

ĐỒNG HÓA SỐ LIỆU BẰNG PHƯƠNG PHÁP BIẾN PHÂN BỐN CHIỀU TRONG DỰ BÁO THỜI TIẾT BẰNG PHƯƠNG PHÁP SỐ TRỊ

Huỳnh Thị Hồng Ngự, La Thị Cang

Trường Đại học Khoa Học Tự Nhiên, ĐHQG –HCM

(Bài nhận ngày 29 tháng 03 năm 2007, hoàn chỉnh sửa chữa ngày 24 tháng 02 năm 2008)

TÓM TẮT: Đồng hóa số liệu trong dự báo thời tiết bằng phương pháp số trị là một quy trình phức tạp và ảnh hưởng rất lớn đến chất lượng dự báo. Trong quy trình này các số liệu quan trắc được kết hợp với kết quả dự báo ngắn hạn trước đây của mô hình để tạo ra điều kiện ban đầu cho một phiên dự báo mới.

Bài báo giới thiệu: kỹ thuật đồng hóa biến phân bốn chiều (4D-Var), là phương pháp đồng hóa số liệu tiên tiến nhất hiện nay; xem xét khả năng áp dụng đồng hóa số liệu trong dự báo thời tiết bằng các mô hình số trị và hướng phát triển trong tương lai của kỹ thuật này ở Việt Nam.

Từ khóa: Đồng hóa biến phân bốn chiều, dự báo thời tiết bằng số toán.

1. GIỚI THIỆU

Đối với dự báo thời tiết bằng phương pháp số trị thì độ chính xác của số liệu đầu vào là rất quan trọng (giả sử mô hình là hoàn hảo). Vì vậy, nếu chúng ta có số liệu đầu vào tốt thì sẽ cho kết quả dự báo tốt và ngược lại. Để cải thiện độ chính xác của số liệu đầu vào người ta dùng các kỹ thuật đồng hóa số liệu (data assimilation). Đồng hóa số liệu trong dự báo thời tiết bằng phương pháp số trị là một quy trình hết sức phức tạp và đòi hỏi khả năng tính toán cao. Trong quá trình đó, các số liệu thu được từ quan trắc và các số liệu có được từ kết quả dự báo ngắn trước đó bằng mô hình sẽ được kết hợp lại với nhau (nhưng vẫn thỏa mãn các ràng buộc về động lực) để tìm ra được kết quả tốt nhất làm điều kiện ban đầu cho một dự báo mới. Hiện nay, kỹ thuật đồng hóa biến phân bốn chiều, thường được viết tắt là 4D-Var (Four Dimensional VARIational data assimilation), là kỹ thuật đồng hóa tiên tiến và phức tạp nhất. Tuy nhiên, để áp dụng 4D-Var trong thực tế thì các cơ quan khí tượng cần có hệ thống máy tính rất mạnh. Vì vậy, hiện tại chỉ có một số trung tâm phát triển mạnh về dự báo số trị trên thế giới như Trung tâm dự báo hạn vừa của Châu Âu (ECMWF - European Centre for Medium-range Weather Forecasts), Cơ quan Khí tượng của Nhật, Cơ quan Khí tượng của Pháp (Météo-France)... mới có đủ khả năng sử dụng 4D-Var cho mô hình toàn cầu.

Bài báo này giới thiệu kỹ thuật 4D-Var đang được thực hiện ở ECMWF, đồng thời cũng đề cập đến mức độ phát triển của đồng hóa số liệu trong hệ thống dự báo thời tiết ở nước ta hiện nay.

2. ĐỒNG HÓA SỐ LIỆU BẰNG PHƯƠNG PHÁP BIẾN PHÂN BỐN CHIỀU

Để đồng hóa số liệu trong dự báo thời tiết bằng phương pháp số trị chúng ta có thể sử dụng một số phương pháp khác nhau như: hiệu chỉnh liên tiếp, nudging, nội suy tối ưu, lọc Kalman, đồng hóa biến phân ba chiều (3D-Var), bốn chiều (4D-Var)... Trong các phương pháp này chỉ có nudging là phương pháp xử lý trực tiếp trong mô hình, các phương pháp còn lại chỉ là các trường hợp riêng của bài toán biến phân. Kỹ thuật biến phân trong đồng hóa số liệu được sử dụng trong dự báo nghiệp vụ từ giữa những năm 1990.

Các thuật toán đồng hóa bằng phương pháp biến phân có trình tự như sau [7]:

• Bắt đầu từ một dự đoán: dự báo ngắn sáu giờ của mô hình số trị (hiện nay ECMWF đã mở rộng cửa sổ đồng hóa lên mười hai giờ);

• Hiệu chỉnh dự đoán trên bằng cách sử dụng thông tin từ các quan trắc. Kết quả cuối cùng là sự thỏa hiệp tốt nhất giữa các quan trắc và dự đoán.

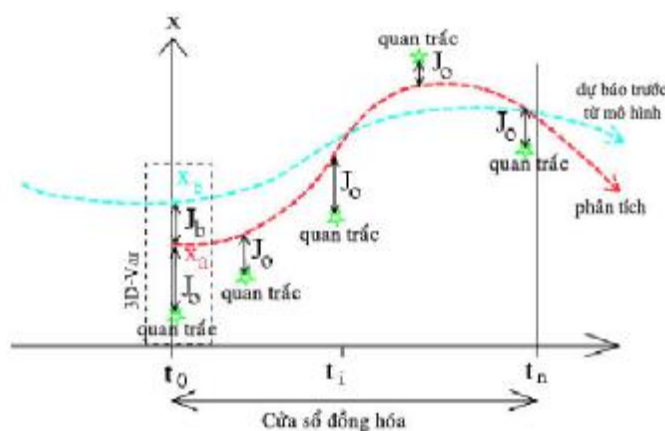
Dự đoán cung cấp $\dot{\mathbf{X}}_b$, quan trắc cung cấp $\dot{\mathbf{Y}}$. Lời giải $\dot{\mathbf{X}}$ (giá số được ký hiệu $d\dot{\mathbf{X}} = \dot{\mathbf{X}} - \dot{\mathbf{X}}_b$) phải cực tiểu hóa hàm giá trị (cost function):

$$J(\dot{\mathbf{X}}) = J_b(\dot{\mathbf{X}}) + J_o(\dot{\mathbf{X}}) + J_c(\dot{\mathbf{X}}) \quad (1)$$

trong đó J_b là độ lệch giữa dự đoán và phân tích, J_o là độ lệch giữa quan trắc và phân tích, J_c là ràng buộc của phép lọc do chúng ta chọn.

Công thức giá số $d\dot{\mathbf{X}}$ thường được sử dụng hơn. Thuật toán đồng hóa tìm $d\dot{\mathbf{X}}$, sau đó cộng $d\dot{\mathbf{X}}$ vào $\dot{\mathbf{X}}_b$.

Trong thuật toán 4D-Var, chúng ta cần lưu ý rằng các quan trắc được thực hiện ở những thời điểm khác nhau. Điểm khác biệt chính của 4D-Var so với 3D-Var là 4D-Var có tính đến thời gian thực của các quan trắc (Hình 1). Vì vậy, để sử dụng phương pháp này trong nghiệp vụ đòi hỏi khả năng tính toán của hệ thống máy tính cao hơn và thời gian tính lâu hơn so với kỹ thuật 3D-Var.



Hình 1. Sơ đồ minh họa sự khác nhau giữa 4D-Var và 3D-Var.

Phần J_o của hàm giá trị có dạng:

$$J_o = \frac{1}{2} \sum_{n_{\text{step}}=0}^{n_{\text{step}}=N_{\text{stop}}} \left(H[\delta\dot{\mathbf{X}}]_{n_{\text{step}}} - \left(\dot{\mathbf{Y}}_{n_{\text{step}}} - H[\dot{\mathbf{X}}_b]_{n_{\text{step}}} \right) \right)^T \left[\mathbf{R}^{-1} \right]_{n_{\text{step}}} \left(H[\delta\dot{\mathbf{X}}]_{n_{\text{step}}} - \left(\dot{\mathbf{Y}}_{n_{\text{step}}} - H[\dot{\mathbf{X}}_b]_{n_{\text{step}}} \right) \right) \quad (2)$$

trong đó H là toán tử được sử dụng trong việc nội suy từ các điểm nút của mô hình về các vị trí quan trắc, $\mathbf{R}$ là ma trận sai số của các quan trắc.

$$[\delta\dot{\mathbf{X}}]_{n_{\text{step}}} = \mathbf{M}_{n_{\text{step}}} [\delta\dot{\mathbf{X}}]_0 \quad (3)$$

$\mathbf{M}$ là ma trận của mô hình tiếp tuyến tuyến tính (the tangent linear model). $\left[\mathbf{X}_b\right]_{n_{step}}$ là dự báo trực tiếp ở các bước thời gian dự đoán $\left[\mathbf{X}_b\right]_0$.

J_0 có thể được viết lại dưới dạng:

$$J_0 = \frac{1}{2} \sum_{n_{step}=0}^{n_{step}=N_{stop}} \left(\mathbf{H}\mathbf{M}_{n_{step}} \left[\delta \mathbf{X} \right]_0 - \left(\mathbf{Y}_{n_{step}} - \mathbf{H} \left[\mathbf{X}_b \right]_{n_{step}} \right) \right)^T \left[\mathbf{R}^{-1} \right]_{n_{step}} \left(\mathbf{H}\mathbf{M}_{n_{step}} \left[\delta \mathbf{X} \right]_0 - \left(\mathbf{Y}_{n_{step}} - \mathbf{H} \left[\mathbf{X}_b \right]_{n_{step}} \right) \right) \quad (4)$$

Phần J_b của hàm giá trị có dạng:

$$J_b = \frac{1}{2} \left(\left[\delta \mathbf{X} \right]_0 \right)^T \mathbf{B}^{-1} \left(\left[\delta \mathbf{X} \right]_0 \right) \quad (5)$$

trong đó $\mathbf{B}$ là ma trận hiệp phương sai của sai số của dự đoán từ mô hình.

Trong 4D-Var, J_c có thể khác không khi chúng ta chọn ban đầu hóa mode chuẩn - NMI (Normal Mode Initialization) hay ban đầu hóa theo phương pháp lọc số - DFI (Digital Filter Initialization).

3. CÁC KẾT QUẢ ĐẠT ĐƯỢC KHI SỬ DỤNG PHƯƠNG PHÁP 4D-VAR TRONG MÔ HÌNH TOÀN CẦU

Ngày 25/11/1997, ECMWF là cơ quan khí tượng đầu tiên trên thế giới áp dụng kỹ thuật đồng hóa biến phân bốn chiều vào trong dự báo thời tiết. Đây là kết quả của nhiều năm nghiên cứu và cộng tác giữa ECMWF và Météo-France. Khí tượng Pháp đã lên kế hoạch triển khai 4D-Var vào năm 2000 nhưng có vài điểm khác biệt so với phiên bản sử dụng ở ECMWF. Khí tượng Pháp là cơ quan khí tượng quốc gia đầu tiên áp dụng kỹ thuật này. ECMWF đã cải tiến đáng kể chất lượng dự báo khi chuyển đổi phương pháp đồng hóa từ 3D-Var sang 4D-Var. Điều này cho phép sử dụng các số liệu quan trắc tốt hơn, phù hợp với động lực học, dẫn đến các dự báo tốt hơn ở tất cả các phạm vi của vùng vĩ độ trung bình. Hơn nữa, một tập hợp các quá trình vật lý qui mô dưới lưới được đưa vào trong phiên bản được tuyến tính hóa của mô hình dự báo, được sử dụng để giải bài toán biến phân 4D-Var. Những quá trình vật lý này mô tả đối lưu sâu của độ ẩm, giáng thủy trên quy mô lớn, khuếch tán theo phương thẳng đứng, bức xạ sóng dài và các ảnh hưởng của địa hình đối với qui mô dưới lưới. Bởi vì chi phí tính toán quá cao, các quá trình vật lý này chỉ được đưa vào trong một phần của quá trình biến phân.

Tuy nhiên, điều này đã có tác động đáng kể đối với trường ẩm trong vùng nhiệt đới nên đã cho kết quả dự báo tốt hơn [5].

Hệ thống đồng hóa nghiệp vụ 4D-Var ở ECMWF được thực hiện như sau [1]:

- Độ phân giải bên ngoài: T511L60 (cắt cụt tam giác ở số sóng 511, 60 mức)
- Độ phân giải gia số: T159L60 (cắt cụt tam giác ở số sóng 159, 60 mức)
- Hai lần cập nhật hóa gia số:
 - o Cập nhật lần thứ nhất: 50 lần lặp với các tham số hóa vật lý đơn giản,
 - o Cập nhật lần thứ hai: 25 lần lặp với các quá trình vật lý phức tạp hơn.
- Tham số hóa vật lý suốt quá trình cập nhật lần thứ hai:
 - o Khuếch tán theo phương thẳng đứng,
 - o Ảnh hưởng của địa hình đối với các qui mô nhỏ hơn độ phân giải của mô hình,
 - o Giáng thủy trên quy mô lớn,
 - o Bức xạ sóng dài,
 - o Đối lưu sâu của độ ẩm.

• Các phân tích riêng lẻ của các đại lượng bề mặt (băng biển, nhiệt độ bề mặt biển, độ ẩm của đất, tuyết, nhiệt độ và độ ẩm ở độ cao 2 m).

Hướng phát triển trong tương lai của đồng hóa số liệu ở ECMWF là phân giải các quá trình vật lý có quy mô nhỏ hơn nữa bằng cách sử dụng thêm nhiều số liệu quan trắc (đặc biệt là các số liệu từ vệ tinh), tăng độ phân giải của mô hình và đưa thêm một số biến mới vào quá trình đồng hóa [4].

4. TÌNH HÌNH PHÁT TRIỂN CỦA ĐỒNG HÓA SỐ LIỆU Ở VIỆT NAM

Tuy dự báo thời tiết bằng phương pháp số trị đã được quan tâm nghiên cứu ở Việt Nam từ nhiều năm nay nhưng sự phát triển thật sự chỉ mới diễn ra trong vài năm gần đây do phát triển chung của công nghệ thông tin, sự đầu tư về máy tính và công nghệ. Năm 1997, chúng ta bước đầu làm quen với các sản phẩm dự báo số trị của Nhật. Từ tháng 5/2002, chúng ta sử dụng mô hình HRM (**H**igh-**r**esolution **R**egional **M**odel) của Đức trong dự báo nghiệp vụ tại Trung tâm Dự báo Khí tượng Thủy văn Trung ương. Tuy nhiên, hệ thống dự báo thời tiết bằng phương pháp số trị ở nước ta có đặc điểm là sử dụng trực tiếp các trường phân tích và dự báo từ mô hình toàn cầu GME (**G**lobal **M**odel for **E**urope - mô hình của Đức) để làm điều kiện ban đầu và điều kiện biên thay vì sử dụng các sơ đồ đồng hóa để cập nhật thêm các thông tin địa phương. Vì thế kết quả dự báo của mô hình HRM phụ thuộc tương đối nhiều vào chất lượng dự báo của mô hình GME. Qua một thời gian tìm hiểu và chạy nghiệp vụ bộ mô hình GME-HRM, các nhà dự báo khí tượng nhận thấy rằng bộ mô hình này còn có nhiều hạn chế cần được nghiên cứu cải tiến. Một phần hạn chế này là do số liệu đầu vào và phương pháp đồng hóa số liệu trong GME. Các biến bề mặt trong trường ban đầu của GME không được lấy từ số liệu quan trắc mà là kết quả phân tích nối tiếp từ khi mô hình này bắt đầu chạy (từ năm 1994) đến nay, với số liệu ban đầu lấy từ trung bình khí hậu. Ngoài ra, trường ẩm do mô hình cung cấp trong nhiều trường hợp không phù hợp với thực tế, đặc biệt trong mùa bão, do các quan trắc về mây nhận được từ vệ tinh địa tĩnh và vệ tinh cực không được cập nhật liên tục trong quá trình đồng hóa số liệu. Phương pháp đồng hóa số liệu của GME là phương pháp nội suy tối ưu ba chiều (3 **D**imensional **O**ptimal **I**nterpolation - 3D-OI) có nhược điểm là làm biến mất các nhiễu động nhỏ.

Để cải tiến trường đầu vào cho HRM, Trung tâm Dự báo Khí tượng Thủy văn Trung ương đã thử nghiệm sử dụng các trường phân tích và dự báo từ các mô hình toàn cầu khác làm trường ban đầu cho mô hình theo ba phương án sau [6]:

- Sử dụng tổ hợp hai trường đầu vào khác nhau từ GME và GSM (**G**lobal **S**pectral **M**odel - mô hình toàn cầu của Nhật): thể hiện tính ưu việt (tuy nhỏ) so với mô hình nguyên gốc HRM 28 km, 20 mức (kí hiệu HRM28) cho ngưỡng mưa lớn.

- Với đầu vào thuần túy từ mô hình toàn cầu AVN (**A**viation **S**pectral **M**odel - mô hình của Mỹ): cũng cho kết quả tương tự như mô hình nguyên gốc HRM 14 km, 31 mức (kí hiệu HRM14) đối với ngưỡng mưa nhỏ. Đối với ngưỡng mưa lớn, mô hình HRM14_AVN thể hiện tính ưu việt (tuy không lớn) so với mô hình nguyên gốc HRM14.

- Với đầu vào thuần túy từ mô hình toàn cầu IFS (**I**ntegrated **F**orecasting **S**ystem - mô hình toàn cầu của ECMWF với sơ đồ đồng hóa hiện đại 4D-Var cho các trường phân tích): tuy trường hợp thử nghiệm cho cơn bão Chanchu (0405) cho kết quả dự báo chưa tốt, song các nhà dự báo hy vọng rằng nếu thử nghiệm trên tập số liệu dài hơn khi đó sẽ cho những kết luận chính xác hơn.

Tất cả những kết quả trên nhận được qua việc thử nghiệm trên tập số liệu chưa dài (hai tháng). Trong thời gian tới Trung tâm Dự báo Khí tượng Thủy văn Trung ương sẽ trực tiếp thử

nghiệm trên tập số liệu dài hơn để có những kết luận chính xác hơn nữa. Việc cải tiến này mang ý nghĩa khoa học và thực tiễn vì các mô hình vùng giới hạn sẽ không phải phụ thuộc hoàn toàn vào trường phân tích và dự báo của một mô hình toàn cầu duy nhất. Ngoài ra, gần đây Trung tâm Dự báo Khí tượng Thủy văn Trung ương bước đầu đã có những nghiên cứu thử nghiệm xây dựng trường ẩm cho mô hình HRM từ số liệu vệ tinh địa tĩnh dựa trên phương pháp 3D-Var. Các kết quả thu được thông qua việc đánh giá chất lượng dự báo mưa của mô hình HRM trong vòng sáu tháng cho thấy rằng dự báo cho kết quả tốt ở Bắc Bộ và Trung Bộ đối với các ngưỡng mưa nhỏ. Khả năng nắm bắt kiểu dáng mưa khi sử dụng số liệu từ vệ tinh địa tĩnh lại tốt hơn so với GME trên toàn Việt Nam dù độ phân bố lượng mưa thấp hơn so với thực tế [2], [3].

5.KẾT LUẬN

Như đã phân tích ở trên, chúng ta thấy rằng đồng hóa số liệu là một thành phần rất quan trọng trong hệ thống dự báo thời tiết. Chúng tôi hy vọng trong tương lai gần ngành Khí tượng Việt Nam sẽ đưa hệ thống đồng hóa số liệu theo phương pháp biến phân ba chiều vào trong dự báo nghiệp vụ, tiến tới nghiên cứu đồng hóa biến phân bốn chiều để nâng cao mức độ chính xác trong dự báo thời tiết, khí hậu, phục vụ nhu cầu dân sinh, kinh tế, xã hội ngày càng tốt hơn.

Cảm ơn: Các tác giả của bài báo đặc biệt cảm ơn cố GS. TSKH. Nguyễn Đình Ngọc, người Thầy vô cùng kính mến, đã gợi mở ý tưởng để chúng tôi bước đầu nghiên cứu về lĩnh vực này.

FOUR-DIMENSIONAL VARIATIONAL DATA ASSIMILATION IN NUMERICAL WEATHER PREDICTION

Huynh Thi Hong Ngu, La Thi Cang
University of Natural Sciences, VNU-HCM

ABSTRACT: *Data assimilation is a very complicated cycle in numerical weather prediction and influences the quality of forecast. In this technique, the observed meteorological data are combined with the results of previous short-range forecast of the model to make an initial condition for a new prediction.*

This paper presents: four-dimensional variational assimilation (4D-Var), the most advanced technique for the data assimilation; the application of data assimilation in numerical weather prediction in Vietnam at present and the development of this technique in future.

Key words: *Four-dimensional variational assimilation, numerical weather prediction.*

TÀI LIỆU THAM KHẢO

- [1]. Fisher M., *Assimilation techniques (4): 4D-Var*. Meteorological Training Course Lecture Series. ECMWF, (2002).
(http://www.ecmwf.int/newsevents/training/lecture_notes/pdf_files/ASSIM/4dVar.pdf).

- [2]. Lê Đức, Đỗ Lệ Thủy, Lương Hồng Trung, *Xây dựng trường ảm cho mô hình HRM từ số liệu vệ tinh địa tĩnh dựa trên phương pháp biến phân ba chiều (3D-Var) (Phần I: cơ sở khoa học và phương pháp thực hiện)*. Tạp chí Khí tượng Thủy văn, 555, 22-32.(2007).
- [3]. Lê Đức, Đỗ Lệ Thủy, Lương Hồng Trung, *Xây dựng trường ảm cho mô hình HRM từ số liệu vệ tinh địa tĩnh dựa trên phương pháp biến phân ba chiều (3D-Var) (Phần II: phương pháp thực hiện và một số kết quả nghiên cứu)*. Tạp chí Khí tượng Thủy văn, 558, 43-49, (2007).
- [4]. Persson A., Grazzini F., *User Guide to ECMWF forecast products*. ECMWF. (www.ecmwf.int/products/forecasts/guide/user_guide.pdf) , (2005).
- [5]. Rabier F., Mahouf J.F., Klinker E., *Une nouvelle technique d'assimilation des données d'observation au CEPMMT: la'assimilation variationnelle quadridimensionnelle*. La Météorologie 8^e série, 30, 87-101, (2000).
- [6]. Tài liệu bồi dưỡng nghiệp vụ dự báo viên Khí tượng Thủy văn, Trung Tâm Dự Báo Khí Tượng Thủy Văn Trung Ương, (2006).
- [7]. Yessad K., *Basics about Arpege/IFS, Aladin and Arome in the cycle 30 of Arpege/IFS*. Météo-France/CNRM/CMAP/ALGO, (2005). (<http://www.cnrm.meteo.fr/gmapdoc/IMG/ps/ykarpbasics.ps>)