

NGHIÊN CỨU ỨNG DỤNG PHƯƠNG PHÁP THỐNG KÊ TỰ CẬP NHẬT ĐỂ CẢI TIẾN CHẤT LƯỢNG DỰ BÁO HẠN THÁNG VÀ MÙA CỦA SỐ LIỆU ECMWF

Trần Hồng Thái¹, Võ Văn Hòa²

¹Trung tâm Khí tượng Thủy văn quốc gia

²Đài Khí tượng Thủy văn khu vực Đồng bằng Bắc Bộ

*B*ài báo này sẽ giới thiệu về một số kết quả nghiên cứu ứng dụng một số phương pháp thống kê có khả năng tự cập nhật các trọng số (AMOS) để cải tiến chất lượng dự báo hạn tháng và hạn mùa cho một số yếu tố khí tượng bề mặt từ hệ thống dự báo tổ hợp hạn mùa - VarEPS của Trung tâm Dự báo thời tiết hạn vừa Châu Âu (ECMWF). Các kết quả đánh giá đã cho thấy tất cả các phương pháp AMOS được thử nghiệm đều cho sai số dự báo nhỏ hơn so với dự báo trực tiếp từ VarEPS. Trong đó, các phương pháp lọc Kalman và hiệu chỉnh trung bình trượt cho mức độ cải thiện chất lượng dự báo tốt nhất.

Từ khóa: thống kê tự cập nhật, dự báo hạn dài, số liệu ECMWF

1. Mở đầu

Để cung cấp thêm các nguồn số liệu tham khảo mang tính khách quan và định lượng trong nghiệp vụ dự báo khí tượng hạn vừa và hạn dài (tháng và mùa) tại Trung tâm Dự báo KTTV trung ương (TTDBTU), các sản phẩm dự báo từ hệ thống dự báo tổ hợp toàn cầu (VarEPS) của Trung tâm Dự báo thời tiết hạn vừa Châu Âu (ECMWF) đã được khai thác tại TTDBTU từ cuối năm 2011. Kể từ khi đưa vào khai thác tại TTDBTU và các Đài KTTV khu vực cho đến nay, các sản phẩm dự báo của ECMWF đã được khai thác hiệu quả và góp phần không nhỏ trong việc nâng cao chất lượng dự báo KTTV nói chung và các hiện tượng KTTV nguy hiểm nói riêng, đặc biệt là công tác dự báo hạn vừa và hạn dài. Tuy nhiên, các kết quả đánh giá chất lượng dự báo cho một số yếu tố khí tượng bề mặt như nhiệt độ, độ ẩm, lượng mưa, gió tại các điểm trạm quan trắc khí tượng đã cho thấy sai số hệ thống tương đối rõ trong dự báo của ECMWF [1]. Nguồn sai số này đã hạn chế phần nào chất lượng dự báo hạn vừa và hạn dài của ECMWF cho khu vực Việt Nam.

Theo Buizza và cộng sự (2005) [2], số liệu dự báo của ECMWF luôn tồn tại các sai số hệ thống nội tại như các mô hình toàn cầu khác do sự không hoàn hảo về động lực, vật lý, phương

pháp số và sai số mô tả địa hình. Do vậy, việc tính toán và dự báo một số yếu tố và hiện tượng dựa trên các yếu tố dự báo trực tiếp có thể bao hàm các sai số hệ thống nói trên, dẫn đến chất lượng dự báo không cao. Do đó, việc nghiên cứu ứng dụng các phương pháp thống kê sau mô hình (MOS) để hiệu chỉnh số liệu ECMWF là rất cần thiết và phù hợp, nhất là cho các hạn dự báo hạn dài (tháng và mùa). Tuy nhiên, do hệ thống VarEPS của ECMWF liên tục được cải tiến về động lực, vật lý, phương pháp số và hệ thống đồng hóa số liệu, do đó việc áp dụng các phương pháp thống kê cổ điển (dựa trên tập số liệu luyện để xây dựng quan hệ thống kê và sử dụng quan hệ tìm được cho tương lai) là không phù hợp do không có khả năng nắm bắt được sự thay đổi của sai số hệ thống. Do vậy, trong nghiên cứu này, chúng tôi sẽ sử dụng các phương pháp thống kê có khả năng tự cập nhật các trọng số (AMOS - Adaptive MOS) để khắc phục các nhược điểm của phương pháp thống kê cổ điển.

Các phần tiếp theo của bài báo sẽ trình chi tiết về các phương pháp AMOS được sử dụng cùng với phương pháp xử lý số liệu ECMWF, phương pháp đánh giá và tập số liệu thử nghiệm. Cuối cùng, các kết quả đánh giá dựa trên bộ số liệu thử nghiệm từ 1/1/2012 đến 31/12/2016 được đưa ra. Các kết quả đánh giá chất lượng dự báo

từ các phương pháp AMOS được so sánh trực tiếp với chất lượng dự báo trực tiếp của ECMWF để đánh giá mức độ cải thiện trong chất lượng dự báo.

2. Phương pháp nghiên cứu và tập số liệu thử nghiệm

2.1. Các phương pháp thống kê tự cập nhật trọng số

Như đã biết, có rất nhiều phương pháp MOS từ đơn giản cho đến phức tạp đã được nghiên cứu và thử nghiệm. Do đó, việc lựa chọn được phương pháp MOS phù hợp là hết sức quan trọng. Theo báo cáo hàng năm của ECMWF, hệ thống dự báo tổ hợp hạn tháng và mùa liên tục được cập nhật cả về động lực, vật lý, phương pháp số, đồng hóa số liệu và cải tiến về độ phân giải. Những cải tiến này làm cho bản chất sai số hệ thống của các hệ thống dự báo hạn dài của ECMWF thay đổi liên tục. Do vậy, cách tiếp cận sử dụng các kỹ thuật thống kê cổ điển như phương pháp hồi quy tuyến tính đa biến là không phù hợp. Trong nghiên cứu này, chúng tôi đề xuất sử dụng các kỹ thuật thống kê dạng cập nhật/thích ứng để xây dựng hệ thống MOS cho số liệu dự báo hạn tháng và mùa của ECMWF, gọi tắt là AMOS. Do số liệu dự báo hạn tháng và hạn mùa của ECMWF được cung cấp theo cả 2 dạng: dự báo tất định (là trung bình cộng đơn giản của 51 dự báo thành phần) và dự báo tổ hợp (DBTH) gồm 51 dự báo thành phần. Nên trong nghiên cứu này, chúng tôi đề xuất thử nghiệm 5 phương pháp AMOS cho cả 2 dạng sản phẩm nói trên. Cụ thể:

1. Tính toán dự báo trung bình tổ hợp (TBTH) dựa trên phương pháp EMES (trung bình tổ hợp có trọng số giảm theo hàm mũ) và EMMV (trung bình tổ hợp có trọng số tỷ lệ nghịch với phương sai sai số)

2. Hiệu chỉnh từng dự báo thành phần trước khi tính toán TBTH là trung bình cộng đơn giản: phương pháp BCMA (hiệu chỉnh bằng trung bình trượt, bias là chỉ là sai số trung bình) và BCES (thêm trọng số giảm theo hàm mũ trong tính toán sai số trung bình)

3. Hiệu chỉnh dự báo TBTH trực tiếp từ

ECMWF (gọi là Raw) thông qua lọc Kalman.

Trước hết, giả thiết có một DBTH bao gồm N dự báo thành phần, khi đó dự báo TBTH có trọng số như nhau có thể được tính thông qua công thức trung bình cộng đơn giản (1) dưới đây:

$$TBTH = \frac{1}{N} \sum_{i=1}^N F_i \quad (1)$$

Trong đó F_i là dự báo thành phần thứ i ($i = 1, N$). Trong công thức (1), trọng số cho mỗi dự báo thành phần là như nhau và bằng $1/N$. Cách tính này có hạn chế là không tính đến được khả năng đóng góp của từng dự báo thành phần tới chất lượng dự báo TBTH. Như đã biết, mỗi dự báo thành phần có chất lượng dự báo khác nhau và chất lượng này liên tục thay đổi theo các phiên dự báo. Do đó, đưa được yếu tố này vào trong công thức (1) sẽ cải thiện được chất lượng dự báo TBTH. Dựa trên giả thiết và luận giải nói trên, phần dưới đây sẽ trình bày chi tiết về mặt toán học của 5 phương pháp AMOS nói trên.

a) *Trung bình có trọng số giảm dần theo hàm mũ (EMES):*

Phương pháp này được Daley (1991) [3] đề xuất trong đó TBTH được tính theo công thức (2) dưới đây:

$$TBTH = \sum_{i=1}^N w_i F_i^{nobias} \quad (2)$$

với F_i^{nobias} là dự báo thành phần thứ i ($i=1, N$) nhưng đã được hiệu chỉnh sai số hệ thống. Các trọng số w_i sẽ được tính theo công thức (3) và (4) dưới đây:

$$w_i = \beta \alpha^{i-1} \quad (3)$$

$$\beta = \frac{1 - \alpha}{1 - \alpha^n} \quad (4)$$

với α là nhân tố làm trơn, giá trị j trong công thức (4) là hạng của dự báo thành phần thứ i (F_i^{nobias}) được tính dựa trên sai số bình phương trung bình (MSE). Cụ thể, từ các giá trị MSE tìm được của từng dự báo thành phần dựa trên tập số liệu phụ thuộc cho trước, tiến hành sắp xếp theo chuỗi trình tự tăng dần. Dự báo thành phần nào có MSE bé nhất sẽ có hạng là 1 và cứ thế tiếp

tục. Nếu các dự báo thành phần có MSE bằng nhau thì sẽ có cùng hạng với nhau. Trong phương pháp này, giá trị MSE được tính toán cho từng dự báo thành phần đã được hiệu chỉnh sai số hệ thống. Theo Yossouf và Stensrud (2006) [4], nhân tố dao động trong khoảng $[0,1 - 0,9]$ và cần thiết phải lựa chọn tối ưu cho từng yếu tố khí tượng khác nhau. Tuy nhiên, rất nhiều nghiên cứu đã chỉ ra giá trị 0,85 là tối ưu cho hầu hết các yếu tố khí tượng liên tục trong đó có nhiệt độ. Theo công thức (3), trọng số sẽ giảm dần theo đường cong hàm mũ khi j tăng lên.

b) *Trung bình có trọng số tính theo phương sai sai số (EMMV):*

Tương tự phương pháp EMES, phương pháp EMMV cũng được Daley (1991) đề xuất trong đó TBTH được tính theo công thức (2) ở trên và các trọng số w_i được tính theo công thức (5) dưới đây với j là các chỉ số chạy theo tổng số dự báo thành phần ($j=1,N$).

$$w_i = \frac{(1/MSE_i)}{\sum_{j=1}^N (1/MSE_j)} \quad (5)$$

c) *Hiệu chỉnh sai số hệ thống bằng trung bình trượt (BCMA):*

Theo phương pháp này, việc hiệu chỉnh cho một dự báo thành phần bất kỳ của DBTH đưa ra được thực hiện theo công thức (6) dưới đây:

$$F_{BCMA} = F_{raw} - \text{bias} \quad (6)$$

trong đó F_{bcma} là dự báo đã được hiệu chỉnh bằng phương pháp BCMA, F_{raw} là dự báo trực tiếp từ DBTH và bias là sai số hệ thống được xác định đơn giản bằng công thức (7):

$$\text{bias} = \frac{1}{n} \sum_{i=1}^n (F_{raw,i} - O_i) \quad (7)$$

với n là tổng số ngày có dữ liệu trước đó, O_i và $F_{raw,i}$ tương ứng là giá trị quan trắc và dự báo thành phần cần hiệu chỉnh của ngày thứ i .

d) *Hiệu chỉnh sai số hệ thống với trọng số theo hàm mũ (BCES):*

Cách hiệu chỉnh theo BCES là tương tự như BCMA nhưng khác ở trong cách tính sai số hệ thống (bias). Cụ thể, bias sẽ được tính như công

thức (8) dưới đây:

$$\text{bias} = \sum_{i=1}^n w_i (F_i - O_i) \quad (8)$$

với n là tổng số ngày có dữ liệu trước đó. Trọng số w_i sẽ giảm dần theo hàm mũ như trong công thức (9):

$$w_i = \beta \alpha^{i-1} \quad (9)$$

$$\beta = \frac{1 - \alpha}{1 - \alpha^N} \quad (10)$$

với i là số thứ tự ngày ($i=1$ cho ngày trước ngày hiện tại và bằng x cho x ngày trước ngày hiện tại). Nhân tố làm trơn $\alpha = 0,85$ như đã trình bày trong phương pháp EMES ở trên. Trọng số w_i sẽ giảm dần theo đường cong hàm mũ khi i tăng lên.

e) *Hiệu chỉnh sai số hệ thống bằng lọc Kalman (KF):*

Trước hết, giả thiết có một quan hệ tuyến tính giữa quan trắc O và dự báo F như sau: $O = a_0 + a_1 F$ với a_0, a_1 là các hệ số được xác định bằng phương pháp hồi quy tuyến tính (BCLR) dựa trên tập số liệu của n ngày trước đó. Khi quan hệ này đã được thiết lập, nó sẽ được áp dụng cho dự báo của ngày kế tiếp. Với giả thiết này, dự báo hiệu chỉnh theo phương pháp hồi quy tuyến tính sẽ được tính theo công thức (11) dưới đây:

$$F_{BCLR} = a_0 + a_1 F_{raw} \quad (11)$$

Để khắc phục nhược điểm của phương pháp hồi quy tuyến tính đơn biến, đó là các trọng số a_0 và a_1 không có khả năng tự cập nhật trong chu kỳ luyện khi có sự đột biến về mặt hình thể thời tiết. Phương pháp lọc Kalman thường được sử dụng để cập nhật các trọng số này trong suốt chu kỳ luyện cho tới thời điểm bắt đầu thử nghiệm dự báo. Quá trình cập nhật các trọng số này được thực hiện thông qua phương trình dự báo và quan trắc của Lọc Kalman như công thức (12) và (13) dưới đây:

$$x_k = A \times x_{k-1} + B \times u_{k-1} + w_{k-1} \quad (12)$$

$$z_k = H \times x_k + v_k \quad (13)$$

trong đó x_k là vector trạng thái thời điểm hiện

tại, x_{k-1} là vector trạng thái thời điểm trước đó, u_{k-1} là vector điều khiển hệ thống tại thời điểm trước đó, w_{k-1} là véc tơ đặc trưng cho độ bất định của hệ thống do quy luật tuyến tính mô tả thông qua ma trận A. Ma trận A kích thước $N \times N$ có tên gọi ma trận chuyển dịch trạng thái (mang thông tin về quy luật hoạt động của hệ thống). Ma trận B được đưa vào phương trình trên đặc trưng cho các quá trình bên ngoài điều khiển hệ thống đang xét tác động thông qua véc tơ điều khiển u_{k-1} . Véc tơ v_k đặc trưng cho sai số hay nhiễu khi thực hiện đo, H là ma trận kích thước $M \times N$ mô tả quan hệ được giả định tuyến tính giữa biến đo được với biến trạng thái cần xác định. H được gọi là ma trận quan trắc. Với tập số liệu luyện đưa ra, phương pháp lọc Kalman sẽ sử dụng thuật toán đệ quy để tìm ra các trọng số tối ưu theo từng thay đổi trong sai số dự báo. Các giá trị a_0 và a_1 được tìm ra bởi phương pháp BCLR sẽ được sử dụng như là giá trị phỏng đoán ban đầu cho bài toán đệ quy để đảm bảo nghiệm hội tụ nhanh và ổn định. Giá trị a_0 và a_1 được tính ra từ lọc Kalman của ngày trước ngày bắt đầu dự báo sẽ được sử dụng lại trong phương trình (11).

Cũng giống như các phương pháp thống kê truyền thống khác, 5 phương pháp AMOS thử nghiệm ở trên cũng cần tập số liệu phụ thuộc (training dataset) để ước lượng. Mỗi một phương pháp AMOS khác nhau sẽ đòi hỏi độ dài của chuỗi số liệu phụ thuộc khác nhau, và cũng tùy thuộc vào từng biến khí quyển được dự báo. Đối với bài toán dự báo thời tiết, do tính biến động lớn nên dung lượng mẫu sử dụng dao động trong khoảng từ 20 - 70 ngày tùy thuộc vào đối tượng nghiên cứu và hạn dự báo. Tuy nhiên, đối với bài toán dự báo hạn tháng và mùa, tính quán tính giữa các lần dự báo liên tiếp khác nhau là tương đối cao do ít khi có sự thay đổi đột biến. Do đó, dung lượng mẫu có thể cần ít hơn so với bài toán dự báo thời tiết. Trong nghiên cứu này, chúng tôi sử dụng cách tiếp cận thực nghiệm để tìm ra lựa chọn độ dài tối ưu cho bộ số liệu phụ thuộc. Cụ thể, các giá trị độ dài chuỗi số liệu phụ thuộc được đưa ra để khảo sát sự biến thiên của sai số quân phương (RMSE) cho từng phương pháp

AMOS. Dung lượng mẫu được thử nghiệm từ 5 - 20 lần dự báo trước đó với khoảng cách tăng là 1 (16 giá trị). Các kết quả khảo sát dựa trên chỉ số RMSE cho thấy chỉ cần 7 lần dự báo trước đó là đủ tạo ra bộ số liệu phụ thuộc tối ưu cho các phương pháp AMOS ở trên. Việc tăng dung lượng mẫu lên không tạo ra bất kỳ sự giảm giá trị RMSE đáng kể nào. Do vậy, trong nghiên cứu này chúng tôi lựa chọn dung lượng mẫu cho tập số liệu phụ thuộc cho 5 phương pháp AMOS được thử nghiệm là 7. Tức là sử dụng 7 dự báo trước đó gần với dự báo hiện tại.

2.2. Mô tả tập số liệu nghiên cứu

Để đánh giá được chất lượng dự báo hạn dài (tháng, mùa) của số liệu ECMWF cho khu vực Việt Nam, đồng thời xây dựng và đánh giá được chất lượng dự báo của hệ thống AMOS, cần thiết phải thu thập các nguồn số liệu như sau:

- Số liệu quan trắc nhiệt độ trung bình ngày (T2m), nhiệt độ điểm sương tại độ cao 2 mét (Td2m), nhiệt độ tối cao ngày (Tmax), nhiệt độ tối thấp ngày (Tmin), tốc độ gió tại độ cao 10 mét (ff10m) của 171 trạm quan trắc khí tượng bề mặt hiện tại từ 1/1/2012 đến 31/10/2016. Sở dĩ phải thu thập số liệu quan trắc ngày là do số liệu dự báo hạn tháng của ECMWF được tính trung bình theo tuần, không phải là trung bình 10 ngày như đang sử dụng ở TTDBTU.

- Số liệu dự báo hạn tháng và hạn mùa của ECMWF cho giai đoạn từ 1/1/2012 - 31/10/2016 cho các biến như số liệu quan trắc.

2.3. Mô tả phương pháp xử lý số liệu và đánh giá

Do mục tiêu của nghiên cứu là đánh giá được khả năng của các phương pháp AMOS trong việc cải thiện chất lượng dự báo hạn tháng và mùa của ECMWF, nên không gian nghiên cứu sẽ thực hiện trên không gian điểm trạm, trong khi số liệu dự báo của ECMWF được cung cấp trên lưới. Do đó, việc nội suy dữ liệu trên lưới về điểm trạm là cần thiết và phương pháp nội suy được sử dụng sẽ tác động đến chất lượng dự báo. Trên thực tế, có rất nhiều các phương pháp nội suy theo không gian được phát triển cho mục đích nội suy dữ liệu trên lưới về điểm bất kỳ bên trong

lưới như phương pháp nội suy song tuyến tính, nội suy spline, ... Trong nghiên cứu này, chúng tôi sử dụng các phương pháp như sau:

- Đối với các yếu tố bề mặt T2m, Td2m, Tmax và Tmin: dùng phương pháp nội suy song tuyến tính. Phương pháp này sử dụng ô lưới 2 x 2 điểm nút lưới để nội suy về điểm trạm nằm bên trong theo nguyên tắc trung bình có trọng số theo khoảng cách

- Đối với yếu tố ff10m: dùng phương pháp nội suy điểm gần nhất. Theo phương pháp này, từ vị trí của điểm cần nội suy, thuật toán sẽ tính toán khoảng cách của điểm nút lưới gần nhất và sử dụng giá trị tại nút lưới này để gán cho điểm nội suy.

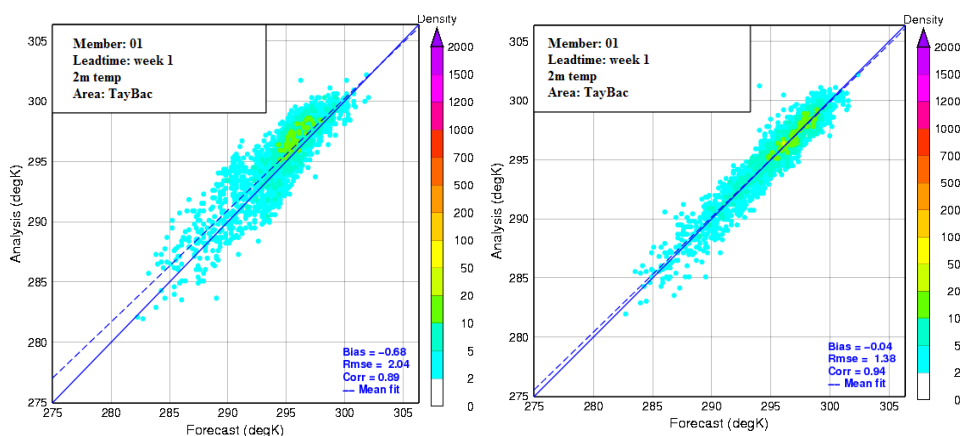
Để hạn chế ảnh hưởng của hiệu ứng làm trơn gradient dọc theo bờ biển, mặt nạ đất/biển được sử dụng để xác định xem các nút lưới được lựa chọn là nằm trên đất liền hay trên biển. Việc sử dụng sai nút lưới để nội suy (nhất là trong phương pháp nội suy điểm gần nhất) có thể dẫn tới sai số lớn. Ví dụ, nếu điểm trạm nằm trên đất liền, trong khi điểm nút lưới gần nhất nằm trên biển có thể gây ra những sai số trong dự báo nhiệt độ, độ ẩm và gió.

Không gian đánh giá được thực hiện tại từng điểm trạm, cụ thể là tại 171 trạm quan trắc khí tượng bề mặt. Để tính giá trị sai số đặc trưng cho từng vùng, thay vì lấy trung bình giá trị đánh giá của các điểm trạm trong vùng đó, đề tài tiếp cận theo cách gộp số liệu của các trạm trong vùng đó

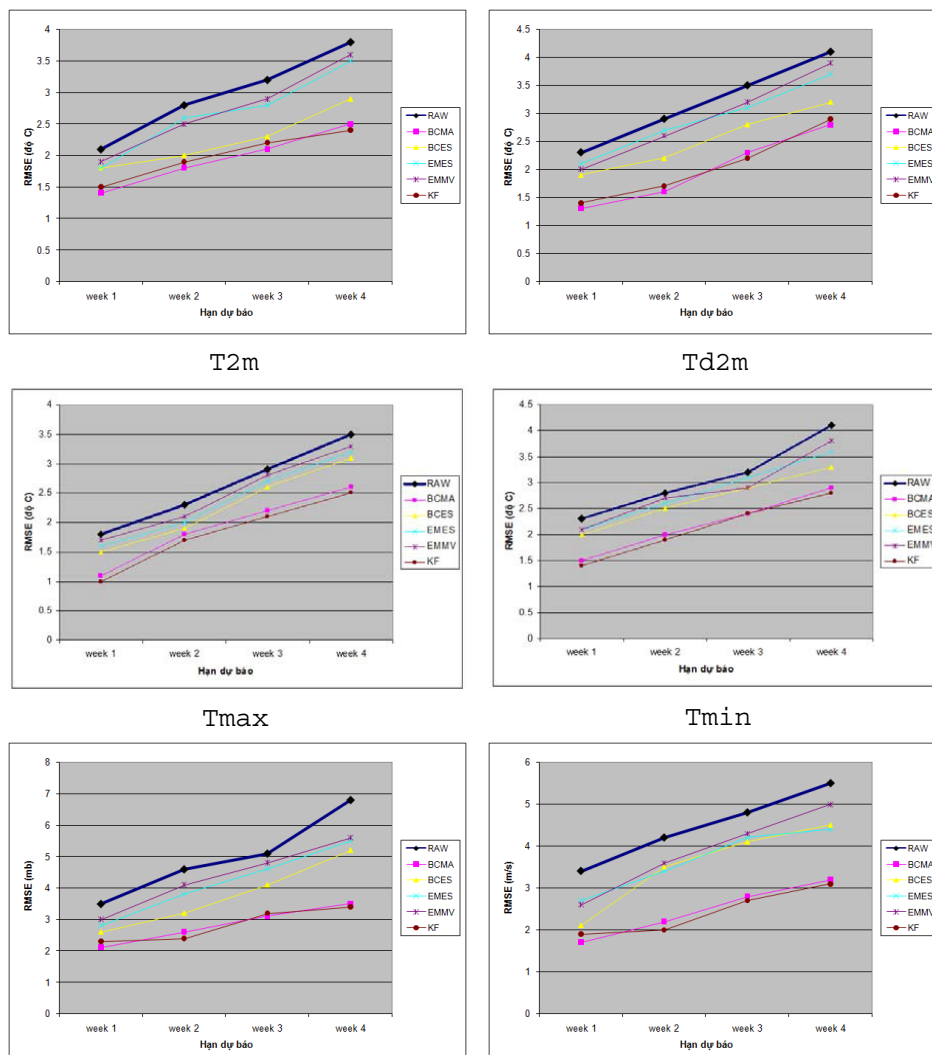
thành 1 chuỗi số liệu để đánh giá. Trong nghiên cứu này, chúng tôi chỉ sử dụng chỉ số sai số quân phương (RMSE). Dự báo đối chứng để kiểm nghiệm khả năng cải thiện chất lượng dự báo của 5 phương pháp AMOS ở trên chính là dự báo TBTH đơn giản (ký hiệu là Raw) của ECMWF.

3. Một số kết quả thử nghiệm và đánh giá

Như đã trình bày ở trên, 5 phương pháp AMOS (MOS thích ứng/tự cập nhật) sẽ được thử nghiệm để cải thiện chất lượng dự báo của một số yếu tố bề mặt có tính liên tục gồm T2m, Td2m, Tmax, Tmin, pmsl và ff10m. Trong 5 phương pháp AMOS được thử nghiệm, có 3 phương pháp là BCMA, BCES và KF hiệu chỉnh trực tiếp các dự báo thành phần trước khi lấy trung bình tổ hợp. Hình 1 đưa ra giản đồ tụ điểm của Raw và BCMA cho dự báo T2m hạn tháng tại khu vực Tây Bắc đối với dự báo thành phần 1 của ECMWF (hệ thống dự báo tổ hợp hạn tháng của ECMWF gồm 51 dự báo thành phần). Từ hình 1 có thể thấy trước khi hiệu chỉnh, dự báo thành phần 1 trực tiếp từ ECMWF có xu thế cho dự báo T2m thiên thấp với giá trị sai số hệ thống là -0,68 và sai số quân phương 2,04°C. Sau khi áp dụng phương pháp BCMA để hiệu chỉnh, sai số hệ thống và sai số quân phương đã giảm đi đáng kể, chỉ còn là -0,04 và 1,38°C. Các kết quả tương tự cũng được tìm thấy cho các phương pháp AMOS khác nhưng mức độ loại bỏ được sai số hệ thống là khác nhau.



Hình 1. Giản đồ tụ điểm cho dự báo T2m trung bình tuần thứ 1 tại khu vực Tây Bắc từ Raw (bên trái) và BCMA (bên phải)



Hình 2. Kết quả tính toán chỉ số RMSE trung bình trên toàn bộ Việt Nam cho dự báo hạn tháng từ Raw và 5 phương pháp AMOS

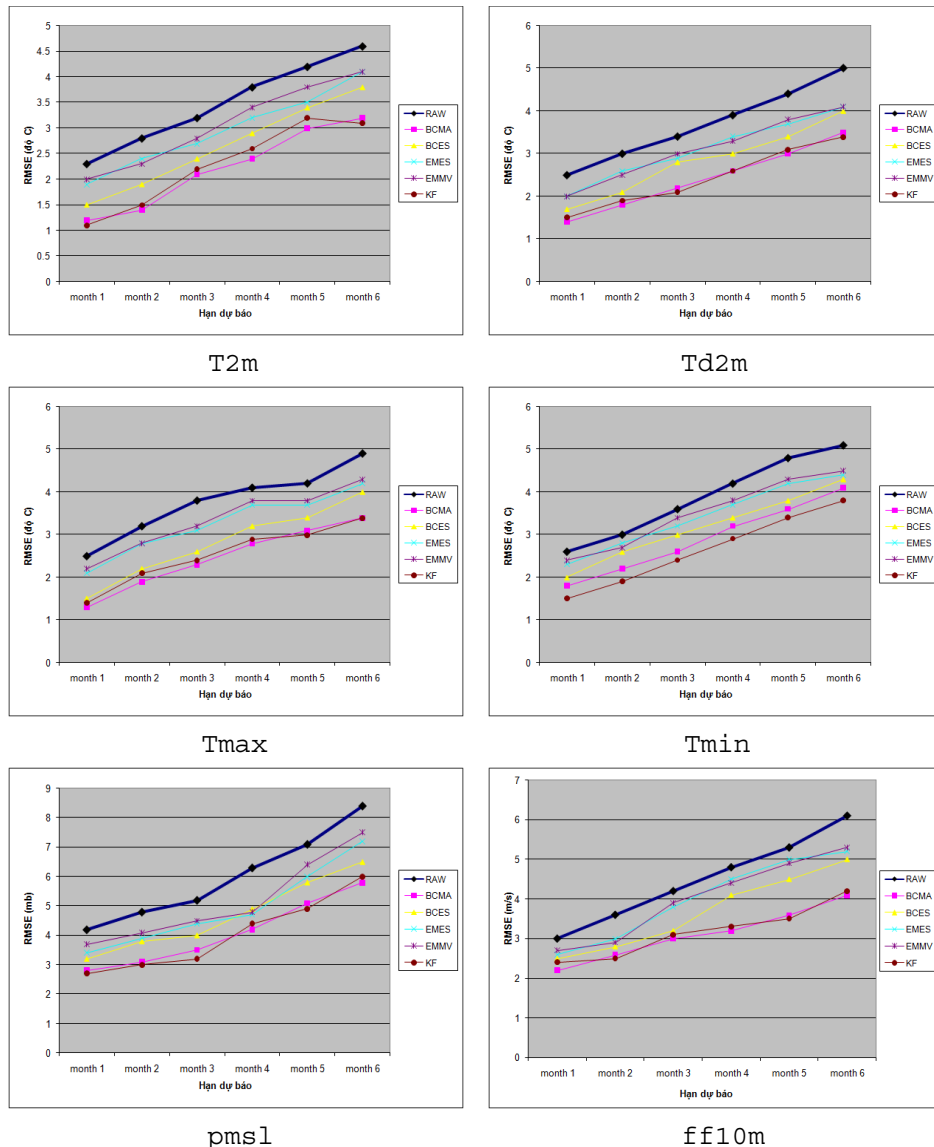
Kết quả minh họa ở trên đã cho thấy tính hiệu quả của các phương pháp AMOS trong việc loại bỏ sai số hệ thống trong các dự báo thành phần Raw, qua đó cải thiện được chất lượng dự báo của dự báo trung bình tổng hợp cuối cùng. Câu hỏi đặt ra là giữa cách tiếp cận tính trung bình có trọng số với hiệu chỉnh sai số hệ thống trong các dự báo thành phần, cách tiếp cận nào sẽ hiệu quả hơn trong việc giảm sai số dự báo ?.

Hình 2 đưa ra kết quả tính toán chỉ số RMSE cho dự báo T2m, Td2m, Tmax, Tmin, pmsl và ff10m trung bình các tuần dự báo của số liệu dự báo hạn tháng của ECMWF (Raw) và 5 phương pháp AMOS được thử nghiệm. Chỉ số RMSE này được tính trung bình trên 171 điểm trạm được nghiên cứu. Từ hình 2 có thể nhận thấy tất

cả các phương pháp AMOS được thử nghiệm cho sai số RMSE nhỏ hơn so với Raw tại các tuần được dự báo. Hay nói cách khác, sai số dự báo đã được giảm sau khi áp dụng AMOS và chất lượng dự báo đã được cải thiện. Trong 5 phương pháp AMOS được thử nghiệm, các phương pháp BCMA và KF cho chất lượng dự báo tốt nhất. Ngoài ra, nếu so sánh giữa 2 nhóm phương pháp, dễ dàng nhận thấy nhóm phương pháp hiệu chỉnh các dự báo thành phần trước khi lấy trung bình tổng hợp có sai số nhỏ hơn so với nhóm phương pháp tính trung bình tổng hợp có trọng số. Các kết quả đánh giá cũng cho thấy mức độ giảm sai số khi áp dụng các phương pháp AMOS tại các hạn dự báo lớn cao hơn so với các hạn dự báo gần (cụ thể sai số dự báo cho tuần 4

giảm được nhiều hơn so với sai số dự báo của tuần 1 khi áp dụng AMOS). Nhìn chung, mức độ cải thiện chất lượng dự báo của Tmax và Tmin là cao hơn so với T2m và Td2m do sai số hệ thống

trong dự báo hạn tháng của Tmax và Tmin có biên độ lớn so với T2m và Td2m. Các kết quả tương tự cũng được tìm thấy khi áp dụng cho số liệu dự báo hạn mùa của ECMWF (Hình 3).



Hình 3. Tương tự hình 2 nhưng cho dự báo hạn mùa

4. Kết luận

Bài báo này đã giới thiệu về cơ sở lý thuyết của 5 phương pháp AMOS được nghiên cứu nhằm hiệu chỉnh số liệu dự báo hạn tháng và mùa của ECMWF. Các kết quả đánh giá đã cho thấy tất cả các phương pháp AMOS được thử nghiệm đều làm giảm đáng kể sai số dự báo, qua đó nâng cao được chất lượng dự báo hạn tháng và hạn mùa cho số liệu ECMWF. Trong đó,

phương pháp BCMA và KF cho thấy sự cải thiện tốt nhất và hoàn toàn có thể triển khai vào nghiệp vụ. Hiện tại, các sản phẩm dự báo từ các phương pháp BCMA và KF đã được đưa vào thử nghiệm nghiệp vụ tại TTDBTU. Trong nghiên cứu tiếp theo, chúng tôi sẽ mở rộng thêm cho các yếu tố như mưa và các hiện tượng khí hậu cực đoan như dự báo số ngày mưa lớn.

Lời cảm ơn: Bài báo này được hoàn thành dựa trên sự hỗ trợ từ Đề tài NCKH cấp Bộ mang mã số TNMT.05.44 và Đề tài NCKH cấp Nhà nước “Nghiên cứu tác động của biến đổi khí hậu tới sự xâm nhập của các đợt lạnh và nóng ẩm bất thường trong mùa đông ở khu vực miền núi phía Bắc phục vụ phát triển kinh tế - xã hội” thuộc chương trình BĐKH/16-20.

Tài liệu tham khảo

1. Võ Văn Hòa và nnk (2016), *Nghiên cứu ứng dụng số liệu dự báo của Trung tâm Dự báo thời tiết hạn vừa Châu Âu để nâng cao chất lượng dự báo hạn tháng và hạn mùa cho khu vực Việt Nam*, Báo cáo tổng kết đề tài NCKH cấp Bộ, 129tr.
2. Buizza R., P. L. Houtekamer, Z. Toth, G. Pellerin, M. Wei, Y. Zhu (2005), *A comparison of the ECMWF, MSC, and NCEP global ensemble prediction systems*, Mon. Wea. Rev., 133, 1076-1097.
3. Daley R. (1991), *Atmospheric Data Analysis*, Cambridge University Press, 457p.
4. Yussouf, N. and D. J. Stensrud (2006), *Prediction of near-surface variables at independent locations from a bias-corrected ensemble forecasting system*, Mon. Rev. Rev., 134, 3415-3424.

APPLICATION OF ADAPTIVE MOS METHODS TO IMPROVE SEASONAL AND MONTHLY PREDICTION OF ECMWF

Tran Hong Thai¹, Vo Van Hoa²

¹National Hydro-Meteorological Services

²Regional hydro-meteorological Center for Northern Delta

Abstract: The paper give out some results in application of adaptive model output statistics methods (AMOS) in order to improve monthly and seasonal forecast of global long-term ensemble prediction system (VarEPS) of ECMWF. The verification pointed out the monthly and seasonal forecast quality of ECMWF at 171 surface meteorological observation stations had been significantly improved after applying AMOS to remove the bias. In which, the BCMA and Kalman filter methods is the best.

Key words: Adaptive MOS, monthly and seasonal forecast, ECMWF.