

Improving WOFOST model to simulate winter wheat phenology in Europe: Evaluation and effects on yield



A. Ceglar^{a,*}, R. van der Wijngaart^b, A. de Wit^b, R. Lecerf^a, H. Boogaard^b, L. Seguini^a, M. van den Berg^a, A. Toreti^a, M. Zampieri^a, D. Fumagalli^a, B. Baruth^a

^a European Commission, Joint Research Centre, via Enrico Fermi 2749, Ispra 21027, Italy

^b Earth Observation and Environmental Informatics, Wageningen Environmental Research (Alterra), P.O. Box 47AA, Wageningen 6700, The Netherlands

ARTICLE INFO

Keywords:

Triticum aestivum
Europe
Phenology
WOFOST
Calibration
Crop yield forecasting

ABSTRACT

This study describes and evaluates improvements to the MARS crop yield forecasting system (MCYFS) for winter soft wheat (*Triticum aestivum*) in Europe, based on the WOFOST crop simulation model, by introducing autumn sowing dates, realistic soil moisture initialization, adding vernalization requirements and photoperiodicity, and phenology calibration. Dataset of phenological observations complemented with regional cropping calendars across Europe is used. The calibration of thermal requirements for anthesis and maturity is done by pooling all available observations within European agro-environmental zones and minimizing an objective function that combines the differences between observed and simulated anthesis, maturity and harvest dates. Calibrated phenology results in substantial improvement in simulated dates of anthesis with respect to the original MCYFS simulations. The combined improvements to the system result in a physically more plausible spatial distribution of crop model indicators across Europe. Crop yield indicators point to better agreement with recorded national winter wheat yields with respect to the original MCYFS simulations, most pronounced in central, eastern and southern Europe. However, model skill remains low in large parts of western Europe, which may possibly be attributed to the impacts of wet conditions.

1. Introduction

Wheat provides nearly 20% of all calories consumed by people worldwide and makes significant contribution to animal feed (Shiferaw et al., 2013). Wheat production in Europe represents 32.2% of the total global wheat production (averaged over the period 1994–2014; Faostat, 2017), making it the second most important producer after Asia.

Early information on crop yields and production in Europe is vital for policy making within the Common Agricultural Policy (CAP). For this purpose, the Monitoring Agricultural Resources Unit (MARS) Crop Yield Forecasting System (MCYFS; Micale and Genovese, 2004; Lazar and Genovese, 2004; Genovese and Bettio, 2004) has been established by the Joint Research Centre (JRC) of the European Commission with the aim to monitor crop growth and provide seasonal yield forecasts of key European annual crops. The mechanistic and dynamic crop model WOFOST (Boogaard et al., 1998; Supit et al., 1994; Van Diepen et al., 1989) represents an essential component of the system as it simulates the impact of weather and crop management on growth and development. The regional application of the WOFOST model within the MCYFS has also often been used also for climate change impact

assessments and yield gap analysis (e.g. Blanco et al., 2017; Boogaard et al., 2013).

Regional application of a crop model in a monitoring system, such as MCYFS, strongly depends on the values of crop model parameters (Ceglar and Kajfež-Bogataj, 2012; Jagtap and Jones, 2002). Crop model parameters are usually calibrated based on location-specific observations and optimized for areas with relatively homogeneous conditions (e.g. Wei et al., 2009). Even though the main aim of regional crop monitoring systems is the reliable estimation of yields at regional scale, observations from local fields and experimental stations remain of utmost importance for calibration.

Accurate simulation of the phenological development is vital for assessment of the impact of weather conditions on crop growth and development. For example, unfavourable weather conditions, such as heat stress and drought, around the short period of flowering, can substantially reduce yield potential (e.g. Barnabás et al., 2008; Porter and Semenov, 2005). High temperatures also affect final yields differently during different phenological stages, by reducing the grain filling period, impacting the photosynthetic rate, reducing water use efficiency, favoring pest and diseases, and eventually damaging the plant

* Corresponding author.

E-mail address: andrej.cegler@ec.europa.eu (A. Ceglar).

cells (Lobell and Gourdji, 2012).

When it comes to regional crop model applications, the most common parameters to be calibrated belong to the phenological model (e.g. Angulo et al., 2013; Harrison and Butterfield, 1996). As for other parameters (e.g. related to leaf area growth, CO₂ uptake, biomass allocation, root growth and soil water balance), sparse spatial distribution of field experimental data often limits their calibration (Jones et al., 2017; Levis, 2014; Iizumi et al., 2014). Literature review and/or expert opinion can, in some circumstances, provide an alternative to proper model calibration (within the plausible range) of related model parameters (e.g. Wallach et al., 2013). Regional crop yield data (i.e. sub-national and/or national level) can be usually retrieved from national statistical yields, but these yields provide insufficient information to calibrate crop model parameters, which were originally developed at field scale under (near-)optimal growing conditions (Therond et al., 2011). The lack of crop measurements often limits the potential of crop models to reproduce regional yields, and often model parameters refer to old varieties (Rötter et al., 2011). Regional datasets of phenological observations that can be used for monitoring and climatological purposes, such as the one produced by COST-action 725 (Koch et al., 2010), can substantially contribute to more reliable simulation of phenological development on regional and continental scales.

MCYFS provides key tools for crop yield monitoring and forecasting on the European level in the support of CAP. The system is regularly reviewed, updated and extended (e.g. Lopez-Lozano et al., 2015; Confalonieri et al., 2010). As a result of several expert meetings, the following critical aspects were identified as a subject to improvement: (i) initializing the WOFOST model at climatically suitable autumn sowing date, (ii) improving phenology simulation by introducing vernalization requirements and photoperiodicity, and (iii) performing a spatial calibration of the phenology of the WOFOST model across Europe. The objective of this study is to describe these improvements and to assess their impact on crop yield simulations at national and European level (see Fig. 1). For these purposes, a robust calibration procedure is developed, suitable for regional crop model simulations with heterogeneous spatial density of field observations. This phenology calibration and validation is performed using approximately 48,000 observational records across Europe. Physical plausibility of simulated crop model variables is assessed and simulated wheat yields are compared with recorded national statistical winter wheat yields.

2. Data and methods

The main building blocks for this study are the original MCYFS system (described in Section 2.2) and the JRC database of observed phenological data across Europe (described in Section 2.1). Fig. 1 presents the schematic overview of the analysis conducted for the purpose of this study; each of the workflow steps is described in the delineated sections.

2.1. Agro-phenological database

The agro-phenological database, constructed and maintained by the JRC, is used as a source of information for crop calendars (sowing and harvest dates) and key phenological stages of winter soft wheat: emergence, anthesis and maturity. The database contains > 48,000 observational data records for winter soft wheat across Europe (Fig. S1), which were retrieved from different projects and/or activities: ASEMARS (Baruth and Kucera, 2007), COST 734 (Orlandini and Nejedlik, 2012), (GISAT, 2003) and (AETS, 2013). Additional data were obtained through bilateral activities with the Deutsche Wetter Dienst (DWD) and INRA (PHETEC database – INRA Agroclim).

Several quality and consistency measures are applied to check physical consistency and plausibility. As several European regions are very scarcely represented by field observations (especially in eastern Europe, Fig. S1), the regional calendars obtained within the GISAT

project are used. Key phenological stages in regional calendars are representative for administrative regions, obtained either from systematic observations or expert opinion (GISAT, 2003).

2.2. Regional implementation of the WOFOST crop model in the MCYFS

The crop growth simulation model WOFOST is a core element of the MCYFS. It calculates daily biomass accumulation and its allocation to different crop components using a photosynthesis approach. Modules in WOFOST include phenological development, light interception, gross CO₂ assimilation, growth and maintenance respiration, dry matter partitioning, source and sink limited leaf area development, soil water balance and soil nutrition balance. WOFOST distinguishes three levels of crop production: potential production (determined by crop variety, radiation and temperature), water limited production (water availability limits potential production) and nutrient limited production (in which nutrient availability limits water limited production). In the MCYFS potential and water limited productions are analyzed.

The regional implementation of WOFOST in the MCYFS simulates crop growth and development for every location with a unique set of weather, soil and crop characteristics. Weather and crop data are assumed to be homogeneous within a 25 km × 25 km grid, whereas soil characteristics are assumed to be homogeneous within a soil typological unit. By overlaying the 25 km × 25 km climatic grid, the soil map (Jones and Buckley, 1996) and the arable land cover map (Baruth and Kucera, 2007), simulation units are determined. More details on the regional implementation of the WOFOST model can be found in Boogaard et al. (2013).

The simulations for winter crops start on the 1st of January, and do not include effects of vernalization and photoperiodicity (Micale and Genovese, 2004). The reasons for doing this are mostly pragmatic and date back to the origins of the development of MCYFS in the early nineties.

The key outputs of the WOFOST crop model, used in subsequent steps of MCYFS, are: leaf area index (LAI), water-limited above-ground biomass weight (BIOM), potential and water-limited weight of storage organs (PSO and WLO), crop water requirement (TWR), crop water consumption (TWC) and relative soil moisture (SM). Even though the crop model simulations are performed on daily time step, the results are temporally aggregated and stored on dekad basis (i.e. 10 days). Simulated yields and other crop model indicators are spatially aggregated over the agricultural areas to the national levels by using the procedure described in Boogaard et al. (2013).

2.3. Zonation of winter wheat varieties in Europe

The Global Agro-Environmental Stratification level 4 (GAES; Mücher et al., 2016) is here used to estimate the zonation of winter wheat varieties across Europe (hereafter agro-environmental zones, see Fig. S1). GAES derives stratified agricultural production zones according to the agro-environmental characteristics, including climate, elevation, agro-management factors (e.g. irrigation, phenological development, crop type) and field size parameters.

2.4. Estimation of sowing dates

For this study, we follow the method described by Waha et al. (2012) by identifying the seasonality type using the MCYFS weather database and conclude that only for the most southern parts of the MCYFS pan-European window, sowing dates are triggered by the main wet season. Temperature seasonality is the prevailing type: farmers tend to plant winter wheat early enough to ensure proper growth before winter dormancy, resulting in a tendency of earlier planting in colder regions. Accordingly, we estimate the sowing date for each climatic grid cell as the first day when the 7-day running long-term-average daily temperature T_d drops below a threshold temperature close to the

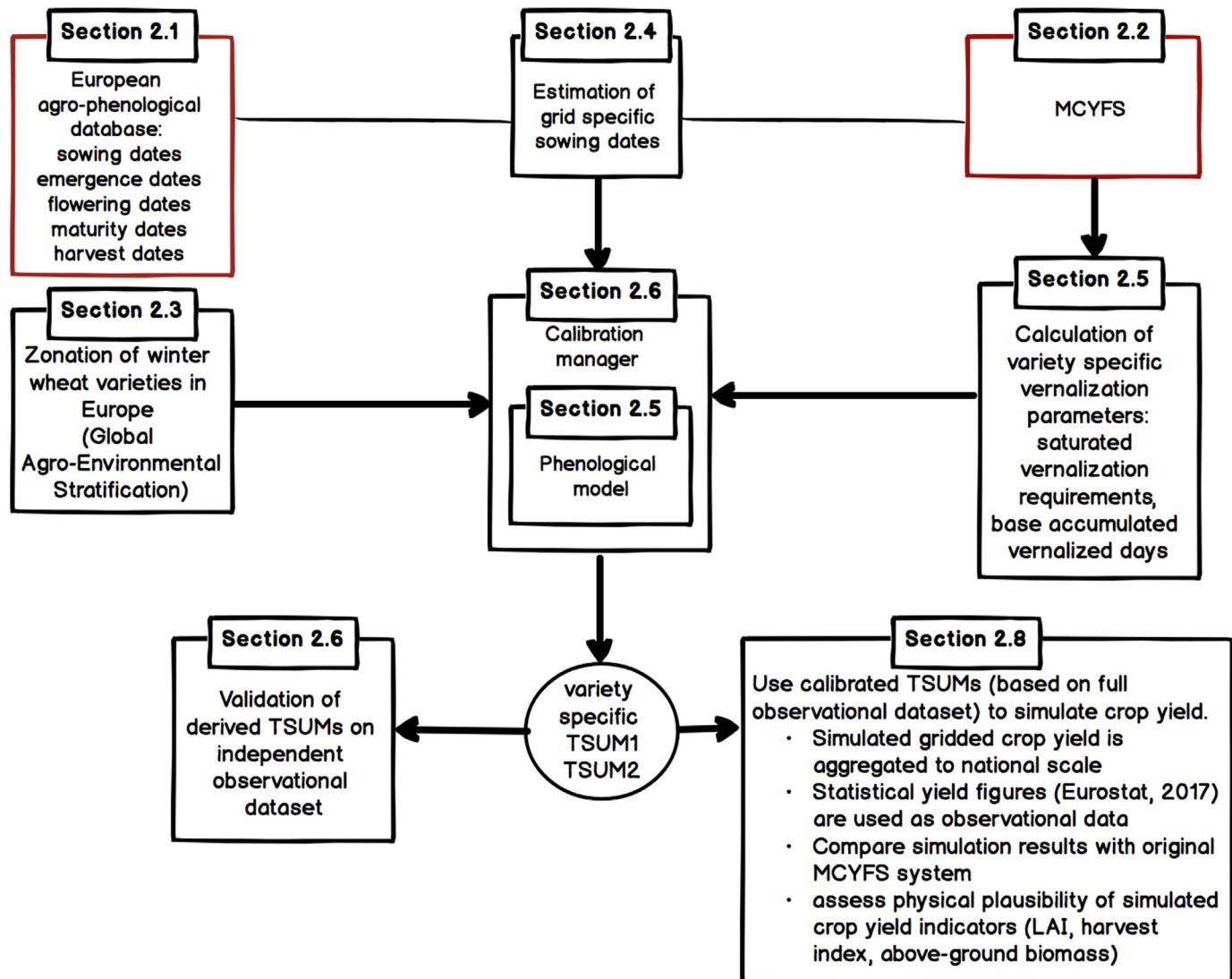


Fig. 1. The workflow of the analysis. Each of the boxes points to the specific section for detailed description.

optimal temperature for vernalization, initially chosen as 12 °C (Waha et al., 2012; Harrison et al., 2000). The 7-day running average smooths the long-term-average daily temperature curve to prevent extremely early or late sowing.

To further align derived sowing dates with the observations, they are fine-tuned by iteratively lowering the original threshold until the bias between observed and derived sowing dates is removed. Zero bias at European level is achieved at 11.1 °C (Table 1), which is the threshold value used to derive final sowing dates for each grid cell. Indeed this is close to the temperature for optimal vernalization (i.e. 10 °C, Porter and Gawith, 1999; Wang and Engel, 1998). For grids in southern part of the Iberian Peninsula and northern Africa, where the saturated vernalization requirement is 0 (see Section 2.5), the above

vernalization related threshold does not apply. For those regions, observations and calendars from the agro-phenological database are used to calculate an average sowing date per zone.

2.5. Simulation of phenological development

To simulate the phenological development of winter wheat, the temperature is accumulated above a base temperature T_b (for winter wheat considered to be 0 °C), development rate remains constant above a certain maximum effective temperature $T_{max,e}$ (for winter wheat considered to be 30 °C) and is corrected for the effects of vernalization and photoperiod (Ewert et al., 1996; Porter, 1993):

$$DVS = \begin{cases} \sum_i \frac{\max(0, \min((T_i - T_b), T_{max,e}))}{TSUM1} \cdot V_{f,i} \cdot P_{f,i} & \text{for } DVS \leq 1 \\ \sum_i \frac{\max(0, \min((T_i - T_b), T_{max,e}))}{TSUM2} & \text{for } 1 < DVS \leq 2 \end{cases} \quad (1)$$

where DVS represents the development stage (0 - emergence, 1 - anthesis and 2 - maturity), T_b the base temperature (Table S2), T_i the daily average temperature, $T_{max,e}$ the maximum effective daily average temperature, $V_{f,i}$ the vernalization factor and $P_{f,i}$ the photoperiod factor on day i ; $TSUM1$ and $TSUM2$ represent the length (in effective degrees days) of vegetative and reproductive periods, respectively.

The photoperiod factor is calculated as:

Table 1
Spatio-temporal averages of observed and simulated sowing dates, bias and absolute difference on European level.

	Observed sowing dates	Simulated sowing dates	Bias	Absolute difference
Average	286	286	0	12
Standard deviation	22.5	19.3	15.8	11.2

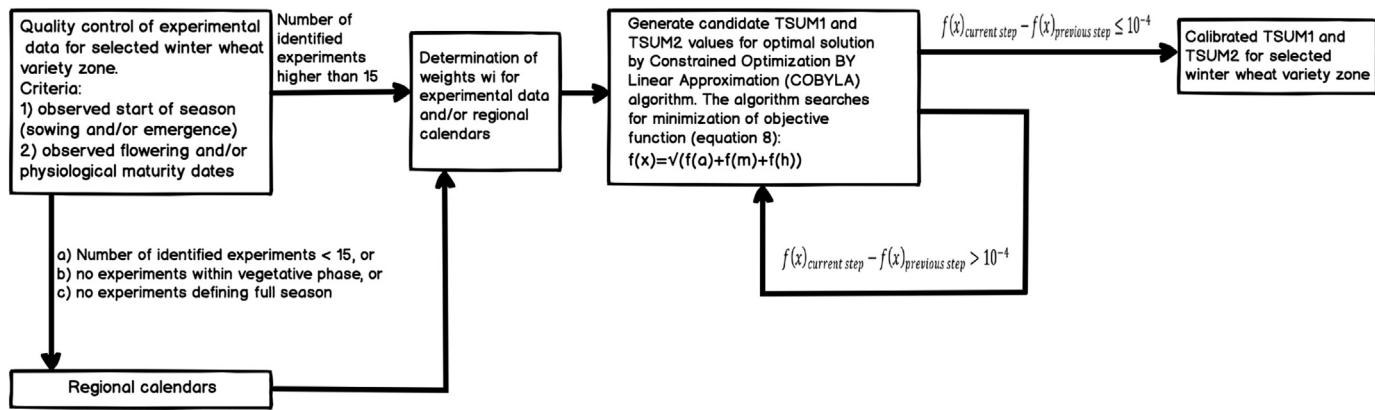


Fig. 2. The workflow of the calibration procedure, which is applied for each agro-environmental zone.

$$P_{f,i} = \begin{cases} \frac{P_i - P_b}{P_{opt} - P_b}; P_b \leq P_i \leq P_{opt} \\ 1; P_i > P_{opt} \\ 0; P_i < P_b \end{cases} \quad (2)$$

where P_i (h d^{-1}) is the daily photoperiod, P_b (h d^{-1}) is the base photoperiod (i.e. the photoperiod below which no development in long-day plants is observed, Table S2), P_{opt} (h d^{-1}) is the optimum photoperiod (i.e. the photoperiod above which no photoperiod-induced delay in long-day plants is observed). The daily photoperiod P_i is calculated based on latitude and Julian day (Monteith and Unsworth, 1990). P_{opt} is assumed to take a value of the local maximum day length occurring on the 21st of June (van Bussel et al., 2015).

The vernalization process depends on daily temperature, and its effectiveness $V_{eff,i}$ is calculated as:

$$V_{eff,i} = \begin{cases} \frac{T_i - T_{v1}}{T_{v2} - T_{v1}}; T_{v1} \leq T_i \leq T_{v2} \\ 1; T_{v2} \leq T_i \leq T_{v3} \\ \frac{T_{v4} - T_i}{T_{v4} - T_{v3}}; T_{v3} \leq T_i \leq T_{v4} \\ 0; T_i < T_{v1} \text{ or } T_i > T_{v4} \end{cases} \quad (3)$$

where T_{v1} and T_{v4} represent, respectively, the minimum and maximum temperature for effective vernalization (Table S2); T_{v2} and T_{v3} represent, respectively, the minimum and maximum temperature for optimal vernalization and T_i is the daily mean temperature. These cardinal temperatures for vernalization vary depending on the variety (Porter and Gawith, 1999; Ewert et al., 1996), resulting in wide plausible range (Table S2). However, given the lack of observational data to properly calibrate these parameters across Europe, equal cardinal temperatures are assumed to characterize all varieties. The vernalization factor is calculated as:

$$V_{f,i} = \begin{cases} 0; V_{DD} < V_b \\ \frac{V_{DD} - V_b}{V_{sat} - V_b}; V_b \leq V_{DD} \leq V_{sat} \\ 1; V_{DD} > V_{sat} \text{ or } DVS > 0.3 \end{cases} \quad (4)$$

$$V_{DD} = \sum_{j=1}^i V_{eff,j} \quad (5)$$

where V_{DD} is the accumulated effective vernalized days from emergence ahead (Eq. (5)), V_{sat} represents the saturated vernalization requirements (required duration of exposure to vernalizing temperatures) and V_b is the base accumulated vernalized days (assumed to be one fifth of V_{sat} Wang and Engel, 1998). The effect of vernalization is stopped when either V_{DD} surpasses the V_{sat} or when the critical development stage of 0.3 is reached; the latter is introduced to improve the model stability in order to avoid the anthesis not being reached in case when saturated vernalization requirements are not met due to potentially too high V_{sat} values.

Vernalization requirements are assumed to vary among cultivars

(i.e. V_{sat} and V_b), whereas vernalization effectiveness V_{eff} is assumed to be equal among cultivars. As for the saturated vernalization requirements, a modified approach of van Bussel et al. (2015) is followed, with the underlying assumption that the frost period has a maximum length of five months in winter wheat growing regions. The saturated vernalization requirement in year y is calculated as:

$$V_{sat,y} = \sum_{m=1}^N V_{sat,m,y} \quad (6)$$

$$V_{sat,m} = \begin{cases} V_{sat,max}; \bar{T}_m \leq T_{v2} \\ V_{sat,max} \cdot \left(1 - \frac{T_m - T_{v2}}{T_{v3} - T_{v2}}\right); T_{v2} < \bar{T}_m < T_{v3} \\ 0; \bar{T}_m \geq T_{v3} \end{cases} \quad (7)$$

where N represents five coldest months, $V_{sat,m}$ the saturated vernalization requirement of month m in days, $V_{sat,max}$ the maximum saturated vernalization requirement per month possible, T_m the long-term-average monthly temperature (originally the long-term-average monthly temperature from preceding year is used), and T_{v2} and T_{v3} ($^{\circ}\text{C}$) the minimum and maximum temperatures for optimal vernalization (Table S2). The value of $V_{sat,max}$ is set to 14 days/month, based on a study by Baloch et al. (2003), who indicated that winter wheat cultivars with high vernalization requirements need at least 70 days of optimum vernalizing temperatures (in this study we assume equal distribution over 5 months). However, this assumption may disqualify late-maturity varieties with large vernalization requirements in temperate regions. Table S2 in supplemental material provides the values of all static parameters of the phenological model that are not subjected to calibration.

Vernalization requirement parameters (V_{sat} and V_b) are calculated for each climatic grid cell and then averaged over the agro-environmental zones, resulting in unique parameter values for each zone (Fig. S2).

2.6. Calibration of the phenological model

Two phenological parameters, $TSUM1$ and $TSUM2$, respectively, are calibrated by pooling together the observations from the agro-phenological database for each agro-environmental zone (Fig. 2). Only experimental data meeting the following requirements are used for calibration: 1) they should have an observed start of the season (sowing or emergence) date, and 2) experiments should have an observed anthesis and/or end-of-season (physiological maturity or harvest) date. The calibration of $TSUM1$ and $TSUM2$ parameters is performed by minimizing the objective function for each zone:

$$f(x) = \sqrt{f(a) + f(m) + f(h)} \quad (8)$$

where $f(a)$, $f(m)$ and $f(h)$ represent the Weighted Mean Squared Difference (WMSD) between observed and simulated anthesis, maturity

and harvest dates, respectively. It should be emphasized that the real observed sowing and/or emergence dates are used when calibrating TSUMs. $f(x)$ is the Weighted Root Mean Square Difference (WRMSD). The minimization of $f(x)$ is performed using two nonlinear optimization algorithms; the global Direct-L algorithm, followed by a local COBYLA algorithm (Steven and Johnson, 2017; Gablonsky and Kelley, 2001; Powell, 1994).

The WMSD for anthesis $f(a)$ is calculated as:

$$f(a) = \frac{\sum_{i=1}^n w_i \cdot (a_{i,o} - a_{i,s})^2}{\sum_{i=1}^n w_i} \quad (9)$$

where w_i is the weight quantifying the contribution of each phenological observation, $a_{i,o}$ the observed anthesis date (field observations or regional calendars), $a_{i,s}$ the simulated anthesis date and n is the number of observations. $f(m)$ and $f(h)$ are calculated similarly using maturity and harvest dates, respectively.

Weights w_i are determined separately for experimental location based data and regional calendars. This is a necessary step to avoid situations with many experiments falling within the single grid cell to have too strong impact on the calibrated TSUM values for the entire agro-environmental zone. If more experiments fall within the same grid cell, their contribution is weighted proportionally :

$$w_i = 1/n_j \quad (10)$$

with n_j representing the number of experimental observations within the j -th grid cell. Regional calendars are used when experimental data are deemed insufficient. This occurs when within a zone a) there are no experiments defining the vegetative phase, b) there are no experiments defining the full season, and c) the total number of experimental observations is < 15 (Fig. 2). A weight is assigned to each of regional calendars, which avoids outnumbering available experimental observations that are only valid for a single grid cell and a single year. The weight is calculated in a way that each calendar as a whole (extending towards several grid cells) has a same weight as a single experimental observation. An example of calibration procedure for selected agro-environmental zone is illustrated in Section S1 of the Supplemental material.

The evaluation of calibration is performed in a two-step procedure. In the first step, the validation is performed on independent dataset for agro-environmental zones with sufficient experimental data available to split the data sample in two parts, subsequently used for calibration and validation. In the second step, full calibration of phenological model is performed (i.e. taking all experimental data and regional calendars). TSUMs derived with full calibration are then used to perform crop yield simulations, which are subsequently compared to recorded national yields (see Section 2.8).

2.7. Soil moisture initialization

Initial soil moisture in the original simulations of WOFOST within MCYFS is set to field capacity. For the purpose of this study, the soil moisture content at autumn sowing dates is calculated by starting the soil water balance three months prior to the grid specific sowing date, assuming a bare soil. This early water balance is initialized using the grid-specific long-term average end-of-season soil moisture content simulated by WOFOST.

2.8. Comparison of simulated with recorded crop yields

Simulated water-limited and potential weights of storage organs are used as proxies to compare with recorded national winter wheat yields. For comparison of simulated with observed crop yields, all available experimental data and regional calendars are used to perform the calibration of the phenological model (full calibration). Resulting zone-specific TSUMs are then used to run the crop model historical

simulations to derive country-specific crop yield values. Historical yield statistics (Eurostat Agriculture, 2016) are used as observed crop yield time series. Besides the inter-annual climatic variability, yield statistics reflect also variations in agro-management and socio-economic factors, not simulated by the process based crop model. Both, modelled yields and recorded yield time series, are therefore de-trended prior to their comparison (see e.g. Ceglar et al., 2016, for further details on detrending). Then, Taylor diagrams (Taylor, 2001) are used to visualize the similarity between de-trended anomalies of modelled and recorded crop yield time series in terms of their correlation, root mean square error and standard deviation.

Additionally, spatial simulations of various simulated crop model indicators, such as LAI, biomass weight and harvest index, are assessed for biophysical plausibility (i.e. whether the simulated values fall within the range suggested by the literature review and/or expert opinion).

3. Results

3.1. Derived sowing dates

Fig. 3 shows the derived long-term average sowing dates. The spatial pattern generally indicates a latitudinal gradient, with earlier sowing dates simulated towards the North, corresponding to increased vernalization requirements of winter wheat varieties (and ensuring an adequate crop stand before the winter frost). The comparison of simulated and observed sowing dates from the agro-phenological database reveals a good agreement at most of the observational locations with average absolute difference of 12 days (Table 1).

Only 0.5% of the observations is affected by an absolute difference larger than 90 days (Fig. 3b,c). The corresponding observed sowing dates are identified in mid winter or early spring in Spain, Italy, Germany and Turkey. These sowing dates most probably indicate spring wheat varieties, and are therefore not further considered.

3.2. Simulation of phenological development

The highest number of observational data to calibrate TSUMs is available in central, western and northern Europe (Fig. S1). Experimental datasets are scarce in south-eastern and eastern Europe, where the calibration mainly relies on regional calendars. Due to quite heterogeneous spatial density of observed phenological data coming from field experiments, the evaluation of calibration is performed in a two-step procedure.

In the first step, the validation is performed on independent dataset for agro-environmental zones with sufficient experimental data available. Only zones with at least 30 pairs of sowing-anthesis dates and 30 pairs of anthesis-maturity dates available from field experiments are selected. Zones with regional calendars are not considered in this step. Overall, these criteria are met in 16 zones across Europe (Fig. 4), located mainly in central parts of Europe, northern France, Spain, western Black Sea areas and Turkey. Half of experimental observations within each of the selected zones are randomly chosen for calibration of the phenological model. The calibrated phenological model is then used to simulate anthesis and maturity dates for the rest of the locations, which are excluded from the calibration procedure (Fig. 4). WRMSE is calculated for both datasets separately: calibration ($WRMSE_{cal}$) and validation ($WRMSE_{val}$).

As expected, $WRMSE_{val}$ is generally higher or comparable to the $WRMSE_{cal}$. The calibrated phenological model is able to reproduce the anthesis and maturity dates more reliably in zones located in central Europe, northern France and western Black sea regions, with $WRMSE_{val}$ and $WRMSE_{cal}$ mainly below 12 days. Exceptions are zones 1976 and 1984 with $WRMSE_{val}$ exceeding two weeks, roughly doubling the values of $WRMSE_{cal}$. Errors higher than 14 days prevail in the Iberian Peninsula and major part of Turkey. Although no clear conclusions can

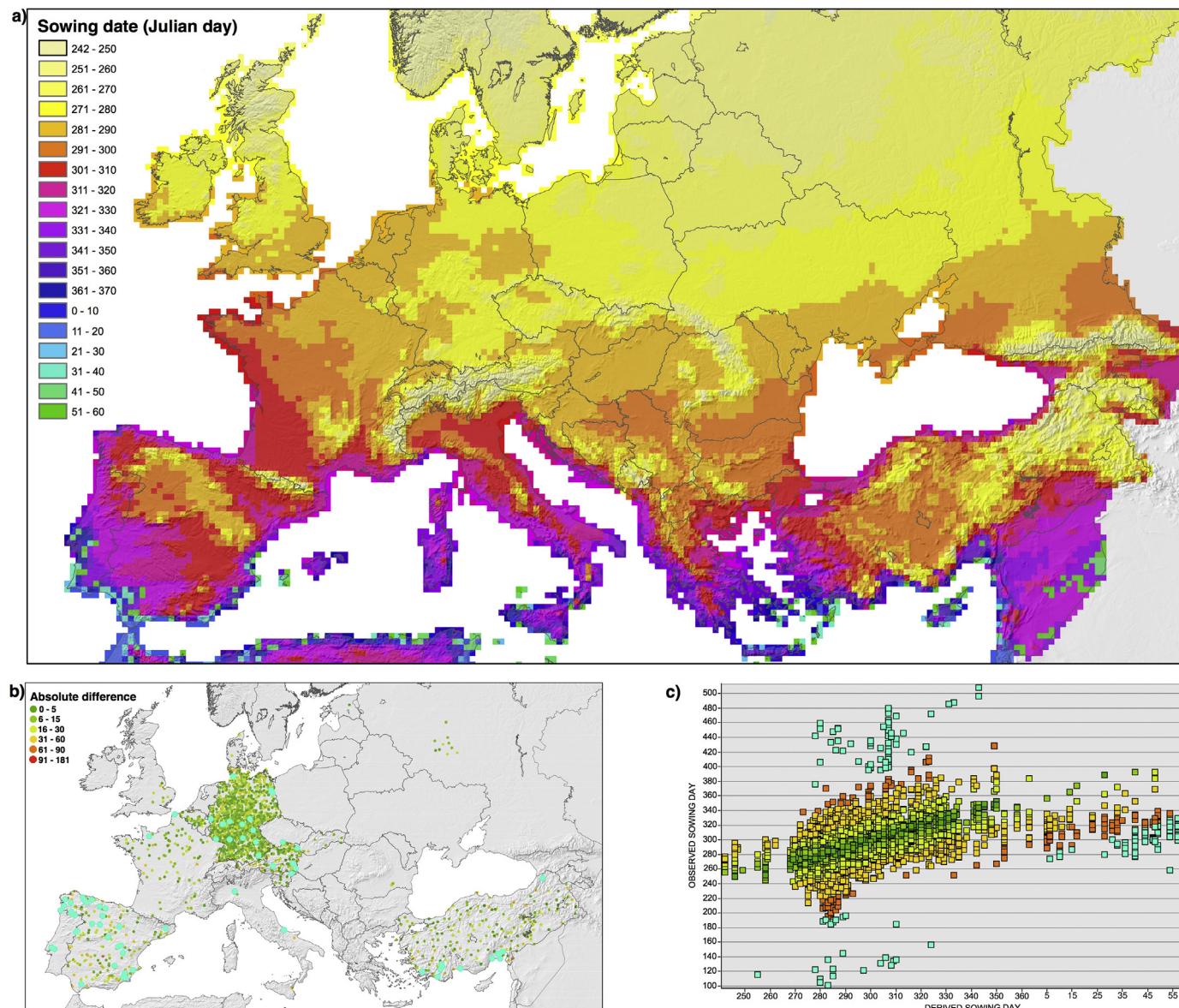


Fig. 3. a) Derived sowing dates in Julian days. b) Average absolute difference between simulated and observed sowing dates (in number of days). Cyan-colored points denote locations, where differences are higher than 90 days, and are excluded from further analyses. c) Scatterplot of observed vs simulated sowing dates. The colour scheme is the same as in b). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

be drawn on the impact of sample sizes on the differences between $WRMSE_{val}$ and $WRMSE_{cal}$, there is a tendency towards larger differences in several zones where overall the number of experimental observations is low (e.g. zones 1976, 1984, 2075 and 2091). Calibration in these zones seems to be more sensitive to the selection of observations; therefore, re-iterated random sub-sampling could provide more robust results.

To take the full advantage of regional calendars, calibration is performed using the entire observational dataset in the second step (full calibration). After the initial zone-specific calibration of thermal requirements, a correction of TSUMs is performed for the zones where the ratio of calibrated TSUM1 and TSUM2 differs substantially from the ones estimated for the neighboring zones, or calibration is not possible due to missing data (Fig. S3). Spurious TSUM1 values could be caused by few or no anthesis observations available within several zones, together with a large number of maturity or harvest dates. In those cases, the overall thermal requirements ($TSUM1 + TSUM2$) are assumed to be more suitable for accurate calibration, while the ratio $TSUM1/TSUM2$ from the neighboring zones (where more anthesis dates are available)

can be used to distribute overall thermal requirements among TSUM1 and TSUM2. Adjusted zones are shown on Fig. S3.

Fig. 5 shows the final estimated TSUM1 and TSUM2. The thermal requirements to reach anthesis (TSUM1) are decreasing with increasing latitude, and reach the highest values in southern Europe. The spatial distribution of thermal requirements during reproductive period (TSUM2) exhibits stronger longitudinal gradient, with lower values observed towards eastern Europe, while values around 900 GDD can be observed in the western half of Europe. Lower thermal requirements in eastern Europe with continental climate correspond to a variety selection, that would avoid drought and heat stress during the reproductive period.

$WRMSE$ under full calibration ($WRMSE_{full}$) is calculated using the entire available phenological dataset. Large error, corresponding to $WRMSE_{full}$ values around 19 days or higher, can be observed in the Iberian Peninsula, North African coast and eastern Turkey for both, simulated anthesis and maturity dates. This corresponds well to the findings based on the independent validation. The differences between $WRMSE_{full}$ and $WRMSE_{cal}$ from the first step are higher in the zones with

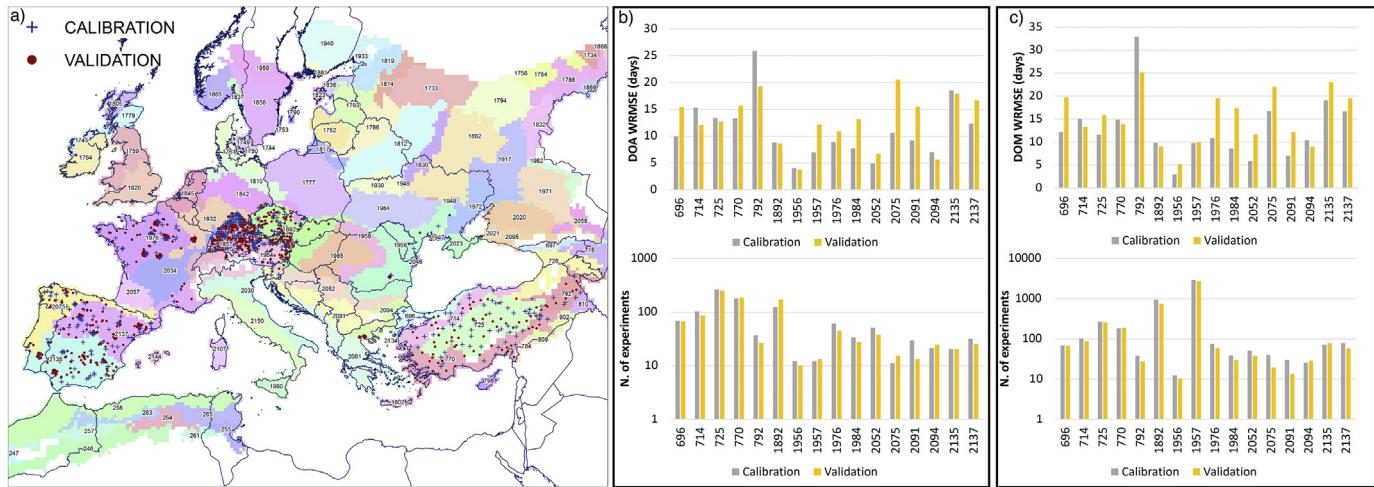


Fig. 4. a) Selected agro-environmental zones for independent validation of phenological simulations of anthesis and maturity dates. The data points within each of the selected zones show the locations of experimental observations that are used for calibration and independent validation of the phenological model. b) WRMSE (in days) for day of anthesis (DOA), calculated from calibration and validation datasets. Additionally, a number of experiments is shown that are used to perform calibration and validation in each of the selected zones. Numbers below the x-axis represent the agro-environmental zone identifier visible on the map in a). c) Same as b), but for maturity dates.

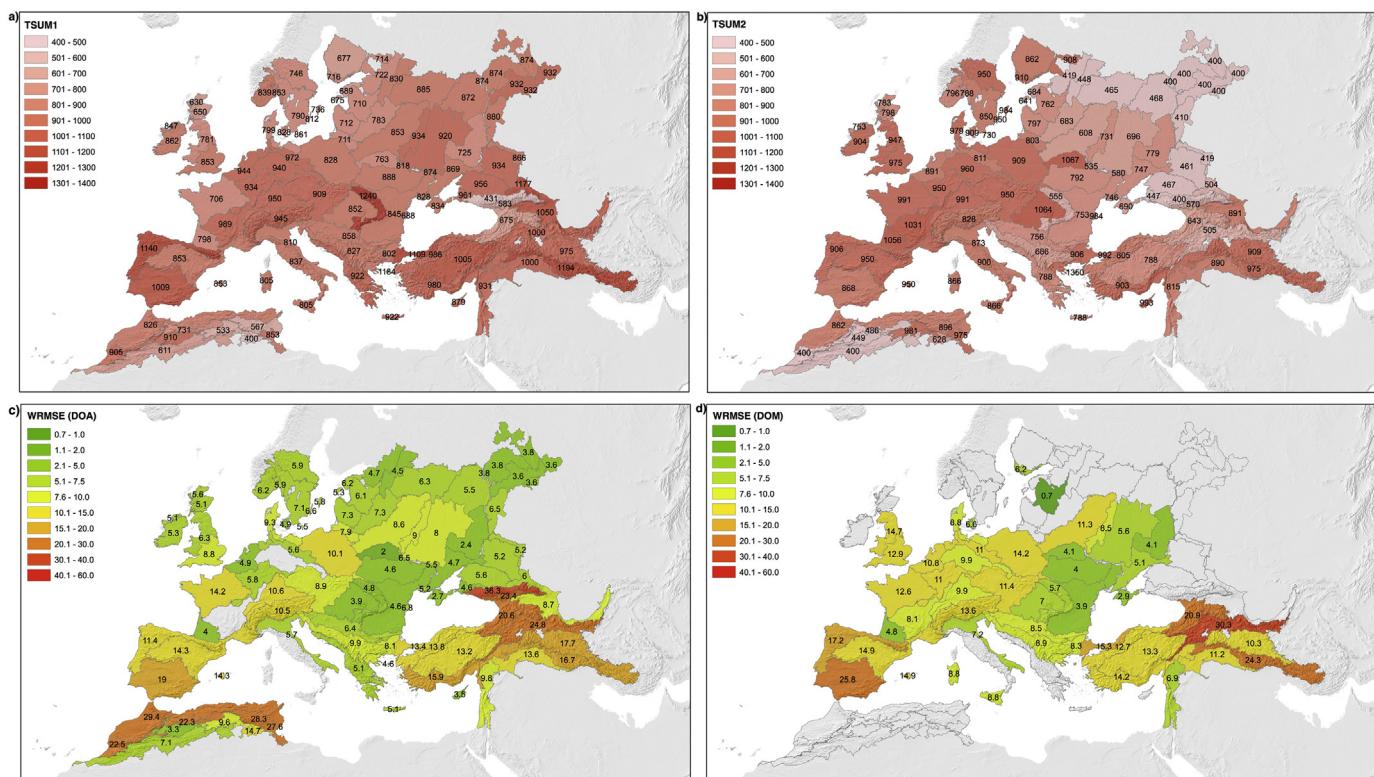


Fig. 5. a) Calibrated TSUM1 for winter wheat in different agro-environmental zones in Europe, reflecting the thermal time between sowing and anthesis. b) Calibrated TSUM2 for winter wheat in different agro-environmental zones in Europe, reflecting the thermal time between anthesis and maturity. c) WRMSE_{full} (in days) for the day of anthesis after calibration. d) WRMSE_{full} for the day of maturity after calibration.

less experimental observations available for calibration (Fig. 3b,c).

To further explore the implications of phenological model calibration, the simulations of anthesis and maturity dates are compared with simulated dates of the original MCYFS system for winter wheat in selected agro-environmental zones (Fig. 4a). Fig. 6 shows scatterplots between simulated and observed dates of anthesis and maturity. The original MCYFS system tends to simulate too early occurrence of anthesis dates, with bias nearly around 20 days averaged over all the analyzed zones (Fig. 4a), whereas the maturity dates are captured

better (Fig. 6). The phenological model and its calibration here proposed lead to substantially lower bias in simulated anthesis dates for the selected agro-environmental zones. As for simulated maturity dates, our results indicate lower spread around 1:1 line than in the case of the original MCYFS system.

3.3. Assessment of plausibility of simulated crop indicators

To further verify the new phenological model calibration, several

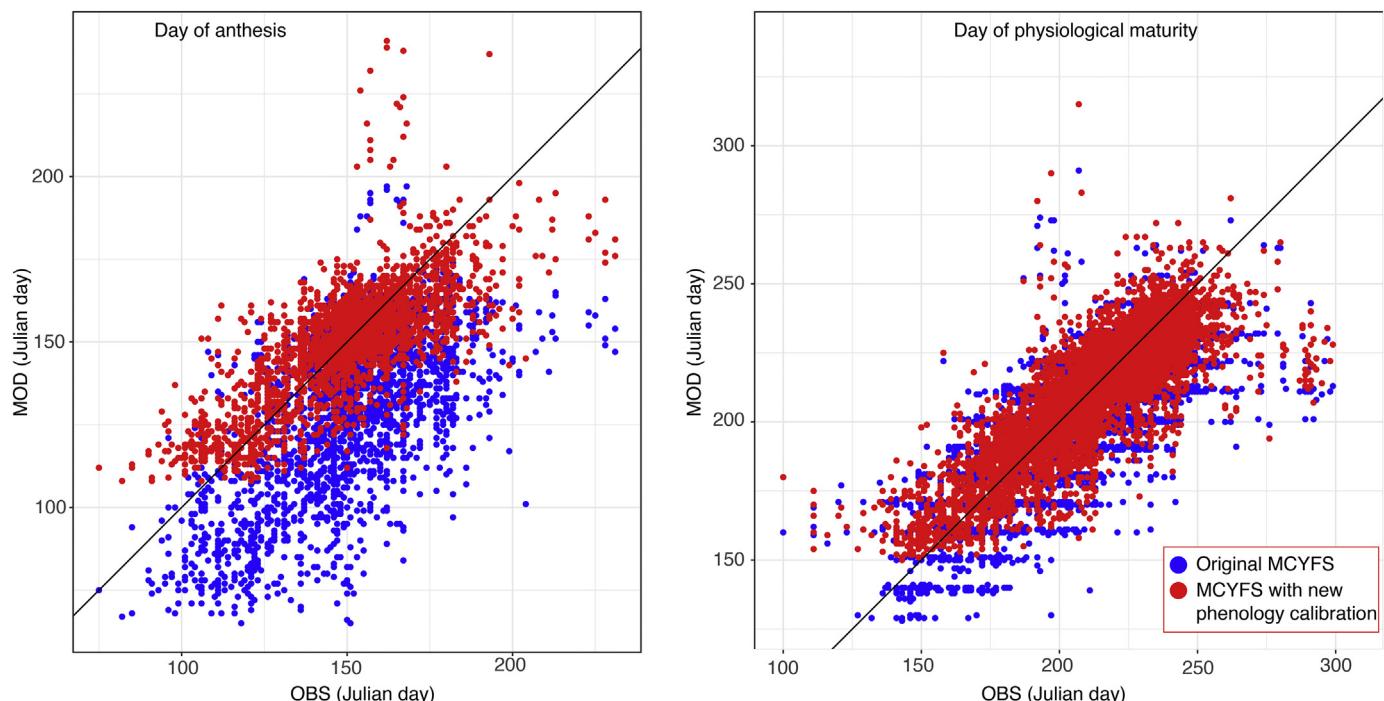


Fig. 6. Scatterplots of simulated (MOD) vs observed anthesis dates (a) and maturity dates (b) for winter wheat in agro-environmental zones, selected for independent validation (Fig. 4a). Colors correspond to the original MCYFS simulations (blue) and new MCYFS simulations with improved and re-calibrated phenological model (red). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

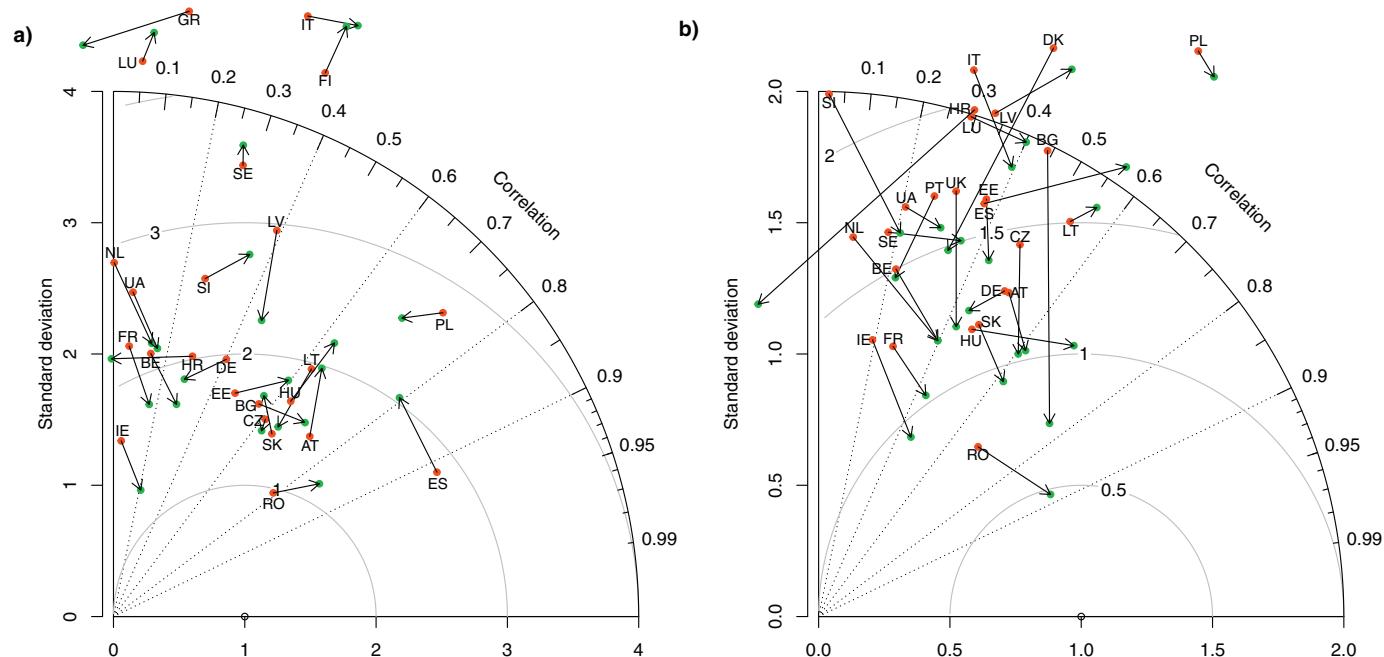


Fig. 7. Taylor diagrams for simulated country level water limited (a) and potential (b) yields. The radial distance from the origin represents the standard deviation of simulated crop model variable, normalized by standard deviation of de-trended national statistical yields. The black dot at unit distance and 0° from the origin would indicate perfect agreement. The distance to that circle represents a normalized root mean square error. For each country, two points are shown on the diagram: starting red point with the name of the country indicates the performance of the original MCYFS system, while the end green point shown by arrow indicates the performance of MCYFS with the improved calibrated phenological model and realistic initialization of soil moisture at autumn sowing dates. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

simulated model variables are assessed for physical plausibility. Under optimal growing conditions, the expected values for *BIOM* should range between 16 and 22 t/ha, maximum *LAI* between 4 and 7, and harvest index (*HI*) between 0.4 and 0.6 (Boogaard et al., 2013; Boons-Prins et al., 1993; Groot and Verberne, 1991). Moreover, the results from the

new phenology calibration is compared to the original MCYFS simulations, starting simulations with the 1st of January ([Micale and Genovese, 2004](#)).

The spatial assessment of simulated values reveals that maximum *LAI* generally falls within the physically plausible range. Without

having calibrated the leaf area dynamics the distribution of maximum *LAI* generally improves with respect to the simulated *LAI* in the original MCYFS system, starting simulations on the 1st of January (Fig. S4). The most relevant improvements can be observed in the central part of Spain, central and south-eastern Europe and Turkey. In the original system, the simulated period of leaf formation is too short, resulting in low simulated maximum *LAI* and light interception (Boogaard et al., 2013). Moreover, the sharp transition in maximum *LAI* (potential and water limited) between south-eastern and central Europe diminished after calibration (Fig. S4). Nevertheless, *LAI* at the end of the season is still too high in eastern Europe, indicating an underestimation of leaf senescence. The spatial distribution of *HI* generally improves with a more narrow distribution under potential growing conditions (Fig. S5). In water limited conditions, extremely low *HI* values occur in several regions of the Mediterranean, mainly where initial soil moisture at sowing is very low. Elsewhere, *HI* values mainly ranges between 0.4 and 0.6 (Fig. S5).

Similar to the original system, the highest *BIOM* ($> 20 \text{ t/ha}$) occurs in coastal regions (e.g. north-western France, the British Isles and the Benelux countries) as a result of a long growing period with mild weather conditions throughout the growing season. Lower values ($15\text{--}20 \text{ t/ha}$) are simulated in central Europe and large areas of eastern and south-eastern Europe (Fig. S6). In eastern Europe, the simulated values are generally lower than the ones of the original system, whereas higher *BIOM* values are simulated in south-eastern Europe. This is mainly caused by lower (higher) *LAI* values, simulated using improved calibration of phenology in eastern (south-eastern) Europe.

3.4. Simulation of crop yields

Fig. 7 shows the Taylor diagrams derived for aggregated water limited and potential yields (*WLO* and *PSO*) in each country. Compared to the original MCYFS system, the simulated *PSO* at the end of the crop cycle better matches the recorded yields in the majority of European countries. As for the *WLO*, the relationship with the recorded yield residuals (after de-trending) slightly improves in Ireland, the Benelux countries, the Baltic countries, Finland, France, the Czech Republic, Slovenia, Bulgaria, Romania and Ukraine; while, the model performance slightly decreases in Spain, Austria, Slovakia, Germany, Portugal and Greece.

3.5. Simulated intra-seasonal crop indicators

Compared to the wheat simulations in the original MCYFS system, a substantial improvement in correlation between crop yield indicators and national crop yields can be observed for the majority of countries, with the most pronounced improvements in south-eastern and central Europe (Fig. 8). Stronger correlations are induced by a shift in the anthesis time, which generally occurs up to 3 dekads earlier with respect to the original MCYFS system, and an earlier senescence. Thus, the main improvement can be observed in the simulation of the leaf area dynamics during the grain filling period, resulting in positive relationship between the potential or water limited *LAI* and the recorded yield residuals in many European countries (especially for dekads between 15 and 23, Fig. 8).

Fig. 9a shows an example of the improved correlations between various crop model indicators during the crop growing season and detrended recorded statistical yields for Romania (considering the period 1996–2016). A good crop model performance in Romania is encouraging, as most of phenological data for the calibration come from regional calendars. It should be emphasized that the analyzed period belongs to the post-communist period in Romania, marked by fundamental transformations in agriculture with poor development of agricultural services, resulting in the intensification of adverse effects of extreme climatic events on crop production (Sima et al., 2015). The MCYFS system with improved crop phenology model is capable of

capturing this dependency in the absence of intensive agro-management practices.

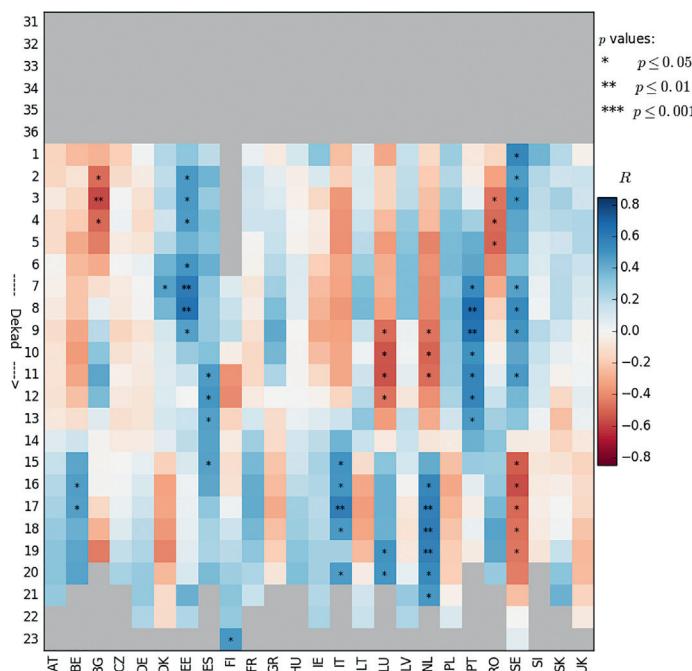
Even though the simulated weight of storage organs at the end of the crop growing season does not correlate well with observed national yields in France in the period 1995–2016, significant correlation between several intra-seasonal crop model indicators and recorded crop yields is revealed using the new phenology calibration with autumn starting date (Fig. 9b). It generally results in a stronger correlation of simulated crop model indicators with recorded yield, especially *LAI* during the senescence phase. Soil moisture plays the most relevant role around the heading development stage, with significant positive correlations. The importance of early season conditions must be also highlighted (Fig. 9b), and the total water requirement in late autumn and early winter appears to be negatively linked to the final recorded yield. Negative correlation between *TWR* and crop yield anomalies suggests that high atmospheric evapotranspirative demand during early winter, which is increasing with amount of intercepted solar radiation, water vapour pressure deficit and wind speed, negatively affects verinalization and phenological development. Additionally, a significant negative correlation can be found between yield anomalies and soil moisture in early winter; negative impact of soil moisture excess during early winter might be related to poor plant establishment and increased disease pressure (Blake et al., 2003). Similar results have been already highlighted by Ceglar et al. (2016) for the northern half of France. Both warm and dry conditions, as well as over-wet conditions, in early season have negative influence on wheat crop establishment and final crop yields. The large diversity of environmental conditions in the French wheat production areas requires further analysis of simulated crop model indicators at sub-national level.

Even though, the WOFOST model does not capture the impact of several extreme weather events on crop growth and development (e.g. frost kill, waterlogging, pests, diseases, etc.), accurate phenological simulations can help to identify and accommodate such events in crop monitoring and forecasting activities. Let us consider the example of the Czech Republic. Fig. 10 shows the observed crop yield anomalies and modelled winter wheat yields (using improved and calibrated phenological model) aggregated at the national level. Despite a modest agreement between simulated and recorded yield series (correlation around 0.6), there are years when the crop model clearly underestimates the extent of negative yield anomalies. Such a case occurred in 2012, when frost kill (not simulated by the crop model) severely affected wheat yields (USDA, 2012). On the other hand, the yield anomalies of 2007 and 2015 represent serious *false alarms*. Weather indicators in both years point to heat stress that affected wheat crops during the grain filling period; yet the recorded yield anomalies at the end of the season are close to 0. According to the model simulations, the heat wave occurred during the grain filling period, whereas in reality (according to observations in agro-phenological database) the crop was already ripening. This case study clearly illustrates the importance of having as accurate phenological simulations as possible in order to provide reliable background for wheat yield monitoring and forecasting.

4. Discussion

The importance of accurately capturing phenological development has often been highlighted in climate change impact studies (e.g. Müller et al., 2017; Boogaard et al., 2013; Balkovič et al., 2013). The lack of observational phenological data and the spatial unit of simulation (in our case 25 km grid cells), representing aggregates of potentially heterogeneous fields, often pose a major challenge to gridded crop modeling (Müller et al., 2017). Our study has (at least partially) bridged this gap in Europe by combining available phenological observations with regional calendars for calibration of winter wheat phenology in the MCYFS. Even though regional calendars bring no information on inter-annual variability of phenological dates, they do serve as valuable set of

a) Operational MCYFS system, with starting date on the 1st of January



b) Simulation starting with autumn sowing date and improved phenology model and calibration

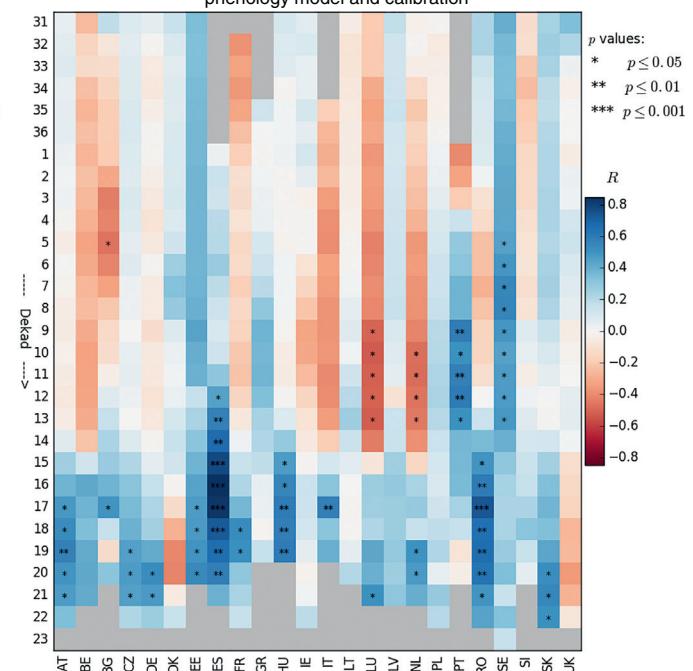


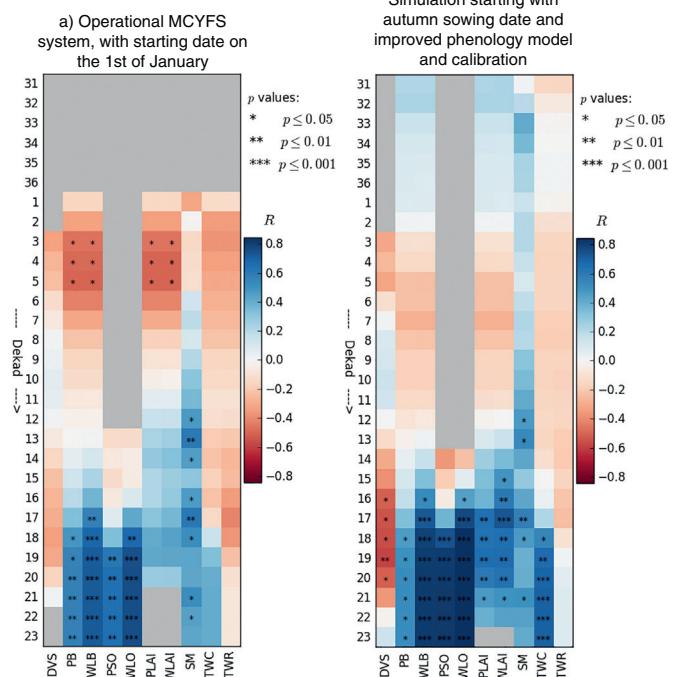
Fig. 8. Correlations between simulated water limited leaf area index and observed national crop yields for individual European countries (for the period between 1996 and 2016), as simulated by the original MCYFS system (a) and the MCYFS with the improved calibrated phenological model and realistic initialization of soil moisture at autumn sowing dates (b).

information over regions with limited or no observed field experimental data. Using less precise but more abundant datasets for phenological model calibration have been shown to improve the prediction skill (e.g. Montesino-San Martin et al., 2018). This is also confirmed in our study,

as additional evaluation on recorded national crop yields shows that over south-eastern Europe, where mainly regional calendars are used for calibration, the crop yield simulations improve substantially.

Introducing autumn sowing date, vernalization and calibration of

Romania



France

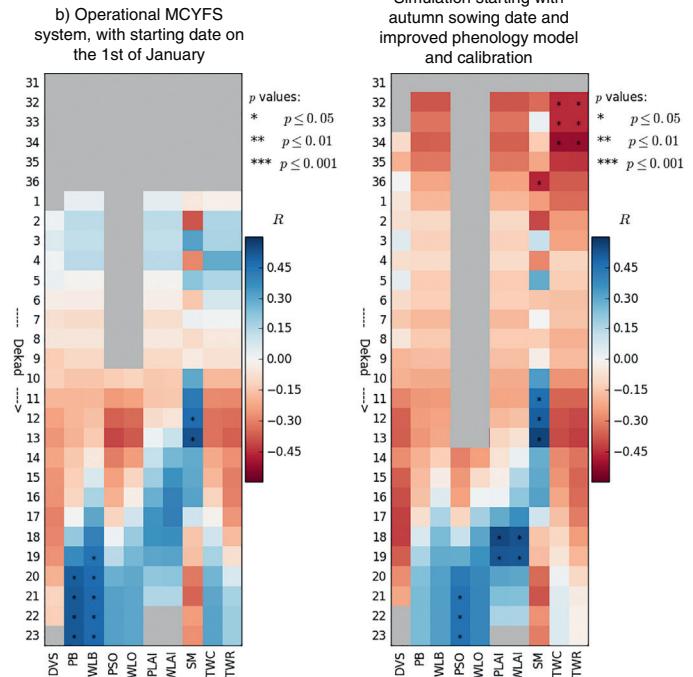


Fig. 9. Correlations between simulated dekadal crop model indicators and de-trended national crop yields for Romania (a) and France (b), calculated over the period between 1996 and 2016. The crop model indicators represent: PB – potential aboveground biomass, PSO – potential storage organs, WLB – water limited above-ground biomass, WLO – water limited storage organs, PLA – potential leaf area index, WLAI – water limited leaf area index, DVS – development stage, SM – soil moisture, TWC – total crop water requirement, TWR – total crop water consumption.

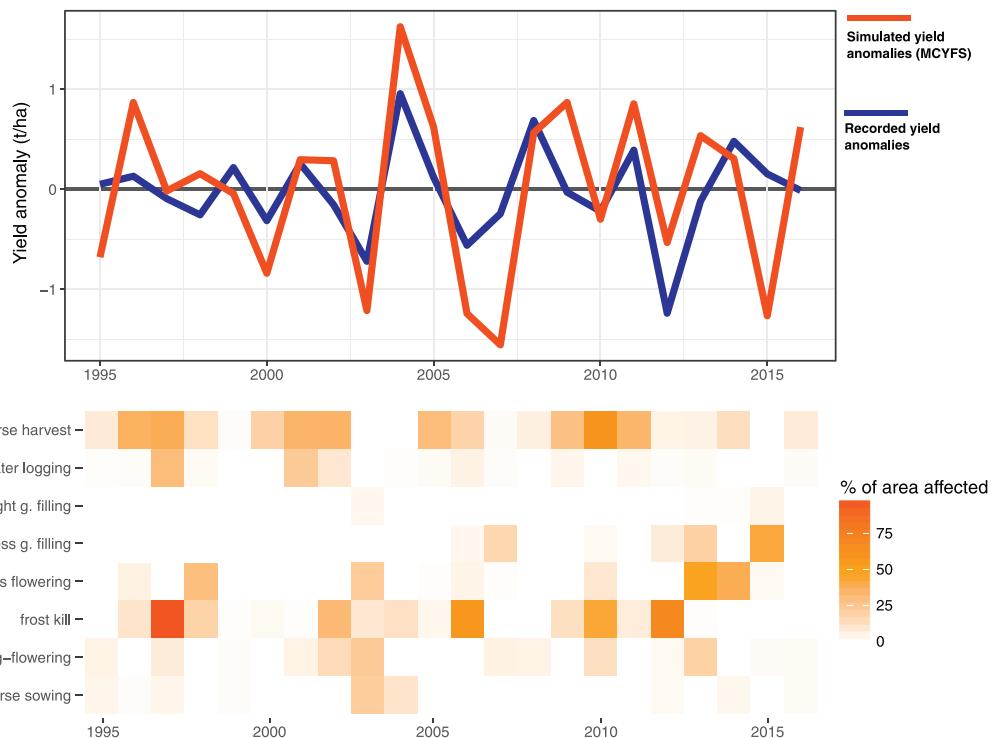


Fig. 10. a) Recorded winter wheat yield anomalies (blue) and simulated yield anomalies (red line) using the MCYFS. b) Percentage of winter wheat area in the Czech Republic affected by adverse weather events during different crop growth stages (Trnka et al., 2014, and Table S3 in supplemental material for additional information on the definition of adverse weather events). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

phenology lead to better performance in the anthesis simulations, especially in central and western Europe. Large errors remain in the Iberian Peninsula and several (mainly Mediterranean) regions in Turkey. These regions could predominantly grow spring wheat varieties with no vernalization requirements, which is not in agreement with our assumptions. However, as the agro-phenological database does not contain full description of varieties related to each recorded experiment, it is not possible to draw clear conclusions with respect to the prevailing reasons.

Changing varieties, as part of the climate adaptation process, represents a big challenge for continental crop growth monitoring systems, such as the MCYFS. As an example, up to the beginning of the 21st century, photoperiod insensitivity has been introduced to most wheat cultivars grown at latitudes below 48°N (Rajaram and van Ginkel, 2001), since it is beneficial for regions with high summer temperatures where farmers would grow crops relatively early in order to avoid heat/drought stress during sensitive stages around anthesis, and to increase the yield potential (Langer et al., 2014). MCYFS system assumes that winter wheat varieties do not change over time; while this assumption is generally valid over shorter time scales (e.g. decadal), the adaptation to changing climate encourages farmers to faster respond with measures such as variety selection to ensure yield stability and performance. The selection of varieties depends on their own experience, exchanges with colleagues and recommendations from the various agricultural and plant breeding organizations (Macholdt and Honermeier, 2017). As an example, Fig. 11 shows the vernalization requirements of French varieties between 1993 and 2016, the main substantial change being the decrease of varieties having the highest vernalization requirements. The simulated vernalization requirements in MCYFS represent the most common varieties in recent years reasonably well.

While the agro-phenological database, used in this study, provides an important source of data for calibration, it should be emphasized that further efforts are necessary to harmonize the available data,

aiming towards a regularly updated database with recent European wheat varieties. This is especially relevant for the improvement of wheat phenology simulations in intensive producing areas, such as the northern half of France. Even though a single wheat variety is assumed to prevail in this area, a wide range of varieties could coexist in reality. Besides prevailing winter climate conditions, the farm type and the main crop rotations can also affect the selection of varieties with different vernalization requirements (Agreste, 2017); varieties with low requirements are used in regions where farms are oriented towards breeding/livestock. A relatively small number of observations available can bias the calibrated thermal requirements due to undersampling the population of prevailing wheat varieties within the region.

The correlation between crop yield simulations and national yield statistics improve with respect to the original MCYFS simulations in Spain, south-eastern Europe and several central European countries (Fig. 7). However, improvements are more substantial for potential than for water-limited simulations. This indicates that the calibration of phenology brings improvement in overall model performance, but the water related aspects of the simulations need further investigation. Moreover, poor crop model performance in terms of regional crop yield simulation remains over western Europe, the UK and Ireland. In these regions, potential crop yield simulations correlate even slightly better with recorded crop yields than the water limited simulations. These results could possibly be attributed to the impacts of water excess rather than drought, which are known to be of importance in those regions (Zampieri et al., 2017; Ceglar et al., 2016), but are not accounted for in the current implementation of the WOFOST model. Lower performance under wet conditions could be related to various processes, such as the impact of water logging on crop growth during the sensitive stages that are not included in the process based crop model simulation. Moreover, additional factors related to wet conditions might reduce the yields, such as occurrence of pests and diseases and difficulties to enter the field for crop treatment.

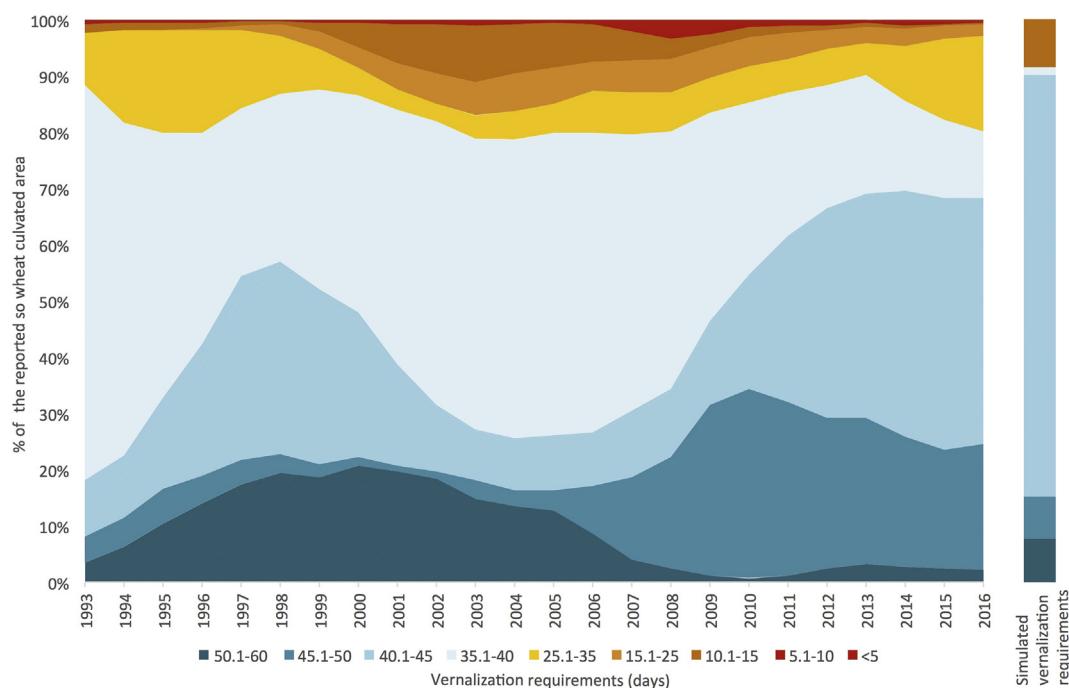


Fig. 11. Evolution of vernalization requirements of winter wheat varieties in France between 1993 and 2016. The data on variety vernalization requirements have been obtained from (Arvalis, 2017) and Variety and Seed Study and Control Group (GEVES, 2017). The column on the right represents the distribution of simulated vernalization requirements in MCYFS.

5. Conclusions

- The comparison of simulated and observed sowing dates from agro-phenological database reveals good agreement at most of the observational locations.
- The vernalization requirements introduced into the WOFOST phenology of the MCYFS show a plausible pattern over Europe and are in good agreement with observed ones in France, especially for prevailing recent varieties.
- A spatial calibration of the phenology of the WOFOST model is performed across Europe, with the winter wheat varieties assumed to be adapted to agro-environmental zones. Calibrated phenology results in substantial improvement in simulated dates of anthesis with respect to the original MCYFS simulations. The calibrated thermal requirements to reach anthesis decrease with increasing latitude. Spatial distribution of thermal requirements from anthesis to maturity exhibit stronger longitudinal gradient, with lower values observed towards eastern Europe. Validation of the calibrated phenological model still revealed large errors over the Mediterranean part of Europe, especially when it comes to simulation of anthesis dates.
- The simulations of the WOFOST crop model with new phenology calibration and realistic soil moisture initialization result generally in a slightly better agreement with recorded national winter wheat yields with respect to the original MCYFS simulations. Spatial assessment of crop model simulations results in physically plausible distribution of different crop indicators across Europe.

Important challenges remain to be addressed in terms of further improvements of the crop model, as the calibration of phenology alone is not sufficient to realistically reproduce inter-annual variation of recorded yields in several regions of Europe. Future efforts in expanding the agro-phenological database need to focus on regular updating with recent observations and on metadata, preferably including information on variety selection. Various processes related to the impacts of water excess crop growth need to be introduced in the crop model simulation. Such processes seem to be of particular relevance in western and

northern Europe.

Acknowledgments

We would like to acknowledge DG AGRI of the European Commission for financing this study under the framework service contract 390604 for supporting MARS crop yield forecasting systems.

Appendix A. Supplemental data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.agys.2018.05.002>.

References

- AETS, 2013. Study on Providing Agro-Pedological and Climatological Knowledge Base of the Main Crops in Turkey and Ukraine as Well as Provision of Agricultural Statistics. European Commission, https://www.aets-consultants.com/en/projects_view/173. Accessed on: 10. 9. 2016.
- Agreste, 2017. Les rotations des cultures. URL: http://agreste.agriculture.gouv.fr/IMG/pdf/dossier21_rotation.pdf. Accessed on: 1. 9. 2017.
- Eurostat – Agriculture, Forestry and Fisheries Database, 2016. URL: <http://ec.europa.eu/eurostat/data/database>, accessed on 30.01.16.
- Angulo, C., Rötter, R., Lock, R., Enders, A., Fronzek, S., Ewert, F., 2013. Implication of crop model calibration strategies for assessing regional impacts of climate change in Europe. *Agric. For. Meteorol.* 170, 32–46.
- Arvalis, 2017. LES FICHES VARIETES : blé tendre, blé dur, orge et pomme de terre. URL: http://www.fiches.arvalis-infos.fr/liste_fiches.php?fiche=var&type=512. Accessed on: 1. 9. 2017.
- Balković, J., van der Velde, M., Schmid, E., Skalský, R., Khabarov, N., Obersteiner, M., Stürmer, B., Xiong, W., 2013. Pan-European crop modelling with EPIC: implementation, up-scaling and regional crop yield validation. *Agric. Syst.* 120, 61–75.
- Baloch, D.M., Karow, R.S., Marx, E., Kling, J.G., Witt, M.D., 2003. Vernalization studies with Pacific northwest wheat. *Agron. J.* 95, 1201–1208.
- Barnabás, B., Jäger, K., Fehér, A., 2008. The effect of drought and heat stress on reproductive processes in cereals. *Plant Cell Environ.* 31, 11–38.
- Baruth, B., Kucera, L., 2007. Crop masking – needs for the MARS crop yield forecasting system. International archives of the photogrammetry, Remote Sensing Spatial Inform. Sci. 36, 3–6.
- Blake, J.J., Spink, J.H., Dyer, C., 2003. Factors affecting cereal establishment and its prediction. *HGCA Res. Rev.* 51, 51.
- Blanco, M., Ramos, F., Van Doorslaer, B., Martinez, P., Fumagalli, D., Ceglar, A., Fernandez, F.J., 2017. Climate change impacts on EU agriculture: a regionalized

- perspective taking into account market-driven adjustments. *Agric. Syst.* 156, 52–66.
- Boogaard, H.L., Van Diepen, C.A., Rötter, R.P., Cabrera, J.M.C.A., Van Laar, H.H., 1998. User's Guide for the WOFOST 7.1 Crop Growth Simulation Model and WOFOST Control Center 1.5. Technical Document 52. Winand Staring Centre, Wageningen, the Netherlands, pp. 144.
- Boogaard, H., Wolf, J., Supit, I., Niemeyer, S., van Ittersum, M., 2013. A regional implementation of WOFOST for calculating yield gaps of autumn-sown wheat across the European Union. *Field Crop Res.* 143, 130–142.
- Boons-Prins, E.R., De Koning, G.H.J., Van Diepen, C.A., Pennin de Vries F.W.T., 1993. Crop specific simulation parameters for yield forecasting across the European Community. *Simulation Reports CABO-TT No. 32, CABO-DLO*, Wageningen, the Netherlands, 43.
- Ceglar, A., Kajfez-Bogataj, L., 2012. Simulation of maize yield in current and changed climatic conditions: addressing modelling uncertainties and the importance of bias correction in climate model simulations. *Eur. J. Agron.* 37, 83–95.
- Ceglar, A., Toreti, A., Lecerf, R., Van der Velde, M., Dentener, F., 2016. Impact of meteorological drivers on regional inter-annual crop yield variability in France. *Agric. For. Meteorol.* 216, 58–67.
- Confalonieri, R., Belloccia, G., Tarantola, S., Acutis, M., Donatelli, M., Genovese, G., 2010. Sensitivity analysis of the rice model WARM in Europe: exploring the effects of different locations, climates and methods of analysis on model sensitivity to crop parameters. *Environ. Model Softw.* 25, 479–488.
- Ewert, F., Porter, J., Honermeier, B., 1996. Use of AFRCWHEAT2 to predict the development of main stem and tillers in winter triticale and winter wheat in North East Germany. *Eur. J. Agron.* 5, 89–103.
- Faostat, 2017. Food and Agriculture Organization of the United Nations. GAEZ: Cereal Production Value per Hectare [ton/ha] (Measure). Accessed on:(10 Oct 2017). URL: <http://data.fao.org/>.
- Gablonsky, J.M., Kelley, C.T., 2001. A locally-biased form of the DIRECT algorithm. *J. Glob. Optim.* 21 (1), 27–37.
- Genovese, G., Bettio, M., 2004. Methodology of the MARS Crop Yield Forecasting System. Statistical Data Collection, Processing and Analysis. Vol. 4 European Communities, Luxembourg 92-894-8183-8.
- GEVES, 2017. Variety and Seed Study and Control Group. URL: <http://www.geves.fr/index.php?lang=en> (Accessed on):1. 9. 2017.
- GISAT, 2003. Studies in support to the MARS project. Crop Monographies on Candidate countries: MOCA study. Final report. GISAT, Praha, Czech Republic, pp. 439.
- Groot, J.J.R., Verberne, E.L.J., 1991. Response of wheat to nitrogen fertilization, a data set to validate simulation models for nitrogen dynamics in crop and soil. *Nutr. Cycling Agroecosyst.* 27, 349–383.
- Harrison, P.A., Butterfield, R.E., 1996. Effects of climate change on Europe- wide winter wheat and sunflower productivity. *Clim. Res.* <https://dx.doi.org/10.3354/cr007225>.
- Harrison, P.A., Porter, J.R., Downing, T.E., 2000. Scaling-up the AFRCWHEAT2 model to assess phenological development for wheat in Europe. *Agric. For. Meteorol.* 101, 167–186.
- Iizumi, T., Tanaka, Y., Sakurai, G., Ishigooka, Y., Yokozawa, M., 2014. Dependency of parameter values of a crop model on the spatial scale of simulation. *J. Adv. Model. Earth Syst.* 6, 1–14.
- Jagtap, S.S., Jones, J.W., 2002. Adaptation and evaluation of the CROPGRO-soybean model to predict regional yield and production. *Agric. Ecosyst. Environ.* 93, 73–85.
- Johnson, S.G., 2017. The NLOpt Nonlinear-Optimization Package, (URL): <http://ab-initio.mit.edu/nlopt>, Accessed on:10. 10. 2016.
- Jones, R.J.A., Buckley, B., 1996. European Soil Database. Information Access and Data Distribution Procedures. EUR 17266 EN, Space Applications Institute, Joint Research Centre of the European Commission, Ispra, Italy, pp. 35.
- Jones, J.W., Antle, J.M., Basso, B., Boote, K.J., Conant, R.T., Foster, I., Charles, H., Godfray, J., Herrero, M., Howitt, R.E., Janssen, S., Keating, B.A., Munoz-Carpena, R., Porter, C.H., Rosenzweig, C., Wheeler, T.R., 2017. Toward a new generation of agricultural system data, models, and knowledge products: state of agricultural systems science. *Agric. Syst.* 155, 269–288.
- Koch, E., Adler, S., Ungerböck, M., Zach-Hermann, S., 2010. PEP725 Pan European Phenological database. *Geophys. Res. Abstr.* 12, 14942 (EGU General Assembly 2010).
- Langer, S.M., Longin, C.F.H., Würschum, T., 2014. Flowering time control in European winter wheat. *Front. Plant Sci.* 5, 537.
- Lazar, C., Genovese, G. (Eds.), 2004. Methodology of the MARS crop yield forecasting system. Agrometeorol. Data Collect. Process. Anal. 2 92-894-8181-1. (European Communities, Luxembourg).
- Levis, S., 2014. Crop heat stress in the context of earth system modeling. *Environ. Res. Lett.* 9, 61002.
- Lobell, D.B., Gourdji, S.M., 2012. The influence of climate change on global crop productivity. *Plant Physiol.* 160.
- López-Lozano, R., Duveiller, G., Seguini, L., Meroni, M., García-Condado, S., Hooker, J., Leo, O., Baruth, B., 2015. Towards regional grain yield forecasting with 1km-resolution EO biophysical products: strengths and limitations at pan-European level. *Agric. For. Meteorol.* 206, 12–32.
- Macholdt, J., Honermeier, B., 2017. Yield stability in winter wheat production: a survey on German Farmers' and Advisors' views. *Agronomy* 7, 45.
- Micale, F., Genovese, G., 2004. Methodology of the MARS crop yield forecasting system. Meteorological data collection. Process. Anal. 1 (European Communities (EUR 21291 EN), Luxembourg).
- Monteith, J.L., Unsworth, M.H., 1990. Principles of Environmental Physics, 2nd Ed. Edward Arnold, New York, pp. 291.
- Montesino-San Martin, M., Wallach, D., Olesen, J.E., Challinor, A.J., Hoffman, M.P., Koehler, A.K., Rötter, R.P., Porter, J.R., 2018. Data requirements for crop modelling—applying the learning curve approach to the simulation of winter wheat flowering time under climate change. *Eur. J. Agron.* 95, 33–44.
- Mücher, S., De Simone, L., Kramer, H., de Wit, A., Roupioz, L., Hazeu, G., Boogaard, H., Schuilting, R., Fritz, S., Latham, J., Cormont, A., 2016. A new global agro-environmental stratification (GAES), Wageningen Environmental Research Report 276. doi:<http://dx.doi.org/10.18174/400815>.
- Müller, C., Elliott, J., Chryssanthacopoulos, J., Arneth, A., Balkovic, J., Ciais, P., Deryng, D., Folberth, C., Glotter, M., Hoek, S., Iizumi, T., Izaurralde, R.C., Jones, C., Khabarov, N., Lawrence, P., Liu, W., Olin, S., Pugh, T.A.M., Ray, D.K., Reddy, A., Rosenzweig, C., Ruane, A.C., Sakurai, G., Schmid, E., Skalsky, R., Song, C.X., Wang, X., De Wit, A., Yang, H., 2017. Global gridded crop model evaluation: benchmarking, skills, deficiencies and implications. *Geosci. Model Dev.* 10, 1403–1422.
- Orlandini, S., Nejedlik, P. (Eds.), 2012. Climate Change Impacts on Agriculture in Europe. Firenze University Press, pp. 206 (Final Report of COST Action 734).
- Porter, J.R., 1993. AFRCWHEAT2: a model of the growth and development of wheat incorporating responses to water and nitrogen. *Eur. J. Agron.* 2, 69–82.
- Porter, J.R., Gawith, M., 1999. Temperatures and the growth and development of wheat: a review. *Eur. J. Agron.* 10, 23–36.
- Porter, J.R., Semenov, M. a, 2005. Crop responses to climatic variation. *Philos. Trans. R. Soc. B* 360, 2021–2035.
- Powell, M.J.D., 1994. A Direct Search Optimization Method that Models the Objective and Constraint Functions by Linear Interpolation. In: Gomez, S., Hennart, J.P. (Eds.), Advances in Optimization and Numerical Analysis. Kluwer Academic, Dordrecht, pp. 51–67.
- Rajaram, S., van Ginkel, M., 2001. Mexico_50 Years of International Wheat Breeding. In: Bonjean, A.P., Angus, W.J. (Eds.), The World Wheat Book - a History of Wheat Breeding. Vol. 1. pp. 579–608.
- Rötter, R.P., Carter, T.R., Olesen, J.E., Porter, J.R., 2011. Crop-climate models need an overhaul. *Nat. Clim. Chang.* 1, 175–177.
- Shiferaw, B., Smale, M., Braun, H.J., Duveiller, E., Reynolds, M., Muricho, G., 2013. Crops that feed the world 10. Past successes and future challenges to the role played by wheat in global food security. *Food Security* 5, 291–317.
- Sima, M., Popovici, E.A., Bălteanu, D., Micu, D.M., Kucsicsa, G., Dragotă, C., Grigorescu, I., 2015. A farmer-based analysis of climate change adaptation options of agriculture in the Bărăgan plain. *Romania Earth Perspect.* 2, 1–21.
- Supit, I., Hooijer, A.A., Van Diepen, C.A. (Eds.), 1994. System Description of the WOFOST 6.0 Crop Simulation Model Implemented in CGMS. European Communities (EUR15956EN), Luxembourg.
- Taylor, K.E., 2001. Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res.* 106, 7183.
- Therond, O., Hengsdijk, H., Casellas, E., Wallach, D., Adam, M., Belhouchette, H., Oomen, R., Russell, G., Ewert, F., Bergez, J.E., Janssen, S., Wery, J., Van Ittersum, M.K., 2011. Using a cropping system model at regional scale: low-data approaches for crop management information and model calibration. *Agric. Ecosyst. Environ.* 142, 85–94.
- Trnka, M., Rötter, R.P., Ruiz-Ramos, M., Kersebaum, K.C., Olesen, J.E., Žalud, Z., Semenov, M.A., 2014. Adverse weather conditions for European wheat production will become more frequent with climate change. *Nat. Clim. Chang.* 4, 637–643.
- USDA, 2012. GAIN Report from Global Agricultural Information Network. (URL): https://gain.fas.usda.gov/Recent%20GAIN%20Publications/Crop%20Situation%20Update_Prague_Czech%20Republic_8-10-2012.pdf, Accessed on:1. 8. 2017.
- van Bussel, L.G.J., Stehfest, E., Siebert, S., Müller, C., Ewert, F., 2015. Simulation of the phenological development of wheat and maize at the global scale. *Glob. Ecol. Biogeogr.* 24, 1018–1029.
- Van Diepen, C.A., Wolf, J., Van Keulen, H., Rappoldt, C., 1989. WOFOST: a simulation model of crop production. *Soil Use Manage.* 5, 16–24.
- Waha, K., Van Bussel, L.G.J., Müller, C., Bondeau, A., 2012. Climate-driven simulation of global crop sowing dates. *Glob. Ecol. Biogeogr.* 21, 247–259.
- Wallach, D., Makowski, D., Jones, J., Brun, F., 2013. Working with Dynamic Crop Models, 2nd edition. Elsevier (25 504).
- Wang, E., Engel, T., 1998. Simulation of phenological development of wheat crops. *Agric. Syst.* 58, 1–24.
- Wei, X., Declan, C., Erda, L., Yinlong, X., Hui, J., Jinhe, J., Ian, H., Yan, L., 2009. Future cereal production in China: the interaction of climate change, water availability and socio-economic scenarios. *Glob. Environ. Chang.* 19, 34–44.
- Zampieri, M., Ceglar, A., Dentener, F., Toreti, A., 2017. Agricultural impacts of climate change wheat yield loss attributable to heat waves, drought and water excess at the global, national and subnational scales. *Environ. Res. Lett.* 12, 064008.