

# KHẢ NĂNG DỰ BÁO SỐ NGÀY GIÁN ĐOẠN MƯA TRÊN KHU VỰC TÂY NGUYÊN BẰNG PHƯƠNG PHÁP THỐNG KÊ

Trần Trung Thành<sup>(1)</sup>, Lê Thị Diệu Thu<sup>(1)</sup>, Nguyễn Thị Thanh<sup>(2)</sup>,  
Tạ Hữu Chính<sup>(2)</sup>, Trương Thị Thanh Thủy<sup>(2)</sup>, Lại Thế Long<sup>(3)</sup>

<sup>(1)</sup>Đài Khí tượng Thủy văn khu vực Tây Nguyên

<sup>(2)</sup>Viện Khoa học Khí tượng Thủy văn và Biến đổi khí hậu

<sup>(3)</sup>Đài Khí tượng Thủy văn tỉnh Vĩnh Phúc

Ngày nhận bài: 25/11/2024; ngày chuyển phản biện: 26/11/2024; ngày chấp nhận đăng: 9/01/2025

**Tóm tắt:** Bài báo trình bày kết quả dự báo thử nghiệm số ngày gián đoạn mưa (NGĐ) trên khu vực Tây Nguyên so sánh giữa phương pháp hồi quy tuyến tính (PTT) và hồi quy phi tuyến (PPT). Nhân tố dự báo sử dụng là nhiệt độ bề mặt biển quan trắc được phân tích trực giao. Kết quả đánh giá chỉ ra, kỹ năng dự báo thời kỳ phụ thuộc tương đối cao, độ chính xác 3 pha từ 60-90% cho PPT vượt trội giá trị 40-60% cho PTT. Kỹ năng dự báo giảm đáng kể trên thời kỳ độc lập ở cả 2 phương pháp PTT và PPT, độ chính xác 3 pha giảm xuống chỉ còn dưới 60%. Điều đáng chú ý, trong giai đoạn dự báo độc lập, mặc dù kỹ năng của PPT có phần cao hơn so với PTT, nhưng không hoàn toàn vượt trội. Kết quả bên chỉ ra, trong nhiều trường hợp, kỹ năng dự báo tại một số trạm thấp hơn với PTT (PPT) nhưng lại cao hơn với PPT (PTT). Điều có thể nâng cấp chất lượng dự báo là sử dụng đồng thời 2 phương pháp này để hỗ trợ nhau.

**Từ khóa:** Tây Nguyên, gián đoạn mưa, mô hình thống kê.

## 1. Mở đầu

Trong một năm, lượng mưa có thể có sự biến động đáng kể, chẳng hạn, có nhiều ngày liên tục không có mưa hoặc có nhiều ngày mưa rất to. Các pha hoạt động hoặc gián đoạn của gió mùa thường xuyên kéo dài, tương ứng với lượng mưa lớn hoặc thấp, có thể dẫn đến các sự kiện gió mùa cực đoan [1]. Do đó, nghiên cứu dự báo các sự kiện hoạt động, gián đoạn của gió mùa có ý nghĩa thực tiễn to lớn trong việc hoạch định các chiến lược phù hợp, giúp tận dụng tốt các giai đoạn, đưa ra các chính sách ngắn hạn tối đa hóa năng suất nông nghiệp và xử lý các sự kiện gió mùa cực đoan.

Quá trình chuyển đổi từ pha gián đoạn sang hoạt động hỗn loạn hơn nhiều so với quá trình chuyển đổi từ pha hoạt động sang gián đoạn nên các sự kiện gián đoạn về bản chất dễ dự báo hơn so với các sự kiện hoạt động của của gió mùa. Tính không chắc chắn trong dự báo động

lực hạn mở rộng chủ yếu phát sinh từ sự không chắc chắn trong các điều kiện ban đầu và tính không chắc chắn của mô hình trong biểu diễn không đầy đủ về mặt vật lý của vấn đề [2], [3], [4]. Trên thế giới, nghiên cứu dự báo các sự kiện hoạt động, gián đoạn gió mùa mùa hè (GMMH) hạn dài và hạn mở rộng bước đầu được quan tâm trong những năm gần đây, đặc biệt là ở Ấn Độ và một số nơi khác như Nam Mỹ.

Saha và cộng sự (2017) dự báo các đợt hoạt động và gián đoạn GMMH Ấn Độ sử dụng bộ phân loại dựa trên phương pháp phân tích phân biệt tuyến tính để phân loại những ngày gió mùa khô, bình thường và ẩm ướt [5]. Kết quả cho thấy, các đợt gián đoạn được dự báo với độ chính xác cao hơn các đợt hoạt động và mở ra một hướng nghiên cứu mới trong việc đánh giá các cực đoan của thời kỳ gió mùa. Kết quả phân loại các ngày gió mùa có thể hỗ trợ việc dự báo năm gió mùa cực đoan (hạn hán hoặc lũ lụt) dựa vào tần suất của những ngày khô hoặc ẩm ướt. Grimm và cộng sự (2018) dự báo các giai đoạn hoạt động và gián đoạn của gió mùa Nam Mỹ sử dụng chỉ tiêu về lượng mưa cho thấy các dị

Liên hệ tác giả: Trần Trung Thành

Email: thanh06kontum@yahoo.com.vn

thường tổng hợp đối với các đợt mưa gió mùa gián đoạn trong vùng lõi gió mùa cho thấy có mối liên quan đến các dị thường gió đông xoáy nghịch quy mô lớn và dao động Madden-Julian (MJO) là một nguồn nhân tố quan trọng cho dự báo [6]. Một số nghiên cứu khác tiếp cận phương pháp tổ hợp sử dụng mô hình động lực [7]. Abilash và cộng sự (2014) dự báo các đợt mưa GMMH Ấn Độ gián đoạn và hoạt động hạn mở rộng bằng cách sử dụng hệ thống dự báo tổ hợp CFS của NCEP [8]. Pattanaik và cộng sự (2020) sử dụng hệ thống dự báo khí hậu phiên bản 2 cùng với hệ thống dự báo tổ hợp của cục Khí tượng Ấn Độ để xem xét hiệu quả của kết quả dự báo hạn mở rộng (extended-range forecast-ERF) thời gian thực của chu kỳ hoạt động gián đoạn gió mùa tại các quy mô không gian khác nhau trong năm 2017-2018. Kết quả chỉ ra rằng ERF nghiệp vụ có độ chính xác cao trong việc dự báo sự chuyển tiếp giữa các pha hoạt động gián đoạn hoạt động của GMMH dựa trên các chỉ số quy mô lớn bao gồm dao động nội mùa gió mùa và hoàn lưu mực thấp [9]. Ngoài ra, các mô hình SeqToSeq (Sequence-to-sequence) và LSTM (Long short-term memory) dựa trên phương pháp tiếp cận học sâu bước đầu được sử dụng trong việc xác định, dự báo các đợt hoạt động, gián đoạn của gió mùa Ấn Độ và cho kết quả tốt hơn các mô hình truyền thống [1]. Đầu tiên, các mô hình dựa vào học sâu phân loại các ngày mưa gió mùa, sau đó tập hợp các ngày gió mùa được phân loại để xác định các đợt gió mùa hoạt động và gián đoạn. Các ngày khô và ướt được phân loại với độ chính xác lần lượt là 0,95 và 0,87. Dự báo các đợt gián đoạn có độ chính xác cao hơn các đợt hoạt động.

Gần đây, Saha và cộng sự (2020) đưa ra mô hình dự báo về các đợt GMMH bình thường và các đợt hoạt động/gián đoạn với thời gian dự báo trước 5 và 10 ngày. Mô hình dự báo nghiên cứu mối quan hệ không gian, thời gian của các biến khí hậu bằng cách sử dụng mạng thần kinh tích chập (Convolutional Neural Network-CNN) [10]. Mô hình dự báo chính xác lượng mưa trong thời kỳ hoạt động và gián đoạn với điểm số AUC (Area Under Curve) là 0,81 và 0,84. Đánh giá hiệu quả của mô hình so với các mô hình tiên tiến nhất cho thấy sự cải thiện đáng kể về kỹ

năng dự báo.

Một số nghiên cứu tại các khu vực Ấn Độ, Biển Đông, Philippine và Tây Thái Bình Dương cho thấy có mối liên hệ chặt chẽ giữa sự xuất hiện của các giai đoạn gián đoạn và pha gián đoạn của dao động nội mùa. [11], [12], [13], [14], [15], [16] các dao động có chu kỳ 30-60 ngày gây ra các giai đoạn gián đoạn mưa trên khu vực Tây Thái Bình Dương dài hơn so với các dao động có chu kỳ 10-20 ngày. Việc chỉ ra mối liên hệ giữa giai đoạn gián đoạn gió mùa và dao động nội mùa có ý nghĩa quan trọng trong bài toán dự báo, đặc biệt dự báo hạn dài.

Ở Việt Nam, các nghiên cứu dự báo các sự kiện hoạt động và gián đoạn của gió mùa hay mưa gió mùa còn khá hạn chế. Hầu hết các nghiên cứu tập trung dự báo ngày bắt đầu và lượng mưa gió mùa sử dụng các phương pháp thống kê truyền thống và phương pháp động lực [17], [18], [19], [20], [21]. Một số các nghiên cứu bước đầu ứng dụng sản phẩm mô hình CFS của NCEP trong việc dự báo ngày bắt đầu gió mùa trên khu vực Tây Nguyên và Nam Bộ và đưa ra kết quả dự báo được đánh giá là có tiềm năng [22].

Ở Tây Nguyên, vấn đề phòng chống thiên tai và an ninh lương thực đặt lên hàng đầu và gắn liền với chất lượng dự báo thời tiết - khí hậu khu vực, trong đó có dự báo mưa lớn và hoạt động của gió mùa, đặc biệt các sự kiện gián đoạn gió mùa liên quan đến sự thiếu hụt nước trong mùa mưa. Nằm ở phía Nam Việt Nam thuộc Bán đảo Đông Dương, kết hợp với địa hình cao và phức tạp nên chế độ mưa ở Tây Nguyên bị ảnh hưởng sâu sắc bởi địa hình và hoàn lưu gió mùa. Trong khi phần sườn Tây của dãy Trường Sơn bị chi phối chủ yếu bởi hệ thống GMMH thì phần sườn Đông của dãy Trường Sơn bị ảnh hưởng bởi hệ thống gió mùa Đông Bắc (đỉnh vào thời kỳ thu đông) [23], [24], [25]. Trong nghiên cứu này, chúng tôi thử nghiệm khả năng dự báo số đợt gián mưa liên quan đến GMMH từ tháng 4 đến tháng 11 trên khu vực Tây Nguyên bằng cách sử dụng phương pháp thống kê truyền thống. Các đợt gián đoạn mưa gió mùa có thể xác định đơn thuần dựa trên lượng mưa, phát xạ sóng dài hoặc sự thay đổi của các hình thể synop [26], [27], [28], [29], [30], [31]. Tuy nhiên, mùa mưa

ở Tây Nguyên thường bắt đầu sớm hơn so với sự xuất hiện của gió mùa [18] và mưa tiền gió mùa thường nhỏ và không duy trì đều đặn, nên chỉ tiêu về lượng mưa mưa được sử dụng trong việc xác định các sự kiện gián đoạn mưa GMMH, giúp phân biệt mưa tiền gió mùa và mưa GMMH [32].

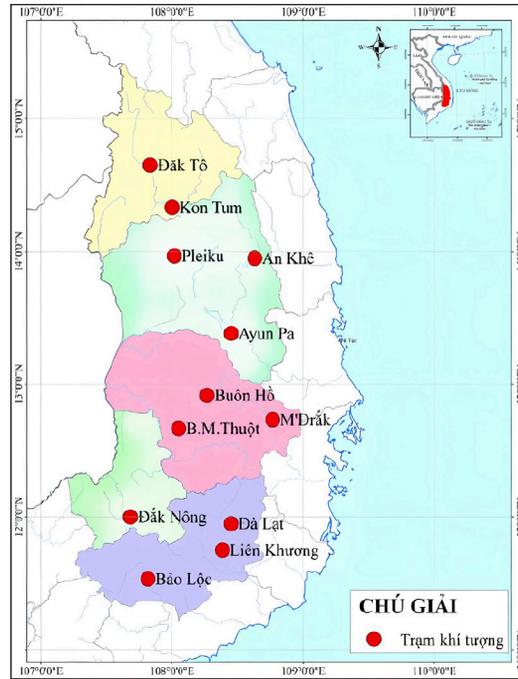
Bài báo được tổ chức như sau: Phần 2 trình bày về số liệu và phương pháp nghiên cứu, Phần 3 trình bày kết quả thử nghiệm.

## 2. Số liệu và phương pháp nghiên cứu

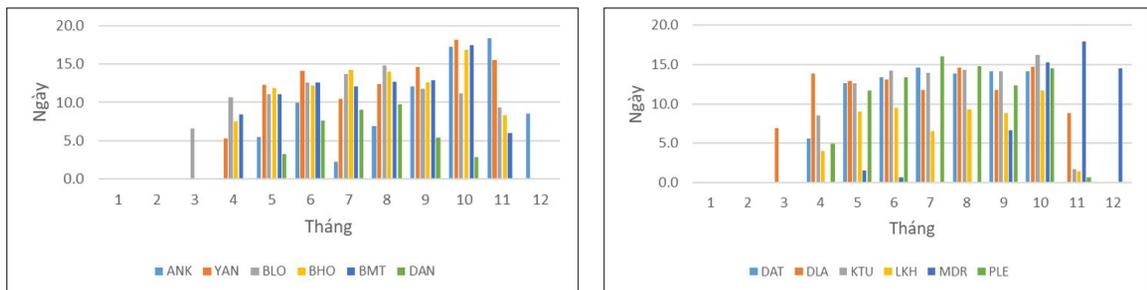
### 2.1. Khu vực nghiên cứu

Khu vực nghiên cứu là phần lãnh thổ Tây Nguyên, đặc trưng về số ngày gián đoạn mưa (NGĐ) được tính toán từ 12 trạm khí tượng

(Hình 1). Mùa mưa khu vực Tây Nguyên bắt đầu vào 5 và kết thúc tháng 11. Dao động mùa của đặc trưng NGĐ (Hình 2) cho thấy sự phù hợp với biến động mưa mùa ở khu vực này. NGĐ phổ biến trên toàn khu vực dao động trong khoảng 8-14 ngày trong tháng 5-tháng 10, 11, một số trạm có biến động mạnh hơn (khoảng 14-18 ngày) vào tháng 10, 11. Riêng trạm M.Drak (kí hiệu MDR, Hình 2) có dao động năm của NGĐ bắt đầu từ khoảng tháng 9 đến tháng 12. Điều này có thể do đặc điểm địa phương của địa hình dẫn tới dao động mưa mùa của M.Drak có cực đại vào tháng 9 đến tháng 12 (thời kỳ gió mùa mùa đông). Vì vậy, trong các kết quả phân tích bên dưới, không có kết quả của trạm M.Drak.



Hình 1. Số trạm quan trắc của khu vực Tây Nguyên lựa chọn



Hình 2. Số ngày gián đoạn mưa (NGĐ) trung bình khí hậu thời kỳ 1981-2021

## 2.2. Số liệu

- Số liệu quan trắc: Số liệu mưa ngày tháng 4 đến tháng 11 giai đoạn 1981-2021 được sử dụng để tính toán số ngày gián đoạn mưa (NGĐ) trong giai đoạn mùa mưa. Số liệu nhận được từ hệ thống trạm quan trắc bề mặt của Tổng Cục Khí tượng Thủy văn, đã được quản lý và kiểm soát chất lượng.

- Số liệu nhiệt độ bề mặt biển (SST) tái phân tích toàn cầu nhận được từ Trung tâm Khí quyển và Đại dương Hoa Kỳ giai đoạn 1981-2024, dùng làm nhân tố dự báo, xây dựng dự báo NGĐ.

- Số liệu sử dụng tại địa chỉ (<https://psl.noaa.gov/data/gridded/index.html>). Trường nhiệt độ bề mặt biển được lấy trong khu vực vùng biển Thái Bình Dương và Ấn Độ Dương (60°E-300°E, -50°S-50°N), để có thể bao gồm trong đó thông tin các dao động khí quyển - đại dương quy mô lớn như: ENSO hay IOD.

## 2.3. Phương pháp nghiên cứu

### 2.3.1. Phương pháp phân tích và dự báo

- Phương pháp xác định các ngày gián đoạn mưa (NGĐ) khu vực Tây Nguyên dựa trên số liệu mưa ngày quan trắc tại trạm. Số liệu mưa đầu tiên được lọc để loại bỏ đi các phổ dao động tần số thấp lớn hơn quy mô mùa (120, 180, 360 ngày). Giá trị dị thường mưa ngày được tính toán trên chuỗi số liệu sau khi được lọc. Giá trị dương (âm) của lượng mưa biểu hiện thời kỳ mưa nhiều (ít). Chỉ những thời kỳ gián đoạn mưa kéo dài từ 5 ngày trở lên, và có ít nhất 1 ngày có giá trị dị thường lượng mưa nhỏ hơn giá trị độ lệch chuẩn của toàn chuỗi, được xem là 1 đợt gián đoạn mưa có tổng NGĐ là 5 ngày. NGĐ được thực hiện phân tích cho từng trạm, nên giữa những trạm có sự khác biệt nhất định. Số liệu sẽ được chia làm 2 giai đoạn, gồm phụ thuộc (từ năm 1981 đến năm 2015) và độc lập (từ năm 2016 đến năm 2021).

- Phương pháp hồi quy tuyến tính (PTT): Được sử dụng để ước lượng dự báo trên chuỗi phụ thuộc và độc lập với yếu tố dự báo là NGĐ, nhân tố dự báo là nhiệt độ bề mặt biển quan trắc (trên lưới). Phương trình dự báo được xây dựng như sau:

$$Y = a_1 \cdot X_1 + a_2 \cdot X_2 + \dots + a_n \cdot X_n + b$$

Trong đó, Y là yếu tố dự báo;  $X_n$  là nhân tố dự báo.

- Phương pháp hồi quy phi tuyến (PPT): Khác với PTT, phương trình hồi quy là 1 đường cong, với hy vọng có thể bắt sát hơn đặc tính của dữ liệu. Có một số phương pháp phi tuyến đã được nghiên cứu áp dụng cho lĩnh vực dự báo, điển hình như mạng thần kinh nhân tạo (ANN). Tuy nhiên, nhược điểm của cách làm này là mẫu dữ liệu lịch sử trong nghiên cứu tương đối nhỏ, nên có thể thuật toán phức tạp của ANN không phát huy được hiệu quả. Một số nghiên cứu đã ứng dụng hồi quy phi tuyến bậc 2 cho dự báo mưa mùa và chỉ ra tính hiệu quả [34]. Nghiên cứu này vì vậy muốn ứng dụng thử nghiệm hồi quy phi tuyến bậc 2 cho dự báo NGĐ.

Phương trình dự báo được xây dựng như sau:

$$Y = a_{11} \cdot X_1 + a_{12} \cdot X_2 + \dots + a_{1n} \cdot X_1^2 + a_{12} \cdot X_1 \cdot X_2 + a_{22} \cdot X_2^2 + b$$

Trong đó, Y là yếu tố dự báo;  $X_n$  là nhân tố dự báo

- Yếu tố dự báo: Trong nghiên cứu này, yếu tố cần thực hiện dự báo là NGĐ được tính toán từ dữ liệu mưa quan trắc từ 12 trạm khu vực Tây Nguyên. Nghiên cứu này chỉ tập trung vào dự báo thử nghiệm trong mùa hoạt động mạnh của hệ thống gió mùa Tây Nam (tháng 6, 7, 8) - nguyên nhân chính gây mưa khu vực Tây Nguyên.

- Nhân tố dự báo: Là trường nhiệt độ bề mặt biển gồm vùng biển Thái Bình Dương và Ấn Độ Dương. Mục đích sử dụng miền tính này là để bao hàm được biến động của các tín hiệu dao động đại dương - khí quyển quy mô lớn như ENSO hay IOD, ....

- Phân tích tích thành phần chính (phân tích trực giao): Phương pháp được ứng dụng để giảm số chiều dữ liệu, điều này cho phép tuyến chọn được các thông tin có ý nghĩa làm nhân tố dự báo, tránh sự trùng lặp và sẽ thuận tiện hơn khi xây dựng các phương trình hồi quy.

- Trường nhiệt độ bề mặt biển được đưa vào phân tích trực giao và chọn khoảng 10 thành phần chính đầu tiên với tổng lượng phương sai khoảng 95% làm nhân tố dự báo.

- Hạn dự báo (hdb): Các thử nghiệm được thực hiện cho 3 hdb khác nhau gồm 1, 2 và 3 tháng. Ví dụ: Để dự báo NGĐ tháng 6, thì với trường nhân tố dự báo SST tại tháng 3, tháng 4, tháng 5 lần lượt gọi là hdb3 (trước 3 tháng), hdb2 (trước 2 tháng), hdb1 (trước 1 tháng)

### 2.3.2. Đánh giá kết quả dự báo

- Bài báo sử dụng các chỉ số đánh giá độ lệch: Sai số trung bình (ME), sai số trung bình quân phương (RMSE), đánh giá mức độ quan hệ của giá trị dự báo và quan trắc - hệ số tương quan (HSTQ), chỉ số kỹ năng (MSSS) - để so sánh mức độ biến động của sai số so với biến động khí hậu để kiểm nghiệm độ chính xác của kết quả dự báo.

$$MSSS = 1 - MSE/MSEc$$

MSE là biến động của sai số; MSEc là biến động khí hậu của yếu tố; MSSS dương và càng gần đến 1, dự báo càng có kỹ năng.

- Chỉ số đánh giá độ chính xác 3 pha (Accu3P) được sử dụng để đánh giá về khả năng dự báo pha của mô hình. Dựa trên chuỗi số liệu lịch sử, giá trị quan trắc của yếu tố được chia làm 3 pha - dưới chuẩn, cận chuẩn và trên chuẩn. Trong đó, pha cận chuẩn được định nghĩa từ phân vị 33

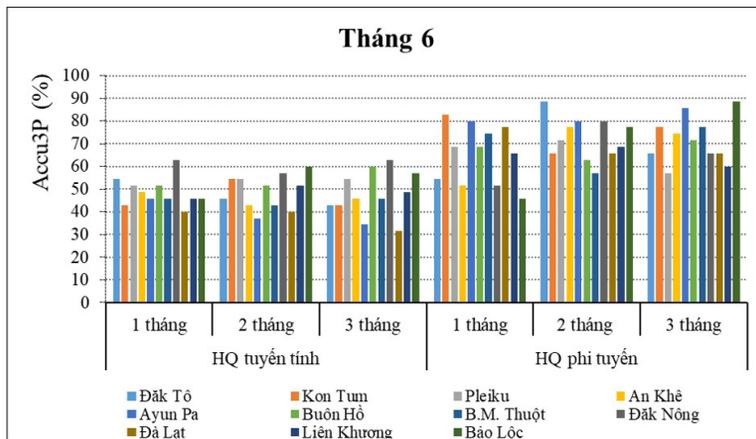
đến 67 của chuỗi dữ liệu lịch sử. Giá trị nhỏ hơn ngưỡng 33 là dưới chuẩn, lớn hơn 67 là trên chuẩn. Nếu giá trị dự báo nằm trong cùng pha với quan trắc được tính là 1 trường hợp đúng, ngược lại là dự báo sai. Giá trị Accu3P được đếm như phần trăm dự báo đúng trên tổng số sự kiện.

## 3. Kết quả nghiên cứu

### 3.1. Kết quả dự báo ngày gián đoạn trong chuỗi số liệu phụ thuộc

#### 3.1.1. Tháng 6

Hình 3 trình bày độ chính xác 3 pha (Accu3P) so sánh giữa kết quả dự báo bằng phương pháp hồi quy tuyến tính (PTT) và phương pháp phi tuyến (PPT). Kết quả cho thấy, Accu3P tháng 6 đều vượt qua giá trị ngẫu nhiên (33%) tại hầu hết các trạm đối với PTT (trừ trạm Đà Lạt tại hdb3 có giá trị nhỏ hơn 33%) và tất cả các trạm đối với PPT. PPT cho kết quả cao hơn vượt trội so với PTT. Cụ thể, với hdb1&2, PTT cho dự báo tốt hơn hdb3 với độ chính xác là 40-63%, 37-60% và 31-63%. PPT, với hdb2&3 nhìn chung cho kỹ năng cao hơn hdb1 (trừ Kon Tum và Đà Lạt), với độ chính xác lần lượt là: 57-83% (hdb2), 57-87% (hdb3), 46-83% (hdb1); độ chính xác cao nhất tại các trạm Đắk Tô, Ayun Pa, Đắk Nông (80-83%) tại hdb2 và tại YaunPa, Bảo Lộc (86-89%) tại hdb3.



Hình 3. Độ chính xác dự báo 3 pha (Accu3P) về dự báo NGĐ trong tháng 6 trên khu vực Tây Nguyên, với thời hạn dự báo 1, 2, 3 tháng (kí hiệu: hdb1, 2, 3)

Kết quả Bảng 1 chỉ ra các chỉ số đánh giá khác (ME, MAE, HSTQ và MSSS). HSTQ dương cao và MSSS dương (0,3-0,7). Lead 2&3 cho dự báo cao hơn lead 1 trên phần lớn các trạm, ngoại trừ

Pleiku có kỹ năng cao nhất tại hdb1. Với hdb1, mô hình dự báo thiên cao hơn quan trắc tại các trạm Đắk Tô, Kon Tum, An Khê, Đắk Nông, Liên Khương; thiên thấp hơn quan trắc (ME âm) tại

các trạm Buôn Hồ, Đà Lạt, Bảo Lộc. HSTQ dương cao xuất hiện ở tất cả các trạm, với giá trị phổ biến từ 0,7 đến 0,9, riêng hai trạm An Khê, Bảo Lộc có giá trị từ 0,5 đến 0,7. Chỉ số RMSE và MSSS lần lượt là 2,2-5,0 và 0,0-0,6. Với hdb2, dự báo thiên thấp hơn quan trắc tại phần lớn các trạm, thiên cao hơn quan trắc tại Pleiku, An Khê, Buôn Ma Thuột và không có xu thế rõ ràng tại Kon Tum, M'drak; HSTQ từ 0,8-1,0, tăng so với hdb1; chỉ số RMSE dao động từ 0,5 đến 3,7 đợt, lớn

nhất tại Kon Tum, M'Đrăk và nhỏ nhất tại Đăk Nông, Bảo Lộc; chỉ số kỹ năng MSSS phổ biến từ 0,5-0,7, ngoại trừ các trạm Kon Tum, Buôn Ma Thuột dưới 0,5. Với hdb3, dự báo thiên cao hơn quan trắc tại Đăk Tô, Kon Tum, Pleiku, Buôn Ma Thuột, Đà Lạt, Liên Khương và thiên thấp hơn tại các trạm còn lại; HSTQ dương và cao, dao động từ 0,7 đến 1,0; chỉ số RMSE dao động từ 0,5 đến 6,1, lớn nhất tại Pleiku và nhỏ nhất tại M'Đrăk; chỉ số kỹ năng MSSS dương, đạt từ 0,2 đến 0,7.

Bảng 1. Các chỉ số đánh giá của PPT dự báo NGĐ trong tháng 6

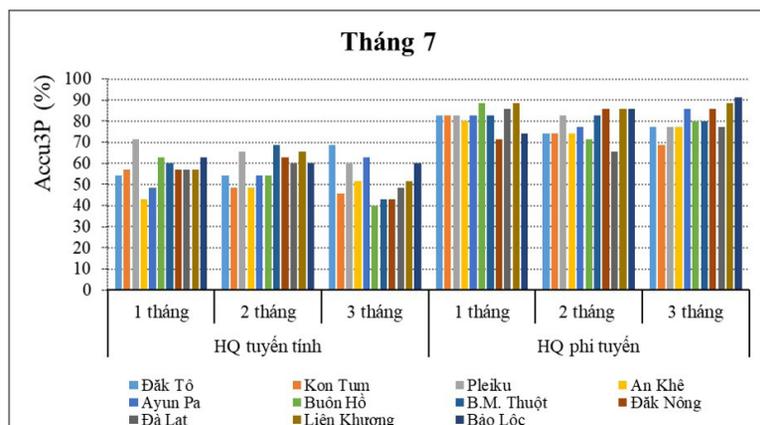
Trạm	Hạn dự báo 1 tháng				Hạn dự báo 2 tháng				Hạn dự báo 3 tháng			
	ME	RMSE	MSSS	HSTQ	ME	RMSE	MSSS	HSTQ	ME	RMSE	MSSS	HSTQ
Đăk Tô	0,8	5,0	0,3	0,7	-0,6	1,9	0,7	0,9	1,0	4,4	0,3	0,8
Kon Tum	3,8	2,3	0,6	0,9	0,0	3,7	0,4	0,8	0,2	2,0	0,7	0,9
Pleiku	0,2	2,8	0,6	0,9	0,8	3,2	0,6	0,9	0,2	6,1	0,2	0,7
An Khê	0,4	4,8	0,1	0,6	0,1	2,3	0,6	0,9	-0,8	2,3	0,6	0,9
Ayun Pa	0,3	2,2	0,6	0,9	-0,2	2,2	0,6	0,9	-0,9	1,6	0,7	0,8
Buôn Hồ	-0,6	3,5	0,4	0,8	-0,9	3,0	0,5	0,9	-0,6	2,9	0,5	0,9
Buôn Ma Thuột	-0,2	2,2	0,6	0,9	0,2	3,7	0,3	0,8	0,2	1,7	0,7	0,8
Đăk Nông	0,1	3,2	0,3	0,8	-0,2	1,6	0,7	0,9	-0,1	3,0	0,4	0,8
Đà Lạt	-0,2	2,8	0,4	0,8	-0,4	2,1	0,6	0,9	0,5	3,2	0,3	0,8
Liên Khương	0,3	3,1	0,4	0,8	-1,3	2,2	0,6	0,9	0,8	4,1	0,3	0,7
Bảo Lộc	-0,6	4,9	0,0	0,5	-0,7	1,8	0,6	0,9	-0,2	1,6	0,7	0,9

### 3.1.2. Tháng 7

Kết quả Hình 4 cho thấy, PPT có độ chính xác vượt trội so với PTT. Cụ thể, Accu3P tại các hdb1, 2, 3 lần lượt là 43-47%, 49-69%, 40-69% cho PTT; trong đó, hdb1&2 cho dự báo cao hơn hdb3, ngoại trừ các trạm Đăk Tô, An Khê, Ayun Pa. Trong khi đó, PPT có Accu3P là 71-89%, 66-86%, 69-91% lần lượt với hdb1, 2, 3; kỹ năng dự báo phổ biến cao hơn trong lead 1 ở phần lãnh thổ phía Bắc, hdb2&3 nhỉnh hơn ở khu vực phía Nam.

Bảng 2 chỉ ra các chỉ số đánh giá thống kê của PPT. Hdb1 có kỹ năng dự báo cao hơn lead 2&3,

ngoại trừ các trạm Đăk Nông, Bảo Lộc. Với hdb1, dự báo thiên cao hơn quan trắc tại các trạm Ayun Pa, Buôn Hồ, Buôn Ma Thuột, Bảo Lộc; thiên thấp hơn quan trắc (ME âm) tại các trạm Đăk Tô, Kon Tum, An Khê, Đăk Nông, Đà Lạt, Liên Khương; và không có xu thế rõ ràng tại Pleiku. HSTQ dương cao xuất hiện ở tất cả các trạm, đạt giá trị từ 0,9 đến 1,0. Sai số dự báo nhỏ, từ 1 đến 3. Chỉ số kỹ năng MSSS có giá trị dương và cao, từ 0,6 đến 0,8. Với hdb2, chỉ số RMSE, MSSS và HSTQ lần lượt là: 1,0-3,1 đợt; 0,5-0,8 và 0,9-1,0. Với hdb3, chỉ số RMSE, MSSS và HSTQ lần lượt là: 0,8-3,7 đợt; 0,3-0,8 và 0,8-1,0.



Hình 4. Độ chính xác dự báo 3 pha (Accu3P) về dự báo NGĐ trong tháng 7 trên khu vực Tây Nguyên, với thời hạn dự báo 1, 2, 3 tháng (kí hiệu: hdb1, 2, 3)

Bảng 2. Các chỉ số đánh giá của PPT về dự báo NGĐ trong tháng

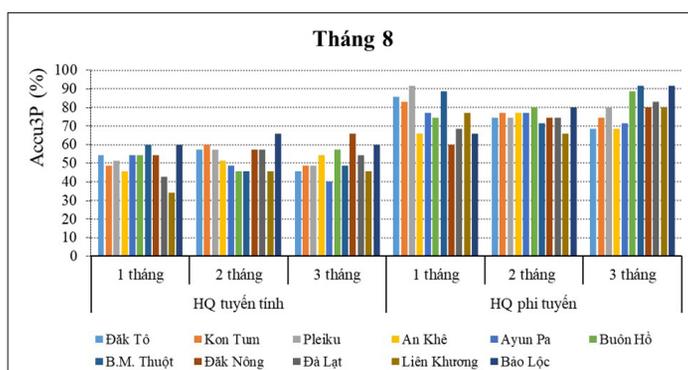
Trạm	Hạn dự báo 1 tháng				Hạn dự báo 2 tháng				Hạn dự báo 3 tháng			
	ME	RMSE	MSSS	HSTQ	ME	RMSE	MSSS	HSTQ	ME	RMSE	MSSS	HSTQ
Đắk Tô	-0,8	1,4	0,8	0,9	0,6	2,4	0,6	0,9	-0,6	2,4	0,6	0,9
Kon Tum	-1,4	1,8	0,7	0,9	0,2	1,9	0,7	0,9	0,2	3,7	0,3	0,8
Pleiku	0,0	2,7	0,6	0,9	0,6	3,1	0,6	0,9	-0,9	2,3	0,7	0,9
An Khê	-0,5	1,0	0,6	0,9	-0,7	1,0	0,6	0,9	-0,7	0,8	0,7	0,8
Ayun Pa	0,4	1,5	0,7	0,8	0,5	1,9	0,6	0,9	-0,3	1,6	0,7	0,9
Buôn Hồ	0,1	2,9	0,6	0,9	0,8	3,1	0,5	0,8	0,0	2,6	0,6	0,9
Buôn Ma Thuật	0,4	1,9	0,7	0,9	0,2	1,9	0,7	0,9	-0,7	2,3	0,6	0,9
Đắk Nông	-0,4	2,1	0,7	0,8	1,6	1,7	0,8	0,9	1,2	1,4	0,8	0,8
Đà Lạt	-0,9	1,2	0,8	0,9	-1,3	1,9	0,7	0,9	0,6	2,1	0,6	0,9
Liên Khương	-0,8	0,9	0,8	0,8	-1,2	1,4	0,7	0,9	0,2	1,4	0,7	0,9
Bảo Lộc	1,2	2,1	0,7	0,9	0,0	1,7	0,7	0,9	-0,7	1,5	0,8	0,8

### 3.1.3. Tháng 8

Kết quả trình bày trên Hình 5 có xu thế tương tự như tháng 6, 7; PPT có kỹ năng cao hơn PTT. Accu3P tại hdb1, 2, 3 theo PTT lần lượt là 34-60%, 46-66%, 40-66%; theo PPT lần lượt là: 60-91%, 66-80%, 69-91%. Với PPT, nhìn chung, hdb2 có kỹ năng đồng đều hơn hdb1&3; trong khi hdb3 có kỹ năng cao hơn ở các tỉnh Trung và Nam Tây Nguyên (Đắk Lắk, Đắk Nông, Lâm Đồng); và hdb1 có kỹ năng nhỉnh hơn ở các tỉnh

phía Bắc.

Kết quả Bảng 3 cho thấy, với PPT, dự báo cao nhất tại hdb1 đối với các tỉnh phía Bắc (Kon Tum, Gia Lai) và hdb3 ở các tỉnh Trung và Nam Tây Nguyên (Đắk Lắk, Đắk Nông, Lâm Đồng). Các chỉ số thống kê ME, RMSE, MSSS và HSTQ tại các hạn dự báo 1 tháng, 2 tháng, 3 tháng lần lượt là: -2,0-2,3, 1,5-4,5; 0,3-0,7; 0,8-1,0; -1,1-0,5, 1,8-3,3; 0,4-0,7, 0,8-0,9 và -2,6-1,1, 1,2-3,5; 0,4-0,8, 0,8-1,0.



Hình 5. Độ chính xác dự báo 3 pha (Accu3P) về dự báo NGĐ trong tháng 8 trên khu vực Tây Nguyên, với thời hạn dự báo 1, 2, 3 tháng (Kí hiệu: hdb1, 2, 3)

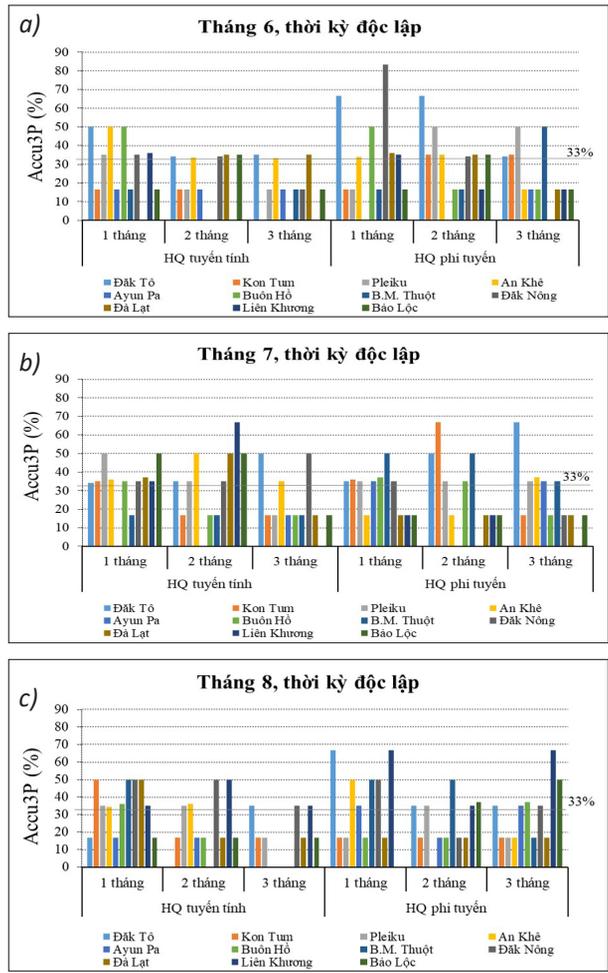
Bảng 3. Các chỉ số đánh giá của PPT về dự báo NGĐ trong tháng 8

Trạm	Hạn dự báo 1 tháng				Hạn dự báo 2 tháng				Hạn dự báo 3 tháng			
	ME	RMSE	MSSS	HSTQ	ME	RMSE	MSSS	HSTQ	ME	RMSE	MSSS	HSTQ
Đắk Tô	0,1	2,0	0,7	0,8	-1,1	2,1	0,7	0,9	0,8	2,4	0,6	0,9
Kon Tum	-2,0	2,3	0,6	0,9	-0,3	3,0	0,5	0,9	-2,0	3,5	0,5	0,9
Pleiku	0,4	1,5	0,7	0,8	0,3	1,9	0,6	0,9	-1,5	2,8	0,5	0,9
An Khê	0,2	2,0	0,6	0,9	0,4	1,8	0,6	0,9	0,5	3,2	0,4	0,8
Ayun Pa	0,7	2,1	0,6	0,9	-0,4	2,4	0,5	0,9	0,1	1,9	0,6	0,9
Buôn Hồ	-1,0	1,6	0,7	0,9	-0,1	1,9	0,7	0,9	-1,4	1,8	0,7	0,9
Buôn Ma Thuột	-0,7	1,8	0,7	0,9	0,1	2,8	0,5	0,9	-0,9	1,4	0,7	1,0
Đắk Nông	0,1	3,4	0,3	0,8	-0,6	1,8	0,6	0,9	-2,6	1,2	0,8	1,0
Đà Lạt	-0,2	2,5	0,6	0,9	0,5	2,5	0,6	0,9	-1,0	2,5	0,6	0,9
Liên Khương	1,3	1,7	0,7	0,9	-0,8	3,3	0,4	0,8	1,1	1,4	0,7	1,0
Bảo Lộc	0,4	4,5	0,4	0,8	-1,0	2,3	0,7	0,9	-0,2	1,4	0,8	1,0

### 3.2. Kết quả dự báo ngày gián đoạn trên số liệu độc lập

Trong tháng 6, độ chính xác dự báo 3 pha của PPT cao hơn so với PTT. Cụ thể, PTT có kỹ năng dự báo cao hơn tại lead 1&2, với khoảng 50% số trạm có trị số vượt giá trị ngẫu nhiên (33%). Trong khi, PPT có khoảng 2/3 số trạm thỏa mãn tiêu chí này (hdb1&2). Như vậy, PPT cho độ chính xác dự báo 3 pha cao nhất tại lead 1&2. Trong tháng 7, trái lại, PTT có kỹ năng nhỉnh hơn ở hdb1&2, nhưng yếu hơn trong hdb3. Điều đáng chú ý, một số trường hợp có kỹ năng yếu hơn trong PPT (PTT) nhưng lại mạnh hơn trong PTT (PPT). Ví dụ, với hdb1&2, các trạm Kon Tum, Liên Khương và Đà Lạt có kỹ năng yếu trong PPT

(nhỏ hơn giá trị ngẫu nhiên) nhưng lại mạnh trong PTT. Với hdb3, PTT có trên 2/3 số trạm nhỏ hơn giá trị ngẫu nhiên, trong khi PPT chỉ ra gần 50% số trạm vượt qua ngưỡng này. Trong tháng 8, tại hdb1&2, kỹ năng dự báo của PPT và PTT có xu hướng cân bằng nhau. Với hdb3 - kỹ năng của PPT nhỉnh vượt trội khi có trên 50% số trạm vượt quá giá trị ngẫu nhiên, trong khi hầu hết các trạm của PTT thấp hơn trị số này. Như vậy, kết quả đánh giá trên số liệu độc lập mùa tháng 6, 7, 8 (Hình 6) cho thấy, mặc dù kỹ năng của PPT có xu hướng cao hơn so với PTT, nhưng không hoàn toàn vượt trội. Trái lại, những trường hợp kỹ năng yếu trong PPT có thể cao hơn trong PTT. Vì vậy, PPT và PTT có thể cùng tồn tại để hỗ trợ nhau khi thực hiện dự báo.



Hình 6. Độ chính xác dự báo 3 pha (Accu3P) về dự báo NGĐ trong các tháng 6, 7, 8 (thời kỳ độc lập) trên khu vực Tây Nguyên

#### 4. Kết luận

Bài báo đã đưa ra kết quả dự báo NGĐ dựa trên nhân tố dự báo là hình thể nhiệt độ bề mặt biển quan trắc được đưa vào phân tích trực giao. Kết quả đánh giá chỉ ra, kỹ năng dự báo thời kỳ phụ thuộc tương đối cao, Accu3P từ 60-90% cho PPT vượt trội giá trị 40-60% cho PTT. Kỹ năng dự báo giảm đáng kể trên thời kỳ độc lập ở cả 2 phương pháp PTT và PPT, Accu3P giảm xuống chỉ còn dưới 60%. Điều đáng chú ý, trong giai đoạn dự báo độc lập, mặc dù kỹ năng của PPT có phần cao hơn so với PTT, nhưng không hoàn toàn vượt trội. Kết quả chỉ ra, trong nhiều

trường hợp, kỹ năng dự báo tại một số trạm thấp hơn với PTT (PPT) nhưng lại cao hơn với PPT (PTT). Kết quả đưa ra kiến nghị, có thể sử dụng đồng thời 2 phương pháp này để hỗ trợ nhau nâng cao chất lượng dự báo. Dù vậy, cần thêm nhiều hơn các nghiên cứu thử nghiệm ứng dụng với các phương pháp học máy, tìm kiếm những điểm mới.

Kỹ năng dự báo cao trên thời kỳ phụ thuộc nhưng thấp hơn trên thời kỳ độc lập của PPT có thể do hiệu ứng “over-fitting”; điều này có thể được xử lý bằng một số phương pháp. Chúng tôi hy vọng có thể thực hiện điều này trong nghiên cứu tiếp theo.

**Đóng góp của từng tác giả trong bài báo:** Xây dựng ý tưởng: Trần Trung Thành; Xử lý số liệu: Lại Thế Long, Nguyễn Thị Thanh; Xây dựng mô hình: Tạ Hữu Chính, Trần Trung Thành, Trương Thị Thanh Thủy, Lê Thị Diệu Thu.

**Lời cảm ơn:** Bài báo hoàn thành nhờ vào kết quả của Đề tài: "Nghiên cứu ảnh hưởng của gián đoạn gió mùa mùa hè đến thiếu hụt nước trong mùa mưa ở khu vực Tây Nguyên" Mã số: TNMT.2023.06.06.

**Lời cam đoan:** Tập thể tác giả cam đoan bài báo này là công trình nghiên cứu của mình, chưa từng công bố trước đó, không sao chép, đạo văn; không có sự tranh chấp lợi ích trong nhóm tác giả.

#### Tài liệu tham khảo

1. Viswanath, S. et al. (2019), "Deep learning-based LSTM and SeqToSeq models to detect monsoon spells of India", in *Computational Science–ICCS 2019: 19th International Conference*, Faro, Portugal, June 12-14, 2019, Proceedings, Part II, 19, 204-218. Springer International Publishing.
2. Goswami, B.N., and Xavier, P.K. (2003), "Potential predictability and extended range prediction of Indian summer monsoon breaks", *Geophysical Research Letters*, 30(18), 9-1-9-4. DOI:10.1029/2003GL017810.
3. Xavier, P.K., and Goswami, B.N. (2007a), "Promising alternative to prediction of seasonal mean all India rainfall", *Current Science*, Vol 93, Issue 2, 195-202. <https://doi.org/xxxx>
4. Xavier, P.K., and Goswami, B.N. (2007b), "An analog method for real-time forecasting of summer monsoon sub-seasonal variability", *Monthly Weather Review*, 135, 4149-4160.
5. Saha, M., Mitra, P., and Nanjundiah, R. S. (2017, December), *Prediction of active and break spells of Indian summer monsoon using linear discriminant analysis*, Proceedings of the Ninth International Conference on Advances in Pattern Recognition (ICAPR), 1-6. IEEE.
6. Grimm, A. et al. (2018, April), *Sub-seasonal prediction of active and break phases of the South American monsoon and the influence of the MJO*, EGU General Assembly Conference Abstracts, 1317.
7. Abhilash, S., et al. (2013), "Predictability during active break phases of Indian summer monsoon in an ensemble prediction system using climate forecast system", *Journal of Atmospheric and Solar-Terrestrial Physics*, 100, 13-23. <https://doi.org/10.1016/j.jastp.2013.03.017>.
8. Abhilash, S., et al. (2013), "Extended range prediction of active-break spells of Indian summer monsoon rainfall using an ensemble prediction system in NCEP Climate Forecast System", *International Journal of Climatology*, 34(1). <https://doi.org/10.1002/joc.3668>.
9. Pattanaik, D.R. et al. (2020), "Active-break transitions of monsoons over India as predicted by coupled model ensembles", *Pure and Applied Geophysics*, 177(3), 4391-4422. DOI:10.1007/s00024-020-02503-2.
10. Saha, M., et al. (2020, September), *CNN-Based Forecasting of Intraseasonal Mean and Active/Break Spells for Indian Summer Monsoon*, Proceedings of the 10th International Conference on Climate Informatics, 15-21.
11. Bagtasa, G. (2020), "Influence of Madden-Julian oscillation on the intraseasonal variability of summer and winter monsoon rainfall in the Philippines", *Journal of Climate*, 33, 9581-9594. DOI: <https://doi.org/10.1175/JCLI-D-20-0305.1>.
12. Guan, W., et al. (2019), "Subseasonal zonal variability of the western Pacific subtropical high in summer: climate impacts and underlying mechanisms", *Climate Dynamics*, 53, 3325-3344. DOI:10.1007/s00382-019-04705-4.
13. Olaguera, L.M.P., et al. (2021), "A climatological analysis of the monsoon break following the summer monsoon onset over Luzon Island, Philippines", *International Journal of Climatology*, 41(3), 2100-2117. DOI:10.1002/joc.6949.
14. Olaguera, L.M.P., et al. (2022), "Influence of boreal summer intraseasonal oscillation on rainfall extremes in the Philippines", *International Journal of Climatology*, 42, 4656-4668. <https://doi.org/10.1002/joc.7495>
15. Xu, K., and Lu, R. (2015), "Break of the western North Pacific summer monsoon in early August",

*Journal of Climate*, 28(8), 3420-3434. DOI:10.1175/JCLI-D-14-00588.1.

16. Xu, K., and Lu, R. (2018), "Decadal change of the western North Pacific summer monsoon break around 2002/03", *Journal of Climate*, 31(1), 177-193. DOI: <https://doi.org/10.1175/JCLI-D-16-0739.1>.
17. Xuan Thanh Pham et al. (2010), "Onset of the summer monsoon over the southern Viet Nam and its predictability", *Theoretical and Applied Climatology*, 99, 105-113. DOI 10.1007/s00704-009-0115-z.
18. Nguyễn Minh Trường và Bùi Minh Tuấn (2013), "Thử nghiệm dự báo ngày bùng nổ gió mùa hè khu vực Nam Bộ năm 2012 bằng mô hình RAMS", *Tạp chí Khoa học Đại học quốc gia Hà Nội, Khoa học Tự nhiên và Công nghệ* 29, Số 1S (2013) 179-186.
19. Bùi Minh Tuấn (2015), "Dự báo ngày bùng nổ gió mùa hè khu vực Nam Bộ sử dụng mô hình WRF", *Tạp chí Khoa học ĐHQGHN: Khoa học Tự nhiên và Công nghệ*, Tập 31, Số 3S tr 238-246.
20. Phạm Thị Châm và cộng sự (2017), "Nghiên cứu ngày bắt đầu gió mùa hè, ngày bắt đầu mùa mưa ở Tây Nguyên và khả năng dự báo thời điểm bắt đầu mùa mưa bằng phương pháp phân tích CANON", *Tạp chí Khí tượng Thủy văn*, Tập 683 Số. 11 (2017), 1-11.
21. Nguyễn Đăng Mậu và cộng sự (2015), *Nghiên cứu dự báo hoạt động của gió mùa hè trên khu vực Việt Nam bằng mô hình động lực*, Báo cáo tổng kết đề tài cấp Bộ, Bộ Tài nguyên và Môi trường, Mã số: TNMT.05.36.
22. Hoàng Đức Cường và cộng sự (2018), *Nghiên cứu xây dựng phương pháp dự báo thời điểm bắt đầu và kết thúc mùa mưa gió mùa hè khu vực Trung Bộ, Tây Nguyên và Nam Bộ*, Báo cáo tổng hợp đề tài TNMT.2016.05.08, Bộ Tài Nguyên và Môi Trường, 142 trang.
23. Nguyễn Đức Ngữ và Nguyễn Trọng Hiệu (2004), *Khí hậu và tài nguyên khí hậu Việt Nam*, Nhà xuất bản Nông nghiệp, Hà Nội.
24. Ngo-Thanh, H. et al. (2018), "A distinction between summer rainy season and summer monsoon season over the Central Highlands of Viet Nam", *Theoretical and Applied Climatology*, 132, 1237-1246.
25. Pham-Thanh, H. et al. (2020), "Predictability of the rainy season onset date in Central Highlands of Viet Nam", *International Journal of Climatology*, 40(6), 3072-3086. <https://doi.org/10.1002/joc.6383>.
26. Chen, T.C. et al. (2004), "Variation of the East Asian summer monsoon rainfall", *Journal of Climate*, 17, 744-762. [https://doi.org/10.1175/1520-0442\(2004\)017<0744:VOTEAS>2.0.CO;2](https://doi.org/10.1175/1520-0442(2004)017<0744:VOTEAS>2.0.CO;2).
27. Olaguera, L.M.P. et al. (2021), "A climatological analysis of the monsoon break following the summer monsoon onset over Luzon Island, Philippines", *International Journal of Climatology*, 41, 2100–2117. <https://doi.org/10.1002/joc.6949>.
28. Raghavan, K. (1973), "Break-monsoon over India", *Monthly Weather Review*, 101(1), 33-43. [https://doi.org/10.1175/1520-0493\(1973\)101<0033:BOI>2.3.CO;2](https://doi.org/10.1175/1520-0493(1973)101<0033:BOI>2.3.CO;2).
29. Ramanadham, R. et al. (1973), "Break in the Indian summer monsoon", *Pure and Applied Geophysics*, 104, 635-647.
30. Rajeevan, M. et al. (2010), "Active and break spells of the Indian summer monsoon", *Journal of Earth System Science*, 119, 229-247. DOI:10.1007/s12040-010-0019-4.
31. Xu, K., & Lu, R. (2022), "Break events of the western North Pacific summer monsoon during 1979-2018", *Journal of Climate*, 35, 463-477. <https://doi.org/10.1175/JCLI-D-21-0419.1>.
32. Trần Trung Thành và cộng sự (2024), "Nghiên cứu phân tích các đặc trưng của gián đoạn mưa gió mùa hè trên khu vực Tây Nguyên", *Tạp chí Khí tượng Thủy văn*, 765(9), 15-26.

# EXPERIMENTAL FORECASTING OF THE NUMBER OF RAINFALL-BREAKING DAYS IN THE CENTRAL HIGHLANDS REGION USING STATISTICAL METHODS

Tran Trung Thanh<sup>(1)</sup>, Le Thi Dieu Thu<sup>(1)</sup>, Nguyen Thi Thanh<sup>(2)</sup>,  
Ta Huu Chinh<sup>(2)</sup>, Truong Thi Thanh Thuy<sup>(2)</sup>, Lai The Long<sup>(3)</sup>

<sup>(1)</sup>Highland region Hydro-Meteorological Center

<sup>(2)</sup>The Viet Nam Institute of Meteorology, Hydrology and Climate Change

<sup>(3)</sup>Vinh Phuc Provincial Hydrometeorological Station

Received: 25/11/2024; Accepted: 9/01/2025

**Abstract:** *The paper presents the experimental forecasting results of the number of rainfall-breaking days (NGD) in the Central Highlands region, Viet Nam, comparing the linear regression method (PTT) and the nonlinear regression method (PPT). The forecasting factor used is the observed sea surface temperature analyzed orthogonally. The results show that forecasting skill in the training period is relatively high, the 3-phase accuracy is from 60-90% for PPT, surpassing the value of 40-60% for PTT. Even though, the forecasting skill in the independent period is significantly reduced in both PTT and PPT, the 3-phase accuracy drops to just below 60%. The skill of PPT is somewhat higher than that of PTT, it is not completely superior. The results show that, in many cases, the forecasting skills at some stations are lower than with PTT (PPT) but higher than with PPT (PTT). Thus, it is essential to use both these two methods simultaneously to support each other.*

**Keywords:** *Highland, rain breaks, statistical model.*