Dr. Pritesh Patil, 2026, 14:3
ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

International Journal of Science, Engineering and Technology
An Open Access Journal

Cloud Native Decision Support Framework for Crop Stress Monitoring Using Random Algorithm

Dr. Pritesh Patil, Vinay Basargekar, Shraddha Thorbole, Saurabh Rai
Department of Information Technology, AISSMS Institute of Information Technology, Pune, Maharashtra, India

Abstract- Satellite imagery has become an essential component of modern precision agriculture, enabling large-scale observation of crop health conditions. However, many existing monitoring solutions primarily provide vegetation indices, stress maps, or threshold-based alerts that can be difficult for farmers to interpret and act upon. To address this limitation, this study presents a cloud-native decision-support framework that integrates multispectral satellite observations, weather information, machine learning, and explainable recommendations. The proposed framework uses key vegetation indicators, including NDVI, NDWI, and SAVI, from Sentinel-2 satellite imagery and combines them with environmental parameters such as rainfall, temperature, and humidity. A Random Forest classifier is utilized to categorize crop conditions into four stress levels: healthy, mild, moderate, and severe. To enhance transparency and usability, an explainability module identifies the primary factors influencing each prediction and translates them into actionable recommendations, such as checking irrigation systems, assessing nutrient deficiencies, or maintaining routine field monitoring. The system follows a modular cloud-native architecture including data retrieval, preprocessing, feature engineering, model prediction, and dashboard visualization. The key contribution of the proposed framework is its ability to convert analytical results into actionable insights. Beyond detecting crop stress levels, it provides explanations for the underlying causes and recommends suitable management practices, enabling farmers to make informed decisions.

Keywords— Decision Support System, Random Forest Algorithm, NDVI , Sentinel-2 Imagery, Cloud-native approach

I. INTRODUCTION

The demand of sustainability in agriculture has never been high as now, as climate change is taking place all across the globe which demands the need for technologies that will foster and increase agricultural production, at the same time enable timely and informed crop management decisions. Many a times problems caused due to water scarcity, climate variability and many more other environmental factors are not detected early which leads to affect crop growth and yield directly. Early detection and mitigation of these problem are crucial for minimizing economic losses and ensuring efficient use of agricultural resources. Recent technologies such as spatial tech has emerged that can help in monitoring all the above aspects of the agriculture and help us detect early.

Freely available datasets such as from sentinel-2 satellite that can massively help in monitoring that provides multispectral bands of any given area on earth. These multispectral bands can be used to analyse vegetation stress in the area, and added with the weather information like rainfall, moisture content, this gives out a clear picture of geographical region and help in prediction of agricultural land yield. Although there has been a significant growth that has been done in this area but there remains a gap of creating a user centric system that can combine this all and directly be used in agricultural fields. The current systems are too complex from technical perspective, so in this proposed paper we are presenting a cloud-native framework that integrates satellite observations, weather context, and machine learning techniques to support crop stress assessment. By incorporating explainable prediction mechanisms and recommendation generation, the framework aims to enhance the accessibility and usefulnes

II. RELATED WORK

This research falls into three key areas: spectral stress sensing, using machine learning for stress recognition, and actually putting monitoring systems to work in real life.

1. Spectral Stress Sensing

The Sentinel-2 satellite is an open-source high resolution Earth Observation Satellite that can provide multispectral imagery for various Earth observation applications. It Samples 13 spectral bands across visible, near infrared (NIR) and short wave infrared (SWIR). The spatial Resolution of Sentinel-2 can go up to 60 meters and it's revisit time is 5 days, it is highly suitable for agricultural monitoring, vegetation analysis, and NDVI-based crop health assessment [1]. Spectral sensing directly means using Satellite images to directly predict the vegetation health of an given area over a period [2]. using these bands several other metrics got emerged that focuses on various aspects of plant health, such as NDWI, it's all about plant hydration rather than general health or greenness [3]. There's also SAVI, which does a better job of reducing soil reflections in sparsely vegetated zones [4].

2. Machine learning & Weather Fusion

Adding Machine learning and deep learning to spectral sensing pipeline improves stress detection massively. Various studies have been conducted that reviews deep-learning approaches for vegetation stress detection strongly shows progress in using image-based phenotyping and remote-sensing analytics [7]. A complementary line of work shows that fusing spectral information with meteorological context improves robustness by reducing ambiguity between visually similar stress states. Also Integrating weather data with vegetation based features improves water stress classification in a precision irrigation setting [8]. These evidence directly motivates our proposed framework where rainfall, temperature, and humidity are added to spectral descriptors so that not only low greenness should become a major factor for determining physiological stress.

3. Operational Deployment and the Explanation Gap

This is the major and final step in our framework that addresses how monitoring can be easily scalable and deployed. Using an Serverless approach for NDVI computations is highly scalable and cost efficient, that decoupling heavy raster processing from fixed infrastructure [9]. Our earlier AWS-based Sentinel-2 pipeline confirmed that automated ingestion, processing, and alerting are technically feasible for agricultural monitoring [10], and a related Sentinel-2 smart-agriculture framework reinforced the usefulness of drought and crop-stress monitoring from satellite imagery [11]. Using an cloud native approach makes system strong on computation and increasingly mature operationally.

The emergence of Human-Centered Artificial Intelligence has highlighted the importance of developing smart farming systems that combine automation with farmer-centric decision support, contributing to the vision of Agriculture 5.0 and sustainable agricultural practices [12]. Random Forest has become one of the most widely used machine learning algorithms due to its robustness, ability to handle non-linear relationships, and strong predictive performance across diverse datasets [13]. More advances in model interpretability have improved the transparency and explainability of machine learning predictions, enabling users to better understand the factors influencing model outcomes and enhancing trust in data-driven decision-making systems [14].

These three together shows that there is a clear gap and need of creating a system that can easily detect vegetation stress at scale and also recommend mitigations for that particular area. This Paper presents a Scalable Framework that shows a system can be scaled easily using cloud native approach and by creating a user centric system. The framework proposed here is positioned precisely at this gap: it retains the spectral and operational strengths of prior work while adding the contextual classification and explanation layers needed to turn monitoring into decision support.

III. METHODOLOGY

In the proposed system, we have taken a Area of Interest(AOI) and a date window as an input along with other required variables that together gives out stress-severity and recommendations. The overall flow is shown in Fig. 1, the pipeline is linear independent component which can also be seen as an replaceable component.

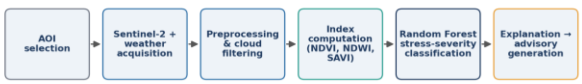

Fig. 1. Flow Diagram of proposed system

1. Data Acquisition & Preprocessing

We Defined the AOI and date window, and collected Sentinel-2 surface reflectance imagery along with daily weather observations covering the same temporal range. Sentinel-2 imagery and vegetation indices are the better choice here as it helps in calculating various metrices using spectral bands which is used in calculating NDVI, NDWI and SAVI. Weather records are also important as it supplies rainfall, temperature and humidity data which provides the contextual signal used later to disambiguate the cause of stress.

In pre-processing that happens over the cloud, it handles the fetching of required AOI sentinel-2 images and cloud masking if any within the defined date window. Later these images are all together are stacked to give a proper definitive angle around the AOI, a cloud-free composite is formed so that subsequent indices are computed on clean, geometrically aligned pixels.

2. Spectral Index Computation

In the proposed system we have majorly used these three indices which are derived from spectral bands obtained via Sentinel-2 data. NDVI defines majorly vegetation stress and is computed from Red and Near-Infrared(NIR) bands; NDWI represents water stress in the AOI and it is computed as short-wave-infrared(SWIR) reflectance, and finally the SAVI indices uses the same Red-NIR pair as NDVI with a soil adjustment factor L to reduce soil-background effects in sparse canopies shown in the equation [1][2][3].The indices are defined as follows:

$$NDVI = (NIR - Red) / (NIR + Red) \quad (1)$$

$$NDWI = (NIR - SWIR) / (NIR + SWIR) \quad (2)$$

$$SAVI = [(1 + L)(NIR - Red)] / (NIR + Red + L) \quad (3)$$

where L is the soil-adjustment factor, typically set to 0.5 for canopies of intermediate density. NDVI and SAVI capture greenness and canopy development, while NDWI adds an explicit moisture dimension that is often the earliest indicator of water stress.

3. Feature Construction & Classification

Each index raster is summarised into field-level descriptors rather than used pixel by pixel. For every index, the mean, standard deviation, percentile spread, and stress-pixel proportion are computed over the clipped AOI, yielding a compact statistical signature of field condition. Weather features rainfall accumulation, a short-term rainfall deficit indicator, mean temperature, a temperature anomaly, and mean humidity are appended to the same vector, along with a temporal NDVI-change term that captures the direction and rate of recent canopy change. The result is a single tabular record per field per date.

A Random Forest classifier is used to predict the severity class from the combined feature vector. Random Forests are well suited to nonlinear relationships among heterogeneous tabular features, are robust to moderate noise and class imbalance, and natively expose feature-importance information that supports explanation [13]. The ensemble learns interactions among indices and weather variables that would be difficult to encode by hand, which is the principal reason it outperforms a fixed-threshold rule.

4. Explanation and Advisory Mapping

This is the final stage where SHAP is applied [14], it's an permutation importance as a lightweight alternative, as there can be possibilities where NDVI or other metrices comes out to be low and rainfall data is proper, this both should contribute into the Explanation and short recommendation, mutual importance to be given to all the aspects such that result doesn't get affected from only one aspect. This part is important in the pipeline as it acts a the middleware between how system converts, the calculated crop health indicators into a stress class and then into farmer recommendations.

Lastly, metrics such as accuracy, precision, recall and F1-score are used for model performance by macro averaging them across four classes so that minority severity level are not masked by the dominant healthy class.

IV. IMPLEMENTATION FRAMEWORK

The framework is articulated in an pipeline which is structured in an cloud native environment using event driven workflows and services, this also ensures the scalability is ensured and mitigates the unnecessary infrastructure complexity. There are three such horizontal layers that depicts our system called data, Processing and delivery layer shown in Fig. 2.

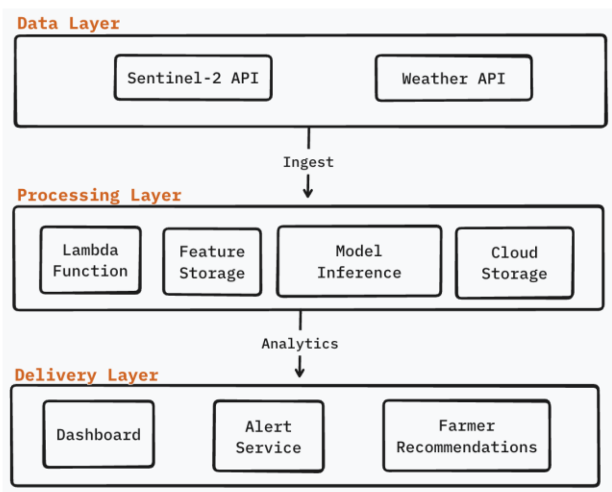

Fig. 2. System architecture of the proposed cloud-native decision-support framework.

The data layer define the sources of area from which the data is been collected, there are two such parallel acquisition steps in the data layer: farm boundary images from Sentinel-2 Satellite and retrieval of corresponding weather records for the same date range. From these parallel outputs the data goes into cloud object storage of the Processing layer. Here in processing layer, pre-processed functions are defined that does AOI clipping, cloud filtering, harmonisation, and vegetation index generation, and a feature-construction stage aggregates of the raster outputs these combining with meteorological variables into a compact tabular record to be run into our trained Random forest model later in the pipeline.

The trained Random forest model predicts the severity classes based upon the given data. The result from here goes into explanation service which extracts top contributing variables for that prediction and passes them into advisory module, which ends up generating short recommendation based upon records. The file layer which is the user faced layer called Delivery Layer, it presents this whole internal calculation into an simple user centric dashboard, and it dispatches a email or raises alert based upon the output, ending not at an index value but at an interpretable advisory.

The design deliberately emphasises modularity. The data layer can use any STAC compatible Sentinel-2 source APIs and respective weather data from any open-source real-time weather database providing it's APIs, while the inference and explanation services remain independent of the data provider. Compared with a pure NDVI alert pipeline, this implementations provides independent sources and contextual prediction and it's explanation without sacrificing the deployment simplicity already demonstrated in earlier serverless work [8], [10].

V. RESULTS AND DISCUSSION

1. Experimental Setup

The evaluation was done under specific configuration over a period shown in Table I, it shows all the parameters used over the samples are the obtained in the rainfall season across a single growing season. These samples were split into 70% / 30% into training and test sets respectively, and the Model comprising of Random Forest as a algorithm was tuned using five-fold cross-validation on the training partition.

Table I. Dataset and Experimental Configuration

Parameter	Value
Study area	~120 ha rainfed cropping cluster (illustrative AOI)
Season / period	Kharif 2023 (June–October)
Satellite source	Sentinel-2 L2A (MSI), 10 m bands
Acquisition cadence	~10-day cloud-free composites
Weather source	Daily rainfall, temperature, humidity (open weather API)
Field-level samples	410
Train / test split	70% / 30% stratified (N_test = 124)
Stress classes	Healthy, Mild, Moderate, Severe
Classifier	Random Forest (300 trees, max_depth 18, min_samples_leaf 3, balanced)
Validation	5-fold cross-validation on training set
Metrics	Accuracy, Precision, Recall, F1 (macro-averaged)

2. Classification Performance

Three Critical configurations shown in Table II and Fig. 5 : an NDVI-threshold baseline, a Random Forest trained on spectral indices only, and the proposed Random Forest trained on spectral indices together with weather Data. Our proposed multi-source model trained on Random Forest performs strongest out of all in performance, increasing the macro F1-score from 0.64 in the threshold baseline to 0.87 in the final framework.

TABLE II. COMPARATIVE CLASSIFICATION PERFORMANCE

Configuration	Accuracy	Precision	Recall	F1 Score
NDVI-threshold baseline	*0.67*	*0.65*	*0.64*	*0.64*
RF (spectral indices only)	0.80	0.79	0.78	0.79
RF (spectral + weather)	0.88	0.87	0.87	0.87

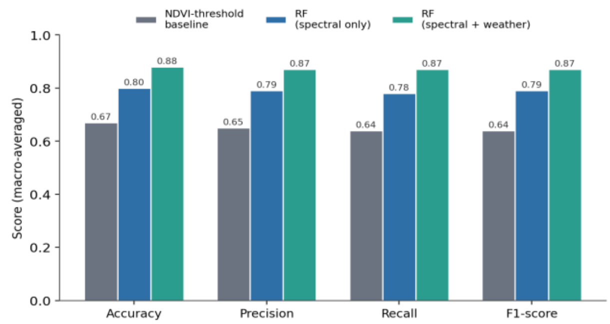

Fig. 3. Comparative performance of the three configurations across all metrics.

The NDVI-threshold configuration is great for rapid screening but sensitive to ambiguity from sparse canopy, exposed soil, atmospheric effects and crop growth stage. Adding Random Forest exponentially improves class separation, as random forest is an tree based algorithm because non-linear interactions among spectral indicators are learned [13], increasing F1 0.64 to 0.79. More additional gain comes from rainfall, temperature, humidity which raises F1 to 0.87. This pattern is consistent with the crop-stress literature, which shows that weather context improves interpretation of vegetation signals and helps separate true physiological stress from merely low greenness [6], [12].

In practical use, the final output simply doesn't show "stress detected" but gives out the severity level based upon satellite data and weather data combined, thus the severity level is used to express the recommendation towards the farmer land. In this sense the framework converts a stress map into decision support, which is the contribution that most clearly differentiates it from index-only workflows.

VI. CONCLUSION AND FUTURE SCOPE

This Paper presents a study that shows our proposed system can use satellite monitoring and make decision based upon that, rather than simply calculating NDVI value and it's map, the proposed system also enables real time alerts and recommendations for farmers to work on, the proposed framework combines Sentinel-2 vegetation Indices, short term weather context trained upon Random-forest based model that results out stress-severity, which factor must have likely caused it, and what action the farmer must take upon. On the illustrative configuration used here, adding weather context raised the macro F1-score from 0.79 to 0.87 over an indices-only model and from 0.64 over the threshold baseline, while the feature-attribution analysis showed that moisture and rainfall variables contribute information beyond greenness alone.

In the proposed System several direction and enhancement can be done. We can Integrate IoT based mechanism to make the quality of data more robust and accurate and also can calculate soil moisture and give out more in-depth analysis and recommendations to the farmers. We can also train crop specific model and evaluate across multiple season and regions to make model better and widely usable. Finally, a multilingual mobile interface with a farmer-feedback loop would improve usability and allow the advisory layer to be corrected and improved over time, helping to bridge the remaining gap between remote-sensing analytics and day-to-day farm action.

REFERENCES

1. M. Drusch et al., “Sentinel-2: ESA’s optical high-resolution mission for GMES operational services,” Remote Sens. Environ., vol. 120, pp. 25–36, 2012.
2. C. J. Tucker, “Red and photographic infrared linear combinations for monitoring vegetation,” Remote Sens. Environ., vol. 8, no. 2, pp. 127–150, 1979.
3. B.-C. Gao, “NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space,” Remote Sens. Environ., vol. 58, no. 3, pp. 257–266, 1996.
4. A. R. Huete, “A soil-adjusted vegetation index (SAVI),” Remote Sens. Environ., vol. 25, no. 3, pp. 295–309, 1988.
5. A. Huete, K. Didan, T. Miura, E. P. Rodriguez, X. Gao, and L. G. Ferreira, “Overview of the radiometric and biophysical performance of the MODIS vegetation indices,” Remote Sens. Environ., vol. 83, no. 1–2, pp. 195–213, 2002.
6. K. Berger et al., “Multi-sensor spectral synergies for crop stress detection and monitoring in the optical domain: A review,” Remote Sens. Environ., vol. 280, Art. no. 113198, 2022.
7. A. Holzinger, I. Fister, I. Fister, H.-P. Kaul, and S. Asseng, “Human-Centered AI in smart farming: Towards Agriculture 5.0,” IEEE Access, vol. 12, pp. 62199–62214, 2024.
8. P. Patil, S. Thorbole, V. Basargekar, and S. Rai, “Spatio-temporal analysis of vegetation stress using serverless approach in precision agriculture,” in Proc. 6th Int. Conf. IoT Based Control Networks and Intelligent Systems, 2025, pp. 828–833.
9. N. Paul, G. C. Sunil, D. Horvath, and X. Sun, “Deep learning for plant stress detection: A comprehensive review of technologies, challenges, and future directions,” Comput. Electron. Agric., vol. 229, Art. no. 109734, 2024.
10. L. Iacono, D. Pacios, and J. L. Vázquez-Poletti, “SNDVI: A new scalable serverless framework to compute NDVI,” Front. High Perform. Comput., vol. 1, 2023.
11. T. Ahmed, N. Javed, M. Faisal, and H. Sadia, “A framework for smart agriculture system to monitor the crop stress and drought stress using Sentinel-2 satellite image,” in Algorithms for Intelligent Systems, Singapore: Springer, 2023, pp. 345–361.
12. K. Wang and Y. Jin, “Mapping walnut water stress with high resolution multispectral UAV imagery and machine learning,” arXiv:2401.01375, 2023.
13. L. Breiman, “Random forests,” Mach. Learn., vol. 45, no. 1, pp. 5–32, 2001.
14. S. M. Lundberg and S.-I. Lee, “A unified approach to interpreting model predictions,” in Adv. Neural Inf. Process. Syst. 30, 2017, pp. 4765–4774.