Satellite Image Based Water Bodies Assessment and Drought Prediction Using Machine Learning

Sidharth Shirsath, Sagar Parhad, Harshal Mali and Nikhil Chavhan Computer Engineering, KCT Late G.N. Sapkal College of Engineering, Nashik

ABSTRACT:

Accurate drought forecasting plays a crucial role in ensuring sustainable water resource management and enhancing disaster readiness. This research focuses on harnessing satellite imagery and machine learning to improve the prediction of drought conditions. By tapping into the vast spatial and temporal insights provided by satellite platforms like Landsat-8, we aim to build predictive models capable of identifying drought onset early and assessing its severity. Critical parameters such as land surface temperature, vegetation vitality, and rainfall distribution are extracted from satellite datasets to serve as key input features. Through the application of advanced machine learning techniques, particularly the Random Forest algorithm, our study captures the complex, non-linear patterns embedded within the data, resulting in more reliable drought forecasts. The effectiveness of the proposed model is assessed using comprehensive validation methods, emphasizing key regression evaluation metrics. The results underscore the promise of blending remote sensing technologies with artificial intelligence to develop sophisticated drought prediction tools, ultimately supporting proactive mitigation and resilience strategies. This interdisciplinary methodology offers valuable contributions to the fields of environmental monitoring and disaster management.

Keywords: Drought forecasting, remote sensing, machine learning, early warning systems, drought severity analysis, environmental monitoring.

INTRODUCTION:

Employing machine learning for drought prediction, particularly through the analysis of satellite imagery from Landsat-8, offers a transformative approach to addressing water scarcity challenges. Landsat-8, equipped with state-of-the-art sensors, provides an extensive range of multispectral data, serving as a foundation for building accurate predictive models. The process begins with the collection of satellite images, accessed from a comprehensive archive that captures Earth's evolving surface conditions over time.

For this study, imagery from three distinct regions—Rajasthan, Nashik, and Cherrapunji—was selected for the month of October, offering a variety of environmental and climatic profiles for thorough drought analysis. To account for seasonal variations, additional data from Rajasthan and Nashik were gathered for February.

Following data acquisition, preprocessing steps are essential to ensure the quality and reliability of the imagery. Techniques such as radiometric calibration and atmospheric correction are applied to remove inconsistencies and standardize the datasets. After preprocessing, crucial features are extracted by calculating spectral indices like NDVI (Normalized Difference Vegetation Index), SAVI (Soil-Adjusted Vegetation Index), and NDII (Normalized Difference Infrared Index), each providing insights into vegetation vitality, soil moisture, and surface water content.

A temporal analysis compares the October and February images, particularly for Rajasthan and Nashik, revealing seasonal shifts and enhancing the understanding of drought progression. These insights are critical for training machine learning models.

The dataset is then divided into training and testing subsets to develop robust predictive algorithms. Using the Random Forest model, the system is trained on patterns from Rajasthan and Nashik to identify drought indicators. Model performance is evaluated using the testing set to ensure it can reliably predict drought across unseen data.

For validation, Landsat-8 images from Jalgaon, covering both October and February, are used as an independent test case. This step assesses the model's ability to generalize across different locations and seasons.

Driven by the increasing urgency of water scarcity, this research highlights the potential of advanced technologies like remote sensing and machine learning in promoting environmental resilience. The ultimate goal is to provide a practical, real-time drought prediction tool for stakeholders, supporting proactive water management and fostering a sustainable future.

Objectives:

1. To detect and map water bodies within the selected study regions using satellite imagery and remote sensing techniques.
2. To develop an advanced drought forecasting model by integrating multispectral satellite data with machine learning algorithms.
3. To assess and classify the severity levels of drought events based on key environmental indicators.
4. To monitor seasonal variations and trends in vegetation health, soil moisture, and surface water availability.
5. To validate the predictive model across diverse geographical regions to ensure robustness and reliability.
6. To provide actionable insights for early warning systems, supporting sustainable water resource management and disaster mitigation strategies.

GEOGRAPHIC INFORMATION SYSTEMS

Geographic Information Systems (GIS) play a vital role in enhancing drought prediction by integrating multiple layers of geographic information, including rainfall distribution, soil moisture content, vegetation condition, land cover, and terrain characteristics. By applying spatial analysis and modeling techniques, GIS enables researchers to pinpoint areas vulnerable to drought, evaluate the intensity of ongoing drought conditions, and forecast future drought events.

GIS technology also excels at simplifying complex environmental data through powerful visualization tools, helping decision-makers formulate effective strategies for water resource management, agricultural operations, and disaster readiness. By overlaying diverse datasets and utilizing geospatial modelling, GIS can highlight drought-prone zones before critical conditions arise. This proactive insight empowers authorities to take preventive actions such as promoting water-saving practices, introducing crop diversification programs, and encouraging the adoption of drought-tolerant crop varieties.

LITERATURE REVIEW:

The Earth faces numerous natural hazards, with droughts being particularly devastating due to their unpredictable nature and wide-reaching impacts on socio-economic systems. Droughts not only reduce water availability but also degrade water quality, intensifying their effects on both human communities and ecosystems. Traditional methods, such as deploying soil sensors for drought monitoring, are limited by their impracticality over large areas. As a result, there is increasing interest in using machine learning techniques, powered by satellite imagery, to predict drought occurrences and their severity more effectively.

Satellite datasets, notably from Landsat-8 and Sentinel-1 (including SMAP data), offer valuable temporal and spatial information critical for drought forecasting. By analyzing variables like soil moisture indices from Sentinel-1, machine learning models such as Random Forest can reliably predict drought events and assess their severity. Water quality also becomes a crucial concern during droughts, with Chlorophyll-a levels serving as key indicators of nutrient imbalances that often result in harmful algal blooms, particularly in warm freshwater environments.

Landsat surface reflectance data—collected by Landsat-5, Landsat-7, and Landsat-8 satellites since 1984 at 16-day intervals—provides a rich resource of over 1,000 images. However, not every image is usable due to interference from factors like cloud cover, snow, or ice. To refine the dataset for analysis using Google Earth Engine (GEE), several masking techniques are employed. Spatial masking isolates pixels within the specific region of interest, classification masking identifies and focuses on water pixels despite changes in waterbody size, and quality masking filters out compromised pixels affected by clouds or atmospheric distortions. Together, these steps ensure the use of high-quality, reliable satellite imagery for environmental analysis.

This document also synthesizes research led by the NOAA Drought Task Force (DTF), a collaboration involving NOAA's Climate Program Office and the National Integrated Drought Information System (NIDIS), along with scientists from academia and other agencies. Their research evaluates recent advancements and remaining challenges in drought monitoring and forecasting. Highlights include the successful application of advanced land-surface hydrology models for objective drought analyses and the improvement of near-real-time satellite monitoring tools for vegetation health and evapotranspiration tracking. Progress has also been made in seasonal drought prediction, particularly through the North American Multimodel Ensemble (NMME) suite. Furthermore, the influence of large-scale climate patterns like La Niña and the internal variability of the atmosphere on drought development, especially in regions like the southern Great Plains, is better understood. These findings mark significant steps toward improving drought resilience while identifying critical areas that still require further research.

DATASET :

Landsat-8, a joint mission by NASA and the U.S. Geological Survey, plays a crucial role in observing and documenting Earth's surface, including both landmasses and polar regions. It captures moderate-resolution imagery across a range of spectral wavelengths, allowing for the detailed monitoring of changes in land cover over time. This data is essential for various applications, such as urban planning, disaster management, agricultural monitoring, and water resource management. As part of the Landsat program, which has been operational for over four decades, Landsat-8 continues to contribute to the longest continuous record of Earth's surface conditions from space.

Equipped with two key instruments—the Operational Land Imager (OLI) and the Thermal Infrared Sensor (TIRS)—Landsat-8 collects seasonal imagery at multiple spatial resolutions. These sensors work together to help scientists accurately track long-term environmental changes. The satellite captures information across the visible, near-infrared, shortwave infrared, and thermal infrared portions of the electromagnetic spectrum, enabling a detailed understanding of Earth's surface dynamics.

Another major advantage of Landsat-8 is that its high-quality, calibrated imagery is freely accessible to the public, fostering research across fields like agriculture, forestry, land use planning, and studies of global environmental change. The combination of accessible data and advanced imaging capabilities makes Landsat-8 a vital tool for managing and protecting natural resources worldwide. The satellite's various spectral bands provide the flexibility needed for a wide range of scientific and practical applications. The bands on Landsat-8 include:

Band Number	Band Name	Wavelength (um)	Resolution (m)
1	Coastal/Aerosol	0.435-0.451	30
2	Blue	0.452-0.512	30
3	Green	0.533-0.590	30
4	Red	0.636-0.673	30
5	NIR	0.851-0.879	30
6	SWIR-1	1.566-1.651	30
7	SWIR-2	2.107-2.294	30
8	Pan	0.503-0.676	15
9	Cirrus	1.363-1.384	30
10	TIR-1	10.60-11.19	100
11	TIR-2	11.50-12.51	100

Figure 1: Bands

NDII (Normalized Difference Infrared Index)
The **Normalized Difference Infrared Index (NDII)** is a vegetation index calculated using satellite imagery, especially from **Landsat-8**, a collaborative mission between NASA and the United States Geological Survey (USGS).

NDII is based on the reflectance values captured by the **Near-Infrared (NIR)** and **Short-Wave Infrared (SWIR)** bands of Landsat-8. It plays a key role in assessing the **water content** of vegetation and soil. NDII values typically range between **-1 and 1**, helping researchers accurately evaluate vegetation health and soil moisture conditions.

This index is widely applied in fields such as **agriculture**, **forestry**, and **environmental monitoring**, supporting informed decision-making related to land and resource management.
The formula for NDII is:

$$\mathbf{NDII} = \frac{\mathbf{NIR - SWIR}}{\mathbf{NIR + SWIR}}$$

Key Variables and Related Indices

1. **NDVI (Normalized Difference Vegetation Index)** – Measures vegetation health by comparing near-infrared and red light reflectance.
2. **EVI (Enhanced Vegetation Index)** – An improved version of NDVI that corrects for atmospheric distortions and canopy background noise.
3. **SAVI (Soil Adjusted Vegetation Index)** – Modifies NDVI to minimize soil brightness effects, ideal for dry or sparse vegetation regions.
4. **NDMI (Normalized Difference Moisture Index)** – Assesses vegetation moisture levels by combining NIR and SWIR bands.
5. **MDWI (Modified Difference Water Index)** –Enhances the detection of surface water bodies by being highly sensitive to water content variations.
6. **MSAVI (Modified Soil Adjusted Vegetation Index)** – A refined version of SAVI that automatically adjusts for soil brightness without assuming a fixed soil line.
7. **SWIR1 and SWIR2 (Short-Wave Infrared Bands 1 and 2)** – Capture moisture and vegetation data; crucial for drought and soil moisture studies.
8. **TIRS1 and TIRS2 (Thermal Infrared Sensor Bands 1 and 2)** – Record thermal radiation, used to estimate land surface temperatures.
9. **NDWI (Normalized Difference Water Index)** – Detects water presence and changes in vegetation water content, commonly used for water body mapping.
10. **NDCI (Normalized Difference Chlorophyll Index)** – Sensitive to chlorophyll content, useful for monitoring plant health and detecting algal blooms.

STUDY AREA:

The regions of **Nashik** and **Jalgaon** in **Maharashtra**, as well as **Rajasthan**, offer distinct environmental challenges that can greatly benefit from the combination of satellite imagery and machine learning techniques.

In **Nashik** and **Jalgaon**, which are semi-arid regions, a potential project could focus on leveraging satellite data and machine learning models to predict drought events. By analyzing key environmental variables such as soil moisture, vegetation health, rainfall patterns, and temperature, machine learning algorithms can be trained to detect drought conditions based on both historical trends and real-time data, offering more accurate forecasts for these regions.

Rajasthan, known for its extreme arid conditions, presents another opportunity to apply similar methods for drought prediction. Satellite imagery can be used to track shifts in land surface temperature, soil moisture levels, and vegetation indices. With these data points, machine learning models can help identify early signs of drought, providing valuable warnings for farmers and local authorities to take preventive measures.

Despite its reputation for high rainfall, **Cherrapunji** faces issues related to vegetation degradation, soil erosion, and the risk of landslides. In this region, a project utilizing satellite imagery and machine learning could focus on monitoring vegetation health, tracking soil erosion patterns, and assessing landslide risk. By analyzing satellite data, machine learning models can predict areas that are most vulnerable to these environmental issues, enabling proactive land conservation and reducing the risk of disasters like landslides.

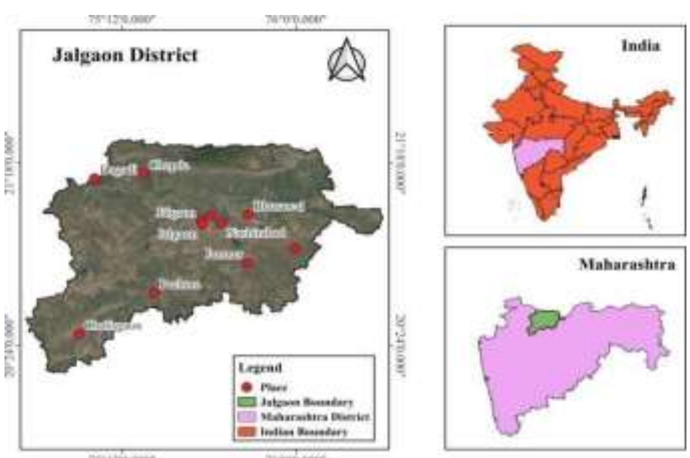

Figure 2: Jalgaon Region

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 06 | June - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

METHODOLOGY:

Random Forest Algorithm

The Random Forest Regressor is a powerful machine learning tool frequently used for predictive modeling tasks, such as predicting drought conditions using satellite imagery. Satellite data provides essential information on environmental factors like land surface temperature, vegetation health, soil moisture levels, and rainfall patterns, which serve as the input variables for this algorithm.

The algorithm works by creating a collection of decision trees. Each tree is trained on a random selection of both the data and the features, and when making predictions, it combines the results from all the trees to generate a final output. This ensemble approach increases the model's accuracy and its ability to generalize, making it particularly effective for complex, multi-dimensional data sets like those derived from satellite images.

In the context of drought prediction, the Random Forest Regressor can learn and identify intricate, non-linear relationships between various environmental factors captured by satellites and the severity of drought conditions. By examining past satellite data and corresponding drought indices, the model uncovers patterns that can predict when and where droughts are likely to occur and their potential intensity.

This algorithm offers multiple benefits for drought forecasting. It is capable of handling large, high-dimensional datasets, captures non-linear interactions between the input features and the target outcomes, and reduces the risk of overfitting by using a random selection of data and features for each tree. Because of these strengths, the Random Forest Regressor is a highly reliable and effective tool for predicting drought events and assessing their severity using satellite-based environmental data.

Linear Regression Algorithm

Integrating satellite imagery with machine learning, particularly through linear regression, offers substantial potential for improving drought forecasting. Satellite data provides essential information on environmental factors like vegetation health, soil moisture, and temperature, all crucial for understanding drought conditions. Using linear regression, these datasets can be effectively analyzed to discover patterns and correlations between different variables and past drought events.

Linear regression is a fundamental statistical approach in machine learning that fits a straight-line model to the data, making it possible to predict future outcomes based on current input features. In the case of drought prediction, linear regression models the relationship between satellite-derived environmental variables—such as NDVI (Normalized Difference Vegetation Index), NDWI (Normalized Difference Water Index), and NDII (Normalized Difference Infrared Index)—and drought severity indices. Through training on historical data, the algorithm identifies patterns that allow it to predict future drought conditions based on new satellite observations.

Despite its strengths, linear regression does have some limitations, especially when it comes to capturing non-linear relationships and complex interactions in the data—issues that are commonly encountered in drought dynamics. In situations where the data includes such complex relationships, more advanced machine learning techniques, like random forest regressors, may provide more accurate predictions. Random forests are particularly good at identifying intricate, non-linear patterns, often leading to superior performance in tasks like drought prediction when the relationships between variables are more complex.

In the Jalgaon region, linear regression has been applied to predict drought severity. The model takes satellite-derived NDVI, NDWI, and NDII values as input and generates predictions of drought conditions. The output is a map illustrating the severity of the predicted drought, which serves as a useful tool for identifying vulnerable areas. This information helps local farmers and policymakers make more informed decisions regarding water resource management and drought mitigation strategies.

Architecture of Drought Prediction Model :

The process starts with collecting satellite data from Landsat-8, which captures detailed images of the Earth's surface. This data is then used to compute various vegetation and moisture indices such as NDVI (Normalized Difference Vegetation Index), NDWI (Normalized Difference Water Index), and NDII (Normalized Difference Infrared Index), which provide crucial insights into the environmental conditions affecting drought.
Next, a machine learning model—specifically a random forest regressor—is used to train the prediction model. The training process involves data from various regions, such as Nashik, Cherrapunji, and Rajasthan for the months of October, and Nashik and Rajasthan in February. The training dataset includes the calculated index values (NDVI, NDWI, NDII) alongside ground truth data indicating actual drought conditions observed during these periods.

The random forest regressor algorithm learns the relationships between the environmental variables (indices) and the drought conditions by analyzing the training data. Through this learning process, the model identifies patterns that help predict drought occurrences. Once trained, the model is ready to make predictions about drought conditions in new, unseen areas.
After the model has been developed, it can be applied to predict drought conditions in other regions not included in the training set, offering valuable insights and helping local authorities and farmers prepare for potential droughts.

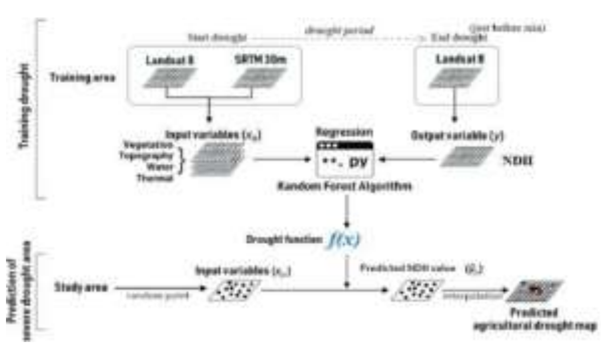

Figure 3: Architecture

Metrics for Evaluating Model Performance:

1. **RMSE (Root Mean Squared Error)**
   - Definition: RMSE measures the difference between the predicted and actual values, providing an average error for the model's predictions.
   - Formula:

$$RMSE = \sqrt{1/n \sum_{i-1}^{n} (y_i - y\hat{}_i)^2}$$

where:

- y_i is the actual value,
- $y\hat{}_i$ is the predicted value,
- n is the number of data points.

   - Interpretation: Lower RMSE values indicate that the model's predictions are closer to the actual values, showing a better fit.

2. **MSE (Mean Squared Error)**
   - Definition: MSE measures the average of the squared differences between actual and predicted values. It's a key metric used to assess the accuracy of regression models.
   - Formula:

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (y_i - y\hat{}_i)^2$$

where:

- y_i is the actual value,
- $y\hat{}_i$ is the predicted value,
- n is the number of data points.

   - Interpretation: Smaller MSE values indicate that the model's predictions are closer to the actual values, signaling better model performance.

3. **R² (R-squared)**
   - Definition: R² is a statistical measure that represents how well the regression model explains the variance in the response data. It is a ratio of the variance explained by the model to the total variance in the data.
   - Formula:

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - y\hat{}_i)^2}{\sum_{i=1}^{n} (y_i - y^{-}_i)^2}$$

where:

- y_i is the actual value,
- $y\hat{}_i$ is the predicted value,
- $\bar{y}$ is the mean of the actual values,
- n is the number of data points.

   - Interpretation: R² ranges from 0 to 1, where:
     - A value closer to 1 indicates that the model explains most of the variance in the data.
     - A value closer to 0 suggests that the model does not explain the variability in the data well.

IMPLEMENTATION:

Drought Prediction Model

Forecasting drought using satellite data and machine learning involves developing an advanced system that blends multiple data streams and intelligent algorithms to predict drought with precision. The first step in this process is collecting satellite images across various spectral bands — such as visible light, near-infrared, and thermal infrared — which capture essential details about land surfaces and vegetation health. These images serve as the core inputs for the predictive model.

Before the images can be used, they undergo preprocessing steps like removing cloud cover, adjusting for atmospheric effects, and normalizing values to create a consistent dataset. The model may also incorporate supplementary information like weather data (rainfall, temperature, humidity) and soil moisture readings to strengthen its predictive capabilities.

Machine learning algorithms, particularly the Random Forest method, are then employed to train the system using historical records of drought events. Through training, the model learns how different features — such as satellite reflectance values and weather conditions — are linked to drought outcomes. To ensure the model's reliability, its predictions are validated by comparing them to actual drought occurrences, using evaluation metrics like accuracy, precision, recall, and the F1-score.

If needed, the model is fine-tuned to improve its predictive strength and adaptability. After successful validation, the model can be deployed for real-world drought forecasting. It will then continuously process fresh satellite and weather data, delivering near real-time drought predictions. These timely forecasts help authorities, farmers, and planners make strategic decisions to minimize drought impacts on agriculture, water supplies, and the environment. Furthermore, integrating these predictions into decision-support systems ensures rapid, informed action in regions vulnerable to drought.

Training Dataset for Drought Prediction

Building a drought prediction model using satellite imagery and machine learning requires a rich and diverse set of training data. This dataset is assembled by pulling features from satellite observations, weather datasets, and historical drought records. In this study, particular attention is given to regions like Nashik, Rajasthan, and Cherrapunji, with data collected for both October and February to capture different seasonal behaviors and drought trends.

From satellite imagery, essential features such as NDVI (Normalized Difference Vegetation Index), land surface temperature, soil moisture levels, and precipitation data are extracted. In addition to these satellite-derived variables, meteorological parameters — like temperature, humidity, wind speed, and atmospheric pressure — are also incorporated to give a fuller picture of environmental conditions.

Historical drought data further strengthens the dataset, providing key insights into drought severity, duration, and frequency across these regions.

This information is crucial for training the model to recognize the distinct patterns that characterize drought events. The dataset is carefully labeled to distinguish between drought and non-drought situations, often using drought severity indices like the NDII (Normalized Difference Infrared Index) for categorization.

Once prepared, the dataset is used to train machine learning models, with the Random Forest algorithm commonly selected due to its effectiveness. After initial training, cross-validation techniques are applied to evaluate how well the model generalizes beyond the training data. To truly test its performance, the model is then validated on a new, unseen region — Jalgaon — ensuring its ability to accurately predict drought conditions outside of the original training areas.

RESULT:

Low-Intensity Drought Analysis

In research aimed at predicting drought using satellite imagery paired with machine learning techniques, scientists compiled datasets from three geographically distinct regions: Nashik, Rajasthan, and Cherrapunji. They utilized the Random Forest Regressor algorithm to model drought conditions specifically for the month of October.

Their analysis revealed interesting trends. Initial observations of the Normalized Difference Infrared Index (NDII) showed noticeable differences between the regions, highlighting their varied environmental states. Nashik recorded moderate NDII values, suggesting relatively healthy vegetation. Rajasthan, on the other hand, displayed much lower NDII readings, indicating dry and arid conditions, while Cherrapunji — true to its reputation — exhibited higher NDII values, representing its rich, dense greenery.

The Random Forest Regressor performed effectively, accurately predicting the subtle regional differences in NDII during October and closely matching the observed data. Nevertheless, the model struggled to fully capture abrupt drought occurrences or sharp changes in vegetation health.

This study emphasized how satellite imagery can play a crucial role in real-time environmental monitoring, offering early signals for necessary interventions in drought management. The flexibility and strength of the Random Forest Regressor make it a powerful tool for navigating the complexities of drought prediction, providing essential data to guide strategic decisions and optimize the use of water and agricultural resources.

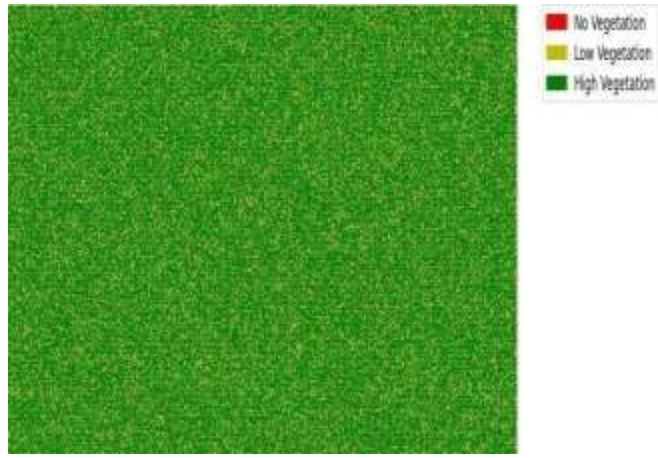

Figure 4: Original NDII

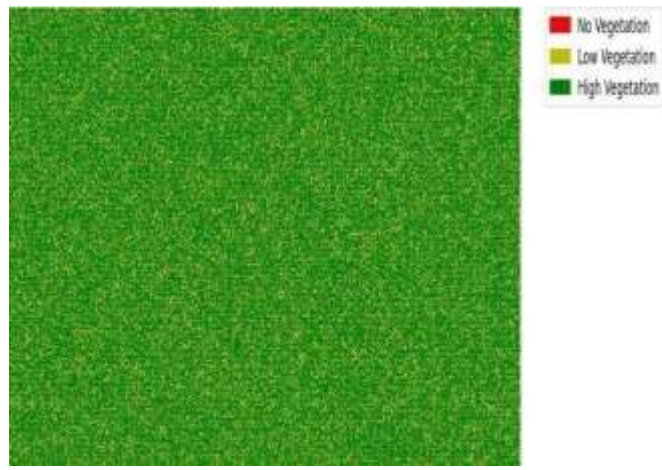

Figure 5: Predicted NDII

Severe Drought Prediction

In research targeting drought forecasting through satellite imagery and machine learning, datasets from **Nashik** and **Rajasthan** for the month of **February** were combined for analysis. Since February typically brings **little to no rainfall** in these regions, the focus shifted to detecting patterns within satellite images to **accurately predict drought stress**.

To build the model, researchers used the **Random Forest Regressor**, a powerful and reliable machine learning algorithm. After extensive training on the collected data, the model generated both **actual** and **predicted** values for the **Normalized Difference Infrared Index (NDII)** — a vital metric that reflects soil moisture and vegetation health, making it essential for drought assessment.

The evaluation showed that the model effectively learned the intricate relationships between satellite-derived features and drought indicators. A comparison between the real and predicted NDII values provided valuable insights, highlighting the model's strengths in capturing environmental conditions and pinpointing areas where predictions could be improved.

This analysis not only confirmed the model's potential for operational drought forecasting but also suggested pathways for enhancing its accuracy and reliability through further refinements.

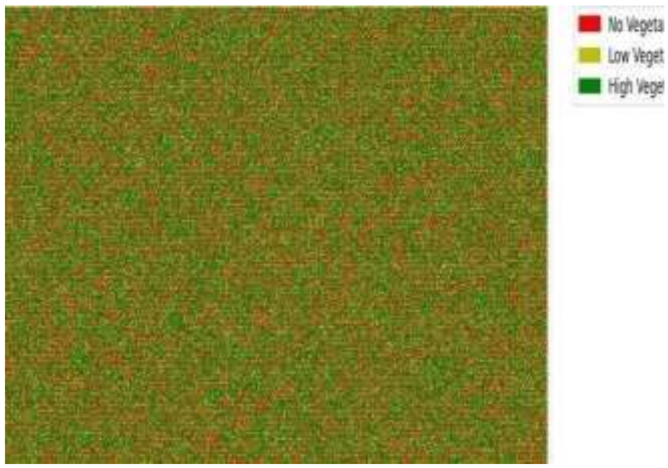

Figure 6: Original NDII

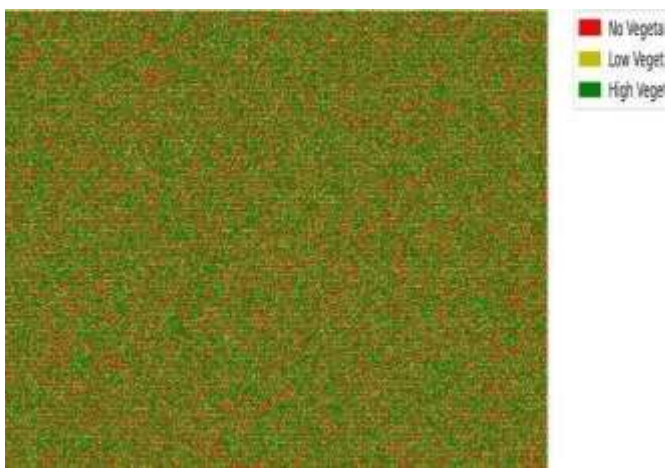

Figure 7: Predicted NDII

Drought Prediction Results

The outcomes of the drought prediction using satellite imagery and machine learning reveal a striking difference between the landscapes observed in October and February. In October, after seasonal rainfall, the satellite images display noticeable vegetation coverage in many areas, suggesting relatively moderate environmental conditions. In contrast, the February images depict a much harsher environment, marked by a significant loss of vegetation.This is particularly evident when analyzing the NDII (Normalized Difference Infrared Index) values:

- NDII values close to 1 represent healthy, green vegetation,
- Values near 0 indicate moderate vegetation health,
- Values approaching -1 signal severe drought.

In February, most NDII values clustered near -1, clearly illustrating the effects of severe drought. This sharp seasonal contrast highlights how variations in rainfall directly influence vegetation health, demonstrating the importance of combining satellite data and machine learning models for effective drought monitoring and early warning.

Moreover, the drought prediction model achieved an impressive accuracy of 99.05%, confirming that the system is both highly reliable and efficient at forecasting drought conditions.

- R2 =99.05
- Root Mean Squared Error=0.100
- Mean Squared Error=0.010

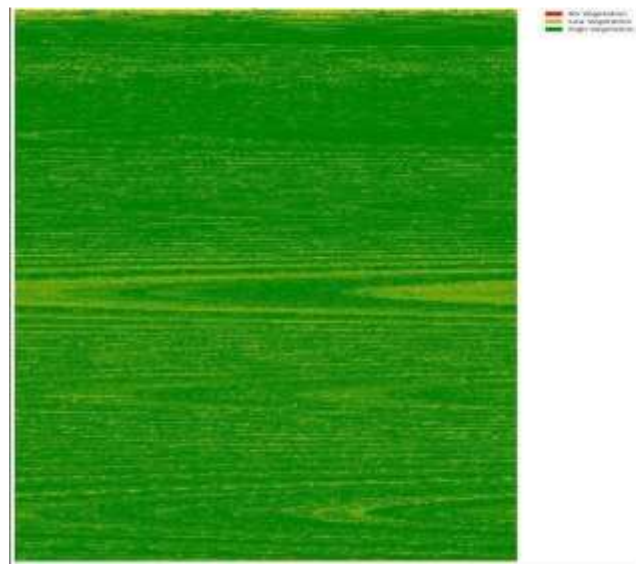

Figure 8: Oct Month

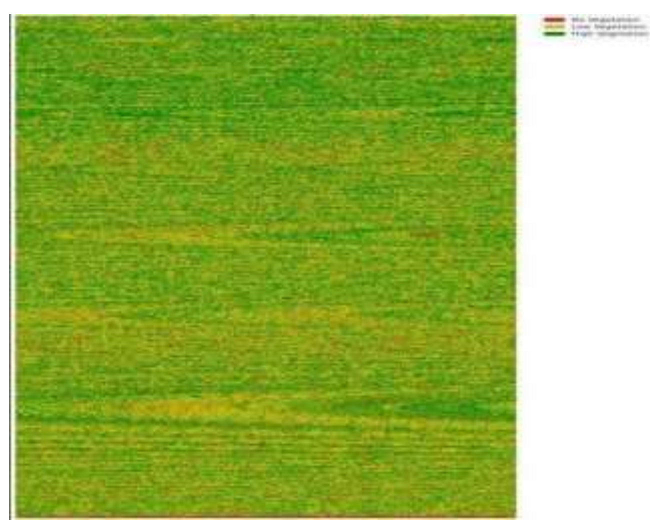

Figure 9: Feb Month

CONCLUSIONS:

This study used machine learning to predict drought conditions from Landsat-8 satellite imagery. We created a detailed dataset by calculating key indices, such as the Normalized Difference Infrared Index (NDII), and carefully preprocessed the data for accuracy. Feature extraction was employed to highlight the most relevant information for drought prediction.

The machine learning model, trained for accuracy and efficiency, produced visual drought prediction maps, helping to identify at-risk areas. These maps provided valuable insights for resource management and mitigation efforts.

We also compared seasonal changes between October and February, revealing that areas experiencing moderate drought in October became more severely affected by February. This highlights the need for timely interventions and adaptive management strategies.

In conclusion, our study demonstrates the effectiveness of combining machine learning with satellite imagery to forecast droughts, offering actionable insights for managing water resources and adapting to evolving environmental conditions.

REFERENCES:

1. Haekyung Park, Kyungmin Kim, and Dong Kun Lee explored the use of Random Forest combined with satellite imagery and topographical data for severe drought area prediction, published in the *MDPI journal Water* in 2019.

2. Getachew Berhan Demisse, Shawndra Hill, Tsegaye Tadesse, and Solomon Atnafu designed a *Drought Prediction System* aimed at improving climate change mitigation efforts, shared through *ResearchGate* in July 2014.

3. Zengchao Hao, Vijay P. Singh, and Youlong Xia discussed the progress, challenges, and future opportunities of *Seasonal Drought Prediction* in an *AGU* publication from January 2018.

4. Zengchao Hao, along with Amir AghaKouchak, Alireza Farahmand, and Navid Nakhjiri, introduced a *Global Integrated Drought Monitoring and Prediction System* in *Scientific Data* in March 2014.

5. Eric F. Wood, Siegfried D. Schubert, and Andrew W. Wood contributed to the *American Meteorological Society* in August 2015 with their insights into improving drought understanding, monitoring, and forecasting.

6. Sachidananda Mishra and Deepak R. Mishra proposed the *Normalized Difference Chlorophyll Index (NDCI)* as a new model for remotely estimating chlorophyll-a concentration in productive but turbid waters, published in *Science Direct*, Volume 117, on February 15, 2012.

7. Luke M. Moseley presented an integrated review of how drought impacts the water quality of freshwater systems, found in *Science Direct*, Volume 140, January 2015, pages 203–214.

8. Amir AghaKouchak developed a baseline probabilistic drought forecasting model based on the Standardized Soil Moisture Index, specifically applying it to the 2012 U.S. drought, published in *Hydrology and Earth System Sciences* in 2014.

9. Zengchao Hao, Fusheng Hao, Vijay P. Singh, Wei Ouyang, and Hecheng Cheng proposed an integrated system for drought monitoring, prediction, and analysis aimed at aiding modeling and assessment, published in *Environmental Modelling & Software* in 2017.

10. S. Park, J. Im, E. Jang, and J. Rhee introduced a machine learning-based approach for drought assessment and monitoring by blending multi-sensor indices across different climate regions, featured in *Agricultural and Forest Meteorology*, Volume 217, in 2016.