

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

**Đinh Bá Duy**

**ĐẶC ĐIỂM HOẠT ĐỘNG VÀ DỰ BÁO HẠN MÙA  
XOÁY THUẬN NHIỆT ĐỚI TRÊN BIỂN ĐÔNG**

**LUẬN ÁN TIẾN SĨ KHÍ TƯỢNG HỌC VÀ KHÍ HẬU HỌC**

**HÀ NỘI - 2021**

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

**Đinh Bá Duy**

**ĐẶC ĐIỂM HOẠT ĐỘNG VÀ DỰ BÁO HẠN MÙA  
XOÁY THUẬN NHIỆT ĐỚI TRÊN BIỂN ĐÔNG**

**Chuyên ngành: Khí tượng học**

**Mã số: 9440222.01**

**LUẬN ÁN TIẾN SĨ KHÍ TƯỢNG HỌC VÀ KHÍ HẬU HỌC**

**NGƯỜI HƯỚNG DẪN KHOA HỌC:**

- 1. PGS. TS. Ngô Đức Thành**
- 2. GS. TS. Phan Văn Tân**

**HÀ NỘI - 2021**

## **LỜI CAM ĐOAN**

Tôi xin cam đoan đây là công trình nghiên cứu của riêng tôi.

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**

**Đinh Bá Duy**

## LỜI CẢM ƠN

Trước tiên, tôi xin chân thành bày tỏ lòng cảm ơn sâu sắc nhất đến PGS. TS. Ngô Đức Thành và GS. TS. Phan Văn Tân, là những người thầy, người hướng dẫn đã tận tình chỉ bảo, định hướng khoa học và tạo mọi điều kiện tốt nhất cho tôi trong thời gian nghiên cứu và thực hiện luận án.

Tôi cũng xin gửi lời cảm ơn đến Khoa Khí tượng Thủy văn và Hải dương học, Phòng Sau Đại học, Trường Đại học Khoa học Tự nhiên đã tạo điều kiện giúp đỡ và tổ chức những hoạt động học tập và nghiên cứu một cách tận tình.

Trong quá trình tiến hành thực hiện một số thí nghiệm cho luận án, tôi đã nhận được sự trợ giúp nhiệt tình, hữu ích của PGS. TS. Trần Quang Đức cùng các thầy, cô tại Bộ môn Khí tượng, tôi xin cảm ơn những giúp đỡ nhiệt tình này.

Luận án này hoàn thành với sự hỗ trợ nguồn số liệu, kinh phí từ Đề tài KC.09.15/16-20, tôi xin trân trọng cảm ơn những hỗ trợ này.

Luận án này sẽ không thể thực hiện được nếu thiếu nguồn động viên, giúp đỡ to lớn từ phía gia đình tôi, tôi xin bày tỏ lòng biết ơn đến những người thân yêu trong gia đình đã có nhiều chia sẻ và giúp đỡ tôi cả về mặt tinh thần và vật chất.

Cuối cùng, đối với bạn bè, đồng nghiệp của tôi ở Viện Sinh thái Nhiệt đới, Trung tâm Nhiệt đới Việt Nga, tôi xin tỏ lòng biết ơn chân thành vì đã tạo điều kiện và có những góp ý hữu ích trong chuyên môn, chia sẻ trong công việc và cuộc sống.

# MỤC LỤC

Danh mục các ký hiệu và chữ viết tắt .....	4
Danh mục hình .....	6
Danh mục bảng .....	9
<b>MỞ ĐẦU</b> .....	11
Tính cấp thiết của đề tài .....	11
Mục tiêu của luận án .....	12
Đối tượng, phạm vi và nội dung nghiên cứu .....	12
Luận điểm bảo vệ của luận án.....	13
Những đóng góp mới của luận án .....	13
Ý nghĩa khoa học và thực tiễn .....	14
Tóm tắt cấu trúc luận án.....	14
<b>CHƯƠNG I. TỔNG QUAN VỀ ĐẶC ĐIỂM HOẠT ĐỘNG XTNĐ VÀ DỰ BẢO HẠN MÙA XTNĐ TRÊN BIỂN ĐÔNG</b> .....	17
1.1 Đặc điểm hoạt động của XTNĐ .....	19
1.2 Nghiên cứu dự báo hạn mùa XTNĐ .....	24
<b>CHƯƠNG II. SỐ LIỆU VÀ PHƯƠNG PHÁP NGHIÊN CỨU</b> .....	39
2.1 Số liệu và khu vực nghiên cứu .....	39
2.1.1 Số liệu XTNĐ .....	39
2.1.2 Số liệu về các chỉ số khí hậu .....	41
2.1.3 Số liệu dự báo lại của mô hình CFS .....	45
2.2 Phương pháp nghiên cứu.....	46
2.2.1 Xử lý số liệu ban đầu .....	46

2.2.2 Các mô hình thống kê .....	47
2.2.3 Tuyển chọn nhân tố dự báo .....	50
2.2.4 Vấn đề kiểm nghiệm phương trình dự báo .....	55
2.3 Đánh giá sai số dự báo của các mô hình .....	56
2.4. Thiết kế thí nghiệm dự báo hạn mùa số lượng XTNĐ .....	58
<b>CHƯƠNG III. MỘT SỐ ĐẶC ĐIỂM CỦA XTNĐ TRÊN BIỂN ĐÔNG VÀ</b>	
<b>MỐI QUAN HỆ VỚI CÁC QUÁ TRÌNH QUY MÔ LỚN .....</b>	<b>61</b>
3.1 Tần suất XTNĐ hoạt động .....	61
3.2 Quỹ đạo và vùng XTNĐ hoạt động .....	65
3.3 Quan hệ giữa số lượng XTNĐ trên Biển Đông và ENSO .....	70
3.4 Quan hệ giữa số lượng và số ngày hoạt động của XTNĐ trên Biển Đông và các chỉ số khí hậu .....	73
<b>CHƯƠNG IV. DỰ BÁO HẠN MÙA SỐ LƯỢNG XTNĐ TRÊN KHU</b>	
<b>VỰC BIỂN ĐÔNG .....</b>	<b>81</b>
4.1 Dự báo hạn mùa số lượng XTNĐ trên Biển Đông bằng phương pháp thống kê truyền thống .....	82
4.1.1 Tuyển chọn tập nhân tố dự báo .....	82
4.1.2 Kết quả dự báo số lượng XTNĐ Biển Đông theo TH1 .....	98
4.1.3 Kết quả dự báo số lượng XTNĐ Biển Đông theo TH2 .....	104
4.2 Dự báo hạn mùa số lượng XTNĐ Biển Đông bằng phương pháp thống kê động lực...109	
4.2.1 Quan hệ giữa các trường sản phẩm CFSv2 với số lượng XTNĐ Biển Đông... 109	
4.2.2 Lựa chọn, thiết lập bộ nhân tố dự tuyển và nhân tố dự báo .....	116
4.2.3 Kết quả dự báo số lượng XTNĐ hoạt động trên khu vực Biển Đông bằng phương pháp thống kê động lực theo TH3 .....	124

<b>KẾT LUẬN</b> .....	132
<b>KIẾN NGHỊ</b> .....	134
<b>DANH MỤC CÔNG TRÌNH KHOA HỌC CỦA TÁC GIẢ LIÊN QUAN ĐẾN LUẬN ÁN</b> .....	135
<b>TÀI LIỆU THAM KHẢO</b> .....	136
<b>PHỤ LỤC</b>	

## DANH MỤC CÁC KÝ HIỆU VÀ CHỮ VIẾT TẮT

ACE	Năng lượng xoáy bão tích lũy (Accumulated cyclone energy)
AGCM	Mô hình hoàn lưu chung khí quyển (Atmospheric general circulation models)
ANN	Mô hình mạng thần kinh nhân tạo (Artificial neural network)
APCC	Trung tâm Khí hậu châu Á – Thái Bình Dương (Asia-Pacific Climate Center)
BĐKH	Biến đổi khí hậu
CFS	Hệ thống dự báo khí hậu (Climate Forecast System)
CMA	Cơ quan khí tượng Trung Quốc (China Meteorological Administration)
CMB	Cuối mùa bão (những XTNĐ hoạt động trong các tháng 9, 10 và 11)
ĐMB	Đầu mùa bão (những XTNĐ hoạt động trong các tháng 6, 7 và 8)
ECMWF	Trung tâm dự báo thời tiết hạn vừa Châu Âu (European Centre for Medium-Range Weather Forecasts)
El Niño	Hiện tượng El Niño
GCACIC	Trung tâm Tác động Khí hậu Châu Á - Thái Bình Dương (Guy Carpenter Asia - Pacific Climate Impact Centre)
IBTrACS	Bộ số liệu tổng hợp về bão được NOAA, NCDC cung cấp (International Best Track Archive for Climate Stewardship)
IOD	Lưỡng cực Ấn Độ Dương (Indian Ocean Dipole)
JMA	Cơ quan khí tượng Nhật Bản (Japan Meteorological Agency)
JTWC	Trung tâm Cảnh báo bão Hải quân Hoa kỳ (Joint Typhoon Warning Center)
La Niña	Hiện tượng La Niña
LAD	Mô hình thống kê theo phương pháp độ lệch tuyệt đối nhỏ nhất (Least absolute deviations method)
LMV	Mô hình thống kê theo phương pháp minimax (Minimax method)

MAE	Sai số tuyệt đối trung bình (Mean Absolute Error)
ME	Sai số trung bình (Mean Error)
MEI	Chỉ số ENSO đa biến (Multivariate ENSO Index)
MLR	Mô hình thống kê theo phương pháp bình phương tối thiểu (Multiple linear regression)
MSSS	Chỉ số đánh giá kỹ năng dự báo (Mean Squared Skill Score)
NCAR	Trung tâm nghiên cứu khí quyển quốc gia của Hoa Kỳ (National Center for Atmospheric Research)
NCDC	Trung tâm Dữ liệu Khí hậu Quốc gia Hoa Kỳ (NOAA's National Climatic Data Center)
NCEP	Trung tâm dự báo môi trường quốc gia của Hoa Kỳ (The United States National Centers for Environmental Prediction)
NOAA	Trung tâm Quản lý Đại dương và Khí quyển Quốc gia (National Oceanic and Atmospheric Administration)
QBO	Dao động tựa hai năm (Quasi-biennial Oscillation)
RMSE	Sai số bình phương trung bình (Root mean square Error)
RSMC	Cơ quan nghiên cứu khí tượng chuyên vùng (Regional Specialized Meteorological Centre)
SOI	Chỉ số dao động nam (Southern Oscillation Index)
SSTA	Dị thường nhiệt độ mực biển (Sea Surface Temperature Anomaly)
TCWC	Trung tâm Cảnh báo Xoáy thuận nhiệt đới (Tropical Cyclone Warning Center)
TBTBD	Tây Bắc Thái Bình Dương
WMO	Tổ chức khí tượng thế giới (World Meteorological Organization)
XTNĐ	Xoáy thuận nhiệt đới

## DANH MỤC HÌNH

Hình 1.1 Các vùng hoạt động chính của XTNĐ trên quy mô toàn cầu .....	17
Hình 2.1 Khu vực nghiên cứu .....	40
Hình 2.2 Sơ đồ mạng ANN với các biến đầu vào, lớp ẩn và lớp đầu ra .....	49
Hình 2.3. Ba trường hợp nghiên cứu, khảo sát khả năng dự báo hạn mùa số lượng XTNĐ hoạt động trên khu vực Biển Đông .....	59
Hình 3.1 Biến động hàng năm của số lượng và số ngày hoạt động của XTNĐ trên Biển Đông giai đoạn 1971-2019 .....	61
Hình 3.2 Biến trình năm số lượng (Hình a), số ngày hoạt động (Hình b) và tần suất xuất hiện (Hình c) của XTNĐ trên Biển Đông giai đoạn 1971-2019 .....	62
Hình 3.3 Biến động hàng năm của các tháng ngoài mùa bão, các tháng trong mùa bão (Hình a), cả năm (Hình b) và biến trình năm (Hình c) số lượng XTNĐ hoạt động trên Biển Đông giai đoạn 1971-2019 .....	63
Hình 3.4 Biến động năm và xu thế biến đổi của tổng lượng XTNĐ trên Biển Đông trong cả mùa bão (MB: các tháng từ 6-11); đầu mùa bão (ĐMB: các tháng 6-7-8); cuối mùa bão (CMB: các tháng 9-10-11) và ngoài mùa bão (NMB: các tháng 1-5 và 12) .....	64
Hình 3.5 Quỹ đạo XTNĐ các tháng trong năm (trái sang phải, trên xuống dưới theo thứ tự từ tháng 1 đến tháng 12) khu vực Biển Đông giai đoạn 1971-2019 .....	65
Hình 3.6 Trung bình năm số lượng (trái) và số ngày hoạt động (phải) của XTNĐ qua từng thập kỷ trên khu vực Biển Đông giai đoạn 1971-2019 .....	67
Hình 3.7 Ba thành phần chính (PCs) và chuỗi thời gian (EOFs) đầu tiên của số lượng XTNĐ khu vực Biển Đông giai đoạn 1971-2019 .....	69
Hình 3.8 Số lượng XTNĐ hoạt động trên Biển Đông trong các năm ENSO phân theo cấp bão (TS/TY) và áp thấp nhiệt đới (TD) .....	71
Hình 3.9 Số lượng XTNĐ Biển Đông trong 3 tháng đầu mùa bão, 3 tháng cuối mùa bão theo các năm ENSO .....	72
Hình 3.10 Tương quan giữa 30 chỉ số khí hậu với số lượng XTNĐ 1, 3 và 6 tháng hoạt động trên Biển Đông giai đoạn 1979-2010 .....	76

Hình 3.11 Tương quan giữa 30 chỉ số khí hậu với số ngày XTNĐ hoạt động trong 1, 3 và 6 tháng hoạt động trên Biển Đông giai đoạn 1979-2010 .....	77
Hình 4.1 Mức độ (%) giải thích thông tin 20 thành phần chính (PC) đầu tiên của bộ dữ liệu nhân tố dự tuyển tháng 6 và tháng 9 .....	83
Hình 4.2 Mối quan hệ giữa 7 EOF đầu tiên của bộ dữ liệu nhân tố dự tuyển tháng 6 với XTNĐ Biển Đông hoạt động trong 3 tháng 6-7-8 .....	86
Hình 4.3 Mối quan hệ giữa 7 EOF đầu tiên của bộ dữ liệu nhân tố dự tuyển tháng 9 với XTNĐ Biển Đông hoạt động trong 3 tháng 9-10-11 .....	87
Hình 4.4 Số lượng các nhân tố dự báo được tuyển chọn bằng kỹ thuật hồi quy từng bước theo các ngưỡng Epsilon khác nhau.....	90
Hình 4.5 Sai số MAE trung bình theo các ngưỡng lọc khác nhau cho giai đoạn dự báo phụ thuộc và độc lập.....	95
Hình 4.6 Quan hệ giữa nhân tố dự báo và yếu tố dự báo tháng 6 ở TH2 .....	97
Hình 4.7 Quan hệ giữa nhân tố dự báo và yếu tố dự báo tháng 9 ở TH2 .....	98
Hình 4.8 Số lượng XTNĐ 3 tháng 6-7-8 và 9-10-11 trên Biển Đông giai đoạn 1980-2010 (hình trên) và giai đoạn độc lập 2011-2017 (hình dưới), số lượng quan trắc (OBS) và kết quả dự báo các mô hình MLR, LAD, LMV, ANN theo TH1 .....	100
Hình 4.9 Số lượng XTNĐ 3 tháng 6-7-8 và 9-10-11 trên Biển Đông giai đoạn 1980-2010 (hình trên) và giai đoạn độc lập 2011-2017 (hình dưới), số lượng quan trắc (OBS) và kết quả dự báo các mô hình MLR, LAD, LMV, ANN theo TH2 .....	106
Hình 4.10 Bản đồ tương quan giữa SST với các độ trễ 1, 2, 3, 6 tháng và số lượng XTNĐ Biển Đông ba tháng (vùng gạch chéo thể hiện giá trị tuyệt đối của hệ số tương quan $\geq 0,4$ và thỏa mãn độ tin cậy thống kê 95%) .....	110
Hình 4.11 Bản đồ tương quan giữa trường xoáy tương đối mực 850mb (VOR850) với các độ trễ 1, 2, 3, 6 tháng và số lượng XTNĐ Biển Đông ba tháng (vùng gạch chéo thể hiện giá trị tuyệt đối của hệ số tương quan $\geq 0,4$ và thỏa mãn độ tin cậy thống kê 95%) .....	113

Hình 4.12 Bản đồ tương quan giữa trường RH500 (hình trái), VWS (hình phải) tại các độ trễ 1, 2, 3, 6 tháng với số lượng XTNĐ Biển Đông ba tháng (vùng gạch chéo thể hiện giá trị tuyệt đối của hệ số tương quan $\geq 0,4$ và thỏa mãn độ tin cậy thống kê 95%).....	115
Hình 4.13 Bản đồ tương quan giữa trường PMSL (hình trái), H500 (hình phải) tại các độ trễ 1, 2, 3, 6 tháng với số lượng XTNĐ Biển Đông ba tháng (vùng gạch chéo thể hiện giá trị tuyệt đối của hệ số tương quan $\geq 0,4$ và thỏa mãn độ tin cậy thống kê 95%) .....	116
Hình 4.14 Các khu vực định hướng lựa chọn nhân tố dự tuyển từ sản phẩm CFSv2.....	117
Hình 4.15 Bản đồ tương quan giữa số lượng XTNĐ Biển Đông 3 tháng với trường SST, VOR850 từ sản phẩm của CFSv2 và lựa chọn nhân tố dự tuyển cho tháng 3 ...	118
Hình 4.16 Số lượng nhân tố dự báo được lựa chọn từ tập các nhân tố dự tuyển từ sản phẩm của mô hình CFSv2 .....	119
Hình 4.17 Sai số MAE trên bộ số liệu phụ thuộc (1983-2009) và độc lập (2012-2016) của TH3 theo các bộ nhân tố dự báo (ngưỡng lọc Epsilon) khác nhau .....	122
Hình 4.18 Mối quan hệ giữa các nhân tố dự báo với XTNĐ Biển Đông 3 tháng 6-7-8...123	
Hình 4.19 Mối quan hệ giữa nhân tố dự báo với XTNĐ Biển Đông 3 tháng 9-10-11 .....	123
Hình 4.20 Số lượng XTNĐ 3 tháng 6-7-8 (hình trên) và 9-10-11 (hình dưới) trên Biển Đông giai đoạn 1983-2009 và giai đoạn độc lập 2012-2016 , số lượng quan trắc (OBS) và kết quả dự báo các mô hình MLR, LAD, LMV, ANN theo TH3 .....	125
Hình 4.21 Tỷ lệ đúng của dự báo cho 3 tháng đầu mùa bão (tháng 6-7-8, hình trái) và 3 tháng cuối mùa bão (tháng 9-10-11, hình phải) theo 3 trường hợp (TH1, TH2, và TH3).....	129
Hình 4.22 Tỷ lệ đúng (tỷ số giữa số lần dự báo đúng và tổng số lần dự báo; %) trung bình ba trường hợp khảo sát của các mô hình MLR, LAD, LMV, ANN và tổ hợp E1234 .....	130

## DANH MỤC BẢNG

Bảng 1.1 Phạm vi không gian của các vùng, khu vực hoạt động chính của XTNĐ trên quy mô toàn cầu .....	18
Bảng 1.2 Một số tổ chức phát hành bản tin dự báo mùa XTNĐ trên thế giới .....	34
Bảng 2.1 Ba nguồn số liệu về XTNĐ .....	41
Bảng 2.2 Bộ chỉ số khí hậu sử dụng trong luận án .....	43
Bảng 3.1 Số lượng và số ngày XTNĐ Biển Đông hoạt động trong các pha ENSO .....	73
Bảng 4.1 Các thí nghiệm khảo sát sai số của việc lựa chọn số lượng thành phần chính làm nhân tố dự báo .....	84
Bảng 4.2 Sai số MAE của ba mô hình thống kê theo số lượng thành phần chính, dự báo cho thời điểm tháng 6 và tháng 9 .....	85
Bảng 4.3 Sai số MAE giai đoạn luyện của mô hình ANN theo các tham số tốc độ luyện và số lớp ẩn.....	89
Bảng 4.4 Sai số MAE giai đoạn dự báo kiểm nghiệm của mô hình ANN theo các tham số tốc độ luyện và số lớp ẩn .....	89
Bảng 4.5 Độ trễ của các nhân tố dự báo tháng 6 được tuyển chọn theo các ngưỡng lọc Epsilon khác nhau trong hồi quy từng bước .....	91
Bảng 4.6 Độ trễ của các nhân tố dự báo tháng 9 được tuyển chọn theo các ngưỡng lọc Epsilon khác nhau trong hồi quy từng bước .....	93
Bảng 4.7 Sai số MAE theo các ngưỡng lọc Epsilon dự báo trên bộ số liệu phụ thuộc (1980-2010).....	94
Bảng 4.8 Sai số MAE theo các ngưỡng lọc Epsilon dự báo trên bộ số liệu độc lập (2011-2017).....	94
Bảng 4.9 Hệ số hồi quy và giá trị Pval của các mô hình dự báo theo TH1 .....	99
Bảng 4.10 Sai số và chỉ số đánh giá kỹ năng dự báo tại thời điểm dự báo tháng 6 ở giai đoạn dự báo phụ thuộc (1980-2010) và độc lập (2011-2017) theo TH1 .....	102

Bảng 4.11 Sai số và chỉ số đánh giá kỹ năng dự báo tại thời điểm dự báo tháng 9 ở giai đoạn dự báo phụ thuộc (1980-2010) và độc lập (2011-2017) theo TH1 .....	103
Bảng 4.12 Hệ số hồi quy và nhân tố dự báo của các phương trình dự báo số lượng XTNĐ hoạt động trong 3 tháng 6-7-8 theo TH2 .....	104
Bảng 4.13 Hệ số hồi quy và nhân tố dự báo của các phương trình dự báo số lượng XTNĐ hoạt động trong 3 tháng 9-10-11 theo TH2 .....	105
Bảng 4.14 Sai số và chỉ số đánh giá kỹ năng dự báo tại thời điểm tháng 6 ở giai đoạn dự báo phụ thuộc (1980-2010) và độc lập (2011-2017) theo TH2 .....	107
Bảng 4.15 Sai số và chỉ số đánh giá kỹ năng dự báo tại thời điểm tháng 9 ở giai đoạn dự báo phụ thuộc (1980-2010) và độc lập (2011-2017) theo TH2 .....	108
Bảng 4.16 Các nhân tố dự tuyển được lựa chọn làm nhân tố dự báo cho các XTNĐ hoạt động trên khu vực Biển Đông các tháng 6-7-8 .....	120
Bảng 4.17 Các nhân tố dự tuyển được lựa chọn làm nhân tố dự báo cho các XTNĐ hoạt động trên khu vực Biển Đông các tháng 9-10-11 .....	120
Bảng 4.18 Hệ số hồi quy của các mô hình dự báo thống kê của TH3 cho tháng 6 ....	124
Bảng 4.19 Hệ số hồi quy của các mô hình dự báo thống kê của TH3 cho tháng 9 ....	124
Bảng 4.20 Sai số và chỉ số đánh giá kỹ năng dự báo tại thời điểm tháng 6 ở giai đoạn dự báo phụ thuộc (1983-2009) và độc lập (2012-2016) theo TH3 .....	126
Bảng 4.21 Sai số và chỉ số đánh giá kỹ năng dự báo tại thời điểm tháng 9 ở giai đoạn dự báo phụ thuộc (1983-2009) và độc lập (2012-2016) theo TH3 .....	127
Bảng 4.22 Kết quả đánh giá dự báo trên chuẩn/ dưới chuẩn so với trung bình khí hậu của 3 trường hợp.....	129

## MỞ ĐẦU

### Tính cấp thiết của đề tài

Đất nước hình chữ S với đặc điểm chiều dài bờ biển lên tới 3.260 km hơn gấp đôi lần chiều dọc và hơn gấp 5 lần chiều ngang lãnh thổ - Việt Nam được biết đến là quốc gia nằm trong trung tâm bão lớn của thế giới thuộc khu vực Tây Bắc Thái Bình Dương. Với đặc điểm này, hàng năm Việt Nam luôn tiềm ẩn nhiều nguy cơ chịu ảnh hưởng mạnh mẽ từ các thiên tai đến từ biển như nước biển dâng, sóng thần, xâm ngập mặn ... trong đó đặc biệt và thường xuyên nhất phải kể đến hiện tượng bão, áp thấp nhiệt đới (được gọi chung là xoáy thuận nhiệt đới – XTNĐ). Ảnh hưởng từ các XTNĐ với sức gió cực mạnh lên tới hàng trăm km/h và lượng mưa lớn hàng trăm mm đã phá hủy, làm hư hại tới nhiều công trình; gây ngập úng trên diện rộng; khiến giao thông và thông tin liên lạc bị chia cắt và ngoài những thiệt hại nặng nề về kinh tế, XTNĐ thậm chí còn gây nên những thiệt hại về người. Những năm gần đây, trong bối cảnh biến đổi khí hậu toàn cầu tình hình hoạt động của XTNĐ ảnh hưởng tới khu vực Đông Nam Á nói chung và nước ta nói riêng ở mức báo động cao, với diễn biến phức tạp khó dự đoán.

Tổng quan các nghiên cứu trong và ngoài nước cho thấy một số đặc điểm XTNĐ trên khu vực Biển Đông như số lượng, số ngày hoạt động, tần suất hoạt động, quỹ đạo và các dạng đường đi của XTNĐ, sự biến đổi thập kỷ và mối quan hệ giữa XTNĐ trên Biển Đông với một số quá trình khí quyển, đại dương quy mô lớn vẫn là những chủ đề còn nhiều tranh luận. Đồng thời các nghiên cứu dự báo XTNĐ cũng cho thấy việc xây dựng được mô hình thống kê dự báo hạn mùa số lượng XTNĐ hoạt động trên khu vực Biển Đông nói chung và Việt Nam nói riêng còn ít và vẫn là thách thức đối với ngành khí tượng.

Xuất phát từ thực tế trên, NCS đã lựa chọn đề tài luận án “Đặc điểm hoạt động và dự báo hạn mùa xoáy thuận nhiệt đới trên Biển Đông” với kỳ vọng các kết quả nghiên cứu sẽ góp phần nâng cao sự hiểu biết về đặc điểm hoạt động của XTNĐ trên Biển Đông cũng như cải thiện chất lượng dự báo hạn mùa số lượng XTNĐ trên

khu vực nghiên cứu. Qua đó nhằm đóng góp một phần nhỏ bé vào công tác phòng tránh, giảm thiểu thiệt hại do bão, áp thấp nhiệt đới đối với kinh tế xã hội và quốc phòng an ninh.

### **Mục tiêu của luận án**

Luận án nhằm đạt được các mục tiêu sau:

1) Làm rõ được một số đặc điểm hoạt động của XTNĐ trên khu vực Biển Đông và mối quan hệ với một số quá trình khí quyển, đại dương quy mô lớn trên cơ sở các bộ số liệu cập nhật mới nhất 1971-2019;

2) Xây dựng được các phương trình dự báo hạn mùa được số lượng XTNĐ hoạt động trên khu vực Biển Đông bằng các phương pháp thống kê trên cơ sở các nhân tố dự báo phù hợp được lựa chọn từ tập chỉ số khí hậu hoặc từ các trường quy mô lớn là sản phẩm của mô hình dự báo khí hậu (CFS).

### **Đối tượng, phạm vi và nội dung nghiên cứu**

**Đối tượng nghiên cứu:** Tổng lượng XTNĐ trên khu vực Biển Đông có cường độ từ áp thấp nhiệt đới trở lên (gió mạnh nhất từ cấp 6 trở lên, theo định nghĩa của Việt Nam) trong 3 tháng đầu mùa bão (các tháng 6, 7 và 8) và 3 tháng cuối mùa bão (các tháng 9, 10 và 11).

**Phạm vi nghiên cứu:** khu vực Biển Đông và Việt Nam được giới hạn bởi vùng không gian từ 100 đến 120° kinh Đông và 5 tới 23° vĩ Bắc theo Phụ lục 1, Quyết định số 03/2020/QĐ-TTg ngày 13/1/2020 của Thủ tướng Chính phủ, số liệu nghiên cứu được sử dụng trong khoảng thời gian từ 1971-2019.

**Hạn dự báo:** 3 tháng.

**Nội dung nghiên cứu:**

- 1) Lựa chọn bộ số liệu về XTNĐ cho khu vực Biển Đông;
- 2) Đặc điểm hoạt động của XTNĐ trên khu vực Biển Đông (tần suất về số lượng, số ngày hoạt động, quỹ đạo, vùng hoạt động và sự biến đổi thập kỷ của XTNĐ);

- 3) Mối quan hệ XTNĐ trên khu vực Biển Đông với ENSO;
- 4) Mối quan hệ XTNĐ trên khu vực Biển Đông với các chỉ số khí hậu;
- 5) Khả năng dự báo hạn mùa (3 tháng) số lượng XTNĐ hoạt động trên khu vực Biển Đông bằng các phương pháp thống kê với nhân tố dự báo được lựa chọn từ tập các chỉ số khí hậu;
- 6) Khả năng dự báo hạn mùa (3 tháng) số lượng XTNĐ hoạt động trên khu vực Biển Đông bằng các phương pháp thống kê với nhân tố dự báo được lựa chọn từ các trường quy mô lớn là sản phẩm của mô hình dự báo khí hậu CFS.

#### **Luận điểm bảo vệ của luận án**

1. Hoạt động của XTNĐ trên Biển Đông có thể có quan hệ với các quá trình khí quyển, đại dương quy mô lớn, được phản ánh qua các chỉ số khí hậu như nhóm chỉ số ENSO và các chỉ số hoàn lưu khí quyển, đại dương cũng như các trường dự báo từ mô hình khí hậu toàn cầu;
2. Việc dự báo hạn mùa (3 tháng) số lượng XTNĐ trên Biển Đông có thể thực hiện được bằng các phương pháp thống kê với nhân tố dự báo được lựa chọn từ tập các chỉ số khí hậu hoặc lựa chọn từ các trường khí quyển, đại dương sẵn có từ các mô hình dự báo hạn mùa toàn cầu.

#### **Những đóng góp mới của luận án**

1. Luận án đã đưa ra được các đặc điểm hoạt động của XTNĐ trên khu vực Biển Đông như: số lượng, số ngày hoạt động, tần suất hoạt động, các vùng hoạt động của XTNĐ, sự biến đổi thập kỷ của các đặc trưng đó và mối quan hệ giữa XTNĐ Biển Đông với các chỉ số khí hậu theo bộ số liệu RSMC (1971-2019);
2. Luận án đã lựa chọn được các bộ nhân tố dự báo thích hợp và xây dựng được các phương trình dự báo hạn mùa (3 tháng) số lượng XTNĐ hoạt động trên khu vực Biển Đông trong 3 tháng đầu mùa bão (tháng 6, 7, 8) và 3 tháng cuối mùa bão (tháng 9, 10, 11) bằng các phương pháp thống kê (MRL, LAD, LMV, ANN) và tổ hợp kết quả.

## **Ý nghĩa khoa học và thực tiễn**

### ***Ý nghĩa khoa học:***

1) Kết quả nhận được của luận án đã góp phần làm chính xác hóa một số đặc trưng thống kê của XTNĐ trên Biển Đông theo bộ số liệu RSMC được cập nhật giai đoạn 1971-2019;

2) Luận án đã chỉ ra được khả năng dự báo hạn mùa số lượng XTNĐ 3 tháng đầu mùa bão (tháng 6, 7, 8) và 3 tháng cuối mùa bão (tháng 9, 10, 11) bằng phương pháp thống kê với nhân tố dự báo được lựa chọn từ tập các chỉ số khí hậu hoặc các trường quy mô lớn là sản phẩm của mô hình dự báo khí hậu CFS.

### ***Ý nghĩa thực tiễn:***

1) Kết quả nhận được của luận án đã góp phần nâng cao sự hiểu biết về một số đặc trưng khí hậu, quy luật phân bố theo không gian, sự biến đổi theo thời gian cũng như mối quan hệ của XTNĐ trên Biển Đông với các quá trình khí quyển, đại dương quy mô lớn;

2) Luận án cũng cho thấy khả năng khai thác, sử dụng các chỉ số khí hậu và sản phẩm dự báo của mô hình CFS để tìm ra được bộ nhân tố dự báo phù hợp cho phương trình dự báo số lượng XTNĐ 3 tháng đầu mùa bão và 3 tháng cuối mùa bão có sai số thấp hơn so với độ lệch chuẩn khí hậu và kỹ năng dự báo tốt hơn so với dự báo trung bình khí hậu. Kết quả này có nhiều triển vọng để ứng dụng vào nghiệp vụ dự báo hạn mùa XTNĐ Biển Đông tại Việt Nam, góp phần cung cấp thông tin hữu ích cho công tác phòng tránh, giảm thiểu thiệt hại do bão, áp thấp nhiệt đới gây ra.

### **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 ĐẶC ĐIỂM HOẠT ĐỘNG XTNĐ VÀ DỰ BÁO HẠN MÙA XTNĐ TRÊN BIỂN ĐÔNG**

Chương này đề cập đến những kết quả nghiên cứu về đặc điểm hoạt động XTNĐ và bài toán dự báo hạn mùa XTNĐ trên thế giới và trong nước từ đó thấy

được những vấn đề chưa được giải quyết trong nghiên cứu về đặc điểm và khả năng dự báo XTNĐ trên khu vực Biển Đông mà đề tài sẽ góp phần giải quyết.

## **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. Chương 2 cũng trình bày về các phương pháp nghiên cứu, phương pháp đánh giá chất lượng dự báo và các thí nghiệm được xây dựng để dự báo hạn mùa XTNĐ trên khu vực Biển Đông.

## **CHƯƠNG 3. MỘT SỐ ĐẶC ĐIỂM CỦA XTNĐ TRÊN BIỂN ĐÔNG VÀ MỐI QUAN HỆ VỚI CÁC QUÁ TRÌNH QUY MÔ LỚN**

Chương này trình bày những kết quả và phân tích để thấy được đặc điểm hoạt động của XTNĐ trên khu vực Biển Đông, các đặc điểm này có thể kể đến như: số lượng, số ngày hoạt động; tần suất; sự biến đổi thập kỷ về số lượng và số ngày hoạt động và quỹ đạo, vùng hoạt động của các XTNĐ trên khu vực Biển Đông. Ngoài ra, để làm cơ sở cho việc lựa chọn các nhân tố dự tuyển, nhân tố dự báo cho việc xây dựng các phương trình dự báo thống kê tại Chương 4, tại đây luận án đã khảo sát mức độ tương quan và phân tích quan hệ giữa số lượng, số ngày của XTNĐ Biển Đông 1, 3 và 6 tháng với 30 chỉ số khí hậu có liên hệ với các nhóm ENSO và hoàn lưu khí quyển, đại dương.

## **CHƯƠNG 4. DỰ BÁO HẠN MÙA SỐ LƯỢNG XTNĐ HOẠT ĐỘNG TRÊN KHU VỰC BIỂN ĐÔNG**

Chương này luận án sẽ tiến hành thử nghiệm 3 trường hợp để lựa chọn nhân tố dự báo cho 4 mô hình thống kê dự báo số lượng XTNĐ hoạt động trên khu vực Biển Đông. Đầu tiên luận án tiến hành khảo sát bộ dữ liệu 300 nhân tố dự tuyển được mở rộng từ bộ 30 chỉ số khí hậu với độ trễ từ tháng thứ 3 tới tháng 12 trên giai đoạn 1979-2010. Trường hợp đầu tiên, luận án sử dụng các nhân tố dự báo được xác định thông qua kỹ thuật phân tích thành phần chính của tập dữ liệu 300 nhân tố trên để lấy ra các EOF phù hợp làm nhân tố dự báo. Tiếp cận theo cách khác, ở trường hợp thứ 2 các nhân tố dự báo được tuyển chọn, xác định bằng kỹ thuật hồi

quy từng bước từ tập dữ liệu 300 nhân tố trên. Kết quả mô phỏng và dự báo số lượng XTND trên khu vực Biển Đông được xây dựng trên bộ số liệu phụ thuộc (1980-2010) và đánh giá, kiểm nghiệm trên bộ số liệu độc lập (2011-2017).

Trường hợp tiếp theo (trường hợp 3), luận án lựa chọn các nhân tố dự tuyển qua việc khảo sát, đánh giá mức độ tương quan giữa bộ sản phẩm đầu ra của mô hình CFSv2 giai đoạn 1983-2009, ở trường hợp này giai đoạn 1983-2009 được lựa chọn để xây dựng phương trình dự báo, bộ số liệu giai đoạn 2012-2016 được sử dụng cho việc kiểm nghiệm độc lập và đánh giá khả năng của phương trình dự báo.

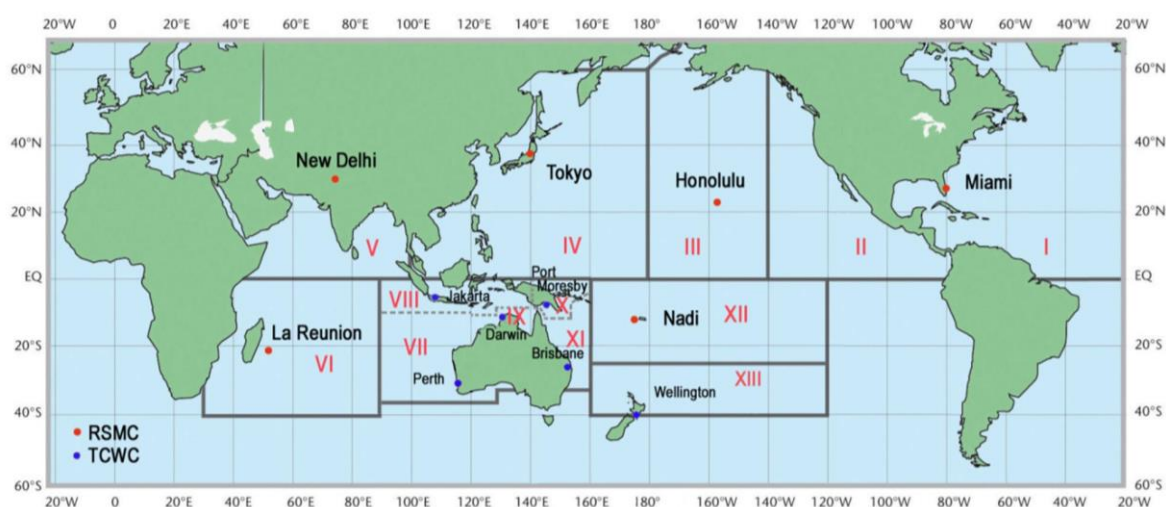
Việc đánh giá khả năng mô phỏng và dự báo hạn mùa số lượng XTND hoạt động trên khu vực Biển Đông của các mô hình được thực hiện thông qua các chỉ số thống kê đánh giá chất lượng và kỹ năng dự báo trên các chuỗi thời gian tương ứng.

### **KẾT LUẬN VÀ KIẾN NGHỊ**

Phần Kết luận sẽ trình bày tóm tắt các kết quả chủ yếu của luận án, những điểm mới đã đạt được 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 I. TỔNG QUAN VỀ ĐẶC ĐIỂM HOẠT ĐỘNG XTNĐ VÀ DỰ BÁO HẠN MÙA XTNĐ TRÊN BIỂN ĐÔNG

Các kết quả nghiên cứu từ Chương trình Bão nhiệt đới của Tổ chức Khí tượng thế giới (WMO) cho thấy, trên phạm vi toàn cầu XTNĐ thường xuyên hoạt động trên 13 vùng biển thuộc bảy khu vực. Cũng theo WMO, các Trung tâm Khí tượng Chuyên ngành Khu vực (RSMC) và Trung tâm Cảnh báo Bão nhiệt đới (TCWC) được chỉ định và có trách nhiệm theo dõi, ban hành các bản tin, cảnh báo, và thông báo về các XTNĐ nằm trong khu vực họ chịu trách nhiệm (Hình 1). Theo phân cấp, TCWC cung cấp thông tin trong phạm vi các vùng nhỏ hơn [103]. Trên thực tế RSMC và TCWC không phải là những tổ chức duy nhất cung cấp thông tin về XTNĐ mà tùy mục đích, nhu cầu và mức độ quan tâm mà các quốc gia riêng lẻ cũng có thể tham gia vào hoạt động này.



Hình 1.1 Các vùng hoạt động chính của XTNĐ trên quy mô toàn cầu (theo WMO, <https://www.wmo.int/pages/prog/www/tcp/Advisories-RSMCs.html>)

Như vậy có thể thấy rằng XTNĐ hoạt động trên khu vực Biển Đông thuộc phạm vi của vùng IV, đây là phần phía tây của khu vực Tây Bắc Thái Bình Dương (TBTBD) (Hình 1.1, Bảng 1.1) do đó các tổng quan về đặc điểm hoạt động và vấn đề nghiên cứu dự báo hạn mùa XTNĐ trong luận án được tập trung chủ yếu ở khu vực TBTBD và Biển Đông để từ đó thấy được những tồn tại mà luận án có thể góp phần làm sáng tỏ.

Bảng 1.1 Phạm vi không gian của các vùng, khu vực hoạt động chính của XTNĐ trên quy mô toàn cầu

(Theo WMO, <https://www.wmo.int/pages/prog/www/tcp/Advisories-RSMCs.html>)

Vùng	Khu vực	Tổ chức theo dõi, cảnh báo	Phạm vi
I-II	Bắc Đại Tây Dương và Đông Thái Bình Dương	Trung tâm Bão Quốc gia Hoa Kỳ (RSMC Miami)	Bắc xích đạo, bờ biển Châu Phi tới 140°W
III	Trung tâm Thái Bình Dương	Trung tâm Bão Thái Bình Dương – Hoa Kỳ (RSMC Honolulu)	Bắc xích đạo, 140 - 180°W
IV	Tây Bắc Thái Bình Dương	Cơ quan Khí tượng Nhật Bản (RSMC Tokyo)	Xích đạo - 60°N, 180 - 100°E
V	Bắc Ấn Độ Dương	Cơ quan Khí tượng Ấn Độ (RSMC New Delhi)	Bắc xích đạo, 100 - 45°E
VI	Tây Nam Ấn Độ Dương	Cơ quan Khí tượng Pháp (RSMC La Réunion)	Xích đạo - 40°N, bờ biển Châu Phi - 90°E
VII-XI	Tây Nam Thái Bình Dương và Đông Nam Ấn Độ Dương	VII: Cơ quan Khí tượng Australia (TCWC Perth)	10 - 36°N, 90 - 160°E
		VIII: Cơ quan Khí tượng Indonesian (TCWC Jakarta)	0 - 10°N, 90 - 141°E
		IX: Cơ quan Khí tượng Australia (TCWC Darwin)	0 - 10°N, 141 - 160°E
		X: Cơ quan Khí tượng Ghi-nê (TCWC Port Moresby)	
		XI: Cơ quan Khí tượng Australia (TCWC Brisbane)	
XII-XIII	Nam Thái Bình Dương	XII: Cơ quan Khí tượng Fiji (RSMC Nadi)	0 - 25°N, 160°E-120°W
		XIII: Cơ quan Khí tượng New Zealand (TCWC Wellington)	25 - 40°N, 160°E - 120°W

## 1.1 Đặc điểm hoạt động của XTNĐ

Trên khu vực TBTBD, khi nghiên cứu mối quan hệ giữa ENSO và hoạt động của XTNĐ trên giai đoạn 1948-1982, sử dụng phương pháp phân tích phổ Chan (1985) đã tìm ra hai đỉnh phổ trong chuỗi thời gian này, tương ứng với dao động Nam ở khoảng thời gian từ 3 tới 3,5 năm và dao động tựa hai năm QBO (Quasi-biennial Oscillation). Hai đỉnh phổ này có mối liên hệ chặt chẽ với chỉ số dao động Nam SOI (Southern Oscillation index), cụ thể kết quả cho thấy SOI đạt đỉnh trước gần 1 năm (11 tháng) trong khi QBO thể hiện dao động cùng pha [26].

Những nghiên cứu về xu thế dài hạn trong hoạt động của XTNĐ trên vùng biển Trung Quốc cũng đã được Chan và Shi (2000) đưa ra khi nghiên cứu đánh giá số liệu quan trắc trên khu vực này và số liệu lịch sử về XTNĐ đổ bộ vào tỉnh Quảng Đông giai đoạn 1470-1931 [31]. Để giảm thiểu mức độ không chắc chắn của bộ dữ liệu trên, các tác giả đã khảo sát thêm 2 bộ số liệu mới, được tạo ra bằng cách lấy trung bình từng giai đoạn 5 năm và 10 năm một, để có được 3 bộ số liệu với dung lượng mẫu tương ứng là 462, 92 và 46. Nghiên cứu sự biến đổi theo thời gian của 3 chuỗi số liệu này, các tác giả đã cho thấy tần suất XTNĐ đổ bộ vào khu vực tỉnh Quảng Đông, Trung Quốc có xu hướng tăng tuyến tính đồng thời cũng chỉ ra sự biến đổi theo quy luật thập kỷ và thập niên tại đây.

Ở một nghiên cứu khác trên khu vực TBTBD, Camargo và cộng sự (2004) đã xem xét ảnh hưởng của ENSO tới cường độ XTNĐ bằng việc phân tích mối quan hệ giữa năng lượng xoáy bão tích lũy ACE (Accumulated cyclone energy) với các yếu tố liên quan khác trong giai đoạn 1950-2002 [24]. Kết quả chỉ ra ACE có tương quan thuận với chỉ số ENSO và trong năm El Niño XTNĐ có xu hướng mạnh hơn, tồn tại lâu hơn ở những năm La Niña. Trong khoảng thời gian giữa mùa bão (tháng 7 tới 10) thậm chí tháng 6 tới tháng 11, ACE thể hiện tương quan khá tốt với ENSO. Giải thích về mối tương quan này, nhóm tác giả đưa ra một số quan điểm có thể kể ra như những biến động lớn trong năm liên quan tới thời gian tồn tại, cường độ và số lượng của các XTNĐ góp phần tạo nên mối quan hệ giữa ENSO và ACE, trong đó thời gian tồn tại của xoáy đường như là quan trọng nhất.

Ở cách tiếp cận khác, Wu và cộng sự (2004) đã sử dụng kỹ thuật tăng dung lượng mẫu (bootstrap) để nghiên cứu về các tác động của ENSO tới những biến đổi trong mẫu hình đổ bộ của XTNĐ ở khu vực TBTBD [93]. Kết quả cho thấy, so với các năm trung tính, số lượng XTNĐ đổ bộ vào khu vực TBTBD giảm đáng kể so với những năm El Niño, ngoại trừ tại khu vực Nhật Bản và bán đảo Triều Tiên. Mặt khác, số lượng XTNĐ ở khu vực TBTBD tăng lên và tăng cao nhất ở khu vực Trung Quốc trong những năm La Niña. Các tác giả cũng đặt ra giả thuyết về sự suy giảm số lượng XTNĐ đổ bộ vào khu vực TBTBD trong những năm El Niño có liên quan đến vị trí (hình thành) khởi phát của XTNĐ dịch ra phía đông và hoạt động yếu đi của áp cao cận nhiệt đới mực 500-hPa ở khu vực 130°E. Ngược lại, sự gia tăng về số lượng các XTNĐ trong những năm La Niña xuất hiện có liên quan đến vị trí hình thành trung bình bị dịch chuyển về phía tây và hoạt động mạnh lên của áp cao cận nhiệt đới mực 500-hPa.

Trên khu vực Đông Bắc Thái Bình Dương, Jien và cộng sự (2015) đã phân loại trạng thái các năm ENSO là El Niño, La Niña hay trung tính bằng cách sử dụng chỉ số ENSO đa biến MEI (multivariate ENSO index) [44]. Nhóm tác giả xác định các chỉ số hoạt động XTNĐ thuần (net tropical cyclone activity index) và chỉ số phân tán năng lượng (power dissipation index) để khái quát hóa vùng hoạt động và cường độ của các XTNĐ trong giai đoạn 1971 – 2012. Kết quả cho thấy, vào các năm El Niño XTNĐ hoạt động thường xuyên và dữ dội hơn so với năm La Niña. Bên cạnh đó, khi xem xét tương quan giữa chỉ số MEI và chỉ số hoạt động XTNĐ thuần; chỉ số phân tán năng lượng nhóm tác giả cũng đã chỉ ra việc tồn tại mối quan hệ chặt chẽ giữa ENSO với các vùng hoạt động cũng như cường độ của XTNĐ. Những mối quan hệ này được kì vọng có thể cải thiện dự báo ngắn hạn vùng hoạt động và cường độ XTNĐ theo mùa tại khu vực Đông Bắc Thái Bình Dương.

Ở Việt Nam, khi nghiên cứu ảnh hưởng của hiện tượng ENSO đến hoạt động của bão và áp thấp nhiệt đới, Nguyễn Đức Ngữ (2002) đã chỉ ra rằng, trong giai đoạn 1956 – 2000 (45 năm) có 311 cơn bão và áp thấp nhiệt đới ảnh hưởng trực tiếp đến Việt Nam, trung bình mỗi năm có 6,9 cơn (0,58 cơn/tháng) [8]. Cùng thời gian trên

nghiên cứu này đã ghi nhận có tổng số 150 tháng El Niño với 63 cơn XTNĐ ảnh hưởng trực tiếp, trong đó trung bình mỗi tháng ở thời kỳ El Niño có 0,42 cơn, ít hơn trung bình nhiều năm khoảng 28%. Trong khi đó, trung bình mỗi tháng ở thời kỳ La Niña có 0,80 cơn (86 cơn/107 tháng), nhiều hơn trung bình nhiều năm khoảng 38%. Trong cả mùa bão (từ tháng 6 đến tháng 12), trung bình nhiều năm có 6,64 cơn, mỗi tháng mùa bão có 0,95 cơn. Bên cạnh đó nghiên cứu này cũng chỉ ra, trong thời kỳ El Niño, trung bình cả mùa bão có 4,83 cơn trong đó trung bình mỗi tháng mùa bão có 0,69 cơn, ít hơn khoảng 27% so với trung bình nhiều năm. Ngược lại, trong thời kỳ La Niña, trung bình mùa bão có 9,17 cơn và 1,31 cơn ở từng tháng mùa bão, nhiều hơn trung bình nhiều năm khoảng 38%. Ngoài ra, trong thời kỳ El Niño, XTNĐ thường tập trung vào giữa mùa bão (tháng 7, 8, 9), trong khi ở thời kỳ La Niña, XTNĐ thường nhiều hơn vào nửa cuối mùa bão (tháng 9, 10, 11) [8].

Khảo sát mối quan hệ giữa biến động theo mùa và nhiều năm của trường nhiệt độ nước mặt biển với hoạt động của XTNĐ tại khu vực Biển Đông và TBTBD, Đinh Văn Ưu và cộng sự (2005) đã chỉ ra có sự biến động đáng kể của trường nhiệt độ nước mặt biển và hoạt động của XTNĐ trên hai khu vực này trong giai đoạn 1960-2000, đặc biệt ở giai đoạn 1990-2000 [19]. Thông qua việc tính toán các chỉ số khí hậu, nghiên cứu này cho thấy khi hiện tượng El Niño hoạt động mạnh thì hoạt động của XTNĐ trên toàn khu vực giảm. Mặc dù nghiên cứu chưa định lượng được mức độ quan hệ giữa các biến động này với chu kỳ dao động mùa song các tác giả cũng đã chỉ ra những tác động đáng kể của một số hiện tượng quy mô toàn cầu như ENSO lên chế độ nhiệt và hoạt động XTNĐ trên khu vực TBTBD, Biển Đông. Nhóm tác giả cũng đề xuất, những mối quan hệ này cần được đi sâu nghiên cứu để xác lập các quy luật biến động từ đó làm căn cứ xây dựng các mô hình dự báo.

Nghiên cứu xu hướng hoạt động của XTNĐ bằng cách phân loại theo vùng ảnh hưởng và theo cường độ từ đó phân tích xu hướng hoạt động, Nguyễn Văn Tuyên (2007) đã chỉ ra trong thời kỳ 1951-2006, hoạt động của XTNĐ trên khu vực TBTBD có xu hướng giảm về số lượng, trong đó số cơn bão yếu và trung bình có xu hướng giảm, còn số cơn bão mạnh lại có xu hướng tăng lên. Bên cạnh đó, những

XTNĐ đổ bộ vào Biển Đông nhưng không vào vùng ven biển và đất liền Việt Nam lại có xu hướng tăng về số lượng [16]. Cũng trong công trình này, tác giả nhận định rằng XTNĐ có xu hướng tăng lên ở hai vùng Trung Bộ và Nam Bộ nhưng ở vùng Bắc Bộ lại có xu hướng giảm.

Khi xem xét đánh giá quy luật biến động dài hạn và xu thế biến đổi số lượng bão và áp thấp nhiệt đới trên khu vực TBTBD, Biển Đông và ven biển Việt Nam, Đinh Văn Ưu (2009) đã cho thấy ở giai đoạn 1959-2008 số lượng trung bình năm của bão và siêu bão (theo phân loại trong nghiên cứu, XTNĐ là bão khi vận tốc gió nằm trong khoảng 17 tới 33m/s hay từ cấp 8 đến cấp 11, siêu bão khi vận tốc gió từ 33m/s hay cấp gió từ 12 trở lên) dao động theo các chu kỳ dài từ hai năm đến nhiều chục năm [20]. Đồng thời kết quả nghiên cứu này cũng cho thấy xu thế gia tăng số lượng bão và siêu bão ở các khu vực trên không rõ ràng, thậm chí số lượng siêu bão còn có xu thế giảm. Trong giai đoạn 1959-2008, số lượng bão gây ảnh hưởng trực tiếp đến ven bờ vịnh Bắc Bộ giảm, trong khi ở Nam Trung Bộ và Nam Bộ lại gia tăng. Những dao động này cho thấy có khả năng hoạt động của XTNĐ trên khu vực TBTBD, Biển Đông và ven biển Việt Nam chịu tác động của các dao động quy mô lớn như tựa 2 năm, El Niño và nhiều chục năm Thái Bình Dương.

Khi khai thác số liệu XTNĐ trong giai đoạn 1945-2007 trên bộ số liệu Unsys để xem xét đặc điểm hoạt động của XTNĐ ở bảy vùng biển gần bờ Việt Nam, Vũ Thanh Hằng và cộng sự (2010) đã nhận thấy số lượng XTNĐ ở các vùng biển gần bờ Việt Nam đều có xu thế tăng lên [3]. Kết quả của nghiên cứu này cũng đã cho thấy số lượng XTNĐ trong những năm La Niña thường nhiều hơn trong những năm El Niño. Giai đoạn 1996-2000 có nhiều XTNĐ nhất và hoạt động của XTNĐ có xu hướng dịch về phía nam nhưng ở mức độ biến động nhỏ.

Năm 2011, Đinh Văn Ưu và cộng sự đã sử dụng số liệu từ Trung tâm Liên hợp Cảnh báo bão của Mỹ với chuỗi thời gian từ năm 1960 đến 2010, sau đó nghiên cứu chia toàn bộ dải bờ biển Việt Nam thành từng đoạn cách nhau 1 vĩ độ, từ 5°N đến 22°N, trong từng đoạn, số lượng XTNĐ nằm trong phạm vi cách bờ 3° về phía đông được lựa chọn để xử lý [21]. Kết quả nghiên cứu cho thấy khả năng bão và áp

thấp nhiệt đới đổ bộ và gây nên các tác động lên đất liền ven biển Việt Nam tập trung ở vùng biển Bắc Bộ và Bắc Trung Bộ, trong đó cực đại trên đoạn bờ Hà Tĩnh - Nam Nghệ An. Dải ven bờ Bình Định - Quảng Ngãi có tần suất bão độ bộ lớn nhất trong khu vực Trung Trung Bộ. Đồng thời kết quả nghiên cứu cũng phát hiện dao động chu kỳ giữa các thập niên, trong đó số lượng XTNĐ ít hơn trong các thập niên 1980-1990, tương ứng pha ấm của dao động thập niên Thái Bình Dương. Tổng số lượng bão, áp thấp nhiệt đới thể hiện xu thế gia tăng ở khu vực Nam Bộ và dường như không đổi ở khu vực Bắc Bộ. Bên cạnh đó tại Trung Bộ, xu thế gia tăng được ghi nhận ở số lượng áp thấp nhiệt đới trong khi số lượng bão tại đây lại giảm.

Sử dụng 3 nguồn số liệu của Cơ quan Khí tượng Nhật Bản, Trung tâm Liên hợp Cảnh báo bão thuộc Hải quân Hoa Kỳ và Trung tâm Dự báo Khí tượng Thủy văn Trung ương trong thời kỳ 1959-2008, Đặng Hồng Nga và cộng sự (2011) đã nghiên cứu, đánh giá xu thế diễn biến tần số XTNĐ ở khu vực TBTBD và Biển Đông [7]. Trên cơ sở phân tích các kết quả thu được, nhóm tác giả đưa ra nhận định rằng khu vực đổ bộ của bão, áp thấp nhiệt đới vào đất liền Việt Nam có sự dịch chuyển về phía nam lãnh thổ đồng thời tần số bão hoạt động trên khu vực TBTBD và Biển Đông giảm nhưng số lượng bão rất mạnh đổ bộ vào Việt Nam có dấu hiệu gia tăng trong những thập kỷ gần đây.

Tổng hợp số liệu từ nhiều nguồn khác nhau, Nguyễn văn Thắng và cộng sự (2016) đã nghiên cứu ảnh hưởng của bão ở Việt Nam thời kỳ 1961-2014 [13]. Nhóm tác giả đã sử dụng nhiều nguồn số liệu như bản đồ đường đi của bão thời kỳ 1961-2014, đặc điểm Khí tượng Thủy văn hàng năm theo báo cáo từ Trung tâm Dự báo Khí tượng thủy văn quốc gia, kết quả nghiên cứu của đề tài cấp Bộ do Dương Liên Châu chủ trì năm 2004 về “Xây dựng cơ sở dữ liệu bão và áp thấp nhiệt đới trên khu vực Biển Đông và ảnh hưởng đến Việt Nam”, tài liệu thống kê ảnh hưởng của bão từ các Đài Khí tượng Thủy văn khu vực trên cả nước và số liệu IBTrACS của NOAA, các bản đồ đường đi của bão do Nhật Bản, Hồng Kông phát hành,.... Kết quả của nghiên cứu này đã chỉ ra trên toàn lãnh thổ Việt Nam, trong thời kỳ 1961-2014 có 364 cơn bão, áp thấp nhiệt đới đổ bộ và ảnh hưởng, trong đó chiếm tỷ

lệ trên 10% tập trung vào 5 tháng từ tháng 7 đến tháng 11. Tổng tỷ lệ % bão ảnh hưởng so với cả năm của 5 tháng này là 87%, tập trung nhiều nhất vào tháng 9, thấp nhất vào tháng 7. Trong thời kỳ 1959-2015, số lượng bão và áp thấp nhiệt đới có xu thế ít biến đổi nhưng có phân bố tập trung hơn vào cuối mùa bão, đây cũng là thời kỳ bão hoạt động chủ yếu ở phía nam. Bão mạnh đến rất mạnh có xu thế gia tăng.

Ở công trình mang tầm quy mô quốc gia, kịch bản biến đổi khí hậu và nước biển dâng cho Việt Nam do Bộ Tài nguyên và Môi trường chủ trì đã xây dựng và phát hành từ năm 2009, cập nhật bổ sung vào các năm 2012 và gần đây nhất là năm 2016 cũng đã trình bày các đặc điểm của XTNĐ trên Biển Đông [14]. Theo kịch bản 2016, trung bình cho thời kỳ 1959-2015, trên Biển Đông hàng năm có khoảng 12 cơn bão và áp thấp nhiệt đới hoạt động, trong đó khoảng 45% số cơn hình thành ngay trên Biển Đông và 55% số cơn hình thành từ Thái Bình Dương di chuyển vào. Mỗi năm có khoảng 7 cơn bão và áp thấp nhiệt đới ảnh hưởng đến Việt Nam, trong đó có 5 cơn đổ bộ hoặc ảnh hưởng trực tiếp đến đất liền nước ta. Nơi có tần suất hoạt động của bão và áp thấp nhiệt đới lớn nhất nằm ở phần giữa của khu vực Bắc Biển Đông. Khu vực bờ biển miền Trung từ 16°N đến 18°N và khu vực bờ biển Bắc Bộ (từ 20°N trở lên) có tần suất hoạt động của bão và áp thấp nhiệt đới cao nhất trong cả dải ven biển Việt Nam. Theo số liệu thời kỳ 1959-2015, bão và áp thấp nhiệt đới hoạt động trên Biển Đông, ảnh hưởng và đổ bộ vào Việt Nam là ít biến đổi. Tuy nhiên, biến động về số lượng bão, áp thấp nhiệt đới là khá rõ và những cơn bão mạnh (sức gió mạnh nhất từ cấp 12 trở lên) có xu thế tăng nhẹ. Mùa bão kết thúc muộn hơn và đường đi của bão có xu thế dịch chuyển về phía nam với nhiều cơn bão đổ bộ vào khu vực phía nam hơn trong những năm gần đây được xem là một trong những biểu hiện của biến đổi khí hậu.

## **1.2 Nghiên cứu dự báo hạn mùa XTNĐ**

Hiện nay, với những biểu hiện bất thường của thời tiết, khí hậu đặc biệt là các hiện tượng thời tiết cực đoan như bão, tố lốc... thì bài toán dự báo hạn mùa (seasonal forecast) XTNĐ đang là vấn đề được nhiều nhà khoa học, nhà quản lý trong và ngoài nước quan tâm bởi các kết quả của nó có ý nghĩa ứng dụng hết sức

thiết thực đối với các hoạt động kinh tế xã hội và an ninh quốc phòng. Nhìn chung sự khác nhau giữa dự báo khí hậu hạn mùa và dự báo thời tiết hạn ngắn không chỉ ở đích thời gian dự báo mà còn ở cách tiếp cận, phương pháp sử dụng và thậm chí định hướng sử dụng kết quả dự báo [10-11, 75]. Khác với dự báo thời tiết, dự báo hạn mùa không đưa ra dự báo chi tiết từng ngày và từ ngày này sang ngày khác. Kết quả của dự báo hạn mùa thường đưa ra những thông tin về đặc trưng thống kê trung bình theo mùa của thời tiết, chẳng hạn như với thời hạn của dự báo là trước 3 tháng [10-11, 68, 90]. Ví dụ, dự báo mùa có thể đưa ra thông tin dự báo về điều kiện khí hậu của mùa tới là sẽ ẩm hơn, khô hơn, nóng hơn hay lạnh hơn so với thường lệ. Cơ sở vật lý để đưa ra những thông tin ước tính này xuất phát từ tính có thể dự báo được các tín hiệu/yếu tố trong đại dương ở qui mô thời gian mùa đến những qui mô nhỏ hơn trên bề mặt lục địa và trong khí quyển [11, 67]. Điểm mấu chốt trong bài toán dự báo mùa là các hiện tượng tương tác giữa đại dương - khí quyển như hiện tượng El Niño là một ví dụ xảy ra ở vùng Thái Bình Dương nhiệt đới và có thể dự báo trước đến 6 tháng hoặc xa hơn nữa [10-11, 45, 91]. Thời hạn trong các dự báo hạn mùa thường từ qui mô tháng cho đến một năm tùy thuộc yếu tố dự báo, trong đó các hạn dự báo phổ biến là 1, 3, 6 và 9 tháng.

Trên thực tế, trong các mô hình số dự báo thời tiết khi thời gian tích phân tăng lên, quá trình tương tác giữa khí quyển và đại dương sẽ đóng một vai trò cực kỳ quan trọng trong việc điều khiển và chi phối hoàn lưu khí quyển [75]. Do vậy, yếu tố nhiệt độ mặt nước biển (SST) và các quá trình vật lý khác trên bề mặt trái đất ảnh hưởng nhiều hơn tới kết quả dự báo hạn mùa so với các điều kiện ban đầu của khí quyển. Trên qui mô toàn cầu, sự thay đổi của yếu tố SST là nguyên nhân chính dẫn đến những thay đổi của khí hậu từ năm này sang năm khác. Đáng chú ý, ở vùng nhiệt đới, dị thường SST liên quan đến hoạt động của đối lưu sâu điều khiển phần lớn hoàn lưu khí quyển toàn cầu do đó yếu tố này là thành phần quan trọng trong bài toán dự báo hạn mùa [79]. Trên cơ sở đó, cho đến nay việc giải quyết các bài toán dự báo mùa thường được tiếp cận theo phương pháp thống kê và phương pháp động lực hoặc là sự kết hợp của cả phương pháp thống kê và động lực.

Đối với bài toán dự báo hạn mùa XTNĐ, có thể phân chia dựa trên thời hạn dự báo hay quy mô thời gian dự báo: 1) Dự báo XTNĐ hạn nội mùa: Thời hạn dự báo thường khoảng trên 10 ngày đến 1 tháng; và 2) Dự báo XTNĐ hạn mùa: Thời hạn dự báo trong khoảng từ 1 tháng đến dưới 12 tháng [10, 69]. Nếu như các dự báo XTNĐ ở hạn ngắn, thông tin dự báo thường bao gồm vị trí, cường độ và phạm vi ảnh hưởng của một cơn XTNĐ cụ thể đang hoạt động với mức độ chi tiết thậm chí đến từng giờ trong thời hạn dự báo thì dự báo XTNĐ hạn mùa và nội mùa chỉ có thể cung cấp thông tin về khả năng có XTNĐ hoạt động hay không trong thời hạn dự báo, và nếu có thì sẽ có khoảng bao nhiêu cơn XTNĐ và khả năng xuất hiện XTNĐ ở vùng nào. Nói cách khác, dự báo hạn mùa cố gắng dự báo điều kiện thời tiết tương lai và đối với dự báo XTNĐ hạn mùa chỉ có thể cung cấp thông tin dự báo về số lượng và vùng hoạt động của các XTNĐ có thể xuất hiện trong tháng tới hoặc mùa tới mà khó có thể thực hiện dự báo cho một XTNĐ cụ thể [79].

- *Dự báo hạn mùa XTNĐ bằng phương pháp động lực*: Tiếp cận theo phương pháp động lực, bài toán dự báo hạn mùa hoạt động của XTNĐ được đưa ra dựa trên cơ sở các tín hiệu, hiện tượng có nguồn gốc đại dương như SST, băng biển, ... hay những thay đổi trong chế độ bức xạ (năng lượng mặt trời, khí nhà kính, sol khí...) và điều kiện bề mặt đất (đặc biệt là độ ẩm đất) có quy mô thời gian từ hạn mùa cho tới năm [64, 76]. Theo Gray (1979), khả năng dự báo hoạt động của các XTNĐ theo mùa phụ thuộc lớn vào các điều kiện khí quyển và đại dương, như trường SST và độ đứt gió vĩ hướng [40]. Ở vùng nhiệt đới, những thay đổi trong hoàn lưu quy mô lớn có liên quan đến thay đổi phân bố SST, hoặc như hiện tượng ENSO, được coi là có thể dự đoán ở quy mô thời gian theo mùa [50, 62]. Từ quan điểm này, hoạt động XTNĐ theo mùa có thể được coi là một quá trình ngẫu nhiên được điều khiển bởi các điều kiện khí hậu theo mùa.

Trên thực tế, dự báo mùa hoạt động của XTNĐ bằng các mô hình hoàn lưu đại dương – khí quyển nói chung đã được nghiên cứu từ những năm cuối thập niên 1990 và Trung tâm Dự báo thời tiết hạn vừa châu Âu (ECMWF) ban hành từ năm 2001 [84]. Những hệ thống dự báo trước đó và nhiều hệ thống dự báo hạn mùa hiện tại sử

dụng sản phẩm mô hình khí hậu toàn cầu (GCM) ở độ phân giải thô để dự đoán sự phát triển theo mùa của các yếu tố khí hậu quy mô lớn. Cùng với sự phát triển của khoa học kỹ thuật, sản phẩm của các mô hình hoàn lưu chung khí quyển toàn cầu (AGCMs) đang ngày càng được cải tiến để mô phỏng chính xác hơn hoạt động của XTNĐ, trong đó bao gồm cả vị trí khởi phát và tần suất hoạt động hay thời gian tồn tại của các XTNĐ [39, 42, 69]. Như vậy, hệ thống dự báo mùa dựa trên mô hình do đó cũng đã trở thành một công cụ hữu ích cho bài toán dự báo XTNĐ hạn mùa.

Sử dụng AGCMs độ phân giải cao, với trường nhiệt độ mực nước biển SST hoặc dị thường nhiệt độ mực biển (SSTA) là một trong số những cách tiếp cận được nhiều tác giả đề cập tới [33-34, 56-57, 101]. Chen, Lin (2011, 2013) đã sử dụng AGCM ở độ phân giải 25 km mô phỏng tần suất hoạt động XTNĐ trên khu vực Bắc Thái Bình Dương trong giai đoạn 1990-2010 nhằm nghiên cứu khả năng dự báo hạn mùa của mô hình động lực với giả định SSTA ổn định tuy nhiên các kết quả lại cho thấy việc mô phỏng còn gặp nhiều sai số. Các tác giả đã kết luận rằng giả định về SSTA ổn định có thể ít được áp dụng cho khu vực Tây Thái Bình Dương và không hoàn toàn thích hợp cho mùa bão năm 2011-2013. Các tác giả cũng cho rằng việc cải thiện hơn nữa các dự đoán theo mùa XTNĐ một phần phụ thuộc vào việc cải thiện độ phân giải mô hình [33-34].

Tại Việt Nam các nghiên cứu dự báo hạn mùa XTNĐ cho đến nay còn khá ít so với tính chất nguy hiểm và quy mô ảnh hưởng của nó. Trong khuôn khổ đề tài cấp Nhà nước KC08.29/06-10, Phan Văn Tân và cộng sự (2010) cũng đã sử dụng các chỉ số khí hậu làm nhân tố dự báo để xây dựng các phương trình dự báo hạn mùa số lượng và số ngày hoạt động của XTNĐ trên khu vực Biển Đông và dọc bờ biển Việt Nam bằng phương pháp thống kê truyền thống [11]. Theo tác giả, mặc dù sai số dự báo còn khá lớn, song có thể thực hiện dự báo hạn mùa số lượng và số ngày bão hoạt động cho khu vực Biển Đông và dọc bờ biển Việt Nam bằng mô hình hồi quy tuyến tính nhiều biến khi sử dụng các chỉ số khí hậu làm nhân tố dự báo. Về vai trò của các nhân tố dự báo, tác giả cho rằng khó có thể lí giải về mặt cơ chế vật lí của chúng đối với sự hoạt động của XTNĐ. Tuy vậy vẫn có thể thấy những nhân

tổ đặc trưng cho dao động quy mô lớn có ý nghĩa nhất định trong việc dự báo hạn mùa XTNĐ. Vai trò của các nhân tố nhóm ENSO là quan trọng, trong đó khu vực Nino4 cần được chú ý nhiều hơn.

Hướng ứng dụng mô hình động lực trong mô phỏng và dự báo bão hạn mùa ở Việt Nam được đề cập đến lần đầu tiên bởi Phan Văn Tân và cộng sự (2011) [10]. Tiếp theo công trình này, năm 2014 nhóm tác giả đã tiến hành thử nghiệm ứng dụng mô hình khí hậu khu vực RegCM4.2 để dự báo số lượng XTNĐ trên khu vực TBTBD và Biển Đông bằng việc hạ qui mô động lực sản phẩm dự báo của mô hình khí hậu toàn cầu NCEP CFSv2 từ độ phân giải ngang  $1^\circ \times 1^\circ$  kinh vĩ về độ phân giải  $36 \text{ km}$  [69]. Kết quả nhận được cho thấy, mô hình RegCM4.2 có khả năng biểu diễn tốt hơn so với NCEP CFSv2 cả về số lượng và quỹ đạo chuyển động của XTNĐ trên khu vực Biển Đông. Trong thí nghiệm, mô hình RegCM4.2 sử dụng số liệu tái phân tích CFSR với độ phân giải ngang  $2.5^\circ \times 2.5^\circ$  kinh vĩ làm dữ liệu đầu vào để mô phỏng cho thời kỳ 1995-2010 và dự báo thời gian thực cho thời kỳ 2012-2013, kết quả cho thấy mô hình RegCM4.2 đã nắm bắt được khá tốt số lượng XTNĐ. Từ đây, nhóm tác giả đã cho rằng việc hạ qui mô động lực sản phẩm dự báo toàn cầu có độ phân giải thô bằng mô hình khí hậu khu vực là quan trọng và cần thiết trong dự báo bão hạn mùa.

Mặc dù giải pháp mô hình có khả năng mô phỏng nhất định hoạt động XTNĐ, tuy nhiên trên thực tế vẫn tồn tại nhiều sai sót liên quan tới đặc điểm khí hậu của XTNĐ, chẳng hạn như tần số, quỹ đạo, cấu trúc và cường độ của XTNĐ [25, 63, 71]. Trên lý thuyết, mô hình động lực với độ phân giải cao là điều cần thiết cho các mô phỏng thực tế của XTNĐ, tuy nhiên thực tế rất ít nghiên cứu trực tiếp kiểm tra ảnh hưởng này vào khả năng dự báo hoạt động XTNĐ theo mùa một phần cũng bởi khả năng tính toán của hệ thống máy tính và phần khác đến từ việc tham số hóa mô tả các quá trình vật lý bên trong, điều kiện ban đầu và điều kiện biên, ... Trong khi đó, theo Lorenz (1963) những điều kiện ban đầu khác nhau quá trình tích phân mô hình có thể thu được các nghiệm số phân kì theo thời gian [60]. Như vậy, những sai số nhỏ không xác định được ở trạng thái ban đầu của khí quyển, đại dương sẽ tạo nên những sai số đủ lớn sau một khoảng thời gian tích phân. Chính

những điều này đã khiến cho bài toán dự báo hạn mùa XTNĐ theo hướng tiếp cận động lực vẫn là một thách thức với trình độ khoa học công nghệ hiện nay.

- *Dự báo hạn mùa XTNĐ bằng phương pháp thống kê*: Tiếp cận theo phương pháp thống kê truyền thống, dự báo mùa hoạt động của XTNĐ cũng được phát triển mạnh mẽ qua nhiều công trình nghiên cứu của Landsea và cộng sự (1994, 1998) [55-56]; Gray và cộng sự (1994) [41]; Nicholls và cộng sự (1998) [66]; Chan và cộng sự (1995, 1998) [28-29]; Elsner và cộng sự (2000) [37]; ...v.v. Trong các công trình này, yếu tố dự báo thường là số lượng và số ngày hoạt động của XTNĐ trên các vùng biển khác nhau, chủ yếu ở Đại Tây Dương và Tây Thái Bình Dương. Các nhân tố dự báo có thể là những nhân tố thuộc nhóm ENSO, như dị thường nhiệt độ bề mặt biển SSTA trên các vùng Nino1+2, Nino3, Nino4, Nino3.4, hoặc các đặc trưng hoàn lưu khí quyển và đại dương qui mô lớn như chỉ số dao động tựa hai năm tầng bình lưu QBO, chỉ số dao động Nam SOI, ...

Trên khu vực TBTBD, hoạt động của XTNĐ cũng đã được nhiều tác giả quan tâm như Chan và cộng sự (1998, 2000, 2001) [29-32]; Chu và cộng sự (2010) [36]; Lu và cộng sự (2010) [61], ...v.v. Trong các công trình này, phương pháp hồi quy từng bước nhiều biến thường được các tác giả sử dụng để lọc nhân tố dự báo. Chẳng hạn, Kwon và cộng sự (2007) đã sử dụng mô hình thống kê để dự báo hạn mùa hoạt động của XTNĐ trên khu vực TBTBD [54]. Để đưa ra được bức tranh toàn diện về hoạt động của XTNĐ, bên cạnh dự báo số lượng XTNĐ nói chung các tác giả còn dự báo cả số lượng các cơn bão mạnh cũng như chỉ số “dị thường bão hoạt động” (Normalized Typhoon Activity). Mô hình dựa trên cơ sở phương pháp hồi quy tuyến tính đa biến trong đó các nhân tố dự báo được lựa chọn theo nguyên tắc làm cực tiểu hoá sai số dự báo. Kết quả của nghiên cứu cho thấy xác suất dự báo tổ hợp các mô hình thống kê là hợp lý hơn so với thực tế.

Ngoài ra, còn có thể kể đến các công trình sử dụng mô hình thống kê để dự báo tần suất hoạt động hoặc vùng hoạt động của XTNĐ như Chan (1995) [28], Chan và cộng sự (1998, 2001) [29-30], Klotzbach (2007) [52], Wang và cộng sự (2009) [86], Zhang và cộng sự (2015) [100], ...v.v.

Kết hợp những ưu điểm của phương pháp thống kê truyền thống và phương pháp động lực, phương pháp thống kê dựa trên sản phẩm của các mô hình động lực (hay còn được viết ngắn gọn là mô hình thống kê động lực) cũng đã được Wang và cộng sự (2009) [86]; Kim và cộng sự (2010) [49]; Vecchi và cộng sự (2011) [83] nghiên cứu, phát triển và ứng dụng để dự báo mùa hoạt động của bão vùng Bắc Đại Tây Dương hay Li, Kim và cộng sự (2013) dự báo bão cho vùng TBTBD [48, 58].

Một mô hình thống kê động lực thường được xây dựng theo một quy trình hai bước. Đầu tiên, một mô hình thống kê được xây dựng dựa trên các mối quan hệ thống kê từ số liệu quan trắc hoặc sản phẩm dự báo theo mùa của mô hình khí hậu. Mô hình thống kê này sau đó được áp dụng để dự đoán theo mùa hoạt động XTNĐ với các nhân tố dự báo là sản phẩm đầu ra của mô hình động lực. Mô hình thống kê động lực cung cấp một cách khác để dự đoán hoạt động XTNĐ theo mùa. Trên thực tế, một số nghiên cứu trước đây Wang và cộng sự (2009) [86]; Vecchi và cộng sự (2011) [83]; Li và cộng sự (2013) [58], Kim và cộng sự (2017) [51] đã chứng minh rằng mô hình thống kê động lực có thể có kỹ năng dự đoán hoạt động XTNĐ tốt hơn so với mô hình thống kê truyền thống hoặc mô hình động lực đơn thuần. Điểm quan trọng để xây dựng được mô hình thống kê hoặc mô hình thống kê động lực có kỹ năng dự báo tốt hoạt động XTNĐ theo mùa là việc lựa chọn được các nhân tố có quan hệ, ảnh hưởng mạnh đến hoạt động XTNĐ trong khu vực nghiên cứu. Thông thường vấn đề này thường được thực hiện bằng cách tính toán và phân tích các mối tương quan giữa yếu tố dự báo và các dữ liệu quan trắc.

Ở cách tiếp cận này, Li và cộng sự (2013) đã sử dụng mô hình thống kê động lực dự báo số lượng XTNĐ hàng năm khu vực TBTBD dựa trên sản phẩm CFSv2 của NCEP [58]. Trong công trình nghiên cứu này, nhóm tác giả sử dụng dữ liệu XTNĐ của JMA cho khu vực TBTBD, dữ liệu quan trắc SST của NOAA và các trường khí quyển khác gồm: gió vĩ hướng mực 200mb (U200), gió vĩ hướng mực 850mb (U850), độ đứt gió giữa hai mực 200mb và 850mb (U200-U850), độ cao địa thế vị mực 500mb (HGT500) từ số liệu tái phân tích CFSR. Bộ dữ liệu dự báo lại quá khứ của mô hình CFSv2 giai đoạn 1982 - 2010 với hạn dự báo 9 tháng. Số liệu dự

báo hạn 9 tháng cho năm 2011 và 2012. Để xác định các nhân tố đối với dự báo XTNĐ, tác giả dựa trên các hệ số tương quan giữa số XTNĐ quan trắc hàng năm và các trường khí quyển và đại dương quy mô lớn là sản phẩm dự báo của CFSv2. Bằng cách này các nhân tố dự báo được lựa chọn gồm: nhiệt độ mặt nước biển vùng nhiệt đới Bắc Đại Tây Dương ở khu vực 10 - 20°N; 280 - 330°W (NATLSST - North Atlantic SST): với hệ số tương quan -0,72; độ đứt gió giữa hai mực 200hPa và 850hPa ở khu vực 10 - 20°N; 130°E - 190°W (WPVZWS - Western Pacific vertical zonal wind shear) với hệ số tương quan -0,73; vận tốc gió vĩ hướng mực 850hPa (U850) khu vực 10 - 20°N; 130°E - 190°W với hệ số tương quan 0,67 và độ cao địa thế vị mực 500hPa (HGT500) ở khu vực (10 - 30°N; 120 - 160°E) với hệ số tương quan -0,69. Phương pháp hồi quy tuyến tính được sử dụng để xây dựng phương trình dự báo với các nhân tố dự báo lựa chọn. Ban đầu nhóm tác giả xây dựng phương trình dự báo cho từng nhân tố kết quả cho thấy phương trình xây dựng với nhân tố dự báo HGT500 cho kỹ năng dự báo tốt nhất với hệ số tương quan và sai số bình phương trung bình quân phương lần lượt là 0,7 và 3,2. Tiếp theo tác giả tiến hành khảo sát sự kết hợp giữa 2 nhân tố bất kỳ trong 4 nhân tố với mục đích xây dựng phương trình dự báo gồm 2 nhân tố. Thử nghiệm xây dựng phương trình dự báo với 2 nhân tố, tác giả nhận thấy mô hình với sự kết hợp hai nhân tố dự báo HGT500 và WPVZWS cho kết quả dự báo lại trong quá khứ tốt nhất. Đây cũng là 2 nhân tố được chọn để xây dựng phương trình dự báo cuối cùng sau khi so sánh với phương trình dự báo sử dụng nhân tố HGT500. Kết quả dự báo đưa ra hệ số tương quan cao nhất đạt 0,79, sai số RMSE tốt nhất là 2,75 và tỷ lệ thành công SRATE cao nhất đạt 77% đối với hạn dự báo từ 0 đến 5 tháng. Khi so sánh kết quả dự báo này với các sản phẩm dự báo thống kê của Trung tâm Cảnh báo bão TSR (Tropical Storm Risk) và Trung tâm Tác động Khí hậu Châu Á - Thái Bình Dương GCACIC (Guy Carpenter Asia-Pacific Climate Impact Centre) cho thấy mô hình thống kê động lực này hoàn toàn cạnh tranh được với các dự báo thống kê của TSR và GCACIC. Trong giai đoạn 10 năm (2003-2012), mô hình dự báo đạt được với RMSE từ 3,08 tới 1,55 và tỷ lệ thành công SRATE từ 70 đến 90% ở hạn dự báo từ 1 đến 3 tháng. Kết quả dự báo trong giai đoạn 10 năm cũng

chỉ ra các mô hình với hạn dự báo 0 tới 1 tháng có RMSE nhỏ hơn dự báo với hạn dự báo 3 tháng. Kỹ năng dự báo ở hạn dự báo 1 tháng và 3 tháng có tỷ lệ thành công đến 70%. Kết quả cũng cho thấy với hạn dự báo 0 tháng, tổng số XTNĐ ghi nhận được trong phạm vi dự báo là 9/10 năm (trừ năm 2009), cho thấy dự báo lại quá khứ dựa trên hạn dự báo 0 tháng cho kết quả tốt hơn.

Ở một nghiên cứu khác, Kim và cộng sự (2017) đã thống kê trên sản phẩm đa mô hình của APCC (Trung tâm Khí hậu châu Á – Thái Bình Dương) để dự báo XTNĐ hạn mùa khu vực TBTBD [51]. Các mô hình trong APCC gồm: 1) CGCM3 và CGCM4 (coupled general circulation models) là mô hình toàn cầu của cơ quan khí tượng Canada; 2) CGCM mô hình toàn cầu của Đại học Quốc gia Pusan; 3) GMAO (Global Modeling and Assimilation Office system) mô hình toàn cầu của NASA và 4) CFSv2 hệ thống dự báo khí hậu phiên bản 2 của NCEP. Nghiên cứu sử dụng nguồn số liệu XTNĐ của trung tâm hỗn hợp cảnh báo bão JTWC. Trong nghiên cứu tập trung vào số XTNĐ hình thành ở các tháng 7, 8, 9, 10 trong năm, giai đoạn chiếm tới 80% số cơn bão hằng năm của khu vực TBTBD. Dữ liệu quan trắc nhiệt độ bề mặt biển SST và dữ liệu gió khai thác từ NOAA và NCEP với lưới  $1^{\circ} \times 1^{\circ}$  và  $0,25^{\circ} \times 0,25^{\circ}$  tương ứng. Nhóm tác giả sử dụng trung bình các tháng 7, 8, 9, 10 đối với gió vĩ hướng ở mực 850mb (U850) và mực 200mb (U200), xoáy tương đối mực 850mb (VOR) và độ đứt gió giữa 2 mực 200mb và 850mb VWS (U200-U850). Số liệu dự báo lại cho giai đoạn 27 năm quá khứ từ 1982-2008 và dự báo thời gian thực giai đoạn 2002-2008 từ sự kết hợp của 5 mô hình khác nhau với thời gian dự báo 6 tháng. Nghiên cứu sử dụng tổ hợp với trọng số của 5 mô hình như nhau. Sử dụng dự báo lại quá khứ và dự báo thời gian thực của mô hình kết hợp đối với mùa bão (các tháng 7, 8, 9 và 10) các hạn dự báo 0 tháng, 1 tháng và 2 tháng. Phương pháp trị tuyệt đối nhỏ nhất LAD (least absolute deviation) được lựa chọn để xây dựng phương trình dự báo, với yếu tố dự báo gồm: tổng số bão (TY - typhoon) và tổng số bão mạnh (ITY - intense typhoon) ở khu vực TBTBD ( $10^{\circ}\text{S} - 20^{\circ}\text{N}$ ,  $160^{\circ}\text{E} - 120^{\circ}\text{W}$ ) trong mùa bão (các tháng 7, 8, 9, 10) gồm: 1) SST nhiệt độ bề mặt biển khu vực giới hạn bờ ( $5^{\circ}\text{S} - 15^{\circ}\text{N}$ ,  $160^{\circ}\text{E} - 140^{\circ}\text{W}$ ); 2) U850 tốc độ gió vĩ hướng mực 850mb khu vực ( $10^{\circ}\text{S} - 20^{\circ}\text{N}$ ,  $130^{\circ}\text{E} - 150^{\circ}\text{W}$ ); 3) U200 tốc

độ gió vĩ hướng mực 200mb khu vực ( $10^{\circ}\text{S} - 20^{\circ}\text{N}$ ,  $130^{\circ}\text{E} - 150^{\circ}\text{W}$ ); 4) VOR độ đứt gió khu vực ( $5^{\circ}\text{N} - 25^{\circ}\text{N}$ ,  $130^{\circ}\text{E} - 180^{\circ}$ ); 5) VWS độ đứt gió giữa hai mực 200mb và 850mb khu vực ( $10^{\circ}\text{S} - 20^{\circ}\text{N}$ ,  $130^{\circ}\text{E} - 150^{\circ}\text{W}$ ). Sử dụng hồi quy từng bước để lựa chọn nhân tố, đối với yếu tố dự báo là TY tác giả lựa chọn ra các nhân tố gồm: U200 đối với hạn dự báo 1 tháng, U200, U850 đối với hạn dự báo 2 tháng, U200 và VOR đối với hạn dự báo 3 tháng. Phương pháp thực hiện cũng được thực hiện tương tự với yếu tố ITY. Kết quả nghiên cứu cho thấy, các phương trình dự báo đã nắm bắt được biến động chung hoạt động của bão trong giai đoạn nghiên cứu, đặc biệt dự báo nắm bắt được các biến động lớn trước và sau năm 1990. Bên cạnh đó, kết quả dự báo còn nắm bắt được 3 năm bão ít hoạt động 1983, 1998, 1999 (La Niña). So sánh kết quả dự báo của APCC với dự báo thống kê TSR trong 7 năm từ 2002-2008 các tác giả đã đưa ra kết luận: khả năng dự báo của APCC tốt hơn so với dự báo thống kê của TSR.

- **Các hệ thống dự báo bão hạn mùa trên thế giới:** Tổng kết công tác dự báo bão hạn mùa trong vài thập kỷ gần đây Klotzbach và các cộng sự (2019) đã đưa ra danh sách gồm 12 tổ chức/đơn vị trên thế giới đã thực hiện và phát hành các dự báo mùa XTNĐ như trình bày tại Bảng 1.2 [53].

Tổng kết tại Bảng 1.2 cho thấy, đối với các nhóm sử dụng mô hình thống kê nhân tố dự báo thường là SST và các đặc trưng trung bình của các trường khí quyển hoặc các chỉ số khí hậu. Sản phẩm dự báo mùa XTNĐ đối với cả phương pháp thống kê và phương pháp động lực thường là số lượng XTNĐ, số ngày hoạt động và một số đặc trưng khác như năng lượng tích lũy. Ngoài ra, khi sử dụng các mô hình động lực thì sản phẩm có thể là vùng hoạt động của XTNĐ. Cũng theo tổng hợp của Klotzbach và các cộng sự (2019) về các nghiên cứu dự báo mùa đối với XTNĐ theo phương pháp thống kê và thống kê động lực cho thấy, nhân tố dự báo phụ thuộc khá lớn vào đối tượng dự báo, vào vùng biển hoặc không gian xuất hiện/ảnh hưởng của XTNĐ. Cũng theo nhóm tác giả này, đôi khi nhân tố dự báo được quyết định tùy theo mức độ hiểu biết về mối quan hệ với đặc điểm XTNĐ hoặc thậm chí quyết định bởi tính chủ quan/kinh nghiệm của các nhà dự báo.

Bảng 1.2 Một số tổ chức phát hành bản tin dự báo mùa XTNĐ trên thế giới

<b>Đơn vị/ Tổ chức</b>	<b>Vùng biển</b>	<b>Loại dự báo</b>	<b>Nhân tố dự báo; Mô hình động lực</b>	<b>Sản phẩm dự báo</b>
Trường ĐH Colorado, Mỹ	Bắc Đại Tây Dương	Thống kê; Thống kê động lực	SST khu vực Bắc và Nam Đại Tây Dương; khí áp Nam Thái Bình Dương; ENSO; Mode kinh hướng Đại Tây Dương	Số lượng bão trên toàn khu vực; xác suất đổ bộ vào lục địa Hoa Kỳ và Caribe
NOAA, Mỹ	Bắc Đại Tây Dương, Đông Bắc Thái Bình Dương	Thống kê; Thống kê động lực; Động lực	ENSO; mode đa thập kỷ nhiệt đới; SST Đại Tây Dương, dao động đa thập kỷ PDO	Số lượng bão trên toàn khu vực
Trung tâm Cảnh báo bão (TSR), Australia	Đại Tây Dương, TBTBD, Australia	Thống kê	Tín phong; MDR SST; ENSO; khí áp mực biển vùng trung tâm Bắc Thái Bình Dương	Số lượng bão theo các cấp; năng lượng tích lũy xoáy bão; năng lượng tích lũy của những cơn bão đổ bộ; số lượng cơn bão đổ bộ theo các cấp cường độ
Cơ quan Khí tượng Anh (MetOffice)	Bắc Đại Tây Dương	Động lực	1. Mô hình động lực kết hợp 2. Mô hình dò tìm và xác định quỹ đạo bão	Số lượng bão

<b>Đơn vị/ Tổ chức</b>	<b>Vùng biển</b>	<b>Loại dự báo</b>	<b>Nhân tố dự báo; Mô hình động lực</b>	<b>Sản phẩm dự báo</b>
Trường ĐH Macquarie, Australia	Australia/ Đông Nam Thái Bình Dương	Thống kê	Chỉ số SOI, Gradient nhiệt độ tiềm năng	Số lượng bão toàn vùng và số lượng bão hoạt động trên vùng biển Coral
Trường ĐH Hong Kong, Trung Quốc	TBTBD	Thống kê	ENSO; áp cao cận nhiệt đới Bắc Thái Bình Dương mở rộng; cường độ rãnh India-Burma	Số lượng bão theo các cấp cường độ
Trung tâm Dự báo hạn vừa Châu Âu (ECMWF)	ĐTD, Australia, Đông Bắc TBD, Bắc Ấn Độ Dương, Nam Ấn Độ Dương, Nam TBD, TBTBD	Động lực	1. Mô hình động lực kết hợp 2. Mô hình dò tìm và xác định quỹ đạo bão	Số lượng bão, vùng bão hoạt động; khu vực bão hình thành trung bình
Cơ quan Khí tượng Cuba (INSMET)	Đại Tây Dương	Thống kê	Gió vùng Bắc Đại Tây Dương; ENSO; Cường độ của sóng áp cao cận nhiệt đới Bắc Đại Tây Dương ; SST Bắc Đại Tây Dương, QBO	Số lượng bão theo các cấp cường độ; số lượng bão được định danh trên Đại Tây Dương, Caribber và đảo Gulf; Ngày đầu tiên (cuối cùng) bão hoạt động; Số lượng bão hình thành tại vùng Đại Tây Dương và ảnh hưởng tới vùng Caribbean

Ở Việt Nam hiện nay, vấn đề dự báo XTNĐ hầu như chỉ mới được quan tâm nhiều ở bài toán dự báo thời tiết trong khi bài toán dự báo hạn mùa XTNĐ vẫn chưa được nghiên cứu nhiều. Có thể điểm qua một vài công trình đáng chú ý nhất về dự báo hạn mùa XTNĐ bằng phương pháp thống kê như nghiên cứu của tác giả Nguyễn Văn Tuyên (2008) về khả năng dự báo mùa hoạt động của XTNĐ trên khu vực TBTBD và Biển Đông [17-18]. Trên cơ sở công cụ thống kê với bộ nhân tố dự báo là các chỉ số khí hậu, tác giả Nguyễn Văn Tuyên đã khảo sát khả năng dự báo hạn mùa số lượng XTNĐ hoạt động trên khu vực Biển Đông. Ở công trình này, tác giả đã khảo sát 36 chỉ số thuộc 4 nhóm chỉ số khí hậu gồm: 1) nhóm nhân tố ENSO; 2) nhóm chỉ số gió mùa; 3) nhóm chỉ số dao động và 4) nhóm các chỉ số khác làm nhân tố dự báo và yếu tố dự báo được xác định theo 16 đặc trưng của XTNĐ hoạt động trên khu vực TBTBD và Biển Đông. Bằng việc sử dụng phép biến đổi với tên gọi “hiệu số 1” và tính toán hệ số tương quan trễ, công trình đã khảo sát và cho thấy khả năng lựa chọn được những nhân tố dự báo cho mô hình thống kê. Trên cơ sở những nhân tố sơ tuyển nêu trên, tác giả đã sử dụng phương pháp hồi quy từng bước cùng thuật toán jackknife để thiết lập các sơ đồ dự báo hoạt động mùa bão. Qua các tính toán cho thấy vai trò của các nhóm chỉ số khí hậu, đặc biệt là các nhóm thuộc chỉ số ENSO, gió mùa, dao động và nhóm các chỉ số quan hệ từ xa đều có mặt trong các sơ đồ dự báo. Nghiên cứu này đã thiết lập được 32 sơ đồ dự báo trong đó 26 sơ đồ có hệ số tương quan giữa giá trị dự báo và số liệu quan trắc dao động trong khoảng 0,70 đến 0,83; 2 sơ đồ có hệ số tương quan xấp xỉ 0,7 và chỉ có 4 sơ đồ cho kết quả hệ số tương quan đạt mức 0,56 đến 0,6. Các phương trình dự báo được thẩm định, kiểm nghiệm và đánh giá độ tin cậy qua dự báo hindcast trên số liệu phụ thuộc, kiểm tra chéo trên số liệu độc lập jackknife và dự báo thử nghiệm. Các kết quả dự báo cuối cùng đã chỉ ra khả năng tổ hợp dự báo và những sơ đồ dự báo có triển vọng áp dụng vào thực tế dự báo nghiệp vụ cũng như các nghiên cứu tiếp sau. Có thể nói đây là một trong những công trình đáng chú ý nhất về dự báo mùa số lượng XTNĐ bằng phương pháp thống kê.

Việc nghiên cứu dự báo hạn mùa số lượng XTNĐ cũng đã được nhóm tác giả Phan Văn Tân và cộng sự (2010) đề cập ở chừng mực nhất định khi nghiên cứu dự báo số lượng XTNĐ trên khu vực Biển Đông [11]. Ở đây các tác giả đã đưa ra được mối quan hệ giữa sự tăng nhiệt độ mặt nước biển trung bình với hoạt động của XTNĐ được đặc trưng bởi chỉ số NetTC và số lượng XTNĐ trên Biển Đông.

Dự báo mùa cũng đã được Viện Khoa học Khí tượng Thủy văn và Biến đổi khí hậu đặc biệt quan tâm, nghiên cứu từ sớm, có thể kể ra một vài công trình liên quan tới dự báo mùa như “Nghiên cứu thử nghiệm dự báo khí hậu ở Việt Nam” của Nguyễn Duy Chinh (đề tài NCKH cấp bộ, 2000); “Xây dựng phương pháp luận để đánh giá dao động và biến đổi khí hậu” của Nguyễn Văn Thắng (Chương trình mục tiêu Quốc gia về Biến đổi khí hậu, 2010); “Nghiên cứu xây dựng hệ thống dự báo, cảnh báo hạn hán cho Việt Nam với thời hạn đến 3 tháng” của Nguyễn Văn Thắng (Đề tài KHCN KC.08.17/11.15, 2014) hay “Nghiên cứu xây dựng hệ thống nghiệp vụ dự báo khí hậu hạn mùa cho Việt Nam bằng các mô hình động lực” của Mai Văn Khiêm (Đề tài KHCN KC.08.01/16-20, 2019), ... Bắt đầu từ năm 2015, sản phẩm dự báo hạn 3 tháng đối với hiện tượng ENSO, nhiệt độ, lượng mưa, XTNĐ và không khí lạnh đã được tổ chức này công bố trên website <http://imh.ac.vn/nghiệp-vu/>. Ở đây, sản phẩm dự báo khí hậu 3 tháng, được xây dựng dựa trên kết quả tổng hợp thông tin từ Viện Nghiên cứu Quốc tế về Xã hội và Khí hậu (IRI), Trung tâm Dự báo Khí hậu Hoa Kỳ (CPC), Cục Khí tượng Úc (BOM), Trung tâm Dự báo Thời tiết Hạn vừa châu Âu (ECMWF) và kết quả dự báo bằng mô hình thống kê của Viện. Đối với sản phẩm XTNĐ, kết quả dự báo đưa ra là số lượng XTNĐ trên Biển Đông và số lượng XTNĐ ảnh hưởng đến Việt Nam dựa trên số liệu trung bình nhiều năm giai đoạn 1971-2000.

#### **Nhận xét chung:**

Tóm lại, có thể thấy rằng các nghiên cứu về đặc điểm khí hậu (như số lượng, tần suất, quỹ đạo và vùng hoạt động) của XTNĐ khu vực TBTBD nói chung và Biển Đông nói riêng đã được thực hiện bởi nhiều nhà khoa học khí tượng trong và ngoài nước. Riêng đối với khu vực Biển Đông, các kết quả nhận được về đặc điểm hoạt động của XTNĐ ở đây rất đáng ghi nhận song cũng có thể nhận thấy rằng các

nghiên cứu về đặc điểm của XTNĐ qua số lượng và số ngày hoạt động, cũng như mối quan hệ của chúng với hiện tượng ENSO và các chỉ số khí hậu để từ đó khai thác, ứng dụng vào các bài toán dự báo XTNĐ trên khu vực Biển Đông vẫn còn ít được đề cập tới.

Bên cạnh đó, tổng kết các nghiên cứu về đặc điểm, các phương pháp dự báo XTNĐ trên các vùng biển khác nhau cho thấy đối với bài toán dự báo hạn mùa XTNĐ có thể được thực hiện theo phương pháp thống kê và động lực. Ở phương pháp thống kê, bài toán này có thể tiếp cận theo hướng đó là: 1) thống kê truyền thống dựa trên các nhân tố dự báo là các chỉ số hoàn lưu khí quyển, đại dương nói chung hay các chỉ số khí hậu nói riêng và 2) thống kê trên các trường quy mô lớn được lấy từ sản phẩm của các mô hình động lực. Riêng đối với khu vực Biển Đông, vấn đề nghiên cứu các đặc điểm về hoạt động của XTNĐ qua số lượng và số ngày hoạt động, cũng như mối quan hệ của chúng với hiện tượng ENSO, các chỉ số khí hậu và các trường quy mô lớn từ sản phẩm đầu ra của mô hình khí hậu toàn cầu, mô hình khí hậu khu vực để từ đó khai thác, ứng dụng vào các bài toán dự báo hạn mùa XTNĐ trên khu vực Biển Đông vẫn còn khiêm tốn, chưa tương ứng với mức độ nguy hiểm của hiện tượng thời tiết cực đoạn này.

Những tổng quan trên là cơ sở, định hướng để luận án thực hiện khảo sát, xác định số bộ liệu về XTNĐ trên Biển Đông, các chỉ số khí hậu, số liệu sản phẩm mô hình động lực cùng các phương pháp để hoàn thành mục tiêu, nội dung luận án đề ra tại Chương 2.

## CHƯƠNG II. SỐ LIỆU VÀ PHƯƠNG PHÁP NGHIÊN CỨU

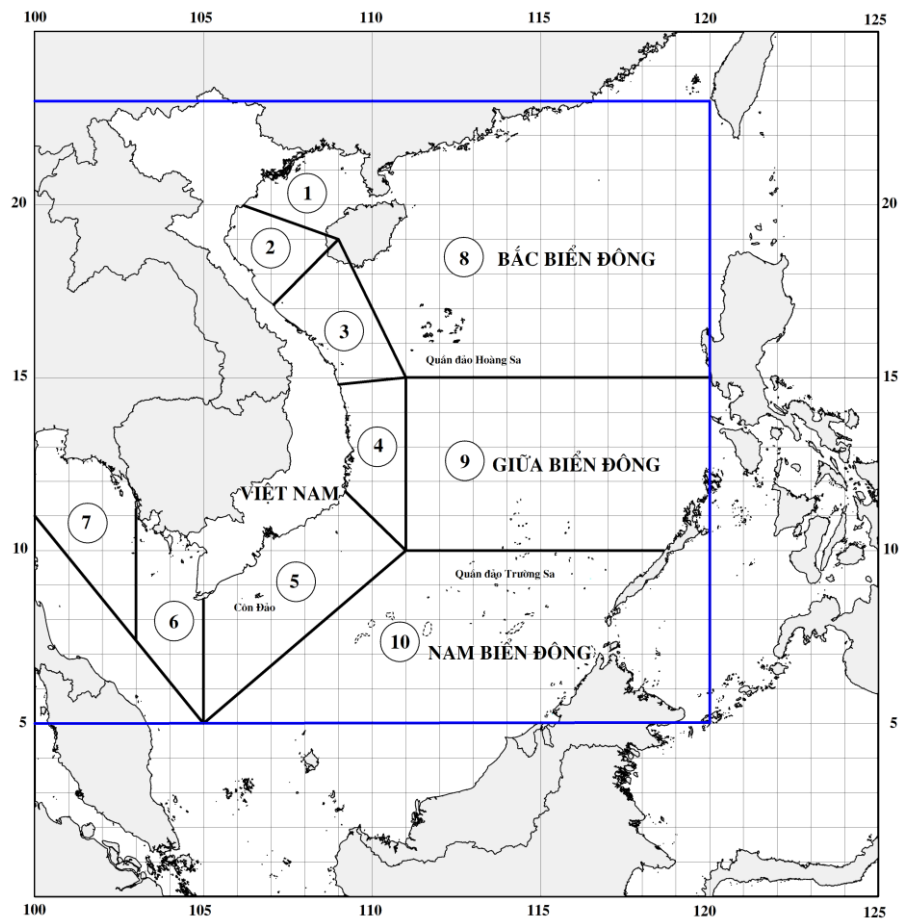
### 2.1 Số liệu và khu vực nghiên cứu

Luận án sử dụng các nguồn số liệu sau: 1) Số lượng XTNĐ hàng tháng trên khu vực Biển Đông, 2) Các chỉ số khí hậu và 3) Các trường trung bình tháng của sản phẩm dự báo lại từ mô hình CFSv2. Về số lượng XTNĐ, luận án chỉ giới hạn xem xét các trường hợp XTNĐ có cường độ gió mạnh nhất đạt từ cấp 6 trở lên, xuất hiện trên khu vực Biển Đông được giới hạn bởi vùng không gian từ  $100^{\circ}$  đến  $120^{\circ}$  kinh Đông và  $5^{\circ}$  đến  $23^{\circ}$  vĩ Bắc như được chỉ ra trên Hình 2.1 (khái niệm áp thấp nhiệt đới, bão; khu vực Biển Đông tại Phụ lục 1 theo Quyết định 03/2020/QĐ-TTg ngày 13/01/2020 quy định về dự báo, cảnh báo và truyền tin thiên tai). Chi tiết về các bộ số liệu này được mô tả dưới đây.

#### 2.1.1 Số liệu XTNĐ

Hiện nay nguồn số liệu về XTNĐ có thể tìm thấy được ở nhiều trung tâm khác nhau, thậm chí các quốc gia có bão hoạt động cũng có thể có cơ sở dữ liệu riêng. Ở Việt Nam cũng đã có một số bộ số liệu về bão Biển Đông và ảnh hưởng đến Việt Nam do các nhà khoa học từ Viện KTTV và BĐKH cũng như Tổng cục KTTV xây dựng cho mục đích nghiên cứu và dự báo nghiệp vụ. Có thể kể đến như công trình “Xây dựng cơ sở dữ liệu bão, áp thấp nhiệt đới trên Biển Đông và ảnh hưởng tới Việt Nam” của Dương Liên Châu và cộng sự (2005), trong đó đã sử dụng ba nguồn dữ liệu từ Trung tâm lưu trữ của Cơ quan Khí tượng Nhật bản (JMA), Trung tâm Liên hiệp Cảnh báo bão của Hải quân Mỹ (JTWC) và có tại Trung tâm Dự báo khí tượng thủy văn quốc gia [1]. Thêm vào đó, công trình này còn tham khảo cơ sở dữ liệu bão Thượng Hải (Trung Quốc) và các số liệu được công bố trong các niên san của Ủy ban Bão vùng TBTBD. Nghiên cứu, khảo sát các bộ dữ liệu cho thấy, trước năm 1986 số liệu về XTNĐ của Việt Nam được lưu trữ dưới nhiều dạng khác nhau và có phần không đầy đủ, do vậy để có được bộ dữ liệu cho giai đoạn 1961-2000 nhóm tác giả đã tiến hành nghiên cứu, thực hiện chỉnh lý theo số liệu của JMA và JTWC trong đó chủ yếu dựa trên JMA. Ở một công trình khác, nghiên cứu của Song và cộng sự (2010) [77] đã chỉ ra rằng giữa các bộ số liệu khác nhau nói chung

không có sự khác biệt lớn về số lượng và quỹ đạo nhưng lại có sự khác nhau đáng kể về cường độ của các XTNĐ. Và điều đó sẽ dẫn đến sự khác nhau về số lượng XTNĐ khi phân chia theo cường độ. Một trong những nguyên nhân gây nên sự khác nhau về cường độ XTNĐ giữa nguồn số liệu có thể liên quan đến phương pháp quan trắc và sử dụng tốc độ gió. Một trong những nguồn số liệu tổng hợp được nhắc đến nhiều trong các nghiên cứu gần đây về bão là IBTrACS (International Best Track Archive for Climate Stewardship). Đây là tập số liệu được tổng hợp từ 14 nguồn dữ liệu của các trung tâm thuộc nhiều quốc gia trên thế giới, bao gồm cả các Trung tâm Khí tượng chuyên vùng (Regional Specialized Meteorological Center - RSMC) [4, 104].



Hình 2.1 Khu vực nghiên cứu (khu vực theo dõi và dự báo áp thấp nhiệt đới, bão tại Biển Đông theo Phụ lục 1, Quyết định 03/2020/QĐ-TTg ngày 13/01/2020)

Trong số các bộ số liệu XTNĐ, liên quan đến vùng Biển Đông có ba nguồn được xem là phù hợp, gồm JMA (Nhật Bản), CMA (Trung Quốc) và JTWC (Mỹ), như được mô tả tại Bảng 2.1 [105-107]. Trong ba nguồn số liệu này, số liệu JMA hay Trung tâm Khí tượng chuyên ngành khu vực Nhật Bản (RSMC Tokyo) được cập nhật thường xuyên hơn và là cơ quan được Tổ chức Khí tượng thế giới phân công theo dõi, phân tích và dự báo XTNĐ cho khu vực giới hạn bởi phạm vi không gian ( $0^{\circ} - 60^{\circ}\text{N}$ ,  $100^{\circ} - 180^{\circ}\text{E}$ ) trong đó bao gồm cả Biển Đông. Bởi vậy, luận án chỉ sử dụng bộ số liệu XTNĐ của trung tâm RSMC Tokyo (từ đây viết tắt là RSMC), trong đó bao gồm thông tin từng 6 giờ một về tọa độ tâm, cường độ theo cấp, tốc độ gió cực đại và khí áp cực tiểu của tất cả các XTNĐ có cường độ từ áp thấp nhiệt đới (ATNĐ) trở lên. Chuỗi số liệu RSMC sẵn có từ 1951-2019 cho toàn khu vực TBTBD, tuy nhiên trong khuôn khổ của luận án, chỉ những cơn bão có ít nhất một lần có tâm xuất hiện trong khu vực Biển Đông mới được tính đến, bất kể chúng được hình thành ở đâu.

Bảng 2.1 Ba nguồn số liệu về XTNĐ

TT	Nguồn (quốc gia)	Yếu tố					Thời gian có số liệu
		Thời gian	Tọa độ	Cấp	Vmax	Pmin	
1	JMA [105] (Nhật Bản)	X	X	X	X <sup>(*)</sup>	X <sup>(**)</sup>	(*) 1977 - 2021 (**) 1951 - 2021
2	CMA [107] (Trung Quốc)	X	X	X	X <sup>(*)</sup>	X <sup>(**)</sup>	(*) 1949 - 2017 (**) 1949 - 2021
3	JTWC [106] (Mỹ)	X	X	X	X <sup>(*)</sup>	X <sup>(**)</sup>	(*) 1945 - 2018 (**) 2001 - 2018

Chú thích: dấu X ký hiệu là có số liệu

### 2.1.2 Số liệu về các chỉ số khí hậu

Chỉ số khí hậu là các đại lượng được tính toán từ các trường khí quyển, đại dương dùng để mô tả trạng thái của hệ thống khí hậu. Các chỉ số khí hậu thường được biểu diễn bằng chuỗi giá trị trung bình tháng. Một số chỉ số là các hàm trực

giao thực nghiệm (EOF) biến thiên theo thời gian của các trường khí quyển, đại dương từ phân tích thành phần chính các trường, như chỉ số dao động Bắc Đại Tây Dương (NAO), chỉ số dao động thập kỷ Thái Bình Dương (PDO),... Một số chỉ số là chuỗi thời gian của trung bình hoặc dị thường hoặc chuẩn hoá trên một vùng không gian nào đó, như dị thường nhiệt độ mặt nước biển (SSTA) các vùng Nino, chỉ số dao động nam (SOI),... Nói chung các chỉ số khí hậu thường được cung cấp miễn phí bởi một số trung tâm trên thế giới như Phòng Thí nghiệm Khoa học Vật lý (<https://psl.noaa.gov/data/climateindices/list/>) và Trung tâm Dự báo Môi trường Quốc gia Hoa Kỳ (NCEP) (<https://www.cpc.ncep.noaa.gov/data/indices/>) của Tổng cục Khí quyển Đại dương Hoa Kỳ (National Oceanic and Atmospheric Administration - NOAA).

Nhiều nghiên cứu đã chỉ ra rằng, các chỉ số khí hậu được xem là thành phần quan trọng trong bài toán dự báo hạn mùa [17-18, 26-27, 53], đặc biệt là dự báo XTNĐ, theo nghĩa chúng có thể được lựa chọn làm nhân tố dự báo cho các mô hình thống kê. Do đó luận án đã tiến hành nghiên cứu sử dụng các chỉ số này. Tuy nhiên, do lịch sử để lại, các chỉ số khác nhau là kết quả nghiên cứu của các tác giả khác nhau nên độ dài của những số liệu này là không giống nhau. Nhằm hướng tới bài toán dự báo nghiệp vụ, luận án chỉ lựa chọn các chỉ số theo một số tiêu chí sau:

1) Các chỉ số được cho là có thể có ảnh hưởng đến hoạt động của XTNĐ trên khu vực Biển Đông và TBTBD (rút ra từ những nghiên cứu trước đây);

2) Chuỗi số liệu đủ dài. Đa số các chỉ số đều có độ dài chuỗi thời gian từ 1948 đến nay, thậm chí một số chỉ số có độ dài dài hơn (tức trước đó), nhưng cũng có chỉ số chỉ có số liệu từ năm 1979 (trên 30 năm). Như vậy các chỉ số được lựa chọn phải có chuỗi thời gian ít nhất từ năm 1979;

3) Có tính cập nhật thường xuyên. Trong số các chỉ số khí hậu sẵn có, nhiều chỉ số chỉ cập nhật sau một khoảng thời gian dài ví dụ nửa năm hoặc một năm, nhưng cũng có chỉ số được cập nhật tương đối thường xuyên, chỉ sau khoảng 1 đến

2 tháng. Với mục đích dự báo mùa, các chỉ số cập nhật sau khoảng thời gian 6 tháng tới 1 năm là chấp nhận được.

Trên cơ sở đó luận án đã lựa chọn 30 chỉ số khí hậu làm nhân tố dự tuyến cho bài toán dự báo hạn mùa XTNĐ trên khu vực Biển Đông. Danh sách các chỉ số khí hậu được mô tả trong Bảng 2.2

Bảng 2.2 Bộ chỉ số khí hậu 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 (0-10S, 90W-80W) <a href="https://www.psl.noaa.gov/data/correlation/NINO1.data">https://www.psl.noaa.gov/data/correlation/NINO1.data</a>	ENSO
2	Niño3	SST vùng nhiệt đới Đông Thái Bình Dương (5N-5S, 150W-90W) <a href="https://www.psl.noaa.gov/data/correlation/NINO3.data">https://www.psl.noaa.gov/data/correlation/NINO3.data</a>	ENSO
3	Niño3.4	Chỉ số SST vùng Niño 3.4 (5N-5S, 170E-120W) <a href="https://www.psl.noaa.gov/data/correlation/NINO34.data">https://www.psl.noaa.gov/data/correlation/NINO34.data</a>	ENSO
4	Niño4	Chỉ số SST vùng Niño 4 (5N-5S, 160E-150W) <a href="https://www.psl.noaa.gov/data/correlation/NINO4.data">https://www.psl.noaa.gov/data/correlation/NINO4.data</a>	ENSO
5	MEIV2	Chỉ số ENSO đa biến phiên bản 2 <a href="https://psl.noaa.gov/data/correlation/meiv2.data">https://psl.noaa.gov/data/correlation/meiv2.data</a>	ENSO
6	CENSO	Chỉ số ENSO kết hợp <a href="https://psl.noaa.gov/data/correlation/censo.data">https://psl.noaa.gov/data/correlation/censo.data</a>	ENSO
7	ONI	Chỉ số Niño đại dương <a href="https://www.psl.noaa.gov/data/correlation/oni.data">https://www.psl.noaa.gov/data/correlation/oni.data</a>	Đại Dương
8	TNI	Chỉ số khuynh hướng Niño hay chỉ số Niño-chuyển dịch (Trans-Niño Index) <a href="https://psl.noaa.gov/data/correlation/tni.data">https://psl.noaa.gov/data/correlation/tni.data</a>	Đại Dương
9	WHWP	Bể nóng bán cầu tây <a href="https://www.psl.noaa.gov/data/correlation/whwp.data">https://www.psl.noaa.gov/data/correlation/whwp.data</a>	Đại Dương

TT	Ký hiệu chỉ số	Tên gọi và địa chỉ lấy số liệu	Nhóm
10	TNA	Chỉ số Bắc Đại Tây Dương nhiệt đới (Dị thường SST trung bình tháng vùng 5.5 - 23.5N và 15 - 57.5W) <a href="https://www.psl.noaa.gov/data/correlation/tna.data">https://www.psl.noaa.gov/data/correlation/tna.data</a>	Đại Dương
11	TSA	Chỉ số Nam Đại Tây Dương nhiệt đới (SSTA trung bình tháng vùng 0 - 20S, 10E - 30W) <a href="https://www.psl.noaa.gov/data/correlation/tsa.data">https://www.psl.noaa.gov/data/correlation/tsa.data</a>	Đại Dương
12	AMONUS	Dao động đa thập kỷ Đại Tây Dương <a href="https://psl.noaa.gov/data/correlation/amon.us.data">https://psl.noaa.gov/data/correlation/amon.us.data</a>	Đại Dương
13	AMM	Mode kinh hướng Đại Tây Dương <a href="https://www.psl.noaa.gov/data/timeseries/monthly/AMM/ammsst.data">https://www.psl.noaa.gov/data/timeseries/monthly/AMM/ammsst.data</a>	Đại Dương
14	NTA	Chỉ số SST nhiệt đới Bắc Đại Tây Dương (SSTA trung bình 2 vùng: 6N-18N, 60W-20W và 6N-10, 20W-10W). <a href="https://www.psl.noaa.gov/data/correlation/NTA_ersst.data">https://www.psl.noaa.gov/data/correlation/NTA_ersst.data</a>	Đại Dương
15	CAR	SST vùng Caribe <a href="https://www.psl.noaa.gov/data/correlation/CAR_ersst.data">https://www.psl.noaa.gov/data/correlation/CAR_ersst.data</a>	Đại Dương
16	REPAC	Dị thường chuẩn hóa khí áp mực biển vùng xích đạo Đông Thái Bình Dương <a href="https://www.cpc.ncep.noaa.gov/data/indices/REPAC.for">https://www.cpc.ncep.noaa.gov/data/indices/REPAC.for</a>	Khí quyển
17	RINDO	Dị thường chuẩn hóa khí áp mực biển vùng biển Indonesia <a href="https://www.cpc.ncep.noaa.gov/data/indices/RINDO.for">https://www.cpc.ncep.noaa.gov/data/indices/RINDO.for</a>	Khí quyển
18	QBO30	Chỉ số dao động tựa 2 năm (theo gió vĩ hướng mực 30mb) <a href="https://www.psl.noaa.gov/data/correlation/qbo.data">https://www.psl.noaa.gov/data/correlation/qbo.data</a>	Khí quyển
19	QBO50	Chỉ số dao động tựa 2 năm (theo gió vĩ hướng mực 50mb) <a href="https://www.cpc.ncep.noaa.gov/data/indices/qbo.u50.index">https://www.cpc.ncep.noaa.gov/data/indices/qbo.u50.index</a>	Khí quyển
20	SOI	Chỉ số dao động nam <a href="https://www.psl.noaa.gov/data/correlation/soi.data">https://www.psl.noaa.gov/data/correlation/soi.data</a>	Khí quyển

TT	Ký hiệu chỉ số	Tên gọi và địa chỉ lấy số liệu	Nhóm
21	REQSOI	Chỉ số dao động nam vùng xích đạo (Equatorial SOI) <a href="https://www.cpc.ncep.noaa.gov/data/indices/reqsoi.for">https://www.cpc.ncep.noaa.gov/data/indices/reqsoi.for</a>	Khí quyển
22	ZWND20	Gió vĩ hướng vùng xích đạo mực 200mb (165W-110W) <a href="https://www.cpc.ncep.noaa.gov/data/indices/zwnd200">https://www.cpc.ncep.noaa.gov/data/indices/zwnd200</a>	Khí quyển
23	WPAC85	Tín phong mực 850 mb (135E-180W; 5N-5S) <a href="https://www.cpc.ncep.noaa.gov/data/indices/wpac850">https://www.cpc.ncep.noaa.gov/data/indices/wpac850</a>	Khí quyển
24	EPAC85	Tín phong mực 850 mb (135W-120W; 5N - 5S) <a href="https://www.cpc.ncep.noaa.gov/data/indices/epac850">https://www.cpc.ncep.noaa.gov/data/indices/epac850</a>	Khí quyển
25	CPAC85	Tín phong mực 850 mb (175W - 140W; 5N – 5S) <a href="https://www.cpc.ncep.noaa.gov/data/indices/cpac850">https://www.cpc.ncep.noaa.gov/data/indices/cpac850</a>	Khí quyển
26	PNA	Chỉ số Thái Bình Dương Bắc Mỹ <a href="https://psl.noaa.gov/data/correlation/pna.data">https://psl.noaa.gov/data/correlation/pna.data</a>	Khí quyển
27	WP	Chỉ số Tây Thái Bình Dương <a href="https://www.psl.noaa.gov/data/correlation/wp.data">https://www.psl.noaa.gov/data/correlation/wp.data</a>	Khí quyển
28	NAO	Chỉ số dao động Bắc Đại Tây Dương <a href="https://www.psl.noaa.gov/data/correlation/nao.data">https://www.psl.noaa.gov/data/correlation/nao.data</a>	Khí quyển
29	NP	Chỉ số Bắc Thái Bình Dương <a href="https://psl.noaa.gov/data/correlation/np.data">https://psl.noaa.gov/data/correlation/np.data</a>	Khí quyển
30	AO	Dao động Bắc Cực <a href="https://psl.noaa.gov/data/correlation/ao.data">https://psl.noaa.gov/data/correlation/ao.data</a>	Khí quyển

### 2.1.3 Số liệu dự báo lại của mô hình CFS

CFS (Climate Forecast System) là mô hình dự báo khí hậu hạn mùa được phát triển bởi Trung tâm Dự báo Môi trường Hoa Kỳ (National Centers for Environmental Prediction - NCEP). CFS là mô hình động lực kết hợp gần như đầy đủ các thành phần

khí quyển, đại dương, đất liền và các quá trình băng, được triển khai vào nghiệp vụ tại NCEP từ tháng 8 năm 2004. Phiên bản CFS hiện nay là phiên bản thứ hai (CFSv2) đã được chạy dự báo nghiệp vụ từ năm 2011 [72]. So với phiên bản đầu tiên (CFSv1), phiên bản CFSv2 đã có nhiều cải tiến đáng kể. Nhằm mục đích phục vụ nghiên cứu phát triển mô hình cũng như đánh giá khả năng ứng dụng mô hình cho các vùng khác nhau trên thế giới, CFSv2 đã được chạy dự báo lại cho giai đoạn quá khứ với tần suất 5 ngày một lần vào các ngày 1, 6, 11, 16, 21 và 26, trong các ngày đó mô hình sẽ được chạy 4 lần/ngày vào các thời điểm 00, 06, 12, 18 UTC. Số liệu được lưu ở các định dạng khác nhau trong đó có các trường trung bình tháng với hạn dự báo đến 9 tháng. Việc nghiên cứu, sử dụng nguồn số liệu này cho bài toán dự báo mùa và nội mùa với các yếu tố dự báo như nhiệt độ, lượng mưa, ... hay là XTNĐ cũng đã được thực hiện ở nhiều nghiên cứu gần đây [50, 53, 58, 68, 81, 96].

Vì vậy, cho mục đích xây dựng các mô hình thống kê dự báo số lượng XTNĐ trên Biển Đông, luận án sẽ sử dụng các trường trung bình tháng dự báo lại của CFSv2 (CFS\_Rfc) làm nhân tố dự báo. Số liệu CFSv2 được lấy trong giai đoạn 1983-2009 và 2012-2016. Các trường CFSv2 được khai thác bao gồm: nhiệt độ bề mặt biển (SST); khí áp mực biển (PMSL); các thành phần gió kinh vĩ hướng và độ cao địa thế vị trên các mực đẳng áp chuẩn 850, 500, 200 mb; tốc độ gió thẳng đứng mực 850mb và độ ẩm tương đối mực 700, 500mb. Trường độ đứt gió được tính toán trên cơ sở các trường gió kinh vĩ hướng và được xem là số liệu ban đầu.

## **2.2 Phương pháp nghiên cứu**

### **2.2.1 Xử lý số liệu ban đầu**

Về số liệu XTNĐ: Số lượng XTNĐ trên Biển Đông được chọn lọc từ bộ số liệu RSMC theo nguyên tắc các XTNĐ có ít nhất một lần quan trắc được (obs 6h) mà tâm rơi vào miền Biển Đông sẽ được tính là một cơn. Số liệu này được sử dụng để khảo sát đặc điểm hoạt động (Chương 3) và xây dựng phương trình dự báo (Chương 4) số lượng XTNĐ. Nói cách khác, tất cả các trường hợp XTNĐ được đề cập đến trong luận án, dù miền không gian biểu diễn bản đồ có mở rộng về phía đông, đều phải quan trắc được ít nhất một lần (obs) xuất hiện trên Biển Đông. Ngoài

ra, XTNĐ được tính trong luận án là tất cả các trường hợp có cường độ đạt từ ATNĐ trở lên, tức tốc độ gió cực đại phải từ 39km/h trở lên. Trong một số trường hợp cụ thể có thể chia XTNĐ thành ATNĐ (tốc độ gió cực đại nằm trong khoảng 39-61km/h) và bão (tốc độ gió cực đại trên 61km/h).

Số liệu các chỉ số khí hậu: Trong một số trường hợp số liệu gốc của các chỉ số là các giá trị các trường, ví dụ SST, khí áp,... nên thứ nguyên không giống nhau và đặc biệt hơn là bậc đại lượng của chúng cũng rất khác nhau. Do vậy, số liệu của các chỉ số khí hậu sẽ được chuẩn hoá theo nguyên tắc lấy giá trị ban đầu trừ đi trung bình nhiều năm và chia cho độ lệch chuẩn.

Đối với số liệu CFSv2, việc xử lý nói chung khá phức tạp vì chúng còn phụ thuộc vào hạn dự báo. Trên thực tế hạn dự báo của các trường CFSv2 lên đến 9 tháng nhưng luận án chỉ khảo sát đến hạn dự báo 6 tháng. Độ phân giải của CFSv2 là 1.0 x 1.0 độ nên mỗi trường sẽ là một ma trận ba chiều.

Ngoài ra, để thuận tiện cho việc tính toán, tất cả các bộ số liệu sẽ được đưa về định dạng NetCDF.

### 2.2.2 Các mô hình thống kê

Cho mục đích dự báo mùa số lượng XTNĐ trên Biển Đông luận án sử dụng hai nhóm mô hình thống kê là mô hình hồi quy tuyến tính và mô hình mạng thần kinh nhân tạo (ANN). Sau đây sẽ trình bày sơ lược lý thuyết các mô hình này.

- Mô hình hồi quy tuyến tính

Dạng tổng quát của mô hình có thể được biểu diễn như sau:

$$\hat{y} = \beta_0 + \sum_{j=1}^m \beta_j x_j \quad (2.1)$$

Trong đó  $\hat{y}$  là yếu tố dự báo (số lượng XTNĐ),  $x_j, j = 1, 2, \dots, m$  là các nhân tố dự báo,  $\beta_0, \beta_1, \dots, \beta_m$  là các hệ số hồi quy mà ước lượng của chúng tương ứng là  $b_0, b_1, b_2, \dots, b_m$  được xác định theo ba phương pháp khác nhau:

- 1) Phương pháp bình phương tối thiểu (MLR):

$$\sum_{i=1}^n (y_i - b_0 - \sum_{j=1}^m b_j x_{ij})^2 \rightarrow \min \quad (2.2)$$

- 2) Phương pháp độ lệch tuyệt đối nhỏ nhất (LAD):

$$\sum_{i=1}^n |y_i - b_0 - \sum_{j=1}^m b_j x_{ij}| \rightarrow \min \quad (2.3)$$

3) Phương pháp minimax (LMV):

$$\text{minimize}(\text{maximum}|y_i - b_0 - \sum_{j=1}^m b_j x_{ij}|, i = 1, 2, \dots, n) \quad (2.4)$$

Trong các công thức từ (2.2) đến (2.4),  $y_i$  và  $x_{ij}$  tương ứng là giá trị quan trắc thứ  $i$  ( $i=1, 2, \dots, n$ ) của yếu tố dự báo  $y$  và các nhân tố dự báo  $x_j$  ( $j=1, 2, \dots, m$ ). Ứng với mỗi phương pháp xác định bộ hệ số hồi quy ở trên, để giản tiện và thống nhất các nghiên cứu và kết quả dưới đây quy ước tên gọi tương ứng là mô hình MLR, mô hình LAD và mô hình LMV. Như vậy, thực chất sẽ có ba mô hình hồi quy tuyến tính. Các hệ số hồi quy của các mô hình được xác định khi sử dụng tập số liệu phụ thuộc (hay còn gọi là tập luyện). Đối với mô hình MLR các hệ số hồi quy được xác định dựa trên cực tiểu hoá “sai số bậc hai” (bình phương của độ lệch giữa ước lượng và quan trắc). Còn các hệ số của mô hình LAD lại được tìm dựa trên cực tiểu hoá “sai số bậc nhất” (trị tuyệt đối của hiệu giữa ước lượng và quan trắc). Trong khi đó mô hình LMV được xây dựng dựa trên nguyên tắc tìm bộ hệ số làm cực tiểu hoá các cực đại của “sai số bậc nhất”. Các hệ số của mô hình MLR được tìm bằng phương pháp giải tích, còn các hệ số của các mô hình LAD và LMV được xác định bằng phương pháp lặp. Khi đã xác định được các hệ số hồi quy, ứng với mỗi bộ số liệu ban đầu mới  $x_j$  ( $j=1, 2, \dots, m$ ) của các nhân tố dự báo, sau khi thay vào phương trình (2.1) ta sẽ nhận được kết quả dự báo của  $y$ .

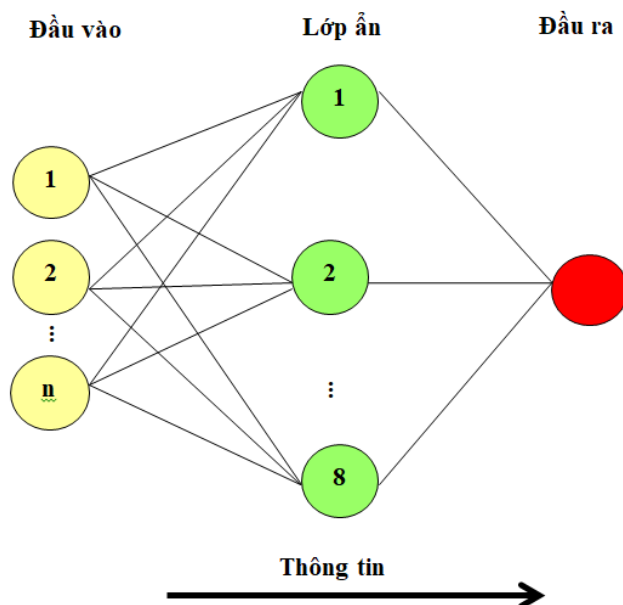
Một số nghiên cứu đã chỉ ra rằng, trong khi mô hình MLR cho kết quả khá ổn định và tốt trong những trường hợp số liệu của biến phụ thuộc ( $y$ ) ít biến động thì các mô hình LAD và LMV có thể thích hợp hơn với trường hợp  $y$  biến động bất thường.

- Mô hình mạng thần kinh nhân tạo (ANN)

Mô hình ANN được nghiên cứu phát triển xuất phát từ những quan sát hoạt động của bộ não và các tế bào thần kinh [47, 97]. ANN là sự mô phỏng cấu trúc và hoạt động của bộ não người, được hình thành từ những đơn vị riêng lẻ gọi là những tế bào thần kinh nhân tạo (hay neuron nhân tạo). Bản chất của phương pháp là tìm

hàm xấp xỉ cho chuỗi số liệu của một hoặc nhiều biến phụ thuộc (như là kết quả phản ứng của hệ thần kinh) dựa trên các biến độc lập làm đầu vào (như là các tín hiệu tác động ban đầu) với sai số cho trước.

Số lớp mạng ANN là rất đa dạng, phân biệt theo cấu hình mạng, phương thức cực tiểu,... Trên hình 2.2 minh họa một ANN với một lớp ẩn và một đầu ra.



Hình 2.2 Sơ đồ mạng ANN với các biến đầu vào, lớp ẩn và lớp đầu ra

Neuron là đơn vị cơ sở của một mạng thần kinh nhân tạo. Một neuron nhận các giá trị đầu vào thực, giá trị đầu ra là giá hàm ngưỡng  $\phi$  với đối số là tổ hợp tuyến tính các giá trị đầu vào. Giả sử các giá trị đầu vào là vectơ  $\vec{x} = \{x_1, \dots, x_n\}$ , các trọng số tương ứng là vectơ  $\vec{w} = \{w_1, \dots, w_n\}$ . Tổ hợp tuyến tính của các giá trị đầu vào là:

$$u = w_1x_1 + \dots + w_nx_n \quad (2.5)$$

Giá trị đầu ra:  $o(\vec{x}) = \varphi(w_0 + u)$ , trong đó  $w_0$  là một giá trị thực thêm vào gọi là độ lệch (bias). Nếu ta đưa thêm ký hiệu  $x_0 = 1$ , ta có  $o(\vec{x}) = \varphi(\sum_{i=0}^n w_ix_i)$

Hàm  $\varphi$ , còn được gọi là hàm kích hoạt (activation), có một số dạng sau đây: Hàm ngưỡng (threshold).

$$\varphi(v) = \begin{cases} 1 & \text{khi } v \geq 0 \\ 0 & \text{khi } v < 0 \end{cases} \quad (2.6)$$

Hàm tuyến tính - tron từng đoạn (piecewise-linear)

$$\varphi(v) = \begin{cases} 1 & \text{khi } v \geq \frac{1}{2} \\ v & \text{khi } -\frac{1}{2} < v < \frac{1}{2} \\ 0 & \text{khi } v \leq -\frac{1}{2} \end{cases} \quad (2.7)$$

Hàm dạng sigma (sigmoid)

$$\varphi(v) = \frac{1}{1 + \exp(-av)} \quad (2.8)$$

với  $a$  là tham số độ dốc.

Sai số  $E$  là tổng các sai số trên tất cả các neuron đầu ra:

$$E(\vec{w}) = \frac{1}{2} \sum_{d \in D} \sum_{k \in \text{outputs}} (t_{kd} - o_{kd})^2 \quad (2.9)$$

ở đây  $D$  là tập số liệu luyện,  $t_{kd}$  là giá trị đích (giá trị mong muốn),  $o_{kd}$  là giá trị đầu ra của ví dụ luyện thứ  $d$ . Mục đích của bài toàn là cực tiểu hóa  $E$  trong không gian  $\vec{w}$ . Quá trình cực tiểu hóa này (còn được hiểu là quá trình học của ANN) được tiến hành theo thuật toán lan truyền ngược. Cụ thể xử lý toán học của quá trình này có thể tìm hiểu trong các tài liệu tham khảo về ANN [47, 97]. Kết quả cuối cùng sẽ là tập các trọng số  $w$ . Trong bài toán dự báo, với một tập số liệu phụ thuộc ta sẽ nhận được một bộ trọng số  $w$  nào đó là kết quả của quá trình “học” của mô hình. Sử dụng bộ trọng số này cho bộ số liệu đầu vào mới ta sẽ nhận được kết quả dự báo.

5) Phương pháp tổ hợp: Với các kết quả dự báo đưa ra từ 3 mô hình hồi quy tuyến tính (MLR, LAD, LMV) và mô hình phi tuyến (ANN) kể trên, luận án xem xét thêm trường hợp tổ hợp trung bình với các trọng số bằng nhau giữa 4 mô hình trên: MLR, LAD, LMV với ANN và được kí hiệu là E1234. Theo đó, E1234 được tính toán theo công thức:  $E1234 = \left( \frac{\text{MLR} + \text{LAD} + \text{LMV} + \text{ANN}}{4} \right)$ .

### 2.2.3 Tuyển chọn nhân tố dự báo

Tuyển chọn nhân tố là một trong những bước quan trọng nhất khi xây dựng các mô hình dự báo thống kê. Thông thường các nhân tố dự báo sẽ được chọn dựa trên một tập khá lớn các nhân tố ban đầu hay còn được gọi là các nhân tố dự tuyển. Về nguyên tắc các nhân tố dự tuyển là những nhân tố tiềm năng có thể được sử dụng làm nhân tố dự báo (là những nhân tố sẽ có mặt trong phương trình dự báo cuối cùng).

Như đã tổng quan trong Chương 1, số lượng XTNĐ trên Biển Đông, yếu tố dự báo mà luận án lựa chọn, phụ thuộc vào rất nhiều quá trình khác nhau, từ quy mô lớn đến quy mô khu vực, và có thể cả các nhân tố địa phương. Các quá trình này có thể có liên quan nhất định tới hoạt động của XTNĐ Biển Đông. Chúng có thể được xem xét như là những nhân tố dự báo tiềm năng. Tuy nhiên, không thể tất cả các nhân tố này đều là nhân tố dự báo, tức là có mặt trong phương trình dự báo cuối cùng. Vấn đề đặt ra là làm thế nào để lựa chọn được các nhân tố tốt nhất có thể trong số các nhân tố dự tuyển để đưa vào phương trình dự báo?

- Phương pháp hồi quy từng bước

Trên thực tế, có nhiều cách để lựa chọn các nhân tố dự báo tùy thuộc vào tính chất của bài toán và tùy thuộc vào từng tình huống cụ thể. Trong trường hợp số nhân tố dự tuyển không nhiều các nhân tố dự báo có thể được lựa chọn trực tiếp bằng phương pháp chủ quan hoặc bán chủ quan dựa vào kinh nghiệm của người nghiên cứu. Nhưng khi số lượng các nhân tố dự tuyển rất lớn, việc lựa chọn trực tiếp là không khả thi, thậm chí mắc sai lầm. Để khắc phục tình trạng đó người ta có thể sử dụng một số phương pháp tuyển chọn khách quan. Trong luận án việc tuyển chọn các nhân tố dự báo từ tập các nhân tố dự tuyển sẽ được thực hiện bằng phương pháp hồi quy từng bước.

Kỹ thuật lọc nhân tố bằng phương pháp hồi quy từng bước cũng có một vài kỹ thuật khác nhau, như dựa trên khái niệm hệ số tương quan riêng hoặc tương quan bội, đưa vào từng nhân tố hoặc loại bỏ từng nhân tố. Chi tiết hơn về kỹ thuật này có thể xem trong [9, 78]. Phương pháp mà luận án chọn là dựa trên hệ số tương quan riêng và đưa vào từng nhân tố. Tóm lược về phương pháp được trình bày như sau.

Ký hiệu  $Y$  là yếu tố dự báo và  $X_1, \dots, X_m$  là các nhân tố dự tuyển. Việc lựa chọn nhân tố được thực hiện theo các bước:

*Bước 1:* Tính các hệ số tương quan toàn phần  $r_{yi}$  giữa  $Y$  và các  $X_i$  ( $i=1..m$ ) và chọn trong chúng hệ số có giá trị tuyệt đối lớn nhất. Giả sử:

$$|r_{y1}| = \max_{1 \leq i \leq m} \{|r_{yi}|\} \quad (2.10)$$

khi đó, biến  $X_1$  được xem là biến có tác động chính lên  $Y$  và ta xây dựng phương trình hồi quy giữa  $Y$  và  $X_1$ :

$$Y^{(1)} = b_0^{(1)} + b_1 * X_1^{(1)} \quad (2.11)$$

Từ phương trình (2.11) ta tính được chuẩn sai thặng dư  $s^{(1)}$  là sai số chuẩn khi sử dụng chỉ một biến  $X_1$ .

*Bước 2:* Tính các hệ số tương quan riêng  $r_{yi.1}$  ( $i=2..m$ ), là hệ số tương quan giữa  $Y$  và các  $X_2, \dots, X_m$  sau khi đã loại bỏ vai trò của  $X_1$ , và chọn hệ số có giá trị tuyệt đối lớn nhất trong chúng. Giả sử:

$$|r_{y2.1}| = \max_{2 \leq i \leq m} \{|r_{yi.1}|\} \quad (2.12)$$

Khi đó ta chọn tiếp biến  $X_2$  (trong số các biến còn lại  $X_2, \dots, X_m$ ) và xây dựng phương trình hồi quy:

$$Y^{(2)} = b_0^{(2)} + b_1^{(2)} * X_1 + b_2^{(2)} * X_2 \quad (2.13)$$

Tương tự như bước 1, từ phương trình (2.13) ta tính được chuẩn sai thặng dư  $s^{(2)}$ . Tại bước 2 này phương trình hồi quy thu được gồm 2 biến với sai số chuẩn là  $s^{(2)}$ . Việc đưa thêm biến  $X_2$  vào phương trình chắc chắn sẽ làm giảm sai số ước lượng đối với  $Y$ . Mức độ suy giảm sai số được xác định bằng cách so sánh các chuẩn sai thặng dư  $s^{(2)}$  với  $s^{(1)}$ . Do đó dẫn đến bước 3.

*Bước 3:* So sánh giá trị chuẩn sai thặng dư  $s^{(2)}$  với  $s^{(1)}$

$$\text{Nếu } \left| \frac{s^{(2)} - s^{(1)}}{s^{(2)}} \right| < \varepsilon \quad (2.14)$$

thì biến  $X_2$  sẽ bị bỏ qua, quá trình tuyển chọn sẽ dừng lại và chỉ có một biến  $X_1$  được lựa chọn. Bởi  $X_2$  là biến có tương quan tốt nhất trong số các biến còn lại nhưng không làm giảm sai số đáng kể thì việc lựa chọn các biến khác là không cần thiết. Tham số Epsilon  $\varepsilon$  là một số dương tùy ý biểu thị mức độ suy giảm sai số khi đưa thêm biến vào, do đó nó có vai trò quyết định liệu biến đang xem xét có được lựa chọn hay không. Nếu  $\varepsilon$  càng nhỏ thì số biến được lựa chọn càng nhiều. Nếu  $\varepsilon = 0$  thì về nguyên tắc tất cả các biến dự tuyển sẽ được chọn.

Nếu (2.14) không thoả mãn, tức  $\left| \frac{s^{(2)} - s^{(1)}}{s^{(2)}} \right| \geq \varepsilon$ , thì biến  $X_2$  được chọn và quy trình được lặp lại từ bước 2 trong đó các hệ số tương quan riêng giữa  $Y$  và các biến còn lại sẽ được tính sau khi đã loại bỏ vai trò của các biến đã được chọn, tức  $r_{y_i.12}$  với  $i=3..m$ . Quá trình cứ tiếp tục như vậy cho đến khi hết tất cả các biến (trường hợp  $\varepsilon = 0$ ) hoặc ở bước thứ  $k$  nào đó (trường hợp  $\varepsilon > 0$ ). Như vậy ở bước thứ  $k$  ta có chuẩn sai thặng dư  $s^{(k)}$  tương ứng với phương trình hồi quy:

$$Y^{(k)} = b_0^{(k)} + b_1^{(k)} * X_1 + \dots + b_k^{(k)} * X_k \quad (2.15)$$

và điều kiện lựa chọn:  $\left| \frac{s^{(k)} - s^{(k-1)}}{s^{(k)}} \right| < \varepsilon \quad (2.16)$

Như đã trình bày trên đây, các nhân tố được chọn theo nguyên tắc xem xét mức độ quan trọng của chúng trong việc làm giảm sai số ước lượng theo thứ tự giảm dần, do đó những nhân tố được chọn trước sẽ quan trọng hơn những nhân tố chọn sau. Do tính chất khác nhau giữa các bộ nhân tố dự tuyển nên việc xác định tham số  $\varepsilon$  cho từng trường hợp cụ thể sẽ được bàn luận sau.

Như đã đề cập trong mục 2.2.2, trong số bốn mô hình thống kê (MLR, LAD, LMV và ANN) chỉ có mô hình MLR là các hệ số có thể tìm được bằng phương pháp giải tích, các mô hình còn lại đều bằng phương pháp lặp. Bài toán lọc nhân tố đối với các mô hình LAD, LMV và ANN rất phức tạp, nằm ngoài khuôn khổ luận án. Do đó trong luận án các nhân tố được lựa chọn bằng phương pháp hồi quy từng bước trên đây cũng được xem là các nhân tố tốt nhất cho các mô hình LAD, LMV và ANN. Nói cách khác, bộ nhân tố dự báo được lựa chọn từ tập các nhân tố dự tuyển sẽ được sử dụng cho cả bốn mô hình MLR, LAD, LMV và ANN.

- Phương pháp phân tích thành phần chính (PCA)

Bên cạnh phương pháp hồi quy từng bước (dựa trên cơ sở hồi quy bình phương trung bình) đề cập trên đây, để lọc các nhân tố dự báo phân tiếp sau đây luận án còn sử dụng phương pháp phân tích thành phần chính (PCA - Principal component analysis). Ưu điểm của phương pháp là cho phép nén thông tin từ một tập hợp khổng lồ các biến (nói chung không hạn chế) về một số rất ít các biến mới

nhưng vẫn bảo tồn được lượng thông tin cần thiết. Nhược điểm chính của phương pháp là đòi hỏi phải tính toán nhiều. Hơn nữa, khi sử dụng vào công tác dự báo đòi hỏi phải có đầy đủ thông tin của các biến theo yêu cầu. Chi tiết về phương pháp này được trình bày trong nhiều tài liệu, chẳng hạn [5, 9, 92]. Ở đây chỉ nêu tóm tắt nguyên lý của phương pháp.

Giả sử  $X(N,M)$  là ma trận số liệu ban đầu của  $M$  biến là tập nhân tố dự tuyến  $x_1, x_2, \dots, x_M$ , với  $N$  là tập mẫu quan trắc theo thời gian và  $x_{ij}, i = 1, 2, \dots, N; j = 1, 2, \dots, M$  là các phần tử của nó.  $M$  biến của  $X$  có thể là các chỉ số khí hậu hoặc các điểm không gian của các trường khí quyển, đại dương.

Phép phân tích thành phần chính (PCA) là phép biến đổi ma trận  $X(N,M)$  thành ma trận  $Z(N,M)$ :

$$Z(N, M) = X(N, M).A(M, M) \quad (2.17)$$

Trong đó  $A(M,M)$  là ma trận biến đổi. Trên thực tế thay cho ma trận  $X(N,M)$  là ma trận  $Y(N,M)$  được chuẩn hoá sao cho các phần tử của  $Y$  là:

$$y_{ij} = \frac{x_{ij} - \bar{x}_j}{\sqrt{N}}, i = 1, \dots, N; j = 1, \dots, M \quad (2.18)$$

với  $\bar{x}_j = \frac{1}{N} \sum_{i=1}^N x_{ij}, j = 1, \dots, M$ . Khi đó thay vai trò của  $X$  bởi  $Y$  trong 2.17 ta có:

$$Z(N, M) = Y(N, M).A(M, M) \quad (2.19)$$

Khi  $A(M,M)$  là ma trận trực chuẩn ta có:

$$Y(N, M) = Z(N, M).A'(M, M) \quad (2.20)$$

Từ đó ta có:

$$Y'(M, N).Y(N, M) = A(M, M).Z'(M, N).Z(N, M).A'(M, M) \quad (2.21)$$

$$\text{Hay } S_x(M, M) = A(M, M).S_z(M, M).A'(M, M) \quad (2.22)$$

Trong các công thức trên, dấu phẩy trên (') chỉ ma trận chuyển vị;  $S_x(M, M)$ ,  $S_z(M, M)$  tương ứng là ma trận covariance của  $Y$  và  $Z$ . Vì các cột của  $Z(N,M)$  là các vector trực giao nên  $S_z(M, M) = \Lambda(M, M)$  là ma trận đường chéo mà các phần tử trên đường chéo chính là các trị riêng  $\lambda_i$  của ma trận  $S_x(M, M)$ , còn các cột của ma trận  $A(M, M)$  là các vector riêng tương ứng với các trị riêng đó. Như vậy phép biến đổi từ  $Y$  sang  $Z$  tương đương với việc chéo hoá ma trận  $S_x(M, M)$ . Nếu các

vector của  $Z$  là trực chuẩn khi đó  $S_z(M, M) = I(M, M)$  là ma trận đơn vị và  $A(M, M).A'(M, M) = \Lambda(M, M)$ .

Thông thường khi ứng dụng PCA vào bài toán dự báo các cột của  $Z(N, M)$  là các biến mới và chúng là những vector trực chuẩn. Còn các phần tử trong các cột của ma trận  $A(M, M)$  là hệ số tương quan giữa các biến mới với các biến ban đầu.

Trên thực tế, biểu thức (2.17) được biểu diễn gần đúng bởi:

$$Z(N, K) \approx Y(N, M).A(M, K) \quad (2.23)$$

Trong đó  $K \ll M$ , nghĩa là thay cho việc lấy tất cả  $M$  biến mới chỉ cần lấy  $K$  biến đầu tiên ứng với  $K$  giá trị riêng lớn nhất và ma trận biến đổi  $A(M, K)$  chỉ giữ lại  $K$  cột đầu tiên. Sai số của phép xấp xỉ này được xác định bởi:

$$\eta = 1 - \frac{\sum_{i=1}^K \lambda_i}{\sum_{i=1}^M \lambda_i} \quad (2.24)$$

Trong đó  $\lambda_i$  là các thành phần của ma trận đường chéo  $\Lambda(M, M)$ . Từ hệ thức (2.23) dễ thấy rằng nếu có một bộ số liệu mới của tập nhân tố ban đầu  $X$  (một hàng mới của  $X$ ), sau khi thực hiện phép biến đổi qua (2.18) ta nhận được các giá trị của  $Y$  tương ứng. Giả thiết rằng ma trận  $A$  không đổi ta có thể nhận được các giá trị mới của biến  $Z$ .

Khi xem xét  $M$  là một biến duy nhất thì kỹ thuật PCA như trình bày ở trên còn được biết đến với tên gọi là phân tích các hàm trực giao thực nghiệm (EOFs) [22, 43]. Tuy vậy, về bản chất hai tên gọi này là giống nhau, do vậy trong luận án này sử dụng thuật ngữ EOF trong trường hợp khi thực hiện phân tích trên một biến theo các xu hướng không gian và thời gian trong khi thuật ngữ PCA phân tích từ hai hoặc nhiều biến số mà mỗi biến phát triển theo không gian riêng và được sắp xếp lại dữ liệu thành các "mode" phát triển theo thời gian.

#### **2.2.4 Vấn đề kiểm nghiệm phương trình dự báo**

Đối với mô hình MLR, sau khi tìm được phương trình hồi quy việc đánh giá thống kê khả năng sử dụng của mô hình có thể được thực hiện hoặc bằng kiểm nghiệm F cho các hệ số hồi quy hoặc kiểm nghiệm t cho hệ số tương quan bội. Đối

với các mô hình LAD, LMV và ANN việc đánh giá đó chỉ thực hiện cho hệ số tương quan bội. Trong mô hình MLR, hệ số tương quan bội là hệ số tương quan giữa tập số liệu ước lượng và quan trắc của  $y$ , tức giữa  $\hat{y}$  và  $y$ . Vì vậy, để thuận tiện trong việc so sánh, luận án chỉ sử dụng phương pháp đánh giá thông qua kiểm nghiệm độ tin cậy của hệ số tương quan giữa  $\hat{y}$  và  $y$ .

Hệ số tương quan được tính bởi:

$$r = \frac{\sum_{i=1}^n (y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum_{i=1}^n (y_i - \bar{y})^2} \sqrt{\sum_{i=1}^n (\hat{y}_i - \bar{\hat{y}})^2}} \quad (2.25)$$

Về bản chất hệ số tương quan  $r$  là ước lượng của hệ số tương quan lý thuyết  $\rho$ , được tính từ tập mẫu nên nói chung khác 0. Tuy nhiên  $r \neq 0$  không đồng nghĩa với  $\rho \neq 0$ , nghĩa là  $r \neq 0$  nhưng có thể không tồn tại mối tương quan giữa  $\hat{y}$  và  $y$ . Vì vậy, giả thiết cần kiểm nghiệm là:

$$H_0: \rho = 0$$

Nếu  $H_0$  sai thì giữa  $\hat{y}$  và  $y$  có tồn tại mối tương quan tuyến tính, nghĩa là phương trình hồi quy dùng được. Ngược lại thì phương trình hồi quy không có ý nghĩa. Để kiểm nghiệm ta cần tính đại lượng  $t$ :

$$t = \frac{r}{\sqrt{1-r^2}/\sqrt{n-2}} \quad (2.26)$$

Đại lượng  $t$  có phân bố Student với  $n-2$  bậc tự do ( $t \in \text{St}(n-2)$ ). Với xác suất phạm sai lầm loại I cho trước bằng  $\alpha$  nào đó (thông thường bằng 5%) ta tính được  $t_\alpha$  từ phân bố  $\text{St}(n-2)$ .

$$P(t > t_\alpha) = \alpha \quad (2.27)$$

Do đó, nếu  $t > t_\alpha$  thì  $H_0$  sai và phương trình hồi quy có ý nghĩa sử dụng. Ngược lại thì phương trình hồi quy không có ý nghĩa.

### 2.3 Đánh giá sai số dự báo của các mô hình

Sai số của các mô hình được đánh giá dựa trên một số chỉ số thống kê cơ bản gồm sai số trung bình (ME - Mean Error), sai số tuyệt đối trung bình (MAE - Mean Absolute Error), sai số bình phương trung bình quân phương (RMSE - Root mean

square Error), tỷ lệ sai số trung bình tuyệt đối so với chuẩn khí hậu (RMAE) và hệ số tương quan (Correlation coefficient) giữa giá trị dự báo và giá trị quan trắc.

Kí hiệu  $F_i$  và  $O_i$  tương ứng là giá trị dự báo của mô hình và giá trị quan trắc của yếu tố dự báo (số lượng XTNĐ),  $n$  là dung lượng mẫu,  $Cl_i$  là giá trị trung bình khí hậu, các đại lượng trên được tính theo các công thức sau:

$$\text{Sai số trung bình: } ME = \frac{1}{n} \sum_{i=1}^n (F_i - O_i) \quad (2.28)$$

$$\text{Sai số tuyệt đối trung bình: } MAE = \frac{1}{n} \sum_{i=1}^n |F_i - O_i| \quad (2.29)$$

Tỷ lệ sai số trung bình tuyệt đối so với chuẩn khí hậu RMAE:

$$RMAE = MAE / Cl_i \quad (2.30)$$

$$\text{Sai số bình phương trung bình: } RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (F_i - O_i)^2} \quad (2.31)$$

$$\text{Hệ số tương quan: } Corr = \frac{\sum_{i=1}^n (F_i - \bar{F})(O_i - \bar{O})}{\sqrt{\sum_{i=1}^n (F_i - \bar{F})^2} \sqrt{\sum_{i=1}^n (O_i - \bar{O})^2}} \quad (2.32)$$

Sai số trung bình ME cho biết xu hướng lệch trung bình của giá trị dự báo so với giá trị quan trắc, nhưng không phản ánh độ lớn của sai số. ME dương phản ánh mô hình dự báo thiên cao và ngược lại. Giá trị lý tưởng là  $ME=0$ . MAE được hiểu là sai số bậc nhất phản ánh mức độ sai lệch trung bình của dự báo so với quan trắc. Giá trị lý tưởng là  $MAE=0$ , tức là dự báo luôn bằng quan trắc. Nếu trị tuyệt đối của ME và MAE tương đương nhau thì có thể dùng ME để hiệu chỉnh sản phẩm dự báo một cách đáng tin cậy. RMSE còn được gọi là sai số bậc hai, dùng để đánh giá mức độ biến động sai số khi sử dụng cùng với MAE. Vì  $RMSE \geq MAE$  nên khi RMSE không quá lớn so với MAE thì sai số được xem là ổn định. Khi RMSE lớn hơn nhiều so với MAE thì sai số dự báo có thể chứa những giá trị đột biến vì RMSE rất nhạy với những giá trị sai số lớn. Mô hình là “hoàn hảo” nếu  $RMSE=0$ . Hệ số tương quan (Corr) cho phép đánh giá mối quan hệ tuyến tính giữa tập giá trị dự báo và tập giá trị quan trắc. Thực chất đó là công thức (2.25) nhưng các thành phần bên trong đã được thay thế để đồng bộ với các công thức sai số từ 2.28 đến 2.31.

Ngoài 5 chỉ số ME, MAE, RMSE, RMAE, CORR, để đánh giá kết quả dự báo của mô hình với dự báo khí hậu, luận án xem xét thêm chỉ số điểm kỹ năng MSSS của các phương trình dự báo, MSSS được tính toán dựa trên sai số bình phương trung bình:

$$MSSS = 1 - \frac{\frac{1}{n} \sum_{i=1}^n (F_i - O_i)^2}{\frac{1}{n} \sum_{i=1}^n (\overline{CL} - O_i)^2} \quad (2.33)$$

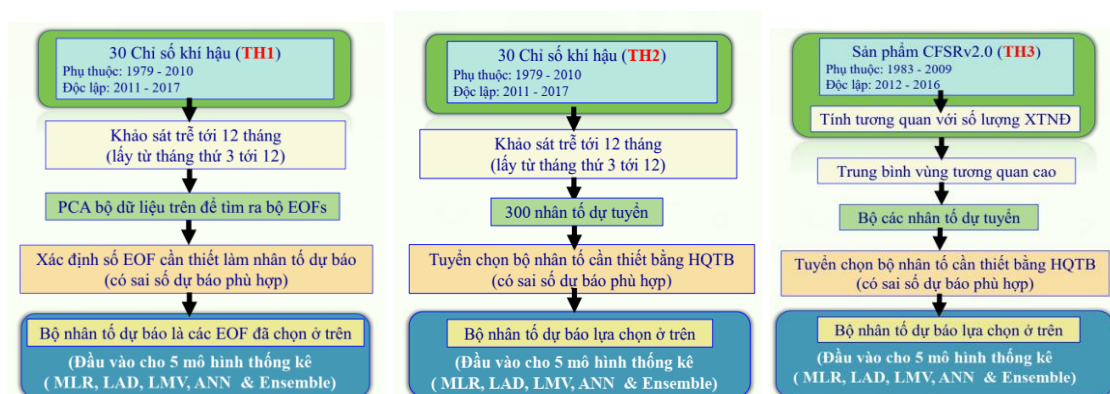
Ở đây,  $\overline{CL}$  là giá trị dự báo dựa trên trung bình khí hậu, nghĩa là luôn luôn đưa ra giá trị dự báo bằng chính giá trị trung bình khí hậu (Cli). Giá trị MSSS nằm trong khoảng  $(-\infty, 1)$ . Từ (2.33) thấy rằng nếu mọi dự báo bằng quan trắc thì tử số ở hạng thứ hai về phải bằng 0 và do đó MSSS=1. Nghĩa là nếu mô hình là hoàn hảo thì MSSS=1. Nếu tử số ở hạng thứ hai về phải lớn hơn mẫu số, tức là sai số dự báo của mô hình lớn hơn sai số dự báo khí hậu, khi đó MSSS<0. Vậy mô hình được xem là tốt hơn dự báo khí hậu nếu MSSS>0. MSSS càng lớn thì mô hình càng có hiệu quả so với dự báo khí hậu.

#### **2.4. Thiết kế thí nghiệm dự báo hạn mùa số lượng XTNĐ**

Với mục đích dự báo hạn mùa số lượng XTNĐ trên Biển Đông bằng các mô hình thống kê khi sử dụng các chỉ số khí hậu và các trường dự báo của CFSv2 làm nhân tố dự báo, luận án đã thực hiện ba phương án thí nghiệm. Trong mỗi phương án luận án xây dựng các phương trình dự báo theo bốn mô hình thống kê (MLR, LAD, LMV và ANN) trên hai tập số liệu: một tập số liệu phụ thuộc dùng để xây dựng các phương trình dự báo và một tập số liệu độc lập để dự báo thử nghiệm bằng các phương trình đã tìm được. Yếu tố dự báo của luận án là tổng số lượng XTNĐ trên khu vực Biển Đông 3 tháng liên tiếp đầu mùa bão (tháng 6, 7, 8) và 3 tháng cuối mùa bão (tháng 9, 10, 11). Hình 2.3 dưới đây minh họa các phương án thí nghiệm của luận án, theo đó:

1) Phương án 1 (ký hiệu là TH1): thống kê truyền thống với các nhân tố được lựa chọn từ tập các chỉ số khí hậu bằng phương pháp phân tích thành phần chính. Từ quan điểm cho rằng tín hiệu từ các chỉ số khí hậu trong các tháng liền trước có thể liên quan tới số lượng XTNĐ của các tháng tiếp sau, với tập 30 chỉ số

khí hậu được xem xét tới độ trễ 10 tháng, từ tháng thứ 3 đến tháng thứ 12 trở về trước tính từ tháng hiện tại, sẽ có tất cả 300 nhân tố dự tuyển. Các nhân tố dự báo tìm được cho trường hợp này không phải chính các chỉ số khí hậu đó mà là các hàm EOFs. Số lượng hàm trực giao EOFs (số thành phần chính được giữ lại) phải đảm bảo đóng góp trên 70% lượng thông tin ban đầu và đảm bảo sai số phù hợp cho bài toán dự báo số lượng XTNĐ Biển Đông (sẽ được trình bày chi tiết tại Chương 4).



Hình 2.3 Ba trường hợp nghiên cứu, khảo sát khả năng dự báo hạn mùa số lượng XTNĐ hoạt động trên khu vực Biển Đông

2) Phương án 2 (ký hiệu là TH2): thống kê truyền thống với các nhân tố được lựa chọn từ tập các chỉ số khí hậu bằng phương pháp hồi quy từng bước. Với bộ chỉ số khí hậu được thành lập như TH1 nhưng các nhân tố dự báo được tìm thông qua kỹ thuật hồi quy từng bước. Các nhân tố dự báo tìm được theo phương án này chứa thông tin của chỉ số khí hậu được chọn và thời gian trễ tương ứng. Việc xác định, lựa chọn các nhân tố dự báo cho TH2 phải đảm bảo có sai số phù hợp cho bài toán dự báo số lượng XTNĐ Biển Đông (sẽ được trình bày chi tiết tại Chương 4).

3) Phương án 3 (ký hiệu là TH3): thống kê động lực từ sản phẩm dự báo của mô hình CFSv2. Khác với hai trường hợp đầu, mô hình CFSv2 có hạn dự báo đến 9 tháng và số liệu được cho trên lưới kinh vĩ của rất nhiều trường, cả các trường bề mặt và trên các mực đẳng áp. Để lựa chọn được các trường hợp lý, luận án đã xây dựng các bản đồ tương quan và qua đó phân tích, đánh giá căn cứ vào độ lớn của hệ số tương quan cũng như ý nghĩa thống kê của chúng. Từ các trường được chọn, các

vùng có tương quan tốt với yếu tố dự báo sẽ được xác định. Giá trị trung bình trường các vùng đó được sử dụng như là những nhân tố dự tuyển. Việc lựa chọn các nhân tố dự báo từ tập nhân tố dự tuyển này được thực hiện bằng phương pháp hồi quy từng bước (sẽ được trình bày chi tiết tại Chương 4).

#### **Nhận xét chung:**

- Về số liệu: sau khi nghiên cứu, khảo sát các nguồn số liệu, để thực hiện mục tiêu, nội dung đề tài, luận án lựa chọn sử dụng số liệu về bão do RSMC phát hành, số liệu dự báo lại từ sản phẩm mô hình CFSv2 và 30 chỉ số khí hậu được khai thác từ NOAA và CPC. Bên cạnh đó, Chương 2 cũng giới hạn đối tượng và phạm vi nghiên cứu là các XTNĐ có cường độ gió mạnh nhất đạt từ cấp 6 trở lên trên khu vực Biển Đông theo Quyết định 03/2020/QĐ-TTg ngày 13/01/2020.

- Về phương pháp: Chương 2 đưa ra 4 mô hình dự báo thông kê (MLR, LAD, LMV và ANN) và tổ hợp trung bình. Bên cạnh đó, 2 kỹ thuật lọc nhân tố dự báo là phân tích thành phần chính và hồi quy từng bước được luận án giới thiệu để định hướng tìm ra bộ nhân tố dự báo phù hợp cho bài toán dự báo hạn mùa số lượng XTNĐ trên Biển Đông. Bên cạnh 5 chỉ số đánh giá sai số (ME, MAE, RMSE, RMAE, CORR), kết quả dự báo của mô hình còn được xem xét với chỉ số điểm kỹ năng MSSS.

- Về các trường thử nghiệm: luận án tiến hành với 3 trường hợp, TH1, TH2 tìm kiếm nhân tố dự báo từ tập các chỉ số khí hậu, tương ứng với cách xác định các thành phần chính của bộ chỉ số khí hậu (TH1) và các chỉ số khí hậu thỏa mãn hồi quy từng bước (TH2) trong khi TH3 sử dụng sản phẩm mô hình động lực CFSv2 để lựa chọn nhân tố dự báo.

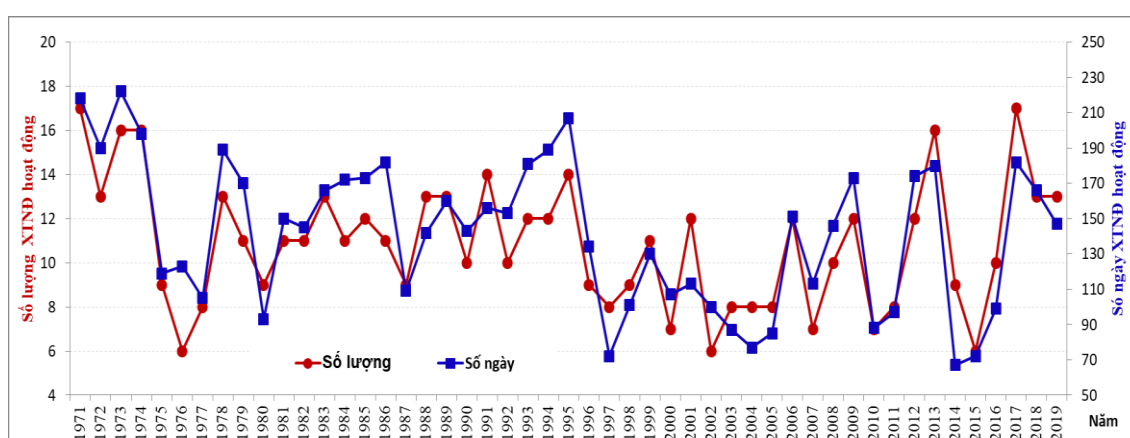
Từ các nguồn số liệu và phương pháp kể trên, Chương 3 và Chương 4 dưới đây sẽ lần lượt trình bày các kết quả của luận án tương ứng với hai nội dung chính là đặc điểm hoạt động và khả năng dự báo hạn mùa số lượng XTNĐ trên Biển Đông.

### CHƯƠNG III. MỘT SỐ ĐẶC ĐIỂM CỦA XTNĐ TRÊN BIỂN ĐÔNG VÀ MỐI QUAN HỆ VỚI CÁC QUÁ TRÌNH QUY MÔ LỚN

Trong chương này sẽ trình bày một số đặc trưng khí hậu điển hình của XTNĐ hoạt động trên Biển Đông, bao gồm tần suất hoạt động, quỹ đạo và vùng hoạt động cũng như mối quan hệ giữa chúng với một số quá trình quy mô lớn.

#### 3.1 Tần suất XTNĐ hoạt động

Hình 3.1 dẫn ra biến động hàng năm của số lượng và số ngày XTNĐ hoạt động trên Biển Đông tính theo bộ số liệu RSMC cho thời kỳ 1971-2019.

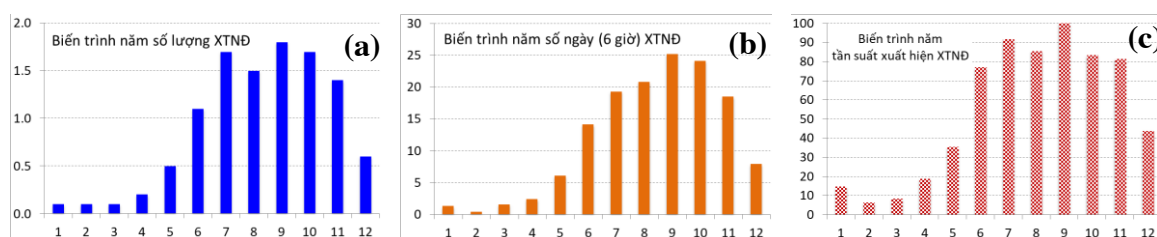


Hình 3.1 Biến động hàng năm của số lượng và số ngày hoạt động của XTNĐ trên Biển Đông giai đoạn 1971-2019

Kết quả tại Hình 3.1 chỉ ra, số lượng (đường đồ thị chấm bi, màu đỏ) và số ngày XTNĐ hoạt động (đường đồ thị ô vuông, màu xanh) trên khu vực Biển Đông trong giai đoạn khảo sát có sự biến động khá rõ nét. Nhìn chung, diễn biến số lượng và số ngày hoạt động của XTNĐ trên Biển Đông là phù hợp với nhau, những năm có nhiều XTNĐ thì tổng thời gian hoạt động cũng tăng tỷ lệ. Trung bình nhiều năm giai đoạn 1971-2019, số lượng XTNĐ theo RSMC là  $10,9 \pm 2,9$  XTNĐ. Năm nhiều nhất là 17 cơn (năm 1971, 2017) và ít nhất là 6 cơn (năm 1976, 2015). Kết quả thống kê tại Hình 3.1 thể hiện XTNĐ hoạt động nhiều vào các giai đoạn 1971-1974, 1988-1995 hoặc các năm 2013, 2017.

Biến trình năm của số lượng, số ngày hoạt động và tần suất xuất hiện của XTNĐ trên Biển Đông được dẫn ra trên Hình 3.2. Nếu xem mùa bão là các tháng liên tục trong năm có tần suất xuất hiện trên 50% và số lượng XTNĐ trung bình ít

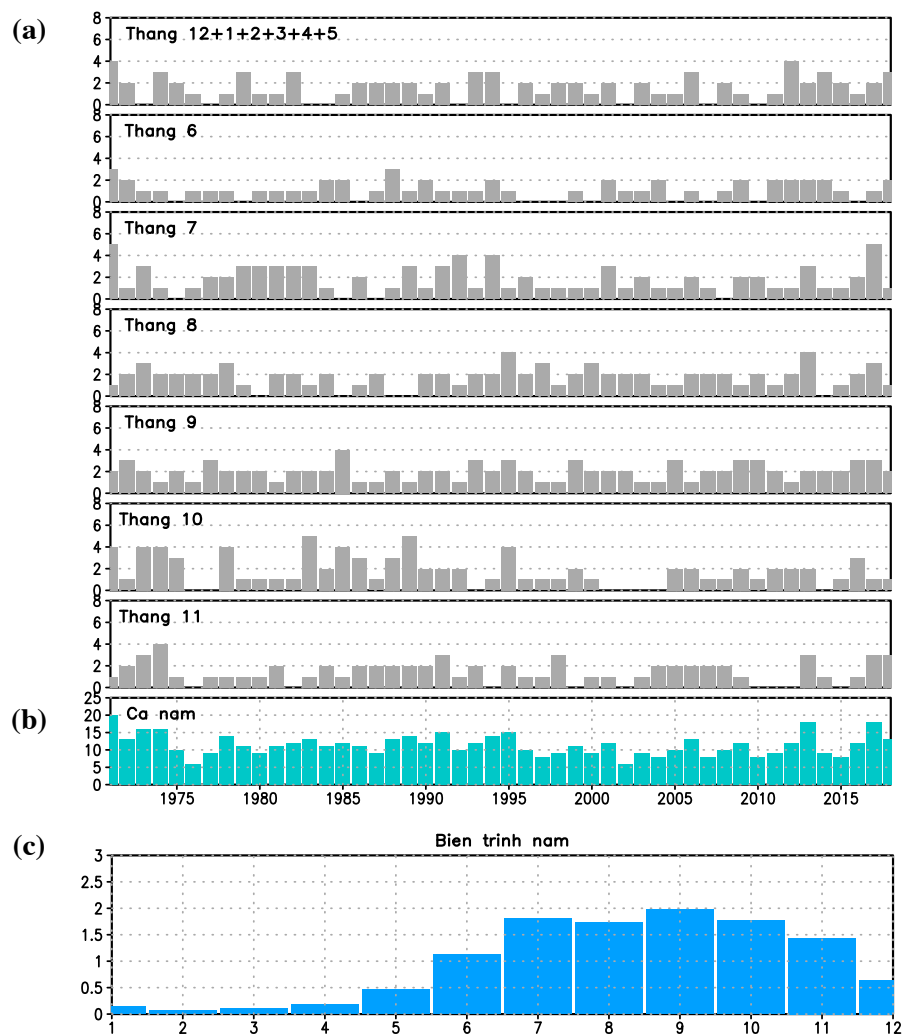
nhất 1 cơn thì mùa bão trên Biển Đông là từ tháng 6 đến tháng 11. Nếu xem các tháng liên tục có trung bình trên 1,5 cơn là mùa bão chính thì mùa bão chính trên Biển Đông rơi vào các tháng 7-10 (4 tháng). Trong các tháng mùa bão trên Biển Đông trung bình mỗi tháng có trên 4 ngày (trên 16 obs 6 giờ) bão hoạt động. Đối với mùa bão chính thời gian bão hoạt động vào khoảng 5-7 ngày (20-30 obs 6 giờ). Trong năm, tháng 1 đến 4 là thời gian có ít bão nhất, trong đó tháng 2 có tần suất bão hoạt động chỉ 6,2% và số lượng trung bình chỉ 0,1 cơn. Các tháng 5 và 12 mặc dù chưa phải là các tháng mùa bão nhưng tần suất xuất hiện cũng khá lớn, khoảng từ 35-45% và trung bình mỗi tháng có khoảng trên 0,5 cơn bão. Tần suất bão xuất hiện trong các tháng mùa bão nói chung vượt quá 75%, trong đó các tháng mùa bão (tháng 6 tới tháng 11) đạt trên dưới 90%. Riêng tháng 9 hầu như 100% số năm trong giai đoạn khảo sát (1971-2019) đều có bão xuất hiện.



Hình 3.2 Biến trình năm số lượng (Hình a), số ngày hoạt động (Hình b) và tần suất xuất hiện (Hình c) của XTNĐ trên Biển Đông giai đoạn 1971-2019

Diễn biến số lượng XTNĐ theo tháng ở các năm giai đoạn nghiên cứu được trình bày trên Hình 3.3. Có thể thấy bậc đại lượng của tổng số XTNĐ trong các tháng ngoài mùa bão tương đương với số lượng XTNĐ từng tháng trong mùa bão (Hình 3.3 a). Trong các tháng mùa bão, tần suất xuất hiện trên 2 XTNĐ dao động trong khoảng từ 31% (tháng 6) đến 73% (tháng 9), trên 3 cơn là 15% (tháng 8) và 29% (tháng 7). Tần suất xuất hiện từ 4 XTNĐ trở lên xảy ra vào các tháng 7 đến tháng 11, trong đó tỷ lệ này cao nhất ở tháng 10 với khoảng 16,3%, tiếp đến là tháng 7 với tỷ lệ khoảng 8,1%, trong khi các tháng còn lại tỷ lệ này chỉ ở mức 2%. Như vậy có thể thấy mặc dù các tháng 7, 8, 9 có tần suất xuất hiện XTNĐ nhiều

nhất (tháng 9 hầu như bằng 100%) nhưng tháng 10 lại là tháng mà XTNĐ thường xuất hiện nhiều nhất.

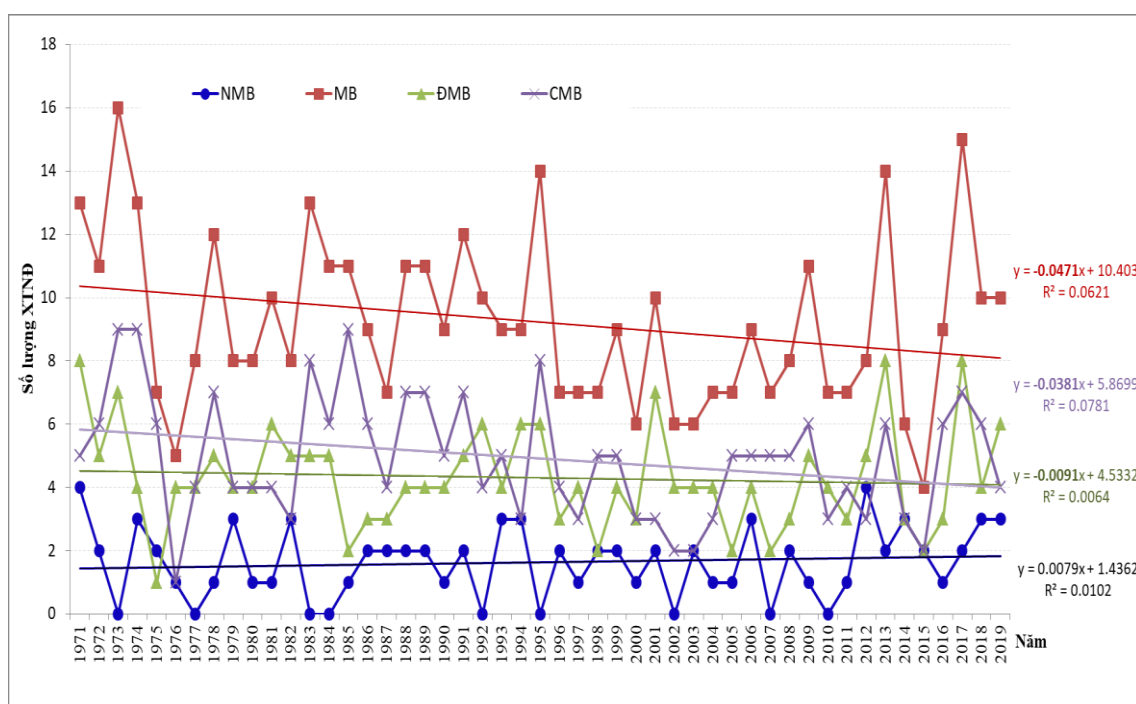


Hình 3.3 Biến động hàng năm của các tháng ngoài mùa bão, các tháng trong mùa bão (Hình a), cả năm (Hình b) và biến trình năm (Hình c) số lượng XTNĐ hoạt động trên Biển Đông giai đoạn 1971-2019

Hình 3.4 biểu diễn biến động hàng năm và xu thế tuyến tính của tổng số XTNĐ trong mùa bão (tháng 6 tới tháng 11) và ngoài mùa bão (các tháng 1 tới tháng 5 và tháng 12). Sự biến động số lượng XTNĐ trong mùa bão được xem xét với 6 tháng mùa bão (tháng 6 tới tháng 11), 3 tháng đầu mùa bão (tháng 6-7-8) và 3 tháng cuối mùa bão (tháng 9-10-11). Kết quả này cho thấy, tổng lượng XTNĐ trong các tháng ngoài mùa bão biến động khá lớn, dao động trong khoảng  $1,6 \pm 1,1$  cơn,

trong đó có những năm không có nhưng có năm có đến 4 XTNĐ (1971, 2012). Số lượng XTNĐ trong 6 tháng mùa bão (tháng 6 tới tháng 11) trung bình khoảng  $9,2 \pm 2,7$  cơn, trong đó 3 tháng cuối mùa bão trung bình khoảng  $(4,9 \pm 1,9)$  cơn thể hiện biến động qua các năm mạnh hơn 3 tháng đầu mùa  $(4,3 \pm 1,6)$  cơn).

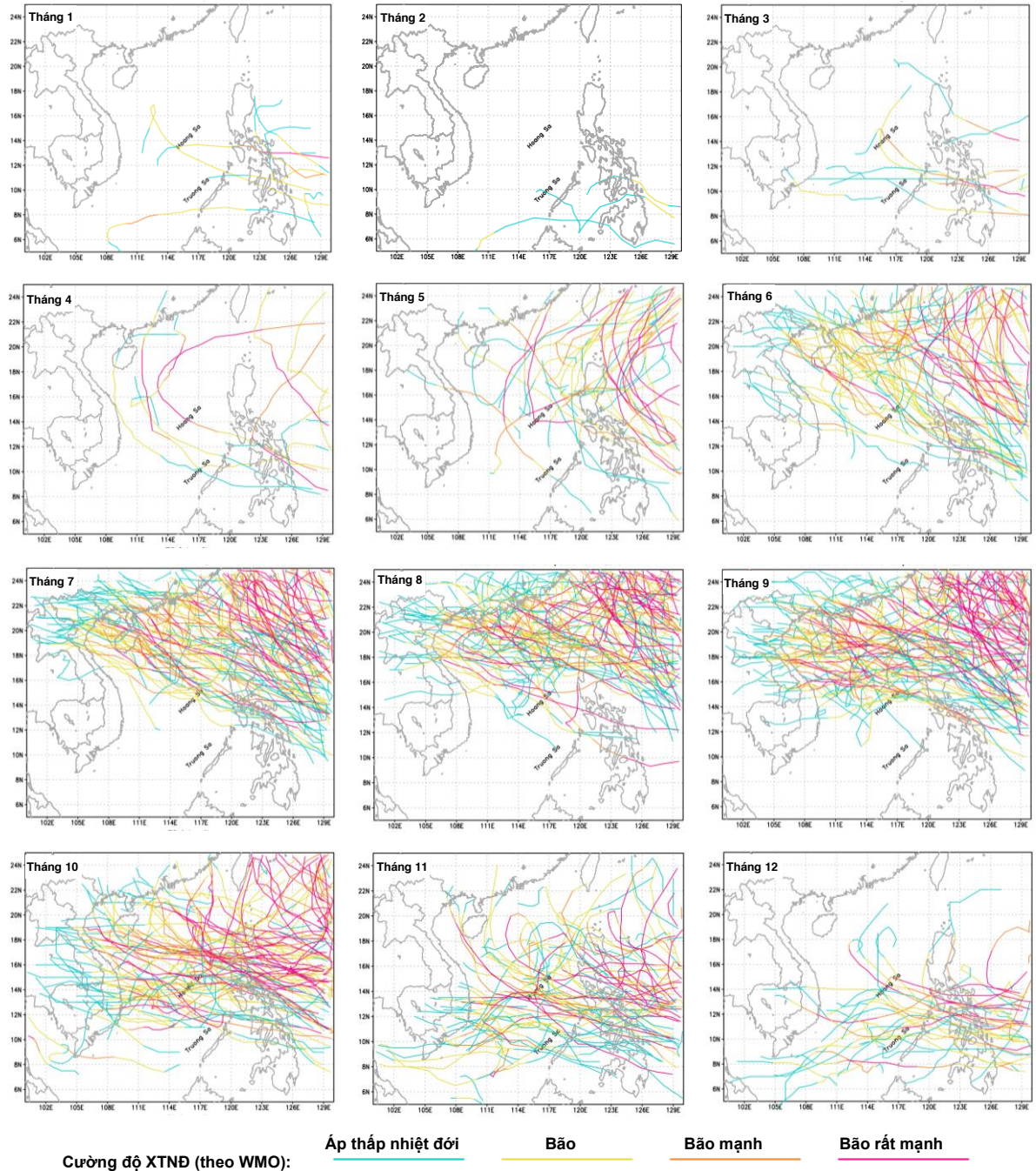
Để đánh giá về xu thế biến đổi số lượng XTNĐ Biển Đông ở đây luận án dựa trên phương pháp tính toán và kiểm nghiệm xu thế phi tham số Mann-Kendall và Sen [46, 74]. Phương pháp này có ưu điểm là loại bỏ được tính nhạy cảm của sai số quan trắc cũng như tính biến động bất thường của số lượng XTNĐ Biển Đông [12]. Kết quả cho thấy, tổng số lượng XTNĐ Biển Đông hoạt động trong mùa bão (tháng 6 tới tháng 11) và trong 3 tháng cuối mùa bão (tháng 9, 10, 11) có xu hướng giảm thoả mãn mức ý nghĩa 10%, tương ứng với mức giảm khoảng 0,5 cơn/thập kỷ ( $P_{val} = 0,06$ ) và 0,3 cơn/thập kỷ ( $P_{val} = 0,07$ ) trong khi đó xu thế tăng, giảm không rõ ràng đối với các XTNĐ hoạt động trong các tháng ngoài mùa bão và 3 tháng đầu mùa bão 6-7-8 (hệ số sen bằng 0 với  $P_{val}$  tương ứng là 0,45 và 0,31).



Hình 3.4 Biến động năm và xu thế biến đổi của tổng lượng XTNĐ trên Biển Đông trong cả mùa bão (MB: các tháng từ 6-11); đầu mùa bão (ĐMB: các tháng 6-7-8); cuối mùa bão (CMB: các tháng 9-10-11) và ngoài mùa bão (NMB: các tháng 1-5 và 12)

### 3.2 Quỹ đạo và vùng XTND hoạt động

Hình 3.5 dẫn ra quỹ đạo theo tháng của tất cả các XTND xuất hiện trên Biển Đông cho giai đoạn 1971-2019 theo số liệu RSMC.



Hình 3.5 Quỹ đạo XTND các tháng trong năm (trái sang phải, trên xuống dưới theo thứ tự từ tháng 1 đến tháng 12) khu vực Biển Đông giai đoạn 1971-2019

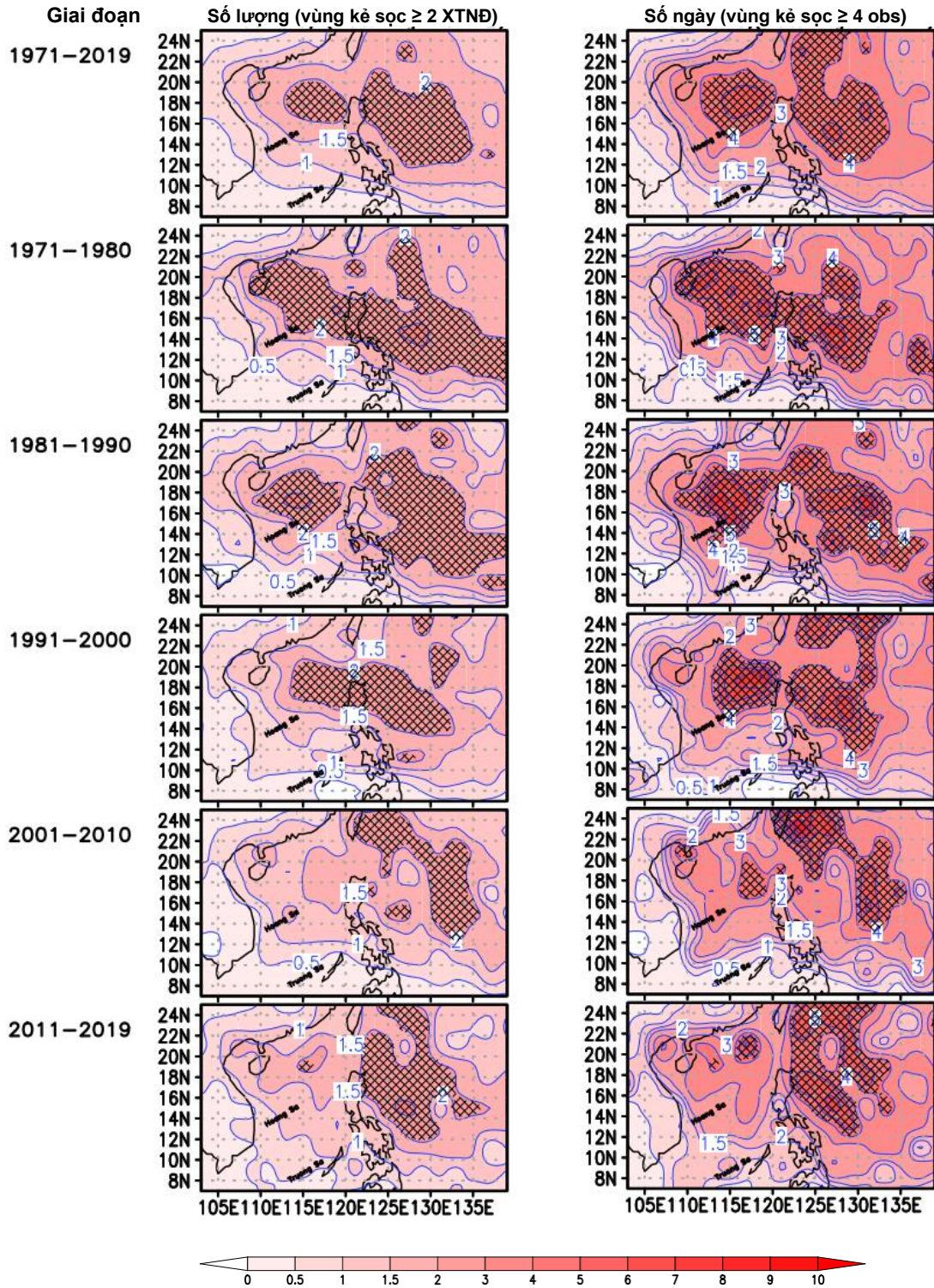
Kết quả tại Hình 3.5 cho thấy rõ, trong khoảng thời gian các tháng 1-4 XTNĐ ít xuất hiện và nếu có xuất hiện thì với cường độ yếu. Vùng XTNĐ hoạt động nói chung cách xa bờ biển Việt Nam và chủ yếu ở các vĩ độ phía nam. Tháng 5 là tháng trước mùa bão, trong đó mặc dù số lượng XTNĐ chưa nhiều nhưng vùng hoạt động chủ yếu ở góc Đông Bắc Biển Đông, ít khi có XTNĐ đổ bộ vào bờ biển Việt Nam.

XTNĐ hoạt động với tần suất tăng dần từ tháng 6 đến tháng 8 với quỹ đạo chủ yếu có hướng tây bắc hoặc tây tây bắc và hầu hết đổ bộ vào vùng bờ biển Bắc Bộ và Đông Nam Trung Quốc. Các tháng 9-11 XTNĐ có xu hướng dịch chuyển vào khu vực Bắc Trung Bộ đến Nam Trung Bộ và Nam Bộ. Hướng quỹ đạo chuyển dần sang tây tây bắc và tây hoặc tây nam. Tháng 10 XTNĐ ảnh hưởng chủ yếu đến vùng Trung Bộ với nhiều cơn bão mạnh. Tháng 11 XTNĐ ảnh hưởng chủ yếu đến khu vực Nam Trung Bộ và Nam Bộ với cường độ yếu hơn các tháng chính mùa bão. Tháng 12 quỹ đạo bão dường như chuyển hẳn sang hướng tây tây nam với cường độ yếu rõ rệt, chủ yếu là ATNĐ và ít có trường hợp đổ bộ vào đất liền.

Một cách khái quát có thể nói trong các tháng 5-6-7-8 XTNĐ thường hoạt động ở khu vực Bắc Biển Đông. Từ tháng 9-10-11-12 vùng hoạt động của XTNĐ dịch chuyển dần vào khu vực miền Trung và Nam Biển Đông. Các tháng ít XTNĐ nhất là từ tháng 1-2-3-4. Các XTNĐ có cường độ mạnh cũng xuất hiện nhiều hơn vào các tháng chính mùa bão.

Hình 3.6 biểu diễn sự biến động qua từng thập kỷ trung bình năm số lượng và số ngày (số obs 6 giờ) hoạt động của XTNĐ trên một phần khu vực TBTBD bao gồm cả Biển Đông giai đoạn 1971-2019 theo số liệu RSMC. Xét trên toàn giai đoạn nghiên cứu trên khu vực Biển Đông, XTNĐ hoạt động nhiều nhất ở vùng Đông Bắc, trung bình với khoảng 2 cơn/năm. Vùng hoạt động và số lượng XTNĐ có sự biến động đáng kể qua từng thập kỷ. Những năm 1971-1980 XTNĐ tập trung nhiều ở khu vực Đông Bắc Biển Đông nhưng đến thập kỷ 1981-1990 vùng nhiều XTNĐ đã dịch vào trung tâm Biển Đông. Thập kỷ 1991-2000 vùng nhiều XTNĐ lại dịch

lên phía đông bắc và có xu hướng giảm. Các năm từ 2001-2019 vùng XTND hoạt động mở rộng xuống khu vực Nam Biển Đông với mật độ dày đặc hơn.



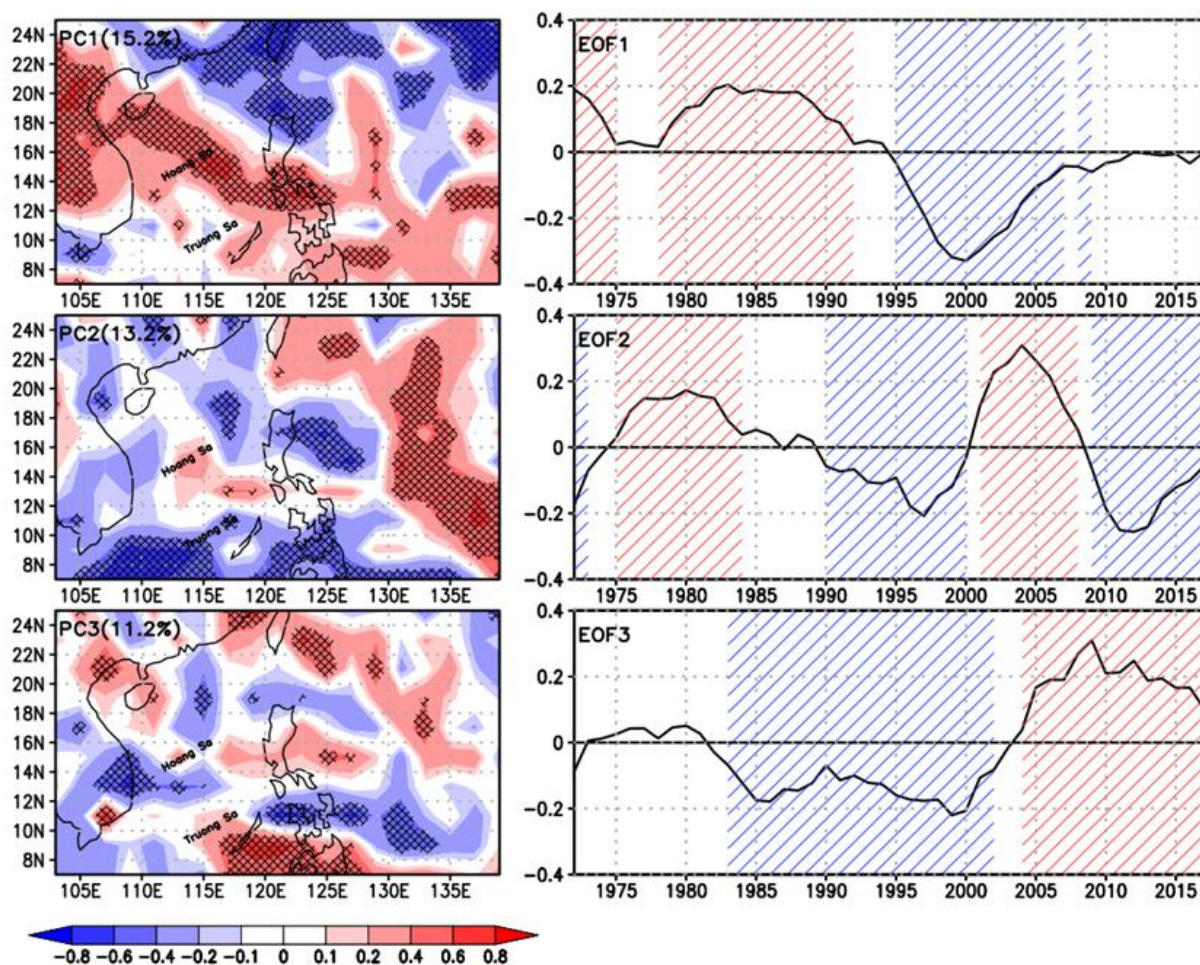
Hình 3.6 Trung bình năm số lượng (trái) và số ngày hoạt động (phải) của bão qua từng thập kỷ trên khu vực Biển Đông giai đoạn 1971-2019

Phân bố của số ngày XTNĐ hoạt động về cơ bản phù hợp với sự phân bố số lượng XTNĐ. Tính chung cho toàn giai đoạn khảo sát, vùng XTNĐ hoạt động trên một ngày (4 obs 6 giờ) trong năm thường nằm ở nửa phía bắc Biển Đông. Cũng như số lượng XTNĐ, số ngày XTNĐ hoạt động tập trung nhiều ở khu vực trung tâm và đông bắc Biển Đông trong những thập niên 1971-2000. Từ những năm 2001 đến nay số ngày có XTNĐ hoạt động phân bố dàn trải trên hầu khắp Biển Đông và có xu hướng gia tăng ở phần phía nam.

Như vậy, ngoài việc phân tích số liệu XTNĐ theo 12 tháng trong năm (Hình 3.5) và theo từng thập kỷ 10 năm (Hình 3.6) về quỹ đạo và vùng hoạt động của XTNĐ Biển Đông giai đoạn nghiên cứu như đã trình bày ở trên thì dưới đây luận án còn áp dụng phương pháp phân tích thành phần chính để nghiên cứu thêm về sự phân bố và biến động này. Phương pháp phân tích thành phần chính cho bộ số liệu RSMC giai đoạn 1971-2019 ở đây được áp dụng theo cách sau: Chia không gian nghiên cứu thành các ô lưới vuông với khoảng cách  $2^{\circ}$  kinh, vĩ, sau đó tiến hành thống kê số lượng XTNĐ di chuyển vào từng ô lưới trong từng năm cụ thể của thời gian nghiên cứu. Để tránh sự chồng chéo mỗi XTNĐ chỉ được tính một lần (một obs quan trắc) cho một ô lưới cụ thể. Nghĩa là nếu một XTNĐ nào đó được ghi nhận nhiều hơn một lần cũng chỉ được tính một lần duy nhất. Như vậy số liệu nhận được cho mỗi ô lưới hàng năm chính là số lượng XTNĐ (tần số) xuất hiện ở ô lưới đó. Áp dụng phân tích thành phần chính cho bộ số liệu này sẽ nhận được sự phân bố không gian của các thành phần chính và các chuỗi thời gian tương ứng (là các hàm trực giao EOF) của chúng. Hình 3.7 dưới đây trình bày phân bố không gian và biến thiên theo thời gian của 3 thành phần chính đầu tiên.

Kết quả cho thấy, 3 thành phần chính đầu tiên đóng góp khoảng gần 40% phương sai tổng cộng, trong đó thành phần thứ nhất (PC1) chiếm khoảng 15,2%, thành phần thứ hai khoảng 13,2% và thành phần thứ ba khoảng 11,2%. Có thể nhận thấy mỗi một thành phần chính biểu thị một dạng (mode) phân bố không – thời gian của XTNĐ trên khu vực, được dùng để giải thích cho mức độ biến động của XTNĐ gây nên bởi những nhân tố chi phối nào đó. Các bản đồ phân bố các thành phần

chính (PCs) là hệ số tương quan giữa các chuỗi thời gian (EOFs) với số lượng XTND đi qua mỗi ô lưới. Sự biến thiên của các thành phần EOFs thể hiện tín hiệu về sự thay đổi của tần số xuất hiện của XTND trong giai đoạn đang xem xét, cụ thể, giá trị dương thể hiện thời kỳ XTND hoạt động nhiều hơn trung bình nhiều năm và ngược lại. Các bản đồ phân bố theo không gian của các thành phần chính (PCs) mô tả hệ số tương quan, hay chính là mối quan hệ giữa tần số XTND xuất hiện và biến trình theo thời gian của các EOFs tại từng ô lưới. Dấu dương (âm) của hệ số tương quan thể hiện mối quan hệ đồng biến (nghịch biến) giữa tần số XTND xuất hiện và diễn biến theo thời gian của EOF. Giá trị tuyệt đối của hệ số tương quan càng tiệm cận đến 1 thì thể hiện mối quan hệ càng chặt chẽ giữa tần số XTND xuất hiện và diễn biến theo thời gian của EOF.

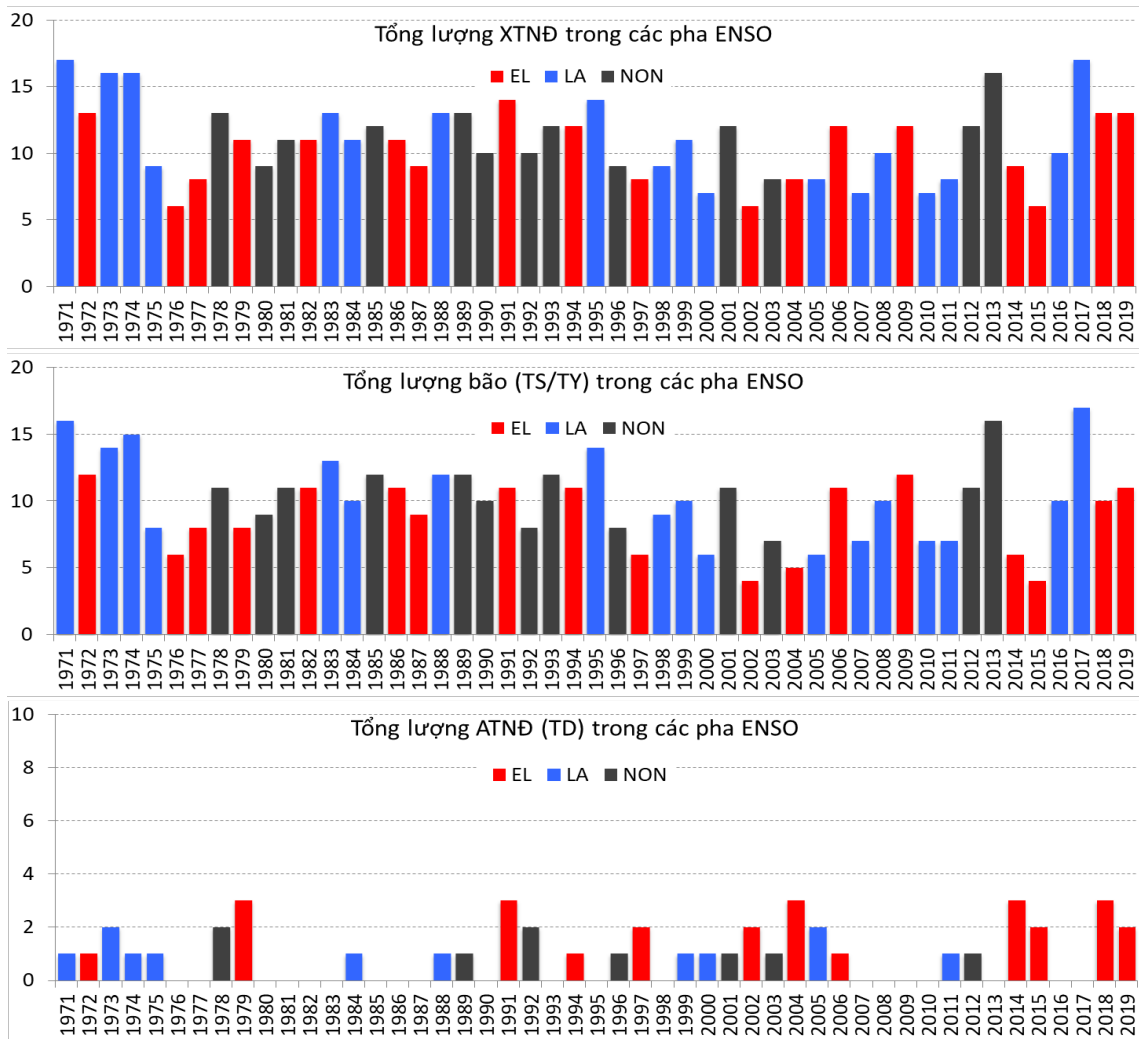


Hình 3.7 Ba thành phần chính (PCs) và chuỗi thời gian (EOFs) đầu tiên của số lượng XTND khu vực Biển Đông giai đoạn 1971-2019

Hình 3.7 cho thấy, ở dạng (mode) thứ nhất XTNĐ di chuyển vào vùng đông bắc Biển Đông và hướng lên phía Đài Loan, Nhật Bản nhiều hơn trung bình nhiều năm trong khoảng thời gian 1980 - 1990 và trước 1975, và đi vào khu vực đông nam Trung Quốc ít hơn trung bình nhiều năm trong những năm 1995-2005. Trong khi đó dạng thứ hai XTNĐ hoạt động nhiều hơn ở phía đông Philippine và phía đông khu vực giữa Biển Đông trong những năm 2000-2010 và 1975-1985, ít hơn trung bình nhiều năm ở phía nam và đông bắc Biển Đông. Dạng thứ ba thể hiện rõ hai giai đoạn trong đó XTNĐ có xu hướng đi vào khu vực Bắc Bộ và Nam Bộ trong thập niên gần đây (sau 2005) và giảm đi ở khu vực phía đông của Nam Trung Bộ. Dạng thứ ba này phù hợp với những kết quả nghiên cứu gần đây về sự dịch chuyển xuống phía nam của XTNĐ trên khu vực Biển Đông [13-14].

### **3.3 Quan hệ giữa số lượng XTNĐ trên Biển Đông và ENSO**

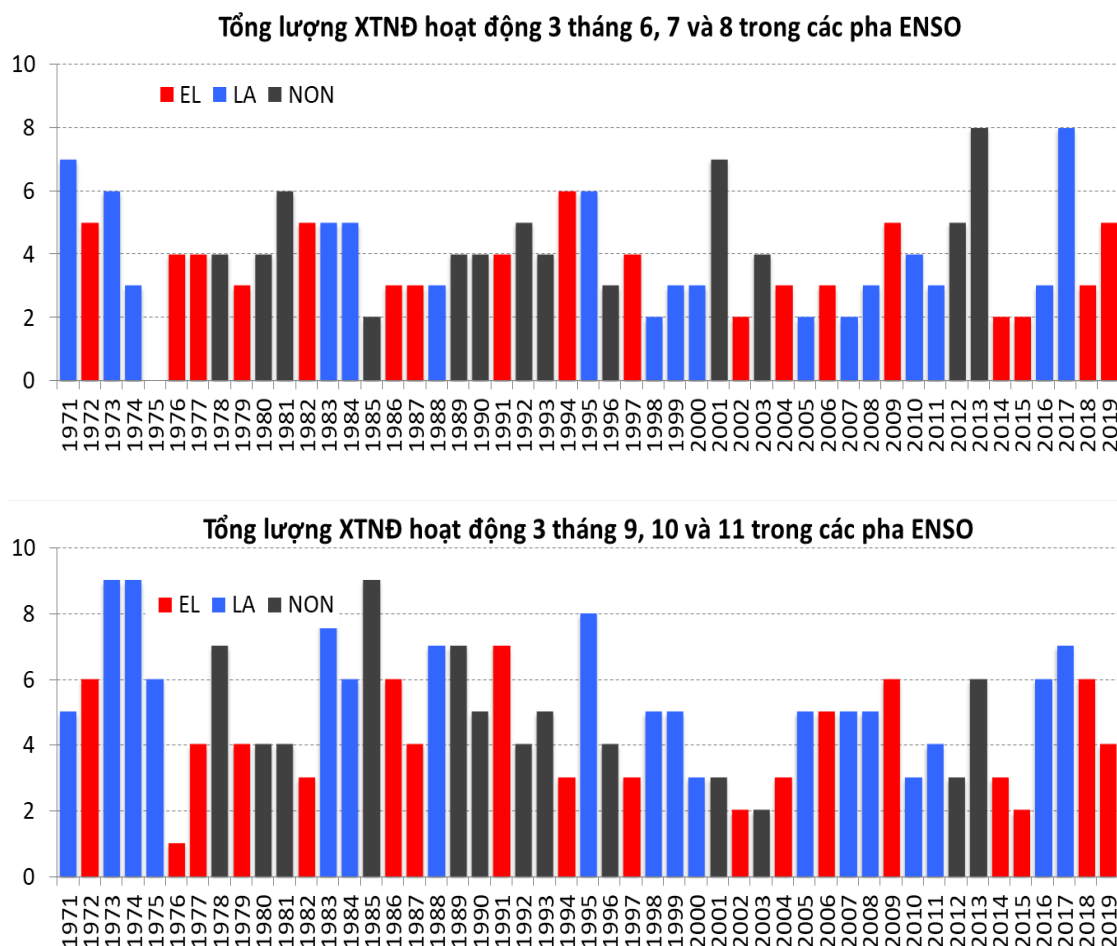
Một trong những nhân tố được cho là có ảnh hưởng đến sự hoạt động của XTNĐ Biển Đông là ENSO. Để có được bức tranh khái quát, trên cơ sở số liệu các năm ENSO được xác định từ NOAA ([https://origin.cpc.ncep.noaa.gov/products/analysis\\_monitoring/ensostuff/ONI\\_v5.php](https://origin.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ONI_v5.php)), số lượng XTNĐ trên khu vực Biển Đông được thống kê cho từng pha ENSO đối với tổng lượng XTNĐ hàng năm và cho từng thời kỳ 3 tháng đầu mùa bão và 3 tháng cuối mùa bão (Hình 3.8, Hình 3.9 và Bảng 3.1). XTNĐ được chia thành hai cấp: ATNĐ (Tropical Depression – TD) và bão (Tropical Storm Typhoon – TS/TY). Tính trung bình giai đoạn 1971-2019, tổng số XTNĐ trong các năm El Niño là  $10,1 \pm 2,7$  cơn/năm, trong các năm La Niña là  $11,3 \pm 3,5$  cơn/năm và trong các năm trung tính (Non-ENSO) là  $11,3 \pm 2,1$  cơn/năm. Như vậy so với trung bình nhiều năm, số lượng XTNĐ trên Biển Đông trong những năm El Niño ít hơn khoảng 1 cơn ( $10,1/10,9$ ) còn trong những La Niña nhiều hơn khoảng 0,4 cơn ( $11,3/10,9$ ). Đối với các XTNĐ đạt cường độ từ bão trở lên (TS/TY) giá trị trung bình tương ứng với các pha ENSO là  $8,7 \pm 2,8$ ;  $10,6 \pm 3,5$  và  $10,6 \pm 2,3$  cơn, còn với các XTNĐ có cường độ ATNĐ các giá trị đó tương ứng là  $1,4 \pm 1,2$ ;  $0,7 \pm 0,7$  và  $0,7 \pm 0,8$  cơn/năm.



Hình 3.8 Số lượng XTNĐ hoạt động trên Biển Đông trong các năm ENSO phân theo cấp bão (TS/TY) và áp thấp nhiệt đới (TD)

Từ Hình 3.8 có thể nhận thấy khá nhiều trường hợp số lượng XTNĐ của những năm sau năm El Niño thường nhiều lên, chẳng hạn các năm 1973, 1978, 1982, 1988, 1995, 1998, 2003, 2005, 2016. Tuy vậy điều đó không hoàn toàn là quy luật, bởi vẫn có những năm sau năm El Niño số lượng XTNĐ lại giảm đi, ví dụ các năm 1992, 2007, 2010. Điều đó cũng tương tự như đối với trường hợp XTNĐ đạt cường độ bão (TS/TY). Một vấn đề cần chú ý là trong những năm El Niño số lượng XTNĐ và bão trên Biển Đông thường ít hơn các năm La Niña nhưng ở các năm Trung tính (Non-ENSO) số lượng ATNĐ lại nhiều hơn. Tính trung bình những năm El Niño có khoảng  $1,4 \pm 1,2$  ATNĐ/năm, năm La Niña và Non-ENSO khoảng

0,7ATNĐ/năm (Bảng 3.1). Số năm El Niño có từ 3 ATNĐ trở lên chiếm tỷ lệ khá lớn so với các năm La Niña và Non-ENSO (Hình 3.8).



Hình 3.9 Số lượng XTNĐ Biển Đông trong 3 tháng đầu mùa bão, 3 tháng cuối mùa bão theo các năm ENSO

Số lượng XTNĐ trong ba tháng đầu mùa bão (tháng 6-7-8) và cuối mùa bão (tháng 9-10-11) biến động khá mạnh theo các năm ENSO và dường như không có quy luật rõ ràng (Hình 3.9). Số lượng XTNĐ ở 3 tháng đầu mùa bão tương đương nhau giữa năm El Niño và La Niña (3,7 tới 3,8 cơn) trong khi ở 3 tháng cuối mùa bão lại chênh lệch nhau khoảng 2 cơn (4,0 và 5,9 cơn) (Bảng 3.1). Như vậy số lượng XTNĐ trên Biển Đông trong những năm La Niña nhiều hơn trong những năm El Niño khoảng gần 2 cơn/năm và thường tập trung vào 3 tháng cuối mùa bão.

Bảng 3.1 Số lượng và số ngày XTNĐ Biển Đông hoạt động trong các pha ENSO

Đặc trưng	Các pha ENSO			
	El Niño	La Niña	Non ENSO	Trung Bình
Số lượng ATNĐ	1,4±1,2	0,7±0,7	0,7±0,8	1,0±1,0
Số lượng bão (TS/TY)	8,7±2,8	10,6±3,5	10,6±2,3	10,0±3,1
Số lượng XTNĐ	10,1±2,7	11,3±3,5	11,3±2,1	10,9±2,9
Số ngày hoạt động của XTNĐ	33,7±11,9	36,0±11,7	37,1±8,4	35,6±10,8
Số lượng XTNĐ ba tháng 6-7- 8	3,7±1,2	3,8±1,8	4,6±1,6	4,0±1,7
Số lượng XTNĐ ba tháng 9-10-11	4,0±1,7	5,9±1,8	4,8±1,9	4,9±1,9
Số ngày XTNĐ hoạt động ba tháng 6-7-8	13,4±5,3	12,5±6,7	15,1±5,0	13,7±5,8
Số ngày XTNĐ hoạt động ba tháng 9-10-11	15,1±7,4	19,6±7,0	16,0±6,9	16,9±7,2

### 3.4 Quan hệ giữa số lượng và số ngày hoạt động của XTNĐ trên Biển Đông với các chỉ số khí hậu

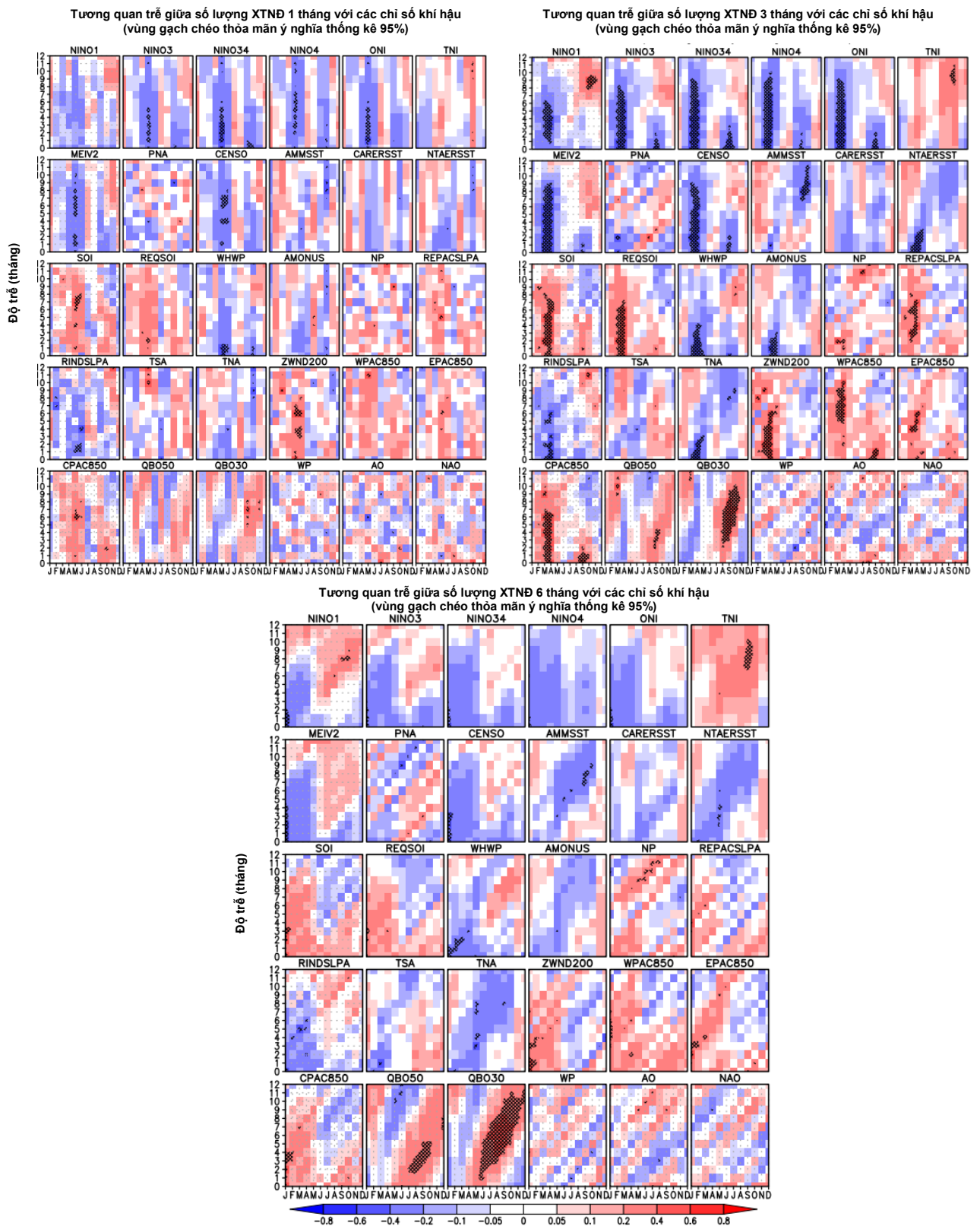
Theo cách phân loại, sắp xếp do NOAA công bố cho thấy các chỉ số khí hậu có thể được phân chia thành 3 nhóm: nhóm 1 gồm các chỉ số có liên quan đến ENSO (Niño1+2, Niño3, Niño3.4, Niño4, MEIV2, CENSO, ONI); nhóm 2 là các chỉ số đại dương (TNI, WHWP, REPAC, RINDO, TNA, TSA, AMONUS, AMM, NTA, CAR); nhóm 3 gồm các chỉ số khí quyển (QBO30, QBO50, SOI, REQSUI, ZWND20, WPAC85, EPAC85, CPAC85, PNA, WP, NAO, NP, AO) (Bảng 2.2) [108, 109].

Có thể thấy rằng, ENSO là một trong những dao động quan trọng chi phối thời tiết khu vực nhiệt đới, cận nhiệt đới và hoạt động của XTNĐ trên khu vực TBTBD và Biển Đông. Các minh chứng về ảnh hưởng của ENSO đến XTNĐ tương đối rõ ràng và đã được đề cập nhiều trong các công trình nghiên cứu trước đây. Vì vậy, nhóm các chỉ số liên quan đến ENSO kỳ vọng sẽ đóng góp thông tin cho việc xây dựng các

phương trình dự báo. Chỉ số Niño3.4 được sử dụng để đánh giá dao động nhiệt độ bề mặt biển ở khu vực nhiệt đới trung tâm Thái Bình Dương, dùng để biểu diễn ENSO truyền thống [82]. Gần đây, Xu và Huang (2015) đã chứng minh được ảnh hưởng khác nhau của sự ấm lên ở phía đông (Niño3) và ấm lên ở trung tâm Thái Bình Dương (Niño4) đã có ảnh hưởng khác biệt tới hoạt động của XTNĐ [95]. Bên cạnh các khái niệm ENSO truyền thống, khái niệm El Niño, Modoki cũng được nhắc đến trong vòng hơn một thập kỷ gần đây. El Niño, Modoki được xác định bởi sử dụng cả chuẩn sai nhiệt độ bề mặt biển ở khu vực phía đông ( $110^{\circ}\text{W} - 70^{\circ}\text{W}$ ,  $15^{\circ}\text{S} - 5^{\circ}\text{N}$ ), phía tây ( $125 - 145^{\circ}\text{E}$ ,  $10^{\circ}\text{S} - 20^{\circ}\text{N}$ ) và trung tâm ( $165^{\circ}\text{E} - 140^{\circ}\text{W}$ ,  $10^{\circ}\text{S} - 10^{\circ}\text{N}$ ) của biển Thái Bình Dương [70, 73]. Nhóm các chỉ số đại dương như: AMM, CAR\_SST, TNA, TSA, WHWP, ... Zhang và cộng sự (2017) đã chỉ ra pha dương (âm) của AMM ngăn cản (kích thích) hoạt động XTNĐ TBTBD [98]. Gao và cộng sự (2018) cũng đưa ra được các bằng chứng cho thấy TNA có tương quan âm với số XTNĐ đổ bộ vào khu vực Việt Nam, Trung Quốc, Bán đảo Đông Dương, Nhật Bản [38]. Nhóm những chỉ số khí quyển như NAO, WP, PNA, ... cũng được Zhou và Cui (2014) đề cập đến quan hệ niên thập kỷ giữa NAO và XTNĐ TBTBD [102]. WP là dao động ở khu vực Tây Thái Bình Dương, có liên quan đến rãnh Đông Á – là một bộ phận của gió mùa Đông Á. Wang Xin và cộng sự (2012), Wang Lei và cộng sự (2016) đã kiểm tra ảnh hưởng của gió mùa Đông Á đến hoạt động của XTNĐ Biển Đông [85, 89]. Kết quả chỉ ra rằng, dòng xiết Đông Á và cao cận nhiệt đới có ảnh hưởng đến dao động thập kỷ của XTNĐ Biển Đông... Như vậy có thể thấy rằng, có rất nhiều các chỉ số khí hậu có liên hệ và thậm chí tác động đến hoạt động của XTNĐ Biển Đông, tuy nhiên thực tế là còn có rất nhiều quan hệ khác chưa được đề cập tới cho tới nay. Với hy vọng có thể tìm ra những liên quan nhất định, luận án tiến hành tính toán hệ số tương quan giữa tổng số lượng và số ngày hoạt động của XTNĐ trên Biển Đông hàng tháng, tích lũy 3 và 6 tháng liên tiếp với bộ 30 chỉ số khí hậu. Các bản đồ tương quan giữa 30 chỉ số khí hậu với số lượng, số ngày XTNĐ hoạt động trên Biển Đông trong thời gian 1, 3 và 6 tháng giai đoạn 1979-2010 lần lượt được trình bày tại các Hình 3.10 và 3.11 dưới đây.

Kết quả trong Hình 3.10, 3.11 chỉ ra giá trị của hệ số tương quan thể hiện được mức ý nghĩa thống kê 95% tập trung nhiều vào nhóm các chỉ số như: Niño1+2, Niño3, Niño4, Niño3.4, ONI, SOI, NTASST, AMMSST, CARERSST, TNA, TSA, WHWP. Trong đó, giá trị nổi bật ở nhóm các chỉ số liên quan đến ENSO (Các chỉ số Niño, SOI, ONI) các chỉ số Đại dương liên quan đến SST như NTASST, WHWP, CARERSST, AMMSST, AMONUS hay các chỉ số khí quyển như QBO30, SOI, REQSUI, ZWND20, WPAC85, CPAC85). Về giá trị, nhóm chỉ số liên quan đến ENSO (Niño1+2, Niño3, Niño3.4, Niño4, ONI) cho giá trị tương quan âm trong khi các nhóm còn lại có trị số dương. Điều này nghĩa là số lượng XTNĐ hoạt động trên khu vực Biển Đông có xu hướng tăng (giảm) trong những năm La Niña (El Niño). Đặc biệt, giá trị tương quan nổi bật vào thời điểm mùa thu (tháng 9 đến tháng 11). Quan hệ giữa XTNĐ hoạt động ở TBTBD (bao gồm cả Biển Đông) và ENSO chỉ khác biệt trong mùa thu. Trong khi đó, các hệ số tương quan thoả mãn ý nghĩa thống kê cũng xuất hiện vào mùa tháng 3, 4. Tương quan thoả mãn ý nghĩa thống kê cũng xuất hiện với độ trễ từ 3 đến 6 tháng với nhóm các chỉ số liên quan đến ENSO. Các kết quả và quan hệ trễ này cũng đã được đề cập đến trong các công trình nghiên cứu trước đây [38, 93, 99]. Với những năm sau năm El Niño, mùa bão thường đến muộn. Ví dụ, năm 1998 và năm 2016 là những năm sau năm El Niño mạnh điển hình 1997 và 2015, mùa bão đến muộn hơn nhiều so với thường lệ. Điều này cũng nghĩa là, số lượng XTNĐ hoạt động trong mùa tháng 1 đến tháng 6 của các năm sau năm El Niño giảm đi so với bình thường. Một số chỉ số khí quyển khác như SOI, REQSUI, ZWND200 (QBO) cũng biểu thị tương quan dương trong thời gian đầu mùa bão (cuối mùa bão).

Các kết quả trên Hình 3.10 và 3.11 cũng cho thấy xuất hiện nhiều vùng tương quan đạt mức ý nghĩa thống kê 95% (vùng màu đỏ đậm hoặc xanh đậm, gạch chéo) ở XTNĐ 3 tháng, 6 tháng hơn so với XTNĐ 1 tháng. Các kết quả tính toán cũng cho thấy quan hệ giữa XTNĐ 3 tháng và 6 tháng với các chỉ số khí hậu nhận đạt được mức ý nghĩa thống kê ổn định hơn so với XTNĐ 1 tháng.



Hình 3.10 Tương quan giữa 30 chỉ số khí hậu với số lượng XTND 1, 3 và 6 tháng hoạt động trên Biển Đông giai đoạn 1979-2010



Cụ thể hơn, với số lượng XTNĐ 3 tháng, bản đồ tương quan Hình 3.10 cho thấy tồn tại nhiều chỉ số khí hậu có tương quan đạt mức ý nghĩa thống kê, điển hình như ở tháng 4-5 và tháng 9-10 đối với các chỉ số thuộc nhóm ENSO, một vài chỉ số thuộc nhóm hoàn lưu khí quyển như QBO, SOI hay tín phong CPAC850, WPAC850. Với XTNĐ 1 tháng, vùng tương quan trên mức 0,2 (về độ lớn tuyệt đối) suy giảm hơn so với XTNĐ 3 tháng, trong khi ở XTNĐ 6 tháng các vùng biểu thị tương quan cao được mở rộng ra hơn trong thời gian từ tháng 5 đến tháng 10. Tuy vậy, các kết quả tại Hình 3.10 cũng cho thấy tồn tại nhiều vùng tương quan của số lượng XTNĐ 6 tháng và 1 tháng không đảm bảo ý nghĩa thống kê. Diễn biến về tương quan giữa số lượng XTNĐ Biển Đông với các chỉ số khí hậu khá tương đồng so với tương quan giữa số ngày XTNĐ Biển Đông hoạt động và các chỉ số khí hậu (Hình 3.11). Tuy nhiên ở đặc trưng này, số ngày hoạt động của XTNĐ 6 tháng thể hiện mức độ tương quan tốt hơn so với 3 tháng trong khi số ngày XTNĐ 1 tháng cho kết quả tương quan kém hơn cả.

Kết quả trên bản đồ tương quan 3 tháng tại Hình 3.10 và 3.11 cho thấy, ở cả đặc trưng số lượng và số ngày hoạt động, 30 chỉ số khí hậu khảo sát đã thể hiện quan hệ với XTNĐ Biển Đông ở các thời điểm (độ trễ) theo các mức độ tương quan khác nhau. Điều này cho thấy tiềm năng của các chỉ số khí hậu để dự báo số lượng XTNĐ hoạt động trong mùa 3 đến 6 tháng với thời hạn dự báo trước từ 3 tới 12 tháng. Xuất phát từ đặc điểm này, bộ 30 chỉ số khí hậu khảo sát ban đầu được luận án sử dụng để tuyển chọn nhân tố cho bài toán dự báo số lượng XTNĐ Biển Đông như được trình bày tại Chương 4 dưới đây.

#### **Nhận xét chung:**

Qua nghiên cứu, phân tích trên bộ số liệu RSMC về đặc điểm hoạt động của XTNĐ trên khu vực Biển Đông cho thấy:

- Giai đoạn 1971-2019 có trung bình khoảng  $10,9 \pm 2,9$  XTNĐ hoạt động trong khoảng 35,5 ngày (142 obs 6 giờ), trong đó XTNĐ nhiều nhất (ít nhất) ở những năm La Niña (El Niño) với số lượng khoảng 11,3 cơn (10,1 cơn). Tính trung bình, ở những năm El Niño (La Niña) thời gian XTNĐ hoạt động trên khu vực Biển

Đông là 33,7 ngày (36 ngày) trong khi ở những năm Trung tính là 37,1 ngày. Nhìn chung số lượng XTNĐ trên Biển Đông trong những năm La Niña nhiều hơn trong những năm El Niño khoảng gần 2 cơn/năm và sự chênh lệch này thường tập trung vào 3 tháng cuối mùa bão (tháng 9, 10 và 11).

- Mùa bão trên Biển Đông kéo dài từ tháng 6-11 nhưng chủ yếu hoạt động mạnh ở thời gian các tháng 7-10. Số lượng XTNĐ Biển Đông giai đoạn 1971-2019 có sự biến động mạnh qua các năm và có dấu hiệu về xu thế giảm thoả mãn mức ý nghĩa 10% ở tổng số lượng XTNĐ trong mùa bão (tháng 6 tới tháng 11) và 3 tháng cuối mùa bão (các tháng 9, 10, 11) tương ứng với mức giảm khoảng 0,5 cơn/thập kỷ ( $P_{val} = 0,06$ ) và 0,3 cơn/thập kỷ ( $P_{val} = 0,07$ ).

- Ở giai đoạn 1971-2019, XTNĐ Biển Đông hoạt động nhiều nhất ở khu vực đông bắc, trung bình khoảng 2 cơn/năm. Vùng và số lượng XTNĐ có sự biến động đáng kể qua từng thập kỷ. Những năm 1971-1980 XTNĐ tập trung nhiều ở khu vực đông bắc Biển Đông nhưng đến thập kỷ 1981-1990 vùng nhiều XTNĐ đã dịch vào trung tâm Biển Đông. Thập kỷ 1991-2000 vùng nhiều XTNĐ lại dịch lên phía đông bắc và có xu hướng giảm so với trung bình giai đoạn khảo sát. Các năm từ 2001-2019 vùng XTNĐ hoạt động mở rộng xuống phía nam Biển Đông với mật độ dần trải hơn.

- Phân tích thành phần chính trên bộ dữ liệu giai đoạn 1971-2019 đưa ra gợi ý rằng các dạng đường đi của XTNĐ trên khu vực Biển Đông tập trung chủ yếu theo các hướng bắc tây bắc, tây bắc đi vào Vịnh Bắc Bộ; hướng tây, tây tây nam đi xuống Nam Bộ và một số XTNĐ hình thành, di chuyển từ ngoài khu vực TBTBD theo hướng bắc tây bắc và bắc đi vào Trung Quốc, Đài Loan; và hướng bắc, bắc đông bắc đi vào khu vực biển Nhật Bản. Kết quả phân tích theo thời gian cho thấy ở những thời điểm thập niên 80 XTNĐ phổ biến theo hướng tây bắc hướng vào khu vực Vịnh Bắc Bộ, tuy nhiên ở thập niên 90 và 2000 dạng quỹ đạo thịnh hành theo hướng bắc tây bắc, bắc đi vào Trung Quốc và bắc đông bắc đi vào khu vực biển Nhật Bản, Hàn Quốc. Từ thập niên 80 trở về sau tồn tại dạng quỹ đạo theo hướng tây, tây nam đi vào khu vực Nam Bộ của Việt Nam phù hợp với những nghiên cứu

gần đây về xu hướng dịch chuyển xuống phía nam của các XTNĐ trên khu vực Biển Đông.

Bên cạnh đó, qua việc khảo sát tương quan giữa 30 chỉ số khí hậu với số lượng và số ngày hoạt động của XTNĐ Biển Đông ở qui mô thời gian tháng đến mùa (1, 3 và 6 tháng) cho thấy tiềm năng trong việc tuyển chọn nhân tố dự báo từ bộ chỉ số khí hậu trên cho bài toán dự báo số lượng XTNĐ Biển Đông. Bên cạnh đó các kết quả này cũng gợi ý về việc khai thác các chỉ số khí hậu tiềm năng cho dự báo mùa XTNĐ hạn 3 và 6 tháng tốt hơn 1 tháng.

Từ các kết quả trên, định hướng luận án tiến hành khảo sát, đánh giá bộ chỉ số khí hậu để lựa chọn các nhân tố dự báo cho bài toán dự báo hạn mùa số lượng XTNĐ hoạt động trong 3 tháng đầu mùa bão (tháng 6-7-8) và 3 tháng cuối mùa bão (tháng 9-10-11) trên Biển Đông. Bên cạnh đó, các kết quả nghiên cứu về đặc điểm khí hậu của XTNĐ Biển Đông được đề cập ở trên cùng với các chỉ số đánh giá chất lượng, kỹ năng dự báo (như đã trình bày ở Chương 2) cũng sẽ góp phần làm rõ hơn kết quả và chất lượng dự báo của các mô hình khảo sát. Các nội dung này sẽ lần lượt được trình bày tại Chương 4 dưới đây.

## CHƯƠNG 4. DỰ BÁO HẠN MÙA SỐ LƯỢNG XTNĐ

### TRÊN KHU VỰC BIỂN ĐÔNG

Trên cơ sở các phương pháp nghiên cứu, nguồn số liệu và các phương án thiết kế thí nghiệm đã được trình bày trong Chương 2, Chương này sẽ trình bày những kết quả nghiên cứu về khả năng sử dụng mô hình thống kê để dự báo hạn mùa số lượng XTNĐ hoạt động trên khu vực Biển Đông. Để có cơ sở trong việc lựa chọn nhân tố dự báo cho các mô hình thống kê, luận án đã tiến hành khảo sát với 3 trường hợp. Trường hợp 1 và 2 (TH1 và TH2) tiếp cận theo phương pháp thống kê truyền thống với các nhân tố dự báo là các chỉ số khí hậu, trong đó TH1 và TH2 khác nhau ở cách xác định nhân tố dự báo, giữa việc xác định nhân tố dự báo theo phương pháp phân tích thành phần chính bộ số liệu các chỉ số khí hậu (TH1) và phương pháp hồi quy từng bước (TH2). Trường hợp 3 (TH3) được tiếp cận theo phương pháp thống kê trên các sản phẩm động lực, với các nhân tố dự báo được lựa chọn từ sản phẩm của mô hình khí hậu CFSv2, theo đó nhân tố dự tuyển được xác định qua giá trị trung bình vùng trên bản đồ tương quan giữa các biến trường sản phẩm CFSv2 với yếu tố dự báo (TH3). Với 3 cách tiếp cận trên, luận án tiến hành xây dựng các mô hình thống kê dự báo số lượng XTNĐ hoạt động trên khu vực Biển Đông trên bộ số liệu phục thuộc (giai đoạn 1980-2010 đối với TH1, TH2 và giai đoạn 1983-2009 đối với TH3) qua đó đánh giá khả năng dự báo số lượng XTNĐ trên khu vực nghiên cứu qua bộ số liệu độc lập (giai đoạn 2011-2017 cho TH1, TH2 và giai đoạn 2012-2016 đối với TH3). Việc đánh giá khả năng mô phỏng và dự báo hạn mùa số lượng XTNĐ của các mô hình được thực hiện thông qua đánh giá các sai số và chỉ số kỹ năng dự báo trên các chuỗi thời gian tương ứng. Ở đây, các sai số và chỉ số đánh giá về khả năng mô phỏng và dự báo hạn mùa số lượng XTNĐ hoạt động trên khu vực Biển Đông cho cả 3 trường hợp bao gồm: sai số trung bình (ME), sai số tuyệt đối trung bình (MAE), sai số bình phương trung bình (RMSE), tỷ lệ (%) giữa sai số tuyệt đối trung bình so với trung bình khí hậu (RMAE), hệ số tương quan giữa số lượng quan trắc và kết quả dự báo (Corr) và kỹ năng dự báo mô hình so với dự báo khí hậu (MSSS).

## **4.1 Dự báo hạn mùa số lượng XTNĐ trên Biển Đông bằng phương pháp thống kê truyền thống**

### **4.1.1 Tuyển chọn tập nhân tố dự báo**

Đối với bài toán dự báo thống kê nói chung và dự báo hạn mùa số lượng XTNĐ nói riêng thì việc phân tích, xác định và lựa chọn nhân tố dự báo có thể chia ra gồm 2 bước: 1) Xác định nhân tố dự tuyển thông qua việc phân tích quan hệ và ý nghĩa vật lý hoặc tính toán hệ số tương quan để định lượng mối quan hệ thống kê giữa yếu tố dự báo (số lượng XTNĐ) và nhân tố dự báo và 2) Xác định các nhân tố dự báo chính thức thông qua phương pháp thành phần chính để sử dụng một số thành phần chính của bộ nhân tố dự tuyển hoặc sàng lọc bộ nhân tố dự tuyển bằng kỹ thuật hồi quy từng bước.

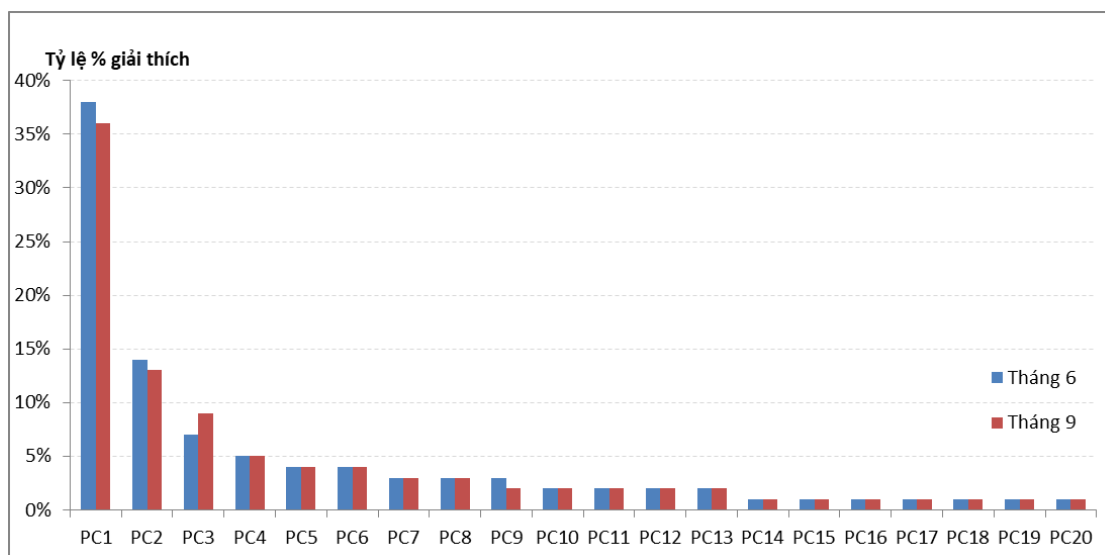
Trên cơ sở khảo sát mối quan hệ giữa các chỉ số khí hậu và số lượng XTNĐ 3 tháng với độ trễ từ 1 tháng tới 12 tháng như đã trình bày tại Chương 3 cho thấy, ở từng thời điểm khác nhau, bộ 30 chỉ số khí hậu khảo sát đã thể hiện tương quan tốt với XTNĐ 3 tháng tại khu vực Biển Đông. Bộ 30 chỉ số khí hậu này được xem xét theo các nhóm: 1) nhóm ENSO; 2) nhóm chỉ số đại dương và 3) nhóm chỉ số khí quyển (Bảng 2.2). Việc phân nhóm này chỉ mang tính chất tương đối, chẳng hạn như chỉ số dao động nam SOI, được tính từ chênh lệch khí áp giữa hai trạm Tahiti-Darwin vừa có thể xếp ở nhóm ENSO vừa có thể xếp vào nhóm hoàn lưu khí quyển, hay như các chỉ số Niño được tính trên dữ liệu về nhiệt độ mặt biển vừa có thể ở nhóm ENSO vừa có thể ở nhóm chỉ số đại dương...

Từ bộ 30 chỉ số khí hậu đã được khảo sát ở trên, bộ dữ liệu các nhân tố dự tuyển mới được mở rộng thêm với độ trễ từ tháng thứ 3 tới tháng thứ 12, tức bao gồm 300 nhân tố dự tuyển được lựa chọn tham gia khảo sát. Việc mở rộng này, một mặt vừa xét đến các mối quan hệ trễ giữa XTNĐ Biển Đông với các chỉ số khí hậu như đã khảo sát tương quan tại Chương 3, mặt khác cũng vừa để tăng dung lượng mẫu khảo sát. Trong khi đó, việc lựa chọn với độ trễ từ tháng thứ 3 tới tháng thứ 12 (bỏ qua trễ 2 tháng đầu) là đảm bảo chắc chắn hơn các chỉ số khí hậu được cập nhật và dữ liệu có sẵn tại các thời điểm dự báo (đảm bảo cho dự báo nghiệp vụ). Tương

ứng với từng tháng dự báo, bộ nhân tố dự báo ở TH1 được xác định là một số thành phần chính của bộ 300 nhân tố dự tuyển thông qua phương pháp thành phần chính. Trong khi bộ nhân tố dự báo ở TH2 được tuyển chọn từ 300 nhân tố dự tuyển qua kỹ thuật hồi quy từng bước như trình bày tại Chương 2.

\* Nhân tố dự báo cho TH1

Hình 4.1 dưới đây dẫn ra mức độ đóng góp/ giải thích thông tin của 20 thành phần chính đầu tiên (từ PC1 tới PC20) của 2 bộ dữ liệu các nhân tố dự tuyển cho các dự báo tại thời điểm tháng 6 (dự báo cho các XTND đầu mùa bão – tháng 6, 7, 8) và tháng 9 (dự báo cho các XTND cuối mùa bão – tháng 9, 10, 11). Kết quả này cho thấy, mức độ đóng góp giải thích thông tin của thành phần đầu tiên (PC1) ở cả 2 bộ dữ liệu nhân tố dự tuyển đều chiếm trên 35%, tiếp đến là thành phần thứ hai (PC2) chiếm trên 13% và giảm dần xuống còn trên dưới 2% ở thành phần chính thứ 10 (PC10), các thành phần chính tiếp sau suy giảm đến dưới 1% ở thành phần chính thứ 20 (PC20).



Hình 4.1 Mức độ (%) giải thích thông tin 20 thành phần chính (PC) đầu tiên của bộ dữ liệu nhân tố dự tuyển tháng 6 và tháng 9

Để có cơ sở cho việc xác định số thành phần chính cần thiết của bộ dữ liệu nhân tố dự tuyển được giữ lại, luận án tiến hành khảo sát 10 thí nghiệm như được mô tả tại bảng 4.1 dưới đây. Các thí nghiệm này tương ứng với sự tham gia lần lượt

của 1 thành phần chính đầu tiên (EOF1); 2 thành phần chính đầu tiên (EOF1, EOF2), 3 thành phần chính đầu tiên (EOF1, EOF2, EOF3), ..., cho tới 10 thành phần chính đầu tiên (EOF1, .. tới EOF10). Các thí nghiệm này lần lượt được khảo sát trên cùng bộ dữ liệu giai đoạn 1980-2010, với yếu tố dự báo là số lượng XTNĐ 3 tháng đầu mùa bão (các tháng 6-7-8 hay TC6-7-8) và 3 tháng cuối mùa bão (các tháng 9-10-11 hay TC9-10-11) thông qua 3 mô hình thống kê MLR, LMV và LAD như được trình bình ở Chương 2.

Bảng 4.1 Các thí nghiệm khảo sát sai số của việc lựa chọn số lượng thành phần chính làm nhân tố dự báo

<b>Thí nghiệm</b>	<b>Nhân tố dự báo</b>	<b>Yếu tố dự báo</b>	<b>Mô hình dự báo thống kê khảo sát</b>	<b>Sai số đánh giá</b>
1	EOF1	TC6-7-8; TC9-10-11	MLR, LMV và LAD	MAE
2	EOF1, EOF2			
3	EOF1, EOF2, EOF3			
4	EOF1, ... , EOF4			
5	EOF1, ... , EOF5			
6	EOF1, ... , EOF6			
7	EOF1, ... , EOF7			
8	EOF1, ... , EOF8			
9	EOF1, ... , EOF9			
10	EOF1, ... , EOF10			

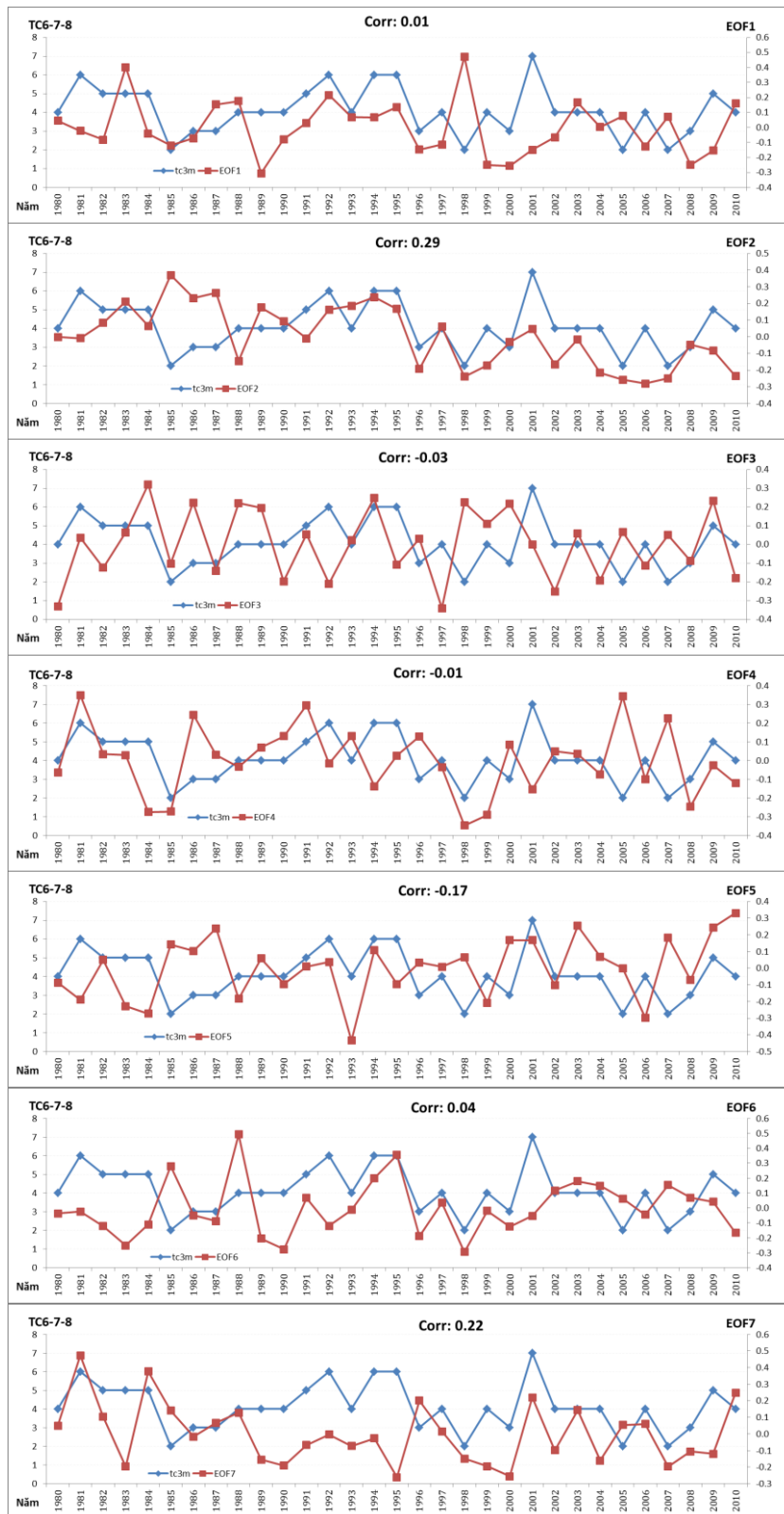
Sai số MAE của 3 mô hình MLR, LAD và LMV theo 10 thí nghiệm trên cho các thời điểm dự báo tháng 6 và tháng 9 được dẫn ra tại Bảng 4.2 dưới đây. Các kết quả này cho thấy, ở cùng thí nghiệm (cùng số lượng thành phần chính) sai số MAE không có sự khác biệt lớn giữa các mô hình tại cùng thời điểm dự báo tháng 6 hoặc tháng 9, trong đó sai số ở thời điểm dự báo tháng 9 có độ lớn cao hơn. Sai số MAE có xu hướng suy giảm tỷ lệ nghịch với số lượng nhân tố dự báo được sử dụng (số lượng thành phần chính). Giá trị trung bình sai số MAE của dự báo tại thời điểm tháng 6 và

tháng 9 cũng phản ánh đặc điểm này, MAE trung bình giảm dần từ 1,23 khi chỉ sử dụng 1 thành phần chính đầu tiên tới 1,03 khi có sự tham gia đóng góp của 7 thành phần chính đầu tiên (EOF1, ..., EOF7). Kết quả tại Bảng 4.2 cũng cho thấy sự có mặt của các thành phần chính thứ 8 trở đi (EOF8, EOF9, EOF10, ...) đã không đóng góp thêm cho việc giảm sai số của các mô hình. Như vậy, 7 thành phần chính đầu tiên của bộ dữ liệu dự tuyển, tương ứng với giải thích 75% (74%) thông tin của bộ dữ liệu các nhân tố dự tuyển cho các dự báo tại thời điểm tháng 6 (tháng 9) có thể xem xét là số lượng nhân tố “đủ” cho các mô hình dự báo thống kê khảo sát.

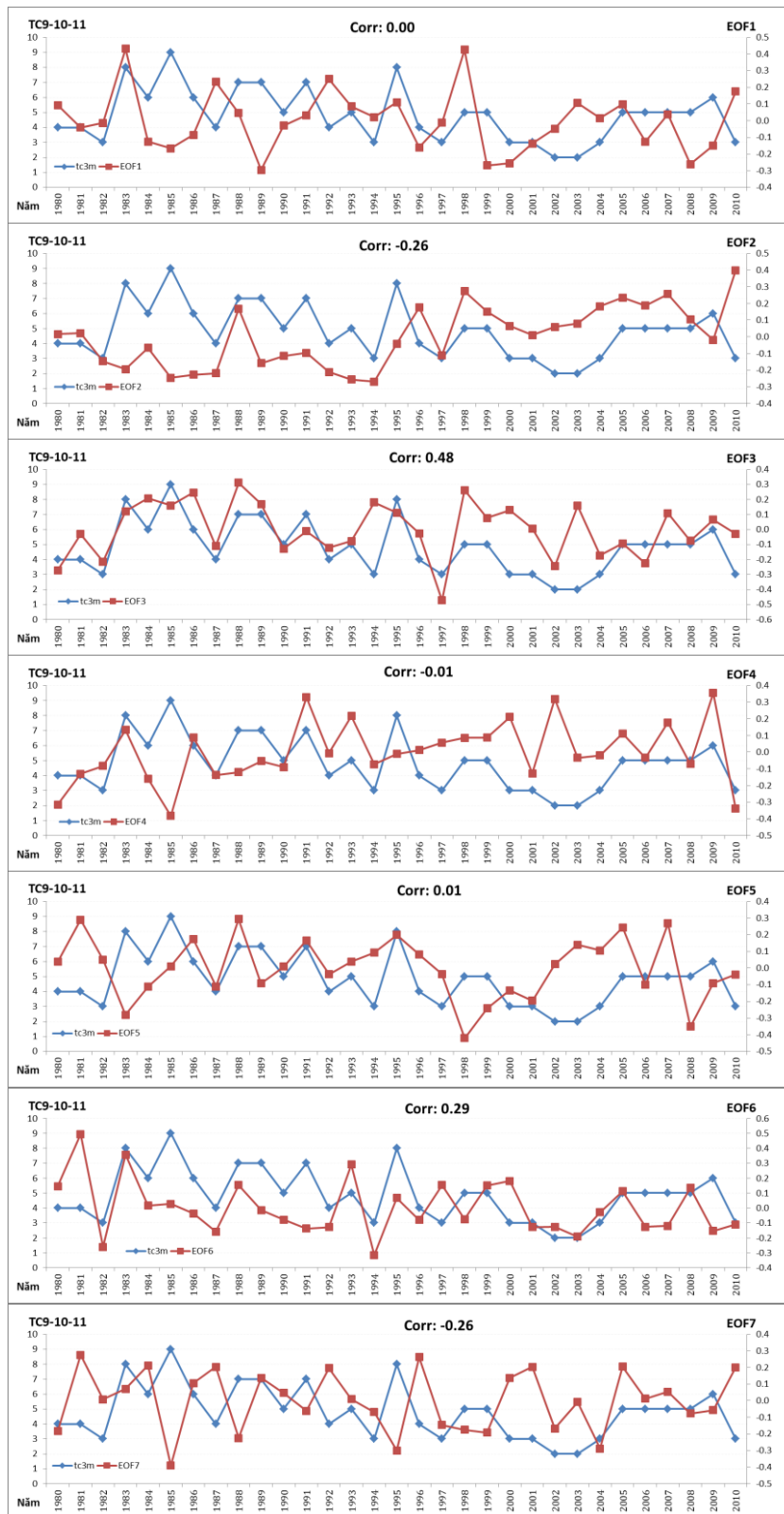
Bảng 4.2 Sai số MAE của ba mô hình thống kê theo số lượng thành phần chính, dự báo cho thời điểm tháng 6 và tháng 9

Số thành phần chính làm nhân tố dự báo	Tháng 6			Tháng 9			Trung bình
	MLR	LAD	LMV	MLR	LAD	LMV	
1	1.0	0.9	1.1	1.4	1.4	1.6	1.23
2	1.0	0.9	1.2	1.5	1.4	1.5	1.25
3	1.0	0.9	1.2	1.2	1.1	1.4	1.13
3	1.0	0.9	1.3	1.2	1.1	1.8	1.22
4	0.9	0.9	1.3	1.2	1.1	1.8	1.20
5	0.9	0.9	1.3	1.1	1.0	1.4	1.10
6	0.9	0.9	1.3	1.1	1.0	1.3	1.08
7	0.9	0.9	1.2	1.0	0.9	1.3	<b>1.03</b>
8	0.9	0.9	1.1	1.0	0.9	1.4	1.03
9	0.9	0.9	1.2	1.0	0.9	1.3	1.03
10	0.9	0.9	1.2	1.0	0.9	1.3	1.03

Khảo sát tương quan giữa 7 EOF tương ứng với 7 thành phần chính đầu tiên đã được khảo sát ở trên (từ PC1 tới PC7) của 2 bộ dữ liệu nhân tố dự tuyển tháng 6 (và tháng 9) với số lượng XTND Biển Đông hoạt động trong 3 tháng 6-7-8 (và 3 tháng 9-10-11) như trình bày tại Hình 4.2 (Hình 4.3) dưới đây. Các kết quả tại Hình 4.2 và Hình 4.3 chỉ ra, ở cả hai thời điểm dự báo đều tồn tại các EOF có quan hệ tương quan với XTND Biển Đông song mức độ tương quan nhìn chung yếu. Ở bộ dữ liệu nhân tố dự tuyển tháng 6 có 2 thành phần EOF có độ lớn tương quan trên 0,2 trong khi số lượng này ở bộ dữ liệu tháng 9 là 4 thành phần EOF, trong đó cá biệt có thành phần EOF thứ 3 có giá trị tương quan tới 0,48.



Hình 4.2 Mối quan hệ giữa 7 EOF đầu tiên của bộ dữ liệu nhân tố dự tuyến tháng 6 với XTND Biên Đông hoạt động trong 3 tháng 6-7-8



Hình 4.3 Mối quan hệ giữa 7 EOF đầu tiên của bộ dữ liệu nhân tố dự tuyển tháng 9 với XTND Biển Đông hoạt động trong 3 tháng 9-10-11

Bên cạnh đó ở cả 2 thời điểm dự báo, các kết quả tại Hình 4.2 và Hình 4.3 cũng chỉ ra tồn tại các EOF có mức độ tương quan rất yếu với XTNĐ Biển Đông, như thành phần EOF1, EOF3, EOF4, EOF6 ở tháng 6 và EOF1, EOF4, EOF5 ở tháng 9. Đặc điểm này có thể ảnh hưởng tiêu cực tới kết quả dự báo và chất lượng dự báo ở thời điểm tháng 6 kém hơn so với tháng 9 (các kết quả này sẽ được trình bày ở phần dưới đây).

\* Khảo sát tham số cho mô hình ANN

Như đã trình bày tại Chương 2, ngoài 3 mô hình dự báo thống kê truyền thống (MLR, LAD và LMV) luận án còn xem xét khảo sát với mô hình mạng thần kinh nhân tạo ANN. Về bản chất, khác với 3 mô hình dự báo thống kê trước, ANN là một mô hình phi tuyến [97]. Một trong các vấn đề gặp phải khi sử dụng ANN cho bài toán dự báo đó là hiện tượng quá khớp (overfitting), đây là hiện tượng mô hình đưa ra các giá trị quá khớp với dữ liệu phụ thuộc (dữ liệu luyện) trong khi lại gặp sai số ở bộ dữ liệu độc lập, đặc biệt hơn khi tập dữ liệu phụ thuộc tồn tại biến động (nhiều) lớn [2, 6, 47, 97]. Để giải quyết vấn đề này, ở đây luận án tiến hành khảo sát sai số MAE của mô hình ANN với cấu hình tham số lớp ẩn thay đổi từ 1 tới 10, tốc độ luyện từ 0,1 tới 1,0 và 7 neuron (bằng với số lượng thành phần chính được giữ lại như đã khảo sát ở trên) để dự báo số lượng XTNĐ ba tháng đầu mùa bão (TC6-7-8) và 3 tháng cuối mùa bão (TC9-10-11) cho giai đoạn độc lập 2011-2017 dựa trên quá trình học của tập dữ liệu luyện giai đoạn 1980-2010. Tập các sai số trung bình cho hai thời điểm dự báo này, tương ứng với giai đoạn luyện và giai đoạn dự báo độc lập được trình bày tại các Bảng 4.3 và 4.4. Kết quả tại Bảng 4.3 cho thấy ở giai đoạn luyện khi số lượng lớp ẩn tăng từ 1 tới 10 lên hoặc tốc độ luyện tăng lên từ 0,1 tới 1,0 thì sai số MAE giảm từ 0,9 xuống 0, tốc độ giảm MAE được tăng cường khi có sự kết hợp sự gia tăng của số lớp ẩn hoặc tốc độ luyện hoặc đặc biệt là khi có sự kết hợp cả hai yếu tố này. Ở giai đoạn dự báo độc lập, kết quả tại Bảng 4.4 cho thấy sai số MAE dường như cũng theo quy luật này, song với mức biến động

nhỏ, sai số tăng lên nhẹ khi số lớp ẩn tăng lên kết hợp với sự gia tăng của tốc độ luyện, đặc biệt rõ nét ở khoảng tốc độ luyện từ 0,1 tới 0,5.

Bảng 4.3 Sai số MAE giai đoạn luyện của mô hình ANN theo các tham số tốc độ luyện và số lớp ẩn

Số lớp ẩn	Tốc độ luyện									
	rate=0.1	rate=0.2	rate=0.3	rate=0.4	rate=0.5	rate=0.6	rate=0.7	rate=0.8	rate=0.9	rate=0.10
1	0.9	<b>0.5</b>	0.3	0.3	0.1	0.1	0.1	0.1	0.1	0.1
2	0.9	0.4	0.3	0.1	0.1	0	0	0	0	0
3	0.7	0.4	0.2	0.1	0	0	0	0	0	0
4	0.7	0.4	0.2	0	0	0	0	0	0	0
5	0.7	0.4	0.2	0	0	0	0	0	0	0
6	0.7	0.4	0.2	0	0	0	0	0	0	0
7	0.7	0.4	0.3	0	0	0	0	0	0	0
8	0.7	0.4	0.3	0.2	0	0	0	0	0	0
9	0.7	0.5	0.3	0	0	0	0	0	0	0
10	0.7	0.5	0.3	0	0	0	0	0	0	0

Bảng 4.4 Sai số MAE giai đoạn dự báo kiểm nghiệm của mô hình ANN theo các tham số tốc độ luyện và số lớp ẩn

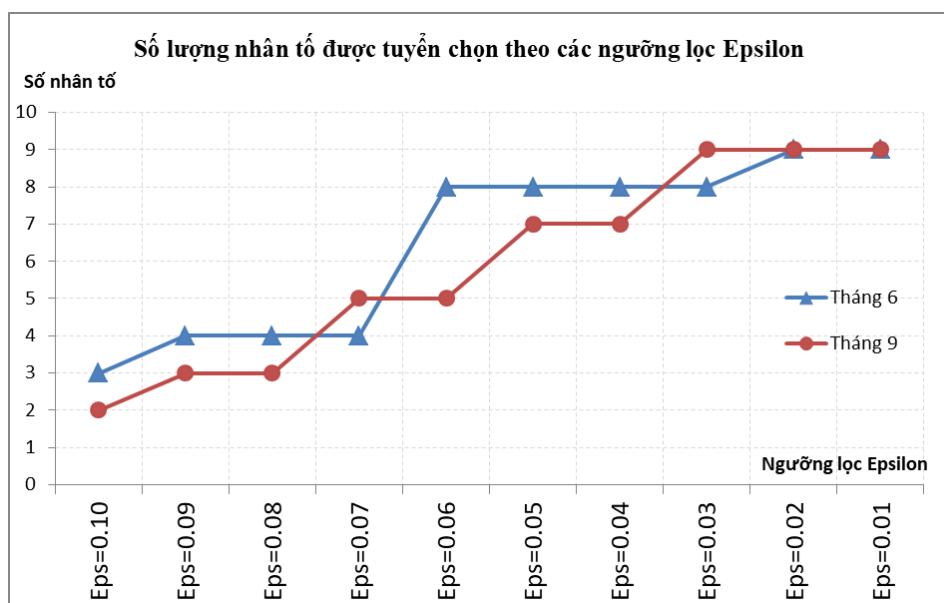
Số lớp ẩn	Tốc độ luyện									
	rate=0.1	rate=0.2	rate=0.3	rate=0.4	rate=0.5	rate=0.6	rate=0.7	rate=0.8	rate=0.9	rate=0.10
1	2.1	<b>1.9</b>	2.0	2.2	2.1	2.0	2.3	2.1	2.1	2.0
2	2.1	2.1	2.2	2.1	2.1	2.0	2.3	2.1	2.1	2.0
3	2.3	2.1	2.3	2.1	2.1	2.0	2.4	2.1	2.1	2.0
4	2.3	2.2	2.3	2.2	2.2	2.0	2.2	2.1	2.1	2.1
5	2.3	2.2	2.4	2.2	2.2	1.9	2.2	2.0	2.1	2.1
6	2.3	2.2	2.4	2.3	2.2	1.9	2.1	2.0	2.1	2.2
7	2.4	2.2	2.4	2.4	2.2	1.9	2.0	2.0	2.1	2.1
8	2.4	2.3	2.3	2.4	2.1	2.2	2.0	2.0	2.2	2.0
9	2.4	2.3	2.3	2.4	2.3	2.1	1.9	1.9	2.2	1.9
10	2.4	2.3	2.3	2.4	1.9	2.4	1.8	1.9	2.3	2.0

Trên thực tế, một mô hình dự báo được cho là tốt khi cả sai số dự báo phụ thuộc và dự báo độc lập đều thấp, mô hình sai hoặc kém ổn định khi sai số phụ thuộc thấp, sai số độc lập cao hoặc sai số độc lập thấp, sai số phụ thuộc cao. Như vậy, qua các khảo sát sai số MAE như trình bày ở trên cho thấy, đối với việc sử dụng ANN cho bài toán dự báo XTND 3 tháng, số lớp ẩn và tốc độ luyện thường được chọn ở các giá trị thấp. Việc lựa chọn một lớp ẩn ở mô hình ANN cũng đã

được một số tác giả đề cập và lựa chọn cho các yếu tố dự báo khác nhau, chẳng hạn như chỉ số nén của đất, độ rỗng hay lưu lượng dòng chảy [2, 6, 15]. Kết quả tại Bảng 4.3 và 4.4 cho thấy, cặp tham số 1 lớp ẩn kết hợp với tốc độ luyện 0,2 cho sai số trung bình ở dự báo phụ thuộc (độc lập) là 0,5 (1,9) và tổng thể chung đây có thể được xem là lựa chọn phù hợp trong số các bộ tham số đã khảo sát cho ANN cho bài toán dự báo số lượng XTNĐ 3 tháng Biển Đông.

**\* Nhân tố dự báo cho TH2**

Theo cách phân chia như đã trình bày bên trên, nhân tố dự báo cho TH2 được tuyển chọn từ bộ dữ liệu 300 nhân tố dự tuyển bằng kỹ thuật hồi quy từng bước. Hình 4.4 dẫn ra số lượng các chỉ số khí hậu được tuyển chọn làm nhân tố dự báo tại tháng 6 và 9 qua kỹ thuật lọc này. Kết quả tại Hình 4.4 cho thấy, ở các tháng dự báo khác nhau, số lượng các nhân tố dự báo có sự khác nhau nhất định tuy nhiên đều cùng phản ánh quy luật: số lượng nhân tố dự báo được lựa chọn tỷ lệ nghịch với ngưỡng lọc Epsilon theo kỹ thuật hồi quy từng bước. Số lượng 2 (3) nhân tố dự báo ở tháng 6 (tháng 9) tại ngưỡng Epsilon 0,1 biến đổi tăng dần lên tới 9 nhân tố dự báo ở ngưỡng Epsilon 0,01 và 0,02 ở cả dự báo tại tháng 6 và tháng 9.



Hình 4.4 Số lượng các nhân tố dự báo được tuyển chọn bằng kỹ thuật hồi quy từng bước theo các ngưỡng Epsilon khác nhau

Tương ứng với số lượng các nhân tố dự báo được lựa chọn ở trên, Bảng 4.5 (Bảng 4.6) dưới đây dẫn ra cụ thể chỉ số khí hậu tại độ trễ xác định được kỹ thuật hồi quy từng bước lọc làm nhân tố dự báo tại tháng 6 (tháng 9).

Bảng 4.5 Độ trễ của các nhân tố dự báo tháng 6 được tuyển chọn theo các ngưỡng lọc Epsilon khác nhau trong hồi quy từng bước

TT	ClimIdx	Ngưỡng lọc									
		Eps=0.01	Eps=0.02	Eps=0.03	Eps=0.04	Eps=0.05	Eps=0.06	Eps=0.07	Eps=0.08	Eps=0.09	Eps=0.10
1	NINO1										
2	NINO3										
3	NINO34	12	12								
4	NINO4										
5	ONI										
6	TNI										
7	MEIV2	9	9	9	9	9	9	9	9	9	9
8	PNA										
9	CENSO	7	7	7	7	7	7				
10	AMMSST	8	8	8	8	8	8				
11	CARERSST	6	6	6	6	6	6				
12	NTAERSST										
13	SOI										
14	REQSOI										
15	WHWP										
16	AMONUS										
17	NP	12	12	12	12	12	12	12	12	12	
18	REPAC	8	8	8	8	8	8				
19	RINDO	9	9	9	9	9	9	9	9	9	9
20	TSA										
21	TNA										
22	ZWND20										
23	WPAC85										
24	EPAC85										
25	CPAC85										
26	QBO50										
27	QBO30										
28	WP										
29	AO	11	11	11	11	11	11	11	11	11	11
30	NAO										

Kết quả tuyển chọn nhân tố dự báo theo TH2 cho tháng 6 tại Bảng 4.5 cho thấy, với ngưỡng lọc Epsilon biến động từ 0,1 tới 1,0 các nhân tố dự tuyển được lựa chọn xuất hiện rải rác ở cả 3 nhóm ENSO, Khí quyển và đại dương, tuy nhiên số lượng thuộc nhóm ENSO và các chỉ số liên quan tới SST chiếm 5 trong số 9 chỉ số khí hậu được lựa chọn (bao gồm: Niño3.4; MEIV2; CENSO; AMM; CAR) trong khi nhóm khí quyển đóng góp 4 chỉ số (gồm: REPAC; RINDO; NP; AO). Xét theo quy mô không gian, ngoài 6 tín hiệu đến từ các chỉ số khí hậu liên quan tới khu vực Thái Bình Dương (Niño3.4; MEIV2; CENSO, REPAC; RINDO, NP) thì cũng tồn tại 3 chỉ số khí hậu có quan hệ xa với số lượng các XTNĐ Biển Đông hoạt động trong ba tháng đầu mùa (tháng 6, 7 và 8).

Các nhân tố dự báo cho thời điểm tháng 9 theo TH2 (Bảng 4.6) có sự khác nhau nhất định so với thời điểm dự báo tháng 6 (Bảng 4.5). Ở thời điểm dự báo tháng 9, nhóm ENSO và các chỉ số liên quan tới SST chiếm 2 trong số 9 chỉ số khí hậu được lựa chọn (bao gồm: CENSO; WHWP) trong khi nhóm khí quyển đóng góp tới 7 chỉ số (gồm: RINDO; QBO30, QBO50; ZWND200; WP; NP; AO). Kết quả này cho thấy, các tín hiệu liên quan tới SST ảnh hưởng tới số lượng XTNĐ Biển Đông hoạt động trong ba tháng cuối mùa bão (tháng 9, 10 và 11) ít hơn so với các chỉ số khí quyển được lựa chọn, điều này cũng có thể giải thích được bởi thực tế trong thời gian này, nhiệt độ mặt biển đã được tích lũy đủ cho điều kiện sinh XTNĐ tại khu vực TBTBD và Biển Đông. Theo quy mô không gian, các chỉ số khí hậu được lựa chọn cho thời điểm dự báo tháng 9 có 7 chỉ số có quan hệ gần và 2 chỉ số có quan hệ xa (WHWP, AO) với số lượng XTNĐ Biển Đông ba tháng cuối mùa bão.

Về độ trễ của các chỉ số khí hậu được lựa chọn, kết quả tại Bảng 4.5 và 4.6 cho thấy các khoảng thời gian trễ từ 6 tháng đến 1 năm có ý nghĩa hơn so với các chỉ số khí hậu trễ từ 2 tới 5 tháng ngoại trừ riêng chỉ số QBO50 ở bộ số liệu dự báo tháng 9 được chọn tại độ trễ 5 tháng.

Bảng 4.6. Độ trễ của các nhân tố dự báo tháng 9 được tuyển chọn theo các ngưỡng lọc Epsilon khác nhau trong hồi quy từng bước

TT	ClimIdx	Ngưỡng lọc									
		Eps=0.01	Eps=0.02	Eps=0.03	Eps=0.04	Eps=0.05	Eps=0.06	Eps=0.07	Eps=0.08	Eps=0.09	Eps=0.10
1	NINO1										
2	NINO3										
3	NINO34										
4	NINO4										
5	ONI										
6	TNI										
7	MEIV2										
8	PNA										
9	CENSO	7	7	7	7	7	7	7	7	7	7
10	AMMSST										
11	CARERSST										
12	NTAERSST										
13	SOI										
14	REQSOI										
15	WHWP	8	8	8	8	8	8	8			
16	AMONUS										
17	NP	7	7	7	7	7					
18	REPAC										
19	RINDO	7	7	7	7	7	7	7	7		
20	TSA										
21	TNA										
22	ZWND20	12	12	12	12	12					
23	WPAC85										
24	EPAC85										
25	CPAC85										
26	QBO50	5	5	5	5	5	5	5	5	5	5
27	QBO30	11	11	11							
28	WP	9	9	9	9	9	9	9			
29	AO	10	10	10							
30	NAO										

Như vậy, có thể thấy rằng ứng với mỗi ngưỡng lọc Epsilon khác nhau tại các thời điểm tháng 6 và 9 đều có thể lựa chọn được bộ nhân tố dự báo cho bài toán dự báo số lượng XTND Biển Đông. Với việc lựa chọn ngưỡng lọc Epsilon thấp (0,01 hay 0,02) ở đây có thể tìm thấy 9 nhân tố dự báo, song với ngưỡng lọc Epsilon tăng lên, số nhân tố dự báo lựa chọn giảm xuống, chẳng hạn giảm xuống chỉ còn 4

nhân tố (3 nhân tố) ở thời điểm dự báo tháng 6 (tháng 9) ở ngưỡng lọc Epsilon 0,08. Để có cơ sở hơn cho việc xác định số lượng nhân tố dự báo, ở đây luận án khảo sát sai số MAE của cả 4 mô hình MLR, LAD, LMV và ANN trên bộ số liệu các chỉ số khí hậu đã được tuyển chọn theo các ngưỡng lọc Epsilon trong khoảng từ 0,01 tới 0,10 như đã mô tả bên trên. Kết quả tổng hợp của sai số MAE cho các khảo sát này được dẫn ra tại Bảng 4.7 và 4.8 dưới đây.

Bảng 4.7 Sai số MAE theo các ngưỡng lọc Epsilon dự báo trên bộ số liệu phụ thuộc (1980-2010)

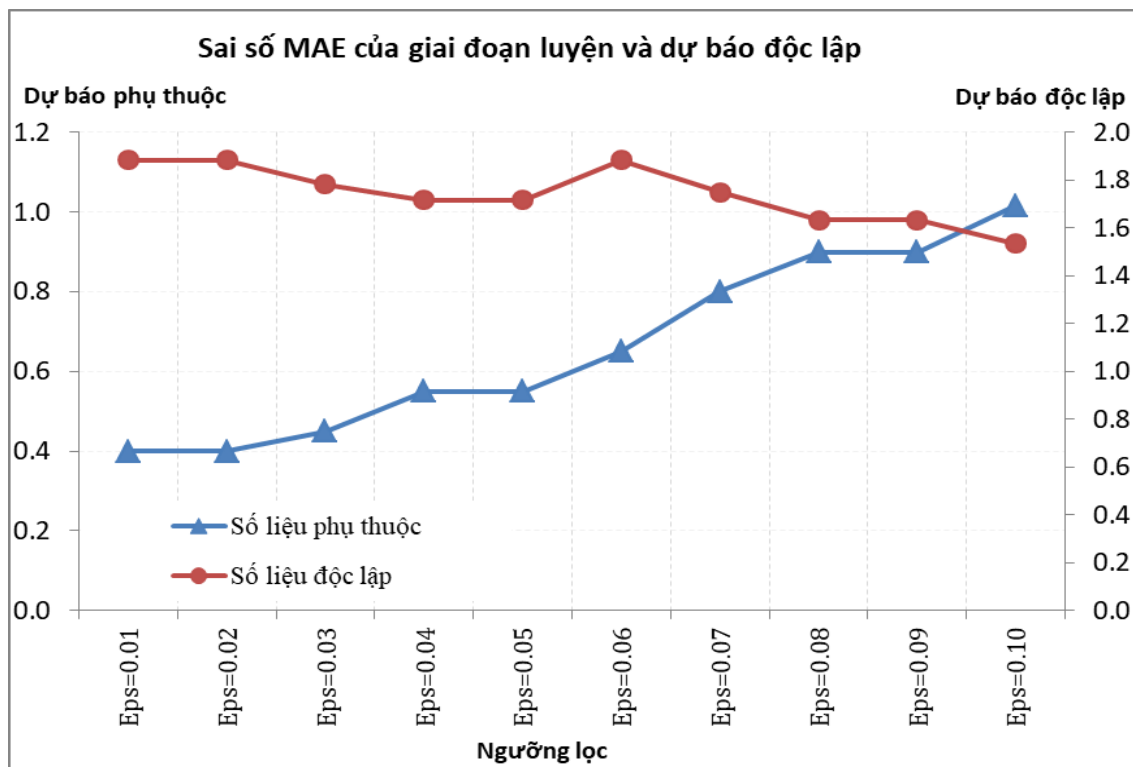
Ngưỡng lọc	Tháng 6			Tháng 9			Trung bình
	MLR	LAD	LMV	MLR	LAD	LMV	
Eps=0.01	0.3	0.3	0.4	0.4	0.4	0.6	0.40
Eps=0.02	0.3	0.3	0.4	0.4	0.4	0.6	0.40
Eps=0.03	0.4	0.4	0.5	0.4	0.4	0.6	0.45
Eps=0.04	0.4	0.4	0.5	0.6	0.6	0.8	0.55
Eps=0.05	0.4	0.4	0.5	0.6	0.6	0.8	0.55
Eps=0.06	0.4	0.4	0.5	0.7	0.7	1.2	0.65
Eps=0.07	0.7	0.7	0.8	0.7	0.7	1.2	0.80
Eps=0.08	0.7	0.7	0.8	1	0.9	1.3	0.90
Eps=0.09	0.7	0.7	0.8	1	0.9	1.3	0.90
Eps=0.10	0.9	0.8	0.9	1.2	1.1	1.2	1.02

Bảng 4.8 Sai số MAE theo các ngưỡng lọc Epsilon dự báo trên bộ số liệu độc lập (2011-2017)

Ngưỡng lọc	Tháng 6			Tháng 9			Trung bình
	MLR	LAD	LMV	MLR	LAD	LMV	
Eps=0.01	2.3	2.3	2.2	1.5	1.5	1.5	1.88
Eps=0.02	2.3	2.3	2.2	1.5	1.5	1.5	1.88
Eps=0.03	2.1	2	2.1	1.5	1.5	1.5	1.78
Eps=0.04	2.1	2	2.1	1.4	1.5	1.2	1.72
Eps=0.05	2.1	2	2.1	1.4	1.5	1.2	1.72
Eps=0.06	2.1	2	2.1	1.7	1.7	1.7	1.88
Eps=0.07	1.8	1.8	1.8	1.7	1.7	1.7	1.75
Eps=0.08	1.8	1.8	1.8	1.5	1.5	1.4	1.63
Eps=0.09	1.8	1.8	1.8	1.5	1.5	1.4	1.63
Eps=0.10	1.8	1.8	1.8	1.3	1.3	1.2	1.53

Kết quả dự báo cho giai đoạn dự báo phụ thuộc tại Bảng 4.7 cho thấy, xu thế gia tăng giá trị sai số MAE đồng thuận với sự gia tăng của tham số Epsilon, điều

đó đồng nghĩa với việc số lượng tham gia của các nhân tố dự báo càng nhiều thì sai số dự báo (trên bộ số liệu phụ thuộc) giảm xuống. Tuy nhiên, đặc điểm này lại thể hiện trái ngược trên giai đoạn dự báo độc lập, như được dẫn ra tại Bảng 4.8. Theo đó, sai số MAE có xu hướng suy giảm cùng với sự gia tăng của tham số Epsilon. Các kết quả này cho thấy, việc lựa chọn Epsilon nhỏ cho giai đoạn luyện sẽ cho kết quả dự báo phụ thuộc tốt tuy nhiên lại gặp phải sai số lớn các dự báo độc lập, hiện tượng này khá giống với hiện tượng “quá khớp” của mô hình ANN như đã khảo sát ở trên. Do vậy, để có được lựa chọn tối ưu giữa việc xác định ngưỡng lọc Epsilon, Hình 4.5 dưới đây dẫn ra kết quả tổng hợp sai số MAE trung bình trên giai đoạn dự báo phụ thuộc và giai đoạn dự báo độc lập đối với các kết quả dự báo tại thời điểm tháng 6 và tháng 9.



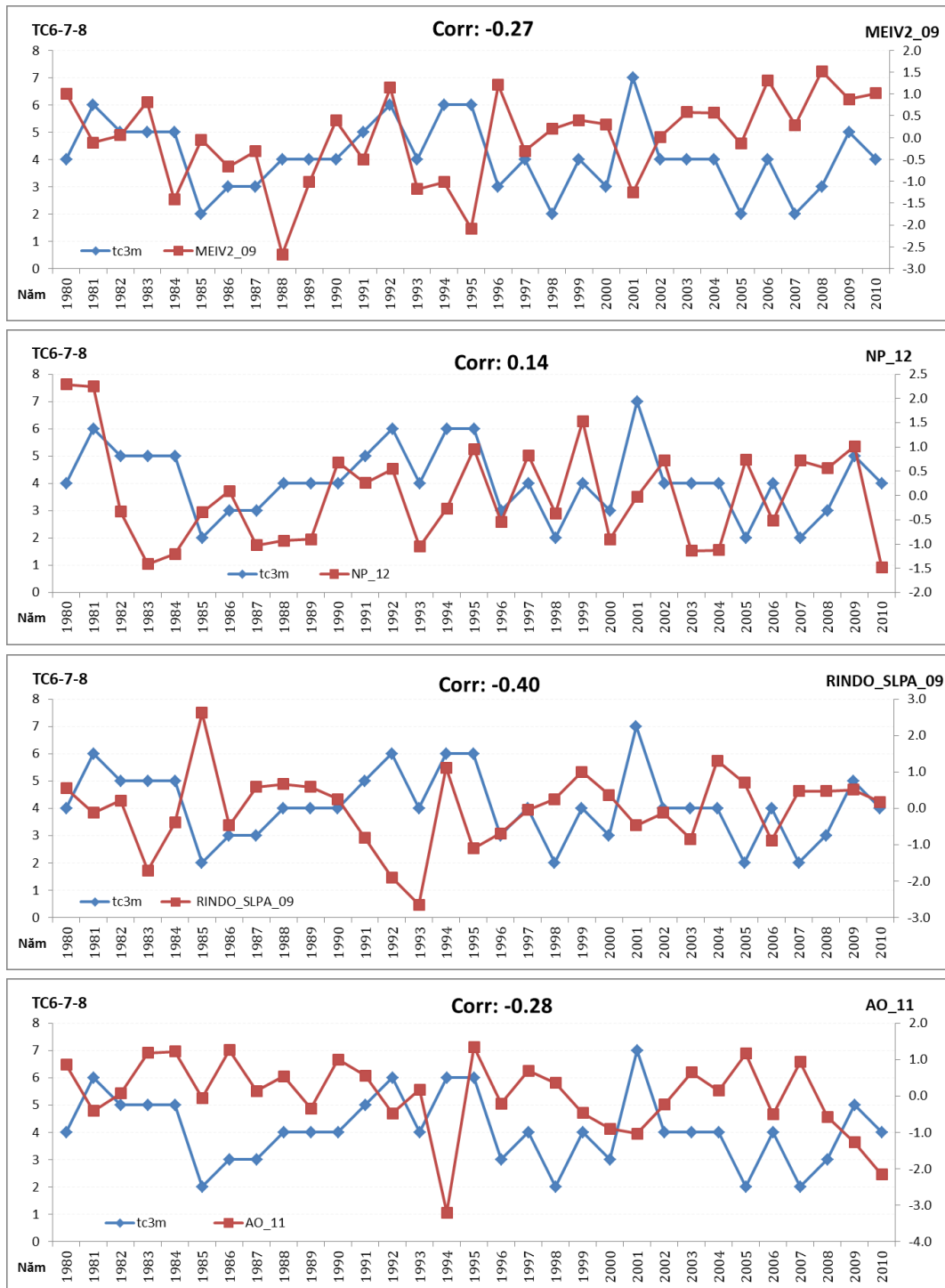
Hình 4.5 Sai số MAE trung bình theo các ngưỡng lọc khác nhau cho giai đoạn dự báo phụ thuộc và độc lập

Kết quả khảo sát tại Hình 4.5 cho thấy, sai số MAE trung bình cho giai đoạn mô hình luyện gia tăng đồng thuận với tham số Epsilon, trong khi đó lại thể hiện trái ngược với xu hướng giảm của giá trị MAE trung bình ở giai đoạn dự báo độc

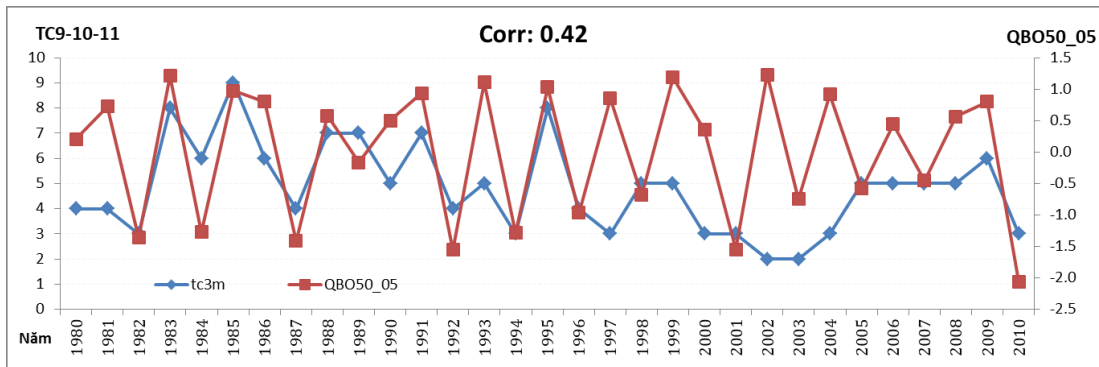
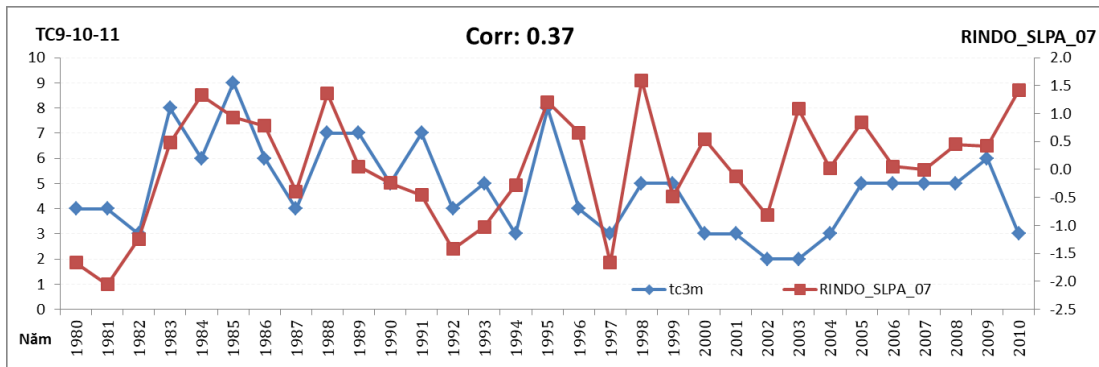
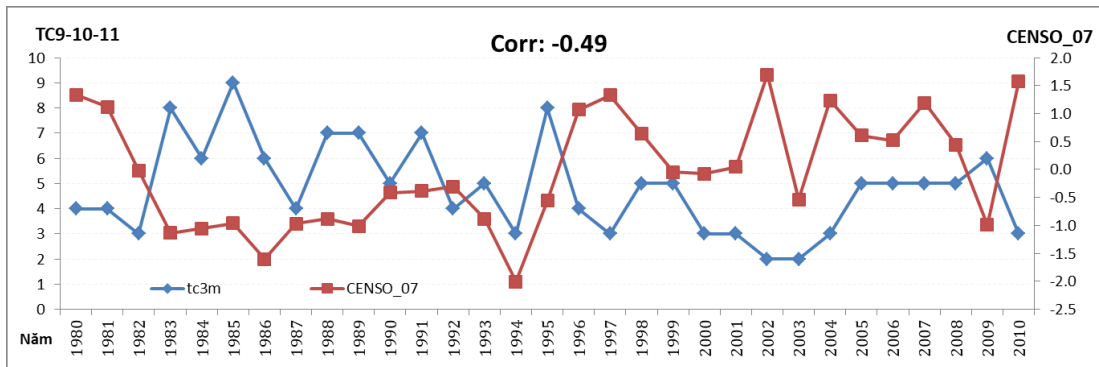
lập. Đường đồ thị MAE của hai giai đoạn phụ thuộc và độc lập cho thấy có sự “cân bằng” tại ngưỡng Epsilon 0,08, với các tham số Epsilon vượt trên và dưới ngưỡng này thì trung bình 4 mô hình khảo sát cho kết quả dự báo kém hơn. Xuất phát từ kết quả này, ở TH2 luận án lựa chọn tham số Epsilon ở mức 0,08 và theo đó có 4 nhân tố dự báo được lựa chọn (MEIV2\_9, NP\_12; RINDO\_9 và AO\_11) cho các dự báo tại tháng 6 và 3 nhân tố dự báo được lựa chọn (CENSO\_7, RINDO\_7, QBO50\_5) cho các dự báo tại tháng 9.

Xem xét kỹ hơn mối quan hệ giữa XTNĐ Biển Đông 3 tháng với các chỉ số khí hậu được lựa chọn, các Hình 4.6 và 4.7 dưới đây dẫn ra các bản đồ thể hiện quan hệ giữa yếu tố dự báo và nhân tố dự báo. Trong số các nhân tố được lựa chọn cho thời điểm dự báo tháng 6, chỉ số RINDO thể hiện quan hệ tốt nhất với XTNĐ Biển Đông 3 tháng 6-7-8, tiếp đến là chỉ số AO, MEIV2 và cuối cùng là NP. Hệ số tương quan giữa các chỉ số này với XTNĐ Biển Đông 3 tháng 6-7-8 tương ứng ở mức -0,4; -0,28, -0,27 và 0,14. Như vậy có thể thấy rằng, ngoài các chỉ số khí hậu có liên hệ gần về mặt không gian tới khu vực Biển Đông thì ở khoảng cách xa hơn, chỉ số Bắc Thái Bình Dương (NP), chỉ số dao động Bắc Cực (AO) cũng cho thấy mức độ ảnh hưởng tới số lượng XTNĐ đầu mùa bão. Kết quả này, phù hợp với nghiên cứu trước đây của Choi và cộng sự (2010) khi đã chỉ ra rằng, AO có ảnh hưởng đáng kể đến các hoạt động của XTNĐ ở TBTBD trong thời gian mùa hè (tháng 7, 8 và 9) [35] hay như nghiên cứu của Wang và cộng sự (2007) đã đưa ra hệ số tương quan giữa NP trung bình các tháng 6-7-8-9 với số lượng XTNĐ hàng năm của TBTBD ở mức 0,37 cho giai đoạn 1949-1998 [87]. Ở thời điểm dự báo tháng 9, chỉ số ENSO kết hợp (CENSO) có quan hệ tốt nhất với XTNĐ Biển Đông 3 cuối mùa bão tháng 9-10-11 (với hệ số tương quan đạt -0,49), tiếp đến là chỉ số QBO50 (0,42) và RINDO (0,37). Trong số 3 nhân tố được lựa chọn này, CENSO được tính toán thông qua sự kết hợp giữa chỉ số dao động nam SOI và SST vùng Nino 3.4, RINDO là dị thường chuẩn hóa khí áp mực biển vùng biển Indonesia (5°N - 5°S, 80°E - 135°W) trong khi QBO50 được tính theo gió vĩ hướng mực 50mb khu vực Singapore do vậy về bản chất, cả 3 nhân tố dự báo được lựa chọn ở thời điểm tháng 9 đều có liên hệ gần với

các điều kiện môi trường đại dương – khí quyển khu vực Biển Đông. Như vậy, so với thời điểm dự báo tháng 6 thì số lượng các nhân tố dự báo lựa chọn được cho thời điểm tháng 9 là ít hơn và thể hiện mức độ tương quan cao hơn.



Hình 4.6 Quan hệ giữa nhân tố dự báo và yếu tố dự báo tháng 6 ở TH2



Hình 4.7 Quan hệ giữa nhân tố dự báo và yếu tố dự báo tháng 9 ở TH2

#### 4.1.2 Kết quả dự báo số lượng XTNĐ Biển Đông theo TH1

Bằng việc khảo sát quan hệ tương quan giữa số lượng XTNĐ 3 tháng hoạt động trên khu vực Biển Đông với bộ 30 chỉ số khí hậu theo các độ trễ khác nhau (như trình bày tại Chương 3) đã cho thấy tiềm năng của bộ chỉ số này trong việc khai thác, lựa chọn làm nhân tố dự báo số lượng XTNĐ Biển Đông. Bộ số liệu 30 chỉ số khí hậu này được mở rộng xem xét với độ trễ từ tháng thứ 3 tới tháng thứ 12 để tạo thành bộ dữ liệu mới với 300 nhân tố dự tuyển tương ứng. Trong TH1 này, các nhân tố dự báo được lựa chọn từ tập 300 nhân tố dự tuyển qua kỹ thuật phân

tích thành phần chính PCA như đã mô tả tại Chương 2. Cơ sở, cách thức tuyển chọn nhân tố cho TH1 được trình bày tại phần 4.1, theo đó 7 nhân tố dự báo chính thức của TH1 tương ứng là 7 thành phần chính đầu tiên (7 EOFs) đóng góp/ giải thích trên 74 % thông tin của tập 300 nhân tố dự tuyển. Trên thực tế, việc sử dụng kỹ thuật PCA ở đây đảm bảo không xuất hiện hiện tượng đa cộng tuyến giữa các nhân tố dự báo bởi thuật toán PCA tìm ra các EOFs trực giao với nhau. Kết quả dự báo trên bộ số liệu phụ thuộc (giai đoạn 1980-2010) và dự báo kiểm nghiệm độc lập (giai đoạn 2011-2017) cùng với các chỉ số đánh giá chất lượng phương trình dự báo tại thời điểm tháng 6 (9) dự báo số lượng XTNĐ hoạt động trên khu vực Biển Đông trong các tháng 6-7-8 (9-10-11) lần lượt được dẫn ra tại các Bảng 4.9 đến 4.11 và Hình 4.8 dưới đây.

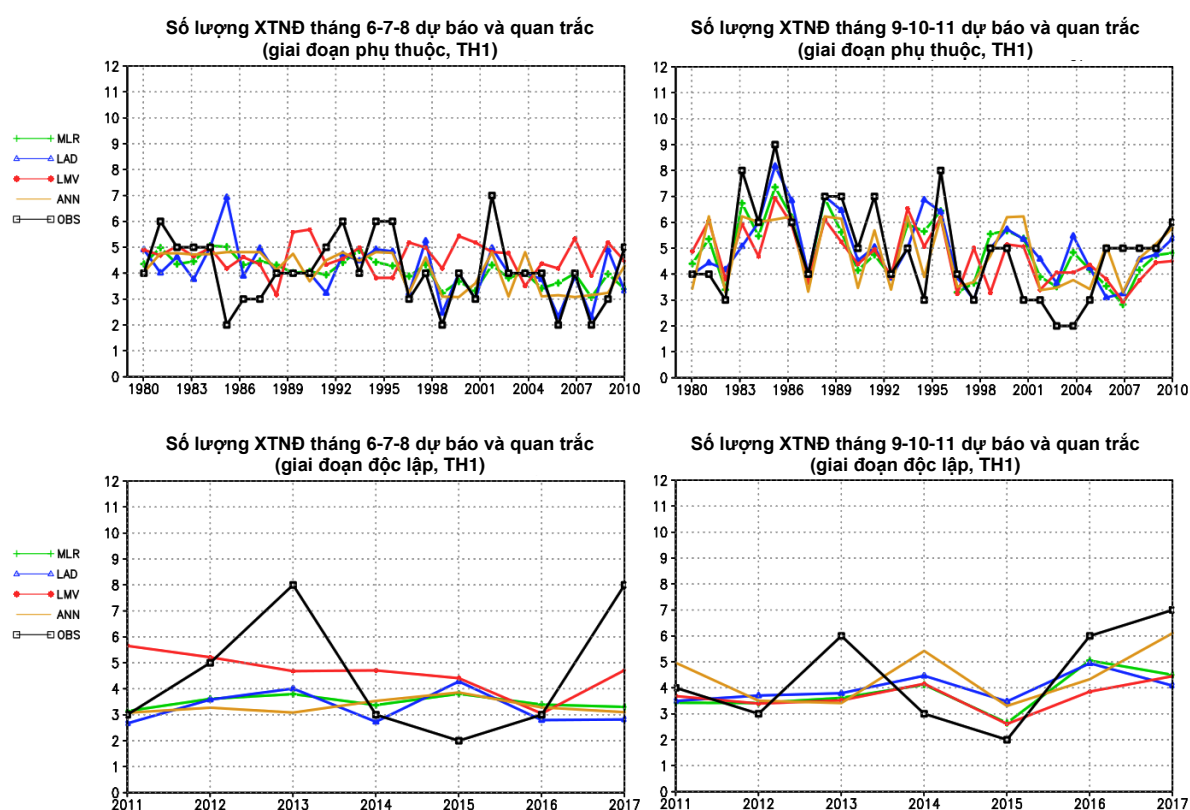
Bảng 4.9 Hệ số hồi quy và giá trị Pval của các mô hình dự báo thống kê theo TH1

Thời điểm dự báo	Mô hình thống kê	Các hệ số hồi quy								Pval
		b0	EOF1	EOF2	EOF3	EOF4	EOF5	EOF6	EOF7	
<b>Tháng 6</b>	MLR	4.10	0.07	2.05	-0.24	-0.10	-1.22	0.27	1.53	0.025
	LAD	4.11	-1.06	3.66	-1.91	-2.33	-0.55	1.33	1.53	0.094
	LMV	4.62	-2.05	0.10	-0.62	-0.16	-1.08	-2.25	-0.34	0.923
<b>Tháng 9</b>	MLR	4.81	0.03	-2.63	4.77	-0.10	0.07	2.91	-2.61	0
	LAD	4.93	-1.74	-2.88	5.57	-0.79	0.56	0.28	-2.65	0
	LMV	4.62	-0.90	-3.67	1.96	0.35	1.54	3.69	-2.48	0.001

Bảng 4.9 dẫn ra bộ hệ số hồi quy của các phương trình dự báo tổng số lượng XTNĐ hoạt động trên Biển Đông tích lũy 3 tháng liên tiếp tại hai thời điểm dự báo tháng 6 và 9. Các giá trị EOF1, ..., EOF7 là hệ số hồi quy tương ứng của 7 thành phần chính được giữ lại, b0 là hệ số chặn (intercept) hay hệ số tự do. Có thể thấy rằng, hệ số chặn “b0” có giá trị qui mô xấp xỉ với giá trị trung bình của số lượng XTNĐ trong mùa 3 tháng đầu mùa bão (tháng 6-7-8) và 3 tháng cuối mùa bão (tháng 9-10-11). Kết quả kiểm định T-test cho các phương trình dự báo thống kê như được thể hiện tại Bảng 4.9 cho thấy, cả 3 mô hình MLR, LAD và LMV đều thỏa mãn ý nghĩa thống kê (giá trị  $p_{val} \leq 0.05$ ) ở thời điểm tháng 9 dự báo cho số lượng XTNĐ ba tháng cuối mùa bão, tuy nhiên ở thời điểm tháng 6 dự báo cho các XTNĐ ba tháng đầu mùa bão chỉ có mô hình MLR thỏa mãn tiêu chí này, các

phương trình xây dựng đạt được theo mô hình LAD và LMV ở trường hợp này không thỏa mãn ý nghĩa thống kê.

Hình 4.8 dưới đây dẫn ra kết quả dự báo của 4 mô hình thống kê MLR, LAD, LMV và ANN trên bộ số liệu phụ thuộc (1980-2010) và độc lập (2011-2017) tương ứng với 2 thời điểm, dự báo các các XTND hoạt động trong 3 tháng đầu mùa bão và 3 tháng cuối mùa bão.



Hình 4.8 Số lượng XTND 3 tháng 6-7-8 và 9-10-11 trên Biển Đông giai đoạn 1980-2010 (hình trên) và giai đoạn độc lập 2011-2017 (hình dưới), số lượng quan trắc (OBS) và kết quả dự báo các mô hình MLR, LAD, LMV, ANN theo TH1

Có thể thấy, ở cả giai đoạn phụ thuộc và dự báo kiểm nghiệm độc lập, 4 mô hình khảo sát nhìn chung đã đưa ra được giá trị dự báo bám khá tốt về xu thế so với số liệu quan trắc. Đồ thị tại Hình 4.8 thể hiện diễn biến của các kết quả dự báo cho ba tháng cuối mùa bão (tháng 9-10-11) tốt hơn so với kết quả dự báo của ba tháng đầu mùa bão (tháng 6-7-8). Trong số 4 mô hình khảo sát, mô hình MLR, ANN cho thấy kết quả dự báo tốt hơn cả trong khi mô hình LAD, LMV ở một số thời điểm

đưa ra các giá trị không, chẳng hạn như tại thời điểm dự báo tháng 6 năm 1985 mô hình LAD đưa ra 7 con cho 3 tháng đầu mùa bão trong khi con số quan trắc thực tế chỉ có 2 con, hay tương tự như mô hình LMV ở năm 1989, 1990 hoặc cũng tại mô hình LMV ở năm 2011, 2014.

Bên cạnh đó, kết quả dự báo ở giai đoạn kiểm nghiệm độc lập cũng cho thấy, tại thời điểm dự báo tháng 6, cả 4 mô hình MLR, LAD, LMV và ANN đều cho kết quả dự báo sai (dự báo của mô hình thấp hơn so với quan trắc) các XTNĐ hoạt động trong 3 tháng đầu mùa bão của năm 2013, 2017 và dự báo không (dự báo của mô hình thấp cao hơn so với quan trắc) ở năm 2015. Các sai số này cũng gặp phải ở kết quả dự báo tại thời điểm tháng 9 tuy nhiên với mức độ sai lệch giữa quan trắc giảm nhiều. Tuy vậy, cũng cần lưu ý rằng 2013, 2017 là những năm ghi nhận số lượng XTNĐ hoạt động trên Biển Đông cao bất thường (năm 2013 có 16 con, 2017 có 17 con) trong khi năm 2015 lại là năm quan trắc được số lượng XTNĐ ít hơn thường lệ (6 con so với trung bình 10,9 con). Các kết quả này cho thấy các mô hình dự báo ở TH1 ít có khả năng nắm bắt được dị thường của XTNĐ Biển Đông tại một số thời điểm nhất định, đặc biệt là đối với các XTNĐ hoạt động trong 3 tháng đầu mùa bão (tháng 6-7-8).

Để đánh giá chất lượng và kỹ năng dự báo của các phương trình dự báo số lượng XTNĐ Biển Đông hoạt động trong thời gian 3 tháng đầu mùa bão và cuối mùa bão, Bảng 4.10 và 4.11 dưới đây dẫn ra các chỉ số sai số ME, MAE, RMAE, RMSE, hệ số tương quan giữa số lượng quan trắc và dự báo (Corr), chỉ số kỹ năng dự báo so với trung bình khí hậu (MSSS) của 4 mô hình và tổ hợp trung bình của 4 mô hình này (với các trọng số bằng nhau). Kết quả tại Bảng 4.10 cho thấy, tại thời điểm dự báo tháng 6, trên giai đoạn phụ thuộc hầu hết các mô hình đã đưa ra các giá trị sai số tương đương với quy mô của độ lệch chuẩn khí hậu. Cụ thể, sai số ME dao động trong khoảng 0 tới 0,5; MAE có giá trị xung quanh 1,0 và sai số RMSE khoảng 1,3 tương đương độ lệch chuẩn của XTNĐ ba tháng đầu mùa bão. Hệ số tương quan ở thời điểm này có giá trị dương ở cả 4 mô hình và mô hình tổ hợp song hai mô hình (LAD, LMV) có hệ số tương quan thấp nhất (tương ứng 0,31 và 0,02) có kỹ năng dự báo MSSS âm (tương ứng -0,1 và -0,37). Cùng tại thời điểm tháng 6,

ở giai đoạn dự báo độc lập (2011-2017), sai số ME, MAE trung bình tương ứng dao động quanh giá trị -1,0; 1,9 trong khi sai số RMSE ở mức 2,5 tới 2,8. Các sai số này có thể so sánh với độ lệch chuẩn (2,3) của yếu tố dự báo trên chuỗi 2011-2017. Hệ số tương quan ở giai đoạn này hầu hết có giá trị dương ngoại trừ mô hình ANN có giá trị âm (-0,66). Kỹ năng dự báo của các mô hình đánh giá trên giai đoạn này hầu hết không đạt (ngoại trừ mô hình LMV). Tuy nhiên cũng cần lưu ý rằng, giai đoạn 2011-2017 có sự biến động mạnh của số lượng XTNĐ Biển Đông (trong khi năm 2013, 2017 có số lượng XTNĐ 16, 17 cơn cao đột biến thì năm 2015 lại xuất hiện 6 cơn so với trung bình nhiều năm 10,9 cơn). Như vậy, có thể thấy rằng, ở thời điểm dự báo tại tháng 6, các sai số ME, MAE và RMSE của cả 4 mô hình và tổ hợp 4 mô hình nhìn chung chỉ tương đương với dự báo khí hậu, nhận định này được củng cố hơn khi chỉ số RMAE (tỷ số giữa sai số tuyệt đối với trung bình khí hậu) chỉ dao động trong khoảng 30% (40%) ở giai đoạn dự báo phụ thuộc (độc lập) đồng thời hệ số tương quan và hệ số kỹ năng dự báo có độ lớn tuyệt đối thấp.

Bảng 4.10 Sai số và chỉ số đánh giá kỹ năng dự báo tại thời điểm dự báo tháng 6 ở giai đoạn dự báo phụ thuộc (1980-2010) và độc lập (2011-2017) theo TH1

Thời điểm dự báo Yếu tố dự báo (Y)	Tháng 6									
	XTNĐ Biển Đông 3 tháng đầu mùa bão (các tháng 6-7-8)									
	Giai đoạn phụ thuộc					Giai đoạn độc lập				
Trung bình Y	4.1					4.6				
Độ lệch chuẩn	1.3					2.3				
Chỉ số đánh giá	Mô hình thống kê									
	MLR	LAD	LMV	ANN	E1234	MLR	LAD	LMV	ANN	E1234
ME:	0	0	0.5	0	0.1	-1.1	-1.3	0.1	-1.3	-0.9
MAE:	0.9	0.9	1.3	0.9	0.9	1.9	2	1.9	2	1.9
RMSE:	1.2	1.3	1.5	1.1	1.2	2.5	2.7	2.3	2.8	2.5
RMAE:	26.3	24.9	39.2	25.5	26.9	37.1	40.6	50.6	40	39.8
CORR:	0.40	0.31	0.02	0.54	0.45	0.14	0.10	0.17	-0.66	0.04
MSSS:	0.16	-0.10	-0.37	0.30	0.19	-0.15	-0.29	0.04	-0.40	-0.11

Bảng 4.11 dưới đây dẫn ra các sai số ME, MAE, RMSE và chỉ số đánh giá RMAE, CORR, MSSS cho thời điểm dự báo tháng 9 theo TH1. Kết quả tại Bảng 4.11 cho thấy, sai số ME xấp xỉ 0 ở giai đoạn luyện của các mô hình thay đổi sang giá trị trên dưới -0,5 ở giai đoạn độc lập. Sai số MAE, RMSE biến động trong khoảng 1,0 - 1,7 trong đó ở giai đoạn phụ thuộc các sai số này có độ lớn thấp hơn

khoảng 0,2 so với giai đoạn dự báo độc lập. Đáng chú ý hơn, các sai số này đều thấp hơn so với độ lệch chuẩn khí hậu (cùng bằng 1,8 ở hai giai đoạn) và chỉ số RMAE biểu thị tỷ lệ % của sai số MAE so với chuẩn khí hậu ở thời điểm dự báo này cũng phản ánh phù hợp với các sai số trên khi chỉ biến động trong khoảng 24-38%. Đồng thuận với các kết quả sai số này, hệ số tương quan dương với giá trị trung bình 0,7 (0,64) và chỉ số kỹ năng dự báo MSSS dương với giá trị trung bình 0,45 (0,2) tương ứng ở giai đoạn phụ thuộc (độc lập). Các kết quả này đã cho thấy, ở cả 4 mô hình và tổ hợp 4 mô hình khảo sát đều có kỹ năng dự báo số lượng các XTNĐ hoạt động trong 3 tháng cuối mùa bão ở Biển Đông đồng thời kết quả dự báo của các mô hình này tốt hơn so với dự báo trung bình khí hậu.

Bảng 4.11 Sai số và chỉ số đánh giá kỹ năng dự báo tại thời điểm dự báo tháng 9 ở giai đoạn dự báo phụ thuộc (1980-2010) và độc lập (2011-2017) theo TH1

Thời điểm dự báo Yếu tố dự báo (Y)	Tháng 9									
	XTNĐ Biển Đông 3 tháng cuối mùa bão (các tháng 9-10-11)									
	Giai đoạn phụ thuộc					Giai đoạn độc lập				
Trung bình Y	4.8					4.4				
Độ lệch chuẩn	1.8					1.8				
Chỉ số đánh giá	Mô hình thống kê									
	MLR	LAD	LMV	ANN	E1234	MLR	LAD	LMV	ANN	E1234
ME:	0	0.1	-0.2	0	0	-0.6	-0.4	-0.8	0	-0.4
MAE:	1.1	1	1.3	1	1.1	1.2	1.5	1.4	1.5	1.3
RMSE:	1.3	1.5	1.5	1.3	1.3	1.5	1.7	1.7	1.6	1.5
RMAE:	27.6	27.5	32.4	23.9	27.2	27.1	36.4	29.1	38.6	31.1
CORR:	0.68	0.60	0.57	0.70	0.70	0.72	0.42	0.64	0.40	0.61
MSSS:	0.45	0.33	0.30	0.49	0.48	0.34	0.13	0.14	0.16	0.28

Như vậy, kết quả đánh giá chất lượng các phương trình thống kê dự báo số lượng XTNĐ Biển Đông trên bộ số liệu phụ thuộc (1980-2010) và kiểm nghiệm trên bộ số liệu độc lập (2011-2017) đã cho thấy cả 4 mô hình thống kê khảo sát (MLR, LAD, LMV, ANN) cùng tổ hợp 4 mô hình nhìn chung đã nắm bắt được xu thế biến động so với số liệu quan trắc. Trong số các mô hình thống kê khảo sát, mô hình MLR cho kết quả tốt nhất tiếp đến là mô hình ANN và kém hơn ở mô hình LAD, LMV. Qua việc phân tích, đánh giá các chỉ số sai số và kỹ năng dự báo của 4 mô hình thống kê và tổ hợp của chúng cho thấy việc nghiên cứu, sử dụng các nhân tố dự báo từ các thành phần chính của bộ chỉ số khí hậu theo TH1 là khả thi. Chất

lượng dự báo số lượng các XTNĐ hoạt động trong 3 tháng đầu mùa bão của các mô hình thống kê xây dựng được có thể so sánh với dự báo trung bình khí hậu, đặc biệt các phương trình dự báo cho các XTNĐ hoạt động trong 3 tháng cuối mùa bão cho kết quả và chất lượng tốt hơn so với dự báo khí hậu. Đánh giá sai số và kỹ năng dự báo các XTNĐ hoạt động trong 3 tháng cuối mùa bão cho thấy, kết quả này đã tiệm cận và có thể so sánh với một số kết quả đã công bố trong nước và quốc tế về dự báo hạn mùa số lượng XTNĐ (Nguyễn Văn Tuyên (2008) [17-18]; Chan và cộng sự (1998) [30]; Li và cộng sự (2013); Klotzbach và cộng sự (2019) [53]).

#### 4.1.3 Kết quả dự báo số lượng XTNĐ Biển Đông theo TH2

Qua khảo sát tương quan giữa số lượng XTNĐ Biển Đông 3 tháng với bộ 30 chỉ số khí hậu tại Chương 3 đã cho thấy tiềm năng để tìm ra bộ nhân tố dự báo cho bài toán dự báo số lượng XTNĐ Biển Đông. Bộ số liệu 30 chỉ số này được mở rộng xem xét với độ trễ từ tháng thứ 3 tới tháng thứ 12 để tạo thành bộ dữ liệu mới với 300 nhân tố dự tuyển tương ứng. TH1 ở trên đã khảo sát, tuyển chọn bộ nhân tố dự báo bằng kỹ thuật phân tích thành phần chính PCA để lọc ra được bộ nhân tố dự báo phù hợp từ tập 300 nhân tố dự tuyển. TH2 dưới đây sẽ sử dụng kỹ thuật hồi quy từng bước như đã mô tả tại Chương 2 để tuyển chọn ra bộ nhân tố dự báo chính thức. Tại tiểu mục 4.1.1, luận án đã trình bày các bước tuyển chọn bộ nhân tố dự báo theo kỹ thuật hồi quy từng bước, theo đó nhân tố dự báo cho thời điểm dự báo tháng 6 gồm 4 nhân tố (MEIV2\_9, NP\_12; RINDO\_9 và AO\_11) và cho thời tháng 9 là 3 nhân tố (CENSO\_7, RINDO\_7, QBO50\_5). Kết quả dự báo của TH3 trên bộ số luyện (giai đoạn 1980-2010) và dự báo kiểm nghiệm trên bộ số liệu độc lập (giai đoạn 2011-2017) cùng với các chỉ số đánh giá chất lượng phương trình dự báo lần lượt được dẫn ra tại các Bảng 4.12 đến 4.14, Hình 4.9 dưới đây.

Bảng 4.12 Hệ số hồi quy và nhân tố dự báo của các phương trình dự báo số lượng XTNĐ hoạt động trong 3 tháng 6-7-8 theo TH2

Mô hình thống kê	Các hệ số hồi quy					Pval
	b0	MEIV2_09	NP_12	RINDO_09	AO_11	
MLR	4.085805	-0.505304	0.398123	-0.688949	-0.649347	0
LAD	4.084788	-0.639327	0.629292	-0.912056	-0.765184	0
LMV	4.060043	-0.314569	0.228803	-0.545319	-0.771103	0

Bảng 4.13 Hệ số hồi quy và nhân tố dự báo của các phương trình dự báo số lượng XTNĐ hoạt động trong 3 tháng 9-10-11 theo TH2

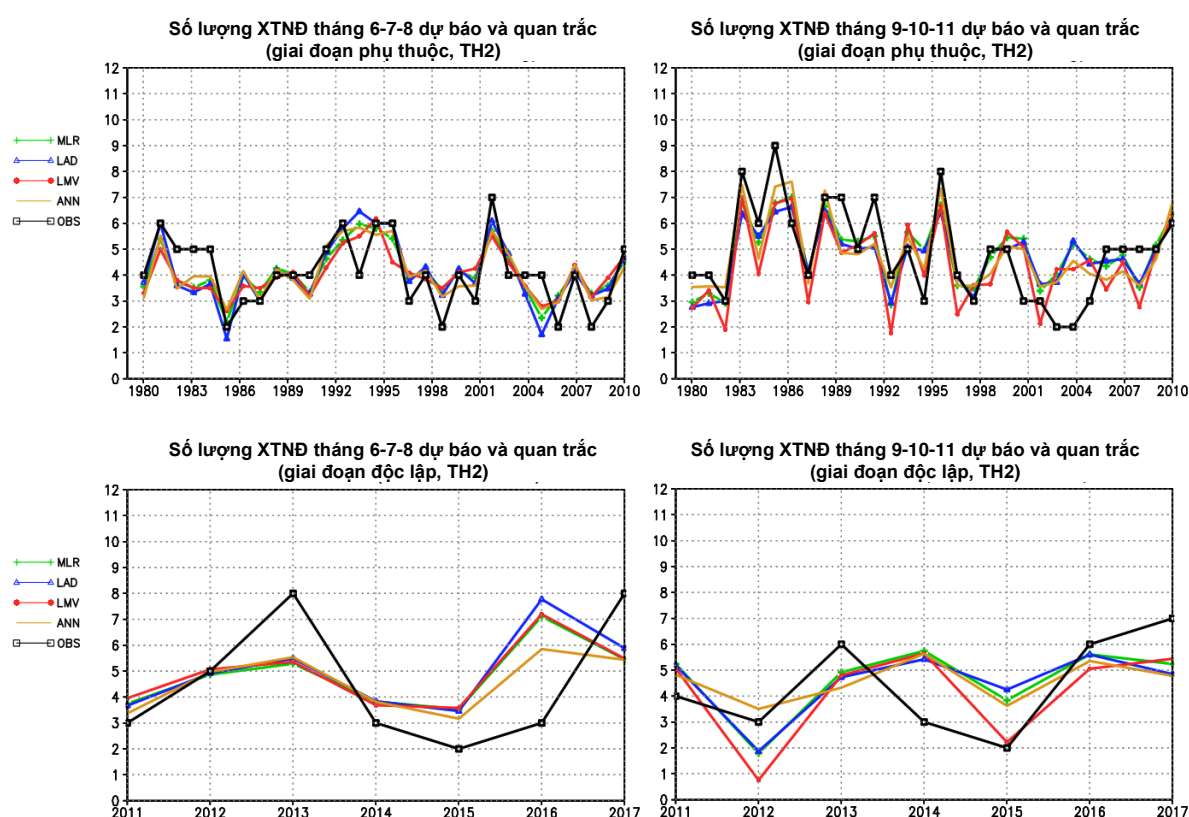
Mô hình thống kê	Các hệ số hồi quy				Pval
	b <sub>0</sub>	CENSO_07	RINDO_07	QBO50_05	
MLR	4.748576	-0.716419	0.59809	0.80673	0
LAD	4.644990	-0.608663	0.710547	0.567789	0
LMV	4.289967	-0.768794	0.478925	1.330423	0

Kết quả tại Bảng 4.13 cho thấy, các phương trình dự báo thống kê xây dựng cho thời điểm tháng 6 để dự báo số lượng 3 tháng đầu mùa bão gồm 4 biến và một hệ số tự do  $b_0$ . Hệ số Pval thể hiện cả 3 mô hình MLR, LAD và LMV đều thỏa mãn ý nghĩa thống kê (giá trị  $p_{val} \leq 0.05$ ). Nhìn chung, hệ số tự do  $b_0$  xấp xỉ với trung bình khí hậu và các hệ số của bộ nhân tố dự báo có cùng quy mô độ lớn với nhau (đều có giá trị tuyệt đối nhỏ hơn 1). Xét về dấu của các hệ số, ngoại trừ hệ số cho nhân tố dự báo NP\_12 mang dấu “ + ” còn lại các hệ số khác mang dấu “ - ” điều này phản ánh những gia tăng (suy giảm) của nhân tố NP\_12 sẽ đóng góp vào việc làm tăng (giảm) số lượng XTNĐ đầu mùa bão ở Biển Đông và ngược lại. Những kết quả này là hoàn toàn phù hợp với các khảo sát về tương quan giữa XTNĐ 3 tháng đầu mùa bão Biển Đông với các nhân tố dự báo ở tiểu mục 4.11 với hệ số tương quan của các nhân tố MEIV2\_9, NP\_12; RINDO\_9 và AO\_11 lần lượt là -0,27; 0,14; -0,4; -0,28 (Hình 4.6).

Các phương trình dự báo thống kê xây dựng cho thời điểm tháng 9 để dự báo số lượng 3 tháng cuối mùa bão tại Bảng 4.13 với số lượng nhân tố và loại nhân tố khác so với các phương trình dự báo cho thời điểm tháng 6. Tuy vậy, các phương trình dự báo thống kê cho thời điểm tháng 9 đều thỏa mãn ý nghĩa thống kê (giá trị  $p_{val} \leq 0.05$ ) và mang các đặc điểm hoàn toàn tương tự như với các phương trình dự báo cho thời điểm tháng 6. Để có cái nhìn trực quan hơn về chất lượng của các phương trình dự báo xây dựng được, Hình 4.9 dưới đây dẫn ra các bản đồ thể hiện kết hợp giữa số liệu quan trắc theo RSMC và kết quả dự báo của các mô hình thống kê về số lượng XTNĐ hoạt động trên Biển Đông trong 3 tháng đầu mùa bão (tháng 6-7-8) và 3 tháng cuối mùa bão (tháng 9-10-11).

Kết quả tại Hình 4.9 cho thấy, trên cả giai đoạn dự báo phụ thuộc (1980-2010) và dự báo kiểm nghiệm độc lập (2011-2017) các mô hình dự báo thống kê

khác nhau đưa ra kết quả dự báo khác nhau về số lượng XTNĐ hoạt động trên khu vực Biển Đông trong 3 tháng liền sau. Ở giai đoạn dự báo phụ thuộc, tồn tại một số năm dự báo sai (3 tháng đầu mùa bão các năm 1982, 1983, 1984 hay 3 tháng cuối mùa bão các năm 1985, 1989, 1991) hoặc các mô hình đưa ra kết quả dự báo không (3 tháng đầu mùa bão năm 1993, 3 tháng cuối mùa bão năm 2003, 2004, 2005). Tuy vậy, nhìn chung các mô hình khảo sát đều nắm bắt được diễn biến theo các số liệu quan trắc, thêm vào đó chênh lệch giữa kết quả mô hình với số liệu quan trắc ở giai đoạn mô hình luyện không quá lớn, trung bình khoảng 1 cơn.



Hình 4.9 Số lượng XTNĐ 3 tháng 6-7-8 và 9-10-11 trên Biển Đông giai đoạn 1980-2010 (hình trên) và giai đoạn độc lập 2011-2017 (hình dưới), số lượng quan trắc (OBS) và kết quả dự báo các mô hình MLR, LAD, LMV, ANN theo TH2

Ở giai đoạn dự báo kiểm nghiệm độc lập, về tổng thể 4 mô hình khảo sát và tổ hợp mô hình đều đã thể hiện được biến động XTNĐ 3 tháng đầu mùa bão (cuối mùa bão) giai đoạn 2011-2017 tương ứng với thời điểm dự báo tháng 6 (9). Kết quả dự báo cho giai đoạn này tại Hình 4.9 (hình dưới) cũng cho thấy hầu hết các mô

hình không nắm bắt được số lượng XTNĐ 3 tháng đầu mùa bão ở những năm có số lượng XTNĐ cao bất thường như năm 2013, 2017.

Bảng 4.14 và 4.15 dưới đây dẫn ra các sai số và chỉ số kỹ năng dự báo đối với XTNĐ Biển Đông hoạt động trong thời gian 3 tháng đầu mùa bão và cuối mùa bão, của 4 mô hình gốc và tổ hợp trung bình E1234.

Bảng 4.14 Sai số và chỉ số đánh giá kỹ năng dự báo tại thời điểm tháng 6 ở giai đoạn dự báo phụ thuộc (1980-2010) và độc lập (2011-2017) theo TH2

Thời điểm dự báo	Tháng 6									
	XTNĐ Biển Đông 3 tháng đầu mùa bão (các tháng 6-7-8)									
Yếu tố dự báo (Y)	Giai đoạn phụ thuộc					Giai đoạn độc lập				
Trung bình Y	4.1					4.6				
Độ lệch chuẩn	1.3					2.3				
Chỉ số đánh giá	Mô hình thống kê									
	MLR	LAD	LMV	ANN	E1234	MLR	LAD	LMV	ANN	E1234
ME:	0	0	0	0	0	0.3	0.4	0.3	0	0.3
MAE:	0.7	0.7	0.8	0.7	0.7	1.8	1.8	1.8	1.5	1.7
RMSE:	0.9	0.9	1	0.8	0.9	2.2	2.3	2.2	1.8	2.1
RMAE:	20.9	19.9	23	19.7	20.1	47.8	48.8	48.5	36.6	45.3
CORR:	0.72	0.71	0.67	0.76	0.73	0.36	0.35	0.35	0.66	0.43
MSSS	0.51	0.45	0.44	0.59	0.54	0.14	0.06	0.12	0.41	0.20

Kết quả đánh giá các sai số và chỉ số kỹ năng dự báo cho các mô hình và tổ hợp mô hình trên bộ số liệu phụ thuộc tại Bảng 4.14 cho thấy, cả 5 mô hình đã đưa ra kết quả dự báo khá tốt số lượng XNĐĐ 3 tháng đầu mùa bão (tháng 6-7-8). Cụ thể, sai số ME đạt giá trị tuyệt đối (0) ở giai đoạn luyện của các mô hình và biến động trong khoảng 0 tới 0,4 ở giai đoạn dự báo kiểm nghiệm độc lập, tương tự vậy sai số MAE, RMSE trung bình ở mức 0,7; 0,9 và biến động ở dự báo độc lập tương ứng là 1,8; 2,1 so với độ lệch chuẩn tương ứng của 2 chuỗi số liệu XTNĐ ba tháng là 1,3 và 2,3. Điều đó cho thấy sai số của các mô hình tương đối ổn định xung quanh các giá trị thấp hơn độ lệch chuẩn khí hậu. Chỉ số đánh giá RMAE cũng củng cố nhận định này khi dao động trên dưới 20% (45%) đối với các kết quả dự báo cho 3 tháng đầu mùa bão ở giai đoạn phụ thuộc (độc lập). Đặc biệt hơn, ở cả 5 mô hình

hệ số tương quan giữa số lượng quan trắc và dự báo của mô hình có giá trị dương cao (trung bình trên 0,7 ở giai đoạn phụ thuộc và 0,4 cho giai đoạn độc lập), thêm vào đó kỹ năng dự báo đều thể hiện đạt với các giá trị lớn (MSSS trung bình đạt 0,5 ở giai đoạn luyện và 0,2 ở giai đoạn dự báo độc lập).

Kết quả dự báo kiểm nghiệm trên chuỗi độc lập đối với các XTNĐ Biển Đông trong 3 tháng cuối mùa bão tại Bảng 4.15 chỉ ra nhiều nét tương đồng so với kết quả dự báo cho 3 tháng đầu mùa bão. Tại đây, sai số ME, MAE và RMSE tương ứng ở hai giai đoạn đều ổn định và nhỏ hơn so với độ lệch của của bộ số liệu quan trắc tương ứng. Tỷ lệ giữa sai số tuyệt đối với trung bình khí hậu ở thời điểm dự báo này chỉ dao động trong khoảng trên dưới 30% (40%) thêm vào đó hệ số tương quan, kỹ năng dự báo đều có giá trị dương, với độ lớn tương ứng trung bình 0,7 và 0,5 (0,5 và 0,2) cho giai đoạn mô hình luyện (dự báo độc lập).

Bảng 4.15 Sai số và chỉ số đánh giá kỹ năng dự báo tại thời điểm tháng 9 ở giai đoạn dự báo phụ thuộc (1980-2010) và độc lập (2011-2017) theo TH2

Thời điểm dự báo	Tháng 9									
	XTNĐ Biển Đông 3 tháng cuối mùa bão (các tháng 9-10-11)									
Yếu tố dự báo (Y)	Giai đoạn phụ thuộc					Giai đoạn độc lập				
Trung bình Y	4.8					4.4				
Độ lệch chuẩn	1.8					1.8				
Chỉ số đánh giá	Mô hình thống kê									
	MLR	LAD	LMV	ANN	E1234	MLR	LAD	LMV	ANN	E1234
ME:	0	-0.1	-0.5	0	-0.1	0.2	0.1	-0.3	0.1	0
MAE:	1	0.9	1.3	0.9	1	1.5	1.5	1.4	1.4	1.4
RMSE:	1.2	1.3	1.4	1.1	1.2	1.6	1.7	1.6	1.6	1.6
RMAE:	26.1	24.8	32.6	24.2	25.3	43.5	45.6	37	39.5	40.2
CORR:	0.72	0.70	0.68	0.77	0.74	0.48	0.40	0.59	0.39	0.50
MSSS	0.51	0.49	0.35	0.59	0.55	0.21	0.10	0.22	0.19	0.26

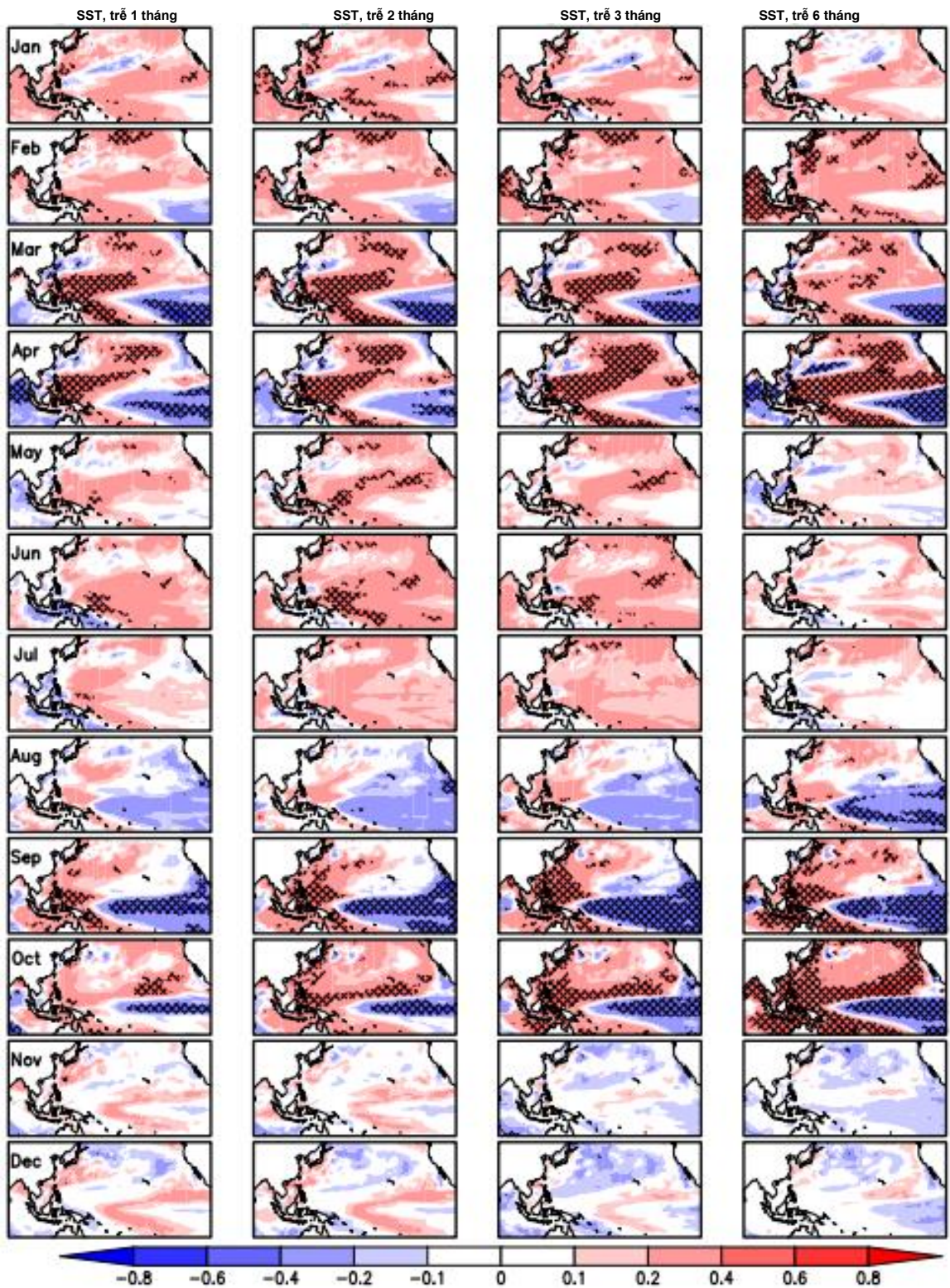
Tóm lại, qua đánh giá các sai số và chỉ số dự báo kỹ năng của 4 mô hình thống kê MLR, LAD, LMV, ANN và tổ hợp của chúng với các nhân tố dự báo được tuyển chọn theo TH2 trên bộ số liệu luyện (1980-2010) và kiểm nghiệm trên bộ số liệu độc lập (2011-2017) cho thấy tổ hợp 4 mô hình (E1234) thể hiện tốt nhất, tiếp

đến là mô hình MLR, ANN và cuối cùng là LMV, LAD. Tuy vậy, cả 5 mô hình đều đưa ra các hệ số kỹ năng và hệ số tương quan với số liệu quan trắc dương và ổn định, điều này khẳng định kỹ năng dự báo ở TH2 tốt hơn so với dự báo trung bình khí hậu cho cả các dự báo đối với các XTNĐ hoạt động trong 3 tháng đầu mùa bão và 3 tháng cuối mùa bão. So sánh với TH1, TH2 đưa ra các giá trị sai số thấp hơn đồng thời thể hiện các kỹ năng dự báo tốt hơn. Kết quả này một phần có thể giải thích bởi các nhân tố được lựa chọn theo TH2 tốt hơn, thể hiện qua các hệ số tương quan riêng giữa nhân tố dự báo và yếu tố dự báo cao hơn nhiều so với TH1, trong khi TH1 các nhân tố dự báo lựa chọn lại thiên về phản ánh/ giải thích nhiều hơn thông tin của bộ nhân tố dự tuyển ban đầu. Bên cạnh đó, các sai số và kỹ năng dự báo các XTNĐ hoạt động trong 3 tháng đầu và cuối mùa bão ở TH2 cho thấy, kết quả này tương đương và thậm chí ở một số thời điểm cụ thể đã đưa ra kết quả tốt hơn so với những kết quả đã công bố trong nước và quốc tế về dự báo hạn mùa số lượng XTNĐ (Nguyễn Văn Tuyên (2008) [17-18]; Chan và cộng sự (1998) [30]; Li và cộng sự (2013); Klotzbach và cộng sự (2019) [53]).

## **4.2 Dự báo hạn mùa số lượng XTNĐ trên Biển Đông bằng phương pháp thống kê động lực**

### **4.2.1 Quan hệ giữa các trường sản phẩm CFSv2 với số lượng XTNĐ Biển Đông**

Để có cơ sở cho việc lựa chọn nhân tố dự báo xây dựng các phương trình dự báo thống kê từ sản phẩm của mô hình khí hậu CFSv2, luận án đã tiến hành khảo sát tương quan giữa các trường quy mô lớn là sản phẩm của CFSv2 với các hạn dự báo 1, 2, 3 và 6 tháng với số lượng XTNĐ hoạt động trên khu vực Biển Đông trong 3 tháng liên tiếp. Các trường quy mô lớn của CFSv2 được luận án khảo sát gồm: 1) Nhiệt độ bề mặt biển (SST); 2) Khí áp mực biển (PMSL); 3) Xoáy tương đối mực 850mb (VOR850); 4) Độ đứt gió theo mực 200mb và 850mb (VWS); 5) Vận tốc thẳng đứng mực 850mb (VVEL850); 6) Độ ẩm tương đối mực 500mb (RH500) và 7) Độ cao địa thế vị mực 500mb (H500). Kết quả các khảo sát tương quan này lần lượt trình bày tại các Hình 4.5 tới Hình 4.10 dưới đây.

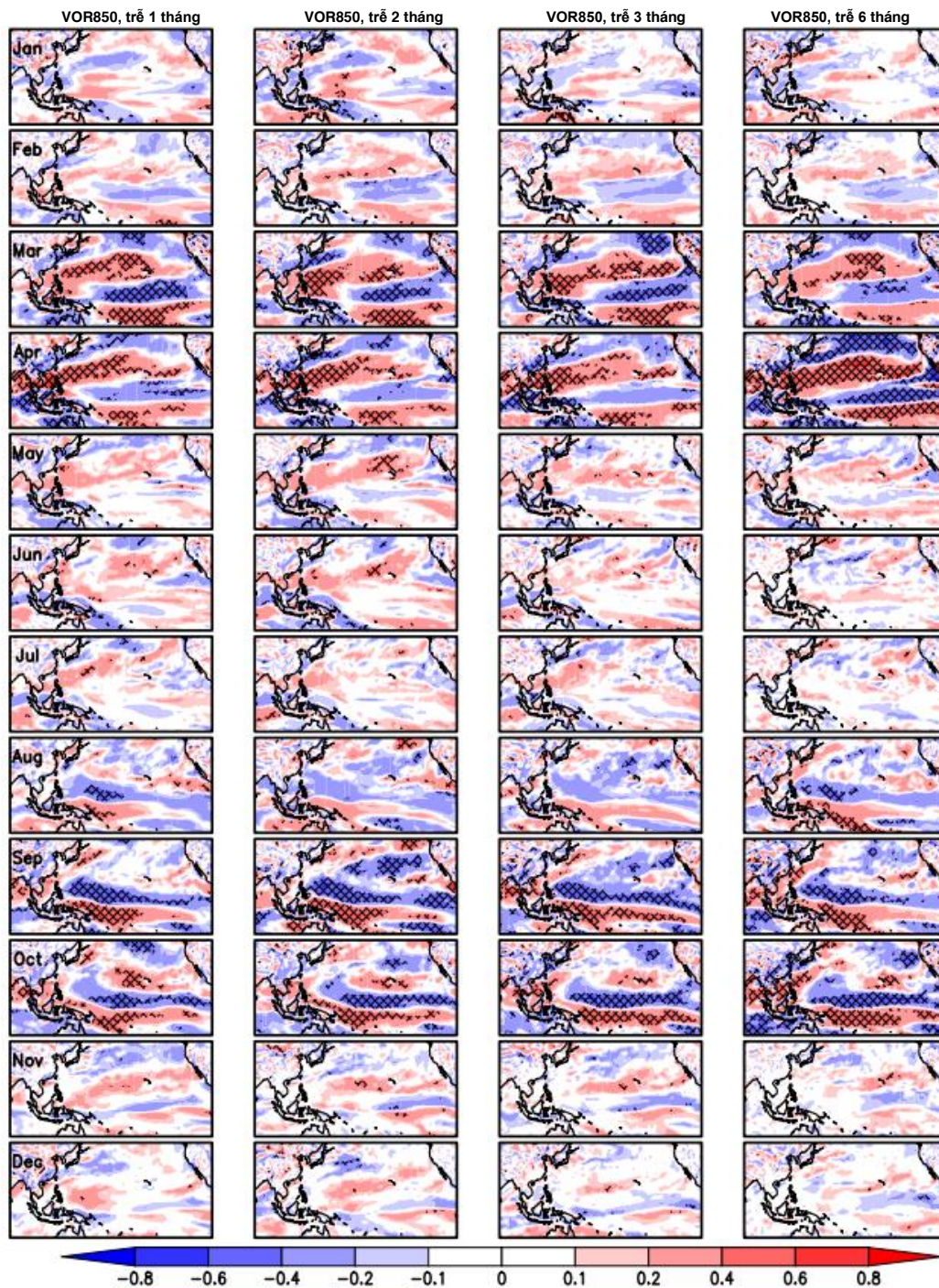


Hình 4.10 Bản đồ tương quan giữa SST với các độ trễ 1, 2, 3, 6 tháng và số lượng XTND Biển Đông ba tháng (vùng gạch chéo thể hiện giá trị tuyệt đối của hệ số tương quan  $\geq 0,4$  và thỏa mãn độ tin cậy thống kê 95%)

Hình 4.5 trình bày bản đồ phân bố không gian của hệ số tương quan giữa tổng số lượng XTNĐ ba tháng và trường nhiệt độ bề mặt biển (SST). Phân bố không gian của hệ số tương quan thể hiện hình thể giống ENSO, đặc biệt là trong các tháng 3, 4, 8, 9, 10, 11. Cụ thể, hệ số tương quan ở phần trung tâm và phía đông của TBD nhiệt đới nhận giá trị âm, trong khi phần phía tây kéo dài đến biển Phillipines nhận giá trị dương. Hình thể dạng vành khuyên của SST này đã được bàn luận ở nhiều nghiên cứu trước đây [59]. Khi vùng biển nhiệt đới ở trung tâm và phía đông của TBD ấm lên (pha ấm – El Niño). Trong pha nóng của ENSO (El Niño), bề mặt nước biển ở vùng trung tâm và Đông Thái Bình Dương xích đạo nóng lên dị thường, tại đây hoạt động đối lưu và dòng thăng được tăng cường. Ngược lại, ở pha lạnh của ENSO (La Niña), vùng đối lưu và dòng thăng dịch chuyển về phía tây của TBD và được tăng cường mà nó liên quan đến sự tăng của nhiệt độ mặt nước biển ở đây. Kết quả gia tăng của hoạt động đối lưu và dòng thăng đã tạo điều kiện thuận lợi hơn cho việc hình thành XTNĐ. Ngoài ra, tương quan âm tồn tại ở vùng biển Ấn Độ Dương có thể biểu hiện quan hệ giữa XTNĐ Biển Đông với IOD (lượng cực Ấn Độ Dương) [23]. Với IOD âm (dương), hoạt động đối lưu mạnh (yếu) ở khu vực gần Indonesia, ảnh hưởng đến chế độ mưa và bão ở các vùng lân cận. Các tháng khác (tháng 12, 1), phân bố không gian của hệ số tương quan có xu hướng ngược lại (dương ở phần trung tâm TBD, âm ở gần Indonesia), tuy nhiên, sự khác biệt là không rõ ràng. Tháng 6, 7, tương quan âm (dương) ở khu vực gần Indonesia (đông và trung tâm TBD). Thực tế, ảnh hưởng của ENSO chỉ được đề cập đến nhiều trong các tháng mùa bão muộn (tháng 9-11). Vì vậy, phân bố không gian của hệ số tương quan tương đối phù hợp với bàn luận của những nghiên cứu trước đây. Như vậy nhìn chung, các bản đồ tương quan trên Hình 4.5 cho thấy tồn tại nhiều vùng biểu thị hệ số tương quan cao trong 12 thời điểm khảo sát trong năm và rõ nét hơn trong các tháng chuyển mùa như tháng 3, 4 và tháng 9, 10. Hình thể không gian và giá trị của hệ số tương quan này không thay đổi nhiều giữa các thời hạn dự báo khác nhau (1, 2, 3, 6 tháng). Điều này cho thấy, sản phẩm SST của CFSv2 có tiềm năng sử

dụng để lựa chọn nhân tố dự báo cho bài toán dự báo thống kê trước các thời hạn 1, 2, 3, thậm chí 6 tháng.

Hình 4.6 trình bày bản đồ phân bố không gian của hệ số tương quan giữa tổng số XTNĐ 3 tháng và xoáy tương đối mực thấp (mực 850-hPa). Trong các tháng 8-9-10, từ khu vực biển Philipines đến khoảng kinh độ 180°E (trung tâm TBD), tương quan thể hiện 2 dải có giá trị đối lập nhau song song với phương vĩ hướng. Dải thứ nhất từ 5°N-20°N thể hiện tương quan âm, dải thứ hai nằm bên dưới 5°N thể hiện tương quan dương. Dải 5°N-20°N phù hợp với vị trí hoạt động của rãnh gió mùa [94]. Tương quan âm ngụ ý, nếu rãnh gió mùa ở phần đông nam của TBTBD hoạt động mạnh (yếu) thì XTNĐ Biển Đông ít (nhiều). Trong khi, hoạt động của rãnh gió mùa bị ảnh hưởng nhiều bởi ENSO. Cụ thể, trong những năm El Niño, di thường dòng gió tây mực thấp (850-hPa) thịnh hành, thổi từ khu vực biển Philippines tới trung tâm TBD. Trong khi đó, phần vĩ độ cao hơn là hệ thống cao cận nhiệt đới. Gió từ rìa đông nam (tín phong) của cao cận nhiệt đới hợp với dòng gió tây ở khu vực gần xích đạo làm cho rãnh gió mùa hoạt động mạnh hơn bình thường. Rãnh gió mùa hoạt động mạnh tạo điều kiện thuận lợi cho XTNĐ hình thành nhiều hơn ở phần đông nam của TBTBD [65]. Tuy nhiên, trong những năm El Niño, một số nghiên cứu đã chỉ ra rằng, có nhiều hơn gấp 2,5 lần so với bình thường những cơn bão có quỹ đạo bị bẻ cong và di chuyển lên phía Nhật Bản, Hàn Quốc. Điều này có thể dẫn tới XTNĐ Biển Đông ít đi [26, 30]. Tương quan dương nhận được trên khu vực Biển Đông hàm ý, xoáy tương đối hoạt động mạnh dẫn tới XTNĐ Biển Đông nhiều, điều này phù hợp với nguyên tắc vật lý. So sánh với tháng 8-9-10, các tháng 2-3-4-5 có bản đồ tương quan tịnh tiến về phía đông. Đặc biệt, tương quan dương xuất hiện rõ nét hơn ở khu vực Biển Đông và biển Phillipines. Có thể, vào thời điểm này của năm, nhân tố tại chỗ quyết định đến hoạt động của XTNĐ Biển Đông nhiều hơn các nhân tố từ xa. Sang thời điểm tháng 6, 7 tương quan dương xuất hiện nhiều hơn trên toàn khu vực TBTBD, tuy nhiên, trị số tương quan không cao.



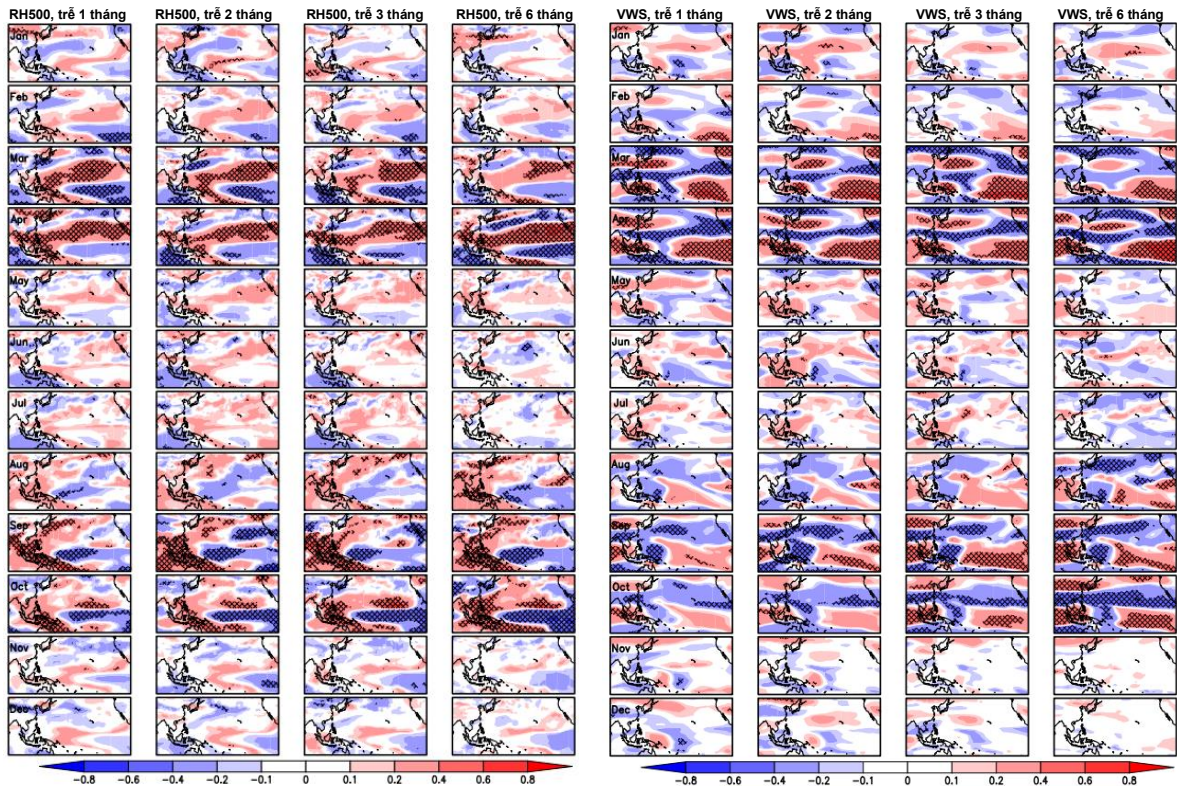
Hình 4.11 Bản đồ tương quan giữa trường xoáy tương đối mực 850mb (VOR850) với các độ trễ 1, 2, 3, 6 tháng và số lượng XTND Biển Đông ba tháng (vùng gạch chéo thể hiện giá trị tuyệt đối của hệ số tương quan  $\geq 0,4$  và thỏa mãn độ tin cậy thống kê 95%)

Hình 4.7 (trái) chỉ ra bản đồ tương quan dương mạnh giữa XTND Biển Đông với độ ẩm tầng đối lưu giữa (500-hPa) ở khu vực Biển Đông và biển Philippines

vào các tháng 2-3-4-5, 8-9-10. Điều này nghĩa là, nếu độ ẩm cao (thấp) ở Biển Đông và biển Philippines dẫn tới nhiều (ít) XTNĐ ở Biển Đông. Một số nghiên cứu đã đề cập đến ảnh hưởng của độ ẩm tầng đối lưu giữa đến sự phát triển lõi bão [85]. Vào năm El Niño, do tương tác của hệ thống cao cận nhiệt đới và dòng xiết (bộ phận của gió mùa Đông Á) làm cho vật chất hội tụ ở tầng khí quyển trên cao (ở mực khoảng 200-hPa) ở cung tây bắc của TBTBD (đông bắc biển Philippines). Điều này, dẫn tới phân kỳ ở mực dưới và làm tầng đối lưu giữa khô [85]. Đây có thể là một trong những nguyên nhân quan trọng làm cho số lượng XTNĐ ở cung tây bắc của TBTBD giảm đi. Hình thành và phát triển của XTNĐ ở cung đông nam của TBTBD có thể liên quan nhiều đến nguyên nhân động lực, trong khi phát triển của XTNĐ ở cung tây bắc của TBTBD có thể liên quan đến nhân tố nhiệt động lực [80]. Trong khi đó, hoạt động của XTNĐ ở cung tây bắc của TBTBD có liên quan chặt với hoạt động của XTNĐ Biển Đông. Điều này có thể giải thích cho sự xuất hiện nhiều (ít) XTNĐ trong Biển Đông trong năm La Niña (El Niño).

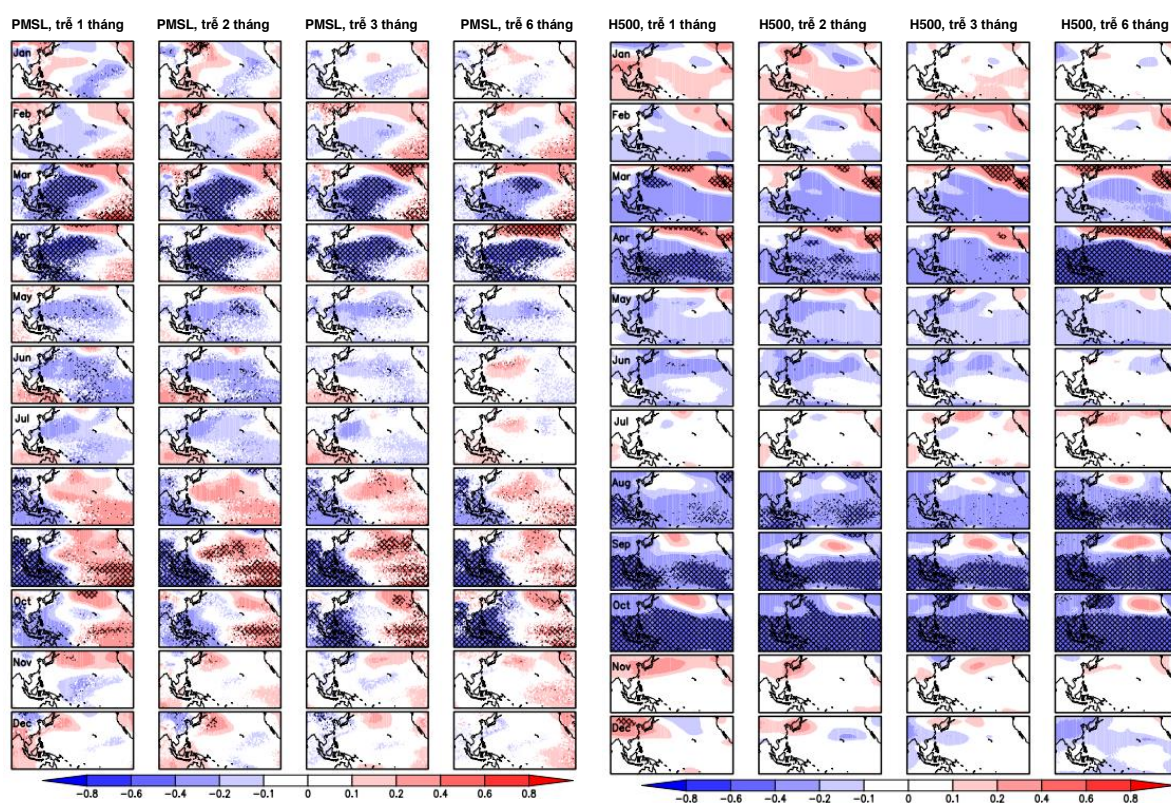
Hình 4.7 (phải) chỉ ra phân bố không gian của tương quan giữa tổng bão 3 tháng và trường độ đứt gió (được tính toán qua thành phần gió mực 200-hPa và 850-hPa). Trong thời gian tháng 2-3-4-5, tháng 8-9-10-11, bản đồ tương quan thể hiện rõ các đới vĩ hướng ở lân cận Biển Đông, biển Philippines và trên lục địa phía nam Trung Quốc. Dải tương quan âm chạy song song theo phương vĩ hướng từ Biển Đông ra qua biển Philippines ngụ ý, độ đứt gió giảm dẫn tới nhiều XTNĐ trong Biển Đông và ngược lại. Hình thế này có thể liên quan đến gió mùa mùa hè Đông Á. Nghiên cứu của Wang và cộng sự (2012) đã chỉ ra sự tương tác giữa dòng xiết trên cao (bộ phận của gió mùa Đông Á) với cao cận nhiệt đới TBTBD có ảnh hưởng đến số lượng XTNĐ trong Biển Đông với qui mô thập kỷ [88]. Kết quả nghiên cứu chỉ ra rằng, cường độ của hệ thống dòng xiết ở tầng đối lưu ảnh hưởng quan trọng đến tạo XTNĐ trong Biển Đông. Cường độ yếu của hệ thống dòng xiết trong thời kỳ 1965-1973 và 1995-2002 dẫn suy giảm độ đứt gió ở khu vực Biển Đông. Điều này cung cấp điều kiện thuận lợi cho phát triển XTNĐ Biển Đông. Tuy nhiên, hình thế này thể hiện không rõ ràng trong một số tháng khác như tháng 6, 7. Như vậy, tương quan của số

lượng XTNĐ với trường độ đứt gió cung cấp thêm thông tin về ảnh hưởng của hệ thống gió mùa Đông Á đến hoạt động của bão Biển Đông.



Hình 4.12 Bản đồ tương quan giữa trường RH500 (trái), VWS (phải) tại các độ trễ 1, 2, 3, 6 tháng với số lượng XTNĐ Biển Đông ba tháng (vùng gạch chéo thể hiện giá trị tuyệt đối của hệ số tương quan  $\geq 0,4$  và thỏa mãn độ tin cậy thống kê 95%)

Bản đồ tương quan với áp suất mực mặt biển (PMSL), độ cao địa thế vị mực 500hPa (H500) cũng thể hiện xu thế tương tự (Hình 4.13). Các vùng tương quan rõ ràng hơn trong các tháng chuyển mùa (tháng 3, 4, 9, 10). Phần tây bắc (đông nam) của biển TBD có tương quan âm (dương). Điều này nghĩa là, trong phần tây bắc của TBD, dị thường áp thấp (cao) sẽ liên quan đến nhiều (ít) XTNĐ trong Biển Đông. Điều ngược lại xảy ra với phần đông nam của TBD. Đặc biệt, trong tháng 8, 9, 10. Khu vực có tương quan dương xuất hiện ở phía đông bắc của biển Philippines. Vị trí bản đồ tương quan giống như lưỡi áp cao cận nhiệt đới. Về mặt vật lý, áp cao cận nhiệt đới hoạt động mạnh sẽ đóng vai trò quan trọng trong việc dẫn đường cho những cơn bão hình thành ở ngoài khơi TBTBD, sau đó di chuyển vào Biển Đông.



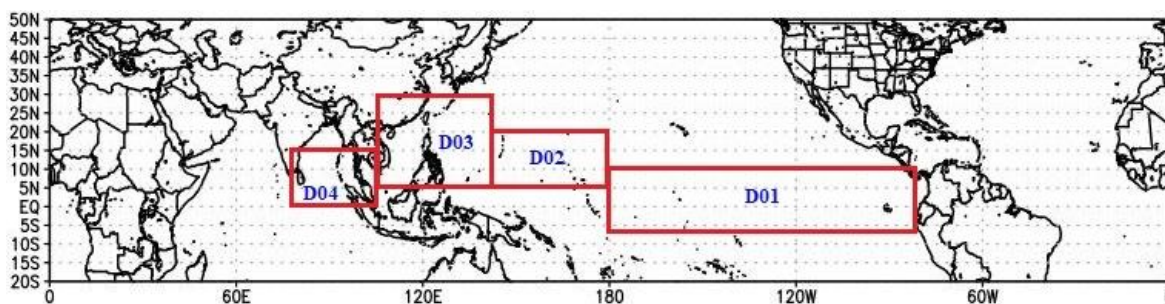
Hình 4.13 Bản đồ tương quan giữa trường PMSL (hình trái), H500 (hình phải) tại các độ trễ 1, 2, 3, 6 tháng với số lượng XTNĐ Biển Đông ba tháng (vùng gạch chéo thể hiện giá trị tuyệt đối của hệ số tương quan  $\geq 0,4$  và thỏa mãn độ tin cậy thống kê 95%)

Như vậy, nhìn chung các bản đồ tương quan cho thấy ở 12 thời điểm dự báo trong năm tồn tại nhiều vùng cho tín hiệu tương quan cao với số lượng XTNĐ hoạt động ở Biển Đông. Nhìn chung, các vùng trên bản tương quan có giá trị tuyệt đối cao thường chủ yếu tập trung ở các trung tâm tác động hoặc các vùng Niño có ảnh hưởng tới sự hình thành, phát triển của XTNĐ trên khu vực TBTBD nói chung và Biển Đông nói riêng. Có thể thấy rằng, đây cũng là các vùng đã được nhiều nhà khoa học lựa chọn khi nghiên cứu dự báo XTNĐ trên khu vực TBTBD bằng phương pháp thống kê động lực [49-54, 58, 83]. Kết quả này là cơ sở cho việc lựa chọn bộ nhân tố dự báo từ sản phẩm của mô hình CFSv2 phục vụ cho việc xây dựng phương trình dự báo số lượng XTNĐ hoạt động ở Biển Đông.

#### 4.2.2. Lựa chọn, thiết lập bộ nhân tố dự tuyến và nhân tố dự báo

Những phân tích về quan hệ tương quan giữa XTNĐ 3 tháng với các trường sản phẩm CFSv2 ở trên đã đưa ra cơ sở và định hướng trong việc tuyến chọn nhân

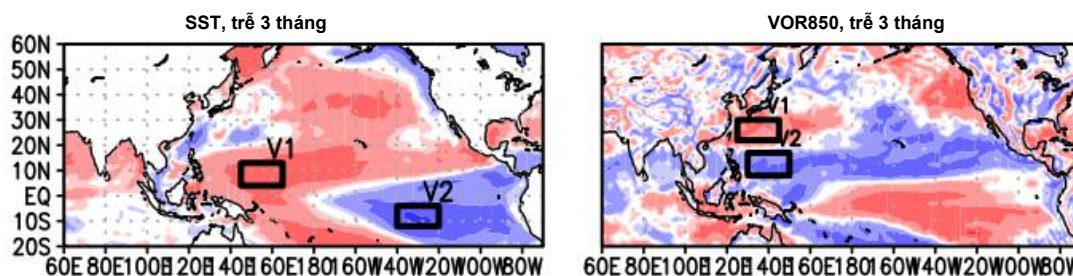
tổ dự báo cho các mô hình thống kê. Phân bố không gian của hệ số tương quan giữa hoạt động của XTNĐ Biển Đông với trường SST có dạng hình vành khuyên giống hình thể ENSO thể hiện tương quan âm (dương) ở phần phía đông (tây) khu vực trung tâm TBD nhiệt đới. Đặc biệt các vùng tương quan rõ ràng hơn trong các tháng chuyển mùa như tháng 8-9-10-11. Trong khu vực Ấn Độ Dương nhiệt đới, tương quan nhận được giá trị âm. Bản đồ tương quan âm ở phía đông và trung tâm TBD và Ấn Độ Dương, và tương quan dương ở khu vực Tây Thái Bình Dương nhiệt đới thể hiện quan hệ ENSO, IOD và XTNĐ Biển Đông [23]. Tương quan với trường xoáy tương đối mực thấp (VOR850) thể hiện với quan hệ với hoạt động của rãnh gió mùa trên TBTBD [94]. Trong khi bản đồ tương quan của độ ẩm tầng đối lưu giữa có thể liên quan đến ảnh hưởng kết hợp của ENSO và gió mùa Đông Á (dòng xiết Đông Á). Từ các phân tích trên, luận án định hướng các khu vực lựa chọn nhân tố dự tuyển như được minh họa trên Hình 4.14.



Hình 4.14 Các khu vực định hướng lựa chọn nhân tố dự tuyển từ sản phẩm CFSv2

Theo bản đồ Hình 4.14, khu vực D01 bao gồm thông tin ENSO (ấm lên hướng đông, hướng tây và trung tâm Thái Bình Dương nhiệt đới). Khu vực D04 đánh giá hoạt động của IOD (nhánh khác của hoàn lưu Walker), D04 cũng thể hiện thông tin về hoạt động của gió mùa tây nam, ảnh hưởng quan trọng đến thời điểm bùng nổ mùa bão. Khu vực D02 đánh giá thông tin hoạt động của rãnh gió mùa trên TBTBD. Cuối cùng, khu vực D03 đánh giá trực tiếp điều kiện môi trường đại dương - khí quyển trong Biển Đông, nó cũng bao gồm một bộ phận của gió mùa Đông Á (rãnh Đông Á). Khu vực lựa chọn bên trên với mục đích bao hàm được những nhân tố quan trọng điều khiển và ảnh hưởng tới hoạt động của XTNĐ trên khu vực Biển Đông.

Phù hợp với mục tiêu của luận án là dự báo hạn mùa (3 tháng) số lượng XTND hoạt động trên khu vực Biển Đông, luận án đã lựa chọn các sản phẩm hạn dự báo 3 tháng (leadtime bằng 3) của mô hình CFSv2 để tiến hành việc khoanh các tiểu vùng biểu thị giá trị tương quan cao, bao quát trên một không gian đủ lớn để xác định các nhân tố dự tuyển cho bài toán dự báo thống kê. Tùy theo từng biến trường sản phẩm CFSv2, các vùng tương quan cao được lựa chọn theo các khu vực không gian định hướng tại Hình 4.14. Theo đó, các vùng được chọn đảm bảo bao quát vùng có giá trị tương quan tuyệt đối lớn hơn khoảng 0,2 trên bản đồ tương quan, với phân bố không gian đủ rộng (khoảng 5 độ kinh, vĩ) để đảm bảo tính ổn định. Tùy theo phân bố trên các bản đồ tương quan, theo các vùng định hướng đã chỉ ra thì số lượng các tiểu vùng được lựa chọn có thể là 0, 1 hoặc 2 tiểu vùng (bằng 0 là vùng định hướng không thể hiện tương quan tức là không chọn được tiểu vùng nào, tương tự bằng 1 (2) tương ứng với vùng định hướng chỉ thể hiện duy nhất 1 tương quan âm hoặc dương (thể hiện cả hai vùng tương quan dương, âm)). Ví dụ như với biến SST tại thời điểm tháng 3, bản đồ tương quan trên Hình 4.10 gợi ý việc lựa chọn 2 tiểu vùng: 1) vùng tương quan dương (trong khoảng từ 125° đến 165°E, 0° đến 13°N) thuộc vùng thông tin về điều kiện môi trường đại dương – khí quyển khu vực Biển Đông (D03); 2) vùng tương quan âm (200° đến 260°W; 8°S đến 4°N) thuộc vùng thông tin về hiện tượng ENSO (D01). Hoặc như biến VOR850, cũng tại thời điểm tháng 3 có thể lựa chọn 2 tiểu vùng trong khu vực lân cận vùng định hướng D03, như minh họa tại Hình 4.15 dưới đây.

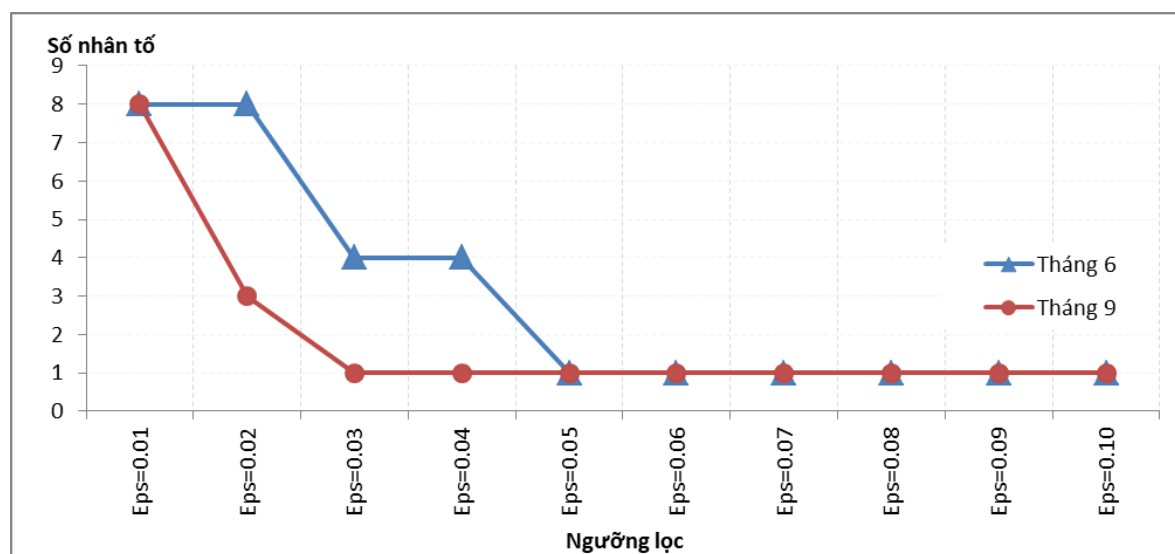


Hình 4.15 Bản đồ tương quan giữa số lượng XTND Biển Đông 3 tháng với trường SST, VOR850 từ sản phẩm của CFSv2 và lựa chọn nhân tố dự tuyển cho tháng 3

Như vậy, tương ứng với mỗi thời điểm dự báo, các nhân tố dự tuyển được xác định bằng giá trị trung bình vùng của từng biến trường được lựa chọn theo cách

trên. Theo cách này, ứng với từng tháng dự báo bộ nhân tố dự tuyển của TH3 có số lượng tối đa 14 biến (7 biến trường sản phẩm CFSv2: SST, PMSL, VOR, VWS, VVEL, RH, H500; mỗi biến trường lấy tối đa 2 vùng tương quan cao). Bộ nhân tố dự tuyển này sau đó được tiếp tục tuyển chọn bằng kỹ thuật hồi quy từng bước như đã trình bày ở Chương 2 để tìm ra được bộ nhân tố dự báo phù hợp cho 4 mô hình thống kê.

Qua việc khảo sát, lựa chọn nhân tố dự báo cho TH2 như đã trình bày tại tiêu mục 4.1.1 cho thấy, việc xác định ngưỡng lọc Epsilon trong kỹ thuật hồi quy từng bước sẽ quyết định số lượng các nhân tố dự báo được giữ lại. Hình 4.16 dưới đây trình bày số lượng các nhân tố dự báo cho thời điểm tháng 6 (dự báo 3 tháng đầu mùa bão) và tháng 9 (dự báo cho 3 tháng cuối mùa bão) của TH3 (được lọc từ bộ 14 nhân tố dự tuyển) theo các ngưỡng lọc (tham số Epsilon) khác nhau từ 0,01 cho tới 0,10 trong kỹ thuật hồi quy từng bước.



Hình 4.16 Số lượng nhân tố dự báo được lựa chọn từ tập các nhân tố dự tuyển từ sản phẩm của mô hình CFSv2

Kết quả tại Hình 4.16 cho thấy, ở ngưỡng lọc 0,01 số lượng nhân tố dự báo cho thời điểm tháng 6 và 9 giống nhau và bằng 8, tuy nhiên khi Epsilon tăng lên 0,02 thì số nhân tố dự báo cho thời điểm tháng 9 giảm xuống chỉ còn 3 nhân tố trong khi số lượng nhân tố cho thời điểm tháng 6 vẫn là 8 nhân tố. Việc gia tăng

Epsilon lên 0,03; 0,04 khiến số lượng nhân tố cho thời điểm dự báo tháng 6 còn 4 trong khi cho thời điểm dự báo tháng 9 chỉ còn 1 nhân tố. Với các ngưỡng lọc lớn hơn 0,04 kỹ thuật hồi quy từng bước chỉ còn tìm thấy 1 nhân tố dự báo cho cả thời điểm dự báo tháng 6 và tháng 9. Như vậy có thể thấy rằng, qua kỹ thuật hồi quy từng bước số lượng nhân tố dự báo của TH3 giảm nhanh theo mức độ gia tăng của ngưỡng lọc Epsilon. Để có cái nhìn trực quan hơn, các Bảng 4.16 và 4.17 dưới đây dẫn ra các nhân tố dự báo được lựa chọn cho 2 thời điểm dự báo tại tháng 6 và tháng 9 theo các ngưỡng lọc Epsilon khác nhau.

**Bảng 4.16 Các nhân tố dự tuyển được lựa chọn làm nhân tố dự báo cho các XTND hoạt động trên khu vực Biển Đông các tháng 6-7-8**

Ngưỡng lọc Epsilon	SST		PMSL		VOR		VWS		VVEL		RH		H500	
	V1	V2	V1	V2	V1	V2	V1	V2	V1	V2	V1	V2	V1	V2
0.01		x	x			x	x	x			x	x		
0.02		x	x			x	x	x			x	x		
0.03	x		x				x	x						
0.04	x		x				x	x						
0.05							x							
0.06							x							
0.07							x							
0.08							x							
0.09							x							
0.10							x							

**Bảng 4.17 Các nhân tố dự tuyển được lựa chọn làm nhân tố dự báo cho các XTND hoạt động trên khu vực Biển Đông các tháng 9-10-11**

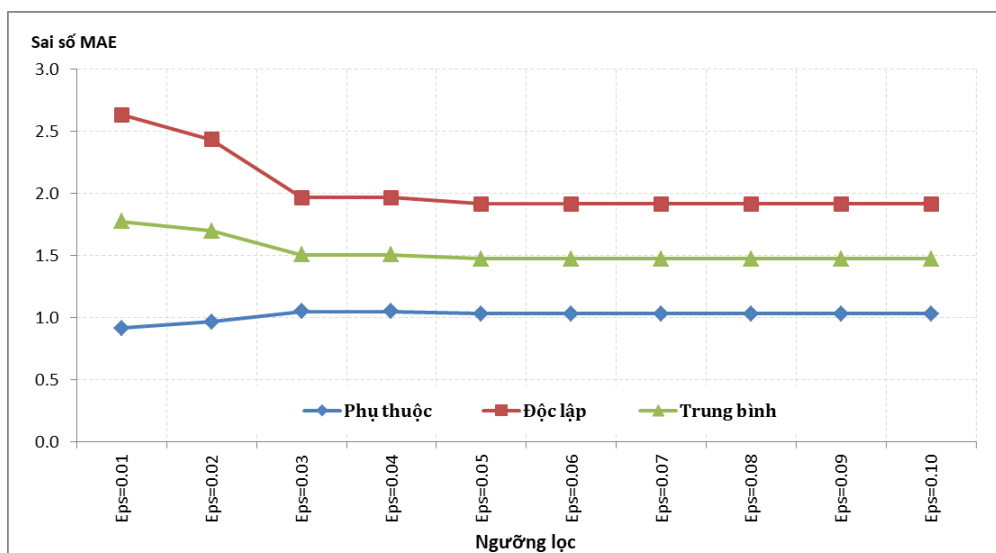
Ngưỡng lọc Epsilon	SST		PMSL		VOR		VWS		VVEL		RH		H500	
	V1	V2	V1	V2	V1	V2	V1	V2	V1	V2	V1	V2	V1	V2
0.01	x	x	x				x		x		x	x	x	
0.02									x		x	x		
0.03									x					
0.04									x					
0.05									x					
0.06									x					
0.07									x					
0.08									x					
0.09									x					
0.10									x					

Kết quả tại các Bảng 4.16 và 4.17 cho thấy, ở cả hai thời điểm dự báo tháng 6 và 9 nhóm nhân tố liên quan tới sự hình thành XTNĐ Biển Đông (SST, VOR, VWS, VVEL, RH) được lựa chọn nhiều hơn so với nhóm liên quan tới yếu tố dòng dẫn (HGT500, PMSL). Cá biệt ở thời điểm dự báo tháng 6, nhân tố VWS có mặt với tất cả các ngưỡng lọc Epsilon, trong khi PMSL, SST chỉ duy trì tới ngưỡng 0,04. Bên cạnh đó, tại thời điểm dự báo tháng 9, nhân tố VVEL có mặt ở tất cả các ngưỡng lọc khảo sát, trong khi hầu hết các yếu tố còn lại chỉ thỏa mãn với ngưỡng lọc 0,02. Như vậy, yếu tố gió thẳng đứng từ sản phẩm của CFSv2 có nhiều ảnh hưởng ý nghĩa tới cả các XTNĐ hoạt động trong 3 tháng đầu mùa bão và 3 tháng cuối mùa bão trên khu vực Biển Đông.

Như đã bàn luận ở phần 4.1 trên, đối với việc xây dựng các phương trình dự báo thống kê nói chung, việc sử dụng nhiều biến giải thích hay cụ thể với TH3 này thì việc đưa tăng thêm biến vào các phương trình dự báo đôi khi không đem lại kết quả dự báo tốt hơn. Để có cơ sở cho việc xác định số lượng nhân tố dự báo cần thiết tại đây luận án tiến hành khảo sát, chuỗi sản phẩm CFSv2 ở hai giai đoạn, giai đoạn luyện được lựa chọn là 1983-2009 và giai đoạn dự báo độc lập được tính toán trên chuỗi 2012-2016. Từ hai bộ số liệu này, luận án khảo sát, đánh giá sai số MAE với các bộ nhân tố dự báo được lựa chọn theo các ngưỡng lọc Epsilon khác nhau từ 0,01 tới 0,10 (như được chỉ ra tại Bảng 4.16 và 4.17). Kết quả về sai số MAE cho hai thời điểm dự báo tháng 6 và tháng 9 trên bộ số liệu phụ thuộc và độc lập được trình bày tại Hình 4.17 dưới đây.

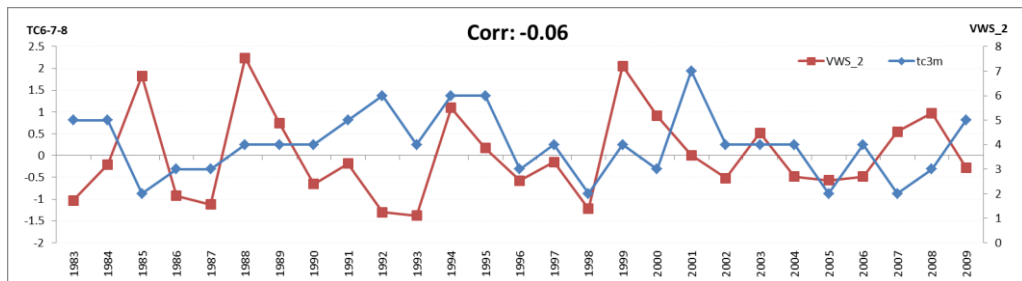
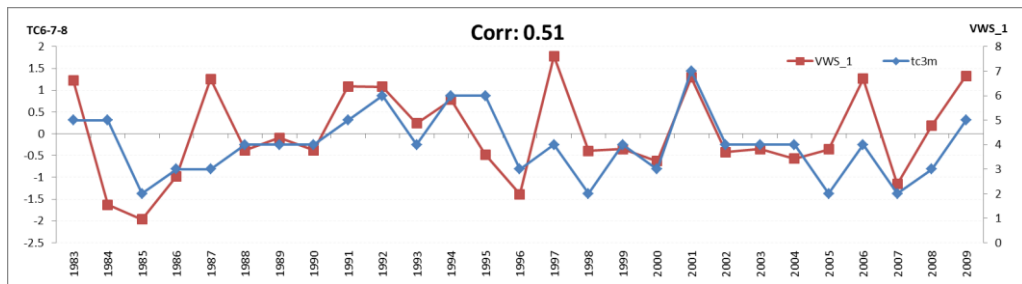
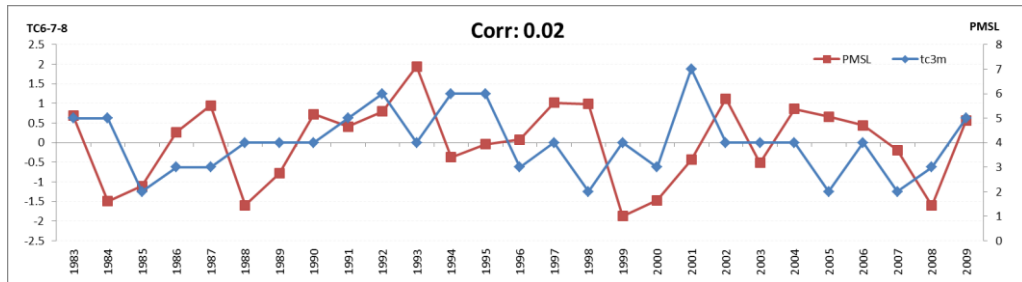
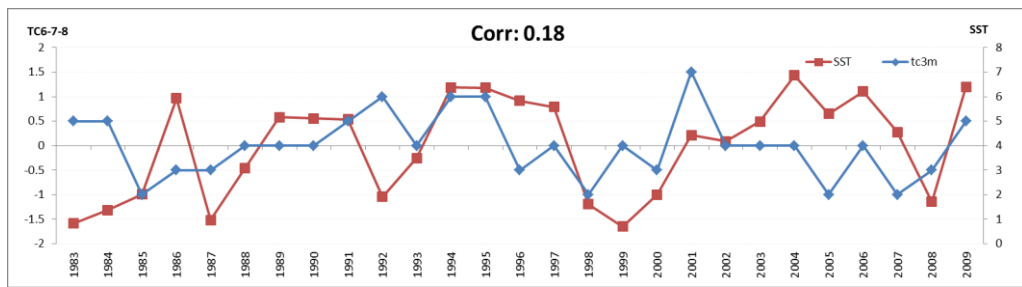
Kết quả tại Hình 4.17 cho thấy, ở ngưỡng lọc 0,01 sai số MAE của dự báo phụ thuộc ở mức 0,9 trong khi sai số này lên tới 2,6, ngưỡng lọc Epsilon tăng lên 0,02 giá trị MAE của hai giai đoạn này tương ứng là gần 1,0 và 2,5. Tại ngưỡng lọc 0,03 MAE cho giai đoạn phụ thuộc (độc lập) có giá trị 1,0 (2,0). Với ngưỡng lọc lớn hơn 0,02 trở đi, sai số MAE ở 2 giai đoạn không đổi do lúc này mỗi thời điểm dự báo chỉ còn lại 1 nhân tố dự báo. Đường đồ thị thể hiện diễn biến sai số MAE trung bình của dự báo phụ thuộc và độc lập trên Hình 4.17 (đường màu xanh, tam giác) cho thấy sự suy giảm dần theo sự gia tăng của ngưỡng lọc Epsilon từ 0,01 tới 0,03. Kết quả này cho thấy ngưỡng lọc 0,03 là điểm “hội tụ” của sai số MAE ở dự báo phụ thuộc và độc

lập. Do vậy, nhân tố dự báo cho TH3 được xác định với ngưỡng lọc 0,03 và khi đó số nhân tố dự báo cho thời điểm tháng 6 để dự báo cho 3 tháng đầu mùa bão (tháng 6-7-8) gồm 4 biến (SST, PMSL, VWS\_1, VWS\_2), trong khi nhân tố dự báo cho thời điểm tháng 9 để dự báo cho 3 tháng cuối mùa bão chỉ gồm 1 biến (VVEL).

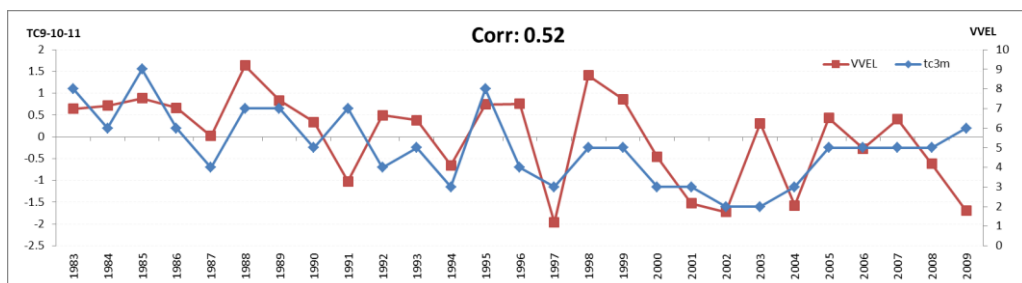


Hình 4.17 Sai số MAE trên bộ số liệu phụ thuộc (1983-2009) và độc lập (2012-2016) của TH3 theo các bộ nhân tố dự báo (ngưỡng lọc Epsilon) khác nhau

Từ bộ nhân tố dự báo cho TH3 đã được lựa chọn ở trên, các Hình 4.18 và Hình 4.19 dưới đây lần lượt dẫn ra các đồ thị biểu thị mối quan hệ giữa các nhân tố được lựa chọn với XTND Biển Đông 3 tháng đầu mùa bão và 3 tháng cuối mùa bão. Trong số các nhân tố dự báo được lựa chọn cho thời điểm tháng 6, độ đứt gió đóng góp 2 thành phần, một thành phần thể hiện tương quan (dương) mạnh với hệ số 0,51, thành phần thứ hai thể hiện tương quan (âm) yếu với hệ số -0,06. Bên cạnh đó, yếu tố sinh XTND trên khu vực TBTBD nói chung và Biển Đông nói riêng là SST cũng thể hiện tương quan dương với giá trị không lớn (0,18) và yếu tố môi trường PMSL thậm chí ở mức không đáng kể (0,02). Ở thời điểm dự báo tháng 9, mặc dù chỉ còn tồn tại duy nhất 1 nhân tố là VVEL song lại thể hiện quan hệ tương quan với XTND Biển Đông 3 tháng cuối mùa bão tương đối lớn, đạt mức 0,52. Từ bộ nhân tố dự báo đã được lựa chọn, phần dưới đây sẽ xem xét khả năng của của các mô hình thống kê trong việc dự báo số lượng XTND trên khu vực Biển Đông trong 3 tháng đầu và cuối mùa bão.



Hình 4.18 Mối quan hệ giữa các nhân tố dự báo với XTND Biển Đông 3 tháng 6-7-8



Hình 4.19 Mối quan hệ giữa nhân tố dự báo với XTND Biển Đông 3 tháng 9-10-11

### 4.2.3 Kết quả dự báo số lượng XTNĐ hoạt động trên khu vực Biển Đông bằng phương pháp thống kê động lực theo TH3

Với bộ nhân tố dự báo cho TH3 đã được lựa chọn bên trên, 4 mô hình thống kê (MLR, LAD, LMV, ANN) và mô hình tổ hợp trung bình của chúng được áp dụng để dự báo số lượng XTNĐ trên khu vực Biển Đông. Lưu ý rằng, các nhân tố dự báo chính thức ở trên được chuẩn hóa trước khi đưa vào xây dựng các phương trình dự báo. Các kết quả dự báo và đánh giá sai số, kỹ năng cho TH3 lần lượt được trình bày tại các Bảng 4.18 đến Bảng 4.21 và Hình 4.20 dưới đây.

Bảng 4.18 Hệ số hồi quy của các mô hình dự báo thống kê của TH3 cho tháng 6

Mô hình thống kê	Các hệ số hồi quy					Pval
	b0	SST	PMSL	VWS_1	VWS_2	
MLR	3.999950	0.384320	-0.87483	0.824688	-0.53044	0.001
LAD	4.117835	0.059892	-0.17434	0.843036	-0.02287	0.003
LMV	3.902123	0.799659	-1.77752	0.842498	-1.32798	0.002

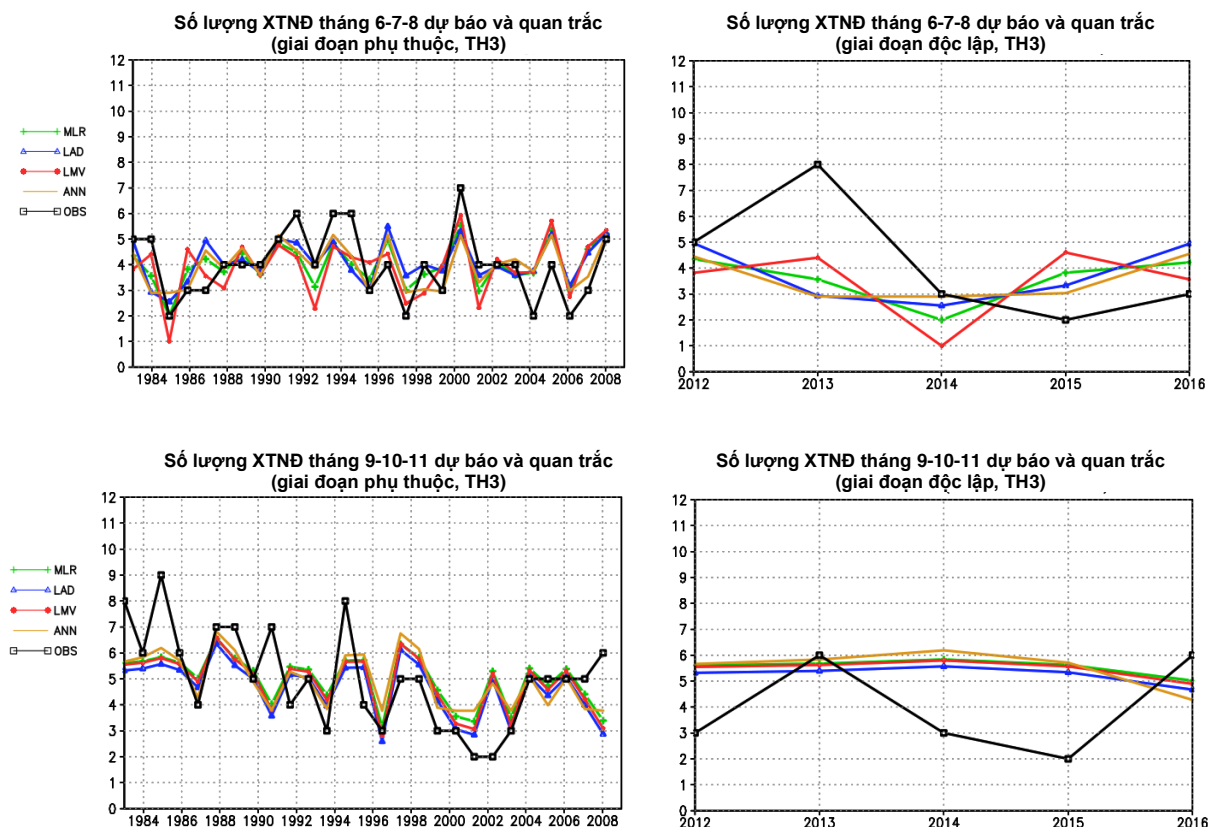
Bảng 4.19 Hệ số hồi quy của các mô hình dự báo thống kê của TH3 cho tháng 9

Mô hình thống kê	Các hệ số hồi quy		Pval
	b0	VVEL	
MLR	4.999997	0.948125	0.006
LAD	4.648608	1.044878	0.006
LMV	4.873152	1.048327	0.006

Bảng 4.18 và Bảng 4.19 ở trên dẫn ra hệ số hồi quy của các phương trình dự báo thống kê tương ứng cho thời điểm tháng 6 và tháng 9 của 3 mô hình. Nhìn chung, hệ số  $p_{val}$  của các mô hình cho thấy các phương trình dự báo xây dựng được đều thỏa mãn độ tin cậy thống kê 95%, thêm vào đó hệ số tự do  $b_0$  của các phương trình đều xấp xỉ với giá trị trung bình khí hậu của yếu tố dự báo. Dấu âm (dương) của các hệ số beta ( $b_1, b_2, \dots$ ) phù hợp với quan hệ tương quan âm (dương) của nhân tố dự báo với yếu tố dự báo như đã trình bày tại Hình 4.18 và Hình 4.19.

Hình 4.20 dưới đây trình bày các bản đồ số lượng XTNĐ Biển Đông 3 tháng đầu mùa và 3 tháng cuối mùa bão theo số liệu quan trắc RSMC và theo các kết quả

dự báo từ các mô hình dự báo đã xây dựng được ở trên. Kết quả này cho thấy, tổng thể chung các mô hình đã nắm bắt được xu thế so với số liệu quan trắc ở cả hai thời điểm dự báo, tuy vậy vẫn còn tồn tại những năm mô hình đưa ra kết quả dự báo sai hoặc dự báo không. Cụ thể, ở giai đoạn dự báo phụ thuộc, đối với số lượng XTND 3 tháng đầu mùa bão các mô hình đưa ra kết quả sai ở những năm 1992, 1994, 1995 và dự báo không ở những năm 2004, 2006, 2007. Trong khi đối với số lượng XTND 3 tháng cuối mùa bão mô hình cho ra các kết quả sai ở những năm 1983, 1985, 1991, 1995 và dự báo không ở năm 2001, 2002. Tuy nhiên có thể thấy là khoảng cách sai lệch giữa giá trị quan trắc và số lượng dự báo của mô hình là không lớn, chênh lệch khoảng 1 con (cá biệt ở kết quả dự báo cho 3 tháng cuối mùa bão, mô hình cho kết quả dự báo sai tới 3 con ở các năm 1985, 1991 và sai 2 con ở năm 1995).



Hình 4.20 Số lượng XTND 3 tháng 6-7-8 (hình trên) và 9-10-11 (hình dưới) trên Biển Đông giai đoạn 1983-2009 và giai đoạn độc lập 2012-2016, số lượng quan trắc (OBS) và kết quả dự báo các mô hình MLR, LAD, LMV, ANN theo TH3

Ở giai đoạn dự báo kiểm nghiệm độc lập, dường như kết quả dự báo cho 3 tháng đầu mùa bão tốt hơn khi chỉ sai lệch với số liệu quan trắc trên dưới 1 cơn, ngoại trừ sai lệch tới trên 4 cơn ở năm 2013 (năm có nhiều XTNĐ trên Biển Đông bất thường, 17 cơn so với trung bình khí hậu 10,9 cơn). Bên cạnh đó, kết quả dự báo XTNĐ cho 3 tháng cuối mùa bão thể hiện tương đối giống nhau giữa các mô hình, điều này xuất phát từ việc các phương trình cho thời điểm dự báo này chỉ có duy nhất 1 biến phụ thuộc. So với số liệu quan trắc, có 2 năm (2013 và 2016) các mô hình đưa ra dự báo đúng trong khi 3 năm còn lại (2012, 2014, 2015) các mô hình đưa ra các kết quả dự báo không từ 2 tới 3 cơn.

Bảng 4.20 và Bảng 4.21 dưới đây dẫn ra các sai số và chỉ số kỹ năng dự báo của các mô hình cho cả giai đoạn dự báo trên bộ số liệu phụ thuộc (1983-2009) và giai đoạn dự báo kiểm nghiệm độc lập (2012-2016). Chú ý rằng, do đặc điểm của bộ số liệu CFSv2 khảo sát mà các khoảng thời gian dự báo phụ thuộc và độc lập ở TH3 ngắn hơn so với 2 trường hợp đầu.

Bảng 4.20 Sai số và chỉ số đánh giá kỹ năng dự báo tại thời điểm tháng 6 ở giai đoạn dự báo phụ thuộc (1983-2009) và độc lập (2012-2016) theo TH3

Thời điểm dự báo Yếu tố dự báo (Y)	Tháng 6									
	XTNĐ Biển Đông 3 tháng đầu mùa bão (các tháng 6-7-8)									
	Giai đoạn phụ thuộc					Giai đoạn độc lập				
Trung bình Y	4					4.2				
Độ lệch chuẩn	1.3					2.1				
Chỉ số đánh giá	Mô hình thống kê									
	MLR	LAD	LMV	ANN	E1234	MLR	LAD	LMV	ANN	E1234
ME:	0	0.1	-0.1	0	0	-0.6	-0.5	-0.7	-0.6	-0.6
MAE:	0.9	0.8	1	0.8	0.8	1.8	1.8	2	1.7	1.8
RMSE:	1	1.1	1.1	1	1	2.3	2.5	2.3	2.4	2.3
RMAE:	24.2	24.1	28.4	22	23.5	46.7	42.1	56.8	36.3	45.5
CORR:	0.62	0.55	0.57	0.64	0.65	0.13	-0.11	0.30	-0.13	0.09
MSSS:	0.40	0.30	0.26	0.42	0.44	-0.12	-0.38	-0.12	-0.31	-0.14

Kết quả đánh giá đối với các dự báo số lượng XTNĐ 3 tháng đầu mùa bão trên Bảng 4.20 cho thấy, sai số ME ở giai đoạn phụ thuộc thấp ở mức xấp xỉ 0 và hạ xuống khoảng -0,6 ở giai đoạn dự báo kiểm nghiệm. Tương tự vậy, sai số MAE và

RMSE trung bình ở mức 0,9 và 1,0 (1,8 và 2,4) ở giai đoạn mô hình luyện (dự báo kiểm nghiệm độc lập). Đáng chú ý, các sai số ở giai đoạn phụ thuộc thấp hơn so với độ lệch chuẩn của yếu tố dự báo trong khi lại cao hơn 1 chút ở giai đoạn dự báo kiểm nghiệm độc lập. Chỉ số RMAE đánh giá mức độ sai số tuyệt đối với độ lệch chuẩn cũng phản ánh điều này, có giá trị thấp với mức khoảng 25% ở giai đoạn luyện của các mô hình và tăng lên trên dưới 45% ở dự báo độc lập. Hệ số tương quan giữa số liệu quan trắc và dự báo của mô hình cùng với chỉ số kỹ năng cho thấy ở giai đoạn luyện tất cả các mô hình đều có kỹ năng mô phỏng số lượng XTNĐ 3 tháng đầu mùa bão, song ở giai đoạn dự báo kiểm nghiệm độc lập các kỹ năng này thể hiện rất yếu thậm chí còn kém hơn so với dự báo trung bình khí hậu. Cũng cần nhắc lại rằng, bộ số liệu kiểm nghiệm cho TH3 khá ngắn, chỉ có 5 năm nhưng riêng số lượng bão cao bất thường ở năm 2013 đóng góp sai số tới 4 cơn đã khiến các chỉ số đánh giá kỹ năng cho giai đoạn độc lập bị ảnh hưởng đáng kể.

Bảng 4.21 Sai số và chỉ số đánh giá kỹ năng dự báo tại thời điểm tháng 9 ở giai đoạn dự báo phụ thuộc (1983-2009) và độc lập (2012-2016) theo TH3

Thời điểm dự báo Yếu tố dự báo (Y)	Tháng 9									
	XTNĐ Biển Đông 3 tháng cuối mùa bão (các tháng 9-10-11)									
	Giai đoạn phụ thuộc					Giai đoạn độc lập				
Trung bình Y	5					4				
Độ lệch chuẩn	1.8					1.7				
Chỉ số đánh giá	Mô hình thống kê									
	MLR	LAD	LMV	ANN	E1234	MLR	LAD	LMV	ANN	E1234
ME:	0	-0.4	-0.1	0	-0.1	1.6	1.3	1.5	1.5	1.5
MAE:	1.2	1.2	1.2	1.2	1.2	2.1	2	2.1	2.3	2.1
RMSE:	1.6	1.6	1.6	1.5	1.5	2.4	2.2	2.4	2.6	2.4
RMAE:	28.9	25.4	27.3	27.8	26.9	76.9	72.5	76.3	82.4	77
CORR:	0.52	0.52	0.52	0.57	0.54	-0.56	-0.56	-0.56	-0.56	-0.56
MSSS:	0.26	0.23	0.26	0.33	0.28	-0.51	-0.32	-0.50	-0.80	-0.52

Kết quả sai số và kỹ năng dự báo đối với các XTNĐ hoạt động trong 3 tháng cuối mùa bão mang nhiều nét tương đồng so với XTNĐ 3 tháng đầu mùa bão (Bảng 4.20). Sai số ME ở giai đoạn phụ thuộc nhỏ xấp xỉ 0 ở các mô hình tuy nhiên biến động mạnh lên tới khoảng 1,5 ở giai đoạn độc lập. Trên bộ số liệu phụ thuộc, sai số

MAE (RMSE) trung bình khoảng 1,2 (1,6) thấp hơn so với độ lệch chuẩn khí hậu 1,8 trong khi đó các giá trị này tương ứng ở giai đoạn dự báo kiểm nghiệm độc lập là 2,1 (2,4) cao hơn so với độ lệch chuẩn khí hậu giai đoạn này (1,7). Diễn biến của các sai số này phù hợp với chỉ số RMAE, khi dao động trong khoảng trung bình 27% ở giai đoạn luyện của mô hình và biến động tăng lên 75% trên bộ số liệu độc lập. Hệ số tương quan và chỉ số kỹ năng dự báo của cả 5 mô hình đều đưa ra giá trị dương tương ứng với mức trung bình trên 0,5 và 0,3 khẳng định khả năng mô phỏng số lượng các XTNĐ hoạt động trong 3 tháng cuối mùa bão. Tuy vậy, kỹ năng dự báo này thể hiện kém ở giai đoạn dự báo kiểm nghiệm độc lập, đây cũng là vấn đề gặp phải như đối với kết quả dự báo cho các XTNĐ đầu mùa bão.

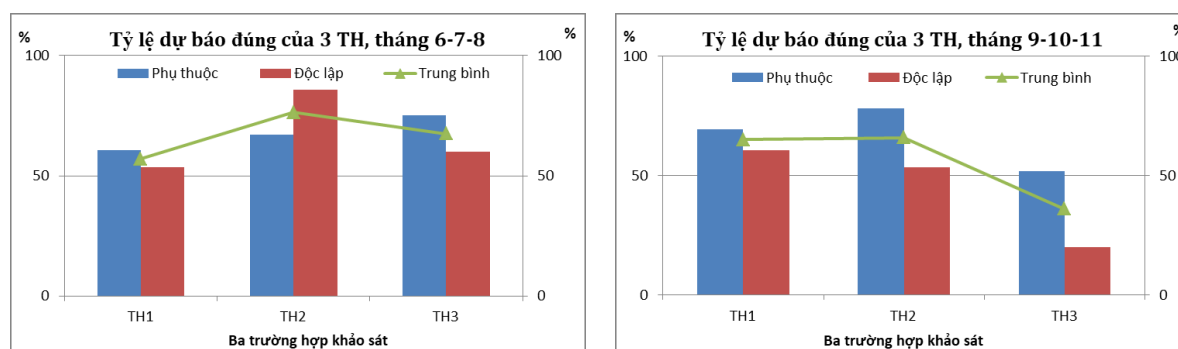
Như vậy, qua kết quả đánh giá sai số và chất lượng của các mô hình khảo sát ở TH3 đã cho thấy, nhìn chung các mô hình đều thể hiện có kỹ năng mô phỏng số lượng XTNĐ 3 tháng đầu mùa bão và 3 tháng cuối mùa bão giai đoạn 1983-2009 với sai số thấp hơn độ lệch chuẩn khí hậu. Tuy vậy ở giai đoạn dự báo kiểm nghiệm độc lập tất cả các mô hình đều không đạt điểm kỹ năng và sai số của các mô hình phần lớn xấp xỉ và vượt so với độ lệch chuẩn khí hậu của thời kỳ. Tuy nhiên khi xem xét kết quả này thì cần lưu ý rằng, bộ số liệu kiểm nghiệm độc lập cho TH3 khá ngắn, với biến động số lượng bão mạnh, tương ứng độ lệch chuẩn 2,1 (1,7) con so với trung bình khí hậu 4,1 (4,0) con đối với 3 tháng đầu (cuối) mùa bão đã khiến các chỉ số đánh giá kỹ năng cho giai đoạn độc lập bị ảnh hưởng đáng kể.

Để có bức tranh chung về sai số và chất lượng của cả 3 trường hợp khảo sát (TH1, TH2, TH3), Bảng 4.22 và Hình 4.21 dưới đây dẫn ra tỷ lệ đúng được tính toán từ dự báo trên chuỗi số liệu phụ thuộc và độc lập tương ứng để dự báo số lượng XTNĐ hoạt động 3 tháng đầu mùa và 3 tháng cuối mùa bão. Tỷ lệ đúng tính toán cho 2 pha (trên chuẩn và dưới chuẩn). Số lượng XTNĐ dự báo được so sánh với giá trị trung bình khí hậu. Nếu hiệu giữa quan trắc và giá trị trung bình khí hậu mang dấu dương (âm) và hiệu giữa dự báo và giá trị trung bình khí hậu mang dấu dương (âm) thì dự báo đúng. Các trường hợp khác được tính là dự báo sai.

Bảng 4.22 Kết quả đánh giá dự báo trên chuẩn/ dưới chuẩn so với trung bình khí hậu của 3 trường hợp

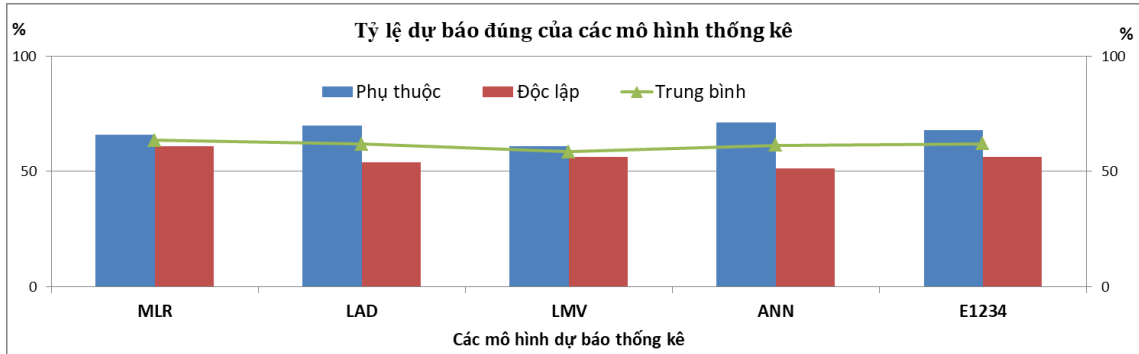
Tháng đứng dự báo	Trường hợp khảo sát	Training					Forecast				
		MLR	LAD	LMV	ANN	E1234	MLR	LAD	LMV	ANN	E1234
Tháng 6	TH1	67.7	67.7	35.5	71.0	58.1	<b>57.1</b>	42.9	<b>57.1</b>	<b>57.1</b>	42.9
	TH2	<b>64.5</b>	<b>71.0</b>	<b>64.5</b>	<b>67.7</b>	<b>71.0</b>	<b>85.7</b>	<b>85.7</b>	<b>85.7</b>	<b>85.7</b>	<b>85.7</b>
	TH3	74.1	74.1	74.1	77.8	74.1	60.0	60.0	60.0	60.0	60.0
Tháng 9	TH1	64.5	74.2	58.1	80.6	71.0	<b>71.4</b>	<b>71.4</b>	<b>57.1</b>	42.9	<b>57.1</b>
	TH2	<b>77.4</b>	<b>80.6</b>	<b>80.6</b>	<b>74.2</b>	<b>80.6</b>	<b>71.4</b>	42.9	<b>57.1</b>	42.9	<b>57.1</b>
	TH3	48.1	51.9	51.9	55.6	51.9	20.0	20.0	20.0	20.0	20.0

Tính trung bình chung, kỹ năng dự báo đánh giá cho 2 pha hầu hết đều trên ngưỡng 50% đối với TH1, TH2 và chỉ đạt khoảng 35% đối ở TH3 (Hình 4.21). Tính riêng đối với các dự báo số lượng XTNĐ 3 tháng đầu mùa bão thì TH2 đưa ra kết quả tốt nhất, tiếp đến là TH3 và TH1. Trong khi đó, nếu xem xét riêng các dự báo XTNĐ cho 3 tháng cuối mùa bão thì ngoài TH2 vẫn đưa ra kết quả tốt nhất thì chất lượng tốt kế tiếp là TH1 và kém nhất là TH3.



Hình 4.21 Tỷ lệ đúng của dự báo cho 3 tháng đầu mùa bão (tháng 6-7-8, hình trái) và 3 tháng cuối mùa bão (tháng 9-10-11, hình phải) theo 3 trường hợp (TH1, TH2, TH3)

Để xem xét kỹ năng của từng mô hình riêng lẻ và tổ hợp trung bình của chúng, Hình 4.22 dưới đây trình bày biểu đồ phản ánh tỷ lệ dự báo đúng trung bình ở cả 3 trường hợp so với dự báo trung bình khí hậu.



Hình 4.22 Tỷ lệ đúng (tỷ số giữa số lần dự báo đúng và tổng số lần dự báo; %) trung bình ba trường hợp khảo sát của các mô hình MLR, LAD, LMV, ANN và tổ hợp E1234

Kết quả tại Hình 4.22 cho thấy mặc tất cả 5 mô hình có tỷ lệ đúng trung bình ở cả dự báo phụ thuộc và độc lập đều trên 50% song mô hình MLR đã thể hiện tính ổn định tốt nhất giữa hai giai đoạn dự báo. Bên cạnh đó, sản phẩm tổ hợp E1234 cũng đưa ra những kết quả tương đối khả quan khi thể hiện được kỹ năng dự báo số lượng XTNĐ Biển Đông ở 3 tháng đầu mùa bão và 3 tháng cuối mùa bão theo TH1, TH2 (Bảng 4.22). Như vậy có thể thấy rằng, tùy theo thời điểm dự báo mà các mô hình riêng lẻ đều có thể cung cấp các thông tin hữu ích để từ đó có thể kết hợp với các phân tích, đánh giá khác nhằm đưa ra phương án dự báo tối ưu.

#### Nhận xét chung:

Trên cơ sở khảo sát mối quan hệ tương quan giữa các chỉ số khí hậu và số lượng XTNĐ 3 tháng theo các độ trễ khác nhau đã cho thấy, ở từng thời điểm cụ thể, các chỉ số khí hậu đã thể hiện tương quan tốt với XTNĐ 3 tháng tại khu vực Biển Đông. Bằng việc áp dụng phương pháp phân tích thành phần chính và hồi quy từng bước đối với bộ số liệu các chỉ số khí hậu khảo sát, luận án đã lựa chọn được hai bộ nhân tố dự báo phù hợp cho bài toán dự báo hạn mùa (3 tháng) số lượng XTNĐ Biển Đông (tương ứng với TH1, TH2). Song song với việc khai thác bộ các chỉ số khí hậu, qua khảo sát phân tích các bản đồ tương quan giữa sản phẩm của mô hình khí hậu CFSv2 với XTNĐ Biển Đông cũng đã cho thấy tồn tại nhiều vùng tương quan cao thể hiện tiềm năng trong việc xác định, lựa chọn làm nhân tố dự báo. Có thể thấy rằng, các vùng này chủ yếu là tập trung ở các trung tâm tác động,

hoặc các vùng Niño có ảnh hưởng tới sự hình thành, phát triển của XTNĐ trên khu vực TBTBD nói chung và Biển Đông nói riêng, đây cũng là các vùng đã được nhiều nhà khoa học lựa chọn trong các nghiên cứu dự báo XTNĐ TBTBD bằng phương pháp thống kê động lực. Đây là cơ sở và định hướng trong việc xác định các nhân tố dự tuyến đồng thời kết hợp với kỹ thuật hồi quy từng bước để tìm ra bộ nhân tố dự báo phù hợp cho bài toán dự báo hạn mùa số lượng XTNĐ trên Biển Đông (TH3). Kết quả đánh giá chất lượng các phương trình dự báo số lượng XTNĐ Biển Đông trong 3 tháng đầu mùa bão (tháng 6-7-8) và 3 tháng cuối mùa bão (tháng 9-10-11) đã cho thấy cả 4 mô hình thống kê khảo sát (MLR, LAD, LMV, ANN) cùng tổ hợp nhìn chung đã nắm bắt được tương đối tốt xu thế biến động của số lượng XTNĐ Biển Đông. Sai số và kỹ năng dự báo ở các trường hợp luận án khảo sát, đặc biệt là TH2 đã cho thấy những kết quả này tiệm cận và có thể so sánh được với một số kết quả đã công bố trong nước và quốc tế.

Kết quả đánh giá, phân tích sai số và kỹ năng dự báo trên bộ số liệu phụ thuộc và độc lập với bốn mô hình thống kê MLR, LAD, LMV, ANN và tổ hợp trung bình cho thấy, mặc dù tất cả các mô hình khảo sát đều thể hiện được kỹ năng dự báo song mô hình MLR có kết quả dự báo ổn định cùng với kỹ năng dự báo tốt nhất. Trong số 3 trường hợp khảo sát, TH2 đã thể hiện kỹ năng dự báo đối với cả các XTNĐ hoạt động trong 3 tháng đầu mùa bão và 3 tháng cuối mùa bão trong khi kết quả của TH1, TH3 không thực sự rõ ràng. Đáng lưu ý, kết quả của 3 trường hợp khảo sát còn cho thấy hạn chế trong việc nắm bắt được các biến đổi bất thường về số lượng XTNĐ Biển Đông chẳng hạn ở những năm nhiều XTNĐ như 2013, 2017, hay năm ít XTNĐ như 2015.

## KẾT LUẬN

Nghiên cứu đặc điểm hoạt động của XTNĐ Biển Đông trên bộ số liệu RSMC giai đoạn 1971-2019, đồng thời tiếp cận theo phương pháp thống kê với nhân tố dự báo được lựa chọn từ bộ 30 chỉ số khí hậu và các trường khí quyển, đại dương sẵn có từ sản phẩm của mô hình CFSv2.0 để dự báo số lượng XTNĐ 3 tháng đầu mùa bão (tháng 6-7-8) và 3 tháng cuối mùa bão (tháng 9-10-11) trên khu vực Biển Đông, cho phép đưa ra một số kết luận sau:

1. Trung bình hàng năm có khoảng  $10,9 \pm 2,9$  XTNĐ hoạt động trong khoảng  $35,6 \pm 10,8$  ngày, trong đó XTNĐ nhiều nhất (ít nhất) ở những năm La Niña (El Niño) với số lượng khoảng  $11,3 \pm 3,5$  cơn ( $10,1 \pm 2,7$  cơn). Ở những năm La Niña (El Niño) thời gian XTNĐ hoạt động khoảng  $36 \pm 11,7$  ngày ( $33,7 \pm 11,9$  ngày) trong khi ở những năm Trung tính là  $37 \pm 8,4$  ngày. Số lượng XTNĐ trong những năm La Niña nhiều hơn so với năm El Niño khoảng gần 2 cơn/năm và sự chênh lệch này thường tập trung vào 3 tháng cuối mùa bão (tháng 9, 10 và 11).

2. Mùa bão Biển Đông kéo dài từ tháng 6-11 nhưng chủ yếu hoạt động mạnh vào các tháng 7-10. Số lượng XTNĐ Biển Đông giai đoạn 1971-2019 có sự biến động mạnh qua các năm và có dấu hiệu về xu thế giảm thoả mãn mức ý nghĩa 10% ở tổng số lượng XTNĐ trong mùa bão (tháng 6 tới tháng 11) và 3 tháng cuối mùa bão (các tháng 9, 10, 11) tương ứng với mức giảm khoảng 0,5 cơn/thập kỷ ( $P_{\text{val}} = 0,06$ ) và 0,3 cơn/thập kỷ ( $P_{\text{val}} = 0,07$ ).

3. Vùng hoạt động và số lượng XTNĐ Biển Đông có sự biến động đáng kể qua từng thập kỷ ở giai đoạn 1971-2019. Những năm 1971-1980 XTNĐ hoạt động tập trung nhiều ở khu vực Đông Bắc Biển Đông nhưng đến thập kỷ 1981-1990 vùng XTNĐ hoạt động nhiều dịch vào trung tâm Biển Đông. Thập kỷ 1991-2000 vùng nhiều XTNĐ có xu hướng giảm và dịch lên phía đông bắc. Hai thập kỷ gần đây 2001-2019 vùng XTNĐ hoạt động mở rộng xuống phía Nam Biển Đông.

4. Có khả năng mô phỏng và dự báo hạn mùa (3 tháng) số lượng XTNĐ 3 tháng đầu mùa bão và 3 tháng cuối mùa bão trên khu vực Biển Đông bằng mô hình

thống kê với bộ nhân tố dự báo được lựa chọn từ bộ 30 chỉ số khí hậu hoặc các trường quy mô lớn từ sản phẩm của mô hình khí hậu CFSv2, trong đó:

- Các phương trình dự báo thống kê với nhân tố dự báo lựa chọn từ bộ chỉ số khí hậu (TH2) dự báo số lượng XTNĐ 3 tháng đầu mùa bão và 3 tháng cuối mùa bão cho sai số thấp hơn so với độ lệch chuẩn khí hậu đồng thời các chỉ số kỹ năng dự báo ở TH2 cũng thể hiện tốt hơn so với dự báo trung bình khí hậu. Trong số 3 trường hợp khảo sát, TH2 có kết quả dự báo tốt nhất.

- Các phương trình dự báo thống kê với nhân tố dự báo lựa chọn từ sản phẩm mô hình khí hậu CFSv2 (TH3) đưa ra kết quả dự báo số lượng XTNĐ 3 tháng đầu mùa bão có sai số và điểm kỹ năng dự báo tốt hơn các phương trình dự báo với nhân tố dự báo lựa chọn từ thành phần chính của bộ chỉ số khí hậu (TH1), trong khi điều này ngược lại khi dự báo cho các XTNĐ 3 tháng cuối mùa bão (TH1 tốt hơn TH3).

- Trong số 5 mô hình khảo sát, bên cạnh mô hình MLR cho kết quả dự báo ổn định, với kỹ năng dự báo tốt nhất thì sản phẩm tổ hợp E1234 cũng thể hiện được kỹ năng dự báo số lượng XTNĐ Biển Đông 3 tháng đầu mùa bão và 3 tháng cuối mùa bão theo TH1, TH2.

## KIẾN NGHỊ

Bão, ATNĐ là hiện tượng thời tiết đặc biệt nguy hiểm do vậy việc nghiên cứu, phát triển công cụ dự báo mang ý nghĩa thực tiễn cao. Để hoàn thiện hơn các kết quả luận án đã trình bày, góp phần đưa các kết quả nghiên cứu này vào công tác dự báo nghiệp vụ, luận án có một số kiến nghị sau:

1. Các khảo sát dự báo số lượng XTNĐ 3 tháng đầu mùa bão và 3 tháng cuối mùa bão theo TH2 đã cho ra kết quả tốt hơn so với dự báo trung bình khí hậu. Bên cạnh đó các sai số và chỉ số kỹ năng dự báo của TH2 tiệm cận và so sánh được với kết quả dự báo cho các vùng biển khác đã được công bố trong khu vực và trên thế giới. Do vậy, cần thiết phải có những đánh giá, nghiên cứu chi tiết hơn để đưa TH2 ứng dụng thực tế dự báo nghiệp vụ.
2. Mặc dù kết quả dự báo của TH1, TH3 kém hơn so với TH2 song đây là cách tiếp cận mới để dự báo XTNĐ cho khu vực Biển Đông, hơn nữa ở một số thời điểm nhất định 2 trường hợp này và sản phẩm dự báo tổ hợp cũng đã thể hiện được kỹ năng dự báo. Điều đó cho thấy cần tiếp tục nghiên cứu, khảo sát thêm TH1, TH3 trong việc tổ hợp các kết quả dự báo với TH2 để có được kết quả dự báo tốt hơn nữa.
3. Trên cơ sở khảo sát quan hệ tương quan giữa XTNĐ Biển Đông 3 tháng, 6 tháng với các chỉ số khí hậu và các trường quy mô lớn là sản phẩm của mô hình CFSv2 cho thấy việc khảo sát bổ sung thêm đối với các yếu tố dự báo như: 1) số lượng XTNĐ Biển Đông 6 tháng; 2) số ngày hoạt động của XTNĐ Biển Đông 3 tháng và 3) số ngày hoạt động của XTNĐ Biển Đông 6 tháng là hoàn toàn khả thi. Điều này sẽ đóng góp và làm rõ hơn về kết quả dự báo mùa hoạt động của XTNĐ Biển Đông trong thời hạn 3 đến 6 tháng.

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

1. Đinh Bá Duy, Ngô Đức Thành, Nguyễn Thị Tuyết, Phạm Thanh Hà, Phan Văn Tân (2016), “Đặc điểm hoạt động của Xoáy thuận Nhiệt đới trên khu vực Tây Bắc Thái Bình Dương, Biển Đông và vùng trực tiếp chịu ảnh hưởng trên lãnh thổ Việt Nam giai đoạn 1978-2015”, *Tạp chí Khoa học ĐHQGHN: Các Khoa học Trái đất và Môi trường*, Tập 32, Số 2 (2016) tr. 1-11.
2. Đinh Bá Duy, Ngô Đức Thành, Phan Văn Tân (2016), “Mối quan hệ giữa ENSO và số lượng, cấp độ Xoáy thuận Nhiệt đới trên khu vực Tây Bắc - Thái Bình Dương, Biển Đông giai đoạn 1951-2015”, *Tạp chí Khoa học ĐHQGHN: Các Khoa học Trái đất và Môi trường*, Tập 32, Số 3S (2016), tr. 43-55.
3. Đinh Bá Duy, Ngô Đức Thành, Trần Quang Đức, Phan Văn Tân (2019), “Dự báo hạn mùa số lượng xoáy thuận nhiệt đới trên Biển Đông bằng các mô hình thống kê”, *VNU Journal of Science: Earth and Environmental Sciences*, Vol. 35, No. 2 (2019) pp. 45-57.
4. Trần Quang Đức, Phạm Thanh Hà, Đinh Bá Duy, Phạm Quang Nam (2020), “Thay đổi hoạt động của bão Biển Đông”, *Tạp chí Khí tượng Thủy văn*, 715, tr. 27-36; doi:10.36335/VNJHM.2020(715).27-36

## TÀI LIỆU THAM KHẢO

### Tiếng Việt

1. Dương Liên Châu và cộng sự (2005), *Xây dựng cơ sở dữ liệu bão, áp thấp nhiệt đới trên Biển Đông và ảnh hưởng tới Việt Nam*, Báo cáo tổng kết đề tài cấp Trung tâm Dự báo khí tượng thủy văn quốc gia, Hà Nội.
2. Tạ Quốc Dũng, Lê Thế Hà, Phạm Duy Khang (2019), “Ứng dụng mạng Neuron nhân tạo (ANN) trong dự báo độ rỗng”, *Tạp chí dầu khí 7*, tr. 18 – 27.
3. Vũ Thị Hằng, Ngô Thị Thanh Hương và Phan Văn Tân (2010), “Đặc điểm hoạt động của bão ở vùng biển ần bờ Việt Nam giai đoạn 1945-2007”, *Tạp chí Khoa học ĐHQGHN, Khoa học Tự nhiên và Công nghệ 26(3S)*, tr. 344 - 353.
4. Nguyễn Văn Hiệp (2016), “Đặc điểm hoạt động của bão ở Tây Bắc Thái Bình Dương và Biển Đông qua số liệu Ibtras”, *Tuyển tập báo cáo tại Hội thảo khoa học năm 2016 của Viện Khoa học KTTV&BDKH*, tr. 9-14.
5. Phạm Văn Huân (2020), *Chuẩn đoán và dự báo thống kê các quá trình hải dương học*, NXB Đại học Quốc gia Hà Nội, Hà Nội
6. Hồ Việt Hùng, Lê Xuân Hiền, Giha Lee (2018), “Ứng dụng mạng thần kinh nhân tạo dự báo lưu lượng dòng chảy sông Hồng tại Sơn Tây dựa trên dữ liệu ở thượng lưu”, *Tuyển tập Công trình Hội nghị khoa học Cơ học Thủy khí toàn quốc lần thứ 2018*, tr. 1-10.
7. Đặng Hồng Nga, Nguyễn Minh Việt và Hoàng Đức Cường (2011), “Xu thế diễn biến của tần số xoáy thuận nhiệt đới ở Tây Bắc Thái Bình Dương và Biển Đông”, *Tạp chí Khí tượng Thủy văn*, tr. 31-37.
8. Nguyễn Đức Ngữ và cộng sự (2002), *Tác động của ENSO đến thời tiết, khí hậu, môi trường và kinh tế- xã hội ở Việt Nam*, Báo cáo kết quả đề tài nghiên cứu khoa học độc lập cấp Nhà nước, Hà Nội.
9. Phan Văn Tân (2005), *Phương pháp thống kê trong khí hậu*, NXB Đại học Quốc gia Hà Nội, Hà Nội.
10. Phan Văn Tân và cộng sự (2011), *Nghiên cứu xây dựng hệ thống mô hình dự báo hạn mùa một số hiện tượng khí hậu cực đoan phục vụ phòng tránh thiên tai ở Việt Nam*, Đề tài cấp Nhà nước, mã số ĐT.NCCN-ĐHUD.2011-G/09, Hà Nội.

11. Phan Văn Tân và cộng sự (2010), *Nghiên cứu tác động của biến đổi khí hậu toàn cầu đến các yếu tố và hiện tượng khí hậu cực đoan ở Việt Nam, khả năng dự báo và giải pháp chiến lược ứng phó*, Báo cáo Tổng kết Đề tài KC08.29/06-10. Bộ Khoa học và Công nghệ, Hà Nội.
12. Ngô Đức Thành, Phan Văn Tân (2012), “Kiểm nghiệm phi tham số xu thế biến đổi của một số yếu tố khí tượng cho giai đoạn 1961-2007”, *Tạp chí khoa học ĐHQG Hà Nội, Khoa học Tự nhiên và Công nghệ* 28 (3S), tr. 129-135.
13. Nguyễn Văn Thắng, Mai Văn Khiêm, Nguyễn Trọng Hiệu, Vũ Văn Thắng, Nguyễn Đăng Mậu, Lê Thị Tuyết (2016), “Ảnh hưởng của bão ở Việt Nam thời kỳ 1961-2014”, *Tạp chí Khoa học ĐHQGHN: Các Khoa học Trái đất và Môi trường* Tập 32(3S), tr. 210-216.
14. Trần Thục, Nguyễn Văn Thắng, Huỳnh Thị Lan Hương, Mai Văn Khiêm, Nguyễn Xuân Hiền, Doãn Hà Phong và nnk (2016), *Kịch bản biến đổi khí hậu và nước biển dâng cho Việt Nam*, NXB Tài nguyên Môi trường và Bản đồ Việt Nam, Hà Nội.
15. Vũ Văn Tuấn (2020), “Lựa chọn cấu trúc mạng Nơ ron nhân tạo (ANN) dự báo chỉ số nén của đất”, *Tạp chí KHCN Xây dựng* (3), tr. 67-75.
16. Nguyễn Văn Tuyên (2007), “Xu hướng hoạt động của xoáy thuận nhiệt đới trên Tây bắc Thái Bình Dương và Biển Đông theo các cách phân loại khác nhau”, *Tạp chí Khí tượng Thủy văn* (559), tr. 4-10.
17. Nguyễn Văn Tuyên (2008), “Khả năng dự báo hoạt động mùa bão Biển Đông Việt Nam Phân tích các yếu tố dự báo và nhân tố dự báo có thể (Phần I).” *Tạp chí Khí tượng Thủy văn* (568), tr.1-8.
18. Nguyễn Văn Tuyên (2008), “Khả năng dự báo hoạt động mùa bão Biển Đông Việt Nam Phân tích các yếu tố dự báo và nhân tố dự báo có thể (Phần II)” *Tạp chí Khí tượng Thủy văn* (571), tr.1-11.
19. Đinh Văn Ưu, Phạm Hoàng Lâm (2005), “Biến động mùa và nhiều năm của trường nhiệt độ nước mặt biển và sự hoạt động của bão tại khu vực Biển Đông”, *Tạp chí Khoa học ĐHQGHN, Khoa học Tự nhiên và Công nghệ* T.XXI (3PT), tr. 127-136.

20. Đinh Văn Ưu (2009), “Đánh giá quy luật biến động dài hạn và xu thế biến đổi số lượng bão và áp thấp nhiệt đới trên khu vực Tây Thái Bình Dương , Biển Đông và ven biển Việt Nam”, *Tạp chí Khoa học ĐHQGHN, Khoa học Tự nhiên và Công nghệ* T.25(3S), tr. 542–550.
21. Đinh Văn Ưu (2011), “Đặc điểm biến động bão và áp thấp nhiệt đới ảnh hưởng trực tiếp đến đất liền Việt Nam”, *Tạp chí Khoa học ĐHQGHN, Khoa học Tự nhiên và Công nghệ* T.27(1S), tr. 266-272. 27.

### **Tiếng Anh**

22. Björnsson H, Venegas SA - CCGCR Report (1997), *A manual for EOF and SVD analyses of climatic data*, McGill university, Canada.
23. Bridhikitti A. (2013), “Connections of ENSO/IOD and aerosols with Thai rainfall anomalies and associated implications for local rainfall forecasts”, *International journal of climatology* 33(13), pp. 2836-2845.
24. Camargo S. J., Sobel A. H. (2004), “Western North Pacific tropical cyclone intensity and ENSO”, *Journal of Climate* 18(15), pp. 2996–3006.
25. Camargo S. J. (2013), “Tropical cyclones in high-resolution climate models”, *US CLIVAR Variations* 11(3), pp. 4–11.
26. Chan J. C. (1985), “Tropical cyclone activity in the Northwest Pacific in relation to El Nino/Southern Oscillation phenomenon”, *Monthly Weather Review* 113(April.) pp. 599–606.
27. Chan J. C. (1995), “Tropical cyclone activity in the western North Pacific in relation to the stratospheric quasi-biennial oscillation”, *Monthly Weather Review* 123(8), pp. 2567-2571.
28. Chan J. C. (1995), “Prediction of annual tropical cyclone activity over the western North Pacific and the South China Sea”, *International Journal of Climatology* 15(9), pp. 1011–1019.
29. Chan J. C., Shi, J. E., Lam C. M. (1998), “Seasonal Forecasting of Tropical Cyclone Activity over the Western North Pacific and the South China Sea”, *Weather and forecasting* 13(4), pp. 997–1004.
30. Chan J. C. (2000), “Tropical Cyclone Activity over the Western North Pacific

- Associated with El Nino and La Nina Events”, *Journal of Climate* 13(16), pp. 2960–2972.
31. Chan J. C., Shi J. E. (2000), “Frequency of typhoon landfall over Guangdong Province of China during the period 1470-1931”, *International Journal of Climatology* 20(2), pp. 183–190.
  32. Chan J. C., Shi J. E., Liu K. S. (2001), “Improvements in the Seasonal Forecasting of Tropical Cyclone Activity over the Western North Pacific”, *Weather and Forecasting* 16(4), pp. 491–498.
  33. Chen J. H., Lin S. J. (2011), “The remarkable predictability of inter-annual variability of Atlantic hurricanes during the past decade”, *Geophysical Research Letters* 38(11), pp. 1–6.
  34. Chen J. H., Lin S. J. (2013), “Seasonal predictions of tropical cyclones using a 25-km-resolution general circulation model”, *Journal of Climate* 26(2), pp. 380–398.
  35. Choi K.S., Byun H.R. (2010), “Possible relationship between western North Pacific tropical cyclone activity and Arctic Oscillation”, *Theor Appl Climatol* 100, pp. 261–274.
  36. Chu P. S., Zhao X., Ho C. H., Kim H. S., Lu M. M., Kim, J. H. (2010), “Bayesian forecasting of seasonal typhoon activity: A track-pattern-oriented categorization approach”, *Journal of climate* 23(24), pp. 6654–6668.
  37. Elsner J. B., Liu K. B., Kocher B. (2000), “Spatial variations in major U.S. hurricane activity: Statistics and a physical mechanism”, *Journal of Climate* 13(13), pp. 2293–2305.
  38. Gao S., Chen Z., Zhang W. (2018), “Impacts of tropical North Atlantic SST on Western North Pacific landfalling tropical cyclones”, *Journal of Climate* 31(2), pp. 853–862.
  39. Goerss J. S., Sampson C. R., Gross J. M. (2004), “A History of Western North Pacific Tropical Cyclone Track Forecast Skill”, *Weather and Forecasting* 19(3), pp. 633–638.
  40. Gray W. M. (1979), *Hurricanes: Their formation, structure and likely role in the tropical circulation*, Royal Meteorological Society, London.

41. Gray W. M., Landsea C. W., Mielke Jr, P. W., Berry K. J. (1994), “Predicting Atlantic Basin Seasonal Tropical Cyclone Activity by 1 June”, *Weather and Forecasting* 9(1), pp. 103-115.
42. Halperin D. J., Fuelberg H. E., Hart R. E., Cossuth J. H., Sura P., Pasch R. J. (2013), “An Evaluation of Tropical Cyclone Genesis Forecasts from Global Numerical Models”, *Weather and forecasting* 28(6), pp. 1423–1445.
43. Hannachi A. (2004), *A Primer for EOF Analysis of Climate Data*, University of Reading, U.K.
44. Jien J. Y., Gough W. A., Butler K. (2015), “The Influence of El Niño-Southern Oscillation on Tropical Cyclone Activity in the Eastern North Pacific Basin”, *Journal of climate* 28(6), pp. 2459–2474.
45. Jin E. K., et al. (2008), “Current status of ENSO prediction skill in coupled ocean-atmosphere models”, *Climate Dynamics* 31(6), pp. 647–664.
46. Kendall M.G. (1975), *Rank Correlation Methods*, Charles Griffin, London.
47. Khashei M., Bijari M. (2010), “An artificial neural network (p,d,q) model for timeseries forecasting”, *Expert Systems with Applications* 37 (2010), pp. 479–489.
48. Kim H. M., Lee M. I., Webster P. J., Kim D., Yoo J. H. (2013), “A physical basis for the probabilistic prediction of the accumulated tropical cyclone kinetic energy in the Western North Pacific”, *Journal of climate* 26(20), pp. 7981–7991.
49. Kim H. M., Webster P. J. (2010), “Extended-range seasonal hurricane forecasts for the North Atlantic with a hybrid dynamical-statistical model”, *Geophysical research letters* 37(21), pp. 1–5.
50. Kim H. M., Webster P. J., Curry J. A. (2012), “Seasonal prediction skill of ECMWF System 4 and NCEP CFSv2 retrospective forecast for the Northern Hemisphere Winter”, *Climate Dynamics* 39(12), pp. 2957–2973.
51. Kim O. Y., Kim H. M., Lee M. I., Min Y. M. (2017), “Dynamical – statistical seasonal prediction for western North Pacific typhoons based on APCC multi-models”, *Climate Dynamics* 48(1), pp. 71–88.
52. Klotzbach P. J. (2007), “Recent developments in statistical prediction of seasonal Atlantic basin tropical cyclone activity”, *Tellus A: Dynamic*

- Meteorology and Oceanography* 59A(4), pp. 511–518.
53. Klotzbach P. J., Blake E., Camp J., Caron L.P., Chan J. C., Kang N.Y, Kuleshov Y., Lee S.M, Murakami H., Saunders M., Takaya Y., Vitart F., Zha R. (2019), “SeaSonal tropical cyclone forecasting”, *SeaSonal tropical cyclone Forecasting* 8(3), pp. 134–149.
  54. Kwon H. J., Lee W. J., Won S. H., Cha E. J. (2007), “Statistical ensemble prediction of the tropical cyclone activity over the western North Pacific”, *Geophysical Research Letters* 34(24), pp. 1-5.
  55. Landsea C. W., Bell G. D., Gray W. M., Goldenberg S. B. (1998), “The Extremely Active 1995 Atlantic Hurricane Season: Environmental Conditions and Verification of Seasonal Forecasts”, *Monthly Weather Review* 126(5), pp. 1174–1193.
  56. Landsea C. W., Gray W. M., Mielke Jr P. W., Berry K. J. (1994), “Seasonal forecasting of Atlantic hurricane activity”, *Weather* 49(8), pp. 273–284.
  57. LaRow T. E., Stefanova L., Shin D. W., Cocke S. (2010), “Seasonal atlantic tropical cyclone hindcasting/forecasting using two sea surface temperature datasets”, *Geophysical Research Letters* 37(2), pp. 1–5.
  58. Li X., Yang S., Wang H., Jia X., Kumar A. (2013), “A dynamical-statistical forecast model for the annual frequency of western Pacific tropical cyclones based on the NCEP Climate Forecast System version 2”, *Journal of Geophysical Research: Atmospheres* 118, pp. 12.061–12.074.
  59. Li Y. X., Yu, J. Y. (2020), “Why rare tropical cyclone formation after maturity of super El Niño events in the western North Pacific?”, *Terrestrial, Atmospheric & Oceanic Sciences* 31(1), pp. 21-32.
  60. Lorenz E. N. (1963), “Deterministic nonperiodic flow”, *Journal of the atmospheric sciences* 20(2), pp. 130-141.
  61. Lu M. M., Chu P. S., Lin Y. C. (2010), “Seasonal Prediction of Tropical Cyclone Activity Near Taiwan Using the Bayesian Multivariate Regression Method”, *Weather and forecasting* 25(6), pp. 1780–1795.
  62. MacLachlan C., et al. (2015), “Global Seasonal forecast system version 5 (GloSea5): A high-resolution seasonal forecast system”, *Quarterly Journal of*

- the Royal Meteorological Society* 141(689), pp. 1072–1084.
63. Manganello J. V., et al. (2012), “Tropical cyclone climatology in a 10-km global atmospheric GCM: Toward weather-resolving climate modeling”, *Journal of Climate* 25(11), pp. 3867–3893.
  64. Manganello J. V., et al. (2016), “Seasonal forecasts of tropical cyclone activity in a high-atmospheric- resolution coupled prediction system”, *Journal of Climate* 29(3), pp. 1179–1200.
  65. Molinari J., Vollaro D. (2013), “What percentage of western North Pacific tropical cyclones form within the monsoon trough?”, *Monthly weather review* 141(2), pp. 499-505.
  66. Nicholls N., Landsea C., Gill J. (1998), “Recent trends in Australian region tropical cyclone activity”, *Meteorology and Atmospheric Physics* 65(3–4), pp. 197–205.
  67. Palmer T. N., Anderson D. L. T. (1994), “The prospects for seasonal forecasting-A review paper”, *Quarterly Journal of the Royal Meteorological Society* 120(518), pp. 755-793.
  68. Phan-Van T. et al. (2014), “Seasonal prediction of surface air temperature across Vietnam using the regional climate model version 4.2 (RegCM4.2)”, *Advances in Meteorology* 2014( ID 245104), pp. 1-13.
  69. Phan-Van T., Trinh-Tuan L., Bui-Hoang H., Kieu C. (2015), “Seasonal forecasting of tropical cyclone activity in the coastal region of Vietnam using RegCM4.2”, *Climate Research* 62(2), pp. 115–129.
  70. Pradhan P. K., Preethi B., Ashok K., Krishnan R., Sahai A. K. (2011), “Modoki, Indian Ocean Dipole, and western North Pacific typhoons: Possible implications for extreme events”, *Journal of Geophysical Research: Atmospheres* 116(18), pp. 1–12.
  71. Roberts M. J., et al. (2015), “Tropical cyclones in the UPSCALE ensemble of high-resolution global climate models”, *Journal of Climate* 28(2), pp. 574–596.
  72. Saha S., et al. (2014), “The NCEP Climate Forecast System Version 2”, *Journal of climate* 27(6), pp. 2185–2208.
  73. Saji N. H., Goswami B. N., Vinayachandran P. N., Yamagata T. (1999), “A dipole mode in the tropical Indian ocean”, *Nature* 401(6751), pp. 360–363.

74. Sen P.K. (1968), “Estimates of the Regression Coefficient Based on Kendall’s Tau”, *Journal of the American Statistical Association* 63(324), pp. 1379-1389.
75. Shaowu W., Jinhong Z. (2001), “A review on seasonal climate prediction”, *Advances in Atmospheric Sciences* 18(2), pp. 197–208.
76. Shukla J., et al. (2000), “Dynamical Seasonal Prediction”, *Bulletin of the American Meteorological Society* 81(11), pp. 2593–2606.
77. Song J. J., Wang Y., Wu L. (2010), “Trend discrepancies among three best track data sets of western North Pacific tropical cyclones,” *Journal of Geophysical Research Atmospheres* 115(12), pp. 1-9.
78. Steinwart I., Christmann A. (2011), “Estimating conditional quantiles with the help of the pinball loss”, *Bernoulli* 17(1), pp. 211–225.
79. Stockdale T. N. (2000), “An overview of techniques for seasonal forecasting”, *Stochastic Environmental Research and Risk Assessment* 14(4), pp. 305–318.
80. Ta-Huu C., Sato T. (2019), “Effect of ENSO phase on the contribution of environmental variables to tropical cyclone genesis in the western North Pacific”, *International Journal of Climatology* 39(4), pp. 2461–2473.
81. Tian D., Wood E. F., Yuan X. (2017), “CFSv2-based sub-seasonal precipitation and temperature forecast skill over the contiguous United States”, *Hydrology and Earth System Sciences* 21(3), pp. 1477–1490.
82. Trenberth K. E. (1997), “The Definition of El Nino”, *Bulletin of the American Meteorological Society* 78(12), pp. 2771–2778.
83. Vecchi G. A., et al. (2011), “Statistical-dynamical predictions of seasonal North Atlantic hurricane activity”, *Monthly Weather Review* 139(4), pp. 1070–1082.
84. Vitart F., Stockdale T. N. (2001), “Seasonal Forecasting of Tropical Storms Using Coupled GCM Integrations”, *Monthly Weather Review* 129(10), pp. 2521–2537.
85. Wang B., Chan, J. C. (2002), “How strong ENSO events affect tropical storm activity over the western North Pacific”, *Journal of Climate* 15(13), pp. 1643-1658.
86. Wang H., et al. (2009), “A statistical forecast model for Atlantic seasonal hurricane activity based on the NCEP dynamical seasonal forecast” *Journal of*

- climate* 22(17), pp. 4481–4500.
87. Wang H., Sun J., Fan K. (2007), “Relationships between the North Pacific Oscillation and the typhoon/hurricane frequencies”, *SCI CHINA SER D* 50, pp. 1409–1416.
  88. Wang X., Zhou W., Li C., Wang D. (2012), “Effects of the East Asian summer monsoon on tropical cyclone genesis over the South China Sea on an interdecadal time scale”, *Advances in Atmospheric Sciences* 29(2), pp. 249–262.
  89. Wang L., Huang R., Wu R. (2016), “Tropical cyclogenesis associated with four types of winter monsoon circulation over the South China Sea”, *Atmospheric Science Letters* 17(5), pp. 326–333.
  90. Weisheimer A., Palmer T. N. (2014), “On the reliability of seasonal climate forecasts”, *Journal of the Royal Society Interface*, 11(96), pp. 1-10.
  91. Weisheimer A., et al. (2009), “ENSEMBLES: A new multi-model ensemble for seasonal to annual predictions - Skill and progress beyond DEMETER in forecasting tropical Pacific SSTs”, *Geophysical research letters* 36(21), pp. 1–6.
  92. Wilks D. (2005), *Statistical Methods in the Atmospheric Sciences*, Vol. 100, 2nd Edition, *Academic Press*, USA.
  93. Wu M. C., Chang W. L., Leung, W. M. (2004), “Impacts of El Niño –Southern Oscillation Events on Tropical Cyclone Landfalling Activity in the Western North Pacific”, *Journal of Climate* 17(6), pp. 1419–1428.
  94. Wu L., Wen Z., Huang R., Wu R. (2012), “Possible linkage between the monsoon trough variability and the tropical cyclone activity over the western North Pacific”, *Monthly Weather Review* 140(1), pp. 140–150.
  95. Xu S. B., Huang F. (2015), “Impacts of the two types of El Niño on Pacific tropical cyclone activity”, *J. Ocean Univ. China* 14, pp. 191–198.
  96. Zhan R., Wang Y. (2016), “CFSv2-Based Statistical Prediction for Seasonal Accumulated Cyclone Energy (ACE) over the Western North Pacific”, *Journal of Climate* 29(2), pp. 525–541.
  97. Zhang G., Patuwo B. E., Hu M. Y. (1998), “Forecasting with artificial neural networks: The state of the art”, *International Journal of Forecasting* 14, pp. 35–62

98. Zhang W., Vecchi G. A., Villarini G., et al (2017), “Modulation of western North Pacific tropical cyclone activity by the Atlantic Meridional Mode”, *Climate Dynamics* 48(1-2), pp. 631-647.
99. Zhao H., Wang C. (2019), “On the relationship between ENSO and tropical cyclones in the western North Pacific during the boreal summer”, *Climate Dynamics* 52, pp. 275–288.
100. Zhang W., Zhang Y., Zheng D., Wang F., Xu L. (2015), “Relationship between lightning activity and tropical cyclone intensity over the northwest Pacific”, *Journal of Geophysical Research: Atmospheres* 120(9), pp. 4072–4089.
101. Zhao M., Held I. M., Vecchi G. A. (2010), “Retrospective Forecasts of the Hurricane Season Using a Global Atmospheric Model Assuming Persistence of SST Anomalies”, *Monthly Weather Review* 138(10), pp. 3858–3868.
102. Zhou B., Cui X. (2014), “Interdecadal change of the linkage between the North Atlantic Oscillation and the tropical cyclone frequency over the western North Pacific”, *Science China Earth Sciences* 57(9), pp. 2148–2155.

**Trang web**

103. <https://www.wmo.int/pages/prog/www/tcp/Advisories-RSMCs.html>
104. <https://www.ncdc.noaa.gov/ibtracs/>
105. <https://www.jma.go.jp/jma/jma-eng/.../besttrack.html>
106. <http://www.usno.navy.mil>
107. <http://tcdata.typhoon.gov.cn/en/index.html>
108. <https://psl.noaa.gov/data/climateindices/list/>
109. <https://www.cpc.ncep.noaa.gov/data/indices/>
110. <http://imh.ac.vn/nghiiep-vu>
111. <https://nchmf.gov.vn/Kttvsite/vi-VN/1/du-bao-mua-9-15.html>

## PHỤ LỤC

Bảng 1. Số lượng XTNĐ 3 tháng đầu mùa bão (6-7-8), TH1

Giai đoạn	Năm	Số liệu quan trắc và kết quả của các mô hình					
		Obs	MLR	LAD	LMV	ANN	E1234
Dự báo phụ thuộc	1980	4	4.4	4.9	4.9	4	4.5
	1981	6	5	4	4.7	4.8	4.6
	1982	5	4.4	4.6	5.1	4.8	4.7
	1983	5	4.5	3.8	4.7	4.7	4.4
	1984	5	5.1	5	5	4.7	4.9
	1985	2	5	6.9	4.2	4.8	5.2
	1986	3	4.3	3.9	4.6	4.8	4.4
	1987	3	4.5	5	4.3	4.8	4.6
	1988	4	4.3	4	3.2	4.1	3.9
	1989	4	4	4	5.6	4.7	4.6
	1990	4	4.1	4	5.7	3.7	4.4
	1991	5	3.9	3.2	4.3	4.5	4
	1992	6	4.4	4.7	4.6	4.8	4.6
	1993	4	4.9	4.5	5	4.5	4.7
	1994	6	4.4	4.9	3.8	4.8	4.5
	1995	6	4.3	4.8	3.8	4.8	4.4
	1996	3	3.9	3.2	5.2	3.2	3.9
	1997	4	4.3	5.2	5	4.6	4.8
	1998	2	3.2	2.5	4.2	3.1	3.2
	1999	4	3.7	4	5.4	3.1	4
	2000	3	3.3	3	5.2	3.6	3.8
	2001	7	4.3	5	4.8	4.8	4.7
	2002	4	3.8	4	4.8	3.1	3.9
	2003	4	4	4	3.5	4.8	4.1
	2004	4	3.4	3.8	4.4	3.1	3.7
	2005	2	3.6	2.3	4.2	3.1	3.3
	2006	4	4	3.9	5.3	3.1	4.1
	2007	2	3.1	2.3	3.9	3.2	3.1
	2008	3	4	4.9	5.2	3.2	4.3
	2009	5	3.4	3.3	4.5	4.3	3.9
2010	4	3.6	3.7	4.3	4.1	3.9	
Dự báo độc lập	2011	3	3.1	2.7	5.7	3.1	3.6
	2012	5	3.6	3.6	5.2	3.3	3.9
	2013	8	3.8	4	4.7	3.1	3.9
	2014	3	3.4	2.7	4.7	3.5	3.6
	2015	2	3.8	4.3	4.4	3.9	4.1
	2016	3	3.4	2.8	3	3.3	3.1
	2017	8	3.3	2.8	4.7	3.1	3.5

Bảng 2. Số lượng XTNĐ 3 tháng cuối mùa bão (9-10-11), TH1

Giai đoạn	Năm	Số liệu quan trắc và kết quả của các mô hình					
		Obs	MLR	LAD	LMV	ANN	E1234
Dự báo phụ thuộc	1980	4	4.4	4	4.9	3.4	4.2
	1981	4	5.4	4.4	6.1	6.2	5.5
	1982	3	3.4	4.2	3.8	3.3	3.7
	1983	8	6.7	5.1	5.9	6.2	6
	1984	6	5.5	6	4.7	6	5.5
	1985	9	7.4	8.2	6.9	6.1	7.1
	1986	6	6.2	6.8	5.9	6.2	6.3
	1987	4	3.9	4	3.7	3.3	3.7
	1988	7	6.9	7	6.1	6.2	6.6
	1989	7	5.6	6.5	5.2	6.1	5.9
	1990	5	4.1	4.5	4.4	3.5	4.1
	1991	7	4.7	5	4.9	5.7	5.1
	1992	4	3.9	3.9	3.9	3.4	3.8
	1993	5	5.9	5	6.5	6.2	5.9
	1994	3	5.6	6.9	5.1	3.9	5.4
	1995	8	6.4	6.4	6.2	6.2	6.3
	1996	4	3.3	3.9	3.2	3.5	3.5
	1997	3	3.7	3	5	3.7	3.8
	1998	5	5.5	5	3.3	4.7	4.6
	1999	5	5.7	5.7	5.1	6.2	5.7
	2000	3	5.4	5.4	5.1	6.2	5.5
	2001	3	3.9	4.6	3.4	3.4	3.8
	2002	2	3.5	3.7	4.1	3.5	3.7
2003	2	4.8	5.5	4.1	3.8	4.5	
2004	3	4.2	4.2	4.4	3.4	4	
2005	5	3.5	3.1	3.8	5.1	3.9	
2006	5	2.8	3.3	2.9	3.3	3.1	
2007	5	4.2	4.6	3.8	4.6	4.3	
2008	5	4.7	4.8	4.4	5.2	4.8	
2009	6	4.8	5.4	4.5	5.8	5.1	
2010	3	2.8	3	1.9	3.3	2.7	
Dự báo độc lập	2011	4	3.4	3.5	3.7	5	3.9
	2012	3	3.4	3.7	3.4	3.5	3.5
	2013	6	3.6	3.8	3.5	3.4	3.6
	2014	3	4.1	4.5	4.2	5.4	4.5
	2015	2	2.6	3.5	2.6	3.3	3
	2016	6	5.1	4.9	3.9	4.3	4.5
	2017	7	4.5	4.1	4.4	6.1	4.8

Bảng 3. Số lượng XTNĐ 3 tháng đầu mùa bão (6-7-8), TH2

Giai đoạn	Năm	Số liệu quan trắc và kết quả của các mô hình					
		Obs	MLR	LAD	LMV	ANN	E1234
Dự báo phụ thuộc	1980	4	3.6	3.7	3.3	3.1	3.4
	1981	6	5.4	6	5	5.5	5.5
	1982	5	3.7	3.6	3.8	3.5	3.7
	1983	5	3.5	3.3	3.5	4	3.6
	1984	5	3.8	3.7	3.5	4	3.7
	1985	2	2.2	1.5	2.6	2.7	2.3
	1986	3	4	4	3.6	4.1	3.9
	1987	3	3.3	3	3.5	3.1	3.2
	1988	4	4.3	4.2	3.9	4.2	4.1
	1989	4	4.1	3.9	4.1	3.8	4
	1990	4	3.3	3.3	3.2	3.1	3.2
	1991	5	4.6	4.9	4.3	5.1	4.7
	1992	6	5.4	5.8	5.2	5.7	5.5
	1993	4	6	6.5	5.5	5.8	5.9
	1994	6	5.8	6	6.2	5.6	5.9
	1995	6	5.4	6	4.5	5.7	5.4
	1996	3	3.9	3.8	4.1	3.9	3.9
	1997	4	4.2	4.3	3.8	4.2	4.1
	1998	2	3.4	3.2	3.5	3.2	3.3
	1999	4	4.1	4.2	4.1	3.6	4
	2000	3	3.9	3.7	4.3	3.6	3.9
	2001	7	5.7	6.1	5.5	5.7	5.8
	2002	4	4.6	4.8	4.5	4.8	4.7
	2003	4	3.5	3.3	3.6	3.6	3.5
	2004	4	2.4	1.7	2.8	2.7	2.4
2005	2	3.2	3.1	3	2.9	3.1	
2006	4	4.1	4.1	4.4	4.4	4.3	
2007	2	3.3	3.2	3.2	3	3.2	
2008	3	3.6	3.5	3.9	3.2	3.5	
2009	5	4.5	4.7	4.7	4.3	4.6	
2010	4	4.3	4	5	4.1	4.3	
Dự báo độc lập	2011	3	3.7	3.7	4	3.4	3.7
	2012	5	4.9	4.9	5.1	5	5
	2013	8	5.3	5.5	5.4	5.5	5.4
	2014	3	3.8	3.8	3.7	3.8	3.8
	2015	2	3.5	3.5	3.6	3.2	3.4
	2016	3	7.1	7.8	7.2	5.8	7
	2017	8	5.5	5.9	5.5	5.4	5.6

Bảng 4. Số lượng XTND 3 tháng cuối mùa bão (9-10-11), TH2

Giai đoạn	Năm	Số liệu quan trắc và kết quả của các mô hình					
		Obs	MLR	LAD	LMV	ANN	E1234
<b>Dự báo phụ thuộc</b>	1980	4	3	2.8	2.7	3.5	3
	1981	4	3.3	2.9	3.4	3.6	3.3
	1982	3	2.9	3	1.9	3.5	2.8
	1983	8	6.8	6.4	7	7.5	6.9
	1984	6	5.3	5.5	4	4.6	4.9
	1985	9	6.8	6.4	6.8	7.4	6.9
	1986	6	7	6.6	7	7.6	7.1
	1987	4	4.1	4.2	3	3.7	3.7
	1988	7	6.7	6.5	6.4	7.3	6.7
	1989	7	5.4	5.2	4.9	4.9	5.1
	1990	5	5.3	5	5.2	4.8	5.1
	1991	7	5.5	5.1	5.6	5.2	5.4
	1992	4	2.9	2.9	1.8	3.5	2.8
	1993	5	5.7	5.1	5.9	5.5	5.5
	1994	3	5	4.9	4	4.2	4.5
	1995	8	6.7	6.4	6.7	7.3	6.8
	1996	4	3.6	3.9	2.5	3.6	3.4
	1997	3	3.5	3.1	3.6	3.6	3.5
	1998	5	4.7	5	3.7	4	4.3
	1999	5	5.4	5	5.7	5.1	5.3
	2000	3	5.4	5.3	5.1	5	5.2
	2001	3	3.4	3.6	2.1	3.6	3.2
	2002	2	4	3.7	4.2	3.7	3.9
	2003	2	5.2	5.3	4.2	4.6	4.8
	2004	3	4.6	4.4	4.6	4	4.4
	2005	5	4.3	4.5	3.5	3.8	4
	2006	5	4.8	4.6	4.5	4.1	4.5
	2007	5	3.5	3.7	2.8	3.6	3.4
2008	5	5.2	5	4.9	4.6	4.9	
2009	6	6.4	6	6.3	6.8	6.4	
2010	3	2.8	3.5	1	3.5	2.7	
<b>Dự báo độc lập</b>	2011	4	5.2	5.2	5.1	4.8	5.1
	2012	3	1.8	1.9	0.8	3.5	2
	2013	6	4.9	4.7	4.8	4.3	4.7
	2014	3	5.7	5.4	5.6	5.6	5.6
	2015	2	3.8	4.3	2.2	3.6	3.5
	2016	6	5.6	5.6	5.1	5.4	5.4
	2017	7	5.2	4.8	5.4	4.8	5.1

Bảng 5. Số lượng XTNĐ 3 tháng đầu mùa bão (6-7-8), TH3

Giai đoạn	Năm	Số liệu quan trắc và kết quả của các mô hình					
		Obs	MLR	LAD	LMV	ANN	E1234
Dự báo phụ thuộc	1983	5	4.3	5	3.8	4.4	4.4
	1984	5	3.6	2.9	4.4	2.9	3.5
	1985	2	2	2.6	1	2.9	2.1
	1986	3	3.8	3.3	4.6	3	3.7
	1987	3	4.2	4.9	3.6	4.6	4.3
	1988	4	3.7	4	3.1	3.9	3.7
	1989	4	4.4	4.2	4.7	4.7	4.5
	1990	4	3.6	3.7	3.6	3.6	3.6
	1991	5	4.8	5	4.8	5.2	4.9
	1992	6	4.5	4.9	4.3	4.5	4.5
	1993	4	3.1	4	2.3	3.8	3.3
	1994	6	4.8	4.9	4.7	5.2	4.9
	1995	6	4	3.8	4.3	4.3	4.1
	1996	3	3.4	3	4.1	2.9	3.4
	1997	4	5	5.5	4.4	5.2	5
	1998	2	3	3.6	2.5	2.9	3
	1999	4	3.6	4	2.9	3	3.4
	2000	3	3.9	3.8	4	3	3.6
	2001	7	5.5	5.3	5.9	5.2	5.5
	2002	4	3	3.6	2.3	3.3	3
2003	4	4.1	3.9	4.2	4	4.1	
2004	4	3.6	3.6	3.7	4.2	3.8	
2005	2	3.7	3.8	3.7	3.7	3.7	
2006	4	5.3	5.2	5.7	5.2	5.3	
2007	2	3	3.2	2.8	3	3	
2008	3	4.6	4.5	4.7	3.5	4.3	
2009	5	5.2	5.2	5.3	5.2	5.2	
Dự báo phụ thuộc	2012	5	4.3	5	3.8	4.4	4.4
	2013	8	3.6	2.9	4.4	2.9	3.5
	2014	3	2	2.6	1	2.9	2.1
	2015	2	3.8	3.3	4.6	3	3.7
	2016	3	4.2	4.9	3.6	4.6	4.3

Bảng 6. Số lượng XTND 3 tháng cuối mùa bão (9-10-11), TH3

Giai đoạn	Năm	Số liệu quan trắc và kết quả của các mô hình					
		Obs	MLR	LAD	LMV	ANN	E1234
<b>Dự báo phụ thuộc</b>	1983	8	5.6	5.3	5.5	5.7	5.5
	1984	6	5.7	5.4	5.6	5.8	5.6
	1985	9	5.8	5.6	5.8	6.2	5.8
	1986	6	5.6	5.3	5.6	5.7	5.6
	1987	4	5	4.7	4.9	4.3	4.7
	1988	7	6.5	6.4	6.6	6.8	6.6
	1989	7	5.8	5.5	5.7	6.1	5.8
	1990	5	5.3	5	5.2	4.9	5.1
	1991	7	4	3.6	3.8	3.8	3.8
	1992	4	5.5	5.2	5.4	5.3	5.3
	1993	5	5.4	5	5.3	5	5.2
	1994	3	4.4	4	4.2	3.8	4.1
	1995	8	5.7	5.4	5.7	5.9	5.7
	1996	4	5.7	5.4	5.7	5.9	5.7
	1997	3	3.1	2.6	2.8	3.8	3.1
	1998	5	6.3	6.1	6.4	6.7	6.4
	1999	5	5.8	5.6	5.8	6.2	5.8
	2000	3	4.6	4.2	4.4	3.9	4.3
	2001	3	3.6	3.1	3.3	3.8	3.4
	2002	2	3.4	2.8	3.1	3.8	3.3
2003	2	5.3	5	5.2	4.8	5.1	
2004	3	3.5	3	3.2	3.8	3.4	
2005	5	5.4	5.1	5.3	5.1	5.2	
2006	5	4.7	4.4	4.6	4	4.4	
2007	5	5.4	5.1	5.3	5	5.2	
2008	5	4.4	4	4.2	3.8	4.1	
2009	6	3.4	2.9	3.1	3.8	3.3	
<b>Dự báo phụ thuộc</b>	2012	3	5.6	5.3	5.5	5.7	5.5
	2013	6	5.7	5.4	5.6	5.8	5.6
	2014	3	5.8	5.6	5.8	6.2	5.8
	2015	2	5.6	5.3	5.6	5.7	5.6
	2016	6	5	4.7	4.9	4.3	4.7