



HAL
open science

Crop Monitoring Using Vegetation and Thermal Indices for Yield Estimates: Case Study of a Rainfed Cereal in Semi-Arid West Africa

Louise Leroux, Christian Baron, Bernardin Zoungrana, Seydou Traore, Danny Lo Seen, Lo Seen, Agnes Begue

► **To cite this version:**

Louise Leroux, Christian Baron, Bernardin Zoungrana, Seydou Traore, Danny Lo Seen, et al.. Crop Monitoring Using Vegetation and Thermal Indices for Yield Estimates: Case Study of a Rainfed Cereal in Semi-Arid West Africa. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2016, 9 (1), pp.347-362. cirad-01951499

HAL Id: cirad-01951499

<https://hal.science/cirad-01951499>

Submitted on 11 Dec 2018

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.

1 Crop monitoring using vegetation and thermal indices for yield estimates:

2 Case study of a rainfed cereal in semi-arid West Africa

3
4 Leroux L^{a,b*}, Baron C^a, Zoungrana B^c, Traoré S.B^c, Lo Seen D^a, Bégué A^a.

5 ^a CIRAD UMR TETIS, Maison de la Télédétection, 500 rue Jean François Breton, Montpellier, 34093,
6 France

7 ^b AgroParisTech, 648 rue Jean François Breton, Montpellier, 34093, France

8 ^c AGRHYMET Regional Center, BP 11011, Niamey, Niger

9 * Corresponding author at: CIRAD UMR TETIS, Maison de la Télédétection, 500 rue Jean François
10 Breton, Montpellier, 34093, France. *Email address:* louise.leroux@teledetection.fr (L. Leroux)

11 **Abstract**

12 For the semi-arid Sahelian region, climate variability is one of the most important risks of food insecurity.
13 Field experimentations as well as crop modelling are helpful tools for the monitoring and the
14 understanding of yields at local scale. However, extrapolation of these methods at a regional scale remains
15 a demanding task. Remote sensing observations appear as a good alternative or addition to existing crop
16 monitoring systems. In this study, a new approach based on the combination of vegetation and thermal
17 indices for rainfed cereal yield assessment in the Sahelian region was investigated. Empirical statistical
18 models were developed between MODIS NDVI and LST variables and the crop model SARRA-H
19 simulated aboveground biomass and harvest index in order to assess each component of the yield
20 equation. The resulting model was successfully applied at the Niamey Square Degree (NSD) site scale
21 with yield estimations close to the official agricultural statistics of Niger for a period of 11 years (2000-
22 2011) ($r=0.82$, $pvalue<0.05$). The combined NDVI and LST indices based model was found to clearly
23 outperform the model based on NDVI alone ($r=0.59$, $pvalue<0.10$). In areas where access to ground
24 measurements is difficult, a simple, robust and timely satellite-based model combining vegetation and

25 thermal indices from MODIS and calibrated using crop model outputs, can be pertinent. In particular, such
26 a model can provide an assessment of the year-to-year yield variability shortly after harvest for regions
27 with agronomic and climate characteristics close to those of the NSD study area.

28
29 Keywords: Remote Sensing, Crop yield, NDVI, Land Surface Temperature, Crop model, MODIS,
30 Rainfed cereal, Niger, Harvest Index

31 **1. Introduction**

32
33 In the Sahelian region of West Africa where traditional rainfed agriculture prevails [1], over 20
34 million people suffered from food insecurity in 2014 [2]. Sahelian rainfed farming systems are known to
35 be at high climatic risk due to a high spatio-temporal variability of rainfall and frequent drought events
36 [3]. Rainfall variability results in large fluctuations in year-to-year crop productivity which leads to
37 episodes of food insecurity. Moreover, the political and socio-economic instability of certain countries in
38 the region also contribute to the variability of agricultural production [4]. These considerations highlight
39 the need for an operational, timely and accurate yield estimation system to assist decision-making [5]–[7].
40 Yield estimation systems based on crop modelling allow accurate quantitative assessments (e.g.
41 AGRHYMET in West Africa; the AGRI4CAST action in Europe), but are confronted with input data
42 availability and spatial consistency constraints [8], [9].

43 For more than two decades, Earth Observation systems have been known to play a significant role
44 in vegetation monitoring by providing synoptic, repetitive, timely, objective and cost-effective
45 information on Earth’s surfaces (e.g.[10]–[12]). They have been acknowledged for their valuable
46 contribution to spatial and temporal monitoring of global vegetation and thus have been used extensively
47 in many parts of the world for crop condition monitoring and yield forecasting [9], [13]–[18]. Combined
48 or not with rainfall data, satellite data is currently being used in early warning systems to assess crop
49 development conditions during the growing season (e.g. FEWS-NET (Famine Early Warning System

50 Network); GIEWS (Global Information and Early Warning System); FOODSEC (Food Security);
51 GEOGLAM (Group on Earth Observation – Global Agricultural Monitoring)), mainly through the
52 production of regularly updated crop growth anomalies' maps based on the Normalized Difference
53 Vegetation Index (NDVI). These systems benefit from timely and synoptic satellite data which
54 compensate the lack of reliable and homogeneous ground data, but which remain mainly qualitative and
55 do not include crop yield monitoring.

56 The empirical relationship between remote sensing spectral vegetation indices and in situ
57 observations to predict crop yields before harvest has been tested for a long time in many studies (for a
58 review see [19]). The simplest approach involves the regression between observed yields and vegetation
59 indices, either on a specific date or through a time integral of vegetation indices between two dates [20]–
60 [22]. Among the vegetation indices, the NDVI has been widely employed due to its close relationship to
61 several vegetation parameters like the Leaf Area Index (LAI), the fraction of Absorbed Photosynthetically
62 Active Radiation (fAPAR) or the green biomass [23]–[25]. Furthermore, several studies have found a
63 good correlation between NDVI and crop yields in many study sites around the world [13], [15], [16],
64 [26]–[28].

65 Nonetheless, there are intrinsic limitations that prevent an operational use of vegetation indices to
66 estimate crop yield. Apart from technical limitations such as low spatial and temporal resolution leading to
67 mixed pixels and incomplete crop growth descriptions respectively, the main limitation is the indirect link
68 between yield and spectral data. Since the 1980's NDVI is known to be a proxy of the aboveground
69 biomass [11], [24], but the ratio between yield and aboveground biomass (referred hereafter as the harvest
70 index, HI) is also known to be highly variable in space and time. [29], [30] showed that biomass
71 production is linearly related to fAPAR for crop with no water stress, while [31] showed that a linear
72 relation between NDVI and fAPAR can be assumed since their functional response to leaf orientation,
73 solar zenith angle and atmospheric optical depth is similar.

74 On the other hand, the harvest index, also known as the reproductive efficiency, is crop-dependent
75 and sensitive to variables that impact the partitioning of the assimilates into grain, such as the genotype,

76 temperature and water/nutrients availability ([32]–[34]). In the Sahelian region, HI is strongly dependent
77 on water conditions during the growing period, in particular during the reproductive stage [35]. Crop
78 water conditions can also be derived from remote sensing data, with indices based on the difference
79 between air and surface temperatures, which are useful indicators of water stress for yield estimation.
80 Indeed, since the 1970's various remote sensing-based studies have shown that final yields can be related
81 to thermal indices [36], [37]. Based on land surface temperature (LST), the Crop Water Stress Index
82 (CWSI) proposed by [38] was found useful for yield estimation and crop assessment (e.g.[39]–[42]).

83 In the framework of current early warning systems for food security, crop yield monitoring would
84 certainly benefit from the consistency in space and time of remote-sensing based crop yield estimations.
85 For this reason, in this study, we investigate the possibility of combining vegetation and thermal indices
86 for crop yield estimation in the Sahelian region, where, to our knowledge, this has not been attempted. Our
87 objective is to build a simple, robust and timely satellite-based model for rainfed cereal yield estimates on
88 the basis that: (i) aboveground biomass can be estimated using NDVI, and that (ii) LST data can provide
89 useful information on the harvest index. Such a model would also provide effective assessments of year-
90 to-year yield variability. The study is conducted in the South-West of Niger (the Niamey Square Degree
91 site) where rainfed pearl millet dominates the agricultural landscape, and soils as well as agricultural
92 practices are relatively homogeneous. We use the SARRA-H (**System for Regional Analysis of Agro-
93 Climatic Risks**) crop model [43], which has already been validated for pearl millet in the Sudano-Sahelian
94 zone [4], to simulate biomass and the corresponding yield for a period of 11 years (2000-2011), using as
95 input data the rainfall measurements from 28 rain gauges and a meteorological station. We derive the
96 NDVI and the CWSI from MODIS data over the same 11-year period to explore their respective
97 relationship with biomass and the harvest index. The model is then validated using crop statistics data at
98 the scale of the Niger Square Degree site. The proposed approach is finally discussed in the framework of
99 a potential operational yield estimation system that would also include data from the upcoming Sentinel-2
100 sensors.

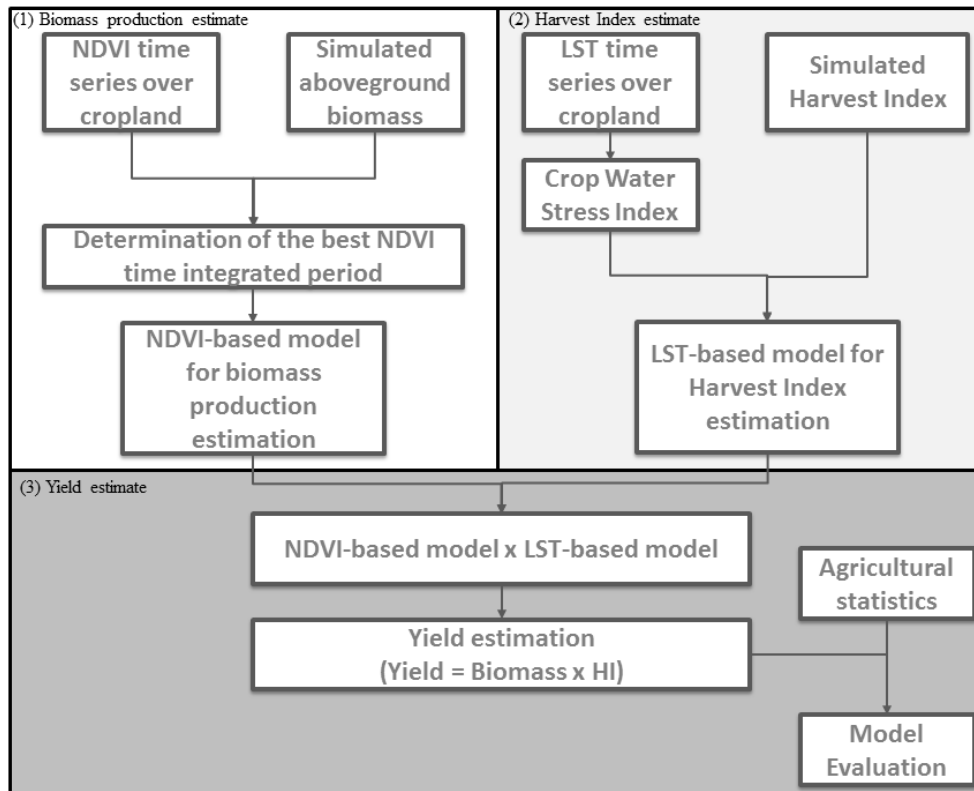
101 2. Material

102 2.1. Overall approach

103 We combine satellite-data with agro-meteorological modeling results to analyze the potential of
104 MODIS derived NDVI and LST time series for pearl millet yield assessment in the Niger Square Degree
105 site. The underlying assumptions of our approach are that:

- 106 (1) Aboveground biomass can be determined from vegetative indices such as the NDVI [44],
- 107 (2) Harvest indices can be significantly reduced under water-limited conditions [45] due to crop water
108 stress. Land surface temperature (LST) observations can be used as an indicator of crop water
109 stress [38] and thus be related to the harvest index,
- 110 (3) The combination of NDVI and LST provides a better estimate of yields than the NDVI on its own
111 in water-limited regions.

112 Fig. 1 summarizes the overall methodology. Empirical statistical relationships are sought (1) between
113 a cropland NDVI integrated over different time periods and aboveground biomass simulated by the crop
114 model SARRA-H, and (2) between a cropland CWSI time series derived from LST data and simulated
115 harvest indices by the same model. **Crop yield is equal to aboveground biomass multiplied by harvest**
116 **index, thus** the relationships obtained are then (3) combined into a simple model for pearl millet yield
117 assessment based on vegetation and thermal indices. Ideally, a remote sensing-based approach has to be
118 calibrated with reliable ground-measurement data. For our study area, the ground truth data currently
119 available are mainly based on farmer's declarative survey and suffer from a lack of consistency both in
120 space and time. Consequently, the choice has been made to use simulated data from SARRA-H crop
121 model to overcome this issue knowing that SARRA-H has been validated for this region [4]. The
122 predictive capacity of the remote sensing-based model is then verified at a regional scale with agricultural
123 statistics.

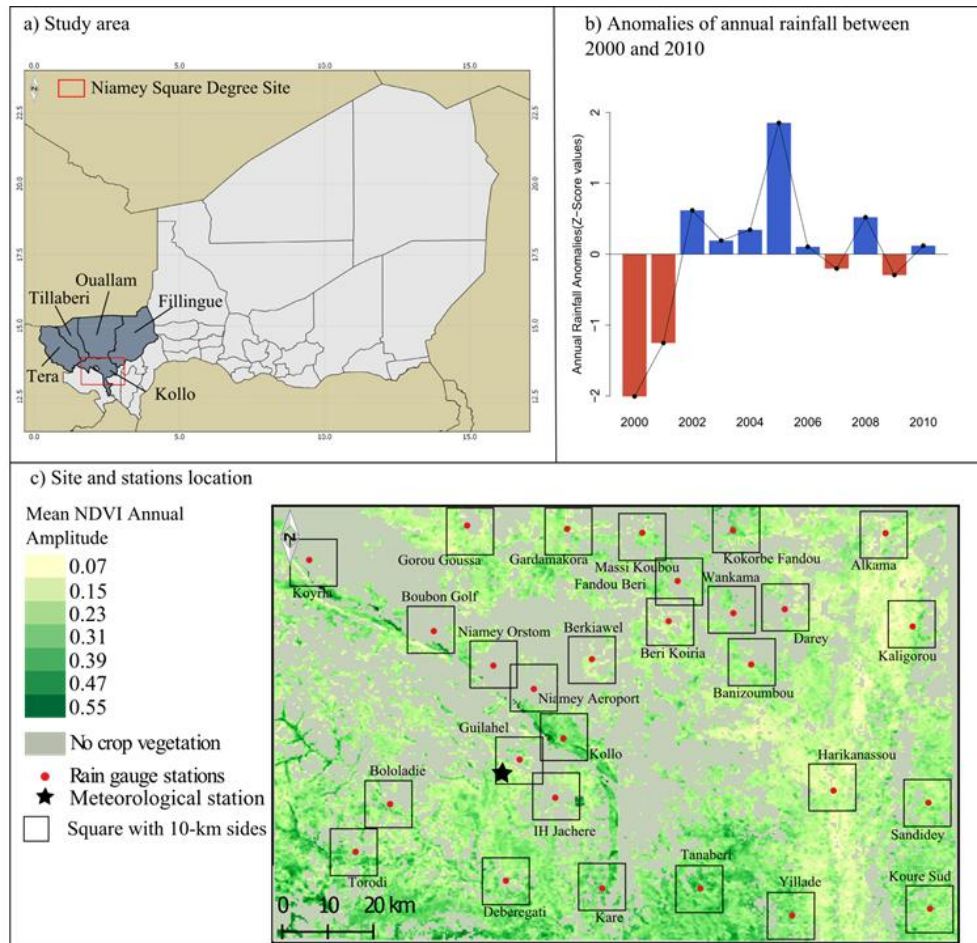


124

125 **Figure 1: Flowchart of the approach adopted corresponding to the three stage of the remote sensed based model**
 126 **development.**

127 **2.2. Study area**

128 The study area (12.9° – 13.9°N; 1.6° – 3.1°E, hereafter referred to as the NSD site) which includes the
 129 Niamey Square Degree site covers about 18,000 km² and is located in South-West Niger (Fig. 2a). The
 130 site is part of the AMMA-CATCH observatory (African Monsoon Multidisciplinary Analysis-Coupling
 131 the Tropical Atmosphere and the Hydrological Cycle; <http://www.amma-catch.org/>) and has been chosen
 132 for two reasons: (1) rainfall is considered as the main driver of crop yield [46], and (2) the site is
 133 instrumented since the early 90's including a dense network of rain-gauges which are continuously
 134 recording rainfall.



135
 136 **Figure 2: The NSD site (Niger) : a) location of the NSD site (red square) and the five departments considered in the study; b)**
 137 **anomalies of annual rainfall (deviation from the mean between 2000 and 2010); c) mean NDVI annual amplitude between**
 138 **2000 and 2010, and location of the 28 rain gauge (red circles) and meteorological (star) stations.**

139 The climate is typically Sahelian. Annual ambient temperatures are high and rainfall distribution
 140 is monomodal during June-September. Rainfall is highly variable spatially [47] and temporally [48] with
 141 10-22% inter-annual variations between 2000-2010 (Fig. 2b). In addition, despite the small size of the
 142 study area (about 160 km x 110 km; Fig. 2c), the regional rainfall pattern shows a high latitudinal gradient
 143 from 480 mm/year (north of the study site) to 630 mm/year (south).

144 The production system is rainfed, dominated by pearl millet [4] which is drought-resistant and well
 145 adapted to the sandy soils predominant in the study area [49]. It is characterized by low inputs [50] and
 146 low yields (generally lower than 700 kg/ha; [49]).

147 **2.3. Satellite data**

148 **2.3.1. The MODIS Vegetation Indices product (MOD13Q1)**

149 The MODIS Vegetation Indices (VI) product (MOD13Q1 collection 5)-was used in this study because
150 of its data consistency, providing spatial and temporal information on vegetation conditions every 16 days
151 at 250-m spatial resolution since 2000 [51]. Even if the MODIS data are pre-processed with the CV-MVC
152 (Constrained view angle-maximum value composites) algorithm, noise still exists in the time series due to
153 cloudiness, sensor problems or Bidirectional Reflectance Distribution Function (BRDF) effects [53]. In
154 consequence, we applied a Savitzky-Golay filter to reduce noise and improve the quality of the NDVI
155 time series towards a more efficient crop yield monitoring [54]. After testing different smoothing
156 parameters, a filter width of 4 and a degree of smoothing polynomial of 6 were retained, which allowed to
157 match the upper envelope of the NDVI time series

158 **2.3.2. The MODIS Land Surface Temperature product (MOD11A2)**

159 The MODIS LST product (MOD11A2, collection 5) is composed of the average value of daily 1-
160 kilometer LSTs under clear sky conditions for an 8-day period [55]. The MODIS LST product was
161 validated with *in situ* temperature measurements recorded at various places and under various surface and
162 atmospheric conditions [56]. According to [56] the MODIS LST accuracy is better than 1 Kelvin. The
163 LST data has been converted to degrees Celsius. As for the MODIS NDVI data, noisy pixels affected by
164 clouds or other atmospheric disturbances were removed when temperatures were below 0°C and the
165 neighboring values in the time series have been linearly interpolated.

166 **2.3.3. The MODIS Land Cover Type product (MCD12Q1)**

167 The MODIS LCP (MCD12Q1, version 51) contains the International Geosphere Biosphere Program's
168 (IGPB) classification, describing 17 land cover classes on a yearly basis at a spatial resolution of 500-m
169 [57], [58]. Two classes are related to agriculture: cropland (class number 12) and cropland/natural
170 vegetation mosaic (class number 14). Assuming that cultivated land cover area did not vary considerably
171 during the 10-year period of study, only "consistent" pixels (i.e. pixels classified as cropland for more than
172 six years between 2001 and 2010) were kept as cropland and the rest masked out. This crop mask was

173 tested against a land cover map based on Landsat images in 2013 and displayed a user accuracy of 73%
174 and a producer accuracy of 50% for the crop classes (not shown here). Because of its availability at a
175 regional scale, we chose to conduct the analysis with the MODIS LCP to ensure the reproducibility of the
176 methodology elsewhere. In this study, we considered that the resulting cropland was approximately
177 equivalent to the pearl millet cultivated area (since pearl millet represents over 70% of the total
178 agricultural production in the study area; [4], [26]).

179 **2.4. Climate data**

180 A set of daily rainfall data recorded throughout the period 2000-2010 at 28 rain-gauges
181 (corresponding to 28 villages) distributed across the study area (Fig. 2c) was used. This dataset was
182 provided by the AMMA-CATCH observing system. Other weather data including daily minimum and
183 maximum air temperature, wind speed, solar radiation and minimum and maximum air relative humidity
184 measurements were obtained from a weather station located south of Niamey (Fig. 2c). According to [50],
185 the variability of other meteorological data is very low compared to rainfall in this area, such that only one
186 weather station was considered necessary.

187 **2.5. Agricultural statistics**

188 Agricultural statistics were used in the validation process of the remotely sensed-based yield model.
189 Pearl millet yield data, collected from ground surveys of major staple crops in Niger, was used. These
190 ground surveys are conducted every year by the Niger Agricultural Statistics service at department level
191 and were therefore available for the 2000 and 2010 period. In this study, yield data for the Kollo
192 department and the four surrounding departments were considered (Fig. 2a).

193 **3. Methods**

194 **3.1. Crop model simulations**

195 **3.1.1. The SARRA-H crop model**

196 SARRA-H V3.3 [43], [50] was used in this study to simulate attainable pearl millet yields under
197 climatic constraint in the NSD site at village level. This model is particularly suited for the analysis of

198 climate impacts on cereal growth and yield in dry environments. It is currently used by AGRHYMET for
199 operational agro-meteorological forecasting across West Africa. It simulates attainable yield under water-
200 limited conditions taking into account potential and actual evapotranspiration, phenology, potential and
201 water-limited assimilation and biomass partitioning (for more details about the SARRA-H crop model see,
202 <http://sarra-h.teledetection.fr>). The crop model SARRA-H has been calibrated and validated for local
203 photoperiod sensitive pearl millet cultivars using ground surveys conducted in various location across
204 West Africa such as in Senegal, Burkina Faso, Mali or Niger [4]. The model was found to perform well
205 over West Africa through comparison with FAO statistics [58]–[60] or in comparison with other crop
206 models in the framework of the Agricultural Model Intercomparison and Improvement Project (AgMIP,
207 [61]).

208 **3.1.2. Aboveground biomass, harvest index and yield simulations**

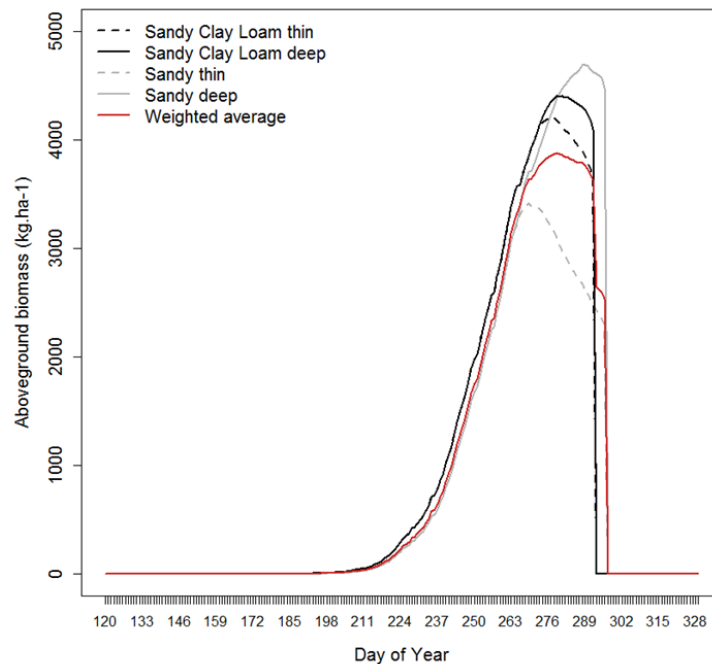
209 Attainable pearl millet aboveground biomass, harvest index and yield were simulated with the
210 SARRA-H crop model for each of the 28 rainfall stations of the NSD site between 2000 and 2010,
211 according to soil type, rainfall regime and agricultural practices (crop varieties and sowing dates). A total
212 of 1276 simulations were conducted. The range of parameters used for the simulation was derived from
213 previous studies and expert knowledge:

- 214 • *Crop varieties*: Two local pearl millet photoperiodic cultivars are found at the NSD site; *Hainy*
215 *Kirey* (90-120 days cycle duration) and *Somno* (120-150 days cycle duration). These two
216 photoperiodic varieties are particularly adapted to spatial and temporal variability of the length as
217 well as the onset of the rainy season of the Sahelian zone [1], [59]. In the NSD site pearl millet
218 HK represents among 80% of the crop [60], [61]. Pearl millet aboveground biomass, harvest index
219 and yields were simulated considering neither fertilization nor irrigation.
- 220 • *Sowing dates*: In Sahelian regions, farmer’s agricultural practices choice is highly determined by
221 the climatic constraints. Farmers generally start sowing photoperiodic millet varieties as soon as
222 possible after the first significant rain, to benefit from the flush of available nitrogen associated
223 with early rains, in spite of a high risk of failure and subsequent need of re-sowing [59], [62]. In

224 the model, the beginning of the time window considered for the search of the satisfying conditions
225 for sowing was set on the 1st May, and the sowing date was automatically generated by the model
226 as the day when simulated soil water available for the plant is greater than 10 mm at the end of the
227 day.

- 228 • *Soil type:* According to the Harmonized World Soil Database [63], 75% of soils in the NSD site
229 are sandy and 25% are sandy clay loam. Since there is no existing data presenting the proportion
230 of each soil type in each of the NSD site's villages respectively, we assumed the proportions
231 proposed for the whole NSD site as being equivalent to the proportion in each village. Yields,
232 aboveground biomass and harvest index were simulated for these two types of soils, weighted
233 according to these proportions and considering two rooting depths (600 mm and 1800 mm) per
234 type of soil.

235 An example of the aboveground biomass output obtained for Torodi village in 2008 is presented in Fig. 3.



236
237 **Figure 3: Example of Somno millet simulated aboveground biomass with the crop model SARRA-H for the village of Torodi in**
238 **2008. The dark curves represent aboveground biomass for a sandy clay loam soil (thin soil in dashed line and deep soil in solid**
239 **line), the gray curves represent a sandy soil (thin soil in dashed line and deep soil in solid line) and red curve represent the**
240 **resulting weighted average.**

241 **3.2. Relationships between crop model simulations and remote sensing** 242 **indices for pearl millet**

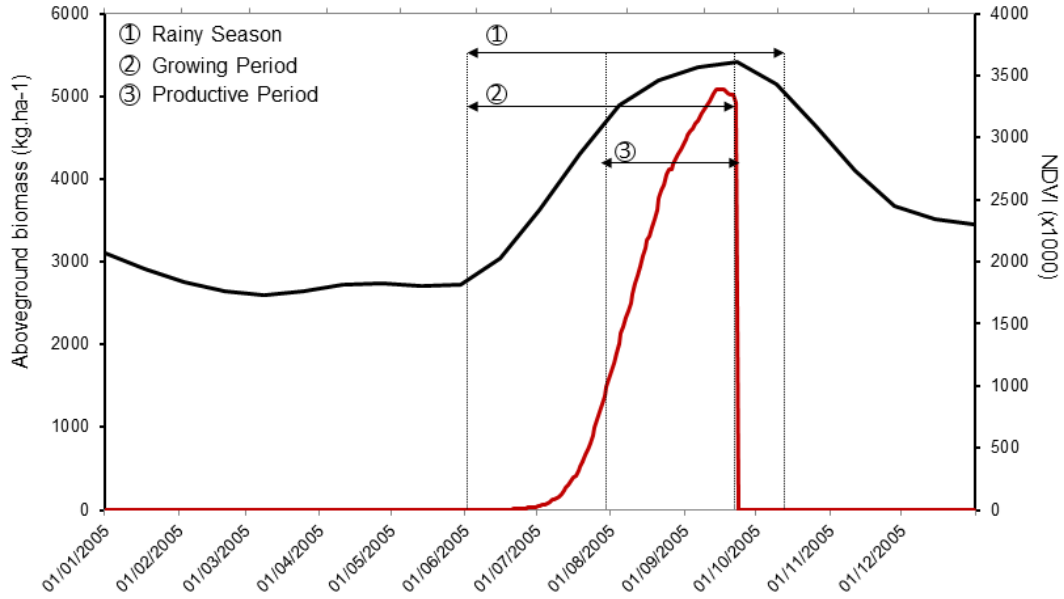
243 **3.2.1. Processing of remote sensing indices**

244 *MODIS NDVI time series*

245 In Niger cultivated areas are principally gathered around villages within a distance of less than 10-km
246 [26]. To compare the NDVI with simulated aboveground biomass, NDVI median values within a square
247 of 10-km by 10-km (corresponding to 1600 MODIS pixels) around each village were extracted in order to
248 limit the analysis to areas with the higher density of crop surfaces. The median value was used to represent
249 the average situation while minimizing the effect of pixels with a significant proportion of natural
250 vegetation as can be expected when working with a broad-scale crop mask. Mean values were also not
251 appropriate because the NDVI values were found not to be normally distributed. In this study three NDVI
252 time integrals (cumulative values) were defined (Fig. 4):

- 253 • *The rainy season* (NDVI_RS) extends from its onset to its retreat. In order to take into
254 consideration the spatial and temporal variability of the length of the rainy season, the onset and
255 retreat of the rainy season was computed for each year and for each village of the NSD site
256 following Sivakumar's definition [64].
- 257 • *The growing period* (NDVI_GP) extends from the onset of the rainy season to the end of
258 September (Fig. 4). The end of the crop growing period corresponds approximatively to the
259 harvesting period which was fixed here to the end of September (270th day of the year) since it
260 generally occurs during September [65].
- 261 • *The productive period* (NDVI_PP) of the crop growing period corresponds to phenological stages,
262 such as the reproductive or the first maturing stages, which are especially sensitive to water stress.
263 Consequently, yield loss becomes significant under water stress conditions during these drought-
264 sensitive stages. The NDVI between the beginning of August (213th day of the year) and the end
265 of September (including the reproductive and the maturation phases as well as the harvesting
266 period) were used to calculate the NDVI integral during the crop productive period.

267 The cropland extent obtained from the MODIS Land Cover Product was used to keep only cropland
 268 classified pixels in the NDVI integral calculation which allows minimizing the influence of natural
 269 vegetation signals.



270
 271 **Figure 4: The three NDVI time integrated periods, example of the Kollo village in 2005. NDVI profile (black line) during the**
 272 **season is compared to the simulated aboveground biomass (red line).**

273 *The Crop Water Stress Index (CWSI)*

274 The CWSI, commonly used as a plant stress detection index, is originally based on canopy-air
 275 temperature difference and their relation to air vapor pressure deficit. It ranges from 0 (ample water) to 1
 276 (maximum stress) [38]. [66] suggest an equivalent approach based only on canopy-air temperature
 277 differences. The CWSI used in this study can be expressed as:

$$CWSI = \frac{(T_c - T_a)_{ref} - (T_c - T_a)_{min}}{(T_c - T_a)_{max} - (T_c - T_a)_{min}} \quad (1)$$

278
 279 where T_c is the canopy-temperature from MODIS LST data and T_a is the air temperature measurement
 280 from the meteorological AGRHYMET station. Subscripts *min*, *max* and *ref* refer to the minimum (non-
 281 stressed crop), maximum (cover no longer transpiring), and observed canopy-air temperature differences

282 respectively, computed for each date within the crop mask over the study area. Since the HI of pearl millet
283 is more sensitive to water stress during the crop productive period of the growing season [35], an integral
284 of CWSI was calculated over the productive period as defined previously (CWSI_PP).

285 3.2.2. Model development for aboveground biomass, HI and yield estimation

286 In Niger, pearl millet is characterized by a LAI generally lower than 2, which suggests that the
287 relationships between NDVI and LAI are below the saturation level explained in [67]. The relationship
288 between simulated aboveground biomass and each of the three NDVI time integrals was modeled with an
289 Ordinary Least Square regression (OLS) through the following expression:

$$SimBiom_{t,n} = b1 + a1 * NDVI_{t,n} + \epsilon_{1,t,n} \quad (2)$$

290
291 where $SimBiom_{t,n}$ represents the simulated aboveground biomass in year t and village n with the crop
292 model SARRA-H, $NDVI_{t,n}$ is the NDVI variable for the same year and village, $b1$ and $a1$ are the
293 parameters to be estimated and $\epsilon_{1,t,n}$ is the error term. An OLS was run at village level for the three NDVI
294 time integrals.

295 As for the aboveground biomass estimation, an OLS regression was applied to derive HI from the CWSI,
296 while the crop model output was used to calibrate the remote sensed based model:

$$SimHI_{t,n} = b2 + a2 * CWSI_{t,n} + \epsilon_{2,t,n} \quad (3)$$

297
298 where $SimHI_{t,n}$ represents the simulated HI in year t and village n with the crop model SARRA-H,
299 $CWSI_{t,n}$ is the CWSI variable for the same year and village, $b2$ and $a2$ are the parameters to be estimated
300 and $\epsilon_{2,t,n}$ is the error term.

301 The basic equation to estimate yield, is:

$$Yield = biomass * HI \quad (4)$$

302 Thus, by replacing each term of Eq. (4) by Eq. (2) and Eq. (3), the following model for yield estimation
303 can be derived (Eq.5):

$$Yield = (b1 + a1 * NDVI_{t,n} + \varepsilon1_{t,n}) \times (b2 + a2 * CWSI_{t,n} + \varepsilon2_{t,n}) \quad (5)$$

304 **4. Results**

305 **4.1. Crop model simulation results**

306 The crop model SARRA-H was run for the 28 villages of the NSD site for a period of 11 years (from
307 2000 to 2010). In these simulations, the mean annual simulated yields at village scale vary from 100
308 kg ha⁻¹ to 1400 kg ha⁻¹ (not shown). The yields are in the same order of magnitude that the ones measured
309 by CIRAD (French agricultural research center for development) and AGRHYMET in the NSD site
310 between 2004 and 2008 (400 to 1100 kg ha⁻¹; [71]). The temporal and spatial variability of the outputs of
311 the simulation protocol are presented in Table 1 and Table 2 respectively. Table 1 shows a general high
312 temporal variability of simulated pearl millet aboveground biomass for the 28 villages with a coefficient
313 of variation (CV) ranging from 31% for Gorou Goussa to 63% for Kollo. Compared to the high year-to-
314 year variability of the aboveground biomass, the temporal variability of the simulated yields (CV ranged
315 from 19% to 46% between 2000 and 2010) and harvest indices (CV below 40% and mean HI = 0.29) are
316 moderate. Given the size of the study area, the aboveground biomass, the HI and the yield's spatial
317 variability could be considered relatively high (CV between 9% and 59% Table 2). The years 2000, 2002,
318 2007 and 2010 are those showing the highest spatial variability between the villages (e.g. 30%, 36%, 30%
319 and 52% respectively, for simulated yields). The analysis of the crop model output during these years (not
320 shown) reveals high water stress conditions at the beginning of the growing period (during the vegetative
321 stage), affecting locally some of the villages and resulting in very low simulated aboveground biomass
322 and yields for those years. We also validate the SARRA-H crop model against agricultural statistics by
323 averaging simulated yield at the NSD site level (Fig.5). The yields are overestimated which is one of the

324 main drawbacks of many crop models since they simulate potential yields limited by water supply which
 325 could be different from the actual yields attained in the field [72].

326 **Table 1: Temporal variability of simulated aboveground biomass, harvest index (HI) and yield. The mean values and the**
 327 **coefficients of variation (CV) are calculated on the 2000-2010 period, and are given for each village. In bold, the values**
 328 **averaged of means and CV over the dataset are given.**

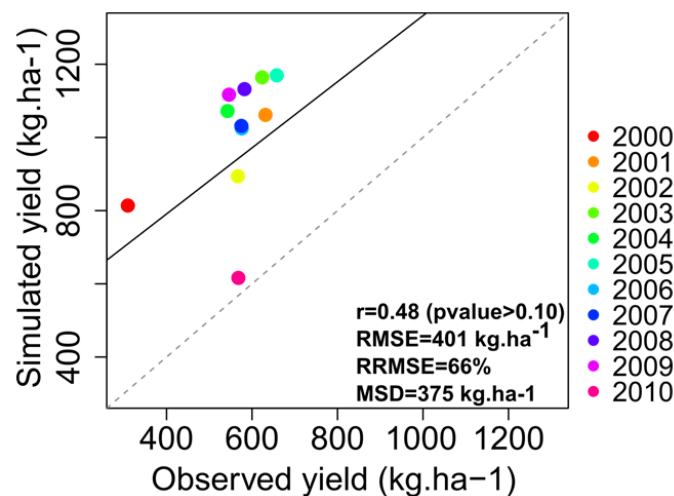
	Aboveground biomass		Harvest Index		Yield	
	Mean (kg ha ⁻¹)	CV (%)	Mean	CV (%)	Mean (kg ha ⁻¹)	CV (%)
Alkama	2063	46	0.29	27	813	33
Banizoumbou	2285	51	0.27	26	768	34
Beri Koira	2290	48	0.31	33	911	19
Berkiawel	2365	52	0.30	30	920	32
Bololadie	2138	58	0.29	28	789	46
Boubon Golf	2387	44	0.31	31	972	17
Darey	2012	40	0.32	21	914	22
Debere Gati	2381	55	0.29	24	888	41
Fandou Beri	2001	43	0.31	22	903	33
Gardamakora	2066	51	0.30	33	808	36
Gorou Goussa	2653	31	0.26	27	956	18
Guilahel	2416	52	0.28	35	855	30
Harikanassou	2732	33	0.27	22	1033	9
IH Jachere	2254	49	0.30	29	902	23
Kaligorou	2349	35	0.28	27	896	25
Kare	2318	51	0.30	25	922	30
Kokorbe Fandou	1936	62	0.33	31	829	37
Kollo	2074	63	0.30	36	754	41
Koure Sud	2321	42	0.29	21	940	25
Koyria	2350	38	0.29	26	924	18
Massi Koubou	2155	49	0.30	35	857	34
Niamey Aeroport	2386	52	0.30	31	892	25
Niamey Orstom	2103	51	0.32	26	907	24
Sandideye	2573	42	0.28	26	963	25
Tanaberi	2302	38	0.29	22	949	25
Torodi	3271	43	0.24	39	934	37
Wankama	1915	49	0.32	23	844	37
Yillade	2674	40	0.27	24	994	19
Mean	2313	47	0.29	28	894	28

329
 330 **Table 2: Spatial variability of simulated aboveground biomass, harvest index (HI) and yield. The mean coefficients of variation**
 331 **(CV) are calculated on the 28-village data set, and are for each year. In bold, the values averaged of means and CV over the**
 332 **dataset are given.**

Aboveground biomass	Harvest Index	Yield
---------------------	---------------	-------

	Mean (kg ha ⁻¹)	CV (%)	Mean	CV (%)	Mean (kg ha ⁻¹)	CV (%)
2000	2332	24	0.22	21	719	30
2001	2536	23	0.25	23	943	18
2002	1501	52	0.35	15	768	36
2003	3054	24	0.24	17	1050	14
2004	2386	34	0.28	16	949	22
2005	3967	23	0.21	17	1082	12
2006	1706	26	0.36	16	911	15
2007	1989	41	0.31	17	879	30
2008	2781	28	0.27	21	1029	11
2009	2365	41	0.31	24	990	19
2010	828	59	0.43	9	518	52
Mean	2313	34	0.29	18	894	24

333



334

335 Figure 5: Observed pearl millet yields from agricultural statistics for the department of Kollo vs simulated yields obtained
336 with SARRA-H aggregated at the NSD site level.

337 4.2. Biomass estimation based on NDVI data

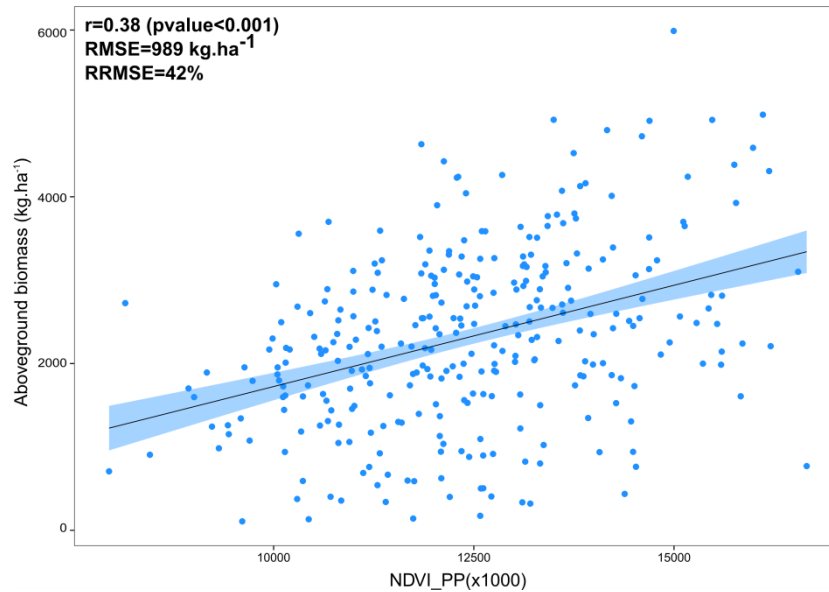
338 4.2.1. Results at village scale

339 In order to support the choice of using median values to extract NDVI around each villages, different
340 descriptive statistics have been extracted for each of the NDVI-integrated variables in order to determine
341 the best NDVI-integrated x descriptive statistics combination for aboveground biomass estimation: the
342 median value, the maximum value, the range (the difference between the maximum and the minimum)
343 and the standard deviation. The results are illustrated in Table 3. The four descriptive statistics x the three
344 NDVI-integrated variables were compared to the simulated aboveground biomass using an OLS

345 regression (Table 3). For all the combinations tested the correlation coefficients are low (below 0.40 but
 346 all highly significant). The Root Mean Square Errors (RMSE) is high with an RMSE equal to 989 kg ha⁻¹
 347 (RRMSE=42%) for the best combination (NDVI median x NDVI_PP), and an RMSE equal to 1060 kg ha.
 348 ₁ (RRMSE=46%) for the less performing combination (NDVI range x NDVI_RS). Fig. 6 shows the
 349 resulting scatterplot of NDVI_PP *versus* simulated aboveground biomass. The dispersion of the points
 350 along the regression lines suggests the low ability of MODIS NDVI to reveal spatial and temporal
 351 aboveground biomass variability at a village scale. According to the Table 3, the best results are observed
 352 for the median NDVI values extracted around villages, thus only the combination NDVI median x NDVI-
 353 integrated variables are considered in the remainder of the study.

354 **Table 3: Elements of the regression analysis obtained between the simulated aboveground biomass and the descriptive**
 355 **statistics x NDVI variables (NDVI integrated during the rainy season, the growing period and the productive period) obtained**
 356 **at the village scale for years 2000-2010.**

Descriptive statistics	NDVI Variables	Intercept	Slope	r	p-value	RMSE (kg ha-1)	RRMSE (%)
Median	NDVI_RS	336	0.07	0.32	6.08E-09	1012	43
	NDVI_GP	255	0.10	0.34	1.20E-09	1006	43
	NDVI_PP	-704	0.24	0.38	5.80E-12	989	42
Max	NDVI_RS	893	0.03	0.26	5.07E-06	1033	45
	NDVI_GP	437	0.06	0.33	5.11E-09	1011	43
	NDVI_PP	-99	0.13	0.31	1.86E-08	1017	44
Range	NDVI_RS	1886	0.01	0.13	0.02	1060	46
	NDVI_GP	1585	0.04	0.21	0.0002	1046	45
	NDVI_PP	1634	0.06	0.18	0.002	1052	46
Standard Deviation	NDVI_RS	1810	0.14	0.16	0.005	1056	45
	NDVI_GP	1389	0.34	0.24	3.07E-05	1039	45
	NDVI_PP	1280	0.59	0.21	0.0001	1045	45



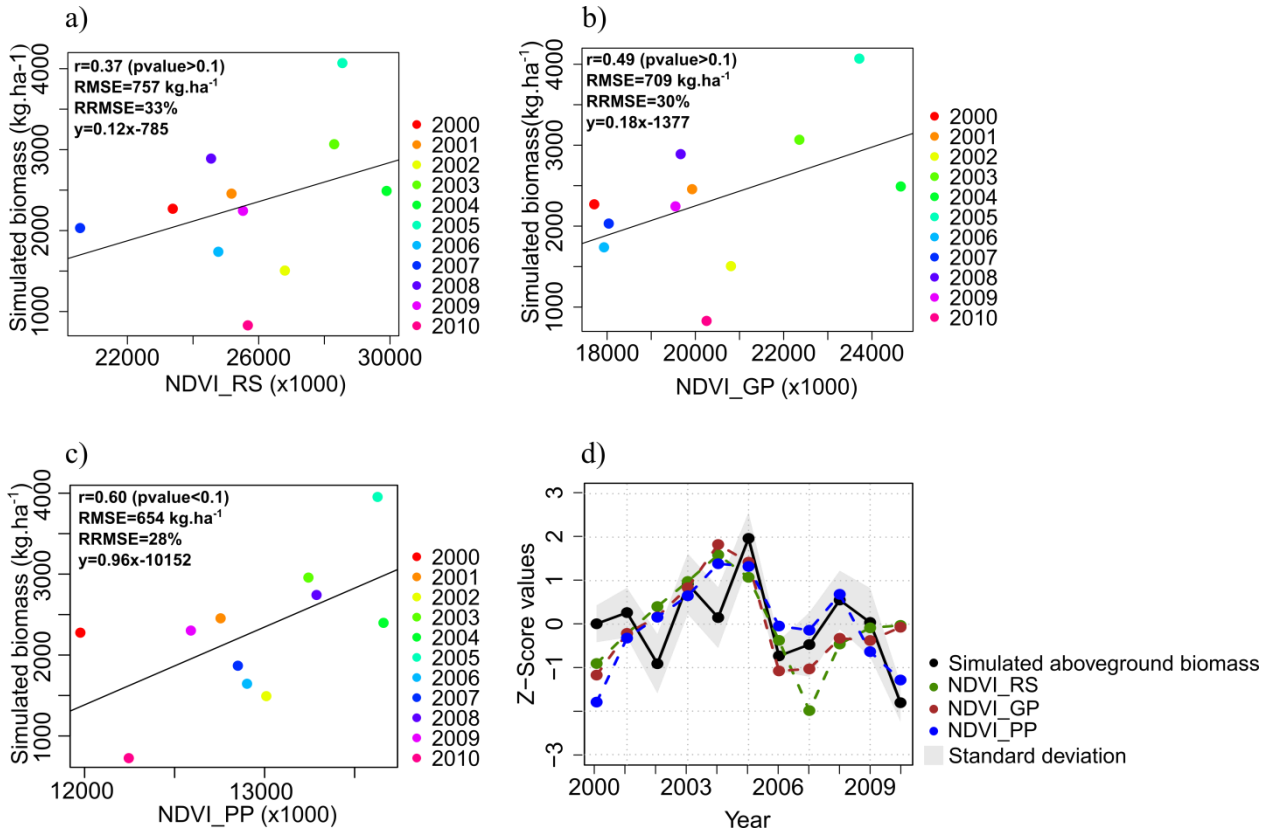
357
 358 **Figure 6: Scatterplot of the simulated aboveground biomass (kg ha⁻¹) and the NDVI integrated over the productive period for**
 359 **the 28 villages of the NSD site and over the 11 years of data. The RMSE of the aboveground biomass is 989 kg ha⁻¹ which is**
 360 **equivalent to a RRMSE of 42%, and the correlation coefficient is 0.38. The solid line is the linear regression line and the blue**
 361 **area is the confidence interval for pvalue<0.1.**

362 4.2.2. Results at the NSD site scale

363 Since neither NDVI observations nor simulated aboveground biomass follow normal distributions,
 364 median values were preferred to mean values to compare NDVI and simulated aboveground biomass over
 365 the 28 villages (NSD site scale). **The aggregated NDVI value at the NSD site scale was computed**
 366 **considering the NDVI median value for all cropped pixels of the 28 villages.** Fig. 7 shows that overall
 367 NDVI observations represent well the magnitude of the simulated aboveground biomass variability (Fig.
 368 7a, 7b and 7c) as well as the global trends and extreme events (Fig. 7d). Among the three NDVI variables,
 369 the NDVI_PP presents the best indicator of pear millet aboveground biomass with a correlation coefficient
 370 **0.60** (significant at 10%) and a RMSE of **654 kg ha⁻¹** which is equivalent to a RRMSE of **28%**, whereas
 371 NDVI_RS appears to be the less reliable indicator (Fig 7c and Fig. 7a, respectively). The year-to-year
 372 variability is correctly displayed, with a positive trend between 2000 and 2005, a negative trend between
 373 2005 and 2010, and NDVI observations differing from simulated aboveground biomass by less than one
 374 standard deviation (Fig. 7d). These NDVI trends coincide with the observed rainfall anomalies at the NSD
 375 site scale (Fig. 2b). At the site scale, the remote sensing based model for aboveground biomass estimation
 376 is expressed as follows:

$$\text{Biomass} = 0.96 * \text{NDVI_PP} - 10152 \quad (6)$$

377 where *Biomass* is the production of pearl millet aboveground biomass estimated at the harvest period in
 378 kg ha^{-1} , and *NDVI_PP* is the NDVI integral during the productive period at the NSD site scale.



379
 380 **Figure 7: SARRA-H simulated aboveground biomass (kg ha^{-1}) vs a) MODIS NDVI integrated during the rainy season, b) MODIS**
 381 **NDVI integrated during the growing season, and c) MODIS NDVI integrated during the productive period. The regression line**
 382 **is in black solid line. d) Comparison of the interannual variability of simulated aboveground biomass and NDVI observations,**
 383 **expressed in z-score values. The grey area is the \pm standard deviation computed from simulated aboveground biomass.**

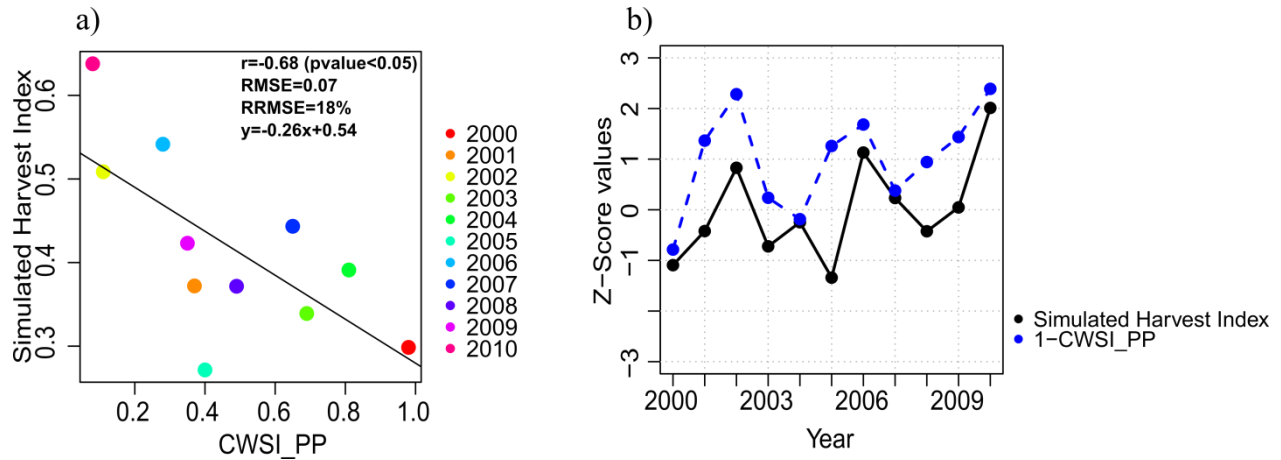
384 4.3. Harvest index estimation based of LST data

385 Since aboveground biomass is estimated at the NSD site scale, the model for HI estimation was
 386 developed at this same scale by taking the median value of the *CWSI_PP* derived from the LST data, and
 387 integrated over the crop productive period. The resulting model is presented in Fig.8, which shows that the
 388 HI and the *CWSI_PP* are linearly and negatively correlated, with a correlation coefficient of **-0.68**
 389 (significant at 5%) and a RMSE of **0.07** (Fig.8a). This relationship may be explain by a new biomass

390 production allocated to grain decreasing as crop water stress increases, leading to a consequent decrease in
 391 yield. In order to better visualize the year-to-year variability of both simulated HI and $CWSI_{PP}$, we have
 392 plotted the $(1-CWSI_{PP})$ value (Fig.8b). The year-to-year variability is generally well represented by the
 393 $CWSI_{PP}$ except for 2005. The model derived for the HI estimation is expressed as follows:

$$HI = -0.26 * CWSI_{PP} + 0.54 \quad (7)$$

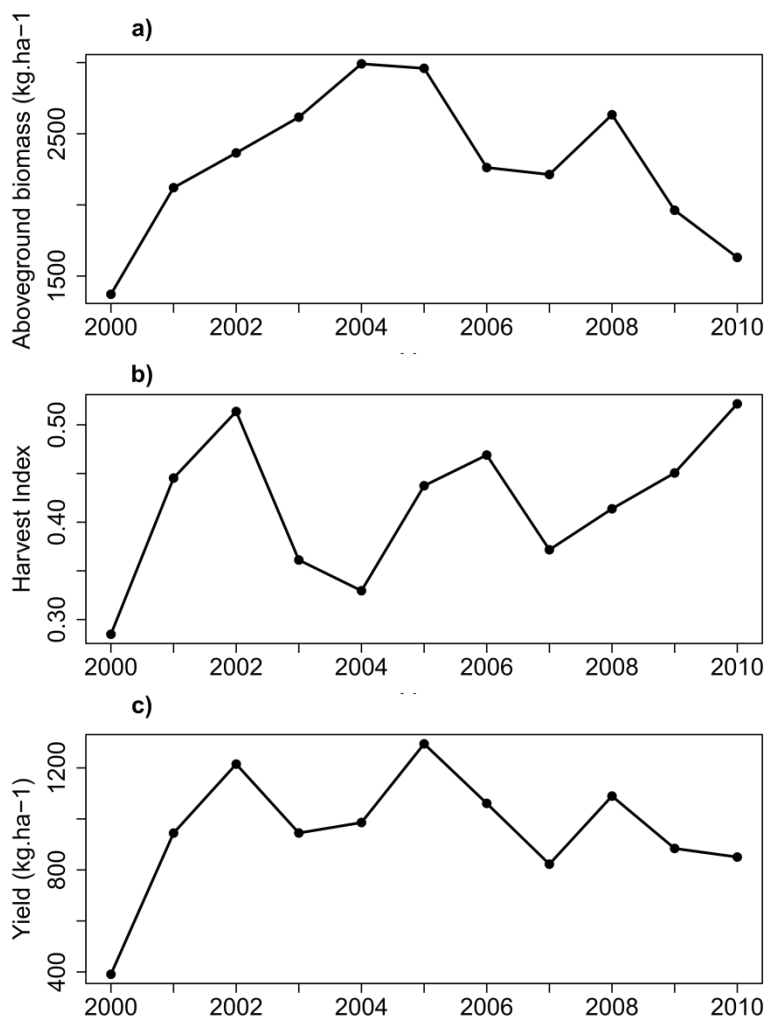
394 where HI is the estimated harvest index and $CWSI_{PP}$ is the Crop Water Stress Index's integrated over
 395 the productive period at the NSD site scale.



396
 397 **Figure 8:** a) SARRA-H simulated harvest index vs $CWSI_{PP}$ estimated from MODIS LST data, over the 2000-2010 period (the
 398 regression line is in black solid line); b) comparison of the interannual variability of SARRA-H simulated harvest index and $(1-$
 399 $CWSI_{PP})$ values, expressed in z-score values. The grey area is the \pm standard deviation computed from simulated harvest
 400 index.

4.4. Yield estimation based on NDVI and LST data and evaluation

401 Pearl millet yields at the NSD site scale were obtained by multiplying the estimated aboveground
 402 biomass (Eq.6; Fig.9a) by the estimated HI (Eq.7; Fig.9b). The estimated yields vary from 390 kg ha⁻¹ to
 403 1294 kg ha⁻¹ (Fig.9c). The estimated yields show an overall stable trend between 2000 and 2010 and a
 404 decline between 2005 and 2007 (Fig.9c).
 405



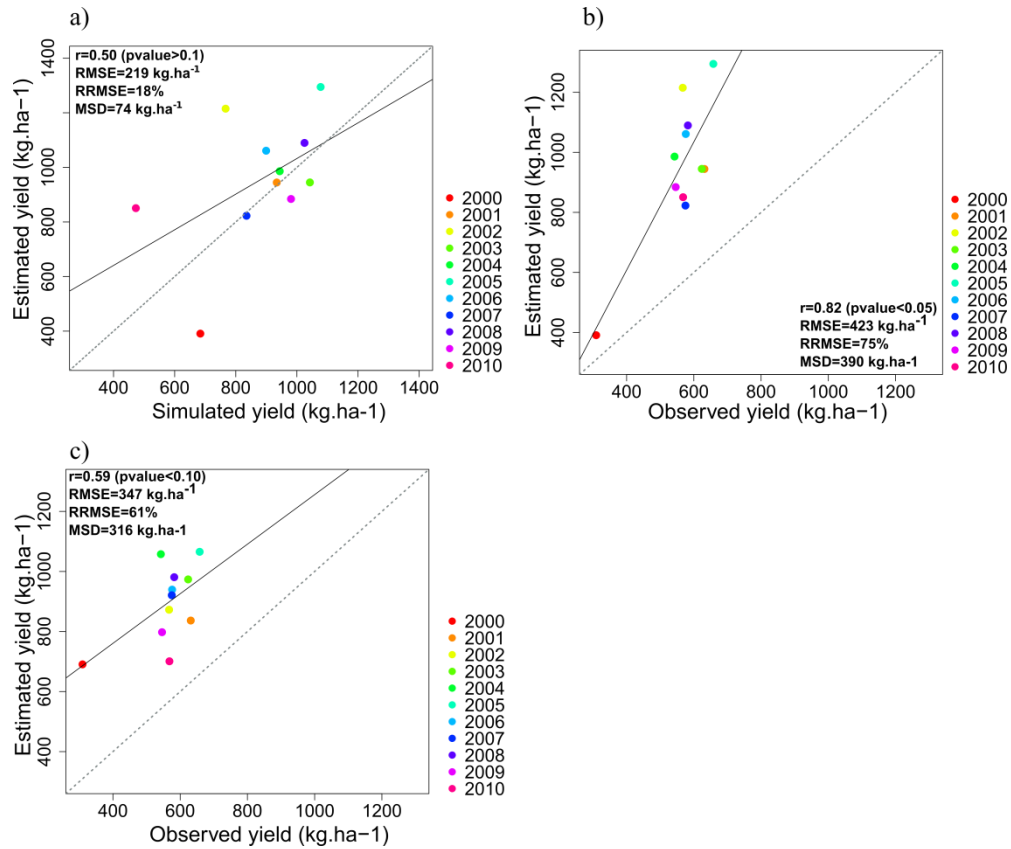
406
 407 **Figure 9: Evolution of a) the aboveground biomass estimated from the MODIS NDVI model (Eq. 6), b) the harvest index**
 408 **estimated from the MODIS-derived CWSI model (Eq. 7), and c) the resulting pearl millet yield derived from the combination**
 409 **of Eq. 6 and Eq. 7, over the study site.**

410 The predictive capacity of the remote-sensing-based model for pearl millet yield estimation is
 411 shown in Fig.10. The combined model based on NDVI and LST data is first evaluated by comparing
 412 simulated crop yield (from SARRA-H) to estimates based on the remote sensing-based model (Fig. 10a).
 413 **The combined model is in moderate agreement with simulated yields ($r=0.50$, RMSE of 219 kg ha⁻¹ and a**
 414 **Mean Signed Difference [MSD] of 74 kg ha⁻¹; Fig.10a).**

415 In order to show the contribution of thermal indices in crop yield estimation, we compared the
 416 results with the estimated yields based only on NDVI data. The model (based on both NDVI and LST
 417 data) results are in good agreement with the official yield statistics ($r=0.82$ significant at 5%, Fig.10b).

418 Furthermore, the combination of NDVI and LST data clearly contributes to improve yield estimation
419 compared to NDVI data alone ($r=0.59$, Fig.10c). However, like the crop model used for the calibration,
420 the remote sensing-based models clearly overestimate yields (Fig.10b and Fig.10c) which leads us to
421 consider the ability of these models to render the yield's year-to-year variability observed by the
422 agricultural statistics. To do so, both estimated and observed yields were normalized. For each year, the
423 absolute differences between agricultural statistics z-score values and those of the models were computed
424 (Fig. 11a). In order to provide an overall indication on the performance of each of the models, the sum of
425 the absolute differences is also assessed. Yield's year-to-year variability from 2000 to 2010 is quite well
426 rendered in both models in Fig. 11a, particularly for the second half of the period (between 2005 and
427 2010). The combined model based on NDVI and LST data is the closest of the agricultural statistics
428 temporal profile (absolute difference sum = 5.61), particularly in extreme dry years such as in 2000
429 (Fig.11b). Nevertheless, the overall trend is also well transcribed, split in a stable period between 2000
430 and 2005, followed by a decrease trend in yields between 2005 and 2010 (Fig.11b).

431 To test the robustness of the remote sensing-based model, yields for the four surrounding
432 departments were computed and compared with the corresponding official yield statistics (Table 4).
433 Overall, computed yields coincide with the yield statistics, with correlation coefficients above 0.50
434 (significant at 10%) for 3 departments (Table 4). As for the NSD site the remote-sensing based model
435 systematically overestimates yields (RMSE ranging from 237 kg ha⁻¹ to 742 kg ha⁻¹).



436

437 **Figure 10: a) Simulated yields from SARRA-H vs estimated yields from the combination of NDVI and LST data, and vs**
 438 **estimated yield from remote sensing with (b) or without (c) LST data. The 1:1 line is given in grey dashed line.**

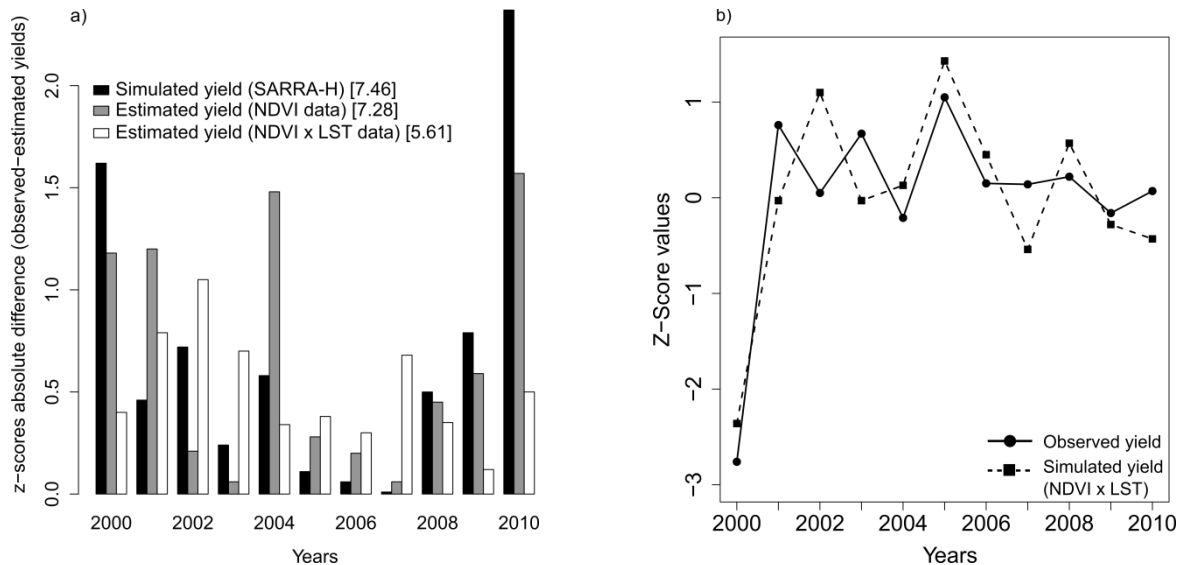


Figure 11: a) Year-to-year yield variability (SARRA-H, NDVI data, NDVI x LST data) comparison with agricultural statistics. The
y-axis indicates the absolute difference between yields anomalies (expressed in z-score) estimated and yield anomalies from
agricultural statistics. In brackets are specified the sum of absolute differences. b) Agricultural statistics and simulated yields
(NDVI x LST data) standardized anomalies (in z-score).

439

Table 4: Estimated yields from the remote-sensed based model vs the agricultural statistics yields

	r	p-value	RMSE (kg ha ⁻¹)
Fillingue	0.45	0.15	238
Kollo	0.82	0.01	423
Ouallam	0.23	0.48	237
Tera	0.58	0.06	505
Tillaberi	0.64	0.03	742

441 5. Discussion

442 5.1. Aboveground biomass estimation based on NDVI time series

443 The first stage of the remote sensing-based model consisted in developing an empirical
 444 relationship between NDVI time series and pearl millet aboveground biomass simulated by the crop
 445 model SARRA-H.

446 The study first highlighted that the ability of the MODIS NDVI time series to estimate
 447 aboveground biomass depends on the scale considered. At the village scale (considering the whole dataset:
 448 28 villages, 11 years) , the study found out that the MODIS NDVI time series are not able to reveal both
 449 the spatial and temporal variability of the simulated aboveground biomass (RRMSE > 40%; Table 3 and
 450 Fig.6). As previously shown by [46], in the semi-arid zone of Niamey, aboveground biomass and final
 451 yields are mainly influenced by the spatio-temporal distribution of rainfall, and so a high variability of
 452 aboveground biomass can be observed between villages which are only a few kilometers apart. Thus, the
 453 low correlation between NDVI and aboveground biomass at the village scale implies that the spatial
 454 variability of NDVI is not as strongly associated with the spatial variability of rainfall. Further analyses
 455 are required on other potential factors that could influence NDVI at this scale. We could assume for
 456 instance that, in semi-arid regions where vegetation cover is relatively sparse, soil may cause high
 457 variations in the NDVI values at such a small scale, causing NDVI values artifacts [74] and therefore
 458 reducing the correlation between NDVI and aboveground biomass. [16], considering a direct relation
 459 between NDVI and yield, found that including soil information improved yield prediction in the Peanut
 460 Basin in Senegal. On the other hand, at the NSD site scale (temporal analysis), a good correlation was

461 found between simulated aboveground biomass and NDVI_PP ($r=0.60$). This improvement could be
462 explained by (1) the reduction of the noise in the NDVI time series when aggregating at a coarser level
463 and (2) a better representativeness of the overall crop growth conditions over the NSD site that is mainly
464 driven by rainfall variability.

465 The capacity of the MODIS NDVI time series to estimate aboveground biomass depends also on
466 the time period used for the integration. On that point our results are different from [11], [75] who found
467 a good correlation between NDVI integrated over the whole growing season and aboveground biomass in
468 Senegal. In these studies, only natural herbaceous vegetation was considered, for which final aboveground
469 biomass is not much different from vegetative biomass, thus justifying NDVI integration over the entire
470 length of the growing season. Our study focuses on a final aboveground biomass that depends on both
471 vegetative biomass and grains. NDVI values were integrated over the crop productive period to account
472 for grains, since it corresponds to the reproductive period and maturation phases, which include grain
473 filling when plants reach their maximum development [26]. Our results corroborate other studies that
474 directly relate NDVI to yields such as [76] who found that the strongest correlation of NDVI with wheat
475 yields is achieved when taking into consideration NDVI values around their maximum which includes the
476 sensitive stages of grain production.[15] then tested the influence of different NDVI integration periods
477 and found a coefficient of determination $R^2=0.50$ (i.e. $r=0.70$) for the productive period. In another
478 analysis, using NOAA AVHRR data between 1982 and 1990 for Niger, [26] concluded that the best time
479 integration period for millet and sorghum yield assessment is from August to September. Finally, more
480 recently in a study conducted in China [77] it was also found that the productive and maturing stages
481 including the heading, flowering and filling of the crops are the best suitable periods for yield estimation
482 of paddy rice, corn and winter wheat due to the stress sensitivity of these periods that would lead to
483 biomass reduction and thus potentially yield losses. In our study, NDVI_RS and NDVI_GP (both
484 determined by the onset of the rainy season) appear to be less correlated to aboveground biomass. A
485 potential explanation for this could be the delay between the NDVI onset of the growing season and the
486 calculated start-of-season, which occur one month apart, as previously shown by [78]. At the beginning of

487 the growing season in a MODIS pixel the proportion of the millet cover is probably lower than the
488 proportion of the surrounding natural vegetation. The latter reacts immediately to the first significant
489 rainfall, whereas crops are sown later, when sufficient water (>10 mm) is available in the soil [1] and have
490 a growth rate lower than natural vegetation.

491 On the year-to-year variability analysis, a decrease in both the simulated aboveground biomass
492 and the NDVI was observed from 2005 to 2010, with an important decline between 2005 and 2006 (Fig.
493 7d.). When comparing this result with annual rainfall anomalies (Fig. 2b) and it can be concluded that
494 both aboveground biomass and NDVI follow the major trends of rainfall anomalies (as seems particularly
495 evident between 2005 and 2006). This comes in support of the previous assumption that rainfall remains
496 the main determinant of NDVI variability at the NSD site scale.

497 **5.2. Harvest index estimation based on an indicator of crop water stress: the** 498 **CWSI**

499 For most crop models, including SARRA-H, DSSAT and CROPWAT [43], [79], [80], water
500 stress during the reproductive and maturation phases is considered a crop yield limiting factor. In the
501 remote-sensing model, we take into account the crop water stress effect on yield through the use of the
502 CWSI, an indicator based on LST. To our knowledge, it is the first time that a link is sought between an
503 indicator of crop stress and HI. An overall good correlation ($r=-0.68$) was found between HI and
504 CWSI_PP at the NSD site scale, meaning that the HI decreased linearly as the water supply became more
505 limited for plants. However, as for the use of vegetation indices in semi-arid zones, the main issue with
506 thermal indices based on canopy temperature is the spatial heterogeneity due to the soil influence when
507 the canopy does not completely cover the ground. Because bare soil is often much warmer than the air, the
508 soil background temperature included in the LST can lead to false detections of crop water stress [81]. To
509 overcome this limitation, a possibility may be to use the Water Deficit Index developed by [69], which
510 considers both the difference between air and surface temperatures and the fraction of crop cover derived
511 from vegetation indices, to estimate the water status. This method was not tested in this study, as some

512 adaptations are ongoing to test the construction of the vegetation index – temperature trapezoid from
513 satellite time series.

514 **5.3. Estimation of pearl millet yields**

515 The two previous approaches for aboveground biomass and HI estimation were combined into a
516 simple, robust and timely satellite-based model of rainfed cereal yield, applicable at the department level.
517 **If in absolute values, yields are overestimated compared to official agricultural statistics of the Kollo**
518 **department, the analysis of the standardized values has shown a good agreement in terms of year-to-year**
519 **variability reproduction, translating into a high correlation with statistics.** In their recent meta-analysis [8]
520 found that for four studies conducted in Senegal, Burkina Faso and Niger using NOAA AVHRR data, the
521 correlation coefficients between NDVI alone and millet yield **were comprised between 0.75 and 0.94**
522 **which is comparable to the present work ($r=0.82$).** However, caution in the interpretations has to be taken
523 particularly because (1) although the size of the study area considered in these studies is similar to that of
524 the present study (i.e. results aggregated at a department level), the time period considered was much
525 shorter (2 years in [15]) and (2) when the time period considered is comparable to ours, results were
526 aggregated at higher administrative levels than for us (several departments or country level; e.g. [16],
527 [28]).

528 The comparison with a model based only on NDVI has highlighted the usefulness of combining
529 vegetation and thermal indices (NDVI and CWSI) for yield estimation. **The ability to render the year-to-**
530 **year variability of pearl millet yield was clearly improved through this combination, with a correlation**
531 **coefficient increasing from 0.59 to 0.82 and the z-score absolute difference sum decreasing from 7.28 to**
532 **6.21.** Indeed, because of the spatial variability of management practices, soil water capacity or nitrogen
533 availability, different yields could be observed for the same amount of biomass. In addition, events such as
534 droughts during the reproductive stage, with potentially drastic yield reduction but negligible effects on
535 vegetative biomass, are certainly poorly detected by a model based only on vegetation indices. Thus the
536 direct relation NDVI/yield mostly allows assessing potential harvestable yields when assuming non-
537 limiting conditions (i.e. when yield is proportional to aboveground biomass). These potential yields could

538 however be reduced by crop water stress during the reproductive stages as shown in this study.
539 Consequently the direct relation NDVI/yield should be considered valid only for specific areas or years
540 without major limiting factors affecting yield.

541 **5.4. Limitations of the method**

542
543 The remote sensing-based model was applied directly to four surrounding departments and the
544 correlation coefficients were globally good despite an overall tendency to yield overestimation by the
545 model. The four departments are situated at the North of Kollo. They are mainly dominated by
546 agropastoral activities, with a mixture of livestock and crop cultivation [82]. Therefore, the probability to
547 have a mixture of crop vegetation and grasslands within a MODIS cropped pixel is high, which may
548 explain a lower performance of the model. Moreover, in these mixed zones of pasture, the seeding rates
549 are also very low leading to a sparse vegetation cover that causes high NDVI variations due to soil effects.
550 This highlights the main limitation of such models, based on empirical relationships between remote
551 sensing indices and yields: they depend on the environmental characteristics of the study area, which
552 restricts their application elsewhere without recalibration. In addition, such models also depend on the
553 farming system considered. For this reason, the model we developed in this study is only valid for a
554 system based on a single crop and should be tested or adapted for other farming systems such as in the
555 cereal-root crop mixed system where a wide range of different cereals is grown (maize, millet, sorghum or
556 cassava among other) including cases of intercropping.

557 Another consideration to take into account concerning our methodology is the need of a crop
558 mask to isolate cropped pixels. Since a pearl millet crop type map is not available for the NSD site, a crop
559 mask from the MODIS LCP was used here. The same approach was also applied to NDVI and CWSI
560 values extracted from the Landsat Crop mask. A coefficient of correlation of **0.80** is obtained when the
561 resulting estimated yields are compared to official statistics (not shown) which is close to the one obtained
562 with the MODIS LCP product. This confirms the relevance of the approach for the NSD site. However,
563 while the MODIS LCP has been validated for our study area, [83] recently spatialized the uncertainties in

564 the localization of cropland in the MODIS LCP over West Africa and showed a high spatial variability
565 with user accuracy varying between 17% to 70% according to farming systems. Thus to extrapolate our
566 methods in other locations, further efforts are needed to develop at least a map locating cultivated zones
567 and if possible the main crop type at a regional scale.

568 The use of a crop model instead of ground measurements to calibrate the remote-sensing model
569 can also be questioned. SARRA-H as most crop models tends to overestimate yields (Fig. 5) since it
570 simulates attainable yields according to agro-meteorological constraints but does not integrate all biotic
571 (e.g. birds, pests, and diseases due to excess moisture) or other non-environmental factors that influence
572 crop management which can lead to yield variations [14], [46], [84] . Remote sensing indices do integrate
573 biotic and non-environmental factors, and because they are calibrated using crop model outputs, an
574 overestimation of yields by the remote sensing-based model could be expected. In addition, since the
575 simulated yields from SARRA-H are overestimated, does that mean that the aboveground biomass and the
576 harvest index are also overestimated? For the latter, the simulated HI as well as the estimated HI are
577 within the range of those measured by [73] over 168 pearl millet plots in the Niamey area. Authors found
578 a mean HI of 0.22, we found a mean simulated and estimated HI of 0.29. For aboveground biomass,
579 reliable measurements in on-farm situations are not available. However, under controlled conditions it has
580 been shown in [46] in Senegal for pearl millet and in [62] for Sorghum in Mali that the aboveground
581 biomass (both yields and growth dynamics) were well simulated by SARRA-H. The same conclusion can
582 be drawn from the study of [4] based on on-farms survey near Niamey (Niger) for pearl millet.
583 Nonetheless, beyond the yields overestimation, our study show that the year-to-year variability is quiet
584 well simulated by the remote sensing based model.

585 Remote sensing indices also present intrinsic limitations. Despite the fact that a filter was applied
586 to reduce noise in the NDVI and LST time series, the presence of clouds, aerosols or dust residues may
587 lead to noise and the downgrading of data quality [85]. Thus, the poor performance of the
588 NDVI/aboveground biomass relation at local scale may also be explained by the 250m x 250m pixel size

589 of MODIS images that integrates a mixture of elements (crops, natural vegetation, bare soils) particularly
590 in the semi-arid region with low and sparse vegetation and where crop fields are often smaller than the
591 pixel size.

592 Finally, our study is limited to a period of eleven years and to 28 sites due to the unavailability of
593 more climatic data from ground observations to run the crop model. Agro-meteorological variables
594 derived from satellite could also be considered as an alternative. However, the correct estimation of these
595 variables from satellite, especially rainfall, remains an open issue. For instance, [86] found in the same
596 area that the TRMM 3B42 product, which delivers rainfall estimates at a daily time step, was not able to
597 accurately detect rainfall temporal pattern at the station level, and particularly the intra-seasonal rainfall
598 distribution. We hope that in a few years, the statistical relationships between aboveground biomass and
599 NDVI, and between HI and CWSI, can be updated and made more robust when more climatic data are
600 available.

601 **6. Conclusion and perspectives**

602

603 The difficulty to access ground measurements in West Africa and to estimate yields over large
604 areas using other monitoring methods such as agrometeorological modelling makes remote sensing
605 observation a good alternative or addition to consider for early warning systems. In this study, we
606 investigated a new approach based on the combination of vegetation and thermal indices for rainfed cereal
607 yield assessment in the Sahelian region. Empirical statistical models were developed between remote-
608 sensing indices (MODIS NDVI and LST), and SARRA-H simulated aboveground biomass and harvest
609 index respectively, and combined for the assessment of crop yield. We demonstrated that the combined
610 model performed better than the one using vegetation index alone. The inclusion of LST improves yield
611 estimations by accounting for the harvest index which is an indicator of the proportion of total
612 aboveground biomass really transformed into grains. In addition, it allows using NDVI as an estimator of
613 aboveground biomass, which is its primary function, rather than an indirect estimator of yield.

614 Furthermore, by using a crop model validated over the study area, this study showed that the combination
615 of satellite data with crop modelling is a good option for yield estimation and its year-to-year variability
616 based on remote sensing, especially for areas where ground measurements, required for the calibration of
617 the remote sensing-based model, are not available.

618 Our study confirms that even in small-holder agriculture such as those of the Sahelian region, the
619 use of coarse resolution satellite information for yield monitoring is possible. As the model proposed is
620 simple, robust and based on empirical relations with vegetation and thermal MODIS indices, there is
621 scope for operational implementation of yield estimation at regional scale in a food security early warning
622 system, in particular for the assessment of the year-to-year yield variability in regions with agronomic and
623 climatologic characteristics close to those of the NSD site. In addition, such a system could provide an
624 early estimation of yield shortly after harvest for an area equivalent to an administrative unit unlike
625 agricultural statistics that are currently available from three to six months after harvest. But that would
626 require addressing the issue of multi-crop type systems on which, to the best of our knowledge, no studies
627 have been conducted in the context of the West African farming systems. That would also require the use
628 of a different model for each broad climatic region and each crop type, and their necessary calibration
629 with appropriate ground measurements or crop model simulations. These in turn point out to the need for a
630 better identification of the crop domain and crop types. For instance, upcoming new sensors such as
631 Sentinel-2 (planned launch in June 2015) are expected to significantly improve yield monitoring by
632 providing high spectral, spatial and temporal data, which will allow more regular information on
633 agricultural land use practices. Consequently, a high quality crop type map as well as a stratification map
634 of West Africa according to crop types will become possible and thus the derivation of a remote sensing
635 model calibrated for each crop type. New optical sensors like Sentinel-2 will probably not resolve the
636 problem of data quality loss due to atmospheric effects. Future research must develop improved methods
637 based on the combination of optical and radar data (e.g. Sentinel 2 and 1) to allow vegetation monitoring
638 under all atmospheric conditions.

639 **Acknowledgments**

640 Leroux L. is supported by the CIRAD and by the Centre National d'Etudes Spatiales (Project CNES-
641 TOSCA "Dynafrique"). The authors thank the MODIS team for sharing the MODIS Vegetation Indices,
642 Land Surface Temperature and Land Cover products. Rainfall data were available thanks to the AMMA
643 program and especially Guillaume Quantin and Théo Vischel (Laboratoire d'Etude de Transferts en
644 Hydrologie et Environnement, Grenoble, France). Agricultural statistics data were kindly provided by the
645 Niger Agricultural Services. We thank Beatriz Belón for a helpful review of the manuscript. Finally, we
646 would like to thank anonymous reviewers for their valuable comments.

647 **References**

- 648 [1] R. Marteau, B. Sultan, V. Moron, A. Alhassane, C. Baron, and S. B. Traoré, "The onset of the
649 rainy season and farmers' sowing strategy for pearl millet cultivation in Southwest Niger," *Agric.
650 For. Meteorol.*, vol. 151, no. 10, pp. 1356–1369, Oct. 2011.
- 651 [2] FAO, "2014-2016 strategic response plan : Sahel region," Rome, Italy, 2014.
- 652 [3] P. B. I. Akponikpè, J. Minet, B. Gérard, P. Defourny, and C. L. Biielders, "Spatial fields '
653 dispersion as a farmer strategy to reduce agro-climatic risk at the household level in pearl millet-
654 based systems in the Sahel : A modeling perspective," *Agric. For. Meteorol.*, vol. 151, pp. 215–
655 227, 2011.
- 656 [4] S. B. Traoré, A. Alhassane, B. Muller, M. Kouressy, L. Somé, B. Sultan, P. Oettli, A. C. Siéné
657 Laopé, S. Sangaré, M. Vaksman, M. Diop, M. Dingkuhn, and C. Baron, "Characterizing and
658 modeling the diversity of cropping situations under climatic constraints in West Africa," *Atmos.
659 Sci. Lett.*, vol. 12, no. 1, pp. 89–95, Jan. 2011.
- 660 [5] C. Justice and I. Becker-Reshef, "Developing a strategy for global agricultural monitoring in the
661 framework of Group on Earth Observations (GEO) Workshop Report," Rome, 2007.
- 662 [6] M. Meroni, D. Fasbender, F. Kayitakire, G. Pini, F. Rembold, F. Urbano, and M. Verstrate, "Early
663 detection of biomass production deficit hot-spots in semi-arid environment using FAPAR time
664 series and a probabilistic approach," *Remote Sens. Environ.*, vol. 142, pp. 57–68, Feb. 2014.
- 665 [7] M. A. White, P. E. Thornton, and S. W. Running, "A continental responses phenology model
666 climatic for monitoring variability vegetation to interannual," vol. 11, no. 2, pp. 217–234, 1997.
- 667 [8] J. Huang and D. Han, "Meta-analysis of influential factors on crop yield estimation by remote
668 sensing," *Int. J. Remote Sens.*, vol. 35, no. 6, pp. 2267–2295, 2014.

- 669 [9] N. T. Son, C. F. Chen, C. R. Chen, L. Y. Chang, H. N. Duc, and L. D. Nguyen, "Prediction of rice
670 crop yield using MODIS EVI-LAI data in the Mekong Delta, Vietnam," *Int. J. Remote Sens.*, vol.
671 34, no. 20, pp. 7275–7292, Oct. 2013.
- 672 [10] C. J. Tucker, C. Vanpraet, E. Boerwinkel, and A. Gaston, "Satellite remote-sensing of total dry-
673 matter production in the Senegalese Sahel," *Remote Sens. Environ.*, vol. 13, no. 6, pp. 461–474,
674 1983.
- 675 [11] C. J. Tucker, "Satellite Remote Sensing of Total Herbaceous Biomass Production in the Senegalese
676 Sahel : 1980-1984," *Remote Sens. Environ.*, vol. 17, pp. 233–249, 1985.
- 677 [12] S. D. Prince, M. J. Eden, and J. T. Parry, "Monitoring the vegetation of semi-arid tropical
678 rangelands with the NOAA-7 Advanced Very High Resolution Radiometer," in *Remote sensing
679 and tropical land management*, John Wiley & Sons Ltd., 1986, pp. 307–334.
- 680 [13] I. Becker-Reshef, E. F. Vermote, M. Lindeman, and C. Justice, "A generalized regression-based
681 model for forecasting winter wheat yields in Kansas and Ukraine using MODIS data," *Remote
682 Sens. Environ.*, vol. 114, no. 6, pp. 1312–1323, Jun. 2010.
- 683 [14] A. K. Prasad, L. Chai, R. P. Singh, and M. Kafatos, "Crop yield estimation model for Iowa using
684 remote sensing and surface parameters," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 8, pp. 26–33, 2006.
- 685 [15] M. S. Rasmussen, "Assessment of millet yields and production in northern Burkina Faso using
686 integrated NDVI from the AVHRR.," *Int. J. Remote Sens.*, vol. 13, no. 18, pp. 3431–3442, 1992.
- 687 [16] M. S. Rasmussen, "Operational yield forecast using AVHRR NDVI data: Reduction of
688 environmental and inter-annual variability," *Int. J. Remote Sens.*, vol. 18, no. 5, pp. 1059–1077,
689 1997.
- 690 [17] A. M. Sibley, P. Grassini, N. E. Thomas, K. G. Cassman, and D. B. Lobell, "Testing Remote
691 Sensing Approaches for Assessing Yield Variability among Maize Fields," *Agron. J.*, vol. 106, no.
692 1, p. 24, 2014.
- 693 [18] C. Yang, J. H. Everitt, and J. M. Bradford, "Evaluating high resolution SPOT 5 satellite imagery to
694 estimate crop yield," *Precis. Agric.*, vol. 10, no. 4, pp. 292–303, Apr. 2009.
- 695 [19] F. Rembold, C. Atzberger, I. Savin, and O. Rojas, "Using Low Resolution Satellite Imagery for
696 Yield Prediction and Yield Anomaly Detection," *Remote Sens.*, vol. 5, no. 4, pp. 1704–1733, Apr.
697 2013.
- 698 [20] D. B. Lobell, "The use of satellite data for crop yield gap analysis," *F. Crop. Res.*, vol. 143, pp.
699 56–64, Mar. 2013.
- 700 [21] M. S. Moran, Y. Inoue, and E. M. Barnes, "Opportunities and Limitations for Image-Based
701 Remote Sensing in Precision Crop Management," *Remote Sens. Environ.*, vol. 61, pp. 319–346,
702 1997.
- 703 [22] L. Wall, D. Larocque, and P. Léger, "The early explanatory power of NDVI in crop yield
704 modelling," *Int. J. Remote Sens.*, vol. 29, no. 8, pp. 2211–2225, 2008.

- 705 [23] C. J. Tucker, "Red and Photographic Infrared linear Combinations for Monitoring Vegetation,"
706 *Remote Sens. Environ.*, vol. 8, pp. 127–150, 1979.
- 707 [24] B. N. Holben, C. J. Tucker, and C. J. Fan, "Spectral assessment of soybean leaf area and leaf
708 biomass," *Photogramm. Eng. Remote Sensing*, vol. 46, pp. 651–656, 1980.
- 709 [25] J. L. Hatfield, E. T. Kanemasu, G. Asrar, R. D. Jackson, P. J. Pinter, R. J. Reginato, and S. B. Idso,
710 "Leaf-area estimates from spectral measurements over various planting dates of wheat," *Int. J.*
711 *Remote Sens.*, vol. 46, pp. 651–656, 1984.
- 712 [26] F. Maselli, S. Romanelli, L. Bottai, and G. Maracchi, "Processing of GAC NDVI data for yield
713 forecasting in the Sahelian region," *Int. J. Remote Sens.*, vol. 21, no. 18, pp. 3509–3523, Jan. 2000.
- 714 [27] F. Maselli and F. Rembold, "Analysis of GAC NDVI Data for Cropland Identification and Yield
715 Forecasting in Mediterranean African Countries," *Photogramm. Eng. Remote Sensing*, vol. 67, no.
716 5, pp. 593–602, 2001.
- 717 [28] S. M. E. Groten, "NDVI—crop monitoring and early yield assessment of Burkina Faso," *Int. J.*
718 *Remote Sens.*, vol. 14, no. 8, pp. 1495–1515, 1993.
- 719 [29] J. L. Monteith and C. J. Moss, "Climate and the efficiency of crop production in Britain," *Philos.*
720 *Trans. R. Soc. Lond. B. Biol. Sci.*, vol. 281, no. 980, pp. 277–294, 1977.
- 721 [30] J. L. Monteith, "Solar radiation and productivity in tropical ecosystems," *J. Appl. Ecol.*, vol. 9, no.
722 3, pp. 747–766, 1972.
- 723 [31] R. B. Myneni and D. L. Williams, "On the relationship between FAPAR and NDVI," *Remote Sens.*
724 *Environ.*, vol. 49, no. 3, pp. 200–211, Sep. 1994.
- 725 [32] J. Lecoœur and T. R. Sinclair, "Harvest index increase during seed growth of field pea," *Eur. J.*
726 *Agron.*, vol. 14, no. 3, pp. 173–180, May 2001.
- 727 [33] R. L. DeLougherty and R. K. Crookston, "Harvest Index of Corn Affected by Population Density,
728 Maturity Rating, and Environment," *Agron. J.*, vol. 71, no. 4, pp. 577–580, 1979.
- 729 [34] M. Unkovich, J. Baldock, and M. Forbes, "Chapter 5 – Variability in Harvest Index of Grain Crops
730 and Potential Significance for Carbon Accounting: Examples from Australian Agriculture," in
731 *Advances in Agronomy*, vol. 105, 2010, pp. 173–219.
- 732 [35] A. N. Misra, "Assimilate partitioning in pearl millet (*Pennisetum glaucum* L.R.Br.)," *Acta Physiol.*
733 *Plant.*, vol. 17, no. 1, pp. 41–46, 1995.
- 734 [36] S. B. Idso, R. D. Jackson, and R. J. Reginato, "Remote-sensing of crop yields.," *Science (80-.)*,
735 vol. 196, no. 4285, pp. 19–25, Apr. 1977.
- 736 [37] R. C. . Smith, H. . Barrs, J. . Steiner, and M. Stapper, "Relationship between wheat yield and
737 foliage temperature: theory and its application to infrared measurements," *Agric. For. Meteorol.*,
738 vol. 36, no. 2, pp. 129–143, Dec. 1985.

- 739 [38] R. D. Jackson, S. B. Idso, R. J. Reginato, and P. J. Pinter, "Canopy temperature as a crop water
740 stress indicator," *Water Resour. Res.*, vol. 17, no. 4, pp. 1133–1138, 1981.
- 741 [39] P. J. Pinter, K. E. Fry, G. Guinn, and J. R. Mauney, "Infrared thermometry: A remote sensing
742 technique for predicting yield in water-stressed cotton," *Agric. Water Manag.*, vol. 6, no. 4, pp.
743 385–395, Aug. 1983.
- 744 [40] D. F. Wanjura, J. L. Hatfield, and D. R. Upchurch, "Crop water stress index relationships with crop
745 productivity," *Irrig. Sci.*, vol. 11, no. 2, pp. 93–99, Apr. 1990.
- 746 [41] A. E. Ajayi and A. A. Olufayo, "Evaluation of Two Temperature Stress Indices to Estimate Grain
747 Sorghum Yield and Evapotranspiration," *Agron. J.*, vol. 96, no. 5, pp. 1282–1287, 2004.
- 748 [42] N. K. Gontia and K. N. Tiwari, "Development of crop water stress index of wheat crop for
749 scheduling irrigation using infrared thermometry," *Agric. Water Manag.*, vol. 95, no. 10, pp. 1144–
750 1152, Oct. 2008.
- 751 [43] M. Dingkuhn, C. Baron, V. Bonnal, F. Maraux, B. Sarr, A. Clopes, and F. Forest, "Decision
752 support tools for rainfed crops in the Sahel at the plot and regional scales," in *Decision support
753 tools for smallholder agriculture in Sub-Saharan Africa : A practical guide*, Muscle Sho., B. T.
754 Struif and M. Wopereis, Eds. 2003, pp. 127–139.
- 755 [44] L. Hatfield, "Remote Sensing Estimators of Potential and Actual Crop Yield," *Remote Sens.
756 Environ.*, vol. 13, pp. 301–311, 1983.
- 757 [45] T. R. Sinclair, C. B. Tanner, and J. M. Bennett, "Water-Use Efficiency Crop Production,"
758 *Bioscience*, vol. 34, no. 1, pp. 36–40, 1984.
- 759 [46] B. Sultan, C. Baron, M. Dingkuhn, B. Sarr, and S. Janicot, "Agricultural impacts of large-scale
760 variability of the West African monsoon," *Agric. For. Meteorol.*, vol. 128, no. 1–2, pp. 93–110,
761 Jan. 2005.
- 762 [47] L. Le Barbé and T. Lebel, "Rainfall climatology of the HAPEX-Sahel region during the years
763 1950-1990," *J. Hydrol.*, vol. 188–189, pp. 43–73, 1997.
- 764 [48] P. Hiernaux, A. Ayantunde, A. Kalilou, E. Mougin, B. Gérard, F. Baup, M. Grippa, and B. Djaby,
765 "Trends in productivity of crops , fallow and rangelands in Southwest Niger : Impact of land use ,
766 management and variable rainfall," *J. Hydrol.*, vol. 375, no. 1–2, pp. 65–77, 2009.
- 767 [49] J. Rockström and A. de Rouw, "Water , nutrients and slope position in on-farm pearl millet
768 cultivation in the Sahel," *Plant Soil*, vol. 195, pp. 311–327, 1997.
- 769 [50] C. Baron, B. Sultan, M. Balme, B. Sarr, S. B. Traoré, T. Lebel, S. Janicot, and M. Dingkuhn,
770 "From GCM grid cell to agricultural plot: scale issues affecting modelling of climate impact.,"
771 *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, vol. 360, no. 1463, pp. 2095–2108, 2005.
- 772 [51] C. Justice, E. F. Vermote, J. Townshend, R. S. DeFries, D. P. Roy, D. K. Hall, V. V. Salomonson,
773 J. L. Privette, G. Riggs, A. H. Strahler, W. Lucht, R. Myneni, Y. Knyazikhin, S. W. Running, R. R.
774 Nemani, A. Huete, W. van Leeuwen, R. E. Wolfe, L. Giglio, J. Muller, P. Lewis, and M. J.
775 Barnsley, "The Moderate Resolution Imaging Spectroradiometer (MODIS): land remote sensing

- 776 for global change research,” *IEEE Trans. Geosci. Remote Sens.*, vol. 36, no. 4, pp. 1228–1249, Jul.
777 1998.
- 778 [52] A. Huete, K. Didan, T. Miura, E. . Rodriguez, X. Gao, and L. . Ferreira, “Overview of the
779 radiometric and biophysical performance of the MODIS vegetation indices,” *Remote Sens.*
780 *Environ.*, vol. 83, no. 1–2, pp. 195–213, Nov. 2002.
- 781 [53] J. Chen, P. Jönsson, M. Tamura, Z. Gu, B. Matsushita, and L. Eklundh, “A simple method for
782 reconstructing a high-quality NDVI time-series data set based on the Savitzky–Golay filter,”
783 *Remote Sens. Environ.*, vol. 91, no. 3–4, pp. 332–344, Jun. 2004.
- 784 [54] Z. Wan, “Modis Land Surface Temperature Products Users’ Guide,” 2013. [Online]. Available:
785 <http://www.icesb.ucsb.edu/modis/LstUsrGuide/usrguide.html>.
- 786 [55] Z. Wan, Y. Zhang, Q. Zhang, and Z.-L. Li, “Quality assessment and validation of the MODIS
787 global land surface temperature,” *Int. J. Remote Sens.*, vol. 25, no. 1, pp. 261–274, Jan. 2004.
- 788 [56] M. A. Friedl, D. Sulla-Menashe, B. Tan, A. Schneider, N. Ramankutty, A. Sibley, and X. Huang,
789 “MODIS Collection 5 global land cover : Algorithm refinements and characterization of new
790 datasets,” *Remote Sens. Environ.*, vol. 114, no. 1, pp. 168–182, 2010.
- 791 [57] M. A. Friedl, D. K. McIver, J. C. F. Hodges, X. Y. Zhang, D. Muchoney, A. H. Strahler, C. E.
792 Woodcock, S. Gopal, A. Schneider, A. Cooper, A. Baccini, F. Gao, and C. B. Schaaf, “Global land
793 cover mapping from MODIS: Algorithms and early results,” *Remote Sens. Environ.*, vol. 83, no. 1–
794 2, pp. 287–302, Nov. 2002.
- 795 [58] B. Sultan, P. Roudier, P. Quirion, A. Alhassane, B. Muller, M. Dingkuhn, P. Ciais, M.
796 Guimberteau, S. B. Traoré, and C. Baron, “Assessing climate change impacts on sorghum and
797 millet yields in the Sudanian and Sahelian savannas of West Africa,” *Environ. Res. Lett.*, vol. 8, no.
798 1, p. 9, Mar. 2013.
- 799 [59] B. Sultan, K. Guan, M. Kouressy, M. Biasutti, C. Piani, G. L. Hammer, G. Mclean, and D. Lobell,
800 “Robust features of future climate change impacts on sorghum yields in West Africa,” *Environ.*
801 *Res. Lett.*, vol. 9, p. 13, 2014.
- 802 [60] J. Ramarohetra, B. Sultan, C. Baron, T. Gaiser, and M. Gosset, “How satellite rainfall estimate
803 errors may impact rainfed cereal yield simulation in West Africa,” *Agric. For. Meteorol.*, vol. 180,
804 pp. 118–131, Oct. 2013.
- 805 [61] S. Bassu, N. Brisson, J.-L. Durand, K. Boote, J. Lizaso, J. W. Jones, C. Rosenzweig, A. C. Ruane,
806 M. Adam, C. Baron, B. Basso, C. Biernath, H. Boogaard, S. Conijn, M. Corbeels, D. Deryng, G.
807 De Sanctis, S. Gayler, P. Grassini, J. Hatfield, S. Hoek, C. Izaurralde, R. Jongschaap, A. R.
808 Kemanian, K. C. Kersebaum, S. H. Kim, N. S. Kumar, D. Makowski, C. Müller, C. Nendel, E.
809 Priesack, M. V. Pravia, F. Sau, I. Shcherbak, F. Tao, E. Teixeira, D. Timlin, and K. Waha, “How
810 do various maize crop models vary in their responses to climate change factors?,” *Glob. Chang.*
811 *Biol.*, vol. 20, no. 7, pp. 2301–2320, 2014.
- 812 [62] M. Kouressy, M. Dingkuhn, M. Vaksman, and A. B. Heinemann, “Adaptation to diverse semi-
813 arid environments of sorghum genotypes having different plant type and sensitivity to
814 photoperiod,” *Agric. For. Meteorol.*, vol. 148, no. 3, pp. 357–371, Mar. 2008.

- 815 [63] G. Bezançon, J.-L. Pham, M. Deu, Y. Vigouroux, F. Sagnard, C. Mariac, I. Kapran, A. Mamadou,
816 B. Gérard, J. Ndjeunga, and J. Chantreau, "Changes in the diversity and geographic distribution of
817 cultivated millet (*Pennisetum glaucum* (L.) R. Br.) and sorghum (*Sorghum bicolor* (L.) Moench)
818 varieties in Niger between 1976 and 2003," *Genet. Resour. Crop Evol.*, vol. 56, no. 2, pp. 223–236,
819 Jul. 2008.
- 820 [64] P. Roudier, B. Sultan, P. Quirion, C. Baron, A. Alhassane, S. B. Traoré, and B. Muller, "An ex-
821 ante evaluation of the use of seasonal climate forecasts for millet growers in SW Niger," *Int. J.*
822 *Climatol.*, vol. 32, no. March 2011, pp. 759–771, 2012.
- 823 [65] M. Vaksman and S. B. Traoré, "Adéquation entre risque climatique et choix variétal du mil : Cas
824 de la zone de Bankass au Mali," in *Bilan hydrique agricole et sécheresse en Afrique Tropicale*,
825 John Wiley., Paris, France, 1991, pp. 113–123.
- 826 [66] FAO/IIASA/ISRIC/ISSCAS/JRC, "Harmonized World Soil Database (version 1.2)." FAO, Rome,
827 Italy and IIASA, Laxenburg, Austria, 2012.
- 828 [67] M. V. . Sivakumar, "Predicting rainy season potential from the onset of rains in Southern Sahelian
829 and Sudanian climatic zones of West Africa," *Agric. For. Meteorol.*, vol. 42, pp. 295–305, 1988.
- 830 [68] FAO, "Crop calendar-An information tool for seed security," 2010. [Online]. Available:
831 <http://www.fao.org/agriculture/seed/cropcalendar/cropcal.do>.
- 832 [69] M. S. Moran, T. R. Clarke, Y. Inoue, and A. Vidal, "Estimating Crop Water Deficit Using the
833 Relation between Surface-Air Temperature and Spectral Vegetation Index," *Remote Sens.*
834 *Environ.*, vol. 49, pp. 246–263, 1994.
- 835 [70] R. Delécolle, S. J. Maas, M. Guérif, and F. Baret, "Remote sensing and crop production models:
836 present trends," *ISPRS J. Photogramm. Remote Sens.*, vol. 47, no. 2–3, pp. 145–161, Apr. 1992.
- 837 [71] CIRAD, "Farmers yield variability assessment and validation of crop model to predict 'average
838 regional' farmers yield for the main cropped varieties of millet, sorghum and maize," Montpellier,
839 France, 2009.
- 840 [72] A. J. Challinor, T. R. Wheeler, P. Q. Craufurd, J. M. Slingo, and D. I. F. Grimes, "Design and
841 optimisation of a large-area process-based model for annual crops," *Agric. For. Meteorol.*, vol.
842 124, no. 1–2, pp. 99–120, Jul. 2004.
- 843 [73] A. Bégué, J. F. Desprat, J. Imbernon, and F. Baret, "Radiation use efficiency of pearl millet in the
844 Sahelian zone," *Agric. For. Meteorol.*, vol. 56, pp. 93–110, 1991.
- 845 [74] A. Huete and C. J. Tucker, "Investigation of soil influences in AVHRR red and near- infrared
846 vegetation index imagery," *Int. J. Remote Sens.*, vol. 12, no. 6, pp. 1223–1242, 1991.
- 847 [75] A. Diouf and E. F. Lambin, "Monitoring land-cover changes in semi-arid regions: remote sensing
848 data and field observations in the Ferlo, Senegal," *J. Arid Environ.*, vol. 48, no. 2, pp. 129–148,
849 Jun. 2001.
- 850 [76] C. J. Tucker, B. . Holben, J. . Elgin, and J. . McMurtrey, "Relationship of spectral data to grain
851 yield variation," *Photogramm. Eng. Remote Sensing*, vol. 46, no. 5, pp. 657–666, 1980.

- 852 [77] J. Huang, H. Wang, Q. Dai, and D. Han, “Analysis of NDVI Data for Crop Identification and Yield
853 Estimation,” *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 7, no. 11, pp. 4374 – 4384,
854 2014.
- 855 [78] E. Vintrou, A. Bégué, C. Baron, S. Alexandre, D. Lo Seen, and S. B. Traoré, “A Comparative
856 Study on Satellite and Model-Based Crop Phenology in West Africa,” *Remote Sens.*, vol. 6, pp.
857 1367–1389, 2014.
- 858 [79] J. . Jones, G. Hoogenboom, C. . Porter, K. . Boote, W. . Batchelor, L. . Hunt, P. . Wilkens, U.
859 Singh, A. . Gijsman, and J. . Ritchie, “The DSSAT cropping system model,” *Eur. J. Agron.*, vol.
860 18, no. 3–4, pp. 235–265, Jan. 2003.
- 861 [80] FAO, “A computer program for irrigation planning and management,” *FAO Irrigation and
862 Drainage Paper*, vol. 46, Rome, Italy, 1992.
- 863 [81] J. L. Hatfield and M. S. Moran, “Agriculture and Remote Sensing,” in *Encyclopedia of Remote
864 Sensing*, E. G. Njoku, Ed. New-York, USA: Springer, 2014, pp. 22–32.
- 865 [82] FEWS NET, “LIVELIHOODS ZONING ‘ PLUS ’ ACTIVITY IN NIGER,” 2011.
- 866 [83] L. Leroux, A. Jolivot, A. Bégué, D. Lo Seen, and B. Zoungrana, “How Reliable is the MODIS
867 Land Cover Product for Crop Mapping Sub-Saharan Agricultural Landscapes?,” *Remote Sens.*, vol.
868 6, pp. 8541–8564, 2014.
- 869 [84] D. Lobell and C. Field, “Global scale climate – crop yield relationships and the impacts of recent
870 warming,” *Environ. Res. Lett.*, vol. 2, p. 7, 2007.
- 871 [85] J. N. Hird and G. J. McDermid, “Noise reduction of NDVI time series : An empirical comparison
872 of selected techniques,” *Remote Sens. Environ.*, vol. 113, pp. 248–258, 2009.
- 873 [86] L. Leroux, C. Baron, S. B. Traoré, D. Lo Seen, and A. Bégué, “Testing satellite rainfall estimates
874 time series for crop yield simulation of a rainfed cereal in West Africa,” in *8th International
875 Workshop on the Analysis of Multitemporal Remote Sensing Images*, 2015.
- 876