

VAI TRÒ CỦA SAI SỐ MÔ HÌNH TRONG BÀI TOÁN ĐỒNG HÓA SỐ LIỆU DỰA TRÊN PHƯƠNG PHÁP BIẾN PHÂN: THỬ NGHIỆM VỚI MÔ HÌNH PHÂN GIẢI CAO WRF-ARW VÀ DỰ BÁO MƯA LỚN TRONG TRÊN KHU VỰC BẮC BỘ

Dư Đức Tiến¹, Hoàng Đức Cường¹, Mai Khánh Hưng¹, Hoàng Phúc Lâm¹

Tóm tắt: Trong bài toán đồng hóa số liệu dựa trên phương pháp biến phân, sai số mô hình đóng vai trò quyết định đến việc lan truyền một cách hiệu quả các thông tin quan trắc khi được bổ sung vào mô hình theo chiều ngang, thẳng đứng và giữa các biến mô hình với nhau. Bài báo sẽ trình bày một số khảo sát khác nhau liên quan đến những lựa chọn thiết lập ma trận sai số của mô hình WRF-ARW thử nghiệm trong dự báo mưa lớn trên khu vực Bắc Bộ với sơ đồ đồng hóa biến phân WRFDA. Các thử nghiệm gồm: i) sử dụng sai số cho trước của NCEP và ii) sai số tính toán dựa trên kết quả dự báo của mô hình WRF-ARW cho riêng khu vực Việt Nam. Hai phương pháp tính sai số mô hình gồm sử dụng dự báo ngắn hạn của NMC và dự báo tổ hợp từ chính mô hình WRF-ARW cho khu vực Việt Nam. Một số thử nghiệm ban đầu với trường hợp mưa lớn điển hình trên khu vực Bắc Bộ cho thấy độ nhạy của kỹ năng dự báo mưa (đánh giá thông qua chỉ số ETS trên trạm quan trắc) phụ thuộc vào việc lựa chọn thiết lập ma trận sai số trường nền B và việc tính toán ma trận B từ dự báo riêng cho khu vực Việt Nam là cần thiết.

Từ khóa: Mô hình WRF-ARW, đồng hóa số liệu, sai số dự báo, sơ đồ biến phân, ma trận hiệp biến sai.

Ban Biên tập nhận bài: 05/01/2019 Ngày phản biện xong: 20/02/2019 Ngày đăng bài: 25/03/2019

1. Mở đầu

Theo Lorenz (1963, 1970) [7], một trong những nguồn gây ra sai số chính cho các mô hình dự báo số trị (*Numerical Weather Prediction - NWP*) là độ chính xác của trường điều kiện biên ban đầu và phương pháp đồng hóa số liệu là phương pháp nâng cao độ chính xác của trường ban đầu cho mô hình dự báo số trị thông qua việc phân tích tối ưu lại trường ban đầu từ số liệu thám sát bổ sung ở quy mô địa phương hoặc các số loại số liệu mới như bức xạ vệ tinh hay độ phản hồi của radar, qua đó giảm thiểu được tối đa những phát sinh do sai số ban đầu gây ra. Tiền thân của phương pháp đồng hóa số liệu là các phương pháp phân tích khách quan với mục tiêu xác định được các giá trị của các biến khí tượng trên lưới rời rạc từ các điểm quan trắc hữu hạn

¹Trung tâm Dự báo khí tượng thủy văn quốc gia
Email: duductien@gmail.com

được cung cấp, điển hình gồm các phương pháp phân tích khách quan (*Objective Analysis*) của Cressman (1959) [7] và phương pháp hiệu chỉnh liên tiếp (*Successive Correction*) của Barnes (1973) [1,7]. Giả sử mô hình số trị NWP gồm các biến gió (thành phần kinh hướng u , vĩ hướng v và thẳng đứng w), nhiệt độ (T), áp suất (P) và độ ẩm (q), khi đó ký hiệu vector x_b còn gọi là trường nền hay trường phân tích ban đầu của mô hình chưa được bổ sung số liệu quan trắc cho trước là vector nhiều chiều có các thành phần bao gồm tất cả các giá trị u, v, w, T, P và q theo không gian, có thể biểu diễn đơn giản thành: $x_b = \{u, v, w, T, P, q\}$. Vì sai số do quá trình nội suy quan trắc về lưới tính, sai số trong quan trắc, độ thừa của quan trắc dẫn tới vector x_b luôn khác so với vector khí quyển thực tế (*truth*), kí hiệu là x_t . Bài toán đồng hóa số liệu dựa trên phương pháp biến phân là việc cực tiểu hóa hàm giá (*cost function*) với dạng như sau:

$$J(x) = \frac{1}{2} [(x - x_b)^T (\mathbf{B})^{-1} (x - x_b) + (y_o - \mathbb{H}(x))^T (\mathbf{R})^{-1} (y_o - \mathbb{H}(x))] \quad (1)$$

Trong đó vector x mà ở đó hàm J cực tiểu sẽ được gọi là trạng thái phân tích (*analysis*) khí quyển tối ưu (x_a) có được khi cho trước vector x_b và quan trắc y_o . Trong phương trình biểu diễn J , kí hiệu T là ma trận chuyển dạng, \mathbf{B} và \mathbf{R} là ma trận sai số hiệp biến tương ứng của trường nền và quan trắc, $\mathbb{H}(x)$ là phép chuyển đổi, nội suy từ các biến mô hình x sang giá trị quan trắc tương ứng y , nếu các quan trắc đưa vào đồng hóa là các biến mô hình như nhiệt độ, độ ẩm, gió hay áp

$$y_o - \mathbb{H}(x) = y_o - \mathbb{H}(x_b + (x - x_b)) = \{y_o - \mathbb{H}(x_b)\} - \mathbf{H}(x - x_b) \quad (2)$$

Khi đó công thức triển khai của đạo hàm $J(x)$ có dạng:

$$\nabla J(x) = \mathbf{B}^{-1}(x - x_b) + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H}(x - x_b) - \mathbf{H}^T \mathbf{R}^{-1} (y_o - \mathbb{H}(x_b)) \quad (3)$$

Giá trị x để J đạt cực đại ứng với vé phải của phương trình $\nabla J(x) = 0$ và biến đổi chuyển vé

$$x_a = x_b + (\mathbf{B}\mathbf{H}^T)(\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^T)^{-1} (y_o - \mathbb{H}(x_b)) \quad (4)$$

Tương tự biểu diễn tường minh của vector $x_b \{u, v, w, T, P, q\}$, ta có thể biểu diễn vector quan trắc $y_o \{u_o, v_o, w_o, T_o, P_o, q_o, \dots Y_o\}$ với Y_o là các quan trắc không trực tiếp và cần chuyển đổi bằng toán tử \mathbb{H} . Trạng thái phân tích (*analysis*) khí quyển tối ưu (x_a) phụ thuộc rất nhiều vào thông tin của \mathbf{B} và \mathbf{R} , đặc biệt ma trận sai số \mathbf{B} sẽ đóng vai trò truyền thông tin theo không gian ngang và thẳng đứng có trọng số đối với quan

$$x_l^a = x_l^b + B_{lk} \frac{y - x_k^b}{B_{kk} + \sigma^2} = x_l^b + \frac{B_{lk}}{B_{kk} + \sigma^2} y - \frac{B_{lk}}{B_{kk} + \sigma^2} x_k^b \quad (5)$$

Vì thế các phần tử không nằm trên đường chéo và khác không của ma trận \mathbf{B} sẽ dẫn tới giá trị thay đổi đến trường phân tích cho thành phần thứ l trong véc tơ trạng thái x . Ngoài ra ta thấy rằng, nếu σ^2 nhỏ hơn nhiều (\ll) phần tử đường chéo B_{kk} khi đó phần tử thứ k trong véc tơ phân tích tối ưu x_k^a sẽ xấp xỉ giá trị quan trắc y , và nếu σ^2 lớn hơn nhiều (\gg) phần tử đường chéo B_{kk} khi đó giá trị phân tích hầu như không khác so với giá trị trường nền ban đầu, hay $x_k^a \sim x_k^b$. Như vậy, nếu ma trận sai số B là quá lớn so với sai số quan trắc, phân tích sẽ gần với quan trắc, ngược lại phân tích tối ưu sẽ gần so với giá trị trường nền ban đầu. Tính phức tạp của ma trận B thể hiện trước hết ở độ lớn của ma trận này, nếu mô

suất thì \mathbb{H} chỉ là quá trình nội suy thuần túy, nếu các quan trắc là dạng phi truyền thống như bức xạ vệ tinh, độ phản hồi radar thì \mathbb{H} là một sơ đồ/mô hình cụ thể, ví dụ sơ đồ truyền bức xạ tính toán và chuyển đổi thông số profile nhiệt ẩm khí quyển cho trước sang giá trị bức xạ ứng với các bước sóng định trước [1]. Để hàm $J(x)$ đạt cực tiểu thì đạo hàm của J theo x phải bằng 0 hay $\nabla_x J(x_a) = 0$. Gọi toán tử \mathbf{H} là toán tử thỏa mãn xấp xỉ sau:

cho x_a ta được công thức cuối cùng về nghiệm phân tích tối ưu:

trắc đưa vào cho trường ban đầu. Để minh họa rõ hơn ý nghĩa của sai số ma trận \mathbf{B} , ta xem xét chỉ có một quan trắc thứ k , có nghĩa véc tơ quan trắc sẽ có dạng $y_o = \{0, 0, 0, \dots y, \dots 0\}$, khi đó ma trận \mathbf{H} và \mathbb{H} cũng sẽ là một véc tơ hàng với các giá trị phần tử bằng 0 trừ phần tử thứ $k = 1$ và $y_o = y$, $\mathbf{R} = \sigma^2$. Khi đó phương trình cho trường phân tích tối ưu sẽ được viết lại dưới dạng:

hình có kích thước theo không gian $\sim 10^2 \times 10^2 \times 10^2 \times$ số biến (~ 5) thì kích thước của ma trận $\mathbf{B} > 10^6 \times 10^6$ là rất lớn để có thể lưu trữ nên trong thực tế ma trận B được đơn giản hóa bằng các giải thiết khác nhau, ví dụ trong phương pháp của WRFDA cho mô hình WRF-ARW, ma trận B được giả thiết là thiết lập bởi ma trận U sao cho $B = U U^T$ và ma trận U lại tiếp tục được giả thiết là có thể phân giải thành các thành phần khác nhau sao cho quá trình cực tiểu hóa hàm giá J có thể thực hiện được. Các phương pháp này còn được gọi là tiền điều kiện hóa (*preconditioning*) [1,5].

Với tầm quan trọng của sai số mô hình (ở đây mà ma trận B) đến việc lan truyền một cách hiệu

quả các thông tin quan trắc được đưa bổ sung vào mô hình, bài báo sẽ trình bày một số khảo sát khác nhau liên quan đến những lựa chọn thiết lập ma trận sai số của mô hình WRF-ARW thử nghiệm trong dự báo mưa lớn trên khu vực Bắc Bộ với sơ đồ đồng hóa biến phân WRFDA. Mục II của bài báo là thiết kế thử nghiệm bao gồm giới thiệu mô hình dự báo WRF-ARW, sơ đồ đồng hóa WRFDA số liệu điều kiện biên, trường hợp thử nghiệm, số liệu quan trắc. Những phân tích kết quả được đưa ra trong phần III của bài báo và một số kết luận chính được tổng kết trong phần VI.

2. Thiết kế thí nghiệm

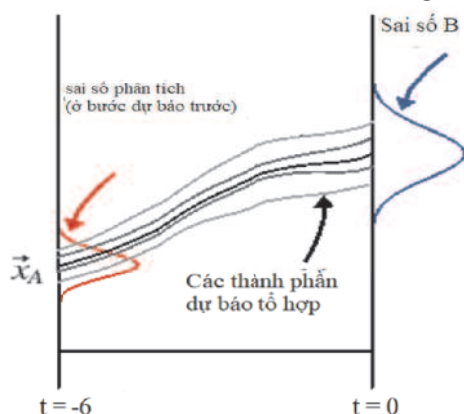
2.1. Mô hình số trị khu vực WRF-ARW

Nghiên cứu sử dụng hệ thống mô hình khu vực WRF với nhân động lực ARW phiên bản 3.9.1.1 do Trung tâm dự báo môi trường quốc gia Mỹ (NCEP) phát triển (gọi tắt là WRF-ARW). Mô hình WRF-ARW được nghiên cứu và áp dụng phổ biến tại Việt Nam những năm vừa qua, các công trình điển hình có thể kể đến gồm trong nghiên cứu dự báo thời tiết và bão hay trong việc áp dụng và dự báo tổ hợp [1,2,3,4]. Chi tiết hơn về mô hình WRF-ARW có thể tham khảo trong [8, 10].

Thử nghiệm thiết lập hai lưới tính gồm 199 điểm lưới theo kinh hướng và 199 điểm lưới theo vĩ hướng với độ phân giải ngang là 15 km (miền

$$B = \langle (\vec{x} - \vec{x}_{true})(\vec{x} - \vec{x}_{true})^T \rangle \sim \alpha \langle (\vec{x}_{48} - \vec{x}_{24})(\vec{x}_{48} - \vec{x}_{24})^T \rangle$$

Đối với phương pháp ước lượng B qua dự báo tổ hợp, ta cần thực hiện dự báo tổ hợp dựa trên mô hình áp dụng (ở đây là WRF-ARW). Độ tán của dự báo tổ hợp tương đương với thông tin của ma trận sai số B, xem minh họa trong hình 1.



Hình 1. Minh họa ước lượng sai số ma trận B từ dự báo tổ hợp ở các lần thực hiện dự báo trước đó

tính bao phủ toàn bộ Việt Nam và Biển Đông) và 5 km (miền tính bao phủ toàn bộ miền Bắc) cho mô hình WRF-ARW, cả hai miền tính đều có số mực thẳng đứng là 41 và dự báo đến hạn 72h. Đối với dự báo chuẩn (*control*), các lựa chọn vật lý cho mô hình WRF-ARW bao gồm: sơ đồ Kain-Fritsch (KF) cho tham số hóa đối lưu, sơ đồ bức xạ sóng ngắn Goddard, sơ đồ tham số hóa lớp biên của Mellor-Yamada-Janjic (MYJ), sơ đồ vi vật lý mây khép kín 6 bậc WSM6 và sơ đồ rời bề mặt là Janjic Eta.

2.2. Sơ đồ đồng hóa biến phân WRFDA và chương trình thiết lập ma trận B

Song song với mô hình WRF-ARW, hệ thống đồng hóa số liệu dựa trên phương pháp biến phân được phát triển với tên gọi là hệ thống WRFDA [5]. Trong hệ thống WRFDA, chương trình thiết lập ma trận B được gọi là *gen_be* trong đó cung cấp việc tính ma trận B theo các cách tiếp cận khác nhau. Chương trình *gen_be* của hệ thống WRFDA đã cung cấp sẵn hai cách ước lượng ma trận B gồm NMC [9] và ENS (sử dụng độ tán của hệ thống tổ hợp [6]). Trong phương pháp NMC, sai số B được ước lượng từ covariance của hai ộp dự báo cách nhau 24h, với kí hiệu $\langle \langle \rangle \rangle$ là phép lấy trung bình, khi đó B có thể xấp xỉ bởi công thức:

Trong nghiên cứu này, để tạo ra dự báo tổ hợp (*ensemble*) trong lựa chọn ENS của chương trình *gen_be*, chúng tôi dựa trên việc thay đổi các lựa chọn vật lý trong mô hình WRF-ARW gồm (a) sơ đồ Kain-Fritsch (KF) hoặc Betts-Miller-Janjic (BMJ) cho tham số hóa đối lưu (*cumulus parameterization scheme-CPS*); (b) sơ đồ bức xạ sóng ngắn Goddard hoặc Dudhia; (d) sơ đồ tham số hóa lớp biên của Yonsei University (YSU) hoặc Mellor-Yamada-Janjic (MYJ) và (e) các sơ đồ vi vật lý mây từ đơn giản đến phức tạp gồm sơ đồ Lin, WSM3, WSM5 đến WSM6 (khép kín 6 bậc). Dựa trên việc tổ hợp các lựa chọn này có thể tạo ra tối đa gồm 32 cấu hình vật lý khác nhau [2]. Lưu ý thêm ở đây đối với sơ đồ lớp biên MYJ thì sơ đồ rời bề mặt sẽ được đặt là sơ đồ Janjic Eta so với sơ đồ khép kín rời dựa trên

giả thiết của Monin-Obukhov. Các dự báo tổ hợp được gọi tắt là **32-ensemble** được thực hiện với mục đích cho thử nghiệm Sử dụng lựa chọn ENS tính từ chương trình *gen_be* từ 1 tuần dự báo tổ

hợp tổ hợp đa vật lý gồm 32 thành phần dựa trên mô hình WRF-ARW cho khu vực Việt Nam. Các thử nghiệm trong nghiên cứu được tổng kết trong bảng 1.

Bảng 1. Các trường hợp thử nghiệm

Kí hiệu thử nghiệm	Thông tin thử nghiệm	Loại thông tin sai số mô hình
EXP_CONTROL	- Không có đồng hóa số liệu	
EXP_CV3_NCP	- Lựa chọn cv=3 trong WRFDA	Sử dụng sai số cho trước của NCEP
EXP_CV5_NMC	- Sai số mô hình tính toán dựa trên kết quả dự báo của mô hình WRF-ARW cho riêng khu vực Việt Nam	Sử dụng lựa chọn NMC tính từ chương trình <i>gen_be</i> từ 2 tuần dự báo trước thời điểm dự báo
EXP_CV5_ENS	(lựa chọn cv=5 trong WRFDA)	Sử dụng lựa chọn ENS tính từ chương trình <i>gen_be</i> từ 1 tuần dự báo tổ hợp tổ hợp đa vật lý gồm 32 thành phần dựa trên mô hình WRF-ARW cho khu vực Việt Nam

2.3. Số liệu điều kiện biên, quan trắc và các trường hợp thử nghiệm

Nghiên cứu sử dụng dự báo từ mô hình toàn cầu GFS (Mỹ) làm điều kiện ban đầu và điều kiện biên cho mô hình WRF-ARW, được cung cấp thông qua địa chỉ sau: <http://www.nco.ncep.noaa.gov/pmb/products/gfs/>. Số liệu GFS có độ phân giải ngang là 55 km với số mực áp suất thẳng đứng của mô hình GFS là 26 và được cập nhật 3 tiếng một cho mô hình WRF-ARW. Các số liệu đồng hóa thử nghiệm gồm bề mặt, cao không và vệ tinh quỹ đạo cực (số liệu đầu đo AMSU, HRIS và MHS) được kiểm soát chất lượng (*quality control*) và hiệu chỉnh tại NOAA, thu nhận trên địa chỉ: <ftp://ftp.ncep.noaa.gov/pub/data/nccf/com/gfs/pr> od.

Thử nghiệm thực hiện dự báo liên quan đến đợt mưa lớn trên khu vực Bắc Bộ trong tháng 5, 6 và 7 năm 2012 và thực hiện dự báo tại các giờ 00z. Ứng với mỗi thời điểm bắt đầu xảy ra mưa lớn được liệt kê trong bảng 1, ví dụ ngày 21/5/2012, dự báo control sẽ được thực hiện từ trước đó ít nhất 15 ngày đối với thử nghiệm EXP_CV5_NMC, và ít nhất 7 ngày (bao gồm cả dự báo control và 32-ensemble) đối với thử nghiệm EXP_CV5_ENS. Ngoài ra sẽ dự báo trước 2 ngày bắt đầu từng đợt mưa đã liệt kê trong bảng 1. Các số liệu quan trắc mưa tại các trạm Synop trên khu vực Bắc Bộ sẽ được sử dụng để đánh giá kĩ năng dự báo mưa giữa các trường hợp khác nhau (>25mm/24h và >50 mm/24h).

Bảng 2. Danh sách đợt mưa lớn ảnh hưởng đến Bắc Bộ năm 2012

Giai đoạn	Khu vực mưa chính	Lượng mưa phổ biến (mm)	Số lượng dự báo (để đánh giá)
21/5 - 26/5	Bắc Bộ và Thanh Hóa	100 – 200	8
14/5 - 15/6	Bắc Bộ	40 – 80	4
1/7 - 2/7	Bắc Bộ và Bắc Trung Bộ	50 – 100	4
17/7 - 19/7	Tây Bắc Bắc Bộ và vùng núi phía Bắc	50 - 100	5
23/7 - 26/7	Vùng núi và trung du Bắc Bộ	80 – 150	6
28/7 - 30/7	Bắc Bộ và Thanh Hóa	50 - 150	5

3. Đánh giá kết quả

Để đánh giá kỹ năng dự báo xảy ra mưa lớn chúng tôi sử dụng chỉ số kỹ năng ETS (*Equitable Threat Score/Gilbert Skill Score*). ETS có giá trị nằm trong khoảng từ -1/3 đến 1 với giá trị bằng 1 được coi là dự báo hoàn hảo. ETS thường được sử dụng trong đánh giá mưa từ mô hình NWP vì

chỉ số này cho phép so sánh công bằng giữa các hình thể hoặc khu vực đánh giá khác nhau. Trong bảng 3 là bảng phân loại (*Contingency table*) theo hiện tượng dự báo (ở đây là mưa với các ngưỡng đánh giá cụ thể, ví dụ > 25mm/24h tại từng trạm quan trắc). Chi tiết về ETS có thể tham khảo trong [4].

Bảng 3. Bảng phân loại tần xuất cho biến dự báo dạng nhị phân

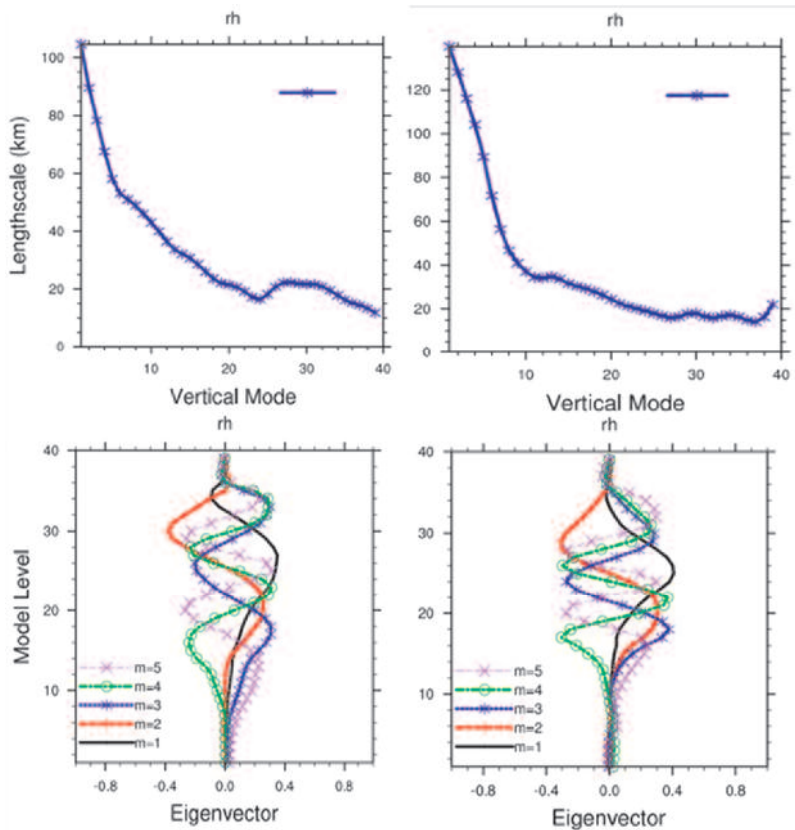
		Quan trắc	
		Có	Không
Dự báo	Có	A	B
	Không	C	D

Khi đó ETS được tính bằng công thức:

$$ETS = \frac{A - A_{\text{random}}}{A + B + C - A_{\text{random}}}$$

Trong đó:

$$(A+B+C+D=N) \text{ và } A_{\text{random}} = \frac{(A + C)(A + B)}{N}$$



Hình 2. Minh họa quy mô ảnh hưởng theo khoảng cách tại từng mực độ cao đối với biến độ ẩm riêng và vector riêng (5 giá trị đầu tiên) trong thử nghiệm NMC (trái) và ENS (phải)

Trong hình 2 là minh họa sự khác biệt của phương pháp NMC (sử dụng control tính từ 21/4/2012-20/5/2012) và ENS (sử dụng 32 - ensemble tính từ 12/5/2012-20/5/2012) liên quan đến thông tin ma trận B. Hình 2 cho thấy quy mô ảnh hưởng theo khoảng cách tại từng mực độ cao đối với thông tin lan truyền cho biến độ ẩm riêng là khác nhau trong thử nghiệm NMC (trái, hàng 1) và ENS (phải, hàng 1), ví dụ dưới mực 10 của mô hình quy mô ảnh hưởng theo khoảng cách đều trên 40km trong phương pháp NMC nhưng ở ENS có thể dưới 35km. Hoặc cũng trong hình 2 (hàng 2) minh họa sự

khác biệt đối với giá trị vector riêng (5 giá trị đầu tiên) trên từng mức mô hình. Những khảo sát kỹ hơn về sự khác biệt này sẽ được chúng tôi thảo luận trong một công trình khác.

Trong bảng 4 là sai số chi tiết của các trường

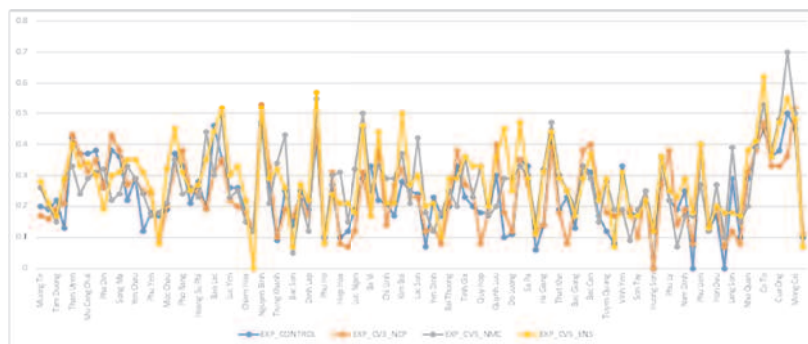
Bảng 4. Chỉ số ETS tại các hạn dự báo 24h, 48h và 72h cho hai ngưỡng đánh giá >25mm/24h và >50mm/24h

Thử nghiệm	Hạn 24h		Hạn 48h		Hạn 72h	
	>25mm/24h	>50mm/24h	>25mm/24h	>50mm/24h	>25mm/24h	>50mm/24h
EXP_CONTROL	0.2552	0.2011	0.2433	0.2172	0.1684	0.1441
EXP_CV3_NCP	0.2298	0.2161	0.2151	0.1754	0.1483	0.1171
EXP_CV5_NMC	0.2688	0.2263	0.2385	0.2075	0.1681	0.1261
EXP_CV5_ENS	0.2768	0.2321	0.2619	0.2095	0.1672	0.1386

Các kết quả cho thấy về mặt trung bình, việc sử dụng sai số mặc định từ NCEP (EXP_CV3_NCP) không cải thiện được chất lượng và kỹ năng dự báo còn giảm đi so với trường hợp dự báo chuẩn. Ở hạn 24h, cả hai phương pháp NMC và ENS đều cho kết quả tốt hơn so với sử dụng mặc định và chưa có đồng hóa và tăng ở cả hai ngưỡng mưa cảnh báo. Một minh họa chi tiết hơn được đưa ra trong hình 3 là tính toán chỉ số ETS tại từng trạm trên khu vực phía Bắc cho bốn trường hợp thử nghiệm cho thấy khá rõ việc cải thiện dự báo tại từng trạm đối với dự báo EXP_CV5_ENS. Ở hạn 48h, dự báo EXP_CV5_ENS sử dụng cập nhật

hợp thử nghiệm khác nhau đánh giá ở 3 hạn dự báo 1,2 và 3 ngày với các ngưỡng >25mm/24h và >50mm/24h trên toàn bộ các trạm quan trắc Synop trên Bắc Bộ (gồm 102 trạm [2]).

từ dự báo tổ hợp cho ma trận B vẫn cho kết quả tốt hơn so với các trường hợp còn lại đối với ngưỡng đánh giá > 25mm/24h, tuy nhiên ở ngưỡng lớn hơn không có sự khác biệt và cải thiện cụ thể. Ở hạn 72h, kỹ năng dự báo ở các trường hợp khá thấp, giảm 30 - 40% so với kỹ năng ở các hạn dự báo 24h và 72h. Cũng ở hạn 72h, sự khác biệt giữa hai trường hợp sử dụng CV5 với dự báo chuẩn là không lớn tại cả hai ngưỡng đánh giá. Đánh giá thêm với sai số trung bình tuyệt đối và trung bình quân phương của mưa dự báo cụ thể tại từng trạm cho thấy biến động về sai số đối với hai phương pháp EXP_CV5_NMC và EXP_CV5_ENS nhỏ hơn



Hình 3. Chỉ số ETS tại từng trạm trên khu vực phía Bắc đối với hạn dự báo 24h, ngưỡng đánh giá > 25mm/24h

4. Kết luận

Nghiên cứu đã trình bày một số khảo sát khác nhau liên quan đến những lựa chọn thiết lập ma trận sai số trong việc sử dụng sơ đồ đồng hóa số liệu WRFDA cho mô hình WRF-ARW trong thử nghiệm trong dự báo mưa lớn trên khu vực Bắc Bộ. Hai phương pháp tính sai số mô hình gồm

sử dụng dự báo ngắn hạn của NMC (EXP_CV5_NMC) và dự báo tổ hợp từ chính mô hình WRF-ARW cho khu vực Việt Nam (EXP_CV5_ENS) được thực hiện và so sánh với việc sử dụng hệ số mặc định của NCEP (EXP_CV3_NCP) và thử nghiệm không có đồng hóa số liệu (CONTROL). Một số thử nghiệm ban

đầu cho thấy hiệu quả của việc tính toán sai số dựa trên những dự báo trên khu vực Việt Nam là cần thiết, thể hiện kết quả dự báo tốt hơn trong các thử nghiệm EXP_CV5_NMC và EXP_CV5_ENS tại hầu hết các ngưỡng và hạn dự báo so với sử dụng hệ số mặc định (EXP_CV3_NCP). Khi tăng hạn dự báo, tính rõ

rệt của việc đồng hóa giảm đi nhưng trường hợp sử dụng dự báo tổ hợp để tính sai số mô hình vẫn cho kết quả tin cậy hơn. Những thử nghiệm tiếp theo sẽ tập trung vào tính hiệu quả ứng với từng loại số liệu quan trắc đưa vào và thực hiện tính toán sai số mô hình trong cả miền tính phân giải cao (5km).

Lời cảm ơn: Các nội dung nghiên cứu trên nằm trong khuôn khổ của đề tài KC.08.06/16-20 “Nghiên cứu xây dựng hệ thống nghiệp vụ dự báo mưa lớn cho khu vực Bắc Bộ Việt Nam”, thuộc Chương trình KC.08/16-20 “Nghiên cứu khoa học và công nghệ phục vụ bảo vệ môi trường và phòng tránh thiên tai”.

Tài liệu tham khảo

1. Dư Đức Tiến và cộng sự (2014), *Nghiên cứu đồng hóa số liệu radar Đông Hà để nâng cao chất lượng dự báo mưa lớn cho khu vực miền Trung*, Tạp chí Khí tượng Thủy văn, (617), tr. 22 -30.
2. Dư Đức Tiến và cộng sự (2019), *Đánh giá tác động của việc sử dụng tham số hóa đối lưu trong dự báo đợt mưa lớn tháng 7 năm 2015 trên khu vực Bắc Bộ bằng mô hình phân giải cao*, Tạp chí KTTV, số 2, 2019, trang 1-8.
3. Hoàng Đức Cường và các cộng sự (2011), *Nghiên cứu ứng dụng mô hình WRF dự báo thời tiết và bão ở Việt Nam*. Báo cáo tổng kết đề tài NCKH cấp Bộ TN&MT.
4. Võ Văn Hòa và cộng sự (2010), *Nghiên cứu phát triển hệ thống dự báo tổ hợp thời tiết hạn ngắn cho khu vực Việt Nam*. Thuyết minh đăng ký đề tài NCKH cấp Bộ, 17 tr.
5. Barker, D.M., Guo, Y.R., Huang, W., Huang, H., Rizvi, S., Xiao, Q., and Lee, M.S. (2005), *WRF-VAR A Unified 4/4D-Var Data Assimilation system for WRF. Preprints. 6th WRF/MM5 Users Workshop*. NCAR. Boulder, CO.
6. Fisher, M. (2003), *Background error covariance modelling, Proceedings of the ECMWF Seminar on Recent Developments in Data Assimilation for Atmosphere and Ocean*, pp. 45-63, ECMWF, Reading, UK.
7. Lorenc, A.C, Neill E. Bowler, Adam M. Clayton, Stephen R. Pring, and David Fairbairn (2015), *“Comparison of Hybrid-4DVar and Hybrid-4DVar Data Assimilation Methods for Global NWP”*, Mon. Wea. Rev., (143), pp. 212-229.
8. Michalakes, J., Dudhia, J., Gill, D., Henderson, T., Klemp, J., Skamarock, W., Wang, W. (2005), *The Weather Research and Forecast Model: Software Architecture and Performance, Proceedings of the Eleventh ECMWF Workshop on the Use of High Performance Computing in Meteorology*, World Scientific, pp. 156-168.
9. Parrish and Derber (1992), *The national meteorological center’s spectral statistical-interpolation analysis system*. Mon. Wea. Rev., 120, 1747-1763.
10. Skamarock, W.C., Klemp, J. B., Dudhia, J., Gill, D.O., Barker, D.M., Duda, M.G., Huang, X.Y., Wang, W., Powers, J.G. (2005), *A description of the Advanced Research WRF Version 3*, NCAR Tech. Note NCAR/TN-475+STR, 113 pp.

THE ROLE OF B-MATRIX IN VARIATIONAL ASSIMILATION SCHEME: CASE STUDY USING WRF-ARW MODEL TO HEAVY RAINFALL FORECAST OVER THE NORTHERN PART OF VIETNAM

Du Duc Tien¹, Hoang Duc Cuong¹, Mai Khanh Hung¹, Hoang Phuc Lam¹

¹National Center of Hydro-Meteorological Forecasting

Abstract: *The data assimilation method based on variational approach, and the model error information or covariance error matrix (B-matrix) of first guess or background forecast plays most important roles in transferring assimilated additional observation data. The case study will present experiments with WRF-ARW and WRFDA assimilation schemes using different information structures of B-matrix on the heavy rainfall forecast over northern part of Vietnam. Two main experiments including: i) using NCEP default background information and ii) using regional background information based on NMC and ensemble forecast approaches. Using quality controlled observation (Synop, TEMP, Low Earth Orbit satellite data from NOAA), the results showed very sensitivities of forecast results and the speed of converging speed in assimilation procedures meaning the most important of selecting B-matrix in assimilation for applying high-resolution regional models in operation. Moreover, the ensemble based B-matrix can provide better skill than other experiments in this study.*

Keywords: *Model forecast error, covariance error matrix, B-matrix, variational assimilation, WRF-ARW, WRFDA.*