Parameter optimisation for a better representation of drought by LSMs: inverse modelling vs. sequential data assimilation

Hélène Dewaele, Simon Munier, Clément Albergel, Carole Planque, Nabil Laanaia, Dominique Carrer, and Jean-Christophe Calvet

CNRM – UMR3589 (Météo-France, CNRS), Toulouse, 31057, France

Correspondence to: Jean-Christophe Calvet (jean-christophe.calvet@meteo.fr)

Received: 2 March 2017 – Discussion started: 17 March 2017
Revised: 16 August 2017 – Accepted: 18 August 2017 – Published: 28 September 2017

Abstract. Soil maximum available water content (MaxAWC) is a key parameter in land surface models (LSMs). However, being difficult to measure, this parameter is usually uncertain. This study assesses the feasibility of using a 15-year (1999–2013) time series of satellite-derived low-resolution observations of leaf area index (LAI) to estimate MaxAWC for rainfed croplands over France. LAI interannual variability is simulated using the CO$_2$-responsive version of the Interactions between Soil, Biosphere and Atmosphere (ISBA) LSM for various values of MaxAWC. Optimal value is then selected by using (1) a simple inverse modelling technique, comparing simulated and observed LAI and (2) a more complex method consisting in integrating observed LAI in ISBA through a land data assimilation system (LDAS) and minimising LAI analysis increments. The evaluation of the MaxAWC estimates from both methods is done using simulated annual maximum above-ground biomass ($B_{ag}$) and straw cereal grain yield (GY) values from the Agreste French agricultural statistics portal, for 45 administrative units presenting a high proportion of straw cereals. Significant correlations ($p$ value < 0.01) between $B_{ag}$ and GY are found for up to 36 and 53 % of the administrative units for the inverse modelling and LDAS tuning methods, respectively. It is found that the LDAS tuning experiment gives more realistic values of MaxAWC and maximum $B_{ag}$ than the inverse modelling experiment. Using undisaggregated LAI observations leads to an underestimation of MaxAWC and maximum $B_{ag}$ in both experiments. Median annual maximum values of disaggregated LAI observations are found to correlate very well with MaxAWC.

1 Introduction

Extreme weather conditions markedly affect agricultural production. The interannual variability of rainfed crop yields is driven to a large extent by the climate variability. Comparing agricultural statistics to climate data shows the impact of atmospheric conditions on vegetation production. For example, lower temperature in northern Europe tends to shorten the period of crop growth. Conversely, persistent high temperatures as well as droughts in southern Europe are linked to negative anomalies of crop yields (Olesen et al., 2011). Li et al. (2010) showed that air temperature tends to influence mean crop yields at small scales (400 to 600 km), whereas rainfall drives crop yields at larger scales (50 to 300 km). Capa-Morocho et al. (2014) also showed the influence of air temperature on crop yields. They established a link between temperature anomalies related to the El Niño phenomenon and potential crop yield anomalies, obtained from reanalysis data and crop model, respectively.

In the context of climate change and of natural climate variability, there is a need to monitor the impacts of droughts on crops and water resources at continental and global scales (Quiroga et al., 2011; Van der Velde et al., 2012; Crow et al., 2012; Bastos et al., 2014). Modelling of continental surfaces into atmospheric and hydrological models has evolved in recent decades towards land surface models (LSMs) able to simulate the coupling of the water, energy and carbon cycles (Calvet et al., 1998; Krinner et al., 2005; Gibelin et al., 2006). In particular, LSMs are now able to simulate photosynthesis and plant growth. A major source of uncertainty in both LSMs and crop models is the maximum available water content of the soil (MaxAWC). This quantity represents the amount of water stored in the soil available for plant transpiration.
ration along the vegetation growing cycle (Portoghese et al., 2008; Piedallu et al., 2011). MaxAWC is constrained by soil parameters and by the plant rooting depth. In regions vulnerable to drought risk exposure, MaxAWC is a key driver of the plant response to the climate variability.

Soil characteristics influence the vegetation response to climate (Folberth et al., 2016). In a model benchmarking study (Eitzinger et al., 2004) simulated evapotranspiration, soil moisture and biomass were compared with observations. Eitzinger et al. used three crop models that differ in the representation of the available soil water content (AWC): WOFOST (WOrld Food Studies model; Van Diepen et al., 1989), CERES (Crop Environment REsource Synthesis model; Ritchie and Otter, 1985) and SWAP (Statewide Agricultural Production model: van Dam et al., 1997). They showed that a better description of rooting depth and evapotranspiration, taking account of soil type and crop type, could significantly improve these models. Tanaka et al. (2004), Portoghese et al. (2008) and Piedallu et al. (2011) have highlighted the important role of the soil characteristics (soil texture, rooting depth) on MaxAWC. Soylu et al. (2011) and Wang et al. (2012) illustrated the major impact of MaxAWC on evapotranspiration. While soil properties such as soil texture determine the soil water holding capacity (in kg m\(^{-2}\)) information on rooting depth is needed to determine MaxAWC (in kg m\(^{-2}\)). A better representation of MaxAWC could improve the simulated interannual variability of both water fluxes and vegetation biomass by LSMs.

The lack of in situ observations of MaxAWC to calibrate and assess LSMs limits the ability of LSMs to represent drought effects on plants. Using satellite observations and data assimilation techniques could be a solution to this problem. A list of atmospheric, oceanic and terrestrial essential climate variables (ECVs) which can be monitored at a global scale from remote sensing observations was proposed by the Global Climate Observing System (GCOS). Leaf area index (LAI), fraction of absorbed photosynthetically active radiation (FAPAR) and soil moisture are key ECVs for land surface modelling. The use of these satellite-derived products to verify LSM simulations or to optimise key LSM parameters has been assessed by several authors (e.g. Becker-Reshef et al., 2010; Crow et al., 2012; Ferrant et al., 2014; Ford et al., 2014; Ghilain et al., 2012; Ichii et al., 2009; Kowalik et al., 2014; Szczypta et al., 2012, 2014). Data assimilation is a field of active research. Data assimilation techniques allow the integration of different observation types (e.g. in situ or satellite derived) into LSMs in order to optimally combine them with model outputs: the correction applied to the model state is called the increment and the corrected model state is the analysis. Previous works have studied the impact of assimilation of LAI observations and found that it can significantly improve the representation of vegetation growth (e.g. Albergel et al., 2010; Barbu et al., 2011, 2014).

The Interactions between Soil, Biosphere and Atmosphere (ISBA) LSM includes a modelling option able to simulate photosynthesis and plant growth (Calvet et al., 1998; Gibelin et al., 2006). ISBA produces consistent surface energy, water and carbon fluxes, together with key vegetation variables such as LAI and the living above-ground biomass (\(B_{\text{ag}}\)). Previous studies showed that this model can represent well the interannual variability of \(B_{\text{ag}}\) over grassland and straw cereal sites in France provided MaxAWC values are tuned (Calvet et al., 2012; Canal et al., 2014). In these studies, MaxAWC for straw cereals was retrieved by maximising the correlation coefficient between simulated annual maximum \(B_{\text{ag}}\) (\(B_{\text{agX}}\)) and grain yield (GY) observations. The MaxAWC values were obtained for 45 French administrative units (‘départements’) representing a large proportion of rainfed straw cereals. For grasslands, dry matter yield observations were used. Significant correlations were found between the simulated \(B_{\text{agX}}\) value of grassland and dry matter yield of grasslands for up to 90% of the administrative units. On the other hand, only 27% of the 45 straw cereals départements (i.e. only 12 départements) had significant correlations. A possible cause of the difficulty to simulate the interannual variability of straw cereals’ GY was that the standard deviation of GY represented less than 10% of the mean GY. This was a relatively weak signal. For grasslands a much larger value, of about 30% of the mean dry matter yield, was observed (Canal et al., 2014).

The main purpose of this study is to estimate MaxAWC for straw cereals using reverse modelling techniques based on satellite-derived LAI observations disaggregated over separate vegetation types. Simulated and observed LAI are compared for a 15-year period (1999–2013) over the same 45 agricultural spots used in the previous studies of Calvet et al. (2012) and Canal et al. (2014). We use LAI observations instead of GY to estimate MaxAWC. The GY observations are used to verify the interannual variability of the simulated \(B_{\text{agX}}\). This can be considered as an indirect validation of the retrieved MaxAWC. In a first experiment, we use a simple inverse modelling technique to estimate MaxAWC together with the mass-based leaf nitrogen content, minimising a cost function based on observed and simulated LAI values. In another experiment, we use a land data assimilation system (LDAS) able to sequentially assimilate LAI observations. In this case, MaxAWC solely is retrieved by minimising the LAI analysis increments. In the following, these two experiments are referred to as inverse modelling and LDAS tuning, respectively.

The main goals of this study are to (1) assess the usefulness of integrating satellite-derived LAI observations into a LSM, (2) compare inverse modelling and LDAS tuning and (3) determine MaxAWC values.

The observation data sets are described in Sect. 2, together with the version of ISBA used in this study and the LDAS. Results obtained from both methods are presented in Sect. 3, and analysed and discussed in Sect. 4. Conclusions and prospects are given in Sect. 5.
Figure 1. Straw cereal sites (35 km × 35 km) in France in 45 administrative units (“départements”). Triangle and diamond symbols show the départements presenting a significant temporal correlation ($R^2 > 0.41$, $F$-test $p$ value < 0.01) between Agreste GY values and LAI$_{\text{max}}$ (blue circles), both inverse modelling and LDAS tuning (red diamonds), LDAS tuning only (yellow upward-facing triangle) and inverse modelling only (green downward-facing triangle). The “×” symbol indicates départements where no significant correlation between biomass simulations and GY could be found.

2 Material and methods

The forcing and validation observations used in this study over the 1999–2013 period are described below. The location of the considered straw cereal spots is presented in Fig. 1.

2.1 Satellite LAI product

We use the GEOV1 global LAI product (Baret et al., 2013) provided in near real time (every 10 days) at a spatial resolution of 1 km × 1 km by the European Copernicus Global Land Service (http://land.copernicus.eu/global/). The GEOV1 LAI product is derived from SPOT-VGT satellite observations starting in 1999. The complete 1999–2013 LAI time series comes from SPOT-VGT and is fully homogeneous. The product is well evaluated against ground observations (see the Supplement). The GEOV1 product is a low-resolution product (1 km × 1 km). At this spatial scale, it is not possible to isolate pure straw cereal pixels and it is preferable to disaggregate the LAI (i.e. compute the LAI of each vegetation type) before integrating it into a straw cereal model. We disaggregated the GEOV1 LAI data following the method developed by Carrer et al. (2014), based on a Kalman filtering technique. This method permits separating the individual LAI of different vegetation types that co-exist in a grid pixel and then provides dynamic estimates of LAI for each type of vegetation within the pixel (Munier et al., 2017). The Kalman filter optimally combines satellite LAI data and prior information from the ECOCLIMAP land cover database (Faroux et al., 2007; Masson et al., 2003). ECOCLIMAP prescribes physiographic parameters (fractional vegetation cover, soil depth, etc.) for several vegetation types including grasslands, forests and C3 crops like straw cereals. Mean annual LAI cycles per vegetation type from ECOCLIMAP are used as a first guess to partition the GEOV1 LAI every time a new satellite observation is available.

2.2 Atmospheric forcing

The global WFDEI data set (Weedon et al., 2014) is used in this study to drive the ISBA simulations. It provides 3-hourly surface atmospheric variables on a 0.5° × 0.5° grid: air temperature, air humidity, wind speed, atmospheric pressure, solid and liquid precipitation, incoming shortwave and longwave radiation. WFDEI is based on the ERA-Interim atmospheric reanalysis (Dee et al., 2011). It includes elevation corrections and seasonal monthly bias corrections from ground-based observations.

2.3 Agricultural GY statistics

The Agreste portal (http://agreste.agriculture.gouv.fr/) provides annual statistical surveys over France which make it possible to establish a database of yearly GY values. The GY estimates are available per crop type and per administrative unit (département). We use GY values for rainfed straw cereals such as barley, oat, rye, triticale and wheat, for the same 45 départements as in Calvet et al. (2012) and Canal et al. (2014). Calvet et al. (2012) and Canal et al. (2014) used Agreste data for the 1994–2008 and 1994–2010 periods, respectively, for both straw cereals and fodder production. We use Agreste data from 1999 to 2013, only for straw cereal GY.

2.4 The ISBA model

The ISBA LSM is included in the SURFEX (SURFace EXternalisée) modelling platform (Masson et al., 2013). The newest version of SURFEX (version 8) is used in this study with the “NIT” biomass option for ISBA. The “C3 crop” plant functional type is considered.

ISBA simulates the diurnal course of heat, water and CO$_2$ fluxes, including gross primary production (GPP). The set of ISBA options we use permits the simulation of LAI and $B_{ag}$ on a daily basis (Calvet et al., 1998, 2008). The model includes a soil moisture stress function ($F_s$) applied to photosynthesis key parameters. For low vegetation such as grass or crops, the parameters related to soil moisture stress (Calvet, 2000) are the mesophyll conductance ($g_m$) and the maximum leaf-to-air saturation deficit ($D_{\text{max}}$). Values of $g_m$ and $D_{\text{max}}$ for straw cereals in well-watered conditions are given in Ta-
ble S1 (in the Supplement), together with other model parameters. It must be noted that this value of $g_m$ was derived using inverse modelling by Canal et al. (2014) for the same straw cereal sites as those considered in this study. In moderately dry conditions, $g_m$ and $D_{max}$ are affected by $F_a$ in such a way as to increase the intrinsic water use efficiency (WUE). This corresponds to a drought-avoiding behaviour (Calvet, 2000). The model is also able to represent a drought-tolerant behaviour (stable or decreasing WUE and Calvet et al. (2012) showed that straw cereals tend to behave as drought avoiding, while grasslands tend to behave as drought tolerant.

The above-ground biomass ($B_{ag}$) consists of two components within ISBA: the structural biomass and the active biomass. The latter corresponds to the photosynthetically active leaves and is related to $B_{ag}$ by a nitrogen dilution allometric logarithmic law (Calvet and Soussana, 2001). The mass-based leaf nitrogen concentration ($N_L$) is a parameter of the model affecting the specific leaf area (SLA), which is the ratio of LAI to leaf biomass (in m$^2$ kg$^{-1}$). The SLA depends on $N_L$ and on plasticity parameters (Gibelín et al., 2006). The $N_L$ parameter is key for LAI simulations and has to be included in any inverse modelling experiment involving LAI.

The net assimilation of CO$_2$ by the leaves ($A_n$) is driven by environmental factors such as the atmospheric CO$_2$ concentration, air humidity, the incoming solar radiation and the leaf surface temperature. To upscale the net assimilation of CO$_2$ and transpiration at the vegetation level, a multilayer radiative transfer scheme is used (Carrer et al., 2013). The daily canopy-scale accumulated value of $A_n$ serves as an input for the vegetation growth and mortality sub-models, and the phenology is completely driven by $A_n$ (no growing degree-day parameterisation is used).

The plant transpiration flux is used to calculate the soil water budget through the root water uptake. The soil hydrology scheme used in this study is referred to as “FR-2L” in SURFEX. It represents two soil layers: a thin surface layer with a uniform depth of 1 cm and a root-zone layer of depth $Z_r$. The latter is used as a surrogate for MaxAWC in the calibration process. Soil texture parameters such as the gravimetric fraction of sand and clay are extracted from the Harmonized World Soil Database (Nachtergaele et al., 2012). Physical soil parameters such as volumetric soil moisture at field capacity ($\theta_{FC}$) and wilting point ($\theta_{Wilt}$) are calculated thanks to pedotransfer functions based on soil texture. The MaxAWC parameter is given by

$$\text{MaxAWC} = \rho (\theta_{FC} - \theta_{Wilt}) \times Z_r. \quad (1)$$

Parameters are defined in the Appendix and model parameter values are summarised in Table S1 in the Supplement.

### 2.5 Land data assimilation system

We used the LDAS described in Barbu et al. (2011, 2014). It consists of a sequential data assimilation system operated offline (uncoupled with the atmosphere). The assimilation is based on a simplified extended Kalman filter (SEKF), able to integrate observations such as LAI and soil moisture in the ISBA model. In this study, only LAI observations are assimilated and the LDAS produces analysed LAI values.

The key update equation of the SEKF is

$$\Delta x = x_f - x_i = K (y^a - y^i), \quad (2)$$

where $\Delta x$ is the analysis increment, $x$ is a control vector of one dimension representing LAI values propagated by the ISBA LSM, and $y_a$ is the observation vector representing the GEOV1 LAI observations. The $t$ superscript stands for time. The initial time is denoted by the 0 superscript. The “a”, “f” and “o” subscripts denote analysis, forecast and observation, respectively. The $y^i$ term of Eq. (2) represents the model value at the analysis time:

$$y^i = h(x^0), \quad (3)$$

where $h$ is the observation operator and

$$K = BH^T (BH^T + R)^{-1}. \quad (4)$$

The Kalman gain is derived from the background error covariance matrix $B$ and from the observation error covariance matrix $R$. Matrix $H$ that appears in Eq. (4) represents the Jacobian of potentially nonlinear $h$ function:

$$H = \frac{\partial y^i}{\partial x^0} \quad (5)$$

which gives the following Jacobian matrix:

$$H = \left( \frac{\partial \text{LAI}^i}{\partial \text{LAI}^0} \right). \quad (6)$$

The initial state at the beginning of an assimilation window is analysed via the information provided by an observation at the end of the assimilation window (Rüdiger et al., 2010). In this approach, the LAI increments (Eq. 2) are applied at the end of 1-day assimilation intervals. The elements of the Jacobian matrix are estimated by finite differences, individually perturbing each components of the control vector $x$ by a small amount $\delta x$:

$$H = \frac{y(x + \delta x) - y(x)}{\delta x}. \quad (7)$$

The background error covariance matrix $B$ is assumed to be constant at the start of each analysis cycle. The covariance matrices $B$ and $R$ are assumed to be diagonal. In the simplified version of the EKF used in this study, namely SEKF, the $B$ matrix does not evolve with time. The standard deviation of errors of GEOV1 LAI is assumed to be 20 % of GEOV1 LAI. The same assumption is made for the standard deviation of errors of the modelled LAI (20 % of modelled LAI) for modelled LAI values higher than 2 m$^2$ m$^{-2}$. For modelled LAI
values lower than 2 m² m⁻², a constant error of 0.4 m² m⁻² is assumed. This assumption is based on option 3 presented in Barbú et al. (2011). They showed that this option gives the best simulated LAI over an instrumented grassland site in southwestern France.

2.6 Upscaling disaggregated LAI observations to département level

Each agricultural spot shown in Fig. 1 corresponds to the area within a département presenting the highest fraction of straw cereals. These 45 locations were chosen by Calvet et al. (2012) on a 8 km × 8 km grid using fractions of vegetation types derived from ECOCLIMAP (Faroux et al., 2013). Disaggregated LAI observations have a spatial resolution of 1 km × 1 km. This represents a small area compared to the size of a département (from 2000 to 10 000 km²). Local values of the straw cereal LAI may not be representative of the straw cereal production at the département level described by Agreste. Preliminary tests showed that averaging the disaggregated LAI on the same 8 km × 8 km grid cell used by Calvet et al. (2012) was not sufficient to represent the interannual variability of the GY observations at the département level. Therefore, an analysis of the consistency of the two observation data sets (in situ GY and disaggregated satellite LAI) is performed. The average maximum annual value of the disaggregated GEOV1 LAI observation (LAIo_max) is calculated for various grid cell sizes for this task. In practice, the LAI_o_max value corresponds to the mean LAI values above a given fraction of the observed maximum annual LAI (LAI_max). We consider five grid cell sizes of 5 km × 5 km, 15 km × 15 km, 25 km × 25 km, 35 km × 35 km and 45 km × 45 km (from 25 to 2025 km²). The five LAI_o_max time series are compared with the GY time series for each département. The area size corresponding to the largest number of départements presenting a significant correlation between LAI_o_max and GY is selected.

2.7 Model calibration/validation

The feasibility of retrieving MaxAWC from LAI satellite data is explored using two different approaches: inverse modelling and LDAS tuning. For the two approaches, this calibration step is followed by a validation step aiming at demonstrating the relevance of the retrieved MaxAWC values and the added value of the retrieval technique.

The satellite LAI observations are available year-round but the sensitivity of straw cereal LAI to MaxAWC may change greatly from one period of the year to another. Prior to calibrating the model, a sensitivity study of the time window used for the MaxAWC retrieval is performed. Three periods are considered: (1) growing period (from 1 March to the date of the observed LAI_max); (2) peak LAI (period for which observed LAI is higher than 50% of observed LAI_max); (3) senescence (from the date when observed LAI_max is reached to 31 July). The ISBA simulations are stopped on 31 July as this date corresponds to the maximum harvest date at most locations.

The validation of the calibrated model consists of comparing the interannual variability of the simulated maximum annual above-ground biomass to the interannual variability of the GY observations. The 1999–2013 period is considered. In drought conditions, modelled Bag can rise to a maximum value and then drop rapidly. Therefore the peak Bag can be dependent on modelling uncertainties and on uncertainties in the atmospheric forcing. In order to limit the impact of model errors, an average value of the simulated Bag is used instead of an instantaneous value. This average value is calculated using all the Bag values above a threshold corresponding to 90% of the maximum annual Bag. It was checked that this threshold value permits the maximisation of the number of départements presenting a significant correlation with GY. Then, scaled anomalies of the average simulated Bag are compared with scaled anomalies of the GY observations, and the R² score is calculated. Scaled anomalies (A) are calculated using the mean and standard deviation of the two variables over the 1999–2013 period:

\[
A_{bag} = \frac{(B_{ag,X} - B_{ag})}{\sigma(B_{ag,X})}, \quad (8)
\]

\[
A_{gy} = \frac{(GY - \mu)}{\sigma(GY)} \quad (9)
\]

The interannual variability of the modelled LAI_max is assessed using the coefficient of variation (CV). The CV is given in % and is calculated according the following formula with σ the standard deviation and μ the mean:

\[
CV = 100 \times \frac{\sigma}{\mu}. \quad (10)
\]

The MaxAWC retrieval is considered to be successful if the Pearson correlation is significant at 1% level (F-test p value < 0.01).

2.8 Design of the experiments

2.8.1 Inverse modelling

Two parameters are estimated: NL and MaxAWC. For a given value of NL, a set of 13 LAI simulations is produced, corresponding to the following MaxAWC values: 44, 55, 66, 77, 88, 99, 110, 121, 132, 154, 176, 198 and 220 mm. Since NL is a key parameter for LAI simulations, it has to be retrieved together with MaxAWC and this simulation process is repeated 5 times, for the following NL values: 1.05, 1.30, 1.55, 2.05 and 2.55 %.

The LAI root mean squared error (RMSE) over the period between the occurrence of the observed LAI_max and 31 July for the 15 years is used to select the best simulation. The
MaxAWC used in the simulation with the lowest RMSE is selected as the optimal one:

\[
\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=0}^{n} (\text{LAI}_i - \text{LAI}_o)^2},
\]

where \(\text{LAI}\) is for simulated LAI, \(\text{LAI}_o\) is for observed LAI and \(n\) is the length of the data vector.

### 2.8.2 LDAS tuning

The LAI observations are integrated into ISBA by the LDAS. The LDAS produces analysed values of LAI and \(B_{ag}\). Therefore, there is no need to estimate \(N_L\) and the only degree of freedom in this case is the value of MaxAWC. Thirteen analyses are made, corresponding to the same MaxAWC values used in the inverse modelling experiment (Sect. 2.8.1).

The median analysis increment (Eq. 2) can present positive or negative values. Small corrections provided by the LDAS indicate that simulation outputs are close to observations and that the dynamics is well represented. The value closest to zero indicates the best simulation and the corresponding MaxAWC value is considered as the retrieved MaxAWC.

### 3 Results

#### 3.1 Disaggregated satellite LAI vs. grain yield observations

In a first step before integrating the disaggregated LAI observations into the ISBA model, we checked the consistency of the interannual variability of \(\text{LAI}_{\text{omax}}\) (Sect. 2.6) with that of the observed GY from Agreste. We investigated several values of the size of the area around each site coordinate to calculate the average of \(\text{LAI}_{\text{omax}}\), from 25 to 2025 km\(^2\). Individual \(\text{LAI}_{\text{omax}}\) values at a spatial resolution of 1 km \(\times\) 1 km correspond to the mean of LAI values above the \(\text{LAI}_{\text{omax}}\) threshold (Sect. 2.6). Several \(\text{LAI}_{\text{omax}}\) threshold values ranging from 40 to 95% of \(\text{LAI}_{\text{max}}\) were investigated together with the grid cell size (see Fig. S1 in Supplement). A \(\text{LAI}_{\text{omax}}\) threshold value of 50% and a grid cell size of 35 km \(\times\) 35 km (1225 km\(^2\)) were selected. In this configuration, a significant temporal correlation (\(F\)-test \(p\) value < 0.01) between the average \(\text{LAI}_{\text{omax}}\) and the observed GY is obtained for 31 départements. The latter are shown in Fig. 1 (empty blue circles). The 45 grid cells of 35 km \(\times\) 35 km are further used to calculate average 10-day LAI observations to be integrated in the ISBA model through either inverse modelling or LDAS tuning. The fraction of straw cereals derived from ECOCLIMAP for these grid cells ranges from 15 to 100%, with a median value of 68% (see Table S2 in Supplement).

The temporal correlation between \(\text{LAI}_{\text{omax}}\) and GY is illustrated in Fig. 2. The two 15-year time series correspond to average annual values of \(\text{LAI}_{\text{omax}}\) and GY across the 31 départements where \(\text{LAI}_{\text{omax}}\) is found to correlate with GY. The two time series present a very good correlation, with \(R^2 = 0.84\). This shows that the disaggregated satellite-derived LAI is able to capture the interannual variability of GY.

#### 3.2 Sensitivity study

Figure 3 presents the impact of ISBA parameters on the simulated annual maximum LAI and on its interannual variability. Two key parameters are considered: MaxAWC and \(N_L\). The same parameter values are applied to all 45 départements, and mean modelled \(\text{LAI}_{\text{omax}}\) are used to calculate CV values (Eq. 10). The CV values are shown in Fig. 3 as a function of these parameters, together with LAI\(_{\text{omax}}\).

It appears that the interannual variability of the modelled LAI\(_{\text{omax}}\) is governed by MaxAWC. The CV values of more than 12% are derived from the ISBA simulations at low values of MaxAWC (e.g. 50 mm). On the other hand, high MaxAWC values (> 200 mm) correspond to limited interannual variability of LAI\(_{\text{omax}}\) (CV < 4%), in relation to a lower sensitivity of plants to drought.

The \(N_L\) parameter has a limited impact on CV and its impact depends on MaxAWC. For large (small) MaxAWC values above (below) the standard average value of 132 mm used in ISBA, the largest values of \(N_L\) tend to cause a decrease (increase) of CV. In the inverse modelling experiment, \(N_L\) mainly impacts the average simulated LAI\(_{\text{max}}\) value. In the ISBA model, \(N_L\) is linearly related to the specific leaf area (SLA) and large \(N_L\) values correspond to large SLA values, i.e. larger LAI values for a given simulated leaf biomass (Gibelien et al., 2006). However, Fig. 3 shows that MaxAWC has a more pronounced impact than \(N_L\) on LAI\(_{\text{omax}}\).

Increasing MaxAWC from 50 to 250 mm triggers a rise in LAI\(_{\text{omax}}\) from about 2 m\(^2\) m\(^{-2}\) at low \(N_L\) values to 3 m\(^2\) m\(^{-2}\) at high \(N_L\) values.
Figure 3. Impact of MaxAWC and $N_L$ parameters on the annual maximum LAI simulated by ISBA and on its interannual variability. The interannual variability is quantified using the coefficient of variation (CV, in %). Mean CV values (across all 45 départements) are plotted as a function of (a) $N_L$ and (b) MaxAWC, for various values of MaxAWC and $N_L$, respectively. Mean LAI$_{max}$ values (across all 45 départements) are plotted as a function of (c) $N_L$ and (d) MaxAWC, for various values of MaxAWC and $N_L$, respectively. The dashed lines are obtained using standard average ISBA parameter values (MaxAWC = 132 mm and $N_L$ = 1.3 %).

at high $N_L$ values. Switching $N_L$ from low to high values at a given MaxAWC level also raises LAI$_{max}$, but not more than 2 m$^2$ m$^{-2}$.

This result confirms that MaxAWC is the key parameter to be retrieved in order to improve the representation of straw cereal biomass, for both inverse modelling and LDAS tuning experiments. The impact of MaxAWC on the cost functions (LAI RMSE and median LAI analysis increments, Eqs. 11 and 2, respectively) may depend on the LAI observation period. We tested the two retrieval methods for three different optimisation periods: start of growing period, peak LAI and senescence (see Sect. 2.7).

This is illustrated in Fig. 4, which shows the average cost function across all 45 départements. In both experiments, MaxAWC has little impact on the cost function during the start of the growing season. The most pronounced response of both LAI RMSE and analysis increments is observed during the senescence. For this period of the growing cycle, both cost functions present a minimum value at MaxAWC = 110 mm. Also, the largest RMSE and increment values are observed during senescence, indicating that the processes at stake during this period are more difficult to simulate. For straw cereals, senescence is related to soil moisture stress (Cabelguenne and Debaeke, 1998) and during this period the value of MaxAWC has a marked impact on the representation of the effect of drought by the model. The peak LAI period is less favourable to the integration of LAI observations into the model, with a reduced accuracy on the retrieved MaxAWC.

3.3 Outcomes of the optimisation

A direct result of the optimisation procedure is the reduction of the cost function value. This is illustrated in Fig. 5 for all 45 départements. Figure 5 presents the impact of the optimisation on the cost functions of inverse modelling and LDAS tuning during the senescence period: LAI RMSE and LDAS LAI increments, respectively.
The RMSE values are systematically reduced by the inverse modelling experiment. For all 45 départements, the median value of the LAI RMSE drops from 1.6 to 1.2 m² m⁻². While LAI RMSE exceeds 1.5 m² m⁻² for 29 départements before the optimisation, this RMSE value is exceeded for only three départements after the optimisation. It must be noted that this is a much better result than the RMSE obtained in Fig. 4 (1.6 m² m⁻²) for the cost function including all 45 départements, with a MaxAWC value of 110 mm. This shows the impact of the spatial variability of MaxAWC.

For LDAS tuning, most of the median daily increment values are sharply reduced: while 17 values are larger (smaller) than 0.2 (−0.2) m² m⁻² before the optimisation, all the values range from −0.1 to 0.1 m² m⁻² after the optimisation. The spatial median value of the LDAS LAI increments varies from −0.03 m² m⁻² for original LDAS to −0.01 m² m⁻² for LDAS tuning, for all 45 départements. Table 1 summarises results showing the impact of the optimisation on indicators such as the number of départements presenting a significant correlation of $B_{agX}$ with GY and the median value of the cost functions. Table 2 also gives median and standard deviation values of the retrieved MaxAWC and of the retrieved $N_L$ in the case of inverse modelling, together with modelled $B_{agX}$ and LAI$_{max}$ values. The results are given for the départements presenting a significant correlation of $B_{agX}$ with GY and for all 45 départements. An extended version of Table 1 (Table S3 in the Supplement) also gives results for undisaggregated LAI and for 15 validated départements for both inverse modelling and LDAS tuning.

In the case of LDAS tuning, the median retrieved MaxAWC (129 ± 44 mm for all 45 départements and
Table 1. Impact of the optimisation (either inverse modelling or LDAS tuning) on parameter values (spatial median values ± standard deviation) of the ISBA model (MaxAWC and $N_L$), on the median value of $B_{agx}$ and LAI$_{max}$, on peak simulated $B_{ag}$ and on the models scores during the senescence period with respect to the disaggregated LAI observations. The results are given for the validated départements, i.e. those presenting a significant correlation ($p$ value < 0.01) of $B_{agx}$ with Agreste straw cereal grain yield observations. Results for all 45 départements are given in brackets and in italics. Parameter values resulting from the optimisation are in bold. Because simulated LAI$_{max}$ varies from one year to another, spatial median values are based on median temporal values across the considered 15-year period.

<table>
<thead>
<tr>
<th></th>
<th>Basic ISBA</th>
<th>Inverse modelling</th>
<th>Original LDAS</th>
<th>LDAS tuning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of départements presenting significant positive correlations ($p$ value &lt; 0.01)</td>
<td>18</td>
<td>16</td>
<td>21</td>
<td>24</td>
</tr>
<tr>
<td>MaxAWC (mm)</td>
<td>132 ± 2</td>
<td>153 ± 40</td>
<td>132 ± 2</td>
<td>133 ± 46</td>
</tr>
<tr>
<td>(132 ± 2)</td>
<td>(111 ± 44)</td>
<td>(132 ± 2)</td>
<td>(129 ± 44)</td>
<td></td>
</tr>
<tr>
<td>$N_L$ (%)</td>
<td>1.30</td>
<td>1.05 ± 0.20</td>
<td>1.30</td>
<td>1.30</td>
</tr>
<tr>
<td>(constant value)</td>
<td>(constant value)</td>
<td>(constant value)</td>
<td>(constant value)</td>
<td></td>
</tr>
<tr>
<td>$B_{agx}$ (kg m$^{-2}$)</td>
<td>0.99 ± 0.05</td>
<td>0.96 ± 0.16</td>
<td>0.96 ± 0.07</td>
<td>0.98 ± 0.17</td>
</tr>
<tr>
<td>(1.01 ± 0.07)</td>
<td>(0.89 ± 0.16)</td>
<td>(0.93 ± 0.11)</td>
<td>(0.97 ± 0.17)</td>
<td></td>
</tr>
<tr>
<td>Peak $B_{ag}$ (kg m$^{-2}$)</td>
<td>1.20 ± 0.05</td>
<td>1.18 ± 0.09</td>
<td>1.20 ± 0.10</td>
<td>1.19 ± 0.18</td>
</tr>
<tr>
<td>(1.22 ± 0.07)</td>
<td>(1.14 ± 0.13)</td>
<td>(1.17 ± 0.14)</td>
<td>(1.17 ± 0.18)</td>
<td></td>
</tr>
<tr>
<td>LAI$_{max}$ (m$^2$ m$^{-2}$)</td>
<td>3.84 ± 0.29</td>
<td>3.83 ± 0.47</td>
<td>4.17 ± 0.26</td>
<td>4.15 ± 0.53</td>
</tr>
<tr>
<td>(3.84 ± 0.30)</td>
<td>(3.59 ± 0.46)</td>
<td>(3.98 ± 0.3)</td>
<td>(3.95 ± 0.52)</td>
<td></td>
</tr>
<tr>
<td>LAI RMSE (m$^2$ m$^{-2}$)</td>
<td>1.6 ± 0.1</td>
<td>1.4 ± 0.2</td>
<td>1.2 ± 0.1</td>
<td>1.1 ± 0.2</td>
</tr>
<tr>
<td>(1.6 ± 0.2)</td>
<td>(1.2 ± 0.2)</td>
<td>(1.3 ± 0.1)</td>
<td>(1.1 ± 0.1)</td>
<td></td>
</tr>
<tr>
<td>Median LAI increments (m$^2$ m$^{-2}$)</td>
<td>0.06 ± 0.28</td>
<td>-0.01 ± 0.03</td>
<td>-0.03 ± 0.33</td>
<td>-0.01 ± 0.03</td>
</tr>
</tbody>
</table>

Table 2. Impact of $N_L$ on MaxAWC retrieval using a single-parameter inverse modelling technique. The retrieved median MaxAWC and LAI RMSE are given for all 45 départements together with their standard deviation.

<table>
<thead>
<tr>
<th>$N_L$ (%)</th>
<th>1.05</th>
<th>1.30</th>
<th>1.55</th>
<th>2.05</th>
<th>2.55</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of départements presenting significant positive correlations ($p$ value &lt; 0.01)</td>
<td>16</td>
<td>13</td>
<td>12</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>MaxAWC (mm)</td>
<td>110 ± 44</td>
<td>110 ± 44</td>
<td>99 ± 43</td>
<td>99 ± 41</td>
<td>88 ± 40</td>
</tr>
<tr>
<td>LAI RMSE (m$^2$ m$^{-2}$)</td>
<td>1.2 ± 0.2</td>
<td>1.2 ± 0.2</td>
<td>1.3 ± 0.2</td>
<td>1.4 ± 0.2</td>
<td>1.5 ± 0.2</td>
</tr>
</tbody>
</table>

133 ± 46 mm for significant départements) is close to the standard value used in ISBA (132 ± 2 mm) but the standard deviation is much larger. This shows that LDAS tuning is able to generate spatial variability in MaxAWC values.

A similar degree of variability is obtained by inverse modelling, but the retrieved MaxAWC presents much lower values for all 45 départements: 111 ± 44 mm. On the other hand, a much larger value of 153 ± 40 mm is found for the 16 validated départements. The retrieved $N_L$ (1.05 ± 0.20) is smaller than the default value of 1.30 %. The role of $N_L$ in the optimisation is discussed in Sect. 4.1.

Figure 6 shows the impact of optimising MaxAWC on the mean annual LAI cycle, with respect to the observed annual LAI cycle over the 45 départements. Inverse modelling tends to produce a smaller LAI$_{max}$ median value (3.59 m$^2$ m$^{-2}$ for all 45 départements) than basic ISBA simulations or LDAS simulations (3.84 and 3.98 m$^2$ m$^{-2}$, respectively). Inverse modelling tends to reduce simulated LAI in May and June, while the LDAS simulations (either original LDAS or LDAS tuning) are much closer to the observations.

The two optimisation methods succeed in reducing the LAI RMSE of the basic ISBA simulations (1.6 m$^2$ m$^{-2}$ for all 45 départements). With optimised MaxAWC, the tuned LDAS annual mean LAI cycle is closer to the observations than LAI resulting from inverse modelling, with LAI RMSE equal to 1.1 m$^2$ m$^{-2}$ for LDAS tuning, as against 1.2 m$^2$ m$^{-2}$ for inverse modelling.

3.4 Validation

Agricultural GY statistics (Sect. 2.3) are used for validation. The optimisation is considered as successful in départements where the correlation between yearly time series of $B_{agx}$ and GY is significant ($p$ value < 0.01). Table 1 shows that even without tuning MaxAWC, the integration of LAI in ISBA by the original LDAS permits an increase in the number of départements where $p$ value < 0.01 from 18 in basic ISBA simulations to 21. LDAS tuning further increases this number to
24 départements. With only 16 validated départements, inverse modelling is not able to outperform the original LDAS simulations.

Time series of mean scaled anomalies of $B_{agX}$ and GY are shown in Fig. 7 for all 45 départements before and after inverse modelling or LDAS tuning. The marked negative anomalies ($<-1$) in 2001, 2003 and 2011 are represented well after LDAS tuning. On the other hand, the impact of sunlight deficit and low temperatures during the growing period of 2001 cannot be represented well after inverse modelling. The marked negative GY anomaly observed in 2007 is not very well represented by the model. Moreover, Fig. 7 shows that parameter tuning does not significantly improve $R^2$ values. Basic ISBA and original LDAS simulations present $R^2$ values of 0.65 and 0.80, as against 0.65 and 0.82 after inverse modelling and LDAS tuning, respectively.

Figure 8 further shows that the interannual variability of $B_{agX}$ is markedly better represented using LDAS tuning. The scaled modelled $B_{agX}$ and the scaled GY observations averaged over 45 départements present a $R^2$ value of 0.82, against 0.65 for inverse modelling. Considering only the successfully validated départements, more similar $R^2$ values are observed: 0.88 and 0.80, respectively. Figure 9 presents the spatial correlation between the scaled $B_{agX}$ and the scaled GY observations averaged over the 15-year period considered in this study. Considering the 45 départements, $R^2 = 0.61$ for LDAS tuning and $R^2 = 0.58$ for inverse modelling. Again, LDAS tuning supersedes inverse modelling, including when the comparison is limited to successfully validated départements, with $R^2$ values of 0.74 and 0.63, respectively.

It must be noted that all the correlations presented in Figs. 8 and 9 are significant, with all $p$ values smaller than 0.001.

In addition to GY data, we made an attempt to use the satellite-derived GLEAM evapotranspiration product (Miralles et al., 2011) but very poor correlations were obtained for most départements (the median $R^2$ value was less than 0.06). On the other hand, good correlations were found for photosynthesis using the GPP FLUXNET-MTE product described in Jung et al. (2009). With respect to basic ISBA simulations, GPP RMSE was nearly systematically improved by the original LDAS simulations, and LDAS tuning drastically reduced the largest RMSE values, observed in southwestern France (see Figs. S2 and S3 in the Supplement).

### 3.5 Impact of the optimisation technique on MaxAWC estimates

Differences in validation results can be caused by uncertainties in Agrete GY observations or by the difficulty in upscaling the observations and the simulations (Sect. 2.6). In order to limit this effect, we further compared the MaxAWC estimates and the simulated vegetation variables for a subset of the départements corresponding to the 15 départements which are validated for both inverse modelling and LDAS tuning. Table S3 shows that for this subset of départements, MaxAWC values are similar: $154 \pm 40$ and $156 \pm 40$ mm, respectively. On the other hand, vegetation variables are more realistically simulated after LDAS tuning: median LAI RMSE is $1.2 \text{ m}^2 \text{ m}^{-2}$ as against $1.4 \text{ m}^2 \text{ m}^{-2}$ for inverse modelling. The median LAI$_\text{max}$ value is much larger for LDAS tuning: $4.35 \text{ m}^2 \text{ m}^{-2}$, against $3.85 \text{ m}^2 \text{ m}^{-2}$ for inverse modelling. However, peak $B_{agX}$ values are similar: $1.26 \text{ kg m}^{-2}$ for LDAS tuning and $1.23 \text{ kg m}^{-2}$ for inverse modelling.
Figure 8. Temporal correlation between $A_{BagX}$ and $A_{GY}$ for (red symbols) départements presenting significant positive correlations ($p$ value < 0.01) between simulated $BagX$ and GY and (red and black symbols) all 45 départements, and for (a) inverse modelling and (b) LDAS tuning.

Figure 9. As in Fig. 8, except for spatial correlation.

4 Discussion

4.1 What is the added value of the LDAS?

The similarity in MaxAWC estimates and the contrasting simulated LAI values are illustrated in Fig. 10. Analysed LAI from LDAS tuning is closer to the LAI observations than the simulated LAI resulting from inverse modelling. The MaxAWC estimates are slightly smaller for inverse modelling and correlate very well with the MaxAWC estimates from LDAS tuning ($R^2 = 0.81$). The latter result is also valid when all 45 départements are considered, with $R^2 = 0.72$.
Improving the simulation of vegetation variables has a positive impact on the quality of simulated hydrological variables such as evapotranspiration and soil moisture (Szczypta et al., 2012). Therefore, the larger MaxAWC values obtained from LDAS tuning (129 ± 44 mm) are likely to be more realistic than those obtained from inverse modelling (111 ± 44 mm).

4.2 Are MaxAWC estimates and simulated peak $B_{ag}$ realistic?

Independent MaxAWC estimates confirm that the MaxAWC values obtained from LDAS tuning (129 ± 44 mm) are more realistic than those obtained from inverse modelling (111 ± 44 mm). On the other hand, the two techniques give similar median $B_{ag,X}$ values (Table 1).

In order to verify the MaxAWC values derived from LAI observations, we extracted MaxAWC values from a map produced by l’Institut National de la Recherche Agronomique (INRA) at a spatial resolution of 1 km 5 1 km. This map was established using pedotransfer functions based on soil physical properties information such as soil texture, soil depth, bulk density and organic matter (Al Majou et al., 2008). A given local MaxAWC value corresponds to a soil typological unit (STU). The 1 km 5 1 km soil mapping units may contain several STUs and the STU fraction is known. We computed weighted-average MaxAWC values for every 35 5 35 km grid cell. The resulting INRA MaxAWC values of the 45 départements present a median value of 151 mm and a standard deviation of 54 mm.

The median peak $B_{ag}$ values are about 1.2 kg m$^{-2}$ in all simulations. This is consistent with total maximum aboveground biomass values for cereals, which range between 1.1 and 1.7 kg m$^{-2}$ (e.g. Loubet et al., 2011). Because $B_{ag,X}$ corresponds to the mean $B_{ag}$ above 90% of the peak $B_{ag}$ value (Sect. 2.7), median $B_{ag,X}$ values are smaller than peak $B_{ag}$ and do not exceed 1 kg m$^{-2}$ (Table 1).

4.3 Are LAI satellite data suitable for the optimisation of MaxAWC?

Our results show that using disaggregated LAI observations is key.

The optimisation methods used in this study are based on disaggregated LAI satellite data and the quality of the results depends on the reliability of the observation data set. The MaxAWC parameter is a crucial parameter for the senescence period, between LAI$_{max}$ and harvesting (Fig. 4). Because LAI$_{max}$ is related to a large extent to MaxAWC (Fig. 3d), an underestimation of observed maximum LAI values would force the retrieval method to underestimate MaxAWC.

Figure 11 compares the mean of annual maximum values of raw LAI and disaggregated LAI for the 45 départements and for 1225 km$^2$ grid cells. Using disaggregated
LAI increases the observed value of maximum LAI by up to 40% with respect to raw LAI. The mean difference is 0.43 m² m⁻². This mitigates a marked underestimation of the MaxAWC estimates. As shown in Table S3, the MaxAWC values obtained from LDAS tuning (110 ± 38 mm) and from inverse modelling (83 ± 30 mm) are much lower (15 to 25%) than those retrieved using disaggregated LAI observations. Moreover, the number of validated départements using GY observations presenting significant positive correlation is reduced: only 10 and 18 for inverse modelling and LDAS tuning, as against 16 and 24 with disaggregated LAI, respectively. Also, peak $B_{ag}$ values (for all 45 départements) are smaller: 1.01 and 1.08 kg m⁻², against 1.14 and 1.17 kg m⁻² with disaggregated LAI, respectively.

### 4.4 Can model simulations predict the relative gain or loss of agricultural production during extreme years?

The continuous constraint on the model applied by the LDAS on simulated vegetation variables allows the indirect representation of adverse effects. This is illustrated in Fig. 7: the negative anomaly of 2007 is much better represented by the LDAS than by simple ISBA simulations.

The observed disaggregated LAI and GY in Fig. 2 show that 2004, 2008, 2009 and 2012 were favourable years for straw cereal production, while 2001, 2003, 2007 and 2011 were unfavourable years. Unfavourable conditions for straw cereal production were caused by droughts, by excess of water, or by a deficit in solar radiation. For example, the 2000–2001 winter was characterised by extensive floods and by a deficit of solar irradiance until the end of the spring. These climate events markedly affected plant growth especially in northern France (Agreste Bilan, 2001). The 2003 and 2011 years were particularly warm, with a marked precipitation deficit at springtime (Agreste Bilan, 2003, 2011). Concerning 2007, although climate conditions were favourable to plant growth during spring, extremely wet conditions occurred at the end of the growing season. This triggered accessibility issues and disease development (Agreste Bilan, 2007). These processes limiting biomass production in response to an excess of water are not represented in the ISBA model.
4.5 Can observed LAI characteristics be used to estimate MaxAWC?

We show that satellite-derived LAI observations have potential to map MaxAWC very simply.

We investigated the use of a simple statistical analysis of the disaggregated LAI observations to estimate MaxAWC. Figure 3 shows that there is a marked relationship between MaxAWC and the simulated LAI$_{\text{max}}$ and LAI CV. To what extent are these relationships observable?

In order to answer this question, we used the LDAS tuning MaxAWC estimates as a reference data set. We compared the observed median annual maximum LAI and LAI CV with MaxAWC. No significant correlation could be shown for LAI CV, with $R^2$ smaller than 0.2. On the other hand, a very good correlation ($R^2 = 0.70$ for all 45 départements) was found for median annual maximum LAI (Fig. 12). Using this simple linear regression model, MaxAWC can be estimated with a RMSE of 28.7 mm. A very similar result is obtained considering only the 24 validated départements for LDAS tuning. The modelled MaxAWC values are given in Table S2 (see Supplement).

5 Conclusions

Satellite data are used to optimise a key parameter of the ISBA land surface model for straw cereals in France: the maximum available soil water content, MaxAWC. Two optimisation methods are used. Inverse modelling consists in minimising the LAI RMSE and LDAS tuning consists in minimising LAI analyses increments. The added value of the optimisation is evaluated using simulated above-ground biomass, through its correlation with in situ grain yield observations.

It is found that disaggregated LAI observations during the senescence are more informative than raw LAI observations and than LAI observations during the growing phase. The best results are obtained using LDAS tuning: the simulated above-ground biomass correlates better with grain yield observations, and the retrieved MaxAWC values are more realistic. It is shown that LDAS simulations can predict the relative gain or loss of agricultural production during extreme years, much better than model simulations even after parameter optimisation.

Finally, it is shown that median annual maximum disaggregated LAI observations correlate with MaxAWC estimates over France. This simple metric derived from LAI observations could be used to map MaxAWC. More research is needed to investigate to what extent this conclusion holds for other regions of the world and other vegetation types.

## Appendix A: Nomenclature

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_{s,B_{ag\times}}$</td>
<td>Scaled anomaly of $B_{ag\times}$ of a given year (–)</td>
</tr>
<tr>
<td>$A_{s,GY}$</td>
<td>Scaled anomaly of GY of a given year ($z$ score) (–)</td>
</tr>
<tr>
<td>AWC</td>
<td>Simulated available soil water content (kg m$^{-2}$)</td>
</tr>
<tr>
<td>$B_{ag}$</td>
<td>Simulated living above-ground biomass (kg of dry matter m$^{-2}$)</td>
</tr>
<tr>
<td>$B_{ag\times}$</td>
<td>Maximum of simulated living above-ground biomass (kg of dry matter m$^{-2}$)</td>
</tr>
<tr>
<td>CV</td>
<td>Coefficient of variation (%)</td>
</tr>
<tr>
<td>$D_{max}$</td>
<td>Maximum leaf-to-air saturation deficit (kg kg$^{-1}$)</td>
</tr>
<tr>
<td>$F_s$</td>
<td>Soil moisture stress function</td>
</tr>
<tr>
<td>$g_m$</td>
<td>Mesophyll conductance in well-watered conditions (mm s$^{-1}$)</td>
</tr>
<tr>
<td>GY</td>
<td>Annual grain yields of crops (kg m$^{-2}$)</td>
</tr>
<tr>
<td>LAI</td>
<td>Leaf area index (m$^2$ m$^{-2}$)</td>
</tr>
<tr>
<td>LDAS</td>
<td>Land data assimilation system</td>
</tr>
<tr>
<td>LSM</td>
<td>Land surface model</td>
</tr>
<tr>
<td>MaxAWC</td>
<td>Maximum available soil water content (mm or kg m$^{-2}$)</td>
</tr>
<tr>
<td>NIT</td>
<td>Photosynthesis-driven plant growth version of ISBA-A-gs</td>
</tr>
<tr>
<td>$N_L$</td>
<td>Leaf nitrogen concentration (% of leaf dry mass)</td>
</tr>
<tr>
<td>SLA</td>
<td>Specific leaf area (m$^2$ kg$^{-1}$)</td>
</tr>
<tr>
<td>WUE</td>
<td>Leaf-level water use efficiency (ratio of net assimilation of CO$_2$ to leaf transpiration)</td>
</tr>
<tr>
<td>$Z_r$</td>
<td>Depth of the root zone layer (m)</td>
</tr>
</tbody>
</table>

### Greek symbols

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho$</td>
<td>Water density (kg m$^{-3}$)</td>
</tr>
<tr>
<td>$\theta$</td>
<td>Volumetric soil water content (m$^3$ m$^{-3}$)</td>
</tr>
<tr>
<td>$\theta_{Fc}$</td>
<td>Volumetric soil water content at field capacity (m$^3$ m$^{-3}$)</td>
</tr>
<tr>
<td>$\theta_{Wilt}$</td>
<td>Volumetric soil water content at wilting point (m$^3$ m$^{-3}$)</td>
</tr>
</tbody>
</table>
Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. The work of Hélène Dewaele was supported by CNES and by Météo-France. The work of Simon Munier was supported by European Union Seventh Framework Programme (FP7/2007–2013) under grant agreement no. 603608, “Global Earth Observation for integrated water resource assessment” (earTH2Observe). The work of Nabil Laanaita was supported by the Belgian MASC BELSPO project under contract no. BR/121/A2/MASC. The authors would like to thank the Copernicus Global Land service for providing the satellite-derived LAI products.

Edited by: Hannah Cloke
Reviewed by: two anonymous referees

References


