

Drought Monitoring Method Using Remote Sensing Data and Artificial Intelligence

TRAN Thi Thu Trang*, DO Thi Phuong Thao, LE Thi Thu Ha, PHAM Thi Lan

Hanoi University of Mining and Geology, Hanoi, Vietnam

* Corresponding author: tranthithutrang@humg.edu.vn

Abstract: Drought is a complex natural disaster with cyclical occurrence, slow onset, and wide-ranging, long-lasting impacts that cause severe damage to socio-economic systems, the environment, and ecosystems. Effective monitoring and forecasting are crucial for early detection, severity assessment, and drought risk management. This study presents a novel approach to drought monitoring by integrating multi-source remote sensing data with artificial intelligence (AI) techniques. Landsat and Sentinel-2 imagery were used to derive ten widely applied drought indices (NDVI, SAVI, VCI, NDWI, LSWI, MSI, NDDI, TCI, TVDI, VHI). Machine learning algorithms, including Random Forest (RF), Support Vector Machine (SVM), and Gradient Tree Boosting (GTB), were employed to classify and predict drought risk based on a training dataset constructed from the TVDI index. The entire workflow was implemented on the Google Earth Engine (GEE) platform, enabling large-scale data processing and automation. The proposed method enhances accuracy, efficiency, and automation in drought monitoring, thereby supporting early warning systems and sustainable water resource management under the increasingly complex challenges of climate change.

Keywords: Drought monitoring; Drought indices; Remote sensing; Machine learning; artificial intelligence

1. Introduction

Drought is an abnormal and prolonged shortage of rainfall, leading to severe water scarcity for domestic use, production, and natural ecosystems. It can arise from multiple combined factors such as reduced precipitation, temperature fluctuations, and especially the impacts of global climate change. In scientific research and practical applications, drought is often classified into four main types: meteorological, agricultural, hydrological, and socio-economic drought. This classification plays an important role in monitoring, impact assessment, and proposing appropriate response measures. Drought monitoring is the process of continuously collecting, updating, analyzing, and evaluating data to track the onset, evolution, spatial extent, intensity, and impacts of drought in real time or near real time. According to Heim (2002), drought monitoring has three main objectives: early detection of drought development, identification of its severity and extent, and support for early warning and the development of drought risk management strategies. Nowadays, drought monitoring is increasingly modernized thanks to the rapid advancement of spatial data acquisition technologies such as GIS, remote sensing, big data analytics, and artificial intelligence (AI).

Drought is among the most pervasive natural hazards, exerting profound and long-lasting impacts on economies, ecosystems, and societies worldwide. Unlike other rapid-onset disasters, drought develops slowly, persists over extended periods, and often lacks clear early warning signals. The UNCCD Global Drought Assessment (2020) reported that between 1998 and 2017, global droughts caused more than USD 124 billion in economic losses and directly affected hundreds of millions of people. Recent studies further indicate that the frequency of severe droughts has doubled since the pre-industrial era, with Sub-Saharan Africa, the Mediterranean, and the western United States identified as global hotspots (Nicholson, 2014; Funk et al., 2015; Spinoni et al., 2015; Williams et al., 2020) (Fig.1a).

In Vietnam, drought increasingly threatens water security and sustainable socio-economic development. National climate assessments reveal declining dry-season rainfall and rising average temperatures, particularly in drought-prone regions such as the South Central Coast, Central Highlands, Mekong Delta, and North Central Coast (MONRE, 2020) (Fig.1b). Moreover, socio-economic pressures, including rapid urbanization, population growth, agricultural expansion, groundwater overexploitation, and forest degradation further intensify water resource vulnerabilities. These trends highlight the urgent need for advanced drought monitoring and management strategies tailored to local conditions under a changing

climate.



Fig.1a. Agricultural drought in India



Fig.1b. Agricultural drought in Ninh Thuan in the dry season of 2024

Conventional drought monitoring methods largely rely on meteorological indices, which only capture precipitation deficits and provide limited spatial coverage due to dependence on sparse ground stations. These methods fail to directly measure critical variables such as soil moisture, vegetation health, or groundwater storage. By contrast, remote sensing has emerged as a powerful alternative, offering objective, large-scale, and long-term observations. It enables simultaneous monitoring of key factors including soil moisture, vegetation dynamics, land surface temperature, and water storage, thereby significantly enhancing drought research, assessment, and monitoring (Fig.2).

Recent advances in big data analytics and artificial intelligence (AI) have significantly transformed drought monitoring. AI’s ability to handle large datasets, capture nonlinear relationships, and predict complex patterns makes it a powerful tool for extracting and analyzing drought information from multi-source remote sensing data. The integration of remote sensing and AI is increasingly regarded as a key direction in modern drought research, particularly in regions with sparse ground-based observation networks such as Vietnam. The aim of this study is to develop a comprehensive framework for drought monitoring by integrating multi-source remote sensing data with AI-based machine learning techniques. The main contributions of the study include: (i) systematic selection and analysis of remote sensing datasets for drought monitoring; (ii) derivation and evaluation of widely used drought indices; (iii) application of advanced machine learning algorithms for classification and prediction; and (iv) implementation of an automated, cloud-based workflow to enhance scalability and operational efficiency. This approach is expected to improve the accuracy, timeliness, and practicality of drought monitoring and early warning under changing climate conditions. This framework is further detailed in the following sections, which describe the study area, datasets, methodological design, and implementation on a cloud-based platform.

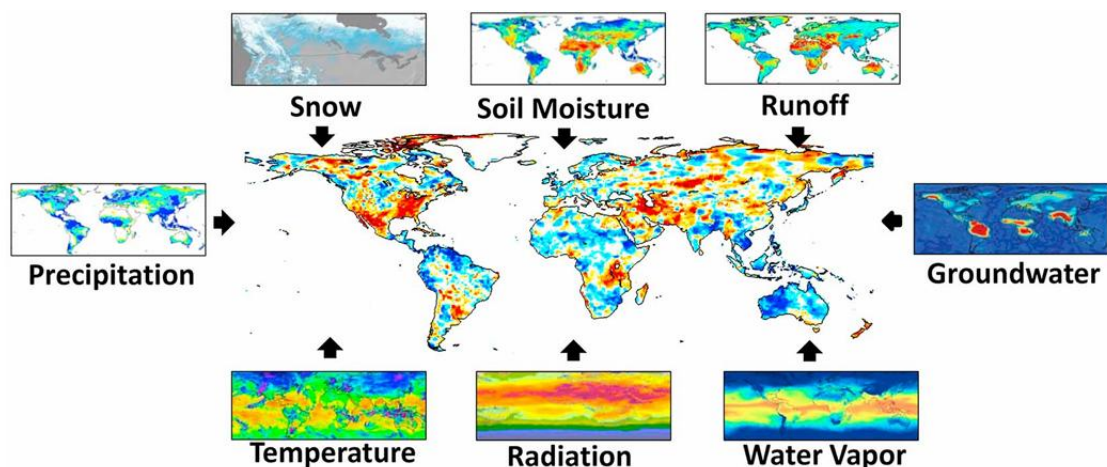


Fig. 2. Multi-indicator integrated drought monitoring using remote sensing data (AghaKouchak et al., 2015)

2. Remote sensing data in drought monitoring

2.1 Analysis of characteristics of remote sensing data sources

In drought forecasting and monitoring, remote sensing data serve as a vital source of information due to their ability to provide large-scale, continuous, and objective environmental observations over time. Commonly used satellite systems include MODIS, Landsat, Sentinel-2, and VIIRS. These sensors offer complementary spatial, temporal, and spectral characteristics, which enable the derivation of diverse drought-related indices such as vegetation condition, soil moisture, and land surface temperature. In this study, multi-source datasets from these satellites were selected and integrated to enhance the reliability and accuracy of drought monitoring.

2.1.1. MODIS (*Moderate Resolution Imaging Spectroradiometer*)

MODIS is one of the most widely used Earth observation remote sensing systems in climate and environmental research (Jaber et al., 2020; Weng, 2009). It is mounted on two satellites, Terra (launched in 1999) and Aqua (launched in 2002) which operate in coordination to provide near-global coverage with a revisit frequency of 1–2 days. The MODIS sensor contains 36 spectral bands ranging from the visible to the thermal infrared region, enabling diverse applications such as land surface temperature estimation, vegetation monitoring, wildfire detection, water turbidity assessment, and land cover classification (NASA, 2025).

A key advantage of MODIS is its wide coverage and high temporal resolution, which are particularly useful for building long-term datasets to analyze climate and large-scale drought trends. MODIS data are continuous, freely available, pre-processed, and easily accessible through platforms such as Google Earth Engine (GEE) and NASA Earthdata, making it suitable for regional, national, or transboundary studies.

However, MODIS is limited by its relatively coarse spatial resolution approximately 250 m for red and near-infrared bands, and 500–1000 m for other spectral bands. This restricts its applicability for detailed analyses at provincial or field scales. Moreover, the large pixel size increases the likelihood of mixed-pixel effects, reducing the accuracy of drought index calculations in fragmented landscapes or heterogeneous land-use regions.

2.1.2 Landsat

Landsat is one of the longest-running and most comprehensive Earth observation programs, jointly operated by the U.S. Geological Survey (USGS) and NASA since 1972 (NASA, 2025). To date, nine generations of Landsat satellites have been launched, providing valuable data for a wide range of scientific studies (Ghaleb et al., 2015; Khosravi et al., 2017). The most recent missions, Landsat 8 (launched in 2013) and Landsat 9 (launched in 2021), are equipped with the Operational Land Imager (OLI) and Thermal Infrared Sensor (TIRS), enabling the simultaneous acquisition of optical and thermal data (USGS, 2025) (Fig.3).

Landsat provides imagery at a spatial resolution of 30 m for optical bands and 100 m for thermal bands, with a 16-day revisit cycle. This makes it suitable for monitoring crop growth cycles and environmental changes at provincial or regional scales. Landsat data are of high quality, atmospherically and geometrically corrected, and freely available through EarthExplorer or Google Earth Engine.

One of the key strengths of Landsat is its long, consistent, and well-calibrated data archive. With records dating back to 1984, Landsat enables long-term drought trend analysis, identification of dry–wet cycles, and assessment of climate change impacts on land and water resources. Thermal data from TIRS are particularly useful for calculating integrated drought indices such as the Temperature Vegetation Drought Index (TVDI), which models the relationship between land surface temperature and vegetation growth.

Furthermore, Landsat's broad spectral coverage facilitates the derivation of multiple drought indices, including NDVI, LSWI, MSI, NDDI, and especially TVDI, which effectively reflects drought risk through the NDVI–LST relationship. However, Landsat also has limitations, particularly its relatively long 16-day revisit cycle, which is often disrupted by cloud cover during rainy seasons or in tropical regions. Despite these constraints, Landsat remains an optimal choice for medium- to long-term drought monitoring, especially when combined with higher-temporal-resolution datasets such as Sentinel-2.

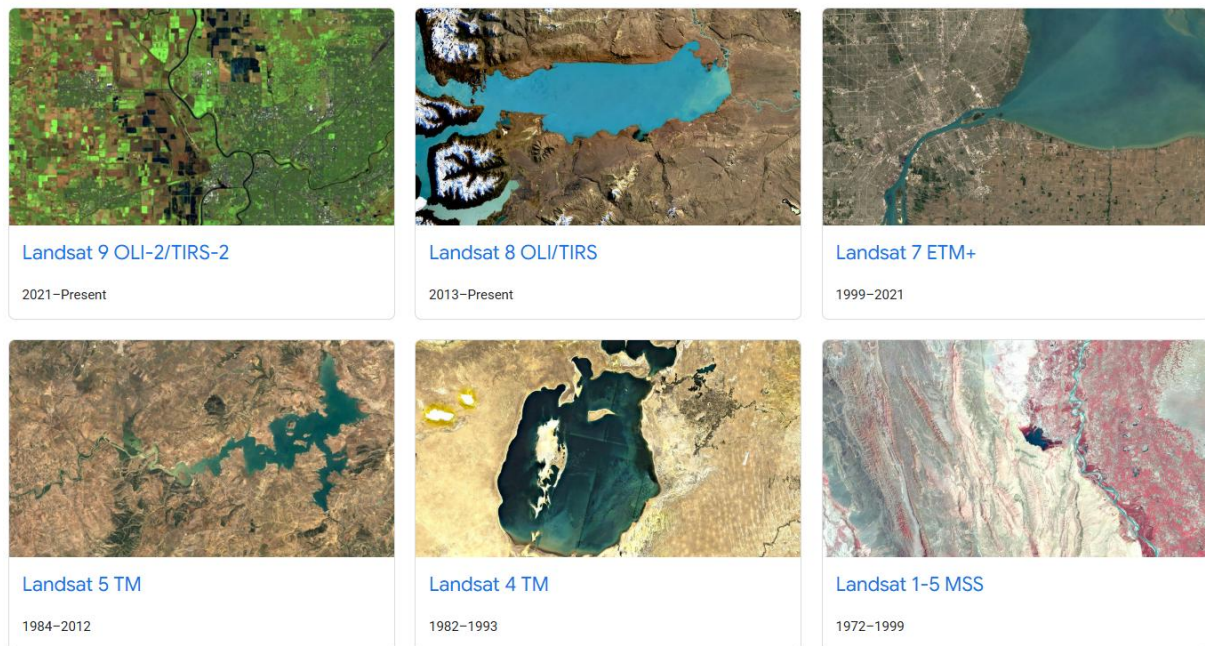


Fig. 3. Landsat datasets on GEE

2.1.3 Sentinel-2

Sentinel-2 is an Earth observation satellite system under the Copernicus Programme, implemented by the European Space Agency (ESA). It consists of two satellites Sentinel-2A (launched in 2015) and Sentinel-2B (launched in 2017) operating in tandem (ESA, 2025). Both satellites carry the MultiSpectral Instrument (MSI), specifically designed for high-resolution monitoring of vegetation, surface water, and terrestrial features with a short revisit cycle (ESA, 2025) (Fig.4).

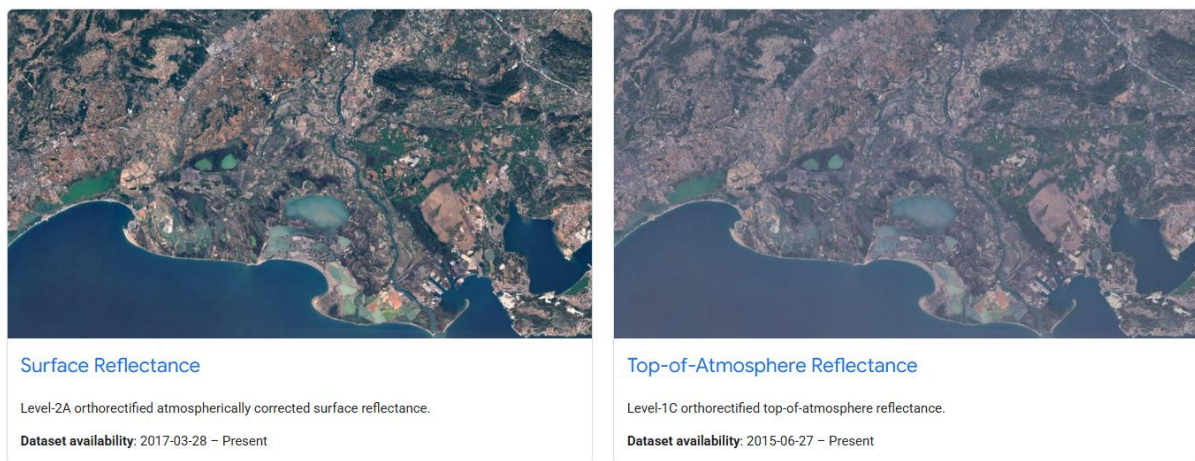


Fig. 4. Sentinel-2 datasets on GEE

Sentinel-2 provides imagery across 13 spectral bands. Among them, the 10 m bands (Blue, Green, Red, NIR) are suitable for vegetation analysis; the 20 m bands (Red Edge, SWIR) support assessments of physiological stress and soil moisture; and the 60 m bands are used for atmospheric correction. With a revisit cycle of only 5 days at the equator (enabled by the tandem operation of the two satellites), Sentinel-2 enables effective monitoring of vegetation dynamics and drought conditions (Varghese et al., 2021; West et al., 2018). Sentinel-2 combines high spatial resolution, short revisit frequency, and diverse spectral coverage tailored for ecological and agricultural studies. Its imagery is freely available at high quality, and notably, the three Red Edge bands (Bands 5, 6, 7) are highly sensitive to changes in chlorophyll content and leaf structure, enabling the early detection of drought-induced vegetation stress (West et al., 2018). Sentinel-2 data are systematically organized and available at multiple pre-processing levels (Level-1C: TOA reflectance; Level-2A: surface reflectance). They can be accessed via the Copernicus Open Access Hub, USGS, or directly through Google Earth Engine (GEE). GEE integration allows large-scale processing

of thousands of Sentinel-2 images for index calculation, time-series extraction, and rapid drought mapping. Although Sentinel-2 lacks thermal bands, unlike Landsat, its high temporal resolution and rich spectral detail make it an ideal dataset for drought trend analysis, machine learning model development, and early drought detection in fine-scale agricultural landscapes.

2.4 VIIRS (Visible Infrared Imaging Radiometer Suite)

The VIIRS is a new-generation sensor developed as the successor to MODIS, deployed on the Suomi NPP satellite (launched in 2011) and NOAA-20 (launched in 2017) (NASA, 2025). VIIRS comprises 22 spectral bands ranging from visible to thermal infrared, enabling measurements of land surface temperature, wildfire detection, nighttime light observation, and climate-related variables. It provides imagery at two spatial resolutions: 375 m for imagery bands (I-bands) and 750 m for moderate bands (M-bands), which is higher than MODIS but coarser compared to Sentinel-2 or Landsat (Fig.5). VIIRS offers reliable daily data acquisition under low cloud conditions, which is valuable for detecting short-term variations in temperature, drought extent, and wildfire risk areas. Additionally, VIIRS generates high-level data products such as Land Surface Temperature (LST), NDVI, EVI, and snow cover with high accuracy. However, VIIRS data have not yet been fully integrated into cloud-based platforms like Google Earth Engine, which presents certain limitations for accessibility and automated data processing. In drought research, VIIRS is often used as a complementary dataset alongside other remote sensing sources for large-scale monitoring and multi-source analysis (Benedict et al., 2021; Kogan et al., 2015). Its capabilities in nighttime imaging, thermal-sensitive spectral bands, and consistent daily coverage make VIIRS a promising dataset for integration with machine learning or deep learning models to rapidly identify drought-affected areas.

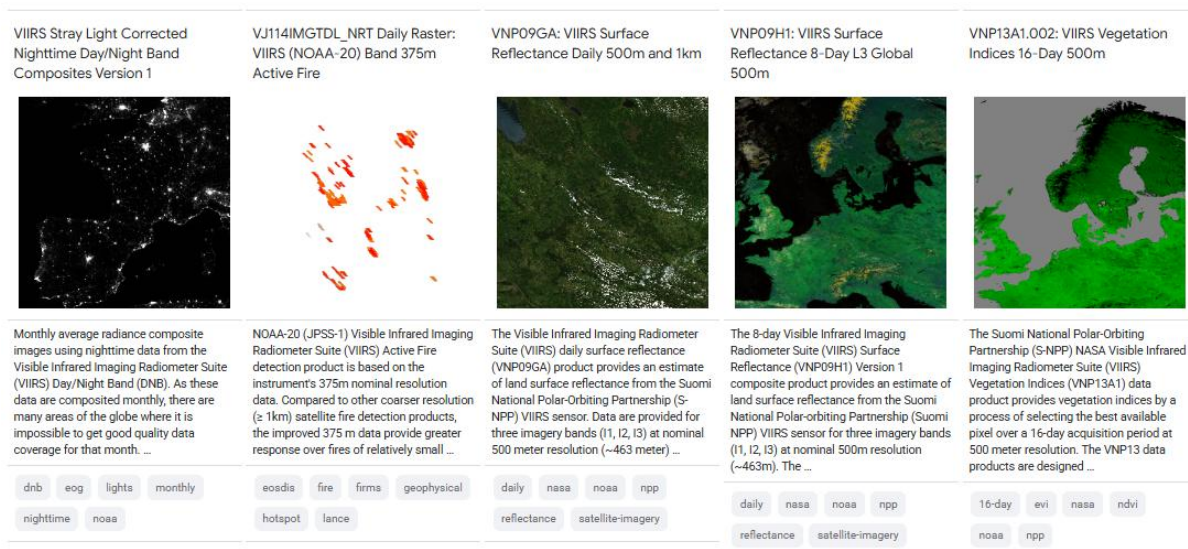


Fig. 5. Some VIIRS datasets on GEE

2.1.5 Select a remote sensing data source

Several satellite-based remote sensing systems are widely used in drought monitoring, including MODIS, Landsat, Sentinel-2, and VIIRS. Each sensor offers unique advantages and limitations in terms of spatial resolution, temporal frequency, spectral characteristics, and data accessibility. Table 1 is the comparison results of major satellite data sources for drought monitoring.

The comparison highlights that MODIS and VIIRS are advantageous for large-scale and daily monitoring, but their coarse resolution constrains detailed regional analyses. In contrast, Landsat and Sentinel-2 provide higher spatial resolution and richer spectral information, making them more suitable for provincial- or field-scale drought monitoring. Landsat’s long-term archive allows the analysis of multi-decadal drought and climate variability, while Sentinel-2 provides high-frequency observations with Red Edge bands that enhance early detection of vegetation stress.

Tab. 1. Comparison of major satellite data sources for drought monitoring

Sensor	Spatial resolution	Temporal resolution	Key advantages	Limitations
MODIS	250 m (red/NIR), 500–1000 m (other bands)	1–2 days	Broad coverage, long time series, free & preprocessed	Coarse resolution, unsuitable for local/regional detail
Landsat	30 m (optical), 100 m (thermal)	16 days	Long-term stable archive (since 1984), optical + thermal data, well-calibrated	Low revisit frequency, cloud contamination in tropics
Sentinel-2	10–20 m (optical/Red Edge/SWIR), 60 m (atmospheric bands)	5 days (with twin satellites)	High spatial resolution, short revisit, Red Edge bands sensitive to vegetation stress, free access	No thermal bands
VIIRS	375 m (I-bands), 750 m (M-bands)	Daily	Nighttime light, LST, fire detection, higher resolution than MODIS	Not fully integrated into cloud platforms, medium resolution

Moreover, the integration of multitemporal remote sensing imagery offers a significant advantage for constructing continuous time series, which are essential for analyzing both seasonal and interannual drought dynamics. Such datasets further support the establishment of near-real-time drought monitoring systems, thereby strengthening the effectiveness of early warning and risk management. For these reasons, this study selects Landsat and Sentinel-2 as the primary data sources for calculating drought indices and developing machine learning models.

2.2 Determining drought indices from multi-source remote sensing data

In drought monitoring and forecasting studies, identifying indices that reflect soil moisture, vegetation status, and thermal conditions is an essential step. Drought indices are calculated from satellite-derived spectral reflectance data, representing the physical and biological characteristics of the Earth’s surface. In this study, ten commonly used remote sensing-based drought indices were employed, categorized into three main groups: vegetation-related indices (NDVI, SAVI, VCI); soil moisture and surface water-related indices (NDWI, LSWI, MSI, NDDI); and composite indices (TCI, TVDI, VHI).

2.2.1 NDVI

The Normalized Difference Vegetation Index (NDVI) is one of the most widely used and important indices for vegetation monitoring and has been extensively applied in drought research (Peters et al., 2002). NDVI reflects vegetation greenness by measuring the difference in reflectance between the near-infrared (NIR) and red bands of the electromagnetic spectrum. Healthy vegetation strongly absorbs red light for photosynthesis and reflects NIR, whereas stressed vegetation (due to drought, pests, or land degradation) tends to reflect more red light and less NIR (Kamble et al., 2010).

The NDVI is defined as follows (Tucker, 1979):

$$NDVI = \frac{NIR-Red}{NIR+Red} \tag{1}$$

Where:

NIR: near-infrared reflectance

Red: red band reflectance

Under drought conditions, NDVI values tend to decrease significantly due to crop water loss, leaf drying, and reduced photosynthetic activity. By comparing NDVI across different time periods or against long-term averages, researchers can identify drought-affected areas and assess the severity of drought impacts.

2.2.2 SAVI

The Soil Adjusted Vegetation Index (SAVI) was developed to overcome some of the limitations of NDVI, particularly the influence of bare soil in areas with sparse or unevenly distributed vegetation (Aşık, 2020). In dryland, sandy, or semi-arid regions—areas that are often heavily affected by drought—NDVI values are easily distorted by strong soil background reflectance, leading to an inaccurate assessment of

vegetation greenness. To address this issue, SAVI incorporates a correction factor (L) to minimize soil background effects.

The SAVI is calculated as follows (Huete, 1988):

$$SAVI = \frac{(1+L)(NIR-Red)}{NIR+Red+L}, L = 0.5 \quad (2)$$

Where:

NIR: near-infrared reflectance

Red: red band reflectance

L: soil adjustment factor, commonly set to 0.5 for areas with intermediate vegetation density (Huete, 1988).

SAVI values have a similar interpretation to NDVI; however, the index provides more accurate information in areas with low vegetation cover (NDVI < 0.3). In regions with high vegetation density (NDVI > 0.5), SAVI and NDVI values are generally comparable. SAVI is particularly useful in drought studies conducted in agricultural areas, dry mountainous regions, or degraded lands, where the influence of barren soil may distort purely spectral indices such as NDVI.

2.2.3 VCI (Vegetation Condition Index)

The VCI is a time-series-normalized index that reflects the current health status of vegetation relative to the historically observed minimum and maximum conditions at the same location. Developed based on NDVI, VCI is capable of detecting abnormal vegetation responses caused by drought.

The VCI is calculated as follows:

$$VCI = \frac{(NDVI-NDVImin)}{NDVImax-NDVImin} \times 100 \quad (3)$$

2.2.4 NDWI (Normalized Difference Water Index)

The NDWI reflects surface moisture and the presence of water within vegetation canopies. It was developed by Gao (1996) to enhance the detection of surface water and vegetation moisture by exploiting the difference in reflectance between the green (Green) and near-infrared (NIR) spectral bands. Physically, green light is strongly absorbed by water, while NIR light is strongly reflected by vegetation. Therefore, the contrast between these two bands makes NDWI highly sensitive to changes in leaf water content and soil moisture.

The NDWI is calculated as follows (Gao, 1996):

$$NDWI = \frac{Green-NIR}{Green+NIR} \quad (4)$$

NDWI typically decreases when vegetation becomes dry due to drought, leaf water loss, or reduced soil moisture. Therefore, this index serves as a sensitive indicator for detecting water stress during dry seasons or in semi-arid regions. Moreover, NDWI is a key component in composite drought indices such as the Normalized Difference Drought Index (NDDI).

2.2.5 LSWI (Land Surface Water Index)

The Land Surface Water Index (LSWI) is widely used to assess surface moisture and water content in vegetation, making it particularly useful for monitoring agricultural drought and soil dryness in cultivated areas. LSWI is calculated based on the contrast between reflectance in the near-infrared (NIR) and shortwave infrared (SWIR) bands. It is highly sensitive to variations in water content within crops and soil, and often reflects drought conditions more clearly before NDVI exhibits a significant decline.

The formula for calculating LSWI is as follows (Xiang et al., 2020):

$$LSWI = \frac{NIR-SWIR}{NIR+SWIR} \quad (5)$$

LSWI is often used to identify transitional stages between crop seasons, distinguish between flooded and dry areas, or detect regions experiencing severe water scarcity during prolonged dry seasons.

2.2.6 MSI (Moisture Stress Index)

The Moisture Stress Index (MSI) is a remote sensing spectral index specifically designed to assess vegetation moisture stress, a phenomenon commonly observed during the early or middle stages of drought. This index reflects changes in leaf water content through the relationship between shortwave infrared (SWIR) and near-infrared (NIR) reflectance. When plants experience water stress, the reduction in leaf cell water content leads to an increase in SWIR reflectance and a decrease in NIR reflectance, resulting in a

higher MSI value.

The formula for calculating MSI is as follows:

$$MSI = \frac{SWIR}{NIR} \quad (6)$$

Where:

NIR: near-infrared reflectance.

SWIR: shortwave infrared reflectance.

The MSI value usually ranges from 0.2 to above 2, depending on vegetation type and climatic conditions. Unlike NDVI or NDWI, which primarily reflect overall vegetation growth or the presence of water, MSI focuses on the intrinsic physiological characteristics of plants. Therefore, MSI is particularly useful for the early detection of drought-induced stress, even when plants still appear green in terms of morphology.

2.2.7 NDDI (Normalized Drought Difference Index)

NDDI is a composite index developed to enhance drought detection by combining two widely used indices: NDVI (Normalized Difference Vegetation Index) and NDWI (Normalized Difference Water Index). This index leverages the correlation between vegetation growth status (NDVI) and water content (NDWI) to more accurately reflect drought severity in the study area.

The formula for NDDI is defined as follows (Gu et al., 2007):

$$NDDI = \frac{NDVI - NDWI}{NDVI + NDWI} \quad (7)$$

Where:

NDVI: Normalized Difference Vegetation Index

NDWI: Normalized Difference Water Index

NDDI values typically range from -1 to +1.

The advantage of NDDI lies in its ability to detect areas experiencing water stress even when NDVI values remain high, a situation commonly observed during the early stages of drought when vegetation still appears green but has already started losing water in leaf tissues. This makes NDDI particularly suitable for early detection and drought risk warning.

2.2.8 TCI (Temperature Condition Index)

TCI reflects the surface temperature condition of vegetation, based on the assumption that when crops or vegetation experience drought-induced stress, their surface temperature increases significantly. This index uses Land Surface Temperature (LST) data normalized over a time series to evaluate thermal stress in vegetation.

The formula for calculating TCI is as follows (Kogan, 1995):

$$TCI = \frac{(T_{max} - T)}{T_{max} - T_{min}} \times 100 \quad (8)$$

Where:

T: Land surface temperature at the time of observation.

Tmin: Minimum land surface temperature in the time series.

Tmax: Maximum land surface temperature in the time series.

The value of TCI ranges from 0 to 100.

TCI is often used in combination with VCI to derive the composite Vegetation Health Index (VHI). This is a useful tool to distinguish between water stress (VCI) and heat stress (TCI), thereby identifying the causes affecting vegetation.

2.2.9 TVDI (Temperature Vegetation Dryness Index)

TVDI is one of the most widely used composite drought indices today, thanks to its ability to integrate information from two important data sources: land surface temperature (LST) and vegetation index (NDVI). This index was developed based on the linear relationship between LST and NDVI under dry conditions, forming a “drought triangle” on the scatterplot (Bian et al., 2023). The core idea is that, for the same level of vegetation greenness, areas with higher surface temperatures are considered drier.

The formula for TVDI (Sandholt et al., 2002):

$$TVDI = \frac{T_s - T_{smin(NDVI)}}{T_{smax(NDVI)} - T_{smin(NDVI)}} \times 100 \quad (9)$$

Where:

T_s : Land surface temperature at each pixel.

$T_{smax(NDVI)}$: Upper boundary (dry edge) – the maximum temperature value corresponding to each NDVI.

$T_{smin(NDVI)}$: Lower boundary (wet edge) – the minimum temperature value corresponding to each NDVI.

The TVDI values range between 0 and 1. The TVDI has the outstanding advantage of eliminating the influence of crop type and background environment, since it is calculated specifically based on the NDVI–LST relationship for each region and time period. This enhances the objectivity and broad applicability of TVDI across different climatic and ecological zones.

2.2.10 VHI (Vegetation Health Index)

The Vegetation Health Index (VHI) was designed to provide an overall view of vegetation health by integrating both temperature conditions (TCI) and growth status (VCI). This composite index helps assess the combined effects of thermal stress and physiological stress due to drought on vegetation.

The VHI is calculated as follows (Kogan, 1997):

$$VHI = \alpha \cdot VCI + (1 - \alpha) \cdot TCI, \alpha = 0.5 \quad (10)$$

Where:

VCI: Vegetation Condition Index.

TCI: Temperature Condition Index.

α : weight factor (commonly set to 0.5).

The main advantage of VHI lies in its ability to combine two critical factors affecting vegetation moisture and temperature, thus providing a more accurate reflection of overall plant health. This is particularly useful in areas where relying on a single index may lead to misleading interpretations.

3. Application of Artificial Intelligence in Drought Monitoring

3.1 The Role of Machine Learning in Drought Risk Zoning and Monitoring

AI, particularly ML, has emerged as a powerful tool for extracting, analyzing, and forecasting drought conditions using multi-source remote sensing data. ML algorithms enable the processing of large datasets with nonlinear relationships, automatically detecting complex patterns and improving the accuracy of classifying different drought severity levels across space. The application of machine learning in drought monitoring not only accelerates data analysis but also enhances forecasting capabilities, thereby supporting timely decision-making for resource and agricultural management agencies (Prodhan et al., 2022). In this study, three machine learning algorithms are employed: RF, SVM, and GTB.

3.2 Typical machine learning algorithms

3.2.1 Random Forest

RF is an ensemble machine learning algorithm developed based on the Bagging technique (Breiman, 2001). The algorithm constructs a "forest" of multiple independent decision trees, each trained on a random subset of the dataset. The combination of these decision trees enhances the model's generalization ability and reduces overall variance. RF performs classification through majority voting or regression tasks through averaging (Parmar et al., 2018).

3.2.2 Support Vector Machine

SVM is a supervised machine learning algorithm that excels in classification tasks due to its ability to generate an optimal hyperplane to separate data classes (Suthaharan, 2016). The main objective of SVM is to identify a hyperplane that maximizes the margin between the hyperplane and the nearest data points from both classes. SVM can also be extended to nonlinear classification problems through the use of kernel techniques, which map the data into a higher-dimensional feature space.

3.2.3 Gradient Tree Boosting

Gradient Tree Boosting is a powerful machine learning algorithm belonging to the boosting family, developed based on the idea of sequentially improving weak models (decision trees) to create a strong

overall model (Friedman, 2002). In GTB, decision trees are built sequentially, with each tree attempting to minimize the errors of the previous trees by learning from the gradient of the loss function.

4. Integrated Drought Monitoring Methodology

This study aims to develop a drought risk prediction model for Quang Tri Province by integrating multi-source remote sensing data with machine learning techniques. The entire workflow is implemented on the Google Earth Engine cloud platform to leverage its capabilities in big data processing and automation. The data processing procedure consists of the following steps:

Step 1: Data assessment and collection

The first step is to conduct an overview assessment of the study area and collect remote sensing data. Quang Tri Province was selected due to its frequent exposure to prolonged drought conditions. The study employs satellite imagery from two primary sources: Landsat and Sentinel-2, covering the period from 2016 to 2025, restricted to the dry season months (March to September).

Step 2: Calculation of drought indices

Ten drought indices, as presented in Section 2.2, are calculated from the pre-processed satellite imagery. Specifically, TVDI and TCI are derived from Landsat imagery, while the remaining indices are computed from Sentinel-2.

Step 3: Sample dataset construction

TVDI is used as the basis for labeling drought severity at the pixel level, as it directly reflects surface dryness (Trinh, 2014). TVDI values are classified into five thresholds, corresponding to five drought levels: No drought (0.0 - 0.2), Mild drought (0.2 - 0.4), Moderate drought (0.4 - 0.6), Severe drought (0.6–0.8), and Extreme drought (0.8 - 1.0) (Chen et al., 2023). This multi-level classification allows the machine learning models not only to distinguish between drought and non-drought areas but also to quantify the severity of drought. From each drought category, 200 pixels (a total of 1,000 samples) are selected to build the training and testing dataset. Each pixel is assigned a label representing drought severity along with a feature set comprising the values of the 10 drought indices.

Step 4: Drought risk modeling using machine learning

Three machine learning models: Random Forest (RF), Support Vector Machine (SVM), and Gradient Tree Boosting (GTB) are employed for performance comparison. The dataset is split into a training set (70%) and a testing set (30%). Model performance is evaluated using overall accuracy and the Kappa coefficient to assess and compare the effectiveness of the algorithms (Maxwell et al., 2018)

Step 5: Drought risk prediction and mapping

The best-performing model is applied to the entire set of remote sensing imagery for Quang Tri Province to generate drought risk maps. The results are presented as annual drought risk maps, illustrating the spatial dynamics of drought across the study area for the period 2016 - 2025.

5. Discussion of effectiveness and challenges

The integrated drought monitoring approach, which combines multi-source remote sensing data with machine learning techniques, demonstrates significant effectiveness in capturing spatiotemporal drought dynamics. By utilizing a diverse set of drought indices, the method provides a comprehensive assessment of vegetation health, soil moisture, and thermal conditions, thereby enabling early detection and classification of drought severity. Furthermore, the application of machine learning models such as Random Forest, Support Vector Machine, and Gradient Tree Boosting enhances predictive accuracy and facilitates the generation of detailed drought risk maps. These outputs are particularly valuable for resource managers and policymakers in drought-prone regions.

However, the methodology also faces several challenges. The quality and availability of remote sensing data may be constrained by cloud cover, atmospheric effects, and spatial resolution limitations, particularly in heterogeneous landscapes. Model performance is highly dependent on the quality and representativeness of training data, which may not fully capture local variability in vegetation and climate conditions. In addition, machine learning models often function as “black boxes,” making it difficult to interpret the underlying drivers of drought stress. These limitations may affect the consistency and transferability of the method across different regions and timescales.

To address these challenges, several future directions are proposed. First, integrating satellite-derived indices with ground-based observations (e.g., soil moisture sensors, meteorological records, and crop yield data) will strengthen model calibration and reliability. Second, multi-source and multi-scale data fusion,

including MODIS, VIIRS, or SMAP, can improve the spatial and temporal continuity of drought monitoring. Third, developing hybrid models that combine machine learning with physical or statistical approaches may enhance both predictive power and interpretability. The incorporation of explainable AI (XAI) techniques can further improve the transparency of model outputs and their utility in decision-making. Finally, the establishment of automated, cloud-based operational frameworks will enable real-time drought monitoring and early warning, thereby supporting adaptive management in the context of increasing climate variability.

6. Conclusion

This study presented a comprehensive and modern approach to drought monitoring, integrating multi-source remote sensing data and artificial intelligence techniques on the Google Earth Engine platform. The method leverages the advantages of Landsat and Sentinel-2 to compute ten drought indices, which are then processed using powerful machine learning algorithms such as RF, SVM, and Gradient Tree Boosting to classify and predict drought risk. This integrated framework addresses the limitations of traditional monitoring approaches by improving accuracy, automation, and large-scale data processing capacity, thereby providing timely and reliable information for drought risk management and water resource planning.

To further enhance this methodology, future research could focus on integrating ground-based observations, meteorological data, and socio-economic factors to improve accuracy and predictive capability. The adoption of more advanced deep learning architectures, such as Convolutional Neural Networks and Long Short-Term Memory networks, could strengthen time-series modeling and long-term forecasting. Moreover, developing interpretable or “white-box” AI models would provide greater insights into the decision-making process, thereby increasing the transparency and usability of predictions. Finally, scaling up the proposed framework could contribute to the development of an intelligent and efficient nationwide drought early warning system.

Acknowledgment

The author thanks unnamed reviewers for their valuable comments, which helped me to improve the quality of the manuscript.

References

- Affandy, N. A., Iranata, D., Anwar, N., Maulana, M. A., Prasetyo, D. D., Wardoyo, W., & Sukojo, B. M. (2024). Assessment of Agricultural Drought Using the Normalized Difference Drought Index (NDDI) to Prediction Drought at Corong River Basin. *International Journal of Integrated Engineering*, 16(1), 378-393.
- AghaKouchak, A., Farahmand, A., Melton, F. S., Teixeira, J., Anderson, M. C., Wardlow, B. D., & Hain, C. R. (2015). Remote sensing of drought: Progress, challenges and opportunities. *Reviews of Geophysics*, 53(2), 452-480.
- Aşık, Ş. (2020). Agricultural Drought Monitoring Using Surface Temperature and Vegetation Indices from Satellite Images. *Ege Üniversitesi Ziraat Fakültesi Dergisi*, 151-160.
- Benedict, T. D., Brown, J. F., Boyte, S. P., Howard, D. M., Fuchs, B. A., Wardlow, B. D., ... & Evenson, K. A. (2021). Exploring VIIRS continuity with MODIS in an expedited capability for monitoring drought-related vegetation conditions. *Remote Sensing*, 13(6), 1210.
- Bian, Z., Roujean, J. L., Fan, T., Dong, Y., Hu, T., Cao, B., ... & Liu, Q. (2023). An angular normalization method for temperature vegetation dryness index (TVDI) in monitoring agricultural drought. *Remote Sensing of Environment*, 284, 113330.
- Breiman, L. (2001). Random forests. *Machine learning*, 45, 5-32
- Friedman, J. H. (2002). Stochastic gradient boosting. *Computational statistics & data analysis*, 38(4), 367-378.
- Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., ... & Michaelsen, J. (2015). The climate hazards infrared precipitation with stations-a new environmental record for monitoring extremes. *Scientific data*, 2(1), 1-21.
- Gao, B. C. (1996). NDWI-A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote sensing of environment*, 58(3), 257-266.
- Ghaleb, F., Mario, M., & Sandra, A. N. (2015). Regional landsat-based drought monitoring from 1982 to 2014. *Climate*, 3(3), 563-577.

- Gu, Y., Brown, J. F., Verdin, J. P., & Wardlow, B. (2007). A five-year analysis of MODIS NDVI and NDWI for grassland drought assessment over the central Great Plains of the United States. *Geophysical research letters*, 34(6).
- Heim Jr, R. R. (2002). A review of twentieth-century drought indices used in the United States. *Bulletin of the American Meteorological Society*, 83(8), 1149-1166.
- Huete, A.R., 1988: A Soil-adjusted Vegetation Index (SAVI). *Remote Sensing of Environment*, 25(3): 295–309. DOI: 10.1016/0034-4257(88)90106-X.
- Kamble, M. V., Ghosh, K., Rajeevan, M., & Samui, R. P. (2010). Drought monitoring over India through normalized difference vegetation index (NDVI). *Mausam*, 61(4), 537-546.
- Khosravi, H., Haydari, E., Shekoohizadegan, S., & Zareie, S. (2017). Assessment the effect of drought on vegetation in desert area using Landsat data. *The Egyptian Journal of Remote Sensing and Space Science*, 20, S3-S12.
- Kogan, F. N. (1997). Global drought watch from space. *Bulletin of the American meteorological society*, 78(4), 621-636.
- Kogan, F., Goldberg, M., Schott, T., & Guo, W. (2015). Suomi NPP/VIIRS: improving drought watch, crop loss prediction, and food security. *International Journal of Remote Sensing*, 36(21), 5373-5383.
- Kogan, F.N. (1995). Application of vegetation index and brightness temperature for drought detection. *Advances in Space Research*, 15(11), 91–100. [https://doi.org/10.1016/0273-1177\(95\)00079-T](https://doi.org/10.1016/0273-1177(95)00079-T).
- MONRE (2020). National Climate Change Report 2020. Ministry of Natural Resources and Environment of Vietnam
- National Aeronautics and Space Administration (NASA). Moderate Resolution Imaging Spectroradiometer. <https://modis.gsfc.nasa.gov/about/>. Accessed April 23, 2025.
- Nicholson, S. E. (2014). A detailed look at the recent drought situation in the Greater Horn of Africa. *Journal of Arid Environments*, 103, 71-79.
- Parmar, A., Katariya, R., & Patel, V. (2018, August). A review on random forest: An ensemble classifier. In *International conference on intelligent data communication technologies and internet of things* (pp. 758-763). Cham: Springer International Publishing.
- Peters, A. J., Walter-Shea, E. A., Ji, L., Vina, A., Hayes, M., & Svoboda, M. D. (2002). Drought monitoring with NDVI-based standardized vegetation index. *Photogrammetric engineering and remote sensing*, 68(1), 71-75.
- Proadhan, F. A., Zhang, J., Hasan, S. S., Sharma, T. P. P., & Mohana, H. P. (2022). A review of machine learning methods for drought hazard monitoring and forecasting: Current research trends, challenges, and future research directions. *Environmental modelling & software*, 149, 105327.
- Spinoni, J., Naumann, G., Vogt, J. V., & Barbosa, P. (2015). The biggest drought events in Europe from 1950 to 2012. *Journal of Hydrology: Regional Studies*, 3, 509-524.
- Suthaharan, S. (2016). Support vector machine. In *Machine learning models and algorithms for big data classification: thinking with examples for effective learning* (pp. 207-235). Boston, MA: Springer US.
- The European Space Agency (ESA). Sentinel-2 User Handbook. https://sentinel.esa.int/documents/247904/685211/sentinel-2_user_handbook. Accessed April 23, 2025.
- Trinh, L.H. (2014). Application of Landsat thermal infrared remote sensing data to study soil moisture based on vegetation temperature drought index. *Journal of Earth Sciences*, 36(3), 262-270.
- Tucker, C. J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote sensing of Environment*, 8(2), 127-150.
- U.S. Geological Survey (USGS). Landsat Satellite Missions <https://www.usgs.gov/landsat-missions/landsat-satellite-missions>. Accessed April 23, 2025.
- UNCCD. (2020). Global Drought Report. United Nations Convention to Combat Desertification.
- Weng, Q. (2009). Thermal infrared remote sensing for urban climate and environmental studies: Methods, applications, and trends. *ISPRS Journal of photogrammetry and remote sensing*, 64(4), 335-344.

- Williams, A. P., Cook, E. R., Smerdon, J. E., Cook, B. I., Abatzoglou, J. T., Bolles, K., ... & Livneh, B. (2020). Large contribution from anthropogenic warming to an emerging North American megadrought. *Science*, 368(6488), 314-318.
- Xiang, K., Yuan, W., Wang, L., & Deng, Y. (2020). An LSWI-Based Method for Mapping Irrigated Areas in China Using Moderate-Resolution Satellite Data. *Remote Sensing*, 12(24), 4181. <https://doi.org/10.3390/rs12244181>

