International Institute for Applied Systems Analysis
IIASA www.iiasa.ac.at

Address: IIASA, Schlossplatz 1, A-2361 Laxenburg, Austria

Email: permissions@iiasa.ac.at

Young Scientists Summer Program

Soil data requirements to improve local fertilization strategies – and how to get there

Marlene Palka (marlene.palka@boku.ac.at)

Approved by:

Mentor(s): Rastislav Skalský and Juan Carlos Laso Bayas
Program: BNR/ASA
Date: September 30, 2022

This report represents the work completed by the author during the IIASA Young Scientists Summer Program (YSSP) with approval from the YSSP mentor.

It was finished by September 30, 2022 and has not been altered or revised since.

Mentor signature:

Rastislav Skalský

Juan Carlos Laso Bayas

Table of contents

Abstract ii
About the author iii
Acknowledgments iii

Introduction 1

Material and methods 2

Description of the process-based crop model SSM-iCrop 2
Description of the reference field experiments 3
Description of soil data sets 4
Introduction of input uncertainty 5
Spatial extension of the prototype application 7
Statistical analysis 7

Results 9

Performance of the prototype 9
Contribution of model input data uncertainty to simulation error 10
Refining of model input data 11
Spatial extension 14

Discussion and Outlook 15

References 18

Appendix 21

SSM-iCrop crop input data 21
Field experimental management details 22
Hydrophysical property scenarios 23

Abstract

Agricultural management practices largely rely on empirical approaches reflecting farmers' experiences from past years. These empirical decisions are afflicted with uncertainty due to incomplete knowledge of the complex crop-environment-management interactions. In the face of climate change, these complexities are expected to become more adverse, leaving empirical management at risk for failure. With this study, we want to contribute to the development of an agricultural decision support system that facilitates current crop management to be more resource efficient and reduce GHG emissions from agriculture. We focus on synthetic nitrogen (N) fertilization as an application case of this tool and integrate a process-based crop model and model input data from different sources and of different availability as a prototype of a potential support tool. We ask how coarse/specific certain input needs to be, and how to attain this level of information, when only little output uncertainty is tolerable for a local management improvement.

To do so, we created a decision support prototype based on a multi-scenario simulation experiment for winter wheat in Eastern Austria as a case study region. We analyzed the simulated output against data collected from four field experiments within our study region. After reducing data uncertainty, we were able to create results potentially relevant for local fertilization improvements. As a cost-efficient means to achieve this reduction in uncertainty, we propose both a targeted involvement of users within our system and the integration of remote sensing data to further improve the performance, accuracy, and applicability of a decision support system such as our prototype.

ZVR 524808900

The authors gratefully acknowledge funding from IIASA and the National Member Organizations that support the institute (The Austrian Academy of Sciences; The Brazilian Federal Agency for Support and Evaluation of Graduate Education (CAPES); The National Natural Science Foundation of China (NSFC); The Academy of Scientific Research and Technology (ASRT), Egypt; The Finnish Committee for IIASA; The Association for the Advancement of IIASA, Germany; The Technology Information, Forecasting and Assessment Council (TIFAC), India; The Indonesian National Committee for IIASA; The Iran National Science Foundation (INSF); The Israel Committee for IIASA; The Japan Committee for IIASA; The National Research Foundation of Korea (NRF); The Mexican National Committee for IIASA; The Research Council of Norway (RCN); The Russian Academy of Sciences (RAS); Ministry of Education, Science, Research and Sport, Slovakia; The National Research Foundation (NRF), South Africa; The Swedish Research Council for Environment, Agricultural Sciences and Spatial Planning (FORMAS); The Ukrainian Academy of Sciences; The Research Councils of the UK; The National Academy of Sciences (NAS), USA; The Vietnam Academy of Science and Technology (VAST).

About the author

Marlene Palka is a PhD student at the Institute of Agronomy, University of Natural Resources and Life Sciences (BOKU), Vienna-Austria, where she focuses on the development and evaluation of decision support tools for adapting crop management to a changing climate in Austria. She received a Bachelor's degree in "Agricultural sciences" from BOKU and completed her Master studies in "Environmental protection and agricultural food production" at the University of Hohenheim, Stuttgart-Germany. (Contact: marlene.palka@boku.ac.at)

Acknowledgments

First, I am sincerely thankful to my two scientific YSSP supervisors, Rastislav Skalský and Juan Carlos Laso Bayas, for their time, thought, patience, and flexibility, supporting me both professionally and personally over the summer. This project would not have been possible without your help and confidence – even conceptually. I thank the Austrian Academy of Sciences for financial support as a YSSP participant. A great thank you to the Agro Drought Austria project team at BOKU, and the SATFARM-Services project team at IIASA (NODES) and BOKU for allowing me to use their datasets. I also thank Ahmad M. Manschadi and Josef Eitzinger for supervising my PhD in general and their help throughout the YSSP application process. I thank Martin Hofer for his technical support. Thank you to Aleksandra Cofala, Brian Fath, Tanja Huber, Fabian Wagner, and the rest of the YSSP team for organizing this program. And last, I am especially thankful to the 2022 YSSP cohort, who made this program an unforgettable experience, in particular Wenjia Cai, Tara Ippolito, and Hyun-Woo Jo, my wonderful and inspiring office colleagues.

Introduction

In recent years, the pressure on cropping systems to provide food in sufficient quantity and quality while moving to more sustainable, "greener" production has increased. Projected increases in frequency and severity of weather extremes such as droughts, heat waves, or frosts, together with other climate change related impacts, including altered precipitation patterns or pathogen infestations, are expected to pose a significant threat to global yields and food production (Leng and Hall, 2019; Savary et al., 2019; Spinoni et al., 2018; Zhao et al., 2017); at the same time, global agricultural production contributes up to 31 % of greenhouse gas emissions (OECD, 2016). In addition to CO^2 emissions from land use and land use change, cropping systems produce greenhouse gases through nitrous oxide emissions from inadequate synthetic nitrogen (N) fertilizer applications (Poore and Nemecek, 2018), and leaching of N to groundwater from over-fertilization can have negative impacts on aquatic ecosystems and other vegetation (Cameron et al., 2013; Liu et al., 2020; Lu et al., 2014). It is estimated that only 30–50 % of the applied N is actually used by crops (Godinot et al., 2016; Ladha et al., 2016; Swaney et al., 2018), leading to significant loss of applied N from agricultural fields, resulting in negative environmental impacts. In addition, the prospect of substantially higher costs for both N fertilizer production and application (due to increases in energy and fuel prices) puts farms at economic risk, when N is applied inadequately.

Consequently, the European Commission explicitly proposed "an obligatory nutrient management tool, designed to help farmers improve their water quality and reduce ammonia and nitrous oxide levels on their farms" (European Commision, 2018), pointing out the importance of improved nutrient (and specifically N) management in agriculture. Currently, agricultural N-fertilization practices largely rely on empirical approaches reflecting farmers' experiences from past years. These empirical decisions are afflicted with uncertainty due to incomplete knowledge of the complex crop-environment-management interactions, however. In the face of climate change, these complexities are expected to become more adverse (Eitzinger et al., 2013; IPCC, 2014; Jägermeyr et al., 2021; Ray et al., 2019), leaving empirical management at risk for failure. Another difficulty related to N management is that the final outcome of a cropping season (yields) determines how much fertilizer has to be applied at the beginning of the season and vice versa, such that the amount of fertilizer applied early during a year determines biomass and yield formation at the time of harvest.
At this point, crop models of different formats have proved versatile tools, translating the multitude of crop-environment-management complexities into more applicable management support. The advantage of process-based crop models compared to empirical, statistical, or any more recent AI related approaches is that by capturing the underlying biophysical input-output processes these models can predict a whole number of outputs, even ones other than the input, making them an integral tool to improve agricultural management in many cases (Laso Bayas et al., 2020; Machwitz et al., 2019; Morari et al., 2020). For example, the SSM-iCrop model (Soltani and Sinclair, 2012) has been extended and calibrated for winter wheat in Eastern Austria in previous studies (Ebrahimi et al., 2016; Manschadi et al., 2021; Manschadi et al., 2022).

Most process-based crop models are based on four categories of input data (at different resolutions and number of associated parameters): weather data, genetic characteristics of the grown crop, soil conditions, and management measures. For the case of N management this means that one can use above-ground dry matter simulations as a convenient proxy to adjust both fertilizer timing and amount, while only defining weather, crop, management, and soil conditions. However, the condition for the system to work is that the necessary input data is available to the user. In a best-case scenario, such a simulation tool is based on perfectly known

conditions: daily available weather data, well calibrated crop characteristics, precise management timing, and measurements of the prevailing soil conditions. Mostly because of resource-intensive measurement procedures, this is rarely the case, meaning that at least some degree of input uncertainty must be expected (Balkovič et al., 2020; Folberth et al., 2019b).
Given that, we want to understand for which input a certain degree of uncertainty can be accepted and what is inevitably required in order to improve agricultural fertilization strategies based on model simulations. To do so, we created a multi-scenario simulation experiment, but gradually moving from a scenario of maximum input certainty to a scenario of maximum input uncertainty. For the scenario set-up, we included the most important soil-related parameters and a range of feasible management strategies as factors, and compared the full range of simulation outcomes to ground-truth measurements from four winter wheat field experiments in the study region of Eastern Austria. This region is known as the breadbasket of the country, with winter wheat being the most important crop grown.

With this study, we want to contribute to the development of an agricultural decision support system that facilitates current crop management to be more resource efficient and reduce GHG emissions from agriculture. We focus on synthetic nitrogen (N) fertilization as an application case of this tool and integrate a process-based crop model and model input data from different sources and of different availability as a prototype of a potential support tool. We ask how coarse/specific certain input needs to be, and how to attain this level of information, when only little output uncertainty is tolerable for a local management improvement.

Material and methods

Description of the process-based crop model SSM-iCrop

For this project, we used the process-based, biophysical crop model SSM-iCrop, a simple simulation model, simulating phenology, growth, and yield formation of legume, maize, and grain crops under water- and N-limited growing conditions. It has been developed Soltani and Sinclair (2012), and tested, extended, and calibrated in previous studies to represent the cultivars and growing conditions of our study region in Eastern Austria (Manschadi et al., 2022). Values of all crop-related model inputs are listed in Table 2 in the Appendix. Compared to other bio-physical crop models like APSIM or EPIC, the SSM-iCrop model is a parsimonious point-scale crop model operating at the canopy level. It requires calibration of a relatively small number of crop input parameters, making it easy to use.
The aim of this study was to understand the importance of input data quality and availability when using SSM-iCrop as a translational tool to provide estimates of crop production and fertilizer demand at spatially extended scales to individual decision makers, rather than calibrating the underlying processes as represented by the model itself.

SSM-iCrop simulates crop phenology through a number of developmental stages using the biological day approach, where a biological day is a day with optimal temperature, photoperiod, and moisture conditions for plant development. Above-ground dry matter production is based on the radiation-use efficiency concept, and adjusted for temperature, water deficit, and atmospheric CO^2 concentration. Depending on the phenological stage, daily dry matter production is partitioned between three sinks: leaves, stems, and grains. Yield formation

is simulated based on a linear increase in harvest index, which is adjusted for pre-seed growth conditions. Crop mass at beginning seed growth is used as an indicator of these conditions. Seed growth rate is limited to crop photosynthesis and mobilized dry matter from vegetative organs.

For the simulation of soil water dynamics and other soil-related processes, SSM-iCrop uses a multilayer sub model, with a simple cascading bucket method. Plant-available water is calculated as the difference between soil water content of each layer and the layer water content at wilting point. SSM-iCrop calculates three water deficit factors from the fraction of the plant-available soil water capacity in the root zone to adjust dry mass production, leaf area expansion, and phenological development under water deficit conditions. Water addition from rainfall or irrigation and water removal due to run-off, evaporation, and drainage are accounted for. Plant water uptake is determined by root growth and plant-available water in the rooting zone. Root growth is simulated using the potential daily increase in root depth, limited by biological day. It is zero when daily dry matter production is zero, the soil layer is dry, and after beginning of seed growth. Transpiration is calculated from the daily rate of crop dry mass production, using a transpiration efficiency coefficient and vapor pressure deficit. Required soil input parameters are the hydrophysical properties of the profile (volumetric soil water content at saturation, drained upper limit (field capacity), crop lower limit (wilting point), air-dry, and soil coarse fraction and bulk density) and initial water content at sowing.

Similar to the soil water balance, SSM-iCrop simulates soil N balance in each layer as the result of processes that add N to the soil (mineralization of organic matter, fertilization) and those that remove it (volatilization, denitrification, leaching and crop N uptake). Crop N uptake from each soil layer depends on the fraction of the layer explored by roots and the ratio of N available to the crop from the soil solution in that layer to total available N in the root zone. The soil-N sub model requires the depth of soil layers, initial soil soluble N and soil organic N available for mineralization, as well as the time, amount, and volatilization fraction of each N fertilizer application as input parameters. For a full model description and code, see Soltani and Sinclair (2012).

Description of the reference field experiments

We used a set of four field experiments from three winter wheat growing seasons in the region of Eastern Austria as reference for this simulation study. The first experiment was a designed plot experiment located in Tulln and conducted during the 2017/2018 growing period. The closest weather station was "Langenlebarn". The second experiment was an on-farm experiment, conducted in Maria Roggendorf, conducted during 2020/21. The closest weather station was located in "Schöngrabern". The third and fourth experiment were both conducted during the 2021/22 growing season and both also on-farm experiments, one located in Hetzmannsdorf, with the closest weather station also in "Schöngrabern" and one in Potzneusiedl (Figure 1), the closest weather station being "Bruckneudorf". All four field experiments were conventionally managed and fertilized. For detailed experimental setup and dates, see Table 3 in the Appendix.

Figure 1: Location of the four winter wheat field experiments in Eastern Austria, used as ground-truth reference.

In addition to regular destructive sampling of above-ground dry matter, the experimental dataset also contains a time series of soil samples drilled until 120 cm depth, that have been analyzed for water retention properties, texture, and water and N content in 5 layers (0-10 cm, 20-30 cm, 30-60 cm, 60-90 cm, 90-120 cm, see Table 4 in the Appendix).

Soil at the Tulln site is classified as an alluvial chernozem, consisting of a silt to sandy loam in the topmost horizon (0–70 cm), a loamy sand with a high content of gravel and rock in the second horizon (70–80 cm), and a bedrock horizon (80-100 cm). The soil in Maria Roggendorf is classified as chernozem, consisting of a sandy clay loam in the first horizon (0-60 cm) and a loamy sand in the second horizon (60-120 cm). In Hetzmannsdorf, the soil is classified as colluvial chernozem, with a silt loam in the topmost horizon (0-30 cm) and a loam in the lower horizon (30-100 cm). At the Potzneusiedl site, the soil is also classified as a chernozem, consisting of a sandy loam (0-20 cm), a sandy loam with a low content of coarse material (20-70 cm), and a loamy sand with a low content of coarse material (70-90 cm), and a loamy sand with moderate content of coarse material (90-100 cm) (BFW, 2007).

Description of soil data sets

Besides point data from destructive soil samples, we also used information from gridded products for this study. The Austrian Digital Soil Map (eBod) is the digital representation of soil properties of Austrian area under agricultural cultivation. It is created, maintained and provided by the Federal Research and Training Center for Forests, Natural Hazards and Landscape (BFW) and available at no charge to the public. Its origins date back to the 1860s, when first analogue soil maps were drawn for areas around Vienna, successively been extended and digitized to date. The map is created from regularly taken, gridded field samples of drill cores (on average one sample per hectare), landform and topography, surface color and structure, and canopy cover; all combined into so-called soil groups. For each 100 hectares of one soil group a full profile is excavated to conduct pedological (e.g., definition of horizons, fraction of coarse material, color, etc.) and laboratory (texture, humus content, pH, etc.) analysis. eBod is provided as georeferenced vector tile cache at a 1 km^2 grid and includes information on soil type, arable quality, depth, humus content, porosity, texture, land use, pH etc. Full access to the data and a detailed description on all parameters available can be found at www.bodenkarte.at.

Within the Agro Drought Austria project (ada.boku.ac.at), the eBod data set has been extended and interpolated to also provide information on respective hydrophysical properties at a 500 x 500 m resolution. This product was created for drought monitoring activities of the project, using eBod in combination with pedotransfer functions to create additional layers, namely the lower limit and field capacity of both a topsoil (0-40 cm) and subsoil (40-100 cm) layer of the profile.

Introduction of input uncertainty

To determine input data requirements for using model simulations as N-management support, we created a multi-scenario simulation experiment using SSM-iCrop as the underlying process-based crop model and put special emphasis on soil-related model input. We created a total number of 13,824 simulated scenarios, gradually differing in uncertainty of the underlying input data (Table 1), moving from best-cases (=measurements, corresponding to scenario A for each factor) to datasets or assumptions of higher uncertainty.

The majority of processes determining crop growth depend on the prevailing soil conditions, while related parameters are cumbersome to determine and hard to collect at spatially extended scales. Available datasets are usually provided at coarse resolutions, giving reason to consider soil data as model input to be afflicted with the highest degree of uncertainty. We included all soil-related SSM-iCrop input parameters as factors, namely the hydrophysical properties, and initial water and N content at sowing.

For each of the hydrophysical scenarios, the entire profile was treated as one coherent, "spongy" complex, meaning that all respective soil hydrophysical properties were changed together. For scenario A (Table 4 in the Appendix), we used values of volumetric soil water content at saturation, field capacity, wilting point, air-dry, and soil coarse fraction and bulk density values as measured at experimental sites for the simulations. In scenario B (Table 5 in the Appendix), we used site-specific values from the combined eBod-Agro Drought Austria digital soil map intercepting experimental sites. In scenario C (Table 6 in the Appendix), we intended to account for regional variability within soils. Therefore, we created 16 "synthetic" soil profiles, as combinations of quartiles of topsoil organic carbon content and plant-available soil water capacity, respectively, using the combined eBod-Agro Drought Austria data within a 10 km radius from each experimental site. The hydrophysical properties of each of the 16 profiles were then based upon means within the quartile combination. For example, for the synthetic soil profile nr.1 in Hetzmannsdorf, we used the median of e.g., field capacity values from grid points that had topsoil organic carbon contents and plant-available soil water capacity values within the lowest quartile of all topsoil organic carbon content and plant-available soil water capacity values within the 10 km radius from the experimental site in Hetzmannsdorf. For the synthetic soil profile nr.2, we used the median of e.g., field capacity values from grid points that had topsoil organic carbon contents within the lowest quartile of all topsoil organic carbon contents and plant-available soil water capacity values within the second-lowest quartile of all plant-available soil water capacity values within the 10 km radius, and so on.

Since even the combined soil dataset only included SSM-iCrop relevant information on field capacity, wilting point, coarse fraction, and humus content in each layer, we further had to include some assumptions for setting up the simulations using hydrophysical properties from scenario B and C: volumetric water content at saturation was calculated as field capacity + 0.05 (He and Wang, 2019), air-dry water content was set wilting point * 0.6

until 30 cm depth and equal to wilting point below 30 cm depth. Soil organic N was calculated as humus/1.7 (conversion of humus content to total organic carbon TOC), and TOC/10 based on a common C/N ratio of 10 of soil organic matter content in arable land. The fraction of soil organic nitrogen available for mineralization was set to an average value of 0.15 (Soltani and Sinclair, 2012).

For the initial water content scenarios, we created conditions within the limits of each profile, reaching from completely dry (scenario B) to unlimited water availability to plants (scenario D). In scenario A, we used measured values at experimental sites. For scenario B, we used 0 % of plant-available soil water capacity, which is equal to the wilting point. This scenario represents growing conditions where no water is available to plants at sowing. Scenario C translates into 50 % of plant-available soil water capacity. For scenario D, we used soil field capacity equal 100 % of plant-available soil water capacity. The set-up of our initial water content scenarios indicates a dependency on the respective hydrophysical property setup of each profile, which was intentional to create more realistic conditions at each location.

We based scenarios of initial N content at sowing upon the range of recorded values across all field experiments. Since this range was already substantial (reaching from 0 kg N/ha to 200 kg N/ha), we believe that the respective values covered most of any potential regional variability, which we wanted to include in our scenario set-up. As usual, scenario A corresponds to measured soil mineral N contents at sowing of each experiment. For scenario B, we used 0 kg N/ha, the minimum recorded value across all field experiments. This scenario represents growing conditions where no N is retained from previous seasons. For scenario B, we used the average recorded soil mineral N value across all field experiments, namely 100 kg N/ha. This amount is also frequently used by farmers when determining their additional fertilizer requirements. We used 200 kg N/ha for scenario C, the maximum recorded value from our field experiments.

To account for different management strategies, we included scenarios of planting dates and the amount of N fertilizer applied within upper and lower limits of what is commonly practiced/allowed within our study region. For scenario A of sowing dates, we used actual dates according to experimental management. We created an early-planting scenario B, where wheat was sown 14 days earlier. Scenario C was the late-planting scenario, where wheat was sown 14 later compared to the actual sowing dates.

Finally, for different N fertilization scenarios, we used N amounts as applied to experiments for scenario A. We used 90 kg N/ha for scenario B, 90 kg N/ha, a low amount according to common agricultural practices in the study region. For scenario C, we used 150 kg N/ha, the average amount according to common agricultural practices in the study region, which is also the amount of N fertilizer recommended by the Austrian Ministry of Agriculture for winter wheat. In scenario D, we used 210 kg N/ha which maximum amount permitted for Austrian cropping regions (BMLFUW, 2017).
The inclusion of different N fertilization scenarios was not only intended to represent the full range of agricultural practices within our multi-scenario simulation study, but ultimately it is also the target variable that our simulation tool should support in determining. The different scenarios of N fertilizer applied therefore also represent potential N fertilization scenarios that can ultimately be translated into improved management measures.

Referring back to the previously mentioned four categories of input data for process-based crop models, weather data and genetic characteristics of the crop remain. Theoretically, any user could install an individual

weather station at each location simulations are run for. Since there already exists a dense network of weather stations across all Austria, which are also frequently maintained by the Austrian Meteorological Agency (ZAMG), we took weather data from each closest station as best-case scenario, and did not include any level uncertainty in this dimension since the data is accessible to the public at any time and at no costs. We also did not include any crop-related uncertainty, first, because SSM-iCrop is well calibrated for winter wheat and any potential changes of the parameters could only be reached through long-lasting breeding activities (which goes beyond the scope of this study).

Spatial extension of the prototype application

To create a spatially extended ground-truth comparison of simulations, where management input data is known, but soil input data is uncertain, we further simulated a dataset from the Austrian Chamber of Agriculture, that has been pre-processed in the SATFARM-Services project, containing 347 georeferenced winter wheat fields in Lower Austria, mostly south of Vienna, and in Burgenland. In addition to the location, the dataset contained information on the crop and cultivar grown, sowing and harvesting date, amount of N fertilizer applied, and final yields from the 2017/18 season until 2020. We used scenario B as best-case of the prevailing the hydrophysical properties, and scenario C as intermediate scenario for the initial water and N content at sowing. Converted yields into total above-ground dry matter (g/m^2, hereafter referred to as "DM") by applying a harvest index of 40 %, we compared simulated and recorded total-above ground dry matter values at the end of season.

As another means of spatial extension, we used remotely sensed data from the Sentinel-2 satellite of the Copernicus mission from the European Space Agency (scihub.copernicus.eu) to track DM development over the growing season non-destructively.

Statistical analysis

All comparisons are quantified on using the root mean squared error (RMSE), calculated as:

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(\hat{y}_i - y_i)^2}{n}} \quad (1)$$

whereby n is the number of simulations, $\hat{y}_i$are the simulated values and y_i are the observed values. Calculations were conducted in R programming language using the RStudio environment (R version 3.6.3). Graphs were also produced within RStudio.

Table 1: Definition of factors for multi-scenario simulations.

Input category	Input factor associated with uncertainty	Scenario A	Scenario B	Scenario C	Scenario D
Soil data	**hydrophysical properties of the soil**	measured values at experimental sites	site-specific values from the Austrian digital soil map (eBod) intercepting experimental sites	16 "synthetic" soil profiles, created within a 10 km radius from each experimental site as combinations of quartiles of topsoil organic carbon content and plant-available soil water capacity, respectively	-
	initial volumetric water content at sowing	measured values at experimental sites	0 % of plant-available soil water capacity, "wilting point"	50 % of plant-available soil water capacity	100 % of plant-available soil water capacity, "field capacity"
	initial soil mineral nitrogen content at sowing	measured values at experimental sites	0 kg N/ha, minimum recorded value for field experiments	100 kg N/ha, average recorded value for field experiments	200 kg N/ha, maximum recorded value for field experiments
Management measures	**Sowing date**	actual dates according to experimental management	14 days earlier	14 days later	-
	Amount of nitrogen fertilizer applied	according to experimental management	90 kg N/ha, low amount according to common agricultural practices in the study region	150 kg N/ha, average amount according to common agricultural practices in the study region	210 kg N/ha, maximum amount permitted for Austrian cropping regions

Results

For the presentation of our results, we will progress along the path for refining input data, starting from scenarios of maximum input uncertainty to best-case conditions with minimum input uncertainty.

Performance of the prototype

Our results in Figure 2 show that the prototype of a decision support tool based on SSM-iCrop as an example, simulated DM over the entire growing season and at all four experimental locations well (R^2=0.93). The hexbin plots also show darker hexagons close to the line of measurements, meaning that the majority of simulations were close to the measured values. However, this came at the expense of comparably large outcome variability when using input data of maximum uncertainty (RMSE=157.91 g/m^2). In some cases, the variability of final simulation outcomes was even twice the range of 30-year averages, represented as dashed gray lines in Figure 2 (in Tulln and Potzneusiedl, for example).

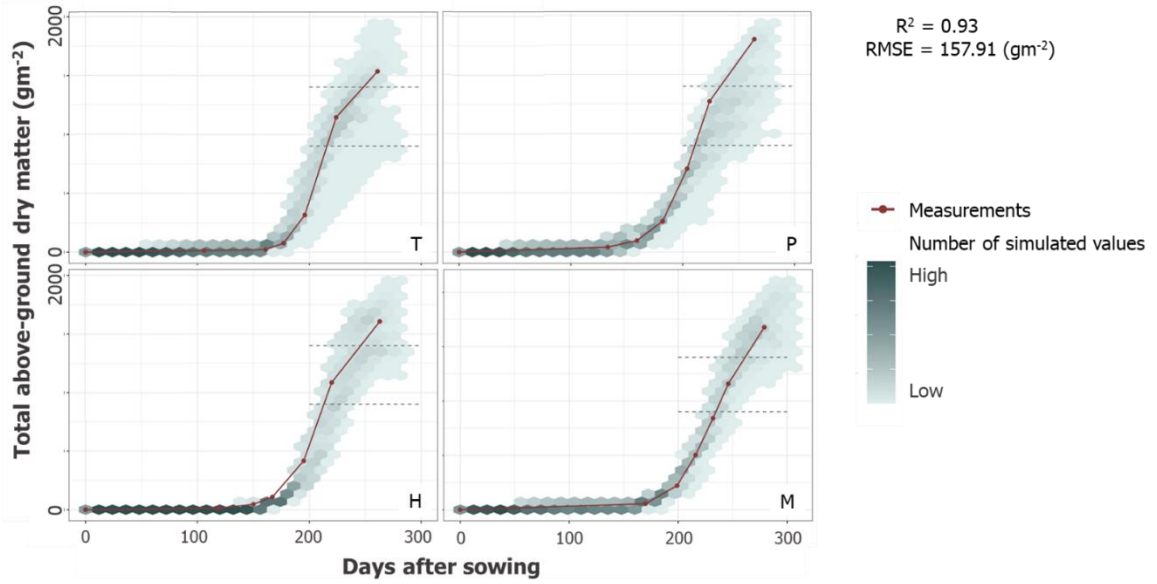

Figure 2: Performance of the prototype simulating DM of winter wheat under maximum input uncertainty across all factors. Individual graphs show results per field experiment (T=Tulln, P=Potzneusiedl, H=Hetzmannsdorf, M=Maria Roggendorf). Red markers and lines represent destructive measurements, tinted hexagons represent simulated values. The darker a hexagon, the higher the number of simulated values within the respective area. Horizontal dashed lines in gray show the 30-year average of DM of winter wheat in the study region.

There were some differences regarding the seasonal performance across experiments. For example, at the site in Maria Roggendorf, measured data were well centered among all potential outcomes, with no general trend of an over-or underestimation. In Hetzmannsdorf, the prototype generally underestimated DM early during the growing season and did not capture measured values during 150 to 200 days after sowing (DAS), while this evened out later towards harvest. Simulations for the Potzneusiedl and Tulln experiments framed measured values well early in the season, but were more likely to underestimate DM later on. Measured DM values at harvest were higher across all locations compared to the 30-year average of final DM production within the study region (900.85 to 1401.80 g/m^2), which can potentially be explained be the four field experiments being conducted within different growing seasons that also differed from the 30-year average. Under the right scenario

setup, the prototype was however able to capture these above-average seasons for all experiments. This result underlined the advantage of using a process-based crop model, namely including weather conditions for simulating crop DM and other outcomes.

Contribution of model input data uncertainty to simulation error

To reduce the previously mentioned variability of final DM simulations and determine which input factor is especially important for increasing the accuracy of our prototype, we determined RMSEs for each input factor individually. The heat map in Figure 3 shows RMSEs of input factors (x-axis) ranked according to highest RMSE across all locations, and experimental locations (y-axis) ranked according to highest RMSE across all input factors. We calculated RMSEs for individual input factors by defining scenario A for all other factors, while considering the full range of potential scenarios for the respective factor. For example, RMSE for initial water content was calculated using scenarios based on measured hydrophysical properties, initial N content, sowing dates, and fertilizer applications, while including all scenarios for initial water content. Results shown in Figure 3 underlined the importance of soil-related input (hydrophysial properties, initial soil water content, and initial mineral N content; shown as darker tiles), and also emphasized a masking effect of potential management benefits from simulated outcomes when soil-related input is unknown. Different scenarios of N fertilizer amounts (ultimately our target variable) accounted for the lowest RMSE values overall, meaning that the use of simulations based on inaccurate soil-related input data does not allow improving any N fertilization strategy.
As before, we observed some differences across experiments. Across all five input factors, simulations for Potzneusiedl showed highest RMSEs. Here the importance of soil-related input was especially pronounced. Simulations for Maria Roggendorf showed lowest RMSEs overall. This was also the site with no general trend of an over-or underestimation over the entire season (Figure 2). Interestingly, in Maria Roggendorf uncertainty related to hydrophysical properties did not contribute to simulation RMSE as much as in the other experiments.

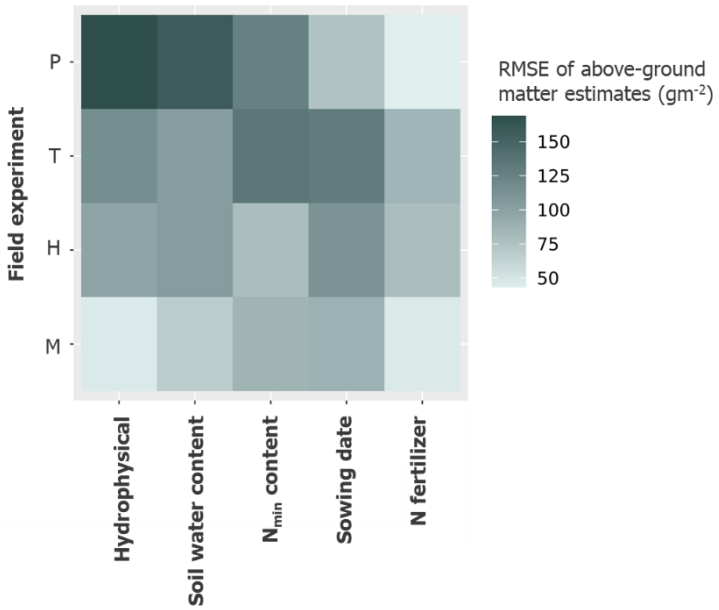

Figure 3: Heat map of RMSE associated to single input factors of the multi-scenario simulation experiment. The darker a tile, the higher the RMSE associated to one of the five factors (for details, see Table 1). Input factors (x-axis) are ranked according to highest RMSE across all locations, and experimental locations (y-axis) are ranked according to highest RMSE across all input factors.

Refining of model input data

In Figure 4, we present seasonal simulation results for a step-wise refining of input data according to the results shown in Figure 3 by the example of Potzneusiedl, the location of maximum RMSE for simulations of maximum uncertainty. We intentionally chose Potzneusiedl as an extreme case to show what degree of accuracy can be attained, even for high initial simulation error and variability. Compared to RMSEs shown in Figure 3, RMSEs shown here were calculated by a stepwise definition of scenario A for the refined factor(s) while considering the full range of potential scenarios for all non-refined factors.

Starting from maximum input uncertainty (RMSE=219.45 g/m^2, Figure 4.1), we first defined hydrophysical properties according to measured conditions in scenario A (Figure 4.2). Although RMSE did not decrease substantially (RMSE=197.74 g/m$^{2)}$, the density of simulated outcomes shifted closer to the line of measurements, shown by darker hexagons along the red line. In a next step, we defined the initial soil water content to be equal to scenario A (Figure 4.3), which resulted in a more pronounced error reduction (RMSE=131.52 g/m^2) and eliminated a group of low DM simulations. When we also refined the initial mineral N content of the soil (Figure 4.4), in other words considering scenarios where all three soil-related input factors used best-case scenario A, DM simulations followed the line of measurements very well already. This refining step eliminated the remainder of underestimating simulations, and RMSE was substantially reduced by almost half (RMSE=75.07 g/m^2). Refining the final two management steps (Figure 4.5 and .6), ultimately targeting an improved N management, further streamlined the simulated seasonal wheat growth, yielding a five-fold reduction of RMSE to 43.35 g/m^2 for the final step.

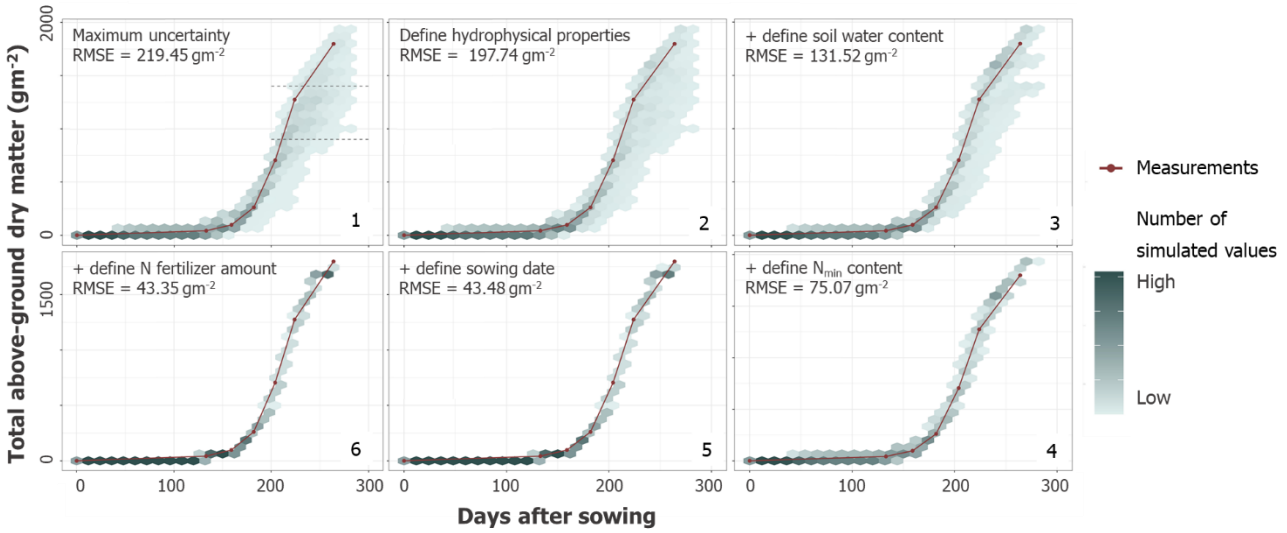

Figure 4: Performance of the prototype simulating DM of winter wheat under increased input certainty. Results are presented for the experiment in Potzneusiedl as example. Individual graphs (1-6) show results for a stepwise definition of input factors according to their importance (see Figure 3). Red markers and lines represent destructive measurements, tinted hexagons represent simulated values. The darker a hexagon, the higher the number of simulated values within the respective area. The horizontal dashed lines in gray show the 30-year average of DM of winter wheat in the study region.

In accordance with results shown in Figure 2, Figure 5 shows simulated DM over the entire growing season and at all four experimental locations when using input data of maximum certainty (scenario A for all factors exclusively). Results of this scenario essentially represent the performance of SSM-iCrop for simulating winter wheat at the experimental locations itself. Outcome variability was eliminated and DM production of winter

wheat reproduced very well. Although R^2 values were very high overall, maximum input certainty further increased R^2 to 0.94 across all for experiments. Compared to the simulation error under maximum input uncertainty, RMSE was reduced to 66.45 g/m^2, a third of the initial value.

Similar to our earlier results, there were some minor differences when comparing simulations of field experiments among each other. In Maria Roggendorf, simulations followed measured data well overall, but slightly underestimated DM at the end of season. Hetzmannsdorf showed some underestimations earlier during the vegetative growth period, which were eliminated later on. Simulations for the Potzneusiedl and Tulln experiments followed measured values well for the better part of the season, but showed a slight underestimation at the end of the season in Potzneusiedl. These results are in line with the simulation trends under maximum input uncertainty (see Figure 2).

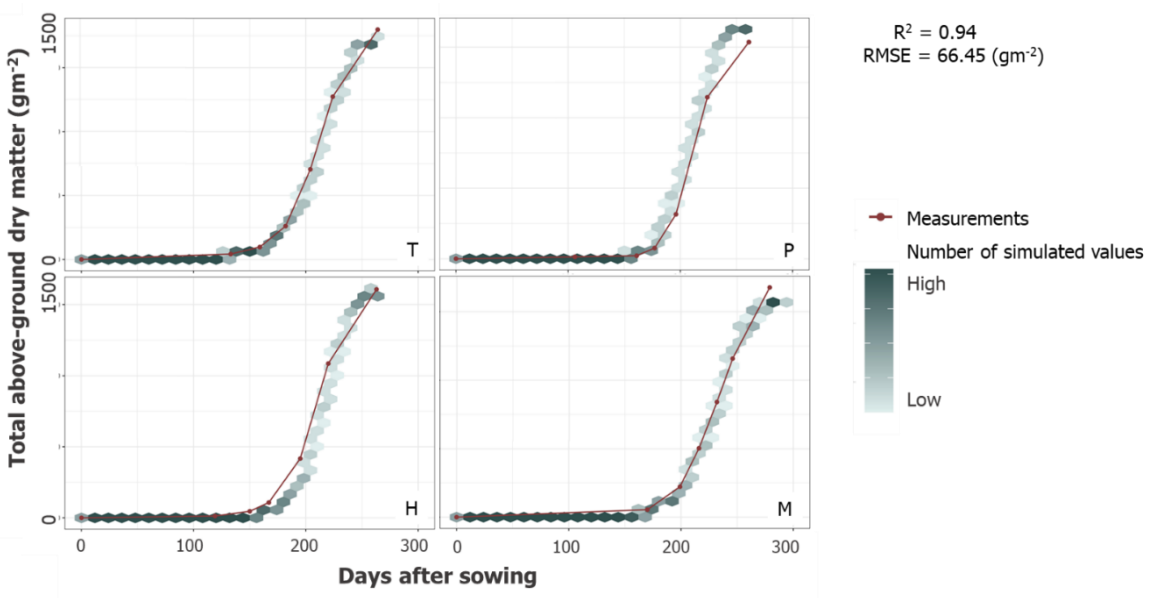

Figure 5: Performance of the prototype simulating DM of winter wheat under maximum input certainty (=best-case) across all factors. Individual graphs show results per field experiment (T=Tulln, P=Potzneusiedl, H=Hetzmannsdorf, M=Maria Roggendorf). Red markers and lines represent destructive measurements, tinted hexagons represent simulated values. The darker a hexagon, the higher the number of simulated values within the respective area.

As indicated, soil-related input uncertainty might disable a simulation-based management improvement. Simulation errors of related input uncertainty overruled management effects predicted by the prototype tool. Since a full laboratory description of field-specific soil conditions is usually not available to users, we included a first step towards refining generally available soil data (in our case the eBod-Agro Drought Austria combined digital soil map) and eventually improve the accuracy of a support system using such soil data as input. We focused on the hydrophysical property factor in this post-processing step, first, because it was contributing to simulation error the most, then, because the initial water content at sowing depends on the hydrophysical setup and is therefore already indirectly accounted for when refining the earlier. The third soil-related soil input factor, initial mineral N content, can only be determined by laboratory analyses and was therefore not included.

For this procedure, we assigned hydrophysical properties to each location that were closest to scenario A conditions a) in geographical space, which translates to scenario B (site-specific values from eBod) and b) in attribute space, which required a post-processing of scenario C. In this first step, we defined the attribute space as topsoil organic carbon content, determining a number of e.g., aeration, microbial, nutrient and water retention characteristics, and plant-available soil water capacity which largely determines DM production in water-limited cropping systems such as winter wheat production in Austria. We determined hydrophysical

properties closest in attribute space by choosing one out of the 16 synthetic profiles where topsoil organic carbon content and plant-available soil water capacity matched measured values best. In the next step, we included a step-wise refining of all remaining input data, but calculated simulation errors per soil scenario.

Figure 6 shows results of the step-wise refining of input data per hydrophysical property scenario. Similar to our previous refining pathway, simulation results gradually improved (RMSE decreased) with the definition of input data. Moving from maximum input uncertainty in step 1 (Figure 6.1), where also no hydrophysical post-processing was included, the definition of hydrophysical properties slightly reduced RMSE (Figure 6.2). In this case, where all other input data was still maximally uncertain, hydrophysical input data from less accurate sources (scenario B and C) yielded even smaller simulation errors compared to scenario A simulations. Scenario A related RMSE outliers could be explained by extreme combinations of scenarios of other input factors, that exceeded the capability of SSM-iCrop representing related biophysical processes when using scenario B and post-processed scenario C hydrphysical property input data. With the definition of initial water content (Figure 6.3) RMSE of scenario A simulations however, dropped compared to scenario B and post-processed scenario C simulations. As before, RMSE further decreased as we consecutively defined the initial N content (Figure 6.4), and the definition of sowing date (Figure 6.5) and the amount of N fertilizer applied (Figure 6.6) yielded last minor improvements.
Regarding the post-processing of generally available soil data, simulations using scenario B (geographically closest) yielded lower RMSEs compared to post-processed scenario C simulations (closest in attribute space). But neither reached the accuracy of using the best-case scenario A of hydrophysical properties as input. These results indicate that we might have oversimplified the matching conditions for defining soil profiles closest in attribute space. Further including soil texture, aggregate composition, or bulk density into the matching procedure might improve these results in the future.

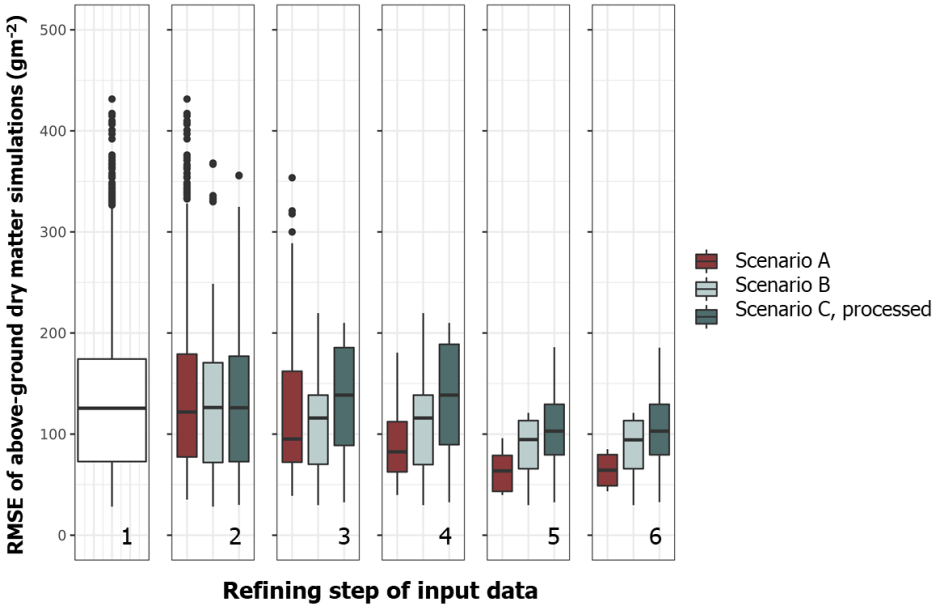

Figure 6: Performance of the prototype simulating DM of winter wheat under increased input certainty. Results are presented for each hydrophysical property scenario individually. Graph 1 shows the full range of potential outcomes. In graph 2, hydrophysical properties are defined. Graph 3 further includes measured initial soil water content. In graph 4, the initial soil mineral N content is defined. In graph 5, also the measured sowing date is used. For graph 6, the amount of applied N fertilizer is finally defined.

Spatial extension

Simulating end-of season DM of the SATFARM-Services dataset, which contained 347 winter wheat fields in Lower Austria and Burgenland, showed, that in cases where management input data is known, but soil input data is uncertain, the prototype failed to estimate measured values (R^2=0.05). Given that –besides the location of fields- the dataset did not contain any information on soil-related input data, this result only confirmed our previous finding of the overruling importance of accurate soil input to the system. In order to simulate the respective DM records, we had to use scenario B as best-case of the prevailing the hydrophysical properties, and scenario C as intermediate scenario for the initial water and N content at sowing. It is also important to mention that the original dataset contained records on grain yield because continuous destructive DM measurements over the season are extremely laborious and in most cases not feasible outside of an experimental setting. We converted yield into DM by applying a harvest index of 40%. This might have introduced another level of uncertainty contributing to the low R^2 value. Finally, compared to the results presented in Figure 5 where final DM estimates deviated from measured values of some field experiments, we expected a decrease in simulation accuracy when simulating end-of-season DM only.

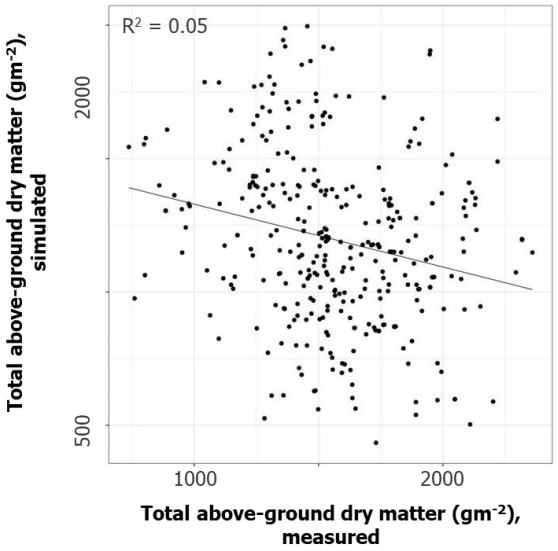

Figure 7: Measured vs. simulated end-of-season DM of winter wheat across 347 fields in Lower Austria and Burgenland, given maximum soil-related input uncertainty. Individual dots represent simulated end-of season DM at each site. The solid line marks the best fit regression line.

To overcome this difficulty of a) working with only one yield record at the end of the season, and b) non-feasible DM measurements over the course of the entire growth period, we fitted an exponential model (R^2=0.94, Equation 2) to remotely-sensed values of the difference vegetation index (DVI, Equation 3, Skakun et al. (2019)) of the three on-farm experiments in Maria Roggendorf, Hetzmannsdorf, and Potzneusiedl. Spectral data for Tulln were not available, since plots were too small for the current Sentinel-2 resolution of 20 m grid size. This allowed us to track DM development over the growing season non-destructively over large areas.

$$DM = 6.6021 * e^{0.0012DVI} \quad (2)$$

$$DVI = NIR - Red \quad (3)$$

One shortcoming of the DVI is, that it works for DM estimation of green canopies only. For the application case of N fertilization, this however less of a disadvantage, since the N is applied early during the vegetative growth stage. Our results showed that during that time, DVI was a capturing destructively measured DM values at all three experimental sites, again with some differences across locations. Interestingly, in Hetzmannsdorf, DVI estimates capture measured DM values even better than the prototype simulations, where early season values were underestimated. In Maria Roggendorf, where the best-case prototype simulations captured early-season DM very well, DVI estimates deviated more substantially, however (see Figure 5 for comparison).

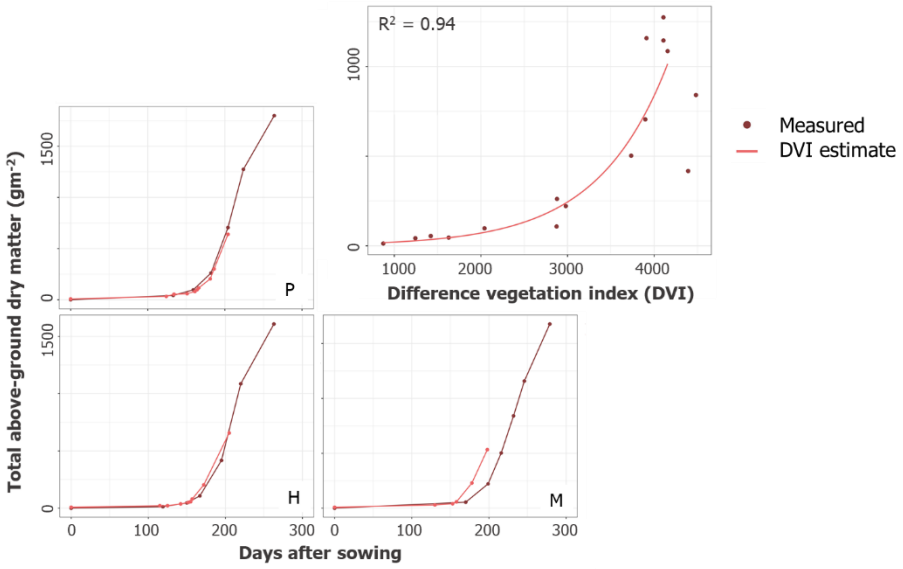

Figure 8: DM production in Potzneusiedl (P), Hetzmannsdorf (H), and Maria Roggendorf (M), based on an exponential DVI model (top right). Markers and lines in red represent destructive measurements, markers and lines in pink DVI estimates.

Discussion and Outlook

The aim of this study was to contribute to the development of an agricultural decision support system that facilitates crop N fertilization to be more resource efficient and reduce GHG emissions from agriculture. As a prototype of a potential support tool, we used a process-based crop model based on model input data of different sources and availability, and put special emphasis on soil-related input.
In the following, we will discuss the strengths and weaknesses of our prototype, and how to attain a required level of information, when the prototype is used to support decisions of N fertilization and only little output uncertainty is tolerable for a local management improvement. We also outline next steps to improve the system in the future.

Under the condition of having access to accurate soil information, our prototype of a decision support tool yielded information that did not only outperform regional long-term average DM production averages frequently used to determine empirical fertilization strategy empirically, but also captured site-specific crop growth over the entire season at various under various pedoclimatic growing environments in Eastern Austria. This is especially important for an improved N management, since fertilizer is applied early in the season when DM production is still low and substantial amounts of N could be lost to the environment when applied inadequately. We used SSM-iCrop as an example of a process-based biophysical crop model and performed our analysis by

the example of N fertilization of winter wheat grown in Austria. Conceptually, SSM-iCrop could be replaced by any other process-based crop model, winter wheat by any other crop in any other region, used to support any other management measure (Balkovič et al., 2020; Beveridge et al., 2018; Wang et al., 2022).
The results of this study however also showed that simulations can support N management only when soil-related input data is available, especially the hydrophysical properties. Other management practices such as irrigation might alter the results. Given the high dependency of most agronomic processes on the prevailing soil conditions (He and Wang, 2019; He et al., 2017), we expect a similar importance of soil-related model input, however. For the common case, that specific information is unknown, we also tested a first, simple approach to post-process generally available hydrophysical data. Although we expect improved simulation results of the prototype as soon as we advance this post-processing, uncertainty related to other soil input such as the initial water and N content at sowing (or even more advanced parameters when using more advanced process-based models) will remain. Some of it could be eliminated using other model simulations or intense data mining (Folberth et al., 2019a; Le Bourgeois et al., 2016; Piniewski et al., 2019; Schauberger et al., 2020; Tóth et al., 2015; Tóth et al., 2017; Vogeler et al., 2019) but even then, this information will come at a certain resolution only and will most probably be afflicted with some remaining uncertainty (Carr et al., 2020). For eBod, the digital soil map of Austria provided at a comparably high resolution of 1 x 1 km, for example, we showed the remaining uncertainty and its implications when translated into applied management measures at individual sites.

At this point, we want to propose two mutually complementing ways of how to attain a required level of information when growing condition and hence input data for the prototype are lesser known compared to our reference field experiments.
First, integrating bottom-up users such as farmers or extension officers into the processing chain might provide a cost-effective means to substantially improve the performance of such a support tool. Some of the factors we tested within the multi-scenario simulation experiment, such as planting day, are usually known by farmers and come with no additional costs. Soil water content or even some hydrophysical properties such as texture could be determined by simple field tests. Determining soil N content requires laboratory analyses. The average price for getting one soil sample analyzed for organic and mineral N content are between 20 and 30 € in Austrian laboratories. By integrating farmers into the processing chain, we could further benefit from their personal experience regarding the prevailing growing conditions. A promising way to do so could be the use of mobile applications, that can integrate all data available and directly translate it into management information. Similar systems have already been developed (Laso Bayas et al., 2020; Laso Bayas et al., 2017; Machwitz et al., 2019).
Second, we propose the use of remote sensing data (Jin et al., 2017). As described, most measurements to accurately determine simulation inputs are cumbersome procedures and might even be impossible as a spatially extended top-down approach. Remotely sensed canopy reflectance data are easily available (Lesiv et al., 2018) and could help define the trajectory of potential simulation outcomes, especially early in the season, when canopies are still green and a variety of spectral models can be used to track crop growth and biomass accumulation. Today, several methods exist to narrow down the range of potential simulation outcomes, for example by the reverse integration of radiative transfer models (Li et al., 2018; Machwitz et al., 2014) or machine learning algorithms (Cai et al., 2019; Evans and Shen, 2021). This has already proved successful in the past (Colaço and Bramley, 2018; Morari et al., 2020). Our spectral model was based on Skakun et al. (2019) who found the DVI as the best estimator when incorporating surface reflectance and phenological fitting into their yield regression models. Other indices might even perform better or over an extended period of time (Aranguren et al., 2020; Fang et al., 2021; Quemada et al., 2019).

We believe that especially the combination of both the bottom-up user involvement and top-down remote sensing technology will help to further improve the performance, accuracy, and applicability of a decision support system such as our prototype.

Finally, some questions remain, for example: Would benefits from using our prototype compensate potential costs to refine input data? What is the value of information provided by such a system, and how to quantify it? Finding answers goes beyond the scope and purpose of this specific study, leaving room for future work, however.

References

Aranguren, M., Castellón, A., and Aizpurua, A. (2020). Crop sensor based non-destructive estimation of nitrogen nutritional status, yield, and grain protein content in wheat. *Agriculture (Switzerland)* **10**.

Balkovič, J., Madaras, M., Skalský, R., Folberth, C., Smatanová, M., Schmid, E., van der Velde, M., Kraxner, F., and Obersteiner, M. (2020). Verifiable soil organic carbon modelling to facilitate regional reporting of cropland carbon change: A test case in the Czech Republic. *Journal of Environmental Management* **274**.

Beveridge, L., Whitfield, S., and Challinor, A. (2018). Crop modelling: towards locally relevant and climate-informed adaptation. *Climatic Change* **147**, 475-489.

BFW (2007). eBod Digitale Bodenkarte. Bundesforschungs- und Ausbildungszentrum für Wald, Naturgefahren und Landschaft.

BMLFUW (2017). "Richtline für die sachgerechte Düngung im Ackerbau und Grünland. Anleitung zur Interpretation von Bodenuntersuchungsergebnissen in der Landwirtschaft." Bundesministerium für Land- und Forstwirtschaft, Umwelt und Wasserwirtschaft, Vienna, Austria.

Cai, Y., Guan, K., Lobell, D., Potgieter, A. B., Wang, S., Peng, J., Xu, T., Asseng, S., Zhang, Y., You, L., and Peng, B. (2019). Integrating satellite and climate data to predict wheat yield in Australia using machine learning approaches. *Agricultural and Forest Meteorology* **274**, 144-159.

Cameron, K. C., Di, H. J., and Moir, J. L. (2013). Nitrogen losses from the soil/plant system: A review. *Annals of Applied Biology* **162**, 145-173.

Carr, T. W., Balkovič, J., Dodds, P. E., Folberth, C., Fulajtar, E., and Skalsky, R. (2020). Uncertainties, sensitivities and robustness of simulated water erosion in an EPIC-based global gridded crop model. *Biogeosciences* **17**, 5263-5283.

Colaço, A. F., and Bramley, R. G. V. (2018). Do crop sensors promote improved nitrogen management in grain crops? *Field Crops Research* **218**, 126-140.

Ebrahimi, E., Manschadi, A. M., Neugschwandtner, R. W., Eitzinger, J., Thaler, S., and Kaul, H. P. (2016). Assessing the impact of climate change on crop management in winter wheat - A case study for Eastern Austria. *Journal of Agricultural Science* **154**, 1153-1170.

Eitzinger, J., Trnka, M., Semerádová, D., Thaler, S., Svobodová, E., Hlavinka, P., Šiška, B., Takáč, J., Malatinská, L., Nováková, M., Dubrovský, M., and Žalud, Z. (2013). Regional climate change impacts on agricultural crop production in Central and Eastern Europe - Hotspots, regional differences and common trends. *Journal of Agricultural Science* **151**, 787-812.

Evans, F. H., and Shen, J. (2021). Long-term hindcasts of wheat yield in fields using remotely sensed phenology, climate data and machine learning. *Remote Sensing* **13**.

Fang, P., Yan, N., Wei, P., Zhao, Y., and Zhang, X. (2021). Aboveground biomass mapping of crops supported by improved casa model and sentinel-2 multispectral imagery. *Remote Sensing* **13**.

Folberth, C., Baklanov, A., Balkovič, J., Skalský, R., Khabarov, N., and Obersteiner, M. (2019a). Spatio-temporal downscaling of gridded crop model yield estimates based on machine learning. *Agricultural and Forest Meteorology* **264**, 1-15.

Folberth, C., Elliott, J., Müller, C., Balkovič, J., Chryssanthacopoulos, J., Izaurralde, R. C., Jones, C. D., Khabarov, N., Liu, W., Reddy, A., Schmid, E., Skalský, R., Yang, H., Arneth, A., Ciais, P., Deryng, D., Lawrence, P. J., Olin, S., Pugh, T. A. M., Ruane, A. C., and Wang, X. (2019b). Parameterization-induced uncertainties and impacts of crop management harmonization in a global gridded crop model ensemble. *PLoS ONE* **14**.

Godinot, O., Leterme, P., Vertès, F., and Carof, M. (2016). Indicators to evaluate agricultural nitrogen efficiency of the 27 member states of the European Union. *Ecological Indicators* **66**, 612-622.

He, D., and Wang, E. (2019). On the relation between soil water holding capacity and dryland crop productivity. *Geoderma* **353**, 11-24.

He, D., Wang, E., Wang, J., and Robertson, M. J. (2017). Data requirement for effective calibration of process-based crop models. *Agricultural and Forest Meteorology* **234-235**, 136-148.

IPCC (2014). "AR5 Climate Change 2014: Impacts, Adaptation, and Vulnerability," Cambridge, United Kingdom and New York, NY , USA.

Jägermeyr, J., Müller, C., Ruane, A. C., Elliott, J., Balkovic, J., Castillo, O., Faye, B., Foster, I., Folberth, C., Franke, J. A., Fuchs, K., Guarin, J. R., Heinke, J., Hoogenboom, G., Iizumi, T., Jain, A. K., Kelly, D., Khabarov, N., Lange, S., Lin, T. S., Liu, W., Mialyk, O., Minoli, S., Moyer, E. J., Okada, M., Phillips, M., Porter, C., Rabin, S. S., Scheer, C., Schneider, J. M., Schyns, J. F., Skalsky, R., Smerald, A., Stella, T.,

Stephens, H., Webber, H., Zabel, F., and Rosenzweig, C. (2021). Climate impacts on global agriculture emerge earlier in new generation of climate and crop models. *Nature Food* **2**, 873-885.
Jin, X., Li, Z., Guijun, Y., Yang, H., Feng, H., Xu, X., Wang, J., li, X., and Luo, J. (2017). Winter wheat yield estimation based on multi-source medium resolution optical and radar imaging data and the AquaCrop model using the particle swarm optimization algorithm. *ISPRS Journal of Photogrammetry and Remote Sensing* **126**, 24–37.
Ladha, J. K., Tirol-Padre, A., Reddy, C. K., Cassman, K. G., Verma, S., Powlson, D. S., van Kessel, C., de B Richter, D., Chakraborty, D., and Pathak, H. (2016). Global nitrogen budgets in cereals: A 50-year assessment for maize, rice, and wheat production systems. *Scientific reports* **6**, 19355-19355.
Laso Bayas, J. C., Gardeazabal, A., Karner, M., Folberth, C., Vargas, L., Skalský, R., Balkovič, J., Subash, A., Saad, M., Delerce, S., Cuaresma, J. C., Hlouskova, J., Molina-Maturano, J., See, L., Fritz, S., Obersteiner, M., and Govaerts, B. (2020). Agrotutor: A mobile phone application supporting sustainable agricultural intensification. *Sustainability (Switzerland)* **12**, 1-10.
Laso Bayas, J. C., Lesiv, M., Waldner, F., Schucknecht, A., Duerauer, M., See, L., Fritz, S., Fraisl, D., Moorthy, I., McCallum, I., Perger, C., Danylo, O., Defourny, P., Gallego, J., Gilliams, S., Akhtar, I. U. H., Baishya, S. J., Baruah, M., Bungnamei, K., Campos, A., Changkakati, T., Cipriani, A., Das, K., Das, K., Das, I., Davis, K. F., Hazarika, P., Johnson, B. A., Malek, Z., Molinari, M. E., Panging, K., Pawe, C. K., Pérez-Hoyos, A., Sahariah, P. K., Sahariah, D., Saikia, A., Saikia, M., Schlesinger, P., Seidacaru, E., Singha, K., and Wilson, J. W. (2017). A global reference database of crowdsourced cropland data collected using the Geo-Wiki platform. *Scientific Data* **4**.
Le Bourgeois, O., Bouvier, C., Brunet, P., and Ayral, P. A. (2016). Inverse modeling of soil water content to estimate the hydraulic properties of a shallow soil and the associated weathered bedrock. *Journal of Hydrology* **541**, 116-126.
Leng, G., and Hall, J. (2019). Crop yield sensitivity of global major agricultural countries to droughts and the projected changes in the future. *Science of the Total Environment* **654**, 811-821.
Lesiv, M., See, L., Bayas, J. C. L., Sturn, T., Schepaschenko, D., Karner, M., Moorthy, I., McCallum, I., and Fritz, S. (2018). Characterizing the spatial and temporal availability of very high resolution satellite imagery in Google Earth and Microsoft Bing Maps as a source of reference data. *Land* **7**.
Li, Z., Jin, X., Yang, G., Drummond, J., Yang, H., Clark, B., Li, Z., and Zhao, C. (2018). Remote sensing of leaf and canopy nitrogen status in winter wheat (Triticum aestivum L.) based on N-PROSAIL model. *Remote Sensing* **10**.
Liu, C., Wang, L., Cocq, K. L., Chang, C., Li, Z., Chen, F., Liu, Y., and Wu, L. (2020). Climate change and environmental impacts on and adaptation strategies for production in wheat-rice rotations in southern China. *Agricultural and Forest Meteorology* **292-293**.
Lu, X., Mao, Q., Gilliam, F. S., Luo, Y., and Mo, J. (2014). Nitrogen deposition contributes to soil acidification in tropical ecosystems. *Global Change Biology* **20**, 3790-3801.
Machwitz, M., Giustarini, L., Bossung, C., Frantz, D., Schlerf, M., Lilienthal, H., Wandera, L., Matgen, P., Hoffmann, L., and Udelhoven, T. (2014). Enhanced biomass prediction by assimilating satellite data into a crop growth model. *Environmental Modelling and Software* **62**, 437-453.
Machwitz, M., Hass, E., Junk, J., Udelhoven, T., and Schlerf, M. (2019). CropGIS – A web application for the spatial and temporal visualization of past, present and future crop biomass development. *Computers and Electronics in Agriculture* **161**, 185-193.
Manschadi, A. M., Eitzinger, J., Breisch, M., Fuchs, W., Neubauer, T., and Soltani, A. (2021). Full Parameterisation Matters for the Best Performance of Crop Models: Inter-comparison of a Simple and a Detailed Maize Model. *International Journal of Plant Production* **15**, 61-78.
Manschadi, A. M., Palka, M., Fuchs, W., Neubauer, T., Eitzinger, J., Oberforster, M., and Soltani, A. (2022). Performance of the SSM-iCrop model for predicting growth and nitrogen dynamics in winter wheat. *European Journal of Agronomy* **135**.
Morari, F., Zanella, V., Gobbo, S., Bindi, M., Sartori, L., Pasqui, M., Mosca, G., and Ferrise, R. (2020). Coupling proximal sensing, seasonal forecasts and crop modelling to optimize nitrogen variable rate application in durum wheat. *Precision Agriculture*.
OECD (2016). "Agriculture and Climate Change: Towards Sustainable, Productive and Climate-Friendly Agricultural Systems. Background note 4." Organisation for Economic Cooperation and Development, Paris, France.
Piniewski, M., Bieger, K., and Mehdi, B. (2019). Advancements in Soil and Water Assessment Tool (SWAT) for ecohydrological modelling and application. *Ecohydrology and Hydrobiology* **19**, 179-181.

Poore, J., and Nemecek, T. (2018). Reducing food's environmental impacts through producers and consumers. *Science* **360**, 987.
Quemada, M., Pancorbo, J. L., Alonso-Ayuso, M., Gabriel, J. L., López-Herrera, J., and Pérez-Martín, E. (2019). Vegetation indices from remote sensing imagery as proxies for yield and grain N in wheat. *In* "Precision Agriculture 2019 - Papers Presented at the 12th European Conference on Precision Agriculture, ECPA 2019", pp. 323-330.
Ray, D. K., West, P. C., Clark, M., Gerber, J. S., Prishchepov, A. V., and Chatterjee, S. (2019). Climate change has likely already affected global food production. *PLOS ONE* **14**, e0217148.
Savary, S., Willocquet, L., Pethybridge, S. J., Esker, P., McRoberts, N., and Nelson, A. (2019). The global burden of pathogens and pests on major food crops. *Nature Ecology and Evolution* **3**, 430-439.
Schauberger, B., Jägermeyr, J., and Gornott, C. (2020). A systematic review of local to regional yield forecasting approaches and frequently used data resources. *European Journal of Agronomy* **120**.
Skakun, S., Vermote, E., Franch, B., Roger, J. C., Kussul, N., Ju, J., and Masek, J. (2019). Winter wheat yield assessment from Landsat 8 and Sentinel-2 data: Incorporating surface reflectance, through phenological fitting, into regression yield models. *Remote Sensing* **11**.
Soltani, A., and Sinclair, T. R. (2012). Modeling physiology of crop development, growth and yield. *Modeling Physiology of Crop Development, Growth and Yield*, 1-322.
Spinoni, J., Vogt, J. V., Naumann, G., Barbosa, P., and Dosio, A. (2018). Will drought events become more frequent and severe in Europe? *International Journal of Climatology* **38**, 1718-1736.
Swaney, D. P., Howarth, R. W., and Hong, B. (2018). Nitrogen use efficiency and crop production: Patterns of regional variation in the United States, 1987–2012. *Science of The Total Environment* **635**, 498-511.
Tóth, B., Weynants, M., Nemes, A., Makó, A., Bilas, G., and Tóth, G. (2015). New generation of hydraulic pedotransfer functions for Europe. *European Journal of Soil Science* **66**, 226-238.
Tóth, B., Weynants, M., Pásztor, L., and Hengl, T. (2017). 3D soil hydraulic database of Europe at 250 m resolution. *Hydrological Processes* **31**, 2662-2666.
Vogeler, I., Carrick, S., Cichota, R., and Lilburne, L. (2019). Estimation of soil subsurface hydraulic conductivity based on inverse modelling and soil morphology. *Journal of Hydrology* **574**, 373-382.
Wang, X., Folberth, C., Skalsky, R., Wang, S., Chen, B., Liu, Y., Chen, J., and Balkovic, J. (2022). Crop calendar optimization for climate change adaptation in rice-based multiple cropping systems of India and Bangladesh. *Agricultural and Forest Meteorology* **315**.
Zhao, C., Liu, B., Piao, S., Wang, X., Lobell, D. B., Huang, Y., Huang, M., Yao, Y., Bassu, S., Ciais, P., Durand, J.-L., Elliott, J., Ewert, F., Janssens, I. A., Li, T., Lin, E., Liu, Q., Martre, P., Müller, C., Peng, S., Peñuelas, J., Ruane, A. C., Wallach, D., Wang, T., Wu, D., Liu, Z., Zhu, Y., Zhu, Z., and Asseng, S. (2017). Temperature increase reduces global yields of major crops in four independent estimates. *Proceedings of the National Academy of Sciences* **114**, 9326.

Appendix

SSM-iCrop crop input data

Table 2: SSM-iCrop crop input data for winter wheat grown in Austria. Adapted and calibrated based on Soltani and Sinclair (2012).

Parameter	Description	Unit	Value
a_plapow_d	A coefficient (exponent) in power relationship between plant leaf area and mainstem node number	-	1.1718
b_plapow_d	A coefficient (exponent) in power relationship between plant leaf area and mainstem node number	-	-0.0006
bdANTPM	Biological days from anthesis to physiological maturity	biol. day	32.4
bdBOTEAR	Biological days from booting to ear emergence	biol. day	2.5
bdEARANT	Biological days from ear emergence to anthesis	biol. day	4
bdEMRTIL	Biological days from emergence to first-tiller	biol. day	1.2
bdPMHM	Biological days from physiological maturity to harvest maturity	biol. day	8
bdSELBOT	Biological days from first-node to booting (ligule of flag leaf visible)	biol. day	12.8
bdSOWEMR	Biological days from sowing to emergence	biol. day	4.1
bdTILSEL	Biological days from first-tiller to first-node (stem-elongation)	biol. day	12.1
CO2RUE	Coefficient to adjust RUE for higher (and lower) atmospheric CO^2 concentration	-	0.8
CPP	Critical photoperiod	hour	16
FLDKL	Killing no. of consecutive flooding	day	20
FLF1A	Partitioning coefficient to leaves during main phase of leaf area development at lower levels of total crop mass	g/g	0.7829
FLF1B	Partitioning coefficient to leaves during main phase of leaf area development at higher levels of total crop mass	g/g	0.2773
FLF2	Partitioning coefficient to leaves from termination leaf growth on mainstem to beginning seed growth	g/g	0.1
FRTRL	Fraction crop mass at beginning seed growth which is translocateble to grains	g/g	0.22
FRZLDR	Fraction of leaf destruction below the critical by each degree centigrade	$m^2/m^2/$ ^oC	0.01
GCC	Grain conversion coefficient	g/g	1
GNCmax	Maximum grain nitrogen concentration	g/g	0.026
GNCmin	Minimum grain nitrogen concentration	g/g	0.016
GRTDP	Potential daily increase (growth) in root depth	mm/ biol. day	50
iDEPORT	Depth of roots at emergence	mm	200
IRUE	physiological potential radiation use efficiency	g/MJ	2.58
KPAR	Extinction coefficient for photosynthetically active radiation (PAR)	-	0.65
leaf_number_phyl2	Number of leafs on main stem for activating PHYL2	-	6.4
MEED	Maximum effective depth of water extraction from soil by roots	mm	1200
MXNUP	Maximum uptake (fixation) rate of nitrogen	$g/m^2/day$	0.57143
PDHI	Potential slope of harvest index (DHI)	g/g/day	0.017
PHYL1	Phyllochron	-	79
PHYL2	Phyllochron	-	93

PLACON	A coefficient (constant) in power relationship between plant leaf area and mainstem node number	-	0.3021
PLAPOW	A coefficient (exponent) in power relationship between plant leaf area and mainstem node number	-	2.7357
PPSEN	Photoperiod sensitivity coefficient	-	0.004
SLA	Specific leaf area	m^2/g	0.021
SLNG	Specific leaf nitrogen in green leaves (target)	g/m^2	2.53
SLNS	Specific leaf nitrogen in senesced leaves (minimum)	g/m^2	0.2
SNCG	Stem nitrogen concentration in green stems (target)	g/g	0.022
SNCS1	Stem nitrogen concentration in senesced stems (minimum)	g/g	0.01
SNCS2	Stem nitrogen concentration in senesced stems (maximum)	g/g	0.002
TBD	Base temperature for development	^oC	0
TBRUE	Base temperature for dry matter production	^oC	0
TCD	Ceiling temperature for development	^oC	40
TCRUE	Ceiling temperature for dry matter production	^oC	35
TEC350	Transpiration efficiency coefficient (350 Pa)	Pa	5.8
TEC700	Transpiration efficiency coefficient (700 Pa)	^oC	1.37
TP1D	Lower optimum temperature for development	^oC	27.5
TP1RUE	Lower optimum temperature for dry matter production	^oC	10
TP2RUE	Upper optimum temperature for dry matter production	^oC	22
VSEN	Vernalization sensitivity coefficient	-	0.03
WDHI1	A critical point for seed growth rate	g/m^2	0
WDHI2	A critical point for seed growth rate	g/m^2	600
WDHI3	A critical point for seed growth rate	g/m^2	1200
WDHI4	A critical point for seed growth rate	g/m^2	3200
WSSD	A coefficient that specifies acceleration or retardation in development in response to water deficit	-	0.4
WSSG	FTSW threshold when dry matter production starts to decline	-	0.3
WSSL	FTSW threshold when leaf area development starts to decline	-	0.4
WSSN	FTSW threshold when nitrogen fixation starts to decline	-	0
WTOPL	Total crop mass when leaf partitioning coefficient turns from FLF1A to FLF1B	g/m^2	127.6

Field experimental management details

Table 3: Experimental details for reference experiments in Eastern Austria.

Experimental location	lat	long	Sowing date	Harvest date	N fertilizer application (kg/ha)			
					1st	2nd	3rd	4th
Tulln	16.0504	48.3106	16.10.2017	04.07.2018	70	70	-	-
Ma.Roggendorf	16.1125	48.6064	06.10.2020	12.07.2021	40	40	30	30
Hetzmannsdorf	16.1046	48.6115	22.10.2021	12.07.2022	40	50	30	40
Potzneusiedl	16.9493	48.0387	13.10.2021	27.06.2022	10	50	50	-

Hydrophysical property scenarios

Table 4: Scenario A, defined as measured for experimental sites (volumetric (%) water content at saturation (SAT), field capacity (DUL), wilting point (LL), air-dry (ADRY); initial volumetric (%) content at sowing, soil organic nitrogen content (NORG), soil mineral nitrogen content (NMIN)).

Experimental location	Layer depth (cm)	SAT	DUL	LL	ADRY	iniWL	BDL	NORG	NMIN
Tulln	10	0.50	0.45	0.15	0.09	0.35	1.36	0.22	14.30
	30	0.56	0.51	0.24	0.14	0.44	1.37	0.22	28.60
	60	0.56	0.51	0.27	0.27	0.45	1.38	0.16	18.60
	90	0.59	0.54	0.29	0.29	0.49	1.42	0.14	8.70
	120	0.59	0.54	0.29	0.29	0.48	1.48	0.11	5.70
Maria Roggendorf	10	0.39	0.34	0.17	0.10	0.16	1.50	0.21	4.85
	30	0.39	0.34	0.18	0.11	0.17	1.50	0.21	10.06
	60	0.40	0.35	0.21	0.21	0.18	1.47	0.18	33.47
	90	0.37	0.32	0.20	0.20	0.28	1.51	0.19	24.41
	120	0.31	0.26	0.20	0.20	0.30	1.59	0.17	7.71
Hetzmannsdorf	10	0.36	0.31	0.15	0.09	0.25	1.52	0.24	9.30
	30	0.32	0.27	0.14	0.08	0.25	1.54	0.23	18.59
	60	0.35	0.30	0.16	0.16	0.28	1.55	0.15	12.20
	90	0.33	0.28	0.16	0.16	0.24	1.57	0.06	15.35
	120	0.28	0.23	0.15	0.15	0.18	1.60	0.04	19.35
Potzneusiedl	10	0.40	0.35	0.21	0.13	0.29	1.58	0.18	26.05
	30	0.40	0.35	0.21	0.13	0.29	1.58	0.17	52.10
	60	0.37	0.32	0.12	0.12	0.27	1.58	0.14	70.25
	90	0.35	0.30	0.12	0.12	0.24	1.62	0.09	24.81
	120	0.32	0.27	0.12	0.12	0.21	1.66	0.04	20.24

Table 5: Scenario B, defined as site-specific values from the Austrian digital soil map intercepting experimental sites (for abbreviations, see caption of Table 4).

Experimental location	Layer depth (cm)	SAT	DUL	LL	ADRY	iniWL	BDL	NORG	NMIN
Tulln	40	0.63	0.58	0.36	0.22	according to initial volumetric water content scenarios	1.50	0.19	according to initial soil mineral nitrogen content scenarios
	100	0.58	0.53	0.37	0.37		1.60	0.19	
Maria Roggendorf	40	0.46	0.41	0.20	0.12		1.50	0.12	
	100	0.40	0.35	0.15	0.15		1.60	0.12	
Hetzmannsdorf	40	0.50	0.45	0.24	0.14		1.50	0.12	
	100	0.46	0.41	0.23	0.23		1.60	0.12	
Potzneusiedl	40	0.47	0.42	0.22	0.13		1.50	0.12	
	100	0.39	0.34	0.15	0.15		1.60	0.12	

Table 6: Scenario C, defined as profile closest in attribute space among 16 "synthetic" soil profiles, created within a 10 km radius from each experimental site (for abbreviations, see caption of Table 4).

Experimental location	Layer depth (cm)	SAT	DUL	LL	ADRY	iniWL	BDL	NORG	NMIN
Tulln	40	0.49	0.44	0.24	0.14	according to initial volumetric water content scenarios	1.50	0.20	according to initial soil mineral nitrogen content scenarios
	100	0.43	0.38	0.18	0.18		1.60	0.20	
Maria Roggendorf	40	0.57	0.52	0.33	0.20		1.50	0.15	
	100	0.46	0.41	0.26	0.26		1.60	0.15	
Hetzmannsdorf	40	0.57	0.52	0.33	0.20		1.50	0.15	
	100	0.46	0.41	0.26	0.26		1.60	0.15	
Potzneusiedl	40	0.50	0.45	0.24	0.14		1.50	0.16	
	100	0.37	0.32	0.13	0.13		1.60	0.16	