

## ỨNG DỤNG KHOA HỌC CÔNG NGHỆ PHÂN TÍCH DỮ LIỆU LỚN (BIG DATA) HƯỚNG TỚI VIỆC BẢO TỒN CÁC DI SẢN ĐỊA CHẤT TẠI KHU VỰC TRÙNG KHÁNH, TỈNH CAO BẰNG

Trịnh Ngọc Như Ánh<sup>1</sup>, Nguyễn Quốc Phi,  
Phan Thị Mai Hoa, Nguyễn Thị Mai Hương  
Khoa Môi trường, Trường ĐH Mỏ - Địa chất.

---

### Tóm tắt:

Khu vực Trùng Khánh, tỉnh Cao Bằng là nơi tập trung các di sản địa chất và văn hóa với nhiều kỳ quan, danh lam thắng cảnh nổi tiếng. Tuy nhiên, đây cũng là một trong những địa phương miền núi thường xuyên xảy ra các tai biến môi trường, đặc biệt là hiện tượng trượt lở, lũ quét, có khả năng ảnh hưởng tới việc bảo tồn các di sản địa chất tại khu vực. Bài báo phân tích các nguồn thông tin khai thác được từ nguồn tư liệu ảnh viễn thám đa thời gian, dữ liệu địa hình, khí tượng thủy văn và hoạt động nhân sinh khác... có ảnh hưởng đến các di sản địa chất tại khu vực nghiên cứu sử dụng các phương pháp phân tích dữ liệu lớn. Kết quả phân tích sử dụng các phương pháp phân tích dữ liệu lớn cho thấy, mô hình mạng trí tuệ nhân tạo (ANN) cho kết quả tốt nhất với độ chính xác toàn cục đạt 89,3%; tiếp theo là phương pháp vectơ hỗ trợ (đạt 87,4%) và cuối cùng là mô hình cây quyết định (đạt 83,6%). Kết quả xây dựng bản đồ các khu vực có nguy cơ xảy ra tai biến môi trường giúp xác định được các điểm di sản địa chất nằm trong các vùng có nguy cơ cao cần được bảo vệ. Các bản đồ phân vùng nguy cơ tai biến môi trường không chỉ áp dụng cho các điểm di sản tại khu vực nghiên cứu mà còn có thể áp dụng cho các khu vực tương tự, hướng tới mục tiêu quản lý và phát triển bền vững tài nguyên thiên nhiên cần được bảo tồn.

**Từ khóa:** Dữ liệu lớn; Di sản địa chất; Bảo tồn di sản; Bảo tồn thiên nhiên; Bảo vệ môi trường.

**Mã số:** 22111801

## APPLICATION OF SCIENCE AND TECHNOLOGY OF ANALYSIS OF BIG DATA TO CONSERVE GEOLOGICAL HERITAGES IN TRUNG KHANH AREA, CAO BANG PROVINCE

### Abstract:

Trung Khanh area, Cao Bang province is a place with geological and cultural heritages with many wonders and famous scenic spots. However, this is also one of the mountainous localities where environmental disasters often occur, especially landslides and flash floods, which are likely to affect the conservation of geological heritages in the area. The article analyzes information sources extracted from multi-temporal remote sensing images, topographical data, hydrometeorology, and other human activities,... in the study area

---

<sup>1</sup> Liên hệ tác giả: anhtn168@gmail.com

*using big data analysis methods. The analysis results using big data analysis methods show that the ANN artificial intelligence network model gives the best results with the global accuracy reaching 89.3%; followed by the support vector method (87.4%) and finally the decision tree model (83.6%). The results of mapping areas at risk of environmental hazards help identify geological heritage sites located in high-risk areas that need to be protected. The hazard zoning maps for environmental hazards are not only applicable to heritage sites in the study area but can also be applied to similar areas, towards the goal of sustainable management and development of natural resources that need to be conserved.*

**Keywords:** *Big data; Geological heritage; Heritage preservation; Environmental protection; Nature conservation.*

## 1. Giới thiệu

Các di sản địa chất chứa đựng những giá trị nổi bật về địa chất, địa mạo, đa dạng sinh học và là một dạng cảnh quan môi trường đặc biệt, mang lại nhiều giá trị về nguồn lợi kinh tế, khoa học, giáo dục, giải trí và thẩm mỹ (UNESCO, 2015). Khu vực Trùng Khánh, tỉnh Cao Bằng có một phần diện tích thuộc Công viên địa chất Non nước Cao Bằng, đây là Công viên địa chất Toàn cầu thứ hai tại Việt Nam sau Công viên Địa chất cao nguyên đá Đồng Văn, Hà Giang. Đồng thời, nơi đây còn chứa đựng nhiều di sản văn hóa vật thể và phi vật thể, đặc biệt là các di tích lịch sử cách mạng, di sản địa chất độc đáo, cảnh quan đá vôi phong phú, đa dạng, hang ngầm karst phản ánh chu kỳ tiến hóa của phía Bắc Việt Nam. Thêm vào đó là các di sản địa chất khác tạo nên kỳ quan với nhiều danh lam thắng cảnh nổi tiếng như Thác Bản Giốc, động Ngườm Ngao, chùa Phật tích Trúc Lâm,...

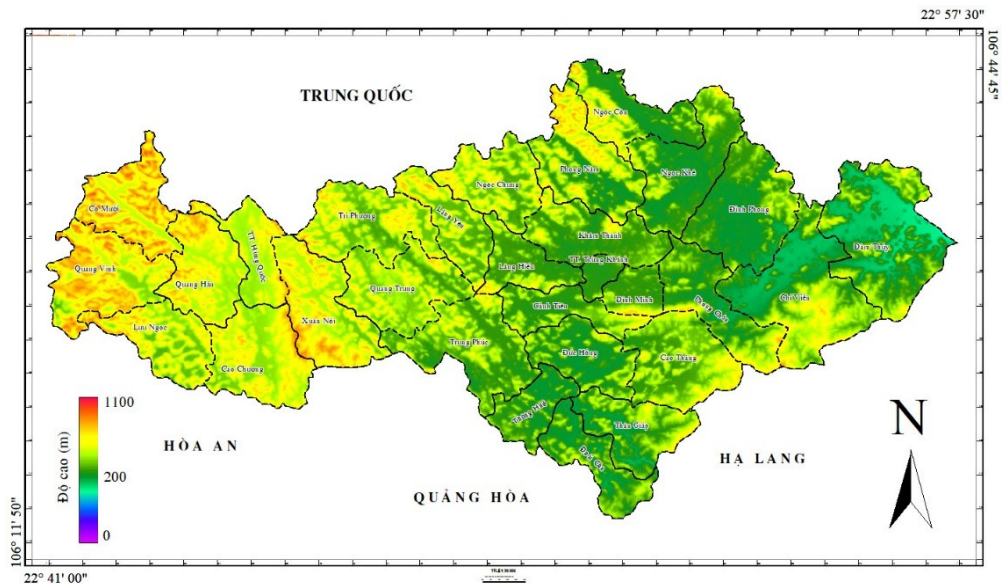
Hiện nay, vấn đề tai biến môi trường đang là một trong những vấn đề nóng không chỉ ở Việt Nam mà còn là vấn đề nổi cộm trên toàn thế giới. Việc bảo vệ các điểm di sản địa chất quý giá đã được công nhận là hết sức cần thiết và rất cần được đầu tư nghiên cứu. Đặc biệt, trong bối cảnh hiện nay, song hành cùng tốc độ phát triển kinh tế nhanh là nhu cầu sử dụng và khai thác tài nguyên lớn, nhiều hoạt động nhân sinh đang diễn ra hàng ngày trên bề mặt và dưới sâu địa hình đã đe dọa và dẫn tới nguy cơ phá vỡ môi trường tự nhiên, phá hủy nguồn tài nguyên quý giá này. Thời gian gần đây, với sự trợ giúp của công nghệ GIS và viễn thám, khả năng đánh giá rủi ro đối với các khu vực có nguy cơ xảy ra tai biến đã tỏ rõ những hiệu quả đáng kể tại các nước phát triển và một số nước trong khu vực như Nhật Bản, Hàn Quốc, Hồng Kông, Đài Loan. Cùng với sự phát triển mạnh mẽ của công nghệ thông tin, hầu hết các công tác quan trắc ở Việt Nam đã được tự động hóa để có thể phát hiện và đưa ra các cảnh báo sự cố một cách chính xác, kịp thời, đã góp phần không nhỏ để giảm thiểu thiệt hại về tính mạng và tài sản do các hiện tượng tai biến gây ra. Cụ thể, hướng nghiên cứu tập trung

trình bày khả năng ứng dụng các phương pháp phân tích dữ liệu lớn (big data) nhằm đánh giá nguy cơ xảy ra tai biến môi trường. Từ đó, đánh giá ảnh hưởng của chúng tới các di sản địa chất, thực hiện cụ thể cho các điểm di sản tại khu vực Trùng Khánh, Cao Bằng.

## 2. Khu vực nghiên cứu, nguồn và phân tích dữ liệu

### 2.1. Khu vực nghiên cứu

Huyện Trùng Khánh, tỉnh Cao Bằng là một trong những khu vực miền núi chịu ảnh hưởng của nhiều thiên tai, đặc biệt là trượt lở đất đá gây ra do tác động của nhiều yếu tố tự nhiên, môi trường và xã hội. Các dạng tai biến tự nhiên trong vùng nghiên cứu khá đa dạng, gồm chủ yếu các tai biến trượt lở đất, lũ quét, lũ bùn đá, xói lở,... Trong đó, các tai biến chủ yếu gồm hiện tượng trượt lở, lũ quét/lũ bùn đá là các dạng tai biến tự nhiên gây thiệt hại nhiều nhất tại khu vực nghiên cứu.



**Hình 1.** Bản đồ nền địa hình tại khu vực nghiên cứu (tỷ lệ 1:50.000)

Thác Bản Giốc nằm trên dòng chảy của sông Quây Sơn. Sông này bắt nguồn từ Trung Quốc, chảy về hướng biên giới hai nước vào lãnh thổ Việt Nam thuộc xã Ngọc Khê, huyện Trùng Khánh. Thác Bản Giốc là thác nước lớn thứ tư thế giới trong các thác nước nằm trên một đường biên giới giữa các quốc gia. Thác có độ cao trên 30m với nhiều khối nước lớn đổ xuống qua nhiều tầng đá vôi.

Động Ngườm Ngao nằm ở bản Gun, xã Đàm Thủy, huyện Trùng Khánh, tỉnh Cao Bằng. Động có chiều dài 2.144m, gồm ba cửa chính là: Ngườm Ngao, Ngườm Lồm và cửa Bản Thuôn.

Sông Quây Sơn này bắt nguồn từ tỉnh Quảng Tây của Trung Quốc chảy qua cửa khẩu Pò Peo. Cảnh sắc của sông Quây Sơn với những cánh đồng lúa chín và những đập nước đặc trưng của miền núi Phía Bắc tựa như những bức tranh tuyệt đẹp trong vắt và xanh ngắt.

Làng Khuổi Ky nằm ở giữa đường từ Bản Giốc đến động Ngườm Ngao. Ngôi làng nổi tiếng với những ngôi nhà sàn bằng đá thơ mộng, nằm cách thác Bản Giốc khoảng 3km, cách trung tâm Thành phố Cao Bằng gần 100km. Làng Khuổi Ky có 14 hộ dân tộc Tày sinh sống tại những ngôi nhà sàn làm bằng đá tồn tại hàng trăm năm. Những ngôi nhà sàn đá này có từ khoảng năm 1594-1677, khi nhà Mạc lên vùng đất Cao Bằng xây dựng thành quách để bảo vệ đất nước.

## **2.2. Nguồn và phân tích dữ liệu**

Các vị trí xảy ra tai biến và cơ sở dữ liệu các yếu tố ảnh hưởng đã được xây dựng trên cơ sở nguồn tư liệu viễn thám và công nghệ GIS. Trong đó, các nguồn dữ liệu mở như nguồn dữ liệu ảnh viễn thám đa thời gian, dữ liệu đo mưa vệ tinh, độ ẩm đất, dữ liệu thảm phủ,... đã được khai thác triệt để. Đây là nguồn thông tin miễn phí, có khả năng sử dụng tối ưu, đặc biệt với các khu vực có độ che phủ kém, địa hình phân cắt mạnh và khó khăn trong quá trình khảo sát như tại vùng nghiên cứu.

Cơ sở dữ liệu các yếu tố môi trường ảnh hưởng đến nguy cơ xảy ra tai biến tại khu vực nghiên cứu bao gồm:

- Nhóm các yếu tố địa hình - địa mạo gồm: Độ cao, độ dốc địa hình, độ ẩm địa hình (TWI), cân bằng khối (MBI)...
- Nhóm các yếu tố địa chất gồm: đặc điểm thạch học, mật độ khe nứt và đứt gãy...
- Nhóm các yếu tố khí tượng thủy văn gồm: dữ liệu về lượng mưa, độ ẩm đất...
- Nhóm các yếu tố nhân sinh gồm: hệ thống giao thông và hiện trạng khai thác mỏ, phân bố dân cư.

Kết quả tổng hợp thu thập được trong toàn vùng nghiên cứu có rất nhiều điểm xảy ra tai biến, với nguồn dữ liệu ảnh Sentinel-2 A/B trước và sau thời điểm xảy ra các trận mưa bão lớn được chồng chập với nhau để nhận diện sự thay đổi của thảm phủ.

Ngoài thông tin về các điểm trượt lở, việc xác định các yếu tố nguy cơ gây nên hiện tượng trượt lở cũng rất quan trọng. Các yếu tố này có thể do độ bền của đất đá bị giảm đi, do trạng thái ứng suất ở sườn dốc bị thay đổi, hoặc do cả hai nguyên nhân trên làm cho điều kiện cân bằng của khối đất đá ở sườn dốc bị phá hủy. Các yếu tố ảnh hưởng có tác dụng hỗ trợ cho quá trình phá hoại sự cân bằng của khối đất đá xảy ra được dễ dàng. Các yếu tố này có thể bao gồm các yếu tố tự nhiên như đặc điểm địa chất (địa tầng, kiến tạo, đứt gãy, nứt nẻ, tính chất cơ lý của đất đá, các quá trình và hiện tượng địa chất động lực công trình,...), đặc điểm của nước ngầm, nước mặt, các điều kiện địa hình, địa mạo, khí hậu, các yếu tố thời gian và các yếu tố nhân sinh. Ảnh hưởng xấu của các yếu tố trên làm giảm khả năng chống trượt của khối đất đá ở sườn dốc, thay đổi trạng thái ứng suất trong khối đất đá. Các yếu tố đó kết hợp đan xen với nhau tạo nên các cơ chế mất ổn định sườn dốc khác nhau.

### **2.3. Phương pháp phân tích dữ liệu lớn**

Các mô hình dữ liệu lớn được sử dụng trong bài viết này gồm: dạng cây quyết định (decision trees C4.5, CART), mạng neural network với cấu trúc mạng suy luận sâu (Deep Feed Forward - DFF) hay phương pháp phân lớp vector hỗ trợ (Support Vector Machines - SVM) để dự báo và so sánh các kết quả với nhau, áp dụng phương pháp tối ưu hóa (optimization) như salp swarm, moth flame... để giúp các mô hình trên hội tụ và cho kết quả nhanh hơn.

Tùy mức độ chi tiết của nguồn số liệu thu thập được, lựa chọn phương pháp phù hợp cho việc đánh giá vai trò ảnh hưởng của các yếu tố liên quan tới các hiện tượng tai biến thể hiện khác nhau và quá trình tích hợp thông tin sẽ tìm ra các vùng hội tụ các yếu tố gây ra tai biến.

#### **2.3.1. Mạng nơron nhân tạo (ANN) với mạng suy luận sâu (Deep Feed Forward - DFF)**

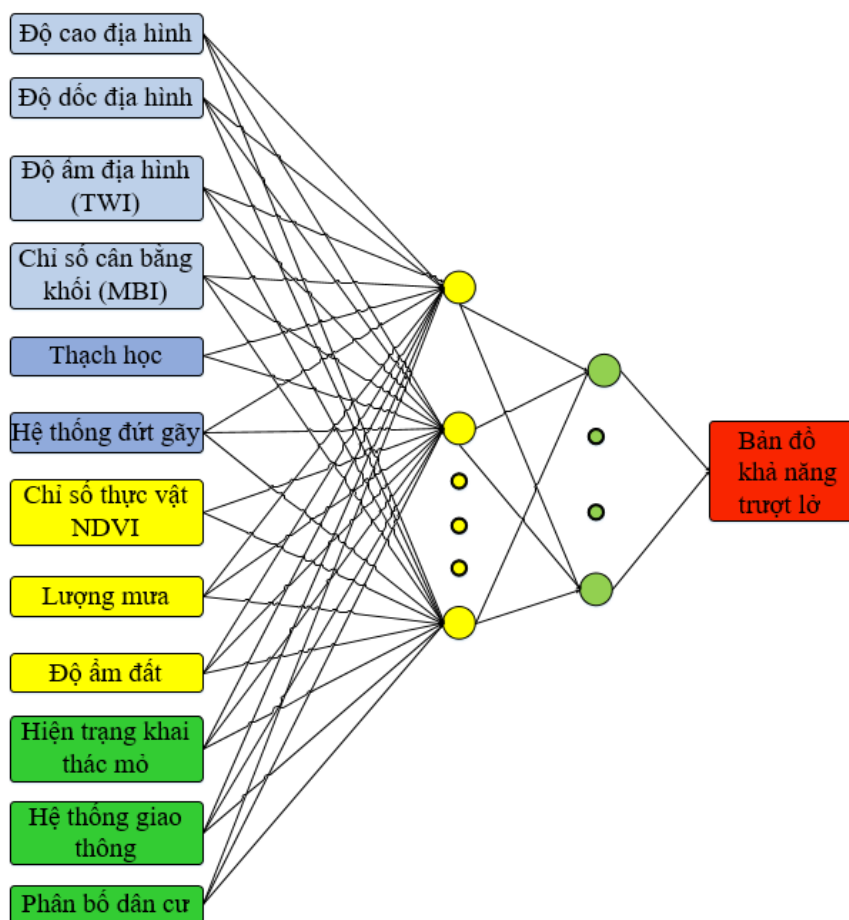
Đây là phương pháp nghiên cứu mới, bắt đầu được đưa vào ứng dụng trong nghiên cứu môi trường. Mạng nơron nhân tạo (artificial neural network - ANN), là thiết kế bằng các mô hình toán học để “bắt chước” hoạt động của bộ não con người. Mỗi mạng nơron là một mô hình tính toán chứa các đơn vị xử lý có khả năng thông tin với nhau bằng cách gửi các tín hiệu đến lẫn nhau thông qua các liên kết có trọng số.

Một mạng nơron bao gồm một chuỗi các “đơn vị” xử lý được kết nối với nhau như tế bào thần kinh của con người. Mạng sẽ bao gồm ba loại lớp: lớp nhập (input), lớp xuất (output) và lớp ẩn (hidden). Mỗi một “đơn vị” trong

lớp nhập và lớp xuất sẽ kết nối với tất cả các “đơn vị” trong lớp ẩn. Điều này khiến cho mạng nơron sẽ là mạng xử lý song song hơn là xử lý tuần tự.

Với các biến đổi đầu vào và đầu ra, mạng nơron sẽ được huấn luyện để bắt đầu tiến trình “học”. Tiến trình học là tiến trình quan trọng của con người, nhờ học mà bộ não ngày càng tích lũy những kinh nghiệm để thích nghi với môi trường và xử lý tình huống tốt hơn. Mạng nơron xây dựng lại cấu trúc bộ não nên cũng cần có khả năng nhận biết dữ liệu thông qua tiến trình học, với các thông số tự do của mạng có thể thay đổi liên tục bởi những thay đổi của môi trường và mạng nơron ghi nhớ giá trị đó.

Khi kết hợp cùng mạng suy luận sâu DFF, mô hình này có khả năng thích nghi, nghĩa là “học sâu” từ các mẫu lặp lại.

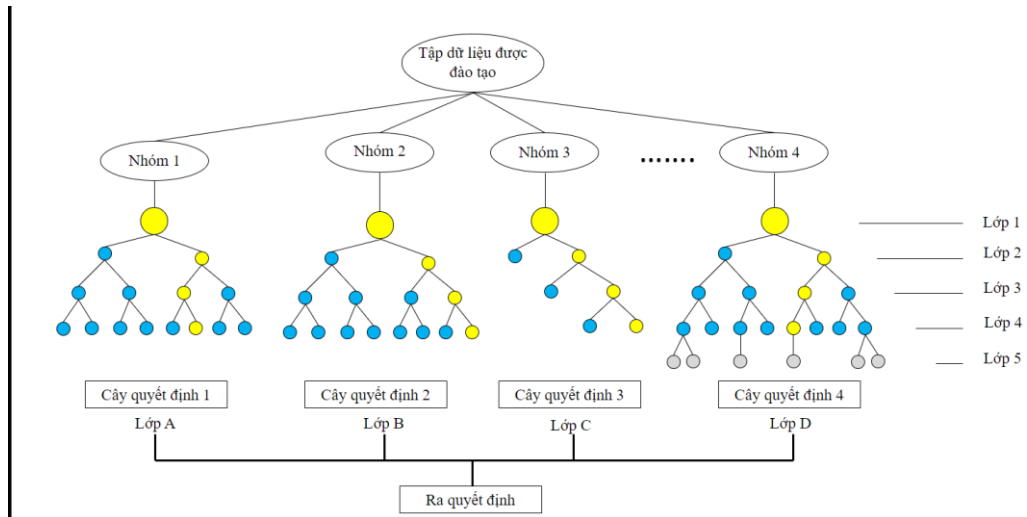


Nguồn: Nhóm nghiên cứu

**Hình 2.** Hình ảnh cấu trúc một mạng nơron các lớp hệ thống

2.3.2. Phương pháp cây quyết định (Decision Tree - DT) (C4.5)

Mô hình cây quyết định là một mô hình được sử dụng khá phổ biến và hiệu quả trong cả hai lớp bài toán phân loại và dự báo. Cây quyết định (Decision Tree) là một cây phân cấp có cấu trúc được dùng để phân lớp các đối tượng dựa vào dãy các luật (series of rules). Khi cho dữ liệu về các đối tượng gồm các thuộc tính cùng với lớp (classes) của nó, cây quyết định sẽ sinh ra các luật để dự đoán lớp của các đối tượng chưa biết (unseen data). Nó mô tả một cấu trúc dạng cây, các lá đại diện cho các phân loại còn cành đại diện cho các kết hợp của các thuộc tính dẫn tới phân loại đó. Một cây quyết định có thể được xây dựng bằng cách chia tập hợp nguồn thành các tập con dựa sự phân loại các giá trị thuộc tính.



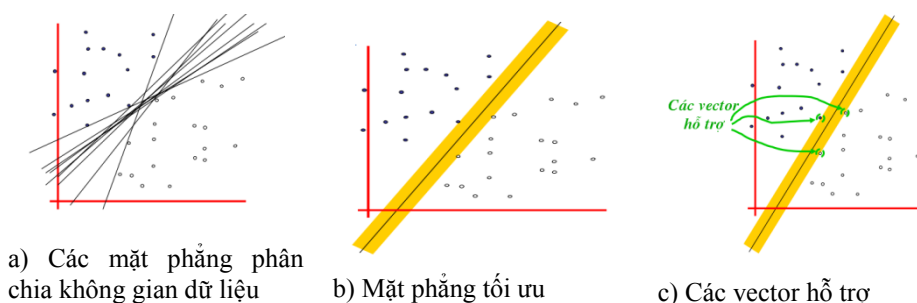
Nguồn: Nhóm nghiên cứu

**Hình 3.** Mô hình cây quyết định

Thuật toán C4.5 (dạng cải tiến viết trên Java là J4.8) cho phép phân tách các nhánh cành cây dưới dạng có/không, dựa trên các điều kiện phân tách (splitting criteria) khác nhau. Ưu điểm của mô hình cây là các điều kiện phân tách rất dễ hiểu, có thể quyết định nhanh chóng, chuẩn bị dữ liệu đơn giản, xử lý nhanh một lượng dữ liệu lớn trong thời gian ngắn. Nhược điểm của cây quyết định là khó giải quyết được những vấn đề có dữ liệu phụ thuộc thời gian liên tục như dữ liệu về dòng chảy hay lượng mưa liên tục và dễ xảy ra lỗi khi cây phát triển có quá nhiều nhánh, làm giảm tính dự báo phổ quát và tăng thời gian tính toán để xây dựng mô hình.

### 2.3.3. Phương pháp vector hỗ trợ (SVM- Support Vector Machines)

Phương pháp phân lớp vector hỗ trợ được Cortes và Vapnik giới thiệu lần đầu tiên vào năm 1995. SVM coi mỗi dữ liệu là một vector với tọa độ là các thông số của chúng. Mục tiêu của SVM là tối ưu hoá sự phân chia không gian dữ liệu thành các miền khác nhau. Ví dụ của việc phân chia dữ liệu thành các miền theo phương pháp SVM trong không gian 2 chiều được biểu diễn như trong Hình 4 dưới đây.



**Hình 4.** Hình ảnh không gian dữ liệu phân chia thành các miền theo SVM

Phương pháp SVM dùng một tập huấn luyện (training set) được biểu diễn trong không gian vector, trong đó, mỗi dữ liệu là một điểm, thuật toán sẽ giúp tìm ra một mặt phẳng quyết định tốt nhất để phân chia các điểm trên không gian này thành hai lớp riêng biệt, tương ứng gọi là lớp dương (+) và lớp âm (-). Chất lượng của siêu phẳng này được quyết định bởi khoảng cách (gọi là biên) của các điểm dữ liệu gần nhất của mỗi lớp đến mặt phẳng này. Khoảng cách biên càng lớn thì mặt phẳng quyết định càng tối ưu, đồng thời việc phân lớp càng chính xác. Mục đích của thuật toán SVM là tìm ra được khoảng cách biên lớn nhất để đạt kết quả phân lớp tốt nhất (Vapnik, 1998).

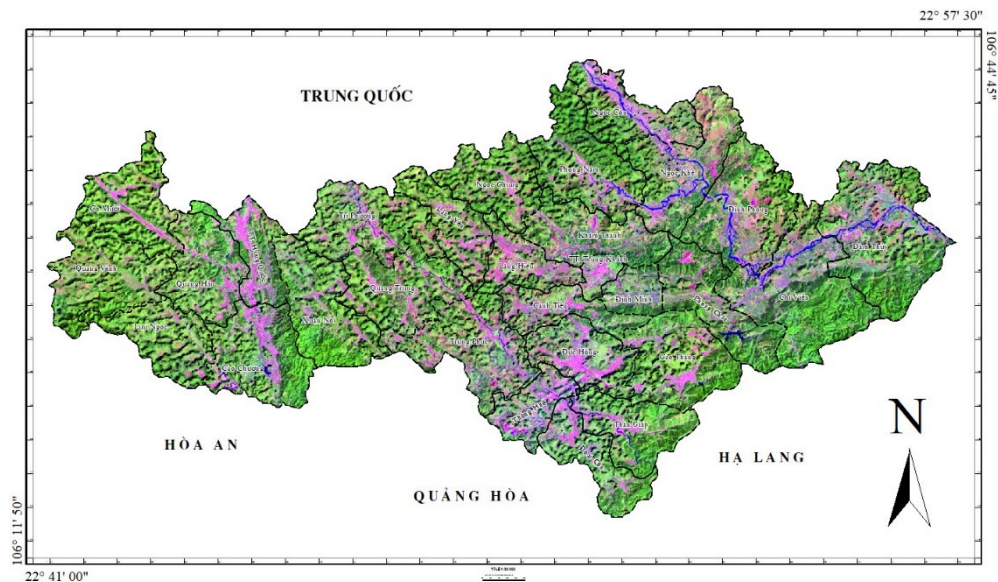
### 3. Các yếu tố ảnh hưởng

Nguồn tư liệu ảnh viễn thám và hệ thống thông tin địa lý (GIS) đã được khai thác nhằm xác định các yếu tố ảnh hưởng đến các di sản địa chất như điều kiện địa hình, hoạt động kiến tạo, đặc điểm địa mạo, các điều kiện thủy văn, công trình... Trong nghiên cứu tai biến môi trường nói chung, công cụ viễn thám thường được sử dụng trong việc xác định vị trí xảy ra tai biến và xây dựng cơ sở dữ liệu các yếu tố ảnh hưởng. Trong đó, việc xác định các vị trí cũng như thời gian xảy ra tai biến là hai thông tin rất quan trọng. Nó là thông tin đầu vào cho các phân tích sau này cũng như là số liệu để đánh giá mức độ chính xác của các kết quả phân tích được thể hiện bằng các bản đồ tai biến và rủi ro.

Tại khu vực nghiên cứu, nguồn ảnh vệ tinh chủ yếu sử dụng là ảnh Landsat



8 của NASA (Mỹ) và ảnh Sentinel-2 của Cơ quan Hàng không Vũ trụ châu Âu (ESA), do sự sẵn có của nguồn ảnh và có thể được sử dụng hoàn toàn miễn phí. Mặc dù nguồn tư liệu ảnh Landsat và Sentinel hoàn toàn miễn phí vẫn là nguồn tư liệu phổ biến nhất cho việc xác định vị trí trượt lở, nhất là tại những nơi xảy ra hiện tượng trượt lở gây mất lớp phủ thực vật có thể phản ánh rõ ràng trên các dải phổ khác nhau, đặc biệt ở dải sóng hồng ngoại và vùng ánh sáng khả kiến. Việc phân tích chủ yếu dựa trên sự so sánh những thay đổi trước và sau một thời điểm cụ thể nào đó, như sau các trận động đất hoặc các cơn bão lớn. Phương pháp này áp dụng có hiệu quả rất cao với các khu vực có độ che phủ kém, địa hình phân cắt mạnh và khó khăn trong quá trình khảo sát.

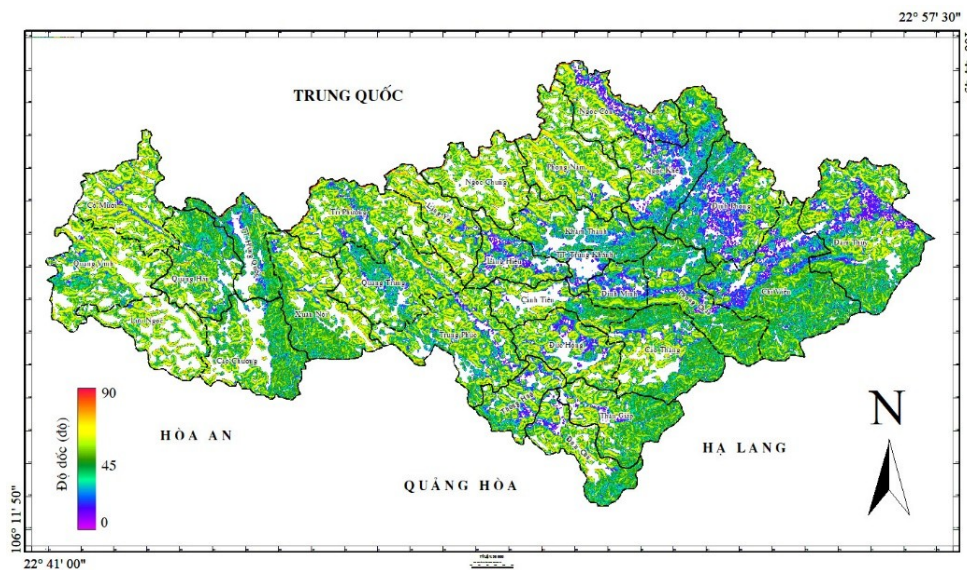


**Hình 5.** Ảnh vệ tinh Landsat 8 OLI tại khu vực nghiên cứu (tỷ lệ 1:50.000)

Các yếu tố trắc lượng hình thái có ảnh hưởng đến nguy cơ xảy ra tai biến địa động lực bao gồm độ dốc, độ cao, hướng sườn hay độ phân cắt sâu, độ phân cắt ngang. Độ cao địa hình khác nhau sẽ làm thay đổi điều kiện ngoại sinh: mưa, ẩm, phong hóa,... từ đó, có thể ảnh hưởng tới các quá trình tai biến. Khu vực nghiên cứu có địa hình phức tạp, có độ cao lớn, phân cắt mạnh và phát triển nhiều hệ thống sông suối khác nhau.

Yếu tố địa hình, trong đó cụ thể là độ dốc sườn là nguyên nhân chính gây ra trượt lở. Địa hình cao và độ phân cắt lớn tạo ra năng lượng địa hình lớn, thuận lợi cho trượt lở có nguồn gốc trọng lực. Kết quả thống kê ngoài thực địa cũng cho thấy số lượng các điểm trượt lở tỷ lệ thuận với độ cao và độ phân cắt của địa hình. Có trên 90% số điểm trượt lở phân bố ở các khu vực có độ dốc địa hình trên  $25^{\circ}$ . Từ mô hình số độ cao của khu vực, nghiên cứu

này còn tính toán bổ sung các chỉ số địa hình nhằm thể hiện sự thay đổi về năng lượng địa hình hoặc mức độ tập trung của dòng chảy. Các chỉ số được sử dụng trong nghiên cứu này bao gồm chỉ số độ ẩm địa hình (Terrain Wetness Index- TWI) và chỉ số cân bằng khối (Mass Balance Index- MBI).



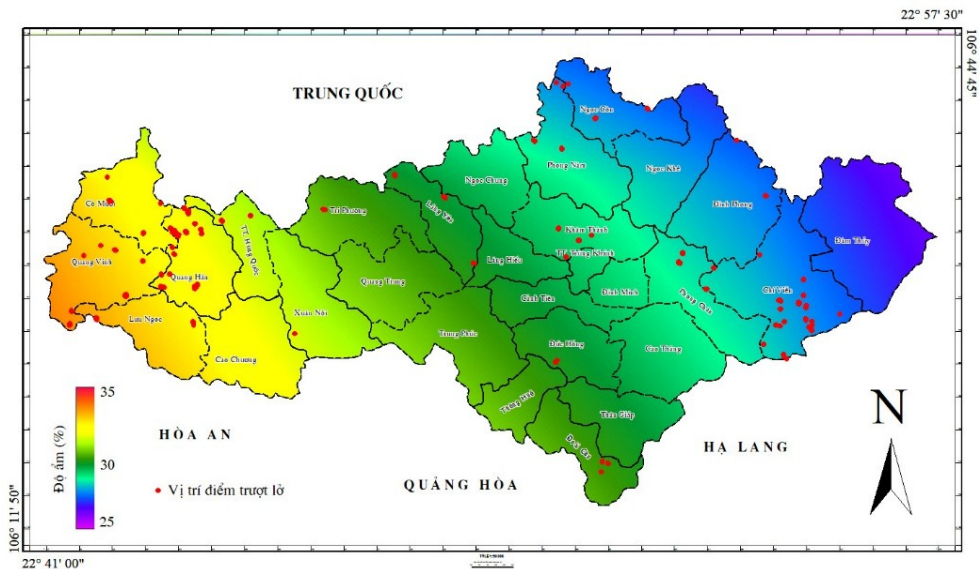
**Hình 6.** Độ dốc địa hình tại khu vực nghiên cứu (tỷ lệ 1:50.000)

Thành phần thạch học của các đá gốc đóng vai trò quan trọng cho sự hình thành hiện tượng trượt lở. Hiện tượng trượt lở có thể xảy ra ở tất cả các loại đá gốc có thành phần khác nhau, tuy nhiên các vật liệu có độ bền thấp biểu hiện mối nguy hiểm lớn nhất. Kết quả thống kê cho thấy các điểm trượt tập trung trên các sườn có độ dốc trên  $20\text{-}30^{\circ}$  và tập trung ở các đới cắt trượt có đá gốc vỡ vụn hoặc những nơi vỏ phong hóa có lẫn sét.

Hiện tượng trượt không những phụ thuộc vào thành phần, tính chất của đá gốc mà còn phụ thuộc vào các hoạt động kiến tạo. Mức độ phá huỷ đứt gãy kiến tạo là điều kiện thuận lợi cho sự phát sinh, phát triển dịch chuyển trọng lực vì đó là những nơi mà đất đá bị vụn nát, các tính chất cơ lý, đặc biệt là góc nội ma sát và lực dính kết giảm đột ngột, là nơi tàng trữ nước, làm giảm sức kháng cắt của đất đá. Các phá huỷ kiến tạo cũng là nơi dễ phát sinh các quá trình địa động lực khác, có khả năng ảnh hưởng và kích thích hiện tượng trượt lở.

Các yếu tố khí tượng - thủy văn cũng quan trọng như thành phần và tính chất của vật liệu mái dốc, hình thành nên các dạng tai biến tự nhiên trong vùng nghiên cứu. Nước hầu như trực tiếp hay gián tiếp liên quan đến các dạng tai biến, chuyển động khối nhanh đều xảy ra trong và sau thời gian mưa dữ dội hoặc mưa lớn kéo dài. Dữ liệu lượng mưa (mm/ngày) được thu

thập từ số liệu đo mưa vệ tinh GSMaP của Cơ quan nghiên cứu và phát triển hàng không vũ trụ Nhật Bản (JAXA) và số liệu đo độ ẩm SMOPS (Soil Moisture Products) của Cơ quan Thông tin, Dữ liệu và viễn thám môi trường quốc gia của Mỹ (NESDIS) vào tháng mưa trung bình cao nhất hàng năm tại khu vực (tháng 7). Kết quả được nội suy về độ phân giải 20x20m từ độ phân giải gốc ~10km (0,1x0,1 độ) để sử dụng trong các bước tính toán tiếp theo.



**Hình 7.** Độ ẩm đất tại khu vực nghiên cứu (tỷ lệ 1:50.000)

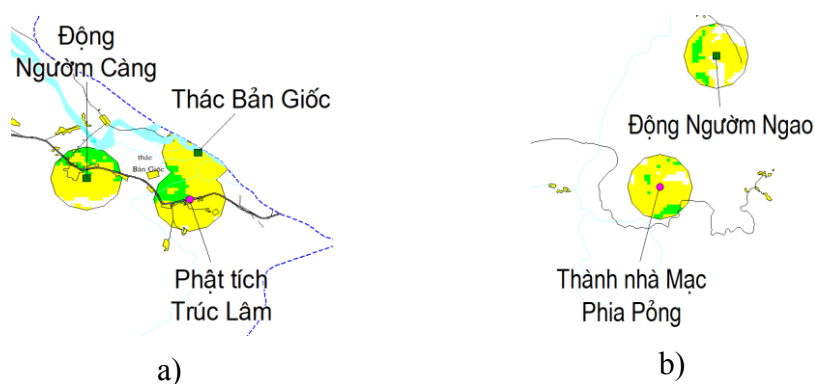
Thảm thực vật là nhân tố quan trọng trên các sườn dốc do thực vật là một màn chắn để hạn chế lượng mưa rơi trên các đỉnh dốc, tạo điều kiện thuận lợi cho sự thấm nước vào đất. Ngoài ra, thực vật có hệ rễ tạo ra sự kết dính các vật liệu trên các sườn dốc và bản thân thực vật cũng thêm trọng lượng vào dốc. Lớp phủ thực vật còn tạo điều kiện bốc hơi ẩm nhanh, lượng bốc hơi từ lớp phủ thực vật vượt rất xa từ đất đá lộ trần và gần bằng lượng bốc hơi từ mặt nước.

Các hoạt động nhân sinh chủ yếu tại khu vực nghiên cứu bao gồm các hoạt động khai thác mỏ, xây dựng công trình nhà cửa, giao thông. Việc nghiên cứu tai biến trượt lở dọc các tuyến đường cho thấy rất rõ mức độ tác động của con người đến sự ổn định của sườn dốc. Bên cạnh các yếu tố môi trường tự nhiên, các hoạt động của con người làm giảm sức chống đỡ của sườn như xây dựng hệ thống đường giao thông, làm thay đổi thể cân bằng sườn, cắt xén chân sườn dốc khi làm đường, cũng như làm tăng tải trọng cho sườn dốc như xây dựng các công trình có tải trọng lớn trên sườn dốc cũng được đưa vào phân tích.

#### 4. Kết quả nghiên cứu

Bộ dữ liệu đưa vào tính toán được lựa chọn ngẫu nhiên từ 12 lớp thông tin gồm và được phân chia thành hai tập dữ liệu với tỷ lệ phân tách là 70% cho tập dữ liệu huấn luyện (training) để xây dựng các mô hình và 30% cho tập dữ liệu kiểm tra (testing) để kiểm tra mức độ chính xác của mô hình tính toán. Kết quả xây dựng 03 mô hình phân tích dữ liệu lớn cho thấy cả ba phương pháp đều cho kết quả dự báo nguy cơ xảy ra tai biến tại khu vực nghiên cứu với độ tin cậy khá tốt. Mô hình mạng trí tuệ nhân tạo ANN cho kết quả tốt nhất (độ chính xác đạt 89,25%), tiếp theo là phương pháp vectơ hỗ trợ SVM (87,38%) và cuối cùng là mô hình cây quyết định C4.5 (83,64).

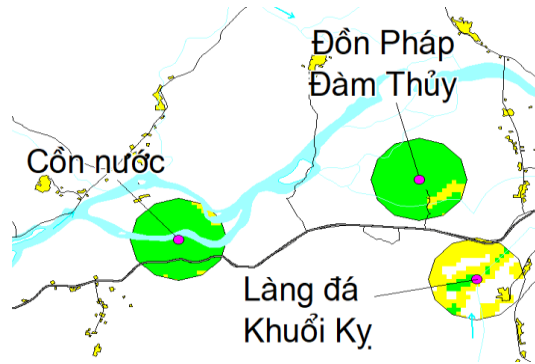
Kết quả phân tích chi tiết cho thấy các khu vực có nguy cơ xảy ra tai biến môi trường mạnh nhất tập trung ở phía Tây và Đông Nam khu vực nghiên cứu, thuộc các xã của huyện Trà Lĩnh cũ và khu vực Hành lang đa dạng sinh học, một phần của khu bảo vệ cảnh quan thác Bản Giốc. Trong đó, tại khu bảo vệ cảnh quan thác Bản Giốc có mức độ nguy cơ không quá lớn song tại đây có 03 điểm di sản địa chất rất có giá trị du lịch là động Ngườm Ngao, thác Bản Giốc, hang Ngườm Càng và 02 điểm di sản văn hóa: Làng Khuổi Ky, chùa Phật Tích Trúc Lâm.



*Nguồn: Nhóm nghiên cứu*

**Hình 8.** Hình ảnh chi tiết khu vực quanh thác Bản Giốc, chùa Phật Tích Trúc Lâm và động Ngườm Càng (a); Động Ngườm Ngao và Thành nhà Mạc Phía Pông (b)

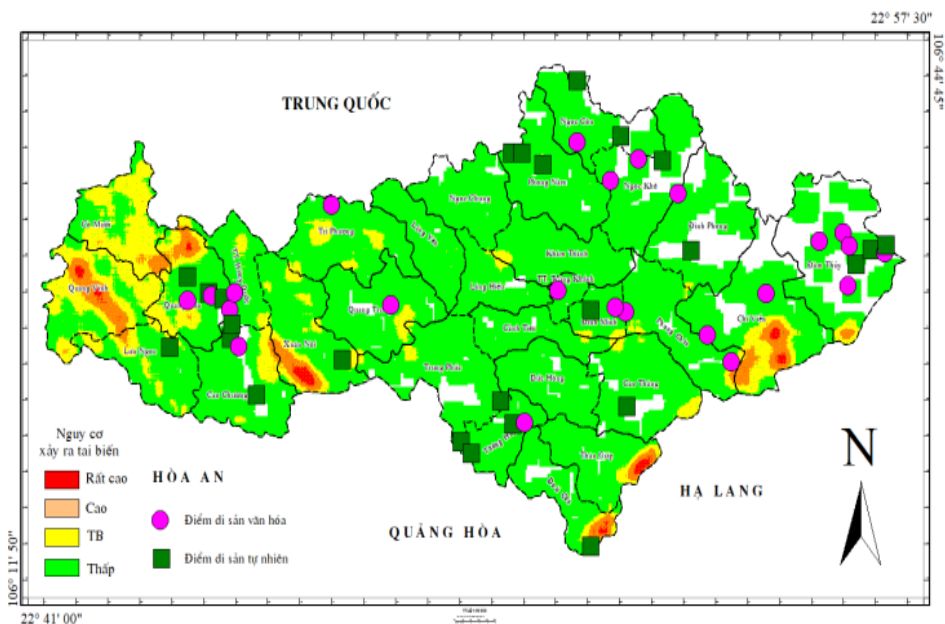
Trong vùng có 02 mỏ khai thác khoáng sản: mỏ mangan Khuổi Ky và mỏ mangan Bản Rạ vẫn còn đang hoạt động gây tác động trực tiếp đến cảnh quan và con người sinh sống trong vùng lõi của khu vực. Các điểm di tích Thành nhà Mạc Phía Cam và di tích Đồn Pháp Đàm Thủy nằm ở phía địa hình thấp cũng có khả năng chịu ảnh hưởng của ngập lụt khi có mưa lớn kéo dài.



Nguồn: Nhóm nghiên cứu

**Hình 9.** Hình ảnh chi tiết khu vực quanh làng đá Khuổi Ky, Đồn Pháp Đàm Thủy và Cồn Nước.

Khu vực Hành lang đa dạng sinh học cũng là nơi có nguy cơ xảy ra tai biến thấp, tuy nhiên, đây là nơi lưu giữ các sinh vật chuẩn hệ gen quý của các loại sinh vật và là khu vực khá nhạy cảm do nằm gần một số điểm di sản địa chất.



Nguồn: Nhóm nghiên cứu

**Hình 10.** Hình ảnh vị trí các điểm di sản trên nền sơ đồ phân vùng nguy cơ xảy ra tai biến (tỷ lệ 1:50.000)



Khu vực phía Tây của vùng nghiên cứu là nơi có nguy cơ xảy ra tai biến mạnh nhất. Khu vực này chủ yếu bao gồm các xã thuộc huyện Trà Lĩnh cũ, được sáp nhập vào huyện Trùng Khánh từ năm 2020 có các điểm di sản địa chất và di sản văn hóa quan trọng như điểm hóa thạch cổ sinh tại xã Lưu Ngọc, ranh giới bất chỉnh hợp tuổi Cambri - Devon - Carbon tại xã Quang Hán, động karst Giác Đầu, di tích cầu đá, đền thờ Nông Thống Lệnh tại thị trấn Hùng Quốc. Quanh khu vực vẫn còn một số mỏ khoáng sản nhỏ lẻ đang hoạt động như các mỏ mangan Bản Mặc, Pác Riếc, các mỏ đá Bản Lò, Cốc Cáng, Phia Chiên,... đây cũng là một phần nguyên nhân gây ra các tác động đến nguy cơ xảy ra tai biến tại khu vực nghiên cứu.

Tình hình mưa không còn ổn định theo mùa, làm tăng nguy cơ trượt lở, lũ quét, lũ bùn đá do vậy mà việc phòng tránh cũng như khắc phục sự cố còn chưa được tốt gây nhiều thiệt hại về người và của. Bên cạnh đó, các hoạt động của con người đóng một vai trò cực kỳ quan trọng kích thích tai biến xảy ra như các hoạt động khai thác khoáng sản, khai thác rừng trái phép, hoạt động nông nghiệp hay các hoạt động xây dựng công trình... gây tác động đến môi trường không hề nhỏ.

## **5. Đề xuất các giải pháp hướng tới mục tiêu quản lý bền vững các di sản địa chất**

Các dạng tai biến môi trường tại khu vực được hình thành có nguyên nhân kết hợp cả từ các điều kiện tự nhiên cũng như các hoạt động của con người và đang diễn ra với tần suất, cường độ ngày càng cao, gây thiệt hại lớn đến sự phát triển bền vững kinh tế-xã hội của khu vực. Kết quả đánh giá mức độ quan trọng của từng yếu tố cho thấy các hiện tượng tai biến môi trường tại khu vực nghiên cứu chủ yếu chịu ảnh hưởng bởi các yếu tố tự nhiên: địa hình, độ dốc, thành phần thạch học, thảm thực vật đóng vai trò là các yếu tố nền và các yếu tố về lượng mưa, độ ẩm đất đóng vai trò là yếu tố kích thích cho sự hình thành tai biến. Các hoạt động nhân sinh như khai thác mỏ, xây dựng công trình nhà cửa, giao thông đóng vai trò thấp hơn song cũng cho thấy mức độ ảnh hưởng tương đối cao của quá trình khai thác khoáng sản, vận chuyển quặng tới các hiện tượng tai biến tại vùng nghiên cứu.

Kết quả khoanh vùng nguy cơ tai biến trượt lở được phân tích dựa trên kết quả tính toán từ các phương pháp: phân tích dữ liệu lớn (big data) gồm mô hình mạng nơron nhân tạo (ANN), phương pháp vectơ hỗ trợ (SVM) và mô hình cây quyết định (DT). Kết quả so sánh mức độ chính xác của các mô hình dự báo trượt lở cho thấy mạng nơron nhân tạo cho kết quả tối ưu nhất.

Dựa trên kết quả tích hợp và phân vùng nguy cơ xảy ra tai biến tại vùng nghiên cứu, bài viết cũng đề xuất một số giải pháp hướng tới mục tiêu

phòng tránh các tai biến môi trường tự nhiên tại khu vực nghiên cứu như sau:

### **5.1. Giải pháp giáo dục**

Nâng cao nhận thức và trách nhiệm của cán bộ, nhân dân và cộng đồng địa phương về công tác quản lý, bảo vệ, bảo tồn di sản, tăng cường công tác tuyên truyền, nâng cao nhận thức, trách nhiệm trong việc gìn giữ và phát huy giá trị khu bảo tồn - là tài sản của nhân loại cần phải giữ gìn cho thế hệ mai sau. Đẩy mạnh công tác phổ biến pháp luật về bảo vệ rừng, bảo vệ môi trường, quy định về du lịch cho cộng đồng địa phương và vùng đệm.

### **5.2. Giải pháp tổ chức quản lý**

Nhà nước cần có những chính sách, quy định phù hợp trong công tác quy hoạch, kế hoạch sử dụng đất để tăng cường hiệu quả quản lý của Nhà nước nhằm phát triển kinh tế-xã hội và đảm bảo an ninh quốc phòng.

Quy hoạch hướng đến mục tiêu phát triển bền vững kinh tế-xã hội, đồng thời, giảm thiểu được nguy cơ xảy ra tai biến trong quá trình khai thác và sử dụng tài nguyên thiên nhiên mà cụ thể ở huyện Trùng Khánh. Do vậy, ngoài chính quyền các cấp, cần huy động và tạo điều kiện thuận lợi cho các bên liên quan như người dân địa phương, hội cựu chiến binh, hội phụ nữ, đoàn thanh niên... tham gia giám sát thực hiện quy hoạch sử dụng tài nguyên thiên nhiên gắn với giảm thiểu tai biến.

Xử lý nghiêm các cá nhân, tổ chức có các hành vi trái với quy định của Nhà nước, của địa phương trong việc khai thác tài nguyên rừng, tài nguyên đất. Xử phạt các hành vi phá rừng làm nhà ở, mở rộng diện tích đất ở trong khu vực cấm khai thác.

### **5.3. Giải pháp về cơ chế chính sách**

Áp dụng chính sách đầu tư phát triển, huy động vốn và nguồn nhân lực để thực hiện phương án quy hoạch; hỗ trợ, cho vay vốn đối với các hộ nghèo, đồng bào dân tộc thiểu số nhằm khuyến khích khai thác và sử dụng đất đai đúng mục đích, có hiệu quả ở vùng núi cao.

Chính sách bảo vệ tài nguyên rừng, đặc biệt là rừng phòng hộ; phát triển kinh tế và phủ xanh đất trống đồi trọc.

Cần nghiên cứu và phát triển cơ chế, chính sách công tác bảo tồn các di sản địa chất dựa trên việc dự báo thông qua bản đồ nguy cơ xảy ra tai biến, bởi lẽ ở đó có những lợi ích thiết thực như tăng trưởng du lịch, tạo công ăn việc làm cho cộng đồng địa phương, công nhân khai thác mỏ nhỏ có cơ hội quay lại làm việc, nhận thức về môi trường và cảnh quan địa chất của người dân

địa phương được gia tăng, cần xác định những vị trí nhạy cảm được bảo vệ tốt hơn.

#### **5.4. Giải pháp về khoa học công nghệ**

Xây dựng hệ thống cơ sở dữ liệu về tai biến nhằm xác định, phân vùng nguy cơ tai biến, tổn thương để đưa ra các chính sách di cư phù hợp cho các khu vực dân cư trong vùng có nguy cơ tai biến, tổn thương cao, dựa vào bộ cơ sở dữ liệu quan trắc tai biến liên tục để có thể phân vùng nguy cơ xảy ra tai biến theo thời gian, phân bố không gian. Các công tác nghiên cứu tiếp theo cần được tiến hành gồm: mở rộng kết quả nghiên cứu cho các dạng tai biến khác như trượt lở, xói mòn các khu vực ven sông, suối... nhằm đánh giá nguy cơ xảy ra tai biến tổng hợp cho tất cả các khu vực phát triển cơ sở hạ tầng khác, đặc biệt tại các khu vực có sự tập trung dân cư cao.

Đào tạo đội ngũ cán bộ khoa học có trình độ chuyên môn cao trong lĩnh vực quản lý, xây dựng, thực hiện và giám sát quy hoạch sử dụng đất và tài nguyên thiên nhiên nhằm giảm thiểu tai biến.

Công tác quản lý và dự báo tai biến môi trường nói chung cần tận dụng tối đa khả năng ứng dụng các công cụ địa không gian hiện có như phân tích ảnh viễn thám và công nghệ GIS kết hợp với các phương pháp tính toán định lượng nhằm đưa ra các kết quả có độ chính xác tốt hơn trong tương lai. Hệ thống các phương pháp và nguồn số liệu phân tích sử dụng trong nghiên cứu này có thể áp dụng được cho các khu vực khác có điều kiện tương tự.

#### **5.5. Giải pháp về quản lý các hoạt động du lịch**

Đẩy mạnh liên kết trong phát triển du lịch, chú trọng phát huy sức mạnh cộng đồng và phát triển du lịch xanh là định hướng cơ bản nhằm khai thác có hiệu quả và quảng bá giá trị của di sản địa chất. Đồng thời, thực hiện định hướng quy hoạch không gian nhằm giảm thiểu tai biến. Tăng cường công tác quản lý hành chính theo luật định, quy hoạch bảo tồn, bảo vệ môi trường, phòng tránh thiên tai. Quy hoạch mạng lưới quan trắc dự báo tai biến trượt lở đất đến cấp thôn bản.

Với giá trị và tiềm năng du lịch của công viên địa chất có thể xây dựng và phát triển các sản phẩm du lịch sau: Du lịch tham quan, thám hiểm hệ thống hang động, tham quan ngắm phong cảnh thiên nhiên của vùng kết hợp du lịch văn hóa lịch sử; Du lịch thể thao, mạo hiểm, vui chơi giải trí; Du lịch trải nghiệm sinh hoạt tại cộng đồng... phát triển du lịch chất lượng cao với nhận thức và hành xử với môi trường một cách thân thiện.

Đề xuất quy hoạch không gian nhằm phòng tránh và giảm nhẹ tai biến là việc tối quan trọng và cần thiết, trước hết cần kiểm tra chặt chẽ, nghiêm



ngặt về đánh giá tác động môi trường của dự án và ảnh hưởng của dự án đến các di sản địa chất.

## 6. Kết luận

Kết quả phân vùng nguy cơ xảy ra tai biến môi trường, các đề xuất nhằm giảm nhẹ rủi ro thiên tai có khả năng ảnh hưởng đến các di sản địa chất tại khu vực nghiên cứu là tài liệu thực tế, có giá trị khoa học về giảm nhẹ rủi ro tai biến môi trường cho địa phương trong định hướng quy hoạch phát triển bền vững kinh tế - xã hội tại các điểm di sản địa chất cần được bảo tồn nói riêng và tại vùng nghiên cứu nói chung.

Đây là hướng nghiên cứu có tính liên ngành về địa chất, môi trường và ứng dụng các phương pháp phân tích dữ liệu lớn tại khu vực nghiên cứu. Các kết quả thu được là cơ sở thông tin hữu dụng cho các nhà quản lý và hoạch định chính sách trong việc bảo vệ các di sản thiên nhiên thế giới nói chung và đặc biệt có giá trị phục vụ cho kỳ tái thẩm định danh hiệu Công viên địa chất toàn cầu UNESCO đối với Công viên địa chất Non nước Cao Bằng./.

## TÀI LIỆU THAM KHẢO

1. La Thế Phúc và Trần Tân Văn (2009). *Báo cáo Nghiên cứu di sản địa chất và xây dựng công viên địa chất ở Việt Nam*. Ủy ban Quốc gia UNESCO Việt Nam, Hà Nội.
2. Phạm Khả Tùy (2004). *Nghiên cứu, đánh giá hiện trạng môi trường karst trên một số vùng trọng điểm ở miền Bắc Việt Nam*. Viện Nghiên cứu Địa chất và Khoáng sản, Hà Nội.
3. Wikipedia (2020). *Công viên địa chất Non Nước Cao Bằng*. Wikipedia <[https://vi.wikipedia.org/wiki/Công\\_viên\\_địa\\_chất\\_Non\\_Nước\\_Cao\\_Bằng](https://vi.wikipedia.org/wiki/Công_viên_địa_chất_Non_Nước_Cao_Bằng)>.
4. Dieu Tien Bui, Binh Thai Pham, Quoc Phi Nguyen, Nhat Duc Hoang, (2016). *Spatial prediction of rainfall-induced shallow landslides using hybrid integration approach of least squares support vector machines and differential evolution optimization: A case study in central Vietnam*. International Journal of Digital Earth. DOI: 10.1080/17538947.2016.1169561.
5. Cortes, C. and Vapnik, V. (1995), *Support Vector Network, Machine Learning*, 20:273-297.
6. Smith K (2001). *Environmental Hazards*. Third Edition. Routledge, London and New York.
7. Deck O. and Verdel T. (2012). *Uncertainties and Risk Analysis Related to Geohazards: From Practical Applications to Research Trends*. Risk Management for the Future - Theory and Cases. InTech Publisher.
8. Quoc Phi Nguyen, Du Duong Bui, SangGi Hwang, Khac Uan Do, Thi Hoa Nguyen, (2018). *Rainfall-triggered landslide and debris flow hazard assessment using data mining techniques: A comparison of Decision Trees, Artificial Neural Network and*

- Support Vector Machines*. Proceedings of the 2018 Vietnam Water Cooperation Initiative (VACI 2018) - Highlights. Science and Technics Publishing House, Hanoi, Vietnam, p.138-141, ISBN 978-604-67-1059-2.
9. Quoc Phi Nguyen and Hoang Bac Bui (2004). *Landslide hazard mapping using Bayesian approach in GIS - Case study in Yangsan area, Korea*. In *GeoInformatics for Spatial-Infrastructure Development in Earth & Allied Sciences GIS-IDEAS 2004*.
  12. UNESCO Global Geoparks Network (2015). *Guidelines and criteria for national geopark seeking UNESCO's assistance to join the Global Geoparks Network*. Cordination Unit, European Geoparks Network, BP 156, F-04005 Digne-les-Bains cedex, France, 10pp.