

ĐẠI HỌC QUỐC GIA HÀ NỘI
TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN

Lại Tuấn Anh

ỨNG DỤNG VIỄN THÁM VÀ GIS TRONG
NGHIÊN CỨU, ĐÁNH GIÁ NGUY CƠ LŨ QUÉT
PHỤC VỤ CẢNH BÁO SỚM TRÊN LƯU VỰC
SUỐI MUỘI TỈNH SƠN LA

DỰ THẢO TÓM TẮT LUẬN ÁN TIẾN SĨ ĐỊA LÝ

Hà Nội - 2023

Công trình được hoàn thành tại:

Người hướng dẫn khoa học:
PGS.TS. Bùi Quang Thành
GS.TS. Nguyễn Ngọc Thạch

Phản biện:

Phản biện:

Phản biện:

Luận án sẽ được bảo vệ trước Hội đồng cấp Đại học Quốc gia chăm
luận án tiến sĩ họp tại
vào hồi giờ ngày tháng năm 2023.

- Có thể tìm hiểu luận án tại:
- Thư viện Quốc gia Việt Nam
 - Trung tâm Thông tin-Thư viện, Đại học Quốc gia Hà Nội

MỞ ĐẦU

Tính cấp thiết của luận án

Nằm trong khu vực Đông Nam Á, Việt Nam là một trong mười quốc gia trên thế giới chịu tác động mạnh nhất của thiên tai do biến đổi khí hậu. Tổng thiệt hại ước tính sơ bộ do bão và lũ thống kê chỉ riêng trong giai đoạn 1996-2015 đã chiếm 0.62% tổng sản phẩm quốc nội GDP, tương đương với khoảng 2,12 tỷ USD. Vào mùa mưa (từ tháng 5 đến tháng 10), các lưu vực sông suối trên khắp miền núi Tây Bắc đều có nguy cơ xảy ra lũ quét khi có mưa lớn. Lưu vực suối Muội tỉnh Sơn La trong 10 năm trở lại đây hầu như năm nào cũng xảy ra tai biến lũ quét gây ảnh hưởng nghiêm trọng đến cuộc sống của người dân. Sự tàn phá của lũ quét sẽ càng lớn hơn đặc biệt ở những nơi chưa phát triển hệ thống dự báo và giảm thiểu tác động của lũ. Để giảm thiểu thiệt hại do lũ quét gây ra, việc xây dựng mô hình đánh giá nguy cơ lũ quét và một hệ thống bản đồ nguy cơ lũ quét dựa trên các yếu tố khác nhau về nguyên nhân và tác động của lũ quét là cần thiết. Với những lợi thế và ưu điểm vượt trội, phương pháp viễn thám và GIS kết hợp với mô hình học máy có thể thực hiện được công việc này một cách hiệu quả. Viễn thám cho phép ta phân tích, giải đoán, chiết tách các thông tin hữu ích như lớp phủ, ngập lụt trên ảnh quang học cũng như ảnh Radar. GIS là công cụ sử dụng hiệu quả trong các bài toán phân tích không gian, chồng chập các lớp bản đồ hỗ trợ ra quyết định trong nghiên cứu lũ quét. Các mô hình học máy rất phổ biến do khả năng xây dựng số lượng phi tuyến tính của lũ dựa trên bộ dữ liệu lịch sử, mang tính khách quan vì vậy các thuật toán học máy đã được cải thiện đều đặn, thể hiện khả năng dự báo lũ vượt trội so với các phương pháp truyền thống.

Xuất phát từ thực tiễn nêu trên, đề tài luận án tiến sĩ: **“Ứng dụng viễn thám và GIS trong nghiên cứu, đánh giá nguy cơ lũ quét phục vụ cảnh báo sớm trên lưu vực suối Muội tỉnh Sơn La”** được đề xuất là cần thiết.

Mục tiêu và nhiệm vụ của luận án

Mục tiêu nghiên cứu bao gồm nghiên cứu các nội dung sau:

- Đánh giá cơ sở khoa học lựa chọn các yếu tố sử dụng trong thành lập bản đồ nguy cơ lũ quét
- Xây dựng bộ dữ liệu bao gồm các điểm lũ quét và các lớp yếu tố từ dữ liệu viễn thám và GIS
- Nghiên cứu mô hình phân tích đa chỉ tiêu (MCA) theo tiếp cận FFPI trong thành lập bản đồ nguy cơ lũ quét
- Nghiên cứu các thử nghiệm mô hình học máy, các thuật toán tối ưu trong thành lập bản đồ nguy cơ lũ quét
- Nghiên cứu mô hình hỗ trợ cảnh báo lũ quét trên lưu vực suối Muội tỉnh Sơn La.

Những điểm mới của luận án

- Đánh giá cơ sở khoa học lựa chọn các yếu tố sử dụng trong thành lập bản đồ nguy cơ lũ quét. Dựa trên mô hình xây dựng, nghiên cứu đã tổng quát hóa được vai trò của các yếu tố trong đánh giá nguy cơ lũ quét
- Tổng quát hóa việc sử dụng mô hình, tối ưu hóa các tham số và xây dựng được mô hình học máy trong nghiên cứu lũ quét. Về mặt ứng dụng có 2 dạng tham số của các mô hình học máy (1) Tham số mô hình được điều chỉnh dựa vào dữ liệu đầu vào sử dụng trong mô hình; (2) Siêu tham số (hyper parameters) được thiết lập trước khi mô hình được đưa vào huấn luyện. Điểm mới của nghiên cứu này là đề xuất thử nghiệm được các thuật toán tối ưu trong lựa chọn các siêu

tham số nhằm nâng cao độ chính xác xây dựng bản đồ nguy cơ lũ quét.

Ý nghĩa khoa học và thực tiễn

- Về mặt khoa học: Luận án đã đề xuất được phương pháp và mô hình học máy trong nghiên cứu lũ quét, đánh giá được nguy cơ lũ quét phục vụ cảnh báo sớm trên lưu vực suối Muội tỉnh Sơn La.

- Về mặt thực tiễn: bổ sung cơ sở khoa học ứng dụng viễn thám, GIS và phương pháp học máy trong nghiên cứu, cảnh báo nguy cơ lũ quét trên lưu vực suối Muội tỉnh Sơn La. Có thể áp dụng cho các khu vực nghiên cứu khác có điều kiện tương đồng.

Chương 1. TỔNG QUAN CÁC NGHIÊN CỨU VỀ LŨ QUÉT

1.1. Khái niệm lũ quét và cơ chế hình thành lũ quét

1.1.1 Khái niệm lũ quét

“Lũ quét (LQ) là hiện tượng thiên tai tự nhiên nguy hiểm, được hình thành do mưa kết hợp các tổ hợp bất lợi về điều kiện mặt đệm (địa hình, địa mạo, lớp phủ..) sinh ra dòng chảy lớn kèm bùn đá trên các sườn dốc (lưu vực, sông suối), xảy ra nhanh trong khoảng thời gian ngắn, bất ngờ và gây ra những tàn phá nghiêm trọng đối với tự nhiên, dân cư và cơ sở hạ tầng”.

1.1.2 Cơ chế hình thành lũ quét

- Lũ quét chỉ có thể xảy ra khi có sự kết hợp của nhiều yếu tố như mưa lớn trong một thời gian ngắn trên các lưu vực sông suối có điều kiện mặt đệm, và bề mặt địa hình thuận lợi (sườn dốc từ 15-30%) cho việc hình thành LQ.

1.2 Tổng quan các nghiên cứu lũ quét trên thế giới và ở Việt Nam

1.2.1 Tổng quan các nghiên cứu lũ quét trên TG

- Các hướng nghiên cứu sử dụng mô hình thống kê, mô hình thủy văn và mô hình học máy và tích hợp GIS, VT và mô hình học máy. Có thể liệt kê một số nghiên cứu điển hình như trong bảng sau:

| Tác giả | Phương pháp | Dữ liệu/ Yếu tố | Kết quả |
|----------------------------|--|---|---|
| Bui et al (2019) | Phương pháp lựa chọn các đối tượng (FSM) theo quy tắc tập mờ dựa vào thuật toán tập mờ quy nạp không theo thứ tự (FURIA), như thuộc tính người đánh giá trong khi thuật toán lan truyền (GA) sử dụng để tìm kiếm nhằm thu được tập biến tối ưu sử dụng trong đánh giá mô hình nguy cơ (nhảy cảm) lũ quét. Thuật toán FURIA-GA là thuật toán của sự kết hợp giữa các thuật toán LogiBoost, Bagging và AdaBoost. | 654 điểm lũ được xác định cùng với 12 yếu tố lũ được lựa chọn như độ cao, độ dốc, hướng sườn, độ cong địa hình, TWI, SPI, toposhade, mật độ sông suối, lượng mưa, NDVI, lớp phủ và thạch học. | FURIA-GA-Bagging – 93.37% FURIA-GA-LogiBoost – 92.35% FURIA-GA-AdaBoost -89.03% |
| Youssef Hegab (2019) | AHP | 232 điểm lũ được xác định, 7 yếu tố lũ gồm độ dốc, độ cao, độ cong địa hình, địa mạo, TWI, thổ nhưỡng và khoảng cách đến sông suối | AHP – 83.3% |
| Ngo et al (2018) | Thuật toán đom đóm (FA), lan truyền ngược Levenberg-Marquardt (LM) và mạng nơ ron | 654 điểm lũ quét được xác định, 12 yếu tố được lựa chọn gồm: độ cao, độ dốc, hướng sườn, độ | FA-LM-ANN -97.0% LM-ANN – 92.6% |

| | | | |
|-----------------------|---|---|--|
| | nhân tạo (FA-LM-ANN), LM-ANN, FA-ANN, SVM và cây phân loại (CT) | cong địa hình, TWI, SPI, toposhade, mật độ sông suối, lượng mưa, NDVI, thổ nhưỡng, và địa mạo | FA-ANN-91.9% SVM-92.9% CT-90.8% |
| Khosravi et al (2018) | Các thuật toán cây quyết định như Logistic Model Trees (LMT), Reduced Error Pruning Trees (REPT), Naïve Bayes Trees (NBT), and Alternating Decision Trees (ADT) | 201 điểm lũ hiện tại và quá khứ được xác định, 11 yếu tố được lựa chọn như độ dốc, độ cao, độ cong địa hình, SPI, TWI, sử dụng đất, lượng mưa, mật độ sông suối, khoảng cách đến sông suối, địa mạo và NDVI | ADT - 97.6% NBT- 97.4% LMT – 97.1% REPT – 81.1% |
| Khosravi et al (2016) | Shannon’s entropy, chỉ số thống kê, mô hình trọng số nhân tố | 211 điểm lũ được xác định, 10 yếu tố được lựa chọn gồm độ dốc, độ cong địa hình, độ cao, TWI, SPI, khoảng cách đến sông suối, lượng mưa, địa mạo, sử dụng đất và NDVI | SI – 98.72% WF – 98.1% SE – 92.53% |
| Youssef et al (2016) | Các mô hình thống kê nhị biến và đa biến (tỷ số tần suất, tích hợp tỷ số tần suất và hồi quy logic) | 127 điểm lũ được xác định, 7 yếu tố được lựa chọn gồm độ dốc, độ cao, độ cong địa hình, địa mạo, sử dụng đất, thổ nhưỡng, khoảng cách đến sông suối | FR – 89.6 % FR+LR -91.3% |

Trên cơ sở các nghiên cứu đã được tổng quan, khả năng tích hợp giữa GIS và các mô hình cho kết quả tốt trong việc xây dựng mô hình nguy cơ lũ quét. Sự kết hợp của các mô hình học máy có thể làm tăng độ chính xác của kết quả. Kết quả nghiên cứu cũng chỉ ra rằng, phương pháp học máy có độ chính xác cao hơn so với phương pháp định tính và định lượng (WoE, FR và AHP). Khả năng tích hợp giữa GIS, viễn thám và các mô hình khác đã chứng minh được là có thể tạo ra mô hình nguy cơ lũ quét.

1.2.2 Tổng quan các nghiên cứu lũ quét ở Việt Nam

Có nhiều nghiên cứu lũ quét được thực hiện như Cao Đăng Dur (2006), Nguyễn Ngọc Thạch (2013), Ưông Đình Khanh (2011), Phạm Thị Hương Lan và nnk (2008). Các phương pháp chủ yếu sử dụng phương pháp AHP và trọng số kinh nghiệm và lũ quét được áp dụng là lũ quét sườn dốc kết hợp với phân tích GIS để thành lập bản đồ phân vùng nguy cơ lũ quét. Về quy mô, nhìn chung các nghiên cứu lũ quét ở VN vẫn tập trung nghiên cứu ở tỷ lệ nhỏ (cấp tỉnh) nên độ chính xác của bản đồ nguy cơ còn thấp. Do vậy việc đánh giá cho các vùng có diện tích nhỏ, các lưu vực nhỏ (miền núi) cần phải được quan tâm nhiều hơn. Đặc biệt chưa có nghiên cứu lũ quét nào với tỷ lệ lớn cho lưu vực suối Muội tỉnh Sơn La.

1.2.3 Tổng quan về các phương pháp cảnh báo sớm lũ quét

Trên thế giới có một số nghiên cứu về cảnh báo sớm lũ quét điển hình như Matkan, A và nnk (2009), Forestieri, A và nnk (2016), Hoedjes, J.C và nnk (2014), Jubach, R và nnk (2016), Wang, Y và nnk (2017) đã sử dụng các phương pháp thủy văn, thủy lực, ngưỡng mưa và các hướng dẫn dự báo lũ quét trong nghiên cứu cảnh báo sớm lũ quét. Đặc biệt hệ thống cảnh báo lũ quét ALERT của tổ chức Khí tượng thế giới (WMO), hệ thống cảnh báo FFGS. Hệ thống này tập

trung phát triển và áp dụng kỹ thuật hiệu chỉnh mưa từ vệ tinh (hay radar) và mô hình thủy văn.

Ở Việt Nam, hệ thống cảnh báo sớm lũ quét vẫn đang còn chưa phát triển. Hiện tại mới chỉ có hệ thống cảnh báo lũ quét VNFFGS đang thực hiện của Viện khoa học Khí tượng Thủy văn và biến đổi khí hậu. Các nghiên cứu về cảnh báo sớm lũ quét đang còn bỏ ngỏ và chưa có một công bố nào khẳng định chắc chắn việc dự báo lũ quét là chính xác tuyệt đối.

1.3 Cách tiếp cận và phương pháp nghiên cứu

1.3.1 Cách tiếp cận

- Tiếp cận lịch sử, tiếp cận kế thừa, tiếp cận tổng hợp, tiếp cận mô hình, tiếp cận công nghệ cao.

1.3.2 Phương pháp nghiên cứu

- Phương pháp điều tra, thu thập số liệu, tài liệu; Phương pháp phân tích thống kê; Phương pháp phân tích và xử lý dữ liệu; Phương pháp xin ý kiến chuyên gia; Phương pháp học máy

Chương 2. PHƯƠNG PHÁP NGHIÊN CỨU VÀ DỮ LIỆU

2.1. Đặc trưng các nhân tố hình thành lũ quét

Trên cơ sở nghiên cứu các tài liệu, dữ liệu thu thập, các nghiên cứu tổng quan về lũ quét, đặc trưng về điều kiện địa lý ở lưu vực suối Muội và phương pháp nghiên cứu, luận án đã đề xuất thành 03 nhóm yếu tố chính liên quan trực tiếp đến nghiên cứu nguy cơ lũ quét:

2.1.1. Nhóm các nhân tố về điều kiện tự nhiên

- Độ cao: địa hình vùng núi Việt Nam nói chung rất dốc, đặc biệt vùng núi phía Tây Bắc do đó độ dốc lòng sông lớn, đó là một trong những điều kiện thuận lợi để phát sinh lũ quét. Ở những nơi có địa hình núi cao như lưu vực suối Muội thường là nơi có lượng mưa lớn

và phân hóa rất mạnh. Các lưu vực sinh lũ thường nhỏ, sông suối bắt nguồn từ các đỉnh núi cao (1000-2000m).

- Độ dốc và độ cong địa hình: Qua khảo sát các khu vực bị lũ quét cho thấy: Các lưu vực đã xảy ra lũ quét thường ở nơi có dạng đường cong lõm, địa hình bị chia cắt dữ dội, sườn núi rất dốc (>30 độ). Độ dốc lòng sông ở đầu nguồn rất lớn, tạo điều kiện thuận lợi hình thành lũ quét. Mặt cắt dọc sông nhiều nơi có điểm gãy mà sau này là vùng thường bị lũ quét ác liệt.

- Mạng lưới thủy văn: Địa hình chia cắt tạo nên mạng lưới sông suối dày đặc. Ở vùng đầu nguồn, nhiều nơi mật độ sông suối lớn hơn 1km/km², thậm chí tới 2km/km². Độ dốc lòng sông, suối lớn nên thời gian tập trung dòng chảy ngắn, tốc độ dòng chảy lớn, năng lượng, sức tải lớn làm gia tăng khả năng lũ quét.

- Địa mạo: Sông suối chảy giữa những kẽ núi, mặt cắt ngang thường có dạng chữ V hoặc chữ U sâu và hẹp. Chảy qua các bậc thềm địa hình, mặt cắt dọc sông, suối thay đổi phức tạp kéo theo sự thay đổi mặt cắt ngang. Nơi thu hẹp, sông sâu thẳng, nơi mở rộng ở các thung lũng, có bãi tràn rộng, thường có các điểm quần cư, phát triển kinh tế mạnh cũng chính là vùng chịu tác động mạnh mẽ của lũ quét.

- Trượt lở đất đá: hình thái lưu vực có tính chất quan trọng góp phần vào quá trình sinh lũ quét. Các lưu vực có diện tích nhỏ (diện tích < 500km²), lưu vực có hình rẻ quạt hoặc tròn, xung quanh có núi cao bao bọc, có hướng thuận lợi đón gió ẩm hình thành những tâm mưa. Sườn dốc được phủ bởi lớp đất đá có độ liên kết kém, dễ xói mòn, sụt lở. Khi có mưa lớn, lũ quét kéo theo nhiều vật rắn như đá, sỏi, cây cối.

2.1.2 Nhóm các nhân tố về điều kiện KTXH

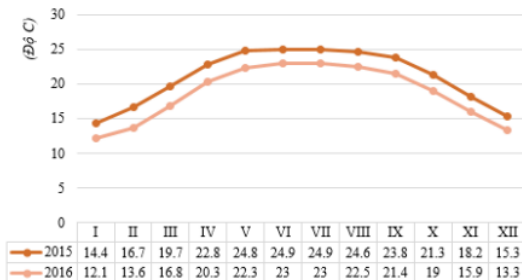
- Rừng, lớp phủ thực vật là những yếu tố biến đổi chậm. Song do tác động của con người, sự suy thoái đến một “ngưỡng” mà vai trò lá chắn của rừng không còn nữa, có nghĩa độ bám đất và khả năng trữ nước của các yếu tố này không còn cao, tổ hợp với các điều kiện khác làm lũ quét xuất hiện nhiều hơn. Kết quả nghiên cứu trong và ngoài nước đều đi đến nhận định rừng có tác dụng điều tiết dòng chảy mặt và dòng chảy lũ.

- Tốc độ đô thị hóa cao đồng nghĩa với quá trình làm thay đổi lớp phủ sử dụng đất khi con người xây dựng các công trình nhà cửa, đường xá... có ảnh hưởng rất lớn đến khả năng thấm nước của bề mặt đất đặc biệt là làm tăng lượng dòng chảy mặt có nghĩa khả năng xảy ra lũ quét trên các khu vực là rất lớn.

2.1.3 Nhóm các nhân tố về khí tượng

Nằm trong vùng khí hậu B1-Vùng khí hậu Tây Bắc, thuộc tiểu vùng khí hậu Nam Tây Bắc (B1-2), khí hậu Thuận Châu mang sắc thái nhiệt đới gió mùa có mùa đông lạnh của vùng núi cao và cao nguyên, một trong những kiểu khí hậu nhiệt đới gió mùa khá đặc biệt ở nước ta.

a. Chế độ nhiệt: chế độ nhiệt cũng bị phân theo hai mùa. Mùa đông lạnh khô kéo dài từ tháng mười đến tháng ba năm sau. Tháng giêng là tháng lạnh nhất. Mùa hè nóng ẩm, thường kéo dài từ tháng tư đến tháng chín. Nóng nhất là tháng sáu và tháng bảy.



Hình 1.2 Nhiệt độ trung bình năm (độ C)

b. Chế độ ẩm: Chế độ mưa - ẩm vùng núi Sơn La phân thành hai mùa Mùa đông là mùa ít mưa, lượng mưa dưới 10% lượng mưa cả năm, độ ẩm không khí thấp, xuống tới 75-76%. Tháng mười hai và tháng một là những tháng có lượng mưa nhỏ nhất.

- Lưu vực suối Muội là lưu vực nhỏ có địa hình đặc biệt sườn núi chắn gió và các thung lũng có tác dụng hút luồng không khí ẩm từ biển vào do vậy khi mưa với cường độ rất lớn tập trung trong vài giờ là nguyên nhân gây ra lũ quét. Mưa lớn còn là động lực chủ yếu gây ra xói mòn, sạt lở tạo thành phần rắn của dòng lũ quét.

2.2. Lựa chọn bộ tiêu chí cho nghiên cứu nguy cơ lũ quét

- Trên cơ sở kế thừa các nghiên cứu, ý kiến các chuyên gia và phân tích các nhóm nhân tố chính ảnh hưởng đến nguy cơ lũ quét trên lưu vực suối Muội tỉnh Sơn La để lựa chọn bộ tiêu chí:

| Các nhóm nhân tố | Chỉ tiêu lựa chọn | Nguồn dữ liệu |
|--------------------|-----------------------------------|--|
| Điều kiện tự nhiên | Độ cao | Bản đồ địa hình tỷ lệ 1:10000 |
| | Độ dốc | Nội suy từ DEM |
| | Hướng sườn | Nội suy từ DEM |
| | Mật độ sông suối | Nội suy từ DEM |
| | Khoảng cách đến sông suối | Nội suy từ dữ liệu sông suối |
| | Chỉ số năng lượng dòng chảy (SPI) | Nội suy từ DEM |
| | Chỉ số ẩm ướt (TWI) | Nội suy từ DEM |
| | Độ cong địa hình | Nội suy từ DEM |
| | Địa mạo | Xây dựng theo nguyên tắc hình thái nguồn gốc |
| | Mật độ trượt lở | Nội suy từ các điểm TL |

| | | |
|-------------------|---------------------------------|--|
| | Chỉ số thực vật (NDVI) | Tính toán từ ảnh vệ tinh Sentinel 2B ngày 18/12/2019 |
| | Rừng | Viện điều tra và quy hoạch rừng |
| Điều kiện KTXH | Thổ nhượng | Viện thổ nhượng nông hóa – Bộ NN&PTNT |
| | Mức độ đô thị hóa (NDBI) | Tính toán từ ảnh vệ tinh Sentinel 2B ngày 18/12/2019 |
| Điều kiện khí hậu | Lượng mưa trung bình 1 ngày max | Số liệu quan trắc của các trạm khí tượng thủy văn |

2.3 Các mô hình nghiên cứu lũ quét

2.3.1 Mô hình phân tích đa chỉ tiêu MCA và phương pháp AHP

Các bước cơ bản:

B1: Xác định các thông số đầu vào (các yếu tố ảnh hưởng đến lũ quét)

B2: Đánh giá mức độ quan trọng của các yếu tố ảnh hưởng tới nguy cơ lũ quét

B3: Tính toán (hoặc gán) trọng số cho các yếu tố ảnh hưởng

B4: Chồng xếp các lớp nhân tố và đánh giá độ chính xác

Lựa chọn biến đầu vào hay xác lập mối liên quan giữa các nhân tố với LQ để có thể thu được kết quả tốt nhất điều đó phụ thuộc rất nhiều vào các biến nhân tố. Sau khi lựa chọn các nhân tố tiến hành phân tích, đánh giá tổng hợp các nguồn số liệu, dữ liệu để đưa vào GIS.

2.3.2 Mô hình tích hợp Gradient Boost và tối ưu hóa Bayesian

a. Thuật toán XGBoost (Extreme Gradient Boosting)

XGBoost nhận đầu vào là tabular datasets với mọi kích thước và dạng dữ liệu bao gồm cả categorical mà dạng dữ liệu này thường được tìm thấy nhiều hơn trong business model, đây là lý do đầu tiên tại sao các cá nhân tham gia Kaggle thường sử dụng.

Ở đây, ta áp dụng hướng tiếp cận “non-parametric” với phép tối ưu hoá Steepest-descent cho không gian hàm số thay vì trong không gian tham số.

b. Tối ưu hóa Bayesian

Các bước thực hiện:

- Lựa chọn giá trị cho các tham số như learning rate, epochs, số layers, hidden units,...

- Huấn luyện mô hình, quan sát kết quả đạt được, đánh giá kết quả, điều chỉnh tham số và lặp lại.

Bayesian Optimization (BO) là một thuật toán giúp tối ưu hiệu quả những hàm mục tiêu có chi phí đánh giá lớn (như training 1 mạng neural) dựa trên định lý Bayesian. BO làm giảm đáng kể số lần thử sai khi tune tham số so với Grid Search hay Random Search, điều mà với một mạng học sâu lớn (ResNet, Inception, Xception,..) cùng bộ data khổng lồ (ImageNet) sẽ tốn hàng giờ thậm chí hàng ngày liền. BO tối ưu hàm mục tiêu dựa trên học máy. Viết dưới dạng công thức:

$$\max_{x \in A} f(x) \quad (2.1)$$

Trong đó A là tập các tham số cần tune

$f(x)$ là hàm mục tiêu để tối ưu lớn nhất/nhỏ nhất

2.4 Mô hình cảnh báo nguy cơ LQ

- Bản đồ cảnh báo của hệ thống sẽ được xây dựng thông qua hàm tính sau:

$$CBLQ = \sqrt{NCLQ * NSLM} \quad (2.23)$$

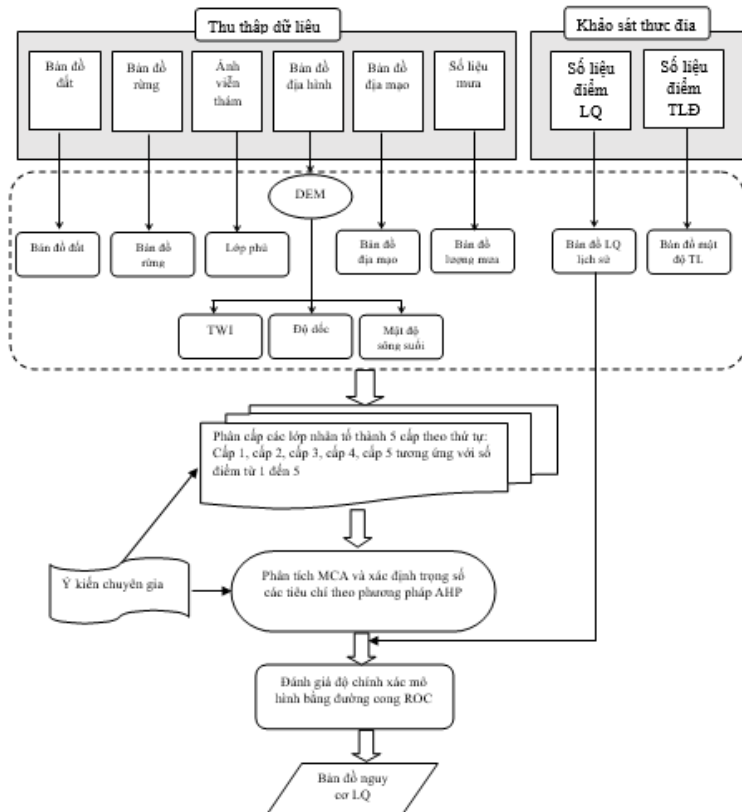
Trong đó: NCLQ - Nguy cơ lũ quét tiềm ẩn (hay FFPI)

NSLM - Dữ liệu phân cấp lượng mưa dự báo theo ngưỡng (số liệu mưa dự báo thu nhận từ trạm quan trắc).

Chương 3. KẾT QUẢ NGHIÊN CỨU VÀ THẢO LUẬN

3.1. Xây dựng bản đồ nguy cơ lũ quét tiềm ẩn áp dụng mô hình MCA và phương pháp AHP

3.1.1 Quy trình thực hiện



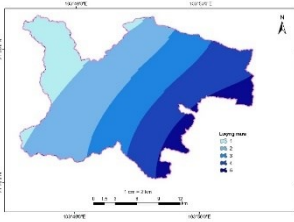
Hình 3.1 Quy trình xây dựng bản đồ nguy cơ LQ theo phương pháp MCA.

3.1.2. Đánh giá các lớp thông tin bản đồ cho nguy cơ lũ quét

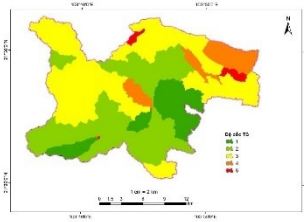
Bảng 3.2. Chỉ tiêu đánh giá nguy cơ LQ cho các lớp thông tin địa mạo (GM), thổ nhưỡng (S), rừng (FR), lượng mưa 1 ngày max (RM), chỉ số ẩm ướt địa hình (TWI), độ dốc (SB), mật độ trượt lở (LS), mật độ sông suối (RD)

| Mức ĐG | FR | GM | S | TWI | LS | SB (độ) | RM (mm) | RD (km ² /mm ²) |
|--------------|--|--|---|----------|-----------|---------|---------|--|
| Cấp 1 | - Rừng hỗn giao tre nứa | - Bề mặt đỉnh cao trên 1000m dạng lượn sóng, bóc mòn tổng hợp yếu -Bề mặt đỉnh có độ cao dưới 1000m, dạng vòm thoải, rửa trôi bề mặt | - Núi đá | -1.5-1.0 | 0.01-0.02 | >25 | <250 | 0.0-0.5 |
| Cấp 2 | - Rừng lá rộng thường xanh trung bình - Rừng tre lú | -Sườn vách dốc rửa lữa trên đá vôi, dốc, rửa lữa, đồ lở - Sườn vách bóc mòn rửa lữa trên đá vôi, vôi xen; rất dốc, đên dốc đứng, chân sườn tích tụ thoải hơn; rửa lữa, đồ lở | -Đất đỏ nâu trên đá vôi -Đất nâu đỏ trên đá macma bazơ và trung bình | 1.0-1.5 | 0.02-0.04 | 21-25 | 250-300 | 0.5-1.5 |
| Cấp 3 | - Rừng lá rộng thường xanh nghèo - Rừng lá rộng tường xanh phục hồi - Rừng trên núi đá - Rừng trồng | - Sườn trượt lở khá mạnh; dạng hơi lõm, hoặc khúc khuỷu, phần trên dốc 25-30 độ, phần dưới 30-40 độ; đồ lở trên đá gốc; trượt lở. -Sườn bóc mòn trên các đá khác nhau; dạng hơi lồi, dốc trung bình 15-25 độ; đôi nơi 20-25 độ, bóc mòn tổng hợp - Sườn trượt lở trung bình; dạng khá thẳng, hoặc hơi lõm, dốc trung bình 25-30 độ, trượt lở trên vỏ phong hóa dày, trượt trôi trên đá gốc | - Đất vàng trên đá cát kết | 1.5-3.5 | 0.04-0.06 | 18-21 | 300-350 | 1.5-2.5 |
| Cấp 4 | - Dân cư - Đất khác | - Bề mặt tích tụ lũ tích, sườn tích; dải ven chân sườn, độ dốc 6-15 độ, thành phần hỗn độn; rửa trôi bề mặt, xâm thực rãnh xói, mương xói | - Đất đỏ vàng biến đổi do trồng lúa nước - Đất đỏ vàng trên đá sét | 3.5-5.5 | 0.06-0.08 | 15-18 | 350-450 | 2.5-3.5 |
| Cấp 5 | - Đất trống - Mặt nước | - Máng trũng dòng chảy xâm thực; trắc diện ngang chữ V; trắc diện dọc thừng hoặc có dạng bậc thể hiện các thềm xâm thực; xâm thực sâu mạnh - Máng trũng xâm thực -tích tụ; trắc diện ngang chữ V, chữ U hẹp; xâm thực tích tụ. | - Đất thung lũng do sản phần dốc tụ - Đất nâu vàng trên phù sa cổ | 5.5-19.4 | 0.08-0.1 | >15 | >450 | 3.5-4.9 |

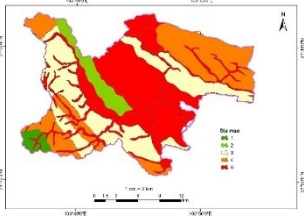
Kết quả đánh giá



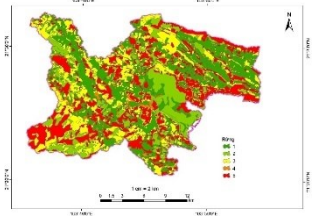
a.Đánh giá lượng mưa 1 ngày max



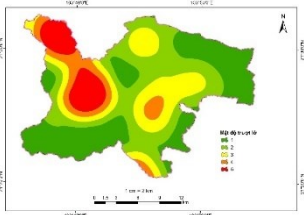
b.Đánh giá độ dốc trung bình



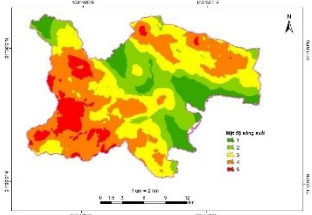
c.Đánh giá theo địa mạo



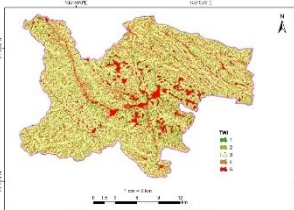
d. Đánh giá lớp phủ rừng



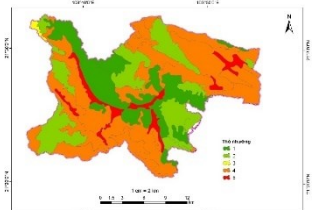
e. Đánh giá theo mật độ trượt lở



f. Đánh giá theo mật độ sông suối



g. Đánh giá theo chỉ số ẩm ướt địa hình



h. Đánh giá theo thổ nhưỡng

Hình 3.4 Các lớp thông tin đánh giá cho bản đồ nguy cơ LQ

Bảng 3.4. Trọng số cho các lớp nhân tố

| | TWI | FR | LS | SB | RM | GM | S | RD | Tổng | Trọng số |
|------------|------|------|------|------|------|------|------|------|------|-------------|
| TWI | 0.33 | 0.26 | 0.23 | 0.25 | 0.45 | 0.34 | 0.18 | 0.22 | 0.27 | 0.28 |
| FR | 0.11 | 0.09 | 0.14 | 0.17 | 0.06 | 0.06 | 0.16 | 0.13 | 0.91 | 0.11 |
| LS | 0.07 | 0.03 | 0.05 | 0.03 | 0.04 | 0.04 | 0.11 | 0.09 | 0.44 | 0.05 |
| SB | 0.11 | 0.04 | 0.14 | 0.08 | 0.08 | 0.06 | 0.13 | 0.14 | 0.77 | 0.10 |
| RM | 0.16 | 0.35 | 0.28 | 0.25 | 0.23 | 0.34 | 0.18 | 0.22 | 2.02 | 0.25 |
| GM | 0.11 | 0.18 | 0.14 | 0.17 | 0.08 | 0.11 | 0.11 | 0.13 | 1.02 | 0.13 |
| S | 0.05 | 0.01 | 0.01 | 0.02 | 0.03 | 0.03 | 0.03 | 0.01 | 0.19 | 0.02 |
| RD | 0.07 | 0.03 | 0.02 | 0.03 | 0.05 | 0.04 | 0.11 | 0.04 | 0.38 | 0.05 |

Ta có $CR = 0.06 < 0.1$ đảm bảo độ nhất quán.

3.1.4. Thành lập bản đồ nguy cơ lũ quét tiềm ẩn

Bản đồ nguy cơ lũ quét tiềm ẩn được thành lập theo hàm sau:

$$FFPI = \sum_{i=1}^n w_i x_i \quad (3.1)$$

Trong đó: FFPI - bản đồ nguy cơ lũ quét tiềm ẩn (Potential Flash flash Index)

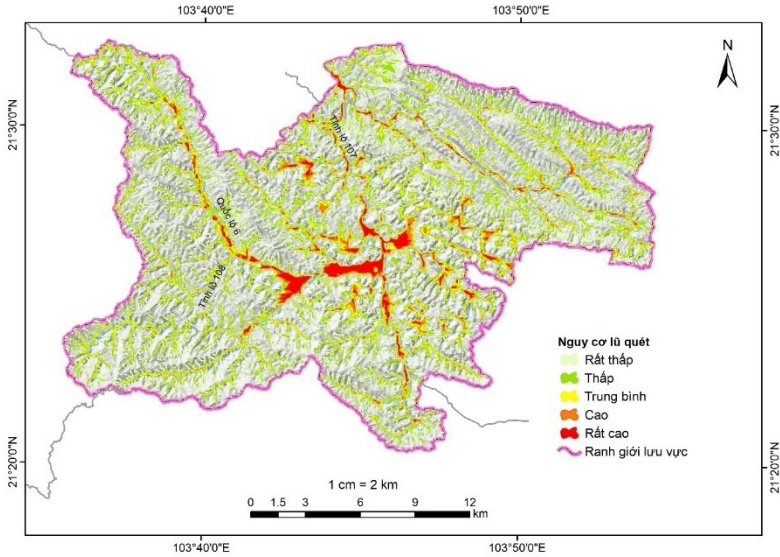
w_i - trọng số của yếu tố thứ (i);

x_i - yếu tố thứ (i);

n - số yếu tố (từ 1 đến n - các lớp thông tin đã nêu ở trên).

Từ kết quả tính toán trọng số kết hợp với bản đồ đánh giá ảnh hưởng áp dụng công thức sau để tìm ra bản đồ nguy cơ lũ quét tiềm ẩn:

$$FFPI = 0,28 * TWI + 0,11 * FR + 0,05 * LS + 0,1 * S + 0,25 * RD + 0,13 * GM + 0,02 * SB + 0,05 * DS \quad (3.2)$$

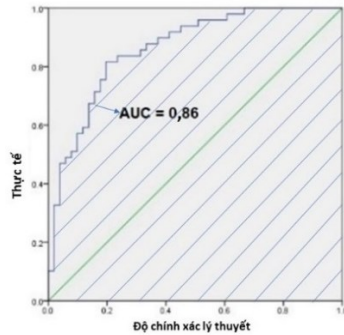
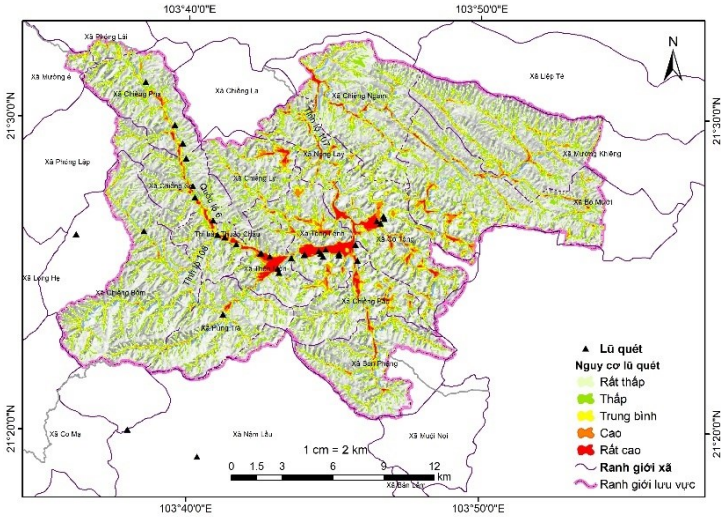


Hình 3.5 Bản đồ nguy cơ lũ quét tiềm ẩn lưu vực suối Muội tỉnh Sơn La áp dụng phương pháp MCA

3.1.5 Đánh giá độ chính xác mô hình

Để kiểm tra đánh giá độ tin cậy của kết quả nghiên cứu, phương pháp xây dựng đường cong thu nhận (receiver operating characteristic hay receiver operating curve-ROC) đã được áp dụng.

Giá trị dưới đường cong (Area Under Curve) càng cao thì sự phân biệt giữa hai trạng thái tin cậy (giá trị 1) và không tin cậy (giá trị 0) là càng tốt. Khi giá trị AUC tăng lên thì đường cong sẽ tiến sát đến trục ngang phía trên có $AUC = 1$. Việc xác định mức độ tin cậy dựa trên hệ thống điểm sau: 0,70 - 0,90 = rất tốt (A), 0,60-0,70 = tốt (B), 0,50-0,60 = sai (C).

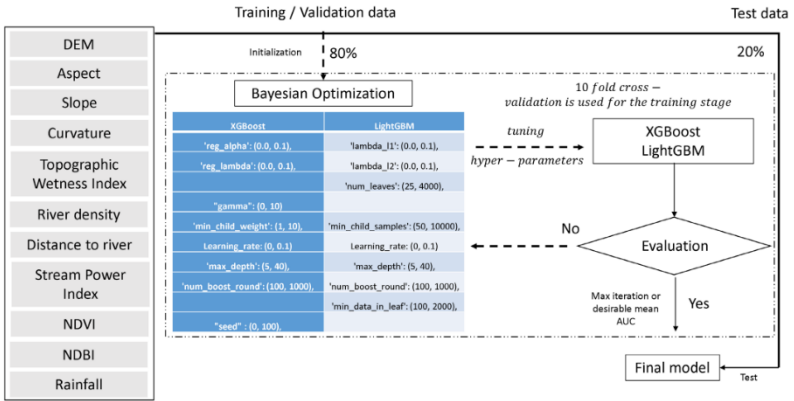


Hình 3.6 Đánh giá độ chính xác mô hình theo các điểm lũ lịch sử

So sánh số liệu các điểm lũ quét đã xảy ra trong quá khứ và bản đồ nguy cơ lũ quét (hình 3.3), đường cong ROC được xây dựng trong hình 3.6 cho giá trị $AUC = 0,86$ tương đối phù hợp với mô hình.

3.2 Thành lập bản đồ nguy cơ lũ quét theo mô hình tích hợp Gradient Boost và tối ưu hóa Bayesian (huấn luyện mô hình dựa trên bộ dữ liệu sẵn có tại Lai Châu)

3.2.1 Quy trình nghiên cứu



Hình 3.7 Quy trình nghiên cứu áp dụng mô hình học máy ML sử dụng thuật toán tối ưu hóa Bayesian

Tóm tắt các bước thực hiện:

B1. Lựa chọn các biến đầu vào (sử dụng bộ dữ liệu mẫu).

11 nhân tố được đưa vào để chạy mô hình gồm (1) độ cao, (2) hướng sườn, (3) độ dốc, (4) độ uốn cong địa hình, (5) TWI/CTI, (6) NDBI, (7) NDVI, (8) mật độ sông suối, (9) khoảng cách đến sông suối, (10) năng lượng địa hình (SPI), (11) lượng mưa.

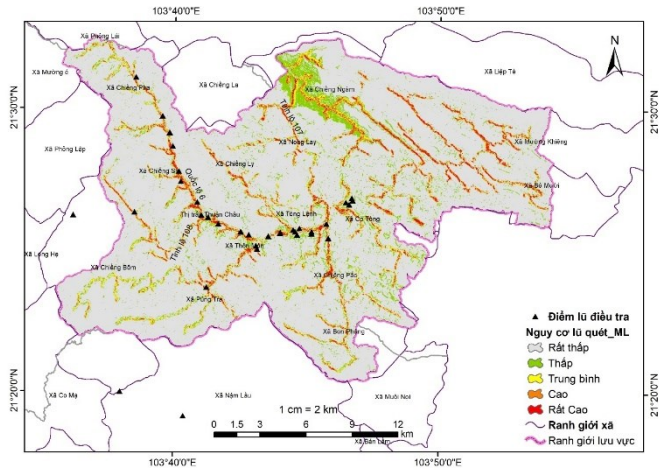
B2. Điều chỉnh tham số và chạy mô hình.

Thực hiện huấn luyện riêng các tham số cho từng thuật toán XGBoost và LightGBM nhằm chọn lựa ra các siêu tham số.

Bảng 3.3 Giá trị các siêu tham số sau quá trình tối ưu

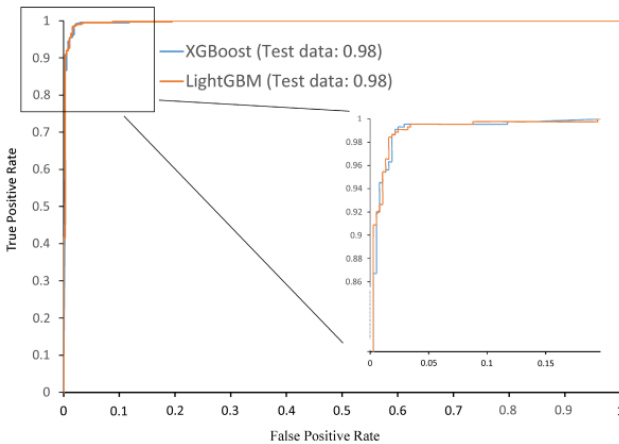
| | AUC for the training set | AUC for test set | Optimal parameter values |
|----------|--------------------------|------------------|--|
| XGBoost | 0.997 | 0.9819 | 'gamma': 1.4818174802938833, 'learning_rate': 0.09819880354454832, 'max_depth': 16.975483858905605, 'min_child_weight': 8.046930170244275, 'num_boost_round': 746.357382107532, 'reg_alpha': 0.07251418997564782, 'reg_lambda': 0.0868171524919831, 'seed': 35.0089988693958 |
| LightGBM | 0.996 | 0.9819 | lambda_l1 = 0.030527340559876517, lambda_l2 = 0.07508630572031713, 'learning_rate': 0.01435, max_depth = 35, min_child_samples = 2471, min_data_in_leaf = 102, num_boost_round = 916, num_leaves = 975 |

3.2.2 Kết quả mô hình



Hình 3.9. Bản đồ nguy cơ lũ quét cho lưu vực suối Muội tỉnh Sơn La áp dụng mô hình học máy (thuật toán tối ưu Bayes)

3.2.3 Đánh giá độ chính xác mô hình



Hình 3.10. Biểu đồ AUC của XGBoost và LightGBM với các siêu tham số được xác định từ quá trình tối ưu hóa

Bảng 3.3 Kiểm định thống kê với Wilcoxon tại mức 0.05

| Kiểm định thống kê với Wilcoxon tại mức 0.05 | Giá trị p |
|---|-----------|
| XGBoost turn bởi BO với Xgboost (mặc định các siêu tham số) | 0.001953 |
| LightGBM turn bởi BO với LightGBM (mặc định các siêu tham số) | 0.001953 |
| Xgboost turn bởi BO với LightGBM turn bởi OA | 0.1602 |

3.3 Mô hình cảnh báo sớm nguy cơ lũ quét

Nguyên lý: lũ quét sẽ chỉ xảy ra ở những khu vực có nguy cơ cao và có lượng mưa vượt ngưỡng phát sinh lũ. Cảnh báo sớm lũ quét là sự báo trước khả năng lũ quét có thể xảy ra trong thời gian ngắn tới ở một vị trí nào đó hoặc trên một lưu vực nhằm giảm tối thiểu thiệt hại về người và tài sản cũng như những tổn hại về kinh tế và xã hội. Cảnh báo nguy cơ căn cứ vào kết quả đánh giá, phân tích dữ liệu từ trạm thời tiết thông minh cũng như đặc điểm của lưu vực suối Muội.

Hàm tính cho mô hình cảnh báo sớm nguy cơ lũ quét:

$$Fr = \sqrt{f \cdot p} \quad (3.3)$$

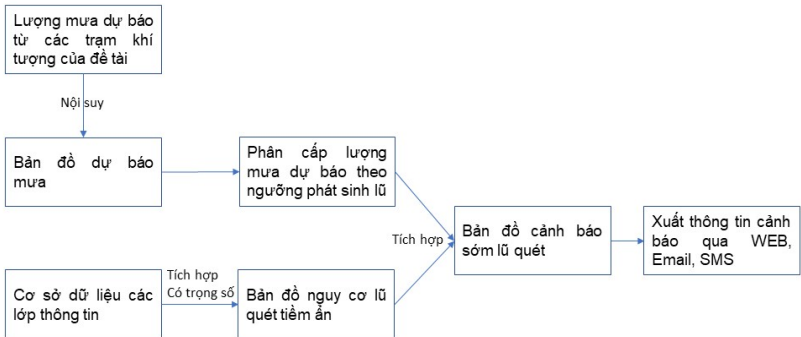
Trong đó: Fr là bản đồ cảnh báo sớm nguy cơ lũ quét

f lượng mưa cực đại dự báo

p là bản đồ nguy cơ lũ quét

Bảng 3.4 Bảng phân cấp nguy cơ xuất hiện lũ quét

| Cấp | Nguy cơ | Đặc điểm lưu vực |
|-----|------------|--|
| I | Rất cao | <ul style="list-style-type: none"> - Lượng mưa tối thiểu 100mm/ngày - Độ dốc bình quân >30% - Độ che phủ <10% - Đất có tính thấm trung bình và kém |
| II | Cao | <ul style="list-style-type: none"> - Lượng mưa tối thiểu 100mm/ngày - Kết hợp 2 trong 3 yếu tố còn lại của cấp I |
| III | Trung bình | <ul style="list-style-type: none"> - Lượng mưa tối thiểu 100mm/ngày - Kết hợp 1 trong 3 yếu tố còn lại ở cấp I |
| IV | Thấp | Không có các đặc điểm trên |



Hình 3.11 Sơ đồ tích hợp thông tin cảnh báo sớm lũ quét trên lưu vực suối Muội tỉnh Sơn La

KẾT LUẬN VÀ KIẾN NGHỊ

1. Kết luận: Trên cơ sở nghiên cứu, xử lý và phân tích kết quả, luận án đã đưa ra kết luận sau đây

- Luận án đã xác lập được cơ sở khoa học trong việc lựa chọn các yếu tố tham gia vào xây dựng bản đồ nguy cơ lũ quét trên lưu vực suối Muội tỉnh Sơn La.

- Đã tổng quát hóa được vai trò của các yếu tố trong đánh giá nguy cơ lũ quét

- Luận án đã áp dụng thành công nhiều mô hình khác nhau để xử lý thông tin nhằm nâng cao độ chính xác trong nghiên cứu lũ quét như phương pháp viễn thám, GIS, phương pháp phân tích đa chỉ tiêu MCA và AHP, mô hình học máy cụ thể là phương pháp tối ưu hóa Bayesian Phương pháp phân tích đa chỉ tiêu MCA và đánh giá trọng số AHP cho phép giải quyết nhanh các bài toán phân tích nhiều tiêu chí hỗ trợ các nhà quản lý ra quyết định nhanh chóng. Phương pháp học máy với các thuật toán tối ưu có hiệu quả trong phân tích dữ liệu khách quan và chi tiết hơn thông qua việc đào tạo dữ liệu đặc biệt là thuật toán LightGBM và Xgboost.

- Thông qua việc so sánh độ chính xác giữa các mô hình, luận án đã đánh giá được vai trò của mô hình học máy trong nghiên cứu lũ quét.

- Kết hợp VT, GIS với các thuật toán tối ưu tuy không trực tiếp giúp người dân có thể tránh được hoàn toàn tai biến thiên nhiên nhưng bản đồ nguy cơ cũng nên được coi là một tài liệu hữu ích cần phổ biến rộng rãi cho người dân nắm được nguyên lý hoạt động cũng như cách phòng chống khi mùa mưa lũ về.

2. Kiến nghị

- Cần thử nghiệm thêm các mô hình học máy khác để kiểm định chất lượng của tập dữ liệu.
- Để áp dụng rộng rãi mô hình cảnh báo lũ quét và sử dụng hiệu quả, đề nghị cần tiếp tục phát triển mô hình và nâng tỷ lệ bản đồ phân vùng lũ quét lên tỷ lệ lớn hơn 1:10000 và lớn hơn nữa. Muốn làm được điều này, cần thiết phải xây dựng, cập nhật, bổ sung với độ chính xác cao, tăng mức độ chi tiết cơ sở dữ liệu như bản đồ địa hình, địa chất, loại đất,... Đồng thời, các trạm khí tượng thủy văn cần thiết phải tăng mật độ trạm đo mưa tự động và cần thiết phải xây dựng các trạm đo mưa tại những khu vực có nguy cơ rất cao xảy ra lũ quét.
- Đồng thời để tăng cường khả năng cảnh báo và dự báo chính xác thời gian, không gian xảy ra lũ quét, nhà nước cần đầu tư thuê hoặc phát triển hệ thống vệ tinh radar chuyên dụng theo dõi mưa có thể cảnh báo chính xác thời gian, địa điểm xảy ra lũ quét trong khu vực và vùng lân cận.
- Luận án mới chỉ dừng lại ở việc xây dựng mô hình cảnh báo sớm và trong tương lai nếu có đủ các dữ liệu về lượng mưa theo thời gian thực từ các trạm khí tượng trong khu vực nghiên cứu kết hợp với bản đồ nguy cơ lũ quét từ kết quả của luận án sẽ có thể cảnh báo sớm được nguy cơ lũ quét.

DANH MỤC CÁC CÔNG TRÌNH KHOA HỌC ĐÃ CÔNG BỐ CỦA TÁC GIẢ LIÊN QUAN ĐẾN LUẬN ÁN

1. **Lại Tuấn Anh**, Nguyễn Ngọc Thạch, Phạm Xuân Cảnh, Lê Như Nga, Vũ Đăng Cường, “Xây dựng hệ thống cảnh báo sớm lũ quét ở vùng núi thử nghiệm tại huyện Thuận Châu, Sơn La”, Tạp chí Khoa học & Công nghệ Việt Nam, số 60, 08/2018, pp.28-35, ISSN 2525-2518.
2. **Lại Tuấn Anh**, “Ứng dụng công nghệ GIS thành lập bản đồ phân chia lưu vực sông phục vụ cảnh báo nguy cơ lũ quét cho huyện Thuận Châu tỉnh Sơn La”, Tuyển tập hội nghị Khoa học thường niên năm 2020, Trường Đại học Thủy Lợi, 011/2020, pp.463-465, ISBN:978-604-82-3869-8.
3. Tien-Yin Chou, Thanh-Van Hoang, Yao-Min Fang, Quoc-Huy Nguyen, **Tuan Anh Lai**, Van-Manh Pham, Van-Manh Vu, Quang-Thanh Bui, “Swam-based optimizer for convolutional neural network: An application for flood susceptibility mapping”, Transactions in GIS, 2020, pp. 1-18. ISSN 1467-9671.
4. **Tuan Anh Lai**, Ngoc-Thach Nguyen, Quang-Thanh Bui, “Hyperparameter optimization of gradient booster for flood susceptibility analysis”, Transactions in GIS, 2022, pp. 1-15. ISSN 1467-9671.