

THỬ NGHIỆM HỆ THỐNG ĐỒNG HÓA GSI TRONG BÀI TOÁN DỰ BÁO ĐỊNH LƯỢNG MƯA TRÊN KHU VỰC NAM BỘ

Phạm Quang Nam¹, Mai Văn Khiêm¹, Nguyễn Quang Trung¹, Vũ Văn Thăng¹

Tóm tắt: Nghiên cứu này đánh giá khả năng mô phỏng 15 ngày mưa lớn tại khu vực TP. Hồ Chí Minh trong năm 2018 của mô hình WRF khi đồng hóa số liệu radar Nhà Bè. Trước đó, ảnh hưởng của quá trình đồng hóa đến trường ban đầu đã được phân tích thông qua khảo sát ba chế độ chạy đồng hóa khác nhau, bao gồm cold start, warm start và cycling. Kết quả cho thấy độ phản hồi ở chế độ cold start trở nên tương đồng với trường hợp không đồng hóa sau một giờ tích phân. Ở chế độ warm start, khác biệt của trường ban đầu so với trường hợp không đồng hóa kéo dài hơn, cho thấy vai trò quan trọng của trường dự báo từ kết quả tích phân trước đó. So sánh với số liệu quan trắc lượng mưa tại 11 trạm quan trắc bề mặt, kết quả cho thấy sự cải thiện của các chỉ số FBI, POD, CSI khi chạy ở chế độ cycling. Điều này có thể thấy qua kết quả đánh giá ở cả ba ngưỡng mưa 1, 5 và 10 mm cũng như ở các hạn dự báo 6h và 12h.

Từ khóa: Đồng hóa số liệu, Mô hình WRF, 3DVar, Radar Nhà Bè.

Ban Biên tập nhận bài: 05/12/2018 Ngày phản biện xong: 14/02/2019 Ngày đăng bài 25/02/2019

1. Mở đầu

Ngày nay, mô hình số trị đóng vai trò quan trọng trong bài toán dự báo thời tiết và đặc biệt trong dự báo định lượng mưa [1]. Mặc dù các hệ thống mô hình số trị đã có những bước tiến vượt bậc nhưng vấn đề dự báo mưa của mô hình vẫn chứa đựng nhiều sai số, đặc biệt ở khu vực nhiệt đới và gió mùa [2, 12]. Độ chính xác của các dự báo thời tiết bằng mô hình số trị phụ thuộc vào nhiều yếu tố khác nhau, tuy nhiên trong đó có vai trò quan trọng của điều kiện ban đầu. Do vậy, việc cải thiện chất lượng trường ban đầu đã và đang được các nhà khoa học và các trung tâm nghiên cứu lớn trên thế giới tập trung nghiên cứu phát triển [2]. Một trong những cách tiếp cận để cải thiện trường ban đầu là cập nhật các số liệu địa phương bị bỏ qua hoặc chưa kịp cập nhật vào mô hình toàn cầu. Kỹ thuật cập nhật số liệu quan trắc cho trường ban đầu của mô hình số trị được gọi là đồng hóa số liệu (*data assimilation*). Đồng hóa số liệu cho mô hình dự báo khu vực được kỳ vọng sẽ nâng cao kỹ năng dự báo của mô hình,

đặc biệt với dự báo lượng mưa [8].

Những thập kỷ gần đây chúng kiến sự phát triển mạnh mẽ của các kỹ thuật đồng hóa số liệu [11]. Các phương pháp đồng hóa cổ điển như nội suy tối ưu (*optimum interpolation*) hay hiệu chỉnh liên tiếp (*successive correction method*) đã được thay thế bằng các kỹ thuật hiện đại với cách tiếp cận biến phân. Ví dụ như phương pháp đồng hóa biến phân 3 chiều (3DVar), 4 chiều (4DVar) và các biến thể lọc Kalman tổ hợp (EnKF) [17]. Trong đó, phương pháp 4DVar yêu cầu nhiều tài nguyên tính toán do số liệu quan trắc được cập nhật theo thời gian đòi hỏi phải sử dụng mô hình tiếp tuyến và liên hợp để tính toán sự phù hợp của các trường phân tích trên của số đồng hóa [7, 8]. Do đó, việc ứng dụng phương pháp 4DVar hiện tại chỉ được thực hiện ở một số trung tâm dự báo lớn, nơi có hệ thống tính toán rất mạnh, như Trung tâm dự báo hạn vừa của Châu Âu, Cơ quan khí tượng Nhật Bản, và Cơ quan khí tượng của Pháp [11].

Với mức độ phát triển cao, các phương pháp

¹Viện Khoa học Khí tượng Thủy văn và Biến đổi khí hậu
Email: vwthang26@gmail.com

đồng hóa số liệu đòi hỏi nhiều bước tính toán cũng như quá trình xử lý số liệu nhiều bước [13]. Một trong những ví dụ điển hình là hệ thống đồng hóa số liệu tích hợp của mô hình WRF (*Weather Research and Forecasting model*), được gọi là WRF-DA (*WRF-Data Assimilation System*). Ở Việt Nam, một số nghiên cứu về đồng hóa số liệu như của Trần Tân Tiến và Nguyễn Thị Thanh (2011) đã sử dụng WRF-DA để thử nghiệm đồng hóa số liệu vệ tinh MODIS bằng phương pháp 3DVar để dự báo mưa lớn ở khu vực Trung Bộ [15]; Trần Tân Tiến và cs (2013) cũng đã ứng dụng phương pháp lọc Kalman tổ hợp vào dự báo cường độ bão 5 ngày [16]. Các kết quả cho thấy chất lượng của dự báo ở các hạn gần đã tốt hơn so với trường hợp không đồng hóa. Tuy nhiên, việc áp dụng WRF-DA còn chưa rộng rãi ở Việt Nam do hạn chế về tài nguyên tính toán cũng như nguồn nhân lực để tiếp cận với các phương pháp đồng hóa phức tạp.

Một trong những nỗ lực để giúp việc nghiên cứu bài toán đồng hóa được đồng bộ và hiệu quả hơn là sự phát triển của hệ thống đồng hóa GSI (*Gridpoint Statistical Interpolation*) [6]. Hệ thống GSI cho phép vận hành với nhiều loại số liệu quan trắc và trường nền từ nhiều mô hình số trị khác nhau. Trải qua hơn một thập kỷ phát triển, hệ thống GSI có thể vận hành với nhiều phương pháp đồng hóa, bao gồm cả đồng hóa lai tổ hợp. Phương pháp đồng hóa lai tổ hợp biến phân 3 chiều được đề xuất lần đầu tiên bởi Hamill và Snyder (2000) [5], để tận dụng những ưu điểm và hạn chế các nhược điểm trong phương pháp đồng hóa biến phân. Một số nghiên cứu sử dụng hệ thống GSI với đồng hóa lai đã đưa ra kết luận rằng phương pháp này cho chất lượng dự báo các trường tốt hơn [10, 12, 17]. Tuy nhiên, theo hiểu biết của nhóm tác giả, hệ thống GSI hiện tại chưa được nghiên cứu và thử nghiệm ở Việt Nam. Do đó, mục đích của nghiên cứu này là lần đầu tiên thử nghiệm áp dụng hệ thống đồng hóa số liệu GSI trong bài toán dự báo mưa ở Việt Nam. Bước đầu, Nam Bộ được chọn làm khu vực nghiên cứu với đặc trưng của khí hậu nhiệt đới gió mùa và cận xích đạo. Mùa mưa

ở Nam Bộ kéo dài từ tháng 5 đến tháng 11, đóng góp trên 70% tổng lượng mưa cả năm [9]. Vào mùa mưa, thường hay xuất hiện những cơn mưa có cường độ lớn trên một số địa phương trong vùng. Đặc biệt, tác động của biến đổi khí hậu gây ra những cơn mưa trái mùa với lượng mưa lớn chưa từng có trong lịch sử. Khi mưa kết hợp với triều cường và lũ sẽ gây ngập úng, ảnh hưởng đến sản xuất và đời sống của dân cư trong vùng. Bên cạnh đó, Nam Bộ có ba hồ thủy lợi lớn và quan trọng là Dầu Tiếng, Trị An và Phước Hòa, có vai trò điều tiết lũ. Do đó, nghiên cứu dự báo mưa cho Nam Bộ có ý nghĩa cả về mặt khoa học và xã hội.

Trong nghiên cứu này, hệ thống GSI sẽ được thử nghiệm với phương pháp đồng hóa 3DVar và đồng hóa lai tổ hợp biến phân 3 chiều (*3DHybEnVar*) cùng với ba loại số liệu quan trắc khác nhau. Mô tả về hệ thống GSI, nguồn số liệu và thiết kế thí nghiệm được trình bày ở phần 2, kết quả và thảo luận ở phần 3, và kết luận ở phần 4.

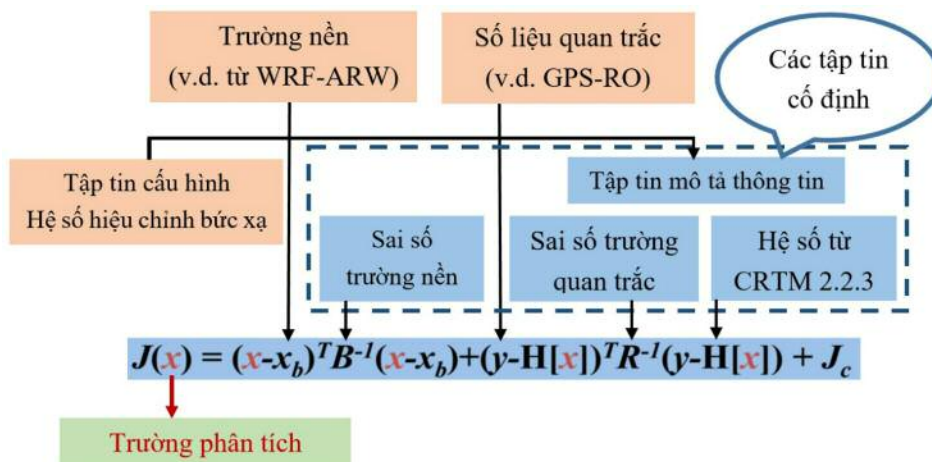
2. Hệ thống GSI, nguồn số liệu và thiết kế thí nghiệm

2.1a Giới thiệu về GSI

Hệ thống đồng hóa GSI ban đầu được phát triển bởi Trung tâm Mô hình hóa Môi trường (*EMC*) trực thuộc Trung tâm Dự báo Môi trường Quốc gia Hoa Kỳ (*NCEP*) [6]. Sau khi được đưa vào sử dụng cho nghiệp vụ tại Hệ thống đồng hóa số liệu tại Bắc Mỹ (*NDAS*) vào tháng 6 năm 2006, hệ thống GSI đã được ứng dụng cho nhiều hệ thống nghiệp vụ khác (v.d. Hệ thống đồng hóa số liệu toàn cầu GDAS, Cơ quan Hàng không và Vũ trụ quốc gia Hoa Kỳ - NASA, hệ thống dự báo bão HWRF). Năm 2007, Trung tâm Thử nghiệm và Phát triển của Hoa Kỳ (*DTC*), kết hợp với nhóm phát triển chính của GSI, đã tiến hành chuyển đổi hệ thống nghiệp vụ GSI sang dạng mã nguồn mở, cũng như hỗ trợ việc phát triển và phân phối [6]. Phiên bản nguồn mở đầu tiên của hệ thống GSI được công bố vào năm 2009. Từ đó cho đến nay đã có nhiều bản nâng cấp được phát hành, với các phương pháp và kỹ thuật đồng hóa mới, cũng như hỗ trợ thêm nhiều loại số liệu quan trắc khác nhau.

Hệ thống GSI cho phép đồng hóa số liệu theo phương pháp biến phân 2 chiều (*2DVar*), biến phân 3 chiều (*3DVar*), tổ hợp biến phân 3 chiều (*3DEnVar*), biến phân 4 chiều (*4DVar*) nếu chạy kết hợp với một mô hình liên hợp, tổ hợp biến phân 4 chiều (*4DEnVar*), lai tổ hợp (*3D/4D hybrid EnVar*) và lọc Kalman tổ hợp [6]. Hệ thống GSI hỗ trợ việc đồng hóa với trường nền từ năm loại mô hình bao gồm (1) mô hình WRF với cả hai lõi động lực ARW và NMM, cũng như mô-đun hóa khí quyển WRF-Chem, (2) mô hình dự

báo toàn cầu GFS, (3) mô hình mô phỏng môi trường NEMS, (4) hệ thống phân tích thời gian thực quy mô vừa RTMA và (5) mô hình chất lượng không khí CMAQ. Các loại số liệu có thể được đồng hóa qua hệ thống GSI rất đa dạng về quy mô bao gồm (1) các số liệu quan trắc truyền thống, (2) số liệu quan trắc và ước lượng từ vệ tinh và (3) các loại số liệu khác như quan trắc từ radar hay GPS-RO (*Global Positioning System - Radio Occultation*) [6].



Hình 1. Mô tả các tập tin đầu vào và đầu ra của hệ thống GSI

Hình 1 mô tả các tập tin đầu vào và đầu ra yêu cầu bởi hệ thống GSI. Trong hầu hết các trường hợp, ba loại số liệu đầu vào cần có bao gồm: (1) trường nền, (2) số liệu quan trắc và (3) các tập tin cố định chứa thông tin về sai số trường nền và trường quan trắc, cũng như hệ số hiệu chỉnh bức xạ. Trong trường hợp chạy thử nghiệm với một số liệu quan trắc giả (lý tưởng), hệ thống GSI không cần số liệu quan trắc. Trong trường hợp chạy 3D hybrid EnVar, cần có sản phẩm dự báo tổ hợp cho toàn cầu hoặc khu vực. Trường nền có thể lấy từ đầu ra của mô hình dự báo chạy độc lập hoặc từ bước đồng hóa trước đó. Số liệu quan trắc có thể lấy từ nhiều nguồn nhưng cần được đưa về định dạng PrepBUFR hoặc BUFR [6]. Trường phân tích được tạo ra bởi hệ thống GSI sẽ được sử dụng để làm trường ban đầu cho bước dự báo tiếp theo của mô hình.

2.1b Ưu điểm của GSI

Có thể nhận thấy, với cấu trúc mang tính phổ

quát cao, hệ thống GSI có thể sử dụng với nhiều loại số liệu cũng như hỗ trợ đồng hóa cho nhiều loại mô hình [11]. Do đó, hệ thống GSI giúp giảm bớt việc phát triển mới hệ thống đồng hóa riêng biệt cho mỗi mô hình. Với cấu trúc nhất quán, hệ thống GSI cho phép ứng dụng ở cả quy mô khu vực và toàn cầu. Một số nghiên cứu đã đánh giá sản phẩm dự báo của hệ thống GSI cho Bắc bán cầu [17], cho khu vực như Bắc Mỹ [10] hay cho Ấn Độ [12]. Hơn thế nữa, hệ thống GSI cung cấp một nền tảng chuẩn cho việc đánh giá và so sánh một cách có hệ thống các mô hình cũng như tác động của các loại số liệu trong bài toán đồng hóa số liệu [10]. Singh và Prasad (2018) đã sử dụng GSI để so sánh khả năng dự báo trong bốn tháng mùa hè năm 2015 tại Ấn Độ giữa hai phương pháp 3DVar và 3D hybrid EnVar [12]. Đánh giá qua các chỉ số thống kê cho thấy phương pháp 3D hybrid EnVar giảm sai số thiên dương trong dự báo mưa trên số khu

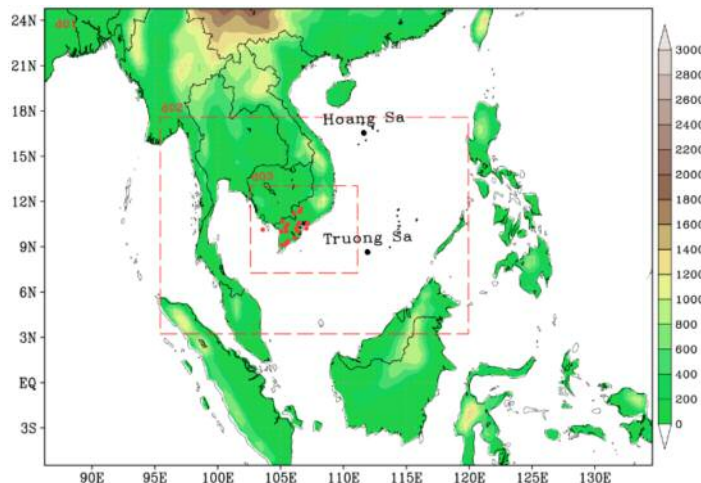
vực của Ấn Độ. Khi phát triển phương pháp tổ hợp lai *3DVar hybrid EnVar* cho hệ thống GSI, Wang và ccs (2013) đã cho thấy tính ứng dụng cao của hệ thống này trong việc nghiên cứu sâu bài toán đồng hóa [17]. Các thí nghiệm như so sánh tốc độ hội tụ của quá trình cực tiểu hóa hay đánh giá mức độ hấp thụ số liệu vào trường phân tích có thể được thực hiện trên nền tảng của GSI. Ứng dụng hệ thống GSI được phổ biến như vậy một phần nhờ GSI tính toán trên không gian điểm lưới thay vì không gian phổ, giúp cho việc áp dụng hiệp phương sai bất đồng nhất dị hướng dễ dàng hơn [11]. Với những ưu điểm này, việc ứng dụng hệ thống GSI ở những trung tâm dự báo tại các quốc gia đang phát triển rất nên được xem xét.

2.2 Nguồn số liệu

Ba loại số liệu quan trắc được sử dụng trong việc đồng hóa bao gồm: (1) số liệu quan trắc tại 19 trạm khí tượng bề mặt (các biến nhiệt độ, khí áp, độ ẩm tuyệt đối, và gió), (2) số liệu quan trắc cao không tại trạm Hồ Chí Minh và (3) số liệu vệ tinh GPS-RO. Các trạm quan trắc bề mặt phân bố ở khu vực Nam Bộ với vị trí được chỉ ra bằng chấm đỏ trên hình 2. GPS-RO là số liệu độ phản

hồi sóng vô tuyến giữa vệ tinh GPS và vệ tinh LEO (low-Earth-orbiting) [4], lý thuyết xây dựng trường số liệu này là sự khúc xạ của tín hiệu GPS khi được truyền xuyên qua khí quyển, và qua thực nghiệm đã xây dựng được một phương trình mà có sự liên hệ với các trường nhiệt độ, khí áp, và áp suất hơi nước [4]. Một điểm đáng lưu ý là GSI yêu cầu số liệu quan trắc dưới định dạng mã chuẩn *BUFR/PrepBUFR* [6] nên trong nghiên cứu này một chương trình Fortran đã được viết nhằm phục vụ quá trình chuyển đổi định dạng số liệu. Số liệu dự báo toàn cầu *Global Forecast System (GFS)* với độ phân giải ngang 0,5x0,5 độ kinh vĩ, 32 mực thẳng đứng, và 84 giờ dự báo, được sử dụng làm điều kiện ban đầu và điều kiện biên cho mô hình WRF. Số liệu được tải về từ địa chỉ <https://www.ncdc.noaa.gov/data-access/model-data/model-datasets/global-forecast-system-gfs>.

Để đánh giá kết quả dự báo mưa của mô hình trong các trường hợp thử nghiệm, số liệu mưa ước lượng từ sản phẩm vệ tinh GSMaP [14] và số liệu quan trắc lượng mưa ngày của 19 trạm tương ứng được sử dụng.



Hình 2. Độ cao địa hình (m), miền tính và vị trí các trạm quan trắc

2.3 Thiết kế thí nghiệm

Ứng dụng hệ thống GSI phiên bản 3.6 với hai phương pháp đồng hóa số liệu *3DVar* và *3D hybrid EnVar*, kết hợp với ba loại số liệu quan trắc khác nhau, bốn trường hợp thử nghiệm đã được tiến hành như mô tả trong bảng 1. Trong đó, hai thí nghiệm với phương pháp *3D hybrid EnVar*

sử dụng số liệu dự báo tổ hợp 21 thành phần của hệ thống dự báo tổ hợp toàn cầu (*Global Ensemble Forecast System – GEFS*). Số liệu này có thể tải về từ địa chỉ: <https://www.ncdc.noaa.gov/data-access/model-data/model-datasets/global-ensemble-forecast-system-gefs>

Bảng 1. Mô tả các thí nghiệm được thực hiện

Tên thí nghiệm	Mô tả	Số liệu dùng để đồng hóa
NoDA	Không đồng hóa	Không có
3DVar	Đồng hóa biến phân 3 chiều	Quan trắc tại trạm và cao không
3DHybEnVar	Đồng hóa lai tổ hợp biến phân 3 chiều	Quan trắc tại trạm, cao không và GPS-RO
3DHybEnVar_GPS-RO		

Trường nền được tạo bởi mô hình WRF phiên bản 3.9.1 với 3 miền tính, như minh họa trên Hình 2. Tâm miền tính ở 10 độ vĩ Bắc và 110,4 độ kinh Đông. Độ phân giải ngang tương ứng của 3 miền là 54 km (100x65 điểm lưới), 18 km (151x91 điểm lưới), 6 km (157x109 điểm lưới) với 36 mực thẳng đứng. Bộ tham số của mô hình được lựa chọn gồm lớp biên hành tinh MYJ, sơ đồ đối lưu Grell-Devenyi, vi vật lý mây WSM6, sơ đồ bức xạ RRTM và sơ đồ đất của NCEP [13]. Năm đợt mưa lớn xảy ra trên khu vực Nam Bộ được thử nghiệm dự báo với hạn dự báo 84 giờ (thời điểm khởi chạy là 12Z) bao gồm: 01-03/08/2016, 09-13/09/2016, 20-22/09/2016, 24-26/09/2016, và 23-26/10/2016.

Các chỉ số thống kê đánh giá dự báo pha được sử dụng để đánh giá kết quả dự báo của mô hình khi so sánh với số liệu quan trắc tại trạm [3]. Trong đánh giá dự báo pha, bảng ngẫu nhiên được xây dựng để đánh giá sự phù hợp giữa sự xảy ra hiện tượng trong dự báo và quan trắc.

Bảng 2. Bảng ngẫu nhiên của pha dự báo
Quan trắc

		Quan trắc	
		Có	Không
Dự báo	Có	H	F
	Không	M	CN

Từ bảng 2, năm chỉ số đánh giá được tính toán bao gồm:

$$FBI = \frac{H + F}{H + M} \quad (1)$$

$$POD = \frac{H}{H + M} \quad (2)$$

$$FAR = \frac{F}{H + F} \quad (3)$$

$$CSI = TS = \frac{H}{M + F + H} \quad (4)$$

$$PC = \frac{H + CN}{M + F + H + CN} \quad (5)$$

Chỉ số FBI (hay Bias score) đánh giá tỷ số giữa số trường hợp dự báo có xảy ra so với số trường hợp thực tế quan trắc có xảy ra. FBI đạt giá trị lý tưởng khi FBI bằng 1. Trong khi đó, FBI < 1 (hoặc > 1) cho thấy mô hình dự báo ít hơn (hoặc nhiều hơn) số lần hiện tượng xảy ra so với quan trắc. Xác suất phát hiện POD (Probability of Detection) cho biết khả năng dự báo thành công của mô hình, có giá trị nằm trong khoảng [0,1]. Giá trị POD càng gần 1 thể hiện chất lượng dự báo tốt càng tốt. Tỷ số dự báo sai FAR (False Alarms Ratio) cho biết tỷ lệ dự báo không của mô hình (mô hình cho kết quả có nhưng thực tế hiện tượng không xảy ra). Giá trị FAR biến đổi trong khoảng [0,1], với giá trị tối ưu bằng 0. Điểm số thành công CSI (Critical Success Index hay Threat Score - TS) và chỉ số dự báo đúng PC (Percentage Correct) có biên thiên từ 0 đến 1. Trong đó, CSI (hoặc PC) bằng 0 nghĩa là mô hình không có kỹ năng và CSI (hoặc PC) bằng 1 mô hình là hoàn hảo. Giá trị của PC càng lớn thì số lần dự báo đúng của mô hình càng cao.

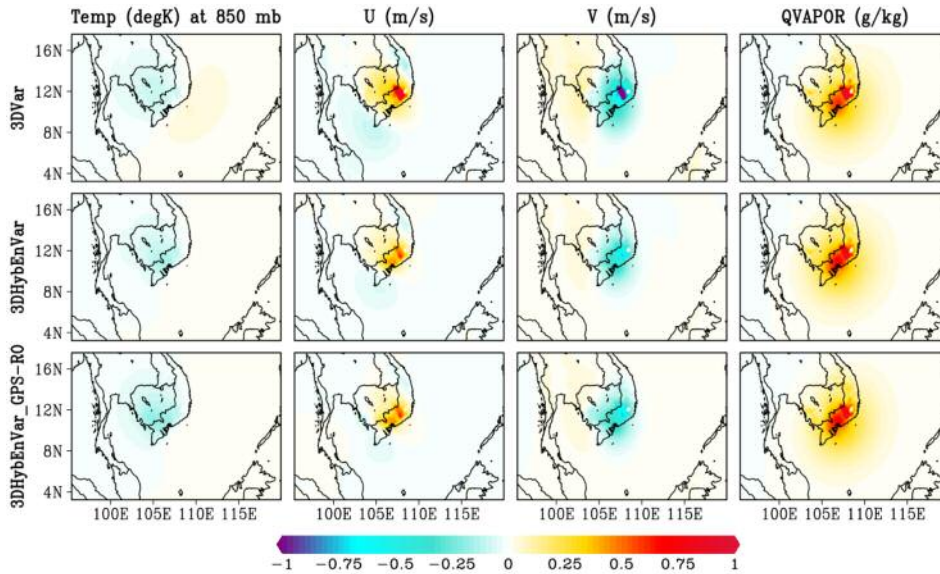
3. Kết quả và thảo luận

3.1 Đánh giá trường ban đầu

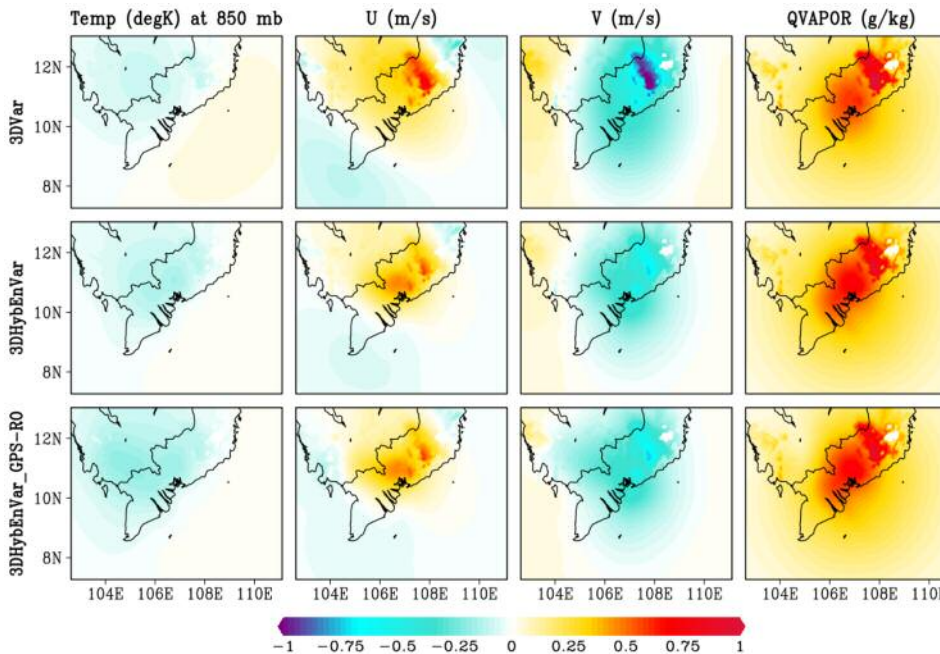
Gia số của trường ban đầu trong các trường hợp đồng hóa so với không đồng hóa được thể hiện trên hình 3 và hình 4 tương ứng với miền tính thứ hai và thứ ba. Các biến được đánh giá tại mực 850 mb bao gồm nhiệt độ, tốc độ gió kinh vĩ hướng và tỷ lệ xáo trộn hơi nước. Nhìn chung, sự thay đổi rõ rệt hơn cả của trường phân tích nằm ở khu vực Lâm Đồng và Đồng Nai (v.d. tốc độ gió kinh vĩ hướng chênh lệch hơn 1 m/s). Đối với biến tỷ lệ xáo trộn hơi nước, khu vực có sự thay đổi lan rộng hơn về phía thành phố Hồ Chí Minh và Long An. Sự thay đổi của các trường phân tích có thể do vị trí có số liệu quan trắc

được đồng hóa, kết hợp với bán kính ảnh hưởng được xác định bởi thuật toán đồng hóa. Trong trường hợp sử dụng phương pháp 3DVar, tồn tại khu vực có sự tăng cục bộ rõ nét hơn vùng xung quanh. Trong khi đó, trường hợp đồng hóa lai có

sự hài hòa hơn, khi độ lớn trường gia số không chênh lệch cao giữa các vùng. Trường hợp 3DHybEnVar và 3DHybEnVar_GPS-RO không có sự khác nhau nhiều. Điều này có thể là do các trạm GPS ở xa khu vực đồng hóa trong miền tính.



Hình 3. Hiệu các trường ban đầu của các trường hợp đồng hóa và không đồng hóa tại thời điểm 12Z 24/10/2016, của miền tính thứ hai (d02)



Hình 4. Tương tự như hình 3 nhưng cho miền tính thứ ba (d03)

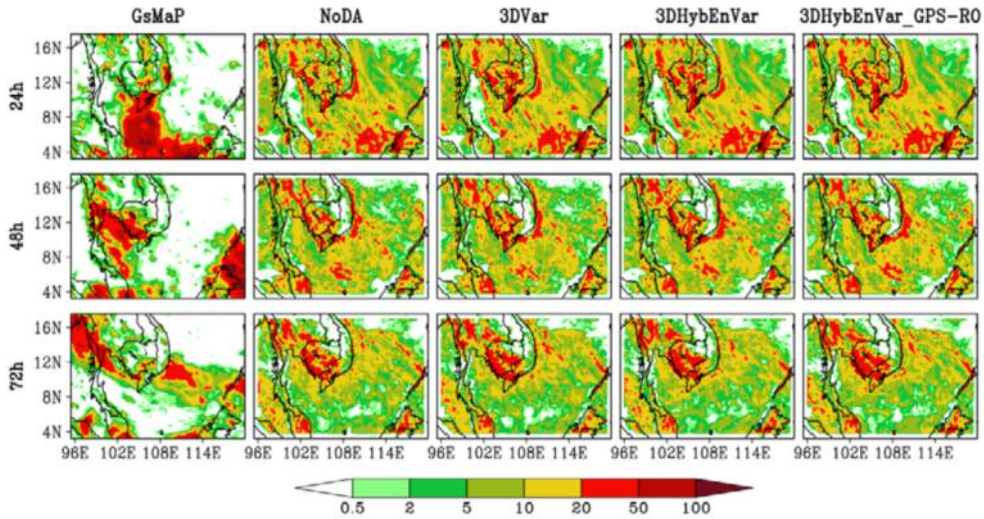
3.2 Đánh giá trường mưa

Kết quả dự báo lượng mưa tích lũy 24h của mô hình WRF chạy không đồng hóa và các trường hợp đồng hóa tại thời điểm 12Z của ngày 24/10/2016 theo 3 hạn dự báo 24h, 48h và 72h

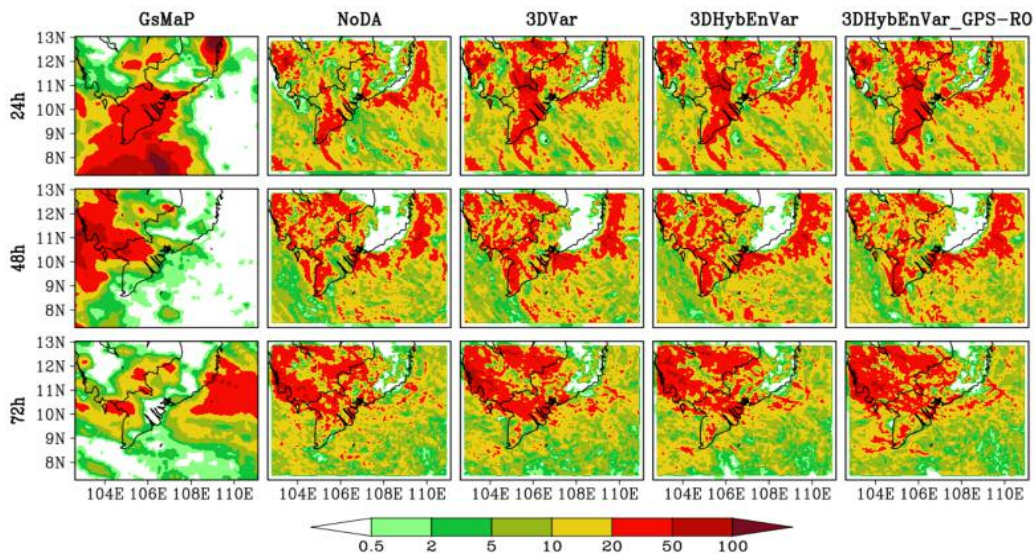
được so sánh với số liệu mưa vệ tinh GSMaP (Hình 5 và Hình 6). Nhìn chung, khi so sánh các trường hợp đồng hóa và không đồng hóa, có thể nhận thấy rằng đã có sự khác nhau về diện mưa dự báo ở xung quanh những vị trí có số liệu quan

trắc được đồng hóa. Ở hạn dự báo 24h, vùng mưa dự báo của các trường hợp đồng hóa có xu hướng mở rộng ra, có thể nhìn thấy tại các vùng lân cận thành phố Tân An, Kiến Tường, Cần Thơ, và Rạch Giá. Trong đó, trường hợp 3DHybEnVar_GPS-RO có sự mở rộng nhiều nhất và

khá phù hợp với GSMaP (Hình 6). Ở hạn dự báo 48h, tâm mưa lớn có xu hướng dịch chuyển sang phía đông và trường hợp 3DHybEnVar_GPS-RO cũng thể hiện điều này rõ hơn. Tuy nhiên, ở hạn dự báo 72h, giữa các trường hợp đồng hóa không còn sự khác nhau nhiều (Hình 6).



Hình 5. Lượng mưa 24h của số liệu GsMaP, và dự báo của mô hình với miền tính d02 trong các trường hợp NoDA (không đồng hóa), 3DVar, 3DHybEnVar và 3DHybEnVar_GPS-RO, ở các hạn dự báo 24h, 48h và 72h, chạy tại thời điểm 12Z 24/10/2016



Hình 6. Tương tự như hình 5 nhưng cho miền tính d03

3.3 Khả năng dự báo mưa định lượng

Lượng mưa dự báo được phân thành ba cấp mưa (mưa vừa, mưa to và mưa rất to) tương ứng với các ngưỡng lượng mưa là 16, 50 và 100 mm/24h. Các chỉ số thống kê FBI, POD, FAR, CSI và PC được tính toán cho mỗi hạn dự báo và trình bày trong Bảng 3. Nhìn chung, từ các

chỉ số thống kê, có thể thấy đối với ngưỡng mưa trên 50mm, cả bốn trường hợp thử nghiệm của mô hình đều cho dự báo với kỹ năng thấp. Nguyên nhân một phần có thể là do dự báo của mô hình cho không quá cao hoặc không quá thấp, và do số đợt mưa lớn chạy thử nghiệm có lượng mưa rơi vào các ngưỡng mưa trên 50 mm

và 100 mm chưa nhiều, do đó, các phân tích sau đây chỉ đánh giá cho ngưỡng mưa vừa.

Chỉ số FBI của các trường hợp, ở cả ba hạn dự báo đều cho lớn hơn 1 hay nói cách khác là số dự báo rơi vào ngưỡng mưa vừa cao hơn quan trắc. Ở hạn dự báo 24h, trường hợp 3DVar cho FBI gần 1 nhất (1,89). Với chỉ số POD và CSI thì trường hợp 3DHybEnVar_GPS-RO cho con số lớn nhất (tương ứng là 0,561 và 0,232), chỉ số PC của trường hợp 3DHybEnVar_GPS-RO cũng cho thấy dự báo tốt hơn các trường hợp đồng hóa

khác, mặc dù nhỏ hơn trường hợp không đồng hóa nhưng sự chênh lệch này là không đáng kể. Ở hạn dự báo 48h, các chỉ số cho thấy hai trường hợp của đồng hóa lai cho dự báo tốt hơn các trường hợp khác, trong đó, trường hợp 3DHybEnVar_GPS-RO có các chỉ số FBI = 1,743, FAR = 0,358, và PC = 0,544 tốt nhất. Còn ở hạn dự báo 72h thì trường hợp 3DHybEnVar cho dự báo tốt hơn, với các chỉ số thống kê tốt nhất là FBI = 2,245, FAR = 0,325, và PC = 0,614

Bảng 3. Điểm số đánh giá kỹ năng dự báo mưa ở các hạn dự báo 24h, 48h và 72h. Số thứ tự 1, 2, 3 và 4 tương ứng với các thí nghiệm NoDA, 3DVar, 3DHybEnVar và 3DHybEnVar_GPS-RO

STT	PC	16-50mm				50-100mm				> 100mm			
		FBI	POD	FAR	CSI	FBI	POD	FAR	CSI	FBI	POD	FAR	CSI
Hạn dự báo 24h													
1	0,503	1,988	0,537	0,458	0,219	0,276	0,069	0,019	0,057	0	0	0	0
2	0,485	1,89	0,5	0,438	0,209	0,724	0,172	0,051	0,111	0	0	0	0
3	0,462	2,061	0,524	0,485	0,207	0,586	0,138	0,042	0,095	0,5	0	0,006	0
4	0,497	1,976	0,561	0,446	0,232	0,621	0,172	0,042	0,119	0,25	0	0,003	0
Hạn dự báo 48h													
1	0,526	1,797	0,446	0,373	0,19	0,065	0	0,006	0	0,25	0	0,003	0
2	0,512	1,811	0,419	0,384	0,175	0,097	0	0,01	0	0	0	0	0
3	0,526	1,932	0,514	0,392	0,212	0,032	0	0,003	0	0	0	0	0
4	0,544	1,743	0,446	0,358	0,194	0,032	0	0,003	0	0	0	0	0
Hạn dự báo 72h													
1	0,573	2,283	0,358	0,353	0,123	0,167	0	0,009	0	0	0	0	0
2	0,556	2,642	0,491	0,394	0,156	0,056	0	0,003	0	0	0	0	0
3	0,614	2,245	0,472	0,325	0,17	0	0	0	0	0	0	0	0
4	0,594	2,302	0,453	0,339	0,159	0,056	0	0,003	0	0	0	0	0

4. Kết luận

Trong nghiên cứu này, hệ thống đồng hóa GSI đã được thiết lập thành công và chạy thử nghiệm cho bài toán dự báo định lượng mưa ở khu vực Nam Bộ. Đây là một hệ thống đồng hóa số liệu có tính đồng bộ cao với khả năng cập nhật nhiều loại số liệu và tương tác với trường nền từ nhiều loại mô hình. Trong nghiên cứu này, hệ thống GSI được thử nghiệm cho các phương pháp đồng hóa 3DVar và 3DHybEnVar cùng với ba loại số liệu (quan trắc bề mặt, cao không và số liệu vệ tinh GPS-RO). Năm đợt mưa lớn xảy ra vào năm 2016 được chọn để thử nghiệm dự báo. Các kết quả bước đầu cho thấy tính ưu việt của hệ thống GSI trong việc chuyển đổi số liệu quan trắc sang định dạng chuẩn PrepBUFR nên thuận

tiện hơn trong việc lưu trữ, cũng như kiểm định chất lượng của số liệu đầu vào. Thử nghiệm đồng hóa lai 3DHybEnVar_GPS-RO đã cho dự báo về diện mưa và ngưỡng mưa vừa (16-50 mm) trong các hạn dự báo 24h và 48h tốt hơn các trường hợp còn lại. Hệ thống GSI đồng hóa lai tổ hợp này kỳ vọng sẽ cải thiện chất lượng dự báo mưa cho khu vực Việt Nam nếu các nguồn số liệu địa phương được cập nhật. Tuy nhiên, cần có những nghiên cứu chuyên sâu và ở quy mô lớn hơn nhằm đưa ra những kết luận mang tính tổng hợp hơn. Bên cạnh đó, một trong những hướng nghiên cứu tiếp theo đối với hệ thống GSI là xem xét thử nghiệm phương pháp đồng hóa 4DVar và số liệu quan trắc từ radar thời tiết

Lời cảm ơn: Bài báo được hoàn thành nhờ sự hỗ trợ của Đề tài “Nghiên cứu xây dựng hệ thống nghiệp vụ dự báo định lượng mưa khu vực Nam Bộ và cảnh báo mưa lớn hạn cực ngắn cho thành phố Hồ Chí Minh”, mã số KC.08.14/16-20.

Tài liệu tham khảo

1. Bauer, H.S., Schwitalla, T., Wulfmeyer, V., Bakhshaii, A., Ehret, U., Neuper, M., & Caumont, O. (2015), Quantitative precipitation estimation based on high-resolution numerical weather prediction and data assimilation with WRF—a performance test, *Tellus A: Dynamic Meteorology and Oceanography*, 67 (1), <https://doi.org/10.3402/tellusa.v67.25047>.
2. Bauer, P., Thorpe, A., Brunet, G. (2015), The quiet revolution of numerical weather prediction. *Nature*, 525, 47-55.
3. Jolliffe, I.T., Stephenson, D.B. (2012), Forecast verification: a practitioner's guide in atmospheric science. *John Wiley & Sons*.
4. Ha, J., Lim, G., Choi, S. (2014), Assimilation of GPS Radio Occultation Refractivity Data with WRF 3DVAR and Its Impact on the Prediction of a Heavy Rainfall Event. *J. Appl. Meteor. Climatol.*, 53, 1381-1398, <https://doi.org/10.1175/JAMC-D-13-0224.1>.
5. Hamill, T.M., Snyder, C. (2000), A Hybrid Ensemble Kalman Filter–3D Variational Analysis Scheme. *Mon. Wea. Rev.*, 128, 2905-2919.
6. Hu, M., Ge, G., Shao, H., Stark, D., Newman, K., Zhou, C., Beck, J., Zhang, X. (2017), Grid-point Statistical Interpolation (GSI) User's Guide v3.6. Developmental Testbed Center, <http://www.dtcenter.org/com-GSI/users/docs/index.php>, 149pp.
7. Huỳnh Thị Hồng Ngự, La Thị Cang (2008), *Đồng hóa số liệu bằng phương pháp biến phân bốn chiều trong dự báo thời tiết bằng phương pháp số trị*. Tạp chí Khoa học và Công nghệ, 12, 98-103.
8. Mazzarella, V., Maiello, I., Capozzi, V., Budillon, G., Ferretti, R. (2017), Comparison between 3D-Var and 4D-Var data assimilation methods for the simulation of a heavy rainfall case in central Italy. *Adv. Sci. Res.*, 14, 271-278, <https://doi.org/10.5194/asr-14-271-2017>.
9. Nguyễn Đức Ngự, Nguyễn Trọng Hiệu (2013), *Khí hậu và Tài nguyên Khí hậu Việt Nam. Nhà xuất bản Khoa học Kỹ thuật, Hà Nội*, 296tr.
10. Pan, Y., Zhu, K., Xue, M., Wang, X., Hu, M., Benjamin, S.G., Weygandt, S.S., Whitaker, J.S. (2014), A GSI-Based Coupled EnSRF-En3DVar Hybrid Data Assimilation System for the Operational Rapid Refresh Model: Tests at a Reduced Resolution. *Mon. Wea. Rev.*, 142, 3756-3780, <https://doi.org/10.1175/MWR-D-13-00242.1>.
11. Shao, H., Derber, J., Huang, X.Y., Hu, M., Newman, K., Stark, D., Lueken, M., Zhou, C., Nance, L., Kuo, Y.H., Brown, B. (2016), Bridging Research to Operations Transitions: Status and Plans of Community GSI. *Bull. Amer. Meteor. Soc.*, 97, 1427-1440, doi: 10.1175/BAMS-D-13-00245.1.
12. Singh, S.K., Prasad, V.S. (2018), Evaluation of precipitation forecasts from 3D-Var and hybrid GSI-based system during Indian summer monsoon 2015. *Meteorology and Atmospheric Physics*, <https://doi.org/10.1007/s00703-018-0580-y>.
13. 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. (2008), A description of the Advanced Research WRF v3. NCAR *Technical Note NCAR/TN-475CSTR*.
14. Kubota, T., Shige, S., Hashizume, H., Aonashi, K., Takahashi, N., Seto, S., Hirose, M., Takayabu, Y.N., Nakagawa, K., Iwanami, K., Ushio, T., Kachi, M., Okamoto, K. (2007), Global Precipitation Map using Satelliteborne Microwave Radiometers by the GSMaP Project: Production and

Validation, *IEEE Trans. Geosci. Remote Sens.*, 45 (7), 2259-2275.

15. Trần Tân Tiến, Nguyễn Thị Thanh (2011), *Đồng hóa dữ liệu vệ tinh MODIS trong mô hình WRF để dự báo mưa lớn ở khu vực Trung Bộ*. Tạp chí Khoa học ĐHQGHN, Khoa học Tự nhiên và Công nghệ, 27, tr. 90-95.

16. Trần Tân Tiến, Hoàng Thị Mai, Công Thanh (2013), *Ứng dụng phương pháp lọc Kalman tổ hợp vào dự báo cường độ bão 5 ngày*. Tạp chí khoa học Đại học Quốc gia Hà Nội. Khoa học Tự nhiên và Công nghệ, 29, 201-206.

17. Wang, X., Parrish, D., Kleist, D., Whitaker, J. (2013), GSI 3DVar-Based Ensemble-Variational Hybrid Data Assimilation for NCEP Global Forecast System: Single-Resolution Experiments. *Mon. Wea. Rev.*, 141, 4098-4117, <https://doi.org/10.1175/MWR-D-12-00141.1>.

AN EXPERIMENT WITH GSI SYSTEM FOR DATA ASSIMILATION TO IMPROVE QUANTITATIVE RAINFALL FORECAST OVER SOUTHERN VIETNAM REGION

Pham Quang Nam¹, Mai Van Khiem¹, Nguyen Quang Trung¹, Vu Van Thang¹

¹Viet Nam institute of Meteorology, Hydrology and climate change

Abstract: *This study represents some experiment results of data assimilation using GSI (Grid-point Statistical Interpolation) system, with the aim to improve quantitative rainfall forecast of WRF (Weather Research Forecasting) model over the Southern region. Observed datasets of surface, sounding and GPS-RO satellite data were collected and converted to the PrepBUFR format which is required by GSI. The GSI is run with cases of 3DVar assimilation and 3DHybEnVar, as well as combined with data observations, along with the non-assimilated to create four cases for the experiment. The results show that 3DHybEnVar with all the observed datasets (3DHybEnVar_GPS-RO) has forecasted the spatial distribution of rainfall and medium threshold (16-50 mm 24h-1h) at 24h and 48h lead times are better than others.*

Keywords: *GSI, 3DVar, Hybrid-EnVar, Rainfall forecast, Southern Vietnam.*