

# MÔ HÌNH HÓA MƯA ĐỘ PHÂN GIẢI CAO KẾT HỢP GIỮA MÔ HÌNH ĐỘNG LỰC KHÍ TƯỢNG VÀ PHƯƠNG PHÁP THỐNG KÊ: ỨNG DỤNG CHO LƯU VỰC SÔNG SÀI GÒN - ĐỒNG NAI

Trịnh Quang Toàn

Viện Sinh thái và Bảo vệ công trình

Đỗ Hoài Nam

Trung tâm Đào tạo và Hợp tác Quốc tế

Nguyễn Kỳ Phùng

Viện Khoa học và Công nghệ Tính toán

Nguyễn Văn Thắng

Trường Đại học Quốc Gia Tp.Hồ Chí Minh

**Tóm tắt:** Mô hình hóa các trận mưa lớn đóng một vai trò quan trọng trong quản lý tài nguyên nước. Trong nghiên cứu này, phương pháp kết hợp giữa chi tiết hóa động lực học và phương pháp hiệu chỉnh thống kê đã được áp dụng. Công nghệ chi tiết hóa đề xuất sử dụng đầu vào được cung cấp từ ba bộ dữ liệu toàn cầu khác nhau bao gồm ERA-Interim, ERA20C và CFSR. Dữ liệu toàn cầu này được chi tiết hóa bằng mô hình Nghiên cứu và Dự báo Thời tiết (WRF), sau đó sử dụng phương pháp thống kê nhằm nâng cao tính chính xác cũng như hiệu quả trong công việc chi tiết hóa mưa có độ phân giải cao (9km) trên lưu vực. Kết quả tính toán mô phỏng cả 03 ba bộ dữ liệu có độ tin cậy tốt và đạt các chỉ tiêu thống kê. Việc thiết lập được hệ thống tính toán khôi phục xu thế diễn biến của các điều kiện khí tượng, thủy văn trên toàn bộ lưu vực tạo điều kiện cho các nhà nghiên cứu đánh giá tổng quát được quá trình hình thành các trận mưa lũ trong quá khứ, từ đó giúp chủ động ứng phó với tình hình thiên tai mưa lũ trên lưu vực.

**Từ khóa:** Mô hình nghiên cứu và dự báo thời tiết (WRF), hiệu chỉnh thống kê, dữ liệu toàn cầu (ERA-Interim, ERA20C và CFSR).

**Summary:** The modeling of large rainfall events play an important role in water resources management. In this study, a blended technique combining dynamical and statistical approaches has been explored. The proposed downscaling technology uses input provided from three different global reanalysis data including ERA-Interim, ERA20C, and CFSR. These reanalysis atmospheric data are downscaled by means of the Weather Research and Forecasting (WRF) model followed by the application of a statistical method to improve accuracy and further downscale high resolution (9km) over the studied basin. Simulations of all three data sets have good reliability and reach the statistical indicators that can be provided as inputs of the hydrological and environmental models. Among the three selected reanalysis datasets, the best calibration and validation results were obtained from the ERA-Interim dataset.

**Key words:** Weather Research and Forecasting Model (WRF), bias correction, reanalysis data (ERA-Interim, ERA20C và CFSR).

## 1. ĐẶT VẤN ĐỀ

Khôi phục và mô phỏng dữ liệu khí tượng-thủy văn trong điều kiện hạn chế về số liệu, hoặc không có số liệu là một vấn đề bức thiết hàng đầu trong công tác quy hoạch và quản lý

các lưu vực sông. Các số liệu khí tượng-thủy văn có độ tin cậy cao trong quá khứ là tiền đề để xây dựng các đánh giá, phân tích các đặc tính biến đổi về điều kiện thủy văn và các ảnh hưởng lên kinh tế xã hội. Trong quá khứ, các nghiên

Ngày nhận bài: 09/7/2021

Ngày thông qua phản biện: 08/9/2021

Ngày duyệt đăng: 04/10/2021

cứ khôi phục số liệu khí tượng-thủy văn chủ yếu sử dụng các dữ liệu mưa thực đo, và tính toán bằng các mô hình toán mưa - dòng chảy, hoặc sử dụng các phương pháp thống kê thông qua các chuỗi số liệu thủy văn trong quá khứ. Tuy nhiên, nếu trong điều kiện không có số liệu đo đạc mưa, dòng chảy, hoặc dữ liệu không liên tục, hoặc ở các bước thời gian quá dài (tháng, năm), thì các phương pháp trước đây không còn thích hợp.

Việc tính toán mô phỏng dữ liệu khí tượng thủy văn trong điều kiện ít số liệu, hoặc không có số liệu thường được bắt đầu bằng việc tính toán khôi phục dữ liệu khí tượng. Một số các phương pháp tính toán đã được áp dụng, như sử dụng các ảnh vệ tinh, hay sử dụng số liệu toàn cầu trong quá khứ (Global climate models-GCMs, và reanalysis data) [1] [2] [3] [17] [18].

Các bộ dữ liệu toàn cầu trong quá khứ (reanalysis) thường ở độ phân giải rất thô. Các dữ liệu nguyên gốc của các GCMs/reanalysis chỉ có thể mô tả được xu thế của một vùng nghiên cứu rộng lớn như 1 quốc gia có diện tích trung bình cỡ Việt Nam. Để nghiên cứu những biến đổi các yếu tố thời tiết khí hậu, hay thủy văn ở phạm vi lưu vực thì phải áp dụng các phương pháp kỹ thuật chi tiết hóa dữ liệu (downscaling) [4].

Thông thường, có hai cách tiếp cận về các phương pháp chi tiết hóa dữ liệu bao gồm phương pháp chi tiết hóa thống kê và phương pháp chi tiết hóa động lực [5] [6] [7] [20]. Phương pháp chi tiết hóa thống kê lại dựa vào mối quan hệ thực nghiệm về cả không gian lẫn thời gian trong quá khứ giữa các số liệu khí tượng với độ phân giải thô từ GCMs với các số liệu thực đo. Sau đó, mối quan hệ tương quan này sẽ được giả định giữ nguyên trong các tính toán tiếp theo cho tương lai hay kéo dài bộ dữ liệu trong quá khứ [8] [7] [9] [10] [11]. Phương pháp chi tiết hóa thống kê cho phép tính toán nhanh chóng và dễ sử dụng hơn nhiều so với phương pháp động lực. Tuy nhiên, các phương pháp này phụ thuộc rất nhiều vào dữ liệu quan

trắc với nhược điểm là thời đoạn đo ngắn, số liệu không liên tục, tần suất đo thưa (ngày, tháng) và mật độ quan trắc theo không gian ít (>25km). Phương pháp chi tiết hóa động lực cho phép mô phỏng chi tiết các thông số khí tượng thủy văn cho toàn bộ lưu vực. Phương pháp chi tiết hóa động lực sẽ sử dụng đầu ra của các bộ dữ liệu toàn cầu để làm các điều kiện biên cũng như điều kiện ban đầu để mô phỏng các dữ liệu khí tượng chi tiết hơn [12] [19]. Tuy nhiên, hạn chế của phương pháp này là cần tài nguyên máy tính và mất nhiều thời gian để mô phỏng chi tiết các yếu tố khí tượng ở độ phân giải cao.

Trong khi cả phương pháp thống kê và phương pháp động lực đều bộc lộ ưu và nhược điểm trong mỗi phương pháp. Việc kết hợp cả hai phương pháp trên, được gọi là phương pháp kết hợp [13] [14], để mô phỏng cũng như là khôi phục dữ liệu khí tượng đang được phát triển và phổ biến. Mục đích chính của nghiên cứu này là ứng dụng công nghệ kết hợp giữa phương pháp chi tiết hóa động lực (dynamical downscaling) và phương pháp thống kê (bias correction), nhằm nâng cao tính chính xác cũng như hiệu quả trong công việc chi tiết hóa mưa có độ phân giải cao (9km) trên lưu vực. Trong đó, phương pháp thống kê được áp dụng để hiệu chỉnh kết quả mô phỏng bởi phương pháp chi tiết hóa động lực sử dụng mô hình WRF. Mô hình WRF được thiết lập có xét tính khả thi về tài nguyên máy tính; đầu vào của mô hình WRF là các số liệu khí tượng, khí hậu toàn cầu, và đầu ra của WRF bao gồm các dữ liệu khí tượng như mưa gió, nhiệt độ, áp suất, bức xạ... Sự kết hợp của hai phương pháp động lực và thống kê sẽ nâng cao độ chính xác trong các mô phỏng khí tượng, qua đó sẽ nâng cao độ tin cậy cho các bước tính toán mô phỏng tiếp theo như mô hình thủy văn lưu vực.

## 2. PHƯƠNG PHÁP NGHIÊN CỨU

### 2.1. Vùng nghiên cứu

Lưu vực hệ thống sông Đồng Nai nằm trong

khoảng vĩ độ 10°20' – 12°20' Bắc, kinh độ 105°45' – 109°15' Đông, bao gồm các tỉnh Lâm Đồng, Bình Thuận, Đồng Nai, Bình Phước, Bình Dương, Tây Ninh, Tp.HCM, Bà Rịa-Vũng Tàu, một phần tỉnh Long An và Đắk- Nông, với phía Bắc và Đông Bắc giáp 2 tỉnh Khánh Hòa và Đắk-Lak, phía tây giáp quốc gia Campuchia, phía Đông giáp phần còn lại của Bình Thuận, Bà Rịa-Vũng Tàu, Ninh Thuận và biển, phía Nam giáp phần còn lại của tỉnh Long An và tỉnh Tiền Giang. Tổng diện tích tự nhiên khoảng 49.644 km<sup>2</sup>. Hệ thống sông Đồng Nai-Sài Gòn gồm các sông chính: Đồng Nai, La Ngà, Bé, Sài Gòn, Vàm Cỏ Đông, Vàm Cỏ Tây.

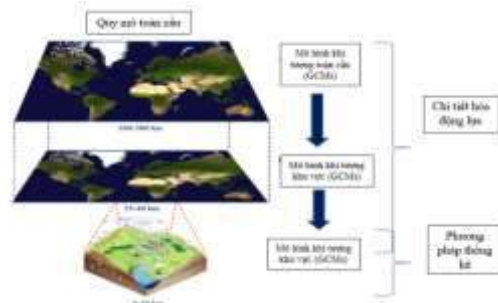
**2.2. Phương pháp nghiên cứu**

Chi tiết hóa dữ liệu mưa (độ phân giải 9km) từ năm 1981 – 2010 ở lưu vực sông Sài Gòn-Đồng Nai sẽ được thực hiện theo các bước sau đây:

- Bước 1: Thu thập dữ liệu
- Bước 2: Thiết lập mô hình chi tiết hóa động lực (WRF) cho 3 bộ dữ liệu khác nhau
- Bước 3: Hiệu chỉnh và kiểm định mô hình chi tiết hóa động lực WRF.
- Bước 4: Hiệu chỉnh và kiểm định phương pháp thống kê với đầu vào được cung cấp từ đầu ra của mô hình WRF (D3).



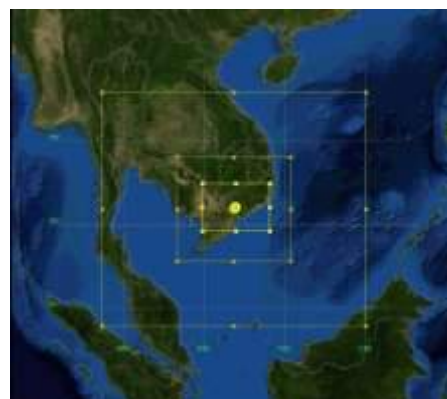
Hình 1: Lưu vực sông Sài Gòn – Đồng Nai



Hình 2: Phương pháp tính toán mô phỏng mưa cho lưu vực sông Sài Gòn-Đồng Nai

**2.2.1. Thiết lập mô hình chi tiết hóa động lực WRF cho ba bộ dữ liệu khác nhau**

Để thiết lập mô hình WRF cho việc mô phỏng chi tiết các quá trình khí tượng trên toàn bộ lưu vực sông Sài Gòn – Đồng Nai thì bước đầu tiên là thiết lập miền tính toán. Để có thể đảm bảo mô phỏng ổn định các quá trình khí tượng, các mô hình RCM, hay cụ thể là mô hình WRF thường được khuyến cáo là chi tiết hóa dữ liệu từ các dữ liệu toàn cầu trong quá khứ xuống các ô lưới tính toán nhỏ hơn, và thông thường bằng 1/3 các ô lưới trước đó. Dữ liệu gốc của ERA-20C có độ phân giải là 125x125 km, do đó trong nghiên cứu này sẽ thiết lập cho ba miền tính toán (D1, D2, D3) như sau: (D1) chi tiết hóa các yếu tố khí tượng từ mô hình toàn cầu có độ phân giải thô (~125km) xuống độ phân giải trung bình (81km); (D2) chi tiết hóa các yếu tố khí tượng từ miền D1 xuống độ phân giải cao (27km); và (D3) chi tiết hóa các yếu tố khí tượng từ miền D2 xuống độ phân giải rất cao (9km) (hình 3).



Hình 3: Thiết lập miền tính toán

### 2.2.2. Thiết lập mô hình chi tiết hóa động lực WRF cho ba bộ dữ liệu khác nhau

Sau khi thiết lập thành công mô hình chi tiết hóa khí tượng WRF, tiến hành mô phỏng các yếu tố khí tượng cho toàn bộ lưu vực nghiên cứu. Hiệu chỉnh kết quả mô phỏng từ mô hình WRF chúng ta cần thay đổi sơ đồ tham số hóa vật lý. Các sơ đồ tham số hóa vật lý trong mô hình WRF có thể chia làm năm loại chính, trong đó mỗi loại bao gồm một số sơ đồ khác nhau. Chúng bao gồm (1) các quá trình vi vật lý, (2) tham số hóa đối lưu, (3) lớp biên hành tinh (PBL), (4) mô hình mặt đất, và (5) sự bức xạ. Sự khuếch tán, cũng có thể coi là một phần của các quá trình vật lý trong mô hình, được xử lý trực tiếp trong mô hình. Quá trình tham số hóa cho các thành phần vật lý được ưu tiên thực hiện ở bước đầu tiên của mô hình. Quá trình này có thể bao gồm việc đọc file số liệu trong đó người sử dụng sẽ tùy chọn sử dụng thành phần vật lý nào. Mỗi sơ đồ tham số hóa vật lý được đóng gói sẵn thành một module riêng biệt trong đó chứa bản thân sơ đồ tham số hóa và các hằng số

của riêng chúng, bên cạnh các hằng số sử dụng chung.

Phân tích độ nhạy sơ đồ tham số hóa mô hình nhằm xác định ảnh hưởng của sự thay đổi từng sơ đồ của mô hình lên kết quả và nhằm chọn ra sơ đồ tham số hóa vật lý có ảnh hưởng lớn để tiết kiệm thời gian và khối lượng tính toán trong bước hiệu chỉnh mô hình.

Bảng 1 thể hiện 12 sơ đồ tham số hóa được xác định dựa vào các nghiên cứu trước đây ở các lưu vực. Từ sự so sánh 12 sơ đồ tham số hóa, sơ đồ tốt nhất được lựa chọn dựa trên việc kiểm tra hệ số tương quan giữa lượng mưa trung bình hàng ngày được mô phỏng trên lưu vực và dữ liệu mưa thực đo VnGP. So sánh giữa 12 sơ đồ tham số hóa sử dụng dữ liệu mưa ngày từ ngày 1 tháng 1 năm 1994 đến ngày 31 tháng 12 năm 1995. Năm 1994 -1995 được lựa chọn để hiệu chỉnh kết quả mô phỏng bởi vì đây là giai đoạn lịch sử xảy ra những trận lũ lụt cực lớn. Cũng cần lưu ý rằng dữ liệu này được so sánh với miền tính D3.

**Bảng 1: Các sơ đồ tham số hóa vật lý để cấu hình mô hình WRF**

STT	Quá trình vi vật lý	Tham số hóa đối lưu	Lớp biên hành tinh	Tham số hóa bức xạ	Mô hình mặt đất
1	WSM3	New SAS	BouLac	New Goddard	RUC Land
2	Eta (Ferrier)	New SAS	BouLac	New Goddard	RUC Land
3	Goddard	New SAS	BouLac	New Goddard	RUC Land
4	Milbrandt 2-mom	New SAS	BouLac	New Goddard	RUC Land
5	CAM 5.1	New SAS	BouLac	New Goddard	RUC Land
6	SBU-YLin	New SAS	BouLac	New Goddard	RUC Land
7	WSM3	Kain-Fritsch	BouLac	New Goddard	RUC Land
8	WSM3	Grell-Freitas	BouLac	New Goddard	RUC Land
9	WSM3	Grell-3	BouLac	New Goddard	RUC Land
10	WSM3	New SAS	MYNN2	New Goddard	RUC Land
11	WSM3	New SAS	YSU	New Goddard	RUC Land
12	WSM3	New SAS	UW	New Goddard	RUC Land

Đánh giá kết quả của hiệu chỉnh của mô hình dựa trên những so sánh lượng mưa ngày trung bình lưu vực (miền tính D3) với thời gian hiệu chỉnh là từ năm 1/1/1994 đến năm 31/12/1995

với 12 bộ sơ đồ tham số hóa (12 mô hình) đã chọn được thể hiện ở bảng 2. Kết quả ở bảng 2 chỉ ra rằng mô hình số 1 sử dụng sơ đồ tham số hóa quá trình vi vật lý WSM3 Hong, Dudhia và

chen (2004, MWR) và sơ đồ tham số hóa đối lưu New SAS Han and Pan (2011, Wea. Forecasting) cho kết quả hệ số tương quan cao nhất cho bộ dữ liệu toàn cầu trong quá khứ ERA-Interim và ERA-20C tại lưu vực mục tiêu (miền tính D3).

**Bảng 2: So sánh hệ số tương quan R<sup>2</sup> của lượng mưa ngày mô phỏng bởi 12 mô hình với dữ liệu mưa thực đo trung bình lưu vực từ 1/1/1994 đến năm 31/12/1995**

STT	ERA-Interim	ERA-20C	CFSR
1	<b>0.72</b>	<b>0.65</b>	0.61
2	0.64	0.57	0.56
3	0.61	0.52	0.53
4	0.60	0.51	0.55
5	0.62	0.502	0.51
6	0.69	0.61	<b>0.69</b>
7	0.59	0.51	0.51
8	0.58	0.50	0.62
9	0.63	0.55	0.65
10	0.63	0.59	0.63
11	0.68	0.62	0.62
12	0.65	0.60	0.61

Mô hình số 6 sử dụng sơ đồ tham số hóa quá trình vi vật lý SBU-YLin, Lin and Colle (2011, MWR) và sơ đồ tham số hóa đối lưu New SAS Han and Pan (2011, Wea. Forecasting) cho kết quả hệ số tương quan cao nhất đối với bộ dữ liệu toàn cầu trong quá khứ CFSR. Kết quả hiệu chỉnh với sơ đồ tham số hóa vật lý đã chọn cho thấy sự phù hợp giữa giá trị mưa thực đo và mô phỏng.

Sau khi thiết lập được các sơ đồ tham số hóa vật lý mô phỏng cho các bộ dữ liệu ERA20C, ERA-Interim và CFSR. Để kiểm tra độ tin cậy và phù hợp của mô hình, tiến hành kiểm định một giai đoạn. Hình 3 đến 8 biểu diễn sự so sánh chuỗi dữ liệu thực đo và mô phỏng theo trung bình lưu vực 1, 3, 5, 7 ngày trên toàn bộ lưu vực sông Sài Gòn – Đồng Nai trong giai đoạn 1986 – 1995 (10 năm). Bằng cách so sánh trực quan giữa mô phỏng WRF tại miền tính D3 và các giá trị mưa thực đo theo lượng mưa trung bình lưu vực 1, 3, 5, 7 ngày tương ứng, kết quả.

cho thấy đạt yêu cầu. Mặc dù, nhìn vào biểu đồ mô phỏng theo thời gian của kết quả mô phỏng và thực đo tại 1 số thời điểm đỉnh mưa lớn của mưa mô phỏng đôi khi cao hơn mưa thực đo, nhưng sự khác biệt không quá lớn, đường quá trình và xu hướng của mưa tính toán và thực đo có sự tương đồng với nhau cả về cường độ và thời gian xuất hiện. Bảng 3, 4, 5 liệt kê các giá trị kiểm tra thống kê của kết quả mô phỏng của mô hình WRF trong giai đoạn kiểm định với 3 bộ dữ liệu ERA-Interim, ERA-20C và CFSR so với dữ liệu mưa thực đo VnGP. Những so sánh này cho thấy mức độ phù hợp của kết quả mô phỏng bởi mô hình với các thực đo tương ứng dựa trên giá trị trung bình, độ lệch chuẩn và hệ số tương quan (R). Giá trị của hệ số tương quan (R) là 1 tương ứng với sự kết hợp hoàn hảo của mô phỏng được mô hình hóa với dữ liệu thực đo. Hệ số tương quan lớn hơn 0,6 có thể được coi là chấp nhận được đối với các khoảng thời gian xác nhận ở bước thời gian ngày.

**Bảng 3: Thông số thống kê kết quả mô phỏng bằng mô hình WRF đối với bộ dữ liệu ERA-Interim giai đoạn 1986- 1995**

Thông số thống kê	ERA-INTERIM			
	1 ngày	3 ngày	5 ngày	7 ngày
Mưa trung bình thực đo [mm]	4.63	13.91	23.18	32.46
Mưa trung bình mô phỏng [mm]	4.60	13.80	23.01	32.20
Độ lệch chuẩn của mưa thực đo [mm]	5.88	15.29	23.62	31.69
Độ lệch chuẩn của mưa mô phỏng [mm]	6.36	16.49	25.17	33.31
Hệ số tương quan	<b>0.67</b>	<b>0.77</b>	<b>0.82</b>	<b>0.85</b>

**Bảng 4: Thông số thống kê kết quả mô phỏng bằng mô hình WRF đối với bộ dữ liệu ERA-20C giai đoạn 1986- 1995**

Thông số thống kê	ERA-20C			
	1 ngày	3 ngày	5 ngày	7 ngày
Mưa trung bình thực đo [mm]	4.63	13.91	23.18	32.46
Mưa trung bình mô phỏng [mm]	5.27	15.81	26.35	36.89
Độ lệch chuẩn của mưa thực đo [mm]	5.88	15.29	23.62	31.69
Độ lệch chuẩn của mưa mô phỏng [mm]	6.84	17.87	27.41	36.37
Hệ số tương quan	<b>0.63</b>	<b>0.74</b>	<b>0.79</b>	<b>0.82</b>

**Bảng 5: Thông số thống kê kết quả mô phỏng bằng mô hình WRF đối với bộ dữ liệu CFSR giai đoạn 1986- 1995**

Thông số thống kê	CFSR			
	1 ngày	3 ngày	5 ngày	7 ngày
Mưa trung bình thực đo [mm]	4.63	13.91	23.18	32.46
Mưa trung bình mô phỏng [mm]	3.85	11.51	19.95	28.79
Độ lệch chuẩn của mưa thực đo [mm]	5.88	15.29	23.62	31.69
Độ lệch chuẩn của mưa mô phỏng [mm]	6.38	16.17	24.25	31.63
Hệ số tương quan	<b>0.60</b>	<b>0.70</b>	<b>0.75</b>	<b>0.79</b>

### 2.2.3. Hiệu chỉnh và kiểm định phương pháp thống kê với đầu vào được cấp từ đầu ra của mô hình WRF (D3)

Nghiên cứu sử dụng phương pháp hệ số khí hậu để hiệu chỉnh lượng mưa ngày được mô phỏng bởi các mô hình khí hậu [16]. Đầu tiên, lượng mưa ngày (quan trắc và mô phỏng) được sắp xếp theo cặp từ lớn đến nhỏ và hệ số khí hậu được xác định cho tập hợp các ngày mưa lớn vượt phân vị thứ 99.5%. Các ngày mưa còn lại được sắp xếp theo chuỗi thời gian và hệ số khí hậu được thiết lập cho từng tháng dựa vào số liệu trung bình nhiều năm. Các bước thực hiện như sau:

- Bước 1: Chuỗi dữ liệu mưa ngày mô phỏng và quan trắc được sắp xếp theo thứ từ lớn đến bé, sau đó lựa chọn cặp số liệu (mô phỏng và quan trắc) tương ứng trên phân vị thứ 99.5%;
- Bước 2: Tính hệ số khí hậu trung bình cho 54 cặp dữ liệu trên phân vị thứ 99.5%;
- Bước 3: Các cặp mưa có phân vị nhỏ hơn 99.5% sẽ được sắp xếp lại theo thời gian xuất hiện và tính giá trị trung bình tháng cho toàn

chuỗi số liệu. Hệ số khí hậu được tính bằng tỉ số giữa giá trị trung bình tháng quan trắc và mô phỏng. Nếu một trong các giá trị trung bình tháng quan trắc và mô phỏng bằng 0 thì hệ số khí hậu sẽ lấy bằng 0;

- Bước 4: Giả thiết các hệ số khí hậu xác định trong bước 2 và 3 là cố định trong điều kiện BĐKH, các công thức dưới đây được sử dụng để hiệu chỉnh kết quả mô phỏng mưa:

$$\alpha_I = \frac{P_{Obs_I}}{Mod_{Pre_I}} \quad (1)$$

$$\alpha_{m_i} = \frac{P_{Obs_I}}{Mod_{Pre_{m_i}}} \quad (2)$$

$$P_{Fut_I} = \alpha_I \times Mod_{Fut_I} \quad (3)$$

$$P_{Fut_{m_i}} = \alpha_{m_i} \times Mod_{Fut_{m_i}} \quad (4)$$

Trong đó:

I: Thứ tự chuỗi dữ liệu mưa ngày mô phỏng và quan trắc được sắp xếp theo thứ từ lớn đến bé;

i: Thứ tự chuỗi dữ liệu mưa ngày mô phỏng và quan trắc được sắp xếp theo thời gian trong tháng;

m: tháng;

$P_{Obs_I}$ : Mưa quan trắc ứng với thứ tự  $I$  cho các cặp dữ liệu trên phân vị thứ 99.5% trong giai đoạn hiệu chỉnh;

$Mod_{Pre_I}$ : Mưa mô phỏng ứng với thứ tự  $I$  cho các cặp dữ liệu trên phân vị thứ 99.5% trong giai đoạn hiệu chỉnh;

$\alpha_I$ : hệ số khí hậu cho các cặp dữ liệu trên phân vị thứ 99.5% trong giai đoạn hiệu chỉnh;

$P_{Obs_{m_i}}$ : giá trị mưa trung bình tháng quan trắc trong giai đoạn hiệu chỉnh;

$Mod_{Pre_{m_i}}$ : giá trị mưa trung bình tháng mô phỏng trong giai đoạn hiệu chỉnh;

$\alpha_{m_i}$ : hệ số khí hậu cho tháng trong năm trong giai đoạn hiệu chỉnh;

$Mod_{Fut_I}$ : Mưa mô phỏng ứng với thứ tự  $I$  trong giai đoạn kiểm định cho các cặp dữ liệu trên phân vị thứ 99.5%

$P_{Fut_I}$ : Giá trị mưa được hiệu chỉnh trong giai đoạn kiểm định cho các cặp dữ liệu trên phân vị

thứ 99.5%;

$Mod_{Fut_{m_i}}$ : giá trị mưa trung bình tháng mô phỏng cho giai đoạn kiểm định.

$P_{Fut_{m_i}}$ : giá trị mưa trung bình tháng đã hiệu chỉnh cho giai đoạn kiểm định

Với bộ dữ liệu 30 năm từ 1981- 2010, lựa chọn 1981-1995 là thời gian hiệu chỉnh, 1996-2010 là thời gian kiểm định để phân tích. Kết quả hiệu chỉnh và kiểm định lượng mưa đối với từng bộ dữ liệu được trình bày sau đây.

**❖ Kết quả hiệu chỉnh và kiểm định bộ dữ liệu ERA-Interim**

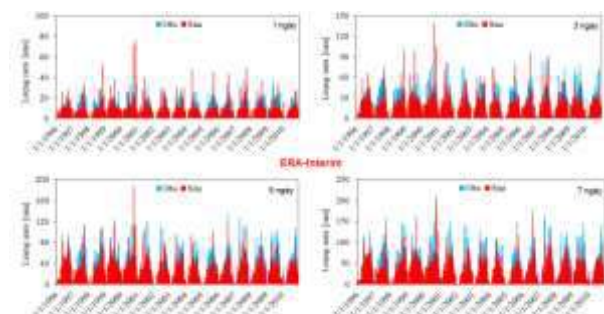
Kết quả hiệu chỉnh và kiểm định lượng mưa đối với bộ dữ liệu ERA-Interim được thể hiện bằng các giá trị thống kê như  $R^2$ , NSE trong bảng 6. Đồng thời giá trị lượng mưa mô phỏng của mô hình WRF sau khi hiệu chỉnh thống kê và dữ liệu thực đo cũng được so sánh sự tương đồng bằng đồ thị như hình 4.

**Bảng 6: Kết quả hiệu chỉnh và kiểm định với bộ dữ liệu ERA-Interim**

Thông số thống kê	Hiệu chỉnh				Kiểm định			
	1 ngày	3 ngày	5 ngày	7 ngày	1 ngày	3 ngày	5 ngày	7 ngày
Mưa trung bình thực đo [mm]	4.56	13.7	22.10	32.02	4.77	14.3	23.8	33.4
Mưa trung bình mô phỏng [mm]	4.01	11.80	19.77	28.10	4.45	13.3	22.1	31.4
Độ lệch chuẩn của mưa thực đo [mm]	5.80	15.20	23.61	31.66	6.05	15.4	23.6	31.5
Độ lệch chuẩn của mưa mô phỏng [mm]	5.01	13.69	20.11	26.21	5.61	14.2	21.9	29.6
$R^2$	0.71	0.81	0.85	0.89	0.73	0.83	0.87	0.90
NSE	0.47	0.651	0.70	0.76	0.50	0.68	0.75	0.78

Hệ số hiệu quả mô hình Nash và Sutcliffe (NSE) cho lượng mưa ngày trong giai đoạn hiệu chỉnh dao động từ 0.47 - 0.76 đồng thời trong giai đoạn kiểm định thì giá trị này tăng là 0.5 - 0.78 theo lượng mưa trung bình lưu vực 1 ngày, 3 ngày, 5 ngày, 7 ngày. Giá trị NSE nằm trong giới hạn mô phỏng từ thỏa mãn đến tốt của tiêu chuẩn thực hiện mô hình cho thấy lượng mưa thực đo và mô phỏng có sự phù hợp cao. Sự tương quan giữa lượng mưa thực đo và mô phỏng được thể hiện bởi hệ số  $R^2$ .  $R^2$  cho lượng mưa ngày trong giai đoạn hiệu chỉnh và kiểm định dao động từ 0.71 -0.90, điều này cho thấy có một mối tương quan tốt giữa giá trị mưa

thực đo và mô phỏng.



**Hình 4: Lượng mưa mô phỏng sau khi sử dụng phương pháp hiệu chỉnh thống kê và thực đo trong giai đoạn kiểm định cho bộ dữ liệu ERA-Interim**

### ❖ Kết quả hiệu chỉnh và kiểm định bộ dữ liệu ERA-20C

Kết quả hiệu chỉnh và kiểm định lượng mưa đối với bộ dữ liệu ERA-20C được thể hiện bằng các

giá trị thống kê như  $R^2$ , NSE trong bảng 7. Đồng thời giá trị lượng mưa mô phỏng của mô hình WRF sau khi hiệu chỉnh thống kê và dữ liệu thực đo cũng được so sánh sự tương đồng bằng đồ thị như hình 5.

**Bảng 7: Kết quả hiệu chỉnh và kiểm định với bộ dữ liệu ERA-20C**

Thông số thống kê	Hiệu chỉnh				Kiểm định			
	1 ngày	3 ngày	5 ngày	7 ngày	1 ngày	3 ngày	5 ngày	7 ngày
Mưa trung bình thực đo [mm]	4.56	13.7	22.10	32.02	4.77	14.3	23.8	33.4
Mưa trung bình mô phỏng [mm]	4.01	12.11	20.11	28.77	4.15	12.5	18.5	28.7
Độ lệch chuẩn của mưa thực đo [mm]	5.80	15.20	23.61	31.66	6.05	15.4	23.6	31.5
Độ lệch chuẩn của mưa mô phỏng [mm]	5.51	12.59	19.87	26.11	4.35	11.5	18.3	25.7
$R^2$	0.65	0.79	0.83	0.86	0.65	0.78	0.79	0.83
NSE	0.40	0.57	0.66	0.71	0.42	0.56	0.61	0.65

Hệ số NSE cho lượng mưa ngày trong giai đoạn hiệu chỉnh dao động từ 0.40 – 0.71 đồng thời trong giai đoạn kiểm định thì giá trị là 0.42 - 0.65 theo lượng mưa trung bình lưu vực 1 ngày, 3 ngày, 5 ngày, 7 ngày. Giá trị NSE thấp hơn so với bộ dữ liệu ERA-Interim tuy nhiên vẫn nằm trong giới hạn mô phỏng thỏa mãn của tiêu chuẩn thực hiện mô hình cho thấy lượng mưa thực đo và mô phỏng có sự phù hợp.  $R^2$  cho lượng mưa ngày trong giai đoạn hiệu chỉnh và kiểm định dao động từ 0.65 - 0.83, điều này cho thấy có một mối

tương quan tốt giữa giá trị mưa thực đo và mô phỏng.

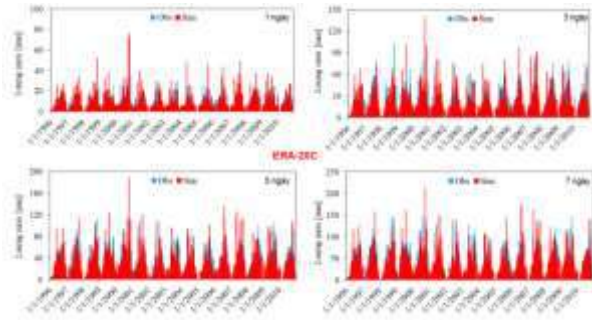
### ❖ Kết quả hiệu chỉnh và kiểm định bộ dữ liệu CFSR

Kết quả hiệu chỉnh và kiểm định lượng mưa đối với bộ dữ liệu CFSR được thể hiện bằng các giá trị thống kê như  $R^2$ , NSE trong bảng 8. Đồng thời giá trị lượng mưa mô phỏng của mô hình WRF sau khi hiệu chỉnh thống kê và dữ liệu thực đo cũng được so sánh sự tương đồng bằng đồ thị như hình 9.

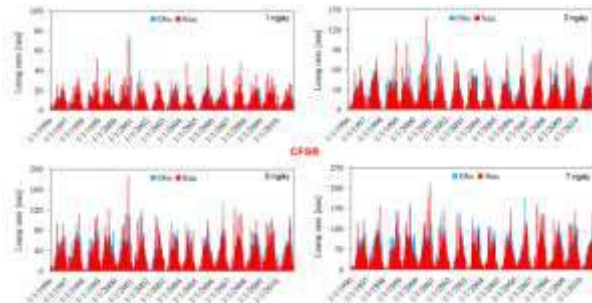
**Bảng 8: Kết quả hiệu chỉnh và kiểm định với bộ dữ liệu CFSR**

Thông số thống kê	Hiệu chỉnh				Kiểm định			
	1 ngày	3 ngày	5 ngày	7 ngày	1 ngày	3 ngày	5 ngày	7 ngày
Mưa trung bình thực đo [mm]	4.56	13.7	22.10	32.02	4.77	14.3	23.8	33.4
Mưa trung bình mô phỏng [mm]	3.81	11.61	17.88	25.42	3.79	11.4	19.1	27.0
Độ lệch chuẩn của mưa thực đo [mm]	5.80	15.20	23.61	31.66	6.05	15.4	23.6	31.5
Độ lệch chuẩn của mưa mô phỏng [mm]	4.88	12.49	19.22	26.71	4.89	11.5	18.8	24.8
$R^2$	0.64	0.76	0.80	0.83	0.68	0.79	0.84	0.88
NSE	0.33	0.52	0.59	0.63	0.42	0.57	0.64	0.70





Hình 5: Lượng mưa mô phỏng sau khi sử dụng phương pháp hiệu chỉnh thống kê và thực đo trong giai đoạn kiểm định cho bộ dữ liệu ERA-20C



Hình 6: Lượng mưa mô phỏng sau khi sử dụng phương pháp hiệu chỉnh thống kê và thực đo trong giai đoạn kiểm định cho bộ dữ liệu CFSR

Hệ số NSE cho lượng mưa ngày trong giai đoạn hiệu chỉnh dao động từ  $\sim 0.63$  đồng thời trong giai đoạn kiểm định thì giá trị là  $\sim 0.67$  theo lượng mưa trung bình lưu vực, 7 ngày. Giá trị NSE thấp hơn so với bộ dữ liệu ERA-Interim tuy nhiên vẫn nằm trong giới hạn mô phỏng thỏa mãn của tiêu chuẩn thực hiện mô hình cho thấy lượng mưa thực đo và mô phỏng có sự phù hợp.  $R^2$  cho lượng mưa ngày trong giai đoạn hiệu chỉnh và kiểm định dao động từ 0.64 - 0.88, điều này cho thấy có một mối tương quan tốt giữa giá trị mưa thực đo và mô phỏng.

Các chỉ số thống kê từ bảng 6 tới 8 cho thấy kết quả hiệu chỉnh và kiểm định sau khi áp dụng phương pháp thống kê cho thấy ERA-Interim vẫn là bộ dữ liệu cho kết quả tốt nhất theo trung bình lưu vực. Đặc biệt với lượng mưa trung bình 7 ngày hệ số tương quan  $R^2 = 0.90$  và  $NSE = 0.78$  nằm trong giới hạn mô phỏng rất tốt của tiêu chuẩn thực hiện mô hình.

### 3. KẾT LUẬN VÀ KIẾN NGHỊ

Bài báo “Mô hình hóa mưa độ phân giải cao kết hợp giữa mô hình động lực khí tượng và phương pháp thống kê: áp dụng cho lưu vực sông Sài Gòn – Đồng Nai” trình bày một số kết quả:

Đã áp dụng thành công phương pháp mô hình WRF và kết hợp với phương pháp hiệu chỉnh thống kê để mô phỏng mưa tới độ phân giải 9km từ ba bộ dữ liệu toàn cầu trong quá khứ: ERA-20C, ERA-Interim, và CFSR cho lưu vực sông Sài Gòn – Đồng Nai. Kết quả thu nhận được là bộ số liệu mưa 30 năm từ năm 1981 – 2010 trên toàn lưu vực sông Sài Gòn – Đồng Nai. Kết quả tính toán mô phỏng cả 03 ba bộ dữ liệu có độ tin cậy tốt và đạt các chỉ tiêu thống kê. Trong đó nếu tính theo trung bình lưu vực thì bộ dữ liệu ERA-Interim (đặc biệt là lượng mưa trung bình 7 ngày theo lưu vực) cho kết quả tốt nhất. Các kết quả của nghiên cứu này đề xuất một phương pháp kết hợp giữa mô hình số trị và thống kê dùng làm tham khảo cho các nhà quản lý, khoa học trong nghiên cứu độ chính xác trong các mô phỏng khí tượng, qua đó sẽ nâng cao độ tin cậy cho các bước tính toán mô phỏng tiếp theo như mô hình thủy văn lưu vực, cũng như lập các kế hoạch thích ứng và ứng phó thiên tai lũ lụt, quản lý tài nguyên nước lưu vực sông.

Kết quả tính toán của nghiên cứu mới chỉ dừng lại ở mức chi tiết hóa đến ô lưới 9km do đó vẫn còn chưa phản ánh hết tác động của điều kiện địa hình đến hình thành chế độ mưa. Chính vì thế hướng phát triển tiếp theo xây dựng mô hình chi tiết hóa ở mức độ chi tiết hơn với bước lưới 6km, 3km. Kết quả của nghiên cứu cũng có thể áp dụng thử nghiệm ở một số lưu vực tương tự.

### LỜI CẢM ƠN

Nghiên cứu này được tài trợ bởi Sở Khoa học và Công nghệ Tp.HCM và được thực hiện bởi Viện Khoa học và Công nghệ tính toán (ICST) thông qua Hợp đồng thực hiện nhiệm vụ khoa học và công nghệ số 16/2020/HĐ-QPTKHCN ngày 22 tháng 04 năm 2020.

## TÀI LIỆU THAM KHẢO

- [1] Krishnamurti T.N., Jha B., Rasch P.J., et al. (1997). A high resolution global reanalysis highlighting the winter monsoon. Part I, reanalysis fields. *Meteorol Atmos Phys*, **64(3-4)**, 123–150.
- [2] Compo G.P., Whitaker J.S., and Sardeshmukh P.D. (2006). Feasibility of a 100-Year Reanalysis Using Only Surface Pressure Data. *Bull Am Meteorol Soc*, **87(2)**, 175–190.
- [3] Fuka D.R., Walter M.T., MacAlister C., et al. (2014). Using the Climate Forecast System Reanalysis as weather input data for watershed models. *Hydrol Process*, **28(22)**, 5613–5623.
- [4] Kjellström E., Bärring L., Nikulin G., et al. (2016). Production and use of regional climate model projections – A Swedish perspective on building climate services. *Clim Serv*, **2-3**, 15–29.
- [5] Wilby R.L. and Wigley T.M.L. (1997). Downscaling general circulation model output: a review of methods and limitations. *Prog Phys Geogr Earth Environ*, **21(4)**, 530–548.
- [6] Prudhomme C., Reynard N., and Crooks S. (2002). Downscaling of global climate models for flood frequency analysis: where are we now?. *Hydrol Process*, **16(6)**, 1137–1150.
- [7] Fowler H.J., Blenkinsop S., and Tebaldi C. (2007). Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling. *Int J Climatol*, **27(12)**, 1547–1578.
- [8] Burlando P. and Rosso R. (2002). Effects of transient climate change on basin hydrology. 1. Precipitation scenarios for the Arno River, central Italy. *Hydrol Process*, **16(6)**, 1151–1175.
- [9] Goyal M.K. and Ojha C.S.P. (2011). Evaluation of linear regression methods as downscaling tools in temperature projections over the Pichola Lake Basin in India. *Hydrol Process*, **25(9)**, 1453–1465.
- [10] Hashmi M.Z., Shamseldin A.Y., and Melville B.W. (2011). Comparison of SDSM and LARS-WG for simulation and downscaling of extreme precipitation events in a watershed. *Stoch Environ Res Risk Assess*, **25(4)**, 475–484.
- [11] Hashmi M.Z., Shamseldin A.Y., and Melville B.W. (2013). Statistically downscaled probabilistic multi-model ensemble projections of precipitation change in a watershed. *Hydrol Process*, **27(7)**, 1021–1032.
- [12] Kure S., Jang S., Ohara N., et al. (2013). WEHY-HCM for Modeling Interactive Atmospheric-Hydrologic Processes at Watershed Scale. II: Model Application to Ungauged and Sparsely Gauged Watersheds. *J Hydrol Eng*, **18(10)**, 1272–1281.
- [13] Gebregiorgis A.S. and Hossain F. (2013). Understanding the Dependence of Satellite Rainfall Uncertainty on Topography and Climate for Hydrologic Model Simulation. *IEEE Trans Geosci Remote Sens*, **51(1)**, 704–718.
- [14] Berg N., Hall A., Sun F., et al. (2015). Twenty-First-Century Precipitation Changes over the Los Angeles Region\*. *J Clim*, **28(2)**, 401–421.

- [15] Nguyen-Xuan T., Ngo-Duc T., Kamimera H., et al. (2016). The Vietnam Gridded Precipitation (VnGP) Dataset: Construction and Validation. *SOLA*, **12**, 291–296.
- [16] Inomata H., Takeuchi K., And Fukami K. (2011). Development Of A Statistical Bias Correction Method For Daily Precipitation Data Of Gcm20. *J Japan Soc Civ Eng Ser B1 (Hydraulic Eng)*, **67(4)**, I\_247-I\_252.
- [17] Trinh, T., Ishida, K., Fischer, I., Jang, S., Darama, Y., Nosacka, J., Brown, K. and Kavvas, M.L., 2016. New methodology to develop future flood frequency under changing climate by means of physically based numerical atmospheric-hydrologic modeling. *Journal of Hydrologic Engineering*, *21(4)*, p.04016001.
- [18] Trinh, T., Ho, C., Do, N., Ercan, A. and Kavvas, M.L., 2020. Development of high-resolution 72 h precipitation and hillslope flood maps over a tropical transboundary region by physically based numerical atmospheric–hydrologic modeling. *Journal of Water and Climate Change*, *11(S1)*, pp.387-406.
- [19] Ho, C., Trinh, T., Nguyen, A., Nguyen, Q., Ercan, A. and Kavvas, M.L., 2019b. Reconstruction and evaluation of changes in hydrologic conditions over a transboundary region by a regional climate model coupled with a physically-based hydrology model: Application to Thao river watershed. *Science of the total environment*, *668*, pp.768-779.
- [20] Jang, S., Kavvas, M.L., Ishida, K., Trinh, T., Ohara, N., Kure, S., Chen, Z.Q., Anderson, M.L., Matanga, G. and Carr, K.J., 2017. A performance evaluation of dynamical downscaling of precipitation over northern California. *Sustainability*, *9(8)*, p.1457.