The Pharma Innovation Journal 2023; 12(3): 2054-2062

www.ThePharmaJournal.com

The Pharma Innovation

ISSN (E): 2277-7695
ISSN (P): 2349-8242
NAAS Rating: 5.23
TPI 2023; 12(3): 2054-2062

www.thepharmajournal.com
Received: 27-12-2022
Accepted: 28-01-2023

Applications of Google earth engine for big data analytics

Ankur P Vora, Dr. Chetan R Dudhagara, Dr. DR Kathiriya and Dr. RS Pundir

Ankur P Vora
Senior Research Fellow, Centre for Agricultural Market Intelligence under NAHEP-CAAST, Anand Agricultural University, Anand, Gujarat, India

Dr. Chetan R Dudhagara
Assistant Professor & Head, Department of Communication and Information Technology, International Agri-Business Management Institute, Anand Agricultural University, Anand, Gujarat, India

Dr. DR Kathiriya
Director IT (I/c) and Principal & Dean, College of Agricultural Information Technology, Anand Agricultural University, Anand, Gujarat, India

Dr. RS Pundir
Principal Investigator, Centre for Agricultural Market Intelligence under NAHEP-CAAST, Anand Agricultural University, Anand, Gujarat, India

Corresponding Author:
Dr. RS Pundir
Principal Investigator, Centre for Agricultural Market Intelligence under NAHEP-CAAST, Anand Agricultural University, Anand, Gujarat, India

Abstract
Massive amounts of data have been gathered by remote sensing (RS) systems for decades, but managing and analyzing them with standard software and desktop computing resources is impractical. To effectively handle the difficulties of big data analytics, Google has created a cloud computing platform called Google Earth Engine (GEE). This platform makes it possible to process big geodata over vast areas and to continuously monitor the environment. The GEE offers access to petabytes of publicly accessible remote sensing datasets and machine learning algorithms using Google's computational infrastructure and a library of Application Programming Interfaces (APIs) with development environments that support well-known coding languages, such as JavaScript and Python. Together, these fundamental components give users the ability to analyze and visualize big data without access to supercomputers or specialized coding knowledge. However, ten years after its inception, the effects of GEE on data analytics have not been properly investigated. As a result, a systematic review of GEE is required to provide readers with the "big picture" of the current state and general trends in GEE. GEE has also been used in a wide variety of applications, including land use/land cover classification, hydrology, urban planning, natural disasters, climate analyses and image processing.

Keywords: Classification and Regression tree, data analytics, Google earth engine, random forest, support vector machine

Introduction
Big data computing, big data collaboration and big data methodologies are some of the more prevalent challenges in handling big data. The appropriate data identification, deployment, representation, fusion, as well as data visualization and interpretation; the individual challenges are related to the big data life cycle in different applications (Chi *et al.*, 2016) [6]. Big data from remote sensing naturally possesses features like dynamic state, multi-scale, and nonlinear. Big data from remote sensing is also nonlinear because time series data are frequently noisy and nonlinear. On the other hand, the extrinsic characteristics of remote sensing big data include the multi-source, high-dimensional, and isomer characteristics (Liu, 2015) [20].

GEE is a free cloud platform that hosts petabyte scales of remotely sensed data from over 40 years, including Landsat, MODIS, National Oceanographic and Atmospheric Administration Advanced Very High-Resolution Radiometer (NOAA AVHRR), Sentinel 1, 2, 3 and 5-P data, and Advanced Land Observing Satellite (ALOS) data. To reduce computational time, the GEE platform makes use of Google's computational infrastructure to enable parallel geospatial data processing (Gorelick *et al.*, 2017) [11]. GEE is the most widely used big geo data processing platform, facilitating scientific discovery by giving users free access to a large number of remotely sensed datasets. GEE is accessible to users through an internet-based Application Programming Interface (API) and a web-based Interactive Development Environment (IDE) such as JavaScript and Python (Tamiminia *et al.*, 2020) [37]. This platform also includes a number of built-in algorithms, such as classification algorithms, for analyzing data on a global scale and assisting researchers in developing their own algorithms with less effort (Amani *et al.*, 2019) [3].

The cloud computing capabilities of GEE eliminate the need to store, process, and analyze the massive amounts of satellite data on a computer by enabling the processing of petabytes of image data alongside other vector data in a cloud environment. Computers with fast processing speeds and large storage capacities are no longer in high demand. Users do not have to rely entirely on specialist remote sensing software, such as Environment for Visualizing Images

(ENVI) and Earth Resources Data Analysis System (ERDAS) Imagine; however, they may be required for special functions that GEE does not provide (such as object-based image analysis). Satellite images do not need to be downloaded, which is a huge benefit for areas with slow internet connections, but an internet connection is still required to use GEE (Kumar and Mutanga, 2018) [18].

GEE Datasets

GEE, as previously stated, contains a massive number of datasets, including raw datasets, preprocessed data, elevation models, and products at global, national, and regional scales. The Google Earth Engine (GEE), which offers high-performance parallel computing, massive remote sensing and geospatial data, is available for free (Wang *et al*., 2020) [40]. Landsat, Sentinel, MODIS, ASTER, SAR, weather, precipitation, atmosphere, rainfall, topography, soil, snow, and other remotely sensed datasets are available in GEE. Some of the datasets that are frequently used by users are discussed in greater detail below.

GEE includes datasets acquired by Sentinel satellites, developed by the European Space Agency (ESA). Sentinel-1 Synthetic Aperture RADAR (SAR) (2014-present), Sentinel-2 multispectral (2015-present), Sentinel-3 Ocean and Land Color (2016-present) and Sentinel-5P Tropospheric Monitoring (2018-present) datasets are part of the Sentinel collection. GEE users have extensively used Sentinel-1 and Sentinel-2 for a variety of applications. Their 10 m spatial resolution allows to analyze objects at a higher resolution than Landsat images. User can also make the training and validation steps in image classification tasks easier (Amani *et al*., 2020) [2]. Inoue *et al*. (2020) [15] extracted paddy fields using Sentinel-1 Ground Range Detected (GRD) images because the Sentinel-1 C-band SAR instrument can obtain data through cloud cover and Sentinel-2 Multispectral Imager (MSI) images because they are more reliable and robust for extracting irrigated areas.

Landsat datasets are useful for performing temporal analysis. Landsat 1-3 (1972-1983), Landsat-4 (1982-1993), Landsat-5 (1984-2012), Landsat-7 (1999-present), and Landsat-8 (2013-present) are the seven multispectral satellites in the Landsat collection. Clouds can obscure the images of optical sensors on Landsat satellites (Amani *et al*., 2020) [2]. As a result, temporal cloud detection, masking, and removal are critical preprocessing steps in a variety of applications, including image classification using multi-temporal imagery (Garcia *et al*., 2018) [10]. Landsat-based datasets from GEE have been used in a variety of applications. Landsat data are among the most commonly used in LULC classification. Landsat-8 Operational Land Imager surface reflectance tier 1 (Landsat-8 OLI SR T1) data from the GEE were used as the primary remote sensing information for LULC classification by Qu *et al*., (2021) [32]. Furthermore, the GEE's various Landsat 8 time-series image collections can be used to generate the NDVI spectral curve. The NDVI is strongly related to the amount of radiation absorbed by actively growing plants; vegetation absorbs strongly in the red portion of the spectrum and strongly reflects in the near-infrared portion of the spectrum (Aghababaei *et al*., 2021) [1].

Data pre-processing

GEE includes several image processing functions that users can use to analyze remote sensing data. Time-series analysis, feature extraction, image color composite-visual interpretation and image pre-processing techniques can be used to analyze satellite imagery in place of machine learning techniques. Image pre-processing includes things like expanding the study area, filtering data based on the purpose of the research, masking and mosaicking and composing various bands, among other things. Furthermore, GEE's rich API documents absorb many superior remote sensing image processing algorithms, which are very beneficial to people who focus on data processing rather than programming (Kandekar *et al*., 2021) [16]. The following are descriptions of various pre-processing methods.

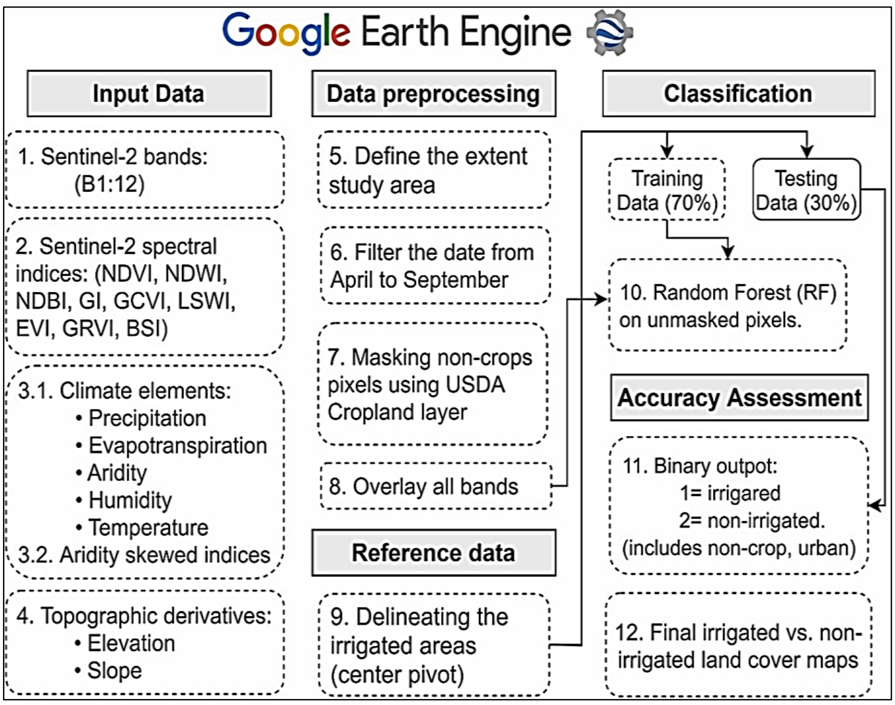

Fig 1: A flow diagram for data processing and the analysis steps

For each pixel, the C Function of Mask (CFMASK) model can be used to process a remote sensing dataset. The "pixel qa" (pixel quality attributes) band can be used to determine the pixel quality of each pixel for water, clouds, snow and cloud shadows. Using GEE, a user can also extract information about clouds and their shadows (Qu *et al*., 2021) [32]. All of the Landsat (TM, ETM+ and OLI) surface reflectance images (16,760 scenes) can be filtered using the GEE platform based on the time period needed for different studies. An annual water map requires the acquisition of high-quality satellite imagery, so bad satellite images resulting from ineffective pixels, clouds, cloud shadows and snow are masked using the Fmask in GEE by Xia *et al*., (2019) [41]. Xiong *et al*., (2016) [42] used Sentinel-2 and Landsat-8 to combine five different bands (blue, green, red, NIR, and NDVI) for each period (January-June 2016, July-December 2015). In addition, he calculated and mosaicked Sentinel-2 TOA reflectance values using median values to create layer stacks for each season separately in GEE. Kandekar *et al*., (2021) [16] used the GEE platform for pre-processing a Sentinel 2 image with the least cloud cover or cloud-free image, as well as Mosaicking a set of images after processing.

Reference data

Data used to classify or categorize other data is referred to as reference data. They are typically static or slowly changing over time. It is necessary to distinguish between reference data and master data. Reference datasets are required for accuracy assessment and validation of remote sensing study results. The user can validate and verify their results using reference data. Furthermore, High precision reference data and an appropriate classification method are required for LULC classification (Qu *et al*., 2021) [32]. Different types of reference data, such as ground data and tertiary data, can be used for class identification and labelling. These data also aided users in improving machine learning classification performance, resulting in optimal results (Oliphanta *et al*., 2019) [25].

Classification in GEE

The GEE platform offers comprehensive support for performing various image classification processes. GEE includes several in-built machine learning classifiers, such as RF, SVM, CART and others that can be used to train the classifier with samples obtained from various sampling designs.

1. **Classification and Regression Tree (CART):** CART is a binary decision tree classifier that makes straightforward decisions for logical if-then questions. The classifier examines the input variables and selects the variable with the greatest information gain, on which the node splits at each level. The input data is randomly divided into a certain number of groups in this technique and trees are generated using all but one of the groups. The pruned tree with the minimum deviations is chosen after the excluded group is used to validate the tree (Breiman *et al*., 1984) [5].

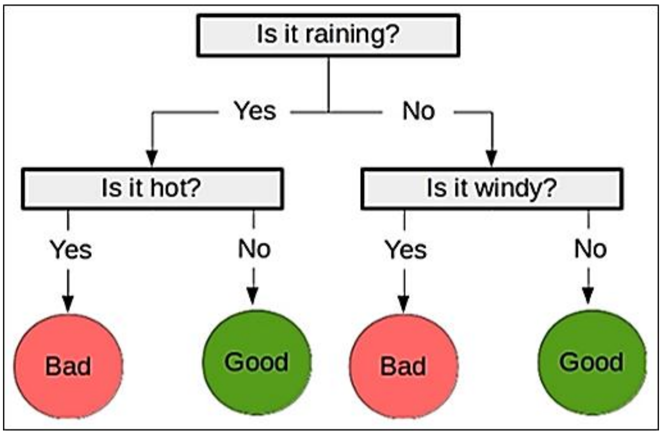

Fig 2: CART algorithm

2. **Support Vector Machine (SVM):** SVM is a popular classifier that works by locating the best hyperplane that separates the decision boundary between different classes. The selection of support vectors is primarily determined by the cost parameters C, Gamma, and kernel functions. The level of punishment for misclassified data is determined by the cost parameter. The greater the value of C, the less misclassified data within a class. Because C is a scale parameter, an exponentially expanding sequence of C provides a better approximation to effective parameter selection, according to Hsu, Chang, and Lin (2003) [14]. Additionally, for training purposes on large datasets, linear kernels are preferred. The gamma parameter is invalid for linear kernels.

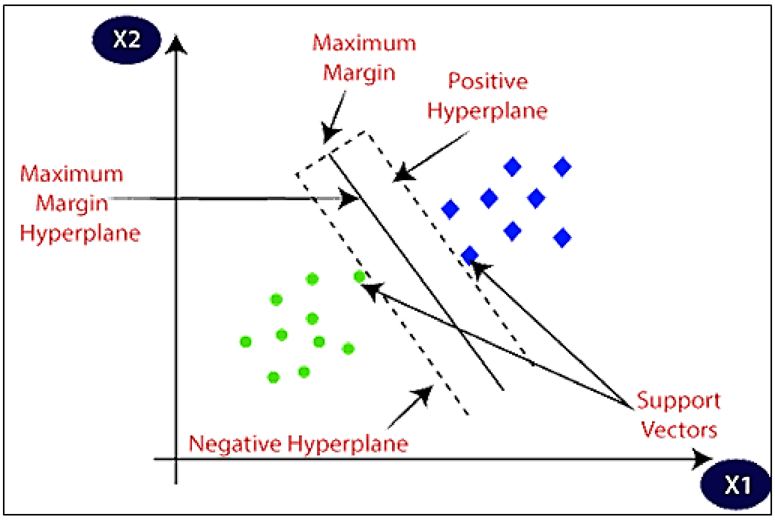

Fig 3: Support Vector Machine classifier

3. **Relevant Vector Machine:** Tipping and Faul (2003) [39] developed RVM, a Bayesian linear model and probabilistic extension of SVM that provides sparse solutions to classification tasks. The Bayesian inference approach considers feature set 'w' (weights) related to input observations to be random variables and understands the distribution of these weights in relation to given input and target data. This posterior probability distribution of w aids in the prediction of target values for new input data. The most useful aspect of this Bayesian approach is that it eliminates all irrelevant variables and generates a simple model that explains the data pattern (Tipping, 2004) [38]. This is an appealing feature for classification in remote sensing, where training samples are scarce.

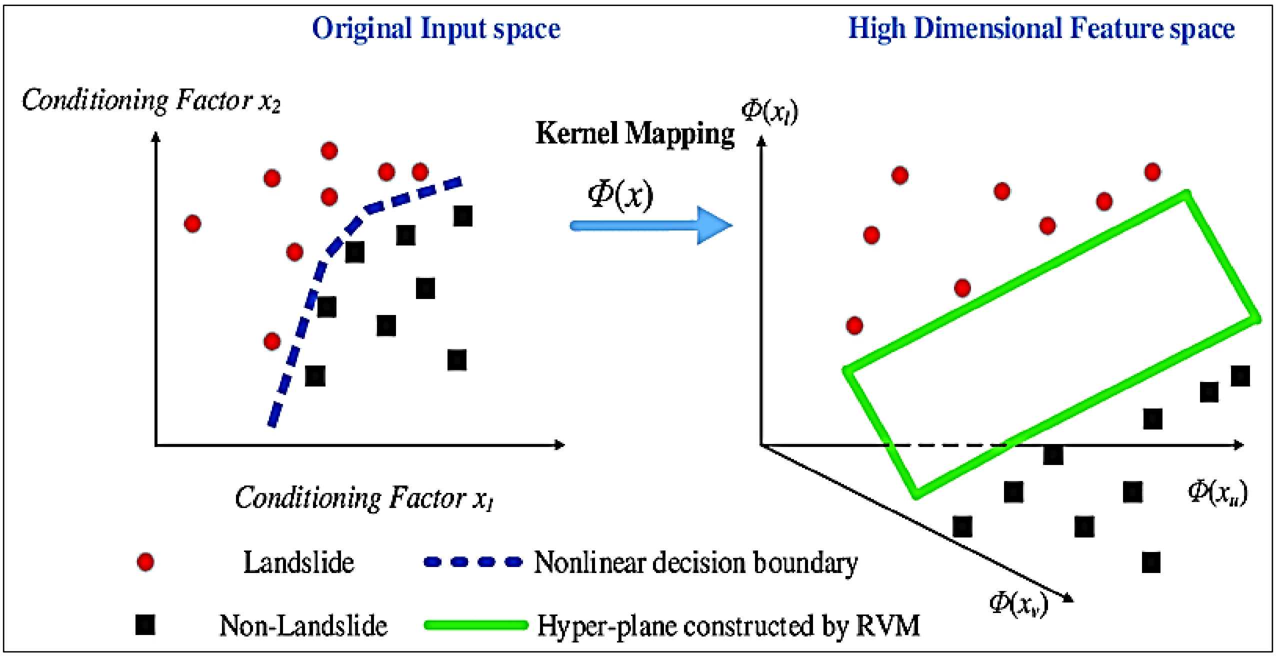

Fig 4: Relevant Vector Machine classifier

4. **Random Forest:** An ensemble of k CART classifiers called a random forest solves the over fitting issue with CART. Since it is the LULC Classification's most widely used classifier, it is a classifier of interest for numerous studies. Random forest employs the bagging technique, which selects a subset of characteristics from the input observations at random for each tree. The number of trees and the variables at each split are the primary input parameters for RF. A very large number of trees does not always mean that the classification is more accurate, as after a certain number of trees, more trees become redundant and do not contribute as much to label prediction. Numerous studies have demonstrated that RF results in LULC classification with relatively high classification accuracy. Furthermore, the benefits of the RF technique include simple parameterization, the capacity to handle collinear features and the processing of high-dimensional data (Pelletier *et al*., 2016) [30].

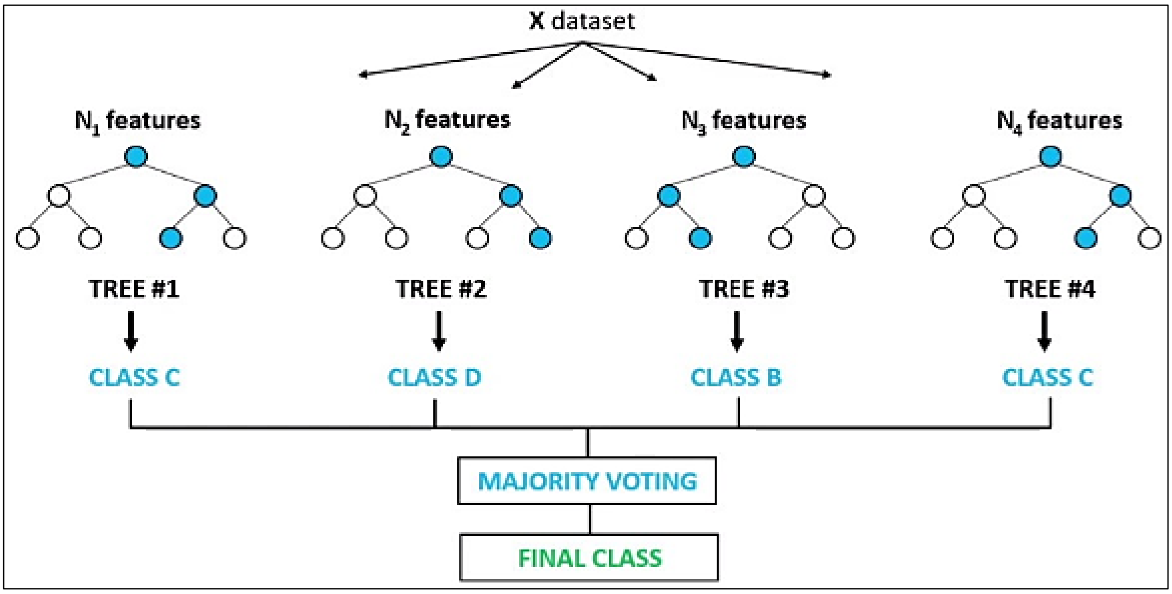

Fig 5: Random Forest Classifier

Accuracy Assessment

The fundamental principle underlying all accuracy assessments is to compare estimates to reality and quantify the difference. It is the proportion of correct predictions to the total number of predictions. The precision of the input raster data on which the classification should be formed, as well as the locational precision of the validation data that can be used to assess the product, must be considered when evaluating the accuracy of a mapping product. A team-based balanced sampling approach utilizing data that the analysts did not have is an ideal way to assess such accuracies. This ensures that the classification cannot be fitted to the validation data and that the accuracy is not artificially inflated as a result of the training data coming from the same source as the validation data (Oliphanta *et al*., 2019) [25].

The Overall Accuracy (OA) measures the model's overall performance and is calculated as the proportion of correctly classified pixels divided by the total number of classified pixels. The agreement between the reference data and the classification is represented by Producer's Accuracy (PA), whereas User's Accuracy (UA) evaluates how well the classified pixels agree with the known reference data. The UA and PA are associated with errors of commission and omission. The Kappa coefficient (Kc) compares the actual agreement between the reference data and the classifier used for classification to the probability of agreement between the reference data and the random classifier (Xia *et al*., 2019) [41]. The PA, UA, OA and Kc Equation is as follows:

$$PA = \frac{Sij}{Si} \times 100\%$$

$$UA = \frac{Sij}{Sj} \times 100\%$$

$$OA = St/n$$

$$Kc = n \times St - \sum_{i=1}^{r} \frac{SiSj}{n^2} - \sum_{i=1}^{r} SiSj$$

Where,

St = Sum of correctly classified pixels.

n = Sum of validation pixels.

r = Number of rows.

Sij = Observation in a row I and column j.

Si = Marginal total of row i.

Sj = Marginal total of column j.

Applications of GEE

Earth Engine is being used in a wide range of disciplines, including global forest change (Hansen *et al*., 2013) [12], global surface water change (Pekel *et al*., 2016) [29], crop yield estimation (Lobell *et al*., 2015) [21], rice paddy mapping (Dong *et al*., 2016) [9], urban mapping (Zhang *et al*., 2015; Patel *et al*., 2015) [45, 28], flood mapping (Coltin *et al*., 2016) [6], fire (Sturrock *et al*., 2014) [36]. Several GEE applications for natural resource management are described in detail below.

Agriculture

Agricultural products are not only important in human life, but they are also important economically. As a result, agriculture can be regarded as a source of livelihood as well as a contributor to national revenue (McArthur and Sachs, 2019) [23]. GEE hosted a large number of publicly available RS datasets that can be used to analyze agricultural productivity, quality, profitability, and sustainability. GEE has been used in plantation mapping and monitoring, cropland mapping, crop condition monitoring, crop yield estimation, irrigation mapping, and other agricultural studies. The GEE platform enabled the fusion of Terra MODIS and Landsat data to estimate Gross Primary Productivity of seven crops in Montana, USA, from 2008 to 2015 at a spatial resolution of 30 m. Cropland productivity patterns and seasonal variations were estimated and compared favorably to country-level crop data (He *et al*., 2018) [13]. Furthermore, in the GEE for the state of Mato Grosso, Brazil, temporal series multispectral images and vegetation indices (near-infrared and red) were used to estimate and map soybean areas. From the identification of soybean areas using the Perpendicular Crop Enhancement Index, a multitemporal algorithm of the Perpendicular Vegetation Index (PVI) of MODIS, OLI, and MSI images was created (PCEI). The authors reported that using MODIS images for soybean area monitoring via the Google Earth Engine platform was a viable and promising automated alternative for large-scale soybean area estimates (Silva Junior, *et al*., 2020) [34].

Vegetation

Vegetation is important in numerous biochemical cycles that interact with water, soil and air, either directly or indirectly (Smith *et al*., 2014) [35]. GEE uses cloud computing services to monitor vegetation covers over an extended period of time. Furthermore, researchers can use GEE's publicly available RS data for vegetation monitoring at various spatial scales. GEE has been extensively used for vegetation mapping, monitoring vegetation dynamics, deforestation, vegetation and forest expansion, monitoring forest health, mapping forests, monitoring pastures and evaluating rangeland. In the Himalayan Upper Khoh River (UKR) Basin, for instance, GEE was used to examine the relationship between spatiotemporal variability of vegetation greenness and related meteorological and hydrological causes at annual and seasonal scales (Kumari *et al*., 2021) [19].

Hydrology

Obtaining accurate information about water resources is extremely important because water, whether in liquid form (such as lakes, rivers and reservoirs) or solid form (such as snow, ice, and glaciers) in the cryosphere, is a necessary component for life. Monitoring inland, coastal and arctic water resources is also useful in studies of climate change (Yao *et al*., 2019) [43]. For surface water dynamics monitoring, bathymetry, shoreline and coastal studies, lake and reservoir mapping and monitoring, glacier studies, snow ablation and snow mapping, suspended sediments and river studies, and water quality assessment, GEE was effectively used. For instance, sentinel imagery and the Google Earth Engine Platform were used to study the surface water dynamics at the Jayakwadi dam in Maharashtra. For this, a number of excellent remote sensing image processing algorithms have been integrated into GEE. The findings demonstrated the validity, uniqueness, and speed of existing methodologies for estimating the maximum and minimum extent of surface water (Kandekar *et al.*, 2021) [16].

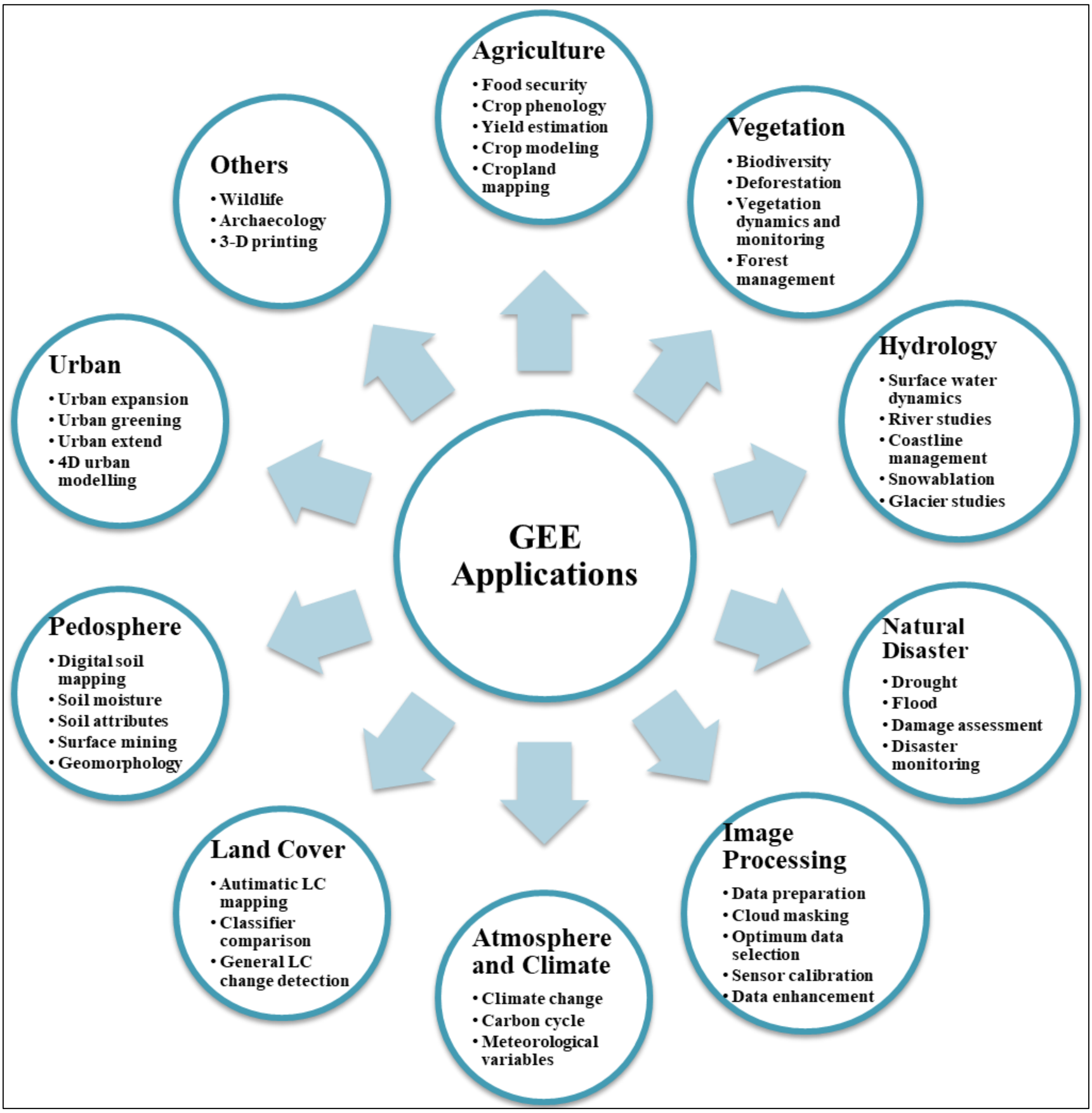

Fig 6: GEE applications (M. Amani *et al.*, 2020) [2]

Natural Disaster

Natural disasters are extreme and unexpected events caused by Earth's natural processes. These occurrences cause problems on the surrounding environment and human life (Bayissa *et al.*, 2017) [4]. GEE was used for drought monitoring, flood mapping and risk assessment, wildfire severity mapping, landslide analyses, hurricane studies, and tsunami studies, among other things. For example, the value of soil moisture data for drought monitoring and crop forecasting was demonstrated using two global soil moisture datasets and a set of soil moisture web-based processing tools developed by GEE. When compared to SPIs, soil moisture anomalies had a shorter drought duration but a higher intensity. The incorporation of global soil moisture data into the GEE data catalogue, as well as the development of the research's web-based tools, will enable a wide range of users to quickly and easily analyze the impact of drought and improve drought risk assessment and early warning planning (Sazib *et al.*, 2018) [33].

Image processing

In present era, almost all EO platforms are equipped with digital sensors, resulting in terabytes of data being generated and stored in digital formats every day. The RS images are widely used in a variety of applications and for a variety of purposes. As a result, it is critical to develop and improve digital image processing algorithms in order to fully realize the potential of digital images. Furthermore, because the quality of each input data directly affects the final accuracy of studies, image processing must be regarded as a requirement. The most important criteria in developing image processing algorithms are precision, level of automation, reliability, computational complexity, and time consumption (Chova *et al.*, 2017) [7] (Kong *et al.*, 2019) [17]. GEE was used by the researchers to create a variety of efficient and useful image processing algorithms, including cloud masking, data selection and enhancement, image-based sensor calibration, and training sample migration. The quality of the image segmentation algorithm determines the accuracy of recognizing changes in area coverages across different Land-use/Land-cover (LULC) classes. To improve the accuracy of Pixel-Based and Object-Based LULC classification, GEE used auxiliary datasets. The image segmentation results for agriculture, woodland, built-up terrain, and water body were all satisfactory, with the exception of grassland (Qu *et al*., 2021) [32].

Pedosphere

The Pedosphere is the Earth's outermost layer, which interacts dynamically with the Biosphere and atmosphere. Monitoring and studying the Pedosphere and its associated categories (e.g., soil, geology, and geomorphology) are prerequisites for sustainable development, particularly in the context of climate modelling. Because RS datasets are available in GEE, it is an appealing platform for Pedosphere studies at various scales. GEE was used for digital soil mapping, geology and mining studies, geomorphology studies, soil topography mapping, soil moisture derivation and estimation of soil carbon and salinity (Amani *et al*., 2020) [2].

Atmosphere and Climate

Constant population growth and human activities cause significant changes in the constituents of the atmosphere (Yeh and Liao, 2017) [44]. Climate change and air pollution are two significant outcomes of these disturbances, both of which have a direct impact on the surrounding environment and human health (Orimoloye *et al*., 2019) [27]. To avoid severe consequences, it is critical to monitor and control air quality and climate conditions. GEE is a great platform for climate studies and air quality monitoring due to the availability of climate and surface products. Because of these benefits, GEE is also used for air pollution analysis, climate change and monitoring, biophysical variable studies, evapotranspiration estimation and precipitation mapping. For instance, the GEE was used to track a sizable changing wetland in Newfoundland, Canada, using time series Landsat data, in the current context of climate change. According to a change detection study, bog, swamp, and fen are the most prevalent wetland types across all time periods, while marsh wetlands are the least prevalent (Mahdianpari *et al*., 2020) [22].

Land cover

The distribution of land covers characterizes the physical interaction between the Earth's surface and its surroundings. Recognizing the significant environmental impacts of land covers and investigating current conditions, as well as monitoring long-term dynamics of land covers, are critical for sustainable development, climate change modelling, biodiversity studies, and natural resource monitoring (Murray *et al*., 2018) [24]. GEE hosted massive publicly available RS datasets with varying spectral and spatial resolutions for land cover mapping, land cover dynamics monitoring, coastal mapping, and wetland classification. The cloud computing platform facilitated the computation of multi-seasonal Landsat spectral vegetation indices as well as illumination normalization algorithms, resulting in successful land cover classification results (He *et al*., 2018) [13].

Urban

Urban areas are areas with concentrated people and human infrastructure that typically expand over time in search of a better way of life. As a result, urban areas may be regarded as the primary site of human interaction with the environment. GEE encourages long-term monitoring of urban conditions in order to effectively study the urban environment from various perspectives. Some of the major urban studies conducted within GEE include urban expansion and extent mapping, urban morphology and local climate zone monitoring, urban green space classification, urban temperature and urban heat island identification. Sand dunes are a major environmental problem that primarily depends on cities, transportation, and population. Sand dune risk was assessed in Libya's Sabha region using Google Earth Engine, Landsat 8 and MODIS images. Furthermore, the author proposed that GEE is an extremely valuable technique for analyzing sand dune risk, and that it can be utilized for land management. (Pradhan *et al*., 2018) [31].

Others

Aside from the previously mentioned applications, there are numerous articles about other GEE applications. These researches are primarily concerned with archaeology, 3-D printing, wildlife, oil platform detection, and crashed Aeroplan detection. GEE, for example, has been reported as a suitable platform for processing high-resolution drone imagery for the identification of pottery shreds. Texture and gradient features from RGB drone imagery were calculated within GEE and ingested into the RF classifier in this regard. The developed algorithm identified pottery shreds with 32.9 percent and 76.8 percent accuracy in two different regions (Orengo and Molsosa, 2019) [26].

Acknowledgement

The authors gratefully acknowledge financial support from the Centre for Agricultural Market Intelligence, NAHEP-CAAST, Anand Agricultural University, Anand under National Agricultural Higher Education Project, Indian Council of Agricultural Research, Pusa, Krishi Anusandhan Bhawan-II, New Delhi-110012.

References

1. Aghababaei M, Ebrahimi A, Naghipour AA, Asadi E, Verrelst J. Vegetation Types Mapping using Multi-Temporal Landsat Images in the Google Earth Engine Platform, Remote Sensing. 2021;13(12):46-83.
2. Amani M, Ghorbanian A, Ahmadi SA, Kakooei M. Google Earth Engine Cloud Computing Platform for Remote Sensing Big Data Applications: A Comprehensive Review, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 2020;(13):5326-5350.
3. Amani M, Mahdavi S, Afshar M, Brisco B, Huang W, Mirzadeh SMJ, *et al*. Canadian wetland inventory using Google Earth Engine: The first map and preliminary results, Remote Sensing. 2019;11(7):842.
4. Bayissa Y, Tadesse T, Demisse G, Shiferaw A. Evaluation of satellite-based rainfall estimates and application to monitor meteorological drought for the Upper Blue Nile Basin. Remote Sensing. 2017;9(7):669.
5. Breiman L, Friedman JH, Olshen RA, Stone CJ. Classification and Regression Trees, Boca Raton: Chapman & Hall/CRC; c1984.
6. Chi M, Plaza A, Benediktsson JA, Sun Z, Shen J, Zhu Y.

Big data for remote sensing: Challenges and opportunities, Proceedings of the IEEE. 2016;104(11):2207-2219.
7. Chova LG, Lopez JA, Garcia GM, Mari JM, Valls GC. Cloud masking and removal in remote sensing image time series, Journal of Applied Remote Sensing, 2017, 11(1).
8. Coltin B, McMichael S, Smith T, Fong T. Automatic boosted flood mapping from satellite data, International Journal of Remote Sensing. 2016;37(5):993-1015.
9. Dong J, Xiao X, Menarguez MA, Zhang G, Qin Y, Thau D, *et al*. Mapping paddy rice planting area in northeastern Asia with Landsat 8 images, phenology-based algorithm and Google Earth Engine, Remote Sensing Environment. 2016;(185):142-154.
10. Garcia GM, Chova LG, Lopez JA, Mari JM, Valls GC. Multi-temporal cloud masking in the Google Earth Engine, Remote Sensing, 2018, 10(7).
11. Gorelick N, Hancher M, Dixon M, Ilyushchenko S, Thau D, Moore R. Google earth engine: Planetary-scale geospatial analysis for everyone, Remote Sensing of Environment. 2017;(202):18-27.
12. Hansen MC, Potapov PV, Moore R, Hancher M, Turubanova SA, Tyukavina A. High-resolution global maps of 21st-century forest cover change, Science. 2013;(342):850-853.
13. He M, Kimball JS, Maneta MP, Maxwell BD, Moreno A, Begueria S, *et al*. Regional Crop Gross Primary Productivity and Yield Estimation Using Fused Landsat-MODIS Data, Remote Sensing, 2018, 10(372).
14. Hsu CW, Chang CC, Lin CJ. A Practical Guide to Support Vector Classification; c2003.
15. Inoue S, Ito A and Yonezawa C. Mapping Paddy Fields in Japan by Using a Sentinel-1 SAR Time Series Supplemented by Sentinel-2 Images on Google Earth Engine, Remote Sensing. 2020;12(10):16-22.
16. Kandekar VU, Pande CB, Rajesh J, Atre AA, Gorantiwar SD, Kadam SA, *et al*. Surface water dynamics analysis based on sentinel imagery and Google Earth Engine Platform: a case study of Jayakwadi dam, Sustainable Water Resources Management, 2021, 7(44).
17. Kong D, Zhang Y, Gu X, Wang D. A robust method for reconstructing global MODIS EVI time series on the Google Earth Engine, ISPRS Journal of Photogrammetry and Remote Sensing. 2019;(155):13-24.
18. Kumar L, Mutanga O. Google Earth Engine applications since inception: Usage, trends and potential, Remote Sensing, 2018, 10(10).
19. Kumari N, Srivastava A, Dumka UC. A Long-Term Spatiotemporal Analysis of Vegetation Greenness over the Himalayan Region using Google Earth Engine, Climate, 2021, 9(109).
20. Liu P. A survey of remote-sensing big data. Frontier in Environmental Science, 2015, 3(45).
21. Lobell D, Thau D, Seifert C, Engle E, Little B. A scalable satellite-based crop yield mapper. Remote Sensing of Environment. 2015;(164):324-333.
22. Mahdianpari M, Jafarzadeh H, Granger JE, Mohammadimanesh F, Brisco B, Salehi B, *et al*. A large-scale change monitoring of wetlands using time series Landsat imagery on Google Earth Engine: a case study in Newfoundland, Gi-Science and Remote Sensing. 2020;57(8):1102-1124.
23. McArthur JW, Sachs JD. Agriculture, aid, and economic growth in Africa, World Bank Economic Review. 2019;33(1):1-20.
24. Murray NJ, Keith DA, Simpson D, Wilshire JH, Lucas RM. Remap: An online remote sensing application for land cover classification and monitoring, Methods in Ecology and Evolution. 2018;9(9):2019-2027.
25. Oliphanta AJ, Thenkabaila PS, Teluguntlaa P, Xionga J, Gummac MK, Congaltond RG, *et al*. Mapping cropland extent of Southeast and Northeast Asia using multi-year time-series Landsat 30-m data using a random forest classifier on the Google Earth Engine Cloud, International Journal of Applied Earth Observation and Geo-information. 2019;(81):110-124.
26. Orengo HA, Molsosa AG. A brave new world for archaeological survey: Automated machine learning-based potsherd detection using high-resolution drone imagery, Journal of Archaeological Science, 2019, (112).
27. Orimoloye IR, Mazinyo SP, Kalumba AM, Ekundayo OY, Nel W. Implications of climate variability and change on urban and human health: A review, Cities. 2019;(91):213-223.
28. Patel N, Angiuli E, Gamba P, Gaughan A, Lisini G, Stevens F, *et al*. Multi-temporal settlement and population mapping from Landsat using Google earth engine, International Journal of Applied Earth Observation and Geo-informatics. 2015;(35):199-208.
29. Pekel JF, Cottam A, Gorelick N, Belward AS. High-resolution mapping of global surface water and its long-term changes, Nature. 2016;(540):418-422.
30. Pelletier C, Valero S, Inglada J, Champion N, Dedieu G. Assessing the Robustness of Random Forests to Map Land Cover with High Resolution Satellite Image Time Series over Large Areas. Remote Sensing of Environment. 2016;(187):156-168.
31. Pradhan B, Moneir AAA, Jena R. Sand dune risk assessment in Sabha region, Libya using Landsat 8, MODIS, and Google Earth Engine images, Geomatics, Natural Hazards and Risk. 2018;9(1):1280-1305.
32. Qu L, Chen Z, Li M, Zhi J, Wang H. Accuracy Improvements to Pixel-Based and Object-Based LULC Classification with Auxiliary Datasets from Google Earth Engine, Remote Sensing. 2021;(13):453.
33. Sazib N, Mladenova I, Bolten J. Leveraging the Google Earth Engine for Drought Assessment using Global Soil Moisture Data, Remote Sensing. 2018;(10):12-65.
34. Silva Junior CAD, Leonel Junior AHS, Rossia FS, Filhob LFC, Santiagoc DDB, Oliveira Junior JFD, *et al*. Mapping soybean planting area in Midwest Brazil with remotely sensed images and phenology-based algorithm using the Google Earth Engine platform, Computers and Electronics in Agriculture. 2020;(169):105-194.
35. Smith B, Warlind D, Arneth A, Hickler T, Leadley P, Siltberg J, *et al*. Implications of incorporating N cycling and N limitations on primary production in an individual-based dynamic vegetation model, Biogeosciences. 2014;11(7):2027-2054.
36. Sturrock HJ, Cohen JM, Keil P, Tatem AJ, Le Menach A, Ntshalintshali NE, *et al*. Fine-scale malaria risk mapping from routine aggregated case data, Malaria Journal. 2014;(13):421.
37. Tamiminia H, Salehi B, Mahdianpari M, Quackenbush L, Adeli S, Brisco B. Google Earth Engine for geo-big data

applications: A meta-analysis and systematic review, ISPRS Journal of Photogrammetry and Remote Sensing. 2020;(164):152-170.
38. Tipping ME. Bayesian inference: An introduction to principles and practice in machine learning. In O. Bousquet, U. von Luxburg and G. Rätsch (Eds.), Advanced Lectures on Machine Learning; c2004. P.41-62.
39. Tipping ME, Faul AC. Fast marginal likelihood maximization for sparse Bayesian models. In Bishop CM and Frey BJ (Eds.), Proceedings of the Ninth International Workshop on Artificial Intelligence and Statistics, Key West, FL; c2003 Jan.
40. Wang X, Xiao X, Zou Z, Chen B, Ma J, Dong J, *et al*. Tracking annual changes of coastal tidal flats in China during 1986-2016 through analyses of Landsat images with Google Earth Engine. Remote Sensing of Environment. 2020;(238):110-987.
41. Xia H, Zhao J, Qin Y, Yang J, Cui Y, Song H, *et al*. Changes in Water Surface Area during 1989-2017 in the Huai River Basin using Landsat Data and Google Earth Engine, Remote Sensing, 2019, 11(1824).
42. Xiong J, Thenkabail PS, Tilton JC, Gumma MK, Teluguntla P, Oliphant A, *et al*. Nominal 30-m Cropland Extent Map of Continental Africa by Integrating Pixel-Based and Object-Based Algorithms using Sentinel-2 and Landsat-8 Data on Google Earth Engine, Remote Sensing, 2016, 9(1065).
43. Yao F, Wang J, Wang C, Cretaux JF. Constructing long-term high-frequency time series of global lake and reservoir areas using Landsat imagery, Remote Sensing Environment, 2019, (232).
44. Yeh JC, Liao CH. Impact of population and economic growth on carbon emissions in Taiwan using an analytic tool STIRPAT, Sustainable Environment Research. 2017;27(1):41-48.
45. Zhang Q, Li B, Thau D, Moore R. Building a better urban picture: combining day and night remote sensing imagery, Remote Sensing. 2015;7(9):11887-11913.