situation. For instance, simulations of yield potential might ignore soil water completely and just assume optimal conditions, while simulations aiming at deriving decision support for irrigation schemes will have to include detailed soil water representations. Equations and algorithms are coded in the soil sub-model in a way to represent soil water infiltration, surface runoff, water distribution, drainage, water and nitrogen uptake by the plant’s roots, nitrogen mineralisation, demineralisation, volatilisation, and more.

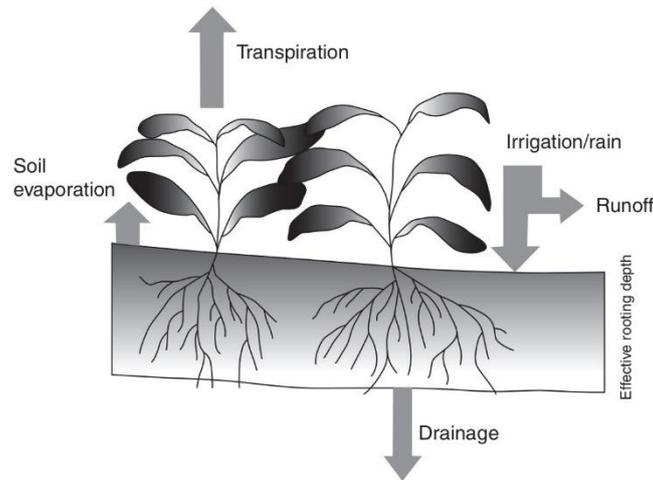


Figure 2 Conceptual overview of the soil and its in- and outputs (figure from "Modeling physiology of crop development, growth and yield", Soltani and Sinclair, 2012).

Implementations of soil water sub-models range from rather simple approaches such as "single-bucket" (Sinclair and Amir, 1992) or "multiple tipping-buckets" (Ritchie, 1998) to complex solutions such as the well-known Richards equation (Richards, 1931). The multiple tipping-buckets approach, also known as cascading layers approach, is among the most frequently applied solutions. It is based on the concept that water that exceeds field capacity (i.e. drained upper limit) in one layer infiltrates the following layer below. Formulas to limit the water flow to soil layers based on hydraulic conductivity also exist (Stöckle and Meza, 2020).

Soil water content highly depends on infiltration of water from precipitation and irrigation. Most crop models allow for both water input from weather data (precipitation) and water input from irrigation as a management option. Water that does not immediately infiltrate the soil may be accumulated on the soil surface or simply be removed as runoff. While only few models account for accumulated water on the soil surface, most do calculate runoff, often using the USDA-SCS curve number method (USDA, 2004) that accounts for slope, soil texture, tillage, amount of precipitation, and more (Stöckle and Meza, 2020). The method may be simplified to require less input parameters (e.g. as used in iCrop (Soltani and Sinclair, 2012) where slope, texture, and tillage are ignored).

Soil water in the topmost layer is prone to evaporation, i.e. water flux from soil to the atmosphere. Several factors affect soil evaporation, including soil water availability, atmospheric evaporative demand, crop cover (leaf area), crop residue, and tillage. According to Stöckle and Meza (2020) only few models use mechanistic approaches which estimate fluxes of heat and water in the soil that result in many finite difference equations. Most crop models apply more simple approaches such as the two-stage evaporation proposed by Ritchie (1972). Stage I is energy limited and captures the atmospheric evaporative demand (including e.g. crop cover, residue cover)

and stage 2 water limited describing the reduction of the soil's evaporative rate (often a function of time or soil water content). Usually, 0.1 m of the topmost layer are affected by evaporation, but this may be increased for sandy soils or decreased for clay-rich soils.

Crop water uptake is often modelled under the assumption of uniform root distribution in the soil layers, meaning that all roots are considered having the same water uptake capacity with always sufficient root length density for water uptake. The simulation of vertical downward root growth incorporates crop/cultivar-specific optimum root extension rate, accounting for limiting factors such as soil temperature and stresses regarding water and nutrients (Wang and Smith, 2004). However, the roots' water uptake capacity in reality is highly dynamic.

Soil water taken up by the roots is transported to the leaves and transpired. Again, approaches with varying complexity exist to model the process of transpiration. Stöckle and Meza (2020) reviewed literature and found that relatively few studies exist which compare different water uptake algorithms and that significant differences may result from different approaches. However, while they argued that nowadays, with high computing power compared to several decades ago, more complex approaches such as the SPAC (soil-plant-atmosphere continuum) framework (e.g. Campbell et al., 1976) would give better results they also acknowledged that most crop models implement rather simple approaches. One such simple method is relating transpiration to daily crop dry matter production ( $\text{g m}^{-2} \text{day}^{-1}$ ), the atmosphere's vapour pressure deficit (VPD, Pa), and a species-specific transpiration efficiency coefficient (TEC, Pa) (Soltani and Sinclair, 2012).

Nitrogen (N) is quantitatively the most important element in plant nutrition (when ignoring hydrogen and oxygen from carbon dioxide assimilation and water uptake). Between 2014 and 2015, most of the world's N fertilisers were applied to maize (19.3%) and wheat (18.5%) (Heffer et al., 2017). Since the invention of the Haber-Bosch process in the early 20<sup>th</sup> century, N fertiliser use has increased highly – and so have wheat yields and grain protein contents (Cao et al., 2018; Follett et al., 2010; Sinclair, 1998). Of course, besides improved N supply to the crop, other advances contributed to this as well (e.g. breeding, technical progress). Nevertheless, the plant available N resource poses a major potential – or limitation. Therefore, crop models need to simulate soil N dynamics and interactions with the crop. Over 50 soil N models for crop models exist, most of them process-based and capable of additionally simulating soil carbon dynamics (Singh and Porter, 2020). The modelling of nitrogen in crop growth models can be divided into crop N demand and soil N supply. The latter comprises of plant available soil N, that is mineral soil N (N<sub>min</sub>; ammonium and nitrate), which is used by most models (Shaffer et al., 2001 cited in Singh and Porter, 2020).

The empirical relation between temperature and the rate of plant physiological processes such as leaf expansion, organ emergence, or the timing of anthesis has been modelled for more than 100 years (Kim et al., 2020). Temperature drives plant development from the very beginning (sowing) until the very end (harvest maturity, crop death). Thermal time is a concept where daily temperatures above a threshold (base temperature) are summed up and used for crop model calculations. The critical question is which temperature to sum up, as the diurnal pattern of temperature is usually far from constant. Also, different crop species react differently to temperature. One of the earliest concepts to condense daily temperature into a single, crop species-specific value was the growing degree day (GDD). McMaster and Wilhelm (1997) used the following formula:

$$GDD = \frac{T_{max} - T_{min}}{2} - T_{base} \quad (1)$$

with  $T_{max}$  and  $T_{min}$  being daily maximum and minimum temperature, respectively, and  $T_{base}$  being a crop specific minimum temperature below which plant development is stopped. GDDs can be cumulated, commonly referred to as thermal time (see below).

While GDD was a big step forward in crop modelling, it did not explain plant phenology sufficiently well. According to Kim et al. (2020), several approaches were found to better describe the influence of temperature on plant development (e.g. Yin et al., 1995). One of them was the “cardinal temperature approach” where  $T_{base}$ ,  $T_{opt}$ , and  $T_{crit}$  are defined: Plant development only starts if temperature is above  $T_{base}$ . The development rate rises until  $T_{opt}$ , then decreases until  $T_{crit}$  after which it ceases completely. While these cardinal temperatures may be interpolated linearly, other solutions were found to work better. One well recognised approach to interpolate between the cardinal temperatures is using a curvilinear relationship (beta function, Figure 3) (Kim et al., 2012; Yan and Hunt, 1999).

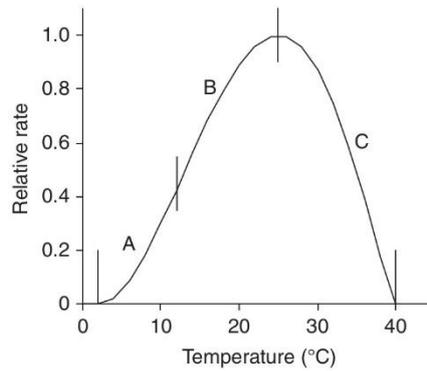


Figure 3 Schematic curvilinear relation of the phenological relative response to temperature with various response regions. Starting point of A:  $T_{base}$  breakpoint B/C:  $T_{opt}$  ending point of C:  $T_{crit}$ . Figure from "Modeling physiology of crop development, growth and yield" (Soltani and Sinclair 2012).

The influence of photoperiod (day-length) and vernalisation (winter chilling) was discovered early in the 19<sup>th</sup> century (e.g. Coville, 1920 cited in Kim et al., 2020). The photoperiod affects induction of anthesis (flowering) of plants: short-day plants require day-length below a certain threshold, long-day plants above a threshold, and day-neutral plants do not react to day-length. The threshold is often defined as 12 hours. Vernalisation represents the need of some plant species for a certain duration of cold temperatures to induce anthesis, otherwise, they stay in a vegetative state.

In terms of crop modelling, vernalisation and photoperiod are often implemented by modifying the accumulated thermal time using factors (0-1) (Kim et al., 2020). For instance, a long-day crop experiencing too short day-lengths will cause the model to reduce the thermal time (photoperiod factor below 1). In the same manner, for a winter wheat crop that experiences rather high temperatures (e.g. 10 °C) the crop model will also reduce thermal time (vernalisation factor below 1), thereby slowing the progress of phenology. Usually, such factors are only active during specific phases of a plant's development, otherwise they are set to 1.

Depending on the purpose of a crop model, different phenological stages will be defined and programmed into the model. For example, Soltani and Sinclair (2012) used the following relevant phenological stages for wheat: emergence (50% of plants have emerged from the soil surface), termination of leaf growth on the main stem (end of the effective leaf production), beginning and termination of seed growth (beginning and end of effective grain mass increase), and harvest maturity (reduction of seed water content to a level where machine harvest is sensible). Soltani and Sinclair (2012) also state that additional stages may be simulated, for instance to support crop management decisions regarding timing of the application of fertilisers and pesticides.

In technical terms, the phenological stages that need to be simulated are defined as thermal time targets. For instance, winter wheat might require 120 °Cd for emergence and a total of 2000 °Cd from emergence until harvest maturity. However, these values are cultivar-specific and require thorough parameterisation.

The simulation of leaf canopy variables, particularly of the development of leaf area index (LAI), has a significant impact on intercepted radiation and dry matter production and, therefore, on grain yield (Sinclair, 1984; Soltani and Sinclair, 2012). Due to the practical importance of yield quantity, the exact simulation of LAI is obviously a key in crop growth models.

Soltani and Sinclair (2012) have reviewed methods for simulating LAI and identified three different approaches: (i) carbon-based, (ii) temperature-based, and (iii) hybrid (carbon and temperature). Carbon-based approaches (i) assume that dry matter allocated to leaves has the most significant influence on LAI development. The daily increase in leaf area is calculated from daily dry matter increase for leaves times the specific leaf area – a coefficient representing leaf area per unit of leaf dry matter. Models using the temperature-based approach (ii) relate LAI growth to temperature. Hybrid methods (iii) use a mixture, for instance calculating both carbon-limited and temperature-limited LAI increase and using the minimum of both.

Main stem node number simulation is usually based on the phyllochron concept. It is important in two ways: first, as a measure of phenological development (e.g. Haun-stage and Zadoks-stage, Haun, 1973; Zadoks et al., 1974) and second as a driver of LAI expansion, both in reality and in the model.

### 5.3. Research Gap Filled by this Thesis

Crop growth models have been applied in various situations all around the planet (e.g. Reynolds et al., 2018). However, models are not “generic”, meaning that a single crop model parameterisation cannot be reliably used anywhere in the world. Rather, each parameterisation is only valid within the ranges of its parameterisation data base, and extrapolation to other environments is very limited. For example, a crop growth model parameterised for winter wheat in a wet environment can hardly be used to simulate barley in a dry environment, or even spring wheat in a semi-dry environment.

As of my knowledge, only few studies have been carried out using crop growth models for cereals in Eastern Austria (e.g. Ebrahimi et al., 2016; Eitzinger et al., 2013a, 2013b). None of them were addressing multiple modern winter wheat varieties with varying N-fertilisation levels using a highly detailed and comprehensive data set as later presented in my study. Also, none of the

previous studies were using the iCrop model in Eastern Austria. While the necessity to parameterise crop models for up-to-date cultivars is obvious, it could be argued that other models are more readily available and more intensely tested than iCrop and should be used instead. However, there are several good reasons to use this specific crop model. First, its source code is freely available without any licensing, thereby enabling its use for scientific, commercial, and combined applications which bridge the gap between research and industry. For instance, the results of this study contributed to the research project COMBIRISK (COMBIned weather related RISK assessment monitor for tailoring climate change adaptation in Austrian crop production, <https://combirisk.boku.ac.at/>) as well as to the Farm/IT project (<https://www.farmit.at/>, Manschadi et al., 2020b, 2019) which focused on implementing the crop modelling research results to software applications for stakeholders in agriculture. Also, iCrop is exceptionally well documented in a book (Soltani and Sinclair, 2012) which can, unfortunately, not be taken for granted (Asseng et al., 2020). Another very handy aspect of iCrop is its ease of use. As described later in the Materials and Methods section, iCrop is written in a rather simple programming language in a macro within Microsoft Excel (Microsoft Corporation, 2018) which enables easy alteration of the model's algorithms and testing scientific hypotheses. Finally, an additional model can in future be used in multi-model ensembles which have received increased attention in recent years for highly significant topics such as climate change impact studies (e.g. Asseng et al., 2019, 2013; Basso et al., 2018; Liu et al., 2019; Rosenzweig et al., 2013). It has been shown that the performance of the single models within the ensemble is a key to decrease the variation of the overall ensemble simulation results (Maiorano et al., 2017).

By parameterisation and evaluation of the iCrop model in a previously rarely tested but agriculturally important environment my study contributes to filling a gap in a highly relevant and up-to-date scientific topic.

## 6. Research Question and Hypotheses

### Research Question

Is the iCrop crop growth model capable of capturing winter wheat canopy growth and development in the environment of Eastern Austria using field data for detailed model parameterisation?

### Hypotheses

In the environment of Eastern Austria iCrop is capable of simulating:

1. Soil mineral N content (N<sub>min</sub>) and soil water content
2. Winter wheat cultivar-specific phenological development
3. Winter wheat cultivar-specific and N-fertilisation-specific crop dry mass and crop-N dynamics at organ-level (leaf, stem, grain)

## 7. Materials and Methods

In this chapter, the materials and methods used to parameterise the iCrop model are described, starting with data acquisition from field experimentation (7.1). Section 7.3 gives a detailed description of the specifics of the iCrop model itself, and section 7.4 explains the calculation of parameters from observed data.

### 7.1. Field Experiments

Plant and soil data were collected from two field experiments conducted in 2017/18 (EXP1) and 2018/19 (EXP2) at the UFT (Universitäts- und Forschungszentrum, Tulln, BOKU university), Lower Austria (48 ° 19' N, 16 ° 04' E, 178 m a.s.l.) (Figure 4). The two experiments were in ca. 500 m distance from each other. The location of EXP2 was changed to another field since the data of EXP1 had shown inhomogeneous soil conditions. Also, the closeness to a windbreak (Figure 5) of EXP1 was unfavourable.

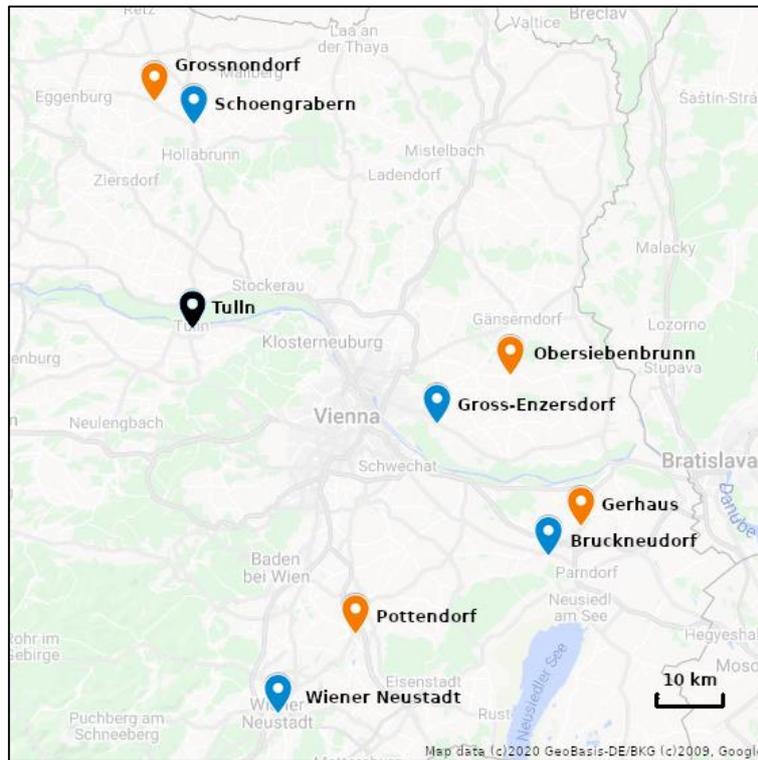


Figure 4 Map of parameterisation experiment location (black symbol), long-term experiment locations (orange symbols) and corresponding weather station locations (blue symbols). The “Langenlebar” weather station (ca. 4 km east of Tulln) is not shown. Map source: [www.google.at/maps](http://www.google.at/maps).

Four widely-grown winter wheat (*Triticum aestivum* L.) cultivars, contrasting in phenology and grain protein content (Arnold, Aurelius, Bernstein, and Emilio), were sown on 16 October 2017 in EXP1 and on 15 October 2018 in EXP2. The sowing density was 375 seeds m<sup>-2</sup> with a row spacing of 120 mm sown in 10 rows (single-plot size: 10 x 1.2 m). Single-plots were sown twice in parallel to increase area for destructive sampling, so the final plot size was ca. 10 x 3 m (including wheel tracks). To minimise border effects from different N-fertiliser treatments, buffer plots of the same size were added. Also, the previous crops (wheat in EXP1, maize in EXP2) were unfertilised to reach low initial soil N levels. Although both experiments were in close proximity to each other, the soil map “eBod” (<https://bodenkarte.at>; Wandl and Horvath, n.d.) classifies the soils of the two locations differently: EXP1 as “Anmoor” (calcaric mollic gleysol) and EXP2 as “Feuchtschwarzerde” (calcaric gleyic phaeozem). The reference weather station was located at Langenlebar (ca. 4 km east of the field experiments). Long term weather averages (period: 1991 to 2015) at Langenlebar during the wheat growing season (October – June) were 408 mm precipitation and 7.9 °C average temperature.

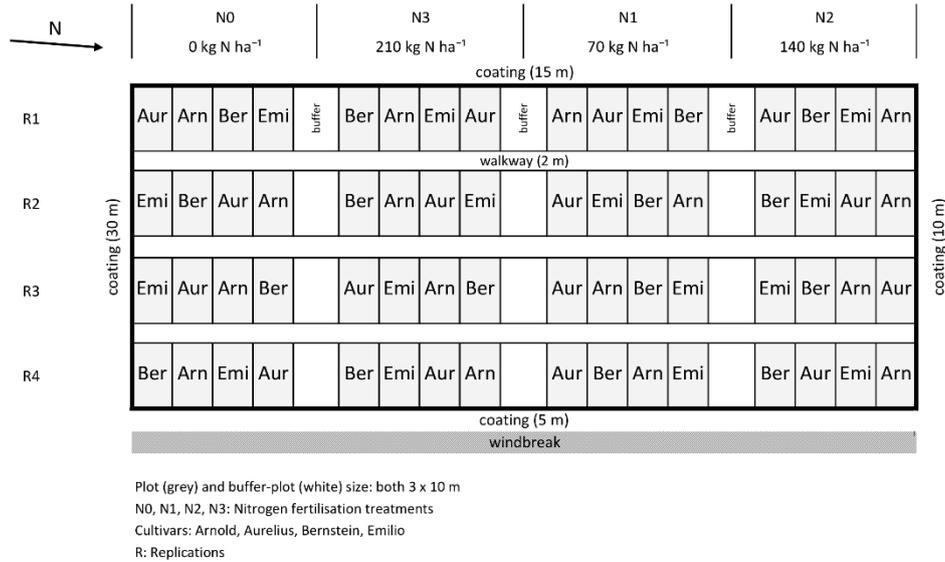


Figure 5 Experimental design of the winter wheat field experiment 2017/18 at Tulln (EXPI). Winter wheat cultivars: Arnold, Aurelius, Bernstein, and Emilio.

The experimental design was a randomised split-plot with fixed N-treatment blocks and four replications (shown for EXP1 in Figure 5). The layout of EXP2 was identical to EXP1, only the (sub) plots within the main-plots (i.e. replication x N-treatment) were newly randomised. Also, there was no windbreak nearby in EXP2. For phenological and destructive crop samples in both experiments, replication R4 was dropped due to obvious inhomogeneity. Both experiments received a nitrogen (N) fertilisation treatment with four different levels: 0 (N0), 70 (N1), 140 (N2), and 210 kg N ha<sup>-1</sup> (N3). N fertiliser was applied in two equal rates on 9 April (BBCH 26, tillering) and 8 May (BBCH 43, booting) in EXP1 and on 8 March (BBCH 22, tillering) and 10 April (BBCH 26-31, tillering/stem elongation) in EXP2. The coating area was sown with Bernstein and received N2 fertilisation, excluding a buffer zone of 2 m next to the N0 plots. Both experiments were rain-fed exclusively. Prior to sowing, a phosphorus-potassium fertiliser (15% P, 40% K) was applied at the rate of 300 kg fertiliser ha<sup>-1</sup>. Control of pests, diseases, and weeds was managed according to common regional practices.

In both experiments, three plants per plot (from all N-treatments in EXP1, and only N0 and N3 in EXP2) were chosen randomly, marked with small plastic markers, and used for weekly scorings of phenology according to the BBCH scale (Meier, 2001) during the main active growing period (i.e. excluding winter, ca. December – February). The phenological stage of a treatment was calculated as follows: If all three plants were in the same BBCH main stage, the average of all three plants was used. If one plant was in a different stage than the other two, the one plant's value was removed and the average of the two other plants was used. This is in compliance with the BBCH scoring method, which states that a main stage is reached if at least 50% of the plants

in the canopy show that stage. For instance, if the three plants were scored 12.3, 13.4, and 21, then 21 was removed and the average of 12.3 and 13.4 was calculated. It did not occur that each of the three plants was in a different main stage. Main stages are: 0 germination, 1 seedling growth, 2 tillering, 3 stem elongation, 4 booting, 5 ear emergence from boot (heading), 6 anthesis, 7 milk development, 8 dough development, and 9 ripening.

Sequential destructive plant samples were cut directly above soil surface and analysed for crop dry mass, leaf area development, N-uptake, and grain yield. Sampling area was 0.25 m<sup>2</sup> for in-season measurements and 1.00 m<sup>2</sup> at final harvest. Plant biomass samples were separated into photosynthetic active leaf blades (i.e. green leaves) and senesced leaf blades (area threshold: 50%), stems (including leaf sheaths), chaff, and grains. For measuring leaf area index (LAI), a subsample of the green leaves was immediately taken and analysed (LI-3100C Area meter, LICOR, USA). All samples were dried for 72 h at 60 °C to retrieve dry mass. The N concentration of plant tissues was measured using a CN elemental analyser (Vario Macro Cube, Elementar Analysensysteme GmbH, Germany).

Table 1 Sampling date, observed phenology (BBCH), and phenological stage name used in text when referencing to specific sampling dates (sowing, anthesis, harvest) during the two winter wheat seasons. Sample types: biomass (i.e. above-ground biomass for measuring dry mass), SW (soil water), and Nmin (soil mineral nitrogen). Seasons: EXP1: 2017/18, EXP2: 2018/19.

Season	Sampling date	Reference in text	BBCH	Sample type		
				Biomass	SW	Nmin
EXP1	13 September 2017				p	p
	18 October 2017	Sowing	<9		p	
	30/31 January 2018			P	p	
	26 March 2018			p		
	11/12 April 2018			p	p	
	30 April/2 May 2018			f	p	
	28/30 May 2018	Anthesis	ca. 71	f	p	
	4 July 2018	Harvest	>89	f	p	p
EXP2	1 October 2018				p	p
	18 October 2018	Sowing	<9		p	
	15 February 2019			f		
	4/5 March 2019			f	p	p
	20/21 March 2019			f	p	
	1/2 April 2019			p	p	
	15/16 April 2019			f	p	
	2 May 2019			f	p	
	17 May 2019			p	p	
	3 June 2019	Anthesis	62-69	f	p	
	11 June 2019			p		
	1 July 2019				p	
	16 July 2019	Harvest	>89	f	p	p

p: partial sample, only selected plots or plant organs were sampled; f: full sample, all plots were sampled

Gravimetric soil water content was measured at sowing, anthesis, harvest, and on several additional dates (Table 1). The sample collection procedure involved driving augers 120 cm deep into the soil. After pulling them out, the soil samples were divided into 0-10, 10-30, 30-60, 60-90, and 90-120 cm soil layers. On each sampling date at least two augers were extracted per plot. The layer-wise corresponding samples from the augers were then mixed homogeneously, packed in plastic bags and immediately stored in dark, cooled containers. Soil fresh weight was then measured as soon as the samples were brought to the laboratory. Afterwards, they were dried at 105 °C and again weighed for dry weight. In EXP 1, in-season soil water samples were taken along a transect, resulting in rather random selection of treatments: Arnold N0, Bernstein N2, Emilio N1, and Emilio N3. At harvest of EXP1, all treatments were sampled, using a mix of replication R2 and R3 to save costs. However, since the transect-method did not prove advantageous, in-season soil water samples of EXP2 were always taken from the same two to three replications of Arnold and Bernstein with treatment N0 and N3 for better comparability between cultivars and

N-treatments. At harvest of EXP2, all treatments were sampled, using two replications (not mixed). Soil bulk density was measured on a nearby (ca. 1 km distance) field and adjusted using publicly available soil data from “eBod” (Wandl and Horvath, n.d.) as well as data from a neighbouring experiment (spatially in between EXP1 and EXP2) (Sethmacher, 2018). The on-field measurement was done by digging a pit and using metal sampling rings (35 mm radius, 50 mm length) to extract undisturbed soil samples from the corresponding soil layers. The soil bulk density samples were dried for 72 h at 105 °C and weighed. Soil bulk density (BD, g cm<sup>-3</sup>) was calculated following a simple division:

$$BD = \frac{m}{V} \quad (2)$$

where m is the soil dry weight (g) and V the cylinder volume (cm<sup>3</sup>).

Volumetric soil water content ( $SW_{VOL}$ ) was calculated from gravimetric soil water content ( $SW_{GRAV}$ ) and bulk density:

$$SW_{VOL} = SW_{GRAV} \cdot BD \quad (3)$$

$$SW_{GRAV} = \frac{m_{fresh} - m_{dry}}{m_{dry}} \quad (4)$$

where  $m_{fresh}$  and  $m_{dry}$  are the soil masses before and after drying, respectively.

Soil mineral nitrogen content ( $N_{min}$ ) was measured pre-sowing and at harvest in both experiments, and additionally at early tillering in EXP2 (ca. BBCH 22, 4 March 2019) for Arnold and Bernstein, each with N0 and N3. In both experiments,  $N_{min}$  samples pre-sowing were taken from four equidistant points along a transect. The values were pooled and used for identical soil  $N_{min}$  initials in the simulations. Harvest samples of both experiments were collected from two replications of all treatments. Only in EXP1 the replications of these samples were mixed before analysis to save costs, therefore only one value per treatment was available. For both experiments, the splitting of soil  $N_{min}$  samples into layers was similar to that for gravimetric soil water samples, only the first two layers were merged: 0-30, 30-60, 60-90, 90-120 cm. Material (augers) and procedure (min. two augers per measurement) was identical to that for soil water (see above). The samples were sent to a commercial soil laboratory for analysing  $N_{min}$  and organic carbon content.

The soil  $N_{min}$  balance was calculated as follows:

$$N_{Bal} = Ini_{Nmin} + Fert_N - N_{UP} - Fin_{Nmin} \quad (5)$$

where  $N_{\text{Bal}}$  ( $\text{kg N ha}^{-1}$ ) is the soil Nmin balance,  $Ini_{\text{Nmin}}$  ( $\text{kg N ha}^{-1}$ ) the soil initial Nmin at sowing,  $Fert_{\text{N}}$  ( $\text{kg N ha}^{-1}$ ) the total fertilised N during the season,  $N_{\text{UP}}$  ( $\text{kg N ha}^{-1}$ ) the N content of the crop at harvest, and  $Fin_{\text{Nmin}}$  ( $\text{kg N ha}^{-1}$ ) the final soil Nmin at harvest.

## 7.2. Long Term Wheat Data Set

Model evaluation was performed using independent field experimental data from AGES (Austrian Federal Agency for Health and Food Safety). Data covered four locations across Eastern Austria (Grossnondorf, Obersiebenbrunn, Gerhaus, Pottendorf), three seasons (2014/15, 2015/16, and 2016/17), and the same four winter wheat cultivars as used for parameterisation (Arnold, Aurelius, Bernstein, Emilio).

Management and soil data availability was limited. Data included the dates of sowing, heading (BBCH 59), and harvest, as well as grain yield and location (city name, but not coordinates). Also, dates and amounts of N fertiliser application were available. All experiments received mineral N fertilisation (no organic fertilisation). On a few occasions, experiments were irrigated (amount and date available).

Weather data was acquired from ZAMG (Central Institution for Meteorology and Geodynamics, a subordinate agency of the Austrian Ministry of Education, Science and Research). The locations of the weather stations vs. field experiments are shown in Figure 4. The closest available weather station was used for simulation of each experiment (ca. 10 to 25 km distance). On a few single days, weather data was missing. In these cases, missing data was substituted with data from the next nearest station.

## 7.3. The “iCrop” Crop Growth Model

### 7.3.1. Overview

iCrop is a process-based crop growth model that simulates phenology, leaf growth and senescence, biomass (above-ground dry mass) production and partitioning, crop nitrogen uptake and partitioning, yield formation, and soil water and nitrogen dynamics. The model operates on a daily time step but also provides the option of hourly time steps for calculating the response of transpiration to vapour pressure deficit. Required model inputs are daily weather data (minimum and maximum temperature, global radiation, precipitation), soil parameters and initials, management settings (sowing date and density, irrigation dates and amounts, fertilisation dates and amounts), and crop/cultivar genetic parameters (phenology, leaf characteristics, dry mass production, nitrogen concentrations) (Soltani et al., 2013; Soltani and Sinclair, 2012).

The iCrop model is implemented as a macro in Microsoft (MS) Excel (Microsoft Corporation, 2018) written in the programming language Visual Basic for Applications (VBA). The implementation into MS Excel allows easy access and usability. All necessary model inputs and outputs are entered and displayed directly in the sheets of a single MS Excel file. The file may be defined to simulate multiple seasons, crops, and locations. It is not necessary to have programming skills for the normal use of iCrop. However, for advanced users it is possible to test the model extensively. Model functionality can be modified easily via the VBA source code, requiring no additional software besides MS Excel. Thereby, scientific hypotheses can be implemented and tested smoothly.

Since the publication of the current version of the model in 2012, iCrop has received some updates and improvements from its developers. Also, the model provides simplified modes for non-limiting conditions regarding soil water and soil nitrogen, where some functions differ from the complete (i.e. water and nitrogen limited) model. However, in the following sections the model is described in its 2012 version including all functionality unless otherwise noted. The main references are the publications of Soltani and Sinclair (2012) and Soltani et al. (2013) and the model code itself, unless explicitly cited otherwise. The contents of the following sections were largely paraphrased from Soltani and Sinclair (2012).

Hereafter, model parameters, variables, and outputs are formatted in italics (e.g. *ppfun*, *WSFD*). In some sections, the italics and normal form of an abbreviation are used next to each other to help distinguishing between model simulations and field measurements/observations (e.g. simulated *LAI* vs. observed LAI).

### 7.3.2. Phenology

iCrop simulates important phenological stages of the wheat crop: emergence, first tiller, first node (i.e. first node on the main stem), flag leaf ligule emergence (i.e. full expansion of the flag leaf), ear emergence, anthesis, and physiological maturity. The development is based on the calculation of biological days (parameter name: *bd<phenology\_phase>*; e.g. *bdEMRTIL* for emergence to tillering, unit: bd), where one bd represents a day with optimal growing conditions (regarding temperature, photoperiod, vernalisation, and soil water availability). Default optimum temperature for wheat is 27.5 °C. For each day, the model calculates

$$bd = tempfun \cdot ppfun \cdot verfun \cdot WSFD \quad (6)$$

using sub-functions for photoperiod (*ppfun*), vernalisation (*verfun*), water stress (*WSFD*, see section Soil Water), and temperature (*tempfun*). When cumulative biological days pass the defined

starting threshold of a phenological phase, that phase is initiated. For instance, the wheat cultivar Tajan requires 5 cumulative biological days from sowing to emergence. The set of  $bd<phenology\_phase>$  thresholds for a cultivar needs to be defined as a part of the parameterisation process (Soltani et al., 2013; Soltani and Sinclair, 2012).

The temperature function ( $tempfun$ ) simulates the effect of daily mean temperature on phenological development (Yan and Hunt, 1999):

$$tempfun = \frac{T_c - T}{T_c - T_o} \cdot \left( \frac{T - T_b}{T_o - T_b} \right)^{\frac{T_o - T_b}{T_c - T_o}} \quad (7)$$

with daily mean temperature ( $T$ ), wheat base temperature ( $T_b$ ), optimum temperature ( $T_o$ ), and critical maximum temperature ( $T_c$ ). Wang et al. (2017; 1998) proposed a more complex temperature function, claiming it improved phenology and biomass simulation of many well-known models. However, the function proposed by Wang et al. is almost identical to the one used in iCrop, so the model appears up to date.

Wheat phenological stages sensitive to day length (photoperiod) are influenced by the photoperiod function (Soltani et al., 2006):

$$ppfun = 1 \text{ if } PP \geq CPP \quad (8)$$

$$ppfun = 1 - ppsen \cdot (CPP - PP)^2 \text{ if } PP < CPP \quad (9)$$

where  $PP$  is the photoperiod ( $\text{h d}^{-1}$ ),  $CPP$  the critical (lower) photoperiod, and  $ppsen$  the cultivar-specific coefficient for photoperiod sensitivity.

The vernalisation function is calculated as

$$verfun = 1 - vsen \cdot (VDSAT - CUMVER_i) \text{ if } CUMVER_i < VDSAT \quad (10)$$

$$verfun = 1 \text{ if } CUMVER_i \geq VDSAT \quad (11)$$

where  $vsen$  is the cultivar-specific sensitivity coefficient,  $VDSAT$  the required cumulative vernalisation demand (50 days for all wheat cultivars), and  $CUMVER_i$  the cumulative vernalisation up to the current day  $i$ :

$$CUMVER_i = CUMVER_{i-1} + VERDAY \quad (12)$$

where  $VERDAY$  is today's vernalisation, which ranges from 0 (no vernalisation) to 1 (optimum vernalisation). Optimum vernalisation temperature in wheat ranges from 0 to 8 °C, while tem-

peratures below -1 °C and above 12 °C do not contribute to vernalisation. Suboptimum vernalisation occurs for temperatures from 0 to -1 and from 8 to 12 °C following a linear decrease (Soltani et al., 2013; Soltani and Sinclair, 2012).

Vernalisation can be reversed under certain circumstances. This is called de-vernalisation and occurs only in early stages when the crop has experienced less than 10 days of cumulative vernalisation and at high maximum air temperatures (above 30 °C).

Both the response to photoperiod and the response to vernalisation occur only from emergence until the appearance of the first node during stem elongation (i.e. ca. BBCH 31).

The water stress deficit factor (*WSFD*) ranges from 1 to 1.4 and accounts for the fastening of crop development during drought stress (Soltani et al., 2013; Soltani and Sinclair, 2012).

### 7.3.3. Leaf Development

Leaf development simulation covers leaf (i.e. node) number on the main stem (*MSNN*, dimensionless) and green leaf area index (*LAI*, m<sup>2</sup> m<sup>-2</sup>, i.e. m<sup>2</sup> leaf area m<sup>-2</sup> soil surface). The environmental factors soil water availability and soil mineral nitrogen availability, photoperiod, vernalisation, and temperature affect leaf development through various processes, as described below (Soltani et al., 2013; Soltani and Sinclair, 2012).

*MSNN* is calculated using the phyllochron approach. A constant, cultivar-specific value of phyllochron is assumed from emergence until the appearance of the flag leaf (note that this has been changed in the iCrop version used in this thesis). During this phenological phase, potential leaf area index (*PLAI*, m<sup>2</sup> m<sup>-2</sup>) directly depends on leaf number, using a simple power function:

$$PLAI = PLACON \cdot MSNN^{PLAPOW} \quad (13)$$

with *PLACON* as a constant (default: 1). The exponent, *PLAPOW* (dimensionless), is a cultivar-specific coefficient which is adjusted for plant density. Daily leaf area increase (*GLAI*, m<sup>2</sup> m<sup>-2</sup>) is then just the difference between today's (i) and yesterday's (i-1) *PLAI* considering water stress:

$$GLAI = (PLAI_i - PLAI_{i-1}) \cdot WSFL \quad (14)$$

where *WSFL* is the water stress factor for leaf area expansion (see 7.3.7 Soil Water).

In addition, *GLAI* is limited to daily available nitrogen to leaves (*INLF*, g m<sup>-2</sup> soil surface) in relation to specific nitrogen of green leaves (*SLNG*, g m<sup>-2</sup> leaf area):

$$GLAI = \max\left(\left(\frac{INLF}{SLNG}\right), (PLAI_i - PLAI_{i-1}) \cdot WSFL\right) \quad (15)$$

After full expansion of the flag leaf and up until the beginning of seed growth, the model assumes a slight continuation of *LAI* increase due to tiller leaf expansion:

$$GLAI = GLF \cdot SLA \quad (16)$$

where *GLF* (g m<sup>-2</sup>) is the dry matter allocated to leaves and *SLA* (m<sup>2</sup> leaf area g<sup>-1</sup>) is the specific leaf area.

Simulated leaf area index is also limited by nitrogen availability via leaf area senescence:

$$DLAI = \frac{XLN}{SLNG - SLNS} \quad (17)$$

where *DLAI* (m<sup>2</sup> m<sup>-2</sup>) is the reduction in leaf area index, *XLN* is nitrogen mobilised from leaves (g m<sup>-2</sup>), *SLNS* (g N m<sup>-2</sup> leaf area) is the specific leaf nitrogen of senesced leaves, and *SLNG* (g N m<sup>-2</sup> leaf area) is the specific leaf nitrogen of green leaves. The current leaf area index on day *i* is eventually calculated as

$$LAI_i = LAI_{i-1} + GLAI - DLAI \quad (18)$$

After beginning of seed growth, no more *LAI* increase is assumed in the model.

At very low and high temperatures, the model also calculates *LAI* senescence due to frost and heat, respectively (Soltani et al., 2013; Soltani and Sinclair, 2012).

#### 7.3.4. Dry Mass Production and Partitioning

iCrop simulates crop biomass (dry mass, *DM*, g m<sup>-2</sup>) production using the radiation use efficiency (*RUE*, g MJ<sup>-1</sup>) approach (Soltani et al., 2013; Soltani and Sinclair, 2012).

$$DM = PAR \cdot FINT \cdot RUE \quad (19)$$

$$FINT = 1 - e^{-KPAR \cdot LAI} \quad (20)$$

$$RUE = IRUE \cdot TCFRUE \cdot WSFG \cdot CO2RUE \quad (21)$$

where *PAR* (MJ m<sup>-2</sup>) is the photosynthetic active radiation which is calculated from the weather data input global radiation (*SRAD*, MJ m<sup>-2</sup>) and a constant conversion factor (0.48) and *FINT* (dimensionless) the fraction of *PAR* that is being intercepted by the crop canopy.

For the calculation of *FINT*, iCrop uses the crop canopy extinction coefficient (*KPAR*, dimensionless) which is a surrogate for crop canopy architectural traits such as the horizontal arrangement of leaves (i.e. their lack of random overlapping), sun angle, and leaf angles (Hay and Porter, 2006). Although *KPAR* changes throughout the day (sun angle) and growing season (*LAI*), the model assumes a constant value (wheat default: 0.65) (Soltani et al., 2013; Soltani and Sinclair, 2012).

The *RUE* calculation uses potential *RUE* (*IRUE*) and includes correction factors (each 0 to 1, dimensionless) for effects of temperature (*TCFRUE*), water stress (*WSFG*, see 7.3.7 Soil Water), and atmospheric CO<sub>2</sub> concentration (*CO2RUE*) (Soltani et al., 2013; Soltani and Sinclair, 2012).

Dry mass partitioning is simulated between leaves (i.e. leaf blades), grains, and everything else except roots (i.e. true stems, leaf sheaths, chaff dry matter). Since the stem dry mass pool comprises most of "everything else", it is referred to as "stem" in the context of the iCrop model (Soltani et al., 2013; Soltani and Sinclair, 2012). iCrop does not simulate root dry mass.

Before the beginning of seed growth (*BSG*) phenological phase, all dry mass is partitioned between leaves and stems. The relation of dry mass partitioning to leaves (leaf fraction) follows a multi-phase curve (Figure 6). In the first phase until a cultivar-specific total dry mass value (default: 160 g m<sup>-2</sup>), leaf fraction is 0.6 (*FLF1A*). Thereafter in phase two, more dry mass is allocated to stems (leaf fraction: 0.3; *FLF1B*). Eventually, in the period between termination of leaf growth on the main stem (*TLM*) and *BSG*, leaf fraction is still 0.1 (*FLF2*) due to delayed development of leaves on tillers. From *BSG* on, all dry mass is allocated to grains, unless grain demand is over-satisfied. In this case, excess dry mass is allocated to stems exclusively (Soltani et al., 2013; Soltani and Sinclair, 2012).

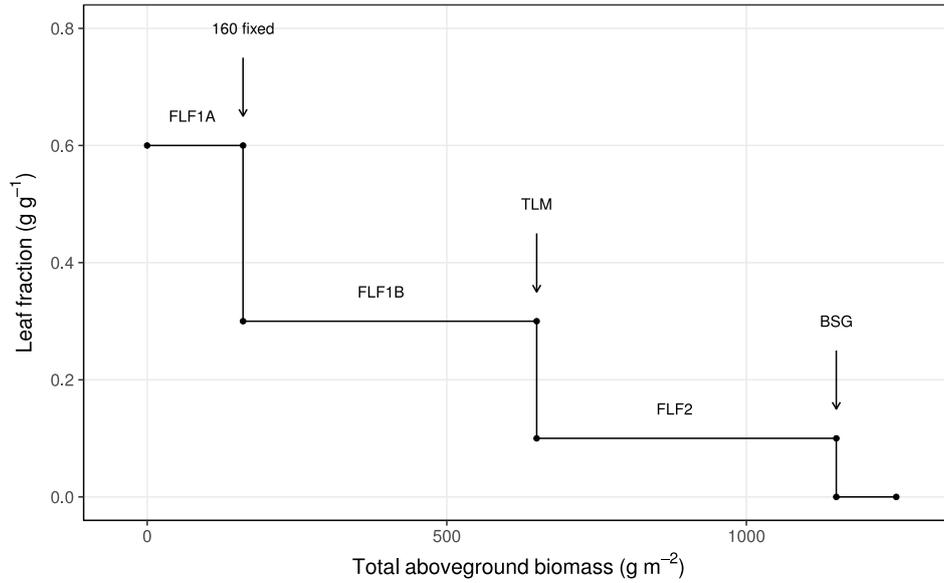


Figure 6 iCrop's default leaf fraction parameters for wheat (phase 1: FLF1A, phase 2: FLF1B, phase 3: FLF2) depending on total above-ground dry mass and key phenological phases (TLM: termination of leaf production on the main stem, BSG: beginning of seed growth). For the switch from FLF1A to FLF1B, a cultivar-specific dry mass threshold is defined, while for TLM and BSG the shown dry mass values are exemplarily only.

### 7.3.5. Nitrogen Uptake and Allocation

Plant N uptake is split in two phases. In the first phase up until BSG (vegetative growth), potential N demand ( $PNUP$ ,  $\text{g N m}^{-2}$ ) is driven by leaf area increase ( $GLAI$ ) and stem dry mass increase ( $GST$ ,  $\text{g m}^{-2}$ ):

$$PNUP = GLAI \cdot SLNG + GST \cdot SNCG \quad (22)$$

where  $SLNG$  is the specific leaf nitrogen of green leaves and  $SNCG$  (%;  $\text{g N g}^{-1}$  dry mass) is the nitrogen concentration of green (i.e. growing) stems. In its default version the iCrop model assumes  $SLNG$  independent from phenology with  $1.5 \text{ g m}^{-2}$  (nitrogen mass per leaf area) and also  $SNCG$  independent from phenology with 1.5% (target stem nitrogen concentration) (Soltani et al., 2013; Soltani and Sinclair, 2012). However,  $SLNG$  and  $SNCG$  can be defined for each cultivar.

Potential crop nitrogen uptake is then limited by several factors to simulate actual N-uptake ( $NUP$ ,  $\text{g N m}^{-2}$ ).  $MXNUP$  ( $\text{g N m}^{-2}$ ) defines a total maximum for daily  $NUP$ , reflecting the plants' energy costs for assimilating nitrate and ammonium into amino acids (Barker and Pilbeam, 2015). Also,  $NUP$  is limited when the soil is close to saturation (flooding) and when the availability of soil mineral nitrogen for crop uptake ( $SNAVL$ ,  $\text{g N m}^{-2}$ ) is below demand (Soltani et al., 2013; Soltani and Sinclair, 2012).

$$NUP = \min(PNUP, MXNUP, NUP_{flooding\_limited}, SNAVL) \quad (23)$$

If N is limited during vegetative growth, i.e. when crop N-uptake does not meet stem and leaf demand, the model first allows stem nitrogen concentration to drop below  $SNCG$  in order to satisfy leaf N demand. Stem N concentration may drop to a minimum defined by  $SNCS$  (senesced stem nitrogen concentration, %). If stem N concentration reaches this minimum and leaf demand still cannot be satisfied by  $NUP$ , then leaf area expansion is limited by N availability, up until the complete inhibition of leaf area increase. At this point of N stress, stem growth continues at minimum N concentration. When  $NUP$  drops even below this minimum, iCrop allows leaf senescence to remobilise N from leaves to stems to continue stem growth at  $SNCS$  (Soltani et al., 2013; Soltani and Sinclair, 2012).

From  $BSG$  on, i.e. in the generative growth phase, the above described N allocation and distribution rules still apply. Obviously, rules regarding leaf area expansion are then obsolete. The major change is that the new sink, grain, is considered primary as well as exclusive by the model. Potential  $NUP$  is, therefore, calculated from seed growth rate ( $SGR$ ,  $g\ m^{-2}$ ) and seed nitrogen concentration ( $\%GN$ , %; i.e.  $g\ N\ g^{-1}$  dry mass):

$$PNUP = SGR \cdot \%GN \quad (24)$$

$\%GN$  ranges within the user-defined boundaries of  $\%GN_{min}$  and  $\%GN_{max}$  (minimum and maximum grain nitrogen concentration).

However, N-uptake from soil requires dry mass production and happens only when there is more dry mass than the demand by  $SGR$ . Otherwise, N-uptake from soil is zero. In case of limited N conditions (higher grain demand for N than soil N-uptake), the model translocates nitrogen from stems and leaves to the grains. The amount of nitrogen translocated from leaves and stems is defined by the fraction of N translocated from leaves ( $FXLF$ , dimensionless) as follows:

$$TRLN = LAI \cdot (SLNG - SLNS) + (NST - WST \cdot SNCS) \quad (25)$$

$$FXLF = LAI \cdot \frac{(SLNG - SLNS)}{TRLN} \quad (26)$$

where  $TRLN$  ( $g\ N\ m^{-2}$ ) is the nitrogen pool translocatable from leaves and stems,  $NST$  ( $g\ N\ m^{-2}$ ) the total nitrogen content in stems, and  $WST$  ( $g\ m^{-2}$ ) the stem dry mass (see above for the description of the other variables). With  $FXLF$ , the daily removal of N from leaves and stems is calculated:

$$XLFN = (SGR \cdot \%GN - NUP) \cdot FXLF \quad (27)$$

$$XNST = (SGR \cdot \%GN - NUP) \cdot (1 - FXLF) \quad (28)$$

where  $XNLF$  and  $XNST$  (both  $\text{g N m}^{-2}$ ) are the nitrogen amounts removed from leaves and stems, respectively. Through the processes described for the vegetative growth (which also apply in the generative phase), reduction of leaf N content eventually leads to leaf area senescence and, hence, reduction of dry mass production (Soltani et al., 2013; Soltani and Sinclair, 2012).

#### 7.3.6. Grain Yield

iCrop uses the harvest index ( $HI$ ,  $\text{g g}^{-1}$ ; grain yield per total above-ground dry mass) approach to simulate wheat grain yield. Growth conditions during the vegetative phase are taken into account by the introduction of several dry mass thresholds at beginning of seed growth. Optimal vegetative growing conditions are assumed when dry mass accumulation reached 600 to 1200  $\text{g m}^{-2}$  at the beginning of seed growth (Soltani et al., 2013; Soltani and Sinclair, 2012).

Low dry mass production can limit seed growth. However, insufficient post-anthesis dry mass production is usually compensated for by mobilisation of dry matter from stems and leaves. Total mobilisable dry mass is by default 0.22 ( $FRTRL$ ,  $\text{g g}^{-1}$ ) of total dry mass at the beginning of seed growth (Soltani et al., 2013; Soltani and Sinclair, 2012).

#### 7.3.7. Soil Water

iCrop simulates soil water (and nitrogen) dynamics using up to 10 user-defined soil layers. Effects of soil water deficiencies and excesses on crop development and growth are accounted for. The version of the iCrop model used in this thesis included some updates (see sub-section 8.2 iCrop Source Code Updates) to the version outlined by Soltani et al. (2012). The following description includes these updates.

Water inputs to the soil are derived from daily weather data (precipitation, mm) and user-definitions (irrigation, mm). iCrop incorporates a simple snow model which depends on maximum temperature. Weather data does not need to define the type of precipitation. Irrigation may be manually defined as single events or automatic irrigation schemes depending on soil water status, crop phenology, or date (Soltani et al., 2013; Soltani and Sinclair, 2012).

Soil water content ( $SW$ ,  $\text{mm mm}^{-3}$ ) is simulated using the cascading layer approach (“tipping bucket approach”). For each layer, the water holding capacity is defined by air-dry limit ( $ADRY$ ), lower limit ( $LL$ ), drained upper limit ( $DUL$ ), and saturation ( $SAT$ ). Plant extractable soil water content ( $EXTR$ ) is calculated as the difference between  $DUL$  and  $LL$ . Further soil input parameters are a drainage factor ( $DRAINF$ , dimensionless; fraction of  $SW$  above  $DUL$  that drains each day) and initial water content ( $iniWL$ ) for each layer. Also, soil albedo (soil reflectance, dimensionless) needs to be defined (Soltani et al., 2013; Soltani and Sinclair, 2012).

The calculation of total soil water (mm) in each layer is as follows:

$$WL(L)_i = WL(L)_{i-1} + FLIN(L) - WU(L) - SE(L) - FLOUT(L) \quad (29)$$

where  $WL(L)_i$  is the total soil water of layer  $L$  on day  $i$  (today),  $WL(L)_{i-1}$  the  $WL$  on the previous day,  $FLIN$  and  $FLOUT$  the incoming and outgoing water fluxes, respectively,  $WU$  the amount of water taken up and transpired by the crop, and  $SE$  the soil evaporation. Incoming water flux in the topmost soil layer includes rainfall and irrigation, while the deeper layers receive incoming water only from the adjacent overlying layer (Soltani et al., 2013; Soltani and Sinclair, 2012).

iCrop calculates actual transpirable soil water ( $ATSW$ , mm) and total transpirable soil water ( $TTSW$ , mm) as follows:

$$ATSW(L) = WL(L) - WLLL(L) \quad (30)$$

$$TTSW(L) = WLUL(L) - WLLL(L) \quad (31)$$

where  $WLUL$  and  $WLLL$  are the total soil water (mm) amounts for a layer  $L$  at drained upper limit and lower limit, respectively. The fraction of actual transpirable soil water ( $FTSW$ ) is then calculated and used as drought stress indicator:

$$FTSW(L) = \frac{ATSW(L)}{TTSW(L)} \quad (32)$$

The simulation of soil evaporation is a two-stage process and uses calculations derived and modified from Priestly and Taylor (Priestley and Taylor, 1972; Ritchie, 1998). Actual evaporation is calculated from potential evaporation and accounts for crop cover ( $LAI$ ), crop and soil albedo, temperature, and global radiation. The soil is considered wet if the fraction of transpirable soil water in the root zone ( $FTSWRZ$ ) is at least 50%. If the soil layer is wet, stage I evaporation occurs. When  $FTSWRZ$  drops below 50%, stage II evaporation occurs. Evaporation is then an exponential function of time, thereby drastically decreasing the amount of evaporated soil water. Rainfall and irrigation events above a certain threshold ( $WETWAT$ , default 10 mm) reset evaporation to stage I. Evaporated water is primarily extracted from the top soil layer but also from deeper layers. Soil evaporation stops when a layer reaches the air-dry lower limit (Soltani et al., 2013; Soltani and Sinclair, 2012).

Transpiration is calculated from daily crop dry mass growth, a transpiration efficiency coefficient ( $TEC$ , Pa), and the vapour pressure deficit ( $VPD$ , kPa).  $VPD$  is simulated from saturated  $VPD$ , using daily minimum and maximum temperatures, assuming a fraction ( $VPDF$ ) of 65 to 75% depending on climate (Soltani et al., 2013; Soltani and Sinclair, 2012).

The effective root depth (hereafter: root depth) is simulated starting with an initial root depth of 200 mm. Growth in root depth is defined by the potential rate of root vertical penetration (*GRTDP*, default: 30 mm bd<sup>-1</sup>). In dry soil layers, i.e. no plant available soil water (wilting point), root growth is stopped. Also, dry mass production must be greater than zero to allow root growth. After the beginning of seed growth (*BSG*), no more root growth is assumed (Soltani et al., 2013; Soltani and Sinclair, 2012).

iCrop calculates water stress factors (each from 0 to 1, dimensionless) for leaf area expansion (*WSFL*), dry mass production (*WSFG*), and phenological development (*WSFD*). *WSFL* and *WSFG* are calculated in the same way, depending on thresholds:

$$WSFL = 1 \text{ if } FT\text{SWRZ} \geq WSSL \quad (33)$$

$$WSFL = \frac{FT\text{SWRZ}}{WSSL} \text{ if } FT\text{SWRZ} < WSSL \quad (34)$$

and

$$WSFG = 1 \text{ if } FT\text{SWRZ} \geq WSSG \quad (35)$$

$$WSFG = \frac{FT\text{SWRZ}}{WSSG} \text{ if } FT\text{SWRZ} < WSSG \quad (36)$$

where *WSSL* and *WSSG* are the threshold *FTSWRZ* values for drought stress effects on leaf area expansion and dry mass production, respectively. The water stress factor affecting crop phenology, *WSFD*, may hasten development. The same threshold as for dry mass production, *WSSG*, is used to calculate this stress factor. However, maximum hastening is 40% (*WSFD* = 1.4), which directly affects the calculation of biological days (Soltani et al., 2013; Soltani and Sinclair, 2012) (see 7.3.2).

Excessive soil water content over 95% of saturation (flooding) sets all water stress factors (*WSFL*, *WSFG*, and *WSFD*) to zero (i.e. maximum stress), thereby restricting any crop growth or development (Soltani et al., 2013; Soltani and Sinclair, 2012).

#### 7.3.8. Soil Nitrogen

iCrop considers ammonium (NH<sub>4</sub><sup>+</sup>) and nitrate (NO<sub>3</sub><sup>-</sup>) as plant available soil mineral nitrogen pool, *NSOL* (soluble nitrogen in the soil solution, g N m<sup>-2</sup>). However, the model does not distinguish between the two forms of nitrogen but only accounts for the pure mineral N amount.

$$NSOL(L)_i = NSOL(L)_{i-1} + NMIN(L) + NFERT(L) - NVOL(L) - NOUT(L) - NDNIT(L) - NUP(L) \quad (37)$$

where *i* (today) and *i-1* (yesterday) represent simulation days, *NMIN* (not to be confused with *Nmin*) is the net mineralisation (including immobilisation) of nitrogen from organic soil matter, *NFERT* the nitrogen applied through fertilisation, *NVOL* the amount of nitrogen volatilisation,

*NOUT* the nitrogen drained to the next lower layer through water mass flow, *NDNIT* the nitrogen transformed through denitrification, and *NUP* the nitrogen taken up by the crop. The plant available nitrogen (*SNAVL*) is calculated considering a minimum required N concentration in the soil solution ( $1 \text{ mg N l}^{-1}$ ) and the rooted fraction of a layer (Soltani et al., 2013; Soltani and Sinclair, 2012).

N mineralisation from soil organic N via ammonium to nitrate is simplified to a single function. N mineralisation is calculated based on the fraction of potentially mineralisable N from organic soil N (*FMIN*, dimensionless), *NSOL*, soil water, and soil temperature (assumed equal to air temperature). N-fertilisation is defined in terms of pure N input ( $\text{g m}^{-2}$ ), regardless of which form (e.g. nitrate, ammonium, urea) of N is applied. Also, the fraction of volatilised N must be specified. iCrop considers fertilised and volatilised N only for the topmost soil layer. N denitrification is simulated when SW in a layer exceeds the drained upper limit (Soltani et al., 2013; Soltani and Sinclair, 2012).

#### 7.4. Parameter Calculation

The iCrop parameterisation for the four winter wheat cultivars in this study was performed in three main steps:

1. Crop phenology parameterisation: cultivar-specific parameters for crop sensitivity to vernalisation and photoperiod were estimated (calibration against observed phenology).
2. Parameterisation of processes which affect dry mass production: iCrop includes several interacting parameters that influence crop dry mass accumulation and partitioning. Among these are crop and soil parameters that control canopy development (i.e. leaf area index (*LAI*), leaf number on the main stem), soil water availability, and soil and plant N dynamics. The parameters were estimated from field data.
3. Final yield: Due to its dependence on dry mass production during the growing season, parameters which directly influence yield formation were parameterised last. These include only *PDHI* (potential daily harvest index) and %GN limits (minimum and maximum).

After parameterisation, the iCrop model was evaluated using an independent long-term data set (four locations across Lower Austria during three seasons, i.e. 12 environments). The crop/cultivar parameters were not changed, only season-specifics (weather data), location-specifics (latitude, soil parameters), and management-specifics (sowing date, fertilisation) were adapted to the evaluation environment. The simulation results were then compared to the observed data sets and the model performance analysed.

The following sections describe how parameters were calculated during the model parameterisation process.

#### 7.4.1. Weather Data Input

Daily weather data (precipitation [mm], daily maximum and minimum temperature [°C], global radiation [ $\text{MJ m}^{-2}$ ]) for parameterisation and evaluation was both acquired from weather stations run by ZAMG (Central Institution for Meteorology and Geodynamics, a subordinate agency of the Austrian Ministry of Education, Science and Research). By default, daily temperature values were referenced to the period from 7 p.m. of the previous day until 7 p.m. of the current day. Also, precipitation data was cumulated from 7 a.m. of the current day until 7 a.m. of the following day. Global radiation was delivered in ( $\text{J cm}^{-2}$ ) and had to be converted to ( $\text{MJ m}^{-2}$ ) by simple multiplication. Although ZAMG continuously verifies all data, I also briefly checked the integrity of the parameterisation data set by simple comparison of yearly cumulative, minimum, average, and maximum global radiation and cumulative rainfall for the period from 1991 to 2019. This comparison did not show any apparent outliers.

The vapour pressure deficit factor (*VPDF*) was calibrated via comparison of simulated and observed plant growth and soil water content. The atmospheric carbon-dioxide ( $\text{CO}_2$ ) concentration (parameter: *CO2*) was taken from literature.

#### 7.4.2. Soil Parameters

Previous experiments (e.g. Fuchs, 2016) have shown that gravimetric soil water measurements in the field are prone to a high variability in the top 0-10 cm layer, while data from deeper layers appears more stable. Therefore, the chosen soil layering (for sampling and model parameterisation) included a thinner top-layer: 0-10, 10-30, 30-60, 60-90, and 90-120 cm. For soil nitrogen, the more common 30 cm intervals were used: 0-30, 30-60, 60-90, and 90-120 cm. For the parameters *NORG* (organic N content), *FMIN* (fraction mineralisable N), and initial *Nmin I* converted the top-soil layering (0-30 cm) to the parameterised soil. For *NORG* and *FMIN*, the 0-30 cm value was directly used for 0-10 and 10-30, while for initial *Nmin* the factors 1/3 (0-10 cm) and 2/3 (10-30 cm) were applied for the conversion.

The model's soil water and *Nmin* parameters were chosen for each experiment separately (EXPI, EXP2) as results showed obvious differences in soil texture, particularly in deeper soil layers (see 8.3.16 and 8.3.17). Initial values for *Nmin* and soil bulk density (*BDL*) were set to the measured values. For initial soil water, the in-season comparison of simulated vs. observed was included in the calibration. The soil water lower limit (*LL*) and drained upper limit (*DUL*) were selected for each layer based on observed data points during the growing season. As a guideline, the 5% quantile (for air-dry *LL* [*ADRY*] for the 0-10 cm layer, and for *LL* for the remaining layers) and

the 95% quantile (for *DUL*) served. For the topmost layer (0-10 cm), *LL* was assumed to be  $0.03 \text{ mm}^3 \text{ mm}^{-3}$  higher than *ADRY*. For the other layers (from 10 cm downwards) I assumed *ADRY* and *LL* equal. Saturated soil water content (*SAT*) was assumed  $0.03 \text{ mm}^3 \text{ mm}^{-3}$  higher than *DUL* across all layers. The drainage factor (*DRAINF*) was set identical for both experiments based on previous experiments and calibration against observed in-season soil water data.

For the calculation of layer-wise soil *NORG* (organic nitrogen content, %) measured total soil organic carbon content (*TOC*, %) was used. *TOC* can roughly be converted to *NORG* by converting it with an average C/N ratio:

$$C/N = \frac{TOC}{NTOT} \quad (38)$$

where *NTOT* is the soil's total nitrogen content (%).

Since *NORG* usually represents at least 95% of *NTOT* in common agricultural soils (Bayerisches Staatsministerium für Ernährung, Landwirtschaft und Forsten (StMELF), 2018) and an estimate was sufficient for estimating the model's parameter, I assumed:

$$NORG = NTOT \quad (39)$$

The C/N ratio of many common agricultural soils usually ranges between 9 and 11 (Baumgarten et al., 2017; Diez and Weigelt, 1991). Assuming an average C/N ratio of 10, which is the optimum according to Baumgarten et al. (2017), together with equation (39), *NORG* was calculated with a simple division:

$$NORG = \frac{TOC}{10} \quad (40)$$

For soil albedo and the runoff curve number, the model's default values were used.

#### 7.4.3. Crop/Cultivar Parameters

The parameters *vsen* (vernalisation sensitivity), *cpp* (critical photoperiod, h), and *ppsen* (photo-period sensitivity) were first calibrated using phenological data from a previous experiment conducted in Eastern Austria with the winter wheat cultivar Capo sown in September, October, November, March, and April in the season 2013/14 (Fuchs, 2016). Based on the pre-calibrated values, the parameters were then estimated to fit simulated phenological stages using phenological data from EXP1 and EXP2.

The phenological parameters which define the duration of phenology phases (bd-phases; i.e. thermal time targets of the phenological phases) were calibrated to match the observed stage occurrence (e.g. beginning of stem elongation, BBCH 31). Observed relative differences between the cultivars received priority during this process. Observed phenological data was interpolated based on thermal time to calculate the occurrence of the exact same main stages as simulated in iCrop.

Calculation of other model parameters based on field observations was always carried out with the highest possible accuracy. Usually, this meant that single measurements were used, e.g. LAI on a specific day of one experiment in a specific plot (i.e. a specific replication of a specific cultivar with a specific N-treatment). Afterwards, this “population of parameters” was mathematically converted into a single parameter value using summary calculations. A suitable summary calculation (e.g. mean) was chosen depending on the parameter’s nature. The summary calculations were either applied to the whole data set or on sub-sets (such as a specific cultivar or N-treatment). Summary-calculations included linear regressions (including segmented linear regressions), power functions, mean, and quantiles (percentiles). Details are described below.

The iCrop parameters affecting leaf canopy development include phyllochron ( $phyl$ , °C leaf<sup>-1</sup>), leaf area expansion parameters ( $PLACON$  and  $PLAPOW$ , both dimensionless), and leaf:stem partitioning parameters ( $WTOPL$ , g m<sup>-2</sup>;  $FLFIA$ , g g<sup>-1</sup>;  $FLFIB$ , g g<sup>-1</sup>;  $FLF2$ , g g<sup>-1</sup>).

Phyllochron was calculated using the linear regression of observed main stem leaf number ( $MSNN$ , dimensionless or “leaf number”) against cumulative thermal time ( $CDTU$ , °Cd) during the leaf appearance phase. While the default version of iCrop defines one phyl parameter, I implemented two phases of leaf number development ( $phyl1$  and  $phyl2$ ) to better reflect observations which showed two linear phases. These were divided by a breakpoint ( $phylBP$ , dimensionless or “leaf number”) based on  $MSNN$ , similar to the approach in APSIM (Holzworth et al., 2014).  $phyl1$  and  $phyl2$  were calculated from the left and right slope of the segmented linear regression of  $MSNN$  versus  $CDTU$  (Figure 7).

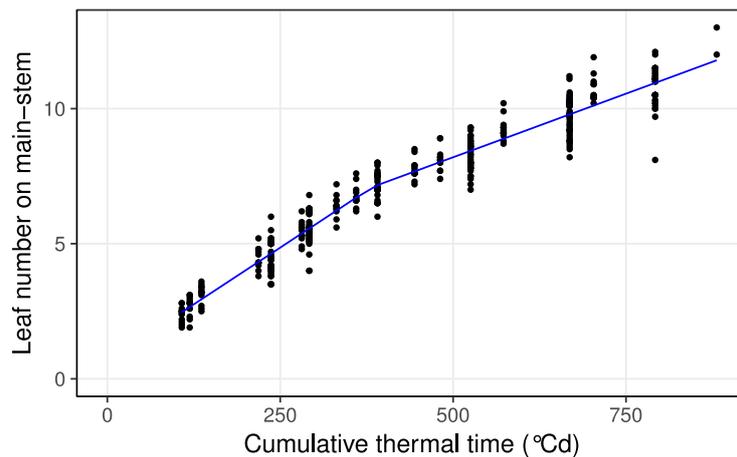


Figure 7 Example for the calculation of the phyllochron parameters ( $phyl1$ : slope of left segment,  $phyl2$ : slope of right segment). Data: Observed main stem leaf number vs. cumulative thermal time from BBCH II (first leaf) to 39 (end of stem elongation), exemplarily shown for the winter wheat cultivar Aurelius in two seasons (2017/18 and 2018/19, Tulln). Lines: two-segment linear regression.

$$MSNN = k \cdot CDTU + d \quad (41)$$

$$k = \frac{1}{phyl} \quad (42)$$

where d is the intercept.

The breakpoint phylBP was converted to the unit main stem leaf number as follows:

$$phylBP [leaf] = \frac{phylBP [^{\circ}C]}{phyl1 [^{\circ}C leaf^{-1}]} \quad (43)$$

Leaf area expansion in iCrop is calculated following a power function (Figure 8); see equation (13). PLACON is assumed 1 (default), and PLAPOW was calculated from field measurements of LA and MSNN by fitting a power function using a non-linear least squares method. The effect of increasing plant density (PD) is also captured. iCrop requires PLAPOW to be normalised to 300 plants per m<sup>2</sup> (PLAPOW<sub>300</sub>):

$$PLAPOW_{300} = \frac{PLAPOW_{PD}}{1.1718 - 0.0006 \cdot PD} \quad (44)$$

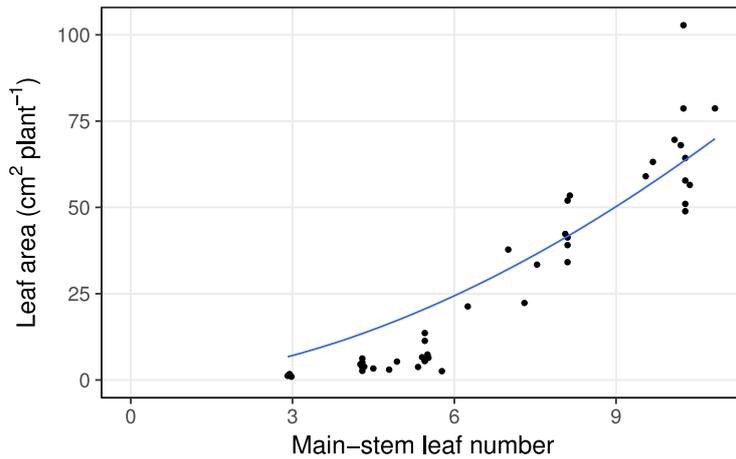


Figure 8 Example for the calculation of the parameter PLAPOW. Data: Leaf area per plant versus main stem leaf number, exemplarily shown for the winter wheat cultivar Aurelius with high N-fertilisation (N2, N3) in two seasons (2017/18 and 2018/19, Tulln). Line: Regression analysis (power).

SLA affects LAI in the model calculations only after termination of the MSNN growth (BBCH 41) up until BSG (BBCH 71) and has a low impact on overall LAI. Therefore, the SLA parameter was calculated as the slope of the linear regression of measured LAI and green leaf dry matter (GLDM, g m<sup>-2</sup>) between BBCH 40 and 71 (intercept assumed zero):

$$SLA = \frac{LAI}{GLDM} \quad (45)$$

The leaf:stem partitioning parameters were calculated from observed dry matter (DM) of leaves and total DM (i.e. leaves plus stems) up until booting, using a two-segment linear model (Figure 9). The model parameters FLFIA and FLFIB represented the slopes before and after the breakpoint (WTOPL, g m<sup>-2</sup>).

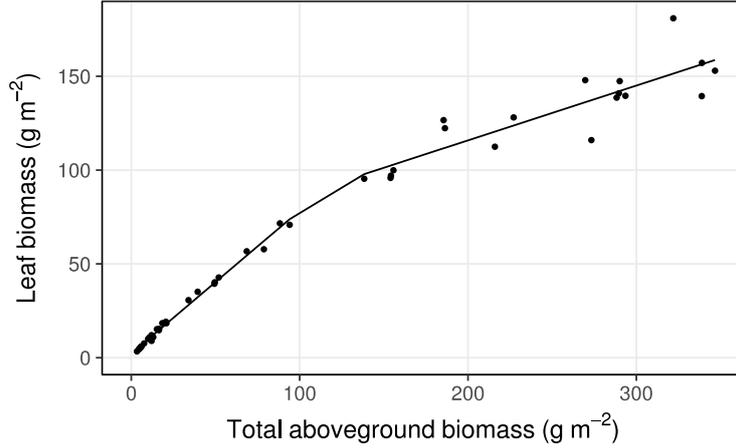


Figure 9 Example for the calculation of the parameters WTOPL (breakpoint at ca. 110 g m<sup>-2</sup>), FLFIA (slope of the left segment), and FLFIB (slope of the right segment). Data: Leaf biomass (dry mass) versus total above-ground biomass (dry mass) from BBCH 9 (emergence) until BBCH 37 (flag leaf emergence), exemplarily shown for winter wheat cultivar Aurelius in well fertilised treatments (N2, N3) in two seasons (2017/18 and 2018/19). Line: Regression analysis (segmented linear regression).

Plant nitrogen parameters included concentrations for stem (senesced: *SNCS*, green: *SNCG*; both g g<sup>-1</sup>), grain (minimum: %GNmin, maximum: %GNmax; both g g<sup>-1</sup>), and specific leaf nitrogen contents (senesced: *SLNS*, green: *SLNG*; both g m<sup>-2</sup> leaf area). The iCrop model calculates leaf N content from *SLNG* and daily leaf area growth. Leaf area growth terminates at *BSG*. Therefore, for the calculation of *SLN* at *BSG* (*SLN*<sub>BSG</sub>) the measurements of green leaf N uptake (*NUPLG*, g m<sup>-2</sup>) at *BSG* and *LAI* (m<sup>2</sup> m<sup>-2</sup>) at *BSG* were used:

$$SLN_{BSG} = \frac{NUPLG_{BSG}}{LAI_{BSG}} \quad (46)$$

$$SLNG = Q_{97.5\%}(SLN_{BSG,i}, \dots, SLN_{BSG,k}) \quad (47)$$

where  $Q_{97.5\%}$  is the 97.5% quantile (percentile), *i* and *k* are the first and last considered single measurements (i.e. factor combination of season, replication, cultivar, and N-treatment).

In the iCrop model, N is mobilised from leaves to grains in relation to *LAI*-dynamics (i.e. *LAI* decrease) and mobilisable *SLN* (i.e. *SLNG* minus *SLNS*) from *BSG* until *TSG*. Simulated leaf N content reaches its maximum at *BSG* and *LAI* may reach zero at *TSG* (for details, see Solatani and Sinclair (2012) and the crop model source code). In other words, iCrop will reduce *LAI* to zero if all mobilisable leaf N is translocated to the grains. Also, as soon as *LAI* is zero, no more N mobilisation from leaves to grains is possible. Therefore, the calculation of the *SLNS* parameter must use measurements of total leaf N uptake (*NUPLTOT*) at *TSG* (which equals harvest) and maximum *LAI*:

$$SLN_{harvest} = \frac{NUPLTOT_{harvest}}{LAI_{max}} \quad (48)$$

$$SLNS = Q_{2.5\%}(SLN_{harvest,i}, \dots, SLN_{harvest,k}) \quad (49)$$

where  $Q_{2.5\%}$  is the 2.5% quantile, i and k: see above.

Stem N content simulation by iCrop is similar to leaf except that it is based on dry mass (g N g<sup>-1</sup> dry matter). Measured stem N concentrations were directly used for parameter calculation:

$$SNCG = Q_{97.5\%}(SNC_{BSG,i}, \dots, SNC_{BSG,k}) \quad (50)$$

$$SNCS = Q_{2.5\%}(SNC_{harvest,i}, \dots, SNC_{harvest,k}) \quad (51)$$

where SNC is the measured stem nitrogen concentration at BSG and at harvest, and Q, i, k: see above.

The 2012 version of the iCrop model defined only one *SNCS* parameter for the whole plant life cycle. I updated the source code and implemented two phases of *SNCS*: *SNCSI* from emergence to *BSG*, and *SNCS2* after *BSG* until maturity (see 8.2 iCrop Source Code Updates). The reasons for this are given in the Discussion chapter (sub-section 9.3.5). The calculation of *SNCS2* is identical to the calculation of the original parameter *SNCS*.

$$SNCS1 = Q_{2.5\%}(SNC_{BSG,i}, \dots, SNC_{BSG,k}) \quad (52)$$

$$SNCS2 = Q_{2.5\%}(SNC_{harvest,i}, \dots, SNC_{harvest,k}) \quad (53)$$

Obviously, both *SNCG* and *SNCSI* were calculated from observed SNC at BSG.

The fraction of above-ground dry mass at *BSG* which is available for translocation to seeds (parameter *FRTRL*, g g<sup>-1</sup>) was estimated from observed data of total above-ground dry mass at *BSG* ( $TDM_{BSG}$ ) and harvest ( $TDM_{harvest}$ ) and grain yield at harvest ( $YLD_{harvest}$ ):

$$FRTRL_n = \frac{(TDM_{BSG,n} - (TDM_{harvest,n} - YLD_{harvest,n}))}{TDM_{BSG,n}} \quad (54)$$

$$FRTRL = mean(FRTRL_i, \dots, FRTRL_k) \quad (55)$$

where n represents a single measurement (i.e. factor combination of season, replication, cultivar, and N-treatment) and i and k are the first and last considered single measurement.

An estimate for the parameter *PDHI* (potential increase in daily harvest index, g g<sup>-1</sup> d<sup>-1</sup>) was calculated as follows:

$$DHI = \frac{HI}{SGD} \quad (56)$$

with

$$HI = \frac{yield_{harvest}}{WTOP_{harvest}} \quad (57)$$

$$SGD = TSG - BSG \quad (58)$$

where  $DHI$  is the actual increase in daily harvest index ( $g\ g^{-1}\ d^{-1}$ ),  $HI$  the measured harvest index ( $g\ g^{-1}$ ),  $SDG$  the measured duration of the seed growth period (d),  $TSG$  the termination of seed growth (days after sowing), and  $BSG$  the beginning of seed growth (days after sowing). For  $BSG$  and  $TSG$ , the corresponding observed BBCH stages were defined as 71 ( $BSG$ ) and 87 ( $TSG$ ). Observed occurrence of these stages (days after sowing) was linearly interpolated between the closest previous and following BBCH observation.

Apparently, to calculate the true PDHI from field data, an experiment with optimal conditions would be required. However, the maximum of the observed  $DHI$ 's was used as an estimate:

$$PDHI = maximum(DHI_i, \dots, DHI_k) \quad (59)$$

where  $i$  and  $k$  are the first and last considered single measurement (see above).

The radiation use efficiency parameter ( $IRUE$ ,  $g\ MJ^{-1}$ ) was changed to the value used in the crop growth model APSIM (Holzworth et al., 2014) using the following conversion: In APSIM, the  $RUE$ -parameter is directly applied to photosynthetic active radiation ( $PAR$ ) for calculating dry matter production. In iCrop, dry matter is calculated using the  $RUE$ -parameter ( $IRUE_{ICROP}$ ) with global radiation ( $SRAD$ ) and a conversion factor to  $PAR$  (0.48). Therefore, the model parameter conversion from  $RUE_{APSIM}$  to  $IRUE_{ICROP}$  was:

$$IRUE_{ICROP} = \frac{RUE_{APSIM}}{0.48} \quad (60)$$

Reasons for changing the  $IRUE$  parameter are given in the Discussion section (9.2.8 Dry Mass).

## 7.5. Statistics

All statistical tests were performed with R (R Core Team, 2015). For data handling and manipulation I used the packages “tidyr” (Wickham and Henry, 2018) and “dplyr” (Wickham et al., 2019). Various statistical tests were performed using the packages “agricolae” (Mendiburu, 2017) (e.g. for ANOVA), “devtools” (Wickham and Chang, 2017), “broom” (e.g. for easier display of grouped ANOVA), and “car” (Fox and Weisberg, 2011). For approximation of missing values the package “zoo” (Zeileis and Grothendieck, 2005) was used. Also graphs were produced using the

R package “ggplot2” (Wickham, 2009) in combination with “segmented” (Muggeo, 2003) for estimation of segmented linear regressions and their breakpoints, and “lubridate” (Grolemund and Wickham, 2011) for easy conversion of date formats. Non-linear least square estimates were calculated with the function `nls()` from the R-base package “stats”.

The root mean square error (RMSE, unit depending on the examined measurement) and mean bias error (MBE) were calculated as described by Salo et al. (2015).

The relative RMSE (RRMSE, %; also called normalised RMSE [NRMSE]) was calculated as:

$$RRMSE = \frac{RMSE}{\bar{o}} \cdot 100 \quad (61)$$

where  $\bar{o}$  is the average of the observed variable.

The percentage error (PE, %) was calculated as follows:

$$PE_n = \frac{s_n - o_n}{o_n} \cdot 100 \quad (62)$$

$$PE = \text{mean}(PE_i, \dots, PE_k) \quad (63)$$

where *s* and *o* are corresponding simulated and observed values, *n* represents a single measurement (i.e. factor combination of season, replication, cultivar, and N-treatment), and *i* and *k* are the first and last considered single measurement.

Taylor et al. (1999) reported an average coefficient of variation (CV) of 13.5% for over 300 wheat field experiments. Also, Asseng et al. (2013) compared model variations using the same threshold (13.5%). However, He et al. (2017) argued that model variation (RRMSE) cannot be expected to be below the variation in the data which was used for calibration. Therefore, if model variation gets close to 13.5% it can be regarded robust. In the following, an RRMSE below 20% is referred to as “robust”.

## 8. Results

This chapter outlines the weather during the field experiments as well as their outcomes (soil observations, in-season measurements, harvest) (8.1), results of model parameter calculations and calibrations from measured data (8.3), simulation results of the field experiments using the derived parameters (8.4), and model evaluation results using long-term data (8.5). Section 8.2 describes changes to algorithms in the iCrop model source code.

### 8.1. Field Experiments

#### 8.1.1. Weather

Total precipitation in the growing season (October to June) at Tulln (Eastern Austria) was 374 mm (EXPI, 2017/18) and 481 mm (EXP2, 2018/19) (Figure 10). Long term average (1991-2015) was 408 mm. Average temperatures in the growing seasons were 9.3 °C in EXPI and 9.4 °C in EXP2 (long term average: 7.9 °C).

During the first three months of the season (Oct – Nov) monthly precipitation decreased in EXPI from ca. 60 mm to 20 mm while it increased in EXP2 from ca. 15 mm to 95 mm. In the same period, average monthly temperatures dropped similarly in both seasons from ca. 13 °C to 3 °C with only slightly lower values in EXPI. Between January and March precipitation in EXPI was low and rather constant at ca. 25 mm while in EXP2 variation was greater (10 to 50 mm). Temperature was rather high in January in EXPI (ca. 4 °C) and cold afterwards (below 0 °C) while in EXP2 it was cooler in January (close to 0 °C) and much warmer February and March (ca 4 and 9 °C, resp.). The last three months of the growing season (Apr – Jun) were warmer and dryer in EXPI than in EXP2. In EXPI, April showed almost no precipitation while May was wet (ca. 120 mm) and June rather dry (ca. 50 mm). In comparison, EXP2 showed much more precipitation in April (ca. 50 mm), a very wet May (ca. 170 mm), followed by a very dry June (ca. 15 mm). Mean temperatures from April to June increased gradually in EXPI from ca. 15 °C to 21 °C and less gradual in EXP2 from ca. 12 °C (April) to 13 °C (May) then jumping to 24 °C (June).

Cumulative global radiation during the growing season was 2644 MJ m<sup>-2</sup> in EXPI and 2729 MJ m<sup>-2</sup> in EXP2 (long term average: 2591 MJ m<sup>-2</sup>).

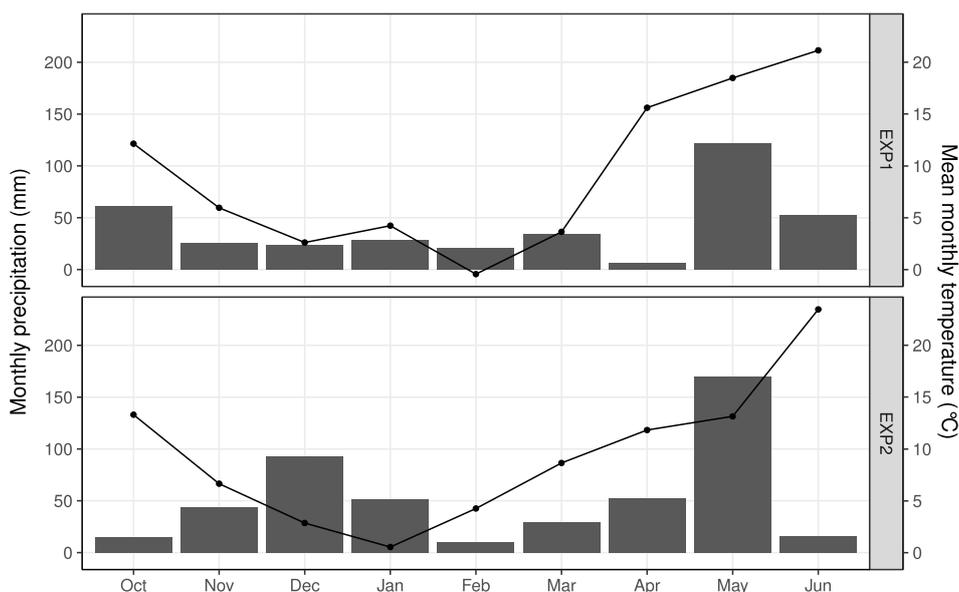


Figure 10 Monthly precipitation (bars) and mean temperatures (dots) during the winter wheat growing season in EXP1 (2017/18) and EXP2 (2018/19) at the meteorological station Langenlebar near Tulln (Lower Austria).

### 8.1.2. Sampling Dates

Phenological stages of the winter wheat experiments were scored 40 times during the two seasons. Destructive plant and soil samples were taken on 29 dates (Table 1). Soil water (SW) samples at sowing were taken 2 days (EXP1) and 3 days (EXP2) after actual sowing. SW and biomass (total above-ground dry mass, TDM) samples closest to observed anthesis were taken during BBCH 71 (EXP1) and BBCH 62-69 (EXP2). At harvest, phenological stages of all treatments were beyond BBCH 89, the exact stage was not scored.

The aim was to collect soil and plant samples at the important stages sowing (BBCH 0), anthesis (BBCH 65), and physiological maturity/harvest (BBCH 89+). Although the actually measured phenological stages deviate from these (especially for anthesis in EXP1), I use the mentioned main stage names in the text when referring to a specific sampling date, as shown in Table 1. The implications of these deviations are discussed in section 9.2 “Calculation of Model Parameters”.

### 8.1.3. Observed Soil Processes

#### *Soil Water*

Soil bulk density was 1.37, 1.50, 1.49, 1.42, and 1.48 g m<sup>-2</sup> for the soil layers 0-10, 10-30, 30-60, 60-90, and 90-120 cm, respectively.

Average total soil water content (full profile: 0-120 cm soil depth) at sowing was 574 mm in EXP1 and 364 mm in EXP2. Soil water contents at harvest were 462 mm in EXP1 and 276 mm in EXP2.

The time course of soil water content in the different soil layers (exemplarily shown for specific treatments in Figure II) showed similar ranges for the top three layers (0-10, 10-30, and 30-60

cm soil depth) in both experiments. The deepest layers (60-90 and 90-120 cm) showed much lower water contents in EXP2 than in EXP1. In EXP1 there was a steep decline of soil water from ca. day 150 (mid-April) onwards, especially in the upper soil layers. In the upper layers, EXP2 showed a decline from ca. day 150 until day 200 (end of April), then an incline for ca. 30 days (until end of May), followed by another decline until the second last measurement (end of June). Across both seasons, variations in soil water content were largest in the topmost soil layer which transferred delayed and with reduced amplitudes to the lower layers.

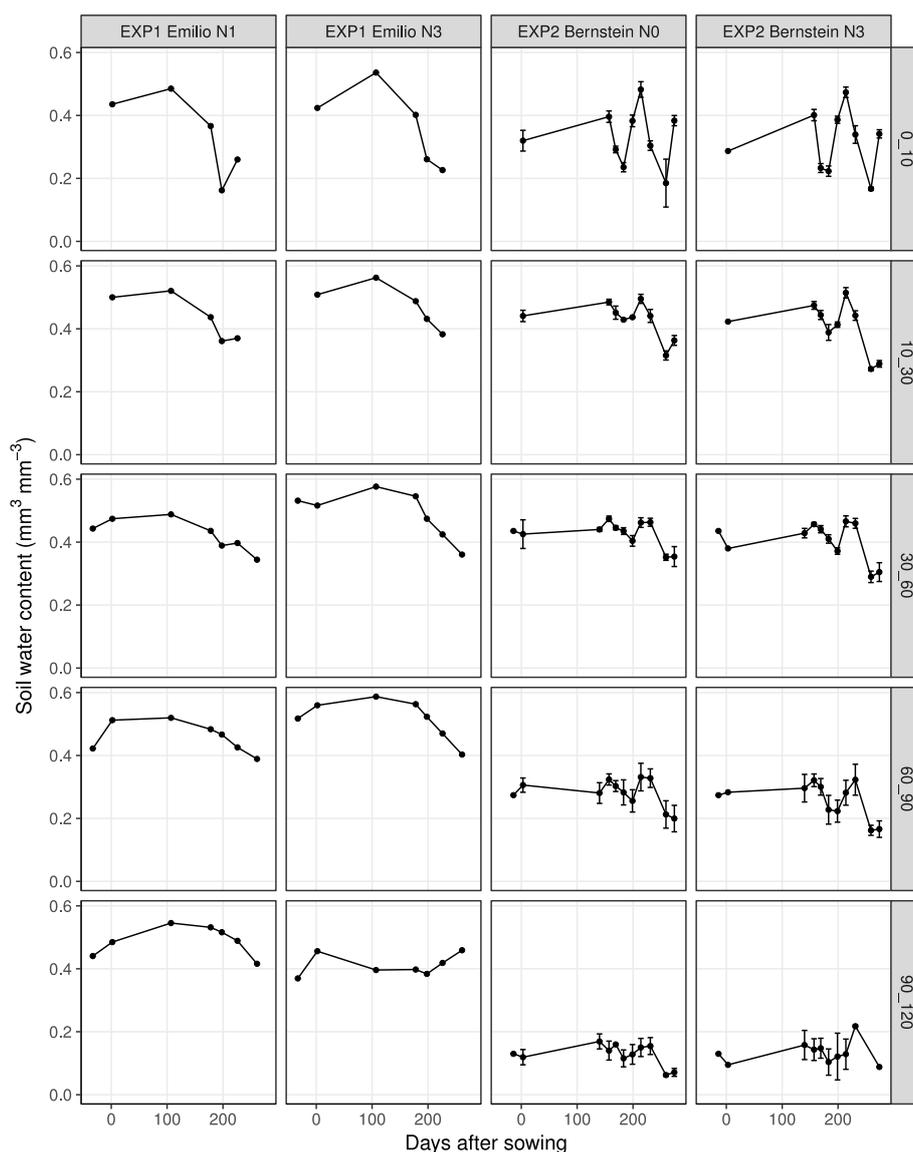


Figure 11 Soil water content (volumetric) for each soil layer of the full soil profile over time, shown exemplarily for winter wheat cultivar Emilio with nitrogen fertilisation treatment N1 and N3 in the season of EXP1 and Bernstein with N0 and N3 in EXP2. Soil layers (cm): 0-10, 10-30, 30-60, 60-90, 90-120. Seasons: EXP1: 2017/18, EXP2: 2018/19. N-treatments: no N (N0), 70 (N1), 140 (N2), and 210 (N3) kg N ha<sup>-1</sup>. Error bars indicate standard error. Points without error bars are based on single measurements.

#### Soil Mineral Nitrogen

Average initial soil mineral nitrogen content (N<sub>min</sub>) of the whole soil profile (0-120 cm soil depth) was 63 kg N ha<sup>-1</sup> one month before sowing in EXP1 and 30 kg N ha<sup>-1</sup> two weeks before sowing in EXP2 (Table 2). Final average N<sub>min</sub> contents at harvest ranged from 23 (N0) to 114 (N3) and from 11 (N0) to 37 (N3) kg N ha<sup>-1</sup> in EXP1 and EXP2, respectively.

The N balance was negative for all treatments, ranging from -10 kg N ha<sup>-1</sup> in Bernstein (N3, EXP2) to -127 in Bernstein (N3, EXP1). There was no consistent trend of the effect of N-fertilisation on the N balance. Absolute values were always higher in EXP1 compared to EXP2.

Table 2 Soil mineral nitrogen (Nmin) balance (Bal.) for both winter wheat growing seasons (EXPI: 2017/18, EXP2: 2018/19). Cultivars (Cv.): Arnold (Ar.), Aurelius (Au.), Bernstein (Be.), and Emilio (Em.). Ini. and Fin. Nmin: Initial (at sowing) and final (at harvest) soil mineral N, FertN: cumulative fertilised N, NUP: crop N-uptake at harvest.

Cv.	N-trt.	EXPI					EXP2				
		Ini. Nmin	FertN	NUP	Fin. Nmin	Bal.	Ini. Nmin	FertN	NUP	Fin. Nmin	Bal.
(kg N ha <sup>-1</sup> )											
Ar.	N0	63.3	0.0	107.5	27.0	-71.3	29.9	0.0	69.0	15.0	-54.1
	N1	63.3	70.0	170.2	57.0	-93.9	29.9	70.0	120.9	16.0	-37.0
	N2	63.3	140.0	194.4	104.0	-95.2	29.9	140.0	181.7	34.0	-45.8
	N3	63.3	210.0	201.5	129.0	-57.2	29.9	210.0	230.8	36.5	-27.4
Au.	N0	63.3	0.0	120.0	25.0	-81.8	29.9	0.0	70.0	9.0	-49.1
	N1	63.3	70.0	197.8	25.0	-89.5	29.9	70.0	123.1	22.0	-45.2
	N2	63.3	140.0	221.0	77.0	-94.7	29.9	140.0	193.3	27.5	-50.9
	N3	63.3	210.0	237.0	78.0	-41.8	29.9	210.0	237.9	36.5	-34.5
Be.	N0	63.3	0.0	112.0	20.0	-68.7	29.9	0.0	65.5	10.5	-46.1
	N1	63.3	70.0	190.7	28.0	-85.5	29.9	70.0	127.6	17.0	-44.7
	N2	63.3	140.0	234.2	52.0	-83.0	29.9	140.0	190.9	20.5	-41.5
	N3	63.3	210.0	231.8	168.0	-126.5	29.9	210.0	215.4	34.5	-10.0
Em.	N0	63.3	0.0	109.9	21.0	-67.6	29.9	0.0	59.2	10.5	-39.8
	N1	63.3	70.0	177.0	43.0	-86.8	29.9	70.0	110.3	19.0	-29.4
	N2	63.3	140.0	207.5	85.0	-89.2	29.9	140.0	184.6	22.0	-36.7
	N3	63.3	210.0	206.7	147.0	-80.5	29.9	210.0	223.6	41.5	-25.2

#### 8.1.4. Phenology

##### *Effect of Season*

For phenology, the seasonal differences (Figure 12) showed emergence 15 days after sowing (DAS) in EXPI and 24 DAS in EXP2. In temperature sums, emergence occurred 123 °Cd (EXPI) and 250 °Cd (EXP2) after sowing. Tillering occurred earlier in EXPI (87 DAS) than in EXP2 (127 DAS) (left arrows in Figure 12). The dynamics of tiller number showed no differences between EXPI and EXP2 (data not shown). The first plants of EXPI had entered the beginning of the anthesis stage 214 DAS (1056 °Cd after emergence) while in EXP2 a similar stage was reached 12 DAS later but 79 °Cd earlier at 226 DAS (977 °Cd) (arrows in Figure 12 and Figure 13). Similarly, physiological maturity (BBCH 87) was observed 6 DAS earlier in EXPI (254 DAS, 1960 °Cd after emergence) than in EXP2 (260 DAS, 1815 °Cd). In general, development in EXP2 was slower than in EXPI.

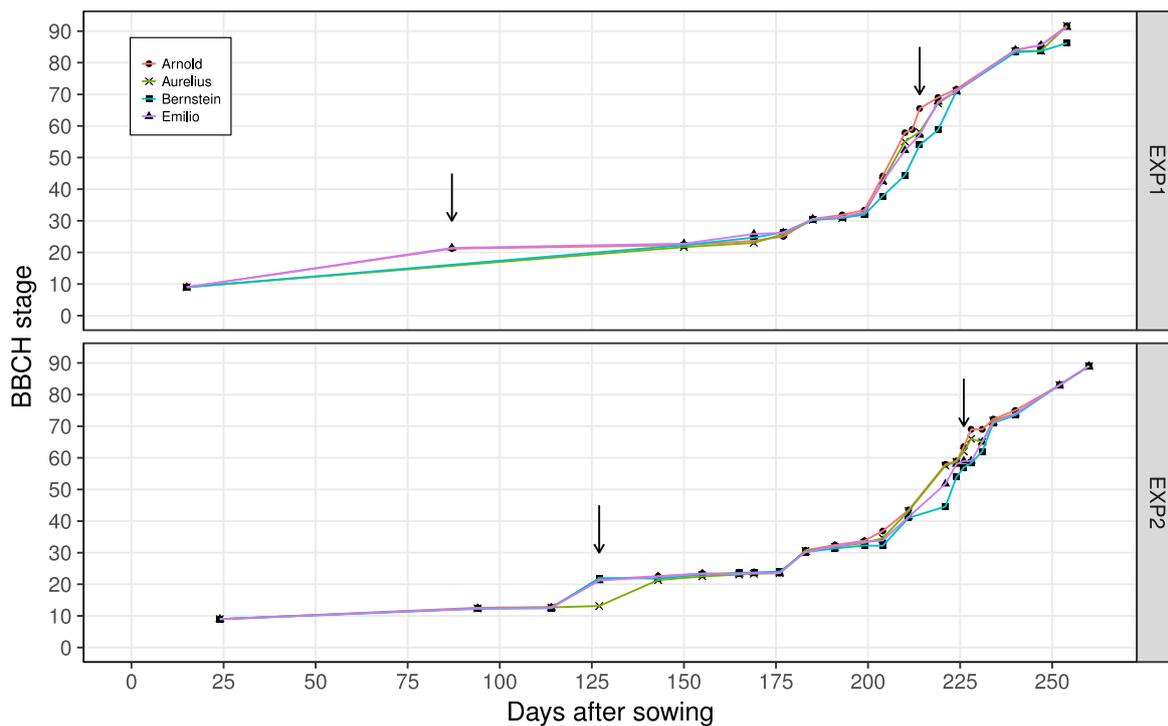


Figure 12 Phenological development (BBCH stages) of the four winter wheat cultivars in both seasons (EXP1: 2017/18, EXP2: 2018/19). Arrows indicate the samplings with first occurrence of (left to right) tillering (BBCH 21+) and anthesis (BBCH 61+).

#### Effect of N-Treatment

The N-treatment appeared to influence development only in EXP2 at a few single measurement dates (data not shown). The largest differences were observed at tiller initiation (ca. mid February 2019) where plants in the N0/N3 treatments were in BBCH 13.2/21.5 (Arnold) and 21.5/12.8 (Bernstein) and at beginning of anthesis (ca. end of May 2019) where plants in N0/N3 were in BBCH stage 64.0/59.0 (Arnold) and 58.3/55.5 (Bernstein). The data indicated a trend of accelerated phenological development of N0, except for one contradictory measurement of Arnold at tiller initiation. However, overall the differences between N0 and N3 rarely exceeded 1 on the BBCH scale. It needs to be mentioned that quantitative comparisons between phenological measurements should be used with caution, as the BBCH scale is not linear among main stages (e.g. between leaf development, BBCH 1x, and tillering, BBCH 2x) and also not necessarily linear within main stages. Still, quantitative comparisons of BBCH give an indication of the direction of different development rates (such as which cultivars or N-treatments were earlier, equal, or later).

#### Effect of Cultivar

In EXP2 the cultivars Arnold, Bernstein, and Emilio started tillering earlier than Aurelius (Figure 12), while in EXP1 measurement intervals during tiller initiation were too large to observe differ-

ences. Tiller dynamics and beginning of stem elongation (BBCH 30) were similar across the cultivars in both experiments. Thereafter until milk development (BBCH 70), the cultivars showed the largest differences in both years (Figure 13). The phenological development of cultivar Arnold was fastest, while Bernstein was slowest (both seasons). The cultivars Aurelius and Emilio ranged in between Arnold and Bernstein but were generally closer to Arnold. In EXP1 Arnold had reached anthesis 214 DAS (observed BBCH: 65.5) while for the other cultivars anthesis was observed 5 days later (Aurelius: BBCH 67.1, Emilio: BBCH 67.6) and even later for Bernstein where an anthesis measurement was missed (224 DAS Bernstein had already reached BBCH 71). Anthesis in EXP2 was first observed 226 DAS for Arnold (BBCH 63.5) and Aurelius (BBCH 62.0) and 5 days later for Bernstein (BBCH 62.0) and Emilio (BBCH 65.0). From grain development (BBCH 71) onwards, differences between cultivars disappeared, only showing a tendency of faster development of Arnold in EXP1 (Figure 12).

Overall, the main observed differences were faster development of cultivar Arnold, intermediate development of the cultivars Aurelius and Emilio, and slowest development of Bernstein, while N-fertilisation had only negligible effects.

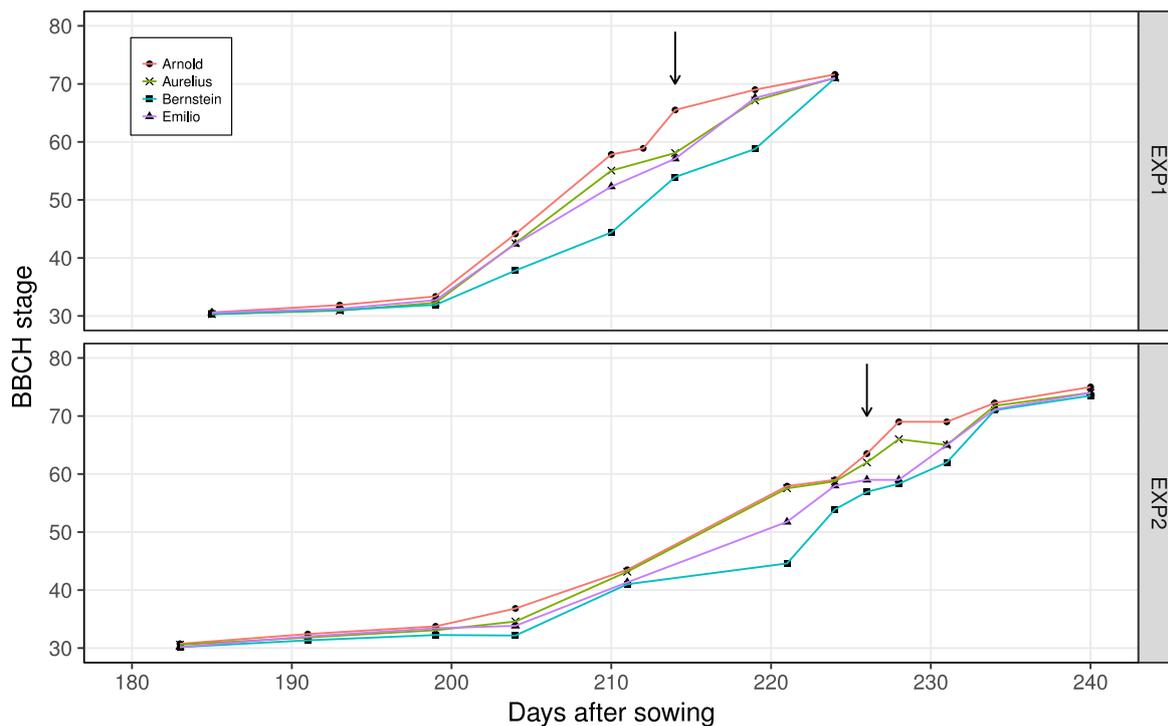


Figure 13 Phenological development from stem elongation (BBCH 30) until end of milk development (BBCH 79) of the four winter wheat cultivars in both seasons (EXP1: 2017/18, EXP2: 2018/19). The arrow indicates the sampling with the first occurrence of anthesis (BBCH 61+). This figure shows the same data as Figure 12 but zoomed to the above mentioned BBCH stages.

#### 8.1.5. Leaf Canopy Development

##### *Tillers*

The number of tillers was measured up until the emergence of the flag leaf (EXP1: 8 May 2018, EXP2: 24 May 2019). Tiller number on that date ranged from 0.5 (Arnold N0, EXP2) to 2.0 (Emilio N2, EXP1). When pooling both experiments and considering only N0 and N3 (N1 and N2 were not measured in EXP2), none of the factors and factor-interactions influenced tiller number significantly. However, when analysing each experiment separately, only the N treatment in EXP2 had significant influence ( $p = 0.01$ ) with N3 showing 1.4 tillers on average and N0 only 0.8. Cultivar averages (both experiments pooled) were 1.1 (Aurelius, Bernstein), 1.2 (Arnold), and 1.4 (Emilio).

##### *Leaf Area Index*

Leaf area index (LAI) at anthesis ranged from 0.7 (Arnold N0, EXP2) to 4.0 (Bernstein N1 and N3, EXP1) (Table 3). LAI was significantly affected by the N-treatment showing highest LAI under highest N-treatments (N3 mean: 3.1 and 3.2 in EXP1 and EXP2, resp.) and a mostly gradual decline towards N0 (mean: 1.8 and 1.0 in EXP1 and EXP2, resp.). Cultivar effect on LAI was significant. Bernstein showed highest average LAI in both years (3.5 and 2.6 in EXP1 and EXP2,

resp.), Emilio (2.3 and 2.0) and Arnold (2.1 and 1.9) the lowest. Between the seasons, LAI of N3 was similar except for Arnold, while LAI of N0 was always lower in EXP2.

Table 3 Leaf area index (LAI), main stem node number (MSNN), leaf fraction, and specific leaf area (SLA) at anthesis for four winter wheat cultivars (Cv.) under different N-treatments in two seasons (EXP1: 2017/18, EXP2: 2018/19). The N-treatments included no N (N0), 70 (N1), 140 (N2), and 210 (N3) kg N ha<sup>-1</sup>.

Cultivar	N-trt.	LAI (mm <sup>2</sup> mm <sup>-2</sup> )		MSNN (-)		Leaf frac. (g g <sup>-1</sup> )		SLA (mm <sup>2</sup> g <sup>-1</sup> )	
		EXP1	EXP2	EXP1	EXP2	EXP1	EXP2	EXP1	EXP2
Arnold	N0	1.4	0.7	10.6	10.5	0.19	0.18	171	164
	N1	2.1	1.7	10.6	NA	0.20	0.19	169	191
	N2	2.4	2.3	10.6	NA	0.20	0.19	179	197
	N3	2.4	2.8	11.1	11.0	0.20	0.19	170	194
Aurelius	N0	2.3	1.4	11.7	11.2	0.21	0.18	207	182
	N1	3.0	2.1	11.7	NA	0.22	0.19	201	196
	N2	3.0	2.7	11.6	NA	0.24	0.20	201	204
	N3	3.4	3.3	11.6	11.2	0.24	0.21	205	205
Bernstein	N0	2.3	1.1	11.2	10.2	0.21	0.17	203	200
	N1	4.0	2.4	10.4	NA	0.23	0.20	216	200
	N2	3.8	3.1	11.3	NA	0.26	0.23	184	206
	N3	4.0	3.9	11.1	10.7	0.23	0.24	215	210
Emilio	N0	1.4	0.8	11.4	10.8	0.20	0.16	155	188
	N1	2.4	1.9	11.3	NA	0.19	0.17	177	196
	N2	2.6	2.3	11.7	NA	0.22	0.19	174	207
	N3	2.8	2.8	11.9	10.7	0.21	0.19	181	204
<b>Source of variation</b>									
Season		*		ns (0.89) <sup>1</sup>		ns (0.24)		ns (0.46)	
Cultivar		***		*** <sup>1</sup>		*		***	
N		***		ns (0.06) <sup>1</sup>		ns (0.37)		ns (0.57)	
Season x cultivar		*		ns (0.13) <sup>1</sup>		**		***	
Season x N		***		ns (0.96) <sup>1</sup>		ns (0.15)		ns (0.08)	
Cultivar x N		ns (0.28)		ns (0.41) <sup>1</sup>		**		ns (0.35)	

\*\*\* Significant at p < 0.001

\*\* significant at p < 0.01

\* significant at p < 0.05; ns: not significant

<sup>1</sup> only N0 and N3 considered

#### Main Stem Leaf Number

The final number of leaves (nodes) on the main stem (MSNN) ranged from 10.2 (Bernstein N0, EXP2) to 11.9 leaves (Emilio N3, EXP1) (Table 3). MSNN was unaffected by N-treatment but varied significantly between the cultivars in both experiments. Bernstein and Arnold showed the lowest (both 10.7) average MSNN, followed by Emilio (11.2) and Aurelius (11.4).

#### Total Number of Green Leaves

The total number of green leaves was only measured in EXP2. At anthesis, it ranged from 946 (Emilio N0) to 2671 leaves per m<sup>2</sup> (Bernstein N3) showing no differences between the cultivars

( $p = 0.46$ ) but significant influence of the N treatment ( $p < 0.001$ ). N3 (2405 leaves  $m^{-2}$ ) and N2 (2213) were significantly higher than N1 (1807) which was higher than N0 (1100) (Tukey post-hoc). While cultivars did not differ significantly, they showed a trend with Bernstein (mean 2070) and Aurelius (2065) having higher average green leaf number than Arnold (1695) and Emilio (1694). When considering only the highly fertilised treatments (N2, N3), neither cultivar nor N treatment were significant factors.

#### *Leaf Fraction*

Leaf fraction, i.e. the ratio of leaf dry mass to total above-ground dry mass, at anthesis ranged from 0.16 to 0.26 with a non-significant tendency of lower values in EXP2 (Table 3). Regarding cultivar differences, Bernstein showed the highest leaf fraction (mean 0.22) followed by Aurelius (0.21) and Arnold and Emilio (both 0.19). The significant interaction effect of season x cultivar showed high values for Aurelius in EXP1 (mean 0.23) but low in EXP2 (mean 0.19) while the other cultivars showed rather similar values across the two seasons. Regarding the cultivar x N-treatment effect the observed response of Emilio to increased N-fertilisation was erratic, showing no change between N0 and N1 (mean: 0.18, both) followed by a jump to N2 (0.20) and then again no response to N3 (0.20). In contrast, the other cultivars responded relatively continuous to increased N-fertilisation between N0 and N2 (e.g. Arnold, mean: 0.18, 0.19, 0.20, 0.19; N0 to N3, resp.) while the response to N3 was mostly minor or slightly negative.

#### *Specific Leaf Area*

The specific leaf area (SLA) at anthesis ranged from 155 (Emilio N0, EXP2) to 216  $mm^2 g^{-1}$  (Bernstein N1, EXP2) (Table 3). The effect of cultivar was significant (means: Bernstein 204, Aurelius 200, Emilio 185, Arnold 179  $mm^2 g^{-1}$ ). Looking at the season x cultivar effect SLA, was higher in EXP2 than in EXP1 for Arnold (mean: 172 and 186  $mm^2 g^{-1}$ , EXP1 and EXP2, resp.) and Emilio (172 and 199) while it was similar in both seasons for Bernstein (204 in both seasons) and Aurelius (203 and 197).

#### 8.1.6. Yield and Dry Mass

Most dry mass related variables measured at harvest showed a stronger influence of the N-treatment in EXP2 (Table 4).

Table 4 Grain yield, total above-ground dry mass (TDM), total above-ground N-uptake (NUP), grain N concentration (%GN), number of ears, thousand kernel weight (TKW), and number of seeds per ear (Seeds) at harvest for four winter wheat cultivars (Cv.) under different N-treatments in two seasons (EXP1: 2017/18, EXP2: 2018/19). The N-treatments included no N (N0), 70 (N1), 140 (N2), and 210 (N3) kg N ha<sup>-1</sup>.

Season	Cv.	N-trt.	Yield (g m <sup>-2</sup> )	TDM (g m <sup>-2</sup> )	NUP (g m <sup>-2</sup> )	%GN (%)	Ear no. (m <sup>-2</sup> )	TKW (g)	Seeds (#)
EXP1	Arnold	N0	424	1064	10.8	2.1	434.7	40	24
		N1	558	1309	17.0	2.6	453.0	42	29
		N2	575	1370	19.4	2.9	488.0	40	30
		N3	583	1340	20.1	3.0	473.7	41	30
	Aurelius	N0	540	1248	12.0	1.8	391.0	44	31
		N1	701	1543	19.8	2.4	454.0	45	34
		N2	706	1559	22.1	2.6	445.0	42	37
		N3	705	1560	23.7	2.7	464.0	42	36
	Bernstein	N0	519	1265	11.2	1.7	400.7	43	30
		N1	676	1572	19.1	2.3	435.7	43	36
		N2	719	1676	23.4	2.7	466.3	43	36
		N3	717	1636	23.2	2.7	477.7	42	36
Emilio	N0	484	1246	11.0	1.8	464.7	36	29	
	N1	625	1541	17.7	2.3	490.7	37	35	
	N2	609	1539	20.7	2.7	507.3	36	34	
	N3	648	1545	20.7	2.6	527.7	35	35	
EXP2	Arnold	N0	268	708	6.9	2.0	367.0	37	20
		N1	435	1035	12.1	2.3	424.0	38	27
		N2	596	1302	18.2	2.6	485.3	38	32
		N3	683	1520	23.1	2.8	531.7	39	33
	Aurelius	N0	324	788	7.0	1.7	352.0	39	23
		N1	545	1165	12.3	1.9	400.0	40	34
		N2	739	1540	19.3	2.2	455.3	39	41
		N3	788	1638	23.8	2.6	479.7	39	43
	Bernstein	N0	303	729	6.6	1.7	318.3	37	26
		N1	553	1226	12.8	1.9	371.3	37	41
		N2	708	1512	19.1	2.2	428.3	36	46
		N3	693	1551	21.5	2.5	463.0	34	44
Emilio	N0	297	734	5.9	1.6	367.0	35	23	
	N1	515	1158	11.0	1.8	393.0	35	37	
	N2	694	1483	18.5	2.2	501.0	34	40	
	N3	741	1632	22.4	2.5	518.0	33	43	
<b>Source of variation</b>									
Season			***	***	***	ns (0.33)	***	***	***
Cultivar			***	*	ns (0.15)	***	**	***	*
N			***	***	***	***	**	*	ns (0.07)
Season x cultivar			*	*	*	*	**	***	**
Season x N			***	***	***	***	***	ns (0.37)	***
Cultivar x N			ns (0.49)	ns (0.83)	ns (0.53)	ns (0.89)	ns (0.86)	**	ns (0.10)

\*\*\* Significant at p < 0.001

\*\* Significant at p < 0.01

\* Significant at p < 0.05

ns: not significant

#### Grain Yield

Observed wheat grain yields ranged from 268 (Arnold N0, EXP2) to 788 g m<sup>-2</sup> (Aurelius N3, EXP2) with both the lowest and highest values (except for Bernstein's highest yield) in EXP2 (Table 4 and Figure 14). Mean yield of EXP1 (619 g m<sup>-2</sup>) was significantly higher than for EXP2 (572 g m<sup>-2</sup>), while mean yield of only N3 was lower in EXP1 (663 g m<sup>-2</sup>) than in EXP2 (726 g m<sup>-2</sup>). Cultivar and N-treatment showed significant effects on yield. For each cultivar, N0 showed always the lowest yield (N0 mean: 395 g m<sup>-2</sup>), and N2 or N3 the highest (mean: N2: 672,

N3: 695 g m<sup>-2</sup>). In EXP1 the response to N-fertilisation was strongest between N0 and N1 while further increases in N-fertilisation (N2, N3) showed only marginal effects. Contrary, in EXP2 the effect of N was more equally distributed from N0 to N3 (except for Bernstein between N2 and N3). Among cultivars, Aurelius showed the highest yields (mean: 631 g m<sup>-2</sup>) followed by Bernstein (611 g m<sup>-2</sup>), Emilio (577 g m<sup>-2</sup>), and Arnold (515 g m<sup>-2</sup>). Regarding the interaction effect season x cultivar, Bernstein showed the strongest response to season with a drop of mean yield between EXP1 and EXP2 of almost 100 g m<sup>-2</sup> (658 to 565 g m<sup>-2</sup>) while Emilio responded with a reduction of less than 30 g m<sup>-2</sup> (591 to 562 g m<sup>-2</sup>). Concerning the season x N-treatment interaction, mean yields of N0 and N1 dropped from EXP1 to EXP2 by ca. 200 and 130 g m<sup>-2</sup>, respectively, while there was a slightly increasing effect of season in N2 (plus ca. 40 g m<sup>-2</sup>) and N3 (plus ca. 60 g m<sup>-2</sup>).

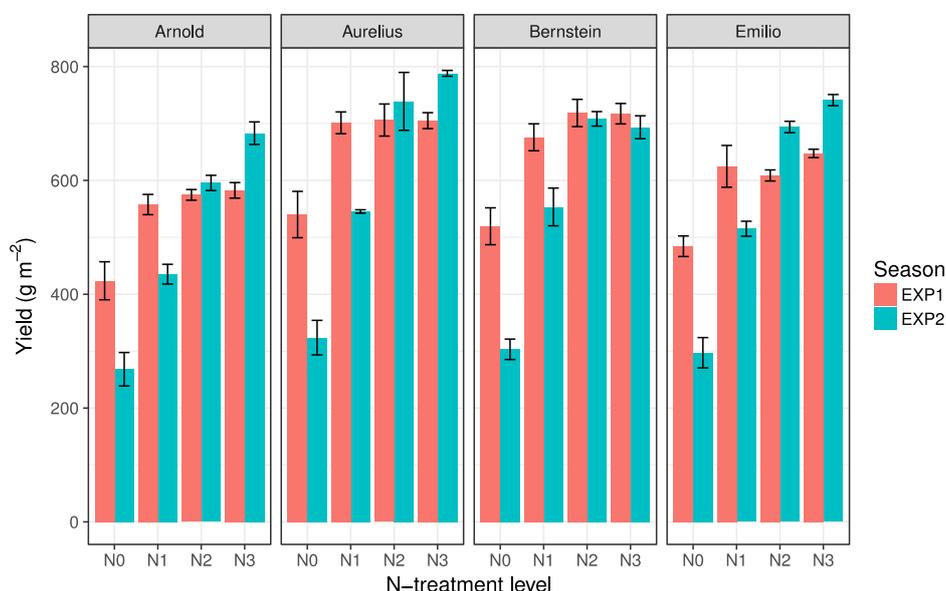


Figure 14 Grain yield of the four winter wheat cultivars under different N-treatments in two seasons (EXP1: 2017/18, EXP2: 2018/19). The N-treatments included no N (N0), 70 (N1), 140 (N2), and 210 (N3) kg N ha<sup>-1</sup>. Bars represent standard errors of the means.

#### Dry Mass

Observed total above-ground dry mass (TDM) at harvest ranged from 708 (Arnold N0, EXP2) to 1676 g m<sup>-2</sup> (Bernstein N2, EXP1) with the lowest values in N0 of EXP2 (Table 4 and Figure 15).

Overall, the response of TDM to season, N-treatment, cultivar, and the tested interactions were similar to those described for yield. Mean TDM in EXP1 (1440 g m<sup>-2</sup>) was significantly higher than in EXP2 (1267 g m<sup>-2</sup>), while N3 TDM was lower in EXP1 (1520 g m<sup>-2</sup>) than in EXP2 (1585 g m<sup>-2</sup>). Contrary, at anthesis N3 TDM was higher in EXP1 (1158 g m<sup>-2</sup>) than in EXP2 (1037 g m<sup>-2</sup>). N-treatment showed significant effect on TDM at harvest. The N0 treatment resulted in the lowest TDM (N0 mean: 973 g m<sup>-2</sup>) followed by N1 (1319), N2 (1494), and N3 (1553).

In EXP1 the response to N-fertilisation was strongest between N0 (mean: 1206 g m<sup>-2</sup>) and N1 (1491) while further increases in N-fertilisation (N2: 1499, N3: 1520) showed only marginal effects. Contrary, in EXP2 the effect of N was more equally distributed (N0: 740, N1: 1146, N2: 1488, N3: 1585). Among cultivars, Bernstein showed the highest mean TDM (1396 g m<sup>-2</sup>) followed by Aurelius (1380 g m<sup>-2</sup>), Emilio (1360 g m<sup>-2</sup>), and Arnold (1206 g m<sup>-2</sup>). Regarding the interaction effect season x cultivar, Bernstein showed the strongest response to season with a drop of mean TDM between EXP1 and EXP2 of almost 300 g m<sup>-2</sup> (1537 to 1254 g m<sup>-2</sup>) while Arnold's response was smallest (1271 to 1141 g m<sup>-2</sup>). The significant season x N-treatment interaction effect showed different mean TDM responses to the season depending on the N-treatment. While mean TDM responded negatively to the season (EXP1 to EXP2) for N0 (minus ca. 460 g m<sup>-2</sup>) and N1 (minus ca. 350 g m<sup>-2</sup>), the response in N2 (minus ca. 10 g m<sup>-2</sup>) and N3 (plus ca. 60 g m<sup>-2</sup>) was positive but much smaller in absolute values.

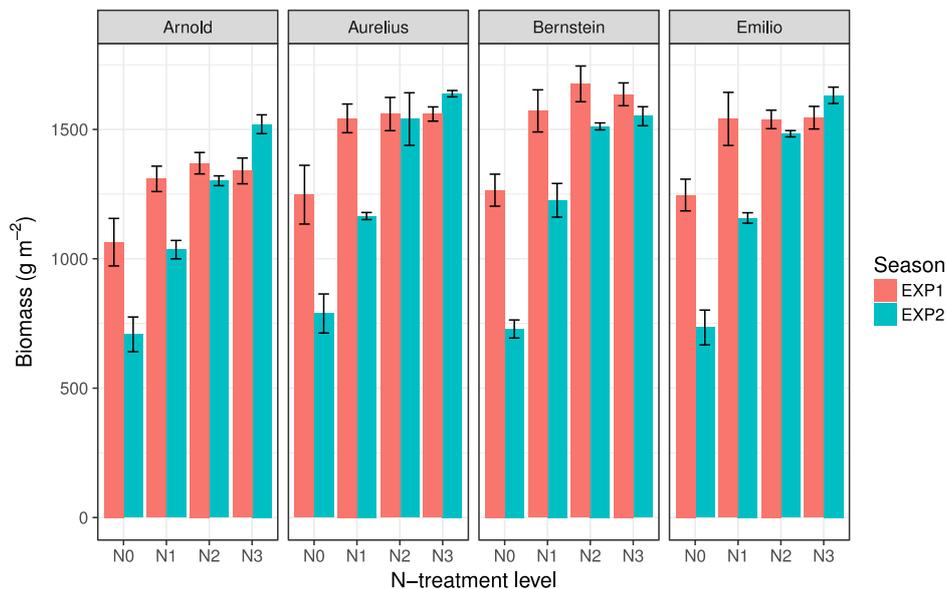


Figure 15 Biomass (total above-ground dry mass) of the four winter wheat cultivars under different N-treatments in two seasons (EXP1: 2017/18, EXP2: 2018/19). The N-treatments included no N (N0), 70 (N1), 140 (N2), and 210 (N3) kg N ha<sup>-1</sup>. Bars represent standard errors of the means.

#### 8.1.7. Crop Nitrogen

Total above-ground N-uptake (NUP) at harvest ranged from 5.9 (Emilio N0, EXP2) to 23.8 g m<sup>-2</sup> (Aurelius N3, EXP2) (Table 4 and Figure 16). Mean NUP in EXP1 was higher (18.2 g N m<sup>-2</sup>) than in EXP2 (15.0 g N m<sup>-2</sup>). There was no significant influence of the cultivar. The effect of N-treatment was significant. NUP increased gradually from N0 (mean: 8.9 g N m<sup>-2</sup>) to N1 (15.2), N2 (20.0), and N3 (22.3). Similar to yield and TDM, the positive effect of N-treatment on N-uptake was larger in EXP2 than in EXP1.

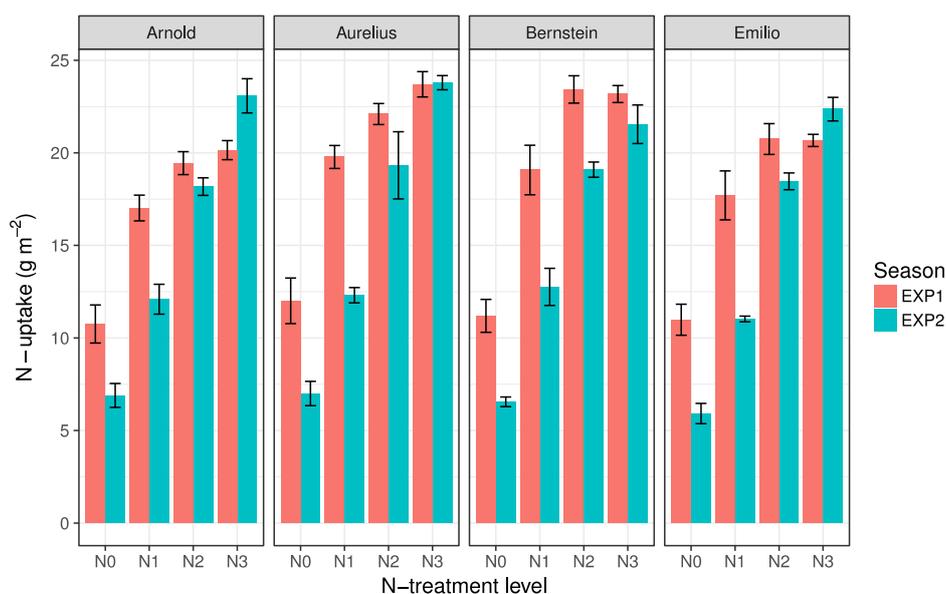


Figure 16 N-uptake of the four winter wheat cultivars under different N-treatments in two seasons (EXPI: 2017/18, EXP2: 2018/19). The N-treatments included no N (N0), 70 (N1), 140 (N2), and 210 (N3) kg N ha<sup>-1</sup>. Bars represent standard errors of the means.

Grain N concentration (%GN) (Table 4) at harvest ranged from 1.6% (Emilio N0, EXP2) to 3.0% (Arnold N3, EXP1). There was no significant influence of the season. The effects of N-treatment and cultivar were significant. Increase of N-fertilisation affected %GN positively (means: N0: 1.8%, N1: 2.2%, N2: 2.5%, N3: 2.7%). Arnold showed highest mean %GN (2.5%) compared to all other cultivars (Aurelius, Bernstein, Emilio; each 2.2%).

#### 8.1.8. Yield Components

The number of ears at harvest ranged from 318 (Bernstein N0, EXP2) to 532 ears m<sup>-2</sup> (Arnold N3, EXP2). Ear number was higher in EXPI (468) than in EXP2 (436). All treatments and their pairwise interactions were significant, except N x cultivar. Highest mean values were observed for the cultivars Emilio (471) and Arnold (457) followed by Aurelius (430) and Bernstein (420). N-treatment had a positive effect on ear number (means: N0: 387, N1: 428, N2: 483, N3: 492).

The thousand kernel weights (TKW) ranged from 33 g (Emilio N3, EXP2) to 45 g (Aurelius N1, EXP1) (Table 4). TKW was lower in EXPI (mean: 34 g) than in EXP2 (37 g). Effects of cultivar, N-treatment, and their interactions were significant. In both seasons, cultivar Emilio had the lowest TKW (mean: 35 g). The effect of increased N-fertilization on TKW was diverse, with negative effects in a few cases (in EXPI: Aurelius and Bernstein, in EXP2: Bernstein and Emilio), a clear positive effect in one case (EXP2: Arnold), and varying responses in the remaining factor combinations (e.g. EXPI: Arnold).

The number of seeds per ear at harvest ranged from 20 (Arnold N0, EXP2) to 46 (Bernstein N2, EXP2). EXPI showed less seeds per ear (mean: 33) than EXP2 (35). This was mainly due to low

seed number of N0 and N1 treatments in EXP1, while N2 and N3 treatments of EXP1 showed higher seed number than EXP2. The effect of cultivar was significant, N-treatment was not ( $p = 0.07$ ). Seed number per ear showed a mostly positive response to N-treatment which was more apparent in EXP2 than in EXP1. Arnold had the lowest mean seed number (mean: 28) followed by Aurelius and Emilio (both 35) and Bernstein (38).

## 8.2. iCrop Source Code Updates

Based on comparisons between model estimates and field experimental data, the crop model code was modified. I introduced functionalities to enable wheat simulation based on (i) two phases of phyllochron and (ii) two phases of minimum stem nitrogen concentration (*SNCS*). The latter was implemented in a similar way as for maize in Manschadi et al. (2020a).

For phyllochron (i), the two phases were separated by a threshold main stem leaf number (*LNP2*) which triggered the switch from the first phase phyllochron (*PHYL1*) to the second phase (*PHYL2*). In the default version of iCrop only one phyllochron was specified for the whole node production period (from emergence to termination of leaf production on the main stem, i.e. flag leaf emergence; BBCH 37).

For *SNCS* (ii), the parameter switch from the first phase (*SNCS1*) to the second phase (*SNCS2*) occurred on the first day of the beginning seed growth phenological phase. By default, iCrop used only one *SNCS* parameter for the entire simulation.

## 8.3. iCrop Setup and Parameters

The important genotype parameters derived for iCrop in this study are presented in Table 5.

Table 5 iCrop parameters derived for the four winter wheat cultivars Arnold, Aurelius, Bernstein, and Emilio based on field data from two seasons (2017/18 and 2018/19) and four N-fertilisation treatments (0, 70, 140, and 210kg N ha<sup>-1</sup>) at one location (Tulln, Eastern Austria).

Parameter	Description	Unit	Cultivar			
			Arnold	Aurelius	Bernstein	Emilio
PLAPOW	Exponent in leaf area expansion calculation	1	1.9	1.9	2	1.9
SLA	Specific leaf area	m <sup>2</sup> g <sup>-1</sup>	0.021	0.021	0.021	0.021
IRUE	Radiation use efficiency under optimum conditions	g MJ <sup>-1</sup>	2.58	2.58	2.58	2.58
FLFIA	Leaf fraction in phase 1	g g <sup>-1</sup>	0.72	0.76	0.80	0.72
FLFIB	Leaf fraction in phase 2	g g <sup>-1</sup>	0.29	0.29	0.37	0.29
WTOPL	Total above-ground dry mass as breakpoint between FLFIA and FLFIB	g m <sup>-2</sup>	123	117	113	123
PHYL1	Phyllochron in phase 1	°C leaf <sup>-1</sup>	55	60	65	50
PHYL2	Phyllochron in phase 2	°C leaf <sup>-1</sup>	130	105	120	110
LNP2	Leaf number as breakpoint between PHYL1 and PHYL2	1	6.5	6.5	6.0	6.5
SNCSI	Minimum stem N concentration in phase 1	g g <sup>-1</sup>	0.0063	0.0063	0.0063	0.0063
SNCS2	Minimum stem N concentration in phase 2	g g <sup>-1</sup>	0.0022	0.0022	0.0022	0.0022
SLNG	Specific leaf N in green leaves	g m <sup>-2</sup>	2.10	2.10	2.10	2.10
SLNS	Specific leaf N in senesced leaves	g m <sup>-2</sup>	0.10	0.10	0.10	0.10
SNCG	Maximum stem N concentration	g g <sup>-1</sup>	0.019	0.019	0.019	0.019
%GNmin	Minimum grain N concentration	g g <sup>-1</sup>	0.016	0.016	0.016	0.016
%GNmax	Maximum grain N concentration	g g <sup>-1</sup>	0.029	0.029	0.029	0.029
MXNUP	Maximum daily N-uptake	g m <sup>-2</sup> d <sup>-1</sup>	0.6	0.6	0.6	0.6
vsen	Vernalisation sensitivity	1	0.018	0.018	0.018	0.018
cpp	Critical photoperiod	h	21	21	21	21
ppsen	Photoperiod sensitivity	1	0.004	0.004	0.004	0.004
bdSOWEMR	Biological days from sowing to emergence	bd	4.5	4.5	4.5	4.5
bdEMRTIL	... emergence to tillering	bd	2	2	2	2
bdTILSEL	... tillering to stem elongation	bd	12.5	13.0	13.5	13.0
bdSELBOT	... stem elongation to booting	bd	7.0	7.0	8.0	7.0
bdBOTEAR	... booting to ear emergence	bd	3.0	3.0	4.0	4.0
bdEARANT	... ear emergence to anthesis	bd	5.0	5.0	6.0	6.0
bdANTPM	... anthesis to plant maturity	bd	33.5	33.5	31.0	33.0

### 8.3.1. Management

The model was set up to start simulations a few weeks before sowing, matching the first measurements of soil water content (EXP1: -33 DAS, 13 Sept. 2017; EXP2: -14 DAS, 1 Oct. 2018). Wheat sowing density (all treatments: 375 seeds m<sup>-2</sup>), N-fertilisation dates (EXP1: 99 and 128 DAS; EXP2: 67 and 100 DAS) and amounts (N0: 0, N1: 3.5, N2: 7, N3: 10.5 g N m<sup>-2</sup> per application) were all set to the exact applied values. Initial soil water content was calibrated based on measured initial soil water and comparison of simulated and observed soil water dynamics.

### 8.3.2. Environment

The vapour pressure deficit factor (VPDF) was calibrated to 0.60 which was slightly below the common range of 0.65 to 0.75 (wet and dry environments, respectively) (Soltani and Sinclair,

2012). Reason for choosing a value outside the common range was the better fit of simulated to observed soil water content.

The atmospheric carbon dioxide (CO<sub>2</sub>) concentration was set to 400 ppm since average global CO<sub>2</sub> levels have exceeded this value in ca. 2016 (Butler and Montzka, 2020). The previous default value in iCrop was 350 ppm.

### 8.3.3. Emergence

The minimum threshold fraction of transpirable soil water (FTSW) that inhibited crop development for germination (i.e. from sowing to emergence) was set to 0.3 (default: 0.0, for details see discussion).

### 8.3.4. Vernalisation and Photoperiod

The vernalisation parameter (vsen) was calibrated to 0.018. The critical photoperiod (cpp) was set to 21 h, photoperiod sensitivity (ppsen) to 0.004. These parameters were set identical for all cultivars in both experiments.

### 8.3.5. Development Phases

The calibrated parameters for the phenological phases defined by iCrop (Table 5) were based on observed differences in development rate between the cultivars (Figure 17). The parameters I chose did not differ between the cultivars until the beginning of tillering (bdSOWEMR, sowing until emergence, and bdEMRTIL, emergence until tillering). In the phase from emergence until stem elongation (BBCH 09 – 31), observations showed faster development of Arnold in both seasons (baseline: average of all four cultivars), as well as slower development of Bernstein in both seasons (Figure 17). This was reflected in the parameter bdTILSEL (phase tillering until stem elongation) (Table 5). The observed slower developments of Bernstein and Emilio in both seasons from BBCH 31 to 61 (the sum of bdSELBOT, bdBOTEAR, and bdEARANT) was also set in the parameters, with Bernstein requiring 18 bd, Emilio 17 bd, and the two remaining cultivars both 15 bd. The phase from beginning of anthesis (BBCH 61) until physiological maturity (BBCH 87) (parameter bdANTPM) showed quickened development of Bernstein in EXP2 only. Emilio showed faster development than the mean in both years. Taking into account measurement difficulties during grain filling, possible effects of drought stress on phenology, and other effects, the parameters were calibrated to a fast development of Bernstein (31 bd) and similarly slower developments of Emilio (33 bd) and the remaining cultivars (both 33.5 bd).

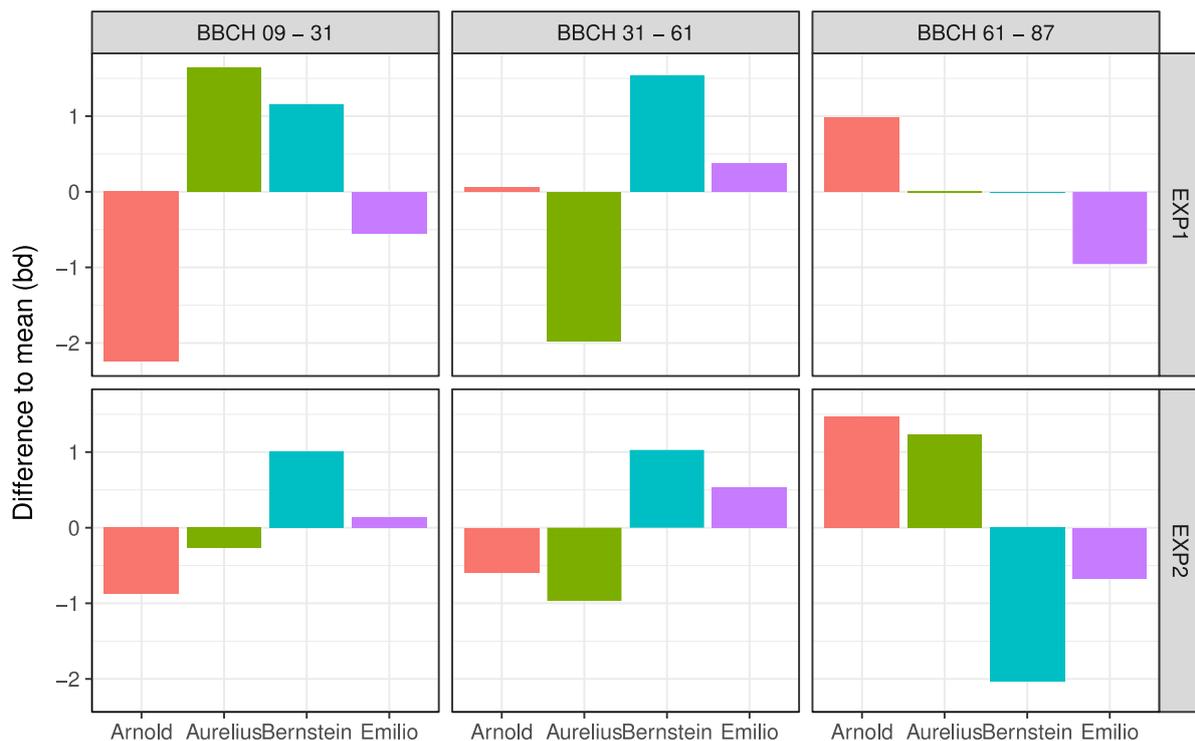


Figure 17 Observed cumulative biological days (bd) differences between each of the four winter wheat cultivars and their mean for the phenological phases from emergence (BBCH 9) to beginning of stem elongation (SEL, BBCH 31), from SEL to beginning of anthesis (ANT, BBCH 61), and from ANT to physiological maturity (BBCH 87), in both seasons (EXP1: 2017/18, EXP2: 2018/19). Data was based on linear interpolation of the phenological stages BBCH 31, 61, and 87 using the actually observed BBCH stages and temperature sums. Conversion to biological days assumed an optimum temperature of 27.5 °C.

### 8.3.6. Phyllochron

The development of leaves on the main stem (main stem node number, *MSNN*) based on temperature sums (phyllochron: required temperature sum to develop one leaf) showed a biphasic pattern for all four cultivars (Figure 18, Table 6). In the first phase, cultivars Emilio and Arnold showed the fastest and Bernstein slowest leaf number increase while in the second phase Aurelius and Emilio were fastest and Arnold slowest.

By default, iCrop used only one phyllochron parameter for the whole leaf development phase. To adequately represent the observed biphasic pattern in iCrop, the model code was updated to support two phyllochron parameters (*PHYLI*, *PHYL2*) and a breakpoint (*LNP2*). For the parameters, calculated phyllochrons were rounded to 5 °C leaf<sup>-1</sup> and breakpoints to 0.5 (Table 5).

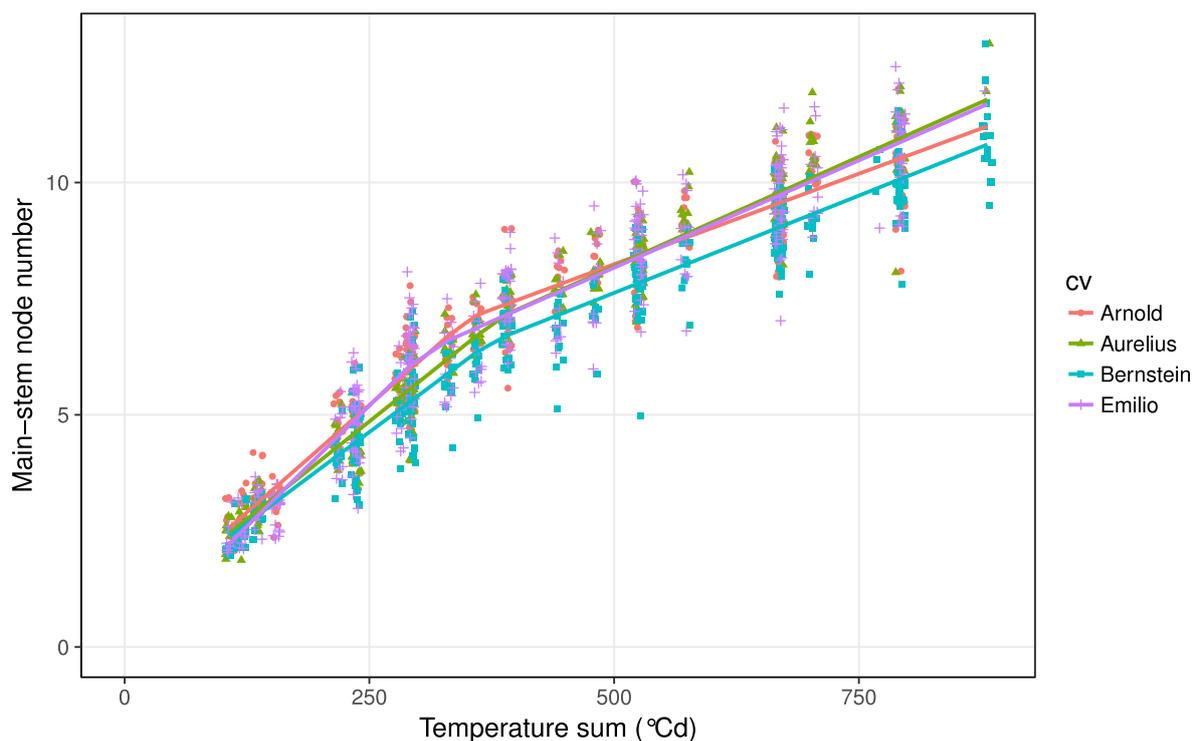


Figure 18 Observed main stem node number vs. temperature sums from BBCH II to 39 for four winter wheat cultivars in two seasons. Lines: two-segment linear regression.

Table 6 Observed phyllochrons (*PHYL1*, *PHYL2*) and breakpoints (*LNP2*) assuming a segmented linear regression for four winter wheat cultivars across two seasons.

Cultivar	Phyllochron (°C leaf <sup>-1</sup> )		<i>LNP2</i> (breakpoint) (leaf no.)
	<i>PHYL1</i>	<i>PHYL2</i>	
Arnold	54.1	129.9	6.5
Aurelius	59.5	106.2	6.4
Bernstein	63.6	119.5	5.9
Emilio	48.2	108.3	6.4

### 8.3.7. Leaf Area

For the calculation of the iCrop parameter *PLAPOW* which defines leaf area expansion in response to main stem leaf number (*MSNN*), I investigated observed plant leaf area and *MSNN* from emergence up until anthesis.

At anthesis, the N-treatment did significantly affect observed LAI but not *MSNN* (Table 3). Due to the effect on LAI I decided to take only well fertilised treatments (N2, N3) into account for the calculation of *PLAPOW*.

The cultivar effect on LAI was significant ( $p < 0.001$ ) and showed a clear trend. Bernstein (3.1 mm<sup>2</sup> mm<sup>-2</sup>) showed highest average LAI followed by Aurelius (2.7), while Emilio (2.1) and Arnold (2.0) were similar with lowest LAI. The plant number was assumed equal to sowing density (370 plants m<sup>-2</sup>; thereby, LAI statistics was representative for leaf area per plant).

Similar to the observed cultivar differences in LAI, calculated PLAPOW was highest for Bernstein ( $2.03 \pm 0.02$  standard error). The cultivars Arnold ( $1.89 \pm 0.02$ ), Aurelius ( $1.88 \pm 0.02$ ), and Emilio ( $1.92 \pm 0.02$ ) were similarly low (Figure 19). For the final model parameter, values were rounded to one decimal.

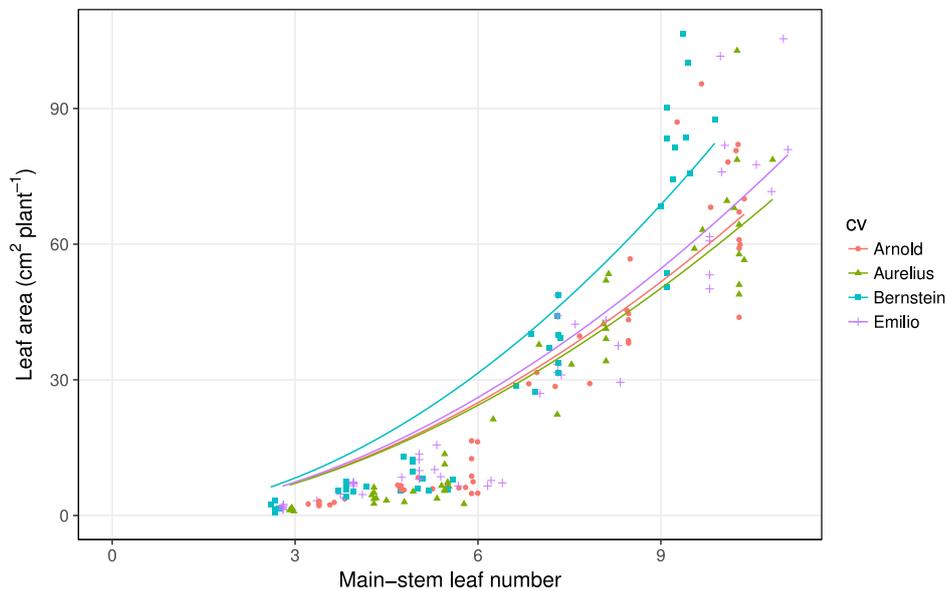


Figure 19 Leaf area per plant versus main stem leaf number for the four winter wheat cultivars with high N-fertilisation (N2, N3) in both seasons. Lines: Regression analysis (power).

For specific leaf area (SLA), data from EXP1 appeared to have outliers (data not shown). Therefore, only EXP2 was investigated. SLA averages in EXP2 ranged from  $0.0191$  to  $0.0240 \text{ m}^2 \text{ g}^{-1}$ . However, this was close to the default value ( $0.0210 \text{ m}^2 \text{ g}^{-1}$ ) and observed data was based on only one sampling (near anthesis). Also, the overall effect of SLA in the crop model is relatively short (only after flag leaf emergence until beginning seed growth) and, therefore, minor. As a result, I decided to keep the default for all cultivars.

#### 8.3.8. Radiation Use Efficiency

The cardinal temperatures for radiation use efficiency (RUE) were kept default ( $^{\circ}\text{C}$ ): base temperature  $TBRUE = 0$ , lower optimum  $TPIRUE = 15$ , upper optimum  $TP2RUE = 22$ , and critical temperature  $TCRUE = 35$ .

The extinction coefficient (KPAR) was also set to the default of 0.65.

I changed the RUE for optimum conditions ( $IRUE$ ) from its default ( $2.2 \text{ g MJ}^{-1}$ ) to  $2.58 \text{ g MJ}^{-1}$ . This value was adopted from the RUE parameter used in the crop growth model APSIM (Holzworth et al., 2014) ( $RUE_{APSIM} = 1.48 \text{ g MJ}^{-1}$ ). The conversion was necessary due to model internal differences in calculating dry matter production (see 7.4).

### 8.3.9. Leaf:Stem Partitioning

Well N-fertilised treatments (N2, N3) of the samplings on 30 April 2018 in EXP1 (196 DAS, ca. BBCH 32) and on 2 May 2019 in EXP2 (198 DAS, ca. BBCH 33) (Figure 20) showed the highest leaf fraction for Bernstein (partly significant) while Arnold and Emilio had lowest. Other samplings in the range BBCH 10 to BBCH 37 showed a mostly similar trend (Figure 21).

The two segmented linear regression of leaf dry mass depending on total dry mass also showed Bernstein clearly differentiated from the other cultivars (Figure 21). I pooled Arnold and Emilio but kept Bernstein and Aurelius separate to calculate iCrop's leaf:stem partitioning parameters *FLFIA*, *FLFIB*, and *WTOPL* (Table 5). Further explanations are given in the Discussion chapter (9.3.4 Dry Mass).

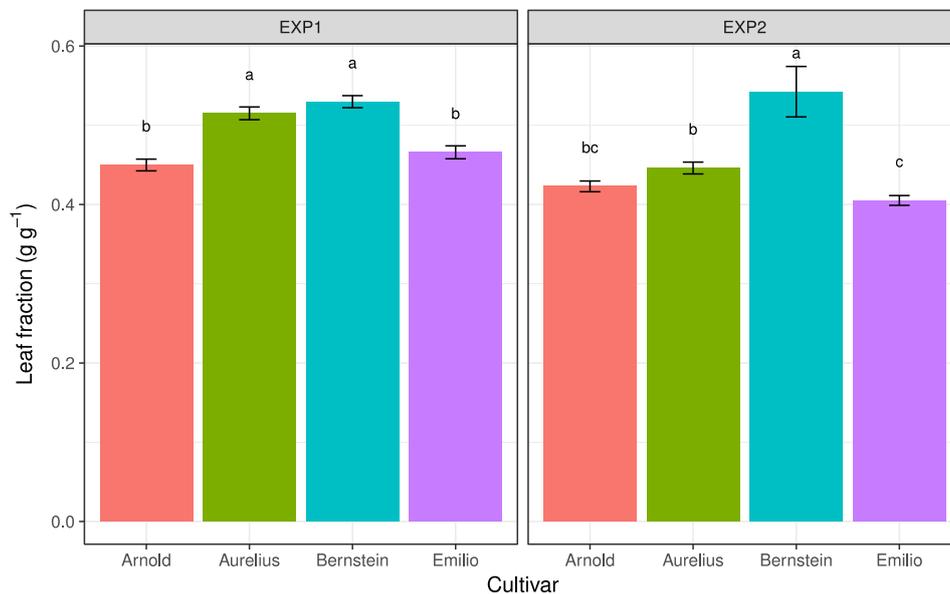


Figure 20 Observed leaf fraction (ratio leaf:total dry mass) for the four winter wheat cultivars on 30 April 2018 in EXP1 (196 DAS, ca. BBCH 32) and on 2 May 2019 in EXP2 (198 DAS, ca. BBCH 33) with high N-fertilisation (average of N2 and N3). Within cultivars, means indicated by the different letters are significantly different at  $p < 0.05$  according to Tukey's post-hoc test

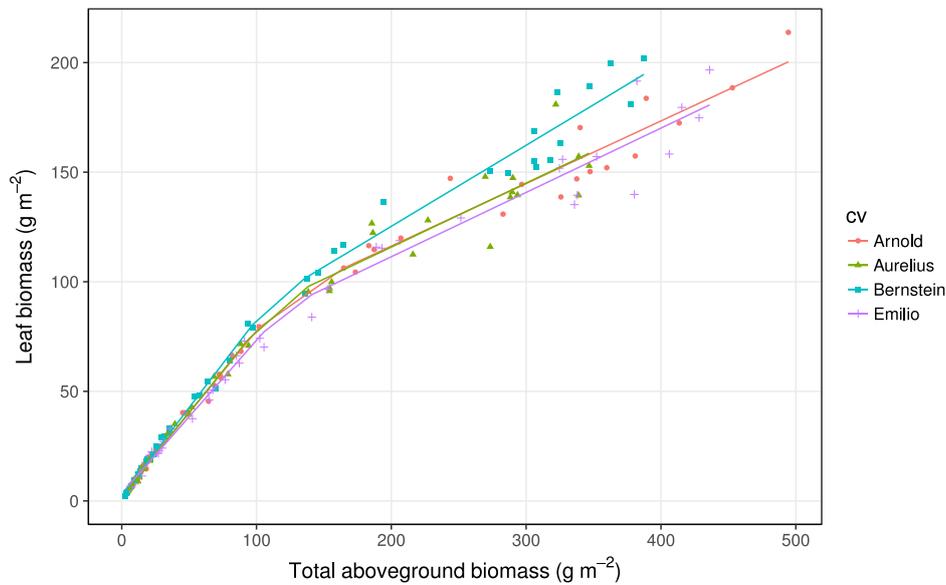


Figure 21 Leaf biomass (dry mass) versus total above-ground biomass from BBCH 9 (emergence) until BBCH 37 (flag leaf emergence) of four winter wheat cultivars in well fertilised treatments (N2, N3) in both experiments (2017/18 and 2018/19). Lines: Regression analysis (segmented linear regression).

For iCrop's leaf fraction parameter *FLF2* (leaf fraction after the flag leaf has fully emerged [BBCH 39] until beginning of seed growth [BBCH 71]) I used the default value of 0.1.

#### 8.3.10. Yield

The observed fraction of translocated dry mass (*FRTRL*, i.e. the part of vegetative crop dry mass available at the beginning of seed growth for translocation to grains until harvest) was 23% on average, ranging from ca. 10% to 35% (Figure 22). While the N-treatment did not show a consistent pattern across cultivars in EXP1, there was a downward trend visible in EXP2. Cultivar averages were 28% (Arnold), 21% (Aurelius), 19% (Bernstein), and 23% (Emilio).

The overall variance of the measurements was high. Also, the *FRTRL* parameter in iCrop defines only the theoretical maximum availability of crop mass for translocation to seeds, not the actual translocated dry mass. Therefore, I used the default parameter value of 22%.

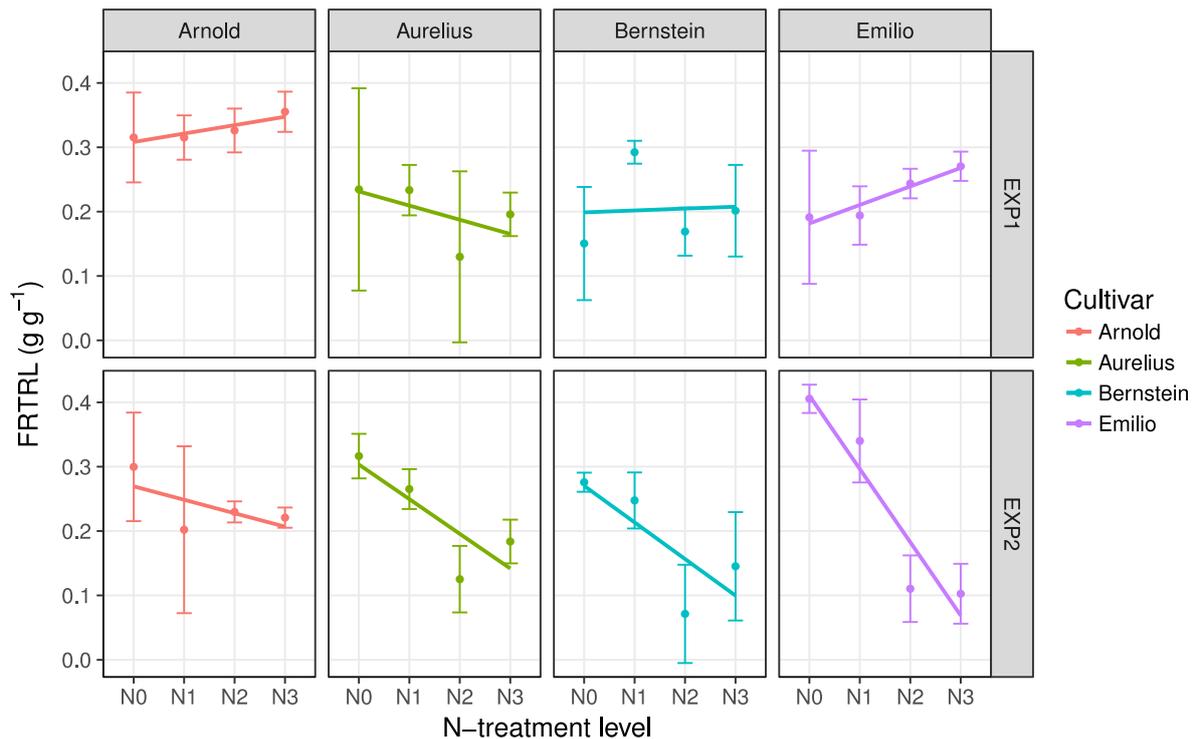


Figure 22 Observed fraction of translocated dry mass (FRTRL) for the four winter wheat cultivars in both seasons (EXP1: 2017/18, EXP2: 2018/19). The N-treatments included no N (N0), 70 (N1), 140 (N2), and 210 (N3) kg N ha<sup>-1</sup>. Lines: linear regression. Error bars indicate standard error.

The observed daily harvest index increase (DHI) was calculated from observed harvest index divided by an estimate of seed growth duration (SGD). SGD was calculated from the time span between BBCH 71 (beginning seed growth) and BBCH 87 (hard dough stage). The dates of these BBCH stages were estimated by linear interpolation versus cumulative thermal time.

DHI showed an increase with higher N-fertilisation rates (Figure 23). Values ranged from 0.014 to 0.020. In EXP1 values were overall lower than in EXP2. Bernstein showed the overall lowest DHI in EXP1. Combined with data from a previous experiment (Fuchs, 2016) I calibrated the parameter *PDHI* (potential DHI) equal for all cultivars to 0.017.

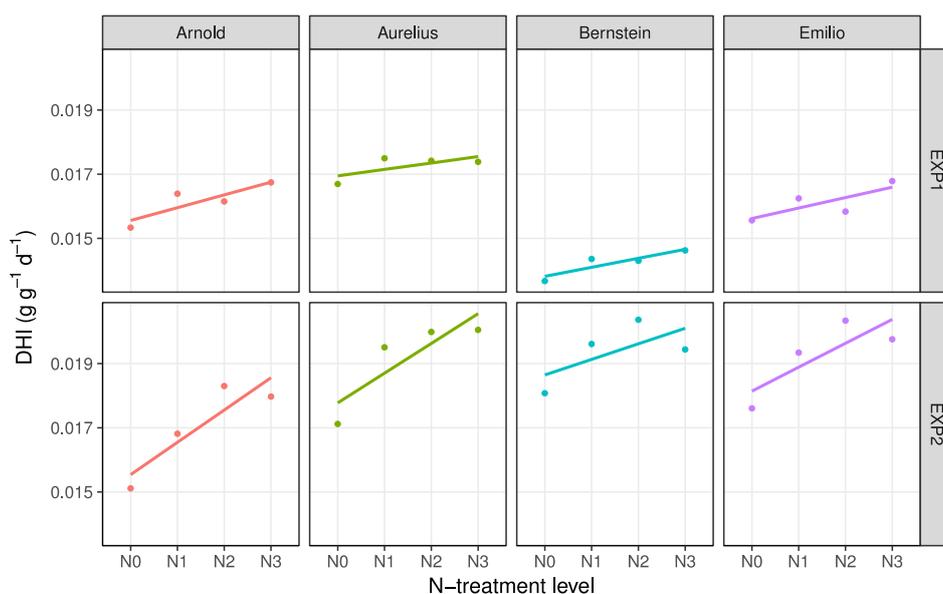


Figure 23 Observed daily harvest index increase (DHI) for the four winter wheat cultivars in both seasons (EXP1: 2017/18, EXP2: 2018/19). The N-treatments included no N (N0), 70 (N1), 140 (N2), and 210 (N3) kg N ha<sup>-1</sup>. Lines: liner regression.

### 8.3.11. Root Growth

The daily rate of root penetration (*GRTDP*) was not measured in the field. Soil water data was too coarse to derive differences in the cultivars' rooting depth rates (data not shown). Therefore, I used the default parameter value (30 mm bd<sup>-1</sup>).

### 8.3.12. Stem Nitrogen

The observed stem nitrogen concentrations (SNC) at anthesis were statistically unaffected by season ( $p = 0.24$ ) and cultivar ( $p = 0.97$ ) but highly dependent on the N-fertilisation treatment ( $p < 0.001$ ) (Figure 24). At harvest, season was significant ( $p < 0.001$ ) but cultivar again did not significantly influence SNC ( $p = 0.28$ ) while N-treatment did ( $p < 0.001$ ). Therefore, I pooled the cultivars and used identical parameters for the four cultivars for each of the parameters *SNCSI*, *SNCS2* (*SNCS*: "senesced", i.e. minimum), and *SNCG* (*SNCG*: "green", i.e. maximum).

The model parameters *SNCSI* and *SNCG* were calculated from measurements at anthesis (Figure 24). Calculated parameters based on both experiments were 0.63% for *SNCSI* (2.5% quantile) and 1.90% for *SNCG* (97.5% quantile). *SNCS2* was 0.22% calculated from data at harvest (2.5% quantile) (Figure 24).

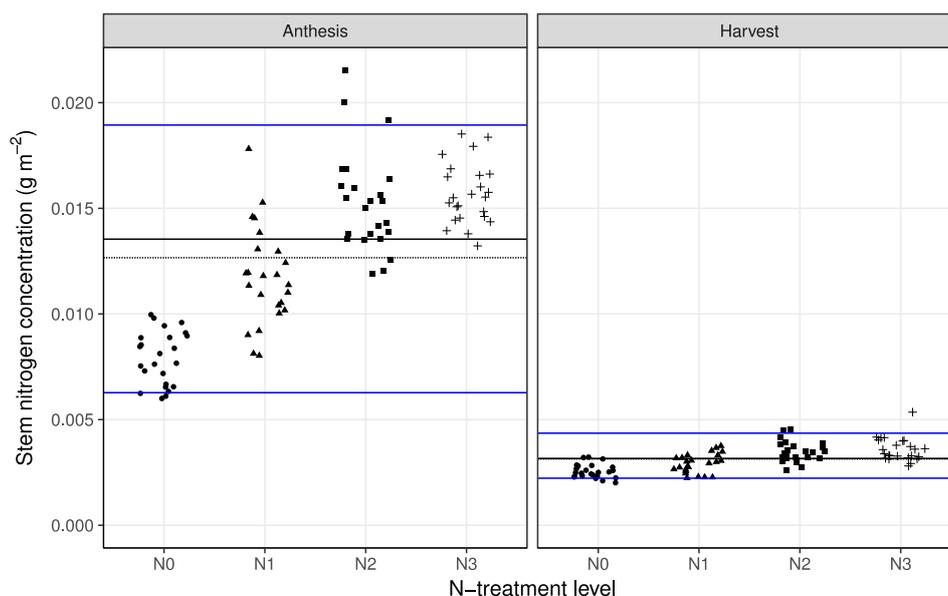


Figure 24 Stem nitrogen concentration at anthesis and harvest. The N-treatments included no N (N0), 70 (N1), 140 (N2), and 210 (N3) kg N ha<sup>-1</sup>. Dashed black line: mean. Full black line: median. Blue lines: 97.5% quantile (upper) and 2.5% quantile (lower).

### 8.3.13. Leaf Nitrogen

The specific leaf nitrogen of green leaves (*SLNG*) was calculated from data sampled at anthesis (Figure 25). There were differences between the seasons ( $p < 0.01$ ) and N-treatments ( $p < 0.001$ ) but not between the cultivars ( $p = 0.06$ ). The 97.5% quantiles of the cultivars were 2.20 (Arnold), 2.06 (Aurelius), 2.23 (Bernstein, outlier removed), and 2.16 g m<sup>-2</sup> (Emilio). For the *SLNG* parameter of all cultivars I used the 97.5% quantile of the pooled data (2.1 g m<sup>-2</sup>, rounded to one decimal) after removal of one extreme value (2.86, Bernstein, EXP1).

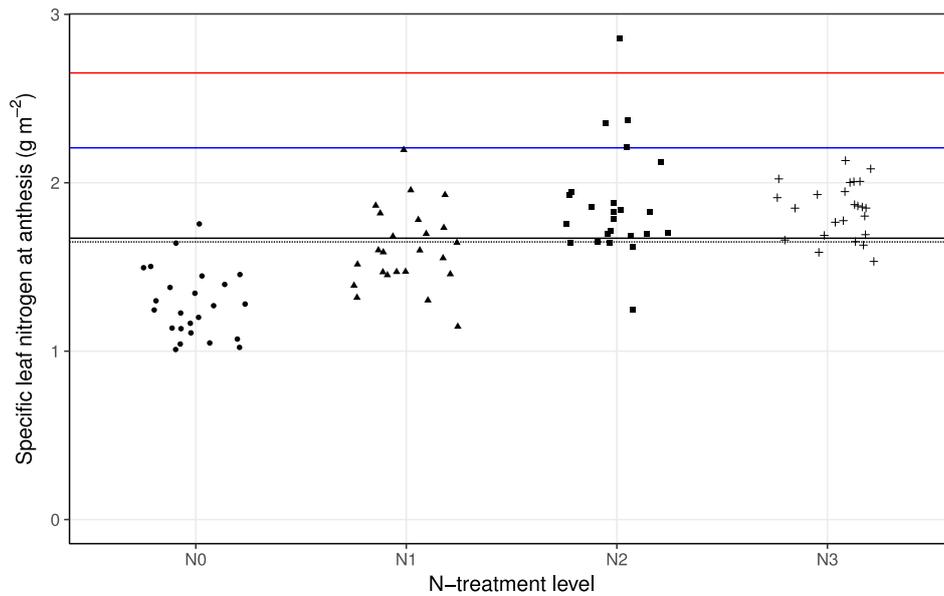


Figure 25 Specific leaf nitrogen (SLN) at anthesis. The N-treatments included no N (N0), 70 (N1), 140 (N2), and 210 (N3) kg N ha<sup>-1</sup>. Dashed black line: mean. Full black line: median. Red line: mean + three standard deviations. Blue line: 97.5% quantile after removal of the highest value (outlier).

SLNS (specific leaf nitrogen of senesced leaves) was calculated from N in senesced leaves at harvest and the highest observed LAI during the season (which was mostly close to or at anthesis). SLN at harvest ranged from roughly 0.1 to 0.4 g m<sup>-2</sup> (Figure 26). There was no effect of season ( $p = 0.09$ ), cultivar ( $p = 0.80$ ), and N-treatment ( $p = 0.36$ ). The 2.5% quantile was 0.10 g m<sup>-2</sup> which I used as the *SLNS* parameter in iCrop for all four cultivars.

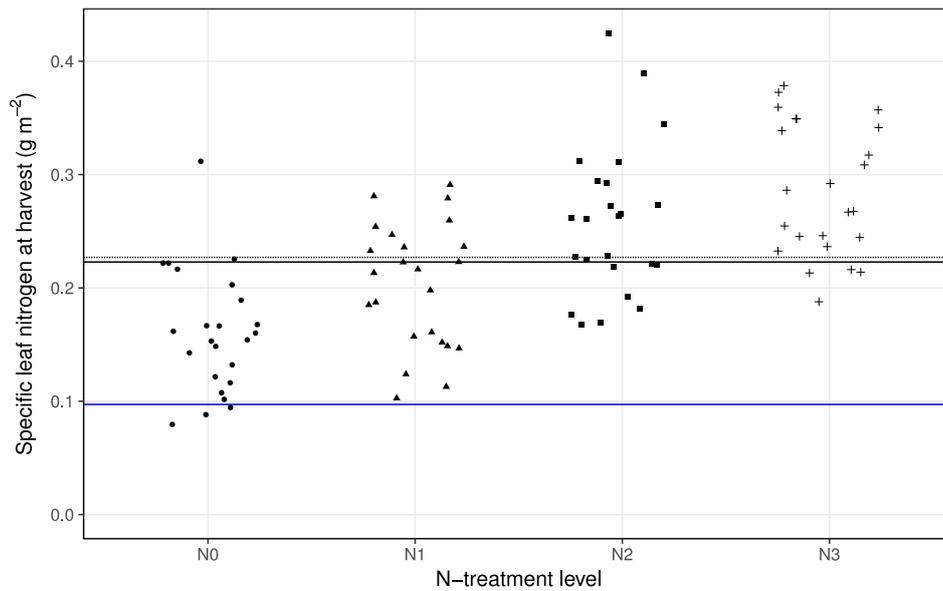


Figure 26 Specific leaf nitrogen of senesced leaves (SLNS,  $g N m^{-2}$ ) at harvest, assuming senesced leaf area at harvest equalled maximum leaf area of green leaves during the season. The N-treatments included no N (N0), 70 (N1), 140 (N2), and 210 (N3)  $kg N ha^{-1}$ . Dashed black line: mean. Full black line: median. Blue line: 2.5% quantile.

#### 8.3.14. Grain Nitrogen

The grain nitrogen concentration (%GN) at harvest did not show seasonal differences (p-value: 0.33). The effects of cultivar and N-treatment were both highly significant ( $p < 0.001$ ). However, the data was pooled and the 2.5% quantile ( $\%GN_{min} = 1.6\%$ , lower blue line in Figure 27) and 97.5% quantile ( $\%GN_{max} = 2.9\%$ , upper blue line) were used for the iCrop model parameters. Reasons for the decision to ignore the significant effect of cultivar and pool the data are explained in the discussion.

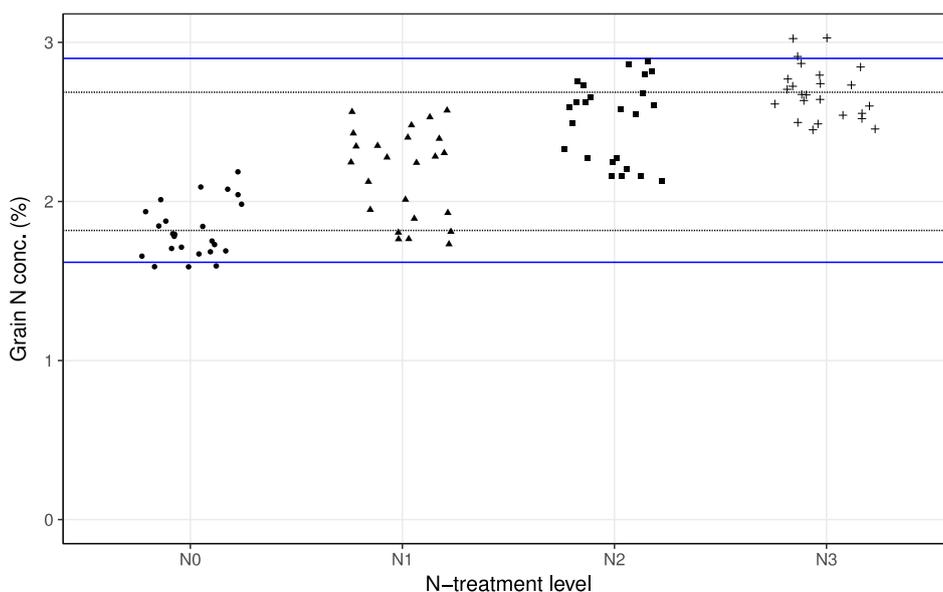


Figure 27 Grain nitrogen concentration at harvest of the different N-treatments. The N-treatments included no N (N0), 70 (N1), 140 (N2), and 210 (N3) kg N ha<sup>-1</sup>. Blue lines: 2.5% and 97.5% quantiles of all data. Dashed black lines: mean of N0 and N3.

### 8.3.15. Nitrogen Uptake Rate

The highest daily N-uptake rates were observed between the samplings on 196 and 224 DAS (EXP1) (ca. beginning of stem elongation and beginning of seed growth, resp.) and between 199 and 214 DAS (EXP2) (ca. beginning of stem elongation and shortly before anthesis, resp.) (Figure 28). The average daily uptake rates were highest in the N3 treatments (except Emilio, EXP2), ranging from 0.30 to 0.53 g N m<sup>-2</sup> d<sup>-1</sup> (Figure 29). I set the iCrop parameter *MXNUP* (maximum daily N-uptake) to 0.60 g N m<sup>-2</sup> d<sup>-1</sup> which is the same value as used in the APSIM wheat model (Holzworth et al., 2014).

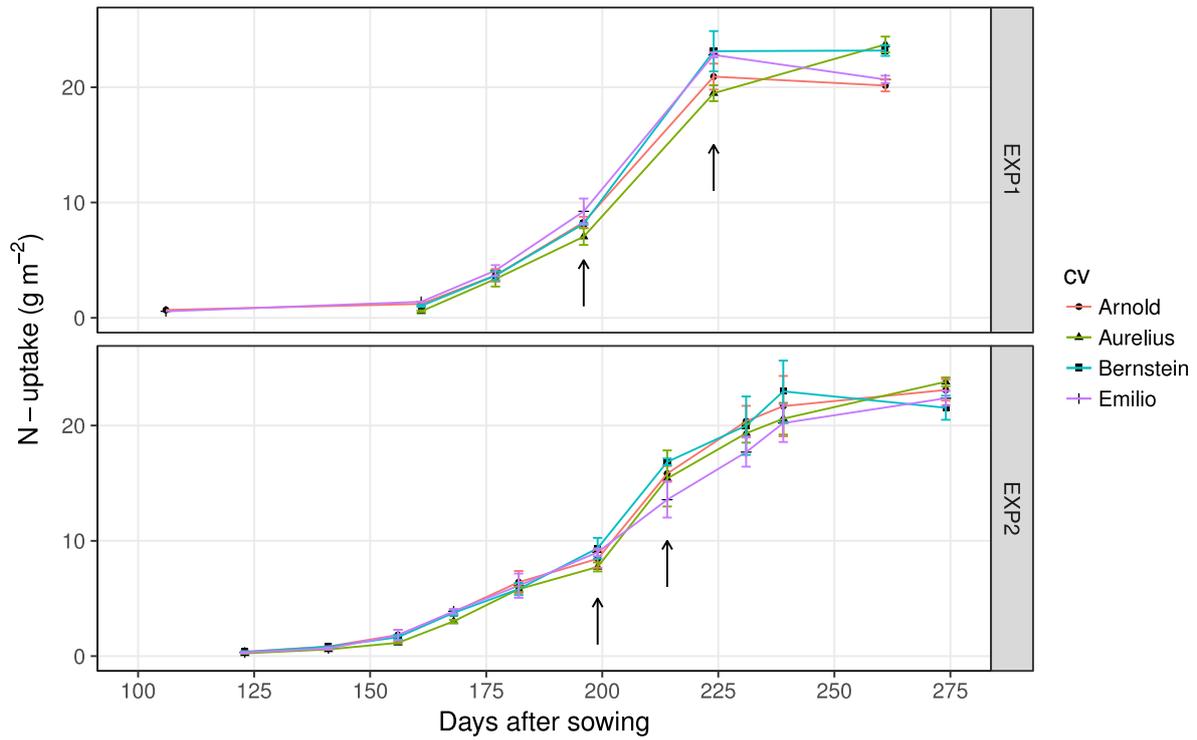


Figure 28 Total nitrogen uptake of the four winter wheat cultivars with the highest N-fertilisation (N3: 210 kg N ha<sup>-1</sup>) in both seasons (EXP1: 2017/18, EXP2: 2018/19). Error bars indicate standard error. Arrows highlight the interval with the highest N-uptake rate in each season.

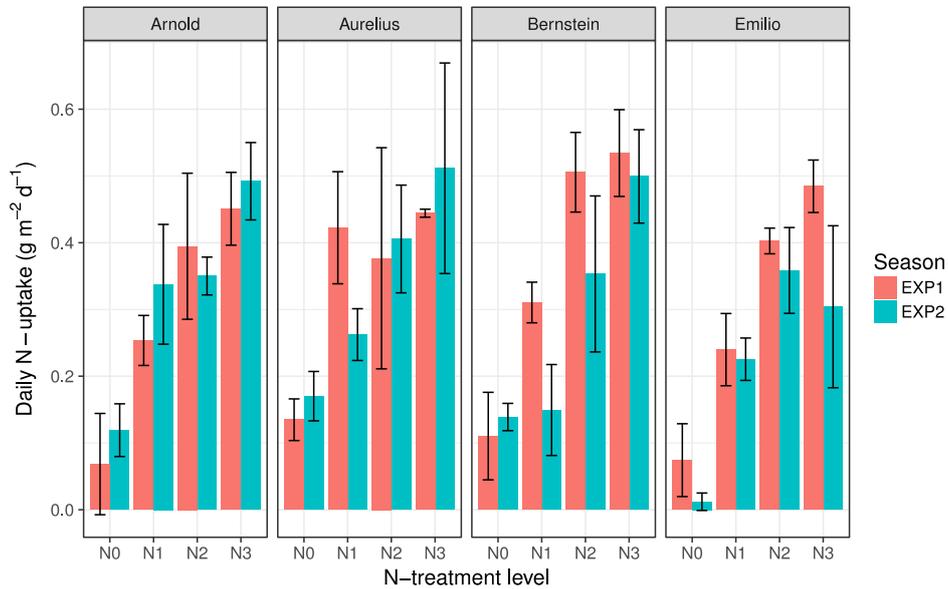


Figure 29 Daily nitrogen uptake of the four winter wheat cultivars from 196 to 224 days after sowing (DAS) in EXP1 and from 199 to 214 DAS in EXP2 (i.e. the periods between the arrows in Figure 28). Seasons: EXP1: 2017/18, EXP2: 2018/19. The N-treatments included no N (N0), 70 (N1), 140 (N2), and 210 (N3) kg N ha<sup>-1</sup>. Error bars indicate standard error.

### 8.3.16. Soil Water

The model was set to simulate runoff.

For calculating soil water content and for the model's *BDL* parameter, measured bulk densities were used. While for EXP1 the exact measured values were used as parameter, bulk density was reduced to  $1.30 \text{ g m}^{-2}$  for the deeper soil layers (60-90 and 90-120 cm) in EXP2 to represent visually observed sandy layers (Figure 30) as well as observed differences in gravimetric soil water content.



Figure 30 Augers filled with soil shortly after being extracted from the soil of the second field experiment (2018/19). The numbers indicate soil depth in cm. The start of the sandy subsoil layer is clearly visible at ca. 70 cm. Some of the soil sample was lost from the top (dry topsoil; re-added manually to the sample) and some from the bottom (gravelly).

Based on previous simulation studies comparing iCrop and APSIM soil water contents (data not shown), the necessary rainfall amount to return from stage II evaporation to stage I was set to 6 mm (default: 10 mm, parameter *WETWAT*). This results in more frequent stage I evaporation and, therefore, generally higher evaporation.

The parameters defining soil water limits (SAT: saturated soil water content, DUL: drained upper limit (i.e. field capacity), LL: lower limit (i.e. wilting point), and ADRY: air-dry lower limit) were calculated for each soil layer from measured soil water data of the two winter wheat growing seasons. Since the two experiments were carried out on different fields with apparent soil differences, I also used different parameter sets for the soils in each season. The observed soil water ranges (Figure 31) and total organic carbon contents (Table 7) of the three topmost layers (0-10 cm, 10-30 cm, and 30-60 cm) were similar in both seasons. Therefore, identical model parameters were defined for these layers in both seasons. With greater soil depth, deviations between observed soil water contents in the two seasons started to increase, with EXP2 showing the lower values. Accordingly, model parameters were calibrated to reflect these differences.

Initial soil water contents were calculated from measurements at simulation start: 33 and 14 days before sowing in EXP1 and EXP2, respectively. Initials in EXP1 amounted to 539 mm and in EXP2 to 362 mm soil water.

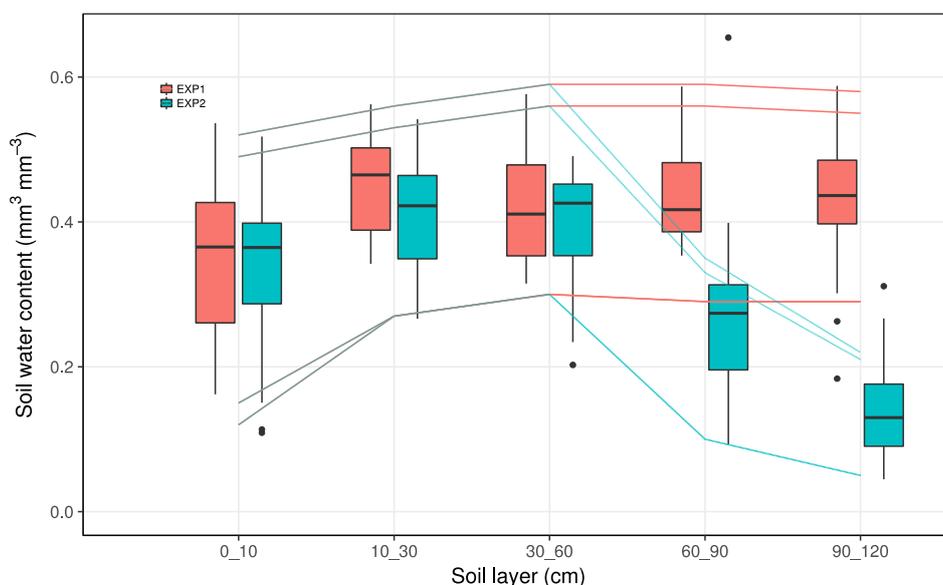


Figure 31 Soil water content (volumetric) measured in the winter wheat seasons 2017/18 (EXP1) and 2018/19 (EXP2) and derived soil water parameters (lines) for saturation (SAT), drained upper limit (DUL), lower limit (LL), and air-dry lower limit (ADRY); lines from top to bottom, respectively.

### 8.3.17. Soil Nitrogen

Soil organic carbon content (TOC) ranged from 0.27% to 2.24% with decreasing values from shallow to deeper soil layers (Table 7). Between the two seasons, TOC was similar in the top soil layers (0-30 and 30-60 cm) while EXP2 showed lower TOC in deeper layers compared to EXP1. Calculated parameter values for soil organic nitrogen content (NORG) reflect these differences directly.

Table 7 Soil total organic carbon content (TOC) measured on three dates and soil C/N ratio used for the calculation of the soil organic N content (NORG) parameter for both seasons (EXP1: 2017/18, EXP2: 2018/19).

Soil layer (cm)	TOC (%) measured			C/N	NORG (%) calculated	
	EXP1	EXP2			EXP1	EXP2
	13/09/2017	1/10/2018	13/05/2019			
0-30	2.33 ± 0.03	2.24 ± 0.04	2.19 ± 0.02	10:1	0.22	0.22
30-60	1.61 ± 0.08	1.62 ± 0.03	1.57 ± 0.04	10:1	0.16	0.16
60-90	1.38 ± 0.11	0.83 ± 0.12	0.89 ± 0.05	10:1	0.14	0.09
90-120	1.09 ± 0.13	0.27 ± 0.09	0.44 ± 0.02	10:1	0.11	0.03

The fraction of NORG available for daily mineralization (FMIN) was calibrated by trial-and-error and comparison between simulated final soil Nmin versus observed. Identical parameter values were used for both seasons, decreasing from 0.050 in the topmost layer to 0.001 in the deepest layer.

Initial soil mineral nitrogen (N<sub>min</sub>) was measured at the same dates as initial soil water (see above). Total initial N<sub>min</sub> contents across the whole soil profile were 63 kg N ha<sup>-1</sup> in EXP1 and 30 kg N ha<sup>-1</sup> in EXP2.

#### 8.4. iCrop Simulation Results

In this section, the results of simulating the parameterisation field experiments (EXP1: 2017/18, EXP2: 2018/19; Tulln, Eastern Austria) using the previously derived parameter set (8.3) are described. Due to their contrasting phenological developments I picked the cultivars Arnold and Bernstein for most comparisons. Unless stated otherwise, the other cultivars (Aurelius, Emilio) did not show any specific behaviour departing largely from the two described cultivars. Descriptive statistics (RMSE, PE) are based on all treatments (all cultivars) unless a cultivar is explicitly referred to.

##### 8.4.1. Phenology

Simulated phenology hardly responded to the N-treatment, showing differences of maximum 2 days only shortly before maturity (data not shown). The simulations showed faster average phenological development of Arnold compared to Bernstein, mainly from BBCH 30 (stem elongation) until ca. BBCH 80 (ripening) (Figure 32). This agreed well to the observed cultivar effects, especially in EXP1.

Overall, the agreement between simulated and observed phenology was good (Table 8). Emergence was simulated very close to the observations. Only the beginning of tillering (BBCH 21) was matched poorly in EXP1, with varying errors for the cultivars. All other stages were simulated within 9 days. In EXP1 the simulations were very close to the observations, showing a trend of being slightly too early. Anthesis in EXP1 was simulated with +/- 2 days, and maturity was within 5 days. EXP2 showed higher absolute errors than EXP1 and was mostly delayed compared to observations. Anthesis (BBCH 61 - 69) was simulated 1 to 5 days later than observed, and maturity with plus 7 to 8 days.

Table 8 Errors for simulating key phenological stages of four winter wheat cultivars in two seasons (EXP1: 2017/18, EXP2: 2018/19). Negative values: simulations earlier than observed. Phenological stages (BBCH): 9 emergence, 21 first tiller, 31 first node (stem elongation), 41 flag leaf sheath extending (booting), 51 tip of ear just visible (ear emergence), 59 ear emergence complete, 61 beginning of anthesis, 65 anthesis half-way, 69 anthesis complete, 71 kernel water ripe (milk development), 81 early dough development, 87 hard dough stage (maturity). The N-fertilisation treatments were pooled.

Season	Cultivar	BBCH:	Error (days)											
			9	21	31	41	51	59	61	65	69	71	81	87
EXP1	Arnold		1	20	3	-3	-2	-1	0	2	0	-3	2	0
	Aurelius		1	-36	-2	-2	-2	-3	-2	-1	-1	-3	1	-1
	Bernstein		1	-31	-1	-3	-4	-2	-1	-1	1	1	3	-5
	Emilio		1	21	1	-2	-2	0	1	2	2	0	5	1
EXP2	Arnold		1	9	9	7	5	3	4	4	4	3	4	7
	Aurelius		1	-7	9	7	6	4	5	1	3	2	4	7
	Bernstein		1	9	8	8	3	3	4	4	5	5	6	7
	Emilio		1	8	8	5	3	3	4	3	4	4	6	8

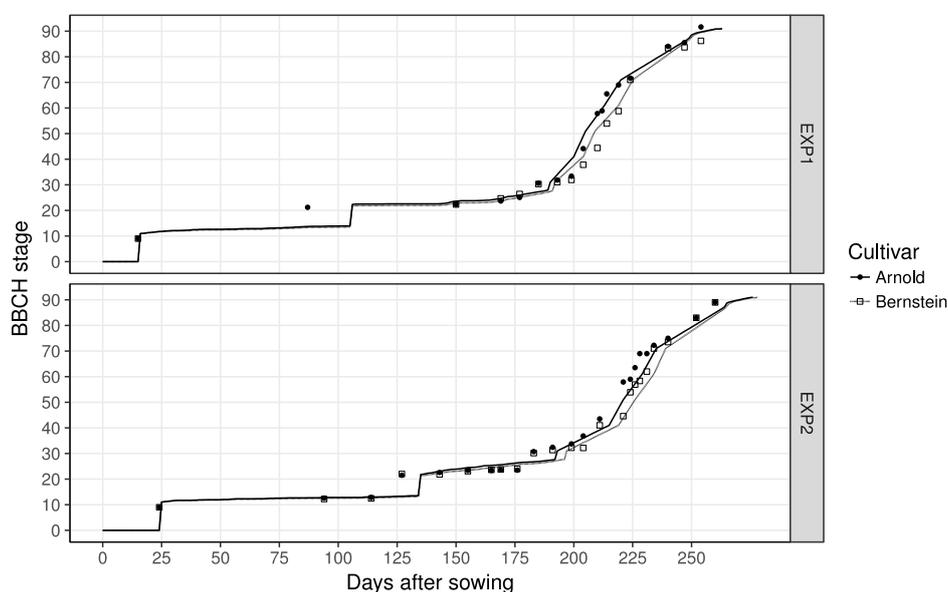


Figure 32 Phenological development (BBCH stages): Observed (symbols) vs. simulated (lines) values of two contrasting winter wheat cultivars (Arnold and Bernstein) in both seasons (EXP1: 2017/18, EXP2: 2018/19). Observed values were averaged across N-treatments.

#### 8.4.2. Leaf Canopy Development

Simulated node number on the main stem (MSNN) was unaffected by N-treatment (data not shown). Accordingly, observed final MSNN was not significantly affected by the N-treatment ( $p = 0.06$ , see also sub-section 8.1.5).

iCrop simulated lower MSNN for Bernstein than for Arnold during the development phase (Figure 33). However, the duration of the MSNN production was simulated longer for Bernstein, resulting in similar final MSNN for both cultivars.

The agreement between simulation and observation of Arnold's and Bernstein's MSNN was especially good in EXP1. In EXP2 iCrop underestimated MSNN in the final phase of the node development phase (ca. 175 until 220 DAS) mainly for Arnold. Still, in both seasons for all four cultivars, estimates of final MSNN were within 1 leaf of observed values and highly robust (RMSE = 0.5, RRMSE = 6%, PE = +3%).

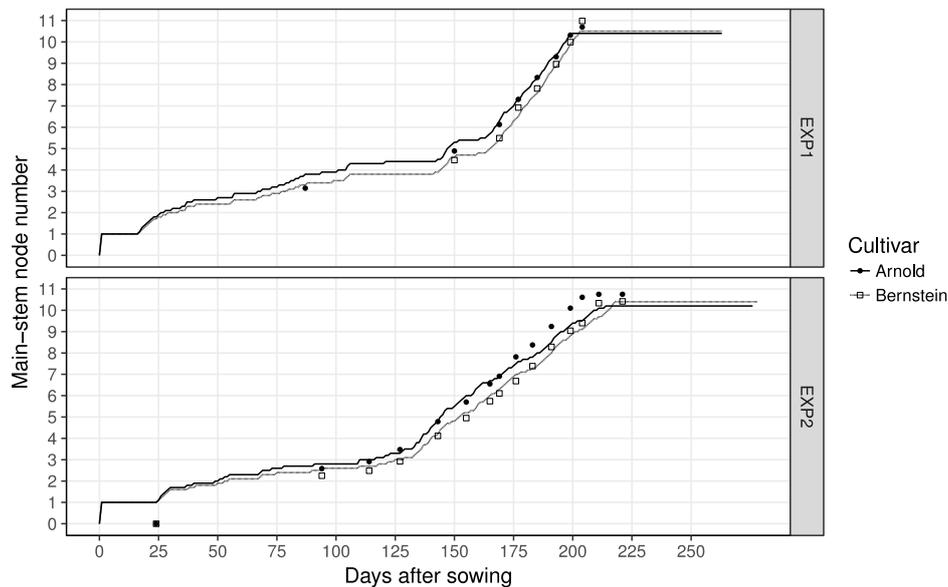


Figure 33 Main stem node number: Observed (symbols) vs. simulated (lines) values of two contrasting winter wheat cultivars (Arnold and Bernstein) in both seasons (EXP1: 2017/18, EXP2: 2018/19). Observed values were averaged across N-treatments

In both seasons, simulated LAI showed response to N-treatment mainly between no fertilisation (N0) and the other N-treatments (N1, N2, N3) (Figure 34). The LAI-reducing effect of N0 was stronger in EXP2 (maximum LAI ca. 2) than in EXP1 (max. LAI ca. 3). The simulations showed hardly any difference between N1, N2, and N3 in both seasons.

Regarding simulated cultivar differences in LAI, Bernstein had consistently higher LAI during most of the growing season across all treatments. Differences between the cultivars were lowest in the N0 treatment. In the other treatments (N1, N2, N3), maximum LAI of Bernstein was higher than Arnold by ca.  $1 \text{ mm}^2 \text{ mm}^{-2}$ .

The model predicted the dynamics of LAI with varying accuracy. The relative trends of the influence of N-treatment and cultivar were captured correctly. At anthesis, LAI simulations were not robust (RMSE = 0.8, NRSME = 32%, PE = +36%). Also, iCrop showed a tendency to overestimate early LAI (up until ca. 150 DAS, i.e. mid-March) as well as LAI at harvest (observed: all zero) for high N-treatments. Maximum LAI of Bernstein (RMSE = 0.5, RRMSE = 17%, PE = +17%) was simulated very close to the observations. In contrast, although Arnold's maximum LAI was simulated lower than Bernstein, it was still overestimated across most treatments (RMSE = 0.9,

RRMSE = 43%, PE = +52%). Aurelius (RMSE = 0.5, RRMSE = 20%, PE = +16%) was similarly well predicted as Bernstein, and Emilio (RMSE = 1.1, RRMSE = 53%, PE = +58%) similarly poor as Arnold.

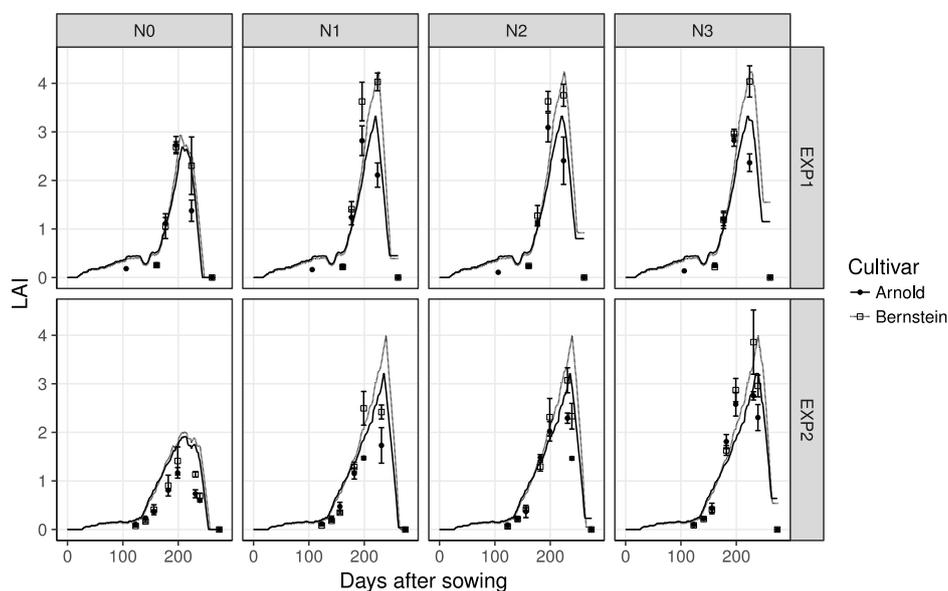


Figure 34 Leaf area index (LAI): Observed (symbols) vs. simulated (lines) values of two contrasting winter wheat cultivars (Arnold and Bernstein) in both seasons (EXP1: 2017/18, EXP2: 2018/19). The N-treatments included no N (N0), 70 (N1), 140 (N2), and 210 (N3) kg N ha<sup>-1</sup>. Bars represent standard errors of the means.

#### 8.4.3. Dry Mass Accumulation and Partitioning

Simulated total above-ground dry mass responded positively to increased N-fertilisation, with the strongest response between N0 and N1 (Figure 35). iCrop consistently simulated higher TDM for Bernstein than for Arnold, with the difference emerging mainly in the last days of TDM growth (shortly before harvest).

While the simulated trends regarding N-treatment and cultivar differences were in agreement with the observations, the predictions of the absolute values ranged from poor to good. Predictions at harvest (all cultivars) were generally better for EXP1 (RMSE = 117, RRMSE = 8%, PE = +/-0%) than for EXP2 (RMSE = 321 g m<sup>-2</sup>, RRMSE = 26%, PE = +28%). iCrop showed a tendency to overestimate TDM for low N-fertilisation treatments while it predicted high N-treatments much better (Figure 35). More generally viewed, low TDM at anthesis and harvest were overestimated, while high TDM at anthesis and harvest were underestimated or estimated well, respectively (Figure 36). While iCrop was able to capture the direction of cultivar differences (Bernstein higher than Arnold), it underestimated these differences at harvest as well as in-season. In almost all treatments, iCrop overestimated in-season TDM between ca. 150 and 200 DAS (ca. the whole stem elongation phase, BBCH 30 to 39). Overall, at anthesis (ca. 225 DAS) iCrop estimates matched observations closely (RMSE = 114 g m<sup>-2</sup>, RRMSE = 12%, PE = +6%).

Also, final TDM estimates matched the observations quite well (overall RMSE of  $241 \text{ g m}^{-2}$ , RRMSE = 18%, PE = +13.7%).

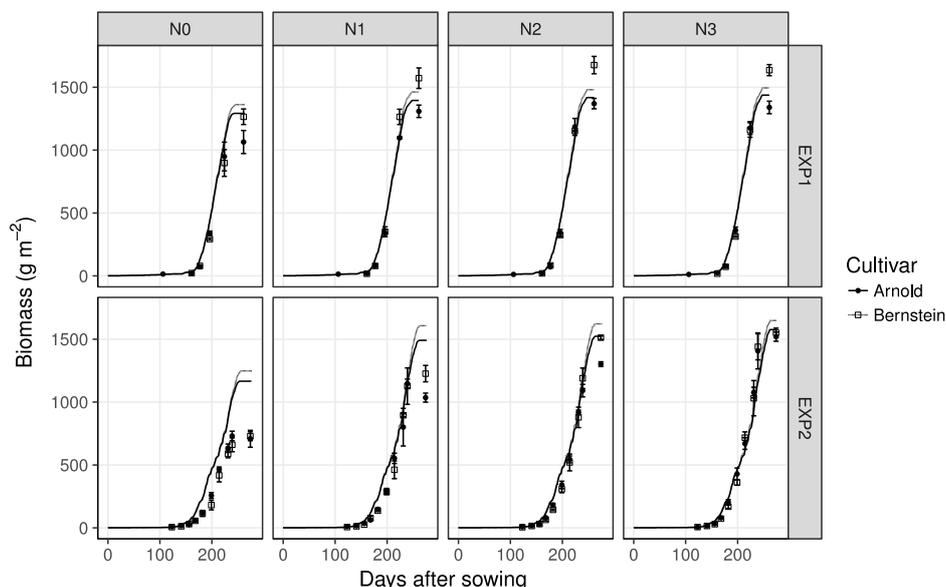


Figure 35 Total above-ground biomass (dry mass): Observed (symbols) vs. simulated (lines) values of two contrasting winter wheat cultivars (Arnold and Bernstein) in both seasons (EXP1: 2017/18, EXP2: 2018/19). The N-treatments included no N (N0), 70 (N1), 140 (N2), and 210 (N3) kg N ha<sup>-1</sup>. Bars represent standard errors of the means.

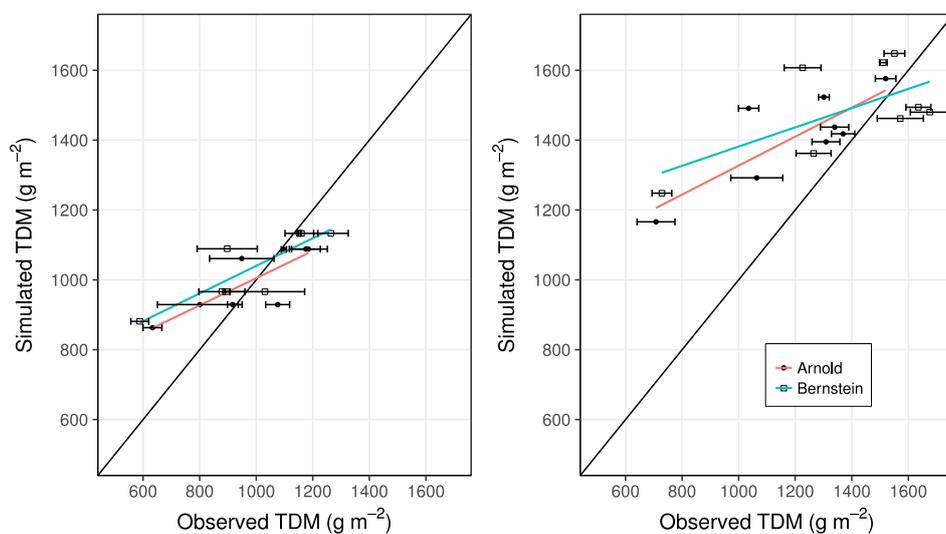


Figure 36 Total above-ground dry mass (TDM) of two contrasting winter wheat cultivars (Arnold and Bernstein) in both seasons (EXP1: 2017/18, EXP2: 2018/19) at anthesis (left) and harvest (right). Coloured lines: linear regression, black line: 1:1. Bars represent standard errors of the means.

Simulated yield showed a stronger positive response to N-fertilisation in EXP2 than in EXP1 (Figure 37). In both seasons, this response was strongest between N0 and N1. Simulated differences between the cultivars were small, with highest N3 yields simulated for Arnold in EXP1 ( $628 \text{ g m}^{-2}$ ) followed by Aurelius (610), Emilio (598), and Bernstein (578), and highest N3 yield for Aurelius in EXP2 ( $748 \text{ g m}^{-2}$ ) followed by Arnold (732), Emilio (726), and Bernstein (700).

Apart from the positive direction of the estimated yield response to N-fertilisation, iCrop had problems capturing seasonal and cultivar differences. In both seasons, the model underestimated the impact of N-fertilisation on yield. For three cultivars (Aurelius, Bernstein, Emilio) of EXP1 the estimate for N0 was good while the higher N-treatments were clearly underestimated. The high N-treatments of the same cultivars were predicted well in EXP2, but N0 and N1 were overestimated. Still, overall final yield simulation was robust (RMSE = 95 g m<sup>-2</sup>, RRMSE = 16%, PE = +6%).

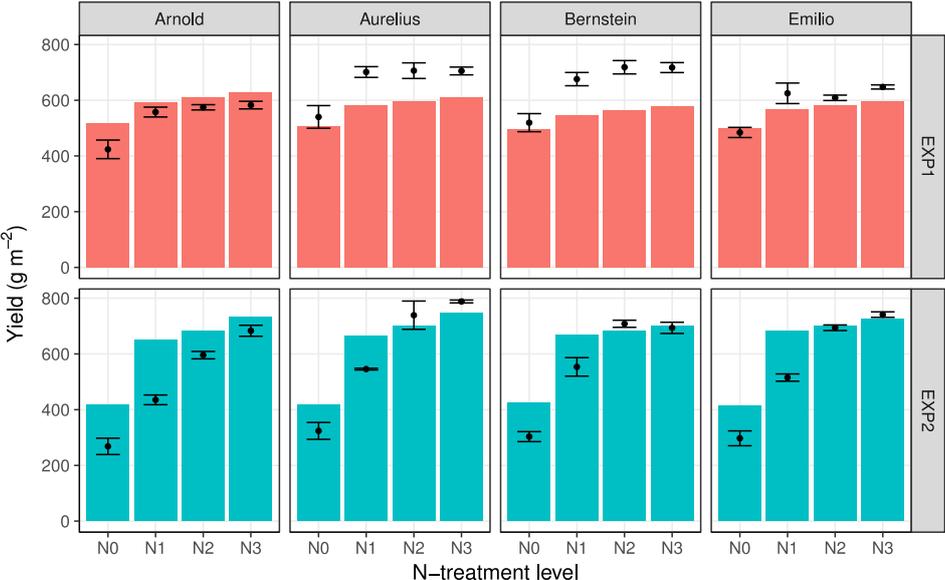


Figure 37 Grain yield: Observed (points) vs. simulated (bars) values of four winter wheat cultivars (Arnold, Aurelius, Bernstein, and Emilio) in both seasons (EXP1: 2017/18, EXP2: 2018/19). The N-treatments included no N (N0), 70 (N1), 140 (N2), and 210 (N3) kg N ha<sup>-1</sup>. Bars represent standard errors of the means.

Simulated leaf dry mass showed hardly any response to N-fertilisation (Figure 38). The model simulated differences between the cultivars, starting ca. 200 DAS, with higher leaf dry mass for Bernstein than for Arnold. At anthesis (ca. 225 DAS) leaf dry mass was largely overestimated (RMSE = 121 g m<sup>-2</sup>, RRMSE = 73%, PE = +78%). The model also overestimated leaf dry mass towards the end of the growing season. However, it needs to be considered that despite iCrop does calculate dry mass retranslocation from vegetative organs (leaves and stems) to the grains, it does not account for retranslocated dry mass separately in the output of leaf and stem dry mass. Therefore, comparisons between simulations and observations for leaf and stem dry mass towards the end of the growing season are not possible.

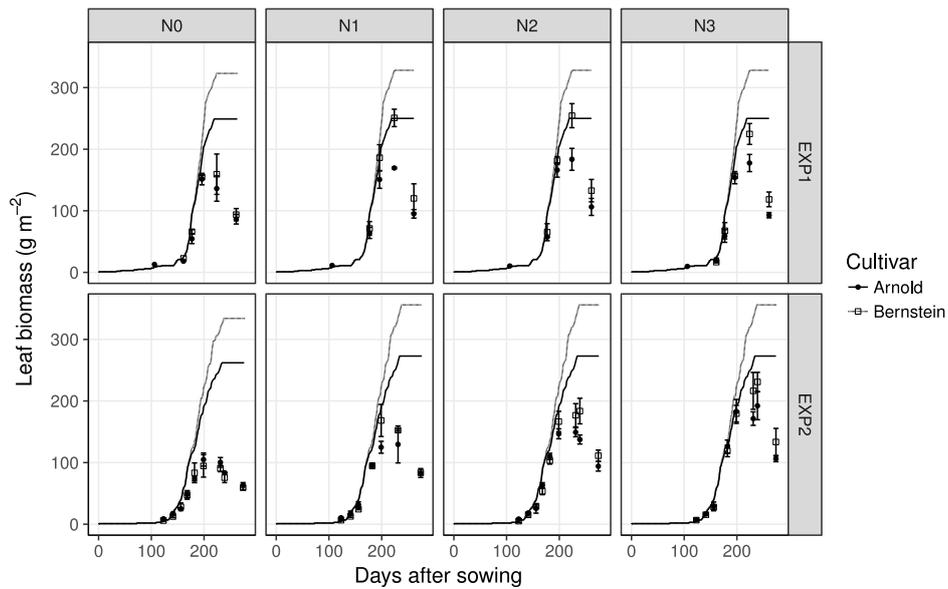


Figure 38 Leaf biomass (dry mass): Observed (symbols) vs. simulated (lines) values of two contrasting winter wheat cultivars (Arnold and Bernstein) in both seasons (EXP1: 2017/18, EXP2: 2018/19). The N-treatments included no N (N0), 70 (N1), 140 (N2), and 210 (N3) kg N ha<sup>-1</sup>. Bars represent standard errors of the means. Note that biomass retranslocation from leaf to grain is calculated in the model, but not shown in the output (i.e. the “plateau” at the end of the season does not represent model simulations correctly).

Simulated stem dry mass showed a weak response to N-fertilisation (Figure 39). The model estimates showed differences between the cultivars only late in the season for final stem weight, with Bernstein showing higher dry mass than Arnold. While the relative direction of this difference matched observations, the second last measurements (anthesis) were predicted with varying precision (RMSE = 131 g m<sup>-2</sup>, RRMSE = 16%, PE = 11%). Observed response to N-treatment was stronger than predicted by the model, thereby resulting in either a good prediction of N0 and poor prediction of the high N-treatments (N1, N2, N3) or the other way around.

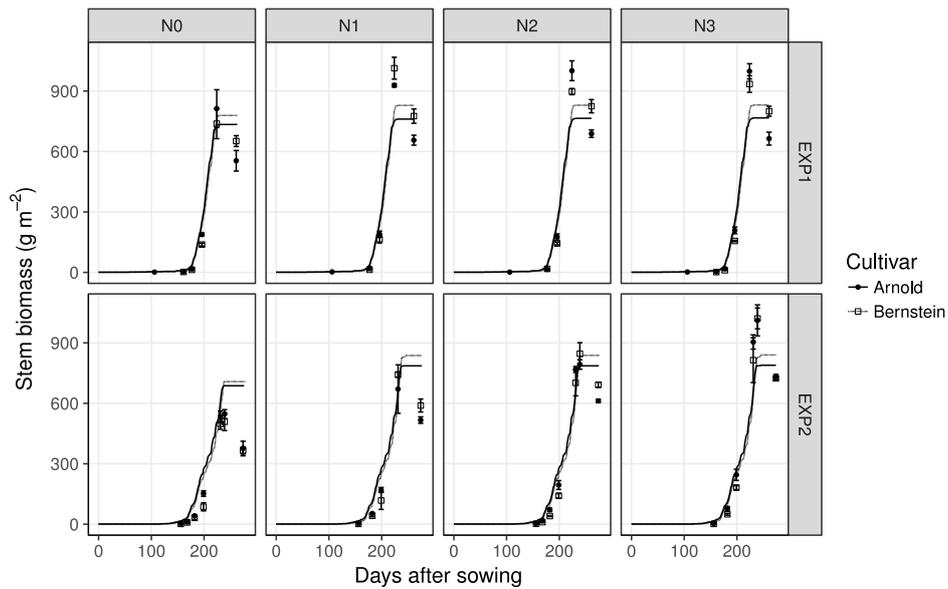


Figure 39 Stem biomass (dry mass): Observed (symbols) vs. simulated (lines) values of two contrasting winter wheat cultivars (Arnold and Bernstein) in both seasons (EXP1: 2017/18, EXP2: 2018/19). The N-treatments included no N (N0), 70 (N1), 140 (N2), and 210 (N3) kg N ha<sup>-1</sup>. Bars represent standard errors of the means. Note that biomass retranslocation from stem to grain is calculated in the model, but not shown in the output (i.e. the “plateau” at the end of the season does not represent model simulations correctly).

Both leaf and stem dry mass (i.e. total dry mass before the beginning of seed growth) of all cultivars were mostly overestimated before anthesis (exemplarily shown for Bernstein in Figure 40). At anthesis and thereafter, N0 was still overestimated, while in N3 stem dry mass was underestimated.

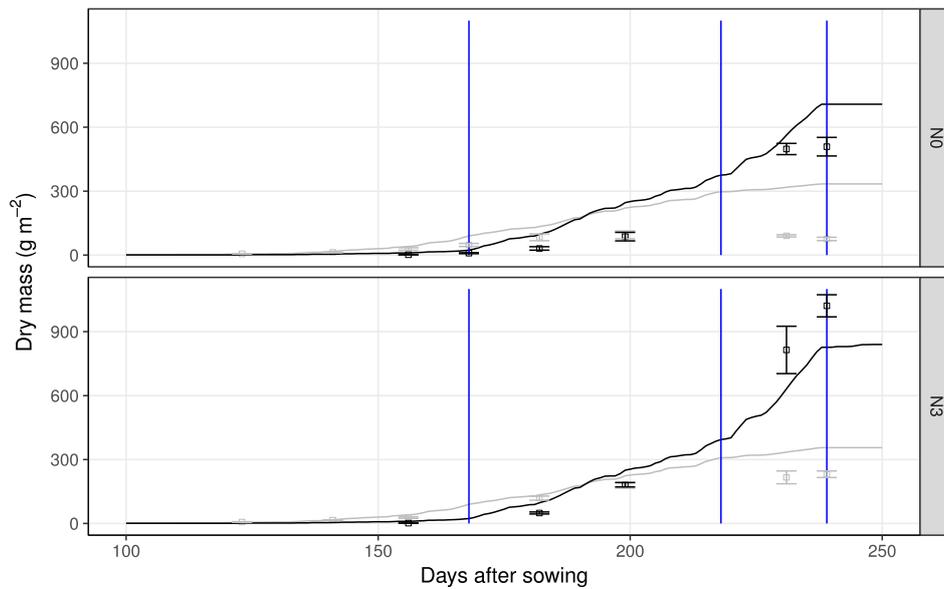


Figure 40 Simulated stem (black lines) and leaf (grey lines) dry mass vs. observations (symbols) for cultivar Bernstein without (N0) and with maximum N-fertilisation (210 kg N ha<sup>-1</sup>) (zoomed to the period 100 to 250 days after sowing). The vertical blue lines indicate the simulated switches between the leaf:stem partitioning parameters (left to right): FLFIA, FLFIB, FLF2, and zero thereafter. Bars represent standard errors of the means.

#### 8.4.4. Crop Nitrogen Uptake and Partitioning

Simulated total shoot N-uptake (TNU) at harvest showed a clear and rather linear response to N-fertilisation (Figure 41). Cultivar differences were marginal. TNU was higher in EXP1 than in EXP2. On average, simulations of TNU at harvest were robust (RMSE = 3.2 g m<sup>-2</sup>, RRMSE = 19%, PE = +14.3%). However, in EXP1 the impact of N-fertilisation was overestimated, expressed by good estimates for N0 across all cultivars but overestimation of N3. In EXP2 all treatments were overestimated.

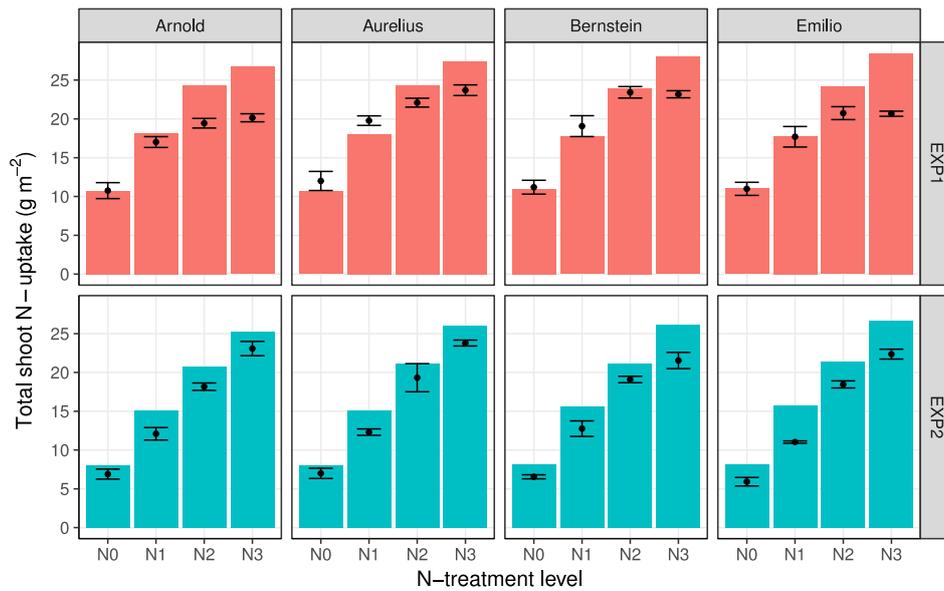


Figure 41 N-uptake at harvest: Observed (points) vs. simulated (bars) values of four winter wheat cultivars (Arnold, Aurelius, Bernstein, and Emilio) in both seasons (EXP1: 2017/18, EXP2: 2018/19). The N-treatments included no N (N0), 70 (N1), 140 (N2), and 210 (N3) kg N ha<sup>-1</sup>. Bars represent standard errors of the means.

Simulated TNU dynamics showed a positive response to N-fertilisation starting ca. 200 DAS (Figure 42). iCrop simulated only small differences between Arnold (lower) and Bernstein (higher), most of those emerging only under higher N-fertilisation rates and at the end of the season. In EXP1 iCrop simulated TNU until 200 DAS very well compared to observations. Thereafter, observed cultivar differences were not matched and N3 was overestimated. In EXP2 estimates showed a similar behaviour. Also, there was a trend towards overestimation. On average across all treatments, TNU at anthesis was simulated well and robust (RMSE = 2.6 g m<sup>-2</sup>, RRMSE = 17%, PE = +14.0%), similar to TNU at harvest (see above).

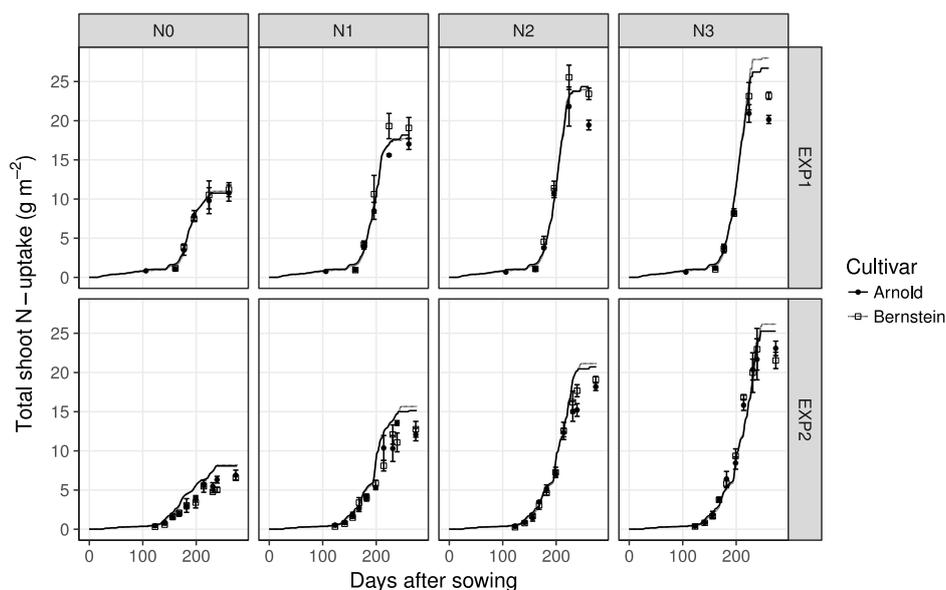


Figure 42 NUP: Observed (symbols) vs. simulated (lines) values of two contrasting winter wheat cultivars (Arnold and Bernstein) in both seasons (EXP1: 2017/18, EXP2: 2018/19). The N-treatments included no N (N0), 70 (N1), 140 (N2), and 210 (N3) kg N ha<sup>-1</sup>. Bars represent standard errors of the means.

The model estimates for grain N-concentration (%GN) showed a positive response to increased N-fertilisation, although the step from N2 to N3 resulted in almost no effect in EXP1 (Figure 43). Simulated cultivar differences were small, showing lowest %GN for Arnold in EXP1 and highest for Bernstein in EXP2. On average, the model estimates for %GN were robust (RMSE = 0.22%, RRMSE = 9%, PE = -1.3%). However, %GN of Arnold was always underestimated (RMSE = 0.32%, RRMSE = 13%, PE = -12.8%) while Aurelius (RMSE = 0.13%, RRMSE = 6%, PE = -0.7%), Bernstein (RMSE = 0.19%, RRMSE = 9%, PE = 5.0%), and Emilio (RMSE = 0.16%, RRMSE = 8%, PE = 3.4%) were simulated better. For all cultivars, there was a trend of N0 underestimation, for Bernstein and Emilio also N3 overestimation, indicating an overestimation of the N-treatment impact.

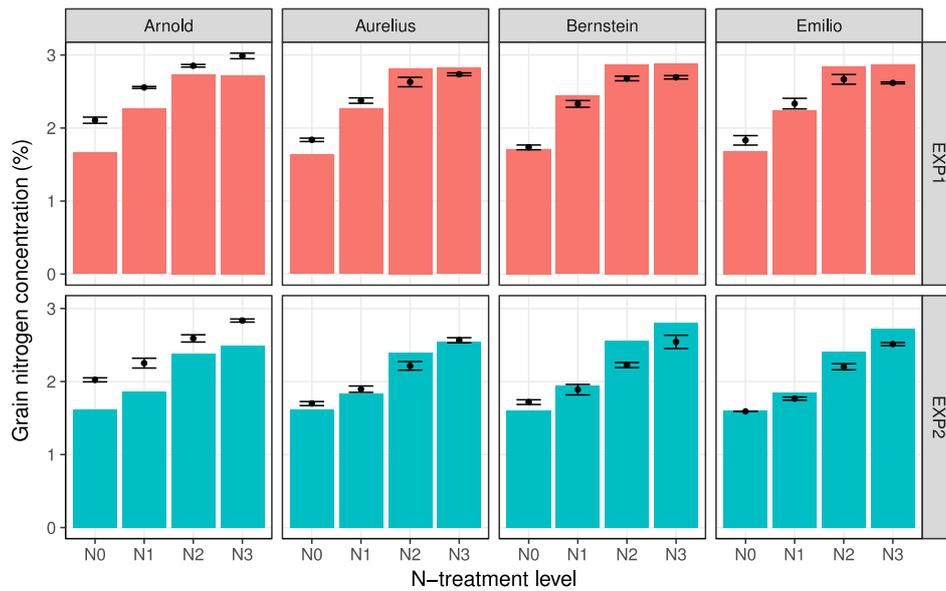


Figure 43 Grain N- concentration: Observed (points) vs. simulated (bars) values of four winter wheat cultivars (Arnold, Aurelius, Bernstein, and Emilio) in both seasons (EXPI: 2017/18, EXP2: 2018/19). The N-treatments included no N (N0), 70 (N1), 140 (N2), and 210 (N3) kg N ha<sup>-1</sup>. Bars represent standard errors of the means.

Simulated leaf N content responded to N-fertilisation only between N0 and N1 (Figure 44). The simulated effect of cultivar (considering only Arnold and Bernstein) started emerging ca. 200 DAS, with Bernstein showing higher values. These cultivar differences reached their maximum ca. at anthesis, where absolute values were at their maximum as well (for N1, N2, N3) or shortly after the maximum (for N0). Afterwards, in most cases, cultivar differences gradually declined towards zero at harvest.

In EXPI, simulations matched observations well until ca. anthesis (RMSE at anthesis: Bernstein 1.3, Arnold 2.0 g N m<sup>-2</sup>). At harvest, all but the N0 treatments were overestimated. For EXP2, estimates of harvest leaf N content were better, but most of the season for N0 was overestimated. iCrop captured the general direction of differences between the cultivars in many treatments. However, in both seasons, Arnold's leaf N content at anthesis was clearly overestimated. Overall, leaf N content was overestimated at both anthesis (RMSE = 2.3 g N m<sup>-2</sup>, RRMSE = 50%, PE = +62.7%) and harvest (RMSE = 1.1 g N m<sup>-2</sup>, RRMSE = 147%, PE = +63.9%).

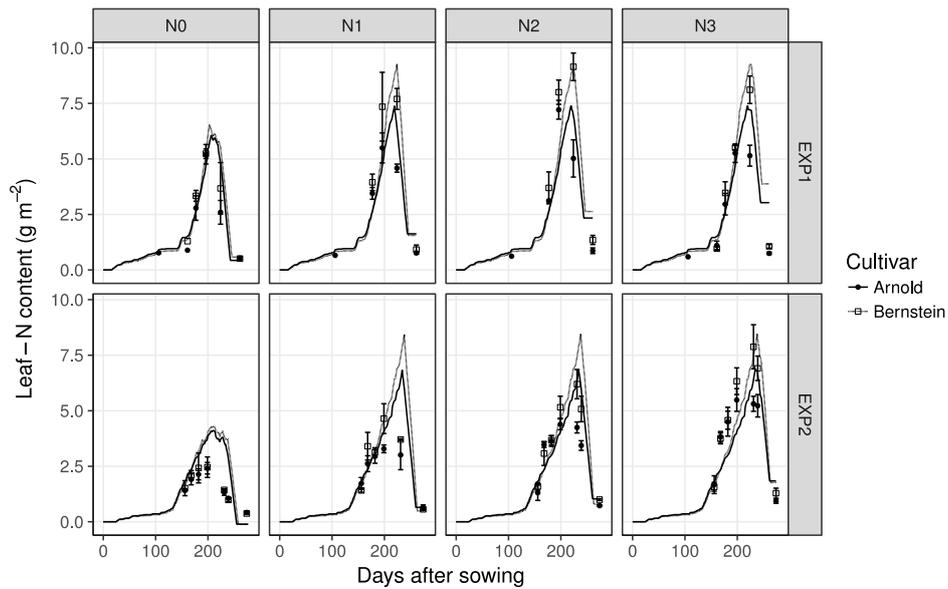


Figure 44 Leaf-N content: Observed (symbols) vs. simulated (lines) values of two contrasting winter wheat cultivars (Arnold and Bernstein) in both seasons (EXP1: 2017/18, EXP2: 2018/19). The N-treatments included no N (N0), 70 (N1), 140 (N2), and 210 (N3) kg N ha<sup>-1</sup>. Bars represent standard errors of the means.

Stem-N content simulated by iCrop showed clear responses from N0 to N2 and a reduced response for N3 (Figure 45). Simulated differences between Arnold and Bernstein were small and showed an interaction with N-fertilisation: For N1, Arnold's maximum was higher than Bernstein's, while this was inverted for N3.

The observed pattern showed a maximum at anthesis and was generally matched by the model. Absolute values at anthesis were underestimated in EXP1 (RMSE = 2.8 g N m<sup>-2</sup>, RRMSE = 23%, PE = -19.5%) and matched well in EXP2 (RMSE = 1.2 g N m<sup>-2</sup>, RRMSE = 13%, PE = -0.7%). At harvest, iCrop overestimated stem N content on average by 69.5% (RMSE = 2.1 g N m<sup>-2</sup>, RRMSE = 99%), showing higher overestimation with increasing N-fertilisation.

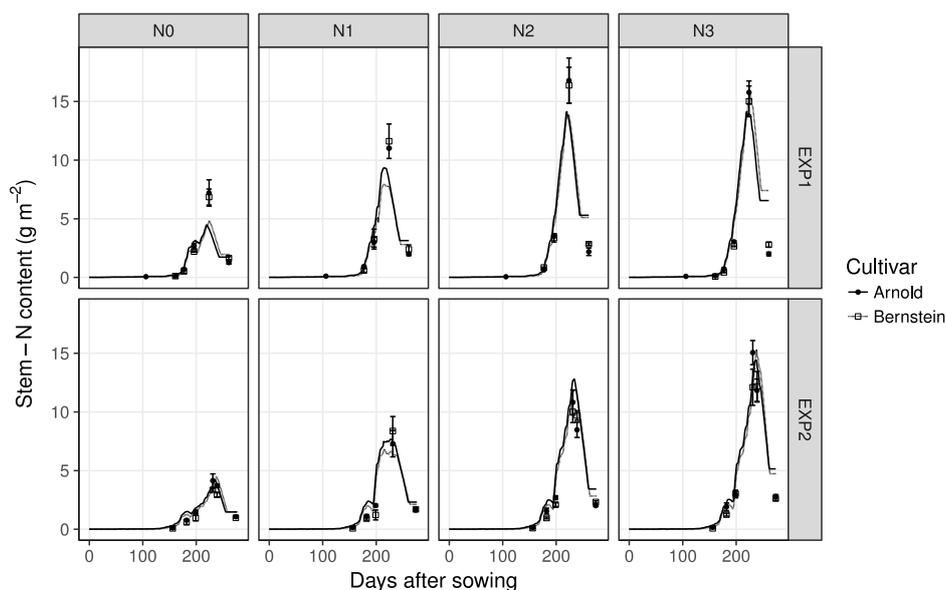


Figure 45 Stem-N content: Observed (symbols) vs. simulated (lines) values of two contrasting winter wheat cultivars (Arnold and Bernstein) in both seasons (EXP1: 2017/18, EXP2: 2018/19). The N-treatments included no N (N0), 70 (N1), 140 (N2), and 210 (N3) kg N ha<sup>-1</sup>. Bars represent standard errors of the means.

#### 8.4.5. Soil Water

Simulated total soil water content showed a clear difference between N0 (higher) and the other N-treatments, starting ca. 230 DAS (Figure 46). Simulated differences between the cultivars were relatively small, total soil water content for N3 at harvest ranged from 400 mm (Bernstein) to 411 mm (Arnold) in EXP1, and from 279 mm (Bernstein) to 293 mm (Arnold) in EXP2.

Overall, in both seasons and across all cultivars, the relative patterns of simulated soil water matched the observations quite well, especially Bernstein in EXP1, where also the absolute values were matched with astonishing precision (except for harvest). The absolute values were also estimated well for Emilio in EXP1, while they were mainly overestimated for Arnold in EXP1. In EXP2, although initial soil water simulations were parameterised to start lower than observations to better match in-season observations, the rest of the season was mainly overestimated (except for harvest). N0 at harvest in EXP2 was clearly overestimated. In most cases where in-season observed values were available, iCrop overestimated soil water decline in the last steep drop of the graph just before harvest. Thereby, either mid-season soil water was overestimated and harvest matched better (e.g. Bernstein N3, EXP2), or mid-season soil water was matched and harvest underestimated (Bernstein N2 and Emilio N3, EXP1). On average, estimates of the measured treatments (only specific treatments were measured, see Materials and Methods) were robust at both anthesis (RMSE = 39 mm, RRMSE = 9%, PE = +2%) and harvest (RMSE = 47 mm, RRMSE = 13%, PE = +4%).

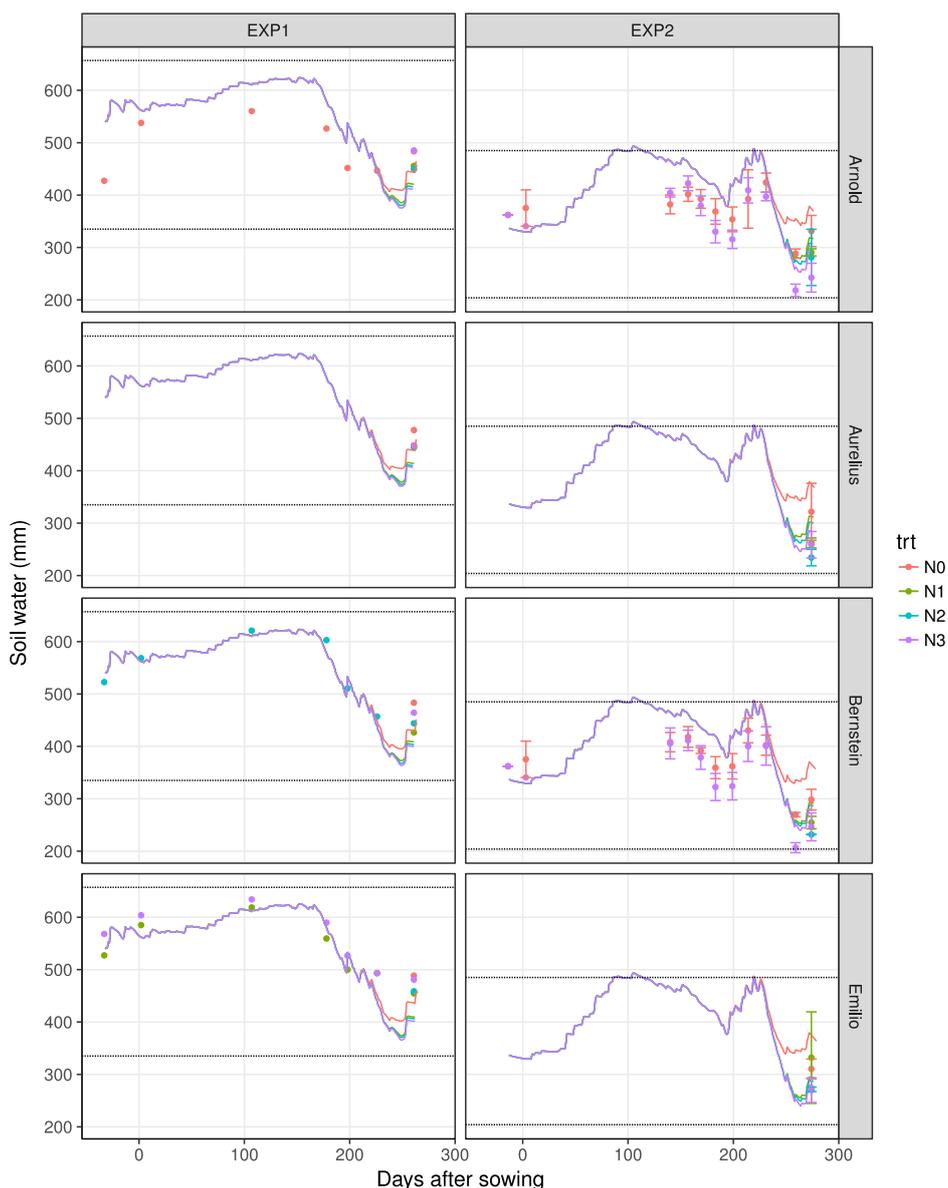


Figure 46 Total soil water content (sum of the whole soil profile, 0-120 cm soil depth): Observed (symbols) vs. simulated (lines) values of four winter wheat cultivars (Arnold, Aurelius, Bernstein, and Emilio) in both seasons (EXP1: 2017/18, EXP2: 2018/19). The N-treatments included no N (N0), 70 (N1), 140 (N2), and 210 (N3) kg N ha<sup>-1</sup>. Bars represent standard errors of the means. Points without bars were single measurements. Dashed lines: drained upper limit (upper line) and lower limit (lower line).

Layer-wise simulated versus observed volumetric soil water (SW) contents are exemplarily shown for specific treatments in Figure 47. The range of simulated SW contents was similar in both seasons in the three topmost layers (0-10, 10-30, 30-60 cm soil depth). In the deeper layers (60-90, 90-120 cm) the simulated amplitude and absolute SW contents were both lower in EXP2. Regarding the effect of N-treatment, lower SW contents were simulated towards the end of the season in the N3 treatments, compared to N1 and N0.

Estimates matched observations variably well. In all layers, estimated SW contents deviated significantly from at least some observation points. However, observed patterns were largely correctly represented by the simulations. For instance, in EXP1 during the winter months (ca. until 150 DAS, mid-March) SW contents increased in the deeper soil layers. Also, in the layers 10-30 and 30-60 cm of EXP2 the sharp decline (from ca. 170-200 DAS) followed by an abrupt rise was predicted well.

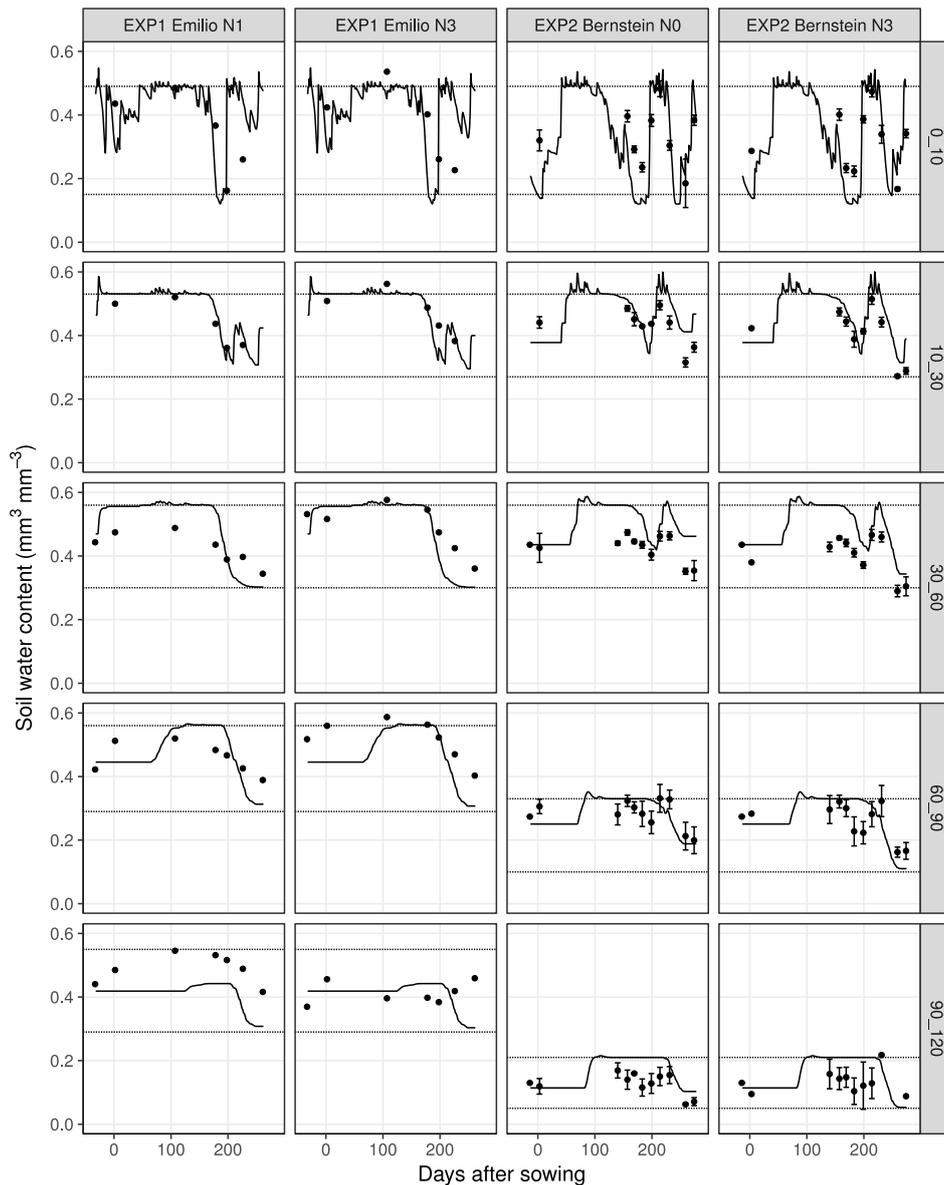


Figure 47 Volumetric soil water content: Observed (symbols) vs. simulated (lines) values of selected treatments (wheat cultivar Emilio in EXP1 [2017/18] and wheat cultivar Bernstein in EXP2 [2018/19]) for each of the defined soil layers (0-10, 10-30, 30-60, 60-90, and 90-120 cm soil depth). The N-treatments included no N (N0), 70 (N1), 140 (N2), and 210 (N3) kg N ha<sup>-1</sup>. Bars represent standard errors of the means. Points without bars were single measurements. Dashed lines: drained upper limit (upper line) and lower limit (lower line).

Simulated cumulative runoff ranged from 44 to 51 mm (EXP1) (data not shown). Much of the runoff (ca. 18 mm) was simulated on a single day (198 DAS) with 64 mm rainfall. In EXP2, runoff

was 42 mm (all simulations identical). Most runoff (ca. 18 mm) occurred between 208 and 219 DAS (cumulative rainfall in this period: 114 mm). Total cumulative drainage was zero in EXP1 and 9 mm in EXP2 (all treatments).

#### 8.4.6. Soil Mineral Nitrogen

In both seasons, initial soil mineral N content (Nmin) was set to the average of observed values. In EXP1, Nmin was measured only at harvest and sowing (Figure 48). In EXP2, Nmin was also measured once during the season (141 DAS), shortly before the first N-fertilisation date.

In EXP1, iCrop underestimated the impact of N-fertilisation. At harvest (EXP1), soil Nmin was predicted well for N0 and N1 (simulated: 2-3, observed: 2-6 kg N ha<sup>-1</sup>), while N2 and N3 were underestimated (simulated: 2-5, observed: 5-15 kg N ha<sup>-1</sup>). In EXP2, the only in-season measurement of Nmin (shortly before the first application of N-fertiliser) was predicted precisely (both simulated and observed: ca. 4 kg N ha<sup>-1</sup>). Also, harvest Nmin of EXP2 was simulated well (2-5 kg N ha<sup>-1</sup> across all treatments). However, iCrop simulations did not represent the observed impact of N-treatments on harvest Nmin correctly, which were higher Nmin with higher N fertilisation (only simulation of N3 in EXP1 followed this trend correctly). The simulations of both seasons resulted in slightly higher Nmin values for N0 than for N1 and N2 (negative influence of N-fertilisation), while observations showed a tendency towards a positive response to increased N-fertilisation.

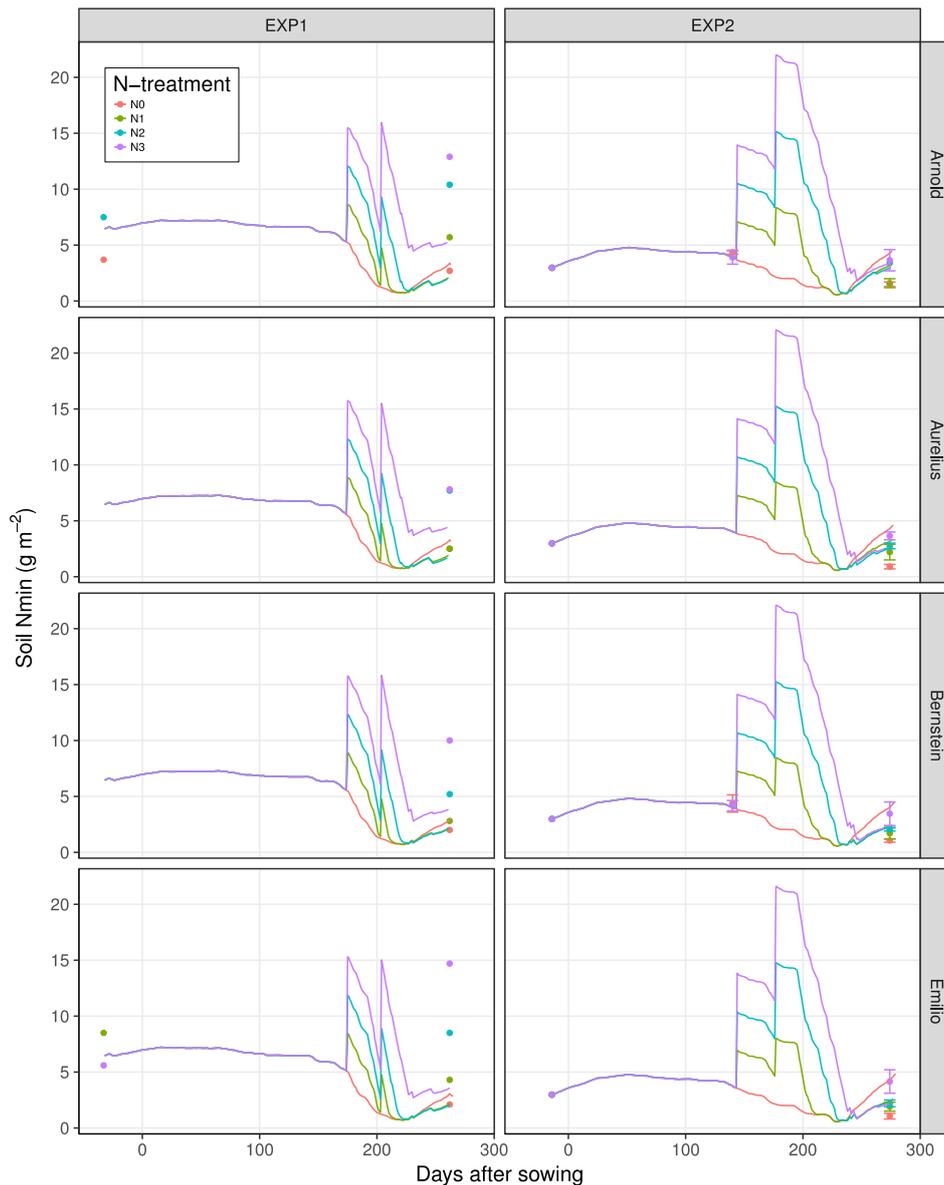


Figure 48 Total soil mineral nitrogen content (Nmin) (sum of the whole soil profile, 0-120 cm soil depth): Observed (points) vs. simulated (lines) values of four winter wheat cultivars (Arnold, Aurelius, Bernstein, and Emilio) in both seasons (EXP1: 2017/18, EXP2: 2018/19). The N-treatments included no N (N0), 70 (N1), 140 (N2), and 210 (N3) kg N ha<sup>-1</sup>. Bars represent standard errors of the means. Points without bars were single measurements

## 8.5. iCrop Evaluation

### 8.5.1. Long Term Data

The available field experimental data covered grain yield, sowing date, heading date, and harvest date, as well as basic management information (dates and amounts of nitrogen fertiliser and irrigation application). No information on soil variables, such as initial/final soil water and mineral nitrogen content, was available.

Cumulative rainfall during the growing season (Oct. to June) across all twelve environments (location x year) ranged from 216 to 450 mm (mean: 341 mm). This was, on average, 10% less

than in EXP1 (374 mm) and 41% less than in EXP2 (481 mm). Average temperatures during the growing season ranged from 7.3 to 9.4 °C (mean: 8.6 °C) which was lower than in EXP1 (9.3 °C) and EXP1 (9.4 °C). Cumulative global radiation during the growing season ranged from 2459 to 2800 MJ m<sup>-2</sup> (mean: 2645 MJ m<sup>-2</sup>). On average, this was almost identical to EXP1 (2644 MJ m<sup>-2</sup>) and 3% lower than in EXP2 (2729 MJ m<sup>-2</sup>).

### 8.5.2. Evaluation Simulation Setup

Due to the evaluation data set being incomplete for the model setup, some required model input data had to be estimated. This included soil parameters and soil initials. The estimates were derived manually by comparison of soil data from EXP1 and EXP2 as well as data from a previous experiment in Eastern Austria (Fuchs, 2016) and freely available soil data in the soil maps “eBod” (Wandl and Horvath, n.d.) and ESDAC (Hiederer, 2013a, 2013b). The same soil parameterisation and initialisation was used for all evaluation simulations, shown in Table 9.

*Table 9 Soil parameters and initials used for the evaluation simulations. #: layer number (top to bottom), DLYER: layer thickness, SAT: saturated soil water content, DUL: drained upper limit, LL: lower limit, ADRY: air-dry lower limit, iniWL: initial soil water content, DRAINF: drainage factor, FG: gravel fraction, BDL: soil bulk density, NORG: soil organic nitrogen content, FMIN: fraction of NORG available for mineralisation, Nmin: soil mineral nitrogen content.*

#	DLYER	SAT	DUL	LL	ADRY	iniWL	DRAINF	FG	BDL	NORG	FMIN	Nmin
	mm			mm <sup>3</sup> mm <sup>-3</sup>			-	-	g cm <sup>-3</sup>	%	-	kg ha <sup>-1</sup>
1	300	0.37	0.32	0.12	0.08	0.32	0.5	0	1.25	0.14	0.097	15
2	300	0.35	0.30	0.08	0.08	0.25	0.3	0	1.23	0.10	0.07	10
3	300	0.33	0.28	0.07	0.07	0.23	0.3	0	1.28	0.06	0.01	5
4	300	0.32	0.27	0.07	0.07	0.22	0.3	0	1.3	0.06	0.01	5

The simulation start dates were set identical for all locations (October 9<sup>th</sup>). While earlier simulation start dates might seem appropriate to have the model estimate soil water content automatically via precipitation from weather input, tests showed that earlier simulation start dates hardly affected simulation output.

Sowing dates, irrigation amounts and their application dates (if any), and fertilisation amounts and application dates were set exactly to the values provided in the data set.

### 8.5.3. Phenology

The date of heading (BBCH 59) was estimated well by the model (Figure 49). Observed vs. simulated average heading occurred 216.1 vs. 215.3 DAS (Arnold), 217.3 vs. 216.3 DAS (Aurelius), 222.3 vs. 220.8 DAS (Bernstein), and 218.3 vs. 218.0 DAS (Emilio) (overall RMSE = 2.9 days, RRMSE = 1%, PE = -0.4%).

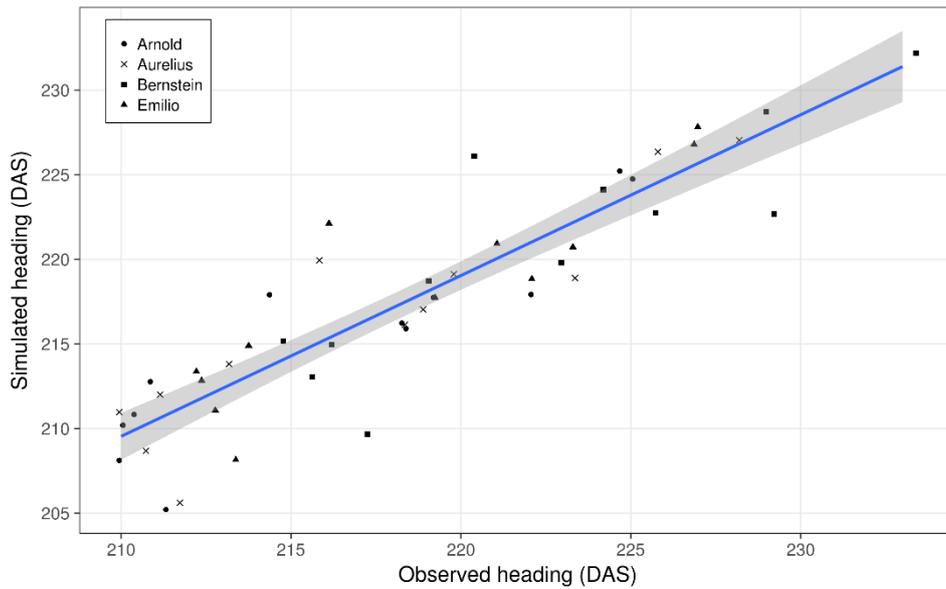


Figure 49 Observed vs. simulated heading date (DAS: days after sowing) of four winter wheat cultivars (Arnold, Aurelius, Bernstein, Emilio) at four locations in three seasons. Blue line: linear regression. Shaded area: standard error of the regression. Points were slightly jittered to avoid over-plotting.

#### 8.5.4. Yield

Long-term yield simulations deviated largely from observations (Figure 50, Figure 51, and Figure 52). Simulations mainly underestimated the yield (RMSE = 198 g m<sup>-2</sup>, RRMSE = 26%, PE = -17.6%). Overall best estimates were found for Arnold (RMSE = 136 g m<sup>-2</sup>, RRMSE = 20%, PE = -9.1%) (Figure 50).

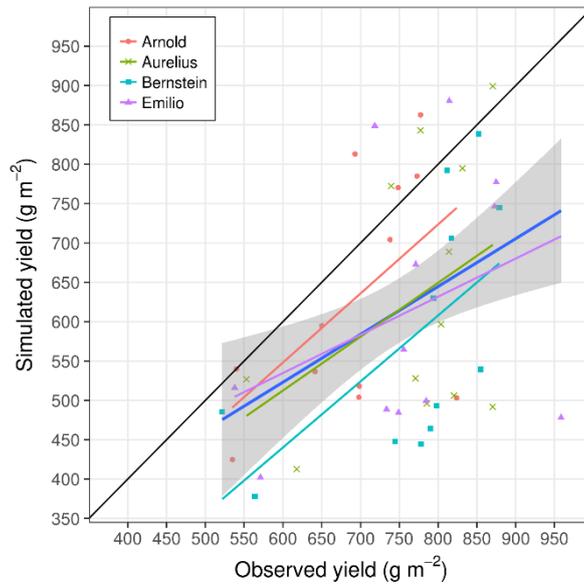


Figure 50 Simulated vs. observed yields of four winter wheat cultivars (Arnold, Aurelius, Bernstein, Emilio) at four locations in three seasons. Lines: linear regression (blue line: average across all cultivars; red: Arnold, green: Aurelius, turquoise: Bernstein, violet: Emilio), black line: 1:1. Shaded area: standard error of the average regression.

The range of observed yields (i.e. range of the 25<sup>th</sup> to 75<sup>th</sup> percentile; boxes excluding whiskers in Figure 51) was not covered by the simulations, except for Arnold. For each cultivar, mean yields were simulated lower than observed.

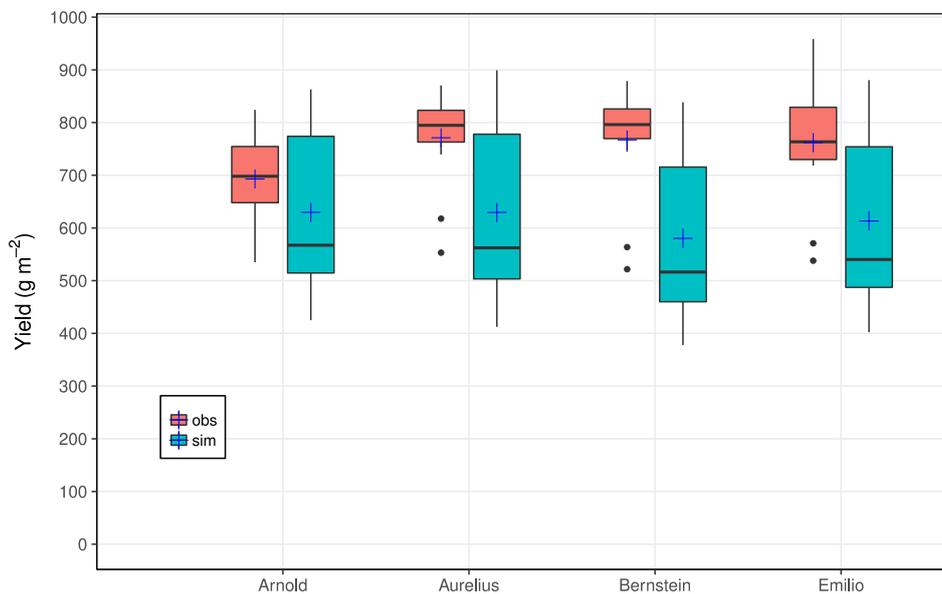


Figure 51 Simulated (red) vs. observed (turquoise) grain yield of four winter wheat cultivars (Arnold, Aurelius, Bernstein, Emilio) at four locations in three seasons. Boxes: 25<sup>th</sup> to 75<sup>th</sup> quartile. Whiskers: value within 1.5 inter-quartile ranges. Points: outliers. Blue crosses: mean. Black horizontal lines in the boxes: median.

Figure 52 shows yield (simulated and observed) versus simulated evapotranspiration (ET). Observed ET was not measured, therefore simulated ET was also used for observed yields. Simu-

lated yields responded much stronger to simulated ET than observed yields. While yield estimates above ca. 500 mm ET were very similar to observed yields, the model underestimated yield increasingly with decreasing ET.

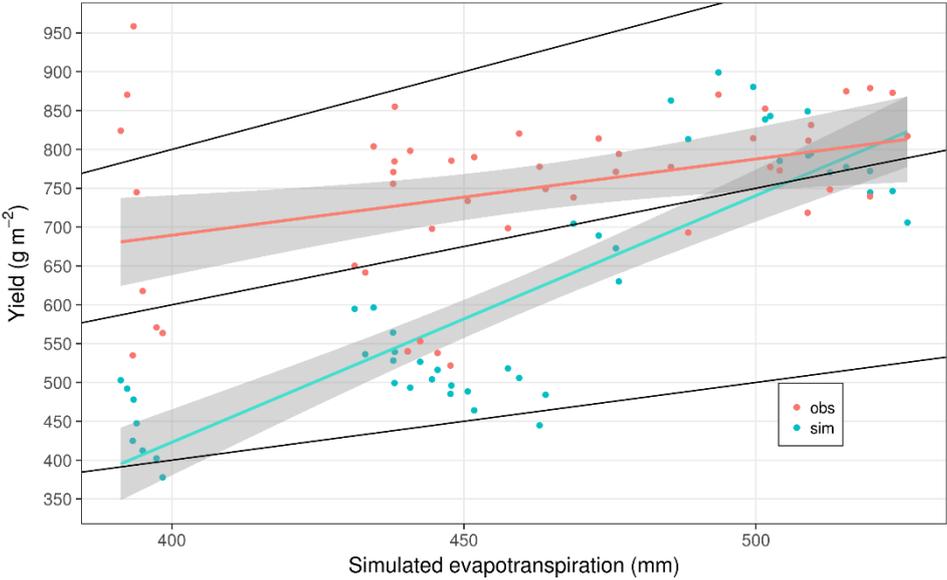


Figure 52 Observed yields (red points) and simulated yields (blue points) versus simulated evapotranspiration of four winter wheat cultivars (Arnold, Aurelius, Bernstein, Emilio) at four locations in three seasons. Black lines: water use efficiency thresholds (top to bottom: 2, 1.5, 1  $\text{g m}^{-2} \text{mm}^{-1}$ ). Coloured lines: linear regression. Shaded areas: standard error of the regression.

## 9. Discussion

The overall objective of this study was the detailed parameterisation and evaluation of the iCrop growth model for the simulation of phenology, growth, canopy development, and N uptake of winter wheat cultivars grown in a temperate environment (Eastern Austria).

For the following discussion, I structured the sections in a similar way as the results: It starts with a relatively technical discussion of the winter wheat parameterisation field experiments (EXPI: 2017/18, EXP2: 2018/19, both at Tulln, Austria) in section 9.1 since they served as an important basis for most of the work in this study. Thereafter, the key topics of this study follow, which are the calculation and estimation of crop model parameters and initials from the parameterisation field experiments (9.2), followed by the simulation results of the parameterised crop model regarding the parameterisation field experiments (9.3), and the evaluation of the parameterised crop model using long-term independent data sets of the same winter wheat cultivars as used in the parameterisation field experiments (9.4). While 9.2 (parameter calculation) and 9.3 (model simulation results) naturally overlap since the simulation results were based on the calculated parameters, I still stuck to separating them in different sections to stress their separation in the modelling process and to provide a clear structure to the reader. Finally, a brief section on the integration of the iCrop model within the scope of the Farm/IT project is presented (9.5), showing the translation of my findings into application.

### 9.1. Field Experiments

This section focuses on various aspects of the parameterisation field experiments, such as the impact of seasonal weather conditions on the winter wheat crop and the influence of different locations of the two parameterisation field experiments. Also, seemingly surprising results are discussed, e.g. clear differences in days to emergence and tillering between the first (EXPI) and the second (EXP2) parameterisation field experiment despite similar weather conditions. Measurement timings and techniques are critically reviewed. Additionally, cultivars are compared regarding their performance (e.g. yield) as well as their interaction with nitrogen fertilisation treatments.

#### 9.1.1. Weather

Growing conditions varied, being advantageous in EXPI (2017/18) until anthesis but better in EXP2 (2018/19) thereafter. While initial soil water content was lower by ca. 200 mm in EXP2, the distribution and amount of rainfall during the growing season (Figure 10) was slightly more favourable in EXP2, thereby overcompensating the initially lacking water. In EXPI rainfall was low in February and March, and almost zero in April. Contrary, rainfall in EXP2 increased from

February (10 mm) until April (50 mm). Also, rainfall in May was higher in EXP2 than in EXP1. While precipitation during winter time is mostly stored in the soil and taken up later in the season by the crop, it is also prone to evaporation due to lack of a closed canopy cover (plants are small, LAI is far below 1). Therefore, precipitation during the highest dry mass production phase, starting ca. in April, might be more efficient. Overall, in EXP2 cumulative precipitation was ca. 100 mm higher than in EXP1. Although EXP2 still had ca. 100 mm less total input water (i.e. precipitation plus initial soil water) than EXP1, the physiologically better distribution of water input was apparently superior to the quantitatively lower total input, as shown by higher average yield (EXP1: 663, EXP2: 726 g m<sup>-2</sup>) and dry mass of the high fertilisation treatment N3 in EXP2 (Figure 14 and Figure 15, resp.).

#### 9.1.2. Soil Water

Observations showed differences in the amplitude of volumetric soil water content in deeper soil layers (60-120 cm soil depth) between the field experiments (lower in EXP2), while the shallower soil layers (0-60 cm) were similar (Figure 11). Although the experiments were only a few hundred meters apart, the soil map eBod (Wandl and Horvath, n.d.) showed different soil types (see also 9.2.3 “Soil Parameters”). Also, visual comparison of the extracted augers showed obvious differences in the depth of increased sand content in soil layers (visible by the contrast between darker soil and brighter sandy soil) between the experiments (shown for EXP2 in Figure 30), with EXP2 showing sand in shallower layers (60 cm) than EXP1 (90 cm). Since the water storage capacity of sand is lowest among common soil textures (e.g. Rawls et al., 1982), this explains lower measured volumetric soil water content in EXP2 well.

#### 9.1.3. Soil Nitrogen

Observed initial soil mineral nitrogen (N<sub>min</sub>) was lower in EXP2 than in EXP1 (30 vs. 62 kg N ha<sup>-1</sup>) (Table 2). Also, final N<sub>min</sub> was lower in EXP2 than in EXP1 (11-37 in EXP2 vs. 23-114 kg N ha<sup>-1</sup> in EXP1; N0-N3, resp.). Both appeared logical as the previous crop in EXP1 was harvested earlier (wheat in July) than in EXP2 (maize in September) and allowed, therefore, more N to be mineralised in the soil of EXP1 before sowing of the wheat experiment. The differences in final N<sub>min</sub> showed the same trend as the initial differences (EXP2 lower than EXP1). In addition, around and after anthesis the wheat crop in EXP2 experienced better growing conditions than EXP1 (higher rainfall in May, see 8.1.1 Weather), thereby enabling the crop to take up more nitrogen.

Moitzi et al. (2020) conducted field experiments in Eastern Austria, with one same cultivar out of the four wheat cultivars used in my experiments (Bernstein), with zero and 160 kg N ha<sup>-1</sup> fertilisation of different types. When recalculating their numbers using the formula presented

here (see equation (5)), their N balances were  $-30.2 \text{ kg N ha}^{-1}$  without N fertilisation and  $-2.9$  to  $22.3 \text{ kg N ha}^{-1}$  with ammonium nitrate fertilisation. In comparison, Bernstein in my experiments with N0 fertilisation balanced to  $-68.7$  and  $-46.1 \text{ kg N ha}^{-1}$  and N2 ( $140 \text{ kg N ha}^{-1}$  fertilisation) balanced to  $-83.0$  and  $-43.5$  (EXP1 and EXP2, resp.). The main reason for lower N balances in my experiments was that Moitzi et al. (2020) found lower crop N uptake for the fertilised treatments (average NUP of  $163 \text{ kg N ha}^{-1}$  for fertilised treatments) compared to N2 in my experiments (average Bernstein NUP:  $213 \text{ kg N ha}^{-1}$ ). This difference might be related to the application scheme used in their experiments (three splits) compared to my experiments (two splits) and, of course, to the different environments (initial soil water content, rainfall, etc.).

The fact that the N balances were exclusively negative ( $-126.5$  to  $-10.0 \text{ kg N ha}^{-1}$ , Table 2) was not surprising since net soil N mineralisation was not accounted for in the calculation. Geisseler et al. (2019) found an average annual net soil N mineralisation in the top 30 cm of Californian (US) fields ranging from 76 to  $123 \text{ kg N ha}^{-1}$ . Risch et al. (2019) investigated 30 grasslands worldwide and found realised net N mineralisation rates during peak plant biomass production of roughly 0.1 to 0.4 (maximum:  $1.4 \text{ mg N kg}^{-1} \text{ soil day}^{-1}$ ), which translates to 11.7 to 46.8 (maximum:  $163.8 \text{ kg N ha}^{-1}$  per month) for the top 30 cm of the soil, assuming a soil bulk density of  $1.3 \text{ g cm}^{-3}$ . Given these ranges, the assumption that the observed negative N balance of up to  $126.5 \text{ kg N ha}^{-1}$  was caused by the soil's net N mineralisation appears logical and realistic.

The effect of varying N fertilisation levels on the N balance (Table 2) was, however, unclear. Higher (i.e. closer to zero) values for the N balance, indicating higher soil-N input/output ratio, were similarly often found in both N3 and N0 treatments. While this seems paradox, both can be argued. For N3, a high N balance makes sense since the N uptake efficiency (ratio of crop NUP to N supply) decreases with increasing N supply. Also, a relatively high N balance for N0 can be argued, since N0 crops likely grow poorly due to N stress due to their complete dependence on soil Nmin resources. However, Soltani and Sinclair (2012) incorporated an inhibition effect of present soil Nmin on the mineralisation rate in iCrop based on previous findings showing reduced N mineralisation with higher soil Nmin contents (Hadas et al., 1986; Sinclair and Amir, 1992). With this assumption, a closer to zero N balance makes more sense for N3 than for N0, since the effect of the (in the N balance calculation) unaccounted N mineralisation is reduced in N3. Furthermore, the measurement technique of Nmin was prone to variation and errors since the intense effort to collect samples resulted in a relatively low sample number (three per plot) and Nmin content is known to possibly vary considerably within close proximities. Overall, the effect of N fertilisation on N balance remained ambiguous.

#### 9.1.4. Phenology

Phenological observations of EXP1 and EXP2 showed fastest development of winter wheat cultivar Arnold, intermediate developments of Aurelius and Emilio, and slowest development of Bernstein (Figure 12). This matched long term observations of AGES (Austrian Federal Agency for Health and Food Safety) that are based on different environments (AGES, 2020) (Table 10).

Observations showed by tendency a slight accelerating effect of the unfertilised treatment (N0) on development. Other studies found no effects of N fertilisation rate on wheat phenology under field conditions (e.g. Basso et al., 2010; Davidson and Campbell, 1983; Salvagiotti and Miralles, 2007).

*Table 10 Rankings of phenological stages (booting; BBCH 41-49, ripening; BBCH 83-89) and yield (dry, wet areas) for the cultivars Arnold, Aurelius, Bernstein, and Emilio. 1: very early/low, 9 very late/high (AGES, 2020).*

	Booting	Ripening	Yield (dry area)	Yield (wet area)
Arnold	2	2	3	3
Aurelius	3	4	7	6
Bernstein	6	7	7	6
Emilio	4	3	7	6

The influence of photoperiod (day length) and vernalisation (requirement of a cold period) can affect phenology and especially cumulative temperature sums to reach a specific stage (e.g. temperature sum from sowing until anthesis). Photoperiod did not influence the two experiments differently since planting date was almost identical (mid-October). Also, the vernalisation demand was satisfied in both years. For instance, based on observations, the APSIM model (Keating et al., 2003) defines 50 days at optimum vernalisation temperature of 2 °C for winter wheat vernalisation, while temperatures between 0 and 15 °C are also effective but prolong the required period. Compared to the observed mean monthly temperatures in EXP1 and EXP2 (Figure 10) it can be assumed that vernalisation was easily met during winter time in both experiments. Therefore, cultivar-specific temperature sums were expected to be equal in both experiments (assuming no influence from drought and N stress).

The delayed observed emergence in EXP2 (24 days after sowing, DAS; 250 °Cd) compared to EXP1 (15 DAS, 123 °Cd) may have been caused by differences in observed initial topsoil water content. The 0-10 cm soil layer contained 0.41 cm<sup>3</sup> cm<sup>-3</sup> volumetric soil water in EXP1 and only 0.31 cm<sup>3</sup> cm<sup>-3</sup> in EXP2. In both experiments, the lowest measured values (i.e. estimate for wilting point) were similar (also the corresponding model parameter lower limit was set identical to 0.15 in both experiments). Due to the higher water content in EXP1 seeds had probably enough available soil water for water absorption and germination. However, in EXP2 initial soil water

was probably too low for germination, but a rainfall event 9 days after sowing (amount: 9 mm) appears to have initiated seed water absorption and the germination process. Calculated from that rainfall event onwards, emergence took 15 days (identical to in EXP1).

Earlier initiation of tillering in EXP1 (87 DAS) compared to EXP2 (127 DAS) was possibly partly caused by slightly differing scoring methods due to changing personnel, error resulting from interpolation of the relatively long measurement intervals in EXP1 (up to two months in winter), and, as mentioned above, delayed emergence in EXP2. Measures were taken to minimise the errors from different scoring-personnel by doing measurements together several times and comparing and adjusting the recorded values. However, the largest error contributor were likely the long measurement intervals in EXP1 which made it impossible to estimate the true beginning of tillering. The estimated date using interpolation very likely included a large error, possibly several weeks.

While, in terms of days after sowing, anthesis occurred 12 days earlier in EXP1 than in EXP2, the relation was the other way around when comparing cumulative temperature sums from sowing (EXP2 56 °Cd earlier than EXP1) or emergence (79 °Cd). However, this difference constituted only two to three biological days (one biological day being a day with 27.5 °C average temperature). This relatively small difference may easily be explained by default experimental and calculation errors (e.g. due to interpolation between measurements). Therefore, in terms of temperature sums, beginning of anthesis can be assumed equal in the two experiments.

The phenological differences between the cultivars were largest around anthesis but these differences mainly disappeared after ca. BBCH 70 (beginning of fruit development) (Figure 13). This was possibly caused by difficult scoring of the phenological stages from that stage onwards (grains must be inspected). Continuing the seasonal trends, EXP1 reached physiological maturity (BBCH 87) 6 days earlier than EXP2. In terms of temperature sums since emergence, EXP1 reached physiological maturity 145 °Cd later than EXP2. Compared to anthesis, the difference between the experiments had grown. A likely explanation for this is terminal drought stress which possibly accelerated phenology in the last days/weeks of EXP1 caused by lower rainfall amounts in April and May in EXP1.

#### 9.1.5. Leaf Development

Among the cultivars, Bernstein showed the highest LAI at anthesis, Arnold the lowest. However, Bernstein and Arnold had the exact same average leaf number on the main stem (MSNN) at anthesis (10.7), while the other cultivars Emilio (11.2) and Aurelius (11.4) had significantly higher MSNN. Tiller number can affect LAI but was not measured at anthesis. While the number of green leaves (only measured in EXP2) did not differ significantly between the cultivars, there

was a trend showing a much higher leaf number for Bernstein (2070 leaves per m<sup>2</sup>) than for Arnold (1695) which could explain the difference in LAI. The possibility of different stay-green behaviours is discussed later (9.3.3).

#### 9.1.6. Dry mass

Total above-ground dry mass (TDM) at harvest for cultivar Bernstein was 729-1636 g m<sup>-2</sup>. This compared well to a study by Moitzi et al. (2020) who conducted field experiments with Bernstein in Eastern Austria (seasons 2016/17 and 2017/18) and found 1023 g m<sup>-2</sup> (no N fertilisation) and 1227-1376 (with 160 kg N ha<sup>-1</sup>) TDM at harvest.

At harvest, the influence of N-treatment on crop dry mass was mostly stronger in EXP2 than in EXP1 (Table 4, Figure 15). This was expected, as the previous unfertilised maize crop was harvested only a few weeks before winter wheat sowing in EXP2, thereby extracting much of the soil N<sub>min</sub> and allowing little time for N mineralisation. Contrary in EXP1, the previous unfertilised crop was wheat which was harvested in July, giving soil N mineralisation ca. three months to accumulate (see above 9.1.3 Soil Nitrogen). As a result, plants in the N0 treatment of EXP2 had less N available than in EXP1. Of course, the same is true for N3, but it is a well-known fact that yield and total dry mass respond less with increasing N-fertilisation, up to a point where the response turns negative. This was described early by Mitscherlich and Boguslawski, named the “law of diminishing yield increment” (Boguslawski, 1958; Mitscherlich, 1954 cited in Marschner, 2012).

#### 9.1.7. Crop Nitrogen

Grain nitrogen concentration (%GN, Table 4) ranged from 1.6 to 3.0% and was similar to reports from other studies in Eastern Austria. Neugschwandtner et al. (2015) measured 2.07% for a relatively old cultivar (Xenos) grown in 2011/12 at Raasdorf (east of Vienna). At the same location using the cultivar Bernstein, Moitzi et al. (2020) found 1.37% (no N fertilisation) and 2.44-3.03% (160 kg N ha<sup>-1</sup>) in the seasons 2016/17 and 2017/18.

Among cultivars, the average %GN was highest for Arnold (2.5%), the other three cultivars were almost identical (2.2%). This appeared to be easily explained by the well-known nitrogen dilution effect (e.g. Justes et al., 1994) since Arnold showed lowest grain yield and lowest total dry mass of all cultivars. However, Arnold might have had a genetic advantage over Bernstein. Results from EXP2 point at this, where the N3 treatments of Arnold and Bernstein showed similar final dry mass and yield, but Arnold had 2.8 %GN while Bernstein had only 2.5. Contrary, in EXP1, Bernstein's yield and dry mass was clearly higher than Arnold's, and the difference in %GN

was still 0.3% (Arnold: 3.0, Bernstein: 2.7). Data from the two field experiments was not sufficient to draw a conclusion regarding any genetic %GN advantages, but further investigation would be interesting.

#### 9.1.8. Yield

Observed yields of Bernstein ranged from 303-717 g m<sup>-2</sup> (Table 4). In a comparable study in Eastern Austria (Raasdorf, east of Vienna), Moitzi et al. (2020) found similar Bernstein yields, ranging from 460 g m<sup>-2</sup> (no N fertilisation) to 535-624 (160 kg N ha<sup>-1</sup>). In a long-term winter wheat experiment at the same location, Neuschwandtner et al. (2015) found yields of different cultivars ranging from 130 to 623 g m<sup>-2</sup> for different pre-crops and tillage systems. The extremely low yield (130 g m<sup>-2</sup>) occurred in 2012 where severe droughts affected the crops (seasonal rainfall 2012: 221 mm; long-term average: 362 mm). Ignoring that year, minimum yield found by Neuschwandtner et al. was 314 g m<sup>-2</sup> which compared well to my experiments (minimum: 268).

Similar to dry mass, yield (Figure 14) showed a stronger response to the N treatment in EXP2 than in EXP1 due to lower initial soil N<sub>min</sub> as well as due to the better growing conditions during anthesis and grain filling phase (discussed previously). Therefore, the yield-range in EXP2 was larger (268 - 788 g m<sup>-2</sup>) than in EXP1 (424 - 719). Also, the stronger yield response to N fertilisation at higher fertilisation levels (N2, N3) in EXP2 compared to EXP1 was affected by the different initial soil N<sub>min</sub> contents in the experiments.

The average yields of the cultivars (Aurelius 631 > Bernstein 611 > Emilio 577 > Arnold 515 g m<sup>-2</sup>) were partly in agreement with publicly available long-term yield-scores (Table 10), where Aurelius, Bernstein, and Emilio were rated equal (score: 7/10) and Arnold (3/10) clearly lower.

### 9.2. Calculation of Model Parameters and Initials

In the following section, different aspects of the calculation and calibration of the iCrop model parameters are discussed. I explain why I chose which method for acquiring a parameter value (e.g. calculation or calibration, using which subset of observation data, pooling cultivars or grouping data by cultivar, etc.). Also, possible error sources and alternative approaches are evaluated. While some effects of the parameter choice on model simulation results are mentioned, the main discussion of crop model simulation results is provided in the subsequent sections (9.3 Crop Model Simulations of the Field Experiments and 9.4 Crop Model Evaluation).

#### 9.2.1. Simulation Initialisation

Simulations were initialised (i.e. the simulation start date was set) on 13.9.2017 (33 days before sowing) in EXP1 and on 1.10.2018 (14 days before sowing) in EXP2. These dates were picked to

match the dates of the first soil water measurements in each season. Although soil water measurements were also available for the date of sowing in both experiments, I used the earlier dates for initiation of the simulation. This ensured “smoothing” of some modelled processes that require a few days after simulation start to achieve a more stable state. Such a process is vertical soil water distribution which shows relatively high water fluxes between neighbouring soil layers for a few days after simulation start due to strong gradients. These gradients are caused by the set of initial soil water values which were measured. Since the model is a simplified representation of reality, observed soil water gradients deviate from simulated gradients which results in the simulated high water fluxes for a few days after initiation. Therefore, starting the simulations very close to sowing date may lead to significant influences on e.g. emergence due to increased or decreased soil water availability during the described water fluxes between soil layers. To avoid such effects, I chose the earlier model initiation dates.

#### 9.2.2. Climate Parameter: CO<sub>2</sub>

White et al. (2011) found that in many climate change impact simulation studies researchers used outdated atmospheric CO<sub>2</sub> levels as baseline. They suggested using climate data close to the date of publication. Translated to the objective of this study, using CO<sub>2</sub> levels corresponding to the seasons when the field experiments were carried out appeared sensible. Therefore, for EXP1 (season 2017/18) and EXP2 (2018/19) I set the model’s CO<sub>2</sub> parameter to 400 ppm as this value corresponds closely to 2016’s average global atmospheric CO<sub>2</sub> levels (Butler and Montzka, 2020). iCrop’s previous default value (350 ppm, ca. corresponding to 1988’s global average) was clearly outdated.

#### 9.2.3. Soil Parameters

Initial soil water contents of the soil layers in EXP1 was set to the average per layer of all observations (excluding the sampling point closest to the wind break). For EXP2, initial soil water was based on the observations and then calibrated to better fit in-season measurements (Figure 46) as well as to avoid simulated growth stress through excess soil water (oversaturation) during winter time. This occurred when using the exact observed initial values (data not shown). While stress through oversaturation was avoided in this way, in-season observations were still overestimated. This is further discussed in 9.3.1 Soil Water.

The choice of soil water parameters (Figure 31) resulted in a total plant-extractable soil water content (EXTR) for the whole profile of 321 mm in EXP1 and 281 mm in EXP2 (1200 mm soil depth). Manschadi et al. (2020a) used iCrop soil parameters resulting in a similar EXTR value (251 mm) as in EXP2 based on a nearby experiment with maize. The eBod soil map (Wandl and Horvath, n.d.) classified EXTR for the locations of both EXP1 and EXP2 as “intermediate” (140

to 220 mm), while only EXP2 was directly adjacent to an area classified as “high” (220 to 300 mm). The range of “intermediate” is considerably lower than the EXTR used for EXP1, and still lower compared to EXTR of EXP2. However, from the documentation of eBod it is not clear which soil depth is used for calculating EXTR. The eBod map is scaled rather coarse into “shallow” (<30 cm), “intermediate” (30-70 cm), and “deep” (>70 cm) soils, with both EXP1 and EXP2 being classified as “deep”. It appears possible that eBod does not account for soil water from layers deeper than commonly used soil depths for sampling (often 90 or 100 cm) for calculating EXTR. When considering only 90 cm of the parameterised soil, EXTR of EXP1 amounts to 245 mm and EXP2 to 233 mm. These reduced values are only slightly higher than eBod’s range for “intermediate” EXTR (140-220 mm). Nevertheless, eBod classified the experiments’ locations differently for the indicator “water supply” (“Wasserverhältnisse”) which describes the “average soil water available to plants”. On a six-parted scale (wet, rather wet, optimal, rather dry, dry, very dry) EXP1 was classified “rather wet” while EXP2 indicated “rather dry”. This fits the relative differences used for the soil parameters and EXTR (EXP1 higher than EXP2). Also, in the comments, eBod noted for the location of EXP1 “very high water storage capability” and for EXP2 “high water storage capability” which also agrees with the chosen parameters.

The use of different soil parameters for EXP1 and EXP2 (Table 7: soil organic N, Figure 31: soil water) might be questioned due to their proximity. However, as mentioned above and in the “Materials and Methods” chapter, the soil map eBod indicated different soil types for the two experiments. In addition, measurements of TOC (soil total organic carbon content) showed clear differences between deeper soil layers of EXP1 and EXP2 (Table 7). Lower TOC of deep layers in EXP2 supported the theory of more sand and lower soil water storage capacity.

#### 9.2.4. Effect of Rooting Parameters on Yield

While *GRTDP* (root depth growth) was an available parameter in iCrop, there was no root exploration factor parameter that defined soil water extraction rate. Other models such as APSIM (Holzworth et al., 2014) use a parameter (*KL*) to define the fraction to which each soil layer is explored by the roots, limiting daily maximum soil water and N uptake. In a previous study, Manschadi et al. (2006) have shown that different wheat genotypes may express very different root architectures which are closely related to drought tolerance. While they found clear differences between the root architectures and water uptake of a drought-tolerant wheat variety (SeriM82) compared to a standard variety (Hartog) in a sophisticated experimental setup using large rooting boxes, they could not measure these differences in water uptake under field conditions. Therefore, there may have been (unidentified) significant differences in the root architectural systems of the cultivars in my experiments, which could explain their differences in yield

and other traits. As an example, by simply decreasing *GRTDP* for Arnold to 20 mm bd<sup>-1</sup> (default: 30), several simulation outputs, including yield, improved substantially (data not shown). However, to justify such parameter modifications, more detailed observed soil data showing clear differences in root architecture would be required. Future experiments could, for instance, investigate root architectural traits of different cultivars using established methods such as in-vitro rooting experiments for root scanning and analysis. Also, innovative techniques to create field-like environments under controlled conditions such as spectral imaging of rhizoboxes (Bodner et al., 2021, 2018) appear promising to gain additional information on cultivar-specific root growth. While such experiments are known to not necessarily translate well into field conditions, they still provide an indication which can serve to better parameterise crop models.

In a study by Christopher et al. (2008) the same wheat cultivars as described above (SeriM82, Hartog) were tested. They found that the drought-tolerant wheat cultivar SeriM82 also exhibited a stay-green phenology, maintaining green leaf area longer into the grain filling period than Hartog. They pointed out that yield advantage of SeriM82 was closely related to the stay-green trait, and also that stay-green was lost when deep soil water was depleted. This indicated that both deep soil water availability and the crop's ability to extract it may be of great importance in dry seasons. However, the cultivars in my study were parameterised with identical *GRTDP* and could, therefore, not reproduce any possible differences in rooting behaviour. Also, iCrop's lack of a root exploration factor such as *KL* in APSIM contributed to this problem. In conclusion, iCrop's inability to capture the clear observed yield differences between Arnold and the other wheat cultivars may be solved by introducing a root exploration parameter (similar to *KL*) and/or conducting experiments to generate high-quality soil data sets as described above to support parameterisation of cultivar-specific rooting differences.

#### 9.2.5. Effect of Soil Runoff on Crop Dry Mass

iCrop assumed runoff by default for rain-fed situations as in the parameterisation field experiments presented in this study (simulated runoff: 44-51 mm in EXP1, 42 mm in EXP2). Simulation of runoff lowered plant available soil water content critically in EXP1. In test-simulations, simply deactivating runoff led to improvement of dry mass and yield simulations in EXP1 (data not shown). However, deactivating runoff had an adverse effect on simulations of EXP2 where an intense rainfall period between 208 and 219 days after sowing led to soil water oversaturation which in turn decreased dry mass and yield production substantially through excess soil water stress (data not shown). While deactivation of runoff for a seemingly flat field seems reasonable, several studies (e.g. Chaplot and Bissonnais, 2003; Fox et al., 1997) have shown that runoff can occur on fields at very small slopes (1.5 ° to 2.3 °). For EXP1 and EXP2 slopes were not measured

while they appeared flat to the “naked eye”. However, considering that runoff is generally possible with even slight slopes which may not be visible to the human eye without measurement, and also the fact that simulations overall worsened without runoff, I decided to stick to the default (i.e. runoff activated).

#### 9.2.6. Phenology Parameters

With the default critical fraction of transpirable soil water threshold for germination ( $FTSW = 0$ ), the date of emergence in EXP2 was simulated too early (data not shown). I calibrated the critical  $FTSW$  threshold for germination in EXP2 manually ( $FTSW = 0.3$ ) for better simulation of the emergence date. While this calibration might be viewed as over-parameterisation, I argue that it was necessary in the scope of this study. The aim here was to parameterise the whole iCrop model and investigate model performance with focus on cultivar differences and N parameters. To be able to parameterise the model with data from the two experiments, good emergence simulation was required. Since all conservative approaches (e.g. calibration of initial soil water content) to simulate emergence correctly in EXP2 failed, I chose to calibrate  $FTSW$  for achieving this.

Another option to calibrate emergence simulation is changing the parameter for sowing depth, which is not available in iCrop but in other models such as APSIM (Keating et al., 2003). While sowing depth on the field was, in theory, identical in both seasons, it is possible that differences in field preparation affected sowing depth and, thereby, emergence date.

The above described results from the simulation of the parameterisation field experiments indicated that iCrop had problems simulating emergence. However, to test the model’s performance in simulating emergence date and the influence of  $FTSW$ , the underlying data set would need to be more comprehensive, covering at least several seasons, locations, and sowing dates. Unfortunately, information on emergence date was not available in the evaluation data set used in this study (see 9.4). Future work should address the question of iCrop’s performance in emergence prediction.

Biological day phase parameters ( $bd<phase\_acronym>$ ) were calibrated based on observed development differences between the cultivars (Table 5). Figure 17 shows the  $bd$ -difference of each cultivar compared to the mean across all cultivars, grouped into three important phenological phases. While the duration of each  $bd$ -phase parameter (e.g. the parameter  $bdSOWEMR$  for the duration from sowing to emergence) can be calculated from observed data, it would not be correct to use this value directly. The reason for this is that environmental factors such as vernalisation, photoperiod, and stresses reduce (sometimes enhance) the effect of the experienced biological day by the crop. Calibrating the  $bd$ -phase parameters was, therefore, challenging as the

observed bd-phase values can hardly be separated from the environmental factors mathematically.

In addition, the observed bd-phase durations had to be calculated using linear interpolation of observed phenological stages to estimate their starting and end points (e.g. BBCH 21, 31, 41 and so on). While this was done based on temperature sums (instead of calendar date) to reflect the strong correlation between phenology and temperature (e.g. Wang and Engel, 1998), phenology expressed in the BBCH scale depending on temperature sums lacks the mathematical properties of continuity and linearity. The BBCH scale is an ordinal scale, meaning that ranking is possible, but not all mathematical operations. While the BBCH scale might be regarded a ratio scale within certain main stages (e.g. BBCH 10-19 where leaves are counted) the numerical difference between two values across the whole scale (BBCH 00-99) is certainly not always representing the same “amount” of phenological progress (e.g. BBCH 25 to 35 vs. BBCH 65 to 75). Therefore, a linear approximation between BBCH values is an option to estimate specific BBCH-stages but also introduces errors which need to be kept in mind when working with such results.

#### 9.2.7. Leaf Development: Phyllochron, LAI

For the simulation of leaf number on the main stem (*MSNN*) in iCrop, I implemented new model source code for adding the functionality to simulate two phases of phyllochron (*phyll*, *phyl2*; Figure 18 and Table 6) instead of only one (default). With only one phyllochron parameter for the whole life cycle of the crop, the model underestimated the in-season course of *MSNN* and reached the observed values only shortly before *MSNN* growth ceased (data not shown). With the implementation of two parameters for phyllochron, both parameterised based on observations, the development of *MSNN* was simulated close to the observations (Figure 33).

While it is largely agreed that phyllochron varies with species and cultivar (e.g. Birch et al., 2003; Cao and Moss, 1989; Frank and Bauer, 1995), the question whether or not phyllochron is generally constant along the whole development of one genotype has been answered differently. While some studies found a rather constant phyllochron across the leaf appearance phase (e.g. Kirby and Perry, 1987), others found that phyllochron was dependent on the phenological stage (e.g. Jamieson et al., 1995; Salvagiotti and Miralles, 2007). Also, the authors of the iCrop model suggested the use of a 2-stage phenology-dependent phyllochron when data supports this (Soltani and Sinclair, 2012). Furthermore, Xue et al. (2004) argued that the phyllochron approach does not at all represent leaf development adequately, and they proposed using non-linear methods instead which they found to give better results. However, using non-linear methods was not considered in my study since the iCrop model should be kept as simple as possible, and using these methods would include complex formulas requiring additional parameters. Xue

et al. (2004) also compared constant phyllochron and a three-phase phyllochron in a two year experiment and found both approaches in different years advantageous above the other. However, in my experiments, while using a two-phase phyllochron did not lead to perfect representation in both years (EXP1 was simulated better than EXP2, Figure 33), in-season predictions of *MSNN* improved compared to the default constant phyllochron. This was the aim, since *MSNN* directly affects *LAI* simulations in iCrop, and *LAI* has several successive effects (e.g. dry matter production, transpiration).

*LAI* growth in iCrop is based on the following equation (see equation (13)):

$$PLAI = PLACON \cdot MSNN^{PLAPOW}$$

The parameter *PLAPOW* has, obviously, a significant impact on leaf area simulations, especially at higher *MSNN*. The calculation of parameters based on the field experiments resulted in higher *PLAPOW* for Bernstein (2.0) than the other cultivars (all 1.9) (Table 5). Similarly, observed *LAI* was highest for Bernstein, while *MSNN* was highest for Aurelius and Emilio (Table 3). Using the formula above for calculating the correlation (Figure 19), the effect of leaf area (*PLAI*) was obviously stronger than that of *MSNN*, resulting in higher *PLAPOW* for Bernstein. A higher *PLAPOW* also leads to increased sensitivity of simulated *LAI* at higher *MSNN*. However, simulations showed good estimates of Bernstein's maximum *LAI* for N3 in both parameterisation experiments (Figure 34), indicating no error from increased sensitivity.

The parameter *PLACON* was assumed 1 as suggested by Soltani and Sinclair (2012). However, the correlation using equation (13) suggested using a lower *PLACON* since leaf areas (*LA*) for *MSNN* between 3 and 6 were overestimated (Figure 19). On the other hand, decreasing *PLACON* to improve the fit of the curve for low *MSNN* and *LA* would result in a steeper incline for higher *MSNN* and *LA*. This would drastically increase the risk of *LAI* overestimation for higher *LA* (around anthesis) due to high sensitivity to *MSNN*. Therefore, I decided to keep the default value.

#### 9.2.8. Dry Mass: Radiation Use Efficiency

iCrop's *RUE* parameter was changed to adapt the value from APSIM (2.58 g MJ<sup>-1</sup>, iCrop default 2.2 g MJ<sup>-1</sup>). With the default value, iCrop underestimated total dry mass (data not shown). Other studies have found *RUE* for wheat of up to 2.47 (Shearman et al., 2005), 2.57 (Rose et al., 2017), or 2.8 g MJ<sup>-1</sup> (Kiniry et al., 1989) which agreed well to APSIM's and iCrop's updated *RUE*.

#### 9.2.9. Crop Nitrogen Parameters

The iCrop model code was updated to support the simulation of stem N content using senesced (minimum) stem N concentration parameters (*SNCS*) depending on crop development stage:

*SNCSI* before BSG (beginning of seed growth, BBCH 71) and *SNCS2* thereafter. This improved simulations of the N partitioning between leaves and stems and, thus, leaf area index, as explained in sub-section 9.3.5 “Effect of Development-dependent *SNCS*”. While the default version of iCrop supported only one *SNCS* parameter for the whole crop growing season, it has been found for wheat that stem minimum nitrogen concentration changes with development stage (e.g. Singh and Porter, 2020; Zhao et al., 2014). Also, other crop models such as APSIM (Holzworth et al., 2014) and AFRCWHEAT2 (Porter, 1993) use crop nitrogen parameters depending on development stage. Manschadi et al. (2020a) also found a development-dependent two-phase *SNCS* necessary to simulate *LAI* of maize correctly with iCrop.

Grain nitrogen concentration parameters (*%GNmin*, *%GNmax*) were calculated from pooled data although statistics found significant effects of the cultivars (Table 4). Therefore, one could argue that these parameters should be calculated per cultivar. However, observed values did clearly not cover the true minimum and maximum values of *%GN* since the experiment was not specifically designed to reach these extremes. While the experimental design offered a large variation in N supply to the crop, various factors diminished the effects of the highest (N3) and lowest (N0) N-treatment to reach maximum and minimum *%GN*, respectively. While the N0 treatment did not receive any mineral N fertiliser (and also the previous crop was not fertilised), a considerable amount of mineral nitrogen was measured in the soil before sowing (63 kg N ha<sup>-1</sup> in EXP1, 30 in EXP2). Also, microbial N release from soil organic N (N mineralisation) added to the N pool of all treatments, leading to increased N-uptake in N0 treatments. Furthermore, stresses such as terminal drought stress and the depletion of soil N towards the end of the season have limited N-uptake also in the high N-treatments. To reach high *%GN* it might be necessary to add an extra N fertilisation around anthesis to boost grain N uptake during grain filling. Overall, it seemed more likely to receive adequate estimates for the *%GN* parameters by pooling the data, generating a larger basis for the calculation.

iCrop limits maximum daily N uptake from the soil (parameter *MXNUP*). The maximum observed values over periods of two (EXP2) to four (EXP1) weeks were similar across the four wheat cultivars (ca. 0.5 g m<sup>-2</sup> d<sup>-1</sup>; Figure 28 and Figure 29). The finally used *MXNUP* parameter value of 0.6 g m<sup>-2</sup> d<sup>-1</sup> (iCrop default: 0.25) was justified by (i) the use of the same value in the APSIM model (Holzworth et al., 2014) and (ii) the fact that the observations were averages over rather long periods (two to four weeks) which likely masked the true daily maximum N uptake capacity of the crop.

#### 9.2.10. Yield Parameters

Bernstein showed the overall lowest observed *DHI* (daily harvest index increase) in EXP1 (Figure 23) due to the longest seed growth period of all cultivars (Figure 12). However, the *PDHI* (potential *DHI*) parameter was calibrated equal for all cultivars (0.017) since the measured cultivar differences (particularly for Bernstein) appeared to have arisen mainly from measurement and interpolation errors. As mentioned earlier, scoring the wheat phenological stages beyond BBCH 71 (beginning of seed growth) is generally challenging. In addition, the linear interpolation of BBCH growth stages to estimate the occurrence of specific BBCH stages was the best option but also introduced errors (as discussed in 9.2.6 Phenology Parameters). Both these sources of error probably resulted in lower calculated *DHI* values for Bernstein in EXP1. To achieve more reliable *DHI* estimates, phenological on-field measurements on a daily basis would be required around BSG (BBCH 71) and TSG (BBCH 87) to be able to exactly determine seed growth duration.

#### 9.2.11. Effect of Sampling Date on Parameter Calculation

Each of the measurements of biomass on the field was done within a single day. The aim was to get crop measurements at specific phenological target-stages, e.g. at sowing, anthesis (BBCH 65), or end of seed growth (BBCH 87). However, due to their genotypic differences the cultivars developed in different speeds. Therefore, they were not exactly in the target-stage when the measurements were done but scattered around the target-stage. For instance, in EXP1 most cultivars had already reached or passed the beginning of seed growth (BSG, BBCH 71) for the anthesis target-stage measurement, while in EXP2 observations scattered from BBCH 62 to 69. Some critical iCrop parameters actually require to be calculated based on field observations at BSG. These parameters are two (out of three) of the stem N concentrations (*SNCSI* and *SNCG*), the specific leaf N for green leaves (*SLNG*), and the fraction of translocatable dry mass (*FRTRL*). Since measured data targeted the anthesis stage (instead of BSG), error was automatically introduced in the parameters. Indeed, for more precise parameter calculations, separate measurements at anthesis and BSG would be desirable. However, since anthesis and BSG occur on the field within a few days, only a low impact on parameter estimates was expected.

The theoretical effect of using measured data from other than the exact optimum phenological stage differs between the parameters. For the stem N concentration parameters *SNCSI* and *SNCG*, the observed SNC (stem N concentration) at BSG was required. Parameters calculated from plant measurements pre-BSG were probably overestimated for SNC and those post-BSG underestimated since the N concentration in stems follows a declining dilution curve. As a result, both *SNCG* and *SNCSI* may be over- or underestimated. Similarly, *SLNG* would be either overestimated (pre-BSG measurements) or underestimated (post-BSG). For *FRTRL*, the error

would be underestimation for pre-BSG, and overestimation for post-BSG, due to the higher influence of  $TDM_{BSG}$  in the denominator of equation (54). Since in most cases biomass samples were rather taken pre-BSG, especially for the slower developing cultivar Bernstein and generally in EXP2, the parameters *SNCG*, *SNCSI*, and *SLNG* may be biased towards overestimation, and *FRTRL* towards underestimation. While keeping the potential biases in mind, the calculated parameters were used in iCrop without calibration or modification. While for *FRTRL* the default model value was used anyway, simulations of stem N (Figure 45) and leaf N content (Figure 44) did not indicate a systematic bias (discussed in more detail for *SNC* in 9.3.6). Therefore, the sampling time can be regarded adequate for the purpose of this study.

#### 9.2.12. Overall Cultivar Differences

Based on observations in the parameterisation field experiments (EXP1, EXP2), the four winter wheat cultivars Arnold, Aurelius, Bernstein, and Emilio were parameterised (overview of the most important parameters in Table 5). The final parameterisation showed no difference between the cultivars for the following parameters (data not shown for some of them): specific leaf area (*SLA*), leaf extinction coefficient (*KPAR*), leaf fraction after flag leaf emergence (*FLF2*), fraction of translocatable dry mass (*FRTRL*), potential daily harvest index increase (*PDHI*), root depth growth (*GRTDP*), minimum stem N concentrations (*SNCSI* and *SNCS2*), specific leaf N concentrations (*SLNG* and *SLNS*), maximum stem N concentration (*SNCG*), grain N concentrations (*%GNmin* and *%GNmax*), maximum N uptake rate (*MXNUP*), vernalisation (*vsen*), critical photoperiod (*cpp*), and photoperiod sensitivity (*ppsen*). Significant differences were observed for other traits, which were reflected in cultivar-specific parameters for leaf area growth (*PLAPOW*), leaf fraction until flag leaf emergence (*FLFIA* and *FLFIB*, *WTOPL*), phyllochron (*PHYLI*, *PHYL2*, *LNP*), and several phenological phase durations (*bdTILSEL*, *bdSELBOT*, *bdBO-TEAR*, *bdEARANT*, *bdANTPM*).

Simulated differences between the cultivars could only originate from parameters which were different between the cultivars. Those were parameters which directly affected traits related to (i) phenology, (ii) LAI, and (iii) dry mass partitioned to leaves. Any simulated cultivar differences found (see 9.3 and 9.4) can be traced back to these three traits.

#### 9.3. Crop Model Simulations of the Field Experiments

While the previous section discussed the determination of the crop model parameters, the focus of the following section lies upon simulation results. iCrop's strengths are pointed out and reasons for poor estimates are discussed where appropriate, giving suggestions to improve the parameterisation for future studies.

### 9.3.1. Soil Water

Total soil water content in EXP1 was clearly overestimated over the whole season for Arnold, while Bernstein was simulated well (Figure 46). Arnold overestimation was likely a result of suboptimal selection of sampling plots in EXP1. As mentioned previously, EXP1 was located relatively close to a wind-break hedge. Soil water samples of Arnold were taken from the replication closest to the wind break, while Emilio and Bernstein samples were taken from replications further away. Therefore, the influence of horizontally spreading roots of the wind-break as well as wind effects have probably caused lower soil water for Arnold. However, the overestimation of Arnold was constant throughout the season and the simulated pattern matched observations. Overall, total soil water content in EXP1 was simulated well by iCrop.

The overestimation of in-season total soil water content in EXP2 remained unclear with several possible explanations. Errors in the soil water measurements seemed unlikely, since all early measurements (before and at sowing) as well as the later measurements (ca. between 140 DAS and harvest) were consistent. The error might have developed during winter time, where unfortunately no measurements were made. As possible error sources remain that either (i) observed weather data was faulty (e.g. too low rain), or that the model overestimated soil water content by poor representation of at least one soil process or parameter. This could be (ii) underestimation of soil water evaporation, (iii) underestimation of drainage, or (iv) overestimation of soil water storage capacity (i.e. the difference between drained upper limit, *DUL*, and lower limit, *LL*). Regarding weather data, there is no indication that rainfall data was erroneous (i). Simulation of soil evaporation (ii) and drainage (iii) was good in EXP1 and the same parameters were used in EXP2. Also, the modification of the *WETWAT* parameter (rainfall threshold to return to stage I evaporation; see 8.3.16 Soil Water) already increased evaporation compared to the default *WETWAT*, but soil water of EXP2 was still overestimated. Further data would be required to assess whether iCrop's evaporation estimation is generally inappropriate for the environment of Eastern Austria. Regarding drainage (iii), the soil water storage capacity (iv) of the deeper soil layers in EXP2 was reduced (compared to EXP1) based on observations (see Figure 31). However, Figure 47 indicates that *DUL* was overestimated in the 30-60 cm soil layer, resulting in overestimation of the water storage capacity of this layer. As a result, water might have been stored in the soil in iCrop but actually drained in the real world. *DUL* overestimation was ca.  $0.1 \text{ mm}^3 \text{ mm}^{-3}$ , which is equivalent to 30 mm total soil water (layer thickness: 300 mm). The overestimation of total soil water (March until harvest) was ca. 50 mm, so the *DUL* overestimation of the 30-60 cm layer could explain most of it. However, drainage was not measured directly, and to avoid over-parameterisation I assumed identical soil parameters from 0-60 cm depth in both experiments. The assumption that sand content in deeper (60-120 cm) layers was

the main contributor to the soil differences between EXP1 and EXP2 appeared sensible at the time of parameterising the model. The fact that total soil organic carbon was almost identical between the experiments in the 0-60 cm layers and only differed clearly in deeper layers support this hypothesis (Table 7). In conclusion, the results indicated that relatively small changes ( $0.1 \text{ mm}^3 \text{ mm}^{-3}$ ) to one parameter (*DUL*) of a single soil layer can lead to a clear offset in total soil water estimates through underestimation of drainage. While this hypothesis would need additional experiments and data to be confirmed, I assume that iCrop's soil water simulation is sensitive to the soil parameters affecting extractable soil water (i.e. *LL* and *DUL*). Great care needs to be taken when parameterising iCrop for soil, but if this is done, precise and accurate soil water simulations can be achieved as shown for EXP1.

### 9.3.2. Phenology

The simulation of emergence was near perfect in both experiments (Table 8). However, emergence in EXP2 was calibrated by manually adjusting one soil water parameter relevant for emergence (see 9.2.6). One reason making this modification necessary could be related to the definition of "emergence" used for field observations. Forcella et al. (2000) criticised the approach of defining emergence as the date where at least 50% of the seedlings have emerged as possibly inappropriate in the context of modelling. They argue that emergence is not normally distributed around the 50% emergence date, and that emergence might take several weeks or even months, depending on the specific genotype and environment. The delayed simulation of emergence in EXP2 may be such an example, since observations showed continued increase of the emergence rate over a few weeks (data not shown). Continuous recordings of emergence rate (%) over several weeks would help to improve estimating the true date of crop emergence on the field.

The poor simulation of BBCH 21 (appearance of the first tiller) (Figure 32, Table 8) was clearly caused by interpolation errors resulting from long measurement intervals, particularly in EXP1. This effect was already discussed in 9.2.7. With more detailed observed data, the model can easily be parameterised to simulate BBCH 21 correctly. However, poor simulation of this stage does generally not affect any other process in the model. Especially leaf development variables (*LAI*, leaf number, leaf weight) in iCrop are independent from the initiation of this stage. Therefore, other simulation results in this study remain valid.

Leaf development in iCrop is affected by the simulated beginning of booting (BBCH 41) where the model stops leaf number growth on the main stem. This affects both *LAI* and leaf dry mass growth. BBCH 41 was simulated well in EXP1 (slightly too early, -2 to -3 days) but moderately delayed in EXP2 (+5 to +8 days) (Table 8). The effect of BBCH 41 simulation on leaf dry mass is

further discussed in 9.3.4 “Dry Mass”. For *LAI*, delayed simulation of BBCH 41 in EXP2 might have contributed to the overestimation, although maximum simulated *LAI* hardly differed between EXP1 and EXP2 and, therefore, suggested no influence (Figure 34).

Another critical phenological stage is anthesis (BBCH 61-69). Numerous studies have investigated anthesis simulation and pointed out its importance (e.g. Asseng et al., 2019, 1998; Eitzinger et al., 2013a; Liu et al., 2016; McMaster et al., 2008; O’Leary et al., 2015; Rötter et al., 2012; Stratonovitch and Semenov, 2015). iCrop simulated the beginning and end of anthesis well in both seasons (both stages within 2 days in EXP1 and within 5 days in EXP2).

iCrop initiates the beginning of seed growth (*BSG*, BBCH 71) five biological days after the beginning of anthesis (BBCH 61). The timely occurrence of *BSG* has a significant impact on seed growth by setting available vegetative crop dry mass for translocation to seeds and by influencing the seed growth duration. Also, leaf production fully terminates at *BSG* and leaf senescence is initiated (*LAI* starts to decline). *BSG* was simulated well in EXP1 (within 3 days) and also moderately well in EXP2 (within 5 days of observations) (Table 8).

Finally, BBCH 87 (termination of seed growth, *TSG*) defines the end of the seed growth period. In EXP1, *TSG* was simulated 5 days early for Bernstein but well for the other cultivars (within 1 day), and in EXP2 *TSG* was simulated 7 to 8 days too late for all cultivars (Table 8). Across all cultivars, both *BSG* and *TSG* were simulated too late in EXP2, resulting in an overall slight overestimation of seed growth duration (2 to 5 days) while EXP1 showed underestimation (-2 to -8 days) of seed-growth duration. While this generally matches the trends of N3 yield under- and overestimation of EXP1 and EXP2, respectively, the results are contradictory for the cultivars (Figure 37). For instance, the simulated slightly delayed *BSG* (+3 days) and clearly too early *TSG* (-5 days) of Bernstein in EXP1 resulted in a shortened seed-growth period which may have contributed to underestimation of Bernstein yield. Contrary, Aurelius’ seed-growth period in EXP1 was estimated well (*BSG* +1, *TSG* -1 day) but yield was still underestimated clearly, similar to Bernstein. It appeared that other factors than the simulated duration of the seed-growth period must have had a higher impact on variously accurate yield estimates (see 9.3.7 Yield). In a multi-model wheat study Asseng et al. (2015) found that phenology had a surprisingly low impact on wheat grain yields. In accordance with this, results from my simulations indicate a low model sensitivity to seed-growth duration regarding yield.

### 9.3.3. Leaf Development

While *LAI* simulations captured the general relative differences between the different N-treatments and cultivars, the most significant error was the overestimation of Arnold and Emilio at anthesis (PE = +52% and +58%, resp.), while Bernstein and Aurelius were simulated well

(PE = +17% and +20%, resp.) (Figure 34 for Arnold and Bernstein). In most N-treatments (particularly the higher fertilisations), observations showed an LAI peak at anthesis for Bernstein and Aurelius, while Arnold and Emilio did not show this peak as well as clearly lower LAI than Bernstein and Aurelius at anthesis. A similar trend as for LAI was found for observed leaf N content for Arnold and Bernstein (Figure 44) and also for Emilio and Aurelius (data not shown). While Bernstein reached anthesis last, the other cultivars (Arnold, Aurelius, Emilio) developed rather similarly (Figure 12), making comparisons between them more robust. With higher LAI and leaf N content at anthesis it can be assumed that Aurelius (and Bernstein) also maintained green leaves longer into the post-anthesis phase than Arnold and Emilio did. Such a prolonged maintenance of green leaves is known as stay-green trait (e.g. Thomas and Ougham, 2014). The different stay-green behaviours of the cultivars were not captured by the simulations. iCrop does not incorporate parameters to explicitly represent stay-green traits. The observed cultivar-specific stay-green behaviours could partly be explained by differences in the cultivars' root systems, leading to prolonged water access for stay-green types (Christopher et al., 2008; Manschadi et al., 2006). However, the only parameter affecting root system directly, *GRTDP*, was set equal for all cultivars (see also 9.2.4 Effect of Rooting Parameters on Yield). Therefore, in this parameterisation, iCrop was unable to simulate differences in stay-green behaviour based on root system differences. The cultivar-specifically parameterised differences in *LAI* development (*PLAPOW* and leaf fraction parameters, see 9.2.12) were apparently not sufficient to represent different stay-green traits.

The overall stronger *LAI* reduction effect of the unfertilised treatment NO in EXP2 than in EXP1 can be explained by lower initial soil mineral N in EXP2, as previously explained for above-ground dry mass and yield.

The model generally overestimated *LAI* at anthesis (RMSE = 0.8, NRSME = 32%, PE = +36%). Basso et al. (2016) reviewed 15 studies that reported on the *LAI* simulation performance of the CERES-Wheat model. These studies found *LAI* estimates with RMSE of 0.108, 0.069-0.075, 0.1-0.9, and 0.87 and RRMSE of 17.9%, 27.8%, 25-35%, 20%, 8%, and 1.27% using a range of treatments (including irrigation and fertilisation) and locations (India, China, USA, and others). Compared to these numbers, iCrop's *LAI* estimation performance ranged among the poorer results. However, iCrop showed an even higher overestimation of *LAI* before the parameterisation (data not shown). The thorough parameterisation in this study based on detailed and comprehensive parameterisation data resulted in substantial improvement of *LAI* estimates. Previous studies (e.g. Bassu et al., 2014; He et al., 2017; Manschadi et al., 2020a; Moeller et al., 2007; Salo et al., 2015; Wallach et al., 2011) have also shown that detailed data sets for parameterisation are required to improve model performance.

iCrop simulated relatively high final *LAI* (at harvest), particularly in EXP1 and in high N treatments, while observed *LAI* at harvest was always zero. This overestimation was caused by the simulation of the *LAI* senescence phase between anthesis and crop harvest. In iCrop, leaf senescence after the beginning of seed growth is directly linked to N retranslocated from leaves to grains. Under optimal conditions, the removal of N from leaves continues until *LAI* is reduced to almost zero, thereby finalising the seed growth duration. However, if the seed growth duration is limited by another factor, *LAI* as well as leaf N content can stay at higher levels. Such a factor was terminal drought stress due to low rainfall in EXP1. As a result, simulated phenology was enhanced and the termination of seed growth was reached a few days earlier than under optimal conditions. Thereby, some leaf area remained at harvest. An improved *LAI* senescence function which compensates for drought-induced leaf area senescence in the seed growth phase would be desirable for iCrop. On the other side, *LAI* overestimation at harvest is probably more of a cosmetic error since transpiration does not depend on *LAI* but only on dry matter production in iCrop, thereby having little effect besides possibly a slight dry mass overestimation post-anthesis.

#### 9.3.4. Dry Mass

Total above-ground dry mass (*TDM*) simulations captured the general relative differences between cultivars and N-fertilisation levels correctly (Figure 35). However, both these differences were underestimated. For instance, the observed difference of harvest *TDM* between Bernstein and Arnold, as well as between N0 (unfertilised) and N3 (210 kg N ha<sup>-1</sup>) was much larger than simulated. The previously described problems with simulating *LAI* (stay-green not simulated) and root growth (no differences between the cultivars) possibly contributed to the underestimation of *TDM* cultivar-differences. Whatever the reason was, *LAI* differences between the cultivars were underestimated (see previous sub-section). Since *LAI* is the main driver of dry matter accumulation in iCrop (see equations (19), (20), and (21)), the congruent underestimation of *TDM* cultivar-differences was a logical consequence. In addition, other growth-influencing parameters which were set equal for all cultivars (e.g. leaf extinction coefficient *KPAR*: direct influence on dry matter production, root growth *GRTDP*: indirect influence via soil water uptake and stress) have contributed to underestimation of cultivar-differences. The underestimation of N-treatment differences, particularly in EXP2, was likely related to soil Nmin simulations (discussed in 9.3.6 Crop and Soil Nitrogen).

Overall, the parameters that were chosen cultivar-specifically (see 9.2.12) were insufficient to adequately represent observed *TDM* cultivar differences. In order to improve cultivar-specific *TDM* simulation in iCrop, I suggest the investigation of stay-green simulation capability (see

previous sub-section) as well as cultivar-specific parameterisation of *KPAR* and *GRTDP*. However, in this study, data was not adequate to estimate these parameters (no data collection regarding *KPAR*, and regarding *GRTDP* see 9.2.4).

On the first glance, the figures showing simulated leaf and stem dry mass (Figure 38, Figure 39) appear to suggest that the leaf:stem partitioning parameters (*FLF*-parameters) were chosen wrong, with too much dry mass in leaves and too little in stems. However, this impression is deceiving. In the vegetative growth phase (emergence until anthesis), the partitioning of dry mass between leaves and stems is calculated based on the parameters *FLFIA*, *FLFIB*, and *FLF2* (“*FLF*”: fraction of dry mass partitioned to leaves) (see Figure 6). These parameters are used in consecutive order: *FLFIA* is used from emergence until the total crop dry mass reaches ca. 113-123 g m<sup>-2</sup> (defined by the cultivar-specific parameter *WTOPL*). Thereafter, *FLFIB* is used until leaf growth terminates on the main stem (which is set equal to booting, BBCH 41). Then, until anthesis, *FLF2* defines the fraction of dry mass that is assigned to leaves. Generally, as well as in the observations of my experiments, more dry mass is allocated to leaves in earlier phenological stages: *FLFIA* (0.72-0.80) > *FLFIB* (0.29-0.37) > *FLF2* (0.1) (see Table 5).

Figure 40 shows the development of stem and leaf dry mass for Bernstein (which is representative for all cultivars), including indicators (blue lines) for the switches from one to another leaf partitioning parameter. Until the right-most blue line in Figure 40, simulated total dry mass comprised only of leaf and stem dry mass (since seed growth started thereafter). Obviously, *TDM* in N0 was overestimated, which lead to overestimation of stem and leaf dry matter. Still, the course of the observed relation of stem to leaf dry matter in N0 was represented well: Leaf dry mass was higher than stem dry mass until ca. 190 DAS after which this relation swapped. N3 showed a very similar trend before anthesis (the second last observed points). From anthesis onwards, N3 stem dry mass was underestimated, while leaf dry mass was overestimated.

Earlier simulation of booting (BBCH 41) would come with earlier initiation of the *FLF2* parameter (the middle blue line in Figure 40). This would result in higher dry mass partitioning to stems (because *FLF2* allocates more dry mass to stems than the previous *FLFIA* does) and, hence, improvement of the simulations at anthesis for N3 as well as the relation of leaf:stem dry matter at anthesis in N0.

The *FLF2* parameter was set identical for all cultivars and could, therefore, not induce cultivar-specific differences. Most treatments showed an increase in leaf dry mass during the *FLF2* period (between second and third blue line in Figure 40, and corresponding observation points in Figure 38). However, *FLF2* might have slightly contributed to leaf dry mass overestimation (too high *FLF2*).

In conclusion, care needs to be taken when parameterising phenology, especially the occurrence of booting (BBCH 41), in order to achieve correct leaf and stem dry mass simulations. However, there are no negative follow-up effects of poor leaf and stem dry mass estimates on other iCrop outputs. Simulated *LAI* is largely independent from leaf dry mass, and the translocation of vegetative dry mass to the seeds during seed growth is based on the sum of stem and leaf dry mass (i.e. vegetative dry mass) rather than a separate translocation from each organ. Still, improved partitioning simulation would be desirable. Additional data on phenology around the switch from *FLFIB* to *FLF2* (BBCH 41) can help to support a more precise parameterisation.

#### 9.3.5. Effect of Development-dependent SNCS

I added to the crop model source code to implement two phases of stem minimum nitrogen concentration (*SNCS*) depending on the development stage (see 8.2 and 9.2.9). Before the update, iCrop used only one *SNCS* parameter and was unable to simulate observed dynamics of stem nitrogen uptake for both parameterisation experiments (data not shown). Stem nitrogen uptake was underestimated before anthesis, especially for low N input treatments (N0) because the model prioritises N distribution to leaves when N supply is limited. As a result, stem N content was underestimated, leaf N content overestimated (no N-stress induced), and, thus, *LAI* growth overestimated. This also led to total dry mass overestimation. With the introduction of two phenology-dependent *SNCS* parameters, this problem was solved. Based on observed data, the two *SNCS* parameters were calculated, resulting in higher first-phase stem minimum N concentration (*SNCSI*, 0.0063 g g<sup>-1</sup>) than second-phase (*SNCS2*, 0.0022 g g<sup>-1</sup>) (Table 5). The relatively high *SNCSI* forced the model to distribute more N to stems before anthesis in N0 treatments, even though N supply was limited (Figure 45). This led to reduced N uptake in leaves of N0 (Figure 44). As a result, *LAI* growth before anthesis was limited in N0 (N-stress). As in the observations, clearly lower maximum *LAI* was simulated in N0 than in the other N-treatments (Figure 34). Manschadi et al. (2020a) also showed the necessity of two phases of *SNCS* to properly simulate *LAI* for maize. Overall, the described “chain reaction” from stem N concentration parameterisation until total dry mass estimation showed nicely how complex the model parameterisation process can become and also that process based models such as iCrop can adequately reflect important crop physiological mechanisms based on a sound parameterisation.

#### 9.3.6. Crop and Soil Nitrogen

Total above-ground crop N-uptake (*TNU*) at harvest was overestimated by iCrop. Simulations were good until ca. 200 DAS (ca. BBCH 33, stem elongation) but started to exceed observations thereafter (N2/N3 in EXP1 and all N-treatments in EXP2, see Figure 42). Overestimation of crop dry mass (*TDM*) can lead to overestimation of *TNU*, since N-uptake from the soil is driven by

dry mass production. This was the case for Arnold EXP2 (N0-N2) where both *TDM* and *TNU* were overestimated (Figure 35 and Figure 42). However, N0 *TDM* of Arnold EXP1 was overestimated, while *TNU* was simulated well. Also, N3 *TDM* of Bernstein EXP1 was underestimated, but *TNU* was overestimated. Still, on average, *TDM* overestimation appeared being the reasonable cause for *TNU* overestimation at harvest (*TDM* PE: +13.7%, *TNU* PE: +14.0%). However, the contrary examples above as well as the comparison of PEs at anthesis (*TDM*: +6.0%, *TNU*: +14.0%) indicate that there were also other factors influencing crop N-uptake overestimation.

One important of those might be related to soil-N supply. Unfortunately, the data for soil Nmin contained only relatively few samples, with a maximum of three in-season measurements (sowing, tillering, harvest) for specific treatments only (Figure 48). However, the data was still sufficient to indicate overestimation of soil N mineralisation in the model: In EXP2 (Arnold, Bernstein), all three observed sampling dates of soil Nmin were simulated well, but final crop *TNU* was overestimated. Both observed and simulated Nmin values were extremely low at harvest (below 5 kg N ha<sup>-1</sup>) but *TNU* overestimation was approximately 50 kg N ha<sup>-1</sup>. This strongly suggests that a significant amount of soil Nmin must have either (i) been lost on the field without being accounted for in the model, or (ii) been overestimated by the model.

Unaccounted field Nmin losses (i) might occur due to N-leaching. iCrop calculates N-leaching based on soil water leaching (i.e. drainage from the deepest soil layer). In both experiments, significant rainfall amounts occurred in May/June but simulated water drainage during that time was very low. However, in EXP2 simulated soil water reached field capacity in May/June (Figure 46). At that point, each additional water input in the model results in more or less direct drainage loss, making drainage very sensitive to slight errors in rain and soil water parameters. Therefore, it seems possible that iCrop underestimated drainage (particularly in EXP2), thereby underestimating N-leaching, resulting in overestimating N-uptake.

Another possible reason for model-unaccounted field Nmin losses (i) is N-volatilisation from N fertiliser (surface application). iCrop accounts for N-volatilisation using a fixed user-defined percentage value which was set to 2%. It is possible that this value was higher in reality. Meisinger and Randall (1991) suggested 2-25% N losses for ammonium nitrate with surface application in subhumid areas (defined as areas with a likely precipitation of 0-6 mm within 7 days of fertiliser application). In a 2-year wheat field experiment in China, Yang et al. (2015) reported N volatilisation of 2.49 kg N ha<sup>-1</sup> for 180 kg N ha<sup>-1</sup> applied fertiliser (i.e. ca. 1.4% loss) in seasons with very low precipitation (ca. 110 and 190 mm in each season). In comparison to these studies, my assumption of 2% N volatilisation was low but reasonable. However, higher volatilisation would be similarly possible, and its underestimation might have contributed to *TNU* overestimation.

For instance, assuming average (12.5%) and maximum (25%) N volatilisation from Meisinger and Randall (1991) for N3 (210 kg N ha<sup>-1</sup> fertilisation) of my experiments, total N loss from volatilisation would be 26.25 and 52.5 kg N ha<sup>-1</sup>, respectively, while the parameterised value of 2% resulted only in 4.2 kg N ha<sup>-1</sup> in both years. The difference (22.05 and 48.3 kg N ha<sup>-1</sup>) might have been lost in reality, but was not accounted for in iCrop. Compared to the RMSE for *TNU* simulation at harvest (3.2 g N m<sup>-2</sup>; i.e. 32 kg N ha<sup>-1</sup>), N volatilisation has the potential to explain much of the simulation error.

Finally, iCrop might have overestimated (ii) soil N mineralisation. The corresponding parameter, *FMIN*, was calibrated for each soil layer based on observed soil N<sub>min</sub> data from the parameterisation field experiments as well as additional data from previous experiments, including a fallow during winter time (data not shown). The fallow experiment was conducted on a field in ca. 2 km distance to the locations of the parameterisation field experiments. According to the soil map eBod (Wandl and Horvath, n.d.) the soil type was the same as in EXP2 (“Feuchtschwarzerde”, calcareic gleyic phaeozem) but different to EXPI. Therefore and due to the distance, it might not translate well to the parameterisation experiments. Additionally, temperature has a significant influence on soil N mineralisation in the field as well as in iCrop. While N<sub>min</sub> on a fallow field during winter time with cold temperatures might be simulated well, there is no guarantee that N<sub>min</sub> dynamics for spring and summer environments with higher temperatures is simulated equally well. Also, soil N mineralisation is related to soil water. Overestimation of soil water content, as previously described, can have contributed to N mineralisation overestimation.

In conclusion, more detailed soil water and soil N<sub>min</sub> data would be required to identify the reasons for *TNU* overestimation. Ideally, this would include multiple years of fallow at several locations, including varying N fertiliser application (for N-volatilisation estimation) and several soil N<sub>min</sub> and water measurements, with shorter measurement intervals during warmer temperatures, around N fertilisation dates, and during/after intense rainfalls (drainage).

The previously mentioned possible errors resulting from early biomass sampling for parameter calculation at beginning seed growth (see 9.2.11) did not occur. Assuming the stem N concentration parameters (*SNCG* and *SNCSI*) were overestimated, this would lead to stem N uptake overestimation in highly fertilised treatments (N2, N3) due to the *SNCG* parameter. Also, stem N uptake for unfertilised treatments (N0) would be overestimated until anthesis (*SNCSI* parameter). However, as shown in Figure 45 for Arnold and Bernstein, stem N uptake was rather underestimated at anthesis. The early overestimation of stem N content was caused by stem N dry mass overestimation at that time (Figure 39). Final stem N content overestimation for the highly fertilised treatments was part of the total N uptake overestimation, which was related to several

possible causes as discussed above (overestimation of soil-N supply) and below (underestimation of N-retranslocation). For leaf N uptake, a possible underestimation of *SLNG* due to too early biomass sampling would induce leaf N content underestimation as well. However, leaf N content estimates reflected observations well (Figure 44). Also, they largely corresponded to *LAI* simulations (Figure 34) which are the main driver of leaf N content in iCrop. Overall, the stem N as well as leaf N content simulations were satisfactory and did not indicate any biases in the parameter calculation.

The simulation of grain nitrogen concentration (%GN) was, on average, robust and satisfactory. However, using identical %GN parameters (%GN<sub>min</sub> and %GN<sub>max</sub>) for all cultivars, the model was not able to simulate the observed differences in %GN between the cultivars, mainly between Arnold (higher) and the three other cultivars (lower %GN) (Figure 43). The reasons for not parameterising %GN cultivar-specific are explained in sub-section 9.2.9 “Crop Nitrogen Parameters”. iCrop’s inability to simulate the high observed %GN of Arnold can be explained by wrongly simulated N-dilution in Arnold due to yield overestimation (Figure 37).

Regarding the effect of N-treatment, %GN of N0 were mostly underestimated. Since total N uptake at harvest (*TNU*, Figure 41) of N0 were mostly simulated well and yields were overestimated (esp. EXP2), the underestimation of %GN can be explained with N-dilution. In addition, N removal from vegetative organs (translocation to grains) was sometimes underestimated (mainly N2 and N3 of EXP1) (see leaf N and stem N content, Figure 44 and Figure 45, resp.). While this contributed to %GN underestimation of Arnold, it has partly “corrected” %GN overestimation of N2 and N3 of the other cultivars.

Overall, while %GN simulation was robust, it can be improved by addressing the following issues: (i) Improvement of yield estimates, thereby reducing errors from grain-N dilution, (ii) improvement of vegetative N translocation simulation, and possibly (iii) the use of cultivar-specific %GN parameters. However, before (iii) %GN parameters can be parameterised for each cultivar, additional data based on field experiments designed to generate a range of %GN contents must be acquired (see 9.2.9). Also, correct (i) yield simulation and (ii) N translocation must be ensured to avoid over-parameterisation of the model.

#### 9.3.7. Yield

While observed yields showed clear responses to N-fertilisation and cultivar, the simulations underestimated these responses (Figure 37). Especially the simulated yield differences between the cultivars were almost negligible and did not capture observations.

The most obvious simulation error occurred in EXP1 where Aurelius and Bernstein yields were clearly underestimated. In the simulations, total soil water contents of all cultivars reached very low levels (close to lower limit) during the grain filling phase (Figure 46). As a result, terminal water stress occurred in the model and caused both total dry mass and yield to decline for all cultivars almost equally. However, in reality, terminal water stress was apparently not as severe as simulated. A range of problems can have caused the overestimation of soil water stress, including non-fitting soil parameters (lower limit) and other reasons, as discussed in 9.3.1 and 9.3.4.

Since the parameters which may affect yield directly (*FRTRL* and *PDHI*) were identical for the cultivars, simulated yield differences could only have come from other, cultivar-specific parameters (these were parameters affecting *LAI*, leaf:stem partitioning, and phenology; see 9.2.12). While the model was able to capture at least the general trends of cultivar differences correctly for simulating *MSNN*, *LAI*, and leaf and stem dry mass (see Figure 33, Figure 34, Figure 38, and Figure 39, resp.), these differences did not translate into correct yield simulation. Therefore, iCrop's inability to capture the observed yield differences between the cultivars based solely on cultivar-specific parameters for *LAI*, leaf:stem partitioning, and phenology suggests that yield parameters (*FRTRL* and *PDHI*) also require cultivar-specific determination for correct yield simulation. Furthermore, other parameters which affect yield indirectly, such as the previously discussed root growth (*GRTDP*) and grain nitrogen concentrations (*%GNmin*, *%GNmax*), should also be considered candidates for cultivar-specific parameterisation. The impact of seed growth duration (BSG to TSG) on yield was discussed previously in sub-section 9.3.2.

#### 9.4. Crop Model Evaluation

The iCrop parameterisation obtained from the field experiments at Tulln (EXP1 and EXP2) was evaluated against an independent long-term data set of field measurements in Eastern Austria.

Average growing conditions during the long-term experiments were overall poorer than in the parameterisation experiments (10-41% less rain, slightly less global radiation). Also, slightly lower average temperatures were observed in the evaluation seasons. These might be both advantageous (less heat stress in summer) and disadvantageous (more frost damage during winter).

The date of heading was simulated well (RMSE = 2.9 days, Figure 49). This indicates that iCrop is robust regarding phenology simulation. Heading date (BBCH 59) is very close to the more commonly observed anthesis (BBCH 65) and, therefore, compares well to other studies showing anthesis date. In a multi-model comparison using 9 crop models, 44 growing seasons at 7 sites, simulating spring barley and winter wheat, Rötter et al. (2012) reported higher RMSE for the

mean model estimates of anthesis (for both crops), ranging from ca. 6 to 14 days. Soltani et al. (2013) evaluated the iCrop wheat model with independent field data from the Gorgan region (Iran) and found anthesis simulation with RMSE of 4.8 days, thereby underlining the quality and accuracy of my results.

Yields of the evaluation data were mostly underestimated (Figure 50 and Figure 51). The RMSE of  $136 \text{ g m}^{-2}$  was much greater than the RMSE found by Soltani et al. (2013) in their iCrop evaluation study (37.7). The reason for this poor estimate was probably related to the lack of initial soil data (water and Nmin), lack of data on soil properties of the different locations (such as lower limit and field capacity), and large distances to weather stations which possibly caused discrepancies between actual field rainfall and weather station rainfall data. Since identical soil parameterisations were used for all locations (see 8.5.2), possible soil differences were not reflected in the evaluation simulations, likely resulting in simulation errors. An indication for poor rainfall data is the relation of yield to evapotranspiration (ET). Simulated yield responded much stronger to simulated evapotranspiration than observed yield (Figure 52). Due to lack of observed ET data I used simulated ET for both simulated and observed yield for the comparison in Figure 52. While this approach clearly does not give the best results, it still provides a reference to assess water use. Simulated ET incorporates interactions with observed rainfall and irrigation and, therefore, reflects observed seasonal differences in water input. However, simulated ET also accounts for effects of simulated crop dry mass (which drives transpiration in iCrop) which can introduce bias when dry mass simulations deviate from observations. Therefore, the following discussion based on simulated ET serves as indication only. Observed rainfall data was taken from weather stations with ca. 10-25 km distance to the experimental fields (Figure 4). Kersebaum et al. (2015) defined guidelines to assess suitability of data for crop modelling. They noted that the relevance of rainfall for a field experiment data drops sharply with distances from 1 to 10 km, being hardly relevant thereafter. Therefore, it can be assumed that the distance surely introduced deviations between actual rainfall at the experimental sites and measured rainfall at the weather stations. However, this probably explains only partly why the estimates for yield were poor. Some extreme points in Figure 52, particularly very high yields ( $700$  to  $1000 \text{ g m}^{-2}$ ) with lowest ET (less than  $400 \text{ mm}$ ) appear unlikely. Almost all points (both observed and simulated) are above the  $1 \text{ g m}^{-2} \text{ mm}^{-1}$  water use efficiency (WUE) threshold line. The simulated yields ranged mainly between WUE 1 and 1.5 (except at high ET values), while the observed yields ranged mainly between 1.5 and 2. Some of the extreme points even exceeded WUE 2 clearly. Moeller et al. (2007) reported WUE for wheat ranging from 0.19-0.98  $\text{g m}^{-2} \text{ mm}^{-1}$ , simulated with APSIM in north-western Syria (annual precipitation:  $340 \text{ mm}$ , mean temperature:  $17.6 \text{ }^\circ\text{C}$ ).

Sadras and Angus (2006) compared data from rain-fed dry environments for wheat (South-Eastern Australia, North American Great Plains, China Loess Plateau, and the Mediterranean Basin). They found a maximum WUE of  $2.2 \text{ g m}^{-2} \text{ mm}^{-1}$  and average WUE of ca. 0.9 to 1. In comparison, the results of my study appear rather high, although parts of that may be explained by higher seasonal rainfall in Eastern Austria. Nevertheless, the different yield responses to ET between simulations and observations indicate a systematic error that could have several reasons. Missing information about initial soil water conditions may have caused underestimation of plant available soil water by choosing too low initial soil water for the simulations. Also, the data set might miss information on applied irrigation. As mentioned above, the distance between weather stations and field locations may have introduced errors. iCrop might systematically underestimate yields in dry environments, although this seems unlikely since iCrop was developed based on data from Gorgan, Iran (Soltani et al., 2013), which is a relatively dry region (December-June: 340 mm average rainfall) (Soltani and Sinclair, 2015). Overall, due to insufficient soil data, the evaluation simulations do neither confirm nor reject the hypothesis of iCrop's capability to simulate cultivar-specific wheat yields in Eastern Austria.

#### 9.5. Application of the iCrop Model

Bringing scientific knowledge into practise often poses a big challenge. Scientists sometimes struggle to effectively transport their findings to a broader, non-expert audience. In agriculture, it seems particularly difficult to generate scientifically-backed information that exceeds the vast practical and intuitive knowledge of farmers, consultants, and other stakeholders. However, crop models have been applied successfully in various scientific and practical fields, including agronomic management, precision agriculture, assessment of environmental impacts, plant breeding, management of climate variability and seasonal forecasting, policy in agriculture, and impacts and adaptations to climate change (Asseng et al., 2013; Basso et al., 2016; Chenu et al., 2017; Reynolds et al., 2018). Hochman et al. (2009) tackled the problem of bringing the advantages of crop models to farmers actively, which resulted in the YieldProphet<sup>®</sup> software tool which is a good example for successful model application in practical farming. Also, this PhD study was conducted within the scope of a “from science to product” project, funded by the Austrian FFG (Forschungsförderungsgesellschaft). The project with the title “ICT for Decision Making in Farming” (Farm/IT; <https://www.farmit.at/>) (Manschadi et al., 2020b, 2019) was tailored to provide web-based tools to stakeholders in the agricultural sector for enhancing informed tactical and strategic decisions. By integrating a range of information and data, including weather, crop, soil, and satellite data, the Farm/IT software is capable of accurately simulating farm systems under changing climatic and environmental conditions. The targeted user-group covers farmers, consultants, agricultural businesses, and government agencies. The main scope

of the Farm/IT project was offering online software tools to forecast crop yield and harvest date, optimising nitrogen fertilisation based on crop modelling and remote sensing, optimising grassland yield and forage quality, optimising irrigation schemes via remote sensing, optimisation and estimation of ecological footprints, and optimisation of crop rotations to improve resource use (Manschadi et al., 2019).

In Farm/IT, the forecast of seasonal crop yield as well as the nitrogen fertilisation demand was identified as most relevant to policy makers and farmers. To address these, the process-based crop growth model iCrop was used in the project. Besides the wheat parameterisation presented here, iCrop is also currently being parameterised and evaluated for maize, potato, and sugar beet. Preliminary results were promising, showing that iCrop can simulate crop development, growth, and yield adequately in response to management (including irrigation, N fertilisation) and weather. However, similarly to my results, one of the biggest challenges for adapting the iCrop model to crops grown and conditions in Austria is the generation and/or acquisition of high-quality weather, soil, and crop data (Manschadi et al., 2019).

The findings of my study presented here have substantially supported the improvement and adaptation of the iCrop model to Eastern Austrian conditions and cultivars. Strengths as well as weak spots in the model which need more thorough parameterisation and data were identified and additional functionality was implemented in the model source code to support correct simulation under critical conditions such as limited N supply. As a result, with additional parameterisation work, the iCrop model can deliver sound estimates for crop yield, harvest date, and N fertilisation demand within the Farm/IT framework. Using established local institutions such as the Maschinenring (an organisation for making agricultural equipment available to farmers at low costs via sharing), Farm/IT aims at delivering scientific and technological advances directly to farmers, thereby giving the outcomes of my study an immediate and relevant value.

## 10. Conclusions

The capabilities and limitations of the iCrop growth model were outlined thoroughly in this work. Based on two seasons of detailed winter wheat field experimentations, iCrop was parameterised for the conditions of Eastern Austria. Focus was set on leaf development, soil and crop nitrogen, phenology, and crop dry mass. Simulation of long-term, independent field experimental data from across Eastern Austria provided a sound basis to evaluate the phenological performance of the model, while evaluation of yield estimates was not possible due to the lack of comprehensive observed soil data.

The overall research question was “Is the iCrop crop growth model capable of capturing winter wheat canopy growth and development in the environment of Eastern Austria using field data for detailed model parameterisation?” (see chapter 0). Based on the results presented, this can be generally answered with yes, although with limitations.

The iCrop model showed its capability to simulate winter wheat growth and development in Eastern Austria regarding several aspects. The simulation of soil water content was shown to be excellent with good parameterisation (Figure 46), while the model appeared sensitive to soil parameters when water content was close to field capacity. Regarding soil N<sub>min</sub>, data scarcity made it difficult to assess iCrop’s abilities. However, overestimation of final crop N uptake (Figure 42) clearly indicated a systematic overestimation of the course of soil N<sub>min</sub> by overestimating N mineralisation and/or underestimating N losses. This was likely a matter of soil parameters rather than a model-intrinsic problem. Overall, iCrop clearly showed its ability to simulate soil water content, while for soil N<sub>min</sub> further investigation is necessary (hypothesis 1).

iCrop impressively demonstrated its capability to deliver accurate and precise cultivar-specific phenology estimates (hypothesis 2) for the two-year parameterisation field experiments (Figure 32) as well as for the long-term independent evaluation data set (Figure 49). Only a few development stages were predicted with errors greater than 5 days which can easily be fixed with more detailed parameterisation data.

Hypothesis 3 presumed cultivar-specific and N-fertilisation-specific dry mass and crop-N simulation capabilities at organ-level. This was, overall, achieved with some limitations. While N-specific differences for total crop dry mass (Figure 35) and yield (Figure 37) were roughly captured, iCrop could not represent cultivar differences or absolute values. This may be linked to some weaknesses in simulation of leaf development and root growth. Regarding leaf development, iCrop was unable to capture clear differences in leaf area index between the cultivars (Fig-

ure 34), possibly due to different stay-green behaviours of the cultivars which were not represented by the model. However, N-specific differences in leaf area were captured well after the implementation of a biphasic stem nitrogen parameter, triggering a cascade of improvements. Concerning root growth simulation, the lack of soil parameters representing root architectural differences (e.g. root exploration factor) between the cultivars might have prevented the model from capturing cultivar and N-fertilisation-specific differences. In regard to yield, cultivar-specific parameterisation of the directly influencing yield formation parameters might be necessary to improve estimates.

Leaf and stem dry mass estimates were biased (Figure 38 and Figure 39). Only the relative differences between the cultivars were captured but not absolute values, and the differences between N-treatments were clearly underestimated. Generally, leaf dry mass was overestimated and stem dry mass underestimated. Improvements of phenology simulation can potentially solve this problem by improving the timing of the activation of leaf:stem partitioning parameters. However, since leaf area increase is decoupled from leaf dry mass growth in iCrop, poor leaf and stem dry mass estimates did not affect leaf area and yield simulations.

Total crop nitrogen simulations were good during the first half of the season but were systematically overestimated towards the end of the season (Figure 42). This was linked to soil N<sub>min</sub> overestimation, as described above. Fixing soil N<sub>min</sub> would likely also improve crop nitrogen simulations.

Grain nitrogen content simulation was overall good, but cultivar-differences were not captured (Figure 43) as a follow-up effect from inaccurate yield estimates resulting in biased grain N dilution. Leaf and stem N contents (Figure 44 and Figure 45) were simulated well, with a tendency to overestimation at harvest, possibly resulting from overall overestimation of crop nitrogen and LAI at harvest.

Numerous previous studies have demonstrated that crop models have reached a stage where they can be used to address a range of important issues such as climate change impact assessment, support of crop breeding, or informed decision making for e.g. optimised fertiliser use. The parameterisation and application of crop models in the important cereal production region in Eastern Austria has been scarce so far (e.g. Ebrahimi et al., 2016; Eitzinger et al., 2013a, 2013b). With this PhD thesis, a relatively simple model has been prepared to be applied in this particular region. Also, my results support the previous finding by Soltani and Sinclair (2015) that simple models with relatively few parameters can deliver accurate simulation results based on a sound parameterisation.

Beyond presenting iCrop's suitability for simulating cultivar-specific growth and development of winter wheat in Eastern Austria, this study also provides information on how to enhance the model's accuracy and precision by giving details on its weak spots and ideas how to address them. Furthermore, with the innovative Farm/IT project (<https://farmit.at/>, Manschadi et al., 2020b), the results of this study were directly put to use through modern software tools aiming at delivering scientific advances to farmers and stakeholders in agriculture.

## 11. References

- AGES, 2020. Österreichische Beschreibende Sortenliste 2020 Landwirtschaftliche Pflanzenarten (No. Schriftenreihe 21/2020, ISSN 1560-635X).
- Asseng, S., Ewert, F., Martre, P., Rötter, R.P., Lobell, D.B., Cammarano, D., Kimball, B.A., Ottman, M.J., Wall, G.W., White, J.W., Reynolds, M.P., Alderman, P.D., Prasad, P.V.V., Aggarwal, P.K., Anothai, J., Basso, B., Biernath, C., Challinor, A.J., De Sanctis, G., Doltra, J., Fereres, E., Garcia-vila, M., Gayler, S., Hoogenboom, G., Hunt, L.A., Izaurrealde, R.C., Jabloun, M., Jones, C.D., Kersebaum, K.C., Koehler, A., Müller, C., Naresh Kumar, S., Nendel, C., O'leary, G., Olesen, J.E., Palosuo, T., Priesack, E., Eyshi Rezaei, E., Ruane, A.C., Semenov, M.A., Shcherbak, I., Stöckle, C., Stratonovitch, P., Streck, T., Supit, I., Tao, F., Thorburn, P.J., Waha, K., Wang, E., Wallach, D., Wolf, J., Zhao, Z., Zhu, Y., 2015. Rising temperatures reduce global wheat production. *Nature Climate Change*; London 5, 143–147. <http://dx.doi.org/10.1038/nclimate2470>
- Asseng, S., Ewert, F., Rosenzweig, C., Jones, J.W., Hatfield, J.L., Ruane, A.C., Boote, K.J., Thorburn, P.J., Rotter, R.P., Cammarano, D., Brisson, N., Basso, B., Martre, P., Aggarwal, P.K., Angulo, C., Bertuzzi, P., Biernath, C., Challinor, A.J., Doltra, J., Gayler, S., Goldberg, R., Grant, R., Heng, L., Hooker, J., Hunt, L.A., Ingwersen, J., Izaurrealde, R.C., Kersebaum, K.C., Muller, C., Naresh Kumar, S., Nendel, C., O'Leary, G., Olesen, J.E., Osborne, T.M., Palosuo, T., Priesack, E., Ripoche, D., Semenov, M.A., Shcherbak, I., Steduto, P., Stockle, C., Stratonovitch, P., Streck, T., Supit, I., Tao, F., Travasso, M., Waha, K., Wallach, D., White, J.W., Williams, J.R., Wolf, J., 2013. Uncertainty in simulating wheat yields under climate change. *Nature Clim. Change* 3, 827–832.
- Asseng, S., Keating, B.A., Fillery, I.R.P., Gregory, P.J., Bowden, J.W., Turner, N.C., Palta, J.A., Abrecht, D.G., 1998. Performance of the APSIM-wheat model in Western Australia. *Field Crops Research* 57, 163–179. [https://doi.org/10.1016/S0378-4290\(97\)00117-2](https://doi.org/10.1016/S0378-4290(97)00117-2)
- Asseng, S., Martre, P., Ewert, F., 2020. Chapter 16: Crop simulation model inter-comparison and improvement, in: *Advances in Crop Modelling for a Sustainable Agriculture*. Burleigh Dodds Science Publishing, Cambridge.
- Asseng, S., Martre, P., Maiorano, A., Rötter, R.P., O'Leary, G.J., Fitzgerald, G.J., Girousse, C., Motzo, R., Giunta, F., Babar, M.A., Reynolds, M.P., Kheir, A.M.S., Thorburn, P.J., Waha, K., Ruane, A.C., Aggarwal, P.K., Ahmed, M., Balkovič, J., Basso, B., Biernath, C., Bindi, M., Cammarano, D., Challinor, A.J., Sanctis, G.D., Dumont, B., Rezaei, E.E., Fereres, E., Ferrise, R., Garcia-Vila, M., Gayler, S., Gao, Y., Horan, H., Hoogenboom, G., Izaurrealde, R.C., Jabloun, M., Jones, C.D., Kassie, B.T., Kersebaum, K.-C., Klein, C., Koehler, A.-K., Liu, B., Minoli, S., Martin, M.M.S., Müller, C., Kumar, S.N., Nendel, C., Olesen, J.E., Palosuo, T., Porter, J.R., Priesack, E., Ripoche, D., Semenov, M.A., Stöckle, C., Stratonovitch, P., Streck, T., Supit, I., Tao, F., Velde, M.V. der, Wallach, D., Wang, E., Webber, H., Wolf, J., Xiao, L., Zhang, Z., Zhao, Z., Zhu, Y., Ewert, F., 2019. Climate change impact and adaptation for wheat protein. *Global Change Biology* 25, 155–173. <https://doi.org/10.1111/gcb.14481>
- Barker, A.V., Pilbeam, D.J., 2015. Chapter 2: Nitrogen, in: *Handbook of Plant Nutrition*. CRC Press, Boca Raton, p. 773.
- Basso, B., Cammarano, D., Troccoli, A., Chen, D., Ritchie, J.T., 2010. Long-term wheat response to nitrogen in a rainfed Mediterranean environment: Field data and simulation analysis. *European Journal of Agronomy* 33, 132–138. <https://doi.org/10.1016/j.eja.2010.04.004>
- Basso, B., Dumont, B., Maestrini, B., Shcherbak, I., Robertson, G.P., Porter, J.R., Smith, P., Paustian, K., Grace, P.R., Asseng, S., Bassu, S., Biernath, C., Boote, K.J., Cammarano, D., De Sanctis, G., Durand, J.-L., Ewert, F., Gayler, S., Hyndman, D.W., Kent, J., Martre, P., Nendel, C., Priesack, E., Ripoche, D., Ruane, A.C., Sharp, J., Thorburn, P.J., Hatfield, J.L., Jones, J.W., Rosenzweig, C., 2018. Soil Organic Carbon and Nitrogen Feedbacks on Crop Yields under Climate Change. *Agricultural & Environmental Letters* 3, 180026. <https://doi.org/10.2134/ael2018.05.0026>

- Basso, B., Liu, L., Ritchie, J.T., 2016. A Comprehensive Review of the CERES-Wheat, -Maize and -Rice Models' Performances, in: Sparks, D.L. (Ed.), *Advances in Agronomy*, Advances in Agronomy. Academic Press, pp. 27–132. <https://doi.org/10.1016/bs.agron.2015.11.004>
- Bassu, S., Brisson, N., Durand, J.-L., Boote, K., Lizaso, J., Jones, J.W., Rosenzweig, C., Ruane, A.C., Adam, M., Baron, C., Basso, B., Biernath, C., Boogaard, H., Conijn, S., Corbeels, M., Deryng, D., De Sanctis, G., Gayler, S., Grassini, P., Hatfield, J., Hoek, S., Izaurrealde, C., Jongschaap, R., Kemanian, A.R., Kersebaum, K.C., Kim, S.-H., Kumar, N.S., Makowski, D., Müller, C., Nendel, C., Priesack, E., Pravia, M.V., Sau, F., Shcherbak, I., Tao, F., Teixeira, E., Timlin, D., Waha, K., 2014. How do various maize crop models vary in their responses to climate change factors? *Global Change Biology* 20, 2301–2320. <https://doi.org/10.1111/gcb.12520>
- Baumgarten, A., Berthold, H., Buchgraber, K., Dersch, G., Egger, H., Egger, R., Eigner, H., Frank, P., Gerzabek, M., Hölzl, F.X., Holzner, H., Janko, M., Pernkopf, G., Peszt, W., Pfundtner, E., Pötsch, E.M., Rohrer, G., Schilling, C., Spanischberger, A., Spiegel, H., Springer, J., Strauss, P., Winkowitsch, C., Zehner, G., 2017. *Richtlinien für die Sachgerechte Düngung* 7. Auflage.
- Bayerisches Staatsministerium für Ernährung, Landwirtschaft und Forsten (StMELF), 2018. Bestandteile von Humus und Humusqualität [WWW Document]. URL <https://www.lfl.bayern.de/iab/boden/031122/index.php>
- Birch, C.J., Vos, J., van der Putten, P.E.L., 2003. Plant development and leaf area production in contrasting cultivars of maize grown in a cool temperate environment in the field. *European Journal of Agronomy* 19, 173–188. [https://doi.org/10.1016/S1161-0301\(02\)00034-5](https://doi.org/10.1016/S1161-0301(02)00034-5)
- BMLRT, 2021. Getreideanbau und Getreidearten in Österreich, [bmlrt.gv.at](https://www.bmlrt.gv.at) [WWW Document]. URL <https://www.bmlrt.gv.at/land/produktion-maerkte/pflanzliche-produktion/getreide/Getreide.html> (accessed 1.26.21).
- Bodner, G., Alsalem, M., Nakhforoosh, A., 2021. Root System Phenotyping of Soil-Grown Plants via RGB and Hyperspectral Imaging, in: Tripodi, P. (Ed.), *Crop Breeding: Genetic Improvement Methods*, Methods in Molecular Biology. Springer US, New York, NY, pp. 245–268. [https://doi.org/10.1007/978-1-0716-1201-9\\_17](https://doi.org/10.1007/978-1-0716-1201-9_17)
- Bodner, G., Nakhforoosh, A., Arnold, T., Leitner, D., 2018. Hyperspectral imaging: a novel approach for plant root phenotyping. *Plant Methods* 14, 84. <https://doi.org/10.1186/s13007-018-0352-1>
- Boguslawski, E. v., 1958. Das Ertragsgesetz. Die Mineralische Ernährung der Pflanze / Mineral Nutrition of Plants 943–976. [https://doi.org/10.1007/978-3-642-94729-2\\_30](https://doi.org/10.1007/978-3-642-94729-2_30)
- Boote, K.J., 2020. Chapter 17: The future of crop modeling for sustainable agriculture, in: *Advances in Crop Modelling for a Sustainable Agriculture*. Burleigh Dodds Science Publishing, Cambridge.
- Boote, K.J., Jones, J.W., Pickering, N.B., 1996. Potential Uses and Limitations of Crop Models. *Agronomy Journal* 88, 704–716. <https://doi.org/10.2134/agronj1996.00021962008800050005x>
- Bouman, B.A.M., van Keulen, H., van Laar, H.H., Rabbinge, R., 1996. The 'School of de Wit' crop growth simulation models: A pedigree and historical overview. *Agricultural Systems* 52, 171–198. [https://doi.org/10.1016/0308-521X\(96\)00011-X](https://doi.org/10.1016/0308-521X(96)00011-X)
- Butler, J.H., Montzka, S.A., 2020. The NOAA Annual Greenhouse Gas Index (AGGI) [WWW Document]. NOAA/ESRL Global Monitoring Laboratory - THE NOAA ANNUAL GREENHOUSE GAS INDEX (AGGI). URL <https://www.esrl.noaa.gov/gmd/aggi/aggi.html> (accessed 6.8.20).
- Campbell, M.D., Campbell, G.S., Kunkel, R., Papendick, R.I., 1976. A model describing soil-plant-water relations for potatoes. *American Potato Journal* 53, 431–442. <https://doi.org/10.1007/BF02852657>

- Cao, H., Hanan, J.S., Liu, Yan, Liu, Yong-xia, Yue, Y., Zhu, D., Lu, J., Sun, J., Shi, C., Ge, D., Wei, X., Yao, A., Tian, P., Bao, T., 2012. Comparison of Crop Model Validation Methods. *Journal of Integrative Agriculture* 11, 1274–1285. [https://doi.org/10.1016/S2095-3119\(12\)60124-5](https://doi.org/10.1016/S2095-3119(12)60124-5)
- Cao, P., Lu, C., Yu, Z., 2018. Historical nitrogen fertilizer use in agricultural ecosystems of the contiguous United States during 1850–2015: application rate, timing, and fertilizer types. *Earth System Science Data* 10, 969–984. <https://doi.org/10.5194/essd-10-969-2018>
- Cao, W., Moss, D.N., 1989. Temperature Effect on Leaf Emergence and Phyllochron in Wheat and Barley. *Crop Science* 29, [cropsci1989.0011183X002900040038x](https://doi.org/10.2135/cropsci1989.0011183X002900040038x)
- Chaplot, V.A.M., Bissonnais, Y.L., 2003. Runoff Features for Interrill Erosion at Different Rainfall Intensities, Slope Lengths, and Gradients in an Agricultural Loessial Hillslope. *Soil Science Society of America Journal* 67, 844–851. <https://doi.org/10.2136/sssaj2003.8440>
- Chenu, K., Porter, J.R., Martre, P., Basso, B., Chapman, S.C., Ewert, F., Bindi, M., Asseng, S., 2017. Contribution of Crop Models to Adaptation in Wheat. *Trends in Plant Science*.
- Christopher, J.T., Manschadi, A.M., Hammer, G.L., Borrell, A.K., 2008. Developmental and physiological traits associated with high yield and stay-green phenotype in wheat. *Australian Journal of Agricultural Research* 59, 354. <https://doi.org/10.1071/AR07193>
- Coville, F.V., 1920. The Influence of Cold in Stimulating the Growth of Plants. *PNAS* 6, 434–435. <https://doi.org/10.1073/pnas.6.7.434>
- Davidson, H.R., Campbell, C.A., 1983. The Effect of Temperature, Moisture and Nitrogen on the Rate of Development of Spring Wheat as Measured by Degree Days. *Canadian Journal of Plant Science*. <https://doi.org/10.4141/cjps83-106>
- de Wit, C.T., 1965. Photosynthesis of leaf canopies (No. 663), *Agricultural Research Reports*. Centre for Agricultural Publications and Documentation, Wageningen.
- Diez, T., Weigelt, H., 1991. *Böden unter landwirtschaftlicher Nutzung, Wissen für die Praxis*. BLV-Verl.-Ges, München.
- Ebrahimi, E., Manschadi, A.M., Neugschwandtner, R.W., Eitzinger, J., Thaler, S., Kaul, H.-P., 2016. Assessing the impact of climate change on crop management in winter wheat – a case study for Eastern Austria. *The Journal of Agricultural Science* 1–18. <https://doi.org/10.1017/S0021859616000083>
- Eck, C., Garcke, H., Knabner, P., 2011. *Mathematische Modellierung, Springer-Lehrbuch*. Springer Berlin Heidelberg, Berlin, Heidelberg.
- Efron, B., 1983. Estimating the Error Rate of a Prediction Rule: Improvement on Cross-Validation. *Journal of the American Statistical Association* 78, 316–331. <https://doi.org/10.2307/2288636>
- Eitzinger, J., Thaler, S., Schmid, E., Strauss, F., Ferrise, R., Moriondo, M., Bindi, M., Palosuo, T., Rötter, R., Kersebaum, K.C., 2013a. Sensitivities of crop models to extreme weather conditions during flowering period demonstrated for maize and winter wheat in Austria. *The Journal of Agricultural Science* 151, 813–835.
- Eitzinger, J., TRNKA, M., SEMERÁDOVÁ, D., THALER, S., SVOBODOVÁ, E., HLAVINKA, P., SISKÁ, B., TAKÁČ, J., MALATINSKÁ, L., NOVÁKOVÁ, M., DUBROVSKÁ, M., ZALUD, Z., 2013b. Regional climate change impacts on agricultural crop production in Central and Eastern Europe - hotspots, regional differences and common trends. *The Journal of Agricultural Science; Cambridge* 151, 787–812. <http://dx.doi.org/10.1017/S0021859612000767>
- Follett, J.R., Follett, R.F., Herz, W.C., 2010. Environmental and human impacts of reactive nitrogen. *Advances in Nitrogen Management for Water Quality* 1–37.
- Forcella, F., Benecch Arnold, R.L., Sanchez, R., Ghera, C.M., 2000. Modeling seedling emergence. *Field Crops Research* 67, 123–139. [https://doi.org/10.1016/S0378-4290\(00\)00088-5](https://doi.org/10.1016/S0378-4290(00)00088-5)

- Fox, D.M., Bryan, R.B., Price, A.G., 1997. The influence of slope angle on final infiltration rate for interrill conditions. *Geoderma* 80, 181–194. [https://doi.org/10.1016/S0016-7061\(97\)00075-X](https://doi.org/10.1016/S0016-7061(97)00075-X)
- Fox, J., Weisberg, S., 2011. *An R Companion to Applied Regression*, Second. ed. Sage, Thousand Oaks CA.
- Frank, A.B., Bauer, A., 1995. Phyllochron Differences in Wheat, Barley, and Forage Grasses. *Crop Science* 35, 19–23. <https://doi.org/10.2135/cropsci1995.0011183X003500010004x>
- Fuchs, W., 2016. Simulating growth and phenology of wheat in Pannonian eastern Austria using APSIM (Master's thesis). University of Natural Resources and Life Sciences, Vienna.
- Gallagher, J.N., 1979. Field Studies of Cereal Leaf Growth I. INITIATION AND EXPANSION IN RELATION TO TEMPERATURE AND ONTOGENY. *J Exp Bot* 30, 625–636. <https://doi.org/10.1093/jxb/30.4.625>
- Geisseler, D., Miller, K.S., Aegerter, B.J., Clark, N.E., Miyao, E.M., 2019. Estimation of Annual Soil Nitrogen Mineralization Rates using an Organic-Nitrogen Budget Approach. *Soil Science Society of America Journal* 83, 1227–1235. <https://doi.org/10.2136/sssaj2018.12.0473>
- Grolemund, G., Wickham, H., 2011. Dates and Times Made Easy with lubridate. *Journal of Statistical Software* 40, 1–25.
- Grüner Bericht, 2019. Grüner Bericht 2019. Die Republik Österreich, vertreten durch die Bundesministerin für Nachhaltigkeit und Tourismus, Wien.
- Grüner Bericht Niederösterreich, 2018. Grüner Bericht Niederösterreich 2018. Amt der NÖ Landesregierung, Abteilung Landwirtschaftsförderung, St. Pölten.
- Hadas, A., Feigenbaum, S., Feigin, A., Portnoy, R., 1986. Nitrification Rates in Profiles of Differently Managed Soil Types. *Soil Science Society of America Journal* 50, 633–639. <https://doi.org/10.2136/sssaj1986.03615995005000030018x>
- Hammer, G.L., Sinclair, T.R., Boote, K.J., Wright, G.C., Meinke, H., Bell, M.J., 1995. A Peanut Simulation Model: I. Model Development and Testing. *Agronomy Journal* 87, 1085–1093. <https://doi.org/10.2134/agronj1995.00021962008700060009x>
- Haun, J.R., 1973. Visual Quantification of Wheat Development. *Agronomy Journal* 65, 116. <https://doi.org/10.2134/agronj1973.00021962006500010035x>
- Hay, R.K.M. [VerfasserIn], Porter, J.R. [VerfasserIn], 2006. *The physiology of crop yield*, 2. ed.. ed. Oxford [u.a.] : Blackwell Publ.
- He, D., Wang, E., Wang, J., Robertson, M.J., 2017. Data requirement for effective calibration of process-based crop models. *Agricultural and Forest Meteorology* 234–235, 136–148. <https://doi.org/10.1016/j.agrformet.2016.12.015>
- Heffer, P., Gruère, A., Roberts, T., 2017. *Assessment of Fertilizer Use by Crop at the Global Level*. International Fertilizer Association (IFA) and International Plant Nutrition Institute (IPNI).
- Hiederer, R., 2013a. *Mapping soil typologies-Spatial decision support applied to European Soil Database*. Publications Office of the European Union, Luxembourg.
- Hiederer, R., 2013b. *Mapping soil properties for Europe—spatial representation of soil database attributes*. Publications Office of the European Union, EUR26082EN Scientific and Technical Research series. Luxembourg.
- Hochman, Z., Rees, H. van, Carberry, P.S., Hunt, J.R., McCown, R.L., Gartmann, A., Holzworth, D., Rees, S. van, Dalgliesh, N.P., Long, W., Peake, A.S., Poulton, P.L., McClelland, T., 2009. Re-inventing model-based decision support with Australian dryland farmers. 4. Yield Prophet® helps farmers monitor and manage crops in a variable climate. *Crop Pasture Sci.* 60, 1057–1070. <https://doi.org/10.1071/CP09020>
- Holzworth, D.P., Huth, N.I., deVoil, P.G., Zurcher, E.J., Herrmann, N.I., McLean, G., Chenu, K., van Oosterom, E.J., Snow, V., Murphy, C., Moore, A.D., Brown, H., Whish, J.P.M., Verrall, S., Fainges, J., Bell, L.W., Peake, A.S., Poulton, P.L., Hochman, Z., Thorburn, P.J., Gaydon, D.S., Dalgliesh, N.P., Rodriguez, D., Cox, H., Chapman, S., Doherty, A., Teixeira, E., Sharp, J., Cichota, R., Vogeler, I., Li, F.Y., Wang, E., Hammer, G.L., Robertson, M.J.,

- Dimes, J.P., Whitbread, A.M., Hunt, J., van Rees, H., McClelland, T., Carberry, P.S., Hargreaves, J.N.G., MacLeod, N., McDonald, C., Harsdorf, J., Wedgwood, S., Keating, B.A., 2014. APSIM – Evolution towards a new generation of agricultural systems simulation. *Environmental Modelling & Software* 62, 327–350. <https://doi.org/10.1016/j.envsoft.2014.07.009>
- Hunt, L.A., Boote, K.J., 1998. Data for model operation, calibration, and evaluation, in: Tsuji, G.Y., Hoogenboom, G., Thornton, P.K. (Eds.), *Understanding Options for Agricultural Production, Systems Approaches for Sustainable Agricultural Development*. Springer Netherlands, Dordrecht, pp. 9–39. [https://doi.org/10.1007/978-94-017-3624-4\\_2](https://doi.org/10.1007/978-94-017-3624-4_2)
- Jamieson, P.D., Brooking, I.R., Porter, J.R., Wilson, D.R., 1995. Prediction of leaf appearance in wheat: a question of temperature. *Field Crops Research* 41, 35–44. [https://doi.org/10.1016/0378-4290\(94\)00102-1](https://doi.org/10.1016/0378-4290(94)00102-1)
- Justes, E., Mary, B., Meynard, J.-M., Machet, J.-M., Thelier-Huche, L., 1994. Determination of a Critical Nitrogen Dilution Curve for Winter Wheat Crops. *Annals of Botany* 74, 397–407. <https://doi.org/10.1006/anbo.1994.1133>
- Keating, B.A., Carberry, P.S., Hammer, G.L., Probert, M.E., Robertson, M.J., Holzworth, D., Huth, N.I., Hargreaves, J.N.G., Meinke, H., Hochman, Z., McLean, G., Verburg, K., Snow, V., Dimes, J.P., Silburn, M., Wang, E., Brown, S., Bristow, K.L., Asseng, S., Chapman, S., McCown, R.L., Freebairn, D.M., Smith, C.J., 2003. An overview of APSIM, a model designed for farming systems simulation. *European Journal of Agronomy* 18, 267–288. [https://doi.org/10.1016/S1161-0301\(02\)00108-9](https://doi.org/10.1016/S1161-0301(02)00108-9)
- Kersebaum, K.C., Boote, K.J., Jorgenson, J.S., Nendel, C., Bindi, M., Frühauf, C., Gaiser, T., Hoogenboom, G., Kollas, C., Olesen, J.E., Rötter, 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. *Environmental Modelling & Software* 72, 402–417. <https://doi.org/10.1016/j.envsoft.2015.05.009>
- Kim, S.-H., Hsiao, J., Kinmonth-Schultz, H., 2020. Chapter 1: Advances and improvements in modeling plant processes, in: *Advances in Crop Modelling for a Sustainable Agriculture*. Burleigh Dodds Science Publishing, Cambridge.
- Kim, S.-H., Yang, Y., Timlin, D.J., Fleisher, D.H., Dathe, A., Reddy, V.R., Staver, K., 2012. Modeling Temperature Responses of Leaf Growth, Development, and Biomass in Maize with MAIZSIM. *Agronomy Journal* 104, 1523–1537. <https://doi.org/10.2134/agronj2011.0321>
- Kiniry, J.R., Jones, C.A., O'toole, J.C., Blanchet, R., Cabelguenne, M., Spanel, D.A., 1989. Radiation-use efficiency in biomass accumulation prior to grain-filling for five grain-crop species. *Field Crops Research* 20, 51–64. [https://doi.org/10.1016/0378-4290\(89\)90023-3](https://doi.org/10.1016/0378-4290(89)90023-3)
- Kirby, E., Perry, M., 1987. Leaf emergence rates of wheat in a mediterranean environment. *Aust. J. Agric. Res.* 38, 455–464.
- Liu, B., Asseng, S., Liu, L., Tang, L., Cao, W., Zhu, Y., 2016. Testing the responses of four wheat crop models to heat stress at anthesis and grain filling. *Global Change Biology* 22, 1890–1903. <https://doi.org/10.1111/gcb.13212>
- Liu, B., Martre, P., Ewert, F., Porter, J.R., Challinor, A.J., Müller, C., Ruane, A.C., Waha, K., Thorburn, P.J., Aggarwal, P.K., Ahmed, M., Balkovič, J., Basso, B., Biernath, C., Bindi, M., Cammarano, D., Sanctis, G.D., Dumont, B., Espadafor, M., Rezaei, E.E., Ferrise, R., Garcia-Vila, M., Gayler, S., Gao, Y., Horan, H., Hoogenboom, G., Izaurrealde, R.C., Jones, C.D., Kassie, B.T., Kersebaum, K.C., Klein, C., Koehler, A.-K., Maiorano, A., Minoli, S., Martin, M.M.S., Kumar, S.N., Nendel, C., O'Leary, G.J., Palosuo, T., Priesack, E., Ripoche, D., Rötter, R.P., Semenov, M.A., Stöckle, C., Streck, T., Supit, I., Tao, F., Velde, M.V. der, Wallach, D., Wang, E., Webber, H., Wolf, J., Xiao, L., Zhang, Z., Zhao, Z., Zhu, Y., Asseng, S., 2019. Global wheat production with 1.5 and 2.0°C above pre-industrial warming. *Global Change Biology* 25, 1428–1444. <https://doi.org/10.1111/gcb.14542>
- Maiorano, A., Martre, P., Asseng, S., Ewert, F., Müller, C., Rötter, R.P., Ruane, A.C., Semenov, M.A., Wallach, D., Wang, E., Alderman, P.D., Kassie, B.T., Biernath, C., Basso, B., Cammarano, D., Challinor, A.J., Doltra, J., Dumont, B., Rezaei, E.E., Gayler, S., Kersebaum,

- K.C., Kimball, B.A., Koehler, A.-K., Liu, B., O'Leary, G.J., Olesen, J.E., Ottman, M.J., Priesack, E., Reynolds, M., Stratonovitch, P., Streck, T., Thorburn, P.J., Waha, K., Wall, G.W., White, J.W., Zhao, Z., Zhu, Y., 2017. Crop model improvement reduces the uncertainty of the response to temperature of multi-model ensembles. *Field Crops Research* 202, 5–20. <https://doi.org/10.1016/j.fcr.2016.05.001>
- Manschadi, A.M., Christopher, J., deVoil, P., Hammer, G.L., 2006. The role of root architectural traits in adaptation of wheat to water-limited environments. *Functional Plant Biology* 33, 823. <https://doi.org/10.1071/FP06055>
- Manschadi, A.M., Eitzinger, J., Breisch, M., Fuchs, W., Neubauer, T., Soltani, A., 2020a. Full Parameterisation Matters for the Best Performance of Crop Models: Inter-comparison of a Simple and a Detailed Maize Model. *Int. J. Plant Prod.* <https://doi.org/10.1007/s42106-020-00116-2>
- Manschadi, A.M., Kaul, H.-P., Eitzinger, J., Friedel, J., Pötsch, E., Bodner, G., Neubauer, T., 2019. Farm/IT – Innovative Digital Technologies for Strengthening the Resilience of Austrian Farming Systems to Climate Risks. Presented at the 20. Österreichischer Klimatag, Climate Change Centre Austria (CCCA), 20. KLIMATAG - Facetten der österreichischen Klimaforschung, Vienna, Austria, p. 139.
- Manschadi, A.M., Soltani, A., Fuchs, W., Ryall, S.P., Koppensteiner, L., Eitzinger, J., Kaul, H.-P., Neubauer, T., 2020b. Integrating Crop Modelling in the Smart Farming Project Farm/IT in Austria: Achievements and Challenges. Presented at the iCROP2020 - Crop Modelling for the Future, Organising and Scientific Committees, Book of Abstracts - Second International Crop Modelling Symposium, Montpellier, FRANCE, p. 599.
- Marschner, P. (Ed.), 2012. *Marschner's Mineral Nutrition of Higher Plants*. Elsevier. <https://doi.org/10.1016/C2009-0-63043-9>
- McMaster, G.S., White, J.W., Hunt, L.A., Jamieson, P.D., Dhillon, S.S., Ortiz-Monasterio, J.I., 2008. Simulating the Influence of Vernalization, Photoperiod and Optimum Temperature on Wheat Developmental Rates. *Annals of Botany* 102, 561. <https://doi.org/10.1093/aob/mcn115>
- McMaster, G.S., Wilhelm, W.W., 1997. Growing degree-days: one equation, two interpretations. *Agricultural and Forest Meteorology* 87, 291–300. [https://doi.org/10.1016/S0168-1923\(97\)00027-0](https://doi.org/10.1016/S0168-1923(97)00027-0)
- Meier, U. (Ed.), 2001. *Entwicklungsstadien mono- und dikotyler Pflanzen - BBCH Monographie*.
- Meisinger, J.J., Randall, G.W., 1991. Estimating Nitrogen Budgets for Soil-Crop Systems, in: *Managing Nitrogen for Groundwater Quality and Farm Profitability*. John Wiley & Sons, Ltd, pp. 85–124. <https://doi.org/10.2136/1991.managingnitrogen.c5>
- Mendiburu, F. de, 2017. *agricolae: Statistical Procedures for Agricultural Research*.
- Microsoft Corporation, 2018. Microsoft Excel.
- Mitscherlich, E.A., 1954. *Bodenkunde für Landwirte, Forstwirte und Gärtner in pflanzenphysiologischer Ausrichtung und Auswertung* (No. 7., neubearb. Aufl.). Parey, Berlin.
- Moeller, C., Pala, M., Manschadi, A.M., Meinke, H., Sauerborn, J., 2007. Assessing the sustainability of wheat-based cropping systems using APSIM: model parameterisation and evaluation. *Australian Journal of Agricultural Research* 58, 75. <https://doi.org/10.1071/AR06186>
- Moitzi, G., Neugschwandtner, R.W., Kaul, H.-P., Wagentristl, H., 2020. Efficiency of Mineral Nitrogen Fertilization in Winter Wheat under Pannonian Climate Conditions. *Agriculture* 10, 541. <https://doi.org/10.3390/agriculture10110541>
- Monteith, J.L., 1996. The Quest for Balance in Crop Modeling. *Agronomy Journal* 88, 695. <https://doi.org/10.2134/agronj1996.00021962008800050003x>
- Monteith, J.L., 1965. Light Distribution and Photosynthesis in Field Crops. *Annals of Botany* 29, 17–37. <https://doi.org/10.1093/oxfordjournals.aob.a083934>
- Muggeo, V.M.R., 2003. Estimating regression models with unknown break-points. *Statistics in Medicine* 22, 3055–3071. <https://doi.org/10.1002/sim.1545>

- Neugschwandtner, R., Kaul, H.-P., Liebhard, P., Wagentristl, H., 2015. Winter wheat yields in a long-term tillage experiment under Pannonian climate conditions. *Plant Soil and Environment* 61, 145–150. <https://doi.org/10.17221/820/2014-PSE>
- Neugschwandtner, R.W., Böhm, K., Hall, R.M., Kaul, H.-P., 2015. Development, growth, and nitrogen use of autumn- and spring-sown facultative wheat. *Acta Agriculturae Scandinavica, Section B—Soil & Plant Science* 65, 6–13. <https://doi.org/10.1080/09064710.2014.958522>
- O’Leary, G.J., Christy, B., Nuttall, J., Huth, N., Cammarano, D., Stöckle, C., Basso, B., Shcherbak, I., Fitzgerald, G., Luo, Q., Farre-Codina, I., Palta, J., Asseng, S., 2015. Response of wheat growth, grain yield and water use to elevated CO<sub>2</sub> under a Free-Air CO<sub>2</sub> Enrichment (FACE) experiment and modelling in a semi-arid environment. *Glob Change Biol* 21, 2670–2686. <https://doi.org/10.1111/gcb.12830>
- Passioura, J.B., 1996. Simulation Models: Science, Snake Oil, Education, or Engineering? *Agronomy Journal* 88, 690. <https://doi.org/10.2134/agronj1996.00021962008800050002x>
- Perry, M.W., Belford, R.K., 2000. Chapter Two. The Structure and Development of the Cereal Plant, in: *The Wheat Book : Principles and Practice*, Bulletin 4443. Department of Agriculture and Food, Western Australia, Perth.
- Porter, J.R., 1993. AFRCWHEAT2: A model of the growth and development of wheat incorporating responses to water and nitrogen. *European Journal of Agronomy* 2, 69–82. [https://doi.org/10.1016/S1161-0301\(14\)80136-6](https://doi.org/10.1016/S1161-0301(14)80136-6)
- Porter, J.R., Gawith, M., 1999. Temperatures and the growth and development of wheat: a review. *European Journal of Agronomy* 10, 23–36. [https://doi.org/10.1016/S1161-0301\(98\)00047-1](https://doi.org/10.1016/S1161-0301(98)00047-1)
- Priestley, C.H.B., Taylor, R.J., 1972. On the Assessment of Surface Heat Flux and Evaporation Using Large-Scale Parameters. *Mon. Wea. Rev.* 100, 81–92. [https://doi.org/10.1175/1520-0493\(1972\)100%3C0081:OTAOSH%3E2.3.CO;2](https://doi.org/10.1175/1520-0493(1972)100%3C0081:OTAOSH%3E2.3.CO;2)
- R Core Team, 2015. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Rawls, W.J., Brakensiek, D.L., Saxton, K.E., 1982. Estimation of Soil Water Properties. *Transactions of the ASAE* 25, 1316–1320. <https://doi.org/10.13031/2013.33720>
- Reynolds, M., Kropff, M., Crossa, J., Koo, J., Kruseman, G., Molero Milan, A., Rutkoski, J., Schulthess, U., Balwinder-Singh, Sonder, K., Tonnang, H., Vadez, V., 2018. Role of Modelling in International Crop Research: Overview and Some Case Studies. *Agronomy* 8, 291. <https://doi.org/10.3390/agronomy8120291>
- Richards, L.A., 1931. Capillary conduction of liquids through porous mediums. *Physics I*, 318–333. <https://doi.org/10.1063/1.1745010>
- Risch, A.C., Zimmermann, S., Ochoa-Hueso, R., Schütz, M., Frey, B., Firn, J.L., Fay, P.A., Hagedorn, F., Borer, E.T., Seabloom, E.W., Harpole, W.S., Knops, J.M.H., McCulley, R.L., Broadbent, A. a. D., Stevens, C.J., Silveira, M.L., Adler, P.B., Báez, S., Biederman, L.A., Blair, J.M., Brown, C.S., Caldeira, M.C., Collins, S.L., Daleo, P., di Virgilio, A., Ebeling, A., Eisenhauer, N., Esch, E., Eskelinen, A., Hagenah, N., Hautier, Y., Kirkman, K.P., MacDougall, A.S., Moore, J.L., Power, S.A., Prober, S.M., Roscher, C., Sankaran, M., Siebert, J., Speziale, K.L., Tognetti, P.M., Virtanen, R., Yahdjian, L., Moser, B., 2019. Soil net nitrogen mineralisation across global grasslands. *Nature Communications* 10, 4981. <https://doi.org/10.1038/s41467-019-12948-2>
- Ritchie, J., 1998. Soil water balance and plant water stress, in: Tsuji, G.Y., Hoogenboom, G., Thornton, P.K. (Eds.), *Understanding Options for Agricultural Production*. Springer, Dordrecht, pp. 41–54.
- Ritchie, J.T., 1972. Model for predicting evaporation from a row crop with incomplete cover. *Water Resources Research* 8, 1204–1213. <https://doi.org/10.1029/WR008i005p01204>
- Rose, T., Nagler, S., Kage, H., 2017. Yield formation of Central-European winter wheat cultivars on a large scale perspective. *European Journal of Agronomy* 86, 93–102. <https://doi.org/10.1016/j.eja.2017.03.003>

- 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., Baigoria, G., Winter, J.M., 2013. The Agricultural Model Intercomparison and Improvement Project (AgMIP): Protocols and pilot studies. *Agricultural and Forest Meteorology*, Agricultural prediction using climate model ensembles 170, 166–182. <https://doi.org/10.1016/j.agrformet.2012.09.011>
- Rötter, R.P., Carter, T.R., Olesen, J.E., Porter, J.R., 2011. Crop-climate models need an overhaul. *Nature Clim. Change* 1, 175–177. <https://doi.org/10.1038/nclimate1152>
- Rötter, R.P., Palosuo, T., Kersebaum, K.C., Angulo, C., Bindi, M., Ewert, F., Ferrise, R., Hlavinka, P., Moriondo, M., Nendel, C., Olesen, J.E., Patil, R.H., Ruget, F., Takáč, J., Trnka, M., 2012. Simulation of spring barley yield in different climatic zones of Northern and Central Europe: A comparison of nine crop models. *Field Crops Research* 133, 23–36. <https://doi.org/10.1016/j.fcr.2012.03.016>
- Sadras, V.O., Angus, J.F., 2006. Benchmarking water-use efficiency of rainfed wheat in dry environments. *Aust. J. Agric. Res.* 57, 847. <https://doi.org/10.1071/AR05359>
- Salo, T.J., Palosuo, T., Kersebaum, K.C., Nendel, C., Angulo, C., Ewert, F., Bindi, M., Calanca, P., Klein, T., Moriondo, M., Ferrise, R., Olesen, J.E., Patil, R.H., Ruget, F., Takáč, J., Hlavinka, P., Trnka, M., Rötter, R.P., 2015. Comparing the performance of 11 crop simulation models in predicting yield response to nitrogen fertilization. *The Journal of Agricultural Science* 1–23. <https://doi.org/10.1017/S0021859615001124>
- Salvagiotti, F., Miralles, D.J., 2007. Wheat development as affected by nitrogen and sulfur nutrition. *Aust. J. Agric. Res.* 58, 39–45. <https://doi.org/10.1071/AR06090>
- 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? *European Journal of Agronomy* 94, 25–35. <https://doi.org/10.1016/j.eja.2018.01.006>
- Sethmacher, V., 2018. Vergleich der Bodenfeuchte und Saugspannung sowie der Transpirationsleistung in Maisparzellen mit verschiedenen Behandlungen während Hitze- und Trockenperioden.
- Shaffer, M.J., Ma, L., Hansen, S. (Eds.), 2001. Modeling carbon and nitrogen dynamics for soil management. CRC Press.
- Shearman, V.J., Sylvester-Bradley, R., Scott, R.K., Foulkes, M.J., 2005. CROP PHYSIOLOGY & METABOLISM. *CROP SCIENCE* 45, 11.
- Sinclair, T.R., 1998. Historical Changes in Harvest Index and Crop Nitrogen Accumulation. *Crop Science* 38, cropscl1998.0011183X003800030002x. <https://doi.org/10.2135/cropsci1998.0011183X003800030002x>
- Sinclair, T.R., 1986. Water and nitrogen limitations in soybean grain production I. Model development. *Field Crops Research* 15, 125–141. [https://doi.org/10.1016/0378-4290\(86\)90082-1](https://doi.org/10.1016/0378-4290(86)90082-1)
- Sinclair, T.R., 1984. Leaf Area Development in Field-Grown Soybeans. *Agronomy Journal* 76, 141–146. <https://doi.org/10.2134/agronj1984.00021962007600010034x>
- Sinclair, T.R., Amir, J., 1992. A model to assess nitrogen limitations on the growth and yield of spring wheat. *Field Crops Research* 30, 63–78. [https://doi.org/10.1016/0378-4290\(92\)90057-G](https://doi.org/10.1016/0378-4290(92)90057-G)
- Sinclair, T.R., Muchow, R.C., 1995. Effect of Nitrogen Supply on Maize Yield: I. Modeling Physiological Responses. *Agronomy Journal* 87, 632–641. <https://doi.org/10.2134/agronj1995.00021962008700040005x>
- Sinclair, T.R., Muchow, R.C., Monteith, J.L., 1997. Model Analysis of Sorghum Response to Nitrogen in Subtropical and Tropical Environments. *Agronomy Journal* 89, 201–207. <https://doi.org/10.2134/agronj1997.00021962008900020009x>
- Sinclair, T.R., Seligman, N.G., 1996. Crop Modeling: From Infancy to Maturity. *Agronomy Journal* 88, 698. <https://doi.org/10.2134/agronj1996.00021962008800050004x>

- Sinclair, T.R., Soltani, A., Marrou, H., Ghanem, M., Vadez, V., 2020. Geospatial assessment for crop physiological and management improvements with examples using the simple simulation model. *Crop Science* 60, 700–708. <https://doi.org/10.1002/csc2.20106>
- Singh, U., Porter, C., 2020. Chapter 3: Improving modeling of nutrient cycles in crop cultivation, in: *Advances in Crop Modelling for a Sustainable Agriculture*. Burleigh Dodds Science Publishing, Cambridge.
- Soltani, A., Alimagham, S.M., Nehbandani, A., Torabi, B., Zeinali, E., Dadrasi, A., Zand, E., Ghassemi, S., Pourshirazi, S., Alasti, O., Hosseini, R.S., Zahed, M., Arabameri, R., Mohammadzadeh, Z., Rahban, S., Kamari, H., Fayazi, H., Mohammadi, S., Keramat, S., Vadez, V., van Ittersum, M.K., Sinclair, T.R., 2020. SSM-iCrop2: A simple model for diverse crop species over large areas. *Agricultural Systems* 182, 102855. <https://doi.org/10.1016/j.agry.2020.102855>
- Soltani, A., Maddah, V., Sinclair, T.R., 2013. SSM-Wheat: a simulation model for wheat development, growth and yield. *International Journal of Plant Production* 7, 711–740.
- Soltani, A., Robertson, M.J., Manschadi, A.M., 2006. Modeling chickpea growth and development: Nitrogen accumulation and use. *Field Crops Research* 99, 24–34. <https://doi.org/10.1016/j.fcr.2006.02.006>
- Soltani, A., Sinclair, T.R., 2015. A comparison of four wheat models with respect to robustness and transparency: Simulation in a temperate, sub-humid environment. *Field Crops Research* 175, 37–46. <https://doi.org/10.1016/j.fcr.2014.10.019>
- Soltani, A., Sinclair, T.R. (Eds.), 2012. *Modeling physiology of crop development, growth and yield*. CABI, Wallingford.
- Soltani, A., Sinclair, T.R., 2011. A simple model for chickpea development, growth and yield. *Field Crops Research*, This Issue Includes a Special Issue Section on: The Use of Field Crop Knowledge in Integrative Inter-Scale Systems Approaches: The Plant, the Field and the Property 124, 252–260. <https://doi.org/10.1016/j.fcr.2011.06.021>
- Stöckle, C., Meza, F., 2020. Chapter 4: Improving modelling of water cycles in crop cultivation, in: *Advances in Crop Modelling for a Sustainable Agriculture*. Burleigh Dodds Science Publishing, Cambridge.
- Stratonovitch, P., Semenov, M.A., 2015. Heat tolerance around flowering in wheat identified as a key trait for increased yield potential in Europe under climate change. *J Exp Bot* 66, 3599–3609. <https://doi.org/10.1093/jxb/erv070>
- Taylor, S.L., Payton, M.E., Raun, W.R., 1999. Relationship between mean yield, coefficient of variation, mean square error, and plot size in wheat field experiments. *Communications in Soil Science and Plant Analysis* 30, 1439–1447. <https://doi.org/10.1080/00103629909370298>
- Thomas, H., Ougham, H., 2014. The stay-green trait. *Journal of Experimental Botany* 65, 3889–3900. <https://doi.org/10.1093/jxb/eru037>
- USDA, 2004. Estimation of direct runoff from storm rainfall. *National Engineering Handbook* 630, 1–51.
- Van Evert, F.K., 2020. Chapter 14: Data for developing, testing, and applying crop and farm models, in: *Advances in Crop Modelling for a Sustainable Agriculture*. Burleigh Dodds Science Publishing, Cambridge.
- Wahbi, A., Sinclair, T.R., 2005. Simulation analysis of relative yield advantage of barley and wheat in an eastern Mediterranean climate. *Field Crops Research* 91, 287–296. <https://doi.org/10.1016/j.fcr.2004.07.020>
- 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. *Environmental Modelling & Software* 26, 386–394. <https://doi.org/10.1016/j.envsoft.2010.09.004>
- Wandl, M., Horvath, D., n.d. Digitale Bodenkarte von Österreich - eBOD [WWW Document]. URL <https://bodenkarte.at/>

- Wang, E., Engel, T., 1998. Simulation of phenological development of wheat crops. *Agricultural Systems* 58, 1–24. [https://doi.org/10.1016/S0308-521X\(98\)00028-6](https://doi.org/10.1016/S0308-521X(98)00028-6)
- Wang, E., Martre, P., Zhao, Z., Ewert, F., Maiorano, A., Rötter, R.P., Kimball, B.A., Ottman, M.J., Wall, G.W., White, J.W., Reynolds, M.P., Alderman, P.D., Aggarwal, P.K., Anothai, J., Basso, B., Biernath, C., Cammarano, D., Challinor, A.J., De Sanctis, G., Doltra, J., Fereres, E., Garcia-Vila, M., Gayler, S., Hoogenboom, G., Hunt, L.A., Izaurrealde, R.C., Jabloun, M., Jones, C.D., Kersebaum, K.C., Koehler, A.-K., Liu, L., Müller, C., Naresh Kumar, S., Nendel, C., O’Leary, G., Olesen, J.E., Palosuo, T., Priesack, E., Eyshi Rezaei, E., Ripoche, D., Ruane, A.C., Semenov, M.A., Shcherbak, I., Stöckle, C., Stratonovitch, P., Streck, T., Supit, I., Tao, F., Thorburn, P., Waha, K., Wallach, D., Wang, Z., Wolf, J., Zhu, Y., Asseng, S., 2017. The uncertainty of crop yield projections is reduced by improved temperature response functions. *Nature Plants* 3, 17102. <https://doi.org/10.1038/nplants.2017.102>
- Wang, E., Smith, C.J., 2004. Modelling the growth and water uptake function of plant root systems: a review. *Aust. J. Agric. Res.* 55, 501–523. <https://doi.org/10.1071/AR03201>
- Webber, H., Martre, P., Asseng, S., Kimball, B., White, J., Ottman, M., Wall, G.W., De Sanctis, G., Doltra, J., Grant, R., Kassie, B., Maiorano, A., Olesen, J.E., Ripoche, D., Rezaei, E.E., Semenov, M.A., Stratonovitch, P., Ewert, F., 2017. Canopy temperature for simulation of heat stress in irrigated wheat in a semi-arid environment: A multi-model comparison. *Field Crops Research* 202, 21–35. <https://doi.org/10.1016/j.fcr.2015.10.009>
- White, J.W., Hoogenboom, G., Kimball, B.A., Wall, G.W., 2011. Methodologies for simulating impacts of climate change on crop production. *Field Crops Research* 124, 357–368. <https://doi.org/10.1016/j.fcr.2011.07.001>
- Wickham, H., 2009. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.
- Wickham, H., Chang, W., 2017. *devtools: Tools to Make Developing R Packages Easier*.
- Wickham, H., François, R., Henry, L., Müller, K., 2019. *dplyr: A Grammar of Data Manipulation*.
- Wickham, H., Henry, L., 2018. *tidyr: Easily Tidy Data with “spread()” and “gather()” Functions*.
- Xue, Q., Weiss, A., Baenziger, P.S., 2004. Predicting leaf appearance in field-grown winter wheat: evaluating linear and non-linear models. *Ecological Modelling* 175, 261–270. <https://doi.org/10.1016/j.ecolmodel.2003.10.018>
- Yan, W., Hunt, L.A., 1999. An Equation for Modelling the Temperature Response of Plants using only the Cardinal Temperatures. *Ann Bot* 84, 607–614. <https://doi.org/10.1006/anbo.1999.0955>
- Yang, Y., Zhou, C., Li, N., Han, K., Meng, Y., Tian, X., Wang, L., 2015. Effects of conservation tillage practices on ammonia emissions from Loess Plateau rain-fed winter wheat fields. *Atmospheric Environment* 104, 59–68. <https://doi.org/10.1016/j.atmosenv.2015.01.007>
- Yin, X., Kropff, M.J., McLaren, G., Visperas, R.M., 1995. A nonlinear model for crop development as a function of temperature. *Agricultural and Forest Meteorology* 77, 1–16. [https://doi.org/10.1016/0168-1923\(95\)02236-Q](https://doi.org/10.1016/0168-1923(95)02236-Q)
- Zadoks, J.C., Chang, T.T., Konzak, C.F., 1974. A decimal code for the growth stages of cereals. *Weed Research* 14, 415–421. <https://doi.org/10.1111/j.1365-3180.1974.tb01084.x>
- Zeileis, A., Grothendieck, G., 2005. zoo: S3 Infrastructure for Regular and Irregular Time Series. *Journal of Statistical Software* 14, 1–27. <https://doi.org/10.18637/jss.v014.i06>
- Zhao, Z., Wang, E., Wang, Z., Zang, H., Liu, Y., Angus, J.F., 2014. A reappraisal of the critical nitrogen concentration of wheat and its implications on crop modeling. *Field Crops Research* 164, 65–73. <https://doi.org/10.1016/j.fcr.2014.05.004>