



UAV-Based Multispectral Time-Series & Weather-Based Hurdle Modeling For Onion Diseases & Pest Estimation

Ajay Pathak¹, Dimple Bhuta², Sagar Wayal³, Rohit Kushwaha², Sanket Bhagde², Ashwini Gajarushi², Suresh Gawande³

¹ Center for Machine Intelligence & Data Science, IIT Bombay, Mumbai, India

² TIH Foundation for IoT & IoE, IIT Bombay, Mumbai, India

³ ICAR–Directorate of Onion & Garlic Research (DOGR), Rajgurunagar, Pune, India



PROBLEM

India produces 24.21 Mt of onions, yet biotic and weather stresses can cause 30–50% yield loss. Early stress detection is critical for effective crop management.

BIOTIC STRESS

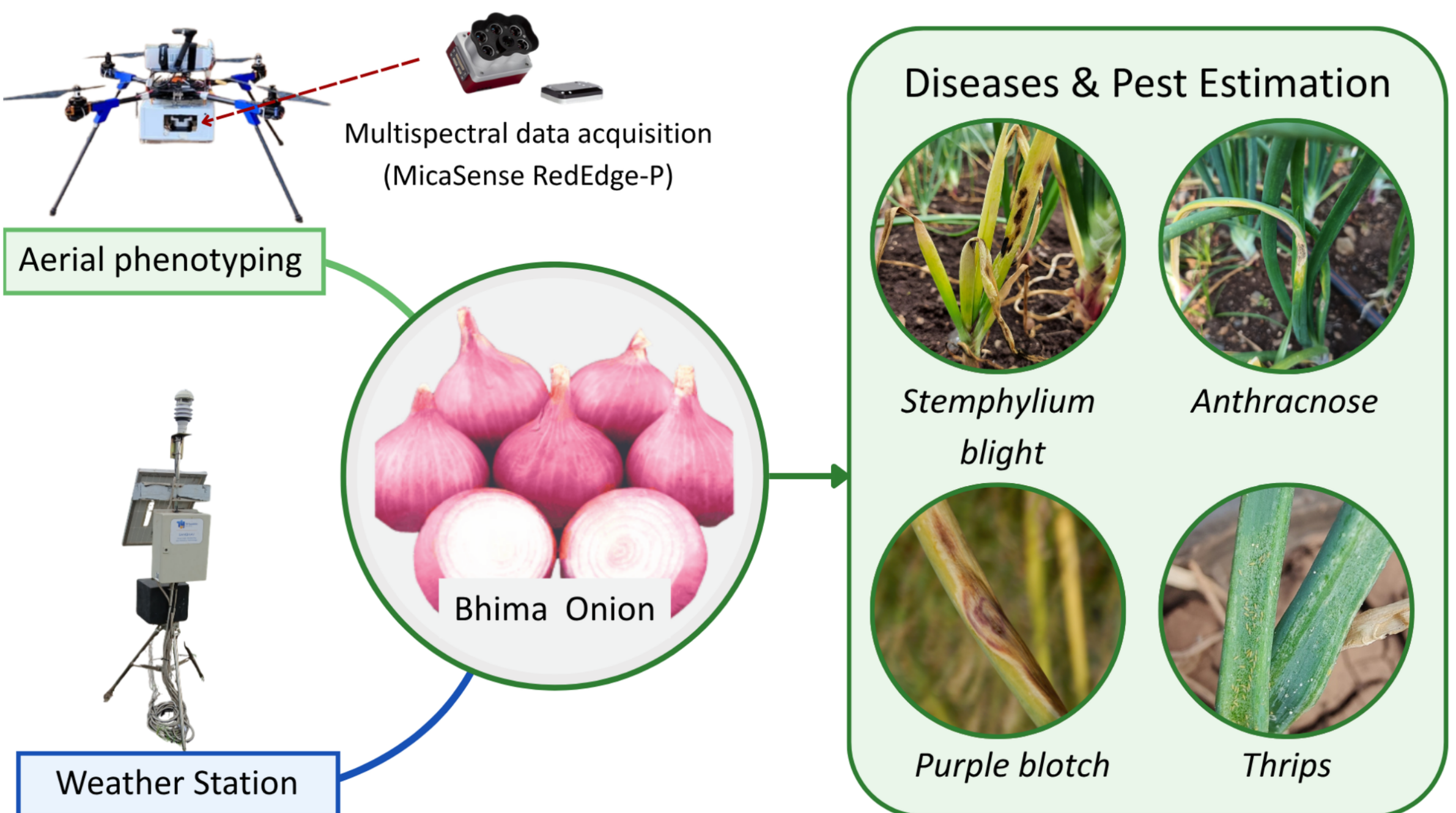
- Stemphylium Blight
- Anthracnose
- Purple Blotch
- Thrips

ABIOTIC STRESS

- Rainfall
- Relative Humidity
- Temperature

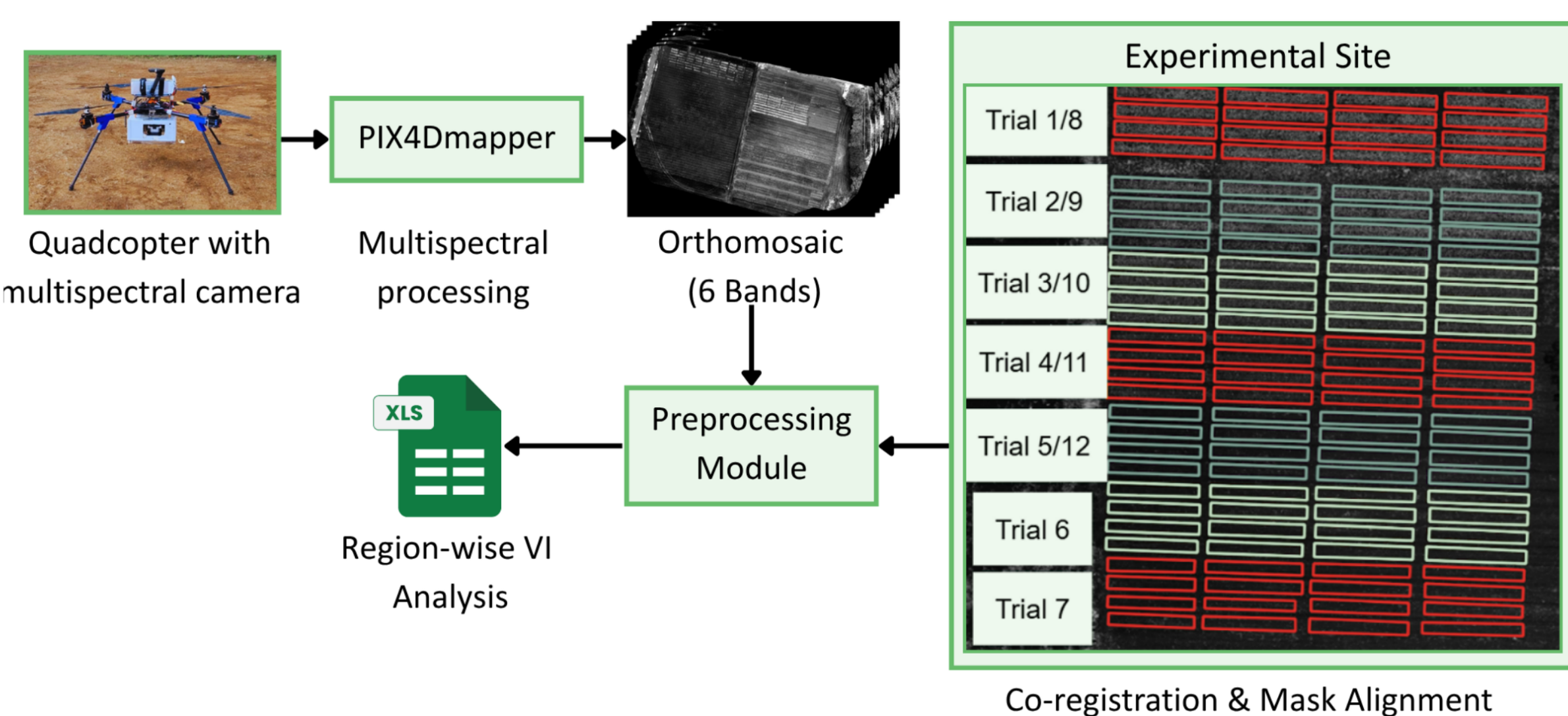
PROPOSED SOLUTION

We propose an integrated predictive framework combining high-resolution **UAV multispectral imagery** and **meteorological data** to estimate **disease and pest incidence** in ‘Bhima’ onions under field conditions.



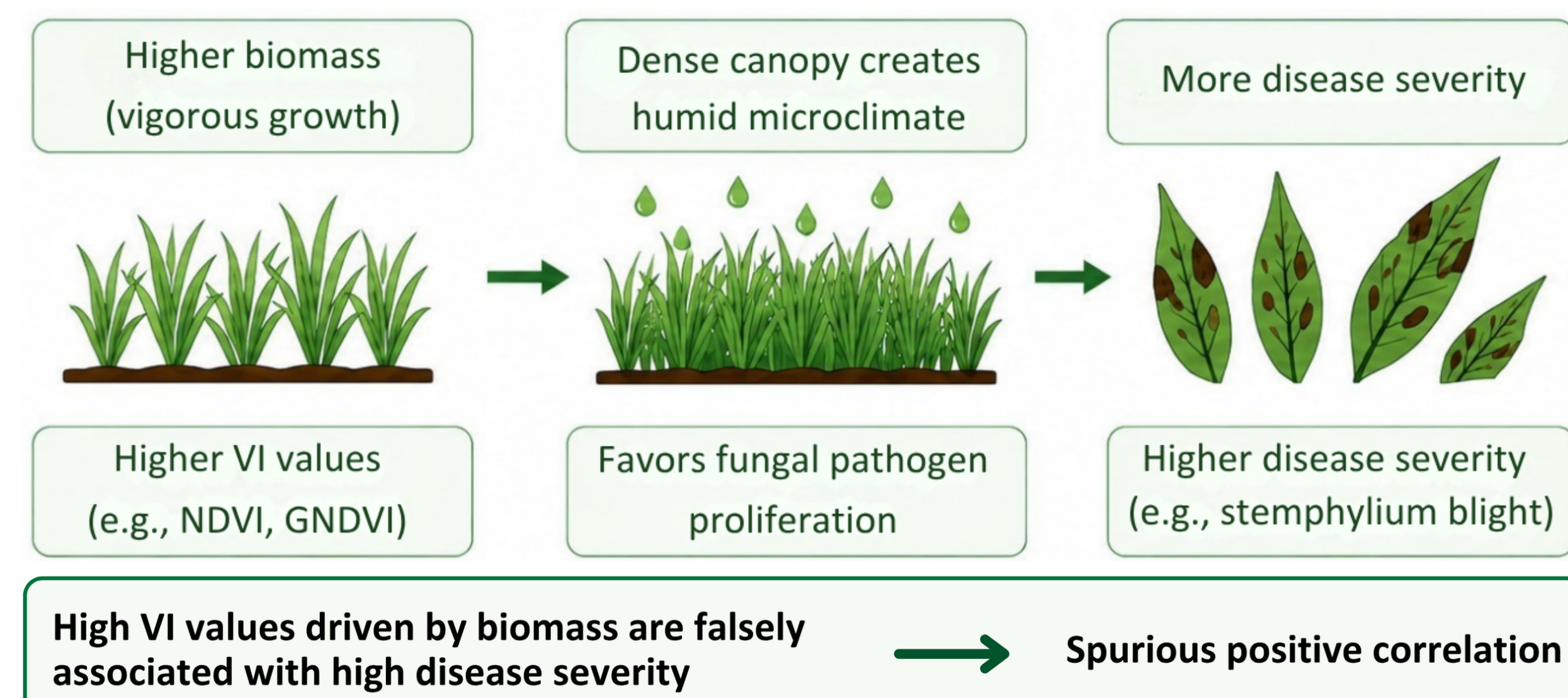
MULTISPECTRAL DATA ACQUISITION & ANALYSIS

Six-band aerial images (2.0 cm/pixel) were acquired using a MicaSense RedEdge-P sensor at 30 m altitude. Calibrated orthomosaics were generated in Pix4Dmapper and co-registered with vegetation masks to isolate canopy pixels. Region-wise Vegetation Index (VI) time-series were then extracted for disease and pest assessment.

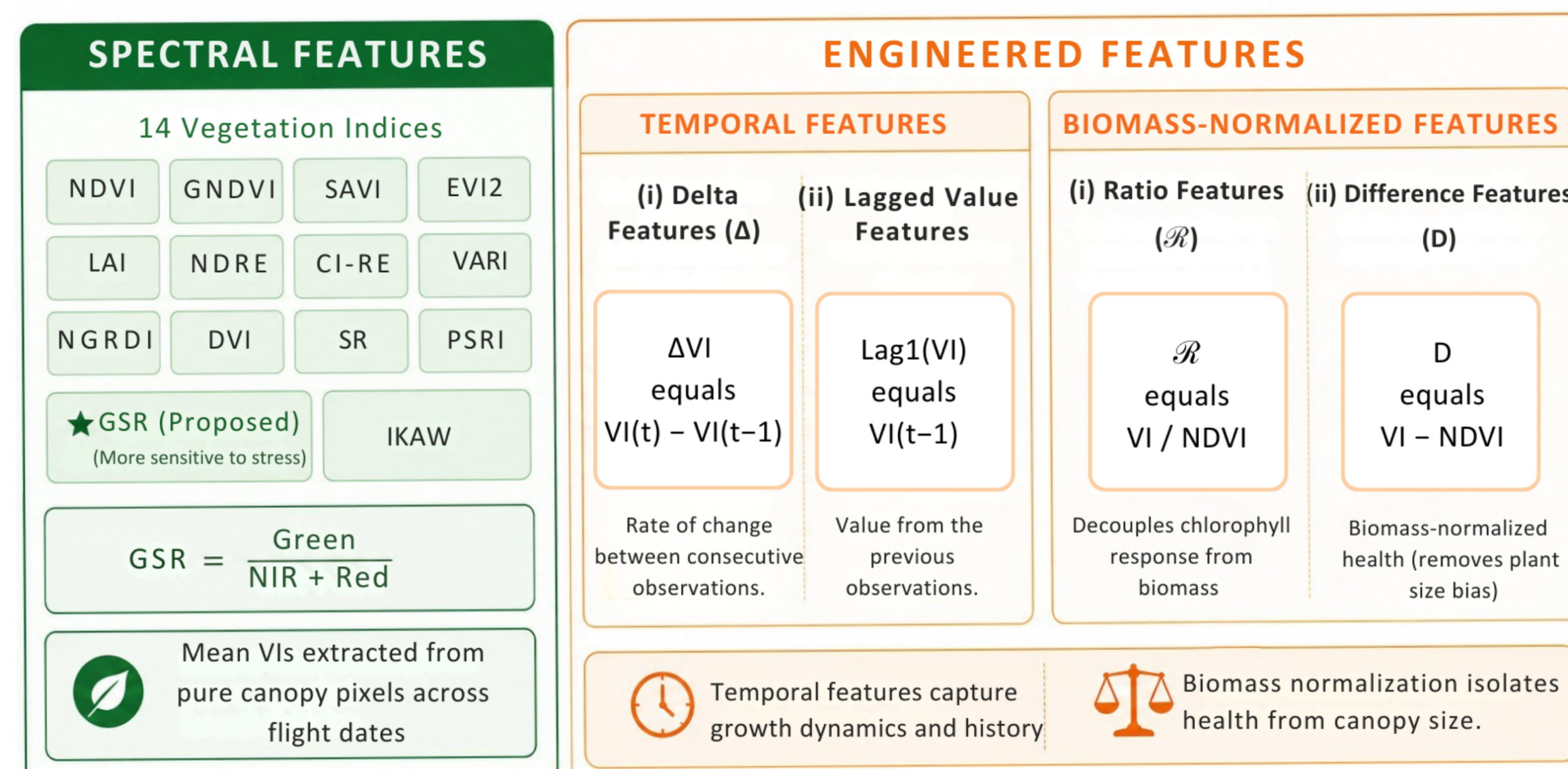


MULTISPECTRAL FEATURES

Standard Vegetation Indices (VIs) were initially extracted to characterize chlorophyll, biomass, & photosynthetic activity. Exploratory analysis revealed a biomass paradox, where dense canopy growth produced misleading positive correlations between VIs & disease severity.

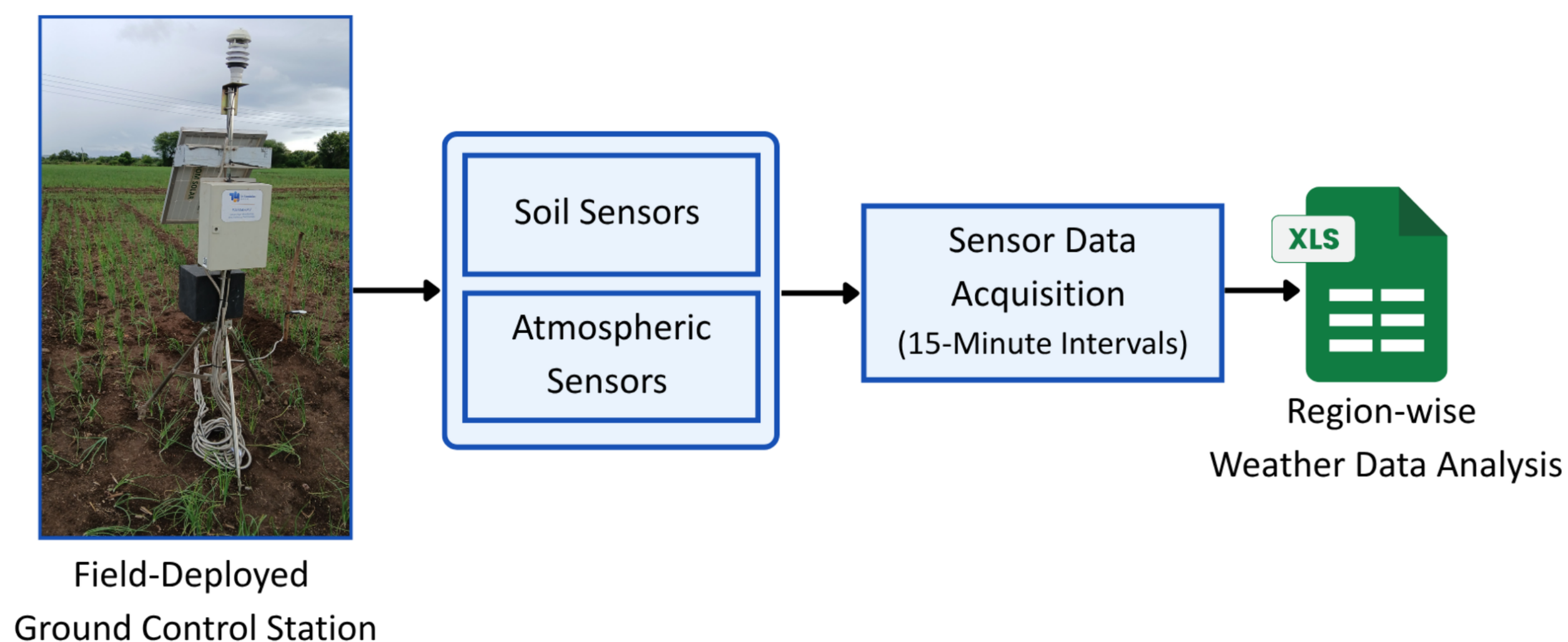


To reduce this confounding effect, temporal (Δ , lagged metrics) and biomass-normalized ($\mathcal{R}$, $\mathcal{D}$ relative to NDVI) features were engineered to capture growth dynamics and minimize canopy density effects, restoring the expected relationship between spectral response and disease severity.



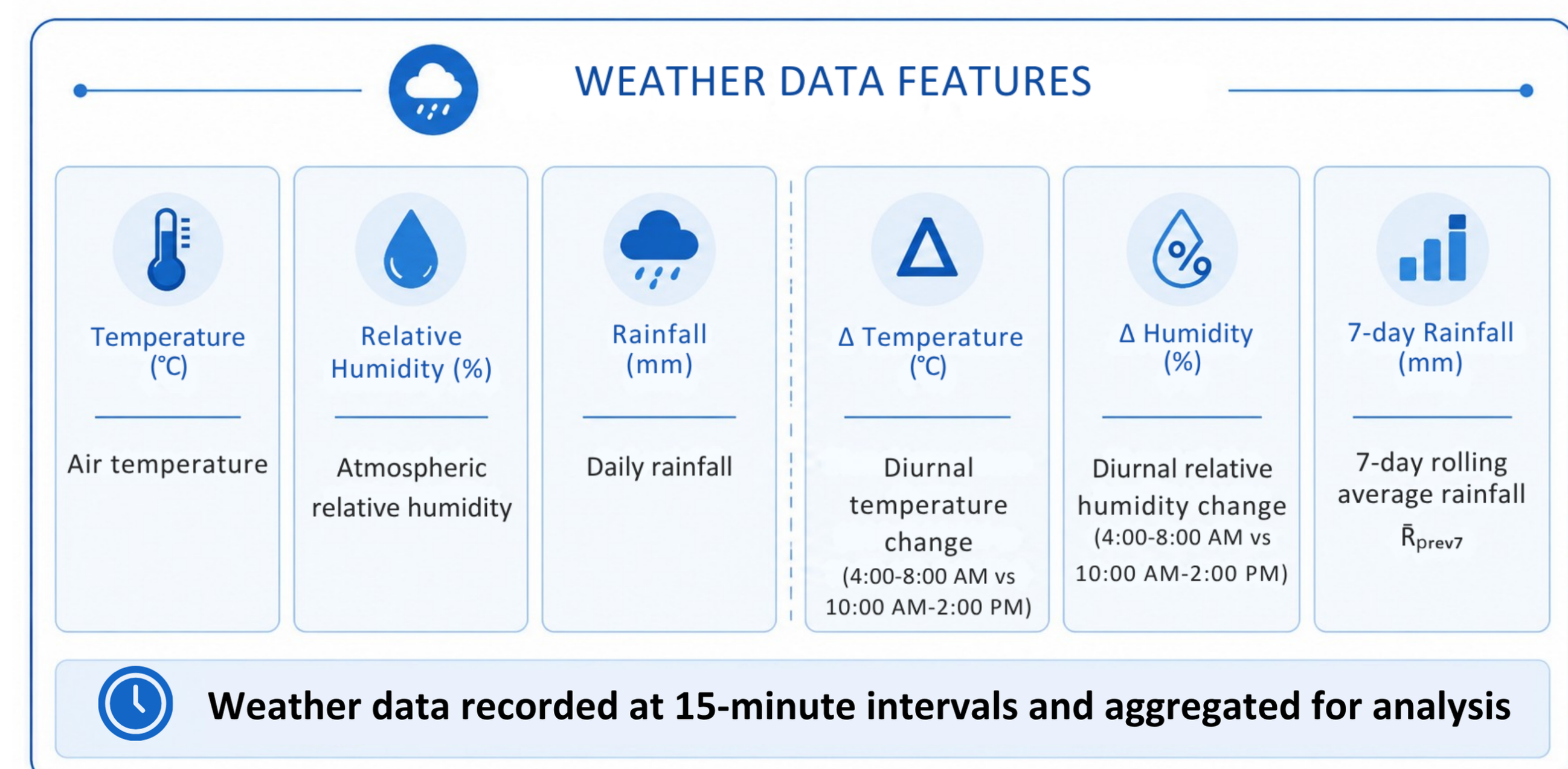
WEATHER DATA ACQUISITION & ANALYSIS

SAMBHAV™, a solar-powered IoT weather station, monitored soil and atmospheric conditions every 15 min, capturing temperatures of 22–31 °C and relative humidity of 35–85%.



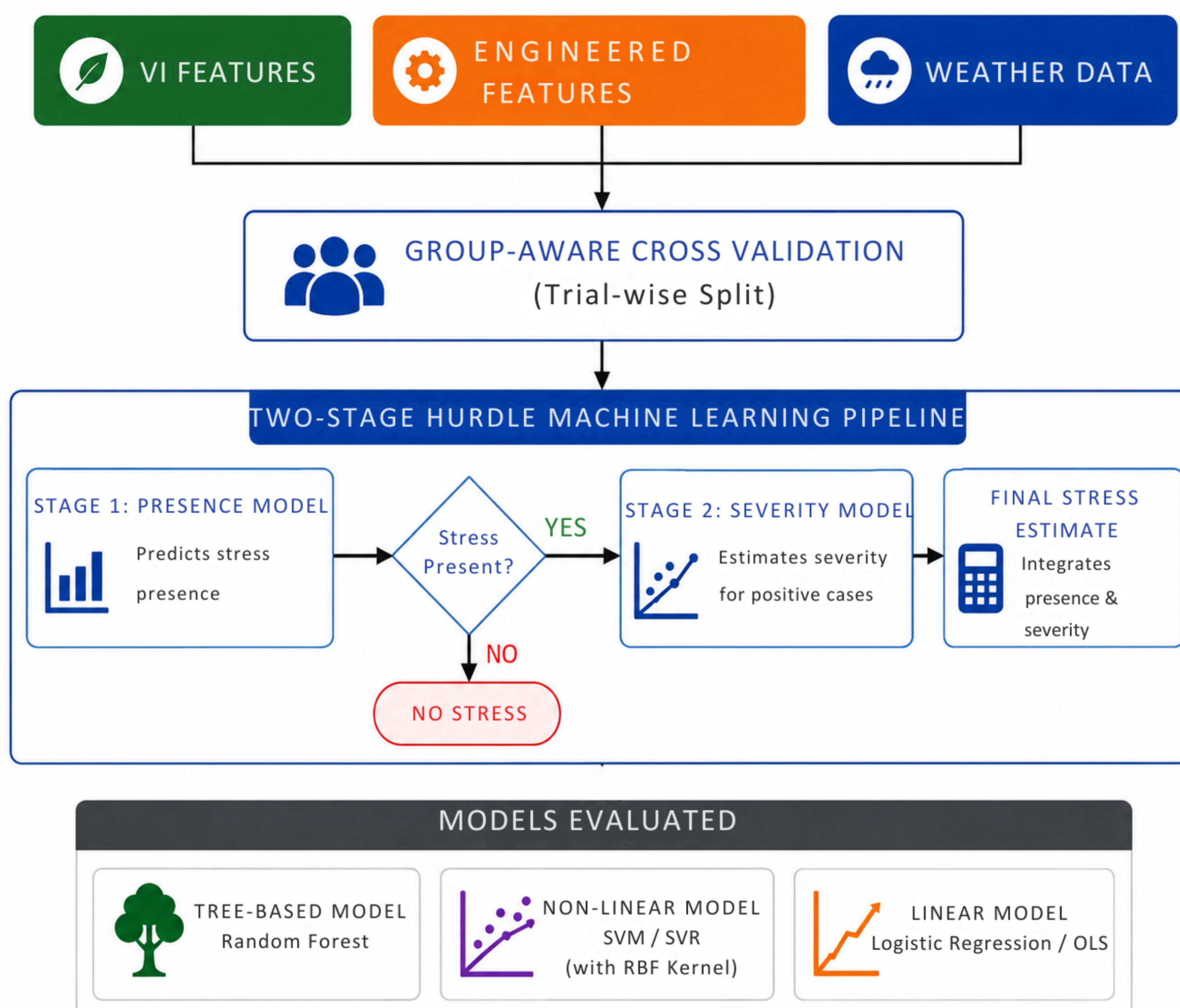
WEATHER FEATURES

Weather data were transformed into targeted features to capture microclimatic drivers of disease and pest incidence. ΔT and ΔRH represented diurnal changes, while R_t and R_{prev7} captured immediate and lagged moisture effects.



TWO STAGE HURDLE MACHINE LEARNING PIPELINE

To address zero-inflated and skewed stress observations, a two-stage hurdle framework integrates multispectral and weather features. Stage 1 predicts the presence or absence of stress, while Stage 2 estimates its severity when present. Random Forest, SVM/SVR, and linear models were evaluated using plot-wise group cross-validation to prevent temporal data leakage.

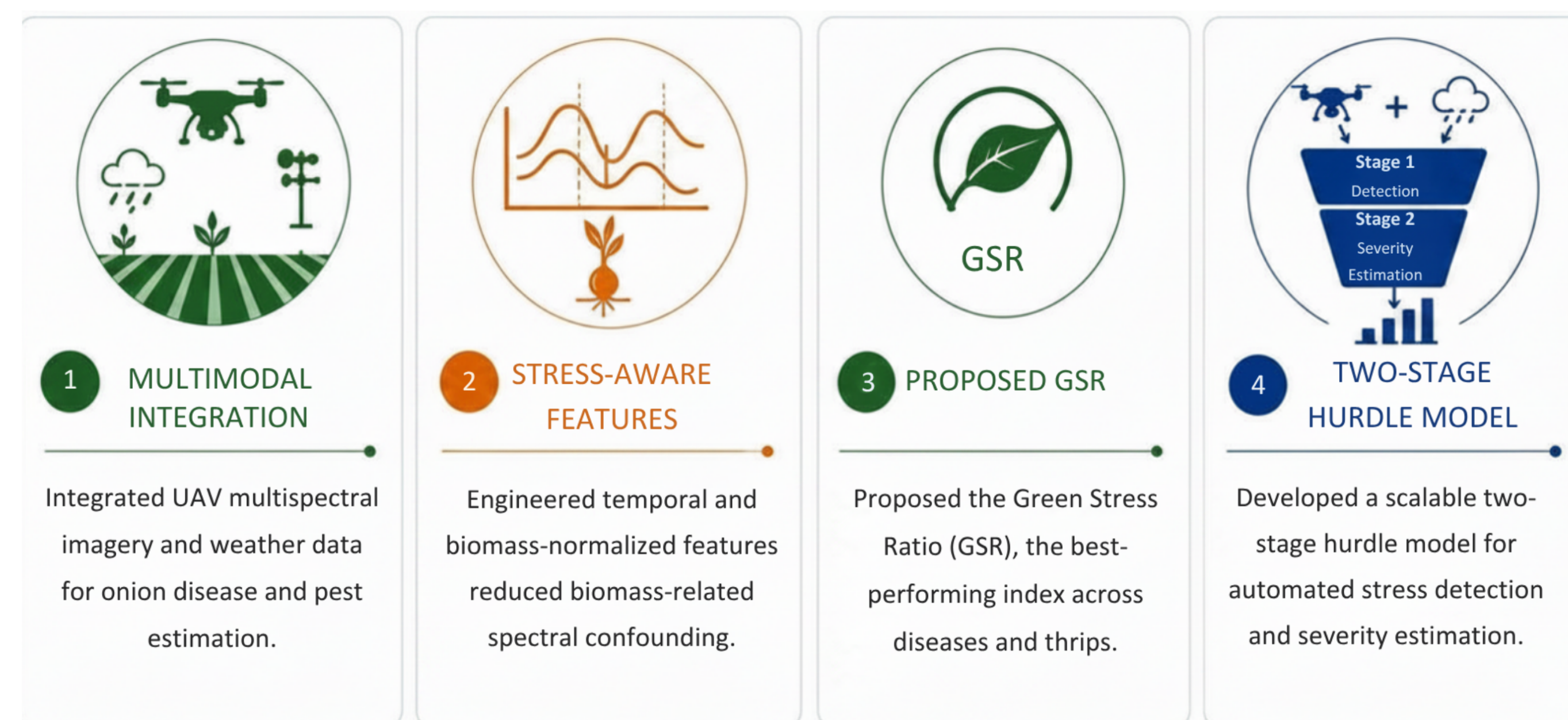


RESULTS

Stage-specific feature selection was performed using trial-independent correlation analysis. Temporal derivatives (Δ , Δ^2) were prioritized for stress detection, while baseline VIs and biomass-normalized features ($\mathcal{R}$, $\mathcal{D}$) were favored for severity estimation. Weather variables and days after transplanting (DAT) were retained across all models.

DETECTION (PRESENCE) Metrics: F1@0.5, AUC (Higher is better)				SEVERITY (NON-ZERO CASES) Metrics: RMSE, MAE (Lower is better), R ² (Higher is better)			
Best model for detection of all stressors: Random Forest				Best model varies by stressor			
Disease/ Pest	Best Model (Classification)	Best Performance		Disease/ Pest	Best Model (Regression)	Best Performance	
		F1@0.5	AUC			RMSE ↓	MAE ↓ R ² ↑
Stemphylium blight	Random Forest	0.80	0.85	Stemphylium blight	Random Forest	12.60	6.91 0.35 ★
Anthracnose	Random Forest	0.80	0.78	Anthracnose	Linear (logistic/OLS)	13.76	22.28 0.19
Purple blotch	Random Forest	0.82 ★	0.93	Purple blotch	Random Forest	7.95	4.06 0.17
Thrips	Random Forest	0.78	0.80	Thrips	Linear (logistic/OLS)	22.28	13.76 0.19
Best overall detection: Purple blotch (F10.8169, AUC = 0.9312)				Best severity overall: Stemphylium blight (R ² = 0.3545 with Random Forest)			

KEY FINDINGS



FUTURE PROSPECTS

