

NGHIÊN CỨU NỀN TẢNG GOOGLE EARTH ENGINE PHỤC VỤ XÂY DỰNG HỆ THỐNG XÁC ĐỊNH CÁC VỊ TRÍ SẠT LỎ ĐẤT TỪ DỮ LIỆU VIỄN THÁM ĐA THỜI GIAN

ĐINH BẢO NGỌC⁽¹⁾, NÔNG THỊ OANH⁽¹⁾, PHẠM QUANG HIỂN⁽¹⁾
TRẦN TUẤN NGỌC⁽²⁾, LÊ NGỌC XUYÊ⁽²⁾

⁽¹⁾Trường Đại học Mở Địa Chất

⁽²⁾Cục Viễn thám Quốc gia

Tóm tắt:

Sạt lở đất đang dần trở thành một trong những loại hình thiên tai nguy hiểm, không những đe dọa trực tiếp tới tính mạng con người mà còn ảnh hưởng trực tiếp tới môi trường và kinh tế xã hội. cần phải có một giải pháp hiệu quả để giảm thiểu mức độ thiệt hại do tai biến thiên nhiên, trong bài báo này nhóm tác giả đã đề xuất quy trình công nghệ và xây dựng hệ thống trên công nghệ WebGIS sử dụng Google Earth Engine để thu thập và phân tích dữ liệu trên ảnh viễn thám Sentinel 2 đa thời gian để xác định các vị trí sạt lở đất, đồng thời đưa ra các đánh giá hỗ trợ ước tính thiệt hại do sạt lở đất gây ra. Hệ thống được xây dựng trên các công nghệ nguồn mở như NodeJS, cơ sở dữ liệu lưu trữ trong hệ quản trị cơ sở dữ liệu PostgreSQL; các phép phân tích ảnh được thực hiện trên bộ thư viện của google earth engine. Hệ thống cho phép người sử dụng có thể lựa chọn vùng phân tích, khoảng thời gian và đưa ra kết quả là các khu vực sạt lở kèm báo cáo một cách nhanh chóng, tiết kiệm thời gian hơn so với với các bước xử lý truyền thống.

Từ khóa: Sạt lở đất, Google earth engine, Sentinel 2, Viễn thám.

1. Đặt vấn đề

Với đặc điểm nước ta có nhiều khu vực đồi núi nên các thiên tai như sạt lở, lũ quét tại các khu vực này thường xuyên diễn ra và gây ra hậu quả lớn về mặt kinh tế cũng như con người cho nhiều địa phương trên cả nước. Điều đó dẫn tới phải có một công cụ cho phép phát hiện, đánh giá thiệt hại và phân vùng sạt lở nhằm trợ giúp các nhà quản lý trong việc quy hoạch hạ tầng giao thông, xây dựng các khu dân cư cách xa các vùng có nguy cơ sạt lở cao.

Hiện nay có rất nhiều các nghiên cứu trợ

giúp việc phát hiện, phân vùng các khu vực có nguy cơ sạt lở; Viễn thám trở thành một công nghệ đang được nhiều nhà nghiên cứu sử dụng cho công việc trên. Với thể mạnh có thể giám sát các đối tượng trên bề mặt trái đất với diện tích rộng (210 km với ảnh Sentinel 2 cảnh ảnh), nhiều độ phân giải ảnh (0,5 m ảnh geoeye, 3 m với ảnh VNRed Sat1, 10 m đến 60 m với ảnh Sentinel), quan sát trong mọi điều kiện thời tiết (ảnh sentinel 1 SAR không bị ảnh hưởng bởi mây và mưa) cho phép viễn thám trở thành công cụ hữu ích trong việc cung cấp dữ liệu phục vụ xác định, phân vùng

các khu vực có nguy cơ sạt lở.

Với những thế mạnh như trên nhưng công nghệ viễn thám truyền thống vẫn tồn tại các vấn đề như mất rất nhiều thời gian trong việc thu thập dữ liệu và xử lý, phân tích dữ liệu bằng các phần mềm chuyên dụng. Để khắc phục các yếu điểm đó, Google đã cung cấp nền tảng Google Earth Engine, nền tảng này cho phép các nhà nghiên cứu có thể truy xuất một kho dữ liệu ảnh viễn thám với nhiều loại khác nhau như Landsat, Sentinel, MODIS, vv... với nguồn tư liệu mới nhất. Ngoài ra, Google Earth Engine còn cung cấp các bộ thư viện cho việc phân tích các dữ liệu ảnh bằng các thuật toán truyền thống hay các thuật toán học máy (machine learning) hoặc học sâu (deep learning) trợ giúp đắc lực cho các nhà nghiên cứu. Google Earth Engine còn cung cấp các API (Application Programming Interface) cho phép các nhà lập trình xây dựng các ứng dụng website dựa trên kho tư liệu và các thư viện xử lý ảnh sẵn có. Mặc dù vậy, Google Earth Engine vẫn tồn tại một số nhược điểm như hạn chế tích hợp các bộ thư viện phân tích không gian khác trên nền tảng của mình, điều đó dẫn tới khó khăn trong một số phân tích.

Đã có một số nghiên cứu sử dụng Google Earth Engine trong quan trắc và phát hiện sạt lở đất như: nghiên cứu xác định bản đồ mức độ nhạy cảm với sạt lở đất dựa vào kỹ thuật học máy Random Forest với hơn 1000 điểm lấy mẫu để xác định các vị trí có yếu tố nhạy cảm dẫn tới sạt lở đất với độ chính xác của kết quả xác định các điểm có nguy cơ sạt lở lên tới 94% (Hafsah Fatihul Ilymy, 2021). Một nghiên cứu khác cũng sử dụng các phương pháp học máy trong Google Earth Engine là random forest với nguồn dữ liệu đầu vào là ảnh Sentinel 1 và 2 kết hợp với dữ liệu địa

hình DEM cho ra bản đồ các điểm sạt lở với độ chính xác lên tới 85% (Teodora Mitroi, 2021). Các nghiên cứu trên cho thấy khả năng ứng dụng Google Earth Engine trong việc xây dựng hệ thống theo dõi sạt lở.

Với các nghiên cứu trên có nhược điểm là thiếu tính linh động về khoảng thời gian theo dõi sạt lở; tập trung phạm vi vào khu vực nhất định; tính ứng dụng của nghiên cứu chưa cao khi chỉ dừng lại ở mức độ đưa ra phương pháp và thực nghiệm tại một khu vực đồng thời đưa ra các báo cáo khuyến nghị; nguồn dữ liệu đầu vào của phân tích dựa chủ yếu từ kho tư liệu của Google chưa sử dụng đến các dữ liệu không gian từ phía người dùng như hiện trạng sử dụng đất, dân cư, vv... dẫn tới tính thời sự của việc thống kê, phân tích chưa cao; chưa cung cấp bộ công cụ theo dõi sạt lở cho các nhà quản lý.

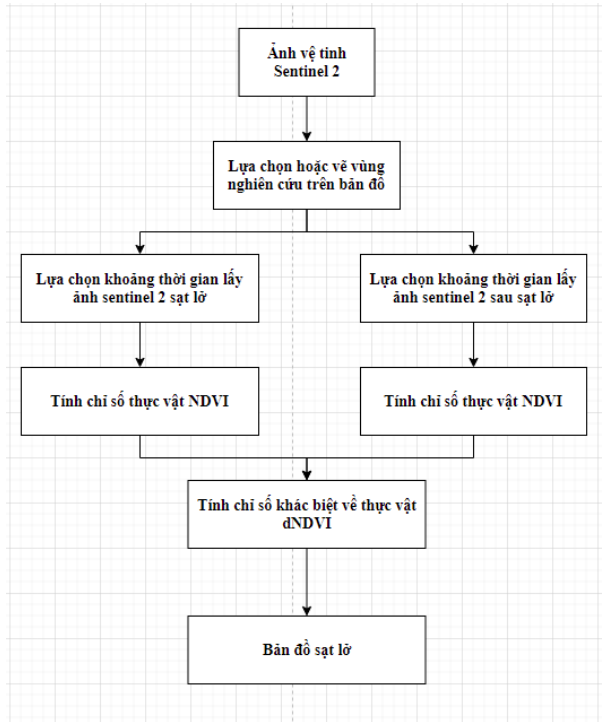
Từ những phân tích trên, nhóm tác giả đã đề xuất xây dựng hệ thống WebGIS cho phép phát hiện, đánh giá sạt lở đất trên nền tảng NodeJS cho phép tích hợp Google Earth Engine API với bộ thư viện TurfJS cho thu nhận, phân tích không gian trên ảnh vệ tinh. Dữ liệu không gian khác của hệ thống như các lớp thủy văn, giao thông hay dân cư, các dữ liệu sau khi phân tích được lưu trữ trong hệ quản trị PostgreSQL. Cần lưu trữ các dữ liệu sau khi phân tích do sạt lở khó dự đoán nhưng có tính lặp lại tại những vị trí đã từng xảy ra nên dữ liệu sẽ cho thấy những khu vực bị tiềm ẩn nguy cơ. Với cơ sở dữ liệu này kết hợp các phương pháp nghiên cứu truyền thống sẽ là nền tảng thiết yếu cho những nghiên cứu chuyên sâu về sạt lở, quy hoạch phát triển trong tương lai. Trong khi, các dữ liệu sông suối giúp xác định sạt lở do lũ quét hay tại các khu vực giao thông sườn núi.

2. Phương pháp nghiên cứu

2.1. Phương pháp theo dõi sạt lở

Việc theo dõi sạt lở được thực hiện trên hệ thống website thông qua Google Earth Engine

API. Phương pháp sử dụng để phát hiện sạt lở thông qua việc xác định sự khác biệt về chỉ số thực vật dNDVI.



Hình 1: Quy trình theo dõi sạt lở

Quy trình được thực hiện với dữ liệu ảnh sentinel 2A với các bước chính gồm: (1) Lựa chọn vùng nghiên cứu; (2) Lựa chọn ảnh trước và sau khi sạt lở; (3) tính toán chỉ số thực vật NDVI trên từng ảnh; (4) tính toán sự khác biệt về NDVI giữa 2 ảnh; (5) đưa ra bản đồ các vị trí sạt lở.

2.2. Xây dựng và thiết kế hệ thống theo dõi sạt lở trên nền tảng WebGIS sử dụng Google Earth Engine

Hệ thống được xây dựng trên môi trường Website với hệ quản trị cơ sở dữ liệu sử dụng là PostgreSQL. Lập trình và môi trường tích hợp Google Earth Engine API sử dụng thư viện Express của NodeJS; Việc trực quan hóa dữ liệu trên bản đồ sử dụng thư viện Mapbox JS.

Google Earth Engine API được sử dụng

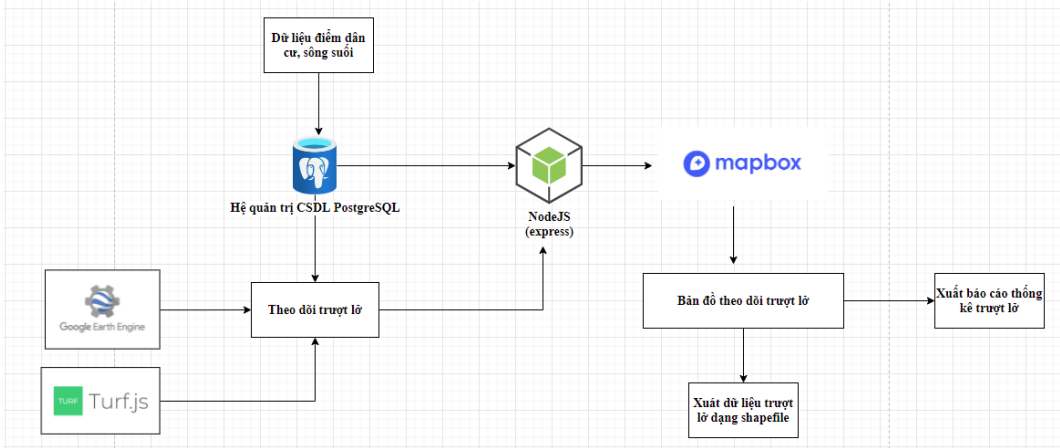
cho việc thu thập các nguồn dữ liệu từ google như ảnh vệ tinh Sentinel 2 theo khoảng thời gian người dùng lựa chọn; Dữ liệu về các điểm dân cư và sông suối được thu thập và đưa vào hệ thống dạng shapefile và được lưu trữ trong hệ quản trị cơ sở dữ liệu PostgreSQL.

Hệ quản trị CSDL PostgreSQL cho phép người dùng lưu trữ các dữ liệu không gian để sử dụng cho việc phân tích ảnh hưởng của sạt lở trên hệ thống

Hệ thống sử dụng NodeJS với thư viện express cho việc kết nối Google Earth Engine API với các bộ thư viện phân tích không gian như TurfJS phục vụ xác định vùng giao nhau ảnh hưởng của ngập lụt với dân cư hoặc đất nông nghiệp; Mapbox JS cho việc trực quan hóa dữ liệu bản đồ trên nền website.

Dữ liệu đầu vào của hệ thống:

STT	Tên dữ liệu	Định dạng
1	Ảnh Sentinel 2A	Geotiff từ Google Earth Engine
2	Điểm dân cư	Shapefile (point) do người dùng cung cấp
3	Sông, suối	Shapefile (polyline) do người dùng cung cấp



Hình 2: Sơ đồ hoạt động và các công nghệ của hệ thống

Hệ thống gồm các chức năng chính như sau:

(1) Thể hiện vị trí các điểm có khả năng do sạt lở trên nền ảnh vệ tinh sentinel 2A.

(2) Theo dõi sạt lở: Modul này có chức năng xác định các vùng sạt lở. Với đề tài ngoài xác định các vùng sạt lở còn phân loại sạt lở do lũ quét hay sạt lở do các yếu tố khác. Để xác định sạt lở do lũ quét đề tài sử dụng bộ thư viện TurfJS để phân tích không gian vị trí các vùng sạt lở có gần các khu vực sông suối hay không. Dữ liệu sau khi được phân tích sẽ cho phép người dùng tải về dưới dạng shapefile.

+ Dữ liệu các điểm dân cư được đưa vào dạng điểm (point) sau đó được xác định phạm vi bằng việc buffer điểm với bán kính 500 m,

việc xác định các điểm dân cư có bị ảnh hưởng bởi sạt lở hay không sẽ thông qua việc tính toán sự giao nhau giữa vùng sạt lở được phát hiện và vùng dân cư được buffer.

+ Việc xác định các vị trí sạt lở có phải do lũ lụt hay không được xác định như sau: Dữ liệu đầu vào của hệ thống là sông suối dạng đường (polyline) dạng shapefile sẽ được xác định có giao cắt với vùng sạt lở hay không. Nếu có giao cắt sẽ đưa ra vị trí các điểm sạt lở có khả năng do lũ lụt.

(3) Báo cáo, thống kê: cho phép xuất báo cáo thống kê diện tích các vị trí có khả năng sạt lở theo ranh giới hành chính cấp huyện và cho phép người dùng tải về dưới dạng file excel.

Dữ liệu ra của hệ thống:

STT	Tên dữ liệu	Mô tả
1	Bản đồ các vùng có nguy cơ sạt lở	Vị trí các vùng sạt lở trên ảnh vệ tinh
2	Vùng sạt lở	Shapefile (polygon) cho phép người dùng download về
3	Báo cáo thống kê	Excel thống kê diện tích vùng có khả năng do sạt lở theo ranh giới cấp huyện

3. Kết quả

Hệ thống cho phép xác định các vị trí trượt lở trên phạm vi các Tỉnh toàn lãnh thổ Việt

Nam. Tuy nhiên, trong bài báo này nhóm tác giả lấy dẫn chứng kết quả của hệ thống với tỉnh Hà Tĩnh.

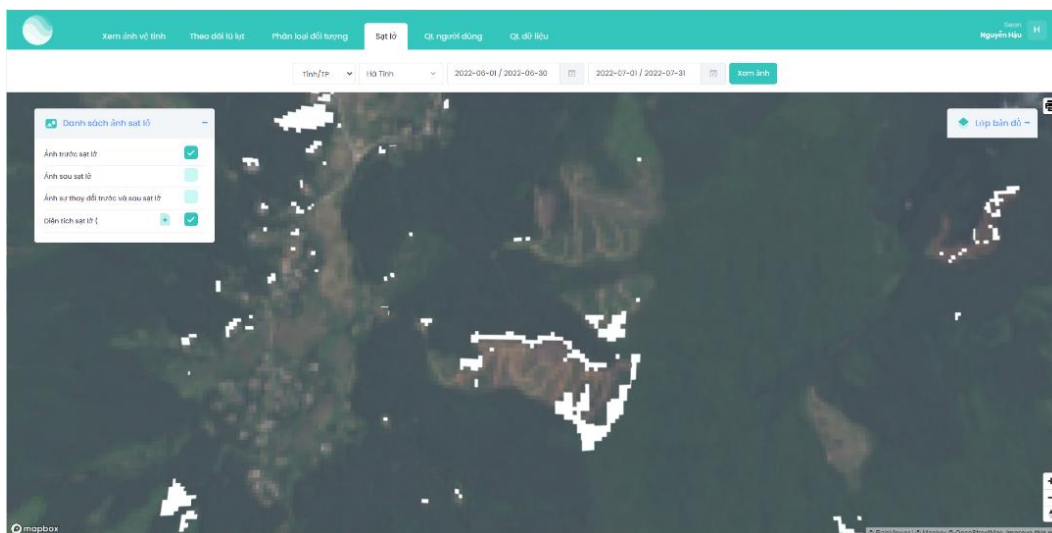
Dữ liệu đầu vào của hệ thống

	Tên dữ liệu	Định dạng
1	Ảnh Sentinel 2A	Khoảng thời gian lấy để tính toán: - Ảnh trước: từ 1 đến 30 tháng 6 năm 2022. - Ảnh sau: từ 1 đến 30 tháng 7 năm 2022.
2	Điểm dân cư	Shapefile (point) do người dùng cung cấp
3	Sông, suối	Shapefile (polyline) do người dùng cung cấp

Ảnh trước theo dõi sạt lở (1) Ảnh sau theo dõi sạt lở (2) Vị trí và khoanh vùng sạt lở (3)



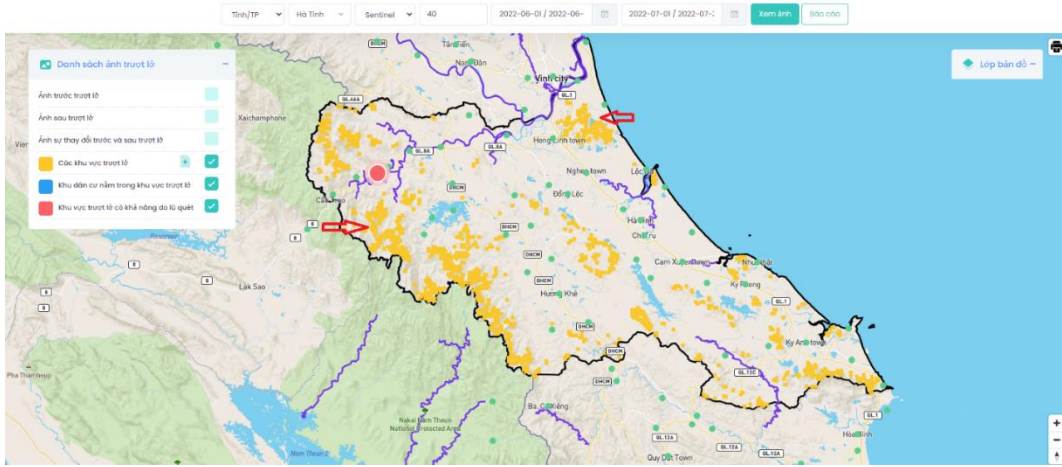
Hình 3: Ảnh vệ tinh trước, sau và kết quả phân vùng khi theo dõi sạt lở



Hình 4: Sau khi chạy chức năng phát hiện và theo dõi sạt lở

Kết quả chạy chức năng theo dõi sạt lở được thực hiện với tỉnh Hà Tĩnh với thời gian trước khi theo dõi sạt lở là tháng 6-2022 và sau

khi theo dõi sạt lở 7-2022 với các vùng màu trắng trên ảnh số 3, hình 3 thể hiện các vùng sạt lở tại khu vực theo dõi.



Hình 5: Bản đồ vị trí các vùng có khả năng sạt lở

Các điểm có khả năng sạt lở được xác định bởi các vùng màu vàng trên bản đồ, các vùng được chỉ bởi các mũi tên màu đỏ. Điểm chấm đỏ trên bản đồ là khu vực sạt lở có khả

năng do lũ gây ra. Dữ liệu sau khi được phân tích sẽ được kết xuất thành báo cáo và cho phép người dùng có thể tải về dưới dạng file excel như hình bên dưới.

STT	Tên huyện	Diện tích trượt lở (HA)	Tên tỉnh
1	Hương Sơn	62682084	Hà Tĩnh
2	Kỳ Anh	42066040	
3	Nghi Xuân	17185894	
4	Thạch Hà	8865079	
5	Vũ Quang	42405599	
6	Đức Thọ	0	
7	Can Lộc	12317980	
8	Cẩm Xuyên	13848254	
9	Hà Tĩnh	0	
10	Hồng Lĩnh	0	
11	Hương Khê	70629352	

Hình 6: Chức năng báo cáo thống kê các vùng có khả năng sạt lở theo ranh giới Huyện (Bảng thống kê trượt lở tỉnh Hà Tĩnh giai đoạn từ tháng 6 đến tháng 7 năm 2022)

4. Kết luận

Nghiên cứu đã đưa ra giới thiệu phương pháp sử dụng Google Earth Engine để trợ giúp theo dõi sạt lở trên tư liệu ảnh vệ tinh Sentinel 2. Nghiên cứu này đã xây dựng hệ thống cho phép người dùng có thể thực hiện các theo dõi tại nhiều khu vực khác nhau trên lãnh thổ Việt

Nam tại các khoảng thời gian khác nhau, ngoài ra dữ liệu sau khi phân tích cũng được hệ thống cho phép người dùng tải về hoặc lưu trữ vào cơ sở dữ liệu phục vụ cho các phân tích chuyên sâu hoặc làm cơ sở so sánh, đối chiếu với thực địa.

Nghiên cứu còn tồn tại một số hạn chế

nhu chức năng giám sát sạt lở chưa được đánh giá và sẽ được đề tài thực hiện trong giai đoạn vận hành và so sánh với dữ liệu thu thập từ thực địa (công việc này được thực hiện trong một phần của đề tài “Nghiên cứu, xây dựng công nghệ trích xuất tự động thông tin bề mặt vùng ảnh hưởng của bão, ngập lụt, sạt lở đất đá, lũ bùn đá và lũ quét từ dữ liệu viễn thám phục vụ dự báo, cảnh báo thiên tai” mã số TNMT.2022.02.22).○

Lời cảm ơn

Các tác giả ghi nhận những đóng góp của đề tài mã số TNMT.2022.02.22 đã cung cấp dữ liệu thực nghiệm giúp hoàn thiện nghiên cứu.

Tài liệu tham khảo

[1]. Hafsah Fatihul Ilmy, M. R. (2021). Application of Machine Learning on Google Earth Engine to Produce Landslide

Susceptibility Mapping (Case Study: Pacitan). *IOP Conference Series: Earth and Environmental Science*.

[2]. Muhammad Fawad Akbar Khan, K. M. (2021). Mapping Allochemical Limestone Formations in Hazara, Pakistan Using Google Cloud Architecture: Application of Machine-Learning Algorithms on Multispectral Data. *ISPRS International Journal of Geo-Information*.

[3]. Teodora Mitroi, K. F. (2021). *AGU Fall Meeting*. From Identifying landslides using SAR, DEMs, optical imagery, spatial texture and machine learning in Google Earth Engine: <https://agu.confex.com/agu/fm21/meetingapp.cgi/Paper/960785>.○

Summary

Research on the Google Earth Engine platform to build a system to identify landslide locations from multi-temporal remote sensing data

Dinh Bao Ngoc, Nong Thi Oanh, Pham Quang Hien

Hanoi University of Mining and Geology

Tran Tuan Ngoc, Le Ngoc Xuyen

Department of National Remote Sensing

Landslides are gradually becoming one of the most dangerous types of natural disasters, not only directly threatening human life but also directly affecting the environment and socio-economy. There needs to be an effective solution to minimise the damage caused by natural disasters. In this article, the authors have proposed a technological process and built a system based on WebGIS technology using Google Earth Engine to collect and analyse data on multi-temporal Sentinel 2 remote sensing images to identify landslide locations and provide assessments to support damage estimates caused by landslides. The system is built on open-source technologies such as NodeJS, a database stored in the PostgreSQL database management system; Image analyses are performed on the Google Earth engine library. The system allows users to select the area and time to analyse and quickly produce results of landslide areas with reports, saving more time than traditional processing steps.○

Keywords: landslide, Google Earth Engine, Sentinel 2, remote sensing.