Simplified plant calibration in the STICS crop model based on remote sensing and yield data applied to spring wheat in semi-arid area

Crop production assessment is a worldwide issue for agriculture in the context of climate change which is expected to cause a shift in the production areas. To address this issue, process28 based models like STICS are being developed to simulate plant-environment interactions. However, adapting varieties in crop models requires a substantial amount of data, which poses as a constraint on their operational application. The objective of this work is to propose a methodology for calibrating STICS plant parameters to address scenarios with limited data access. This scenario, referred to as the basic optimization scenario (BOS), relies only on farmers' knowledge (soil properties, management, and yield) and access to open remotely sensed data. The methodology is applied to a spring wheat variety grown in the Central Plateau of Spain during 2015 – 2016 and is evaluated against a reference optimisation scenario (ROS) that also includes phenology and biomass field sampling data. In the BOS approach, the remotely sensed biophysical variable green leaf area dynamics (GAIRS) is used to describe the canopy dynamics. Additionally, phenological vegetation indices (PVI) were derived from specific temporal features of GAIRS curves which were used as an alternative to phenological field observations. The evaluation of the PVIs against phenological field observations showed a bias of between 9 and 1 days, depending on the nutritional status of crops, with a lower prediction error when crops are growing under non-stress conditions. The results show comparable results between ROS and BOS, with rmse ranging between 6 and 14 days for phenology, 1.2 and 1.4 m2 m-2 for GAI and 1.7 and 1.8 t ha-1 for yield, considering the uncertainties of the non-calibrated plant parameters, the STICS formulation, and the uncertainty in the description of the environmental conditions of crop growth (soil and climate). The similarities between the simulations obtained from the implementation of ROS and BOS over a wide range of climate conditions over a period of 20 years indicate the feasibility of using the parameters calibrated with the BOS approach for climate change studies. Finally, some technical issues related to the consideration of biomass data and the inclusion of a module to account for the contribution of ears to the GAI are analysed. The calibration of biomass is required for accurate simulations, while the lack of calibration had little impact on the other variables such as GAI, yield, and growth stages. The EAI module is not significant when canopies are well developed (maximum GAI >6.5 m2 m-2), but it improves simulations with less canopy, being essential for the assimilation of remotely sensed data into crop models. It would be important to acquire data sets on the contribution of ears, which are still scarce at present.

1. Introduction:

Crop production assessment is a global concern in the context of climate change. The Intergovernmental Panel on climate change Fifth Synthesis Assessment Report (IPCC, 2014) notes climate change impacts in agriculture include shifts in production areas around the world, which, combined with increasing food demand, pose significant risks to food security and intensify competition for water. In the past decades, several process-based models have been developed in agriculture to represent plant-environment-management interactions such as CERES (Ritchie and Otter, 1985), EPIC (J. R. Williams et al., 1984), DSSAT (Jones et al., 2003), APSIM (Keating et al., 2003), CropSyst (Stöckle et al., 2003), STICS (Brisson and Launay, 2008), MONICA (Nendel et al., 2011), WOFOST (Boogaard et al., 2014), AquaCrop (Steduto et al., 2009), and DAISY (Abrahamsen and Hansen, 2000). Literature shows that crop models have the potential to understand the vulnerability of regions to climate change by predicting crop growth and yield under different growing conditions, including water and nutritional stress (Katerji et al., 79 2013; Jin et al., 2015; Kassie, 2016; Rahimi-Moghaddam et al., 2021) as well as variations in temperature, CO2, and irradiance (Asseng et al., 2013; Rosenzweig et al., 2013; Ewert et al., 2015; Tao et al., 2020). Other authors applied crop models for mitigation and adaptation purposes, exploring adapted management practices (Hadria et al., 2007; Nendel et al., 2014; Ngwira et al., 2014; Bécel et al., 2015; Kadiyala et al., 2015; Bai and Gao, 2021; Justino et al., 2022) and supporting crop breeding (Rötter et al., 2015; He et al., 2017; Yang et al., 2022). Crop models represent plant-environment-management interactions through equations driven by plant parameters, which may vary for each variety of a given species (Manivasagam and Rozenstein, 2020). Therefore, plant parameters as well as the inputs describing crop growing conditions, i.e. weather, soil, management, and initialisation, are the main sources of uncertainty when running crop models (Jin et al., 2018; Hao et al., 2021). They are usually adapted based on literature, in situ measurements, and calibration procedures (Attia et al., 2016; Razzaghi et al., 2017; Seyoum et al., 2018; Falconnier et al., 2019). In calibration procedures, simulations of crop development and growth are matched to the corresponding observations by fitting plant parameters. Jégo et al. (2011) show some limitations in the calibration in dry conditions, highlighting the importance of mastering the environmental conditions both in the calibration phase and in the model implementation phase. Therefore, it is necessary to mobilise large data sets describing crop development and growth as well as growing conditions, which are not always easy to collect, being one of the limiting factors for the calibration and, consequently, the implementation of crop models. Furthermore, the fact that the selection of parameters and the order of the steps in the calibration procedures can have an impact on the results cannot be neglected. Guillaume et al. (2011) showed that focusing on a variable such as yield gives better results for yield than attempting to represent both yield and leaf area index development. Jégo et al. (2011, 2012) suggested that calibration can be simplified to one variety per region, thus limiting to a reduced number of parameters while maintaining good predictive quality over the area considered. Therefore, when facing a calibration process, we are confronted with the dual challenge of data availability and standardisation of calibration protocols. The use of remote sensing can be a way to overcome these two limitations, as will be shown in this work. Remote sensing data allows for the derivation of key biophysical variables for crop monitoring, such as the area of green leaves in a canopy (known as green area index or GAI) (Weiss et al., 2020), the fraction of absorbed radiation (fAPAR), biomass or phenological stages (Zhang et al., 2003; Sakamoto et al., 2005; Lobell et al., 2013; González-Gómez et al., 2018; Zhou, 2019; Meroni et al., 2021; Liu et al., 2022). Recent advances in remote sensing technology have significantly improved the potential for accurate and reliable estimates of canopy biophysical variables (Jin et al., 2018), such as remote sensors on-board the satellites Landsat8 and Sentinel2, which offer a wide range of open-access spectral observations of the Earth's surface. The launch of Sentinel2 in 2015 and 2017 has boosted remote sensing capabilities by providing unprecedented high spatial and temporal resolution (decametric and 5-day revisit time) (Weiss et al., 2020), resulting in a wealth of open data sets suitable for working at the field scale and witha good revisit frequency, enabling effective monitoring of crop dynamics throughout the growing season simultaneously over wide areas, and offering a powerful alternative to sampling data.

The coupling between remote sensing data and crop models has been explored for a long time. There exists an abundant literature on the use of remotely sensed data to improve crop models simulations by mitigating the impact of uncertainties associated to model plant parameters. Reinitialisation of some of the parameters or the initial conditions of the models is the most common approach. Maas (1988) proposed to improve the yield predictions of a simple crop growth model, reinitializing the model's state variable GAI through tuning the initial values to match simulation with remote sensing observations. Guérif and Duke, (2000) and Jégo et al. (2012) worked on the initial stages of the vegetation. They use the time series of GAI to estimate the emergence date, which is a critical stage for the simulation of the subsequent phenological stages. Prévot et al. (2003), meanwhile, focused on the sowing date. An alternative way of using remotely sensed data is by forcing crop model simulations (Huang Jingfeng et al., 2001). For instance, Casa et al. (2012) used STICS model as a transformation function to infer performance from GAI and showed the interest in having frequent and accurate measurements overcoming some uncertainties in the management data. Data assimilation schemes are an improvement of the forcing method. For instance, the ensemble Kalman Filter (EnKF) is an optimal way to correct a model simulation by taking into account model and observation errors as well as the correlation between variables (de Wit and van Diepen, 2007; Nearing et al., 2012; Zhao et al., 2013; Linker and Ioslovich, 2017; Fattori Junior et al., 2022; Zhuo et al., 2022). Most of these studies used a plant-calibrated crop model and have shown a gain in simulation accuracy. Fattori Junior et al. (2022) underline the fact that the gain is maximum with non-calibrated plant genotypes. However, some limitations were highlighted when the link between the target variable and observable quantities is not well simulated by the model due to either the model parameters or process parameterization (Nearing et al., 2012; Fattori Junior et al., 2022), highlighting the need of well calibrated crop models. While the literature is abundant on the use of remote sensing data to correct a simulation by attenuating the impact of uncertainties in the model's inputs, there are relatively few studies on the use of remote sensing for calibration purposes. For instance Prévot et al. (2003) used airborne data to demonstrate the benefits of having large number of measurements to calibrate two key STICS parameters regulating phenology and leaf development, obtaining a good simulation of GAI. Rodriguez et al. (2004) focused on the calibration of the parameter describing GAI development in STICS and showed a positive impact of this calibration on processes such as water transfer and surface energy balance. Liao et al. (2019) calibrated parameters related to phenology and leaf area development in the SAFY simple crop model. They used phenological stages inferred from the application of the TSF shape model developed by Sakamoto et al. (2005) to time series of vegetation indices as observations of phenology. After calibration, they found a positive impact on biomass and yield simulations even if the biomass to yield conversion parameters were not calibrated. In the literature, several works can be found that show the possibility of deriving phenological stages from canopy reflectance dynamics such as Zhang et al. (2003), Sakamoto et al. (2005) or Zhou (2019). These results showed the potential of using remote sensing data for the calibration of not only parameters related to canopy development, but also parameters related to crop phenology. In these studies, a good estimate of leaf development enables a good estimate of yield even if the parameters enabling biomass and yield conversion are not calibrated. For this, in situ yield data are required. The aim of this work is to propose a methodology for the calibration of plant parameters of the STICS crop model under limited access to data taking advantage of open-access remote sensing data as an alternative to field sampling. The proposed calibration approach is based on easily available in situ data based on farmers' knowledge, i.e. management and yield, and remote sensing data to infer green leaf area dynamics and phenology. An optimisation scenario with field measurements of phenology and biomass was considered as a benchmark. The study was implemented on a spring wheat variety grown in the semi-arid conditions of the Central Plateau of Spain. Thus, to overcome the interaction with the environmental conditions that make the delineation with the plant development characteristics difficult, well-irrigated crops are used for the calibration of plant parameters. The validation of the calibrated parameters was conducted to assess their ability to simulate crop development and growth under different stress conditions.

2. Materials and methods

2.1. Study sites

A total of 12 study sites of commercial irrigated spring wheat fields grown during 2015 and 2016 are included in this study (Figure 1). The study sites are located on the Central Plateau of Spain and define a study area of 3600 km2 (39º21´N 2º6´W; 38º40´N 1º32´W) with around 1000km2 of irrigated crops. The climate zone is characterised as Mediterranean with continental effects, with an annual average air temperature of 14°C, annual cumulative potential evapotranspiration of 1242 mm and annual cumulative precipitation of 363 mm based on records from 2002 to 2021 of 4 weather stations distributed along the study area and operated by the Servicio Integral de Asesoramiento al Regante (SIAR) of Spanish Ministerio de Agricultura, Pesca y Alimentación (MAPA) (open data: http://eportal.mapama.gob.es/websiar/SeleccionParametrosMap.aspx?dst=1). Soils are calcareous classified as Calcixerepts (Soil Survey Staff, 2014) typical of the Mediterranean region. Figure 1. Study area: distribution of fields and weather stations considered in this study and Sentinel2 satellite tiles.

2.2. STICS crop model

STICS model (Brisson et al., 1998) is a generic crop model which simulates crop development, carbon, water and nitrogen dynamics in the soil–crop–atmosphere system for multiple crops (Guillaume et al., 2011). The version V8.50 (http://www6.paca.inrae.fr/stics_eng) is used in this work. The model is divided into modules that simulate crop development and biochemical and hydrological processes (Brisson and Launay, 2008) for a specie-variety growing under potential conditions that can be reduced according to thermal or nutritional constrictions (water or nitrogen stress). It runs daily over a site-year-management unit, termed as the Unit of SiMulation (usm), characterised by climate conditions, soil properties and management practices. The modules and plant parameters on which this study is focused are briefly described next: crop development, growth and yield formation. Crop development is variety dependent and is driven between growth stages by the daily accumulation of photothermal units (upvt, ºCd). upvt corresponds to daily cardinal degree-days (ºCd) from emergence date (lev) corrected by the required days of vernalisation jvc (days) and the sensitivity to the photoperiod sensiphot (dimensionless). The crop development is defined by the time intervals in upvt units between wheat growth stages, hereafter referred to as pheno-phases: stlevamf (from emergence lev to the beginning of maximal growth amf), stamflax (from amf to flag leaf lax), stlevdrp (from lev to the onset of grain filling drp) and stdrpmat (from drp to physiological maturity mat). Crop growth formulation is divided into light interception driven by green leaf area expansion and aerial drymass production. The simulation of green leaf area expansion is divided into the growth and the senescence phase. The growing period is simulated by a logistic function driven by a normalised leaf development unit in upvt units and scaled to plant population by adens (dimensionless). adens determines the ability of a variety to withstand increasing densities. Green leaf area senescence is based on the concept of leaf lifespan (Maas, 1993) that simulates the loss of greenness of leaves once the lifetime has elapsed driven by the variety parameter durvieF (cumulative Q10 unit). In order to make green leaf area index (GAI) simulations comparable with the GAI as observed by the satellite (GAIRS), the contribution of the green area index of ears (EAI) is simulated apart in STICS and added to the GAI simulations of the model in a post-processing phase. In the EAI simulations, it is assumed that ears appear linearly from lax until a maximum EAI value (eaimax) during a prescribed period stmax. Then, a plateau period is simulated of stplat duration, followedby a linear decrease to a value EAI=0 with the senescence of the green leaf area during stsen. An illustration of the process can be seen in Figure S3 in Section 3 of Supplementary Material (SM). As a few data exist in literature about the growth of the ears, measurements made by Casals, (1996) were used for the module parameterisation: stmax= 400 ºCd; stplat= 120 ºCd; stsen= 370 ºCd; eaimax = 1 m2 m-2. Measurements showed a maximum EAI of 0.6 m2 m-2, which was rounded to 1 m2 m-2 to take into account the bearbs, which are important in the case of the varieties considered in this study and agree with the upper limit reported by Weiss and Baret (2016) for wheat. The simulation of growth of aerial drymass, hereafter referred to as biomass, is daily computed based on specie maximum use efficiency of intercepted radiation (RUE, gMJ-1) during the juvenile phase efcroijuv (from lev to amf), the vegetative phase efcroiveg (amf – drp) and the reproductive phase efcroirep (drp – mat). Yield formation is simulated by assuming that a variety can reach a maximum grain weight (pgrainmax) and develops a maximum number of grains (nbgrmax). Yield formation is represented assuming a progressive carbon allocation from biomass to a grain by applying a harvest index (vitircarb, g grain g biomass-1 d-1) during the grain filling period, stdrpmat, not exceeding pgrainmax. The number of grains is determined at the time of the onset of grain filling, drp, considering the mean growth rate of the crop during the period preceding to drp driven by cgrain (g-1 m2).

2.3. Field data Field sampling were made on 75 sampling areas (plots) of 1 ha size distributed in the 12 fields considered in this study (Figure 1). Three spring wheat varieties (Califa, Atomo, and Galera) were cropped during 2015 and 2016 in the frame of an experiment designed for the evaluation of the impact of the variable rate nitrogen fertilisation management on yield (Campos et al., 2019). In this study, the three varieties were considered as a unique variety since they have similar morphogenetic characteristics classified as short-cycle, semi-early spring wheat of medium-low height, and then a single plantfile was calibrated. The wheat season covered a period from mid December (sowing period from 8th December to 22th January) to early July. Plant phenology, aerial biomass and soil properties were sampled on 3 geolocated areas of 1.5m×2 crop rows on each plot (sampling points) from 2 leaves developed until physiological maturity. Plot values were obtained from the average of the sampling points. The management of each plot was reported by farmers. Phenological observations were made during 2015 and 2016 on BBCH scale (Lancashire et al., 1991) at intervals of 7 – 20 days. Noted that, observations were not taken at the exact growth stages as required by STICS and very few emergence dates were observed. Moreover, sinceobservations were not regularly observed and the transition from one phenological stage to the next might vary, some delay on the days noted with respect to the exact date of the onset of thestage is expected. The adaptation approach of the phenological field observations to the growth stages as required by STICS adopted in this study is described in section 2.4. Aerial biomass dry weight, hereafter referred to as biomass, was determined from the above ground biomass field sampled and dried at 80 °C until a constant weight was reached. At maturity, yields were determined by weighing grains. Soils samples were taken at two depth layers, 0 – 20 and 20 – 40 cm depth. Granulometry (Bouyoucos method), calcium carbonate content (Bernard method), soil organic matter content (Walkley-Black method), organic and mineral nitrogen content (Kjeldahl method), pH (pH meter) and depth of surface calcareous horizons when it appears above 40 cm were determined. Soil characteristics are presented in Table 1.

Table 1. Pedoclimatic spring wheat growing conditions. Environmental indicator Parameter/definition Range Soil Soil clay content (%) (g kg-1) 16 – 51 Soil organic content (%) (g kg-1) 6.0 – 13.5 Soil organic nitrogen content (%) (g kg-1) 0.1 – 0.3 Water amount of field capacity (%w.) 15.9 – 29.4 Climate1 Average daily mean temperature (ºC) 11.1 Occurrence of days with temperature < 0ºC 22 – 68* Mean daily global radiation (MJ m-2 d-1) 15 – 17 Cumulative rainfall wheat season (mm) 124 – 235 Cumulative evaporative demand (mm) 581 – 694 Daylength (h/d) 10.4 – 16.0 1: arithmetic mean of daily observations along the wheat growing season. *Significant differences between years: (2015) 30 – 68; (2016) 22 – 48;

2.4. Adaptation of phenological field observations to STICS growth stages STICS growth stages were derived from the phenological field observations (section 2.3) based on the correspondences presented in Table 2. Instead of using a single BBCH index, a range of BBCH indices (enclosed in brackets in Table 2) was considered around the BBCH index corresponding to each growth stage (highlighted in bold in Table 2). This approach was adopted to prevent bias in the observed field dates (section 2.3). It is important to note that BBCH31 was also included, although it was aligned with any STICS growth stage, because of its interest in the determination of the photoperiod sensitivity of the variety (refer to section 2.6 for details). lev dates, not observed in field, were inferred from the annotations of the appearance of the first leaves by applying the phyllochron concept. Phyllochron is defined as the thermal time interval between the appearance of two successive leaves (Gate, 1995; Slafer and Rawson, 1997), which is assumed constant up to 8 leaves (Jamieson et al., 1995). The phyllochron can be affected by both genetic and environmental factors (Slafer, 2015). Consequently, a phyllochron value was computed for each field. Values in the range of 70 – 100 growing degree-days (gdd, ºCd) were obtained, with lower values for more intensively managed, i.e. unstressed, fields.

2.5. Green Area Index and phenology from remote sensing data In this work, GAIRS derived from satellite observations was used to monitor: i) the growth dynamics of the canopy and ii) the crop development. GAIRS was computed using the BV-NET algorithm (Weiss and Baret, 1999) applied to cloud free multispectral images of Landsat 8 OLI (L8, 30 m bands: B3, B4, and B8) and Sentinel-2 (S2, 301 10 m bands: B3, B4, and B8). L8 was used for 2015 monitoring and S2 for 2016. The time interval of acquisition dates varied between 3 days (path overlapping) to more than 1 month (cloudy periods), leading to 6 L8 dates over the wheat season in 2015 and 12 S2 in 2016. The growth dynamics of the wheat canopy were monitored by considering the average GAIRS value of the pixels in a plot. Crop development was monitored by detecting phenology through the timing of specific temporal features in the seasonal dynamics of GAIRS. This approach led to the definition of Phenological Vegetation Indices (PVI) corresponding to the times when the features of interest occurred. A seasonal relationship was established by fitting the analytical model proposed by Baret (1986) on GAIRS time series. The model combines a logistic function for the growth phase, and an exponential function for the senescence phase, based on daily growing degree-days (gdd, ºCd) (Equation 1). GAI(gdd) ~ $GAImax$ [1 1 + $er\ g(gddm{-}gdd) - er\ s(gdds{-}gdd)$] Equation 1 gdd are computed as the daily average air temperature from a base (0ºC) to a maximum (33ºC), which are also values used in STICS for wheat. In Equation 1, GAImax represents the maximum GAI, rg is the parameter of the logistic function that describes the relative growth rate at the inflection point gddm, and rs is the parameter of the exponential function that describes the relative senescence rate when all leaves have senesced gdds, i.e. when GAI = 0. A rmse = 0.30 m2 m-2 was obtained between the simulated GAI and GAIRS observations, which is comparable to the value reported by Casa et al. (2012). Subsequently, the following five PVIs were characterized using threshold and analytical methods: Beginning of Growth Rate (BGR), Mid Growth Rate (MGR), Beginning Of Plateau (BOP), Mid Of Plateau (MOP) and End Of Season (EOS) (Figure 2). Figure 2. Example of the methodology to derive the phenological vegetation indices (PVI) (dotted) from GAIRS dynamics and their 1st (long dashed line) and 2nd (short dashed lines) derivatives. BGR and EOS were derived using threshold methods to detect the green-up and the end of the season, near the maturity, respectively. They were defined as the points when daily GAI values reach 10% of the seasonal maximum, as used in previous studies by Lobell et al. (2013) and González-Gómez et al. (2018). MGR, BOP, and MOP were derived using analytical methods. MGR was determined as the time of maximum GAI growth rate (rg), as proposed by Zhang et al. (2003) and Zhou (2019), indicating an intermediate stage during the vegetation growth, around the onset of the stem elongation. BOP was identified as the second local minimum of the second derivative of the daily GAI values, indicating a slowdown in vegetative growth rate after period of rapid growth and just before vegetation stops growing around the time of the flag leaf stage. MOP was derived as the second local maximum of the second derivative of the daily GAI values, as proposed by Sakamoto et al. (2005). This point indicates the time when the crop reaches its maximum GAI values and undergoes a transition from the vegetative to the reproductive phase around flowering, accompanied by the onset of leaf senescence. The correspondences between PVI and STICS growth stages were established based on the comparison of their time courses as a functions of STICS green leaf area (GAISTICS). To achieve this, growth stages and green leaf area (GAISTICS) were simulated using STICS with different bread wheat varieties commonly cultivated in the northern France (49.5ºN, 3.7ºE) and Yecora Rojo spring wheat variety adapted to semi-arid areas. Subsequently, GAISTICS was used as inputfor the PVI derivation approach, and the resulting PVIs (PVISTICS) were compared with the STICS growth stages. The corresponding relationships are presented in Table 2. An rmse of 11 days and bias of 1 day were computed between PVISTICS and STICS simulated growth stages. Most of the stages of interest exhibited an rmse ≤ 8 days, except BOP (lax) which showed larger errors with an rmse = 19 days. Furthermore, the PVIs were evaluated against BBCH field observations (section 2.4). To delineate the differences in the GAI dynamics between crops the assessment was performed by grouping the data set into the calibration and the validation sets (section 2.8). To facilitate the comparison at the dates of remote sensing observations, BBCH observations were interpolated by fitting the observed BBCH time series to analytical functions described in Section S1 in Supplementary Materials.

2.6. Optimisation scenarios In this work, three calibration scenarios, referred to as optimisation scenarios, were considered 360 to assess the proposed methodology for calibrating STICS plant parameters under varying conditions of data availability. The proposed methodology is designed under the assumption that the minimum available data include typical farmer knowledge (crop management, harvested yield, and general soil properties such as soil texture, soil depth, and occasional soil mineral nitrogen content) along with open access remote sensing data. Based on this assumption, the least documented optimisation scenario is referred to as the Basic Optimisation Scenario (BOS). This scenario utilizes both GAIRS to describe canopy dynamics and PVI to describe phenological development. BOS is structured in 5 successive steps, organized according to the influence of fitted parameters from earlier stages on subsequent ones. The calibrated parameters, target output variables, and methods employed in each step are presented in Table 3. A total of 12 plant parameters were calibrated, selected for their sensitivity in simulating each respective target variable. A comprehensive description is available in section 2.7.

Table 3. Optimisation scenarios for the calibration of STICS plant parameters. Step Module Parameter Target variable Optimization scenarios ROS ROSnB BOS 1 Photothermal time units (upvt) jvc, sensiphot Phenology X X X 2 Crop development stlevamf, stamflax, stlevdrp, stdrpmat Phenology dates corresponding to amf,lax,drp and mat X X X 3 Biomass formation efcroiveg, efcroirepro Biomass X 4 Green leaf area formation adens, durvieF GAIRS X X X 5 Potential yield pgrainmax, nbgrmax Yield X X X 6 Yield formation cgrain, vitircarb Yield X X X The Reference Optimisation Scenario (ROS), in contrast to BOS, represents the most extensively documented optimisation scenario. It incorporates field measurements of phenology and biomass to the data used for BOS, serving as the benchmark against which BOS is evaluated. Consequently, the calibration process for ROS involves an additional step for calibrating biomass parameters, resulting in a total of 6 steps and 14 parameters (Table 3). Finally, biomass data are not always available due to the effort required for field sampling. However, its role within STICS is of high importance for simulating both green leaf area and yield. Therefore, in order to assess the impact of biomass in the calibration process, an intermediate optimisation scenario based on ROS but without the biomass step, the Reference Optimisation Scenario without biomass (ROSnB, Table 3) was also included.

2.7. Optimisation approach The general optimisation approach considered in this work, presented in Table 3, is described in this section. Step1. The parameters jvc and sensiphot, which play a role in the simulation of photothermal units (upvt), were initially calibrated due to their significant impact on development parameters (step2). Optimal values were selected to minimize the variability in upvt, allowing the achievement of various growth stages across different climatic conditions of the plots within the calibration data set. Variability was quantified using the coefficient of variation (CV%) which measures the deviation to the mean. jvc was determined by spanning its range of values between 0 and 38 days. The CV was evaluated exclusively at the amf as vernalization requirements are not expected at a later stage. sensiphot was calibrated at amf, lax and drp stages, as no photoperiod effect is expected beyond drp. Additionally, BBCH31 was included in the assessment, despite not being a STICS growth stage. This inclusion aligns with Slafer (1996) and Ochagavía et al. (2018) who identify it as a photoperiod-sensitive period. Step2. The pheno-phases were derived by summing the daily upvt values between growth stages of the plots in the calibration data set.

Step3. efcroiveg and efcroirepro parameters were adjusted to align biomass simulations with biomass measurements. This step was performed before calibrating the green leaf area parameter (step4) due to the significant role of biomass in simulating leaf senescence.
Step4. adens and durvieF parameters were adjusted to align green leaf area simulations with GAIRS observations. Step5. The parameters that define the maximum potential yield of the variety nbgrmax and pgrainmax were first determined. pgrainmax was derived from the correlation between yield and grain weight, based on data collected from official trials conducted by different organisations across Spain. This correlation was extrapolated to estimate the potential yield (refer to Section S2 in SM). Subsequently, nbgrmax was calculated from the relationship between potential yield and pgrainmax.

Step6. cgrain and vitircarb parameters were adjusted to align yield simulations with measurements. The adjustment of parameters to align STICS simulations with the target observations (step3, step4, and step6) was performed by minimizing the goodness-of-fit criterion using the Nelder Mead simplex algorithm. This algorithm is used for multidimensional minimisation of any parameter within the STICS crop model (Wallach et al., 2011). The initial plantlife for the Yecora Rojo variety, already available in STICS, was employed. Its parameters were updated throughout the calibration process, while keeping uncalibrated parameters unchanged. Yecora Rojo (YR) corresponds to a spring wheat variety adapted to a semi-arid environment. It was calibrated as part of the crop model comparison within the AgMIP exercise, using an extensive database collected from an agronomic trial (Asseng et al., 2015). The soil properties and the initialisation conditions for STICS simulations were primarily described according to the main assumption of this work, which involves considering typical farmer knowledge. Consequently, soil profiles were characterized in three layers of 20 cm depth each, in line with the typical soils of the study area. These soils are mainly shallow, with a calcareous bedrock near the surface. Hydrological properties for each layer were derived using pedotransfer functions (Wösten et al., 1999). It is expected that this approach might introduce some uncertainty into the calibration process. In simulations, the interactions between environmental streses related to water and nitrogen nutrition and plant development heavily rely on a precise characterization of the soil and initialisation conditions. To mitigate this, two approaches were taken. Firstly, the calibration set included plots with the highest yield, well irrigated and well-fertilised throughout the growing season, under the assumption that environmental stress was negligible. Secondly, simulations were initialised just after substantial rainfall events, allowing the initialisation of the simulations at field soil capacity, reducing the impact of initial soil conditions. The plant parameters obtained from each optimisation scenario were assessed using prediction errors metrics, namely rmse (Root Mean Square Error) and bias. These metrics were calculated between the phenology, green area index, biomass, and yield data, and the STICS simulations.

rmse and bias were computed separately for the calibration and validation datasets, as outlined in section 2.8. 2.8. Optimisation and validation data sets To address the uncertainties arising from the inaccurate characterization of soil and initial conditions, specific calibration and validation data sets were established. As a result, plots with the highest yield, well-irrigated and well-fertilised throughout the growing season, were selected for calibration purposes. The plots included in the calibration data set (DBOpti) were identified based on both the maximum GAIRS during the whole season and highest yield. A total of 6 plots were selected for DBOpti. The remaining 63 plots, independent from DBOpti, were employed for evaluating the calibrated plant parameters (DBValid). Furthermore, DBValid was divided into three validation subsets (DBValid1 – 3) also based on GAIRS dynamics and yield to represent three different levels of water and nitrogen stress attributed to crop intensification. A comprehensive description of the data sets, encompassing the range of GAIRS values, yields, as well as the total nitrogen and water doses applied, is provided in Table 4.

Table 4. Calibration (DBOpti) and validation (DBValid) data sets. DBOpti DBValid1 DBValid2 DBValid3 n 6 24 25 14 Sowing date (Julian days1) 358±8 369±17 372±16 359±15 N fertilisation (kgN/ha) 304±32 226±69 211±56 137±35 Irrigation (mm) 499±4 440±42 421±54 298±71 Yield (t/ha) 9.6±2.0 9.7±1.4 7.4±1.3 5.1±0.9 GAIRS (m2/m2) 7.6±0.5 6.8±0.8 5.6±0.6 4.7±0.5 1: Julian days of two consecutive years. n: number of usms included in each group; arithmetic mean±standard deviation of the usms of each group of the total amount of nitrogen fertilized (N fertilization) and irrigation water (Irrigation) applied during the season, yield field measured (Yield) and maximum GAIRS (GAIRS).

3. Results

3.1. Phenological vegetation indices (PVI) The PVIs, used in BOS as phenology observations, were evaluated against the BBCH field observations to determine their accuracy in predicting STICS growth stages. The prediction errors of each PVI in DBOpti showed a moderate bias of <5 days, except for BOP which showed the largest deviation of -9 days (Table 5). BOP, intended to detect the lax stage (flag leaf stage BBCH39), occurred earlier than expected, during stem elongation (BBCH34). This advance can be attributed to the fact that remote sensors introduce uncertainties when GAIRS values exceed 6 m2 m-2, owing to the saturation of the remote sensing signals in densely vegetated areas (Weiss and Baret, 2016). This phenomenon accelerates the rapid development phase and leads to the slowdown of the vegetative growth rate before the flag leaf, as observed in high-nutrient DBOpti plots. The accuracy identified in this study is consistent with results presented in section 2.5, where the PVIs derived from simulated STICS GAISTICS time series were compared with simulated phenological stages. Hence, the discrepancy between PVIs and phenology, particularly with BOP and its relation to the lax stage, could potentially influence the calibration of plant parameters. This impact will be further evaluated.

The results presented in Table 5 also highlight the importance of paying attention to BGR, as a significant bias was observed in DBValid3, which comprises the less irrigated plots. These particular plots exhibited the shortest canopy greenness dynamics throughout the season, which was further influenced by lower crop density, resulting in delayed predictions of vegetation phase stages and earlier predictions of maturity. Zhou (2019) reported that the temporal patterns of crops, as described by NDVI time series, can be influenced by environmental factors like temperature, precipitation, and soil moisture. Consequently, the selection of plots that have grown under optimal conditions becomes crucial for appropriately parameterizing crop development in STICS, supported by remote sensing observations. This is also important to ensure that STICS parameterization adequately replicates GAI as observed by satellites, which should be achieved under non-stress conditions.

3.2. Calibration of plant parameters The parameters obtained in each calibration scenario, ROS, ROSnB and BOS, are presented in this section following the different steps (Table 3). Step 1 (Photothermal time unit). In the upvt parameterization, the jvc value was calibrated initially. In the case of ROS, the minimum CV% of accumulated upvt at the amf stage, calculated from the plots of DBOpti was obtained with jvc values of 0 and 10 days (Table 6). These findings indicate that the variety does not require days of vernalization, as expected for spring wheat varieties. This conclusion aligns with the parameters of Yecora Rojo (YR) and with the results reported by Jégo et al. (2010) for spring wheat in Canada. Similarly, with BOS, the minimum CV% was also achieved with jvc = 0 (Table 6). This consistency in the results of ROS and BOS was maintained despite the larger values of CV% attributed to the broader range of dates corresponding to BGR stage. Moving forward to subsequent calibration steps, a jvc value of 0 was retained for both ROS and BOS scenarios (Table 7).

In relation to sensiphot, the CV% within upvt was computed for each growth stage between lev and drp and summed to provide a comprehensive indicator. This approach is chosen since a single sensiphot value is used in STICS for upvt calculation. Notably, while only one year was included in DBOpti, advantage was taken of the difference between the dates of emergence of approximately one month to have variations in photoperiod. In the case of ROS, the results were consistent regardless of the tested sensiphot values (Table 6). Our results do not decisively establish the significance of photoperiod on wheat development. The results in the literature (Slafer, 1996; Ochagavía et al., 2018) tend to suggest that sensitivity to photoperiod is lower at the early stages of the season (before BBCH31) than during later stages. However, upon considering the CV% from BBCH31 to drp, it becomes apparent that the best results were obtained with sensiphot=1, suggesting no substantial impact of photoperiod. For BOS, similarly to ROS, no significant effect of photoperiod is observed between the values tested and the stages considered. These results lead us to adopt a sensiphot value of 1 for both ROS and BOS. Step 2 (Pheno-phases). The durations of the pheno-phases obtained in ROS were significantly different from those in YR (Table 7), due to the different sensiphot values (0.1 for YR while 1 in our study). Similar results were reported by Jégo et al., (2010) for stlevamf and stdrpmat, where values of 235 ºCd and 700 ºCd were reported respectively. When considering BOS, the pheno

phases obtained were generally similar to ROS, except for stamflax, which yielded a shorter value of 116 ºCd, due to the dates derived for BOP (section 3.1). The comparison of the pheno-phase parameters influenced by sensiphot with other works in the literature proved challenging, as the same parameters are not always calibrated and the complete plantfile is not frequently reported. Step 3 (Biomass). Aerial biomass parameters obtained with ROS were significantly different from YR, with efcroirep higher than efcroiveg, 3.5 g MJ-1 and 3.1 g MJ-1 respectively (Table 7). This indicates that the values used in YR, which correspond to the generic values proposed in the STICS model distribution for wheat plant group, were unsuitable for simulating biomass dynamics under the calibration conditions. It is important to note that these generic values are used in both ROSnB and BOS. Step 4 (Green leaf area). The durvieF value in ROS was lower at 238 compared to YR. Conversely, very similar values to YR were obtained in both BOS and ROSnB (Table 7). The common point between BOS and ROSnB was the absence of biomass parameters calibration, leading to overestimated biomass simulations (section 3.3). Regarding adens, the impact of non calibration of biomass was limited, as the fitted value of ROSnB was similar to that of YR and ROS. In BOS, a lower adens value was derived, which can be attributed to the shorter duration of the vegetative period indicated by the stlevdrp parameter. Step 5 (Potential yield). The maximum yield potential threshold, parameterised by pgrainmax and nbgrmax, was increased to 12.5 t ha-1 to match the maximum yield measurements (Table 7). These revised values were uniformly applied across all optimisation scenarios, as they were established based on accessible reports from variety trials, assumed accessible at all levels. Step 6 (Yield formation). In terms of the calibrated parameters, cgrain was larger in ROS than in YR (Table 7), primarily aimed at adjusting the yield simulations to lower biomass values. In STICS, cgrain is determined as dependent on biomass during 30 days preceding the onset of grain filling, drp. Consequently, the lower value obtained in ROSnB was calibrated to compensate the higher biomass simulations. With BOS, similar values to ROS were obtained, and this coincidence was to balance the dynamics of growth of the biomass of the vegetative period during cgrain calibration. Regarding vitircarb, similar values were obtained for all scenarios closely aligning with YR. This consistency highlights the stability of viticarb as a parameter, in accordance with its characterization as a specie parameter. 3.3 Optimisation scenarios evaluation The plant parameters calibrated for the different scenarios were evaluated using prediction error (rmse, bias) of key variables linked to the phenology, GAI, biomass, and yield. The evaluation was made on both DBOpti and DBValid data sets.

The STICS simulations of growth stages were first evaluated against the simulations when lev dates were forced (Table 8). lev is a crucial stage for simulating growth stages and is significantly influenced by water conditions at sowing. Therefore, the errors in the simulation of lev may have a strong effect on the subsequent stages. When lev was simulated by STICS, an rmse of 3 days was obtained in DBOpti and 14 days in DBValid regardless of the optimisation scenario, showing influenced by errors in water stress simulations. In terms of the subsequent stages, a slight increase proportional to the lev error was observed in both rmse and bias for amf and, to a lesser extent, for lax when lev was simulated by STICS, highlighting the impact of lev in the later stages. With ROS, the evaluation of crop development simulations using DBOpti and forcing lev showed similar rmse values to those obtained with YR forcing lev (rmse < 7 days). However, there was a reduction in bias from amf to drp, highlighting the improvement introduced with the ROS parameterisation. It is noteworthy that there is a positive impact on drp simulations, which were more accurately simulated (positive bias) compared to YR (negative bias). When ROSnB was applied to DBOpti, differing from ROS by not forcing lev and not calibrating biomass, a slight increase in rmse was observed, influenced by errors in stress simulations resulting from inaccurate biomass simulations. When BOS was applied to DBOpti, comparable results to ROSnB were obtained, although some differences were observed due to biases introduced by the use of the PVIs (Table 8). In particular, lax was now estimated earlier with a bias of -3 day due to the shorter period defined for stamflax. On the other hand, drp showed a reduction in both rmse and bias to 3 days, due to the longer stlevdrp period. These results highlight the limitations of using PVIs to parameterise crop development in contrast to using field observations. Such limitations should be carefully assessed for their potential impact on GAI and yield simulations. The assessment of crop development simulations on DBValid with the three optimisation scenarios showed a significant increase in the prediction errors compared to those reported on DBOpti with rmse greater than 6 days levels, potentially making its simulations susceptible to inaccurate simulations of stress conditions due to poor characterisation of soil properties and initialisation conditions. Biomass. The simulations of biomass dynamics throughout the growing season using ROS onDBOpti plots showed a reduction in rmse of about 2 t ha-1 (rmse = 1.6 t ha-1). This improvement is mainly attributed to a significant reduction in bias (bias = -0.5 t ha-1) compared to the optimisation scenarios where parameters controlling biomass dynamics were not calibrated, ROSnB and BOS, which resulted in a significant overestimation of biomass (bias = 2.5 t ha-1) (Table 8, Figure 3a,dError! Reference source not found.). It is noteworthy from Table 8 that a more accurate characterization of biomass has minimal impact on other simulated characteristics, such as crop growth stage simulation, GAI and yields. In the DBValid plots, the results closely resembled those obtained with DBOpti. Although the rmse increased slightly, the bias was nearlyzero, showing that the calibrated parameters were effective over a wide range of environmental conditions.

Green leaf area. The simulations of green leaf area dynamics throughout the growing season using ROS on DBOpti plots showed a reduction in both rmse (0.9 m2 m-2) and bias (-0.1 m2 m-2) compared to those obtained with YR parameters. It can also be noted that forcing lev dates had minimal impact on the overall accuracy of GAI (Table 8). These results are comparable to previous calibration exercises (Prévot et al., 2003; Guillaume et al., 2011). The comparison between ROS simulations and in situ observations showed good estimations along the growing season, as presented in Figure 3b with a scatter plot around the 1:1 line, and demonstrated by the temporal dynamics shown in Figure 4a. a) b) Figure 4. STICS green leaf area daily simulations with YR (dashed grey line), ROS (solid black line) and BOS (dashed black line) plus GAIRS observations (dots) of a) one plot of the calibration set and b) one plot of the validation set. With BOS the prediction errors were higher than those obtained with ROS and ROSnB and comparable to those with YR (Table 8). The comparison between BOS simulations and in situ observations indicated a tendency to underestimate GAI, particularly when intermediate values were reached (Figure 3e). This feature is illustrated in Figure 4a, which shows the underestimation during the senescence phase. BOS simulations showed an overall advancement of the growing season compared to the simulations of the other optimisation scenarios. This might result from the approximation used to determine the phenology and its subsequent impact on STICS pheno phase

When considering DBValid plots, GAI results showed a reduction of the bias across the three calibration scenarios compared to YR, while all scenarios resulted in comparable rmse values (Table 8, Error! Reference source not found.b,e). In all cases, STICS performance on GAI was only slightly degraded for validation plots, even though the stress conditions encountered were much more significant than those in the calibrated plots. This is very encouraging for further STICS implementations. Yield. The evaluation of yield formation simulations using ROS, BOS and ROSnB on DBOpti yielded similar results to those obtained with YR (Table 8). However, significant differences were found in terms of bias, which can be attributed to the increase in the maximum yield potential threshold of the variety under consideration. This increase is highlighted by the strong negative bias observed with YR. Figure 3c shows that the observed yield variability was not well simulated by STICS. The factors explaining such variability may not be fully accounted for in the observed STICS inputs (e.g. local variations in irrigation doses) or modelled by STICS (crop health). Regarding the impact of lev, simulations with both forced and non-forced lev showed no impact. With ROSnB, the prediction errors were slightly higher than with ROS. Ruget et al. (2002) conducted an intermodule sensitivity analysis of STICS and reported that all parameters of the shoot production modules have an effect on yield. However, there was no impact on yield when considering DBValid, which tends to temper such a statement. In contrast, BOS showed slightly better results than ROS, with rmse = 1.0 t ha-1 and bias = 0.5 658 t ha-1, indicating the benefits of simulating earlier drp dates and a longer grain filling period, stdrpmat. The evaluation of yield simulations on DBValid significantly increased the rmse to around 1.7 661 t ha-1, regardless of the optimisation scenario considered, and these results were still better than those of YR. In Figure 3c and Figure 3f, the three circled points of DBValid1 data set (comprising intensively cultivated plots) show a strong bias. These points correspond to a fertilisation trial where the final nitrogen dressing was omitted. STICS was very sensitive to such a reduction in N fertilization, while the actual impacts were moderate. This illustrates the effect of inadequately described environmental properties (soil depth, water storage capacity, organic matter, initial conditions) that hampers the capacity of the model to accurately determine stress.

4. Discussion : In this study, three calibration scenarios were proposed to determine STICS plant parameters of a spring wheat variety. The calibration scenarios were based on assumptions that need to be discussed, such as the correction function used to account for the ears in STICS green leaf area simulations, or the use of within-field yield maps which are often unavailable in operational contexts. Furthermore, the calibration strategies yielded notably different parameter sets, raising questions about the potential divergence of STICS simulations under climatic conditions that differ from the calibration growing conditions.

4.1. STICS green leaf area adaptation incorporating ear contribution Green leaf area as derived from remote sensing incorporates all the green contributors of the canopy (Weiss and Baret, 2016), which in the case of wheat, includes different organs like leaves, stems and ears. Ears, which develop above the canopy, contribute significantly to the green area even though they are not considered in STICS simulations (Weiss et al., 2001). The proposed EAI module is driven by parameters such as eaimax, stmax, stplat and stsen, whose values were obtained from measurements by Casals (1996) on a different variety from the one considered in this study. EAI simulations modify the GAI simulations of STICS and may impact the calibration process. This was evaluated with the ROS scenario. Initially, the prediction errors (rmse and bias) were computed by comparing STICS green leaf area simulations obtained with a wide range of values for the EAI module parameters (Table S3 in Section3 of SM) with the GAIRS observations in DBOpti. Among the EAI parameters, eaimax presented the strongest impact on GAI simulations (Figure S4 in Section3 of SM). Surprisingly, the best results were obtained when the EAI correction was minimised or absent. The STICS simulations were almost unbiased when the addition of the ears was not considered, eaimax = 0. 691 Drawing from these findings, it can be inferred that the integration of the EAI module into STICS is not necessary to reproduce the GAIRS as observed by remote sensors. However, this might be a consequence of dense vegetation cover with a saturation of the remote sensing signal, as reported by Weiss and Baret (2016) who showed that GAI product derived from Sentinel2 has uncertainties with GAI values >6 m2 m-2. Additionally, Li et al. (2015) reported that the spectral reflectance of the canopy in the near-infrared (760 - 1150 nm) showed a limited response to spike removal under conditions of excessive nitrogen content. This is consistent with the selection criteria applied to the DBOpti plots, which were the highest GAIRS values and the highest nitrogen fertilisation rates. GAIRS dynamics of the DBOpti plots reached their maximum at the flag leaf with values of around 7.5 m2 m-2, and remained constant and insensitive to full ear development (Figure S5 in Section3 of SM). In contrast, DBValid plots, showed maximum GAIRS at flowering with values ≤ 6.5 m2 m-2. Consequently, the EAI module is not necessary when GAIRS values are over 6.5 m2 m-2 and highly fertilised because it introduces an artefact not expressed by the GAIRS. However, the EAI module improves the STICS simulations when GAIRS values < 6.5 m2 m-2 and considering eaimax values between 0.5 – 1.0 m2 m-2 (Figure S5 in Section3 of SM). Finally, to assess the impact of EAI parameters on the calibration, we replicated the ROS calibration approach under three scenarios of EAI parameterisation: No ear (ST0), low ear contribution (ST1), and high ear contribution (ST2) (Table S4 in Section3 of SM). The calibrated parameters were comparable with variations of about 1%, which are considerably smaller than the differences observed between the ROS and BOS scenarios (Table S5 in Section3 of SM). These differences resulted in minimal changes in terms of prediction errors, remaining within the range of errors found with ROS in DBOpti (rmse: 0.9 m2 m-2, bias: 0.4 m2 m-2). Furthermore, predictions errors were slightly lower with ST0 and ST1 (rmse: 0.8 m2 m-2, bias: 0.3 m2 m-2).

4.2. Impact of the source of data for yield optimisation: local yield vs. field average yield 716 Local yield measurements with the spatial resolution of GAIRS observations are not consistently available. An alternative approach could be to use average field yields, commonly known to farmers, in the calibration process, even though it may impact the calibration results. To explore this, step6 of the ROS approach was performed using the average field yield reported by farmers instead of using the plots sampled values. A lower range of values was then considered: from 9.0 – 11.8 t ha-1 (plots) to 10.0 – 10.6 t ha-1 (field average). The calibrated parameters showed slightly lower values for cgrain, 0.053, compared to that obtained in ROS, 0.056, while no change was observed for vitircarb, maintaining a value of 0.014. The prediction errors for yield simulations in BDValid using the new parameters showed an rmse of 1.6 t ha-1 and bias of 0.6 t ha-1, comparable to those obtained with ROS, 1.7 t ha-1 and 0.7 t ha-1 respectively. Therefore, it can be concluded that the average field yield commonly known by farmers can be used in the calibration process, making the adaptation of STICS to a variety more 7 accessible.

4.3. Equifinality of ROS and BOS plant parameters ROS and BOS calibration led to significantly different parameter sets, although the simulations for green leaf area and yield simulations were quite comparable. This brings up the concept of equifinality, where multiple parameter combinations can produce similar model outputs. One may wonder to what extent STICS simulations might diverge under climatic conditions other than those considered in this study. Addressing this question is of great importance for climate change studies. The aim is to determine whether the differences in plant parameters induced by the calibration strategy generate comparable results whatever the climatic conditions, or whether, on the contrary, the simulations can diverge significantly under certain climatic conditions. For this purpose, the discrepancies between ROS and BOS simulations for both green leaf and yield were examined for the climate period 2002 – 2020, considering DBOpti and DBValid plots together (Figure 5). a) b)

Figure 5 Interannual variability of rmse (solid line) and bias (dashed line) over the period 2002 – 2020 of: a) green leaf area, and b) yield. Grey horizontal lines present the average values of the values of rmse and bias. On average, the differences in GAI simulations showed little bias of -0.11 m2 m-2 and an rmse of 0.73 m2 m-2. Although the differences are significant, they remain lower than the prediction errors computed against observed GAIRS (Table 8). However, the interannual variability is small compared to the average differences (Figure 5a), thus both plant files can be used to study the behaviour of GAI under different climatic conditions. Regarding the simulations of yield, the differences using ROS and BOS parameters showed an rmse of 0.92 t ha-1 and a bias of -0.45 t ha-1. These values are comparable to those obtained after the calibration processes (Table 8) and confirm the bias obtained on yields with the BOS scenario. The interannual variability of bias displayed in Figure 5b is large, with values ranging from -1 t ha-1 in 2006 to 0 in 2013. The processes that led to such differences are complex and difficult to analyse. However, the lowest temperatures during the grain filling phase were in 2013, which may have levelled out the differences in the yield simulated by each scenario. Thus, the differences in parameter obtained for ROS and BOS have an impact on yields under atypical temperatures.

5. Conclusions: The methodology proposed for the calibration of a wheat variety in STICS, based on farmers' knowledge and remote sensing information, gave satisfactory results. The advantage of using remote sensing data is twofold: on the one hand, it overcomes the difficulties of access to data and, on the other hand, it ensures consistency between the calibration data and the data that can then be used in assimilation schemes, thus minimising the effects of differences in determination methods, such as in the

determination of leaf area data (remote sensing versus in situ). To mitigate the impact of inaccurately described environmental crop growing conditions on the calibrated parameters, such as soil properties like soil water content, which could lead to the simulation of uncontrolled stress conditions, we performed the calibration on well-irrigated and well-fertilised plots. This was possible in the area of La Macha (Spain), where we found intensively irrigated wheat fields. The implementation of STICS using the calibrated plant parameters was successfully performed under different environmental conditions spanning a wide range of water and nitrogen stresses. This means that the impact of stress on plant development and yield is well addressed by the STICS formalism once the plant parameters are calibrated. The obtained accuracy (rmse) in GAI ranged between 1.2 and 1.5 m2 m-2 and in yield between 1.7 and 1.8 t ha-1 t ha-1, across the optimisation scenarios considered. These values include uncertainties in the in-plant parameterisation of those parameters not calibrated, in the STICS formulation, and in the description of the environmental crop growing conditions (soil and climate). These results represent an improvement compared to the simulations obtained with the Yecora Rojo plant parameters, with a smaller bias, which is the closest variety already calibrated using an extensive field data set. In this study, some technical issues concerning phenology, biomass and GAI were also considered. Firstly, we compared two optimisation scenarios, with and without phenological field observations. When phenological observations were not available, we proposed to use PVIs that were derived from analytical properties of the GAIRS time series and were therefore unambiguously defined. Some shifts were observed compared to the expected STICS key growth stages, leading to differences in the calibrated parameters. However, the accuracy of the simulations of yield and GAI were comparable, being also consistent between them over a wide range of climatic conditions encountered over the last 20 years. Regarding biomass, while the calibration of the parameter governing biomass dynamics proved necessary to accurately simulate biomass (strong biases were found when it was not included in the calibration process), not including it had low impact on the other variables such as GAI, yield and growth stages. Finally, we considered an empirical module for estimating the EAI. We showed that with well-developed crop canopies, such as those used for calibration, the inclusion of the EAI module had little impact on the GAI simulations and the calibration process. However, when canopies were less developed (maximum GAI <6.5 m2 m-2), the inclusion of the ears in GAI simulations improved the agreement between STICS and remote sensing observations. This is essential for assimilating remotely sensed data into crop models, highlighting the importance of obtaining data sets on the contribution of ears, which remain scarce at present.