Forecasting Extreme Weather Events Using Statistical Methods Up to 6 Months in Vietnam

Tran, N. V.,[1] Hoang, T. V.,[2*] Mai, V. K.,[1] Nguyen, T. H.,[1] Hoang, T. M.[1] and Trinh, T. N.[2]
[1]Vietnam National Center for Hydrology Meteorology Forecasting, Hanoi, Vietnam
E-mail: tranngocvan281285@gmail.com, tvhoang@o365.fcu.edu.tw, maikhiem77@gmail.com
ms.hoa2510@gmail.com, hoangmaik52dubao@gmail.com, nguyentt607@gmail.com
[2]Feng Chia University, No 100 Wenhua Rd, Situn 40724 Taichung city, Taiwan
**Corresponding Author*

DOI: https://doi.org/10.52939/ijg.v21i8.4373

Abstract
This study evaluates the forecasting performance of three statistical methods Canonical Correlation Analysis (CCA), Principal Component Regression (PCR), and Multivariate Linear Regression (MLR) in predicting extreme weather events in Vietnam up to six months in advance. Among the methods assessed, PCR consistently outperforms both CCA and MLR across various climatic regions and event durations (i.e., 2–4 days and over 5 days), particularly in forecasting cold waves and moderate to heavy rainfall during the October–December period. By transforming correlated predictors into orthogonal components, PCR effectively addresses issues of multicollinearity, facilitates dimensionality reduction, and captures dominant climate signals with greater precision. Notably, PCR demonstrates superior adaptability to multiple climate models (e.g., CanCM4i, GFDL-SPEAR, CanSIPSv2) and maintains stable predictive accuracy without heavy reliance on post-processing techniques such as quantile mapping an advantage not observed with CCA and MLR. A key contribution of this research is the integration of PCR into a long-range climate forecasting framework tailored for Vietnam, enabling more robust and reliable predictions of extreme events across diverse climatic zones. Furthermore, this study introduces a seasonal drought forecasting system that produces probabilistic maps of average conditions, extreme climate indices, and drought occurrences up to six months in advance. By addressing the limitations of conventional statistical approaches and enhancing prediction accuracy, this research offers critical insights for disaster risk management and climate adaptation planning in Vietnam.

Keywords: Canonical Correlation Analysis, Climate models, Cold wave, Extreme weather events, Heat wave prediction, Heavy rainfall, Multivariate Linear Regression, Principal Component Regression, Seasonal climate forecasting, Vietnam

1. Overview

Numerical forecasting models play a vital role in seasonal prediction due to their ability to effectively simulate atmospheric conditions [1]. Their use allows researchers to conserve time and resources [2]. These models are commonly employed to forecast seasonal averages of temperature [3], total precipitation [4], and hazardous weather phenomena [5]. Incorporating multiple predictive factors particularly those related to storm activity has been shown to yield more accurate forecasts than relying on single-variable models [6]. For instance, early forecasts from the ECMWF system issued in May provide valuable insights into potential summer heatwave conditions [7].

Despite these advancements, climate forecasting centers still face challenges in providing accurate predictions for extreme weather events, such as severe cold spells, heatwaves, and intense rainfall. Presently, most centers offer only probabilistic forecasts for temperature and precipitation levels, without specific guidance on extreme phenomena. To address this limitation, statistical techniques such as Canonical Correlation Analysis (CCA) [8] and Principal Component Regression (PCR) [9] have been utilized to develop targeted forecast products.

Perfect Prog was introduced in 1959 [10], followed by the development of the Model Output Statistics (MOS) approach in 1972 [11]. Both approaches leverage multivariate linear regression models to enhance forecast accuracy [12]. Regression equations play a critical role in quantifying the relationship between each predictor and the target variable following model calibration [13].

Post-processing statistical methods further refine seasonal forecasts by integrating deterministic model outputs with observed data [14] and [15]. Given the influence of climate change, rainfall patterns are projected to shift, with increased precipitation in high-latitude regions and decreases in many subtropical zones [16]. Understanding the seasonal behavior of extreme weather phenomena has thus become increasingly important [17]. However, forecasting both general and extreme rainfall remains a complex challenge. Consequently, researchers have applied hybrid approaches to improve predictive performance. For example, long-term forecasts of summer monsoon rainfall over northwestern India and the Indian Peninsula were developed using a combination of Multivariate PCR, Neural Networks (NN), and Linear Discriminant Analysis (LDA) [18]. Similarly, monsoon rainfall in Pakistan has been predicted using both Multiple Linear Regression (MLR) and PCR [19].

Post-model statistical approaches that incorporate diverse predictor variables have also been widely adopted. Global sea surface temperatures (SSTs) have been shown to be effective for forecasts up to five months in advance [20]. Rainfall in East Asia including Korea and Japan can be forecast using global SST data in combination with 700-hPa geopotential height [21]. Spring has been identified as the optimal season for rainfall prediction. For instance, sea surface temperature anomalies, integrated vapor transport (IVT), and 250-hPa geopotential height (Z250) were used as key predictors [22]. Meanwhile, mean sea level pressure (MSLP) outputs from the GloSea5 and SEAS5 dynamical models have been applied for forecasting winter and summer rainfall [23]. Forecast horizons of up to four months have shown reasonably high accuracy when applying MLR and Artificial Neural Network (ANN) models.

Extreme rainfall has been defined using the 99th percentile threshold, calculated separately for each grid point across the entire dataset [24] and [25]. Reliability diagrams have been used to illustrate forecast errors. Forecasts require datasets spanning over 20 years from the initial start date. The probability of accurately forecasting heavy rainfall within 1–2 days is approximately 30–40%, decreasing with longer lead times. Long-range forecasting is crucial for extreme climate phenomena such as severe cold spells. Extreme cold days are correlated with large-scale circulation patterns, such as the Arctic Oscillation (AO) [26] and [27]. During AO phases, more active Siberian high pressure allows cold air to penetrate further south, resulting in increased extreme cold days during winter. Cold air incursions also follow the ENSO cycle, with higher frequencies in East Asia and the Northwest Pacific during La Niña phases and lower frequencies during El Niño [28]. In Taiwan, cold air intrusions are associated with the North Atlantic Oscillation (NAO), particularly during its negative phases [29]. One study investigated the sources and limits of seasonal predictability in the total number of extremely cold days (NECD) across China using physics-based empirical models (PEMs), with predictability influenced by tropical Pacific SST anomalies and snow cover in September and October [30].

Another model based on ENSO-related predictive factors showed a 65% success rate in forecasting low-temperature events at the 10th percentile (Tn10) during June to August 2019, while its performance for high-temperature events at the 90th percentile during February to April 2019 was only 18% [30]. In 2023, a forecasting system using the GFDL SPEAR platform was shown to be capable of predicting extreme cold events in North America months in advance [31]. The effects of global warming are also increasing the frequency and severity of heatwaves, affecting health and productivity worldwide. Forecasting their seasonality is therefore essential. Using global climate models from Phase 5 of the Coupled Model Intercomparison Project, three emission scenarios (RCPs 2.6, 4.5, and 8.5) were assessed, showing probabilities of heatwaves in 2100 ranging from 36% to 88% under the high emission scenario [32]. Forecasts of the start and end dates of heatwave episodes have been produced using four S2S models (ECMWF, CMA, UKMO, and NCEP) [33]. Early warning signs of extreme heat conditions in West Africa have also been provided using the CNRM-CM global climate model [34]. However, forecasts of extreme heat events still face limitations, as many atmospheric factors show limited predictive skill beyond three months [35].

A forecast of extreme phenomena in Vietnam will be presented in this paper, including the number of heat waves, cold spells, and moderate to heavy rainfall events for seven regions, as well as four regions for cold spells, with a forecasting period of one year.

2. Data, Method and Study Area

2.1 Study Area

Vietnam is geographically divided into seven distinct climatic regions, each characterized by unique weather patterns shaped by topographical and seasonal influences (Figure 1). The Northwest and Northeast regions exhibit a humid subtropical climate, typified by cold, dry winters and hot, humid summers, with frequent occurrences of severe cold waves and heavy rainfall. The Red River Delta shares

similar climatic characteristics but is particularly vulnerable to flooding and tropical storms during the rainy season, which extends from May to October.

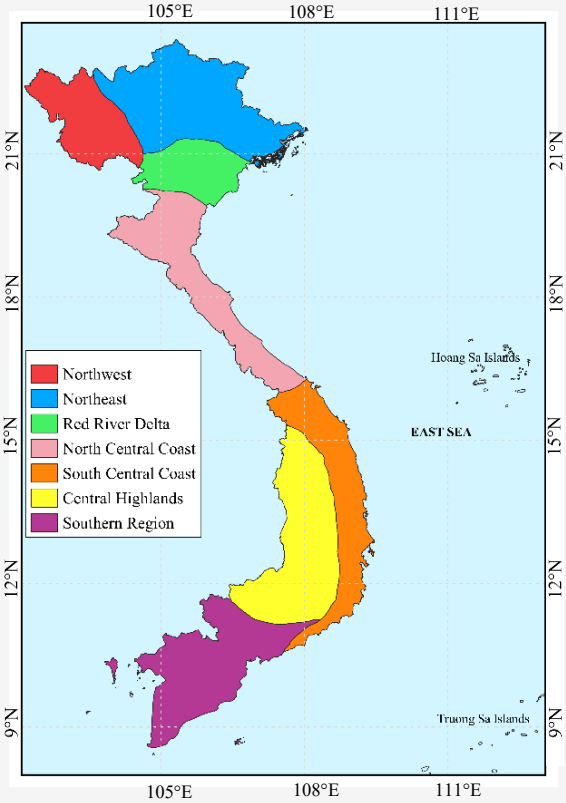


Figure 1: Seven regions of Vietnam

The North Central Coast functions as a transitional climatic zone, experiencing cold, rainy winters and summers that are often affected by typhoons and tropical storms. In contrast, the South Central Coast is characterized by a tropical savanna climate, marked by an extended dry season from January to August and a concentrated rainy season from September to December, during which droughts and occasional typhoons are common. The Central Highlands, due to its higher elevation, experiences a tropical monsoon climate with a cooler temperature regime and well-defined wet (May–November) and dry (December–April) seasons. Southern Vietnam also follows a tropical monsoon pattern, with consistently high temperatures throughout the year and pronounced wet and dry seasons. Flooding is a recurrent issue during the rainy season in this region, particularly in low-lying areas.

2.2 Data

Table 1 shows forecasts of minimum temperatures, maximum temperatures, and rainfall from six global models covering 1983 to 2022. We use six different models for our 6-month forecasting: Cansipsv2, Cancm4i, Gfdlspear [36], Ccsm4 [37], Cfsv2 [38] and Nasageoss2s [39]. Using the predicted maximum temperature, predicted minimum temperature, and rainfall data from the six aforementioned models, we forecast heat waves, cold spells, and moderate to heavy rain events.

2.2.1 Number of heatwaves

In this study, heatwaves in Vietnam are categorized based on the daily maximum temperature (Tx). A day is classified as a heat day when $35°C \leq Tx < 37°C$, as an intense heat day when $37°C \leq Tx < 39°C$, and as a particularly severe heat day when $Tx \geq 39°C$. The term broad heat day refers to a situation in which at least half of the weather stations within a designated forecasting region (e.g., the Northwestern region of Vietnam) record a maximum daily temperature of 35°C or higher. A heatwave is defined as a period during which broad heat conditions persist for two or more consecutive days within the same forecasting region.

2.2.2 Number of cold spells

Cold days are defined as days in which the average daily temperature is less than or equal to 15°C. A cold spell is characterized by the persistence of such temperatures for a minimum duration of two consecutive days.

Table 1: Overview of six global climate prediction

No	Model	Description
1	CFSv2	NOAA-NCEP climate forecast model 1.0x1.0 degree resolution,
2	CanSipsv2	The multi-model ensemble system uses two climate models developed by the Canadian Climate Model Analysis Center and the Canadian Meteorological Center, with a resolution of 1.0x1.0 degree.[40]
3	CanCM4i	North American Multi-Model Ensemble (NMME) 1x1 resolution
4	CCSM4	North American Multi-Model Ensemble (NMME) 1x1 resolution
5	Nasageoss2s	North American Multi-Model Ensemble (NMME) 1x1 resolution
6	Gfdlspear	North American Multi-Model Ensemble (NMME) 1x1 resolution

2.2.3 Number of heavy rain events

A day of moderate to heavy rainfall is defined as any day with a recorded precipitation amount of 16 mm or more. A moderate to heavy rainfall event occurs when such conditions persist for at least two consecutive days and are observed at more than half of the weather stations within a given forecasting region.

2.3 Method

This study employs a range of statistical methods, including Canonical Correlation Analysis (CCA) [41], Principal Component Regression (PCR) [42], and Multivariate Linear Regression (MLR), to analyze and forecast extreme weather events [43], as illustrated in Figure 2, which outlines the workflow from data collection and preprocessing to model application and the generation of seasonal forecasts. In this article, three methods were employed to predict the seasonal duration in Vietnam. Methods such as CCA, PCR, and MLR offer a straightforward and interpretable approach for identifying the linear relationships between large-scale climate drivers (e.g., ENSO, sea surface temperature anomalies) and regional drought-related variables (e.g., rainfall, temperature, dry days) [44]. These models are particularly well-suited for Vietnam's climate, where many key climate–drought interactions, such as the influence of ENSO on summer rainfall [45] or the impact of Western Pacific Sea surface temperature anomalies on droughts in Central Vietnam, exhibit predominantly linear or near-linear patterns. Another advantage of these methods lies in their modest data requirements and ease of training. CCA and PCR, for instance, perform effectively even with relatively small datasets, such as the 30–50 years of climate records typically available in Vietnam. This is crucial for minimizing the risk of overfitting when working with limited historical data. In contrast, advanced machine learning models, including artificial neural networks and deep learning approaches, typically require much larger datasets and involve complex hyperparameter tuning, which may not be feasible given the current data availability and infrastructure in Vietnam.

Furthermore, CCA provides an effective framework for integrating large-scale climate signals into regional forecasts by identifying spatial correlation patterns between global climate indices and local meteorological variables. PCR aids in dimensionality reduction, allowing for the retention of key predictors that explain the majority of variance in the target drought indicators. Lastly, these methods are highly practical for operational use. MLR models, in particular, can be readily implemented within the existing systems of Vietnam's meteorological agencies, such as the National Center for Hydro-Meteorological Forecasting, without the need for extensive computational resources or specialized infrastructure.

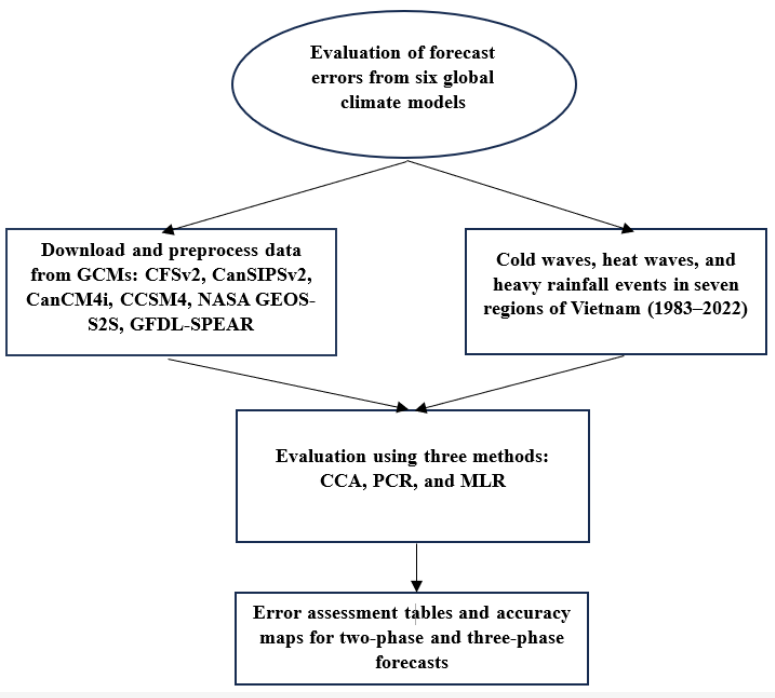


Figure 2: Statistical post-processing framework for seasonal forecasts of extreme events in Vietnam

2.3.1 Principal component regression (PCR)

Principal component regression (PCR) applies an orthogonal transformation to project a dataset from a high-dimensional space into a lower-dimensional space typically two or three dimensions in order to better capture the variability of the data. This approach can be summarized as follows: consider a data field represented as a two-dimensional matrix defined over time (t) and space (s). Through PCR, the data field is decomposed into two components. The space-dependent component is expressed as EOFs(s), a matrix of empirical orthogonal functions (i.e., eigenvectors), while the time-dependent component is represented as *PCs(t)*, or principal components. This decomposition is presented in Equation (1).

$$F(s,t) = EOF_{1(s)}PC_{1(t)} + EOF_{2(s)}PC_{2(t)} + \ldots + EOF_{n(s)}PC_{n(t)}$$

Equation 1

From the original field *F(s,t)*, each *EOF-PC* pair carries a distinct amount of information. In the first component, (*EOF1, PC1*), the majority of the significant information from the original field is captured, which gradually diminishes in successive components (*EOF2, PC2*), (*EOF3, PC3*). The initial components are regarded as the main signal, containing substantial information from the original field, while later components are seen as noise, which can be eliminated. Specifically, the initial field *F(s,t)* represents the global model's predictions for temperature. There are six global models here, and "*t*" indicates the time frame (for instance, 1983-2022). The diagram presents the procedures for constructing a regression forecasting model based on the *PC* coefficients from principal component analysis, which includes the following steps in Figure 3. The following procedure describes the development of a statistical forecasting model using principal component regression. The key steps include dimensionality reduction, model construction, and selection of principal components based on forecast error.

Step 1: Identify the forecast variables (e.g., temperature, precipitation) and corresponding predictors derived from global forecasting models. Perform Principal Component Analysis (*PCA*) on these predictor datasets to obtain principal component (*PC*) coefficients. These PCs are used as forecasting factors in place of the original climate indices. Initially, only the first principal component (PC_1) is used to construct the forecasting model, which takes the form presented in Equation (2).

$$Y = a_1PC_1 + b$$

Equation 2

Step 2: Apply the newly constructed regression equation to the dependent variable time series to perform a reforecast, and calculate the mean forecast error.

Step 3: Repeat Steps 1 and 2 using additional *PCs* to build equations with two, three, or more components. For each case, compute the forecast error on the dependent series. The model that yields the smallest forecast error is selected, and the corresponding combination of *PC* coefficients is retained for forecasting on independent datasets.

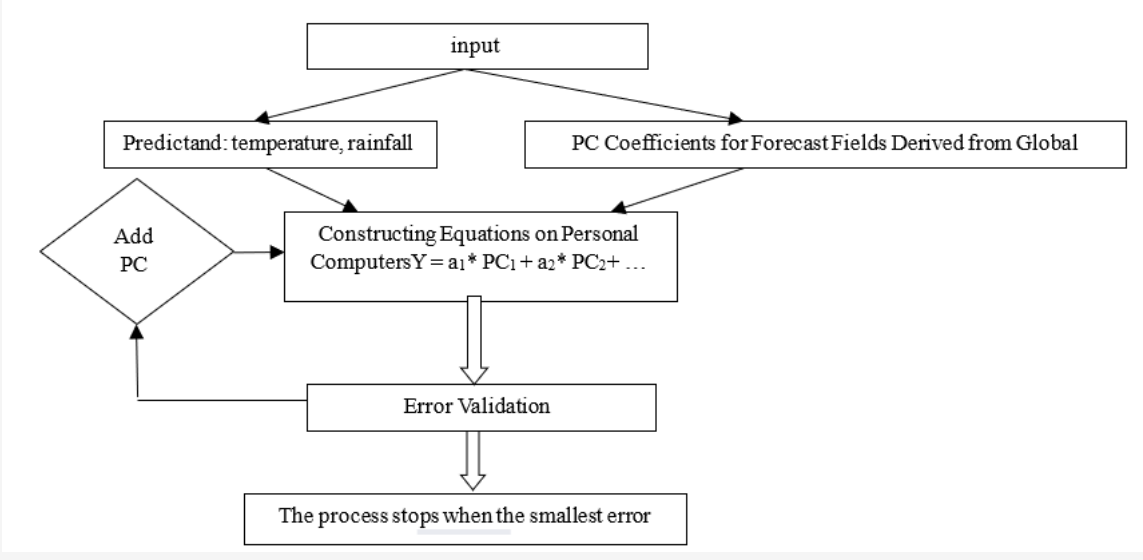


Figure 3: Regression on PC coefficients diagram

2.3.2 Canonical correlation analysis (CCA)

Canonical Correlation Analysis (CCA) is employed to explore relationships between two multidimensional datasets using the *PCA/EOF* framework as defined in Equation 3:

$$F_{1(s,t)} = EOF_{\{1,1\}(s)}PC_{\{1,1\}(t)} + EOF_{\{1,2\}(s)}PC_{\{1,2\}(t)} + \ldots + EOF_{\{1,n\}(s)}PC_{\{1,n\}(t)}$$

Equation 3

$$F_{2(s,t)} = EOF_{\{2,1\}(s)}PC_{\{2,1\}(t)} + EOF_{\{2,2\}(s)}PC_{\{2,2\}(t)} + \ldots + EOF_{\{2,n\}(s)}PC_{\{2,n\}(t)}$$

Equation 4

CCA analyses aim to identify pairs of principal components (*PCs*) that exhibit the strongest correlation between the two datasets, $F_{1(s,t)}$ and $F_{2(s,t)}$, as presented in Equations (3) and (4). In this research, $F_{2(s,t)}$ represents the predicted temperature data from global models, where s indicates the number of global models and t indicates the time dimension. In contrast, $F_{1(s,t)}$ represents climate factors such as heavy rainfall, cold spells, and heatwaves, where *s* represents the number of monitoring stations, and t represents the time period. The Canonical Correlation Analysis (CCA)-based forecasting procedure consists of the following steps (Figure 4):

Step 1: Apply Principal Component Analysis (PCA) to both the predictor fields (e.g., temperature and precipitation variables from global models) and the predictand field (e.g., observed temperature and precipitation). This process yields eigenvector matrices and their corresponding principal component (PC) coefficient matrices. The predictor PCs are denoted as PCtan, and the predictand PCs are denoted as PC^{tor}. Forecasting equations are then constructed to predict PC^{tan} as a function of PC^{tor}. Initially, these equations are formulated using only the first PCtor component, resulting in Equations 5 and 6:

$$PC_1^{\tan} = a_1 PC_1^{tor} + b$$

Equation 5

$$PC_2^{\tan} = a_2 PC_2^{tor} + b$$

Equation 6

Step 2: The newly constructed forecasting equations are applied to the dependent dataset to perform reforecasts, and the mean forecast error is computed to assess model accuracy.

Step 3: Steps 1 and 2 are iteratively repeated, each time constructing forecasting equations using an increasing number of principal components e.g., 2 *PCs*, 3 *PCs*, or 4 *PCs*. The forecast errors from each version of the model (based on 1, 2, 3, or more PCs) are compared. The model configuration that yields the smallest forecast error is selected, and the corresponding set of principal component coefficients is fixed for application to the independent (validation or real-time) dataset.

Step 4: The selected forecasted principal component coefficients PC_1^{tan} are used to recreate the original data field.

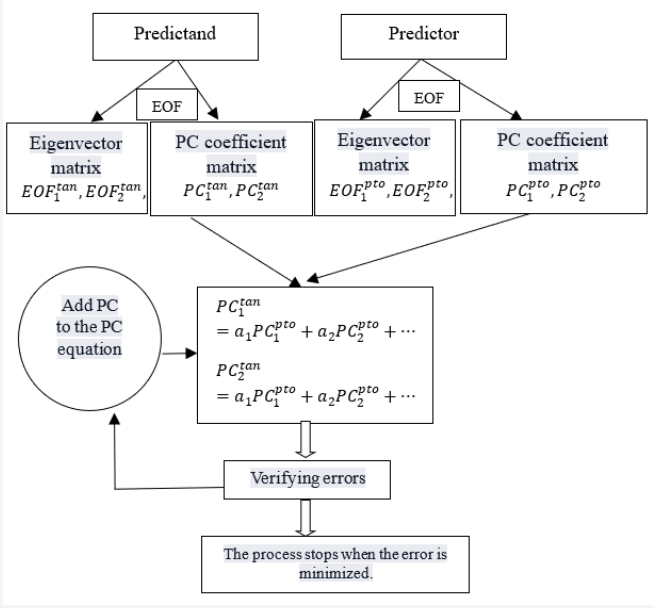


Figure 4: Forecast using CCA correlation diagram

2.3.3 Multivariate linear regression (MLR)
Multivariable linear regression involves establishing a linear relationship between the dependent variable Y and the independent variables X_i, as presented in Equation 7:

$$Y = a_1X_1 + a_2X_2 + \ldots a_nX_n + b \quad \text{Equation 7}$$

To enhance the spatial relationship between observational stations and model grid points, the following procedure is employed to construct station-specific regression models based on correlation analysis:

Step 1: Input data from observational stations and global model outputs are collected and preprocessed.

Step 2: For stations with incomplete time series data, missing values are imputed using a stochastic (random function) method to preserve temporal variability.

Step 3: For each observation station, correlation coefficients are calculated between the station's time series and the time series at every grid point of the model dataset.

Step 4: Grid points with correlation coefficients exceeding a predefined threshold are identified. These points are considered to have a statistically significant relationship with the station data.

Step 5: A regression equation is then constructed, linking the observational station data to the selected grid points with high correlation. This regression model is tailored to each individual station.

The same procedure is repeated for all monitoring stations to generate a set of localized regression models across the study region. In this study, three statistical methods Canonical Correlation Analysis (CCA), Principal Component Regression (PCR), and Multivariate Linear Regression (MLR) were employed to predict seasonal climate patterns in Vietnam. These methods provide a straightforward and interpretable means of identifying linear relationships between large-scale climate drivers (e.g., ENSO, sea surface temperature anomalies) and regional drought-related variables (e.g., rainfall, temperature, and number of dry days). Such models are particularly well-suited to Vietnam's climate system, where many critical climate drought interactions such as the influence of ENSO on summer precipitation or the impact of Western Pacific Sea surface temperature anomalies on drought occurrences in Central Vietnam tend to exhibit predominantly linear or near-linear characteristics.

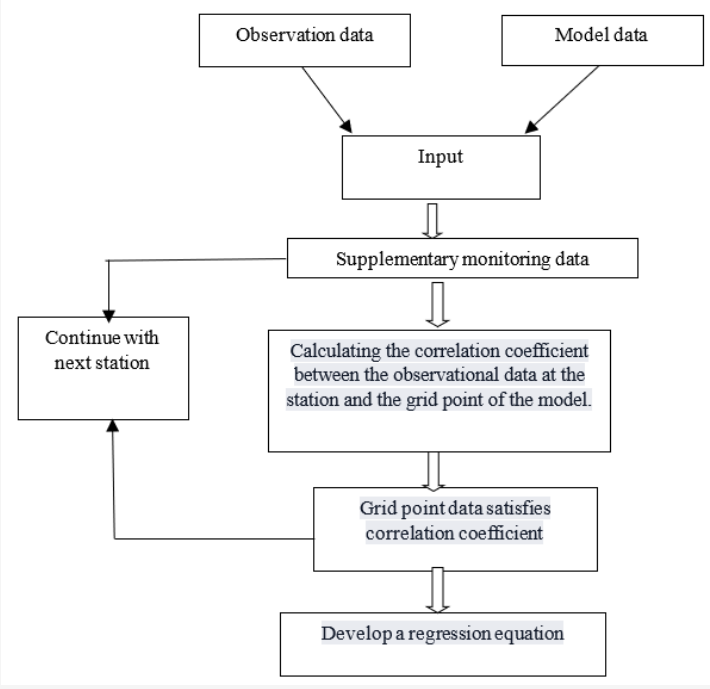


Figure 5: Diagram of constructing a multivariate linear regression equation based on the correlation coefficient

A key advantage of these statistical approaches lies in their relatively low data requirements and computational simplicity. For example, CCA and PCR have been shown to perform effectively even with limited historical datasets, such as the 30–50 years of climate records typically available in Vietnam. This feature is essential for mitigating the risk of overfitting, particularly in data-constrained settings. In contrast, more advanced machine learning techniques including artificial neural networks and deep learning methods often require large volumes of training data and involve complex hyperparameter tuning processes, which may not be feasible given current data availability and computational infrastructure in Vietnam.

Moreover, CCA offers a robust framework for integrating large-scale climate signals into regional forecasts by identifying spatial correlation patterns between global climate indices and local meteorological variables. PCR further enhances predictive capacity through dimensionality reduction, retaining key components that account for the majority of variance in target drought indicators. Finally, MLR models are especially practical for operational deployment. These models can be readily implemented within the existing infrastructure of national agencies such as the National Center for Hydro-Meteorological Forecasting, without requiring significant computational resources or specialized technical expertise.

2.4 Quantile Mapping (QM)

Quantile mapping (QM) [46] is used to transform modeled data into probability distributions similar to observations, as presented in Equation 8:

$$Q_m(t)=F_0^{-1}[F_s[Q_s(t)]$$

Equation 8

Here, $Q_m(t)$, $Q_s(t)$ represent the pre-adjustment and post-adjustment data of the model, respectively, while Fs is the cumulative distribution function of the model's forecasted data, and F_0^{-1} is the inverse of the cumulative distribution function of the observed data.

To forecast and evaluate extreme phenomena over a one-year period, the research team follows the following framework. Based on dependent and independent data series from 1983-2022, the computational program assesses the error.

Errors used for evaluation include: Two-phase forecast (Accu2P): events are divided into two phases: below normal and above normal. Out of the total number of two-phase events, *Accu2P* represents the number of correctly forecasted events.

In a three-phase forecast (*Accu3P*), events are divided into three phases: below normal, above normal, and near normal. *Accu3P* is defined as the percentage of three-phase events that are correctly forecasted. After calibration, *AccuQ2p* and *AccuQ3p* represent the accuracy of 2-phase and 3-phase. Ranked Probability Skill Score (RPSS) [47] The RPS index combines three forecast phases with three observation phases. The observation phase is counted at 100% in the actual occurrence phase, but 0% in the other two phases. The formula for calculating this index is presented in Equations 9 and 10:

$$RPS=\sum_{cat=1}^{cat=N}\left(Pcum_{F(cat)}-Pcum_{0(cat)}\right)^2$$

Equation 9

$$RPSS=1-\frac{RPS_{forecst}}{RPS_{reference}}$$

Equation 10

In view of this formulation, and considering that an $RPS_{forecast}$ close to zero represents better forecasts, an *RPSS* value of closer to 1 represents greater skill in the forecast system (compared to the reference, usually climatology). As the area under the ROC curve [48], the ROC skill score is calculated. A seasonal climate prediction with the skill equivalent to climatology will correspond to a diagonal curve with a ROC area of 0.5. Therefore, the ROC score should be higher than 0.5 for a skillful prediction.

3. Results and Discussion

Numerous studies conducted in Vietnam and across Southeast Asia have utilized dynamical models such as CFSv2, ECMWF SEAS5, and ACCESS-S for seasonal climate forecasting. However, these models often exhibit limitations related to coarse spatial resolution, systematic biases, and limited skill in capturing local and regional climate variability. For example, [47] assessed the performance of the CFSv2 model in forecasting seasonal precipitation in Vietnam and reported low predictive skill beyond two-month lead times. Similarly, [48] observed that while SEAS5 and NCEP models were able to capture broad warming trends across Southeast Asia, their forecast reliability remained low to moderate for inland regions of Indochina. Given these constraints, recent efforts by national and regional institutions have emphasized the need to integrate advanced statistical and machine learning techniques into forecasting systems to improve regional applicability and accuracy. For instance, [49] has advocated for the adoption of hybrid approaches that combine dynamical and statistical models.

Furthermore, initiatives such as [50] have demonstrated the practical utility of statistical forecasting methods in supporting agricultural planning and disaster risk reduction efforts across Vietnam and neighboring countries.

In response to these limitations, the present study first conducts a comparative evaluation of forecasting skill before applying the three statistical methods Canonical Correlation Analysis (CCA), Principal Component Regression (PCR), and Multivariate Linear Regression (MLR) to forecast extreme weather events. Specifically, the study assesses the Accu2P accuracy of six global climate models relative to the post-processed outputs derived from the three statistical methods. This evaluation is performed for both 3-month and 6-month lead times using a consistent historical dataset spanning 1983 to 2013. Each climate model's performance is analyzed individually. This comparative step is essential for quantifying the added value of statistical post-processing in improving the raw outputs of dynamical models, particularly in the context of Vietnam's complex and regionally diverse climate system.

3.1 Comparative Evaluation of Forecast Skill (Accu2p) between Raw Climate Models and Statistical Post-Processing Methods

Prior to applying the three statistical methods Canonical Correlation Analysis (CCA), Principal Component Regression (PCR), and Multivariate Linear Regression (MLR) for forecasting extreme weather events, the research team conducted a comparative evaluation of the Accu2P accuracy of six global climate models relative to the post-processed outputs produced by the statistical methods. This assessment was carried out for both 3-month and 6-month lead times, using a consistent historical dataset spanning the period from 1983 to 2013, with each of the six models evaluated individually. Based on the evaluation of the performance of climate models using 3-month forecasts from 1983 to 2013 across seven climate regions of Vietnam, with forecast initiation in February, May, August, and November, the GFDL-SPEAR model is identified as the most stable and high-performing model, as shown in Figure 6(a)-6(f).

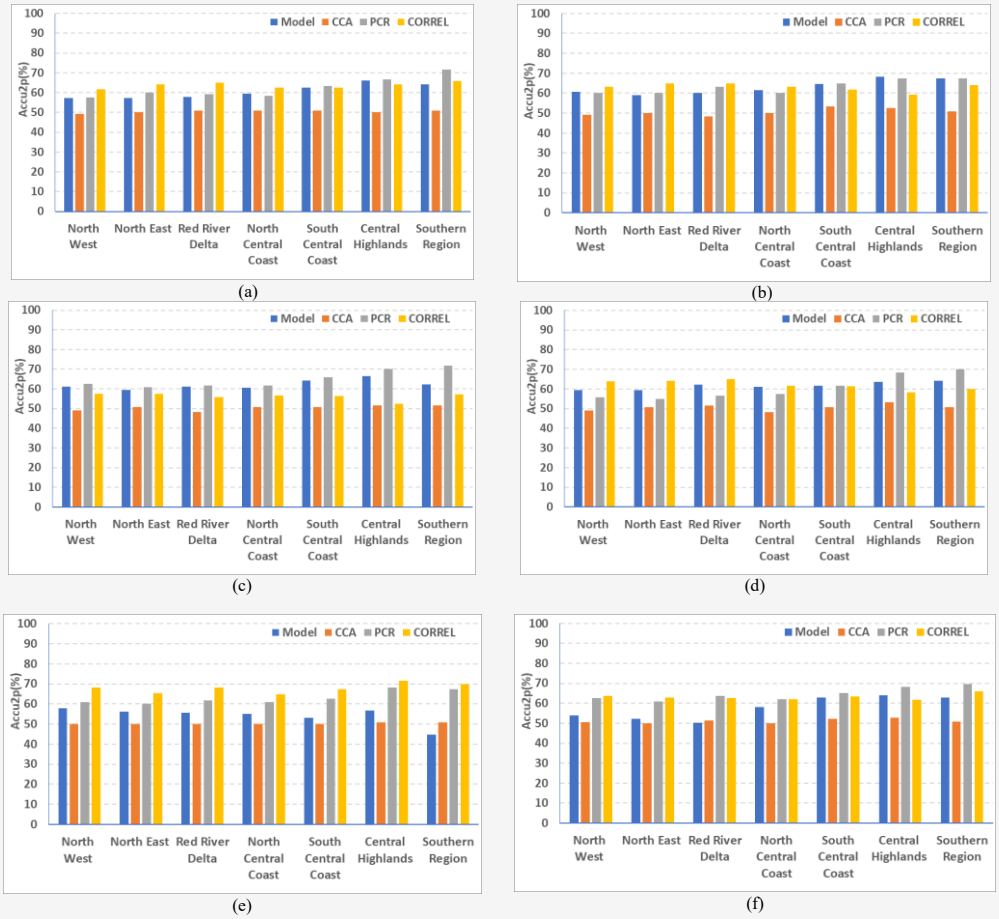


Figure 6: Comparison of 3-month Accu2p skill across models and post-processing methods in seven regions of Vietnam: (a) Cancm4i, (b) Cansips, (c) Cfsv2, (d) CCSM4, (e) Gfdlspear, (f) Nasageos2s

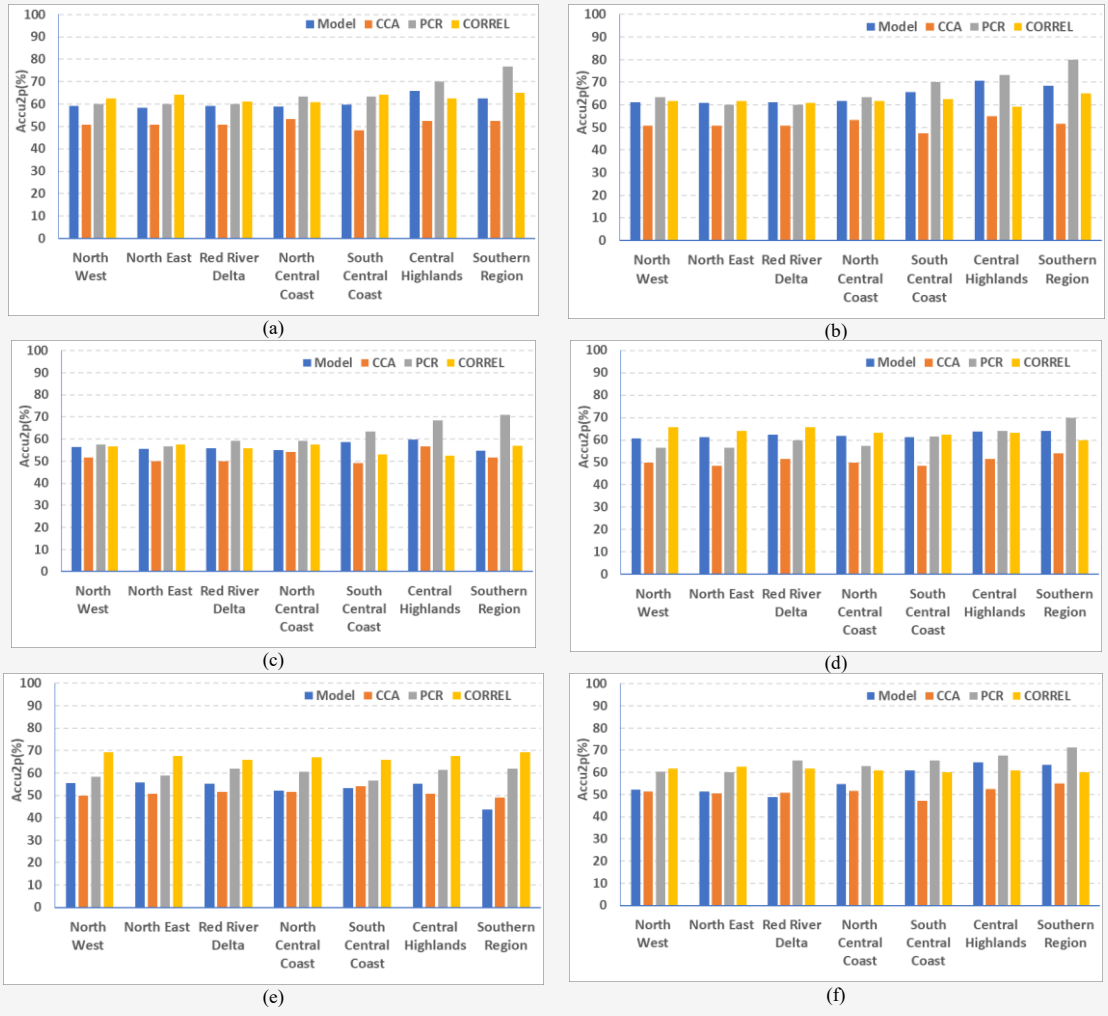


Figure 7: Comparison of 6-month Accu2p skill across models and post-processing methods in seven regions of Vietnam: (a) Cancm4i, (b) Cansips, (c) Cfsv2, (d) CCSM4, (e) Gfdlspear, (f) Nasageos2s

It shows consistently high correlation values across regions, especially after statistical bias correction methods such as PCR and CORREL, with correlations reaching around 70–80%. The CanCM4i model also performs well, particularly in central and southern regions, and shows significant improvement after correction. In contrast, while the CFSv2 model performs reasonably well, it does not outperform the other two models, especially in regions like the Northwest and Southern Vietnam. Among the correction methods, Principal Component Regression (PCR) proves to be the most effective, substantially enhancing the forecast accuracy for most models. Overall, all bias correction methods improve the original model performance, with the combination of GFDLSpear and PCR delivering the best results.

Figure 7(a)-7(f) illustrates the performance of six global climate models in producing 6-month lead forecasts across seven climatic regions of Vietnam during the period 1983–2013. Among the models evaluated, GFDL-SPEAR demonstrated the most consistent and highest forecasting skill, particularly after the application of statistical bias correction methods such as Principal Component Regression (PCR) and CORREL, with correlation values ranging from 65% to 85%. The CanCM4i model also exhibited substantial improvement following correction, especially in the central and southern regions. By contrast, although the CFSv2 model maintained relatively stable performance, it generally underperformed compared to GFDL-SPEAR and CanCM4i in most regions. Among the correction techniques assessed, PCR consistently yielded the greatest enhancement in forecast accuracy. Overall, statistical post-processing improved the performance of all models, with the combination of GFDL-SPEAR and PCR providing the most reliable results for long-lead seasonal forecasting in Vietnam. As part of Section 3, the research team presents the

results of the error assessment associated with forecasting extreme weather phenomena specifically cold spells, heatwaves, and moderate to heavy rainfall across seven climatic regions of Vietnam. This analysis is based on two datasets: dependent (training) and independent (validation) time series. Cold spells and heatwaves are evaluated using two duration-based categories: short-duration events lasting 2 to 4 days, and long-duration events lasting more than 5 days. In contrast, moderate to heavy rainfall events are assessed only for the 2–4 day category, as rainfall exceeding five consecutive days is relatively rare in certain regions of Vietnam and thus not statistically significant for this analysis.

3.2 Evaluation of the Cold Spells Forecast Results

3.2.1 Assessment of the forecast results for the cold spell lasting 2-4 days

During the core winter months in Vietnam typically from December to February and occasionally extending into March harsh cold weather conditions are most frequently observed. These severe cold spells are primarily driven by the southward extension of the Siberian High and the intensification of the East Asian trough at upper atmospheric levels. The combination of these large-scale circulation patterns contributes to the development of intense and prolonged cold air outbreaks. Notably, the most severe cold spell ever recorded in Vietnam occurred in 2008, persisting for an unprecedented 38 to 40 days across the northern provinces. The objective of this section is to evaluate the forecasting performance for severe cold spells during the winter periods of November–January, December–February, and January–March, with forecast lead times ranging from three to six months. Among the forecasting methods evaluated, three commonly used approaches exhibit Accu3P values ranging from 0.3 to 0.4, whereas the Principal Component Regression (PCR) method achieves higher Accu3P values of 0.5 for phases 10–11121 and 7–11121, indicating superior performance. In general, the post-calibration accuracy metric AccuQ3P shows an improvement over Accu3P, increasing by approximately 0.1 to 0.2 across most forecast lead times. As illustrated in Figure 8(c)-8(d), the Accu2P for the three common methods reaches 0.6 in certain cases. In contrast, AccuQ2P exhibits minimal variation after statistical adjustment.

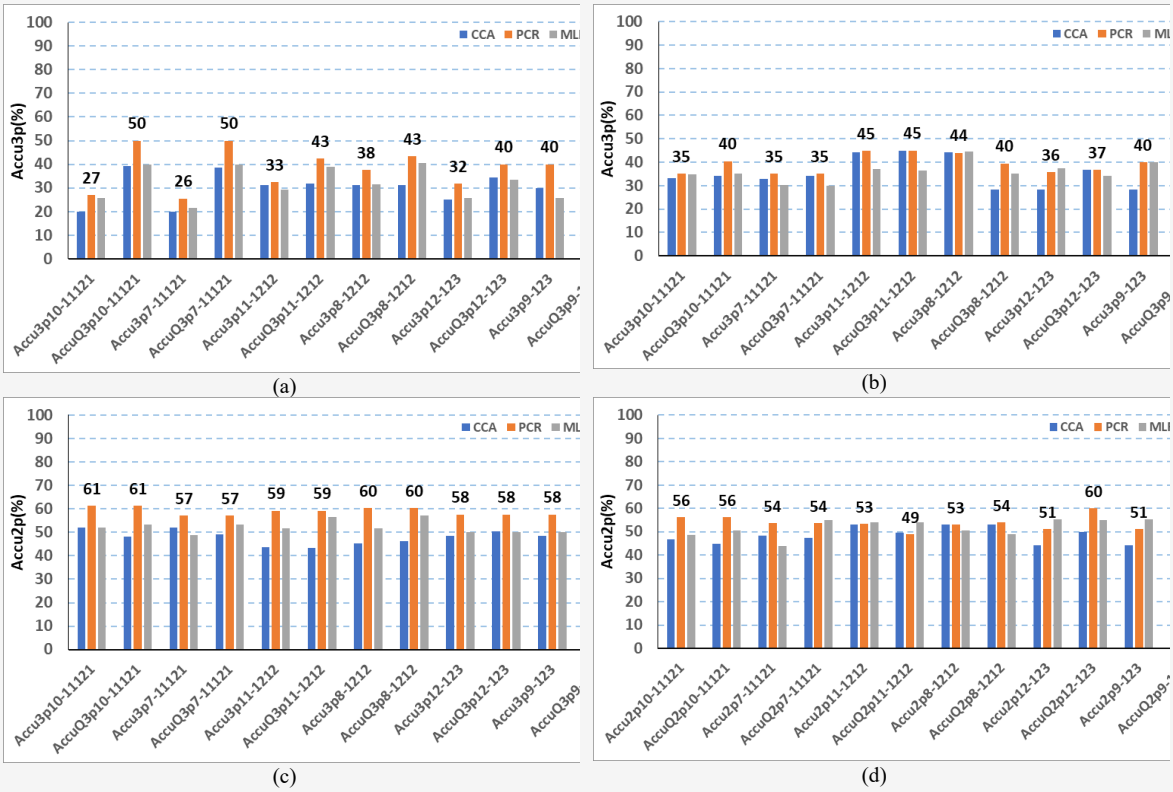


Figure 8: Accu3p and Accu2p of dependent and independent data series of the cold spells from 2-4 days in 6 months forecast results: (a) independent Accu3p, (b) dependent Accu3p, (c) independent Accu2p, (b) dependent Accu2p

Based on Figures 8(a)-8(d), Accu3P typically falls within the 0.3 to 0.4 range, while Accu2P generally ranges from 0.5 to 0.6. Across most forecasting scenarios, the PCR method consistently delivers the highest accuracy, reinforcing its effectiveness for long-lead seasonal forecasting in Vietnam. Based on the frequency of models achieving the highest Accu3P and Accu2P values over the 6-month forecast horizon (Figures 9(a)–9(d)), the GFDL-SPEAR model-when combined with CCA and PCR methods-exhibits the most consistent performance, attaining top accuracy in approximately 25% to 42% of the cases. For the CanCM4i and NASA-GEOS models, the MLR method yields strong results, with top accuracy frequencies ranging from 20% to 40%. In contrast, other models show limited effectiveness, with frequencies of highest accuracy generally falling below 20%. For each station point, the RPSS index is calculated and then the positive RPSS ratio (%) is calculated. As shown in Figure 10(a) and Figure 10(b), the Principal Component Regression (PCR) method yields the highest Ranked Probability Skill Score (RPSS) among all methods evaluated, with a positive RPSS ratio of approximately 70% for the months of September and December.

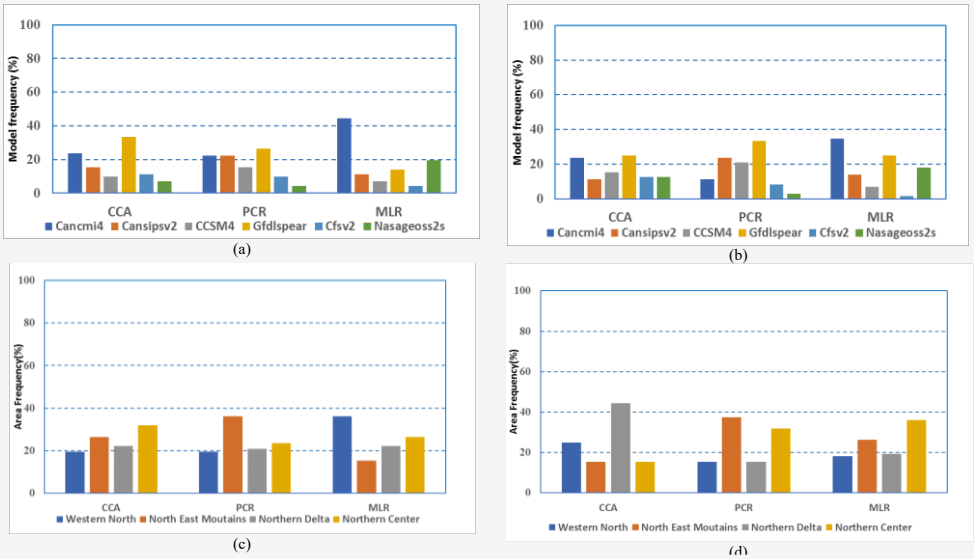


Figure 9: The frequency of appearance of the model and area with the highest Accu3p(a) and (c) and Accu2p (b) and (d) for 2-4day cold spells in 6 months forecast results

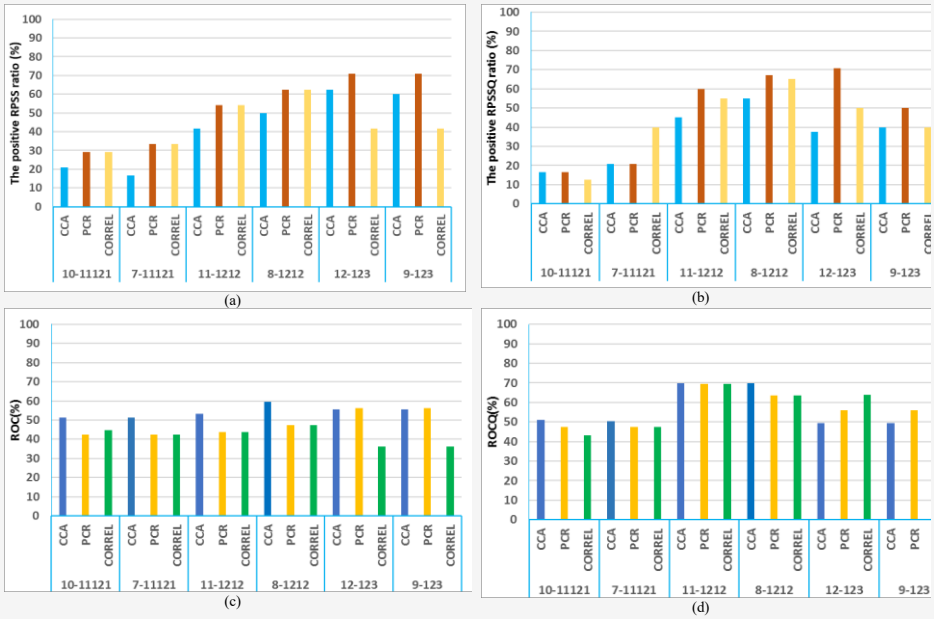


Figure 10: The positive RPSS ratio and ROC for 2-4day cold spells in 6 months forecast results: (a)The positive RPSS ratio, (b) The positive RPSSQ ratio, (c) ROC, (d) ROCQ

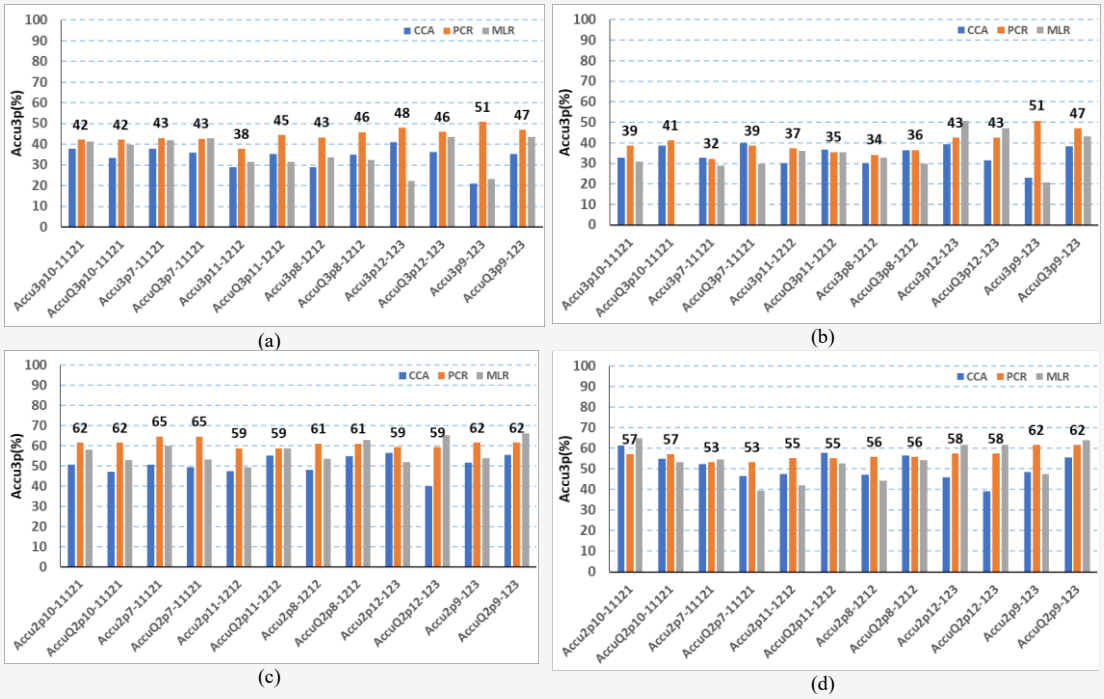


Figure 11: Accu3p and Accu2p for independent and dependent data series of cold spell forecasts over 5-day periods in a 6-month lead time: (a) Accu3p for independent data; (b) Accu3p for dependent data; (c) Accu2p for independent data; and (d) Accu2p for dependent data

Both the PCR and CORREL methods exhibit slight improvements in forecast performance following statistical adjustments. Furthermore, post-adjustment results for the Receiver Operating Characteristic (ROC) index presented in Figure 10(c) and Figure 10(d) show a notable enhancement, with values increasing from the 0.4–0.6 range to approximately 0.5–0.7, indicating improved discrimination ability of the forecast models.

3.2.2 Assessment of the forecast results for the cold spell lasting 5 days

As illustrated in Figures 11(a)-(b), both the Multivariate Linear Regression (MLR) and Principal Component Regression (PCR) methods demonstrate higher Accu3P values compared to the Canonical Correlation Analysis (CCA) method, with values ranging between 0.4 and 0.5. According to Figures 11(c)–11(d), the Accu2P metric attains its highest performance with the PCR method, yielding values in the range of 0.62–0.65 on the dependent dataset, whereas performance on the independent dataset slightly decreases to 0.53–0.58. Notably, both datasets exhibit the highest Accu2P values in September, with PCR reaching 0.62, indicating its superior performance during that forecast cycle. As illustrated in Figures 12(a)–12(b), the frequency of models achieving the highest Accu3P/Accu2P scores in 6-month forecasts reveals that the GFDL-SPEAR model consistently exhibits strong performance across all methods, with top accuracy rates ranging from 25% to 42%. Notably, the model attains particularly high Accu2P values when combined with the MLR method, reaching up to 58%. In contrast, the remaining models exhibit lower performance, with top accuracy frequencies generally falling below 20%. Figures 12(c)-12(d) present the frequency of regions attaining the highest Accu3P/Accu2P in 6-month forecasts. Results show that four regions consistently reach accuracy levels of 20–25% across all three methods. Notably, the PCR method achieves exceptional performance in the Northeast Mountainous region, with top accuracy frequencies exceeding 50%. Figures 13(a)-13(b) illustrate that the Ranked Probability Skill Score (RPSS) fluctuates between 50% and 60% across evaluated phases. For phases 10–11121 and 7–11121, the PCR method yields the highest positive RPSS ratios. Additionally, the Receiver Operating Characteristic (ROC) index generally ranges from 0.4 to 0.6, with PCR again achieving the highest values. Phases 11–1212 and 8–1212 show modest improvements in ROC index after statistical adjustment (Figure 13(c)-13(d)). The PCR method provided the best forecasting results regarding the number of cold spells. Cancmi4 and Gfdlspear performed better than the other models.

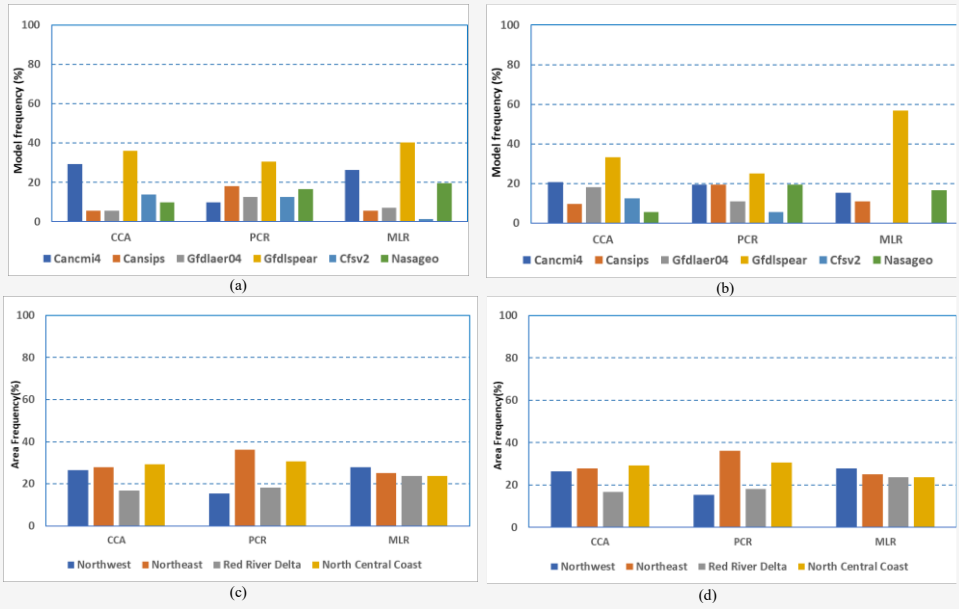


Figure 12: The frequency of appearance of the model and area with the highest Accu3p (a) and (c) and Accu2p (b) and (d) for 5day cold spells in 6 months forecast results

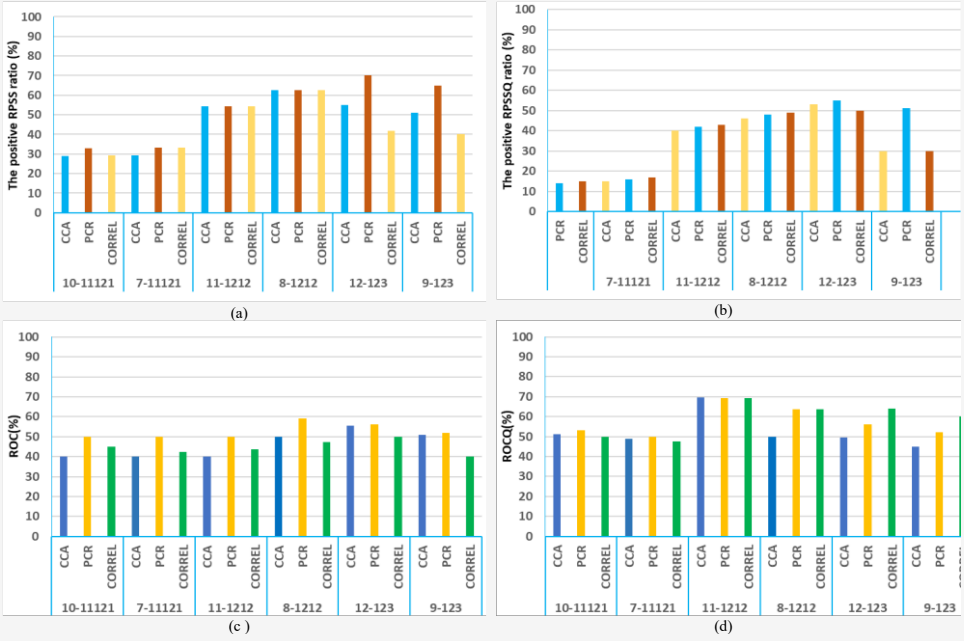


Figure 13: The positive of RPSS ratio and ROC for 5 day cold spells in 6 months forecast results: (a)The positive RPSS ratio, (b) The positive RPSSQ ratio, (c) ROC, (d) ROCQ

3.3 Evaluation of the Heat Waves Forecast Results

3.3.1 Assessment of the forecast results for the heat waves lasting 2–4 days

In Vietnam, hot weather typically begins in March, first affecting the Northwestern and Southern regions, and progressively intensifies throughout the year. From May to August, heatwaves are most frequently observed in the Northern and Central regions. Notably, record-breaking temperatures often occur during the early months of the hot season. For instance, in April 2024, 110 out of 186 weather stations recorded temperatures exceeding historical maxima. In this study, with a forecast lead time of three to six months, an assessment is conducted to evaluate the predicted number of heatwave events for the months of February, March, and April. Heatwaves are categorized based on their duration into two groups: short-duration events (lasting 2–4

days) and long-duration events (exceeding 5 days). As illustrated in Figures 14(a)-14(b), the Accu3P values for the dependent and independent data series range from 0.4 to 0.5 and 0.3 to 0.4, respectively. During the forecast months of March and April, the MLR method achieves an Accu3P value of 0.5, indicating relatively strong performance in those periods. According to Figures 14(c)-14(d), Accu2P values generally fall between 0.5 and 0.6 across models and lead times. However, following post-processing adjustments, both AccuQ3P and AccuQ2P show a slight decline in accuracy compared to their unadjusted counterparts, Accu3P and Accu2P, suggesting potential overfitting or calibration limitations in some scenarios. According to Figures 15(a)-15(b), the models with the highest frequencies of achieving top Accu3P and Accu2P scores are as follows: for the CCA method, the GFDL-SPEAR model stands out as the most frequently successful; for the PCR method, both CanSIPSv2 and GFDL-SPEAR exhibit superior performance; and for the MLR method, the CanCM4i and NASA-GEOS S2S models show the highest accuracy frequencies. As shown in Figures 15(c)-15(d), the regions with the most frequent occurrences of top Accu3P and Accu2P values vary by method. The Central Highlands region shows the highest performance under the CCA method; the South Central region is most frequently associated with the highest scores under the PCR method; and both the South Central and Central Highlands regions perform best when using the MLR method. As shown in Figures 16(a)-16(b), the average ROC index across all methods is approximately 0.5, indicating moderate forecast skill.

However, during the March forecast period, all three methods CCA, PCR, and MLR achieved elevated ROC values of 0.6, suggesting improved discrimination performance during this specific lead time. Notably, the ROC indices for the MLR method increased after calibration, whereas those for CCA and PCR exhibited a slight decline post-calibration. In terms of the Ranked Probability Skill Score (RPSS), the PCR method demonstrated strong performance during the March forecast, with a positive RPSS ratio of approximately 90%. However, following calibration, this ratio declined slightly, indicating a marginal reduction in probabilistic forecast skill despite the application of post-processing adjustments (Figures 16(c) - 16(d)).

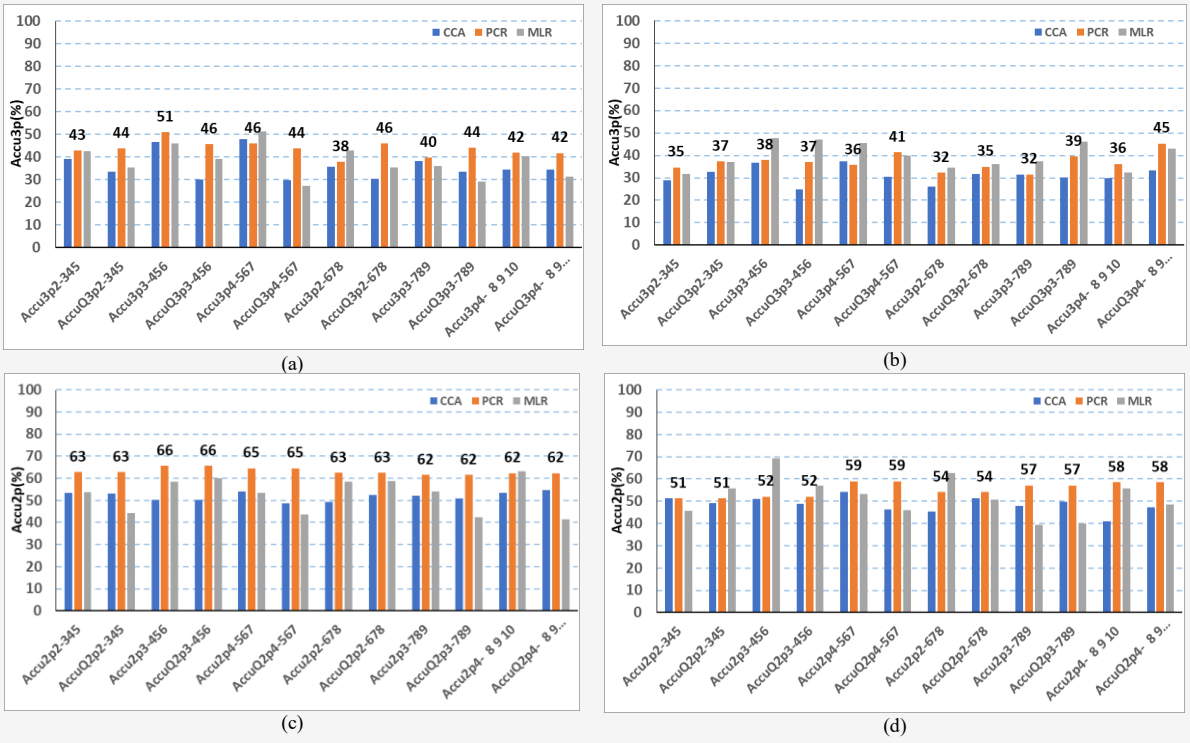


Figure 14: Accu3p and Accu2p of dependent and independent data series of the heat wave from 2-4 days in 6 months forecast results: (a) independent Accu3p, (b) dependent Accu3p, (c) independent Accu2p, (b) dependent Accu2p

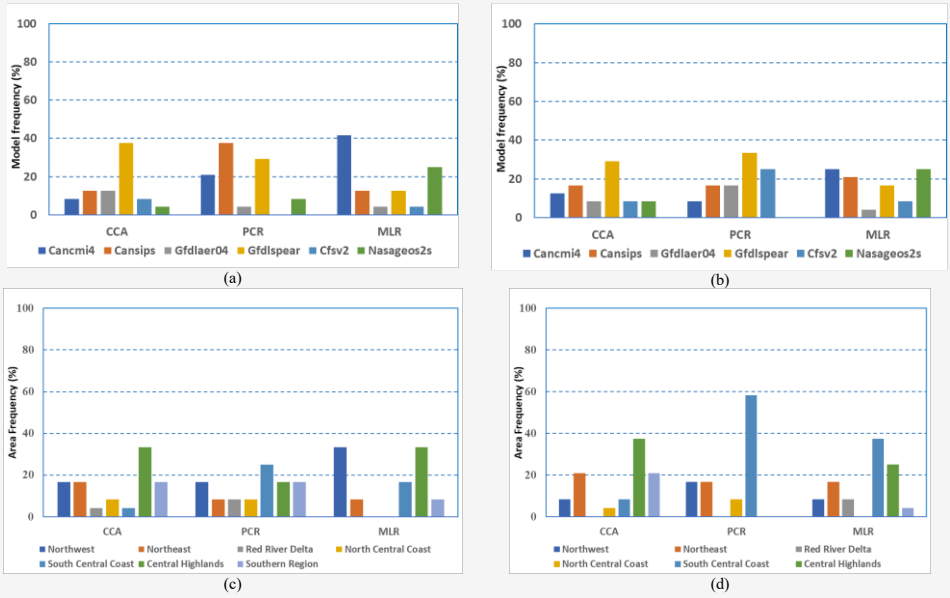


Figure 15: The frequency of appearance of the model and area with the highest Accu3p (a) and (c) and Accu2p (b) and (d) for 2-4day heat waves in 6 months forecast results

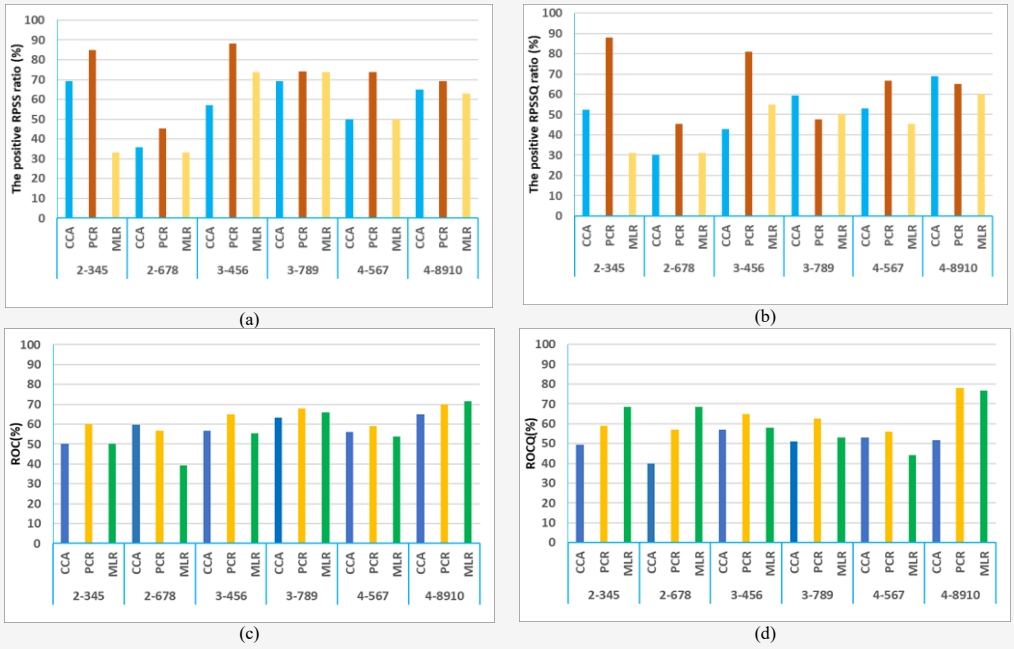


Figure 16: The positive of RPSS ratio and ROC for 2-4 days heat waves in 6 months forecast results: (a)The positive RPSS ratio, (b) The positive RPSSQ ratio, (c) ROC, (d) ROCQ

3.3.2 Assessment of the forecast results for the heat waves lasting 5 days

Heat waves lasting longer than 5 days are quite common in Vietnam during the summer, severely affecting daily life and economic activities. Figure 17 illustrates the forecast accuracy (Accu3p and Accu2p) of heat waves beyond 5 days using independent and dependent data. Overall, the PCR method consistently outperforms CCA and MLR across most stations in both independent (Figures 17(a), 17(c)) and dependent datasets (Figures 17(b), 17(d)). Notably, dependent data generally yield higher accuracy than independent data, highlighting the influence of overfitting or data consistency in model calibration. Among the metrics, Accu2p tends to show more stable performance across methods and stations compared to Accu3p. As illustrated in Figure 18, the frequency of appearance of the most accurate model and region varies across methods and forecast metrics.

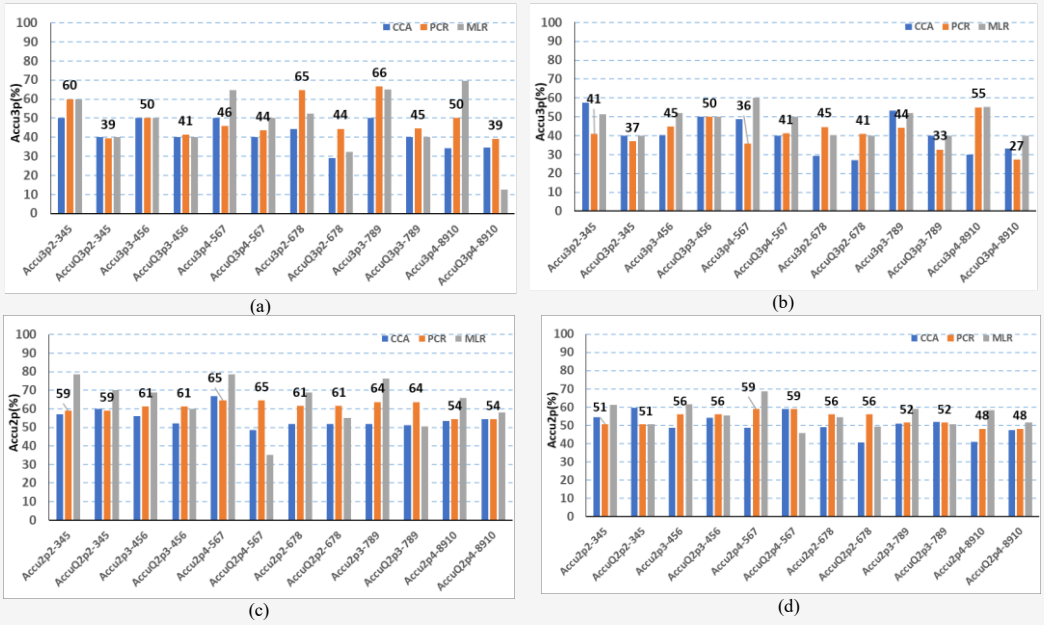


Figure 17: Accu3p and Accu2p of dependent and independent data series of the heat wave beyond 5 days in 6 months forecast results:
(a) independent Accu3p, (b) dependent Accu3p, (c) independent Accu2p, (d) dependent Accu2p

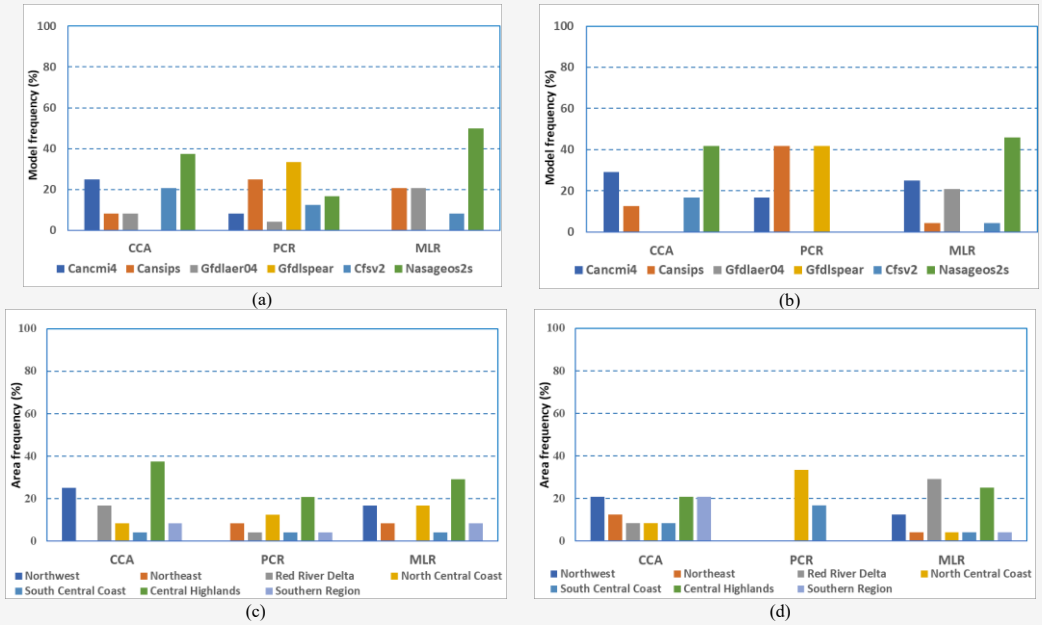


Figure 18: The frequency of appearance of the model and area with the highest Accu3p (a) and (c) and Accu2p (b) and (d) beyond 5 days heat waves in 6 months forecast results

Specifically, Figures 18(a)-18(b) show the distribution of highest Accu3p values across models and regions for independent and dependent data, respectively, while Figures 18(c) and 18(d) present the corresponding results for Accu2p. The PCR method consistently dominates in both independent and dependent forecasts, especially in the North Central Coast and Red River Delta.

In contrast, MLR exhibits localized strengths in the Central Highlands, and CCA contributes less significantly across most regions. These patterns reflect both regional and methodological differences in forecast performance beyond 5 days. According to Figures 19(c) and 19(d), the ROC index generally ranges from 0.6 to 0.7 across most forecast periods.

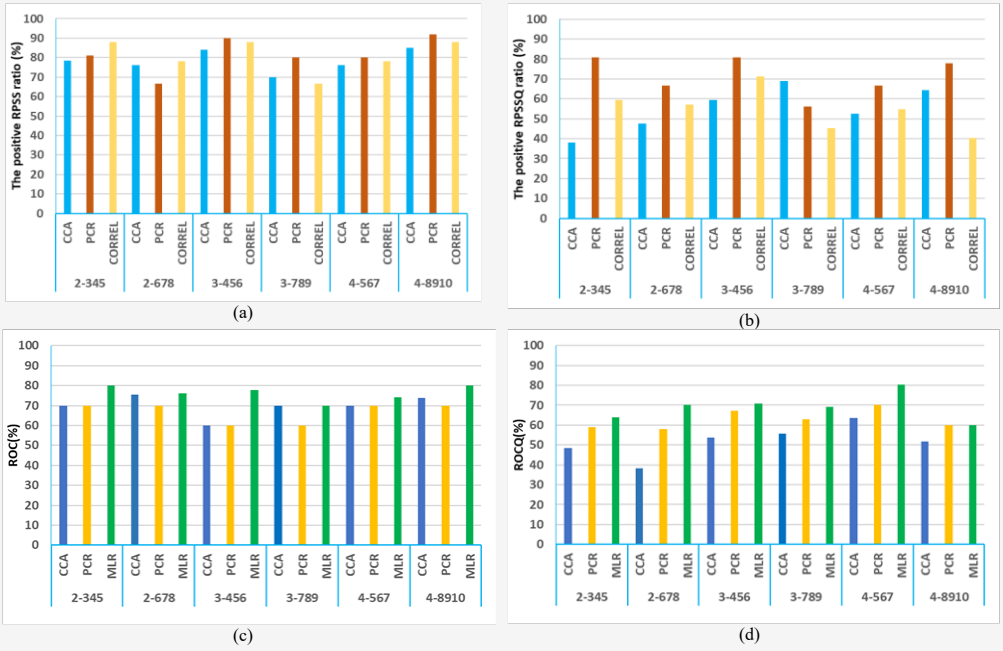


Figure 19: RPSS and ROC for 5 days heat waves in 6 months forecast results: (a)The positive RPSS ratio, (b) The positive RPSSQ ratio, (c) ROC, (d) ROCQ

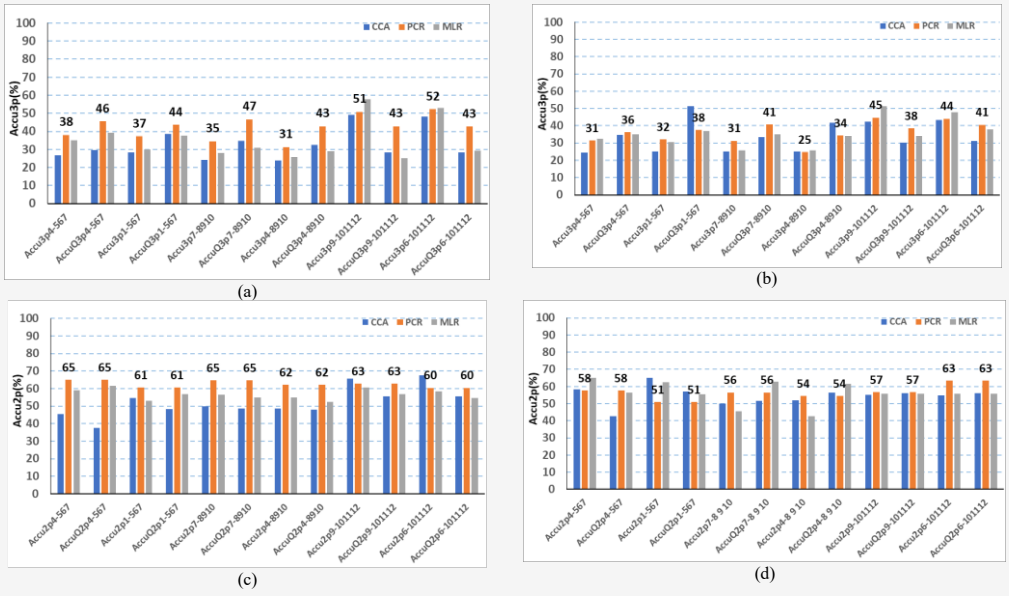


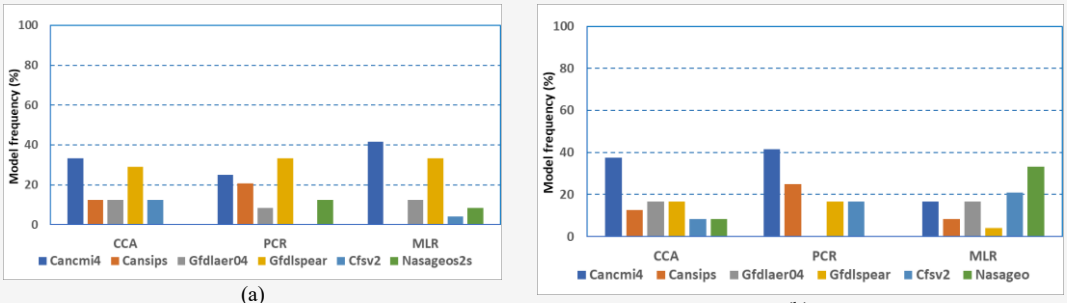
Figure 20: The accuracy of 3 phase (a) and (b), 2 phase (c) and (d) of dependent and independent data series forecasts heavy rainfall for 6 months

Notably, at forecast points in February and April, the MLR method attains an ROC index of 0.8, indicating strong discriminative capability (Figures 19(c)–19(d)). In addition, the positive RPSS ratio generally ranges between 60% and 90% (Figures 19(a)–19(b)). Nevertheless, after calibration, both the RPSSQ ratio and ROCQ index exhibit a slight reduction compared to the original RPSS and ROC values. Among the methods evaluated, MLR consistently exhibits the highest RPSS and ROC scores, highlighting its superior performance in forecasting heatwave events. In comparison with CCA, both PCR and MLR deliver more accurate forecasting results. Regionally, the South Central Coast and Central Highlands of Vietnam demonstrate the highest forecast ac curacy for heatwaves across all methods.

3.4 Evaluation of the Heavy Rain Spells Forecast

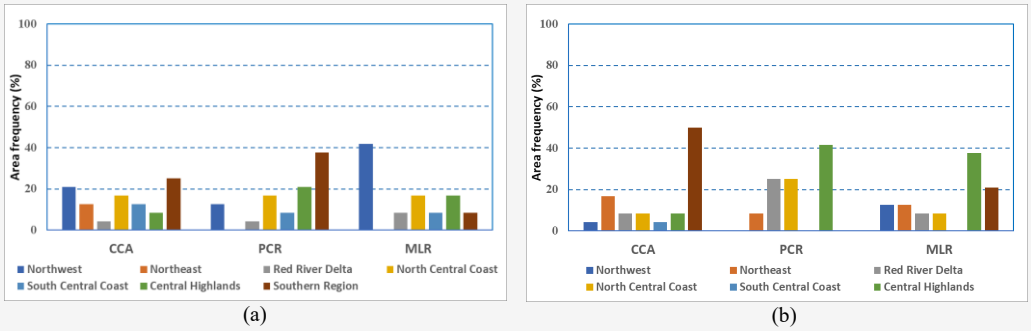
Due to Vietnam's diverse topography and climatic variation, the rainy season extends from May to December. In the northern provinces, significant rainfall typically occurs from May to October. The Central region experiences its peak rainy season between September and November, while the Central Highlands and Southern regions receive the majority of their rainfall from May until mid-December. A sliding seasonal forecast was conducted for the periods May–July (MJJ), August–October (ASO), and October–December (OND), with forecast lead times of three and six months, focusing on moderate to heavy rainfall events. According to Figure 20(a)-20(b), the Principal Component Regression (PCR) method performed best for forecasting heavy rainfall during the MJJ and ASO seasons, achieving Accu3P values of 0.46 and 0.47 with model initializations in April and July, respectively. In contrast, the Multivariate Linear Regression (MLR) method showed the best performance for the OND season, reaching an Accu3P of 0.52 with a June initialization. Additionally, MLR remained effective with a later initialization in September, achieving an Accu3P of 0.45. For the Accu2P dependent dataset (Figure 20(c)), the PCR method again proved most effective in the MJJ and ASO seasons, with accuracy values ranging from 0.61 to 0.65. In the OND season, the CCA method yielded the highest two-phase forecast accuracy, with Accu2P values between 0.60 and 0.63. The Accu2P results for the independent dataset (Figure 20(d)) showed a similar trend, though the accuracy of the PCR method decreased slightly, with values ranging from 0.54 to 0.61. Figure 21(a) illustrates the frequency with which each of the six global climate models achieved the highest forecast accuracy for three-phase heavy rainfall events at a six-month lead time. The CanCM4i model consistently outperforms others, with its highest accuracy frequency ranging from 33% to 42% under the CCA and MLR methods. The GFDL-SPEAR model also demonstrates strong performance, particularly when used with the PCR method, reaching a maximum frequency of 33%.

Figure 21(b) presents the corresponding results for two-phase forecasts of heavy rainfall events at the same six-month lead time. Once again, CanCM4i emerges as the most consistently accurate model under the CCA and PCR methods, with frequency values ranging from 38% to 42%. In contrast, the NASA GEOS-S2S model shows the best performance under the MLR method, achieving a top forecast accuracy frequency of 33%. Figure 22(a) illustrates the frequency with which each of Vietnam's seven climatic regions achieved the highest forecast accuracy for three-phase heavy rainfall events at a six-month lead time.



(a) (b)

Figure 21: Frequency of occurrence of the model with the highest accuracy in three-phase: (a) and two-phase (b) forecasts of 6-month heavy rainfall



(a) (b)

Figure 22: Frequency of occurrence of the region with the highest forecast accuracy in three-phase: (a) and two-phase (b) forecasts of heavy rainfall events at a six-month lead time

Under the CCA and PCR methods, the Southern region demonstrated the most consistent performance, with frequencies ranging from 25% to 38%. In contrast, when using the MLR method, the Western Northern region recorded the highest frequency of top accuracy, reaching 42%. Figure 22(b) presents the corresponding results for two-phase forecasts. According to the CCA method, the Southern region again achieved the highest frequency, peaking at 50%. Meanwhile, under the PCR and MLR methods, the Central Highlands emerged as the most accurate forecast region, with frequencies ranging from 38% to 42%.

As shown in Figures 23(a)-23(b), the ROC index for forecasting heavy rainfall spells generally fluctuates between 0.55 and 0.60, indicating moderate forecast skill. Notably, during the October–December (OND) season, with initializations in September and June, the ROC index increases slightly, reaching up to 0.68. When the Quantile Mapping (QM) method is applied, the ROC index further improves, ranging from 0.70 to 0.73 during the same OND season forecast periods. As illustrated in Figures 23(c)–23(d), the positive Ranked Probability Skill Score (RPSS) ratios vary between 52% and 66% across the baseline methods Under the MLR approach, the positive RPSS ratio increases significantly during the OND season with model initializations in June and September, reaching values between 84% and 86%, indicating a substantial gain in probabilistic forecast accuracy. Among the evaluated methods, Principal Component Regression (PCR) demonstrates the highest accuracy in forecasting heavy rainfall spells for both three-phase and two-phase classification schemes. Furthermore, when using either CCA or PCR, the CanCM4i model consistently emerges as the most effective among the six tested climate models.

3.5 Several Applications of Research Findings Assess Extreme Phenomena in Vietnam

This study integrates research on the forecasting and evaluation of extreme seasonal drought events into a comprehensive operational forecasting system. Through this system, researchers and forecasters can produce statistical forecast products and seamlessly integrate them with outputs from numerical forecasting models. The system also provides reanalysis maps based on JRA-3Q data, covering the period from 1981 to the present, enabling continuous monitoring of temperature and precipitation trends on monthly, seasonal, and annual timescales. Statistical forecast products are developed using three methods: Canonical Correlation Analysis (CCA), Principal Component Regression (PCR), and Multivariate Linear Regression (MLR). These methods focus on predicting average, extreme, and severe drought-related factors. The system produces three-phase probability forecast maps for these variables, as well as two-phase and three-phase probability evaluation maps. In addition, the system provides forecast evaluation metrics including Mean Error (ME), Correlation Coefficient (Corr), Ranked Probability Skill Score (RPSS), and Receiver Operating Characteristic (ROC) for seven climatic regions across Vietnam. The visual interface and features of the system are illustrated in Figures 24 and 25.

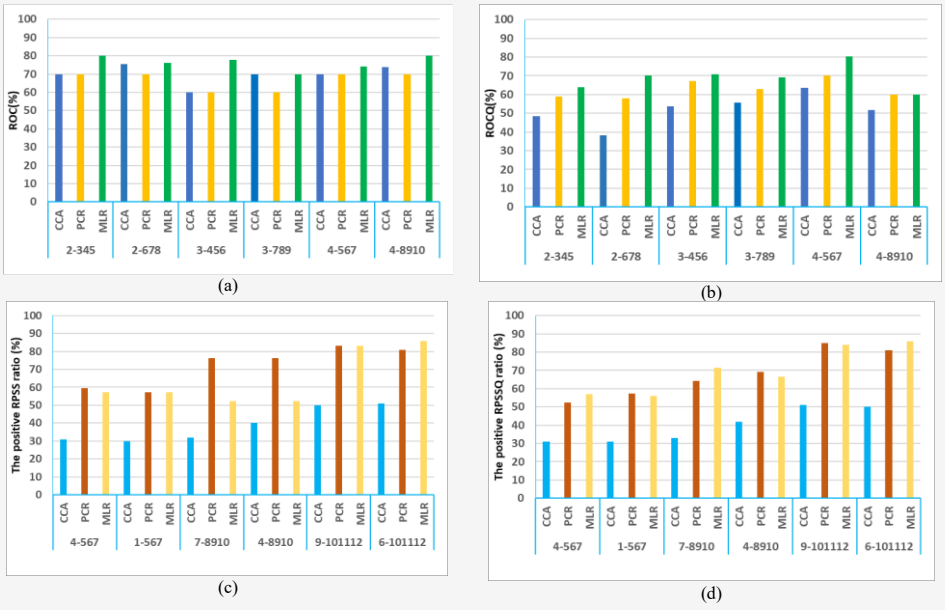
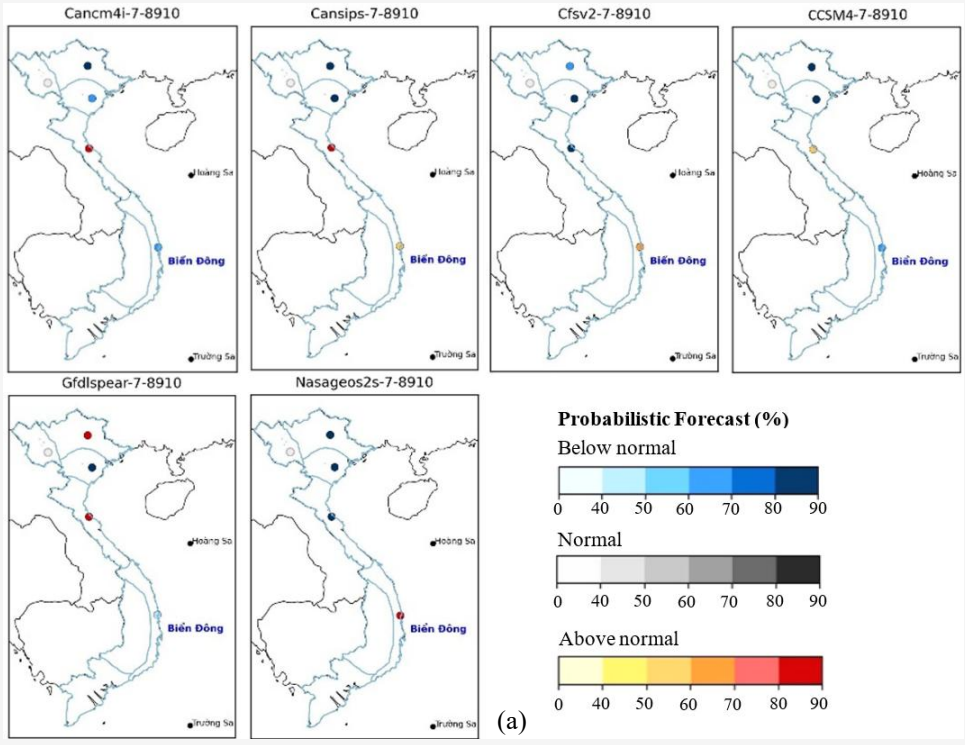


Figure 23: Evaluation of 6-month forecasts for heavy rainfall spells:
(a) ROC, (b) ROCQ, (c)The positive RPSS ratio, (d) The positive RPSSQ ratio



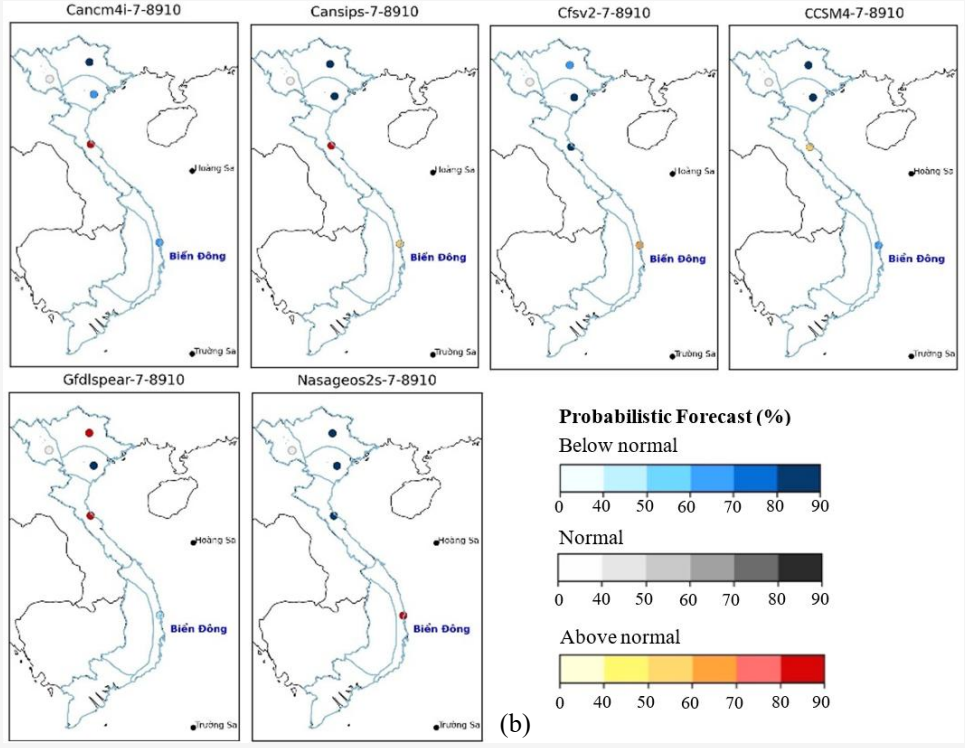


Figure 24: Probability forecast maps of moderate and heavy rainfall events:
(a) heatwaves (b) for the August–October 2025 period (forecast issued in July 2025), and severe cold waves
(c) for the November 2025 – January 2026 period (Continue next page)

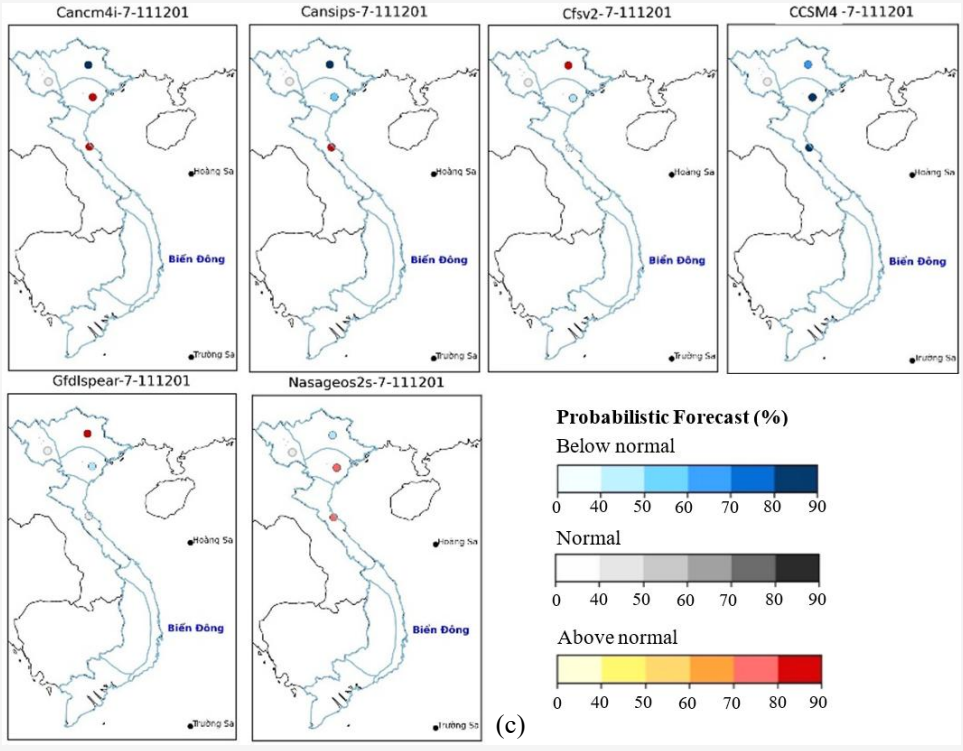


Figure 25: Probability forecast maps of moderate and heavy rainfall events: (a), heatwaves (b) for the August–October 2025 period (forecast issued in July 2025), and severe cold waves (c) for the November 2025 – January 2026 period (Continue from previous page)

4. Conclusions

This study provides a comprehensive evaluation of three statistical post-processing methods Principal Component Regression (PCR), Canonical Correlation Analysis (CCA), and Multivariate Linear Regression (MLR) for improving seasonal forecasts of extreme weather events in Vietnam. The events examined include cold spells, heatwaves, and moderate to heavy rainfall, classified by duration (2–4 days and over 5 days) across seven distinct climatic regions. Among the methods tested, PCR consistently demonstrated superior forecasting skill, outperforming CCA and MLR in most cases and across various seasons and lead times. The strong performance of PCR can be attributed to its ability to resolve multicollinearity among predictors, reduce dimensionality, and retain only the dominant climate signals from large-scale atmospheric fields. This property is particularly advantageous in Vietnam, where topographic complexity such as the Trường Sơn Mountain Range, foehn wind effects, and the contrast between coastal lowlands and highland plateaus plays a central role in modulating local climate extremes. Notably, PCR yielded Accu3P and Accu2P values exceeding 0.5–0.6 in multiple contexts, particularly for cold spells and heavy rainfall events during the October–December season. For heatwaves, PCR demonstrated especially high accuracy in the South Central Coast and Central Highlands during the early hot season (March–April), performing comparably or better than MLR. Its performance advantage is further emphasized by its low reliance on post-processing techniques, such as quantile mapping (QM), making it more efficient for operational integration.

The effectiveness of PCR contrasts sharply with earlier regional studies relying on CCA. For instance, applications of CCA in East Asia reported relatively modest gains in forecast skill, typically with correlation values below 0.3 [51] and [52]. In comparison, the present study demonstrates that PCR can achieve significantly higher accuracy, consistent with previous findings that highlighted the method's robustness in nonlinear and multivariate contexts, especially under the influence of complex seasonal signals [53] and [54]. These results reinforce the value of region-specific calibration in seasonal climate forecasting. In Vietnam, the interplay of ENSO, Madden–Julian Oscillation (MJO), Arctic Oscillation (AO), and Indian Ocean Dipole (IOD) alongside local drivers such as monsoonal wind regimes, land–sea temperature gradients, and altitudinal variation creates forecasting challenges that demand methods capable of distinguishing dominant from spurious signals.

From an operational standpoint, PCR's low computational demand and scalability make it a strong candidate for inclusion in national and regional early warning systems. Its consistent performance across different event types and regions indicates that it could be effectively integrated into Vietnam's National Center for Hydro-Meteorological Forecasting (NCHMF) systems with minimal modifications, serving as a valuable complement to existing dynamical models. Furthermore, enhancing seasonal forecast capacity is increasingly critical in the context of climate change, as Vietnam faces growing exposure to climate extremes such as heatwaves, prolonged droughts, and intense rainfall. Reliable and timely forecasts are vital for climate-sensitive sectors, including agriculture, energy, public health, and disaster risk management. In summary, this study provides robust evidence that Principal Component Regression is a statistically sound, computationally efficient, and regionally adaptable tool for seasonal climate forecasting in Vietnam. Its superior performance relative to CCA and MLR especially in topographically complex and climatically diverse regions highlights its strong potential for both research applications and operational deployment.

5. Limitations

Despite the promising results, this study also underscores several important limitations related to the input data and observational coverage. First, the coarse spatial resolution (~1°) of the six dynamical climate models used in this study may underrepresent fine-scale variability, especially in regions with steep elevation gradients. This is critical in areas such as the Central Highlands, where microclimatic conditions are influenced by terrain-induced circulations and local hydrological processes. The limitations of model resolution can result in reduced spatial specificity and potentially limit the accurate identification of localized extreme events. Second, the observational dataset used for model training and validation consists of only 186 meteorological stations across Vietnam, many of which are concentrated in lowland and coastal areas. This sparse distribution results in incomplete spatial representation, particularly in remote or mountainous regions, where weather variability is often greatest. Such gaps can introduce sampling biases, which may affect both the training of statistical models and the accuracy of verification metrics such as Accu2P, Accu3P, ROC, and RPSS. Additionally, observational uncertainties such as instrumental errors, missing records, and inconsistent maintenance practices can reduce the reliability of both reanalysis and real-time datasets. These challenges can hinder the calibration of statistical models, particularly PCR, which depends on clean and stable temporal patterns to extract meaningful principal components.

6. Recommendations for Future Works

To address these issues and further improve seasonal forecast performance, future research should prioritize the use of higher-resolution datasets, such as regional climate model (RCM) outputs, statistical downscaling products, or satellite-based reanalysis (e.g., CHIRPS, ERA5-Land). These datasets can better capture localized climate variability and provide more detailed predictors for input into statistical models. Moreover, there is strong potential for integrating additional physical predictors, such as sea surface temperature (SST) anomalies, upper-tropospheric humidity, surface pressure, zonal/meridional wind anomalies, and outgoing longwave radiation (OLR) all of which are early indicators of large-scale climatic oscillations and teleconnection patterns. Incorporating these variables could further enhance model interpretability and sensitivity, particularly in seasonal-to-subseasonal (S2S) forecasting frameworks. Another promising research direction involves the integration of machine learning (ML) and deep learning techniques, including artificial neural networks (ANNs), support vector machines (SVMs), random forests, and deep architectures like convolutional neural networks (CNNs) and long short-term memory (LSTM) models. These techniques offer advanced capabilities for capturing nonlinear relationships and temporal dependencies within climate systems. Hybrid approaches that blend dynamical model outputs with ML-based post-processing techniques may yield more adaptive, accurate, and computationally scalable forecasting tools. For Vietnam, such systems could bridge the gap between global climate signals and localized weather responses, improving operational readiness and disaster preparedness.

Moving forward, the integration of high-resolution data, additional physical predictors, and advanced machine learning methods will be essential for building more resilient, accurate, and localized forecasting systems. These efforts will be critical in supporting Vietnam's long-term climate adaptation, disaster preparedness, and sustainable development goals.

References

[1] Jahyun, C., Sung-Ho, W., Jin-Ho, Y., Jin-Young, C., Daegyun, L. and Jee-Hoon, J., (2024). Dynamical-Statistical Method for Seasonal Forecasting of Wintertime PM10 Concentration in South Korea Using Multi-

Model Ensemble. *Environmental Research Letters*, Vol. 19(6). https://doi.org/10.1088/1748-9326/ad5030

[2] Michael, T. W. T., Mitchell, S. B., Chen, J. and Williams, J., (2016). Numerical Modelling Approach for the Management of Seasonal Influenced River Channel Entrance. *Ocean and Coastal Management*, Vol. 130, 79–94. https://doi.org/10.1016/j.ocecoaman.2016.06.004

[3] Ju-Young, S., Kyu-Rang, K. and Jong-Chul, H., (2020). Seasonal Forecasting of Daily Mean Air Temperatures Using a Coupled Global Climate Model and Machine Learning Algorithm for Field-Scale Agricultural Management. *Agricultural and Forest Meteorology*, Vol. 281. https://doi.org/10.1016/j.agrformet.2019.107858.

[4] Otieno, G. L., Opijah, F. J., Mutemi, J. N., Ogallo, L. A., Anyah, R. O., Ongoma, V. and Sabiiti, G., (2014). Seasonal Rainfall Forecasting Using the Multi-Model Ensemble Technique Over the Greater Horn of Africa. *International Journal of Physical Sciences*, Vol. 2(6), 095–104. http://academeresearchjournals.org/journal/ijps.

[5] Sarah, J. M. and Richard, W., (2001). Seasonal Forecasting for Climate Hazards: Prospects and Responses. *Natural Hazards*, Vol. 23, 171–196.

[6] Fred, E. V., (2006). Seasonal Forecasting of Tropical Storm Frequency Using a Multi-Model Ensemble. *Quarterly Journal of the Royal Meteorological Society*, Vol. 132, 647–666. https://doi.org/10.1256/qj.05.65.

[7] Chloé, P., Stefano, M., Constantin, A., Rachel, H. W., Lauriane, B., Virginie, G., Georgios, F. and Javier, G. S., (2022). Seasonal Prediction of European Summer Heatwaves. *Climate Dynamics*, Vol. 58, 2149–2166. https://doi.org/10.1007/s00382-021-05828-3.

[8] Barnston, A. G. and He, Y., (1996). Skills of CCA Forecasts of 3-Month Mean Surface Climate in Hawaii and Alaska. *Journal of Climate*, Vol. 9, 2579–2605. https://doi.org/10.1175/1520-0442(1996)009<2579:SOCCAF>2.0.CO;2.

[9] Ruping, M. and David, M. S., (2002). Statistical–Dynamical Seasonal Prediction Based on Principal Component Regression of GCM Ensemble Integrations. *Monthly Weather Review*, Vol. 130, 2167–2187. https://doi.org/10.1175/1520-0493(2002)130<2167:SDSPBO>2.0.CO;2.

[10] Klein, W. H., Lewis, B. M. and Enger, I., (1959). Objective Prediction of Five-Day Mean Temperature During Winter. *Journal of Meteorology*, Vol. 16, 672–682. https://doi.org/10.1175/1520-0469(1959)016<0672:OPOFDM>2.0.CO;2.

[11] Glahn, H. R. and Lowry, D. A., (1972). The Use of Model Output Statistics (MOS) in Objective Weather Forecasting. *Journal of Applied Meteorology*, Vol. 11, 1203–1211. https://doi.org/10.1175/1520-0450(1972)011<1203:TUOMOS>2.0.CO;2.

[12] Daniel, S. W., (2020). Statistical Methods in the Atmospheric Sciences. *Elsevier*, Vol. 9, 359–428. https://doi.org/10.1016/C2017-0-03921-6.

[13] Jim, F., (2019). Making Predictions with Regression Analysis. *The American Astronomical Society*, Vol. 157. https://statisticsbyjim.com/regression/predictions-regression.

[14] Fowler, H. J. and Wilby, R. L., (2007). Beyond the Downscaling Comparison Study. *International Journal of Climatology*, Vol. 27, 1543–1545. https://doi.org/10.1002/joc.1616.

[15] Fowler, H. J., Blenkinsop, S. and Tebaldi, C., (2007). Review: Linking Climate Change Modeling to Impacts Studies Recent Advances in Downscaling Techniques for Hydrological Modelling. *International Journal of Climatology*, Vol. 27. https://doi.org/10.1002/joc.1556.

[16] Intergovernmental Panel on Climate Change (IPCC). (2021). *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press.

[17] Alve, J., Broksa, L. H., Rubbelke, D. T. G. and Vogele, S., (2024). Studying Extreme Events: An Interdisciplinary Review of Recent Research. *Heliyon*, Vol. 10. https://doi.org/10.1016/j.heliyon.2024.e41024.

[18] Rajeevan, M., Guhathakurta, P. and Thapliyal, V., (2000). New Models for Long Range Forecasts of Summer Monsoon Rainfall Over North West and Peninsular India. *Meteorologische Atmosphärische Physik*, Vol. 73, 211–225. https://doi.org/10.1007/s007030050074.

[19] Muhammad, A., Nadia, R., Shaukat, A., Shahbaz, M., Kaleem, A. M., Aftab, A. K. and Bushra, K., (2017). Prediction of Summer Rainfall in Pakistan from Global Sea-Surface Temperature and Sea-Level Pressure. *Weather*, Vol. 72, 76–84. https://doi.org/10.1002/wea.2784.

[20] Landman, W. A. and Mason, S. J., (1999). Operational Long-Lead Prediction of South African Rainfall Using Canonical Correlation Analysis. *International Journal of Climatology*, Vol. 19, 1073–1090. https://doi.org/10.1002/(SICI)1097-0088(199908)19:10<1073::AID-JOC415>3.0.CO;2-J.

[21] Hwang, S., Schemm, J. E., Barnston, A. G. and Kwon, W., (2001). Long-Lead Seasonal Forecast Skill in Far Eastern Asia Using Canonical Correlation Analysis. *Journal of Climate*, Vol. 14, 3005–3016. https://doi.org/10.1175/1520-0442(2001)014<3005:LLSFSI>2.0.CO;2.

[22] Gaoyun, W., Yizhou, Z., Rong, F., Siyu, Z. and Hongqing, W., (2021). Improving Seasonal Prediction of California Winter Precipitation Using Canonical Correlation Analysis. *Journal of Geophysical Research: Atmospheres*. https://doi.org/10.1029/2021JD034848.

[23] Golian, S., Murphy, C., Wilby, R. L., Matthews, T., Donegan, S., Quinn, D. F. and Harrigan, S., (2022). Dynamical–Statistical Seasonal Forecasts of Winter and Summer Precipitation for the Island of Ireland. *International Journal of Climatology*, Vol. 42(11), 5714–5731. https://doi. org/10.1002/joc.7557.

[24] Endo, N., Matsumoto, J. and Lwin, T., (2009). Trends in Precipitation Extremes Over Southeast Asia. *SOLA*, Vol. 5, 168–171. https://doi.org/10.2151/sola.2009-043.

[25] David, M., (2021). Verification of Forecasts for Extreme Rainfall, Tropical Cyclones, Flood and Storm Surge Over Myanmar and the Philippines. *Weather and Climate Extremes*, Vol. 33. https://doi.org/10.1016/j.wace.2021.100325.

[26] Luo, X. and Wang, B., (2017). Predictability and Prediction of the Total Number of Winter Extremely Cold Days Over China. *Climate Dynamics*, Vol. 50, 1769–1784. https://doi.org/10.1007/s00382-017-3720-z.

[27] Chaoxia, Y. and Wenmao, L., (2019). Variations in the Frequency of Winter Extreme Cold Days in Northern China and Possible Causalities. *Journal of Climate*, Vol. 32, 8127–8141. https://doi.org/10.1175/JCLI-D-18-0771.1.

[28] Chen, T. C., Huang, W. R. and Yoon, J. H., (2004). Interannual Variation of the East Asian Cold Surge Activity. *Journal of Climate*, Vol. 17, 401–413. https://doi.org/10.1175/1520-0442(2004)017<0401:IVOTEA>2.0.CO;2.

[29] Hong, C. C., Hsu, H. H., Chia, H. H. and Wu, C. Y. (2008). Decadal Relationship between the North Atlantic Oscillation and Cold Surge Frequency in Taiwan. *Geophysical Research Letters*, Vol. 35(24). https://doi.org/10.1029/2008GL034766.

[30] Soledad, C., Mariana, B. and Matilde, R., (2022). Seasonal Forecast of the Percentage of Days with Extreme Temperatures in Central-Northern Argentina: An Operational Statistical Approach. *Climate Services*, Vol. 26. https://doi.org/10.1016/j.cliser.2022.100293.

[31] Liwei, J., Thomas, L. D., Xiaosong, Y. and William, F. C., (2023). Seasonal Prediction of North American Wintertime Cold Extremes in the GFDL SPEAR Forecast System. *Climate Dynamics*, Vol. 61, 1769–1781. https://doi.org/10.1007/s00382-022-06655-w.

[32] Rasmus, E. B., Bob, V. O., Flavio, J., Frode, S., Kajsa, M. P., Abdelkader, M., Helene, B. E., Jana, S. and Milton, E. F., (2018). Downscaling Probability of Long Heatwaves Based on Seasonal Mean Daily Maximum Temperatures. *Advances in Statistical Climatology, Meteorology and Oceanography (ASCMO)*, Vol. 4, 37–52. https://doi.org/10.5194/ascmo-4-37-2018.

[33] Zhang, W., Gao, J., Qiaozhen, L., Yanzhen, C. and Tonghua, S., (2022). Probabilistic Forecast of the Extended Range Heatwave Over Eastern China. *Atmospheric Science Letters*. https://doi.org/10.3389/feart.2021.810579.

[34] Batté, L., Ardilouze, C. and Déqué, M., (2018). Forecasting West African Heat Waves at Sub–Seasonal and Seasonal Time Scales. *Monthly Weather Review*, Vol. 146, 889–907. https://doi.org/10.1175/MWR-D-17-0211.1.

[35] Katsafados, P., Anastasios, P., George, V., Papadopoulou, E. and Mavromatidis, E., (2014). Seasonal Predictability of the 2010 Russian Heat Wave. *Natural Hazards and Earth System Sciences*, Vol. 14, 1531–1542. https://doi.org/10.5194/nhess-14-1531-2014.

[36] Delworth, T. L., Adcroft, A., Anderson, W., Balaji, V., Benson, R., Dixon, K., Dunne, J., Griffies, S., Hallberg, R., Harrison, M., Horowitz, L., Krasting, J., Malyshev, S., Milly, C., Shevliakova, E., Stouffer, R., Winton, M., Wittenberg, A. and Zhang, R., (2020). SPEAR: The Next Generation GFDL Modeling System for Seasonal to Multidecadal Prediction and Projection. *Journal of Advances in Modeling Earth Systems*, Vol. 12(3). https://doi.org/10.1029/2019MS001895.

[37] Peter, R. G., Hurrell, J. W., Gent, P. R., Danabasoglu, G., Donner, L. J., Holland, M. M., Large, W. G., Lawrence, D. M., Lindsay, K., Lipscomb, W. H., Long, M. C., Mahowald, N. M., Marsh, D. R., Neale, R. B., Rasch, P. J., Vertenstein, M., Bader, D., Collins, W. D., Hack, J. J., Kiehl, J. T. and Marshall, S., (2011). The Community Climate System Model Version 4. *Journal of Climate*, Vol. 24, 4973–4991. https://doi.org/10.1175/2011JCLI4083.1.

[38] Suranjana, S., Jeffrey, P., Huug, M. V. D. D., Arun, K., Phillip, P., Lawrence, L., Michael, J. and Wanqiu, W., (2014). The NCEP Climate Forecast System Version 2. *Journal of Climate*, Vol. 27(6), 2185–2208. https://doi.org/10.1175/JCLI-D-12-00823.1.

[39] Andrea, M., Hackert, E. and Li, L., (2020). GEOS-S2S Version 2: The GMAO High Resolution Coupled Model and Assimilation System for Seasonal Prediction. *Journal of Geophysical Research: Atmospheres*, Vol. 125(5). https://doi.org/10.1029/2019JD031767.

[40] Kharin, V. V., Qiaobin, T., Francis, W. Z., George, J. B., Jacques, D. and Juan, S. F., (2009). Skill Assessment of Seasonal Hindcasts from the Canadian Historical Forecast Project. *Atmosphere-Ocean*, Vol. 47, 204–223. https://doi.org/10.3137/AO1101.2009.

[41] Anthony, G. B. and Chester, F. R., (1992). Prediction of ENSO Episodes Using Canonical Correlation. *Journal of Climate*, Vol. 5, 1316–1320. https://doi.org/10.1175/1520-0442(1992)005<1316:POEEUC>2.0.CO;2.

[42] Cornillon, P. A., Imam, W. and Matzner, E., (2008). Forecasting Time Series Using Principal Component Analysis with Respect to Instrumental Variables. *Computational Statistics and Data Analysis*, Vol. 52, 1269–1280. https://doi.org/10.1016/j.csda.2007.06.017.

[43] Andrew, F. S. and Michael, R. W., (2022). Chapter 12 – Multiple Regression: Predicting One Variable from Several Others. *Practical Business Statistics (Eighth Edition)*; 371–431. https://doi.org/10.1016/B978-0-12-820025-4.00012-9.

[44] Viet, L. V., (2022). Analysis of the Relationship Between the Standard Precipitation Index in South Vietnam and Sea Surface Temperature. *VNU Journal of Science: Earth and Environmental Sciences*, Vol. 38, 2588–1094. https://doi.org/10.25073/2588-1094/vnuees.4721.

[45] Thang, V. V., Hieu, T. N., Thang, V. N. and Hiep, V. N., (2015). Effects of ENSO on Autumn Rainfall in Central Vietnam. *Advances in Meteorology*, Vol. 2015, 1–12. https://doi.org/10.1155/2015/264373.

[46] Jun-Haeng, H., Hyunjun, A., Ju-Young, S., Thomas, R. K. and Changsam, J., (2019). Probability Distributions for a Quantile Mapping Technique for a Bias Correction of Precipitation Data: A Case Study Under Climate Change. *Water*, Vol. 11. https://doi.org/10.3390/w11071475.

[47] Nguyen-Le, D., Ho, T. L. and Pham, H. V., (2021). Evaluation of CFSv2 Seasonal Forecasts in Vietnam. *Climate*, Vol. 9(7). https://doi.org/10.3390/cli9070135.

[48] APEC Climate Center. (2023). APEC Climate Outlook for May–July 2023. *APEC Climate Center*. https://www.climate.go.kr/home/eng/apec_view.php?bbsId=eng_apec_news&dataId=343&currentPageNo=1

[49] Do, H. X., Nguyen, T. H., Tran, M. D. and Le, T. T., (2024). Future Changes in Hydro-Climatic Extremes across Vietnam: Evidence from a Semi-Distributed Hydrological Model Forced by Downscaled CMIP6 Climate Data. *Water*, Vol. 16(5), https://doi.org/10.3390/w16050674.

[50] DeRISK Southeast Asia. (2021). Applying Seasonal Climate Forecasting to Support Climate Risk Management in Vietnam. *Climate Change Vietnam*. https://www.climatechange.vn/applying-seasonal-climate-forecasting-to-support-climate-risk-management-in-vietnam.

[51] Hwang, S. O., Kang, I. S. and Jin, K., (2001). Leading Large-Scale Modes of Objectively Analyzed 500 hPa Heights and Their Use in Long-Range Forecasts in Korea and Japan. *Journal of Climate*, Vol. 14(13), 3005–3015. https://doi.org/10.1175/1520-0442(2001)014.

[52] Guo, Y., Zhang, Q., Huang, H., Li, J. and Zhang, H. (2020). Improving Early Summer Precipitation Forecasts in Southern China Based on Statistical Downscaling of CFSv2. *Weather and Forecasting*, Vol. 35(4), 1273–1291. https://doi.org/10.1175/WAF-D-19-0230.1.

[53] Kim, H. M., Webster, P. J. and Curry, J. A., (2017). Data-Adaptive Principal Component Regression Analysis for Seasonal Prediction. *International Journal of Climatology*, Vol. 37(8), 3356–3369. https://doi.org/10.1002/joc.4939.

[54] Wang, B., Xiang, B. and Lee, J. Y., (2013). Subtropical High Predictability Establishes a Prominent Impact of the Indian Ocean Dipole on East Asian Summer Monsoon Rainfall. *Journal of Climate*, Vol. 26(9), 3167–3181. https://doi.org/10.1175/JCLI-D-12-00476.1.