Machine Learning for Sustainable Rice Production: Region-Scale Monitoring of Water-Saving Practices in Punjab, India

Ando Shah[1*†], Rajveer Singh[2], Akram Zaytar[3], Girmaw Abebe Tadesse[3], Caleb Robinson[3], Negar Tafti[2], Stephen A. Wood[2,4], Rahul Dodhia[3], Juan M. Lavista Ferres[3]

[1]University of California, Berkeley
[2]The Nature Conservancy
[3]Microsoft AI for Good Research Lab
[4]Yale School of the Environment
ando@berkeley.edu

Abstract

Rice cultivation supplies half the world's population with staple food, while also being a major driver of freshwater depletion–consuming roughly a quarter of global freshwater–and accounting for ∼48% of greenhouse gas emissions from croplands. In regions like Punjab, India, where groundwater levels are plummeting at 41.6 cm/year, adopting water-saving rice farming practices is critical. Direct-Seeded Rice (DSR) and Alternate Wetting and Drying (AWD) can cut irrigation water use by 20–40% without hurting yields, yet lack of spatial data on adoption impedes effective adaptation policy and climate action. We present a machine learning framework to bridge this data gap by monitoring sustainable rice farming at scale. In collaboration with agronomy experts and a large-scale farmer training program, we obtained ground-truth data from ∼1,400 fields across Punjab. Leveraging this partnership, we developed a novel dimensional classification approach that decouples sowing and irrigation practices, achieving F1 scores of 0.8 and 0.74 respectively, solely employing Sentinel-1 satellite imagery. Explainability analysis reveals that DSR classification is robust while AWD classification depends primarily on planting schedule differences, as Sentinel-1's 12-day revisit frequency cannot capture the higher frequency irrigation cycles characteristic of AWD practices. Applying this model across 3 million fields reveals spatial heterogeneity in adoption at the state level, highlighting gaps and opportunities for policy targeting. Our district-level adoption rates correlate well with government estimates (Spearman's ρ=0.69 and Rank Biased Overlap=0.77). This study provides policymakers and sustainability programs a powerful tool to track practice adoption, inform targeted interventions, and drive data-driven policies for water conservation and climate mitigation at regional scale.

Code & Dataset — https://github.com/microsoft/rice-irrigation-mapping-s1s2

Extended version — https://arxiv.org/abs/2507.08605

*Work done while at Microsoft AI for Good Lab, Kenya.
†Corresponding author

1 Introduction

Rice is a staple food for half of the world's population, but traditional paddy cultivation requires enormous quantities of water—2,000–5,000 liters per kilogram of rice produced (Bouman 2007), consuming 24-30% of world's freshwater (Surendran et al. 2021). As climate change intensifies water scarcity in many regions, sustainable water management in rice cultivation becomes essential for ensuring long-term food security, critical to addressing the United Nations Zero Hunger Sustainable Development Goal (SDG). Furthermore, rice cultivation accounts for ∼48% of greenhouse gas (GHG) emissions from croplands as a direct consequence of water-intensive paddy cultivation (Qian et al. 2023). Two water-saving practices have emerged as particularly promising: Direct Seeded Rice (DSR) and Alternate Wetting and Drying (AWD). DSR eliminates the need to raise and transplant seedlings by directly sowing seeds into the field, reducing water consumption by 20-40% (Bhushan et al. 2007; Bouman 2007), relative to traditionally practiced puddled transplanted rice (PTR). AWD involves periodically allowing fields to dry before re-flooding, reducing water use by up to 30% compared to traditional continuous flooding (CF) (Richards and Sander 2014). When properly managed, these techniques maintain or even enhance yields, presenting a balanced approach for sustainable rice production (Lansing et al. 2023).

Despite the promise of these practices, their scale of adoption remains poorly understood. This knowledge gap severely hampers evidence-based policymaking, resource allocation, and intervention planning for several reasons:

- Ground surveys are labor-intensive, expensive, and often constrained by limited access to remote fields
- Manual methods struggle to differentiate between traditional and newly adopted practices
- Adoption patterns are not routinely tracked in national or sub-national surveys
- Understanding current adoption helps target resources where they're most needed

Remote sensing offers a potential solution for monitoring water management practices at scale. While satellite-

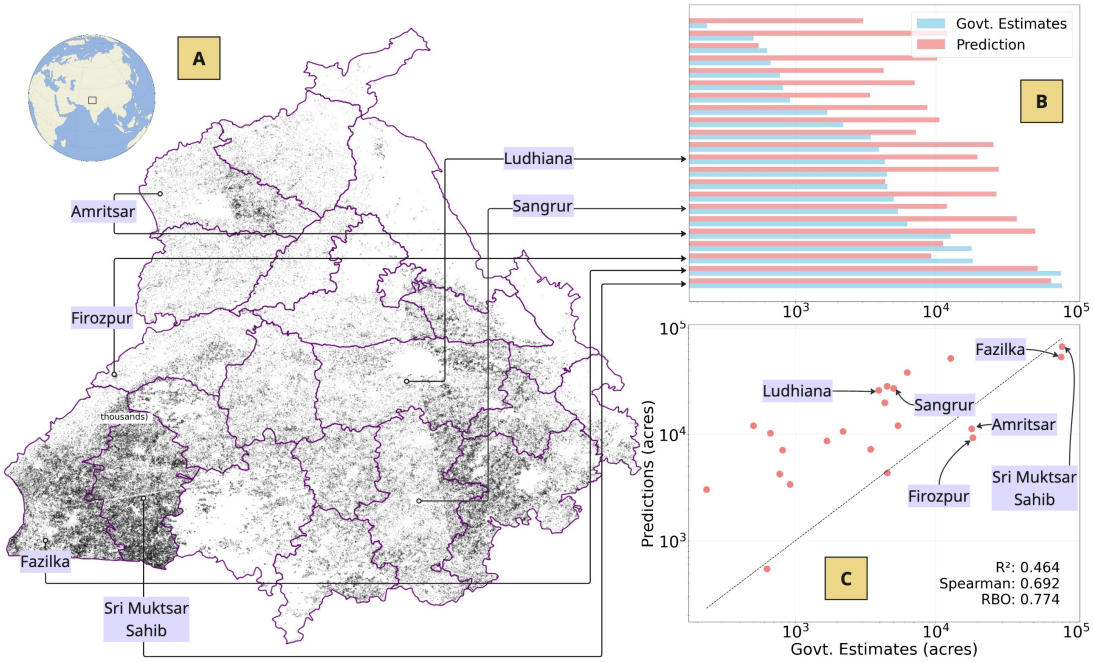

Figure 1: DSR predictions for Punjab in 2024. A) Map of Punjab with predicted DSR adoption distributions, shown with district boundaries; darker shades denote higher DSR plot-density. B) Comparison between model predictions and Punjab government estimates of DSR adoption by district, ordered by acres of DSR activity. X-axis is area under DSR in acres. C) District-level adoption compared to government records; evaluated using Spearman's ρ=0.69, R^2=0.46, Rank Biased Overlap=0.77 and MAE$\approx$ 12 thousand acres. Only key districts are labeled for visual clarity. Please refer to the Appendix for detailed figures.

based methods have successfully mapped rice cultivation using multispectral optical and synthetic aperture radar (SAR) approaches (Nguyen, Gruber, and Wagner 2016; Fatchurrachman et al. 2022; Singha et al. 2019a), detecting specific water management practices remains challenging. Existing studies rely heavily on prior knowledge of planting dates (Fikriyah et al. 2019; Villano et al. 2019) or use coarse resolution data (Gumma et al. 2015), limiting their scalability in regions where planting dates vary significantly, or smallholder-dominated regions.

We focused our analysis on Punjab, India, a major rice-producing region facing severe water stress. Punjab contributes 10.42% of India's rice production (Department of Agriculture and Farmers Welfare, Govt. of India 2024), but faces groundwater depletion of 41.6 cm/year (Baweja et al. 2017). This crisis stems from the region's semi-arid climate, rice-wheat cropping systems that increased irrigation demand, and groundwater extraction facilitated by free electricity (Gupta 2023; Punjab State Farmers and Farm Workers Commission 2023). Approximately 73% of irrigation water comes from rapidly depleting groundwater sources (Sidhu, Sharda, and Singh 2021). To address these challenges, we leveraged data from The Nature Conservancy's Promoting Regenerative and No-burn Agriculture (PRANA) project[1]. PRANA trained approximately 150,000 farmers across Punjab on sustainable water management methods, while collecting detailed field-level data from about 1,400 participants. This included documentation of sowing methods, irrigation practices, and precise field boundaries–critical ground truth data for our remote sensing approach. Our study introduces a dimensional framework for monitoring water management practices using Sentinel-1 (S1) satellite imagery. By separating practices along sowing (DSR vs. PTR) and irrigation (AWD vs. CF) dimensions, we leverage their distinct temporal signatures while avoiding the limitations of simultaneous detection. Conceptually, this framing clarifies that agronomic processes driving sowing and irrigation decisions are distinct and temporally separable. Importantly, our approach does not require prior knowledge of cropping calendars, enabling large-scale monitoring, with results shown in Figure 1. Our main contributions include:

- *Methodologically*, we develop a comparative evaluation framework that operationalizes this decomposition at state scale, integrating features extracted from S1 using traditional methods and pretrained Earth Observation (EO) models.
- *Empirically*, we show that this formulation achieves higher accuracy than a three-class model (0.798 v/s 0.616). We conduct explainability analysis of model outputs to provide evidence that high irrigation classification performance (0.742 F1) stems from planting schedule artifacts rather than irrigation patterns of frequencies higher than can be currently captured by a single S1 satellite.

[1]https://www.nature.org/en-us/about-us/where-we-work/india/our-priorities/prana/

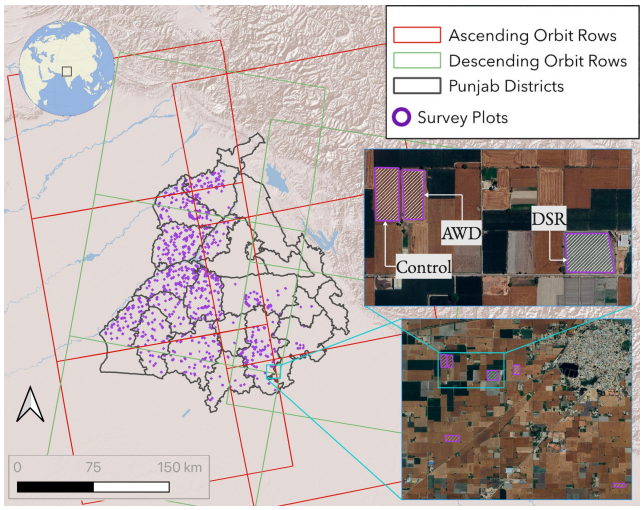

Figure 2: Map of Punjab showing the study area, Sentinel-1 satellite coverage (ascending and descending orbits), and locations of surveyed plots. Insets show sample plots at different zoom levels, with the innermost inset displaying delineated fields from three classes: Control (plots that use puddled transplanting and continuous flooding), alternate wetting and drying (AWD), and direct seeded rice (DSR). Background satellite imagery from Bing Maps.

- *Practically* we demonstrate scalable and near-operational monitoring of >3 million agricultural plots in Punjab, with aggregated district-level predictions correlating strongly with government records.

2 Background

Study Area

Our study focuses on Punjab, India (Fig. 2), a major rice-producing region facing severe water stress. Punjab contributes 10.42% of India's rice production and approximately 3.3% of global output (Department of Agriculture and Farmers Welfare, Govt. of India 2024), but faces rapid groundwater depletion. This crisis stems from wheat-rice double-cropping system using PTR and CF methods (Han et al. 2022) and groundwater extraction facilitated by free electricity, with 73% of irrigation extracted from rapidly depleting underground aquifers. A 2022 Government report classified 114 of Punjab's 150 blocks as over-exploited (Central Ground Water Board 2022). To address these challenges, we partnered with The Nature Conservancy's PRANA project, which trained approximately 150,000 farmers on sustainable water management while collecting detailed field-level data from more than 1,400 participants.

Challenges

While satellite-based monitoring offers a potential solution, several technical challenges have limited its application:

Water Management Practices and Separability: Water-saving practices in rice cultivation occur at different crop stages and create distinct water level patterns. DSR affects water management during early sowing and tillering phases, while AWD impacts water dynamics from tillering through harvest. These practices create unique temporal signatures in water levels and soil moisture that can be detected via SAR remote sensing (Lovell 2019; Fikriyah et al. 2019; Villano et al. 2019; Gumma et al. 2015). As shown in Figure 3, PTR maintains consistently high water levels, while DSR shows a sharp initial decline after pre-sowing irrigation. Similarly, CF maintains flooded fields throughout the season, whereas AWD creates cyclical wet-dry patterns.

Dimensional Classification Approach: Attempting to classify all water management combinations simultaneously poses substantial challenges due to overlapping characteristics. We address this by separating practices into two independent dimensions: **sowing dimension**, distinguishing between DSR and PTR practices and **irrigation dimension**, separating AWD from CF practices.

Planting Date Heterogeneity: Satellite-based monitoring faces significant variation in planting schedules. Previous studies relied on synchronizing satellite data with known planting dates (Fikriyah et al. 2019; Villano et al. 2019) – a critical challenge for monitoring millions of plots. In our study area, planting dates span up to 110 days, with DSR plots typically sown first, followed by control plots and AWD plots (Fig. S3).

Plot-level Predictions: To provide fine-grained insights, we must be able to both identify individual plot boundaries as well as make predictions about these monolithic entities. We employ Fields of The World (FTW) algorithm (Kerner et al. 2025) to automate the extraction of plot boundaries, enabling us to make predictions at an individual plot level.

3 Materials and Methods

Let X = $[x_t]_{t=1}^{T}$ be the seasonal S1 feature sequence for a plot. We want to learn two predictive functions: sowing $f_s : X \rightarrow \hat{y}_s \in \{\text{DSR,PTR}\}$ and irrigation $f_i : X \rightarrow \hat{y}_i \in \{\text{AWD,CF}\}$ as parallel binary tasks, additionally fused into a joint prediction, $\hat{y}_{joint} = g(\hat{y}_s, \hat{y}_i)$. We aim to learn these predictors under limited ground truth and heterogeneous planting dates.

Ground Truth Data Collection

We partnered with the PRANA project to obtain reliable training data. Field teams collected detailed plot-level data from intervention sites across 18 districts in Punjab during the 2024 *kharif* rice season. The resulting dataset includes 452 AWD plots, 420 DSR plots and 411 PTR plots employing CF, with mean areas of 4.2, 2.7 and 2.1 acres, respectively (see Figure 2 for samples). Sowing and transplantation dates, irrigation schedules, and precise field boundaries were documented through on-site visits. Field polygons were digitized from 2024 Google Earth Pro imagery (30–50 cm) enforcing a minimum area of 2000 m^2, and validated in-person for a subset of plots. A negative 10m buffer was applied

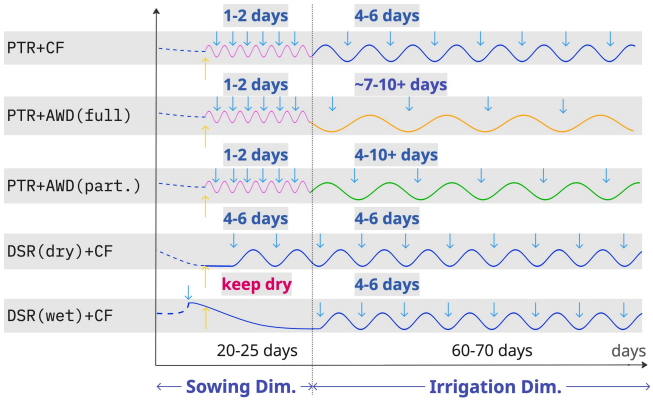

Figure 3: Water management practices across classes, created in collaboration with rice agronomy experts. The y-axis represents water level in a rice paddy, and x-axis shows growth stages. The growing season naturally separates into two phases: initial (left of dotted line) and main (right), which have distinct irrigation needs. Blue arrows indicate irrigation events, while yellow arrows denote sowing. The different practices create distinct temporal signatures, with a clear boundary between sowing and irrigation dimensions.

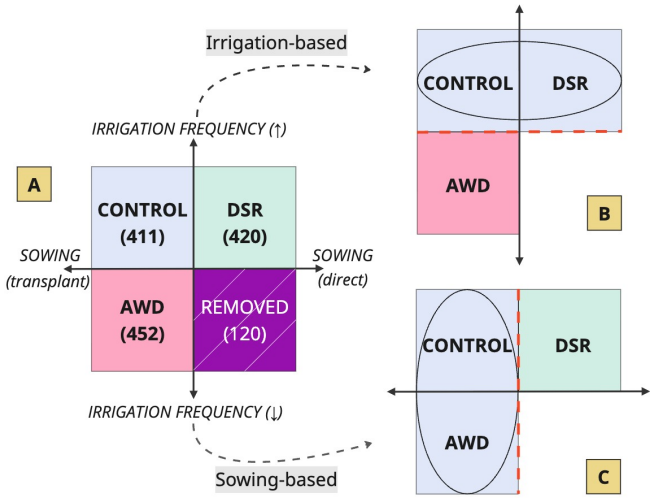

Figure 4: Dimensional divisions of dataset by practice type. A) Plots divided by sowing and irrigation dimensions, with number of plots in parentheses. B) Classification of irrigation practices (grouping Control and DSR as CF). C) Classification of sowing practices (grouping Control and AWD as PTR). Dotted red lines in B) and C) show classes are separated in a task-dependent manner.

to mitigate co-registration and boundary errors between the Google Earth basemap and co-registered S1 tiles.

Satellite Data and Processing

We leveraged S1 SAR data to monitor water management practices across the growing season (April - December 2024). We acquired 108 images from ascending orbit covering Punjab, focusing on May 1 - December 15[2]. Previous studies have shown SAR's effectiveness in capturing diverse rice agriculture metrics, including cultural type classification, phenological parameter detection, growing season estimation (Clauss et al. 2018; Yang et al. 2021; Sah et al. 2023), and differentiation between multiple cropping cycles and field-level tillage practices (Singha et al. 2019a; Liu et al. 2022a). SAR data was also selected for its cloud-penetrating capabilities during the monsoon season.

All imagery was processed using the SNAP toolbox (Filipponi 2019) to extract time series at 10-meter resolution. We extracted VV (Vertical transmit, Vertical receive) and VH (Vertical transmit, Horizontal receive) polarization bands, applying standard preprocessing: orbit file application, noise removal, radiometric calibration, multi-temporal speckle filtering (Yommy, Liu, and Wu 2015), and terrain correction to obtain terrain-geocoded backscatter coefficient, γ_0. For each plot, we calculated mean backscatter values per timestep, generating plot-level timeseries. To handle temporal gaps, we applied temporal smoothing with a 1-D Gaussian filter followed by third-degree polynomial spline fitting, enabling regular sampling despite variable acquisition dates, crucial for dealing with missing timeseries data. We additionally create a VV/VH band ratio time series, which has proven effective for rice mapping in previous studies (Villano et al. 2019; Bazzi et al. 2019; Singha et al. 2019b).

[2]Through experimentation, the ascending orbit dataset was shown to have higher performance and was selected for training and inference

Feature Engineering

We derive complementary feature families that summarize both explicit temporal structure and learned representations of radar backscatter dynamics throughout the rice-growing season. All features are computed at the plot level from smoothed S1 VV, VH, and VV/VH ratio time series.

Analytical Features: For each polarization (VV, VH, VV/VH), we extract temporal characteristics that include the position and amplitude of the first k troughs, crests, and inflection points; the total number of troughs and crests; summary statistics (mean, minimum, maximum amplitude); parameters from Gaussian kernel fitting to the VV/VH ratio (amplitude, peak day, standard deviation and R^2(Bazzi et al. 2019)); and radar vegetation index (RVI) values (Nasirzadehdizaji et al. 2019) for a total of 61 features, which we refer to as *hand-crafted* (HC) features. To complement these direct time-domain descriptors, we compute Fourier-domain features using fast Fourier transform (FT) that capture periodic irrigation-drainage patterns and overall signal smoothness.

Learnt Features: We include pretrained representations from two EO foundation models. The Pretrained Remote

Features							3-Class				2-Class			
Group	HC	DTW	FT	AE	P	#Feat.	One-shot		Fused		Sowing		Irrigation	
							Acc	F1	Acc	F1	Acc	F1	Acc	F1
Random	×	×	×	×	×		0.333	0.332	0.333	0.334	0.559	0.559	0.542	0.540
Baselines	✓					61	0.574	0.567	0.527*	0.534*	0.775*	0.769*	0.698*	0.687*
		✓				3	0.450	0.450	0.434	0.431	0.636	0.639	0.612	0.612
			✓			96	0.496	0.496	0.496	0.502	0.744	0.742	0.721*	0.707*
AE				✓		64	0.450	0.450	0.450	0.453	0.721	0.715	0.635	0.625
	✓			✓		125	**0.620**	**0.616**	0.574	0.578	**0.798**	**0.796**	0.713*	0.697*
			✓	✓		160	0.589	0.588	0.597	0.601	0.783	0.774	**0.744**	**0.742**
Presto					✓	6272	0.535*	0.526*	0.581	0.583	0.760	0.755	0.721	0.724
	✓				✓	6333	0.550	0.541	0.567	0.567	0.744	0.745	0.736	0.739
			✓		✓	6368	0.543*	0.534	0.589	0.589	0.752	0.750	0.721	0.725
AE + P				✓	✓	6336	0.581*	0.575*	0.581	0.583	0.752	0.750	0.736	0.736
			✓	✓	✓	6432	0.581*	0.576*	**0.605**	**0.603**	0.752	0.750	0.729	0.734

Table 1: Classification performance by dimension of separation. Dimensional separation consistently outperforms 3-class classification for all tasks. Features: HC = hand crafted, DTW = Dynamic Time Weighting + kNN, FT = Fourier transform, P = Presto, AE = AlphaEarth. Results are from LightGBM unless *=RF. **Bold** indicates best results per task; underlined, second-best. All tasks use the best date interval from ablations, while AE features are fixed for the calendar year 2024. 'One-shot' refers to a 3-class classification which will mechanically have lower scores, and hence we compare to 'fused' which is the combination of making a DSR prediction first, followed by AWD for a given plot. Acc = accuracy, F1 = F1-weighted. #Feat. indicates the number of features used for that task.

Sensing Transformer (Presto) (Tseng et al. 2023) encodes up to 24 Sentinel-1 timesteps into 128-dimensional temporal embeddings that capture non-linear phenological trajectories. The Google AlphaEarth (AE) embeddings product (Brown et al. 2025) provides lightweight 64-dimensional embeddings summarizing annual multisensor (S1, S2, Landsat) variability, capturing landscape level dynamics at 10-meter resolutions. Both these embeddings are inexpensive to acquire, and therefore very attractive to practitioners.

Classification Approach

We test two ensemble learning methods—Random Forest (RF) and Gradient Boosted Trees (GB)–for their proven effectiveness in agricultural classification tasks (Liu et al. 2022b; Fang et al. 2024). These models effectively handle the heterogeneous feature sets while remaining computationally efficient. We also include a Dynamic Time Warping (DTW) baseline to benchmark performance without explicit feature extraction. DTW directly compares raw S1 time series using dynamic alignment and classifies them via a 1-nearest-neighbor scheme implemented using the tslearn package (Zhang et al. 2017).

Following our dimensional framework, we structure three classification tasks: a) AWD vs. DSR vs. Control (null hypothesis: all practices on same dimension), b) DSR vs. PTR (sowing dimension), and c) AWD vs. CF (irrigation dimension). For each task, we split data into non-overlapping train/test sets (90:10 ratio) using stratified sampling. We optimize hyperparameters using Optuna (Akiba et al. 2019) and evaluate performance using overall accuracy and F1 scores (globally weighted averages). We test a set of combinations of features and classification methods, summarized in Table 1. This approach tests our hypothesis that separating water management practices along their natural dimensions (sowing and irrigation) produces more accurate results than attempting simultaneous multi-class classification (i.e. the null hypothesis, 3-class one-shot). For comparison, we also evaluate a fused three-class setup where the sowing and irrigation outputs are combined into a single categorical label via rule-based fusion (fused := DSR $\rightarrow$ DSR, else irrigation decides AWD vs CF) - illustrating compounded error from sequential binary decisions and serves as a fairer baseline.All results use stratified random splits within the same region-year (Punjab-2024). Please see Figure A3 for the complete training workflow.

4 Results

Dimensional Classification Outperforms Combined Approach

The dimensional approach–classifying sowing (DSR vs. PTR) and irrigation (AWD vs. CF) separately–significantly outperforms the combined three-class formulation, as shown in Table 1. DSR and AWD achieve F1 scores of 0.796 and 0.742, respectively, supporting our hypothesis that practices are more effectively identified along their distinct dimensions, relative to the 3-class formulations. For those same situations, fusing their outcomes reduces the joint accuracy to ≈ 0.6, confirming that dimensional decomposition avoids error compounding while better reflecting the underlying agronomic processes.

Temporal Ranges										Lag (%)	Sowing		Irrigation		F1 (weighted)		
Row	Dates	M	J	J	A	S	O	N	D		DSR	PTR	CF	AWD	ALL	DSR	AWD
1	May 1 - Aug 15*									∼50%	✓	✓	>	×	0.519	0.740	0.718
2	Jun 1 - Aug 30									∼50%	>	✓	✓	<	0.539	0.720	0.723
3	Jun 1 - Sep 5*									∼50%	>	✓	✓	<	0.539	**0.755**	0.723
4	Jun1 - Oct 15									∼60%	>	✓	✓	✓	**0.543**	**0.751**	0.731
5	July 1 - Oct 15									∼50%	×	✓	✓	✓	0.531	0.728	0.719
6	Aug 1 - Oct 15									∼33%	×	×	✓	✓	0.489	0.696	0.709
7	Aug 1 - Nov 15									∼50%	×	×	✓	✓	0.434	0.728	0.707
8	Sep 1 - Dec 15									∼50%	×	×	✓	✓	0.499	0.677	0.662
9	Oct 1 - Dec 15									∼33%	×	×	<	✓	0.519	0.658	0.643
10	Aug 1 - Dec 15									∼60%	×	×	✓	✓	0.470	0.701	0.683
11	Jun 1 - Dec 15									∼33%	>	✓	✓	✓	**0.541**	0.713	**0.737**
12	May 1 - Dec 15**									∼90%	✓	✓	✓	✓	0.537	0.713	**0.739**

Table 2: Effect of temporal range selection on accuracy. Features are generated for different temporal ranges to analyze the effect on outcomes by manipulating the availability of lag and practice features for Presto+ features. AE features are not used, since its temporal range is fixed to a calendar year. Cell colors for months represent how much of the month is covered for that row (blank=none, green=full, orange=half). Values in bold indicate the top-2 performers for each classification task. Nominal sampling frequency of 7 days; *= a sampling frequency of 4 days, **=10 days. ✓ and × indicate the availability of a feature type, while > & < imply that most & some features available, respectively. Sowing and irrigation feature availability is decided by virtue of which part of the growing cycle the temporal range reveals.

Foundation Model Embeddings Enable Robust Classification

Both Presto and AE yield the highest performance among single feature sources, achieving sowing F1 scores of 0.755 and 0.715, respectively, comparable to the best handcrafted features. Their full value emerges in combination with complementary feature families, particularly Fourier features, where AE+FT attains the best sowing performance (F1=0.798). A similar pattern holds for irrigation, underscoring that foundation model embeddings capture transferable temporal dynamics that enhance orthogonal feature spaces. These results highlight the utility of such embeddings for practitioners, as they are inexpensive to obtain and require no specialized training. Further details are provided in Appendix D.

Lag Features Drive Classification Performance

Lag features encode planting-date offsets between plots. The independent PRANA dataset (Figure A1) shows wide state-level variance but compact practice-specific planting windows, enabling these offsets to act as strong discriminators—though their utility depends on temporal stability across seasons. **Practice features** capture water-management dynamics intrinsic to each technique. Sowing features (DSR/PTR) are most salient in early growth, while AWD/CF differences emerge later as irrigation frequency diverges (typically every 4–10 days; Figure 3).

To isolate their effects, we trained models on restricted temporal windows (Table 2). Sowing models perform best when both lag and early-season features are available (rows 1, 3, 10–11), confirming that they rely on both signal types and remain robust to sampling frequency (4–7 days). In contrast, irrigation models depend almost entirely on lag information: performance peaks when all features are present (row 11) and collapses once lag cues are reduced (row 8) or sowing phases are absent (rows 5–9). This dependence indicates that AWD detection leverages schedule shifts rather than true irrigation cycles, which Sentinel-1A's 12-day revisit cannot resolve. Further analyses of feature importance, orbit effects, and errors are provided in Appendix F.

Inference at Scale

A key challenge in promoting sustainable agriculture is understanding adoption patterns of water-saving practices. Despite extensive farmer training, PRANA lacked state-wide visibility due to the high cost of surveys. Our satellite-based approach overcomes this limitation, enabling monitoring of Punjab's 3.2 million hectares of rice cultivation and providing critical insights for food and water policy. We delineate field boundaries using the FTW algorithm (Kerner et al. 2025) applied to bi-temporal Sentinel-2 RGB-NIR imagery (May & Oct 2024). FTW polygons are filtered to 0.2-10 Ha, matching Punjab's typical rice-plot size—means: 0.5±0.7 Ha (PRANA internal survey), 0.6±0.5 Ha ((Wang, Waldner, and Lobell 2022) Punjab subset), and 2.7 Ha (Singh et al. 2024). Validation of FTW polygons over the PRANA training area shows that 57% of sub-1 Ha plots were detected, with mIoU 0.41 ±0.2 over all plots. FTW often splits large fields (≈1.4 sub-plots per label) and misses 36% of plots, reflecting both its resolution limits and ability to capture smaller sub-plots. We provide further details and qualitative samples of this step in Appendix G and Figure A9.

Given sparse training labels (∼1,400 labeled plots vs. 3 million during inference), we employed an ensemble of top-performing RF and GB models, aggregating predictions via the modal class. Results are summarizes in Figure 1A, showing marked spatial heterogeneity in DSR adoption across and within districts. Our district-level predictions

correlate strongly with government records (Sidhu 2025) reporting 253,328 acres under DSR (Spearman's ρ=0.69, RBO=0.77 (Webber, Moffat, and Zobel 2010)), despite regional discrepancies. Since our training data included only rice plots, we applied a rice-cultivation mask (Han et al. 2022) from 2021 to exclude non-rice fields. We observe a residual diagonal gradient arises from minor temporal offsets between adjoining S-1 rows, mitigated via radiometric terrain-geocoded correction.

The model generally overestimates DSR area relative to government figures, except in the top three districts. These discrepancies may stem from both model error and noise in the self-reported government data[3], which, as seen in neighboring states (Deswal 2024), may include unverifiable claims. Notably, for Sri Muktsar Sahib and Fazilka, masking introduced large differences ($\sim$40k acres each), likely due to post-2021 expansion in rice cultivation. We therefore report unmasked estimates for these districts (see Table A1 and Appendix G). Further in-field validation will be important to confirm these large-scale adoption estimates. Despite such variations, the high RBO score indicates strong preservation of district rankings, making the system well-suited for policy prioritization by highlighting both lagging and leading districts for targeted action and learning best practices.

5 Limitations

Our approach, while effective, faces several constraints that merit consideration. First, the current S1 revisit frequency of 12 days limits our ability to capture high-frequency irrigation variability characteristic of AWD practices. This temporal resolution challenge could be addressed by integrating data from ascending and descending orbits or incorporating forthcoming SAR missions.

Second, field boundary extraction using the FTW algorithm is constrained by Sentinel-2's spatial resolution, which imposes a minimum detectable plot size. This limitation is particularly salient in smallholder farming contexts where field sizes may fall below this detection threshold, potentially excluding smaller agricultural parcels from analysis.

Third, explainability analysis revealed that our models have degrees of reliance on temporal lag features that capture distinctions between water management practices. Additionally, the small dataset size ($\sim$1,400 plots) relative to the large inference dataset underscores the need for careful validation. Although Punjab has relatively low diversity in its agroecological zones, other regions may present greater variation in cropping systems, climate, or practice clustering, affecting generalization. Deployment of such methods would require systematic post-training field validation to verify model predictions against ground truth before operational use, particularly given our reliance on temporal patterns that may shift across agricultural contexts. Therefore, cross-year and cross-region generalization remain future extensions to assess temporal and spatial transferability.

Finally, our analysis depends on external estimates of rice-growing areas. Temporal mismatches between these masks and the study period introduced complications, particularly in regions where rice cultivation has expanded. Future work should explore explicitly classifying non-rice plots to reduce reliance on exteral and outdated maps.

6 Conclusion & Future Work

Our study demonstrates the feasibility of satellite-based monitoring for rice water management, offering a scalable and data-driven framework for tracking sustainable agriculture adoption. By decoupling classification into sowing (DSR) and irrigation (AWD) dimensions, the approach outperforms combined formulations. Co-designed with agronomy experts, this dimensional structure proved essential: DSR detection notably surpassed joint classification and enabled statewide mapping of DSR adoption across 3 million fields in Punjab, India. Explainability analysis further showed that while AWD classification achieves an relatively high F1 scores, it depends primarily on differential planting schedules rather than irrigation signals, underscoring the importance of interpretability for reliable deployment. AWD performance remains limited by Sentinel-1A's 12-day revisit, though upcoming missions (Sentinel-1C/D, NISAR) will improve temporal resolution, potentially improving irrigation detection.

Operating without prior knowledge of planting dates, the framework generalizes to real-world systems with heterogeneous cropping calendars. Embeddings from remote sensing foundation models–Presto (Tseng et al. 2023) and AlphaEarth (Brown et al. 2025)–were key to this generalization, delivering strong accuracy from limited labeled data. A new class of efficient and highly performant EO foundation models, including OlmoEarth (Herzog et al. 2025), Panopticon (Waldmann et al. 2025) and Galileo (Tseng et al. 2025), have recently emerged, with precomputed embeddings increasingly available, further extending the potential for scalable, low-cost monitoring of agricultural practices. Finally, future work should examine why AlphaEarth's year-long embeddings retain strong sensitivity to rice sowing signals despite spanning multiple cropping seasons.

District-level predictions correlate strongly with government records, supporting utility for food and water policy through: (1) evidence-based resource allocation, (2) large-scale impact measurement, (3) identification of low-adoption areas, and (4) statewide monitoring without costly field surveys. At finer scales, plot-level irrigation mapping provides timely, high-resolution data on water management practices, enabling biogeochemical models to more accurately simulate methane emissions and quantify the mitigation benefits of sustainable agriculture. Extending this framework to other rice-growing regions and adapting it to local conditions–crop varieties, irrigation regimes, and management practices–will enhance generalization and policy relevance. As climate change accelerates water scarcity, scalable monitoring of sustainable irrigation will be vital for ensuring food security and advancing the Zero Hunger Sustainable Development Goal.

[3]Data sourced from a DSR incentive program offering Rs. 1500 per acre, pre-verification; c.f. see (Sidhu 2025) Table 1

References

Akiba, T.; Sano, S.; Yanase, T.; Ohta, T.; and Koyama, M. 2019. Optuna: A next-generation hyperparameter optimization framework. In *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining*, 2623–2631.

Baweja, S.; Aggarwal, R.; Brar, M.; and Lal, R. 2017. Groundwater depletion in Punjab, India. *Encyclopedia of soil science*, 2017: 1–5.

Bazzi, H.; Baghdadi, N.; El Hajj, M.; Zribi, M.; Minh, D. H. T.; Ndikumana, E.; Courault, D.; and Belhouchette, H. 2019. Mapping paddy rice using Sentinel-1 SAR time series in Camargue, France. *Remote sensing*, 11(7): 887.

Bhushan, L.; Ladha, J. K.; Gupta, R. K.; Singh, S.; Tirol-Padre, A.; Saharawat, Y.; Gathala, M.; and Pathak, H. 2007. Saving of water and labor in a rice–wheat system with no-tillage and direct seeding technologies. *Agronomy Journal*, 99(5): 1288–1296.

Bouman, B. 2007. *Water management in irrigated rice: coping with water scarcity*. Int. Rice Res. Inst. (IRRI).

Brown, C. F.; Kazmierski, M. R.; Pasquarella, V. J.; Rucklidge, W. J.; Samsikova, M.; Zhang, C.; Shelhamer, E.; Lahera, E.; Wiles, O.; Ilyushchenko, S.; et al. 2025. Alphaearth foundations: An embedding field model for accurate and efficient global mapping from sparse label data. *arXiv preprint arXiv:2507.22291*.

Central Ground Water Board. 2022. Central Ground Water Board Report 2022. https://www.cgwb.gov.in/old_website/GW-Assessment/GWR-2022-Reports%20State/Punjab.pdf, Accessed:2025-02-11.

Clauss, K.; Ottinger, M.; Leinenkugel, P.; and Kuenzer, C. 2018. Estimating rice production in the Mekong Delta, Vietnam, utilizing time series of Sentinel-1 SAR data. *International journal of applied earth observation and geoinformation*, 73: 574–585.

Department of Agriculture and Farmers Welfare, Govt. of India. 2024. Market Intelligence Report - Rice. https://upag.gov.in, Accessed:2025-02-11.

Deswal, D. 2024. 69% of Claims Filed by Farmers for DSR Incentive Found Fake in Haryana. *The Tribune, India*. www.tribuneindia.com/news/haryana/69-of-claims-of-haryana-farmers-for-dsr-incentive-found-fake/, Accessed:2025-11-11.

Fang, H.; Liang, S.; Chen, Y.; Ma, H.; Li, W.; He, T.; Tian, F.; and Zhang, F. 2024. A comprehensive review of rice mapping from satellite data: Algorithms, product characteristics and consistency assessment. *Science of Remote Sensing*, 100172.

Fatchurrachman; Rudiyanto; Soh, N. C.; Shah, R. M.; Giap, S. G. E.; Setiawan, B. I.; and Minasny, B. 2022. High-resolution mapping of paddy rice extent and growth stages across peninsular malaysia using a fusion of sentinel-1 and 2 time series data in google earth engine. *Remote Sensing*, 14(8): 1875.

Fikriyah, V. N.; Darvishzadeh, R.; Laborte, A.; Khan, N. I.; and Nelson, A. 2019. Discriminating transplanted and direct seeded rice using Sentinel-1 intensity data. *International Journal of Applied Earth Observation and Geoinformation*, 76: 143–153.

Filipponi, F. 2019. Sentinel-1 GRD preprocessing workflow. In *Proceedings*, volume 18, 11. MDPI.

Gumma, M. K.; Uppala, D.; Mohammed, I. A.; Whitbread, A. M.; and Mohammed, I. R. 2015. Mapping direct seeded rice in Raichur district of Karnataka, India. *Photogrammetric Engineering & Remote Sensing*, 81(11): 873–880.

Gupta, D. 2023. Free power, irrigation, and groundwater depletion: Impact of farm electricity policy of Punjab, India. *Agricultural Economics*, 54(4): 515–541.

Han, J.; Zhang, Z.; Luo, Y.; Cao, J.; Zhang, L.; Zhuang, H.; Cheng, F.; Zhang, J.; and Tao, F. 2022. Annual paddy rice planting area and cropping intensity datasets and their dynamics in the Asian monsoon region from 2000 to 2020. *Agricultural Systems*, 200: 103437.

Herzog, H.; Bastani, F.; Zhang, Y.; Tseng, G.; Redmon, J.; Sablon, H.; Park, R.; Morrison, J.; Buraczynski, A.; Farley, K.; et al. 2025. OlmoEarth: Stable Latent Image Modeling for Multimodal Earth Observation. *arXiv preprint arXiv:2511.13655*.

Kerner, H.; Chaudhari, S.; Ghosh, A.; Robinson, C.; Ahmad, A.; Choi, E.; Jacobs, N.; Holmes, C.; Mohr, M.; Dodhia, R.; et al. 2025. Fields of the world: A machine learning benchmark dataset for global agricultural field boundary segmentation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, 28151–28159.

Lansing, J.; Kremer, J.; Suryawan, I.; Sathiakumar, S.; Jacobs, G.; Chung, N.; and Artha Wiguna, I. W. A. 2023. Adaptive irrigation management by Balinese farmers reduces greenhouse gas emissions and increases rice yields. *Philosophical Transactions of the Royal Society B*, 378(1889): 20220400.

Liu, Y.; Rao, P.; Zhou, W.; Singh, B.; Srivastava, A. K.; Poonia, S. P.; Van Berkel, D.; and Jain, M. 2022a. Using Sentinel-1, Sentinel-2, and Planet satellite data to map field-level tillage practices in smallholder systems. *Plos one*, 17(11): e0277425.

Liu, Y.; Rao, P.; Zhou, W.; Singh, B.; Srivastava, A. K.; Poonia, S. P.; Van Berkel, D.; and Jain, M. 2022b. Using Sentinel-1, Sentinel-2, and Planet satellite data to map field-level tillage practices in smallholder systems. *PLoS One*, 17(11): e0277425.

Lovell, R. J. 2019. Identifying alternative wetting and drying (AWD) adoption in the Vietnamese Mekong River Delta: A change detection approach. *ISPRS International Journal of Geo-Information*, 8(7): 312.

Nasirzadehdizaji, R.; Balik Sanli, F.; Abdikan, S.; Cakir, Z.; Sekertekin, A.; and Ustuner, M. 2019. Sensitivity analysis of multi-temporal Sentinel-1 SAR parameters to crop height and canopy coverage. *Applied Sciences*, 9(4): 655.

Nguyen, D. B.; Gruber, A.; and Wagner, W. 2016. Mapping rice extent and cropping scheme in the Mekong Delta using Sentinel-1A data. *Remote Sensing Letters*, 7(12): 1209–1218.

Punjab State Farmers and Farm Workers Commission. 2023. Punjab State Agricultural Policy. {https://psfc.org.in/wp-content/uploads/2024/10/Punjab-State-Agricultural-Policy-English.pdf}, Accessed:2025-02-11.

Qian, H.; Zhu, X.; Huang, S.; Linquist, B.; Kuzyakov, Y.; Wassmann, R.; Minamikawa, K.; Martinez-Eixarch, M.; Yan, X.; Zhou, F.; et al. 2023. Greenhouse gas emissions and mitigation in rice agriculture. *Nature Reviews Earth & Environment*, 4(10): 716–732.

Richards, M. B.; and Sander, B. O. 2014. Alternate wetting and drying in irrigated rice. *CSA Practice Brief*.

Sah, S.; Haldar, D.; Chandra, S.; and Nain, A. S. 2023. Discrimination and monitoring of rice cultural types using dense time series of Sentinel-1 SAR data. *Ecological Informatics*, 76: 102136.

Sidhu, A. S. 2025. Progressive Farming. *Progressive Farming*, 61(6). A Monthly Journal of Punjab Agricultural University.

Sidhu, B. S.; Sharda, R.; and Singh, S. 2021. Spatio-temporal assessment of groundwater depletion in Punjab, India. *Groundwater for Sustainable Development*, 12: 100498.

Singh, A.; Kaur, M.; Arora, K.; and Kingra, H. 2024. Economics of Basmati Rice Cultivation in Punjab. *Journal of Agricultural Development and Policy*, 33(2): 269–273.

Singha, M.; Dong, J.; Zhang, G.; and Xiao, X. 2019a. High resolution paddy rice maps in cloud-prone Bangladesh and Northeast India using Sentinel-1 data. *Scientific data*, 6(1): 26.

Singha, M.; Dong, J.; Zhang, G.; and Xiao, X. 2019b. High resolution paddy rice maps in cloud-prone Bangladesh and Northeast India using Sentinel-1 data. *Scientific data*, 6(1): 26.

Surendran, U.; Raja, P.; Jayakumar, M.; and Subramoniam, S. R. 2021. Use of efficient water saving techniques for production of rice in India under climate change scenario: A critical review. *Journal of Cleaner Production*, 309: 127272.

Tseng, G.; Cartuyvels, R.; Zvonkov, I.; Purohit, M.; Rolnick, D.; and Kerner, H. 2023. Lightweight, pre-trained transformers for remote sensing timeseries. *arXiv preprint arXiv:2304.14065*.

Tseng, G.; Fuller, A.; Reil, M.; Herzog, H.; Beukema, P.; Bastani, F.; Green, J. R.; Shelhamer, E.; Kerner, H.; and Rolnick, D. 2025. Galileo: Learning Global & Local Features of Many Remote Sensing Modalities. *arXiv preprint arXiv:2502.09356*.

Villano, L.; Raviz, J.; Paguirigan, N.; Gutierrez, M.; Mabalay, M.; and Laborte, A. 2019. Separability of Transplanted and Direct Seeded Rice Using Multi-Temporal SENTINEL-1A Data. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 42: 471–478.

Waldmann, L.; Shah, A.; Wang, Y.; Lehmann, N.; Stewart, A.; Xiong, Z.; Zhu, X. X.; Bauer, S.; and Chuang, J. 2025. Panopticon: Advancing any-sensor foundation models for earth observation. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, 2204–2214.

Wang, S.; Waldner, F.; and Lobell, D. B. 2022. Unlocking large-scale crop field delineation in smallholder farming systems with transfer learning and weak supervision. *Remote Sensing*, 14(22): 5738.

Webber, W.; Moffat, A.; and Zobel, J. 2010. A similarity measure for indefinite rankings. *ACM Transactions on Information Systems (TOIS)*, 28(4): 1–38.

Yang, H.; Pan, B.; Li, N.; Wang, W.; Zhang, J.; and Zhang, X. 2021. A systematic method for spatio-temporal phenology estimation of paddy rice using time series Sentinel-1 images. *Remote Sensing of Environment*, 259: 112394.

Yommy, A. S.; Liu, R.; and Wu, S. 2015. SAR image despeckling using refined Lee filter. In *2015 7th International Conference on Intelligent Human-Machine Systems and Cybernetics*, volume 2, 260–265. IEEE.

Zhang, Z.; Tavenard, R.; Bailly, A.; Tang, X.; Tang, P.; and Corpetti, T. 2017. Dynamic time warping under limited warping path length. *Information Sciences*, 393: 91–107.