



Calibrating and testing APSIM for wheat-faba bean pure cultures and intercrops across Europe

Herman N.C. Berghuijs^{a,b,*}, Martin Weih^a, Wopke van der Werf^c, Alison J. Karley^d,
Eveline Adam^e, Angel M. Villegas-Fernández^f, Lars P. Kiær^g, Adrian C. Newton^d,
Christoph Scherber^{h,j}, Stefano Tavolettiⁱ, Giulia Vico^a

^a Department of Crop Production Ecology, Swedish University of Agricultural Sciences, Ulls väg 16, 75651, Uppsala, Sweden

^b Plant Production Systems Group, Wageningen University & Research, Droevendaalsesteeg 1, 6708 PB, Wageningen, The Netherlands

^c Centre for Crop Systems Analysis, Wageningen University & Research, Droevendaalsesteeg 1, 6708 PB, Wageningen, The Netherlands

^d Ecological Sciences, The James Hutton Institute, Invergowrie, Dundee, DD2 5DA, United Kingdom

^e Saatzucht Gleisdorf, Am Tieberhof 33, 8200, Gleisdorf, Austria

^f Instituto de Agricultura Sostenible, Consejo Superior de Investigaciones Científicas, Avenida Menéndez Pidal s/n, Campus Alameda del Obispo, 14004, Córdoba, Spain

^g Department of Plant and Environmental Sciences, Thorvaldsensvej 40, 1871, Frederiksberg C, Denmark

^h Institute of Landscape Ecology, University of Münster, Heisenbergstraße 2, 48149, Münster, Germany

ⁱ Dipartimento di Scienze Agrarie, Alimentari ed Ambientali, Università Politecnica delle Marche, via Brecce Bianche, 60131, Ancona, Italy

^j Centre for Biodiversity Monitoring, Zoological Research Museum Alexander Koenig, Adenauerallee 160, 53113, Bonn, Germany

ARTICLE INFO

Keywords:

Intercrop
Mixtures
Modeling
APSIM
Europe
Wheat
Faba bean
Triticum aestivum
Vicia faba

ABSTRACT

Cereal-legume intercropping can increase yields, reduce fertilizer input and improve soil quality compared with pure culture. Designing intercropping systems requires the integration of plant species trait selection with choice of crop configuration and management. Crop growth models can facilitate the understanding and prediction of the interactions between plant traits, crop configuration and management. However, currently no existing crop growth model has been calibrated and tested for cereal-legume intercrops throughout Europe. We calibrated the Agricultural Production Systems sIMulator (APSIM) for pure cultures of wheat and faba bean using data from Dutch field trials, and determined the phenological parameters to simulate pure cultures and intercrops from seven field experiments across Europe. APSIM successfully reproduced aboveground dry matters and, for wheat only, grain yields in pure cultures. In intercrops, APSIM systematically overestimated the aboveground dry matter and grain yield of faba bean and underestimated those of wheat. APSIM was reasonably capable of simulating plant heights in pure cultures, but respectively overestimated and underestimated the height of faba bean and wheat in intercrops. In order to simulate wheat-faba bean intercrops better, APSIM should be improved regarding the calculation of biomass partitioning to grains in faba bean and of height growth in both species.

1. Introduction

Global food production needs to satisfy increasing demands while reducing its environmental footprint by lowering anthropogenic inputs and spillovers. Diversification in agriculture can support achieving this complex goal (Tamburini et al., 2020). One way to diversify is intercropping, i.e. cultivating two or more species in the same field for a significant part of their growing period (Willey and Rao, 1980). Intercropping can increase and stabilize crop yield and reduce the environmental impact of arable farming (Malezieux et al., 2009). Yield advantages can only occur if at least one of the species in the intercrop

experiences less intra- and inter-specific competition in an intercrop than it would experience from intraspecific competition in pure cultures (Loreau, 2010). For instance, cereals and legumes grown in intercrops may complement each other regarding nitrogen use. While legumes are capable of fixing nitrogen from the atmosphere, cereals obtain all their nitrogen from the soil (Jensen, 1996). Therefore, the cereal experiences less competition for nitrogen from a neighboring legume than it would from a neighbor of its own species. As a result of such mechanisms for complementarity, intercrops often – but not always – show overyielding (Yu et al., 2015, 2016; Martin-Guay et al., 2018; Li et al., 2020; Xu et al., 2020). Here, overyielding is defined as a situation where the land area

* Corresponding author at: Plant Production Systems Group, Wageningen University & Research, Droevendaalsesteeg 1, 6708 PB, Wageningen, The Netherlands.
E-mail address: herman.berghuijs@wur.nl (H.N.C. Berghuijs).

<https://doi.org/10.1016/j.fcr.2021.108088>

Received 28 June 2020; Received in revised form 12 January 2021; Accepted 28 January 2021

Available online 22 February 2021

0378-4290/© 2021 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

that is required for two species to obtain a certain yield is lower if these species are intercropped than if those two species were grown in a pure culture. Moreover, intercropping may result in an increase in yield stability and improved resource use (Raseduzzaman and Jensen, 2017). Intercropping can reduce the abundance of herbivores and the extent of crop damage by herbivory and enhance the abundance of natural enemies (Letourneau et al., 2011). Intercropping can also result in a stronger weed suppression than a pure culture (Liebman and Dyck, 1993). Further, organic carbon and nitrogen in soil was shown to be higher in soils with a history of intercropping (Cong et al., 2015), due to increased organic matter input and improved nutrient retention. Because of all these potential advantages, intercropping can support the ecological intensification of agriculture (Bommarco et al., 2013).

Global agriculture is dominated by cereals, with maize, wheat and rice representing the crops with the largest growing areas, but they need high inputs of nitrogen via fertilizer application. Incorporation of legumes in cropping systems can reduce the need for anthropogenic nitrogen input. One way to incorporate legumes in cereal systems, and meet the nitrogen demand of cereals, is by using cereal-legume mixtures. In such mixtures, the legume can predominantly utilize nitrogen that is fixed biologically from air through symbiotic bacteria in the root nodules, thus reducing the competition with the cereal, relying on inorganic nitrogen from soil (Jensen, 1996). Designing sustainable and productive cereal-legume intercropping systems is, however, challenging. It is necessary to account for aboveground and belowground interaction between the cereal and legume species, including competition for light, water and nutrients, and the mechanisms behind such competition.

The extent to which one species dominates another is determined by various crop traits, including differences in rooting systems (Corre-Hellou et al., 2007), canopy structure, height and leaf angles (Goudriaan, 1988; Keating and Carberry, 1993; Pronk et al., 2003; Gou et al., 2017b), and nutrient uptake capacity by the cereal and the legume (Corre-Hellou et al., 2006). Additionally, the degree of interspecific competition on both species is also determined by management decisions like row configuration, plant density, fertilization timing and amount, and sowing dates (Yu et al., 2016).

Crop growth models quantitatively integrate the interactions among intercropped species over time and provide predictions of production outcomes (Evers et al., 2019; Gaudio et al., 2019). They can thus help effectively exploring the net outcome of all the aspects affecting the growth of intercrops - pedoclimatic conditions, crop and cultivar traits, and management practices. As such they can support choices of species and management for the ecological intensification of agriculture. Several crop growth models have been developed and successfully applied to simulate the growth, development and yield of a wide variety of crop species and cultivars in pure cultures in various crop production systems and environments (Wallach et al., 2018). Some crop growth models can also simulate intercropping systems. For instance, the crop growth models FASSET (Jacobsen et al., 1998) and STICS (Brisson et al., 2003, 2004) were used to simulate pea-barley intercrop systems in Denmark (Berntsen et al., 2004) and France (Corre-Hellou et al., 2009), respectively. Wheat-maize intercropping was modelled by Gou et al. (2017b) and Tan et al. (2020) added competition for water to this model. The M³ model was used to stimulate the nitrogen-limited growth of wheat-fababean strip intercrops in The Netherlands (Berghuijs et al., 2020). Yet, thorough parameterization and validation of crop growth models for cereal-legume intercrops remain rare, limiting the applicability of models to design the most effective intercropping systems.

The Agricultural Production Systems sIMulator (APSIM) (Keating et al., 2003) has also been used for simulating intercropping systems (Carberry et al., 1996; Knörzer et al., 2011; Chimonyo et al., 2016). APSIM is potentially interesting as a tool for designing sustainable cereal-legume production systems because it has modules for a broad range of crop species, and can simulate their growth under diverse conditions. However, APSIM has been infrequently tested and applied under European conditions, in particular when considering

intercropping systems. To the best of our knowledge, APSIM has been applied to cereal-legume intercropping systems in Europe only once (Knörzer et al., 2011), i.e. wheat-field pea and maize-field pea mixtures grown in Southern Germany. Furthermore, there are only few studies available in which APSIM WHEAT (crop module in APSIM for modelling wheat) has been calibrated and validated for pure cultures of wheat in Europe (Asseng et al., 2000; Knörzer et al., 2011) and none relative to APSIM FABABEAN (crop module in APSIM for modelling faba bean) outside Australia. Given the potential broad applicability of APSIM, it is of interest to further assess its power in simulating cereal-legume intercrops under a variety of pedoclimatic conditions.

To advance our capabilities to evaluate intercropping as a tool to support ecological intensification of agriculture, here we assess to what extent APSIM can be used to simulate intercrops of cereals and legumes in Europe. Specifically, we consider spring wheat-faba bean intercrops grown under European growing conditions. This study aims to 1) calibrate and validate APSIM for pure cultures of spring wheat and faba bean under temperate European conditions, exploiting detailed data from experiments in the Netherlands (Kropff, 1989; Boons-Prins et al., 1993; Gou et al., 2016); and 2) evaluate the performance of APSIM in predicting the crop yield in various European sites for both pure cultures and intercrops of spring wheat and faba bean, based on several field experiments recently carried out throughout Europe.

2. Material and methods

2.1. Field experimental data

We exploited two groups of experimental data. The first group consists of previously published crop measurements on pure cultures of spring wheat (Gou et al., 2017b) and faba bean (Kropff, 1989) from experiments conducted in Wageningen, the Netherlands. These data are comparatively detailed, including results from periodic harvests and measurements of leaf area index (LAI) and plant height at different times within the growing season. Thanks to this higher level of detail, these data (Table 1) lend themselves to calibrating and validating APSIM WHEAT and APSIM FABABEAN for pure cultures of wheat and faba bean. The second group of data consists of final aboveground dry matter and yield measurements of spring wheat and faba bean in pure culture and intercrops. The data were collected in 2017 and 2018 at seven locations along a latitudinal gradient in Europe, as part of the EU Horizon 2020-funded project DIVERSify (Tables 2 and 3). The experiments in Dundee (UK) in 2018 and in Gleisdorf in both 2017 and 2018 were unfertilized. The experiment in Ancona in 2018 was fertilized, but it comprised only a single, low fertilization regime. In all the other cases, two different management treatments were considered, hereafter referred to as "conventional management" and "low input management". Input levels differed among the locations, but at each location nitrogen was applied at a higher rate at the conventional management level than at the low input level. In all intercropping experiments, the crop species were sown and harvested at the same date and both crop species were mixed within the row (i.e. not alternate row intercropping). We did not consider the experiment conducted in 2017 in Córdoba, because the crops were heavily affected by aphids, most likely as the result of an unusually late sowing date. Aphid infestation cannot be reproduced by APSIM. We also excluded the data from Dundee collected in 2017, because measurements of aboveground dry matter and grain yield separated per species were not available.

2.2. Pedoclimatic inputs

2.2.1. Soil water module

We used the SoilWat module to simulate the soil water dynamics for each combination of experiment and treatment. This module requires as input for each layer the thickness (Δz), the bulk density (ρ) and the volumetric soil water contents at air dry (θ_{ad}), field capacity (θ_{du}),

Table 1
Overview of experiments that were used to calibrate and validate APSIM.

Reference	Location	Crop Species	Cultivar	Years calibration	Years validation	Treatment
Kropff (1989) and Boons-Prins et al. (1993)	Wageningen (the Netherlands)	Faba bean	Monica	1985	1986, 1988	No exposure to SO ₂
Gou et al. (2017a)	Wageningen (the Netherlands)	Wheat	Tybalt	2013	2014	Pure culture

saturation (θ_s) and the 15-bar lower limit soil moisture content (θ_{LL15}) (Probert et al., 1998).

Each of these volumetric water contents can range from 0 to soil porosity. In order to determine these parameters, we first obtained ρ and mass fractions of clay (f_{clay}), sand (f_{sand}), silt (f_{silt}), and organic carbon (f_{oc}) from the freely-available SoilGrids database (Hengl et al., 2014, 2017). We calculated the fraction of organic matter f_{om} as (Pribyl, 2010):

$$f_{om} = 1.72 \cdot f_{oc} \quad (1)$$

We used the pedotransfer function described by Wösten et al. (1999) and Wösten and Nemes (2004) to calculate the Van Genuchten-Mualem parameters α , n , and θ_s (Mualem, 1976). Those parameters describe a soil water retention curve (Van Genuchten, 1980) as:

$$\theta = \theta_r + \frac{\theta_s - \theta_r}{(1 + (\alpha \cdot |\psi|)^n)^{1-\frac{1}{n}}} \quad (2)$$

where θ_r is the residual volumetric soil moisture content and ψ is the water potential (hPa). We assumed that $\theta_r = 0.025$.

We further adopted the assumptions from Leffelaar (2014) that the pF value (10-base logarithm of $|\psi|$) of a soil layer equals 5.0 at air dry, 4.2 at the wilting point and 2.0 at field capacity; and that the soil moisture content at wilting point is the same as θ_{LL15} . With these assumptions, we used Eq. 2 to calculate θ_{ad} , θ_{LL15} , and θ_{dul} .

2.2.2. Meteorological input data

APSIM requires daily global radiation, minimum and maximum temperature ($T_{min}(t)$ and $T_{max}(t)$), and precipitation as input data. For each field trial and growing season, these data were obtained from the NASA Langley Research Center Atmospheric Science Data Center Surface meteorological and Solar Energy (SSE) web portal supported by the NASA LaRC POWER Project (<https://power.larc.nasa.gov/data-access-viewer/>).

2.2.3. Management data

APSIM requires sowing dates, nitrogen application amounts, and harvest dates as inputs. Tables 2 and 3 shows the input values of these variables.

2.3. APSIM calibration and validation procedure

2.3.1. Calibration procedure for spring wheat

Table 1 shows an overview of the data that were used to calibrate and validate the crop module APSIM WHEAT (Meinke et al., 1997, 1998a; Meinke et al., 1998b; Wang et al., 2003; Zheng et al., 2015). APSIM's cultivar Hartog was used as a starting point for the calibration procedure. Sensitivities for photoperiodicity ($photop_sens$) and vernalization ($vern_sens$) were set to 0. The parameter $shoot_lag$ was calculated from the sowing depth, the default value of $shoot_rate$ ($^{\circ}C$ d), the observed emergence dates and daily temperature data. Phenological parameters (i.e. thermal times from one stage to the next) were determined based on phenological observations and the corresponding temperatures in the experimental data as:

$$tt_stage = \Delta t \cdot \sum_{t_i=t_{start}}^{t_{end}} \max\left(\frac{T_{min}(t_i) + T_{max}(t_i)}{2} - T_b, 0\right) \quad (3)$$

where tt_stage is the thermal time ($^{\circ}C$ d) from the start of a certain phenological stage (at time t_{start}) until the start of the next stage (at time

t_{end}). Δt is the time step (d). T_b is the base temperature ($^{\circ}C$ d), i.e., the temperature below which the thermal time does not increase with temperature. In this way, we determined $tt_end_juvenile$, $tt_floral_initiation$, $tt_flowering$, $tt_start_grainfill$, and $tt_end_grain_fill$, i.e., the thermal times from the end of the juvenile stage until floral initiation, floral initiation to flowering, flowering to the start of grain filling and from the start of grain filling until the end, respectively. After determining the parameters values mentioned above, we ran APSIM to determine the stem weight at the day of flowering for wheat for the experiment from Wageningen conducted in 2013. From the simulated stem weight and the measured yield components, we calculated the parameters max_grain_size (maximum grain size; $g\ seed^{-1}$) and $grain_per_gram_stem$ (number of kernels per gram of stem at flowering). Finally, we manually optimized parameter y_frac_leaf (fraction of newly produced biomass allocated to leaves).

2.3.2. Calibration procedure for faba bean

Table 1 shows an overview of the data that were used to calibrate the crop module APSIM FABABEAN (Robertson et al., 2002; Turpin et al., 2002, 2003). We used parameters for the cultivar Fjord as starting point. We switched off the sensitivity to vernalization ($vern_sens = 0$). We also switched off the sensitivity to photoperiodicity, by assuming that the phenological parameter values at short daylengths equal their values at long day lengths. We calculated the values of $y_tt_end_of_juvenile$, $y_tt_flowering$ and $y_tt_start_grain$ from phenological observations and temperature data using Eq. 3. We thereby assumed that the ratio of the $y_tt_emergence$ and $y_tt_floral_initiation$ to the thermal time from emergence to flowering, and the ratio of $tt_flowering$ and $tt_start_grain_fill$ to the thermal time from flowering to maturity, were the same in the cultivars that were used in this study and in the cultivar Fjord that was used as the starting point of calibration. We calculated the parameter $shoot_lag$ for faba bean in the same way as for the cereals, assuming a sowing depth of 40 mm. Unfortunately, although the emergence dates for the Wageningen data from Kropff (1989) were known, the sowing dates were not. Therefore, we ran the simulations for the Kropff (1989) experiments using the emergence date as the start date of simulation in such a way that the plant emerged immediately with sowing depth equal to 0.

We estimated the radiation use efficiency y_rue as the slope of the relationship between the sum of the daily intercepted radiation and the observed biomass, calculated as explained by Gou et al. (2017a). Finally, we adjusted the values of the maximum harvest index that can be reached ($y_hi_max_pot$) and the daily growth rate of the harvest index (y_hi_incr) to higher values than those representing the cultivar Fjord.

2.3.3. Validation procedure

We validated APSIM WHEAT and APSIM FABABEAN by comparing measured and simulated values of LAI, biomass and, in the case of APSIM WHEAT, the height of the plants. Table 1 shows the data that were used for validation procedure.

2.4. Local model recalibration and model evaluation on DIVERSify data

We used the newly calibrated and validated APSIM crop modules for each species based on the Wageningen data to predict the grain yields, total aboveground dry matters and plant heights at the end of the growing season that were measured in pure cultures and intercrops in the DIVERSify experiments (Tables 2 and 3). Since the cultivars differed from one location to the other and sometimes also between different

Table 2

Overview of the DIVERSify experiments conducted in 2017. The treatment levels consisted of a combination of management (C: conventional management; L: low input management) and cropping system (PC: pure culture; IC: intercrop). Fertilization rates and sowing densities are reported for wheat (W) and faba bean (F).

Location	Treatment	Fertilization (kg N ha ⁻¹)	Sowing date (yy-mm-dd)	Sowing density (seeds ha ⁻¹)	Harvest date (yy-mm-dd)
Ancona (Italy) 43° 32'42" N 13° 21'34" E	C-PC	W: 160, F: 20 ¹	2017-02-16	W: 450, F: 60	2017-07-20
	L-PC	W: 80, F: 0 ¹	2017-02-16	W: 450, F: 60	2017-07-20
	C-IC	80	2017-02-16	W: 225, F: 30	2017-07-20
Córdoba (Spain) 37° 47'15" N 5°3' 13" W	L-IC	25	2017-02-16	W: 225, F: 30	2017-07-20
	C-PC	32	2017-03-01	W: 440, F: 50	2017-06-12
	L-PC	16	2017-03-01	W: 440, F: 50	2017-06-12
Dundee (UK) 56° 48'17" N 3° 11' 17" W	C-IC	32	2017-03-01	W: 220, F: 25	2017-06-12
	L-IC	16	2017-03-01	W: 220, F: 25	2017-06-12
	C-PC	40	2017-03-29	W: 440, F: 50	2017-08-30
Gleisdorf (Austria) 47° 6'49" N 15° 42'0" E	L-PC	0	2017-03-29	W: 440, F: 50	2017-08-30
	C-IC	40	2017-03-29	W: 220, F: 25	2017-08-30
	L-IC	0	2017-03-29	W: 220, F: 25	2017-08-30
Gleisdorf (Austria) 47° 6'49" N 15° 42'0" E	L-PC	0	2017-03-16 ² 2017-04-11	W: 440, F: 40	2017-08-09
	L-IC	0	2017-03-16 ² 2017-04-11	W: 220, F: 20	2017-08-09
Münster (Germany) 51° 58'32" N 7° 33'59" E	C-PC	75	2017-04-11	W: 320, F: 80	2017-09-11
	L-PC	0	2017-04-11	W: 320, F: 80	2017-09-11
	C-IC	75	2017-04-11	W: 160, F: 40	2017-09-11
Taastrup (Denmark) 55° 40' 10" N 12° 18'17" E	L-IC	0	2017-04-11	W: 160, F: 40	2017-09-11
	CPC	60	2017-04-07	W: 528, F: 48	2017-09-06
	L-PC	20	2017-04-07	W: 528, F: 48	2017-09-06
Uppsala (Sweden) 59° 50'6" N 17° 42'0" E	C-IC	60	2017-04-07	W: 264, F: 24	2017-09-06
	L-IC	20	2017-04-07	W: 264, F: 24	2017-09-06
	C-PC	140	2017-05-05	W: 490, F: 60	2017-09-05
Uppsala (Sweden) 59° 50'6" N 17° 42'0" E	L-PC	0	2017-05-05	W: 490, F: 60	2017-09-05
	C-IC	140	2017-05-05	W: 245, F: 30	2017-09-05
	L-IC	0	2017-05-05	W: 245, F: 30	2017-09-05

¹The rates of nitrogen application differed between wheat (W) and faba bean (F) in the conventional management treatment.

² The experiments in Gleisdorf included a sowing date treatment.

years at the same location, we recalculated the phenological parameters for each combination of crop species, cultivar and location from phenological observations in these field trials. All other parameters remained the same as calibrated based on the Wageningen data. The field trial of Córdoba lacked phenological measurements. Therefore, we determined the phenological parameters from estimates that were done in Spain on wheat monocultures in a previous study (Boons-Prins et al., 1993). We assumed that the phenological parameters from Córdoba were equal to those from Italy.

Following Salo et al. (2016), we consider three indices to quantify the performance of APSIM: i) the mean bias error (MBE); ii) the root-mean squared error (RMSE); and iii) the index of agreement (IA) (Willmott and Wicks, 1980; Willmott, 1981). The MBE represents the average difference between the model predictions and the measured values and is usually called bias in the statistical literature. It is an average over all the locations, cultivars and treatments and is calculated as:

$$MBE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i) \quad (4)$$

where y_i is either the observed average yield, average aboveground dry matter, or maximum plant height for a certain combination i of location and management treatment; \hat{y}_i is the corresponding simulated yield, aboveground dry matter, or plant height for combination i ; N is the total number of combinations of location and management treatment.

The RMSE quantifies the absolute deviation between observed and simulated values:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (5)$$

RMSE is a standard deviation, and is expressed in the same units as the variable of interest.

The Index of Agreement (IA) is used as an indicator of model efficiency:

$$IA = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (|\hat{y}_i - \bar{y}| + |y_i - \bar{y}|)^2} \quad (6)$$

where \bar{y} is the average yield or average aboveground dry matter of all observations. IA is an index that varies from 0 to 1, with higher values representing better predictions. Its interpretation is similar to that of the coefficient of determination R^2 (proportion of variance explained) in ordinary regression.

3. Results

3.1. Calibration and validation of APSIM crop modules using within-season data

3.1.1. APSIM WHEAT

The calibration procedure started with considering the effects of vernalization (exposure to low temperature) and daylength on the temperature sums required to reach different phenological stages (Zheng et al., 2015). We set the sensitivity to vernalization (vern_sens) to zero because this process is largely irrelevant for spring wheat. Moreover, even if it would have some effect, it cannot be quantified separately from the parameter tt_end_juvenile easily, as this would require data from various experiments during which the cultivar of interest was exposed to different low temperature regimes. We varied the parameter photop_sens between its minimum (0) and maximum value (5.0) and found that this parameter did not have any appreciable effect on the

Table 3

Overview of the DIVERSify experiments conducted in 2018. The treatment levels consisted of a combination of management (C: conventional management; L: low management) and cropping system (PC: pure culture; IC: intercrop). Fertilization rates and sowing densities are reported for wheat (W) and faba bean (F).

Location	Treatment	Fertilization (kg N ha ⁻¹)	Sowing date (yy-mm-dd)	Sowing density (seeds ha ⁻¹)	Harvest date (yy-mm-dd)
Ancona (Italy) 43° 32'42" N 13° 21'34" E	L-PC	48	2018-01-31	W: 450, F: 60	2018-07-16
	L-IC	48	2018-01-31	W: 225, F: 30	2018-07-16
Córdoba (Spain) 37° 47'15" N 5°3' 13" W	C-PC	84	2017-11-27	W: 440, F: 50	2018-06-08
	L-PC	42	2017-11-27	W: 440, F: 50	2018-06-08
	C-IC	84	2017-11-27	W: 220, F: 25	2018-06-08
	L-IC	42	2017-11-27	W: 220, F: 25	2018-06-08
Dundee (UK) 56° 48'17" N 3° 11'17" W	L-PC	0	2018-04-02	W: 440, F: 50	2018-09-03
	L-IC	0	2018-04-02	W: 220, F: 25	2018-09-03
Gleisdorf (Austria) 47° 6'49" N 15° 42'0" E	L-PC	0	2018-04-12	W: 440, F: 40	2018-08-08
	L-IC	0	2018-04-12	W: 220, F: 20	2018-08-08
Münster (Germany) 51° 58'32" N 7° 33'59" E	C-PC	32.6	2018-04-23	W:320, F:80	2018-09-11
	L-PC	0	2018-04-23	W: 320, F:80	2018-09-11
	C-IC	32.6	2018-04-23	W:160, F:40	2018-09-11
	L-IC	0	2018-04-23	W:160, F:40	2018-09-11
Taastrup (Denmark) 55° 40' 10" N 12° 18'17" E	C-PC	65	2018-04-19	W: 528, F: 48	2018-09-15
	L-PC	23	2018-04-19	W: 264, F: 24	2018-09-15
	C-IC	65	2018-04-19	W: 264, F: 24	2018-09-15
	L-IC	23	2018-04-19	W: 490, F: 60	2018-09-15
Uppsala (Sweden) 59° 50'6" N 17° 42'0" E	C-PC	140	2018-04-30	W: 490, F: 60	2018-09-06
	L-PC	0	2018-04-30	W: 490, F: 60	2018-09-06
	C-IC	140	2018-04-30	W: 245, F: 30	2018-09-06
	L-IC	0	2018-04-30	W: 245, F: 30	2018-09-06

simulated phenology. Therefore, this parameter was also set to zero.

The next step in the calibration procedure of the spring wheat was to determine two of the main parameters known to affect yield: the number of grains per unit of stem weight (*grains_per_stem*), driving how many kernels are formed; and the maximum grain weight (*max_grain_size*), defining the maximum weight a grain can reach. This calibration was necessary because the default values for these parameters were not high enough to allow APSIM WHEAT to simulate the high yields that were reported in the experiments of Gou et al. (2016).

Figs. 1–3 summarize the results of the calibration and validation, and Table 4 shows the original parameter values of the APSIM WHEAT crop module and their re-estimated values after the first calibration step using the Wageningen data (Gou et al. (2016)). There was a good agreement between the simulated and measured biomass partitioning in the calibration experiment from 2013 (Fig. 1A), but LAI was somewhat overestimated after flowering (Fig. 2). There was also good agreement between the measured and simulated biomass of leaf and ear as well as LAI in the validation experiment from 2014 (Fig. 1B and 2B, but the stem biomass was underestimated after flowering. Finally, although there was good agreement between the simulated and measured plant height in the second half of the season in the calibration, the model systematically underestimated plant height in the validation (Fig. 3).

3.1.2. APSIM FABABEAN

APSIM FABABEAN uses thermal times to determine the timing of key phenological stages (*y_tt_end_of_juvenile*, *y_tt_floral_initiation*, *y_tt_flowering*, and *y_tt_start_grain_fill*). The thermal times are defined as a function of day length (*x_pp_end_of_juvenile*, *x_pp_floral_initiation*, *x_pp_flowering*, and *x_pp_grain_fill*). Consequently, APSIM assumes a rapid decline of the value of these parameters with increasing daylength. Although these daylength-dependent functions have been successfully applied to simulate the growth of faba bean in Australia (Turpin et al., 2003), the use of these functions resulted in a growing season duration that was too short under European conditions, which include longer daylengths than at lower latitudes, like in Australia. Consequently, the yields and the aboveground dry weights of faba bean were underestimated substantially when the default APSIM parameter values for faba bean were used. Therefore, in the calibration procedure, we assumed that the listed parameters were independent of daylength. We calculated the thermal times from the emergence dates, flowering dates, and maturity dates reported by Boons-Prins et al. (1993). Finally, we

calculated the radiation use efficiency (*y_rue*; g MJ⁻¹) from the assumed extinction coefficients in APSIM, the measured aboveground biomass and LAIs, and radiation using the method described by Gou et al. (2017a).

Table 5 summarizes the original parameter values of the APSIM FABABEAN crop module and their re-estimated values after calibration. Figs. 4 and 5 compare observed and simulated aboveground dry matter, partitioned over different organs, LAI and crop height for faba bean. In general there was a good agreement between the simulated and measured biomass partitioning, although the stem biomass was slightly underestimated in 1985 and 1988 (Fig. 4). There was a good agreement between the measured and simulated LAI for both the dataset used for calibration (from 1985) and that used for validation (1986 and 1988). For both 1985 and 1988, APSIM FABABEAN overestimated plant height early in the growing season and underestimated it at the end of the growing season (Fig. 5).

3.2. Evaluation of APSIM using DIVERSIFY data in pure crops and intercropping

3.2.1. Biomass and yield in pure cultures and intercrops

Tables 6 and 7 show the adjusted values of the phenological parameters for each location in the DIVERSify experiments. APSIM WHEAT performed reasonably well in reproducing most of the measurements on pure stands of wheat in the seven locations across Europe of the DIVERSify experiments in 2017 and 2018, although with a few outliers (Fig. 6). The simulated aboveground dry matter in Dundee (UK) in 2018 was about 6.5 times larger than the observations for both wheat and faba bean. The observed biomass production of the crop and the yield were particularly low at this location in 2018 due to severe drought stress. Nevertheless, APSIM WHEAT was not capable of fully reproducing these low yields, despite including the soil water balance and at least some aspects of the effects of low water availability on crop development. A possible explanation is that the interpolated soil data that we obtained from SoilGrids for this location were not representative of this site and overestimated the soil water holding capacity.

Similar to APSIM WHEAT, APSIM FABABEAN overestimated the aboveground dry matter in Dundee (UK). However, APSIM FABABEAN was also not able to reproduce the high total aboveground biomass as measured in Taastrup (Denmark) in 2017. While the performance of APSIM FABABEAN in simulating the aboveground biomass was

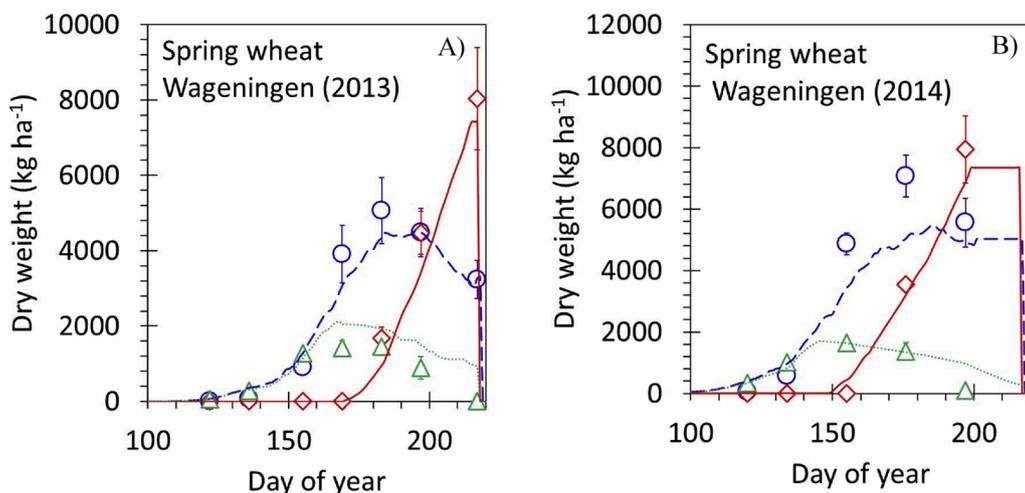


Fig. 1. Simulated (lines) and measured (symbols) biomass in spring wheat for the experiments from Gou et al. (2016), used to calibrate (A) and validate (B) APSIM WHEAT. Measured ear, stem and leaf biomass are represented as red diamonds, blue circles, and green triangles respectively. Simulated ear, stem and leaf biomass are represented as red solid lines, blue dashed lines, and green dotted lines, respectively. The length of an error bar is one standard deviation. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

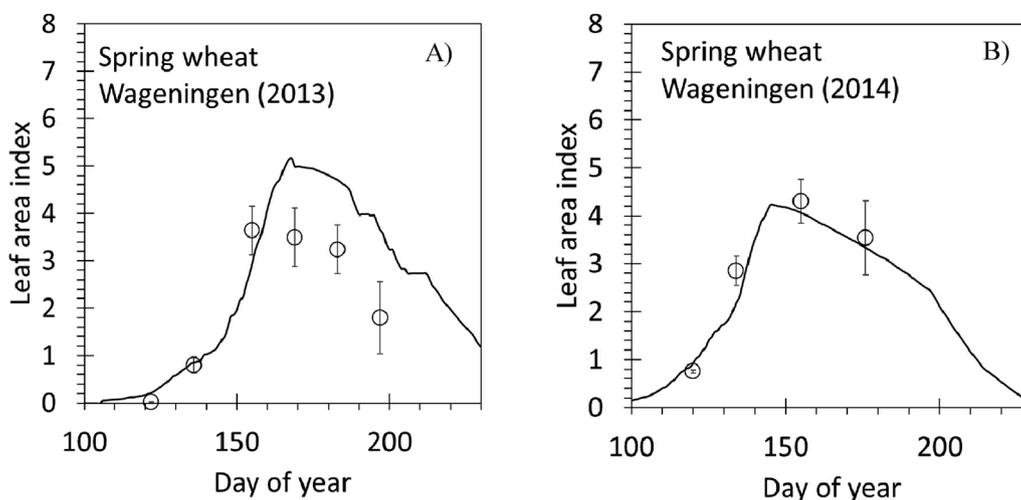


Fig. 2. Simulated (lines) and measured (symbols) LAI in spring wheat for the experiments from Gou et al. (2016), used to calibrate (A) and validate (B) APSIM WHEAT. Circles represent measurements and lines simulated values. The length of an error bar is one standard deviation.

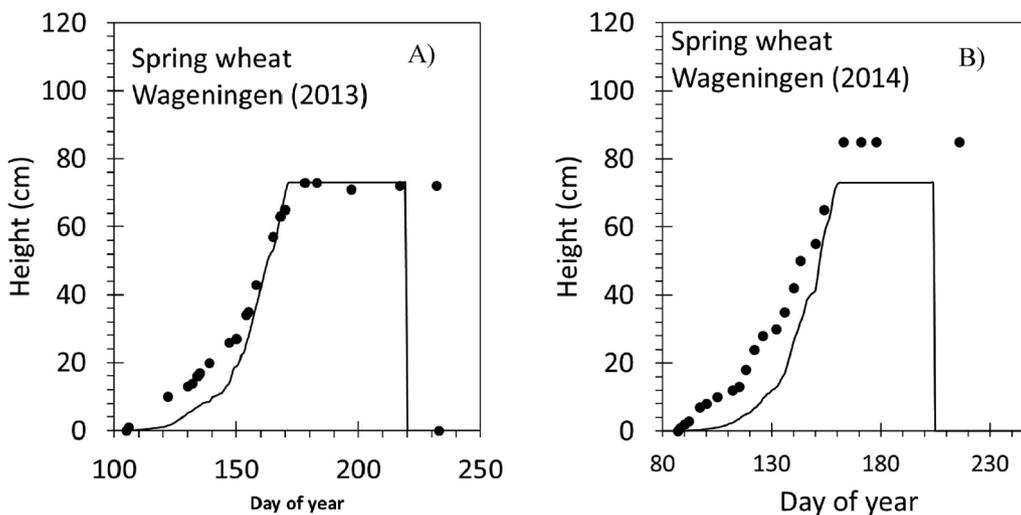


Fig. 3. Simulated (lines) and measured (symbols) heights in spring wheat for the experiments from Gou et al. (2016), used to calibrate (A) and validate (B) APSIM WHEAT. Dots represent measurements and lines simulated values.

Table 4
Overview of default APSIM WHEAT parameters values and their values after recalibration.

Parameter	Definition	Default value	Re-estimated value	Unit
vern_sens	Sensitivity of thermal times to vernalization	1.5	0	–
photop_sens	Sensitivity of thermal times to photoperiod	3.0	0	–
shoot_lag	Thermal time between sowing and start of linear coleoptile growth	56	67	°C d
tt_end_of_juvenile	Thermal time from the the end of juvenile to floral initiation	400	730	°C d
tt_floral_initiation	Thermal time from floral initiation until to flowering	555	124	°C d
tt_flowering	Thermal time from flowering to the start of grain filling	120	118	°C d
tt_start_grain_fill	Thermal time from the start to the end of grain filling	545	598	°C d
max_grain_size	Maximum grain size	0.041	0.056	g
grain_per_gram_stem	Number of grains per stem weight at flowering	35.0	36.0	g ⁻¹
x_frac_leaf	Vector with phenological stages for which y_frac leaf contains a corresponding value.	1,2,3,4,5,5,4,6,6,9,7,8,9,10,11	1,2,3,4,5,5,4,6,6,9,7,8,9,10,11	–
y_frac_leaf	Vector with phenology dependent fractions of newly produced biomass partitioned to leaves	0,0,0,6,0,6,0,6,0,42,0,0,0,0,0,0	0,0,0,45,0,45,0,45,0,1,0,0,0,0,0,0	–
x_stem_wt	Vector with phenological stages for which y_stem_wt contains a corresponding value.	0,6	0,1.56	g
y_stem_height	Vector with stem weight dependent heights	0,1500	0,730	mm

Table 5
Overview of default APSIM FABABEAN parameters values and their values after recalibration.

Parameter	Defintion	Default value	Re-estimated value	Unit
tt_emerge	Thermal time from emergence until the end of the juvenile stage	438	357	°C d
y_tt_end_of_juvenile	Vector with day length dependent thermal times from the end of the juvenile stage until floral initiation	381,1	316, 316	°C d
y_tt_floral_initiation	Vector with day length dependent thermal times from floral initiation to flowering	16,16	19,19	°C d
y_tt_flowering	Vector with day length dependent thermal times from flowering to the start of grain filling.	500, 100	414, 414	°C d
y_tt_start_grain_fill	Vector with day length dependent thermal times from flowering to the start of grain filling.	1300,100	1066,1066	°C d
y_rue	Vector with phenology dependent radiation use efficiencies.	0, 0,1.15, 1.15, 1.15, 1.15, 1.15, 1.15	0, 0, 1.32, 1.32, 1.32, 1.32, 1.32, 1.32	g MJ ⁻¹

mediocre, APSIM FABABEAN performed poorly in simulating faba bean yield. This was, among others, reflected in a low value of the IA (0.35).

Regarding intercrops, the model substantially underestimated the aboveground dry matter of wheat (MBE = –3878 kg ha⁻¹) and overestimated the aboveground dry matter of faba bean (MBE =7295 kg ha⁻¹), although some underestimated values occurred for some locations (Fig. 7).

3.2.2. Plant height in pure cultures and intercrops

APSIM performed poorly in explaining the variation of the observed plant heights both in pure culture and intercrop (Fig. 8). For example, the index of agreement was low for the plant heights in wheat pure cultures (IA = 0.43). Nevertheless, the model error was small (MBE = 19 cm, RMSE =20 cm) relative to the simulated heights, which were all between 79 cm and 85 cm. The IA was considerably higher in faba bean pure cultures (0.67) but, also there, the model error was relatively small (MBE =3 cm, RMSE =25 cm). Similarly to the results of the aboveground dry matter (Fig. 6C), APSIM FABABEAN strongly overestimated the plant height in pure cultures in Taastrup under both low input and conventional input management. In contrast to the pure cultures, APSIM WHEAT strongly underestimated the plant height of wheat in intercrops (MBE = –54 cm, RMSE =55 cm), while APSIM FABABEAN strongly overestimated the plant height of faba bean (MBE =45 cm, RMSE =49 cm).

4. Discussion

This study is the first application of APSIM to simulate cereal-legume intercrops throughout Europe. We calibrated and validated APSIM WHEAT and APSIM FABABEAN on wheat and faba bean pure cultures based on detailed datasets from Dutch field trials. Next, we adjusted the phenological parameters for different locations in Europe and simulated pure cultures and intercrops of wheat and faba bean throughout Europe. We evaluated the performance of APSIM to simulate these systems.

4.1. Performance of APSIM WHEAT for wheat monocultures

APSIM WHEAT could be successfully applied to simulate the yield and the aboveground dry matter of spring wheat and faba bean in the Netherlands under well fertilized conditions and pure cultures. This required adjustments of the phenological and yield component parameters with respect to those of the APSIM WHEAT default cultivar. Given that APSIM had mostly been calibrated and validated in subtropical and tropical areas for wheat, our results demonstrate the robustness of the APSIM WHEAT module in other climatic regions. The newly calibrated APSIM WHEAT was also capable of reproducing the biomass and yields of wheat in pure cultures of the DIVERSify field-trials across Europe reasonably well, after the phenological parameters were adjusted according to the location.

4.2. Performance of APSIM FABABEAN in pure cultures

Similarly to APSIM WHEAT, APSIM FABABEAN mainly required adjustments of the phenological parameters to simulate the biomass production of faba bean in the Netherlands. Making phenology

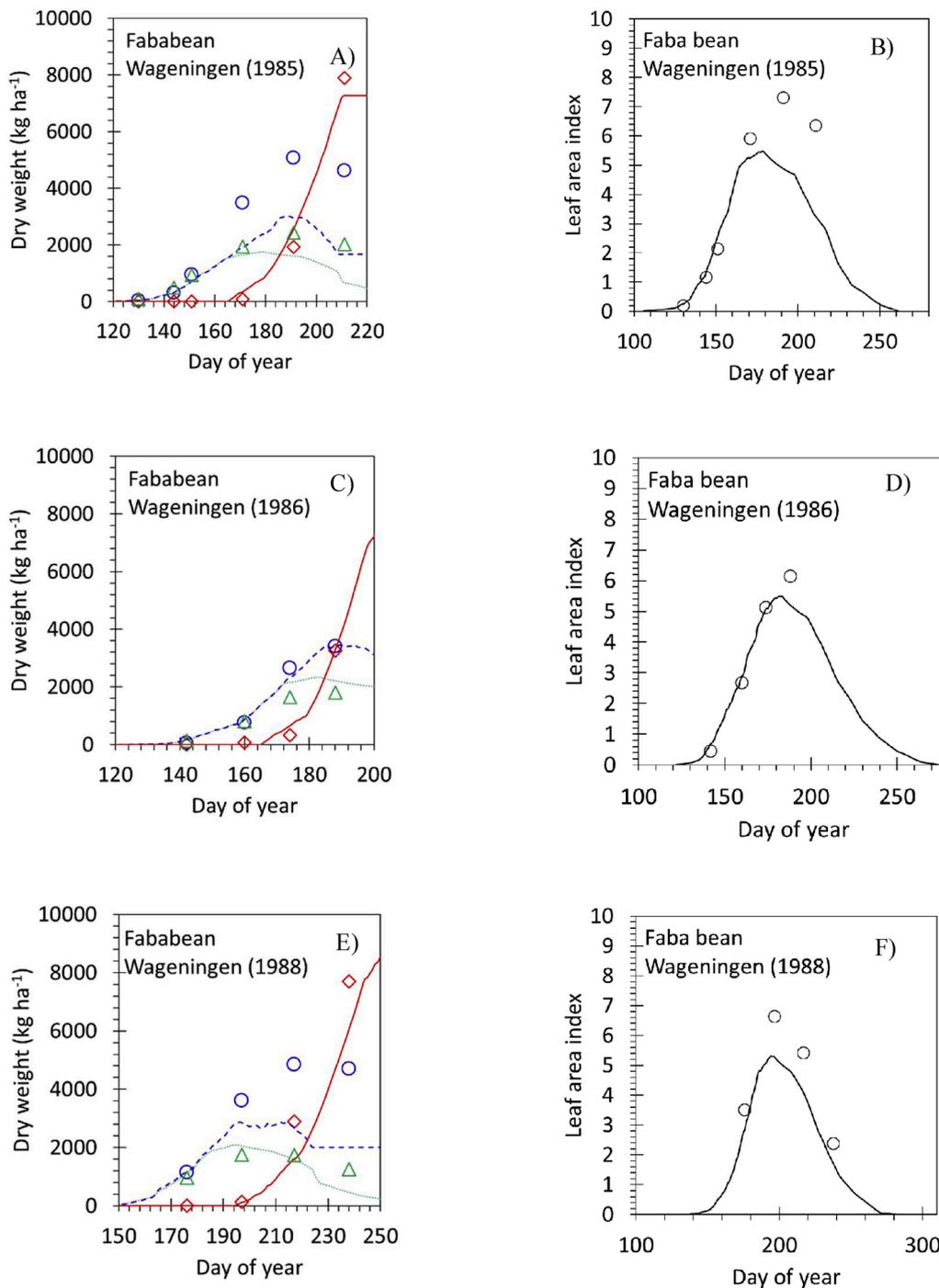


Fig. 4. Simulated and measured biomass partitioning and leaf area index in faba bean for the experiments from Kropff (1989) that were used to calibrate (A-B) and validate (C-F) APSIM FABABEAN. Measured fruit (combined pod and grain), stem and leaf biomass are represented as red diamonds, blue circles, and green triangles respectively. Simulated fruit, stem and leaf biomass are represented as red solid lines, blue dashed lines, and green dotted lines, respectively (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

independent of day length was an essential part of the calibration. Despite the calibration, the resulting APSIM FABABEAN performance was mediocre at best in reproducing the aboveground dry matter from pure cultures of the DIVERSify field-trials across Europe and poor for grain yields (Fig. 6). This indicates that the model for biomass partitioning to the grains of APSIM FABABEAN needs to be further improved. We note that APSIM WHEAT calculates the number of grains from the stem weight per plant at flowering and then determines the increase in the weight per grain from the grain filling rate. Grain filling stops once either the thermal requirement of grain filling has been fulfilled or a certain weight per grain has been reached (Zheng et al., 2015). In contrast, APSIM FABABEAN assumes that harvest index is 0 until flowering and then starts to increase linearly with time until either the

thermal requirement of grain filling or a preset maximum harvest index has been reached (Robertson et al., 2002). Hence, unlike APSIM WHEAT, this approach does not consider the grain number and requires parameters (like daily increase of the harvest index and the maximum harvest index) that are difficult to obtain and can depend on both cultivar and the environment where the plant is grown. In the light of these results, the APSIM FABABEAN parameters that determine the harvest index and the biomass partitioning need to be determined for a broader range of cultivars and growing conditions. Additionally, it may be necessary to adjust the way in which APSIM FABABEAN calculates the grain dry matter production.

APSIM FABABEAN performed considerably better in predicting the plant heights in pure cultures than APSIM WHEAT at most locations,

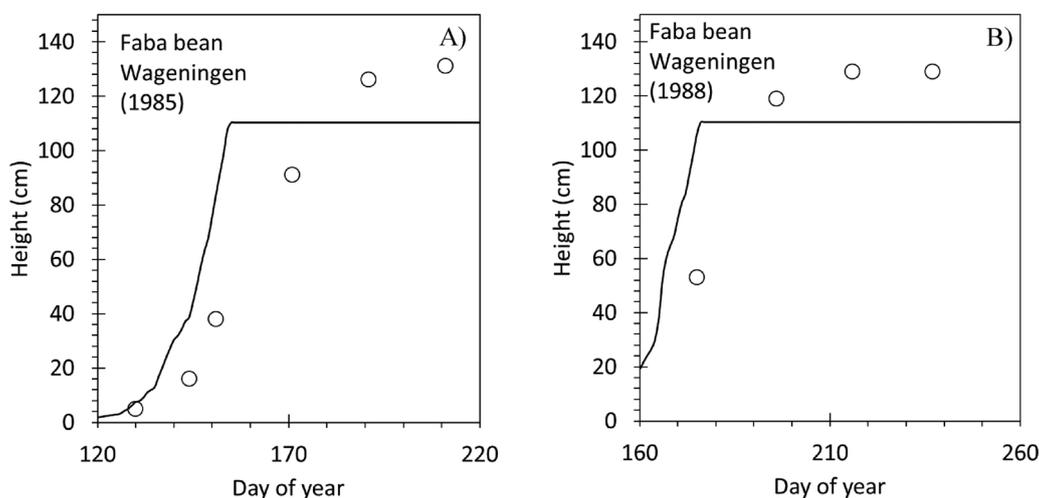


Fig. 5. Simulated and measured heights in faba bean for the experiments from Kropff (1989) that were used to calibrate (A) and validate (B) APSIM FABABEAN. Circles represent measurements and lines represent simulated values.

Table 6

Overview of wheat phenological parameters per location; all parameters are expressed in °C d.

Phenological parameters	Ancona (Italy)	Córdoba (Spain)	Dundee (UK)	Gleisdorf (Austria)	Münster (Germany)	Taastrup (Denmark)	Uppsala (Sweden)	Wageningen (Netherlands)
tt_end_of_juvenile	1001	1240	723	901	558	515	628	730
tt_floral_initiation	170	210	122	153	85	87	107	124
tt_flowering	534	661	386	481	298	275	335	118
tt_start_grain_fill	2704	3351	1955	2435	1509	1393	1699	598

Table 7

Overview of faba bean phenological parameters per location; all parameters are expressed in °C d.

Phenological parameters	Ancona (Italy)	Córdoba (Spain)	Dundee (UK)	Gleisdorf (Austria)	Münster (Germany)	Taastrup (Denmark)	Uppsala (Sweden)	Wageningen (Netherlands)
y_tt_emergence	393	393	248	237	215	201	237	357
y_tt_end_of_juvenile	342	342	216	106	187	175	206	316
tt_flowering	208	208	131	125	114	106	125	181
tt_start_grain_fill	535	535	338	223	293	273	322	467

except for Taastrup in 2017, where it strongly underestimated the plant height at both high and low management input treatments (Fig. 8).

4.3. Performance of APSIM WHEAT and APSIM FABABEAN in intercrops

APSIM had limited capabilities to reproduce the yields of wheat-faba bean intercrops. This is not surprising as APSIM FABABEAN poorly reproduced the faba bean yields in pure cultures. However, APSIM also performed poorly in reproducing the total aboveground dry matter of both species when intercropped. It underestimated the total aboveground dry matter of wheat for each DIVERSify field trial and overestimated the total aboveground dry matter of faba bean. A possible explanation of this pattern is the way that APSIM simulates crop height. APSIM assumes that the crop height is proportional to the stem dry weight per plant until grain filling starts or until the plant has reached its maximum height (i.e. the last element of the parameter vector *y_height*; Table 4). After that, the plant height remains constant. This can work reasonably well for pure cultures. However, the direct link between crop height and stem production may be problematic in intercrops, as the plant heights of the two species grown in intercrop can differ, unlike in a pure culture. Since APSIM lacks a mechanism that allows plants to enhance the growth of the stem if they are shaded (Knörzer et al., 2011), the shorter crop species, which was almost always wheat in this study,

will be increasingly shaded by the taller species over time without having a mechanism to adapt. This will result in an overestimation of the competitive ability of the taller crop (faba bean in our case) to intercept radiation relative to wheat (Fig. 8). This likely explains why wheat aboveground dry matter is systematically underestimated and faba bean aboveground dry matter is systematically overestimated.

The performance of APSIM in intercrops could thus be substantially improved by assuming that plant height growth is independent of the stem weight per plant in both APSIM WHEAT and APSIM FABABEAN, for instance by implementing in the source code a logistic relation between thermal time from emergence and height growth (Kropff and Van Laar, 1993; Gou et al., 2017b). Another solution would be to implement a relationship in APSIM that increases the biomass partitioning of dry matter of one species to the stem, if that species is shaded by its companion species. This solution would require changes in the APSIM source code. An alternative approach not requiring a change to the way in which APSIM simulates plant growth would be directly calibrating the model in intercrops (Chimonyo et al., 2016). While potentially improving APSIM performances in intercrop, this approach, however, does still not explicitly consider a potential adjustment in plant features caused by interspecific competition. Then, crop and cultivar parameters do not only depend on the crop and cultivar traits, but also on the cropping system in which they are grown. This will likely make it hard to use the same parameters values of a cultivar in APSIM that is calibrated

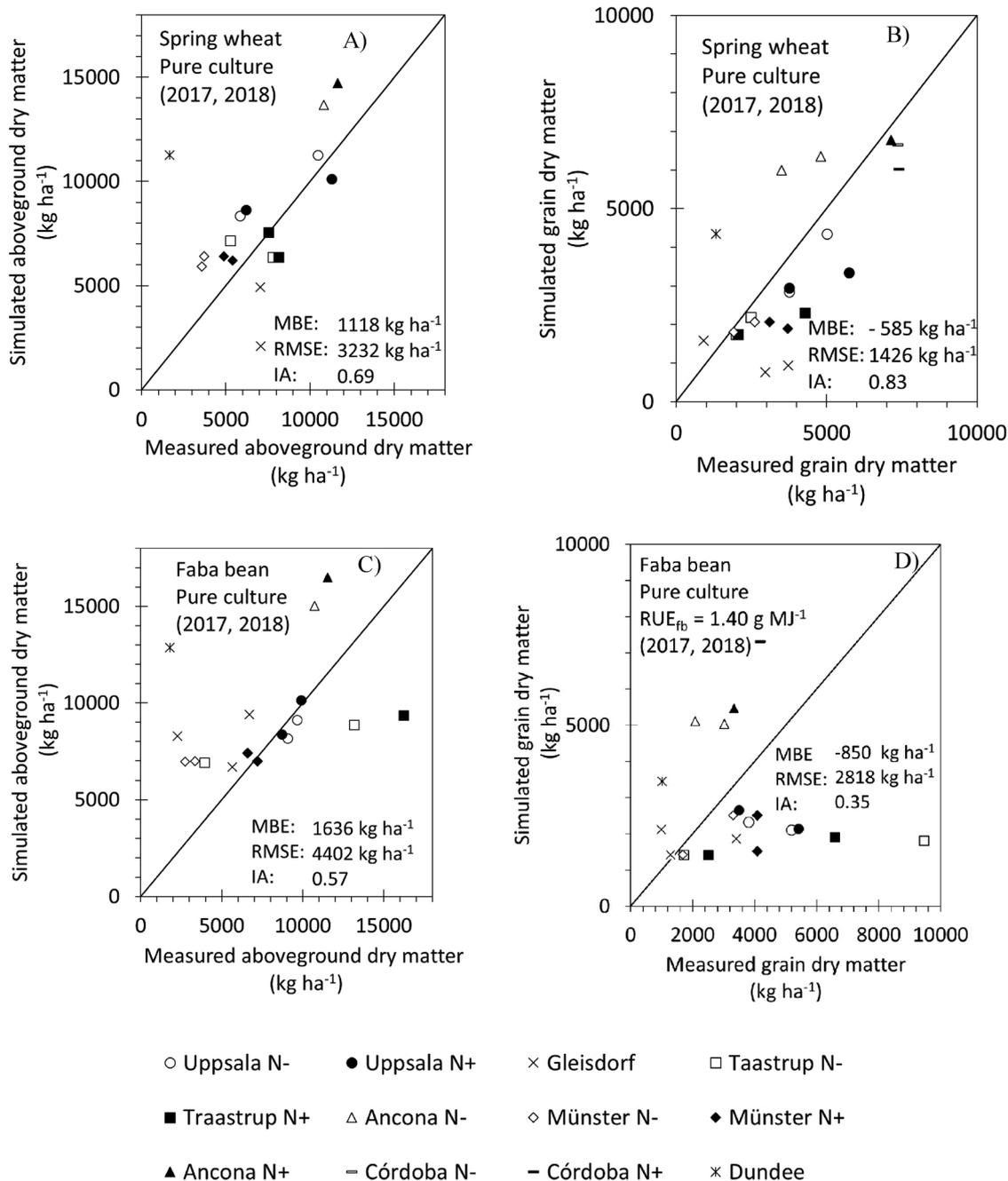


Fig. 6. Comparison of measured (symbols) and simulated (lines) aboveground dry matter (A, C) and grain dry matter (B, D) of pure cultures of wheat (A-B) and faba bean (C-D). Data were collected in Córdoba (Spain), Uppsala (Sweden), Gleisdorf (Austria), Taastrup (Denmark), Ancona (Italy), Münster (Germany) and Dundee (United Kingdom), as indicated in the legend. Open symbols indicate low management (N-) treatments and closed symbols indicate conventional management treatments (N+).

for one cropping system to simulate another system.

5. Conclusions and recommendations for future research

Crop growth models offer a powerful tool to explore the performance of a variety of species and management choices, under different pedoclimatic conditions, including future climates and the use of still to-be-bred varieties. For intercropping, this means quantifying the net effects of intra- and inter-specific competition between the two species and varieties, and determining when facilitation effects prevail over competition, leading to higher yields than the corresponding pure crop and, in general, reducing the negative environmental effects of agriculture. This is a necessary step when aiming at using intercropping as

one measure to support the ecological intensification of agriculture. Clearer understanding of the advantages and disadvantages can also facilitate the uptake of intercrop by farmers – currently, this is not a preferred diversification practice for most farmers (Kleijn et al., 2019). Further, by testing different parameters and aspects of crop interactions, models can provide further insight into the mechanisms driving intra- and inter-specific competition in field crops. Yet, relatively few crop growth models have been tested for their ability to reproduce intercrops under a variety of pedoclimatic conditions.

We successfully calibrated and validated APSIM WHEAT and APSIM FABABEAN on field trials of pure cultures of wheat and faba bean in the Netherlands. We then adjusted the phenological parameters of APSIM WHEAT and APSIM FABABEAN to various locations in Europe, in order

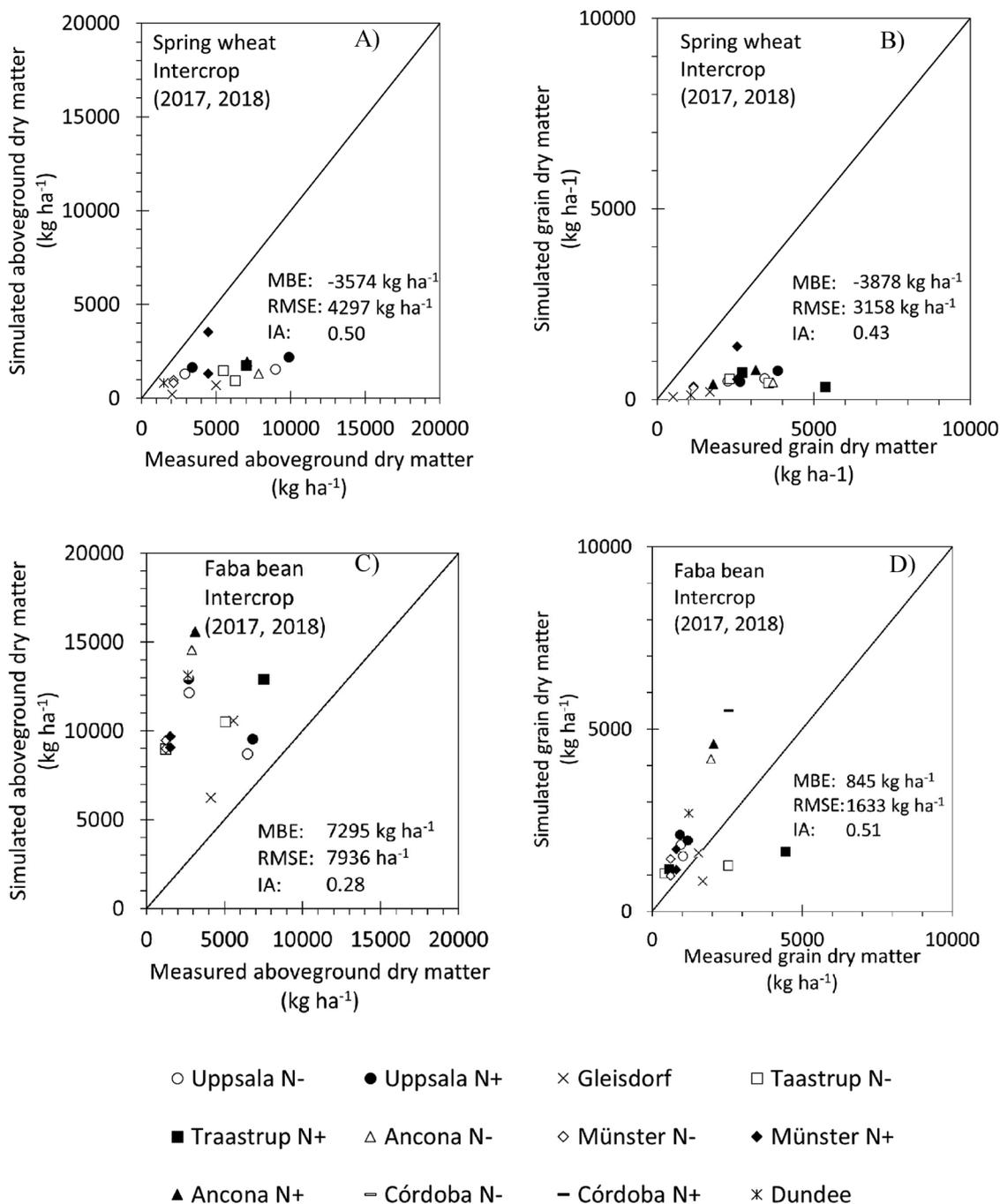


Fig. 7. Comparison of measured (symbols) and simulated (lines) aboveground dry matter (A, C) and grain dry matter (B, D) of intercrops of wheat (A-B) and faba bean (C-D). Data were collected in Córdoba (Spain), Uppsala (Sweden), Gleisdorf (Austria), Taastrup (Denmark), Ancona (Italy), Münster (Germany) and Dundee (United Kingdom), as indicated in the legend. Open symbols indicate low input management (N-) treatments and closed symbols indicate conventional management treatments (N+).

to simulate field trials of pure cultures and intercrops of wheat and faba bean.

In pure cultures, APSIM WHEAT performed well in reproducing the observed grain yields and aboveground dry matter. APSIM FABABEAN was, to a certain extent, capable of reproducing most of the aboveground dry matter observations of pure cultures. But it had limited performances regarding grain yields, indicating limited capability of APSIM FABABEAN to simulate partitioning to grain under European growing conditions. In wheat-faba bean intercrops, APSIM WHEAT systematically underestimated the aboveground dry matter of wheat and APSIM FABABEAN systematically overestimated the aboveground dry matter.

Further evaluation of the results suggested some possible

explanations for the limited performance of APSIM in simulating growth and yield of faba bean in pure stands and intercropping. We recommend further investigation of the following aspects:

1) Can the performance of APSIM FABABEAN to simulate the yield be improved by examining how the harvest index related parameters differ among cultivars and regions in Europe and/or by making the model for biomass partitioning to grains more mechanistic?

2) As competition for light is a key driver in the model, simulation of crop height is decisive for the partitioning of light and the growth of species in intercrops. In APSIM, height growth is linked to biomass growth. A model for height growth using a logistic growth function could improve the capability of APSIM WHEAT and APSIM FABABEAN

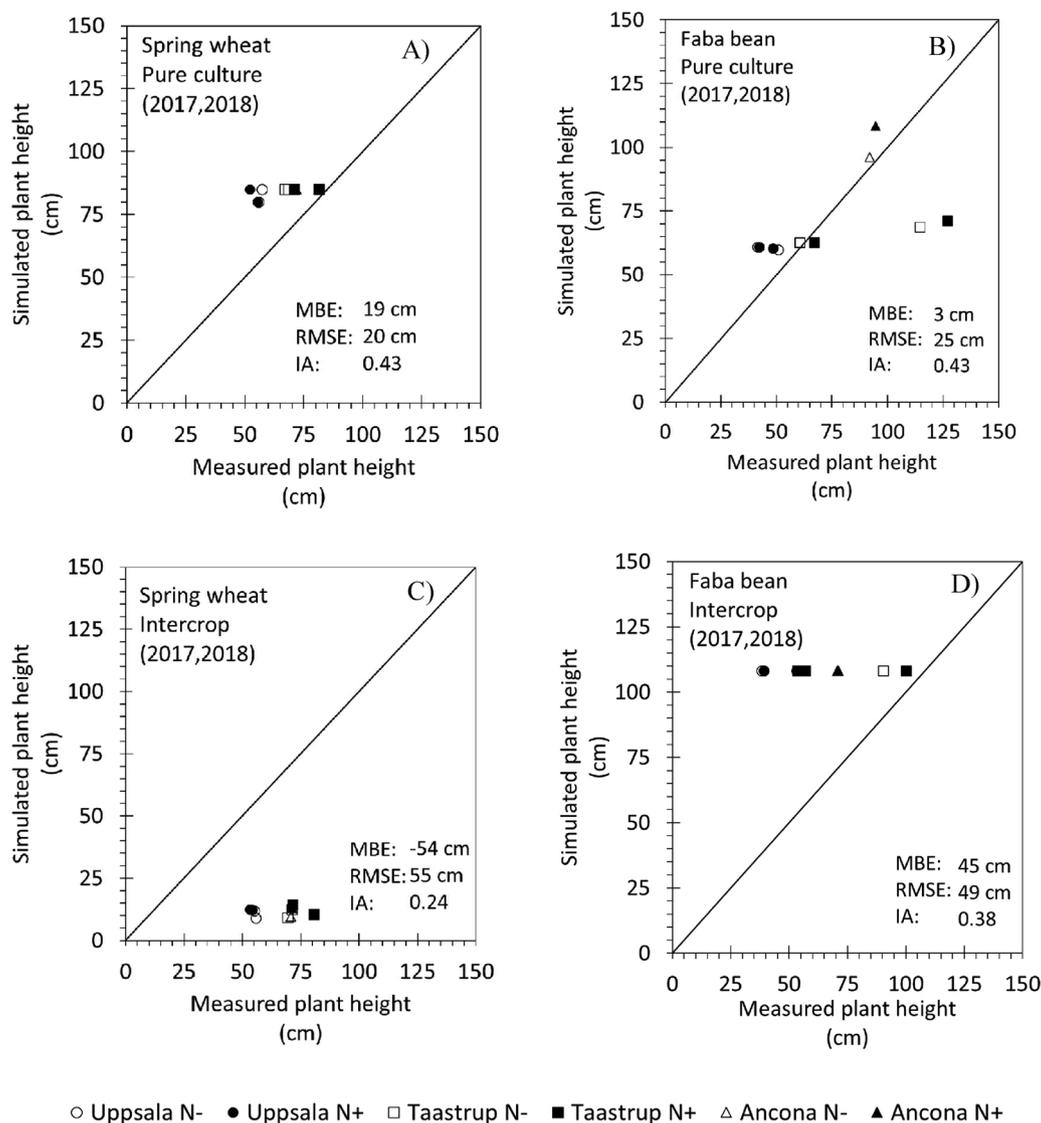


Fig. 8. Comparison of measured (symbols) and simulated (lines) plant heights at the end of the growing season of wheat (A, C) and faba bean (B, D) of pure cultures (A, B) and faba bean (C-D). Data were collected in Uppsala (Sweden), Taastrup (Denmark), and Ancona (Italy), as indicated by the legend. Open symbols indicate low input management (N-) treatments and closed symbols indicate conventional management treatments (N+).

to simulate dry matter production in wheat-faba bean intercrops. Alternatively, APSIM should be extended with a module that increases the biomass partitioning to the stem to one species in an intercrop, if it gets shaded by its companion species.

These aspects directly link to mechanisms currently not well captured by APSIM. Their proper inclusion is a necessary step to improve APSIM performance at least in the cereal-legume intercrop system and under the environmental conditions considered here. Only a model able to realistically represent pure cultures and intercrops using different cultivars and under a variety of pedoclimatic conditions such as those covered by our data would allow the design and evaluation of ecologically intensive cropping systems based on intercrops.

CRedit authorship contribution statement

Herman N.C. Berghuijs: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing - original draft, Writing - review & editing, Visualization. **Martin Weih:** Conceptualization, Resources, Investigation, Writing - original draft, Writing - review & editing, Supervision, Project administration, Funding acquisition. **Wopke van der Werf:** Resources, Writing - original draft,

Writing - review & editing. **Alison J. Karley:** Conceptualization, Investigation, Data curation, Resources, Writing - review & editing, Project administration, Funding acquisition. **Eveline Adam:** Conceptualization, Investigation, Data curation, Resources, Writing - review & editing, Project administration, Funding acquisition. **Ángel M. Villegas-Fernández:** Conceptualization, Data curation, Investigation, Resources, Writing - review & editing, Funding acquisition. **Lars P. Kiær:** Conceptualization, Data curation, Investigation, Resources, Writing - review & editing, Project administration, Funding acquisition. **Adrian C. Newton:** Conceptualization, Data curation, Investigation, Resources, Writing - review & editing, Funding acquisition. **Christoph Scherber:** Conceptualization, Data curation, Investigation, Resources, Writing - review & editing, Project administration, Funding acquisition. **Stefano Tavoletti:** Conceptualization, Data curation, Investigation, Resources, Writing - review & editing, Funding acquisition. **Giulia Vico:** Conceptualization, Investigation, Methodology, Writing - original draft, Writing - review & editing, Supervision, Project administration, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no conflict of interest

Acknowledgements

HNCB, MW, AJK, EA, ÁMV, LPK, ACN, CS, ST, and GV received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No. 727284, project DIVERSify: Designing InnoVative plant teams for Ecosystem Resilience and agricultural Sustainability, <https://www.plant-teams.eu/>. WvdW received funding from the European Union Horizon 2020 Research and Innovation Programme under grant agreement no. 727217, project ReMIX: Redesigning European cropping systems based on species MIXtures, <https://www.remix-intercrops.eu>.

We would like to thank Fang Gou (formerly at Wageningen UR, the Netherlands) for sharing her data with us and for her explanations of her experiments. We would also like to thank Henk Wösten (Wageningen UR, the Netherlands) for advice on the calculations on soil textural data and Marcos Lana (Swedish University of Agricultural Sciences, Sweden) for organizing two brainstorming sessions on crop model calibration.

References

- Asseng, S., van Keulen, H., Stol, W., 2000. Performance and application of the APSIM n/wheat model in the Netherlands. *Eur. J. Agron.* 12 (1), 37–54.
- Berghuijs, H.N.C., Wang, Z.S., Stomph, T.J., Wei, M., Van der Werf, W., Vico, G., 2020. Identification of species traits enhancing yield in wheat-faba bean intercropping: development and sensitivity analysis of a minimalist mixture model. *Plant Soil* 455, 203–226.
- Berntsen, J., Hauggard-Nielsen, H., Olesen, J.E., Petersen, B.M., Jensen, E.S., Thomsen, A., 2004. Modelling dry matter production and resource use in intercrops of pea and barley. *Field Crops Res.* 88 (1), 69–83.
- Bommarco, R., Kleijn, D., Potts, S.G., 2013. Ecological intensification: harnessing ecosystem services for food security. *Trends Ecol. Evol.* 28 (4), 230–238.
- Boons-Prins, E.R., De Koning, G.H.J., Van Diepen, C.A., De Vries, F.W.T., 1993. In: Van Keulen, H., Goudriaan, J. (Eds.), *Crop Specific Simulation Parameters. Simulation Reports CABO-TT*. DLO Centre for Agrobiological Research, Wageningen.
- Brisson, N., Gary, C., Justes, E., Roche, R., Mary, B., Ripoche, D., Zimmer, D., Sierra, J., Bertuzzi, P., Burger, P., et al., 2003. An overview of the crop model STICS. *Eur. J. Agron.* 18 (3–4), 309–332.
- Brisson, N., F. B. Ozier-Lafontaine, H., Tournebise, R., Sinoquet, H., 2004. Adaptation of the crop model STICS to intercropping. Theoretical basis and parameterisation. *Agronomie* 24, 409–424.
- Carberry, P.S., Adiku, S.G.K., McCown, R.L., Keating, B.A., 1996. Application of the APSIM cropping systems model to intercropping systems. In: Ito, C., Johansen, C., Adu-Gyamfi, K., Katayama, K., Kumar-Rao, J.V.D.K., Rego, T.J. (Eds.), *Dynamics of Roots and Nitrogen in Cropping Systems of the Semi-Arid Tropics: Japan International Resource Centre of Agricultural Sciences*, 663–648.
- Chimonyo, V.G.P., Modi, A.T., Mabhauthi, T., 2016. Simulating yield and water use of a sorghum-cowpea intercrop using APSIM. *Agric. Water Manag.* 177, 317–328.
- Cong, W.F., Hoffland, E., Li, L., Six, J., Sun, J.H., Bao, X.G., Zhang, F.S., van der Werf, W., 2015. Intercropping enhances soil carbon and nitrogen. *Glob. Chang. Biol.* 21 (4), 1715–1726.
- Corre-Hellou, G., Fustec, J., Crozat, Y., 2006. Interspecific competition for soil N and its interaction with N₂ fixation, leaf expansion and crop growth in pea-barley intercrops. *Plant Soil* 282 (1–2), 195–208.
- Corre-Hellou, G., Brisson, N., Launay, M., Fustec, J., Crozat, Y., 2007. Effect of root depth penetration on soil nitrogen competitive interactions and dry matter production in pea-barley intercrops given different soil nitrogen supplies. *Field Crops Res.* 103 (1), 76–85.
- Corre-Hellou, G., Faure, M., Launay, M., Brisson, N., Crozat, Y., 2009. Adaptation of the STICS intercrop model to simulate crop growth and N accumulation in pea-barley intercrops. *Field Crops Res.* 113 (1), 72–81.
- Evers, J.B., van der Werf, W., Stomph, T.J., Bastiaans, L., Anten, N.P.R., 2019. Understanding and optimizing species mixtures using functional-structural plant modelling. *J. Exp. Bot.* 70 (9), 2381–2388.
- Gaudio, N., Escobar-Gutierrez, A.J., Casadebaig, P., Evers, J.B., Gerard, F., Louarn, G., Colbach, N., Munz, S., Launay, M., Marrou, H., et al., 2019. Current knowledge and future research opportunities for modeling annual crop mixtures. A review. *Agron. Sustain. Dev.* 39 (2).
- Gou, F., van Ittersum, M.K., Wang, G.Y., van der Putten, P.E.L., van der Werf, W., 2016. Yield and yield components of wheat and maize in wheat-maize intercropping in the Netherlands. *Eur. J. Agron.* 76, 17–27.
- Gou, F., van Ittersum, M.K., Simon, E., Leffelaar, P.A., van der Putten, P.E.L., Zhang, L.Z., van der Werf, W., 2017a. Intercropping wheat and maize increases total radiation interception and wheat RUE but lowers maize RUE. *Eur. J. Agron.* 84, 125–139.
- Gou, F., van Ittersum, M.K., van der Werf, W., 2017b. Simulating potential growth in a relay-strip intercropping system: model description, calibration and testing. *Field Crops Res.* 200, 122–142.
- Goudriaan, J., 1988. The bare bones of leaf-angle distribution in radiation models for canopy photosynthesis and energy exchange. *Agric. For. Meteorol.* 43 (2), 155–169.
- Hengl, T., de Jesus, J.M., MacMillan, R.A., Batjes, N.H., Heuvelink, G.B.M., Ribeiro, E., Samuel-Rosa, A., Kempen, B., Leenaars, J.G.B., Walsh, M.G., et al., 2014. SoilGrids1km - Global soil information based on automated mapping. *PLoS One* 9 (8).
- Hengl, T., de Jesus, J.M., Heuvelink, G.B.M., Gonzalez, M.R., Kilibarda, M., Blagotic, A., Shangquan, W., Wright, M.N., Geng, X.Y., Bauer-Marschallinger, B., et al., 2017. SoilGrids250m: global gridded soil information based on machine learning. *PLoS One* 12 (2).
- Jacobsen, B.H., Olesen, J.E., Petersen, B.M., Bernsten, J., Boye, C., 1998. FASSET - a dynamic whole farm simulation model. In: Van Keulen, H., Lantinga, E.A., Van Laar, H.H. (Eds.), *Mixed Farming Systems in Europe*. Dronter, The Netherlands: Grafisch Service Centrum Van Gils B.V. Wageningen.
- Jensen, E.S., 1996. Grain yield, symbiotic N₂ fixation and interspecific competition for inorganic N in pea-barley intercrops. *Plant Soil* 182 (1), 25–38.
- Keating, B.A., Carberry, P.S., 1993. Resource capture and use in intercropping - solar radiation. *Field Crops Res.* 34 (3–4), 273–301.
- 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., et al., 2003. An overview of APSIM, a model designed for farming systems simulation. *Eur. J. Agron.* 18 (3–4), 267–288.
- Kleijn, D., Bommarco, R., Fijen, T.P.M., Garibaldi, L.A., Potts, S.G., Van der Putten, W.H., 2019. Ecological intensification: bridging the gap between science and practice. *Trends Ecol. Evol.* (Amst.) 34 (2), 154–166.
- Knörzer, H., Lawes, R., Robertson, M., Graeff-Hönniger, S., Claupein, W., 2011. Evaluation and performance of the APSIM crop growth model for German winter wheat, maize and fieldpea varieties within monocropping and intercropping systems. *J. Agric. Sci. Technol.* 1, 698–717.
- Kropff, M.J., 1989. Quantification of SO₂ Effects on Physiological Processes, Plant Growth and Crop Reproduction. PhD Dissertation. Wageningen Agricultural University Wageningen.
- Kropff, M.J., Van Laar, H.H., 1993. Modelling Crop-weed Interactions. BPCC Wheatons Ltd, Exter, UK.
- Leffelaar, P.A., 2014. Lintul-2: Water limited crop growth. A simple general crop growth model for water-limited growing conditions (example: spring wheat). In: Van Oijen, M., Leffelaar, P.A. (Eds.), *Quantitative Aspects of Crop Production*. Wageningen, pp. 6B1–6B47.
- Letourneau, D.K., Ambrecht, I., Rivera, B.S., Lerma, J.M., Carmona, E.J., Daza, M.C., Escobar, S., Galindo, V., Gutierrez, C., Lopez, S.D., et al., 2011. Does plant diversity benefit agroecosystems? A synthetic review. *Ecol. Appl.* 21 (1), 9–21.
- Li, C.J., Hoffland, E., Kuyper, T.W., Yu, Y., Li, H.G., Zhang, C.C., Zhang, F.S., van der Werf, W., 2020. Yield gain, complementarity and competitive dominance in intercropping in China: a meta-analysis of drivers of yield gain using additive partitioning. *Eur. J. Agron.* 113.
- Liebman, M., Dyck, E., 1993. Crop-rotation and intercropping strategies for weed management. *Ecol. Appl.* 3 (1), 92–122.
- Loreau, M., 2010. Chapter 2: the maintenance and functional consequences of species diversity. In: Loreau, M. (Ed.), *From Populations to Ecosystems: Theoretical Foundations for a New Ecological Synthesis (MBP-46)*. Princeton University Press.
- Malezieux, E., Crozat, Y., Dupraz, C., Laurans, M., Makowski, D., Ozier-Lafontaine, H., Rapidel, B., de Tourdonnet, S., Valantin-Morison, M., 2009. Mixing plant species in cropping systems: concepts, tools and models. A review. *Agron. Sustain. Dev.* 29 (1), 43–62.
- Martin-Guay, M.O., Paquette, A., Dupras, J., Rivest, D., 2018. The new Green Revolution: sustainable intensification of agriculture by intercropping. *Sci. Total Environ.* 615, 767–772.
- Meinke, H., Hammer, G.L., van Keulen, H., Rabbinge, R., Keating, B.A., 1997. Improving wheat simulation capabilities in Australia from a cropping systems perspective: water and nitrogen effects on spring wheat in a semi-arid environment. *Eur. J. Agron.* 7 (1–3), 75–88.
- Meinke, H., Hammer, G.L., van Keulen, H., Rabbinge, R., 1998a. Improving wheat simulation capabilities in Australia from a cropping systems perspective - III. The integrated wheat model (IWHEAT). *Eur. J. Agron.* 8 (1–2), 101–116.
- Meinke, H., Rabbinge, R., Hammer, G.L., van Keulen, H., Jamieson, P.D., 1998b. Improving wheat simulation capabilities in Australia from a cropping systems perspective - II. Testing simulation capabilities of wheat growth. *Eur. J. Agron.* 8 (1–2), 83–99.
- Mualem, Y., 1976. A new model for predicting the hydraulic conductivity of unsaturated porous media. *Water Resour. Res.* 12 (3), 513–523.
- Pribil, D.W., 2010. A critical review of the conventional SOC to SOM conversion factor. *Geoderma* 156 (3–4), 75–83.
- Probert, M.E., Dimes, J.P., Keating, B.A., Dalal, R.C., Strong, W.M., 1998. APSIM's water and nitrogen modules and simulation of the dynamics of water and nitrogen in fallow systems. *Agric. Syst.* 56 (1), 1–28.
- Pronk, A.A., Goudriaan, J., Stilma, E., Challa, H., 2003. A simple method to estimate radiation interception by nursery stock conifers: a case study of eastern white cedar. *NJAS* 51 (3), 279–295.
- Raseduzzaman, M., Jensen, E.S., 2017. Does intercropping enhance yield stability in arable crop production? *Eur. J. Agron.* 91, 25–33.
- Robertson, M.J., Carberry, P.S., Huth, N.I., Turpin, J.E., Probert, M.E., Poulton, P.L., Bell, M., Wright, G.C., Yeates, S.J., Brinsmead, R.B., 2002. Simulation of growth and

- development of diverse legume species in APSIM. *Aust. J. Agric. Res.* 53 (4), 429–446.
- Salo, T.J., Palosuo, T., Kersebaum, K.C., Nendel, C., Angulo, C., Ewert, F., Bindi, M., Calanca, P., Klein, T., Moriondo, M., et al., 2016. Comparing the performance of 11 crop simulation models in predicting yield response to nitrogen fertilization. *J. Agric. Sci.* 154 (7), 1218–1240.
- Tamburini, G., Bommarco, G., Wanger, T.C., Kremen, C., Van der Heijden, M.G.A., Liebman, M., Hallin, S., 2020. Agricultural diversification promotes multiple ecosystem services without compromising yield. *Sci. Adv.* 6 (45), eaba1715.
- Tan, M., Gou, F., Stomph, T.J., Jing, W., Wen, Y., Zhang, L., Qiang, C., Van der Werf, W., 2020. Dynamic process-based modelling of crop growth and competitive water extraction in relay strip intercropping: model development and application to wheat maize intercropping. *Field Crops Res.* 243.
- Turpin, J.E., Robertson, M.J., Hillcoat, N.S., Herridge, D.F., 2002. Fababean (*Vicia faba*) in Australia's northern grains belt: canopy development, biomass, and nitrogen accumulation and partitioning. *Aust. J. Agric. Res.* 53 (2), 227–237.
- Turpin, J.E., Robertson, M.J., Haire, C., Bellotti, W.D., Moore, A.D., Rose, I., 2003. Simulating fababean development, growth, and yield in Australia. *Aust. J. Agric. Res.* 54 (1), 39–52.
- Van Genuchten, M.T., 1980. A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Sci. Soc. Am. J.* 44 (5), 892–898.
- Wallach, D., Makowski, D., Jones, J., Brun, F., 2018. *Working With Dynamic Crop Models: Methods, Tools and Examples for Agriculture and Environment*, second edition.
- Wang, E., Van Oosterom, E., Meinke, H., Asseng, S., Robertson, M.J., Huth, N.I., Keating, B.A., Probert, M.E., 2003. The new APSIM-wheat model: performance and future improvements. In: 11th Australian Society of Agronomy Conference. Geelong, Australia: Australian Society of Agronomy.
- Willey, R.W., Rao, M.R., 1980. Intercroppings - its importance and research needs. Part 1. Competition and yield advantages. *Field Crops Abstracts* 32 (1), 1–10.
- Willmott, C.J., 1981. On the validation of models. *Phys. Geogr.* 2, 184–194.
- Willmott, C.J., Wicks, D.E., 1980. An empirical method for the spatial interpolation of monthly precipitation within California. *Phys. Geogr.* 1, 59–73.
- Wösten, J.H.M., Nemes, A., 2004. Pedotransfer functions in Europe. In: Pachepsky, Y., Rawls, W.J. (Eds.), *Developments in Soil Science*. Elsevier, pp. 431–435.
- Wösten, J.H.M., Lilly, A., Nemes, A., Le Bas, C., 1999. Development and use of a database of hydraulic properties of European soils. *Geoderma* 90 (3–4), 169–185.
- Xu, Z., Li, C.J., Zhang, C.C., Yu, Y., van der Werf, W., Zhang, F.S., 2020. Intercropping maize and soybean increases efficiency of land and fertilizer nitrogen use; A meta-analysis. *Field Crops Res.* 246.
- Yu, Y., Stomph, T.J., Makowski, D., van der Werf, W., 2015. Temporal niche differentiation increases the land equivalent ratio of annual intercrops: a meta-analysis. *Field Crops Res.* 184, 133–144.
- Yu, Y., Stomph, T.J., Makowski, D., Zhang, L.Z., van der Werf, W., 2016. A meta-analysis of relative crop yields in cereal/legume mixtures suggests options for management. *Field Crops Res.* 198, 269–279.
- Zheng, B., Chenu, K., Doherty, A., Chapman, S., 2015. The APSIM-Wheat Module (7.5 R3008). CSIRO, Australia.