Maize yield estimation in West Africa from crop process-induced combinations of multi-domain remote sensing indices

Louise Leroux, Mathieu Castets, Christian Baron, Maria-José Escorihuela, Agnès Bégué, Danny Lo Seen

Maize yield estimation in West Africa from crop process-induced combinations of multi-domain remote sensing indices

Louise Leroux [a,b*], Mathieu Castets [c,d], Christian Baron [c,d], Maria-Jose Escorihuela [e], Agnès Bégué [c,d], Danny Lo Seen [c,d]

[a] CIRAD, UPR AIDA, Dakar, Sénégal.

[b] AIDA, Univ Montpellier, CIRAD, Montpellier, France.

[c] CIRAD, UMR TETIS, F-34398 Montpellier, France.

[d] TETIS, Univ Montpellier, AgroParisTech, CIRAD, CNRS, IRSTEA, Montpellier, France.

[e] isardSAT, Advanced Industry Park, Carrer Marie Curie 8-14, 08042 Barcelona, Spain.

[*] Corresponding author. Tel. : +221778665322

E-mail address: louise.leroux@cirad.fr (L. Leroux)

Abstract

Remote sensing data, crop modelling, and statistical methods are combined in an original method to overcome current limitations of crop yield estimation. It is then tested for timely estimation of maize grain yields and their year-on-year variability in Burkina Faso. Outputs from the SARRA-O crop model were used as a proxy for observed data for calibration. The final remote sensing-based yield model was constructed on the interaction between aboveground biomass at flowering (AGB-F) and crop water stress (Cstr) over the reproductive and maturation phases. Various vegetation and drought-related indices were derived from different spectral domains and tested. Model performance was evaluated by cross-validation against (a) simulated yields and (b) independent yields from ground surveys aggregated at village level. The results showed that the RF (Random Forest) model outperformed the MLR (Multiple Linear Regression) model for year-on-year yield estimation at the end of the season when compared to simulated yields. Surface soil moisture (SSM) information, as a proxy for soil water available for plant growth, together with information on the temperature of the canopy cover, helped to improve the RF maize yield model, impacting more particularly the estimation of crop water stress. Lastly, two months before harvest the RF model predicted 46% of the observed end-of-season maize grain yield variability. The combined remote sensing, crop model and machine learning method is thus an effective approach for estimating and forecasting inter-annual maize

crop yields in environments where field data are scarce, such as in most parts of the African continent. However, more research is needed to better retrieve the spatial variability of yields in order to strengthen current agricultural monitoring systems, and to address societal challenges, such as declining food security.

Keywords: Crop yield estimation and forecast; Food security; MODIS NDVI and LST; SMOS SSM; Statistical model; SARRA-O crop model, uncalibrated approach

1. INTRODUCTION

With more than a billion tons per year, maize (*Zea mays* l.) is the most widely grown crop in the world and is thus considered as key in supporting global food security. In addition, most of the maize produced in the developing world comes from low income countries, with the livelihoods of the most vulnerable populations strongly dependent on maize production and its fluctuations (Bassu et al., 2014; Shiferaw et al., 2011). In West Africa, maize as a staple crop plays a central role in fulfilling population food requirements (Chivasa et al., 2017; Shiferaw et al., 2011), which were estimated at more than 30 kg/capita/year in 2013 by FAOSTAT. However, important climate and demographic trends combine to worsen an already difficult situation in the region. In particular, in the first decade of the 21st century, maize yield barely increased, or even stagnated (Ray et al., 2012). Climate change has already been seen to impact maize productivity globally, with a 3.8% reduction in maize yields since 1980 (Lobell et al., 2011b). With global warming, a significant decline in maize yield is further foreseen for most parts of West Africa (Lobell et al., 2011a). In addition, the population is expected to outgrow yield improvement, portending a decline in per capita food production in the coming years (Ray et al., 2013). Among the current and future societal challenges brought about by climate change, food security in regions with dominant smallholder farming systems remains a pressing priority, whereby timely and reliable information on maize crop yields and their inter-annual variability is urgently needed for effective decision-making.

There are several ways of estimating crop yields, from a plot to continent scale. Direct methods based on field surveys are expected to give reliable yield estimates, but they have significant weaknesses, including the cost, in terms of time and labour, and the difficulty of upscaling to large areas (Burke and Lobell, 2017).

Another way of estimating crop yields is to use crop growth models that incorporate ecophysiological processes to simulate crop growth, development and yields according to soil characteristics, agricultural practices and meteorological data. For most of these crop models, water and/or heat stress during the reproductive and maturation phases are considered as crop yield limiting factors. For instance, in the FAO AquaCrop model, the impact of water stress is taken into account in canopy growth, senescence and transpiration, as well as in the Harvest Index that is used to derive crop yield from simulated aboveground biomass (Akumaga et al., 2017). In the SARRA-H crop model the impact of water stress during the reproductive and maturation phases is taken into account through the reduction of potential yields using a crop water stress factor (Cstr), with daily values ranging from 0 for no stress to 1 for full stress (Dingkuhn et al., 2003). These models require a large amount of input data on a field scale, which limits their scalability over large spatial scales. Conversely, some of these crop models, like SARRA-O, allow crop monitoring over large areas using spatialized agrometeorological data, such as rainfall estimates, as inputs, but in that case they are only able to represent broad processes.

Compared to crop growth models, remote sensing provides a physical measurement of crop areas, and their temporal and spatial development, which implicitly integrates underlying determinants of crop productivity, such as sowing dates, pest attacks, irrigation or intensification levels, which are not accessible or too expensive at these scales or are not always represented in crop growth models (Duncan et al., 2015). Owing to its large area and repetitive coverage at relatively low cost, satellite remote sensing has been widely used to estimate and forecast yields, but mainly for large homogeneous crop plots in developed countries (Becker-Reshef et al., 2010; Bolton and Friedl, 2013; Johnson, 2014; Johnson et al., 2016; Sibley et al., 2014). Some promising results have nevertheless been obtained for African smallholder farming (Jin et al., 2017) such as for maize grain yield estimations in East and southern Africa (e.g., Mkhabela et al., 2005; Unganai and Kogan, 1998), and for millet and/or sorghum (e.g., Groten, 1993; Leroux et al., 2016; Maselli et al., 2000; Rasmussen, 1992). However, the main limitations of such approaches can be categorized into three groups. First, as pointed out in Leroux et al. (2016) there are limitations in using vegetation index alone (e.g. Normalized Difference Vegetation Index, NDVI) as a direct indicator of final crop yields, particularly when aboveground biomass is not proportional to the final harvestable yield. Arguments put forward include the fact that: (1) different yields could be observed for the same amount of aboveground biomass due to spatial variability in management practices or environmental conditions, and (2) droughts during sensitive phases, such as the reproductive stage, can lead to significant yield reductions, but with negligible effects on vegetative aboveground biomass. To overcome these limitations, Leroux et al. (2016) proposed an approach based on MODIS NDVI and CWSI (Crop Water Stress Index) to assess each component of the yield equation, namely the aboveground biomass and the Harvest Index, to estimate pearl millet yields in Niger. Other approaches have been tested for estimating and forecasting yields based on canopy temperatures (Jackson et al., 1981; Johnson, 2014; Kogan, 1995; Unganai and Kogan, 1998) or soil moisture information (Chakrabarti et al., 2014; Holzman et al., 2014), given that plant heat stress or soil moisture availability has negative impacts on photosynthetic activity, development rates and reproductive processes, and thus on crop yields. In particular, since rainfall does not necessarily reflect the actual water available for plant growth, soil water content is considered as a better driver in explaining crop yields (Holzman et al., 2014). Secondly, most of these studies rely on statistical linear models to predict crop yields, although several underlying processes are nonlinear. Thus, some studies, mainly from the field of machine learning, have also tested the use of nonlinear models to predict crop yields. For instance, Johnson et al. (2016) compared model-based recursive partitioning and Bayesian neural networks to predict barley, canola and wheat yields in Canada, while Fieuzal et al. (2017) used artificial neural networks to estimate corn yields in France using optical and radar image time series. Lastly, at least in developing countries where reliable *in-situ* yield measurements are scarce, all these approaches rely on empirical relationships between remote sensing indices and national agricultural statistics. Agricultural statistics are generally available several months after the end of the cropping season and thus do not allow for timely and reliable yield estimations. Recently, several studies have proposed using an “uncalibrated approach”, meaning that remote sensing-based models are calibrated using outputs from crop models validated for the targeted crop and areas. Promising results were obtained when compared to ground data (Azzari et al., 2017; Burke and Lobell, 2017; Jin et al., 2017; Sibley et al., 2014), suggesting that the method could be an alternative in environments where field data are scarce.

In this context, the objective of this study is to develop an original method that overcomes the current limitations of crop yield estimation by combining recent research on remote sensing data, crop modelling, and statistical methods. More specifically we conduct a benchmarking analysis between a Multiple Linear model and a Random Forest model to estimate early and end-of-season maize yields in Burkina Faso (West Africa), using remote sensing indicators based on vegetation indices, canopy temperature, and surface soil moisture. The models training rely on proxy of observed crop variables ("uncalibrated approach") that are simulated by the crop growth model SARRA-O which was calibrated and verified for the Sahelian rainfed cereals.

2. MATERIALS

MODIS NDVI, MODIS Land Surface Temperature (LST) and SMOS (Soil Moisture and Ocean Salinity) Surface Soil Moisture (SSM), were used to derive phenological metrics as well as vegetation vigour, drought and heat stress related indices. In addition, the SARRA-O crop model (Baron et al., 2005) was used to simulate AGB-F (aboveground biomass at the flowering), Cstr (water stress coefficient) and final maize yields in order to calibrate remote sensing-based models, and ground-based data were employed to validate our results. All remote sensing processing, statistical analysis and graphical outputs were carried out with R software version 3.5.2 (R Core Team, 2018). The full list of the R packages and the main functions used are given in Supplementary Materials (S1).

2.1. STUDY AREA AND FIELD DATA SURVEY

The study was conducted from 2011 to 2016 around Koumbia and neighbouring villages. The whole study area was located in Tuy province, south-west Burkina-Faso, the main cotton production zone in the country (Figure 1a). The climate there is Sudanian, characterized by a unimodal rainfall season, with annual rainfall ranging from 650 mm to 850 mm (Figure 1b) and a rainy season lasting from June to September, with the highest cumulative rainfall being recorded in July and August. According to the Harmonized World Soil Database (FAO/IIASA/ISRIC/ISSCAS/JRC, 2012), the soils are mainly loamy sand and loam. The landscape is highly fragmented with small to very small fields, often smaller than 1 ha (Fritz et al., 2015) with high inter and intra-field spatial variability due to different soil conditions, farming practices and the presence of trees within fields. Maize and cotton are the main crops, about 90% of the cultivated area, and are often cultivated in rotation allowing the maize to benefit from inputs for cotton supplied on credit by the cotton company (Diarisso et al., 2015). Other crops encountered are sorghum, millet, sesame and groundnut. All crops are rainfed.

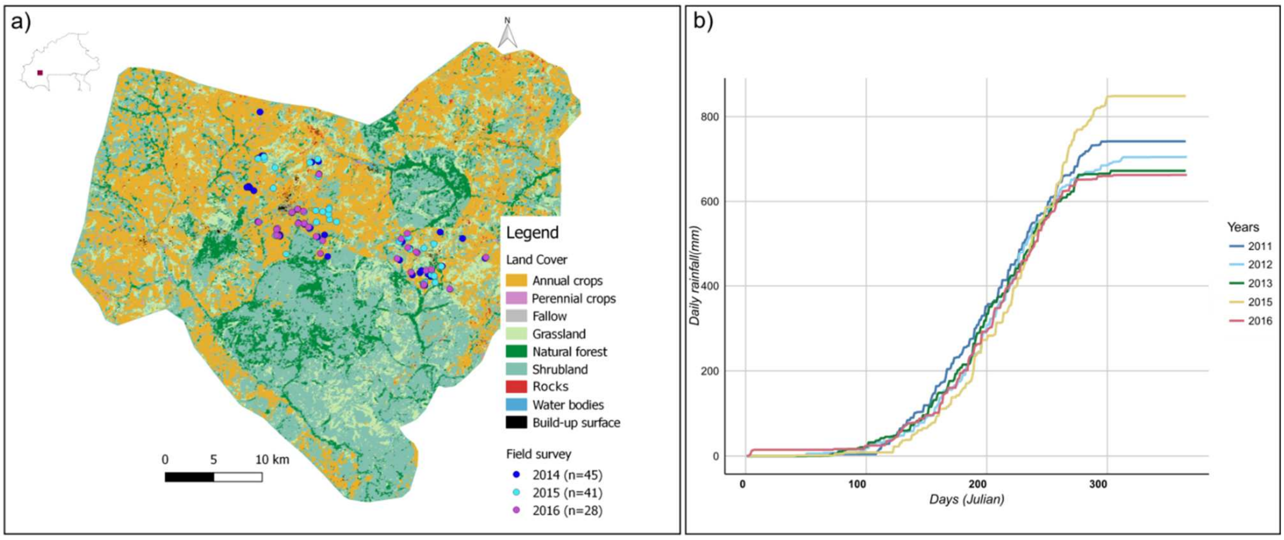

Figure 1: The study site: a) Main land cover (adapted from Gaetano et al. 2016) and location of field survey (dots), b) Mean daily cumulated rainfall over the study area for the 2011-2016 period, from TAMSAT rainfall estimates. 2014 was not included due to inconsistencies when compared to rain gauge data (see Section 2.3. Sarra-O crop model).

Field and rainfall data were collected in 2014, 2015 and 2016 for 3 villages located in the study area as part of the FP7 SIGMA project (http://www.geoglam-sigma.info). A network of 114 geolocalized maize fields (Figure 1a) under farmed conditions was surveyed to monitor agricultural practices (sowing dates, crop rotations, sowing density, etc.) and to measure biomass components (grain, stalk and leaf) at harvest. Measurements were made on three 25 m² randomly located plots within each field. Measured yields ranged from 415 kg/ha to 5840 kg/ha revealing high spatial and temporal variability in maize yields (Figure 2). Lastly, a fertility class was assigned to each field assuming maize grain yields were well correlated to the level of field fertility (Adiku et al., 2015), with four possible classes (Table 1). Maize yield data were then aggregated at village level on an annual basis, accounting for the weight of each fertility class. Field data were used to check the robustness of the remote-sensing maize yield model against independent data.

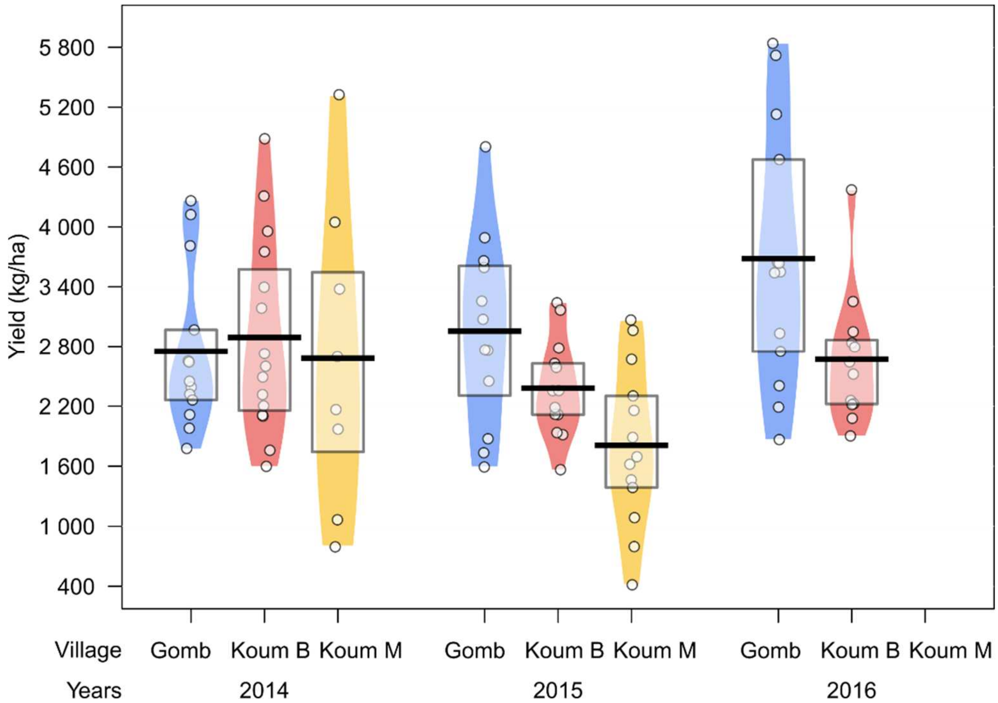

Figure 2. Variability in measured maize yields (white dots) according to villages and year, where Gomb = Gombeledougou (45 fields), Koum B = Koumbia Bwaba (46 fields) and Koum M = Koumbia Mossi (23 fields). Horizontal black lines indicate the mean value and the rectangles the interquartile range.

Table 1. Level of fertility according to measured maize yields in Tuy province (Burkina Faso).

Fertility class	F1	F2	F3	F4
Maize yield (kg/ha)	<1800	1800-3200	3200-4600	>4600
Weight (% of fields)	27	53	15	5

2.2. Remote Sensing data and preprocessing

2.2.1. Normalized Difference Vegetation Index

For this study, the NDVI time series from the MODIS MOD13Q1 product (collection 6; Didan, 2015) was used for the 2011-2016 period. MODIS products are freely distributed by the U.S. Land Processes Distributed Active Archive Center (https://lpdaac.usgs.gov/). The MOD13Q1 product consists of 16-day NDVI average values at a spatial resolution of 250 m. To reduce any remaining atmospheric effects, we applied a Savitzky-Golay temporal filter algorithm, a simplified least-square fit convolution used for smoothing time series. In order to account for areas with a very low vegetation density, or bare soils, and thus characterized by no vegetation seasonality, a mask was applied by excluding pixels with NDVI standard deviation values below 0.125.

In addition, a cropland mask was used to select MODIS cropped pixels and extract the corresponding NDVI values. The cropland mask was derived from a 2014 land cover classification (Gaetano et al., 2016) made with ground surveys and very high spatial resolution multi-source images (Pléiades, Deimos, RapidEye, Landsat). The classification was achieved using a Random Forest algorithm (Breiman, 2001) and produced a cropland map with 92% overall accuracy. A binary mask was then created (with annual cropped pixels set to 1 and other pixels to 0) and used to obtain a map with the percentage annual cropped pixels at MODIS resolution (250 m). The MODIS pixels with at least 50% of their area within the crop mask were kept as annual cropped pixels at 250 m.

2.2.2. Land Surface Temperature

The Land Surface Temperature (LST) MODIS MOD11A2 product (collection 6) was used in this study. The MOD11A2 product consists of a simple average of clear-sky LST values over an 8-day period at 1 km (Wan, 2015). The LST data were converted to degrees Celsius. One of the main limitations of MODIS LST data is that they are highly prone to contamination by clouds or other atmospheric disturbances. Thus, noisy pixels were removed when LST values were below 0°C and the missing values were filled by also applying a Savitzky-Golay filter. As for the NDVI time series, non-cropped pixels were masked out taking the same approach. Lastly, the temporal resolution of the MODIS LST time series was reduced to 16-day resolution in order to match that of the NDVI time series.

2.2.3. Surface Soil Moisture

The Soil Moisture and Ocean Salinity (SMOS) satellite mission provides global Surface Soil Moisture (SSM) at a spatial resolution of approximately 40 km and a revisit of less than 3 days with a high target accuracy of 4% volumetric soil moisture (Kerr et al., 2010). SMOS SSM provides soil moisture estimation for the top 5 cm of soil based on the relationship between soil moisture and dielectric constant that influences microwave brightness temperature (Gruhier et al., 2010). Soil moisture is highly variable spatially, so we applied a disaggregation approach in order to obtain more relevant soil moisture information for the monitoring of rainfed crops in heterogeneous agricultural landscapes. The data were disaggregated at a 1 km spatial resolution using the DisPATCh method (Merlin et al., 2010) based on MODIS NDVI and LST time series. A 16-day time series was obtained by averaging the daily SM over 16-day periods to match the MODIS NDVI temporal resolution. Lastly, as for NDVI and LST, the non-annual cropping pixels were masked out.

2.2.4. Derived vegetation vigour and drought remote sensing indices

In this study, indicators of vegetation productivity and of plant water or heat stress (referred to hereafter as ‘drought indices’) were calculated and investigated as explanatory variables for maize yields to take into account the impact of agricultural drought as a limiting factor in final crop yields. The indicators were based on NDVI, LST and SSM or a combination of them. Table 2 gives the different vegetation and drought indices used in the study. These include NDVI and SSM alone, TCI (Temperature Condition Index), CWSI (Crop Water Stress Index), TVDI (Temperature Vegetation Dryness Index) and SMADI (Soil Moisture Agricultural Drought Index).

Table 2. Vegetation and drought indices selected as explanatory variables. Equations are given in Supplementary Materials (S2).

Indices	Definition	Meaning	Remote sensing data	References
Vegetation indices				
NDVI	Normalized Difference Vegetation Index	Aboveground biomass production	MODIS NDVI	Tucker et al. (1980)
Drought indices				
SSM	Soil Surface Moisture	Soil water content	SMOS SSM	Kerr et al. (2010)
TCI	Temperature Condition Index	Temperature related to vegetation stress over time	MODIS LST	Kogan (1995)
CWSI	Crop Water Stress Index	Variation in water deficit through space	MODIS LST	Jackson et al. (1981) Son et al. (2012)

TVDI	Temperature Vegetation Dryness Index	Soil water availability and vegetation conditions on the surface	MODIS NDVI, MODIS LST	Sandholt et al. (2002)
SMADI	Soil Moisture Agricultural Drought Index	Soil moisture drought conditions	MODIS NDVI, MODIS LST, SMOS SSM	Sánchez et al. (2016)

Several studies suggested that accumulated vegetation or drought indices are more closely related to vegetation growth and crop production than instantaneous measurements (e.g., Meroni et al., 2013; Tucker, 1985). On the other hand, Bolton and Friedl (2013) also found that the inclusion of information related to crop phenology significantly improved the prediction of maize crop yields. Therefore, each vegetation vigour and drought indicator was integrated over two phenological phases in order to account for spatial and temporal variations in crop growth due to environmental characteristics and management strategies: (1) the vegetative phase determining the aboveground biomass and (2) the productive phase, starting with the reproductive phase and ending with the ripening phase, and including heading, flowering and development of fruit, which are sensitive periods. To do so, three seasonal phenological metrics were derived from MODIS NDVI time series: (1) SOS, the timing of the start of the growth phase, (2) EOS, the timing of the end of the senescence phase and (3) TOS, the timing of the peak of the growing season. The R software "greenbrown" package (http://greenbrown.r-forge.r-project.org/index.php) was used to derive phenological metrics on a pixel basis with a fixed threshold computed from the long-term mean values of the NDVI time series. All vegetation or drought-related indices were then integrated over the vegetative period corresponding to the period between SOS and TOS and the productive period corresponding to the period between TOS and EOS. In order to match the output of the SARRA-O crop model, all indicators were resampled from their respective original spatial resolution to 4 km spatial resolution using the nearest neighbour method.

2.3. SARRA-O CROP MODEL SIMULATIONS

The process-based crop model SARRA-O was used in this study. SARRA-O is the spatialized version of the SARRA-H crop model (Baron et al., 2005) implemented under the Ocelet modelling platform (Degenne and Lo Seen, 2016). SARRA-H is a process-based crop model, designed to simulate attainable agricultural yields under tropical conditions. It takes into account potential and actual evapotranspiration, phenology, potential and water-limited assimilation, and biomass partitioning. The SARRA-H crop model was calibrated and verified for different millet, sorghum and maize cultivars based on agronomic trials and on-farm surveys conducted in different West African countries (Senegal, Burkina-Faso, Mali and Niger; Traoré et al. 2011) and showed good agreement with FAO statistics (Sultan et al., 2014) and other crop model results (Bassu et al., 2014; Durand et al., 2018). SARRA-H has been used in several studies in West Africa, particularly to assess the impacts of climatic change on future cereal yields (Guan et al., 2015; Oettli et al., 2011; Sultan et al., 2013) or to assess the impact of climate on farmers' cropping practices (Guan et al., 2017; Marteau et al., 2011; Roudier et al., 2016). In SARRA-H, the water constraints impact the potential yield reduction in different ways, depending on the phenological phases. During the anthesis to flowering stages, water constraints will induce a reduction in biomass and yields through a reduction in the number of grains. After flowering, assimilates are distributed to grains as a priority.

The SARRA-O crop model uses as inputs TAMSAT rainfall data (daily, 4 km; Tarnavsky et al., 2014) and ECMWF (European Centre for Medium-range Weather Forecast) agrometeorological data (minimum and maximum temperature, global radiation and evapotranspiration) provided on a 10-day frequency at 0.25° spatial resolution. However, due to an underestimation of cumulative rainfall over the rainy season and an overestimation of the daily intensity of rainfall events at the beginning of the rainy season by TAMSAT compared to ground measurements (Guillemot, 2016), 2014 was not considered in this study. A local maize cultivar, adapted from the DMR-ESR-W cultivar in Benin (Allé et al., 2014) with a 120-day cycle duration and lowest harvest index, was used for all simulations. Soils were defined based on the Harmonized World Soil Database (FAO/IIASA/ISRIC/ISSCAS/JRC, 2012). To catch soil fertility variability due to different management practices, simulations were conducted considering four fertility levels (F1 to F4; see Section 2.1.) for each type of soil which is translated in SARRA-O by a decrease in the optimum radiation to biomass conversion capacity. In the model, the sowing date was automatically generated starting from May 1. The sowing date is defined as the day when simulated plant soil water availability is greater than 10 mm at the end of the day, followed by a 20-day period during which crop establishment is monitored. The juvenile stage of the crop is considered failed, triggering automatic re-sowing, if the simulated daily total biomass decreases for 3 of the 20 days.

3. Methodology

3.1.1. Overall approach

The proposed methodology is illustrated by the flowchart presented in Figure 3 and is organised around two main phases: calibration and validation. In this study, two yields models are tested. The "end of season" yield model is developed combining a remote sensing model for each component of the yield equation, namely the aboveground biomass at flowering (AGB-F, Figure 3a) and the Cstr (water stress coefficient, Figure 3b) using remote sensing indices integrated over the whole cropping season (Figure 3d). The "early estimate" yield model is built using remote sensing indices integrated over the vegetative period only to have yield forecast two-months before the end of season (Figure 3c). The SARRA-O crop model is used in this study to simulate vegetation AGB-F (Figure 3a), the Cstr (Figure 3b) and attainable final maize yields over the study area, for each growing season between 2011 and 2015, according to the soil type, rainfall regime and agricultural practices (crop varieties, sowing dates and fertility classes). The simulated SARRA-O data are then used as a proxy for ground data in the calibration phase of the remote sensing-based model, in a so-called "uncalibrated approach". In addition, in order to account for drought-related stress over sensitive phenological phases of the maize cropping development cycle, the Cstr values were integrated over the productive period (i.e. reproductive and maturation phases, corresponding to phase 4 and phase 5 in SARRA-O, starting with inflorescence emergence stage and ending with ripening stage), when the final grain yields were set (Eq.1). The maize AGB-F, Cstr and final grain yields for each fertility class were aggregated using a weighted average according to the proportion of each fertility class observed in the field network (Table 1).

$$Cstr = \sum_{Inflorescenceemergence}^{Ripening} Cstr_{daily} \quad (1)$$

Where $Cstr$ (unitless) is the Cstr value integrated during the productive phase and $Cstr_{daily}$ the SARRA-O daily simulated Cstr values.

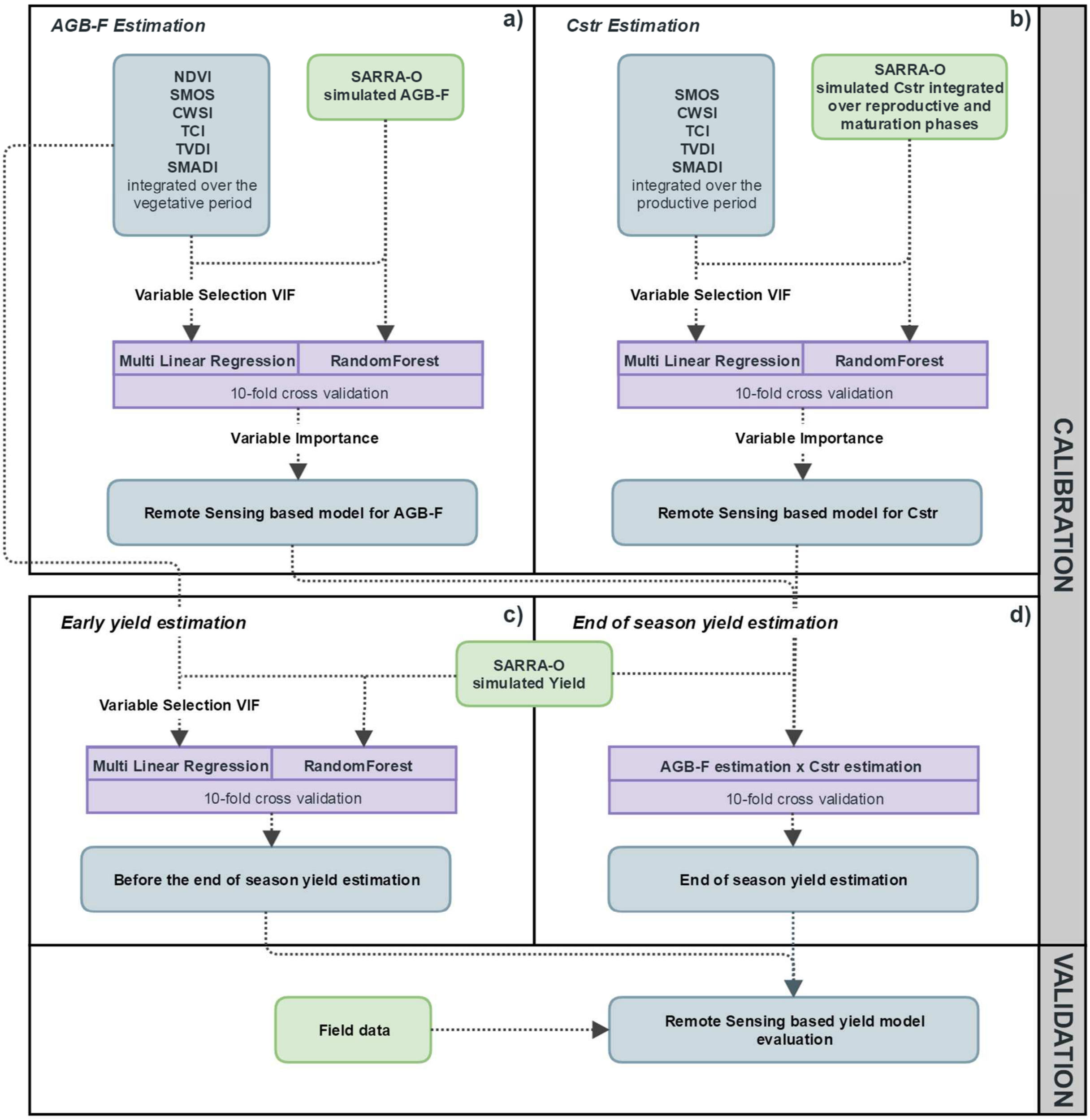

Figure 3. Flowchart of the methodology used to estimate maize grain yield.

3.1.2. Statistical models development and accuracy measurements

Linear (MLR) and nonlinear (RF) models were tested and compared. Each model was trained with vegetation vigour and drought related indices as candidate explanatory variables for AGB-F and Cstr (Table 2).

The Multiple Linear Regression (MLR) model is a regression with two or more predictors that are linearly related to the dependent variable. To avoid an overfitting of the MLR model, the VIF (Variable Inflation Factor) was used to remove highly correlated predictors. A stepwise approach was used in which VIF was recalculated at each step and the predictors with the highest VIF were dropped until all VIF values were smaller than 2 (see Zuur et al. 2010). The remaining predictors were then used in the MLR model. The MLR model was established for different combinations of predictors (i.e. vegetation or drought-related remote sensing indices) and cross terms were also considered to take into account possible interactions between predictors (e.g. impact of SSM on photosynthetic activity and thus on NDVI). Lastly, in order to help in understanding and to give an ecophysiological meaning to the resulting model, the contribution of each predictor to the final MLR model was assessed using the Lindeman, Merenda and Gold (LMG) method. The contribution of each predictor was expressed as a percentage (Grömping, 2006).

The Random Forest (RF) model is a non-parametric algorithm proposed by Breiman (2001) and is an ensemble learning method based on the combination of decisions from multiple decision trees. RF is a method that is now widely used in crop monitoring, both for crop yield modelling or cropland mapping (e.g., Forkuor et al., 2017; Lebourgeois et al., 2017). It is particularly valuable for its capacity to assess the relative importance of each predictor (i.e. remote sensing indices) used for regression. For the former point, we focused on an RF internal variable importance measurement, namely the mean decrease in Mean Square Error (MSE) where the larger the decrease in MSE, the higher is the predictor variable. In this study, the RF algorithm was implemented using the RandomForest package available in R (Liaw and Wiener, 2002).

The MLR and RF models were calibrated separately for AGB-F and Cstr. (Figure 3a and Figure 3b) Vegetation vigour and drought-related remotely sensed indicators integrated over the vegetative (i.e. from SOS to TOS) and productive (i.e. from TOS to EOS) periods for AGB-F and Cstr, respectively, were used as potential candidates each time the MLR and RF models were calibrated and compared over the 2011-2015 period (2014 excluded). In order to test the added-value of soil water content information in estimating and predicting maize crop yields, each model was calibrated with and without SSM predictors in the input dataset. The model was assessed using a 10-fold cross validation approach and through five statistical metrics (Table 3).

Table 3. Definition of the accuracy metrics used in the study, where y_{obs} is the observed data (i.e., the data simulated with the SARRA-O crop model), y_{pred} is the predicted data with the MLR or RF models and y_{mean} is the multi-annual average of observed data and the number of years. "cv" stands for "cross-validation".

Metric	**Formula**	**Definition**
cv-R² (coefficient of determination)	$1-\frac{\sum_{2011}^{2015}\left(y_{obs}-y_{pred}\right)^2}{\sum_{2011}^{2015}\left(y_{obs}-y_{mean}\right)^2}$	Expresses how accurately the vegetation vigour or drought-related remote sensing indices can predict AGB-F or Cstr. Its values vary between 0 (none of the observed data variability is explained by the model) to 1 (all of the observed data variability is explained by the model).
cv-RMSE (Root Mean Square Error)	$\sqrt{\frac{\sum_{2011}^{2015}\left(y_{pred}-y_{obs}\right)^2}{n}}$	Gives the error between the observed and predicted values. Values can range from 0 to ∞ with lower values indicating better fitting of the model.
cv-RRMSE (Relative Root Mean Square Error)	$\frac{cvRMSE}{y_{mean}}$	Is a normalization of cv-RMSE by the multi-annual average observed data. Lower values indicate better fitting of the model.
cv-MAE (Mean Absolute Error)	$\frac{1}{n}\sum_{2011}^{2015}\left\lvert y_{obs}-y_{pred}\right\rvert$	Measures the average error in the predicted values when compared to the observed values. Values can range from 0 (no average absolute difference between observed and predicted data) to ∞ (large average absolute difference).
Willmott d (index of agreement d)	$d=1-\frac{\sum_{2011}^{2015}\left(y_{obs}-y_{pred}\right)^2}{\sum_{2011}^{2015}\left(\left\lvert y_{pred}-y_{mean}\right\rvert+\left\lvert y_{obs}-y_{mean}\right\rvert\right)^2}$	Is a standardized measurement of the degree of model prediction error and is very useful for cross-comparisons

		between models. Its values vary between 0 (no agreement between observed and predicted data) and 1 (perfect agreement).

In order to be in line with the main ecophysiological processes involved in final grain yield formation as implemented in most crop models, conversion from the aboveground biomass at flowering (AGB-F) and crop water stress (Cstr) previously estimated by remote sensing, up to final end-of-season maize grain yields, was done using a MLR model with interaction between AGB-F and Cstr (Figure 3d, Eq.2).

$$Yield_{SARRA-O} = \alpha + \beta_1 AGB_{F_{estimated}} + \beta_2 Cstr_{estimated} + \beta_3 AGB_{F_{estimated}} Cstr_{estimated} \quad (2)$$

Where $Yield_{SARRA-O}$ is the final maize grain yield simulated by SARRA-O; $AGB_{F_{estimated}}$ and $Cstr_{estimated}$ are aboveground biomass at flowering and the crop water stress index integrated over the productive period estimated by remote sensing models; β_1, β_2, β_3 are the regression coefficients associated with each term of the equation and α the error term. In addition, in order to assess whether the final maize grain yields could be accurately predicted before the end of the season, the MLR and RF models were also calibrated using vegetation vigour and drought-related indices integrated over the vegetative period only, with final maize yields simulated by SARRA-O as the variable to be predicted (Figure 3c). The robustness of each remote-sensing model for maize yields was verified using independent field data for 2014, 2015 and 2016 aggregated on a village scale.

4. RESULTS

4.1. SARRA-O CROP MODEL RESULTS

The results of SARRA-O simulations for AGB-F, Cstr and final grain yields for the 2011-2015 period (2014 excluded) are shown in Figure 4. Aboveground biomass at the flowering stage simulated by the SARRA-O crop model exhibited high temporal variability, with values ranging from 3400 kg/ha in 2015 to almost 7000 kg/ha in 2011 (Figure 4a). Spatial variability was also observed with extreme annual ranges of 800 kg/ha to 1800 kg/ha observed in 2013 and 2015, respectively. On average, 2015 was the year with the lowest values for vegetative biomass at the flowering stage, mainly due to a late onset of the rainy season and irregular first rainfall events, with a low daily intensity (Figure 1b). Compared to AGB-F, the spatial and temporal variability of Cstr integrated over the reproductive and maturation phases (SARRA-O phases 4 and 5) was relatively low, except for 2015, which experienced lower water stress during phases 4 and 5 (Cstr around 81), explained by good rainfall over the productive phase of crop growth (Figure 1b and Figure 4b). Thus, the final grain yields simulated by SARRA-O mainly reflected the level of AGB-F, with simulated maize yields ranging from 2400 kg/ha in 2015 to more than 4000 kg/ha in 2011 (Figure 4c).

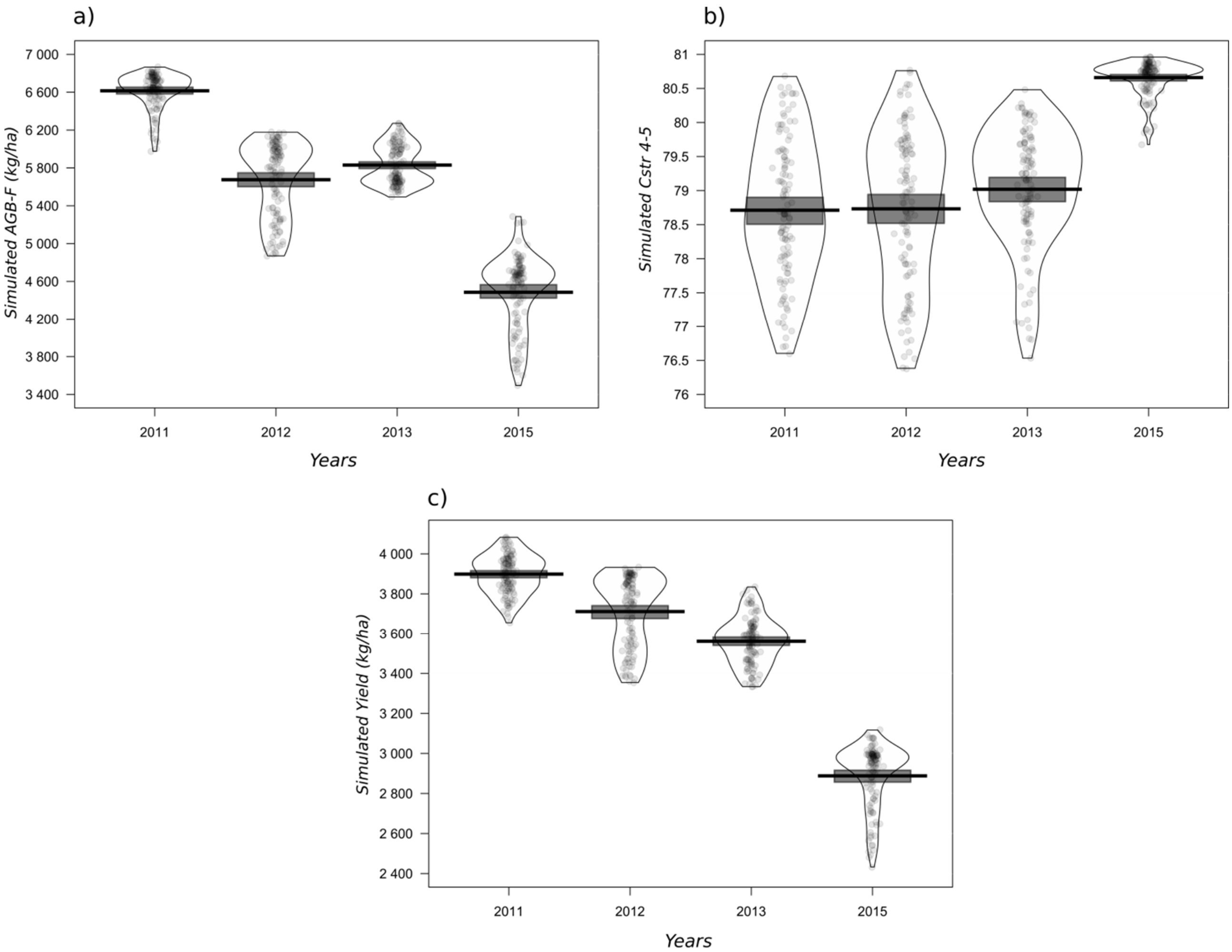

Figure 4: Annual and spatial variability in biomass at flowering simulated by the SARRA-O crop model (a), Cstr for phase 4 (reproductive phase) and phase 5 (maturation phase including development of fruit and ripening) (b) and final maize grain yields (c). Horizontal black lines indicate the mean value. Cstr is unitless, with lower values indicating higher water stress, while higher values indicate lower water stress.

4.2. Model for the evaluation of aboveground biomass at flowering

The AGB-F simulated by SARRA-O was used to calibrate a remote sensing-based model over the 2011-2015 period using vegetation vigour and drought-related variables (Table 2) integrated over the vegetative period (SOS-TOS) as predictors. Table 4 presents the accuracy metrics for the final MLR and RF models considering either the full dataset or the dataset without predictors derived from SMOS SSM (raw SSM and SMADI integrated over the vegetative period). Figure 5 shows the results (accuracy assessment, cumulative distribution and variable importance) for the full MLR and RF model. For the MLR model, a prior selection of the most relevant predictive variables was done by removing predictors causing VIF>2 (NDVI, TVDI and SMADI). The remaining predictors in the MLR model were variables related to drought conditions (TCI, CWSI) and soil water content (SSM), with an interaction between TCI and CWSI resulting in the highest prediction accuracy (Figure 5a and Table 4). The final MLR model had a moderate but highly significant predictive power (cv-R^2 = 0.46 and cv-RRMSE of 10.6%) with a tendency to overestimate AGB-F for low values when compared to simulated values, and to underestimate for high AGB-F values (Figure 5a and 5c). The RF model was significantly better than the previous one for ABG-F estimation with a cv-R^2 of 0.57, a relative error in cross validation of 9.3% (Figure 5b and 5d) and an index of agreement d of 0.84 compared to 0.71 for the MLR model. The resulting scatterplot of points revealed clear specific patterns for each year (Figure 5a and 5b) which means that, for both the MLR and RF models, the explained variance came from the inter-annual variation rather than the spatial variation. The variability distributions for estimated and simulated AGB-F differed but followed approximately a normal distribution in both cases, particularly for estimated AGB-F (Figure 5c and 5d). In addition, estimated and simulated AGB-F were in relatively good agreement around median values, though with better fitting of the RF model to the simulated probability distribution curve (Figure 5d). For the MLR model, among the predictors in the model, the variable related to the soil water content appeared to be highly relevant in predicting AGB-F and its annual variability, with 50% of the cv-R^2 explained by SSM, while for the RF model the most important variables were TCI (31% mean decrease in MSE) and NDVI (26% mean decrease in MSE), a proxy for cover temperature and photosynthetic activity, respectively. When the MLR and RF models were calibrated without SMOS SSM variables, unsurprisingly the predictive accuracy was significantly lower for the MLR model (cv-R^2 of 0.23) while for the RF model no significant improvement was observed (Table 4).

*Table 4. Cross-validation metrics for the AGB-F MLR and RF models. For MLR, the final predictor combination is indicated in square brackets. * NDVI, SMOS, TCI, TVDI, CWSI, SMADI. *** All cv-R^2 are significant for p-value<0.001.*

	MLR		***RF***	
	Full [CWSIxTCI+SMOS]	Without Soil Moisture predictors [CWSIxTCI]	Full*	Without Soil Moisture predictors
cv-R^2 [0-1]***	0.46	0.23	0.57	0.58
cv-RMSE (kg/ha)	610	723	537	529
cv-RRMSE (%)	10.6	12.5	9.3	9.2
cv-MAE (kg/ha)	475	582	397	393
d [0-1]	0.71	0.45	0.84	0.85

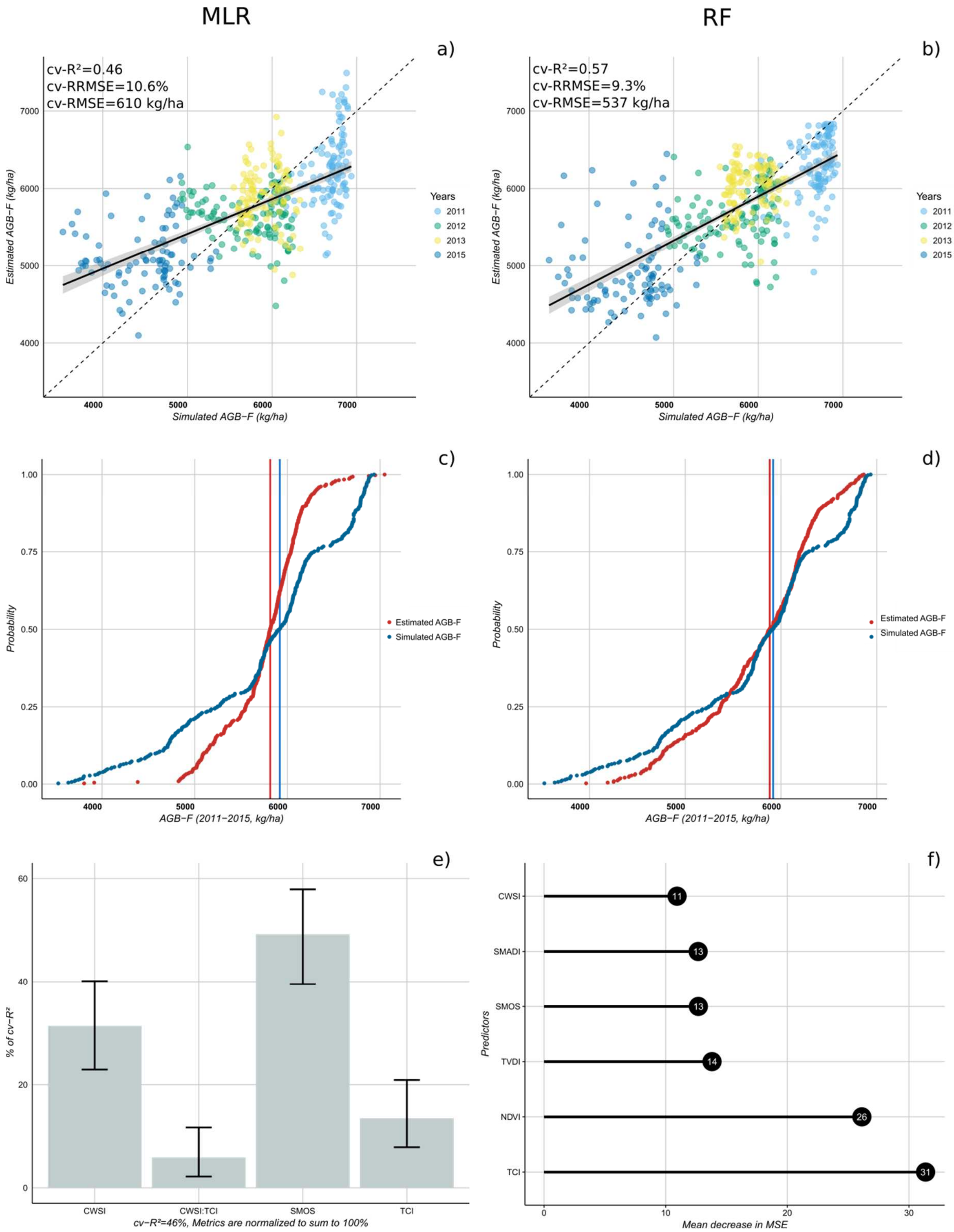

Figure 5: Cross-validation accuracy metrics for the AGB-F remote sensing-based model for the full MLR (a) and RF (b) models. Cumulative probability distribution between estimated and simulated AGB-F for the full MLR (c) and RF (d) models. Vertical lines represent median values. (e) Relative importance of the predictors used in the AGB-F full MLR model with a 95% bootstrap confidence interval using the Lindeman, Merenda and Gold method. (f) Relative importance of the predictors used in the AGB-F full RF model using the mean decrease in Mean Square Error.

4.3. EVALUATION OF CSTR DURING REPRODUCTIVE-MATURATION PHASES OF THE MODEL

Remote sensing drought indices (Table 2) integrated over the productive period (from TOS to EOS) were then used as potential predictors of a final grain yield reducing factor, which was a water stress coefficient (Cstr) calculated during sensitive phases of crop growth (flowering, development of fruit and ripening phases). The results of Cstr model calibration are presented in Table 5 and Figure 6. For the MLR model, the variables remaining as predictive variables after VIF selection were TCI, TDVI and SMADI. The MLR model exhibited a low but still significant predictive power with a cv-R^2 of 0.24 with CWSI, TCI and SMADI remaining in the final model, with interaction between TCI and SMADI, and TCI explaining more than 50% of cv-R^2. For Cstr estimation, the RF model was also better with a cv-R^2 of 0.43 with the most important variables being variables related to temperature conditions (TCI and CWSI), as for the MLR model (Figure 6e and 6f). However, for both models, the cv-RRMSE value was below 2%, meaning that, on average, the values predicted by remote sensing were very close to the Cstr values simulated by the SARRA-O crop model. This could be attributed to the low variability observed in the Cstr values simulated by SARRA-O. When analysing variability distribution (Figure 6c and 6d), we found quite a similar distribution curve for the MLR and RF model estimations with, in both cases, an overestimation of low Cstr values and, conversely, an underestimation of the highest Cstr values. Calibration of the Cstr models without SSM-related predictors led to a stronger impact on the predictive accuracy of the RF model, resulting in a drop in cv-R^2 of 28% explained mostly by the fact that SSM and SMADI accounted together for 43% of the mean decrease in MSE (Figure 6f).

4.4. EVALUATION AND VALIDATION OF MAIZE GRAIN YIELD MODELS (END OF SEASON AND EARLY END ESTIMATIONS)

Firstly, the end-of-season maize grain yield model was established using a linear relationship between AGB-F estimated by remote sensing during the vegetative period, Cstr during the productive period and yields simulated by SARRA-O (Eq.2). Table 6 and Figure 7 present the accuracy evaluation of the final remote sensing-based MLR and RF models for yields, considering either the full dataset or the dataset without SMOS SSM-related variables.

*Table 5. Cross-validation metrics for Cstr Phase 4-5 MLR and RF models. For MLR, the final predictor combination is indicated in square brackets. * SMOS, TCI, TVDI, CWSI, SMADI. *** All cv-R^2 are significant for p-value<0.001.*

	MLR		***RF***	
	Full [CWSI+TCIxSMADI]	Without Soil Moisture predictors [CWSI+TCI]	Full*	Without Soil Moisture predictors
cv-R² [0-1]***	0.24	0.17	0.43	0.15
cv-RMSE	1.05	1.10	0.90	1.11
cv-RRMSE (%)	1.31	1.37	1.13	1.39
cv-MAE	0.87	0.90	0.69	0.89
d [0-1]	0.60	0.50	0.77	0.50

The final remote sensing-based models both had good potential for estimating end-of-season maize yields and their inter-annual variability with a cv-R^2 of 0.54 for the MLR model and cv-R^2 of 0.59 for the RF model between simulated SARRA-O and predicted maize yields and a cross-validated RMSE below 300 kg/ha (Table 6 and Figure 7a and 7b). Average conditions, and extreme conditions like those in 2015, were properly captured by both models, while spatial yield variability was poorly rendered by both models, as suggested by the absence of a clear pattern within each year (Figure 7a and 7b). This results in overall good fitting of the estimated yields to the simulated probability distribution curve, particularly around low and median values, with median simulated yields of 3634 kg/ha and 3648 kg/ha for the MLR model and 3659 kg/ha for the RF model (Figure 7c and 7d). Lastly, the importance of the SMOS SSM-related variables in the predictive accuracy of the remote sensing-based model mirrored that of the AGB-F biomass estimation, with a higher impact on the MLR model than on the RF model (Table 6).

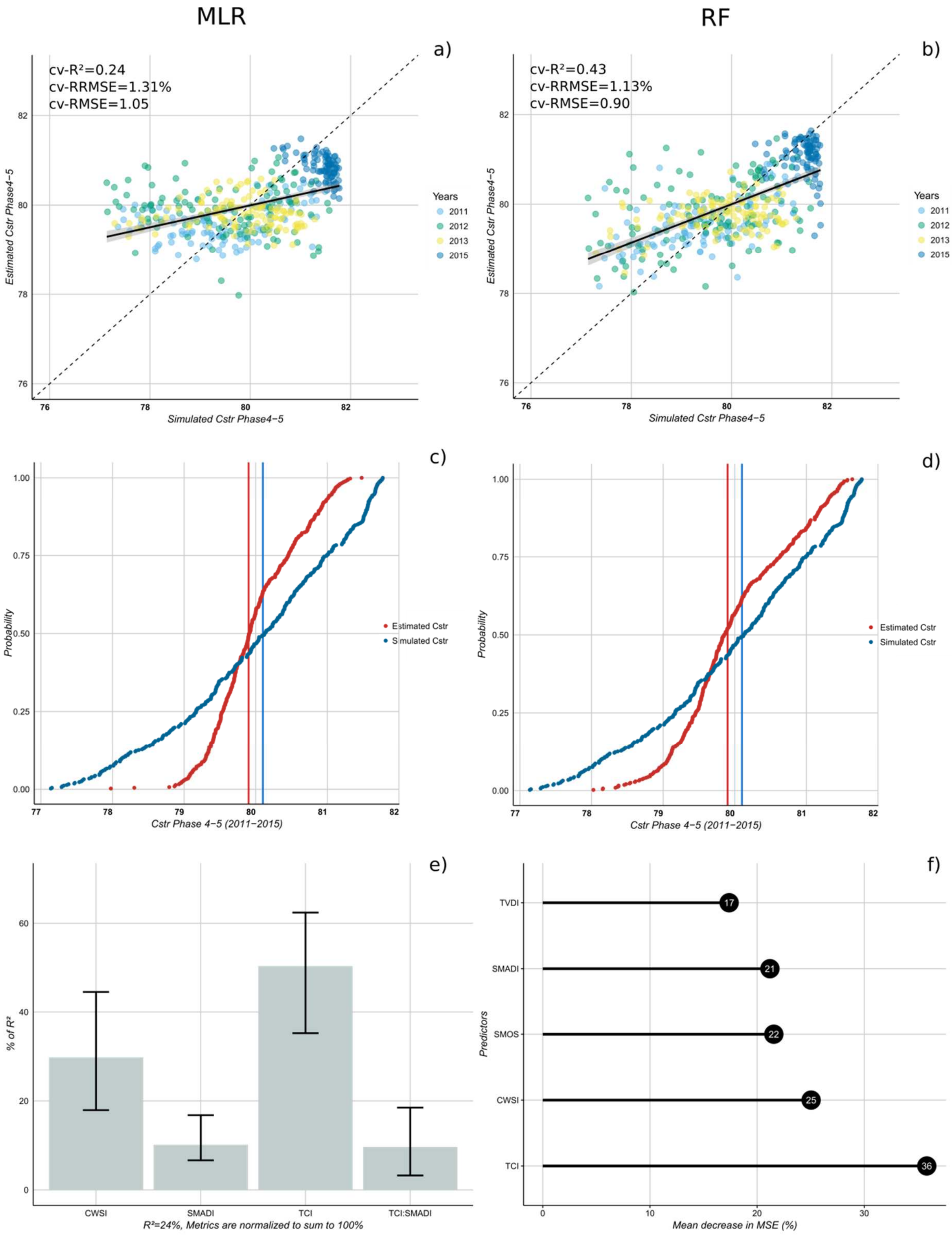

Figure 6: Cross-validation accuracy metrics for the Cstr Phase 4-5 remote sensing-based model, for the full MLR (a) and RF (b) models. Cumulative probability distribution between estimated and simulated Cstr Phase 4-5 for the full MLR (c) and RF (d) models. Vertical lines represent median values. (e) Relative importance of the predictors used in the Cstr Phase 4-5 full MLR model with a 95% bootstrap confidence interval using the Lindeman, Merenda and Gold method. (f) Relative importance of the predictors used in the Cstr Phase 4-5 full RF model using the mean decrease in Mean Square Error.

This implies that incorporating soil moisture adds little information to the RF model, while most of the yield variability seems to be linked to the soil water content information in the MLR model.

*Table 6. Cross-validation metrics for final maize grain yield models based on the combination of AGB-F and Cstr Phase 4-5 estimated by the MLR and RF models. *** All cv-R^2 are significant for p-value<0.001).*

	MLR		***RF***	
	Full	Without Soil Moisture predictors	Full	Without Soil Moisture predictors
cv-R^2 [0-1]***	0.54	0.17	0.59	0.52
cv-RMSE (kg/ha)	271	363	258	273
cv-RRMSE (%)	7.65	10.3	7.29	7.72
cv-MAE (kg/ha)	213	291	201	214
d [0-1]	0.80	0.48	0.82	0.79

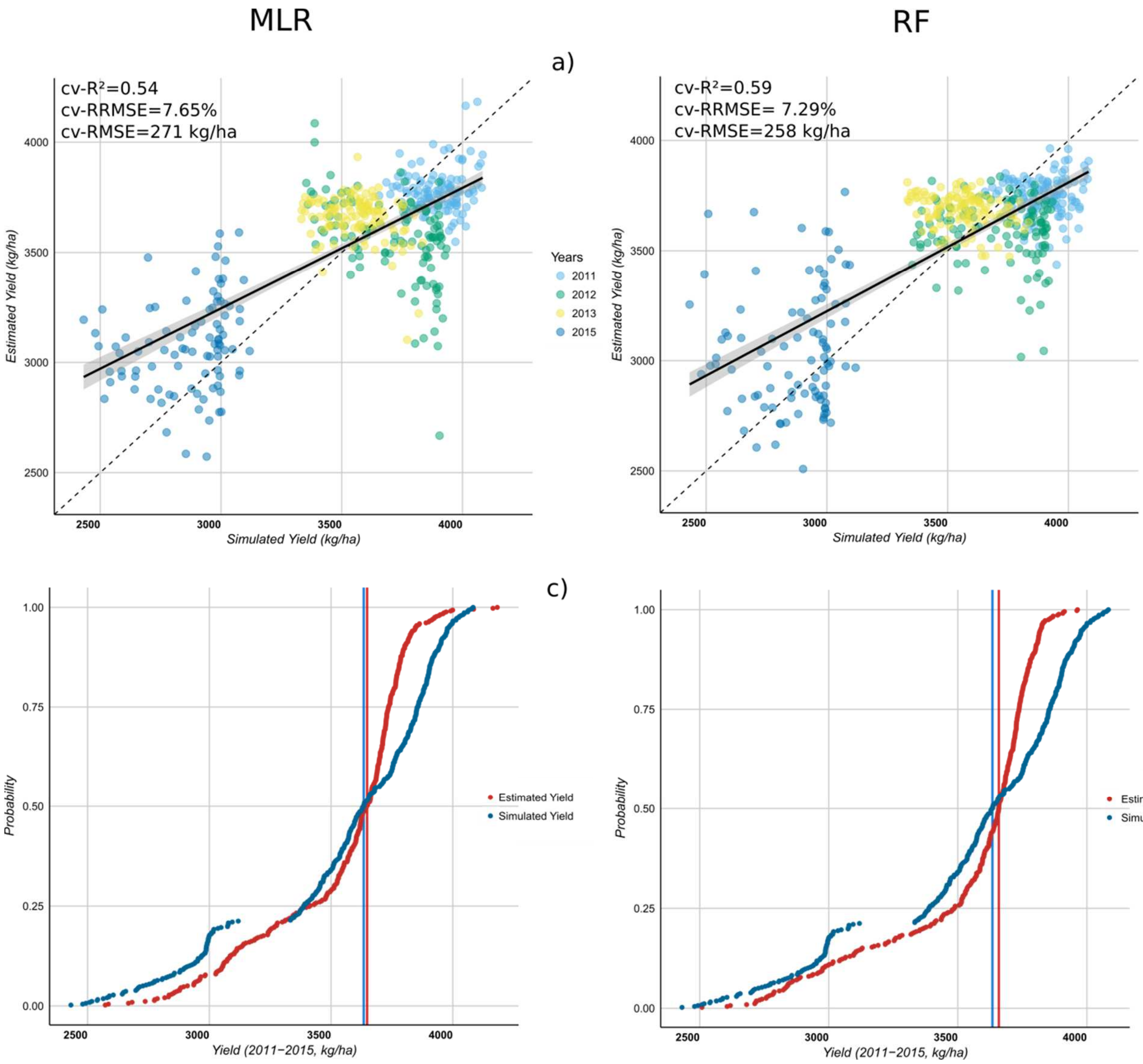

Figure 7: Cross-validation accuracy metrics for the maize yield remote sensing-based model obtained from AGB-F and Cstr Phase 4-5 for the full MLR (a) and RF (b) models. Cumulative probability distribution between estimated maize yields and simulated maize yields for the full MLR (c) and RF (d) models are presented. Vertical lines represent median values.

Secondly, the final maize grain yield model was established using estimated AGB-F and Cstr values obtained with vegetation vigour and drought indices over the vegetative period (SOS-TOS, Table 7 and Figure 8) in order to have an assessment of maize yields prior to the end of the cropping season. When compared to simulated SARRA-O maize yields, the results showed that between 42% and 53% of yield variance was predicted by the MLR and RF models, respectively, with cv-RMSE values similar to the models calibrated over the full cropping season (cv-RRMSE below the threshold of 10%) and an overall agreement when the probability distribution curves were compared, particularly around median values for the RF model (3629 kg/ha). All in all, this means that roughly half of the maize yield variability could be explained about two months before the end of the cropping season.

*Table 7. Cross-validation metrics for early assessment of maize grain yield models using remote sensing indices integrated over the vegetative period only. *** All cv-R² are significant for p-value<0.001.*

	MLR [CWSIxTCIxSMOS]	***RF***
cv-R² [0-1]***	0.42	0.53
cv-RMSE (kg/ha)	303	271
cv-RRMSE (%)	8.57	7.65
cv-MAE (kg/ha)	236	202
d [0-1]	0.71	0.82

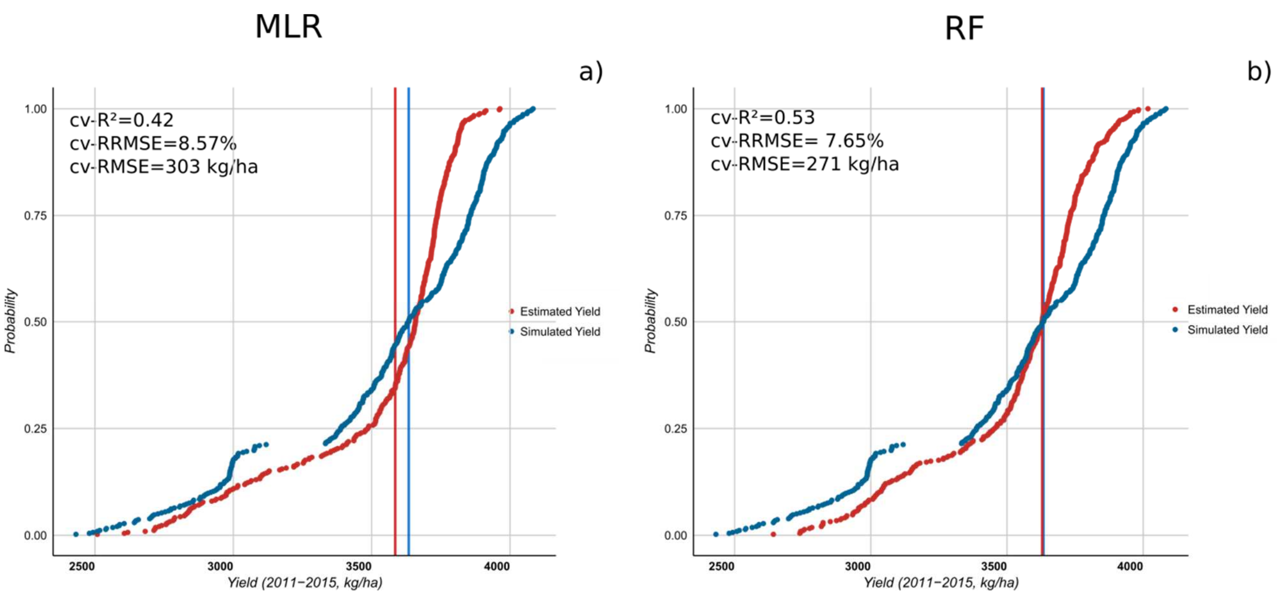

Figure 8: Cross-validation accuracy metrics and cumulative probability distribution between early estimated maize yields and simulated maize yields at the end of the season for the full MLR (c) and RF (d) models. The remote sensing-based models were calibrated using the full dataset over the vegetative period (SOS-TOS). Vertical lines represent median values.

Lastly, the robustness of the remote sensing-based models was assessed using independent maize yield data from field surveys for 2014, 2015 and 2016. Thus, maize yields were estimated from the previous MLR and RF models for 2014 and 2016, which were years not used in the calibration steps. The results are illustrated in Table 8. Overall, when compared to an independent dataset, the RF models outperformed the MLR models and depicted overall good agreement with field data, accounting for roughly 60% of the yield variability when all the cropping season was considered and 46% in early assessment with, however, high error in the second configuration (1117 kg/ha).

*Table 8. Accuracy metrics for the validation of the MLR and RF models with field yield data for 2014, 2015 and 2016. The results were compared on an annual basis on a village scale. * R^2 significant for p-value<0.1. ** R^2 significant for p-value<0.01.*

	MLR		***RF***	
	Estimation	Early assessment	Estimation	Early assessment
R^2 [0-1]	0.39*	0.10	0.59**	0.46*
RMSE (kg/ha)	823	824	637	1117
RRMSE (%)	33.4	33.4	25.9	45
MAE (kg/ha)	806	768	592	1096

5. DISCUSSION

5.1. MAIN FINDINGS

In most crop models, the basic ecophysiological processes involved in final grain yield build-up rely on an empirical reduction function applied to potential yield based on a crop water stress factor calculated during the phenological phases. The present study proposed a remote sensing-based model that went beyond traditional methods, by taking into account essential ecophysiological processes implemented in a crop growth model. Overall, we found that with the machine learning model (RF) we were able to obtain a reliable estimation of year-on-year variability in maize yields, both at the end of the season (R^2 of 0.59) and approximately two months prior to harvest (R^2 of 0.49), with a significant impact of soil water content information for Cstr estimation from the flowering to ripening phases. Our results are in line with and comparable to previous studies, both in intensive agriculture and in tropical smallholder farming systems. For instance, by also using an "uncalibrated approach", in the USA, Sibley et al. (2014) and Lobell et al.

(2015) were able to explain 37% and an average of 35% of the maize yield variance in the USA with a model using a MODIS and a Landsat-derived vegetation index, respectively. Meanwhile, in a tropical context characterized by a fragmented agricultural landscape, Mkhabela et al. (2005) obtained accuracies ranging from 5% to 68% on the scale of an ecological zone in Zimbabwe using NOAA's-AVHRR NDVI. Azzari et al. (2017) obtained a R^2 of 0.55 on a province scale using MODIS data in Zambia. As mentioned in the introduction, such studies have not yet been conducted for maize in West Africa. This suggests that low spatial resolution, such as MODIS, can be considered as a reliable alternative to high spatial resolution images for rainfed crop yield estimates and their inter-annual variability in heterogeneous agricultural landscape, as there is often a lack of sufficient dense time series of high spatial resolution.

5.2. CAN NONLINEAR STATISTICAL MODELS IMPROVE MAIZE GRAIN YIELD ESTIMATIONS COMPARED TO LINEAR STATISTICAL MODELS?

In the calibration phase, we found that the RF model was only slightly better than the MLR model, both for estimating and predicting maize yields (Table 7). However, when compared to an independent dataset, and including an additional year, we found that the machine learning model outperformed the linear model (Table 8). Despite the supremacy of linear models in the field of crop yield forecasting relying on remote sensing observations (e.g. Bolton and Friedl, 2013; Jin et al., 2017; Rasmussen, 1998), this study brings evidence in support of agro-ecosystems functioning with complex interactions between biophysical, ecological and physiological processes, and management practices, that can be far from linear. The results of this study are in line with the recent study by Jeong et al. (2016). However, as pointed out by the authors, RF is a "black box" approach in the sense that the relationships between the response variable and the predictors are not easily readable, since the algorithm is based on a set of decision trees, where each single tree is not accessible. While the results of this study seem to show promising prospects of the machine learning algorithm for crop yield forecasting in smallholder agriculture, a more in-depth analysis would be necessary to test the robustness of the method using a longer time series, including more variability in terms of agrometeorological conditions and, particularly, extreme drought or excess moisture years. Indeed, as already observed by Jeong et al. (2016) we also found a systematic tendency to underestimate high yield values and to overestimate low yield values (Figure 7 and Figure 8). This effect was attributed by the authors to an imbalance in the variance distribution of the variables, which implied a tendency for the algorithm to underestimate or overestimate extreme years, such as 2015 (Figure 7). Thus, when predicting extreme conditions, our RF model may result in a loss of accuracy. Increasing the size of the calibration dataset, with more balanced predictor variance, may help to minimize this issue. In addition, the specific patterns observed in the RF estimated AGB-F vs simulated SARRA-O AGB-F (Figure 5b) and the RF estimated yield vs simulated SARRA-O yield (Figure 7) depicted a good capacity of the RF model to render year-on-year variability, but a limited ability to retrieve spatial variability. In a context marked by high between-field yield variability due to rainfall variability, soil fertility and management practices, this is the main limitation of our study. It has been shown, for instance, for millet in Niger that spatial yield patterns are greatly determined by the variability in sowing dates (Akponikpè et al., 2011). Integrating information on spatio-temporal rainfall distribution and sowing conditions, together with vegetation and drought-related indices, can certainly help to predict spatial yield variability more precisely.

5.3. Can remote sensing information relative to soil moisture and soil water content help in better accounting for water stress in the final maize yields?

Our results showed a stronger impact of SSM-related variables in the MLR model, mainly due to a high contribution of SSM integrated over the vegetative period to maize AGB-F. This can be explained by the fact that herbaceous vegetation seed germination is highly dependent on the amount of moisture available for the seeds. In SARRA-O, germination is triggered when simulated soil water available for the plant is greater than 10 mm at the end of the day. On the other hand, for the RF model, SMOS SSM and SMADI were of greater importance for Cstr over the sensitive phenological phases for maize and went hand in hand with temperature condition indicators (TCI and CWSI), but with a limited impact on final maize grain yields. This limited impact on final maize yields can also be explained by a near absence of Cstr simulated by SARRA-O over the study period. Here, the interaction between temperature and soil moisture suggested that crop water stress depended either on a reduction in soil moisture or an increase in drought, with both having an impact on how maize copes with heat, such as evaporative cooling (Lobell et al., 2011a). While processes determining crop yield are mainly limited by the soil water content in the root zone depth, our study highlighted that information on near-surface soil moisture (0-5 cm) is already a good proxy for the water effectively available for developing vegetation. Besides the SSM variables, the Temperature Condition Index (TCI) was also revealed to be an important variable, particularly for Cstr estimation, in both the MLR and RF models (see Figure 6e and 6f). This tallied with the study by Unganai and Kogan (1998), where a strong correlation was found between TCI and maize yields during the grain filling period in Zimbabwe. While most process-based crop models rely on ambient air temperature to drive various processes such as photosynthesis, the development rate or reproductive development (Eyshi Rezaei et al., 2015), it has been shown that canopy temperature more effectively explains grain losses, particularly when heat stress occurs around sensitive phases (Siebert et al., 2014). In particular, heat stress induces a decrease in transpiration rate and thus an increase in crop canopy temperature resulting from stomatal closure. For maize, photosynthesis and reproductive processes are altered, particularly when heat stress occurs at the flowering stage or during grain filling, with a significant reduction in the grain number, a key element of final yields (Eyshi Rezaei et al., 2015). Thus, our results are along the same lines as the results of studies conducted on a field scale as well as a regional scale (Lobell et al., 2011a), and they highlighted the need to account for canopy temperature conditions in remote sensing-based yield models.

5.4. Can an early estimation of final maize yields be obtained before harvest with a remote sensing-based model?

When compared to the yields simulated with the SARRA-O crop model, both the MLR and RF models allowed an early estimation of maize grain yields (two months on average; Table 7). However, when compared to field data, only the RF model was able to explain 46% of the observed yield variability, but with an overestimation of roughly 1120 kg/ha (RRMSE of 45%, Table 8). For the MLR model, the low predictive power regarding observed maize yields came from a limited ability to accurately estimate yields for 2016. When 2016 was discarded from the analysis, the MLR model was able to account for more than 90% of the measured yield variability (not shown). This illustrates one of the main limitations of a parametric model such as MLR, which is not easily extendable to periods outside the one used for regression. For instance, variations in agrometeorological conditions may not have been included in the population from which the MLR model was derived. For the RF model, the overestimation was mainly due to the fact that relying solely on remote sensing indicators taken over the vegetative period amounted to estimating potential yields (similar to AGB-F). However, these potential yields can be drastically reduced if a heat or drought stress occurs during sensitive phases. In addition, this overestimation was not disconnected from the "uncalibrated approach", since the SARRA-O crop model simulates attainable yields according to agrometeorological constraints, but does not integrate all biotic or non-environmental factors that may lead to yield variations (Sultan et al., 2005). Despite this overestimation, whenever temporal variability is well represented, it is already useful information and constitutes an improvement in risk forecasting of a maize grain yield deficit.

5.5. Towards a scalability of remote sensing-based maize yield estimations

The approach proposed in this study relies on the calibration of a remote sensing model based on the output of a crop model, considered as a proxy for observation data, thus not requiring any ground data. For areas such as in West Africa, where ground measurements are either unreliable or not available at the right time, the performance of the "uncalibrated approach" is undoubtedly a definite option for the scalability of maize yield estimations over wider areas. However, several issues have to be addressed before extending to other areas or other types of crops.

Firstly, besides the need for reliable ground data observations for model calibration, the estimation of maize yields over larger areas also requires an accurate crop mask on a regional scale, so as to have a crop signal that is as pure as possible and to avoid bias introduced by natural or semi-natural vegetation. The launch of new sensors, such as Sentinel-2 with high spatial, temporal and spectral resolution (Drusch et al., 2012), or progress in cloud computing solutions, open up new perspectives for enhancing cropland monitoring in fragmented landscapes and mapping on a wide scale (Fritz et al., 2015; Pérez-Hoyos et al., 2017).

A second improvement that would be required is the use of vegetation vigour and drought-related indices at a higher spatial resolution, in order to be more consistent with the spatial complexity of smallholder farming systems. Here, we obtained good results with MODIS and downscaled SMOS aggregated at 4 km for a system dominated by maize with small fields in a relatively flat environment. For highland farming systems characterized by rugged terrain, very small fields, a large variety of crop species and/or the presence of mixed crops or agroforestery systems, such results probably could not be achieved without an adaptation of the proposed approach. Again, a lot of hope is placed in the new generation of high-resolution sensors, such as Sentinel-2, as well as the improvement of spatial disaggregation techniques for low-resolution product sensors, such as SMOS, making it possible to mitigate the impact of mixed-pixels on the spectral signatures of cropping areas and thus to improve crop yield estimating and forecasting in heterogeneous African farming systems (Chivasa et al., 2017).

One caveat of our study is the estimation of phenological metrics using and automatic method based on a threshold computed from the long-term mean values of the NDVI time series. However, as we used coarse resolution data, the bias in estimating crop phenology can be huge and the start of the season was more probably a translation of the response of surrounding trees and shrubs, which tended to start a month before crop vegetation (Vintrou et al., 2014). Thus, a third requirement would be not to rely on phenological metrics, but rather detect and take into account farming practices, such as sowing dates, in a spatially explicit fashion.

Last, but not least, while the focus of the study was on maize crop yields, what is even more important for many users, such as early warning systems, national agricultural statistics departments, or other institutes in charge of agricultural management and planning, is total crop production. With a view to improving food security it is, for instance, important to have reliable information on food availability, which implies food production as a function of total crop area and yield per unit area. Thus, focusing solely on final crop yields may lead to a not insubstantial underestimation of food availability. Thus, a last requirement to be met is now to promote research on food production estimations considering crop areas and yield estimates together.

6. Conclusion

A timely and robust maize grain yield model based on remote sensing, crop modelling and statistical approaches was developed in a context where ground measurements are either unreliable or not available at the right time. To this end, we adopted an “uncalibrated approach” as defined and tested recently by Burke and Lobell (2017), Lobell et al. (2015), where the output of the SARRA-O crop model, validated over our study area, was used as pseudo-ground data to estimate vegetative biomass at the flowering stage and a crop water stress index to restrain the conversion from aboveground biomass to final grain yield. Three different approaches were tested: (1) a linear (MLR) *vs* nonlinear (RF) statistical model, (2) the use of soil water content information to improve the performance of the maize yield model and (3) an estimation *vs* a forecasting approach. This study showed that a nonlinear model, such as Random Forest, outperformed a traditional linear model for maize yield estimates making it more possible to account for underlying ecophysiological processes involved in vegetation development. In addition, soil moisture information as a proxy for the soil water actually available for vegetation growth contributed to improving the RF maize yield model, particularly by impacting mainly on crop water stress (Cstr) over sensitive phases of maize development, such as the reproductive and maturation phases. Furthermore, we found that the year-on-year variability of end-of-season maize grain yields can be predicted with a good level of confidence two months before the end of the season, when only data from the vegetative period are used in the remote sensing model. The early assessment of main crop yield reduction is of great importance for improving early warning systems for food security, by mitigating the impact of food shortages on population food security and livelihoods, as well as helping in drawing up strategic planning to meet food demands. This is strengthened by the use of an "uncalibrated approach", which did not require ground measurements for calibration of the remote sensing-based yield model, which are usually considered as a significant curb to the effectiveness of crop monitoring systems in the region. However, continued efforts are needed to validate the approach presented in this study, particularly by extending the analysis to other smallholder farming systems around the world, and to move towards scalability over larger areas. Such efforts can also be supported by the use of time series from multiple new high spatial and temporal resolution sensors such as Sentinel, Venµs or Planet, which would not only significantly improve the estimation of maize grain yields and their year-on-year variability over large areas, but would also make it possible to capture more precisely the variability in yields between and within fields.

ACKNOWLEDGEMENTS

This study was supported by the VERSAO project (DAR-TOSCA 4800000915) funded by the French Space Agency (CNES) and by the European FP 7 SIGMA project (Stimulating Innovation for Global Monitoring of Agriculture) (project No. 603719). MODIS data were obtained from NASA's Earth Observing System Data and Information System (EOSDIS) (http://reverb.echo. nasa.gov), TAMSAT data were provided by the University of Reading (UK), and ECMWF agrometeorological data were obtained from EC-JRC-MARS and were created by MeteoConsult based on ECWMF (European Centre for Medium Range Weather Forecasts) model outputs. SMOS disaggregated data were kindly provided by isardSAT and were developed in the framework of the SMELLS project partially funded by the European Space Agency (contract No. 4000113413/15/I-NB). The authors are very thankful to Peter Biggins for English revisions, to colleagues based in Burkina Faso, Patrice Kouakou, Medina Karambiri and Kalifa Coulibaly, for their supervision of the field surveys, to MSc students (Akakpo Kolade, Stella Guillemot and Cyrille Ahmed Midingoyi) for their screening of the ground database and their initial work on the crop model, and to the tree anonymous reviewers for useful suggestions on earlier version of the manuscript.

REFERENCES

Adiku, S.G.K., MacCarthy, D.S., Hathie, I., Diancoumba, M., Freduah, B.S., Amikuzuno, J., Traore, P.C.S., Traore, S., Koomson, E., Agali, A., Lizaso, J.I., Fatondji, D., Adams, M., Tigana, L., Diarra, D.Z., N'diaye, O., Valdivia, R.O., 2015. Climate Change Impacts on West African Agriculture: An Integrated Regional Assessment (CIWARA), in: Handbook of Climate Change and Agroecosystems. pp. 25–73. doi:10.1142/9781783265640_0014

Akponikpè, P.B.I., Minet, J., Gérard, B., Defourny, P., Bielders, C.L., 2011. Spatial fields ' dispersion as a farmer strategy to reduce agro-climatic risk at the household level in pearl millet-based systems in the Sahel : A modeling perspective. Agric. For. Meteorol. 151, 215–227. doi:10.1016/j.agrformet.2010.10.007

Akumaga, U., Tarhule, A., Yusuf, A.A., 2017. Validation and testing of the FAO AquaCrop model under different levels of nitrogen fertilizer on rainfed maize in Nigeria, West Africa. Agric. For. Meteorol. 232, 225–234. doi:10.1016/J.AGRFORMET.2016.08.011

Allé, C.S.U.Y., Baron, C., Guibert, H., Agbossou, E.K., Afouda, A.A., 2014. Choice and risks of management strategies of agricultural calendar: application to the maize cultivation in south Benin. Int. J. Innov. Appl. Stud. 7, 1137–1147.

Azzari, G., Jain, M., Lobell, D.B., 2017. Towards fine resolution global maps of crop yields: Testing multiple methods and satellites in three countries. Remote Sens. Environ. In press. doi:10.1016/J.RSE.2017.04.014

Baron, C., Sultan, B., Balme, M., Sarr, B., Traoré, S.B., Lebel, T., Janicot, S., Dingkuhn, M., 2005. From GCM grid cell to agricultural plot: scale issues affecting modelling of climate impact. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 360, 2095–2108.

Bassu, S., Brisson, N., Durand, J.-L., Boote, K., Lizaso, J., Jones, J.W., Rosenzweig, C., Ruane, A.C., Adam, M., Baron, C., Basso, B., Biernath, C., Boogaard, H., Conijn, S., Corbeels, M., Deryng, D., De Sanctis, G., Gayler, S., Grassini, P., Hatfield, J., Hoek, S., Izaurralde, C., Jongschaap, R., Kemanian, A.R., Kersebaum, K.C., Kim, S.H., Kumar, N.S., Makowski, D., Müller, C., Nendel, C., Priesack, E., Pravia, M.V., Sau, F., Shcherbak, I., Tao, F., Teixeira, E., Timlin, D., Waha, K., 2014. How do various maize crop models vary in their responses to climate change factors? Glob. Chang. Biol. 20, 2301–2320. doi:10.1111/gcb.12520

Becker-Reshef, I., Vermote, E.F., Lindeman, M., Justice, C., 2010. A generalized regression-based model for forecasting winter wheat yields in Kansas and Ukraine using MODIS data. Remote Sens. Environ. 114, 1312–

1323. doi:10.1016/j.rse.2010.01.010
Bolton, D.K., Friedl, M., 2013. Forecasting crop yield using remotely sensed vegetation indices and crop phenology metrics. Agric. For. Meteorol. 173, 74–84. doi:10.1016/j.agrformet.2013.01.007
Breiman, 2001. Random Forest. Mach. Learn. 45, 5–32.
Burke, M., Lobell, D.B., 2017. Satellite-based assessment of yield variation and its determinants in smallholder African systems. Proc. Natl. Acad. Sci. U. S. A. 114, 2189–2194. doi:10.1073/pnas.1616919114
Chakrabarti, S., Bongiovanni, T., Judge, J., Zotarelli, L., Bayer, C., 2014. Assimilation of SMOS Soil Moisture for Quantifying Drought Impacts on Crop Yield in Agricultural Regions. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 7, 3867–3879. doi:10.1109/JSTARS.2014.2315999
Chivasa, W., Mutanga, O., Biradar, C., 2017. Application of remote sensing in estimating maize grain yield in heterogeneous African agricultural landscapes: a review. Int. J. Remote Sens. 38, 6816–6845.
Degenne, P., Lo Seen, D., 2016. Ocelet: Simulating processes of landscape changes using interaction graphs. SoftwareX 5, 89–95. doi:10.1016/j.softx.2016.05.002
Diarisso, T., Corbeels, M., Andrieu, N., Djamen, P., Tittonell, P., 2015. Biomass transfers and nutrient budgets of the agro-pastoral systems in a village territory in south-western Burkina Faso. Nutr. Cycl. Agroecosystems 101, 295–315. doi:10.1007/s10705-015-9679-4
Didan, K., 2015. MOD13Q1 MODIS/Terra Vegetation Indices 16-Day L3 Global 250m SIN Grid V006. NASA EOSDIS Land Processes DAAC. http://doi.org/10.5067/MODIS/MOD13Q1.006.
Dingkuhn, M., Baron, C., Bonnal, V., Maraux, F., Sarr, B., Clopes, A., Forest, F., 2003. Decision support tools for rainfed crops in the Sahel at the plot and regional scales, in: Struif, B.T., Wopereis, M. (Eds.), Decision Support Tools for Smallholder Agriculture in Sub-Saharan Africa : A Practical Guide. pp. 127–139.
Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F., Bargellini, P., 2012. Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services. Remote Sens. Environ. 120, 25–36. doi:10.1016/j.rse.2011.11.026
Duncan, J.M.A., Dash, J., Atkinson, P.M., 2015. Elucidating the impact of temperature variability and extremes on cereal croplands through remote sensing. Glob. Chang. Biol. 21, 1541–1551. doi:10.1111/gcb.12660
Durand, J.L., Delusca, K., Boote, K., Lizaso, J., Manderscheid, R., Weigel, H.J., Ruane, A.C., Rosenzweig, C., Jones, J., Ahuja, L., Anapalli, S., Basso, B., Baron, C., Bertuzzi, P., Biernath, C., Deryng, D., Ewert, F., Gaiser, T., Gayler, S., Heinlein, F., Kersebaum, K.C., Kim, S.H., Müller, C., Nendel, C., Olioso, A., Priesack, E., Villegas, J.R., Ripoche, D., Rötter, R.P., Seidel, S.I., Srivastava, A., Tao, F., Timlin, D., Twine, T., Wang, E., Webber, H., Zhao, Z., 2018. How accurately do maize crop models simulate the interactions of atmospheric CO2 concentration levels with limited water supply on water use and yield? Eur. J. Agron. 100, 67–75. doi:10.1016/J.EJA.2017.01.002
Everingham, Y., Sexton, J., Skocaj, D., Inman-Bamber, G., 2016. Accurate prediction of sugarcane yield using a random forest algorithm. Agron. Sustain. Dev. 36, 1–9. doi:10.1007/s13593-016-0364-z
Eyshi Rezaei, E., Webber, H., Gaiser, T., Naab, J., Ewert, F., 2015. Heat stress in cereals: Mechanisms and modelling. Eur. J. Agron. 64, 98–113. doi:10.1016/j.eja.2014.10.003
FAO/IIASA/ISRIC/ISSCAS/JRC, 2012. Harmonized World Soil Database (version 1.2).
Fieuzal, R., Marais Sicre, C., Baup, F., 2017. Estimation of corn yield using multi-temporal optical and radar satellite data and artificial neural networks. Int. J. Appl. Earth Obs. Geoinf. 57, 14–23. doi:10.1016/j.jag.2016.12.011
Forkuor, G., Conrad, C., Thiel, M., Zoungrana, B., Tondoh, J., 2017. Multiscale Remote Sensing to Map the Spatial Distribution and Extent of Cropland in the Sudanian Savanna of West Africa. Remote Sens. 2017, Vol. 9, Page 839 9, 839. doi:10.3390/RS9080839
Fritz, S., See, L., Mccallum, I. a N., You, L., Bun, A., Moltchanova, E., Duerauer, M., Albrecht, F., Schill, C., Perger, C., Havlik, P., Mosnier, A., Thornton, P., Wood-sichra, U., Herrero, M., Becker-Reshef, I., 2015. Mapping global cropland and field size. Glob. Chang. Biol. 21, 1–13. doi:10.1111/gcb.12838
Gaetano, R., Lebourgeois, V., Dupuy, S., Jolivot, A., Butler, S., Baron, C., Guillemot, S., Lo Seen, D., Bégué, A., 2016. Presentation of the Burkina Faso (Koumbia) site activities, in: JECAM/GEOGLAM Science Meeting. Kiev, Ukraine, p. Poster.
Grömping, U., 2006. Relative Importance for Linear Regression in R: The Package relaimpo. J. Stat. Softw. 17, 1–27. doi:10.18637/jss.v017.i01
Groten, S.M.E., 1993. NDVI—crop monitoring and early yield assessment of Burkina Faso. Int. J. Remote Sens. 14, 1495–1515. doi:10.1080/01431169308953983
Gruhier, C., de Rosnay, P., Hasenauer, S., Holmes, T., de Jeu, R., Kerr, Y., Mougin, E., Njoku, E., Timouk, F., Wagner, W., Zribi, M., 2010. Soil moisture active and passive microwave products: intercomparison and evaluation over a Sahelian site. Hydrol. Earth Syst. Sci. 14, 141–156. doi:10.5194/hess-14-141-2010
Guan, K., Sultan, B., Biasutti, M., Baron, C., Lobell, D.B., 2017. Assessing climate adaptation options and uncertainties

for cereal systems in West Africa. Agric. For. Meteorol. 232, 291–305. doi:10.1016/J.AGRFORMET.2016.07.021

Guan, K., Sultan, B., Biasutti, M., Baron, C., Lobell, D.B., 2015. What aspects of future rainfall changes matter for crop yields in West Africa? Geophys. Res. Lett. 42, 8001–8010. doi:10.1002/2015GL063877

Guillemot, S., 2016. Utilisation de données d'estimation des précipitations par satellite avec un modèle de culture, application à la province de Tuy (Burkina Faso). Montpellier, France.

Holzman, M.E., Rivas, R., Piccolo, M.C., 2014. Estimating soil moisture and the relationship with crop yield using surface temperature and vegetation index. Int. J. Appl. Earth Obs. Geoinf. 28, 181–192. doi:10.1016/j.jag.2013.12.006

Jackson, R.D., Idso, S.B., Reginato, R.J., Pinter, P.J., 1981. Canopy temperature as a crop water stress indicator. Water Resour. Res. 17, 1133–1138.

Jeong, J.H., Resop, J.P., Mueller, N.D., Fleisher, D.H., Yun, K., Butler, E.E., Timlin, D.J., Shim, K.M., Gerber, J.S., Reddy, V.R., Kim, S.H., 2016. Random Forests for Global and Regional Crop Yield Predictions. PLoS One 11, e0156571. doi:10.1371/journal.pone.0156571

Jin, Z., Azzari, G., Burke, M., Aston, S., Lobell, D., 2017. Mapping Smallholder Yield Heterogeneity at Multiple Scales in Eastern Africa. Remote Sens. 9, 931. doi:10.3390/rs9090931

Johnson, M.D., 2014. An assessment of pre- and within-season remotely sensed variables for forecasting corn and soybean yields in the United States. Remote Sens. Environ. 141, 116–128. doi:10.1016/J.RSE.2013.10.027

Johnson, M.D., Hsieh, W.W., Cannon, A.J., Davidson, A., Bédard, F., 2016. Crop yield forecasting on the Canadian Prairies by remotely sensed vegetation indices and machine learning methods. Agric. For. Meteorol. 218–219, 74–84. doi:10.1016/j.agrformet.2015.11.003

Kerr, Y., Waldteufel, P., Wigneron, J.P., Delwart, S., Cabot, F., Boutin, J., Escorihuela, M.J., Font, J., Reul, N., Gruhier, C., Juglea, S.E., Drinkwater, M.R., Hahne, A., Martín-Neira, M., Mecklenburg, S., 2010. The SMOS Mission: New Tool for Monitoring Key Elements ofthe Global Water Cycle. Proc. IEEE 98, 666–687. doi:10.1109/JPROC.2010.2043032

Kogan, F., 1995. Application of vegetation index and brightness temperature for drought detection. Adv. Sp. Res. 15, 91–100.

Lebourgeois, V., Dupuy, S., Vintrou, E., Ameline, M., Butler, S., Bégué, A., 2017. A Combined Random Forest and OBIA Classification Scheme for Mapping Smallholder Agriculture at Different Nomenclature Levels Using Multisource Data (Simulated Sentinel-2 Time Series, VHRS and DEM). Remote Sens. 9, 259. doi:10.3390/rs9030259

Leroux, L., Baron, C., Zoungrana, B., Traoré, S.B., Lo Seen, D., Bégué, A., 2016. Crop Monitoring Using Vegetation And Thermal Indices For Yield Estimates: Case Study Of A Rainfed Cereal In Semi-Arid West Africa. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 9, 347–362. doi:10.1109/JSTARS.2015.2501343

Liaw, A., Wiener, M., 2002. Classification and Regression by randomForest. R news 2, 18–22. doi:10.1177/154405910408300516

Lobell, D.B., Bänziger, M., Magorokosho, C., Vivek, B., 2011a. Nonlinear heat effects on African maize as evidenced by historical yield trials. Nat. Clim. Chang. 1, 42–45. doi:10.1038/nclimate1043

Lobell, D.B., Schlenker, W., Costa-Roberts, J., 2011b. Climate trends and global crop production since 1980. Science 333, 616–20. doi:10.1126/science.1204531

Lobell, D.B., Thau, D., Seifert, C., Engle, E., Little, B., 2015. A scalable satellite-based crop yield mapper. Remote Sens. Environ. 164, 324–333. doi:10.1016/j.rse.2015.04.021

Marteau, R., Sultan, B., Moron, V., Alhassane, A., Baron, C., Traoré, S.B., 2011. The onset of the rainy season and farmers' sowing strategy for pearl millet cultivation in Southwest Niger. Agric. For. Meteorol. 151, 1356–1369. doi:10.1016/j.agrformet.2011.05.018

Maselli, F., Romanelli, S., Bottai, L., Maracchi, G., 2000. Processing of GAC NDVI data for yield forecasting in the Sahelian region. Int. J. Remote Sens. 21, 3509–3523. doi:10.1080/014311600750037525

Merlin, O., Al Bitar, A., Walker, J.P., Kerr, Y., 2010. An improved algorithm for disaggregating microwave-derived soil moisture based on red, near-infrared and thermal-infrared data. Remote Sens. Environ. 114, 2305–2316. doi:10.1016/J.RSE.2010.05.007

Meroni, M., Marinho, E., Sghaier, N., Verstrate, M., Leo, O., 2013. Remote Sensing Based Yield Estimation in a Stochastic Framework — Case Study of Durum Wheat in Tunisia. Remote Sens. 5, 539–557. doi:10.3390/rs5020539

Mkhabela, M.S., Mkhabela, M.S., Mashinini, N.N., 2005. Early maize yield forecasting in the four agro-ecological regions of Swaziland using NDVI data derived from NOAA's-AVHRR. Agric. For. Meteorol. 129, 1–9. doi:10.1016/J.AGRFORMET.2004.12.006

Oettli, P., Sultan, B., Baron, C., Vrac, M., 2011. Are regional climate models relevant for crop yield prediction in West Africa? Environ. Res. Lett. 6, 014008. doi:10.1088/1748-9326/6/1/014008

Pérez-Hoyos, A., Rembold, F., Kerdiles, H., Gallego, J., 2017. Comparison of Global Land Cover Datasets for Cropland Monitoring. Remote Sens. 9, 1118. doi:10.3390/rs9111118

R Core Team, 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rasmussen, M.S., 1998. Developing simple, operational, consistent NDVI-vegetation models by applying environmental and climatic information. Part II: Crop yield assessment. Int. J. Remote Sens. 19, 119–139. doi:10.1080/014311698216468

Rasmussen, M.S., 1992. Assessment of millet yields and production in northern Burkina Faso using integrated NDVI from the AVHRR. Int. J. Remote Sens. 13, 3431–3442. doi:10.1080/01431169208904132

Ray, D.K., Mueller, N.D., West, P.C., Foley, J.F., 2013. Yield Trends Are Insufficient to Double Global Crop Production by 2050. PLoS One 8, e66428. doi:10.1371/journal.pone.0066428

Ray, D.K., Ramankutty, N., Mueller, N.D., West, P.C., Foley, J.A., 2012. Recent patterns of crop yield growth and stagnation. Nat. Commun. 3, 1293. doi:10.1038/ncomms2296

Roudier, P., Alhassane, A., Baron, C., Louvet, S., Sultan, B., 2016. Assessing the benefits of weather and seasonal forecasts to millet growers in Niger. Agric. For. Meteorol. 223, 168–180. doi:10.1016/j.agrformet.2016.04.010

Sánchez, N., González-Zamora, Á., Piles, M., Martínez-Fernández, J., 2016. A New Soil Moisture Agricultural Drought Index (SMADI) Integrating MODIS and SMOS Products: A Case of Study over the Iberian Peninsula. Remote Sens. 8, 1–25. doi:10.3390/rs8040287

Sandholt, I., Rasmussen, K., Andersen, J., 2002. A simple interpretation of the surface temperature/vegetation index space for assessment of surface moisture status. Remote Sens. Environ. 79, 213–224. doi:10.1016/S0034-4257(01)00274-7

Shiferaw, B., Prasanna, B.M., Hellin, J., Bänziger, M., 2011. Crops that feed the world 6. Past successes and future challenges to the role played by maize in global food security. Food Secur. 3, 307–327. doi:10.1007/s12571-011-0140-5

Sibley, A.M., Grassini, P., Thomas, N.E., Cassman, K.G., Lobell, D.B., 2014. Testing Remote Sensing Approaches for Assessing Yield Variability among Maize Fields. Agron. J. 106, 24. doi:10.2134/agronj2013.0314

Siebert, S., Ewert, F., Eyshi Rezaei, E., Kage, H., Graß, R., 2014. Impact of heat stress on crop yield—on the importance of considering canopy temperature. Environ. Res. Lett. 9, 1–8. doi:10.1088/1748-9326/9/4/044012

Son, N.T., Chen, C.F., Chen, C.R., Chang, L.Y., Minh, V.Q., 2012. Monitoring agricultural drought in the Lower Mekong Basin using MODIS NDVI and land surface temperature data. Int. J. Appl. Earth Obs. Geoinf. 18, 417–427. doi:10.1016/j.jag.2012.03.014

Sultan, B., Baron, C., Dingkuhn, M., Sarr, B., Janicot, S., 2005. Agricultural impacts of large-scale variability of the West African monsoon. Agric. For. Meteorol. 128, 93–110. doi:10.1016/j.agrformet.2004.08.005

Sultan, B., Guan, K., Kouressy, M., Biasutti, M., Piani, C., Hammer, G.L., Mclean, G., Lobell, D., 2014. Robust features of future climate change impacts on sorghum yields in West Africa. Environ. Res. Lett. 9, 13. doi:10.1088/1748-9326/9/10/104006

Sultan, B., Roudier, P., Quirion, P., Alhassane, A., Muller, B., Dingkuhn, M., Ciais, P., Guimberteau, M., Traoré, S.B., Baron, C., 2013. Assessing climate change impacts on sorghum and millet yields in the Sudanian and Sahelian savannas of West Africa. Environ. Res. Lett. 8, 9. doi:10.1088/1748-9326/8/1/014040

Tarnavsky, E., Grimes, D., Maidment, R., Black, E., Allan, R.P., Stringer, M., Chadwick, R., Kayitakire, F., 2014. Extension of the TAMSAT satellite-based rainfall monitoring over Africa and from 1983 to present. J. Appl. Meteorol. Climatol. 53, 2805–2822. doi:10.1175/JAMC-D-14-0016.1

Traoré, S.B., Alhassane, A., Muller, B., Kouressy, M., Somé, L., Sultan, B., Oettli, P., Siéné Laopé, A.C., Sangaré, S., Vaksmann, M., Diop, M., Dingkuhn, M., Baron, C., 2011. Characterizing and modeling the diversity of cropping situations under climatic constraints in West Africa. Atmos. Sci. Lett. 12, 89–95. doi:10.1002/asl.295

Tucker, C.J., 1985. Satellite Remote Sensing of Total Herbaceous Biomass Production in the Senegalese Sahel : 1980-1984. Remote Sens. Environ. 17, 233–249.

Tucker, C.J., Holben, B.., Elgin, J.., McMurtrey, J.., 1980. Relationship of spectral data to grain yield variation. Photogramm. Eng. Remote Sensing 46, 657–666.

Unganai, L.S., Kogan, F.N., 1998. Drought Monitoring and Corn Yield Estimation in Southern Africa from AVHRR Data. Remote Sens. Environ. 63, 219–232. doi:10.1016/S0034-4257(97)00132-6

Vintrou, E., Bégué, A., Baron, C., Alexandre, S., Lo Seen, D., Traoré, S.B., 2014. A Comparative Study on Satellite and Model-Based Crop Phenology in West Africa. Remote Sens. 6, 1367–1389. doi:10.3390/rs6021367

Wan, Z., 2015. Modis Land Surface Temperature Products Users' Guide. MOD11A2 MODIS/Terra L. Surf. Temp. Emiss. 8-Day L3 Glob. 1 km Grid SIN V006. NASA EOSDIS L. Process. DAAC. doi:10.5067/MODIS/MOD11A2.006

Zuur, A.F., Ieno, E.N., Elphick, C.S., 2010. A protocol for data exploration to avoid common statistical problems. Methods Ecol. Evol. 1, 3–14. doi:10.1111/j.2041-210X.2009.00001.x

Supplementary Material S1. List of the R software packages used in the study : references, the main uses and functions employed are provided.

Package	**Description**	**Reference**	**Use**	**Function**
car	A companion to Applied Regression	Fox, J. and Weisberg, S. (2011)	Statistical analysis	vif
DAAG	Data Analysis and Graphics Data and Functions	Maindonald, J.H and Braun, W.J. 2015	Statistical analysis	cv.lm
doParallel	Foreach Parallel Adaptor for the 'parallel' Package	Microsoft Corporation and Wetson, S. 2017	Remote sensing processing	foreach
gdalUtils	Wrappers for the Geospatial Data Abstraction Library (GDAL) Utilities	Greenberg, J.A. and Mattiuzzi, M. 2015	Remote sensing processing	gdalwarp
ggplot2	Elegant Graphics for Data Analysis	Wickham, H. 2009	Graphic	Ggplot, geom_point, geom_smooth, geom_abline, geom_vline, stat_ecdf, geom_bar
ggthemes	Extra Themes, Scales and Geoms for 'ggplot2'	Arnold, J.B. 2017	Graphic	theme, scale_color, scale_fill
greenbrown	Land surface phenology and trend analysis	Forkel et al. 2013	Remote sensing processing	PhenologyRaster
hydroGOF	Goodness-of-fit functions for comparison of simulated and observed hydrological time series	Zambrano-Bigiarini, M. 2017	Statistical analysis	rmse, rd, mae
quantreg	Quantile Regression	Koenker, R. 2017	Statistical analysis	rq
signal	Signal processing	Signal developers. 2013	Remote sensing processing	sgolayfilt
randomForest	Breiman and Cutler's Random Forests for Classification and Regression	Liaw, A. and Wiener, M. (2002)	Statistical analysis	randomForest
raster	Geographic data analysis and modelling	Hijmans, R.J. 2016	Remote sensing processing	raster, stack, extent, crop, calc, writeRaster, shapefile, stackApply, values, projectRaster
relaimpo	Relative Importance for Linear Regression in R	Grömping, U. 2006	Statistical analysis	calc.relimp, boot.relimp
reshape2	Reshaping Data with the reshape Package	Hickham, H. 2007	Graphic	melt
stats	The R Stats package	R Core Team. 2017	Statistical analysis	lm
yarrr	A Companion to the e-Book "YaRrr!: The Pirate's Guide to R"	Phillips, N. 2017	Graphic	piratplot

Supplementary Material S2. Equation of the vegetation and drought indices used as explanatory variables.

Vegetation Index

$$NDVI = (\rho_{NIR} - \rho_R) \div (\rho_{NIR} + \rho_R)$$

Where ρ_R and ρ_{NIR} are the surface reflectance in the Red and Near Infra Red bands.

Drought Indices

$$TCI = (LST_{multi\,max} - LST_i) \div (LST_{multi\,max} - LST_{multi\,min})$$

Where LST_i is the smoothed weekly Land Surface Temperature (LST), and $LST_{multi\,max}$ and $LST_{multi\,min}$ its multi-year maximum and minimum, respectively.

$$CWSI = (LST_i - LST_{mini}) \div (LST_{maxi} - LST_{mini})$$

Where LST_i is the smoothed weekly Land Surface Temperature (LST), and LST_{maxi} and LST_{mini} its maximum and minimum (i.e. maximum and minimum for the week i), respectively.

$$TVDI = (LST_i - LST_{min}) \div (a + bNDVI_i - LST_{min})$$

Where LST_i is the smoothed be-weekly Land Surface Temperature (LST), $NDVI_i$ is the smoothed be-weekly Normalised Difference Vegetation Index, LST_{min} is the minimum temperature observed in the NDVI/LST space regression (wet edge) and $LST_{max} = a + bNDVI$ is the maximum LST temperature for a given value of NDVI, a and b are the intercept and the slope of the dry edge, modeled as a linear fit to the data.

$$SMADI = SMOS_i \times (LST_i \div NDVI_i)$$

Where $SMOS_i$ is the smoother be-weekly soil surface moisture, LST_i and $NDVI_i$ its smoother be-weekly LST and NDVI, respectively.