Journal of Namibian Studies, 35 S1 (2023): 2016-2028 ISSN: 2197-5523 (online)

Coimbatore Meteorlogical Data Based Phenology Stage Prediction With NDVI With Respect Vegetables Of Horticulture Crops

Dhanya K[1] , B.Jayanthi[2]

[1]Research Scholor,Department of Computer Science,Rathnavel Subramanian College of Arts and Science,Sulur. Mail id: dhanyakmail@gmail.com
[2]HOD,Department of Computer Science,Rathnavel Subramanian College of Arts and Science,Sulur. Mail id: jayanthi@rvsgroup.com

Abstract:

The monsoon season has a significant impact on Indian agriculture. As a result, the yield of food grains varies from year to year. Every variety of crop interacts with the nutrients in the soil in a unique way, releasing and absorbing different nutrients. Horticulture is the production, use, and improvement of horticultural crops, including fruits, vegetables, spices, and sauces as well as ornamental, plantation, medicinal, and aromatic plants. Crop phenology is crucial for overseeing farming operations, calculating crop yields, and assessing agroecosystems. The phases of crop growth can frequently be observed from the ground; however this takes time and has minimal regional variation. An artificial intelligence (AI) algorithm to manage the live streamed spatial spectral meteorological data and estimate the crop phenology through Normalized Difference Vegetation Index (NDVI) has been offered as a solution to the problem of climate change and weather pattern understanding.

Keywords: - Crop Yield, Meteorological data, Phenology, Prediction, Gradient Boosting Regression, Horticulture.

1. Introduction

Indian economy is based on its agriculture, which contributes significantly to the GDP. To meet the needs of its citizens for food, a government must cultivate healthy, productive crops. Crop monitoring with human intervention becomes very difficult due to the variety of the land, weather patterns, geographic locations, defenses against diseases, and natural calamities [1]. As the world's population rises and extreme weather events like droughts, heat waves, and major storms become more often or more violent owing to climate change, the significance of crop yield forecasts for food security is anticipated to rise. India's agriculture is heavily reliant on the monsoon season. Food grain production as a result varies year after year. A year of abundant cereal production is frequently followed by one of severe shortage.

Specific nutrients in the soil are depleted when the same crop variety is planted repeatedly. Every type of crop interacts differently with the soil's nutrients and releases and absorbs various kinds of nutrients. Because crop rotation replaces nutrients that are unavailable or absorbs nutrients that are abundant, it promotes soil fertility by regulating deficiency or excess nutrients.

1.1. Major Crops in India

In India, the agricultural season runs from July to June. Based on the monsoon, the Indian cropping season is divided into two primary seasons: (i) Kharif and (ii) Rabi. Rabi cropping season runs from October to March (winter), while the kharif cropping season runs from July to October during the south-west monsoon. Summer crops are those that are cultivated from March to June. Table 1 show the different variety crop and type of each crop grown in India.

Table 1: Major Crops Grown in India

Types of Crops	Meaning	Major Crops
Food grains	Plants grown for human consumption	Rice, Wheat, Maize, Millets, Pulses and Oilseeds
Commercial Crops	Crops that are raised for sale and are either sold in their raw or semi-processed forms	Cotton, Jute, Sugarcane, Tobaccco and Oilseeds
Plantation Crops	Crops raised on vast estates' worth of plantations	Tea, Coffee, Coconut and Rubber

Horticulture	Areas of agriculture that are used to raise fruits and vegetables	Fruits and Vegetables

1.2. India's Horticulture

Horticulture is the production, use, and enhancement of horticultural crops like decorative, plantation, medicinal, and fragrant plants as well as fruits, vegetables, spices, and sauces. Pomology and olericulture, two subfields of horticulture, deal with the production of plants for food, whereas floriculture and landscape horticulture deal with ornamental plants. The horticulture business in India has established itself as a successful and productive industry that is expanding quickly. India is the world's second-largest producer of fruits and vegetables, behind China, according to the Agricultural and Processed Food Products Export Development Authority (APEDA). Horticulture's productivity has risen significantly from 8.8 tons per hectare (TPH) in 2001-02 to 12.1 TPH in 2020-21, causing a strong growth in production and acreage that has outpaced the production of food grains since 2012-2013.

1.3. Challenges of Horticulture Sectors

Small operational landholdings, a lack of irrigation, and inadequate soil management are just a few of the many issues the horticulture industry faces, which prevent it from reaching its full potential.

- Reduced soil fertility can result in lower yields and worse-quality food when it is combined with poor soil management techniques like over-tilling, over-fertilizing, and mono cropping.
- Another major issue that could result in crop failures and losses is climate change, which includes shifting weather patterns, droughts, floods, and other natural catastrophes. This could have an impact on the sector's overall output.
- Constraints on land leasing are a problem, especially for small farmers who might not have access to enough land for farming.

India is known as one of the world's top producers of fruits and vegetables, and the horticulture industry has constantly produced well. Predicting the climate changes and maintaining adequate water supply will result in good yield of the crop. By putting these procedures in place, the horticulture industry may grow even more and boost farmer's profitability. Crop phenology is important for managing agriculture, estimating crop yields, and evaluating agro ecosystems. Crop growth phases are often seen from the ground, which takes time and has little geographic variation. Land surface phenology (LSP) and agricultural growth stages have been mapped using remote sensing Vegetation Index (VI) time data, primarily after the growing season [8]. To overcome the climate change problem and understanding the weather pattern proposed an AI model to handle the live streamed spatial spectral meteorological data and estimating the crop phenology through NDVI.

Further sections of this paper are Literature review (Section 2), Methodology (Section 3), Result and Discussion (Section 4) and Conclusion and Future work (Section 5).

2. Literature Review

In [2], the study evaluates climate impacts on maize, sorghum, and soybean yields in the region using extensive crop yield and climate datasets from 1968-2013. Climate variability significantly impacts global crop yields, with location, crop type, and irrigation playing a crucial role. The U.S Great Plains, with its significant role in national food production, is an ideal area for investigating climate impacts on food production. Results show that temperature and precipitation factors explain a quarter of regional average yield variability, with irrigated yields showing increased resilience and effective mitigation strategies.

In [3] states Crop yields are heavily influenced by weather, and climate change impacts are increasingly being studied in recent years. A new approach to yield modeling uses a semi parametric variant of a deep neural network, which can account for complex nonlinear relationships and parametric structure. This method outperforms classical statistical methods and fully-nonparametric neural networks in predicting yields during model training. The approach shows significant negative impacts on corn yield, but less severe than classical statistical methods.

In [4], a machine learning prediction system is proposed to predict the yield of six crops in West African countries, including rice, maize, cassava, seed cotton, yams, and bananas. The system uses climatic, weather, agricultural, and chemical data to help decision-makers and farmers predict crop yields. The system uses decision tree, multivariate logistic regression, and k-nearest neighbor models. The decision tree model performs well with a coefficient of determination (R^2) of 95.3%, while the K-Nearest Neighbor model and logistic regression perform similarly with R^2 = 93.15% and 89.78%, respectively. The prediction results of the decision tree and K-nearest neighbor model are correlated to the expected data, proving the model's efficacy.

In [5], the study explores a phenological approach using a remote sensing vegetation index to predict corn yield in 314 US Corn Belt counties. The study used the Moderate Resolution Imaging Spectroradiometer (MODIS) data product MOD09Q1 to calculate the normalized difference vegetation index (NDVI) time series. The study divided the corn growing season into four growth phases and calculated phenological information metrics (duration and rate) for each growth phase. The maximum correlation NDVI (Max-R2) was found to be the most significant correlation with corn yield. Three groups of yield regression models were built, including univariate, multivariate, and multivariate models. The results showed that most phenological metrics had a statistically significant relationship with corn yield, and models with phenological metrics could predict yield before harvest in three regions.

In [6], the study used PhenoCams to track crop phenology and compare the effects of organic vs. conventional farming with intensive or conservation tillage on a pea-barley mixture and winter wheat. Crop phenology is a crucial aspect of plant performance and yield, but the impact of cropping systems (CS) on this is less understood. The study found that CS significantly affected phenological metrics of both crops, with less pronounced effects in unfertilized pea-barley mixtures and stronger effects for early-season harvests. Organic farming caused an initial growth lag and a shorter duration of the period of stable GCC in winter wheat. Winter wheat in reduced or no-tillage systems showed a tendency toward delayed phenology compared to intensive tillage. The study demonstrated that delayed phenology was an important factor causing lower yields in organic farming compared to conventional farming.

In [7], a new approach for within-season phenology estimation for cotton at the field level uses Earth observation vegetation indices and numerical simulations of atmospheric and soil parameters. Crop phenology is crucial for crop yield estimation and agricultural management. The method is unsupervised to address sparse ground truth data, making most supervised alternatives impractical in real-world scenarios. Fuzzy c-means clustering is applied to identify principal phenological stages of cotton and predict transitional phases between adjacent stages. The model was tested against a baseline model, showing it significantly outperforms the baseline one, indicating its promising nature.

In [9], the study uses Time-Weighted Dynamic Time Warping (TWDTW) to generate crop samples automatically in two target areas.Crop type mapping is crucial for food security applications, but traditional methods often lack training samples. The resulting samples are refined using proximity measures from Random Forests (RF) and Sentinel-2 time-series. The methodology works well for classes with reduced inter-classes similarity, such as sugar beets or grains, but less for high inter-classes similarity crops like potatoes. The methodology works well for balanced crop samples but favors prevalent classes in areas with imbalanced crops, resulting in low accuracy for minority crops. Despite these shortcomings, the proposed methodology offers a viable option for generating crop samples in regions with few ground labels.

3. Methodology

Proposed a statistical approach to monitor the vegetation and determine the crop phenology which improves the vegetation. Numerical Analysis performed to identify the causes (Climate Variables) affecting NDVI. As shown fig.1 the proposed architecture flows.

Figure: 1. Proposed Architecture

3.1. Dataset

Particularity focused on horticulture sector and concentrated Coimbatore region. Fruits and vegetables including tomato, brinjal, bhendi, and onion, as well as spices like turmeric and flowers like tube rose and jasmine, are the main horticulture crops grown in Coimbatore district. Planning farm schedules for planting seeds and seedlings, applying fertiliser, irrigating crops, and monitoring crops can benefit from meteorological data. Radiation, air temperature, humidity, and wind speed are the main climatic variables influencing evapotranspiration [10]. There are different Weather data retrieved for 2 years using Meteosat API. Moving averages are computed 2 weeks period applying a rolling window. Temperature, dew point, wind direction, wind speed, cloud cover, cloud layer(s), ceiling height, visibility, the present weather, and the precipitation are the data retrieved through the API.

3.2. Preprocessing

Data streamed from the weather website contain data dropping, due to that missing values and outlier present in the dataset as shown in fig.2. To handle the difficulty of above state proposed an Enhanced Recursive Random Outlier Imputing Algorithm (ERROI). The ERROI technique divides a dataset into numerous partitions at random in order to pinpoint abnormalities. The process is repeated until every data point is a component of a partition. Typically, fewer partitions are needed to separate an anomaly than are necessary to separate a regular point. An anomalous data point is supposed to be farther apart from other points and therefore simpler to distinguish (separate). A typical data point, on the other hand, is likely to be aggregated closer to the bigger set and will consequently need more partitions (splits) for separating that point.

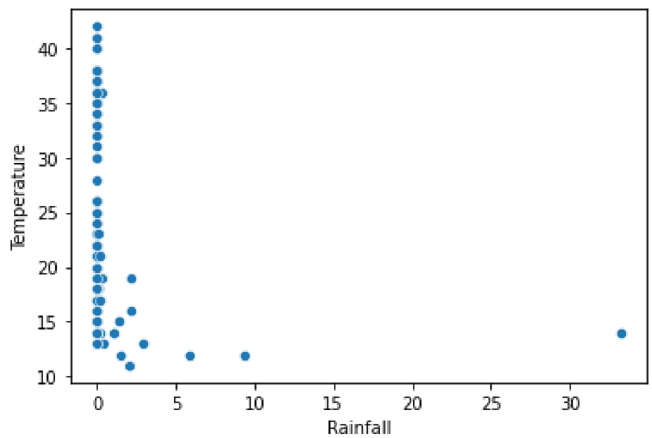


Figure: 2. Outlier

3.3. Feature Selection

Temperature, dew point, wind direction, wind speed, cloud cover, cloud layer(s), ceiling height, visibility, the present weather, and the precipitation are the data retrieved through the API. Important Features selected using SS-NMRA algorithm which was proposed in [11] earlier work. Features such as Temperature, dew points, humidity, precipitation, rainfall are selected. Person correlation depicted to identify the features of high important.

3.4. Prediction

Prediction of phenology performed with Gradient Boosting Regression used than the regular regression algorithms. A potent boosting method is gradient boost.

It creates a strong tree by gradually joining weak trees to increase the model's accuracy. It achieves low bias and low variance in this way. Gradient Boosting is a potent boosting approach that turns numerous weak learners into strong learners (as shown in fig.3). Each successive model is trained using gradient descent to minimise the loss function of the preceding model, such as mean square error or cross-entropy. Most of them include support for handling categorical features, some of them accept missing values natively, and train quicker than other modes, especially on bigger datasets.

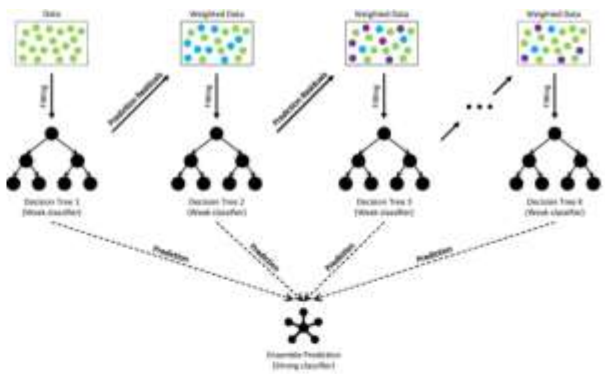


Figure: 3. Gradient Boosting Regression Architecture

4. Result and Discussion

Phenology Estimation performed based on the NDVI with respect to rate of climate, temperature, and rainfall. Vegetation stage identified and estimated the yield. Nearly 8 stage of phenology for vegetation has to be estimated. As per the stage of thread happen since used Last 5 stage of vegetation here i) Stem Elongation, ii) Inflorescence Emergence, iii) Flowering, iv) Fruit Development, and v) Ripening. NDVI plotted based on the weekly average and Daily average on phenology dates interpolated, which is depicted in the fig.4.

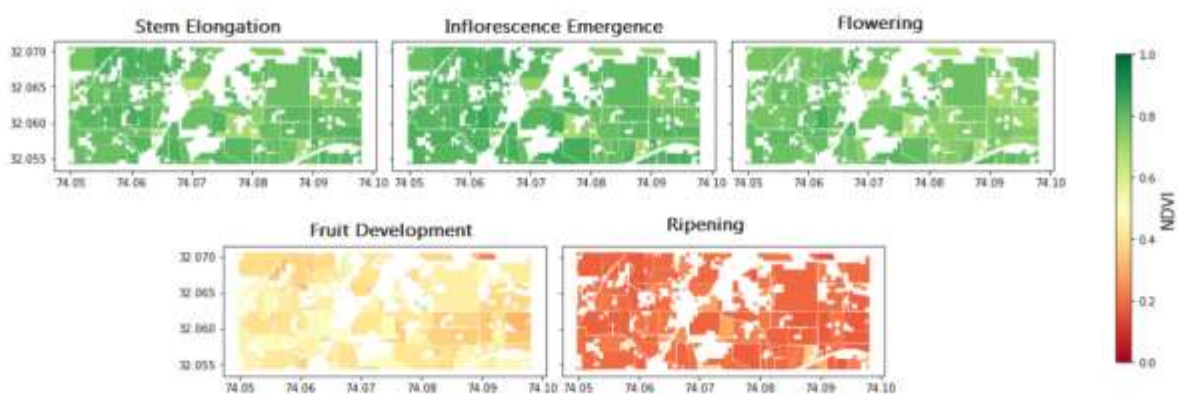


Figure: 4 NDVI based Phenology

The regional level depiction NDVI based ground truth explored for Coimbatore region on April month as shown in fig.5.

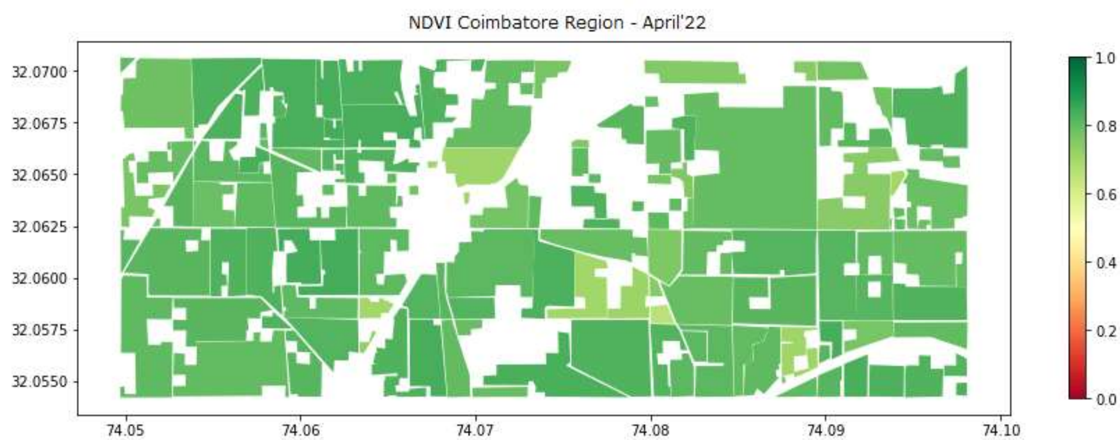
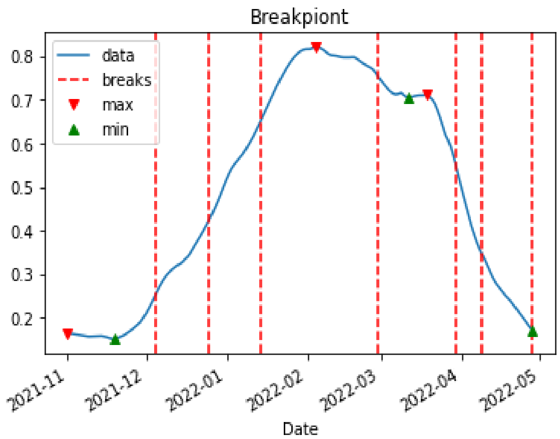


Figure: 5 NDVI Based Region Map

Breakpoints determined based on the 5 phenology stages and determined the date of Maxima's and Minima's of data points. The model fit on the prediciton based on the breaks and NDVI. Plot generate with respect to data, breaks, max, and min values as shown in fig. 6. Eliminated the smaller of the two discovered maxima as well as the minimum between those maxima if they occur within 15 days of one another. Avoided choosing a min-max-min section with a relevant maximum and two irrelevant minima on either side of it. Calculate the area of the triangle formed by the minimum and the two maxima. With the above constrain monthly average of climate change depictd basd on the NDVI as shown in fig. 7



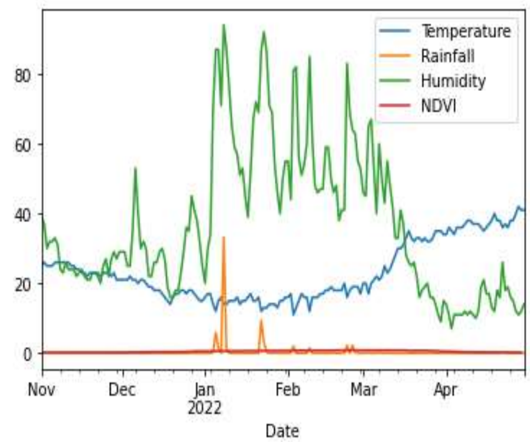


Figure: 6. Breakpoints

Figure: 7. Monthly Climate Condition

With respect to the number of days necessary for vegetation development, displayed the derivative, or rate of change of temperature and rainfall, in coimbatore area. The function coefficient determines the equation for the rate of change of climate variables in relation to the days that are essential for vegetation development as shown in fig.8.

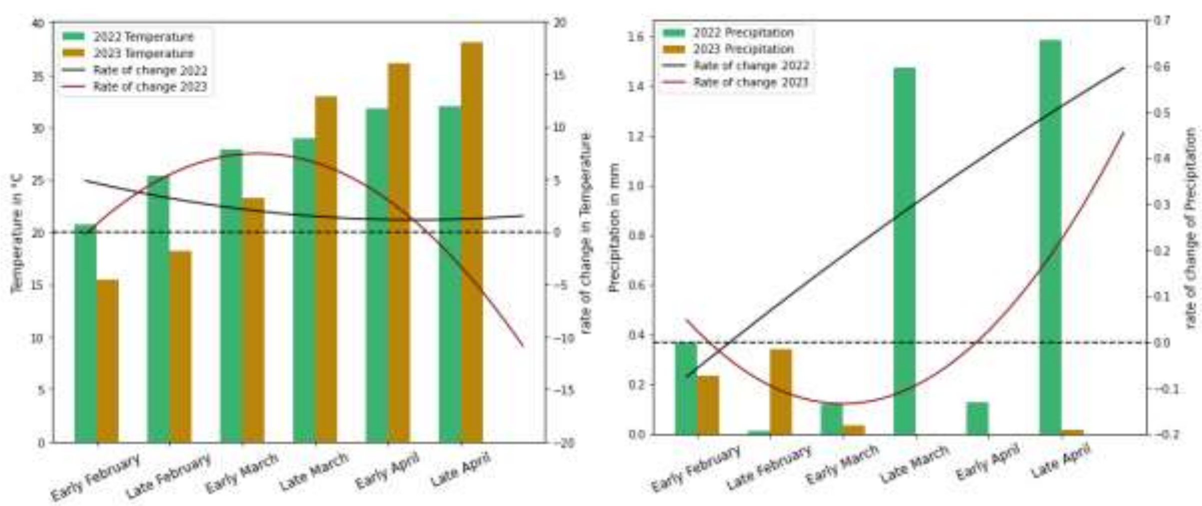


Figure: 8. Rate of change

5. Conclusion

The majority of statistical and mechanical phenology models have historically been created for tree species as opposed to non-woody species. Although phenological modelling has advanced significantly in recent years, the possibility of modelling vegetation plants with estimation of cross-scalar phenology is still underappreciated or just occasionally used. The methodology used in this paper demonstrates the high potential of employing machine learning models to close geographical and temporal gaps in ground-based observations as well as forecasting certain phenological phases using meteorologically data with NDVI.

Reference:-

[1] Ilyas, Qazi, et al. “Automated Estimation of Crop Yield Using Artificial Intelligence and Remote Sensing Technologies.” Bioengineering, vol. 10, no. 2, Jan. 2023, p. 125. Crossref, https://doi.org/10.3390/bioengineering10020125.

[2] Kukal, M.S. and Irmak, S. (2018) ‘Climate-driven crop yield and yield variability and climate change impacts on the U.S. great plains agricultural production’, Scientific Reports, 8(1). doi:10.1038/s41598-018-21848-2.

[3] Crane-Droesch, A. (2018) ‘Machine learning methods for crop yield prediction and climate change impact assessment in agriculture’, Environmental Research Letters, 13(11), p. 114003. doi:10.1088/1748-9326/aae159.

[4] Cedric, L.S. et al. (2022) ‘Crops yield prediction based on machine learning models: Case of West African countries’, Smart Agricultural Technology, 2, p. 100049. doi:10.1016/j.atech.2022.100049.

[5] Ji, Z. et al. (2021) ‘Prediction of crop yield using phenological information extracted from Remote Sensing Vegetation index’, Sensors, 21(4), p. 1406. doi:10.3390/s21041406.

[6] Liu, Y. et al. (2022) ‘Using phenocams to track crop phenology and explain the effects of different cropping systems on yield’, Agricultural Systems, 195, p. 103306. doi:10.1016/j.agsy.2021.103306.

[7] Sitokonstantinou, V. et al. (2023) ‘Fuzzy clustering for the within-season estimation of cotton phenology’, PLOS ONE, 18(3). doi:10.1371/journal.pone.0282364.

[8] Gao, F. and Zhang, X. (2021) 'Mapping crop phenology in near real-time using satellite remote sensing: Challenges and opportunities', Journal of Remote Sensing, 2021. doi:10.34133/2021/8379391.

[9] Belgiu, M. et al. (2021) 'Phenology-based sample generation for supervised crop type classification', International Journal of Applied Earth Observation and Geoinformation, 95, p. 102264. doi:10.1016/j.jag.2020.102264.

[10] Kim, S.-J., Bae, S.-J. and Jang, M.-W. (2022) 'Linear regression machine learning algorithms for estimating reference evapotranspiration using limited climate data', Sustainability, 14(18), p. 11674. doi:10.3390/su141811674.

[11] Dhanya K1 , B.Jayanthi2 (2023) 'Crop REcommendation system for identifying the better yielding crops using hybrid evolutionary algorithm SS-NMRA' . DOI:10.31838/ecb/2023.12.s1-B.446