

# ĐÁNH GIÁ KHẢ NĂNG DỰ BÁO NHIỆT ĐỘ THẤP NHẤT VÀ CAO NHẤT TRUNG BÌNH HẠN ĐẾN 01 NĂM BẰNG PHƯƠNG PHÁP THỐNG KÊ SAU MÔ HÌNH

Trần Ngọc Vân<sup>(1)</sup>, Hoàng Phúc Lâm<sup>(1)</sup>, Nguyễn Đăng Mậu<sup>(2)</sup>, Nguyễn Đức Hòa<sup>(1)</sup>, Nguyễn Văn Hưởng<sup>(1)</sup>, Nguyễn Văn Huấn<sup>(3)</sup>, Hoàng Thị Mai<sup>(1)</sup>, Nguyễn Thanh Hoa<sup>(1)</sup>, Trịnh Thùy Nguyên<sup>(1)</sup>, Lê Lan Phương<sup>(1)</sup>, Nguyễn Thanh Thủy<sup>(4)</sup>, Lê Văn Phong<sup>(2)</sup>

<sup>(1)</sup>Trung tâm Dự báo khí tượng thủy văn quốc gia

<sup>(2)</sup>Viện khoa học khí tượng thủy văn và biến đổi khí hậu

<sup>(3)</sup>Đài khí tượng thủy văn khu vực Tây Nguyên

<sup>(4)</sup>Đài khí tượng thủy văn hu vực Đồng bằng và Trung du Bắc Bộ

Ngày nhận bài: 23/9/2024; ngày chuyển phản biện: 24/9/2024; ngày chấp nhận đăng: 24/10/2024

**Tóm tắt:** Trong khuôn khổ bài báo nghiên cứu các tác giả đã sử dụng ba phương pháp thống kê sau mô hình bao gồm: Phương pháp tương quan Canon, phương pháp hồi quy thành phần chính (HQTPC), phương pháp hồi quy tuyến tính đa biến (HQTTĐB) để dự báo yếu tố nhiệt độ thấp nhất trung bình và nhiệt độ cao nhất trung bình hạn đến 01 năm. Số liệu đầu vào bao gồm số liệu dự báo của 6 mô hình toàn cầu và số liệu quan trắc từ năm 1983-2022. Kết quả dự báo cho thấy, với yếu tố dự báo nhiệt độ thấp nhất trung bình phương pháp hồi quy tuyến tính đa biến có ưu thế so với hai phương pháp còn lại với kết quả độ chính xác 3 pha đạt 0,5-0,6, độ chính xác 2 pha đạt 0,7-0,8, trong khi đó với yếu tố nhiệt độ cao nhất trung bình với thời điểm dự báo tháng 2 và tháng 5 phương pháp hồi quy thành phần chính chiếm ưu thế, vào thời điểm tháng 8 và tháng 11 phương pháp hồi quy tuyến tính đa biến chiếm ưu thế. Với cả hai yếu tố dự báo trên mô hình Cancm4i và mô hình Gfdlspear có ưu thế hơn so với các mô hình khác. Khu vực Tây Nguyên và Nam Bộ có kết quả dự báo chính xác hơn các khu vực còn lại. Với chỉ số RPSS, phương pháp HQTPC cho kết quả dự báo tốt nhất.

**Từ khóa:** Dự báo thống kê sau mô hình, hạn đến 01 năm, 7 khu vực, nhiệt độ thấp nhất trung bình, nhiệt độ cao nhất trung bình.

## 1. Mở đầu

Để tiến hành dự báo mùa, với phương pháp dự báo số trị tổ chức khí tượng thế giới WMO đã đưa ra bản cập nhật khí hậu theo mùa toàn cầu (GSCU) cung cấp thông tin về điều kiện khí hậu hiện tại và dự báo theo mùa trên quy mô toàn cầu [1], bên cạnh đó phương pháp dự báo thống kê cũng được sử dụng rộng rãi đặc biệt với hạn trên 6 tháng [2]. Phương pháp thống kê sau mô hình, đã bắt đầu được áp dụng với phương pháp dự báo hoàn hảo PP (Perfect Prog) do Klein và cộng sự (1959) [3] đề xuất. Sau đó một thập niên, Glahn và Lowry (1972) [4] đề xuất phương

pháp thứ hai thống kê sau mô hình MOS (Model Output Statistics). Hai phương pháp này đều có mục đích làm chính xác hơn các kết quả dự báo từ mô hình và cùng dựa trên phương pháp hồi quy tuyến tính đa biến [5]. Phương pháp thống kê sau mô hình MOS dựa trên thuật toán như tương quan Canon [6], [7] hay phương pháp thành phần chính PCR [8], [9]. Đây là phương pháp thông qua việc xây dựng phương trình hồi quy nhằm làm tăng tính độc lập giữa các nhân tố dự báo, giảm thiểu được khối lượng phép tính. Nghiên cứu của Jim Frost [10] cũng chỉ ra phương trình hồi quy là một phần quan trọng trong kết quả thống kê sau bước hiệu chỉnh mô hình, các hệ số trong phương trình để xác định mối quan hệ giữa từng biến độc lập và biến phụ

Liên hệ tác giả: Trần Ngọc Vân

Email: tranngocvan281285@gmail.com

thuộc. Phương pháp thống kê sau mô hình sẽ kết hợp được dữ liệu của sản phẩm mô hình số trị và dữ liệu quan trắc thực tế để đưa ra các sản phẩm dự báo mùa theo nhu cầu quan tâm của người dùng [11], [12].

Nhiệt độ trung bình được coi là một trong những biến khí tượng bề mặt dễ dự đoán nhất [13]. Với các yếu tố cực trị cần nhiều nghiên cứu sâu hơn về các nhân tố dự báo, hoàn lưu quy mô lớn. Theo Soledad và cộng sự (2012) [14] đã sử dụng các mô hình thống kê khí hậu để dự báo số ngày có cực trị nhiệt độ theo mùa với kỹ năng tốt, đặc biệt mùa hè, trong khi đó, kỹ năng thấp lại xuất hiện ở vùng ngoại nhiệt đới. Theo Ali Assani (2017) [15] sử dụng phân tích tương quan Canon nghiên cứu nhiệt độ cực trị, trong đó nhiệt độ cực trị không tương quan với lượng mưa hoặc tuyết mùa đông ở miền Nam Quebec. Nghiên cứu của Kamarainen và cộng sự (2019) [16] đã trình bày dự báo thống kê nhiệt độ hạn mùa cho khu vực Tây Âu và Scandinavi khi xem xét các thành phần chính là nhiệt độ bề mặt biển (SST) và độ cao địa thế vị ở mức 150 hPa. Dự báo này có ý nghĩa hơn khi chuỗi thời gian đủ dài qua đó gia tăng khả năng dự báo của các mô hình thống kê. Một số tác giả đã áp dụng tương quan Canon vào dự báo mùa như dự báo nhiệt độ ở Canada [17]. Trong đó thường sử dụng trường SST của thời kỳ quá khứ, đôi khi sử dụng thêm các trường trong khí quyển để làm nhân tố dự báo.

Bên cạnh đó một số tác giả khác đã sử dụng phương pháp hồi quy thành phần chính để dự báo mùa như tác giả Robert Vautard và các cộng sự [18] dựa trên nhân tố dự báo là nhiệt độ bề mặt biển SST, kết quả đáng tin cậy và có tương quan cao vào các mùa đông, mùa xuân, mùa hè. Đối với mùa đông, khu vực hữu ích nhất là tại Canada và biên giới Bắc Mỹ. Nghiên cứu của Dennett (1977) [19] sử dụng phương pháp phân tích thành phần chính với bộ số liệu mưa 11 trạm tại Mauritius cho thấy, lượng mưa trong mùa hè từ tháng 11-tháng 4 có liên quan chặt chẽ đến vị trí của áp cao cận nhiệt đới ảnh hưởng đến khu vực Đông Úc trong những tháng trước đó, còn dao động Nam ít ảnh hưởng đến lượng mưa tại Mauritius. Với yếu tố cực trị R.Sneyers (1992) [20] đã sử dụng phương pháp phân tích

thành phần chính để phân tích nhiệt độ cực trị trung bình tháng tại mạng lưới đo nhiệt độ của Bờ. Đối với cả nhiệt độ tối cao và nhiệt độ tối thấp, hai thành phần đầu tiên được coi là đóng góp nhiều nhất và có ý nghĩa về khí hậu. Đối với nhiệt độ tối cao, thành phần chính thứ hai phụ thuộc trực tiếp vào khoảng cách gần biển, trong khi ở với nhiệt độ tối thấp, thành phần thứ hai đưa ra sự điều chỉnh do một phần do gần biển và một phần do cấu hình địa điểm của trạm (cao nguyên hoặc thung lũng). Theo Mapitsi Rangata và ccs (2020) [21] đã nghiên cứu sự biến động của nhiệt độ tối cao tháng theo không gian và thời gian ở phía Nam Châu Phi tại 16 điểm thời kỳ 1965-2010, với khoảng thời gian 5 năm, sử dụng phương pháp thành phần chính (PCA). Nghiên cứu chỉ ra rằng, khi sử dụng phương pháp thành phần chính tại 16 điểm ở Nam Phi, mô hình nhiệt độ thay đổi nhiều hơn trong mùa hè ở Úc và ít thay đổi hơn trong mùa đông ở Úc.

Theo Zbigniew Zuska và các cộng sự (2019) [22] đã ứng dụng phương pháp thành phần chính để đánh giá tác động của các yếu tố khí tượng bao gồm: Nhiệt độ tối cao, nhiệt độ tối thấp, nhiệt độ trung bình, tốc độ gió trung bình và cực đại, trong các mùa thu, mùa đông và mùa xuân. Phân tích cho thấy, vào mùa thu, mùa đông và mùa xuân, ảnh hưởng của nhiệt độ tối cao, nhiệt độ tối thấp và nhiệt độ trung bình chiếm ưu thế trong thành phần chính thứ nhất.

Với phương pháp hồi quy tuyến tính đa biến, theo Chul-Gyum Kim (2023) [23] xem xét khả năng dự đoán khí hậu hạn dài ở lưu vực sông Seomjin bằng mô hình hồi quy tuyến tính bội, sử dụng 44 chỉ số khí hậu. Phân tích tỷ lệ trùng pha đối với lượng mưa là 34,7%. Ngược lại, đối với nhiệt độ tỷ lệ trùng pha 40%.

Bài toán dự báo hạn mùa cũng đã và đang được quan tâm nghiên cứu tại Việt Nam. Đã có nhiều nghiên cứu trong nước về sử dụng phương pháp thống kê sau mô hình cho các bài toán hạn mùa như tác giả Tạ Hữu Chinh (2015) [24] đã sử dụng phương pháp hiệu chỉnh thống kê hạ thấp quy mô của các trường dự báo, sau đó xây dựng phương trình dự báo và tiến hành trung bình tổ hợp có trọng số. Kết quả của nghiên cứu cho thấy khả năng dự báo trường nhiệt độ vượt qua dự báo ngẫu nhiên nhiều lần

(HITSS 15-40). Kỹ năng dự báo mưa có xu hướng thấp. Với yếu tố cực trị, trong đề tài KC08.29 [25] đã sử dụng phương pháp phân tích thành phần chính (PCA) đưa ra kết quả dự báo khả năng xuất hiện nắng nóng bằng biểu đồ tin cậy và chỉ số kỹ năng BSS (Brier Skill Score). Trong đề tài TNMT.2022.05.21 [26] bằng các phương pháp thống kê sau mô hình đã đưa ra biến động số ngày nắng nóng có quan hệ mạnh với các mode nhiệt độ bề mặt biển ở khu vực nhiệt đới Thái Bình Dương, Ấn Độ Dương và một số biến đại dương/khí quyển quy mô lớn. Với yếu tố số đợt không khí lạnh (KKL), tác giả Đinh Hữu Dương, Võ Văn Hòa (2018) [27], trình bày kết quả nghiên cứu với nhân tố dự báo được lựa chọn là các chỉ số gió mùa mùa đông dựa trên nghiên cứu của Li Yueqing và Yang Song (2010). Về mặt trung bình, tất cả các phương trình dự báo có xu thế dự báo số lượng đợt KKL trong các tháng mùa đông cao hơn so với thực tế từ 3-5 đợt. Phương trình dự báo đa biến dựa trên 4 chỉ số LiYang, IJhun, IYang và ISUN cho chất lượng dự báo tốt nhất.

Như vậy các tác giả trong và ngoài nước đã

áp dụng nhiều phương pháp thống kê sau mô hình để dự báo yếu tố trung bình và cực trị, thành phần chính được phân tích từ nhân tố dự báo có thể phụ thuộc theo vị trí địa lý hoặc theo mùa. Trong khuôn khổ bài báo này, nhóm tác giả tiến hành đánh giá khả năng dự báo hạn đến 01 năm của nhiệt độ thấp nhất và nhiệt độ cao nhất trung bình bằng ba phương pháp thống kê sau mô hình. Phương pháp và số liệu nghiên cứu sẽ được giới thiệu trong phần 2 và phần 3 sẽ trình bày các kết quả thu được.

## 2. Phương pháp và số liệu nghiên cứu

### 2.1 Số liệu nghiên cứu

Số liệu dự báo nhiệt độ từ 6 mô hình toàn cầu từ năm 1983-2022 theo Bảng 1.

Số liệu quan trắc nhiệt độ và lượng mưa của 186 trạm tại Việt Nam được lấy từ nguồn của Trung tâm Dự báo khí tượng thủy văn quốc gia từ năm 1983-2022. Các trạm được phân thành 7 khu vực: Tây Bắc Bộ (TB Bộ), Việt Bắc Đông Bắc (VB-ĐB), Đồng bằng Bắc Bộ (ĐB Bắc Bộ), Bắc Trung Bộ (BT Bộ), Nam Trung Bộ (N T Bộ), Tây Nguyên (T Nguyên), Nam Bộ (N Bộ).

Bảng 1. Các mô hình dự báo khí hậu toàn cầu

STT	Tên mô hình	Diễn giải mô hình
1	CFSv2	Mô hình dự báo khí hậu của NOAA-NCEP độ phân giải 1,0x1,0 độ
2	CanSipsv2	Hệ thống tổ hợp đa mô hình sử dụng 02 mô hình khí hậu được phát triển bởi Trung tâm phân tích mô hình khí hậu Canada và Trung tâm khí tượng Canada, độ phân giải 2,5x2,5 độ.
3	CanCM4i	Mô hình dự báo khí hậu của Trung tâm khí tượng Canada độ phân giải 1x1 độ
4	GFDL-CM2p1-ae04	Mô hình tổ hợp Bắc Mỹ độ phân giải 1x1 độ
5	NASA-GEOS2S	Tổ hợp đa mô hình Bắc Mỹ độ phân giải 1,0x1,0 độ
6	GFDL_SPEAR	Mô hình khí hậu kết hợp trên nền tảng của mô hình dự báo khí hậu - thập kỷ GFDL, với các mô phỏng theo phương ngang của khí quyển. Độ phân giải 1,0x1,0 độ

### 2.2 Phương pháp nghiên cứu

+ Các phương pháp thống kê sau mô hình được sử dụng trong nghiên cứu bao gồm: - Phương pháp tương quan Canon (CCA) sử dụng hàm trực giao tự nhiên - EOF (empirical orthogonal function) phân tích cả nhân tố dự báo X và yếu tố dự báo Y thành phần theo không gian và thời gian được mô tả bằng các vectơ riêng

trực giao với nhau. Phương pháp này giúp tìm ra mối quan hệ tuyến tính giữa 2 trường yếu tố và cũng có thể kết hợp với hồi quy để thực hiện dự báo.

- Phương pháp hồi quy thành phần chính (PCR) là một kỹ thuật thống kê phân tích đa biến được sử dụng rộng rãi mà để giảm tập hợp các biến phụ thuộc bằng cách phép biến đổi trực giao để biến đổi một tập hợp dữ liệu từ một

không gian nhiều chiều sang một không gian mới ít chiều hơn (2 hoặc 3 chiều) nhằm tối ưu hóa việc thể hiện sự biến thiên của dữ liệu.

- Phương pháp hồi quy tuyến tính đa biến: Hồi quy tuyến tính biến Y theo các biến X<sub>j</sub> (j=1,..., m) chính là xây dựng mối quan hệ giữa biến Y với tổ hợp tuyến tính của các biến X<sub>j</sub>. Trước hết tính toán hệ số tương quan tại từng điểm trạm với số liệu trên từng điểm lưới của mô hình sau đó tìm kiếm những điểm lưới có hệ số tương quan thỏa mãn 1 ngưỡng lựa chọn. Cuối cùng sẽ xây dựng phương trình hồi quy giữa điểm trạm quan trắc với các điểm lưới thỏa mãn hệ số tương quan

- Phương pháp hiệu chỉnh phân vị (quantile mapping-QM): Trong khuôn khổ nghiên cứu bài báo các tác giả đưa ra phương pháp hiệu chỉnh trên nhằm chuyển đổi các dữ liệu mô hình hóa để dữ liệu trên có được các phân bố xác suất giống như các dữ liệu thực đo thông qua một hàm chuyển đổi:

$$Q_m(t) = FO^{-1}[F_s[Q_s(t)]] \quad (1)$$

Với  $Q_m(t)$ ,  $Q_s(t)$  lần lượt là số liệu trước hiệu chỉnh và số liệu sau hiệu chỉnh của mô hình,  $F_s$  là hàm phân bố tích lũy của số liệu dự báo của mô hình,  $FO^{-1}$  nghịch đảo của hàm phân bố tích lũy của số liệu quan trắc.

Tiến hành đánh giá dự báo số liệu dự báo của 6 mô hình toàn cầu như sau:

Bước 1: Tải số liệu mô hình và chuyển file số liệu mô hình theo số liệu lưới trạm và theo năm.

Bước 2: Xây dựng chương trình dự báo bằng ba phương pháp thống kê sau mô hình tương quan Canon (TQ Canon), hồi quy thành phần chính (HQTPC), hồi quy tuyến tính đa biến (HQTTĐB).

Bước 3: Xây dựng chương trình tính toán đánh giá sai số dựa trên chuỗi số liệu phụ thuộc từ năm 1983-2012 và chuỗi số liệu độc lập từ năm 2013-2022.

Ở bước 1, số liệu quan trắc sau khi được tải về sẽ được thiết lập chương trình để đưa về dạng sau: Số liệu theo phương ngang là 186 trạm quan trắc trên phạm vi toàn quốc. Số liệu theo phương dọc là từ năm 1983-2022.

Số liệu của 6 mô hình dự báo toàn cầu: CFSv2, CanSips, CanCM4i, GFDL-CM2p1-ae04, NASA-

GEOSS2S, GFDL\_SPEAR được tải về và phân bố theo từng điểm lưới như hình theo từng điểm lưới theo phương dọc từ 0-90 độ Bắc và theo phương ngang từ 0-360 độ Đông

Ở bước 2 chương trình tính toán thống kê sau mô hình được xây dựng theo 3 phương pháp tương quan Canon, phân tích thành phần chính và hồi quy tuyến tính đa biến. Sản phẩm của chương trình này bao gồm các bản đồ dự báo xác suất và phương trình dự báo

Bước 3 Xây dựng chương trình tính toán đánh giá sai số dựa trên chuỗi số liệu phụ thuộc từ năm 1983-2012 và chuỗi số liệu độc lập từ năm 2013-2022.

Các sai số được sử dụng để đánh giá bao gồm: Độ chính xác dự báo pha (Accuracy; %)- Dự báo 2 pha (Accu2P): Các sự kiện được phân chia làm 2 pha: Dưới chuẩn, trên chuẩn. Accu2P định nghĩa là số sự kiện dự báo đúng trong tổng số sự kiện 2 pha. - Dự báo 3 pha (Accu3P), các sự kiện được phân chia làm 3 pha: Dưới chuẩn, trên chuẩn và cận chuẩn. Accu3P định nghĩa là số sự kiện dự báo đúng trong tổng số sự kiện 3 pha. AccuQ3p và AccuQ2p lần lượt độ chính xác 3 pha và 2 pha có hiệu chỉnh bằng phương pháp QM.

- Chỉ số kỹ năng RPSS (The Ranked Probability Skill Score):

Chỉ số RPS là chỉ số bình phương tích lũy giữa 3 pha dự báo và 3 pha quan trắc. Ở đây pha quan trắc được tính 100% ở pha xảy ra thực tế và 0% ở hai pha còn lại. Chỉ số này được tính theo công thức sau:

$$RPS = \sum_{cat=1}^{N_{cat}} (Pcum_{F(cat)} - Pcum_{O(cat)})^2 \quad (2)$$

RPS càng nhỏ thì dự báo càng tốt, tốt nhất khi RPS=0.

Chỉ số kỹ năng RPSS so sánh kỹ năng dự báo RPS với kỹ năng dự báo của nền khí hậu trung bình. Chỉ số RPSS càng lớn chứng tỏ kỹ năng dự báo tốt hơn so với kỹ năng của nền khí hậu trung bình.

$$RPSS = 1 - \frac{RPS_{forecast}}{RPS_{reference}} \quad (3)$$

Kết quả chương trình đánh giá sai số bao gồm độ chính xác 3pha chuỗi phụ thuộc (Accu\_3p), độ chính xác 3 pha có hiệu chỉnh chuỗi phụ thuộc (AccuQ\_3p), độ chính xác 2pha chuỗi phụ thuộc (Accu\_2p), độ chính xác 2 pha có hiệu chỉnh chuỗi phụ thuộc (AccuQ\_2p), độ chính xác 3pha chuỗi độc lập (Accu\_3pFCT), độ chính xác 3 pha có hiệu chỉnh chuỗi độc lập (AccuQ\_3pFCT), độ chính xác 2pha chuỗi độc lập (Accu\_2pFCT), độ chính xác 2 pha có hiệu chỉnh chuỗi phụ thuộc (AccuQ\_2pFCT), chỉ số kỹ năng RPSS.

### 3. Kết quả và thảo luận

#### 3.1. Đánh giá kết quả dự báo nhiệt độ cao nhất trung bình

##### 3.1.1. Đánh giá kết quả dự báo nhiệt độ cao nhất hạn 6 tháng

Nhóm nghiên cứu đã đánh giá độ chính xác 3 pha (Accu3p) và độ chính xác 2 pha (Accu2p) trên 2 chuỗi phụ thuộc và chuỗi độc lập tại các thời điểm tháng 2, tháng 5 tháng 8 và tháng 11. Ngoại trừ trường hợp tháng 2 dự báo cho tháng 6-8 và tháng 5 dự báo cho tháng 9-11, phương pháp HQTĐB có kết quả tốt nhất với Accu3p đạt 0,5-0,6, Accu2p đạt 0,7-0,8. Những trường hợp còn lại HQTTPC có kết quả dự báo tốt nhất Accu3p đạt 0,5-0,6, Accu2p đạt 0,7-0,8.

Dựa trên độ chính xác 3pha Accu3p và độ chính xác 2 pha Accu2p, nhóm tác giả tính toán và lọc ra trường hợp mô hình và khu vực có kết quả cao nhất.

Với phương pháp tương quan Canon mô hình Cancmi4 có tần suất cao nhất từ 27-42%. Với hai phương pháp HQTTPC và HQTĐB mô hình Gfdlaer04 và mô hình Gfdlspear có tần suất cao nhất từ 22-33%. Với phương pháp TQ Canon, khu vực Tây Bắc có tần suất xuất hiện cao nhất đạt 37%. Với hai phương pháp còn lại khu vực Nam Bộ có tần suất cao nhất từ 50-67%.

Chỉ số RPSS được tính toán cho từng điểm trạm và tính toán tỷ số RPSS dương trung bình cho khu vực và các mô hình. Với phương pháp HQTTPC có tỷ số RPSS dương cao nhất dao động từ 40-60%.

##### 3.1.2. Đánh giá kết quả dự báo nhiệt độ cao nhất trung bình hạn từ 7-12 tháng

Với phương pháp HQTTPC và HQTĐB, độ

chính xác 3 pha Accu3p đạt 0,5, độ chính xác 2 pha Accu2p đạt 0,6-0,7, trong khi đó phương pháp TQ Canon, Accu3p đạt 0,3-0,4, Accu2p đạt 0,5.

Với phương pháp TQ Canon mô hình Gfdlspear có kết quả tốt nhất với tần suất có Accu3p và Accu2p cao nhất đạt từ 25-35%. Với phương pháp HQTTPC và phương pháp HQTĐB mô hình Cansips tần suất đạt 50-73% và mô hình Gfdlspear tần suất đạt 21-41%. Hầu hết với 3 phương pháp khu vực Nam Bộ có tần suất xuất hiện cao nhất đạt từ 39-80%.

Phương pháp HQTTPC có chỉ số RPSS cao nhất với tỷ số từ 30-58%, phương pháp HQTĐB phổ biến tỷ số RPSS dương phổ biến 20-40%, riêng phương pháp tương quan Canon chỉ số RPSS rất thấp.

#### 3.2. Đánh giá kết quả dự báo nhiệt độ thấp nhất trung bình

##### 3.2.1. Đánh giá kết quả dự báo nhiệt độ thấp nhất hạn từ 3-6 tháng

Với yếu tố dự báo nhiệt độ thấp nhất trung bình hạn 6 tháng phương pháp HQTĐB có kết quả dự báo tốt nhất với độ chính xác 3 pha đạt 0,5-0,7, độ chính xác 2pha đạt 0,8-0,9.

Dựa trên độ chính xác 3pha, với phương pháp TQ Canon mô hình Gfdlspear có tần suất cao nhất đạt 37%, với phương pháp HQTTPC và phương pháp HQTĐB: Mô hình Cansips với tần suất 29%. Dựa trên độ chính xác 2pha, với phương pháp TQ Canon mô hình Cancm4i có tần suất cao nhất đạt 39%, với phương pháp HQTTPC và phương pháp HQTĐB: Mô hình Gfdlspear với tần suất 25-29%. Dựa trên Accu3p và Acc2p với phương pháp TQ Canon khu vực Tây Bắc có kết quả tốt nhất với tần suất từ 20-30%. Với hai phương pháp còn lại khu vực Tây Nguyên và Nam Bộ có tần suất cao nhất đạt 27-39%.

Chỉ số ROCB đạt từ 0,5-0,9. Với phương pháp HQTTPC, tỷ số RPSS dương đạt 20-45%, với phương pháp HQTĐB tỷ số RPSS dương đạt 20-60%.

##### 3.2.2. Đánh giá kết quả dự báo nhiệt độ thấp nhất hạn từ 7-12 tháng

Độ chính xác Accu3p của phương pháp HQTTPC cao nhất đạt 0,5, riêng chuỗi độc lập với Accu3p của phương pháp này đạt 0,6-0,7, hai

phương pháp còn lại có Accu3p thấp hơn đạt 0,3-0,5. Sau hiệu chỉnh với chuỗi độc lập Accu3p của phương pháp HQTPC và tương quan Canon tăng khoảng 0,1. Với chuỗi phụ thuộc độ chính xác Accu2p đạt từ 0,6-0,7 với phương pháp HQTPC và HQ TTĐB với phương pháp TQ Canon đạt 0,5. Sau hiệu chỉnh Accu2p hầu như ít thay đổi.

Với phương pháp TQ Canon mô hình Gfdlspear tần suất xuất hiện Accu3p và Accu2p cao nhất đạt 20-42%, mô hình Cancm4i với tần suất từ 22-33%. Với phương pháp HQTPC và phương pháp HQTTĐB mô hình Cansip có tần suất cao nhất đạt từ 41-63%, mô hình Gfdlspear có tần suất từ 22-33%. Với phương pháp TQ Canon khu vực Tây Bắc có tần suất xuất hiện Accu3p và Accu2p cao nhất đạt từ 39-42%, trong khi đó hai phương pháp còn lại khu vực Tây Nguyên có tần suất cao nhất đạt 40-59%.

Với tỷ số RPSS, phương pháp tương quan Canon chỉ số chỉ ở ngưỡng dưới 10%, trong khi 2 phương pháp còn lại đạt từ 30-50%.

#### 4. Kết luận

Kết quả nghiên cứu cho thấy với hạn dự báo 6 tháng với yếu tố dự báo nhiệt độ cao nhất trung bình đánh giá trên hai chuỗi phụ thuộc và chuỗi độc lập, phương pháp HQTPC có Accu3p và Accu2p cao nhất với giá trị lần lượt 0,5 và 0,6-0,7. Trong khi đó với yếu tố nhiệt độ thấp nhất trung bình, phương pháp HQTTĐB có ưu thế hơn với giá trị Accu3p đạt 0,5-0,7 và Accu2p đạt 0,7-0,8. Với hạn dự báo từ 7 tháng đến 01 năm với yếu tố dự báo nhiệt độ cao nhất trung bình thời điểm dự báo tháng 2 và tháng 5 phương pháp HQTTĐB có ưu thế hơn, trong khi đó thời điểm tháng 8 và tháng 11 phương pháp HQTPC có ưu thế hơn. Trong khi đó với yếu tố nhiệt độ thấp nhất trung bình phương pháp HQTTĐB có ưu thế nhất, riêng thời điểm dự báo tháng 8 và tháng 11

hạn dự báo 6 tháng, Accu2p đạt 0,8 và 0,9.

Với yếu tố nhiệt độ cao nhất trung bình dự báo bằng phương pháp TQ Canon, mô hình Cancm4i chiếm ưu thế tất cả các hạn dự báo, với phương pháp HQTPC, mô hình Gfdlspear có ưu thế hơn so với các mô hình khác, phương pháp HQTTĐB mô hình Gfdlspear có ưu thế ở hạn từ 6 tháng. Với yếu tố nhiệt độ thấp nhất trung bình hạn 6 tháng mô hình Gfdlspear có ưu thế hơn ở phương pháp tương quan Canon và HQTPC, trong khi đó ở hạn 7-12 tháng mô hình Cancm4i và mô hình Gfdlspear có ưu thế hơn so với các mô hình khác.

Với yếu tố nhiệt độ cao nhất trung bình dựa trên chỉ số Accu3p và Accu2p ở cả 3 phương pháp, khu vực Nam Bộ là nơi có tần suất xuất hiện cao hơn so với các khu vực khác. Với yếu tố nhiệt độ thấp nhất trung bình, khu vực Tây Nguyên và Nam Bộ có chỉ số Accu2p và Accu3p tốt nhất với phương pháp HQTTĐB và HQTPC, còn với phương pháp TQ Canon khu vực Tây Bắc có kết quả dự báo tốt nhất. Với yếu tố nhiệt độ cao nhất trung bình và nhiệt độ thấp nhất trung bình, tỷ số RPSS dương của phương pháp HQTPC tốt nhất đạt từ 40-55%.

Ở hai yếu tố dự báo nhiệt độ thấp nhất trung bình và nhiệt độ cao nhất trung bình phép hiệu chỉnh ánh xạ phân bố có tác động rõ rệt ở chuỗi độc lập, độ chính xác 3 pha Accu3p ở cả ba phương pháp đều tăng khoảng 0,1-0,2.

Như vậy với yếu tố nhiệt độ thấp nhất trung bình phương pháp HQTTĐB chiếm ưu thế hơn, trong khi đó với yếu tố nhiệt độ cao nhất trung bình phương pháp HQTPC chiếm ưu thế hơn. Có thể sử dụng phương pháp thống kê sau mô hình để dự báo nhiệt độ cao nhất và nhiệt độ thấp nhất trung bình hạn từ 3 tháng đến 01 năm với độ chính xác 3 pha đạt 0,5, độ chính xác 2 pha đạt 0,6-0,7.

**Đóng góp của tác giả:** Xây dựng ý tưởng nghiên cứu: Hoàng Phúc Lâm, Trần Ngọc Vân, Nguyễn Đăng Mậu, Nguyễn Đức Hòa, Nguyễn Văn Hưởng, Nguyễn Văn Huấn, Hoàng Thị Mai, Nguyễn Thanh Hoa, Trịnh Thùy Nguyên; Xử lý số liệu: Trần Ngọc Vân, Hoàng Thị Mai, Nguyễn Thanh Hoa, Trịnh Thùy Nguyên, Nguyễn Thanh Thủy; Lấy mẫu: Trần Ngọc Vân, Hoàng Thị Mai, Nguyễn Thanh Hoa, Trịnh Thùy Nguyên, Nguyễn Thanh Thủy, Lê Lan Phương, Lê Văn Phong; Viết bản thảo bài báo: Hoàng Phúc Lâm, Trần Ngọc Vân, Nguyễn Đăng Mậu, Nguyễn Đức Hòa, Nguyễn Văn Huấn, Nguyễn Văn Hưởng, Hoàng Thị Mai, Lê Lan Phương, Nguyễn Thanh Hoa, Nguyễn Thanh Thủy, Trịnh Thùy Nguyên, Lê Văn Phong; Chỉnh sửa bài báo: Hoàng Phúc Lâm,

Trần Ngọc Vân, Nguyễn Đức Hòa, Nguyễn Văn Hưởng, Nguyễn Văn Huấn, Nguyễn Thanh Hoa, Hoàng Thị Mai.

**Lời cảm ơn:** Nghiên cứu này được thực hiện dưới sự tài trợ của đề tài nghiên cứu khoa học cấp Bộ Tài Nguyên Môi Trường, mã số TNMT.2023.06.02.

**Lời cam đoan:** Tập thể tác giả cam đoan bài báo này là công trình nghiên cứu của tập thể tác giả, chưa được công bố ở đâu, không được sao chép từ những nghiên cứu trước đây; không có sự tranh chấp lợi ích trong nhóm tác giả.

### Tài liệu tham khảo

1. WMO, (2020), *Guidance on Operational Practices for Objective Seasonal Forecasting*, WMO-No 1246
2. Ester Salium et al. (2015), "Skill evaluation of the CanCM4 and its MOS for seasonal rainfall forecast in Malaysia during the early and late winter monsoon periods", *International Journal of Climatology*, <https://doi.org/10.1002/joc.4361>
3. Klein, W.H., B.M. Lewis and I. Enger, (1959), "Objective prediction of five-day mean, temperature during winter", *J. Meteor.*, 16, 672-682.
4. Glahn, H.R., Lowry, D.A. (1972), "The use of model output statistics (MOS) in objective weather forecasting", *Journal of applied meteorology*, 11, 1203-1211.
5. Wills, D, (2006), *Statistical Methods in the Atmospheric sciences*.
6. Barnston, A. G., He Y. (1996), "Skills of CCA forecasts of 3-month mean surface climate in Hawaii and Alaska", *J. Clim.*, 9, 2579 - 2605.
7. Amir Shabbar, Anthony G.Barnston, (1996), "Skill of seasonal climate forecasts in Canada using Canonical Correlation Analysis", *monthly weather review*, volume 124, 2370-2380
8. G. Louis Smith (2012), "Potter A technique using principal component analysis to compare seasonal cycles of Earth radiation from CERES and model computations", *Journal of Geophysical Research Atmospheres*, p1120-1130, <https://doi.org/10.1029/2011JD017343>
9. Deanna Wilson-Diaz (2001), *A principal component analysis of sea-surface temperature in the Arabian Sea, Deep Sea Research Part II: Topical Studies in Oceanography*, Volume 45, Issues 6-7, 2001, Pages 1097-1114.
10. Jim Fort (2019), *Making Predictions with Regression Analysis*, The American Astronomical Society. All rights reserved. The Astronomical Journal, Volume 157.
11. Fowler, H. J., and Wilby, R. L. (2007), "Beyond the downscaling comparison study", *International Journal of Climatology*, 27: 1543-1545 (2007).
12. Fowler, H.J. et al. (2007), Review: Linking climate change modeling to impacts studies-Recent advances in downscaling techniques for hydrological modelling: *International Journal of Climatology*, v. 27, p. 1547-1578
13. Haiden, T, et al. (2015), *Evaluation of ECMWF forecasts*, including 2014-2015 upgrades ECMWF Techn. Memoranda 2015(1.1).
14. Soledad Collazo (2021), "Seasonal forecast of the percentage of days with extreme temperatures in central-northern Argentina: An operational statistical approach", *Climate Services*, Volume 26, April 2022, <https://doi.org/10.1016/j.cliser.2022.100293>
15. Ali Assani, Nadjat Guerfi (2017), "Analysis of the Joint Link between Extreme Temperatures, Precipitation and Climate Indices in Winter in the Three Hydroclimate Regions of Southern Quebec", *Atmosphere* 8(4):1-13, article no 75 DOI:10.3390/atmos8040075 (1.1)
16. Kamarainen, M. et al. (2019), "Statistical learning methods as a basis for skillful seasonal temperature forecasts in Europe", *J. Clim*, 32, pp. 5363-5579.
17. Anthony G Barnston and Chester F.Ropelewski, (1981), "Prediction of ENSO Episodes Using

*Canonical Correlation", Journal of Climate, Volume 5, 1316-1320*

18. Robert Vautard, Guy Plaut, Risheng Wang, Seasonal Prediction of North American Surface Air Temperatures Using Space-Time Principal Components,
19. M. D. Dennett, Variation of Rainfall and seasonal forecasting in Mauritius –, 8-1977, Arch. Met. Geoph. Biokl., Ser. B, 35, 359-370
20. R. Sneyers, M. Vandiepenbeeck, R. Vanlierde, Principal component analysis of air temperature of Belgium, March 1992, Volume 5, 1992, theoretical and Applied Climatology (2.1)
21. Mapitsi Rangata & Sonali Das & Montaz Ali, 2020. "Analysing Maximum Monthly Temperatures in South Africa for 45 years Using Functional Data Analysis," Advances in Decision Sciences, International Journal of Climate, vol. 24(3), pages 1-27, September. <https://doi.org/10.1002/joc.4361>
22. Zbigniew Zu'ska et al. (2019), "Application of the Principal Component Analysis (PCA) Method to Assess the Impact of Meteorological Elements on Concentrations of Particulate Matter (PM10): A Case Study of the Mountain Valley (the Sącz Basin, Poland)", Sustainability, Volume 11, Issue 23 MDPI, 11, 6740; doi:10.3390/su11236740 [www.mdpi.com/journal/sustainability](http://www.mdpi.com/journal/sustainability)
23. Chul-Gyum Kim et al. (2023), "Evaluation of Seasonal Climate Predictability Considering the Duration of Climate Indices", Water 2023, 15(18), 3291; <https://doi.org/10.3390/w15183291>
24. Tạ Hữu Chính (2015), Nghiên cứu xây dựng hệ thống dự báo tổ hợp hạn mùa dựa trên kết quả dự báo của các mô hình nghiệp vụ toàn cầu, Đề tài mã số: TNMT.05.25, Cơ quan chủ trì đề tài: Trung tâm Dự báo khí tượng thủy văn Trung ương, Trung tâm KTTV quốc gia.
25. Phan Văn Tân (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, Đề tài cấp Nhà nước, mã số KC08.29/06-10.
26. Tạ Hữu Chính (2022), Nghiên cứu ứng dụng các chỉ số dao động khí hậu quy mô lớn vào dự báo khí hậu cho Việt Nam, Đề tài cấp Bộ, mã số TNMT.2021.02.05.
27. Đinh Hữu Dương và Võ Văn Hòa (2018), "Thử nghiệm dự báo hạn mùa số đợt không khí lạnh trong các tháng chính đông dựa trên cách tiếp cận dự báo hoàn hảo", Tạp chí Khí tượng Thủy văn, 696, 54-61.

## EVALUATING THE ABILITY OF THE AVERAGE MINIMUM AND MAXIMUM TEMPERATURE FORECASTS UP TO 1-YEAR USING STATISTICAL METHODS

Tran Ngoc Van<sup>(1)</sup>, Hoang Phuc Lam<sup>(1)</sup>, Nguyen Dang Mau<sup>(2)</sup>, Nguyen Duc Hoa<sup>(1)</sup>,  
Nguyen Van Huong<sup>(1)</sup>, Nguyen Van Huan<sup>(3)</sup>, Hoang Thi Mai<sup>(1)</sup>, Nguyen Thanh Hoa<sup>(1)</sup>,  
Trinh Thuy Nguyen<sup>(1)</sup>, Phuong Le Lan<sup>(1)</sup>, Nguyen Thanh Thuy<sup>(4)</sup>, Le Van Phong<sup>(2)</sup>

<sup>(1)</sup>National Centre For Hydro-Meteorological Forecasting

<sup>(2)</sup>The Viet Nam Institute of Meteorology, Hydrology and Climate Change

<sup>(3)</sup>Centre Highland Regional Hydrometeorological Center

<sup>(4)</sup>Northern Delta and Midland Regional Hydro-Meteorological Center

Received: 23/9/2024; Accepted: 24/10/2024

**Abstract:** In this study, the author used three statistical methods: The canonical correlation analysis method (CCA), the principal component regression method (PCR), and the multivariate linear regression method (MLR) for 1-year the average minimum and maximum temperature forecasts. The data from six global models and observations from 1983 to 2022 served as initial data. The results show that the MLR has an advantage over the other two methods with 3-phase accuracy reaching 0.5-0.6 and 2-phase accuracy



reaching 0.7-0.8 for average minimum temperature forecast. On the other hand, for the mean maximum temperature forecast, the PCR dominates the forecasting periods of February and May, while the MLR has better performance in August and November. The Cancm4i and the Gfdlspear models prevail for both forecasting factors. The Central Highlands and the South have more accurate forecasting results than the remaining regions. With the RPSS index, the PCR method gives the best forecasting results.

**Keywords:** Model output statistic, up to 1 year, 7 regions, the average minimum and maximum temperature.