



Article

Annual Field-Scale Maps of Tall and Short Crops at the Global Scale Using GEDI and Sentinel-2

Stefania Di Tommaso ¹, Sherrie Wang ^{2,3}, Vivek Vajiyeey ¹, Noel Gorelick ^{4,5}, Rob Strey ⁶ and David B. Lobell ^{1,*}

¹ Department of Earth System Science & Center on Food Security and the Environment, Stanford University, Stanford, CA 94305, USA; sditom@stanford.edu (S.D.T.); vvajiyeey@stanford.edu (V.V.)

² Goldman School of Public Policy, University of California, Berkeley, CA 94720, USA

³ Department of Mechanical Engineering & Institute for Data, Systems, and Society, MIT, Cambridge, MA 02139, USA

⁴ Google, 8002 Zurich, Switzerland; gorelick@google.com

⁵ Remote Sensing Laboratories, University of Zürich, 8006 Zurich, Switzerland

⁶ Progressive Environmental & Agricultural Technologies, 10435 Berlin, Germany; rob@plantix.net

* Correspondence: dlobell@stanford.edu

Abstract: Crop type maps are critical for tracking agricultural land use and estimating crop production. Remote sensing has proven an efficient and reliable tool for creating these maps in regions with abundant ground labels for model training, yet these labels remain difficult to obtain for many regions and years. NASA's Global Ecosystem Dynamics Investigation (GEDI) spaceborne LiDAR instrument, originally designed for forest monitoring, has shown promise for distinguishing tall and short crops. In the current study, we leverage GEDI to develop wall-to-wall maps of short vs. tall crops on a global scale at 10 m resolution for 2019–2021. Specifically, we show that (i) GEDI returns can reliably be classified into tall and short crops after removing shots with extreme view angles or topographic slope, (ii) the frequency of tall crops over time can be used to identify months when tall crops are at their peak height, and (iii) GEDI shots in these months can then be used to train random forest models that use Sentinel-2 time series to accurately predict short vs. tall crops. Independent reference data from around the world are then used to evaluate these GEDI-S2 maps. We find that GEDI-S2 performed nearly as well as models trained on thousands of local reference training points, with accuracies of at least 87% and often above 90% throughout the Americas, Europe, and East Asia. A systematic underestimation of tall crop area was observed in regions where crops frequently exhibit low biomass, namely Africa and South Asia, and further work is needed in these systems. Although the GEDI-S2 approach only differentiates tall from short crops, in many landscapes this distinction is sufficient to map individual crop types (e.g., maize vs. soy, sugarcane vs. rice). The combination of GEDI and Sentinel-2 thus presents a very promising path towards global crop mapping with minimal reliance on ground data.

Keywords: Google Earth Engine; cropland; agriculture; LiDAR; classification; maize



Citation: Di Tommaso, S.; Wang, S.; Vajiyeey, V.; Gorelick, N.; Strey, R.; Lobell, D.B. Annual Field-Scale Maps of Tall and Short Crops at the Global Scale Using GEDI and Sentinel-2. *Remote Sens.* **2023**, *15*, 4123. <https://doi.org/10.3390/rs15174123>

Academic Editors: Sidike Paheding and Ashraf Saleem

Received: 25 July 2023

Revised: 10 August 2023

Accepted: 18 August 2023

Published: 22 August 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Farmer livelihoods and food production are affected by myriad ongoing changes in climate, markets, and policies. Accurate data on cropping systems are essential to monitor and understand the effects of these changes, yet such data are often lacking. One major aspect of cropping systems are the crops that farmers choose to plant, which typically change from season to season as farmers rotate crops or shift into new crops [1]. Information on crop choice is helpful for various applications, including modeling land use decisions, mapping yield variations, and forecasting regional food production.

Given the widespread demand for crop type information, maps of crop types have been developed from a variety of sources and at a range of spatial and temporal resolutions [2,3].

In a small number of countries, such as the United States [4], Canada [5], and France [6], detailed crop type maps at the field scale are publicly available for each growing season, based either on farmer surveys or a combination of ground and satellite sources. In most countries, however, timely data is much harder to obtain. Several gridded datasets with global coverage have been developed, but these are often based on census data more than a decade old [2]. Given the dynamic nature of agriculture, including evidence of rapid cropland expansion in some regions and cropland abandonment in others [7], decades old data are insufficient for many uses. Moreover, global products typically have a resolution of 10 km or coarser [2], which limits their utility for applications requiring field-scale data.

As a result, there is a continued need for improved approaches to mapping crop types [2,3]. This need is recognized, for example, by the new WorldCereal effort that aims to create global, annual maps for wheat and maize at 10 m resolution (<https://esa-worldcereal.org/>, accessed on 1 March 2023). Remote sensing offers clear advantages for large-scale crop type mapping, with proven success in many local or national scale studies [4,8–12]. Yet a major challenge remains that models require large amounts of training data, and models trained in one region for a single season often do not transfer well to other regions or seasons. One sensible way to address this challenge, as in the WorldCereal project, is to invest in large amounts of field data collection around the world, so that models can be locally trained anywhere. Other efforts have focused on developing models that are better able to maintain performance in years or locations outside of their training domain [13–15].

A third, complementary approach has been to seek training data derived without the need for field data collection. In recent work [16], we demonstrated the promise of one such source of data—LiDAR measurements acquired by the Global Ecosystem Dynamics Investigation (GEDI) [17]. GEDI LiDAR returns provide information on canopy heights with a nominal spatial resolution of 25 m and a vertical precision of roughly 50 cm [17]. Although many crops have similar heights, some of the key commodity crops grown throughout the world, especially maize and sugarcane, are typically 1 m taller than other common crops such as wheat, rice, or soybean (Figure 1).

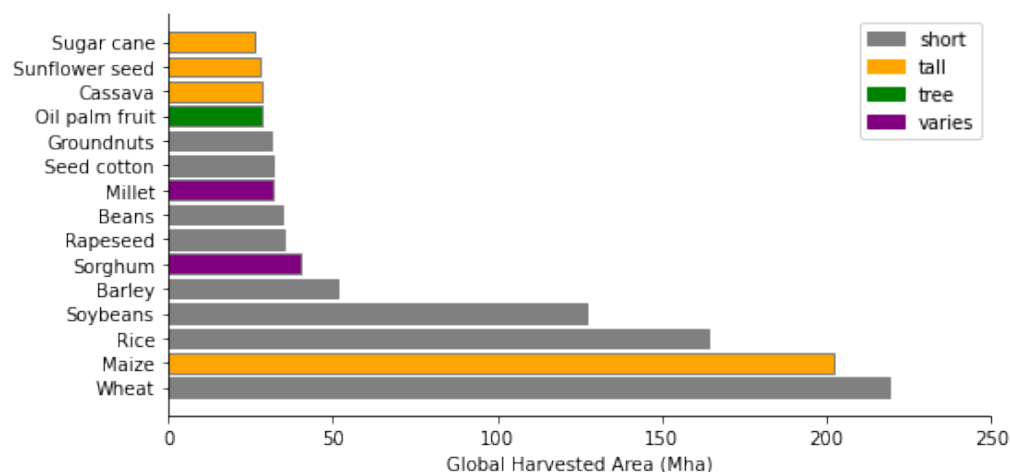


Figure 1. Most commonly grown crops in the world based on FAO crop statistics for 2019, color coded by crop heights base on U.S. National Plan Germplasm System. Crops are considered tall if most common varieties exceed 2 m in peak height. Data source: <http://www.fao.org/faostat> (accessed on 1 February 2023) for global crop areas, <https://npgsweb.ars-grin.gov> (accessed on 1 February 2023) for crop heights.

Indeed, in many landscapes the two main crops are one tall crop and one short crop (e.g., maize and soybean, or sugarcane and rice), such that the ability to discern tall from short crops goes a long way toward mapping individual crops (Figure 2). For example, roughly two-thirds of the area sown to maize in the world lies in regions where maize is 90% or more of the area covered by tall crops.

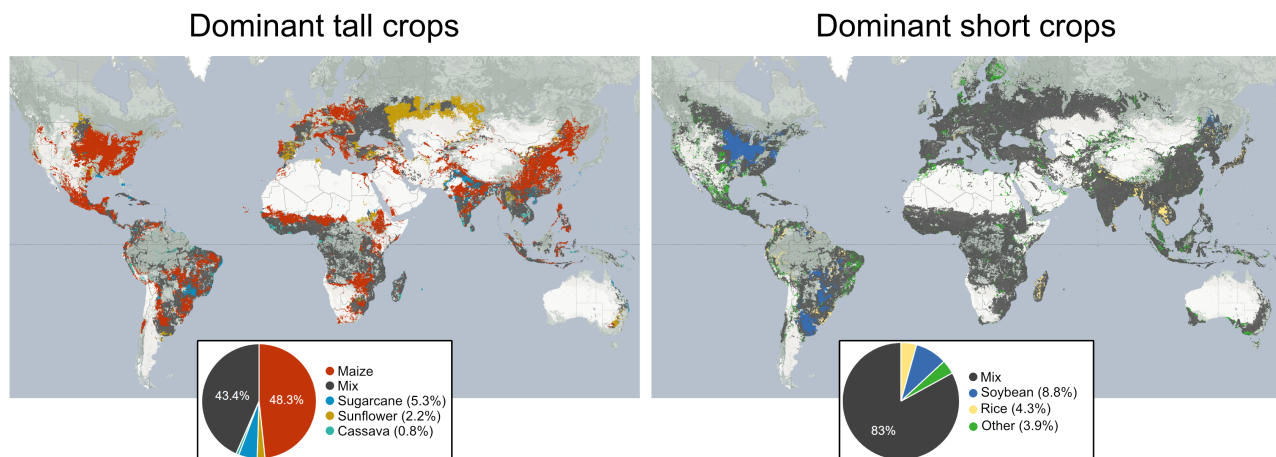


Figure 2. Maps illustrating dominant crops based on IFPRI SPAM-2010 v2.0 [18] (using the layer corresponding to physical area—all technologies together). In these maps, a crop is deemed dominant in a region if it accounts for more than 90% of the tall or short crop total area. The map on the left displays tall crops (i.e., maize, sugarcane, sunflower, and cassava), with maize dominating in 48% of areas with tall crops. Put another way, 68% of the total global area where maize is grown consists of regions where maize is the dominant crop; for sugarcane, this fraction is 46%. The map on the right represents short crops. It includes all the short crops listed in the SPAM dataset (with the exception of wheat, barley, and rapeseed grown primarily during winter). A total of 50% of the global area where soybean is grown consists of regions where soybean is the dominant crop, while for rice this fraction is around 20%.

GEDI alone, however, only samples a very small fraction of the landscape, so therefore rather than use GEDI directly, Di Tommaso et al. [16] use GEDI to train a random forest model that predicts crop height class based on Sentinel-2 (S2) optical data. This combined GEDI-S2 approach was found to map crop types nearly as well as a model trained on thousands of local ground training points in the US, France, and China.

Here, we develop and test an approach that combines GEDI and S2 to map tall and short crops throughout the world for a three-year period (2019–2021). We extend the initial insight from Di Tommaso et al. [16], namely that GEDI signals are informative in cropped landscapes, in several important ways. These include a method to automatically identify the most appropriate months for tall crop delineation, an investigation into the effects of view angle and topography on GEDI signals in the context of crop discrimination, and global scale implementation of the GEDI-S2 approach. We also conduct an evaluation of GEDI-S2 models in a much broader set of countries and cropping systems, using various independent datasets on crop types during the study period. Overall, we find that GEDI is a useful resource for advancing the goal of low-cost, timely, and accurate global mapping of crop types. At the same time, we identify some important areas for improvement to guide future research efforts.

The following section describes the various datasets used in the study, including any initial processing steps for the data. Section 3 then describes the methods used to map crop height class and evaluate the predictions. Section 4 presents the main results, while Section 5 discusses various sources of errors and potential future directions for improvement. Finally, Section 6 briefly summarizes the main conclusions.

2. Datasets

This study utilized five main data sources: a global cropland mask used to define cropland areas, GEDI shot returns for cropland areas, Sentinel-2 optical imagery, reference data on crop types from three regions used to train the GEDI classification model, and reference data on crop types from throughout the world used to evaluate the performance of our GEDI-S2 tall/short crop predictions. Below we describe each of these, as well as supplementary datasets used to analyze and interpret our results, including a global map that defines the number of growing seasons in each location, and reference crop type maps used to analyze the relationship between peak crop biomass and model errors.

2.1. Crop Mask

To identify cropped areas we used the European Space Agency (ESA) [19] and ESRI [20] global Sentinel-based 10 m global land cover maps available in the Google Earth Engine (GEE) [21] official and community data catalogs, respectively [22]. Both the ESA World-Cover 2020 product and ESRI 2020 Global Land Use Land Cover provide a global land cover map for 2020 at 10 m resolution, the former based on Sentinel-1 and Sentinel-2 data, and the latter based on Sentinel-2 alone. We primarily used the ESA mask, which, based on visual inspection using Google's high-resolution basemap, better captured cropland in most areas where the two maps disagreed. However, for Kenya and Uganda the ESA mask tended to greatly underestimate cropland area, and to better capture cropland for these countries we therefore merged the two masks, defining a pixel as cropland if either of the two classified it as cropland.

2.2. GEDI Data

GEDI is a sensor onboard the International Space Station (ISS) that acquires LiDAR waveforms between 51.6°N and 51.6°S to observe the Earth's surface in 3D. It is the first spaceborne LiDAR instrument specifically optimized to measure vegetation structure [17]. It contains three lasers emitting near-infrared (1064 nm) light. Two of the lasers are full-power lasers, with the other coverage laser split into two beams, producing a total of four beams. Each beam is then optically dithered across-track resulting in eight ground tracks (four full power and four cover tracks) spaced 600 m on the ground. Shots have an average footprint of 25 m in diameter and are separated 60 m along the track.

GEDI spatial coverage changes in time. In particular, in early 2020 the ISS lifted its orbit, causing GEDI to have "orbital resonance" which means it goes over the same tracks repeatedly while leaving big gaps in between (Figure 3a–c). While orbital resonance does not change the number of shots acquired in a time period, it reduced the spatial coverage of GEDI in 2020 relative to 2019. When GEDI samples of agricultural areas are less geographically uniform and more clustered, we expect GEDI-based crop type classification accuracy to decrease.

Another important aspect of GEDI is that while its viewing angle is typically near-nadir, it can be rotated by up to 6°, allowing the lasers to be pointed up to 40 km on either side of the ISS ground track. This capability is used to sample the Earth's land surface as completely as possible, but can also complicate interpretation of the GEDI returns [23].

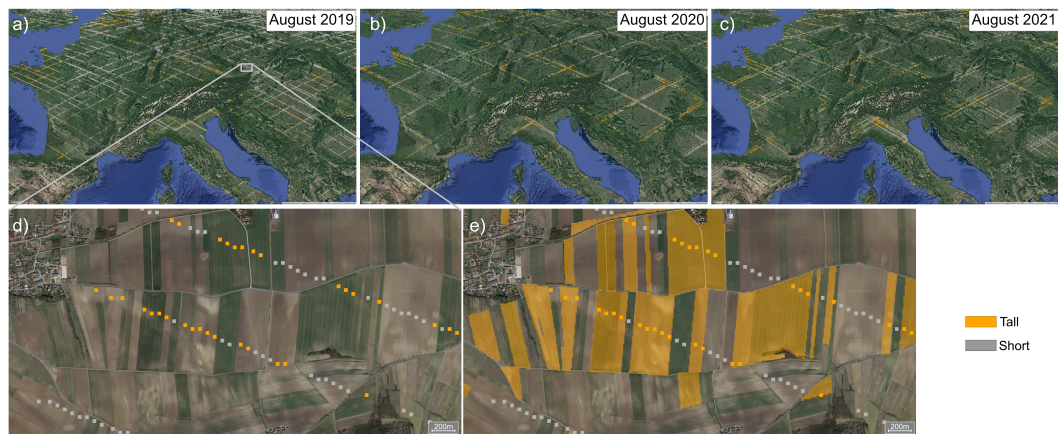


Figure 3. Spatial locations of GEDI shots over croplands in Europe. (a–c) Show spatial distribution for 2019–2021, respectively. Starting in 2020 higher International Space Station (ISS) altitudes caused the clustering of GEDI observations along its orbital track (i.e., “resonance”) leaving bigger gaps across tracks. (d) Zoom in of GEDI shots for 2019 over fields in Austria, color coded by GEDI model predicted class, gray for short, orange for tall crop. (e) Same as (d) but also showing the fields growing maize according to ground truth from Austria parcels dataset, illustrating that GEDI predicted class agrees well with the ground truth.

For this study, we used the GEDI dataset Level 2A (L2A) and Level 2B (L2B) from April 2019 to December 2021, available in GEE data catalog. Level 2 data provide information about the vertical distribution of the canopy retrieved from the waveform return at footprint level. The main GEDI product used is GEDI’s L2A Geolocated Elevation and Height Metrics Product, which is primarily composed of Relative Height (RH) metrics, which collectively describe the waveform collected by GEDI. Relative Height (RH) metrics give the height at which a certain percentile of energy is returned relative to the ground. RH are reported at 1% intervals, resulting in 101 metrics. The GEDI L2A dataset (LARSE/GEDI/GEDI02_A_002_MONTHLY) is a rasterized version of the original GEDI product, with each GEDI shot footprint represented by a 25 m pixel [24]. This rasterization process can introduce an additional geolocation error to the initial GEDI shot error. The raster images are organized as monthly composites of individual orbits in the corresponding month. RH values and their associated quality flags and metadata are preserved as raster bands.

A secondary dataset L2B was used to retrieve the GEDI view angle (i.e., local beam elevation property). This is available in GEE as a table of points (LARSE/GEDI/GEDI02_B_002) with a spatial resolution (average footprint) of 25 m. At the time of writing, the raster version of the L2B dataset (LARSE/GEDI/GEDI02_B_002_MONTHLY) is only partially ingested in GEE, and we therefore used the table.

2.3. Sentinel-2

We used S2 surface reflectance data (Level-2A) present in GEE and filtered out clouds using the S2 Cloud Probability dataset provided by SentinelHub in GEE. The Sentinel-2A/B satellites acquire images with a spatial resolution of 10 m (Blue, Green, Red, and NIR bands) and 20 m (Red Edge 1, Red Edge 2, Red Edge 3, Red Edge 4, SWIR1, and SWIR2 bands), and together they provide images at a 5-day interval.

To capture crop phenology, we extracted S2 imagery for 2019–2021 from 1 January to 31 December for the northern hemisphere, and from 1 July of one year to 30 June of the next for the southern hemisphere. Features were extracted from S2 time series by fitting harmonic regressions to all cloud-free observations in cropped areas. For each spectral band or vegetation index $f(t)$, the harmonic regression takes the form

$$f(t) = c + \sum_{k=1}^n [a_k \cos(2\pi\omega kt) + b_k \sin(2\pi\omega kt)]$$

where a_k are cosine coefficients, b_k are sine coefficients, and c is the intercept term. The independent variable t represents the time an image is taken within a year expressed as a fraction between 0 and 1. The number of harmonic terms n and the periodicity of the harmonic basis controlled by ω are hyperparameters of the regression.

To determine n and ω , we sampled multiple locations around the world and compared the harmonic fit of the time series by varying the hyperparameters. We found that third order harmonics ($n = 3$) with $\omega = 1$ were a good fit for both regions with one or multiple growing seasons.

We computed harmonic coefficients for four bands and one vegetation index: NIR, SWIR1, SWIR2, RDED4 and GCVI. GCVI is the green chlorophyll vegetation index [25] computed as

$$\text{GCVI} = \text{NIR}/\text{Green} - 1$$

This yields seven features per band, for a total of 35 coefficients. Previous crop type classification studies [26,27] have reported the efficacy of using these four bands and VI, demonstrating performance comparable to classification models using all optical bands and a variety of other VIs.

2.4. GEDI Model Training Dataset

To train the GEDI model to distinguish tall from short crops we used high-accuracy crop type labels from 2019 from the three regions used in prior work [16] and mapped in red in Figure 4: Jilin province in China, Grand Est region in France, and Iowa state in USA. These regions are major agricultural production areas containing a mix of tall and short crops and have accurate, field-scale crop type maps that are publicly available. Although at similar latitudes, these regions are located in three separate continents and management practices do differ. Maize in France in particular exhibits a wide range of GCVI, and China exhibits very small fields. Differences in agricultural practices across regions for the same classes could translate to differences in the GEDI waveforms, helping the GEDI model to be more flexible and adaptable in other regions as well.

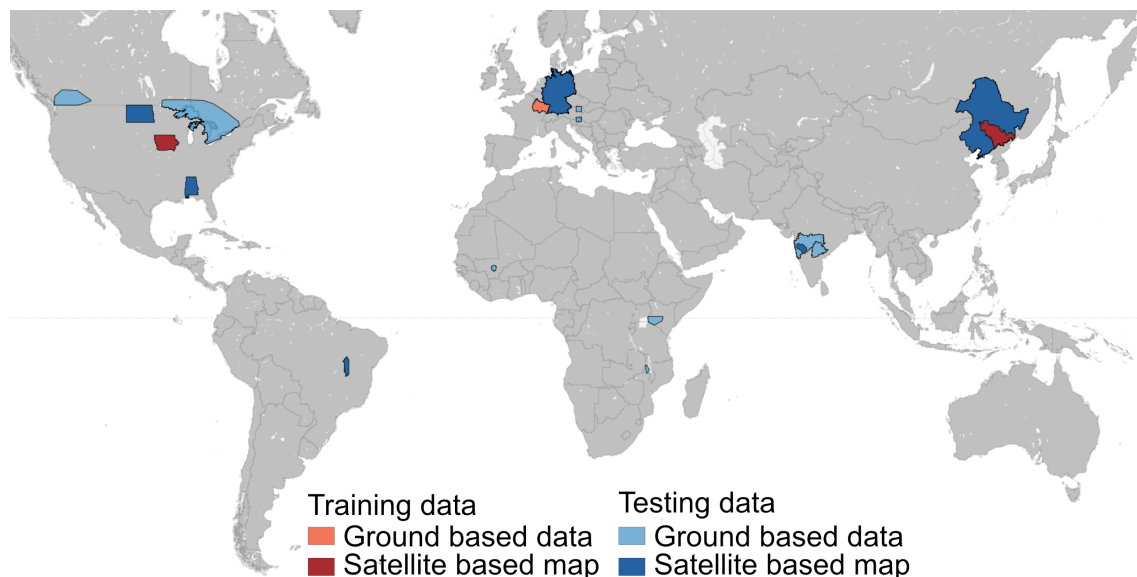


Figure 4. Distribution of data used for training the GEDI model in red, and independent evaluation of the final maps in blue. Ground-based refers to either point or polygon data collected on the ground. Satellite-based refers to maps typically created by combining ground data with remote sensing datasets, such as the Cropland Data Layer in the United States.

For Jilin, China we used the 2019 crop type map produced by You et al. [9]. It maps three major crops in the area (maize, soybean, and rice) at 10 m with an accuracy of 87%, and F1-scores of 85% for maize. For Grand Est, France we used the Registre Parcellaire Graphique (RPG) 2019 dataset downloaded from <https://www.data.gouv.fr/> (accessed on 1 July 2021). It is a public georeferenced vector product derived via survey. For Iowa, USA we used the U.S. Department of Agriculture’s 2019 Cropland Data Layer (CDL) at 30 m resolution available in GEE [4]. It has an overall accuracy of 90%, and precision and recall for maize exceed 95%.

2.5. Evaluation Datasets

To evaluate our product, we sought high-quality crop type datasets for a diverse set of cropping systems and regions for the 2019–2021 period. We used a combination of field data and crop type maps that were produced by combining field and satellite data. In regions with multiple growing seasons, we filtered for crop type labels matching the growing season that the GEDI model predicted for. A map of location and type of reference data is shown in Figure 4 and a summary of data characteristics including sample size are given in Table 1.

Table 1. Characteristics of the regional datasets used to evaluate GEDI-S2. For each dataset we report the year, the type of reference data, the number of samples, the percentage of reference data labelled as a tall crop and the main labels. *Point* and *polygon* data type refers to ground-based data, while *map* refers to satellite-based maps. US ND and AL stand for North Dakota and Alabama. Brazil BA for Bahia. India UBB for Upper Bhima Basin, TG for Telangana, and MH for Maharashtra.

Region	Year	Type	Samples	% Tall Crop Labels	Main Labels
Austria	2019	polygons	159,528	16.1	maize, pasture, wheat
Slovenia	2019	polygons	122,792	40	maize, wheat, barley
Germany	2019	map	2278	24.1	maize, wheat, barley
Germany	2020	map	2864	15.1	maize, wheat, barley
Canada (BC)	2019	points	704	31.3	mixed forage, maize
Canada (ON)	2019	points	29,960	33.9	soybean, maize, mixed forage
Canada (BC)	2020	points	871	39.6	mixed forage, maize, alfalfa
Canada (ON)	2020	points	14,960	31.2	soybean, maize, mixed forage
Canada (BC)	2021	points	15,384	31.2	mixed forage, maize, alfalfa
US (ND)	2019	map	1847	18.8	soybean, wheat, maize
US (ND)	2020	map	1860	11.4	soybean, wheat, maize
US (ND)	2021	map	1882	22.4	soybean, wheat, maize
US (AL)	2019	map	1085	24.1	cotton, maize, soybean
US (AL)	2020	map	1088	24.7	cotton, maize, soybean
US (AL)	2021	map	1078	25	cotton, maize, soybean
Brazil (BA)	2020	map	1992	0	soybean
China	2019	map	2736	56.5	maize, soybean, rice
India (U.B.B.)	2020	map	1211	50.6	sugarcane, cotton, rice
India (TG)	2020	points	4844	4.6	rice, cotton, peanut, maize
India (TG)	2021	points	28,562	4.9	rice, cotton, peanut, maize
India (MH)	2020	points	8639	27.7	cotton, maize, rice, sugarcane
Malawi	2021	polygons	719	31.4	groundnut, maize, soybean
Mali	2019	polygons	73	26	sorghum, millet, maize, rice
Kenya	2021	points	1423	58.1	maize, tea, sugarcane

2.5.1. Ground-Based Reference Data

Europe

Schneider et al. [28] contains harmonized agricultural parcels information data from regions in Austria (2019), Denmark (2020), and Slovenia (2019). The parcel data are based on publicly available self-declared crop reporting datasets, gathered for the purposes of

subsidy payments. We focus on Austria and Slovenia, since the Denmark dataset is outside the GEDI latitude coverage.

Canada

Agriculture and Agri-Food Canada [29] is a collection of thousands of points identifying crops types and occasionally other land cover types across Canada from 2011 to 2021. These point sources are used by Agriculture and Agri-Food Canada (AAFC) as training or reference sites for the creation of the Annual Crop Type map.

Malawi

Field boundaries for three crops—groundnut, maize, and soybean—were collected in five districts of Malawi (Lilongwe, Ntchisi, Kasungu, Salima, and Mzimba) for 2021 as part of research on the groundnut value chain conducted by the AgroHitech Innovation and Advisory Consortium for the Peanut Innovation Lab.

Mali

Field boundaries in Mali were collected by the NASA Harvest team during the 2019 growing season [30]. The crop type growing in each field was observed by a surveyor. In total, the dataset contains 148 fields. The data were released as part of the CropHarvest dataset and is also available on Radiant MLHub.

Kenya

The Global Agriculture Monitoring initiative of the Group on Earth Observation, called Copernicus4GEOGLAM, collected ground reference data during field surveys in three countries—Uganda, Tanzania, and Kenya [31]—for both the 2021 long rain and 2021–2022 short rain seasons. The georeferenced ground data were used by Copernicus4GEOGLAM to train random forest models to map crop type using S2 imagery as input. Given the fairly low accuracies of the resulting maps (e.g., maize F1 scores were often below 0.6), we utilized only the field data for our evaluation. We focused on the Kenya point dataset for the 2021–2022 short rain season, which had the greatest number of points overlapping with the season of our GEDI-S2 predictions.

India

India crop type labels are crowdsourced from farmers via Plantix, a free Android application created by Progressive Environmental and Agricultural Technologies (PEAT). The Plantix app is used by farmers who submit photos of their crops seeking help to diagnose and treat crop diseases. As part of the disease diagnosis, PEAT uses a convolutional neural network to assign crop labels based on the submitted photos. We used these data in the Indian states of Maharashtra and Telangana, where the accuracy of Plantix crop type labels exceeds 0.90 for most major crops. These data have been cleaned to remove location inaccuracy (keeping only submissions with GPS accuracy better than 10 m), as suggested by previous work by [32]. To match the timing of the GEDI-S2 predictions, we filtered the Plantix data for the 2021 kharif season based on photo submission timing.

2.5.2. Satellite-Based Reference Data

United States

The Cropland Data Layer (CDL) produced by the United States Department of Agriculture (USDA) provides yearly crop type maps across the conterminous US at 30 m spatial resolution [4]. Maps are based on Landsat and other satellite imagery using training data from the Farm Service Agency (FSA). For validation we chose two states, North Dakota and Alabama, that were far from the conditions and locations of the Iowa locations used in the training data. Accuracy of CDL on FSA labels are available in the CDL metadata, with precision and recall for maize for 2019–2021 higher than 81% and 85% in North Dakota and Alabama, respectively.

Germany

National scale crop type maps for Germany were recently produced for 2017–2019 [33] and 2020 [34]. These maps are generated using a random forest classifier based on Sentinel-1, Sentinel-2 and Landsat time series, with parcel data used for training. More details about the underlying data and methods can be found in Blickensdörfer et al. [12]. Overall accuracy for 2019 is 78%, with precision and recall for the maize class of 90% and 83%, respectively.

Brazil

Annual soybean maps were recently produced for South America at 30 m resolution between 2000 and 2020 by combining Landsat and MODIS satellite observations and sample field data [11]. These maps are available in GEE as (projects/glad/soy_annual_SA). For evaluation, we focused on western Bahia in 2020, since this region grows maize and soy in the main season and has only one primary growing season per year. We evaluated only recall on soy, since other short crops, e.g., cotton, are also grown in the same season but are not distinguished from other non-soy crops in their study. Accuracy for 2020 is not reported, but for the years 2017–2019 they report overall accuracies of 96%, 94%, and 96%, respectively, with high and balanced producer's and user's accuracies.

China

For China, we used the same 2019 crop type map described by You et al. [9] in Section 2.4 used in training the GEDI model. For validation, we used a random sample over the four northeast regions (Liaoning, Nei Mongol, Jilin, and Heilongjiang), which span a much larger area than used in the training sample from Jilin.

India

Lee et al. [35] produced a map of sugarcane area in the Upper Bhima Basin, a major sugarcane producing region in Maharashtra, India. Their 10 m resolution map is based on crowdsourced Plantix data and a neural network applied to S2 data. Reported overall accuracy for sugarcane vs. not sugarcane was 77% (85% precision and 67% recall).

2.6. Number of Growing Seasons per Year

The Anomaly hotspots of Agricultural Production (ASAP) system is an online decision support tool for early warning about production anomalies developed by the Joint Research Center (JRC) of the European Commission. ASAP has produced several maps including satellite-based phenology information, which are computed from the long-term average of MODIS NDVI data at 0.01° resolution [36]. We downloaded the phenology layer that defines the number of growing seasons (1 or 2 seasons) [37], and aggregated this information at 5° based on the majority of the crop pixels' seasonality.

2.7. Digital Elevation Model (DEM)

We used a DEM to investigate the effect of topography on the usability of GEDI shots for tall crop classification. The Shuttle Radar Topography Mission SRTM V3 (SRTM Plus) [38] digital elevation data product is provided by NASA JPL at a resolution of 1 arc-second (approximately 30 m) and is available in GEE. We calculated the slope in degrees from the terrain DEM in GEE.

2.8. Reference Maps for Error Analysis

As described below, we hypothesize that errors in our GEDI-S2 predictions were often related to low biomass of the tall crop. To further investigate this, we utilized two additional crop type maps that provided wall-to-wall coverage in countries where our preferred data for evaluation covered only a subset of fields. Widespread coverage was needed to ensure a wide range of biomass values for pixels in the reference map that overlap with the GEDI shot locations.

2.8.1. Canada

The Earth Observation Team of the Science and Technology Branch at Agriculture and Agri-Food Canada (AAFC) have created Annual Crop Inventory maps that are accessible in GEE. These maps are generated using a combination of crop type labels from crop insurance data and ground-truth information collected across the country to train a decision tree model based on optical and radar satellite images. Maps have a spatial resolution of 30 m and an accuracy of at least 85%.

2.8.2. Kenya

As described above, in addition to field data, the Copernicus4GEOGLAM produces end-of-season crop type maps for each country and season where field data was performed [31]. In our error analysis, we used the long rains map for Kenya, which possesses the highest F1 score for maize among the various countries and seasons. For this map, overall crop type accuracy is 80%, and F1 for maize is 0.64.

3. Methods

Here, we describe the steps taken to create and evaluate wall-to-wall maps of crop type height, using a combination of GEDI and S2 as input. Figure 5 provide a graphical overview of the methods presented in this paper.

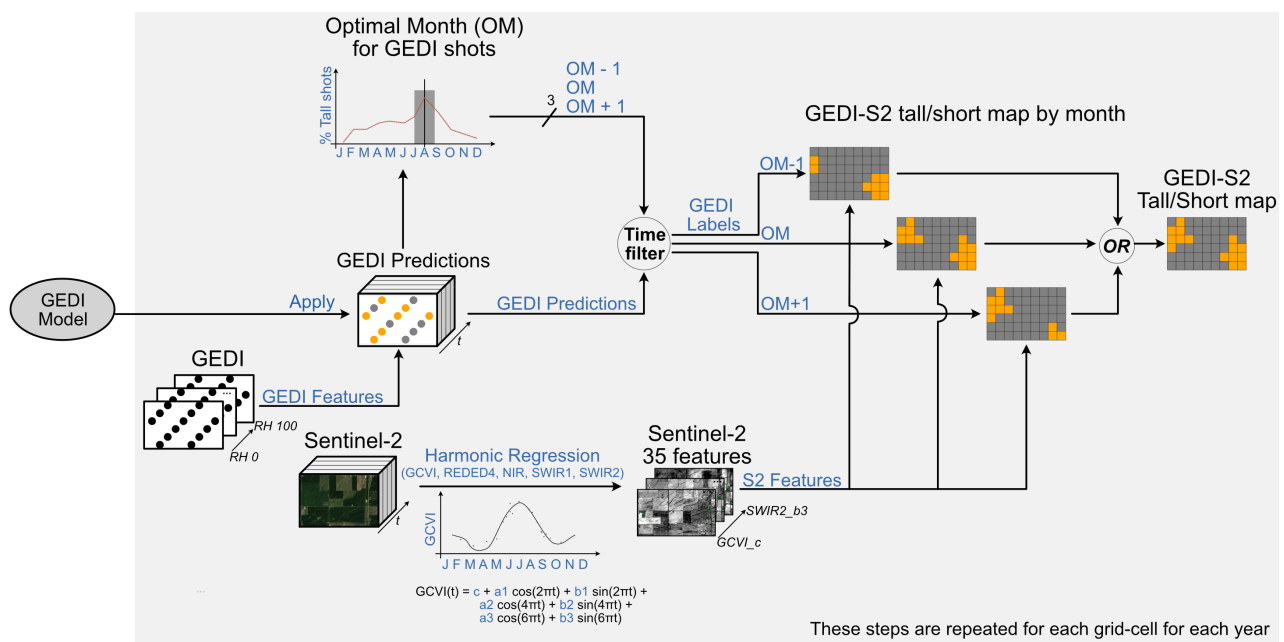


Figure 5. Flowchart of the steps for creating the GEDI-S2 tall/short map for each grid-cell in each year. The GEDI model is trained a single time using the GEDI RH metrics and corresponding crop type maps in three training regions (see Section 3.1). The steps in the gray box are repeated for each $5^\circ \times 5^\circ$ grid cell. To create the global map for a specific year, all the individual grid cell maps are mosaicked.

The sections below describe in detail each of the six steps in this process:

1. Train a single model, which we refer to as the *GEDI model*, that uses GEDI features to classify locations as having short crops, tall crops, or trees;
2. Apply the GEDI model to GEDI shots acquired from cropland areas globally for three years of 2019–2021;
3. Tile the globe into $5^\circ \times 5^\circ$ grid cells;
4. Determine the optimal month to predict tall crops for each grid-cell;
5. Train a local *GEDI-S2 model* for each grid-cell based on GEDI predictions in the 3-month time window around the optimal month;

6. Evaluate results against local reference data.

3.1. GEDI Model Training

Following Di Tommaso et al. [16], we began by defining a random forest model to classify GEDI shots in three crop height classes: short, tall or tree. The decision to use the random forest model was primarily motivated by its high accuracy, advantageous computational efficiency and seamless implementation at a large scale in GEE. To train the model we used labels from three areas with high-quality crop type maps for 2019: Jilin in China, Grand Est in France, and Iowa in the United States (see Section 2.4). Crop type labels were sampled at GEDI shot locations and assigned a *tall* label for maize class and a *short* label to remaining short crops. We also defined a third *tree* class, for shots with RH100 greater than 10 m. The choice of the 10 m threshold was made empirically, relying on visual assessment of GEDI shots over trees in Google's high-resolution basemap.

We used all GEDI shots in August 2019 over the three regions since our previous study [16] showed August to be a good time to distinguish maize from other short crops in these regions. This resulted in a total of approximately 253 k samples, with 47 k samples in Jilin, 23 k in Grand Est and 183 k in Iowa.

Although we are interested in the crop height, we found that our model worked best when multiple RHs were included to fully capture the GEDI returned waveform. To reduce the number of features, since consecutive RH metrics are highly correlated with each other, we sampled a metric every 5% and omitted RH in the middle of the RH profile based on feature importance analysis. In total, 11 RH metrics were used: RH0, RH5, RH10, RH15, RH20, RH25, RH30, RH35, RH40, RH45, and RH50.

The features and labels were used in a random forest model, implemented in GEE. Data were split into 80% training and 20% test points. To minimize spatial correlation across the training and test sets, we binned the shots by their lat/lon into $0.5^\circ \times 0.5^\circ$ bins and GEDI shots in each bin were placed entirely in either the training set or test set. The overall test accuracy across the three regions was 0.885, with F1 scores for short, tall and tree classes of 0.863, 0.898, and 1.000, respectively. The very high F1 score for the tree class is explained by the definition of tree class, as based on GEDI RH100 metric directly and not field labels.

3.2. GEDI Model Predictions

The random forest model described above was then applied to all GEDI shots in cropland pixels, according to the crop mask described in Section 2.1. The predicted class was saved along with the prediction probabilities (the fraction of trees in the random forest model that predicted the class) as a measure of confidence. Figure 3 illustrates these predictions for a selection of GEDI orbits, with shots colored orange for tall and gray for short based on the GEDI model predictions.

The predicted shots were filtered to retain only high quality shots to use as labels in subsequent steps. First, we removed shots with a quality flag value of zero in the original GEDI returns, which indicates poor quality, as well as shots with a non-zero degrade flag, which indicates poor geolocation. We then removed low confidence predictions (lower than 0.8) to have more confidence in the GEDI-generated labels.

Another step that proved essential was to filter out shots with low view angle and on high slope terrain since both factors can affect the accuracy of the GEDI model predictions. We refer to view angle as the angle between the off-nadir beam and the ground. Prior work has revealed that small changes in view angle can increase errors for models based on GEDI returns [23,39]. In particular, existing analysis recommends removing observations where the view angle was below 1.5 rad, or roughly 86° [39].

To explore the appropriate threshold for our application, we considered shots for the US Corn Belt where we have confidence in the reference data from CDL, and where the view angle property is available in GEE at the GEDI shot level. The GEDI model prediction errors (treating CDL as truth) were evaluated for different levels of view angle, as shown

in Figure 6a. At low view angles, errors are as high as 60%. Above the recommended threshold of 1.5 rad, however, errors are below 10% and fairly insensitive to additional increases in view angle. We therefore adopted a threshold of 1.51 rad for further analysis.

The GEDI view angle varies over time as shown in Figure 6b. View angles were particularly low in June and July of 2020, causing the removal of most shots during the peak of the growing season in many regions. Other periods of frequent observations with low view angles include late 2019 and late 2021. Unfortunately, at the time of writing, information on the GEDI view angle was not yet available at the shot level for all shots globally in the GEE catalog. To create a view angle filter, we sampled orbits over a longitudinal transect and aggregated these data, averaging the view angle for each beam on each day. We then removed all shots from beams and days with an average view angle above 1.51 rad. Although this was a pragmatic way to filter out data with low view angle globally, future versions would likely benefit from accounting for view angle at the shot level, to account for variation by latitude and over time within the day.

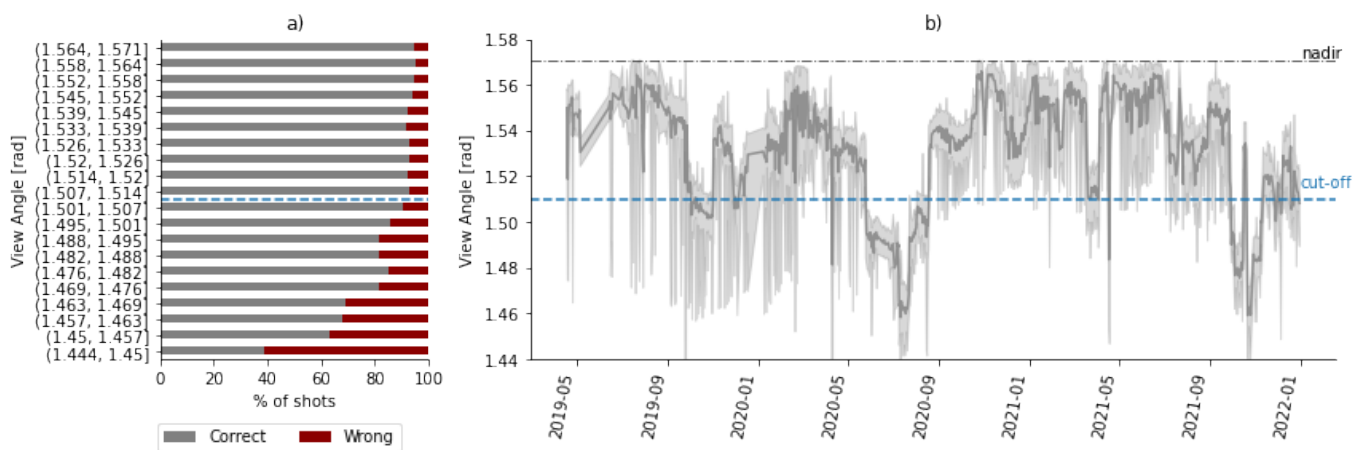


Figure 6. (a) Effects of GEDI view angle on the accuracy of the GEDI model predictions in the Central United States. (b) The temporal variation of view angle over the study period, with shading showing the min-max values. GEDI can be rotated up to 6° (0.1 rad) from nadir. In some periods, such as summer 2020, GEDI was systematically targeting reference ground tracks that were further off nadir than other times. Dashed line labeled “cut-off” indicates the threshold used to filter GEDI shots in this study.

We also removed shots on high slope terrain, defined as areas with slope higher than 5° . GEDI metrics are dependent on topographic slope [40], and given the relatively small height signal being used by our model to classify tall vs. short crops, the effect of topographic slope are potentially important. Based on analysis of CDL in the United States, similar to the view angle analysis presented in Figure 6a, a slope below 5° was deemed sufficient to avoid artifacts from the terrain. As cropland is typically situated on flat or nearly-flat land, this filter removed only a small fraction of GEDI shots.

3.3. Model Grids

The filtered GEDI model predictions provide labels with which to train a model that takes S2 data as input. However, we did not expect a single model to be applicable globally, since the timing of growing season and mix of crops differs across the world. Building on prior approaches [24,41] we instead sought to develop locally-calibrated models. We defined a grid within the GEDI coverage (between 51.6°N and 51.6°S) with $5^\circ \times 5^\circ$ cells. Although more localized models would potentially improve performance in some regions and years, the choice of grid cell size was dictated by the orbital resonance of GEDI in 2020 and 2021. That is, a finer grid would often have cells that have very few GEDI observations because of the large gaps in GEDI coverage in those years. Furthermore, moving from the

pole towards lower latitudes, the spacing between GEDI tracks becomes more substantial. This necessitated at lower latitude the adoption of larger grid cells in terms of real area to ensure sufficient data coverage. By employing 5-degree grid cells, we struck a balance between capturing the heterogeneity within each cell and ensuring an adequate sample size of GEDI data for model training and analysis.

To reduce computation, we only processed grid cells for which more than 5% of S2 pixels were classified as cropland, yielding 238 cells. The grid cells that we kept cover an area that comprises 93% of the total crop area within the latitude bands of GEDI coverage.

3.4. Optimal Timing

For each grid cell, we defined the optimal month to classify tall vs. short crops as the month in which the highest percentage of GEDI shots were predicted as tall. Specifically, we combined the 3 years of GEDI model predictions by month, computed the percentages of tall and short shots by grid cell and month, and then selected the month for each cell when the percentage of tall shots was highest. We interpret this month as the period during which tall crops have reached their peak height. Consequently, this is the time when GEDI is most likely to detect a contrast with other crops within that particular cell.

3.5. GEDI-S2 Models

For each grid cell and for each year, we separately trained a local 2-class (tall vs. short) S2 model using the GEDI predictions for the relevant time as labels and harmonic coefficients as features. Data were randomly split into 80% training and 20% test, to evaluate model accuracy. We refer to these as GEDI-S2 models, with a unique model for each grid cell and year. To account for variations in the timing of the growing season within the 5° grid cells, we considered a three-month window centered on the optimal month. We created GEDI-S2 predictions for individual months and then combined the predictions on a pixel basis, with pixels classified as tall if the predicted class was tall in any of the three months.

The result of this process was a wall-to-wall 10 m resolution map of tall and short crops for all cropland pixels in the 5° grid cells. To reduce computation, we only applied the GEDI-S2 models to grid cells where the percentage of tall shots is higher than 4%, i.e., 201 grid cells per year. Since GEDI data was not always available in all the regions in the time window of interest, the number of grid cells processed is 1562, less than the expected $1809 = 201 \text{ (grid cells)} \times 3 \text{ (months)} \times 3 \text{ (years)}$. This resulted into 189 unique locations in 2019, 457 in 2020 and 201 in 2021, for a total of 590 grid cells for the 3 years.

For this local training, we omitted all shots where the GEDI model predicted the tree class, as these were viewed as likely to be a mixture of crops and trees within the GEDI footprint, which at 25 m diameter is more than four times larger than the 10 m S2 pixel. Thus, predictions of tree were viewed as unreliable labels for a 2-class model focused on S2 pixels classified as cropland.

To minimize spatial artifacts when mosaicking adjacent cells, we created predictions for pixels in a 0.5° buffer around each cell and mosaicked the overlapping predictions taking the predictions in the cell with higher GEDI-S2 accuracy.

3.6. Evaluation of GEDI-S2 Predictions

The first evaluation of GEDI-S2 predictions is against reference data from around the globe (Table 1). All reference data were ingested in GEE for comparison with GEDI-S2 predictions. For regions with ground-based point or polygon data, we used all fields for evaluation. In the case of polygons, the centroid of the polygon was used to define the relevant pixel from the GEDI-S2 predictions for comparison. For regions where crop type maps were used, we randomly sampled the maps using 2000 to 4000 points and removed the ones without a specific crop type label to create a reference dataset.

Because some datasets contain as many as 100 crop types, for each reference dataset we selected the 10 most common crops for evaluation, which typically represents more than 90% of the crop areas in the reference regions evaluated. From specific crop type labels, we generated a binary tall/short classification, with maize and sugarcane defined as tall and all other crops defined as short (none of our evaluation data had sunflower or cassava among the 10 most common crops). For each evaluation dataset, we report the accuracy, precision, recall, F1 and Kappa scores, using the following equations [42]:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$F1\ score = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

$$Kappa\ score = \frac{P_o - P_e}{1 - P_e}$$

where

- True Positive, TP , is the number of samples labelled as positive by the model that are actually positive
- False Positive, FP , is the number of samples labelled as positive by the model that are actually negative
- True Negative, TN , is the number of samples labelled as negative by the model that are actually negative
- False Negative, FN , is the number of samples labelled as negative by the model that are actually positive
- P_o is the proportion of observed agreement, i.e., the accuracy achieved by the model
- P_e is the proportion of agreements expected by chance

Our second evaluation compares GEDI-S2 predictions against an S2 model trained locally within each reference region. These *S2-Local* models provide benchmarks that represent how well a model trained on local field data would perform. To conduct this analysis, we exported S2 harmonic features at the same reference field locations and trained a local S2 model based on binarized tall/short field labels. We used the `RandomForestClassifier` implemented in Python's `scikit-learn` package, setting similar hyperparameters to the GEDI-S2 random forest models implemented in GEE. Reference data were binned by their lat/lon into $0.5^\circ \times 0.5^\circ$ bins, and data in each bin were placed entirely in either the training set or test set using a 80%/20% train/test split. We ran the S2 classifier multiple times using each time a different train/test split and reported the average S2-Local model performance metrics.

4. Results

4.1. GEDI Predictions during Optimal Months

The fraction of GEDI shots classified as a tall crop generally reaches its peak during the latter months of the growing season, such as August, throughout much of the Northern Hemisphere (Figure 7a). This period coincides with the grain filling stage after flowering, when maize has reached its peak height, but before fall months when harvesting begins. We refer to this month of peak tall crop percentage as the optimal month for using GEDI to distinguish tall and short crops. In most regions, the optimal month is stable across space, with most neighboring cells differing by no more than one month. Exceptions to this pattern are evident in cases where tall crops are a very small percentage of the total crop area (e.g., western Canada) or where two or more growing seasons occur throughout the

year (e.g., Brazil and India) (Figure 7b). In both cases, this can lead to two or more months having very similar tall percentages, making the optimal month less stable.

The fraction of GEDI shots from croplands classified as tall crops during these peak months exhibits large spatial variation, with a pattern that coincides closely with known areas of maize production (e.g., eastern United States, eastern China, Brazil). The fraction reaches as high as 75% in eastern China and parts of Central America, where maize dominates the summer growing seasons (Figure 7c). Beyond these regions, shots classified as short crops are generally the dominant class, although a sizable fraction of shots in Africa and Asia were classified as trees (i.e., $RH100 > 10$ m) (Figure 7d). The presence of trees in areas classified as cropland likely reflects both a higher proportion of trees in crop fields in these regions, as well as a lower precision of the ESA cropland map in smallholder systems common in these regions.

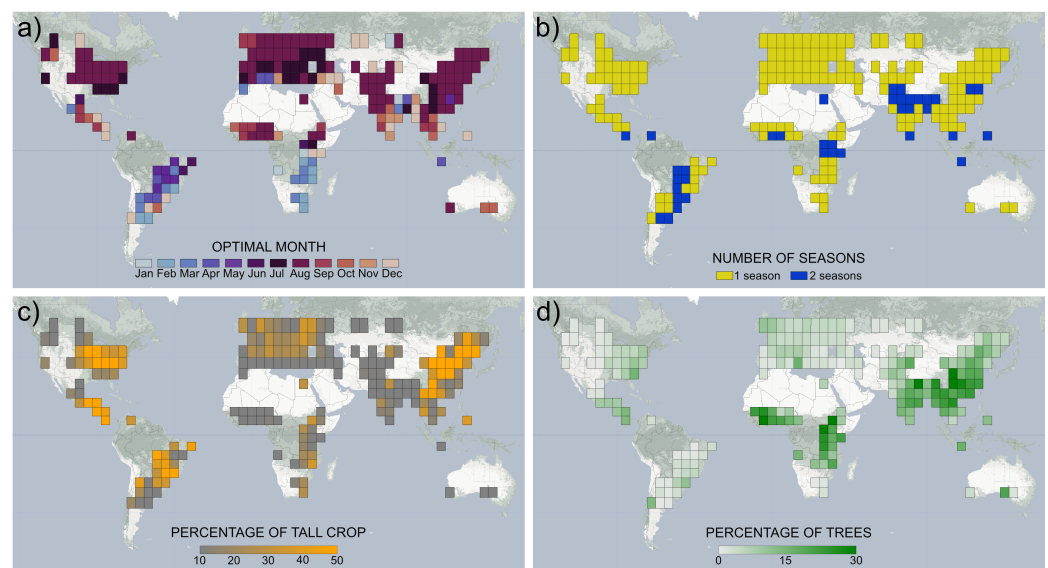


Figure 7. Characteristics of $5^{\circ} \times 5^{\circ}$ grid cells where GEDI-S2 was applied. (a) The optimal month to identify tall crops, defined as the month with the greatest proportion of shots classified as tall, (b) the number of growing seasons per year based on the Anomaly hotspots of Agricultural Production (ASAP) phenology information dataset [37], (c) the percentage of shots classified as tall, and (d) the percentage of shots within the cropland classified as trees.

4.2. GEDI-S2 Model Training

When using the GEDI predicted class (i.e., tall vs. short crop) as labels to train a model based on S2 harmonics, we find that the S2 models are generally able to explain a very high fraction of variability in the GEDI class. In 95% of grid cells (1488 out of 1562 total grid cells), the test accuracy for the model on GEDI shots held out of training was over 0.85. This indicates that tall crops in a region (e.g., maize or sugarcane) are typically distinct enough from other crops in the feature space of the S2 harmonics. In a small number of locations (21 out of 590 cells over the 3 years), the GEDI-S2 training was relatively poor with a test accuracy averaged across all months of less than 0.85. These cells typically occur in regions with high topographic variation, a factor that is known to affect GEDI returns and reduce the accuracy of tree height models based on GEDI [40]. Fortunately, the small number of locations with poor GEDI-S2 training performance indicates that, for most agricultural settings, topographic variation or other sources of error are not a major impediment to using GEDI to identify tall crops. We emphasize that this statement only applies to GEDI shots that have first been filtered for view angle, as including shots with high view angles leads to substantial degradation of performance.

4.3. GEDI-S2 Model Evaluation

The predicted crop class from the GEDI-S2 model is able to closely reproduce reference maps of tall crops in many (but not all) cases (Table 2, Figure 8). Performance was similar for both reference data from ground-based measurements (either as points or field polygons) or from raster maps of crop type developed through a combination of ground and satellite data. In both cases, we compare the GEDI-S2 model performance to the performance of a model trained on the local reference data with the same S2 features as used in the GEDI-S2 model (S2-Local). This comparison identifies the impact of substituting GEDI measurements for ground observations.

Given the large number of validation regions, we discuss the results in terms of clusters of similar behavior rather than discussing each region in detail. The first cluster includes regions where both the local S2 and GEDI-S2 models show high performance, with accuracies typically above 0.9 and F1 scores for both short and tall crops that typically exceed 0.8. Among the regions in this category are North America, Europe, and China (Figure 9a,d). Visual inspection of maps generated by training on ground data vs. GEDI corroborate the strong performance of GEDI (Figure 8). Performance in Brazil appears similarly strong (Figure 8), although because the reference map only identified soybean locations, we cannot calculate total accuracy or F1 scores but only recall for the short crop class.

Table 2. Summary statistics of performance of the GEDI-S2 map. For each dataset we report the S2-Local model and GEDI-S2 model performance in terms of accuracy, precision, recall, F1 and Kappa scores. Precision, recall and F1 scores are presented as (short, tall) pairs. Canada BC and ON stand for British Columbia and Ontario, respectively. US ND and AL stand for North Dakota and Alabama. Brazil BA for Bahia. India UBB for Upper Bhima Basin, TG for Telangana, and MH for Maharashtra.

Region	Year	Accuracy	F1	S2 Local Precision	Recall	K-Score	Accuracy	F1	GEDI-S2 Precision	Recall	K-Score
Austria	2019	0.96	0.97, 0.90	0.97, 0.89	0.97, 0.91	0.87	0.94	0.96, 0.86	0.95, 0.89	0.97, 0.83	0.82
Slovenia	2019	0.9	0.92, 0.87	0.93, 0.87	0.91, 0.88	0.79	0.88	0.90, 0.83	0.88, 0.86	0.91, 0.80	0.73
Germany	2019	0.97	0.98, 0.94	0.98, 0.92	0.97, 0.95	0.92	0.96	0.97, 0.92	0.97, 0.95	0.98, 0.89	0.89
Germany	2020	0.96	0.97, 0.85	0.97, 0.88	0.98, 0.81	0.82	0.94	0.96, 0.83	0.99, 0.74	0.94, 0.94	0.79
Canada (BC)	2019	0.97	0.98, 0.94	0.97, 0.96	0.99, 0.92	0.92	0.97	0.98, 0.92	0.96, 0.98	0.99, 0.88	0.9
Canada (ON)	2019	0.94	0.96, 0.90	0.94, 0.95	0.98, 0.86	0.86	0.93	0.95, 0.89	0.93, 0.95	0.98, 0.84	0.84
Canada (BC)	2020	0.93	0.92, 0.88	0.90, 0.88	0.94, 0.89	0.8	0.9	0.88, 0.85	0.83, 0.95	0.98, 0.78	0.75
Canada (ON)	2020	0.92	0.94, 0.85	0.93, 0.88	0.95, 0.83	0.8	0.88	0.92, 0.80	0.91, 0.81	0.92, 0.79	0.71
Canada (BC)	2021	0.94	0.96, 0.90	0.94, 0.93	0.97, 0.87	0.86	0.94	0.95, 0.90	0.96, 0.88	0.94, 0.92	0.85
US (ND)	2019	0.94	0.96, 0.81	0.94, 0.93	0.99, 0.73	0.78	0.96	0.97, 0.87	0.96, 0.94	0.99, 0.81	0.84
US (ND)	2020	0.95	0.97, 0.65	0.95, 0.91	0.99, 0.52	0.63	0.95	0.97, 0.70	0.97, 0.73	0.97, 0.68	0.68
US (ND)	2021	0.9	0.94, 0.70	0.90, 0.88	0.98, 0.58	0.64	0.91	0.94, 0.75	0.92, 0.86	0.97, 0.66	0.69
US (AL)	2019	0.93	0.95, 0.85	0.94, 0.91	0.97, 0.80	0.81	0.94	0.96, 0.86	0.94, 0.95	0.99, 0.80	0.83
US (AL)	2020	0.95	0.97, 0.89	0.96, 0.92	0.98, 0.86	0.85	0.87	0.91, 0.76	0.96, 0.68	0.86, 0.88	0.68
US (AL)	2021	0.94	0.96, 0.88	0.95, 0.92	0.97, 0.85	0.84	0.94	0.96, 0.88	0.94, 0.96	0.99, 0.82	0.85
Brazil (BA)	2020									0.97	
China	2019	0.91	0.89, 0.92	0.88, 0.92	0.90, 0.91	0.81	0.92	0.91, 0.93	0.90, 0.94	0.92, 0.93	0.84
India (U.B.B.)	2020	0.87	0.85, 0.87	0.84, 0.88	0.87, 0.87	0.73	0.7	0.74, 0.63	0.62, 0.87	0.92, 0.50	0.41
India (TG)	2020	0.96	0.98, 0.16	0.96, 0.38	0.99, 0.10	0.15	0.93	0.96, 0.01	0.96, 0.01	0.97, 0.01	−0.02
India (TG)	2021	0.94	0.97, 0.19	0.94, 0.69	0.99, 0.11	0.18	0.82	0.90, 0.03	0.93, 0.02	0.87, 0.04	−0.06
India (MH)	2020	0.84	0.89, 0.68	0.87, 0.71	0.90, 0.65	0.57	0.6	0.74, 0.15	0.70, 0.19	0.77, 0.13	−0.1
Malawi	2021	0.72	0.82, 0.43	0.76, 0.59	0.89, 0.36	0.27	0.7	0.78, 0.46	0.77, 0.53	0.81, 0.43	0.26
Mali	2019	0.74	0.83, 0.31	0.78, 0.30	0.91, 0.36	0.23	0.73	0.84, 0.07	0.73, 0.18	0.99, 0.05	0.04
Kenya	2021	0.62	0.40, 0.71	0.51, 0.65	0.35, 0.80	0.15	0.42	0.50, 0.30	0.39, 0.55	0.74, 0.21	−0.04

In a second set of regions, the performance of S2-Local and GEDI-S2 models were lower, but were similar to each other (Figure 9b,e). This situation, which occurs in parts of India and Africa, indicate where both approaches struggle to accurately map crop classes, particularly the tall class. One plausible reason for this is that the phenological and spectral differences between different crops are smaller in these regions, so that harmonics-based features are less informative. More sophisticated features, such as based on convolutional neural networks, could help to improve both models but are beyond the scope of the study.

A third and final category of performance includes regions where GEDI-S2 performs notably worse than S2-Local models (Figure 9c,f). In these cases, which we observe primarily in India, the use of labels from GEDI rather than local ground data incurs a loss of accuracy. Here, the problem is unlikely to be either uninformative harmonic features or noisy reference data, both of which would also affect the local S2 model. Instead, our GEDI

model appears to be mislabeling many points, with a tendency in particular to overstate the percentage of short crops. We further analyze the sources of these errors in the discussion Section 5.1.

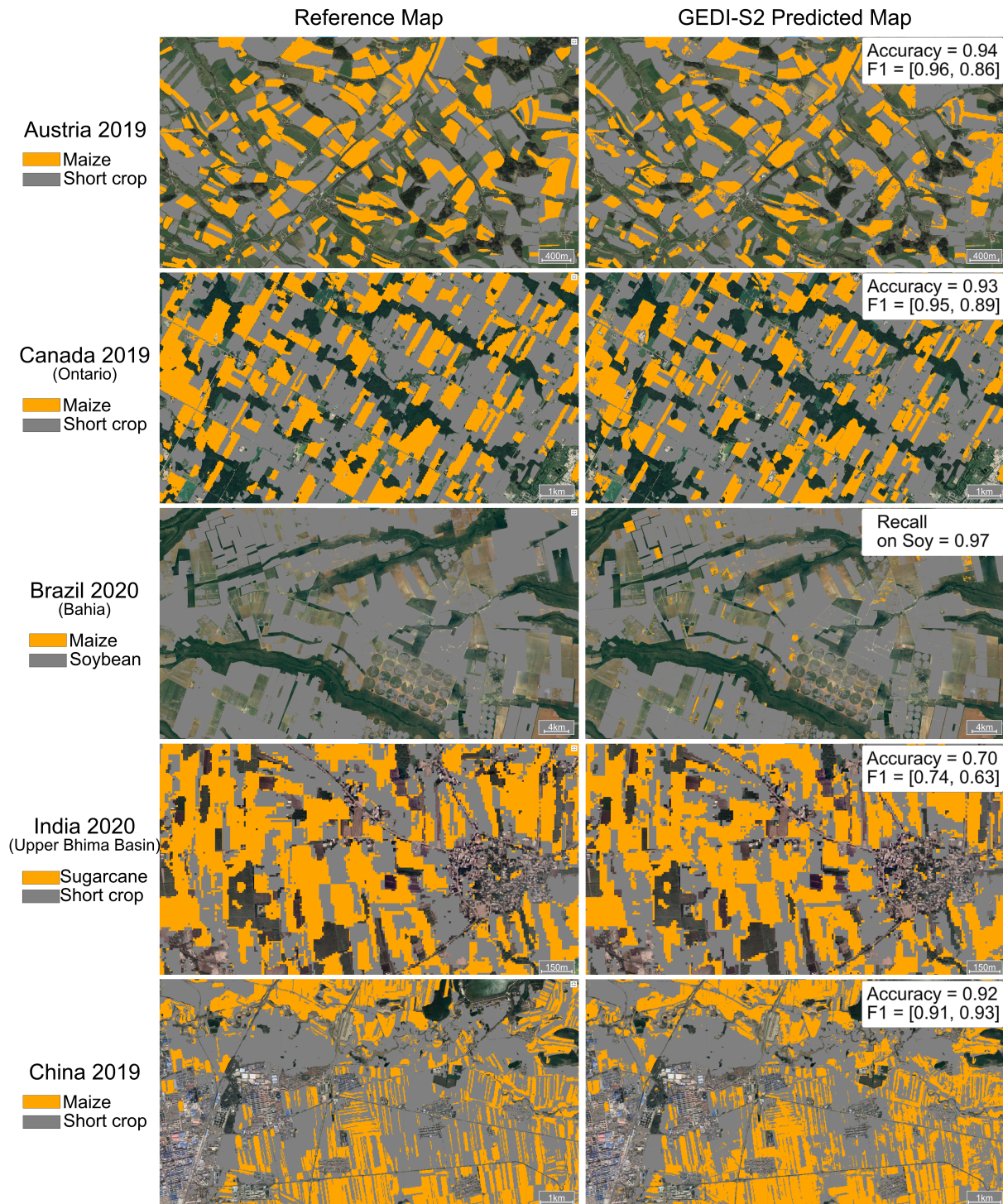


Figure 8. Visual comparison of reference maps (left) and predicted crop classes from the GEDI-S2 model (right). The accuracy and F1 scores in **top-right** of each row refer to the entire region, not just the small areas displayed in the figure. F1 scores are presented as [F1-short, F1-tall]. In Brazil, only recall for short crops is reported since the reference map contained only soybean locations.

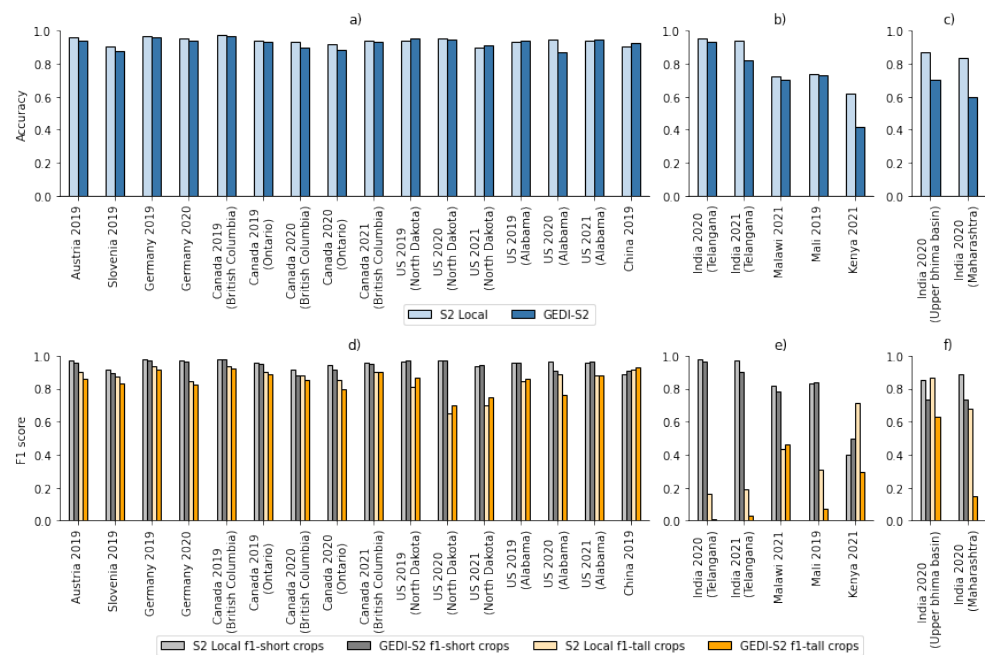


Figure 9. Comparison of GEDI-S2 model performance in terms of accuracy (a–c) and F1 scores (d–f) with the S2-Local model (i.e., a Sentinel 2 model trained on the local reference labels). In most regions, the performance of GEDI-S2 was very close to and occasionally even exceeded that of the S2-Local model (a,d). In some areas, both approaches struggled with tall crops (b,e), while in others GEDI-S2 performed worse than the S2-Local model (c,f).

4.4. The Global Distribution of Tall Crops

The overall pattern of tall crop area estimated by our model is shown for each year in Figure 10. In general, this map corresponds to known areas of maize production, as expected because maize is the most widespread tall crop in the world (Figure 1). However, it also indicates areas with significant sugarcane area (e.g., western Uttar Pradesh in India, the eastern coast of South Africa, and the Philippines) as well as sunflower area (e.g., Romania and Ukraine).

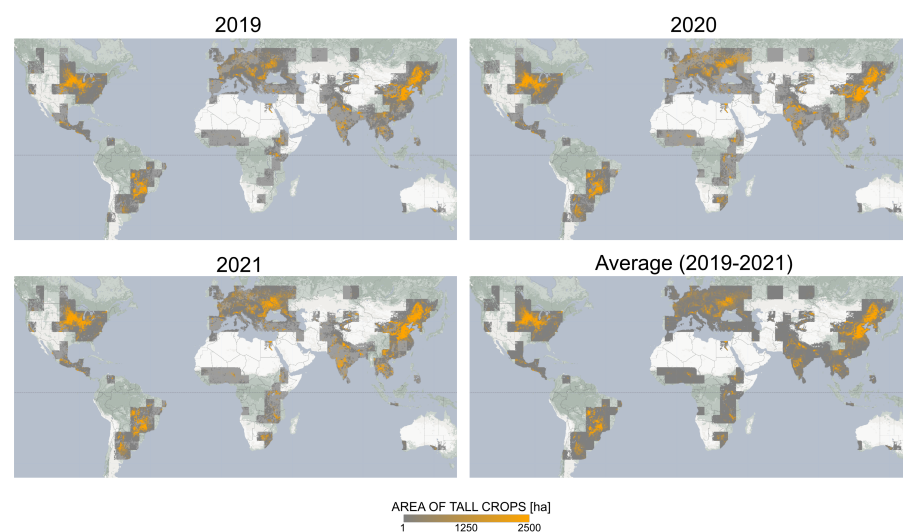


Figure 10. GEDI-S2 global maps gridded at 10 km resolution. For the individual years, cropped areas with peak GCVI above 4 are mapped using the color scale shown, while those with peak GCVI below 4 are mapped with a lighter shade of gray. Low peak GCVI is used as an indicator of where GEDI-S2 is prone to underestimating tall crop area.

5. Discussion

5.1. Sources of Error

Despite the overall encouraging performance of the GEDI-S2 approach, some regions show a clear underestimation of tall crop area. Visual inspection of the GEDI estimates used to train the GEDI-S2 models in these regions indicates that several of the training points are incorrectly labeled as short crops. An example for Canada illustrates this phenomenon well (Figure 11). One likely cause of GEDI falsely predicting a short crop is that the biomass of the tall crop (in this case maize) is significantly less than the typical biomass in regions used to train the GEDI crop height model. This lower biomass results in a greater fraction of photon returns within the 25 m GEDI footprint coming from close to the ground rather than the top of the canopy. One might argue that these maize fields should not, in fact, be considered a tall crop because they have insufficient biomass above the height of typical small crops. However, as the goal of this work is to reliably map crop types regardless of their biomass, it is important that all fields with maize be included in the same class.

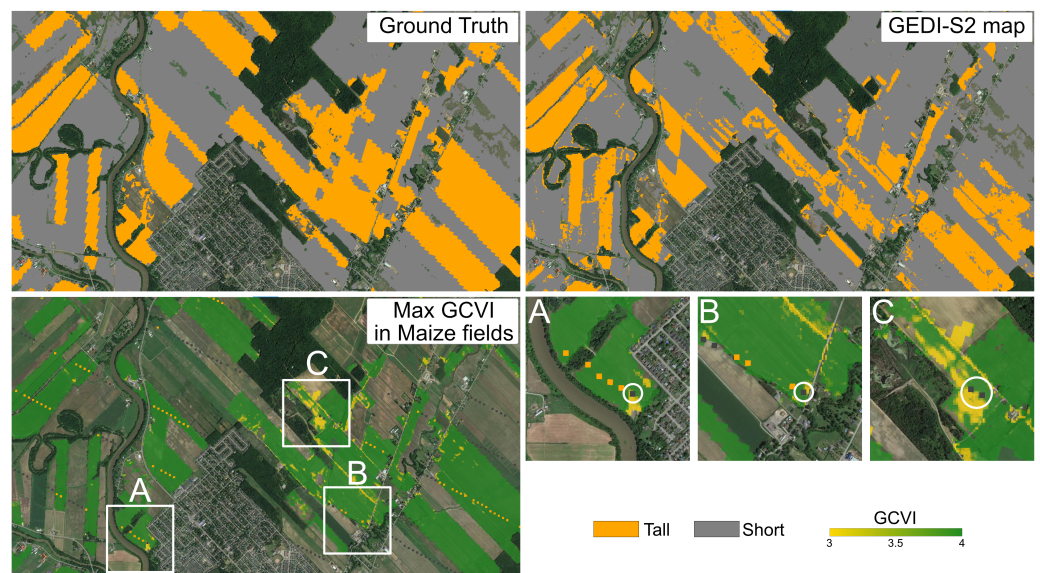


Figure 11. An example of GEDI-S2 errors in low biomass maize fields in Canada. Ground truth (top left) represents the Canada AAFC 2019 map. The corresponding GEDI-S2 2019 predictions (top right) miss several tall fields. S2 peak GCVI in the 2019 growing season (bottom left) indicates that omitted fields often have a peak GCVI below 4 (yellow to light green). A zoom into three examples (panels A, B and C) (lower right) shows the individual GEDI shots and the predicted class (orange = tall, gray = short). Circles indicate examples where the shots were classified as short over low GCVI areas. Since the GEDI-S2 model is then trained on the predicted classes for these shots, it also incorrectly classifies some tall crop area as short.

To further explore the hypothesis that GEDI struggles are related to low biomass of the tall crop, we consider three regions for which maize spans a range of low to high biomass—Kenya, Malawi, and Canada. For each region, we take all GEDI shots that fall onto pixels predicted to be maize by either the local reference data (in the case of Kenya and Canada) or the S2-Local model (in the case of Malawi), and then split these GEDI shots into four groups based on the peak GCVI of the S2 pixel that overlaps the GEDI shot. Peak GCVI is used as a proxy for biomass, given that GCVI has been widely shown to correlate well with maize biomass [25]. We then calculate the fraction of maize GEDI shots predicted to be a tall crop. Consistent with our hypothesis, we find that the GEDI recall is much higher for fields with higher peak GCVI (Figure 12). Recall increases monotonically as the GCVI increases, and recall for pixels with a peak above 4 is at least double that for pixels with a peak VI below 3. Based on this analysis, we consider pixels with peak GCVI

below 4 to be less reliable for GEDI-S2 predictions, and therefore provide a quality flag for these pixels in our final estimates.

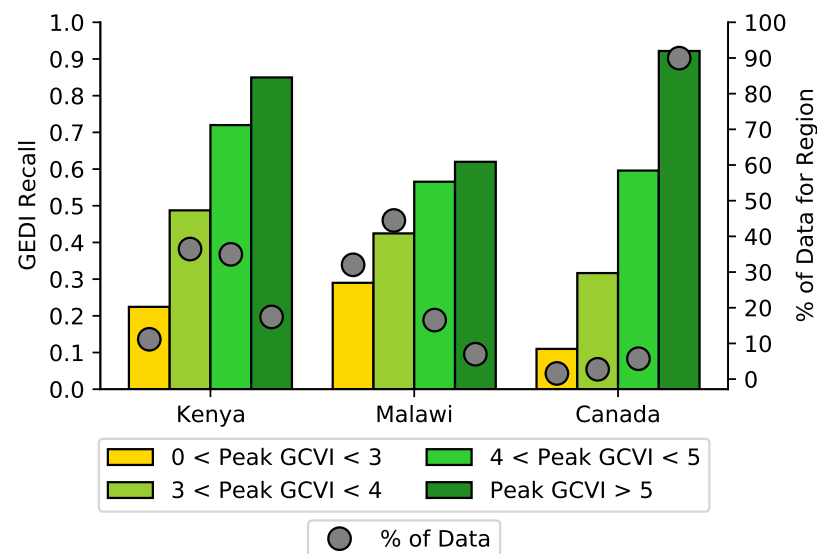


Figure 12. The effect of peak GCVI on GEDI model recall for maize. Peak GCVI, a proxy for biomass, is computed from S2 time series as the maximum GCVI in the optimal month: July for Kenya, February for Malawi and August for Canada. Maize fields in the three regions are defined by their respective reference maps: JRC 2021 Long Rains crop type map for Kenya, 2019 AAFC for Canada, and 2021 S2-Local map for Malawi. Recall is much higher on fields with peak GCVI above 4. In Kenya and Malawi, more than 75% of fields have peak GCVI below 4 in these years.

In general, the regions performing poorly in both the field-level and global evaluation are those for which peak VI is frequently below 4 (Figures 11 and 12). A region such as Canada still performs well overall because the frequency of fields with peak GCVI below 4 is very low (Figure 12). In contrast, more than half of maize fields in Kenya, Malawi, India, and many other regions have peak GCVI below this value, resulting in poor overall performance.

5.2. Future Improvements

The strong agreement with independent reference data in many regions (Table 2, Figure 8) indicates that the GEDI-S2 approach is a promising tool for global crop type mapping. Fully realizing its potential will require progress on several fronts, all of which are beyond the scope of the current paper but for which we anticipate progress is likely. First and foremost is the need to improve performance in areas with lower crop biomass, where the distinction between tall and short crops is reduced. One approach could be to retrain a separate GEDI model for these regions if enough high-quality reference data are available, although the small disparities between tall and short crop returns in these settings make that unlikely to work. Another approach could be to use semi-supervised methods where the GEDI shots predicted for high GCVI fields within low biomass regions are used as high quality labels for fine tuning [43,44]. These semi-supervised approaches may also benefit from using S2 features that are less sensitive to peak biomass [14,45].

A second area for future work is to extend the GEDI-S2 approach to map multiple seasons in regions where more than one crop per year is typically grown. In the current work we focused only on the season with the highest proportion of tall crops, but many areas have two seasons that each have a significant fraction of tall crops (e.g., Eastern Africa, Northern India). Performing the GEDI-S2 training for multiple seasons would be a straightforward extension of the current work, with the main requirement likely being a shift in the window over which the S2 harmonic features are calculated.

A third extension could be to further discriminate among crops beyond the two categories of tall and short. With additional training data, it may be possible for GEDI returns to distinguish very short crops, such as legumes, from slightly taller crops such as rapeseed or cotton. For example, many areas in India are dominated in the monsoon season by rice, cotton, and sugarcane [32], each of which would fall into separate height classes. Even if the GEDI predictions from this approach are noisy, they could be effective labels for training a GEDI-S2 model.

Alternatively, refining the crop classification could utilize complementary features that are unrelated to height. For example, distinguishing short winter or spring crops such as winter and barley from short summer crops is fairly simple based on phenological differences in optical or radar data [12,46], and rice can be accurately distinguished from other crops based on flooding patterns detected in radar imagery [47,48]. Many agricultural landscapes will possess only one other main crop beyond wheat, maize, and rice, and so the ability to map these three could effectively map all major crop types (if the fourth crop is estimated as fields not in one of these three crops). In much of North and South America, for example, maize and soybeans are by far the most common summer crops, and so a large fraction of non-maize area is in soybean (Figure 2). Further work is needed to test how far the tall vs. short crop distinction can help to solve the more general problem of mapping crop species.

Finally, we note that the recent incorporation of additional GEDI data in GEE should improve performance, in at least two ways. The most recent GEDI data from after 2021 is not affected by the orbit resonance problem, which provides an opportunity to reduce grid cell size below the 5 degree spacing used here. By doing so, one could train more localized models. In addition, view angle information is now available globally at the level of individual shots, allowing one to implement a more precise filter for shots with a low view angle.

6. Conclusions

In this study we sought to test the general applicability of an approach that uses GEDI returns to train local crop type mapping models that use S2 data as input. Our general conclusion is that the approach exhibits considerable promise for advancing crop mapping. Tall and short crops were mapped with high accuracy in the majority of maize production systems, including most of the Americas, Europe, and East Asia. Specifically, we showed that GEDI returns can first be classified into tall and short crops, that the frequency of tall crops over time can be used to identify the appropriate months for S2 training (when tall crops are at their peak height), and that S2 models trained on these GEDI shots can accurately predict the GEDI crop height class in nearly all regions. Only in rare cases, such as areas with high topographic variation, did S2 features fail to predict the GEDI crop height class. We then showed that the predictions from the GEDI-S2 agree remarkably well with independent reference data at the field scale.

At the same time, we uncovered cases where the current implementation of GEDI-S2 is problematic. The most common cause for low accuracy appears to be low biomass of tall crops, which occurs frequently in Africa and South Asia. In these regions, the GEDI classification model consistently underestimated the frequency of tall crops. S2 models trained on these shots then inherit this under-prediction of tall crop area. Although this is a notable limitation of the current approach—particularly because these regions are among those with the most limited ground data, and thus where an approach that relied on GEDI for training would be most valuable—we anticipate that future work can greatly improve the performance in low biomass regions. Progress seems most likely for semi-supervised methods that can leverage the fact that even low biomass areas typically have a significant number of fields with high biomass that are accurately captured by GEDI.

Author Contributions: Conceptualization, S.D.T., S.W. and D.B.L.; methodology, S.D.T., S.W. and D.B.L.; validation, S.D.T. and V.V.; formal analysis, S.D.T.; resources, N.G. and R.S.; data curation, S.D.T.; writing—original draft preparation, S.D.T. and D.B.L.; writing—review and editing, S.D.T., S.W. and D.B.L.; visualization, S.D.T. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the NASA Harvest Consortium (NASA Applied Sciences Grant No. 80NSSC17K0652, sub-award 54308-Z6059203 to DBL). SW was supported by the Ciriacy-Wantrup Postdoctoral Fellowship at the University of California, Berkeley.

Data Availability Statement: The tall/short 10 m global maps for 2019–2021 generated in this study are available publicly as assets on Google Earth Engine. The code used to train the GEDI model, generate the grid-cells, compute harmonics features, and train the GEDI-S2 models is available publicly as Google Earth Engine scripts. Links to scripts, data, and an interactive map can be found on GitHub repository at: https://github.com/LobellLab/GEDI-S2_tall_short_global_map.

Acknowledgments: We thank Rick Brandenburg for sharing the Malawi crop reference data, Marcel Schwieder for access to the Germany crop type maps, and the Google Earth Engine team for making large-scale computational resources available to researchers.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Bégué, A.; Arvor, D.; Bellon, B.; Betbeder, J.; De Aballeyra, D.; Ferraz, R.P.D.; Lebourgeois, V.; Lelong, C.; Simões, M.; Verón, S.R. Remote sensing and cropping practices: A review. *Remote Sens.* **2018**, *10*, 99. [CrossRef]
2. Kim, K.H.; Doi, Y.; Ramankutty, N.; Iizumi, T. A review of global gridded cropping system data products. *Environ. Res. Lett.* **2021**, *16*, 093005. [CrossRef]
3. Nakalembe, C.; Becker-Reshef, I.; Bonifacio, R.; Hu, G.; Humber, M.L.; Justice, C.J.; Keniston, J.; Mwangi, K.; Rembold, F.; Shukla, S.; et al. A review of satellite-based global agricultural monitoring systems available for Africa. *Glob. Food Secur.* **2021**, *29*, 100543. [CrossRef]
4. Boryan, C.; Yang, Z.; Mueller, R.; Craig, M. Monitoring US agriculture: The US Department of Agriculture, National Agricultural Statistics Service, Cropland Data Layer Program. *Geocarto Int.* **2011**, *26*, 341–358. [CrossRef]
5. Agriculture and Agri-Food Canada. Annual Crop Inventory. Available online: <https://open.canada.ca/data/en/dataset/ba2645d5-4458-414d-b196-6303ac06c1c9> (accessed on 1 January 2021).
6. Agence de Services et de Paiement. Registre Parcellaire Graphique (RPG): Contours des Parcelles et Îlots Culturels et Leur Groupe de Cultures Majoritaire. 2019. Available online: <https://www.data.gouv.fr/en/datasets/registre-parcellaire-graphique-rpg-contours-des-parcelles-et-ilots-culturels-et-leur-groupe-de-cultures-majoritaire/> (accessed on 1 January 2021).
7. Potapov, P.; Turubanova, S.; Hansen, M.C.; Tyukavina, A.; Zalles, V.; Khan, A.; Song, X.P.; Pickens, A.; Shen, Q.; Cortez, J. Global maps of cropland extent and change show accelerated cropland expansion in the twenty-first century. *Nat. Food* **2022**, *3*, 19–28. [CrossRef]
8. Lobell, D.B.; Asner, G.P.; Ortiz-Monasterio, J.I.; Benning, T.L. Remote sensing of regional crop production in the Yaqui Valley, Mexico: Estimates and uncertainties. *Agric. Ecosyst. Environ.* **2003**, *94*, 205–220. [CrossRef]
9. You, N.; Dong, J.; Huang, J.; Du, G.; Zhang, G.; He, Y.; Yang, T.; Di, Y.; Xiao, X. The 10-m crop type maps in Northeast China during 2017–2019. *Sci. Data* **2021**, *8*, 41. [CrossRef]
10. Han, J.; Zhang, Z.; Luo, Y.; Cao, J.; Zhang, L.; Cheng, F.; Zhuang, H.; Zhang, J. AsiaRiceMap10m: High-resolution annual paddy rice maps for Southeast and Northeast Asia from 2017 to 2019. *Earth Syst. Sci. Data Discuss* **2021**, *211*, 1–27.
11. Song, X.P.; Hansen, M.C.; Potapov, P.; Adusei, B.; Pickering, J.; Adami, M.; Lima, A.; Zalles, V.; Stehman, S.V.; Di Bella, C.M.; et al. Massive soybean expansion in South America since 2000 and implications for conservation. *Nat. Sustain.* **2021**, *4*, 784–792. [CrossRef]
12. Blickensdörfer, L.; Schwieder, M.; Pflugmacher, D.; Nendel, C.; Erasmi, S.; Hostert, P. Mapping of crop types and crop sequences with combined time series of Sentinel-1, Sentinel-2 and Landsat 8 data for Germany. *Remote Sens. Environ.* **2022**, *269*, 112831. [CrossRef]
13. Kluger, D.M.; Wang, S.; Lobell, D.B. Two shifts for crop mapping: Leveraging aggregate crop statistics to improve satellite-based maps in new regions. *Remote Sens. Environ.* **2021**, *262*, 112488. [CrossRef]
14. Lin, C.; Zhong, L.; Song, X.P.; Dong, J.; Lobell, D.B.; Jin, Z. Early-and in-season crop type mapping without current-year ground truth: Generating labels from historical information via a topology-based approach. *Remote Sens. Environ.* **2022**, *274*, 112994. [CrossRef]
15. Luo, Y.; Zhang, Z.; Zhang, L.; Han, J.; Cao, J.; Zhang, J. Developing High-Resolution Crop Maps for Major Crops in the European Union Based on Transductive Transfer Learning and Limited Ground Data. *Remote Sens.* **2022**, *14*, 1809. [CrossRef]

16. Di Tommaso, S.; Wang, S.; Lobell, D.B. Combining GEDI and Sentinel-2 for wall-to-wall mapping of tall and short crops. *Environ. Res. Lett.* **2021**, *16*, 125002. [\[CrossRef\]](#)
17. Dubayah, R.; Blair, J.B.; Goetz, S.; Fatoyinbo, L.; Hansen, M.; Healey, S.; Hofton, M.; Hurtt, G.; Kellner, J.; Luthcke, S.; et al. The Global Ecosystem Dynamics Investigation: High-resolution laser ranging of the Earth's forests and topography. *Sci. Remote Sens.* **2020**, *1*, 100002. [\[CrossRef\]](#)
18. International Food Policy Research Institute. *Global Spatially-Disaggregated Crop Production Statistics Data for 2010*; Version 2.0; Harvard Library: Cambridge, MA, USA, 2019. [\[CrossRef\]](#)
19. Zanaga, D.; Van De Kerchove, R.; De Keersmaecker, W.; Souverijns, N.; Brockmann, C.; Quast, R.; Wevers, J.; Grosu, A.; Paccini, A.; Vergnaud, S.; et al. *ESA WorldCover 10 m 2020 v100*; Zenodo: Geneva, Switzerland, 2021.
20. Karra, K.; Kontgis, C.; Statman-Weil, Z.; Mazzariello, J.C.; Mathis, M.; Brumby, S.P. Global land use/land cover with Sentinel 2 and deep learning. In Proceedings of the 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, Brussels, Belgium, 11–16 July 2021; pp. 4704–4707.
21. Gorelick, N.; Hancher, M.; Dixon, M.; Ilyushchenko, S.; Thau, D.; Moore, R. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* **2017**, *202*, 18–27. [\[CrossRef\]](#)
22. Roy, S.; Swetnam, T.; Robitaille, A.; Trochim, E.; Pasquarella, V. *Samapriya/Awesome—Gee—Community—Datasets: Community Catalog (1.0.1)*; Zenodo: Geneva, Switzerland, 2022. [\[CrossRef\]](#)
23. Fayad, I.; Baghdadi, N.; Bailly, J.S.; Frappart, F.; Zribi, M. Analysis of GEDI elevation data accuracy for inland waterbodies altimetry. *Remote Sens.* **2020**, *12*, 2714. [\[CrossRef\]](#)
24. Healey, S.P.; Yang, Z.; Gorelick, N.; Ilyushchenko, S. Highly local model calibration with a new GEDI LiDAR asset on Google Earth Engine reduces landsat forest height signal saturation. *Remote Sens.* **2020**, *12*, 2840. [\[CrossRef\]](#)
25. Gitelson, A.A.; Vina, A.; Ciganda, V.; Rundquist, D.C.; Arkebauer, T.J. Remote estimation of canopy chlorophyll content in crops. *Geophys. Res. Lett.* **2005**, *32*, L08403. [\[CrossRef\]](#)
26. Wang, S.; Azzari, G.; Lobell, D.B. Crop type mapping without field-level labels: Random forest transfer and unsupervised clustering techniques. *Remote Sens. Environ.* **2019**, *222*, 303–317. [\[CrossRef\]](#)
27. Song, X.P.; Huang, W.; Hansen, M.C.; Potapov, P. An evaluation of Landsat, Sentinel-2, Sentinel-1 and MODIS data for crop type mapping. *Sci. Remote Sens.* **2021**, *3*, 100018. [\[CrossRef\]](#)
28. Schneider, M.; Broszeit, A.; Körner, M. Eurocrops: A pan-european dataset for time series crop type classification. *arXiv* **2021**, arXiv:2106.08151.
29. Agriculture and Agri-Food Canada. Annual Crop Inventory Ground Truth Data. 2021. Available online: <https://open.canada.ca/data/en/dataset/503a3113-e435-49f4-850c-d70056788632> (accessed on 15 November 2022).
30. Tseng, G.; Zvonkov, I.; Nakalembe, C.; Kerner, H. CropHarvest: A global dataset for crop-type classification. In Proceedings of the Neural Information Processing Systems Track on Datasets and Benchmarks, Paris, France, 8 December 2021; Volume 1.
31. European Commission, Joint Research Centre (JRC). Kenya AOI. European Commission, Joint Research Centre (JRC) [Dataset]. 2021. Available online: <https://data.jrc.ec.europa.eu/dataset/5b6245d3-e561-4f6c-8c09-627888063d11> (accessed on 15 November 2022).
32. Wang, S.; Di Tommaso, S.; Faulkner, J.; Friedel, T.; Kennepohl, A.; Strey, R.; Lobell, D.B. Mapping Crop Types in Southeast India with Smartphone Crowdsourcing and Deep Learning. *Remote Sens.* **2020**, *12*, 2957. [\[CrossRef\]](#)
33. Blickensdörfer, L.; Schwieder, M.; Pflugmacher, D.; Nendel, C.; Erasmi, S.; Hostert, P. *National-Scale Crop Type Maps for Germany from Combined Time Series of Sentinel-1, Sentinel-2 and Landsat 8 Data (2017, 2018 and 2019)*; Zenodo: Geneva, Switzerland, 2021. [\[CrossRef\]](#)
34. Schwieder, M.; Erasmi, S.; Nendel, C.; Hostert, P. *National-Scale Crop Type Maps for Germany from Combined Time Series of Sentinel-1, Sentinel-2 and Landsat 8 Data (2020)*; Zenodo: Geneva, Switzerland, 2022. [\[CrossRef\]](#)
35. Lee, J.Y.; Wang, S.; Figueroa, A.J.; Strey, R.; Lobell, D.B.; Naylor, R.L.; Gorelick, S.M. Mapping Sugarcane in Central India with Smartphone Crowdsourcing. *Remote Sens.* **2022**, *14*, 703. [\[CrossRef\]](#)
36. Rembold, F.; Meroni, M.; Urbano, F.; Csak, G.; Kerdiles, H.; Perez-Hoyos, A.; Lemoine, G.; Leo, O.; Negre, T. ASAP: A new global early warning system to detect anomaly hot spots of agricultural production for food security analysis. *Agric. Syst.* **2019**, *168*, 247–257. [\[CrossRef\]](#) [\[PubMed\]](#)
37. European Commission, Joint Research Centre (JRC). Global Land Surface Phenology—Number of Growing Seasons [Dataset]. 2018. Available online: <http://data.europa.eu/89h/jrc-10112-10008> (accessed on 15 November 2022).
38. Farr, T.G.; Rosen, P.A.; Caro, E.; Crippen, R.; Duren, R.; Hensley, S.; Kobrick, M.; Paller, M.; Rodriguez, E.; Roth, L.; et al. The shuttle radar topography mission. *Rev. Geophys.* **2007**, *45*, RG2004. [\[CrossRef\]](#)
39. Fayad, I.; Baghdadi, N.; Frappart, F. Comparative Analysis of GEDI's Elevation Accuracy from the First and Second Data Product Releases over Inland Waterbodies. *Remote Sens.* **2022**, *14*, 340. [\[CrossRef\]](#)
40. Fayad, I.; Baghdadi, N.; Alcarde Alvares, C.; Stape, J.L.; Bailly, J.S.; Scolforo, H.F.; Cegatta, I.R.; Zribi, M.; Le Maire, G. Terrain slope effect on forest height and wood volume estimation from GEDI data. *Remote Sens.* **2021**, *13*, 2136. [\[CrossRef\]](#)
41. Potapov, P.; Li, X.; Hernandez-Serna, A.; Tyukavina, A.; Hansen, M.C.; Kommareddy, A.; Pickens, A.; Turubanova, S.; Tang, H.; Silva, C.E.; et al. Mapping global forest canopy height through integration of GEDI and Landsat data. *Remote Sens. Environ.* **2021**, *253*, 112165. [\[CrossRef\]](#)
42. Grandini, M.; Bagli, E.; Visani, G. Metrics for multi-class classification: An overview. *arXiv* **2020**, arXiv:2008.05756.

43. Weinstein, B.G.; Marconi, S.; Bohlman, S.; Zare, A.; White, E. Individual tree-crown detection in RGB imagery using semi-supervised deep learning neural networks. *Remote Sens.* **2019**, *11*, 1309. [[CrossRef](#)]
44. Wu, H.; Prasad, S. Semi-supervised deep learning using pseudo labels for hyperspectral image classification. *IEEE Trans. Image Process.* **2017**, *27*, 1259–1270. [[CrossRef](#)] [[PubMed](#)]
45. Qiu, B.; Huang, Y.; Chen, C.; Tang, Z.; Zou, F. Mapping spatiotemporal dynamics of maize in China from 2005 to 2017 through designing leaf moisture based indicator from Normalized Multi-band Drought Index. *Comput. Electron. Agric.* **2018**, *153*, 82–93. [[CrossRef](#)]
46. Veloso, A.; Mermoz, S.; Bouvet, A.; Le Toan, T.; Planells, M.; Dejoux, J.F.; Ceschia, E. Understanding the temporal behavior of crops using Sentinel-1 and Sentinel-2-like data for agricultural applications. *Remote Sens. Environ.* **2017**, *199*, 415–426. [[CrossRef](#)]
47. Nelson, A.; Setiyono, T.; Rala, A.B.; Quicho, E.D.; Raviz, J.V.; Abonete, P.J.; Maunahan, A.A.; Garcia, C.A.; Bhatti, H.Z.M.; Villano, L.S.; et al. Towards an operational SAR-based rice monitoring system in Asia: Examples from 13 demonstration sites across Asia in the RIICE project. *Remote Sens.* **2014**, *6*, 10773–10812. [[CrossRef](#)]
48. Singha, M.; Dong, J.; Zhang, G.; Xiao, X. High resolution paddy rice maps in cloud-prone Bangladesh and Northeast India using Sentinel-1 data. *Sci. Data* **2019**, *6*, 26. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.