OGC Stack Winter School,
15-31 Dec, 2022

Can Geospatial data and GeoAI help improve our understanding of Agroecosystem?

Prof. K S RAJAN

Professor and Head, Lab for Spatial Informatics, IIIT Hyderabad

Chairman, KAIINOS Geospatial Tech Pvt. Ltd.

ISRS, ISG, OSGeo, IEEE GRSS-CIS,
ISPRS-WG V/8, CSI-SIG-BDA, FSMI, Geo4All, ISSE

RAJAN@IIIT.AC.IN

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

About Me

- PhD in Land Use Land Cover Modelling (IIS/Seiken, 東京大学 U of Tokyo)
  - Integrating Biophysical (including Crop Modelling) and Socioeconomic to understand key Environmental Processes
- Started Research as an Environmental Process Modelling researcher
- Modelling → Analysis → Data Generation
- Data → Information → Knowledge
- Research Interests
  - From using "What is Available" → "What can we Extract from the Data" → "Improve our Understanding of Earth Systems"
    - Agricultural Cropping Practices;
    - Terrain Modelling;
    - Water Quality
    - Urban Landscapes and Urbanisation

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

Lab for Spatial Informatics - Research Overview

IoT and Air Pollution Monitoring / Evaluation
Remote Sensing and IoT for Water Quality

GeoSpatial Science and Technology

- Spatial Data Science/ Geo-AI
    - Spatio-temporal Analytics / Data Mining / Databases
    - Health, Crop Yield analysis, Climate, Crime
    - Platforms developed
- Geo-Visualization
    - 4D Flood Visualization
    - Deep Learning on Terrain Super Resolution
    - 3D realistic terrains/topography
- Geo-Governance - Web/Mobile GIS
    - Air Pollution and Human Health
    - Water Quality and Quantity

FOSS4G

Environmental / Policy

- Optical Imagery
    - Feature Extraction – Roads, Buildings
    - Time-series Analysis – Crop phenology & Season Calendar
    - OBIA
- HyperSpectral Sensing
    - Key Parameter Characterisation – Vegetation
- RS – Parameter Estimation
    - Water Quality in Inland waterbodies

Remote Sensing

- ✓ Spatial Modelling
- ✓ Strategic Environmental Assessments
- ✓ River Basin Hydrology
- ✓ Climate Extremes and Impacts

Spatial Modelling & Simulations

Land Use Modelling – Urban, Agricultural
Forest Dynamics: Crop Modelling

Faculty at LSI:


Prof. KS Rajan


Dr. RC Prasad


Dr. S Rehana


Prof. R Nagaraja

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

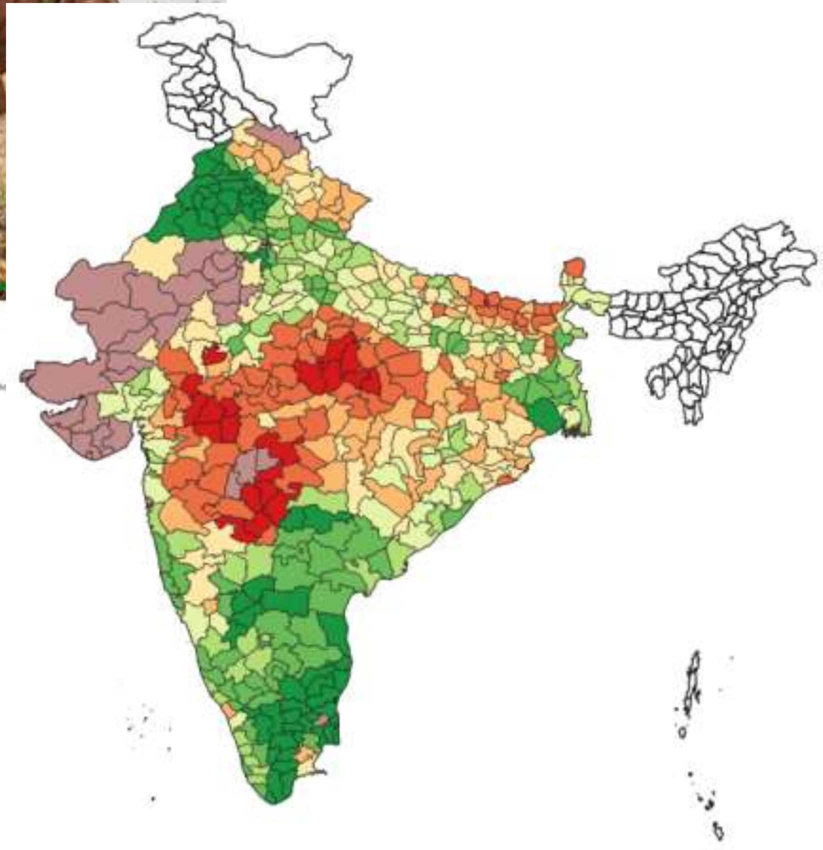
Related Ongoing works

- Spatial Data Science
  - Spatio-Temporal Mining of Crop Yields - MiSTIC algorithm to identify regions of Good/Bad Performance
    - How Analytics and Visualizations helps

- Spatial Data Generation and Knowledge Discovery
  - Data Analysis / Pattern Extraction with a Purpose
  - Crop Growth pattern and Crop Mapping – What does it tell us about Extreme Events?

- AMRT Dashboard
  - Integrating the data and services over a Geo-platform

INTERNATIONAL INSTITUTE OF
INFORMATION TECHNOLOGY
HYDERABAD



Spatio-Temporal Data Analytics and Visualization of Crop Yields

- Are all FOOD PRODUCING Regions/Districts in India Sustainable?
- How can we Assess Food Production / Crop Yields ?
- Are current methods/tools useful/sufficient to do this?

- Are they consistent Performers over Space and time?

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY

HYDERABAD

FACTORS AFFECTING YIELD

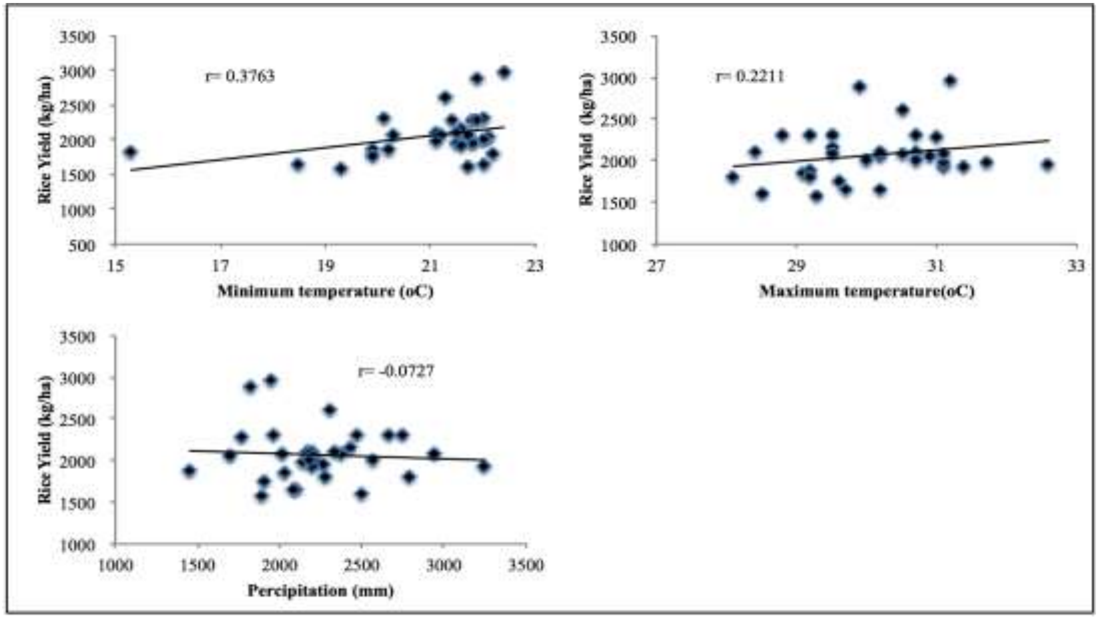
- Ecological factors
  - Availability of water resources (Rainfall, Irrigation mechanisms)
  - Type and quality of Soil
  - Temperature
  - Fertilizers, etc.

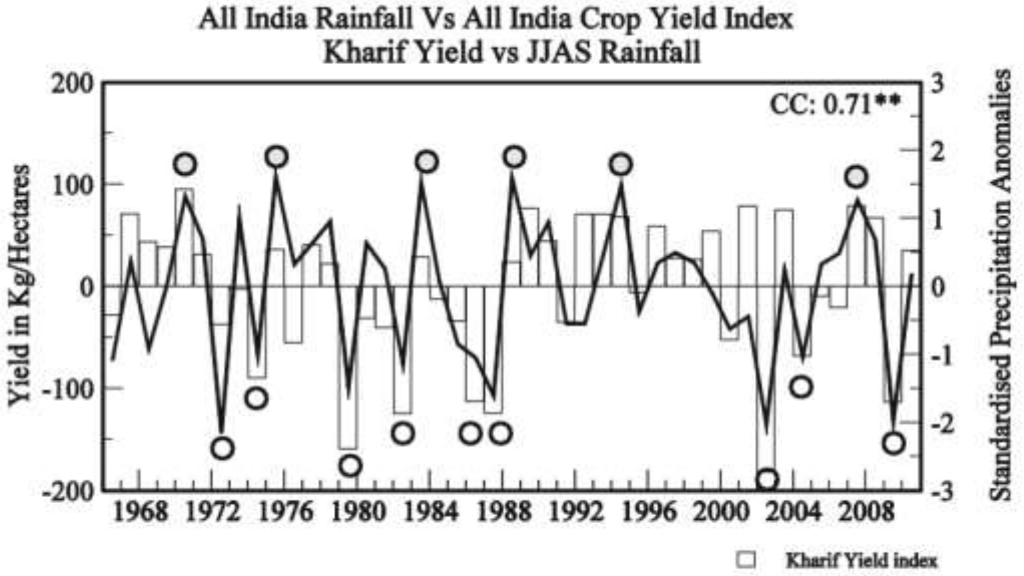
- Non-ecological factors – Management Practices
  - Cropping practices
  - Allocation of funds and resources
  - Training and education by government
  - Other socio-economic factors

Source: https://seepnetwork.org/Thematic-Areas-Agriculture-Food-Security

EXISTING METHODS

- Input based studies
  - Sensitivity analysis; Cause-Effect Models
  - feature -v- feature based
  - Ex.: changes in yield with rainfall/soil etc.





Source:
1) Relationship between Climate Variability and Crop Yield, Shobha P., R. Shaw
2) Impact of monsoon rainfall on the total foodgrain yield over India, Prasanna V.

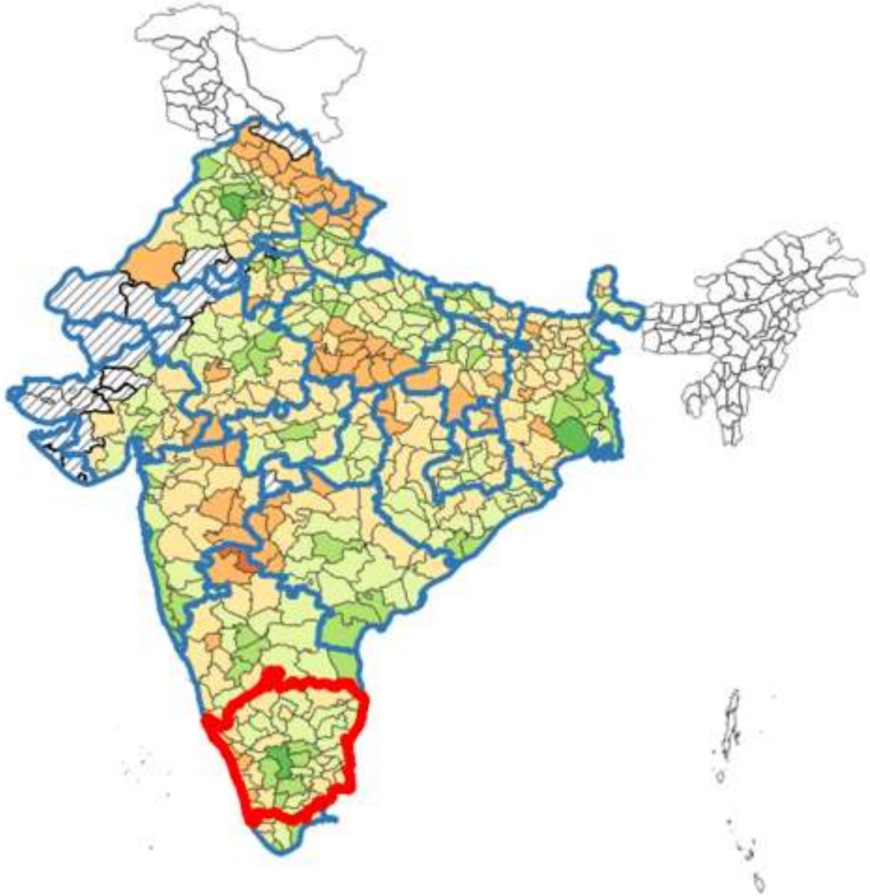
INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

FINAL DATA FOR ANALYSIS

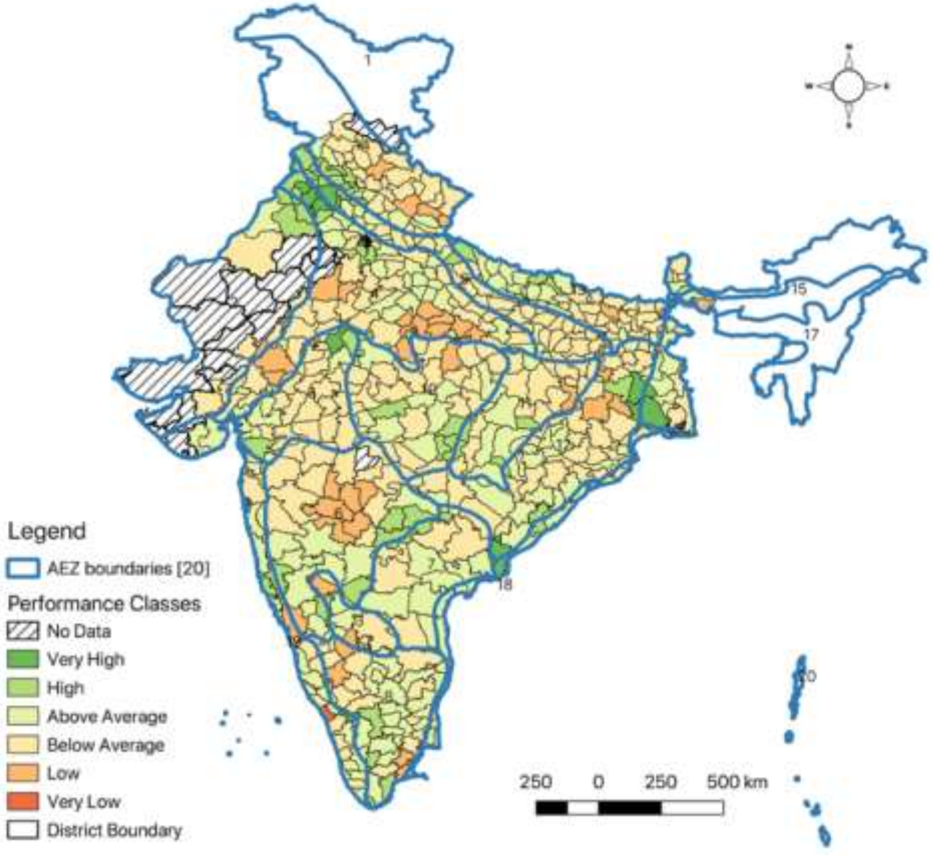
- Inputs: Crop yield for rice in kg/Ha (499 districts; 2000 – 2010)

	A	B	C	D	E	F	G	H	I	J	K	L
1	Name	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
2	Adilabad	2069.99788	2231.99917	1031.57726	2467.99642	1519.01114	2571.99353	2398.99294	2400	2440.99732	1987.01093	1672.02511
3	Agra	1892.58699	1878.60395	1675.48501	1742.71845	2347.31085	2266.40549	1893.186	1980.64516	2198.99425	1758.09801	2167.89498
4	Ahmadabad	1368.78307	1759.7561	1379.56172	2375	2230.67174	2484.17722	1837.61562	2035.1711	1826.22269	1885.61321	2643.47826
5	Ahmadnagar	690.909091	1269.63718	722.640917	776.31579	725	1156.94943	785.714286	1244.94842	1076.40355	1110.80507	1143.07141
6	Ajmer	935.483871	1254.90196	1000	1653.84615	1500	1461.53846	1600	2000	1750	1571.42857	2000
7	Akola	0	0	0	0	0	0	0	0	0	0	0
8	Alappuzha	3001.82763	2725.09847	3096.50709	1801.26901	2352.43823	2773.24393	2928.6653	2125.33572	3027.24646	3026.16124	2386.38311
9	Aligarh	1972.26391	1935.49058	1760.99858	2058.24737	2251.66098	2086.98957	2060.99306	2007.00688	2154.00471	1732.00096	2041.00178
10	Allahabad	1829.9563	1999.55519	1902.00269	2332.56035	1790.65999	2016.00043	2115.00222	2152.00198	2443.99819	2169.99649	2325.00212
11	Almora	986.195239	1254.78353	669.002854	1023.7374	979.667972	1037.14878	1254.04889	1139.46814	983.750549	980.350278	1030.68311
12	Alwar	936.868687	1248.17518	807.692308	1656.25	1482.01439	1462.12121	1574.80315	2032.25807	1807.54717	1514.45087	2026.45503
13	Ambala	3020.60134	3111.60375	2977.30416	3099.39924	3166.62137	3318.52626	3446.13597	3685.27487	3296.21491	3718.87267	3444.52723
14	Ambedkar N.	2473.65711	2418.69514	2137.83957	2617.86979	2007.93555	2384.99572	1797.99968	2551.00044	2700	2590.00251	2637.99756
15	Amravati	494.382023	941.747573	654.545455	588.235294	752.14845	1047.31362	541.176471	825.396825	967.465993	1441.86047	736.842105
16	Amreli	0	1000	0	0	0	0	0	0	0	0	0
17	Amritsar	3047.02194	3003.1348	2826.36656	2674.84663	3095.80838	2907.24638	5993.57018	6219.46251	6013.6127	5789.50346	5611.27782
18	Anand	0	0	0	1984.70097	2177.8607	2168.89429	1729.28177	2326.36469	1977.32426	2562.5	2111.74458
19	Anantapur	2949.00388	2907.99319	2282.01335	2563.00314	3495.00719	2797.98465	3017.003	3166.66667	3121.98905	3150.0031	2916.73577
20	Anugul	902.766515	1654.86726	932.186647	1513.76147	1340.65934	1647.05882	1422.22222	891.566265	1378.04878	1072.28916	1436.20951
21	Araria	1308.70197	2130.80331	1160.64594	1415.5215	1004.53401	1083.30659	930.174276	1127.21939	1157.98146	1223.40131	991.339189
22	Ariyalur	0	0	0	0	0	0	0	0	0	3097.31311	3881.99595
23	Auraiya	1834.95126	2433.85596	2207.93123	2268.08826	2821.36087	2518.9962	2769.00591	2740.99318	3067.99792	2823.01009	2684.00496
24	Aurangabad	1250	0	1894.21589	1807.00376	1141.8786	1340.33603	2598.27152	2516.10404	1449.3942	1409.13432	1416.83976
25	Aurangabad	666.666667	600	771.374314	827.477224	666.666667	1000	1000	0	0	0	0
26	Azamgarh	1852.58564	2068.31605	1684.8002	2033.27457	883.423609	1752.99908	1468.99781	1929.9981	1967.99809	1746.00204	2050.9979
27	Bagalkot	1818.71345	2403.50877	2053.57143	2258.06452	2792.96875	2817.46032	2897.4359	3029.41177	3058.96552	2869.02256	3206.61157
28	Bageshwar	1965.05376	1232.74106	987.189092	1506.18949	1331.54395	1420.72744	1229.88589	1270.55786	1375.25232	1343.89104	1205.76642
29	Baghpat	2691.71289	2282.06421	2228.08268	2705.43375	2618.88545	2468.10631	2596.89041	3024.11576	2315.94999	2716.95008	2351.03825

Comparison of Crop Yields with different zonation approaches



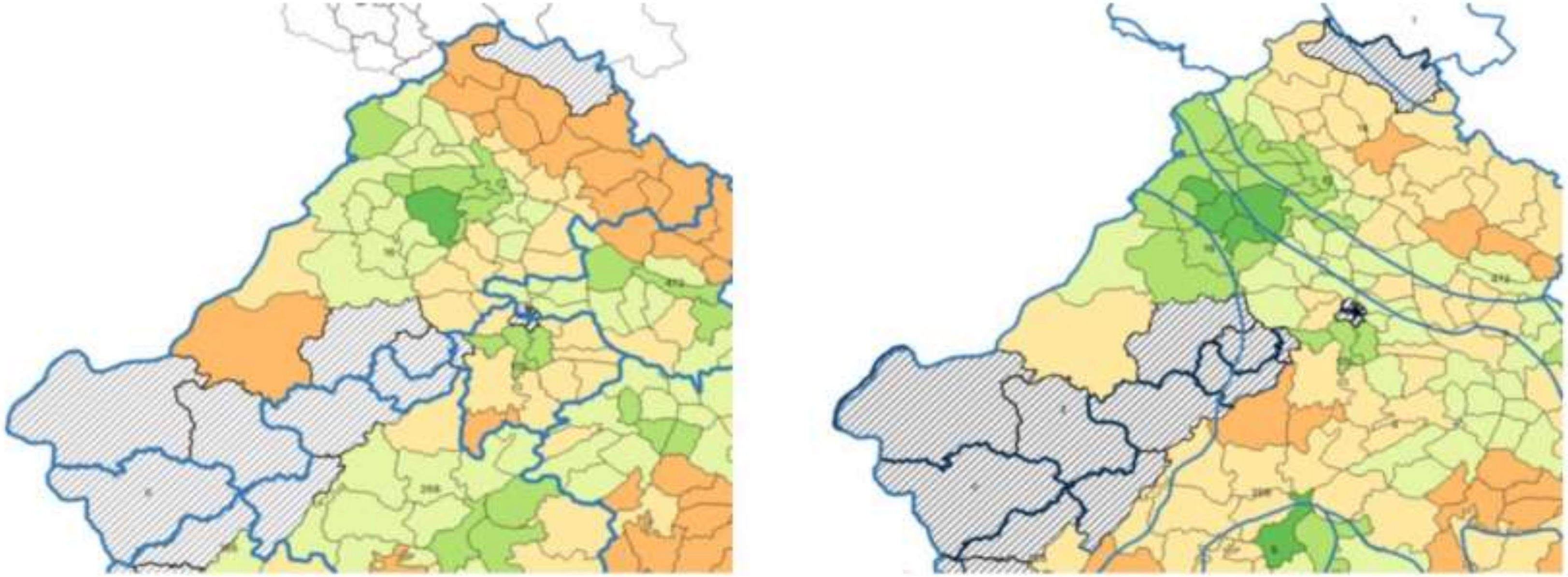
MiSTIC maps



AEZ maps

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

Which of these better explains the performance pattern ?



(a) MiSTIC Zone

(b) Agro-Ecological Zone

Figure 5.16: Comparison of performance classes in MiSTIC v/s AEZ

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

But, What we need ?

- Have been helpful in showing correlations:

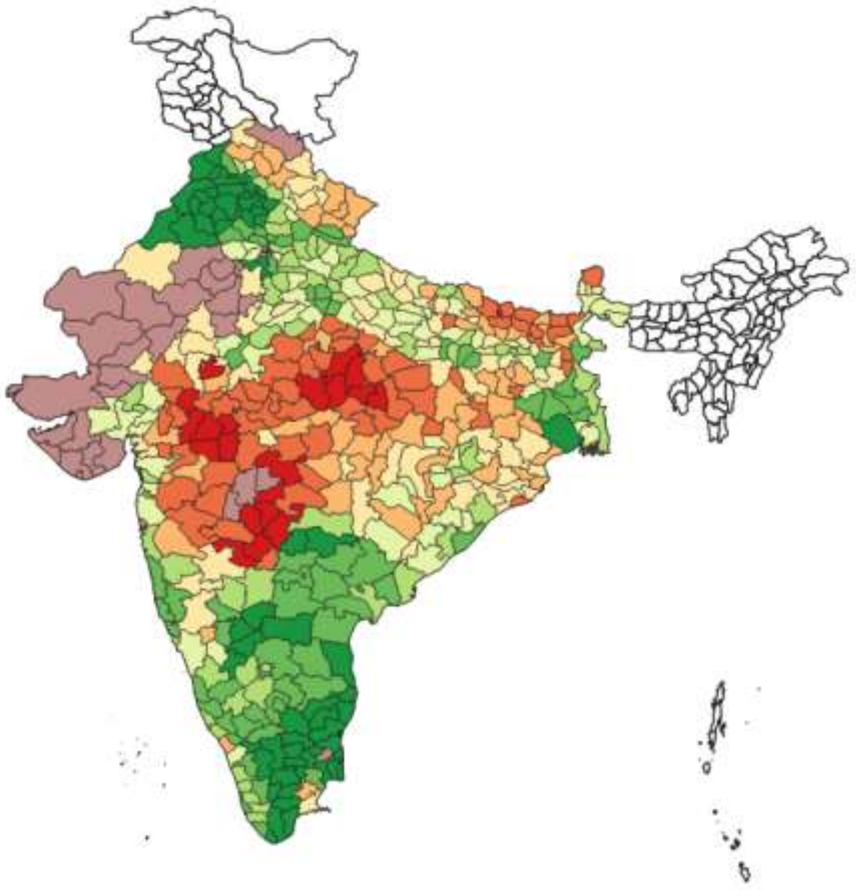
But,

- Not a complete representation of yield values
  - Limited to that factor (single/multi variable); Cannot generalize
  - Other features may feed into actual output yield numbers
- **Spatial and Temporal characteristics not well captured**
  - **Spatial: Neighborhood effects**
  - **Temporal: Periodic cycles**
- Inferences bases on direct correlation may or may not be effective
- Complex to formulate an approach with each additional feature
  - Unable to fully account for the factors
- Modelling challenges :
  - Scale
  - Correctness

INTERNATIONAL INSTITUTE OF
INFORMATION TECHNOLOGY
HYDERABAD

STUDY AREA

- District level Crop Yield data of India
- Crop: Paddy
- Season: Kharif (Rainy Season)
- Spatial Coverage : 499 districts
  - Newly created districts merged into old ones
  - taking districts from the year 2000 as reference
- Temporal Coverage : 11 years [2000 – 2010]
  - Based on availability
- Total data points: 499 * 11 = 5489



- Source of Data: https://data.gov.in/catalog/district-wise-season-wise-crop-production-statistics

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

METHODOLOGY

STEP I. SPATIO-TEMPORAL (ST) PARTITIONING

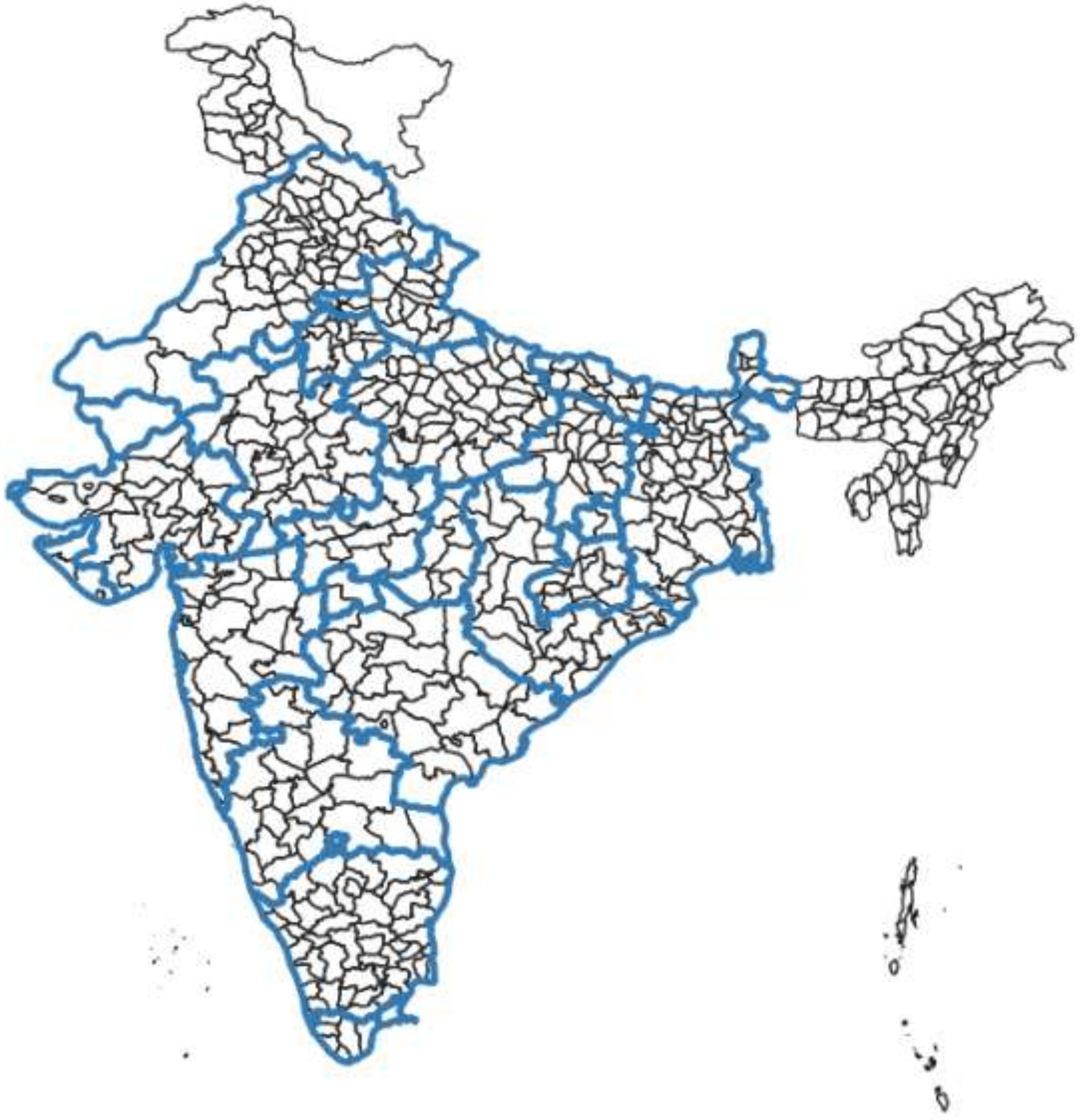
- Segregate the area into contiguous zones
  - Zones comprising of a set of districts
  - Each zone captures intrinsic ST interactions

- How ?
  - Using a ST data mining algorithm: MiSTIC
    - **Mi**ning **S**patio-**T**emporally **I**nvariant **C**ores

MiSTIC Zones

STEP I. SPATIO-TEMPORAL PARTITIONING

Figure: 18 zone boundaries obtained from MiSTIC

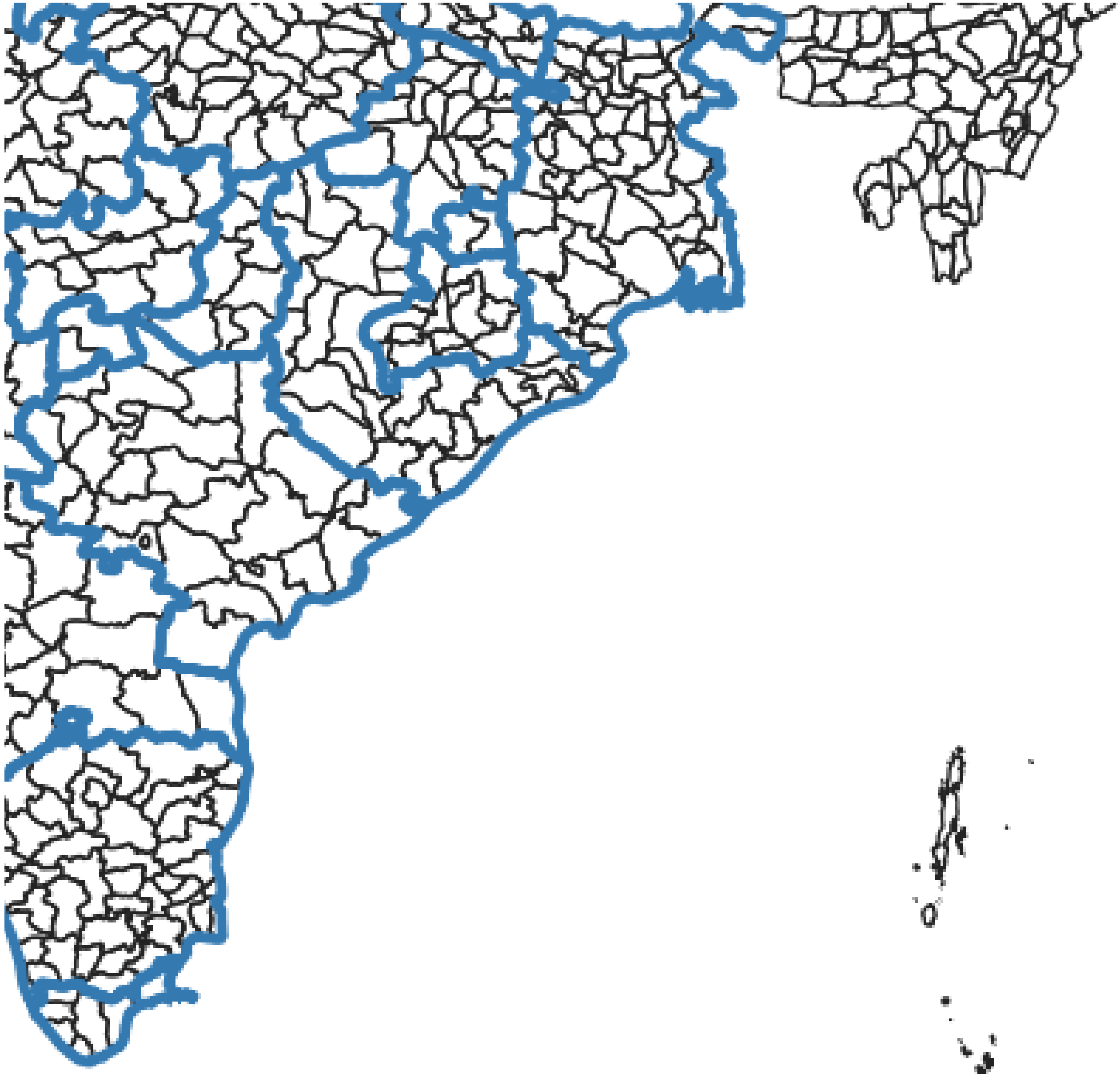


INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

MiSTIC Zones – based on Crop Yield distribution

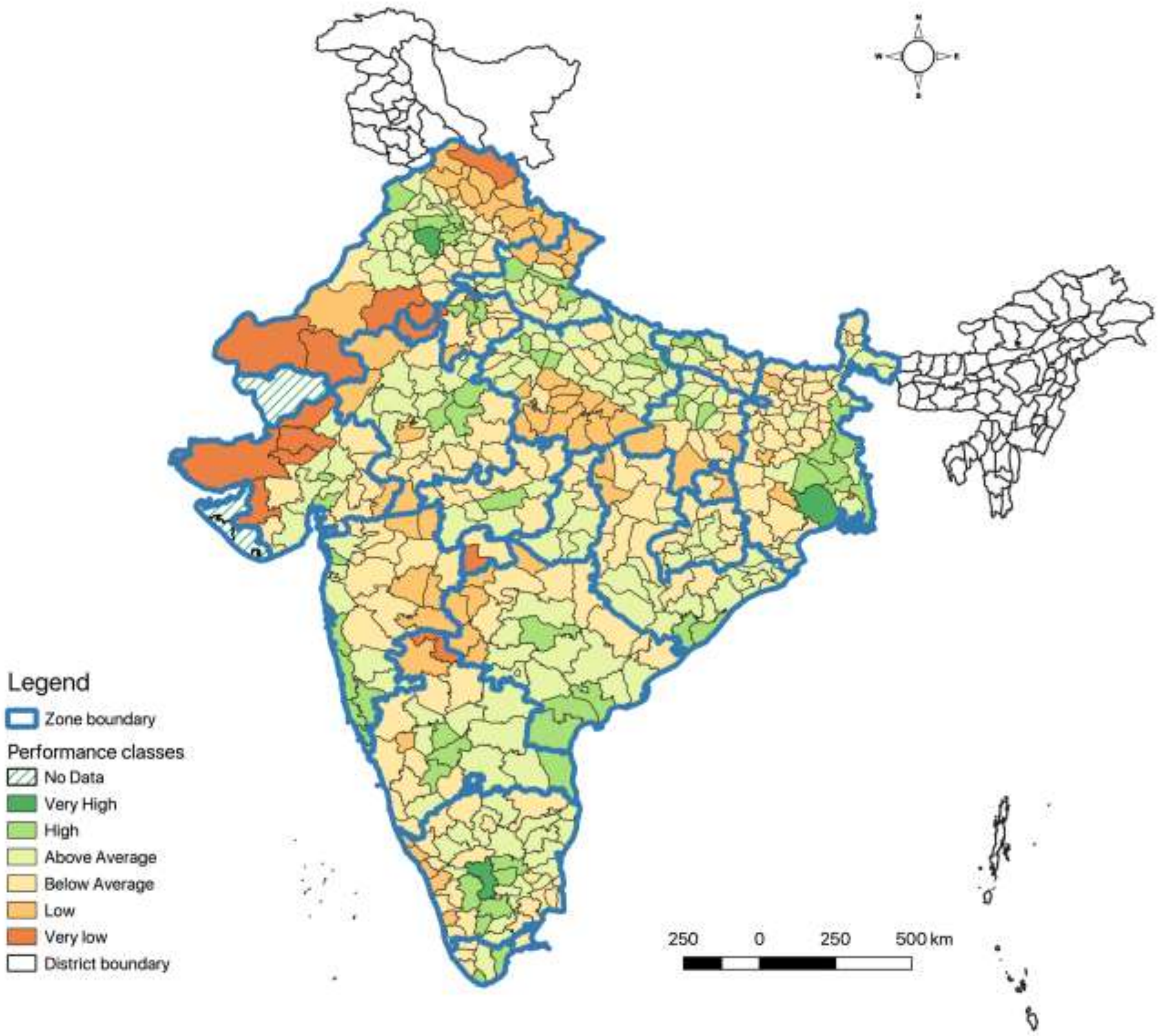
STEP I. SPATIO-TEMPORAL PARTITIONING

Figure: Close-up of zone boundaries



INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

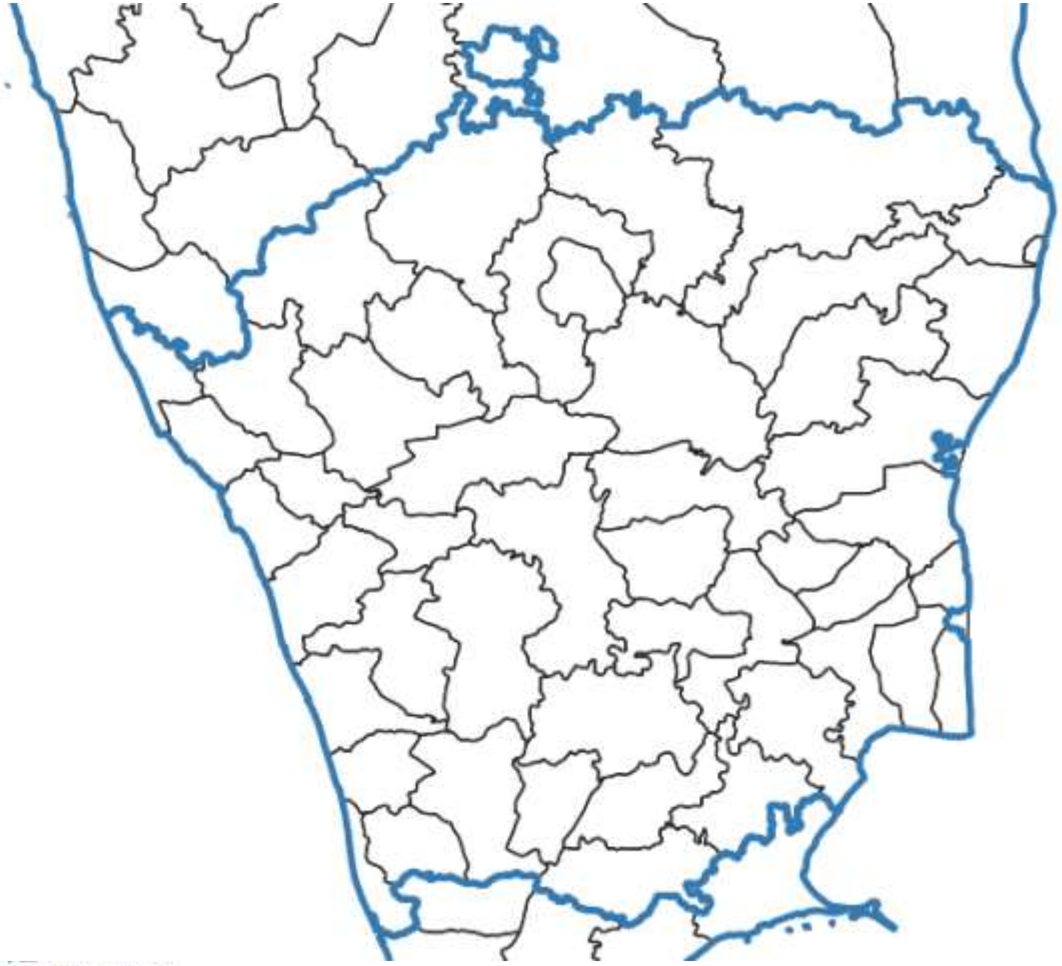
PERFORMANCE CLASSES WITHIN ZONES

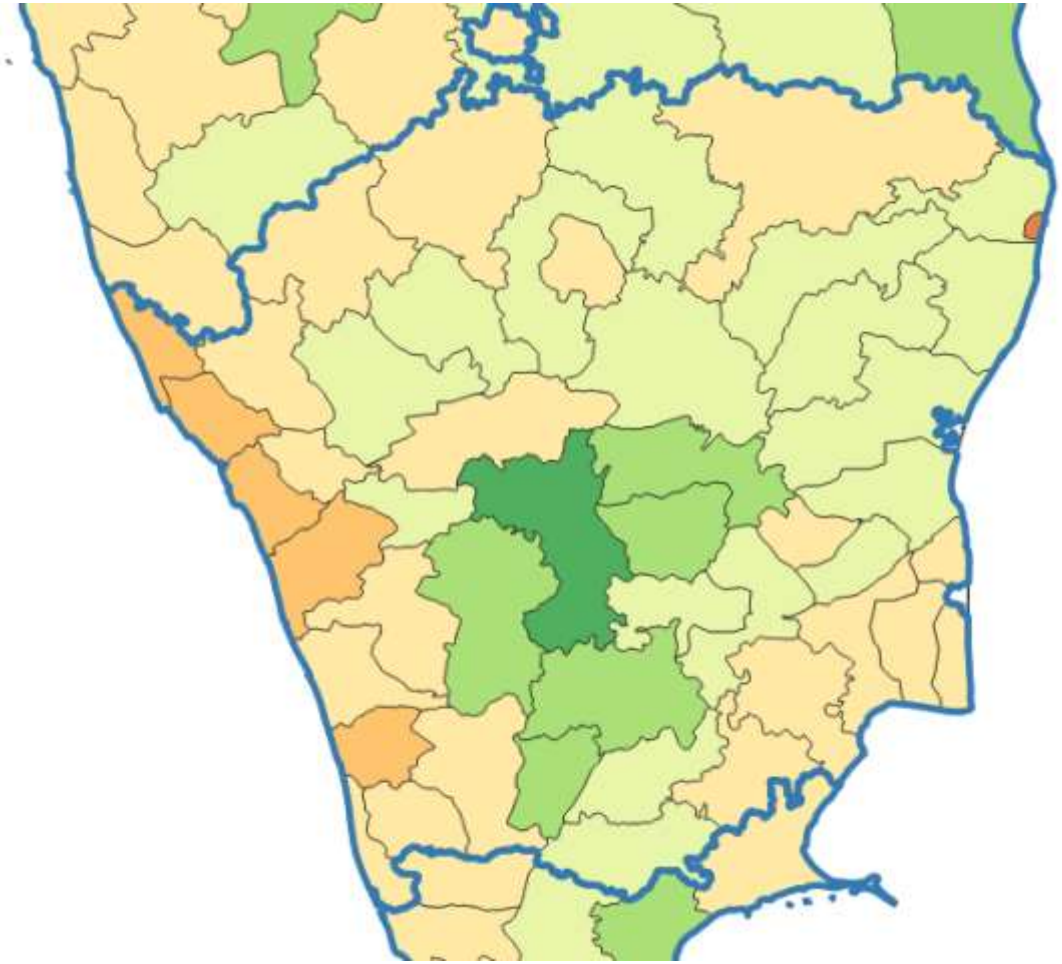


INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

PERFORMANCE CLASSES WITHIN ZONES
(zoomed view)

R2. Classification for a zone





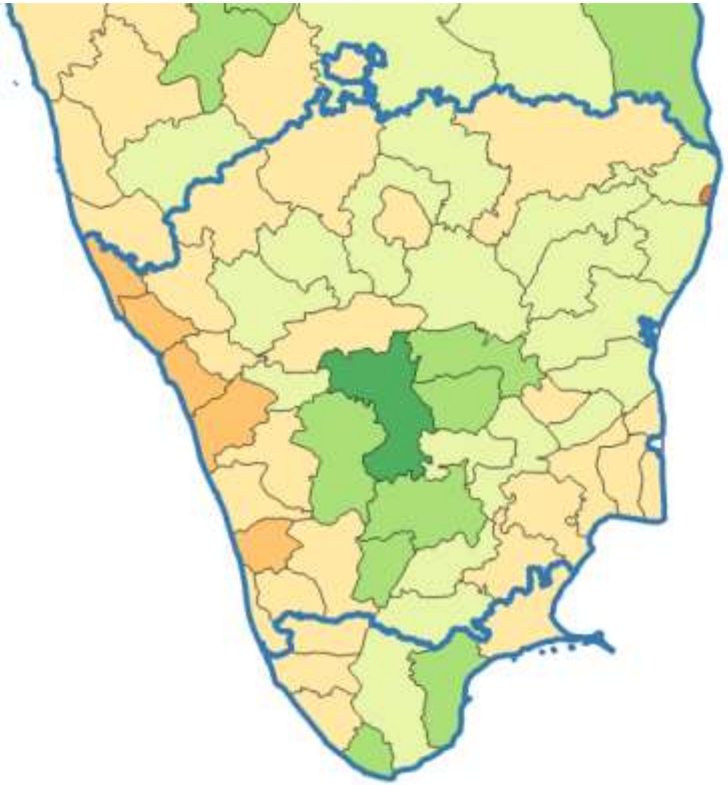
INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

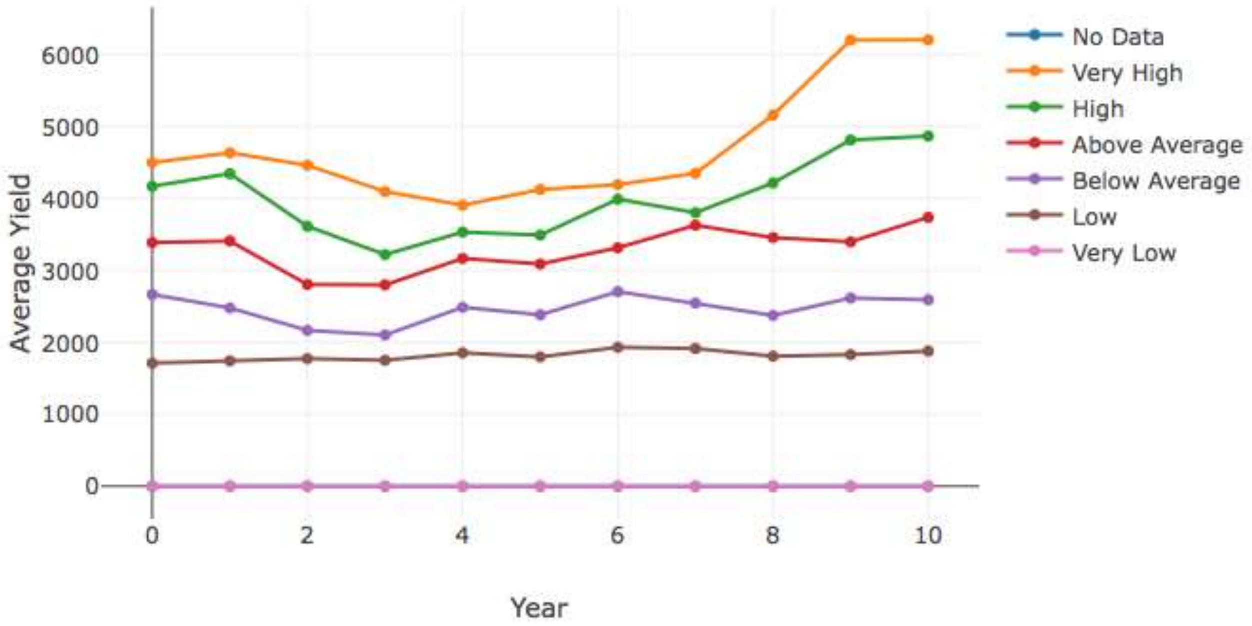
Re-Analysis of TEMPORAL TRENDS

- For a specific zone,
  - We have performance classes
  - i.e., groups of homogeneous districts
- Performance may
  - remain consistent
  - Increase/decrease
- Observe effects on performance of each class
  - Drought/Flood years
- Within a zone
  - For the districts of a performance class
  - Plot their annual average with time

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

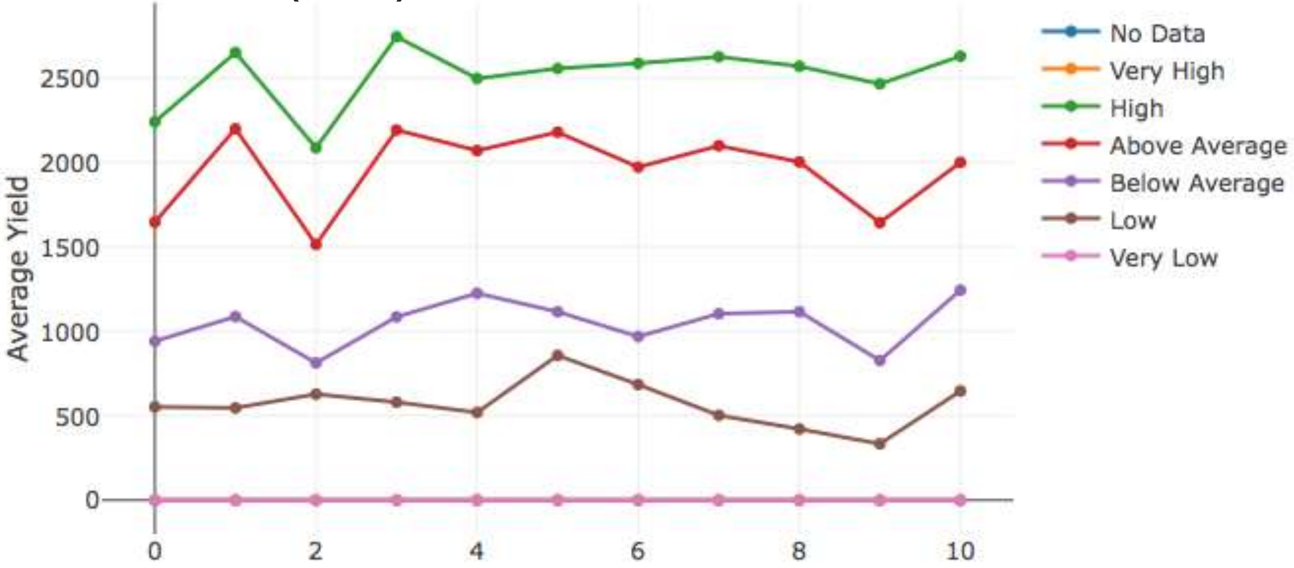
TEMPORAL TRENDS (1/2)





INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

TEMPORAL TRENDS (2/2)



Class ID	No of districts	Min (in kg/ha)	Max (in kg/Ha)
1	0	-	-
2	6	1889.2	3171.4
3	7	1080	2827.4
4	9	489.8	2420.6
5	5	319.1	1014.4
6	0	-	-

Table 6.3: General statistics of performance classes of zone ID 379

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

CONCLUSION

- Good and poor performing districts
  - Preserve High Performers
  - High performers close to the center of the zones
  - Respond appropriately to low performing regions for a crop
    - Might become unsustainable over time

- Generic method of analysis
  - Can be done for any crop
  - Different hotspots and different zones

- Scale Independent
  - Better results if done on spatially contiguous sub-district level data

Related Ongoing works

- Spatial Data Science
  - Spatio-Temporal Mining of Crop Yields - MiSTIC algorithm to identify regions of Good/Bad Performance
    - How Analytics and Visualizations helps

- Spatial Data Generation and Knowledge Discovery
  - Data Analysis / Pattern Extraction with a Purpose
  - Crop Growth pattern and Crop Mapping – What does it tell us about Extreme Events?

- AMRT Dashboard
  - Integrating the data and services over a Geo-platform

What is needed for a Good Monitoring System ?

- A good baseline data
  - Coverage, periodic updates, record of causes of changes, if any
- Is Crop-calendar a good baseline?
  - esp. if it is one calendar for the whole district
- What about uncertainties in the crop calendar?
- Can Phenology provide the right clues?

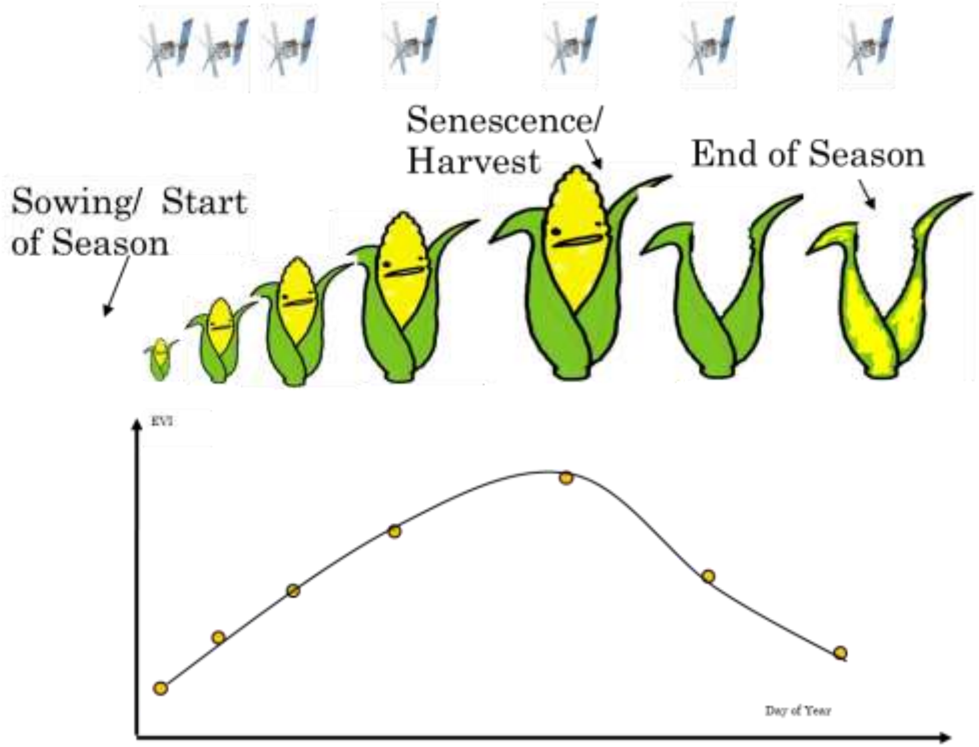
How Events like droughts affect Cropping patterns in a region?

- ➢ All are areas affected similarly ?
- ➢ Can such analysis help us identify the **Causative and underlying factors**?

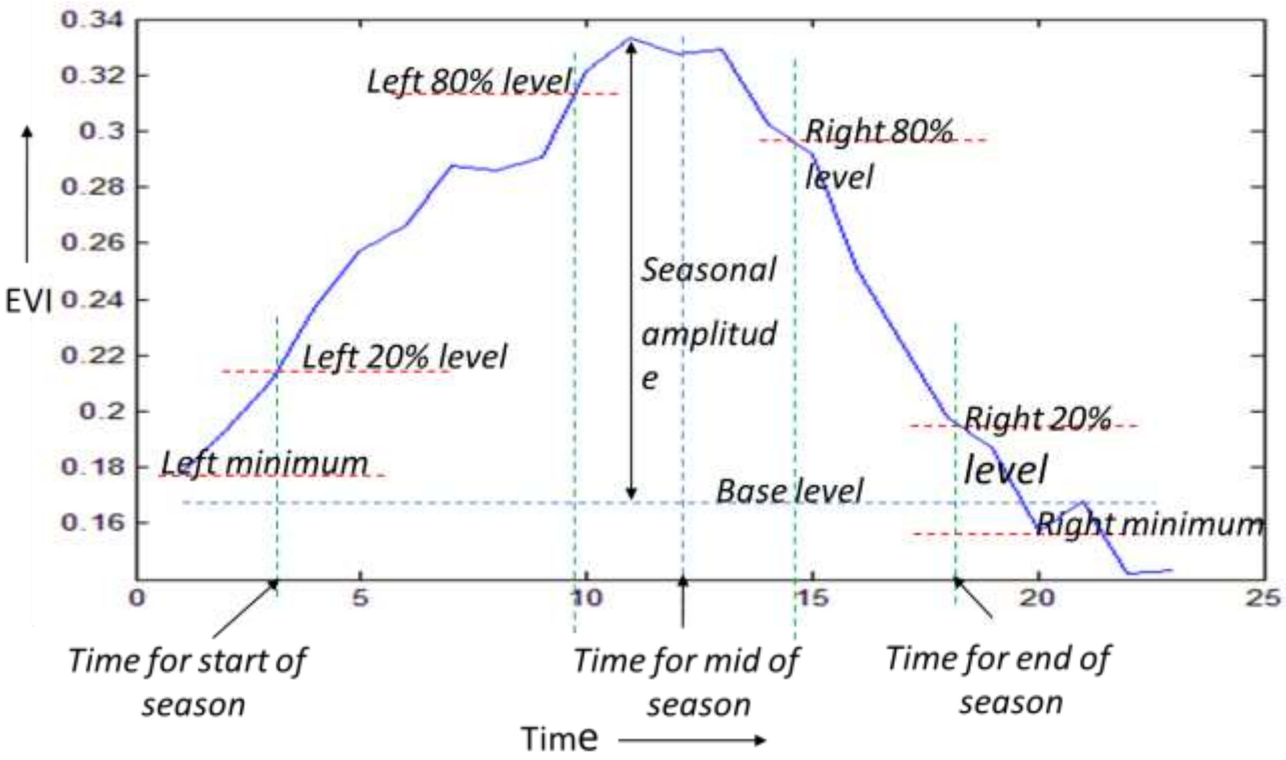
INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

Generation of Location specific Crop Phenology – Pixel wise Crop Calendar

EVI and Vegetation Profile



Ideal EVI Time Series



INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

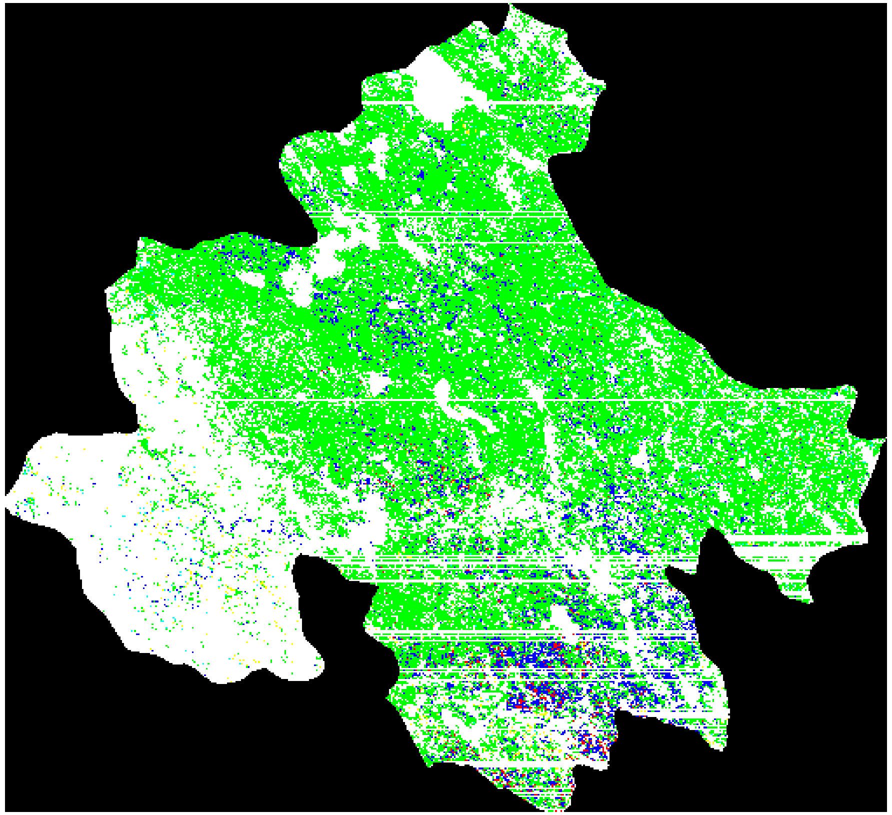
Season calendar results for Chamrajnagar

Type	Start	End	Mid	10-Jun	26-Jun	12-Jul	28-Jul	13-Aug	29-Aug	14-Sep	30-Sep	15-Oct	1-Nov	17-Nov	3-Dec	19-Dec	1-Jan	17-Jan	2-Feb	18-Feb	4-Mar
1	12-Jul	15-Oct	14-Sep																		
2	1-Nov	4-Mar	1-Jan																		
3	15-Oct	18-Feb	19-Dec																		
4	15-Oct	17-Jan	3-Dec																		
5	30-Sep	1-Jan	17-Nov																		

Derived season calendar for Chamrajnagar District

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

Crop Season Map :
SVM and DTW/CDTW based Time Series Classification



Hassan District in Karnataka

- Kharif + Rabi + Zaid
- Kharif + Rabi
- Kharif + Zaid
- Kharif only
- Rabi + Zaid
- Rabi only
- Zaid only

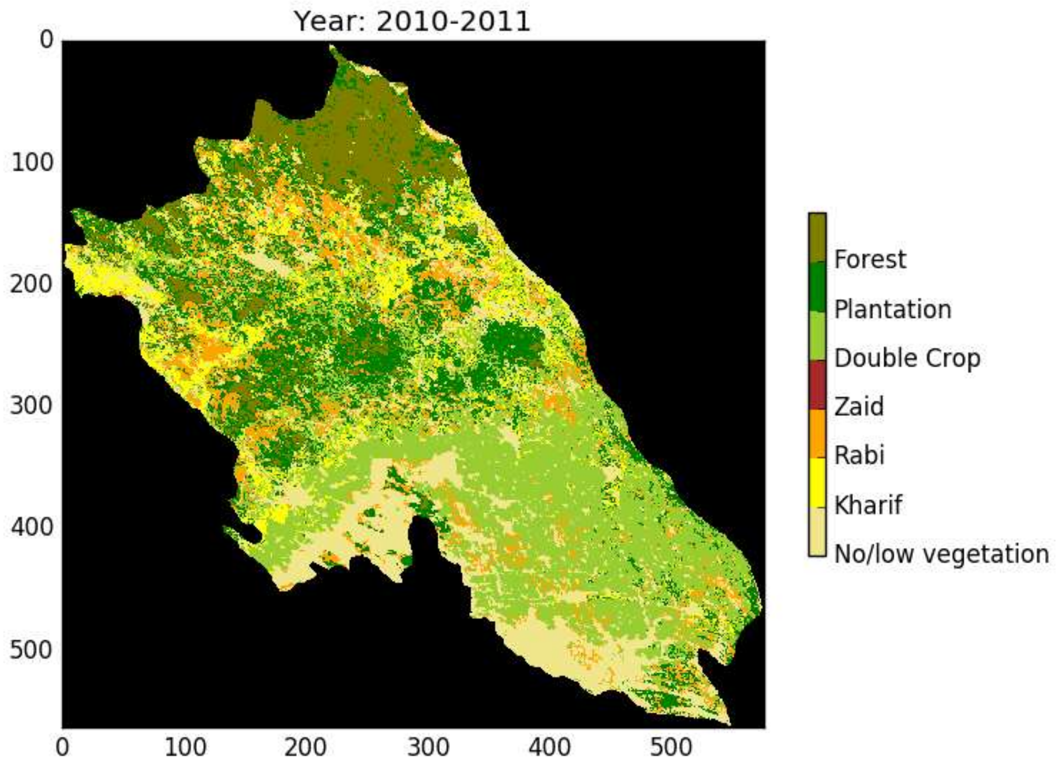
	Statistics	Our results
	hectares	hectares
Kharif	393481	367010
Rabi	66244	40756

S Gupta and K.S.Rajan. 2011. Extraction of Training Samples from Time-Series MODIS imagery and its utility for Land Cover Classification. International Journal of Remote Sensing - 32 (24) 9397-9413.
Gupta, S., and K. S. Rajan. 2010. "Temporal Signature Matching for Land Cover Classification." International Society for Photogrammetry and Remote Sensing – Technical Commission VIII Symposium, August 9–12

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

Results in West Godavari District, India

- **Vegetation Cover Map**
  - The time-series curve of **each pixel** was labelled using the proposed approach.
  - Vegetation cover map generated for the
    agricultural-year **2010-11** over the West
    Godavari district.
  - Shows the **extent and spatial distribution**
    of different vegetation covers.



INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

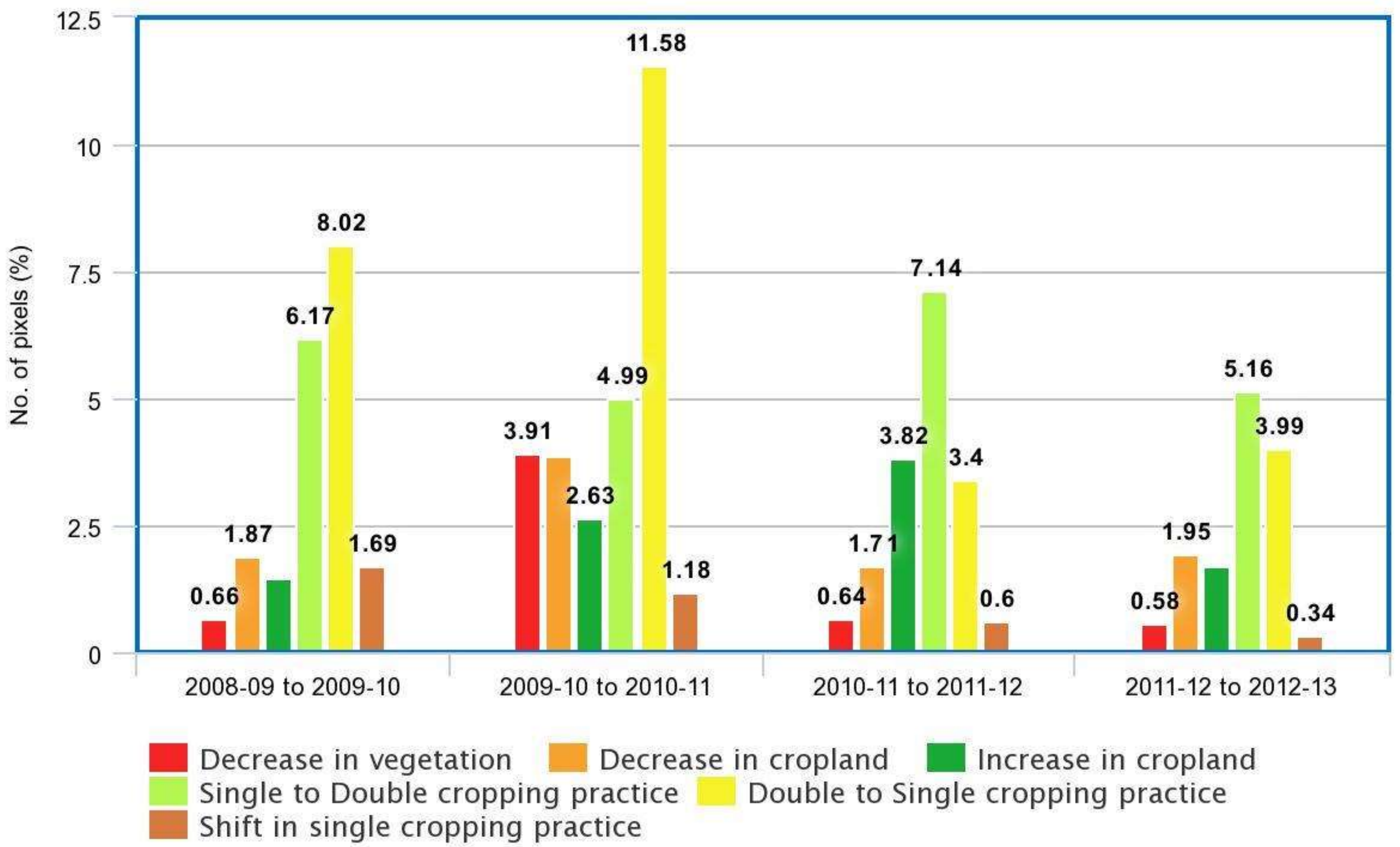
Change Detection

- Performed classification for four agricultural-year periods between **2008 and 2012.**

- To capture **significant changes** in agricultural practices and corresponding region of interests, we define seven major changes:
  - (1) Decrease in vegetation - regions that lost their forest or plantation cover.
  - (2) Decrease in cropland - agricultural regions that converted to fallow land.
  - (3) Increase in cropland - fallow regions that started agriculture.
  - **(4) Single to Double (5) Double to Single**
  - **(6) Shift in single (monsoon to winter crop or vice versa) - denoting shifts in cropping practices, and**
  - (7) No changes - regions that showed no significant change.

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

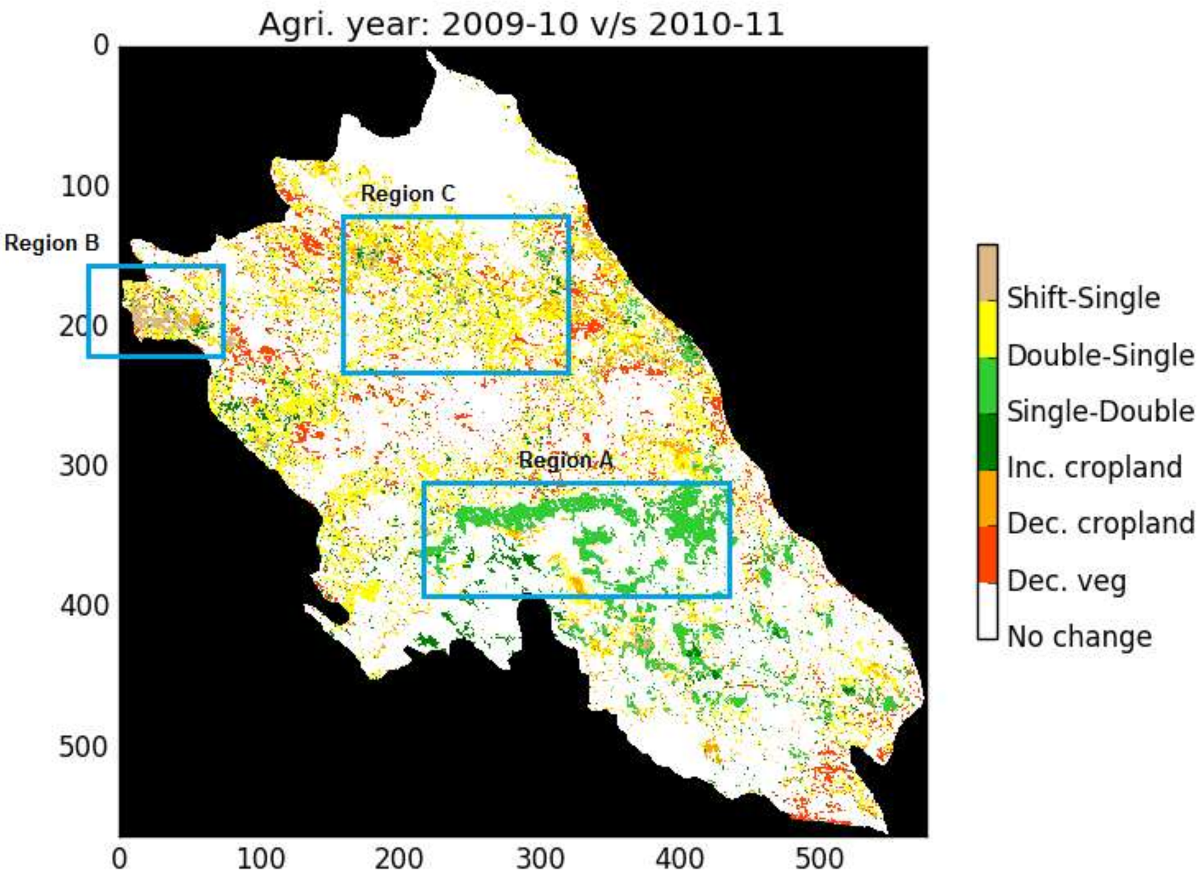
Table 2.a Change analysis for each transition (for agriculture years from 2008-09 to 2012-13).



INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

Analysis of Cropping Transition
from agricultural-year 2008-09 to 2011-12 (4 crop years)

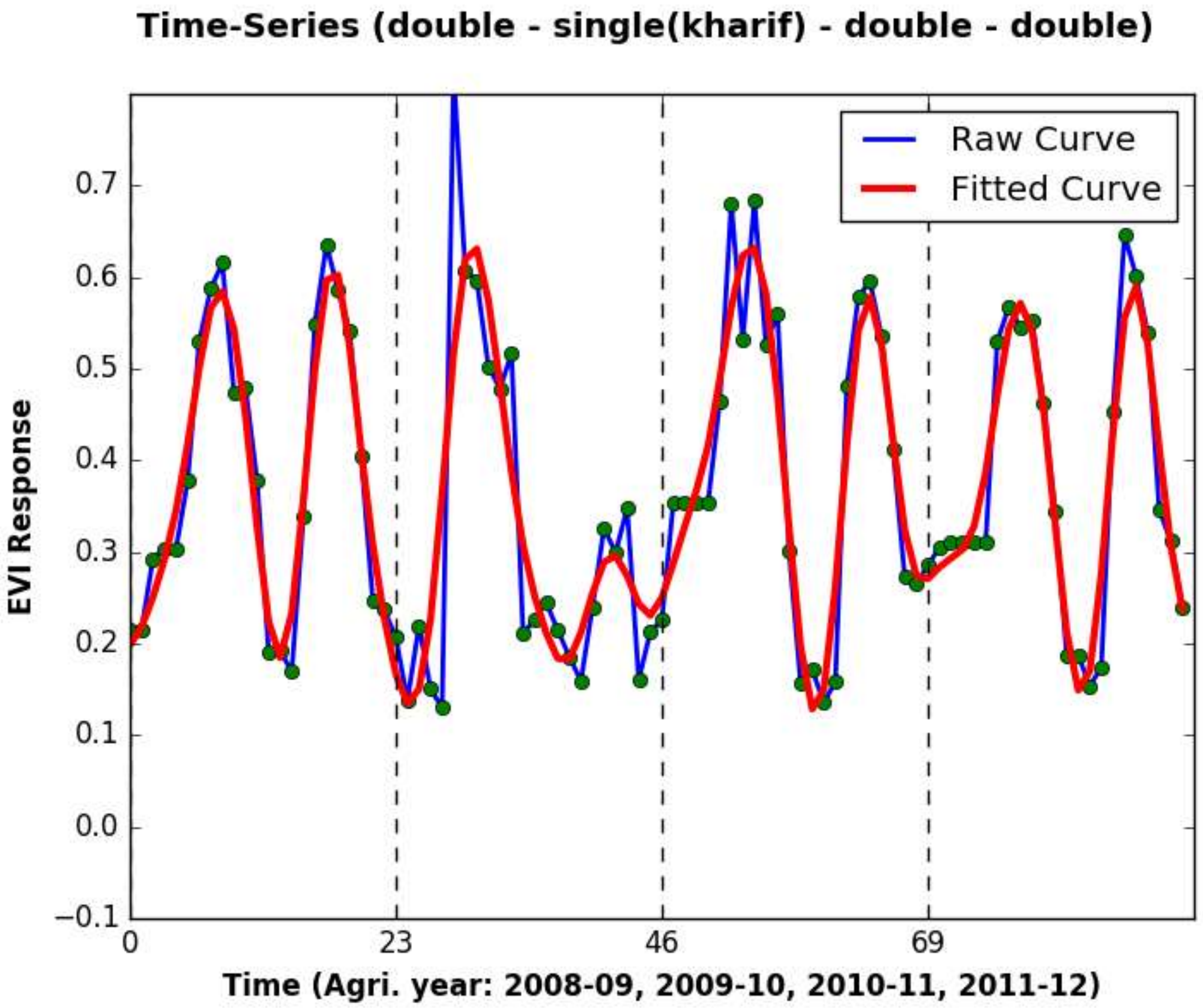
- Three **Spatial clusters** are analysed here -
  - Regions that showed significant changes in their corresponding vegetation covers.
- 2009 was a Drought year – 22 to 42% rainfall deficiency

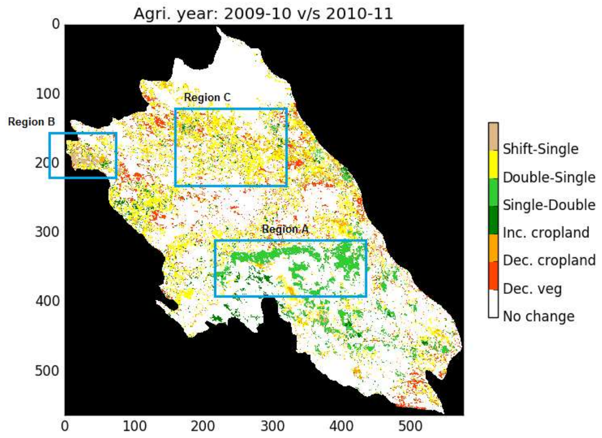


INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

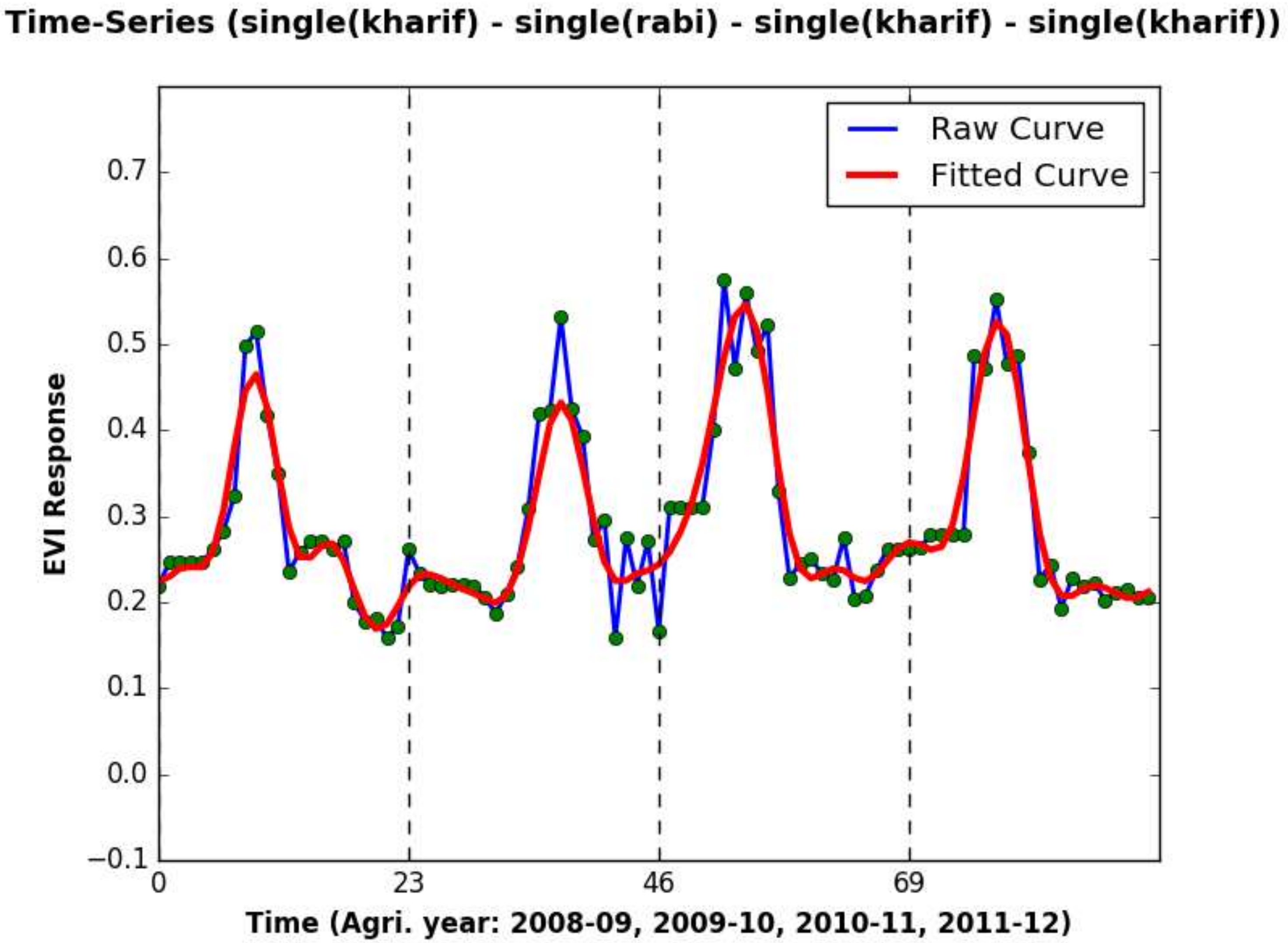
D Baheti, KS Rajan. 2017. A Shape-Based Approach to Spatio-Temporal Data Analysis Using Satellite Imagery. Proceedings of the 4th IEEE International Conference on Data Science and Advanced Analytics (DSAA), Tokyo, Japan. Oct 19-21, 2017.

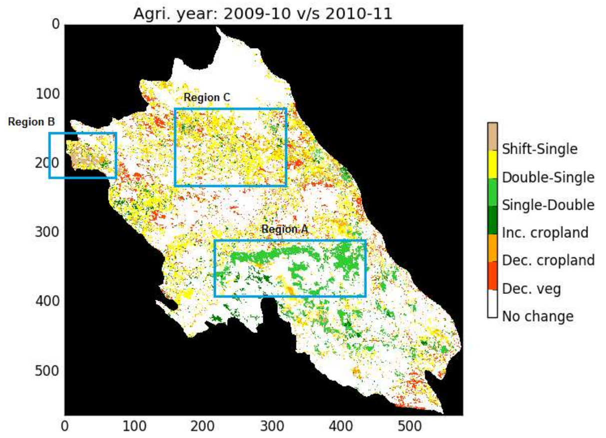
Region A: Downstream location, water surplus



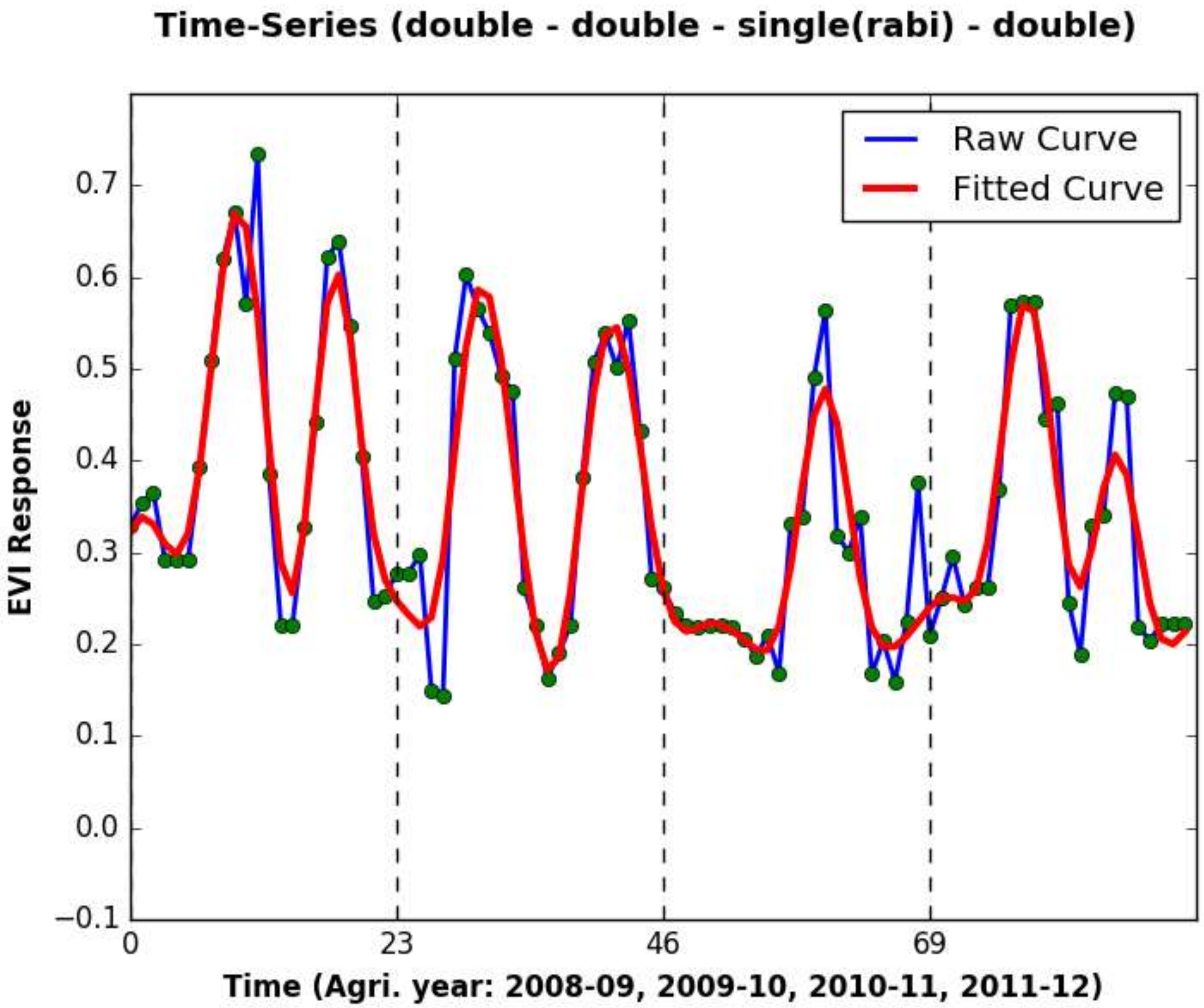


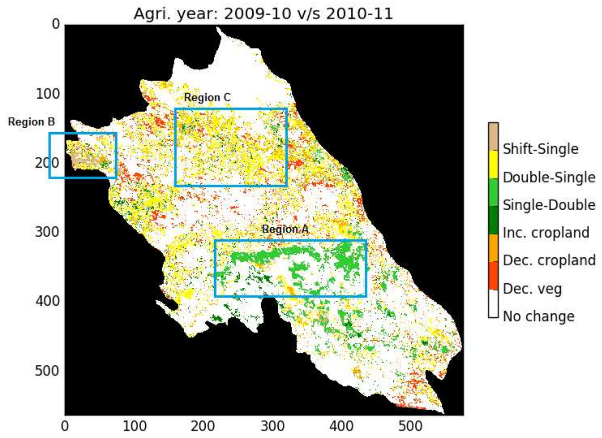
Region B: Away from River, Rainfed-Ground water irrigation





Region C: Upstream, Mostly Irrigation (canal/ground water) based





Drought Event impact on Year-on-Year Growth Pattern

- Detected spatial clusters - regions under which **similar trends** and changes are exhibited.
    - **Region A and Region B**, where anomalies were detected in **agri. year 2009-10**.
    - **Negative Impact.**
        - Region A: (double - single - double - double)
        - Region B: (monsoon - winter - monsoon - monsoon)
    - Abrupt changes (in region A and region B) were triggered due to **lack of irrigation water** in drought year. **Rainfall dependent regions**

    - **Region C** showed different trend (anomaly detected in **agri. year 2010-11**) capturing two major changes :
    - **Delayed** Negative **Impact**
        - (double - double - single - double) and
        - (winter - winter - fallow -winter)
    - May **not be rainfall dependent** and the **source of irrigation could be ground water**. Time lag in Impact as **rainfall patterns can affect groundwater levels**.

- Indian Meteorological department, reported agricultural-year **2009-10** as a **drought year.**

Related Ongoing works

- Spatial Data Science
  - Spatio-Temporal Mining of Crop Yields - MiSTIC algorithm to identify regions of Good/Bad Performance
    - How Analytics and Visualizations helps
- Spatial Data Generation and Knowledge Discovery
  - Data Analysis / Pattern Extraction with a Purpose
  - Crop Growth pattern and Crop Mapping – What does it tell us about Extreme Events?
- AMRT Dashboard
  - Integrating the data and services over a Geo-platform

An Example of an Ongoing Project which brings together the Analysis onto a Geo-platform for effective Consumption of the data and services

AMRT

Agri Monitored Re-engineering and Transformation Methodology, Activities approved by RKVY/ICAR

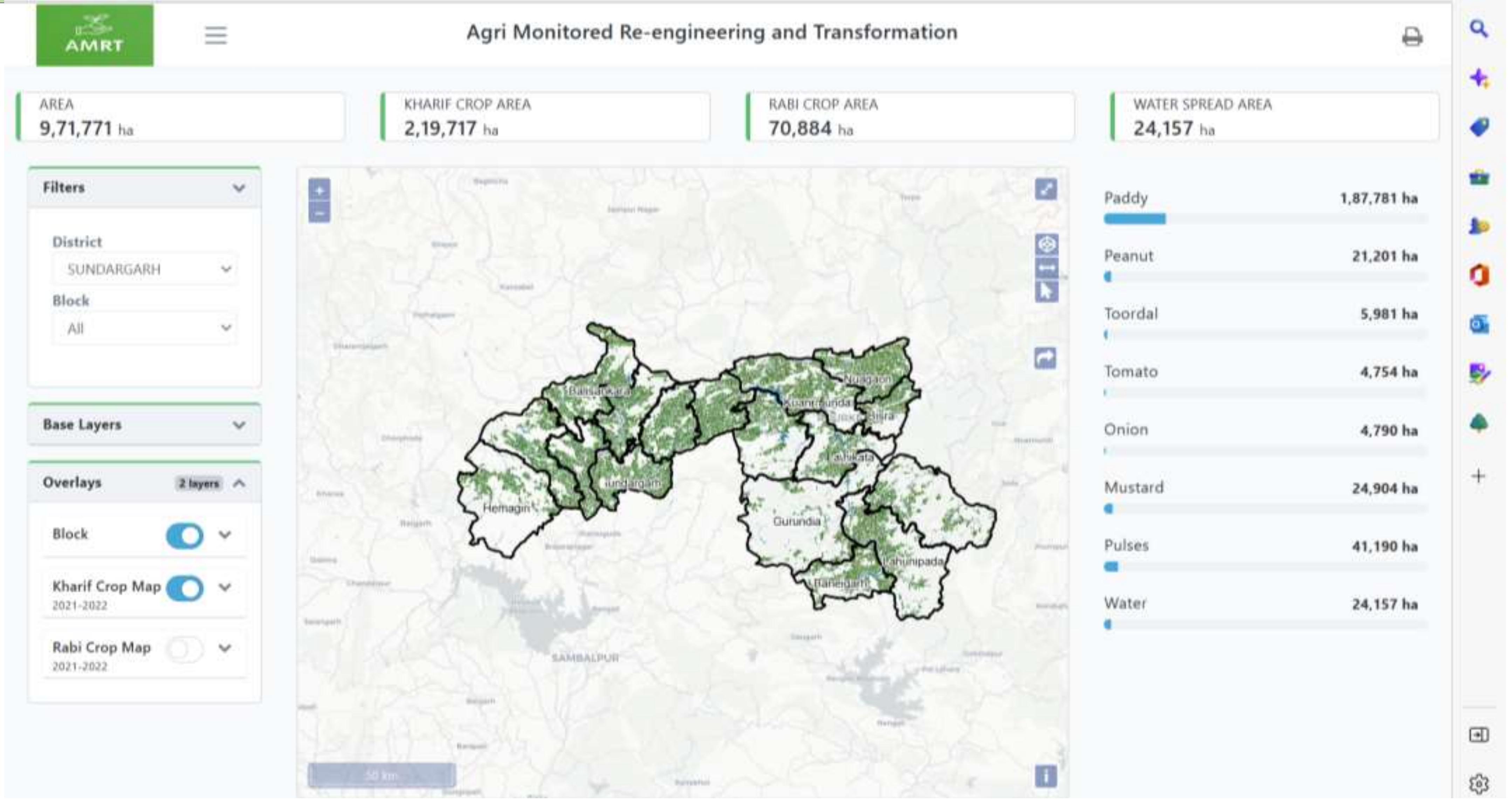
By

A Consortium of

AMRT Dashboard Implemented

Single view to monitoring crops landholding-wise

Historical analysis



INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

धन्यावाद

Thank You!

Contact Info:

K S Rajan (rajan@iiit.ac.in)

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD