



Original Article

Subseasonal Rainfall Forecasting Skill of S2S-ECMWF and Its Potential for Bias Correction Using ANN over the Central Highlands of Vietnam During the Transition from the Dry to the Rainy Season

Nguyen Thi Kim Chi*, Phan Van Tan

VNU University of Science, 334 Nguyen Trai, Thanh Xuan, Hanoi, Vietnam

Received 30th May 2026

Revised 15th June 2026; Accepted 25st June 2026

Abstract: This study evaluates the subseasonal rainfall forecasting skill of the S2S-ECMWF model over the Central Highlands of Vietnam during the transition from the dry to the rainy season (March–April) for the period 2003–2022. It also examines the potential of applying an Artificial Neural Network (ANN) for correcting biases in the raw forecast products. The results show that the original S2S-ECMWF forecasts capture some broad features of rainfall occurrence, particularly for light rainfall events and short lead times. However, their quantitative forecasting skill remains limited, as indicated by large errors, low or even negative correlations, and a clear decline in skill at higher rainfall thresholds. The model's ability to detect moderate and heavy rainfall events decreases substantially, and its discrimination between rainfall and non-rainfall events above the selected thresholds is very weak. After ANN-based bias correction, forecast performance improves markedly, especially during weeks 1–2. The corrected forecasts also exhibit much stronger discriminatory skill, particularly at short lead times. Although forecast skill still decreases with increasing lead time and heavy rainfall prediction remains challenging, the results demonstrate the potential of ANN-based post-processing for improving subseasonal rainfall forecasts over the Central Highlands of Vietnam.

Keywords: Sub-seasonal forecasting; ANN; ECMWF; Central Highlands.

* Corresponding author.

E-mail address: nguyenchikttv@gmail.com

<https://doi.org/10.25073/2588-1094/vnuces.5524>

Kỹ năng dự báo mưa hạn nội mùa của S2S-ECMWF và khả năng hiệu chỉnh bằng ANN cho khu vực Tây Nguyên trong giai đoạn chuyển tiếp từ mùa khô sang mùa mưa

Nguyễn Thị Kim Chi*, Phan Văn Tân

*Trường Đại học Khoa học Tự nhiên, Đại học Quốc gia Hà Nội,
334 Nguyễn Trãi, Thanh Xuân, Hà Nội, Việt Nam*

Nhận ngày 30 tháng 5 năm 2026

Chỉnh sửa ngày 15 tháng 6 năm 2026; Chấp nhận đăng ngày 25 tháng 6 năm 2026

Tóm tắt: Nghiên cứu này đánh giá kỹ năng dự báo mưa hạn nội mùa của mô hình S2S-ECMWF trên khu vực Tây Nguyên, Việt Nam, trong giai đoạn chuyển tiếp từ mùa khô sang mùa mưa (tháng 3–4) thời kỳ 2003–2022. Nghiên cứu cũng khảo sát khả năng ứng dụng mạng thần kinh nhân tạo (ANN) để hiệu chỉnh sai số của sản phẩm dự báo gốc. Kết quả cho thấy dự báo gốc của S2S-ECMWF nắm bắt được một số đặc điểm cơ bản của sự xuất hiện mưa, đặc biệt đối với các sự kiện mưa nhỏ và các hạn dự báo ngắn. Tuy nhiên, kỹ năng dự báo định lượng mưa vẫn còn hạn chế, thể hiện qua sai số lớn, hệ số tương quan thấp hoặc thậm chí âm, và kỹ năng dự báo giảm rõ rệt khi ngưỡng mưa tăng lên. Khả năng phát hiện các sự kiện mưa vừa và mưa lớn của mô hình giảm đáng kể, đồng thời khả năng phân biệt giữa các trường hợp có và không có mưa vượt ngưỡng rất kém. Sau khi hiệu chỉnh bằng ANN, chất lượng dự báo được cải thiện rõ rệt, đặc biệt trong các tuần 1–2. Dự báo sau hiệu chỉnh cũng thể hiện kỹ năng phân biệt các sự kiện có và không có mưa vượt ngưỡng tốt hơn nhiều, nhất là ở các hạn dự báo ngắn. Mặc dù kỹ năng dự báo vẫn giảm khi hạn dự báo tăng và dự báo mưa lớn vẫn còn là một thách thức, kết quả nghiên cứu cho thấy có thể ứng dụng ANN trong hậu xử lý để cải thiện dự báo mưa hạn nội mùa cho khu vực Tây Nguyên, Việt Nam.

Từ khóa: Dự báo nội mùa; ANN; ECMWF; Tây Nguyên.

1. Mở đầu

Dự báo mưa đóng vai trò quan trọng đối với nhiều lĩnh vực kinh tế, xã hội, đặc biệt là sản xuất nông nghiệp và quản lý tài nguyên nước. Trong đó, dự báo hạn nội mùa (khoảng từ 2 tuần đến 2 tháng), đang ngày càng được quan tâm do có khả năng lấp đầy khoảng trống giữa dự báo thời tiết (vài ba ngày đến một tuần) và dự báo hạn mùa (từ 1 tháng đến 12 tháng) [1]. Ở Việt Nam, Tây Nguyên là một trong những vùng kinh tế quan trọng, đặc biệt đối với sản xuất các loại cây công nghiệp chủ lực, tạo ra các sản phẩm xuất khẩu có

giá trị kinh tế cao, như cà phê, hồ tiêu và cao su. Các tháng chuyển tiếp từ mùa khô sang mùa mưa (tháng 3–4) là giai đoạn quan trọng, quyết định sự sinh trưởng và phát triển của cây trồng, lập kế hoạch tưới tiêu cũng như quản lý rủi ro do các hiện tượng thời tiết cực đoan [2]. Do đó, việc dự báo chính xác diễn biến mưa, bao gồm sự xuất hiện mưa, lượng mưa, trong giai đoạn này có ý nghĩa quan trọng đối với cả sản xuất nông nghiệp cũng như quản lý tài nguyên nước.

Ở Việt Nam hiện nay vẫn chưa có hệ thống dự báo hạn nội mùa nào được đưa vào vận hành

* Tác giả liên hệ.

Địa chỉ email: nguyenchikittv@gmail.com

<https://doi.org/10.25073/2588-1094/vnuees.5524>

ng nghiệp vụ, trong khi trên thế giới đã có nhiều hệ thống được phát triển và ứng dụng rộng rãi. Vì vậy, việc khảo sát, đánh giá khả năng ứng dụng sản phẩm dự báo sẵn có của các Trung tâm trên thế giới cho Việt Nam là vấn đề cần thiết. Trong số các nguồn sẵn có đó, sản phẩm dự báo nội mùa của Trung tâm Dự báo thời tiết hạn vừa Châu Âu (S2S-ECMWF) được xem là một trong những nguồn tương đối tốt [1–6] và có thể khai thác miễn phí. Tuy nhiên, chất lượng dự báo của mô hình thường phụ thuộc vào khu vực, mùa và hạn dự báo. Tây Nguyên là nơi có địa hình phức tạp, chịu ảnh hưởng của các hệ thống gió mùa, việc đánh giá khả năng dự báo mưa hạn nội mùa của mô hình S2S-ECMWF cho khu vực này trước khi ứng dụng vào nghiệp vụ là hết sức cần thiết.

Dự báo nội mùa hiện đang là một trong những hướng nghiên cứu thu hút nhiều sự quan tâm của cả cộng đồng các nhà khoa học cũng như người dùng. Cho đến nay, đã có nhiều công trình tập trung đánh giá kỹ năng dự báo của các hệ thống khác nhau cho các khu vực khác nhau trên thế giới. Chẳng hạn, các nghiên cứu của Olaniyan và cs (2018) [3] cho Nigeria và De Andrade và cs (2021) [4] cho Châu Phi nói chung cho thấy sản phẩm của S2S-ECMWF có kỹ năng dự báo tốt hơn so với một số hệ thống khác như UKMO (United Kingdom Meteorological Office) hoặc NCEP (National Centers for Environmental Prediction). Tuy nhiên, De Andrade và cs (2021) [4], Vitart (2013) [5] và Liu và cs (2013) [7] lại chỉ ra rằng S2S-ECMWF vẫn có những hạn chế nhất định, đặc biệt trong các giai đoạn chuyển mùa, kỹ năng dự báo có xu hướng giảm theo hạn dự báo.

Ngoài việc đánh giá kỹ năng của mô hình, một số nghiên cứu đã áp dụng các phương pháp học máy, học sâu như Random Forest, XGBoost,... để hiệu chỉnh sai số mô hình, nhằm cải thiện chất lượng dự báo [6–8]. Worsnop và cs (2024) [9] đã phát triển mô hình RUFCO (kết hợp giữa các thành phần ResU-Net, FiLM và Climatological-Offset) để hiệu chỉnh sản phẩm dự báo mưa của hệ thống dự báo tổ hợp toàn cầu NOAA (GEFSv12) cho Hoa Kỳ. Kết quả chỉ ra rằng mô hình học sâu có khả năng cải thiện kỹ năng dự báo tốt hơn so với các phương

pháp hiệu chỉnh thống kê truyền thống. Gần đây, bằng việc sử dụng kiến trúc ResU-Net để hiệu chỉnh sản phẩm dự báo nhiệt độ và lượng mưa từ hệ thống S2S-ECMWF, Mengist và Seo (2026) [10] cho thấy chất lượng dự báo trong vòng 6 tuần tăng lên đáng kể.

Ở Việt Nam, đã có một số công trình đánh giá khả năng dự báo mưa hạn nội mùa của S2S-ECMWF trên phạm vi cả nước hoặc tại một vùng khí hậu cụ thể. Kết quả cho thấy S2S-ECMWF có khả năng dự báo tương đối tốt lượng mưa (Đặng Hải Yến, 2023) [11] và ngày bắt đầu mùa mưa (Phạm Thanh Hà, 2022) [2] ở các hạn dự báo ngắn, đặc biệt trong các tuần thứ nhất và thứ hai. Tuy nhiên, kỹ năng dự báo vẫn còn hạn chế tại một số khu vực, trong đó có Tây Nguyên và Nam Bộ. Hoàng Thị Thu Hương và cs (2023) [1] cũng đã thử nghiệm các phương pháp hiệu chỉnh sai số dự báo mưa của S2S-ECMWF cho khu vực Bắc Trung Bộ và cho thấy hiệu quả nhất định trong việc giảm sai số và cải thiện chất lượng dự báo.

Mặc dù đã có nhiều nghiên cứu về đánh giá kỹ năng dự báo hạn nội mùa và hiệu chỉnh sản phẩm dự báo của các mô hình nói chung, S2S-ECMWF nói riêng, bằng các phương pháp học máy, học sâu ở trên thế giới cũng như tại Việt Nam, nhưng chưa có nghiên cứu nào tập trung riêng cho khu vực Tây Nguyên trong giai đoạn chuyển tiếp từ mùa khô sang mùa mưa (tháng 3, 4). Xuất phát từ đó, nghiên cứu này nhằm trả lời hai câu hỏi: i) Có thể sử dụng trực tiếp sản phẩm gốc của S2S-ECMWF để dự báo mưa hạn nội mùa cho khu vực Tây Nguyên trong các tháng chuyển từ mùa khô sang mùa mưa (tháng 3–4) hay không?; và ii) Có thể cải thiện chất lượng dự báo mưa hạn nội mùa cho khu vực Tây Nguyên trong giai đoạn này bằng việc ứng dụng mô hình mạng thần kinh nhân tạo (ANN) để hiệu chỉnh sản phẩm của S2S-ECMWF hay không?.

2. Số liệu và phương pháp

2.1. Số liệu

Nghiên cứu sử dụng hai nguồn số liệu, bao gồm số liệu quan trắc mưa ngày trên lưới và số

liệu dự báo lại của mô hình S2S-ECMWF giai đoạn 2003–2022.

Số liệu quan trắc lượng mưa ngày trên lưới giai đoạn 2003–2022 được xây dựng dựa trên số liệu quan trắc từ mạng lưới trạm khí tượng và trạm đo mưa trên toàn lãnh thổ Việt Nam trong khuôn khổ dự án “Cập nhật kịch bản biến đổi khí hậu cho Việt Nam” của Viện Khoa học Khí tượng Thủy văn và Biến đổi khí hậu (nay là Viện Khoa học Khí tượng Thủy văn, Môi trường và Biển) thực hiện (2023–2025) [12–14], và đã được Le và Phan (2025) [15] sử dụng trong công trình công bố gần đây. Độ phân giải không gian của số liệu là $0,25 \times 0,25$ độ kinh vĩ. Miền nghiên cứu nằm trong phạm vi $10^{\circ}\text{N} - 17^{\circ}\text{N}$ và $105^{\circ}\text{E} - 110^{\circ}\text{E}$, bao phủ toàn bộ khu vực Tây Nguyên.

Số liệu dự báo của S2S-ECMWF được sử dụng là sản phẩm dự báo lại (reforecast) phiên bản 2023 cho giai đoạn 2003–2022 với hạn dự báo đến 45 ngày, trừ ngày đứng làm dự báo (ngày thứ 1, hạn dự báo bằng 0). Nếu tính cả ngày đứng làm dự báo là 46 ngày (1–46). Vì mỗi tuần chỉ có 2 phiên dự báo nên mỗi năm sẽ có 17 phiên dự báo trong các tháng 3 và 4. Để phù hợp với số liệu quan trắc, độ phân giải không gian của S2S-ECMWF được chọn là $0,25^{\circ} \times 0,25^{\circ}$. Các biến dự báo của mô hình bao gồm hai loại chính là số liệu ngày (daily data) và số liệu tức thời tại các lát cắt thời gian 00, 06, 12 và 18 UTC. Đối với lượng mưa, số liệu được cho là tổng lượng mưa tích lũy từ thời điểm làm dự báo đến ngày cuối cùng của hạn dự báo, với các khoảng thời gian từng 6 giờ một. Ngoài lượng mưa ngày, nghiên cứu còn sử dụng nhiều biến khác, mà chúng sẽ được dùng làm đầu vào cho mô hình ANN. Các biến này hoặc sẵn có từ số liệu gốc, hoặc được tính ra từ các biến sẵn có, bao gồm các dòng hiền nhiệt, ẩn nhiệt từ bề mặt, bức xạ sóng ngắn, sóng dài tại bề mặt, nhiệt độ và độ ẩm các lớp đất gần bề mặt, độ dày các lớp khí quyển giữa các mực 1000–850 hPa, 975–700 hPa, 950–500 hPa, 1000–500 hPa và 850–700 hPa, chênh lệch nhiệt độ giữa các lớp khí quyển tầng thấp (850–500 hPa và 700–500 hPa, đặc trưng cho gradient thẳng đứng của nhiệt độ), năng lượng đối lưu tiềm năng (CAPE), tổng

lượng nước của cột khí quyển, thông lượng ẩm, các thành phần gió ngang,...

2.2. Phương pháp

Trong nghiên cứu này, hai nội dung chính sẽ được thực hiện: 1) Đánh giá kỹ năng của S2S-ECMWF trong dự báo mưa hạn nội mùa các tháng 3–4 và 2) Đánh giá khả năng ứng dụng ANN để hiệu chỉnh sản phẩm dự báo mưa hạn nội mùa của S2S-ECMWF trong các tháng 3–4 cho khu vực Tây Nguyên. Với hạn dự báo của S2S-ECMWF là 0–45 ngày, ở đây chỉ thực hiện việc đánh giá cho các hạn từ ngày thứ 4 đến ngày thứ 45, và được chia theo tuần: tuần 1 (W1): từ ngày thứ 4 đến ngày thứ 10, tuần 2 (W2) từ ngày thứ 11 đến ngày thứ 17, ..., tuần 6 (W6) từ ngày thứ 39 đến ngày thứ 45. Cũng cần lưu ý rằng, thông thường các “tuần” trong dự báo nội mùa được tính từ ngày thứ nhất sau ngày đứng làm dự báo, tức là tuần 1 từ ngày 1 đến ngày 7, ... tuần 6 từ ngày 36 đến ngày 42. Với hạn dự báo 45 ngày, cách làm đó sẽ bỏ mất 3 ngày cuối cùng (43–45). Do đó, thay vì bỏ 3 ngày cuối cùng, ở đây sẽ bỏ 3 ngày đầu tiên, thuộc dự báo thời tiết hạn ngắn.

2.2.1. Ứng dụng ANN trong hiệu chỉnh sản phẩm mưa dự báo

Mô hình ANN [16] sử dụng trong nghiên cứu này được thiết kế với nhiều phương án số lượng biến đầu vào, số lớp ẩn và số nơ thần của mỗi lớp khác nhau nhằm tạo ra nhiều sản phẩm hiệu chỉnh mà chúng được dùng như là các thành phần tổ hợp. Số lượng biến đầu vào biến thiên từ ít nhất 8 biến đến nhiều nhất 26 biến, được lựa chọn dựa trên cơ sở ý nghĩa vật lý của các nhóm biến đối với biến dự báo là lượng mưa. Số lượng 26 biến này được xác định từ một tập trên 60 biến ban đầu, bao gồm các biến gốc (original variables) của S2S-ECMWF, các biến phái sinh (derived variables), các hàm trực giao thực nghiệm từ một số trường khí quyển, trên cơ sở đánh giá mức độ quan trọng của các biến dựa trên hoán vị (permutation feature importance).

Trong quá trình xây dựng mô hình, tập mẫu 20 năm được chia một cách ngẫu nhiên thành 3 tập con không giao nhau, trong đó 15 năm làm

tập huấn luyện (training), 3 trong 5 năm còn lại làm tập kiểm định (validating) và 2 năm còn lại cuối cùng làm tập đánh giá độc lập (testing). Kiến trúc của mô hình ANN, bao gồm số lớp ẩn, số nơ-đon trong mỗi lớp, tốc độ học (learning rate), hệ số quán tính (momentum), hàm kích hoạt (Activation function),... được thực hiện cho một số điểm lưới ngẫu nhiên trong vùng nghiên cứu với bộ các biến đầu vào khác nhau. Trên cơ sở đó, các bộ tham số điều khiển mô hình (hyperparameters – siêu tham số) sẽ được lựa chọn. Trong đó, một số tham số hầu như không ảnh hưởng đến kết quả nên sẽ được đặt cố định cho tất cả các cấu hình. Cụ thể, $\text{learning_rate} = 0,001$, $\text{activation_hidden} = \text{"leaky_relu"}$, $\text{momentum} = 0,85$, $\text{loss_type} = \text{"MAE"}$. Các tham số ảnh hưởng rõ nhất là số lớp ẩn và số nơ-đon trong mỗi lớp. Số lớp ẩn được chọn tối thiểu là 1 lớp và tối đa là 3 lớp (Bảng 1). Số nơ-đon cũng biến thiên trong khoảng từ 8–52 cho lớp ẩn thứ nhất, 4–30 cho lớp ẩn thứ hai, 8–30 cho lớp ẩn thứ ba (Bảng 1). Tổng cộng có tất cả 38 trường hợp được thử nghiệm. Vì vậy việc đánh giá kết quả hiệu chỉnh, sẽ được thực hiện dựa trên sản phẩm tổ hợp từ 38 thành phần, được tính lại cho toàn bộ 20 năm số liệu, và sẽ bỏ qua quá trình kiểm định và đánh giá độc lập cho toàn bộ khu vực nghiên cứu.

Bảng 1. Cấu hình số lớp ẩn và số nơ-đon của mỗi lớp cho mô hình ANN

STT	Số lớp	Số nơ-đon/lớp		
		1	2	3
1	1	8		
2	1	16		
3	2	12	6	
4	2	16	8	
5	2	16	8	
6	2	18	9	
7	2	20	10	
8	2	20	10	
9	2	22	11	
10	2	28	14	
11	2	30	15	
12	2	32	16	
13	2	32	16	
14	3	8	16	8
15	3	16	4	16

16	3	18	12	6
17	3	20	14	8
18	3	20	8	20
19	3	24	16	8
20	3	30	20	10
21	3	30	30	30
22	3	40	20	10
23	3	52	26	13

2.2.2. Đánh giá dự báo định lượng

Để đánh giá kỹ năng dự báo định lượng mưa của S2S-ECMWF, cả trước và sau khi hiệu chỉnh, các chỉ số ME (sai số trung bình), RMAE (sai số tuyệt đối trung bình tương đối), và CORR (hệ số tương quan) [17] sẽ được sử dụng. ME biến thiên trong khoảng $(-\infty, +\infty)$ cho biết xu hướng thiên lệch của dự báo so với quan trắc nhưng không phản ánh độ lớn của sai số, với giá trị lý tưởng là $ME = 0$ (không thiên lệch về phía nào). $RMAE (\geq 0)$ phản ánh sai số trung bình của dự báo so với quan trắc được tính theo % so với trung bình quan trắc, nhưng không nói lên xu hướng thiên lệch của dự báo so với quan trắc. Giá trị lý tưởng là $RMAE = 0$. Ưu điểm của RMAE là có thể so sánh sai số của mô hình trong các điều kiện mưa ít và mưa nhiều. Ngoài ra, để đánh giá hiệu quả của việc hiệu chỉnh so với sản phẩm dự báo của mô hình trước khi hiệu chỉnh, chỉ số giá trị thặng dư (Added Value - AV), tính theo công thức (1), sẽ được sử dụng. Trong (1), $RMAE_{mod}$ và $RMAE_{hc}$ tương ứng là lượng mưa dự báo của mô hình trước và sau hiệu chỉnh. AV biến thiên trong khoảng $[-1, 0; 1, 0]$ với $AV > 0$ phản ánh sai số của lượng mưa sau hiệu chỉnh nhỏ hơn trước khi hiệu chỉnh. Giá trị lý tưởng là $AV = 1, 0$.

$$AV(RMAE) = \frac{RMAE_{mod} - RMAE_{hc}}{\max\{RMAE_{mod}, RMAE_{hc}\}} \quad (1)$$

2.2.3. Đánh giá dự báo pha

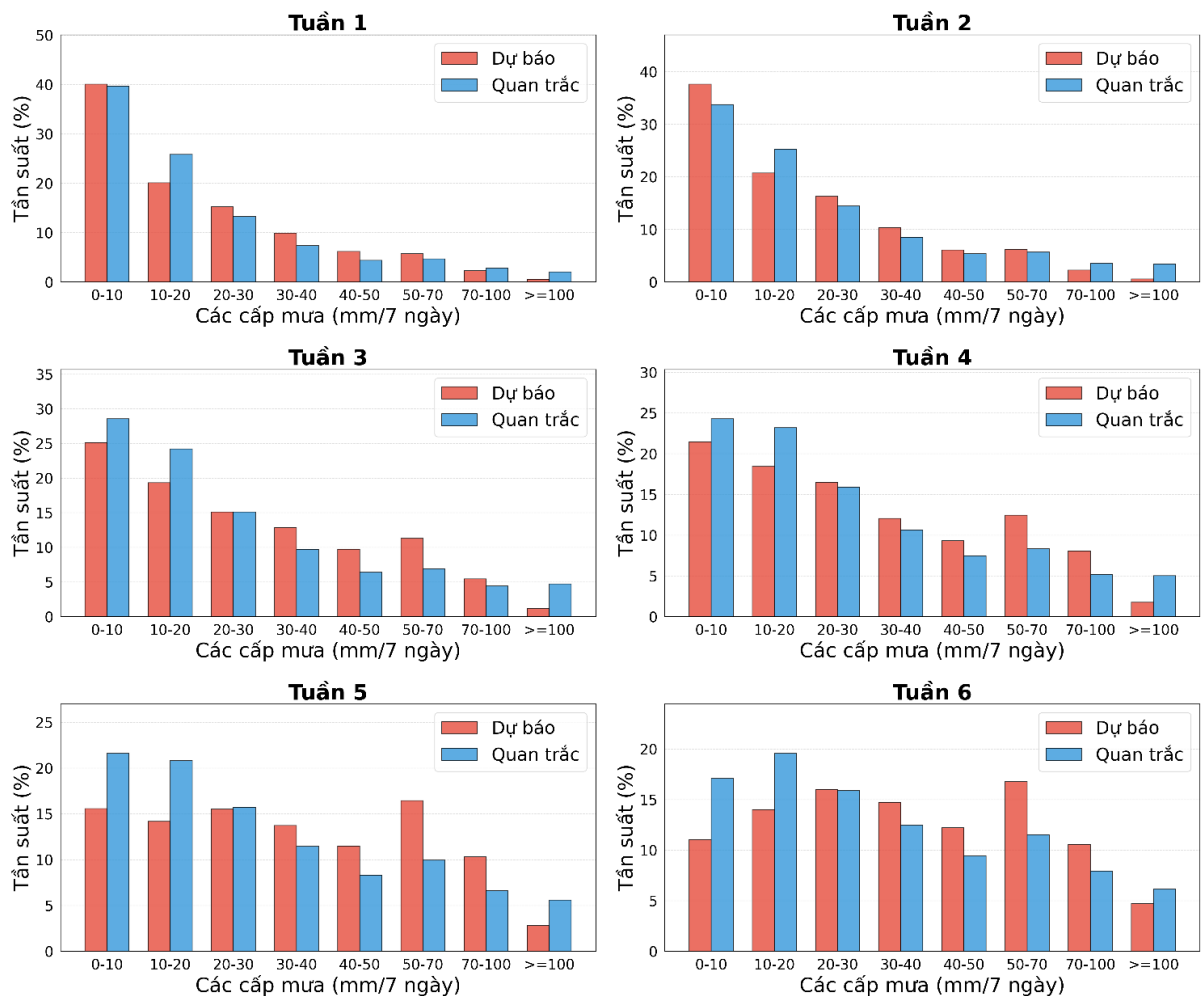
Trong đánh giá dự báo pha, việc đánh giá thường được quy về bài toán nhị phân, trong đó lượng mưa được phân chia theo các ngưỡng để xác định sự kiện *lượng mưa vượt ngưỡng đã cho* có xảy ra (1) hay không (0). Khi đó, nếu ký hiệu: i) số lần dự báo đúng sự kiện có xảy ra là H; ii) số lần dự báo sự kiện có xảy ra nhưng thực tế không xảy ra (dự báo không) là F; iii) số lần không dự báo được sự kiện có xảy ra (dự báo sai

hay dự báo sót) là M; và iv) số lần dự báo đúng hiện tượng không xảy ra là CN, thì $H+F+M+CN = N$ là số lần dự báo. Trong nghiên cứu này, các chỉ số POD (xác suất dự báo đúng sự kiện có xảy ra), POFD (xác suất dự báo sai sự kiện không xảy ra), FAR (tỷ lệ dự báo không), TS hay CSI (chỉ số thành công tối hạn), PC (tỷ lệ dự báo đúng), và BIAS (chỉ số thiên lệch) sẽ được sử dụng [15]. Việc sử dụng nhiều chỉ số đánh giá nhằm xem xét một cách đầy đủ kỹ năng dự báo của mô hình.

3. Kết quả và thảo luận

3.1. Kỹ năng dự báo mưa của S2S-ECMWF

Hình 1 trình bày phân bố tần suất lượng mưa theo các cấp của dự báo S2S-ECMWF và quan trắc cho các hạn dự báo từ W1 đến W6. Nhìn chung, ở các W1-W3, cả dự báo và quan trắc, tần suất mưa giảm dần khi cấp mưa tăng lên. Các cấp mưa nhỏ, đặc biệt là 0–10 mm/7 ngày và 10–20 mm/7 ngày, chiếm tần suất lớn nhất, trong khi các cấp mưa lớn hơn có tần suất thấp hơn.



Hình 1. Tần suất xuất hiện lượng mưa các cấp của S2S-ECMWF và quan trắc trên khu vực Tây Nguyên theo hạn dự báo với các phiên dự báo trong tháng 3 và 4 giai đoạn 2003–2022.

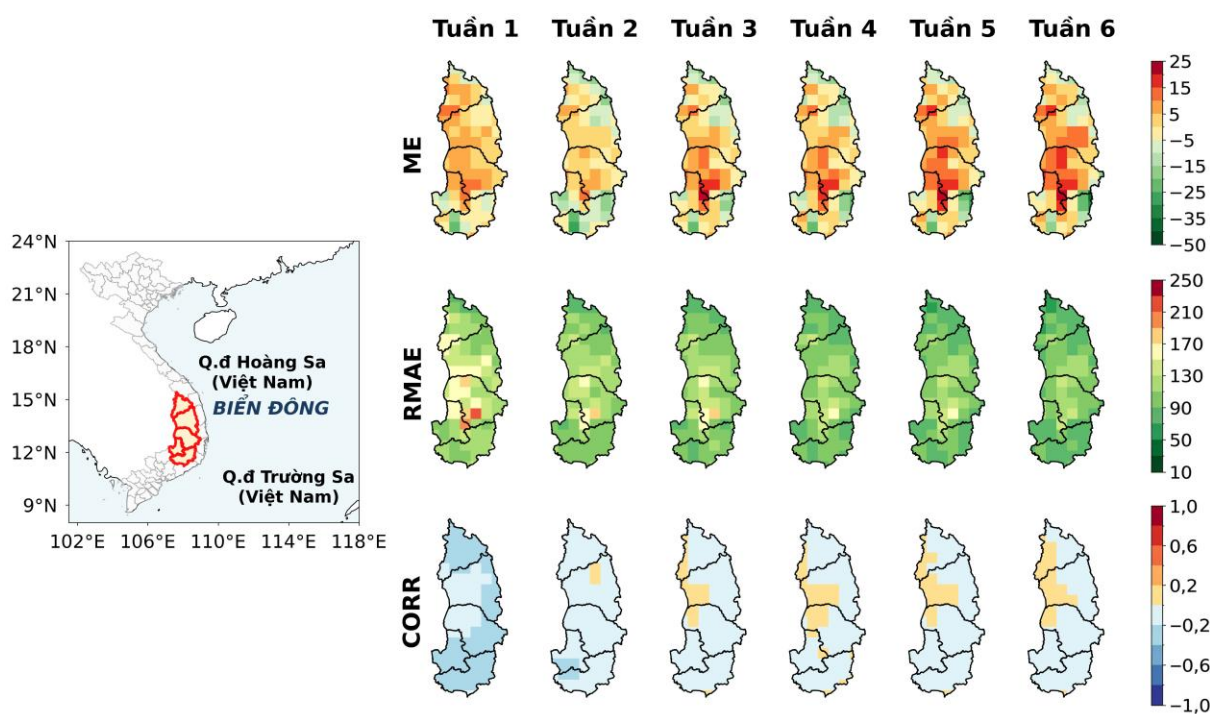
Điều này phản ánh đặc điểm phổ biến của phân bố mưa trong giai đoạn chuyển mùa, khi

các ngày mưa nhỏ và mưa vừa vẫn chiếm ưu thế. Ở W1, mô hình dự báo khá tốt tần suất cấp mưa

0–10 mm/7 ngày, nhưng có xu hướng thấp hơn quan trắc ở cấp 10–20 mm/7 ngày. Ngược lại, tại các cấp mưa vừa, như 20–30, 30–40 và 40–50 mm/7 ngày, tần suất dự báo cao hơn quan trắc. Sang W2 và W3, dạng phân bố giảm dần theo cấp mưa vẫn được thể hiện; tuy nhiên chênh lệch giữa dự báo và quan trắc tăng lên. Nhìn chung, mô hình dự báo thấp hơn quan trắc ở các cấp mưa nhỏ đến mưa vừa, trong khi lại có xu hướng dự báo cao hơn ở một số cấp mưa vừa đến mưa lớn. Từ W4–W6, phân bố tần suất mưa có sự thay đổi rõ rệt hơn. Tần suất quan trắc ở các cấp 10–20 mm/7 ngày và 20–30 mm/7 ngày tăng lên, cho thấy thời điểm dự báo có thể đã rơi vào giai đoạn đầu mùa mưa ở Tây Nguyên. Đây là giai đoạn hoạt động đối lưu và các nhiễu động gây mưa bắt đầu gia tăng, làm cho phân bố mưa dịch chuyển từ các cấp mưa rất nhỏ sang các cấp mưa lớn hơn. Tuy nhiên, mô hình chưa nắm bắt được những

điều kiện này. Trong các W4–W6, dự báo thường thấp hơn quan trắc ở một số cấp mưa nhỏ và mưa vừa, nhưng lại cao hơn ở các cấp mưa 50–70 mm/7 ngày hoặc 70–100 mm/7 ngày.

Đáng chú ý, ở hầu hết các hạn dự báo, tần suất mưa rất lớn, trên 100 mm/7 ngày, mô hình thường thấp hơn quan trắc. Điều đó chứng tỏ S2S-ECMWF còn hạn chế trong việc dự báo các sự kiện mưa lớn, đặc biệt ở các hạn dự báo dài. Nguyên nhân có thể liên quan đến khả năng của mô hình trong việc nắm bắt các quá trình đối lưu quy mô nhỏ, tác động địa hình và các nhiễu động gió mùa trong giai đoạn chuyển mùa. Nhìn chung, mô hình có khả năng nắm bắt tương đối tốt hình thái phân bố mưa, nhất là ở các tuần đầu, nhưng vẫn chưa đáp ứng được yêu cầu, đặc biệt đối với các cấp mưa lớn và các hạn dự báo từ W4 trở đi.



Hình 2. Phân bố không gian của sai số trung bình (ME), sai số tuyệt đối trung bình tương đối (RMAE) và hệ số tương quan (CORR) giai đoạn 2003–2022 của S2S-ECMWF so với quan trắc trên khu vực Tây Nguyên theo các hạn dự báo.

Hình 2 trình bày phân bố không gian của các chỉ số ME, RMAE và CORR giữa lượng mưa dự báo S2S-ECMWF và quan trắc trên khu vực Tây

Nguyên từ W1 đến W6. Có thể thấy kỹ năng dự báo của mô hình còn hạn chế, sai số tương đối

lớn và hệ số tương quan thấp ở hầu hết các hạn dự báo.

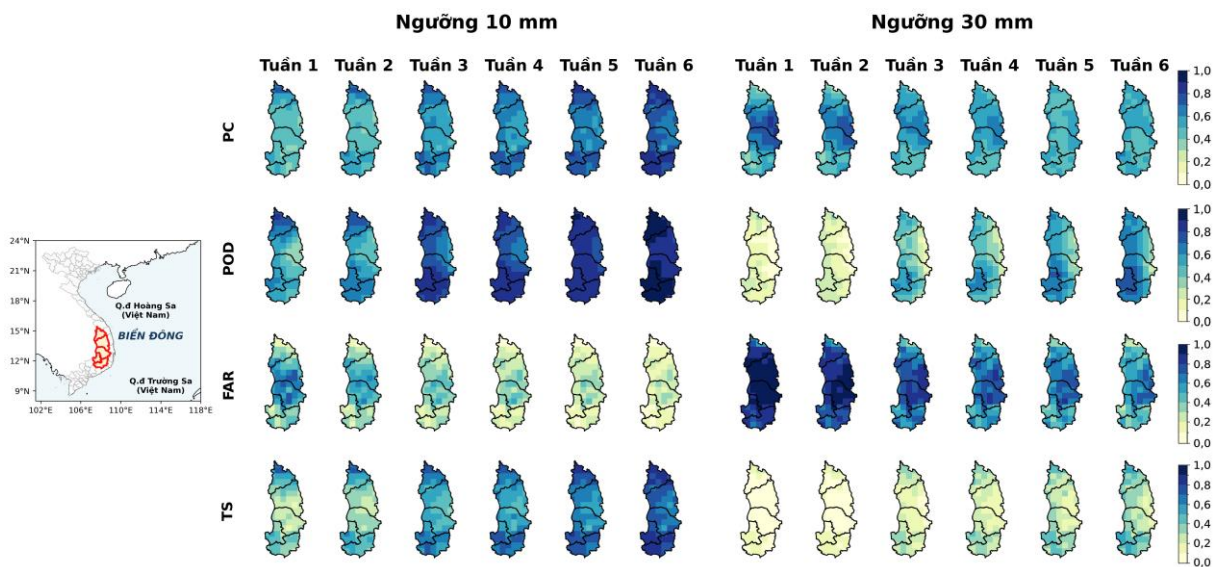
Mô hình có xu hướng thiên dương ($ME > 0$) trên phần lớn khu vực, rõ nhất ở khu vực trung tâm và phía Bắc Tây Nguyên. Xu thế này tăng lên ở các hạn dự báo dài hơn, nhất là từ W4 đến W6, chứng tỏ xu hướng dự báo cao hơn quan trắc trong khoảng thời gian trước và đầu mùa mưa. Tuy nhiên, ở một số khu vực phía Nam Tây Nguyên, ME có giá trị nhỏ hơn hoặc chuyển sang âm cục bộ, phản ánh sự khác biệt không gian của sai số trung bình.

Giá trị của RMAE nhìn chung khá cao trên toàn khu vực, phổ biến trong khoảng 80–110%, thậm chí trên 150%, cho thấy sai số của mô hình còn lớn. RMAE cao hơn ở các tuần đầu, đặc biệt W1 và W2, sau đó có xu hướng giảm nhẹ ở các hạn dự báo dài hơn. Điều này có thể liên quan đến đặc điểm mưa trong giai đoạn chuyển mùa, khi lượng mưa quan trắc còn nhỏ và biến động mạnh trong các tuần rơi vào mùa khô. Mặc dù RMAE giảm dần từ W3 đến W6, giá trị sai số

vẫn ở mức đáng kể, cho thấy mô hình chưa dự báo tốt lượng mưa trên khu vực Tây Nguyên.

Giá trị của hệ số tương quan CORR nhìn chung khá thấp, dao động xung quanh 0. Ở W1, nhiều khu vực có CORR âm, tức là mô hình chưa nắm bắt được biến động thời gian của lượng mưa quan trắc. Điều này có thể do thời điểm được dự báo chủ yếu rơi vào các tháng 3–4 hoặc tuần đầu của tháng 5, khi mà mùa mưa nói chung chưa bắt đầu, và do đó số sự kiện mưa (nếu có) là khá ít hoặc mưa có tính cục bộ, và biến động mạnh về lượng, dẫn đến mô hình không dự báo được. Từ W2 đến W6, tương quan có xu hướng tăng nhẹ ở một số khu vực, đặc biệt tại phần phía Bắc và trung tâm Tây Nguyên, nhưng vẫn chưa đạt mức đủ cao. Nghĩa là khả năng dự báo định lượng mưa của S2S-ECMWF còn hạn chế.

Hình 3 trình bày phân bố không gian các chỉ số đánh giá PC, POD, FAR và TS tại các ngưỡng 10 mm và 30 mm trên khu vực Tây Nguyên trong các hạn dự báo từ W1 đến W6, trong đó quan trắc được sử dụng làm số liệu tham chiếu.



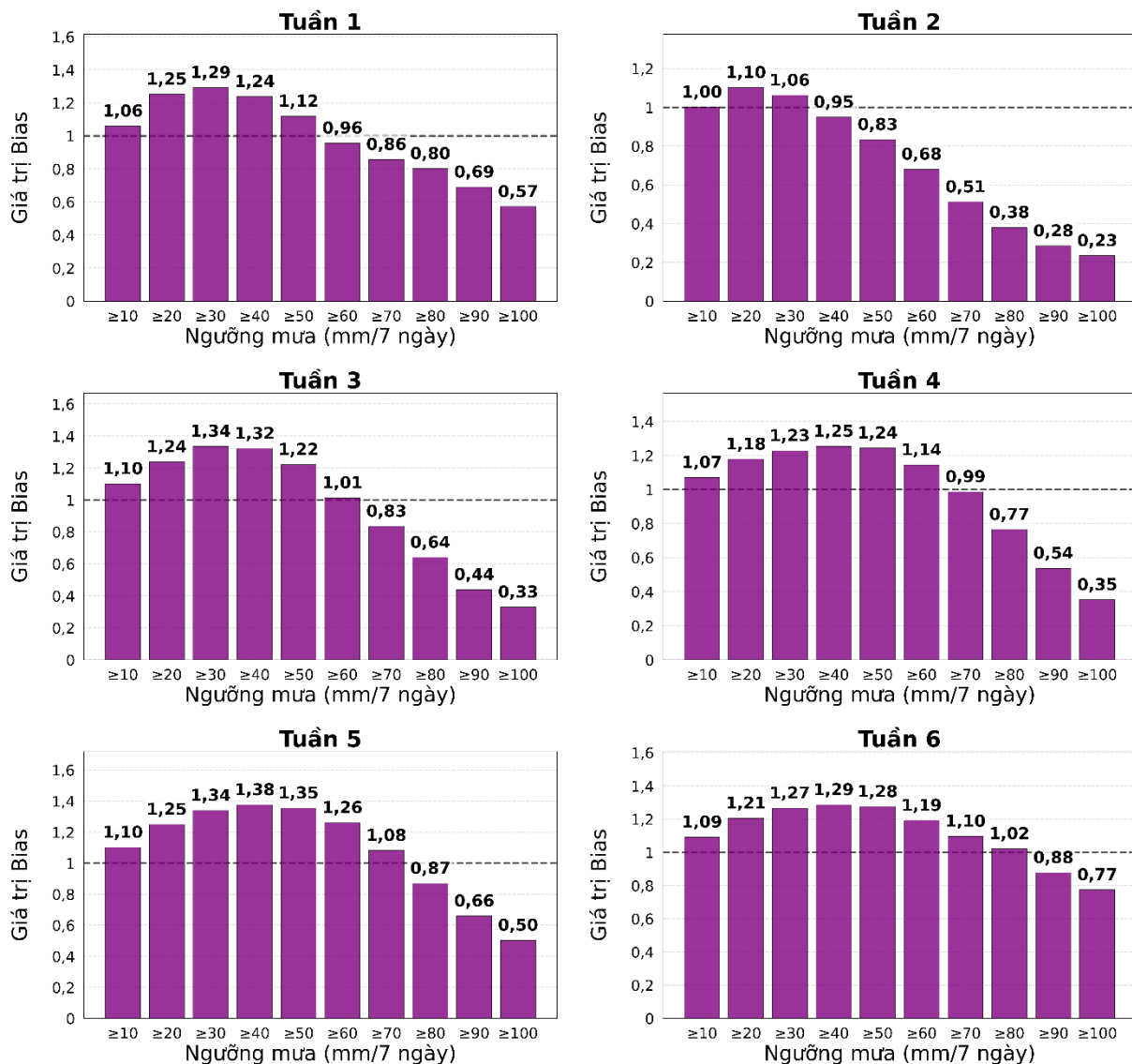
Hình 3. Phân bố không gian của PC, POD, FAR, TS ở ngưỡng mưa 10 mm (bên trái) và 30 mm (bên phải) giữa lượng mưa của S2S-ECMWF so với quan trắc giai đoạn 2003–2022 trên khu vực Tây Nguyên.

Ở ngưỡng 10 mm, mô hình có kỹ năng dự báo tương đối tốt trên phần lớn các khu vực Tây Nguyên. Cụ thể, giá trị của POD khá cao, đạt trên 0,7 (W5–W6), trong khi FAR có xu hướng giảm

dần theo hạn dự báo, từ khoảng 0,19–0,61 (W1–W2) xuống còn 0,08–0,37 (W5–W6). Chỉ số TS tăng từ khoảng 0,25–0,63 lên 0,56–0,84. Khu vực phía Nam Tây Nguyên, kỹ năng dự báo cao

hơn so với khu vực phía Bắc và trung tâm. Với ngưỡng 30 mm, kỹ năng dự báo giảm rõ rệt: POD giảm mạnh, đặc biệt trong các W1–W3, phản ánh khả năng phát hiện sự kiện mưa còn hạn chế. Chỉ số FAR có xu hướng tăng và TS chỉ dao động khoảng 0,02–0,49. Tình huống tương

tự cũng xuất hiện ở các ngưỡng 20 mm và 40 mm (không trình bày ở đây). Nói chung, S2S-ECMWF có khả năng dự báo tương đối tốt các sự kiện mưa nhỏ, nhưng với các sự kiện mưa vừa đến mưa lớn còn kém.



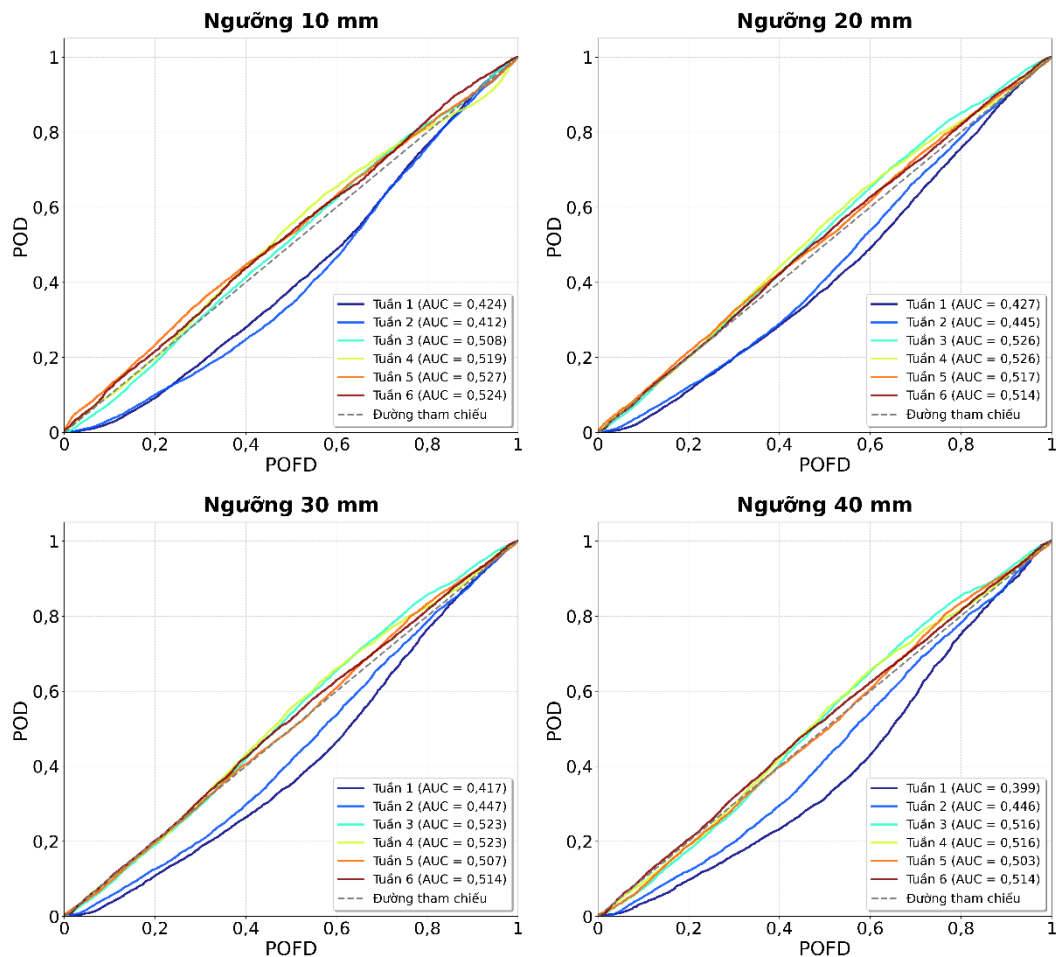
Hình 4. Chỉ số BIAS của S2S-ECMWF so với quan trắc giai đoạn 2003–2022.

Hình 4 trình bày chỉ số BIAS tính cho các ngưỡng mưa khác nhau và cho các hạn dự báo từ W1 đến W6, phản ánh mức độ sai khác giữa số sự kiện dự báo và số sự kiện quan trắc. Qua đó nhận thấy ở các ngưỡng mưa nhỏ đến vừa, S2S-

ECMWF có xu hướng dự báo vượt quá quan trắc ($\text{BIAS} > 1$), còn ở các ngưỡng mưa lớn lại cho dự báo ít hơn ($\text{BIAS} < 1$). Ở các ngưỡng từ 10–50 mm, $\text{BIAS} > 1$ trong hầu hết các hạn dự báo, nghĩa là mô hình có xu hướng dự báo số sự kiện

mưa nhỏ, mưa vừa nhiều hơn quan trắc. Khi ngưỡng mưa tăng lên, đặc biệt từ 70 mm trở lên, BIAS giảm rõ rệt và thường nhỏ hơn 1. Mặc dù ở ngưỡng 70 mm, BIAS xấp xỉ 1,02 tại W6, các ngưỡng cao hơn vẫn có xu hướng thiên thấp, với BIAS khoảng 0,88 ở ngưỡng 90 mm và 0,77 ở

ngưỡng 100 mm, chứng tỏ khá nhiều sự kiện mưa lớn xảy ra nhưng mô hình không dự báo được. Như vậy, S2S-ECMWF chỉ có khả năng nắm bắt được các sự kiện mưa nhỏ đến trung bình và thường bỏ sót các sự kiện mưa lớn.



Hình 5. Đường cong ROC (Receiver Operating Characteristic curve – tạm dịch là đường cong tiếp nhận) và chỉ số AUC (Area Under the Curve – diện tích dưới đường cong ROC) của dự báo S2S-ECMWF trên khu vực Tây Nguyên giai đoạn 2003–2022.

Hình 5 trình bày đường cong ROC và chỉ số AUC, được xây dựng từ tập mẫu tất cả các điểm lưới trên khu vực và các hạn dự báo cho các ngưỡng lượng mưa 10, 20, 30 và 40 mm trên khu vực Tây Nguyên, trong đó quan trắc được sử dụng làm tham chiếu. ROC cho biết mô hình nắm bắt được bao nhiêu sự kiện quan trắc so với số sự kiện dự báo sai. Còn AUC là diện tích dưới

đường cong ROC, cho biết mô hình có khả năng phân biệt giữa sự kiện có xảy ra và không xảy ra tốt đến mức nào. Qua đó nhận thấy, các đường ROC chủ yếu nằm gần đường chéo 45°, thậm chí một số trường hợp nằm dưới đường chéo, còn AUC chủ yếu xấp xỉ hoặc nhỏ hơn 0,5, chứng tỏ mô hình hầu như không có kỹ năng phân biệt các sự kiện mưa và không mưa theo các cấp.

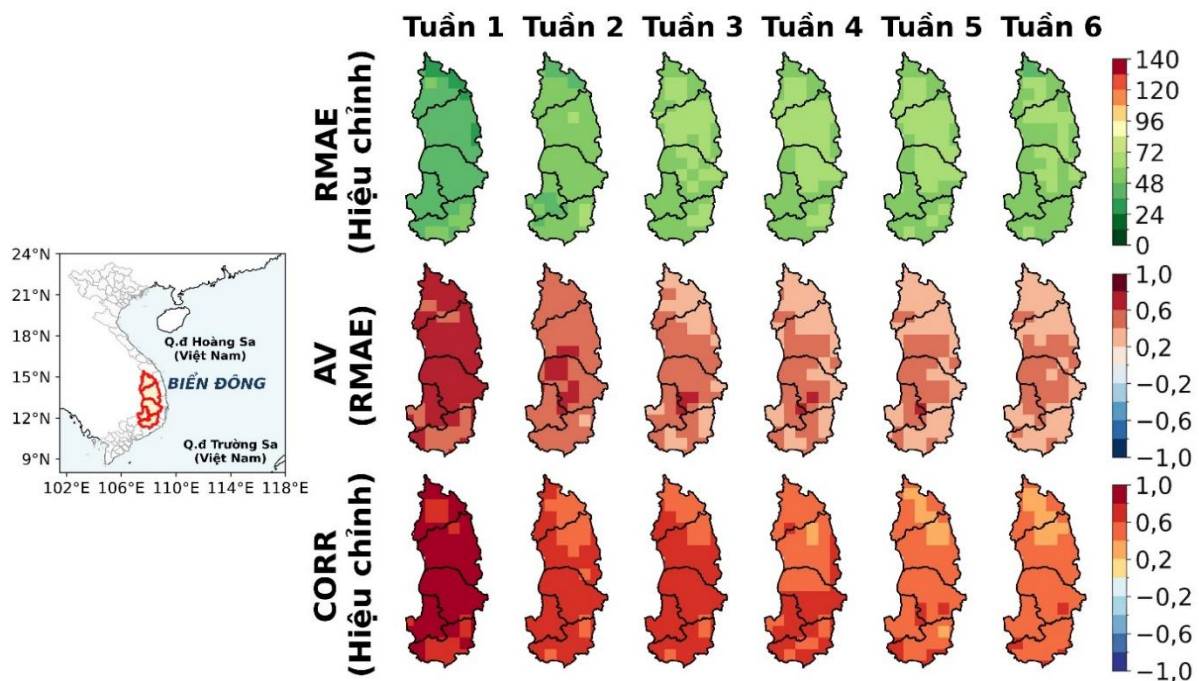
3.2. Kết quả thử nghiệm hiệu chỉnh sản phẩm S2S-ECMWF

Trong mục này sẽ trình bày kết quả ứng dụng ANN để hiệu chỉnh sản phẩm dự báo gốc của S2S-ECMWF. Hình 6 trình bày phân bố không gian của các chỉ số RMAE, AV và CORR đối với dự báo mưa hạn nội mùa trên khu vực Tây Nguyên cho các hạn dự báo từ W1 đến W6. Khi so sánh giá trị của RMAE và CORR trên các Hình 2 và 6 có thể thấy việc hiệu chỉnh đã cải thiện rõ rệt chất lượng dự báo: RMAE giảm đáng kể, còn CORR tăng lên rõ rệt. RMAE sau hiệu chỉnh chỉ dao động trong khoảng 30–70%, nhỏ hơn nhiều so với 80–110% thậm chí trên 150% trước khi hiệu chỉnh. CORR sau hiệu chỉnh phần

lớn nằm trong khoảng 0,4–0,9 so với xấp 0,0 thậm chí âm trước khi hiệu chỉnh.

Giá trị AV phổ biến lớn hơn 0,5 trên phần lớn diện tích khu vực, chứng tỏ việc hiệu chỉnh đã làm giảm RMAE một cách đáng kể so với trước khi hiệu chỉnh. Mức cải thiện thể hiện rõ ở tất cả các hạn dự báo, đặc biệt trong các tuần đầu. Mặc dù vậy, sai số vẫn có xu hướng tăng theo hạn dự báo, nghĩa là hiệu quả hiệu chỉnh giảm dần khi hạn dự báo càng dài. Hệ số tương quan CORR sau hiệu chỉnh đạt khoảng 0,8–0,9 ở W1, và giảm dần đến khoảng 0,3–0,6 ở các W5–W6.

Nhìn chung, kết quả thể hiện trên Hình 6 đã khẳng định việc hiệu chỉnh đã cải thiện một cách đáng kể chất lượng dự báo định lượng mưa của S2S-ECMWF cho khu vực Tây Nguyên, dù hiệu quả vẫn giảm dần theo hạn dự báo.



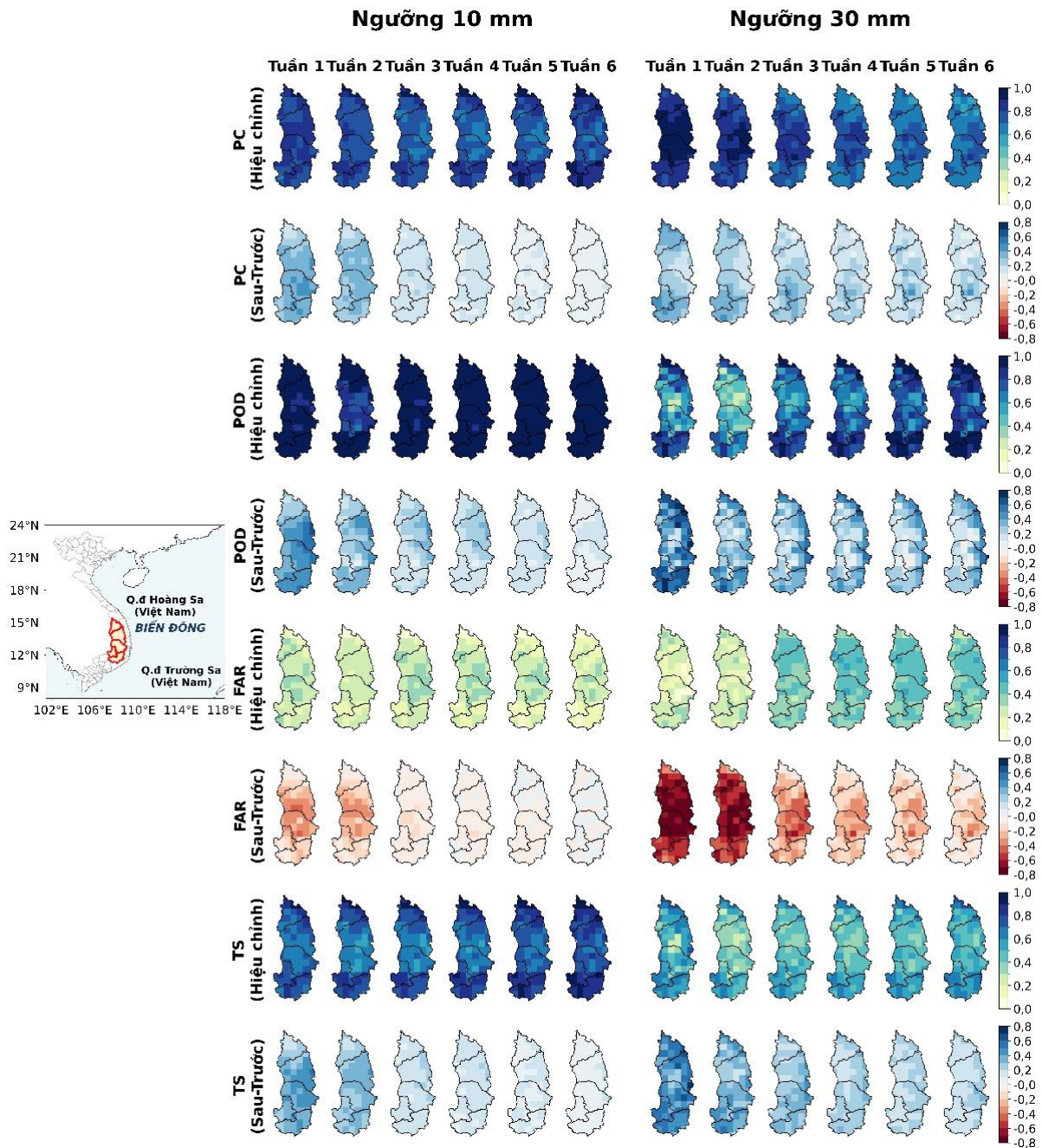
Hình 6. Phân bố không gian của RMAE, AV và CORR của S2S-ECMWF sau hiệu chỉnh trên khu vực Tây Nguyên giai đoạn 2003–2022.

Hình 7 trình bày phân bố không gian của các chỉ số đánh giá dự báo các sự kiện mưa của S2S-ECMWF sau khi hiệu chỉnh bằng ANN, và hiệu của chúng so với trước khi hiệu chỉnh, cho khu vực Tây Nguyên tại hai ngưỡng 10 mm và 30 mm và các hạn dự báo từ W1 đến W6.

Có thể thấy, sau khi hiệu chỉnh kỹ năng dự báo của mô hình cho ngưỡng 10 mm khá tốt. PC dao động trong khoảng 0,66–0,92, POD đạt 0,75–1,00 và TS đạt 0,58–0,92, trong khi FAR giảm xuống còn khoảng 0,08–0,37. Mức cải thiện thể hiện rõ nhất ở các W1–W2, với POD

tăng khoảng 0,27–0,58 và FAR có nơi chỉ khoảng 0,1–0,2. Điều này cho thấy việc hiệu chỉnh đã làm tăng khả năng phát hiện các sự kiện

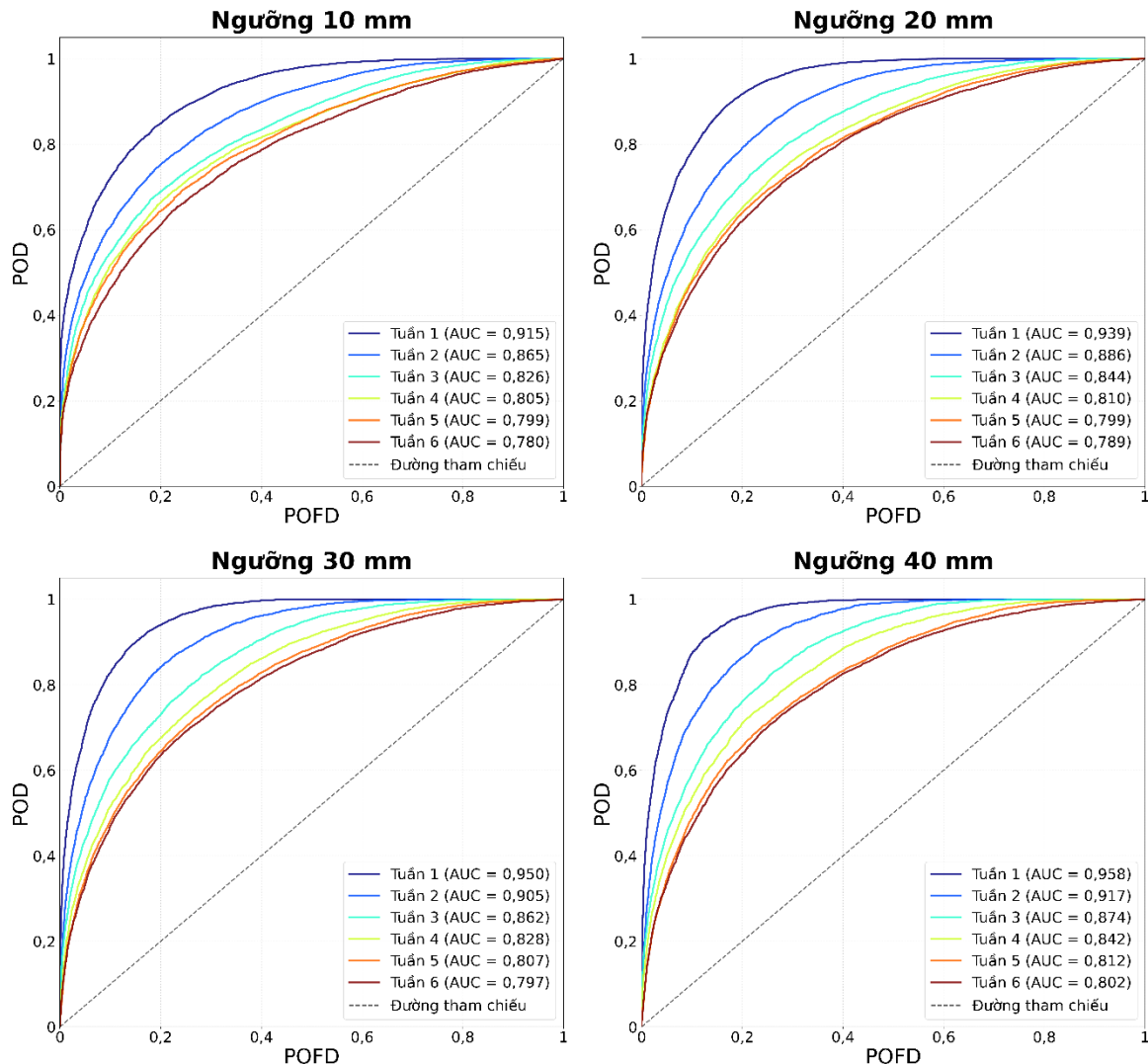
mưa nhỏ, đồng thời làm giảm số trường hợp dự báo không so với sản phẩm của mô hình gốc.



Hình 7. Phân bố không gian của các chỉ số PC, POD, FAR, TS sau khi hiệu chỉnh và chênh lệch của chúng so với trước khi hiệu chỉnh ở ngưỡng mưa 10 mm (bên trái) và 30 mm (bên phải) của mô hình S2S-ECMWF trên khu vực Tây Nguyên giai đoạn 2003–2022.

Tại ngưỡng 30 mm, hiệu quả của việc hiệu chỉnh vẫn được ghi nhận nhưng thấp hơn so với ngưỡng 10 mm. Cụ thể, sau hiệu chỉnh, giá trị POD dao động trong khoảng 0,24–0,97, TS đạt khoảng 0,23–0,68, trong khi FAR giảm đáng kể trên toàn khu vực, đặc biệt trong các tuần đầu. Tuy nhiên, kỹ năng dự báo tại ngưỡng này còn

phân hóa mạnh theo không gian và hạn dự báo. Hình 7 cũng cho thấy hiệu của PC, POD và TS sau hiệu chỉnh so với trước hiệu chỉnh đều dương, còn hiệu của FAR chủ yếu âm, chứng tỏ việc hiệu chỉnh đã cải thiện rõ rệt chất lượng dự báo các sự kiện mưa của S2S-ECMWF, đặc biệt đối với ngưỡng mưa nhỏ và các hạn dự báo ngắn hơn.



Hình 8. Tương tự hình 5 nhưng cho trường hợp sau hiệu chỉnh.

Hình 8 trình bày các đường cong ROC và giá trị của AUC, được xây dựng cho các ngưỡng mưa 10, 20, 30 và 40 mm sau khi hiệu chỉnh cho toàn khu vực Tây Nguyên. Qua đó nhận thấy, việc hiệu chỉnh đã cải thiện rõ rệt khả năng phân

biệt các sự kiện có và không có mưa vượt ngưỡng so với trước khi hiệu chỉnh (Hình 5). So với dự báo gốc (Hình 5), các đường ROC sau hiệu chỉnh đều nằm phía trên đường chéo 45° và dịch chuyển rõ về góc trên bên trái, phản ánh sự

gia tăng xác suất phát hiện đúng đồng thời giảm xác suất phát hiện sai.

Giá trị AUC sau hiệu chỉnh phần lớn vượt quá 0,8, thậm chí trên 0,9. Ở W1, AUC dao động trong khoảng 0,92 tại ngưỡng 10 mm đến 0,96 tại ngưỡng 40 mm. Khi hạn dự báo tăng lên, AUC có xu hướng giảm dần. Tuy nhiên, ngay cả ở W6, AUC vẫn duy trì khoảng 0,78–0,80, cao hơn rõ rệt so với trước khi hiệu chỉnh. Sau hiệu chỉnh kỹ năng dự báo khá ổn định đối với tất cả các ngưỡng mưa. Điều này, chứng tỏ việc hiệu chỉnh đã nâng cao khả năng phát hiện các sự kiện mưa, từ mưa nhỏ đến mưa vừa và mưa lớn.

4. Kết luận

Nghiên cứu này đã đánh giá khả năng dự báo mưa hạn nội mùa của sản phẩm S2S-ECMWF cho khu vực Tây Nguyên trong giai đoạn chuyển tiếp từ mùa khô sang mùa mưa (tháng 3–4), trên cơ sở so sánh với số liệu quan trắc giai đoạn 2003–2022. Đồng thời, nghiên cứu cũng đã thử nghiệm hiệu chỉnh lượng mưa dự báo bằng phương pháp mạng thần kinh nhân tạo (ANN) nhằm đánh giá khả năng cải thiện chất lượng sản phẩm dự báo.

Kết quả nhận được cho thấy sản phẩm dự báo gốc của S2S-ECMWF có khả năng nắm bắt được một số đặc điểm phân bố của lượng mưa, đặc biệt đối với các sự kiện mưa nhỏ. Tuy nhiên, chất lượng dự báo định lượng mưa còn khá hạn chế. Sai số ME có xu hướng thiên dương ở nhiều khu vực, nhất là ở các hạn dự báo dài; RMAE nhìn chung còn cao, trong khi hệ số tương quan CORR thấp, thậm chí mang giá trị âm ở một số nơi và một số hạn dự báo. Các chỉ số đánh giá dự báo pha cũng cho thấy mô hình có kỹ năng tốt hơn ở các ngưỡng mưa thấp, nhưng suy giảm rõ rệt khi ngưỡng mưa tăng. Cụ thể, mô hình có khả năng phát hiện mưa ở ngưỡng 10 mm tương đối tốt trong một số hạn dự báo. Tuy nhiên, khi các ngưỡng mưa tăng lên, đặc biệt từ 30–40 mm, xác suất phát hiện các sự kiện mưa (POD) và kỹ năng dự báo (TS) của mô hình giảm rõ rệt, còn tỷ lệ báo động sai (FAR) thì tăng lên. Việc xem xét chỉ số BIAS cũng cho thấy mô hình còn bỏ sót nhiều sự kiện mưa vừa đến mưa lớn. Kỹ năng

phân biệt các trường hợp có và không có mưa vượt ngưỡng của mô hình, được phản ánh qua đường cong ROC và chỉ số AUC, rất kém.

Sau khi hiệu chỉnh bằng ANN, chất lượng dự báo của S2S-ECMWF đã được cải thiện rõ rệt. Sai số RMAE giảm đáng kể, hệ số tương quan CORR tăng mạnh, đặc biệt ở các hạn dự báo gần. So với dự báo của mô hình gốc, việc hiệu chỉnh đã làm tăng đáng kể POD, TS và làm giảm FAR rõ rệt. Việc phân tích đường cong ROC và chỉ số AUC cho thấy, việc hiệu chỉnh đã làm tăng khả năng phân biệt các sự kiện mưa vượt ngưỡng một cách rõ rệt so với khi chưa hiệu chỉnh.

Qua đó có thể nói, việc ứng dụng phương pháp ANN để hiệu chỉnh sản phẩm mưa của S2S-ECMWF là hướng tiếp cận có tiềm năng trong việc nâng cao chất lượng dự báo mưa hạn nội mùa cho khu vực Tây Nguyên trong các tháng chuyển từ mùa khô sang mùa mưa. Việc sử dụng trung bình tổ hợp từ kết quả hiệu chỉnh của ANN với bộ các biến đầu vào khác nhau, kiến trúc mạng khác nhau là cách tiếp cận hợp lý, giúp làm giảm tính bất định của sản phẩm dự báo.

Mặc dù vậy, hiệu quả của việc hiệu chỉnh vẫn có xu hướng giảm khi hạn dự báo kéo dài, khả năng dự báo các sự kiện mưa lớn vẫn còn là thách thức. Hơn nữa, do hạn chế về độ dài chuỗi số liệu (20 năm, mỗi năm có 17 phiên dự báo), việc đánh giá cho từng ô lưới, từng hạn dự báo và từng ngưỡng lượng mưa sẽ dẫn đến, trong một số trường hợp, dung lượng mẫu sẽ quá nhỏ để đảm bảo độ ổn định thống kê, đặc biệt đối với các ngưỡng lượng mưa cao. Điều đó có thể làm tăng tính không chắc chắn trong kết quả nhận được. Bởi vậy, cần phải có những nghiên cứu sâu hơn trước khi đưa vào ứng dụng nghiệp vụ. Chẳng hạn, ngoài ANN có thể sử dụng các mô hình học máy, học sâu khác cũng như làm tăng dung lượng mẫu bằng các giải pháp khác nhau. Ngoài ra, để đảm bảo tính khách quan, cũng cần chạy dự báo thử nghiệm độc lập trên bộ số liệu đủ dài.

Tài liệu tham khảo

- [1] H. T. T. Huong, N. V. Luong, P. V. Vinh, P. T. Ha, Study on Bias Correction of Sub-seasonal Rainfall

- Forecast Products for the North Central Region Using Quantile Mapping with Gamma Approximation (QM-G), *Journal of Hydro-meteorology*, No. 748, 2023, pp. 77-89 (in Vietnamese).
- [2] P. T. Ha, Seasonal and Sub-seasonal Prediction of the Rainy Season Onset Date in Vietnam Based on Numerical Model Products, Doctoral Dissertation, VNU University of Science, Hanoi, Vietnam, 2022 (in Vietnamese).
 - [3] E. Olaniyan, E. A. Adefisan, F. Oni, E. Afiesimama, A. A. Balogun, K. A. Lawal, Evaluation of the ECMWF Sub-seasonal to Seasonal Precipitation Forecasts During the Peak of West Africa Monsoon in Nigeria, *Frontiers in Environmental Science*, Vol. 6, No. 4, 2018, <https://doi.org/10.3389/fenvs.2018.00004>.
 - [4] F. M. De Andrade, M. P. Young, D. MacLeod, L. C. Hirons, S. J. Woolnough, E. Black, Subseasonal Precipitation Prediction for Africa: Forecast Evaluation and Sources of Predictability, *Weather and Forecasting*, Vol. 36, No. 1, 2021, pp. 265-284, <https://doi.org/10.1175/WAF-D-20-0054.1>.
 - [5] F. Vitart, Evolution of ECMWF Sub-seasonal Forecast Skill Scores Over the Past 10 Years, ECMWF Technical Memorandum, European Centre for Medium-Range Weather Forecasts, Reading, United Kingdom, 2013.
 - [6] E. W. Benet, M. Pyrina, B. J. Esteve, E. Fraenkel, J. Cohen, D. I. V. Domeisen, Subseasonal Prediction of Central European Summer Heatwaves with Linear and Random Forest Machine Learning Models, *Artificial Intelligence for the Earth Systems*, Vol. 2, No. 2, 2023, pp. e220038, <https://doi.org/10.1175/AIES-D-22-0038.1>.
 - [7] X. Liu, S. Yang, A. Kumar, S. Weaver, X. Jiang, Diagnostics of Subseasonal Prediction Biases of the Asian Summer Monsoon by the NCEP Climate Forecast System, *Climate Dynamics*, Vol. 41, No. 5-6, 2013, pp. 1453-1474, <https://doi.org/10.1007/s00382-012-1553-3>.
 - [8] S. He, X. Li, T. DelSole, P. Ravikumar, A. Banerjee, Sub-Seasonal Climate Forecasting Via Machine Learning: Challenges, Analysis, and Advances, *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 35, No. 1, 2021, pp. 169-177.
 - [9] R. P. Worsnop, M. Scheuerer, T. M. Hamill, T. A. Smith, J. Schlör, RUFEO: A Deep Learning Framework to Postprocess Subseasonal Precipitation Accumulation Forecasts, *Artificial Intelligence for the Earth Systems*, 2024, <https://doi.org/10.1175/AIES-D-24-0020.1>.
 - [10] C. K. Mengist, K. H. Seo, Improving Subseasonal to Seasonal Global 2-m Temperature and Precipitation Forecasts Using a Deep Learning-Based Postprocessing Approach, *Monthly Weather Review*, Vol. 154, No. 3, 2026, pp. 391-405, <https://doi.org/10.1175/MWR-D-25-0179.1>.
 - [11] D. H. Yen, Assessment of Rainfall Forecasting Skill of the ECMWF Model over Vietnam. Master's Thesis, VNU University of Science, Hanoi, Vietnam, 2023 (in Vietnamese).
 - [12] Q. Tran-Anh, T. Ngo-Duc, E. Espagne, L. Trinh-Tuan, A 10-km CMIP6 Downscaled Dataset of Temperature and Precipitation for Historical and Future Vietnam Climate, *Scientific Data*, Vol. 10, 2023, pp. 257.
 - [13] X. T. Nguyen, D. T. Nguyen, H. Kamimera, T. L. Trinh, J. Matsumoto, T. Inoue, V. T. Phan, The Vietnam Gridded Precipitation (VnGP) Data Set: Construction and Validation, *Scientific Online Letters on the Atmosphere*, Vol. 12, 2016, pp. 291-295.
 - [14] A. Q. Tran, D. T. Nguyen, E. Espagne, T. L. Trinh, A High-Resolution Projected Climate Dataset for Vietnam: Construction and Preliminary Application in Assessing Future Change, *Journal of Water and Climate Change*, Vol. 13, No. 9, 2022, pp. 3379-3396.
 - [15] V. T. Le, V. T. Phan, Recent Changes in Temperature Extremes During Cold Season Over Northern Vietnam and Teleconnections With ENSO, IOD, NAO and AO, *International Journal of Climatology*, Vol. 46, No. 2, 2025, pp. e70190.
 - [16] M. Kubat, *An Introduction to Machine Learning* (2nd ed.), Springer International Publishing, Cham, Switzerland, 2017.
 - [17] I. T. Jolliffe, D. B. Stephenson (Eds.), *Forecast Verification: A Practitioner's Guide in Atmospheric Science* (2nd ed.), Wiley-Blackwell, Chichester, United Kingdom, 2012.