



INTERNATIONAL CONFERENCE GIS-IDEAS 2023

PROCEEDINGS

Geospatial Integrated Technologies for Natural Hazards and Environmental Problems

HUNRE, Ha Noi, Viet Nam
07-09 November 2023



PUBLISHING HOUSE FOR SCIENCE AND TECHNOLOGY



INTERNATIONAL CONFERENCE

GEOSPATIAL INTEGRATED TECHNOLOGIES FOR NATURAL HAZARDS AND ENVIRONMENTAL PROBLEMS

**PUBLISHING HOUSE FOR SCIENCE AND TECHNOLOGY
HA NOI - 2023**

Organized by

Hanoi University of Natural Resources and Environment (HUNRE)
Osaka Metropolitan University (OMU)
The Japan - Viet Nam Geoinformatics Consortium (JVGC)

Co-organizers

Ha Noi University of Mining and Geology (Vietnam)
Japan Geotechnical Consultant Association
Japan Society of Geoinformatics

With the support and collaboration of

Department of National Remote Sensing (Vietnam)
Department of Survey, Mapping and Geographic Information (Vietnam)
Ho Chi Minh University of Natural Resources and Environment (Vietnam)
National Centre for Hydro-Meteorological Forecasting (Vietnam)
Vietnam Institute of Geosciences and Mineral Resources
Institute of Resources Geography, Ho Chi Minh (Vietnam)
Vietnam Disaster and Dyke Management Authority
Ho Chi Minh City Space Technology Application Center (Vietnam)
Institute of Marine Geology and Geophysics (Vietnam)
Nong Lam University, Ho Chi Minh (Vietnam)
Vietnam National University, Hanoi
Vietnam Academy for Water Resources
Vietnam National Forestry University
Association of Vietnam Geographers
Military Technology Academy (Vietnam)
Asian Institute of Technology (Thailand)
Naresuan University (Thailand)
Thuy Loi University (Vietnam)
University of Phayao (Thailand)
NextTech Company Vietnam
Everest Schools (Vietnam)

CONFERENCE FOUNDERS

Late Prof. Takashi FUJITA (Japan)

Dr. Nghiem Vu KHAI (Vietnam)

SYMPOSIUM CHAIRS

Prof. Dr. Huynh Thi Lan HUONG (Vietnam)

Prof. Venkatesh RAGHAVAN (Japan)

ORGANIZING SECRETARIES

Assoc. Prof. Dr. Le Thi TRINH (Vietnam)

Dr. Go Yonezawa (Japan)

CONFERENCE COORDINATORS

Ms. Tanyaluck CHANSOMBAT (Thailand)

Dr. Tatsuya NEMOTO (Japan)

Ms. Vu Thi Thuy NGAN (Vietnam)

Dr. Truong Xuan QUANG Vietnam)

STEERING COMMITTEE

Dr. Ho Dinh DUAN (Vietnam)
Dr. Duong Van HAI (Vietnam)
Prof. Dr. Tran Thanh HAI (Vietnam)
Assoc. Prof. Hoang Anh HUY (Vietnam)
Dr. Mai Van KHIEM (Vietnam)
Prof. Yasuyuki KONO (Japan)
Prof. Nguyen Kim LOI (Vietnam)
Prof. Truong Xuan LUAN (Vietnam)
Prof. Shinji MASUMOTO (Vietnam)
Prof. Alaa A MASOUD (Vietnam)
Prof. Muneki MITAMURA (Japan)
Dr. Doan Thi Tuyet NGA (Vietnam)
Dr. Lam Dao NGUYEN (Vietnam)
Assoc. Prof. Huynh QUYEN (Vietnam)
Dr. Trinh Hai SON (Vietnam)
Assoc. Prof. Nguyen Canh THAI (Vietnam)
Prof. Nitin TRIPATHI (Thailand)
Dr. Chu Hai TUNG (Vietnam)
Prof. Yasushi YAMAGUCHI (Japan)

WORKSHOP COMMITTEE

DAO Hoang Tung (Vietnam)
Natraj VADDADI (India)

SCIENTIFIC COMMITTEE

Dr. Tran Thi AN (Vietnam)
Dr. Tran Van ANH (Vietnam)
Dr. Vansarochana CHAIWIWAT (Thailand)
Dr. Sittichai CHOOSUMRONG (Thailand)
Dr. Nguyen Dai DONG (Vietnam)
Dr. Nguyen Duc HA (Vietnam)
Prof. Dr. Pham Hoang HAI (Vietnam)
Assoc. Prof. Dr. Nguyen Hai HOA (Vietnam)
Dr. Pham Thi HOA (Vietnam)
Assoc. Prof. Dr. Nguyen Thanh HUNG (Vietnam)
Assoc. Prof. Dr. Trinh Le HUNG (Vietnam)
Dr. Atsushi KAJIYAMA (Japan)
Prof. Dr. Vo Chi MY (Vietnam)
Dr. Tatsuya NEMOTO (Japan)
Dr. Sarawut NINSAWAT (Thailand)
Dr. Susumu NONOGAKI (Japan)
Assoc. Prof. Dr. Pham Quy NHAN (Vietnam)
Dr. Nguyen Quoc PHI (Vietnam)
Dr. Vinayaraj POLIYAPRAM (Japan)
Assoc. Prof. Dr. Nguyen An THINH (Vietnam)
Assoc. Prof. Dr. Le Thi TRINH (Vietnam)
Assoc. Prof. Dr. Pham Thi Mai THAO (Vietnam)
Assoc. Prof. Dr. Nguyen Tien THANH (Vietnam)
Assoc. Prof. Dr. Phi Truong THANH (Vietnam)
Dr. Pham Thi Mai THY (Vietnam)
Dr. Pham Hong TINH (Vietnam)
Assoc. Prof. Dr. Nguyen Nhu TRUNG (Vietnam)
Dr. Le Quang TUAN (Vietnam)
Dr. Nghiem Van TUAN (Vietnam)
Assoc. Prof. Dr. Ho Thanh VAN (Vietnam)
Assoc. Prof. Dr. Pham Thi Mai THAO (Vietnam)

TABLE OF CONTENT

1.	DETERMINATION OF THE OPTIMAL DENSITY CONTRAST IN THE SEAFLOOR DEPTH INVERSION FROM GRAVITY ANOMALIES USING THE GRAVITY-GEOLOGIC METHOD ON THE CENTRAL EAST SEA Nguyen Dinh Hai, Nguyen Van Sang, Tran Tuan Dung.....	1
2.	IMPACT OF URBANIZATION ON SEAWATER QUALITY IN QUANG NINH PROVINCE Nguyen Tran Dinh, Le Thanh Son, Nguyen Tran Dien, Vu Anh Tuan, Nguyen Hong Quang, Le Cao Khai	9
3.	FORECAST OF SUBMARINE LANDSLIDES RELATED TO THE REACTIVE PROBABILITY OF THE FAULT SYSTEM IN THE SOUTHEAST VIETNAM CONTINENTAL SHELF AND ADJACENT AREAS Tran Tuan Duong, Tran Tuan Dung, Nguyen Quang Minh, Tran Trong Lap.....	20
4.	APPLYING MACHINE LEARNING ALGORITHMS TO CLASSIFY FOREST COVER TYPE FROM SENTINEL 2 MSI DATA Trinh Le Hung, Tran Xuan Bien, Pham The Trinh, Le Van Phu.....	29
5.	EFFECTS OF THE STRONG EL NINO 2015-2016 ON THE SEA SURFACE SALINITY OVER SOUTHERN VIETNAM AS OBSERVED BY REMOTE SENSING DATA Le Van Thien.....	41
6.	RESEARCH THE APPLICATION OF TERRESTRIAL LASER SCANNING TECHNOLOGY FOR MINING MANAGEMENT (A CASE STUDY AT KHE CHAM MINE SITE, QUANG NINH PROVINCE) Nguyen Ba Dzung, Dang Tuyet Minh, Vu Quoc Lap	49
7.	APPLICATION FOR GIS TECHNOLOGY TO COMPARISON OF ORDINARY LEAST SQUARES AND GEOGRAPHICALLY WEIGHTED REGRESSION MODEL IN THE ASSESSMENT OF THE MARKET RESIDENCE LAND IN THUY VAN WARD, HUE CITY, VIETNAM Le Huu Ngoc Thanh, Nguyen Huu Ngu, Pham Thi Thao Hien, Duong Quoc Non.....	58
8.	QUANTITATIVE CORRELATION OF FRACTURE ORIENTATION DISTRIBUTION AND THEIR RELATIONSHIP WITH TECTONIC CHARACTERISTICS IN NAM DU ARCHIPELAGO AREA, KIEN HAI DISTRICT, KIEN GIANG PROVINCE, VIETNAM Phi Truong Thanh, Vu Ngoc Binh, Van Duc Tung, Nguyen Quang Minh, Do Manh Tuan, Vu Thi Hong Cam	67
9.	APPLICATION OF RANDOM FOREST ALGORITHM AND GOOGLE COLAB FOR LAND COVER CLASSIFICATION Hoa Thanh Thi Pham, Ngoc Quang Vu, Nghi Thanh Le, Nam Phuong Thi Doan.....	75

10.	ESTIMATION OF LAND SURFACE TEMPERATURE AND VEGETATION DRYNESS INDEX (TVDI) IN BAC BINH - BINH THUAN USING REMOTE SENSING IMAGES Linh Nguyen Thi Thuy, Minh Hoang Thi Nguyet, Linh Phung Thi	84
11.	COLORIZATION OF BLACK-AND-WHITE AERIAL PHOTOGRAPHS USING DEEP LEARNING FOR OBJECT-BASED IMAGE ANALYSIS LAND USE CLASSIFICATION Arunothai Waesonthea, Sarawut Ninsawat, Nitin Kumar Tripathi, Sanit Arunplod , Thantham Khamyai.....	99
12.	SPATIAL-TEMPORAL ANALYSIS OF NO2 CONCENTRATION IN BINH DUONG PROVINCE, VIETNAM Nguyen Thi Bich Ngoc, Pham Thi Mai Thy, Nguyen Le Tan Dat, Tran Thi An.....	109
13.	DEVELOPMENT OF WEB GIS FOR INVENTORY OF VIETNAM MARINE PROTECTED AREAS Van Ngoc Truc Phuong.....	117
14.	MONITORING SURFACE WATER BODIES CHANGES FROM SENTINEL-2A IMAGERY WITH MODIFIED NORMALIZED DIFFERENCE WATER INDEX: APPLICATION IN DALAT, LAM DONG, VIET NAM Trung Van Nguyen, Ha Thu Thi Le.....	125
15.	METHOD FOR FORECASTING LANDSLIDE RISKS AND IDENTIFYING LANDSLIDE FORMS FOR RESETTLEMENT AREAS OF SONLA HYDROPOWER PLANT Phung Vinh An, Nguyen Van Thang.....	135
16.	APPLICATION OF AHP MODEL TO ESTABLISH A LANDSLIDE PROBABILITY ZONING MAP IN A LUOI DISTRICT, THUA THIEN-HUE PROVINCE, VIETNAM Nguyen Thi Thuy Hanh, Quach Thi Chuc	146
17.	GEOSPATIAL OVERVIEW OF THE VEGETATION ENVIRONMENT SUGGESTED BY COMMUNITY LEVEL PLACE NAMES IN THE AREA OF NORTHEASTERN THAILAND AND NORTHERN CAMBODIA Nagata Yoshikatsu.....	160
18.	PRESENTING OF THE MILITARY TERRAIN ANALYSIS BETWEEN THAILAND AND MYANMAR'S HISTORICAL BATTLEFIELD, OVER 200 YEARS (THE 9 -ARMIES WAR) WITH GEO-INFORMATIC TECHNIQUES Kittitouch Naksri, Rangsak Ket-ord, Gistada Panumonwatee, Chaiwiwat Vansarochana	165
19.	OPTIMIZING BROILER HOUSE MANAGEMENT TO REDUCE COSTS AND REDUCE LOSSES WITH THE INTERNET OF THINGS TECHNOLOGY AND WIRELESS SENSOR NETWORK Sittichai Choosumrong, Rhutairat Hataitara, Kampanart Piyathamrongchai, Tossaporn Incharoen, Venkatesh Raghavan, Thanwamas Phasinam, Khongdet Phasinam.....	172
20.	DEVELOPMENT OF WATER LEVEL MEASURING SENSOR PROTOTYPE EQUIPMENT FOR ANALYSIS AND ALERTING FLOOD RISK SITUATIONS IN THE FORM OF A 3D SIMULATOR ON A WEB GIS APPLICATION Natima Udon, Kampanart Piyathamrongchai, Sittichai Choosumrong.....	179

21. IMPACTS OF URBANIZATION AND POPULATION GROWTH ON LAND COVER AND COASTLINE CHANGES BASED ON REMOTE SENSING AND GIS TECHNIQUE FOR VIETNAMESE COASTAL REGIONS
Le Thi Thu Hang, Nguyen Hong Quang, Vu Anh Tuan, Nguyen Manh Hung, Nguyen Thi Phuong Hao 186
22. THREE-DIMENSIONAL BUILDING MODEL USING DRONE POINT CLOUDS
Sawarin Lerk-u-suke, Phaisarn Jeefoo, Pornthep Rojanavas, Jirabhorn Chaiwongsai, Nakarin Chaikaew, Bowonsak Srisungsittisunti, Niti lamchuen, Wipop Paengwangthong, Boonsiri Sukpromsun, Watcharaporn Preedapirom Jeefoo, Pranorm Khruewan, Jiraporn Kulsoontornrat, Suchatri Prasomsuk, Chatpong Pachanaparn, Nootchararat Thawadee..... 198
23. WEBGL-BASED VISUALIZATION TOOL FOR 3D GEOLOGICAL STRUCTURES IN SHALLOW SUBSURFACE IN URBAN AREAS
Susumu Nonogaki, Tsutomu Nakazawa 208
24. BUILDING A 3D INFORMATION MODEL FOR UNDERGROUND COAL MINES
Pham Van Chung, Cao Xuan Cuong, Le Van Canh, Nguyen Quoc Long, Le Thi Thu Ha, Nguyen Viet Nghia 213
25. DETECTION OF LAND SURFACE TEMPERATURE CHANGE IN COAL MINING AREA USING REMOTE SENSING AND GIS TECHNIQUES - A CASE STUDY IN QUANG NINH PROVINCE, VIETNAM
Le Thi Thu Ha, Nguyen Van Trung..... 228
26. A COMBINATION OF ALOS-2, SENTINEL-1 IMAGERY FOR RAPID DEFORESTATION DETECTION IN VIETNAM
Ngo Duc Anh, Vu Anh Tuan, Truong Tuan Nghia..... 237
27. TSUNAMI EVACUATION SIMULATION USING MULTI-AGENT SYSTEM: A CASE STUDY OF WAJIMA CITY, ISHIKAWA PREFECTURE, JAPAN
Tatsuya Nemoto, Aoi Sato, Venkatesh Raghavan 248
28. REMOTE SENSING APPLICATION USING GOOGLE EARTH ENGINE PLATFORM TO ASSESS CROP BURN AREAS IN WINTER-SPRING RICE CROP IN THE MEKONG DELTA, VIETNAM
Tran Van Dung, Lam Dao Nguyen, Hoang Phi Phung, Dang Pham Bao Nghi, Phung Chi Sy, Pham Thi Mai Thy 253
29. SWASH MODEL APPROACH FOR FLOWS INDUCED BY LOW-FREQUENCY WAVES OVER WOODEN FENCES
Hoang Tung Dao, Ngan Vu Thi Thuy, Ngo Thi Thuy Anh 259
30. DEM GENERATION AND TOPOGRAPHIC CHANGES OF CENTRAL HANOI, VIETNAM
Go Yonezawa, Tatsuya Nemoto, Xuan Luan Truong, Susumu Nonogaki, Do Thi Hang, Muneki Mitamura, Venkatesh Raghavan 268

31. DISTRIBUTION, FATE AND ECOLOGICAL TOXICITY OF SOME HEAVY METALS IN SEDIMENT: A CASE STUDY FROM DAY RIVER DOWNSTREAM, VIETNAM
Nguyen Khanh Linh, Trinh Thi Tham, Kieu Thi Thu Trang, Trinh Thi Thuy, Nguyen Thi Linh Giang, Luu Duc Hai, Tu Binh Minh, Le Thi Trinh 274
32. MANGROVE DEGRADATION ASSESSMENT USING WORLDVIEW-2 IMAGERY FOR MEKONG DELTA, VIETNAM
Pham Hong Tinh, Tran Dang Hung, Richard A. MacKenzie, Truong Van Vinh, Bui Thanh Huyen, Mai Huong Lam, Nguyen Thi Hong Hanh 285
33. UTILIZATION OF BOREHOLE DATA FOR CORRELATION SYSTEM OF STRATA: A CASE STUDY OF HANOI, VIETNAM
Kenichi Sakurai, Go Yonezawa, Luan Xuan Truong, Tatsuya Nemoto, Shinji Masumoto.. 293
34. DROUGHT MONITORING USING MODIS DATA AND THE GOOGLE EARTH ENGINE PLATFORM, CASE STUDY IN DAK LAK PROVINCE
Nguyen Ngoc Anh, Hoang Ngoc Khac, Tran Thi Ngoc Lam, Vu Thi Thuy Ngan 301
35. THAILAND COVID-19 CASE FATALITY SPATIAL CLUSTER AND DIRECTION ANALYSIS
Athitaya Sakunmungmee, Tanyaluck Chansombat, Pathana Rachavong 309
36. APPLICATION OF REMOTE SENSING AND GIS TO ESTABLISH SURFACE TEMPERATURE MAP OF PHU THO PROVINCE
Tran Thi Ngoan..... 316
37. THE ROLE OF OROGRAPHIC EFFECTS ON HEAVY RAINFALL EVENT OVER CENTRAL VIETNAM IN OCTOBER 2021
Phong Le Van, Phong Nguyen Binh, Thuong Le Thi 322
38. INVESTIGATE SAR INTENSITY AND OPTICAL IMAGES TO RAPIDLY DETECT SMALL AND MEDIUM LANDSLIDES IN MU CANG CHAI DISTRICT YEN BAI PROVINCE
Xuan Quang Truong, Van Anh Tran, Cam Chi Nguyen, Manh Dat Truong, Chi Cong Nguyen, Thi Phuong Anh Dao, Thi Thanh Thuy Pham..... 333
39. REGIONAL POST-EARTHQUAKE DAMAGE ESTIMATION USING GAUSSIAN GEOSTATISTICAL SIMULATION
Sunanthacha Pholkerd, Pathana Rachavong, Tanyaluck Chansombat..... 341
40. BA LAT DELTA EVOLUTION IN RESPONSE TO CHANGING FLUVIAL SEDIMENT SUPPLY BY THE RED RIVER, VIETNAM
Nguyen Hao Quang, Ha Nam Thang, Nguyen Van An, Nguyen Thanh Luan..... 348
41. SPATIAL PATTERN ANALYSIS OF THE SPREAD OF SARS-COV-2 VARIANT IN HANOI CITY, VIETNAM
Thi Quynh Nguyen..... 357
42. A NOVEL APPROACH OF NEURAL NETWORKS AND USLE IN SMART SOIL EROSION MODELING, CASE STUDY IN SOUTHERN COASTAL OF VIET NAM
Tran Thi Hoa, Tran Thanh Ha, Luan Cong Doan, Thu Minh Phan, Hung Van Nguyen 369

43. ASSESSMENT OF THE DYNAMIC IMPACT OF DEBRIS FLOW ON STRUCTURE SAFETY: A CASE STUDY OF CHECK DAM IN VIETNAM Nguyen Chi Thanh, Vu Quoc Cong, Vu Le Minh, Do Van Chinh, Tran Thi Nga, Luyen Le Dieu Linh	380
COLLECTION OF ABSTRACTS.....	390

APPLICATION OF RANDOM FOREST ALGORITHM AND GOOGLE COLAB FOR LAND COVER CLASSIFICATION

Hoa Thanh Thi Pham^{1,2*}, Ngoc Quang Vu³, Nghi Thanh Le^{1,2}, Nam Phuong Thi Doan^{1,2}

¹Faculty of Geomatics and Land Administration,
Hanoi University of Mining and Geology, Vietnam

²Geomatics in Earth Sciences Research Group,
Hanoi University of Mining and Geology, Vietnam

³Engineering Faculty, University of Transport Technology, Vietnam

*Corresponding author. Email: phamthithanhhoa@humg.edu.vn

ABSTRACT

Land cover is important information for the management and monitoring environment including disaster assessment in terms of flood, erosion, forest fire, landslide, and so on. Nowadays, many machine learning methods have been used to map land cover. Scientists assessed Random Forest (RF) to have more advantages than other image classification methods in its accuracy, its flexibility, and it is an easy-to-use algorithm. Especially, the appearance of Google Colab has brought the potential for RF in land cover classification. Google Colab allows researchers to run Python code through the browser, which not only requires no installation to use or upgrade computer hardware but also provides access free of charge to computing resources. Therefore the research is to provide the approach RF with Google Colab environment for classifying land cover in Quang Binh province, Vietnam using sentinel - 2 image in August 2021. The results have indicated the ability of this research direction with an overall accuracy above 80 %.

1. INTRODUCTION

Land cover is the observed (bio)physical cover on the earth's surface (FAO, 1998). It is important information for the management and monitoring environment including disaster assessment in terms of flood, erosion, forest fire, landslide, and so on. Many methods have been used for land cover classification. Remote sensing is still an effective tool for studying land cover at large scales (Friedl et al., 2010). Many commonly used open-access satellite imagery for mapping land cover are Landsat (Yang and Lo, 2002; Yuan et al., 2005), Sentinel 2 (Majidi Nezhad et al., 2019),...

Today, technological advances in storage and processing have enabled the creation of innovative algorithms for remote sensing applications based on machine learning. One of the most widely used algorithms is Random Forest (RF) (Breiman, 2001). It is the result of the development of the decision trees algorithm. In the world, scientists have used the Random Forest (Alonso Martínez et al., 2021; Gislason et al., 2006; Horning, 2010; Jin et al., 2018; Tokar et al., 2018). (Tokar et al., 2018) used this algorithm for land cover classification of Landsat 8 and gave an overall accuracy of 84,6 % and Kappa 0,808 for classes including cloud, water, urban and grassland (accuracy more than 90 %), forests (accuracy nearly 75 %). (Kulkarni and Lowe, 2016) assessed Random Forest to have more advantages than other image classification methods (Maximum likelihood, Minimum distance, Decision trees, Neural networks and Support Vector

Machine) in accuracy, although they had the same training and testing sample.

In Vietnam, RF has started to be used recently (Le et al., 2022; Nguyen et al., 2018; Tran et al., 2021). (Tran et al., 2021) established a land cover map for the study area of Ca Mau, Vietnam with the application of RF and Sentinel 1 Radar, (Nguyen et al., 2018) approached RF in the R software environment for classifying with Landsat 8. Therefore, it can be seen that RF has brought the potential in land cover classification.

Especially, the appearance of Google Colab (GC) which runs entirely on the cloud platform, allows researchers to execute code for image processing through the browser and is especially well suited to machine learning (according to Google Research). However, only a few studies have used GC (Lilay and Taye, 2023) for mapping land cover. Therefore, the research aims to provide the approach RF with Google Colab environment for classifying land cover in Quang Binh province, Vietnam using Sentinel-2 images.

2. STUDY AREA AND MATERIAL

2.1 Study area



Figure 1. The study area.

Quang Binh province is in the North Central area of Vietnam with coordinates $16^{\circ}55'$ to $18^{\circ}05'$ North and $105^{\circ}37'$ to $107^{\circ}00'$ East. It borders Ha Tinh province on the North, Quang Tri province on the South, Laos on the West, and the East Sea on the East. The provincial terrain is narrow and sloping from the West to the East. The area is divided into specific zones: 85 % high mountainous area and hill in the West, while the remaining is plain and coastal sandy area in the East of the province. The seaside sand dunes belt is a natural dam that protects the land from ocean tides (according to Quang Binh Portal).

2.2 Data resources

The study used Sentinel-2 (S2) with cloud cover of less than 30 % in August 2021.

S2 Multispectral Instrument (MSI) is a European wide-swath, high-resolution, multi-spectral imaging mission. The full mission with twin satellites flying in the same orbit but phased at 180°, is designed to give a high revisit frequency of 5 days at the Equator. The orbital swath width is 290 km. S2 carries an optical instrument payload that samples 13 spectral bands: Four bands at 10 m, six bands at 20 m and three bands at 60 m spatial resolution (ESA). The characteristics of the S2 are presented in Table 1.

Tab 1. The characteristics of Sentinel-2.

Band Number	Pixel Size	Description
B1	60 meters	Aerosols
B2	10 meters	Blue
B3	10 meters	Green
B4	10 meters	Red
B5	20 meters	Red Edge 1
B6	20 meters	Red Edge 2
B7	20 meters	Red Edge 3
B8	10 meters	Near infrared (NIR)
B8A	20 meters	Red Edge
B9	60 meters	Water vapor
B10	60 meters	Shortwave infrared/Cirrus
B11	20 meters	Shortwave infrared 1 (SWIR 1)
B12	20 meters	Shortwave infrared 2 (SWIR 2)

Source: <https://sentinels.copernicus.eu>

3. METHODOLOGY

3.1 Google Colab

The study used Google Colab (GC) in land cover classification. Google Colab (short for Colaboratory) is a powerful cloud platform and is based on the Jupyter Notebook open-source environment. It is designed to help scientists work more efficiently, especially in machine learning fields.

Google Colab allows researchers to run Python code through the browser, which not only requires no installation to use or upgrade computer hardware but also provides access free of charge to computing resources. They are also the advantage of GC.

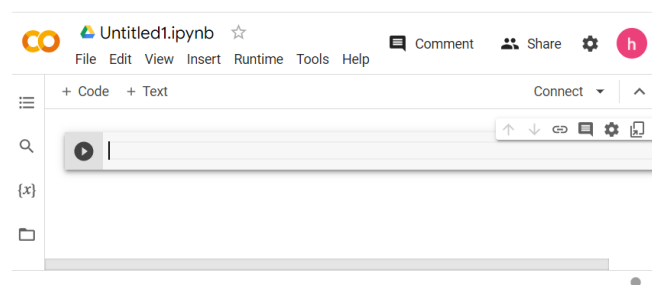


Figure 2. Google Colab Interface.

Google Colab integrates Python with a full of popular and pre-installed data science libraries for data analysis and visualization. Python is an interpreted, object-oriented, high-level programming

language. It is assessed as simple, easy to learn and mostly applies on many different platforms.

Generally, Google Colab is appropriate for people who need to work with multiple devices, for instance, computers or laptops, tablets, and even work in many places such as home, office,... because it has the ability to sync seamlessly between devices.

3.2 Random Forest

Random forest is one of the most popular supervised machine-learning algorithms which can be used to solve both classification and regression issues. Random forest is a combination of tree predictors such that each tree depends on the values of a random vector sampled independently and with the same distribution for all trees in the forest (Breiman, 2001). Theoretically, a forest obtains numerous trees, and the more trees more it will be stronger. Thus, one of the major advantages of RF is that it comprises a great number of trees leading to high accuracy and preventing the problem of overfitting. Feature bagging also makes the Random Forest classifier an effective tool for estimating missing values as it maintains accuracy when a portion of the data is missing (IBM).

With random forest classification, decision trees are created by using different random subsets of the given data and features. Each tree will provide its prediction to classify the data. RF is based on the majority votes of predictions and takes the most popular result to the final output.

The below diagram illustrates how the Random Forest algorithm works.

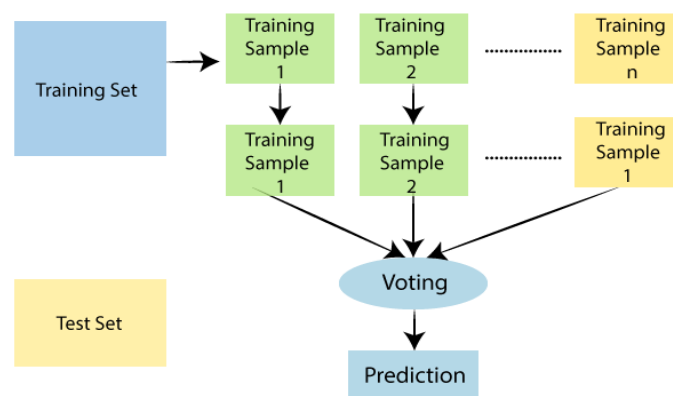


Figure 3. Random Forest.

Source: <https://www.javatpoint.com/machine-learning-random-forest-algorithm>

There are 4 steps:

Step 1: Select random samples from a given data

Step 2: Construct a decision tree for every sample and get the prediction result from every decision tree.

Step 3: Perform the votes for every predicted target

Step 4: Choose the most voted prediction result as the final prediction result.

Random forest algorithms have three main hyperparameters which need to be set before training. They are node size, the number of trees, and the number of features sampled.

3.3 Image Processing and Classification

Firstly, we installed and imported some libraries and maps in Google Colab such as Numpy,

Geemap,... Because Google Earth Engine platform has provided a large of remote sensing imagery, so we used it to import Sentinel-2 for this study.

```

# Time
START_DATE = ee.Date('2021-08-01');
END_DATE = ee.Date('2021-08-31');
MAX_CLOUD_PROBABILITY = 30;

# Study Area
provinceName = 'Quang Binh'
bound_vn = ee.FeatureCollection("projects/ee-thanhhoa/assets/gadm41_VNM_1")
province = bound_vn.filter(ee.Filter(ee.Filter.eq('VARIABLE_1', provinceName)))
Map.addLayer(province, {'color': 'grey', 'fillColor': '00000000'}, 'Study Area')
Map.centerObject(province, 8)

# Load Sentinel 2 data
s2Sr = ee.ImageCollection('COPERNICUS/S2_SR')
s2Clouds = ee.ImageCollection('COPERNICUS/S2_CLOUD_PROBABILITY')

#Vizualize
bands = ['B4', 'B3', 'B2'];
vizParams = {'bands': bands, 'min': 0, 'max': 2000, 'gamma': 1.3};

```

Figure 4. Set time and cloud coverage, select the study area and Sentinel-2.

Then, a set of Sentinel-2 images was filtered in August 2021 with less than 30 % cloud coverage (Figure 4). A cloud mask was applied using the “probability” band. The composite images were created by determining the median values of each pixel (in all bands) in the images. The final image was clipped within the study area boundary.

Besides, adding spectral indices can increase the information and improve the results of the classification (Praticò et al., 2021). Therefore, two indices NDVI (Normalized Difference Vegetation Index) and NDWI (Normalized Difference Water Index) were used in the study and added as bands in images. NDVI is a simple index but it can help recognize the difference between the forest and cropland (Tran et al., 2022). The NDVI is calculated as formula (1) (Tucker, 1979). While NDWI is used to highlight water bodies against the soil and vegetation: positive values for water features and negative ones (or zero) for soil and vegetation (McFeeters, 1996).

$$NDVI = \frac{NIR - RED}{NIR + RED} \quad (1)$$

$$NDWI = \frac{GREEN - NIR}{GREEN + NIR} \quad (2)$$

Where RED, GREEN and NIR stand for the spectral reflectance measurements acquired in the red, green and near-infrared bands.

We used Random Forest for land cover classification in Quang Binh. Sample data for 5 classes: Forest, Cropland, Fallow, Waterbody, and Build-Up were created based on the geodatabase of the study area, visually integrating information from Bing Map, Google Satellite, Sentinel-2 in QGIS software. They were divided randomly with 80 % for training and 20 % for validation. For each class, we assign a number to the property named ‘landcover’.

Figure 5 presents code for RF method with a tree number of 300 and bands used for the classification (band 2, band 3, band 4, band 8, NDVI and NDWI at 10m spatial resolution).


```

# Select bands for training
bands = ['B2', 'B3', 'B4', 'B8', 'NDVI', 'NDWI'];
### RF Classifier Model Building
# ee.Classifier.smileRandomForest(numberOfTrees, variablesPerSplit, minLeafPopulation, bagFraction, maxNodes, seed)
RFclassifier = ee.Classifier.smileRandomForest(300).train(**{
  'features': trainingsample,
  'classProperty': 'Class',
  'inputProperties': bands,
})
# Classify the image
Classified = S2_StudyArea.select(bands).classify(RFclassifier)

```

Figure 5. Random Forest code in Google Colab.

4. RESULTS AND DISCUSSION

Google Colab allows displaying results in its interface via Geemap.

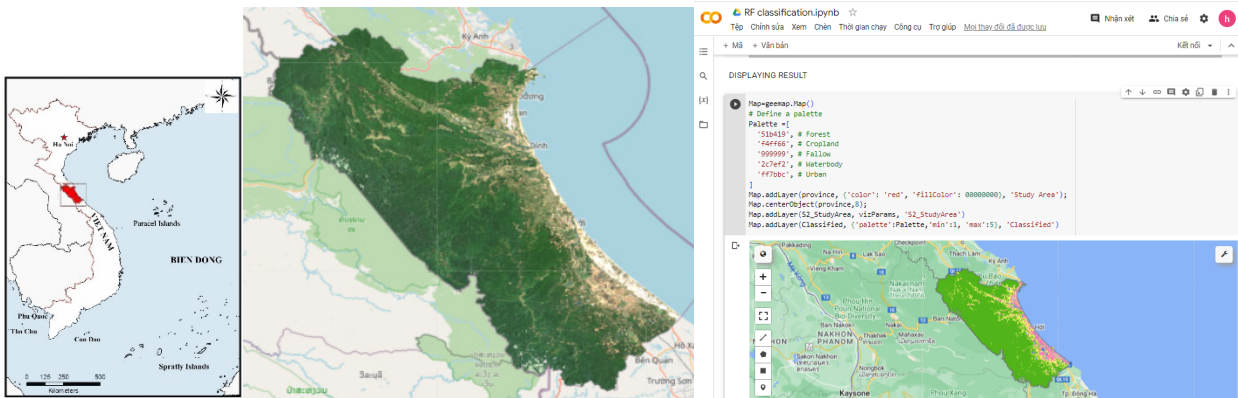


Figure 6. a) Sentinel-2 (True RGB color), b) Display results in Google Colab.

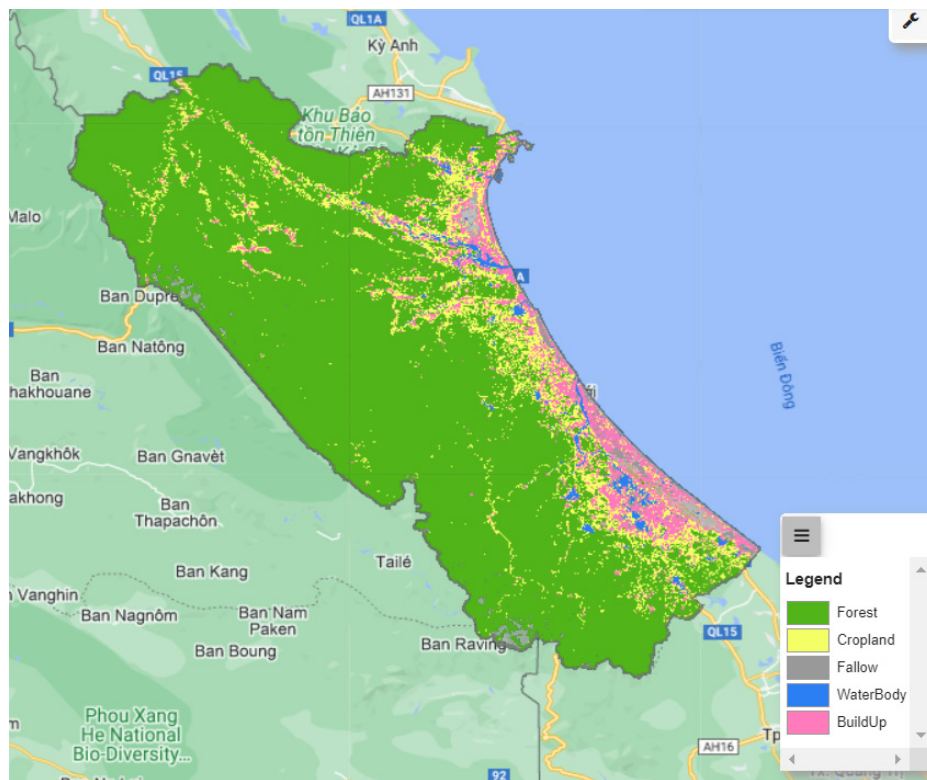


Figure 7. Land cover of Quang Binh province in Google Colab.

Figure 7 shows the land cover of the study area in August 2021 with 5 classes: Forest, Cropland (also known as Agricultural land), Fallow (including Sand, Barren Land), Waterbody, and Build-Up (also including Resident/homestead land). Generally, most of the forest area was concentrated in the west of Quang Binh province, and part of the forest was in the coastal area in the east. With the large forest area in the west, there is a famous forest Ke Bang within Phong Nha-Ke Bang National Park which was recognized as a World Natural Heritage Site by UNESCO in 2003. In addition, other land cover was scattered throughout the entire area and the eastern area. Fallow, especially sand, was distributed in the coastal area. The water bodies contain rivers such as Gianh, Nhat Le,... and numerous ponds.

Finally, the result of the classification is evaluated.

- Overall Accuracy (OA): OA is calculated by dividing the sum of the number of correct classifications (the diagonal of the matrix) by the total number of samples taken (Story et al., 1986).

- Producer's Accuracy (PA): is calculated by dividing the number of correctly classified reference points by the total number of reference points for that class.

It corresponds to the map accuracy from the point of view of the map maker. It represents how often real features on the ground are correctly shown on the classified map or the probability that a certain land cover of an area on the ground is properly classified (Alonso Martínez et al., 2021).

- User's Accuracy (UA): is computed by dividing the total number of classified pixels that agree with the reference data by the total number of classified pixels for that class. This value represents the reliability of the map or the probability that a pixel classified into a given category actually represents that category on the ground (Alonso Martínez et al., 2021).

- Kappa Coefficient: It is generated to evaluate the accuracy of classification. It compares the accuracy obtained in the classification to the accuracy that would be obtained randomly. A negative number indicates the classification is significantly worse than random. A value close to 1 indicates that the classification is significantly better than random.

Table 1 shows the results of the accuracy assessment including the Consumer's accuracy and the Producer's accuracy of each class in the land cover. Besides, the overall error (OA) = 86.7 %, Kappa = 0.8.

Table 2. Accuracy of Land cover classification result.

Class	UA (%)	PA (%)
Forest	75	88.2
Cropland	85.3	78.3
Fallow	81.8	99
Waterbody	75	99
Build-Up	94.2	89.1

The UA and PA of cropland were lower than others. The main reason is that, in the image of the study area, pixels distinguished between Cropland, Forest, and a part of Water were quite similar. Therefore, this study used NDVI to add in Random Forest algorithm, helping to improve the accuracy of results (above 75 %). Water bodies were recognized with high accuracy due to

adding NDWI. The OA above 80 % and Kappa coefficient above 0.8 show that the result of classification is generally good with Sentinel-2 and Random Forest.

Unlike other similar studies based on commercial software, Google Colab has brought the point of novelty of our research in this field. It is an effective tool because it allows users to run Python in the browser without the need for complex configuration, and the results were obtained in a short time.

5. CONCLUSION

In this study, based on the Sentinel-2 and Random Forest Algorithm, 5 classes of land cover were classified in Quang Binh province in 2021: Forest, Cropland, Fallow, Waterbody, and Build-Up. The accuracy of classification is generally high. Furthermore, the study also indirectly highlights how the possibility of exploiting the potential of Google Colab in this field. The problem is the time limit of free Google Colab sessions, the internet connection and the limited availability of data. Therefore, the improvement in this direction will bring outstanding results. Future research will apply different remote sensing imagery, more spectral indices, and other data such as DEM in land cover classification.

6. ACKNOWLEDGEMENTS

The study is supported and funded by Project No T23-41 of Hanoi University of Mining and Geology, Vietnam.

7. REFERENCES

- Alonso Martínez, L., Picos, J., and Armesto, J., 2021. Forest Land Cover Mapping at a Regional Scale Using Multi-Temporal Sentinel-2 Imagery and RF Models. *Remote Sensing*, 13, 2237. Doi:10.3390/rs13122237.
- Breiman, L., 2001. Random Forests. *Machine Learning*, 45(1), 5-32. Doi:10.1023/A:1010933404324.
- FAO, 1998. Land cover and Land use. *The FAO AFRICOVER Programme*.
- Friedl, M., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., and Huang, X., 2010. MODIS Collection 5 Global land cover: Algorithm refinements and Characterization of new datasets. *Remote Sensing of Environment*, 114, 168-182. Doi:10.1016/j.rse.2009.08.016
- Gislason, P. O., Benediktsson, J. A., and Sveinsson, J. R., 2006. Random Forests for land cover classification. *Pattern Recognition Letters*, 27(4), 294-300. Doi: <https://doi.org/10.1016/j.patrec.2005.08.011>
- Horning, N., 2010. Random Forests : An algorithm for image classification and generation of continuous fields data sets. *International Conference on Geoinformatics for Spatial Infrastructure Development in Earth and Allied Sciences 2010*.
- IBM, (2023). What is random forest? Retrieved from <https://www.ibm.com/topics/random-forest#:~:text=Random%20forest%20is%20a%20commonly,both%20classification%20and%20regression%20problems>. [Accessed on 12 June 2023].
- Jin, Y., Liu, X., Chen, Y., and Liang, X., 2018. Land-cover mapping using Random Forest classification and incorporating NDVI time-series and texture: a case study of central Shandong. *International Journal of Remote Sensing*, 39(23), 8703-8723. Doi:10.1080/01431161.2018.1490976.
- Kulkarni, A. D., and Lowe, B., 2016. Random Forest Algorithm for Land Cover Classification. *International*

- Le, H., Pham, L., Hoang, T., and Dinh, T., 2022. Land-cover classification using Random Forest and incorporating NDVI time-series and topography: a case study of Thanh Hoa province, Vietnam. *VNUHCM Journal of Earth Science and Environment*, 5(S3), 40-53. Doi: <https://doi.org/https://doi.org/10.32508/stdjsee.v5iSI2.681>.
- Lilay, M. Y., and Taye, G. D., 2023. Semantic segmentation model for land cover classification from satellite images in Gambella National Park, Ethiopia. *SN Applied Sciences*, 5(3), 76. Doi:10.1007/s42452-023-05280-4.
- Majidi Nezhad, M., Heydari, A., Fusilli, L., and Laneve, G. (2019). Land Cover Classification by using Sentinel-2 Images: A case study in the city of Rome. *Proceedings of the 4th World Congress on Civil, Structural and Environmental Engineering (CSEE'19)*.
- McFeeters, S. K., 1996. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *International Journal of Remote Sensing*, 17(7), 1425-1432. Doi:10.1080/01431169608948714.
- Nguyen, H. T. T., Doan, T. M., and Radeloff, V., 2018. Applying random forest classification to map land use/land cover using Landsat 8 OLI. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLII-3/W4, 363-367. Doi:10.5194/isprs-archives-XLII-3-W4-363-2018.
- Praticò, S., Solano, F., Di Fazio, S., and Modica, G., 2021. Machine Learning Classification of Mediterranean Forest Habitats in Google Earth Engine Based on Seasonal Sentinel-2 Time-Series and Input Image Composition Optimisation. *Remote Sensing*, 13, 586. Doi:10.3390/rs13040586.
- Story, M., Congalton, R. G. J. P. E., and Sensing, R., 1986. *Accuracy assessment: A user's perspective*, 52, 397-399.
- Tokar, O., Olena, V., Lubov, K., Havryliuk, S., and Korol, M., 2018. Using the Random Forest Classification for Land Cover Interpretation of Landsat Images in the Prykarpattya Region of Ukraine. *2018 IEEE 13th International Scientific and Technical Conference on Computer Sciences and Information Technologies (CSIT)*, 241-244. Doi:10.1109/STC-CSIT.2018.8526646.
- Tran, V. A., Le, M. H., Tran, H. H., Le, T. N., Tran, T. A., Nguyen, C. C., and Ha, T. K., 2022. Land cover mapping in Camau province by machine learning algorithms using Sentinel-2 imagery. *The 43th Asian Conference on Remote Sensing (ACRS 2022)*.
- Tran, V. A., Le, T. L., Tran, H. H., and Le, T. N., 2021. Monitoring Vegetation Cover Changes by Sentinel-1 Radar Images Using Random Forest Classification Method. *Journal of the Polish Mineral Engineering Society*, 1, 441-452.
- Tucker, C. J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment*, 8(2), 127-150. Doi: [https://doi.org/10.1016/0034-4257\(79\)90013-0](https://doi.org/10.1016/0034-4257(79)90013-0).
- Yang, X., and Lo, C. P., 2002. Using a time series of satellite imagery to detect land use and land cover changes in the Atlanta, Georgia metropolitan area. *International Journal of Remote Sensing*, 23(9), 1775-1798. Doi:10.1080/01431160110075802.
- Yuan, F., Sawaya, K. E., Loeffelholz, B. C., and Bauer, M. E., 2005. Land cover classification and change analysis of the Twin Cities (Minnesota) Metropolitan Area by multitemporal Landsat remote sensing. *Remote Sensing of Environment*, 98(2), 317-328. Doi: <https://doi.org/10.1016/j.rse.2005.08.006>.

PUBLISHING HOUSE FOR SCIENCE AND TECHNOLOGY
A16, 18 Hoang Quoc Viet Road, Cau Giay, Ha Noi
Marketing & Distribution Department: **024.22149040**;
Editorial Department: **024.37917148**
Administration Support Department: **024.22149041**
Fax: **024.37910147**, Email: nxb@vap.ac.vn; Website: www.vap.ac.vn

GIS IDEAS
PROCEEDINGS INTERNATIONAL CONFERENCE GIS-IDEAS 2023
GEOSPATIAL INTEGRATED TECHNOLOGIES
FOR NATURAL HAZARDS AND ENVIRONMENTAL PROBLEMS

Hanoi University of Natural Resources and Environment,
Vietnam 07-09 November 2023

Hanoi University of Natural Resources and Environment (HUNRE)
Osaka Metropolitan University (OMU)
Japan-Vietnam Geoinformatics Consortium (JVGC)

Responsible for Publishing

Director, Editor in Chief

PHAM THI HIEU

Editors: **Nguyen Van Vinh, Nguyen Thi Chien**
Ha Thi Thu Trang

Computing Technique: **Nguyen Duc Manh**

Cover design: **Tran Thu Hien**

Corporate publishing:

Hanoi University of Natural Resources and Environment

Address: 41A Phu Dien Road, Phu Dien Ward, Bac Tu Liem District, Ha Noi

ISBN: 978-604-357-207-0

Printing 400 copies, size 20,5 × 29,5 cm, printed at Consulting Publishing and Media Viet Joint Stock Company. Address: No. 4/20, Lane 156 Hong Mai, Bach Mai Ward, Hai Ba Trung District, Ha Noi City, Vietnam.

Registered number for Publication: 3759-2023/CXBIPH/01-42/KHTNVN.

Decision number for Publication: 87/QĐ-KHTNCN was issued on 15 December 2023.

Printing and copyright deposit were completed in the 4th quarter, 2023.



Organized by

Hanoi University of Natural Resources and Environment
The Japan - Vietnam Geoinformatics Consortium (JVGC)
Osaka Metropolitan University
<https://gis-ideas.org/2023/>

ISBN: 978-604-357-207-0



9 786043 572070

NOT FOR SALE