

## NGHIÊN CỨU ỨNG DỤNG LỌC KALMAN HIỆU CHỈNH MỘT SỐ TRƯỜNG DỰ BÁO BỀ MẶT TỪ MÔ HÌNH HRM.

Phần II. Một số kết quả nghiên cứu ban đầu

ThS. Võ Văn Hòa, CN. Nguyễn Thu Hằng, ThS. Nguyễn Chi Mai  
Trung tâm Dự báo Khí tượng Thủy văn Trung ương

**P**hương pháp lọc Kalman được nhà toán học nổi tiếng Kalman đề xuất từ năm 1960 và trở thành một công cụ thống kê hữu ích trong bài toán hiệu chỉnh sau mô hình. Phần I của bài báo [6] đã trình bày cơ sở lý thuyết của lọc Kalman dựa trên lý thuyết xác suất thống kê và đặc trưng biến ngẫu nhiên. Phần II này sẽ trình bày các kết quả nghiên cứu ứng dụng lọc Kalman hiệu chỉnh các kết quả dự báo trực tiếp từ mô hình HRM [1] cho một số biến bề mặt tại 25 trạm synôp dựa trên chuỗi số liệu 7 tháng từ 01/07/2005 cho đến 31/01/2006. Các kết quả đánh giá đã cho thấy sự cải thiện đáng kể trong chất lượng dự báo khi hiệu chỉnh các dự báo trực tiếp từ mô hình HRM bằng lọc Kalman.

### 1. Đặt vấn đề

Trong 20 năm trở lại đây, dự báo thời tiết bằng phương pháp số đã đạt được nhiều thành tựu khả quan và góp phần nâng cao chất lượng dự báo. Tuy nhiên do bản chất rối của khí quyển cũng như tính phi tuyến trong các bài toán khí tượng mà các kết quả dự báo định lượng trực tiếp từ mô hình (DMO-Direct Model Output) luôn chứa một sai số nào đó, đặc biệt là dự báo cho các biến khí tượng bề mặt. Để cải thiện DMO của các mô hình, các phương pháp hiệu chỉnh sau mô hình đã được phát triển và đưa vào ứng dụng trong dự báo nghiệp vụ như phương pháp thống kê sau mô hình MOS (Model Output Statistics) [7], [11], dự báo hoàn hảo PP (Perfect Prognostic) [7], GAM (Generalized Additive Models) [14], mạng thần kinh nhân tạo ANN (Artificial Neural Network) [9] và lọc Kalman [8],[10]. Về cơ bản, các phương pháp MOS và PP là hồi quy tuyến tính đa biến đơn giản. Tuy nhiên, MOS và PP đòi hỏi một cơ sở dữ liệu lớn khi tính toán các hệ số hồi quy và chỉ có hiệu quả khi các đặc trưng sai số là không thay

đổi hoặc có thay đổi một cách có hệ thống. Với sự phát triển mạnh mẽ của khoa học máy tính, các phương pháp phi tuyến như GAM và ANN đã được phát triển và ứng dụng rộng rãi. Nhược điểm lớn nhất của hai phương pháp phi tuyến này là đòi hỏi số liệu lớn và phức tạp trong tính toán.

Lọc Kalman được đề xuất bởi nhà toán học Kalman vào năm 1960 và đã được ứng dụng khá rộng rãi trong các bài toán không gian như hiệu chỉnh quỹ đạo của tàu vũ trụ, hiệu chỉnh các kết quả dự báo khí tượng thủy văn. Phương pháp lọc Kalman thực chất là một kỹ thuật lọc tối ưu với các hệ số hiệu chỉnh được cập nhật liên tục theo thời gian dựa trên việc phân tích bản chất sai số dự báo của một vài chu kỳ dự báo trước. Ngoài ra, lọc Kalman không đòi hỏi nhiều về chi phí tính toán và bộ số liệu phụ thuộc. Tại nước ta, lọc Kalman đã được nghiên cứu cho bài toán dự báo dòng chảy thủy văn từ năm 1990 [2], và gần đây là cho bài toán hiệu chỉnh các sản phẩm dự báo từ mô hình dự báo thời tiết số và các dự báo quỹ đạo bão từ các trung tâm dự báo quốc tế [3],[5]. Các kết quả

Người phản biện: ThS. Đỗ Lê Thủy

nghiên cứu lọc Kalman từ các công trình này đã cho thấy tính ưu việt của lọc Kalman trong việc loại bỏ các sai số hệ thống và tính khả thi trong bài toán dự báo nghiệp vụ.

Tại Việt Nam, dự báo thời tiết số bắt đầu được nghiên cứu và ứng dụng từ năm 2000 và bước đầu thu được nhiều kết quả khả quan. Các mô hình dự báo số khu vực lãnh thổ hạn chế đã và đang được nghiên cứu phải kể đến như mô hình HRM [1], MM5, ETA, RAM và gần đây là WRF. Trong đó, mô hình khu vực phân giải cao HRM đã được chạy nghiệp vụ tại Trung tâm Dự báo Khí tượng thủy văn Trung ương (TT DBKTTV TŨ) từ năm 2002 cho đến nay. Hiện tại, mô hình HRM thực hiện dự báo thời hạn 48 giờ cho hai khu vực: khu vực lớn  $5^{\circ}\text{S}-35^{\circ}\text{N}$ ,  $80^{\circ}-130^{\circ}\text{E}$  với độ phân giải ngang là 28km và khu vực nhỏ:  $7.125^{\circ}-27.125^{\circ}\text{N}$ ,  $97.25^{\circ}-117.25^{\circ}\text{E}$  với độ phân giải ngang là 14 km. HRM sử dụng các trường phân tích và dự báo từ mô hình toàn cầu GME. Mô hình HRM với mã nguồn mở là mô hình số đầu tiên được đưa vào nghiệp vụ tại TT DBKTTV TŨ, góp phần quan trọng trong công tác dự báo tác nghiệp. Tuy nhiên, do những hạn chế trong động lực, vật lý, tham số hóa và ban đầu hóa của mô hình HRM, đặc biệt là những thiếu sót trong các trường phân tích bề mặt và ẩm của mô hình toàn cầu GME, nên chất lượng dự báo của mô hình HRM vẫn còn nhiều hạn chế, đặc biệt là trong dự báo các biến khí tượng bề mặt. Để khắc phục những hạn chế này, nhóm nghiên cứu đề xuất thử nghiệm ứng dụng lọc Kalman để hiệu chỉnh các dự báo trực tiếp cho các biến bề mặt nhằm mục đích cải tiến chất lượng dự báo từ mô hình HRM. Các kết quả khảo sát sai số dự báo trực tiếp từ HRM và đánh giá kỹ năng cải thiện chất lượng dự báo sau khi áp dụng lọc Kalman cho các hạn dự báo và 25 trạm nghiên cứu sẽ được lần lượt trình bày trong các mục tiếp theo.

## 2. Mô tả tập số liệu nghiên cứu

Trong nghiên cứu này, tác giả sử dụng mô hình khu vực phân giải cao HRM

(High resolution Regional Model) cho khu vực lớn. Về mặt động lực, phiên bản HRM sử dụng ở đây là mô hình thủy tĩnh với đầy đủ tham số hóa đối lưu, vi vật lý mây, bức xạ và lớp biên. Để nội suy kết quả dự báo trên lưới về tọa độ trạm, chúng tôi sử dụng phương pháp đơn giản nhất và sẵn có trong mô hình HRM, coi giá trị dự báo từ mô hình tại một trạm cần quan tâm là giá trị dự báo của nút lưới gần trạm đó nhất. Tập số liệu sử dụng trong nghiên cứu này bao gồm các thông tin dự báo của các biến: nhiệt độ không khí tại độ cao 2m (T2m), nhiệt độ điểm sương tại độ cao 2m (Td2m) và tốc độ gió tại độ cao 10m (SP10m) cho các hạn dự báo +12h, +24h, +36h và +48h và các giá trị quan trắc tương ứng tại 25 trạm synop phát báo quốc tế (xem hình 2) cho chuỗi thời gian từ 01/07/2005 cho đến 31/01/2006 (chỉ lấy phiên dự báo 00GMT). Trong nghiên cứu này, chúng tôi chưa áp dụng cho biến hướng gió bởi vì đại lượng này mang tính ngẫu nhiên cao và không ổn định. Để đánh giá bản chất sai số của mô hình HRM và hiệu quả của việc áp dụng lọc Kalman theo mùa đông và mùa hè, 2 bộ số liệu con được tách ra từ bộ số liệu trên gồm: bộ số liệu từ 01/07/2005 cho đến 30/09/2005 và từ 01/10/2005 cho đến 31/01/2006. Do lọc Kalman cần một bộ số liệu phụ thuộc (training dataset) để ước lượng, nên các trường dự báo và quan trắc được lấy lùi lại 7 ngày tương ứng cho mỗi chuỗi số liệu con (từ ngày 23/06/2005 cho chuỗi số liệu mùa hè và từ 23/09/2005 cho chuỗi số liệu mùa đông). Các đặc trưng thống kê được sử dụng để đánh giá bao gồm: sai số trung bình ME (Mean Error hay còn gọi là BIAS), sai số tuyệt đối trung bình MAE (Mean Absolute Error) và căn bậc hai của sai số bình phương trung bình RMSE (Root of Mean Square Error). Công thức tính các đại lượng này như sau:

$$ME = \frac{1}{N} \sum_{i=1}^N (F_i - O_i) \quad (1)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |F_i - O_i| \quad (2)$$



TT	Tên trạm	Mã Trạm	Vĩ độ	Kinh độ	Độ cao thực tế (m)	Độ cao trong mô hình (m)
11	Đồng Hới	48848	17.48	106.6	8	133.49
12	Huế	48852	16.43	107.58	9	28.86
13	Đà Nẵng	48855	16.03	108.2	7	14.39
14	Plêiku	48866	13.98	108.0	779	729.1
15	Quy Nhơn	48870	13.77	109.22	6	60.17
16	Nha Trang	48877	12.22	109.22	4	0.01
17	Sông Tử Tây	48892	11.42	114.33	5	0.01
18	Phan Thiết	48887	10.93	108.1	5	106.87
19	Hồ Chí Minh	48900	10.82	106.67	10	7.25
20	Phủ Quốc	48917	10.22	103.97	4	48.75
21	Thổ Chu	48916	9.28	103.47	0	0.92
22	Cà Mau	48914	9.18	105.15	3	1.03
23	Côn sơn	48918	8.68	106.6	9	8.67
24	Trường Sa	48920	8.65	111.92	3	0.01
25	Huyền Trân	48919	8.02	110.62	19	0.01

### 3. Các đặc trưng sai số trong dự báo các trường bề mặt của mô hình HRM

Các kết quả khảo sát sai số dựa trên 3 đại lượng thống kê là BIAS, MAE và RMSE trung bình trên 25 synôp nghiên cứu cho 3 biến khí tượng: T2m, Td2m và SP10m tại các hạn dự báo +12h (F12h), +24h (F24h), +36h (F36h) và +48h (F48h) tương ứng cho mùa hè và mùa đông được minh họa trong các hình vẽ từ hình vẽ 2 cho đến hình vẽ 4. Từ hình 2 có thể thấy mô hình HRM có BIAS cho dự báo T2m và Td2m thấp hơn so với thực tế trong cả mùa hè và mùa đông. Ngược lại, dự báo SP10m lại cho cao hơn so với thực tế. BIAS của mỗi biến dự báo là cùng dấu cho tất cả các hạn dự báo. Giá trị BIAS lớn tìm thấy trong dự báo SP10m và khá nhỏ đối với dự báo nhiệt độ T2m. Từ các chỉ số thống kê MAE và RMSE, có thể nhận thấy sai số dự báo cho cả 3 yếu tố trong mùa đông là lớn hơn so với mùa hè. Hai chỉ số này dao động trong khoảng từ 1 cho đến 2.5. Ngoài ra, khi nghiên cứu kỹ 3 đặc trưng thống kê nói trên cho từng biến dự báo tại các hạn dự báo, trạm nghiên cứu và mùa dự báo khác nhau chúng tôi nhận thấy BIAS tại mỗi trạm nghiên cứu và từng biến dự báo là rất khác nhau. Sự khác nhau trong giá trị BIAS theo không gian (từng trạm) và thời gian (từng

mùa) phản ánh sự tác động của các đặc tính địa phương tới chất lượng dự báo từ mô hình HRM. Tuy nhiên, BIAS theo các hạn dự báo cho từng biến dự báo và trạm nghiên cứu là tương đối thống nhất (cùng dấu). Việc khảo sát tính chất sai số dựa trên chỉ số BIAS trong nghiên cứu này thu nhận được một số kết quả chính như sau:

- Có một sự liên hệ giữa sai số độ cao địa hình và BIAS trong dự báo nhiệt độ không khí tại độ cao 2m (T2m): những trạm có BIAS âm (dự báo thấp hơn so với thực tế) là những trạm có sai số địa hình dương (tức là độ cao trong mô hình cao hơn so với độ cao thực tế) và ngược lại. Sai số địa hình mà càng lớn thì BIAS càng lớn.

- Đối với dự báo T2m và Td2m, vào mùa hè giá trị BIAS tại các hạn dự báo +12h và +36h thường lớn hơn các hạn dự báo +24h và +48h. Ngược lại, vào mùa đông, BIAS cho các hạn dự báo +24h và +48h lại lớn hơn so với hai hạn dự báo còn lại. Điều này ngụ ý rằng dự báo nhiệt độ tối cao có BIAS lớn hơn so với dự báo nhiệt độ tối thấp vào mùa hè và ngược lại.

- Đối với dự báo SP10m, các trạm đảo thường có BIAS lớn hơn so với các trạm trên

đất liền.

- Nói chung BIAS giữ nguyên xu thế khi chuyển từ mùa hè sang mùa đông đối với dự báo T2m, Td2m và SP10m tại tất cả các hạn dự báo. Ngoài trừ trường hợp của biến T2m và Td2m cho khu vực Nam Bộ là có BIAS thay đổi gần như ngược lại khi chuyển từ mùa hè sang mùa đông.

Từ các kết quả đánh giá dựa trên chỉ số thống kê MAE, chúng tôi cũng thu được một số kết quả như sau:

- Về mặt trung bình, giá trị MAE biến đổi trong khoảng từ 1-2,5<sup>o</sup>C đối với các biến nhiệt độ T2m và Td2m (ngoại trừ một số trạm như 48803, 48808 thuộc Bắc Bộ, 48840 ở Trung Bộ và 48887 ở Nam Bộ) và 1,5-3m/s đối với tốc độ gió tại độ cao 10m. Đối với T2m và Td2m, giá trị MAE cho các trạm thuộc Nam Bộ là nhỏ hơn so với các trạm Trung Bộ và Bắc Bộ, các trạm Bắc Bộ nói chung có giá trị MAE lớn. Sự biến đổi giá trị MAE từ mùa hè sang mùa đông thể hiện rõ ở các trạm Bắc Bộ, các trạm thuộc Trung Bộ và Nam Bộ nói chung thay đổi không đáng kể, đặc biệt là

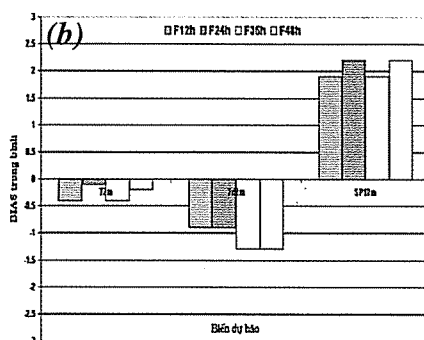
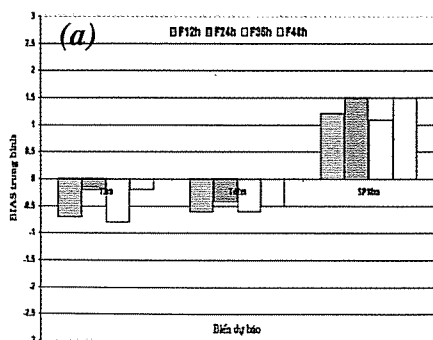
vùng Nam Bộ.

- Các trạm trên biển thường có MAE cho trường nhiệt và ẩm thấp hơn so với các trạm trên đất liền.

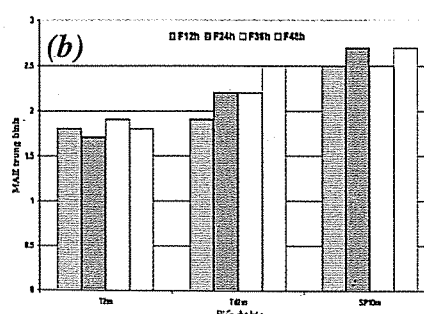
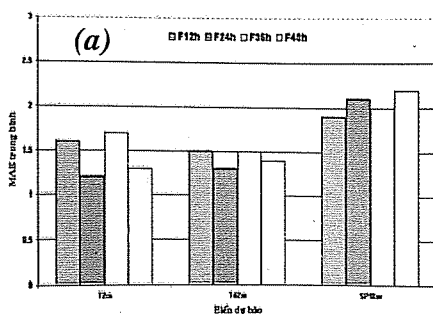
- Đối với dự báo nhiệt độ T2m, giá trị MAE của các hạn dự báo +12h và +36h thường lớn hơn so với hai hạn dự báo +24h và +48h trong cả hai mùa.

- Sự biến thiên mạnh của MAE cho dự báo tốc độ gió theo mùa thể hiện rõ rệt qua các trạm ở khu vực Nam Bộ.

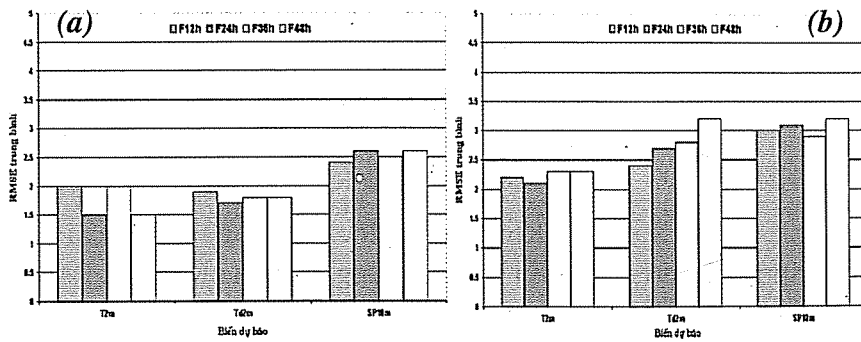
Đối với đại lượng RMSE, nói chung RMSE biến đổi từ 1-2,5<sup>o</sup>C trong mùa hè và 1,5-3<sup>o</sup>C trong mùa đông đối với các biến T2m và Td2m, từ 1-3m/s đối với biến SP10m. Ngoài ra, các đặc trưng biến đổi của RMSE theo không gian (theo từng vùng) và thời gian (từ hè sang đông) trong nghiên cứu này là gần tương tự như các đặc trưng đã được nêu trong phần đánh giá đối với đại lượng thống kê MAE (toàn bộ phần khảo sát sai số này, có thể tham khảo chi tiết chi [4], bài báo này chỉ đưa ra những kết quả chính).



Hình 2. Giá trị BIAS trung bình trên 25 trạm nghiên cứu của 3 biến dự báo tại 4 hạn dự báo tương ứng cho mùa hè (a) và mùa đông (b)



Hình 3. Giá trị MAE trung bình trên 25 trạm nghiên cứu của 3 biến dự báo tương ứng cho mùa hè (a) và mùa đông (b)



Hình 4. Giá trị PMSE trung bình trên 25 trạm nghiên cứu của 3 biến dự báo tại 4 hạn dự báo tương ứng cho mùa hè (a) và mùa đông (b)

#### 4. Một số kết quả ứng dụng lọc Kalman hiệu chỉnh các dự báo trực tiếp từ mô hình HRM

Trong nghiên cứu này, chúng tôi sử dụng bộ chương trình mã nguồn lọc Kalman của Trung tâm Dự báo Hạn vừa Châu Âu (ECMWF) với một số thay đổi nhỏ để phù hợp với bài toán Việt Nam. Để đánh giá hiệu quả của việc áp dụng lọc Kalman hiệu chỉnh các dự báo T2m, Td2m và SP10m so với dự báo trực tiếp từ mô hình (control forecast), trong nghiên cứu này chúng tôi sử dụng chỉ số thống kê RSS (Relative Skill Score) được định nghĩa như sau:

$$RSS = \frac{Err_{control} - Err_{Kalman}}{Err_{control} + Err_{Kalman}} * 100\% \quad (4)$$

Trong đó  $Err_{control}$  là sai số của dự báo trực tiếp từ mô hình,  $Err_{Kalman}$  là sai số của dự báo đã được hiệu chỉnh qua lọc Kalman và giá trị của chỉ số RSS được tính bằng %. Từ công thức trên có thể thấy nếu RSS mang giá trị dương, thì có nghĩa là dự báo có áp dụng lọc Kalman cho sai số dự báo nhỏ hơn so với dự báo trực tiếp. Hay nói cách khác, lọc Kalman đã đưa đến một sự cải thiện trong chất lượng dự báo, mức độ cải thiện tương đối sẽ được biểu hiện qua giá trị của chỉ số RSS. Trong trường hợp ngược lại, tức là RSS âm, thì không có bất kỳ sự cải thiện nào, thậm chí việc áp dụng lọc Kalman còn làm cho sai số dự báo tăng lên. Trong nghiên cứu này, chúng tôi áp dụng chỉ số RSS cho 2 chỉ số thống kê là MAE và RMSE và quá trình đánh giá được

áp dụng cho 2 tùy chọn lọc Kalman là sơ đồ một tham số và sơ đồ hai tham số (xem phần I của bài báo [6]). Từ các đặc trưng phân bố theo không gian (theo từng trạm) và theo thời gian (hạn dự báo) của chỉ số RSS chúng tôi thu được một số kết quả như sau:

- Hiệu chỉnh các kết quả dự báo T2m, Td2m và SP10m bằng lọc Kalman cho chỉ số RSS có giá trị dương tại hầu hết các trường hợp (theo từng hạn dự báo, tại từng trạm nghiên cứu, theo mùa và đối với từng tùy chọn sơ đồ một tham số hay hai tham số). Ngoại trừ một vài trường hợp có chỉ số RSS âm, nhưng giá trị không lớn (không vượt quá 5%). Các kết quả này chứng tỏ việc áp dụng lọc Kalman đã mang lại một sự cải thiện chất lượng dự báo so với dự báo trực tiếp từ mô hình với mức độ cải thiện tùy thuộc vào từng biến dự báo, hạn dự báo, mùa, tùy chọn sơ đồ và từng trạm nghiên cứu.

- Đối với dự báo nhiệt độ không khí tại độ cao 2m (T2m): các kết quả cho thấy mức độ cải tiến đối với các trạm vùng Bắc Bộ lớn hơn các trạm vùng Trung Bộ và Nam Bộ trong cả mùa hè lẫn mùa đông. Việc áp dụng lọc Kalman hiệu chỉnh dự báo T2m cho các trạm Nam Bộ mang lại sự cải tiến không đáng kể. Đặc biệt là các trạm trên biển (do có sai số hệ thống nhỏ). Trong mùa đông, sự cải tiến là lớn hơn so với mùa hè. Đối với các trạm thuộc Trung Bộ và Nam Bộ, trong hầu hết các trường hợp sơ đồ một tham số cho sai số nhỏ hơn sơ đồ hai tham số. Tuy nhiên, đối với các

trạm thuộc Bắc Bộ, vào mùa hè sơ đồ một tham số là tốt hơn, nhưng vào mùa đông thì sơ đồ hai tham số lại cho sự cải thiện lớn hơn. Một số trạm cho thấy sự cải thiện đáng kể (trên 40%) trong chất lượng dự báo T2m là 48808, 48803 và 48848.

- Đối với dự báo nhiệt độ điểm sương tại độ cao 2m (Td2m): sự cải thiện trong chất lượng dự báo trong mùa hè tại hầu hết các trạm và hạn dự báo là không nhiều ngoại trừ một số trạm như 48803, 48830. Sự cải thiện lớn tìm thấy vào mùa đông, đặc biệt là đối với các trạm thuộc Bắc Bộ. Cũng giống như trường hợp biến T2m, việc so sánh mức độ cải tiến giữa hai sơ đồ lọc Kalman trong trường hợp này cho thấy vào mùa hè sơ đồ một tham số là tốt hơn, nhưng vào mùa đông sơ đồ hai tham số lại cho sự cải thiện lớn hơn. Một số trạm cho chỉ số RSS lớn phải kể đến như 48803, 48887. Sự cải thiện trong chất lượng dự báo của các trạm đất liền là lớn hơn nhiều so với các trạm trên biển.

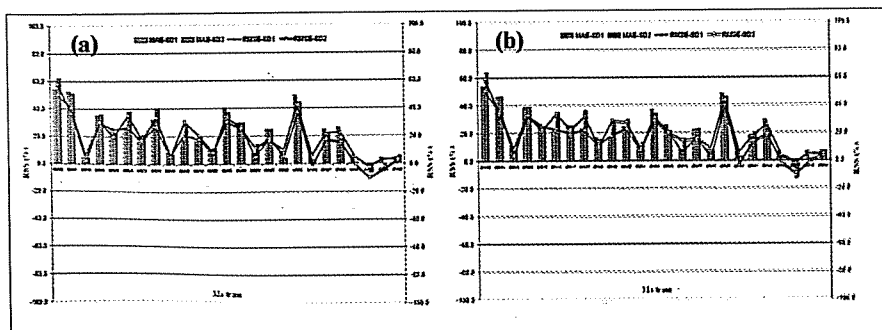
- Đối với dự báo tốc độ gió tại độ cao 10m (SP10m): sự cải thiện đáng kể tìm thấy tại hầu hết các trạm nghiên cứu và hạn dự báo. Các trạm trên đất liền có chỉ số RSS dương và lớn hơn nhiều so với các trạm trên biển. Khác hẳn với dự báo T2m và Td2m, sơ đồ hai tham số là tốt hơn sơ đồ một tham số trong hầu hết trường hợp và mùa dự báo. Như trường hợp của các trạm 48803, 48870 và 48877. Nói chung, việc áp dụng lọc Kalman hiệu chỉnh dự báo SP10m cho các trạm thuộc Trung Bộ, cho sự cải thiện lớn nhất

- Sự cải thiện lớn trong chất lượng dự báo tại một trạm nào đó (có chỉ số RSS dương và lớn) có một mối liên hệ chặt chẽ với đặc trưng sai số tại trạm đó. Tức là, khi BIAS dự báo mà càng lớn thì chỉ số RSS dương sẽ càng lớn.

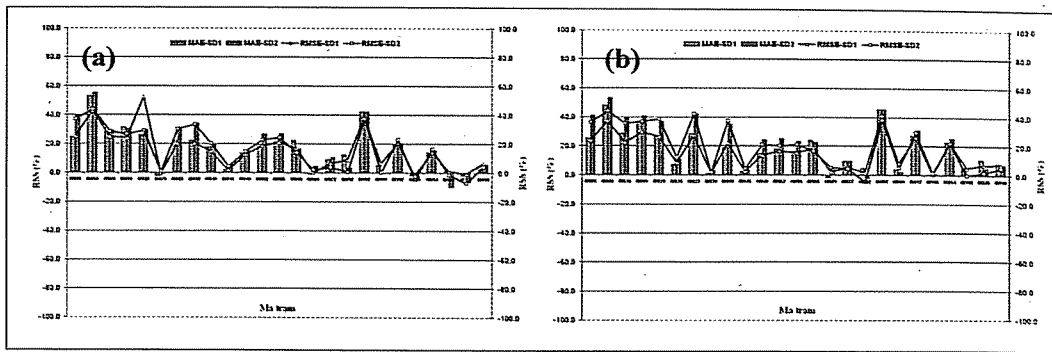
- Khi áp dụng lọc Kalman, mức độ cải tiến đối với các đặc trưng thống kê MAE và RMSE là gần như nhau.

- Sau khi áp dụng lọc Kalman, sai số hệ thống là gần như được loại bỏ, tức là BIAS gần bằng không.

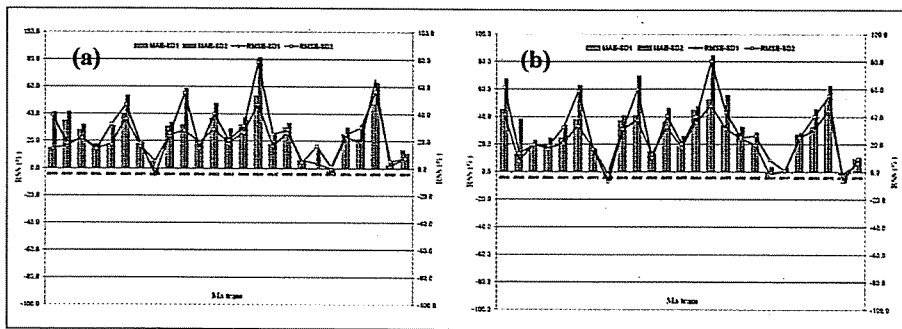
Các hình vẽ 5-7 minh họa một số kết quả so sánh chỉ số RSS giữa các sơ đồ lọc Kalman một tham số (SD1) và hai tham số (SD2) cho các biến dự báo T2m, Td2m và SP10m cho hai đại lượng thống kê MAE và RMSE cho 25 trạm nghiên cứu vào mùa đông. Các hình vẽ 8 và 9 minh họa cho một số kết quả nghiên cứu ứng dụng lọc Kalman hiệu chỉnh dự báo T2m tương ứng tại các trạm Cao Bằng (48808) và Lào Cai (48803) cho chuỗi số liệu mùa đông. Có thể nhận thấy, trước khi áp dụng lọc Kalman, sai số hệ thống là khó rõ. Tuy nhiên, sau khi qua lọc Kalman, sai số hệ thống hầu như được khử hoàn toàn như trong trường hợp trạm Lào Cai. Sự khử được sai số hệ thống này dẫn đến sự cải thiện đáng kể trong chất lượng dự báo, tức là giảm đáng kể sai số tuyệt đối trung bình MAE và căn bậc hai của sai số bình phương trung bình RMSE. Từ việc so sánh giữa giá trị quan trắc, giá trị dự báo trực tiếp từ mô hình và đã hiệu chỉnh qua Kalman cho toàn bộ chuỗi số liệu ta có thể nhận thấy sự cải tiến là khá ổn định.



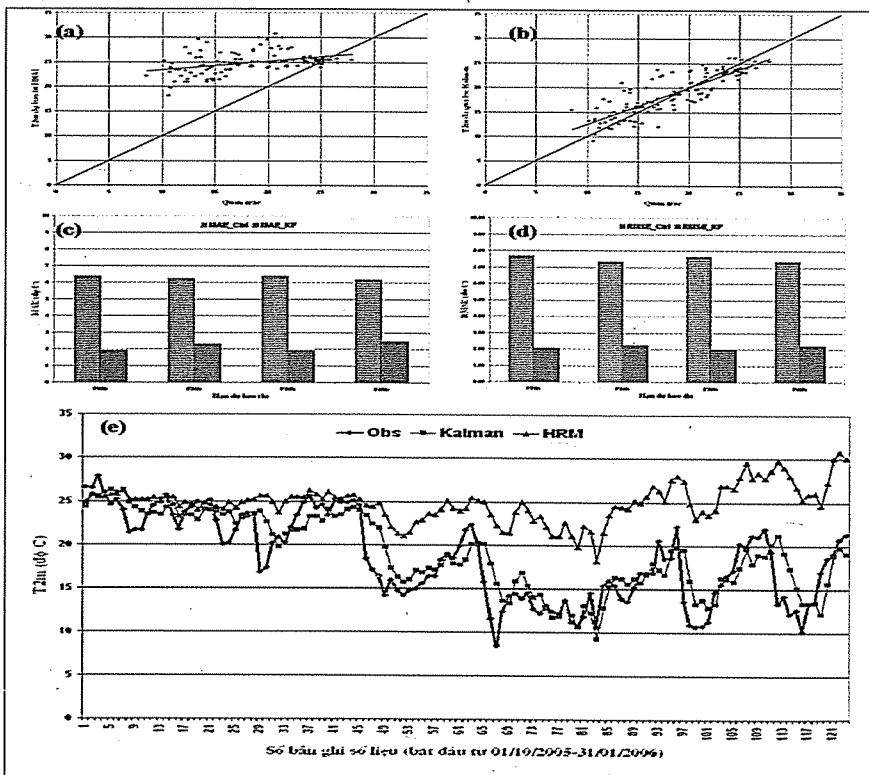
Hình 5. So sánh chỉ số RSS(%) của biến T2m tại các hạn dự báo +12h (a) và +36h (b) cho chỉ số MAE (dạng cột) và RMSE (dạng đường) giữa sơ đồ 1 tham số (SD1) và hai tham số (SD2) cho 25 trạm nghiên cứu vào mùa hè.



Hình 6. So sánh chỉ số RSS(%) của biến Td2m tại các hạn dự báo +12h (a), +36h (b), cho chỉ số MAE (dạng cột) và RMSE (dạng đường) giữa sơ đồ 1 tham số (SD1) và hai tham số (SD2) cho 25 trạm nghiên cứu vào mùa đông.

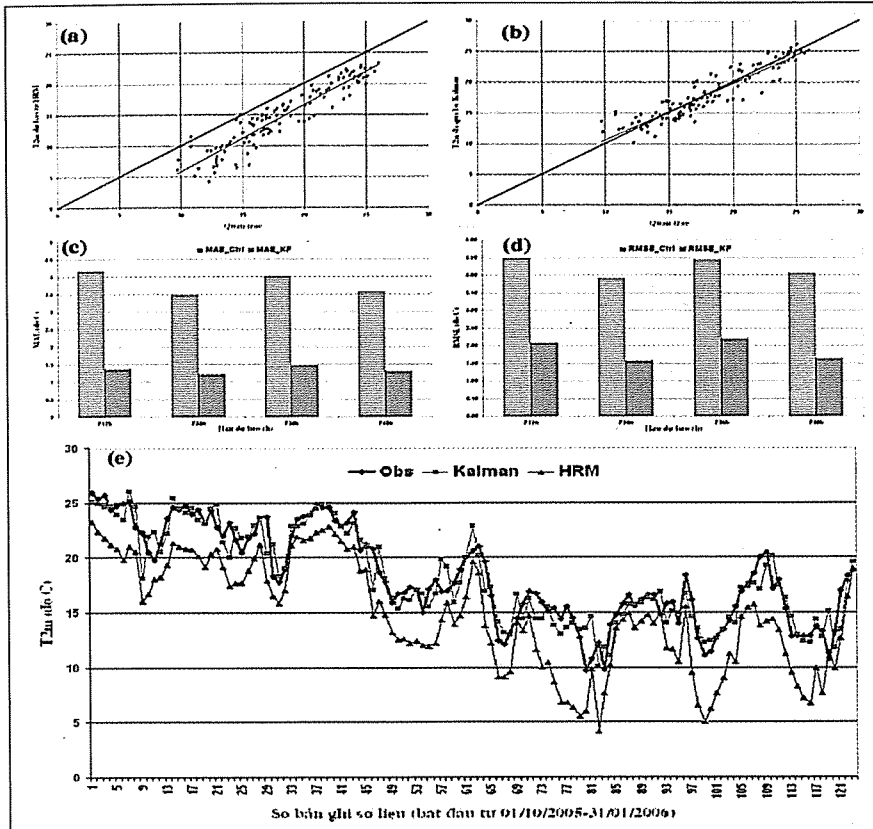


Hình 7. So sánh chỉ số RSS(%) của biến SP10m tại các hạn dự báo +12h (a) và +24h (b) cho chỉ số MAE (dạng cột) và RMSE (dạng đường) giữa sơ đồ 1 tham số (SD1) và hai tham số (SD2) cho 25 trạm nghiên cứu vào mùa đông.



Hình 8. Một số đặc trưng thống kê dự báo T2m hạn +12h tại trạm Cao Bằng (48808) cho chuỗi số liệu mùa đông. Các biểu đồ khuếch tán (a) và (b) mô tả sai số hệ thống của dự báo trực tiếp từ HRM và sau khi lọc Kalman. Hình nhỏ (c) và (d) so sánh giá trị MAE và RMSE trước và sau khi lọc Kalman. Hình nhỏ (e) so sánh giá trị quan trắc (Obs), dự báo trực tiếp (HRM) và đã hiệu chỉnh (Kalman) cho toàn bộ chuỗi số liệu mùa đông.





Hình 9. Một số đặc trưng thống kê dự báo T2m hạn +24h tại trạm Lào Cai (48803) cho chuỗi số liệu mùa đông. Các biểu đồ khuếch tán (a) và (b) mô tả sai số hệ thống của dự báo trực tiếp từ HRM và sau khi lọc Kalman. Hình nhỏ (c) và (d) so sánh giá trị MAE và RMSE trước và sau khi lọc Kalman. Hình nhỏ (e) so sánh giá trị quan trắc (Obs), dự báo trực tiếp (HRM) và đã hiệu chỉnh (Kalman) cho toàn bộ chuỗi số liệu mùa đông.

### 5. Kết luận

Thông qua việc tìm hiểu cơ sở lý thuyết của lọc Kalman từ tổng quan cho đến cách tiếp cận cho bài toán hiệu chỉnh sau mô hình khí tượng, trong nghiên cứu này chúng tôi đã thử nghiệm áp dụng bộ chương trình lọc Kalman của Trung tâm dự báo hạn vừa Châu Âu (ECMWF) để hiệu chỉnh các kết quả dự báo nhiệt độ không khí (T2m) và nhiệt độ điểm sương (Td2m) tại độ cao 2m, tốc độ gió tại độ cao 10m (SP10m) từ mô hình khu vực phân giải cao HRM cho 25 trạm synop phát báo quốc tế với các chuỗi số liệu mùa hè và mùa đông trong năm 2005. Quá trình khảo sát và đánh giá sai số dự báo của HRM cho các biến, các trạm và hai chuỗi số liệu nói trên đã cho thấy một sai số hệ thống trong dự báo trực tiếp từ mô hình HRM. Các dự báo nhiệt độ tối cao và tối thấp từ mô hình HRM thường cho sai số lớn. Sai số lớn thường tìm thấy đối với các trạm thuộc Bắc Bộ hoặc trên đất liền. Sai số dự báo vào mùa đông cũng thường lớn hơn so với mùa hè. Các kết quả đánh giá hiệu quả

của lọc Kalman dựa trên chỉ số thống kê RSS đã cho thấy sự cải thiện đáng kể trong chất lượng dự báo (giảm sai số dự báo) tại hầu hết các hạn dự báo và trạm nghiên cứu. Mức độ cải thiện tùy thuộc vào từng biến, hạn dự báo và từng trạm. Đặc biệt, sau khi áp dụng lọc Kalman, sai số hệ thống hầu như được khử hoàn toàn. Việc so sánh kỹ năng hiệu chỉnh giữa sơ đồ lọc Kalman một tham số và hai tham số cho thấy đối với các dự báo T2m và Td2m, sơ đồ một tham số khả quan hơn sơ đồ hai tham số cho chuỗi số liệu mùa hè (đặc biệt đối với các trạm thuộc Bắc Bộ). Tuy nhiên sang mùa đông thì sơ đồ hai tham số lại là tối ưu hơn. Đối với dự báo tốc độ gió SP10m, sơ đồ hai tham số là tối ưu nhất trong cả hai mùa. Từ những thành công ban đầu trong việc nghiên cứu ứng dụng lọc Kalman để hiệu chỉnh các kết quả dự báo từ mô hình HRM, nhóm nghiên cứu mạnh dạn đề xuất đưa bộ chương trình lọc Kalman vào dự báo nghiệp vụ cho 25 trạm nghiên cứu nói trên. Ngoài ra, tiếp tục nghiên cứu khảo sát thêm cho một số

biến dự báo khác và tiến tới áp dụng cho toàn bộ các trạm trên lãnh thổ Việt Nam. Bên cạnh mô hình nghiệp vụ HRM, lọc Kalman cũng sẽ được thử nghiệm áp dụng cho các mô hình khu vực bán nghiệp vụ khác như ETA và WRF,

hoặc cho các sản phẩm dự báo trực tiếp từ một số mô hình toàn cầu như GME, ECMWF, GFS, GSM, ... đang được thu nhận hàng ngày tại Trung tâm Dự báo KTTV Trung ương.

### Tài liệu tham khảo

1. *Kiều Thị Xin, Lê Công Thành, Phan Văn Tân, Đỗ Lệ Thủy, Nguyễn Văn Sáng. Mô hình dự báo số phân giải cao HRM và thử nghiệm áp dụng dự báo thời tiết khu vực Đông Nam Á- Việt Nam. Tạp chí KTTV, số 8(488), tr 36 - 44. 2001.*
2. *Lê Bắc Huỳnh. Ứng dụng kỹ thuật lọc Kalman vào tính toán và dự báo dòng chảy. Tạp chí KTTV, số 11, tr 15 - 25. 1990.*
3. *Nguyễn Chi Mai, Nguyễn Thu Hằng. Ứng dụng lọc Kalman cho hiệu chỉnh dự báo bão. Báo cáo tại Hội nghị Khoa học Công nghệ Dự báo và phục vụ Dự báo KTTV lần thứ VI, tr 164 - 176. 2005.*
4. *Nguyễn Thu Hằng, Võ Văn Hòa, Nguyễn Thị Anh Đào, 2006. Nghiên cứu ứng dụng lọc Kalman hiệu chỉnh một số trường dự báo từ mô hình HRM. Báo cáo tổng kết đề tài cơ sở, tr 1-43.*
5. *Võ Văn Hòa, Nguyễn Chi Mai, Nguyễn Thu Hằng, 2005. Áp dụng thử nghiệm lọc Kalman cho nhiệt độ và độ ẩm dự báo bằng mô hình số trị. Báo cáo tại hội nghị khoa học công nghệ dự báo và phục vụ dự báo KTTV lần thứ VI, tr 153 - 163.*
6. *Võ Văn Hòa, Nguyễn Thu Hằng, Nguyễn Chi Mai. Nghiên cứu ứng dụng lọc Kalman hiệu chỉnh một số trường dự báo bề mặt từ mô hình HRM. Phần I, Khái quát về lọc Kalman. Tạp chí KTTV, số 557 tháng 5, tr 49-57 năm 2007.*
7. *Daniel. S. W., 1995: Statistical methods in the atmospheric sciences. An introduction. Academic Press, tr 159-209.*
8. *Greg, W. and Gary, B., 2004: An introduction to the Kalman filter, tr 1-16*
9. *Hall, T., H. E. Brooks and C. A. Doswell, 1997: Precipitation Forecasting using a Neural Network, Submitted to Weather and Forecasting, November 1997.*
10. *Joel Le Roux, 2003: An introduction to Kalman filter: probabilistic and deterministic approaches, tr 1-12.*
11. *M. Casaioli, R. Mantovani, F Proietti Scorzonti, S. Puca, A. Speranza, and B. Tirozzi, 2002. Linear and nonlinear post-processing of numerically forecasted surface temperature. Nonlinear process in Geophysics, 10, tr 373-383.*
12. *Person A., 1991: Kalman filtering-A new approach to adaptive statistical interpretation of numerical meteorological forecasts. Lectures and papers presented at the WMO training on the interpretation of NWP products in terms of local weather phenomena and their verification, WMO, Wageningen, the Netherlands, XX-27-XX-32.*
13. *Simonsen, C., 1991: Self adaptive model output statistics based on Kalman filtering. Lectures and papers presented at the WMO training on the interpretation of NWP products in terms of local weather phenomena and their verification, WMO, Wageningen, the Netherlands, XX - 28 - XX - 33.*
14. *Vislocky, R. L. and J. M. Fritch, 1995: Generalized additive models versus linear regression in generating probabilistic MOS forecast of aviation weather parameters. Weather and Forecasting, 10, 669 - 680.*