



**HAL**  
open science

# Combining UAV and Sentinel-2 Imagery for Estimating Millet FCover in a Heterogeneous Agricultural Landscape of Senegal

Ibrahima Diack, Serigne Mansour Diene, Leroux Louise, Diouf Abdoul Aziz, Heuclin Benjamin, Roupsard Olivier, Letourmy Philippe, Audebert Alain, Sarr Idrissa, Diallo Moussa

## ► To cite this version:

Ibrahima Diack, Serigne Mansour Diene, Leroux Louise, Diouf Abdoul Aziz, Heuclin Benjamin, et al.. Combining UAV and Sentinel-2 Imagery for Estimating Millet FCover in a Heterogeneous Agricultural Landscape of Senegal. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2024, 17, pp.7305-7322. 10.1109/jstars.2024.3373508 . hal-04562624

**HAL Id: hal-04562624**

**<https://hal.inrae.fr/hal-04562624>**

Submitted on 29 Apr 2024








**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

# Combining UAV and Sentinel-2 Imagery for Estimating Millet FCover in a Heterogeneous Agricultural Landscape of Senegal

Ibrahima Diack , Serigne Mansour Diene , Leroux Louise , Diouf Abdoul Aziz , Heuclin Benjamin , Rouspard Olivier , Letourmy Philippe , Audebert Alain , Sarr Idrissa , and Diallo Moussa 

**Abstract**—In recent decades, remote sensing has been shown to be useful for crop cover monitoring over smallholder agricultural landscapes, such as agroforestry parklands. However, the fraction of green vegetation cover (FCover) has received little attention. Indeed, the collection of FCover ground data representative of the within-field heterogeneity is time-consuming. Thus, this article aims to bridge this gap by proposing an original methodological framework combining FCover data derived from unmanned aerial vehicle (UAV) and Sentinel-2 (S2) images for estimating millet FCover at the landscape scale in an agroforestry parkland of the groundnut basin of Senegal during the 2021 and 2022 cropping seasons. UAV-based FCover was computed over a 3 m × 3 m grid using a thresholding approach for six dates over the cropping seasons and then used as ground observation for the upscaling of millet FCover at the landscape scale with S2 data. Various spectral vegetation

indices and textural features were derived from S2, and several modeling approaches based on machine learning algorithms were benchmarked. Our results showed that the modeling approach using the full-time series in combination with a random forest algorithm was able to explain 73% (root mean square error = 12.13%) of the UAV-FCover variability after validation in 2021 and 2022. In addition, UAV images are suitable for consistent monitoring of millet FCover over heterogeneous agricultural landscapes by training S2 satellite images. To further check its robustness, this approach should be tested for different crops and practices across a variety of agricultural landscapes in sub-Saharan Africa.

**Index Terms**—Drone, fraction of green vegetation cover (FCover), machine learning, satellite, smallholder agriculture, sub-Saharan Africa (SSA).

## I. INTRODUCTION

**M**ORE than one-third of people suffering from hunger in the world in 2021 were in Sub-Saharan Africa (SSA) [1]. The SSA population is expected to reach 3.1 billion in 2100 [2] while the continent is also challenged by the need to reduce its footprint on the environment and increase food production without further encroachment of natural ecosystems [3] and comply with the sustainable development goals (SDG) [4]. In addition, agricultural productivity is still far below its potential, which is mainly a result of inherent low soil fertility, nutrient limitations and suboptimal agricultural practices [5], [6]. Consequently, this trend is expected to worsen in the coming years. That is why the sustainable intensification of SSA smallholder agriculture, i.e., the enhancement of agricultural productivity together with the maintenance of ecosystem services and resilience to shocks, is recognized as a critical component to reach a proper balance across all dimensions of the SDG [6].

Agroforestry is defined as the combination of trees or woody shrubs and crops or pastures on the same land space [7]. This practice is recognized as a sustainable intensification practices [8] and is considered one of the most promising strategies to increase food production on land that currently has low tree cover without additional deforestation [9], [10], [11], [12]. Trees and woody shrubs have been an integral part of SSA agricultural landscapes for centuries [13], and there exists a wide diversity of agroforestry systems shaped by environmental conditions as well as farming strategies and practices (e.g., [14]). Some prominent examples of SSA agroforestry systems cited by Mbow et al. [11] are multistory home gardens on Mt. Kilimanjaro in Tanzania, cocoa systems in Ivory Coast, and rotational

Manuscript received 8 October 2023; revised 23 January 2024; accepted 19 February 2024. Date of publication 5 March 2024; date of current version 3 April 2024. The work of Ibrahima Diack and Serigne Mansour Diene was also supported by Cirad. This work was supported in part by the SustainSahel Project, in part by the European Union's Horizon 2020 research and innovation program under Grant 861974 (SustainSAHEL), and in part by the ground data collection has been supported by the ELISA project funded by Cirad. (Corresponding authors: Ibrahima Diack; Leroux Louise.)

Ibrahima Diack is with the Cheikh Anta DIOP University, 15532 Dakar, Senegal, also with the CIRAD, UPR AIDA, 15532 Dakar, Senegal, and also with the Centre de Suivi Ecologique, 15532 Dakar, Senegal (e-mail: diack.ib@gmail.com).

Serigne Mansour Diene is with the Cheikh Anta DIOP University, 15532 Dakar, Senegal, also with the CIRAD, UMR AGAP Institut, 1386 This Escale, Senegal, and also with the LMI IESOL, Centre IRD-ISRA de Bel Air, 15532 Dakar, Senegal.

Leroux Louise is with the CIRAD, UPR AIDA, 30772-00100 Nairobi, Kenya, also with the AIDA, Univ Montpellier, CIRAD, 34398 Montpellier, France, and also with the IITA, 30772-00100 Nairobi, Kenya (e-mail: louise.leroux@cirad.fr).

Diouf Abdoul Aziz is with the Centre de Suivi Ecologique, 15532 Dakar, Senegal.

Heuclin Benjamin and Letourmy Philippe are with the CIRAD, UPR AIDA, F-34398 34398 Montpellier, France, and also with the AIDA, University Montpellier, CIRAD, 34398 Montpellier, France.

Audebert Alain is with the CIRAD, UMR AGAP Institut, 34398 Montpellier, France, and also with the UMR AGAP Institut, Univ Montpellier, CIRAD, INRAE, Institut Agro, 34398 Montpellier, France.

Rouspard Olivier is with the CIRAD, UMR Eco&Sols, Dakar 1386, Senegal, also with the Eco&Sols, Univ Montpellier, CIRAD, INRAE, Institut Agro, IRD, Montpellier, France, and also with the LMI IESOL, Centre IRD-ISRA de Bel Air, 1386 Dakar, Senegal.

Sarr Idrissa is with the Faculty of science, Cheikh Anta DIOP University, 5085 Dakar, Senegal.

Diallo Moussa is with the Polytechnic Institute of the University Cheikh Anta Diop, 5085 Dakar, Senegal.

This article has supplementary downloadable material available at <https://doi.org/10.1109/JSTARS.2024.3373508>, provided by the authors.

Digital Object Identifier 10.1109/JSTARS.2024.3373508

woodlots in Kenya. A specific case of agroforestry systems well known in dry tropical Africa, particularly in Senegal, is the *Faidherbia albida* (*F. albida*) parklands [15]. By providing food, fuelwood or income diversification options, *F. albida* parklands can contribute to improving farm household food security, livelihoods, and resilience (e.g., [16]). In addition, *F. albida* parklands have been shown to increase crop productivity (e.g., [17], [18], and [19]) through an improvement in nutrient concentrations [17], soil carbon [20], water infiltration [21], and a modification of the microclimate below the crown [22]. However, *F. albida* parklands are also highly diversified in terms of species composition and structure [23], which introduces a very large within-field heterogeneity in terms of crop productivity, with a decay-gradient effect with distance from the trunk [19]. For instance, using a geostatistical approach and multispectral indices from unmanned aerial vehicle (UAV) imagery, Roupard et al. [24] identified that the distance of influence of *F. albida* trees on millet productivity was 17 m, while in more recent studies, Leroux et al. [16] showed that the provision of multiple ecosystem services by *F. albida* trees is optimal within a range of 10 m from the trunk. Hence, to define effective policy interventions to meet the SDG, reliable estimates of crop productivity in this type of challenging agricultural landscape are needed.

Because of its large and repetitive coverage, satellite remote sensing has been widely used for crop cover monitoring and crop productivity assessment over large areas (e.g., [25], [26], [27]). However, until very recently, estimation of crop productivity by satellite remote sensing was still challenging for smallholder agricultural landscapes characterized by a small field size [28], important cloud cover during the crop growing season [29] and high intrafield heterogeneity due to the presence of isolated trees, diversity in soil conditions and crop management practices [30]. Some promising results have been obtained for smallholder farming systems by exploiting the capabilities of new satellites, such as PlanetScope or Sentinel-2 (S2) (e.g., [31], [32], [33], [34]). However, these studies failed to reproduce the variability observed in farmers' fields. In addition, despite their high spatiotemporal resolution, these optical systems are still hampered by the atmospheric conditions during the growing season.

In the case of heterogeneous agricultural landscapes such as agroforestry parklands, the combined use of satellite images with UAV images offers a tangible solution since subdecimeter UAV images can be acquired at low-altitudes below clouds [35]. In addition, UAV images can assess within-field variability even in the heterogeneous agricultural landscape of smallholder farming systems [35], [36]. However, Pdua et al. [37] reviewed the use of UAVs for agroforestry system monitoring and showed that most of the studies focused on northern systems. Few examples focusing on crop monitoring with UAVs in agroforestry systems in SSA can be found, such as the study of [38] for mango orchards in Senegal or the study of [24] in an *F. albida* parkland in Senegal. Consequently, the potential of UAVs for crop monitoring in SSA agroforestry parklands remains to be fully exploited.

In recent years, visible and multispectral imaging sensors have been integrated into UAVs to estimate crop vegetation cover because of their low cost, efficiency, convenience, and spatial accuracy [39], [40], [41]. At the same time, the recent advances made in the past 10 years in terms of satellite spatiotemporal

resolution, such as the one offered by the S2 constellation, have opened new avenues for the monitoring of smallholder agriculture by reducing mixed pixel issues [35], [42]. However, the S2 constellation is still limited in terms of spatial resolution [i.e., 10 m for the visible and near infrared (NIR) bands] to capture the within-field variability in smallholder agricultural landscapes. Consequently, we can assume that one way of optimizing spatial resolution issues is to combine several sources of information, such as UAV images and freely available satellite images, such as S2. This represents an unprecedented opportunity to improve and strengthen crop monitoring in heterogeneous agricultural landscapes. Even if the processing of images from multiple sources is challenged by the different radiometric characteristics and temporal and spatial resolutions, their combination can add value by cumulating their respective advantages. Specifically, the potential of UAV images to fill the gap between ground data information and satellite information for habitat mapping [43], [44], invasive plant species mapping [45] or agricultural monitoring [35], [46] has been recently demonstrated.

One of the vegetation cover biophysical parameters that are important for agricultural monitoring but have raised little attention in the scientific literature on satellite-based agricultural monitoring in SSA is the fraction of green vegetation cover (FCover). FCover is defined as the fraction of ground covered by green vegetation and is a canopy architecture parameter and hence is used to quantify the spatial extent of photosynthetically active vegetation. FCover plays a crucial role in ecological or biophysical processes [47]. It is, for instance, an important parameter for predicting crop yield or other canopy biophysical variables, such as the leaf area index or gross primary productivity, and it is also an indicator of surface phenology (e.g., [35], [48], [49]). Satellite images have been shown to be very efficient for crop FCover estimation in various ecosystems [45]. Likewise, other studies have also shown the potential of UAV-based images to offer spatially continuous quasi-ground truth FCover data [35], [44], [50].

Many approaches have been developed in the literature to quantify biophysical variables, such as FCover from (proxy-) remote sensing data, which can be categorized into statistical (variable-driven) and physical-based (radiometric data-driven) categories [51]. However, most of the studies rely on vegetation indices (VIs) to estimate crop biophysical parameters. [52] indicated that VIs from UAVs have good correlations with vegetation cover parameters. Some authors have also shown the benefit of considering textural indices [35], [50], [53], [54]. For instance, in a study on mapping the fractional cover of invasive species combining S2 and UAV images, [45] observed that textural indices allowed a higher accuracy for the mapping of species occurrence. The same conclusion was reached by Hall et al. [35] for mapping maize FCover with UAV images in heterogeneous smallholder farming systems of Ghana and by Kwak and Park [54] for different crop type cover mapping in Korea. In terms of statistical approaches, machine learning algorithms are now increasingly used in crop monitoring studies mainly due to their abilities to deal with nonlinear relationships between the spectral information and the targeted crop variables. In Burkina Faso, Leroux et al. [55] demonstrated the ability

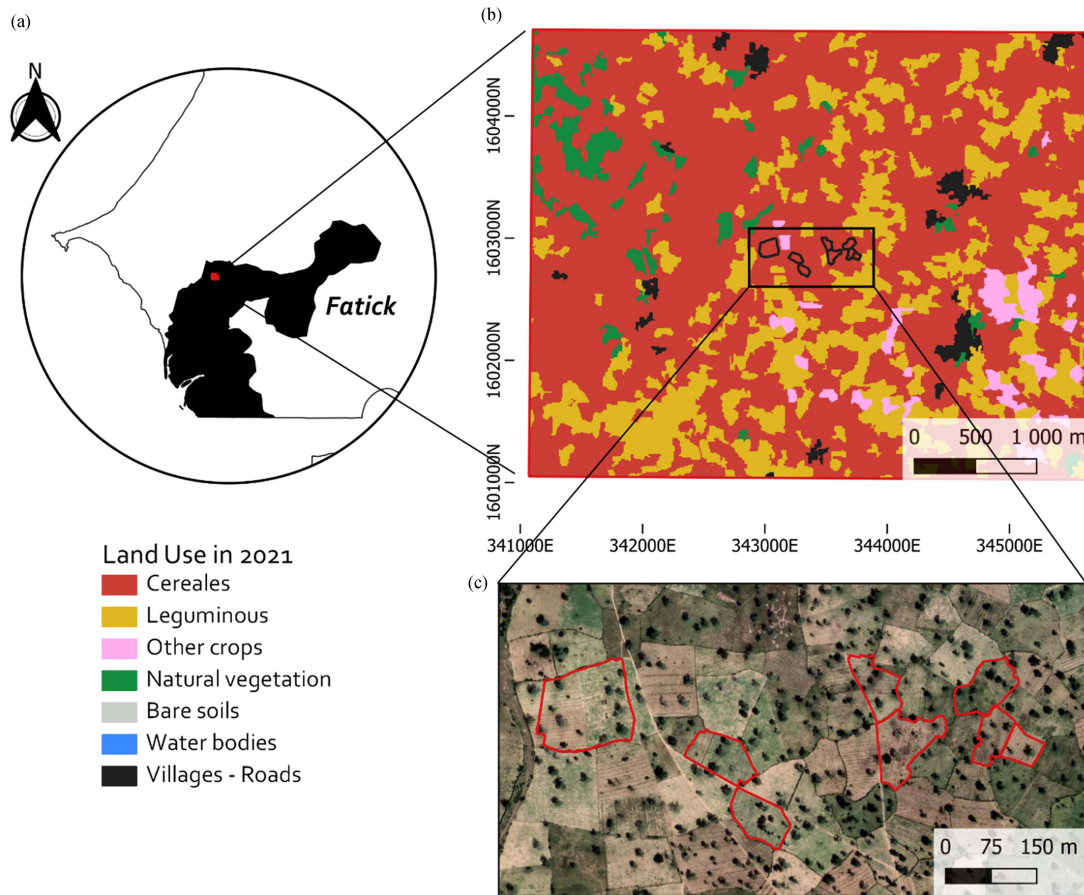


Fig. 1. Study area. (a) Location of the study area (red square) in the Fatick region of Senegal, (b) location of the 8 farmers' plots monitored over the 2021 cropping season (black boundaries) on the land use map of 2021 derived from S2 images and a RF algorithm, and (c) zoom over the 8 farmers' plots of 2021 (red boundaries) displayed on a Google Satellite ©background.

of the random forest (RF) algorithm in maize yield estimation. Later, in the *F. albida* parklands of Senegal, Gbodjo et al. [31] showed that deep learning algorithms provide promising options for millet yield estimation, while Leroux et al. [33] used a gradient boosting machine (GBM) algorithm to disentangle the millet yield variability drivers. Under lower latitudes, in Ghana, Hall et al. [35] also demonstrated the robustness of support vector machine (SVM) for maize FCover estimation with UAV images. Among the possibilities to further increase the accuracy of machine learning approaches, the use of multitemporal images over the season instead of using only one single image has been demonstrated [54]. Model stacking is another way of improving machine learning approaches. Model stacking is an ensemble method that allows one to improve the model predictions by combining the output of different individual models [56]. While model stacking has been used in various applications linked to agricultural monitoring using remote sensing data (e.g., [57], [58], [59]), to the best of the authors knowledge, it has received very little attention for the estimation of crop cover parameters in heterogeneous smallholder agricultural landscapes. However, we can mention the work of the authors in [60], [61], and [62] on using a stacking approach for crop type mapping in South Africa, Zimbabwe, and Mali, respectively.

As one additional step toward the improvement of crop productivity monitoring to meet the challenges of smallholder

farmers in SSA, this study aims to propose an original methodological framework combining UAV and S2 images using up-to-date machine learning approaches for estimating crop FCover at the landscape scale. In particular, we aimed to (1) benchmark the traditional approach based on one machine learning algorithm applied on a unique date to two alternative approaches considering a full-time series over the crop cycle and model stacking, (2) compare the performance of the combined use of UAV and S2 images in estimating FCover with FCover estimated with S2 data only, and (3) assess the crop FCover at the landscape scale using the best approaches resulting from (1) and (2). We argue that agroforestry systems inherently offer a wide range of conditions to test those methods; therefore, this study was conducted in a parkland in Senegal with pearl millet under *F. albida* trees (see Fig. 1).

## II. MATERIALS AND METHODS

### A. Study Area

The study was conducted in 2021 and 2022 in a *F. albida* agroforestry parkland near the village of Niakhar in the Fatick department located in the groundnut basin of Senegal [see Fig. 1(a)]. The study area covers approximately 16 km<sup>2</sup> [see Fig. 1(b)]. The climate is Sudano-Sahelian, with one rainy season lasting from July to October and with August and September being

TABLE I  
VIS USED IN THIS STUDY AND CALCULATED FROM OPTICAL BANDS OF UAV  
AND S2 IMAGES

Index name	Formula	Reference
NDVI	$(\text{NIR}-\text{Red})/(\text{NIR}+\text{Red})$	[68]
GCVI	$(\text{NIR}/\text{Green})-1$	[67]
GNDVI	$(\text{NIR}-\text{Green})/(\text{NIR}+\text{Green})$	[69]
ExGI	$\frac{(2 \times \text{RGB})}{(\text{Green}+\text{Red}+\text{Blue})}$	[70]

the wettest months. The annual rainfall in our area of interest calculated from the weather station in the “Faidherbia-Flux” experiment<sup>1</sup> was 482 mm in 2021 and 822 mm in 2022. Soils are mainly sandy soils located over dunes and flat areas, while slightly more clayish soils are located in lowland areas and interdunes [63]. In the study area, as in most parts of the country, the livelihoods of rural populations are centered on small-scale rainfed agriculture with low external input use and low soil fertility and nutrient availability, with fertility decreasing with increasing distance from homestead fields where farmers tend to concentrate their nutrient resources [64]. The main crops cultivated in the region are pearl millet (*Pennisetum glaucum* (L.) R. Br); used as a staple crop, and groundnut (*Arachis hypogaea* L.), used as cash crop. Pearl millet and groundnut are cultivated in a typical biennial rotation in bush fields, while homestead fields are continuously cultivated with pearl millet. In a typical cropping season, pearl millet is sown around May-June, before the first rain and harvested from October to November. The study area is also characterized by a tree-based cropping system dominated by *F. albida*, representing 42% of the trees among the 60 tree species inventoried in the study area [23]. *F. albida* is a leguminous nitrogen-fixing species with phreatophytic behavior and a reverse phenology [65]. *F. albida* is well known to improve water and nutrient availability [19] and to foster microclimate modification under the tree crown [22] and the concentration of animal dejections. *F. albida* sheds leaves at the end of the dry season, which provides green manure. Overall, these features contribute to increasing soil fertility and crop yield under tree crowns, creating a “fertility island” also termed the “albida effect” [19], [66].

## B. Data and Preprocessing

1) *S2 Data*: The S2 satellites (S2A and S2B) compose the European Space Agency (ESA) optical high-resolution mission for the Copernicus Program. Time series of S2 level-2 (S2-L2A) were downloaded from the THEIA Platform.<sup>2</sup> Level 2 A (L2A) products provide surface reflectances corrected for atmospheric effects and are supplied with a mask of clouds and their shadows. The S2 multispectral imager provides a wide range of multispectral observations from 13 bands in the visible, NIR and shortwave infrared parts of the electromagnetic spectrum

with spatial resolutions of 10, 20, and 60 m depending on the wavelength (see Section III). The repetitiveness of sensors S2A and S2B alternates over a ten-day period, offering a revisit of five days at the equator under the same viewing conditions [67]. In this study, six S2 images were acquired over the 2021 cropping season and two images during the 2022 cropping season (see Fig. 2) and only the visible spectral bands [red, green, and blue (RGB)] and the NIR (band 8) with 10-m spatial resolution were considered. For each acquisition date, four VIs and eight textural features (TFs) were calculated.

a) *Computation of VIs*: The four selected bands (i.e., visible and NIR) were used to calculate four VIs, namely, the normalized difference vegetation index (NDVI), green normalized difference vegetation index (GNDVI), green chlorophyll vegetation index (GCVI), and excess green index (ExGI) (see Table I). These four VIs have been chosen due to their sensitivity to chlorophyll contents of crop cover, and hence the presence of green vegetation. NDVI is by far the most common VI used for crop vegetation monitoring. In addition, in a recent study conducted in the same area, GNDVI has been shown to be more sensitive to pearl millet aboveground biomass than traditional VIs, such as NDVI [33]. The main strength of GCVI compared with the other VIs is that it can be used for a wide range of plant species [68]. Lastly, excess green is a color VI that is therefore very sensitive to the presence of the green color on (UAV or satellite) images. In addition, it relies only on R, G, and B bands, consequently it is a cost-effective solution for vegetation monitoring and mapping since it can be derived from any sensors. VIs derived from the red-edge band (i.e., normalized difference red-edge index) have been shown to be good proxies of chlorophyll content toward the end of the growing season when crops reach their maturity, and for high-density canopy cover. Due to (1) the very low sowing density of pearl millet over our study area, with a maximum leaf area index on average below 2, and (2) similar performance of NDRE compared with NDVI to estimate pearl millet aboveground biomass at harvest over our study area (not shown), no index based on the red-edge band has been included in this study to allow the applicability of the methodological workflow to a broader range of sensors (e.g., Planet).

b) *Calculation of TFs*: For each VI, eight TFs were extracted in all directions and on  $3 \times 3$  pixel windows using the gray-level co-occurrence matrix (GLCM) developed by Haralick et al. [72] for remote sensing processing. The GLCM has already been applied in many domains, such as invasive species monitoring [54]. TFs are also used to reduce noise effects. The extracted TFs are mean, variance, homogeneity, contrast, dissimilarity, entropy, second moment, and correlation.

2) *UAV Data*: Eight and four farmers’ millet fields were monitored over the 2021 and 2022 cropping seasons, respectively. The UAV system used in this study is a DJI Phantom 4 equipped with a multispectral camera acquiring information on six spectral bands (RGB NIR, and red edge, see Section IV). Since it was not possible to cover all the fields on the same day due to battery limitations, UAV image acquisition was performed over a five-day period centered on the S2 acquisition dates (see Fig. 2). UAV flights were carried out in open sky conditions and low wind speeds between noon and 4 P.M. local time. The DJI Gs Pro application was used to define the flight missions.

<sup>1</sup>[Online]. Available: <https://iped.info/wikiObsSN/?Faidherbia-Flux>

<sup>2</sup>[Online]. Available: <https://theia.cnes.fr/atdistrib/rocket#/search?collection=SENTINEL2>

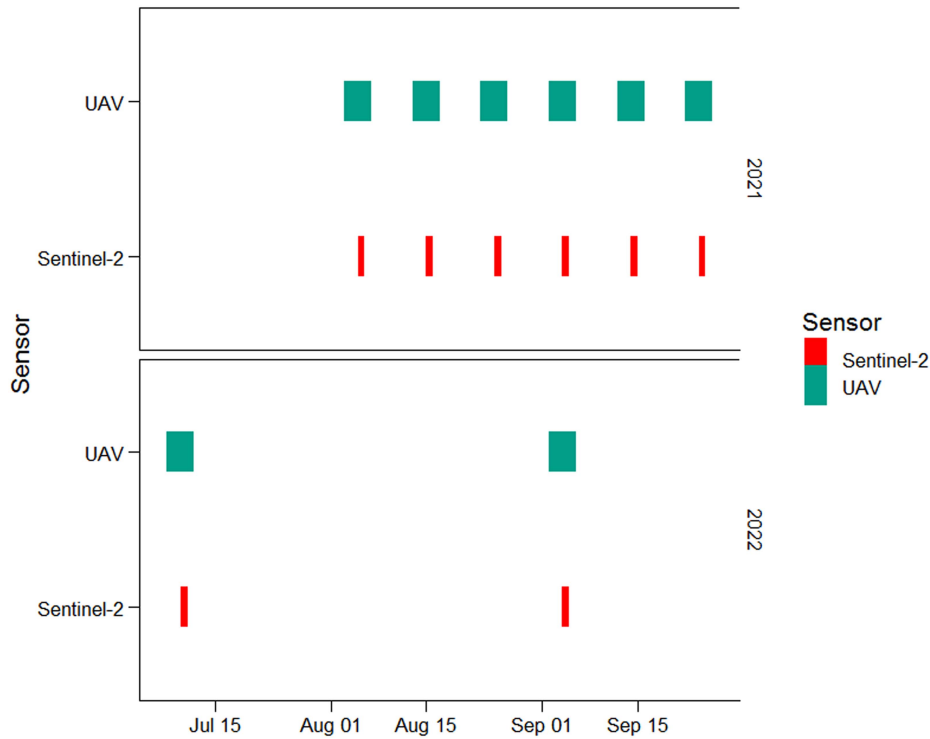


Fig. 2. Overview of the acquisition dates of S2 and UAV images over the 2021 and 2022 cropping seasons. UAV acquisitions were performed over a five-day period centered on the S2 acquisitions.

Images were acquired every 2 s at a flight altitude of 25 m with a longitudinal overlap rate of 80% and lateral overlap rate of 70% to obtain gapless orthoimages. Based on these flight settings, a total of 1300–2000 images were acquired per flight, and the UAV image spatial resolution was 1.3 cm. Then, the UAV images were mosaicked to obtain multispectral orthoimages, and these orthoimages were coregistered with S2 images to align them. Since the UAV image is already aligned with the S2 image, the corresponding pixels should be relatively superposed.

3) *Land Use and Land Cover Data (LULC)*: A LULC [see Fig. 1(b)] of the study area was used to locate millet fields in 2021 and to extrapolate the FCover over millet areas. The LULC map was obtained from a ground survey conducted before the harvest time of the 2021 cropping season and a S2 time series. A RF algorithm [73] implemented on the Google Earth Engine cloud-based platform was used to implement the LULC classification [74]. The global accuracy was 88%, and the user and producer accuracy for the cereal class were 80% and 87%, respectively. Then, we assumed that individual or contiguous pearl millet fields, having similar biophysical and management characteristics, would have similar spectral signatures and consequently could be grouped under a unique “millet patch.” Those millet patches were derived from the intersection of (1) an object-based segmentation into homogeneous patches using the multitemporal S2 NDVI data and (2) the 2021 LULC map. Object-based large-scale segmentation was performed using the MeanShift algorithm implemented in the Orfeo Toolbox.<sup>3</sup> Finally, to extract the main LULC class in each patch and select only patches of millet, majority voting was applied.

<sup>3</sup>[Online]. Available: <https://www.orfeo-toolbox.org/>

### C. Methodology

For the sake of clarity, the methodological workflow for the S2-UAV FCover modeling and assessment is presented in Fig. 3. First, the observed FCover was derived from the UAV images (see Section II-C1). Then, the S2-UAV FCover was modeled and assessed by analyzing the statistical relationships between S2 VIs and TFs and the UAV-based FCover (see Section II-C2). To that end, three different machine learning algorithms were tested: RF, GBM, and SVM. In addition, 3 modeling options were compared: 1) a baseline option where the models were calibrated on a date-by-date basis, and two optimized options 2) model stacking and 3) modeling considering the full-time series (see Section II-C2). Finally, the S2-UAV FCover was assessed against accuracy metrics and an FCover derived from the SNAP algorithm (see Section II-C1) and spatialized over the study area (see Section II-C4).

1) *UAV-Based Fcover*: UAV-based FCover (hereafter UAV-FCover) was used in this study as the calibration and validation observation to upscale FCover estimation at the landscape scale. When applied to remote sensing images, FCover represents the percentage of green pixels out of the total pixels. FCover determination was performed based on the following four steps.

- 1) Coregistration between UAV orthomosaic images and S2 images to correct the misalignment between UAV and S2 data. Coregistration is a way to increase the accuracy of directly georeferenced UAV orthoimages to other high-resolution aerial or satellite imagery [75]. To that end each individual S2 image (i.e., date) was considered as a reference while each corresponding UAV image as the

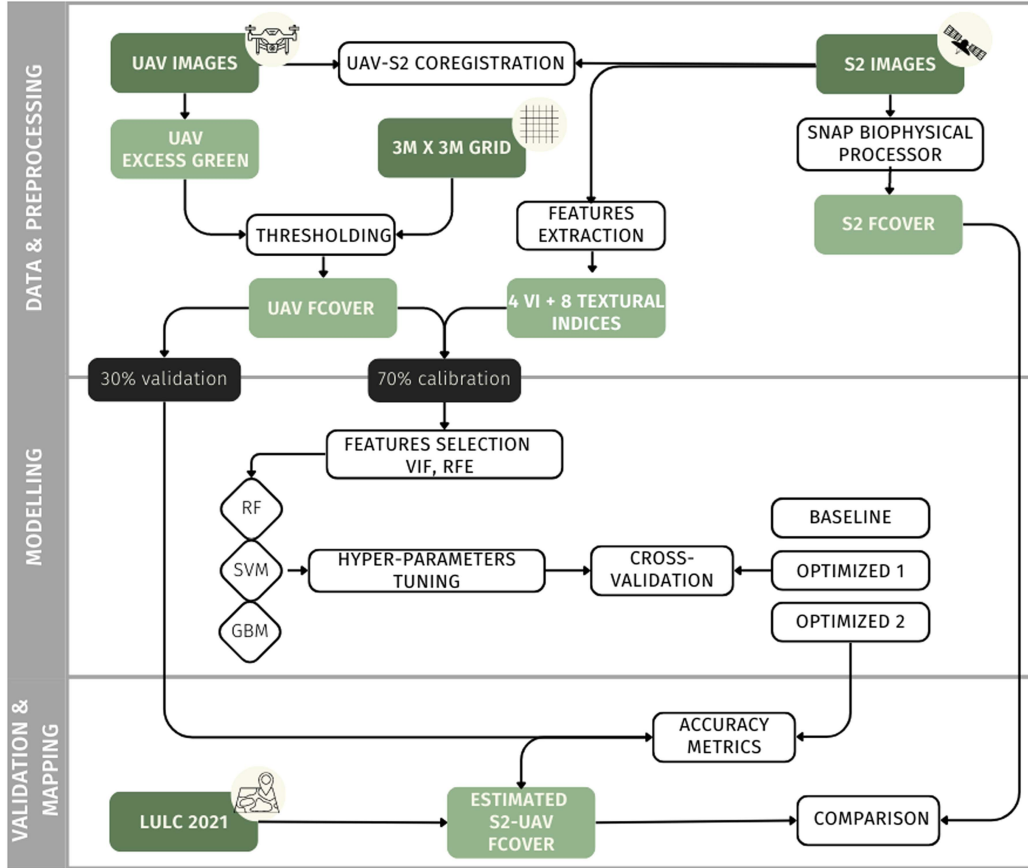


Fig. 3. Methodological workflow used to estimate pearl millet FCover from UAV and S2 images where baseline corresponds to the modeling on a date-by-date basis, optimized 1 corresponds to the first optimized option for modeling, the model stacking, and optimized 2 to the second optimized option, the modeling on the full-time series.

image with an offset to be corrected with respect to S2. To that end, a maximum absolute radius (shift) was set to 3 with an increment equal to 1.

- 2) Calculation of the ExGI (see Table I) on each UAV image. Snchez et al. [76] demonstrated that ExGI is a powerful index to discriminate vegetation due to its simplicity and its satisfactory mean accuracy and SD accuracy at 30 and 60 m flight altitudes for any image acquisition date.
- 3) Binarization of UAV-ExGI images based on a threshold approach defined iteratively for each acquisition date to discriminate green vegetation from bare soil. The threshold was defined based on visual inspection of RGB composite images. The basic principle of this method is to set the threshold of the ExGI and compare it with each pixel to obtain the vegetation information. When the pixel value is larger than this threshold, it is a vegetation pixel; otherwise, it is a nonvegetation pixel [77].
- 4) A 3 m grid was constructed, and then the percentage of green vegetation in each grid was computed to obtain the FCover (see Fig. 4).

2) *S2-UAV FCover Modeling*: The mean of each S2 VI and TF (see Section 2.2.1) was extracted over the 3-m grid. In our modeling framework, the VIs and TFs were used as independent variables, while the UAV-FCover was used as the dependent variable (see Fig. 3).

a) *Variables selection*: To avoid overfitting of our models, the number of VIs and TFs was reduced. To that end, two techniques were applied. First, multicollinearity is known to influence the parameters of variable selection. The variable inflation factor [VIF; (1)] was calculated and used to test all the VIs and TFs for multicollinearity and remove highly correlated variables. A stepwise approach was used, in which VIF was recalculated at each step, and the variables with the highest VIF were dropped until all VIF values were smaller than 10. Then, the remaining variables were used in a recursive feature elimination (RFE) algorithm to finally select the optimal number of variables and increase the computational efficiency. RFE is a backward feature selection algorithm, meaning that the number of variables is reduced at each step by removing the variables that decrease the model quality

$$VIF_i = \frac{1}{1 - R_i^2} \quad (1)$$

where  $VIF_i$  stands for the VIF of the variable  $i$  and  $R_i^2$  stands for unadjusted coefficient of determination for regressing the  $i$ th independent variable on the remaining ones.

b) *Tested machine learning algorithms*. *RF*: The RF is a popular and powerful machine learning algorithm proposed by Breiman [73] and based on the ensemble learning technique

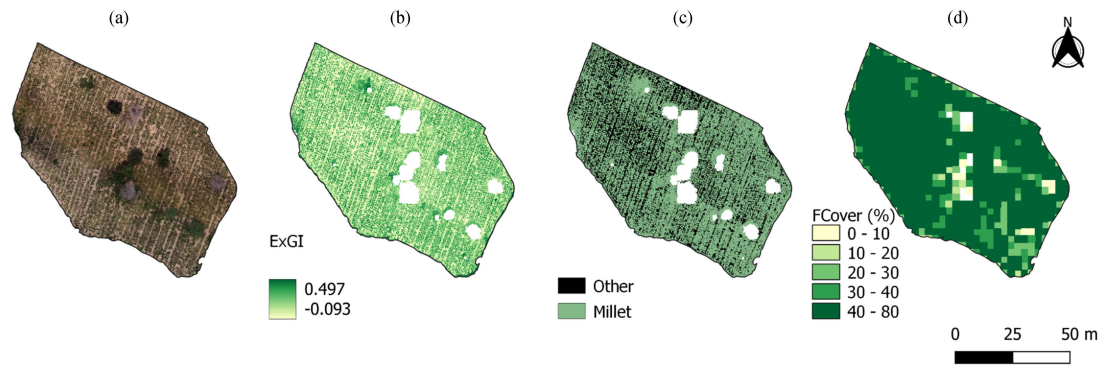


Fig. 4. Illustration of the UAV-FCover steps for one field and for the image of 04-09-2021, with (a) UAV RGB image, (b) ExGI used for the binarization based on a threshold approach, (c) resulting binarization, and (d) FCover estimated from (c) on a 3-m grid, where the percentage of millet pixels within a 3-m grid has been computed.

(bagging). It is a nonparametric algorithm based on the combination of decisions from multiple decision trees. Among its main advantages, RF is easy to implement, robust to outliers, and stable and limits overfitting, which explains that it is now widely used in crop monitoring, both for crop yield modeling and cropland mapping included in smallholder agricultural landscapes (e.g., [31], [55], [78]).

*Gradient boosting machine:* The GBM is a machine learning algorithm that is less popular so far in crop monitoring than RF (e.g., [78]). GBM is an ensemble learning technique that minimizes overfitting risks and optimizes the predictive performance through the combination of a large number of simple trees [28]. GBM is also a nonparametric approach that is relatively insensitive to outliers and able to account for nonlinear interactions between dependent and independent variables or between independent variables.

*Support vector machine:* SVM also belongs to the machine learning algorithm category. While SVM is widely used to solve classification problems [79], [80], it is rarely applied to regression problems. When applied to a regression problem, SVM is called support vector regression (SVR). SVR is also a nonparametric algorithm that seeks to find a hyperplane that best fits the observation data points in a continuous space. To that end, the independent variables are mapped to a high-dimensional feature space to find the hyperplane that maximizes the distance between the hyperplane and the closest data points while minimizing the prediction error. SVR is also able to address nonlinear and complex relationships between the dependent variable and independent variables by using a kernel function to map the data to a higher-dimensional space. To improve the modeling accuracy, the hyperparameters of each algorithm were tuned based on a grid search assessing the top-performing combination: *mtry* for RF, *C* and *Sigma* for SVM and *interaction.depth*, and *n.trees*, *shrinkage*, and *n.minobsinnode* for GBM (see Table S3).

*c) Modeling options. Date-by-date modeling (Baseline option):* In the baseline option, each machine learning algorithm was calibrated and validated on a date-by-date basis, resulting in the calibration-validation of 15 models during the 2021 cropping season (3 algorithms  $\times$  5 dates). By crossing all the possible algorithms  $\times$  dates, this is the most detailed modeling framework that can be setup. In addition, it will help in the determination

of the best algorithm but also of the most sensitive period for FCover estimation over the cropping season.

*Model stacking (Optimised option 1):* Model stacking is an ensemble method that allows improving the model predictions by combining the output of different individual models, called base learners, and running them through a meta-learner to generate final predictions that are generally more robust than the predictions coming from the individual base learners [56]. In this study, the base learners were the S2-UAV FCover predicted by RF, GBM, and SVM, and the UAV-FCover was used as the dependent variable. The meta-learner was based on a multiple linear regression algorithm. Linear regression is usually used for the meta-learner because it does not require tuning hyperparameters [81]. The model stacking was constructed for the date reaching the best accuracies in the baseline option.

*Modeling on full-time series (Optimised option 2):* Instead of considering each date as an individual dataset for modeling, all the dates were pulled together to form a unique dataset, and the date information was considered an additional independent variable. By doing this, we increased the size of the dataset and the range of FCover variability. Hence, we assumed that it will improve the robustness of the modeling framework while taking better advantage of the potential of machine learning algorithms. Consequently, date information, VIs and TFs time series were used together for FCover modeling and to gain insight into the impacts of date information on the model accuracy. The modeling on the full-time series was applied on the best performing algorithm in the baseline option.

### 3) Assessment and Validation of S2-UAV FCover Modeling:

*a) Accuracy metrics:* The performance of the different modeling options was assessed using the coefficient of determination  $R^2$  and the root mean square error (RMSE) when compared with UAV-based FCover. In the calibration phase, a fivefold cross validation with five repetitions was adopted to avoid overfitting of the models since all the data were used for both training and validation (70% of the total dataset). This means that the data are equally assigned to five equally sized subdatasets, called folds. Each fold is subsequently removed, while the remaining data are used to fit the regression model and to predict the missing data. This step is repeated five times. In the



validation phase (30% of the total dataset), the S2-UAV FCover estimations were directly compared with the UAV-FCover. The best modeling option, i.e., the one with the largest  $R^2$  and the smallest RMSE, was selected for comparison with the FCover estimated from the SNAP algorithm and for the spatialization over the study area (see Section II-C4).

*b) Comparison with the FCover estimated from the SNAP algorithm:* The S2-UAV FCover was compared against a FCover obtained from the Sentinel Simplified Level 2 Product Prototype Processor embedded in the ESA Sentinel Application Platform (SNAP) that has been designed to retrieve biophysical parameters from S2 observations. The SNAP biophysical processor was chosen in this study due to its simplicity and efficiency for estimating vegetation biophysical variables across different ecosystems (e.g., [82], [83], [84]). The biophysical modules of SNAP allow for the retrieval of biophysical parameters (i.e., leaf area index, fraction of absorbed photosynthetically active radiation, canopy chlorophyll, canopy water content, and FCover) based on artificial neural networks trained with radiative transfer model simulations from PROSPECT+SAIL [85]. We applied the SNAP algorithm to estimate the FCover of the six dates over the 2021 cropping season using the eight S2 reflectance bands (B3, B4, B5, B6, B7, B8A, B11, and B12, see Table S1). The SNAP-FCover was rescaled at a 10-m spatial resolution using the nearest neighbor resampling technique. The FCover values were then extracted on a 3-m grid to match the S2-UAV FCover estimation resolution, and the agreement between the two FCover estimations was examined using  $R^2$  and RMSE accuracy metrics (see Section II-C3).

#### Estimation of the S2-UAV FCover in 2022

To test the robustness of the approach, the best option (i.e., baseline option, optimized option 1 and optimized option 2) was applied for the two dates of the 2022 cropping season and compared with the 2022 UAV-based FCover.

*4) S2-UAV FCover Spatialization Over the Study Area:* The modeling option with the best accuracy ( $R^2$  and RMSE) was used to estimate the S2-UAV FCover over the study area by using the pearl millet patch map (see Section II-B3). The model was applied on the five dates of the 2021 cropping season to estimate the spatial and temporal distribution of pearl millet FCover. The “raster” [86], “sp” [87], “rgdal” [88], ranger [89], “RF” [90], “rpart” [91], “car” [92], and tidyverse [93] packages from R software version 4.1 [94] were used during the process of FCover model calibration, validation and spatialization. QGIS 3.18 was used to produce the different maps, while all the statistical graphics were plotted with the ggplot2 [95] package from R software.

### III. RESULTS

#### A. UAV-FCover

The UAV-FCover for the six acquisition dates along the 2021 cropping season are shown in Fig. 5. The UAV-FCover exhibited high temporal variability following the pearl millet development cycle, with coefficient of variation values of 71.64%. The UAV-FCover increased from 5 August, 2021, corresponding to the emergence date, up to 4 September, 2021, corresponding to the

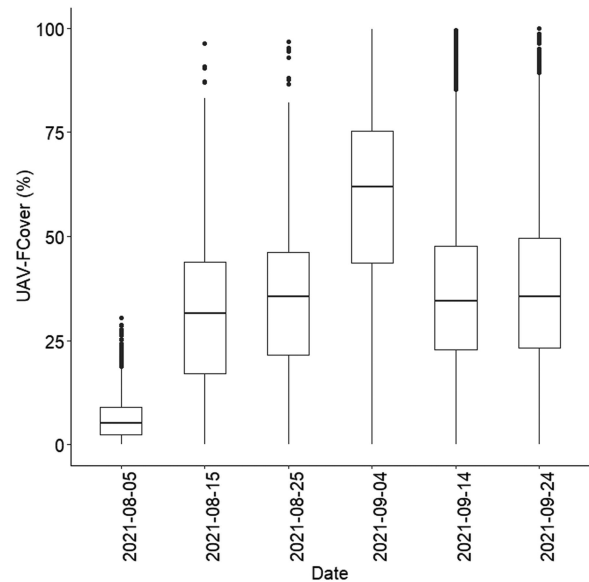


Fig. 5. Temporal variability of the UAV-based FCover over the pearl millet 2021 cropping season. The UAV-based FCover was obtained from a threshold approach applied to ExGI UAV images.

TABLE II  
 $R^2$  AND RMSE OBTAINED BY THE MODEL STACKING (I.E., OPTIMIZED OPTION 1) RUN FOR THE DATE WITH THE HIGHEST ACCURACY IN THE BASELINE OPTION—5 AUGUST, 2021

Algorithms	Date	$R^2$ (Val)	RMSE (Val)
RF	2021/08/05	0.69	2.76
GBM	2021/08/05	0.67	2.83
SVM	2021/08/05	0.65	2.91
Mean		<b>0.67</b>	<b>2.83</b>
Model stacking	2021/08/05	<b>0.68</b>	<b>2.7</b>

Only the results of the validation (Val) phase are presented.

start of the maturing phase with the maximum aboveground biomass, and then decreased over the last two dates, corresponding to the maturing phase up to harvest. Spatial variability was also observed with a higher range of values for the date of maximum vegetation (i.e., 4 September, 2021).

#### B. S2-UAV FCover Date-by-Date Modeling

Table S4 presents the results of the variable selection steps. Overall, GCVI and ExGI were the two VIs consistently selected for the six dates. For TFs no clear pattern were observed since all TFs were selected at one time or another. On the 32 potential TFs (8 TFs  $\times$  4 VIs), between 12 and 14 TFs were selected depending on the date. Fig. 6 presents the results of the calibration and validation phase for the baseline option, i.e., estimation of the S2-UAV FCover on a date-by-date basis. On the basis of  $R^2$  and RMSE, Fig. 6 shows that the performance ranged between 0.60 and 0.83 for  $R^2$  (2.5%–12% for RMSE) in calibration and between 0.48 and 0.69 for  $R^2$  (2.76–14.28% for RMSE) in validation. Among the machine learning algorithms tested, RF showed higher accuracies. GBM and SVM performed similarly. When the performance is analyzed regarding the growing season

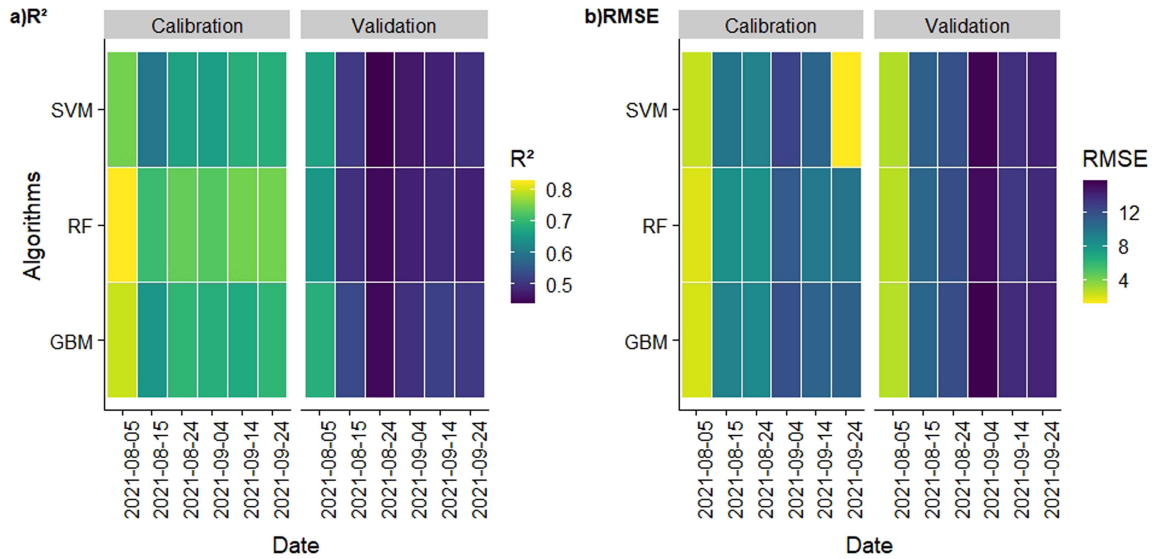


Fig. 6. S2-UAV FCover estimation for the baseline option where the S2-UAV FCover has been estimated on a date-by-date basis. (a)  $R^2$  and (b) RMSE are provided for the calibration and validation phases for each tested algorithm (RF, SVM, and RF) and for each of the five dates over the 2021 cropping season.

the beginning of the growing season (i.e., around the emergence stage) showed the highest accuracy.

### C. S2-UAV FCover Modeling Based on the Optimized Options

1) *Model Stacking*: The results of the model stacking (i.e., optimized option 1) are presented in Table II. The model stacking was tested for 5 August, 2021, the date with the highest accuracy in the baseline option. The  $R^2$  and RMSE values for each of the base learners (i.e., each algorithm) were similar to the  $R^2$  and RMSE values for the meta learner. The  $R^2$  and RMSE values for each of the base learners (i.e., each algorithm) were similar to the  $R^2$  and RMSE values for the meta learner ( $R^2 = 0.68$ , RMSE = 2.73%). Hence, overall, our results showed that the use of model stacking did not significantly improve the accuracy, suggesting that RF, GBM, and SVM are robust enough for pearl millet FCover estimation in our study area.

2) *Modeling on the Full-Time Series*: Fig. 7 presents the results of optimized option 2, i.e., using the full-time series as a unique dataset and based on the RF algorithm since it gave the best result in the baseline option. The model achieved an accuracy of 0.86% in calibration (RMSE = 8.84%) and 0.73% in validation (RMSE = 12.13%) [see Fig. 7(a)], hence outperforming the baseline option and optimized option 1. The density distribution of the predicted S2-UAV FCover showed an increase in the FCover values from the beginning of the cropping season to early September and then a decrease in the FCover values corresponding to leaf senescence. The FCover distribution was also more heterogeneous when moving toward the end of the cropping season (see Fig. 7). The visual comparison of the UAV-FCover and the predicted S2-UAV FCover showed overall good agreement in terms of both spatial and temporal variability [see Fig. 7(c)].

### D. Assessment of S2-UAV FCover Modeling

1) *Comparison With FCover Estimates From the SNAP Algorithm*: The results obtained by the best modeling approach of this study (i.e., optimized option 2: RF + full-time series)

was compared against the FCover computed with the SNAP algorithm (see Fig. 8). The SNAP-based FCover showed a weak relationship when compared with the reference UAV-FCover [Fig. 8(a);  $R^2=0.18$ ; RMSE = 21.47]. A systematic overestimation of the SNAP-FCover was observed, particularly for high FCover values. Consequently, the comparison between the S2-UAV FCover and the SNAP-FCover showed low agreement as well [Fig. 8(b);  $R^2=0.15$ , RMSE = 19.17%]. The S2-UAV FCover seemed more suited than the SNAP-FCover when the vegetation density was low, with FCover  $\leq 40\%$ . Above this threshold, the UAV FCover and SNAP-FCover performed more similarly (see Supplementary Material Table S5).

2) *Estimation of the FCover in 2022*: The best modeling approach of this study (i.e., optimized option 2: RF + full-time series) was then applied over the 2022 cropping season (see Fig. 9). When considering the whole cropping season, the calibration and validation results were similar to those obtained for the 2021 cropping season ( $R^2 = 0.73$  and RMSE = 13.16% in validation). Analysis of Fig. 9 shows two main groups of data ranging from 0% to 12% corresponding to the emergence stage (10 July, 2022) and ranging from 35% to 75% corresponding to the preheading stage (04 September, 2022). This result also showed that the method can be used for FCover estimation throughout the cropping season ( $R^2 = 0.64$  and RMSE = 14.8% for 10 July, 2022 and 04 September, 2022, respectively, in validation, see Fig. S1 in Supplementary Material).

### E. Spatial and Temporal Variability of S2-UAV FCover Over the Study Area

Optimized option 2 (RF + full-time series) was used to map the spatiotemporal variability in the FCover over our study area (see Fig. 10) based on the S2-UAV FCover and LULC data (see Section II-B3). High spatiotemporal variability in the FCover was observed. Around the emergence date (2021-08-15), low FCover values ( $<16\%$ ), were observed, followed by an increase in FCover values up to the start of the maturing stage (4 September, 2021), where the maximum FCover values were

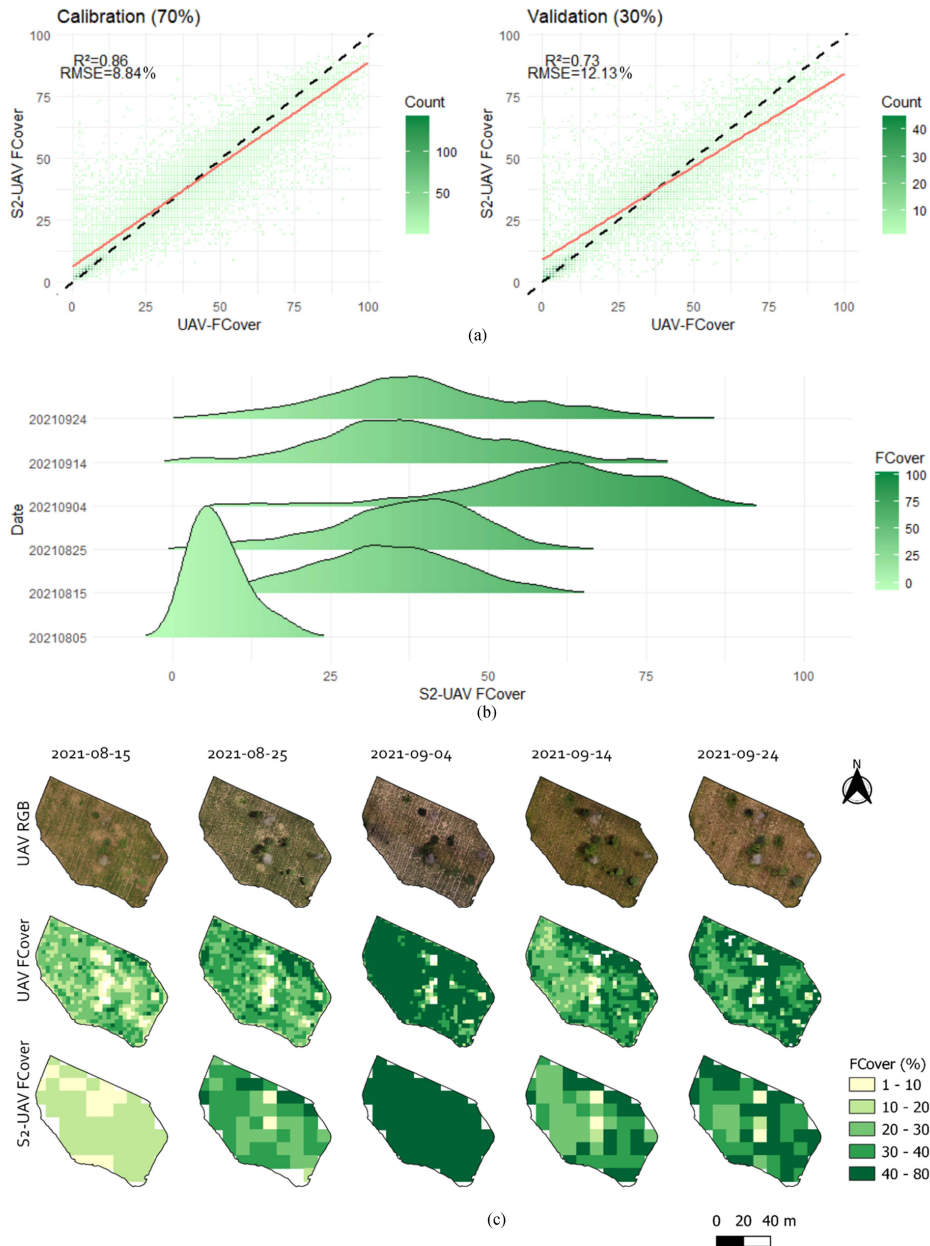


Fig. 7. S2-UAV FCover estimation based on the use of the full-time series (i.e., optimized option 2). (a) Comparison between the predicted S2-UAV FCover values and the UAV-FCover values for the calibration and validation phases. The red solid line is the regression line between the observed and predicted FCover, while the dashed line is the 1:1 line. The RF algorithm was applied based on the output of the baseline option. (b) Density distribution of the predicted S2-UAV FCover. (c) Details showing the UAV RGB image, the UAV-FCover and the predicted S2-UAV FCover for one field.

reached ( $\leq 64\%$ ). From the maturing stage to the harvest time, FCover values decreased (32%–48%) together with leaf senescence and fall. A higher spatial heterogeneity was also observed toward the end of the cropping season (04 September, 2021), where FCover values were between  $<16\%$  and  $>64\%$  suggesting that a large-scale difference in the timing of the cropping season can be observed for the same crop.

#### IV. DISCUSSION

##### A. Modeling on the Full-Time Series With RF Allows Improvement of the FCover Estimation

In this study, the following three modeling options have been compared.

- 1) A baseline option where the estimation of the S2-UAV FCover has been performed on a date-by-date basis and for three machine learning algorithms (RF, GBM, and SVM), and two optimized options.
- 2) Model stacking.
- 3) Modeling on the full-time series.

Our results showed that modeling using multitemporal information (i.e., using the full-time series) combined with an RF algorithm was the option allowing us to capture the millet FCover variability with highest accuracy [ $R^2 = 0.86$  in calibration and  $R^2 = 0.73$  in validation; Fig. 7(a)] when compared with the baseline option ( $R^2$  in validation ranging from 0.48 and 0.69 for RF depending on the dates; Fig. 6). We considered that these figures were within the range of validity for upscaling millet

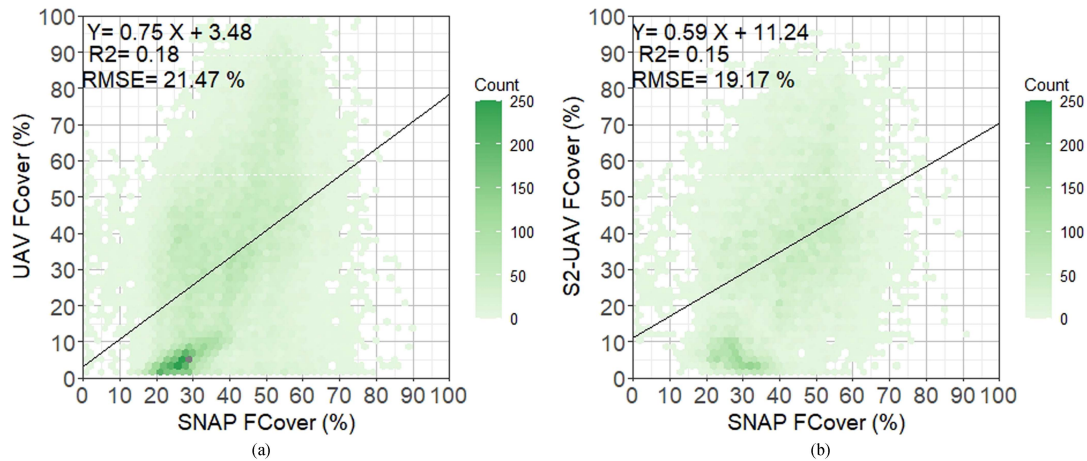


Fig. 8. Comparison of FCover estimated from the SNAP algorithms for all the dates over the cropping season in 2021. (a) SNAP-FCover vs. UAV-FCover; (b) S2-UAV FCover versus SNAP-FCover. The black line shows the linear regression line. The color gradient represents the number of points in each hexagon.

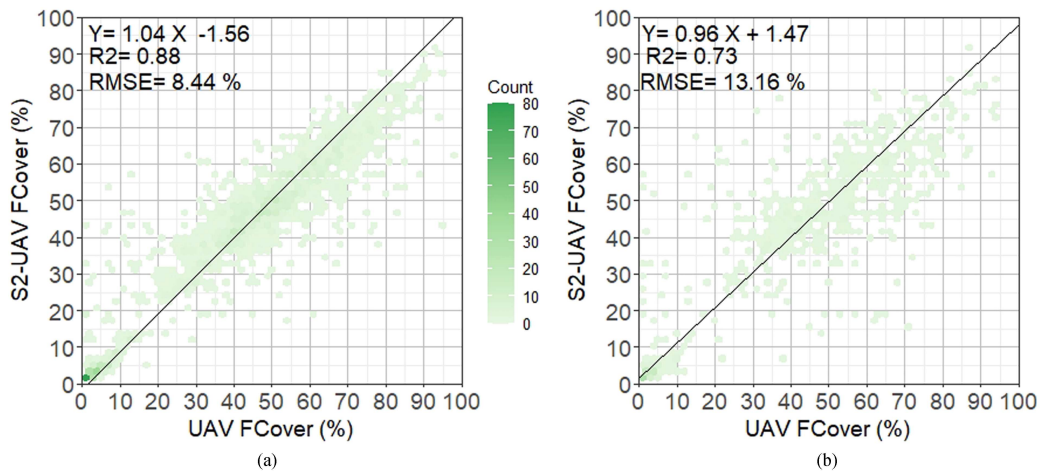


Fig. 9. Comparison between the predicted S2-UAV FCover values and the UAV-FCover values in the 2022 season for (a) calibration and (b) validation. The black line is the regression line between the observed and predicted FCover, while the dashed line is the 1:1 line.

FCover at the landscape scale with S2. Hence, our results suggest that using a multitemporal approach can better distinguish the phenological development of pearl millet than a date-by-date analysis (baseline option). Our results are in line with those found by Kamenova and Dimitrov [96] for mapping wheat FCover in Bulgaria using S2 data or [45] for mapping the FCover of the invasive shrub *Ulex europaeus* using UAV and S2 images. The modeling framework based on the full-time series is also more flexible in that it can be used to estimate FCover regardless of the developmental stage of pearl millet between emergence and harvesting. This is all the more convenient when reasoning at the landscape scale, where substantial differences can be found in the phenological development of the same crop. The second optimized option we tested was model stacking, which uses a diversity of models in an attempt to improve the overall ensemble performance. Surprisingly, model stacking performed similarly to the baseline option and hence did not show any additional value. This can be explained by the fact that each base learner performed relatively similarly ( $R^2$  and RMSE values in the same range of values) and by the fact that the three algorithms are closed in the way they operate, hence limiting the diversity in

terms of base learners [97]. In the future, more machine learning algorithms could be used as base learners to increase the diversity. In addition, a machine learning algorithm can be used as a meta-learner, instead of the linear regression used here, to better account for the potential nonlinear relationships between VIs and TFs with the UAV-FCover [98]. Our results also showed the higher performance of RF compared with SVM and GBM. This is in line with a previous study conducted over the same study area showing that RF outperformed other machine learning/deep learning approaches for millet yield estimations based on S2 and Sentinel-1 data [31]. However, an underestimation of high FCover values ( $> 50\%$ ) was observed, which can be explained by an overall tendency of RF to overestimate low values and to underestimate higher values, as already shown in previous studies [45], [55], [99], [100].

#### B. S2 Data Have Good Potential for Estimating FCover When Combined With UAV Data

To validate our methodological approach, the S2-UAV FCover obtained with the best modeling option (i.e., optimized option 2

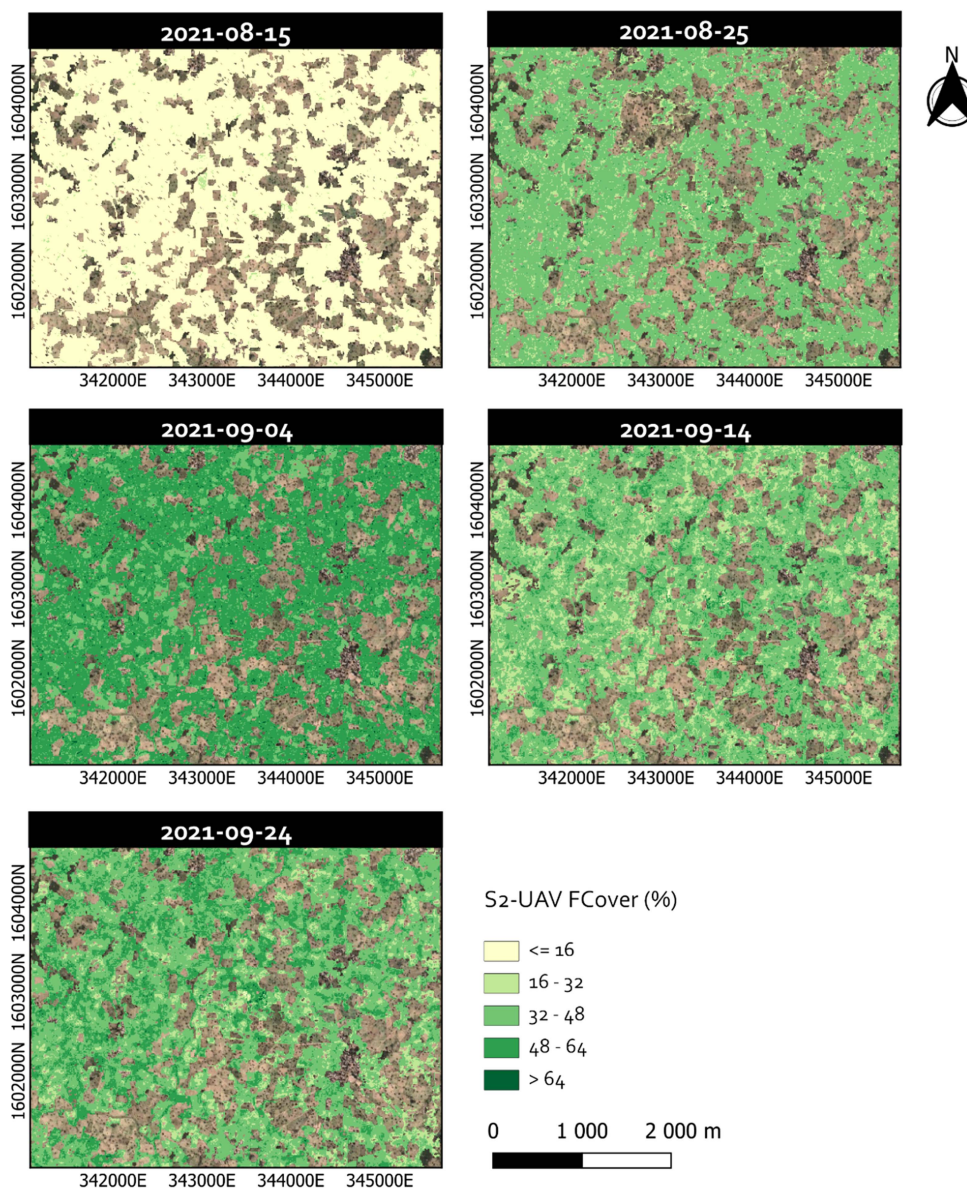


Fig. 10. S2-UAV FCover estimation over the study area for millet crop patches for five dates along the 2021 cropping season, displayed on a Google Satellite ©background.

using the full-time series) was compared against the FCover obtained from the SNAP algorithm using S2 data only (see Fig. 8). We have shown that the FCover estimated using the combination of S2 and UAV images clearly outperformed the one estimated through the SNAP biophysical model (PROSAIL) using S2 only ( $R^2 = 0.73$  in validation versus  $R^2 = 0.18$  when compared to UAV-based FCover). The SNAP FCover overestimated full range of values measured by UAV. Furthermore the overestimation affect mostly higher values ( $> 50\%$ ). Our results are in line with the study in [96] and [101] showing a higher accuracy for PROSAIL when estimating the FCover of row crops than empirical approaches. However, the authors agreed on the strong influence of the background on the capacity of PROSAIL to retrieve the FCover. Hence, in our case, this can be explained by the higher proportion of mixed pixels at the beginning of the cropping season when the FCover is low and the

background reflectance dominates the S2 spectral signal, which might strongly influence the variable retrieval by the SNAP algorithm. This is all the more true in our case because the level of analysis and calibration of the S2-UAV FCover model is over a grid of  $3\text{m} \times 3\text{m}$ , thus better taking into account the variability in the vegetation cover, compared with the SNAP algorithm, which works only at the native spatial resolution of S2 of 10 m. This is reinforced by the fact that UAV data, even aggregated at a spatial resolution of 3 m, also allow a better discrimination of the vegetation signature due to the use of the thresholding approach to remove background pixels, particularly at the beginning of the season when the vegetation cover is low. In addition, to test the robustness of the methodological framework over time, the best modeling option was applied to the 2022 dataset (see Fig. 9). Similar performances to those observed in 2021 (see Fig. 7) were found. Hence, even if the combined S2-UAV approach added

more steps in the methodological workflow (e.g., orthoimage processing and coregistration), this method added value for FCover estimation by improving accuracy, confirming in other studies (e.g. [44]) and was robust over time. In addition, for moderate spatial resolution satellite sensors, such as S2, for which it is almost impossible to find a pure pixel of vegetation in heterogeneous agricultural fields, we have shown here that using UAV data as pseudoground data can significantly improve the ground data collection for satellite images.

### C. S2-UAV FCover is Able to Depict the Spatiotemporal Variability of Millet FCover

Figs. 7 and 10 illustrate the ability of the S2-UAV FCover to depict the spatiotemporal variability of millet FCover. The analysis of the temporal profiles [see Figs. 7(b) and (c), and 10] showed a clear temporal pattern linked to the state of pearl millet development that matched the one observed with the UAV-based FCover (see Fig. 5). In the early season, when the vegetation cover is low and homogeneous, a normal distribution of FCover values was observed, with maximal values of approximately 10%. Then, it is followed by the core of the growing season (15 August, 2021 and 25 August, 2021), corresponding to an increase in the FCover values and a unimodal distribution. The peak of the growing season was reached on approximately 4 September, 2021, with FCover values of approximately 60%, followed by a decrease in FCover values toward the end of the cropping season but with a more heterogeneous distribution of FCover values. This heterogeneous pattern corresponds to the transition between the maximum of the growing phase and the maturation phase and was probably due to three simultaneous events, the fall of some millet leaves, the change in color of some leaves from green to yellow and the peak of weed development. For that analysis, the timing of acquisition was defined based on the schedule of S2 acquisition to obtain information every 10 days and hence to be able to capture the phenological development of the millet crop. Our results highlighted that the high repetition rate of the S2 constellation (i.e., ten days per satellite and five days in combination) is highly valuable to increase the chance of having a clear view during the key phenological stages of millet development. Conversely, analysis at the plot level [see Fig. 7(b) and (c)] and landscape level also showed high spatial heterogeneity within fields toward the end of the cropping season. This has already been shown for millet yield at harvest at the plot scale using UAV images [24] and landscape level using satellite images [33] over our study area and has been mainly explained by the presence of *F. albida* and, more broadly, the composition and structure of the tree layers at the landscape scale. In addition, it has been demonstrated for various crops, including millet, that the spatial variability of weed distribution can affect the crop yield at the field scale and that this effect is not constant throughout the field, as weeds develop as patches [102].

### D. Limits of the Study

While new satellite such as S2 have been designed for global mapping of canopy variables such as FCover, the moderate

spatial resolution (i.e., 10 m) they offered remains adapted to homogeneous agricultural landscape but still fails to reproduce the intrafield variability of the vegetation development observed in smallholder farming systems. Meanwhile, subcentimeter UAV images have been successfully used in heterogeneous agricultural landscape of smallholder farming systems for crop cover monitoring (e.g., [36]). In this study, UAV-based FCover images were used as an alternative to ground data to train satellite data and examine pearl millet FCover over our study area. In line with other studies (e.g. [35], [44], [45]), our results have shown the high potential of UAV-based FCover to scale-up FCover estimation in complex agricultural landscapes instead of using traditional field data that are more time-consuming to obtain. However, the accuracy of our approach is tightly dependent on the accuracy of the FCover derived from UAV images. First, the quality of the UAV-based FCover depends on the method used to compute the FCover. In that analysis, a threshold approach based on the ExGI was used to obtain a binary classification of millet-nonmillet pixels. However, even if the threshold has been carefully adjusted according to the date of acquisition, it might still be difficult to discriminate pure vegetation pixels due to high spatial variability within and between fields. That is why an alternative approach is to use machine learning algorithms to derive the FCover from UAV images because of their larger ability to account for large observed variations than the threshold approach [47]. The weeds within the fields, particularly those present at the end of the growing season, can also lead to a mixed signal and an overestimation of the millet FCover with UAV images. The ExGI is based on the visible-RGB channels that are limited in discriminating between small differences in the colors or shapes of plants. That is why some studies recommended using hyperspectral sensors to overcome those issues and be able to discriminate weeds [103]. Due to the different shapes between crops and weeds (weeds forming generally large continuous patches), Hall et al. [35] also recommended using object-based segmentation to discriminate maize from weeds. At the end of the cropping season during the senescence phase, we can also expect an underestimation of UAV-based FCover since nonphotosynthetic vegetation cannot be discriminated by the green vegetation index used [44], [104]. Second, the quality of the UAV-based FCover also depends on the spatial resolution of the data, as shown by [44]. In this study, a grid of 3m × 3m was chosen as an intermediate resolution between the UAV resolution and S2 resolution since it corresponds to the spatial resolution of PlanetScope images that are now available worldwide and which can be considered an alternative sensor to map FCover over a large scale. However, the aggregation of the data from a finer spatial resolution to a coarser spatial resolution leads to a spatial effect called the modifiable area unit problem (MAUP) by geographers. One way of coping with this issue is to rely on the aggregation of UAV-based FCover over different grid sizes to examine the (in)stability of the relationships for different spatial resolutions [105]. However, since the RF algorithm we used relies on bootstrapping (i.e., successively randomly removing several pixels from the analysis), we can assume that we limited the influence of MAUP by stabilizing the relationships. Last, the quality of the UAV processing workflow can also have an influence on the S2-UAV FCover estimation,

particularly linked to coregistration errors between UAV images and S2 images. These coregistration errors are mainly due to uncertainty in the GNSS positioning for both Sentinel and UAV images and in the coregistration step to align the UAV images across the different dates. Very few studies have considered these satellite UAV image coregistration issues when combining the two sources of information. Recently [45] proposed solving the issue of the spatial offset of both sensors by translating it into an optimization problem and hence allowing the identification of the optimal coalignment between S2 images and UAV images. Despite these potential limitations, the good agreement found using an independent dataset for 2022 suggests that the methodological framework proposed here is already robust for millet FCover estimation in the *F. albida* parkland. Future studies must be conducted to solve the aforementioned uncertainties related to UAV image processing, UAV-S2 coregistration and UAV-based FCover estimation.

## V. CONCLUSION

Mapping the FCover of staple food crops such as pearl millet at the field and landscape scales is a crucial step toward the improvement of crop productivity monitoring in smallholder heterogeneous agricultural landscapes such as in *F. albida* agroforestry parklands. The improvements made in terms of spatial, temporal and spectral resolution by the S2 constellation have opened new opportunities for crop monitoring at the landscape scale but failed so far to properly capture the within-field variability typical of smallholder agricultural landscapes. In this study, we assumed that the combination of low-cost UAV data and freely available satellite data, such as S2 can play an important role in refining crop monitoring in SSA. Our main objective was to propose an original methodological framework using UAV images and S2 data for estimating millet FCover using the example of a *F. albida* parkland in 2021 and 2022. UAV-based FCover over a 3 m × 3 m grid was used as ground observation for the upscaling of millet FCover at the landscape scale with S2 data. Six UAV and S2 images were acquired at different times of the millet growing season for our study area. After benchmarking different modeling approaches, the resulting best model was able to catch 73% of the UAV-FCover variability in validation in 2021 and 2022. In particular, we demonstrated that 1) the use of the full-time series allowed us to improve the FCover estimation over the cropping season compared with a date-based approach, and 2) satellite sensors such as S2 had good potential for estimating the spatiotemporal variability of crop FCover when associated with UAV images in smallholder agricultural landscapes. The results of this study show that combining UAV and satellite sensors can be a valuable tool to map crop FCover in heterogeneous agricultural landscapes, such as the Senegalese *F. albida* parklands. Thus, we recommend that users rely on UAV images as an alternative to ground data in such complex environments for FCover monitoring. However, the robustness of our approach needs to be tested under various smallholder agricultural landscapes with different crops and practices, especially at higher crop development, such as occurring in well fertilized and/or irrigated fields. To that

end, the applicability of our approach to new agricultural landscapes can be tested by analyzing the “Area of Applicability” as suggested by Meyer and Pebesma [106] when predictions are made for heterogeneous areas and with limited field data as it is most of the time the case in smallholder agricultural landscapes.

## ACKNOWLEDGMENT

The authors would like to thank the three Master’s students (P.O. Bouso, A. Dabo, and M. Faye), Ibou Diouf, Robert Diatta, and the farmers. The authors would also like to thank the CSE and Cirad, which provided technical and financial support for this study.

## REFERENCES

- [1] FAO, *The State of Food Security and Nutrition in the World: 2022: Repurposing Food and Agricultural Policies to Make Healthy Diets More Affordable*. Rome, Italy: FAO.
- [2] S. E. Vollset et al., “Fertility, mortality, migration, and population scenarios for 195 countries and territories from 2017 to 2100: A forecasting analysis for the Global Burden of Disease Study,” *Lancet*, vol. 396, no. 10258, pp. 1285–1306, 2020.
- [3] K. G. Cassman and P. Grassini, “A global perspective on sustainable intensification research,” *Nature Sustainability*, vol. 3, no. 4, pp. 262–268, 2020.
- [4] U. Nations, “The sustainable development goals report 2016,” United Nations, United Nations Pub., 300 East 42nd Street, New York, NY, 10017, USA, Tech. Rep. E.16.I.10, 2016.
- [5] J. V. Silva, P. Reidsma, F. Baudron, A. G. Laborte, K. E. Giller, and M. K. Van Ittersum, “How sustainable is sustainable intensification? Assessing yield gaps at field and farm level across the globe,” *Glob. Food Secur.*, vol. 30, 2021, Art. no. 100552.
- [6] B. Vanlauwe et al., “Sustainable intensification and the african smallholder farmer,” *Curr. Opin. Environ. Sustainability*, vol. 8, pp. 15–22, 2014.
- [7] P. R. Nair, B. M. Kumar, and V. D. Nair, 2021.
- [8] Fao and Faad, “The state of world fisheries and aquaculture,” in *Opportunities and Challenges*. Rome, Italy: Food and Agriculture Organization of the United Nations, 2012.
- [9] J. Bayala, A. Kalinganire, Z. Tchoundjeu, F. Sinclair, and D. Garrity, “Conservation agriculture with trees in the West African Sahel—A review,” *ICRAF Occasional Paper*, vol. 14, pp. 1–72, 2011.
- [10] D. P. Garrity et al., “Evergreen Agriculture: A robust approach to sustainable food security in Africa,” *Food Secur.*, vol. 2, pp. 197–214, 2010.
- [11] C. Mbow, M. Van Noordwijk, E. Luedeling, H. Neufeldt, P. A. Minang, and G. Kowero, “Agroforestry solutions to address food security and climate change challenges in Africa,” *Curr. Opin. Environ. Sustainability*, vol. 6, pp. 61–67, 2014.
- [12] T. S. Rosenstock et al., “A planetary health perspective on agroforestry in Sub-Saharan Africa,” *One Earth*, vol. 1, no. 3, pp. 330–344, 2019.
- [13] R. J. Zomer, A. Trabucco, R. Coe, F. Place, M. Van Noordwijk, and J. C. Xu, “Trees on farms: An update and reanalysis of agroforestry’s global extent and socio-ecological characteristics,” (Ed WACISARPD WP14064. PDF). Bogor, Indonesia, Tech. Rep., Working Paper 179, 2014.
- [14] A. Sambou, B. Sambou, and A. Ræbild, “Farmers’ contributions to the conservation of tree diversity in the Groundnut Basin, Senegal,” *J. Forestry Res.*, vol. 28, no. 5, pp. 1083–1096, 2017.
- [15] G. E. Wickens, “A study of *Acacia albida* Del.(Mimosoideae),” *Kew Bull.*, vol. 23, pp. 181–202, 1969.
- [16] L. Leroux et al., “Exploring the agricultural landscape diversity-food security nexus: An analysis in two contrasted parklands of Central Senegal,” *Agricultural Syst.*, vol. 196, 2022, Art. no. 103312.
- [17] R. M. Kho et al., “Separating the effects of trees on crops: The case of *Faidherbia albida* and millet in Niger,” *Agroforestry Syst.*, vol. 52, no. 3, pp. 219–238, 2001.
- [18] D. Louppe, B. N’Dour, and A. Samba, “Influence de *Faidherbia albida* sur l’arachide et le mil au Sénégal. Méthodologie de mesure et estimations des effets d’arbres émondés avec ou sans parage d’animaux,” 1996, pp. 123–139.

- [19] G. W. Sileshi, "The magnitude and spatial extent of influence of *Faidherbia albida* trees on soil properties and primary productivity in drylands," *J. Arid Environ.*, vol. 132, pp. 1–14, 2016.
- [20] J. Bayala, J. Sanou, H. R. Bazié, R. Coe, A. Kalinganire, and F. L. Sinclair, "Regenerated trees in farmers' fields increase soil carbon across the Sahel," *Agroforestry Syst.*, vol. 94, pp. 401–415, 2020.
- [21] J. Bayala, L. K. Heng, M. Van Noordwijk, and S. J. Ouedraogo, "Hydraulic redistribution study in two native tree species of agroforestry parklands of West African dry savanna," *Acta oecologica*, vol. 34, no. 3, pp. 370–378, 2008.
- [22] T. S. Sida, F. Baudron, H. Kim, and K. E. Giller, "Climate-smart agroforestry: *Faidherbia albida* trees buffer wheat against climatic extremes in the Central Rift Valley of Ethiopia," *Agricultural Forest Meteorol.*, vol. 248, pp. 339–347, 2018.
- [23] B. Ndao, L. Leroux, A. Hema, A. A. Diouf, A. Bégué, and B. Sambou, "Tree species diversity analysis using species distribution models: A *Faidherbia albida* parkland case study in Senegal," *Ecological Indicators*, vol. 144, 2022, Art. no. 109443.
- [24] O. Roupsard et al., "How far does the tree affect the crop in agroforestry? New spatial analysis methods in a *Faidherbia* parkland," *Agriculture, Ecosystems Environ.*, vol. 296, 2020, Art. no. 106928.
- [25] M. D. Johnson, W. W. Hsieh, A. J. Cannon, A. Davidson, and F. Bédard, "Crop yield forecasting on the Canadian Prairies by remotely sensed vegetation indices and machine learning methods," *Agricultural Forest Meteorol.*, vol. 218, pp. 74–84, 2016.
- [26] A. Cai et al., "Manure acts as a better fertilizer for increasing crop yields than synthetic fertilizer does by improving soil fertility," *Soil Tillage Res.*, vol. 189, pp. 168–175, 2019.
- [27] E. Kamir, F. Waldner, and Z. Hochman, "Estimating wheat yields in Australia using climate records, satellite image time series and machine learning methods," *ISPRS J. Photogrammetry Remote Sens.*, vol. 160, pp. 124–135, 2020.
- [28] S. Fritz et al., "Mapping global cropland and field size," *Glob. Change Biol.*, vol. 21, no. 5, pp. 1980–1992, 2015.
- [29] A. K. Whitcraft, E. F. Vermote, I. Becker-Reshef, and C. O. Justice, "Cloud cover throughout the agricultural growing season: Impacts on passive optical earth observations," *Remote Sens. Environ.*, vol. 156, pp. 438–447, 2015.
- [30] Z. Jin, G. Azzari, M. Burke, S. Aston, and D. B. Lobell, "Mapping smallholder yield heterogeneity at multiple scales in Eastern Africa," *Remote Sens.*, vol. 9, no. 9, 2017, Art. no. 931.
- [31] Y. J. E. Gbodjo, D. Ienco, and L. Leroux, "Benchmarking statistical modelling approaches with multi-source remote sensing data for millet yield monitoring: A case study of the groundnut basin in central Senegal," *Int. J. Remote Sens.*, vol. 42, no. 24, pp. 9285–9308, 2021.
- [32] Z. Jin et al., "Smallholder maize area and yield mapping at national scales with Google Earth Engine," *Remote Sens. Environ.*, vol. 228, pp. 115–128, 2019.
- [33] L. Leroux et al., "Using remote sensing to assess the effect of trees on millet yield in complex parklands of Central Senegal," *Agricultural Syst.*, vol. 184, 2020, Art. no. 102918.
- [34] D. B. Lobell, S. Di Tommaso, C. You, I. Y. Djima, M. Burke, and T. Kilic, "Sight for sorghums: Comparisons of satellite- and ground-based sorghum yield estimates in Mali," *Remote Sens.*, vol. 12, no. 1, 2019, Art. no. 100.
- [35] O. Hall, S. Dahlin, H. Marstorp, M. F. Archila Bustos, I. Öborn, and M. Jirstrom, "Classification of maize in complex smallholder farming systems using UAV imagery," *Drones*, vol. 2, no. 3, 2018, Art. no. 22.
- [36] A. G. Schut, P. C. S. Traore, X. Blaes, and R. A. de By, "Assessing yield and fertilizer response in heterogeneous smallholder fields with UAVs and satellites," *Field Crops Res.*, vol. 221, pp. 98–107, 2018.
- [37] L. Pádua et al., "UAS, sensors, and data processing in agroforestry: A review towards practical applications," *Int. J. Remote Sens.*, vol. 38, no. 8–10, pp. 2349–2391, 2017.
- [38] J. Sarron, E. Malézieux, C. A. B. Sané, and E. Faye, "Mango yield mapping at the orchard scale based on tree structure and land cover assessed by UAV," *Remote Sens.*, vol. 10, no. 12, 2018, Art. no. 1900.
- [39] D. Fawcett et al., "Multi-scale evaluation of drone-based multispectral surface reflectance and vegetation indices in operational conditions," *Remote Sens.*, vol. 12, no. 3, 2020, Art. no. 514.
- [40] J. L. Martinez et al., "Comparison of satellite and drone-based images at two spatial scales to evaluate vegetation regeneration after post-fire treatments in a mediterranean forest," *Appl. Sci.*, vol. 11, no. 12, 2021, Art. no. 5423.
- [41] R. J. Nuijten, N. C. Coops, C. Watson, and D. Theberge, "Monitoring the structure of regenerating vegetation using drone-based digital aerial photogrammetry," *Remote Sens.*, vol. 13, no. 10, 2021, Art. no. 1942.
- [42] J. E. Böhrer, M. E. Schaepman, and M. Kneubühler, "Crop classification in a heterogeneous arable landscape using uncalibrated UAV data," *Remote Sens.*, vol. 10, no. 8, 2018, Art. no. 1282.
- [43] I. Alvarez et al., "A winter upwelling event in the Northern Galician Rias: Frequency and oceanographic implications," *Estuarine, Coastal Shelf Sci.*, vol. 82, no. 4, pp. 573–582, 2009.
- [44] H. Riihimäki, M. Luoto, and J. Heiskanen, "Estimating fractional cover of tundra vegetation at multiple scales using unmanned aerial systems and optical satellite data," *Remote Sens. Environ.*, vol. 224, pp. 119–132, 2019.
- [45] T. Gränzig, F. E. Fassnacht, B. Kleinschmit, and M. Förster, "Mapping the fractional coverage of the invasive shrub *Ulex europaeus* with multi-temporal Sentinel-2 imagery utilizing UAV orthoimages and a new spatial optimization approach," *Int. J. Appl. Earth Observ. Geoinformation*, vol. 96, 2021, Art. no. 102281.
- [46] S. Wang et al., "Cross-scale sensing of field-level crop residue cover: Integrating field photos, airborne hyperspectral imaging, and satellite data," *Remote Sens. Environ.*, vol. 285, 2023, Art. no. 113366.
- [47] L. Li et al., "Review of ground and aerial methods for vegetation cover fraction (fCover) and related quantities estimation: Definitions, advances, challenges, and future perspectives," *ISPRS J. Photogrammetry Remote Sens.*, vol. 199, pp. 133–156, 2023.
- [48] F. Baret et al., "GEOV1: LAI and FAPAR essential climate variables and FCOVER global time series capitalizing over existing products. Part1: Principles of development and production," *Remote Sens. Environ.*, vol. 137, pp. 299–309, 2013.
- [49] G. Song et al., "Monitoring leaf phenology in moist tropical forests by applying a superpixel-based deep learning method to time-series images of tree canopies," *ISPRS J. Photogrammetry Remote Sens.*, vol. 183, pp. 19–33, 2022.
- [50] T. Kattenborn, J. Lopatin, M. Förster, A. C. Braun, and F. E. Fassnacht, "UAV data as alternative to field sampling to map woody invasive species based on combined Sentinel-1 and Sentinel-2 data," *Remote Sens. Environ.*, vol. 227, pp. 61–73, 2019.
- [51] F. Baret and S. Buis, "Estimating canopy characteristics from remote sensing observations: Review of methods and associated problems," *Adv. Land Remote Sens.: Syst., Model., Inversion Appl.*, vol. 2008, pp. 173–201, 2008.
- [52] A. Beniaich, M. L. Naves Silva, F. A. P. Avalos, M. D. Menezes, and B. M. Candido, "Determination of vegetation cover index under different soil management systems of cover plants by using an unmanned aerial vehicle with an onboard digital photographic camera," *Semina-Ciencias Agrarias*, vol. 40, no. 1, pp. 49–66, 2019.
- [53] Y. Feng, K. Fennel, G. A. Jackson, S. F. DiMarco, and R. D. Hetland, "A model study of the response of hypoxia to upwelling-favorable wind on the northern Gulf of Mexico shelf," *J. Mar. Syst.*, vol. 131, pp. 63–73, 2014.
- [54] G.-H. Kwak and N.-W. Park, "Impact of texture information on crop classification with machine learning and UAV images," *Appl. Sci.*, vol. 9, no. 4, 2019, Art. no. 643.
- [55] L. Leroux, M. Castets, C. Baron, M.-J. Escorihuela, A. Bégué, and D. L. Seen, "Maize yield estimation in West Africa from crop process-induced combinations of multi-domain remote sensing indices," *Eur. J. Agronomy*, vol. 108, pp. 11–26, 2019.
- [56] O. Sagi and L. Rokach, "Ensemble learning: A survey," *Wiley Interdiscipl. Rev.: Data Mining Knowl. Discov.*, vol. 8, no. 4, 2018, Art. no. e1249.
- [57] S. Tao, X. Zhang, R. Feng, W. Qi, Y. Wang, and B. Shrestha, "Retrieving soil moisture from grape growing areas using multi-feature and stacking-based ensemble learning modeling," *Comput. Electron. Agriculture*, vol. 204, 2023, Art. no. 107537.
- [58] K. Li, L. Wang, and D. Yin, "Deriving corn and soybeans fractions with Land Remote-Sensing Satellite (System, Landsat) imagery by accounting for endmember variability on Google Earth Engine," *Int. J. Remote Sens.*, vol. 42, no. 12, pp. 4493–4513, 2021.
- [59] D. C. Corrales, A. Figueroa, A. Ledezma, and J. C. Corrales, "An empirical multi-classifier for coffee rust detection in colombian crops," in *Proc. 15th Int. Conf. Comput. Sci. Appl.*, 2015, pp. 60–74.
- [60] W. Masiza, J. G. Chirima, H. Hamandawana, and R. Pillay, "Enhanced mapping of a smallholder crop farming landscape through image fusion and model stacking," *Int. J. Remote Sens.*, vol. 41, no. 22, pp. 8739–8756, 2020.



- [61] J. Useya and S. Chen, "Comparative performance evaluation of pixel-level and decision-level data fusion of Landsat 8 OLI, Landsat 7 ETM and Sentinel-2 MSI for crop ensemble classification," *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.*, vol. 11, no. 11, pp. 4441–4451, Nov. 2018.
- [62] G. Aguilar, A. P. López-Monroy, F. A. González, and T. Solorio, "Modeling noisiness to recognize named entities using multitask neural networks on social media," 2019, *arXiv:1906.04129*.
- [63] A. Lericollais, "Paysans Sereer: Dynamiques Agraires et Mobilités au Sénégal: IRD, À travers champs, 1999, 668 p., 190 F," *Nature Sci. Sociétés*, vol. 8, no. 1, p. 86, 2000.
- [64] P. Tittonell, A. Muriuki, C. J. Klapwijk, K. D. Shepherd, R. Coe, and B. Vanlauwe, "Soil heterogeneity and soil fertility gradients in smallholder farms of the East African highlands," *Soil Sci. Soc. Amer. J.*, vol. 77, no. 2, pp. 525–538, 2013.
- [65] O. Roupsard et al., "Reverse phenology and dryseason water uptake by *Faidherbia albida* (Del.) A. Chev. in an agroforestry parkland of Sudanese west Africa," *Funct. Ecol.*, vol. 13, no. 4, pp. 460–472, 1999.
- [66] G. F. Félix, J. Scholberg, C. Clermont-Dauphin, L. Cournac, and P. Tittonell, "Enhancing agroecosystem productivity with woody perennials in semi-arid West Africa. A meta-analysis," *Agronomy Sustain. Develop.*, vol. 38, no. 6, pp. 1–21, 2018.
- [67] M. Drusch et al., "Sentinel-2: ESA's optical high-resolution mission for GMES operational services," *Remote Sens. Environ.*, vol. 120, pp. 25–36, 2012.
- [68] A. A. Gitelson, Y. Gritz, and M. N. Merzlyak, "Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves," *J. Plant Physiol.*, vol. 160, no. 3, pp. 271–282, 2003.
- [69] J. W. Rouse Jr, R. H. Haas, J. A. Schell, and D. W. Deering, "Paper a 20," in *Proc. 3rd Earth Resour. Technol. Satell.-1 Symp.: Proc. Symp. Held Goddard Space Flight Center*, 1973, Art. no. 309.
- [70] A. A. Gitelson, Y. J. Kaufman, and M. N. Merzlyak, "Use of a green channel in remote sensing of global vegetation from EOS-MODIS," *Remote Sens. Environ.*, vol. 58, no. 3, pp. 289–298, 1996.
- [71] D. M. Woebbecke, G. E. Meyer, K. Von Bargen, and D. A. Mortensen, "Color indices for weed identification under various soil, residue, and lighting conditions," *Trans. ASAE*, vol. 38, no. 1, pp. 259–269, 1995.
- [72] R. M. Haralick, K. Shanmugam, and I. H. Dinstein, "Textural features for image classification," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-3, no. 6, pp. 610–621, Nov. 1973.
- [73] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, pp. 5–32, 2001.
- [74] N. Gorelick, M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau, and R. Moore, "Google Earth Engine: Planetary-scale geospatial analysis for everyone," *Remote Sens. Environ.*, vol. 202, pp. 18–27, 2017.
- [75] X. Zhuo, T. Koch, F. Kurz, F. Fraundorfer, and P. Reinartz, "Automatic UAV image geo-registration by matching UAV images to georeferenced image data," *Remote Sens.*, vol. 9, no. 4, 2017, Art. no. 376.
- [76] J. Torres-Sánchez, J. M. Peña, A. I. De Castro, and F. López-Granados, "Multi-temporal mapping of the vegetation fraction in early-season wheat fields using images from UAV," *Comput. Electron. Agriculture*, vol. 103, pp. 104–113, 2014.
- [77] J. Chen, S. Yi, Y. Qin, and X. Wang, "Improving estimates of fractional vegetation cover based on UAV in alpine grassland on the Qinghai-Tibetan Plateau," *Int. J. Remote Sens.*, vol. 37, no. 8, pp. 1922–1936, 2016.
- [78] G. Forkuor, C. Conrad, M. Thiel, B. J. Zoungrana, and J. E. Tondoh, "Multiscale remote sensing to map the spatial distribution and extent of cropland in the Sudanian Savanna of West Africa," *Remote Sens.*, vol. 9, no. 8, 2017, Art. no. 839.
- [79] S. R. Debats, D. Luo, L. D. Estes, T. J. Fuchs, and K. K. Caylor, "A generalized computer vision approach to mapping crop fields in heterogeneous agricultural landscapes," *Remote Sens. Environ.*, vol. 179, pp. 210–221, 2016.
- [80] J. Xiong et al., "Nominal 30-m cropland extent map of continental Africa by integrating pixel-based and object-based algorithms using Sentinel-2 and Landsat-8 data on Google Earth Engine," *Remote Sens.*, vol. 9, no. 10, 2017, Art. no. 1065.
- [81] K. M. Ting and I. H. Witten, "Nominal 30-m cropland extent map of continental Africa by integrating pixel-based and object-based algorithms using Sentinel-2 and Landsat-8 data on Google Earth Engine," *Remote Sens.*, vol. 179, no. 10, 2017, Art. no. 1065.
- [82] Q. Xie et al., "Retrieval of crop biophysical parameters from Sentinel-2 remote sensing imagery," *Int. J. Appl. Earth Observ. Geoinformation*, vol. 80, pp. 187–195, 2019.
- [83] M. Kganyago, P. Mhangara, T. Alexandridis, G. Laneve, G. Ovakoglou, and N. Mashiyi, "Validation of sentinel-2 leaf area index (LAI) product derived from SNAP toolbox and its comparison with global LAI products in an African semi-arid agricultural landscape," *Remote Sens. Lett.*, vol. 11, no. 10, pp. 883–892, 2020.
- [84] L. A. Brown et al., "Validation of baseline and modified Sentinel-2 Level 2 Prototype Processor leaf area index retrievals over the United States," *ISPRS J. Photogrammetry Remote Sens.*, vol. 175, pp. 71–87, 2021.
- [85] M. Weiss, F. Baret, and S. Jay, "S2toolbox level 2 products: Lai, Fapar, Fcover," *Inst. Nat. de la Recherche Agronomique Avignon*, 2016.
- [86] R. J. Hijmans, *Raster: Geographic Data Analysis and Modeling*, 2022, R. package version 3.5-15. [Online]. Available: <https://CRAN.R-project.org/package=raster>
- [87] E. Pebesma and R. S. Bivand, "S classes and methods for spatial data: The SP package," *R News*, vol. 5, no. 2, pp. 9–13, 2005.
- [88] R. Bivand, T. Keitt, and B. Rowlingson, "rgdal: Bindings for the 'Geospatial' Data Abstraction Library," R. Package Version 1.5-32, 2022. [Online]. Available: <https://CRAN.R-project.org/package=rgdal>
- [89] M. N. Wright and A. Ziegler, "ranger: A fast implementation of random forests for high dimensional data in C and R," *J. Stat. Softw.*, vol. 77, no. 1, pp. 1–17, 2017.
- [90] A. Liaw and M. Wiener, "Classification and regression by randomforest," *R News*, vol. 2, no. 3, pp. 18–22, 2002.
- [91] T. Therneau and B. Atkinson, "rpart: Recursive partitioning and regression trees," *R Package Version*, vol. 4, pp. 1–15, 2019.
- [92] J. Fox and S. Weisberg, *An R Companion to Applied Regression*, 3rd ed. Thousand Oaks, CA, USA: SAGE, 2019.
- [93] H. Wickham et al., "Welcome to the tidyverse," *J. Open Source Softw.*, vol. 4, no. 43, 2019, Art. no. 1686.
- [94] R Core Team, *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R. Foundation for Statistical Computing, 2021.
- [95] H. Wickham, *ggplot2: Elegant Graphics for Data Analysis*. New York, NY, USA: Springer-Verlag, 2016.
- [96] I. Kamenova and P. Dimitrov, "Evaluation of Sentinel-2 vegetation indices for prediction of LAI, fAPAR and fCover of winter wheat in Bulgaria," *Eur. J. Remote Sens.*, vol. 54, no. sup1, pp. 89–108, 2021.
- [97] M. Barton and B. Lennox, "Model stacking to improve prediction and variable importance robustness for soft sensor development," *Digit. Chem. Eng.*, vol. 3, 2022, Art. no. 100034.
- [98] Y. Zhang, J. Ma, S. Liang, X. Li, and J. Liu, "A stacking ensemble algorithm for improving the biases of forest aboveground biomass estimations from multiple remotely sensed datasets," *GIScience Remote Sens.*, vol. 59, no. 1, pp. 234–249, 2022.
- [99] J. Lopatin, K. Dolos, H. J. Hernández, M. Galleguillos, and F. E. Fassnacht, "Comparing generalized linear models and random forest to model vascular plant species richness using LiDAR data in a natural forest in central Chile," *Remote Sens. Environ.*, vol. 173, pp. 200–210, 2016.
- [100] G. Zhang and Y. Lu, "Bias-corrected random forests in regression," *J. Appl. Statist.*, vol. 39, no. 1, pp. 151–160, 2012.
- [101] X. Ma, L. Lu, J. Ding, F. Zhang, and B. He, "Estimating fractional vegetation cover of row crops from high spatial resolution image," *Remote Sens.*, vol. 13, no. 19, 2021, Art. no. 3874.
- [102] N. Colbach et al., "The pitfalls of relating weeds, herbicide use, and crop yield: Don't fall into the trap! a critical review," *Front. Agronomy*, vol. 2, 2020, Art. no. 615470.
- [103] Y. Li et al., "Identification of weeds based on hyperspectral imaging and machine learning," *Front. Plant Sci.*, vol. 11, 2021, Art. no. 611622.
- [104] C. Li et al., "Maize yield estimation in intercropped smallholder fields using satellite data in Southern Malawi," *Remote Sens.*, vol. 14, no. 10, 2022, Art. no. 2458.
- [105] M. Buzzelli, "Modifiable areal unit problem," *Int. Encyclopedia Hum. Geography*, 2020, pp. 169–173, doi: [10.1016/B978-0-08-102295-5.10406-8](https://doi.org/10.1016/B978-0-08-102295-5.10406-8).
- [106] H. Meyer and E. Pebesma, "Predicting into unknown space? estimating the area of applicability of spatial prediction models," *Methods Ecol. Evol.*, vol. 12, no. 9, pp. 1620–1633, 2021.



**Ibrahim Diack** is Senegal citizen. He received the master's degree in marine ecology from university Cheikh Anta DIOP, Senegal. He is currently working toward the Ph.D. degree with the Mathematic and Informatics Department, Cheikh Anta DIOP University, Dakar, Senegal, in collaboration with French agricultural research and international cooperation organization (CIRAD) and Centre de Suivi Ecologique (CSE), Senegal.

He worked with different kind of oceanographic and atmospheric data from remote sensing (Oceana color, Ascat, QuikSCAT) and oceanographic in-situ data (CTD, acoustic, nutrients, etc). He works mainly on the development of crop monitoring solution using UAV and satellite combination together with empirical approaches.



**Heuclin Benjamin** received the Ph.D. degree in Bayesian statistics applied to genetics from the University of Montpellier, Montpellier, France, in 2021.

He is currently a biostatistician with CIRAD, Montpellier, France. He joined the AIDA unit, where his research focus on survey and experimental trial design and analyses, sensitivity analysis of crop models, and machine learning approaches for remote sensing data.



**Serigne Mansour Diene** received the master's degree in topography from University Iba Der Thiam, Thies, Senegal, in 2021. He is currently working toward the Ph.D. degree in computer science, specializing in remote sensing applied to agriculture, from Computer Science Department, Cheikh Anta Diop University of Dakar (UCAD), Dakar, Senegal, in collaboration with the French agricultural research and international cooperation organization (CIRAD).

As part of his research work, he is posted at the LMI IESOL (IRD - UCAD - ISRA), Dakar, and work on estimating crop yields in agroforestry systems and the spatialized assessment of the effect of trees on crops using aerial imagery obtained by UAV and statistical approaches. In addition, during his degree, he worked on sorghum phenotyping for varietal creation programs using thermal images from UAV to assess the crop's level of water stress.



**Roupsard Olivier** received the Ph.D. degree in forest biology from the University of Nancy, Nancy, France, in 1997.

He is a French Senior Scientist with CIRAD, Montpellier, France. He research focuses on tropical agroforestry, bioclimatology and modeling, hydrology, and proxy-sensing. He has been the P.I. of several FLUXNET eddy-covariance sites in the tropics. He is currently at Senegal to work on sahelian agro-silvo-pastoral systems.



**Leroux Louise** is currently with CIRAD, UPR AIDA, Nairobi, Kenya; AIDA, University of Montpellier, CIRAD, Montpellier, France; IITA, Nairobi, Kenya, working on agroecology and sustainable intensification of annual crops production in terms of quantity and also quality where its relevant, in a particularly constrained tropical environment. She is a Geographer with a strong background in remote sensing applied to agricultural monitoring. Over the last years, she worked mainly in West Africa, with a focus on agroforestry systems. She was previously seconded

to Centre de Suivi Ecologique, Dakar, Senegal, and she is currently with IITA, Nairobi, Kenya, where she conducts her research mainly in Ethiopia and Rwanda. Her research focuses on the use of remote sensing technologies combined with statistical or biophysical modeling to improve the cropping systems descriptions (where are the crop areas and what kind of crops, what are the agricultural practices, what are the interactions with the surrounding landscape, etc.) and improve the assessment of agronomical and environmental performances of smallholder cropping systems.



**Letourmy Philippe** received the Ph.D. degree in biometry applied to dairy technology from AgroParis-Tech, France, 1988.

He is currently a Senior Biostatistician with CIRAD, Montpellier, France. After an experience in Ivory Coast as Agrostatistician for a special Project on upland rice, he has been working with CIRAD for 35 years. Since 2014, he has been with AIDA unit. His research interests include sampling and design of experiments and statistical analyses, and different approaches in data science.



**Diouf Abdoul Aziz** received the Ph.D. degree in environmental sciences and management from the University of Liège, Liège, Belgium, in 2016.

He is currently with the Centre de Suivi Ecologique (CSE), Dakar, Senegal, as a Research Scientist and Coordinator of the research and development programme. He is also Researcher of the Pastoralism and Drylands Pole (PPZS) in West Africa. His research interests include acquisition, processing, and analysis of drone and temporal series of satellite images, statistical modeling of forage biomass production and

crop yields using earth observation data, and landscape ecology and mapping of natural ecosystem variables.



**Audebert Alain** received the Ph.D. degree in plant Physiology from the University of Montpellier, Montpellier, France, in 1993.

He is a French Senior Scientist with CIRAD Montpellier, France. For several years now, his field of activity has been focused on high-throughput phenotyping methods, notably using UAV onboard imagery. His research focuses on ecophysiology for a better understanding of adapting tropical crops to abiotic constraints.



**Sarr Idrissa** received the master's degree in computer science from Cheikh Anta Diop University, Dakar, Senegal, in 2006, the Ph.D. degree in large-scale distributed databases from Université Pierre et Marie Curie, Paris, France, in 2010, and the

He is currently a Teacher and Researcher with Cheikh Anta Diop University. He was a Postdoc in social data analysis with Université du Québec en Outaouais, Gatineau, QC, Canada, between 2011 and 2012. His research interests include datascience and artificial intelligence.



**Diallo Moussa** received the master's degree in electronics and telecommunications from the University of Limoges, Limoges, France, in 2007, and the Ph.D. degree in telecommunications from the University of Rennes, Rennes, France, in 2010.

He is currently a Full Professor with the Polytechnic Institute, University Cheikh Anta Diop de Dakar, Dakar, Senegal. His research interests include wireless communication, IoT, and wireless sensor networks.