

Application of Remote Sensing & GIS in Agriculture of Nalgonda District, Telangana State

Balduri Harish, A. Balakishan

¹Department of Geography, University College of Science, Osmania University, Hyderabad 500007, India

²Professor, Department of Geography, University College of Science, Osmania University, Hyderabad 500007, India

Abstract: Remote sensing and GIS can also be used very effectively in land use / land cover analysis as well as damage assessment because of drought, floods and other extreme weather events. Information on meteorology and vegetation are the two major important inputs into agricultural meteorology. Applications of remote sensing technologies are important and effective method to identify pest-infested and disease. It is one of the effective tools for assessing and monitoring the water resources. This article provides an overview of some of the recent research in agriculture involving remotesensing and GIS. Attention focuses on application of remote sensing and GIS specially in agriculture including geography, land surveying, most Earth Science disciplines, parent child relationship, unique identification, attributes, technical parameters, 2D/3D view and any other requirement customized. These advances have been made over recent years and foundations for future research established and can be efficiently used in Agriculture for better results.

Keywords: Remotesensing and GIS, Agriculture Meteorology, Technical Parameters, Earth sciences

1. Introduction

Agriculture plays a vital role in every nation economy. It represents a substantial trading industry for an economically strong country. Production of food in a cost-effective manner is the essential goal of every farmer, large-scale farm manager and regional agricultural agency. Remote sensing and Geographic information system used to analyze and visualize agricultural environments has proved to be very beneficial to farming community as well as industry. It plays great role in agriculture throughout the world by helping farmers in increasing production, reducing costs and managing their land more efficiently. Geographic information systems (GIS) has been widely applied and been recognized as effective and powerful tool in detecting land cover and land use change [1]. Using remote sensing and GIS are important to understand the health of crop, extent of infestation, potential yield and soil conditions. It applied to explore agricultural applications such as crop identification, area estimation, crop condition assessment, soil moisture estimation, yield estimation, agriculture water management, agro meteorological etc.

Applications of remote sensing in agriculture including major important things such as; biomass and yield estimation, vegetation vigor and drought stress monitoring, assessment of crop phenological development, crop acreage estimation and cropland mapping, mapping of disturbances and land use land cover changes in addition to precision agriculture and irrigation management [2]. GIS based mapping application can help to identify location of crops growing across the country and to adapt different variables, monitor the health of individual crops, estimate yields from a given field, and maximize crop production. By using land-use and primary food crop statistics, along with data

collected by different tools including mobile devices able to identify areas in need and underlying causes of food insecurity, GIS is an instrumental in the efforts to end global hunger and it is an integral part of automated field operations.

Using data collected from remote sensors, and from sensors mounted directly on farm machinery, farmers have improved decision-making capabilities for planning their cultivation to maximize yields. Previous crop yields, terrain specifics, organic matter content, pH, moisture, and nutrient levels of the soil all aid in proper preparation for precise farming. Combine harvesters equipped with GPS tracking units can measure crop yields along with crop quality values like plant water content and chlorophyll levels in real time and at the exact location in the field from which they are harvested. Rapidly emerging remote sensing and geospatial technology can play vital role for crop growth monitoring, identification and management of different types of stresses, regional yield estimations, to sustain the natural resources and agricultural productivity [3].

2. Study Region

Nalgonda is situated in the central stretch of the eastern seaboard of the Indian Peninsula., Nalgonda district is a district in the Telangana state of India. The district was situated between 78° 40' and 80° 05'E, of the eastern longitudes and 16° 25' and 17° 60'N, of northern latitudes. It has a Area (km²) 2449.79, population of 1631399 and 31 mandals. The district shares boundaries with Suryapet district, Rangareddy, Yadadri, Nagarkurnool districts and with State Boundary. The rivers Krishna, Musi River, Aleru, Peddavagu, and Dindiflow through the Nalgonda district.

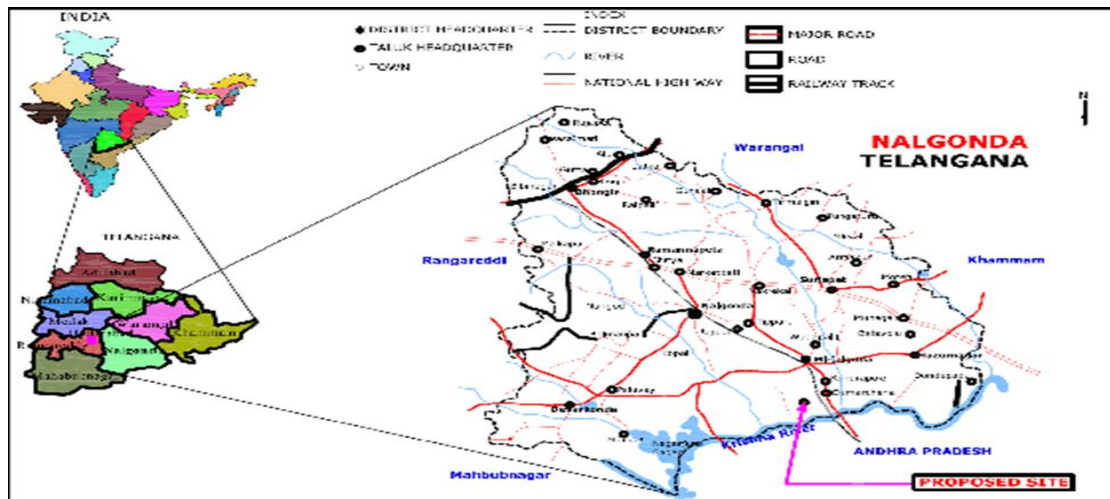


Figure 1: Location Map of Study area Nalgonda district

Telangana borders Andhra Pradesh to the northwest, and until 2014. Today it is the eleventh largest Indian state (112,077 sq km). Situated on the Deccan Plateau in central-south India, its climate is semi-arid and drier than that of Nalgonda, with district average annual precipitation more variable from 500 mm to 1200 mm and temperatures from 15 °C to 45 °C. Forty-three percent of the state is cropped, and 24% is forest [40]; cotton, rice, and maize are major kharif crops, while rice, maize, and peanut comprise the major rabi crops [9]. The average operational holding size in Telangana was 1.00 ha in 2015–2016 [28].

3. Datasets

3.1 Plantix User Submissions

Plantix is a free Android application created by Progressive Environmental and Agricultural Technologies (PEAT) in 2015 to help farmers identify pests, diseases, and nutrient deficiencies using a mobile phone camera and image recognition software. The user—usually a farmer, sometimes a hired plant expert—takes a photo of his or her crop with a mobile phone and uploads the photo to PEAT servers for a diagnosis of plant health. The photo is then run through a deep neural network, which returns predicted plant ailments. This information, along with corresponding treatments, are sent back to the user's Plantix app.

Between 1 January 2017 and 1 January 2019, the Plantix app received 8.6 million geolocated submissions from India, 1.8 million of which were in Nalgonda and Telangana. When a photo was taken via Plantix, the time of capture and the location of the phone were recorded. Figure 1 shows that most of the submissions were logged between September and November, which is during the harvest of the kharif season. While submissions were not tagged with a crop type by the farmer, crop scientists at PEAT assigned a crop type label to a subset of the submissions based on the uploaded photos, and a deep convolutional neural network (Plantix-DNN) was trained on expert labels to predict the crop type of all million submissions.

Details of how Plantix submissions were filtered and used to construct training, validation, and test sets for crop type classification are described in Section 4.1.

3.2 Sentinel-2 Time Series

Sentinel-2 was chosen for its high spatial resolution and public availability, and prior work has shown that optical features can be used to distinguish crop types [14,18,27,41]. Sentinel-2A was first launched by the European Space Agency (ESA) in June 2015 as part of the European Union's Copernicus Programme for Earth observation, and captures high-resolution (10–60 m) optical imagery to serve a wide range of scientific applications on land and in coastal waters. Since March 2017, with the launch of Sentinel-2B, images have been collected on a 5-day cycle. ESA distributes a top of atmosphere reflectance product (Level-1C) for Sentinel-2, while a higher-level surface reflectance product (Level-2A) can be derived using a toolbox provided by ESA [42]. At the time of this work, pre-generated Level-2A imagery was not available for download from either ESA or Google Earth Engine (GEE) over India before December 2018. Since the ESA Toolbox was also not available in GEE, one would have to compute the Level-2A product and ingest it into GEE in order to obtain time series of surface reflectance. Such an approach is hugely expensive computationally and storage-wise, and does not scale well to a study region as large as Nalgonda and Telangana. Furthermore, prior work showed that land cover classification using top-of-atmosphere reflectance is comparable to using surface reflectance, since relative spectral differences drive classification [43,44]. For these reasons, we used the Level-1C product, with the recognition that this imperfect input will still place some limits on the performance of a crop classifier.

Using Google Earth Engine [45], we exported all Sentinel-2 Level-1C images at each submission coordinate for the corresponding crop year, where a crop year is defined to be from 1 April of one year to 31 March of the next. For example, a Plantix submission from 1 September 2017—during the kharif season—generates a time series of Sentinel-2 readings from 1 April 2017 to 31 March 2018, which encompasses the fall 2017 kharif season and winter 2017–2018 rabi season. 1 April was chosen as the cutoff date to avoid truncating early kharif or late rabi satellite data that could be relevant for crop type classification. All spectral bands were sampled at 10 m ground resolution; the

20 m and 60 m bands were resampled using the GEE default nearest neighbor algorithm. In addition to the 13 spectral bands, we also computed the green chlorophyll vegetation index ($GCVI = NIR/green - 1$) [46] as previous work has shown GCVI to correlate well with leaf area index [47] and be a strong feature for crop type classification [18, 48]. Figure 2 visualizes GCVI time series of the 10 crop types and shows the high levels of noise due to clouds in Sentinel-2 imagery.

Across the dataset of Plantix submissions, the median number of Sentinel-2 images for the 2017–2018 crop year was 42, with minimum and maximum image counts of 36 and 143. In 2018, the median image count was 71, with minimum and maximum of 67 and 143. Of these images, many are affected by clouds and cloud shadows, especially during the monsoon season; since the Sentinel-2 Level 1-C cloud mask has a high omission error at the time of writing [49], we describe methods to minimize the impact of clouds in Section 4.3.2. Although we will show that it is possible to make progress in distinguishing crop types in the presence of cloudy imagery, the lack of high-quality cloud mask remains a major challenge to mapping crop types in smallholder systems, and we expect the improvement of such masks to enable better classification performance in the future.

3.3 Sentinel-1 Time Series

The Sentinel-1 constellation is composed of two satellites, Sentinel-1A and Sentinel-1B, launched April 2014 and April 2016, respectively. The satellites carry a C-band synthetic-aperture radar (SAR) instrument that captures 5–40 m resolution imagery with a revisit period of 12 days. We used the Interferometric Wide swatch mode, which acquires images with dual polarization (vertical transmit, vertical receive (VV) and vertical transmit, horizontal receive (VH)). In addition to using backscatter coefficients as features, we computed their ratio and difference ($RATIO = VH/VV = VH_{dB} - VV_{dB}$ and $DIFF = VV - VH$) based on previous work that indicated the suitability of such transformations for agricultural applications [18,50,51]. Unlike optical imagery, radar is not affected by weather conditions and can monitor the Earth's surface through clouds, making it a useful complement to optical imagery during the wet season. The median number of Sentinel-1 images available was 30 in both 2017 and 2018, with minimum and maximum of 21 and 78.

Sentinel-1 Ground Range Detected (GRD) scenes are available in GEE, which processes the imagery using ESA's

Sentinel-1 Toolbox to reduce noise and standardize bands at 10 m spatial resolution. We note that after this processing Sentinel-1 imagery still contains speckle-noise, which is caused by backscatter interference. While speckle is highly disruptive for interpreting individual images taken at one point in time [52], its effect on a SAR time series over the course of a year is dampened by the feature extraction methods described in Section 4.3.2 and Section 4.3.3. We used GEE to export Sentinel-1 GRD observations at each submission coordinate for the corresponding crop year, where a crop year is defined to be from 1 April of one year to 31 March of the next (Figure A2).

3.4 DigitalGlobe Static Satellite Imagery

Due to Android phone location inaccuracy and users taking photos of crops while not standing in their fields (e.g., from the road next to their field or in an urban area with internet access), we employed very high-resolution satellite imagery to filter for points within a crop field. A single DigitalGlobe image was downloaded using the company's Web Map Service application for each submission location, for all 72,000 kharif season submissions (see Section 4.1 for how we arrived at 72,000). Images have 0.3 m ground sample distance, are RGB, and span 256×256 pixels (76.8 m \times 76.8 m) around the submission location. This ground resolution is high enough for humans to visually determine whether any pixel in the image belongs to a crop field or not.

Since we used time series of Sentinel imagery for crop type classification, we were interested in whether the Sentinel pixel covering each submission location is within a crop field. We therefore drew a 10 m \times 10 m square centered within each DigitalGlobe image to approximate the bounds of the corresponding Sentinel-2 pixel.

3.5 Sub-National Crop Area Statistics

District-wise, season-wise crop production statistics were available from 1997 to 2016 on India's Crop Production Statistics Information System and Open Government Data (OGD) Platform [9,53]. We compared the kharif crop area data in and 30 districts of Telangana against our aggregated crop type predictions. Note that the number of districts in Telangana increased from 10 to 33 upon redistricting in 2016. In the latest government crop statistics (2016–2017), data were only available for 30 of the 33 districts.

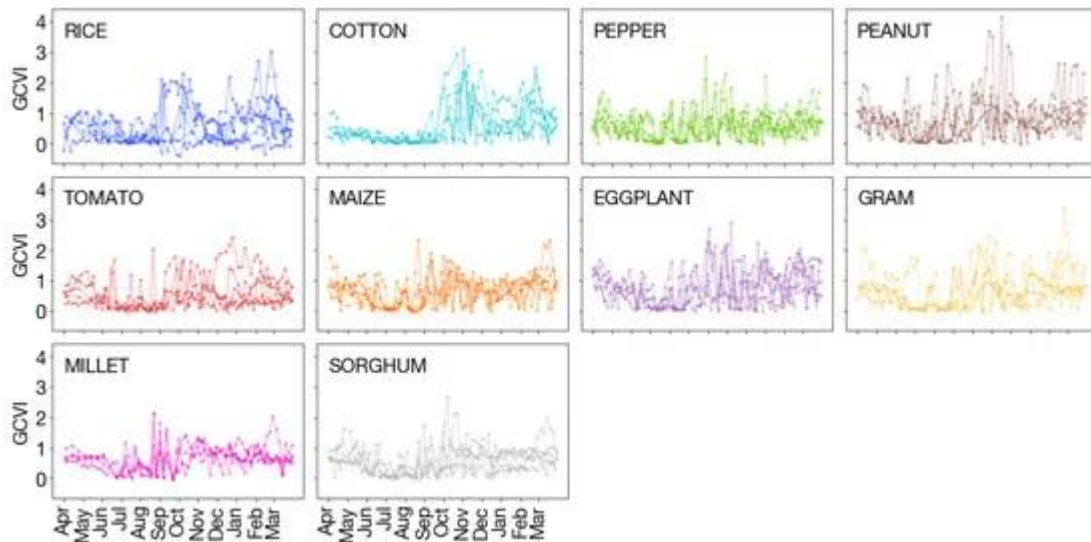


Figure 2: Raw Sentinel-2 time series. For each crop type, the green chlorophyll vegetation index (GCVI) of Sentinel-2 time series at 5 randomly sampled submissions are shown from 1 April of the crop year to 31 March of the next year. GCVI is defined as $NIR/green - 1$ and measures chlorophyll concentration in vegetation.

4. Methods

For a graphical overview of the methods presented in this paper, please see Figure 3.

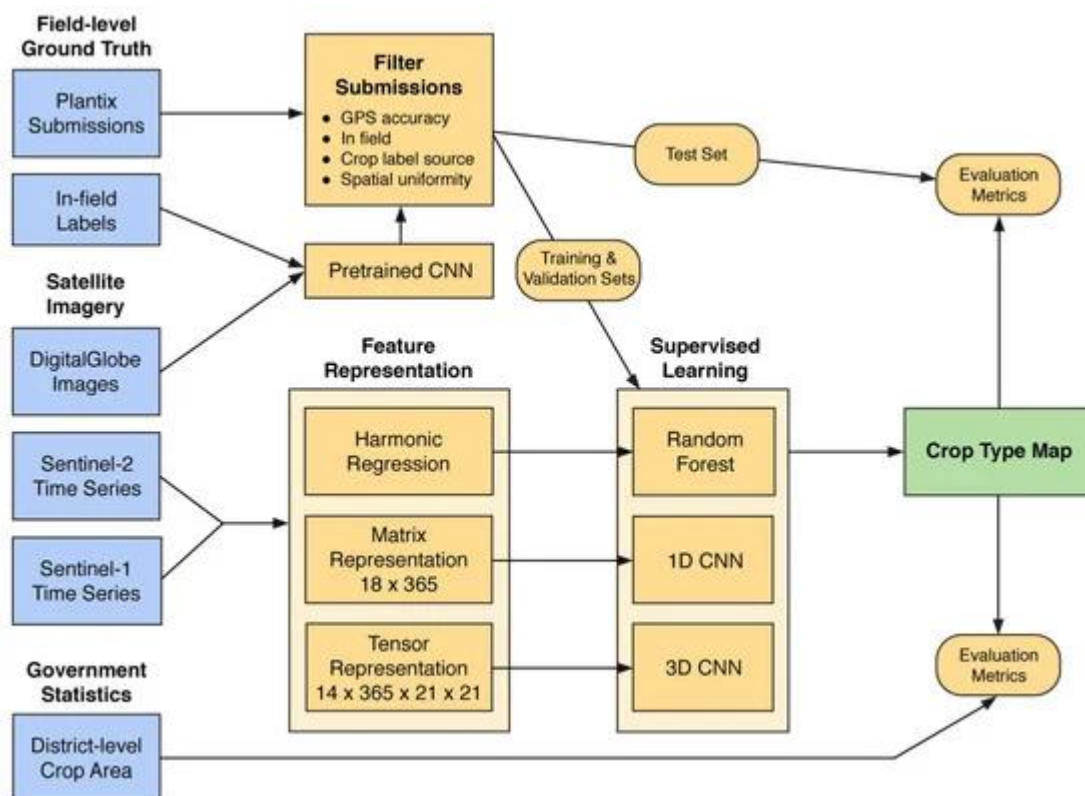


Figure 3: Explanatory diagram of the methods presented in this paper. Raw datasets are shown on the left; the data are cleaned and input to feature extraction and machine learning models, and the methods output crop type predictions that are validated against both a hold-out set of ground labels and government statistics on crop area

4.1 Measuring and Reducing Noise in Crowdsourced Data

4.1.1 Initial Filtering by PEAT

To extract usable data and minimize noise, our collaborators at PEAT applied several filtering steps to the original 1.8 million submissions. First, photos had to show a crop (1.3

million samples) and be labeled by an expert or with a Plantix-DNN prediction consistent with disease prediction (1.0 million samples). Second, the predicted crop type had to be one of a pre-selected list of crops: rice, cotton, peanut, pepper, tomato, eggplant, maize, gram, millet, and sorghum (620,000 samples), which account for about 70% of the region’s kharif crops (Table 1). Third, Android location

accuracy had to be available and within 200 m (213,000 samples). Lastly, only one image of each crop type was permitted per user (102,000 samples). This minimized the overrepresentation of very active users and submissions from the same field.

Table 1. Kharif season crop type distribution. Submission counts are shown for the filtered Plantix dataset and compared to government crop area statistics for Andhra Pradesh and Telangana in 2016–2017. For comparison, Indian state agriculture departments collected 11,469 points nationally in both kharif and rabi seasons in 2017–2018 [33].

Since there are two growing seasons in southeast India, we chose to focus on kharif crops to simplify the classification task. We defined kharif samples as those submitted between 1 June and 1 December, which filtered the dataset further to 72,000 samples. A geographic and temporal distribution of Plantix submissions in our study region is shown in Figure 1, and a quantitative summary by crop type, with comparisons to government statistics, can be found in Table 1. Rice is by far the most highly represented crop type in Plantix submissions with over 32,000 points, followed by cotton with over 10,000. All other crop types have below 10,000 submissions. It is worthwhile to note that, between the selection bias of submission through the Plantix app and filtering for usable samples, this dataset is likely to contain biases; we discuss these in more detail in Section 6.

The dataset of 72,000 kharif samples, though filtered, still included many noisy submissions. Below, we describe methods to further filter the dataset on Android location

accuracy, how much of the Sentinel-2 pixel is within a crop field, whether crop type was assigned by a human expert or by the Plantix-DNN, and spatial uniformity across the study region.

4.1.2. Android Location Accuracy

Information on the location accuracy of submissions was available via the Android platform. The Android Location API supplied the horizontal accuracy of a location as the radius of 68% confidence. In general, higher accuracy in location required higher battery drain, so preservation of battery life compromised location accuracy and therefore the majority of submissions' usefulness for crop type mapping. Filtering for submissions with accurate locations is a crucial step in de-noising crowdsourced data.

A histogram of submission location accuracy across a random sample of the 1.8 million dataset is shown in Figure 4a. The distribution is bimodal, with one small peak at 5 m and a second, much larger one at 3000 m. In comparison, the average operational holding size in AP and Telangana is 0.94 ha and 1.00 ha, respectively [28], and we observed individual fields as small as 10×10 m in DigitalGlobe imagery. This implies that a location accurate to 200 m is unlikely to still be inside the field of the submitted photo. We filtered for location accuracies at the 10 m, 20 m, and 50 m level in our sensitivity analyses; submissions with accuracies >50 m was discarded. Note that, since the Android location service seeks to preserve battery at the expense of accuracy, the ≤50 m criterion alone removes 61% of all submissions.

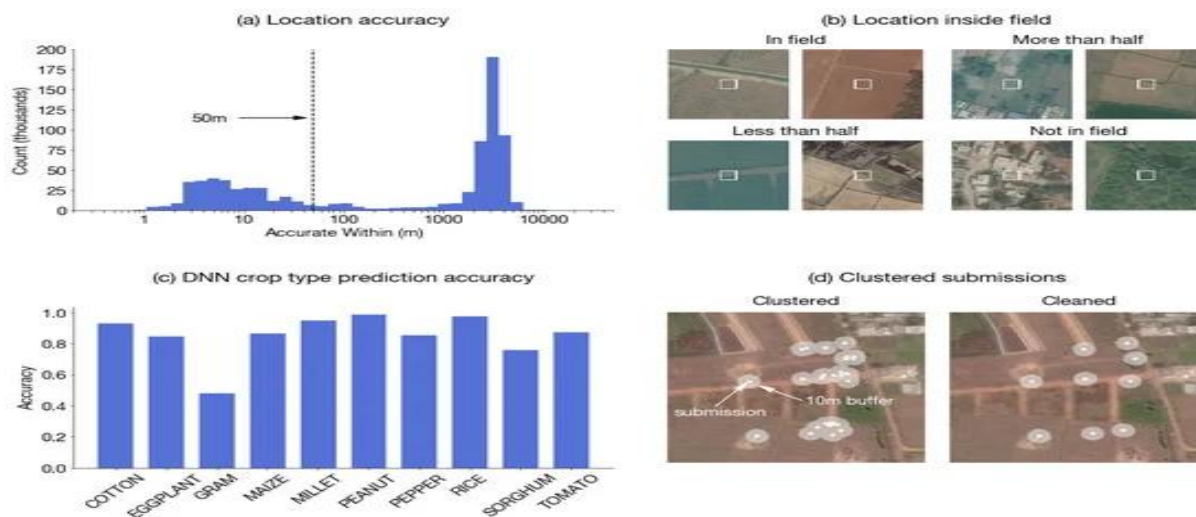


Figure 4: Sources of error in crowdsourced locations and crop type labels. (a) Android location accuracy varies from 1 m to 10 km depending on battery life preservation. Histogram shows accuracy distribution with 50 m as cutoff for inclusion in our analyses. (b) Sentinel-2 pixel (denoted by 10×10 m white box) may range from entirely inside to entirely not inside a crop field. (c) Plantix photo-based deep convolutional neural network (DNN) crop type prediction is imperfect and varies by crop type. (d) Clustered submissions from active locations overrepresent some geographic subregions

4.1.3. Pretrained CNN for In-Field Classification

A second source of label error is the farmer taking a photo of a plant without standing inside their field. They often take the photo from a road at the edge of their field. They often take the photo from a road at the edge of their field, pick off a diseased leaf, and photograph it in an urban area (i.e., in a place with WiFi), or stand under trees in fields while using the app. Though the crop type labels may be correct for the

submitted photos, the satellite time series at these non-field locations do not correspond to those labels.

We used the DigitalGlobe (DG) imagery described in Section 3.4 and a 2D CNN to classify each submission into one of four classes: “in field”, “more than half”, “less than half”, and “not in field” (Figure 4b). The rationale for finer-

grained labels “more than half” and “less than half” is that the spectral reading at a pixel that is mostly in a field may be dominated by the crop in that field, whereas pixels mostly not in one field are too mixed or contaminated.

Since the DG images are RGB, we classified them using well-established network architectures designed for natural images. We tried two commonly-used CNN architectures: VGG and ResNet [54,55], each with two network depths (VGG-11, VGG-19, ResNet-18, and ResNet-50). Their weights were initialized by pretraining on the large RGB image database ImageNet to boost classification accuracy with small training sets [56]. All four pretrained initializations were available off-the-shelf in the deep learning framework PyTorch [57]. We also tried two resolutions of DG imagery: 0.3 m and 0.6 m, the latter of which was downsampled from the 0.3 m imagery. At both resolutions, images were cropped to 224×224-pixels to fit the pretrained models.

We sampled 3000 submissions from the 72,000 kharif points obtained at the end of Section 4.1.1 in a geographically uniform manner (Figure A3). Through manual inspection of the DG images, which took a total of 4 hours, two human labelers generated in-field labels to train the CNNs. Figure A4 shows the agreement between the two labelers to give a sense of the task difficulty. Of these 3000 DG images, 2000 were placed in the training set, 500 in the validation set, and 500 in the test set; each split was constructed so that no points in one split were within 200 m of any points in the other two splits (to ensure non-overlapping DG images). Table 2a summarizes the distribution of labels.

The CNNs were trained to minimize cross entropy loss (Equation (2)) with $C=4$ for the four classes. During training, common data augmentation strategies were used to increase the diversity of the training set: random horizontal flips, vertical flips, rotations, and color jitters. The best model was selected via the highest validation set accuracy, and test set accuracy was evaluated using this model. Details of the network architectures and optimization parameters are shown in [Table A1](#), training loss and accuracy over epochs are shown in [Figure A5](#), and pretrained networks’ validation set accuracies are shown in [Table A2](#).

We note that this data cleaning step significantly reduces noise found in the Plantix dataset, but remains imperfect. CNN misclassifications aside, it is unclear which field a submission is from when it is on the boundary between two fields, but the submission may still be considered “more than half” in a field. Geolocation errors of Sentinel-2 images (especially Sentinel-2B prior to mid-2018), which we did not correct, may also add noise to the time series of submissions near field boundaries [58]. We therefore tested different in-field thresholds for inclusion in the training set and report their effect on performance in the results.

4.1.4 Expert vs. Plantix-DNN Labeling

Since the vast majority of crop type labels were assigned by Plantix’s deep neural network (DNN) trained on expert labels, the use of these imperfect labels introduces another source of noise. [Figure 4c](#) shows the DNN accuracy by crop

type, evaluated on the expert labels. The overall accuracy is 97%, while crop-specific recalls range from 48% for gram to 99% for peanut. Accuracies for rice and cotton are 97% and 93%, respectively. In the sensitivity analyses, we compared the performance of models trained on expert-labeled submissions to those trained on DNN-labeled submissions.

4.1.5 Spatial Distribution of Submissions

The geographic distribution of submissions is heavily concentrated around urban areas (e.g., Hyberadad) and locations of highly-active users ([Figure 1](#)), likely due to differences in smartphone ownership, internet access, and knowledge of the Plantix app. One field may generate multiple submissions when its farmer is an active user. We introduced another filtering step in which submissions within 20 m of another were considered duplicates and removed ([Figure 4d](#)).

4.2 Constructing Training, Validation, and Test Sets

Identifying a good crop type classifier and providing an unbiased out-of-sample estimate of classification performance requires validation and test sets that (1) have high label accuracy and (2) are representative samples of the region. To generate cleaned validation and test sets, we first filtered for the set of points that satisfied the following criteria.

- 1) The location accuracy is deemed to be ≤ 10 m by the Android platform.
- 2) The Sentinel-2 pixel at the submission location has been classified as either “in field” or “more than half” inside a field.
- 3) A crop scientist, not the DNN, assigned the crop type label based on the submission photo.

The resulting dataset was heavily skewed toward the eastern part of the study region ([Figure A8a](#)), undermining the ability of the validation set to select a good model for the entire region and of the test set metrics to represent the entire region. To achieve greater spatial uniformity, we started with an empty set, randomly sampled coordinates within the study region, and added the “clean” submission closest to the sampled coordinate until the validation and test sets each reached 400 samples ([Figure A8b](#)).

In these cleaned validation and test sets, the median location accuracy was 4.6 m and 4.1 m, respectively. This means that the sample has on average a 68% chance of being within 4.6 m (or 4.1 m) and a 95% chance of being within 9.2 m (or 8.2 m) of the submission location. Forty-one percent of the validation set were classified as completely inside a field, while 43% of the test set were completely in-field.

To see how much validation set noise affects the ability to choose good training data and models, we also constructed a noisy validation set comprised of 400 points sampled at random from the original dataset. It therefore includes submissions with GPS accuracy from 10–50 m, submissions not taken inside a crop field, and submissions labeled by the Plantix-DNN. Note that, to avoid overfitting and inflated metrics, no samples in the validation or test sets were within 500 m of samples in the other two sets.

The training set was derived from the remaining points in the dataset not in the validation and test sets (Figure A8c). Training points were filtered to be at least 500 m from all points in the validation and test sets (45,000 samples), to have Sentinel-2 pixels in field or more than half in a field (21,000 samples), and to not contain points within 20 m of each other. The final training set used to train the classifiers contained 9079 samples, a very large reduction from the 1.8 million raw submissions.

4.9 Crop Type Classification

As a result that smallholder systems are highly heterogeneous and simple rules to separate crop types are not immediately apparent (Figure 2), we tested the performance of three machine learning algorithms for feature extraction and crop type classification (Figure 5).

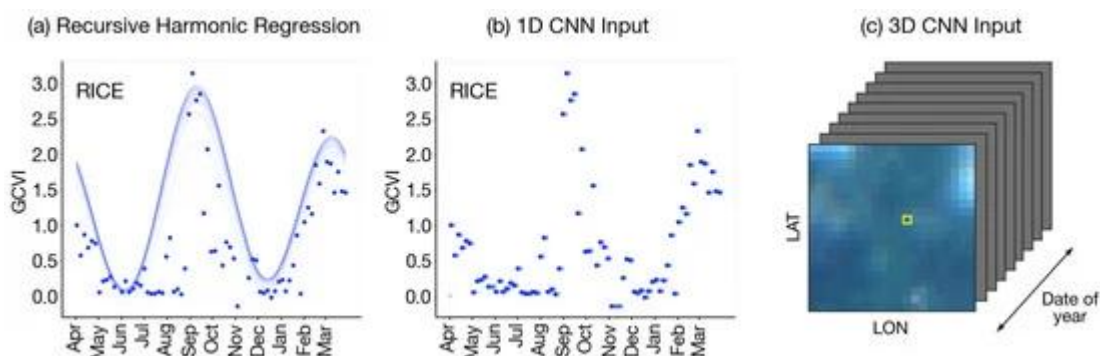


Figure 5: Feature extraction methods for Sentinel-2 time series. For the same example rice submission, (a) third-order harmonic regression with six recursive fits, (b) 1D convolutional neural network (CNN) time series input, and (c) 3D CNN time series schematic for an entire tile surrounding the submission are shown for the green chlorophyll vegetation index (GCVI) band. In (b,c), days without a Sentinel-2 image are filled in with the most recent previous image. In (c), the yellow box encircles the labeled submission pixel; all other pixels are unlabeled

4.3.1 Choosing Crop Types to Predict

The Plantix dataset contains ten crop types, whose distribution is shown in Table 1. Ideally, a classifier would achieve high precision and recall on all ten crops; in reality, the minor crop types do not have enough label quantity or signal in their time series for accurate classification. We first show results for 10-crop classification to demonstrate the difficulty of mapping minor crops over large geographic extents with only small label quantities. We then simplify the classification to a 3-class problem of distinguishing rice and cotton from all other crops (lumped into one “other” class). Rice and cotton were chosen because they are the two major kharif crops in and Nalgonda (Table 1), had label accuracy exceeding 90% from the Plantix DNN (Figure 4c), and had high precision and recall in the 10-crop classification task (Figure A13).

4.3.2. Random Forest with Harmonic Features

In order to use the phase and amplitude of plant phenology to differentiate crop types, a method is needed to transform variable-length discrete time series into features that can be input into machine learning algorithms. One such method whose features have been successfully used to classify crop types is the harmonic regression, or regression using a Fourier basis [18,34,35,36]. The harmonic regression decomposes a function of time into its frequencies, yielding a compact representation of the time series at each satellite

The first is a random forest using features derived from harmonic regressions on satellite time series, an algorithm that has performed well at crop type classification in previous studies [18,36]. The second is a 1D CNN with kernels convolving over the temporal dimension of the time series. The advantage of a CNN is that features are learned, not prescribed, and can take useful forms that the harmonic coefficients cannot. The third model is a 3D CNN with kernels convolving over two spatial and one temporal dimension of the time series for an entire image tile. The 3D CNN can see a broad spatial context that the previous two methods cannot, but it contains more parameters, is much more computationally intensive to train, and is more prone to overfitting on small datasets. A comparison of the data storage and computational runtime required for each model is provided in Table A3.

band or vegetation index (VI). Mathematically, it is equivalent to performing a discrete Fourier transform [59].

We viewed each satellite band or VI as a time-dependent function $f(t)$ and performed the harmonic regression

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

(1) for each band/VI independently, where a_k are cosine coefficients, b_k are sine coefficients, c is the intercept term, and n is the order of the harmonic series. The independent variable t represents the time an image is taken within a crop year expressed as a fraction between 0 (1 April) and 1 (next 31 March).

Due to the presence of clouds, the harmonic regression fit naively to a Sentinel-2 time series does not capture crop phenology well. In the absence of an accurate cloud mask, we followed a recursive curve fitting procedure similar to that implemented in the TIMESAT program [60], which has been shown to reduce the bias introduced by clouds. The algorithm recursively fits the harmonic regression to the time series, and then imputes the cloud-free values of the time series by taking the maximum (or minimum) of the band/VI and the regression fit if clouds appear as low (or high) values for that band/VI. For example, since clouds appear as low GCVI values, one iteration of the recursive algorithm would regress the raw GCVI values on the harmonic terms, then take the maximum of the fitted curve

and the raw GCVI values. This can be repeated for a total of r recursive fits (Figure 5a).

The values of n and r are hyperparameters that must be tuned via cross-validation for a given dataset and task. A larger n (more cosine and sine terms) increases model flexibility but risks the model overfitting to spurious patterns. Meanwhile, r should minimize the influence of clouds on the coefficients without obscuring real phenological signal. We picked $n=3$ and $r=2$ by minimizing crop type classification error on a hold-out set (Table A4).

After performing this regression recursively, we extracted coefficients $a_1, a_2, a_3, b_1, b_2, b_3$, and c for each of the 18 bands and VIs, giving us a total of 126 features on which to classify crop types. Since the model has seven parameters to fit, it requires at least seven cloud-free observations at a pixel to extract meaningful coefficients, a criteria that is met at all Sentinel-1 and Sentinel-2 pixels (Section 3.2). An example regression is shown in Figure 5a for GCVI on a rice submission time series.

Finally, to perform crop type classification, we trained random forest models with the harmonic coefficients as input and Plantix-labeled crop type as output. Random forest is an ensemble machine learning method comprised of many decision trees in aggregate [37], and has frequently been used in the field of remote sensing to perform land cover classification and crop type mapping [61,62]. It often yields higher accuracy than maximum likelihood classifiers, support vector machines, and other methods for crop type mapping [14,34,63,64]. We used the default parameters of Python's scikit-learn [65] package Random Forest Classifier with the exception of increasing $n_{\text{estimators}}$ (the number of decision trees in the random forest) from 10 to 500 to reduce model variance. Error bars on classification metrics were obtained by fitting the classifier on multiple bootstrapped training sets (sampled with replacement).

4.3.3. 1D Convolutional Neural Network

While random forests are commonly used for crop type mapping in the literature [14,18,36], obtaining features for the model still require the user to assume a functional form to summarize time series data. In contrast, neural networks learn both the feature representations from the raw data as well as how to use them to perform classification. If harmonic coefficients fail to capture some information that is helpful for classifying crop types, or random forests are not well-suited to learn the types of nonlinearities that characterize decision boundaries, a neural network has the potential to perform better.

To classify satellite time series, we constructed a 1D convolutional neural network. The time series for each sample was represented as an $18\text{row} \times 365\text{column}$ matrix, where each row is a band/VI and each column is a day of the year. The first 14 rows are Sentinel-2 bands and GCVI, and the last 4 rows are Sentinel-1 VV, VH, RATIO, and DIFF. This encoding was chosen to standardize Sentinel-1 and Sentinel-2 time series with different observation dates to the same neural network input size. If a satellite took an observation on day D , then column D in the matrix will be filled with that satellite's band values (Figure 5b). Since the

revisit times of Sentinel-1 and Sentinel-2 are 6 and 5 days, respectively, the values on days with no observation were imputed with values from the previous observation (known as "forward imputation" [66]). Lacking a high-quality cloud mask for Sentinel-2, we again were not able to remove cloudy observations. As a result, that neural networks can learn which parts of an input are relevant to the task at hand [67,68,69] and have previously been shown capable of ignoring cloudy observations [70], we did not further process the time series to reduce the influence of clouds, as we did for the random forest classifier. An occlusion sensitivity analysis, shown in Figure A15, shows that the 1D CNN indeed learns to rely largely on clear observations for classification.

In image classification, convolutions are 2D and are performed across the two spatial dimensions of the image. Here, each submission is comprised of one pixel and there are no spatial dimensions; the CNN instead convolves over the temporal dimension with kernels of size 3. The 1D CNN architecture is comprised of multiple convolutional blocks, each of which is a stack of 1D convolution, batch normalization, rectified linear unit (ReLU), and 1D max pooling layers (Figure A11). The final prediction is output by a few fully connected layers.

Training was performed by minimizing the cross-entropy loss, defined as a function of the i th input sample as

$$\ell(\theta, x, y) = -\sum_{c=1}^C \text{Cyclog} \hat{y}^c$$

(2) for model parameters θ , number of classes C , the input time series x , crop type probabilities $\hat{y} = f_{\theta}(x)$, and one-hot ground truth label y . The notation y^c denotes the c th element of the vector y , which is equal to 1 if the sample belongs to class c and 0 otherwise. The element \hat{y}^c is the predicted probability that the sample belongs to class c . Minimizing cross entropy incentivizes the network to maximize the value of \hat{y}^c for the correct class c . Figure A12 shows a typical training curve for the 1D CNN.

Hyperparameters, such as the number of convolutional blocks and the number of filters per layer, were chosen to maximize prediction performance on a validation set and are shown in Table A6. The model that performed the best on the validation set was a CNN with 4 layers, 64 filters in the first layer, a learning rate of 0.001, and a batch size of 16. Other implementation details can be found in Table A5.

4.3.4. 3D Convolutional Neural Network

A limit of the 1D CNN is that it is only able to use temporal information to classify crop types; it does not "see" the spatial context that includes clues like field size and shape, surrounding vegetation, and proximity to buildings and roads. To see whether spatial information can improve crop type classification, we built a 3D U-Net, modeled after the popular 2D U-Net for image segmentation [71].

A 21×21 -pixel tile of each Sentinel-2 image was exported around each submission coordinate to allow spatial features to inform crop type classification. We did not export Sentinel-1 tiles, as the additional storage and computational runtime (Table A3) was high compared to the marginal benefit SAR brought to CNNs in this setting (Table A7). The input to the 3D U-Net is therefore a 4D tensor of size

14×365×21×21, where days with no Sentinel-2 observation are imputed with the previous observation. Thus, the network sees not only the labeled pixel's time series, but also the time series of pixels up to 200 m away (Figure 5c). The tensor is first down sampled (encoded) via blocks of 3D convolution, batch normalization, ReLU, and 3D max pooling operations in the first half of the network, then upsampled (decoded) back to its original resolution in the second half. A diagram of the network is shown in Figure A16.

The output of the network is a 21×21-pixel segmented prediction in which every pixel is assigned a crop type label. Since we only observed the label at one pixel in the image (Figure 5c), we only computed the loss and performance metrics at that pixel. Like the 1D CNN, the 3D U-Net was trained using a cross entropy loss with C classes (Equation (2)). Figure A17 shows a typical training curve for the 3D CNN.

4.4. Feature Importance via Permutation

We performed experiments to determine the relative importance of features to crop classification by permuting each feature across all samples, as suggested in [37]. The algorithm is as follows.

- 1) Fit a classifier to the training set (e.g., harmonic coefficients and random forest, 1D CNN).
- 2) Record the baseline predictive performance of the model on the validation set (i.e., accuracy).
- 3) For each feature j , randomly permute feature j in the validation set, thereby breaking the association between feature j and the label y . Apply the model to this modified validation set, and record the model performance.
- 4) The feature importance is the difference between the baseline performance and the permuted performance.

We applied this algorithm to the 126 harmonic coefficient features with the random forest classifier to see which bands and Fourier frequencies decrease classification accuracy the most when permuted. With the 1D CNN, the spectral and temporal dimensions were permuted independently to study which bands and times of year are most important to distinguishing crops. Band b was permuted across all time steps with the same band from another sample, or, for each time step t , all bands were permuted with those of another sample from 1 April to date t .

Note that a feature's importance via permutation is not the same as how much worse a model would perform if trained without that feature, due to correlations between features. That is, a model trained without a particular feature can rely more on other correlated features to compensate. For a correlation matrix of the harmonic coefficients, see Figure A9.

4.5 Assessing the Additional Value of Sentinel-1

As a result that the use of Sentinel-1 for crop type mapping is relatively recent and its utility is not fully known, we performed experiments in which we classified crop types using only Sentinel-2 time series and compared performance

to using both Sentinel-2 and Sentinel-1. We compared results for both the harmonics/random forest model and the 1D CNN.

4.6 Validation against District-Wise Production Data

We sampled ten thousand points uniformly at random from the study region in areas classified as cropland by the Global Food Security-support Analysis Data (GFSAD) Cropland Extent 30 m map [72] and exported all images taken at these points by Sentinel-2 and Sentinel-1 in the period 1 April 2018–31 March 2019. We then used the highest-performing 1D CNN trained on Plantix labels to classify the unlabeled samples into 3 classes (rice, cotton, and other), where rice and cotton had the highest precision and recall among all 10 crops. We compared the percent of samples classified as each crop to the percent of cropland devoted to each crop in 2016–2017 district-level statistics. At the time of writing, statistics were not yet available for 2017–2019, so they predate our classified samples by two years. We exported samples for the 2018–2019 kharif season instead of 2017–2018 because the former has more frequent Sentinel imagery.

4.7 Study Region-Wide Crop Type Map

Using Google Earth Engine, we computed harmonic features on Sentinel-1 and Sentinel-2 bands (Section 4.3.2) across Nalgonda and Telangana for the crop year 1 April 2018 to 1 April 2019. Ten thousand Plantix points from the 2017–2019 kharif seasons were sampled as described in Section 4.2 and used as training points in an Earth Engine random forest classifier with 500 trees. The points were labeled with one of three kharif crop classes: rice, cotton, and other crops. The trained random forest was then applied to all Sentinel pixels in the rest of the region to predict among the three crop classes. We used the harmonic features and random forest for map creation due to their ease of scalability in Earth Engine, as the entire study region contains over 2.7 billion Sentinel pixels. Lastly, we masked out pixels that were not deemed to be cropland by the GFSAD cropland product [72].

5. Results

5.1 DigitalGlobe Images for In-Field Classification

Pretrained convolutional neural networks fit to our in-field dataset ($n=2000$) were able to classify whether the center boxes are in-field with high accuracy. The best model, a pretrained Resnet with 18 layers using 0.3 m resolution DigitalGlobe imagery (Table A2), distinguished between center boxes that are completely, more than half, less than half, and not in a field with 74.2% test set accuracy (Table 2b). Comparison to the baseline accuracy (guessing majority class) of 38.4% and human labeler agreement of 82.5% (Figure A4) indicates that the Resnet performed considerably better than the baseline on a task that can often be confusing to humans.

When classification errors occurred, they came mostly from difficulty telling adjacent classes apart, rather than confusing boxes not in a field with those entirely in a field (Figure A6).

Grouped together into a binary “more than half” versus “less than half” classification, the corresponding test set accuracy was 89.0%. Examples of correctly and incorrectly classified images are displayed in Figure A7. We see that boxes in urban areas and boxes entirely in rectangular fields were easy to identify as “not in field” and “in field”, respectively, while irregularly-shaped fields, lone trees, and dirt roads occasionally confused the classifier.

We applied the trained ResNet-18 model to predict whether Sentinel-2 pixels are in a field on the remaining submissions without in-field labels, thereby allowing all submissions to be filtered on this attribute. Of the unlabeled submissions, over 55% had Sentinel-2 pixels more than half in a field. These “more than half” in field Sentinel-2 pixels were used to train the crop type classification models.

5.2. Crop Type Classification with Multi-Temporal Satellite Imagery

Both neural networks and harmonic coefficients with random forests were able to distinguish rice and cotton from other crops with overall accuracy above 70% (Table 3b); for comparison, a baseline model that classifies everything as the most common class (rice) would achieve 39% accuracy. However, both models struggled to classify minor crops for which there is less data. The 1D CNN, with its highly flexible feature-learning algorithm, consistently outperformed the harmonics and random forest classifier by a small but statistically significant amount (3-class test set accuracy: $74.2 \pm 1.4\%$ CNN versus $71.5 \pm 0.7\%$ harmonics/random forest). On the 10-crop task, however, the recall of non-rice/cotton crops ranged from 20–50% (maize, peanut, pepper, tomato) to 0% (eggplant, gram, millet, and sorghum) (Table 4, full confusion matrix in Figure A13). For all models, the precision and recall were positively correlated with the number of samples available in the training set (Figure 6). The 3D CNN, which has a 21×21 -pixel contextualized view of each sample, achieved accuracy lower (3-class task: 72.7%) than the 1D CNN at the cost of much greater data storage and computational runtime (Table A3), which may be due to the much larger number of parameters in the 3D CNN (6.1×10^6) compared to the 1D CNN (2.6×10^6).

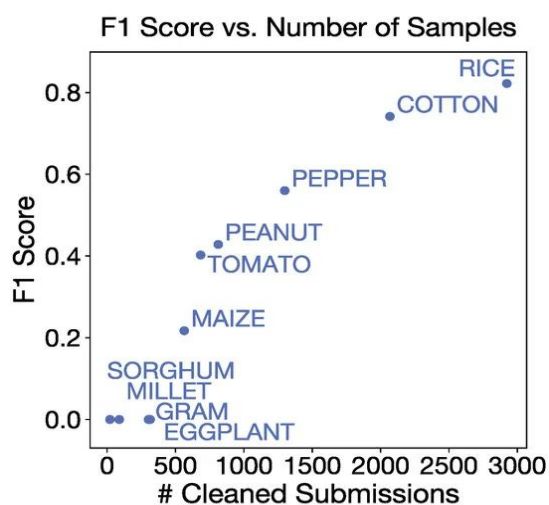


Figure 6: F1 score versus training set size by crop type. Precision and recall from Table 4 are summarized as one F1 score for each crop type.

Table 3. Summary of crop type classification results. Overall accuracy and precision, recall, and F1 scores (weighted average across classes) are shown for models trained on Sentinel-1 and -2 combined features to classify (a) all 10 crops and (b) simplified 3-class task of rice, cotton, and other crops.

5.3 Combined Sentinel-1 and -2 Imagery Versus Sentinel-2 Only

In our comparison of classification accuracy under models using both Sentinel-1 and -2 imagery versus only Sentinel-2 imagery, we found that adding Sentinel-1 improves performance, especially when features are harmonic coefficients (Table A7). Indeed, on the 3-crop task, adding Sentinel-1 improves overall accuracy on the validation set from $69.4 \pm 0.6\%$ to $75.0 \pm 0.8\%$ when using harmonic features. Permutation experiments also show a number of VV and VH coefficients in the top 30 most important harmonic features (Figure A10). However, the additional value of Sentinel-1 bands on 1D CNN accuracy is not statistically significant relative to the variance introduced by bootstrapping the training set. This may be because the neural network is able to extract more information from the raw Sentinel-2 time series, thereby diminishing the returns from adding Sentinel-1 data. Permutation experiments on the 1D CNN reveal that the neural network relies strongly on the red edge bands (especially around 740 nm, RDED2), short wave infrared, and the difference of VV and VH polarizations (DIFF) to differentiate crop types (Figure 7). This is consistent with prior work showing that red-edge bands are sensitive to leaf and canopy structure [73,74,75] and SWIR bands are sensitive to leaf and soil water content [75,76]. Surprisingly, the VH/VV ratio did not emerge as the top SAR feature despite prior evidence suggesting its sensitivity to crop growth [50]. Instead, the DIFF feature, which is sparsely documented to date, was favored by the model. Permutation of times of year also shows that the most important months to have satellite data are October and November, which correspond to the harvest of kharif crops.

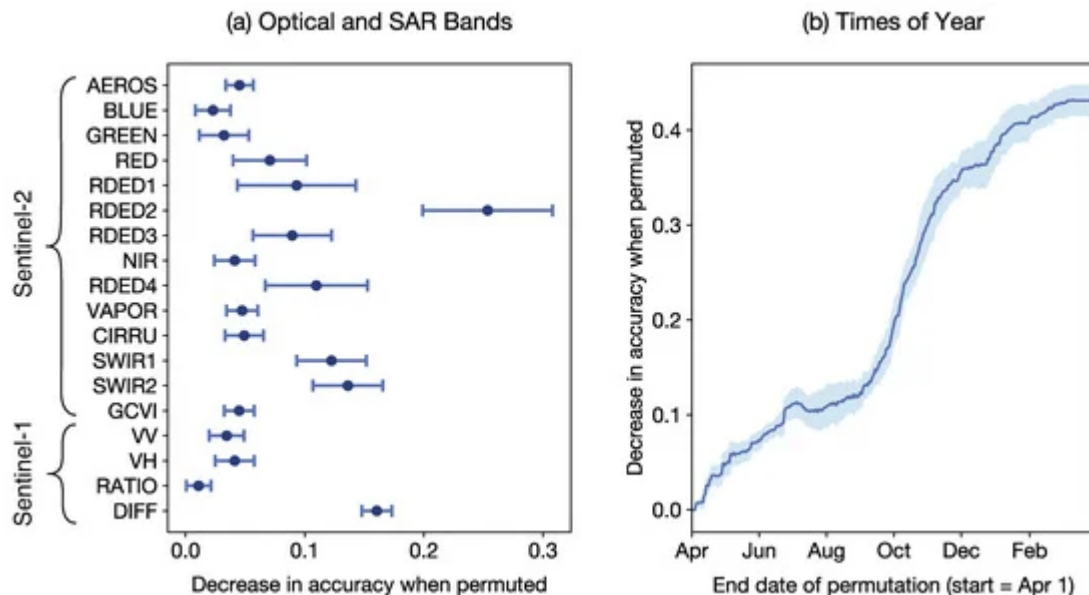


Figure 7: CNN feature importance. Importance was computed as the decrease in validation set accuracy when the feature was permuted across samples, breaking the association between the feature and the label. For the 1D CNN, permuted features were (a) optical and SAR bands and (b) times of year. Error bars were obtained over 10 runs with 10 different bootstrapped training sets

5.4 Robustness to Location and Label Noise

Crowd sourced data can contain many sources of error, and efforts to reduce noise can be costly. To understand whether a noisy Plantix dataset can still be useful for crop type mapping and how much investment should be made to clean it, we tested the sensitivity of 1D CNN performance to both training and validation set quality. Figure 8a shows the effect of increasing training set noise along three axes (GPS

accuracy, label origin, and in-field threshold) on a cleaned validation set, while Figure 8b shows the same effects as they appear on a noisy validation set. We use overall accuracy as the metric, but obtained the same analysis for F1 scores.

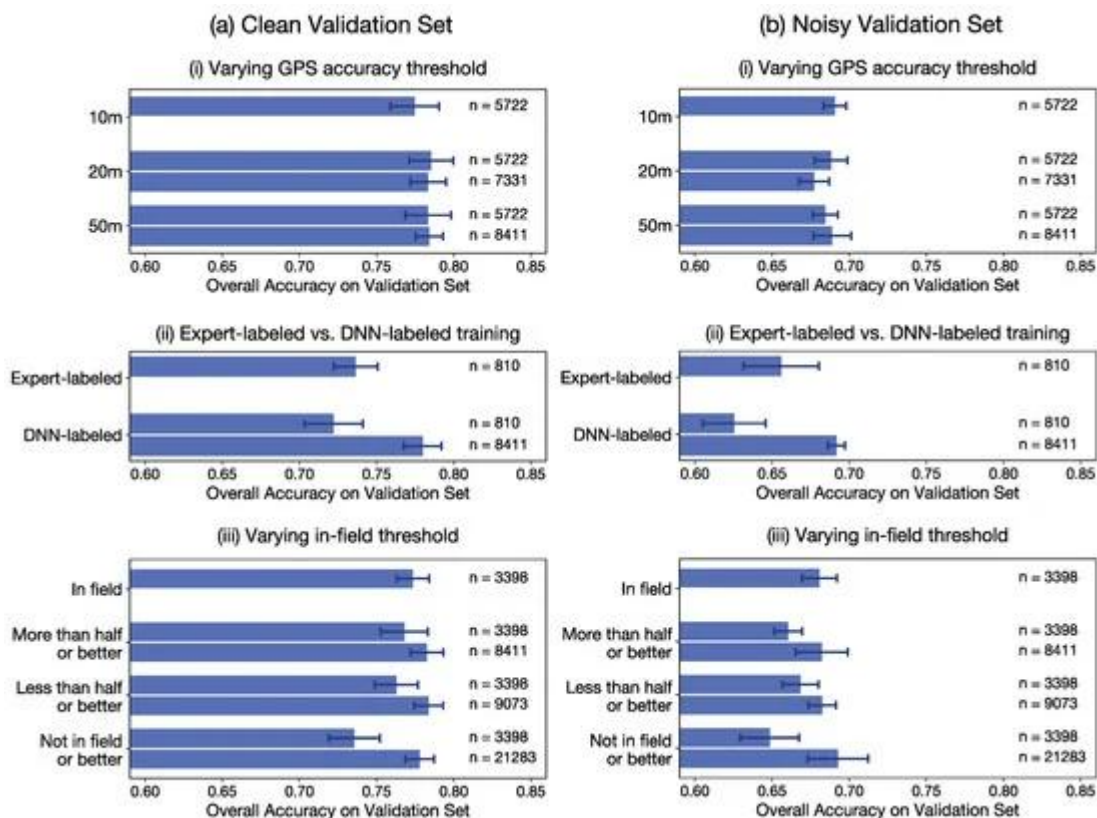


Figure 8: Sensitivity of classification results to dataset noise

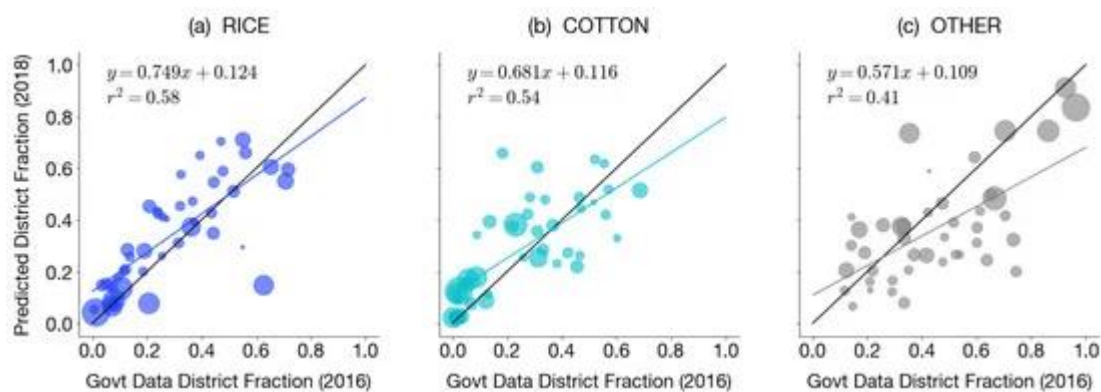
The effect of increasing label error is shown along three dimensions of noise: GPS accuracy threshold (10 m, 20 m, 50 m), label source (expert vs. DNN), and in-field threshold (in field, more than half, less than half, and not in field) for (a) a clean validation set and (b) a noisy validation set.

The conclusions concerning training set quality are similar from both validation sets. First, holding training set sizes constant, training sets with moderate levels of noise did not yield significantly worse validation set accuracies than the highest quality training sets. For example, classifiers trained on samples with GPS locations accurate to 50 m performed as well as those trained on samples with GPS locations accurate to 10 m. Second, adding samples whose labels are noisy to the training set can still boost classification performance. Adding samples labeled by the Plantix-DNN, samples with GPS locations accurate to 50 m, and samples whose Sentinel-2 pixels are only partially in a crop field all increased validation set accuracy (though not to a statistically significant extent in the last two examples). Lastly, high levels of label noise do decrease classification performance, as seen when submissions not in crop fields were added to the training set. However, even in this last setting, validation accuracy degraded only a few percentage points, and the decrease disappeared when all available data was used for training instead of a subsample with the same training set size.

While the CNN is robust to some training set noise, it requires a high quality validation set to yield accurate error estimates on unseen data. Accuracies on the noisy validation set are consistently 10% lower than those on the clean validation set, so that a naive interpretation of results on a noisy validation set would underestimate true classifier performance.

5.5 Comparison to District-Level Data

In addition to validating the classifiers on a hold-out set of Plantix submissions, we also compared the 3-class 1D CNN predictions on 2018–2019 GFSAD cropland samples (see Section 4.6) to the 2016–2017 district-level kharif season crop area statistics from the Indian Department of Agriculture [9]. Across 43 districts, R2 values between the fraction of samples predicted and the fraction of cropped area in the official statistics were 0.58, 0.54, and 0.41 for rice, cotton, and other crops, respectively (Figure 9). The classifier's aggregated predictions captured broad district-level characteristics correctly—districts like East Godavari, West Godavari, Krishna, and Srikakulam are dominated by rice; Adilabad, Nalgonda, and Warangal are dominated by cotton, and Anantapur and Chittoor grow mostly peanut (groundnut). However, our rice prediction for SPSR Nellore was much lower than the official statistic, and our cotton prediction for Vikarabad was much higher. These discrepancies may be due to a combination of classifier bias, error in the district statistics, and mismatch between the year of our samples and the year of district statistics. Note that comparatively few training samples came from SPSR Nalgonda (Figure A8), and historical changes in crop area statistics in this district have also been large, suggesting statistics could also have changed between 2016–2017 and 2018–2019 (Figure A18). Meanwhile, cotton dominated training samples in Nalgonda and could have biased the classifier toward predicting cotton on similar time series. While the latitude and longitude of samples were not explicitly provided as features to any classifiers, the 1D CNN could have learned to associate crop type with, say, the satellite image acquisition schedule or regional phenology in addition to with meaningful phenological characteristics. Without more updated district statistics, it is difficult to diagnose the main source of discrepancy between our aggregated predictions and the statistics.



6. Discussion

6.1 Contributions and Shortcomings of This Work

We draw the following lessons from cleaning Plantix submissions to supervise crop classification, which are applicable to both crop type mapping in smallholder systems and land use mapping more generally.

First, crowdsourced data, though very high in noise, can be cleaned and used to supervise remote sensing tasks if data

quality is measured for each sample. In this study, location accuracy, whether the submission came from inside a field, and whether crop type was labeled by human experts were important metadata on which to filter Plantix submissions. We discuss the challenges and potential of crowdsourced data in detail in the next section.

Second, a high-quality hold-out set is crucial for assessing model performance accurately. For the same classifier, our results indicate that a higher level of noise in the validation set biases the validation accuracy downward, since the noise

is random and follows no learnable pattern. Given that our cleaned validation and test sets still contain non-zero noise, our reported metrics may in fact be underestimates of model performance. At the same time, since conclusions drawn about training set noise sensitivity were similar for both the clean and noisy validation sets, a noisy validation set can still be useful for data filtering and model selection.

7. Conclusions

This is the first study to explore the potential of crowdsourced data to augment or replace ground surveys for land use mapping at a large scale. We derived a large but noisy crowdsourced dataset from the Plantix mobile app to train and validate a crop type map in southeast India. Two million farmer submissions were filtered to 10,000 higher-quality labels, and three machine learning models trained on multi-temporal satellite imagery were able to differentiate rice and cotton from other crops with 70+% accuracy. We found classification performance to be robust to moderate levels of the label and location noise common to crowdsourced data. Our 3-crop prediction (rice, cotton, other) for the 2018 kharif season was validated against a hold-out set of Plantix data and district-level crop area statistics from the Indian Department of Agriculture, and is available upon request.

References

- [1] Khalil, C.A.; Conforti, P.; Ergin, I.; Gennari, P. *Defining Small Scale Food Producers to Monitor Target 2.3. of the 2030 Agenda for Sustainable Development*; Technical Report; Food and Agriculture Organization of the United Nations: Rome, Italy, 2017. [Google Scholar]
- [2] Lowder, S.K.; Skoet, J.; Raney, T. The Number, Size, and Distribution of Farms, Smallholder Farms, and Family Farms Worldwide. *World Dev.* 2016, 87, 16–29. [Google Scholar] [CrossRef]
- [3] Rapsomanikis, G. *The Economic Lives of Smallholder Farmers*; Technical Report; Food and Agriculture Organization of the United Nations: Rome, Italy, 2015. [Google Scholar] [CrossRef]
- [4] Ricciardi, V.; Ramankutty, N.; Mehrabi, Z.; Jarvis, L.; Chookolingo, B. How much of the world's food do smallholders produce? *Glob. Food Secur.* 2018, 17, 64–72. [Google Scholar] [CrossRef]
- [5] Samberg, L.H.; Gerber, J.S.; Ramankutty, N.; Herrero, M.; West, P.C. Subnational distribution of average farm size and smallholder contributions to global food production. *Environ. Res. Lett.* 2016, 11, 124010. [Google Scholar] [CrossRef]
- [6] Burke, M.; Lobell, D.B. Satellite-based assessment of yield variation and its determinants in smallholder African systems. *Proc. Natl. Acad. Sci. USA* 2017, 114, 2189–2194. [Google Scholar] [CrossRef]
- [7] Plourde, J.D.; Pijanowski, B.C.; Pekin, B.K. Evidence for increased monoculture cropping in the Central United States. *Agric. Ecosyst. Environ.* 2013, 165, 50–59. [Google Scholar] [CrossRef]
- [8] Espey, J. *Data for Development: A Needs Assessment for SDG Monitoring and Statistical Capacity Development*; Technical Report; Sustainable Development Solutions Network: New York, NY, USA, 2015. [Google Scholar]
- [9] Ministry of Agriculture and Farmers' Welfare. Crop Production Statistics Information System. 2017. Available online: <https://aps.dac.gov.in/APY/Index.htm> (accessed on 28 September 2019).
- [10] Christiaensen, L. Agriculture in Africa—Telling myths from facts: A synthesis. *Food Policy* 2017, 67, 1–11. [Google Scholar] [CrossRef]
- [11] Atzberger, C. Advances in Remote Sensing of Agriculture: Context Description, Existing Operational Monitoring Systems and Major Information Needs. *Remote Sens.* 2013, 5, 949–981. [Google Scholar] [CrossRef]
- [12] USDA National Agricultural Statistics Service Cropland Data Layer. Published Crop-Specific Data Layer [Online]. 2018. Available online: <https://nassgeodata.gmu.edu/CropScape/> (accessed on 29 August 2019).
- [13] Fisette, T.; Rollin, P.; Aly, Z.; Campbell, L.; Daneshfar, B.; Filyer, P.; Smith, A.; Davidson, A.; Shang, J.; Jarvis, I. AAFC annual crop inventory. In Proceedings of the 2013 Second International Conference on Agro-Geoinformatics (Agro-Geoinformatics), Fairfax, VA, USA, 12–16 August 2013; pp. 270–274. [Google Scholar] [CrossRef]
- [14] Inglada, J.; Arias, M.; Tardy, B.; Hagolle, O.; Valero, S.; Morin, D.; Dedieu, G.; Sepulcre, G.; Bontemps, S.; Defourny, P.; et al. Assessment of an Operational System for Crop Type Map Production Using High Temporal and Spatial Resolution Satellite Optical Imagery. *Remote Sens.* 2015, 7, 12356–12379. [Google Scholar] [CrossRef]
- [15] Kremen, C.; Iles, A.; Bacon, C. Diversified Farming Systems: An Agroecological, Systems-based Alternative to Modern Industrial Agriculture. *Ecol. Soc.* 2012, 17, 44. [Google Scholar] [CrossRef]
- [16] Sheahan, M.; Barrett, C.B. Ten striking facts about agricultural input use in Sub-Saharan Africa. *Food Policy* 2017, 67, 12–25. [Google Scholar] [CrossRef]
- [17] Defourny, P.; Bontemps, S.; Bellemans, N.; Cara, C.; Dedieu, G.; Guzzonato, E.; Hagolle, O.; Inglada, J.; Nicola, L.; Rabaute, T.; et al. Near real-time agriculture monitoring at national scale at parcel resolution: Performance assessment of the Sen2-Agri automated system in various cropping systems around the world. *Remote Sens. Environ.* 2019, 221, 551–568. [Google Scholar] [CrossRef]
- [18] Jin, Z.; Azzari, G.; You, C.; Tommaso, S.D.; Aston, S.; Burke, M.; Lobell, D.B. Smallholder maize area and yield mapping at national scales with Google Earth Engine. *Remote Sens. Environ.* 2019, 228, 115–128. [Google Scholar] [CrossRef]