TẠP CHÍ
KHÍ TƯỢNG THỦY VĂN

Check for updates

Bài báo khoa học

Nghiên cứu, đánh giá khả năng ứng dụng sản phẩm dự báo hạn nội mùa của mô hình IFS CY48R1 và CFSv1 cho khu vực Bắc Trung Bộ

Đào Anh Công[1]*, Nguyễn Văn Lượng[1], Phan Văn Vinh[1], Phan Như Xuyến[1], Trịnh Tuấn Long[2]

[1] Đài KTTV khu vực Bắc Trung Bộ; daoanhcong.k55.hus@gmail.com; luongnvkttv@gmail.com; vinhpv@gmail.com; phannhuxuyen@gmail.com
[2] Trường Đại học KHTN Hà Nội; trinhtuanlong@gmail.com

*Tác giả liên hệ: daoanhcong.k55.hus@gmail.com; Tel: +84–948946895

Ban biên tập nhận bài: 15/6/2023 Ngày phản biện xong: 27/7/2023 Ngày đăng bài: 25/8/2023

Tóm tắt: Dự báo hạn nội mùa, hay còn gọi là khoảng dự báo từ 2 tuần đến 2 tháng, trước đây vẫn được coi là "sa mạc của dự báo", nhưng đây lại là giai đoạn lý tưởng để lập các kế hoạch trung hạn cho công tác quản lý nguồn nước, điều tiết hồ chứa. Ở Việt Nam, những nghiên cứu về bài toán này vẫn chỉ đang nằm ở những bước đi đầu tiên. Để lấp đầy chỗ hổng dự báo này, nhóm tác giả đã tiến hành đánh giá kỹ năng dự báo của mô hình IFS CY48R1 và CFSv1 thông qua việc tính toán và phân tích các chỉ số thống kê ME, MAE, RMSE, CORR, Brier score và ROC curve cho hạn dự báo nói trên sử dụng số liệu mưa vệ tinh TRMM và số liệu dự báo lại của hai mô hình. Kết quả chỉ ra rằng cả 2 mô hình đều cho thấy tiềm năng trong việc ứng dụng vào công tác dự báo hạn nội mùa cho khu vực Bắc Trung Bộ. Mô hình CFS có kỹ năng dự báo tốt hơn IFS vào các tháng 11 - tháng 3, nhưng lại thể hiện kỹ năng dự báo kém hơn khá nhiều so với IFS vào các tháng mùa mưa (tháng 7 - tháng 10). Cả 2 mô hình đều tiềm ẩn khả năng hiệu chỉnh sai số hệ thống vào các tháng 12 - tháng 4 và tháng 8 - tháng 10 ở phía Nam khu. Cả 2 mô hình đều có kỹ năng phát hiện những đợt mưa vừa, mưa to đến rất to, tuy nhiên mô hình IFS cho thấy khả năng vượt trội so với CFS khi có tỷ lệ dự báo khống thấp hơn nhiều.

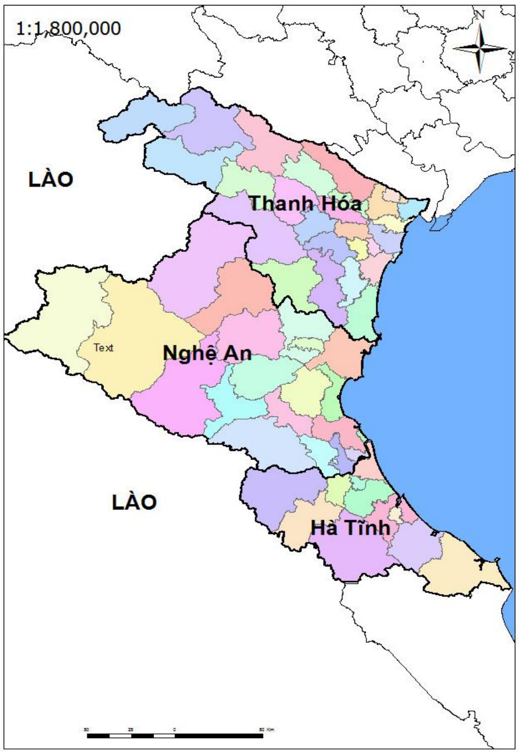
Từ khóa: ECMWF; NCEP; IFS CY48R1; CFSv1; Hạn nội mùa; S2S; Bắc Trung Bộ; Dự báo.

1. Mở đầu

Dự báo hạn nội mùa [1] (từ 2 tuần đến 2 tháng) là giai đoạn nằm giữa dự báo thời tiết (dự báo đến 15 ngày) và dự báo hạn mùa (3-6 tháng), trước đây được coi là "sa mạc của dự báo" [2–3] bởi khoảng thời gian này quá dài khiến phần lớn bộ nhớ về các điều kiện ban đầu của khí quyển bị mất đi và quá ngắn để cho sự biến đổi của đại dương có ảnh hưởng đến. Từ năm 2013, dự án S2S (*sub–seasonal to seasonal prediction project*) [3–4] đã được khởi động với mục tiêu nâng cao khả năng dự báo đối với hạn nội mùa và nâng cao hiểu biết về động lực cũng như các yếu tố khí hậu ảnh hưởng đến hạn dự báo này. Cũng từ khoảng thời gian này, các nghiên cứu về dự báo hạn nội mùa đã có nhưng bước phát triển đáng kể. Năm 2020, tác giả [5] đã đánh giá khả năng dự báo mưa từ tuần 1 đến tuần thứ 4 của các mùa ở châu Phi với 3 mô hình của các trung tâm ECMWF, UKMO và NCEP, kết quả cho thấy mô hình của ECMWF có kỹ năng dự báo vượt trội so với các mô hình khác và tốt hơn ở đông phi so với các khu vực khác, kỹ năng dự báo cho tuần 3-4 cũng giảm mạnh so với tuần 1-2. Năm 2018, nghiên cứu [6] cũng chỉ ra tiềm năng dự báo các hiện tượng thời tiết cực đoan với việc phát hiện trước 3 tuần dị thường nhiệt độ không khí tăng mạnh trong đợt sóng nóng kỷ lục ở Nga năm 2010, hay phát hiện khả năng xảy ra xoáy thuận nhiệt đới bằng việc dự báo dao động Madden-Julian (MJO).

Trong những năm trở lại đây, các nghiên cứu về dự báo hạn nội mùa ở Việt Nam cũng bắt đầu có những bước đi đầu tiên. Năm 2022, tác giả [7] đã nghiên cứu xác định ngày bắt đầu mùa mưa (NBĐMM) sử dụng số liệu dự báo lại của mô hình IFS và CFS đối với dự báo hạn mùa và hạn nội mùa, kết quả cho thấy sai số trung bình của ngày bắt đầu mùa mưa với hạn nội mùa dao động khá lớn (-35 đến +11 ngày) và phụ thuộc vào hạn dự báo cũng như chỉ tiêu xác định NBĐMM, sai số trung bình tuyệt đối cũng dao động tương đương với giá trị tuyệt đối của sai số trung bình (0-30 ngày), cho thấy tiềm năng về cải thiện chất lượng NBĐMM. Năm 2023, tác giả [8] đã tiến hành hiệu chỉnh số liệu dự báo lại (*reforecast*) của mô hình IFS, kết quả nhận được cho thấy sai số ME, MAE, RMSE giảm xuống rõ rệt từ tháng 12 đến tháng 3, các tháng 7-9 sai số tuy giảm nhưng vẫn cao, giá trị tương quan sau hiệu chỉnh hầu như không cải thiện.

Khoảng thời gian từ 2 tuần đến 2 tháng cũng được coi là giai đoạn lý tưởng để xây dựng những kế hoạch trung hạn cho công tác quản lý nguồn nước, điều tiết hồ chứa, đảm bảo sản xuất nông nghiệp và an ninh năng lượng nhưng đây lại là giai đoạn thiếu thông tin dự báo. Chính vì lý do đó, bài báo dưới đây sẽ tập trung nghiên cứu đánh giá kỹ năng dự báo hạn nội mùa của mô hình IFS CY48R1 và CFSv1 cho tổng lượng mưa từng 7 ngày từ tuần thứ 3-6 và nhận định về khả năng ứng dụng kết quả dự báo này cho khu vực Bắc Trung Bộ [9] với mục tiêu lấp đầy chỗ hổng thông tin này.



Hình 1. Bản đồ khu vực Bắc Trung Bộ.

2. Số liệu và phương pháp nghiên cứu

2.1. Số liệu nghiên cứu

Các bộ số liệu sử dụng trong nghiên cứu này bao gồm số liệu dự báo lại (*reforecast*) của IFS CY48R1 từ 2000-2019, bộ số liệu dự báo lại của CFSv1 từ 2000-2010 được tải về từ trang web chính thức của dự án S2S và số liệu mưa vệ tinh TRMM từ 2000-2019 được dùng để đánh giá như số liệu quan trắc.

Sản phẩm dự báo hạn nội mùa từ mô hình IFS CY48R1 (sau đây sẽ gọi tắt là IFS) [10–11] của ECMWF được cung cấp với tần suất 2 lần/tuần vào mỗi thứ 2 và thứ 5, sản phẩm bao gồm 11 members (thành phần) dự báo, có hạn dự báo lên đến 46 ngày với bước thời gian dự báo (*time step*) là 6 giờ và độ phân giải ngang là 0.25°×0.25°. Sản phẩm được cung cấp miễn phí thông qua dự án S2S và có thể tải về từ trang chủ của dự án S2S: https://apps.ecmwf.int/datasets/data/.

CFS (*Climate Forecast System*) là mô hình dự báo khí hậu hạn mùa được phát triển bởi Trung tâm Dự báo Môi trường Hoa Kỳ (NCEP) bao gồm 2 phiên bản CFSv1 (đã kết thúc nghiệp vụ từ 2011) và CFSv2 đã có nhiều cải tiến đáng kể so với phiên bản trước đó. Sản phẩm dự báo lại được sử dụng trong nghiên cứu này thuộc về phiên bản CFSv1 [10, 12] với độ phân giải từ 0.25° đến 0.5°, hạn dự báo là 44 ngày với 4 members. Bộ số liệu này cũng được phân phối miễn phí qua dự án S2S: https://apps.ecmwf.int/datasets/data/ và đã được nội suy về độ phân giải 0.25°×0.25° để đồng bộ về mặt không gian với mô hình IFS.

Sản phẩm mưa TRMM [13] (có số hiệu TMPA 3B42_Daily) là sản phẩm mưa 3 giờ được nội suy từ vệ tinh bởi NASA GES DISC (*Goddard Earth Sciences Data and Information Services Center*) dưới dạng lưới có độ phân giải 0.25ox0.25o. Sản phẩm được tải về từ trang web: https://disc.gsfc.nasa.gov/datasets/TRMM_3B42_Daily_7/summary.

2.2. Phương pháp nghiên cứu

Trong nghiên cứu này, nhóm tác giả sẽ đánh giá hạn dự báo từ 3 đến 6 tuần với từng độ phân giải thời gian là 7 ngày trong 20 năm (2000-2019) đối với mô hình IFS và 11 năm (2000-2010) đối với mô hình CFSv1. Kết quả dự báo nội mùa của IFS có hạn dự báo đến 46 ngày, kết quả dự báo của mô hình CFSv1 và số liệu mưa TRMM sẽ được chia thành 46/7 = 6 lead-times (hạn dự báo) với độ phân giải 7 ngày và được đánh giá tương ứng với nhau. Tại mỗi ô lưới tương ứng với mỗi hạn dự báo có hai chuỗi số liệu quan trắc và dự báo. Các chỉ số đánh giá thống kê ME, MAE, RMSE, CORR và chỉ số đánh giá pha brier score và ROC curve sẽ được tính cho từng ô lưới.

Với n là số cặp giá trị quan trắc - dự báo, f_i và O_i lần lượt là giá trị dự báo và giá trị quan trắc thứ i (i=1, 2,…,n–1, n), $\overline{f}$ và $\overline{O}$ lần lượt là giá trị trung bình của dự báo và quan trắc. Các chỉ số đánh giá được tính bằng các công thức sau:

- ME (*Mean error* - Sai số trung bình):

$$ME = \frac{1}{n}\sum_{i=1}^{n}\left(f_i - O_i\right) \quad (1)$$

Giá trị ME cho biết khuynh hướng thiên cao hay thiên thấp của giá trị dự báo (f_i) so với giá trị quan trắc (O_i). Giá trị ME nằm trong khoảng (–∞; +∞). ME âm cho thấy giá trị dự báo thấp hơn giá trị quan trắc và ngược lại. Khả năng dự báo được xem là “hoàn hảo” (không thiên lệch về một phía nào cả) nếu ME = 0. Tuy nhiên, do ME không thực sự phản ánh độ lớn của sai số, nên các chỉ số thống kê khác được sử dụng để đánh giá một cách toàn diện hơn.

- MAE (*Mean absolute error* - Sai số tuyệt đối trung bình) [14]:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left(\left|f_i - O_i\right|\right) \quad (2)$$

Giá trị MAE mô tả mức độ lệch của giá trị dự báo so với giá trị quan trắc tuy nhiên lại không thể đánh giá được xu thế thiên dương hay thiên âm của dự báo. Giá trị MAE nằm trong khoảng (0; +∞). Khả năng dự báo của mô hình là tốt nhất khi MAE = 0 tương đương với giá trị dự báo và quan trắc hoàn toàn không sai khác.

- RMSE (*Root mean square error* - Sai số trung bình quân phương) [15]:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(f_i - O_i\right)^2} \quad (3)$$

Tương tự như MAE, RMSE cũng biểu thị mức độ chênh lệch của dự báo so với quan trắc và có giá trị nằm trong khoảng (0; +∞), tuy nhiên khác với MAE khi các sai số đều có vai trò tương đương nhau thì RMSE phóng đại các sai số lớn của dự báo. Giá trị RMSE luôn lớn hơn hoặc bằng với MAE, nếu khoảng cách giữa RMSE và MAE càng lớn thì chứng tỏ mô hình càng có nhiều sai số khác nhau. Khi RMSE = MAE, tất cả sai số đều có chung mức độ.

- CORR (*Correlation Coefficient* - Hệ số tương quan):

$$\mathrm{CORR} = \frac{\frac{1}{n}\sum_{i=1}^{n}\left(f_i - \bar{f}\right)\left(O_i - \bar{O}\right)}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(f_i - \bar{f}\right)^2\left(O_i - \bar{O}\right)^2}} \tag{4}$$

Hệ số tương quan đánh giá tính nhất quán của sự thay đổi theo thời gian hoặc không gian giữa giá trị dự báo và giá trị quan trắc, giá trị của CORR nằm trong khoảng (–1, +1). Khi CORR nhận giá trị dương (âm), mối quan hệ giữa dự báo và quan trắc là đồng biến (nghịch biến), CORR càng tiến đến ±1 thì tương quan giữa 2 biến càng mạnh, ngược lại khi CORR càng tiến đến 0 thì tương quan càng yếu.

- Brier score (BS)

$$\mathrm{BS} = \frac{1}{n}\sum_{i=1}^{n}\left(f_i - O_i\right)^2 \tag{5}$$

Chỉ số Brier score [16–17] được tính bằng cách lấy trung bình của n lần dự báo, với f_i là dự báo xác suất xảy ra sự kiện và O_i nhận giá trị 1 hoặc 0 tương ứng với sự kiện có xảy ra hay không. Brier score được sử dụng để đánh giá độ chính xác của một dự báo xác suất và có giá trị nằm trong khoảng (0,1) với giá trị 0 tương ứng dự báo chính xác hoàn toàn và 1 tương ứng dự báo hoàn toàn sai. Brier score có thể cho chúng ta biết mức độ dự báo chính xác là bao nhiêu, nhưng không nói lên được nó có tốt hơn so với trung bình khí hậu hay không.

- ROC curve (*Receiver operating characteristic curve* - đường cong đặc tính vận hành tiếp nhận) [18–20]

Bảng 1. Bảng khả năng xảy ra của dự báo/quan trắc.

		Quan trắc		
		Positive	Negative	Tổng
Dự báo	Positive	True Positive (TP)	False positive (FP)	TP + FP
	Negative	False Negative (FN)	True Negative (TN)	FN+TN
	Tổng	TP + FN	FP + TN	

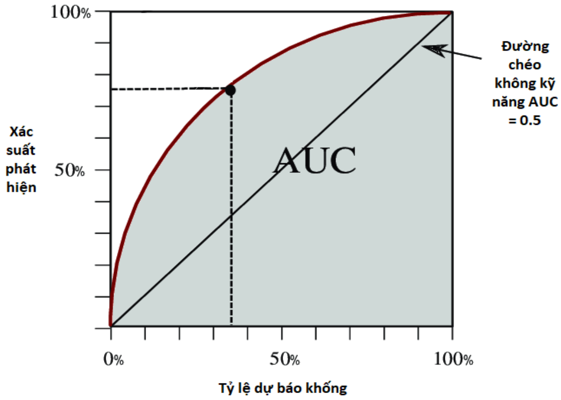
- Xác suất phát hiện POD (*probability of detection*) = TPR (*true positive rate*): cho biết tỷ lệ dự báo đúng (dự báo có và quan trắc có) so với tổng tất cả các lần quan trắc có, có ngưỡng giá trị nằm trong khoảng (0,1) với giá trị 0 thể hiện không có kỹ năng và 1 thể hiện mô hình chính xác hoàn toàn.

$$\mathrm{POD} = \mathrm{TPR} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{6}$$

- Tỷ lệ dự báo khống FAR (*False alarm rate*) = FPR (*False positive rate*): cho biết tỷ lệ dự báo khống (dự báo có nhưng quan trắc không) so với tổng số lần quan trắc không có hiện tượng xảy ra, giá trị của FAR cũng nằm trong ngưỡng (0,1) với giá trị 0 thể hiện kỹ năng dự báo hoàn hảo.

$$\mathrm{FAR} = \mathrm{FPR} = \frac{\mathrm{FP}}{\mathrm{FP} + \mathrm{TN}} \tag{7}$$

Đường cong ROC là đồ thị miêu tả mối quan hệ giữa POD được đặt làm trục tung và FAR đóng vai trò là trục hoành. Đường cong thể hiện kỹ năng dự báo tốt khi nó càng hướng lên góc phía trên bên trái - biểu thị sự vượt trội của POD đối với FAR, nếu đường cong trùng với đường chéo no discrimination (không xác định/không kỹ năng), chứng tỏ dự báo không có kỹ năng (POD = FAR), cuối cùng, nếu đường cong nằm dưới đường chéo, thể hiện kỹ năng dự báo kém hơn cả "đoán mò" (POD < FAR). Giá trị AUC (*Area under ROC curve* - diện tích dưới đường cong) là phần diện tích được tính từ đường cong ROC đến hết phần dưới của đồ thị và có giá trị nằm trong khoảng (0,1). Mô hình dự báo được coi là tối ưu khi AUC ≥ 0,8, và mô hình đánh giá không có kỹ năng khi UAC = 0,5. Tuy nhiên, với các sự kiện hiếm, việc đánh giá pha có xu hướng không chính xác.



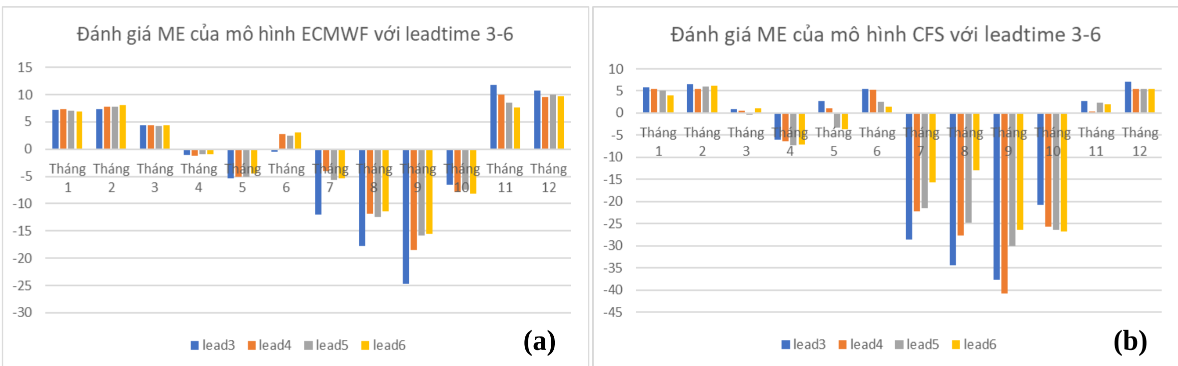
Hình 2. Đồ thị đường cong ROC [20].

3. Kết quả và thảo luận

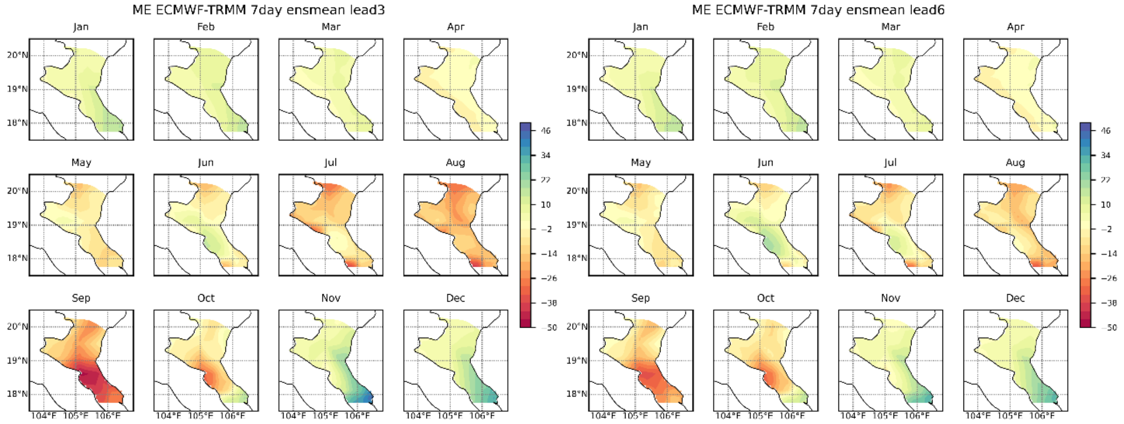
Nhóm nghiên cứu sẽ tiến hành đáng giá kết quả tính toán các giá trị ME, MAE, RMSE, CORR, Brier score và đường cong ROC của trung bình các thành phần dự báo (*ensemble mean*) theo phân bố về thời gian và phân bố về không gian cho độ phân giải 7 ngày và lead-time 3-6 của hai mô hình IFS và CFSv1.

3.1. Phân tích và đánh giá kết quả sai số trung bình ME

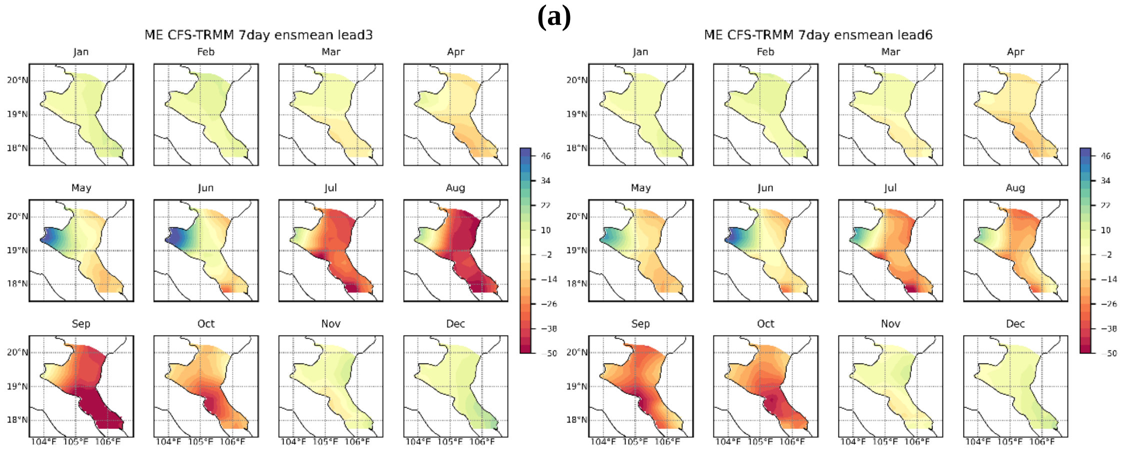
Có thể nhận thấy từ hình 3a và 3b rằng, trong giai đoạn mùa khô từ tháng 11 đến tháng 3, chỉ số ME ở cả hai mô hình IFS và CFSv1 đều có xu hướng thiên dương với giá trị tương đối thấp (5,10) mm. Sai số giữa các lead-time là tương đối đồng đều và ME của IFS thấp hơn một chút so với ME của CFSv1. Giá trị ME của IFS đạt cực tiểu vào tháng 4 trong khi ở mô hình CFSv1 là tháng 3. Từ tháng 4 đến tháng 5, ME ở cả 2 mô hình chuyển sang thiên âm và quay lại thiên dương trong tháng 6 với giá trị trong khoảng (-5,+5) mm. Bắt đầu từ tháng 7 đến tháng 10, ME chuyển sang thiên âm khá lớn ở cả 2 mô hình, khi giá trị ME của IFS nằm trong khoảng (-25, -5) mm thì ME của CFSv1 thiên âm mạnh hơn (-41, -15) mm, ở các tháng 7-9, sai số cao hơn trong các lead-time 3-4 và giảm dần ở lead-time 6-7. Trong tháng 10, sau số giảm dần và các lead-time có sai số gần bằng nhau.



Hình 3. (a) Phân bố theo thời gian giá trị ME của mô hình IFS so với TRMM từ lead-time 3-6; (b) Phân bố theo thời gian giá trị ME của mô hình CFSv1 so với TRMM từ lead-time 3-6.



(a)



(b)

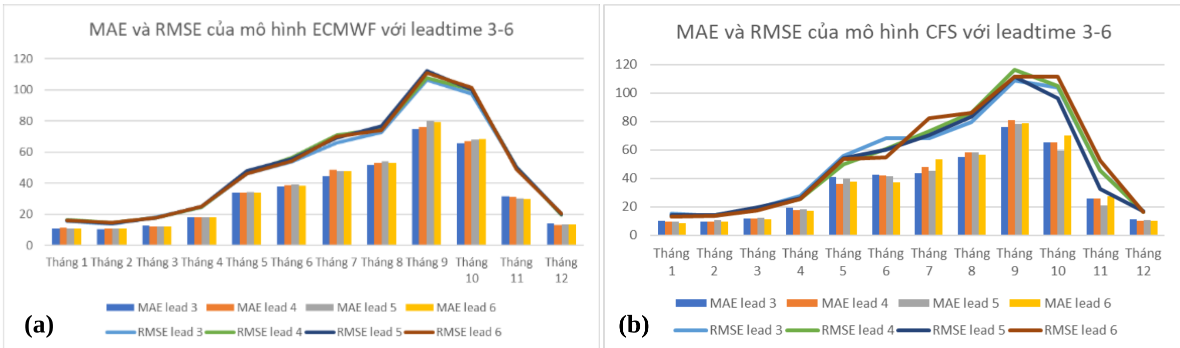
Hình 4. (a) Phân bố theo không gian giá trị ME của mô hình IFS so với TRMM ở lead-time 3 và 6; (b) Phân bố theo không gian giá trị ME của mô hình CFSv1 so với TRMM ở lead-time 3 và 6.

Hình 4a và 4b, có thể thấy giá trị ME ở lead-time 3 và lead-time 6 của cả hai mô hình đều tương tự như nhau, khác biệt duy nhất ở đây là ME ở lead-time 6 đã giảm đi đáng kể so với lead-time 3. Ở cả 2 mô hình, ME từ tháng 11 đến tháng 4 là tương đối thấp (-5, +5), tuy nhiên trong các tháng 11-12, trong khi CFSv1 cho sai số ME khá đồng đều ở trên toàn khu vực thì IFS cho thấy mức độ thiên dương cao hơn khá rõ ở phía nam của khu vực. Từ tháng 5 đến tháng 6, IFS cho kết quả thiên âm nhỏ và khá đồng đều trên cả khu vực, CFSv1 lại thể hiện dự báo thiên dương tương đối cao ở khu vực vùng núi phía tây Nghệ An. Từ tháng 7-10, CFSv1 cho thiên dương nhỏ ở vùng núi phía tây Nghệ An và thiên âm rất lớn ở những vùng còn lại, nhất là phía nam khu vực trong tháng 9. Trong khoảng thời gian đó, IFS cho kết quả thiên âm trung bình trên toàn khu vực, đến tháng 9-10 chuyển sang thiên âm khá rõ ở khu vực phía nam.

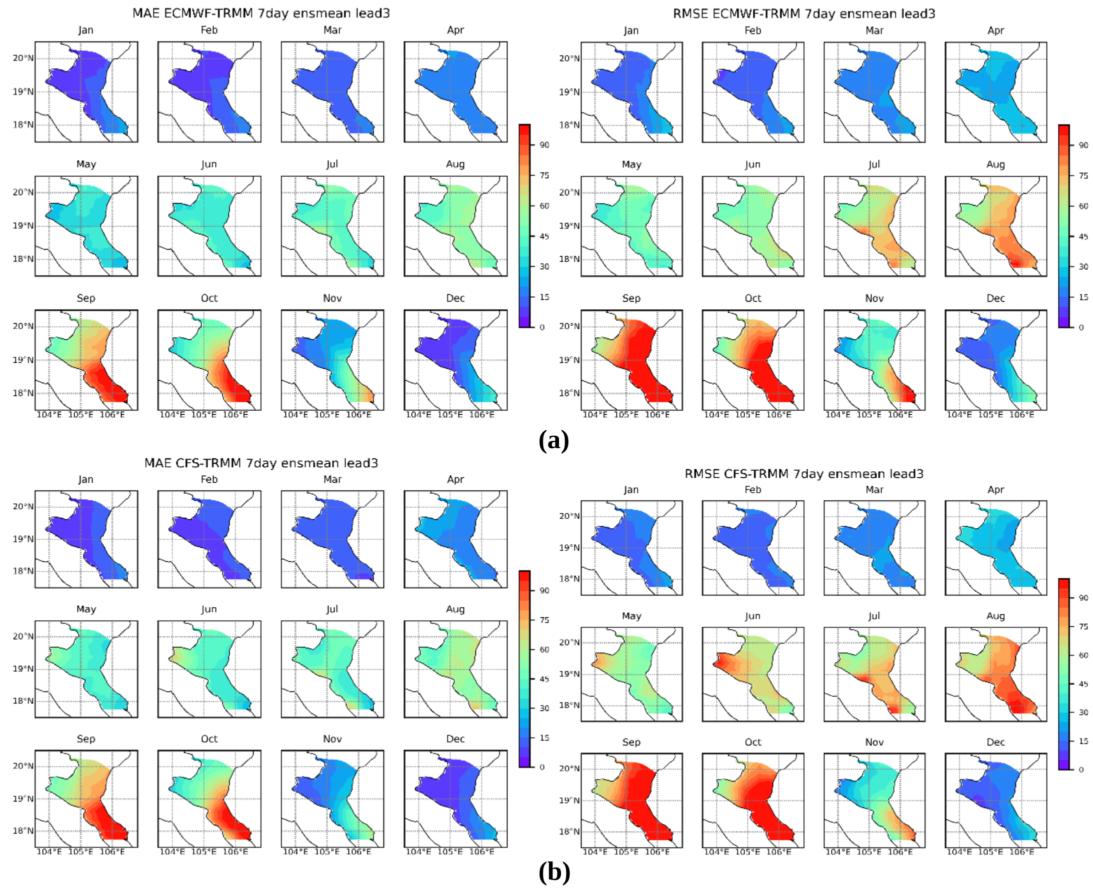
3.2 Phân tích và đánh giá kết quả sai số trung bình tuyệt đối MAE và sai số quân phương RMSE

Hình 5a và 5b có thể thấy ở cả 2 mô hình, giá trị MAE (biểu diễn dưới dạng cột) hầu như không có sự chênh lệch đáng kể và RMSE (biểu diễn dưới dạng đường) cũng vậy, khác biệt duy nhất là sai số ở các lead-time của IFS khá tương đồng nhau, trong khi ở mô hình CFSv1, các lead-time có sai số chênh lệch trong một khoảng nhỏ. Từ tháng 12 đến tháng 4, MAE và RSME ở cả 2 mô hình có giá trị tương đối thấp (10-20 mm) và gần như xấp xỉ nhau. Nhưng từ tháng 5 đến tháng 10, MAE và RMSE tăng lên khá nhanh ở cả 2 mô hình và đạt giá trị cao nhất vào tháng 9 (100-120 mm), sau đó giảm dần trong các tháng 10-11. Giá trị của RMSE trong giai đoạn này cũng cao hơn hẳn so với MAE, tuy nhiên khoảng cách của RMSE so với

MAE trong các tháng ở mô hình IFS có xu hướng đồng đều hơn, cho thấy khả năng sai số của mô hình có tính "hệ thống". Trong khi đó, cách biệt giữa MAE và RSME ở mô hình CFSv1 có vẻ bất ổn định hơn, cho thấy sai số là tương đối tản mạn.



Hình 5. (a) Phân bố theo thời gian giá trị ME của mô hình IFS so với TRMM từ lead-time 3-6; (b) Phân bố theo thời gian giá trị ME của mô hình vv1 so với TRMM từ lead-time 3-6.

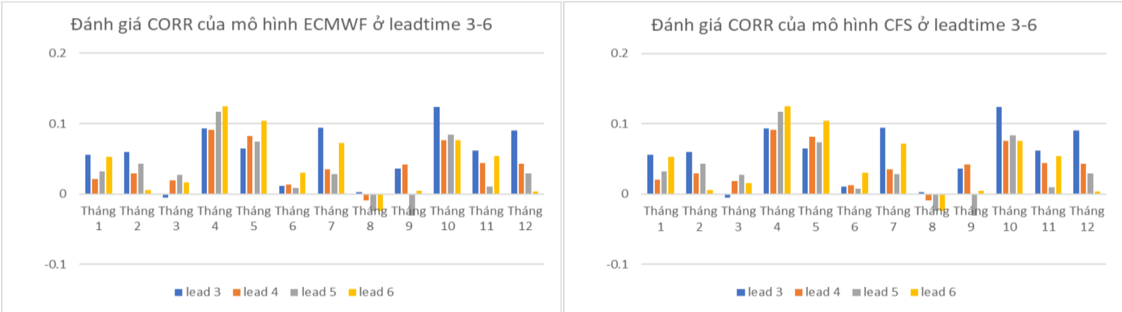


Hình 6. (a) Phân bố theo không gian MAE và RMSE của IFS so với TRMM ở lead-time 3; (b) Phân bố theo không gian MAE và RMSE của CFSv1 so với TRMM ở lead-time 3.

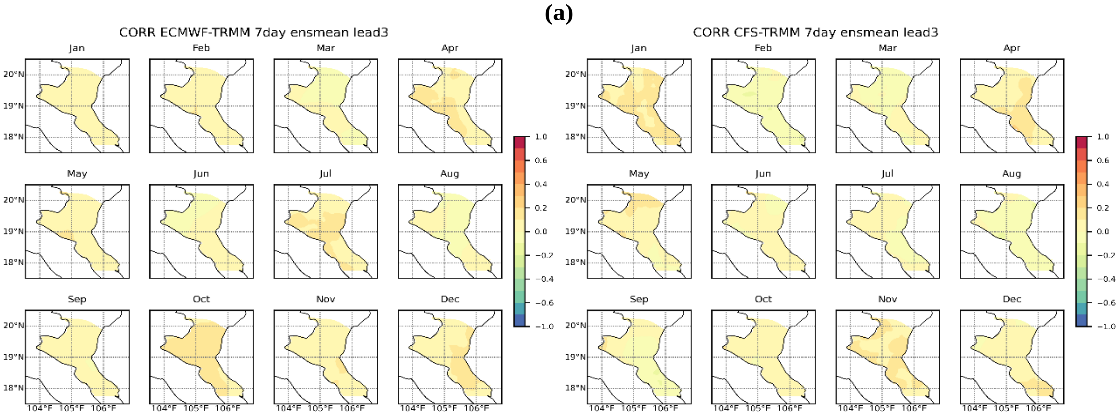
Hình 6a và 6b thấy phân bố theo không gian của MAE và RMSE của cả hai mô hình IFS và CFSv1 là khá tương đồng, cụ thể như sau: Từ tháng 12 đến tháng 3, cả MAE và RSME đều có giá trị nhỏ (dưới 20mm) và tương đối gần nhau. Mọi thứ đảo ngược khi bước vào các tháng 5-8, khi MAE đạt giá trị trung bình (30-60 mm) và dàn trải đều trên toàn bộ khu vực, thì RSME có xu hướng tăng cao hơn ở vùng đồng bằng ven biển và phía nam khu vực (60-90 mm). Trong tháng 9-10, trong khi ở phía nam của khu vực RMSE và MAE tương đồng nhau và đạt giá trị cực đại (90-120) tiềm ẩn khả năng về tính hệ thống trong sai số của mô hình, thì ở phía bắc khu vực, MAE và RSME chênh lệch nhau là khá nhiều, cho thấy sự phân tán trong sai số. Điều này cho thấy phân bố nam - bắc và đồng bằng - vùng núi của MAE và RMSE là tương đối khác biệt.

3.3. Phân tích và đánh giá kết quả hệ số tương quan CORR

Hình 7a và 7b thấy rằng hệ số tương quan ở cả 2 mô hình là rất thấp, cho thấy dự báo và quan trắc hầu như không có quan hệ tuyến tính nào.



(a)

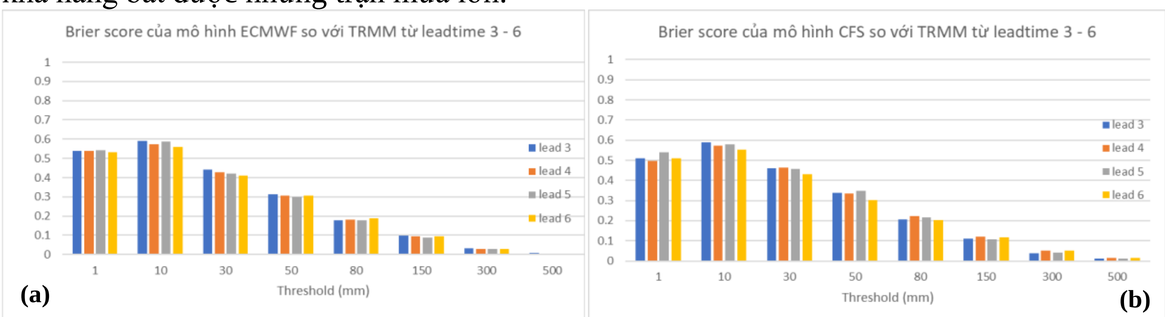


(b)

Hình 7. (a) Phân bố theo thời gian CORR của IFS (trái) và CFSv1 (phải) so với TRMM ở lead-time 3; (b) Phân bố theo không gian CORR của IFS (trái) và CFSv1 (phải) so với TRMM ở lead-time 3.

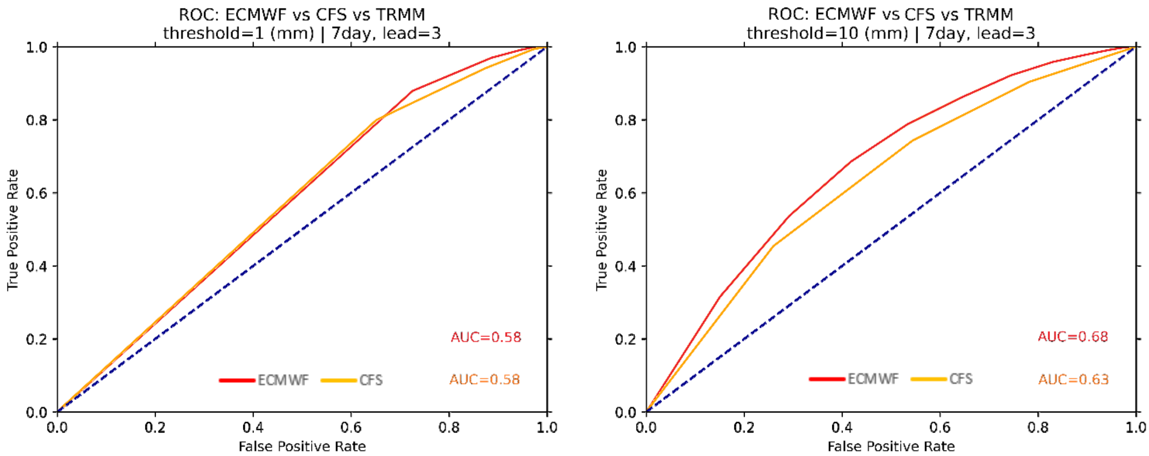
3.4. Phân tích và đánh giá chỉ số Brier score (BS)

Từ hình 8a và 8b, có thể thấy giá trị BS ở các lead-time hầu như có rất ít sự khác biệt. Với lượng mưa dưới 30mm/7 ngày ở cả 2 mô hình, giá trị của BS nằm trong khoảng 0.4 – 0.6, cho thấy kỹ năng dự báo chỉ ở mức trung bình. Với lượng mưa 50mm/7 ngày, cả hai mô hình đã cho thấy sự cải thiện trong kỹ năng dự báo khi có BS xấp xỉ 0.3. Đối với lượng mưa từ 80-150 mm/7 ngày, BS có giá trị dưới 0.2, chứng tỏ khả năng phát hiện những đợt mưa vừa và mưa to của cả IFS và CFSv1 là khá tốt. Với những đợt mưa rất to có tổng lượng 300 đến 500mm/7 ngày, chỉ số BS có giá trị tiệm cận và bằng 0, cho thấy cả 2 mô hình đều có khả năng bắt được những trận mưa lớn.



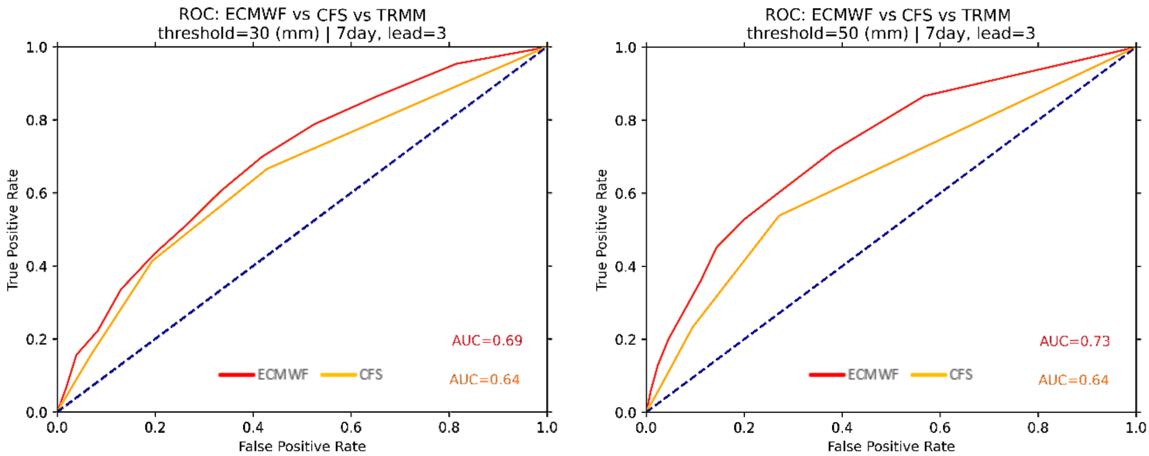
Hình 8. (a) Brier score của mô hình IFS so với TRMM với các ngưỡng mưa khác nhau từ lead-time 3-6; (b) Brier score của mô hình CFSv1 so với TRMM với các ngưỡng mưa khác nhau từ lead-time 3-6.

3.5 Phân tích và đánh giá đường cong ROC và AUC



Hình 9. (a) ROC của IFS và CFS tại ngưỡng mưa 1mm; (b) ROC của IFS và CFS tại ngưỡng mưa 10mm.

Hình 9a cho thấy ở ngưỡng mưa rất nhỏ (1 mm/7 ngày), đường cong ROC của hai mô hình IFS và CFSv1 đều chỉ nhỉnh một chút so với đường chéo không phân biệt, giá trị AUC của cả 2 mô hình đều là 0.58 cho thấy kỹ năng dự báo pha tương đối kém. Hình 9b đánh giá kỹ năng dự báo với ngưỡng mưa nhỏ (10 mm/7 ngày và 30 mm/7 ngày), cả 2 mô hình đều cho kỹ năng ở mức trung bình, tuy nhiên mô hình IFS thể hiện kỹ năng dự báo nhỉnh hơn với AUC = 0,68 so với AUC của CFSv1 là 0,63.

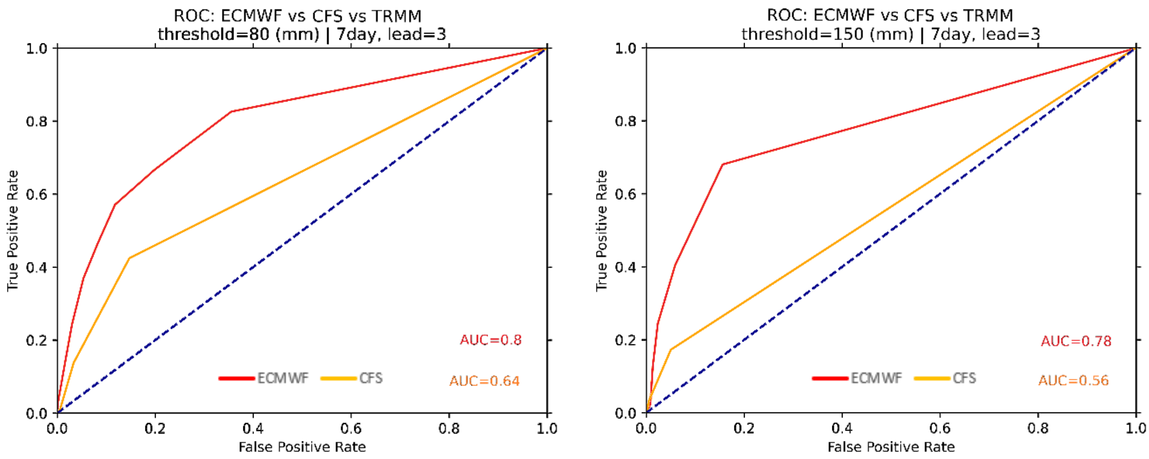


Hình 9. (c) ROC của IFS và CFS tại ngưỡng mưa 30mm; (d) ROC của IFS và CFS tại ngưỡng mưa 50mm.

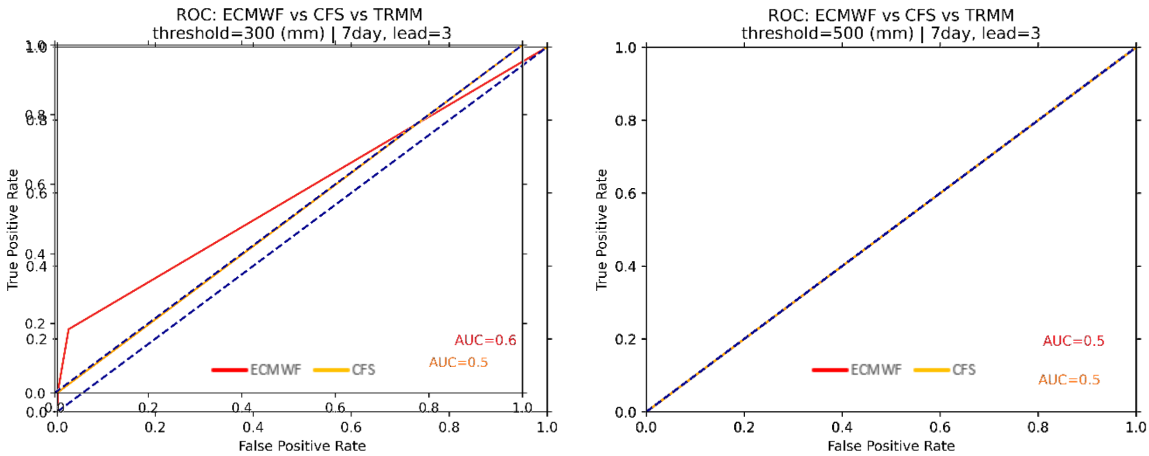
Đánh giá ở ngưỡng mưa 30 mm/7 ngày, hình 9c cho thấy một lần nữa cả 2 mô hình đều thể hiện kỹ năng dự báo trung bình, và IFS vẫn có kỹ năng nhỉnh hơn so với CFSv1. Kỹ năng dự báo ở ngưỡng mưa 30 mm/7 ngày cũng cao hơn so với 10 mm/7 ngày. Ở ngưỡng mưa 50 mm/7 ngày, trong khi mô hình CFSv1 chỉ thể hiện kỹ năng dự báo trung bình với AUC = 0,64, thì IFS cho thấy kỹ năng dự báo tốt hơn với AUC có giá trị 0,73 (Hình 9c, 9d).

Hình 9e và 9f nhận ra sự vượt trội đáng kể đối với kỹ năng dự báo của mô hình IFS với lượng mưa vừa (80 mm/7 ngày) và mưa to (150 mm/7 ngày), mô hình IFS đã thể hiện kỹ năng dự báo rất tốt với AUC lần lượt là 0,8 và 0,78. Trong khi đó, mô hình CFSv1 cho thấy kỹ năng dự báo trung bình và kém khi giá trị AUC chỉ lần lượt là 0,64 và 0,56.

Đánh giá kỹ năng dự báo mưa rất to (300-500 mm), cả 2 mô hình đều cho kỹ năng dự báo kém và cực kỳ kém khi đường cong ROC tiệm cận và trùng với đường chéo không phân biệt, AUC của cả 2 mô hình cũng có giá trị cực kỳ thấp, chỉ ở trong ngưỡng 0,5–0,6.



Hình 9. (e) ROC của IFS và CFS tại ngưỡng mưa 80mm; (f) ROC của IFS và CFS tại ngưỡng mưa 150mm.



Hình 9. (g) ROC của IFS và CFS tại ngưỡng mưa 300mm; (h) ROC của IFS và CFS tại ngưỡng mưa 500 mm.

4. Kết luận

Thông qua quá trình đánh giá kỹ năng dự báo của mô hình IFS và CFSv1 so với số liệu mưa TRMM, có thể rút ra một vài kết luận như sau:

Cả 2 mô hình đều cho thấy tiềm năng trong việc ứng dụng vào công tác dự báo hạn nội mùa cho khu vực Bắc Trung Bộ. Mô hình CFSv1 có kỹ năng dự báo tốt hơn IFS vào các tháng 11 - tháng 3, nhưng lại thể hiện kỹ năng dự báo kém hơn khá nhiều so với IFS vào các tháng mùa mưa (tháng 7 - tháng 10). Cả 2 mô hình đều tiềm ẩn khả năng hiệu chỉnh sai số hệ thống vào các tháng 12 - tháng 4 và tháng 8 - tháng 10 ở phía Nam khu vực, khi MAE và RMSE của cả 2 mô hình trong những giai đoạn này có giá trị tương đồng với nhau. Đối với hệ số tương quan CORR, cả 2 mô hình đều cho giá trị rất thấp, cho thấy hầu như không có quan hệ tuyến tính giữa dự báo và kết quả quan trắc. Cả 2 mô hình đều có kỹ năng phát hiện những đợt mưa vừa, mưa to đến rất to, tuy nhiên mô hình IFS cho thấy khả năng vượt trội so với CFSv1 khi có tỷ lệ dự báo khống thấp hơn nhiều.

Đóng góp của tác giả: Xây dựng ý tưởng nghiên cứu: D.A.C.; Xử lý số liệu: T.T.L.; Viết bản thảo bài báo: P.V.V., D.A.C., N.V.L.; Chỉnh sửa bài báo: P.N.X., D.A.C.

Lời cảm ơn: Tập thể tác giả xin trân trọng cảm ơn Đề tài cấp bộ “Nghiên cứu xây dựng công nghệ dự báo mưa hạn nội mùa từ một số mô hình toàn cầu kết hợp với mô hình thủy văn để phục vụ quản lý nguồn nước, điều tiết hồ chứa cho khu vực Bắc Trung Bộ”, mã số TNMT.2022.02.17 đã hỗ trợ về số liệu và phương pháp luận để thực hiện bài báo này.

Lời cam đoan: Tập thể tác giả cam đoan bài báo này là công trình nghiên cứu của tập thể tác giả, chưa được công bố ở đâu, không được sao chép từ những nghiên cứu trước đây; không có sự tranh chấp lợi ích trong nhóm tác giả.

Tài liệu tham khảo

1. Vitart, F.; Ardilouze, C.; Bonet, A.; Brookshaw, A.; Chen, M.; Codorean, C.; Déqué, M.; Ferranti, L.; Fucile, E.; Fuentes, M.; Hendon, H.; Hodgson, J.; Kang, H.S.; Kumar, A.; Lin, H.; Liu, G.; Liu, X.; Malguzzi, P.; Mallas, I.; Manoussakis, M.; Mastrangelo, D.; MacLachlan, C.; McLean, P.; Minami, A.; Mladek, R.; Nakazawa, T.; Najm, S.; Nie, Y.; Rixen, M.; Robertson, A.W.; Ruti, P.; Sun, C.; Takaya, Y.; Tolstykh, M.; Venuti, F.; Waliser, D.; Woolnough, S.; Wu, T.; Won, D.J.; Xiao, H.; Zaripov, R.; Zhang, L. The subseasonal to seasonal (S2S) prediction project database. *Bull. Amer. Meteor. Soc.* **2017**, *98*, 163–173.
2. Vitart, F.; Andrew, W.R. Sub–seasonal to seasonal prediction, XIII–XIV, 2019.
3. Christopher, J.W. et al. Potential applications of subseasonal-to-seasonal (S2S) predictions. *Meteorol. Appl.* **2017**, 24(3), 315–325.
4. Pegion, K. et al. The subseasonal experiment (SubX): A multimodel subseasonal prediction experiment. *Bull. Amer. Meteor. Soc.* **2019**, *100*, 2043–2060. https://doi.org/10.1175/BAMS-D-18-0270.1.
5. de Andrade, F.M.; Young, M.P.; MacLeod, D.; Hirons, L.C.; Woolnough, S.J.; Black, E. Subseasonal precipitation prediction for africa: Forecast evaluation and sources of predictability. *Wea. Forecasting* **2021**, *36*, 265–284.
6. Vitart, F.; Robertson, A.W. The sub–seasonal to seasonal prediction project (S2S) and the prediction of extreme events. *Clim. Atmos. Sci* **2018**, *1*, 3. https://doi.org/10.1038/s41612-018-0013-0.
7. Hà, P.T. Dự báo hạn mùa và nội mùa ngày bắt đầu mùa mưa ở Việt Nam trên cơ sở sản phẩm mô hình số. Luận án Tiến sỹ, Đại học Quốc Gia Hà Nội, Việt Nam, 2022.
8. Hương, H.T.T.; Lượng, N.V.; Vinh, P.V.; Hà, P.T. Nghiên cứu hiệu chỉnh sản phẩm dự báo mưa hạn nội mùa cho khu vực Bắc Trung Bộ bằng phương pháp phân vị với xấp xỉ Gamma (QM–G). *Tạp chí Khí tượng Thủy văn* **2023**, *748*, 77–89. Doi: 10.36335/VNJHM.2023(748).77-89.
9. Trực tuyến: https://dkvbtb.gov.vn/gioi-thieu/gioi-thieu-ve-dai-khi-tuong-thuy-van-khu-vuc-bac-trung-bo-16.html.
10. Trực tuyến: https://confluence.ecmwf.int/display/S2S/Models.
11. Trực tuyến: https://confluence.ecmwf.int/display/FCST/Implementation+of+IFS+Cycle+48r1.
12. Saha, S. et al. The NCEP climate forecast system reanalysis. *Bull. Amer. Meteor. Soc.* **2010**, *91*, 1015–1058. https://doi.org/10.1175/2010BAMS3001.1.
13. George, J.H.; David, T.B. TRMM and other data precipitation data set documentation, 2018, pp. 44.
14. Cort, J.W.; Kenji, M. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Clim. Res.* **2005**, *30*, 79–82.
15. Hyndman, R.J.; Koehler, A.B.. Another look at measures of forecast accuracy. *Int. J. Forecasting* **2006**, 22(4), 679–688.
16. Glenn, W.B. Verification of forecasts expressed in terms of probability. *Mon. Weather Rev.* **1950**, *78(1)*, 1–3.
17. Christopher, A.T.F. Comparing probabilistic forecasting systems with the brier score. *Wea. Forecasting* **2007**, *22*, 1076–1088. https://doi.org/10.1175/WAF1034.1.
18. Marzban, C. The ROC curve and the area under it as performance measures. *Wea. Forecasting* **2004**, *19*, 1106–1114. https://doi.org/10.1175/825.1.

19. Fawcett, T. An introduction to ROC analysis. *Pattern Recognit. Lett.* **2005**, 27(8), 861–874. Doi:10.1016/j.patrec.2005.10.010.
20. Rodríguez–Hernández, M.M.; Rosa, E.P.; Juan, M.R. Statistical analysis of the evolutive effects of language development in the resolution of mathematical problems in primary school education. *Mathematics* **2021**, *9(10)*, 1081. Doi:10.3390/math9101081.

Research and evaluate the applicability of IFS CY48R1 and CFSV1 models for subseasonal forecasting for the North Central region

Dao Anh Cong[1]*, Nguyen Van Luong[1], Phan Van Vinh[1], Phan Như Xuyen[1], Trinh Tuan Long[2]

[1] North Central regional hydro-meteorology center; daoanhcong.k55.hus@gmail.com; luongnvkttv@gmail.com; vinhpv@gmail.com; phannhuxuyen@gmail.com
[2] Ha noi university of science; trinhtuanlong@gmail.com

Abstract: The subseasonal to seasonal forecast, also known as the time span from two weeks to two months, was previously considered the “desert of predictability”, but this is the ideal period for making medium-term plans for water resource management and reservoir regulation. In Vietnam, the research on this problem is still only in the first steps. To fill this forecast gap, the authors conducted a study and evaluated the forecasting skills of the IFS CY48R1 and CFSv1 models for the aforementioned forecast period using TRMM satellite rainfall data and re–forecasting data of ECMWF and CFSV1 models to calculate statistical indicators ME, MAE, RMSE, CORR, Brier score and ROC curve. The results show that both models show the potential for application in subseasonal forecasting for the North Central region. The CFSv1 model has better forecasting skills than IFS in November - March but shows much worse forecasting skills than IFS in the rainy months (July - October). Both models have the potential to correct systematic errors in the months of December - April and August - October in the South of the area. Both models can detect moderate, heavy to very heavy rains, however, the IFS model shows superior ability to CFS when it has a much lower false positive rate.

Keywords: ECMWF; NCEP; IFS CY48R1; CFSv1; Subseasonal to seasonal; S2S; North central region; Forecast.