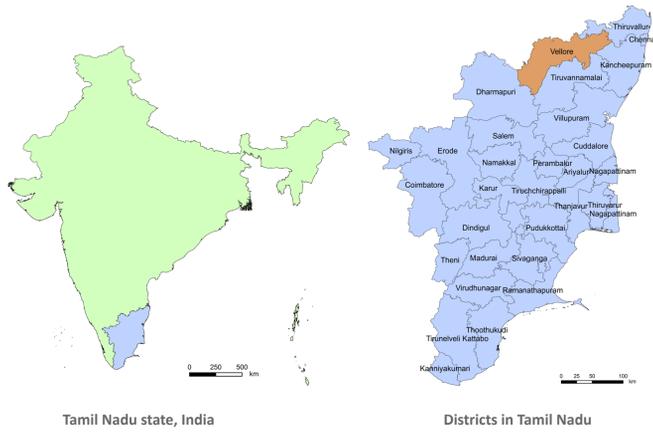


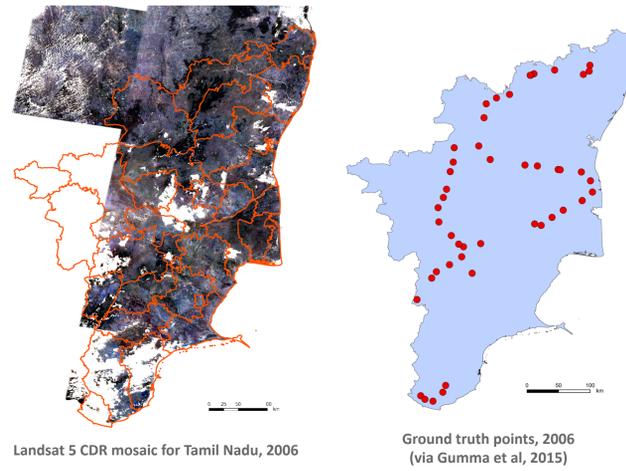
# SUPERVISED CLASSIFICATION OF AGRICULTURAL CROPS

## A CASE STUDY IN VELLORE, INDIA

### PROJECT LOCATION



### DATASETS



### RESEARCH QUESTIONS:

What is the spatial distribution of crops grown in Vellore, India?

Which vegetative or chlorophyll indices are most suitable for crop identification?

### INTRODUCTION:

Agriculture in Vellore district in Tamil Nadu, India, utilizes 29.5% of the district's land and employs 27% of its population. Vellore farmers used 85,578 megatons of fertilizers, 105 megatons of dusted pesticides, and 9,205.6 liters of liquid pesticides in 2012-2013 to support this growth. Groundwater irrigation through open wells and tube wells irrigates 50% of Vellore's net sown area. However, no groundwater source in Vellore has been assessed to be safe for future use; due to severe depletion, all resources have been classified as Semi-Critical, Critical, or Overexploited since 2004. Increased groundwater contamination by pesticides in this region will result in increased exposure and health risk for the population, especially for those living in close proximity to fertilized agricultural regions. This analysis aims to use remote sensing data to characterize the spatial distribution of crops in Vellore for further development of a pesticide use map.

### METHODS:

**Data Collection** | Landsat 4-5 Climate Data Record (CDR) data was downloaded from USGS Earth Explorer. CDR data provides pre-processed cloud masks and corrected band reflectance values for optimal processing.

**Masking** | Masks were applied to exclude clouds, cloud shadows, ocean, and inland water bodies from analysis

**Mosaicking** | Images were mosaicked via the Seamless Mosaicking tool in ENVI to generate a state-wide mosaic for September/October 2006

**Vegetative Indices** | Band Math was used to calculate NDVI, OSAVI, TCARI, MTVI, and TCI indices to detect differences in crop type

**Supervised Classification** | Ground truth points from an agricultural survey by Gumma et al (2015) were used to classify crops. 50% of the original data was used as the training set, and 50% was used for validation

**Yield Approximation** | Rough crop distribution for Tamil Nadu state was approximated

**Validation** | Vellore district crop area estimates was used to validate classification for 2006

### RESULTS:

CROP	TRUE AREA CROPPED (km <sup>2</sup> )	TCARI CROP AREA (km <sup>2</sup> )		GROUND TRUTH CLASSIFICATION CROP AREA (km <sup>2</sup> )	
		Classified	Error %	Classified	Error %
Coconut	230.98	1527.00	561.09	362.20	56.81
Cotton	70.13	441.50	529.54	601.60	757.84
Groundnut	513.94	1063.10	106.85	379.10	26.24
Maize	2.73	365.60	13291.94	1,221.30	44636.26
Plantains	3.00	747.20	24806.66	660.30	21910.00
Rice	199.42	922.20	362.44	366.20	83.63
Sorghum	104.62	819.60	683.40	436.10	316.84
Sugarcane	172.02	535.40	211.24	1059.80	516.09

A significant percent error exists in these classifications, particularly in classifying maize and plantain crops. This is likely a combination of various factors. Firstly, the ROIs from Gumma et al (2015) were recorded as 500m x 500m plots, with more than one crop grown in each area. In the absence of a pure spatial reference for each type of crop grown in the ROI, the dominant crop was chosen as the primary classifier. This likely resulted in a large range of pixels to be included in the classification rule for each crop, resulting in increased error.

The Maximum Likelihood classification method accounts for spectral signatures of each ROI pixel between all six bands. Therefore, it is more sensitive to spectral differences, which are key to accurate classification. However, the Maximum Likelihood algorithm requires at least two bands for execution, which were not available from index calculations. Parallelepiped classification, which was used for the single-band index classifications, performs poorly in areas with overlapping ROIs. Thus, error from both the ROIs and the classification methodology, confounded the classification, leading to misclassification and high error.

Based on these results, TCARI is a marginally better classifier than the original image, but further analysis is necessary to determine its role in agricultural crop classification and identification.

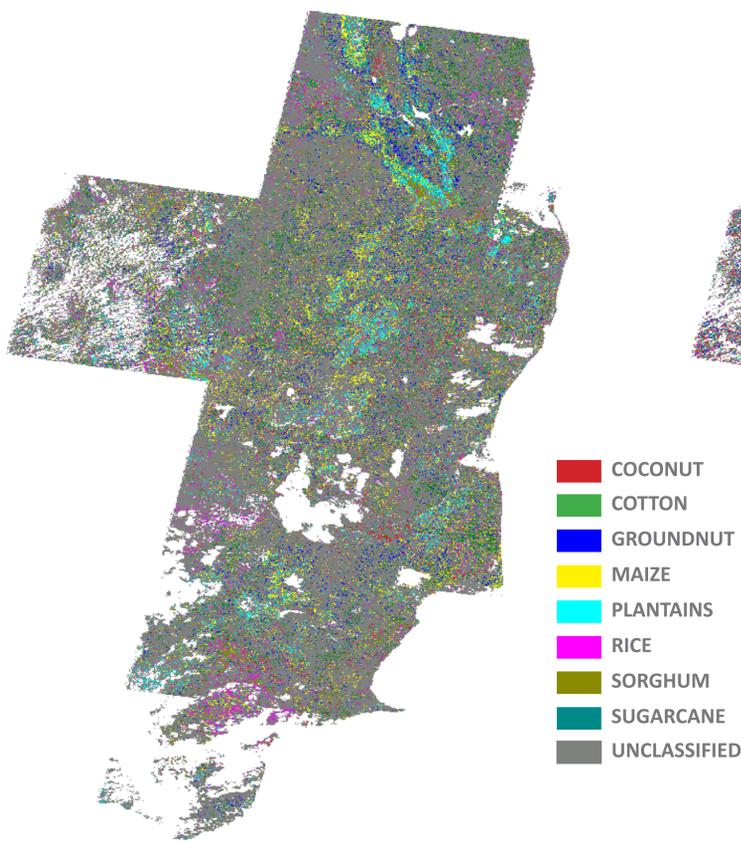
### NEXT STEPS:

An improved classification routine including probability of discriminating "true crop" vs. "false crop" pixels must be developed. This classification scheme must include a high probability of non-agricultural pixels, and narrow, pure ROIs for each crop type. This information may be obtained by further refining the Gumma et al (2015) dataset, or by collecting new data from a field survey. Additionally, to account for missing and masked pixels in each region, an experimental interpolation method is required to map the spatial distribution of crops throughout the district and state.

An important goal of this work is to extend aforementioned methods to Indian Space Research Organization (ISRO) products. ISRO provides LISS III imagery, and IMS-1 hyperspectral imagery, which offer improved spatial resolution and less cloud cover compared to Landsat images. Being able to utilize these products will allow for improved development of site-specific crop management techniques, regional natural resources management, and resilience planning in the move towards precision agriculture.

### CLASSIFICATION

#### "Ground Truth" Image Classification

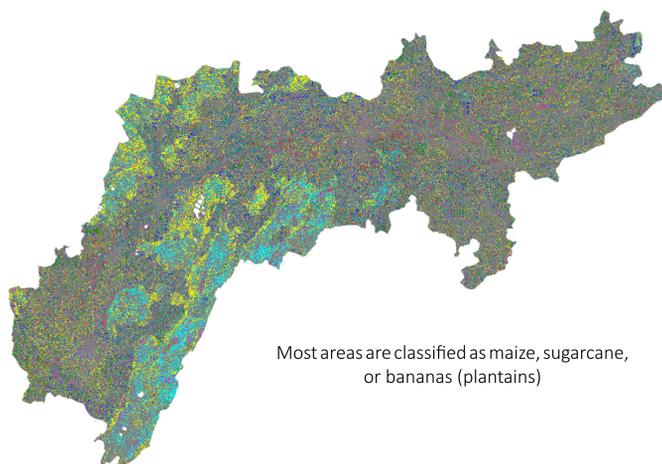


#### Transformed Chlorophyll Absorption in Reflectance Index (TCARI)

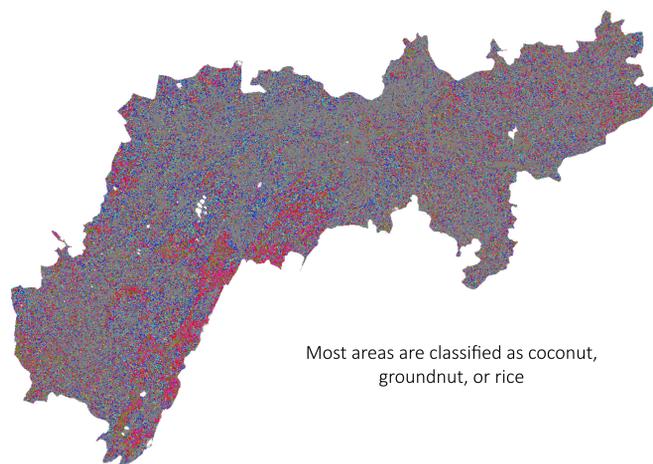


### VALIDATION

#### Vellore Ground Truth Image



#### Vellore TCARI Image



#### "GROUND TRUTH" MAXIMUM LIKELIHOOD CLASSIFICATION

Overall Accuracy: 4.51%      Kappa coefficient: 0.0137

	Coconut	Maize	Plantains	Rice	Sorghum	Sugarcane	Total
Unclassified	68.99	94.07	82.62	87.93	91.47	75.42	83.21
Coconut	4.97	0.4	0.08	1.13	0	2.96	1.72
Maize	3.57	0	5.29	2.22	0	3.89	2.77
Plantains	0.61	0.2	10.92	4.23	4.1	4.37	4.37
Rice	1.48	4.35	0	2.62	4.35	4.48	3.04
Sorghum	0.87	0.59	0	0.65	0	1.2	0.72
Sugarcane	19.51	0.4	1.09	1.23	0.09	7.68	4.16
Total	100	100	100	100	100	100	100

#### TCARI PARALLELEPIPED CLASSIFICATION

Overall Accuracy: 6.48%      Kappa coefficient: 0.0117

	Coconut	Maize	Plantains	Rice	Sorghum	Sugarcane	Total
Unclassified	77.84	82.63	37.84	70.15	46.89	71.52	67.62
Coconut	4.68	4.66	18.22	8.78	16.29	8.37	9.43
Maize	3.82	0.64	5.07	1.35	1.04	2.86	2.15
Plantains	4.49	1.69	12.77	3.14	5.18	4.9	4.47
Rice	2.48	3.39	11.46	6.43	11.49	4.2	6.15
Sorghum	2.67	5.3	6.57	7.73	16.67	4.57	6.98
Sugarcane	4.01	1.69	8.08	2.41	2.45	3.57	3.2
Total	100	100	100	100	100	100	100