

NGHIÊN CỨU ĐÁNH GIÁ CHẤT LƯỢNG DỮ LIỆU MƯA LƯỚI MỘT SỐ TRẬN MƯA LỚN TRÊN KHU VỰC THÀNH PHỐ HÀ TĨNH

Nguyễn Thị Liên, Nguyễn Quang Hưng, Trịnh Minh Ngọc
Trường Đại học Khoa học Tự nhiên, Đại học Quốc gia Hà Nội

Ngày nhận bài: 3/11/2022; ngày chuyển phản biện: 4/11/2022; ngày chấp nhận đăng: 28/11/2022

Tóm tắt: Số liệu mưa lưới, sản phẩm ước tính mưa từ vệ tinh kết hợp với giá trị mưa từ mô hình số trị và số liệu mưa thực đo mặt đất đang ngày càng phát triển do khắc phục được nhiều yếu điểm của mưa mặt đất. Mưa lưới có phân giải không gian, thời gian cao, bao phủ nhiều vùng rộng lớn, tần suất lên đến 30 phút một số liệu và phần nhiều cung cấp miễn phí, do đó có tính ứng dụng rất cao. Các sản phẩm mưa lưới dù sử dụng cùng một loại cảm biến, nhưng áp dụng thuật toán khác nhau để xử lý số liệu và hiệu chỉnh với các số liệu đo mưa khác nên có sự khác biệt khá lớn. Nghiên cứu này xem xét dữ liệu mưa giờ từ 04 nguồn gồm mưa GPM, GSMaP, PERSIANN-CCS và ERA5-Land cung cấp cho khu vực thành phố Hà Tĩnh để đánh giá chất lượng và khả năng ứng dụng thực tế. Bốn trận mưa được sử dụng để đánh giá với mưa giờ thực đo tại trạm khí tượng Hà Tĩnh. Các chỉ số thống kê RMSE, R^2 , PBIAs, POD, FAR và CSI được sử dụng để đánh giá so sánh. Kết quả cho thấy nhìn chung cả 4 dữ liệu mưa đều có chất lượng tốt, PERSIANN-CCS có tương quan cao nhất với mưa quan trắc mặt đất, chỉ số sai số quân phương trung bình tốt nhất, trong khi đó các chỉ số khác đều chỉ ra dữ liệu GSMaP kém chất lượng nhất trong 4 nguồn dữ liệu.

Từ khóa: Mưa lưới, mưa vệ tinh, thành phố Hà Tĩnh.

1. Giới thiệu chung

Mưa là quá trình quan trọng trong chu trình thủy văn, đóng vai trò cầu nối trong tương tác khí quyển và mặt đất [20]. Mặc dù các số liệu mưa thu thập được bằng các trạm quan trắc mặt đất vẫn được công nhận là có tính chính xác cao, tuy nhiên vẫn còn nhiều vấn đề xảy ra như trang thiết bị hỏng hóc dẫn đến gián đoạn thu thập số liệu, sai số do chính trang thiết bị, sai số do gió, vị trí đặt trạm, và nhất là đang bị giới hạn về mặt không gian [15]. Có thể nói, các vấn đề về số liệu mưa đo đạc trên mặt đất là một trong những rào cản lớn trong các ứng dụng thực tế của Thủy văn và Tài nguyên nước [7]. Thời gian gần đây, các sản phẩm quan trắc của vệ tinh cũng như mưa tái phân tích (gọi chung là sản phẩm mưa lưới - GPP_s - Gridded Precipitation Products) đang được quan tâm và triển khai trong các ứng dụng mô hình thủy văn. Các đặc tính của mưa vệ tinh và mưa tái phân

tích về không gian, độ bao phủ, tính liên tục, tính kịp thời, độ chính xác đã chứng tỏ hướng tiếp cận sử dụng mưa vệ tinh là hướng đi đúng đắn, hỗ trợ mạnh mẽ cho các ứng dụng tính toán và mô hình hóa [6].

Chính vì vậy, nghiên cứu này hướng tới việc đánh giá các sản phẩm mưa vệ tinh và mưa tái phân tích (mưa lưới) với mưa thực đo của trạm đo mưa mặt đất. Kết quả đánh giá sẽ là cơ sở khoa học để tiến hành tiếp các nghiên cứu về tính ứng dụng của sản phẩm mưa lưới khi liên kết trực tiếp làm đầu vào cho các mô hình toán thủy văn. Nghiên cứu này sẽ tập trung vào việc phân tích số liệu mưa lưới ở các đặc điểm: Sự xuất hiện mưa, và độ chính xác so với mưa thực đo trên mặt đất.

Trong nghiên cứu đánh giá lựa chọn mưa vệ tinh cho nghiên cứu dòng chảy trên lưu vực sông Cả, tác giả Bùi Tuấn Hải và Nguyễn Văn Tuấn [2] đã phân tích và so sánh giữa các dữ liệu GPM, TRMM, CHIRPS, CMORPH với dữ liệu mưa tại 12 trạm đo mưa trên lưu vực sông Cả. Nhóm nghiên cứu đã sử dụng bộ số liệu trong ba năm từ 2015 đến 2017 để so sánh lượng mưa ngày

Liên hệ tác giả: Nguyễn Thị Liên
Email: liennt1@hus.edu.vn

(khả năng nhận diện mưa, tương quan lượng mưa), tổng lượng và phân bố lượng mưa năm, đánh giá lượng mưa tháng (tương quan tổng lượng mưa và tính toán các hệ số tương quan R , R^2 , sai số RMSE, MAE). Kết quả nghiên cứu đã chỉ ra khả năng nhận diện mưa ngày ở các dữ liệu mưa vệ tinh là khá tốt, tuy nhiên tương quan lượng mưa ngày lại không cao; tương quan lượng mưa tháng cho thấy dữ liệu CHIRPS cho kết quả ổn định nhất. Tuy nhiên nghiên cứu này mới dừng ở giới hạn mưa ngày, do đó cần có các nghiên cứu sâu hơn, phân tích mưa giờ để có thể lựa chọn nguồn dữ liệu cho các mục đích với yêu cầu mưa giờ.

Năm 2013, Ngô Đức Thành và các đồng nghiệp đã sử dụng trí tuệ nhân tạo để hiệu chỉnh mưa trung bình tháng cho các nguồn dữ liệu mưa lưới trên lưu vực sông Vu Gia Thu Bồn với số liệu đo mưa từ 8 trạm đo mưa mặt đất. mưa lưới GSMaP_NOW, APHRODITE được thu thập và đánh giá chất lượng trên lưu vực nghiên cứu. Kết quả cho thấy số liệu mưa lưới thu thập trong thời gian 2001 - 2007 cho kết quả thấp đặc biệt trong thời gian gió mùa Đông Bắc. Kết quả phân tích cho thấy cao độ cũng có ảnh hưởng đến độ chính xác của mưa lưới. Do đó nhóm tác giả đã ứng dụng mô hình trí tuệ nhân tạo để hiệu chỉnh mưa lưới và kết quả đã cải thiện chất lượng mưa lưới khá tốt [21].

Trước đó, nhóm nghiên cứu của tác giả Bùi Chí Nam đã triển khai đánh giá, dữ liệu mưa quan trắc vệ tinh từ GPM và PERSIANN phục vụ cảnh báo mưa thành phố Hồ Chí Minh thông qua các thông số thống kê [5]. Kết quả tính toán cho thấy độ chính xác của 3 loại dữ liệu này đều cao, phân bố trung bình của PC (độ chính xác - Percentage Correct) lần lượt của các loại dữ liệu TMPA, IMERG và PERSIANN-CCS là 76,44%; 71,51% và 73,70%. Sai số lượng mưa TMPA, PERSIANN-CCS có xu thế cao hơn lượng mưa mặt đất có sai số tuyệt đối lần lượt là trong khoảng 13,14 mm và 12,94 mm. Trong khi đó IMERG có xu thế thấp hơn so với trạm, sai số tuyệt đối là 10,86 mm. Về không gian vùng mưa từ dữ liệu PERSIANN-CCS là rộng nhất, kể đến là IMERG và cuối cùng là TMPA có phân bố phù hợp nhất so với các trạm. Sau khi phân tích đánh giá kết quả, nghiên cứu khuyến nghị sử

dụng PERSIANN-CCS, do có sự phân hạng theo các tiêu chí đánh giá ở mức cân bằng so với hai loại dữ liệu còn lại.

Trong nghiên cứu gần đây, năm 2021, Phạm Thành Hưng và cộng sự đã công bố kết quả đánh giá khả năng sử dụng các sản phẩm mưa vệ tinh MSWEP, TMPA 3B42, TMPA 3B42RT, PERSIANN CDR, PERSIANN RT thông qua việc so sánh với giá trị mưa đo tại 06 trạm mưa trên lưu vực Nông Sơn - Quảng Nam. Kết quả cho thấy sai số của các sản phẩm mưa vệ tinh so với mưa trạm đo là khá lớn, theo đó gây ra sai số trong kết quả mô phỏng dòng chảy. Các sản phẩm mưa vệ tinh dựa vào thông tin sóng siêu cao tần (MW) như TMPA cho kết quả mô phỏng dòng chảy và thời gian xuất hiện đỉnh lũ tốt hơn các sản phẩm mưa vệ tinh dựa vào thông tin sóng hồng ngoại (IR) như PERSIANN [3].

Một nghiên cứu mang tính cập nhật hơn đã được Mai Khánh Hưng và các đồng nghiệp công bố năm 2020 so sánh đánh giá các sản phẩm mưa vệ tinh thời đoạn 6 tiếng với 184 trạm đo mưa trên lãnh thổ Việt Nam, cụ thể là các sản phẩm mưa lưới Himawari-8, GSMaP_NOW and GSMaP_MVK. Dữ liệu mưa được thu thập từ tháng 10 năm 2019 đến tháng 1 năm 2020 và được đánh giá thông qua các chỉ số thống kê, chỉ số dự báo. Kết quả cho thấy GSMaP_MVK đạt các giá trị thống kê tốt nhất trong số các nguồn mưa lưới, nhưng lại thấp nhất với năng lực phát hiện mưa. Dữ liệu Himawari-8 thường cho kết quả thiên cao, đặc biệt trong trường hợp nhiệt độ đỉnh mây thấp. Nhìn chung dữ liệu mưa lưới Himawari và GSMaP_NOW có thể sử dụng kết hợp đồng thời với nhau để cung cấp thông tin thời gian thực cho dự báo cảnh báo mưa lớn, lũ quét và sạt lở đất [11].

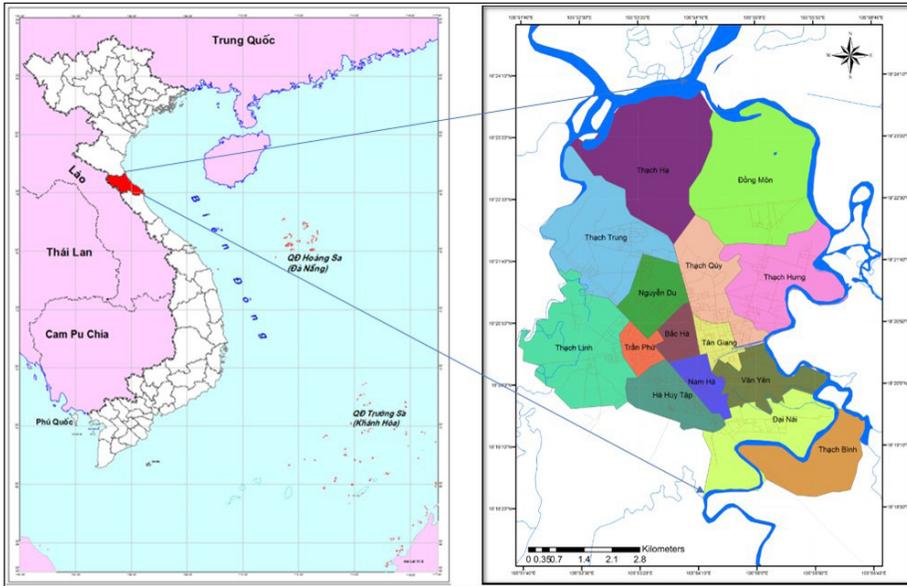
Mưa thay đổi theo không gian, phụ thuộc nhiều vào điều kiện địa hình và khí hậu của từng khu vực. Các vùng lãnh thổ khác nhau thì khả năng ứng dụng và chất lượng mưa lưới từ các nguồn quốc tế cũng có chất lượng dao động. Chất lượng sản phẩm mưa lưới còn phụ thuộc vào các dữ liệu phát báo tại địa phương tới mô hình số trị toàn cầu và hiệu chỉnh kết quả tính toán mưa lưới từ các dữ liệu ảnh vệ tinh... Các nghiên cứu trước đây nhìn chung đang giới hạn xem xét đánh giá mưa trung bình tháng, 7 ngày,

hoặc gần đây nhất là mưa thời đoạn 6 tiếng, và hoàn toàn chưa có nghiên cứu nào về mưa giờ. Vì thế, việc nghiên cứu đánh giá chất lượng sản phẩm mưa lưới và khả năng ứng dụng thực tế của nó trong bài toán thủy văn và tài nguyên nước là hoàn toàn cần thiết và cần triển khai độc lập trên từng khu vực, nhất là với mưa thời đoạn 1 giờ.

2. Dữ liệu và phương pháp nghiên cứu

2.1. Khu vực nghiên cứu

Thành phố Hà Tĩnh trải dài từ 18°18' đến 18°24' vĩ Bắc và từ 105°53' đến 105°56' kinh Đông, nằm trên trục đường Quốc lộ 1A, cách thủ đô Hà Nội 340 km, cách thành phố Vinh 50 km về phía Bắc; cách thành phố Huế 314 km về phía Nam và cách biển Đông 12,5 km (Hình 1).



Hình 1. Khu vực nghiên cứu - thành phố Hà Tĩnh, tỉnh Hà Tĩnh

Thành phố Hà Tĩnh nằm trong vùng đồng bằng ven biển miền Trung, địa hình dốc nghiêng từ Tây sang Đông (độ dốc trung bình 1,2% có nơi 1,8%) và bị chia cắt mạnh bởi hệ thống sông suối của dãy Trường Sơn. Tỉnh Hà Tĩnh có 2 mùa rõ rệt: Mùa nắng: Kéo dài từ tháng 4 đến tháng 10, khí hậu khô nóng, nhiệt độ trung bình tháng từ 24,7°C (tháng 4) đến 32,9° (tháng 6). Nhiệt độ cao nhất có thể lên đến 38,5 - 40°C. Mùa mưa kéo dài từ tháng 10 đến tháng 4 năm sau, nhiệt độ trung bình tháng từ 18,3°C (tháng 1) đến 21,8°C (tháng 11) với nhiều ngày ở một số khu vực có nhiệt độ 7°C (tháng 11, 12). Hà Tĩnh có lượng mưa khá lớn, trung bình trên 2.000 mm/năm, cá biệt có vùng lên đến 3.500 mm/năm. Lượng mưa phân bố không đồng đều trong năm, mùa Đông - Xuân lượng mưa thấp, chiếm 25% lượng mưa hàng năm, chủ yếu là mưa phùn kết hợp gió mùa Đông Bắc. Mưa lớn tập trung vào mùa Hạ và mùa Thu, chiếm 75%

lượng mưa cả năm, đặc biệt cuối thu thường mưa rất to.

2.2. Nguồn dữ liệu mưa được sử dụng trong nghiên cứu

Sự ra đời và phát triển của sản phẩm mưa lưới (GPP_s) đã mở ra một kỷ nguyên mới trong giám sát, tính toán và dự báo tài nguyên nước, đặc biệt là những vùng không có hoặc thiếu số liệu. Tuy nhiên, có khá nhiều sản phẩm mưa lưới khác nhau, giữa các sản phẩm có độ chính xác và độ phân giải khác nhau cả về không gian và thời gian trên quy mô toàn cầu do thiết bị sử dụng, thuật toán xử lý khác nhau hoặc do thuộc tính vật lý của mây ở vùng nghiên cứu [2, 3]. Các sản phẩm mưa lưới có thể được chia ra theo dữ liệu đầu vào gồm: (1) Dữ liệu quan trắc tại trạm mặt đất; (2) Dữ liệu mưa tái phân tích; (3) Dữ liệu mưa từ tín hiệu vệ tinh/radar.

(1) GPP_s dựa trên dữ liệu quan trắc tại trạm mặt đất được xây dựng qua các thuật toán nội

suy (Bảng 1).

(2) Dữ liệu mưa tái phân tích là sản phẩm từ mô hình số trị dự báo thời tiết (NWP), mô hình khí quyển. Những sản phẩm này được đồng bộ với dữ liệu vệ tinh và trạm quan trắc mặt đất nhằm tăng độ chính xác [9] (Bảng 2).

(3) Dữ liệu mưa vệ tinh/radar bắt đầu xuất hiện từ những năm 1960, khi vệ tinh TIROS (Television and IR Observation Satellite) chụp ảnh đám mây từ trên cao. Sản phẩm mưa vệ tinh thường từ 2 loại cảm biến: cảm biến hồng ngoại (IR) từ vệ tinh địa tĩnh GEO và cảm biến vi sóng (PMW) từ vệ tinh quỹ đạo Trái đất tầng thấp LEO. Dữ liệu IR có tính ưu việt về độ phủ không gian và thời gian (1 - 4 km, 10 - 30 phút). Các ảnh này sẽ cho bức tranh về trường mây và bề mặt mây được hiển thị bằng tông độ ánh sáng của ảnh, mây càng dày thì ảnh càng sáng, mây càng

mỏng thì ảnh càng tối. Vì thế sản phẩm IR không có khả năng phát hiện các tín hiệu từ cấu trúc đám mây, điều này ảnh hưởng đến khả năng xác định lượng mưa. Cảm biến PMW có thể xác định lượng mưa tương đối chính xác theo nguyên lý tán xạ (trên đất liền) và phát xạ (trên đại dương). Tuy nhiên, các vệ tinh LEO chỉ cung cấp 1 - 2 ảnh trong 1 ngày và có khoảng trống lớn về không gian. Do đó, để tận dụng ưu điểm của mỗi loại dữ liệu, nhiều kỹ thuật kết hợp IR và PMW hoặc mưa trạm hoặc mưa tái phân tích đã được phát triển. Kỹ thuật này dẫn đến sự ra đời của một loạt các sản phẩm mưa có độ phân giải cao và chính xác hơn. Bảng 3 thể hiện các sản phẩm mưa được phân tích từ vệ tinh, tuy nhiên còn nhiều sản phẩm lượng mưa khác ở các độ phân giải không gian khác nhau rất đa dạng và phong phú.

Bảng 1. Thông tin các sản phẩm mưa lưới nội suy từ dữ liệu quan trắc mặt đất

Dữ liệu mưa	Thời kỳ	Độ phân giải không gian	Độ phân giải thời gian	Vùng có số liệu	Nguồn
VnGP	1980 - 2010	0,1° và 0,25°	Ngày	Việt Nam	[17]
APHRODITE	1951 - 2015	0,25° và 0,5°	Ngày	50°N - 50°S	[23]
GPCC	1900 - 2019	0,5° và 1°	Tháng	50°N - 50°S	[19]
CRU	1901 - 2013	0,5°	Tháng	50°N - 50°S	[10]
CPC	1979 - 2019	0,5°	Ngày	60°N - 60°S	[8]

Bảng 2. Thông tin một số sản phẩm mưa tái phân tích

Dữ liệu mưa	Thời kỳ	Độ phân giải không gian	Độ phân giải thời gian	Vùng có số liệu	Nguồn
NCEP1	1948 - nay	2,5°	Tháng/ngày/6 giờ	Toàn cầu	NCEP/NCAR
NCEP2	1979 - nay	1,875°	Tháng/6 giờ	Toàn cầu	NCEP/NCAR
ERA 40	1957 - 2002	2,5° và 1,125°	Tháng/6 giờ	Toàn cầu	ECMWF
ERA Interim	1979 - 2019	1,5° và 0,75°	Tháng/6 giờ	Toàn cầu	ECMWF
ERA-20C	1900 - 2014	125 km	Tháng/3 giờ	Toàn cầu	ECMWF
ERA5 - Land	1950 - nay	0,1°	Tháng/ngày/1 giờ	Toàn cầu	ECMWF

Bảng 3. Thông tin một số sản phẩm mưa vệ tinh

Dữ liệu mưa	Vệ tinh liên quan	Thời kỳ	Độ phân giải không gian	Độ phân giải thời gian	Vùng có số liệu	Nguồn	Độ trễ
PGF	EUMETSAT GEOS, MTSAT NOAA-19	1948 - 2010	0,25°	3 giờ	90°N - 90°S	Princeton Univ	

Phương pháp nghiên cứu

Số liệu mưa quan trắc trực tiếp tại trạm là dữ liệu điểm, mưa lưới là dữ liệu lưới, do đó để so sánh 2 loại dữ liệu này thường có 2 phương pháp bao gồm: (1) Nội suy mưa điểm tại trạm về mưa lưới có cùng độ phân giải với sản phẩm mưa lưới; (2) Nội suy mưa lưới về điểm quan trắc. Tuy nhiên, với cách tiếp cận (1) chỉ nên sử dụng ở những khu vực có mạng lưới quan trắc dày và các sản phẩm nội suy cần được đánh giá kiểm định về phương pháp sử dụng và lượng mưa tại mỗi ô lưới để đảm bảo số liệu mưa tại mỗi ô lưới sẽ đại diện cho số liệu mưa tại đó.

Với cách tiếp cận (2) nội suy mưa lưới về điểm quan trắc thường được sử dụng trong nhiều nghiên cứu, cách tiếp cận này có ưu điểm có thể áp dụng được tại những nơi có mật độ trạm thưa và khi so sánh các sản phẩm mưa lưới trực tiếp với các giá trị mưa quan trắc tại trạm sẽ đảm bảo không mất đi đặc điểm mang tính cực trị của dữ liệu quan trắc. Cụ thể, xét xem trạm quan trắc nằm trong ô lưới nào, sau đó lấy giá trị mưa của ô lưới đó so sánh với giá trị quan trắc của trạm. Kết quả so sánh được đánh giá dựa trên các chỉ số thống kê. Ngoài so sánh trực tiếp hai giá trị tại một ô lưới, một số các tác giả khác còn so sánh giá trị mưa mặt đất với giá trị trung bình của mưa trên 9 ô lưới đối với mưa vệ tinh.

Các chỉ số thống kê đánh giá bao gồm: Độ lệch quân phương trung bình (RMSE), hệ số tương quan (R^2), chỉ số sai số độ lệch (PBIAS). Ngoài ra nghiên cứu còn sử dụng chỉ số thống kê định tính: Xác suất phát hiện đúng (POD), phát hiện

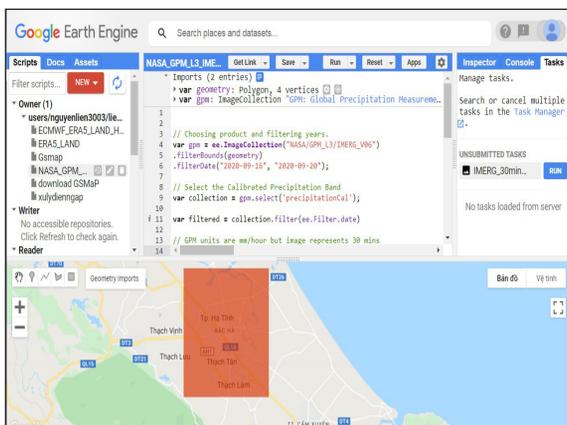
khống (FAR) và chỉ số thành công (CSI). Trong đó, CSI là tỷ số giữa số lần cảnh báo đúng có xảy ra hiện tượng và tổng số lần cảnh báo, FAR để đánh giá khả năng phát hiện khống (không mưa nhưng cảnh báo là mưa) các hiện tượng mưa. Chỉ số POD là tỷ số giữa số lần phát hiện đúng (hiện tượng có mưa và không có mưa) trên tổng số lần cảnh báo hiện tượng [1, 16].

3. Phân tích đánh giá dữ liệu mưa

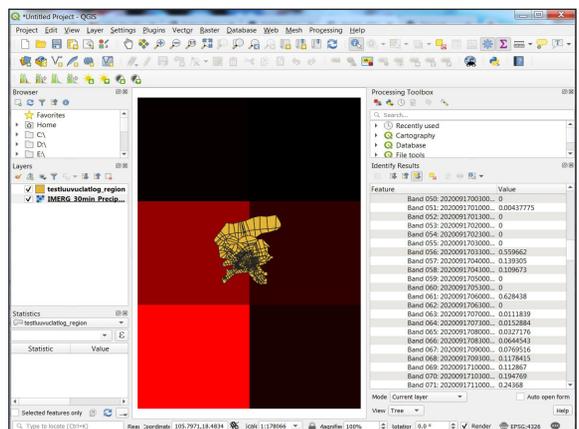
3.1. Thu thập và xử lý dữ liệu mưa lưới

Dữ liệu mưa GPM được tải trên trang chủ <https://gpm.nasa.gov/data/imerg>, với tần suất 30 phút một bộ số liệu. Nhóm nghiên cứu sử dụng code trong công cụ Google Earth Engine (GEE) để tải nhanh dữ liệu mưa vệ tinh với chuỗi thời gian dài (Hình 3). Dữ liệu mưa tải về chủ yếu ở định dạng tif, sau đó được tiến hành xử lý bằng các công cụ trong phần mềm QGIS (Hình 4). Nhìn chung công việc xử lý dữ liệu mưa phải làm thủ công và tốn nhiều thời gian để đảm bảo chất lượng của bộ dữ liệu.

Tương tự như thế, các dữ liệu mưa khác đều được tải về máy tự động bằng các đoạn mã nhóm nghiên cứu tự phát triển trong môi trường GEE và được xử lý thủ công trong phần mềm QGIS. Dữ liệu mưa GSMaP được tải trên trang <https://sharaku.eorc.jaxa.jp/GSMaP/>. Dữ liệu mưa PERSIANN-CCS có thể truy cập tại địa chỉ <https://chrsdata.eng.uci.edu/>. Dữ liệu mưa ERA5-Land có tại trang web: <https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-land?tab=form>.



Hình 3. Tải dữ liệu mưa GPM bằng Google Earth Engine (GEE)



Hình 4. Xử lý dữ liệu mưa GPM bằng QGIS

3.2. Phân tích, đánh giá các dữ liệu mưa

a. Đánh giá khả năng phát hiện mưa

Các chỉ số POD và CSI được sử dụng để phân tích khả năng phát hiện lượng mưa của tất cả các sản phẩm mưa lưới được chọn và 0,1 mm/giờ được xác định là ngưỡng phát hiện mưa/không mưa. Nhìn chung, POD của tất cả các dữ liệu mưa lưới cho thấy giá trị khá tốt, lớn hơn 0,7 (Bảng 4), chỉ có POD của CCS ở trận mưa

cuối tháng 10/2020 là cho giá trị tương đối thấp, khoảng 0,47. Chỉ số POD của GSMaP và ERA5-Land có giá trị cao nhất, hầu hết đều trên 0,9, kết quả này cho thấy dữ liệu mưa GSMaP và ERA5-Land thường có nhiều giờ mưa hơn so với thực tế. Mặc dù mưa GSMaP và ERA5-Land có chỉ số POD cao hơn các loại mưa khác, nhưng cũng có chỉ số FAR đều cao hơn so với mưa CCS, điều đó có nghĩa là mưa GSMaP và ERA5-Land có xu hướng xác không trường hợp mưa.

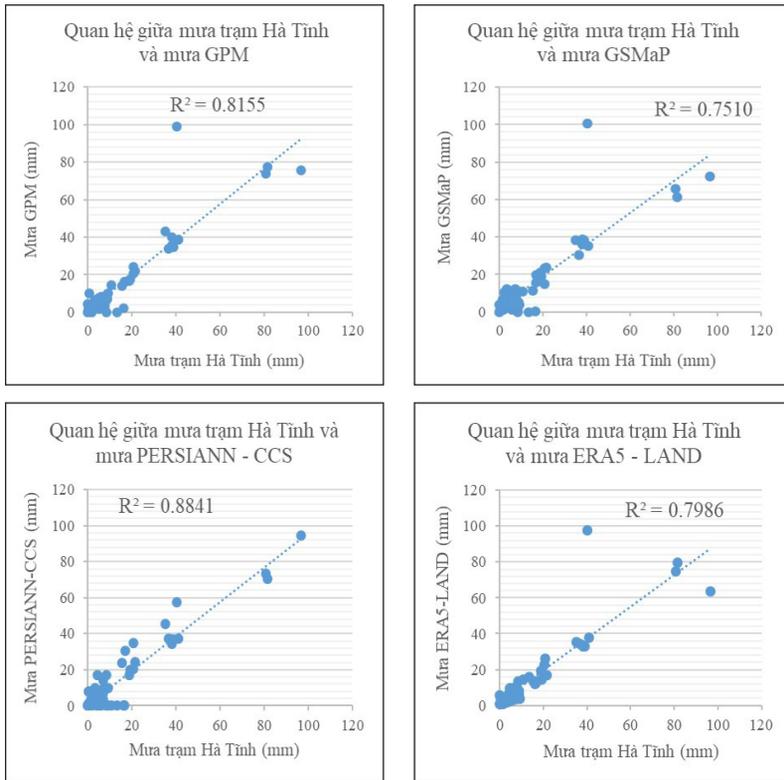
Bảng 4. Các chỉ số thống kê giữa dữ liệu mưa quan trắc tại trạm Hà Tĩnh và các sản phẩm mưa lưới

Thời gian	Sản phẩm mưa	R ²	RMSE (mm)	MAE (mm)	PBIAS (%)	POD	FAR	CSI
13 - 16/10/2016	GPM	0,816	8,060	3,163	-0,78	0,94	0,12	0,84
	GSMaP	0,751	9,091	4,263	-1,29	0,96	0,20	0,78
	CCS	0,884	6,188	4,030	0,71	0,74	0,00	0,74
	ERA5-Land	0,799	8,173	3,295	-1,25	1,00	0,26	0,74
17 - 19/09/2020	GPM	0,742	4,981	2,319	-16,84	0,93	0,48	0,50
	GSMaP	0,639	6,340	1,214	-17,66	0,93	0,29	0,68
	CCS	0,843	3,879	1,351	-2,74	0,76	0,33	0,55
	ERA5-Land	0,816	4,199	1,098	-1,88	0,97	0,53	0,46
16 - 19/10/2020	GPM	0,881	5,205	3,654	0,61	0,81	0,15	0,71
	GSMaP	0,872	5,294	3,626	-1,34	0,95	0,09	0,87
	CCS	0,888	5,010	3,453	-0,90	0,86	0,07	0,81
	ERA5-Land	0,876	5,207	3,451	-1,87	0,99	0,13	0,86
29 - 31/10/2020	GPM	0,815	7,618	4,554	-1,78	0,70	0,06	0,67
	GSMaP	0,831	7,018	3,595	-4,97	0,98	0,08	0,90
	CCS	0,849	7,128	4,158	1,66	0,47	0,00	0,47
	ERA5-Land	0,841	6,825	3,432	-2,20	1,00	0,15	0,85

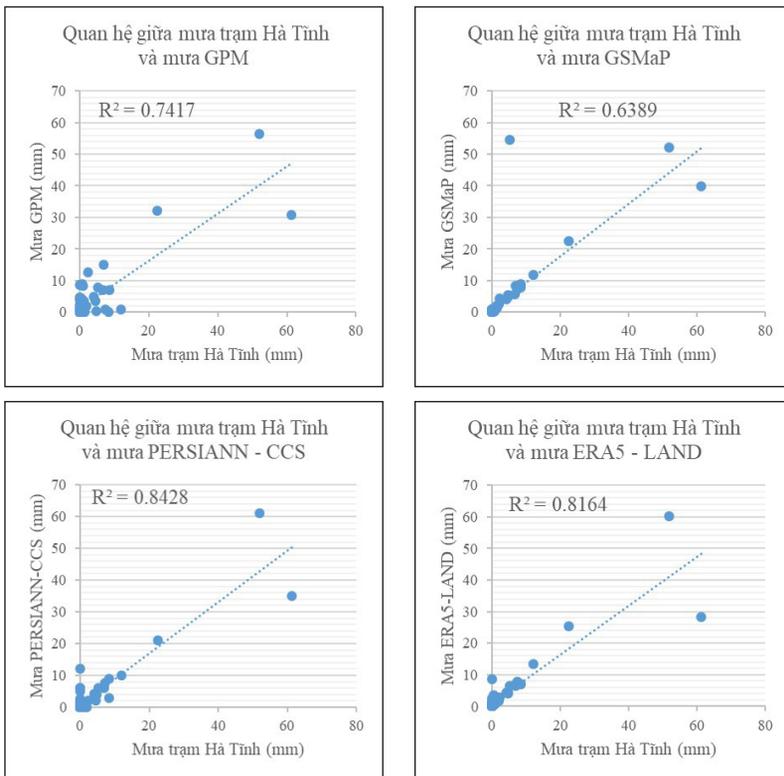
b. Đánh giá sai số lượng mưa

Hệ số tương quan giữa lượng mưa giờ tại trạm quan trắc Hà Tĩnh và các sản phẩm mưa lưới khá cao khoảng 0,64 đến 0,88; trong đó mưa PERSIANN-CCS có tương quan cao nhất với

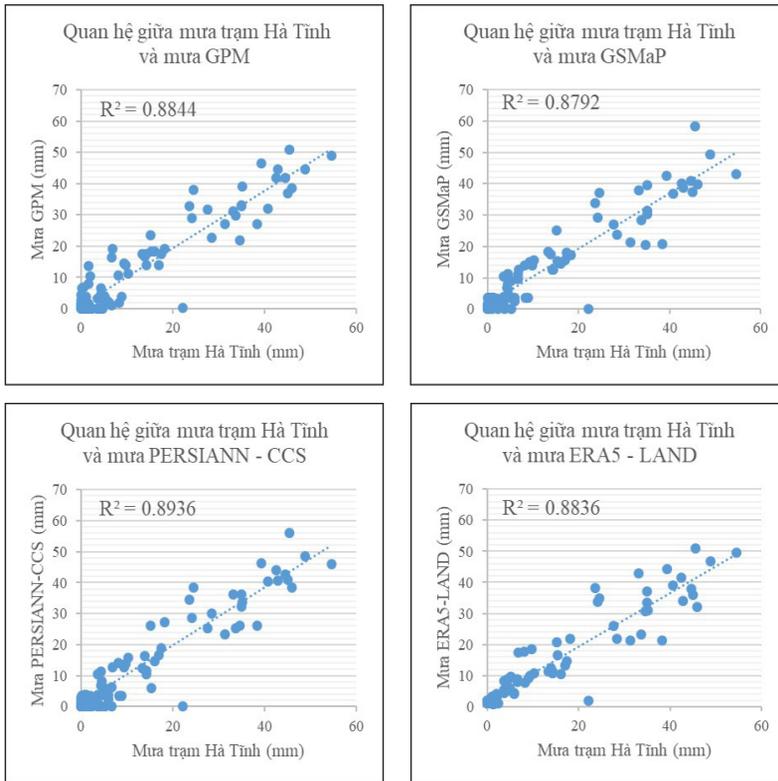
mưa quan trắc mặt đất với hệ số trên 0,8 ở tất cả 4 trận mưa, mưa GSMaP có hệ số tương quan thấp nhất với hệ số của trận mưa 17 - 19/9/2020 chỉ đạt 0,64 so với mưa quan trắc mặt đất (Bảng 4 và các Hình 5, Hình 6, Hình 7, Hình 8).



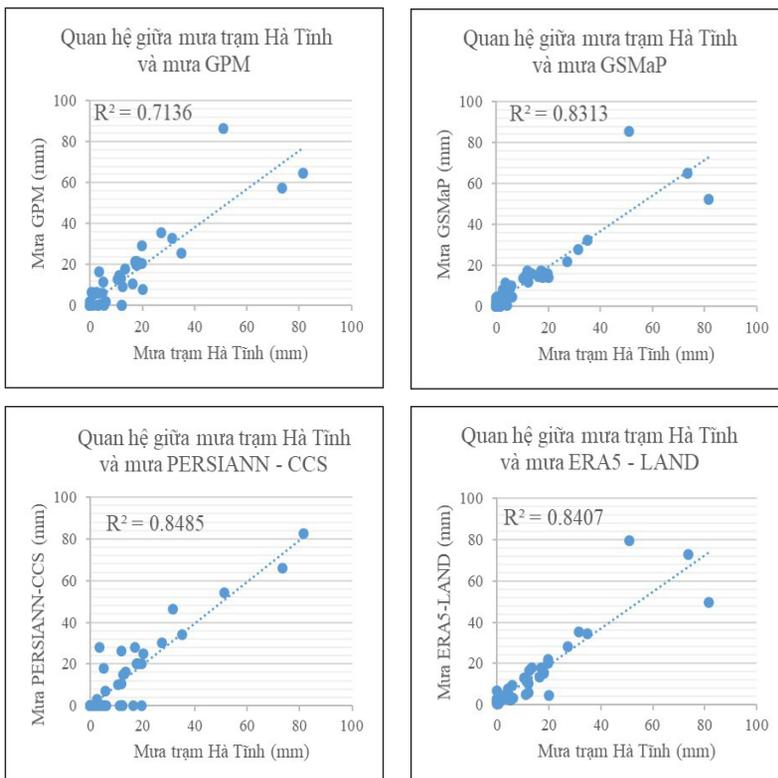
Hình 5. Tương quan giữa mưa trạm Hà Tĩnh và mưa lưới trận mưa 13 - 16/10/2016



Hình 6. Tương quan giữa mưa trạm Hà Tĩnh và các mưa lưới trận mưa 17 - 19/9/2020



Hình 7. Tương quan giữa mưa trạm Hà Tĩnh và mưa lưới trận mưa 16 - 19/10/2020



Hình 8. Tương quan giữa mưa trạm Hà Tĩnh và mưa lưới trận mưa 29 - 31/10/2020

Chỉ số sai số quân phương trung bình (RMSE) của dữ liệu mưa CCS là thấp nhất, dao động từ 3,87 - 7,13 mm; sản phẩm mưa GSMaP có RMSE cao nhất (nghĩa là độ chính xác kém nhất), dao động từ 5,29 - 9,09 mm (Bảng 4). Còn đối với chỉ số MAE thì dữ liệu mưa GPM thường có chỉ số cao nhất, sau đó đến dữ liệu mưa GSMaP và cuối cùng là dữ liệu mưa ERA5-Land. Với 2 chỉ số này có thể thấy rằng: Chỉ số MAE của mưa GSMaP chỉ xếp sau mưa GPM và có trệch cao hơn cả GPM. Còn RMSE của mưa GSMaP lại cao nhất. Vì vậy xét theo 2 chỉ số này thì dữ liệu mưa PERSIANN-CCS cho kết quả ở mức cân bằng, còn dữ liệu mưa GSMaP cho độ chính xác kém nhất trong 4 loại mưa.

Phân bố sai số độ lệch (PBIAs) cho thấy hầu hết các sản phẩm mưa đều có sai số nhỏ, hầu hết đều dao động trong khoảng 10% (nằm ở mức tốt). Chỉ có trệch mưa 17 - 19/9/2020 là có PBIAs trên 15% (nằm ở mức đạt) đối với dữ liệu mưa GPM và GSMaP (Bảng 4).

Kết quả phân tích chỉ số thống kê của các dữ liệu mưa lưới cho thấy, nhìn chung dữ liệu mưa PERSIANN-CCS có sai lệch nhỏ hay có thể coi nó là sản phẩm ước tính tốt nhất trong số 4 dữ liệu so với số liệu quan trắc.

4. Kết luận

Nghiên cứu đã sử dụng các chỉ số đánh giá để đánh giá lượng mưa giờ của các dữ liệu mưa lưới GPM, GSMaP, PERSIANN-CCS, ERA5-Land tương ứng với dữ liệu tại trạm Hà Tĩnh. Dựa trên kết quả tính toán các chỉ số, nghiên cứu đã phân tích và đánh giá cho thấy: Các dữ liệu mưa lưới đều cung cấp đầu vào khá tốt, hệ số tương quan so với trạm đo mặt đất khá cao (đều trên 0,7), đặc biệt là dữ liệu mưa PERSIANN-CCS cho kết quả tốt nhất (hệ số tương quan trên 0,8). Khả năng phát hiện mưa của mưa lưới là khá tốt, chỉ số POD của tất cả các dữ liệu mưa lưới đều lớn hơn 0,7, chỉ có riêng trệch mưa cuối tháng

10/2020 thì dữ liệu mưa CCS cho chỉ số tương đối thấp (đạt 0,47). Chỉ số POD của GSMaP và ERA5-Land có giá trị cao nhất, hầu hết đều trên 0,9; kết quả này cho thấy dữ liệu mưa GSMaP và ERA5-Land thường có nhiều giờ mưa hơn so với thực tế. Tuy nhiên, mưa GSMaP và ERA5-Land lại có chỉ số FAR cao hơn so với các dữ liệu mưa khác, có nghĩa là 2 loại mưa này có xu hướng xác định không trường hợp mưa lớn hơn. Chỉ số sai số quân phương trung bình (RMSE) của dữ liệu mưa CCS là thấp nhất, dao động từ 3,87 - 7,13 mm; sản phẩm mưa GSMaP có RMSE cao nhất, dao động từ 5,29 - 9,09 mm.

Kết quả cho thấy đánh giá mưa giờ cho các nguồn mưa lưới tại khu vực Hà Tĩnh có nhiều điểm khác biệt so với những kết quả nghiên cứu đi trước về đánh giá chất lượng mưa lưới trung bình tháng, trung bình năm. Điều này khẳng định thêm nhận định ban đầu, tuy cùng một nguồn dữ liệu gốc, nhưng với các cách tiếp cận xử lý số liệu, tùy theo từng vùng mà chất lượng của mỗi loại mưa lưới sẽ khác nhau và luôn cần có đánh giá cụ thể trước khi sử dụng chúng trong các ứng dụng tính toán, mô hình. Qua các kết quả tính toán, đánh giá, nghiên cứu kiến nghị có thể sử dụng dữ liệu mưa PERSIANN-CCS để phục vụ công tác dự báo, cảnh báo ngập lụt đô thị cho thành phố Hà Tĩnh, do kết quả tính toán đánh giá có sự phân hạng theo các chỉ tiêu đánh giá ở mức cân bằng.

Hướng nghiên cứu tiếp theo sẽ đánh giá chất lượng thực tế của các nguồn mưa lưới khi đưa vào trong mô hình thủy văn đô thị, xem xét sự phân bố về không gian của mưa đầu vào có ảnh hưởng như thế nào đến kết quả mô phỏng dòng chảy trong hệ thống thoát nước và ngập lụt trên bề mặt đô thị. Với đặc điểm vượt trội là mưa giờ, dữ liệu mưa lưới PERSIANN-CCS sẽ hỗ trợ được việc thiết lập mô hình dự báo cảnh báo ngập lụt đô thị, đảm bảo được cho yêu cầu của ứng dụng thủy văn đô thị.

Tài liệu tham khảo

Tài liệu tiếng Việt

1. Bộ Tài nguyên và Môi trường (2017), *Thông tư số 41/2017/TT-BTNMT: Quy định kỹ thuật đánh giá chất lượng dự báo, cảnh báo khí tượng*, TT số 41/2017/TT-BTNMT. Bộ Tài nguyên và Môi trường, Việt Nam.

2. Bùi Tuấn Hải, Nguyễn Văn Tuấn (2018), "Nghiên cứu đánh giá và so sánh các dữ liệu mưa vệ tinh độ phân giải cao lưu vực sông Cả", *Tạp chí Khí tượng Thủy văn* số tháng 11 - 2018.
3. Phạm Thành Hưng, Nguyễn Quang Bình, Võ Nguyễn Đức Phước (2021), "Đánh giá khả năng sử dụng các sản phẩm mưa vệ tinh để mô phỏng dòng chảy bằng mô hình thủy văn", *Tạp chí Khoa học Kỹ thuật Thủy lợi và Môi trường*, số 74 (6/2021).
4. Nguyễn Tiến Kiên (2020), "Nghiên cứu ứng dụng số liệu mưa vệ tinh mô phỏng lũ khu vực trung lưu sông Mã", *Tạp chí khí tượng thủy văn*, số tháng 1-2020.
5. Bùi Chí Nam (2017), "Nghiên cứu đánh giá dữ liệu mưa quan trắc vệ tinh từ GPM và PERSIANN phục vụ cảnh báo mưa thành phố Hồ Chí Minh", *Tạp chí Khí tượng Thủy văn* số tháng 7-2017.

Tài liệu tiếng Anh

6. Amorim, J. et al (2020), "Evaluation of satellite precipitation products for hydrological modeling in the Brazilian Cerrado Biome", *Water* 12, 2571. <https://doi.org/10.3390/w12092571>.
7. Behrangi, A., et al (2011), "Hydrologic evaluation of satellite precipitation products over a mid-size basin", *J. Hydrol.* 397, 225-237. <https://doi.org/10.1016/j.jhydrol.2010.11.043>.
8. Chen M. and Xie P, (2007), *CPC Unified Gauge-based Analysis of Global Daily Precipitation* [Internet].
9. Ebert E.E., et al (2007), "Comparison of Near-Real-Time Precipitation Estimates from Satellite Observations and Numerical Models", *Bulletin of the American Meteorological Society*, Volume 88: Issue 1, pp. 47–64. <https://doi.org/10.1175/BAMS-88-1-47>.
10. Harris I., et al (2014), "Updated high-resolution grids of monthly climatic observations - the CRU TS3.10 Dataset", *International Journal of Climatology*, 34, pp. 623–42. <https://doi.org/10.1002/joc.3711>
11. Hung, M.K. et al (2020), "Application of GSMaP Satellite data in precipitation estimation and nowcasting: evaluations for October 2019 to January 2020 period for Vietnam", *VN J. Hydrometeorol.* 5, 80-94.
12. <https://chrsdata.eng.uci.edu/>.
13. <https://gpm.nasa.gov/data/directory>
14. Kaptue, A.T. et al (2015), "Spatial and temporal characteristics of rainfall in africa: Summary statistics for temporal downscaling", *Water Resour. Res.* 51, 2668–2679.
15. Maggioni, V. et al (2016), "A review of merged high-resolution satellite precipitation product accuracy during the Tropical Rainfall Measuring Mission (TRMM) era", *J. Hydrometeorol.* 17, 1101–1117. <https://doi.org/10.1175/JHM-D-15-0190.1>
16. Mohamed Salem Nashwan et al (2019), "Assessment of Satellite-Based Precipitation Measurement Products over the Hot Desert Climate of Egypt", *Remote Sens.* 11 (5), 555; doi:10.3390/rs11050555, www.mdpi.com/journal/remotesensing
17. Nguyen-Xuan T. et al. (2016), "The Vietnam Gridded Precipitation (VnGP) Dataset: Construction and Validation", *Sola*, 12, pp. 291-6. <https://doi.org/10.2151/sola.2016-057>
18. Saimy Davis et al (2022), "A Flood Forecasting Framework Coupling a High Resolution WRF Ensemble With an Urban Hydrologic Model", *Frontiers in Earth Science*, doi:10.3389/feart.2022.883842.
19. Schneider U., et al (2014), "GPCC's new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle", *Theoretical and Applied Climatology*, 115, pp. 15–40. <https://doi.org/10.1007/s00704-013-0860-x>
20. Talchabhadel et al (2021), "Evaluation of precipitation elasticity using precipitation data from ground and satellite-based estimates and watershed modeling in Western Nepal", *J. Hydrol. Reg. Stud.* 33, 100768 <https://doi.org/10.1016/j.ejrh.2020.100768>
21. Thanh, N.D. et al (2013), "Monthly adjustment of Global Satellite Mapping of Precipitation (GSMaP) data over the Vu Gia – Thu Bon River Basin in Central Vietnam using an artificial neural network", *Hydrol. Res. Lett.* 2013, 7, 85–90.

22. Tze Huey Tam et al (2019), "Application of Satellite Rainfall Products for Flood Inundation Modelling in Kelantan River Basin, Malaysia", *Hydrology*, 6, 95; doi:10.3390/hydrology6040095.
23. Yatagai A. et al (2012), "APHRODITE: Constructing a Long-Term Daily Gridded Precipitation Dataset for Asia Based on a Dense Network of Rain Gauges", *Bulletin of the American Meteorological Society*, 93, pp. 1401–15. <https://doi.org/10.1175/BAMS-D-11-00122.1>

EVALUATION OF THE QUALITY OF GRIDDED HEAVY RAINFALL OVER THE HA TINH CITY

Nguyen Thi Lien, Nguyen Quang Hung, Trinh Minh Ngoc
University of Science, Vietnam National University, Hanoi

Received: 3/11/2022; Accepted: 28/11/2022

Abstract: Gridded rainfall products which is the production generated by satellite together with weather numerical models combined with ground rain are increasingly developed based on its advantages outstanding ground rainfall measurement. Gridded rainfall products provide spatial information and its interval time is up to 30 minutes, and freely available, that makes it became very high applicability. Although the products are from the same sensors, still the algorithms used for correction, interpolation and matching are different, so there is a huge difference between the products. This study considers rain data from 04 sources including GPM, GSMaP, PERSIANN-CCS and ERA5-Land over Ha Tinh city area to assess their quality and practical applicability. Four rainfall events are used to assess the hourly rainfall compare to the data at rain gauge station in Ha Tinh city. Statistical indicators RMSE, R^2 , PBIAs, POD, FAR and CSI were used for the evaluation. The results show that in general, all four rain data have good quality, PERSIANN-CCS has the highest correlation with ground-observed rain, the best mean square error index, and other indicators show that GSMaP data is generally the worst quality

Từ khóa: Gridded rainfall products, satellite rainfall, Hà Tĩnh city.