BỘ CÔNG THƯƠNG
TRƯỜNG ĐẠI HỌC CÔNG NGHIỆP THÀNH PHỐ HỒ CHÍ MINH

HUỲNH CÔNG LỰC

NGHIÊN CỨU ĐÁNH GIÁ HẠN HÁN LƯU VỰC SÔNG SRÊPỐK BẰNG PHƯƠNG PHÁP VIỄN THÁM VÀ HỌC MÁY

Ngành: QUẢN LÝ TÀI NGUYÊN VÀ MÔI TRƯỜNG

Mã ngành : 9850101

TÓM TẮT LUẬN ÁN TIẾN SĨ

THÀNH PHỐ HỒ CHÍ MINH, NĂM 2026

DANH MỤC CÔNG TRÌNH CÔNG BỐ CỦA TÁC GIẢ

*** Bài báo quốc tế:**

1. **Luc, Huynh Cong**, and Luong Van Viet*. "Early forecasting of vegetation health index in the Srepok basin based on meteorological and hydrological drought indices," *Environmental Research Communications*, vol. 10, no. 3, 2025. https://doi.org/10.1088/2515-7620/add861 **(ISI-Q1)**
2. **Luc, Huynh Cong*,** Luong Van Viet, and Bui Dang Hung. "Determining suitable meteorological drought and vegetation indices for monitoring drought and crop yield in Srepok River basin, Vietnam". *Agriculture and Natural Resources* , vol. 059, no. 2, 2025. https://doi.org/10.34044/j.anres.2025.59.2.09 **(Scopus-Q3)**
3. **Luc, Huynh Cong**, & VIET, Luong Van.*. "Developing machine learning models for vegetation health index forecasting in the Srepok basin, Vietnam." *Ecological Engineering & Environmental Technology (EEET)*, vol. 26, no. 11, 2025. https://doi.org/10.12912/27197050/212962 **(Scopus-Q3)**

*** Bài báo trong nước:**

1. **Huỳnh Công Lực*;** Lương Văn Việt. Diễn biến hạn hán lưu vực Srêpốk giai đoạn 2000-2022 thông qua phân tích chỉ số sức khỏe thực vật (VHI) và chỉ số lượng mưa chuẩn hóa (SPI). *VNU Tạp chí khoa học: Các khoa học Trái đất và Môi trường*, Vol. 42, No. 1, pp. 100–115, 2026. https://doi.org/10.25073/2588-1094/vnuees.5314

*** Bài báo quốc tế đã tham gia:**

1. N. V. Hiep, N. T. T. Thao, L. V. Viet, **H. C. Luc**, and L. H. Ba, "Affecting of nature and human activities on the trend of vegetation health indices in Dak Nong province, Vietnam,"

Sustainability, vol. 15, no. 7, p. 5695, 2023. https://doi.org/10.3390/su15075695 **(ISI-Q1)**

*** Báo cáo tại hội thảo:**

1. **Huỳnh Công Lực**, "Xây dựng mô hình học máy dự báo sớm chỉ số sức khỏe thực vật (VHI) tại lưu vực sông Srêpốk." Báo cáo Seminar chuyên đề Tài nguyên nước của Viện Khoa học Công nghệ và Quản lý Môi trường, 2025

*** Đề tài Nghiên cứu khoa học:**

Huỳnh Công Lực, Đồng Phú Hảo, Lê Thị Thùy Trang, Trần Bùi Ngọc Phương, "Ứng dụng Viễn thám và GIS đánh giá hạn hán trong nông nghiệp ở lưu vực sông Srêpốk." Báo cáo tổng kết đề tài cơ sở, trường Đại học Công Nghiệp Thành phố Hồ Chí Minh, 2025

*** Giáo trình:**

Lương Văn Việt, **Huỳnh Công Lực,** Lê Bá Long. (2023). GIS trong Quản lý tài nguyên môi trường (978-604-920-203-2). Đại học Công nghiệp Thành phố Hồ Chí Minh

MỞ ĐẦU

1. Đặt vấn đề

Hạn hán là dạng thiên tai, có phạm vi ảnh hưởng rộng và ngày càng gia tăng về tần suất, cường độ trong bối cảnh biến đổi khí hậu toàn cầu, đặc biệt tại các vùng nhiệt đới gió mùa. Tại Việt Nam, các đợt hạn nặng tiêu biểu giai đoạn 1997–1998, 2004–2005, 2014–2016 và 2019–2020 đã gây thiệt hại nghiêm trọng cho sản xuất nông nghiệp, riêng đợt hạn 2014–2016 khiến 15 tỉnh phía Nam công bố tình trạng thiên tai với thiệt hại ước tính 5.572 tỷ đồng. Lưu vực sông Srêpốk là vùng sản xuất cà phê, hồ tiêu và cây công nghiệp lớn nhất cả nước đang chịu tác động đặc biệt nặng nề: đợt hạn 2014–2016 khiến hơn 170.000 ha cây trồng bị ảnh hưởng, sản lượng cà phê sụt giảm nghiêm trọng và 28.300 hộ dân thiếu nước sinh hoạt. Mặc dù công nghệ viễn thám và mô hình học máy đã chứng minh hiệu quả trong giám sát và dự báo hạn hán, lưu vực Srêpốk vẫn còn những khoảng trống chưa được giải quyết: đặc điểm hạn hán và tác động đến năng suất cây trồng chưa được phân tích hệ thống; bộ chỉ số hạn phù hợp với điều kiện sinh thái và khí hậu đặc thù chưa được xác định; và chưa có công cụ dự báo sớm dựa trên tích hợp đa nguồn dữ liệu. Xuất phát từ thực trạng đó, luận án "Nghiên cứu đánh giá hạn hán lưu vực sông srêpốk bằng phương pháp viễn thám và học máy" được thực hiện. Nhằm xây dựng khung phương pháp tích hợp viễn thám, dữ liệu quan trắc mặt đất và mô hình học máy cho giám sát và dự báo sớm hạn hán nông nghiệp trên lưu vực Srêpốk, góp phần cung cấp cơ sở khoa học phục vụ cảnh báo sớm và quản lý sản xuất nông nghiệp bền vững trong bối cảnh biến đổi khí hậu.

2. Mục tiêu nghiên cứu

2.1 Mục tiêu tổng quát

Nghiên cứu này đánh giá đặc điểm hạn hán, xác định chỉ số thực vật phù hợp trong giám sát hạn hán và xây dựng các mô hình dự báo hạn trên lưu vực sông Srêpốk dựa trên công nghệ viễn thám, dữ liệu quan trắc mặt đất

và các mô hình học máy, nhằm nâng cao hiệu quả quản lý hạn hán và góp phần ứng phó với biến đổi khí hậu.

2.2 Mục tiêu cụ thể

- Làm rõ mức độ ảnh hưởng và thời gian ảnh hưởng của hạn hán đến thảm thực vật và năng suất cây trồng trên lưu vực nghiên cứu.
- Xác định chỉ số viễn thám và chỉ số hạn khí tượng cùng quy mô thời gian phù hợp trong giám sát hạn hán, vùng nhạy cảm với hạn hán và các biến tiềm năng cho dự báo hạn hán.
- Xây dựng các mô hình học máy cho dự báo chỉ số thực vật có độ tin cậy cao và thời gian dự báo sớm từ 1–3 tháng..

3. Đối tượng và phạm vi nghiên cứu

- Giới hạn không gian:

Luận án lựa chọn lưu vực sông Srêpốk thuộc Việt Nam.

- Giới hạn thời gian:

Dữ liệu thực hiện trong luận án từ năm 2000 đến năm 2022.

- Giới hạn đối tượng:

Các thông số khí tượng, thủy văn, chỉ số thực vật từ viễn thám tại lưu vực và các mô hình học máy được ứng dụng trong đánh giá hạn hán nông nghiệp.

- Giới hạn nội dung:

Luận án tập trung nghiên cứu hạn hán nông nghiệp, bao gồm việc lựa chọn chỉ số hạn phù hợp, phân tích đặc điểm hạn hán theo không gian và thời gian, đánh giá tác động đến sản xuất nông nghiệp, và xây dựng mô hình dự báo sớm bằng học máy. Nghiên cứu không đề cập đến các ảnh hưởng của biến đổi khí hậu, thay đổi sử dụng đất và phát triển hồ chứa đến hạn hán.

4. Ý nghĩa khoa học và thực tiễn

4.1 Ý nghĩa khoa học

Luận án phát triển khung phương pháp tích hợp viễn thám, dữ liệu khí tượng thủy văn và học máy để đánh giá, dự báo hạn hán, có thể áp dụng cho các khu vực có điều kiện tương tự. Nghiên cứu làm rõ đặc điểm hạn

hán và phản ứng của thảm thực vật trước điều kiện khô hạn, đồng thời cung cấp cơ sở khoa học cho việc lựa chọn chỉ số hạn và tối ưu hóa mô hình dự báo trong điều kiện sinh thái và khí hậu nhiệt đới gió mùa.

4.2 Ý nghĩa thực tiễn

Luận án cung cấp công cụ dự báo sớm hạn hán đáng tin cậy, hỗ trợ cơ quan quản lý và người sản xuất chủ động ứng phó thông qua điều tiết nguồn nước, bố trí mùa vụ hợp lý, góp phần giảm thiểu thiệt hại nông nghiệp. Kết quả nghiên cứu có tính linh hoạt cao, có thể chuyển giao cho các khu vực nhiệt đới gió mùa khác, đóng góp trực tiếp vào quản lý tài nguyên nước, hoạch định chính sách và phát triển nông nghiệp bền vững.

5. Tính mới của đề tài

Giá trị đóng góp của luận án thể hiện ở 2 điểm:

- Luận án đã phát triển một khung phương pháp luận nghiên cứu tích hợp dữ liệu viễn thám, dữ liệu quan trắc khí tượng thủy văn và các mô hình học máy trong việc đánh giá và dự báo hạn hán lưu vực sông Srêpốk, làm rõ được đặc điểm hạn hán tại lưu vực và xác định được cặp chỉ số chỉ hạn khí tượng - chỉ số thực vật từ viễn thám phù hợp nhất cho giám sát hạn.
- Luận án đã xây dựng các mô hình học máy nhằm dự báo sớm chỉ số sức khỏe thực vật (VHI) cho lưu vực sông Srêpốk dựa trên các chỉ số hạn khí tượng (SPI), hạn thủy văn (SSI) và dữ liệu VHI của các thời điểm trước đó.

Tính mới của luận án không chỉ thể hiện ở việc áp dụng công nghệ và phương pháp tiên tiến, mà còn ở cách tiếp cận tích hợp. Những đóng góp này vừa có ý nghĩa khoa học vừa mang tính ứng dụng thực tiễn cao, đáp ứng nhu cầu cấp thiết về giám sát và dự báo hạn hán trong bối cảnh biến đổi khí hậu.

CHƯƠNG 1. TỔNG QUAN VỀ CÁC VẤN ĐỀ NGHIÊN CỨU

1.1 Tổng quan về hạn hán

Hạn hán được phân loại thành bốn dạng chính: hạn khí tượng (thiếu hụt lượng mưa so với trung bình nhiều năm), hạn thủy văn (suy giảm dòng chảy và mực nước ngầm), hạn nông nghiệp (thiếu hụt độ ẩm đất ảnh hưởng đến sinh trưởng cây trồng) và hạn kinh tế - xã hội (tác động đến cung cầu tài nguyên nước). Các dạng hạn có mối quan hệ nhân quả theo chuỗi thời gian: hạn khí tượng khởi phát trước, tiếp đến là hạn nông nghiệp và cuối cùng là hạn thủy văn. Trong đó, hạn nông nghiệp phản ánh trực tiếp mức độ ảnh hưởng đến sức khỏe thực vật và năng suất cây trồng là đối tượng trung tâm của luận án.

Nguyên nhân gây ra hạn hán bao gồm cả yếu tố tự nhiên (biến động hoàn lưu khí quyển, hiện tượng ENSO, suy giảm thảm phủ rừng) và nhân tạo (khai thác quá mức tài nguyên nước, mở rộng diện tích canh tác). Tác động của hạn hán rất đa chiều: giảm năng suất cây trồng, suy giảm nguồn nước, gia tăng nguy cơ cháy rừng, suy thoái đất và ảnh hưởng nghiêm trọng đến sinh kế cộng đồng nông thôn.

Trên thế giới và tại Việt Nam, hạn hán ngày càng gia tăng về tần suất và cường độ dưới tác động của biến đổi khí hậu. Tại Việt Nam, các đợt hạn nặng điển hình giai đoạn 1997–1998, 2004–2005, 2014–2016 và 2019–2020 chủ yếu xảy ra trong mùa khô, trùng với thời kỳ canh tác vụ Đông Xuân và đầu vụ Hè Thu, gây thiệt hại nặng nề cho sản xuất nông nghiệp, đặc biệt tại khu vực Tây Nguyên và Nam Trung Bộ.

1.2. Các chỉ số thực vật từ viễn thám trong giám sát hạn hán

Công nghệ viễn thám, đặc biệt là dữ liệu MODIS với chu kỳ quan sát 8–16 ngày và độ phân giải không gian 250–1.000 m, đã mở ra khả năng giám sát hạn hán liên tục trên quy mô lớn mà không phụ thuộc vào mạng lưới trạm quan trắc mặt đất. Từ dữ liệu MODIS, ba chỉ số thực vật được sử dụng phổ biến trong giám sát hạn nông nghiệp gồm VCI, TCI và VHI. VHI là chỉ số tổng hợp kết hợp VCI và TCI theo trọng số bằng nhau (α = 0,5), phản ánh đồng thời cả khả năng phản xạ và phát xạ của thảm thực

vật. Nhiều nghiên cứu quốc tế và trong nước đã khẳng định VHI vượt trội hơn VCI và TCI trong giám sát hạn nông nghiệp nhờ tính toàn diện này. Tuy nhiên, hiệu quả của các chỉ số viễn thám phụ thuộc vào loại thảm phủ và điều kiện khí hậu địa phương, đòi hỏi phải có đánh giá và lựa chọn hệ thống trước khi áp dụng cho từng khu vực cụ thể.

1.3. Các chỉ số hạn khí tượng và thủy văn

Bên cạnh các chỉ số viễn thám, các chỉ số hạn khí tượng và thủy văn đóng vai trò quan trọng trong việc phản ánh nguyên nhân và diễn biến hạn hán. SPI (Standardized Precipitation Index) là chỉ số chuẩn hóa lượng mưa theo phân phối xác suất, có thể tính toán ở nhiều quy mô thời gian từ 1 đến 24 tháng, dễ áp dụng và cho phép so sánh giữa các vùng khí hậu khác nhau. EDI (Effective Drought Index) phản ánh lượng mưa hiệu quả tích lũy có tính đến sự suy giảm theo thời gian, nhạy hơn SPI trong phát hiện hạn ngắn hạn. SSI (Standardized Streamflow Index) là chỉ số hạn thủy văn tính từ dữ liệu dòng chảy, phản ánh tình trạng suy giảm nguồn nước mặt và bổ sung thông tin thủy văn quan trọng cho dự báo hạn hán nông nghiệp. Mỗi chỉ số có ưu điểm và hạn chế riêng; việc lựa chọn quy mô thời gian phù hợp cho từng loại thảm phủ và mục đích sử dụng là yếu tố then chốt quyết định hiệu quả giám sát.

1.4. Các phương pháp học máy trong dự báo hạn hán

Các mô hình học máy đã chứng minh hiệu quả vượt trội so với các phương pháp thống kê truyền thống trong dự báo hạn hán nhờ khả năng mô phỏng các mối quan hệ phi tuyến phức tạp giữa các biến đầu vào. Ba nhóm phương pháp được lựa chọn trong luận án gồm SMLR, RF và ANN. SMLR (Stepwise Multiple Linear Regression) là hồi quy tuyến tính bội với quy trình lựa chọn biến tự động, ưu điểm là đơn giản, dễ diễn giải và phù hợp khi mối quan hệ giữa các biến mang tính tuyến tính. RF (Random Forest) là mô hình tổ hợp nhiều cây quyết định, có khả năng xử lý dữ liệu phi tuyến, kháng nhiễu tốt và cung cấp thông tin về mức độ quan trọng của các biến đầu vào. ANN (Artificial Neural Network) là mạng nơ-ron nhân tạo nhiều lớp, có khả năng học các mối quan hệ phi tuyến phức tạp,

đặc biệt hiệu quả với dữ liệu chuỗi thời gian thủy văn. Tổng quan quốc tế cho thấy hiệu quả của từng mô hình phụ thuộc mạnh vào thời hạn dự báo và điều kiện khí hậu khu vực; ANN và RF thường vượt trội ở thời hạn ngắn, trong khi SMLR có ưu thế ở thời hạn dài hơn khi mối quan hệ tuyến tính chiếm ưu thế.

1.5. Tổng quan khu vực nghiên cứu

Lưu vực sông Srêpốk có diện tích khoảng 18.480 km², trải rộng trên địa bàn các tỉnh Đắk Lắk, Đắk Nông và Lâm Đồng. Địa hình phân hóa thành ba vùng chính: vùng núi cao phía Nam và Đông Nam (1.000–2.400 m), vùng cao nguyên bazan Buôn Ma Thuột và Đắk Nông (400–800 m) và vùng đất thấp ven sông (200–450 m). Chế độ khí hậu nhiệt đới gió mùa với mùa mưa tập trung từ tháng 5 đến tháng 10 (chiếm 85% lượng mưa năm) và mùa khô kéo dài từ tháng 11 đến tháng 4 tạo nên sự phân hóa nguồn nước gay gắt. Lưu vực có hệ thống sông ngòi phát triển với hai nhánh chính Krông Ana và Krông Knô, cùng hệ thống 30 công trình thủy điện với tổng công suất trên 900 MW.

Về sử dụng đất, cây lâu năm (chủ yếu là cà phê Robusta, hồ tiêu, điều) chiếm 28,9% diện tích, tiếp theo là rừng đặc dụng (20,1%), rừng sản xuất (16,8%), cây hàng năm (7,7%) và lúa (4,3%). Hạn hán xảy ra thường xuyên, đặc biệt nghiêm trọng trong các giai đoạn 2004–2005, 2014–2016 và 2019–2020, gây ảnh hưởng nặng nề đến sản xuất nông nghiệp và đời sống người dân.

1.6. Các tồn tại và định hướng nghiên cứu

Tổng quan các công trình nghiên cứu tại lưu vực Srêpốk cho thấy các nghiên cứu trước đây chủ yếu sử dụng mô hình thủy văn SWAT để đánh giá hạn thủy văn, chưa tích hợp dữ liệu viễn thám để giám sát trực tiếp trạng thái thảm thực vật. Ngoài ra, mối quan hệ định lượng giữa chỉ số hạn và năng suất các cây trồng chính chưa được làm rõ; bộ chỉ số hạn phù hợp với điều kiện sinh thái–khí hậu đặc thù của lưu vực chưa được xác định một cách hệ thống; và chưa có nghiên cứu nào ứng dụng học máy

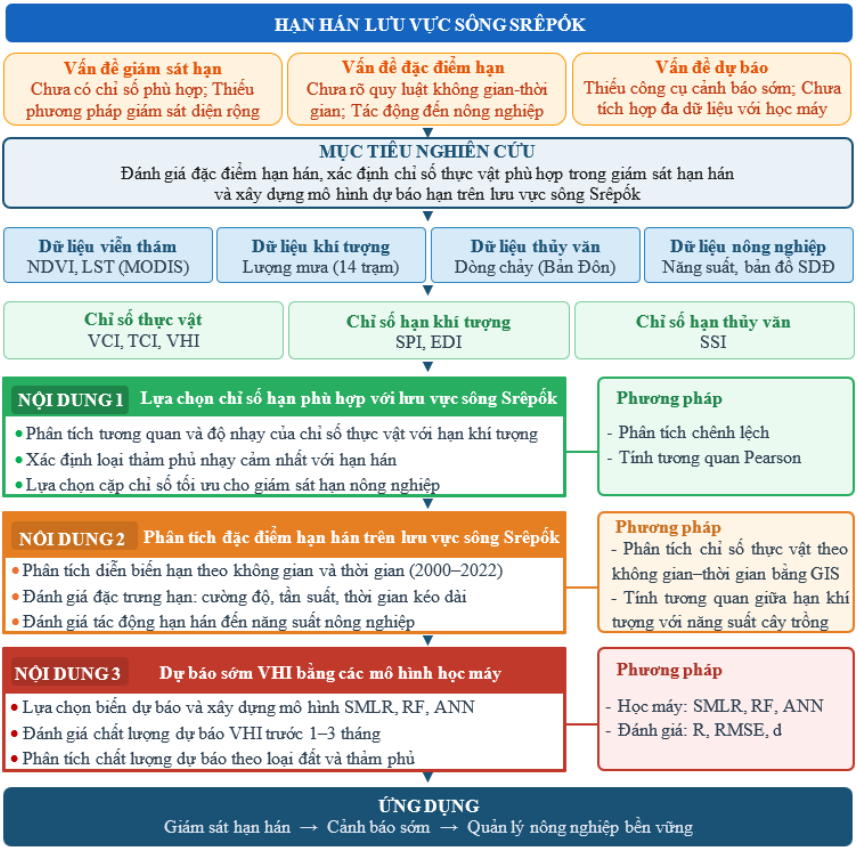
kết hợp dữ liệu viễn thám và khí tượng–thủy văn để dự báo sớm hạn nông nghiệp cho khu vực này.

Từ những khoảng trống đó, luận án xác định ba hướng nghiên cứu nối tiếp: (i) lựa chọn bộ chỉ số hạn phù hợp; (ii) phân tích đặc điểm hạn hán theo không gian–thời gian và đánh giá tác động đến năng suất cây trồng; (iii) xây dựng mô hình học máy dự báo sớm chỉ số thực vật trước 1–3 tháng trên cơ sở tích hợp dữ liệu viễn thám với dữ liệu quan trắc khí tượng–thủy văn.

CHƯƠNG 2. NỘI DUNG VÀ PHƯƠNG PHÁP NGHIÊN CỨU

2.1. Khung nghiên cứu và nội dung nghiên cứu

Luận án được thiết kế theo ba nội dung nghiên cứu nối tiếp nhau, trong đó kết quả của mỗi nội dung trước tạo tiền đề cho nội dung tiếp theo (Hình 2.1).



Hình 2.1 Khung nghiên cứu tổng thể của luận án

Nội dung 1 tập trung tính toán, đánh giá và lựa chọn bộ chỉ số hạn hán phù hợp với đặc điểm khí hậu–sinh thái của lưu vực. Nội dung 2 kế thừa bộ chỉ số đã xác định để phân tích toàn diện đặc điểm hạn hán theo không gian và thời gian, đồng thời đánh giá tác động đến năng suất cây trồng

chính. Nội dung 3 phát triển mô hình học máy dự báo sớm chỉ số sức khỏe thực vật VHI trước 1–3 tháng, tích hợp dữ liệu viễn thám với dữ liệu quan trắc khí tượng–thủy văn phục vụ cảnh báo sớm hạn hán nông nghiệp.

2.2. Dữ liệu nghiên cứu

Nghiên cứu khai thác các nguồn dữ liệu chính được tổng hợp trong Bảng 2.1.

Bảng 2.1 Tổng hợp dữ liệu sử dụng trong nghiên cứu

Dữ liệu	Mô tả dữ liệu	Giai đoạn	Nguồn
NDVI (MOD13A1)	463m x 463m, 16 ngày	2000–2022	Cơ quan Khảo sát Địa chất Hoa Kỳ
LST (MOD11A2)	926 m x 926m, 8 ngày	2000–2022	Cơ quan Khảo sát Địa chất Hoa Kỳ
Lượng mưa	14 trạm, ngày	2000–2022	Tổng cục Khí tượng Thủy văn Việt Nam
Lưu lượng dòng chảy	1 trạm, ngày	2000–2022	Tổng cục Khí tượng Thủy văn Việt Nam
Bản đồ hiện trạng sử dụng đất	Vector	2015	Phân viện QH và TK Nông nghiệp
Bản đồ đất	Vector	2015	Phân viện QH và TK Nông nghiệp
Năng suất cây trồng	Cấp huyện; năm	2000–2022	Niên giám Thống kê 3 Tỉnh

Dữ liệu viễn thám MODIS gồm chỉ số thực vật NDVI (sản phẩm MOD13A1, độ phân giải 463 m × 463 m, chu kỳ 16 ngày) và nhiệt độ bề mặt đất LST (sản phẩm MOD11A2, độ phân giải 926 m × 926 m, chu kỳ 8 ngày), giai đoạn 2000–2022. Dữ liệu khí tượng gồm lượng mưa ngày tại 14 trạm quan trắc (5 trạm trong lưu vực và 9 trạm lân cận) do Tổng cục Khí tượng Thủy văn Việt Nam quản lý, với tỷ lệ dữ liệu thiếu khoảng 6,4% được bổ sung bằng hồi quy tuyến tính từng bước. Dữ liệu thủy văn

gồm lưu lượng dòng chảy ngày tại trạm Bản Đôn, phụ trách 92% diện tích lưu vực. Ngoài ra, nghiên cứu còn sử dụng bản đồ hiện trạng sử dụng đất và bản đồ đất năm 2015 từ Phân viện Quy hoạch và Thiết kế Nông nghiệp, cùng dữ liệu năng suất cà phê và lúa cấp huyện giai đoạn 2000–2022 từ niên giám thống kê ba tỉnh Đắk Lắk, Đắk Nông (cũ) và Lâm Đồng. Dữ liệu MODIS được tiền xử lý bằng phương pháp Maximum Value Composite (MVC) để tổng hợp ảnh tháng từ ảnh 16 ngày, đồng thời LST được nội suy lại theo phương pháp Bilinear về độ phân giải 463 m × 463 m để đồng nhất với NDVI.

2.3. Phương pháp tính toán các chỉ số hạn hán

Từ các nguồn dữ liệu trên, ba nhóm chỉ số hạn hán được tính toán. Nhóm chỉ số hạn khí tượng gồm SPI (Standardized Precipitation Index) và EDI (Effective Drought Index) được tính cho các quy mô thời gian từ 1 đến 12 tháng tại từng trạm, sau đó nội suy không gian theo phương pháp IDW về lưới raster 463 m × 463 m. Nhóm chỉ số hạn thủy văn gồm SSI (Standardized Streamflow Index) tính từ dữ liệu lưu lượng tại trạm Bản Đôn với quy trình tương tự SPI. Nhóm chỉ số thực vật viễn thám gồm VCI từ NDVI chuẩn hóa, TCI từ LST chuẩn hóa, và VHI là chỉ số tổng hợp kết hợp VCI và TCI. Cấp độ hạn được phân loại theo ngưỡng VHI: hạn nhẹ (<40), hạn vừa (<30), hạn nặng (<20) và hạn nghiêm trọng (<10).

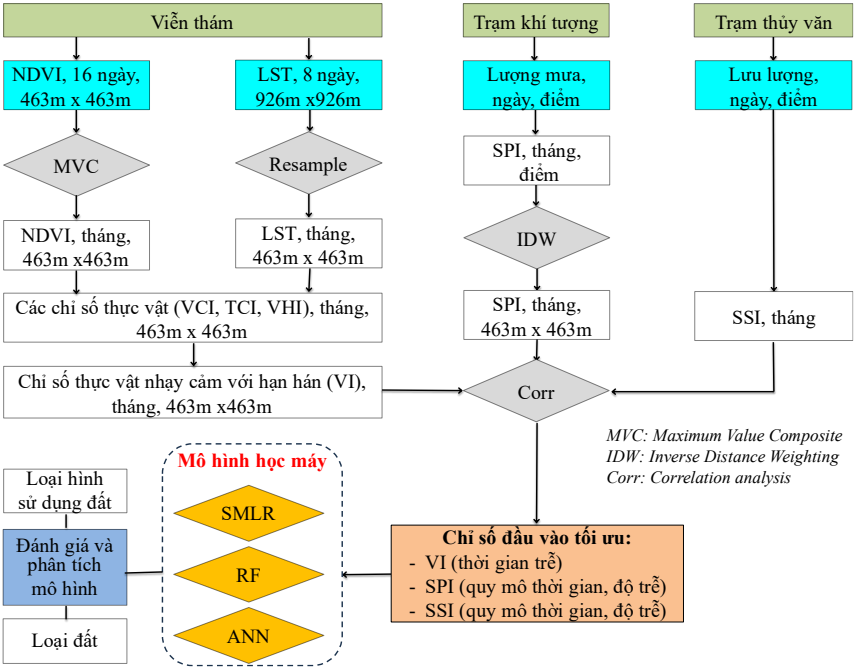
2.4. Phương pháp lựa chọn chỉ số hạn phù hợp

Để xác định bộ chỉ số hạn tối ưu cho lưu vực, nghiên cứu thực hiện theo các bước: (1) phân tích mức chênh lệch của các chỉ số thực vật giữa điều kiện khô hạn (phân vị ≤25%) và ẩm ướt (phân vị ≥75%) theo từng loại thảm phủ; (2) phân tích hệ số tương quan Pearson giữa các chỉ số thực vật (VCI, TCI, VHI) và chỉ số hạn khí tượng (SPI, EDI) ở các quy mô thời gian và độ trễ khác nhau, thực hiện riêng biệt cho từng loại hình sử dụng đất; (3) xác định cặp chỉ số tối ưu và loại thảm phủ có độ nhạy cảm cao nhất với hạn hán; và (4) đánh giá mối quan hệ giữa chỉ số hạn và năng suất cà phê, lúa thông qua hệ số tương quan bội có loại bỏ ảnh hưởng xu

thế bằng biến thời gian, với dữ liệu năng suất chuẩn hóa từ 19 huyện được ghép nối thành chuỗi chung.

2.5. Phương pháp dự báo sớm chỉ số thực vật nhạy cảm với hạn hán bằng các mô hình học máy

Quá trình xây dựng mô hình dự báo được thực hiện theo quy trình tại Hình 2.4.



Hình 2.1 Sơ đồ nghiên cứu dự báo chỉ số thực vật nhạy cảm với hạn hán

- **Các bước thực hiện chính bao gồm:** (1) Xây dựng cơ sở dữ liệu đầu vào, bao gồm NDVI và LST từ MODIS, lượng mưa tại các trạm khí tượng, lưu lượng dòng chảy tại trạm Bản Đôn, cùng bản đồ đất và bản đồ sử dụng đất; (2) Tiền xử lý và tính toán các chỉ số như VCI, TCI, VHI, SPI và SSI, bảo đảm đồng nhất về độ phân giải không gian và chu kỳ thời gian; (3) Phân tích mối quan hệ giữa chỉ số thực vật và các chỉ số hạn để

xác định biến đầu vào tối ưu cho mô hình; (4) Xây dựng và hiệu chỉnh các mô hình dự báo bao gồm SMLR, RF và ANN; và (5) Đánh giá ảnh hưởng của loại đất và loại hình sử dụng đất đến kết quả dự báo chỉ số thực vật nhạy cảm với hạn hán.

- **Ba mô hình học máy được xây dựng và so sánh.**

+ Mô hình SMLR sử dụng hồi quy tuyến tính bội với quy trình lựa chọn biến tự động theo mức ý nghĩa thống kê.
+ Mô hình RF được triển khai bằng Python/scikit-learn (RandomForestRegressor), các siêu tham số (n_estimators, max_depth, min_samples_split, min_samples_leaf) được tối ưu hóa bằng GridSearchCV kết hợp 5-fold cross-validation.
+ Mô hình ANN được xây dựng theo kiến trúc mạng truyền thẳng nhiều lớp (MLP) bằng TensorFlow/Keras, với cơ chế Early Stopping để tránh quá khớp.

Các mô hình được xây dựng riêng cho từng tháng nhằm loại bỏ tác động của mùa vụ.

Dữ liệu được phân chia theo thời gian thành tập huấn luyện (70%, giai đoạn 2000–2014) và tập kiểm định độc lập (30%, giai đoạn 2015–2022).

- **Đánh giá kết quả dự báo:**

Độ chính xác của kết quả dự báo chỉ số thực vật nhạy cảm với hạn hán được đánh giá qua ba tham số thống kê là hệ số tương quan (R), sai số bình phương trung bình (RMSE) và chỉ số Willmott (d). Các tham số này được xác định như sau:

$$R = \frac{\sum_{i=1}^{n}(Y_i-\bar{Y})(Ys_i-\overline{Ys})}{\sqrt{\sum_{i=1}^{n}(Y_i-\bar{Y})^2}\sqrt{\sum_{i=1}^{n}(Ys_i-\overline{Ys})^2}} \quad (1)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Ys_i - Y_i)^2} \quad (2)$$

$$d = 1 - \frac{\sum_{i=1}^{n}(Y_i-Ys_i)^2}{\sum_{i=1}^{n}\left(|Ys_i-\overline{Y}|+|Y_i-\overline{Y}|\right)^2} \quad (3)$$

Trong các phương trình này thì n là độ dài của chuỗi số liệu; Y là giá trị thực tế của chỉ số thực vật và $\overline{Y}$ là giá trị trung bình của Y; Ys là kết quả dự báo.

Để tăng chất lượng dự báo, các biến trước khi đưa vào tuyển chọn bằng kỹ thuật LOOCV cần được tuyển chọn sơ bộ. Trong nghiên cứu này việc tuyển chọn sơ bộ dựa việc phân tích hệ số tương quan, chỉ có các biến có quan hệ tốt với biến được dự báo mới được giữ lại.

- **Đánh giá sự đóng góp của các biến:**

Mức độ đóng góp của các biến vào mô hình được đánh giá một cách tương đối. Mức đóng góp được xác định thông qua mức giảm của hệ số xác định hiệu chỉnh (R^2) khi biến phân tích được rút khỏi mô hình. Mức đóng góp được tính cho một biến bất kỳ được rút ra như sau:

$$C_j = \frac{R^2 - R_j^2}{\sum_{i=1}^{m}(R^2 - R_i^2)} 100\% \quad (4)$$

Trong đó: Cj chỉ mức đóng góp của biến thứ j, j = 1, 2 … ,m, m là số biến được chọn; R^2 là hệ số xác định hiệu chỉnh khi sử dụng tất cả các biến được chọn; R_i^2 và R_j^2 tương ứng là hệ số xác định hiệu chỉnh khi biến thứ i và biến j được rút ra.

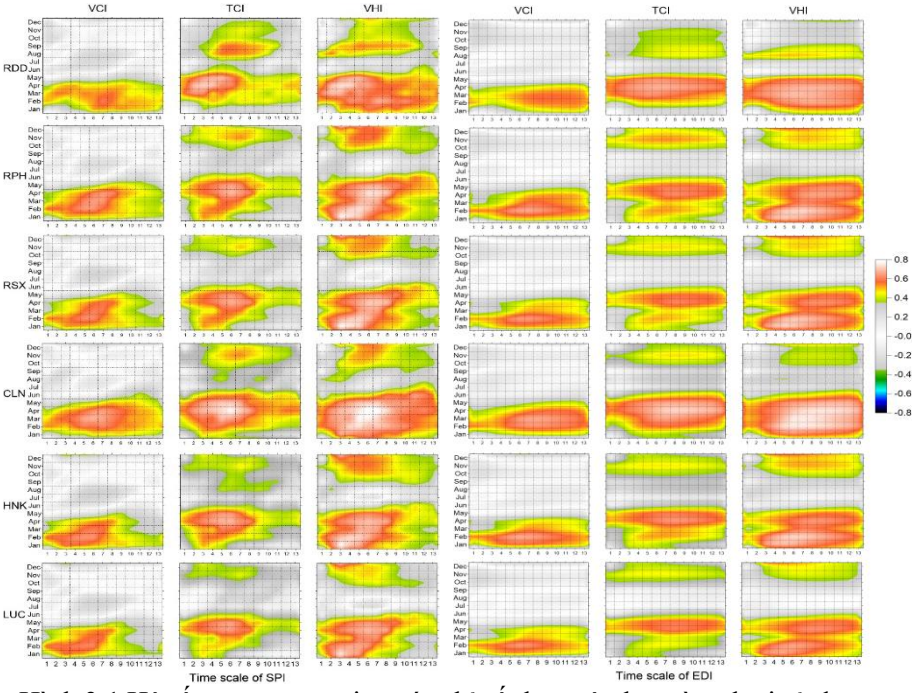
- Chất lượng dự báo cũng được phân tích theo từng loại hình sử dụng đất và loại đất nhằm đánh giá tính ứng dụng thực tiễn của mô hình.

CHƯƠNG 3. KẾT QUẢ NGHIÊN CỨU VÀ THẢO LUẬN

3.1. Lựa chọn chỉ số hạn hán phù hợp cho lưu vực sông Srêpốk

3.1.1. Chênh lệch chỉ số thực vật giữa điều kiện khô hạn và ẩm ướt

Phân tích mức chênh lệch các chỉ số thực vật giữa điều kiện khô hạn (phân vị ≤25%) và ẩm ướt (phân vị ≥75%) theo từng loại thảm phủ cho thấy mức chênh lệch theo SPI thường cao hơn EDI, và chênh lệch theo VHI là rõ rệt nhất trong ba chỉ số thực vật. Cây lâu năm (CLN) là nhóm nhạy cảm nhất với điều kiện khô hạn, trong khi rừng tự nhiên ít chịu tác động nhất.



Hình 3.1 Hệ số tương quan giữa các chỉ số thực vật theo từng loại sử dụng và chỉ số hạn ứng với các quy mô thời gian từ 1 tháng đến 13 tháng

Kết quả này nhất quán với phân tích tương quan Pearson (Hình 3.2): các chỉ số thực vật trên vùng CLN có hệ số tương quan cao nhất với chỉ số

hạn khí tượng, tập trung trong khoảng thời gian từ tháng 1 đến tháng 5, đặc biệt rõ nhất từ tháng 2 đến tháng 4, đây giai đoạn cuối mùa khô khi lượng mưa gần như không đáng kể và nhiệt độ đạt đỉnh. VHI cho hệ số tương quan với SPI cao hơn VCI và TCI ở tất cả các quy mô thời gian do phản ánh đồng thời cả tác động của độ ẩm đất và nhiệt độ lên thực vật.

i. Xác định cặp chỉ số tối ưu và loại thảm phủ chỉ thị

Phân tích hệ số tương quan giữa các chỉ số thực vật với SPI và EDI ở các quy mô thời gian từ 1 đến 13 tháng cho giai đoạn tháng 2–5 cho thấy SPI cho hệ số tương quan cao hơn và đỉnh rõ rệt hơn EDI. Trong các quy mô thời gian của SPI, SPI6 đạt hệ số tương quan cao nhất với VHI, phản ánh tác động tích lũy của lượng mưa từ tháng 9 năm trước đến tháng 2 năm sau đến tình trạng thảm thực vật cuối mùa khô.

Bảng 3.1 Thống kê giá trị trung bình và độ lệch chuẩn của hệ số tương quan SPI6 và VHI

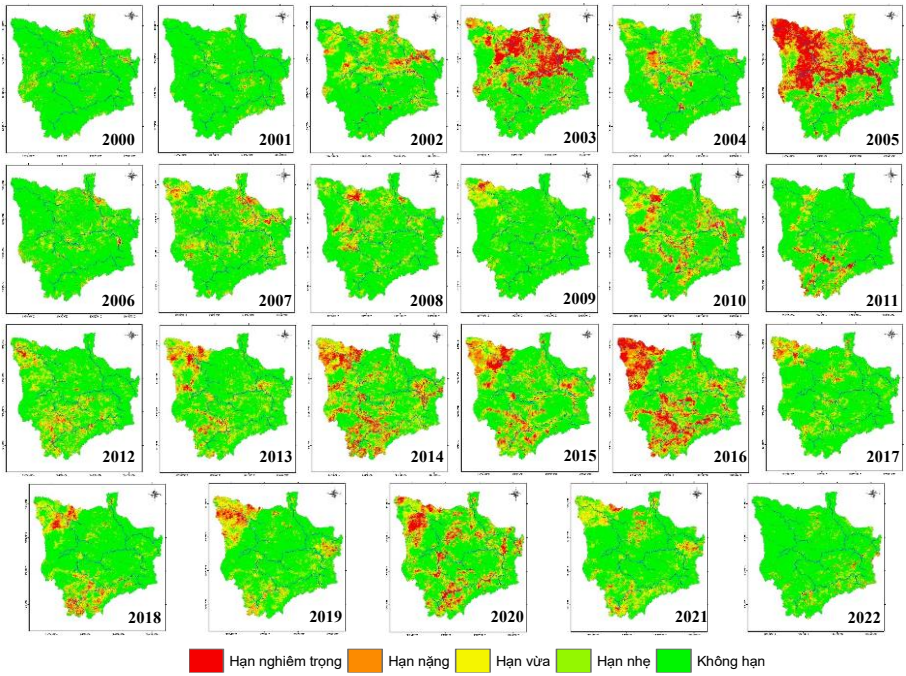
Loại hình sử dụng đất/Thảm phủ	**Kí hiệu**	**Trung bình**	**Độ lệch chuẩn**
Rừng đặc dụng	RDD	0,31	0,06
Rừng phòng hộ	RPH	0,36	0,1
Rừng sản xuất	RSX	0,39	0,12
Lúa	LUC	0,44	0,11
Cây hàng năm	HNK	0,46	0,13
Cây lâu năm	CLN	0,49	0,11

Phân tích phân bố không gian hệ số tương quan giữa SPI6 và VHI theo từng ô lưới được tổng hợp theo Bảng 3.2 cho thấy CLN là nhóm nhạy cảm nhất (R trung bình = 0,49), tiếp theo là HNK (R = 0,46), LUC (R = 0,44), RSX (R = 0,39), RPH (R = 0,36) và RDD (R = 0,31). Kết hợp với diện tích lớn nhất (28,9% lưu vực) và giá trị kinh tế cao nhất, CLN được xác định là đối tượng chỉ thị tối ưu nhất cho giám sát hạn hán. Như vậy, cặp chỉ số SPI6 và VHI trên vùng cây lâu năm trong giai đoạn tháng 2–5 là phù hợp nhất cho giám sát hạn nông nghiệp trên lưu vực.

3.2. Đặc điểm hạn hán trên lưu vực sông Srêpốk giai đoạn 2000–2022

3.2.1. Diễn biến của hạn hán theo không gian và thời gian

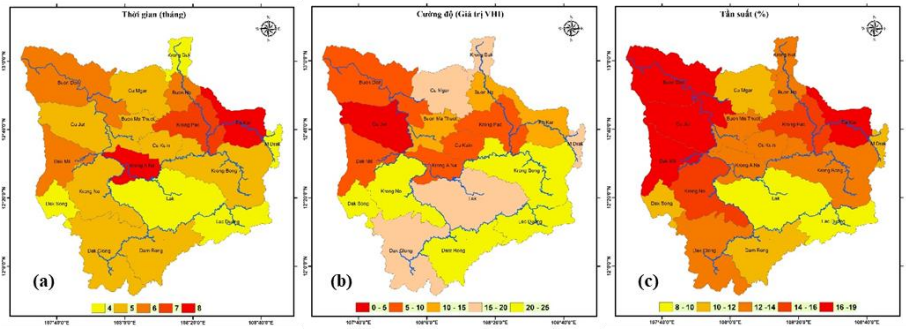
Phân tích chuỗi VHI tháng 4 trong 23 năm (Hình 3.7) cho thấy hạn hán trên lưu vực xuất hiện với chu kỳ khoảng 5 năm, gắn liền với các đợt El Niño mạnh. Hai đợt hạn nghiêm trọng nhất được ghi nhận vào năm 2005 và 2016, với diện tích bị ảnh hưởng từ hạn vừa đến hạn nghiêm trọng lần lượt đạt 42% và 41% tổng diện tích lưu vực. Các năm 2003, 2010 và 2020 cũng có tỷ lệ diện tích bị hạn nặng đáng kể (28%, 11% và 18%).



Hình 3.2 Kết quả tính toán chỉ số VHI theo cấp hạn giai đoạn 2000 – 2022

3.2.2. Phân bố các đặc trưng hạn theo không gian

Phân tích ba đặc trưng hạn theo đơn vị hành chính cấp huyện dựa trên ngưỡng VHI < 40 và thời gian kéo dài tối thiểu 3 tháng (Hình 3.10) cho thấy sự phân hóa rõ rệt theo hướng Bắc–Nam và Tây–Đông. Về thời gian kéo dài, khu vực phía Bắc và Tây Bắc (Buôn Đôn, Cư Jút, Đắk Mil, Ea Kar, Krông Pac) chịu hạn 6–8 tháng liên tiếp, trong khi phía Nam (Lạc Dương, Đam Rông, Krông Bông) chỉ 4–5 tháng. Về cường độ, các huyện Tây Bắc và trung tâm (Buôn Đôn, Cư Jút, Đắk Mil, Krông Ana, Cư Kuin, Krông Pắc) có VHI thấp nhất (0–10, hạn nghiêm trọng), trong khi phía Đông Nam có mức ảnh hưởng trung bình (VHI = 20–25). Về tần suất, hạn hán xảy ra với tỷ lệ cao nhất ở khu vực Tây Bắc (14–19% thời gian nghiên cứu), trung bình ở trung tâm lưu vực (10–14%) và thấp nhất ở Lạc Dương và Lắk (8–10%).



Hình 3.3 Phân bố theo không gian của đặc trưng hạn: a) thời gian kéo dài, b) cường độ và c) tần suất

3.2.3. Tác động đến sản xuất nông nghiệp

Phân tích diện tích bị ảnh hưởng trong hai đợt hạn nghiêm trọng 2005 và 2016 cho thấy sự phân hóa rõ rệt theo nhóm cây trồng. Đất trồng cây hàng năm (HNK) chịu tổn thương cao nhất với tổng tỷ lệ diện tích bị hạn nghiêm trọng và hạn nặng lên đến 42,57% (khoảng 38.777 ha). Đất trồng lúa (LUC) cũng bị ảnh hưởng nặng với 43,04% (21.866 ha), chủ yếu do suy giảm nguồn nước hồ chứa trong các năm hạn khốc liệt. Đất trồng cây

lâu năm (CLN) có tỷ lệ bị hạn nghiêm trọng và hạn nặng thấp hơn (31,48%), nhưng do diện tích canh tác lớn (340.000 ha - gấp 3,7 lần HNK và 6,7 lần LUC), tổng diện tích bị ảnh hưởng lớn nhất về giá trị tuyệt đối: hơn 60.409 ha bị hạn nghiêm trọng, trong đó cà phê và hồ tiêu ghi nhận rụng quả, giảm năng suất mạnh khi hạn trùng với thời kỳ ra hoa, nuôi quả.

3.1.3. Quan hệ giữa chỉ số hạn và năng suất cây trồng

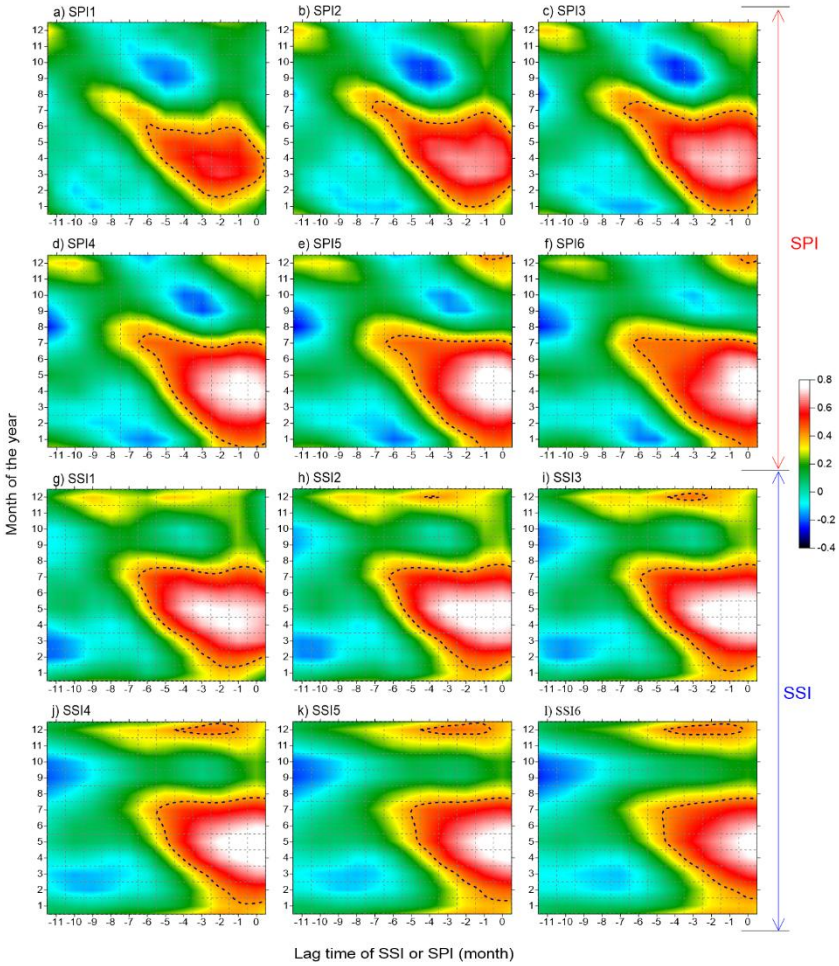
Hệ số tương quan bội (loại bỏ ảnh hưởng xu thế) giữa năng suất cà phê và chỉ số hạn khí tượng cho thấy SPI5 lấy trước thời gian thu hoạch 9 tháng ($SPI5_9$) cho hệ số tương quan cao nhất (R = 0,73), xác định lượng mưa tích lũy từ tháng 9 năm trước đến tháng 1 năm thu hoạch là yếu tố quyết định năng suất cà phê, trong đó lượng mưa tháng 11 đóng vai trò quan trọng nhất. Đối với lúa Đông Xuân, SPI5 lấy trước 3 tháng ($SPI5_3$) cho hệ số tương quan cao nhất (R = 0,74), tương ứng với lượng mưa từ tháng 8 đến tháng 12 của năm trước. Kết quả này mở ra khả năng dự báo sớm tác động của hạn hán đến năng suất trước 3–9 tháng, phục vụ lập kế hoạch sản xuất.

3.3. Dự báo sớm chỉ số VHI bằng các mô hình học máy

3.3.1. Lựa chọn biến dự báo và xây dựng tập dữ liệu

Phân tích hệ số tương quan giữa VHI với SPI (quy mô 1–6 tháng) và SSI (quy mô 1–6 tháng) ở các độ trễ (lag) từ 0–11 tháng (Hình 3.14) cho thấy VHI có hệ số tương quan cao nhất với SPI4 và SPI5 và SSI1–SSI3. Phân tích tự tương quan VHI cho thấy VHI tháng trước (lag-1) là biến dự báo mạnh với tất cả các tháng, đặc biệt lag-2 và lag-3 hiệu quả cho giai đoạn tháng 4–7. Từ đó xác định 31 biến đầu vào gồm: VHI trễ, SPI4, SPI5, SSI1, SSI2, SSI3 lấy từ 1–5 tháng trước tháng dự báo và biến xu thế thời gian.

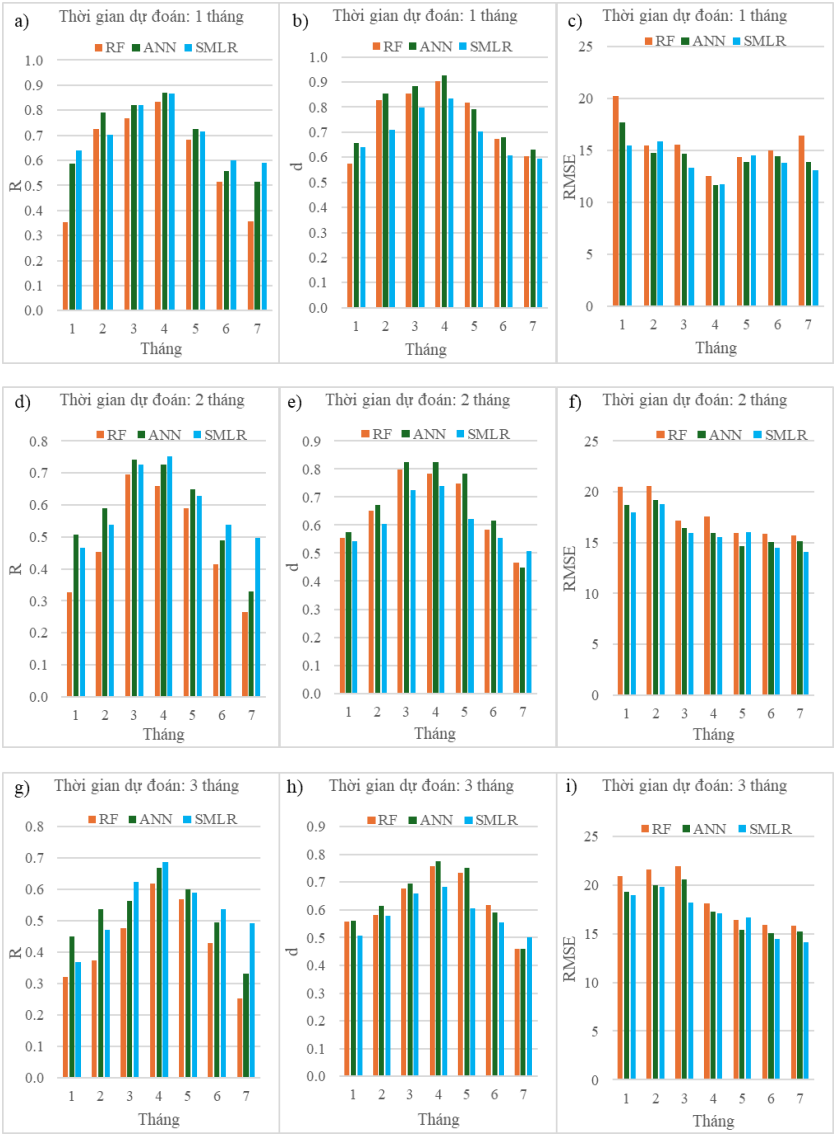
Tập dữ liệu dự báo được xây dựng từ tập hợp 55.610 ô lưới trong 23 năm tạo ra chuỗi có độ dài 1.279.030. Tập này được phân chia thành tập huấn luyện (2000–2014, 70%) và tập kiểm định độc lập (2015–2022, 30%). Dữ liệu VHI được chuẩn hóa về giá trị trung bình bằng 0 và độ lệch chuẩn bằng 1 của tập huấn luyện để tránh rò rỉ thông tin.



Hình 3.4 Hệ số tương quan giữa VHI của các tháng trong năm với SPI và SSI có quy mô thời gian từ 1 đến 6 tháng. Trong đó SPI và SSI được lấy theo thời gian trễ trước VHI từ 0 đến 11 tháng.

3.3.2. So sánh hiệu quả dự báo của ba mô hình

Kết quả đánh giá trên tập kiểm định độc lập (Hình 3.16) cho thấy sự phân hóa rõ rệt về hiệu suất giữa ba mô hình theo thời hạn dự báo và mùa vụ.



Hình 3.5 Kết quả đánh giá chất lượng dự báo qua a) R, b) d, và c) RMSE cho thời gian dự báo 1 tháng; d) R, e) d, và f) RMSE cho thời gian dự báo 2 tháng; g) R, h) d, và i) RMSE cho thời gian dự báo 3 tháng

Với dự báo trước 1 tháng, ANN đạt chất lượng tốt nhất: tháng 4 đạt R = 0,87, d = 0,93, RMSE = 11,5; tháng 2–3 đạt R = 0,79–0,81, d = 0,86–0,88. ANN dẫn đầu về R ở 57,1% số tháng và về d ở 100% số tháng; SMLR chiếm ưu thế về RMSE (57,1%). RF không đứng đầu ở bất kỳ chỉ tiêu nào.

Với dự báo trước 2 tháng, ANN tiếp tục vượt trội về d (85,7% số tháng), trong khi SMLR có ưu thế về RMSE (85,7%), SMLR phù hợp hơn khi ưu tiên giảm sai số tuyệt đối.

Với dự báo trước 3 tháng, SMLR trở thành lựa chọn phù hợp hơn, dẫn đầu về R (57%) và RMSE (85,7%); ANN vẫn ưu thế về d (71,4%). RF có hiệu suất thấp nhất, đặc biệt tháng 7 với dự báo 3 tháng chỉ đạt R = 0,25, do dữ liệu chưa đủ lớn để khai thác quan hệ phi tuyến. Nhìn chung, mức độ suy giảm hiệu suất khi kéo dài thời hạn dự báo của ANN (15–20%/tháng) thấp hơn SMLR (18–25%) và RF (25–35%), khẳng định ANN có tính ổn định cao nhất.

Như vậy các kết quả cho thấy ANN là mô hình tối ưu cho dự báo trước 1–2 tháng, đặc biệt trong mùa khô khi biến động VHI lớn và tác động của hạn hán rõ rệt. SMLR phù hợp hơn cho dự báo dài hạn 3 tháng hoặc trong mùa mưa, khi mối quan hệ tuyến tính giữa biến dự báo và VHI chiếm ưu thế. Ngược lại, RF có chất lượng dự báo thấp nhất và độ ổn định kém, do đó không được khuyến nghị áp dụng cho dự báo VHI tại lưu vực Srêpốk.

3.3.3. Vai trò đóng góp của các nhóm biến

Phân tích mức đóng góp tương đối của các biến vào mô hình ANN cho thấy: với dự báo 1 tháng, VHI trễ đóng góp lớn nhất (54–88%), đỉnh điểm tháng 3 và 4 (74% và 88%) khi thảm thực vật chịu áp lực hạn hán cao nhất; SPI đứng thứ hai trong mùa mưa (44–46%); SSI đóng góp ổn định 17–20% trong mùa khô đầu năm. Khi tăng lên 2–3 tháng, vai trò VHI giảm dần, SSI nổi bật hơn trong giai đoạn tháng 1–3 (40–72%) và SPI chiếm ưu thế ở tháng 6–7 (59–75%), phản ánh sự chuyển dịch từ tín hiệu thực vật sang tín hiệu khí tượng–thủy văn khi thời hạn dự báo kéo dài.

3.3.4. Ảnh hưởng của loại hình sử dụng đất và loại đất đến chất lượng dự báo

Kiểm định T-test ($p = 0,01$) xác nhận chất lượng dự báo khác biệt có ý nghĩa thống kê giữa các nhóm sử dụng đất.

Nhóm cây trồng (HNK, CLN, LUC) đạt hệ số R cao nhất (0,83–0,87 ở dự báo 1 tháng), $d > 0,89$ và RMSE ≈ 12, do độ nhạy cảm mạnh với hạn hán. Nhóm rừng (RSX, RDD, RPH) có R thấp hơn (0,57–0,73 ở dự báo 3 tháng) và RMSE cao hơn, phản ánh khả năng tự điều tiết nước tốt hơn của thực vật rừng.

Về loại đất, đất nâu đỏ trên đá bazan (Fu, Ru, Fs, Fk) cho chất lượng dự báo tốt nhất, nhờ khả năng giữ ẩm cao hỗ trợ mối quan hệ tuyến tính ổn định giữa chỉ số hạn và VHI; đất xám (X) và nhóm đất feralit trên đá macma axit và đá cát (Fa, Ha, Fq) có chất lượng dự báo kém hơn do thoát nước nhanh và độ ẩm biến động mạnh và điều kiện canh tác không thuận lợi.

KẾT LUẬN VÀ KIẾN NGHỊ

1. Kết luận

Luận án đã hoàn thành mục tiêu trong đánh giá đặc điểm hạn hán, xác định chỉ số thực vật phù hợp trong giám sát hán hán, xác định được các biến tiềm năng và xây dựng các mô hình có khả năng dự báo sớm hạn hán và độ chính xác tin cậy cho lưu vực Srêpốk.

Nghiên cứu đã xác định được cặp chỉ số VHI-SPI6 là phù hợp nhất cho giám sát hạn hán. SPI6 có quan hệ chặt chẽ với VHI từ tháng 2 đến tháng 5, phản ánh vai trò quyết định của lượng mưa tích lũy 6 tháng đến sức khỏe thực vật cuối mùa khô. Mức độ nhạy cảm với hạn hán giảm dần từ cây lâu năm, cây hàng năm, lúa, rừng trồng đến rừng tự nhiên, trong đó cây lâu năm là đối tượng chỉ thị tối ưu cho giám sát hạn hán trên lưu vực. Phân tích chuỗi dữ liệu VHI từ 2000-2022 cho thấy hạn hán có tính tựa chu kỳ với các đợt hạn nặng là 2004-2005, 2014-2016, 2019-2020 nghiêm trọng nhất. Khu vực trồng cà phê phía Bắc và Tây Nam chịu tác động nặng nhất. Hệ số tương quan khá cao giữa chỉ số hạn và năng suất cà phê, lúa phản ánh tác động chính của thiếu hụt nguồn nước đến năng suất cây trồng. Phân tích này đã chỉ ra SPI5 là phù hợp cho dự báo năng suất cà phê và lúa vụ Đông Xuân, và lượng mưa các tháng cuối mùa mưa giữ vai trò quan trọng đến việc hình thành năng suất.

Ba mô hình học máy ANN, SMLR và RF đã được xây dựng để dự báo VHI với thời gian từ 1 đến 3 tháng. ANN thể hiện khả năng vượt trội khi thời gian dự báo là từ 1 đến 2 tháng, SMLR ưu thế với thời gian dự báo 3 tháng, RF không được khuyến nghị. Chất lượng dự báo tốt nhất vào các tháng 2-5 là khoảng thời gian hạn nặng thường xuất hiện. Mức độ đóng góp của các biến vào chất lượng dự báo cho thấy với thời gian dự báo 1 tháng thì VHI giữ vai trò chính với mức đóng góp từ 74-88%, khi thời gian dự báo 2-3 tháng thì các chỉ số hạn khí tượng-thủy văn trở thành yếu tố chính. Độ chính xác dự báo VHI cũng phụ thuộc nhóm đất và cây trồng. Nhóm đất đỏ/bazan (Fu, Ru, Fs, Fk) cho độ chính xác cao hơn so với đất feralit-đất xám (Fq, X, Fa, Ha) do đặc tính thổ nhưỡng và điều kiện canh

tác thuận lợi. Các khu vực cây trồng nông nghiệp đạt độ chính xác cao hơn so với rừng, do thảm thực vật nông nghiệp nhạy cảm hơn với biến động nguồn nước.

Luận án đã phát triển khung phương pháp tích hợp viễn thám–khí tượng–thủy văn–học máy cho giám sát và dự báo sớm hạn nông nghiệp. Quy trình này có thể nhân rộng cho các lưu vực có điều kiện tương đồng, cung cấp công cụ cảnh báo sớm hạn hán phục vụ quản lý tài nguyên nước và phát triển nông nghiệp bền vững.

2. Kiến nghị

Luận án tập trung hạn hán nông nghiệp, chưa đánh giá cho các loại hạn khác. Dữ liệu MODIS sử dụng có độ phân giải chưa cao (463m) và do dữ liệu này có từ năm 2000 nên thời gian phân tích mới chỉ thực hiện trong giai đoạn 2000-2022. Chưa tích hợp được các sản phẩm của dự báo khí hậu và các yếu tố như độ ẩm đất, bốc hơi vào dự báo. Thời gian dự báo tối đa mới chỉ đạt 3 tháng. Chưa đánh giá sâu tác động kinh tế-xã hội, nguồn nước ngầm và hệ sinh thái.

Trong các nghiên cứu tiếp theo, cần bổ sung thêm các biến dự báo (độ ẩm đất, bốc hơi tiềm năng, chỉ số giám sát biến động khí hậu) và tích hợp kịch bản biến đổi khí hậu vào mô hình dự báo hạn hán để đánh giá xu thế hạn nông nghiệp trong tương lai. Đồng thời, sử dụng dữ liệu viễn thám có độ phân giải cao hơn như Sentinel-2 và Landsat thay thế MODIS, áp dụng thêm các mô hình học máy và mô hình lai ghép để nâng cao chất lượng và kéo dài thời gian dự báo. Bên cạnh đó, cần phân tích chi tiết hơn cho từng nhóm cây trồng cụ thể, bổ sung các yếu tố ảnh hưởng ngoài hạn hán như phát triển nguồn nước, giống và phân bón, đồng thời mở rộng phạm vi nghiên cứu trên không gian lớn hơn để ứng dụng kết quả vào thực tế.