

ĐỒNG HÓA MƯA VỆ TINH BẰNG PHƯƠNG PHÁP 3DVAR, ÁP DỤNG THỬ NGHIỆM CHO LƯU VỰC SÔNG HỒNG

Trần Thanh Huyền⁽¹⁾, Kiều Quốc Chánh⁽²⁾, Trần Quang Đức⁽¹⁾

⁽¹⁾Đại học Khoa học Tự nhiên, Đại học Quốc gia Hà Nội

⁽²⁾Đại học Bloomington, Indiana, USA

Ngày nhận bài: 2/4/2019; ngày chuyển phản biện: 4/4/2019; ngày chấp nhận đăng: 15/5/2019

Tóm tắt: Đồng hóa số liệu mưa đầu vào phục vụ mục đích dự báo thủy văn đã và đang được các nhà nghiên cứu trong lĩnh vực khí tượng thủy văn quan tâm. Đặc biệt đối với những khu vực có địa hình phức tạp và số liệu phụ thuộc do tính chất liên quốc gia như khu vực thượng nguồn sông Hồng và địa giới Trung Quốc, số liệu đo đạc thường không liên tục và hạn chế ở nhiều khu vực, trong khi đó các sản phẩm mưa từ các vệ tinh quan sát trái đất dù có độ phủ không gian lớn nhưng khó có thể xác định được mưa đối lưu, mưa cục bộ với quy mô nhỏ. Nghiên cứu ứng dụng thử nghiệm phương pháp đồng hóa 3DVar cho mưa vệ tinh trên lưu vực sông Hồng với việc sử dụng 2 nguồn số liệu mưa vệ tinh GSMaP (Global Satellite Mapping) và GPM (Global Precipitation Measurement) được thu thập và xử lý bằng công cụ hỗ trợ dự báo Delft-FEWS. Số liệu mưa đồng hóa sau đó được kiểm định bằng số liệu thực đo cho chuỗi thời gian trong tháng 8/2018. Kết quả sau khi đồng hóa đã trình diễn một số điểm mưa cục bộ không phát hiện được bởi 2 nguồn mưa vệ tinh, đem lại bức tranh hoàn thiện hơn về phân bố mưa theo không gian cho lưu vực sông Hồng.

Từ khóa: 3DVar, mưa vệ tinh, GPM, GSMaP, lưu vực sông Hồng.

1. Đặt vấn đề

Trên quy mô lưu vực, mưa là một yếu tố khí tượng quan trọng ảnh hưởng trực tiếp tới các quá trình thủy văn như phát sinh dòng chảy, mưa lớn dẫn tới lũ quét phía thượng nguồn và ngập lụt phía hạ du. Hiện nay, có nhiều nguồn số liệu mưa đã được sử dụng và tham khảo trong dự báo thủy văn bao gồm số liệu thực đo tại trạm, số liệu mưa vệ tinh và số liệu radar. Do tính bất định cao của yếu tố mưa, các nguồn mưa vệ tinh còn bị hạn chế và không quan trắc được các hình thái mưa tầm thấp như mưa đối lưu hoặc mưa mang tính cục bộ do ảnh hưởng địa hình. Do đó, hiện tại các dự báo viên chỉ sử dụng số liệu mưa tại trạm để dự báo thủy văn và các nguồn số liệu khác chỉ mang tính tham khảo. Tuy nhiên, số liệu mưa trạm hạn chế về mặt không gian, đặc biệt những khu vực không có số liệu hoặc số liệu không liên tục do yếu tố liên quốc gia như

trường hợp của vùng nghiên cứu - lưu vực sông Hồng một phần lưu vực thuộc lãnh thổ Trung Quốc. Hạn chế này cần được khắc phục, đặc biệt trong những điều kiện thời tiết cực đoan như mưa lớn hay khi có bão đổ bộ.

Một số nghiên cứu đã được thực hiện để xây dựng và đánh giá trường mưa không gian từ nội suy mưa trạm như nghiên cứu của Trần Anh Đức, Ngô Đức Thành và cộng sự (2007) đã xây dựng bộ số liệu lượng mưa ngày trên lưới 1×1 độ cho toàn Việt Nam [2] hay nghiên cứu so sánh các phương pháp nội suy mưa không gian cho số liệu mưa vệ tinh TRMM, MOD07, ASTER DEM và mưa trạm trong nghiên cứu của Đỗ Khắc Phong và cộng sự (2015). Tuy nhiên, các nghiên cứu này mới chỉ nghiên cứu trên phạm vi rộng với bước lưới khá thô (1×1 độ kinh vĩ) nên không có tính ứng dụng cao trong mô phỏng thủy văn. Nghiên cứu này sẽ mở ra hướng nghiên cứu mới về đồng hóa số liệu mưa vệ tinh với độ phân giải vừa ($0,1 \times 0,1$ độ kinh vĩ) cho lưu vực sông Hồng, để từ đó đánh giá khả năng

Liên hệ tác giả: Trần Thanh Huyền
Email: tranthanhhhuyen@hus.edu.vn

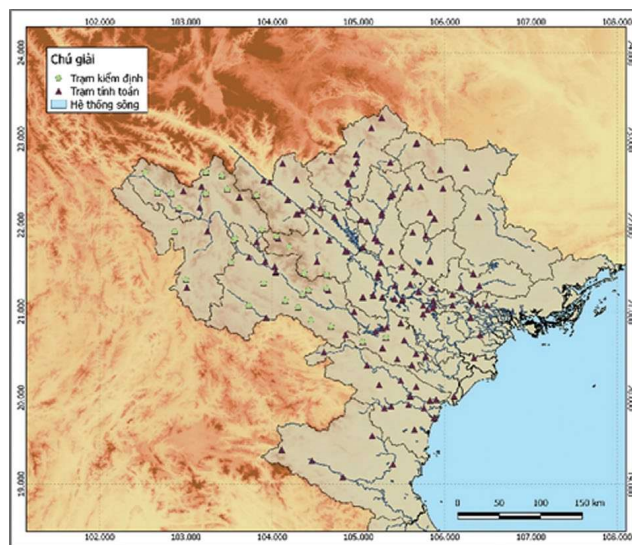
cải thiện độ chính xác của số liệu mưa vệ tinh để đảm bảo đầu vào cho mô hình thủy văn.

2. Đối tượng và phương pháp nghiên cứu

2.1. Đối tượng nghiên cứu

Khu vực nghiên cứu bao phủ toàn bộ lưu vực sông Hồng với miền tính nằm trong khoảng 101,45-108,0 độ kinh Đông và

19,05-25,5 độ vĩ Bắc, đặc trưng bởi hệ thống núi cao hướng Tây Bắc - Đông Nam ảnh hưởng tới hình thái mưa tại khu vực. Khu vực này có khoảng hơn 150 trạm mưa phân bố trên toàn miền nghiên cứu và chỉ có 30 trạm thuộc lưu vực sông Hồng - sông Đà có đầy đủ số liệu cho tháng 8/2018 sử dụng để kiểm định (Hình 1).



Hình 1 Phân bố các trạm mưa trên lưu vực sông Hồng

2.2. Phương pháp tiếp cận

2.2.1. Phương pháp đồng hóa số liệu

Đồng hóa số liệu là thuật ngữ được sử dụng cho mục đích tổng hợp các nguồn thông tin/dữ liệu khác nhau để đưa ra trạng thái tốt nhất của một hệ thống hoặc một trường giá trị nào đó [3]. Cho tới nay, nhiều phương pháp đồng hóa số liệu đã được giới thiệu và ứng dụng rộng rãi trong lĩnh vực khí tượng thủy văn. Một số phương pháp theo hướng phân tích chuẩn đoán như nội suy tối ưu (Optimal Interpolation - OI), lọc Kalman, lọc Kalman tổ hợp; trong khi đó một số phương pháp theo hướng phân tích biến phân như 3DVar và 4DVar [3]. Mục đích của các phương pháp này là tạo ra một trường ban đầu tốt cho mô hình tính toán.

Nghiên cứu quan tâm đến việc chuẩn bị đầu vào tốt nhất cho mô hình thủy văn cho các nghiên cứu tiếp theo, do đó không xét đến sai số mô hình hoặc các bước thời gian trong mô hình. Với mục đích đó, nghiên cứu chọn lựa phương pháp đồng hóa biến phân 3DVar để tính

toán trường mưa phân tích. Cơ sở lý thuyết của phương pháp 3DVar là tính toán cực tiểu của hàm giá (J) bằng cách sử dụng vòng lặp sao cho vi phân $\nabla J(x)=0$ [8]. Cụ thể:

Hàm giá

$$J(x) = (x - x_b)^T B^{-1} (x - x_b) + (y - H[x])^T R^{-1} (y - H[x]) \quad (1)$$

và vi phân hàm giá

$$\nabla J(x) = 2B^{-1} (x - x_b) - 2H^T R^{-1} (y - H[x]) \quad (2)$$

Trong đó, x là giá trị của trường phân tích; x_b là giá trị trường nền; B là ma trận tương quan sai số trường nền; H là ma trận tương quan giữa trường quan trắc và trường nền; R là ma trận tương quan sai số trường quan trắc; y là giá trị của trường quan trắc. Đồng hóa 3DVar, so với các phương pháp đồng hóa biến phân khác như 4DVar, đòi hỏi thời gian và năng lực tính toán ít hơn và không xét đến sai số mô hình và bước thời gian. Một nhược điểm của phương pháp 3DVar đó là việc tính toán ma trận B đòi hỏi năng lực và thời gian tính toán cao [5]. Để khắc phục nhược điểm này, nghiên cứu thiết lập một số giả thiết và giới hạn miền tính toán để đảm bảo khả năng tính toán của 3DVar.

2.2.2. Thiết lập giả thiết cho bài toán

Trong đồng hóa biến phân, tương quan sai số trường nền là tham số quan trọng nhất quyết định chất lượng của kết quả đồng hóa [9]. Giả thiết tương quan về không gian giữa các điểm giá trị mưa trên khu vực nghiên cứu được tính toán dựa trên phương pháp của Cressman về nội suy không gian. Phương pháp này đặt vấn đề rằng các điểm trên ô lưới sẽ có tương quan lẫn nhau với các trọng số khác nhau và phụ thuộc vào khoảng cách của điểm đó tới các điểm còn lại trên lưới [2]. Ma trận B được tính toán dựa trên phương pháp này. Cụ thể:

$$W_{ik} = \begin{cases} \frac{R^2 - r_{ik}^2}{R^2 + r_{ik}^2} & r_{ik}^2 < R \\ 0 & r_{ik}^2 \geq R \end{cases} \quad (3)$$

Trong đó, w_{ik} là hệ số tương quan giữa nút lưới thứ i và nút lưới thứ k ; R là bán kính ảnh hưởng tâm là điểm nút lưới; r_{ik} là khoảng cách tính từ nút lưới thứ i tới nút lưới thứ k . Sử dụng tính toán tương quan trong bài toán đồng hóa bằng phương pháp Cressman sẽ đảm bảo được tính phi tuyến và phân bố bất đồng đều theo phương không gian của bộ số liệu so với phương pháp nội suy tuyến tính do phương pháp này xác định bán kính ảnh hưởng của các điểm trên lưới đối với các khu vực lân cận. Tuy nhiên, phương pháp này cũng có những hạn chế nhất định vì cần đòi hỏi một số lượng trạm quan trắc nhất định nằm trong các ô lưới để đảm bảo tính tương quan của các giá trị [2]. Bên cạnh đó, việc xác định bán kính ảnh hưởng (R) cũng rất quan trọng phải được đưa vào bước hiệu chỉnh cho quá trình đồng hóa để chọn ra R phù hợp.

Nghiên cứu giả thiết sai số trung bình của trường quan trắc (σ_o) có giá trị là 0,2mm theo quy định trong Thông tư số 30/2018/TT-BTNMT về Quy định kỹ thuật về quan trắc và cung cấp thông tin, dữ liệu khí tượng thủy văn đối với trạm khí tượng thủy văn chuyên dùng [7]. Sai số trung bình của trường nền - trường mưa vệ tinh (σ_b) được ước lượng dựa trên phương pháp sai số bình phương (Root-mean-square error - RMSE). Để đơn giản hóa bài toán, σ_b được giả thiết không đổi trên toàn miền nghiên cứu và không phụ thuộc vào địa hình, khu vực địa lý và hình thái thời tiết. σ_o và σ_b là hai tham số đầu

vào quan trọng để tính toán ma trận B và R.

2.2.3. Nguồn số liệu

Số liệu mưa tại trạm được thu thập cho tháng 8/2018 được sử dụng để tính toán đồng hóa và kiểm định. Số liệu mưa vệ tinh GPM được trích xuất từ bộ số liệu tổ hợp sản phẩm thu hồi từ đa vệ tinh GPM IMERG/Early V05B, là số liệu mưa gần với thời gian thực nhất so với các bộ số liệu khác của IMERG (Final, Late, Early) [4]. Ưu điểm của nguồn số liệu GPM IMERG/Early là độ trễ tương đối nhỏ, có thể áp dụng cho khu vực thượng lưu để ứng dụng cho thủy văn, khả năng cung cấp số liệu độ phân giải cao cả không gian và thời gian (0,5 giờ - 11km) so với bộ sản phẩm cũ TRMM (3 giờ - 25km) [5] và được hiệu chỉnh bởi các đo đạc từ hệ thống quan sát cốt lõi GPM (GPM Core Observatory) [10]. Tuy nhiên, bộ số liệu GPM này vẫn chưa đáp ứng được nhu cầu dự báo đối với các khu vực cần tính tức thời. Bên cạnh GPM, nghiên cứu còn sử dụng bộ số liệu GSMaP NOW của Cơ quan Hàng không Vũ trụ Nhật Bản (JAXA). Bộ số liệu này cũng có ưu điểm về độ phân giải với bước thời gian 0,5 giờ và bước lưới 0,1 độ (11km) nhưng vượt trội hơn các sản phẩm mưa vệ tinh khác ở khả năng cung cấp số liệu mưa thời gian thực (độ trễ = 0) [6]. Hai nguồn mưa vệ tinh này được thu thập tương ứng với chuỗi thời gian của số liệu tại trạm quan trắc để tính toán đồng hóa và kiểm định.

3. Kết quả và bàn luận

3.1. Tính toán tương quan và sai số trung bình trường nền

Sai số trung bình của trường mưa vệ tinh GPM và GSMaP được tính qua mỗi bước thời gian đồng hóa. Một số phân tích thống kê của các sai số tính toán được của trường mưa GPM và GSMaP được thể hiện trong Bảng 1. Giá trị trung bình sai số của trường mưa GPM là 7,10mm, trong khi đó của trường GSMaP là 6,22mm. Có thể thấy giá trị sai số cực đại của cả hai nguồn mưa này khá lớn (37,86mm). Các giá trị sai số lớn này thường xuất hiện ở những khu vực có mưa lớn do tính bất định của số liệu mưa cũng tăng lên khi lượng mưa tăng. Mayor và cộng sự (2017) đã chỉ ra rằng, số liệu mưa GPM thường thiên thấp khi có các trận mưa lớn xuất hiện [10].

Bảng 1. Các đặc trưng thống kê sai số bình phương của 2 tập giá trị mưa GPM và GSMaP

RMSE	Cực tiểu	Phân vị thứ nhất	Trung vị	Trung bình	Phân vị thứ ba	Cực đại
GPM	0,00	1,09	6,50	7,10	10,27	37,86
GSMaP	0,00	0,28	4,93	6,22	9,29	37,86

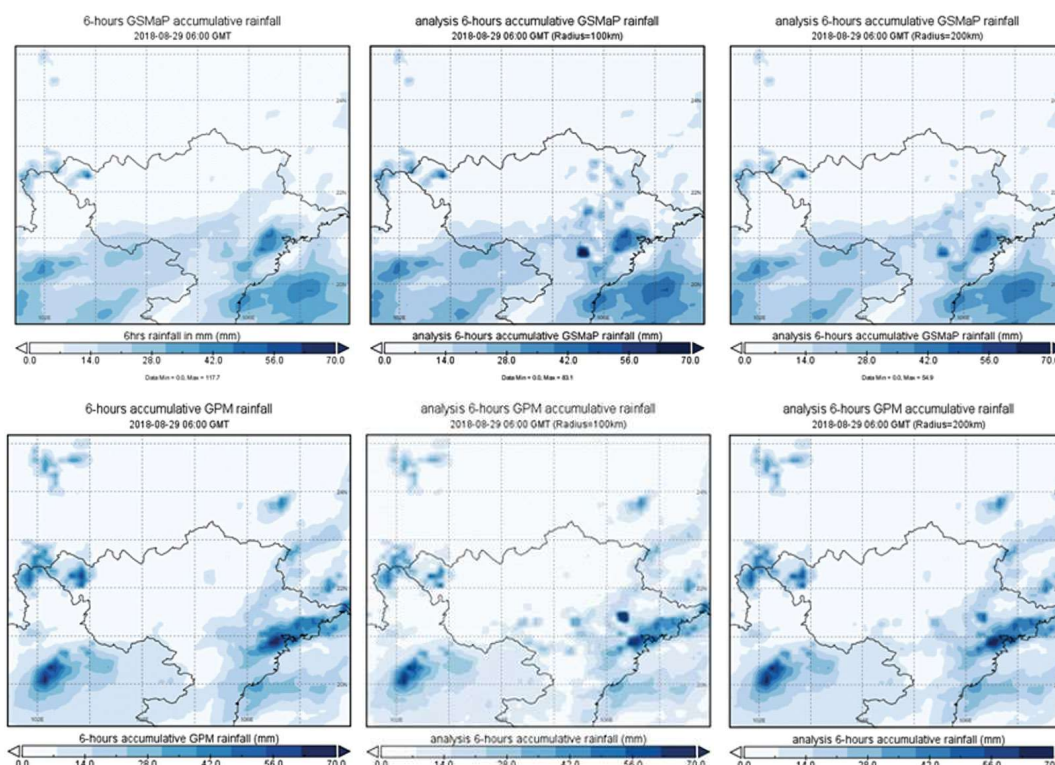
3.2. Đồng hóa 3DVar cho GSMaP và GPM

3.2.1. Chọn bán kính ảnh hưởng R

Các biểu đồ trong Hình 2 cho thấy phân bố mưa thay đổi đáng kể khi thay đổi tham số R.

Ví dụ cho thấy, với bán kính R=100km, trường phân tích đối với cả GPM và GSMaP đều

thu được thể hiện tốt hơn sự ảnh hưởng của các trạm mưa trong các ô lưới so với sử dụng bán kính 200km, thể hiện được những vùng có mưa (hoặc không có mưa) theo quan trắc từ thực đo. Trường phân tích với bán kính R=100km thể hiện rõ rệt khu vực chịu ảnh hưởng của số liệu thực đo (gây ra hiệu ứng “mắt trâu”) (Hình 2).



Hình 2. Trái qua phải: Hình trên: Trường nền GSMaP; Trường phân tích GSMaP (R=100km); Trường phân tích GSMaP (R=200km); Hình dưới: Trường nền GPM; Trường phân tích GPM (R=100km); Trường phân tích GPM (R=200km)

3.2.2. Đồng hóa 3DVar

Các biểu đồ phân bố mưa không gian trong Hình 3 đã cho thấy, trường phân tích của cả hai bộ số liệu đều thể hiện được sự cải thiện rõ rệt so với trường ban đầu nhờ sử dụng mạng lưới trạm thực đo vào đồng hóa. Một số khu vực có xuất hiện mưa cục bộ mà vệ tinh không phát hiện được cũng đã được biểu diễn trên trường phân tích sau đồng hóa.

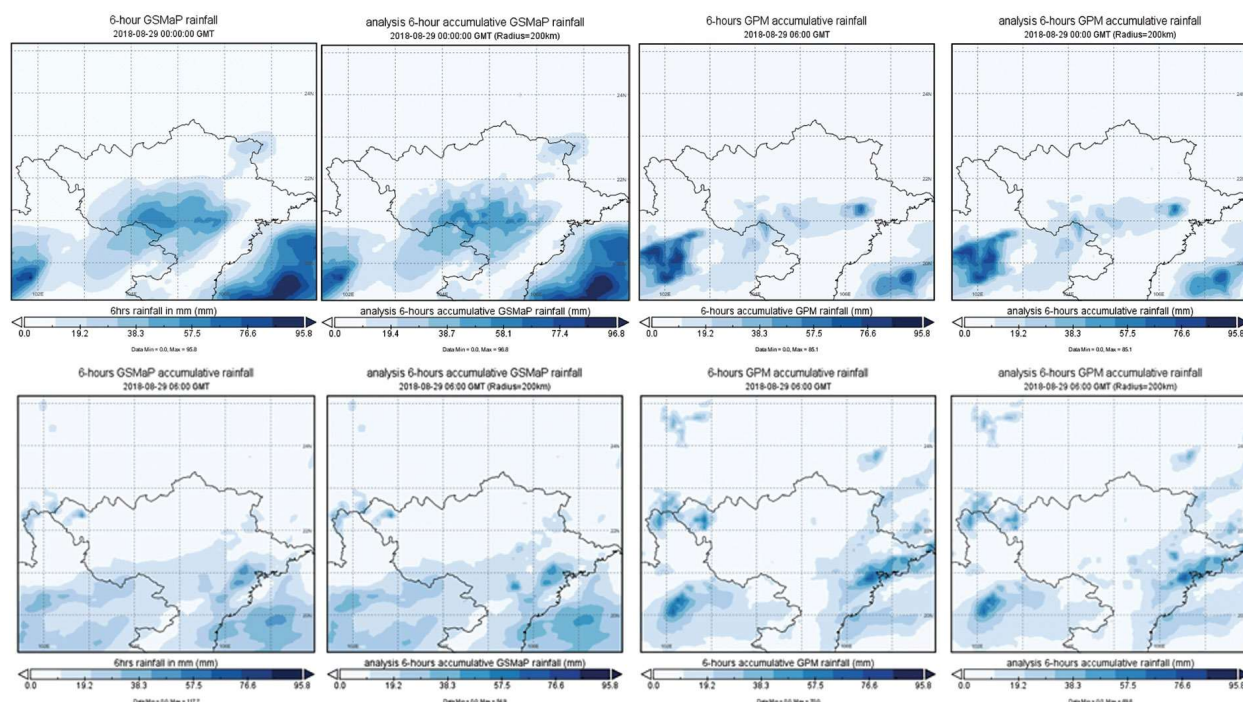
Chuỗi số liệu 20 ngày tháng 8 (từ ngày

1-20/8/2018) được sử dụng để kiểm định chương trình đồng hóa 3DVar. Cụ thể, số liệu trường phân tích sau tính toán được đưa về tọa độ trạm để so sánh với số liệu thực đo tại các trạm độc lập.

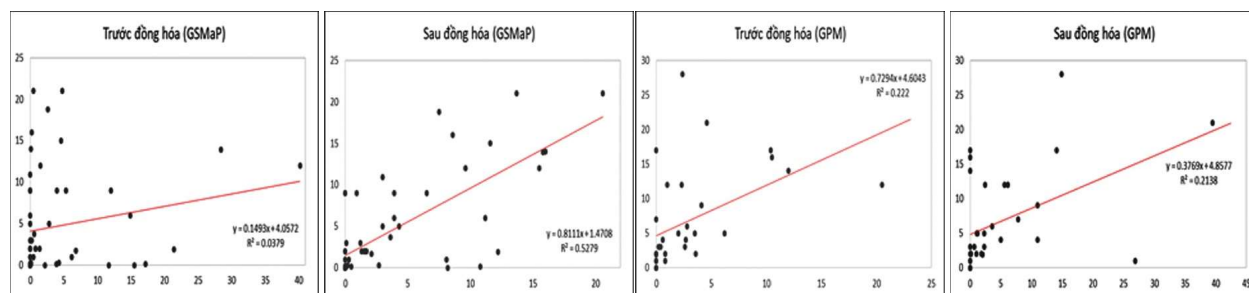
Phân tích tương quan cho thấy, mưa phân tích của GSMaP cho kết quả khá tốt với hệ số tương quan $R^2=0,53$, trong khi đó GPM chỉ là 0,21 (Hình 4).

4. Kết luận

Nghiên cứu sử dụng phương pháp đồng hóa



Hình 3. Trái qua phải: Trường ban đầu GSMaP, Trường phân tích GSMaP, Trường ban đầu GPM, Trường phân tích GPM vào lúc 00:00 GMT ngày 29/8/2018 (trên); và lúc 06:00 GMT ngày 29/8/2018 (dưới)



Hình 4. Tương quan giữa số liệu mưa GSMaP trước và sau đồng hóa so với thực đo

3DVar tạo ra trường mưa phân tích cho GSMaP và GPM. Kết quả cho thấy trường mưa ban đầu được cải thiện đáng kể, biểu diễn được một số điểm gây mưa cục bộ không quan sát được bởi vệ tinh. Nghiên cứu cũng chỉ ra bán kính ảnh hưởng R là một tham số rất quan trọng, có thể quyết định chất lượng của sản phẩm phân tích. Độ tin

cậy và mật độ trạm quan trắc là yếu tố quan trọng để tạo ra một sản phẩm đồng hóa tốt. Đây là một hướng tiếp cận mới sử dụng mưa vệ tinh độ phân giải không gian và thời gian cao kết hợp đồng hóa số liệu giải quyết bài toán mô phỏng dòng chảy ở quy mô lưu vực. Cụ thể trong nghiên cứu, có thể áp dụng tốt cho lưu vực sông Hồng.

Lời cảm ơn: Nghiên cứu là sản phẩm thuộc Đề tài TN.18.24 “Phát triển hệ thống dự báo lũ tích hợp trên nền tảng Delft-FEWS để nâng cao hiệu quả dự báo lũ cho lưu vực sông Hồng”, Trường Đại học Khoa học Tự nhiên, Đại học Quốc gia Hà Nội, do Th.S Trần Thanh Huyền là chủ nhiệm đề tài và đồng thời là tác giả chính. Tập thể tác giả xin cảm ơn sự hỗ trợ của Trung tâm Động lực học Thủy khí Môi trường và Tiểu dự án: “Hiện đại hóa hệ thống quan trắc và mô phỏng/dự báo các điều kiện khí tượng hải văn - môi trường biển và đới ven bờ độ phân giải cao phục vụ khai thác bền vững tài nguyên biển và giảm thiểu rủi ro thiên tai”, trong khuôn khổ Dự án FIRST góp phần xây dựng nghiên cứu được hoàn chỉnh hơn.

Tài liệu tham khảo

Tài liệu tiếng Việt

1. Bộ Tài nguyên và Môi trường (2018), “Thông Tư 30/2018/TT-BTNMT Quy định kỹ thuật về quan trắc và cung cấp thông tin, Dữ liệu Khí tượng thủy văn đối với trạm Khí tượng thủy văn chuyên dùng”.
2. Trần Anh Đức, Phan Văn Tân, Ngô Đức Thành (2010), Xây dựng bộ số liệu mưa ngày VNGP_1DEG trên lưới 1°x1° kinh vĩ cho Việt Nam, Tạp chí Khí tượng Thủy văn, Hà Nội, 2 (590), tr. 42-48.

Tài liệu tiếng Anh

3. E. Blayo, E. Cosme, and A. Vidard (2011), “Introduction to Data Assimilation,” vol. 2, pp. 1–35.
4. G. J. Huffman, D. T. Bolvin, E. J. Nelkin, and J. Tan (2019), “Integrated Multi-satellitE Retrievals for GPM (IMERG) Technical Documentation,” no. March, pp. 1–64.
5. H. Beria, T. Nanda, D. S. Bisht, and C. Chatterjee (2017), “Does the GPM mission improve the systematic error component in satellite rainfall estimates over TRMM? An evaluation at a pan-India scale,” Hydrol. Earth Syst. Sci., vol. 21, no. 12, pp. 6117–6134.
6. Kim Hyungjun; T. Kubota; N. Utsumi (2017), “Development and Applications of the GSMaP: Overview & Lessons learned in a real-world case for Hydrological Status and Outlook System,” WMO Glob. Hydrol. Status Outlook Syst.
7. K. P. Do et al., (2015), “Spatial interpolation and assimilation methods for satellite and ground meteorological data in Viet Nam” J. Inf. Process. Syst., vol. 11, no. 4, pp. 556–572.
8. P. Courtier and F. Bouttier (1999), “Data assimilation concepts and methods”, in Data assimilation concepts and methods.
9. V. Rakesh; P. Goswami (2011), “Impact of background error statistics on 3D-Var assimilation: Case study over the Indian region Impact of background error statistics on 3D-Var assimilation: Case study over the Indian region,” no. April.
10. Y. G. Mayor, I. Tereshchenko, M. Fonseca-Hernández, D. A. Pantoja, and J. M. Montes (2017), “Evaluation of error in IMERG precipitation estimates under different topographic conditions and temporal scales over Mexico,” Remote Sens., vol. 9, no. 5, pp. 1–18.

SATELLITE-BASED RAINFALL DATA ASSIMILATION USING 3DVAR TECHNIQUE - CASE STUDY IN RED RIVER BASIN

Tran Thanh Huyen⁽¹⁾, Kieu Quoc Chanh⁽²⁾, Tran Quang Duc⁽¹⁾

⁽¹⁾University of Science, Viet Nam National University

⁽²⁾Indiana University Bloomington, Indiana, USA

Received: 2/4/2019; Accepted: 15/5/2019

Abstract: Data assimilation of rainfall as input for hydrological forecasting has been intensively researched by many scientists in hydro-meteorological fields. Especially, in areas that characterized by complex terrains and rainfall that are highly data dependent due to transboundary issues. This is true for the upstream area of Red River and Chinese territories. Of which rainfall data is not continuously measured and is not highly available in some remote areas. Meanwhile, in spite of the large domain cover rainfall products from the Earth Observation Satellites (EOS), there still exist a challenge in detecting small scale convective rainfall and local rainfall. This study focuses on applying a 3DVar technique for satellite-based rainfall data assimilation using two satellite rainfall sources: GSMaP (Global Satellite Mapping) and GPM (Global Precipitation Measurement) that are collected and preprocessed by Delft-FEWS. Analyzed rainfall data was then validated using observed rainfall for August, 2018. The results showed that the analysis data could detect some local rainfall points which were not detected by the two satellite rainfall sources thus provides a more complete image of spatial rainfall distribution for the Red River Basin.

Keywords: 3DVar, Satellite rainfall, GPM, GSMaP, Red River Basin.