The impact of conflict-driven cropland abandonment on food insecurity in South Sudan revealed using satellite remote sensing

Olsen, Victor Mackenhauer; Fensholt, Rasmus; Olofsson, Pontus; Bonifacio, Rogerio; Butsic, Van; Druce, Daniel; Ray, Deepak; Prishchepov, Alexander V.

Published in:
Nature Food

DOI:
10.1038/s43016-021-00417-3

Publication date:
2021

Document version
Early version, also known as pre-print

Citation for published version (APA):
Revealing Cropland Abandonment and Food Insecurity in War-ravaged South Sudan

Victor Mackenhauer Olsen¹*, Rasmus Fensholt¹, Rogerio Bonifacio², Van Butsic³, Deepak Ray⁴, Alexander V. Prishchepov¹

¹ Department of Geosciences and Natural Resource Management (IGN), University of Copenhagen, Øster Voldgade 10, DK-1350 København K, Denmark
² Vulnerability Analysis and Mapping (VAM), United Nations World Food Programme (WFP), Via Cesare Giulio Viola, 68, 00148 Roma RM, Italy
³ Department of Environmental Science, Policy, & Management, Berkeley University of California, 130 Mulford Hall 3114, CA 94720, Berkeley, United States of America
⁴ Institute on the Environment (IonE), University of Minnesota, 1954 Buford Ave St. Paul, MN 55108, Minnesota, United States of America

* corresponding author:
Victor Mackenhauer Olsen
victor.olsen@hotmail.com
+45 60 12 57 82
Abstract

Armed conflicts often result in cropland abandonment with significant impacts on food security. Moreover, conflicts restrict the collection of on-the-ground information required for organizing targeted relief distribution. Satellite remote sensing provides a way of observationally gathering information about disruptions during armed conflicts and food-security status in conflict zones. Using ~7500 multisource satellite images, we implemented a data-driven approach that showed a reduction in cultivated croplands in war-ravaged South Sudan by 23% from 2016 to 2018. Propensity score matching revealed a close relationship between cropland abandonment and armed conflicts that lead to drastic decreases in the available food supply. If war-induced abandonment had not occurred, our analysis shows that the abandoned croplands could have supported at least a quarter of the population in the southern states of South Sudan. Here, we demonstrate that remote-sensing technology can play a crucial role in rapid assessments of cropland abandonment in food-insecure regions, thus improving the basis for timely aid provision.
Introduction

Food insecurity remains a prevalent problem around the world. Approximately 821 million people were facing food insecurity in 2017, of whom 124 million had an acute need of food. Food insecurity signifies a lack of access to sufficient nutrients essential to support life and ranges from nutritional deficiency to famine. With economic and technological developments, the relative number of people facing food insecurity has declined significantly over the 20th century. However, the absolute number of food-insecure people has remained high over the past decades and has even increased over the last three years, in part due to climate change and violent conflict. The importance of addressing food insecurity is articulated in the United Nations’s second sustainable development goal, which aims at ending hunger by 2030.

However, to reverse the current trend and fulfill the purpose of eradicating hunger, more timely and accurate information on food production, food availability, and land use is needed. Food insecurity is closely linked with armed conflict and land abandonment. Armed conflicts are among the major drivers of cropland abandonment, which may have multiple implications for the environment and may serve as a proximate cause of food insecurity. In 2019, all 19 countries suffering from protracted food crises were also affected by armed conflicts.

Understanding the impacts of armed conflict on cropland abandonment is, therefore, important for addressing the root causes of food insecurity. To achieve this, and to provide timely and accurate information on food production, better means of agricultural assessments are needed, as field surveys are often restricted due to safety concerns and inaccessibility. Earth observation techniques have potential as a means of extracting agricultural information and better understanding the drivers of food insecurity.
The challenge of food insecurity and adequate aid provision is particularly pressing in South Sudan. The country has suffered from several years of violent conflict, which has disrupted agricultural production, resulting in cropland abandonment and restricted access to markets. Ultimately, the conflicts in South Sudan resulted in nationwide food insecurity and periodic famine. In South Sudan, a large proportion of the population relies on local markets and smallholder farming to meet their nutritional needs, making them vulnerable to local fluctuations in food availability. The conflict restricted access for agricultural surveyors and resulted in limited information on agricultural production in the region, which is necessary to inform humanitarian response. Here, we used a fusion of open-access multisource high-resolution data from the ESA Copernicus and USGS satellite systems to gather information about the state of agricultural production in South Sudan in relation to conflicts.

We mapped cropland in three states of South Sudan: Western Equatoria, Central Equatoria, and Lakes, from 2016 to 2018. We used multisource satellite imagery, data fusion, and machine-learning supervised image classification techniques. Agriculture in South Sudan is characterized by smallholder farming and small agricultural plots. To detect cropland, we took advantage of the improved spatial resolution (as high as 10 meters for some bands) of freely accessible optical Sentinel-2 and radar Sentinel-1 images. Sentinel data were combined with 30-m Landsat-8 and 250-m Aqua/Terra MODIS products to fill cloud-induced gaps in monthly image composites. Finally, land cover was classified using the random forest classifier trained on an extensive in situ dataset collected by the UN World Food Programme (WFP) during several field campaigns and augmented with observations from very-high-resolution (VHR) commercial imagery available from PlanetScope™ and GoogleEarth™.
We statistically compared the extent of cropland in conflict and non-conflict areas to assess the conflict's impact during 2017 and 2018. To attribute the abandonment patterns to the ongoing conflicts, we conducted a comparative analysis of cropland abandonment in conflict and non-conflict areas and reduced confounder bias using propensity score matching \(^{13,14}\). Propensity score matching reduces bias in the comparative analysis by subsampling (or “matching”) the data set to only include samples comparable to the causal variable under investigation \(^{15}\). By matching samples on the likelihood of conflict, based on the distance to roads, distance to settlements, and county-level population density, we satisfied the conditional independence assumption and compared cropland abandonment in the two groups to assess the impact of conflict on cropland abandonment.

Finally, we demonstrate the impact of war-induced cropland abandonment on kilocalorie availability. Unattained cropland yields were measured using the mapped abandonment extent in 2017 and 2018 and the proportions and average yields of the major crops reported at the provincial and state level. To convert the unattained crop yields into kilocalories, we used extensive nutrient profile data on various crops from the Food Data Central of the U.S. Department of Agriculture (USDA) \(^{16}\).
Results

Cropland Dynamic in South Sudan for 2016-2018

Land-cover maps produced with combined multisource satellite data at 10-meter resolution (Figure 1) reached overall accuracies of between 0.81 and 0.90 (Table 2) and revealed significant cropland abandonment from 2016 to 2018 (Table 1). From 2016 to 2017, croplands declined by 10%, with a net loss of approximately 39,000 ha. From 2017 to 2018, cropland declined by another 14%, with a net loss of approximately 49,000 ha. In total, from 2016 to 2018, croplands declined by approximately 23%, from 390,000 hectares of active croplands in 2016 to 302,000 hectares in 2018. Cropland abandonment primarily occurred in Western Equatoria and Central Equatoria states, whereas the Lakes state underwent an overall cropland increase from 2016 to 2018.
Figure 1. Land-cover map for 2018 produced with supervised image classification and the examples of land-cover changes in three agricultural areas—Rumbek, Yambio, and Yei (A, B, and C, respectively).
Table 1. Non-error adjusted cropland extent (thousand ha).

<table>
<thead>
<tr>
<th>State</th>
<th>2016</th>
<th>2017</th>
<th>2018</th>
<th>Change 2016-2018</th>
</tr>
</thead>
<tbody>
<tr>
<td>Western Equatoria</td>
<td>92</td>
<td>90</td>
<td>68</td>
<td>-24</td>
</tr>
<tr>
<td>Central Equatoria</td>
<td>165</td>
<td>88</td>
<td>66</td>
<td>-99</td>
</tr>
<tr>
<td>Lakes</td>
<td>133</td>
<td>173</td>
<td>167</td>
<td>+34</td>
</tr>
<tr>
<td>Total</td>
<td>390</td>
<td>351</td>
<td>302</td>
<td>-88</td>
</tr>
</tbody>
</table>

Table 2. Accuracy assessment results.

<table>
<thead>
<tr>
<th></th>
<th>Savanna</th>
<th>Forest</th>
<th>Bare</th>
<th>Water</th>
<th>Cropland</th>
</tr>
</thead>
<tbody>
<tr>
<td>2016</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Producer's Accuracy</td>
<td>0.79</td>
<td>0.85</td>
<td>0.73</td>
<td>0.95</td>
<td>0.87</td>
</tr>
<tr>
<td>User's Accuracy</td>
<td>0.89</td>
<td>0.9</td>
<td>0.97</td>
<td>0.9</td>
<td>0.59</td>
</tr>
<tr>
<td>Overall Accuracy</td>
<td>0.81</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2017</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Producer's Accuracy</td>
<td>0.71</td>
<td>0.94</td>
<td>0.63</td>
<td>1</td>
<td>0.92</td>
</tr>
<tr>
<td>User's Accuracy</td>
<td>0.91</td>
<td>0.82</td>
<td>0.96</td>
<td>0.95</td>
<td>0.65</td>
</tr>
<tr>
<td>Overall Accuracy</td>
<td>0.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2018</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Producer's Accuracy</td>
<td>0.98</td>
<td>0.69</td>
<td>0.17</td>
<td>0.63</td>
<td>0.22</td>
</tr>
<tr>
<td>User's Accuracy</td>
<td>0.92</td>
<td>0.77</td>
<td>0.92</td>
<td>0.94</td>
<td>0.66</td>
</tr>
<tr>
<td>Overall Accuracy</td>
<td>0.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: These are non-adjusted classification accuracies. Details on accuracy assessment can be found in the Methods section.

Impact of Armed Conflict on Cropland Abandonment

The comparative analysis with propensity score matching revealed a significantly larger reduction of cropland extent between 2016 and 2018 in proximity to locations of violent conflicts (Table 3). Within a 5-km radius around conflict and non-conflict locations, the average annual
decline in cropland extent was -346 ha (-27.7%) for the conflict group (out of an average cropland extent of 1276 ha) and -79 ha for the non-conflict group (out of an average cropland extent of 450 ha). This yielded a difference in mean decline of 266 ha. Similarly, at a 10-km radius, the average annual decline in cropland extent was -488 ha (-15.7%) for the conflict group (out of an average cropland extent of 3,108 ha), and an increase of 46 ha (+3.49%) for the non-conflict (out of an average cropland extent of 1309 ha), yielding a difference in mean decline of -554 ha. Thereby, both the 5-km and 10-km buffer samples indicate a statistically significant impact of conflict on the cropland abandonment.

Table 3. The average change of cropland extent in the conflict and non-conflict groups between 2016 and 2018.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Conflict zones</th>
<th>Non-conflict zones</th>
<th>Diff. (ha)</th>
<th>T-Val.</th>
<th>DF</th>
<th>P-Val</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 km</td>
<td>Absolute: -345.5, +/- 68</td>
<td>Relative: +/- 68, -27.1</td>
<td>Absolute: -79, +/- 35.3, -17.7</td>
<td>-266.1</td>
<td>3.47</td>
<td>145.78</td>
</tr>
<tr>
<td>10 km</td>
<td>Absolute: -487.9, +/- 145</td>
<td>Relative: +/- 145, -15.7</td>
<td>Absolute: 45.6, +/- 98, 3.5</td>
<td>-533.6</td>
<td>3.05</td>
<td>126.35</td>
</tr>
</tbody>
</table>

Impact of cropland abandonment on kilocalorie availability

Cropland abandonment significantly reduced food production and the local supply needed to meet the population's required kilocalorie needs. If abandonment had not taken place, the assessment of the proportion of major crops and yields before the abandonment revealed that an additional 172,403 tonnes of various crops could have been produced in the 2016-2017 crop calendar year (Figure 2). These crops included 16,700 tonnes of sorghum, 7600 tonnes of maize, and 125,350 tonnes of cassava, representing the bulk of staple food in the studied region. Similarly, if abandonment had not occurred in the 2017-2018 crop calendar year, an additional
12,644 tonnes of sorghum, 5,040 tonnes of maize, and 73,808 tonnes of cassava could have been produced. Such production could have met the kilocalorie needs of approximately 54,150 households and 44,080 households in the 2016-2017 and 2017-2018 crop calendar years, respectively. In sum, if abandonment had not occurred, crop production could have satisfied an additional 382,800 people in 2016-2017 and 327,600 people in 2017-2018, representing a quarter of the population.

**Figure 2. Missed crop production due to cropland abandonment.**
Armed conflicts are trigger events that rapidly change land systems and often cause detrimental short- and long-term impacts on food security. Contemporary remote sensing systems provide unprecedented technology to evaluate the implications of armed conflicts on food production systems and can serve as an attractive complementary data source to conventional field-based surveying methods. Based on a multisource satellite remote sensing imagery approach, propensity score matching, and calculation of missed kilocalories, we unraveled how armed conflict had a detrimental impact on South Sudan's food security. The results demonstrated the crucial role of freely available satellite imagery, particularly the ESA Copernicus Sentinel-1 and Sentinel-2 missions, for estimating agricultural production. The imagery was analyzed using a multisource satellite fusion technique, implemented to limit clouds' impacts on optical satellite data, to ultimately map smallholder farming at 10-meter spatial resolution. The demonstrated utility of multisource image fusion techniques and the role of Sentinel-1 and Sentinel-2 data can be relevant beyond South Sudan to many other parts of the world, where smallholder farming is common, particularly in areas where armed conflicts occur.

The analysis showed a statistically significant impact of conflict on cropland extent, which was indicated by reduced areas under agriculture in proximity to armed conflict locations (both within 5 km and 10 km buffer zones around the conflict locations) compared to non-conflict areas, after accounting for the confounders. The observed cropland abandonment dramatically limited the potential for local food production. Given the reliance on subsistence smallholder farming and local markets, it resulted in unattained kilocalories and increased food insecurity. Additionally, some local markets ceased to operate and, therefore, abandonment resulted in a strong need for farming and non-farming groups to rely explicitly on food aid. As quantified in
our analysis, at least one-quarter of the population in Western Equatoria, Central Equatoria, and Lakes could have been supported by the agricultural production of abandoned cropland, thus highlighting the impact of conflict on food security.

Our findings on cropland abandonment and its implications for food security are worrying. By the end of 2018, approximately 74% of the South Sudan population was estimated to be suffering from food insecurity at the post-harvest time, and 26% of the population was severely food insecure. However, FAO/WFP also revealed that the number of food-insecure people decreased slightly compared to the beginning of 2018, albeit primarily due to substantial humanitarian assistance. Therefore, local food production and combatting the ongoing abandonment of croplands are critical to alleviating a chronic demand for food aid, closure of local markets, and the distortion of trade routes.

Here, we also established a link between war-induced cropland abandonment and food insecurity. With the expectation of adverse impacts of climate change in already food insecure regions often found in the Global South, more violent conflicts and temporary or permanent cropland abandonment may occur. Given already widespread cropland abandonment with complex implications for the environment and food security, we also recommend revisiting global abandonment patterns from a food-security perspective, particularly in already food-insecure regions of the world.
Methods

Study Area

This study focuses on three states in South Sudan, which are of high importance to national food security: Western Equatoria, Central Equatoria, and Lakes (Figure 1)\textsuperscript{11,21,29}. Geographically, the study area encompasses the so-called “greenbelt”, which is the most productive agricultural zone in South Sudan, stretching along the southern and western national border\textsuperscript{21}. The greenbelt was a traditional net exporter of food until the escalation of the conflict\textsuperscript{21}. The study covered two growing seasons, 2016–2017 and 2017–2018.

Cropland is predominantly managed by smallholder farming with an average cereal area per household ranging from 0.4 to 1.3 ha and with very limited mechanization and irrigation\textsuperscript{21}. Animal traction is rarely used and the fields are cleared by human power\textsuperscript{21}. The application of mineral fertilizers, pesticides, and herbicides is rare. Soil fertility is predominantly managed by shifting cultivation and dung from household animals or pastoralist herds\textsuperscript{21}. These farming practices provide low yields and the crops are prone to the infestation of pests and droughts.

Several crops are grown in the study area. Cereals comprise most cultivated areas, with sorghum and maize being the most important cereals, followed by bulrush millet and finger millet\textsuperscript{21}. Two types of sorghum are commonly cultivated, the short-season sorghum with a < 90-day growing period, and the long-season sorghum with a > 220-day growing period\textsuperscript{21}. Maize is the second most common cereal in the greenbelt and is often cultivated twice per season on the same plot. The most common non-cereal crops are groundnuts and cassava that together comprise up to half of all crops in the greenbelt. The growing season is 200-220 days in the study area which allows for two to three harvests annually\textsuperscript{21}. 
**Acquisition and Processing of Satellite Imagery**

Imagery from Sentinel-1, Sentinel-2, Landsat-8, and MODIS, as well as a Copernicus land-cover product, were combined and used for land-cover classification in Google Earth Engine (GEE) \(^{30}\). From the optical sensors, the Red, Green, Blue, NIR, and SWIR 1 and 2 bands were used (Supplementary Table 1 and 2). The optical datasets were processed and combined following five steps: 1) cloud masking, 2) mosaicking and compositing, 3) reflectance calibration, 4) data fusion, and 5) deriving additional metrics (see a flowchart in Supplementary Figure 1). The process was repeated for each classified year (2016, 2017, and 2018), using all images available from Sentinel-2, Landsat-8, and MODIS. In total, approximately 7,500 multiband images from Sentinel-2, Landsat-8, and MODIS were reduced to 13 multiband images per year (Supplementary Table 2).

Clouds and shadows were masked out from Sentinel-2 MSI and Landsat-8 OLI imagery with quality flag bands. Sentinel-2 MSI was further masked by applying thresholds on VNIR and SWIR bands (bands number 4, 5, 10, and 12) to detect both dense clouds and cirrus clouds \(^{31,32}\). The cloud-free Sentinel-2 MSI, Landsat-8 OLI, and MODIS time-series were mosaicked and composit ed independently. Before data fusion, the Landsat-8 and MODIS time-series were calibrated to match the reflectance values of Sentinel-2. For that, we calculated the ratio between overlapping pixels from Sentinel-2, Landsat-8, and MODIS for each monthly composite. The calculated ratio was then used as a multiplication factor for each pixel in Landsat-8 and MODIS composites to match with Sentinel-2 data sets. Additionally, Landsat and MODIS images were resampled to 10 m to match the resolution of Sentinel-2. In total, we fused the three optical time-series.
series into 13 monthly image composites—January through December, plus January of the following year. When multiple cloud-free pixels from different optical sensors were present for a location and month, the pixels with the highest NDVI were prioritized. The cloud-masked Sentinel-2 composites were used first, whereas Landsat and MODIS were used only to fill cloud-induced gaps in the Sentinel-2 time-series. Approximately 85% of the study area was covered by Sentinel-2 composite data for each period. The remaining cloud-masked area was first filled with the pre-processed Landsat-8 composites and, if any cloud-masked area remained, MODIS was used to fill final data gaps (see Supplementary Figure 2 and Table 3).

We also calculated and included NDVI, NDWI, and time-series seasonality metrics—minimum, maximum, amplitude, and standard deviation—because seasonality metrics can boost classification accuracies. Additionally, we included the 100-m resolution Copernicus Land Cover Map in the classification to help the random forest classifier classify crops in areas with a high probability of crops being present (Supplementary Table 2).

The Sentinel-1 VV and VH time-series were processed through two steps: 1) mosaicking/compositing, and 2) deriving additional metrics. The process was repeated for each classified year, using all images that had been acquired within the given year. The Sentinel-1 VV and VH annual time-series were mosaicked and composited into 12 monthly composites separately for VV and VH. For the SAR imagery, compositing was achieved by condensing the time-series into the average reflectance values in monthly intervals (differing from the optical time-series which was composited through gap-filling), which could help to reduce speckle and noise common in SAR datasets. The 12-month SAR time-series was then also used to calculate radar seasonality metrics (for the VV-polarization only), which were incorporated into the layer stack for classification. To increase abandoned and non-abandoned crop separability,
we used the Haralick technique and calculated entropy and inertia (‘grey-level co-occurrence matrix’) texture measures for both the VV and VH monthly composites with 5*5 pixels in GEE. In total, we had 136 inputs/bands for each annual layer stack, which was later classified with the random forest algorithm (Supplementary Table 2).

**Training Data for Land-cover Classification**

Reference data for training the random forest classifier were obtained from the United Nations WFP for 2017 and 2018. Over 2,000 fields were allocated with non-differential GPS during the field campaigns in 2017 and 2018, and the crop types were labeled by WFP experts (Supplementary Table 4). We used the centroids from digitized fields as training samples to mitigate potential error of imprecise digitizing of surveyed plots. The reference field data for croplands were not collected in 2016 by WFP experts. Therefore, cropland training data were instead digitized manually from multi-temporal VHR imagery available for 2016 via Google Earth™, Bing™ (e.g., 1.84-meter WorldView) and Planet Explorer™ (5-meter RapidEye and 3-meter PlanetScope). Training data for all non-cropland land-cover classes (savanna, forest, bare, grassland, water, and impervious surface combined) were assigned by visual expert interpretation for all years using VHR imagery from the sources listed above.

**Land-cover Classification**

To produce land-cover classifications for 2016, 2017, and 2018, we used the non-parametric machine-learning random forest classifier available in GEE. We utilized the SciKit feature
selection package in Python, including precursive feature elimination and cross-validation in order to run sensitivity tests for random forest parameters and feature selection based on training data. We used 10-fold cross-validation and 500 trees for the random forest. The feature selection results were used to identify the top 60 performing features for each year, thus reducing the layer stack's size and ensuring the utility of the best features (performance plateaued after approximately 60 features). The image classifications were conducted in GEE using the top-performing features. Again, the random forest classifier was parameterized with 500 trees and input features were equal to the square root of the number of features. Land-cover was initially classified into seven classes: cropland, savanna, forest, bare, grassland, water, and artificial surface (the description of land-cover classes can be found in Supplementary Table 5).

Accuracy Assessment

To validate the land-cover maps, we produced an independent validation dataset of 1000 points following the recommended accuracy assessment practices. A stratified random sample design was implemented to target smaller land-cover classes, thereby meeting the randomized sample requirement while maintaining an acceptable standard error for our classes. We digitized 35 strata of 15 x 15 km that targeted smaller land-cover classes (Supplementary Figure 3). We sampled validation points with a minimum distance of 1 km, which returned a Moran’s I at <0.40 based on the 2017 classified map. The validation sample was labeled following visual interpretation of high-resolution imagery available from Digital Globe™ and Planet™, accessed through Google Earth™, Bing™, and Planet Explorer™ online platforms. Each validation point was revisited and thematic land-cover classes were assigned separately for 2016, 2017, and 2018.
Error-adjusted area corrections based on the accuracy assessment matrix and initial area estimates would have required the sample to represent land-cover proportions of the entire study area. However, these criteria could not be met, because the crops covered only a small proportion of the study area. Therefore, we calculated both the error-adjusted and unadjusted area estimates (Supplementary Table 6, 7, 8).

Revealing the Impacts of Armed Conflict on Cultivated Areas with Propensity Score Matching

Three covariates were selected to account for the occurrence of violent conflict: distance to roads, distance to settlements, and county-level population density, which have been proven relevant in explaining the spatial determinants of land-cover change, including cropland abandonment and the occurrence of conflict. The distance variables were produced based on Open Street Maps road and settlement data, and the population density per county was calculated based on population statistics for 2017. The georeferenced conflict dataset was obtained from the Armed Conflict Location and Event Data Project (ACLED) on political violence around the globe, which is one of the most comprehensive and widely used datasets for quantitative conflict studies. Approximately 50% of the events were recorded explicitly for the settlements in which the events occurred (ACLED’s category 1), whereas another 50% of events were georeferenced and spatially linked to the nearest settlement (ACLED’s category 2). The dataset was limited to political violence, defined as violence that “occurs within civil wars and periods of instability, public protest and regime breakdown.” We retained only conflict occurring within the study area and between the annual agricultural seasons. Therefore,
we used conflict events from November 1, 2016 to March 31, 2017 and from November 1, 2017
to March 31, 2018 (Table 4). We only kept ‘battles’, ‘violence against civilians’, and
‘explosions/remote violence’, and excluded ‘protests’, ‘riots’, and ‘strategic development’
(Supplementary Table 9).

**Table 4. Conflict sample periods.**

<table>
<thead>
<tr>
<th>Year</th>
<th>2016</th>
<th>2017</th>
<th>2018</th>
</tr>
</thead>
<tbody>
<tr>
<td>Month Season</td>
<td>123456789101112</td>
<td>123456789101112</td>
<td>123456789101112</td>
</tr>
<tr>
<td>Season</td>
<td>Growing Season</td>
<td>2016/17 Sample</td>
<td>Growing Season</td>
</tr>
</tbody>
</table>

Conflict samples were generated as circles with a 5 or 10 km radius for each conflict point. Non-
conflict samples of the same size were randomly sampled in the remaining area (Supplementary
Figure 4). Sampling was performed separately in the two sample periods, 1) November 2016 to
March 2017 and 2) November 2017 to March 2018, but after extraction and linkage of the
cropland extent and covariate statistics for the respective year, the samples from the two time
periods were combined into one sample. The minimum sample size was chosen to account for
georeferencing errors, whereas the maximum size was limited by the need to fit enough samples
within the study area for propensity score matching. The entire sampling process and analysis
were conducted separately for the 5 km and 10 km samples, thus providing statistics on the
impact of conflict at two spatial scales.

Two logistic regression models (separate models for the 5 km and 10 km sample) were fitted
with the explanatory variables to calculate propensity scores that would represent the likelihood
of conflict occurrence. The model was tested for goodness of fit with $R^2$, Receiver Operating
Characteristic Area Under the Curve (AUC), Akaike Information Criterion (AIC), null deviance,
and residual deviance. Additionally, the model was checked for multicollinearity with variance
inflation factor (VIF) and the statistical significance of covariate coefficients was assured (a summary can be found in Supplementary Table 10)\(^\text{47}\). The conflict and non-conflict samples were matched based on the similarity of propensity scores using the MatchIt package in R, parameterized with “nearest neighbor” and restricted to the area of common support (a summary can be found in Supplementary Table 11)\(^\text{48,49}\). This yielded subsamples of both the 5 and 10 km samples with a similar likelihood of the occurrence of conflict, thus reducing bias in the subsequent group comparison (Supplementary Figure 4, Table 12 and 13). The comparison was conducted to measure “the average effect of treatment on the treated”, in this case meaning the effect of conflict on the annual change in cropland extent. In addition to absolute change, change was also measured and compared as a proportion of the pre-existing cropland extent.

Evaluation of the Impact of Cropland Abandonment on Kilocalorie Availability

We summarized abandoned lands at the county level (administrative level-3) and linked abandoned areas with expected crop type proportions, yield, and kilocalorie estimates. Then, we calculated the unattained kilocalories. Crop type proportions were available at the sub-national level (Western Equatoria, Central Equatoria, and Lakes) circa 2018, for the major crops comprising 88% of produced crops in the study area: sorghum, maize, rice, finger millet, pearl millet, groundnuts, cassava, beans, and sesame\(^\text{11,21,29,50}\). Average yields (2014-2016) for cassava and groundnuts were available at the province level\(^\text{29,51,52}\) (WFP/FAO, 2014-2017) and for other crops at the national level\(^\text{53}\). To convert the unattained crop yields into kilocalories, we used extensive nutrient profile data on various crops from the Food Data Central of the USDA (Supplementary Table 13)\(^\text{16}\). We assessed how many households and people could be supported
if abandonment had not occurred by using the reported average of 2180 kcal per day per capita in South Sudan⁵, which was slightly above the Minimum Dietary Energy Requirements (MDER) of 1751 kcal per day per capita.
References


9. FAO, IFAD, UNICEF, WFP & WHO. *The State of Food Security and Nutrition in the*


18. Valero, S. et al. Production of a dynamic cropland mask by processing remote sensing


28. Chaudhary, S. *et al.* A Synopsis of Farmland Abandonment and Its Driving Factors in
Nepal. Land 9, 84 (2020).


Figures

Figure 1

Land-cover map for 2018 produced with supervised image classification and the examples of land-cover changes in three agricultural areas—Rumbek, Yambio, and Yei (A, B, and C, respectively). Note: The designations employed and the presentation of the material on this map do not imply the expression of
any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

Figure 2

Missed crop production due to cropland abandonment.

**Supplementary Files**

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementaryinformation.pdf](#)