

**ĐẠI HỌC QUỐC GIA HÀ NỘI
TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN**

Phạm Thanh Hà

**DỰ BÁO HẠNG MÙA VÀ NỘI MÙA NGÀY BẮT ĐẦU
MÙA MƯA Ở VIỆT NAM TRÊN CƠ SỞ
SẢN PHẨM MÔ HÌNH SỐ**

LUẬN ÁN TIẾN SĨ KHÍ TƯỢNG VÀ KHÍ HẬU HỌC

Hà Nội - 2022

**ĐẠI HỌC QUỐC GIA HÀ NỘI
TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN**

Phạm Thanh Hà

**DỰ BÁO HẠN MÙA VÀ NỘI MÙA NGÀY BẮT ĐẦU
MÙA MƯA Ở VIỆT NAM TRÊN CƠ SỞ
SẢN PHẨM MÔ HÌNH SỐ**

**Chuyên ngành: Khí tượng học
Mã số: 9440222.01**

LUẬN ÁN TIẾN SĨ KHÍ TƯỢNG VÀ KHÍ HẬU HỌC

**NGƯỜI HƯỚNG DẪN KHOA HỌC:
GS.TS. Phan Văn Tân**

Hà Nội - 2022

LỜI CAM ĐOAN

Tôi xin cam đoan đây là công trình nghiên cứu của riêng tôi và được sự hướng dẫn khoa học của GS.TS. Phan Văn Tân.

Các số liệu và kết quả được trình bày trong luận án là hoàn toàn trung thực và chưa từng được ai công bố trong bất kỳ công trình nghiên cứu nào khác.

Tác giả

Phạm Thanh Hà

LỜI CẢM ƠN

Trước tiên, tôi xin bày tỏ sự kính trọng và biết ơn sâu sắc nhất đến người thầy đã luôn tận tâm, tận tình hướng dẫn và định hướng tôi trong suốt hành trình làm luận án và trong công việc đó là GS.TS. Phan Văn Tân. Thầy là “Người đưa đò” thật đặc biệt mà tôi luôn trân quý. Thầy là người luôn khích lệ, tôn trọng và ủng hộ những ý tưởng sáng tạo của tôi, và cũng là người góp phần làm những ý tưởng đó trở nên hoàn thiện hơn.

Xin được gửi lời cảm ơn chân thành nhất tới những người thầy và người bạn đã luôn đồng hành, giúp đỡ, động viên, chia sẻ kiến thức và kinh nghiệm cho tôi trong suốt quá trình thực hiện và hoàn thiện luận án, đó là PGS.TS. Ngô Đức Thành, GS.TS. Andreas H. Fink, TS. Roderick van der Linden và ThS. Phạm Quang Nam.

Xin cảm ơn những người đặc biệt đã luôn theo sát, là nguồn động lực giúp tôi vững tâm và bền chí khi thực hiện luận án, đó là: PGS.TS. Nguyễn Thanh Sơn, PGS.TS. Nguyễn Minh Trường, PGS.TS. Vũ Thanh Hằng, TS. Bùi Minh Tuấn, NCS. Đào Nguyễn Quỳnh Hoa.

Tôi xin chân thành cảm ơn các thầy cô, đồng nghiệp, bạn bè trong và ngoài Khoa Khí tượng - Thủy văn và Hải dương học, Trường Đại học Khoa học Tự nhiên đã luôn giúp đỡ và tạo điều kiện thuận lợi cho tôi trong suốt thời gian thực hiện luận án.

Cuối cùng, luận án này sẽ không thể hoàn thành được nếu thiếu nguồn động viên, giúp đỡ về tinh thần và vật chất từ phía gia đình. Tôi xin bày tỏ lòng biết ơn đến những người thân yêu trong gia đình, đặc biệt là mẹ tôi - người bạn đồng hành cùng tôi trên mọi hành trình.

“When you want something, all the universe conspires in helping you to achieve it” - Paulo Coelho

Tác giả

Phạm Thanh Hà

MỤC LỤC

LỜI CAM ĐOAN

LỜI CẢM ƠN

MỤC LỤC

1 DANH MỤC CÁC KÝ HIỆU VÀ CHỮ VIẾT TẮT

4 DANH MỤC HÌNH

4 DANH MỤC BẢNG

8 MỞ ĐẦU

9 Chương 1. TỔNG QUAN VỀ PHƯƠNG PHÁP XÁC ĐỊNH VÀ

DỰ BÁO NGÀY BẮT ĐẦU MÙA MƯA 141.1. Xác định ngày bắt đầu mùa mưa

15

1.2. Biến động theo không gian và thời gian của ngày bắt đầu mùa mưa

tại Việt Nam

21

1.3. Dự báo ngày bắt đầu mùa mưa

24

1.3.1. Dự báo ngày bắt đầu mùa mưa bằng các mô hình thống kê 24

1.3.2. Dự báo ngày bắt đầu mùa mưa dựa trên sản phẩm đầu ra

của mô hình số 26

Chương 2. SỐ LIỆU VÀ PHƯƠNG PHÁP NGHIÊN CỨU 332.1. Số liệu và khu

vực nghiên cứu

34

2.1.1. Số liệu quan trắc 34

2.1.2. Số liệu các chỉ số khí hậu 35

2.1.3. Số liệu dự báo lại của mô hình CFSv2 38

2.1.4. Số liệu dự báo lại của mô hình S2S-ECMWF 39

2.2. Xác định ngày bắt đầu mùa mưa và xu thế biến đổi từ số liệu quan trắc

tại trạm

40

2.2.1. Xác định ngày bắt đầu mùa mưa tại trạm 40

2.2.2. Xác định ngày bắt đầu mùa mưa tại vùng khí hậu 45

2.2.3. Xu thế biến đổi của ngày bắt đầu mùa mưa 46

2.3. Dự báo hạn mùa ngày bắt đầu mùa mưa bằng phương pháp thống kê truyền thống	47
2.4. Dự báo hạn mùa ngày bắt đầu mùa mưa bằng phương pháp thống kê - động lực	48
2.5. Dự báo ngày bắt đầu mùa mưa sử dụng sản phẩm mưa dự báo của mô hình số	49
2.5.1. Dự báo hạn mùa NBĐMM dựa trên sản phẩm mưa của CFSv2.....	49
2.5.2. Dự báo hạn nội mùa ngày bắt đầu mùa mưa dựa trên sản phẩm mưa của S2S-ECMWF	51
2.6. Phương pháp đánh giá	53
2.6.1. Các chỉ số đánh giá	53
2.6.2. Phương pháp kiểm chứng chéo - OLOO	54
Chương 3. ĐẶC ĐIỂM CỦA NGÀY BẮT ĐẦU MÙA MƯA	
TRÊN KHU VỰC VIỆT NAM	
573.1. Tính bất định của ngày bắt đầu mùa mưa xác định cho từng trạm	57
3.2. Xác định chỉ tiêu ngày bắt đầu mùa mưa	59
3.2.1. Tín hiệu vùng của ngày bắt đầu mùa mưa.....	59
3.2.2. Chỉ tiêu ngày bắt đầu mùa mưa dựa trên tín hiệu vùng và trạm.....	60
3.3. Phân bố không gian và xu thế biến đổi của ngày bắt đầu mùa mưa	66
3.4. Mối quan hệ giữa ngày bắt đầu mùa mưa và ENSO	73
Chương 4. DỰ BÁO HẠN MÙA VÀ NỘI MÙA NGÀY BẮT ĐẦU MÙA MƯA TRÊN KHU VỰC VIỆT NAM	
794.1. Dự báo hạn mùa ngày bắt đầu mùa mưa bằng phương pháp thống kê truyền thống sử dụng các chỉ số khí hậu	79
4.2. Dự báo hạn mùa ngày bắt đầu mùa mưa bằng phương pháp thống kê dựa trên sản phẩm của mô hình CFSv2	84
4.2.1. Lựa chọn nhân tố dự báo.....	84
4.2.2. Kết quả dự báo và đánh giá.....	88

4.3. Dự báo hạn mùa ngày bắt đầu mùa mưa từ sản phẩm mưa của CFSv2	94
4.4. Dự báo nội mùa Ngày bắt đầu mùa mưa từ sản phẩm mưa của S2S-ECMWF	98
4.4.1. Tính bất định của ngày bắt đầu mùa mưa dự báo nội mùa từ sản phẩm mưa của S2S-ECMWF	98
4.4.2. Khả năng dự báo hạn nội mùa (7-40 ngày) ngày bắt đầu mùa mưa dựa trên sản phẩm mưa của S2S-ECMWF	100
4.4.3. Sự phụ thuộc của ngày bắt đầu mùa mưa dự báo theo hạn dự báo của S2S-ECMWF	103
KẾT LUẬN	110
NGHỊ	114
LIÊN QUAN ĐẾN LUẬN ÁN	115
KHẢO	116
ĐANH MỤC CÔNG TRÌNH KHOA HỌC CỦA TÁC GIẢ	KIẾN
TÀI LIỆU THAM	PHỤ LỤC

DANH MỤC CÁC KÝ HIỆU VÀ CHỮ VIẾT TẮT

Kí hiệu	Nội dung
CFS	Climate Forecast System (Hệ thống dự báo khí hậu)
cs	cộng sự
ECMWF	European Centre for Medium-Range Weather Forecasts (Trung tâm dự báo thời tiết hạn vừa Châu Âu)
El Niño	Hiện tượng El Niño
ENSO	El Niño Sounthern Oscillation (Dao động nam El Niño)
EOF	Empirical Orthogonal Function (Hàm trực giao nghiệm)
$\bar{F}$	Giá trị trung bình dự báo
GM	Gió mùa
GMMH	Gió mùa mùa hè
La Niña	Hiện tượng La Niña
MAE	Mean Absolute Error (Sai số trung bình tuyệt đối)
mb	millibar
ME	Sai số trung bình (Mean Error)
NBĐMM	Ngày bắt đầu mùa mưa
NBĐGM	Ngày bắt đầu gió mùa
$\bar{Q}$	Giá trị trung bình quan trắc
ONI	Oceanic Nino Index (Chỉ số Nino đại dương)
OLR	Outgoing Longwave Radiation (Bức xạ sóng dài đi ra)
OLOO	One leave one out (phương pháp kiểm chứng chéo)
PCA	Principal Component Analysis (Phân tích thành phần chính)
MPCA	Multi variable Principal Component Analysis (Phân tích thành phần chính nhiều biến)
PRMSL	Mean Sea Level Pressures (Khí áp mực biển trung bình)
QT	Quan trắc
RMSE	Root Mean Square Error (Sai số quân phương)
SOI	Southern Oscillation Index (Chỉ số dao động Nam)
SS _{MAE}	Điểm kỹ năng (skill score) của MAE
SST	Sea Surface Temperatures (Nhiệt độ mặt nước biển)
U850	850-hPa Zonal Winds (Gió vĩ hướng mực 850hPa)

DANH MỤC HÌNH

- Hình 2. 1:** Tổng lượng mưa trung bình tại 131 trạm quan trắc trong giai đoạn 1979-2019. Đường màu đỏ thể hiện ranh giới phân chia giữa các vùng khí hậu
- 34**Hình 2. 2:** Định dạng mẫu của số liệu đầu vào cho phương pháp MPCA
- 43**Hình 2.3:** Sơ đồ mô tả quy trình dự báo NBĐMM hạn mùa (dưới 6 tháng) đối với sản phẩm mưa của CFSv2 cho một vùng khí hậu
- 50**Hình 2.4:** Sơ đồ mô tả quy trình dự báo NBĐMM nội mùa (7-40 ngày) đối với sản phẩm mưa của S2S-ECMWF cho một vùng khí hậu
- 51**Hình 2.5:** Sơ đồ mô tả phương pháp kiểm chứng chéo -OLOO
- 54
- Hình 3.1:** Phân bố không gian giá trị của chỉ số SI tại 131 trạm quan trắc trên khu vực Việt Nam. Đường màu đen thể hiện ranh giới giữa các vùng khí hậu
- 58
- Hình 3.2:** (a-c) Phân bố theo không gian các giá trị tối ưu của P, N, và D để xây dựng bộ chỉ tiêu xác định NBĐMM quy mô địa phương (tại trạm) trên khu vực Việt Nam; (d) hệ số tương quan giữa tín hiệu quy mô khu vực của NBĐMM (EOF1) với NBĐMM quy mô địa phương xác định dựa trên bộ chỉ tiêu tối ưu được lựa chọn. Đường màu đen thể hiện ranh giới của 3 khu vực R123, R45 và R67. Các hình tam giác (vòng tròn) trong hình (a-c) chỉ ra các trạm mà việc thay đổi các ngưỡng giá trị của điều kiện có ảnh hưởng có (không có) ý nghĩa thống kê, dựa trên kiểm nghiệm f với mức độ tin cậy 95%. Các hình tam giác (vòng tròn) trong hình (d) thể hiện mối quan hệ tương quan có (không có) ý nghĩa thống kê, dựa trên kiểm nghiệm t với mức độ tin cậy 95%.
- 63**Hình 3.3:** (Trái) Đồ thị hình hộp của NBĐMM xác định bằng bộ chỉ tiêu mới tại các trạm trên khu vực R123 (R1, R2, R3). Dấu (+) thể hiện các giá trị lớn hơn (nhỏ hơn) $1.5 * IQR$. (Phải) Lượng mưa trung bình hàng tháng với các chấm đỏ thể hiện tháng tương ứng của NBĐMM được xác định bằng chỉ tiêu mới
- 68**Hình 3.4:** Tương tự hình 3.3 nhưng trên khu vực R45 (R4 và R5)
- 71**Hình 3.5:** Tương tự Hình 3.3 nhưng trên khu vực R67 (R6 và R7)
- 71**Hình 3.6:** Tổng lượng mưa tháng trung bình tại 7 vùng khí hậu trong giai đoạn 1979-2019.

Các chấm đỏ thể hiện tháng tương ứng của NBĐMM được xác định bằng chỉ tiêu mới

Hình 3.7: (a) Giá trị trung vị của NBĐMM; (b) xu thế thay đổi của NBĐMM trong giai đoạn 1979-2019 tại 131 trạm trên khu vực Việt Nam (ngày/thập kỷ). Hình tam giác (tròn) thể hiện trạm có xu thế thay đổi thoả mãn (không thoả mãn) mức độ tin cậy 95% của kiểm nghiệm Mann-Kendall. Đường viền màu đen thể hiện đường biên của các khu vực R123, R45 và R67

Hình 3.8: Đồ thị tụ điểm giá trị chuẩn sai NBĐMM và giá trị trung bình chỉ số ONI các tháng 12-1-2 trong giai đoạn 1979-2019 tại 7 vùng. Dấu * thể hiện giá trị hệ số tương quan thoả mãn mức độ ý nghĩa 95% (t-test)

Hình 3.9: Đồ thị hình cột thể hiện xác suất xuất hiện các pha (sớm/muộn) của NBĐMM trong những năm El Niño và La Niña tại 7 vùng khí hậu

Hình 4.1: So sánh biến trình giữa các năm của NBĐMM trong giai đoạn 1979-2019 giữa quan trắc (QT) với dự báo phụ thuộc (DBPT) và dự báo độc lập (DBĐL).

Phần màu ghi xám thể hiện miền giá trị $QT \pm$ phương sai

Hình 4.2: Các khu vực được lựa chọn làm nhân tố dự báo NBĐMM tại các hạn dự báo 1-2-3-6 tháng cho khu vực R3 bằng phương pháp hồi quy từng bước (dựa trên số liệu CFSv2 giai đoạn 1983-2011). Màu sắc của các khu vực tương ứng với các biến PRMSL (xanh), STT (đỏ), U850 (tím)

Hình 4.3: Tương tự Hình 4.2 nhưng cho khu vực R6

Hình 4.4: Tương tự Hình 4.2 nhưng cho khu vực R7

Hình 4.5: Các chỉ số đánh giá chất lượng dự báo NBĐMM bằng phương pháp thống kê động lực trên cơ sở sản phẩm CFSv2 với các hạn dự báo 1-2-3-6 tháng (trên tập số liệu phụ thuộc trong giai đoạn 1983-2011), bao gồm: a) CORR, b) ME, c) MAE, d) SS_{MAE}

Hình 4.6: Các chỉ số đánh giá chất lượng dự báo NBĐMM bằng phương pháp thống kê động lực trên cơ sở sản phẩm CFSv2 với các hạn dự báo 1-2-3-6 tháng (trên tập số liệu độc lập trong giai đoạn 2012-2019), bao gồm: a) ME, b) MAE, c) SS_{MAE}

Hình 4.7: Các chỉ số đánh giá chất lượng dự báo NBĐMM bằng phương pháp thống kê động lực trên cơ sở sản phẩm CFSv2 với các hạn dự báo 1-2-3-6 tháng (trên tập số liệu độc lập trong giai đoạn 1983-2011 - phương pháp kiểm chứng chéo OLOO), bao gồm: a) CORR, b) ME, c) MAE, d)

SS_{MAE} 93**Hình 4.8:** Các chỉ số đánh giá chất lượng dự báo NBĐMM khi sử dụng trực tiếp sản phẩm mưa của CFSv2, bao gồm CORR (hàng 1, màu nền xanh ứng với khoảng giá trị không thoả mãn mức độ ý nghĩa thống kê 95%), ME (hàng 2, màu nền xanh ứng với khoảng giá trị âm), MAE (hàng 3), SS_{MAE} (hàng 4, màu nền xanh ứng với khoảng giá trị âm) tại hạn dự báo 1-2-3-4-5-6 tháng tại 7 vùng khí hậu, với các chỉ tiêu xác định NBĐMM khác nhau 96**Hình 4.9:** Chỉ số SIM (ngày) cho các hạn dự báo khác nhau tại 5 vùng khí hậu khảo sát (R1, R2, R3, R6 và R7) dựa trên các chỉ tiêu xác định NBĐMM khác nhau 99**Hình 4.10:** Chỉ số AV cho các hạn dự báo khác nhau (7-40 ngày) tại 5 vùng khí hậu khảo sát (R1, R2, R3, R6 và R7) dựa trên các chỉ tiêu xác định NBĐMM khác nhau 100**Hình 4.11:** Các chỉ số đánh giá chất lượng dự báo NBĐMM khi sử dụng sản phẩm mưa của S2S-ECMWF, bao gồm CORR (hàng 1), ME (hàng 2, đường kẻ ngang màu nâu thể hiện giá trị = 0), MAE (hàng 3), SS_{MAE} (hàng 4, đường kẻ ngang màu nâu thể hiện giá trị = 0) tại hạn dự báo 7 đến 40 ngày tại 5 vùng khí hậu (R1, R2, R3, R6 và R7), với các chỉ tiêu xác định NBĐMM khác nhau 101**Hình 4.12:** NBĐMM quan trắc (QT) và dự báo dựa trên sản phẩm S2S-ECMWF tại các hạn dự báo khác nhau trong giai đoạn 2000-2019 tại các vùng khí hậu. Khu nền màu xám cho biết phạm vi của một độ lệch chuẩn của NBĐMM QT 105

DANH MỤC BẢNG

Bảng 1.1: Thống kê một số chỉ tiêu xác định NBĐMM/NBĐGM	15
Bảng 2.1: Các chỉ số khí hậu được sử dụng trong luận án	36
Bảng 3.1: Phần trăm phương sai được giải thích bởi PC1-4 tại 7 vùng khí hậu	59
Bảng 3.2: Một số chỉ số thống kê thể hiện mức độ tương đồng giữa NBĐMM tại trạm với vùng khí hậu tương ứng	60
Bảng 3.3: Phần trăm (%) các trạm trên từng vùng mà việc điều chỉnh các ngưỡng giá trị của tham số P, N, và D làm thay đổi đáng kể giá trị NBĐMM (dựa trên kiểm nghiệm f với độ tin cậy 95%). Giá trị được gạch chân và in đậm thể hiện tham số có ảnh hưởng lớn nhất (tức là có giá trị phần trăm lớn nhất) và nên được quan tâm đối với từng vùng và toàn bộ khu vực Việt Nam	64
Bảng 3.4: Tần suất xuất hiện của các ngưỡng giá trị tương ứng với các tham số P, N, D cho từng vùng khí hậu và cho toàn khu vực Việt Nam. Giá trị được gạch chân và in đậm thể hiện giá trị có tần suất xuất hiện lớn nhất trong số các giá trị tham chiếu được lựa chọn cho từng ngưỡng P, N, và D tại từng vùng khí hậu và toàn khu vực Việt Nam	66
Bảng 3.5: Giá trị NBĐMM trung bình và độ lệch chuẩn (biến động giữa các năm) tại 7 vùng khí hậu trong giai đoạn 1979-2019	69
Bảng 4.1: Một số chỉ số đánh giá kỹ năng dự báo NBĐMM bằng mô hình thống kê truyền thống tại 7 vùng khí hậu đối với DBPT và DBĐL (bằng phương pháp kiểm nghiệm chéo -OLOO) trong giai đoạn 1979-2019. Dấu * thể hiện giá trị CORR thoả mãn mức ý nghĩa thống kê 95% (kiểm nghiệm t)	83
Bảng 4.2: Thông tin về kinh độ, vĩ độ, trường của các nhân tố dự báo được lựa chọn từng hạn dự báo 1-2-3-6 tháng của R3, R6 và R7 (dựa trên số liệu CFSv2 giai đoạn 1983-2011). Các nhân tố dự báo được sắp xếp theo thứ tự tuyển chọn của phương pháp hồi quy từng bước	89
Bảng 4.3: Chỉ tiêu xác định NBĐMM có kết quả dự báo với SS_{MAE} cao nhất ứng với các hạn dự báo 7-14-21-28-35-40 ngày	104
Bảng 4.4: Một số chỉ số đánh giá kỹ năng dự báo NBĐMM dựa trên	

sản phẩm S2S-ECMWF tại các hạn dự báo khác nhau (7-14-21-28-35-40 ngày) trong giai đoạn 2000-2019 tại các vùng khí hậu 106 **Bảng 4.5:** Thống kê khả năng dự báo NBĐMM tại từng vùng khí hậu với các phương pháp dự báo và hạn dự báo khác nhau. Kỹ năng dự báo được dựa trên chỉ số SS_{MAE} , theo đó dự báo có kỹ năng khi $SS_{MAE} > 0$ và không có kỹ năng khi $SS_{MAE} < 0$ 109

MỞ ĐẦU

Tính cấp thiết của đề tài

Việt Nam là một quốc gia đang phát triển với nền kinh tế phụ thuộc chủ yếu vào sản xuất nông nghiệp. Đối với lĩnh vực nông nghiệp, lượng mưa là một trong những nhân tố quan trọng ảnh hưởng trực tiếp đến năng suất và thu nhập của người nông dân [Lambert, 2014]. Tuy nhiên, nền nông nghiệp ở Việt Nam dường như dễ bị tổn thương bởi những biến đổi bất thường của thời tiết nói chung, mưa nói riêng. Điều này đã được đề cập tới trong nhiều nghiên cứu gần đây [Endo và cs, 2009; Ngô Đức Thành và Phan Văn Tân, 2012; Viện Khí tượng Thủy văn và Môi trường, 2015; Ngo-Duc và cs, 2014; Trinh-Tuan và cs, 2014; Pham-Thanh và cs, 2020].

Sự biến đổi theo thời gian và không gian của các đặc trưng mưa ở Việt Nam khá phức tạp. Mùa mưa ở miền Bắc, miền Nam và Tây Nguyên trùng với mùa gió mùa mùa hè (GMMH) châu Á (tháng 5 đến tháng 10), trong khi ở miền Trung mùa mưa dịch chuyển sang mùa thu - mùa đông (tháng 8 đến tháng 12) [Phan-Van và cs, 2018]. Trong bối cảnh nóng lên toàn cầu, xu hướng thay đổi của các đặc trưng mưa năm tại Việt Nam có sự khác biệt giữa các trạm và giữa các vùng khí hậu. Tổng lượng mưa năm có xu hướng tăng lên ở miền Nam, nhưng lại có xu hướng giảm ở miền Bắc Việt Nam trong những thập kỷ gần đây [Endo và cs, 2009; Phạm-Thanh và cs, 2020]. Trong số các đặc trưng mưa, ngày bắt đầu mùa mưa (NBĐMM) đã nhận được sự quan tâm lớn vì ý nghĩa quan trọng của nó không chỉ đối với nông nghiệp mà còn đối với kinh tế xã hội do tác động của nó đến nhiều mặt của đời sống như hoạt động nông nghiệp, sản xuất điện, quản lý tài nguyên nước, sức khỏe con người, v.v. [Bombardi và cs, 2020]. Hơn nữa, trong những năm gần đây, các nghiên cứu đã chỉ ra rằng có sự thay đổi về ngày bắt đầu gió mùa (NBĐGM), cũng như NBĐMM trên khu vực Việt Nam [Kajikawa và cs, 2012; Ngo-Thanh và cs, 2018]. Do đó, những thông tin

dự báo hạn nội mùa đến mùa về NBĐMM có ý nghĩa thực tiễn rất lớn trong việc góp phần nâng cao khả năng ứng phó và giảm thiểu thiệt hại do những tác động bất lợi gây nên, phục vụ phát triển kinh tế-xã hội của Việt Nam.

Xuất phát từ thực tế trên, luận án đã lựa chọn đề tài ***“Dự báo hạn mùa và nội mùa ngày bắt đầu mùa mưa ở Việt Nam trên cơ sở sản phẩm mô hình số”***. Với hướng nghiên cứu này, luận án kỳ vọng sẽ góp phần nâng cao chất lượng dự báo hạn nội mùa đến mùa NBĐMM ở Việt Nam. Đây sẽ là thông tin có ý nghĩa đối với ngành nông nghiệp cũng như trong nhiều lĩnh vực của đời sống kinh tế, xã hội.

Mục tiêu của luận án

- 1) Xác định được NBĐMM cho các vùng khí hậu Việt Nam từ số liệu quan trắc tại trạm.
- 2) Đánh giá được mối quan hệ giữa NBĐMM tại các vùng khí hậu Việt Nam với một số quá trình khí quyển, đại dương quy mô lớn.
- 3) Đánh giá được khả năng dự báo hạn nội mùa và mùa NBĐMM cho các vùng khí hậu Việt Nam bằng phương pháp thống kê và thống kê động lực.

Đối tượng, phạm vi và nội dung nghiên cứu

Đối tượng nghiên cứu: ngày bắt đầu mùa mưa.

Phạm vi nghiên cứu: 7 vùng khí hậu Việt Nam.

Thời hạn dự báo:

- Hạn nội mùa: 7 - 40 ngày
- Hạn mùa: dưới 6 tháng

Nội dung nghiên cứu:

- 1) Xác định NBĐMM trên khu vực Việt Nam.
- 2) Phân tích mối quan hệ giữa NBĐMM với các quá trình khí quyển - đại dương quy mô lớn.
- 3) Xây dựng phương trình và khảo sát khả năng dự báo NBĐMM với hạn dưới 6 tháng dựa trên thông tin về sự biến đổi của các trường khí quyển - đại dương từ sản phẩm dự báo hạn mùa của CFSv2.
- 4) Khảo sát khả năng dự báo NBĐMM với hạn dưới 6 tháng bằng việc sử dụng trực tiếp sản phẩm mưa của CFSv2.

- 5) Khảo sát khả năng dự báo NBĐMM với hạn 7 - 40 ngày bằng việc sử dụng trực tiếp sản phẩm mưa của S2S-ECMWF.

Luận điểm bảo vệ của luận án

- 1) Do bị chi phối bởi hoạt động của gió mùa và các hệ thống thời tiết ở các mức độ và khía cạnh khác nhau nên NBĐMM ở các vùng khí hậu có sự khác biệt nhau và có sự biến động theo thời gian từ năm này qua năm khác.
- 2) NBĐMM ở các vùng khí hậu khác nhau có mối quan hệ nhất định với các quá trình khí quyển - đại dương quy mô lớn và có thể sử dụng mối quan hệ này để dự báo hạn mùa NBĐMM.
- 3) Có thể dự báo được NBĐMM ở các vùng khí hậu Việt Nam trên cơ sở sản phẩm dự báo mưa cũng như các trường khác của các mô hình số.

Những đóng góp mới của luận án

- Luận án đã đề xuất được phương pháp khách quan xác định NBĐMM cho các trạm khí tượng ở Việt Nam dựa trên mối quan hệ NBĐMM ở quy mô địa phương và khu vực và áp dụng phương pháp này xác định cho các trạm khí tượng và các vùng khí hậu ở Việt Nam.
- Luận án đã xây dựng được các phương trình dự báo NBĐMM cho các vùng khí hậu Việt Nam bằng phương pháp thống kê, thống kê động lực từ các chỉ số khí hậu, sản phẩm của mô hình CFSv2 và mưa của S2S-ECMWF.

Ý nghĩa khoa học và thực tiễn

- Kết quả nhận được của luận án đã góp phần làm rõ mối quan hệ giữa NBĐMM ở Việt Nam với các quá trình tương tác khí quyển - đại dương quy mô lớn cũng như với các quá trình quy mô nhỏ được phản ánh qua các đặc điểm địa phương như vị trí địa lý, địa hình.
- Kết quả của luận án có thể được sử dụng để xây dựng, phát triển các phương trình dự báo nghiệp vụ NBĐMM với các hạn dự báo khác nhau, từ hạn nội mùa đến hạn mùa.

Tóm tắt cấu trúc luận án

Ngoài các mục mở đầu, tài liệu tham khảo và phụ lục,... những nội dung chính của luận án bao gồm:

CHƯƠNG 1: TỔNG QUAN VỀ PHƯƠNG PHÁP XÁC ĐỊNH VÀ DỰ BÁO NGÀY BẮT ĐẦU MÙA MƯA

Chương này đề cập đến các kết quả nghiên cứu về NBĐMM trên thế giới và trong nước liên quan đến: 1) Phương pháp xác định NBĐMM; 2) Đặc điểm của NBĐMM, gió mùa và ngày bắt đầu gió mùa trên khu vực Việt Nam; 3) Dự báo hạn nội mùa - mùa NBĐMM. Từ đó, những vấn đề còn tồn tại trong nghiên cứu về NBĐMM trên khu vực Việt Nam sẽ được nêu ra và giải quyết trong luận án.

CHƯƠNG 2: SỐ LIỆU VÀ PHƯƠNG PHÁP NGHIÊN CỨU

Chương này mô tả chi tiết các nguồn số liệu được sử dụng, phạm vi và giới hạn nghiên cứu của luận án. Phương pháp nghiên cứu, thiết kế thí nghiệm, phương pháp đánh giá chất lượng đối với dự báo hạn nội mùa - mùa NBĐMM trên khu vực Việt Nam cũng sẽ được trình bày trong chương này.

CHƯƠNG 3: ĐẶC ĐIỂM CỦA NGÀY BẮT ĐẦU MÙA MƯA TRÊN KHU VỰC VIỆT NAM

Chương này trình bày những kết quả tính toán và phân tích các đặc điểm của NBĐMM trên toàn lãnh thổ Việt Nam theo số liệu quan trắc tại trạm. Các khía cạnh được đề cập đến bao gồm: chỉ tiêu xác định NBĐMM; tính bất định trong việc xác định NBĐMM; phân bố không gian và xu thế biến đổi của NBĐMM; mối quan hệ giữa sự biến động hàng năm của NBĐMM với ENSO. Các kết quả này sẽ là cơ sở để dự báo NBĐMM trong Chương 4. *Theo đó, NBĐMM xác định từ số liệu quan trắc tại trạm sẽ được coi là NBĐMM quan trắc trong thực tế và sẽ đóng vai trò làm yếu tố dự báo cũng như giá trị tham chiếu tại Chương 4.* Bên cạnh đó, mối quan hệ giữa sự biến động hàng năm của NBĐMM với ENSO sẽ là tiền đề cho việc thử nghiệm và đánh giá khả năng dự báo NBĐMM bằng phương pháp thống kê truyền thống dựa trên các chỉ số khí hậu.

CHƯƠNG 4. DỰ BÁO HẠN NỘI MÙA - MÙA NGÀY BẮT ĐẦU MÙA MƯA TRÊN KHU VỰC VIỆT NAM

Chương này trình bày quá trình xây dựng và đánh giá khả năng dự báo NBĐMM hạn nội mùa và hạn mùa cho các vùng khí hậu Việt Nam bằng các phương pháp khác nhau, bao gồm thống kê truyền thống, thống kê động lực và mô hình số.

Bốn trường hợp sẽ được xem xét là 1) Dự báo hạn mùa (dưới 6 tháng) NBĐMM bằng phương pháp thống kê truyền thống sử dụng các chỉ số khí hậu làm nhân tố dự báo; 2) Dự báo hạn mùa (dưới 6 tháng) NBĐMM bằng phương pháp thống kê động lực dựa trên các trường dự báo của mô hình CFSv2; 3) Dự báo hạn mùa (dưới 6 tháng) NBĐMM sử dụng trực tiếp sản phẩm mưa dự báo của mô hình CFSv2; 4) Dự báo hạn nội mùa (7 - 40 ngày) NBĐMM dựa trên sản phẩm mưa dự báo của ECMWF.

KẾT LUẬN VÀ KIẾN NGHỊ

Trong phần này sẽ trình bày tóm tắt các kết quả chủ yếu và những điểm mới đã đạt được của luận án. Bên cạnh đó, phần kiến nghị sẽ nêu những tồn tại, kiến nghị những vấn đề cần tiếp tục nghiên cứu tiếp sau và việc sử dụng kết quả luận án.

Chương 1. TỔNG QUAN VỀ PHƯƠNG PHÁP XÁC ĐỊNH VÀ DỰ BÁO NGÀY BẮT ĐẦU MÙA MƯA

Thông tin về NBĐMM đóng vai trò rất quan trọng trong đời sống, kinh tế và xã hội. Đối với quốc gia có nền nông nghiệp lúa nước và xuất khẩu gạo hàng đầu thế giới như Việt Nam, sự biến động của NBĐMM có thể ảnh hưởng đến năng suất, sản lượng và chất lượng sản phẩm. Bên cạnh đó, sự biến động của ngày bắt đầu mùa mưa cũng như độ dài mùa mưa có thể ảnh hưởng đến việc quản lý tài nguyên nước, như điều tiết các hồ thủy lợi cũng như các hồ thủy điện. Do đó, việc xác định và dự báo NBĐMM là hết sức cần thiết nhưng cũng tồn tại nhiều thách thức [Bombardi và cs, 2020].

1.1. XÁC ĐỊNH NGÀY BẮT ĐẦU MÙA MƯA

NBĐMM là một khái niệm dùng để chỉ thời điểm trong năm mà từ đó mưa xảy ra thường xuyên hơn, được đặc trưng bởi lượng mưa đủ lớn và có thể kéo dài từng đợt sao cho tổng lượng mưa tháng phải lớn hơn hoặc bằng một ngưỡng nào đó, và phải kéo dài liên tục trong nhiều tháng. Ở Việt Nam, mùa mưa trong một năm nào đó được xem là các tháng liên tục có tổng lượng mưa tháng lớn hơn hoặc bằng 100mm/tháng [Nguyễn Đức Ngữ và Nguyễn Trọng Hiệu, 2004]. Tuy nhiên, NBĐMM không phải là một biến quan trắc, do đó không có chuỗi số liệu lịch sử.

Thông thường, NBĐMM chỉ được xác định thông qua các chỉ tiêu. Mặc dù, có rất nhiều chỉ tiêu khác nhau được xây dựng để xác định đại lượng này, nhưng chúng có thể chia thành hai nhóm chính: nhóm thứ nhất, dựa trên sự thay đổi về đặc trưng mưa; nhóm thứ hai, dựa trên sự thay đổi của hoàn lưu khí quyển. Tuy nhiên, việc xác định NBĐMM dựa trên các chỉ tiêu thuộc cả hai nhóm này vẫn còn tồn tại những hạn chế nhất định. Bảng 1.1 thống kê một số phương pháp xác định NBĐMM đã được sử dụng trong các nghiên cứu trước đây tại các khu vực khác nhau trên thế giới và tại Việt Nam.

Bảng 1.1: Thống kê một số chỉ tiêu xác định NBDMM/NBDGM

Tên nghiên cứu	Biến sử dụng	Quy mô	Khu vực nghiên cứu
Stern và cs (1981)	Mưa	Địa phương (trạm)	Tây Phi
Nguyễn Hương Điền và cs (2000)	Mưa	Địa phương (trạm)	Nam Bộ - Việt Nam
Liebmann và Marengo (2001)	Mưa	Địa phương (điểm lưới)	Amazon
Wang và LinHo (2002)	Mưa	Địa phương (điểm lưới)	Châu Á
Nguyễn Thị Hiền Thuận và cs (2007)	Mưa	Địa phương (trạm)	Nam Bộ - Việt Nam
Dunning và cs (2016)	Mưa	Địa phương (điểm lưới)	Châu Phi
Nguyen-Le và cs (2015)	Mưa	Khu vực	Phía đông Bán đảo Đông Dương
Lê Thị Xuân Lan và cs (2017)	Mưa	Địa phương (trạm)	Nam Bộ - Việt Nam
Ngo-Thanh và cs (2018)	Mưa	Địa phương (trạm)	Tây Nguyên-Việt Nam
Wester và Yang (1992)	Gió vĩ hướng mực 850mb và 200mb	Khu vực	40-110°E; 0 -20°N
Silva và Carvalho (2007)	Độ ẩm tuyệt đối mực 850mb, nhiệt độ và gió	Khu vực	Nam Mỹ
Tong và cs (2009)	Gió vĩ hướng mực 850mb	Khu vực	110-120°E; 5 -20°N
Garcia và Kayano (2009)	OLR	Khu vực	Miền trung Amazon
Nguyen và cs (2014)	Gió và khí áp mực biển	Khu vực	Khu vực gió mùa điển hình: Á-Úc Châu Phi, Châu Mỹ.

Cách tiếp cận thứ nhất (dựa trên các đặc trưng về mưa) được áp dụng phổ biến trên quy mô địa phương (các trạm) tại nhiều khu vực khác nhau trên thế giới [Stern và cs, 1981; Sivakumar, 1988; Omotosho, 1992; Moron và cs, 2009]. Một trong những chỉ tiêu phổ biến thường thấy đối với cách tiếp cận này có thể được khái quát như sau:

NBĐMM là ngày đầu tiên trong năm của một đợt 5 ngày liên tiếp thoả mãn:

- (1) Có ít nhất N ngày có mưa;
- (2) Tổng lượng mưa của cả đợt lớn hơn P mm;
- (3) Không có sự xuất hiện của bất kỳ một đợt 7 ngày không mưa liên tiếp trong D ngày kể từ ngày bắt đầu mùa mưa.

Có thể nhận thấy, việc xác định NBĐMM bằng các chỉ tiêu này phụ thuộc hoàn toàn vào các ngưỡng giá trị của N, P và D. Trong các nghiên cứu trước đây, các ngưỡng giá trị của N, P, D thường là các ngưỡng thực nghiệm và được xác định dựa trên các đặc trưng về mưa tại khu vực nghiên cứu. Chính vì thế, việc áp dụng các ngưỡng giá trị N, P và D cho các khu vực khác cần được điều chỉnh cho hợp lý, nếu không sẽ có thể gây nên những sai số đáng kể đối với NBĐMM. Liebmann và cs (2001) đã đề xuất một cách tiếp cận thay thế cho việc lựa chọn các ngưỡng cố định N, P và D khi xác định ngày bắt đầu hoặc kết thúc mùa mưa bằng việc phân tích biến trình năm chuẩn sai tích lũy của lượng mưa ngày. Theo đó, ngày bắt đầu/kết thúc mùa mưa được coi là ngày đầu tiên đánh dấu thời điểm lượng mưa ngày tăng lên/giảm đi so với giá trị lượng mưa trung bình ngày nhiều năm. Ưu điểm của phương pháp này chính là việc xác định ngày bắt đầu/kết thúc mùa mưa thông qua việc tìm cực tiểu/cực đại của biến trình năm giá trị chuẩn sai tích lũy của lượng mưa ngày. Điều này giúp hạn chế được việc sử dụng các ngưỡng cố định trong các chỉ tiêu. Bởi vậy phương pháp này có thể áp dụng được trên quy mô các trạm quan trắc, quy mô các vùng và trên cả số liệu lưới. Tuy nhiên, phương pháp này đòi hỏi chuỗi số liệu quá khứ phải đủ dài khi xác định giá trị lượng mưa trung bình nhiều năm. Điều này đã làm giảm tính ứng dụng của phương pháp khi áp dụng vào công tác nghiệp vụ. Một hạn chế khác của phương pháp này là trong những năm mà lượng mưa quá dư thừa/thiếu hụt sẽ dẫn tới tình trạng lượng mưa ngày sẽ cao/thấp hơn nhiều so với giá trị trung bình

mưa ngày nhiều năm, làm cho NBĐMM xác định được có xu hướng đến sớm/muộn hơn nhiều so với thực tế [Bombardi và cs, 2017]. Mặc dù vậy, phương pháp xác định ngày bắt đầu/kết thúc mùa mưa được đưa ra bởi Liebmann và cs (2011) vẫn đang được điều chỉnh và sử dụng rộng rãi trong nhiều nghiên cứu, nhất là trong đánh giá khả năng mô phỏng/dự báo ngày bắt đầu/kết thúc mùa mưa bằng các mô hình động lực [Liebmann và cs, 2007; Dunning và cs, 2016; Bombardi và cs, 2017]. Cũng cách tiếp cận tương tự nhưng thay vì sử dụng số liệu mưa ngày trực tiếp, một số nghiên cứu lại tái cấu trúc lại các số liệu này bằng phương pháp phân tích Fourier theo chuỗi thời gian. Phương pháp này đã được áp dụng cho khu vực gió mùa Châu Á bởi Wang và Linho (2002), khu vực Đông Phi bởi Seregina và cs (2019).

Cách tiếp cận thứ hai để xác định ngày bắt đầu hoặc kết thúc mùa mưa là dựa trên hoàn lưu khí quyển hoặc các đặc trưng nhiệt - động lực. Đáng lưu ý là phương pháp này thường được áp dụng cho các khu vực gió mùa điển hình, nơi mà mùa gió mùa gần như trùng với mùa mưa, như khu vực Ấn Độ, Tây Phi, Châu Á, v.v.... Thông qua việc phân tích synop, Davidson và cs (1983) đã chỉ ra rằng giữa NBĐMM khu vực gió mùa Châu Úc với các quá trình hoàn lưu quy mô lớn khu vực vĩ độ trung bình có mối quan hệ chặt chẽ. Đối với khu vực Tây Phi, những thay đổi cả về hướng gió và nhiệt độ bề mặt phản ánh sự thay đổi của độ ẩm trong dòng gió mùa, đã làm tiền đề cho việc xác định NBĐMM [Omotosho, 1992; Omotosho và cs, 2000]. Trên thực tế, các chỉ tiêu xác định NBĐMM hoặc GM thuộc nhóm này tương đối phong phú, được xây dựng bằng sự kết hợp đa dạng các đặc trưng liên quan đến sự thay đổi của hoàn lưu khí quyển. Chẳng hạn, Holland (1986), Tong và cs (2009) đã dựa trên sự thay đổi của hướng gió thịnh hành; Hendon và Liebmann (1990) sử dụng đồng thời trường gió và mưa; Li và Yanai (1996), Mao và cs (2004) dựa trên của sự tương phản gradient nhiệt độ phía trên tầng đối lưu (mức 200-500hPa); Trần Trung Trực và cs (1999) sử dụng mưa và gió vĩ hướng mức 850mb; Nguyễn Đức Ngũ và Nguyễn Thị Hiền Thuận (2006) sử dụng gió vĩ hướng mức 850mb; Nguyễn Đăng Mậu và cs (2016) sử dụng gió vĩ hướng mức 850mb; còn Nguyen và cs (2014b) lại kết hợp gió và áp suất mực biển.

Nhìn chung, cách tiếp cận thứ nhất (sử dụng lượng mưa ngày) thường được áp dụng để xác định NBĐMM tại quy mô địa phương, trong khi đó cách tiếp cận thứ hai (dựa trên sự thay đổi của hoàn lưu khí quyển) thường được áp dụng để xác định NBĐGM ở quy mô không gian lớn hơn (Bảng 1.1). Tuy nhiên, cần lưu ý tới sự khác biệt về kết quả giữa hai cách tiếp cận này, nhất là tại các khu vực mà mùa mưa và mùa gió mùa không trùng nhau. Ví dụ, đối với khu vực Tây Nguyên và Nam Bộ của Việt Nam, Nguyễn Hướng Điền và cs (2000) cho rằng NBĐMM và NBĐGM mùa hè có thể không trùng nhau. Điều đó được thể hiện trong kết quả thu được từ nghiên cứu của Ngo-Thanh và cs (2018), trong đó tác giả đã chỉ ra rằng NBĐMM có thể đến sớm hơn NBĐGM khoảng 2-3 tuần ở khu vực Tây Nguyên và Nam Bộ. Do đó, thông tin về NBĐMM dựa vào các chỉ tiêu liên quan đến sự thay đổi của hoàn lưu (thường được sử dụng với mục đích nắm bắt NBĐGM) cho những khu vực này có thể đưa ra kết quả không chính xác, thậm chí có thể sai lệch lớn so với thực tế.

Trong những năm gần đây, nhiều nghiên cứu về NBĐMM cho khu vực Việt Nam đã sử dụng các chỉ tiêu khác nhau theo cả hai phương pháp tiếp cận trên nhằm mục đích nắm bắt được sự thay đổi của lượng mưa hàng ngày trong giai đoạn chuyển tiếp từ mùa khô sang mùa mưa cho các vùng khác nhau [Nguyễn Hướng Điền và Trần Công Minh, 2000; Nguyễn Thị Hiền Thuận, 2007; Nguyen-Le và cs, 2014; Lê Thị Xuân Lan và cs, 2017; Ngo-Thanh và cs, 2018]. Trong đó, nhiều nghiên cứu tập trung phân tích các đặc điểm của NBĐMM cho khu vực Nam Bộ. Đáng chú ý, trong các nghiên cứu này phương pháp xác định NBĐMM hay cụ thể là các bộ chỉ tiêu được sử dụng trong các nghiên cứu này hoàn toàn khác nhau. Ví dụ, Nguyễn Hướng Điền và cs (2000) cho rằng NBĐMM trên khu vực Nam Bộ là ngày có tổng lượng mưa 5 ngày lớn hơn ngưỡng 25 mm, và là ngày thứ 3 trong 5 ngày đó. Trong khi đó, Nguyễn Thị Hiền Thuận (2007) lại cho rằng NBĐMM ở Nam Bộ là ngày đầu tiên trong năm sau ngày 1/4 thỏa mãn các điều kiện sau: 1) là ngày có lượng mưa lớn hơn 5 mm, 2) 2 đợt 5 ngày liên tiếp theo sau phải có lượng mưa trung bình mỗi đợt $\geq 5 \text{ mm}$, 3) có ít nhất 5 ngày trong tổng hai đợt 5 ngày liên tiếp có lượng mưa $\geq 5 \text{ mm}$. Lê Thị Xuân Lan và cs (2017) lại cho rằng NBĐMM tại khu vực Nam Bộ là ngày đầu tiên trong năm thỏa mãn các điều kiện sau: 1) có lượng mưa $\geq 5 \text{ mm}$, 2) tổng lượng mưa trượt

10 ngày sau đó lớn hơn 50 mm, 3) với ít nhất 5 ngày có mưa và sau thời kỳ này không có chuỗi ngày gián đoạn mưa liên tục quá 5 ngày. Hay Ngô Thị Thanh Hương (2018) trong luận án của mình tác giả lại áp dụng chỉ tiêu xác định NBĐMM trên khu vực Nam Bộ theo nghiên cứu của Zhang và cs (2002). Theo đó, NBĐMM hàng năm là 1) ngày đầu tiên trong 5 ngày liên tiếp có lượng mưa trung bình trượt 5 ngày lớn hơn 5 mm, 2) trong 20 ngày liên tiếp có ít nhất 10 ngày có lượng mưa trung bình trượt 5 ngày lớn hơn 5 mm. Có thể thấy rằng, các chỉ tiêu xác định NBĐMM đã áp dụng cho khu vực Nam Bộ trong các nghiên cứu trước đây về cơ bản đều có những điểm chung là sử dụng các đặc trưng mưa ngày gồm: 1) tổng lượng mưa và số ngày mưa của đợt mưa chứa NBĐMM; 2) số ngày không mưa liên tiếp sau NBĐMM. Sự khác biệt giữa các chỉ tiêu chủ yếu liên quan đến việc lựa chọn các ngưỡng, và việc lựa chọn này thường mang tính chủ quan dựa trên đặc điểm mưa tại khu vực nghiên cứu. Mặc dù vậy, các kết quả thu được từ các bộ chỉ tiêu khác nhau có sự tương đồng khi cho rằng NBĐMM trên khu vực Nam Bộ thường xảy ra vào thời điểm cuối tháng 4 đầu tháng 5. Nếu so sánh giá trị NBĐMM được xác định bằng các chỉ tiêu khác nhau đối với từng năm cụ thể thì cũng có những sự khác biệt nhất định.

Từ đó nhận thấy rằng, việc lựa chọn chỉ tiêu xác định NBĐMM phù hợp đối với từng khu vực là một thách thức lớn do nó phụ thuộc rất nhiều vào đặc điểm và cơ chế mưa tại khu vực nghiên cứu. Đối với toàn lãnh thổ Việt Nam, nơi đặc điểm mưa giữa các trạm và giữa các vùng khí hậu có tính phân hoá phức tạp theo cả không gian và thời gian do ảnh hưởng của các nhân tố hoàn lưu và sự chi phối của các nhân tố địa phương [Nguyen và cs, 2014a; Nguyen-Le và cs, 2014; Nguyen-Le và cs, 2015; Ngo-Thanh và cs, 2018], nên việc xây dựng một bộ chỉ tiêu khách quan và phù hợp để xác định NBĐMM cho từng vùng và từng địa phương (trạm) là hết sức cần thiết. Đây sẽ tiền đề cho các nghiên cứu khác liên quan như xác định và dự báo độ dài của mùa mưa phục vụ công tác điều hành sản xuất nông nghiệp, vận hành các hồ chứa thủy lợi, thủy điện v.v...

1.2. BIẾN ĐỘNG THEO KHÔNG GIAN VÀ THỜI GIAN CỦA NGÀY BẮT ĐẦU MÙA MƯA TẠI VIỆT NAM

Trên thực tế, có sự phân hoá rõ rệt theo không gian và thời gian của NBĐMM và ngày bắt đầu gió mùa trên khu vực Việt Nam. Nếu xem xét theo không gian, biến trình mưa hàng năm có hai dạng chính: ở các vùng khí hậu phía bắc (Tây Bắc, Đông Bắc, Đồng bằng Bắc Bộ), một phần phía bắc của Bắc Trung Bộ, Nam Bộ và Tây Nguyên mùa mưa trùng với mùa GMMH (tháng 5 đến tháng 10), trong khi phần còn lại của Bắc Trung Bộ và Nam Trung Bộ mùa mưa dịch chuyển về các tháng cuối mùa hè và đầu mùa đông (tháng 8 đến tháng 12) [Nguyễn Đức Ngữ và Nguyễn Trọng Hiệu, 2004; Nguyễn Thị Hiền Thuận, 2007; Nguyen-Le và cs, 2014; Lê Thị Xuân Lan và cs, 2017; Ngo-Thanh và cs, 2018]. Tuy nhiên, mùa mưa thường bắt đầu sớm nhất tại khu vực Tây Bắc, vào khoảng 25/4 [Nguyễn Đức Ngữ và Nguyễn Trọng Hiệu, 2004; Nguyen-Le và cs, 2015]. Đối với khu vực bán đảo Đông Dương (8.5-23.5°N, 100-110°E), bằng việc sử dụng phân tích hàm trực giao cho số liệu mưa vệ tinh APHRODITE trong giai đoạn 1958-2007, Nguyen-Le và cs (2015) đã chỉ ra rằng thời điểm bắt đầu mùa mưa mùa hè trung bình là ngày 6 tháng 5, với độ lệch chuẩn là 13 ngày. Trong khi đó, ngày bắt đầu mùa mưa mùa thu có thời gian bắt đầu trung bình là 16 tháng 9 và độ lệch là 12 ngày.

Sự biến động theo thời gian của NBĐMM cũng như ngày bắt đầu gió mùa tại nhiều khu vực trên thế giới trong đó có Việt Nam đã được nhiều nghiên cứu chỉ ra có liên quan tới ENSO. Chẳng hạn, đối với những năm El Niño, mùa mưa có xu thế đến muộn hơn trên khu vực Indonesia [Moron và cs, 2010], Philippines [Moron và cs, 2009], Australia [Hendon và Liebmann, 1990], ... Trong khi đó, đối với khu vực Tanzania và khu vực Horn của Châu Phi, hiện tượng El Niño lại có liên hệ chặt chẽ với sự đến sớm của mùa mưa. Đối với khu vực Biển Đông, Lau và cs (1997) đã tìm ra sự xuất hiện muộn hơn (sớm hơn) của GMMH có mối liên hệ với sự nóng lên (lạnh đi) ở Thái Bình Dương và Ấn Độ Dương. Zhou và cs (2007) đã khảo sát mối liên hệ giữa ngày bắt đầu gió mùa Đông Nam Á (hay còn gọi là gió mùa Nam Hải, tức gió mùa Biển Đông) và ENSO khi sử dụng số liệu tái phân tích NCEP (Trung tâm dự báo môi trường Hoa Kỳ) và ECMWF (Trung tâm dự báo hạn vừa Châu Âu). Theo tác giả,

NBĐGM là thời điểm mà gió vĩ hướng mực 850mb trên khu vực Biển Đông chuyển từ gió đông sang gió tây kéo dài liên tục 2 pentad. Kết quả cho thấy trong những năm thuộc pha nóng (lạnh) hoặc năm tiếp theo sự kiện ENSO gió mùa có xu hướng bắt đầu muộn hơn (sớm hơn) với cường độ yếu hơn (mạnh hơn). Các kết quả này tương đồng với nghiên cứu Chu Thị Thu Hường và Trần Đình Linh (2019) khi xem xét hoạt động của gió mùa trên lãnh thổ Việt Nam, theo đó, trong những năm El Niño, GMMH thường bắt đầu muộn và kết thúc sớm hơn trong những năm La Niña. Cụ thể, trong các năm La Niña như 1996, 2000, 2001, 2008, 2012, GMMH bắt đầu trước ngày 10/5. Ngược lại, trong các năm El Niño 1987, 1991, 1993, 2014, 2015, GMMH bắt đầu muộn khoảng đầu đến giữa tháng 6 trên toàn khu vực Việt Nam. Ngoài ra, Nguyễn Thị Hiền Thuận và Chiêu Kim Quỳnh (2007) cũng đã chỉ ra rằng NBĐMM ở Nam Bộ sẽ đến muộn hơn trong những năm El Niño và sớm hơn trong những năm La Niña. Tương tự như khu vực Nam Bộ, xu thế biến động giữa các năm của NBĐMM khu vực Tây Nguyên cũng có mối quan hệ với ENSO [Phạm Thị Châm và cs, 2017]. Bên cạnh đó, khi nghiên cứu mối quan hệ giữa nhiệt độ bề mặt biển (SST) trên khu vực nhiệt đới Thái Bình Dương (28N-28S; 120E-85W) và Ấn Độ Dương (28N-28S; 30E-105E) với lượng mưa tháng ở Tây Nguyên, tác giả Nguyen và cs (2007) cũng đã chỉ ra sự thay đổi của SST có ảnh hưởng rõ rệt đến ngày bắt đầu và kết thúc GMMH.

Thực tế, bên cạnh việc nghiên cứu các đặc trưng mưa và mối quan hệ của nó với hoàn lưu khí quyển-đại dương, xu thế biến đổi, đặc biệt về tổng lượng mưa hàng năm cũng đã được tiến hành khảo sát do tầm quan trọng của nó đối với đời sống kinh tế-xã hội. Các nghiên cứu này đa phần dựa trên số liệu quan trắc tại các trạm bề mặt với phạm vi nghiên cứu đa dạng từ toàn cầu cho tới khu vực. Đáng chú ý, các kết quả thu được đều cho thấy xu thế thay đổi của tổng lượng mưa năm có sự khác biệt, thậm chí trái ngược nhau giữa các khu vực khác nhau trên thế giới [Alexander và cs, 2006]. Đối với khu vực Việt Nam, trong những năm gần đây nhiều nghiên cứu đã chỉ ra xu thế thay đổi không chỉ của tổng lượng mưa mà còn của các đặc trưng mưa khác như các chỉ số mưa cực đoan, hạn hán,... [Endo và cs, 2009; Ngo-Duc, 2014; Trinh-Tuan và cs, 2019; Pham-Thanh và cs, 2020 etc.]. Tuy nhiên, sự phân bố theo không gian về xu thế thay đổi của các đặc trưng này lại có sự khác biệt rõ rệt giữa các trạm và

giữa các vùng. Ví dụ, trong những thập kỷ gần đây tổng lượng mưa có xu thế tăng ở các trạm phía Nam và giảm ở các trạm phía Bắc của Việt Nam [Endo và cs, 2009; Ngô Đức Thành và Phan Văn Tân, 2012; Pham-Thanh và cs, 2020].

Mối quan hệ giữa tổng lượng mưa năm và độ dài mùa mưa cũng đã được khảo sát [Boyard-Micheau và cs, 2013; Ngo-Thanh và cs, 2018]. Sự thay đổi của tổng lượng mưa hàng năm ở Việt Nam có thể liên quan đến sự thay đổi của độ dài mùa mưa, mà cụ thể là sự thay đổi của ngày bắt đầu và kết thúc mùa mưa. Đặc biệt, sự thay đổi của NBDMM và NBDGM đã được quan sát thấy trên khu vực Việt Nam trong nhiều nghiên cứu gần đây. Cụ thể, Kajikawa và cs (2012) đã chỉ ra NBDGM trên khu vực vịnh Bengal, Bán đảo Đông Dương, Tây Thái Bình Dương có xu hướng xuất hiện sớm hơn từ 10-15 ngày khi so sánh hai giai đoạn 1979-1993 và 1994-2008. Trần Quang Đức (2010) cho rằng có sự thay đổi của một số đặc trưng hoạt động của GMMH trên khu vực Việt Nam trong thời kỳ 1950-2010, trong đó ngày bắt đầu GMMH có xu thế tới sớm hơn, độ dài mùa gió mùa có thể dài hơn, nhưng cường độ hoạt động của GMMH lại có xu thế yếu đi. Dựa trên bộ số liệu quan trắc trong giai đoạn 1981-2014, trong luận án Ngô Thanh Hương (2018) cũng chỉ ra xu thế đến sớm hơn của ngày bắt đầu GMMH ở khu vực Tây Nguyên và Nam Bộ tương ứng là 2.5 ngày/thập kỷ và 3.3 ngày/thập kỷ. Tuy nhiên, xu thế thay đổi của ngày kết thúc GMMH tại 2 khu vực trên là nhỏ hơn và không rõ rệt. Với bộ số liệu cập nhật hơn cho thời kỳ 1981-2015, Chu Thị Thu Hường và Trần Đình Linh (2019) cũng cho kết quả tương tự với các nghiên cứu trước đây trong đó mùa GMMH có xu thế đến sớm từ 1-2 ngày/thập kỷ, và thời gian hoạt động của GMMH nhìn chung có xu thế tăng lên trên khu vực Việt Nam. Các kết quả trên cho thấy sự cần thiết của công tác dự báo ngày bắt đầu/kết thúc mùa mưa trên toàn khu vực Việt Nam, nhất là trong bối cảnh các đặc trưng liên quan đến tổng lượng mưa và hoạt động của GMMH đã ghi nhận những sự biến đổi phức tạp theo thời gian và không gian trong những năm gần đây.

1.3. DỰ BÁO NGÀY BẮT ĐẦU MÙA MƯA

Thông tin dự báo về NBĐMM đóng vai trò quan trọng trong nhiều lĩnh vực của đời sống như hoạch định kế hoạch sản xuất, lập nông lịch, quản lý tài nguyên nước, điều tiết hồ chứa thủy lợi, thủy điện, v.v. Do đó, để đáp ứng yêu cầu của xã hội, nhiều nghiên cứu đã được tiến hành với mục đích dự báo NBĐMM bằng nhiều cách tiếp cận khác nhau. Có thể chia các nghiên cứu này thành hai nhóm tiếp cận chính là tiếp cận theo phương pháp thống kê và tiếp cận theo phương pháp động lực (tức là sử dụng trực tiếp sản phẩm đầu ra của các mô hình số).

1.3.1. Dự báo ngày bắt đầu mùa mưa bằng các mô hình thống kê

Việc dự báo ngày bắt đầu/kết thúc mùa mưa bằng phương pháp thống kê được áp dụng rộng rãi trên các khu vực khác nhau trên thế giới [Omotosho, 1992, 2000; Lala và cs, 2020]. Đáng chú ý, trong các mô hình thống kê dự báo ngày bắt đầu hoặc kết thúc mùa mưa (là yếu tố dự báo), việc lựa chọn nhân tố dự báo đóng vai trò cực kỳ quan trọng và thường được thực hiện dựa trên mối quan hệ giữa yếu tố dự báo với nhân tố dự báo. Trong các nghiên cứu trước đây, mối quan hệ giữa ngày bắt đầu/kết thúc mùa mưa với sự thay đổi của các trường hoàn lưu quy mô lớn thông qua các đại lượng như nhiệt độ bề mặt biển, gió, áp suất mực biển, bức xạ sóng dài v.v.... thường được xem xét trong việc xây dựng các mô hình thống kê dự báo.

Kết quả về khả năng dự báo NBĐMM và NBĐGM dựa trên các mô hình thống kê với các nhân tố dự báo liên quan đến sự thay đổi của hoàn lưu khí quyển-đại dương đã đạt được những thành công khi nắm bắt được sự biến đổi hàng năm của NBĐMM thực tế, và chứng minh được tính khả thi của phương pháp tại nhiều khu vực khác nhau trên thế giới. Ví dụ, Odekunle và cs (2005) đã sử dụng các biến như nhiệt độ mặt biển khu vực Đại Tây Dương, nhiệt độ bề mặt tại các trạm, và sự tương phản nhiệt độ đất-biển giữa các trạm và khu vực nhiệt đới Đại Tây Dương đã được xem xét để đưa vào xây dựng phương trình dự báo NBĐMM cho khu vực Nigeria. Theo tác giả, các kết quả thu được dựa trên mô hình dự báo thống kê rất khả quan. Theo đó, 75% số trường hợp thử nghiệm cho hệ số tương quan giữa NBĐMM dự báo và thực tế trên mức 0.5, thậm chí trong nhiều trường hợp lên tới 0.9. Như vậy, thông tin dự báo NBĐMM bằng phương pháp thống kê có thể nắm bắt tốt những biến động

hàng năm của NBĐMM thực tế trên khu vực Nigeria. Điều này chứng minh rằng việc lựa chọn các nhân tố có liên quan đến sự tăng mưa hay sự phát triển của đối lưu làm các nhân tố để xây dựng các phương trình dự báo NBĐMM là khả thi. Ngoài ra, khả năng dự báo hạn mùa NBĐMM bằng phương pháp thống kê trên khu vực Indonesia cũng đã được Moron và cs (2009) chỉ ra. Trong nghiên cứu này, tác giả đã chứng minh được mối liên hệ chặt chẽ giữa NBĐMM với các pha của ENSO thông qua việc xem xét giá trị nhiệt độ mặt biển tháng 7. Qua đó tác giả đã sử dụng nhiệt độ mặt biển khu vực nhiệt đới Thái Bình Dương và Ấn Độ Dương (20°S - 20°N , 80° - 280°E) làm nhân tố dự báo NBĐMM cho Indonesia. Kết quả dự báo lại NBĐMM giai đoạn 1979-2004 cho thấy hệ số tương quan giữa NBĐMM dự báo và thực tế tại trạm khá cao (~ 0.8). Phương pháp dự báo NBĐMM bằng các mô hình thống kê cũng được áp dụng rộng rãi ở khu vực gió mùa Tây Phi, trong đó có thể kể đến nghiên cứu của Omotosho và cs (2000) hay Fontaine và cs (2008). Với các nhân tố dự báo thể hiện cho giá trị gradient của OLR giữa khu vực vịnh Guinea và bán đảo Châu Phi, và sự dịch chuyển theo hướng bắc của gió mùa Tây Phi, Fontaine và cs (2008) cũng đã cho thấy khả năng dự báo NBĐGM ở khu vực này bằng phương pháp hồi quy đa biến. Ngoài ra, một cách tiếp cận thống kê khác đã được Stolbova và cs (2016) đưa ra với mục đích dự báo ngày bắt đầu và kết thúc gió mùa Ấn Độ với hạn trên 30 ngày. Theo đó, ngày bắt đầu và kết thúc GMMH Ấn Độ được xác định dựa trên việc phân tích sự thay đổi theo thời gian của nhiệt độ và độ ẩm tương đối trên khu vực phía đông Ghats và phía bắc Pakistan. Kết quả thu được cho thấy có đến 73% số năm thử nghiệm cho kết quả sai số dưới 7 ngày so với thực tế.

Đối với khu vực Việt Nam, Pham và cs (2010), Phạm Thị Châm và cs (2017) cũng đã cho thấy khả năng dự báo ngày bắt đầu GMMH bằng mô hình thống kê cho khu vực Nam Bộ và Tây Nguyên. Cụ thể, Pham và cs (2010) đã sử dụng phương pháp hồi quy từng bước để lựa chọn và xây dựng phương trình dự báo sử dụng bộ 10 nhân tố tiềm năng, bao gồm: (1) năng lượng tĩnh ẩm khu vực Trung Quốc-Nhật Bản, (2) MSE khu vực Australia, (3) áp suất mực biển trung bình khu vực Ấn Độ Dương, (4) áp suất mực biển khu vực biển Hoa Đông, (5) gió trên biển Ả Rập, (6) gió trên vịnh Bengal, (7) gió trên khu vực phía nam Việt Nam, (8) OLR trên khu vực vịnh Bengal,

(9) OLR trên khu vực Biển Đông, (10) OLR trên khu vực Indonesia. Kết quả cho thấy 4 nhân tố hoàn lưu có mối liên hệ chặt chẽ nhất với ngày bắt đầu gió mùa trên khu vực Nam Bộ gồm: 1) áp suất mực biển khu vực biển Hoa Đông ($20-140^{\circ}\text{E}$, $30-35^{\circ}\text{N}$, tại pentad thứ 18), 2) gió trên khu vực phía nam Việt Nam ($105-115^{\circ}\text{E}$, $10-15^{\circ}\text{N}$, tại pentad thứ 20), 3) OLR trên khu vực vịnh Bengal ($90-100^{\circ}\text{E}$, $10-20^{\circ}\text{N}$, tại pentad thứ 18) và 4) OLR trên khu vực Indonesia ($100-120^{\circ}\text{E}$, $0-10^{\circ}\text{S}$; tại pentad thứ 20). Hệ số tương quan giữa ngày bắt đầu GMMH dự báo và thực tế lên tới 0.75. Đối với khu vực Tây Nguyên, Phạm Thị Châm và cs (2017) đã áp dụng phương pháp phân tích tương quan Canon với các nhân tố dự báo được tuyển chọn là nhiệt độ bề mặt biển, bức xạ sóng dài và gió vĩ hướng trên mực 850mb trung bình ba tháng đầu năm. Kết quả thu được cho thấy mặc dù còn tồn tại sai số dự báo nhưng nhìn chung NBĐMM dự báo đều năm bắt được xu thế xảy ra sớm hay muộn hơn so với trung bình nhiều năm trên hầu hết các trạm trong ba năm gần đây nhất là năm 2015, 2016 và năm 2017.

Như vậy, việc sử dụng phương pháp thống kê để dự báo NBĐMM đã được áp dụng và cho kết quả tốt khi so sánh với thực tế quan trắc tại nhiều khu vực trên thế giới, trong đó có Tây Nguyên và Nam Bộ của Việt Nam. Điều này cho thấy việc thử nghiệm áp dụng phương pháp này đối với việc dự báo NBĐMM dựa trên những đặc trưng thay đổi của trường hoàn lưu cho khu vực Việt Nam là hoàn toàn khả thi.

1.3.2. Dự báo ngày bắt đầu mùa mưa dựa trên sản phẩm đầu ra của mô hình số

Việc sử dụng trực tiếp các sản phẩm đầu ra của mô hình số để mô phỏng, dự báo NBĐMM đã được tiến hành trong nhiều nghiên cứu. Cần lưu ý, NBĐMM hay NBĐGM không phải là sản phẩm trực tiếp từ các mô hình số, do đó cần phải sử dụng các chỉ tiêu để xác định. Chính vì vậy, chất lượng của việc mô phỏng hay dự báo NBĐMM sẽ phụ thuộc rất nhiều vào việc lựa chọn các chỉ tiêu này.

Một số nghiên cứu trước đây đã chỉ ra những hạn chế nhất định của các mô hình động lực trong việc mô phỏng hoạt động của gió mùa, đặc biệt là NBĐGM. Ví dụ, Rajendran và cs (2004) đã sử dụng mô hình MRI-CGCM2 để mô phỏng ngày bắt đầu và kết thúc mùa mưa gió mùa Châu Á. Các kết quả thu được chỉ ra mô hình mô phỏng NBĐGM trên khu vực Đông Á đến sớm hơn rõ rệt so với thực tế. Sự khác biệt này có thể do sự thiếu sót của mô hình trong khả năng mô phỏng những khía cạnh cơ

bản trong hoạt động của gió mùa, như sự đổi chiều của hướng gió thịnh hành trong giai đoạn chuyển mùa bằng các mô hình hoàn lưu chung trong khí quyển (AGCMs). Li và Zhang (2009) đã đánh giá khả năng mô phỏng ngày bắt đầu và kết thúc gió mùa châu Á dựa trên kết quả đầu ra từ 7 mô hình trong khuôn khổ dự án AMIP (Atmospheric Model Intercomparison Project). Các kết quả thu được cho thấy các mô hình này có khả năng mô phỏng ngày kết thúc tốt hơn NBDGM. Cụ thể, trong khi hầu hết các mô hình có thể nắm bắt được sự thay đổi của hướng gió tại thời điểm kết thúc gió mùa, thì chỉ có 3/7 mô hình cho kết quả mô phỏng NBDGM khá sát với thực tế. Do đó, việc mô phỏng và dự báo hoạt động của gió mùa hay mưa gió mùa là một trong những thách thức lớn đối với các mô hình khí hậu [Becker và cs, 2014] vào thời điểm đó.

Tuy nhiên, trong những năm gần đây, nhiều nghiên cứu đã cho thấy các mô hình khí hậu có khả năng mô phỏng lại hoạt động của ngày bắt đầu và kết thúc gió mùa tại các khu vực gió mùa điển hình khác nhau trên thế giới. Đối với khu vực gió mùa châu Á, Cherchi và Nvarra (2003) và Li và Zhang (2009) đã đánh giá khả năng mô phỏng ngày bắt đầu GMMH bằng các mô hình hoàn lưu chung khí quyển. Kết quả thu được cho thấy, kỹ năng mô phỏng của mô hình phụ thuộc rất nhiều vào việc lựa chọn tiêu chí xác định ngày bắt đầu và kết thúc. Cụ thể, cả hai nghiên cứu đều cho thấy việc xác định ngày bắt đầu và kết thúc dựa trên sự thay đổi các đặc trưng liên quan đến hoàn lưu (như gió, phát xạ sóng dài,...) cho kết quả mô phỏng tốt hơn so với việc dựa trên sự thay đổi các đặc trưng của mưa. Điều này có thể liên quan đến khả năng mô phỏng các đặc điểm mang tính chất địa phương như tương tác địa hình dẫn đến hiện tượng mưa mang tính chất cục bộ của các mô hình động lực vẫn còn hạn chế, nhất là đối với những mô hình có độ phân giải thô. Bên cạnh đó, cả 2 nghiên cứu đều chỉ ra hoạt động của gió mùa từ một mô hình đơn lẻ là kém hơn so với từ tổ hợp đa mô hình. Do đó, việc sử dụng phương pháp tổ hợp trong mô phỏng hoạt động của gió mùa nên được quan tâm, kết quả mô phỏng tổ hợp sẽ được cải thiện rõ rệt khi đưa thêm nhiều thành phần. Ví dụ, Cherchi và Nvarra (2003) cho thấy cần ít nhất 16 thành phần của mô hình ECHAM4 (các thành phần này giống nhau ở SST, khác nhau ở điều kiện ban đầu) để nắm bắt được sự biến động của hoạt động GMMH Ấn Độ.

Đối với khu vực gió mùa Nam Mỹ, Bombardi và Carvalho (2009), Jones và Carvalho (2013) đã đánh giá khả năng mô phỏng ngày bắt đầu và kết thúc của gió mùa trên khu vực này bởi mô hình CMIP3 và CMIP5 tương ứng. Những nghiên cứu này đã chỉ ra rằng các mô hình khí hậu trên có thể nắm bắt được thời gian hoạt động của gió mùa Nam Mỹ, cũng như các điều kiện quy mô lớn liên quan với hoàn lưu gió mùa đặc biệt, so sánh một cách tương đối, việc mô phỏng hoạt động của gió mùa Nam Mỹ, đặc biệt là lượng mưa cực đại bằng mô hình CMIP5 có phần tốt hơn so với mô hình CMIP3 [Jones và Carvalho, 2013].

Các nghiên cứu gần đây cho thấy những tín hiệu khả quan trong công tác dự báo hạn mùa NBDGM dựa trên cơ sở của các sản phẩm số, đặc biệt là tại các khu vực gió mùa khác nhau trên thế giới như Tây Phi, Ấn Độ, Nam Mỹ, Bắc Úc,...Theo đó, NBDGM có thể được dự báo với các hạn dự báo từ vài tuần cho tới vài tháng bằng các mô hình động lực [Vellinga và cs, 2013; Alessandri và cs, 2015; Bombardi và cs, 2017; Chevuturi và cs, 2018]. Chẳng hạn, đối với khu vực Tây Phi, Vellinga và cs (2013) đã sử dụng các sản phẩm dự báo nghiệp vụ của Cơ quan Khí tượng Anh (GloSea4) trong giai đoạn 1996-2009 để xem xét khả năng dự báo NBDGM trên khu vực Sahel. Kết quả thu được cho thấy, việc dự báo NBDGM trước 2-3 tháng dựa trên thông tin về nhiệt độ mặt biển (SST) vùng nhiệt đới vào tháng 6 là hoàn toàn khả thi. Đặc biệt chất lượng của các dự báo này là cao hơn so với dự báo khí hậu, giá trị của ROC có thể đạt từ 0.6 đến 0.8.

Đối với khu vực Ấn Độ, Alessandri và cs (2015) đã chỉ ra rằng ngày bắt đầu gió mùa tại đây có thể được dự báo với hạn dự báo trước 1 tháng dựa trên kết quả của mô hình động lực, mà cụ thể ở đây là mô hình dự báo được phát triển bởi Centro Euro-Mediterraneo sui Cambiamenti Climatici (CMCC). Đáng chú ý, ngày bắt đầu GMMH trong nghiên cứu này đã được định nghĩa theo 2 cách khác nhau: (1) dựa trên chỉ tiêu liên quan tới gió mực 850mb, và (2) dựa trên chỉ tiêu liên quan đến vận tải ẩm theo phương thẳng đứng. Kết quả thu được cho thấy, giá trị dự báo NBDGM mùa hè từ mô hình được xác định bằng 2 chỉ tiêu trên đều nắm bắt được sự thay đổi hàng năm của giá trị quan trắc. Tuy nhiên, giá trị tương quan giữa NBDGM mùa hè dự báo và quan trắc là tốt hơn trong trường hợp sử dụng chỉ tiêu liên quan tới gió mực 850mb

(đạt 0.65) khi so sánh với trường hợp còn lại (chỉ đạt 0.52). Bên cạnh đó, các kết quả về dự báo pha sự xuất hiện GMMH cũng cho thấy các kết quả trong trường hợp sử dụng chỉ tiêu liên quan tới gió mực 850mb là tốt hơn, được thể hiện thông qua các giá trị BSS (Brier Skill Score) lớn hơn trong cả pha trên chuẩn và dưới chuẩn. Các kết quả trên đã nhấn mạnh vai trò của việc lựa chọn chỉ tiêu xác định GMMH cũng ảnh hưởng tới chất lượng dự báo của mô hình. Pradhan và cs (2017) cũng đã chỉ ra khả năng dự báo ngày bắt đầu GMMH khu vực Ấn Độ với hạn dự báo 2-3 tuần, bằng việc sử dụng số liệu của Viện Khí tượng Nhiệt đới Ấn Độ. Kết quả thu được cho thấy các mô hình động lực đã nắm bắt được sự thay đổi hàng năm của ngày bắt đầu gió mùa. Nghiên cứu cũng chỉ ra rằng khả năng dự báo ngày bắt đầu gió mùa có thể được cải thiện khi thay đổi độ phân giải của mô hình. Cụ thể, khi tăng độ phân giải, khả năng nắm bắt sự thay đổi trước và sau thời kỳ giao mùa của hoàn lưu khi chuyển mùa thông qua hoạt động của đối lưu, của độ ẩm cũng tăng lên. Bên cạnh đó, Chevuturi và cs (2018) cũng đã chỉ ra mô hình GloSea5-GC2 có khả năng đưa ra các dự báo ngày bắt đầu gió mùa Ấn Độ với độ tin cậy cao trong hạn dự báo một tháng, đặc biệt là trong các dự báo pha (dưới chuẩn-chuẩn-trên chuẩn). Trong nghiên cứu này, các tác giả cũng quan tâm đến sự ảnh hưởng của các chỉ tiêu xác định ngày bắt đầu gió mùa đối với chất lượng dự báo. Điều này được thể hiện thông qua việc so sánh các kết quả dự báo dựa trên 5 chỉ tiêu khác nhau đại diện cho các quy mô khác nhau từ địa phương tới khu vực, liên quan tới: hoàn lưu gió mùa, sự thay đổi về mặt thủy văn, gradient nhiệt độ tầng đối lưu, lượng mưa trung bình toàn vùng, lượng mưa trên từng điểm lưới. Các kết quả dự báo ngày bắt đầu gió mùa khu vực Ấn Độ đều nắm bắt được sự biến động hàng năm khi so sánh với quan trắc trong tất cả 5 trường hợp khảo sát. Tuy nhiên, các chỉ tiêu xác định dựa trên sự thay đổi về mặt động lực và nhiệt động lực quy mô lớn cho kết quả dự báo ngày bắt đầu gió mùa tốt hơn so với các chỉ tiêu động lực quy mô địa phương hay chỉ số thủy văn. Điều này cho thấy mô hình GloSea5-GC2 có khả năng mô phỏng các quá trình quy mô lớn tốt hơn so với các quá trình quy mô nhỏ (địa phương).

Đối với khu vực Việt Nam, việc mô phỏng hoàn lưu quy mô lớn thời kỳ bắt đầu mùa GMMH bằng các mô hình số cũng đã được nghiên cứu. Có thể kể đến,

nguyên cứu của Bùi Minh Tuấn và Nguyễn Minh Trường (2011) đã mô phỏng sự phát triển hoàn lưu quy mô lớn thời kỳ bắt đầu GMMH bằng mô hình RAMS. Các đặc trưng cơ bản của ngày bắt đầu GMMH đã được chỉ ra, theo đó, GMMH xuất hiện sớm nhất ở vịnh Bengal, rồi xuất hiện đồng thời trên bán đảo Đông Dương và Biển Đông. Đối với khu vực Nam Bộ, ngày bắt đầu GMMH xảy ra sau khoảng 3 ngày kể từ thời điểm có sự xuất hiện của xoáy kép Sri Lanka và sự dịch chuyển của dải áp thấp xích đạo. Ngoài ra, việc lựa chọn chỉ tiêu phù hợp đối với việc xác định ngày bắt đầu GMMH cho mô hình số cũng đã được đề cập trong nhiều nghiên cứu. Nguyễn Minh Trường và cs (2012) đã sử dụng các chỉ tiêu khác nhau để xác định ngày bắt đầu GMMH trên khu vực Nam Bộ dựa trên sản phẩm của mô hình RAMS, bao gồm: chỉ số gió tây, chỉ số mưa và chỉ số gradient nhiệt độ mực cao. Dựa trên các kết quả mô phỏng GMMH trong các năm 1998, 2001, 2004, 2010, tác giả cho rằng chỉ số gió tây, được xác định bằng việc trung bình gió vĩ hướng mực 850mb khu vực ($10-15^{\circ}\text{N}$, $100-110^{\circ}\text{E}$) là phù hợp nhất để xác định ngày bắt đầu GMMH trên khu vực Nam Bộ do chỉ số này vừa thể hiện được đặc trưng của hoàn lưu quy mô lớn, lại vừa có quan hệ tốt với lượng mưa. Gần đây, việc dự báo NBĐMM/gió mùa trên khu vực Tây Nguyên và Nam Bộ dựa trên sản phẩm của mô hình CFSv2 là có tiềm năng. Đặc biệt, đối với phương pháp động lực việc sử dụng chỉ tiêu NRM mô tả sự thay đổi của trường áp suất mực biển và gió vĩ hướng mực 850mb cho kết quả dự báo ngày bắt đầu GMMH tốt tại các hạn dự báo 2-3 tháng, sai số dự báo tăng lên khi hạn dự báo tăng [Hoàng Đức Cường và cs, 2018].

Như vậy, đối với phương pháp sử dụng các mô hình động lực, NBĐMM/NBĐGM ở quy mô cục bộ (địa phương) chỉ có thể dự đoán được nếu nó bị chi phối nhiều bởi hoàn lưu quy mô lớn [Bombardi và cs, 2020]. Theo đó, sự tác động của các hoàn lưu quy mô lớn làm thúc đẩy hay kìm hãm sự phát triển của đối lưu trên một khu vực rộng lớn, được coi là nguyên nhân chính của sự đến sớm hoặc muộn của mùa mưa/mùa gió mùa ở quy mô địa phương. Do đó, trong trường hợp này, khả năng dự báo thời điểm dịch chuyển mùa của gió mùa/mùa mưa sẽ phụ thuộc rất nhiều vào khả năng của mô hình trong việc mô tả sự thay đổi của các hoàn lưu quy mô lớn, cũng

như sự tương tác giữa chúng với các nhân tố mang tính chất địa phương. Điều này có thể lý giải phần nào việc cải thiện chất lượng dự báo ngày bắt đầu gió mùa/mùa mưa khi tăng độ phân giải của mô hình như đã đề cập ở trong nghiên cứu của Pradhan và cs (2017) cho khu vực gió mùa Ấn Độ.

Bên cạnh đó, nhiều nghiên cứu đã chỉ ra khả năng dự báo hạn nội mùa của ngày bắt đầu và kết thúc mùa mưa bằng các sản phẩm của mô hình động lực, đặc biệt là tại các khu vực gió mùa khác nhau trên thế giới. Cụ thể, Bombardi và cs (2017) đã nghiên cứu khả năng dự báo ngày bắt đầu và kết thúc mùa mưa tại 7 khu vực gió mùa điển hình trên thế giới bao gồm: Bắc Mỹ, Nam Mỹ, Tây Phi, Đông Phi, Ấn Độ, Đông Á, Bắc Úc. Các kết quả đánh giá được dựa trên sản phẩm của 3 mô hình: mô hình của CMA, mô hình của JMA, mô hình của NCEP thuộc khuôn khổ của dự án S2S. Kết quả thu được cho thấy chất lượng dự báo tại các khu vực nghiên cứu là khác nhau, theo đó chỉ những khu vực gió mùa như Bắc Mỹ, Đông Á, Bắc Úc, ngày bắt đầu và kết thúc mùa mưa có thể được dự báo với hạn dự báo nội mùa (khoảng 30 ngày). Đối với khu vực gió mùa Tây Phi, Kumi và cs (2020) đã cho thấy kết quả dự báo từ mô hình động lực có thể nắm bắt được những đặc trưng chính liên quan mùa mưa do gió mùa (ngày bắt đầu, đỉnh, ngày kết thúc) trên khu vực Tây Phi với hạn dự báo từ 20 đến 60 ngày. Bên cạnh đó, nghiên cứu cũng chỉ ra việc lựa chọn chỉ tiêu phù hợp xác định NBĐMM đóng vai trò quan trọng và ảnh hưởng lớn tới kết quả đánh giá khả năng mô phỏng của mô hình bằng việc so sánh kết quả dự báo khi sử dụng 2 cách xác định NBĐMM khác nhau.

Nhận xét chung:

Tóm lại, có thể thấy rằng NBĐMM đã được nhiều nghiên cứu trong và ngoài nước tiến hành với các chủ đề khác nhau (khảo sát đặc trưng, dự báo) và với sự đa dạng về các nguồn số liệu (như quan trắc tại trạm, tái phân tích, mô hình) và tại các quy mô khác nhau (địa phương, khu vực). Đối với khu vực Việt Nam, các nghiên cứu hiện vẫn đang được tiến hành chủ yếu tập trung đến khu vực Tây Nguyên và Nam Bộ, trong khi đó, vai trò và thông tin về đặc trưng và khả năng dự báo NBĐMM là quan trọng và cần thiết tại tất cả các vùng khí hậu.

Tại Việt Nam, các nghiên cứu trước đây cho thấy các đặc trưng mưa có sự bất đồng nhất không chỉ giữa các vùng mà còn giữa các trạm trong cùng một vùng. Do đó, thông tin về các đặc trưng của NBĐMM, đặc biệt là tại quy mô địa phương (từng trạm) là cần thiết và hữu ích cho nhiều lĩnh vực đời sống kinh tế-xã hội. Tuy nhiên, thách thức với các nghiên cứu liên quan đến việc xác định các đặc trưng khí hậu của NBĐMM đó chính là việc lựa chọn chỉ tiêu xác định phù hợp. Có thể thấy rằng, có sự không thống nhất giữa các nghiên cứu trước đây cho khu vực Việt Nam về chỉ tiêu xác định NBĐMM. Điều này khiến cho NBĐMM được xác định bằng các chỉ tiêu khác nhau có sự tương đồng về mặt khí hậu (giá trị trung bình trong nhiều năm), nhưng lại có sự khác biệt nhất định nếu đi vào một năm cụ thể. Cần lưu ý, các chỉ tiêu xác định NBĐMM trong các nghiên cứu trước đây được lựa chọn mang tính chủ quan, đa phần chỉ dựa trên đặc trưng chung về chế độ mưa của vùng khí hậu đang nghiên cứu mà chưa tính toán đến sự khác biệt giữa các trạm. Do đó, trong một số trường hợp, việc sử dụng chỉ tiêu xác định NBĐMM chung cho một vùng khí hậu là chưa phù hợp với một số trạm có sự khác biệt lớn với những đặc trưng mưa mang tính cục bộ.

Bên cạnh đó, mặc dù sử dụng các chỉ tiêu xác định khác nhau nhưng mối quan hệ chặt chẽ giữa sự biến động hàng năm của NBĐMM tại khu vực Tây Nguyên và Nam Bộ với ENSO đã được chỉ ra trong nhiều nghiên cứu. Tuy nhiên, việc dựa trên mối quan hệ này để dự báo NBĐMM vẫn chưa được thực hiện. Do đó, việc thử nghiệm dự báo NBĐMM bằng phương pháp thống kê truyền thống với nhân tố dự báo là các chỉ số khí hậu liên quan tới ENSO cho một số vùng trên khu vực Việt Nam là có cơ sở.

Đối với bài toán dự báo, các nghiên cứu trên thế giới đã chỉ ra rằng có khả năng dự báo hạn nội mùa-mùa NBĐMM/gió mùa dựa trên các sản phẩm của mô hình số tại nhiều khu vực gió mùa điển hình trên thế giới như: Châu Á, Bắc Úc, Nam Mỹ, Ấn Độ, Tây Phi, ... Tuy nhiên, khả năng dự báo NBĐMM/gió mùa cũng phụ thuộc vào chỉ tiêu xác định. Theo đó, đa phần các nghiên cứu đều cho thấy các chỉ tiêu mô phỏng sự thay đổi của hoàn lưu quy mô lớn cho kết quả tốt hơn so với các chỉ tiêu mô phỏng sự thay đổi mang tính địa phương. Điều này cho thấy các mô hình số hiện tại dường như cho kết quả mô phỏng các điều kiện hoàn lưu quy mô lớn tốt hơn so

với các điều kiện mang tính địa phương. Đây cũng có thể coi là một tín hiệu tốt cho bài toán dự báo NBĐMM/gió mùa bằng phương pháp thống kê-động lực dựa trên sản phẩm về các trường hoàn lưu của mô hình số. Trên thực tế, phương pháp này đã ghi nhận được kết quả dự báo tốt hơn so với dự báo khí hậu tại các khu vực khác nhau trên thế giới. Đối với khu vực Việt Nam, phương pháp này mới chỉ được xem xét cho khu vực Nam Bộ và thu được các kết quả khả quan.

Bên cạnh đó, việc sử dụng trực tiếp sản phẩm mưa từ các mô hình số để dự báo NBĐMM với hạn nội mùa cũng đã thu được những tín hiệu tốt tại nhiều khu vực gió mùa điển hình trên thế giới như Tây Phi, Bắc Mỹ, Đông Á, Bắc Úc. Một điểm lưu ý của cách tiếp cận này mà các nghiên cứu trước đây đề cập đến đó là việc lựa chọn các chỉ tiêu xác định NBĐMM cũng đóng vai trò quyết định tới chất lượng của các dự báo đưa ra. Tuy nhiên, đối với khu vực Việt Nam, việc dự báo NBĐMM theo cách tiếp cận này hiện vẫn chưa được khảo sát trong khi việc cung cấp thông tin dự báo sớm trước 1-2 tháng đóng vai trò quan trọng trong nhiều lĩnh vực của xã hội, đặc biệt là trong canh tác nông nghiệp, quản lý hồ chứa.

Những tổng quan trên là cơ sở, định hướng để luận án thực hiện khảo sát các bộ dữ liệu khác nhau và lựa chọn phương pháp phù hợp để hoàn thành mục tiêu. Nội dung chi tiết về số liệu và phương pháp sử dụng trong luận án sẽ được đề cập tại Chương 2.

Chương 2. SỐ LIỆU VÀ PHƯƠNG PHÁP NGHIÊN CỨU

2.1. SỐ LIỆU VÀ KHU VỰC NGHIÊN CỨU

Luận án sử dụng 4 bộ số liệu, bao gồm: 1) Số liệu mưa ngày quan trắc tại trạm; 2) Các chỉ số khí hậu; 3) Sản phẩm dự báo lại của mô hình CFSv2 (bao gồm các trường khí áp mực biển (PRMSL), gió vĩ hướng mực 850mb (U850), nhiệt độ mặt nước biển (SST) trung bình tháng 3 và tháng 7, lượng mưa hàng ngày); 4) Lượng mưa ngày dự báo của mô hình S2S-ECMWF. Chi tiết về các bộ số liệu này được mô tả dưới đây.

2.1.1. Số liệu quan trắc

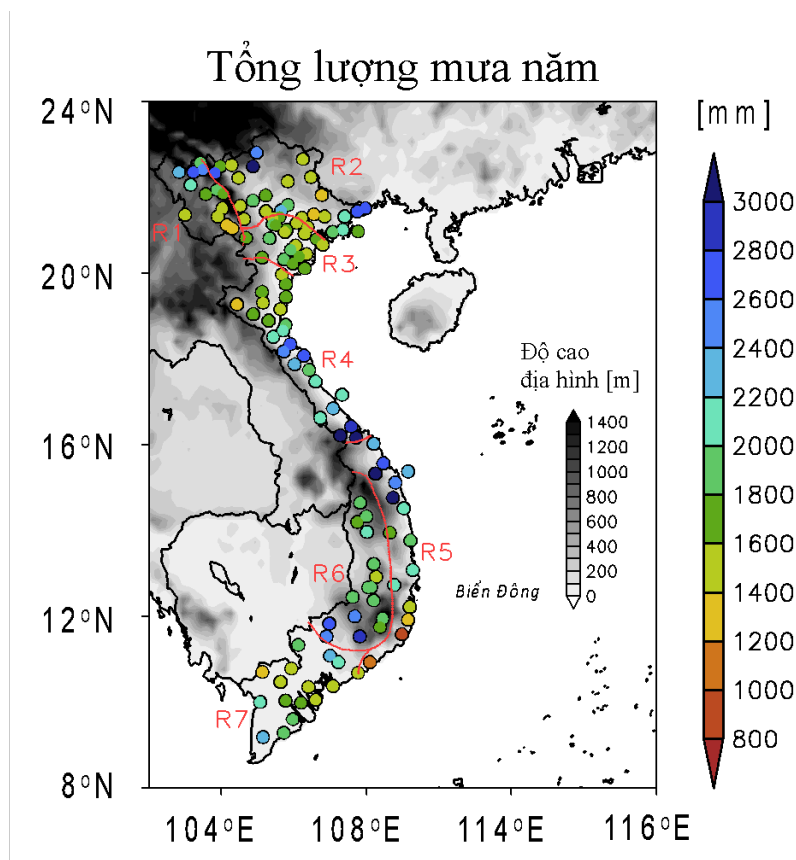
Số liệu mưa ngày tại các trạm quan trắc sẽ được sử dụng để xác định các đặc trưng của NBĐMM ở trong Chương 3. Để tránh những sai sót có thể xảy ra, NBĐMM của một năm bất kỳ chỉ được xác định khi chuỗi số liệu của năm đó thoả mãn điều kiện:

- (1) Số liệu khuyết không quá 25 ngày (tức là 1 năm có ít nhất 340 ngày);
- (2) Trong giai đoạn xung quanh thời điểm xuất hiện NBĐMM (tháng 3-tháng 6), số liệu mưa ngày quan trắc không được thiếu quá 3 ngày liên tiếp.

Đối với những trường hợp thoả mãn đồng thời cả 2 điều kiện trên mà vẫn còn tồn tại số liệu khuyết thì những ngày mà không có số liệu trên sẽ được bổ khuyết bằng giá trị trung bình nhiều năm của ngày tương ứng.

Kết quả sau quá trình tiền xử lý trên đã lựa chọn được 131 trạm phân bố tương đối đồng đều trên 7 vùng khí hậu với độ dài chuỗi từ 1979-2019. Sơ đồ phân vùng khí hậu Việt Nam được sử dụng trong luận án dựa theo Nguyễn Đức Ngữ và Nguyễn Trọng Hiệu (2004), trong đó lãnh thổ Việt Nam được chia thành 7 vùng, bao gồm Tây Bắc (ký hiệu R1), Đông Bắc (R2), Đồng bằng Bắc Bộ (R3), Bắc Trung Bộ (R4),

Nam Trung Bộ (R5), Tây Nguyên (R6), Nam Bộ (R7). Số lượng các trạm tương ứng với từng vùng khí hậu có thể xem tại Bảng 3.1. Địa hình Việt Nam, tọa độ của 131 trạm và ranh giới các vùng khí hậu được sử dụng trong nghiên cứu được thể hiện tại Hình 2.1.



Hình 2. 1: Tổng lượng mưa trung bình tại 131 trạm quan trắc trong giai đoạn 1979-2019. Đường màu đỏ thể hiện ranh giới phân chia giữa các vùng khí hậu

2.1.2. Số liệu các chỉ số khí hậu

Chỉ số khí hậu là các đại lượng được xây dựng cho mục đích nghiên cứu dao động, biến đổi khí hậu và cung cấp thông tin về một số dạng dao động khí hậu. Các chỉ số này được tính toán từ các trường khí quyển, đại dương và thường được biểu diễn bằng chuỗi giá trị trung bình tháng. Có nhiều cách khác nhau để thể hiện các chỉ số khí hậu nhưng thông thường là các chuỗi thời gian các tháng liên tiếp. Ví dụ, các chỉ số liên quan tới ENSO thường được biểu diễn dưới dạng chuỗi thời gian của trung bình trượt hoặc dị thường hoặc chuẩn hoá trên một vùng không gian nào đó, chẳng hạn dị thường nhiệt độ mặt nước biển (SSTA) các vùng Nino, chỉ số dao động nam (SOI),... Hiện nay, các chỉ số khí hậu được cung cấp miễn phí bởi nhiều trung tâm trên thế giới như Phòng Thí nghiệm Khoa học Vật lý và có thể dễ dàng tiếp cận được thông qua trang web: <https://psl.noaa.gov/data/climateindices/list/>.

Để thử nghiệm khả năng dự báo NBĐMM bằng phương pháp thống kê truyền thống dựa trên các chỉ số khí hậu, luận án sẽ chỉ tập trung vào một số chỉ số được lựa chọn. Việc lựa chọn này được dựa trên kết quả từ các nghiên cứu trước đây khi chỉ ra rằng sự biến động hàng năm của NBĐMM tại một số vùng khí hậu của Việt Nam có mối quan hệ chặt chẽ với các pha của ENSO, hoạt động của gió mùa, cũng như sự biến đổi của một số trường khí quyển, đại dương [Nguyễn Thị Hiền Thuận và Chiêu Kim Huỳnh, 2007; Nguyen và cs, 2007; Phạm Xuân Thành và cs, 2010; Phạm Thị Châm và cs, 2017; Ngo-Thanh và cs, 2018; Chu Thị Thu Hường và Trần Đình Linh, 2019]. Nhằm hướng tới bài toán dự báo nghiệp vụ, luận án chỉ lựa chọn các chỉ số đảm bảo một số tiêu chí sau:

- Chuỗi số liệu đủ dài. Độ dài chuỗi thời gian ít nhất 30 năm. Để phù hợp với độ dài của chuỗi số liệu quan trắc, các chuỗi này phải sẵn có từ 1979-2019.
- Có tính cập nhật thường xuyên. Thực tế thời gian cập nhật các chỉ số khí hậu là khác nhau dao động từ 1-2 tháng cho tới 1 năm. Với mục đích dự báo hạn mùa NBĐMM, các chỉ số cập nhật thường xuyên trong khoảng thời gian dưới 6 tháng tính đến thời điểm làm dự báo là chấp nhận được.

Thông tin về các chỉ số khí hậu được lựa chọn trong luận án được thể hiện ở Bảng 2.1.

Bảng 2.1: Các chỉ số khí hậu được sử dụng trong luận án

TT	Ký hiệu chỉ số	Tên gọi và địa chỉ lấy số liệu	Nhóm
1	Niño1+2	Chỉ số SST vùng Niño 1+2 (0-10°S, 90°W-80°W) https://www.psl.noaa.gov/data/correlation/NINO1.data	ENSO
2	Niño3	SST vùng nhiệt đới Đông Thái Bình Dương (5°N-5°S, 150°W-90°W) https://www.psl.noaa.gov/data/correlation/NINO3.data	ENSO
3	Niño3.4	Chỉ số SST vùng Niño 3.4 (5°N-5°S, 170°E-120°W) https://www.psl.noaa.gov/data/correlation/NINO34.data	ENSO
4	Niño4	Chỉ số SST vùng Niño 4 (5°N-5°S, 160°E-150°W)	ENSO

		https://www.psl.noaa.gov/data/correlation/NINO4.data	
5	MEIV2	Chỉ số ENSO đa biến phiên bản 2 https://psl.noaa.gov/data/correlation/meiv2.data	ENSO
6	CENSO	Chỉ số ENSO kết hợp https://psl.noaa.gov/data/correlation/censo.data	ENSO
7	ONI	Chỉ số Niño đại dương https://www.psl.noaa.gov/data/correlation/oni.data	Đại Dương
8	TNI	Chỉ số khuynh hướng Niño hay chỉ số Niño-chuyển dịch (Trans-Niño Index) https://psl.noaa.gov/data/correlation/tni.data	Đại Dương
9	WHWP	Bề nóng bán cầu tây https://www.psl.noaa.gov/data/correlation/whwp.data	Đại Dương
10	TNA	<i>Chỉ số Bắc Đại Tây Dương nhiệt đới</i> (Dị thường SST trung bình tháng vùng 5.5°N - 23.5°N, 15°W - 57.5°W) https://www.psl.noaa.gov/data/correlation/tna.data	Đại Dương
11	TSA	<i>Chỉ số Nam Đại Tây Dương nhiệt đới</i> (SSTA trung bình tháng vùng 0 - 20°S, 10°E - 30°W) https://www.psl.noaa.gov/data/correlation/tsa.data	Đại Dương
12	REPAC	Dị thường chuẩn hóa khí áp mực biển vùng xích đạo Đông Thái Bình Dương https://www.cpc.ncep.noaa.gov/data/indices/REPAC.for	Khí quyển
13	RINDO	Dị thường chuẩn hóa khí áp mực biển vùng biển Indonesia https://www.cpc.ncep.noaa.gov/data/indices/RINDO.for	Khí quyển
14	QBO30	Chỉ số dao động tựa 2 năm (theo gió vĩ hướng mực 30mb) https://www.psl.noaa.gov/data/correlation/qbo.data	Khí quyển
15	QBO50	Chỉ số dao động tựa 2 năm (theo gió vĩ hướng mực 50mb) https://www.cpc.ncep.noaa.gov/data/indices/qbo.u50.index	Khí quyển
16	SOI	Chỉ số dao động nam https://www.psl.noaa.gov/data/correlation/soi.data	Khí quyển
17	REQSOI	Chỉ số dao động nam vùng xích đạo (Equatorial SOI) https://www.cpc.ncep.noaa.gov/data/indices/reqsoi.for	Khí quyển

18	ZWND20	<i>Gió vĩ hướng vùng xích đạo mực 200mb (165°W-11°0W)</i> https://www.cpc.ncep.noaa.gov/data/indices/zwnd200	Khí quyển
19	WPAC85	<i>Tín phong mực 850 mb (135°E-180°W, 5°N-5°S)</i> https://www.cpc.ncep.noaa.gov/data/indices/wpac850	Khí quyển
20	EPAC85	<i>Tín phong mực 850 mb (135°W-120°W, 5°N - 5°S)</i> https://www.cpc.ncep.noaa.gov/data/indices/epac850	Khí quyển
21	CPAC85	<i>Tín phong mực 850 mb (17°5W - 140°W, 5°N - 5°S)</i> https://www.cpc.ncep.noaa.gov/data/indices/cpac850	Khí quyển
22	PNA	Chỉ số Thái Bình Dương Bắc Mỹ https://psl.noaa.gov/data/correlation/pna.data	Khí quyển
23	WP	Chỉ số Tây Thái Bình Dương https://www.psl.noaa.gov/data/correlation/wp.data	Khí quyển

2.1.3. Số liệu dự báo lại của mô hình CFSv2

Số liệu dự báo lại CFSv2 là sản phẩm dự báo lại của hệ thống dự báo Khí hậu (CFS) được nghiên cứu và phát triển bởi Trung tâm Dự báo Môi trường Hoa Kỳ (National Centers for Environmental Prediction - NCEP). Phiên bản đầu tiên có tên gọi là CFSv1, đã được khai thác thử nghiệm trong công tác động nghiệp vụ tại NCEP vào tháng 8 năm 2004. CFSv1 được xây dựng và phát triển bằng cách tổ hợp 4 thành phần độc lập sau: (1) điều kiện ban đầu của khí quyển và bề mặt được cung cấp bởi NCEP-DOE-Reanalysis 2 [Kanamitsu và cs, 2002]; (2) trạng thái ban đầu của đại dương được cung cấp bởi hệ thống đồng hoá số liệu toàn cầu GODAS đã được sử dụng trong các công tác nghiệp vụ tại NCEP từ năm 2003 [Behringer và cs, 2007]; (3) mô hình khí quyển của hệ thống dự báo toàn cầu của NCEP (GFS); (4) mô hình đại dương MOM3 của phòng thí nghiệm động lực học địa-vật lý (Geophysical Fluid Dynamics Laboratory (GFDL)). Thực tế, các sản phẩm dự báo CFSv1 có chất lượng rất tốt, đến mức có rất nhiều nghiên cứu vẫn sử dụng nó ngay cả khi các sản phẩm dự báo CFSv2 được đưa vào sử dụng từ tháng 3 năm 2011. Tới tháng 9 năm 2012, NCEP chính thức ngừng cung các sản phẩm dự báo CFSv1. CFSv2, phiên bản cập nhật của

CFSv1, là mô hình kết hợp các thành phần khí quyển-đại dương-đất-băng biển, hiện đang được sử dụng chạy dự báo nghiệp vụ hạn mùa tại NCEP [Saha và cs, 2014].

CFSv2 là phiên bản đã được phát triển và cải thiện cả 4 thành phần, bao gồm, 2 mô hình dự báo và 2 hệ thống đồng hoá số liệu. Ngoài ra, CFSv2 cũng có một số điểm khác biệt sau: mô hình đất 4 mực, mô hình tương tác băng biển 3 lớp, và có tính đến sự thay đổi của nồng độ khí CO₂ trong quá khứ (ví dụ: tăng/giảm nồng độ qua từng thời kỳ). Nhưng trên hết, CFSv2 đã cải thiện được sự bất nhất quán giữa các mô hình trạng thái và số liệu trạng thái ban đầu từ hệ thống đồng hoá số liệu. Hạn dự báo của CFSv2 là 9 tháng tính từ tháng tiếp theo sau tháng mà mô hình bắt đầu chạy dự báo vào một ngày nào đó. Như vậy, nếu tính bằng ngày thì hạn dự báo bằng số ngày từ ngày bắt đầu chạy mô hình đến hết tháng hiện tại cộng với số ngày của 9 tháng tiếp theo. Sản phẩm dự báo của mô hình được trích xuất từng 6 giờ một cho tất cả các biến trường dự báo và các biến thứ cấp (hay biến cảnh báo - diagnostic variables).

Đối với mục đích đánh giá khả năng dự báo của mô hình số CFSv2, bộ số liệu được sử dụng sẽ lấy trong giai đoạn 1983-2019, với độ phân giải 1.0°. Bên cạnh đó, luận án cũng tiến hành đánh giá chất lượng dự báo hạn mùa NBĐMM với 02 cách tiếp cận phổ biến đã được áp dụng trong các nghiên cứu trước đây (xem mục 1.3). Vì vậy, các biến sử dụng sẽ có sự khác biệt trong từng trường hợp, cụ thể:

- Dự báo NBĐMM bằng việc xây dựng mô hình thống kê dựa trên sản phẩm về các trường hoàn lưu khí quyển và nhiệt độ bề mặt biển của mô hình CFSv2: trường dự báo trung bình tháng 3 và tháng 7 của PRMSL, U850, SST giới hạn trong miền từ 40°E-100°W và từ 40°S-40°N.
- Dự báo NBĐMM bằng việc sử dụng trực tiếp sản phẩm mưa của mô hình CFSv2: lượng mưa ngày dự báo.

2.1.4. Số liệu dự báo lại của mô hình S2S-ECMWF

S2S-ECMWF là hệ thống dự báo hạn nội mùa của ECMWF với hạn dự báo đến 46 ngày. Các sản phẩm dự báo này được cung cấp với tần suất 2 lần/tuần vào thứ Hai và thứ Năm, trong đó bao gồm cả các sản phẩm dự báo lại trong 20 năm trước tính từ thời điểm có sản phẩm dự báo nghiệp vụ mới nhất. Vì vậy mà phiên bản của

các mô hình được sử dụng cho các sản phẩm dự báo lại cũng được cập nhật và điều chỉnh thường xuyên như đối với sản phẩm dự báo nghiệp vụ.

Tuy nhiên, với mục đích đánh giá khả năng sử dụng trực tiếp sản phẩm mưa của mô hình của ECMWF (ký hiệu là S2S-ECMWF) để dự báo hạn nội mùa NBĐMM, luận án sẽ chỉ sử dụng các sản phẩm dự báo lại. Cụ thể là luận án sẽ sử dụng bộ dữ liệu mưa dự báo lại cho giai đoạn 2000-2019 từ 2 phiên bản mô hình được sử dụng trong năm 2020, trong đó phiên bản CY46R1 được dùng cho đến ngày 29 tháng 6 năm 2020 và phiên bản CY47R1 được dùng từ ngày 30 tháng 6 năm 2020 trở đi. Sản phẩm mưa dự báo lại của S2S-ECMWF bao gồm 11 thành phần (1 thành phần gốc và 10 thành phần nhiễu đối với điều kiện ban đầu). Hạn dự báo là 46 ngày. Khi thực hiện dự báo mô hình chạy với độ phân giải không gian thay đổi theo hạn dự báo, trong đó 15 ngày đầu có độ phân giải Tco639 L91 (khoảng 16 km) và từ sau ngày thứ 15 trở đi có độ phân giải Tco319 (khoảng 32 km). Tuy nhiên khi cung cấp cho người dùng các số liệu này đã được quy về cùng độ phân giải cho tất cả các hạn dự báo. Ngoài ra, để đáp ứng nhu cầu người dùng, các bộ số liệu có độ phân giải khác nhau cũng đã được tạo ra. Luận án lựa chọn độ phân giải cao nhất có thể là $0.125^\circ \times 0.125^\circ$. Thông tin cũng như dữ liệu về bộ số liệu này được cung cấp miễn phí trực tuyến tại địa chỉ: <https://apps.ecmwf.int/datasets/data/s2s/>.

2.2. XÁC ĐỊNH NGÀY BẮT ĐẦU MÙA MƯA VÀ XU THẾ BIẾN ĐỔI TỪ SỐ LIỆU QUAN TRẮC TẠI TRẠM

2.2.1. Xác định ngày bắt đầu mùa mưa tại trạm

Trong luận án này, NBĐMM xác định từ số liệu mưa quan trắc tại 131 trạm trên khu vực Việt Nam (mục 2.1.1) được định nghĩa một cách tổng quát là ngày đầu tiên của năm đáp ứng các điều kiện sau:

- Tổng lượng mưa của 5 ngày liên tiếp lớn hơn P mm.
- NBĐMM và ít nhất N ngày trong chuỗi 5 ngày liên tiếp có lượng mưa lớn hơn 0.1 mm/ngày.
- Không quá 7 ngày không mưa liên tiếp xảy ra trong D ngày tiếp theo sau NBĐMM.

Trong đó, các tham số P, N, D sẽ được xác định một cách khách quan dựa trên tập giá trị có thể nhận của chúng. Trên thực tế, giá trị của các tham số này trong các nghiên cứu trước đây không giống nhau [Stern và cs, 1981; Sivakumar, 1988; Omotosho, 1992; Nguyễn Hướng Điền và Trần Công Minh, 2000; Nguyễn Thị Hiền Thuận và Chiêu Kim Quỳnh, 2007; Moron và cs, 2009; Phạm Thị Châm và cs, 2017; Lê Thị Xuân Lan, 2017] và phần lớn được lựa chọn một cách chủ quan dựa trên đặc điểm mưa tại khu vực áp dụng. Vấn đề đặt ra là làm thế nào để lựa chọn được bộ các ngưỡng giá trị của P, N, D một cách khách quan và phù hợp cho việc xác định NBĐMM?

Để giải quyết bài toán này các ngưỡng giá trị phù hợp của P, N, D có thể được xây dựng dựa trên các tín hiệu NBĐMM tại các trạm (quy mô địa phương) và trên toàn vùng (quy mô khu vực) nơi mà các trạm trong vùng có sự tương đồng về đặc điểm của NBĐMM. Nói cách khác, bộ tham số P, N, D nên được lựa chọn sao cho NBĐMM phản ánh được tính đồng nhất tương đối trong vùng và tính địa phương của trạm. Phương pháp này gồm 3 bước chính sau đây khi sử dụng số liệu quan trắc tại trạm giai đoạn 1979-2019:

Bước 1: Xác định tín hiệu NBĐMM cho từng trạm.

Tín hiệu về NBĐMM cho từng trạm (quy mô địa phương) được xác định dựa trên việc điều chỉnh bộ các ngưỡng giá trị P, N, D được lựa chọn. Về nguyên tắc các ngưỡng giá trị này có thể chọn tùy ý, tuy nhiên để giảm bớt khối lượng tính toán chúng có thể được lựa chọn trong phạm vi hợp lý. Trong luận án tập các giá trị của P, N và D được lựa chọn trên cơ sở tổng hợp từ các nghiên cứu trước đây [Stern và cs, 1981; Sivakumar, 1988; Omotosho, 1992; Nguyễn Hướng Điền và Trần Công Minh, 2000; Nguyễn Thị Hiền Thuận và Chiêu Kim Quỳnh, 2007; Moron và cs, 2009; Phạm Thị Châm và cs, 2017; Lê Thị Xuân Lan, 2017] cũng như kết quả phân tích số liệu quan trắc tại trạm. Tập giá trị của các tham số P, N và D đó là:

- P=10, 15, 20, 25, 30, 40, hoặc 50 mm;
- N=2, 3, hoặc 4;
- D=20 hoặc 30 ngày.

Bộ các ngưỡng giá trị này khi kết hợp lại với nhau sẽ tạo ra $7 \times 3 \times 2 = 42$ chỉ tiêu NBĐMM khác nhau. Như vậy, mỗi chỉ tiêu sẽ xác định một giá trị NBĐMM cho mỗi năm, và do đó mỗi trạm sẽ có 42 chuỗi giá trị NBĐMM hay chính là 42 tín hiệu khác nhau về NBĐMM.

Sự khác biệt của NBĐMM giữa các chỉ tiêu khác nhau sẽ được đánh giá thông qua chỉ số độ nhạy SI (sensitivity index) cho từng trạm:

$$SI = \frac{1}{L \times \frac{M \times (M-1)}{2}} \sum_{k=1}^L \sum_{j=1}^{M-1} \sum_{e=j+1}^M |x_{kj} - x_{ke}|$$

với L là tổng số năm (L=41 năm, từ 1979-2019); M là tổng số các chỉ tiêu (M=42); x_{kj} là NBĐMM của năm k xác định theo chỉ tiêu j; x_{ke} là NBĐMM của năm k xác định theo chỉ tiêu e. Độ nhạy SI là thước đo mức độ biến động của NBĐMM qua các năm xác định được theo các chỉ tiêu khác nhau. Giá trị của SI lớn (nhỏ) cho biết sự khác biệt lớn (nhỏ) giữa các kết quả xác định NBĐMM theo các chỉ tiêu khác nhau. Các trạm có SI càng lớn thì việc lựa chọn các ngưỡng P, N và D (chỉ tiêu xác định NBĐMM) càng phải hết sức thận trọng. Nói cách khác, SI cho ta biết những trạm nào thuộc vùng nào tại đó việc lựa chọn chỉ tiêu xác định NBĐMM (các ngưỡng P, N và D) sẽ khó khăn hơn và do đó cần được xem xét một cách cẩn thận.

Bước 2: Xác định tín hiệu NBĐMM cho vùng.

Trong mỗi vùng khí hậu, NBĐMM có thể đến sớm hơn (muộn hơn) ở một số trạm do sự chi phối của điều kiện địa phương (mà chủ yếu là địa hình). Tuy nhiên, do tính đồng nhất tương đối, mỗi vùng khí hậu sẽ có những đặc điểm chung về NBĐMM mà các trạm trong vùng đều chịu ảnh hưởng. Đặc điểm chung đó có thể được hiểu là tín hiệu vùng của NBĐMM mà nó được thể hiện qua sự biến động hàng năm của NBĐMM trong toàn vùng. Sự biến động này có thể được xác định dựa trên chuỗi thời gian của các thành phần chính (ký hiệu là EOFs) thu được từ kết quả phân tích thành phần chính nhiều biến (Multi variable Principal Component Analysis, ký hiệu MPCA) áp dụng cho bộ số liệu đầu vào là các giá trị NBĐMM tại các trạm được xác định từ 42 chỉ tiêu kết hợp ở bước 1. Chi tiết về phương pháp MPCA có thể xem tại Wilks (2006). Số lượng EOFs được chọn phụ thuộc vào mức độ đóng góp vào tổng

phương sai cũng như mức ý nghĩa của chúng. Việc lựa chọn các EOFs sẽ được phân tích ở mục 3.2.

Hình 2.2 mô tả cấu trúc của số liệu đầu vào cho phương pháp MPCA. Cụ thể, NBĐMM được xác định từ 42 chỉ tiêu sẽ được coi như là các biến cho Nsta trạm trong vùng nghiên cứu của giai đoạn 1979-2019 (41 năm). Kích thước mảng đầu vào x tương ứng sẽ là 42 chỉ tiêu và Nsta trạm. Nsta thành phần đầu tiên của x tương ứng với NBĐMM trong 41 năm được xác định bằng chỉ tiêu thứ nhất; Nsta thành phần thứ hai của x tương ứng với NBĐMM trong 41 năm được xác định bằng chỉ tiêu thứ hai; tương tự như vậy, cuối cùng Nsta thành phần thứ 42 của x tương ứng với NBĐMM trong 41 năm được xác định bằng chỉ tiêu thứ 42. Cần lưu ý, khi áp dụng phương pháp PCA cho nhiều trường (tức MPCA), ma trận hệ số tương quan $[R]$ và của ma trận vector riêng $[E]$ có kích thước là $([42 \times N_{sta}] \times [42 \times N_{sta}])$. Việc áp dụng MPCA sẽ tạo ra các thành phần chính mà đã tối đa hóa giá trị phương sai chung của 42 chỉ tiêu dựa trên việc xem xét mối tương quan giữa các biến này tại Nsta trạm.

RSODs		Years						
		1979	1980	1981	1982	...	...	2019
1 st criterion	1	122	133	131	178	...	...	163
	2	142	167	195	246	...	...	163
	3	142	134	258	245	...	...	216
	...	148	169	173	246	...	...	209
	Nsta	208	201	258	245	...	...	252
2 nd criterion	1	149	199	183	230	...	...	163
	2	142	170	184	174	...	...	163
	3	150	242	275	245	...	...	252
	...	135	124	139	246	...	...	163
	Nsta	135	242	258	246	...	...	164
....	1	135	135	258	246	...	...	252
	2	129	207	145	140	...	...	139
	3	220	211	175	236	...	...	173
	...	147	258	175	130	...	...	173
	Nsta	123	275	132	141	...	...	133
42 nd criterion	1	221	254	248	238	...	...	224
	2	183	258	133	130	...	...	229
	3	121	277	143	149	...	...	132
	...	219	256	219	219	...	...	225
	Nsta	221	257	134	140	...	...	140

Hình 2. 2: Định dạng mẫu của số liệu đầu vào cho phương pháp MPCA

Chú ý rằng, các tín hiệu về NBĐMM trên quy mô vùng sẽ rõ ràng hơn khi biến trình mưa năm có sự tương đồng giữa các trạm; và việc thay đổi các chỉ tiêu không làm ảnh hưởng quá lớn đến kết quả xác định NBĐMM.

Bước 3: Kết hợp các tín hiệu NBĐMM tại trạm và vùng

Sự kết hợp tín hiệu NBĐMM tại trạm và vùng khí hậu tương ứng được thể hiện thông qua giá trị hệ số tương quan giữa NBĐMM được xác định theo từng chỉ tiêu (các ngưỡng giá trị P, N, D) (NBĐMM tại từng trạm, đại diện quy mô địa phương) và các EOFs xác định được từ bước 2 (tín hiệu của NBĐMM tại vùng khí hậu, đại diện quy mô vùng).

Theo đó, chỉ tiêu phù hợp nhất cho từng trạm là bộ giá trị P, N, D mà tương quan giữa NBĐMM quy mô địa phương và NBĐMM quy mô vùng cao nhất. Điều này đảm bảo ảnh hưởng của các đặc trưng khí hậu mang tính địa phương, như mưa do địa hình, cũng sẽ được tính đến trong biến động hàng năm của NBĐMM. Bên cạnh đó, sự tương đồng lớn (nhỏ) theo thời gian giữa NBĐMM tại trạm và NBĐMM của vùng khí hậu được phản ánh thông qua giá trị của hệ số tương quan giữa các đại lượng này. Những trạm có giá trị hệ số tương quan cao cho thấy NBĐMM tại đây chịu sự chi phối của hoàn lưu quy mô lớn rõ rệt hơn so với ảnh hưởng của điều kiện địa phương. Qua ba bước thực hiện đã trình bày sẽ nhận được bộ chỉ tiêu tối ưu (đã bao gồm cả tín hiệu quy mô địa phương và quy mô vùng) tại 131 trạm trên toàn khu vực Việt Nam, từ đó xác định được NBĐMM cho từng trạm dựa trên số liệu quan trắc.

Để nhận diện vai trò của các tham số P, N, D khi xây dựng chỉ tiêu xác định NBĐMM, phân tích phương sai một chiều (one-way analysis of variance, f-statistic, hay kiểm nghiệm f) đối với hệ số tương quan cũng được tiến hành riêng biệt cho từng trạm và từng ngưỡng của P, N, D. Mục đích của việc này là để kiểm tra giá trị phương sai của hệ số tương quan thay đổi như thế nào khi điều chỉnh các ngưỡng giá trị của P, N, D. Lý thuyết về kiểm nghiệm f được diễn giải như sau. Giả sử ta có K nhóm độc lập, tương ứng nhóm thứ i có n_i giá trị quan trắc. Ta ký hiệu Y_{ij} là giá trị quan

trắc thứ j của nhóm thứ i. Giả thiết H_0 được đặt ra là giá trị phương sai giữa các nhóm là bằng nhau (K nhóm).

Giá trị F được xác định theo công thức:

$$F_{K-1, N-K} = \frac{MSB}{MSE}$$

với

$$MSB = \frac{1}{K-1} \sum_{i=1}^K n_i (\bar{Y}_i - \bar{Y})^2$$

$$MSE = \frac{1}{N-K} \sum_{i=1}^K \sum_{j=1}^{n_j} (Y_{ij} - \bar{Y}_i)^2$$

$$\bar{Y} = \frac{1}{N} \sum_{i=1}^K \sum_{j=1}^{n_i} Y_{ij}$$

$$N = \sum_{i=1}^K n_i$$

$$\bar{Y}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} Y_{ij}$$

Giá trị F sau đó được so sánh với F_α được xác định ứng với xác suất phạm sai lầm loại I (α) với F tuân theo phân bố Fisher, $f(K-1, N-K)$. Trong nghiên cứu này, giá trị α được lựa chọn là 0.05, tương ứng với mức độ tin cậy của kiểm nghiệm là 95%. Theo đó, nếu:

- $F \geq F_\alpha \Rightarrow$ Bác bỏ giả thiết H_0
- $F < F_\alpha \Rightarrow$ Chấp nhận giả thiết H_0

Kết quả tiến hành 3 bước trên đây là bản đồ phân bố NBĐMM trên khu vực Việt Nam dựa trên các bộ chỉ tiêu phù hợp cho từng trạm. Thông tin về NBĐMM này sẽ được sử dụng như là NBĐMM thực tế trong giai đoạn 1979-2019.

2.2.2. Xác định ngày bắt đầu mùa mưa tại vùng khí hậu

Trong luận án, giá trị NBĐMM đại diện cho các vùng khí hậu sẽ được sử dụng để phục vụ cho bài toán đánh giá khả năng dự báo hạn mùa và nội mùa NBĐMM.

Xin lưu ý ở đây là NBĐMM không phải là đại lượng quan trắc trực tiếp mà nó chỉ được xác định gián tiếp thông qua các chỉ tiêu. Tuy nhiên, để thuận tiện cho việc trình bày, *NBĐMM được xác định từ số liệu quan trắc* (trong luận án là số liệu mưa ngày trên mạng lưới trạm) *sẽ được gọi là NBĐMM quan trắc*. NBĐMM quan trắc của một vùng khí hậu được định nghĩa là giá trị trung vị của NBĐMM từ tất cả các trạm trong vùng, trong đó NBĐMM tại các trạm được xác định bằng phương pháp đã trình bày ở mục 2.2.1.

2.2.3. Xu thế biến đổi của ngày bắt đầu mùa mưa

Xu thế biến đổi của NBĐMM tại 131 trạm quan trắc trong giai đoạn 1979-2019 sẽ được xác định thông qua giá trị hệ số góc Sen [Sen, 1968]. Độ tin cậy của xu thế được xác định bằng kiểm nghiệm phi tham số Mann-Kendall [Kendall, 1975] với mức tin cậy 95% (hay mức ý nghĩa 5%). Phương pháp xác định xu thế theo hệ số góc Sen và kiểm nghiệm xu thế Mann-Kendall đã được áp dụng trong nhiều nghiên cứu trước đây do những điểm nổi bật của chúng. Cụ thể là các phương pháp này (1) không cần biết trước quy luật phân bố của tập mẫu, và (2) giảm thiểu được sự ảnh hưởng của các giá trị cực trị đối với xu thế của chuỗi số liệu. Do đó, chúng được gọi là phương pháp phi tham số. Tóm tắt phương pháp tính hệ số góc Sen và kiểm nghiệm Mann-Kendall được trình bày sau đây.

Ký hiệu $\{x_1, x_2, \dots, x_n\}$ là chuỗi thời gian NBĐMM của một trạm nào đó, với x_i là NBĐMM của năm thứ i . Từ chuỗi này ta lập chuỗi mới gồm $n(n-1)/2$ phần tử theo công thức $y_t = \frac{x_j - x_k}{j - k}$; với $j, k = 1, 2, \dots, n; j > k; t = 1, 2, \dots, n(n-1)/2$. Khi đó hệ số góc Sen là trung vị của chuỗi $\{y_t\}$.

Với chuỗi $\{x_1, x_2, \dots, x_n\}$ trên đây, ta thành lập đại lượng thống kê Mann-Kendall (S):

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^n \text{sign}(x_j - x_k)$$

và định nghĩa biến ngẫu nhiên τ :

$$\tau = \begin{cases} \frac{S+1}{\sqrt{Var(S)}} & \text{khi } S < 0 \\ 0 & \text{khi } S = 0 \\ \frac{S+1}{\sqrt{Var(S)}} & \text{khi } S > 0 \end{cases}$$

Với $Var(S)$ là phương sai của S , được xác định bởi:

$$Var(S) = \frac{1}{18} \left[n(n-1)(2n+5) - \sum_{p=1}^g t_p(t_p-1)(2t_p+5) \right]$$

Trong đó g là số nhóm, trong đó mỗi nhóm là một tập các phần tử của chuỗi có cùng giá trị, và t_p là số các phần tử thuộc nhóm p . Với định nghĩa như trên, Q sẽ cùng dấu với τ .

Biến τ sẽ có phân bố chuẩn chuẩn hóa, $\tau \in N(0,1)$. Khi đó ứng với xác suất phạm sai lầm loại I (α) nào đó ta sẽ xác định được giá trị τ_α sao cho $P(|\tau| > \tau_\alpha) = \alpha$. Nếu $|\tau| > \tau_\alpha$ thì xu thế xác định theo hệ số góc Sen có ý nghĩa với mức ý nghĩa α (hay độ tin cậy $1 - \alpha$). Trong nghiên cứu này, độ tin cậy của xu thế được chọn là 95%, ứng với xác suất phạm sai lầm loại I là 5% (hay $\alpha = 0.05$).

2.3 DỰ BÁO HẠN MÙA NGÀY BẮT ĐẦU MÙA MƯA BẰNG PHƯƠNG PHÁP THỐNG KÊ TRUYỀN THỐNG

Việc dự báo NBĐMM bằng phương pháp thống kê truyền thống được dựa trên mối quan hệ giữa NBĐMM quan trắc tại các vùng khí hậu với các chỉ số khí hậu trong tháng 1 và tháng 2 đối với khu vực R1, R2, R3, R6 và R7; trong tháng 5 và tháng 6 đối với khu vực R4 và R5 (các chỉ số liệt kê tại mục 2.1.2). Về mặt kỹ thuật để đưa ra thông tin dự báo, phương pháp hồi quy từng bước (chi tiết xem tại cuốn “Các phương pháp thống kê trong khí hậu” của Phan Văn Tân, 2003) được sử dụng để lựa chọn các nhân tố dự báo phù hợp và xây dựng phương trình dự báo bằng mô hình hồi quy đa biến. Như vậy, đối với trường hợp này:

- Hạn dự báo: 2-3 tháng
- Số liệu cho tập luyện (số liệu phụ thuộc): giai đoạn 1979-2019
- Yếu tố dự báo: NBĐMM (xem mục 2.2.2)
- Nhân tố dự tuyển: 46 nhân tố, tương ứng với 23 chỉ số khí hậu (Bảng 2.1) của hai tháng. Giá trị của các nhân tố dự tuyển đều được quy tâm trước khi sử dụng phương pháp hồi quy từng bước.

2.4. DỰ BÁO HẠN MÙA NGÀY BẮT ĐẦU MÙA MƯA BẰNG PHƯƠNG PHÁP THỐNG KÊ-ĐỘNG LỰC

Dự báo hạn mùa NBĐMM bằng phương pháp thống kê-động lực là dựa trên mối quan hệ giữa NBĐMM quan trắc với các trường dự báo của CFSv2 (hoàn lưu khí quyển, SST). Mô hình thống kê sử dụng là mô hình hồi quy từng bước, tương tự như đối với trường hợp dự báo NBĐMM bằng phương pháp thống kê truyền thống (mục 2.3). Cụ thể:

- Số liệu cho tập luyện (số liệu phụ thuộc): 1983-2011
- Số liệu độc lập: 2012-2019 (dùng cho đánh giá)
- Yếu tố dự báo: NBĐMM (xem mục 2.2.2)
- Nhân tố dự tuyển: Các trường dự báo lại của CFS được lựa chọn dựa trên mối quan hệ với NBĐMM. Trong luận án, 3 trường trung bình tháng 3 (tháng 7) của PRMSL, U850 và SST từ số liệu dự báo lại của CFSv2 được sử dụng để dự báo NBĐMM cho vùng R1, R2, R3, R6 và R7 (cho vùng R4 và R5). Giá trị của các nhân tố dự tuyển được chuẩn hoá trước sử dụng phương pháp hồi quy từng bước.
- Hạn dự báo:

Đối với các vùng R1, R2, R3, R6 và R7:

- + 1 tháng: khi đứng làm dự báo tại tháng Hai
- + 2 tháng: khi đứng làm dự báo tại tháng Một
- + 3 tháng: khi đứng làm dự báo tại tháng Mười Hai
- + 4 tháng: khi đứng làm dự báo tại tháng Mười Một
- + 5 tháng: khi đứng làm dự báo tại tháng Mười
- + 6 tháng: khi đứng làm dự báo tại tháng Chín

Đối với các vùng R4 và R5:

- + 1 tháng: khi đứng làm dự báo tại tháng Sáu
- + 2 tháng: khi đứng làm dự báo tại tháng Năm

- + 3 tháng: khi đứng làm dự báo tại tháng Tư
- + 4 tháng: khi đứng làm dự báo tại tháng Ba
- + 5 tháng: khi đứng làm dự báo tại tháng Hai
- + 6 tháng: khi đứng làm dự báo tại tháng Một

2.5. DỰ BÁO NGÀY BẮT ĐẦU MÙA MƯA SỬ DỤNG SẢN PHẨM MƯA DỰ BÁO CỦA MÔ HÌNH SỐ

Bên cạnh việc dự báo NBĐMM bằng các mô hình thống kê, luận án sẽ đánh giá khả năng dự báo hạn nội mùa và hạn mùa NBĐMM sử dụng trực tiếp sản phẩm mưa dự báo của mô hình số trị. Trong phạm vi luận án, hai trường hợp sẽ được xem xét, bao gồm: (1) Dự báo hạn mùa NBĐMM dựa trên sản phẩm mưa của mô hình CFSv2 và (2) Dự báo hạn nội mùa NBĐMM dựa trên sản phẩm mưa của mô hình S2S-ECMWF.

Tương tự đối với NBĐMM quan trắc, trong trường hợp này NBĐMM dự báo *không phải là sản phẩm trực tiếp của mô hình* mà được xác định gián tiếp từ lượng mưa dự báo của mô hình (và được gọi là NBĐMM mô hình). Do đó *NBĐMM mô hình không chỉ phụ thuộc vào lượng mưa dự báo của mô hình mà còn phụ thuộc vào chỉ tiêu xác định NBĐMM từ lượng mưa mô hình*. Trên cơ sở kết quả nghiên cứu của Boyard-Micheau và cs (2013) cũng như các kết quả khảo sát của tác giả luận án cho khu vực Việt Nam, ở đây 5 chỉ tiêu xác định NBĐMM khác nhau sẽ được áp dụng cho lượng mưa mô hình để khảo sát chất lượng dự báo trong cả 2 trường hợp thử nghiệm. Các chỉ tiêu được xây dựng dựa trên việc điều chỉnh ngưỡng tổng lượng mưa của chuỗi 5 ngày liên tiếp mà trong đó NBĐMM là ngày đầu tiên. Thông tin cụ thể về hai trường hợp dự báo trên sẽ được trình bày dưới đây.

2.5.1. Dự báo hạn mùa NBĐMM dựa trên sản phẩm mưa của CFSv2

Hình 2.3 dẫn ra sơ đồ mô tả quy trình dự báo NBĐMM hạn mùa (dưới 6 tháng) sử dụng sản phẩm CFSv2 cho một vùng khí hậu. NBĐMM quan trắc trong hình 2.3 được xác định bằng phương pháp nêu ở mục 2.2.2. NBĐMM dự báo nhận được từ số liệu mưa của CFSv2 sẽ được xác định là ngày đầu tiên trong chuỗi giá trị mưa dự báo thoả mãn:

[1] Tổng lượng mưa của 5 ngày liên tiếp phải lớn hơn **P** (mm)

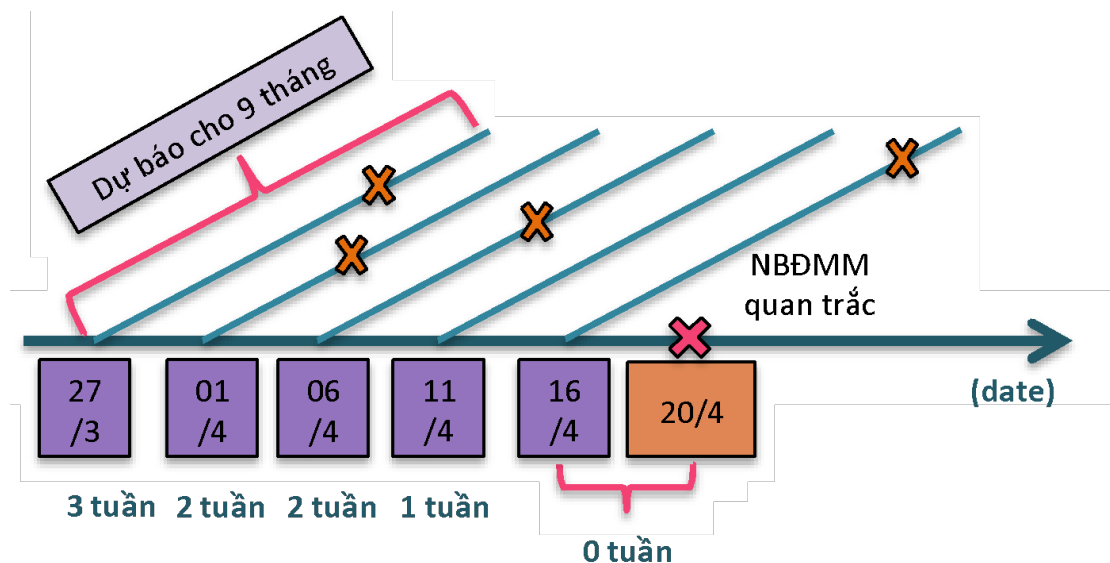
[2] Ngày bắt đầu và ít nhất 2 ngày trong 5 ngày liên tiếp phải đạt lượng mưa ngày lớn hơn **1** mm/ngày

[3] Không có quá 7 ngày liên tiếp không có mưa trong 30 ngày tiếp theo kể từ NBĐMM

Tham số **P** trong [1] sẽ nhận các giá trị 20, 25, 30, 40, 50 mm, ký hiệu tương ứng là P20, P25, P30, P40, P50, nhằm khảo sát ảnh hưởng của chỉ tiêu đến khả năng dự báo của mô hình, qua đó lựa chọn chỉ tiêu hợp lý nhất.

Việc xác định giá trị dự báo NBĐMM ứng với các hạn 1-6 tháng sẽ dựa trên khoảng thời gian từ NBĐMM quan trắc với ngày tiến hành dự báo (tính theo tuần tương ứng). Ví dụ: NBĐMM quan trắc là ngày 20/4, nếu ngày tiến hành dự báo là ngày 16/4 thì sẽ cách nhau 0 tuần và nếu ngày tiến hành dự báo là ngày 27/3 thì sẽ cách nhau 3 tuần (Hình 2.3). Theo đó, giá trị dự báo NBĐMM của hạn dự báo tương ứng với thời gian 1-6 tháng sẽ là giá trị trung vị của các dự báo nằm trong khoảng thời gian quy định như sau:

- Hạn 1 tháng: 5 - 8 tuần
- Hạn 2 tháng: 9 - 12 tuần
- Hạn 3 tháng: 13 - 16 tuần
- Hạn 4 tháng: 17 - 20 tuần
- Hạn 5 tháng: 21 - 24 tuần
- Hạn 6 tháng: 25 - 28 tuần



Hình 2.3: Sơ đồ mô tả quy trình dự báo NBĐMM hạn mùa (dưới 6 tháng) đối với sản phẩm mưa của CFSv2 cho một vùng khí hậu

2.5.2. Dự báo hạn nội mùa ngày bắt đầu mùa mưa dựa trên sản phẩm mưa của S2S-ECMWF

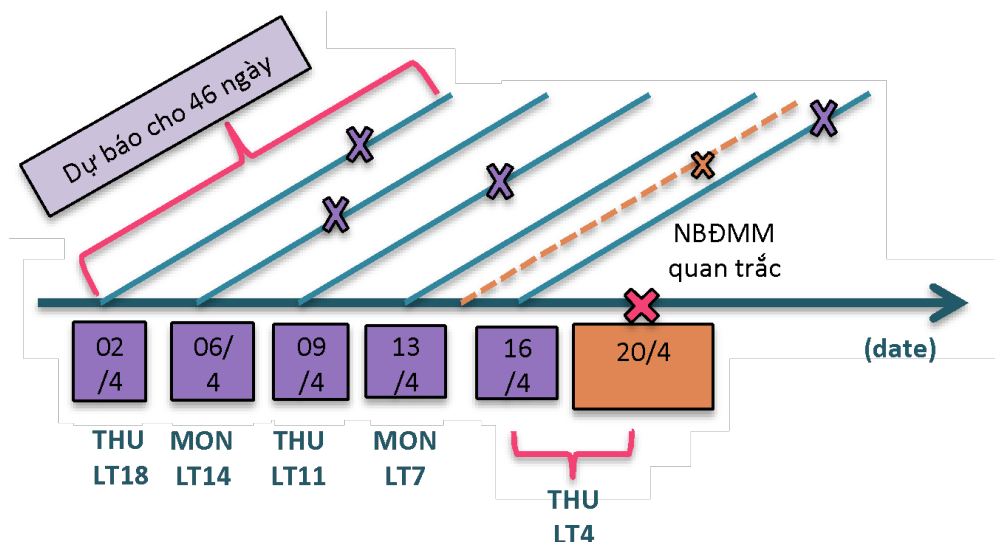
Hình 2.4 dẫn ra sơ đồ mô tả quy trình dự báo NBĐMM hạn nội mùa (7-40 ngày) sử dụng sản phẩm mưa S2S-ECMWF cho một vùng khí hậu. NBĐMM quan trắc trong Hình 2.4 là giá trị NBĐMM được xác định bằng phương pháp nêu ở mục 2.2.2. Về cơ bản chỉ tiêu xác định NBĐMM dự báo trong trường hợp này tương đồng với trường hợp dự báo hạn mùa trên sản phẩm CFSv2, tuy nhiên có sự điều chỉnh ở điều kiện [3] sao cho phù hợp với đặc điểm của chuỗi số liệu dự báo chỉ kéo dài 46 ngày của S2S-ECMWF.

Giá trị NBĐMM dự báo từ số liệu S2S-ECMWF sẽ được xác định dựa trên chỉ tiêu như sau: NBĐMM của một sản phẩm dự báo kéo dài 46 ngày sẽ là ngày đầu tiên trong chuỗi giá trị mưa dự báo thoả mãn:

- [1] Tổng lượng mưa của 5 ngày liên tiếp phải lớn hơn **Pmm**
- [2] Ngày bắt đầu và ít nhất 2 ngày trong 5 ngày liên tiếp phải đạt lượng mưa ngày lớn hơn 1mm/ngày
- [3'] Không có quá 7 ngày liên tiếp không có mưa theo kể từ NBĐMM

Việc khảo sát ảnh hưởng của chỉ tiêu xác định NBĐMM đến khả năng dự báo của mô hình cũng được tiến hành tương tự như trong trường hợp dự báo mùa bằng

sản phẩm mưa của CFSv2 (mục 2.5.1). Cụ thể, tham số **P** nhận các giá trị 20, 25, 30, 40, 50 mm, và được ký hiệu là P20, P25, P30, P40, P50 tương ứng.



Hình 2.4: Sơ đồ mô tả quy trình dự báo NBĐMM nội mùa (7-40 ngày) đối với sản phẩm mưa của S2S-ECMWF cho một vùng khí hậu

Việc xác định giá trị dự báo NBĐMM ứng với các hạn 7-40 ngày sẽ dựa trên khoảng thời gian tính theo ngày tương ứng tính từ NBĐMM quan trắc với ngày tiến hành dự báo. Ví dụ: NBĐMM quan trắc là ngày 20/4, nếu ngày tiến hành dự báo là ngày 16/4 thì sẽ cách nhau 4 ngày (tương ứng với hạn dự báo là 4 ngày) và nếu ngày tiến hành dự báo là ngày 2/4 thì sẽ cách nhau 18 ngày (tương ứng với hạn 18 ngày) (Hình 2.4). Tuy nhiên, các bản tin dự báo của ECMWF chỉ được đưa ra với tần suất 2 lần một tuần vào các ngày thứ Hai và thứ Năm hàng tuần, nên có thể thấy trên Hình 2.4, tại một số hạn dự báo (ví dụ hạn 8 ngày) không có thông tin dự báo. Do đó, trong các trường hợp này, giá trị dự báo sẽ là trung bình của NBĐMM được xác định bởi hai dự báo có thông tin liên kề hai bên. Ví dụ như đối với thông tin NBĐMM tại hạn dự báo 8 ngày (bị khuyết), giá trị dự báo NBĐMM sẽ là trung bình NBĐMM của dự báo ứng với hạn 4 ngày và 7 ngày (Hình 2.4). Do đó, các kết quả đánh giá NBĐMM với các hạn dự báo từ 7-40 ngày sẽ là thông tin đại diện cho hạn dự báo đó trong khoảng ± 3 hoặc 4 ngày.

Bên cạnh đó, do số liệu dự báo S2S-ECMWF gồm 11 thành phần khác nhau (xem mục 2.1.4). Do đó, luận án cũng sẽ tiến hành khảo sát sự khác biệt của NBĐMM

giữa các thành phần khác nhau thông qua chỉ số độ nhạy SIM (Sensitivity Index for Member) cho từng thành phần dự báo NBĐMM (tức là cho P20, P25, P30, P40, P50):

$$SIM = \frac{1}{L \times \frac{M \times (M-1)}{2}} \sum_{k=1}^L \sum_{j=1}^{M-1} \sum_{e=j+1}^M |x_{kj} - x_{ke}|$$

với L là tổng số năm (L=20 năm từ 2000-2019); M là tổng số các thành phần (M=11); x_{kj} là NBĐMM của năm k và được xác định bằng thành phần dự báo j; x_{ke} là NBĐMM của năm k và được xác định bằng thành phần dự báo e. Giá trị của SIM lớn (nhỏ) cho biết sự khác biệt lớn (nhỏ) giữa các kết quả thu được khi sử dụng các thành phần dự báo khác nhau. Ở các vùng và hạn dự báo tại đó chỉ số SIM cao (độ nhạy lớn) có nghĩa là kết quả dự báo tổ hợp NBĐMM từ 11 thành phần sẽ biến động mạnh, do đó cần được xem xét cẩn thận.

Để đánh giá tính hiệu quả của NBĐMM dự báo tổ hợp so với NBĐMM dự báo của một thành phần đơn lẻ ở đây sẽ sử dụng chỉ số AV (Added Value):

$$AV = \frac{RMSE_{mod} - RMSE_{ens}}{RMSE_{ens}}$$

với

$$RMSE_{mod} = \sqrt{\frac{1}{n} \sum_{i=1}^n (F_{mod,i} - O_i)^2}$$

$$RMSE_{ens} = \sqrt{\frac{1}{n} \sum_{i=1}^n (F_{ens,i} - O_i)^2}$$

Trong đó n là số năm, $F_{mod,i}$, $F_{ens,i}$, và O_i lần lượt là NBĐMM được xác định từ thành phần dự báo đơn lẻ, NBĐMM được xác định bằng việc tổ hợp 11 thành phần dự báo (cụ thể ở đây là trung vị tổ hợp) trong năm thứ i tương ứng. Do đó, AV là thước đo mức độ suy giảm sai số quân phương (RMSE) của dự báo tổ hợp so với dự báo của các thành phần đơn lẻ. Giá trị dương của AV có nghĩa là việc sử dụng dự báo tổ hợp từ 11 thành phần tốt hơn so với khi chỉ sử dụng một thành phần dự báo duy nhất.

2.6. PHƯƠNG PHÁP ĐÁNH GIÁ

2.6.1. Các chỉ số đánh giá

Để đánh giá chất lượng dự báo NBĐMM, các chỉ số thống kê thông dụng, bao gồm sai số trung bình (ME - Mean Error), sai số tuyệt đối trung bình (MAE-Mean Absolute Error), sai số quân phương (RMSE - Root Mean Square Error), hệ số tương quan (CORR-Correlation coefficient), điểm kỹ năng dự báo SS_{MAE} giữa giá trị dự báo và giá trị quan trắc. Cụ thể như sau:

$$\begin{aligned}ME &= \frac{1}{n} \sum_{i=1}^n (F_i - O_i) \\MAE &= \frac{1}{n} \sum_{i=1}^n |F_i - O_i| \\SS_{MAE} &= 1 - \frac{MAE}{MAE_{obs}} \\MAE_{obs} &= \frac{1}{n} \sum_{i=1}^n |O_i - \underline{O}| \\\underline{O} &= \frac{1}{n} \sum_{i=1}^n O_i \\CORR &= \frac{\sum_{i=1}^n (F_i - \underline{F})(O_i - \underline{O})}{\sqrt{\sum_{i=1}^n (F_i - \underline{F})^2} \sqrt{\sum_{i=1}^n (O_i - \underline{O})^2}}\end{aligned}$$

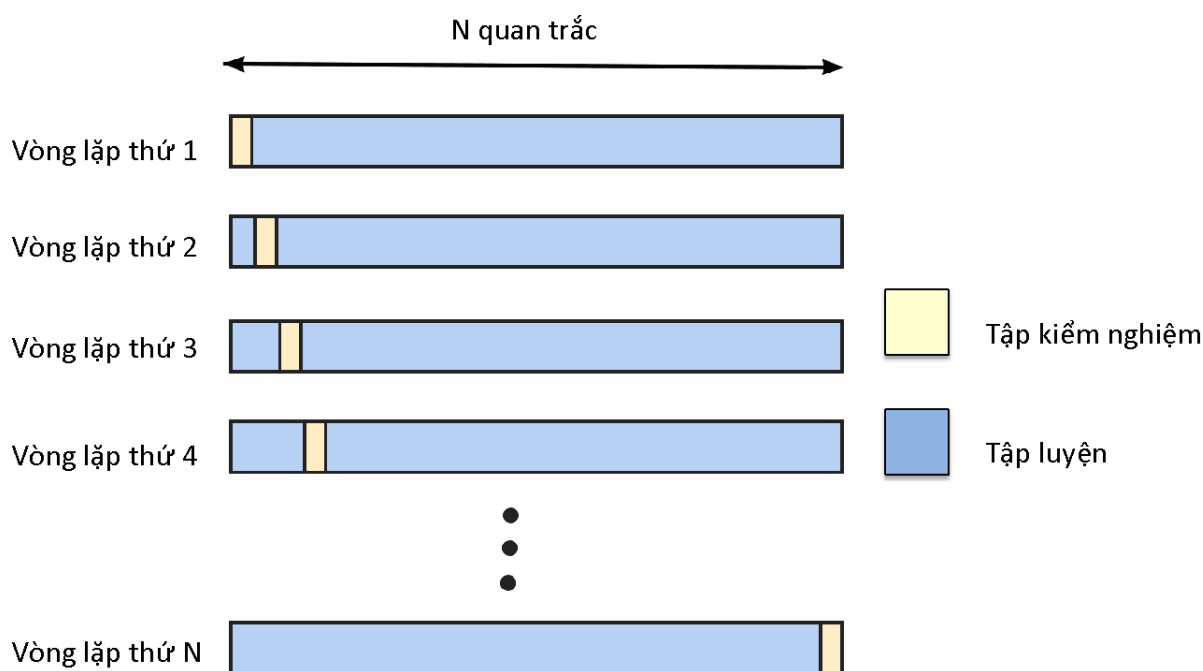
Trong đó, n là số năm; F_i là giá trị dự báo năm thứ i ; O_i là giá trị quan trắc năm thứ i ; $\underline{O}$ là giá trị trung bình khí hậu của NBĐMM dựa trên số liệu quan trắc tại trạm; MAE_{obs} là sai số trung bình tuyệt đối với $\underline{O}$ được coi là giá trị tham chiếu. Giá trị $\underline{O}$ được coi là các giá trị dự báo khí hậu.

2.6.2. Phương pháp kiểm chứng chéo - OLOO

Đối với bài toán dự báo NBĐMM bằng phương pháp thống kê, việc đánh giá trên tập số liệu độc lập là cần thiết. Tuy nhiên do tập số liệu sẵn có không quá dài để có thể chia thành hai bộ số liệu phụ thuộc và độc lập. Do đó, luận án sẽ áp dụng

phương pháp OLOO để đánh giá chất lượng của các mô hình dự báo và khảo sát sự ổn định của các mô hình dự báo NBDMM. Phương pháp OLOO là một trong những phương pháp thông dụng được áp dụng trong nhiều nghiên cứu trước đây [Shao và cs, 2010; Hur và Ahn, 2015; Agou và cs, 2019].

Về bản chất, phương pháp OLOO là một trường hợp đặc biệt của kiểm chứng chéo. Trong đó, quá trình dự báo được lặp lại N lần, mỗi lần thực hiện với tập mẫu có kích thước là $N - 1$ (Hình 2.5). Giả sử tập mẫu có N năm, bỏ ra năm đầu tiên, dùng số liệu của $N-1$ năm còn lại để xây dựng phương trình dự báo. Sau đó sử dụng phương trình tìm được để dự báo cho năm đầu tiên bỏ ra đó. Thủ tục được lặp lại như vậy cho các năm tiếp theo, khi lần lượt từng năm một cho đến năm thứ N . Như vậy, kết thúc N lần lặp, ta sẽ thu được bộ gồm N giá trị dự báo độc lập tương ứng với N phần tử bỏ ra (không tham gia xây dựng phương trình dự báo) và N bộ hệ số hồi quy khi thay đổi tập số liệu luyện.



Hình 2.5: Sơ đồ mô tả phương pháp kiểm chứng chéo -OLOO

Nhận xét chung:

Về số liệu: luận án sử dụng 4 bộ số liệu, bao gồm: 1) Số liệu quan trắc; 2) Các chỉ số khí hậu; 3) Các trường khí áp mực biển (PRMSL), gió vĩ hướng mực 850mb

(U850) và nhiệt độ mặt nước biển (SST) trung bình tháng 3 và tháng 7; giá trị mưa hàng ngày của sản phẩm dự báo lại từ mô hình CFSv2; 4) Giá trị mưa hàng ngày được dự báo từ mô hình S2S-ECMWF.

Về phương pháp:

- *Đối với việc xác định NBĐMM từ số liệu quan trắc*, luận án đưa ra phương pháp lựa chọn chỉ tiêu khách quan xác định NBĐMM cho từng trạm bằng phương pháp MPCA trong đó đã kết hợp vai trò của các nhân tố địa phương và ảnh hưởng của quy mô vùng khí hậu. Bên cạnh đó, sự ảnh hưởng của các tham số trong chỉ tiêu xác định NBĐMM tại từng trạm cũng được xem xét thông qua việc áp dụng kiểm nghiệm f. Xu thế thay đổi của NBĐMM trong giai đoạn 1979-2019 sẽ được phân tích dựa vào kiểm nghiệm Man-Kendall và hệ số góc Sen. Kết quả xác định NBĐMM tại các trạm sẽ được sử dụng để xác định NBĐMM cho vùng khí hậu tương ứng thông qua giá trị đại diện là trung vị (phương pháp được trình bày ở mục 2.2.2). NBĐMM tại từng vùng khí hậu xác định được từ số liệu mưa ngày quan trắc tại trạm sẽ được coi là NBĐMM quan trắc và được sử dụng như là yếu tố dự báo trong phương pháp dự báo thống kê và được dùng làm tham chiếu để đánh giá chất lượng dự báo.
- *Đối với việc dự báo NBĐMM*, luận án sẽ xem xét và đánh giá khả năng dự báo với 4 trường hợp khác biệt về hạn dự báo và phương pháp dự báo, bao gồm: 1) Dự báo hạn mùa NBĐMM bằng phương pháp thống kê truyền thống dựa trên các chỉ số khí hậu; 2) Dự báo hạn mùa NBĐMM bằng phương pháp thống kê dựa trên các trường dự báo của mô hình CFSv2; 3) Dự báo hạn mùa NBĐMM sử dụng trực tiếp sản phẩm dự báo mưa của CFSv2; 4) Dự báo hạn nội mùa NBĐMM sử dụng trực tiếp sản phẩm dự báo mưa của S2S-ECMWF. Các kết quả dự báo sẽ được đánh giá thông qua các chỉ số thống kê thông dụng như ME, MAE, CORR, SS_{MAE} và phương pháp kiểm chứng chéo OLOO.

Từ các nguồn số liệu và phương pháp nêu trên, Chương 3 và Chương 4 dưới đây sẽ lần lượt trình bày các kết quả của luận án tương ứng với 2 nội dung chính là đặc điểm và khả năng dự báo hạn nội mùa và hạn mùa của NBĐMM trên khu vực Việt Nam.

Chương 3. ĐẶC ĐIỂM CỦA NGÀY BẮT ĐẦU MÙA MƯA TRÊN KHU VỰC VIỆT NAM

Trong chương này, luận án sẽ khảo sát tính bất định của NBĐMM do việc lựa chọn chỉ tiêu thông qua chỉ số SI được áp dụng cho từng trạm với 42 bộ giá trị khác nhau của các tham số P, N, và D. Phân bố không gian của SI sẽ cung cấp một bức tranh tổng thể tương đối rõ ràng về các trạm/khu vực mà việc lựa chọn chỉ tiêu phù hợp để xác định NBĐMM cần được quan tâm.

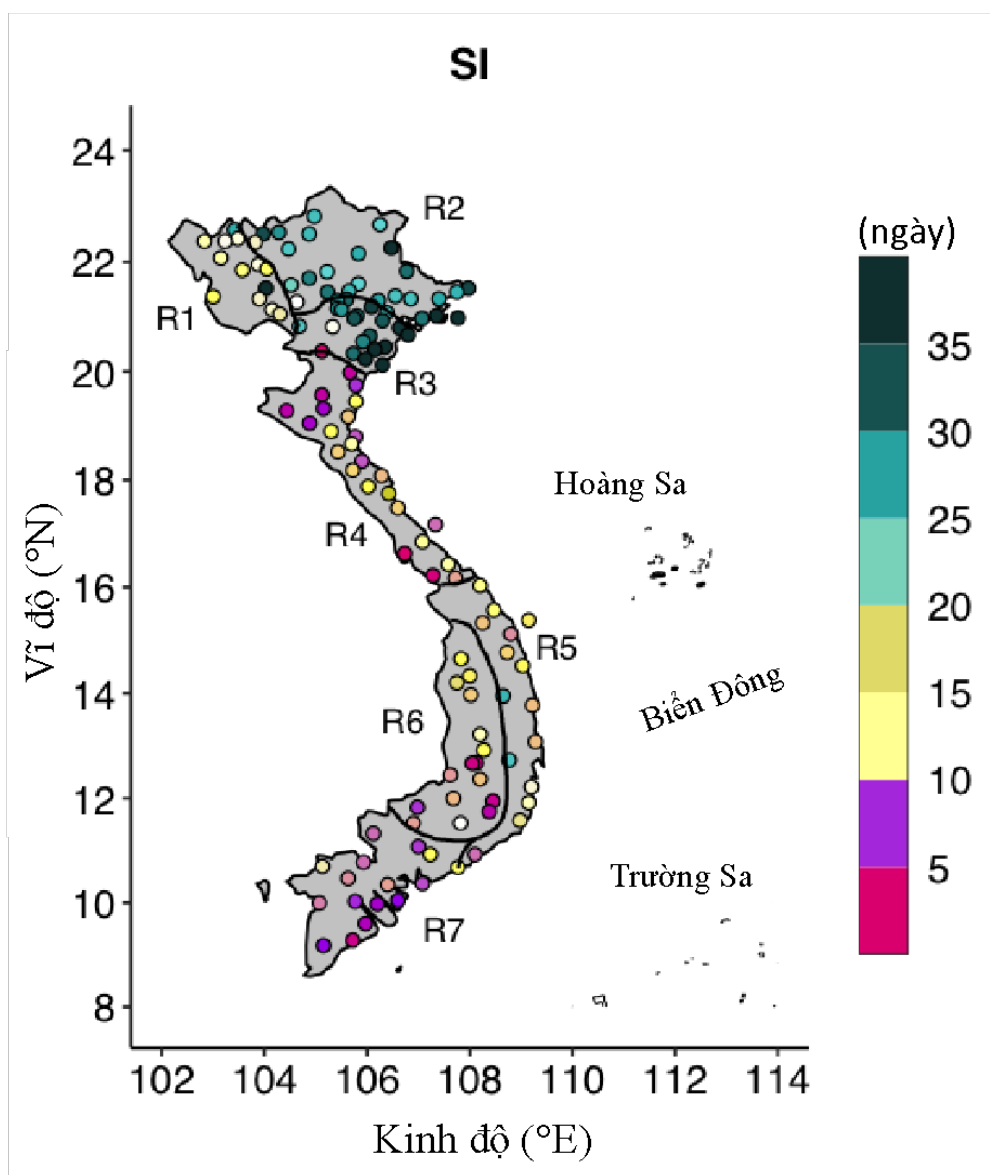
Bên cạnh đó, dựa trên tiêu chí về sự tương đồng giữa NBĐMM tại mỗi trạm (tín hiệu địa phương) và tín hiệu NBĐMM trên vùng khí hậu tương ứng, bộ chỉ tiêu xác định NBĐMM cho từng trạm trên khu vực Việt Nam cũng được đưa ra. Dựa vào chỉ tiêu được đề xuất, các đặc trưng cơ bản, xu thế biến đổi và mối quan hệ với hiện tượng ENSO của NBĐMM tại 7 vùng khí hậu cũng sẽ được khảo sát và trình bày.

3.1. TÍNH BẤT ĐỊNH CỦA NGÀY BẮT ĐẦU MÙA MƯA XÁC ĐỊNH CHO TỪNG TRẠM

Cũng như một số hiện tượng khác, NBĐMM là một thời điểm ước lệ và trên thực tế hầu như không thể nói chính xác là một ngày cụ thể. Vì vậy giá trị xác định được của NBĐMM phụ thuộc vào việc lựa chọn bộ chỉ tiêu quy ước và vì vậy nó sẽ khác nhau khi sử dụng các chỉ tiêu khác nhau. Hình 3.1 cho thấy tính bất đồng nhất của chỉ số SI tại 131 trạm trên khu vực Việt Nam. Nhìn chung, ngưỡng giá trị SI cao trên 10 ngày xuất hiện phổ biến, cụ thể tại 90/131 trạm (68.7%). Đáng chú ý, tại hầu khắp các trạm thuộc R2 và R3, SI nhận giá trị phổ biến trong khoảng 20-35 ngày. Trong khi đó, các giá trị SI thấp (khoảng 5-10 ngày) chỉ được quan sát thấy tại phía bắc của R4, phía nam của R6 và toàn bộ R7.

Về mặt ý nghĩa, độ lớn của chỉ số SI thể hiện độ lệch trung bình tuyệt đối giữa NBĐMM xác định từ 42 bộ chỉ tiêu thử nghiệm cho từng trạm (xem phần 2.2.1). Do đó, giá trị SI cao (thấp) đồng nghĩa với việc thay đổi các bộ giá trị P, N và D sẽ ảnh hưởng rõ rệt (không rõ rệt) tới kết quả xác định NBĐMM. Như vậy, bản đồ phân bố giá trị SI (Hình 3.1) đã cung cấp một bức tranh tổng thể, rõ ràng về các khu vực mà

việc lựa chọn chỉ tiêu phù hợp để xác định NBDMM cần được xem xét cẩn thận hơn. Thêm vào đó, sự khác biệt của chỉ số SI giữa các trạm trong cùng một vùng khí hậu cho thấy việc lựa chọn chỉ tiêu phù hợp nên được quan tâm không chỉ cho từng vùng mà còn cho cả từng trạm.



Hình 3.1: Phân bố không gian giá trị của chỉ số SI tại 131 trạm quan trắc trên khu vực Việt Nam. Đường màu đen thể hiện ranh giới giữa các vùng khí hậu

3.2. XÁC ĐỊNH CHỈ TIÊU NGÀY BẮT ĐẦU MÙA MƯA

3.2.1. Tín hiệu vùng của ngày bắt đầu mùa mưa

Kết quả phân tích MPCA cho thấy phần trăm đóng góp của bốn thành phần chính đầu tiên (PC1 đến PC4) vào phương sai tổng cộng (Bảng 3.1). Tại cả 7 vùng khí hậu, mức độ đóng góp thông tin của PC1 là lớn hơn một cách đáng kể khi so sánh với PC2, PC3 và PC4. Phần trăm đóng góp vào phương sai tổng cộng của PC1 có thể lên tới 43%, trong khi đó giá trị này đối với PC2 là dưới 15%, với PC3 là dưới 12% và dưới PC4 là 11%. Kiểm nghiệm North-test [North và cs, 1982] cho thấy tại cả 7 vùng khí hậu chỉ có PC1 là thành phần chính duy nhất thoả mãn ý nghĩa thống kê. Do đó, tín hiệu vùng của NBĐMM sẽ được đại diện bởi PC1. Như vậy, sự biến động hàng năm của NBĐMM quy mô vùng sẽ được mô tả bởi biến trình theo chuỗi thời gian của thành phần chính thứ nhất (EOF1).

Bảng 3.1: Phần trăm phương sai được giải thích bởi PC1-4 tại 7 vùng khí hậu

Vùng	Số trạm/vùng	PC1	PC2	PC3	PC4
R1	12	35	14	12	9
R2	34	37	10	9	6
R3	24	43	9	8	6
R4	24	24	15	10	8
R5	17	16	14	12	11
R6	14	28	14	11	9
R7	14	37	15	12	8

Về mặt độ lớn, phần trăm đóng góp của PC1 vào phương sai tổng cộng của PC1 có sự khác biệt rõ rệt giữa các vùng khí hậu. Lượng thông tin được giải thích bởi PC1 đạt giá trị lớn nhất tại R3 (43%) và thấp nhất tại R5 (16%). Các giá trị cao (thấp) sẽ phản ánh phần nào mức độ tương đồng lớn (nhỏ) về NBĐMM được xác định bằng 42 chỉ tiêu dự tuyển giữa các trạm trong cùng một vùng. Một cách tổng quát, tại miền khí hậu phía Bắc (R1, R2 và R3), và phía Nam (R6 và R7), phần trăm đóng góp của PC1 đạt trên 35%. Trong khi đó, đối với miền Trung (R4 và R5), các giá trị này nhỏ

hơn 28%. Như vậy, thông qua kết quả phân tích MPCA có thể thấy phần nào sự bất đồng nhất lớn theo không gian của NBĐMM giữa các trạm trong vùng khí hậu R4 và R5.

3.2.2. Chỉ tiêu ngày bắt đầu mùa mưa dựa trên tín hiệu vùng và trạm

Hình 3.2d dẫn ra hệ số tương quan giữa NBĐMM tại từng trạm (quy mô địa phương) được xác định bằng bộ chỉ tiêu được lựa chọn khách quan theo phương pháp mà luận án đã đề xuất ở mục 2.2.1 và giá trị của EOF1 đại diện cho tín hiệu vùng của NBĐMM tại vùng khí hậu tương ứng. Nhìn chung, bằng kiểm nghiệm t ($\alpha = 0.05$) mối quan hệ chặt chẽ về NBĐMM giữa hai quy mô đã được tìm thấy tại 116/131 trạm (88.5%). Các trạm có giá trị hệ số tương quan thấp dưới ngưỡng thoả mãn mức độ ý nghĩa thống kê chủ yếu tập trung rải rác tại khu vực miền núi thuộc R2, R6, và tương đối dày đặc tại khu vực R5. Việc 88.5% tổng số trạm có NBĐMM quan hệ chặt chẽ với tín hiệu vùng của NBĐMM tại vùng tương ứng cho thấy tính hợp lý của phương pháp được đề xuất để lựa chọn một cách khách quan chỉ tiêu phù hợp để xác định NBĐMM cho từng trạm.

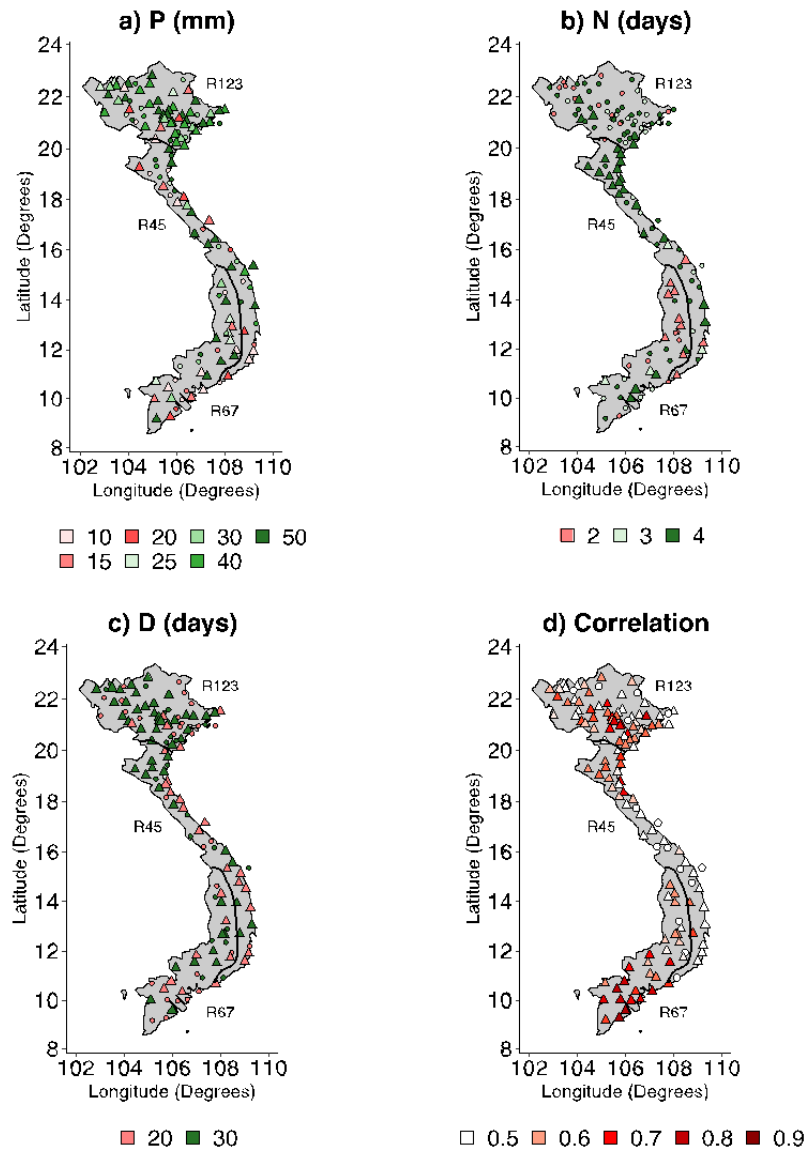
Bảng 3.2: Một số chỉ số thống kê thể hiện mức độ tương đồng giữa NBĐMM tại trạm với vùng khí hậu tương ứng

	Trung vị của CORR	Phần trăm số trạm có CORR thoả mãn mức độ ý nghĩa thống kê
R1	0.52	100
R2	0.50	85
R3	0.62	95
R4	0.51	84
R5	0.40	69
R6	0.58	89
R7	0.74	100

Phân bố theo không gian của hệ số tương quan có sự khác biệt rõ rệt giữa các vùng (Hình 3.2d, Bảng 3.2). Có thể thấy, tại các khu vực mà mùa mưa bị chi phối nhiều (ít) bởi hoạt động của GMMH như R1, R2, R3, R6 và R7 (R4 và R5), CORR có xu hướng nhận các giá trị cao (thấp), điều này được phản ánh thông qua phần trăm số trạm tại từng vùng khí hậu có giá trị CORR thoả mãn mức độ tin cậy của kiểm nghiệm t. Theo đó, có ít nhất 85% số trạm trên các vùng khí hậu chịu ảnh hưởng của gió mùa có giá trị CORR thoả mãn mức độ tin cậy 95%, trong khi, đối với các vùng khí hậu còn lại, tức là R4 và R5, giá trị này tương ứng chỉ 84% và 69%. Mức độ tương đồng giữa NBĐMM quy mô vùng và quy mô địa phương (trạm) cao nhất tại R7, với giá trị trung vị của CORR trong toàn vùng là 0.74 và 100% các trạm trong vùng có giá trị CORR thoả mãn mức độ tin cậy 95% của kiểm nghiệm t (Bảng 3.2). Trong khi đó, sự khác biệt về ngày NBĐMM quy mô vùng và quy mô địa phương (trạm) được thể hiện rõ nhất ở R5, với giá trị trung vị của hệ số tương quan trong toàn vùng là 0.40 và chỉ có 69% các trạm trong vùng có giá trị CORR thoả mãn mức độ tin cậy 95% của kiểm nghiệm t. Các kết quả trên cho thấy việc xác định NBĐMM nên được thực hiện trên quy mô địa phương (từng trạm) đối với khu vực miền Trung, đặc biệt là R5.

Hình 3.2a-c thể hiện vai trò của việc điều chỉnh các ngưỡng giá trị và ngưỡng giá trị phù hợp của từng tham số P, N và D đối với kết quả xác định NBĐMM cho từng trạm. Các kết quả này thu được dựa trên kiểm nghiệm F-test. Nhìn chung, vai trò của các tham số (P, N và D) đối với việc xác định NBĐMM giữa các vùng và trên toàn Việt Nam là khác nhau. Bảng 3.3 cho thấy trong ba tham số P, N và D thì việc điều chỉnh các ngưỡng giá trị của tham số N có ảnh hưởng ít nhất tới kết quả xác định NBĐMM trên hầu khắp các vùng khí hậu, trừ R4. Có thể thấy rằng, các vùng thuộc miền khí hậu phía bắc, trừ R4 (tức là R1, R2 và R3), phần trăm số trạm bị ảnh hưởng đáng kể đến giá trị NBĐMM khi chỉ tiêu xác định có sự điều chỉnh ngưỡng giá trị của tham số N chỉ dao động ở dưới ngưỡng 11.1%, đặc biệt con số này là 0% tại R1. Trong khi đó, đối với miền khí hậu phía nam, bao gồm R5, R6 và R7, tại mỗi vùng đều có ít nhất 31.3% số trạm có NBĐMM bị thay đổi đáng kể khi điều chỉnh ngưỡng

giá trị của tham số N trong chỉ tiêu. So sánh giữa 3 tham số, P đóng vai trò là tham số quan trọng nhất trong việc xác định NBĐMM trên khu vực Việt Nam. Dựa vào kiểm nghiệm F-test, sự ảnh hưởng của việc điều chỉnh ngưỡng giá trị của P làm thay đổi đáng kể kết quả xác định NBĐMM tại 87/131 trạm (66.4%). Đặc biệt, vai trò của tham số P càng trở nên rõ rệt hơn ở các vùng khí hậu phía bắc như R1, R2 và R3, với phần trăm số trạm bị ảnh hưởng dao động trong khoảng 74.1% đến 84.2% (Bảng 3.3). Tham số P cũng được coi là nhân tố quyết định đến việc xác định NBĐMM trên khu vực nhiệt đới gió mùa Châu Phi [Boyard-Micheau và cs, 2013]. Bên cạnh đó, 73/131 trạm (55.7%) trên khu vực Việt Nam có sự thay đổi đáng kể NBĐMM khi điều chỉnh giá trị tham số D. Vai trò của tham số D thực sự nổi bật tại R6, khi nó chi phối đáng kể giá trị NBĐMM tại 13/18 trạm (72.2%) thuộc vùng này. Như vậy, Hình 3.2a-c và Bảng 3.3 đã cung cấp một bức tranh đầy đủ và chi tiết về vai trò của từng tham số P, N và D trong việc xác định NBĐMM cho từng trạm, và trong từng vùng và trên toàn khu vực nghiên cứu. Đây là thông tin hữu ích vì nó cho biết tham số nào trong các tham số P, N và D cần được quan tâm trong quá trình lựa chọn, xây dựng chỉ tiêu xác định NBĐMM dựa trên đặc trưng mưa trên khu vực Việt Nam.



Combination of figure a, b, c is
the optimal 3-tuple values of P, N, and D at each station

Hình 3.2: (a-c) Phân bố theo không gian các giá trị tối ưu của P, N, và D để xây dựng bộ chỉ tiêu xác định NBĐMM quy mô địa phương (tại trạm) trên khu vực Việt Nam; (d) hệ số tương quan giữa tín hiệu quy mô khu vực của NBĐMM (EOF1) với NBĐMM quy mô địa phương xác định dựa trên bộ chỉ tiêu tối ưu được lựa chọn. Đường màu đen thể hiện ranh giới của 3 khu vực R123, R45 và R67. Các hình tam giác (vòng tròn) trong hình (a-c) chỉ ra các trạm mà việc thay đổi các ngưỡng giá trị của điều kiện có ảnh hưởng có (không có) ý nghĩa thống kê, dựa trên kiểm nghiệm f với mức độ tin cậy 95%. Các hình tam giác (vòng tròn) trong hình (d) thể hiện mối quan hệ tương quan có (không có) ý nghĩa thống kê, dựa trên kiểm nghiệm t với mức độ tin cậy 95%.

Bảng 3.3: Phần trăm (%) các trạm trên từng vùng mà việc điều chỉnh các ngưỡng giá trị của tham số P , N , và D làm thay đổi đáng kể giá trị NBĐMM (dựa trên kiểm nghiệm f với độ tin cậy 95%). Giá trị được gạch chân và in đậm thể hiện tham số có ảnh hưởng lớn nhất (tức là có giá trị phần trăm lớn nhất) và nên được quan tâm đối với từng vùng và toàn bộ khu vực Việt Nam

	P	N	D
R1	<u>76.9</u>	0.0	61.5
R2	<u>74.1</u>	11.1	55.6
R3	<u>84.2</u>	10.5	57.9
R4	52.0	<u>68.0</u>	48.0
R5	<u>53.8</u>	46.2	<u>53.8</u>
R6	61.1	44.4	<u>72.2</u>
R7	<u>62.5</u>	31.3	43.8
VN	<u>66.4</u>	31.3	55.7

Bảng 3.4 cho thấy tần suất xuất hiện các ngưỡng giá trị tương ứng với các tham số P , N , D cho từng vùng khí hậu và cho toàn khu vực Việt Nam. Đối với P - tham số được coi là quan trọng nhất trong việc xác định NBĐMM trên hầu khắp các trạm thuộc Việt Nam, việc lựa chọn ngưỡng giá trị cao cho tham số ngày ($P > 25\text{mm}$) chiếm ưu thế tại 6/7 vùng khí hậu (trừ R7), cụ thể tại 83/131 trạm (63.4%). Các ngưỡng $P > 25\text{mm}$ tập trung nhiều tại các trạm thuộc khu vực phía bắc, với mật độ cao lên tới trên 80% như ở vùng R2 (88.9%) và R3 (89.5%). Đáng lưu ý, các trạm thuộc khu vực núi cao cũng là nơi mà $P > 25\text{mm}$ là phù hợp hơn so với $P \leq 25\text{mm}$ (Hình 3.1a). Trong khi đó, ngưỡng $P \leq 25\text{mm}$ lại xuất hiện phổ biến trên vùng R7 với 81.3% số trạm. Sự phân bố bất đồng nhất về mặt giá trị của tham số P có thể liên quan đến sự khác biệt về cường độ mưa giữa các trạm, giữa các vùng khí hậu trong giai đoạn chuyển giao từ mùa khô sang mùa mưa.

Đối với N, tham số phản ánh mức độ ổn định của đợt mưa đầu tiên của mùa mưa, ngưỡng $N=4$ chiếm ưu thế ở 5/7 vùng khí hậu trừ các vùng R6 và N7 (Bảng 3.4). Có thể thấy, tham số N có xu thế nhận các giá trị >3 tại các vùng khí hậu phía Bắc và các giá trị ≤ 3 tại các vùng khí hậu phía Nam. Đáng chú ý, 84% số trạm thuộc vùng R4 lựa chọn $N=4$ trong chỉ tiêu xác định NBĐMM. Điều này có thể liên quan đến sự ngắt quãng của mùa mưa do thời kỳ Foehn khô nóng diễn ra vào tháng 6-7, nên đối với R4, mùa mưa chỉ được coi thực sự bắt đầu khi số ngày có mưa trong giai đoạn chuyển tiếp phải đủ dài (điều kiện $N=4$).

Đối với tham số D, ngưỡng $D=30$ có phần phù hợp với đa số các trạm, cụ thể, ngưỡng giá trị này xuất hiện tại 77/131 trạm (58.8%) trên toàn Việt Nam. Đặc biệt, đối với các trạm thuộc miền khí hậu phía Bắc (R1-R4), giá trị $D=30$ là phù hợp hơn cho chỉ tiêu xác định NBĐMM tại từng trạm, với tần suất trên 60% số trạm trong từng vùng. Như vậy, đối với các vùng này, việc lựa chọn ngưỡng $D=30$ là hợp lý do nó có khả năng loại bỏ được các đợt mưa trái mùa xảy ra trên quy mô cục bộ trong khoảng thời gian trước thời kỳ hoạt động của GMMH [Nguyễn Đức Ngữ và Nguyễn Trọng Hiệu, 2004]. Trong khi đó, ngưỡng $D=20$ phân bố chủ yếu ở R5 và R7 với tần suất xuất hiện tương ứng là 61.5% và 75.0% số trạm trong vùng (Bảng 3.4, Hình 3.2c). Việc ngưỡng thấp $D=20$ là phù hợp hơn khi xác định NBĐMM ở R5 và R7 có thể liên quan tới thời điểm bắt đầu mùa mưa trên hai vùng này. Đó là lúc mà các nhân tố gây mưa như GMMH đối với R7, và ITCZ đối với R5 có hoạt động ổn định [Nguyễn Hương Điền và Trần Công Minh, 2000; Nguyễn Đức Ngữ và Nguyễn Trọng Hiệu, 2004; Phạm và cs, 2010; Nguyen-Le và cs, 2014]. Điều này khiến biến trình mưa ngày trong giai đoạn chuyển tiếp ít chịu ảnh hưởng bởi các đợt giảm mưa bất thường tại R5 và R7.

Bảng 3.4: Tần suất xuất hiện của các ngưỡng giá trị tương ứng với các tham số P, N, D cho từng vùng khí hậu và cho toàn khu vực Việt Nam. Giá trị được gạch chân và in đậm thể hiện giá trị có tần suất xuất hiện lớn nhất trong số các giá trị tham chiếu được lựa chọn cho từng ngưỡng P, N, và D tại từng vùng khí hậu và toàn khu vực Việt Nam

Tham số	Ngưỡng giá trị	R1	R2	R3	R4	R5	R6	R7	VN
P	>25	<u>61.5</u>	<u>88.9</u>	<u>89.5</u>	<u>56.0</u>	<u>53.8</u>	<u>55.6</u>	18.7	<u>63.4</u>
	<=25	38.5	11.1	10.5	44.0	46.2	44.4	<u>81.3</u>	36.6
N	2	<u>46.1</u>	18.5	5.1	4.0	23.1	<u>61.1</u>	18.7	22.9
	3	7.7	22.2	31.6	12.0	23.1	5.6	<u>43.8</u>	20.6
	4	<u>46.2</u>	<u>59.3</u>	<u>63.2</u>	<u>84.0</u>	<u>53.8</u>	33.3	37.5	<u>56.5</u>
D	20	23.1	37.0	26.3	36.0	<u>61.5</u>	38.9	<u>75.0</u>	41.2
	30	<u>76.9</u>	<u>63.0</u>	<u>73.7</u>	<u>64.0</u>	38.5	<u>61.1</u>	25.0	<u>58.8</u>

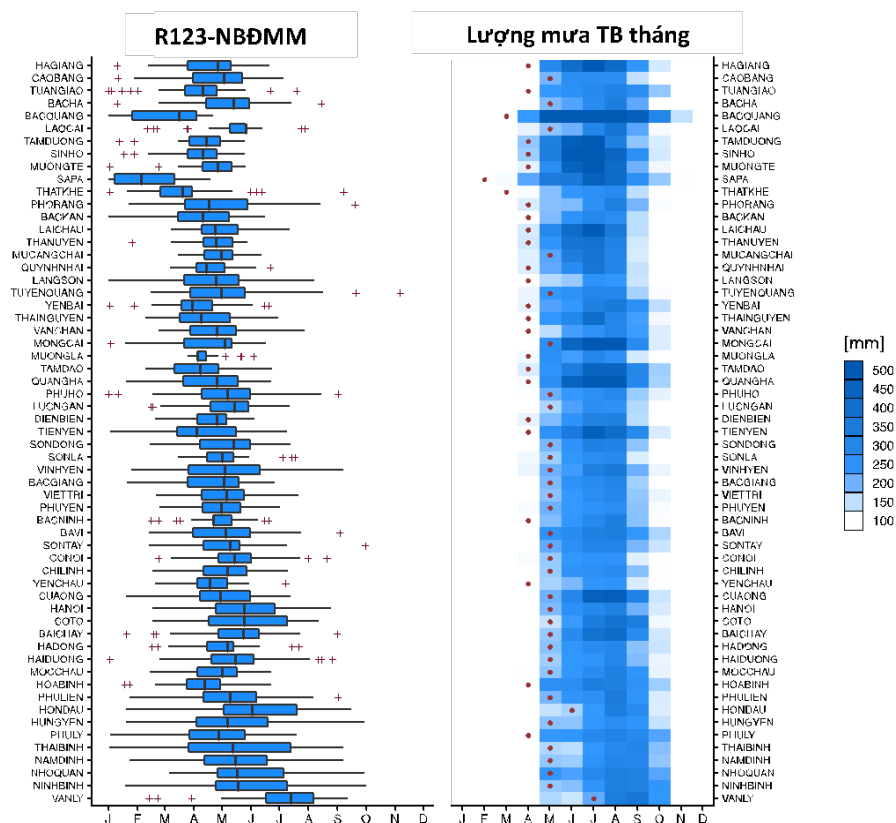
Như vậy, sự khác biệt về phân bố không gian của các giá trị tham số P, N, D đã phản ánh phần nào sự bất đồng nhất về đặc trưng mưa trong giai đoạn chuyển mùa giữa các trạm trong khu vực nghiên cứu.

3.3. PHÂN BỐ KHÔNG GIAN VÀ XU THẾ BIẾN ĐỔI CỦA NGÀY BẮT ĐẦU MÙA MƯA

Để đánh giá tính hợp lý của phương pháp tiếp cận mới được đề xuất trong luận án, NBĐMM tại 131 trạm trên khu vực Việt Nam trong giai đoạn 1979-2019 đã được xác định theo các chỉ tiêu mới. Biến trình mưa năm và NBĐMM trong các nghiên cứu trước đây sẽ được sử dụng làm thông tin tham chiếu về tính phù hợp của các kết quả nhận được. Để thuận tiện cho việc phân tích, luận án sẽ nhóm các vùng R1, R2 và R3 thành khu vực R123; các vùng R4 và R5 thành khu vực R45; các vùng R6 và R7 thành khu vực R67. Việc phân chia các vùng như trên được dựa trên khoảng cách

về mặt địa lý, sự tương đồng về biến trình mưa năm và NBĐMM tham khảo trong các nghiên cứu trước đây [Nguyễn Hường Điền và Trần Công Minh, 2000; Nguyễn Đức Ngữ và Nguyễn Trọng Hiệu, 2004; Pham và cs, 2010; Nguyen-Le và cs, 2014], và đặc điểm của một số trường hoàn lưu khí quyển (xem Phụ lục -Hình 1). Theo đó, các khu vực R123 và R67 được đặc trưng bởi mùa mưa gần như trùng với thời kỳ hoạt động mạnh của đới gió tây mực 850mb (xem Phụ lục - Hình 1). Tuy nhiên, thời điểm đới gió tây chiếm ưu thế là sớm hơn rõ rệt tại vùng R123 từ những tháng đầu năm (khoảng tháng 2), khi so sánh với vùng R67 (khoảng tháng 5). Đối với R45, trường mưa dường như ít chịu ảnh hưởng bởi hoạt động của gió mùa (trái ngược với R123 và R67), khi mà có sự suy giảm đáng kể vào thời kỳ đầu cũng như hoạt động mạnh của đới gió tây mực 850mb (xem Phụ lục - Hình 1). Điều này đã phản ánh phần nào sự khác biệt về đặc trưng của mùa mưa cũng như hình thái chi phối nó tại 7 vùng khí hậu trên khu vực Việt Nam.

Hình 3.3, 3.4 và 3.5 dẫn ra kết quả tính biến trình năm của tổng lượng mưa tháng (phải) và đồ thị hộp của NBĐMM (trái) tại từng trạm. Theo Nguyễn Đức Ngữ và Nguyễn Trọng Hiệu (2004), mùa mưa được coi là thời gian (số tháng) liên tiếp trong năm mà tổng lượng mưa tháng đạt từ 100mm trở lên. Theo đó, biến trình mưa hàng năm ở Việt Nam có thể chia thành hai dạng: ở khu vực R123 và R67 mùa mưa trùng với mùa GMMH (tháng 5 đến tháng 10), trong khi ở R45 mùa mưa dịch chuyển về các tháng cuối mùa hè (thu) và đầu mùa đông (tháng 8 đến tháng 12) (Hình 3.3-3.5). Như vậy, Hình 3.3, 3.4 và 3.5 (phải) đã chỉ ra sự phù hợp giữa NBĐMM dựa trên các chỉ tiêu mới và tháng bắt đầu mùa mưa trên cả quy mô khu vực và quy mô địa phương (tại trạm). Cụ thể, trung bình nhiều năm NBĐMM tại hầu hết các trạm thuộc khu vực R123 và R67 xảy ra từ giữa tháng 4 đến giữa tháng 5, trong khi đó, mùa mưa tại khu vực R45 lại diễn ra muộn hơn từ đầu tháng 7-tháng 9 (Hình 3.3, 3.5).



Hình 3.3: (Trái) Đồ thị hình hộp của NBĐMM xác định bằng bộ chỉ tiêu mới tại các trạm trên khu vực R123 (R1, R2, R3). Dấu (+) thể hiện các giá trị lớn hơn (nhỏ hơn) $1.5 * IQR$. (Phải) Lượng mưa trung bình hàng tháng với các chấm đỏ thể hiện tháng tương ứng của NBĐMM được xác định bằng chỉ tiêu mới

Mặc dù vậy, mùa mưa xuất hiện sớm nhất ở khu vực R123, cụ thể là tại vùng R1 và R2 với giá trị trung bình nhiều năm của NBĐMM là 23/4 (Hình 3.6a, Bảng 3.5). Kết quả này là tương đồng với các nghiên cứu trước đây. Thực tế, NBĐMM có thể xuất hiện sớm nhất tại các trạm nằm ở phía tây bắc của khu vực R123 (Hình 3.3 và 3.6a) vào cuối tháng 3 [Nguyen-Le và cs, 2014]. Việc mùa mưa xuất hiện sớm trước thời kỳ thịnh hành của GMMH trên khu vực bán đảo Đông Dương có thể liên quan đến sự xuất hiện của các front lạnh gây mưa lớn cục bộ ở miền Bắc Việt Nam và phía đông nam Trung Quốc [Wang và LinHo, 2002; Nguyễn Đức Ngữ và Nguyễn Trọng Hiệu, 2004]. Đây cũng là nguyên nhân giải thích cho sự biến động của NBĐMM giữa các năm tại các vùng thuộc R123 là lớn rõ rệt hơn (trên 21 ngày) so với R45 và R67 (Bảng 3.5, Hình 3.6). Trong 7 vùng khí hậu, sự biến động của NBĐMM giữa các năm lớn nhất xuất hiện tại vùng R3 (lên tới 35 ngày). Nguyên

nhân có thể liên quan đến sự xuất hiện của các đợt mưa do gió mùa đông bắc trong giai đoạn chuyển mùa trước khi bắt đầu mùa mưa.

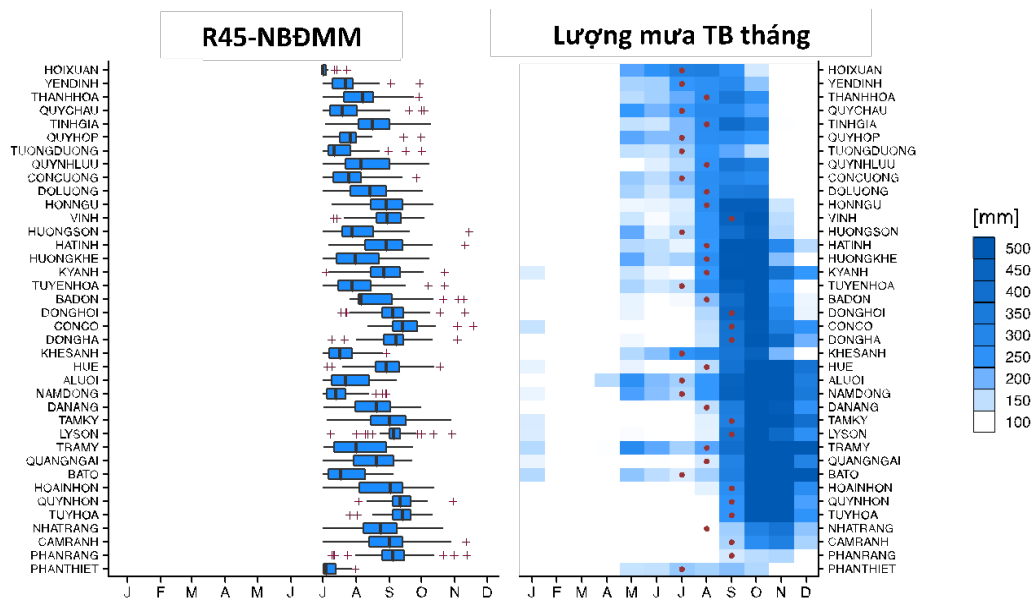
Bảng 3.5: Giá trị NBĐMM trung bình và độ lệch chuẩn (biến động giữa các năm) tại 7 vùng khí hậu trong giai đoạn 1979-2019

	Trung bình	Độ lệch chuẩn
R1	23/4	21
R2	23/4	23
R3	9/5	35
R4	8/8	16
R5	23/8	18
R6	26/4	12
R7	6/5	12

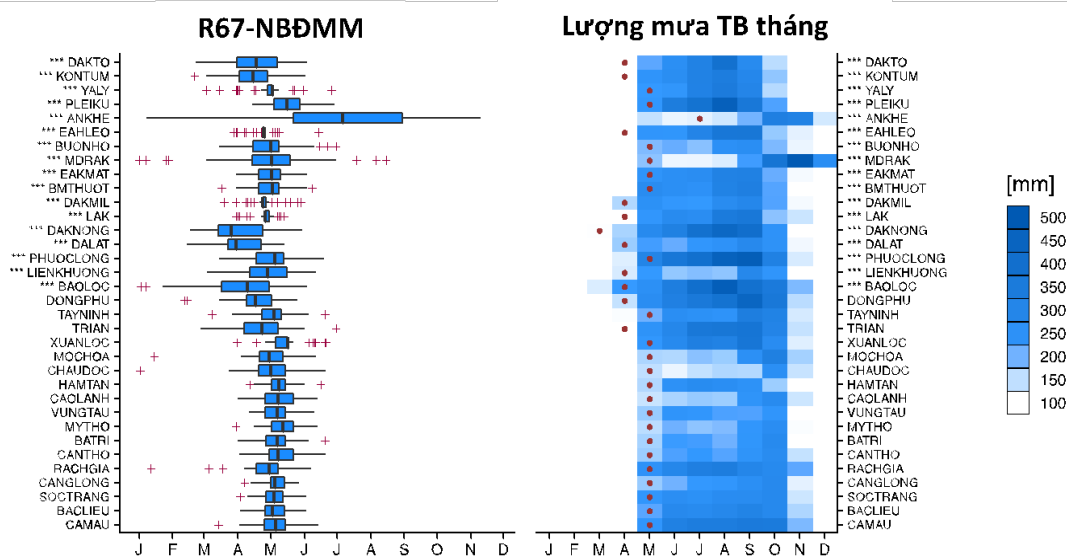
Đối với khu vực R67, NBĐMM có sự khác biệt giữa các vùng. Cụ thể, trung bình nhiều năm mùa mưa xuất hiện ở R6 vào ngày 26/4 và sớm hơn so với R7 (ngày 6/5) (Bảng 3.5). Đặc biệt, tại một số trạm ở phía nam của vùng R6, mùa mưa xuất hiện sớm hơn (từ đầu tháng 4), với tổng lượng mưa tháng vượt quá 100mm (Hình 3.5). Việc mùa mưa đến sớm hơn mùa GMMH từ 2~3 tuần đã được đề cập đến trong các nghiên cứu trước đây của Ngo-Thanh và cs (2018). Trong một số năm, mùa mưa xuất hiện sớm vào cuối tháng 3 - đầu tháng 4 (Hình 3.5), điều này khiến một số trạm tại R6 như BAOLOC, DAKNONG..., có sự biến động mạnh về NBĐMM giữa các năm. Cần lưu ý, một số trạm như ANKHE, MDRAK có vị trí nằm hơi lệch hoặc thuộc sườn đông dãy Trường Sơn Nam, chịu ảnh hưởng của mùa mưa vùng R5, và do đó có mùa mưa khác biệt hơn so với các trạm trong cùng vùng R6. Đối với vùng R7, phần lớn các trạm có mùa mưa xuất hiện vào cuối tháng 4-đầu tháng 5 (Hình 3.5) và gần như trùng với thời điểm hoạt động của GMMH [Nguyen-Le và cs, 2014; Zhang và cs, 2002]. Tuy nhiên, mùa mưa dường như đến sớm hơn ở phía tây của vùng vào cuối tháng 4, và lan dần sang phần phía đông vào đầu tháng 5 (Hình 3.7). Biến động giữa các năm của NBĐMM tại các trạm ở vùng R7 nhỏ hơn so với các trạm thuộc vùng R6 (Hình 3.4-trái). Tuy nhiên, khi xem xét trên toàn khu vực (R67) biến động

này là tương đồng nhau giữa R6 và R7 (khoảng 12 ngày, Bảng 3.5). Các kết quả này là phù hợp với những gì nhận được trong các nghiên cứu trước đây, trong đó các chỉ tiêu xác định NBĐMM là khác nhau nhưng đều dựa trên lượng mưa ngày tại trạm [Nguyễn Hương Điền và Trần Công Minh, 2000; Pham và cs, 2010; Nguyen-Le và cs, 2014; Phạm Thị Châm và cs, 2017; Lê Thị Xuân Lan, 2017; Ngo-Thanh và cs, 2018].

Đối với khu vực R45, mùa mưa tập trung vào các tháng mùa thu và NBĐMM có xu hướng xảy ra muộn hơn ở phần phía nam (vùng R5) so với phần phía bắc (vùng R4) (Hình 3.4). Điều này có thể liên quan đến sự dịch chuyển về phía nam của ITCZ. Trung bình nhiều năm trong giai đoạn 1979-2019, NBĐMM tại R4 là ngày 8/8, và tại R5 là ngày 23/8. Đáng chú ý, một số trạm ở phía bắc thuộc R4 có biến trình mưa năm được đặc trưng bởi hai mùa mưa: một mùa mưa phụ vào khoảng tháng 5 đến tháng 6 và một mùa mưa chính bắt đầu từ tháng 8 đến tháng 9 [Nguyen-Le và cs, 2014]. Do đó, sự khác biệt của NBĐMM trung bình giữa các trạm riêng lẻ trong khu vực R45 là khá lớn. Cụ thể, mùa mưa bắt đầu sớm nhất là ngày 1 tháng 7 tại Hồi Xuân nằm ở phía bắc R4, trong khi NBĐMM muộn nhất là ngày 13 tháng 9 tại Tuy Hòa nằm ở phía nam của R5 (Hình 3.4). Về mặt địa lý, khu vực R45 nằm trải dài theo hướng bắc nam (trên 10°) nhưng lại hẹp theo hướng đông tây. Hầu hết các trạm thuộc khu vực R45 đều nằm dọc theo bờ biển và ở phía sườn khuất gió của dãy Trường Sơn khi có hoạt động của GMMH. Do đó, sự xuất hiện không đồng thời của mùa mưa tại các trạm trong khu vực này không chỉ liên quan đến sự khác biệt về hoàn lưu khí quyển trong thời gian đầu của mùa mưa mà còn liên quan tới các đặc trưng mang tính chất địa phương. Kết quả là, sự biến động lớn theo không gian và thời gian của các sự kiện mưa trong thời kỳ chuyển tiếp từ mùa khô sang mùa mưa có thể được quan trắc thấy tại một số trạm thuộc khu vực R45. Điều này được phản ánh thông qua sự biến động lớn về NBĐMM. Các kết quả này hoàn toàn phù hợp với những kết luận rút ra từ việc phân tích MPCA đối với NBĐMM (xem mục 3.2, Bảng 3.1). Như vậy, có thể thấy rằng việc xác định NBĐMM cả quy mô địa phương (tại từng trạm) và quy mô vùng ở các vùng R4 và R5 là phức tạp và thách thức do đặc điểm mưa tại đây mang tính chất cục bộ nhiều.



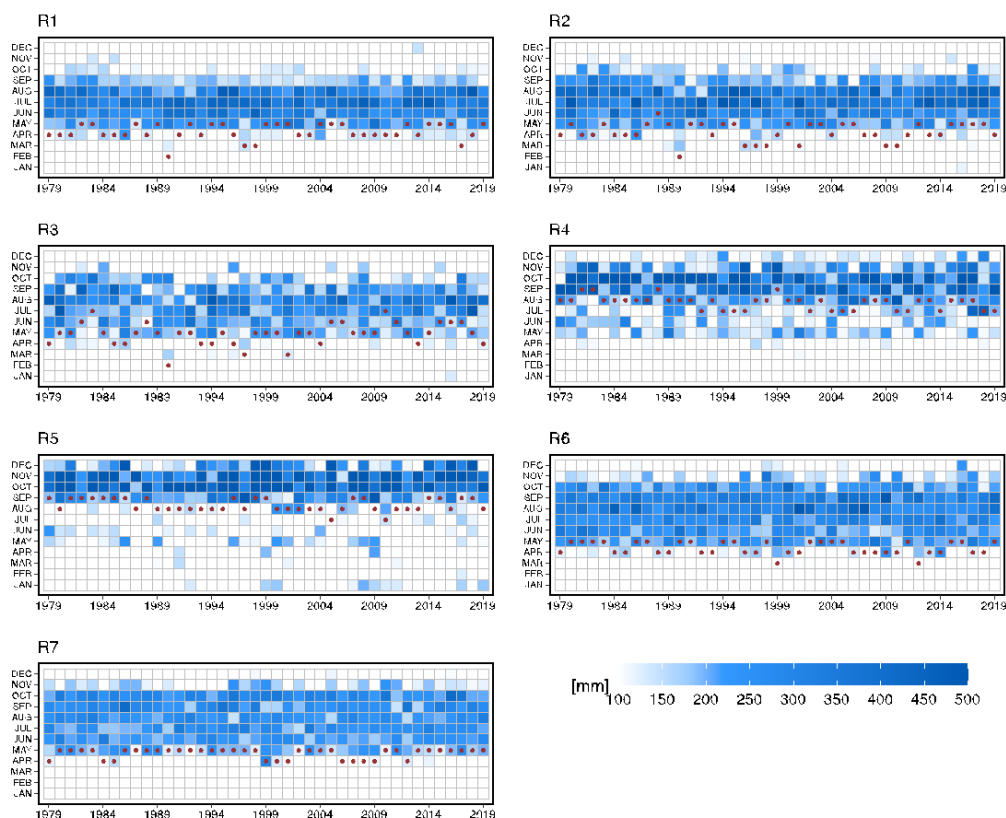
Hình 3.4: Tương tự hình 3.3 nhưng trên khu vực R45 (R4 và R5)



Hình 3.5: Tương tự Hình 3.3 nhưng trên khu vực R67 (R6 và R7)

Trên quy mô vùng, sự biến động hàng năm của NBĐMM trong giai đoạn 1979-2019 được thể hiện phần nào trên Hình 3.6. Nhìn chung, trên cả 7 vùng khí hậu, NBĐMM được xác định bằng chỉ tiêu mới đã nắm bắt tốt sự biến động hàng năm của tổng lượng mưa trong các tháng chuyển tiếp từ mùa khô sang mùa mưa, như tháng 4-5 đối với vùng R1, R2, R3, R6 và R7; tháng 7-9 đối với vùng R4 và R5. Sự biến động của NBĐMM giữa các năm có sự khác biệt giữa các vùng khí hậu. Cụ thể, NBĐMM

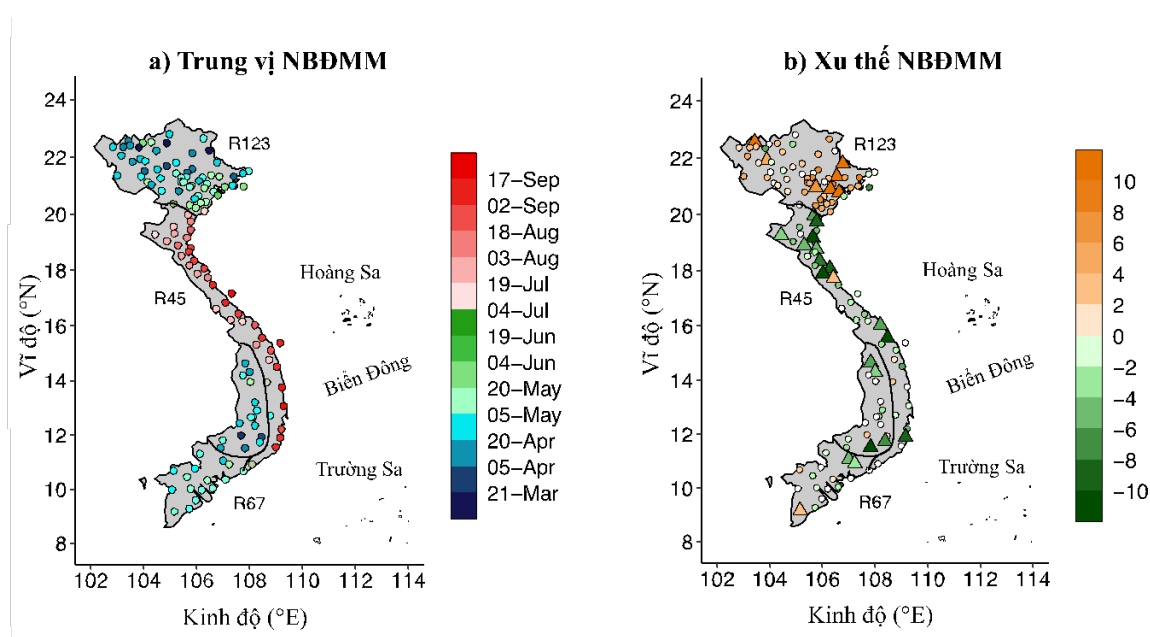
có phần ổn định hơn ở các vùng khí hậu R4-7 với độ lệch chuẩn dưới 18 ngày, trong khi đó, độ lệch chuẩn dao động trong khoảng 21 đến 35 ngày đối với các vùng khí hậu R1-3 (Bảng 3.5).



Hình 3.6: Tổng lượng mưa tháng trung bình tại 7 vùng khí hậu trong giai đoạn 1979-2019. Các chấm đỏ thể hiện tháng tương ứng của NBĐMM được xác định bằng chỉ tiêu mới

Xu thế biến đổi của NBĐMM trên khu vực Việt Nam được thể hiện ở Hình 3.7b. Nhìn chung, tại hầu khắp các trạm thuộc khu vực R45 và R67, mùa mưa có xu hướng bắt đầu sớm hơn một ít, khoảng 1-2 ngày/thập kỷ, nhưng xu thế này chỉ rõ rệt tại một số nơi. Tuy nhiên, các trạm ở phía bắc của R4 và ở phía đông nam của R6, NBĐMM có xu thế đến sớm hơn rõ rệt, từ 3 đến 8 ngày/thập kỷ với mức độ tin cậy 95%. Ngược lại, xu thế đến muộn của NBĐMM lại tập trung chủ yếu ở các trạm thuộc khu vực R123. Tuy nhiên, xu thế này chỉ thực sự rõ rệt tại một số trạm phía đông của R3 với mùa mưa có xu hướng bắt đầu muộn hơn khoảng 5-10 ngày/thập kỷ. Về cơ bản, các kết quả này tương đồng với xu thế thay đổi của ngày bắt đầu GMMH nhận được bởi Kajikawa và cs (2012). Cụ thể, khi cách so sánh hai giai đoạn 1979-

1993 và 1994-2008 các tác giả đã chỉ ra rằng ngày bắt đầu GMMH trên Vịnh Bengal, bán đảo Đông Dương và phía tây Thái Bình Dương xảy ra sớm hơn khoảng 10-15 ngày trong những thập kỷ gần đây. Ở phía bắc của R6, xu thế đến sớm của cả NBĐMM và mùa GMMH trong những năm gần đây cũng đã được chỉ ra bởi Ngo-Thanh và cs (2018). Xu thế đến sớm của NBĐMM là một tín hiệu tốt vì thời gian khô hạn, thiếu nước trong mùa khô có xu hướng kết thúc sớm hơn. Bên cạnh đó, các kết quả về sự biến động hàng năm lớn và xu thế biến đổi NBĐMM trong giai đoạn 1979-2019 cũng nhấn mạnh sự cần thiết của công tác dự báo NBĐMM đối với các lĩnh vực khác nhau của đời sống.



Hình 3.7: (a) Giá trị trung vị của NBĐMM; (b) xu thế thay đổi của NBĐMM trong giai đoạn 1979-2019 tại 131 trạm trên khu vực Việt Nam (ngày/thập kỷ).

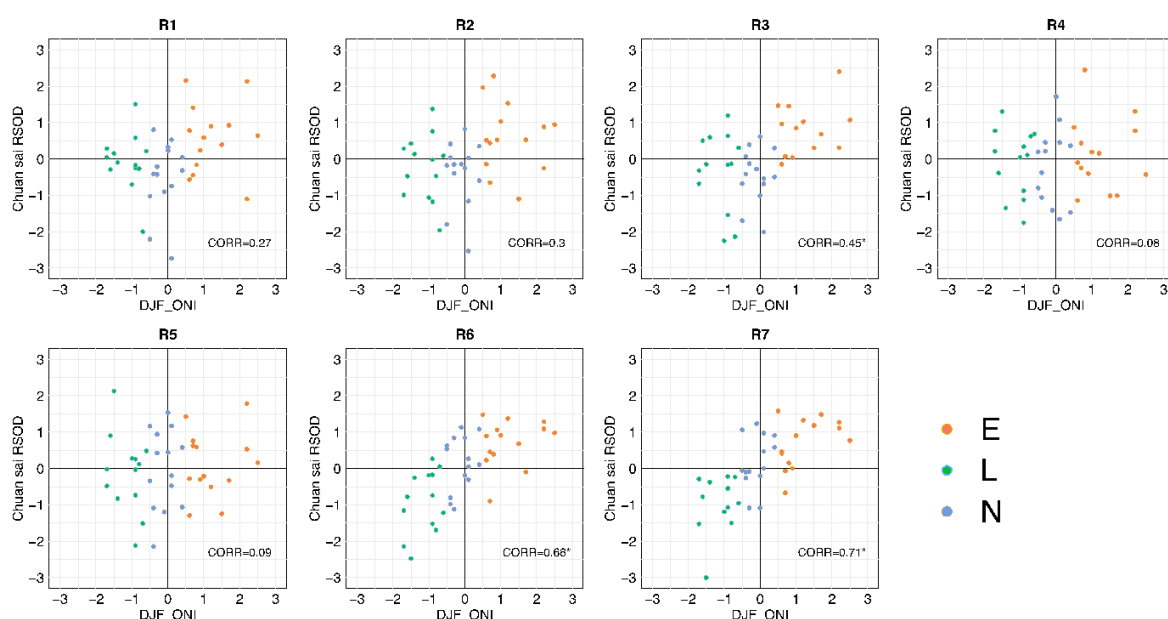
Hình tam giác (tròn) thể hiện trạm có xu thế thay đổi thỏa mãn (không thỏa mãn) mức độ tin cậy 95% của kiểm nghiệm Mann-Kendall. Đường viền màu đen thể hiện đường biên của các khu vực R123, R45 và R67

3.4. MỐI QUAN HỆ GIỮA NGÀY BẮT ĐẦU MÙA MƯA VÀ ENSO

Như đã đề cập ở mục 3.3, sự biến động giữa các năm của NBĐMM ở các vùng khí hậu trên khu vực Việt Nam là khác nhau. Ảnh hưởng của ENSO đến sự biến động của mưa và ngày bắt đầu GMMH tại khu vực Việt Nam đã được chỉ ra trong nhiều nghiên cứu trước đây [Lau và cs, 1997, Zhou và cs, 2007, Nguyễn Thị Hiền Thuận và Chiêu Kim Quỳnh, 2007, Nguyen và cs, 2007, Phạm Thị Châm và cs, 2017]. Do

đó, trong mục này, luận án sẽ khảo sát quan hệ giữa ENSO và NBĐMM trên các vùng khí hậu thông qua hệ số tương quan giữa trung bình ONI trong các tháng 12-1-2 và chuẩn sai NBĐMM.

Kết quả tính toán cho thấy ra mối quan hệ giữa NBĐMM với các pha của ENSO chỉ thực sự rõ rệt tại 3/7 vùng khí hậu là R3, R6, và R7. Điều này được thể hiện thông qua hệ số tương quan giữa trung bình ONI các tháng 12-1-2 và chuẩn sai NBĐMM đạt trên 0.45 và thoả mãn độ tin cậy thống kê 95% (t-test) (Hình 3.8). Phân bố không gian của hệ số tương quan cho thấy sự khác biệt rõ rệt về mối quan hệ giữa NBĐMM với ENSO giữa các vùng khí hậu. Các giá trị dương của hệ số tương quan xuất hiện tại các khu vực R123 và R67, trong khi các giá trị âm lại xuất hiện tại khu vực R45. Cụ thể, hệ số tương quan đạt trên 0.67 tại R67, khoảng 0.27-0.45 tại R123, và chỉ -0.09 đến -0.08 tại R45.

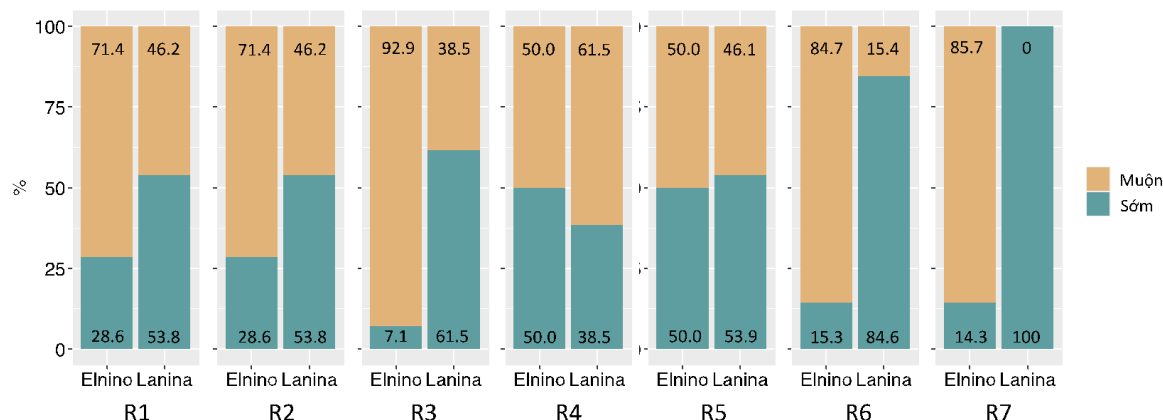


Hình 3.8: Đồ thị tụ điểm giá trị chuẩn sai NBĐMM và giá trị trung bình chỉ số ONI các tháng 12-1-2 trong giai đoạn 1979-2019 tại 7 vùng. Dấu * thể hiện giá trị hệ số tương quan thoả mãn mức độ ý nghĩa 95% (t-test)

Xác suất xuất hiện các pha (sớm/muộn) của NBĐMM trong những năm El Niño và La Niña tại 7 vùng khí hậu cũng được thể hiện trên Hình 3.8. Nhìn chung, NBĐMM có xu thế đến muộn hơn trong những năm El Niño và đến sớm hơn trong những năm La Niña, xu thế này đặc biệt rõ rệt tại các khu vực R123 và tại R45 - những nơi mà mùa mưa chịu ảnh hưởng rõ rệt bởi hoạt động GMMH. Quan hệ này

là tương đồng với mối quan hệ giữa ngày bắt đầu GMMH với ENSO đã được đề cập trong các nghiên cứu trước đây. Chẳng hạn, Lau và cs (1997) đã cho thấy sự xuất hiện muộn hơn (sớm hơn) của GMMH trên khu vực Biển Đông có mối liên hệ với sự nóng lên (lạnh đi) ở Thái Bình Dương và Ấn Độ Dương. Zhou và cs (2007) đã khảo sát mối liên hệ giữa ngày bắt đầu gió mùa Đông Nam Á (hay còn gọi là gió mùa Nam Hải, tức gió mùa Biển Đông) và ENSO khi sử dụng số liệu tái phân tích NCEP (Trung tâm dự báo môi trường Hoa Kỳ) và ECMWF (Trung tâm dự báo hạn vừa Châu Âu). Ngày bắt đầu gió mùa được xác định trên cơ sở gió vĩ hướng mực 850mb trên khu vực Biển Đông chuyển từ gió đông sang gió tây kéo dài liên tục 2 pentad. Kết quả nhận được chỉ ra rằng trong những năm thuộc pha nóng (lạnh) hoặc năm tiếp theo sự kiện ENSO gió mùa có xu hướng bắt đầu muộn hơn (sớm hơn) với cường độ yếu hơn (mạnh hơn). Nguyễn Thị Hiền Thuận và Chiêu Kim Quỳnh (2007) lại cho thấy NBĐMM ở Nam Bộ sẽ đến muộn hơn trong những năm El Niño và sớm hơn trong những năm La Niña. Khi nghiên cứu mối quan hệ giữa nhiệt độ bề mặt biển (SST) trên khu vực nhiệt đới Thái Bình Dương (28N-28S; 120E-85W) và Ấn Độ Dương (28N-28S; 30E-105E) với lượng mưa tháng ở Tây Nguyên tác giả Nguyen và cs (2007) cũng đã chỉ ra sự thay đổi của SST có ảnh hưởng rõ rệt đến ngày bắt đầu và kết thúc GMMH.

Như vậy, trên toàn khu vực Việt Nam, ảnh hưởng của ENSO đến sự xuất hiện (sớm/muộn) của NBĐMM được thể hiện rõ nhất ở trên các khu vực mà mùa mưa bị chi phối bởi hoạt động của gió mùa (R123 và R67). Trên hai khu vực này, mặc dù mức độ quan hệ giữa NBĐMM và chỉ số ONI trung bình các tháng 12-1-2 chỉ thể hiện sự chặt chẽ tại một số vùng (R3, R6 và R7) (Hình 3.9), nhưng khả năng sử dụng thông tin về các pha của ENSO làm “tín hiệu” dự báo NBĐMM trên tất cả các vùng thuộc R123 và R67 là hoàn toàn khả thi, đặc biệt là trong những năm El Niño. Điều đó được thể hiện thông qua xác suất NBĐMM đến muộn trong những năm El Niño tại các vùng thuộc khu vực R123 đạt trên 70%, cá biệt vùng R3 với xác suất lên tới 92.9%.



Hình 3.9: Đồ thị hình cột thể hiện xác suất xuất hiện các pha (sớm/muộn) của NBĐMM trong những năm El Niño và La Niña tại 7 vùng khí hậu

Nhận xét cuối chương 3

Khảo sát chỉ số SI cho thấy có 90/131 trạm trên khu vực Việt Nam (chiếm 68.7%) SI có giá trị trên 10 ngày. Điều này chứng tỏ việc lựa chọn chỉ tiêu phù hợp để xác định NBĐMM là quan trọng và cần thiết đối với hầu hết các trạm. Việc phân hoá của giá trị chỉ số SI giữa các vùng khí hậu và giữa các trạm trong cùng một vùng khí hậu cho thấy việc lựa chọn chỉ tiêu phù hợp nên được quan tâm trên cả quy mô vùng và trên cả quy mô địa phương (cho từng trạm).

Để lựa chọn một cách khách quan chỉ tiêu xác định NBĐMM phù hợp trên quy mô địa phương (tại trạm), luận án đã đề xuất và thử nghiệm một phương pháp mới dựa trên sự tương đồng giữa NBĐMM tại trạm và tín hiệu NBĐMM quy mô vùng (tại vùng khí hậu tương ứng) được xác định bằng phương pháp MPCA. Kết quả nhận được cho thấy các chỉ tiêu được lựa chọn nắm bắt tốt thời điểm bắt đầu mùa mưa trên cả quy mô vùng khí hậu và quy mô địa phương (trạm). NBĐMM xác định được cho khu vực Việt Nam phù hợp với các nghiên cứu trước đây.

Đối với việc lựa chọn chỉ tiêu xác định NBĐMM cho từng trạm, luận án đã đề xuất một phương pháp xác định các ngưỡng phù hợp cho 3 tham số P, N và D dựa trên sự tương đồng về biến động giữa các năm của NBĐMM tại từng trạm và NBĐMM tại từng vùng khí hậu tương ứng. Kết quả cho thấy, mỗi một tham số đóng vai trò khác nhau đến việc xác định NBĐMM tại từng trạm, từng vùng khí hậu và

cho toàn Việt Nam. Nhìn chung, trên toàn khu vực Việt Nam, tham số P là tham số quan trọng nhất với 66.4% số trạm bị ảnh hưởng đáng kể khi điều chỉnh ngưỡng giá trị của tham số này. Đây cũng là tham số quan trọng nhất tại 5/7 vùng khí hậu, trừ R4 và R6. Trong khi đó, chỉ 31.3% và 43.8% số trạm trên khu vực Việt Nam bị ảnh hưởng khi thay đổi ngưỡng của hai tham số N và D. Vai trò của hai tham số này chỉ nổi bật tại 1 số vùng: tham số N tại R4, và tham số D tại R6.

Bên cạnh đó, sự phân bố theo không gian của các ngưỡng tham số cũng có sự khác biệt giữa các vùng và giữa các trạm trong cùng một vùng. Cụ thể, ngưỡng $P > 25$ mm có xu hướng phổ biến ở 6/7 vùng khí hậu trừ (R7), đặc biệt tại khu vực miền núi (chiếm 63.4% số trạm trên khu vực Việt Nam). Đối với tham số D, ngưỡng cao hơn (=30 ngày) có xuất hiện ở 58.8% số trạm trên tổng 131 trạm, và phổ biến ở 5/7 vùng khí hậu (trừ R5 và R7). Đối với tham số N, ngưỡng N cao nhất (=4 ngày) phổ biến tại 56.5% số trạm và chiếm ưu thế tại 5/7 vùng khí hậu (trừ R6 và R7). Như vậy, có sự khác biệt về vai trò cũng như ngưỡng phù hợp của từng tham số để xác định NBĐMM giữa các vùng, và giữa các trạm trong cùng một vùng. Điều này phản ánh phần nào sự phức tạp về đặc điểm mưa giai đoạn chuyển giao từ mùa khô sang mùa mưa trên các vùng khí hậu Việt Nam.

NBĐMM tại từng trạm, và từng vùng khí hậu được xác định dựa trên bộ chỉ tiêu mới là phù hợp với số liệu mưa quan trắc. Kết quả thu được của luận án cũng phù hợp với những nghiên cứu trước đây. Theo đó, NBĐMM xảy ra sớm nhất ở phía tây bắc của khu vực R123 vào khoảng nửa đầu tháng 4 (trước thời kỳ GMMH thịnh hành). NBĐMM ở R123 và R67 xảy ra vào giữa tháng 4 đến giữa tháng 5. Tuy nhiên, sự biến động giữa các năm của NBĐMM tại R123 lớn hơn (khoảng 21-35 ngày) so với ở R67 (12 ngày). Đối với R45, NBĐMM có sự khác biệt lớn giữa các trạm với mùa mưa xuất hiện sớm hơn ở phía bắc, vào ngày 1 tháng 7 và muộn nhất vào ngày 13 tháng 9 ở phần phía nam. Sự khác biệt lớn này có thể liên quan đến điều kiện địa lý và cơ chế gây mưa trong thời kỳ chuyển tiếp từ mùa khô sang mùa mưa. Bên cạnh đó, sự biến động giữa các năm tại R45 cũng khá lớn với độ lệch chuẩn khoảng 16-18 ngày.

Xu thế biến đổi của NBĐMM trong giai đoạn 1979-2019 có sự khác biệt giữa các vùng khí hậu, và giữa các trạm trong cùng một vùng. Nhìn chung, mùa mưa có xu hướng bắt đầu sớm hơn, khoảng 1-2 ngày/thập kỷ, (nhưng không rõ rệt) tại hầu khắp các trạm thuộc khu vực R45 và R67. Tuy nhiên, tại một số nơi như phía bắc của R4 và ở khu vực phía đông nam của R6, ghi nhận xu thế đến sớm hơn rõ rệt của NBĐMM, từ 3-8 ngày/thập kỷ. Trong khi đó, xu thế đến muộn của NBĐMM lại tập trung chủ yếu ở các trạm thuộc khu vực R123, nhưng xu thế này chỉ thực sự rõ rệt với mức khoảng 5-10 ngày/thập kỷ tại một số trạm thuộc khu vực phía đông của R3.

Mối quan hệ đồng biến giữa ENSO và NBĐMM được tìm thấy tại R123 và R67 - nơi mà mùa mưa chịu sự chi phối mạnh mẽ của GMMH. Theo đó, mùa mưa có xu thế đến muộn (sớm) trong những năm El Niño (La Niña). Tuy nhiên, mối quan hệ này có sự khác biệt giữa các vùng khí hậu. Cụ thể, chỉ 3/7 vùng khí hậu có mối quan hệ chặt chẽ thỏa mãn mức độ tin cậy thống kê 95% là R3 (khoảng 0.45), R6 (khoảng 0.68) và R7 (khoảng 0.71). Mặc dù vậy, khả năng sử dụng thông tin về các pha của ENSO làm “tín hiệu” dự báo sự biến động của NBĐMM tại R123 và R67 là hoàn toàn khả thi, đặc biệt là trong những năm El Niño.

Chương 4. DỰ BÁO HẠN MÙA VÀ NỘI MÙA NGÀY BẮT ĐẦU MÙA MƯA TRÊN KHU VỰC VIỆT NAM

Chương 4 của luận án sẽ trình bày các nội dung dự báo hạn mùa và nội mùa NBĐMM cho từng vùng khí hậu Việt Nam với 4 trường hợp theo 2 cách tiếp cận:

Phương pháp thống kê:

- Dự báo hạn mùa NBĐMM bằng phương pháp thống kê truyền thống sử dụng các chỉ số khí hậu làm nhân tố dự báo (mục 4.1)
- Dự báo hạn mùa NBĐMM bằng phương pháp thống kê sử dụng các trường dự báo khí quyển, đại dương của mô hình CFSv2 (mục 4.2)

Dựa trên sản phẩm dự báo mưa của mô hình số:

- Dự báo hạn mùa NBĐMM sử dụng sản phẩm mưa dự báo của mô hình CFSv2 (mục 4.3)
- Dự báo hạn nội mùa NBĐMM sử dụng sản phẩm mưa dự báo của mô hình S2S-ECMWF (mục 4.4)

Đối với các mô hình thống kê, các phương trình dự báo được xây dựng với yếu tố dự báo là NBĐMM xác định được từ số liệu quan trắc cho từng vùng khí hậu tại Chương 3 (theo phương pháp nêu ở mục 2.2.2). Trong khi đó, NBĐMM dự báo từ mô hình số sẽ được xác định trực tiếp từ sản phẩm mưa dự báo của mô hình trên cơ sở các chỉ tiêu được xây dựng cho từng mô hình riêng biệt. Kết quả dự báo thu được sẽ được đánh giá khi so sánh với NBĐMM quan trắc.

4.1. DỰ BÁO HẠN MÙA NGÀY BẮT ĐẦU MÙA MƯA BẰNG PHƯƠNG PHÁP THỐNG KÊ TRUYỀN THỐNG SỬ DỤNG CÁC CHỈ SỐ KHÍ HẬU

Trong trường hợp này các phương trình dự báo được xây dựng với nhân tố dự báo là các chỉ số khí hậu mô tả sự thay đổi của hoàn lưu khí quyển và đại dương, yếu tố dự báo là NBĐMM tại các vùng khí hậu xác định từ số liệu quan trắc. Mô hình thống kê được sử dụng là mô hình hồi quy từng bước. Nhân tố dự báo được lựa chọn tự động từ tập 46 nhân tố dự tuyển của 23 chỉ số khí hậu (xem Bảng 2.1) cho 02 tháng:

- Tháng 1 và Tháng 2 đối với vùng khí hậu R1, R2, R3, R6 và R7
- Tháng 5 và Tháng 6 đối với vùng khí hậu R4 và R5

Như vậy, trong trường hợp này, hạn dự báo NBĐMM là 1-2 tháng. Kết quả thu được các phương trình dự báo NBĐMM cho 7 vùng khí hậu là:

$$NBĐMM_{(R1)} = 113.95 + 4.96 \times Ni\tilde{n}o3_{(jan)}$$

$$NBĐMM_{(R2)} = 110.41 + 6.01 \times PNA_{(feb)}$$

$$NBĐMM_{(R3)} = 133.02 + 33.88 \times Ni\tilde{n}o3_{(jan)} - 19.86 \times CENSO_{(jan)}$$

$$NBĐMM_{(R4)} = 223.00 + 3.34 \times TNI_{(jun)} + 4.98 \times EPAC85_{(jun)}$$

$$NBĐMM_{(R5)} = 233.56 + 2.68 \times TNI_{(jun)}$$

$$NBĐMM_{(R6)} = 118.41 + 15.36 \times Ni\tilde{n}o4_{(feb)}$$

$$NBĐMM_{(R7)} = 126.41 + 11.77 \times Ni\tilde{n}o4_{(feb)} - 7.17 \times CPAC85_{(feb)} + 2.35 \times SOI_{(feb)}$$

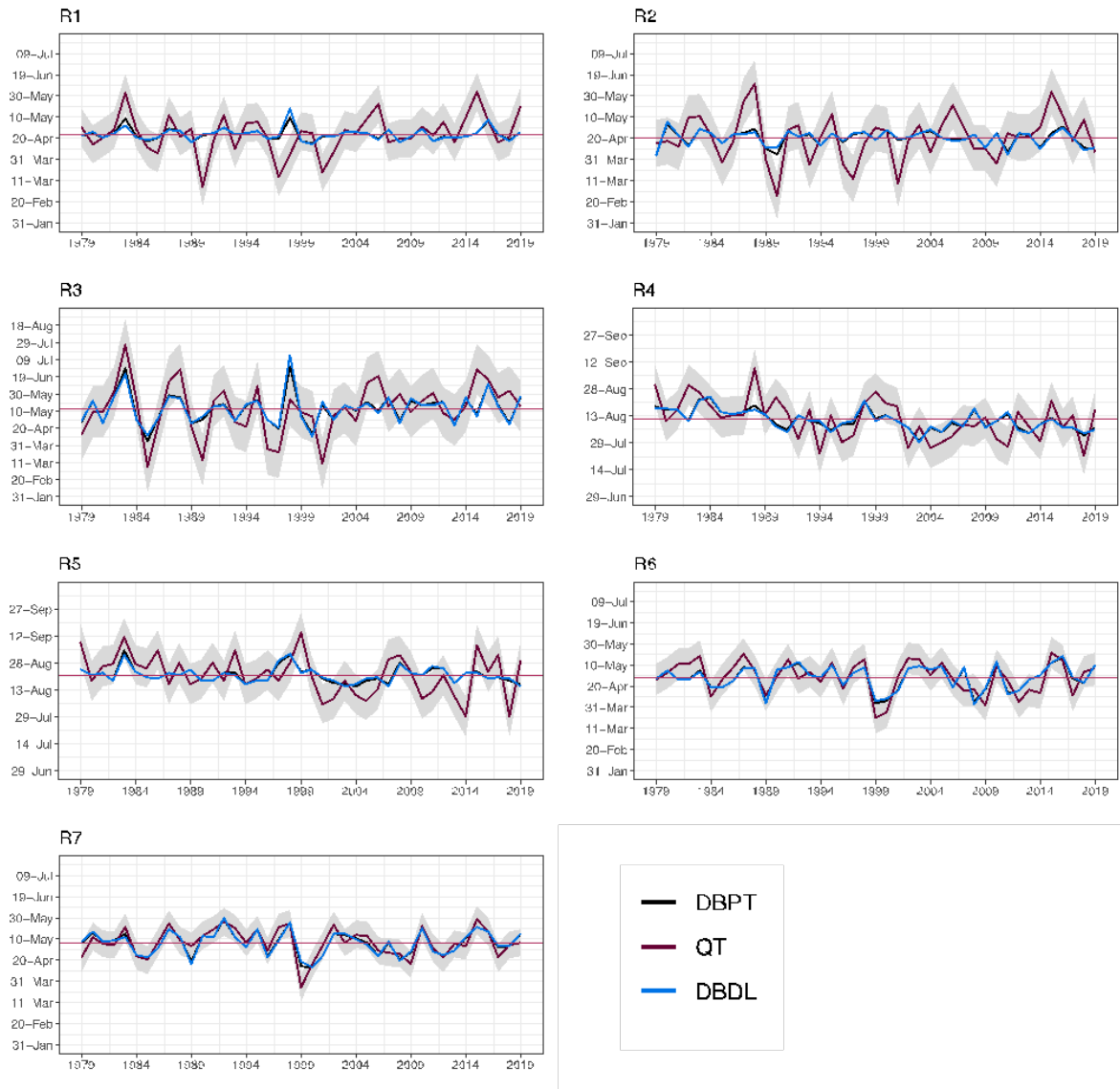
Nhìn chung, các phương trình dự báo NBĐMM cho từng vùng đều có giá trị hệ số tự do tương ứng với giá trị NBĐMM trung bình nhiều năm trong giai đoạn 1979-2019. Cụ thể, đối với vùng R1 là 113.95 (tương ứng ngày 25/4), với R2 là 110.41 (tương ứng với ngày 21/4), với R3 là 133.02 (tương ứng ngày 14/5), với R4 là ngày 223.00 (tương ứng với ngày 12/8), với R5 là ngày 233.56 (tương ứng ngày 22/8), với vùng R6 là 118.41 (tương ứng ngày 29/4), với R7 là 126.41 (tương ứng ngày 7/5).

Có thể thấy rằng có sự khác biệt về sự có mặt của các nhân tố dự báo trong các phương trình ở các vùng khí hậu. Điều đó thể hiện sự biến động hàng năm của NBĐMM tại các vùng khí hậu bị chi phối bởi các quá trình quy mô lớn khác nhau được phản ánh thông qua các chỉ số khí hậu. Mặc dù vậy, các chỉ số liên quan tới ENSO như Niño3, Niño4, TNI, CENSO, SOI có mặt ở hầu hết các phương trình dự báo NBĐMM ở các vùng khí hậu trên khu vực Việt Nam. Hệ số hồi quy ở các phương trình dự báo NBĐMM có cùng nhân tố dự báo đều nhất quán với nhau về dấu. Chẳng hạn, hệ số hồi quy trong các phương trình dự báo cho các vùng R1, R3, R6 và R7 đều

xuất hiện nhân tố Niño3 hoặc Niño4 là dị thường SST ở các vùng Niño3 và Niño4 đều có giá trị dương (+) có nghĩa là khi ENSO thiên về pha nóng (El Niño) thì NBĐMM có xu hướng đến muộn hơn và ngược lại. Điều này phù hợp với các kết luận đã đưa ra ở Chương 3 về mối quan hệ giữa NBĐMM với hiện tượng ENSO trên khu vực Việt Nam. Theo đó, NBĐMM có xu hướng đến sớm hơn trong những năm La Niña, và đến muộn hơn trong những năm El Niño.

Hình 4.1 dẫn ra biến động qua các năm của NBĐMM quan trắc (QT) và NBĐMM dự báo từ các tập số liệu phụ thuộc (DBPT) và độc lập (DBĐL) cho giai đoạn 1979-2019. Trong đó, các kết quả DBĐL được xác định bằng phương pháp kiểm chứng chéo-OLOO (chi tiết xem mục 2.2.4.2). Nhìn chung, các kết quả dự báo NBĐMM từ DBPT và DBĐL chỉ nắm bắt được khá tốt sự thay đổi của NBĐMM QT giữa các năm tại 3/7 vùng khí hậu (vùng R3, R6 và R7), với giá trị hệ số tương quan (CORR) trong khoảng 0.50 đến 0.84 (thoả mãn độ tin cậy thống kê 95%). Tuy nhiên, giá trị CORR có sự khác biệt rõ rệt giữa các vùng (Bảng 4.1). Giá trị này đạt trên 0.72 đối với cả DBPT và DBĐL ở vùng R6 và R7. CORR thấp hơn ở vùng R3, chỉ đạt 0.60 đối với DBPT và 0.50 đối với DBĐL. Trong khi đó, đối với các vùng khí hậu R1, R2, R4 và R5, khả năng nắm bắt sự thay đổi của NBĐMM quan trắc của các sản phẩm dự báo từ DBPT và DBĐL là tương đối kém, với giá trị CORR dao động trong khoảng 0.11 đến 0.49 (đa số không thoả mãn độ tin cậy thống kê 95%).

Ở tất cả các trường hợp xem xét (DBPT và DBĐL) tại 7 vùng khí hậu đều có giá trị ME=0, điều này cho thấy các kết quả dự báo bằng các mô hình thống kê truyền thống dựa trên các chỉ số khí hậu cho kết quả NBĐMM trung bình nhiều năm sát với giá trị trung bình nhiều năm QT thực tế. Sai số dự báo được thể hiện qua chỉ số MAE ở Bảng 4.1 cho thấy tại cùng một vùng khí hậu, các giá trị của MAE của DBPT và DBĐL là tương đối bằng nhau. Tuy nhiên, giá trị của MAE có sự khác biệt rõ rệt giữa các vùng khí hậu. MAE nhận giá trị khoảng 6 ngày đối với R7, khoảng 8~9 ngày đối với R4, R5 và R6, và trên 12 ngày đối với R1, R2 và R3. Đáng chú ý, MAE~20 ngày đối với khu vực R3.



Hình 4.1: So sánh biến trình giữa các năm của NBĐMM trong giai đoạn 1979-2019 giữa quan trắc (QT) với dự báo phụ thuộc (DBPT) và dự báo độc lập (DBDL). Phần màu ghi xám thể hiện miền giá trị $QT \pm$ phương sai

Như vậy, sự khác biệt của các chỉ số CORR, MAE giữa các vùng khí hậu đã cho thấy chất lượng của các mô hình dự báo thống kê truyền thống phụ thuộc nhiều vào mức độ quan hệ giữa yếu tố dự báo (NBĐMM) và nhân tố dự báo (các chỉ số khí hậu). Xét về kỹ năng dự báo có thể thấy, nhìn chung các mô hình thống kê truyền thống cho kết quả dự báo NBĐMM tốt hơn so với với dự báo khí hậu, cả đối với DBPT và DBDL và tại 7 vùng khí hậu, được thể hiện bởi giá trị dương (+) của SS_{MAE} . Chỉ duy nhất trường hợp DBDL cho khu vực R2 có SS_{MAE} âm, nhưng khá nhỏ là -

0.06. Đa phần, giá trị SS_{MAE} dao động trong khoảng +0.1 đến +0.50 và có sự khác biệt giữa các vùng khí hậu. Tại R6 và R7, giá trị SS_{MAE} dao động trong khoảng +0.25 đến +0.50, cao hơn so với các vùng khí hậu còn lại (dưới +0.17) trong cả DBPT và DBDL. Do đó, sự cải thiện về chất lượng dự báo ở R6 và R7 là rõ rệt hơn so với các vùng khí hậu khác (tức là R1-5).

Qua đó có thể nhận định rằng, việc ứng dụng phương pháp thống kê truyền thống sử dụng các chỉ số khí hậu làm nhân tố dự báo để dự báo hạn mùa NBĐMM tại R6 và R7 là có tiềm năng, nhất là khi các chỉ số hồi quy của phương trình dự báo có tính ổn định cao (xem Phụ lục - Hình 2).

Bảng 4.1: Một số chỉ số đánh giá kỹ năng dự báo NBĐMM bằng mô hình thống kê truyền thống tại 7 vùng khí hậu đối với DBPT và DBDL (bằng phương pháp kiểm nghiệm chéo -OLOO) trong giai đoạn 1979-2019. Dấu * thể hiện giá trị CORR thỏa mãn mức ý nghĩa thống kê 95% (kiểm nghiệm t)

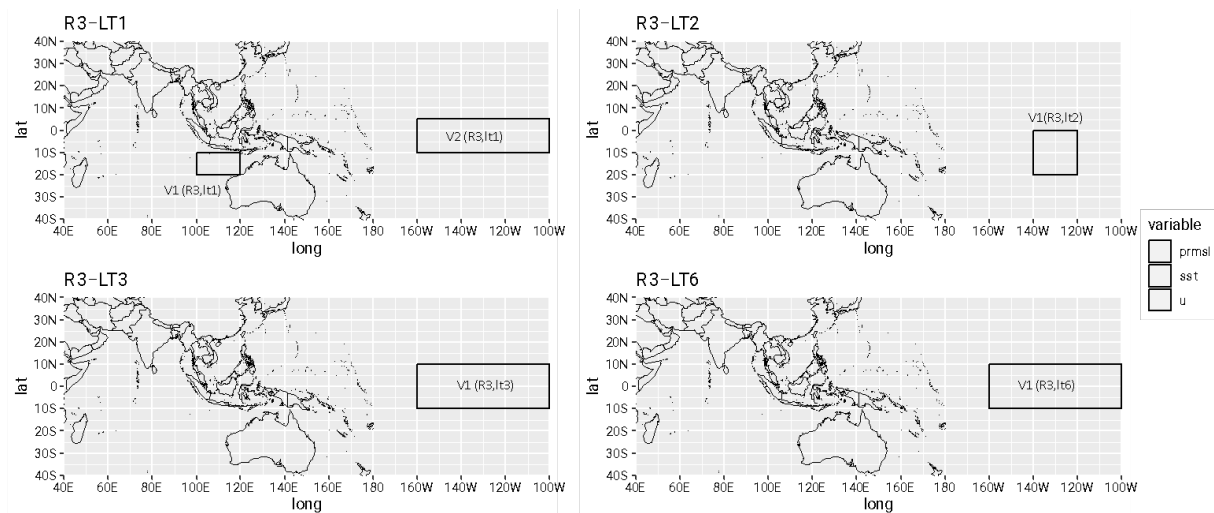
Vùng	Trường hợp	CORR	ME	MAE	SS_{MAE}
R1	DBPT	0.29	0	12	0.08
	DBDL	0.11	0	13	0
R2	DBPT	0.33	0	16	0
	DBDL	0.17	0	17	-0.06
R3	DBPT	0.60*	0	19	0.17
	DBDL	0.50*	0	20	0.13
R4	DBPT	0.49*	0	8	0.11
	DBDL	0.38	0	8	0.11
R5	DBPT	0.37	0	8	0.11
	DBDL	0.26	0	9	0
R6	DBPT	0.75*	0	8	0.33
	DBDL	0.72*	0	9	0.25
R7	DBPT	0.84*	0	5	0.5
	DBDL	0.79*	0	6	0.4

4.2. DỰ BÁO HẠN MÙA NGÀY BẮT ĐẦU MÙA MƯA BẰNG PHƯƠNG PHÁP THÔNG KÊ DỰA TRÊN SẢN PHẨM CỦA MÔ HÌNH CFSv2

Trong mục này sẽ trình bày việc xây dựng các phương trình dự báo NBĐMM khi sử dụng các trường khí quyển, đại dương dự báo của mô hình CFSv2 làm nhân tố dự báo. Do số lượng các trường dự báo của CFSv2 rất nhiều nên trong khuôn khổ luận án không thể khảo sát hết tất cả nên ở đây chỉ lựa chọn một số trường cơ bản, đó là các trường PRMSL, SST, U850. Kết quả khảo sát mối quan hệ giữa NBĐMM quan trắc và các trường dự báo PRMSL, SST, U850 cho tháng 3 của CFSv2 trên khu vực Việt Nam giai đoạn 1983-2011 cho thấy mối quan hệ này chỉ thực sự rõ rệt ở khu vực R3, R6 và R7 (xem Phụ lục - Hình 3-8). Do đó, luận án sẽ chỉ bàn luận đến chất lượng dự báo NBĐMM bằng phương pháp thông kê dựa trên sản phẩm CFSv2 tại các vùng R3, R6, và R7.

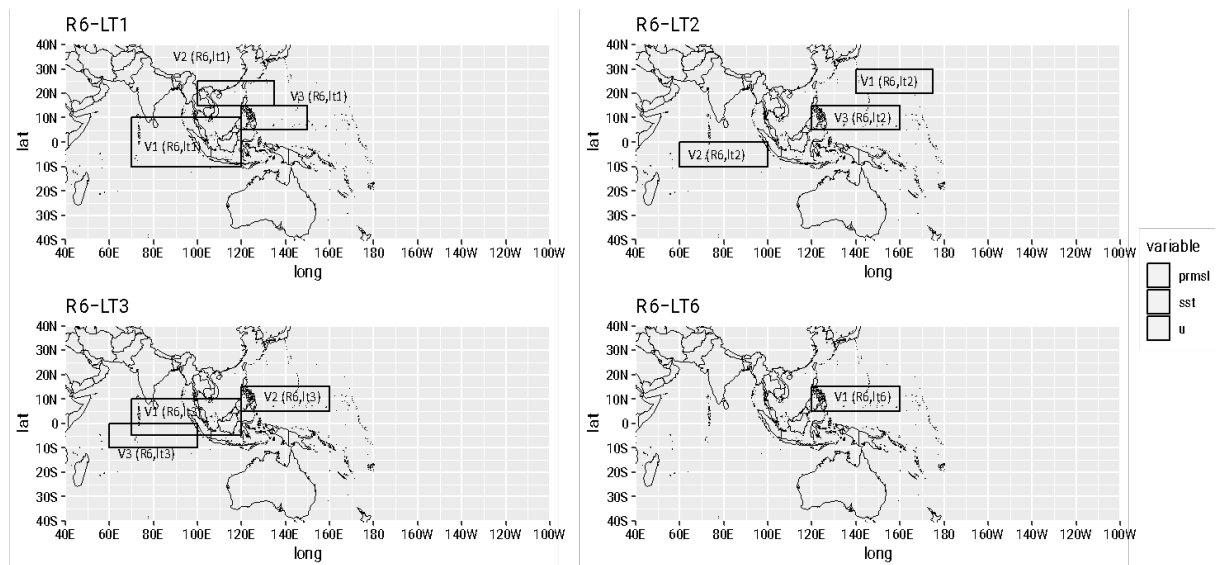
4.2.1. Lựa chọn nhân tố dự báo

Hình 4.2 đến 4.4 thể hiện các khu vực được lựa chọn làm nhân tố dự báo NBĐMM với các hạn dự báo 1-2-3-6 tháng cho các khu vực R3, R6 và R7 bằng phương pháp hồi quy từng bước. Nhìn chung, có sự khác biệt về nhân tố dự báo được lựa chọn giữa các vùng khí hậu. Điều này phản ánh sự khác biệt về vai trò chi phối của các nhân tố dự báo đối với sự biến động của NBĐMM tại từng vùng khí hậu. Sự khác biệt về nhân tố được lựa chọn (khu vực và biến) giữa các hạn dự báo của cùng một vùng khí hậu cũng phản ánh sự khác biệt về vai trò của các nhân tố dự báo đến các hạn dự báo khác nhau.



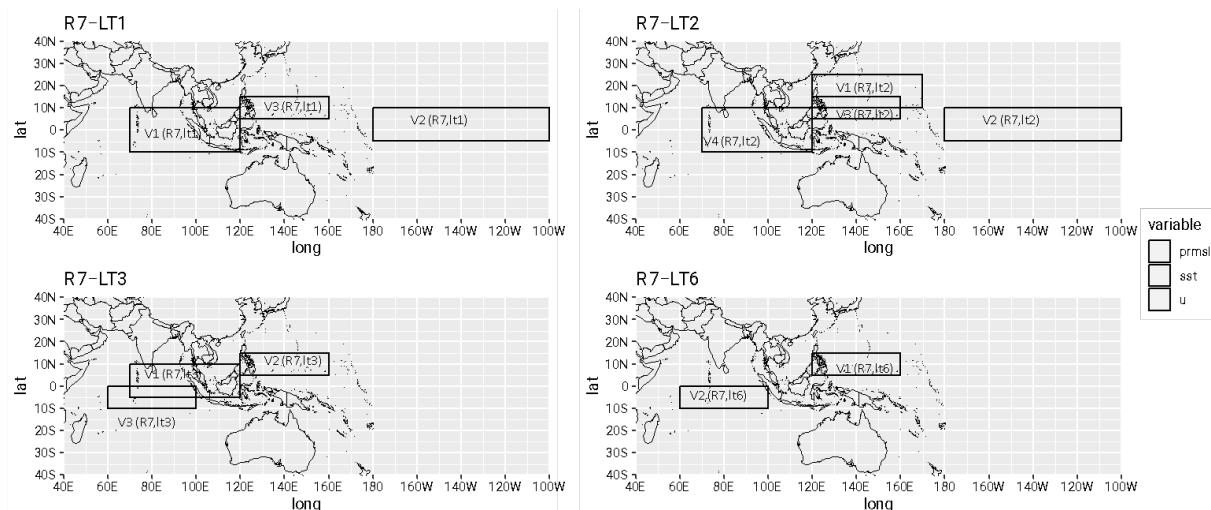
Hình 4.2: Các khu vực được lựa chọn làm nhân tố dự báo NBĐMM tại các hạn dự báo 1-2-3-6 tháng cho khu vực R3 bằng phương pháp hồi quy từng bước (dựa trên số liệu CFSv2 giai đoạn 1983-2011). Màu sắc của các khu vực tương ứng với các biến PRMSL (xanh), SST (đỏ), U850 (tím)

SST khu vực trung tâm Thái Bình Dương xích đạo có thể được coi là nhân tố quan trọng nhất đối với việc dự báo NBĐMM cho vùng R3 (Hình 4.2). Điều này được thể hiện khi giá trị SST tại khu vực này là nhân tố xuất hiện ở cả 4 hạn dự báo (1-2-3-6 tháng) và là nhân tố duy nhất được lựa chọn đối với hạn dự báo 2-3-6 tháng. Mối quan hệ giữa SST ở khu vực trung tâm Thái Bình Dương xích đạo và NBĐMM là đồng biến không chỉ đối với vùng R3, mà còn với R6 và R7 (xem Phụ lục - Hình 5). Về mặt nhiệt động lực, sự ấm lên hoặc lạnh đi của SST tại khu vực trung tâm Thái Bình Dương xích đạo được coi là tín hiệu xác định các pha ENSO. Trong những năm El Niño, nhiệt độ mặt nước biển ấm hơn ở khu vực này làm cho NBĐMM có xu hướng đến muộn hơn, và ngược lại, trong những năm La Niña, nhiệt độ mặt nước biển lạnh hơn, NBĐMM có xu hướng tới sớm hơn (Chương 3).



Hình 4.3: Tương tự Hình 4.2 nhưng cho khu vực R6

Đối với khu vực R6 (Hình 4.3) và R7 (Hình 4.4), có thể thấy bên cạnh SST khu vực trung tâm Thái Bình Dương xích đạo, SST ở khu vực tây Thái Bình dương xích đạo và SST khu vực nam Ấn Độ Dương xích đạo cũng được lựa chọn làm nhân tổ dự báo. Tuy nhiên, quan hệ giữa SST ở các khu vực kể trên với NBĐMM là trái ngược nhau. Giá trị hệ số tương quan âm tại tây Thái Bình dương xích đạo, dương tại trung tâm Thái Bình Dương xích đạo và tại khu vực nam Ấn Độ Dương xích đạo (xem Phụ lục - Hình 5). Điều đó chứng tỏ sự nóng lên (lạnh đi) của SST ở khu vực trung tâm Thái Bình Dương xích đạo và khu vực nam Ấn Độ Dương xích đạo (khu vực tây Thái Bình Dương xích đạo) sẽ kéo theo NBĐMM đến muộn hơn (sớm hơn). Về mặt vật lý, SST khu vực trung tâm Thái Bình Dương xích đạo càng cao sẽ làm cho nhánh đi lên của hoàn lưu Walker ở tây Thái Bình dương có xu thế dịch về phía đông, còn SST ở khu vực tây Thái Bình Dương xích đạo càng thấp có thể kéo theo sự kìm hãm thiết lập gió mùa tây nam, hệ quả là NBĐMM sẽ xảy ra muộn hơn.



Hình 4.4: Tương tự Hình 4.2 nhưng cho khu vực R7

Bên cạnh đó, đối với khu vực R6 và R7, các nhân tố khác liên quan đến PRMSL, và U550 cũng được lựa chọn khi xây dựng mô hình dự báo NBĐMM (Hình 4.3-4.4). Đáng chú ý, tại hai khu vực này, trường U550 tại khu vực nam Ấn Độ Dương - tây Thái Bình Dương xích đạo là nhân tố dự báo được lựa chọn phổ biến đối với các hạn dự báo dưới 3 tháng. Giá trị của U550 tại khu vực nam Ấn Độ Dương - tây Thái Bình Dương xích đạo có mối tương quan âm với NBĐMM (xem Phụ lục - Hình 4). Ở một góc độ nào đó, có thể coi trường U550 tại khu vực này như là một trung tâm chi phối quá trình thiết lập gió mùa tây nam. U550 trên khu vực nam Ấn Độ Dương - tây Thái Bình Dương xích đạo càng mạnh càng thúc đẩy quá trình này do dòng khí vượt xích đạo từ Nam bán cầu lên khu vực vịnh Bengal, làm thúc đẩy và tăng cường hoạt động của gió mùa tây nam, góp phần làm mùa mưa đến sớm hơn.

Ngoài ra, trường PRMSL tại khu vực tây bắc Thái Bình Dương cũng được lựa chọn làm nhân tố dự báo với hạn dự báo 2 tháng cho các vùng R6 và R7. Mối quan hệ giữa NBĐMM tại hai vùng này và trường PRMSL là đồng biến. Theo đó, khi giá trị của PRMSL tại khu vực tây bắc Thái Bình Dương tăng thì NBĐMM tại R6 và R7 có xu hướng tới muộn hơn.

4.2.2. Kết quả dự báo và đánh giá

Bằng việc sử dụng phương pháp hồi quy từng bước, phương trình dự báo NBĐMM tại các vùng R3, R6 và R7 cho từng hạn dự báo (1-2-3-6 tháng), dựa trên sản phẩm của CFSv2 trong giai đoạn 1983-2011 thu được như sau:

❖ Tại R3:

$$NBĐMM_{(R3,lt=1)} = 130.41 + 12.24 * V1_{(R3,lt=1)} + 9.88 * V2_{(R3,lt=1)}$$

$$NBĐMM_{(R3,lt=2)} = 130.41 + 15.96 * V1_{(R3,lt=2)}$$

$$NBĐMM_{(R3,lt=3)} = 130.41 + 14.61 * V1_{(R3,lt=3)}$$

$$NBĐMM_{(R3,lt=6)} = 130.41 + 14.13 * V1_{(R3,lt=6)}$$

❖ Tại R6:

$$NBĐMM_{(R6,lt=1)} = 118.14 - 6.5 * V1_{(R3,lt=1)} + 4.6 * V2_{(R6,lt=1)} - 2.91 * V3_{(R6,lt=1)}$$

$$NBĐMM_{(R6,lt=2)} = 118.14 - 5.16 * V1_{(R3,lt=2)} + 11.4 * V2_{(R6,lt=2)} - 10.76 * V3_{(R6,lt=2)}$$

$$NBĐMM_{(R6,lt=3)} = 118.14 - 2.16 * V1_{(R3,lt=3)} - 7.64 * V2_{(R6,lt=3)} + 5.81 * V3_{(R6,lt=3)}$$

$$NBĐMM_{(R6,lt=6)} = 118.14 - 10.83 * V1_{(R3,lt=6)}$$

❖ Tại R7:

$$NBĐMM_{(R7,lt=1)} = 127.00 - 5.61 * V1_{(R7,lt=1)} + 5.45 * V2_{(R7,lt=1)} - 3.25 * V3_{(R7,lt=1)}$$

$$NBĐMM_{(R7,lt=2)} = 127.00 - 7.59 * V1_{(R7,lt=2)} + 10.20 * V2_{(R7,lt=2)} - 8.76 * V3_{(R7,lt=2)} - 4.10 * V4_{(R7,lt=2)}$$

$$NBĐMM_{(R7,lt=3)} = 127.00 - 5.56 * V1_{(R7,lt=3)} - 5.69 * V2_{(R7,lt=3)} + 3.58 * V3_{(R7,lt=3)}$$

$$NBĐMM_{(R7,lt=6)} = 127.00 - 8.83 * V1_{(R7,lt=6)} + 5.23 * V2_{(R7,lt=6)}$$

Thông tin cụ thể về nhân tố dự báo của từng phương trình được trình bày cụ thể ở trong Bảng 4.2 và Hình 4.2-4.4. Có thể thấy, giá trị của hệ số tự do trong các phương trình dự báo tương ứng với giá trị NBĐMM QT trung bình trong giai đoạn 1983-2011 tại từng vùng khí hậu. Như vậy, sự biến động giữa các năm của các giá trị dự báo NBĐMM sẽ chịu sự chi phối bởi giá trị của các nhân tố dự báo. Sự chi phối này được thể hiện qua dấu (+/-) của các hệ số hồi quy tương ứng với từng nhân tố. Các giá trị của các hệ số hồi quy trong phương trình dự báo NBĐMM tại các vùng

R3, R6 và R7 cho hạn dự báo dưới 6 tháng là hợp lý, do chúng phản ánh đúng mối quan hệ giữa sự biến động NBDMM và các nhân tố hồi quy như đề cập ở mục 4.2.1.

Bảng 4.2: Thông tin về kinh độ, vĩ độ, trường của các nhân tố dự báo được lựa chọn từng hạn dự báo 1-2-3-6 tháng của R3, R6 và R7 (dựa trên số liệu CFSv2 giai đoạn 1983-2011). Các nhân tố dự báo được sắp xếp theo thứ tự tuyển chọn của phương pháp hồi quy từng bước

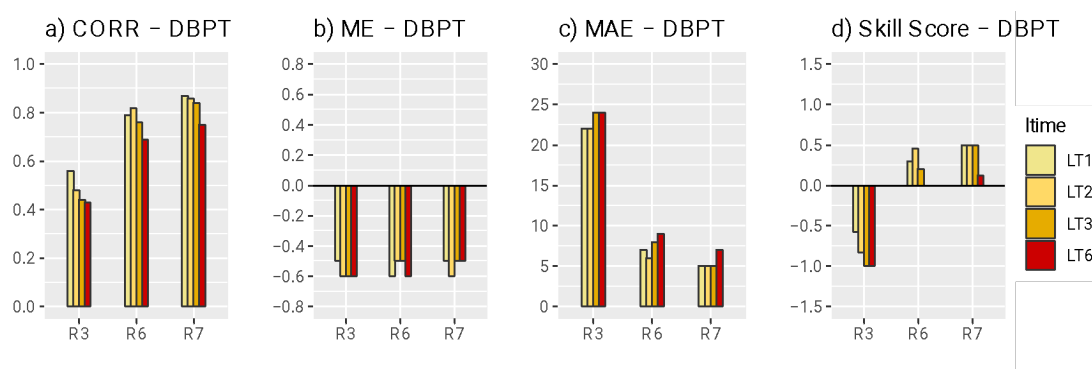
Vùng	Hạn dự báo (tháng)	Ký hiệu nhân tố	Tên biến	Kinh độ - vĩ độ
R3	1	$V1_{(R3,lt=1)}$	SST	(100°E -120°E; 20°S-10°S)
		$V2_{(R3,lt=1)}$	SST	(160°W -100°W; 10°S-5°N)
	2	$V1_{(R3,lt=2)}$	SST	(140°W -120°W; 20°S-0°N)
	3	$V1_{(R3,lt=3)}$	SST	(160°W -100°W; 10°S-10°N)
	6	$V1_{(R3,lt=6)}$	SST	(160°W -100°W; 10°S-10°N)
R6	1	$V1_{(R3,lt=1)}$	U850	(70°E -120°E; 10°S-10°N)
		$V2_{(R6,lt=1)}$	U850	(100°E -135°E; 15°N-25°N)
		$V3_{(R6,lt=1)}$	SST	(120°E -150°E; 5°N-15°N)
	2	$V1_{(R3,lt=2)}$	PRMSL	(140°E -175°E; 20°N-30°N)
		$V2_{(R6,lt=2)}$	SST	(60°E -100°E; 10°S-0°N)
		$V3_{(R6,lt=2)}$	SST	(120°E -160°E; 5°N-15°N)
	3	$V1_{(R3,lt=3)}$	U850	(70°E -120°E; 5°S-10°N)
		$V2_{(R6,lt=3)}$	SST	(120°E -160°E; 5°N-15°N)
		$V3_{(R6,lt=3)}$	SST	(60°E -100°E; 10°S-0°N)
	6	$V1_{(R3,lt=6)}$	SST	(120°E -160°E 5°N-15°N)
R7	1	$V1_{(R7,lt=1)}$	U850	(70°E -120°E; 10°S-10°N)
		$V2_{(R7,lt=1)}$	SST	(180°E -100°W; 5°S-10°N)
		$V3_{(R7,lt=1)}$	SST	(120°E -160°E; 5°N-15°N)
	2	$V1_{(R7,lt=2)}$	PRMSL	(120°E -170°E; 10°N-25°N)

		$V2_{(R7,l=2)}$	SST	(180°E -100°W; 5°S-10°N)
		$V3_{(R7,l=2)}$	SST	(120°E -160°E; 5°N-15°N)
		$V4_{(R7,l=2)}$	U850	(70°E -120°E; 10°S-10°N)
	3	$V1_{(R7,l=3)}$	U850	(70°E -120°E; 5°S-10°N)
		$V2_{(R7,l=3)}$	SST	(120°E -160°E; 5°N-15°N)
		$V3_{(R7,l=3)}$	SST	(60°E -100°E; 10°S-0°N)
	6	$V1_{(R7,l=6)}$	SST	(120°E -160°E; 5°N-15°N)
		$V2_{(R7,l=6)}$	SST	(60°E -100°E; 10°S-0°N)

Để đánh giá khả năng dự báo hạn mùa NBĐMM bằng phương pháp thống kê dựa trên sản phẩm dự báo của mô hình CFSv2, luận án đã tiến hành khảo sát các chỉ số thống kê thông dụng bao gồm CORR, ME, MAE, SS_{MAE} . Việc đánh giá này được tiến hành trên 3 tập số liệu: 1) tập số liệu phụ thuộc giai đoạn 1983-2011, được sử dụng để xây dựng phương trình dự báo NBĐMM (Hình 4.5); 2) tập số liệu độc lập giai đoạn 2012-2019, không tham gia vào quá trình xây dựng phương trình dự báo NBĐMM (Hình 4.6); và 3) tập số liệu độc lập giai đoạn 1983-2011, được xây dựng bằng phương pháp kiểm chứng chéo OLOO (chi tiết xem mục 2.2.4.2, Hình 4.7).

Hình 4.5 dẫn ra kết quả tính các chỉ số đánh giá trên tập số liệu phụ thuộc trong giai đoạn 1983-2011 cho các vùng R3, R6 và R7 với các hạn dự báo 1-2-3-6 tháng. Có thể nhận thấy rằng giá trị CORR tại cả 3 vùng khí hậu đều đạt trên 0.43 (thoả mãn độ tin cậy thống kê 95%) cho tất cả các hạn dự báo. Như vậy NBĐMM dự báo từ các mô hình thống kê động lực đã nắm bắt tốt sự biến động qua các năm của NBĐMM quan trắc thực tế. Tuy nhiên, giá trị CORR có sự khác nhau giữa các vùng khí hậu và giữa các hạn dự báo. Nhìn chung, trên cùng một vùng khí hậu, CORR không chênh lệch nhiều giữa các hạn dự báo nhưng giá trị của CORR có xu thế giảm khi tăng hạn dự báo, nghĩa là khả năng nắm bắt sự biến động của NBĐMM bị suy giảm khi hạn dự báo càng xa. Giá trị CORR cũng có sự khác biệt đáng kể giữa các vùng khí hậu. CORR có giá trị lớn tại R7 (0.75~0.86) và R6 (0.69~0.82), trong khi đó tại R3 giá trị CORR chỉ dao động dưới mức 0.56. Về sai số ME, các mô hình thống kê cho kết quả

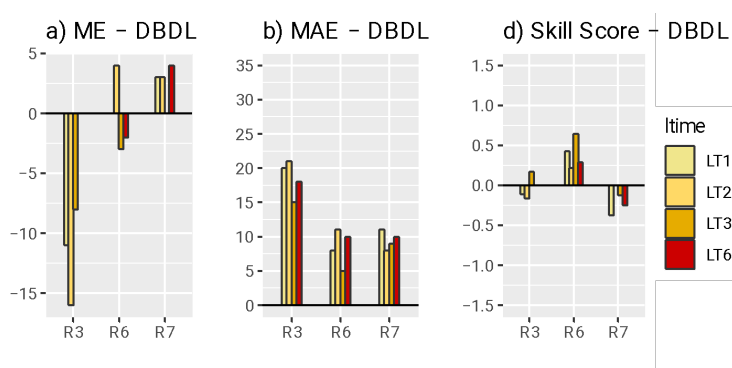
dự báo NBĐMM trung bình nhiều năm sớm hơn so với giá trị này tại quan trắc thực tế, nhưng không đáng kể (dưới 0.6 ngày). Xu thế này nói chung tương đồng giữa các vùng R3, R6 và R7 và trong tất cả các hạn dự báo dưới 6 tháng. Đối với sai số tuyệt đối (MAE), kết quả đánh giá trên tập phụ thuộc giai đoạn 1983-2011 cho thấy sai số MAE có xu thế lớn hơn ở các hạn dự báo dài hơn (ví dụ 6 tháng). Sự khác biệt về MAE giữa các hạn dự báo trong cùng một vùng khí hậu là nhỏ, dao động ổn định ở mức dưới 2 ngày. Sự khác biệt của MAE rõ rệt hơn khi so sánh giữa các vùng khí hậu. MAE lớn nhất ở R3, dao động trong khoảng 22~24 ngày, trong khi đó, giá trị này chỉ trong khoảng 6~8 ngày đối với các vùng R6 và R7. Chỉ số SS_{MAE} trên tập phụ thuộc giai đoạn 1983-2011 cũng có sự khác biệt rõ rệt giữa các vùng. Theo đó, R3 nhận giá trị âm (-) trong khi các vùng R6 và R7 nhận giá trị (+) tại tất cả các hạn dự báo (1-2-3-6 tháng). Điều này có nghĩa rằng, các mô hình thống kê có kỹ năng dự báo NBĐMM cho các vùng khí hậu là khác nhau khi so sánh với dự báo khí hậu. Tại vùng R3, SS_{MAE} dao động trong khoảng -1.0 đến -0.57. Chất lượng dự báo ở vùng này được cải thiện với các hạn dự báo ngắn (ví dụ 1 tháng). Đối với R6 và R7, giá trị SS_{MAE} dao động trong khoảng 0.13~0.5. Đáng chú ý, ở R7, SS_{MAE} của các dự báo hạn dưới 3 tháng khá ổn định và nhận giá trị cao (khoảng 0.5).



Hình 4.5: Các chỉ số đánh giá chất lượng dự báo NBĐMM bằng phương pháp thống kê động lực trên cơ sở sản phẩm CFSv2 với các hạn dự báo 1-2-3-6 tháng (trên tập số liệu phụ thuộc trong giai đoạn 1983-2011), bao gồm: a) CORR, b) ME, c) MAE, d) SS_{MAE}

Hình 4.6 dẫn ra kết quả đánh giá khả năng dự báo hạn mùa (dưới 6 tháng) NBĐMM bằng phương pháp thống kê sử dụng sản phẩm của CFSv2 trên tập số liệu

độc lập cho giai đoạn 2011-2019. Nhìn chung, xu hướng sai số (ME) có sự khác biệt giữa các hạn dự báo và giữa các vùng khí hậu. Theo đó, tại R3, NBĐMM dự báo với hạn dưới 6 tháng có xu hướng sớm hơn 7~16 ngày so với quan trắc. Trong khi đó, xu hướng này là ngược lại đối với R7. NBĐMM dự báo cho R7 muộn hơn khoảng 3~4 ngày so với thực tế. Đối với R6, xu hướng sai số phụ thuộc vào hạn dự báo. Cụ thể, NBĐMM dự báo có xu hướng muộn hơn quan trắc thực tế khoảng 0~4 ngày tại các hạn dự báo ngắn (dưới 2 tháng), và sớm hơn quan trắc thực tế khoảng 2~3 ngày tại các hạn dự báo dài hơn (trên 3 tháng).

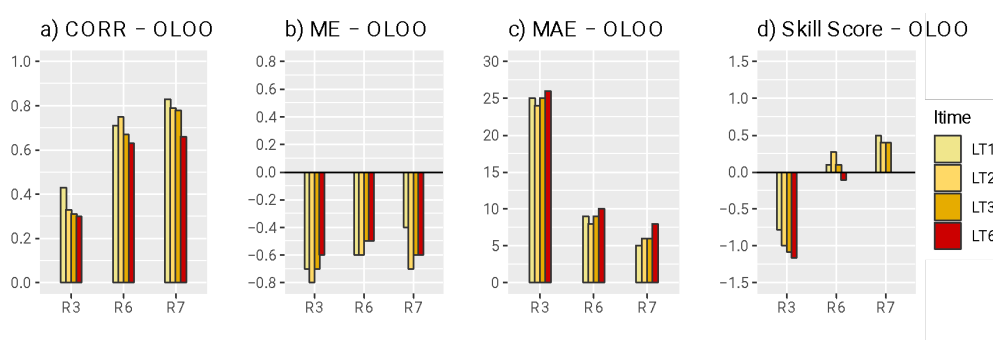


Hình 4.6: Các chỉ số đánh giá chất lượng dự báo NBĐMM bằng phương pháp thống kê động lực trên cơ sở sản phẩm CFSv2 với các hạn dự báo 1-2-3-6 tháng (trên tập số liệu độc lập trong giai đoạn 2012-2019), bao gồm: a) ME, b) MAE, c) SS_{MAE}

Sai số tuyệt đối (MAE) của mô hình thống kê dựa trên tập số liệu độc lập trong giai đoạn 2012-2019 cho các vùng R6 và R7 (Hình 4.6) nói chung có giá trị lớn hơn so với các kết quả thu được dựa trên tập phụ thuộc trong giai đoạn 1983-2011 (Hình 4.5). Cụ thể, sai số MAE trong trường hợp DBĐL tại hai vùng khí hậu này khá tương đồng nhau và dao động trong khoảng 8~10 ngày so với 6~8 ngày trong trường hợp DBPT. Trong khi đó, đối với R3, giá trị MAE cho DBĐL dao động trong khoảng 15~21 ngày, thấp hơn so với DBPT (22~24 ngày). Kết quả là chỉ số SS_{MAE} tại R3 trong trường hợp DBĐL chỉ trong khoảng -0.17 đến +0.17, cải thiện hơn so với trong trường hợp DBPT. So sánh giữa ba vùng khí hậu, giá trị SS_{MAE} tại R6 đạt giá trị cao nhất (trong khoảng 0.21~0.64). Đặc biệt, giá trị SS_{MAE} tại R6 trong trường hợp DBĐL cao hơn so với trong trường hợp DBPT khi so sánh ở cùng một hạn dự báo. Trong khi đó, giá trị SS_{MAE} tại R7 lại mang giá trị âm, điều này có nghĩa là trên tập độc lập

trong giai đoạn 2012-2019, các giá trị dự báo hạn mùa NBĐMM tại đây cho kết quả kém hơn so với dự báo khí hậu. Có thể thấy rằng, mặc dù sai số tuyệt đối MAE tại R6 và R7 trong trường hợp DBĐL cho giai đoạn 2012-2019 là xấp xỉ nhau, nhưng dấu giá trị SS_{MAE} lại có sự trái ngược giữa hai vùng này. Điều này có thể liên quan đến sự khác biệt về mức độ biến động giữa các năm của NBĐMM tại R6 và R7. Cụ thể, NBĐMM tại R7 có sự biến động giữa các năm nhỏ hơn so với R6. Do đó, việc cải thiện chất lượng dự báo hạn mùa so với dự báo khí hậu bằng các mô hình thống kê-động lực dựa trên sản phẩm CFSv2 gặp nhiều thách thức tại các vùng mà NBĐMM có sự ổn định giữa các năm như tại R7.

Hình 4.7 dẫn ra kết quả đánh giá bằng phương pháp kiểm chứng chéo OLOO kết quả dự báo hạn mùa (dưới 6 tháng) NBĐMM của phương pháp thống kê sử dụng sản phẩm của CFSv2 trên tập số liệu độc lập giai đoạn 1983-2011. Nhìn chung, các chỉ số CORR, ME, MAE, SS_{MAE} trong trường hợp DBĐL cho kết quả tương đồng với DBPT (Hình 4.5). Theo đó, khả năng nắm bắt biến động giữa các năm của NBĐMM tốt nhất ở các dự báo hạn ngắn (ví dụ 1 tháng) và tốt nhất ở vùng R7, với giá trị CORR lên tới trên 0.8. Trên cả 3 vùng khí hậu, NBĐMM dự báo có xu hướng sớm hơn so với thực tế khoảng 0.5 ngày tại tất cả hạn dự báo dưới 6 tháng. Bên cạnh đó, chỉ các NBĐMM dự báo cho các vùng R6 và R7 có cải thiện hơn so với dự báo khí hậu.



Hình 4.7: Các chỉ số đánh giá chất lượng dự báo NBĐMM bằng phương pháp thống kê động lực trên cơ sở sản phẩm CFSv2 với các hạn dự báo 1-2-3-6 tháng (trên tập số liệu độc lập trong giai đoạn 1983-2011 - phương pháp kiểm chứng chéo OLOO), bao gồm: a) CORR, b) ME, c) MAE, d) SS_{MAE}

Tóm lại, từ việc khảo sát các chỉ số CORR, ME, MAE, SS_{MAE} trên các tập dữ liệu phụ thuộc và độc lập khác nhau, kỹ năng dự báo NBĐMM bằng phương pháp

thống kê-động lực trên sản phẩm dự báo của mô hình CFSv2 với các hạn dự báo dưới 6 tháng đã được chỉ ra. Theo đó, trên cả 3 vùng khí hậu (R3, R6 và R7), các dự báo hạn ngắn (1 tháng) cho kết quả nói chung tốt hơn so với các dự báo hạn dài hơn (ví dụ 6 tháng). Chất lượng của các dự báo có sự khác biệt giữa các vùng khí hậu, có thể liên quan tới sự khác biệt về đặc trưng của NBĐMM tại mỗi vùng, bao gồm: mối quan hệ giữa NBĐMM với các trường quy mô lớn (nhân tố dự báo) và sự biến động giữa các năm của NBĐMM. Các kết quả chỉ ra rằng, chất lượng dự báo NBĐMM của mô hình thống kê được cải thiện hơn so với dự báo khí hậu chỉ ở các vùng R6 và R7. Điều này cho thấy tiềm năng của việc ứng dụng phương pháp thống kê-động lực trên cơ sở sản phẩm CFSv2 để dự báo hạn mùa NBĐMM tại các vùng này. Các chỉ số hồi quy của phương trình dự báo cho các vùng này có tính ổn định cao (xem Phụ lục - Hình 10-11). Tuy nhiên, việc cải thiện chất lượng dự báo có thể gặp nhiều thách thức tại các vùng mà NBĐMM có sự ổn định cao giữa các năm như tại R7.

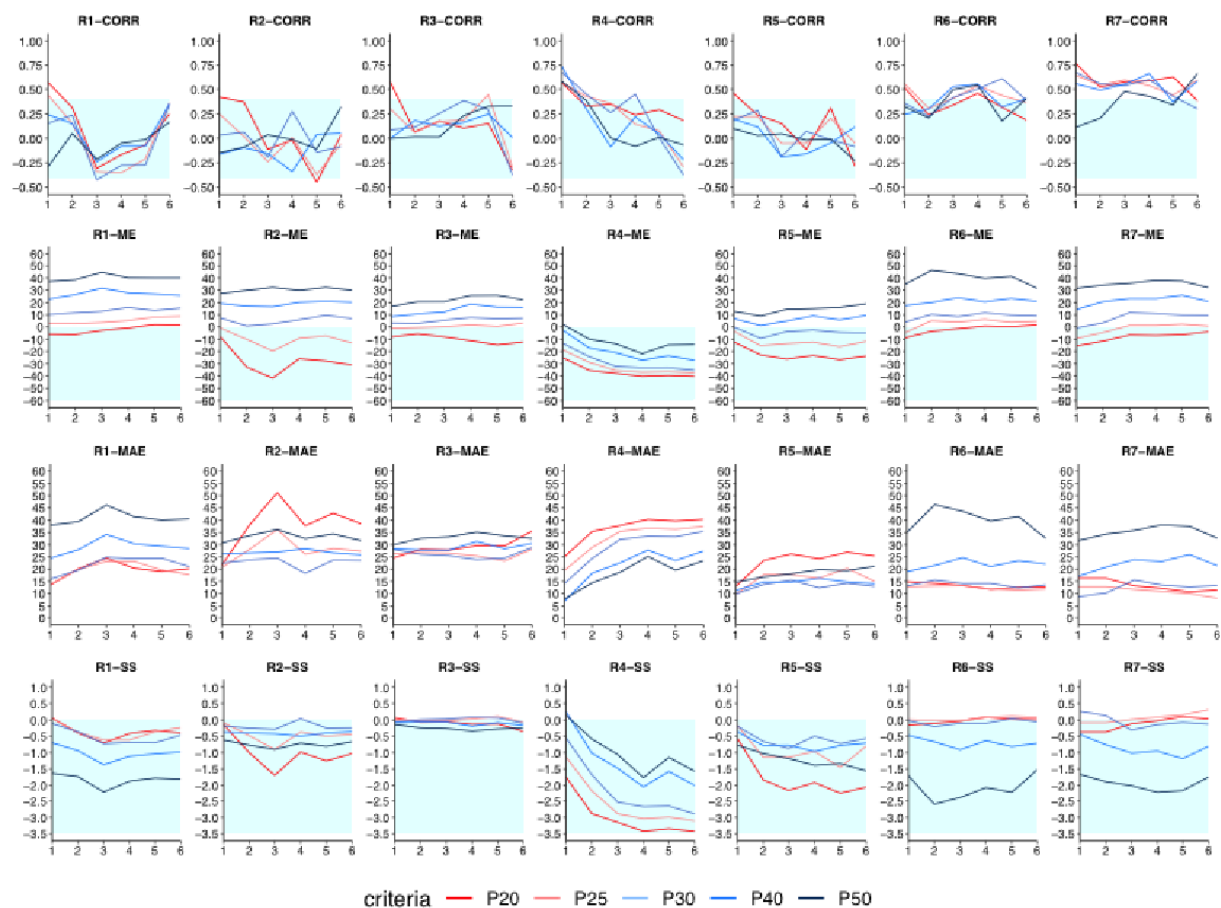
4.3. DỰ BÁO HẠN MÙA NGÀY BẮT ĐẦU MÙA MƯA TỪ SẢN PHẨM MƯA CỦA CFSV2

Trong những năm gần đây, việc dự báo NBĐMM sử dụng trực tiếp sản phẩm mưa của mô hình số trị đã nhận được nhiều sự quan tâm của các nhà khoa học trên thế giới. Điểm mạnh của phương pháp này so với phương pháp thống kê (bao gồm thống kê động lực và thống kê truyền thống) là khu vực áp dụng không bị hạn chế do không đòi hỏi mối quan hệ chặt chẽ giữa yếu tố dự báo và nhân tố dự báo giống như phương pháp thống kê. Tuy nhiên, thách thức không nhỏ trong trường hợp này là chất lượng dự báo NBĐMM phụ thuộc chủ yếu vào độ chính xác dự báo lượng mưa của mô hình động lực và chỉ tiêu xác định NBĐMM từ lượng mưa dự báo của mô hình. Trong mục này, luận án sẽ tiến hành đánh giá khả năng dự báo NBĐMM khi sử dụng trực tiếp sản phẩm mưa của CFSv2 trên 7 vùng khí hậu của Việt Nam với 5 bộ chỉ tiêu xác định NBĐMM khác nhau (xem Mục 2.5.1)

Hình 4.8 dẫn ra kết quả tính các chỉ số CORR, ME, MAE, SS_{MAE} được dùng để đánh giá khả năng dự báo NBĐMM cho giai đoạn 1983-2019 với 5 chỉ tiêu xác

định NBĐMM khác nhau (P20, P25, P30, P40, P50). Nhìn chung, trong tất cả các trường hợp khảo sát liên quan đến hạn dự báo và chỉ tiêu xác định NBĐMM, giá trị CORR chủ yếu dao động trong khoảng -0.4 đến +0.4 tại 6/7 vùng khí hậu (trừ R7). Đây là khoảng ngưỡng không thoả mãn độ tin cậy 95% của kiểm nghiệm t-test. Điều này cho thấy khả năng nắm bắt được sự biến động hàng năm của NBĐMM quan trắc khi sử dụng trực tiếp mưa dự báo của mô hình CFSv2 mà không thông qua hiệu chỉnh về cơ bản là chưa tốt và có sự không tương đồng giữa các vùng khí hậu.

Giá trị CORR có sự khác biệt khi so sánh giữa các hạn dự báo và giữa các chỉ tiêu xác định NBĐMM tại mỗi vùng. Tuy nhiên, mối quan hệ này không có quy luật rõ ràng tại đa số các vùng khí hậu. Đối với khu vực R7 - nơi duy nhất có các giá trị CORR thoả mãn mức độ tin cậy 95%, hệ số CORR nhìn chung có xu hướng giảm khi hạn dự báo tăng. Ảnh hưởng của việc lựa chọn chỉ tiêu xác định NBĐMM đến giá trị CORR tại R7 cũng khá rõ, đặc biệt ở hạn dự báo ngắn. Theo đó, các kết quả dự báo trong trường hợp sử dụng chỉ tiêu P50 cho giá trị CORR thấp hơn tại tất cả các hạn dự báo, khi so sánh với trường hợp sử dụng các chỉ tiêu khác.



Hình 4.8: Các chỉ số đánh giá chất lượng dự báo NBĐMM khi sử dụng trực tiếp sản phẩm mưa của CFSv2, bao gồm CORR (hàng 1, màu nền xanh ứng với khoảng giá trị không thoả mãn mức độ ý nghĩa thống kê 95%), ME (hàng 2, màu nền xanh ứng với khoảng giá trị âm), MAE (hàng 3), SS_{MAE} (hàng 4, màu nền xanh ứng với khoảng giá trị âm) tại hạn dự báo 1-2-3-4-5-6 tháng tại 7 vùng khí hậu, với các chỉ tiêu xác định NBĐMM khác nhau

Việc lựa chọn chỉ tiêu xác định NBĐMM ảnh hưởng rõ rệt đến giá trị ME và MAE tại tất cả các vùng khí hậu. Nhìn chung, ngoại trừ vùng R4, tại đa số các hạn dự báo, các chỉ tiêu P30, P40, P50 cho kết quả dự báo có xu hướng muộn hơn so với quan trắc thực tế, trong khi đó, các chỉ tiêu P20 và P25 lại cho các kết quả dự báo đến sớm hơn. Đối với vùng R4, tất cả các chỉ tiêu khảo sát (P20, P25, P30, P40, P50) đều cho kết quả dự báo sớm hơn so với quan trắc. Sai số trung bình (ME) tương đối ổn định giữa các hạn dự báo tại tất cả các vùng khí hậu và các chỉ tiêu khảo sát. Tuy vậy, sự khác biệt về giá trị ME liên quan đến việc lựa chọn chỉ tiêu xác định NBĐMM là

tương đối lớn, dao động trong khoảng 20 đến 50 ngày. Điều đó nói lên tầm quan trọng của việc lựa chọn chỉ tiêu phù hợp để xác định NBĐMM từ sản phẩm mưa dự báo của mô hình động lực.

Về sai số trung bình tuyệt đối, nhìn chung MAE có xu hướng tăng khi tăng hạn dự báo. Đối với các vùng R1, R2, R3, R6, và R7, việc sử dụng các chỉ tiêu có ngưỡng P thấp (P20, P25) sẽ cho giá trị MAE nhỏ hơn so với trường hợp ngưỡng P cao (P40, P50). Ngược lại, đối với các vùng R4 và R5, việc sử dụng ngưỡng P cao (P40 và P50) lại cho giá trị MAE nhỏ hơn so với trường hợp sử dụng ngưỡng P thấp. Tuy nhiên, kết quả dự báo hạn mùa NBĐMM khi sử dụng trực tiếp sản phẩm mưa của CFSv2 chưa mang lại hiệu quả đáng kể so với dự báo khí hậu trên khu vực Việt Nam. Điều này được thể hiện bởi chỉ số kỹ năng SS_{MAE} nhận giá trị âm (-) tại đa số các trường hợp khảo sát cho 7 vùng khí hậu với các hạn dự báo và các chỉ tiêu xác định NBĐMM khác nhau. SS_{MAE} chỉ nhận giá trị (+) đối với một số trường hợp tại R7 với giá trị $SS_{MAE} < 0.3$.

Tóm lại, thông qua việc khảo sát các chỉ số đánh giá CORR, ME, MAE, SS_{MAE} , có thể nhận thấy kỹ năng dự báo NBĐMM khi sử dụng trực tiếp sản phẩm mưa dự báo của mô hình CFSv2 nói chung kém hơn so với dự báo khí hậu tại tất cả các vùng khí hậu của Việt Nam. Kỹ năng này là khác nhau giữa các vùng khí hậu. Kết quả dự báo về cơ bản chưa nắm bắt được sự biến động giữa các năm của NBĐMM quan trắc tại hầu hết các vùng khí hậu, trừ R7. Kỹ năng dự báo NBĐMM khi sử dụng trực tiếp sản phẩm mưa của CFSv2 kém hơn so với dự báo khí hậu về cơ bản là do sai số dự báo lượng mưa ngày của CFSv2 cũng như do độ phân giải của mô hình quá thô ($1.0^\circ \times 1.0^\circ$) mà nó cũng đã được đề cập tới trong một số nghiên cứu trước đây [Ren và cs, 2017; Ren và cs, 2018;...]. Đối với khu vực Việt Nam, Phan-Van và cs (2018) đã chỉ ra rằng việc sử dụng mô hình khu vực để cải thiện độ phân giải từ sản phẩm của CFSv2 sẽ làm tăng chất lượng của sản phẩm dự báo trường mưa trong các giai đoạn chuyển tiếp từ mùa mưa sang mùa khô và từ mùa khô sang mùa mưa tại hầu khắp các vùng khí hậu. Ngoài ra, việc lựa chọn chỉ tiêu phù hợp để xác định NBĐMM từ lượng

mưa mô hình cũng là một trong những nhân tố ảnh hưởng đến kỹ năng dự báo khi sử dụng trực tiếp sản phẩm mưa của CFSv2.

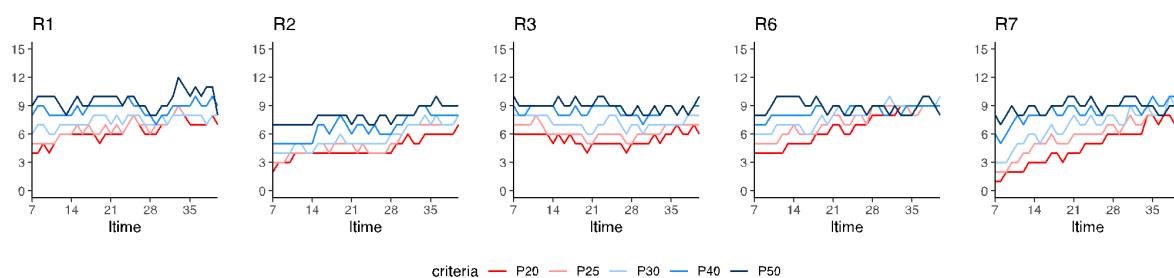
4.4. DỰ BÁO NỘI MÙA NGÀY BẮT ĐẦU MÙA MƯA TỪ SẢN PHẨM MƯA CỦA S2S-ECMWF

Gần đây, các nghiên cứu đã chỉ ra rằng NBĐMM có thể được dự báo với hạn nội mùa khi sử dụng trực tiếp sản phẩm của mô hình số trị tại nhiều khu vực trên thế giới, đặc biệt là tại các khu vực gió mùa. Do đó, trong mục này, luận án sẽ tiến hành đánh giá khả năng dự báo NBĐMM bằng phương pháp động lực dựa trên sản phẩm mưa của S2S-ECMWF cho 5 vùng khí hậu (R1, R2, R3, R6 và R7), nơi mà các đặc trưng mưa có liên quan chặt chẽ với hoạt động của gió mùa, khi sử dụng 5 bộ chỉ tiêu xác định NBĐMM khác nhau (chi tiết xem tại mục 2.5.2).

4.4.1. Tính bất định của ngày bắt đầu mùa mưa dự báo nội mùa từ sản phẩm mưa của S2S-ECMWF

Hình 4.9 dẫn ra kết quả tính chỉ số SIM đối với NBĐMM tại các hạn dự báo 7-40 ngày khi sử dụng 5 chỉ tiêu xác định NBĐMM khác nhau. Có thể thấy giá trị của SIM dao động trong khoảng từ 1~12 ngày và phụ thuộc vào cả hạn dự báo và chỉ tiêu sử dụng. Với mọi chỉ tiêu xác định NBĐMM, giá trị của chỉ số SIM có xu hướng tăng lên theo hạn dự báo, rõ nhất ở vùng R7. Tại R3, nơi SIM dao động nhẹ xung quanh một giá trị nhất định trong tất cả các hạn dự báo. Đối với hầu hết các hạn dự báo tại tất cả các vùng khí hậu, SIM nhận giá trị nhỏ hơn khi sử dụng các chỉ tiêu xác định NBĐMM với ngưỡng P nhỏ (P20, P25). Khi tăng ngưỡng P cao hơn (P40 và P50), giá trị của SIM tăng lên, nằm trong khoảng từ 5 đến 12 ngày nhưng khác biệt không nhiều giữa các hạn dự báo. Giá trị SIM lớn thể hiện kết quả dự báo NBĐMM cho một vùng khí hậu có sự khác biệt lớn giữa các thành phần dự báo tại cùng một hạn dự báo. Do đó, việc sử dụng giá trị tổ hợp (trung vị của kết quả dự báo từ 11 thành phần, sau đây gọi là dự báo tổ hợp), thay vì kết quả dự báo của từng thành phần riêng lẻ nên được xem xét, bởi nó sẽ giúp giảm bớt ảnh hưởng của các giá trị cực trị (quá nhỏ hoặc quá lớn).

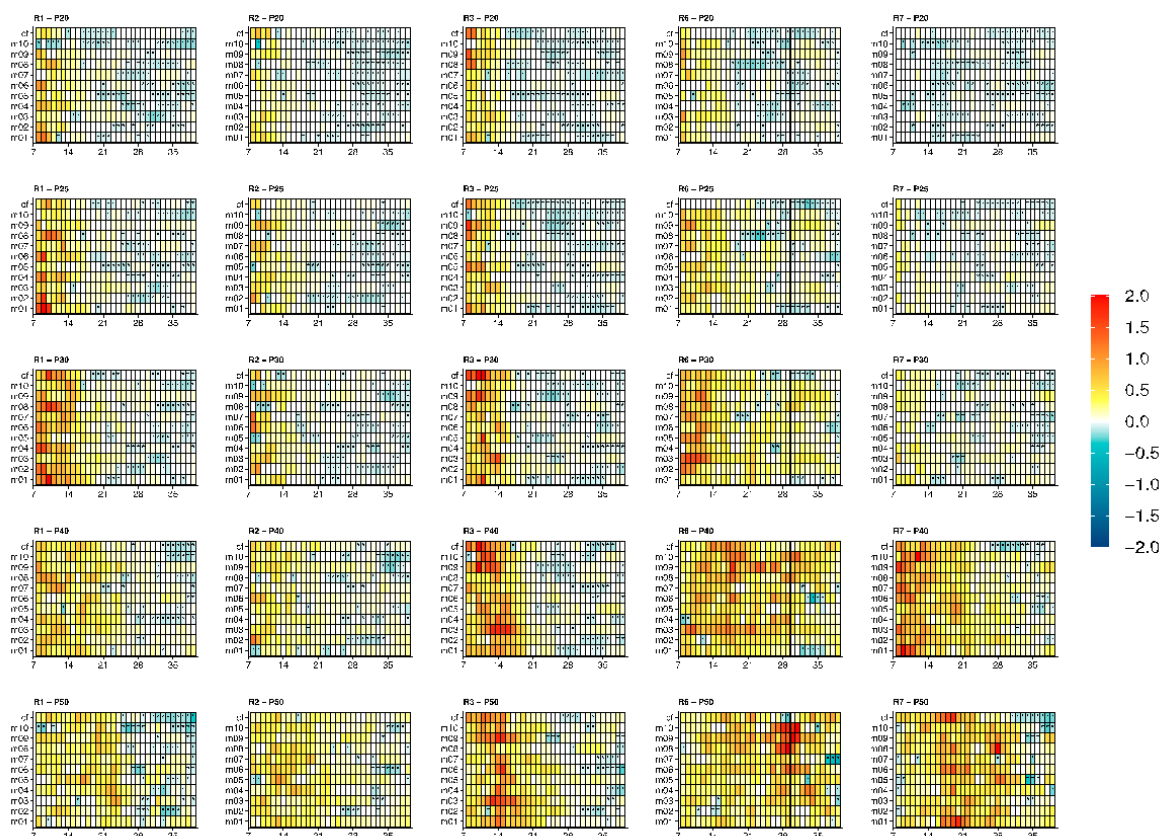
NBĐMM dự báo phụ thuộc đồng thời vào cả độ chính xác của mưa dự báo bởi mô hình và chỉ tiêu xác định NBĐMM. Trong một số trường hợp, khi sử dụng các chỉ tiêu P40, P50 và/hoặc hạn dự báo quá xa (trên 28 ngày), các thành phần dự báo có thể không đưa ra được kết quả dự báo NBĐMM (xem Phụ lục - Hình 13). Khi đó, việc sử dụng dự báo tổ hợp sẽ cho phép dự báo được NBĐMM cho số năm ít hơn 20 năm (giai đoạn 2000-2019).



Hình 4.9: Chỉ số SIM (ngày) cho các hạn dự báo khác nhau tại 5 vùng khí hậu khảo sát (R1, R2, R3, R6 và R7) dựa trên các chỉ tiêu xác định NBĐMM khác nhau

Hình 4.10 dẫn ra kết quả tính chỉ số giá trị gia tăng AV cho dự báo tổ hợp khi sử dụng 11 thành phần với hạn dự báo 7-40 ngày cho 5 vùng khí hậu. AV phản ánh sự cải thiện của dự báo tổ hợp so với dự báo từ các thành phần riêng lẻ. Nhìn chung, giá trị AV dao động tản mạn, phụ thuộc vào cả hạn dự báo lẫn chỉ tiêu xác định NBĐMM. Tuy nhiên, đối với các dự báo hạn ngắn, giá trị AV dương (+) chiếm ưu thế ở tất cả các vùng (R1, R2, R3, R6 và R7) khi sử dụng các chỉ tiêu xác định NBĐMM khác nhau. AV có xu hướng nhận giá trị cao tại các dự báo hạn ngắn và các chỉ tiêu có ngưỡng P cao (P40, P50).

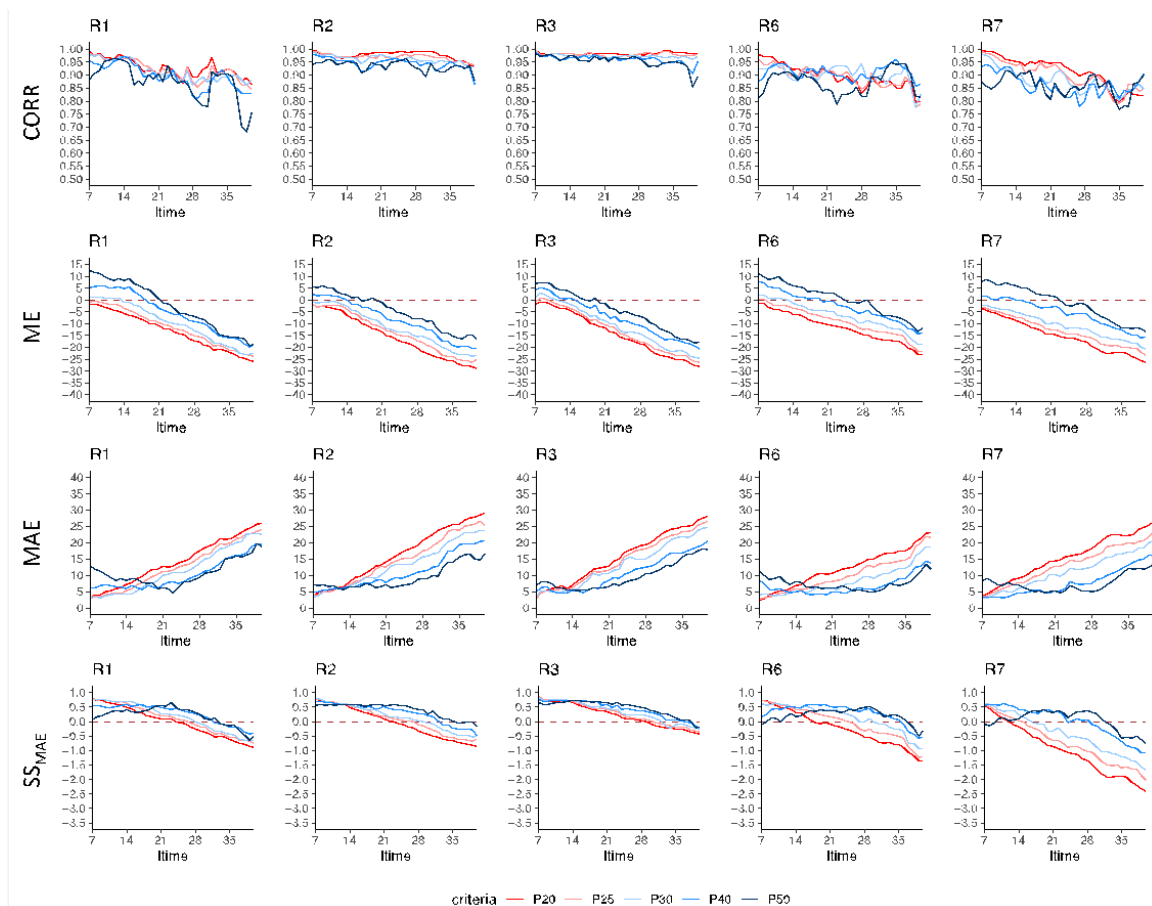
Với những ưu điểm đã nêu của dự báo tổ hợp, là giá trị trung vị của NBĐMM dự báo từ 11 thành phần, trong nội dung tiếp theo luận án sẽ tiến hành đánh giá khả năng dự báo hạn nội mùa NBĐMM khi sử dụng trực tiếp sản phẩm mưa dự báo của S2S-ECMWF.



Hình 4.10: Chỉ số AV cho các hạn dự báo khác nhau (7-40 ngày) tại 5 vùng khí hậu khảo sát (R1, R2, R3, R6 và R7) dựa trên các chỉ tiêu xác định NBĐMM khác nhau

4.4.2. Khả năng dự báo hạn nội mùa (7-40 ngày) ngày bắt đầu mùa mưa dựa trên sản phẩm mưa của S2S-ECMWF

Hình 4.11 dẫn ra kết quả tính các chỉ số CORR, ME, MAE, SS_{MAE} cho giai đoạn 2000-2019 với 5 chỉ tiêu xác định NBĐMM khác nhau.



Hình 4.11: Các chỉ số đánh giá chất lượng dự báo NBĐMM khi sử dụng sản phẩm mưa của S2S-ECMWF, bao gồm CORR (hàng 1), ME (hàng 2, đường kẻ ngang màu nâu thể hiện giá trị = 0), MAE (hàng 3), SS_{MAE} (hàng 4, đường kẻ ngang màu nâu thể hiện giá trị = 0) tại hạn dự báo 7 đến 40 ngày tại 5 vùng khí hậu (R1, R2, R3, R6 và R7), với các chỉ tiêu xác định NBĐMM khác nhau

Giá trị CORR (hàng trên cùng của Hình 4.11) cho thấy NBĐMM dự báo nắm bắt tốt sự biến động hàng năm của NBĐMM quan trắc với hạn dự báo 7-40 ngày. Hệ số tương quan CORR giữa NBĐMM dự báo và NBĐMM quan trắc khá cao, nằm trong khoảng từ 0.60 đến 0.99 (thoả mãn độ tin cậy 95%) tại tất cả các vùng (R1, R2, R3, R6 và R7) và cho tất cả các chỉ tiêu (P20, P25, P30, P40 và P50). Về cơ bản, giá trị CORR có xu hướng giảm đi khi hạn dự báo tăng lên. Việc lựa chọn chỉ tiêu xác định NBĐMM có ảnh hưởng rõ rệt đến NBĐMM dự báo về khả năng nắm bắt sự biến động giữa các năm của NBĐMM quan trắc. Điều này được thể hiện khá rõ tại các vùng R1 và R2. CORR có xu hướng nhận các giá trị cao hơn với các chỉ tiêu có

ngưỡng P nhỏ hơn (ví dụ P20 và P30). Tại hầu hết các hạn dự báo, CORR là thấp nhất khi sử dụng chỉ tiêu P50. Tuy nhiên, ở R6 và R7, trong một số dự báo với hạn trên 30 ngày, việc sử dụng chỉ tiêu có ngưỡng P lớn (ví dụ P40, P50) cho giá trị CORR cao hơn so với các chỉ tiêu khác. Trên thực tế, khi sử dụng chỉ tiêu có ngưỡng P nhỏ (ví dụ P20) thì một đợt mưa bất kỳ có thể được coi là NBĐMM. Do đó, trong một số trường hợp, đặc biệt tại các hạn dự báo dài đối với các vùng R6 và R7, NBĐMM dự báo xuất hiện vào những ngày đầu của chuỗi 46 ngày dự báo sẽ có thể không sát với NBĐMM thực tế. Sai số của dự báo trong các trường hợp này có thể làm tăng biên độ dao động hàng năm của NBĐMM dự báo, hoặc có thể làm cho dấu của dị thường NBĐMM giữa dự báo và quan trắc trái ngược nhau mà hệ quả là tương quan giữa NBĐMM dự báo và NBĐMM quan trắc bị giảm xuống.

Hình 4.11 cho thấy giá trị của ME dao động khá lớn, trong khoảng từ -35 (sớm hơn so với quan trắc) tới +11 ngày (muộn hơn so với quan trắc). Giá trị của ME phụ thuộc vào hạn dự báo và chỉ tiêu xác định NBĐMM. Đối với các chỉ tiêu P20, P25 và P30, NBĐMM dự báo có xu hướng sớm hơn so với thực tế tại hầu hết các hạn dự báo. Trong khi đó, đối với chỉ tiêu P40 và P50, giá trị ME phụ thuộc vào hạn dự báo và có sự đổi dấu từ dương (+) tại hạn dự báo ngắn (dưới 21 ngày) sang âm (-) tại hạn dự báo dài hơn. Trên tất cả các vùng khí hậu, đối với một hạn dự báo, sự khác biệt của ME giữa các chỉ tiêu xác định NBĐMM dao động trong khoảng từ 5 đến 15 ngày. Các chỉ tiêu sử dụng ngưỡng P cao (ví dụ P40 và P50) có xu thế cho giá trị NBĐMM dự báo muộn hơn so với các chỉ tiêu còn lại (tức là P20, P25 và P30). Đáng chú ý, NBĐMM dự báo có xu hướng đến sớm hơn khi hạn dự báo càng dài. Đối với hạn dự báo dưới 20 ngày, giá trị tuyệt đối của ME là dưới 7 ngày, trong khi đó giá trị này có thể lên tới trên 20 ngày với các hạn dự báo trên 30 ngày.

Kỹ năng dự báo NBĐMM phụ thuộc nhiều vào hạn dự báo. Trên cả 5 vùng khí hậu (R1, R2, R3, R6 và R7), với hạn càng dài thì kỹ năng dự báo NBĐMM càng kém, được phản ánh bởi MAE tăng lên theo hạn dự báo (Hình 4.11) với giá trị MAE dao động trong khoảng từ 0 đến 30 ngày. Đáng lưu ý là giá trị MAE xấp xỉ bằng giá trị tuyệt đối của ME trong hầu hết các trường hợp khảo sát ở các vùng khí hậu và đối

với các chỉ tiêu xác định NBĐMM. Điều đó cho phép nhận định về khả năng cải thiện chất lượng dự báo NBĐMM hạn nội mùa bằng một hiệu chỉnh đơn giản.

4.4.3. Sự phụ thuộc của ngày bắt đầu mùa mưa dự báo theo hạn dự báo của S2S-ECMWF

Kỹ năng dự báo NBĐMM khi sử dụng sản phẩm mưa của S2S-ECMWF so với dự báo khí hậu được thể hiện thông qua giá trị SS_{MAE} (Hình 4.11). Có thể thấy, giá trị SS_{MAE} giảm khi hạn dự báo tăng lên. SS_{MAE} mang giá trị dương (âm) tương ứng với dự báo bằng mô hình số có kỹ năng (không có kỹ năng) so với dự báo khí hậu. Qua đó thấy rằng việc sử dụng sản phẩm mưa của S2S-ECMWF cho kết quả dự báo NBĐMM tốt hơn so với dự báo khí hậu trên cả 5 vùng (R1, R2, R3, R6 và R7) với hạn dự báo lên tới 30 ngày (Hình 4.11). Kỹ năng dự báo của mô hình là tốt hơn ở các hạn dự báo từ 7~14 ngày, với giá trị SS_{MAE} đạt trên 0.5. Đối với các dự báo hạn dưới 14 ngày, các chỉ tiêu với ngưỡng P thấp (ví dụ P20 và P30) cho kỹ năng cao hơn. Trong khi đó, với các dự báo hạn trên 14 ngày các chỉ tiêu có ngưỡng P cao (ví dụ P40 và P50) lại cho kỹ năng tốt hơn. Như vậy, việc lựa chọn chỉ tiêu phù hợp để xác định NBĐMM là hết sức quan trọng vì nó có ảnh hưởng đáng kể tới chất lượng dự báo.

Mặc dù NBĐMM dự báo dựa trên sản phẩm mưa của S2S-ECMWF đã có thể nắm bắt tốt NBĐMM quan trắc ở tất cả vùng khí hậu (R1, R2, R3, R6 và R7) với hạn dự báo từ 7 đến 40 ngày, chất lượng dự báo nói chung phụ thuộc nhiều vào hạn dự báo và vào chỉ tiêu xác định NBĐMM. Về cơ bản kết quả dự báo NBĐMM từ sản phẩm S2S-ECMWF cho là tốt hơn so với dự báo khí hậu tại các hạn dự báo từ 1~6 tuần. Trong phần này, các chỉ tiêu xác định NBĐMM cho kết quả dự báo tốt nhất (dựa vào giá trị SS_{MAE}) tại hạn dự báo 7, 14, 21, 28, 35 và 40 ngày sẽ được chọn để đánh giá định lượng kỹ năng dự báo NBĐMM cho từng vùng (bao gồm R1, R2, R3, R6 và R7). Giá trị dự báo NBĐMM là tổ hợp từ 11 dự báo thành phần, được xác định bằng trung vị của các thành phần này.

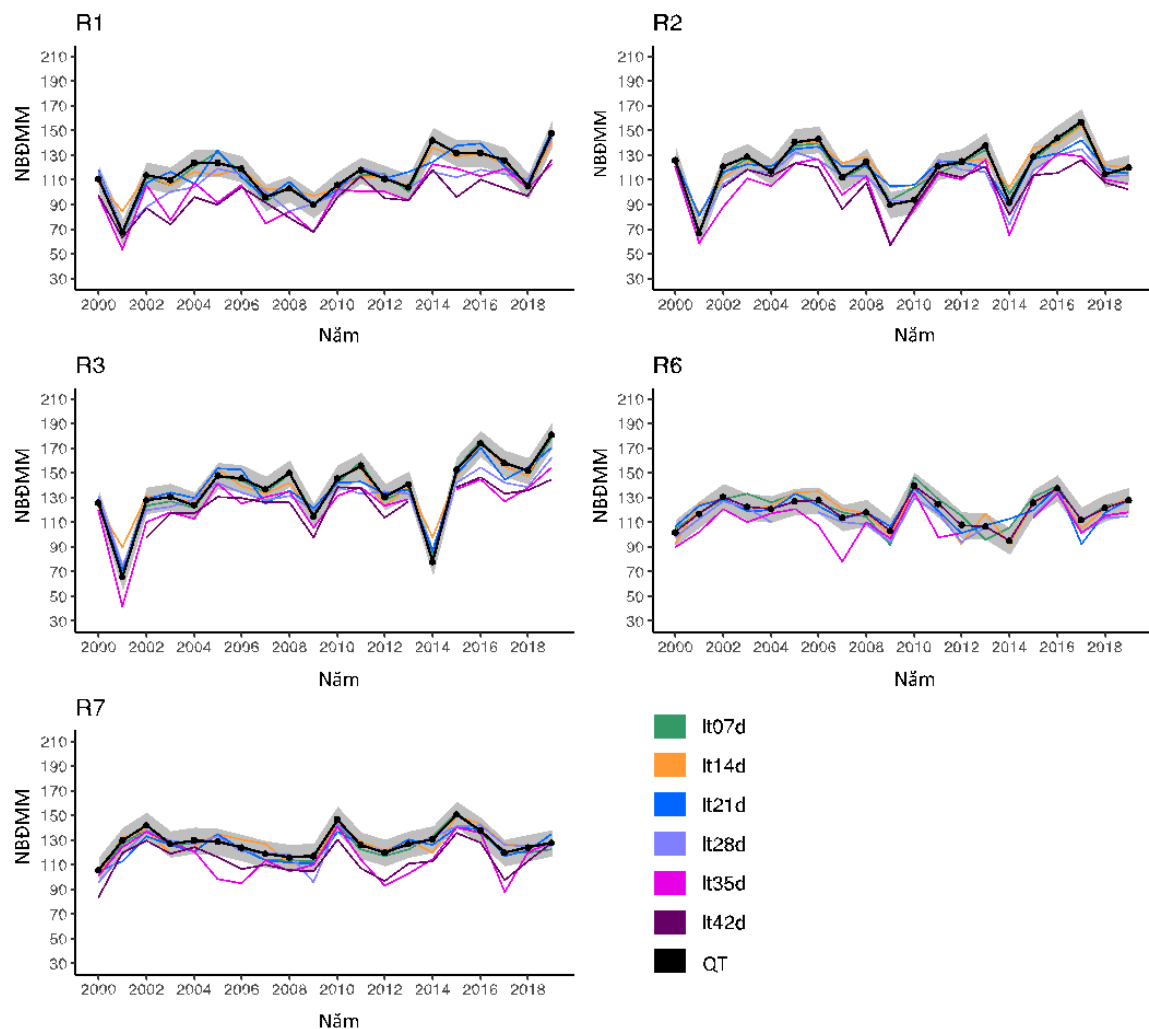
Bảng 4.3 cho thấy các chỉ tiêu xác định NBĐMM cho kết quả dự báo có điểm kỹ năng SS_{MAE} cao nhất ứng với các hạn dự báo 7-14-21-28-35-40 ngày. Có thể thấy

rằng các chỉ tiêu với ngưỡng P nhỏ (P20 và P30) là phù hợp hơn với các hạn dự báo ngắn, trong khi các chỉ tiêu sử dụng ngưỡng P lớn hơn (P40 và P50) sẽ phù hợp hơn với các hạn dự báo dài hơn.

Bảng 4.3: Chỉ tiêu xác định NBĐMM có kết quả dự báo với SS_{MAE} cao nhất ứng với các hạn dự báo 7-14-21-28-35-40 ngày

Hạn dự báo						
Vùng	7 ngày	14 ngày	21 ngày	28 ngày	35 ngày	40 ngày
R1	P25	P30	P50	P50	P50	P40
R2	P30	P50	P50	P50	P50	P50
R3	P25	P40	P50	P50	P50	P40
R6	P20	P40	P40	P50	P50	P50
R7	P30	P40	P40	P50	P50	P50

Hình 4.12 thể hiện biến động giữa các năm của NBĐMM QT và dự báo dựa trên S2S-ECMWF giai đoạn 2000-2019 cho các hạn dự báo khác nhau (7-14-21-28-35-40 ngày) tại các vùng khí hậu, với bộ chỉ tiêu sử dụng tương ứng trong Bảng 4.3. Nhìn chung, biến động giữa các năm của NBĐMM QT tại các vùng được nắm bắt tốt bởi NBĐMM dự báo tại tất cả các hạn dự báo. Hệ số tương quan CORR giữa NBĐMM quan trắc và dự báo khá cao, dao động trong khoảng 0.76~0.99 (thoả mãn mức độ tin cậy 95% của kiểm nghiệm t, Bảng 4.4). Giá trị CORR có xu hướng giảm khi hạn dự báo tăng lên. NBĐMM dự báo có xu hướng sớm hơn so với thực tế, được thể hiện bởi các giá trị ME dao động trong khoảng -19 đến +2 ngày. Tuy nhiên, đối với các hạn dự báo dưới 21 ngày, giá trị ME tương đối nhỏ, dao động trong khoảng -3 đến +2 ngày.



Hình 4.12: NBĐMM quan trắc (QT) và dự báo dựa trên sản phẩm S2S-ECMWF tại các hạn dự báo khác nhau trong giai đoạn 2000-2019 tại các vùng khí hậu. Khu nền màu xám cho biết phạm vi của một độ lệch chuẩn của NBĐMM QT

Giá trị MAE dao động trong khoảng 3~19 ngày, và phụ thuộc đáng kể vào hạn dự báo (Bảng 4.4). Hạn dự báo càng dài, giá trị MAE càng lớn, khoảng 3~7 ngày đối với hạn dự báo dưới 21 ngày, khoảng 10 ngày với hạn dự báo 28 ngày, và khoảng 7~19 ngày với hạn dự báo trên 35 ngày. Giá trị MAE tại vùng R6 và R7 dường như ít thay đổi, chỉ dao động trong khoảng 3~5 ngày, ở các hạn dự báo dưới 28 ngày. Điều này cho thấy kỹ năng dự báo NBĐMM cho các vùng R6 và R7 là tốt và ổn định hơn so với các vùng khác (R1, R2 và R3).

Từ Bảng 4.4 có thể thấy, giá trị SS_{MAE} dao động trong khoảng -0.74 đến +0.86, phụ thuộc vào vùng và hạn dự báo. Dự báo NBĐMM có kỹ năng cao hơn tại các hạn dự báo ngắn hơn. Với các dự báo hạn dưới 7 ngày SS_{MAE} đạt trên 0.62 tại tất cả các vùng (R1, R2, R3, R6 và R7), thậm chí đạt tới 0.86 tại R3. Mặc dù tại R6 và R7 các giá trị MAE là nhỏ, nhưng giá trị của SS_{MAE} ở những vùng này lại có phần nhỏ hơn so với các vùng khác. Nhìn chung, SS_{MAE} nhận giá trị dương (+) đối với các hạn dự báo dưới 28 ngày. Đối với hạn dự báo 35 ngày, các giá trị dương (nhỏ) chỉ xuất hiện tại các vùng R2, R3 và R6. Đối với hạn dự báo 40 ngày, SS_{MAE} đều nhận giá trị âm (-), nằm trong khoảng từ -0.74 đến -0.20 (tức không có kỹ năng). Điều này có nghĩa là dự báo NBĐMM dựa trên sản phẩm của S2S-ECMWF chỉ cho kết quả tốt hơn dự báo khí hậu tại hạn dự báo dưới 40 ngày.

Bảng 4.4: Một số chỉ số đánh giá kỹ năng dự báo NBĐMM dựa trên sản phẩm S2S-ECMWF tại các hạn dự báo khác nhau (7-14-21-28-35-40 ngày) trong giai đoạn 2000-2019 tại các vùng khí hậu

Vùng	Hạn dự báo	CORR	ME	MAE	SS_{MAE}
R1	7	0.98	-1	3	0.79
	14	0.97	-1	4	0.69
	21	0.91	0	6	0.56
	28	0.83	-7	9	0.32
	35	0.90	-16	16	-0.15
	40	0.83	-19	19	-0.40
R2	7	0.99	-1	3	0.79
	14	0.95	2	6	0.60
	21	0.94	-1	7	0.54
	28	0.93	-8	9	0.42
	35	0.93	-15	15	0.01
	40	0.88	-16	16	-0.20
R3	7	0.99	-1	3	0.86

	14	0.97	1	5	0.74
	21	0.97	-1	6	0.67
	28	0.96	-7	10	0.47
	35	0.95	-16	16	0.12
	40	0.89	-18	18	-0.21
R6	7	0.98	-2	3	0.74
	14	0.91	2	5	0.50
	21	0.91	0	5	0.51
	28	0.82	-1	5	0.50
	35	0.94	-7	7	0.18
	40	0.82	-12	12	-0.35
R7	7	0.98	-2	3	0.62
	14	0.89	0	5	0.41
	21	0.84	-3	5	0.33
	28	0.85	-2	5	0.37
	35	0.76	-12	12	-0.57
	40	0.90	-13	13	-0.74

Nhận xét cuối chương 4

Trong chương này luận án đã trình bày việc đánh giá khả năng dự báo hạn mùa và nội mùa NBĐMM cho các vùng khí hậu Việt Nam theo 4 trường hợp khác nhau. Kết quả đánh giá được tổng hợp lại trong Bảng 4.5. Qua đó nhận thấy chất lượng dự báo NBĐMM của các phương pháp và hạn dự báo là khác nhau giữa các vùng khí hậu.

Với cách tiếp cận theo hướng thống kê, bao gồm cả thống kê truyền thống và thống kê động lực, mối quan hệ giữa NBĐMM tại từng vùng với các nhân tố dự báo đóng vai trò quyết định đến khả năng áp dụng cũng như chất lượng của các phương trình dự báo. Các kết quả thu được cho thấy:

- Đối với phương pháp thống kê truyền thống sử dụng các chỉ số khí hậu làm nhân tố dự báo, kết quả đánh giá DBPT và DBĐL tại 7 vùng khí hậu đều cho thấy về cơ bản các mô hình thống kê dự báo NBĐMM thu được là tốt hơn so với dự báo khí hậu, được thể hiện thông qua giá trị dương (+) của điểm kỹ năng SS_{MAE} . Tuy nhiên, giá trị SS_{MAE} có sự khác biệt giữa các vùng: tại R6 và R7, SS_{MAE} dao động trong khoảng +0.25 đến +0.50, cao hơn so với 5 vùng khí hậu còn lại (dưới +0.17) trong cả DBPT và DBĐL. Nghĩa là kỹ năng dự báo của mô hình thống kê truyền thống so với dự báo khí hậu ở R6 và R7 là cao hơn so với các vùng khí hậu còn lại. Các phương trình dự báo có độ ổn định cao nên việc áp dụng phương pháp thống kê truyền thống dựa trên chỉ số khí hậu để dự báo NBĐMM cho khu vực R6 và R7 là hoàn toàn khả thi.
- Đối với phương pháp thống kê-động lực, chất lượng dự báo NBĐMM có sự khác biệt giữa các vùng và giữa các hạn dự báo. Nhìn chung, trên cả 3 vùng khí hậu (R3, R6 và R7), các dự báo hạn ngắn (ví dụ 1 tháng) cho kết quả tốt hơn so với các dự báo hạn dài hơn (ví dụ 6 tháng). Dự báo NBĐMM bằng phương pháp thống kê-động lực dựa trên sản phẩm của CFSv2 cho kết quả kém hơn dự báo khí hậu ở R3 và tốt hơn ở R6 và R7. Các phương trình dự báo có độ ổn định cao nên việc áp dụng phương pháp thống kê-động lực để dự báo NBĐMM cho khu vực R6 và R7 là khả thi. Tuy nhiên, việc cải thiện chất lượng dự báo NBĐMM bằng phương pháp thống kê-động lực dựa trên sản phẩm mô hình CFSv2 so với dự báo khí hậu sẽ gặp nhiều thách thức tại các khu vực mà NBĐMM có sự ổn định giữa các năm như tại R7.

Với cách tiếp cận theo hướng sử dụng trực tiếp sản phẩm mưa của mô hình số, luận án đã tiến hành đánh giá khả năng dự báo NBĐMM với hạn mùa (dưới 6 tháng) dựa trên sản phẩm CFSv2 và hạn nội mùa (7-40 ngày) dựa trên sản phẩm ECMWF. Các kết quả thu được cho thấy:

- Khi sử dụng trực tiếp sản phẩm mưa từ CFSv2 kết quả dự báo NBĐMM kém hơn so sánh với dự báo khí hậu tại tất cả các vùng khí hậu của Việt Nam (R1,

R2, R3, R4, R5, R6 và R7). Kỹ năng dự báo NBĐMM bằng phương pháp này là khác nhau giữa các vùng khí hậu. NBĐMM dự báo về cơ bản chưa nắm bắt được sự biến động giữa các năm của NBĐMM quan trắc tại phần lớn các vùng khí hậu, trừ R7. Do mối quan hệ giữa NBĐMM quan trắc và dự báo thiếu chặt chẽ nên việc hiệu chỉnh kết quả dự báo NBĐMM sẽ gặp nhiều khó khăn. Ngoài sự phụ thuộc vào chất lượng dự báo mưa của mô hình, việc lựa chọn chỉ tiêu phù hợp để xác định NBĐMM từ sản phẩm mưa của CFSv2 đóng vai trò quan trọng đối với chất lượng dự báo NBĐMM. Nói chung, việc lựa chọn chỉ tiêu xác định NBĐMM nên được tiến hành cho từng vùng khí hậu riêng biệt.

- Việc sử dụng trực tiếp sản phẩm mưa của S2S-ECMWF để dự báo NBĐMM trên các vùng khí hậu R1, R2, R3, R6 và R7 cho thấy các kết quả dự báo nắm bắt tốt sự thay đổi hàng năm của NBĐMM quan trắc với hạn dự báo lên tới 40 ngày. NBĐMM dự báo có xu hướng sớm hơn so với quan trắc khi tăng hạn dự báo. Kỹ năng dự báo NBĐMM bằng phương pháp này là tốt hơn so với dự báo khí hậu với hạn dự báo đến 28~35 ngày, phụ thuộc vào từng khu vực. Tuy nhiên, các dự báo hạn ngắn cho kết quả có phần tốt hơn so với các dự báo hạn dài. Việc lựa chọn chỉ tiêu xác định NBĐMM cũng ảnh hưởng đáng kể đến kỹ năng dự báo của mô hình. Kết quả phân tích giá trị SS_{MAE} cho các vùng khí hậu chỉ ra rằng, các chỉ tiêu sử dụng ngưỡng P thấp (ví dụ P20, P25, P30) phù hợp hơn ở các dự báo hạn ngắn hơn, ngược lại, các ngưỡng P40 và P50 lại có xu hướng phù hợp hơn đối với các dự báo hạn dài hơn. Tuy nhiên, các giá trị của MAE, ME, CORR cho thấy có thể cải thiện chất lượng dự báo bằng một hiệu chỉnh đơn giản. Do đó, với các kết quả nhận được, việc sử dụng sản phẩm của S2S-ECMWF để dự báo NBĐMM với hạn nội mùa cho các khu vực R1, R2, R3, R6 và R7 là khả thi.

Bảng 4.5: Thống kê khả năng dự báo NBĐMM tại từng vùng khí hậu với các phương pháp dự báo và hạn dự báo khác nhau. Kỹ năng dự báo được dựa trên chỉ số SS_{MAE} , theo đó dự báo có kỹ năng khi $SS_{MAE} > 0$ và không có kỹ năng khi

$$SS_{MAE} < 0$$

Vùng khí hậu Phương pháp dự báo	R1	R2	R3	R4	R5	R6	R7
1. Phương pháp thống kê truyền thống dự báo hạn mùa (dưới 6 tháng) NBĐMM dựa trên các chỉ số khí hậu	Có kỹ năng (DBPT)	Không có kỹ năng	Có kỹ năng (DBPT & DBĐL)	Có kỹ năng (DBPT & DBĐL)	Có kỹ năng (DBPT)	Có kỹ năng (DBPT & DBĐL)	Có kỹ năng (DBPT & DBĐL)
2. Phương pháp thống kê-động lực dự báo hạn mùa (dưới 6 tháng) NBĐMM dựa trên sản phẩm CFSv2			Không có kỹ năng			Có kỹ năng (DBPT & DBĐL)	Có kỹ năng (DBPT)
3. Phương pháp sử dụng trực tiếp sản phẩm mưa của CFSv2 để dự báo hạn mùa (dưới 6 tháng) NBĐMM	Không có kỹ năng	Không có kỹ năng	Không có kỹ năng	Không có kỹ năng	Không có kỹ năng	Không có kỹ năng	Không có kỹ năng
4. Phương pháp sử dụng trực tiếp sản phẩm mưa của S2S-ECMWF để dự báo hạn nội mùa (7-40 ngày) NBĐMM	Có kỹ năng dự báo	Có kỹ năng dự báo	Có kỹ năng dự báo			Có kỹ năng dự báo	Có kỹ năng dự báo

KẾT LUẬN

Dự báo NBĐMM là một trong những bài toán quan trọng của ngành khí tượng thủy văn. Tuy nhiên NBĐMM không phải là biến được quan trắc trực tiếp mà chỉ

được xác định một cách gián tiếp dựa trên các bộ chỉ tiêu nào đó. Vì vậy kết quả dự báo NBĐMM hàm chứa tính bất định cao. Mặc dù vậy, đã có nhiều công trình nghiên cứu giải quyết bài toán này vì kết quả của nó được ứng dụng trong nhiều lĩnh vực kinh tế xã hội như phục vụ sản xuất nông nghiệp, quản lý tài nguyên nước, v.v. Trong khuôn khổ luận án này hai vấn đề chính đã được đặt ra và giải quyết là: 1) Xác định được NBĐMM một cách khách quan và chỉ ra được một số đặc điểm phân bố theo không gian và biến động theo thời gian của NBĐMM trên các vùng khí hậu Việt Nam dựa trên bộ số liệu quan trắc tại 131 trạm trong giai đoạn 1979-2019; 2) Xây dựng được các phương án dự báo NBĐMM hạn nội mùa và mùa cho các vùng khí hậu Việt Nam bằng các phương pháp khác nhau. Kết quả thu được của luận án cho phép rút ra một số kết luận sau:

- Việc xác định NBĐMM phụ thuộc nhiều vào việc lựa chọn chỉ tiêu. Vai trò của các tham số P, N và D mô tả sự thay đổi của các đặc trưng mưa trong giai đoạn chuyển tiếp từ mùa khô sang mùa mưa là khác nhau giữa các vùng khí hậu cũng như giữa các trạm trong cùng một vùng. Tham số P có thể được coi là tham số quan trọng nhất, được thể hiện bởi 66.4% tổng số trạm trong đó NBĐMM thay đổi đáng kể khi điều chỉnh ngưỡng của tham số này. Do đó khi xây dựng chỉ tiêu xác định NBĐMM đối với khu vực Việt Nam cần đặc biệt lưu ý khi lựa chọn giá trị của tham số này. Với bộ tham số được lựa chọn trong luận án, NBĐMM trên khu vực Việt Nam có sự khác biệt giữa các vùng khí hậu và giữa các trạm trong cùng một vùng. NBĐMM xảy ra sớm nhất ở phía tây bắc của khu vực R123, vào khoảng nửa đầu tháng 4 (trước thời kỳ GMMH thịnh hành). NBĐMM ở R123 và R67 xảy ra vào khoảng giữa tháng 4 đến giữa tháng 5. Tuy nhiên, sự biến động giữa các năm của NBĐMM tại R123 lớn hơn (khoảng 21-35 ngày) so với ở R67 (12 ngày). Đối với R45, NBĐMM có sự khác biệt lớn giữa các trạm, sớm nhất vào ngày 1 tháng 7 (phần phía bắc) và muộn nhất vào ngày 13 tháng 9 (phần phía nam). Điều đó có thể liên quan tới điều kiện địa lý và cơ chế mưa trong thời kỳ chuyển tiếp từ mùa khô sang mùa mưa ở khu vực này. Sự biến động giữa các năm của NBĐMM tại R45 cũng khá lớn, với độ lệch chuẩn khoảng 16-18 ngày. Sự biến động hàng năm của NBĐMM tại các vùng khí hậu R3, R6 và R7 có mối liên hệ chặt chẽ

với hoạt động của ENSO, theo đó NBĐMM có xu thế đến muộn (sớm) trong những năm El Niño (La Niña). Xu thế biến đổi của NBĐMM giữa các vùng khí hậu và các trạm là khác nhau. Nói chung xu thế biến đổi của NBĐMM không rõ rệt nhưng có dấu hiệu đến sớm hơn tại khu vực R45 và R67 và muộn hơn ở R123. Chỉ một số trạm NBĐMM thể hiện rõ xu thế, chẳng hạn NBĐMM đến sớm hơn từ 3-8 ngày/thập kỷ tại phía bắc của R4, phía đông nam của R6 và đến muộn hơn 5-10 ngày/thập kỷ tại phía đông của R3.

- Việc dự báo hạn mùa NBĐMM bằng phương pháp thống kê truyền thống sử dụng các chỉ số khí hậu làm nhân tố dự báo phần lớn cho kết quả tốt hơn dự báo khí hậu tại cả 7 vùng khí hậu, đặc biệt là đối với R6 và R7. Kết quả đánh giá trên các tập số liệu phụ thuộc và độc lập cho thấy tiềm năng của việc áp dụng phương pháp thống kê truyền thống này vào dự báo hạn mùa NBĐMM cho các vùng R6 và R7.
- Dự báo hạn mùa NBĐMM bằng phương pháp thống kê-động lực sử dụng các trường dự báo của CFSv2 chỉ áp dụng được đối với ba vùng R3, R6 và R7. Chi phối mạnh mẽ đến sự biến đổi hàng năm của NBĐMM tại các vùng khí hậu này là các trường SST, PRMSL, U850. Trên cả 3 vùng khí hậu R3, R6 và R7 kết quả dự báo có phần tốt hơn đối với hạn dự báo ngắn hơn. Kỹ năng dự báo của mô hình kém hơn dự báo khí hậu ở R3 và tốt hơn dự báo khí hậu ở R6 và R7. Việc kiểm chứng trên tập độc lập, và khảo sát mức độ ổn định của phương trình dự báo cho thấy tính khả thi của việc áp dụng phương pháp thống kê-động lực để dự báo NBĐMM cho các vùng R6 và R7.
- Dự báo NBĐMM sử dụng trực tiếp sản phẩm mưa từ CFSv2 cho kết quả kém hơn khi so sánh với dự báo khí hậu tại tất cả các vùng khí hậu Việt Nam (R1, R2, R3, R4, R5, R6 và R7). Kỹ năng dự báo NBĐMM bằng phương pháp này là khác nhau giữa các vùng khí hậu. Kết quả dự báo với hạn dưới 6 tháng về cơ bản chưa nắm bắt được sự biến động giữa các năm của NBĐMM quan trắc tại hầu hết các vùng khí hậu, trừ R7. Ngoài chất lượng dự báo mưa của CFSv2, việc lựa chọn chỉ tiêu phù hợp để xác định NBĐMM có vai trò quan trọng khi sử dụng trực tiếp sản phẩm mưa dự báo của CFSv2. Việc lựa chọn chỉ tiêu này nên được tiến hành cho từng vùng khí hậu riêng biệt. Một trong những nguyên

nhân làm cho kết quả dự báo NBĐMM kém còn có thể do độ phân giải của mô hình quá thấp (1.0 độ).

- Dự báo hạn nội mùa NBĐMM từ phẩm mưa của S2S-ECMWF nắm bắt tốt sự biến động hàng năm của NBĐMM quan trắc với hạn dự báo lên tới 40 ngày trên các vùng khí hậu khảo sát (R1, R2, R3, R6 và R7). NBĐMM dự báo có xu hướng sớm hơn so với thực tế khi tăng hạn dự báo. Kỹ năng dự báo NBĐMM bằng phương pháp này là tốt hơn so với dự báo khí hậu với hạn dự báo lên tới 28~35 ngày, phụ thuộc vào từng khu vực. Tuy nhiên, các dự báo hạn ngắn cho kết quả dự báo có phần tốt hơn so với các dự báo hạn dài. Việc lựa chọn chỉ tiêu xác định NBĐMM cũng có ảnh hưởng đáng kể đến kỹ năng dự báo. Kết quả phân tích giá trị SS_{MAE} cho thấy, các chỉ tiêu sử dụng ngưỡng P thấp (ví dụ P20, P25, P30) phù hợp cho việc xác định NBĐMM tại các dự báo hạn ngắn, ngược lại, các ngưỡng P40 và P50 lại có xu hướng phù hợp hơn đối với các dự báo hạn dài. Việc phân tích các chỉ số MAE, ME, CORR cho phép nhận định rằng có thể cải thiện chất lượng dự báo bằng một hiệu chỉnh đơn giản. Từ các kết quả nhận được có thể nói việc sử dụng trực tiếp sản phẩm mưa dự báo của S2S-ECMWF để dự báo NBĐMM với hạn nội mùa cho các khu vực R1, R2, R3, R6 và R7 là khả thi.

KIẾN NGHỊ

NBĐMM đóng vai trò quan trọng trong nhiều lĩnh vực của đời sống kinh tế-xã hội do vậy việc nghiên cứu, phát triển công cụ dự báo mang ý nghĩa thực tiễn cao. Để có thể ứng dụng những kết quả của luận án vào công tác dự báo nghiệp vụ cần thiết phải tiếp tục hoàn thiện theo một số hướng sau:

- Phương pháp thống kê truyền thống sử dụng các chỉ số khí hậu làm nhân tố dự báo là phương pháp đơn giản, dễ sử dụng nhưng cho kết quả dự báo khá tốt tại vùng R6 và R7. Do đó, phương pháp này nên được xem xét để thử nghiệm ứng dụng trong nghiệp vụ. Trong quá trình thử nghiệm có thể khảo sát thêm các chỉ số khí hậu khác cũng như có thể sử dụng các mô hình thống kê khác, như ANN, LAD,...
- Phương pháp thống kê-động lực dự báo hạn mùa NBĐMM sử dụng các trường dự báo của CFSv2 có thể là cách tiếp cận tiềm năng và có thể cho kết quả dự báo tốt. Tuy nhiên, để ứng dụng phương pháp này cần có sự khảo sát đầy đủ, chi tiết hơn mối quan hệ giữa yếu tố dự báo với các trường khác nữa. Trong phạm vi luận án, phương pháp thống kê-động lực thể hiện khả năng ứng dụng cho các vùng R6 và R7.
- Việc sử dụng trực tiếp sản phẩm mưa của S2S-ECMWF để dự báo hạn nội mùa (7-40 ngày) NBĐMM là rất khả quan. Do đó nên tiến hành thử nghiệm ứng dụng nghiệp vụ càng sớm càng tốt.

DANH MỤC CÔNG TRÌNH KHOA HỌC CỦA TÁC GIẢ LIÊN QUAN ĐẾN LUẬN ÁN

- 1) Phan Văn Tân, Phạm Thanh Hà, Nguyễn Đăng Quang, Nguyễn Văn Hiệp, Ngô Đức Thành, (2016): “Sự biến đổi của ngày bắt đầu mùa mưa ở Tây nguyên và khả năng dự báo”, *Tạp chí Khoa học ĐHQGHN: Các Khoa học Trái đất và Môi trường*, Tập 32, Số 3S, tr.1-18.
- 2) Pham-Thanh, H., van der Linden, R, Ngo-Duc, T, Nguyen-Dang, Q, Fink, AH, Phan-Van, T., (2019): “Predictability of the rainy season onset date in Central Highlands of Vietnam”, *International Journal of Climatology*, 40, 3072- 3086. <https://doi.org/10.1002/joc.6383>.
- 3) Pham-Thanh, H., T. Phan-Van, A. H. Fink, and R. van der Linden, (2021): “Local-Scale Rainy Season Onset Detection: A New Approach Based on Principal Component Analysis and its Application to Vietnam”, *International Journal of Climatology*, doi: 10.1002/joc.7441.
- 4) Pham-Thanh, H., T. Phan-Van, R. van der Linden, and A. H. Fink, (2021): “The Performance of ECMWF sub-seasonal forecasts to predict the Rainy Season Onset Dates in Vietnam”, *Weather and Forecasting*, doi: 10.1175/WAF-D-21-0144.1.
- 5) Phạm Thanh Hà, Phan Văn Tân, (2022): “Xây dựng chỉ tiêu khách quan xác định ngày bắt đầu mùa mưa cho khu vực Tây Nguyên-Nam Bộ”, *Tạp chí Khoa học ĐHQGHN: Các Khoa học Trái đất và Môi trường*, Tập 38, số 1, tr.85-94.

TÀI LIỆU THAM KHẢO

Tiếng Việt

1. Phạm Thị Châm, Trần Thị Thuý Nga, Hoàng Đức Cường, Nguyễn Đăng Quang (2017), “Nghiên cứu ngày bắt đầu gió mùa mùa hè, ngày bắt đầu mùa mưa ở Tây Nguyên và khả năng dự báo thời điểm bắt đầu mùa mưa bằng phương pháp phân tích Canon”, *Tạp chí Khí tượng Thủy văn*, 683, tr. 1-11.
2. Hoàng Đức Cường và cs (2018), “Nghiên cứu xây dựng phương pháp dự báo thời điểm bắt đầu và kết thúc mùa mưa gió mùa mùa hè khu vực Trung Bộ, Tây Nguyên và Nam Bộ”, *Báo cáo tổng hợp đề tài TNMT.2016.05.08*, Bộ Tài Nguyên và Môi Trường, 142 trang.
3. Nguyễn Hương Điền, Trần Công Minh (2000), “Tính chu kỳ và tương quan giữa lượng mưa và độ kéo dài của thời kỳ gió mùa mùa hè trên lãnh thổ Việt Nam”, *Báo cáo tổng kết đề tài mã số QT 98-13*, Đại học Quốc gia Hà Nội.
4. Trần Quang Đức (2010), “Ảnh hưởng của ENSO đến dao động và biến đổi nhiều năm của mưa gió mùa mùa hè trên khu vực Việt Nam”, *Đề tài ĐHQG nhóm B*.
5. Ngô Thị Thanh Hương (2018), *Biến đổi một số đặc trưng gió mùa mùa hè ở Tây Nguyên và Nam Bộ*, Luận án tiến sĩ khí tượng và khí hậu học, Trường Đại học Khoa học Tự nhiên, ĐHQGHN.
6. Chu Thị Thu Hường, Trần Đình Linh (2019), “Đặc điểm hoạt động của gió mùa mùa hè trên lãnh thổ Việt Nam”, *Tạp chí Khí tượng Thủy văn*, 705, tr. 56-63.
7. Lê Thị Xuân Lan (2017), “Đặc điểm mùa mưa ở khu vực Nam Bộ, Đài khí tượng thủy văn khu vực Nam Bộ”, *sách chuyên khảo*, chương II, tr. 16-24.
8. Nguyễn Đăng Mậu, Nguyễn Văn Thắng, Mai Văn Khiêm, Lưu Nhật Linh, Nguyễn Trọng Hiệu (2016), “Nghiên cứu chỉ số gió mùa mùa hè cho khu vực Việt Nam”, *Tạp chí Khí tượng Thủy văn* 2, tr. 1-7.

9. Nguyễn Đức Ngữ, Nguyễn Trọng Hiệu (2004), *Khí hậu và Tài nguyên Khí hậu Việt Nam*, Nhà xuất bản Nông nghiệp Hà Nội, 295 trang.
10. Nguyễn Đức Ngữ, Nguyễn Thị Hiền Thuận (2006), “Đề xuất chỉ số hoàn lưu gió mùa để nghiên cứu tính biến động của gió mùa mùa hè ở Nam Bộ”, *Tạp chí KTTV* 545, tr. 1-10.
11. Ngô Đức Thành, Phan Văn Tân (2012), “Kiểm nghiệm phi tham số xu thế biến đổi của một số yếu tố khí tượng giai đoạn 1961-2007”, *Tạp chí ĐHQGHN, Khoa học Tự nhiên và Công nghệ* 28(3S), tr. 129-135.
12. Nguyễn Thị Hiền Thuận (2007), *Nghiên cứu ảnh hưởng của ENSO đến gió mùa mùa hè và mưa ở Nam Bộ*, Luận án tiến sĩ địa lý, Viện khoa học Khí tượng Thủy văn và Môi trường.
13. Nguyễn Thị Hiền Thuận, Chiêu Kim Quỳnh (2007), “Nhận xét về sự biến động của các đặc trưng mưa mùa hè ở khu vực Nam Bộ trong các năm ENSO”, *Tuyển tập báo cáo Hội thảo khoa học lần thứ 10*, Viện KH KTTV và MT, tr. 314-322.
14. Bùi Minh Tuấn, Nguyễn Minh Trường (2010), “Hoàn lưu qui mô lớn thời kỳ bùng nổ gió mùa mùa hè trên khu vực Nam Bộ năm 1998”, *Tạp chí Khoa học ĐHQG, Khoa học Tự nhiên và Công nghệ* 26 (3S), tr. 470-478.
15. Trần Trung Trực, Phạm Thị Thanh Hương (1999), “Các hình thể synôp xác định quá trình thiết lập gió mùa Tây Nam trên khu vực Tây Nguyên - Nam Bộ”, *Tạp chí Khí tượng Thủy văn* 2/1999, tr. 16-24.
16. Nguyễn Minh Trường và cs (2012), “Đặc điểm hoàn lưu và thời tiết thời kỳ bùng nổ gió mùa mùa hè trên khu vực Việt Nam”, *Báo cáo tổng kết đề tài QG-10-07*, Đại học Quốc gia H. Nội, 64 trang.
17. Phan Văn Tân (2003), *Các phương pháp thống kê trong khí hậu*, NXB ĐHQGHN, Hà Nội, 208 trang.
18. Viện khí tượng thủy văn và môi trường (2015), *Báo cáo đặc biệt của Việt Nam về quản lý rủi ro thiên tai và các hiện tượng cực đoan nhằm thúc đẩy thích ứng với biến đổi khí hậu*, NXB Tài nguyên - môi trường và bản đồ Việt Nam.

Tiếng Anh

19. Agou, V.D., Varouchakis, E.A. and Hristopulos, D.T. (2019), “Geostatistical analysis of precipitation in the Island of Crete (Greece) based on a sparse monitoring network”, *Environmental Monitoring and Assessment*, 191(353), 353. <https://doi.org/10.1007/s10661-019-7462-8>
20. Alessandri, A., Borrelli, A., Cherchi, A., Materia, S., Navarra, A., Lee, J.-Y., et al (2015), “Prediction of Indian summer monsoon onset using dynamical subseasonal forecasts: effects of realistic initialization of the atmosphere”, *Mon. Weather Rev.* 143, 778-793. doi: 10.1175/MWR-D-14-00187.1
21. Alexander, L. V., et al. (2006), “Global observed changes in daily climate extremes of temperature and precipitation”, *J. Geophys. Res.*, 111, D05109, doi:10.1029/2005JD006290.
22. Becker, E., den Dool, H. and Zhang, Q. (2014), “Predictability and forecast skill in NMME”, *Journal of Climate*, 27(15), 5891-5906.
23. Behringer, D. W. (2007), “The Global Ocean Data Assimilation System (GODAS) at NCEP. Preprints”, *11th Symp. on Integrated Observing and Assimilation Systems for Atmosphere, Oceans, and Land Surface*, San Antonio, TX, Amer. Meteor. Soc., 3.3. [Available online at <http://ams.confex.com/ams/pdfpapers/119541.pdf>.]
24. Bombardi, R. J., Pegion, K. V., Kinter, J. L., Cash, B. A., & Adams, J. M. (2017), “Sub-seasonal predictability of the onset and demise of the rainy season over monsoonal regions”, *Frontiers in Earth Science*, 5, 1- 17. <https://doi.org/10.3389/feart.2017.00014>
25. Bombardi, R.J. and Carvalho, L.M.V. (2009), “IPCC global coupled model simulations of the South America monsoon system”, *Climate Dynamics*, 33(7-8), 893-916.
26. Bombardi, RJ, Moron, V, Goodnight, JS (2020), “Detection, variability, and predictability of monsoon onset and withdrawal dates: A review”, *Int J Climatol*, 40: 641- 667, <https://doi.org/10.1002/joc.6264>

27. Boyard-Micheau, J., P. Camberlin, N. Philippon, and V. Moron (2013), “Regional-Scale Rainy Season Onset Detection: A New Approach Based on Multivariate Analysis”, *J. Climate*, 26, 8916–8928, <https://doi.org/10.1175/JCLI-D-12-00730.1>.
28. Cherchi, A. and Navarra, A. (2003), “Reproducibility and predictability of the Asian summer monsoon in the ECHAM4-GCM”, *Climate Dynamics*, 20(4), 365–379.
29. Chevuturi A, Turner AG, Woolnough SJ, Martin GM, MacLachlan C (2019), “Indian summer monsoon onset forecast skill in the UK Met Office initialized coupled seasonal forecasting system (GloSea5-GC2)”, *Clim Dyn*, 52(11), 6599–6617.
30. Davidson, N.E., McBride, J.L. and McAvaney, B.J. (1983), “The onset of the Australian monsoon during winter MONEX: synoptic aspects”, *Monthly Weather Review*, 111(3), 496–516, [https://doi.org/10.1175/1520-0493\(1983\)111<0496:TOOTAM>2.0.CO;2](https://doi.org/10.1175/1520-0493(1983)111<0496:TOOTAM>2.0.CO;2).
31. Dunning, C.M., Black, E.C.L. and Allan, R.P. (2016), “The onset and cessation of seasonal rainfall over Africa”, *Journal of Geophysical Research-Atmospheres*, 121(19), 11405–11424.
32. Endo, N., J. Matsumoto, and T. Lwin (2009), “Trends in precipitation extremes over Southeast Asia”, *SOLA*, 5, 168–171.
33. Fontaine, B., Louvet, S. and Roucou, P. (2008), “Definition and predictability of an OLR-based west African monsoon onset”, *International Journal of Climatology*, 28(13), 1787–1798.
34. Garcia, S.R. and Kayano, M.T. (2009), “Determination of the onset dates of the rainy season in Central Amazon with equatorially antisymmetric outgoing longwave radiation”, *Theoretical and Applied Climatology*, 97(3–4), 361–372.
35. Hendon, H.H. and Liebmann, B. (1990), “A composite study of onset of the Australian summer monsoon”, *Journal of the Atmospheric Sciences*, 47(18),

2227-2240, [https://doi.org/10.1175/1520-0469\(1990\)047<2227:ACSOOO>2.0.CO;2](https://doi.org/10.1175/1520-0469(1990)047<2227:ACSOOO>2.0.CO;2).

36. Holland, G.J. (1986), “Interannual variability of the Australian summer monsoon at Darwin: 1952-82”, *Monthly Weather Review*, 114 (3), 594-604, [https://doi.org/10.1175/1520-0493\(1986\)114<0594:IVOTAS>2.0.CO;2](https://doi.org/10.1175/1520-0493(1986)114<0594:IVOTAS>2.0.CO;2).
37. Hur, J. and Ahn, J.-B. (2015), “Seasonal prediction of regional surface air temperature and first-flowering date over South Korea”, *International Journal of Climatolology*, 35, 4791-4801. <https://doi.org/10.1002/joc.4323>.
38. Jones, C. and Carvalho, L.M.V. (2013), “Climate change in the south American monsoon system: present climate and CMIP5 projections”, *Journal of Climate*, 26(17), 6660-6678.
39. Kajikawa, Y., Yasunari, T., Yoshida, S. and Fujinami, H. (2012), “Advanced Asian summer monsoon onset in recent decades”, *Geophysical Research Letters*, **39**(3), 1033, <https://doi.org/10.1029/2011GL050540>.
40. Kanamitsu, Masao, Wesley Ebisuzaki, Jack Woollen, Shi-Keng Yang, J. J. Hnilo, M. Fiorino, and G. L. Potter (2002), “NCEP-DOE AMIP-II Reanalysis (R-2)”, *Bulletin of the American Meteorological Society*, 83.11, 1631-1644. < <https://doi.org/10.1175/BAMS-83-11-1631>>.
41. Kendall, M.G. (1975), “Rank Correlation Methods”, *Charles Griffin, London*, 272 pp, 1975.
42. Kumi, N., Abiodun, B. J., & Adefisan, E. A. (2020), “Performance evaluation of a Subseasonal to Seasonal model in predicting rainfall onset over West Africa”, *Earth and Space Science*, **7**, e2019EA000928-T. <https://doi.org/10.1029/2019EA000928>
43. Lala, J., Tilahun, S., & Block, P. (2020), “Predicting Rainy Season Onset in the Ethiopian Highlands for Agricultural Planning”, *Journal of Hydrometeorology*, 21(7), 1675-1688. Retrieved Jun 3, 2021, from <https://journals.ametsoc.org/view/journals/hydr/21/7/jhmD200058.xml>

44. Lambert, D.L. (2014), "Historical Impacts of Precipitation and Temperature on Farm production in Kansas", *Journal of Agricultural & Applied Economics*, 46(4), 439-456
45. Lau K.M. và Yang S. (1997), "Climatology and interannual variability of the southeast asian summer monsoon", *Adv Atmos Sci*, 14(2), 141-162
46. Li, C. and Yanai, M. (1996), "The onset and interannual variability of the Asian summer monsoon in relation to Land-Sea thermal contrast", *Journal of Climate*, 9(2), 358-375.
[https://doi.org/10.1175/1520-0442\(1996\)009<0358:TOAIVO>2.0.CO;2](https://doi.org/10.1175/1520-0442(1996)009<0358:TOAIVO>2.0.CO;2).
47. Li, J. and Zhang, L. (2009), "Wind onset and withdrawal of Asian summer monsoon and their simulated performance in AMIP models", *Climate Dynamics*, 32(7-8), 935-968.
48. Liebmann, B. and Marengo, J.A. (2001), "Interannual variability of the rainy season and rainfall in the Brazilian Amazon Basin", *Journal of Climate*, 14(22), 4308-4318.
49. Liebmann, B., Camargo, S.J., Seth, A., Marengo, J.A., Carvalho, L.M. V., Allured, D., Fu, R. and Vera, C.S. (2007), "Onset and end of the rainy season in South America in observations and the ECHAM 4.5 atmospheric general circulation model", *Journal of Climate*, 20(10), 2037-2050.
50. Mao, J., Chan, J.C.L. and Wu, G. (2004), "Relationship between the onset of the South China Sea summer monsoon and structure of the Asian subtropical anticyclone", *Journal of the Meteorological Society of Japan*, 82(3), 845-859,
<https://doi.org/10.2151/jmsj.2004.845>.
51. Moron, V., Robertson, A.W. and Boer, R. (2009), "Spatial coherence and seasonal predictability of monsoon onset over Indonesia", *Journal of Climate*, 22(3), 840-850, <https://doi.org/10.1175/2008JCLI2435.1>.
52. Moron, V., Robertson, A.W. and Qian, J.-H. (2010), "Local versus regional-scale characteristics of monsoon onset and post-onset rainfall over Indonesia", *Climate Dynamics*, 34(2-3), 281-299.

53. Ngo-Duc, T. (2014), "Climate change in the coastal regions of Vietnam", *Coastal Disasters and Climate Change in Vietnam - Engineering and Planning Perspectives*, N. D. Thao, H. Takagi, and M. Esteban, Eds., Elsevier, Butterworth-Heinemann, ISBN: 978-0-12-800007-6, 175-198, doi:10.1016/B978-0-12-800007-6.00008-3.
54. Ngo-Thanh, H., Ngo-Duc, T., Nguyen-Hong, H., Baker, P. and Phan-Van, T. (2018), "A distinction between summer rainy season and summer monsoon season over the central highlands of Vietnam", *Theoretical and Applied Climatology*, 132(3-4), 1237-1246, <https://doi.org/10.1007/s00704-017-2178-6>.
55. Nguyen-Le, D., Matsumoto, J. and Ngo-Duc, T. (2014), "Climatological onset date of summer monsoon in Vietnam", *International Journal of Climatology*, 34(11), 3237-3250, <https://doi.org/10.1002/joc.3908>.
56. Nguyen-Le, D., Matsumoto, J. and Ngo-Duc, T. (2015), "Onset of the rainy seasons in the eastern Indochina peninsula", *Journal of Climate*, 28(14), 5645-5666, <https://doi.org/10.1175/JCLI-D-14-00373.1>.
57. Nguyen, D.-Q., Renwick, J. and McGregor, J. (2014a), "Variations of surface temperature and rainfall in Vietnam from 1971 to 2010", *International Journal of Climatology*, 34(1), 249-264, <https://doi.org/10.1002/joc.3684>.
58. Nguyen, D.-Q., Renwick, J. and McGregor, J. (2014b), "Variations of monsoon rainfall: a simple unified index", *Geophysical Research Letters*, 41(2), 575-581, <https://doi.org/10.1002/2013GL058155>.
59. Nguyen, T. D., Uvo, C. and Rosbjerg, D. (2007), "Relationship between the tropical Pacific and Indian Ocean sea-surface temperature and monthly precipitation over the central highlands, Vietnam", *Int. J. Climatol.*, 27: 1439-1454. doi:10.1002/joc.1486

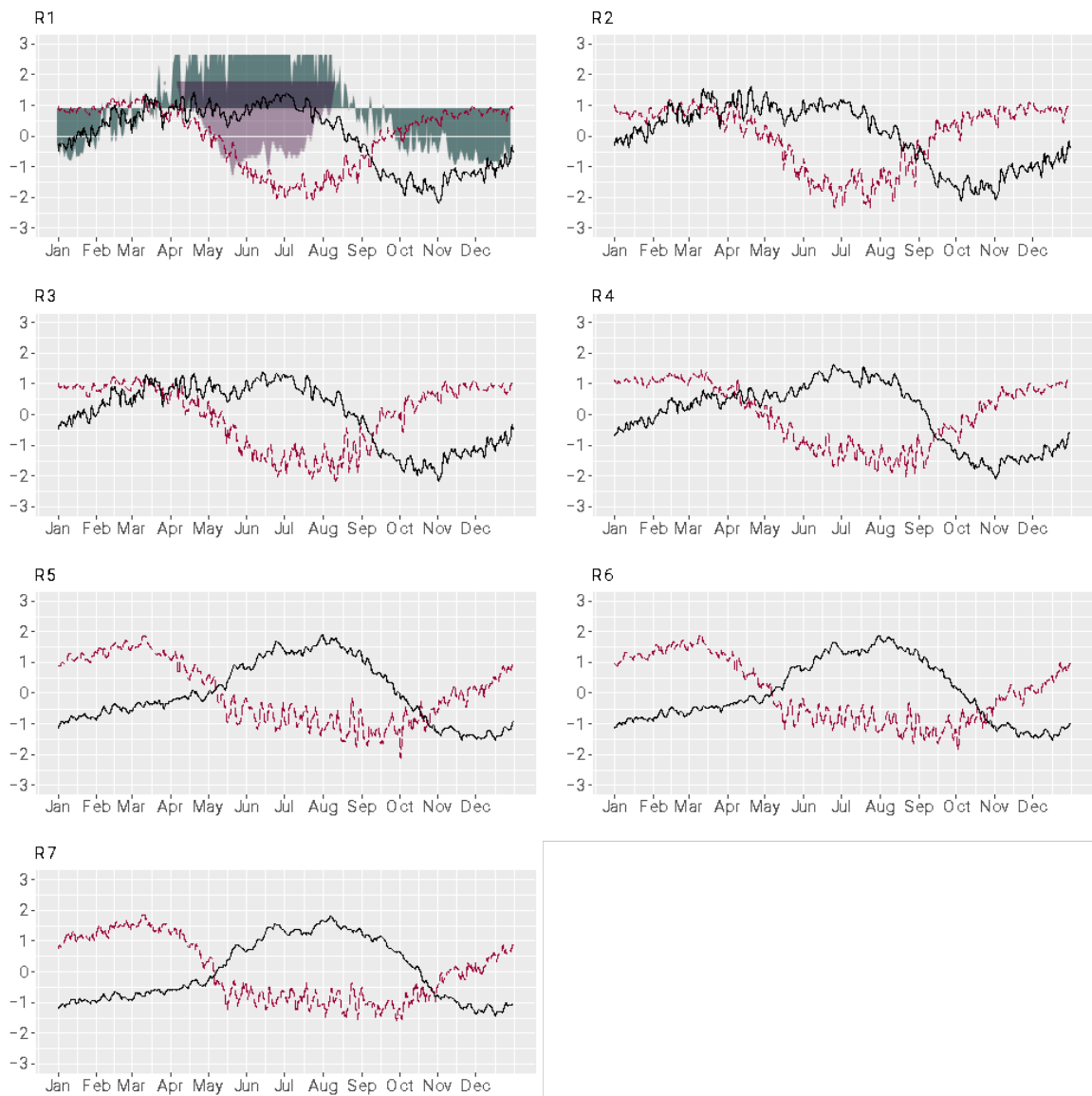
60. North, G.R, T.L. Bell, R.F. Cahalan, and F.J. Moeng. (1982), "Sampling errors in the estimation of empirical orthogonal functions", *Mon. Wea. Rev.*, 110, 699-706.
61. Odekunle, T.O., Balogun, E.E. and Ogunkoya, O.O. (2005), "On the prediction of rainfall onset and retreat dates in Nigeria", *Theoretical and Applied Climatology*, 81(1-2), 101-112. <https://doi.org/10.1007/s00704-004-0108-x>.
62. Omotosho, J.B. (1992), "Long-range prediction of the onset and end of the rainy season in the west African Sahel", *International Journal of Climatology*, 12(4), 369-382, <https://doi.org/10.1002/joc.3370120405>.
63. Omotosho, J.B., Balogun, A.A. and Ogunjobi, K. (2000), "Predicting monthly and seasonal rainfall, onset and cessation of the rainy season in West Africa using only surface data", *International Journal of Climatology*, 20(8), 865-880, [https://doi.org/10.1002/1097-0088\(20000630\)20:8<865::AID-JOC505>3.0.CO;2-R](https://doi.org/10.1002/1097-0088(20000630)20:8<865::AID-JOC505>3.0.CO;2-R).
64. Pham Xuan Thanh, Bernard Fontaine, Nathalie Philippon (2010), "Onset of the summer monsoon over the southern Vietnam and its predictability", *Theor Appl Climatol* (2010) 99:105-113 doi 10.1007/s00704-009-0115-z
65. Pham-Thanh, H., T. Ngo-Duc, J. Matsumoto, T. Phan-Van, and H Vo-Van (2020), "Rainfall Trends in Vietnam and their Associations with Tropical Cyclones during 1979-2019", *SOLA*, doi:10.2151/sola.2020-029.
66. Phan-Van, T., Nguyen-Xuan, T., Nguyen-Van, H., Laux, P., Pham- Thanh, H. and Ngo-Duc, T. (2018), "Evaluation of the NCEP climate forecast system and its downscaling for seasonal rainfall prediction over Vietnam", *Weather and Forecasting*, 33(3), 615-640. <https://doi.org/10.1175/WAF-D-17-0098.1>.
67. Pradhan, M., Rao, A.S., Srivastava, A., Dakate, A., Salunke, K. and Shameera, K.S. (2017), "Prediction of Indian summer-monsoon onset variability: a season in advance", *Scientific Reports*, 7(1), 14229.
68. Rajendran, K., Kitoh, A., & Yukimoto, S. (2004), "South and East Asian Summer Monsoon Climate and Variation in the MRI Coupled Model (MRI-CGCM2)", *Journal of Climate*, 17(4), 763-782. Retrieved Sep 4, 2021, from

https://journals.ametsoc.org/view/journals/clim/17/4/15200442_2004_017_0763_saeasm_2.0.co_2.xml

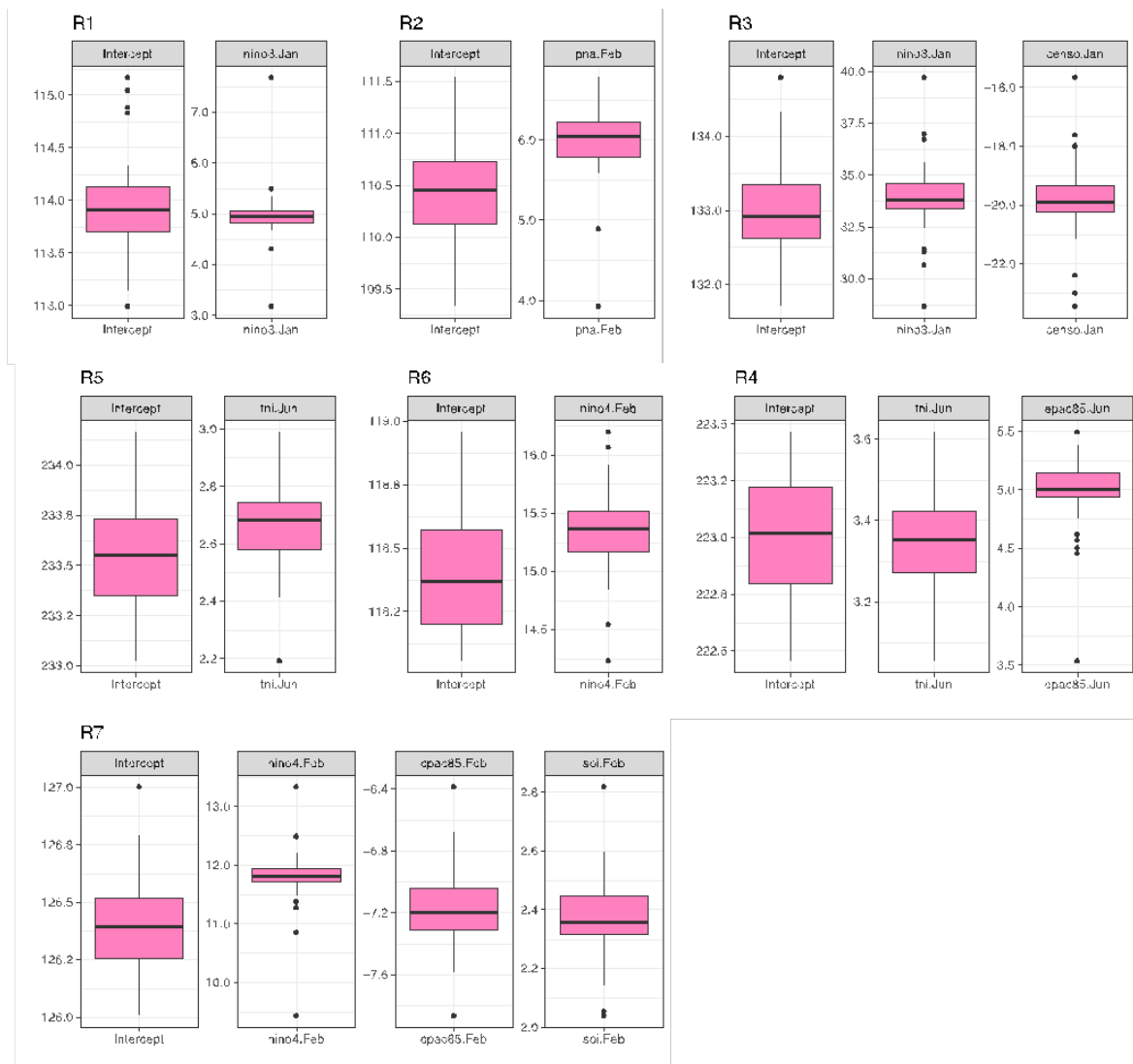
69. Ren, Hong-Li, and Pengfei Ren (2017), "Impact of Madden-Julian Oscillation upon Winter Extreme Rainfall in Southern China: Observations and Predictability in CFSv2", *Atmosphere* 8, no. 10: 192.
<https://doi.org/10.3390/atmos8100192>
70. Ren, P., Ren, H.-L., Fu, J.-X., Wu, J., & Du, L. (2018), "Impact of boreal summer intraseasonal oscillation on rainfall extremes in southeastern China and its predictability in CFSv2", *Journal of Geophysical Research: Atmospheres*, 123, 4423-4442. <https://doi.org/10.1029/2017JD028043>
71. Saha S., et al. (2014), "The NCEP Climate Forecast System Version 2", *Journal of climate* 27(6), pp. 2185-2208.
72. Sen, P.K. (1968), "Estimates of the Regression Coefficient Based on Kendall's Tau", *Journal of the American Statistical Association*, 63(324) (1968) 1379-1389.
73. Seregina, LS, Fink, AH, van der Linden, R, Elagib, NA, Pinto, JG. (2019), "A new and flexible rainy season definition: Validation for the Greater Horn of Africa and application to rainfall trends", *Int J Climatol.*, 39, 989- 1012, <https://doi.org/10.1002/joc.5856>
74. Shao, X., Kang, J. and Cai, W. (2010), "Quantitative determination by temperature dependent near-infrared spectra", *Talanta*, 82, 1017-1021. <https://doi.org/10.1016/j.talanta.2010.06.009>.
75. Silva, A.E. and Carvalho, L.M.V. (2007), "Large-scale index for South America monsoon (LISAM)", *Atmospheric Science Letters*, 8(2), 51-57.
76. Sivakumar, M.V.K. (1988), "Predicting rainy season potential from the onset of rains in southern Sahelian and Sudanian climatic zones of West Africa", *Agricultural and Forest Meteorology*, 42(4), 295-305.

77. Stern, R.D., Dennett, M.D. and Garbutt, D.J. (1981), “The start of the rains in West Africa”, *International Journal of Climatology*, 1(1), 59-68, <https://doi.org/10.1002/joc.3370010107>.
78. Stolbova, V., Surovyatkina, E., Bookhagen, B. and Kurths, J. (2016), “Tipping elements of the Indian monsoon: prediction of onset and withdrawal”, *Geophysical Research Letters*, 43(8), 3982-3990.
79. Tong, H.W., Chan, J.C.L. and Zhou, W. (2009), “The role of MJO and mid-latitude fronts in the South China Sea summer monsoon onset”, *Climate Dynamics*, 33(6), 827-841.
80. Trinh-Tuan, L., R. T. Konduru, T. Inoue, T. Ngo-Duc, and J. Matsumoto (2019), “Autumn rainfall increasing trend in South Central Vietnam and its association with changes in Vietnam’s East Sea surface temperature”, *Geographical reports of Tokyo Metropolitan University*, **54**, 11–22.
81. Vellinga, M., Arribas, A., & Graham, R. (2013), “Seasonal forecasts for regional onset of the West African monsoon”, *Climate Dynamics*, 40(11-12), 3047- 3070. <https://doi.org/10.1007/s00382-012-1520-z>
82. Wang, B. and LinHo. (2002), “Rainy season of the Asian-Pacific summer monsoon”, *Journal of Climate*, 15(4), 386-398.
83. Webster, P.J. and Yang, S. (1992), “Monsoon and Enso: selectively interactive systems”, *Quarterly Journal of the Royal Meteorological Society*, 118(507), 877-926.
84. Wilks, D. (2006), *Statistical Methods in the Atmospheric Sciences*.
85. Zhang, Y., Li, T., Wang, B. and Wu, G. (2002), “Onset of the summer monsoon over the Indochina peninsula: climatology and interannual variations”, *Journal of Climate*, 15(22), 3206-3221, [https://doi.org/10.1175/1520-0442\(2002\)015<3206:](https://doi.org/10.1175/1520-0442(2002)015<3206:)
86. Zhou Wen and Johnny C. L. Chan (2007), “ENSO and the South China Sea summer monsoon onset”, *Int. J. Climatol.* 27: 157-167.

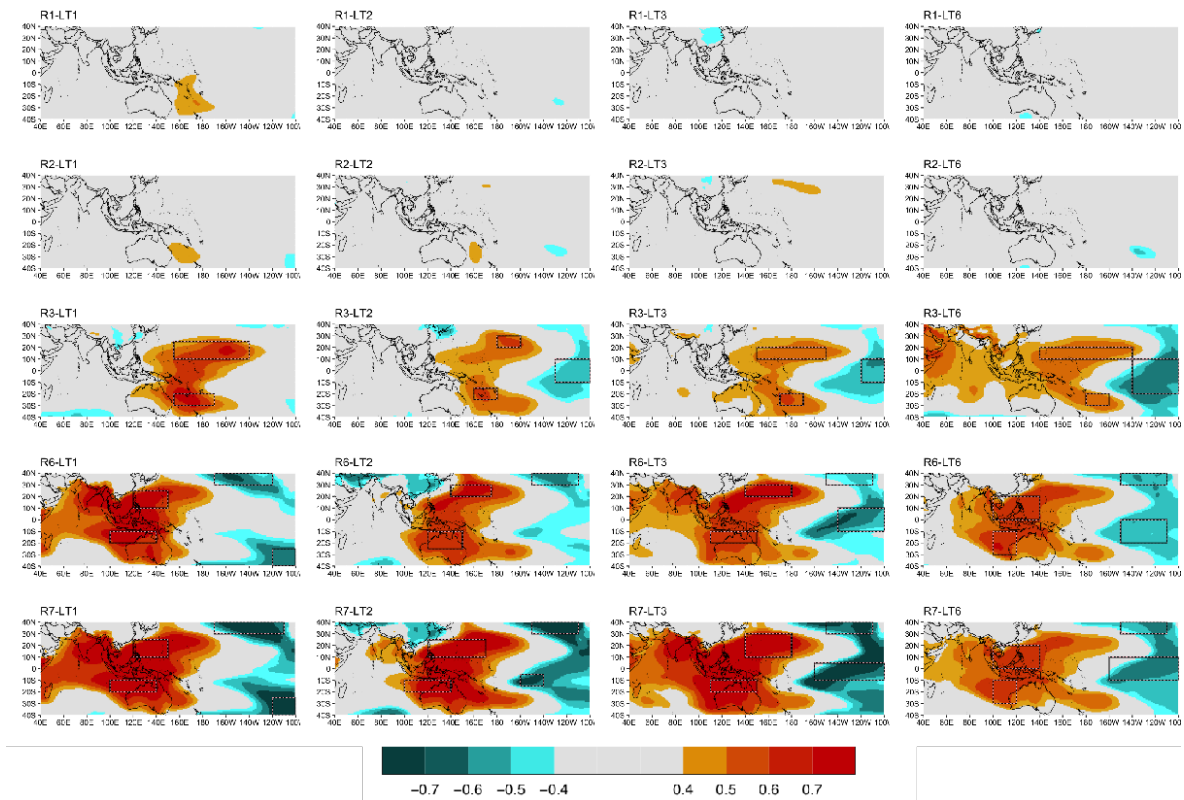
PHỤ LỤC



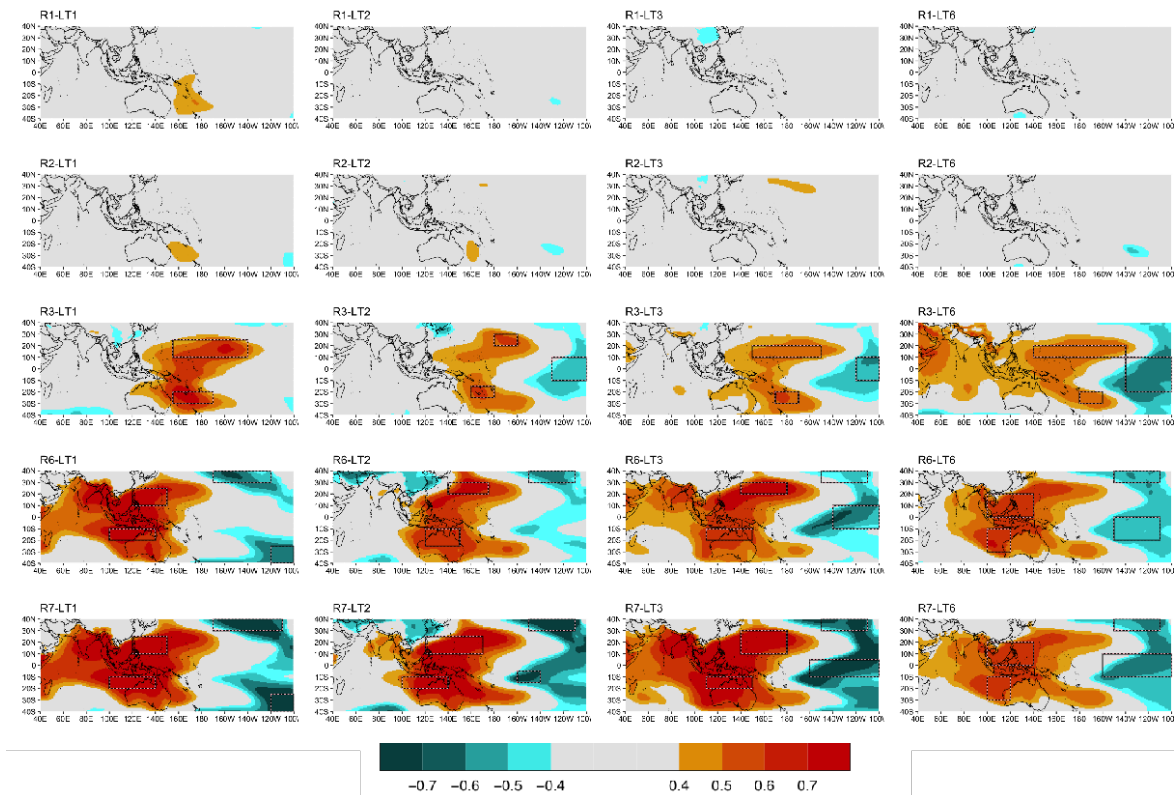
Hình 1: Trung bình nhiều năm giá trị chuẩn hoá trường mưa (màu lá), trường khí áp mực biển (màu tím), trường gió vĩ hướng mực 850mb (màu đen), trường OLR (màu đỏ) tại 7 vùng khí hậu



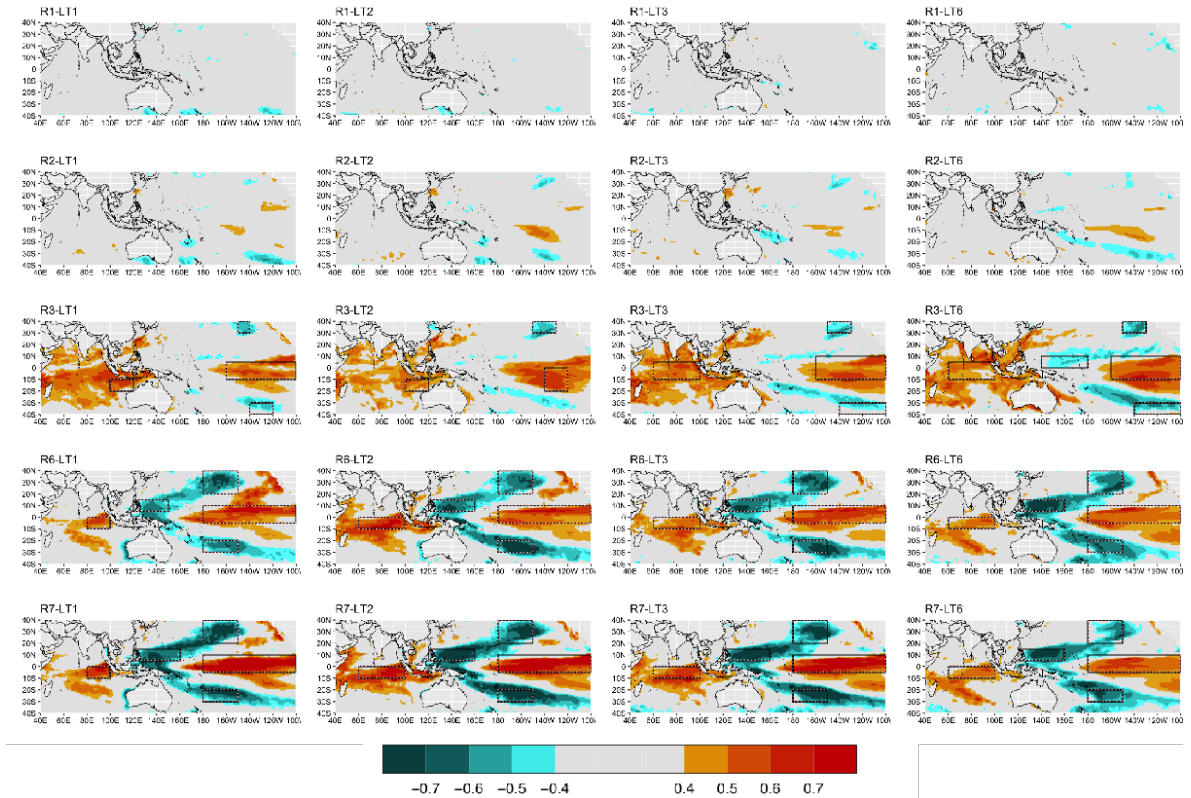
Hình 2: Khảo sát sự thay đổi của các hệ số hồi quy trong các phương trình dự báo NBĐMM dựa trên các chỉ số khí hậu tại 7 vùng khí hậu bằng phương pháp kiểm chứng chéo -OLOO



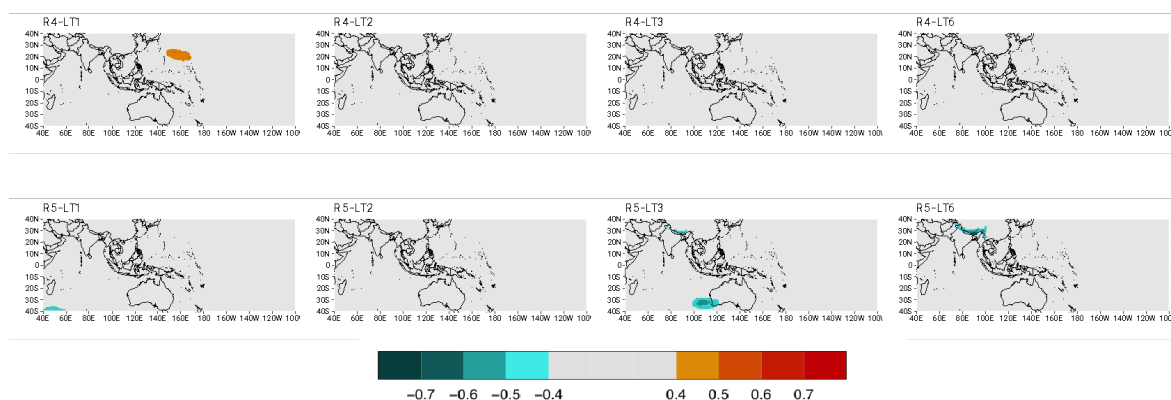
Hình 3: Hệ số tương quan giữa NBĐMM tại vùng R1 (hàng 1), R2 (hàng 2), R3 (hàng 3), R6 (hàng 4) và R7 (hàng 5) với trường dự báo PRMSL trung bình tháng 3 của CFSv2 với hạn 1 tháng (cột 1), hạn 2 tháng (cột 2), hạn 3 tháng (cột 3) và hạn 6 tháng (cột 4) trong giai đoạn 1983-2011. Chú ý: màu xám thể hiện khu vực có giá trị hệ số tương quan không thỏa mãn mức độ tin cậy 95% của kiểm nghiệm t. Những vùng đóng khung thể hiện khu vực có giá trị hệ số tương quan cao, đồng thời cũng là khu vực được lựa chọn làm nhân tố dự báo.



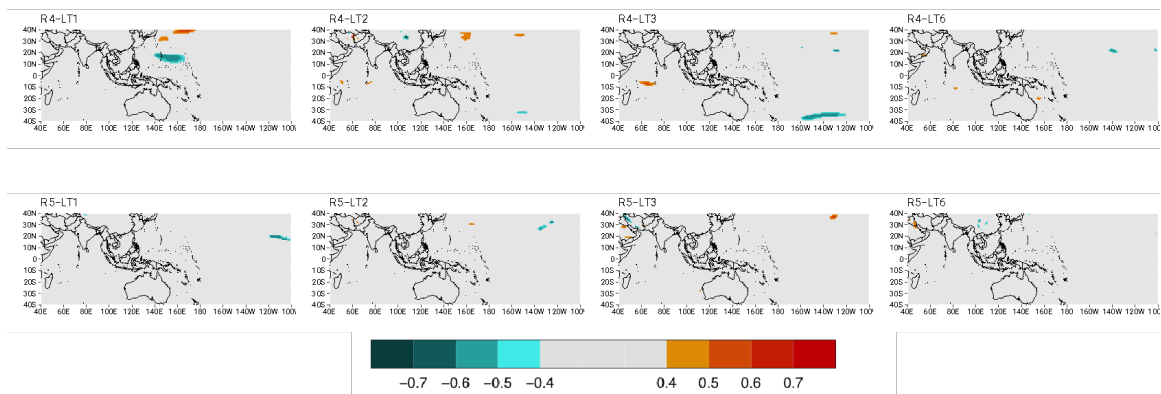
Hình 4: Hệ số tương quan giữa NBDMM tại vùng R1 (hàng 1), R2 (hàng 2), R3 (hàng 3), R6 (hàng 4) và R7 (hàng 5) với trường dự báo U850 trung bình tháng 3 của CFSv2 với hạn 1 tháng (cột 1), hạn 2 tháng (cột 2), hạn 3 tháng (cột 3) và hạn 6 tháng (cột 4) trong giai đoạn 1983-2011. Chú ý: màu xám thể hiện khu vực có giá trị hệ số tương quan không thỏa mãn mức độ tin cậy 95% của kiểm nghiệm t. Những vùng đóng khung thể hiện khu vực có giá trị hệ số tương quan cao, đồng thời cũng là khu vực được lựa chọn làm nhân tố dự báo.



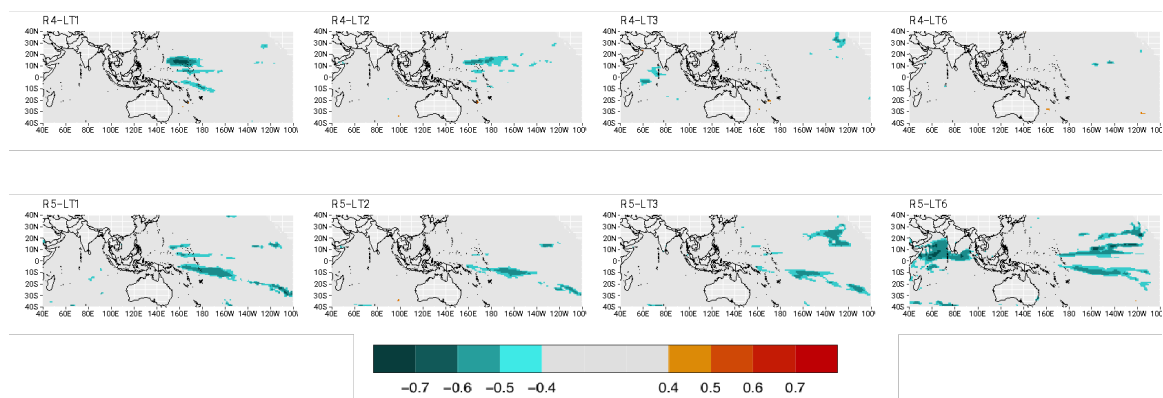
Hình 5: Hệ số tương quan giữa NBĐMM tại vùng R1 (hàng 1), R2 (hàng 2), R3 (hàng 3), R6 (hàng 4) và R7 (hàng 5) với trường dự báo SST trung bình tháng 3 của CFSv2 với hạn 1 tháng (cột 1), hạn 2 tháng (cột 2), hạn 3 tháng (cột 3) và hạn 6 tháng (cột 4) trong giai đoạn 1983-2011. Chú ý: màu xám thể hiện khu vực có giá trị hệ số tương quan không thỏa mãn mức độ tin cậy 95% của kiểm nghiệm t. Những vùng đóng khung thể hiện khu vực có giá trị hệ số tương quan cao, đồng thời cũng là khu vực được lựa chọn làm nhân tố dự báo.



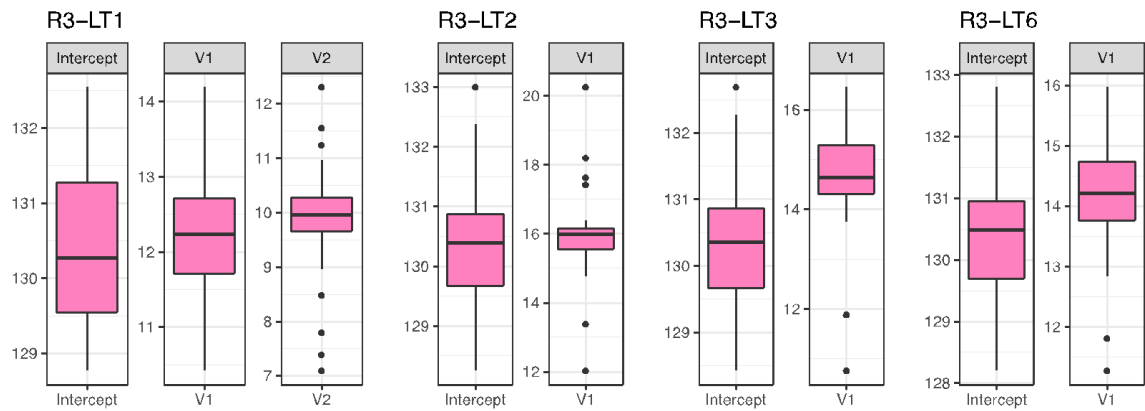
Hình 6: Hệ số tương quan giữa NBDMM tại vùng R4 (hàng 1), R5 (hàng 2) với trường dự báo PRMSL trung bình tháng 7 của CFSv2 với hạn 1 tháng (cột 1), hạn 2 tháng (cột 2), hạn 3 tháng (cột 3) và hạn 6 tháng (cột 4) trong giai đoạn 1983-2011. Chú ý: màu xám thể hiện khu vực có giá trị hệ số tương quan không thỏa mãn mức độ tin cậy 95% của kiểm nghiệm t. Những vùng đóng khung thể hiện khu vực có giá trị hệ số tương quan cao, đồng thời cũng là khu vực được lựa chọn làm nhân tố dự báo.



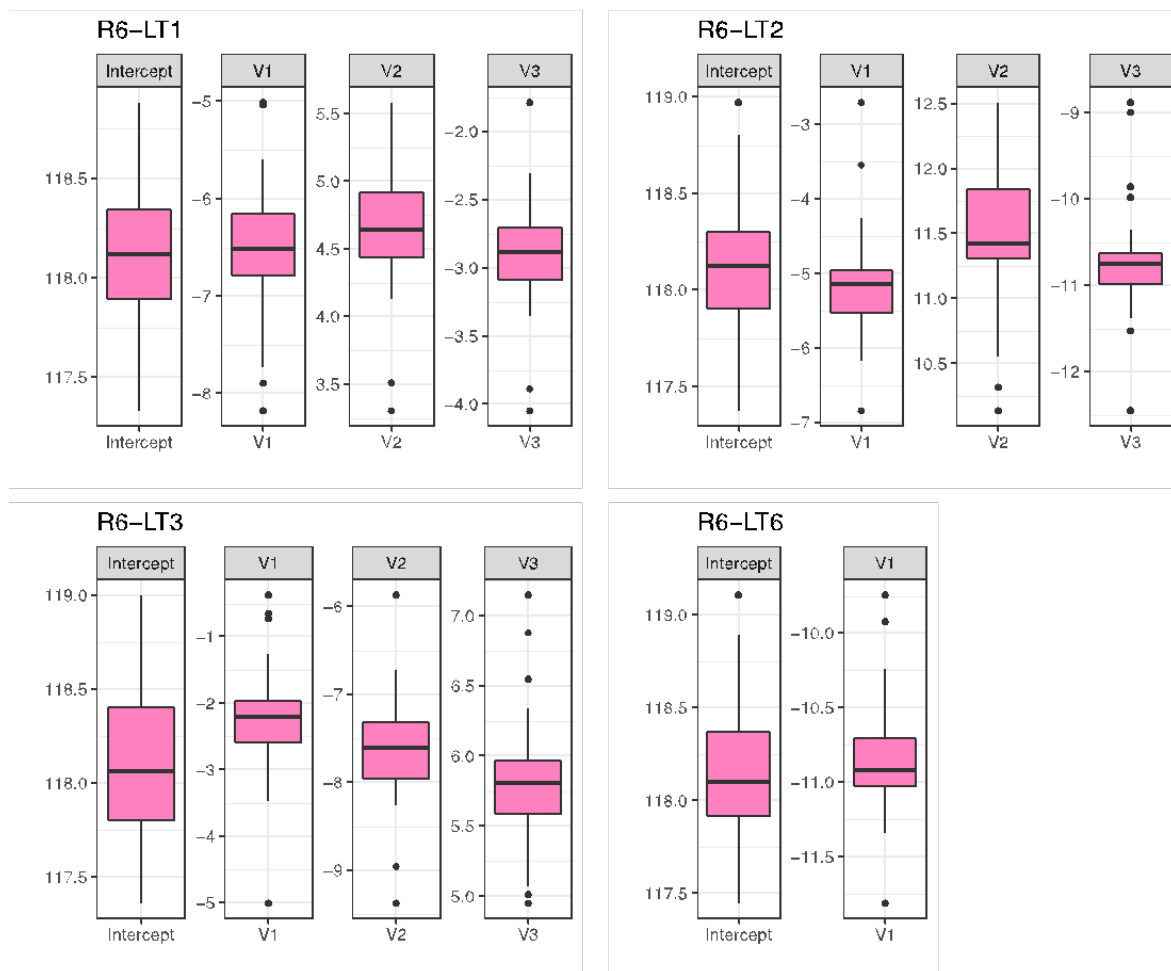
Hình 7: Hệ số tương quan giữa NBĐMM tại vùng R4 (hàng 1), R5 (hàng 2) với trường dự báo U850 trung bình tháng 7 của CFSv2 với hạn 1 tháng (cột 1), hạn 2 tháng (cột 2), hạn 3 tháng (cột 3) và hạn 6 tháng (cột 4) trong giai đoạn 1983-2011. Chú ý: màu xám thể hiện khu vực có giá trị hệ số tương quan không thỏa mãn mức độ tin cậy 95% của kiểm nghiệm t. Những vùng đóng khung thể hiện khu vực có giá trị hệ số tương quan cao, đồng thời cũng là khu vực được lựa chọn làm nhân tố dự báo.



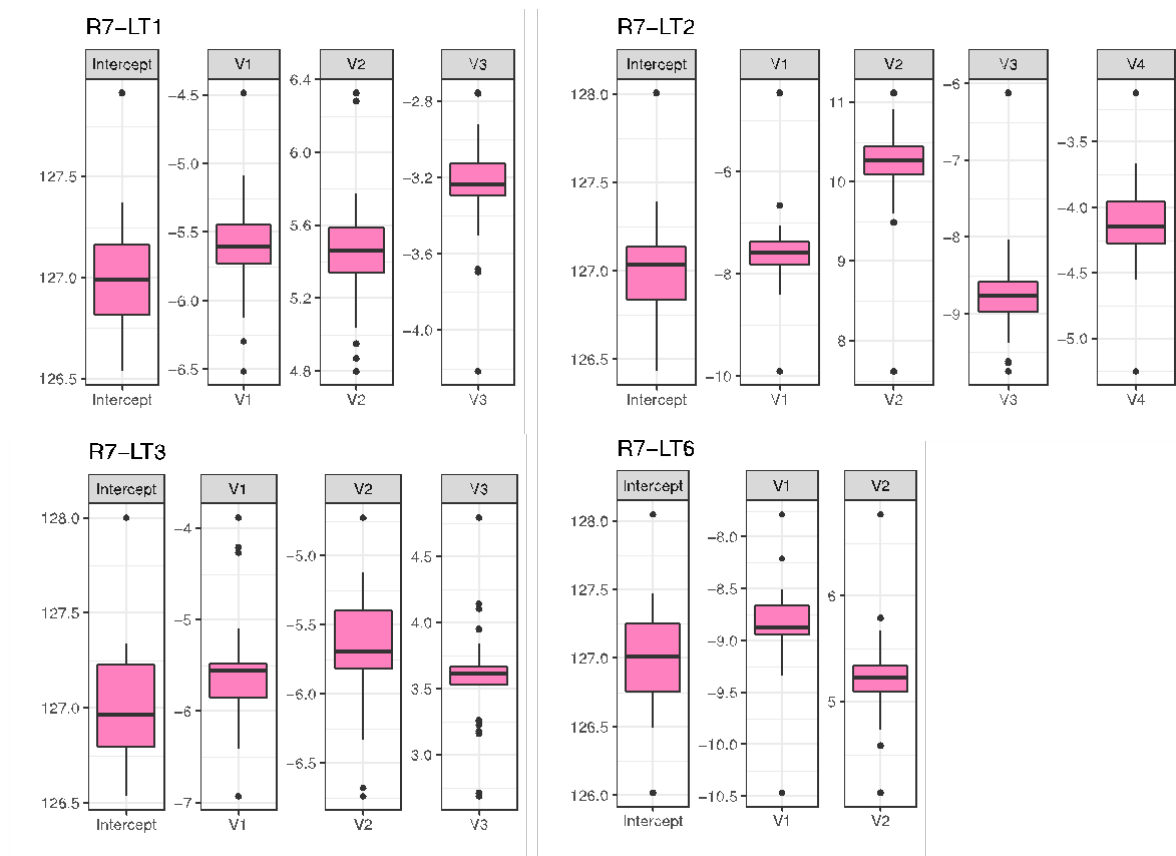
Hình 8: Hệ số tương quan giữa NBDMM tại vùng R4 (hàng 1), R5 (hàng 2) với trường dự báo SST trung bình tháng 7 của CFSv2 với hạn 1 tháng (cột 1), hạn 2 tháng (cột 2), hạn 3 tháng (cột 3) và hạn 6 tháng (cột 4) trong giai đoạn 1983-2011. Chú ý: màu xám thể hiện khu vực có giá trị hệ số tương quan không thỏa mãn mức độ tin cậy 95% của kiểm nghiệm t. Những vùng đóng khung thể hiện khu vực có giá trị hệ số tương quan cao, đồng thời cũng là khu vực được lựa chọn làm nhân tố dự báo.



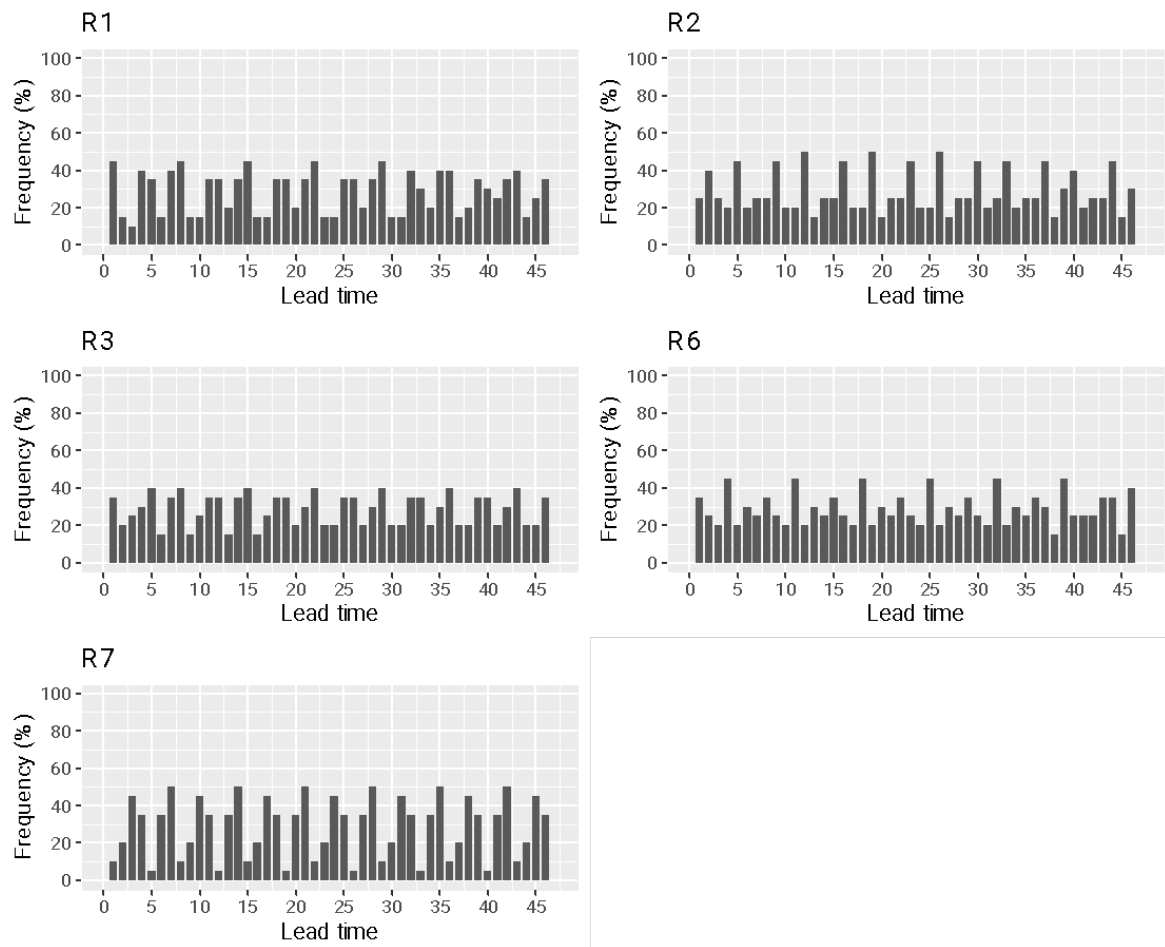
Hình 9: Khảo sát sự thay đổi của các hệ số hồi quy trong các phương trình dự báo NĐMM dựa trên sản phẩm CFSv2 cho các hạn dự báo khác nhau (1-2-3-6 tháng) tại vùng R3 bằng phương pháp kiểm chứng chéo -OLOO



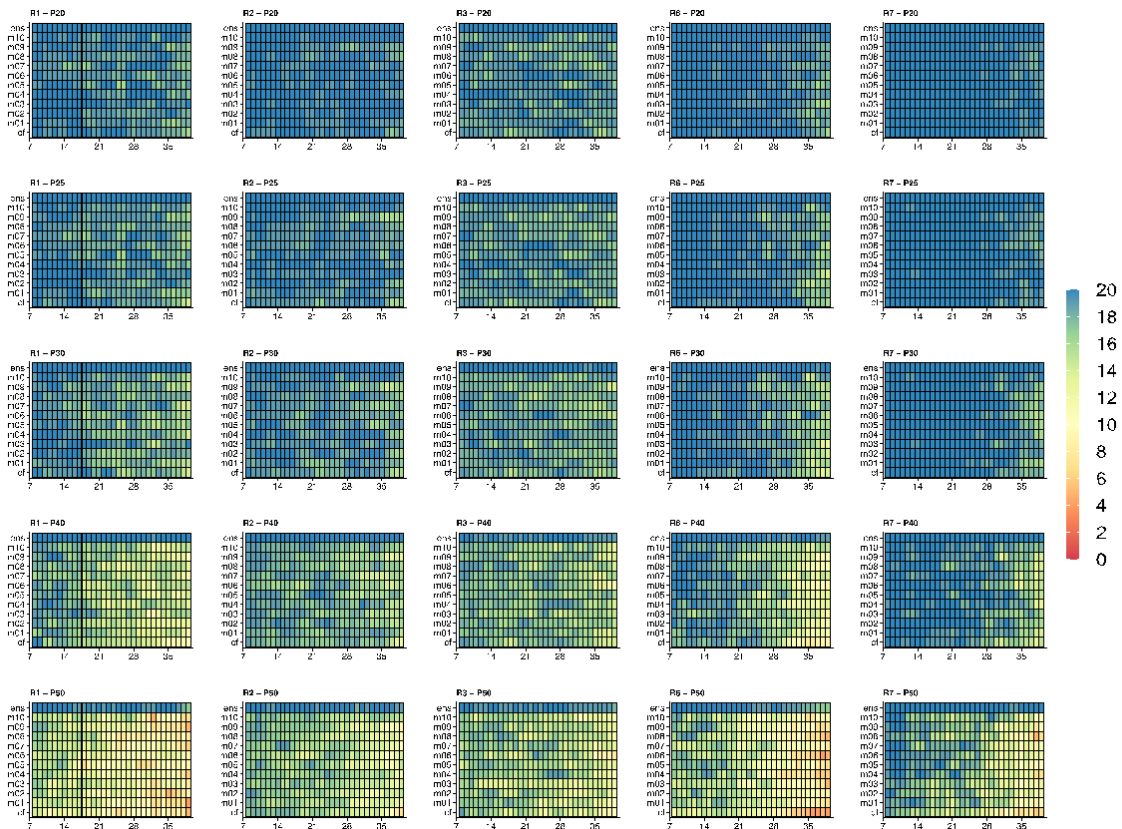
Hình 10: Khảo sát sự thay đổi của các hệ số hồi quy trong các phương trình dự báo NBĐMM dựa trên sản phẩm CFSv2 cho các hạn dự báo khác nhau (1-2-3-6 tháng) tại vùng R6 bằng phương pháp kiểm chứng chéo -OLOO



Hình 11: Khảo sát sự thay đổi của các hệ số hồi quy trong các phương trình dự báo NBĐMM dựa trên sản phẩm CFSv2 cho các hạn dự báo khác nhau (1-2-3-6 tháng) tại vùng R7 bằng phương pháp kiểm chứng chéo -OLOO



Hình 12: Tần suất của các dự báo NBĐMM mà không cần phải nội suy giá trị tại các hạn dự báo khác nhau (0-46 ngày) trong giai đoạn 2000-2019



Hình 13: Số năm mà NBDMM có thể được xác định từ thông tin dự báo của 11 thành phần và từ trung vị tổ hợp (ens) tại các hạn dự báo và vùng khí hậu khác nhau trong giai đoạn 2000-2019

Bảng 1: Giá trị NBĐMM giai đoạn 1979-2019 tại 7 vùng khí hậu
trên khu vực Việt Nam

Năm	R1	R2	R3	R4	R5	R6	R7
1979	1/05	16/04	13/04	31/08	9/09	27/04	23/04
1980	15/04	18/04	10/05	11/08	19/08	3/05	13/05
1981	21/04	13/04	10/05	17/08	27/08	12/05	6/05
1982	29/04	10/05	2/06	31/08	28/08	12/05	5/05
1983	3/06	11/05	27/07	27/08	12/09	19/05	22/05
1984	29/04	23/04	24/05	20/08	28/08	11/04	25/04
1985	12/04	29/03	6/03	13/08	26/08	27/04	21/04
1986	6/04	18/04	23/04	14/08	5/09	9/05	7/05
1987	12/05	25/05	14/06	14/08	17/08	21/05	25/05
1988	22/04	11/06	28/06	9/09	29/08	5/05	9/05
1989	30/04	31/03	23/04	14/08	17/08	12/04	4/05
1990	6/03	25/02	14/03	24/08	20/08	30/04	14/05
1991	19/04	29/04	23/05	16/08	29/08	16/05	20/05
1992	12/05	3/05	4/06	1/08	19/08	28/04	27/05
1993	11/04	27/03	28/04	17/08	5/09	4/05	21/05
1994	5/05	22/04	23/04	24/07	17/08	25/04	7/05
1995	6/05	14/05	9/06	14/08	20/08	13/05	20/05
1996	21/04	27/03	28/03	30/07	25/08	18/04	30/04
1997	15/03	13/03	23/03	3/08	19/08	8/05	22/05
1998	5/04	16/04	24/05	21/08	29/08	16/05	25/05
1999	28/04	1/05	10/05	27/08	15/09	22/03	26/03
2000	26/04	28/04	4/05	21/08	22/08	27/03	17/04
2001	19/03	9/03	10/03	19/08	6/08	30/04	4/05
2002	8/04	18/04	6/05	27/07	9/08	17/05	24/05

2003	29/04	3/05	15/05	8/08	19/08	16/05	7/05
2004	26/04	8/04	29/04	27/07	11/08	1/05	15/05
2005	9/05	3/05	13/06	30/07	8/08	13/05	14/05
2006	23/05	22/05	20/06	2/08	14/08	27/04	30/04
2007	17/04	1/05	16/05	9/08	31/08	16/04	28/04
2008	20/04	11/04	30/05	8/08	2/09	17/04	27/04
2009	20/04	11/04	10/05	13/08	24/08	3/04	17/04
2010	2/05	28/03	23/05	1/08	9/08	10/05	23/05
2011	23/04	25/04	1/06	28/07	13/08	26/04	2/05
2012	6/05	21/04	9/05	16/08	22/08	6/04	23/04
2013	17/04	22/04	1/05	8/08	10/08	17/04	6/05
2014	10/05	1/05	16/05	31/07	30/07	14/04	4/05
2015	4/06	4/06	28/06	22/08	8/09	22/05	29/05
2016	7/05	12/05	16/06	7/08	24/08	15/05	18/05
2017	17/04	18/04	26/05	14/08	2/09	12/04	23/04
2018	22/04	8/05	3/06	23/07	30/07	4/05	5/05
2019	21/05	7/04	16/05	17/08	30/08	7/05	7/05

