HYDROLOGIC FORECASTING BASED ON STATISTICAL AND PHYSICAL APPROACHES FOR THE UPPER CHAO PHRAYA RIVER BASIN, THAILAND

by

Nkrintra Singhrattna

A dissertation submitted in partial fulfillment of the requirements for the
degree of Doctor of Philosophy in
Water Engineering and Management

Examination Committee:	Dr. Mukand S Babel (Chairperson) Dr. Sutat Weesakul (Member) Dr. Sylvain Roger Perret (Member) Dr. Kiyoshi Honda (Member)
External Examiner:	Prof. Stefan Uhlenbrook Processor of Hydrology, UNESCO-IHE Institute of Water Education, The Netherlands
Nationality:	Thai
Previous Degree:	Master of Science in Water Resources University of Colorado, U.S.A.
Scholarship Donor:	Royal Thai Government – AIT Fellowship

Asian Institute of Technology
School of Engineering and Technology
Thailand
May 2012

Acknowledgements

Apart from the efforts of myself, this dissertation would not have been successful without Dr. Mukand S Babel who not only served as my academic supervisor, but he also supported, encouraged and challenged me from the initial to the final steps of dissertation process. Other committee members: Dr. Sutat Weesakul, Dr. Sylvain Roger Perret and Dr. Kiyoshi Honda guided me through the whole process. I would like to thank them all.

Prof. Pierluigi Claps (Politecnico di Torino, Italy) is acknowledged for his professional guidance and useful comments. I would like to extend my gratitude to Dr. Elisa Bartolini (Politecnico di Torino, Italy) who shared her academic experience and reviewed on my draft report.

I would like to show my greatest appreciation to Prof. Stefan Uhlenbrook (UNESCO-IHE Institute of Water Education, The Netherlands) who reviewed on my dissertation. Without doubt, his detailed comments improved report.

The Department of Public Works and Town & Country Planning is acknowledged for the kind support. I am grateful to the Royal Thai Government for AIT fellowships and thankful to CIRAD and Erasmus Mundus EuroAsia Program for the scholarships that offered me the great academic experience in Europe.

I am forever indebted to my mother for her unconditional love enabled me to complete this dissertation. I am most grateful to my father for his endless patience and his support from kindergarten to graduate levels. Without two important persons, I would not have achieved PhD. I am also thankful to my sister for her support and understanding. My thanks go to all members of my family whose encouragement helped me to handle with difficulties.

Special thanks to my best friend: Sasirin Chumchuensuk and her family for their encouragement and for feeding me when it was most required. Big thanks go to Piyawat Wuttichaikitcharoen for his data support. I also would like to thank my friends and colleagues in Thailand and Italy who helped me along the way.

Abstract

Hydroclimates like precipitation and streamflow are related to large-scale atmospheric variables via oceanic-atmospheric circulations. Unusual patterns of circulation are connected to changes in climatic conditions due to factors like population and technology growth, urbanization and economic development. El Niño and La Niña, defined as anomalous sea surface temperatures over the tropical Pacific Ocean, interrupt the Walker circulation and cause variability in hydroclimates and anomalous weather events. The case study (i.e. the Ping River Basin which is a sub-basin of the Upper Chao Phraya River Basin) is located in northern Thailand. The Ping River Basin covers an area of 33,899 km^2. The climate is classified as tropical monsoon with all average monthly temperatures greater than 18°C, and the highest temperature occurs in a period prior to the monsoon season. Moreover, precipitation being less than 61 mm per month is found in one or more months. The Ping River Basin experiences dry season rainfall (from November to April) which is inversely related to air temperature in the summer season (i.e. March-April-May or MAM). From 1951 to 2007, an increasing trend in dry season rainfall (by 16.3 mm over 57 years) is consistent with a decreasing trend of MAM temperature (drops by 0.6°C over 57 years). Furthermore, a higher MAM temperature influences the land-sea temperature gradient and strengthens the monsoon. The pre-monsoon season rainfall (i.e. May-June-July or MJJ) is inversely correlated to MAM temperature, whereas the monsoon season rainfall (i.e. August-September-October or ASO) is positively related to MAM temperature. In El Niño years, the MJJ and ASO rainfalls tend to decrease and vice versa in La Niña years

Using correlation maps, seasonal rainfall (i.e. MJJ, ASO, NDJ and FMA) of the study basin can be statistically related to large-scale atmospheric variables (sea surface temperature, sea level pressure, surface zonal and meridian winds) at long lead times, varying from 4 to 15 months prior to the start of the season. Atmospheric predictors are identified over different regions (such as the Pacific and Indian Oceans) based on significant relationships with rainfall at 95% confidence levels. The gridded monthly data of identified predictors from 1961 to 2100 is obtained from a general circulation model (GCM) called GFDL-R30 and used in a statistical model to forecast and determine the effects of future climate on seasonal rainfall of the study basin.

A modified k-nearest neighbor (k-nn) model is used to downscale the rainfall of the study basin from large-scale atmospheric variables. The modified k-nn model is a nonparametric approach, which locally fits a regression between dependent (e.g. rainfall) and independent variables (e.g. atmospheric predictors) using a small set of neighbors (k) at any given point. k and the order of polynomial (p) are selected using a generalized cross validation (GCV) method. In terms of the effects of future climate under two scenarios (A2 and B2), the 2011-2100 MJJ and ASO rainfall of the Ping River Basin is predicted to decrease by 0.11-6.16 mm per year. Increasing trends of 0.02-5.91 mm per year are associated with the 2011-2100 dry season rainfall (i.e. NDJ and FMA). Furthermore, the monsoon season rainfall will have more chances of being dry and less chances of being wet. In contrast, future climate will affect more chances of wetness and less chances of dryness for the dry season rainfall. The wet season will tend to shift by two seasons, from ASO to FMA, under A2 and by one season, from ASO to NDJ, under B2.

To compare two algorithms of rainfall-runoff models, the SIMHYD and HEC-HMS models have been studied. Both models are calibrated from April 1999 to March 2003 and validated from April 2003 to March 2007. Four efficiency indexes (the deviation of volume (D_v), correlation coefficient (r), normalized root mean square error (NRMSE) and the Nash-Sutcliffe efficient index (E)) are used to evaluate the model performance. The SIMHYD model shows poor performance, due to the homogeneity of basin characteristics, in capturing average monthly streamflow at the stations that cover a large drainage area. Comparing the performance of these models at six gauging stations, the HEC-HMS model performs better than the SIMHYD model in capturing average monthly streamflow. Moreover, the HEC-HMS model can capture low flow better than the SIMHYD model. Although its performance is not consistent at all stations, more efficiency in high flow simulation is associated with the HEC-HMS model as it shows less NRMSE and greater E. Therefore, the HEC-HMS model has been selected to simulate the 2011-2100 daily streamflow using rainfall ensembles obtained from the multisite daily rainfall generator.

The effects of future climate under both scenarios present decreasing average discharges in the dry and wet seasons. The shift in peak discharge from mid-September to the end of September or the beginning of October is expected to be observed. With the exception of Station P75, P67 and 061302, the dry spells will be shorter in the future compared to historical records. Less severity of shortage during the dry spell is also predicted for all stations except P67 and 061302. Under A2, with the exception of Station P24A, 061302 and P14, the wet spells will be shorter. Under Scenario B2, wet spells will be shorter at all stations. The intensity of abundance is less than that seen in historical records. Anomalous low flow in the wet season and anomalous high flow in the dry season have also been examined using thresholds of the observed Q_{90} in the wet season ($Q_{90,wet}$) and the observed Q_{10} in the dry season ($Q_{10,dry}$). Due to future climate alterations, the magnitudes of simulated $Q_{90,wet}$ and $Q_{10,dry}$ will likely be lower than those of observations. With the exception of Station P75 and P21, dry spells in the wet season will be shortened. Anomalous low flow in the wet season during 2011-2100 is less severe than historical records. In terms of anomalous high flow in the dry season, a shorter duration of wet spells with lesser intensity of abundance will be found at four stations under A2 and at three stations under B2.

Table of Contents

Table of Contents

List of Tables

List of Tables

List of Figures

List of Figures

Title **Page**

List of Figures

List of Figures

Chapter 1
Introduction

1.1 Background and statement of the problem

The hydroclimates (e.g. precipitation) at a basin scale are influenced by oceanic-atmospheric circulations. Seen over the eastern equatorial Pacific Ocean, the El Niño-Southern Oscillation (ENSO) is an anomalous oceanic-atmospheric circulation in terms of sea surface temperature (SST), and it has been linked with changes in climate in the region. The changes in climate are caused by the increase in greenhouse gas concentration in the atmosphere due to population growth, urbanization, economic and technology development. Climate change causes variability of hydroclimates such as temperature over land and the sea, precipitation and the amount of water present in other forms (e.g. streamflow, underground water and ice glaciers) in a hydrologic cycle. A report of the Intergovernmental Panel of Climate Change (IPCC) estimated the variability of annual global surface temperature from 1956 to 2005 as varying from 0.10 to 0.16°C per decade (IPCC, 2007a). However, due to global warming, increasing trends in annual surface temperature were found in several regions. For average land temperatures (Figure 1.1), the temperature departures estimated with respect to 1961-1990 average temperature show a warmer trend after 1980.

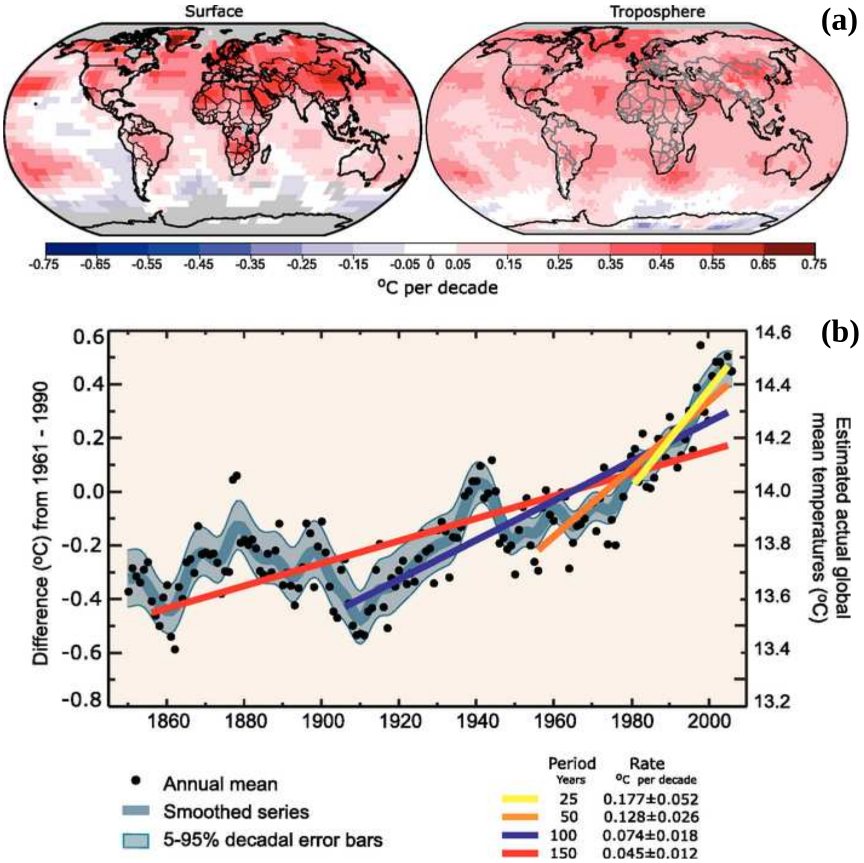


Figure 1.1: (a) Linear trends in global annual temperature from 1979 to 2005; and (b) Annual temperature anomalies estimated with respect to 1961-1990 average.
Source: http://www.ipcc.ch/publications_and_data/ar4/wg1/en/tssts-3-1-1.html

Precipitation variability, as seen in the trends of annual global precipitation (Figure 1.2) from 1979 to 2005, shows a variation from -60% to +60% per decade (IPCC, 2007a). Several developing countries, especially in the arid and semi-arid regions of southern

Africa, the Mediterranean and part of southern Asia, observed decreasing trends. However, in some regions such as northern Europe, northern and central Asia and North and South America, increasing trends were found. The variability in hydroclimates encourages engineers, researchers and scientists to work on several related topics aiming at understanding the relationships between local hydroclimates and large-scale atmospheric variables, and at mitigating anomalous events like floods and droughts. In this study, the variability of hydroclimates is a motivation to address the problems stated below:

(i) How does the variability of local hydroclimates relate to large-scale atmospheric variables?

(ii) How to develop a model with the incorporation of large-scale atmospheric information to forecast rainfall?

(iii) What could be the effects of future climate on rainfall and streamflow at the basin scale?

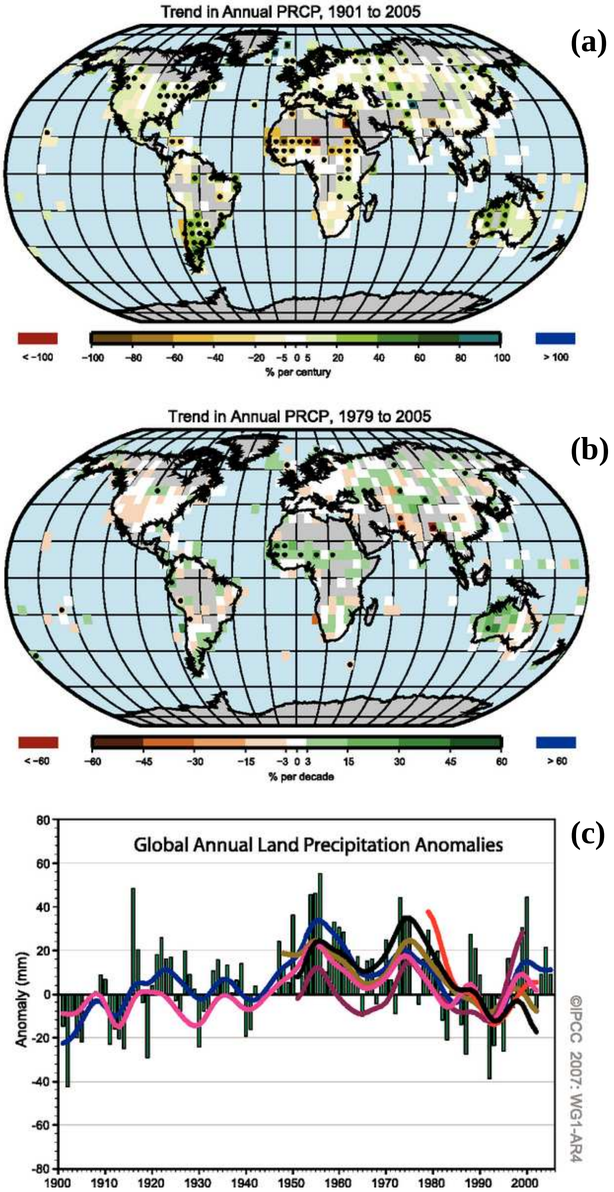


Figure 1.2: (a) Linear trends in global annual precipitation from 1901 to 2005; (b) Linear trends in global annual precipitation from 1979 to 2005; and (c) Annual precipitation anomalies estimated with respect to 1961-1990 average.

Source: http://www.ipcc.ch/publications_and_data/ar4/wg1/en/tssts-3-1-3.html

Variability in hydroclimates at the basin scale can cause social and economic problems, especially in developing countries, because their economy is dependent upon the rain-fed agriculture. The total annual losses from extreme weather events are estimated more severe in the present (i.e. since 1990) than were in the past (IPCC, 2007a) because of more severe disasters in terms of intensity and frequency. These disasters have been associated with changes in climate, population growth, economic development and urbanization. Both tangible and intangible damages influence a discontinuity of economic growth and development. Moreover, the effects on the quantity and quality of water are directly connected to deteriorating living standards and degrading natural environment because water availability is also related to uses by the ecosystem and the environment. Therefore, the rationale of this study is to:

(i) Understand the effects of oceanic-atmospheric circulation on the hydroclimates in a study basin.
(ii) Implement large-scale atmospheric variables into a forecasting model of hydroclimates.
(iii) Examine the effects of future climate on anomalous weather events.

1.2 Objectives of the study

The broad objectives of this study are to develop a model to forecast rainfall and streamflow with a set of selected predictors of large-scale atmospheric variables, and also to determine the effects of future climate on rainfall and streamflow using the developed model. The specific objectives of the study are as follows:

(i) To understand and develop the statistical relationships between rainfall and large-scale atmospheric variables.
(ii) To propose a statistical model to forecast rainfall with a set of selected large-scale atmospheric variables.
(iii) To downscale local rainfall from the large-scale atmospheric variables obtained from a general circulation model (GCM) and to determine the effects of future climate using the developed statistical model.
(iv) To simulate streamflow using a rainfall-runoff model with the historical and downscaled rainfalls.
(v) To determine anomalous streamflow events (i.e. low and high flow) that will occur due to future climate using downscaled rainfall and simulated streamflow.

1.3 Scope of the study

The objectives outlined above are pursued according to the following steps:

(i) Data collection of hydroclimates (e.g. rainfall, streamflow and temperature) from several stations in the study basin (the Upper Chao Phraya River Basin, Thailand) and large-scale atmospheric information (e.g. SST, sea level pressure (SLP) and wind) over several locations in the Pacific and Indian Oceans.
(ii) Correlation and composite analysis to develop statistical relationships between rainfall and large-scale atmospheric variables and to identify predictors at long lead times that can be used to forecast rainfall.
(iii) Development of a statistical model to forecast rainfall with the predictor sets of large-scale atmospheric variables identified from (ii).
(iv) Downscaling rainfall from the large-scale atmospheric variables to determine the effects of future climate using the developed statistical model.

(v) Development of a conditioning daily rainfall generator to resample historical daily rainfall based on the probability density function (PDF) of downscaled rainfall obtained from (iv).

(vi) Calibration and validation of rainfall-runoff models using observed streamflow in the study area.

(vii) Simulation of daily streamflow using downscaled rainfall under future climate scenarios.

(viii) Estimation of anomalous hydrological events based on the rainfall and streamflow simulations obtained from (v) and (vii).

1.4 Organization of the report

This study proposes two approaches of the hydrological simulation. The first one is a statistical model with the integration of large-scale atmospheric variables as the predictors to forecast rainfall. The criteria for potential predictors of atmospheric variables are: (i) significant correlation with rainfall; and (ii) a long leading relationship between rainfall and atmospheric variables. The second approach is a physical model to simulate streamflow. The streamflow simulations will respond to rainfall ensembles obtained from the statistical model. Data gleaned from both approaches to rainfall and streamflow simulation will aid in long-term planning of water resources and reservoirs in the study area.

This report firstly introduces a review of literature in Chapter 2 (Figure 1.3) where variability of local hydroclimates in various areas and their links to atmospheric variables is shown. In addition, the implements of large-scale atmospheric variables in the forecasting models are also reviewed here. The case study (i.e. the Upper Chao Phraya River Basin) is described in Chapter 3. The development of a statistical relationship between rainfall in the study basin and large-scale atmospheric variables is also presented. Chapter 4 covers the selection of atmospheric predictors for the rainfall forecasting model based on significant statistical relationships. After identifying the potential predictors, a statistical stochastic model is proposed in Chapter 5. This chapter also includes: (i) a conditioning daily rainfall generator to resample the historical daily rainfall based on the PDF obtained from the statistical model; (ii) the evaluation of model performance of the statistical model and rainfall generator; and (iii) the downscaling of local rainfall from the atmospheric variables obtained from a general circulation model (GCM) using developed statistical model. The downscaling aims to determine the effects on rainfall under future climate scenarios proposed by IPCC. The streamflow simulation achieved from a rainfall-runoff model using downscaled rainfall and the estimation of anomalous conditions associated with the downscaled rainfall and simulated streamflow are shown in Chapter 6 and Chapter 7 respectively. Lastly, Chapter 8 addresses the conclusions of the entire work and suggests extensions of this research.

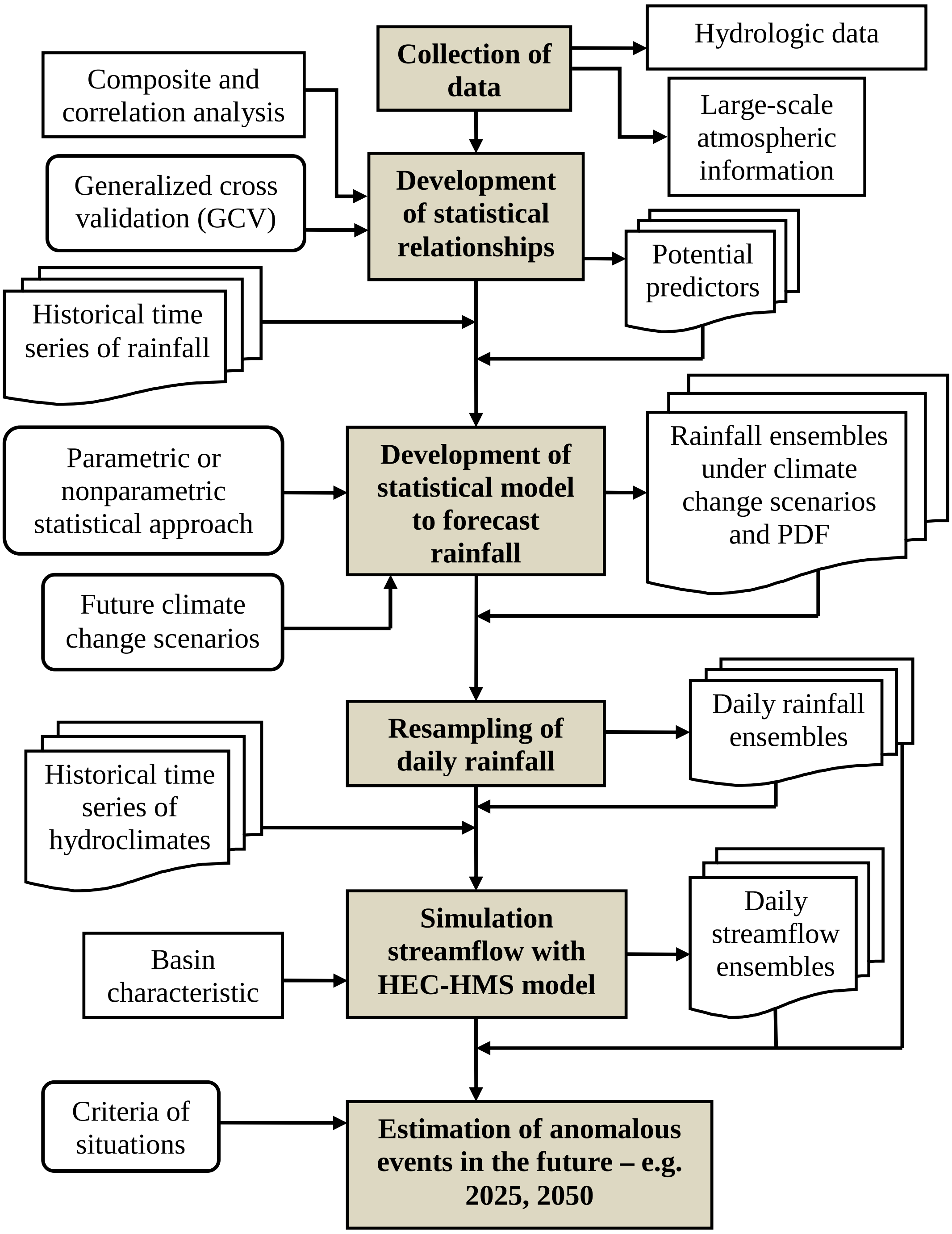


Figure 1.3: A flow chart of this study.

Chapter 2
Review of Literature

2.1 Anomalies in atmospheric conditions over the Pacific Ocean

Under normal condition (Figure 2.1), the Walker circulation consists of trade winds moving warm-moist surface air from the east to the west, due to temperature gradients, across the tropical Pacific Ocean. The moist air rises above the western tropical Pacific Ocean, forms clouds, and falls down as precipitation in this region. After losing its moisture, the drier air blows from the west to the east and descends back to the surface in the eastern tropical Pacific Ocean to complete the Walker circulation. As per oceanic circulation, trade winds gather warm surface moisture on the west coast of the tropical Pacific Ocean or Indonesia, so the water surface temperature in the west is normally higher than in the east (Figure 2.2). In terms of thermocline (see also Figure 2.1), the sea levels along the west coast of the tropical Pacific Ocean are higher than the coast of South America, which causes an occurrence of nutrient-rich cold water along the coasts of Peru and Ecuador by upwelling the cold water from deeper levels. The maximum temperature gradients between sea surface temperature (SST) along the west and the east coasts of the tropical Pacific Ocean are observed during September and October due to minimum temperatures on the east coast.

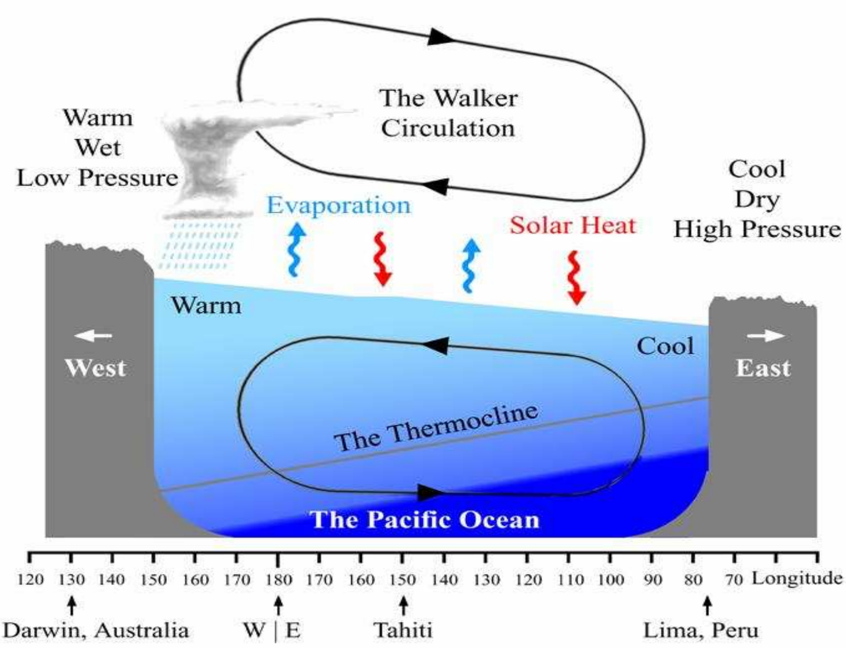


Figure 2.1: The Walker circulation under normal conditions.
Source: www.absoluteastronomy.com/topics/Walker_circulation

El Niño and La Niña are those oceanic-atmospheric phenomena which indicate anomalies of SST in the eastern equatorial Pacific Ocean. They are sometimes called El Niño-Southern Oscillation (ENSO) because the effects of SST anomalies on oceanic circulation and atmospheric circulation are observed in the Southern Oscillation. The air pressure difference between Tahiti (southern Pacific Ocean) and Darwin (northern Australia) is one example of SST anomalies. During the El Niño phase, the weakening trade winds of the Walker circulation develop an unusual pattern of oceanic-atmospheric circulation (Figure 2.3). Warm SST does not cover only the west coast, as during normal conditions, but also extends to the central and eastern tropical Pacific Ocean. As a consequence, SST anomalies are observed in these regions (Figure 2.4). These SST anomalies can interrupt oceanic circulation and the upwelling of nutrient-rich cold water along the west coast of South America. The El Niño is defined as the averaged SST over the date line and the eastern tropical Pacific Ocean being above its normal temperature by more than 0.5°C for five consecutive months or longer. For a strong El Niño, the average SST anomaly is estimated around 2.0-3.5°C warmer than normal.

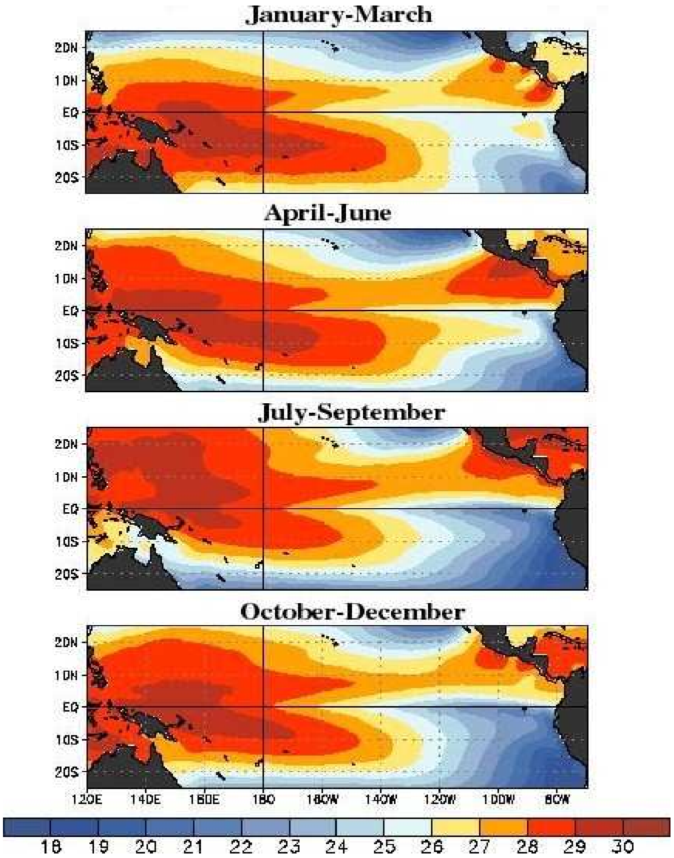


Figure 2.2: Average sea surface temperatures (°C) under normal conditions.
Source: www.cpc.noaa.gov/products/analysis_monitoring/ensocycle/meansst.shtml

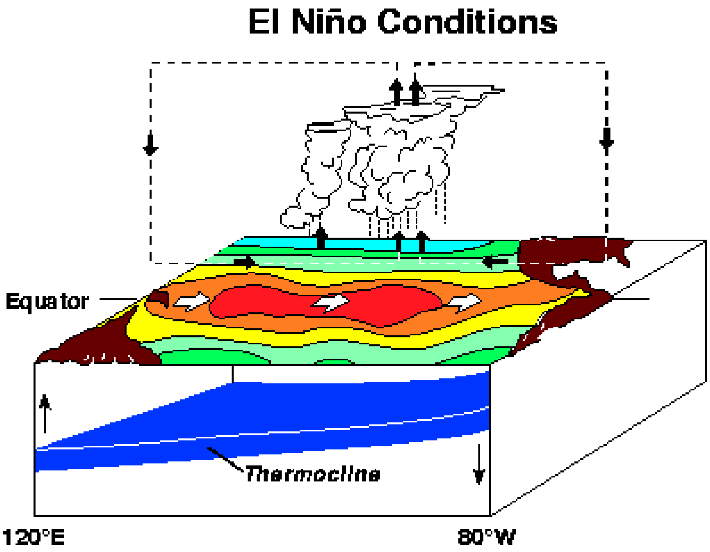


Figure 2.3: The Walker circulation in the presence of El Niño.
Source: http://www.pmel.noaa.gov/tao/proj_over/diagrams/gif/nina_normal_nino.gif

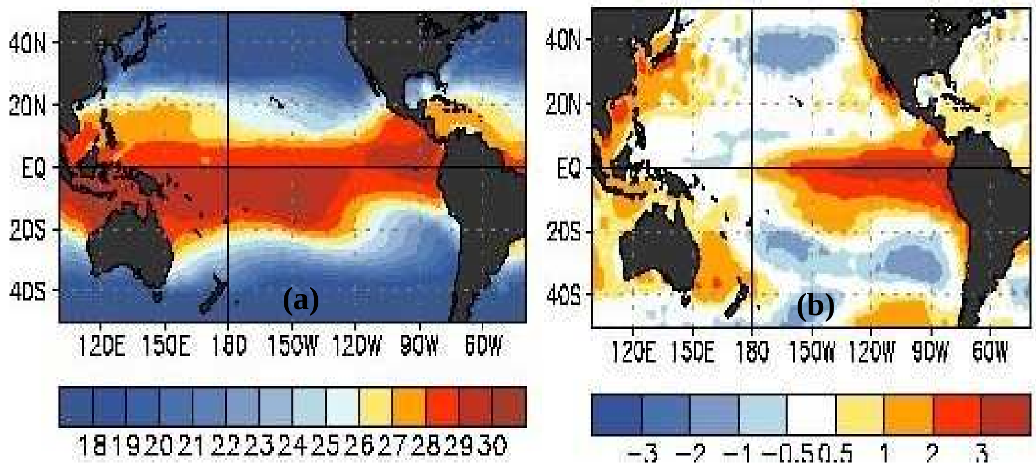


Figure 2.4: (a) Average sea surface temperatures (°C) from January to March 1998; and (b) Temperature departures (°C) under El Niño.
Source: www.cpc.noaa.gov/products/analysis_monitoring/ensocycle/ensocycle.shtml

The El Niño, based on historical records, can be observed once in 2-7 years; for example, it occurred in 1951, 1953, 1957-58, 1965, 1969, 1972-73, 1976, 1982-83, 1986-87, 1991-92, 1994 and 1997-98. The regions used to define an anomalous condition can be divided into four locations between the date line and the eastern tropical Pacific Ocean (Figure 2.5). The different variables (e.g. SST and sea level pressure) measured over these regions are used to estimate the anomalies. The four regions located are described below (Schöngart and Junk, 2007):

(i) NINO 1+2 covers the regions of the coast of Peru, Ecuador and Galápagos Islands in the equatorial Pacific Ocean. Its location is between 0°-10°S latitude and 90°-80°W longitude.

(ii) NINO 3 is located between 5°N-5°S latitude and 150°-90°W longitude, corresponding with the central equatorial Pacific Ocean.

(iii) NINO 3.4 is the overlapping region between NINO 3 and NINO 4, between 5°N-5°S latitude and 170°-120°W longitude (as shown in Figure 2.5).

(iv) NINO 4 is located between 5°N-5°S latitude and 160°E-150°W longitude, which lies in the western equatorial Pacific Ocean.

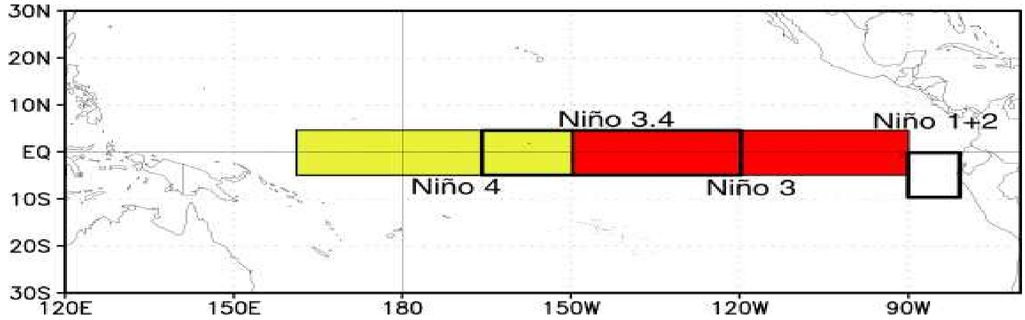


Figure 2.5: El Niño regions.
Source: www.cpc.noaa.gov/products/analysis_monitoring/ensostuff/nino_regions.shtml

During La Niña events (Figure 2.6 and 2.7), the Walker circulation is disrupted by below-normal SST in the eastern tropical Pacific Ocean. This causes the atmospheric circulation to be stronger than normal, which promotes more rising air, cloudiness and precipitation over the western Pacific Ocean and Indonesia. Hence, there is more descending air over the eastern side of the Pacific Ocean. La Niña is said to be occurring basically if the average

SST over the eastern tropical Pacific Ocean is below the normal temperature by more than 0.5°C for five consecutive months or longer. La Niña was observed in 1954-56, 1961-62, 1964, 1970-71, 1973-76, 1988-89, 1995 and 1998-2000. Due to a strong La Niña, SST over the eastern Pacific Ocean is found to be 1.0-3.0°C colder than normal. The onset period of El Niño and La Niña is between June and August, and the peak often occurs from December to April. Then, they decay in the months of May, June and July of the following year. The total duration from the onset to decay covers 9-12 months.

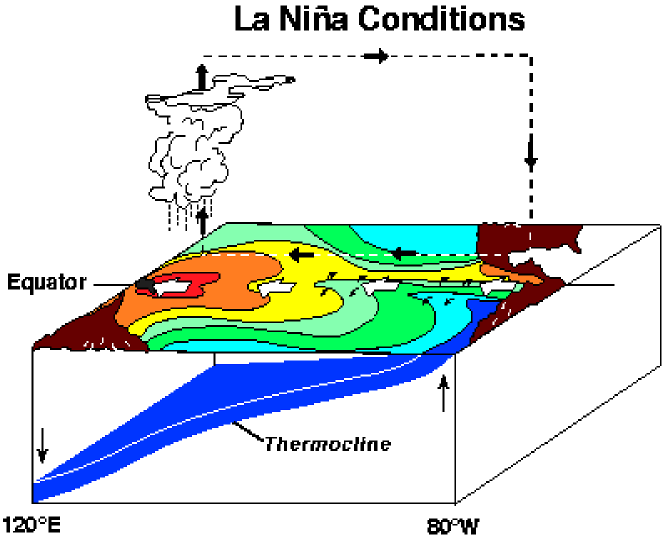


Figure 2.6: The Walker circulation in the presence of La Niña.
Source: www.pmel.noaa.gov/tao/proj_over/diagrams/gif/nina_normal_nino.gif

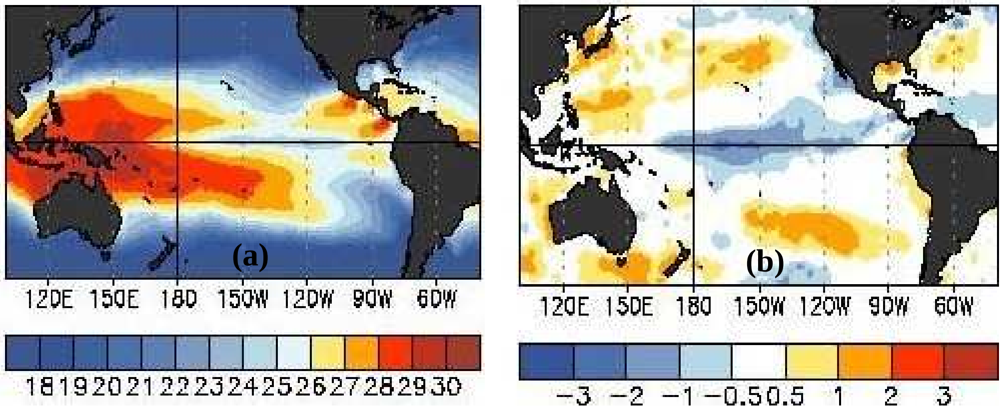


Figure 2.7: (a) Average sea surface temperatures (°C) from January to March 1989; and (b) Temperature departures (°C) under La Niña.
Source: www.cpc.noaa.gov/products/analysis_monitoring/ensocycle/ensocycle.shtml

The unusual pattern of oceanic-atmospheric circulation, i.e. ENSO, has to be carefully observed and its effects determined because ENSO has negative and positive relationships with regional hydroclimates, and this can affect various related activities directly or indirectly. Several researchers have attempted to study ENSO effects in different regions; however, it is difficult to determine with certainty the corresponding effects in the Pacific Ocean because of the nonlinear relationships and the asymmetrical responses of regional hydroclimates under El Niño and La Niña.

2.2 Effects of anomalous atmospheric conditions on hydroclimates

In order to define an anomalous condition of the oceanic-atmospheric circulation, large-scale atmospheric variables (e.g. sea surface temperature (SST), wind, sea level, changes in depth of thermocline, sea level pressure (SLP) and outgoing longwave radiation (OLR)) have to be monitored. Depending on measurable variables and locations, the anomalies associated with climatic circulation, known as teleconnections (Ropelewski and Halpert, 1987; Garreaud and Battisti, 1999; Glantz, 2001; Chou et al., 2003; Frederiksen and Branstator, 2005; An et al., 2007; Anderson, 2007; Frankignoul and Sennéchael, 2007; Li et al., 2007; L. Wu et al., 2007), are defined by several indices such as El Niño-Southern Oscillation (ENSO), Pacific Decadal Oscillation (PDO), Pacific North American (PNA) pattern, North Pacific (NP) index, Arctic Oscillation (AO) and North Atlantic Oscillation (NAO). Moreover, the anomalous conditions can be segregated under warm and cold phases, i.e. El Niño and La Niña, respectively.

The oceanic-atmospheric circulation (e.g. Walker circulation) influences local hydroclimates such as temperature, precipitation and streamflow. Although the short-term variability of hydroclimates is influenced by various local factors (e.g. human activities and land uses), long-term variability experiences the influences of anomalous atmospheric conditions. Several studies have focused on anomalous SST over the tropical Pacific Ocean (E. M. Rasmusson and Carpenter, 1983; Mantua et al., 1997; Gong and Wang, 1999; Whitaker et al., 2001). To understand the relationships between local hydroclimates and large-scale atmospheric variables, the interdecadal (Y. Zhang et al., 1997; Mestas-Nuñez and Enfield, 2001), interannual and intraseasonal variability (Frederiksen and Branstator, 2005; Keenlyside and Latif, 2007; Masih et al., 2011) of hydroclimates have been investigated.

Moreover, variability and trends in terms of spatial coverage such as a global scale (Mason and Goddard, 2001; Frederiksen and Branstator, 2005), a regional scale (Orlanski, 2005; Goswami et al., 2006) and a basin scale (Mendoza et al., 2005; Yang et al., 2007) have also been estimated. Due to atmospheric circulation, the anomalous condition of large-scale atmospheric variables could affect hydroclimates over the regions which are located near or even at a distance from the anomalous sources (Saravanan and Chang, 2000; Harshburger et al., 2002; Tereshchenko et al., 2002). Under anomalous conditions, both positive and negative relationships between local hydroclimates and atmospheric variables have been witnessed, which were also associated with anomalous weather events like floods and droughts. The effects of the warm and cold phases of the anomalous conditions respond asymmetrically; El Niño may cause extremely dry conditions, whereas La Niña may affect wet conditions with a different level of anomaly (Gershunov, 1998; Mason and Goddard, 2001; A. Shrestha and Kostaschuk, 2005; An et al., 2007). Furthermore, in some regions, the relationship is hardly identifiable because local hydroclimates are also influenced by several oceans at a time; in this case, the Pacific, Indian and the Atlantic Oceans (R. Wu and Kirtman, 2004; Nagura and Konda, 2007). In addition, the lag from an anomalous event can be found several months after an anomalous event has occurred (Gutiérrez and Dracup, 2001; Sourza Filho and Lall, 2003; Grantz et al., 2005), which make the occurrence difficult to understand in isolation. However, lagging effects can be estimated by forecasting models of local hydroclimates and this forecasting can improve the effectiveness of water resources management and planning.

This chapter presents a review of past studies that show the effects of anomalous atmospheric conditions on hydroclimates at global, regional and basin scales. A review of hydroclimate forecasting models, with the incorporation of large-scale atmospheric variables to improve model performance, is also presented.

2.2.1 Effects on global hydroclimates

As for the influences of atmospheric circulation on global hydroclimates, several approaches such as the empirical method and the atmospheric general circulation model (AGCM) have been adopted to show ENSO-related global precipitation in Australia, North America, South America, the Indian subcontinent, Africa and Central America (Ropelewski and Halpert, 1987; Enfield and Alfaro, 1999; Giannini et al., 2000). For the 11 strongest ENSO events, the effects of ENSO are recorded on a gridded monthly precipitation. These data, obtained from almost 12,000 stations around the world in post-1951, show that seasonal land precipitation tends to be below normal when associated with El Niño and vice versa for La Niña (Mason and Goddard, 2001). Furthermore, the influence of La Niña on monthly precipitation covers more area than El Niño. From September to November, the influence of El Niño and La Niña covers the most widespread areas of 22% and 25% of total land areas respectively. The regions indicating a high frequency of below-normal precipitation due to El Niño are eastern Australia and Indonesia; the regions obtaining a high frequency of above-normal precipitation due to La Niña are the Middle East, east Africa and the U.S.

On the other hand, using the empirical method, Ropelewski and Halpert (1987) found that the high frequencies of above-normal precipitation associated with La Niña can be also found in Indonesia and northern Australia, whereas below-normal precipitation can be observed in southeastern South America, northeastern Argentina, western Saudi Arabia, Kyrgyzstan and the equatorial Pacific islands. Subsequently, the relationships between global streamflow and ENSO have been investigated (Dettinger et al., 2000; Chiew and McMahon, 2002) using statistical, correlation and harmonic analyses. Anomalous weather events such as floods and droughts in some regions (e.g. Mexico and Colombia) show significant links to ENSO (Jain and Lall, 2001).

2.2.2 Effects on regional hydroclimates

The influence of anomalous atmospheric conditions such as ENSO on regional hydroclimates has been investigated over northeastern South America (NSA), southeastern South America (SSA), the U.S., Canada, Latin America and Asia. Ropelewski and Halpert (1987) found stronger effects of ENSO over NSA than over SSA. The below-normal precipitation associated with ENSO has been observed in north equatorial Brazil, French Guiana, Surinam, Guyana and Venezuela. However, above-normal precipitation has been observed in SSA, Uruguay and parts of northeastern Argentina.

From theoretical arguments and modeling studies, hurricanes or cyclones over the Atlantic Ocean have been associated with SST anomalies, which in turn are related to the global mean surface air temperature (Emanuel, 2005; Webster, 2005). Using partial correlation and regression analysis, Elsner et al. (2006) found that the increasing hurricane intensity over the Atlantic Ocean due to a higher SST is partially compensated by a greater atmospheric stability. Moreover, greater atmospheric stability responds to the warm troposphere temperatures which are related to SST. The SST over the tropical Atlantic Ocean is remotely linked to NAO and ENSO (L. Wu et al., 2007). However, through the PNA pattern (Saravanan and Chang, 2000), the SST over the tropical Atlantic Ocean can be linked to ENSO solely, in particular, over the regions of central and eastern equatorial Pacific Ocean (Curtis and Hastenrath, 1995; Enfield and Mayer, 1997). Using AGCM, Saravanan and Chang (2000) concluded that the warm tropical Atlantic Ocean is related to warm SST over the eastern tropical Pacific Ocean. Hence, the intensity and frequency of hurricanes over the Atlantic Ocean are influenced by ENSO, which can increase tropospheric temperature and subsequently increase SST over the Atlantic Ocean.

The air temperatures in the U.S. and Canada are linked to both ENSO and PDO (E.M. Rasmusson and Carpenter, 1982; Mantua et al., 1997). Ropelewski and Halpert (1986) investigated ENSO effects on temperature and precipitation over North America. They found positive temperature anomalies over Alaska and western Canada from December to March in ENSO years, and negative temperature anomalies over southeastern U.S. from October to March. In addition, four regions; the mid Atlantic (MA), High Plains (HP), Great Basin (GB) and the Gulf and the Mexican Area (GM) present evidence of ENSO-related precipitation. For about 81% of ENSO events, GM (GB) experiences above-normal precipitation from October to March (April to October). On the other hand, the HP precipitation from April to October is inconsistent with ENSO events although it indicates a link with ENSO. Giannini et al. (2001) found remote effects of El Niño over the tropical Pacific Ocean in decreasing trends of precipitation over the tropical Americas via the direct atmospheric bridge and the delayed response of SST in the tropical North Atlantic Ocean.

In terms of the frequencies of anomalous weather events in the U.S., temperature and precipitation are correlated to SST and SLP (Gershunov, 1998) under the oceanic-atmospheric circulation (Hu and Feng, 2001). During El Niño winters, the extreme precipitation frequency (EPF) tends to increase in the coastal Southeast, parts of the Southwest and the central plains of the U.S. Moreover, the extreme warm frequency (EWF) shows a significant and consistent decreasing trend in southern and eastern U.S. A significant and inconsistent decreasing trend of the extreme cold frequency (ECF) is observed in northern and northwestern U.S. On the other hand, during La Niña, EWF is found in Texas, New Mexico and parts of the surrounding states, but weak increases of ECF have been observed in the Northwest and the northern Rockies.

To improve the seasonal snow forecasts in the Pacific Northwest and the Rockies of the U.S., Smith and O'Brien (2001) studied the composite snowfall quartiles associated with the warm, cold and neutral phases of ENSO during early, middle and late winters. The effect of ENSO on snowfall over the Pacific Northwest is found in early and mid-winter, but the influence over northern Rockies is observed only in mid-winter. More snowfall is expected in both regions during the cold phase rather than the neutral and warm phases of ENSO. In addition, the northern Great Lakes, southwest Montana and Wyoming experience ENSO-related snowfall during late winter. The northern Great Lakes indicate an increase in (decrease in) snowfall associated with a cold (warm) phase of ENSO. Surprisingly, both the warm and cold phases can decrease (increase) snowfall in southwest Montana (Wyoming).

In Latin America, Mendoza et al. (2005) reported that the El Niño causes anomalous droughts in central Mexico. The historical data from 1450 to 1900 are used to define the region-specific drought situation which occurred in 1483, 1533, 1571, 1601, 1650, 1691, 1730, 1818 and 1860. All the droughts in central Mexico are associated with strong and extremely strong El Niño. Furthermore, using the precipitation data obtained from approximately 1,000 stations distributed over Mexico, Pavia et al. (2006) showed that with the combination of PDO and El Niño, the wet condition in Mexico during summers (winters) is linked to low PDO (high PDO). In terms of ENSO and PDO effects on mean temperatures, the colder condition is associated with the summer La Niña and the winter El

Niño without PDO, but the warmer condition is related to the summer El Niño and high PDO.

In central Chile and central-western Argentina, the water supply management is dependent upon streamflow due to the snowpack in the central Andes. Masiokas et al. (2006) used the time series of snowpack from 1951 to 2005 and applied multiple regression to relate these with large-scale climate. More snow has a tendency to occur during El Niño years; however, the relationship between snowpack and atmospheric circulations is complicated and needs more analysis for better understanding and forecasting.

The observed discharge obtained from 35 different Colombian rivers is used to correlate with five ENSO indices – i.e. US Southern Oscillation Index (SOI), Multivariate Index (MEI), and tropical Pacific Ocean SST indices, e.g. NINO 1+2, NINO 3 and NINO 4. The below- (above-) normal discharges are associated with El Niño (La Niña). In addition, from the cross-correlation analysis, the discharges of Colombian rivers are significantly related to MEI, SOI and NINO 4 by lag times of 4-6 months from the indices (Gutiérrez and Dracup, 2001).

In Asia, the correlation between hydroclimates and large-scale atmospheric variables has been identified in several regions such as China (Gong and Wang, 1999; Y. Liu et al., 2002; Jiang et al., 2006; Q. Zhang et al., 2007), India (Whitaker et al., 2001; Fasullo and Webster, 2002; Nagura and Konda, 2007), Bangladesh (Whitaker et al., 2001), Nepal (M. L. Shrestha, 2000; A. Shrestha and Kostaschuk, 2005), and Sri Lanka (Zubair, 2003a, 2003b; Chandimala and Zubair, 2007). Significant parts of these relationships depend on the anomalous phases of atmospheric circulations. In China, decreasing precipitation is related to the warm phase of ENSO – i.e. El Niño (Gong and Wang, 1999). This study, based on statistical analysis, shows significant correlations between ENSO and rainfall in eastern China during winter and autumn.

In Nepal, Shrestha (2000) and Shrestha and Kostaschuk (2005) studied the influence of ENSO on rainfall and streamflow respectively. Using harmonic analysis, aggregate composite and index time series analysis, the above- and below-normal rainfall and streamflow are associated with ENSO indices, e.g. MEI, SOI and NINO 3.4. The regions and periods of significant ENSO effects are also identified. According to the conclusions from both studies, El Niño brings a stronger effect over Nepal than La Niña, and La Niña effects cover wider areas than El Niño elsewhere. The significant influences of El Niño (La Niña) on streamflow are observed in the regions where a weak (strong) monsoon occurs. The temporal consistency of El Niño is less than that of La Niña. Compared to normal conditions, the decreasing discharges in western (eastern) Nepal are related to El Niño from July (June) to December, and the increasing discharges in western Nepal are correlated to La Niña from June to January.

The Indian Ocean and the dipole pattern (i.e. the SST oscillation over western and eastern Indian Ocean) are affected by El Niño as an anomalous seasonal development of surface wind and SST over the equatorial eastern Indian Ocean (EEIO). When La Niña switches to the El Niño phase, the westerly zonal wind over the EEIO in spring changes towards the easterly direction. The anomalous wind produces negative SST anomalies over the EEIO during winter and over the eastern pole of the dipole pattern during fall. When the anomalous zonal wind delays until late summer or fall, negative SST anomalies over the EEIO are still observed during winter. However, SST anomalies over the eastern pole are not found in fall because of the decreasing amplitude of SST anomalies in summer (Nagura and Konda, 2007). Subsequently, based on the land-ocean temperature gradient (Goswami et al., 2006) and the dipole pattern of the Indian Ocean, the effects of El Niño on the Asian summer monsoon are to be expected. The interannual and interdecadal variability of the Asian summer monsoon is strongly related to the dipole pattern of the Indian Ocean, anomalous SST gradients between the eastern and central equatorial Pacific Ocean, and the zonal vertical integrated moisture transport (Fasullo and Webster, 2002).

2.2.3 Effects on basin hydroclimates

In the Colorado River Basin (CRB), using the point frequency analysis and the empirical orthogonal function (EOF), the extremely wet and dry conditions can be related to the variability of atmospheric variables over the Pacific Ocean during El Niño years. However, during La Niña years, only extremely dry conditions are observed (Cañón et al., 2007). In terms of spatial coverage, the anomalous weather events, which indicate a higher tendency during the warm than cold season, are found in a small area of the CRB. However, from year to year, the spatial coverage is inconsistent. In terms of a temporal scale, the precipitation shows a 1-season lag relationship with PDO-SOI. Due to disparities in station records (e.g. inhomogeneity and inconsistency of observations in space and time), the Standardized Precipitation Index (SPI), which is a method of nonparametric approach to define precipitation anomalies, has been adopted (Kim et al., 2006). In the upper basin of the CRB, increasing summer precipitation is linked to a low ENSO phase. On the other hand, decreasing winter precipitation in the lower basin is related to a high ENSO phase.

In Idaho, Harshburger et al. (2002) developed the relationship between hydroclimates (e.g. precipitation and streamflow) and standard ENSO indices (e.g. NINO 3, NINO 3.4, NINO 4, SOI and PDO) using cross-correlation analysis. The negative correlations between NINO 3.4 and winter precipitation and spring discharge are obtained. The precipitation and streamflow are increased due to La Niña-negative PDO and are decreased due to El Niño-positive PDO. However, the cold phase of ENSO more affects winter precipitation and spring streamflow than the warm phase, which indicates the asymmetrical responses mentioned earlier. In addition, topography may play a role in the relationship because significant correlations are obtained in mountainous areas. Yang et al. (2007) reported the opposite result in correlation between precipitation over the Great Plains in the U.S. and ENSO, showing a precipitation increase (decrease) during El Niño (La Niña). This is an instance of a positive correlation. The strongest relationship is found associated with SST anomalies over the tropical central-eastern Pacific Ocean (i.e. NINO 3.4) at one month lead time.

By combining two basins (i.e. Sacramento River (SAC) in California and Blue River (BLU) in western Colorado), the joint drought condition, which is defined as low streamflow in both rivers, presents inconsistent relationships with ENSO or PDO in the 20th century but shows significant relationships during some periods in the past (Meko and Woodhouse, 2005) as shown in Figure 2.8. This shows us nonlinear relationships.

The influence of ENSO on precipitation variability in the Amazon River Basin (Latin America) is relative to flood levels in the basin (Sombroek, 2001). Lower (higher) flood levels are observed during the warm (cold) phase of ENSO. Schöngart and Junk (2007) studied the flood-pulse of Central Amazonia related to SOI, NINO 1+2, NINO 3, NINO 3.4 and NINO 4. During El Niño years, the aquatic phase of streamflow is shortened by 44 days on an average due to lower maximum flood levels. In contrast, the aquatic phase during La Niña years is extended by 31 days because of higher flood levels.

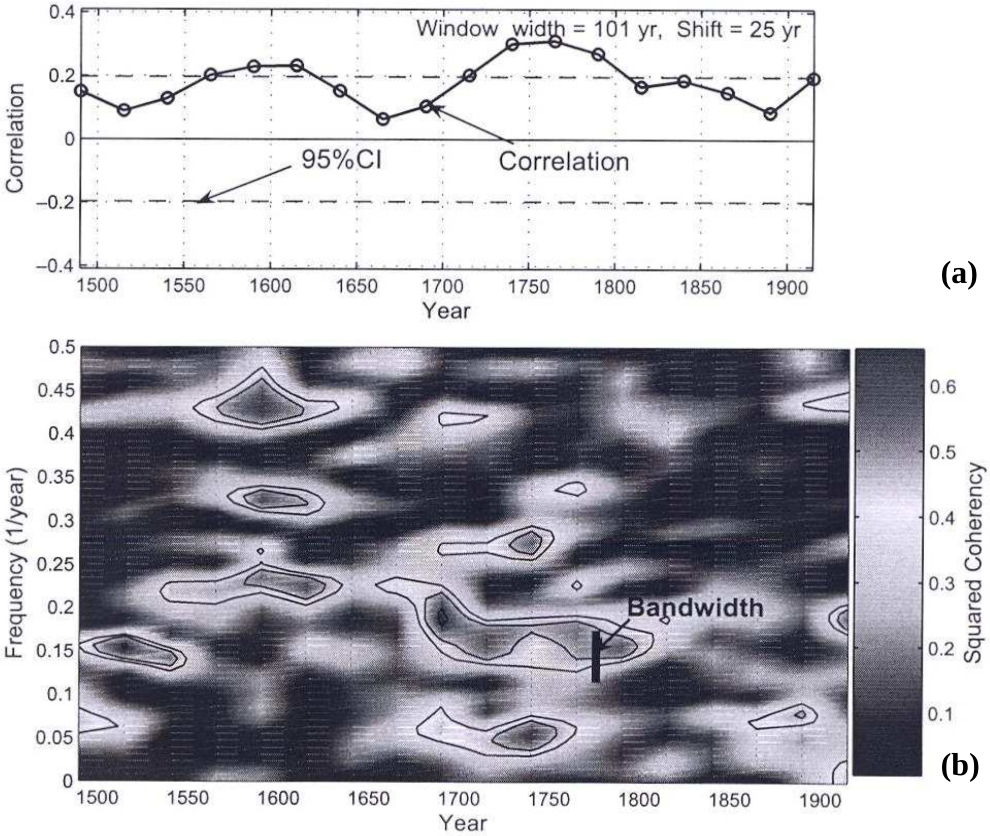


Figure 2.8: Correlations and squared coherency* of reconstructed SAC and BLU as a function of time; (a) the product-moment correlations for 101-year periods offset by 25-year periods; and (b) squared coherency (C^2) from cross-spectral analysis using the same setting for sliding windows.
*Squared coherency: correlation coefficient as a function of frequency.
Source: Meko and Woodhouse (2005)

On the other hand, the average streamflow of the Paraná River flowing through several countries in Latin America (Brazil, Paraguay and Uruguay) is higher due to El Niño than the streamflow due to La Niña (Berri et al., 2002). Furthermore, streamflow anomalies stay relevant to El Niño in the spring and autumn seasons of the following year (Camilloni and V.R., 2003). The discharge anomalies are positively correlated with spring SST anomalies over the eastern tropical Pacific Ocean and with autumn SST anomalies over the NINO 3 region. However, any significant correlation between discharge anomalies and SST during summer is hardly found.

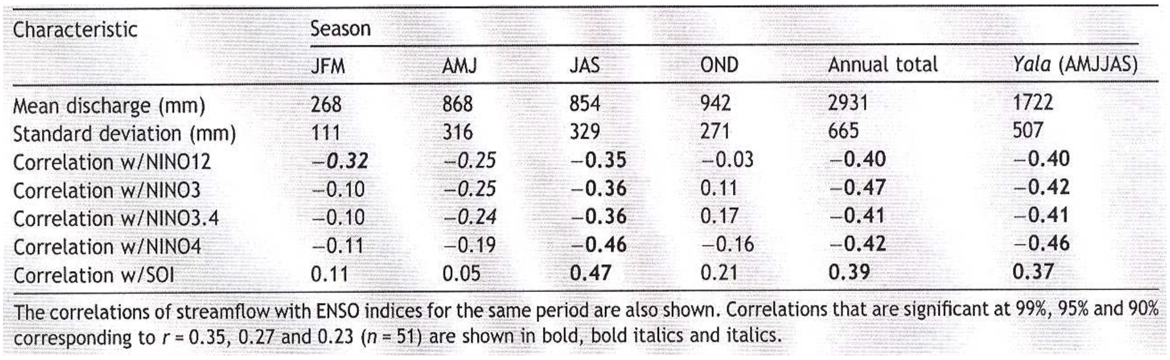
Using the correlation analysis, the streamflow of the Nile River, which is the most important river in Africa and the longest river in the world, is correlated with Guinea precipitation in the previous year and 1-year leading SST over a few regions in the Pacific Ocean. As a result, these two variables can be used as the best climatic predictors of streamflow (Eldaw et al., 2003).

In China, the streamflow of the two longest rivers (i.e. Yangtze River and Yellow River) is significantly related to ENSO. From the cross-wavelet analysis and wavelet coherence, the links between annual maximum streamflow in the Yangtze River and ENSO are dependent upon phases of relationships and locations of sub-basins. The in-phase and anti-phase relationships can be observed in the Lower and Upper Yangtze River Basins respectively;

however, an ambiguous relationship is seen in the Middle Yangtze River Basin (Jiang et al., 2006). Decreasing streamflow in the Upper Yellow River is associated with El Niño, and increasing streamflow corresponds to La Niña (Z. Liu and Yang, 2001). Hu et al. (2011) found the significant climate linkages between temperature and streamflow, and precipitation and streamflow in the source region of the Yellow River.

The rainfall over the Kelani River Basin (Sri Lanka) shows a non-linear relationship with ENSO. A stronger correlation is observed from October to December, and a weaker relationship is found in January and February (Table 2.1 and 2.2). Moreover, from April to September, the relationship between streamflow and ENSO is stronger than that of rainfall and ENSO (Chandimala and Zubair, 2007).

Table 2.1: Quarterly Characteristics of the Kelani Streamflow at Glencourse based on Records from 1950 to 2000

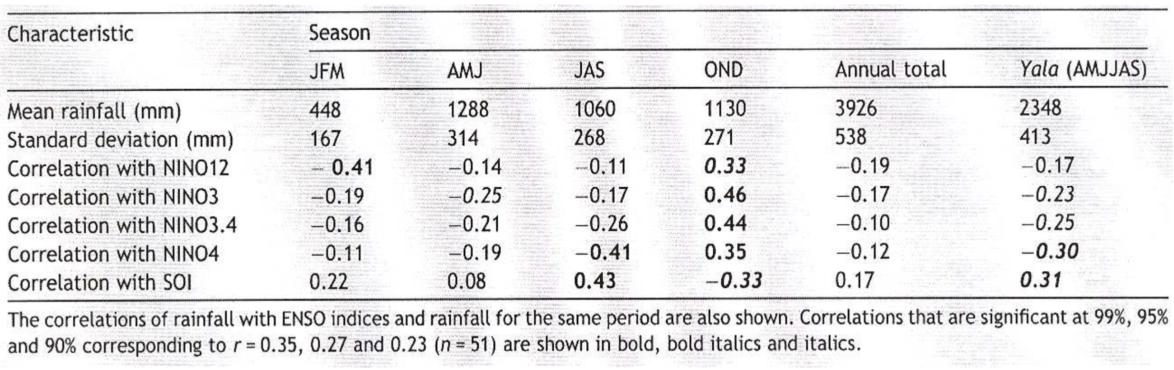


Characteristic	Season					
	JFM	AMJ	JAS	OND	Annual total	*Yala* (AMJJAS)
Mean discharge (mm)	268	868	854	942	2931	1722
Standard deviation (mm)	111	316	329	271	665	507
Correlation w/NINO12	***−0.32***	*−0.25*	**−0.35**	−0.03	**−0.40**	**−0.40**
Correlation w/NINO3	−0.10	*−0.25*	**−0.36**	0.11	**−0.47**	**−0.42**
Correlation w/NINO3.4	−0.10	*−0.24*	**−0.36**	0.17	**−0.41**	**−0.41**
Correlation w/NINO4	−0.11	−0.19	**−0.46**	−0.16	**−0.42**	**−0.46**
Correlation w/SOI	0.11	0.05	**0.47**	0.21	**0.39**	**0.37**

The correlations of streamflow with ENSO indices for the same period are also shown. Correlations that are significant at 99%, 95% and 90% corresponding to $r = 0.35$, 0.27 and 0.23 ($n = 51$) are shown in bold, bold italics and italics.

Source: Chandimala and Zubair (2007)

Table 2.2: Quarterly Characteristics of Rainfall in the Glencourse Catchment based on Averaged Rainfall Records for 10 Stations from 1950 to 2000



Characteristic	Season					
	JFM	AMJ	JAS	OND	Annual total	*Yala* (AMJJAS)
Mean rainfall (mm)	448	1288	1060	1130	3926	2348
Standard deviation (mm)	167	314	268	271	538	413
Correlation with NINO12	− **0.41**	−0.14	−0.11	***0.33***	−0.19	−0.17
Correlation with NINO3	−0.19	*−0.25*	−0.17	**0.46**	−0.17	*−0.23*
Correlation with NINO3.4	−0.16	−0.21	−0.26	**0.44**	−0.10	*−0.25*
Correlation with NINO4	−0.11	−0.19	**−0.41**	**0.35**	−0.12	***−0.30***
Correlation with SOI	0.22	0.08	**0.43**	***−0.33***	0.17	***0.31***

The correlations of rainfall with ENSO indices and rainfall for the same period are also shown. Correlations that are significant at 99%, 95% and 90% corresponding to $r = 0.35$, 0.27 and 0.23 ($n = 51$) are shown in bold, bold italics and italics.

Source: Chandimala and Zubair (2007)

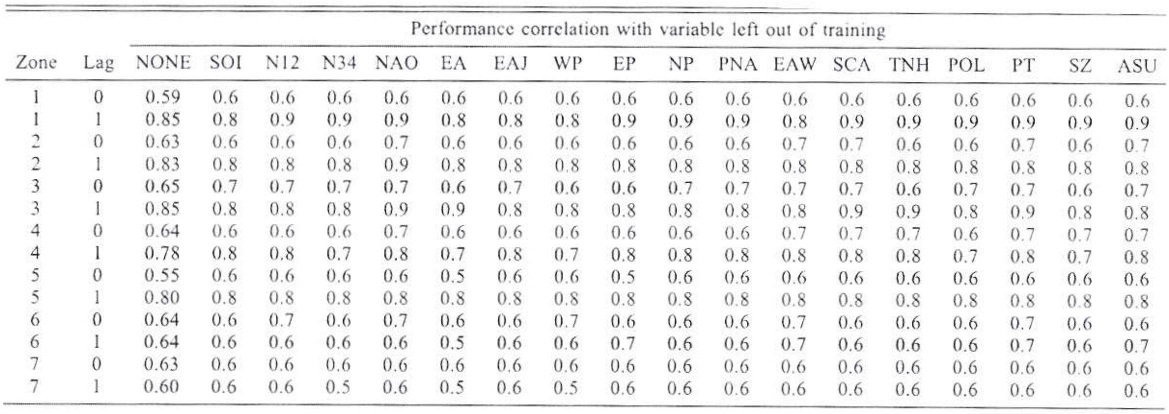
In conclusion, the variability of local hydroclimates such as surface temperature, precipitation and streamflow are linked with anomalies in large-scale atmospheric variables. Via oceanic-atmospheric circulations, anomalous conditions differently affect local hydroclimates. Some regions witness below-normal precipitation under El Niño and vice versa under La Niña. Anomalous weather events (flood and drought) in several basins are also affected by ENSO. The developed relationships between large-scale atmospheric variables with leading time and local hydroclimates present the leading predictability of a model. The long-range leading forecasts of hydroclimates are a useful tool to manage and plan water resources if an anomalous event were to occur.

2.3 Forecasting models with the cooperation of identified atmospheric variables

To apply the development of relationships between large-scale atmospheric variables and hydroclimates to forecasting models, large-scale atmospheric variables are identified as the predictors of these models. Forecasting models, especially those based on the probabilistic (stochastic) approach, along with the incorporation of large-scale atmospheric predictors (Basson and Rooyen, 2001; Krzysztofowicz, 2001; Trenberth et al., 2006) show significant performance and give reliable forecasts. Precipitation and streamflow forecasts can be used in reservoir management (Hamlet et al., 2002; Steinemann, 2006), agricultural schedules and mitigation plans for anomalous weather events.

Gershunov (1998) applied the statistical approach to 60-year daily data obtained from 168 stations in the US to predict the frequencies of daily anomalous temperature and precipitation. The model performance depends on the season of ENSO and geographic locations of weather stations. From the predictions of anomalous precipitation frequency, a significant performance of the developed model is obtained along the Gulf Coast, the central plains, the Southwest and in the Ohio River Valley under the winter El Niño. The predictability of extreme warm frequency is associated with the southern and eastern regions also under the winter El Niño. Furthermore, artificial neural networks (ANNs) using standard indices of anomalous atmospheric variables (Table 2.3) such as SOI, NAO and PNA at several leading periods is adopted to forecast California precipitation (Silverman and Dracup, 2000). The performance of ANNs is significantly improved compared to the results at no lead.

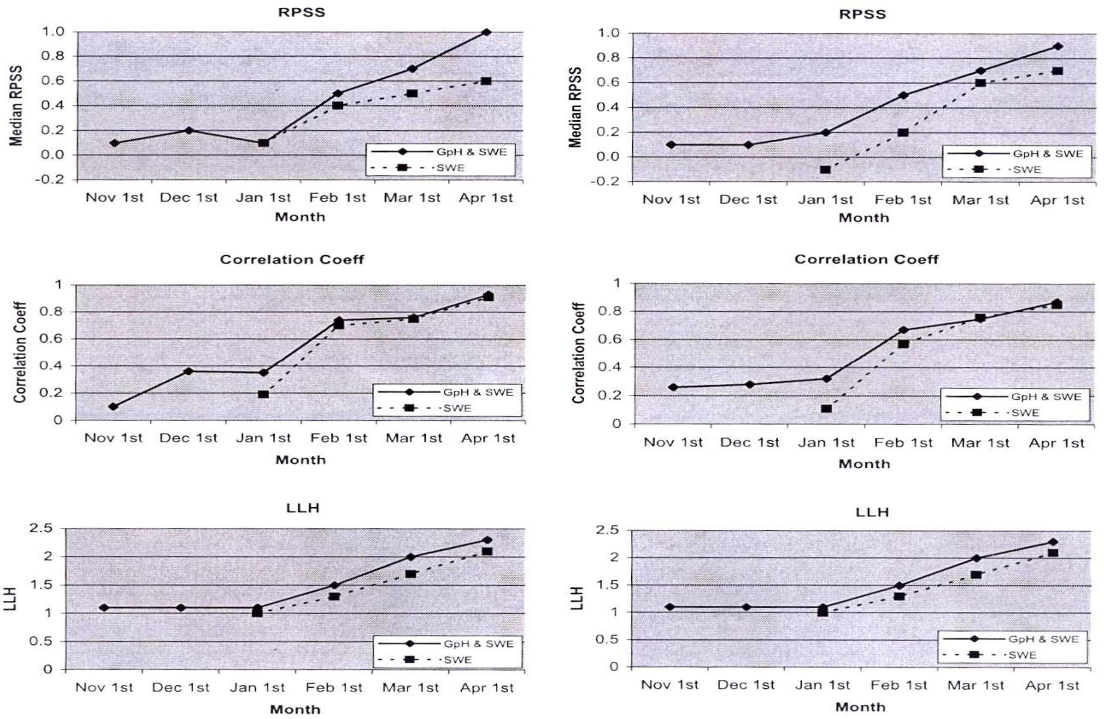
Table 2.3: Correlation between Observed and Predicted Precipitations with the Listed Variable Removed from the Training Set



		Performance correlation with variable left out of training																	
Zone	Lag	NONE	SOI	N12	N34	NAO	EA	EAJ	WP	EP	NP	PNA	EAW	SCA	TNH	POL	PT	SZ	ASU
1	0	0.59	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6
1	1	0.85	0.8	0.9	0.9	0.9	0.8	0.8	0.8	0.9	0.9	0.9	0.8	0.9	0.9	0.9	0.9	0.9	0.9
2	0	0.63	0.6	0.6	0.6	0.7	0.6	0.6	0.6	0.6	0.6	0.6	0.7	0.7	0.6	0.6	0.7	0.6	0.7
2	1	0.83	0.8	0.8	0.8	0.9	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8
3	0	0.65	0.7	0.7	0.7	0.7	0.6	0.7	0.6	0.6	0.7	0.7	0.7	0.7	0.6	0.7	0.7	0.6	0.7
3	1	0.85	0.8	0.8	0.8	0.9	0.9	0.8	0.8	0.8	0.8	0.8	0.8	0.9	0.9	0.8	0.9	0.8	0.8
4	0	0.64	0.6	0.6	0.6	0.7	0.6	0.6	0.6	0.6	0.6	0.6	0.7	0.7	0.7	0.6	0.7	0.7	0.7
4	1	0.78	0.8	0.8	0.7	0.8	0.7	0.8	0.7	0.8	0.8	0.8	0.8	0.8	0.8	0.7	0.8	0.7	0.8
5	0	0.55	0.6	0.6	0.6	0.6	0.5	0.6	0.6	0.5	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6
5	1	0.80	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8
6	0	0.64	0.6	0.7	0.6	0.7	0.6	0.6	0.7	0.6	0.6	0.6	0.7	0.6	0.6	0.6	0.7	0.6	0.6
6	1	0.64	0.6	0.6	0.6	0.6	0.5	0.6	0.6	0.7	0.6	0.6	0.7	0.6	0.6	0.6	0.7	0.6	0.7
7	0	0.63	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6
7	1	0.60	0.6	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6

Source: Silverman and Dracup (2000)

The performance of streamflow forecasting models also shows significant improvement when large-scale atmospheric predictors are used (Clark et al., 2001; Hamlet et al., 2002; McCabe and Dettinger, 2002; Wood et al., 2002). Grantz et al. (2005) identified the predictors of seasonal streamflow in the Truckee-Carson River (the U.S.) and applied them in a local regression model. The predictors identified by correlation maps are the snow water equivalent (SWE), 500-millibar (mb) geopotential height index and SST. The model performance increases corresponding to decreasing forecasting lead times; however, the model performance is significantly improved by up to a 5-month forecasting lead time (Figure 2.9).



(a) **(b)**

Figure 2.9: Model performance scores of the forecasts issued on the first of each month from November to April for (a) the Truckee; and (b) the Carson Rivers.
Source: Grantz et al. (2005)

The streamflow of various rivers in Latin America is linked to large-scale atmospheric variables. The streamflow of the Cearra River (Brazil) develops a significant link with the Atlantic and Pacific Oceans (Sourza Filho and Lall, 2003). The streamflow of the Columbian River is correlated to MEI, SOI and NINO 4 at lead times of 4-6 months (Gutiérrez and Dracup, 2001). The flood-pulse of the Negro River (central Amazonia) is related to the February SOI and NINO 3.4 (Figure 2.10), which is a 4-month lead time from the occurrence of maximum flood levels in June (Schöngart and Junk, 2007). The lead time of identified predictors is a benefit for forecasting models (like the multiple regression model) aiming to forecast the flood-pulse of the Negro River (Figure 2.11). Simulated flood levels can be used to plan the annual water uses of this basin for fishery, agriculture and timber extraction.

In Asia, Chandimala and Zubair (2007) used the March SST over the Indo-Pacific Ocean as the predictor of a principle component regression to forecast the streamflow of the Kelani River (Sri Lanka). Using a cross-validated method, a correlation coefficient between streamflow forecasts and observed data was estimated to be 0.5 after 1960. The results suggest that the streamflow should be divided into two halves of the season because of the nonlinear relationships between streamflow and large-scale atmospheric predictors.

Consequently, using the identified large-scale atmospheric predictors, the forecasting models based on various approaches (e.g. the statistical model, artificial neural networks and local regression) present significant predictability of hydroclimates, particularly precipitation and streamflow. The model performance depends upon significant levels of developed relationships between hydroclimates and identified predictors, geographical locations of hydroclimate stations and forecasting lead time.

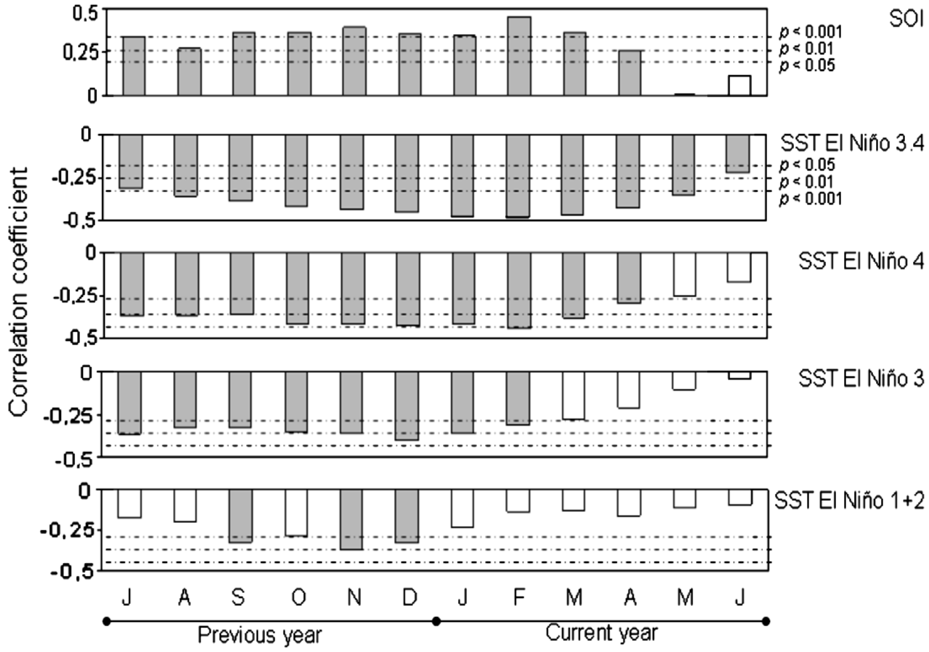


Figure 2.10: Correlation between the maximum water level of the Negro River and ENSO indices for a 12-month period before the occurrence of the maximum flood level in June.
Source: Schöngart and Junk (2007)

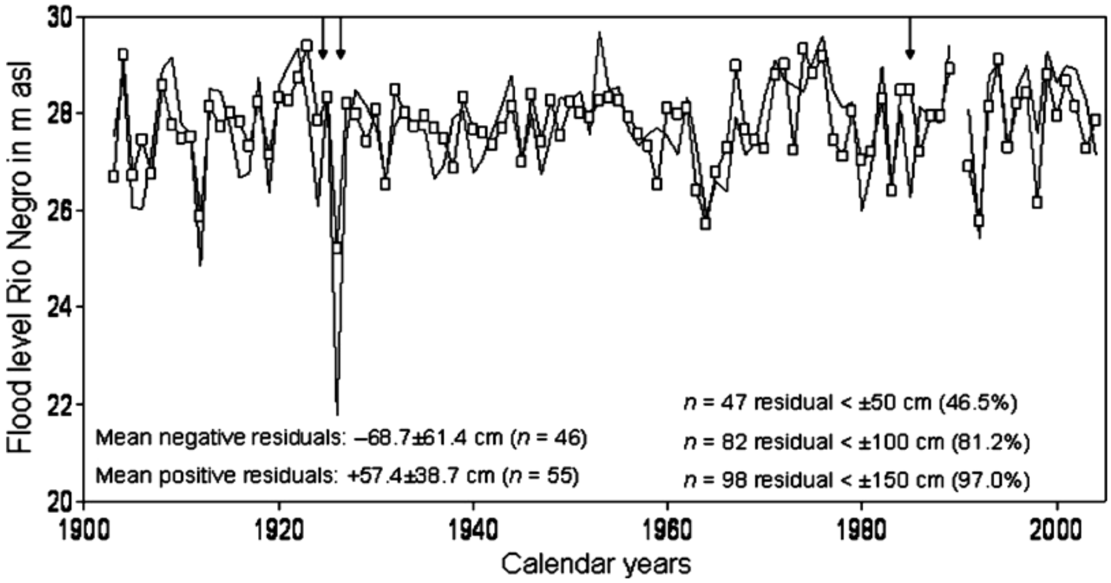


Figure 2.11: Comparison between the observed (black line) and predicted (dotted line) maximum flood level of the Negro River.
Source: Schöngart and Junk (2007)

2.4 Studies of the effects of atmospheric teleconnection on Thailand hydroclimates

The variability of hydroclimates over Asian regions has been investigated by many (Gutman et al., 2000; Ho and Wang, 2002; Meehl and Arblaster, 2002; Fasullo and Webster, 2003; Krishnan et al., 2003; Nodzu et al., 2006; Ogino et al., 2006; Krishnamurthy and Shukla, 2007). The possible forcing factors of hydroclimate variability include internal variables like land use changes (Croke et al., 2004) and deforestations (Kanae et al., 2001), and external variables like atmospheric circulation (Chu et al., 1999; Z. Liu and Yang, 2001). To determine hydroclimate variability, several studies cover regions of interest over the Indian subcontinent, South Asia, Indochina and Southeast Asia (Y. Liu et al., 2002; Meehl and Arblaster, 2002; Yaremchuk and Qu, 2004), and Thailand is a part of the study areas. The variability of Asian hydroclimates shows links with anomalous atmospheric variables over the Indian and Pacific Oceans, and it is also related to the Indian monsoons, and the Asian and Pacific summer monsoons.

Chang et al. (2005) reported that the annual cycle of Southeast Asian climate is dominated by interactions between a simple annual circulation of the surface monsoonal winds passing from the Indian Ocean to the South China Sea and the equatorial western Pacific Ocean, and the complex land terrain, and seascape.

Krishnan et al. (2000) noted that a weakening convection over the Bay of Bengal, the eastern Indian Ocean, Indonesia, Southeast Asia and the equatorial western Pacific Ocean is influenced by the non-convective anomalies over the equatorial Indian Ocean.

The climate of Indochina (i.e. Cambodia, Laos, Vietnam and Thailand), which is dependent upon the Indian and southeast-east Asian monsoon components, shows a significant relation between interannual variability of monsoon rainfall and NINO 3 (Chen and Yoon, 2000). The monsoon rainfall over Indochina is above (below) normal under the cold (warm) phase of the eastern tropical Pacific Ocean. In addition, the results suggest large-scale atmospheric variables such as NINO 3, outgoing longwave radiation (OLR) and velocity potential can be used for short-term forecasting of hydroclimates over the Indochina region.

Arnell (1999) adopted the Hadley Centre climate models (HadCM2 and HadCM3) and a macro-scale hydrological model to simulate the streamflow of global rivers. By 2050, the average annual runoff in high latitude regions, equatorial Africa and Asia, and Southeast Asia including Thailand will increase, and the runoff will decrease in the mid-latitudes and subtropical regions. Also, the increasing trends in global temperature will decrease the intensity of snowfall and the duration of snow cover in several areas. Shifts in the wet season of streamflow due to snow melt will also be observed.

Due to the geographical location and physical mechanisms of climate, Thailand hydroclimates respond to the anomalous conditions over the Indian and Pacific Oceans. The interannual and interdecadal variability of Thailand summer monsoon rainfall during the recent decades (post-1980) points to a significant correlation with standard anomalous indices (SOI and ENSO) and atmospheric variables over both oceans (Singhrattna et al., 2005b). The relationship is largely dependent upon the phases and regions of ENSO. The below- (above-) normal monsoon rainfall is related to El Niño (La Niña) over the eastern equatorial Pacific Ocean.

Further, the climatologic monsoon break (CMB) over Thailand in late June is associated with a drastic change of large-scale monsoon circulation in the seasonal duration. The Southeast Asian summer monsoon can be divided into two periods – the early and later monsoon (Takahashi and Yasunari, 2006). From the wavelet analysis using daily rainfall in 1998, the duration of the early monsoon indicates a range of 30 to 60 days, and the duration of later monsoon ranges from 10 to 20 days. The later monsoon season is associated with the horizontal structures of atmospheric circulation called the Rossby wave and vertical structures in the troposphere (Yokoi and Satomura, 2005). The atmospheric variables (e.g. surface temperature and pressure) are identified as the potential predictors of a forecasting model for Thailand summer monsoon rainfall. The incorporation of large-scale atmospheric predictors in a statistical stochastic model shows a high level performance for a forecasting lead times of 2-5 months (Singhrattna et al., 2005a).

Hence, Thailand hydroclimates have significant correlations with large-scale atmospheric variables. The monsoon rainfall tends to decrease under the warm phase of ENSO and vice versa for the cold phase. Using the identified large-scale atmospheric predictors, the forecasting models of Thailand monsoon rainfall give a reliable performance depending on the forecasting lead time.

2.5 Summary

The significant relationship between large-scale atmospheric variables and local hydroclimates based on the global, regional and basin scales have been examined in this chapter. The anomalous conditions of oceanic-atmospheric circulations are responsible for hydroclimate variability and anomalous weather events. Due to spatial and temporal coverage, the asymmetrical and inconsistent responses of hydroclimates can be observed under different phases, i.e. warm and cold phases of the anomalies.

The interannual and interdecadal variability of Thailand summer monsoon rainfall is dependent on the anomalous conditions over the Indian and Pacific Oceans. The warm phase of ENSO is associated with below-normal monsoon rainfall and vice versa for the cold phase. For Thailand summer monsoon rainfall, several atmospheric variables such as NINO 3, OLR, velocity potential and SST over the eastern equatorial Pacific Ocean have been identified as potential predictors for forecasting models.

Using identified atmospheric predictors, various approaches (e.g. statistical models, ANNs and multiple regressions) give a good performance in hydroclimate forecasting. Although the models are successful for short forecasting lead times (e.g. up to 4-5 months prior to the monsoon season), the long leading predictability by a forecasting model remains a gap. Therefore, subsequent steps of this study are (i) to detect a significant relationship between large-scale atmospheric variables and local hydroclimates like rainfall; (ii) to identify large-scale atmospheric predictors; and (iii) to develop a forecasting model using the identified predictors to forecast and downscale local hydroclimates from large-scale atmospheric variables for long lead times, i.e. more than 6 months prior to the start of the season.

Chapter 3
Study Area Description and Climate Diagnostics

3.1 Introduction

Thailand is located between 5°-20°N latitude and 97°-106°E longitude, and between the water bodies of the Indian Ocean and the Gulf of Thailand, which is connected to the Pacific Ocean (Figure 3.1). Thailand covers an area of 513,115 km^2 with a population of 62.4 million people in 2005 (NSO, 2005). The major occupation in the country is agriculture, which accounts for 50-60% of the national economy. The natural water supply for agriculture is either direct rainfall or irrigation by water stored in reservoirs, which are also dependent on rainfall. Rainfall occurs during the annual monsoon season, from August to October, with the average annual rainfall in the country ranging from 1,200 to 1,600 mm per year. In terms of climate, Thailand is influenced by the Indian and Pacific Oceans in the form of land-ocean circulation. The summer season, lasting from mid-February to mid-May, is responsible for developing the land-ocean temperature gradient that strengthens the Southwest monsoon from the Indian Ocean in the rainy season to follow. The rainy season, from mid-May to mid-October, is caused by the Inter Tropical Convergence Zone (ITCZ) and by the Southwest monsoon, which force heavy rainfall to occur in Thailand from August to October. During the winter season, Thailand gets dry and cool winds brought on by the Northeast monsoon from the mid-latitudes, between mid-October and mid-February. In terms of streamflow, Thailand receives about 289,000 million m^3 (MCM) year^{-1} of the average runoff from 25 major river basins. However, the capacity of reservoirs in the country to store and supply water for various uses is estimated at 38,000 MCM year^{-1} or 13.1% only of the annual runoff.

Figure 3.1: The Kingdom of Thailand.

The Chao Phraya River Basin, the largest basin among the 25 major basins in the country, covers an area of 178,000 km^2 or 35% of the country land area (Figure 3.2). Four major tributaries – i.e. the Ping, Wang, Yom and Nan Rivers merge at Nakhon Sawan and form the Chao Phraya River. The Upper Chao Phraya River Basin is the portion above the confluence at Nakhon Sawan, whereas the portion below the confluence is called the Lower Chao Phraya River Basin. The upper basin covers an area of 102,635 km^2 or 58% of the watershed area, whereas the lower basin covers the capital (i.e. Bangkok), and government, business and agricultural areas. The average area irrigated by the Chao Phraya River Basin is 6,878 km^2. The domestic and total water demands are estimated by 2,240 and 11,000 MCM year^{-1}, respectively.

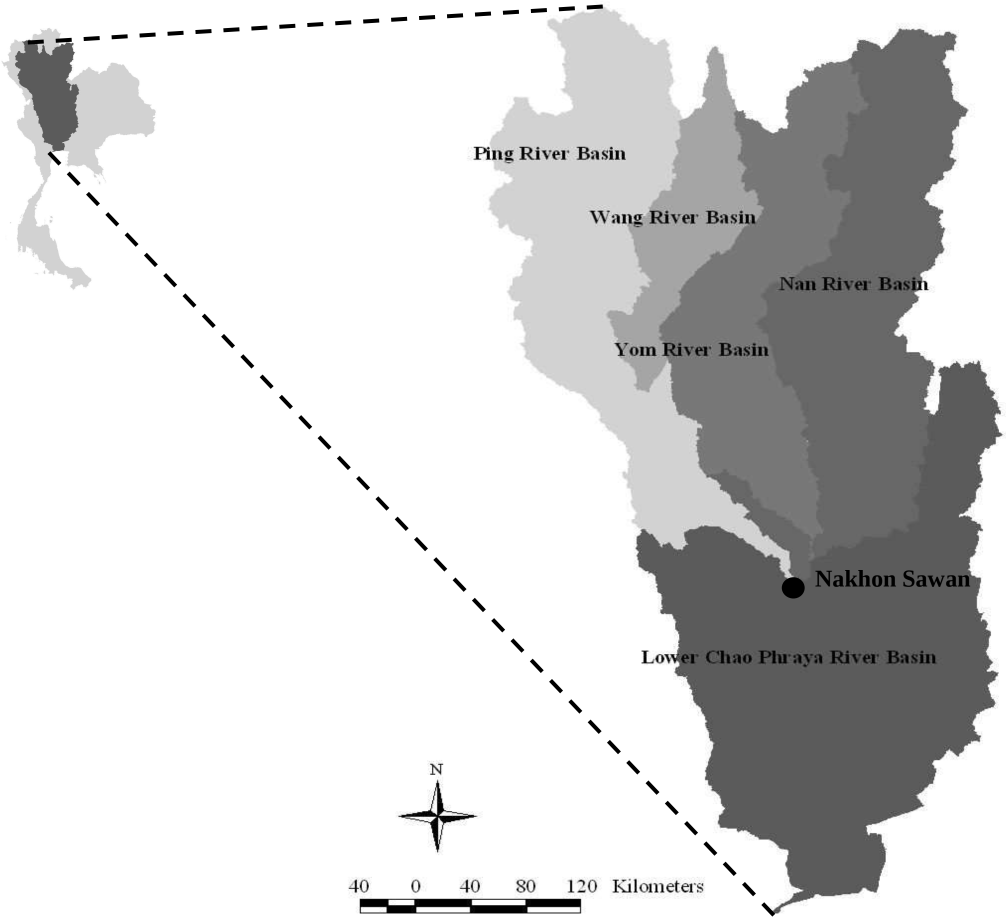


Figure 3.2: The Chao Phraya River Basin.

The Ping River Basin, located in northern Thailand, lies between 15°-19°N latitude and 98°-100°E longitude and covers an area of 33,899 km^2 in five provinces (Chiang Mai, Lam Phun, Tak, Kamphaeng Phet and Nakhon Sawan). The river is 740 km long and originates from Pee Pan Nam mountain range in Chiang Dao district (Chiang Mai). The streamflow flows to the south, passing Chiang Mai to Lam Phun, and then to the Bhumipol dam in the Sam Ngao district (Tak). The Lower Ping River, downstream of the Bhumipol dam, flows to the plain areas in Tak and merges with the Wang River. Subsequently, the Ping passes Kamphaeng Phet before joining the Yom and Nan Rivers to form the Chao Phraya River at the Pak Nam Pho district (Nakhon Sawan). The elevation of the Ping River Basin is shown in Figure 3.3.

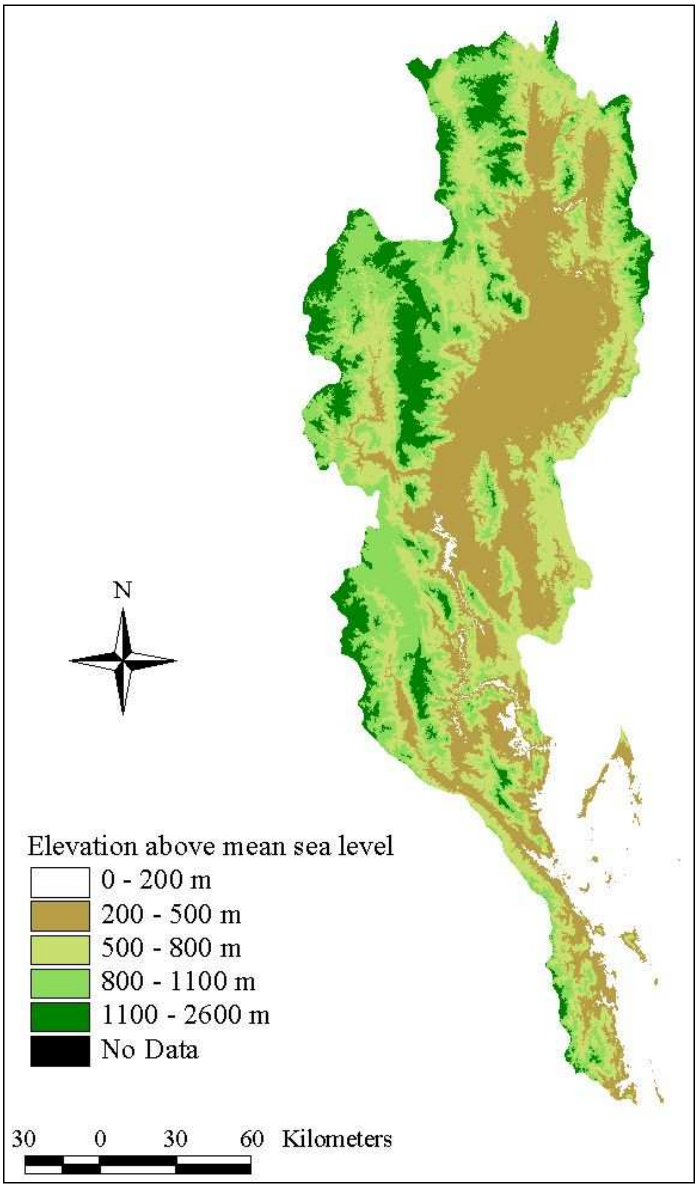


Figure 3.3: Topographic map of the Ping River Basin.

The five main tributaries of the Ping River are Mae Ngad, Mae Taeng, Mae Kuang, Mae Li and Mae Cham. The description of the tributaries is as follows:

1) The Mae Ngad originates from the Dan Laos mountain range and joins the Ping River in the Mae Taeng district (Chiang Mai). The Mae Ngad Storage Dam (also called the Mae Ngad Somboonchol Dam), located on the Mae Ngad, was initiated in 1977 by the Royal Irrigation Department (RID) of Thailand and started to operate in 1984. The dam is 59 m high and 1,950 m long with a maximum storage capacity of 265 MCM. In 1985, a hydropower plant with an annual generation capacity of 19 million kWh was constructed here by the Electricity Generating Authority of Thailand (EGAT).
2) The Mae Taeng also originates from the Dan Laos mountain range and joins the Ping River at the Mae Taeng district (Chiang Mai). The total length of the Mae Taeng is 154 km.
3) The Mae Kuang also originates from the Dan Laos mountain range and joins the Ping River in Lam Phun.

4) The Mae Li flows from the Li district (Chiang Mai) up north and joins the Ping River in the Chom Thong district (Chiang Mai).
5) The Mae Cham originates from the Thanon Thongchai mountain range in the northwest of the Ping River and joins the Ping in the Hot district (Chiang Mai).

The Ping River Basin receives an annual runoff of 8,700 MCM to serve an annual water demand of 6,127 MCM. The Ping River Basin can be divided into 20 sub-basins (Figure 3.4). The area coverage of each sub-basin is shown in Table 3.1. There are many storage dams located on the Ping River and its tributaries. The most important dam, the Bhumipol dam was constructed for the multiple purposes of generating hydropower, irrigating agriculture, fishery, water transportation and flood mitigation. The dam height is 154 m, with a length of 486 m. The maximum storage capacity is 13,462 MCM, and it receives an annual average inflow of 5,900 MCM from the Ping River Basin. Its hydropower capacity is about 780 MW. In terms of land use, 71.46% of the basin area is covered by forests located in the upstream where the Ping River originates. The remaining area is water sources and the plain areas on both banks of the Ping River and the flood plain downstream, which is covered by agricultural and residential users. The irrigated area of the Ping River Basin is estimated to be 2,332 km^2.

3.2 Data collection

The daily data of hydroclimates (like rainfall, streamflow, temperature and evaporation) used in this study are provided by the Royal Irrigation Department (RID) of Thailand, the Thailand Meteorology Department (TMD), the Department of Water Resources (DWR), and the Electricity Generating Authority of Thailand (EGAT). The list of 208 rainfall stations located in and around the Ping River Basin is presented in Appendix A1. Figure 3.5(a) shows the locations of the rainfall stations. Out of the 208 rainfall stations, the 50 stations that have been selected for the study are the ones that have data of more than 30 recent years with less than 5% incomplete data. The list of the 50 selected rainfall stations is presented in Appendix A2. The time series of these 50 stations range from 31 to 86 years.

The streamflow data are obtained from 45 gauging stations, which are operated by the RID, DWR and EGAT, and are listed in Appendix A3. The 12 streamflow stations selected from 45 stations are the ones that have no incomplete data during a consecutive period which is consistent among all the selected stations. The consecutive period which has no incomplete data in the 12 selected stations runs from April 1999 to March 2007 (i.e. eight years). The list of these 12 selected stations is shown in Appendix A4, and the locations are presented in Figure 3.5(b).

For temperature and evaporation, the daily data are obtained from 11 and 18 stations respectively (Appendix A5), which are located in and around the Ping River Basin (Figure 3.5(c)). The stations are operated by the RID and TMD. The data length of daily temperature (evaporation) varies from 16 to 57 years (12 to 38 years).

3.3 Climate diagnostics

Since the cross-correlations of monthly temperature (rainfall) from 11 (50) selected stations are significant at a 95% confidence level by Fisher’s Transformation (Haan, 2002), the averages over the selected stations have been estimated fairly accurately. Appendix B1 (B2) shows the average monthly and annual temperature (rainfall) from 1951 (1950) to 2007.

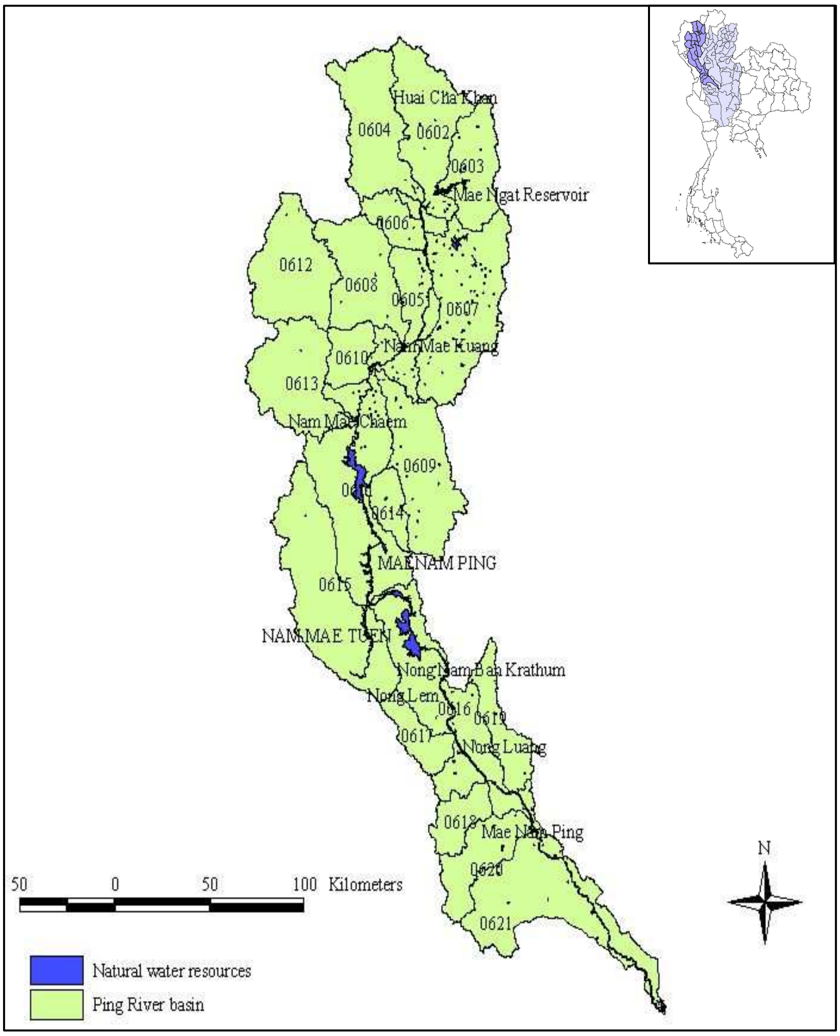


Figure 3.4: Sub-basins of the Ping River Basin.

Table 3.1: Descriptions of the selected 20 Sub-basins of the Ping River Basin

Sub-basin code	Sub-basin name	Basin area (km^2)	Sub-basin code	Sub-basin name	Basin area (km^2)
0602	Upper Ping Part	2,018	0612	Upper Mae Cham	1,912
0603	Mae Ngad	1,260	0613	Lower Mae Cham	1,926
0604	Mae Taeng	1,761	0614	Mae Had	535
0605	2nd Ping Part	1,624	0615	Mae Tuen	3,143
0606	Mae Rim	584	0616	4th Ping Part	2,940
0607	Mae Kuang	1,165	0617	Huai Mae Toa	2,151
0608	Mae Ngan	1,711	0618	Klong Wang Chao	647
0609	Mae Li	1,956	0619	Klong Mae Ra Ka	882
0610	Mae Klang	600	0620	Klong Suan Mak	1,069
0611	3rd Ping Part	3,071	0621	Lower Ping Part	2,944

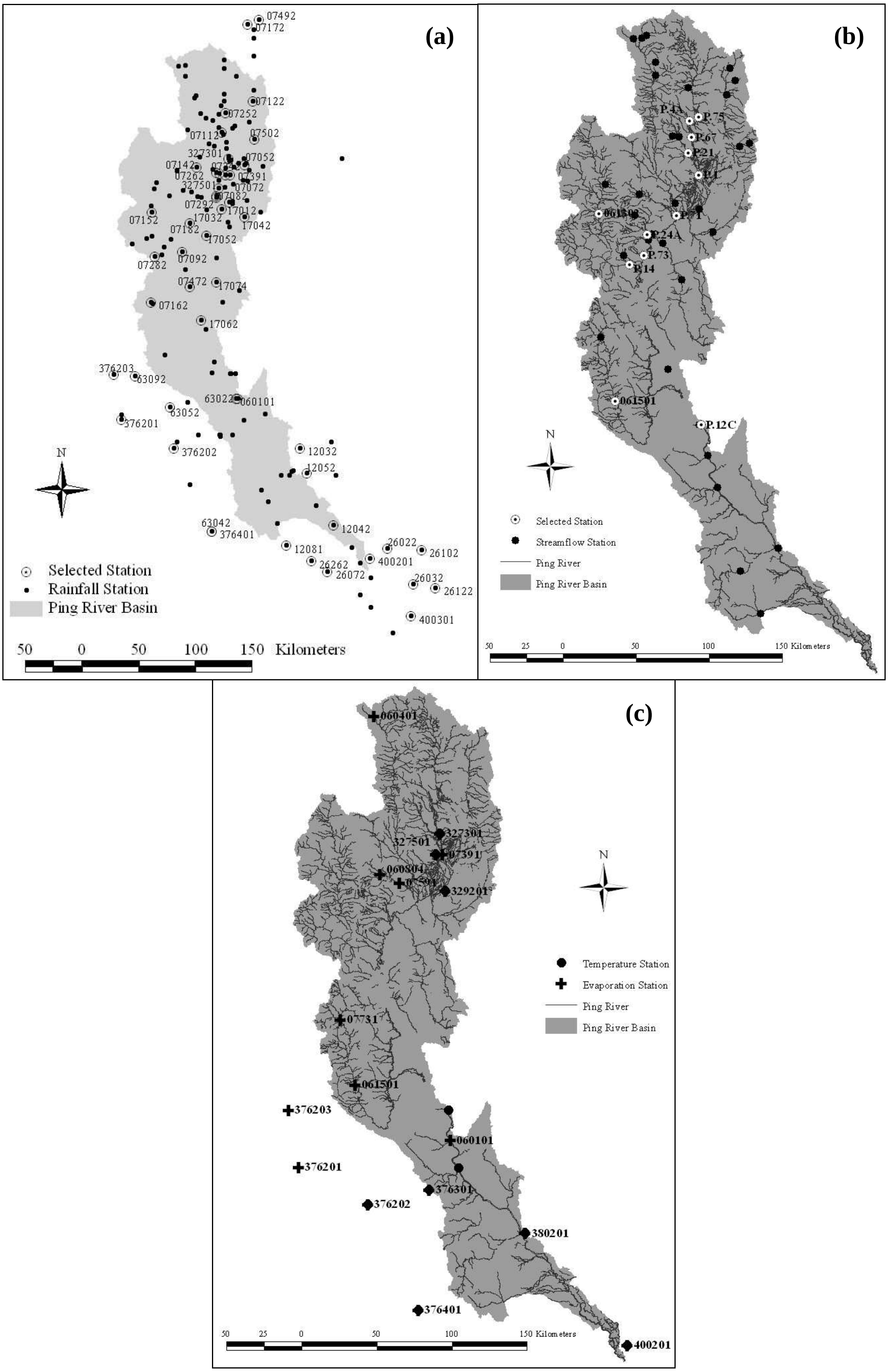


Figure 3.5: Locations of (a) rainfall stations; (b) streamflow gauging stations; and (c) temperature and evaporation stations.

3.3.1 Interannual variability

The annual cycles of air temperature and rainfall of the Ping River Basin are shown in Figure 3.6. The summer season occurs during March-April-May (MAM) with a maximum temperature of 30.2°C in April. The minimum temperature, 22.9°C, occurs in December. From 1951 to 2007, the maximum MAM temperature of 30.7°C was observed in 1958, and the minimum of 28.0°C was observed in 2000 and 2007 (see also Appendix B1).

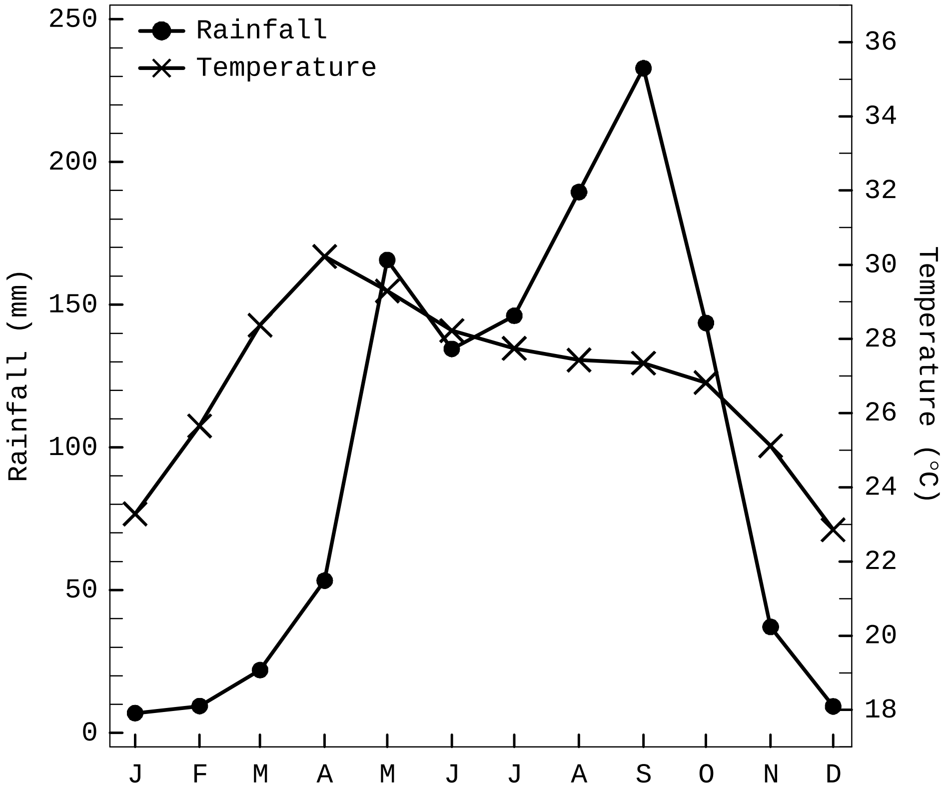


Figure 3.6: The annual cycle of temperature and rainfall.

As for the annual cycle of rainfall, Figure 3.6 shows the bi-modal regime with two peaks: one in May and another in September. The primary peak occurs during August-September-October (ASO), which is the monsoon or rainy season in Thailand. The secondary peak occurs during May-June-July (MJJ), which is the pre-monsoon season or the transition period from the summer to the rainy season. Two peaks of the annual cycle are caused by ITCZ and the Southwest monsoon. The secondary peak corresponds with ITCZ and the Southwest monsoon moving from the Indian Ocean to Thailand in May and passing to the South China Sea and central China in mid-June. The primary peak is associated with ITCZ as it moves back to cover Thailand during ASO. From the rainfall data of 58 years (1950-2007), the maximum MJJ and ASO rainfall have been observed in 1950, at 651.8 and 948.3 mm respectively. The minimum MJJ (ASO) rainfall at 254.0 (387.7) mm was recorded in 1997 (2004). The total annual rainfall varies from 843.0 to 1,605.6 mm per year (see also Appendix B2 and B3). The pre-monsoon (MJJ) and monsoon (ASO) seasons have about 88% of the total annual rainfall. The remaining, 12%, is the dry season rainfall, which falls from November to April of the following year and ranges from 49.4 to 295.2 mm.

3.3.2 Interseason variability

Air temperature is related to rainfall in terms of a developing land-sea temperature gradient which subsequently strengthens a monsoon. In this study, the summer season temperature is averaged over MAM (Appendix B1). The MAM temperature develops in an inverse relationship with the dry season (November to April) rainfall (Figure 3.7). The more (less) the dry season rainfall; the cooler (warmer) the land and atmosphere, and this decreases (increases) MAM air temperature. The increment of the dry season rainfall by 100 mm decreases MAM air temperature by 0.1°C. The MAM temperature subsequently strengthens or weakens a monsoon due to the development of the land-sea temperature gradient.

Figure 3.8(a) and (b) show the relationship between MAM temperature and MJJ rainfall, and MAM temperature and ASO rainfall, respectively. As expected, the temperature gradient is not fully developed during the pre-monsoon season (MJJ). As a result, an inverse relationship between MAM temperature and MJJ rainfall can be observed. Increasing (decreasing) MAM temperature brings about less (more) MJJ rainfall. On the other hand, when the temperature gradient is fully developed during the monsoon season (ASO), the increasing MAM temperature can strengthen a monsoon and bring more ASO rainfall to the study basin. Increasing MAM temperature by 1.0°C affects the MJJ and ASO rainfalls by -16.1 and +18.3 mm respectively. The relationship between MAM temperature and rainfall during MJJ and ASO is confirmed by the correlation maps (Figure 3.9). The MJJ and ASO rainfalls are correlated with MAM air temperature over the study basin with a significance of 95% confidence level (i.e. the upper and lower bounds of the significant correlations are +0.3 and -0.3 respectively) which is consistent with the developed relationship shown in Figure 3.8. Under the full development of the temperature gradient, the positive (negative) relationship between ASO rainfall and surface land (sea) MAM temperature is stronger than that between MJJ rainfall and surface land (sea) MAM temperature. Furthermore, during the monsoon season, a significant relationship between rainfall and MAM temperature over the South China Sea can be derived.

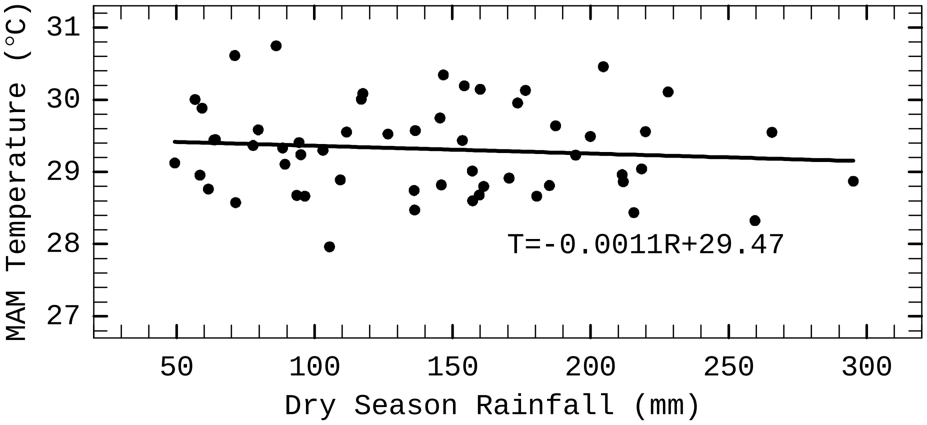


Figure 3.7: Scatter plots between dry season (November to April) rainfall (R) and MAM air temperature (T).

To investigate the variability of hydroclimates, a linear trend is adopted. Over 57 years (i.e. 1950-2006), the dry season rainfall shows a slightly increasing trend of 16.3 mm (Figure 3.10(a)). Due to the inverse relationship between the dry season rainfall and MAM temperature, the increasing trend of the dry season rainfall is responsible for the decreasing trend of MAM temperature, as shown in Figure 3.11. From 1951 to 2007, MAM temperature shows a decreasing trend (by 0.6°C) which is statistically significant at 95% confidence level by the standard *t*-test (Haan, 2002). Although the pre-monsoon season (MJJ) rainfall develops in an inverse relationship with MAM temperature, Figure 3.10(b) shows a slightly decreasing trend of MJJ rainfall by 35.4 mm over 58 years. For the monsoon season (ASO) rainfall (Figure 3.10(c)), due to positive relationship with MAM temperature, a decreasing trend of 105.8 mm is obtained from 1950 to 2007 which is significant at a 95% confidence level. The decreasing trends of MAM temperature and ASO rainfall over the study basin are consistent with global trends as shown in Figure 1.1. These are also corroborated by Trenberth et al. (2007), who present a decreasing trend of precipitation since 1970 over the area of 10° to 30°N latitude based on historical data from 1900 to 2005.

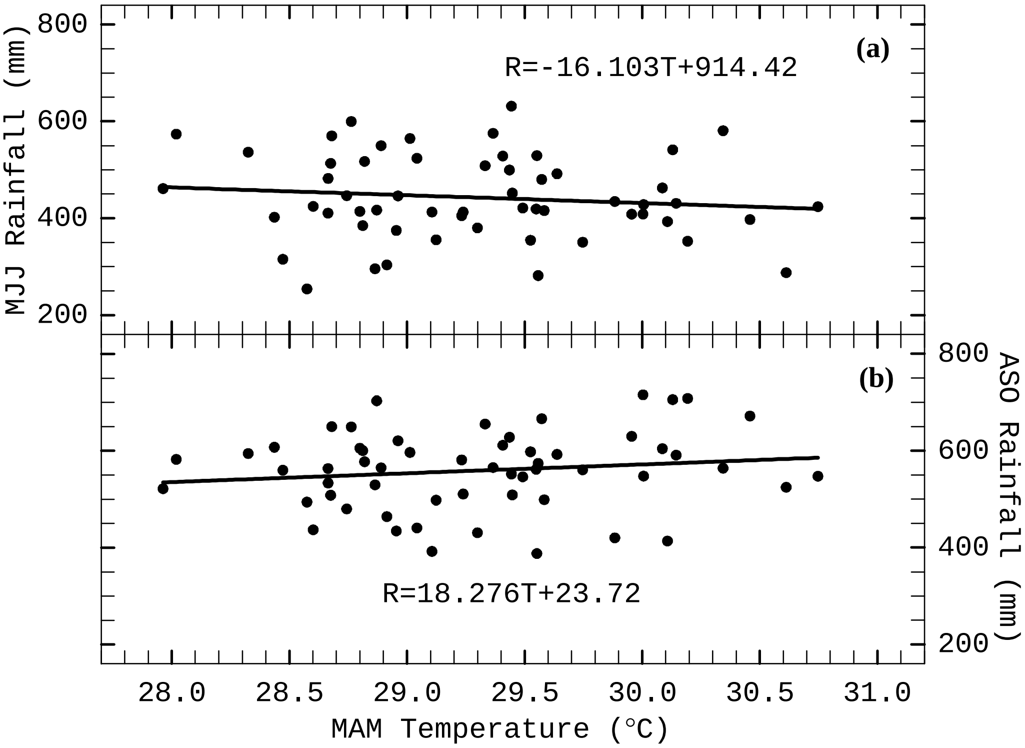


Figure 3.8: Scatter plots between MAM temperature (T) and rainfall (R) during (a) the pre-monsoon season (MJJ); and (b) the monsoon season (ASO).

To initially investigate the effects of anomalous atmospheric circulations, e.g. ENSO, on the variability of rainfall over the study basin, Figure 3.12 shows the standardized rainfall anomalies associated with ENSO years (Table 3.2). Out of the 20 events of El Niño, the dry season rainfall experiences 11 (9) events of above- (below-) normal rainfall. Out of the 15 events of La Niña, the dry season rainfall experiences 8 (7) events of above- (below-) normal rainfall. On the other hand, the MJJ (ASO) rainfall indicates 16 (11) events of below-normal rainfall during El Niño years. Under La Niña, there are 11 events of above-normal rainfall during MJJ and ASO. Therefore, the effects of ENSO on dry season rainfall are inconsistent. However, the below-normal pre-monsoon and monsoon season rainfall coincide with the warm phase of ENSO, i.e. El Niño, and vice versa for the cold phase. Furthermore, the MJJ and ASO rainfalls show different and asymmetrical responses with both phases of ENSO.

3.4 Summary

The study basin, the Ping River Basin, is located in northern Thailand (Southeast Asia). The basin has an area of 33,899 km^2, and 71.46% of this area is covered with forests. In terms of climate, the temperature during the summer season (i.e. MAM) helps to develop the land-sea temperature gradient, which in turn strengthens the monsoons over this region. The annual cycle of rainfall shows a bi-modal variability with two peaks. One peak is in

May which is associated with the pre-monsoon season during MJJ. Another peak is in September and corresponds with the monsoon season during ASO. The annual cycle of rainfall is caused by ITCZ and the Southwest monsoon. The remaining period (i.e. November to April of the following year) is defined as the dry season.

The dry season rainfall develops an inverse relationship with MAM air temperature where more (less) dry season rainfall tends to decrease (increase) MAM air temperature. In addition, MAM temperature is negatively (positively) correlated with MJJ (ASO) rainfall. Over 57 years, dry season rainfall has tended to increase by 16.3 mm, which is consistent with a decreasing trend of MAM air temperature by 0.6°C due to their inverse relationship. Furthermore, the pre-monsoon (MJJ) and monsoon (ASO) season rainfall show decreasing trends (by 35.4 and 105.8 mm over 57 years respectively). It is also important to note that the below-normal MJJ and ASO rainfall is associated with El Niño and vice versa for La Niña. However, the effects of warm and cold phases of ENSO on MJJ and ASO rainfall are inconsistent and asymmetrical.

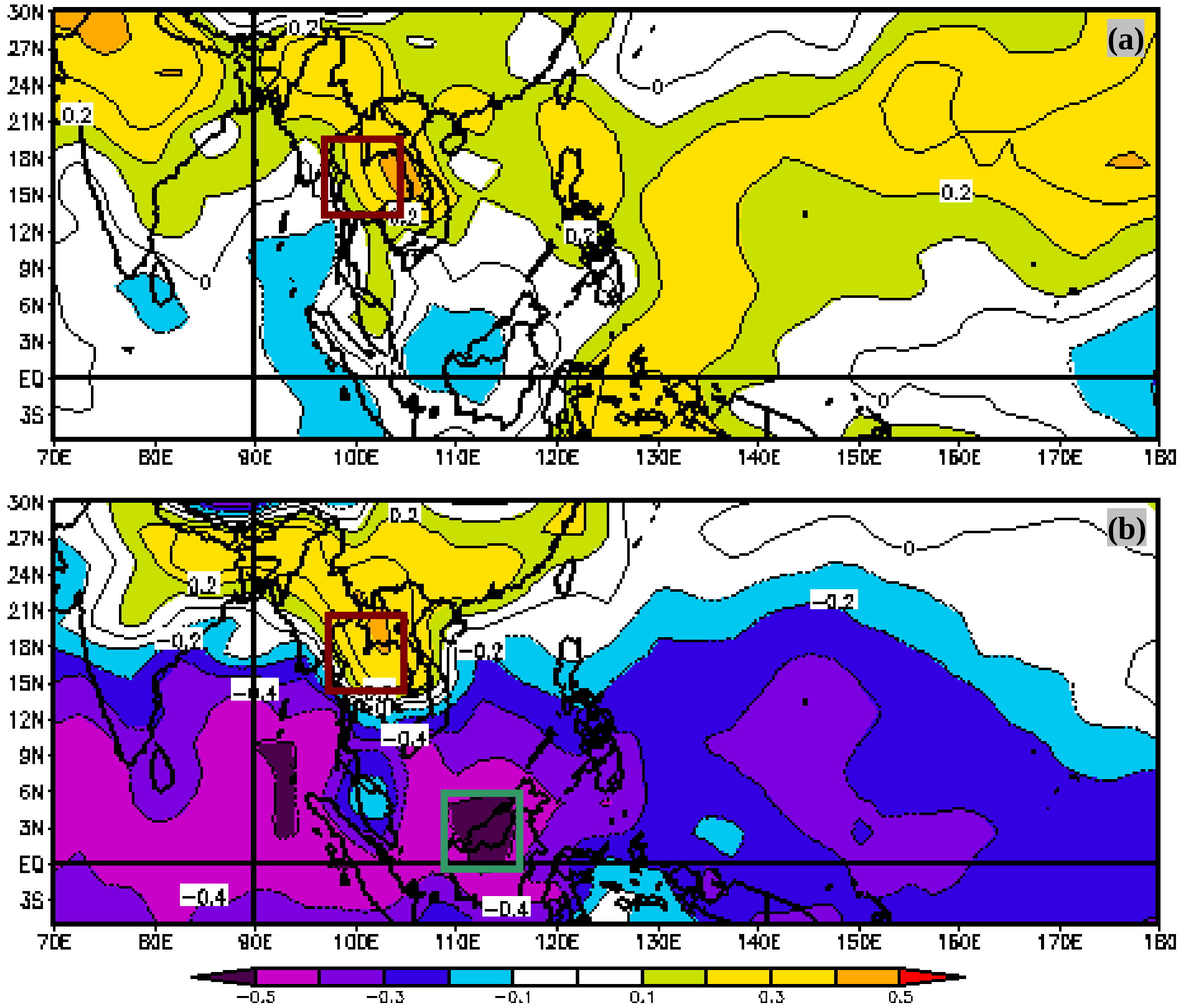


Figure 3.9: Correlation maps between (a) MAM air temperature and MJJ rainfall; and (b) MAM air temperature and ASO rainfall from 1950 to 2007. The 95% confidence levels of the correlations are ±0.3.

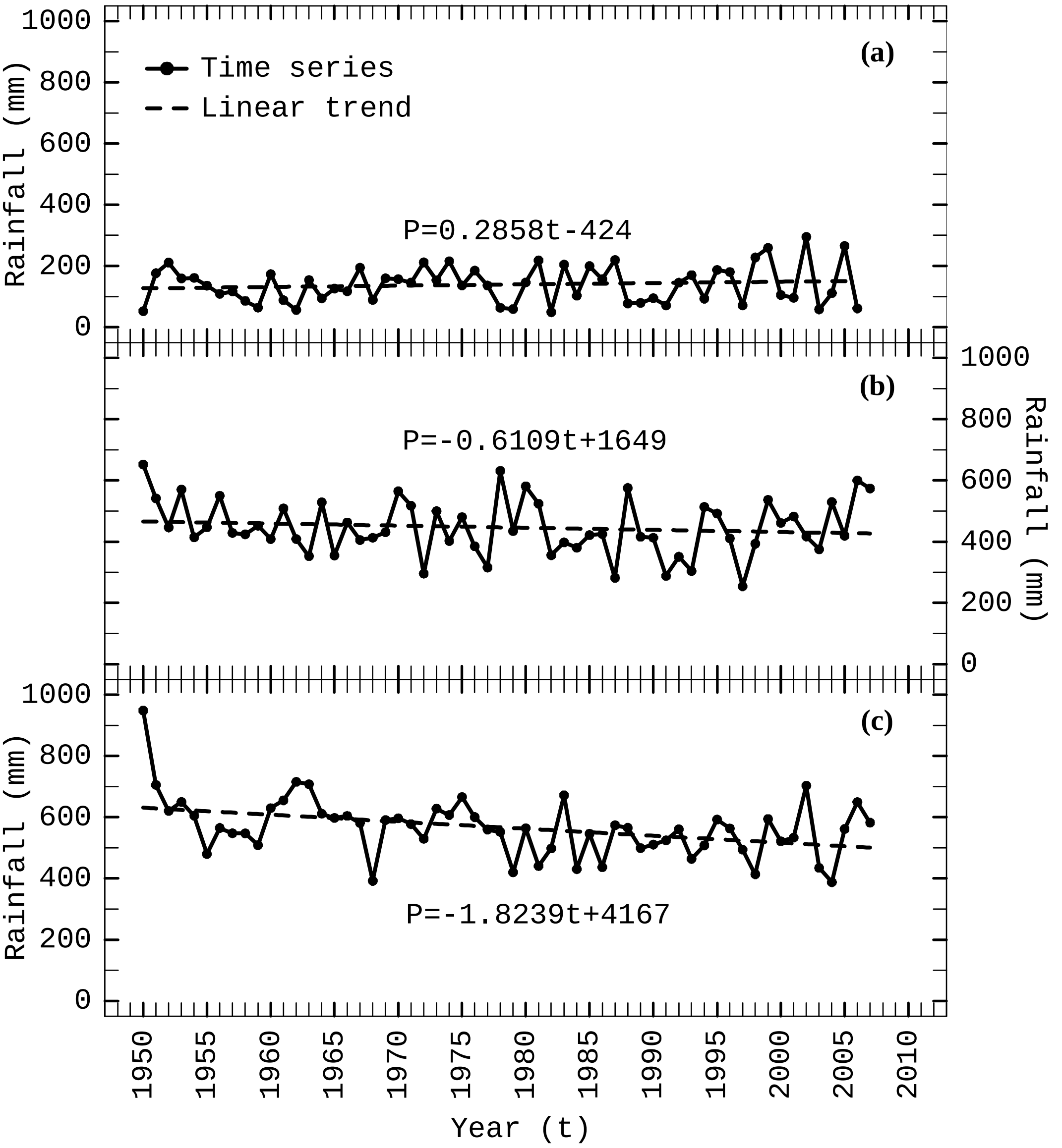


Figure 3.10: Time series of rainfall (P) during (a) the dry season (November to April); (b) the pre-monsoon season (MJJ); and (c) the monsoon season (ASO).

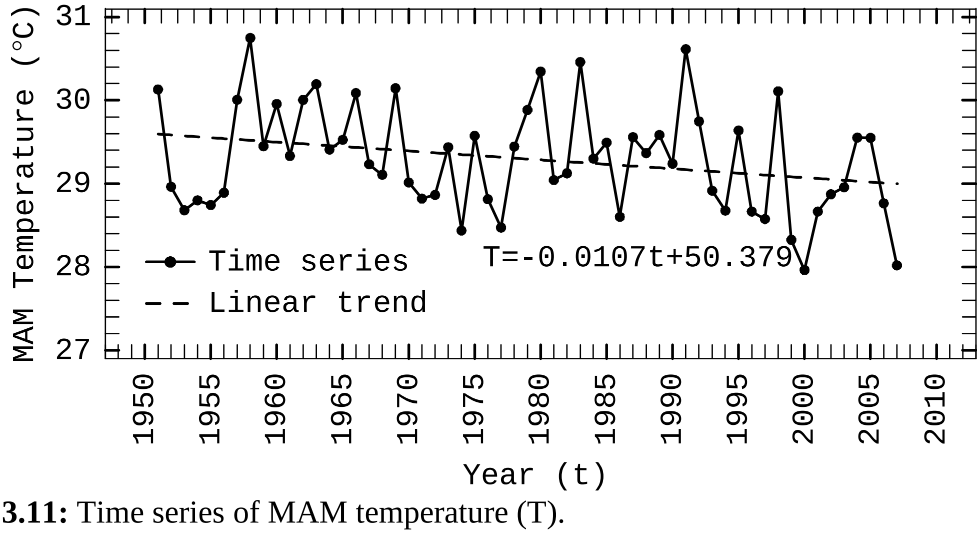


Figure 3.11: Time series of MAM temperature (T).

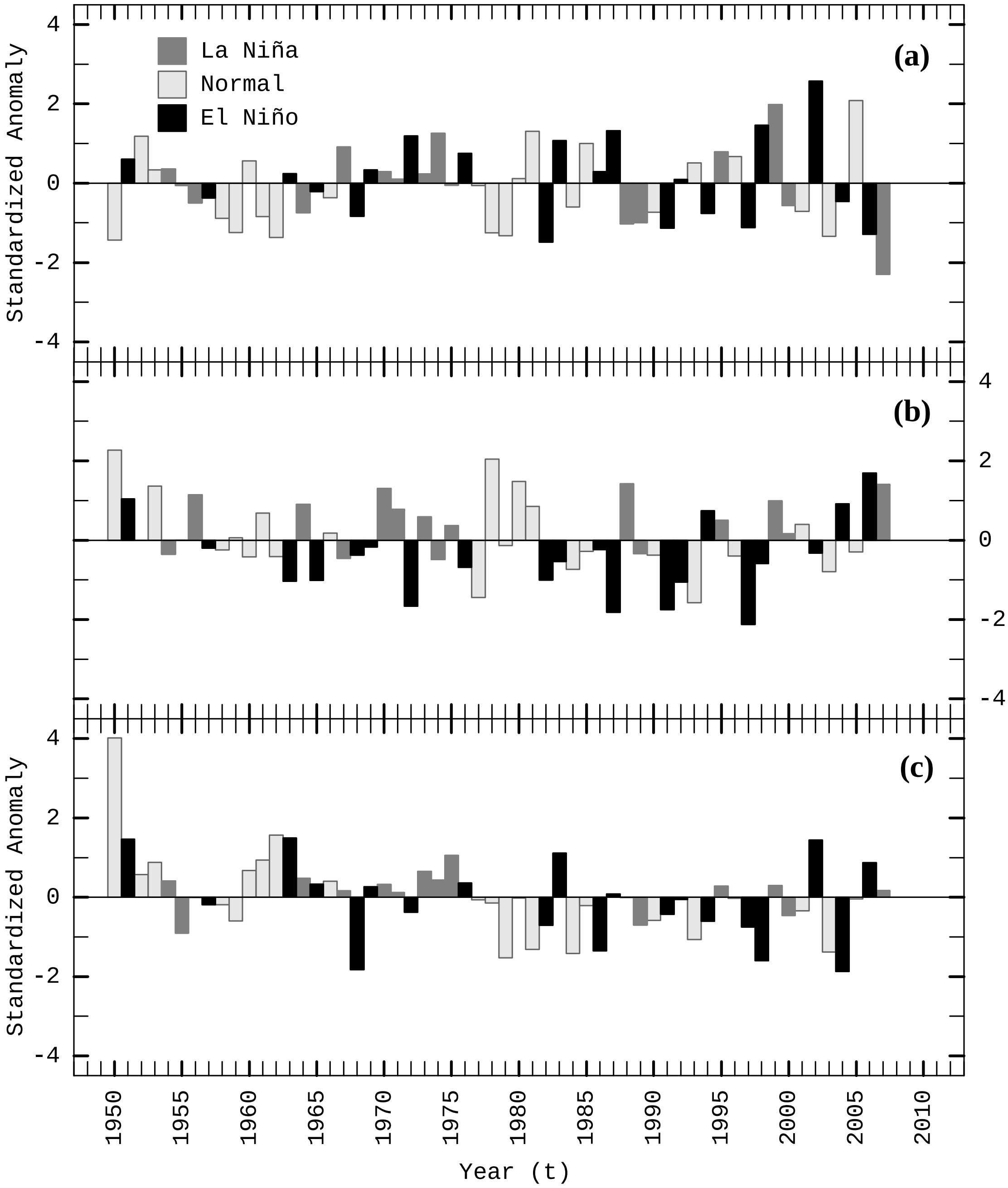


Figure 3.12: Time series of standardized anomalies of rainfall during (a) the dry season (November to April); (b) the pre-monsoon season (MJJ); and (c) the monsoon season (ASO).

Table 3.2: El Niño and La Niña Years Defined by the Climate Prediction Center (CPC) and the Center for Ocean-Atmospheric Prediction Studies (COAPS)

Phase	Year
El Niño	1951, 1957, 1963, 1965, 1968-69, 1972, 1976, 1982-83, 1986-87, 1991-92, 1994, 1997-98, 2002, 2004, 2006
La Niña	1954-56, 1964, 1967, 1970-71, 1973-75, 1988-89, 1995, 1999, 2000

Chapter 4
Predictor Identification for Rainfall in the Study Basin

4.1 Introduction

The objective of this chapter is to identify large-scale atmospheric predictors and optimal sets of predictors at long lead times to develop a statistical model which can simulate and downscale rainfall from atmospheric variables. First, a statistical relationship between rainfall and large-scale atmospheric variables has been developed using the correlation maps provided by the Physical Sciences Division (PSD). Based on a seasonal temporal scale, the correlation maps are created for different lead times of large-scale atmospheric variables from rainfall in order to investigate the long leading relationship between atmospheric variables and rainfall and to provide long-range predictability of the forecasting model. Second, the identified predictors are selected from correlation maps based on a significant relationship at long lead times. Then, the annual and decadal variability of identified predictors is determined using the observed data of the NCEP/NOAA and the projected data for the future by simulating a general circulation model (GCM) under several future climate scenarios. Finally, the optimal subsets of predictors, which are composed of the minimum members of mutually exclusive variables, are identified by an objective function like generalized cross validation (GCV), likelihood and Akaike criterion (AIC). In this study, GCV has been adopted. The identified combinations of predictors are subsequently used to develop a statistical model to forecast rainfall.

4.2 Predictor identification by correlation maps

4.2.1 Atmospheric variable description

To identify the predictors for rainfall using the statistical approach, the large-scale atmospheric variables used in this study are obtained from the reanalysis derived data provided by the NCEP/NOAA (Kalney et al., 1996). The daily and monthly data from 1948 to the present of several atmospheric variables (e.g. temperature, pressure, geopotential height and outgoing longwave radiation (OLR)) are available online (PSD, 2007a). The observed data obtained from different sources (e.g. ships, satellites and ground stations) cover the global grid of 2.5°latitude x 2.5°longitude. Data are also provided for various vertical levels such as surface, two velocity levels (at 0.2101 and 0.995 sigma) and 17 pressure levels from 10 to 1,000 millibar (mb). Correlation maps are developed between rainfall and the four principal atmospheric variables, which play a vital role in influencing the convection over Thailand and in strengthening the monsoon downpour. Monthly data recorded from 1948 to 2007 of surface air temperature (SAT), sea level pressure (SLP), surface zonal wind (SXW) – i.e. wind blowing in the latitudinal direction, and surface meridian wind (SYW) – i.e. wind blowing in the longitudinal direction, are used in this study.

4.2.2 Methodology

The correlation analysis in terms of Pearson's product moment correlation coefficient (r) (shown in Equation 4.1) is the determination of linear dependency between two variables: dependent (y_i) and independent (x_i) variables. The coefficients vary from -1.0 to +1.0. A value of -1.0 (+1.0) indicates a strong linear negative (positive) relationship between the two variables – i.e. increasing y_i corresponds to decreasing (increasing) x_i and vice versa for decreasing y_i.

Based on Pearson's *r*, correlation maps are adopted to develop a statistical relationship between rainfall averaged over 50 selected stations and large-scale atmospheric variables. The correlation maps are online interactive plots and analyses provided by the Earth System Research Laboratory (ESRL) of PSD (ESRL, 2008). The rainfall during the pre-monsoon season (MJJ), monsoon season (ASO) and dry seasons (i.e. NDJ and FMA) from 1950 to 2007 has been cross-correlated to the gridded SAT, SLP, SXW and SYW provided by the NCEP/NOAA. Analysis has also been carried out at several lead times of atmospheric variables ranging from 4 to 15 months prior to the start of the season. Predictors are selected over potential regions (e.g. the study basin, the Pacific Ocean, the Indian Ocean and the South China Sea) based on two criteria: (i) a significant relationship with rainfall at 95% confidence levels; and (ii) the long lead periods of each predictor.

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}} \qquad \text{Equation 4.1}$$

where $\bar{x}$ and $\bar{y}$ are the averages of x_i and y_i respectively, and *n* is the total number of pairs obtained from the data.

To test the criterion of significant relationships, the significance of a correlation coefficient is tested using Fisher's Transformation (z'), as shown in Equation 4.2 (Haan, 2002). Using Fisher's Transformation, the data set, which may not be normally distributed, is converted into a modified data set, which is Gaussian distributed. Subsequently, the deviation from the mean (i.e. z'_{upper} and z'_{lower}) of the modified data set can be estimated by standard normal distribution (*z*) based on a desired significance level (*p*), as shown in Equation 4.3. Then, the upper and lower bounds of the correlation coefficient can be calculated by Equation 4.2 to convert the values of z'_{upper} and z'_{lower} to *r*. For example, the upper and lower bounds of a significant correlation at a 95% confidence level for *n*=58 (i.e. the data from 1950 to 2007) are +0.26 and -0.26 respectively.

$$z' = 0.5 * \ln\left(\frac{1+r}{1-r}\right) \qquad \text{Equation 4.2}$$

$$z'_{upper} = z' + z_{p/2}\frac{1}{\sqrt{n-3}} \qquad \text{Equation 4.3(a)}$$

$$z'_{lower} = z' - z_{p/2}\frac{1}{\sqrt{n-3}} \qquad \text{Equation 4.3(b)}$$

where *r* is the correlation coefficient between two variables, $z_{p/2}$ is the standard normal distribution at a significance level (*p*), and *n* is the amount of data.

According to the correlation analysis at several long lead periods of predictors, the consistent and slow development of linear relationships between rainfall and large-scale atmospheric variables, in particular SAT, can be observed (Sahai et al., 2003). The developed long leading relationships can provide long-range predictability in a forecasting model which is, in turn, useful in water resources planning, agricultural practices and insurance policy making for anomalous weather events.

4.2.3 Predictor selection

From the correlation maps, the non-linear relationship between four large-scale atmospheric variables and rainfall during MJJ, NDJ and FMA are observed but no significant relationships can be found from 1950 to 2007. However, the relationships developed in post-1980 are significant at a 95% confidence level (i.e. the upper and lower bounds for n=28 from 1980 to 2007 are +0.37 and -0.37 respectively) over potential regions such as the study basin and the South China Sea. The non-linear relationships over decades are influenced by the shift in the spatial coverage by anomalous SST in the Pacific Ocean (i.e. ENSO regions). Anomalous SST, which is observed over the dateline Pacific Ocean during pre-1980, tends to expand over the equatorial eastern Pacific Ocean during post-1980. As a result, the convection of the Walker circulation brings a greater effect of SST anomalies from the eastern Pacific Ocean to Thailand and Southeast Asia in the post-1980 period (Singhrattna et al., 2005b; Krishna Kumar et al., 1995). However, the influence of shifted ENSO regions on the monsoon season rainfall (i.e. ASO rainfall) is not as strong as the significant relationships that can be observed during 1950 to 2007. Therefore, a statistical relationship between four large-scale atmospheric variables and MJJ, NDJ and FMA rainfalls is developed for the time period of 1980 to 2007, whereas the relationship with ASO rainfall is established for the period of 1950 to 2007.

Figure 4.1 to 4.4 show the correlation maps between atmospheric variables and rainfall during MJJ, ASO, NDJ and FMA respectively. More correlation maps, done for 12 lead times of atmospheric variables, are presented in Appendix C1 to C15. Based on significant correlations at 95% confidence levels, the predictors for monsoon rainfall (i.e. MJJ and ASO rainfall) are identified over the study basin and nearby seas, such as the South China Sea and the Andaman Sea. The SAT over the study basin (i.e. northern Thailand) is positively correlated with MJJ rainfall, whereas decreasing SLP over the Gulf of Thailand is associated with increasing MJJ rainfall. Furthermore, significant correlations between the pre-monsoon season (MJJ) rainfall and SXW (SYW) can be found over the equatorial Indian Ocean (the eastern equatorial Pacific Ocean). For the monsoon season (ASO) rainfall, a higher SAT (SLP) over the South China Sea (northern Thailand) is associated with decreasing ASO rainfall. Moreover, stronger winds blowing in the latitudinal and longitudinal directions (SXW and SYW) from the Gulf of Thailand and the Andaman Sea respectively, bring more moisture and convection to the study basin and subsequently increase ASO rainfall.

On the other hand, the predictors of the dry season rainfall (i.e. NDJ and FMA rainfall) are identified over more distant regions like northeastern India and Java (Indonesia). This shows the influence of remote atmospheric circulations on local hydroclimates (Harshburger et al., 2002; Tereshchenko et al., 2002). Higher SAT over the southeast coast of Sumatra (Indonesia) corresponds to an increasing NDJ rainfall. SXW and SYW over the Indian Ocean and northeastern India respectively present significant remote influences on NDJ rainfall. Furthermore, positive significant relationships between FMA rainfall and SAT (SLP) are observed over the regions of Java in Indonesia (the western Pacific Ocean). A stronger SXW over the eastern Pacific Ocean and a weaker SYW over the Indian Ocean are also significantly related to increasing FMA rainfall.

In terms of lead periods of large-scale atmospheric predictors, monsoon rainfall (i.e. MJJ and ASO rainfall) is correlated with atmospheric variables at longer lead times than the dry season rainfall (i.e. NDJ and FMA rainfall). A significant relationship between MJJ rainfall and the four atmospheric variables can be found at lead times varying from 5 to 14 months, whereas ASO rainfall is significantly correlated with SAT, SLP, SXW and SYW at the longest lead time of 15 months. For NDJ rainfall, a long leading relationship is observed at 6-12 months. Moreover, significant links between FMA rainfall and large-scale atmospheric variables have lead times of 7-14 months.

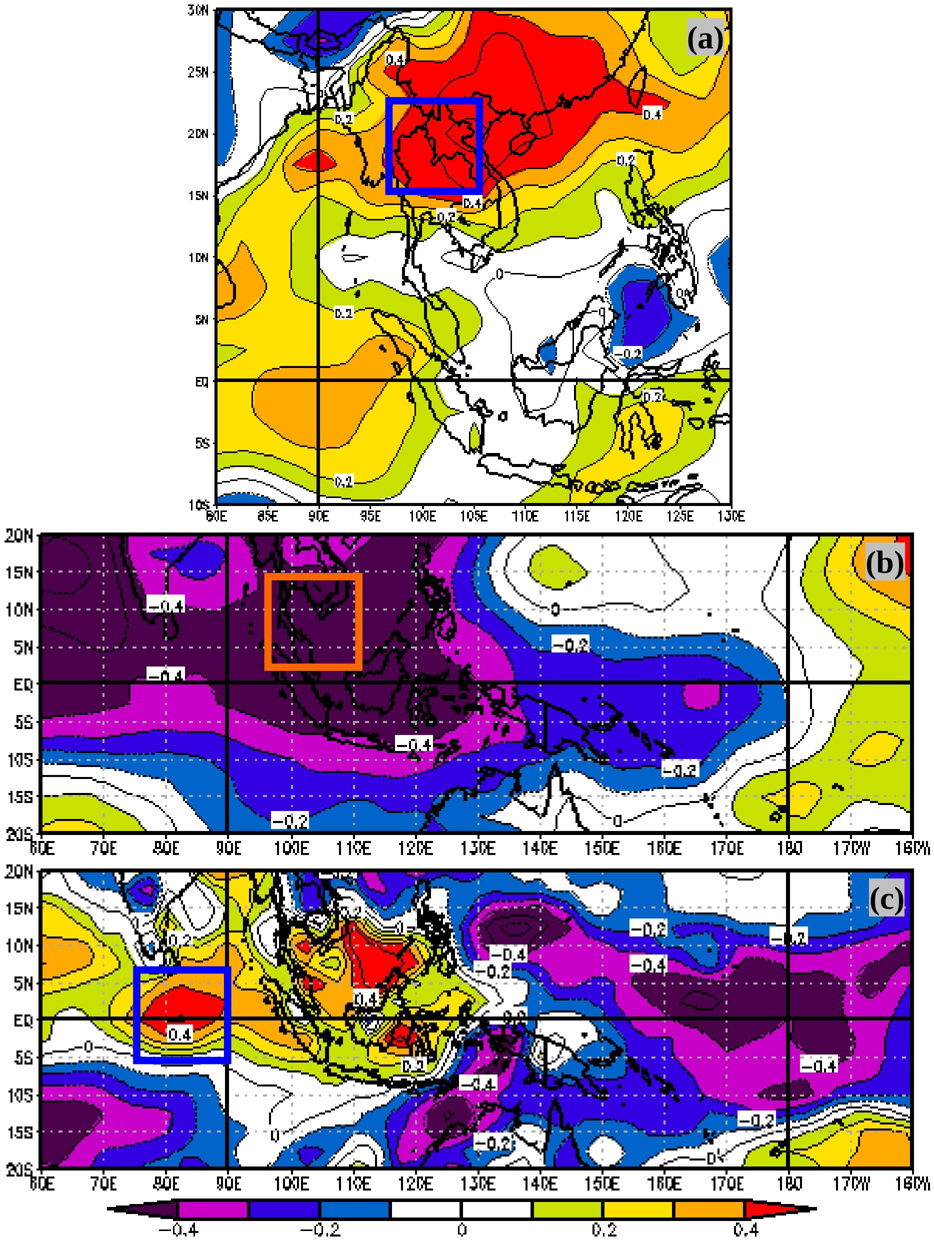


Figure 4.1: Correlation maps between MJJ rainfall and (a) MJJ SAT; (b) NDJ SLP; (c) OND SXW; and (d) MJJ SYW. The 95% confidence levels of the correlations are ±0.37 (*n*=28 from 1980 to 2007).

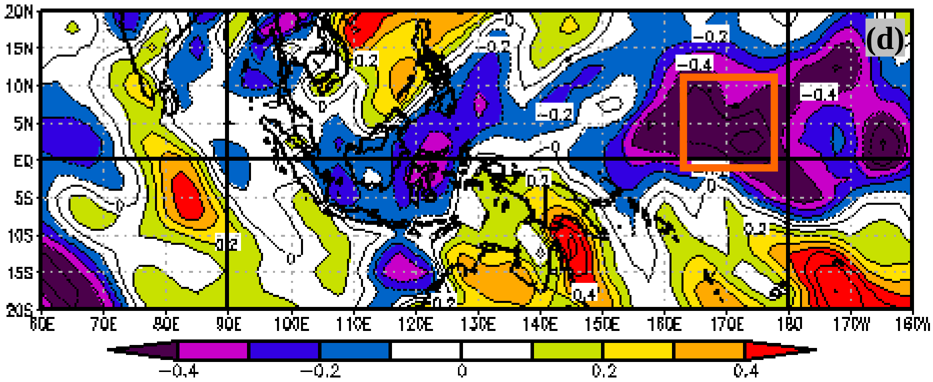


Figure 4.1 (cont).

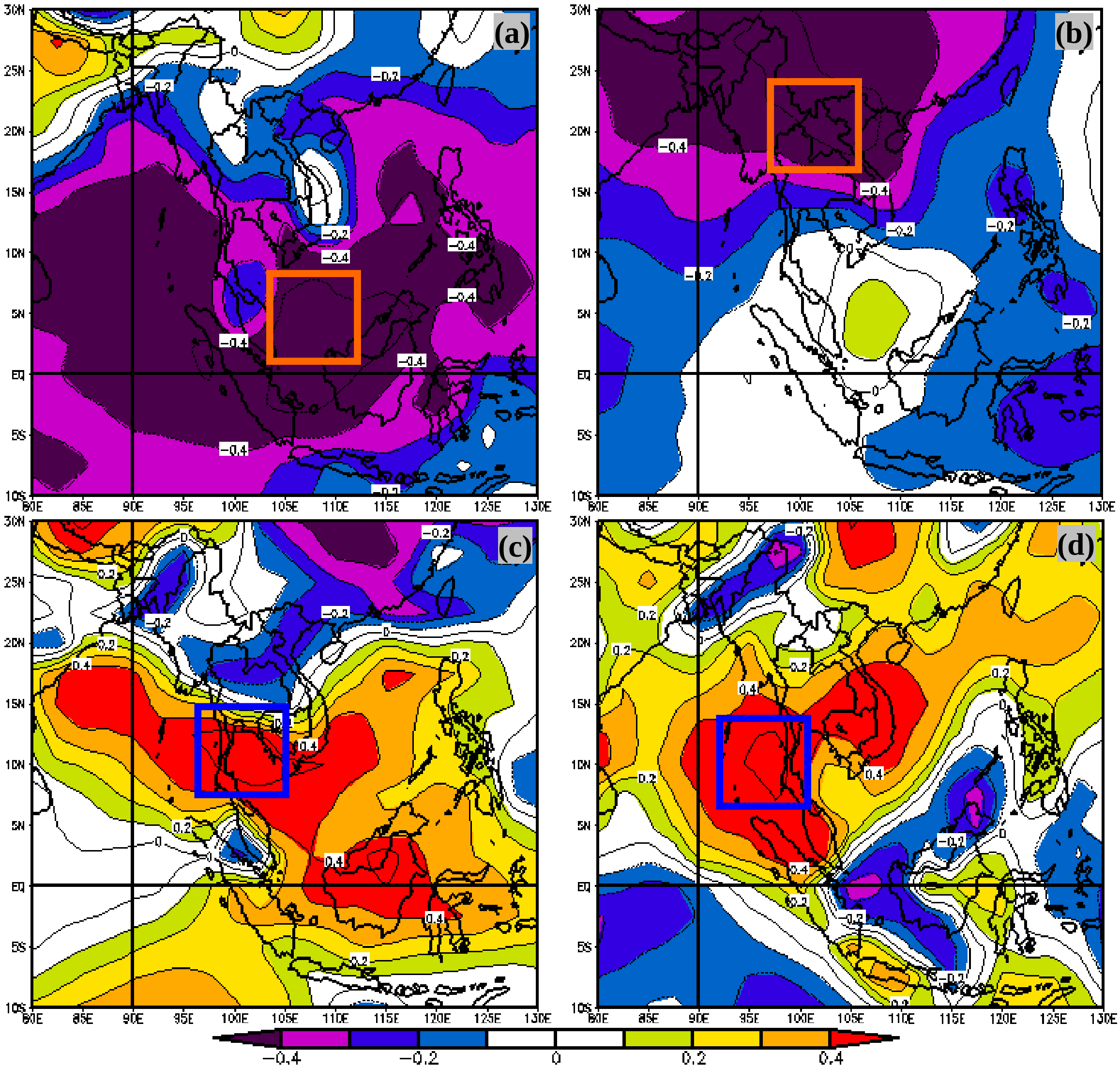


Figure 4.2: Correlation maps between ASO rainfall and (a) MJJ SAT; (b) JJA SLP; (c) OND SXW; and (d) SON SYW. The 95% confidence levels of the correlations are ±0.26 (*n*=58 from 1950 to 2007).

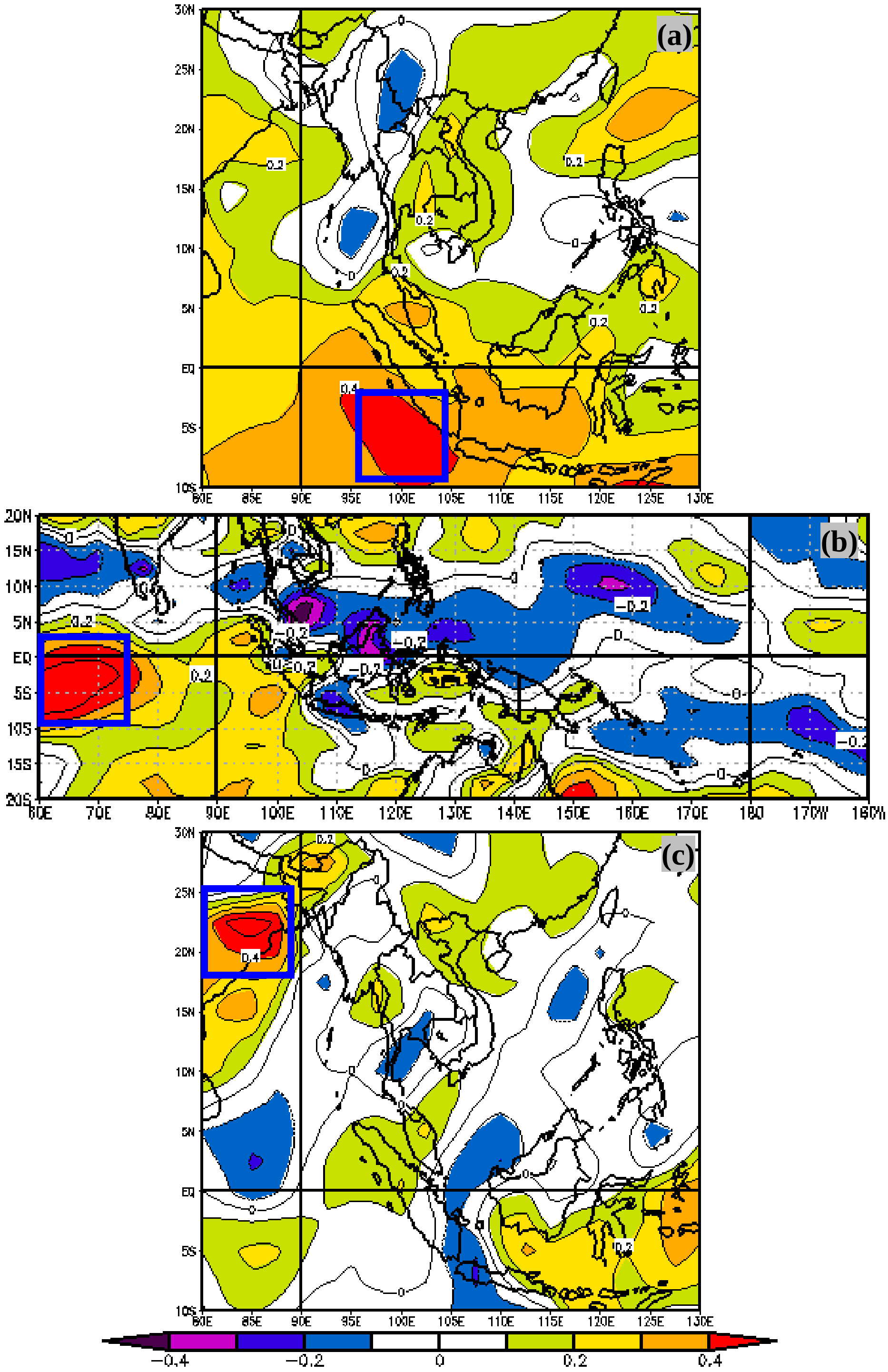


Figure 4.3: Correlation maps between NDJ rainfall and (a) FMA SAT; (b) MJJ SXW; and (c) NDJ SYW. The 95% confidence levels of the correlations are ±0.37 (*n*=28 from 1980 to 2007).

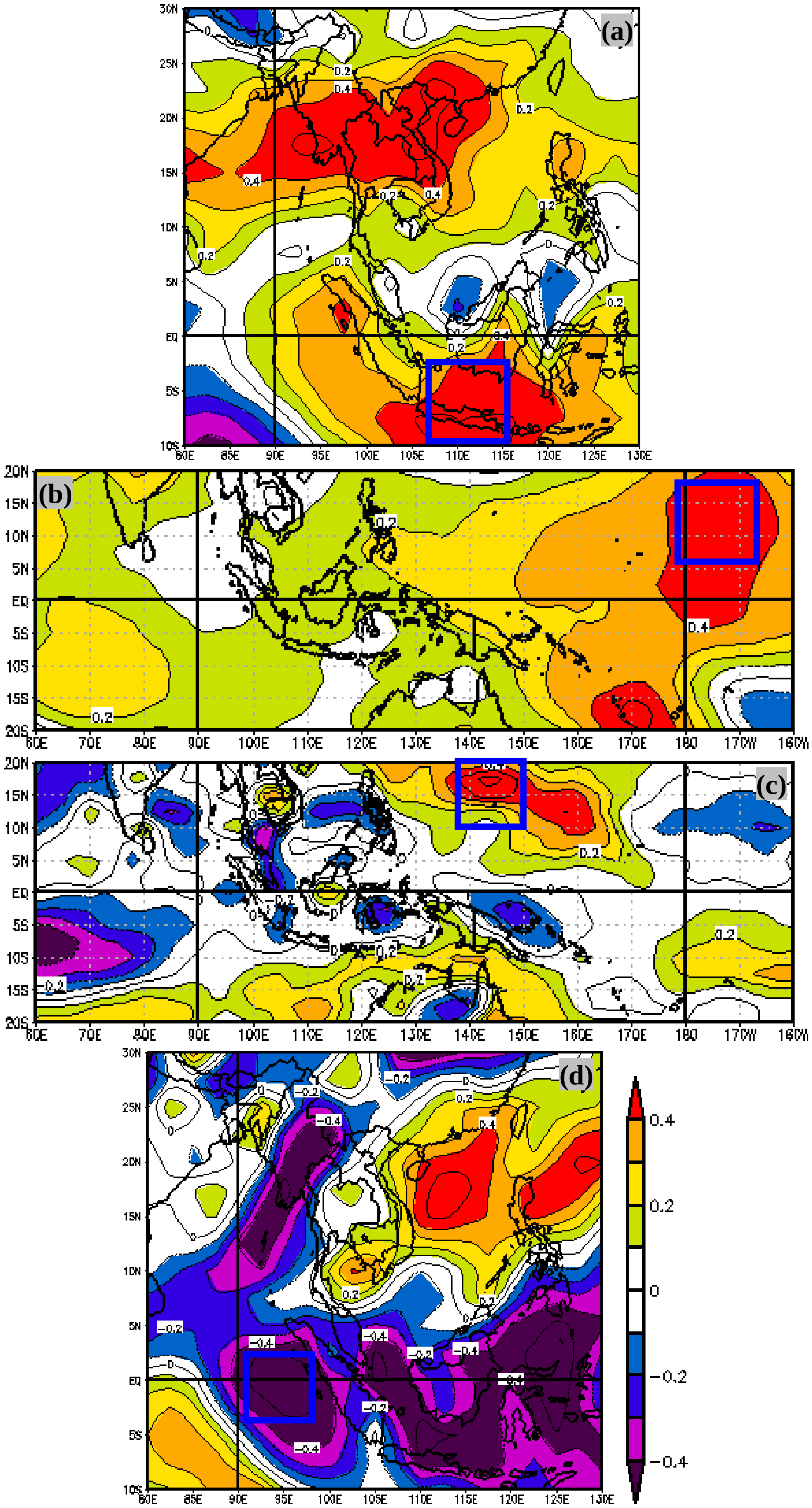


Figure 4.4: Correlation maps between FMA rainfall and (a) JJA SAT; (b) JFM SLP; (c) DJF SXW; and (d) MJJ SYW. The 95% confidence levels of the correlations are ±0.37 (*n*=28 from 1980 to 2007).

Hence, atmospheric predictors have been selected over the identified regions based on their significant relationships with rainfall at 95% confidence levels and at long lead times of the predictors. The developed statistical relationship between rainfall and large-scale atmospheric variables shows the influence of atmospheric-oceanic circulations on seasonal rainfall in the Ping River Basin. This confirms the observation that the variability of atmospheric-oceanic circulations brings its effects on local hydroclimates over those regions located near or even at a distance from these sources.

Moreover, a consistent and significant relationship associated with monsoon rainfall (MJJ and ASO) can be developed at long lead times of large-scale atmospheric variables, in particular of surface temperature, which is corroborated by Sahai et al. (2003) and Nicholls (1983). Sahai et al. (2003) showed the significant lead time relationships between Indian summer monsoon rainfall (ISMR) and SST. The significant relationships show a slow and consistent temporal evolution, which indicates a lead time of four years prior to the monsoon season. Likewise, Nicholls (1983) found a lead time of 16 months from the relationships between ISMR and SST near Indonesia. However, the dry season rainfall (NDJ and FMA) of the study basin is influenced by unstable local conditions for a finer time scale. These local conditions include increasing surface temperature and humidity on an hourly and daily basis. The significant relationship between atmospheric variables and rainfall during NDJ and FMA are observed at shorter lead periods than those of MJJ and ASO rainfall (TMD, 2007). Table 4.1 summarizes the identified predictors of MJJ, ASO, NDJ and FMA rainfall based on significant relationships at long lead times. It is also important to note that significant relationships between NDJ rainfall and SLP are hardly found over the regions under study here. Furthermore, observed rainfall (RAIN) is included in the list of predictors because the lag autocorrelations or correlogram of RAIN (Figure 4.5) show significant correlation coefficients at 95% confidence levels associated with 6- and 12-month lags. This suggests a long leading predictability of forecasting models using RAIN as a predictor. Therefore, five predictors (i.e. SAT, SLP, SXW, SYW and RAIN) are identified for MJJ, ASO and FMA rainfall, and four predictors (i.e. SAT, SXW, SYW and RAIN) are selected for NDJ rainfall.

4.3 Predictors from a General Circulation Model (GCM)

This study aims to determine the effects of future climate on hydroclimates like rainfall in the study basin. A statistical model has been developed using large-scale atmospheric variables as predictors to simulate rainfall under various conditions of future climate. The IPCC presents the variability of oceanic and atmospheric variables such as precipitation, temperature and pressure achieved from different general circulation models (GCMs) which run under various scenarios of future climate. GCMs are of two types: the atmospheric general circulation model (AGCM) and the oceanic general circulation model (OGCM). For this reason, GCMs are also called coupled atmosphere-ocean general circulation models (AOGCMs) and are used to simulate the climate under several scenarios of increasing greenhouse gas (GHG) concentrations. The emission scenarios of GHG are obtained by keeping in consideration possible climatic, socio-economic and environmental changes. The results from a GCM represent future climate projections that are expected in feasible environmental systems and to chart out human activities in terms of economic, demographic and technological growth. However, working under the assumption that GCMs show better performance in simulating atmospheric variables at the upper air level rather than at the surface level, downscaling approaches are adopted to spatially downscale local hydroclimates from the large-scale atmospheric variables. The downscaled models also aim at determining the effects of future climate on hydroclimates such as precipitation and streamflow.

Table 4.1: The Identified Predictors for MJJ, ASO, NDJ and FMA Rainfall

Season of rainfall	Atmospheric variables	Identified region	
		Latitude	Longitude
MJJ	SAT	20.0°N	97.5°-102.5°E
	SLP	7.5°-10.0°N	102.5°-107.5°E
	SXW	0°	82.5°-87.5°E
	SYW	0°-2.5°N	172.5°-175.0°E
	RAIN	15.30°-19.36°N	98.05°-100.70°E
ASO	SAT	2.5°-5.0°N	107.5°-110.0°E
	SLP	17.5°-20.0°N	97.5°-100.0°E
	SXW	10.0°N	100.0°-102.5°E
	SYW	10.0°N	95.0°-97.5°E
	RAIN	15.30°-19.36°N	98.05°-100.70°E
NDJ	SAT	2.5°-7.5°S	97.5°-102.5°E
	SXW	2.5°-5.0°S	62.5°-67.5°E
	SYW	20.0°-22.5°N	85.0°E
	RAIN	15.30°-19.36°N	98.05°-100.70°E
FMA	SAT	7.5°S	110.0°-112.5°E
	SLP	15.0°N	187.5°-192.5°W
	SXW	17.5°N	140.0°-150.0°E
	SYW	2.5°S-2.5°N	95.0°-97.5°E
	RAIN	15.30°-19.36°N	98.05°-100.70°E

MJJ: May-June-July; ASO: August-September-October; NDJ: November-December-January; FMA: February-March-April.

SAT: surface air temperature; SLP: sea level pressure; SXW: surface zonal or latitudinal wind; SYW: surface meridian or longitudinal wind; RAIN: observed rainfall over the study basin.

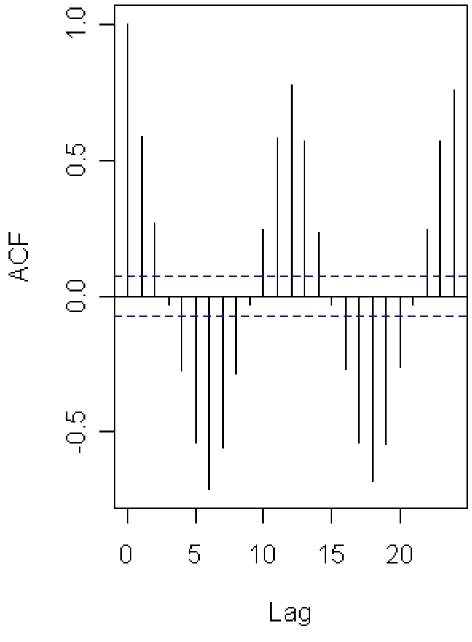


Figure 4.5: Correlogram or lag autocorrelation (ACF) of rainfall time series from 1950 to 2007. The blue-dotted lines represent significant correlations at 95% confidence levels.

In this study, GCM data used are based on the IPCC Third Assessment Report (TAR) conducted in 2001 (IPCC, 2001). The descriptions of GCMs provided in the IPCC TAR are presented in Table 4.2. The smallest model resolution (i.e. 2.8°latitude x 2.8°longitude) is provided by ECHAM4/OPYC3, NCAR-CSM and NCAR-PCM. In contrast, the largest dimension (i.e. 5.6°latitude x 5.6°longtitude) is presented by CCSR/NIES AGCM+CCSR OGCM, which also provides the longest duration of simulation from 1890 to 2100. In general, the length of simulation by GCMs varies from 100 to 211 years. In this study, a GCM is selected, and its simulated results are used to downscale and to determine the effects of future climate on seasonal rainfall in the Ping River Basin.

Table 4.2: Description of the GCMs corresponding to the IPCC TAR

Model name	Center	Country	Scenario	Model resolution	Temporal coverage
ECHAM4/ OPYC3	Max Planck Institüt für Meteorologie (MPlfM)	Germany	A2, B2	2.8° x 2.8°	1990-2100
HadCM3	Hadley Centre for Climate Prediction and Research (HCCPR)	UK	A1Fl, A2, A2b, A2c, B1, B2, B2b	3.75° x 2.5°	1950-2099
CSIRO-MK2	Australia’s Commonwealth Scientific and Industrial Research Organisation (CSIRO)	Australia	A1, A2, B1, B2	5.6° x 3.2°	1961-2100
NCAR-CSM	National Centre for Atmospheric Research (NCAR)	USA	A2	2.8° x 2.8°	2000-2099
NCAR-PCM	NCAR	USA	A1B, A2, B2	2.8° x 2.8°	1980-2099
GFDL-R30	Geophysical Fluid Dynamics Laboratory (GFDL)	USA	A2, B2	3.75° x 2.25°	1961-2100
CGCM2	Canadian Center for Climate Modelling and Analysis (CCCma)	Canada	A2, A2b, A2c, B2, B2b, B2c	3.75° x 3.75°	1900-2100
CCSR/NIES AGCM+CCSR OGCM	Center for Climate System Research/National Institute for Environmental Studies (CCSR/NIES)	Japan	A1, A1Fl, A1T, A2, B1, B2	5.6° x 5.6°	1890-2100

4.3.1 Description of atmospheric variables from the GFDL-R30

As mentioned above, large-scale atmospheric predictors for rainfall during MJJ, ASO, NDJ and FMA identified by the correlation maps include SAT, SLP, SXW, SYW and RAIN (see also Table 4.1). Based on the criterion of the availability of identified predictors (Table 4.3), GFDL-R30, which provides simulated data of all the identified predictors, has been selected.

Table 4.3: Availability of Data (●) from the GCMs according to the Identified Predictors for Rainfall

Model	Scenario	SAT	SLP	SXW	SYW	RAIN
ECHAM4/OPY3	A2	●	●			●
	B2			●	●	
HADCM3	A2	●	●	a	a	●
	A2b	●		a	a	●
	A2c	●		a	a	●
	B2	●	●	a	a	●
CSIRO-Mk2	A1	●	●	a, b	a, b	●
	A2	●	●	a, b	a, b	●
	B1	●	●	a, b	a, b	●
	B2	●	●	a, b	a, b	●
NCAR-CSM	A2	●	●			●
NCAR-PCM	A1b	●	●	a	a	●
	A2	●	●	a	a	●
	B2	●	●	a	a	●
GFDL-R30	A2	●	●	●	●	●
	B2	●	●	●	●	●
CGCM2	A2	●	●	a, b	a, b	●
	B2	●	●	a, b	a, b	●
CCSR/NIES AGCM+CCSR OGCM	A1	●	●	b	b	●
	A1Fl	●	●	b	b	●
	A1T	●	●	b	b	●
	A2	●	●	b	b	●
	B1	●	●	b	b	●
	B2	●	●	b	b	●

SAT: surface air temperature; SLP: sea level pressure; SXW: surface zonal or latitudinal wind; SYW: surface meridian or longitudinal wind; RAIN: observed rainfall over the study basin.
a: only mean wind at 10 m is available; b: only wind at 200 hPa and higher levels is available.

GFDL-R30, which is a coupled AOGCM of the Geophysical Fluid Dynamics Laboratory (GFDL), is composed of four components: an atmospheric spectral GCM, an OGCM, a simple model of sea ice, and a model of land surface processing. The atmospheric component has a spatial resolution of 2.25°latitude x 3.75°longtitude covering 7,680 global grids. It also covers 14 vertical levels. The ocean component has double the number of global grids with a resolution of 1.125°latitude x 3.75°longtitude and 18 vertical levels (Delworth et al., 2002). Functioning with the prediction that atmospheric CO_2 concentration will double by 2100, GFDL-R30 simulates two scenarios of future climate: A2 and B2 scenarios. In term of demographic growth, both scenarios are characterized by a continuously increasing human population. However, the population under A2 changes at a rate faster than under B2 (Table 4.4). Based on the population projection of the International Institute for Applied System Analysis (IIASA), the population growth under A2 is defined as slow fertility transition projection with high fertility and mortality rates. The 1995-2100 population growth under B2 is based on the UN 1998 medium long range projection (Arnell, 2004). Both scenarios emphasize improving per capita income and additional slow technological changes. In terms of economic growth, regionally oriented development is used to describe Scenario A2 with self-reliant and independent nations, whereas the economic growth under B2 is more focused on locally oriented development rather than global economic stability.

Table 4.4: Summary of the IPCC Emission Scenarios

Emi	Economic development	Global population (million)	Technology changes	Theme
A2	**Regionally-oriented**	**Continuously increasing;**	**Slower and more fragmented th**	**Self-reliance and preservat**

		2 8,7 2 11, 2 14,		ion of local identities
B2	Intermediate levels	Continuously increasing but	Less	Local economic, soci

		slower than A2;		**al and environmental sustainability**
		2 8,0		
		2 9,5		
		2 10,		

Source: Arnell (2004) and NIC (2009)

The IPCC-Data Distribution Center (IPCC-DDC) provides the results of atmospheric variables simulated by GFDL-R30 from 1961 to 2100 (IPCC-DDC, 2009). Table 4.5 shows the details of GFDL-R30 grid coverage corresponding to the identified regions of predictors for MJJ, ASO, NDJ and FMA rainfall. The largest spatial coverage (i.e. eight grids between 3.35397°S-3.35397°N latitude and 93.75°-97.50°E longitude) is associated with SYW of FMA rainfall predictors. In contrast, the identified region of SXW of ASO rainfall predictors covers only one grid of GFDL-R30, that between 10.06192°N latitude and 101.25°E longitude. The monthly data from 1961 to 2100, which are averaged over selected grids, are used in this study.

Table 4.5: Grid Coverage of the GFDL-R30 corresponding to the Identified Predictors for Rainfall

Predictor	Identified region by correlation maps		Numbers of grid	Grid coverage of GFDL-R30	
	Latitude	Longitude	(Lat. x Long.)	Latitude	Longitude
MJJ rainfall					
SAT	20.0°N	97.5°-102.5°E	2 (1 x 2)	19.00583°N	97.50°-101.25°E
SLP	7.5°-10.0°N	102.5°-107.5°E	6 (2 x 3)	7.82514°-10.06192°N	101.25°-108.75°E
SXW	0°	82.5°-87.5°E	4 (2 x 2)	1.11799°S-1.11799°N	82.50°-86.25°E
SYW	0°-2.5°N	172.5°-175.0°E	6 (3 x 2)	1.11799°S-3.35397°N	172.50°-176.25°E
RAIN	15.30°-19.36°N	98.05°-100.70°E	6 (3 x 2)	14.53387°-19.00583°N	97.50°-101.25°E
ASO rainfall					
SAT	2.5°-5.0°N	107.5°-110.0°E	2 (2 x 1)	3.35397°-5.58995°N	108.75°E
SLP	17.5°-20.0°N	97.5°-100.0°E	4 (2 x 2)	16.76985°-19.00583°N	97.50°-101.25°E
SXW	10.0°N	100.0°-102.5°E	1 (1 x 1)	10.06192°N	101.25°E
SYW	10.0°N	95.0°-97.5°E	2 (1 x 2)	10.06192°N	93.75°-97.50°E
RAIN	15.30°-19.36°N	98.05°-100.70°E	6 (3 x 2)	14.53387°-19.00583°N	97.50°-101.25°E
NDJ rainfall					
SAT	2.5°-7.5°S	97.5°-102.5°E	6 (3 x 2)	7.82594°S-3.35397°N	97.50°-101.25°E
SXW	2.5°-5.0°S	62.5°-67.5°E	4 (2 x 2)	5.58995°-3.35397°S	63.75°-67.50°E
SYW	20.0°-22.5°N	85.0°E	3 (3 x 1)	19.00583°-23.47778°N	86.25°E
RAIN	15.30°-19.36°N	98.05°-100.70°E	6 (3 x 2)	14.53387°-19.00583°N	97.50°-101.25°E
FMA rainfall					
SAT	7.5°S	110.0°-112.5°E	2 (1 x 2)	7.82594°N	108.75°-112.50°E
SLP	15.0°N	187.5°-192.5°W	2 (1 x 2)	14.53387°N	187.50°-191.25°W
SXW	17.5°N	140.0°-150.0°E	4 (1 x 4)	16.76985°N	138.75°-150.00°E
SYW	2.5°S-2.5°N	95.0°-97.5°E	8 (4 x 2)	3.35397°S-3.35397°N	93.75°-97.50°E
RAIN	15.30°-	98.05°-100.70°E	6 (3 x 2)	14.53387°-	101.25°E

	19.36°N			19.00583°N	

MJJ: May-June-July; ASO: August-September-October; NDJ: November-December-January; FMA: February-March-April.

SAT: surface air temperature; SLP: sea level pressure; SXW: surface zonal or latitudinal wind; SYW: surface meridian or longitudinal wind; RAIN: observed rainfall over the study basin.

4.3.2 Model performance of the GFDL-R30

To evaluate the performance of GFDL-R30, criteria such as annual statistics, the coefficient of determination (R^2) and the normalized root mean square error (NRMSE) are adopted. The monthly data from 1961 to 2007 of the observed and modeled predictors obtained from GFDL-R30 are used to estimate all the criteria. The monthly observed data consist of two data sets: (i) observed SAT, SLP, SXW and SYW provided by NCEP/NOAA; and (ii) observed RAIN obtained from historical data averaged over 50 selected stations. The annual statistics adopted to evaluate the performance of the GFDL-R30 are the arithmetic mean ($\overline{x}$) and standard deviation (SD), as shown in Equation 4.4 and 4.5 respectively.

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad \text{Equation 4.4}$$

$$SD = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2} \qquad \text{Equation 4.5}$$

where x_i is the time series of monthly data from 1961 to 2007, and n is the total amount of data.

The confidence intervals (CI) for $\overline{x}$ and SD of the observed data are calculated using Equation 4.6 and 4.7 respectively. A good performance of GFDL-R30 on capturing the statistics of historical data can be presented by $\overline{x}$ and SD estimated from the modeled data falling into the upper and lower bounds of CI for $\overline{x}$ and SD of the observed data respectively.

$$CI_{\overline{x}} = \overline{x} \pm \frac{t_{(\alpha/2,n-1)}SD}{\sqrt{n}} \qquad \text{Equation 4.6}$$

$$CI_{SD,upper} = SD\sqrt{\frac{n-1}{\chi^2_{(\alpha/2,n-1)}}} \qquad \text{Equation 4.7(a)}$$

$$CI_{SD,lower} = SD\sqrt{\frac{n-1}{\chi^2_{(1-\alpha/2,n-1)}}} \qquad \text{Equation 4.7(b)}$$

where α is the significance level – i.e. 95% in this case, $t_{(\alpha/2,n-1)}$ is the critical value from *t*-distribution with *n-1* degree of freedom, and $\chi^2_{(\alpha/2,n-1)}$ and $\chi^2_{(1-\alpha/2,n-1)}$ are the upper and lower critical values respectively of the Chi Squared distribution with *n-1* degree of freedom.

Furthermore, R^2 and NRMSE are estimated using Equation 4.8 and 4.9 respectively. R^2 is a method to measure the goodness-of-fit of a model. It presents the proportion of the variation of one data set (e.g. modeled results) over the variation of another set (e.g. observed data). R^2 varies from 0 to +1.0. On the other hand, NRMSE stands for the global error or residuals of the model simulation and is calculated by measuring the difference between simulated results and observations. An R^2 value tending towards +1.0 and a small NRMSE indicate the good performance of a model.

$$R^2 = \left(\frac{n\sum_{i=1}^{n} x_i y_i - (\sum_{i=1}^{n} x_i)(\sum_{i=1}^{n} y_i)}{\sqrt{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2}\sqrt{n\sum_{i=1}^{n} y_i^2 - (\sum_{i=1}^{n} y_i)^2}} \right)^2 \qquad \text{Equation 4.8}$$

$$NRMSE = \frac{\sqrt{\frac{\sum_{i=1}^{n} (y_i - x_i)^2}{n}}}{x_{\max} - x_{\min}} \qquad \text{Equation 4.9}$$

where x_i and y_i are observed and modeled data respectively, and *n* is the number of pairs of data. x_{max} and x_{min} are, respectively, the maximum and minimum values of observed data.

From Figure 4.6(a), the $\overline{x}$ of 1961-2007 simulated variables by GFDL-R30 under A2 and B2 falls outside the CI for $\overline{x}$ of the observed data except when the $\overline{x}$ of SYW is associated with the predictor of ASO rainfall – i.e. SYW over the Andaman Sea.

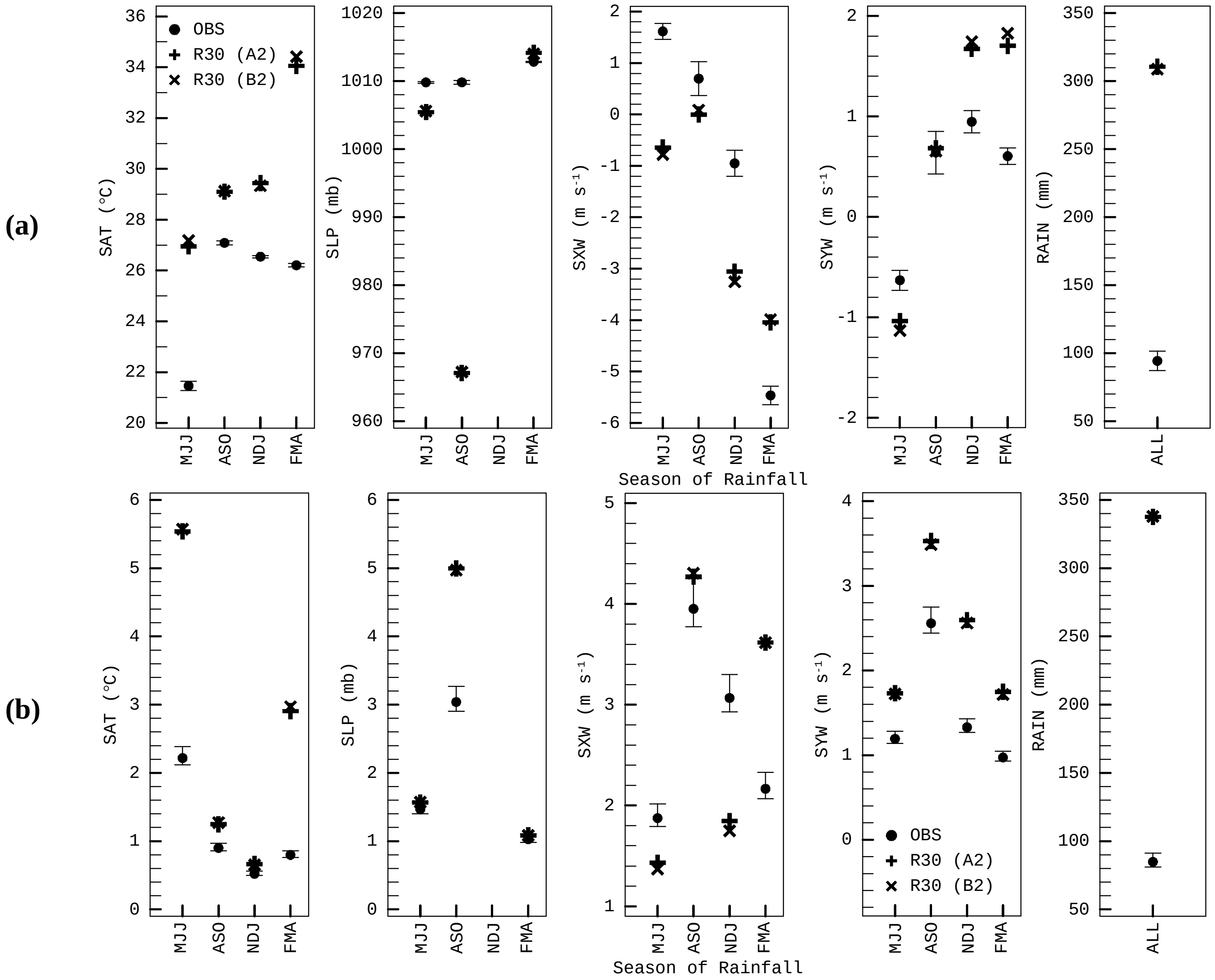


Figure 4.6: Annual statistics of the observed and modeled data (a) $\overline{x}$; and (b) SD. The vertical lines extending from $\overline{x}$ and SD of observed data represent the upper and lower bounds of CI.

Based on significant lag autocorrelations shown in Figure 4.5, RAIN, which is the average rainfall over the study basin, is also selected to be a predictor of rainfall. The GFDL-R30 gives poor performance in simulating RAIN under A2 (B2) with a mean estimated 310.5 (308.7) mm per month compared to a mean of 94.3 mm per month of the 1961-2007 observed rainfall. In terms of SD (Figure 4.6(b)), the modeled SLP under A2 and B2 associated with the predictors of MJJ and FMA rainfall can capture the SD of 1961-2007 observed SLP. The SD of simulated RAIN, which is estimated at 337.6 (337.8) mm corresponding to A2 (B2), exhibits a large difference from the SD of historical data, i.e. 84.8 mm.

Based on the maximum R^2 of SAT, SLP, SXW and SYW under A2 and B2 (Figure 4.7), a better performance of GDFL-R30 is associated with the predictors of ASO rainfall. Under A2 (B2), the R^2 corresponding to these predictors of ASO rainfall indicates that 70-78% (67-78%) of the observed data can be explained by the modeled results from GFDL-R30. However, the minimum R^2 of SAT, SLP and SYW, ranging from 3 to 8%, can be found corresponding to the predictors of FMA rainfall, whereas the minimum R^2 of SXW is associated with the predictor of MJJ rainfall which is identified over the equatorial Indian Ocean. The R^2 of RAIN indicates that only 47% (46%) of historical rainfall can be explained by GFDL-R30 simulation under A2 (B2).

As expected, minimum NRMSE is consistent with maximum R^2. The minimum NRMSE of SAT, SXW and SYW under A2 (Figure 4.8(a)), varying from 0.16 to 0.49, corresponds to the ASO rainfall predictors. However, the minimum NRMSE of SLP under A2, estimated at 0.28, is associated with the predictor of FMA rainfall.

Moreover, from Figure 4.8(b), the minimum NRMSE from B2 is consistent with the minimum NRMSE from A2, except the minimum NRMSE of SYW which is associated with the predictor of MJJ rainfall. NRMSE presents the error of the GFDL-R30 on RAIN simulation by 1.03 under both A2 and B2.

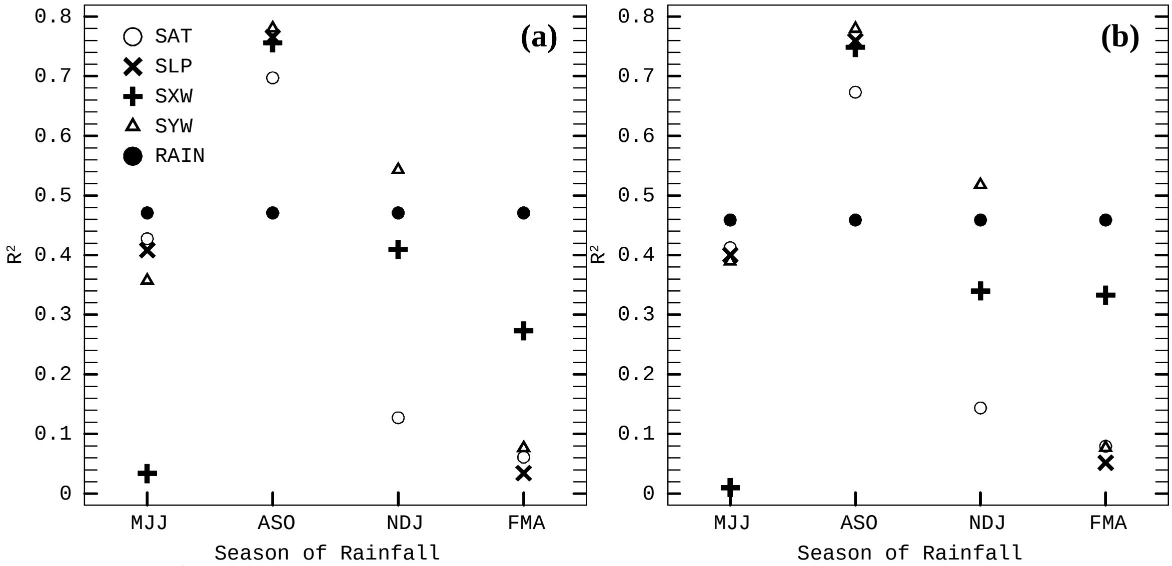


Figure 4.7: R^2 between observed and GFDL-R30 data under (a) A2; and (b) B2.

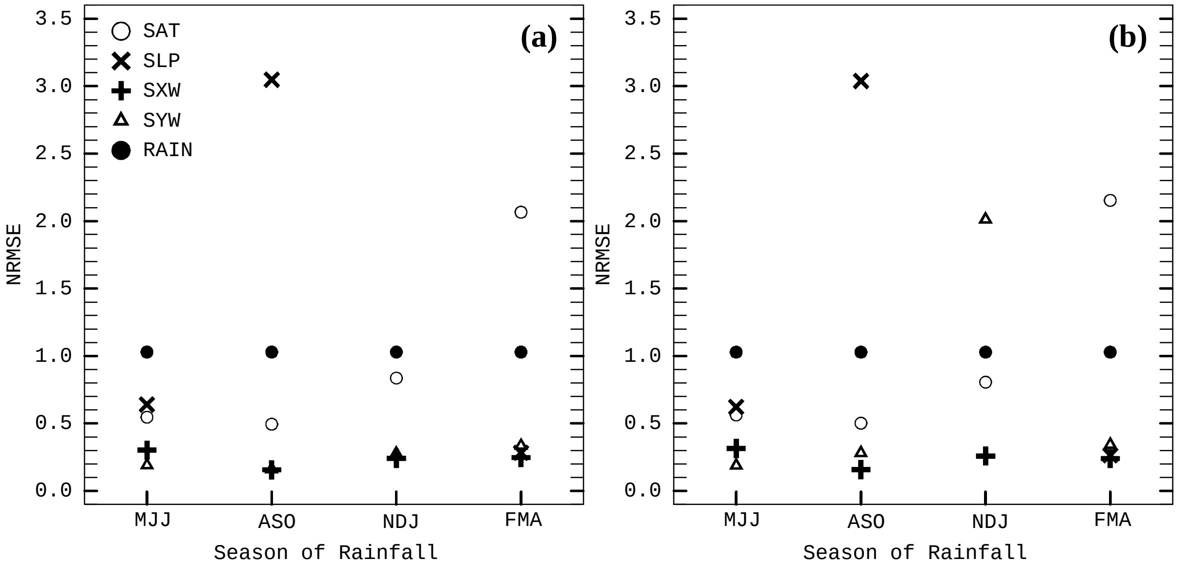


Figure 4.8: NRMSE of GFDL-R30 under (a) A2; and (b) B2.

To summarize, the simulated data of identified predictors from 1961 to 2100 can be collected from a GCM named GFDL-R30. GFDL-R30 is a coupled AOGCM which runs under two scenarios of future climate: A2 and B2. To evaluate the performance of the GFDL-R30, the observed and simulated data from 1961 to 2007 are used to calculate the annual statistics (i.e. mean and standard deviation), R^2 and NRMSE. The model lacks the capability to capture the means and standard deviations of observed atmospheric variables over the identified regions. However, the modeled results associated with the predictors of ASO rainfall (e.g. SAT over the South China Sea and SLP over northern Thailand) present maximum R^2 and minimum NRMSE when these values are compared among the predictors of rainfall during other seasons. 67 to 78% of the observed data of the ASO rainfall predictors can be explained by the GFDL-R30 data. Furthermore, as expected, GFDL-R30 shows better performance in simulating the atmospheric variables at the upper air level (e.g. temperature) than those at the surface level (e.g. rainfall). Under both scenarios, the GFDL-R30 presents large differences in means and standard deviations between observed and simulated rainfall over the study basin. Only 46 to 47% of historical rainfall can be explained by the simulated rainfall of GFDL-R30.

4.3.3 Annual and decadal variability of the predictors

The variability of large-scale atmospheric predictors is determined using the observed data from 1948 to 2007 and the projected data from 2011 to 2100 simulated by the GFDL-R30 under Scenario A2 and B2. For the predictors of MJJ rainfall (Figure 4.9), the annual observed SAT anomalies estimated with respect to the observed SAT values averaged from 1961 to 1990 show that during 1990s, the surface temperature over the identified region (i.e. northern Thailand) was warmer than in the earlier century.

From 1948 to 2007, the trend of the annual observed SAT is estimated at +0.0041°C per year or +0.41°C per century. It is less than the trend in global surface temperature, which indicates a range from +0.01 to +0.02°C per year (IPCC, 2007b; Hansen et al., 2010; Jenkins et al., 2008). Moreover, the GFDL-R30 suggests that by the end of the 21st century, SAT over northern Thailand will be warmer by 2 to 5°C with an increasing linear trend of 3.47 °C (1.93°C) per century according to A2 (B2). Both trends are significant at a 99.9% confidence level by the standard *t*-test (Haan, 2002), as shown in Equation 4.10.

$$t = \frac{\hat{\beta} - \beta_0}{SE} \qquad \text{Equation 4.10}$$

where $\hat{\beta}$ is the slope of a linear trend or a fitting regression $(y_i = \alpha + \hat{\beta}x_i + \varepsilon_i)$, β_0 is a specific value for testing, i.e. 0 in this case, and SE is the standard error (ε_i) of the least-squares of the estimates.

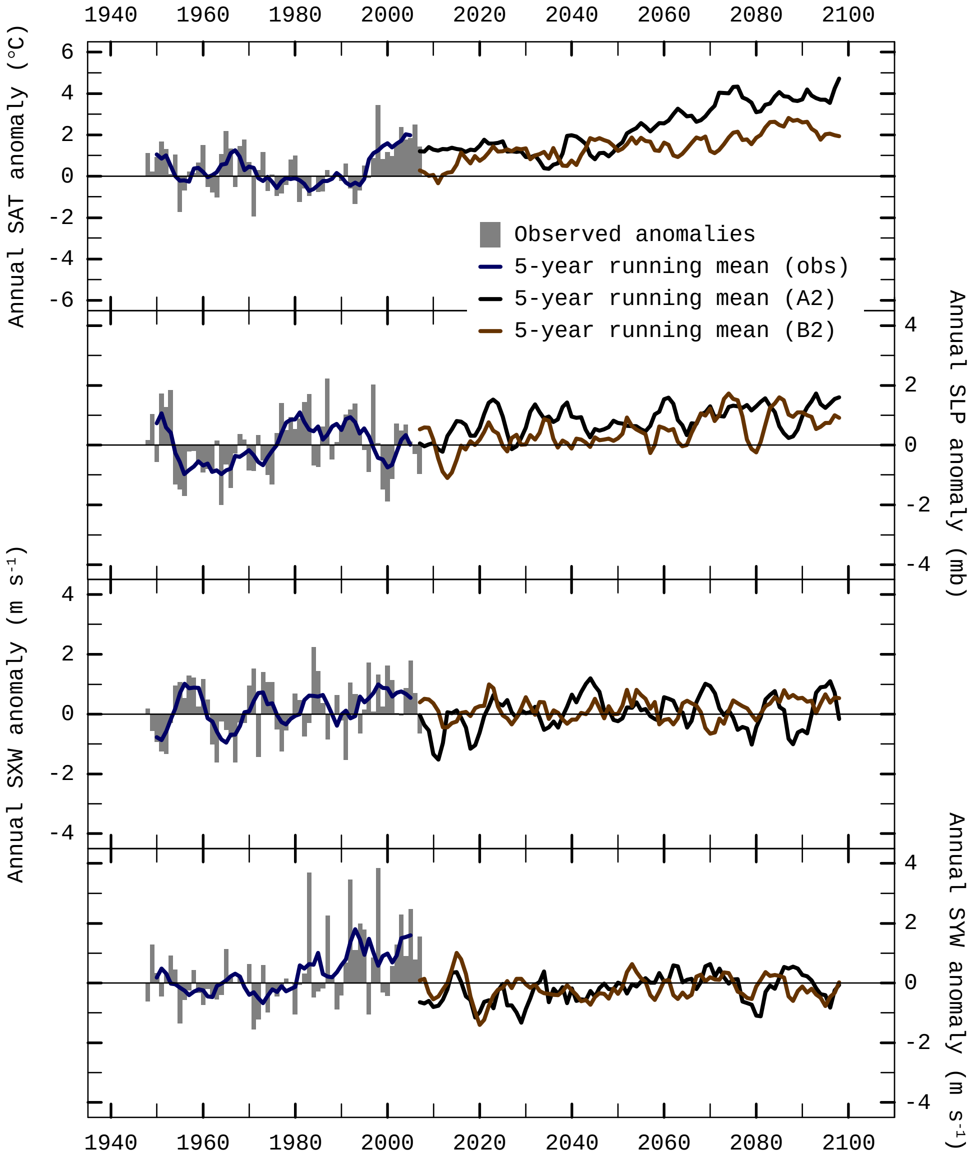


Figure 4.9: Annual anomalies of the four observed predictors of MJJ rainfall (1948-2007), and the 5-year running mean from observed and GFDL-R30 simulated data (2011-2100). The anomalies are estimated with respect to 1961-1990 observed or simulated average values.

The annual observed SLP identified over the Gulf of Thailand as the predictor of MJJ rainfall shows an increasing trend from 1948 to 2007 by 0.32 mb per century. The negative anomalies of SLP are found during pre-1980, whereas during post-1980, positive anomalies can be observed. In addition, projected SLP values by the GFDL-R30 tend to gradually increase with a significant rate at a 99% (99.5%) confidence level by 0.40 (0.54) mb per century under A2 (B2).

The identified regions for SXW and SYW are over the equatorial Indian Ocean and the eastern equatorial Pacific Ocean respectively. The increasing trends of observed SXW and SYW are estimated at 0.87 and 0.76 m s^{-1} per century respectively. These trends are also significant at 90 and 99.5% confidence levels. However, from GFDL-R30, the annual SXW under A2 (B2) from 2011 to 2100 shows slightly increasing trends by 0.03 (0.08) m s^{-1} per century and 0.18 (0.03) m s^{-1} per century for the annual SYW.

For the predictors of ASO rainfall (Figure 4.10), in the post-1980, the observed SAT, which is identified over the South China Sea, is warmer than in the pre-1980 period with an increasing trend of 0.0209°C per year or 2.09°C per century, and significant at a 99.9% confidence level. Due to changes in future climate, the projected SAT from GFDL-R30 shows that the temperature over the South China Sea from 2011 to 2100 will increase with linear trends of 2.57 and 1.65°C per century in A2 and B2 scenarios respectively. These trends are significant at a 99.9% confidence level. In addition, by the end of 21st century, the annual SAT over the South China Sea will be 9 to 10°C warmer than the 1961-1990 average annual temperature.

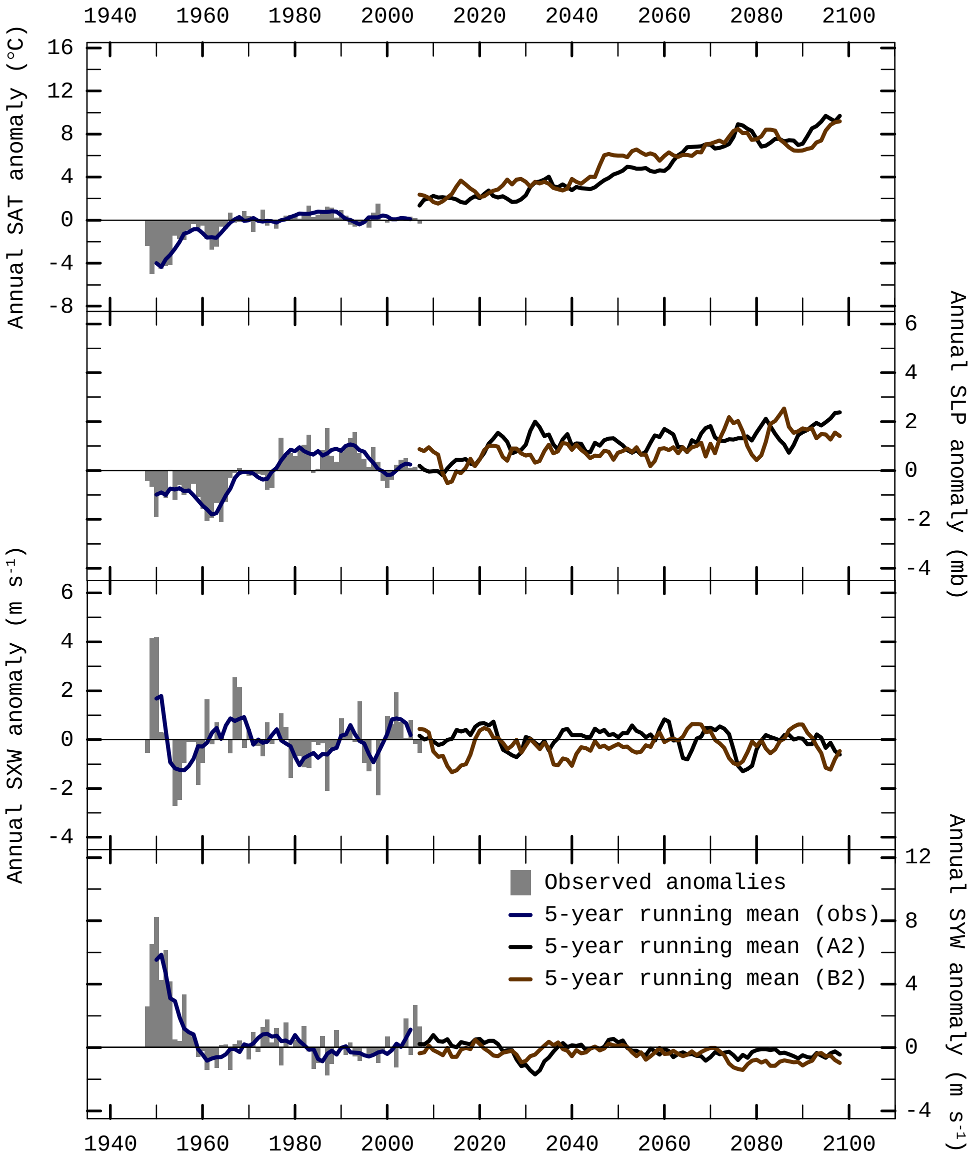


Figure 4.10: Same as Figure 4.9 but for the predictors of ASO rainfall.

For the SLP identified over northern Thailand, the annual observed data from 1948 to 2007 indicate an increasing trend of 3.03 mb per century, which is significant at a 99.9% confidence level. Below- (above-) normal SLP over northern Thailand found in pre- (post-) 1980 are consistent with the SLP anomalies over the Gulf of Thailand – i.e. the predictor of MJJ rainfall. From 2011 to 2100, an increasing trend of projected SLP under A2 (B2) from the GFDL-R30 is significant at a 99.9% (99.5%) confidence level with an increase of 0.83 (0.71) mb per century.

On the other hand, the observed SXW, which is identified over the Gulf of Thailand, shows a deceasing trend by 0.18 m s^{-1} per century. Likewise, the GFDL-R30 indicates that SXW will decrease (increase) with a linear trend of 0.14 (0.18) m s^{-1} per century with respect to A2 (B2). For the SYW over the Andaman Sea, a decreasing trend, estimated at 1.00 m s^{-1} per century with the significance at 99.9% confidence level, is calculated from observed data. SYW anomalies indicate that the annual winds during pre-1960 were stronger than post-1960. Under Scenario A2 (B2) of future climate, GFDL-R30 presents a significant decreasing trend at a 90% (99%) confidence level in SYW by 0.20 (0.29) m s^{-1} per century from 2011 to 2100.

The identified predictors of NDJ rainfall are SAT over the east coast of Sumatra (Indonesia), SXW over the Indian Ocean and SYW over northeastern India. From Figure 4.11, the annual observed SAT over the east coast of Sumatra from 1948 to 2007 shows an increase by a linear rate of 1.31°C per century; however, from 2011 to 2100, the increasing trends in SAT are estimated at 2.48 and 1.53°C per century corresponding to the simulated SAT from GFDL-R30 under A2 and B2 respectively. These trends in observed and modeled data are significant at 99.9% confidence levels. By the end of the 21st century, the air temperature over the east coast of Sumatra will become warmer than the average annual temperature recorded from 1961 to 1990 by 5-9°C, and this increase will be caused by the doubling of atmospheric CO_2 concentration.

An increasing trend of 0.85 m s^{-1} per century (significant at 99% confidence levels) in SXW over the Indian Ocean over the period of 1948-2007 is also observed. From 2011 to 2100, the projected SXW derived from GFDL-R30 under A2 (B2) presents a slightly increasing trend by 0.26 (0.18) m s^{-1} per century. In terms of SYW over northeastern India, the annual observed SYW from 1948 to 2007 tends to decrease by 1.28 m s^{-1} per century, which is significant at a 99.9% confidence level. With respect to the 1961-1990 average annual SYW, the positive and negative anomalies from the observed SYW are exhibited during pre- and post-1980 respectively. From GFDL-R30, the projected SYW under A2 and B2 from 2011 to 2100 presents an increasing trend by 0.52 and 0.18 m s^{-1} per century respectively.

Figure 4.12 shows the variability and trends of predictors of FMA rainfall. The SAT is identified over a region of Java, Indonesia. The 1948-2007 observed temperature over this region tends to increase by 0.49°C per century. Under a condition of doubled atmospheric CO_2 concentration by 2100, the increasing trends in SAT during 2011 to 2100 are estimated at 4.27 and 3.09°C per century under A2 and B2 respectively. These trends are significant at 99.9% confidence levels. By 2100, the SAT over Java will be 4 to 6°C higher than 1961-1990 average temperature.

In terms of SLP over the western Pacific Ocean, a slightly increasing trend of 0.35 mb per century is observed in the historical data during 1948 to 2007. However, the 2011-2100 projected SLP from the GFDL-R30 tends to decrease by 0.37 and 0.28 mb per century corresponding to Scenario A2 and B2 respectively.

The SXW, which is identified over the eastern Pacific Ocean, shows negative anomalies with respect to the 1961-1990 average observed SXW. The 1948-2007 observed SXW presents a linear trend of +0.60 m s^{-1} per century. Under A2 (B2), the decreasing trend is estimated by 0.14 (0.20) m s^{-1} per century. The SYW of the FMA rainfall predictor is identified over the Indian Ocean. The decreasing trend of 1948-2007 observed SYW is calculated to be 0.53 m s^{-1} per century. In the period of 2011 to 2100, a decreasing trend is also found under A2 by 0.14 m s^{-1} per century. However, with respect to B2, the 2011-2100 simulated SYW over the Indian Ocean tends to slightly increase by 0.04 m s^{-1} per century.

In terms of RAIN (Figure 4.13) over the study basin, the annual observed RAIN obtained from 50 selected stations from 1950 to 2007 shows an annual variability with a decreasing trend of 2.42 mm per year. The negative anomalies estimated with respect to the 1961-1990 annual observed rainfall are exhibited during 1980 to 2000. Simulated RAIN from GFDL-R30 indicates that from 2011 to 2100, the rainfall over the study basin will tend to increase by 3.84 mm per year corresponding to A2, but a deceasing trend (at 0.31 mm per year) in annual rainfall will be observed under B2.

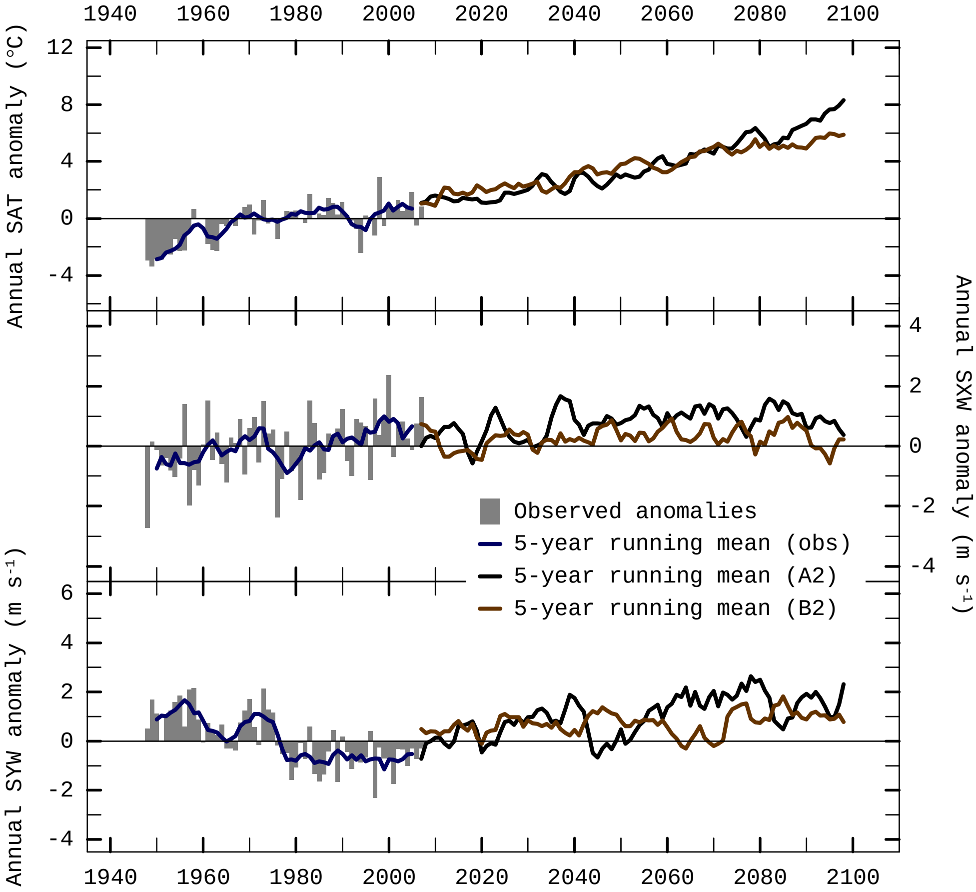


Figure 4.11: Same as Figure 4.9 but for the predictors of NDJ rainfall.

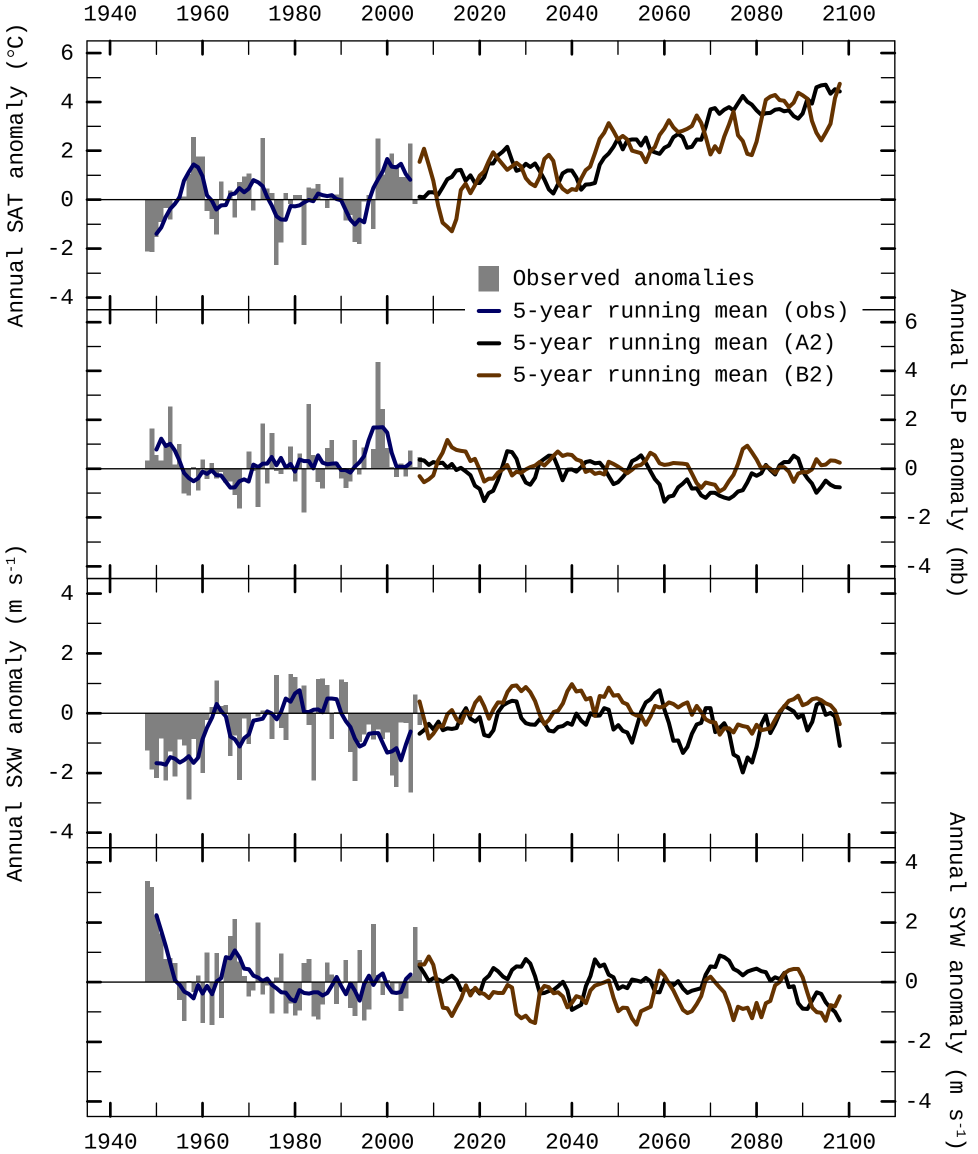


Figure 4.12: Same as Figure 4.9 but for the predictors of FMA rainfall.

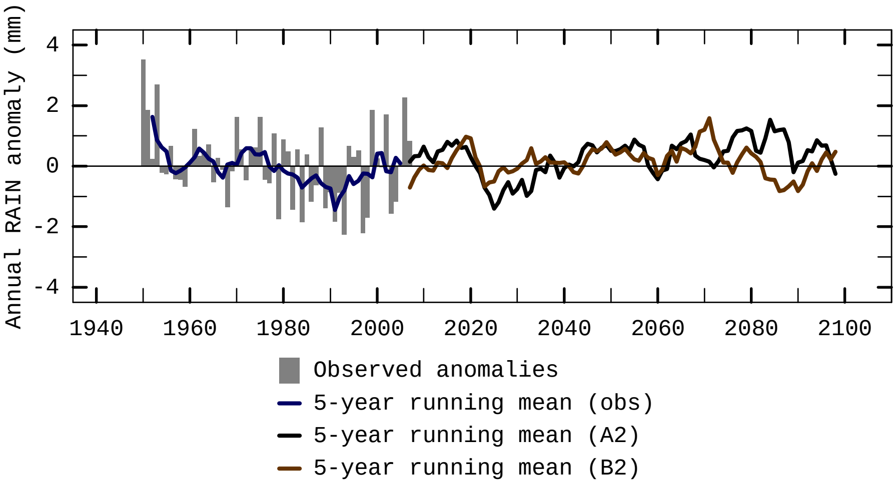


Figure 4.13: Annual anomalies of observed RAIN (1950-2007), and the 5-year running mean from observed and GFDL-R30 simulated data (2011-2100). The anomalies are estimated with respect to the period of 1961-1990 observed or simulated average values.

Therefore, the variability and trends of large-scale atmospheric variables, which are identified as the predictors of rainfall, are observed in the 1948-2007 historical data and in the 2011-2100 GFDL-R30 data. The 1948-2007 observed SAT over the four identified regions (i.e. northern Thailand for MJJ rainfall, the South China Sea for ASO rainfall, the east coast of Sumatra for NDJ rainfall and Java for FMA rainfall) tends to increase ranging from 0.41 to 2.09°C per century. From 2011 to 2100, the increasing trends in modeled SAT under A2 vary from 2.48 to 4.27°C per century and are observed with significance at 99.9% confidence levels. Under B2, the SAT over the four regions obtained from GFDL-R30 is also found to have significant trends at 99.9% confidence levels, and the SAT varies from 1.53 to 3.09°C per century. The 1948-2007 observed SLP over the Gulf of Thailand, northern Thailand and the western Pacific Ocean associated with the predictors of MJJ, ASO and FMA rainfall respectively, tends to increase from 0.32 to 3.03 mb per century. With respect to A2 and B2 scenarios, the increasing trends of simulated SLP between 2011 and 2100 over these regions is also observed, except for modeled SLP over the western Pacific Ocean – i.e. the predictor of FMA rainfall. The annual trends in observed and modeled SXW and SYW are inconsistent, and they depend upon the identified regions and scenarios of future climate. The selection of optimal combinations of predictors based on an objective function will be presented in next section.

4.3.4 Combination cases of the predictors

Based on significant relationships with rainfall, five predictors (i.e. SAT, SLP, SXW, SYW and RAIN) have been selected. The simulated data of atmospheric predictors from 1961 to 2100 are achieved from the GFDL-R30 associated with the identified regions of these predictors. To avoid the redundancy of predictors, an optimal combination is selected. The optimal subset is a predictor set which is composed of the minimum number of mutually exclusive variables. For *k* multiple independent variables, there are 2^k-1 combination cases in total. Hence, from five identified predictors, a total of 31 combination cases can be set. Among these 31 cases, one combination set is selected as the optimal subset of predictors using a criterion function like cross validation (CV), likelihood or the Akaike criterion (AIC). In this study, generalized cross validation (GCV) with the leave-one-out technique is applied to select the optimal subset. GCV estimates the error from a fitting regression following the equation shown below.

$$GCV = \frac{\sum_{i=1}^{n} \frac{(y_i - y_i')^2}{n}}{(1-m/n)^2} \qquad \text{Equation 4.11}$$

where y_i is observed data (i.e. seasonal rainfall in this case) at the dropped point (x_i) based on the leave-one-out technique, y_i' is the estimation from the fitting regression at the dropped point x_i, *n* is the total number of data, and *m* is the number of independent variables or predictors used to fit the regression.

In this case, GCV is calculated using the 1961-2007 simulated data by GFDL-R30 as the independent variable (*x*) of fitting regression and the observed rainfall averaged over 50 selected stations as the dependent variable (*y*). GCV estimation is done separately between Scenario A2 and B2. The GCV of each combination case of predictors is also calculated under a condition of varying lead periods of predictors from 4 to 15 months prior to the start of the rainfall season. Appendix D1 to D4 show the GCV scores associated with all the combination cases of predictors for MJJ, ASO, NDJ and FMA rainfall corresponding to

A2 and Appendix D5 to D8 show these scores corresponding to B2. Based on a minimum GCV, a best combination case is selected. Table 4.6 summarizes the selected optimal subset of predictors. According to A2, all selected optimal subsets consist of two predictors, except the selected combination case for NDJ rainfall, which has one predictor – SXW. The predictors of MJJ and ASO rainfall (i.e. monsoon rainfall) are identified with respect to lead times of 6 and 4 months respectively. However, the lead times of selected predictors for the dry season rainfall (i.e. NDJ and FMA rainfall) are longer than those of monsoon rainfall. They can be indicated 15 months prior to onset of the season. On the other hand, the selected optimal subsets of predictors associated with B2 are identified as cases with one predictor, except the selected subset for ASO rainfall which includes SAT and SLP. The long lead time of predictors (i.e. 14 months) is associated with the selected subset of ASO rainfall. In contrast, the short lead times are found corresponding to the predictors of MJJ and the dry season rainfall.

Table 4.6: Summary of the Optimal Subset of Predictors for Rainfall during MJJ, ASO, NDJ and FMA

Rainfall	Optimal subset of predictors	Season of predictors	Lead time (month)
A2			
MJJ	SLP and SXW	NDJ	6
ASO	SXW and SYW	AMJ	4
NDJ	SXW	ASO	15
FMA	SAT and SLP	NDJ	15
B2			
MJJ	SLP	NDJ	6
ASO	SAT and SLP	JJA	14
NDJ	SAT	MAM	8
FMA	SXW	ASO	6

SAT: surface air temperature; SLP: sea level pressure; SXW: surface zonal or latitudinal wind; SYW: surface meridian or longitudinal wind.

4.4 Summary

The predictors of MJJ, ASO, NDJ and FMA rainfall are identified by correlation maps based on significant linear relationships at long lead times with seasonal rainfall. The developed relationships show the influence of atmospheric circulations on local hydroclimates, e.g. rainfall, over the regions located nearby and distant from the source. The leading relationships of predictors suggest the predictability of a forecasting model. Five predictors each for MJJ, ASO and FMA rainfall, which include SAT, SLP, SXW, SYW and RAIN, are selected over different regions (namely the study basin, the South China Sea, the Pacific and the Indian Oceans). Moreover, there are four identified predictors for NDJ rainfall (i.e. SAT, SXW, XYW and RAIN). To determine the effects of future climate on rainfall, gridded monthly data from 1961 to 2100 is obtained of identified predictors, which are simulated by a GCM called the GFDL-R30 under the condition of doubled atmospheric CO_2 concentration by 2100, and the GFDL-R30 is used to develop a statistical model. The optimal subsets of predictors are selected using generalized cross validation with the leave-one-out technique. The selected subsets are composed of 1-2 predictors for lead times varying from 4-15 months.

Chapter 5
Development of a Statistical Downscaling Model to Simulate Rainfall

5.1 Introduction

The objectives of this chapter are to develop a statistical model to simulate and downscale seasonal rainfall from large-scale atmospheric predictors, and to develop a multisite daily rainfall generator to resample daily rainfall from historical data. The seasonal rainfall of the Upper Chao Phraya River Basin shows a significant relationship with the large-scale atmospheric variables presented in Chapter 4. The identified predictors at long-range lead times are selected based on their significant correlations with the regions at a 95% confidence level. Optimal combinations of predictors from GFDL-R30 are identified based on GCV. The methodology of the statistical model, which uses optimal combinations of predictors as independent variables, is described in 5.2.1, and the conditioning rainfall generator and the multisite daily rainfall generator are explained in sections 5.2.2 and 5.2.3 respectively. The performance of both models is evaluated by statistical criteria such as goodness-of-fit and likelihood skill score. The obtained results describe the effects of future climate on seasonal rainfall in the Ping River Basin.

5.2 Methodology

5.2.1 The modified k-nearest neighbor (k-nn) model

The nonparametric approach, which is a function to fit the relationship between dependent (*y*) and independent (*x*) variables, has been developed to improve the performance of fitting regression in the parametric approach. Although the parametric approach is widely adopted, due to the ease of use when fitting a regression between two variables, there are some drawbacks of this fitting regression. One drawback is that a prior assumption of relationships between two variables (i.e. linear regression) is required, which causes difficulty in fitting some arbitraries such as the bivariate and multivariate regressions. Another drawback is global fitting, in which all points of data are used to fit a regression. An individual point of regression gets heavily influenced by other points of data to minimize the residuals from fitting.

Nonparametric regression, on the other hand, does not carry these drawbacks. The nonparametric regression function is shown in Equation 5.1.

$$y = f(x_1, x_2, x_3, ..., x_k) + e \qquad \text{Equation 5.1}$$

where *f* is the regression function to fit independent variables (i.e. univariate or multivariate): x_1, x_2, x_3,…., x_k; *y* is the dependent variable; and *e* is the error or residual of the fitting regression which is assumed to be normally distributed with mean=0 and variance=σ.

Nonparametric regression does not require a prior assumption of relationship between two data sets. The fitting function (*f*) can locally capture the relationship using a small set of neighbors (*k*) at a given point (x_i). So, the function is flexible and able to describe the relationship better than parametric regression. Moreover, the drawback of parametric regression in terms of global fitting can be solved by the nonparametric approach using a small set of neighbors. There are several approaches of nonparametric regression. One, which has been developed for a derivative curve which contains discontinuities, is the spline approach. Another, which can locally apply a regression at a given point (x_i) of data and its neighbors, is called local polynomials. This approach includes locally weighted polynomials (Loader, 1999) and k-nearest neighbor (k-nn) local polynomials (Owosina, 1992; Rajagopalan and Lall, 1999). Two parameters (i.e. the size of the neighborhood (*k*) called bandwidth and the order of polynomials (*p*)) are required for the development of a fitting regression. A criteria objective method such as GCV and likelihood can be used to determine both parameters.

The steps of fitting regression and ensemble simulation of the modified k-nn model are described as follows:

1) For the fitting process, the size of the neighbors (*k*) and the order of the polynomial (*p*) (Figure 5.1), which is normally 1 or 2, are selected and associated with the combination of *k* and *p* so as to obtain minimum GCV. The GCV is estimated by Equation 5.2.

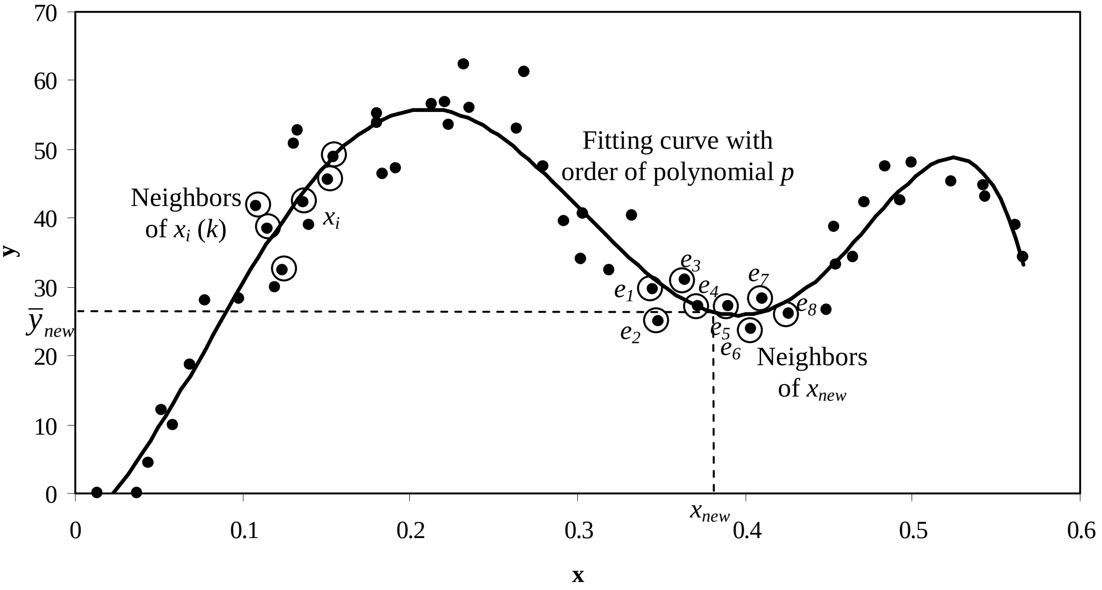


Figure 5.1: Fitting process and ensemble simulation of the modified k-nn model.

$$GCV(k,p) = \frac{\sum_{i=1}^{n} \frac{e_i^2}{n}}{(1-m/n)^2} \qquad \text{Equation 5.2}$$

where e_i is the error from the developed regression using *k* and *p*, *n* is the number of data points, and *m* is the number of parameters.

2) The regression is locally fitted with the obtained *k* and *p*.
3) The dependent variables (*y*) according to the developed fitting regression are then estimated and called mean estimations ($\bar{y}_1, \bar{y}_2, \bar{y}_3, ..., \bar{y}_n$). Then, the residuals ($e_1, e_2, e_3, ..., e_n$) are computed.
4) The simulation or forecast of a dependent variable is required at each new point of an independent variable (x_{new}). The mean estimation ($\bar{y}_{new}$) is calculated from the developed regression.
5) A simulating ensemble is obtained by adding a residual (e_i) to $\bar{y}_{new}$. The residual (e_i) is associated with one of the k-nearest neighbors (k-nn) of x_{new} which is randomly selected using a weight function, presented in Equation 5.3.

$$W(j) = \frac{1/j}{\sum_{i=1}^{k}(1/i)} \qquad \text{Equation 5.3}$$

where $W(j)$ is a weight of a neighbor of x_{new} and its distance from x_{new} falls in the j^{th} rank, and k is the size of neighbors which can be different from k for the fitting process. The formula $\sqrt{n-1}$ is, in practice, used to estimate k where n is the total number of x_i. It is also noted from Equation 5.3 that the nearest neighbor has more weight, and the farthest neighbor has less weight. Moreover, the distance between x_{new} and all the points of x_i needs to be estimated. There are several methods to calculate the distance between two points of data such as the Euclidean distance and the Mahalanobis distance. In this study, Euclidean distance has been adopted (Equation 5.4(a) and 5.4(b)) for its simplicity.

For univariate data; $d_i = \sqrt{(x_{new} - x_i)^2}$ Equation 5.4(a)

For multivariate data; $d_i = \sqrt{\sum_{j=1}^{m}(x_{new,j} - x_{i,j})^2}$ Equation 5.4(b)

where i=1, 2, 3, …, n, and m is the number of independent variables.

6) Repeat step 5) as many times as required to achieve the N ensembles or a number of simulations.
7) Repeat step 4) to 6) for each simulation point (x_{new}).

In this study, the dependent variable (y) is the rainfall averaged over 50 selected stations in the Ping River Basin during the pre-monsoon (MJJ), monsoon (ASO) and dry seasons (i.e. NDJ and FMA). The independent variables (x) are the optimal combinations of predictors during the seasons, as shown in Table 4.6, and are univariate (e.g. SXW predictor for NDJ rainfall under A2) or multivariate data (e.g. SLP and SXW predictors for MJJ rainfall under A2). Rainfall simulation is done separately for each rainfall season and for the two future climate scenarios: A2 and B2. Seasonal rainfall is simulated for the period of 2011 to 2100 in order to determine the effects of climate in the future. Furthermore, the 300 ensembles are simulated for each year in order to estimate a probability density function (PDF). The non-exceedence and exceedence probabilities of anomalous events (e.g. dry and wet) can be calculated from the PDF based on the defined threshold of events (described in 5.4.2). The probabilities of anomalous events are a useful tool for decision making for water resource planning, agricultural practices, reservoir operations and insurance policies.

5.2.2 The conditioning daily rainfall generator

The conditioning daily rainfall generator is a stochastic model aiming to generate or resample a series of daily observed rainfall which is used as the input for a multisite daily rainfall generator. Firstly, the historical seasonal rainfall averaged over 50 selected stations is divided into three categories based on the defined thresholds, which, in this case, are the 20th and 80th percentiles. The 20th (80th) percentiles of MJJ, ASO, NDJ and FMA rainfall from 1950 to 2007 are estimated at 381.7 (528.9), 498.2 (624.8), 15.8 (79.0) and 53.9 (114.5) mm respectively. Rainfall which is less than the 20th percentile falls into a category of below-normal rainfall or a dry condition (D). Rainfall which is greater than the 80th percentile is defined as above-normal rainfall or a wet condition (W). The rainfall which does not fall into either category is termed normal rainfall (N). Appendix B3 shows the classifications of historical rainfall from 1950 to 2007.

Historical rainfall is then randomly selected from three categories. Out of a number of total required samples (G), the number of samples selected from each category is associated with categorical probabilities, which are estimated from the PDF of N rainfall ensembles simulated by the modified k-nn model. Note that the categorical probabilities of ensembles are calculated with respect to the same thresholds – i.e. the 20th and 80th percentiles of historical seasonal rainfall. Then, a series of G samples of daily rainfall developed by the conditioning daily rainfall generator is used in the multisite daily rainfall generator. For example, in this study, a series of 100 samples (i.e. G=100) is developed separately for each rainfall season. Assuming that the categorical probabilities obtained from the PDF of the 300 MJJ rainfall ensembles are D:N:W=0.6:0.1:0.3, then 60, 10 and 30 samples out of the total 100 required samples are randomly selected with replacement from the dry, normal and wet categories respectively of historical MJJ rainfall. Hence, the 100 samples of daily rainfall during MJJ (i.e. 92 days: $[\mathbf{R}]_{ND\mathrm{x}G}=[\mathbf{R}]_{92\mathrm{x}100}$) are applied in the multisite daily rainfall generator. It is important to note that in this case, the daily rainfall in the developed series is the averaged value from the 50 selected stations. In other words, this is spatially averaged daily rainfall.

5.2.3 The multisite daily rainfall generator

The historical daily rainfall at several selected rainfall stations in the study basin is resampled by the multisite daily rainfall generator. The resampling is based on the state of rainfall, which is also bootstrapped by the model. The multisite daily rainfall generator has two components: (i) a Markov Chain to generate the state of daily rainfall (e.g. dry, wet and extremely wet) based on the transition probabilities of climatology; and (ii) a Monte Carlo approach to resample the amount of daily rainfall (i.e. spatially averaged values over several rainfall stations) associated with the rainfall state obtained from (i). Daily rainfall at multiple stations is also achieved from the multisite daily rainfall generator.

Firstly, the daily observed rainfall averaged over 50 selected stations (see also Figure 3.5) is divided into three states by the thresholds which can be defined by the users of the model. In this study, the thresholds are defined at the 10th and 90th percentiles in order to focus on anomalous events of daily rainfall. Table 5.1 presents the thresholds at the 10th and 90th percentiles which are estimated separately for each month using the daily observed data from 1950 to 2007. Daily rainfall less than the 10th percentile is defined as a dry (d) state, whereas rainfall in a single day greater than the 90th percentile is denoted as an extremely wet (e) state. Daily rainfall which does not fall under either of these states is defined as the wet (w) state.

Table 5.1: Defined Thresholds (mm d^{-1}) at the 10th and 90th Percentiles of the 1950-2007 Daily Rainfall

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
$P_{10}{}^{th}$	0.0	0.0	0.0	0.0	0.5	0.7	0.7	1.1	1.1	0.0	0.0	0.0
$P_{90}{}^{th}$	0.2	0.7	1.9	5.3	13.1	9.4	10.3	13.4	16.9	12.4	3.4	0.4

Then, the unconditional and transition probabilities of the three states are calculated. The unconditional probability of a state (i.e. P_d, P_w and P_e) is the proportion of daily rainfall falling into that state. Table 5.2 shows the unconditional probabilities of the three states (d, *w* and *e*) of the 1950-2007 daily observed rainfall based on the thresholds outlined in Table 5.1.

Table 5.2: Unconditional Probabilities of the Three States of the 1950-2007 Daily Rainfall

Month	Probability (P_i)			Month	Probability (P_i)		
	P_d	P_w	P_e		P_d	P_w	P_e
Jan	0.844	0.556	0.100	Jul	0.101	0.800	0.099
Feb	0.763	0.139	0.098	Aug	0.111	0.789	0.100
Mar	0.619	0.283	0.098	Sep	0.105	0.796	0.099
Apr	0.326	0.575	0.099	Oct	0.124	0.776	0.100
May	0.100	0.800	0.100	Nov	0.537	0.366	0.097
Jun	0.113	0.788	0.099	Dec	0.841	0.070	0.089

P_d: probability of dry state; P_w: probability of wet state; P_e: probability of extremely wet state.

The transition probability (i.e. $P_{i\text{-}j}$) from state *i* at the previous time step (*t-1*) to state *j* at the current time step (*t*) is estimated based on conditional probability and the total probability theorem. Conditional probability (i.e. $P(S_t|S_{t-1})$) is the probability of the state of a current time step (S_t) which occurs when a state of the previous time step (S_{t-1}) has also occurred (Equation 5.5). If there are three states each of previous (*t-1*) and current (*t*) time steps (i.e. dry (*d*), wet (*w*) and extremely wet (*e*)), the transition probability of each state at a current time step (Equation 5.6) can be calculated using the total probability theorem. Note that transition probability can be estimated for each month as shown in this study (Table 5.3), or for other temporal scales like a week, bi-weekly, or a season.

$$P(S_t|S_{t-1}) = \frac{P(S_t \cap S_{t-1})}{P(S_{t-1})} \text{ for } P(S_{t-1}) > 0 \qquad \text{Equation 5.5}$$

$$\begin{aligned} P(S_{t,d}|(S_{t-1,d} \cup S_{t-1,w} \cup S_{t-1,e})) &= P(S_{t,d}|S_{t-1,d}) * P(S_{t-1,d}) + \\ P(S_{t,d}|S_{t-1,w}) * P(S_{t-1,w}) &+ P(S_{t,d}|S_{t-1,e}) * P(S_{t-1,e}) \end{aligned} \qquad \text{Equation 5.6(a)}$$

$$\begin{aligned} P(S_{t,w}|(S_{t-1,d} \cup S_{t-1,w} \cup S_{t-1,e})) &= P(S_{t,w}|S_{t-1,d}) * P(S_{t-1,d}) + \\ P(S_{t,w}|S_{t-1,w}) * P(S_{t-1,w}) &+ P(S_{t,w}|S_{t-1,e}) * P(S_{t-1,e}) \end{aligned} \qquad \text{Equation 5.6(b)}$$

$$\begin{aligned} P(S_{t,e}|(S_{t-1,d} \cup S_{t-1,w} \cup S_{t-1,e})) &= P(S_{t,e}|S_{t-1,d}) * P(S_{t-1,d}) + \\ P(S_{t,e}|S_{t-1,w}) * P(S_{t-1,w}) &+ P(S_{t,e}|S_{t-1,e}) * P(S_{t-1,e}) \end{aligned} \qquad \text{Equation 5.6(c)}$$

Subsequently, a state of daily rainfall is generated by the three-state, first-order Markov Chain. Figure 5.2 shows the schematics of a Markov Chain. The steps require generating a state of daily rainfall are:

1) A series of uniform random number from 0 to 1 are generated (Q_i where *i*=1, 2, 3, …, *ND*, and *ND* is the number of days required for generating a state. For example, *ND*=92 days for a series of daily rainfall during the pre-monsoon season (MJJ).
2) The rainfall state of day 1 (S_1) is dry, corresponding to $Q_1 \le P_d$ Table 5.2 or wet, corresponding to $P_d < Q_1 \le (P_d + P_w)$, or extremely wet otherwise. For example, if a

rainfall state in the month of May is required and Q_1=0.580, then S_1 is wet since 0.100<Q_1≤0.900.

3) The rainfall state of day 2 (S_2) depends on S_1 and Q_2. As seen in Table 5.3, if S_1 is dry, then $P_{d\text{-}d}$, $P_{d\text{-}w}$ and $P_{d\text{-}e}$ are compared to Q_2. If $Q_2 \leq P_{d\text{-}d}$, then S_2 is dry, or if $P_{d\text{-}d} < Q_2 \leq (P_{d\text{-}d}+P_{d\text{-}w})$, then S_2 is wet. Otherwise, S_2 is extremely wet. On the other hand, if S_1 is wet, then $P_{w\text{-}d}$, $P_{w\text{-}w}$ and $P_{w\text{-}e}$ are compared to Q_2. If $Q_2 \leq P_{w\text{-}d}$, then S_2 is dry, or if $P_{w\text{-}d} < Q_2 \leq (P_{w\text{-}d}+P_{w\text{-}w})$, then S_2 is wet. Otherwise, S_2 is extremely wet. Likewise, if S_1 is extremely wet, then $P_{e\text{-}d}$, $P_{e\text{-}w}$ and $P_{e\text{-}e}$ are compared to Q_2. If $Q_2 \leq P_{e\text{-}d}$, then S_2 is dry. If $P_{e\text{-}d} < Q_2 \leq (P_{e\text{-}d}+P_{e\text{-}w})$, then S_2 is wet; otherwise, S_2 is extremely wet.
4) Repeat step 3) to generate rainfall states for a required sequence of days (S_1, S_2, S_3,…, S_{ND}).
5) Repeat step 1) to 4) for G simulations ($[\mathbf{S}]_{NDxG}$), i.e. 100 simulations in this case.

Table 5.3: Transition Probabilities of the Three States of the 1950-2007 Daily Rainfall

Transition probability (P_{i-j})				Transition probability (P_{i-j})			
State	*d*	*w*	*e*	State	*d*	*w*	*e*
d	$P_{d\text{-}d}$	$P_{d\text{-}w}$	$P_{d\text{-}e}$	*d*	$P_{d\text{-}d}$	$P_{d\text{-}w}$	$P_{d\text{-}e}$
w	$P_{w\text{-}d}$	$P_{w\text{-}w}$	$P_{w\text{-}e}$	*w*	$P_{w\text{-}d}$	$P_{w\text{-}w}$	$P_{w\text{-}e}$
e	$P_{e\text{-}d}$	$P_{e\text{-}w}$	$P_{e\text{-}e}$	*e*	$P_{e\text{-}d}$	$P_{e\text{-}w}$	$P_{e\text{-}e}$
Jan	0.924	0.030	0.046	Jul	0.380	0.585	0.035
	0.546	0.212	0.243		0.079	0.832	0.088
	0.324	0.145	0.531		0.035	0.630	0.335
Feb	0.870	0.086	0.044	Aug	0.336	0.644	0.020
	0.540	0.284	0.176		0.096	0.816	0.088
	0.274	0.294	0.432		0.022	0.646	0.333
Mar	0.795	0.168	0.037	Sep	0.410	0.548	0.042
	0.430	0.439	0.131		0.078	0.832	0.090
	0.143	0.451	0.406		0.030	0.656	0.314
Apr	0.675	0.302	0.023	Oct	0.587	0.344	0.070
	0.197	0.700	0.103		0.060	0.865	0.075
	0.056	0.637	0.307		0.069	0.555	0.377
May	0.401	0.566	0.034	Nov	0.766	0.218	0.016
	0.074	0.851	0.074		0.334	0.584	0.083
	0.016	0.622	0.362		0.045	0.442	0.513
Jun	0.301	0.650	0.050	Dec	0.933	0.043	0.024
	0.100	0.818	0.081		0.574	0.158	0.269
	0.055	0.589	0.357		0.295	0.196	0.509

$P_{i\text{-}j}$: transition probability from state *i* at the previous time step to state *j* at the current time step; *d*: dry state; *w*: wet state; *e*: extremely wet state.

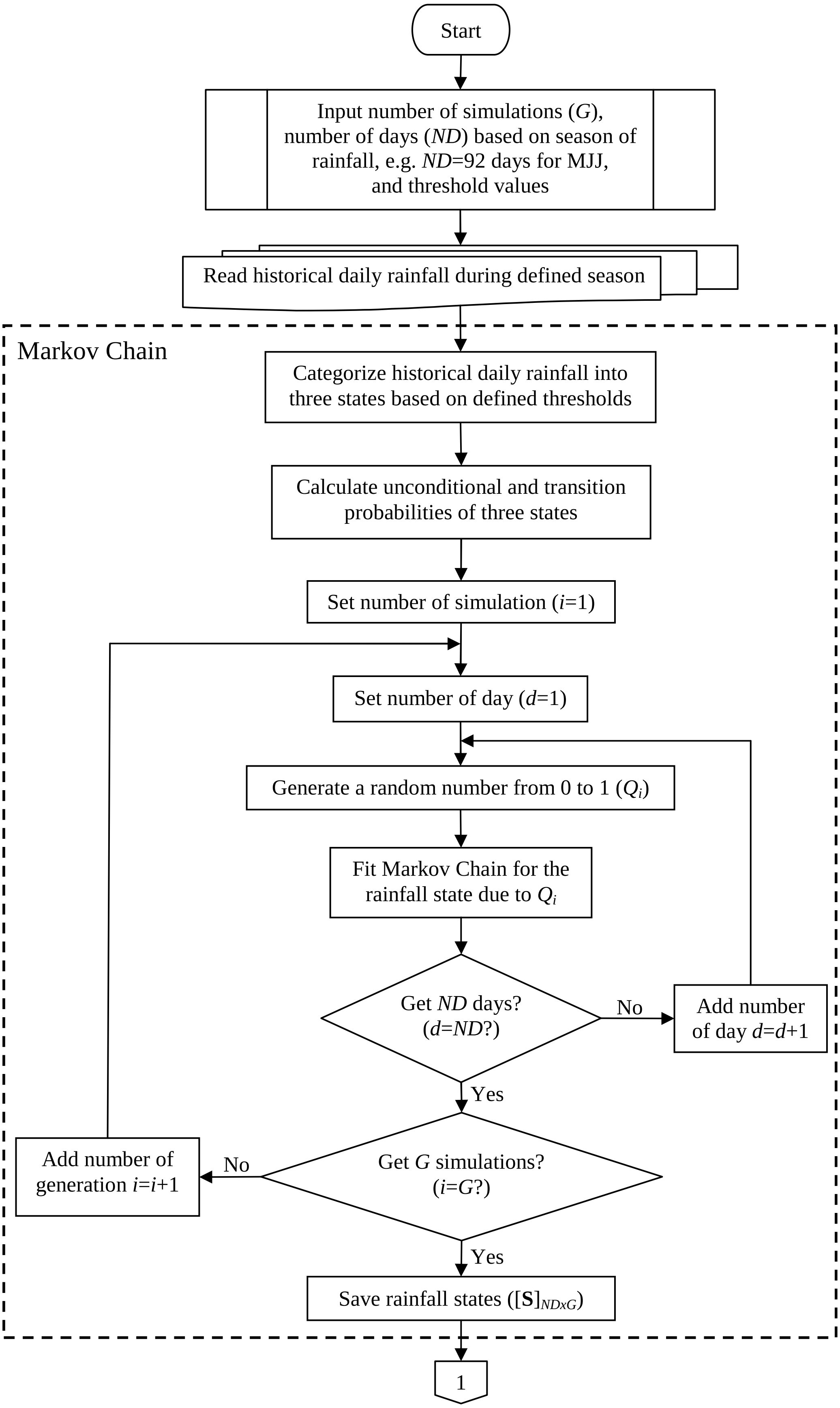


Figure 5.2: Schematic of the multisite daily rainfall generator.

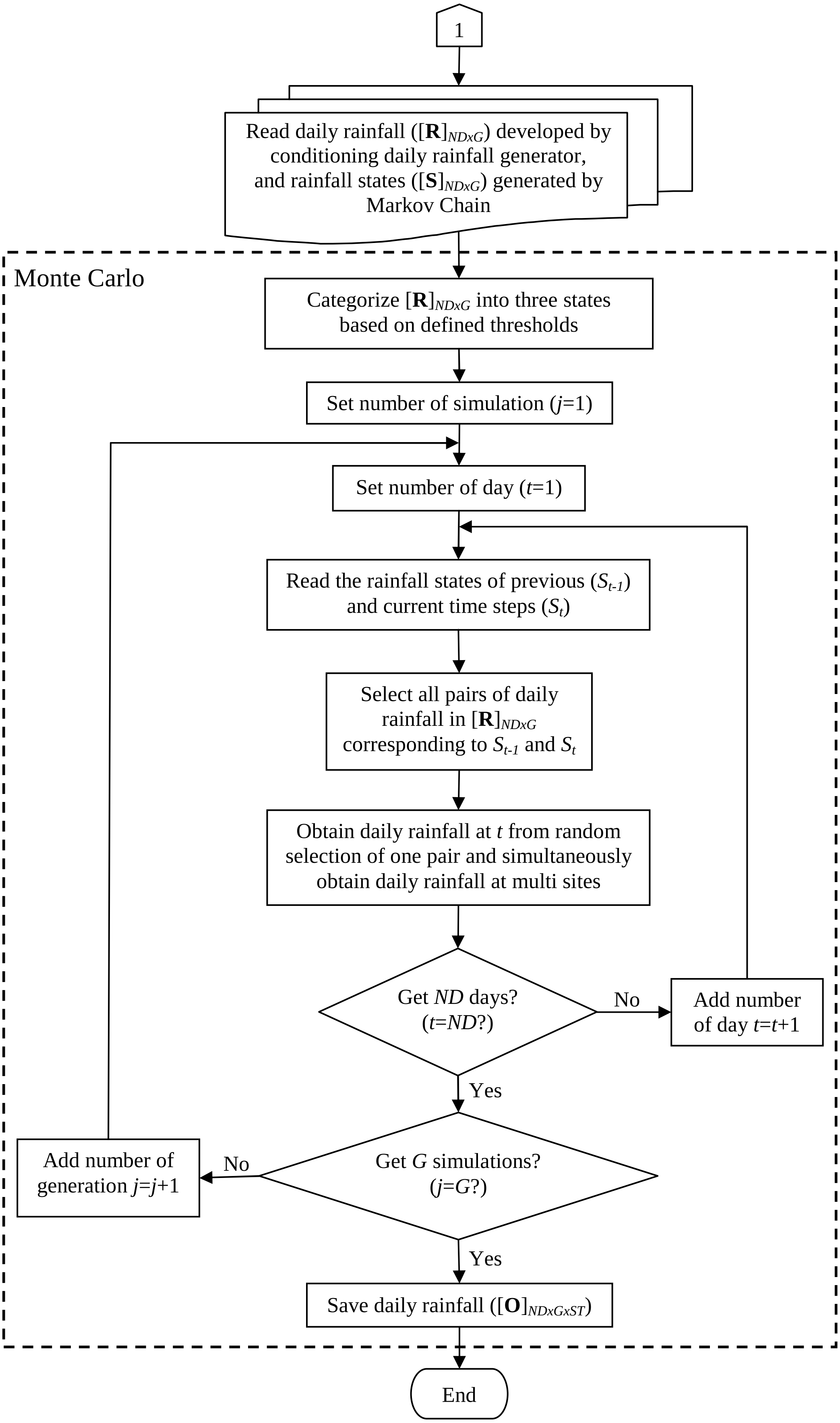


Figure 5.2 (cont).

After obtaining a series of rainfall states from the Markov Chain, the amount of daily rainfall associated with the generated states are bootstrapped using the Monte Carlo approach. The algorithm of Monte Carlo is presented in Figure 5.2. It is important to note that the input of Monte Carlo leads to a series of spatially averaged daily rainfall ($[\mathbf{R}]_{ND\text{x}G}$) developed by the conditioning daily rainfall generator (described in 5.2.2). So, the initial results from Monte Carlo are the averaged daily rainfall over the 50 selected rainfall stations. The amount of daily rainfall at each station is then obtained based on the following algorithm.

1) The daily rainfall developed by the conditioning daily rainfall generator ($[\mathbf{R}]_{ND\text{x}G}$) is defined under a rainfall state based on the thresholds outlined in Table 5.1.
2) The amount of rainfall at a current time step (O_t) is bootstrapped corresponding to the states of previous (S_{t-1}) and current (S_t) time steps which are obtained by the Markov Chain. For examples, the states of previous and current time steps from the Markov Chain indicate wet and dry (*w-d*) respectively. All pairs of daily rainfall in $[\mathbf{R}]_{ND\text{x}G}$ which also indicate the state as *w-d* have been chosen. One pair is randomly selected, and O_t is obtained corresponding to the selected pair.
3) The daily rainfall at all stations (*ST*), i.e. the 50 selected stations in this case, is simultaneously obtained corresponding to the historical day of averaged daily rainfall selected in step 2).
4) Repeat step 2) and 3) to bootstrap the amount of daily rainfall for a required sequence of days (O_1, O_2, O_3,.., O_{ND}) at multiple stations.
5) Repeat step 2) to 4) for *G* simulations ($[\mathbf{O}]_{NDxGxST}$), i.e. 100 simulations in this case.

5.3 Evaluation of the model performance

5.3.1 The modified k-nn model

The modified k-nn model is developed and evaluated from 1962 to 2007 using identified predictors (from the GFDL-R30 model under Scenario A2 and B2) as independent variables. Using the leave-one-out cross validation, the modified k-nn model is evaluated separately for each rainfall season (i.e. MJJ, ASO, NDJ and FMA) and each future climate scenario. Based on the leave-one-out cross validation, one pair of observations is dropped out from the data set of dependent and independent variables. Then, the regression is fitted using the remaining data. Using the developed regression, the rainfall at the dropped point is estimated. The leave-one-out cross validation is applied at all points of observation for the duration of 1962 to 2007 (i.e. 46 years). Note that there are 300 simulations of each year obtained from the modified k-nn model. The criteria used to evaluate the performance of the modified k-nn model are (i) the annual variability of seasonal rainfall; (ii) the annual statistics of seasonal rainfall including mean, median, and standard deviation (SD), interquartile range (IQR) and coefficient of skew (skew); (iii) the absolute bias; and (iv) the likelihood skill score (LLH).

Under A2 and B2, the annual variability of observed rainfall from 1962 to 2007, along with the box plots of rainfall ensembles, are presented separately for MJJ, ASO, NDJ and FMA rainfall. The box plot of each year is estimated from the 300 simulated members obtained from the modified k-nn model. A box is defined by the quartile range of ensembles between the upper quartile (QU) or the 75th percentile, and the lower quartile (QL) or the 25th percentile. The median value is presented by the horizontal line within each box. The caps of upper and lower whiskers indicate outlier values as calculated by Equation 5.7. The solid lines with marks represent the observed values.

$$cap_{upper} = QU + 1.5 * (QU - QL) \qquad \text{Equation 5.7(a)}$$

$$cap_{lower} = QL - 1.5 * (QU - QL)$$ Equation 5.7(b)

The annual variability of observed MJJ rainfall from 1962 to 2007 ranges from 254.0 to 631.3 mm, with the minimum and maximum found in 1997 and 1978 respectively. The annual variability of observed ASO rainfall varies from 387.6 to 715.4 mm. The minimum ASO rainfall is observed in 2004, whereas the maximum was seen in 1962. For the observed NDJ (FMA) rainfall, the minimum and maximum were 0.0 mm in 1979 and 184.4 mm in 2002 (14.5 mm in 1983 and 187.1 mm in 2006) respectively. Hence, a wide variation in the amounts of seasonal rainfall and the timing of anomalous occurrences among the four seasons can be detected from historical observation.

From Figure 5.3, the rainfall ensembles of all the four seasons from 1962 to 2007 under A2 fairly well capture the annual variability of observations. Out of the 46 validating years, the modified k-nn model can capture historical observation for 30-35 years. The simulations under B2 (Figure 5.4) can preserve the annual variability of seasonal observed rainfall.

A better performance of the modified k-nn model is associated with B2, rather than A2. The large spread of ensembles (i.e. the quartile range in the box plots) is found in some years' simulation under both scenarios. As the result, the modified k-nn model performs moderately well as far as capturing the annual variability of seasonal rainfall is concerned.

The annual statistics (mean, median, SD, IQR and skew) of 1962-2007 seasonal rainfall are calculated separately for the observed and modeled rainfall. The annual statistics of observations are plotted in black dots overlaying the box plots of annual statistics estimated from 300 simulations corresponding to A2 (Figure 5.5) and B2 (Figure 5.6). Under A2 and B2, the annual means and medians of observed rainfall in the four seasons are well captured by the modified k-nn model. The consistency in capturing both annual statistics is highlighted by the small spread of values (i.e. the quartile range in the box plots). Furthermore, the modified k-nn model under both scenarios can well preserve the annual SD and IQR of observations. The model performance indicates more consistency in capturing the annual SD than IQR. The coefficients of skew are also well captured by the modified k-nn model under A2 and B2. Hence, the modified k-nn model presents a good performance in capturing the annual statistics of seasonal rainfall under both scenarios of future climate.

An additional index, i.e. the bias, is computed to evaluate the model performance and is intended as the absolute difference in annual mean of the 1962-2007 estimations from the observations (Equation 5.8).

$$bias_g = \left| \frac{\sum_{m=1}^{nyears} Y_{m,g} - \sum_{m=1}^{nyears} O_m}{nyears} \right|$$ Equation 5.8

where $Y_{m,g}$ is the estimation during year m of simulation g with m=1, 2, 3,…., n *years* and g=1, 2, 3,…., N simulations, and O_m is the observation at year m.

The absolute bias expressed as the percentage of annual averaged rainfall ranges from 0 to +∞. A smaller value of bias indicates a better performance of model and vice versa for a greater value.

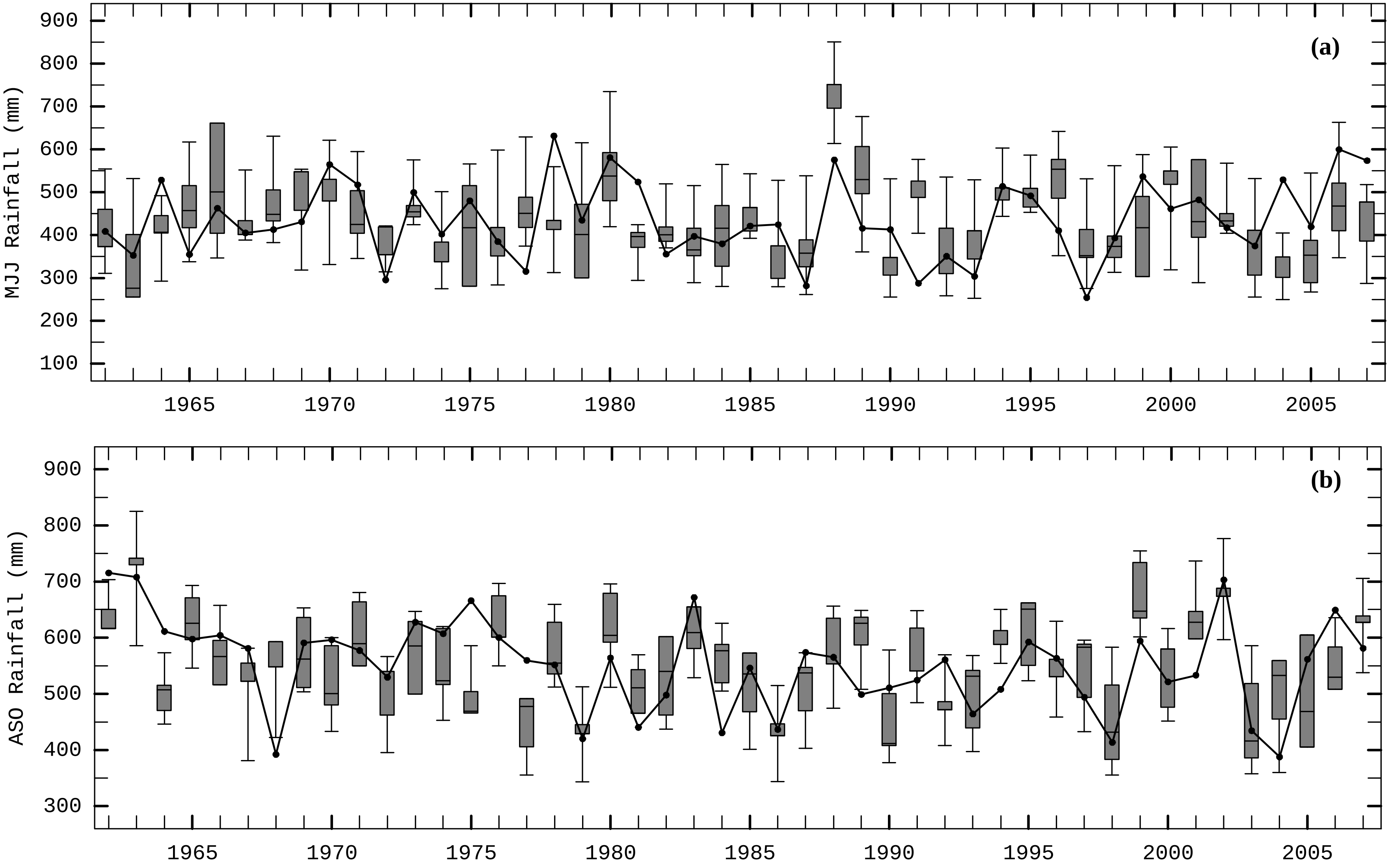


Figure 5.3: Box plots of 1962-2007 rainfall during (a) MJJ; (b) ASO; (c) NDJ; and (d) FMA estimated from 300 ensemble members simulated by the modified k-nn model under A2. The solid lines with marks represent the annual observations.

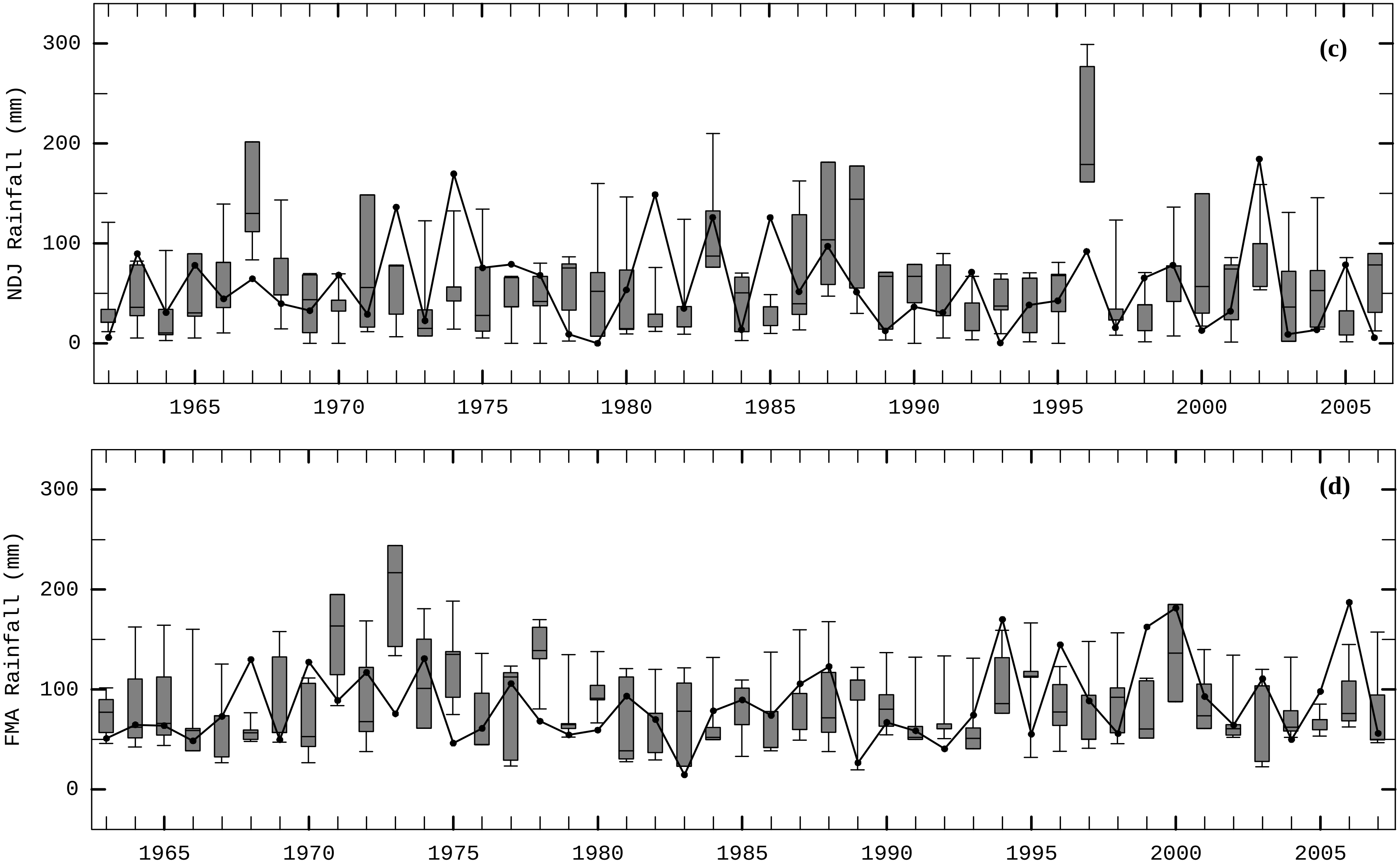


Figure 5.3 (cont).

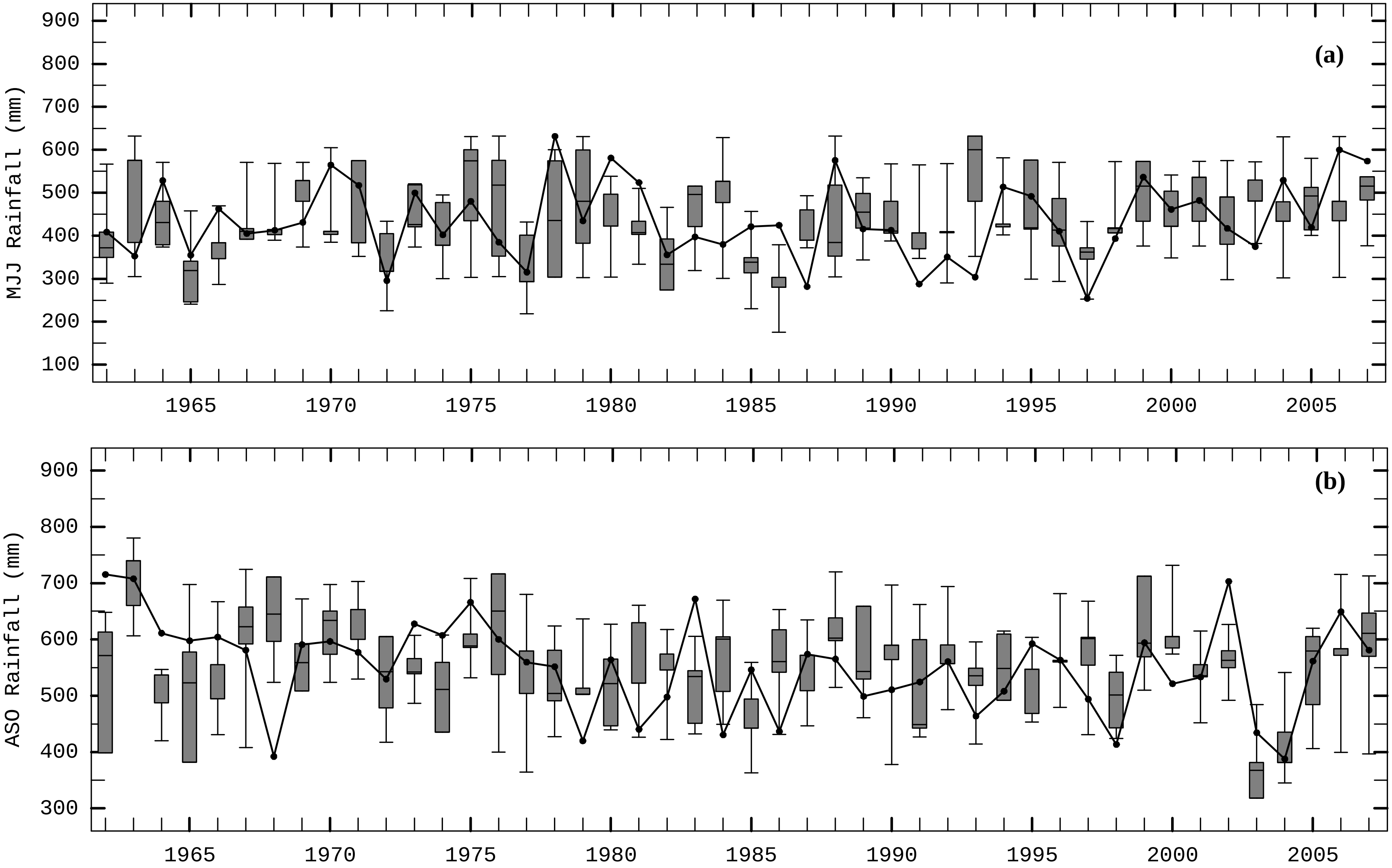


Figure 5.4: Same as Figure 5.3 but for simulations under B2.

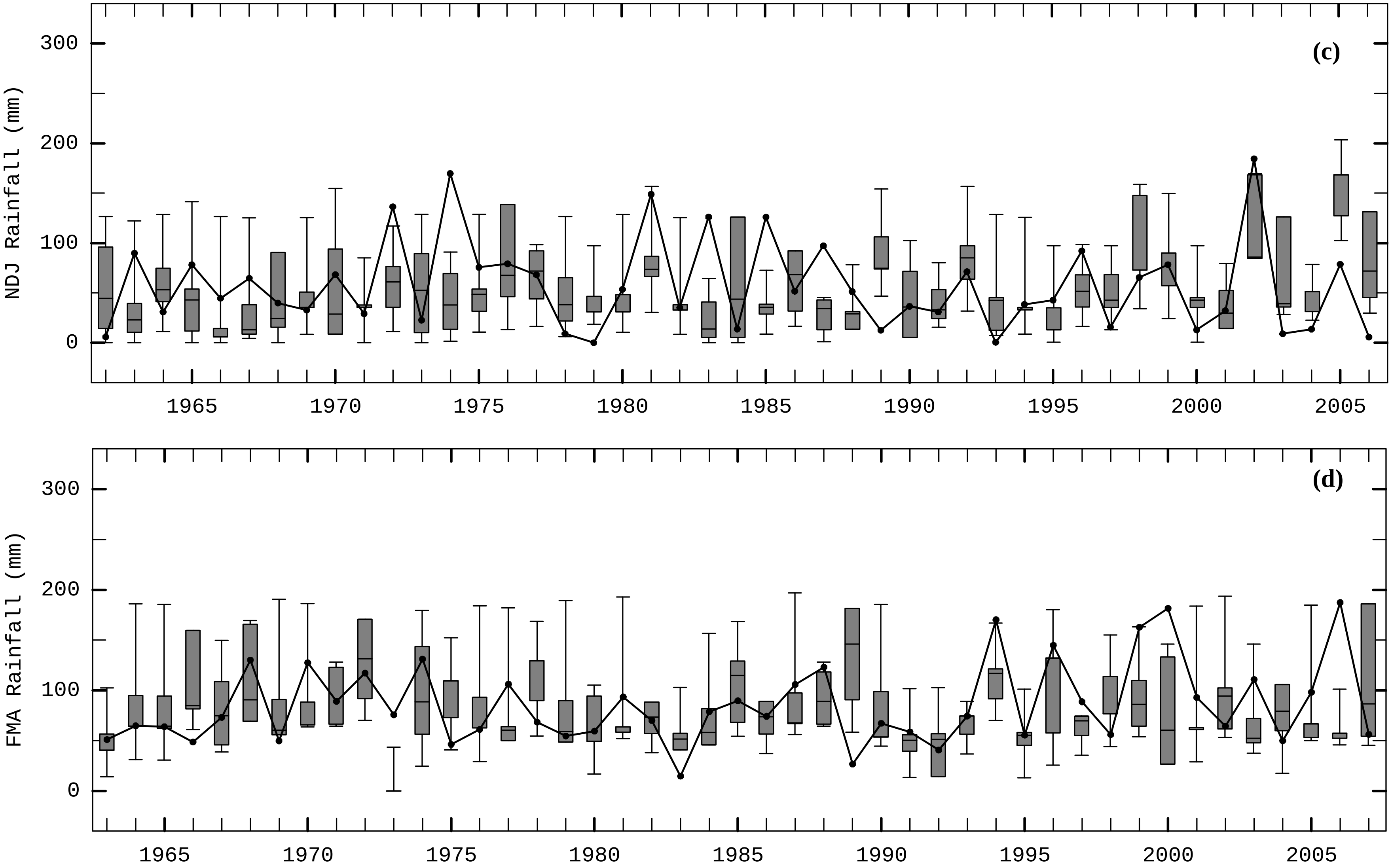


Figure 5.4 (cont).

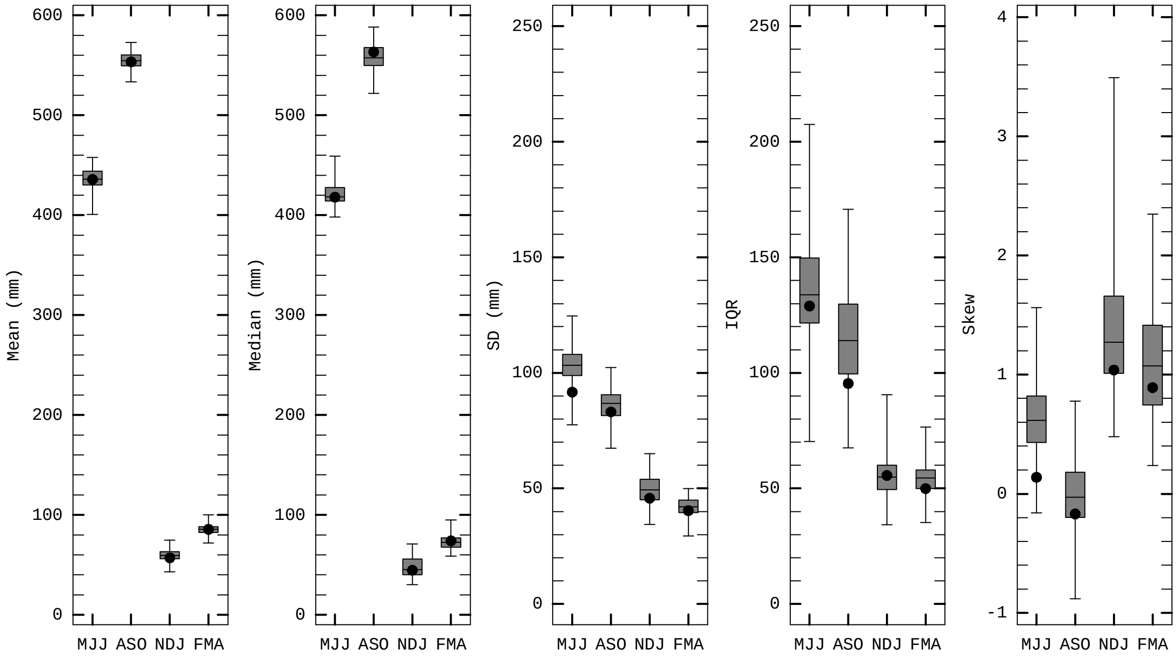


Figure 5.5: Box plots of annual statistics of 1962-2007 seasonal rainfall from 300 simulated members under A2. The black dots represent the annual statistics of observations.

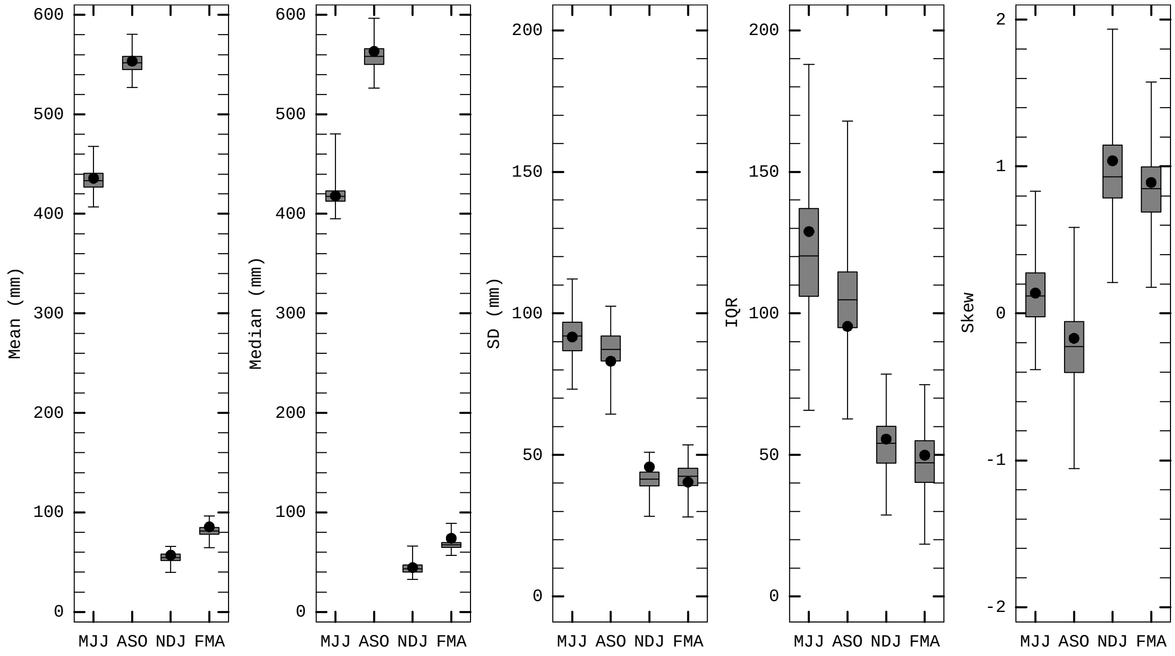


Figure 5.6: Same as Figure 5.5 but for simulations under B2.

The box plots of absolute biases estimated from the 300 ensemble members are presented for each season of rainfall. Under A2 (Figure 5.7(a)), the absolute biases are under 32% for all four seasons. The smallest absolute biases (ranging from 0.01% to 3.59%) are associated with ASO rainfall simulations, whereas the absolute biases of NDJ rainfall ensembles, varying from 0.06% to 31.35%, are the largest.

On the other hand, the absolute biases of simulations under B2 (Figure 5.7(b)) are within 30%. The absolute biases corresponding to ASO rainfall simulations, which are estimated to be 0.02-5.33%, are the smallest. As is the case with A2, the largest absolute bias corresponds to NDJ rainfall simulations with a range of 0.06-29.92%. Based on the absolute bias, the modified k-nn model performs better when simulating monsoon rainfall (i.e. MJJ and ASO rainfall) rather than the dry season rainfall (NDJ and FMA rainfall). This is because of a stronger relationship between large-scale atmospheric predictors and monsoon rainfall than that between the predictors and the dry season rainfall (as has been presented in Chapter 4).

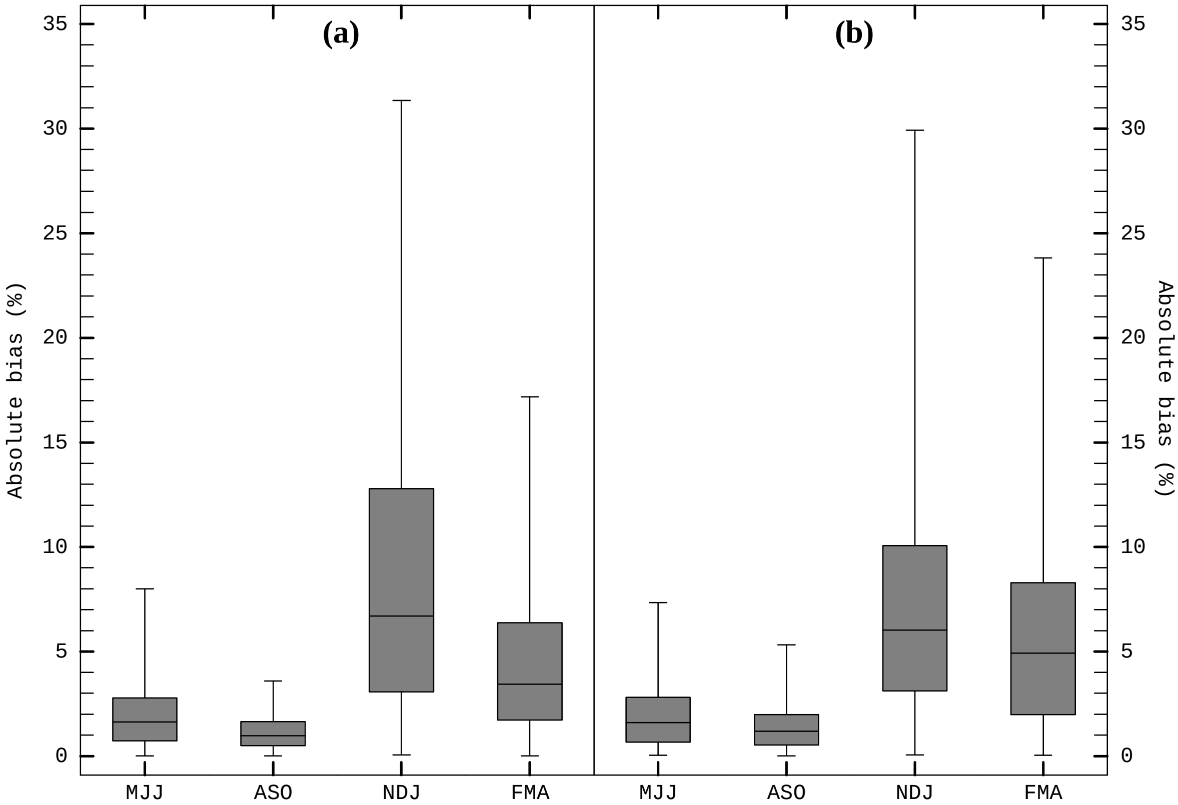


Figure 5.7: Box plots of absolute biases of annual mean from 300 simulated members under (a) A2; and (b) B2.

The likelihood skill score (LLH) is used to evaluate a statistical stochastic model for capturing the PDF of climatology. First, the observations are divided into three categories based on the defined thresholds, which in this case are set at the 33rd and 67th percentiles. Rainfall below the 33rd percentile is defined as below-normal rainfall while rainfall above the 67th percentile falls into the category of above-normal rainfall. Rainfall which does not fall into either category is denoted as normal rainfall. The second step is to calculate the categorical probabilities of climatology, which are the proportion of historical rainfall in each category. In this case, since historical rainfall is divided at the 33rd and 67th percentile, the categorical probability of all three categories is 1/3. Then, in a given year, the N simulated ensembles are also divided into three categories using the same thresholds. The categorical probabilities of ensembles in a given year, which are the proportion of rainfall ensembles in each category, are computed. Subsequently, LLH is estimated using Equation 5.9.

$$LLH = \frac{\prod_{t=1}^{n} \hat{P}_{j,t}}{\prod_{t=1}^{n} P_{cj,t}}$$ Equation 5.9

where n is the number of years, j is the category of the observed value in the year t, $\hat{P}_{j,t}$ with $\hat{P}_{j,t} = (\hat{P}_{1,t}, \hat{P}_{2,t}, \hat{P}_{3,t}, ..., \hat{P}_{k,t})$ is the probability of rainfall ensembles for category j in the year t, where k is the number of categories, and $P_{cj,t}$ is the categorical probability of climatology for category j in the year t, which in this case is the same value for all three categories, i.e. 1/3 each.

LLH varies from 0.0 to a number of categories, like +3.0 in this study. The score of +1.0 indicates no difference between the model performance and reference simulated climatology. A score of less than +1.0 indicates a weaker performance of the model compared to climatology. On the other hand, a better performance than climatology is associated with LLH greater than +1.0. The score of +3.0 indicates the perfect performance of the model. In the study case, LLH is estimated separately for each season of rainfall, validating year and scenario of future climate.

Figure 5.8 and 5.9 show the LLH of rainfall ensembles during the four seasons from 1962 to 2007 under A2 and B2 scenarios respectively. The darker shading represents a better performance of the modified k-nn model. A dry year (D) is defined by rainfall below the 20th percentile from the 1950-2007 observations, whereas a wet year (W) is denoted by rainfall above the 80th percentile. The threshold values at the 20th (80th) percentiles of 1950-2007 MJJ, ASO, NDJ and FMA rainfall are computed as 381.7 (528.9), 498.2 (624.8), 15.8 (79.0) and 53.9 (114.5) mm respectively. As seen in Figure 5.8, under A2, the median LLH of 1962-2007 MJJ (ASO) rainfall simulations is calculated to be 1.12 (1.23). The 1962-2007 NDJ and FMA rainfall ensembles indicate median LLH as 1.29 and 1.13 respectively. Based on LLH greater than +1.0, the modified k-nn model performs well over anomalous events (i.e. dry and wet), in particular for ASO and NDJ rainfalls.

On the other hand, the modified k-nn model under B2 (Figure 5.9) performs worse than under A2. The median LLH of 1962-2007 MJJ and ASO rainfall simulations are 1.12 and 0.79 respectively. The dry season rainfall has a median LLH less than +1.0, i.e. 0.79 for NDJ rainfall and 0.99 for FMA rainfall. The model performance over anomalous events is not consistent among the four seasons.

5.3.2 The multisite daily rainfall generator

The performance evaluation of a multisite daily rainfall generator involves two separate simulations. First, the daily rainfall generator is evaluated by applying its results to the daily rainfall averaged over 50 selected stations for the period of 1950 to 2007. In this study, 100 daily rainfall ensembles for each year of 58 years are generated. Second, from 1950 to 2007, the multisite generator simulates 100 ensembles of daily rainfall separately for each selected station. The evaluation of the model performance is done in conjunction with daily generations in all the 50 stations.

The first kind of simulation (i.e. the daily rainfall generator) is evaluated using three indexes: (i) the transition probabilities of daily rainfall which can be used to evaluate the Markov Chain on fitting the transition probabilities of observations; (ii) the dry- and wet-spell lengths of daily rainfall by month which can present the performance of the Markov Chain on capturing the number of dry and wet days, and the lengths of dry and wet spells of historical daily rainfall; and (iii) the statistics of daily rainfall month-by-month in terms of mean, median, SD, IQR, skew and lag-1 autocorrelation, which are then used to diagnose the performance of the Monte Carlo approach on capturing the statistics of the daily observations. All indexes are estimated separately for each month using simulated results and presented by box plots. A box plot is computed from 100 ensembles. A box is defined by the 75th percentile (the upper quartile or QU) and the 25th percentile (the lower quartile or QL). The horizontal line within the box presents a median of 100 ensembles. The caps of upper and lower whiskers indicate the extreme values (Equation 5.7). The outliers are represented by dots above or below the caps. The historical values are also shown along with the box plots by the solid lines with marks.

Figure 5.10 shows the evaluation of the model in terms of transition probabilities. From 100 daily rainfall ensembles for each year of 58 years, the Markov Chain can capture well the nine transition probabilities of observations in all the 12 months. The consistency of performance corresponding to the small spread of values (i.e. the quartile range in a box plot) is observed especially for $P_{w\text{-}d}$, $P_{w\text{-}w}$ and $P_{w\text{-}e}$ in the months from May to October, which cover the pre-monsoon and monsoon seasons.

As expected, the Markov Chain can also perform well when capturing the average dry and wet days of historical rainfall (Figure 5.11(a) and (b)). However, since the dry and wet spells of observations can be smoothened out by the three-state, first-order Markov Chain, the average dry- and wet-spell lengths (Figure 5.11(c) and (d)) are underestimated in some months, like in May and June. From Figure 5.11(e) and (f), the maximum dry-spell lengths of some months: in particular July, August and September, are not well reproduced. However, the model can capture well maximum wet-spell lengths.

The Monte Carlo algorithm generates the amount of daily rainfall based on the rainfall states obtained from the three-state, first-order Markov Chain. Figure 5.12 shows the performance of the daily rainfall generator using the Monte Carlo method on preserving the statistics of daily observations. The model can capture well the means and medians of daily rainfall. However, the means corresponding to the daily rainfall ensembles in April and December are slightly underestimated, whereas an overestimation associated with the October simulations has been found. Similarly, underestimation in medians is observed in the simulations of April, May and July. For the daily rainfall simulations in June, the medians are slightly overestimated. The SD, IQR and skew of 1950-2007 daily observations is captured accurately by Monte Carlo. On the other hand, the lag-1 autocorrelations in all 12 months are underestimated because, as previously mentioned, daily rainfall during any consecutive days is resampled based on the rainfall states generated by the Markov Chain.

For the multisite generator, the model performance is evaluated at 50 rainfall stations using 100 simulated ensembles of daily rainfall from 1950 to 2007. Two criteria are applied: one is the statistics of the 1950-2007 daily rainfall, and another is spatial cross-correlations among the 50 selected stations. In terms of statistics, the mean, SD and skew of the 1950-2007 daily rainfall at each station are estimated using both, the observations and simulations. The box plots of statistics of simulated results are used to diagnose the performance of the multisite generator on capturing the spatial distribution of the statistics of daily observed rainfall. Moreover, the spatial cross-correlations, which are the correlations between daily rainfall recorded in a pair of stations, are used to evaluate the model performance in preserving the linear relationship of daily rainfall among the 50 stations. The spatial cross-correlations of observations are estimated using daily historical rainfall, whereas the cross-correlations of simulations are calculated using averaged values over 100 ensembles of daily modeled results.

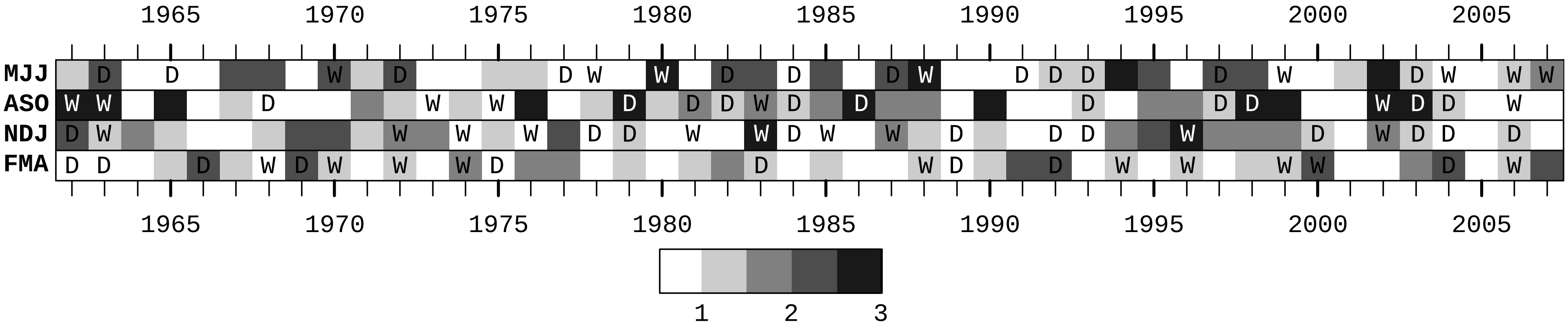


Figure 5.8: LLH of rainfall ensembles during the four seasons from 1962 to 2007 under A2. D: dry year; W: wet year.

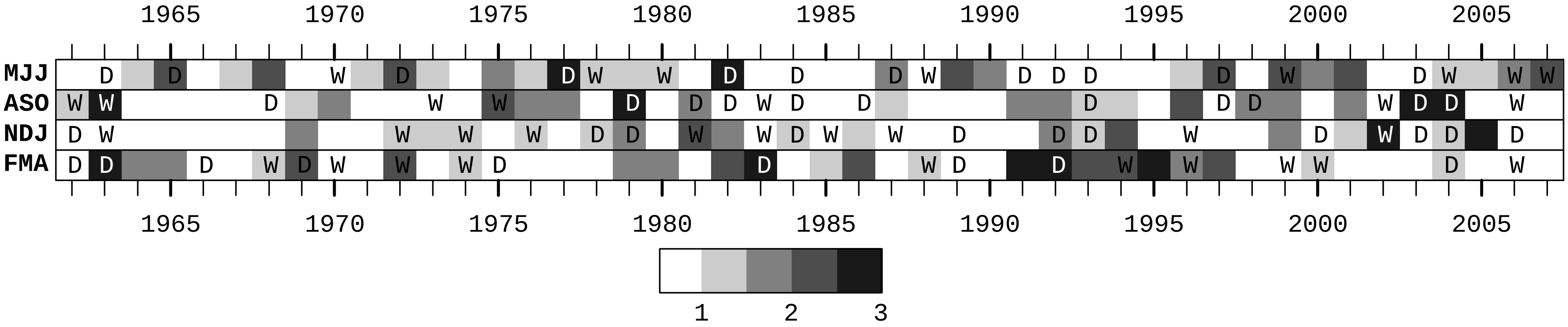
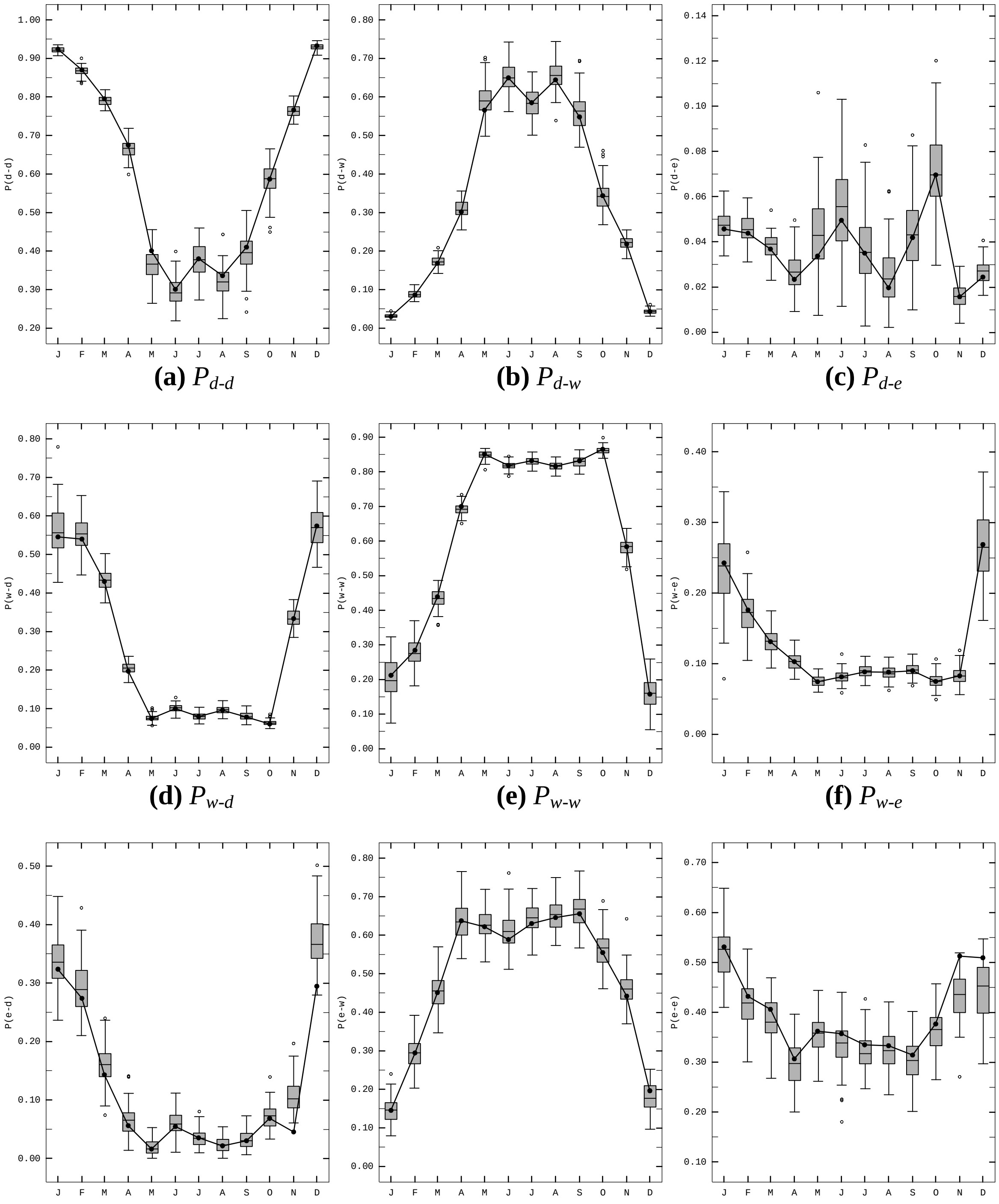


Figure 5.9: Same as Figure 5.8 but for simulations under B2.

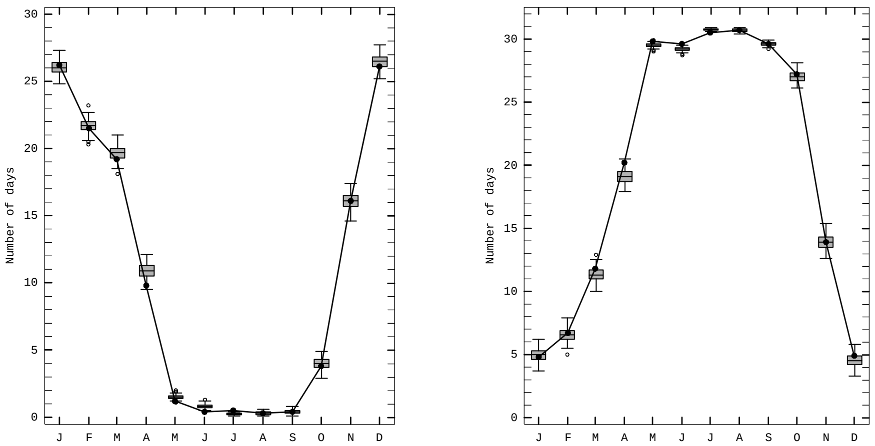


(a) $P_{d\text{-}d}$ **(b)** $P_{d\text{-}w}$ **(c)** $P_{d\text{-}e}$

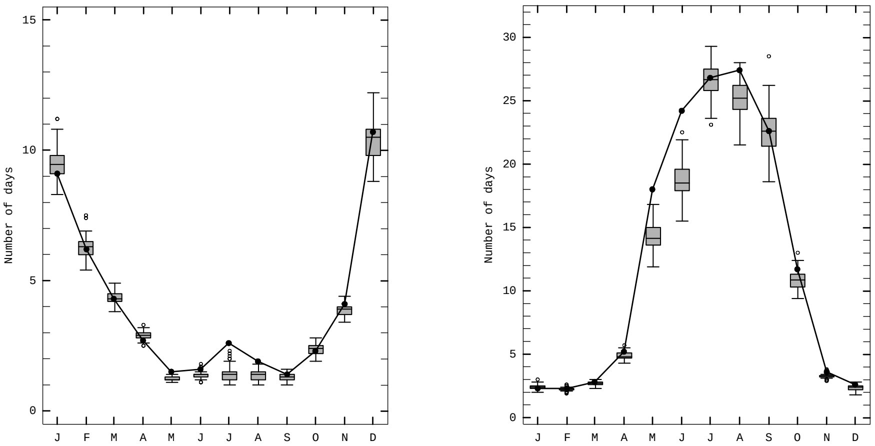
(d) $P_{w\text{-}d}$ **(e)** $P_{w\text{-}w}$ **(f)** $P_{w\text{-}e}$

(g) $P_{e\text{-}d}$ **(h)** $P_{e\text{-}w}$ **(i)** $P_{e\text{-}e}$

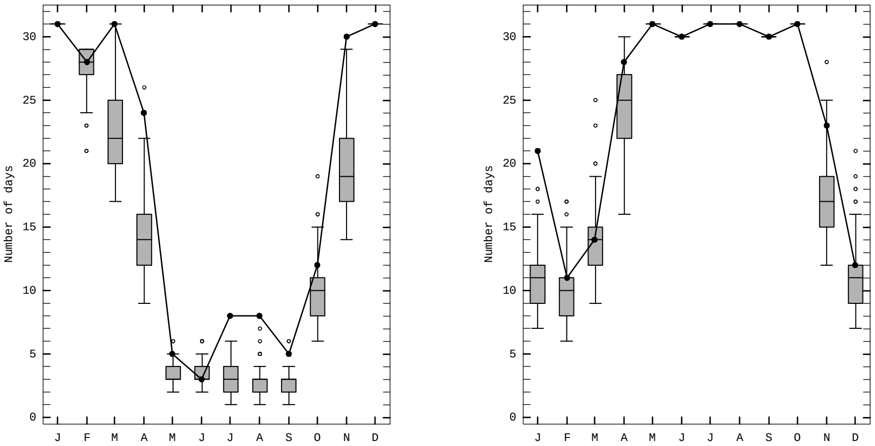
Figure 5.10: Box plots of transition probabilities of 1950-2007 daily rainfall from 100 ensembles simulated by the daily rainfall generator. The dots above or below the caps of whiskers represent the outlier values. The solid lines with marks represent the transition probabilities of 1950-2007 daily observations.



(a) Average dry days **(b)** Average wet days

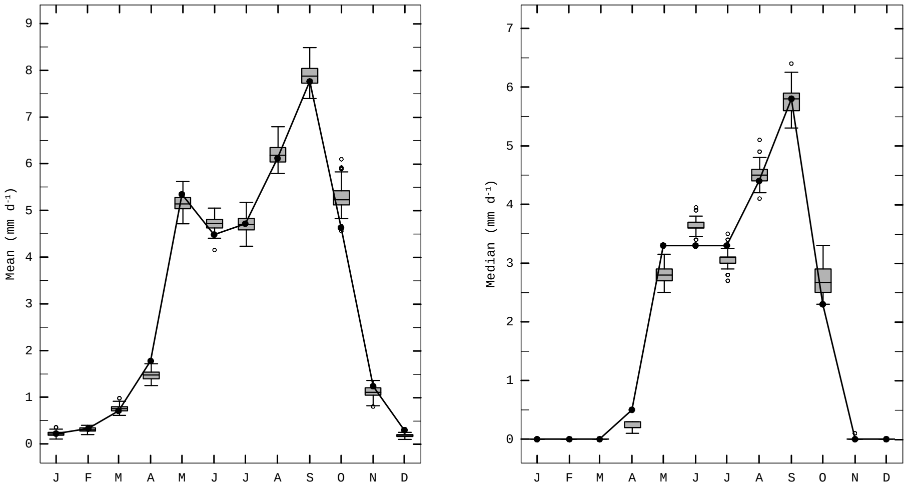


(c) Average dry-spell lengths **(d)** Average wet-spell lengths

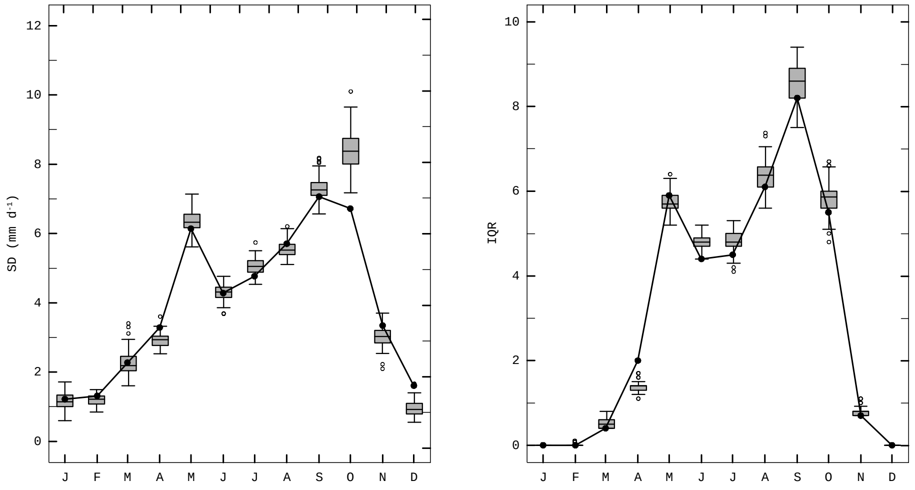


(e) Maximum dry-spell lengths **(f)** Maximum wet-spell lengths

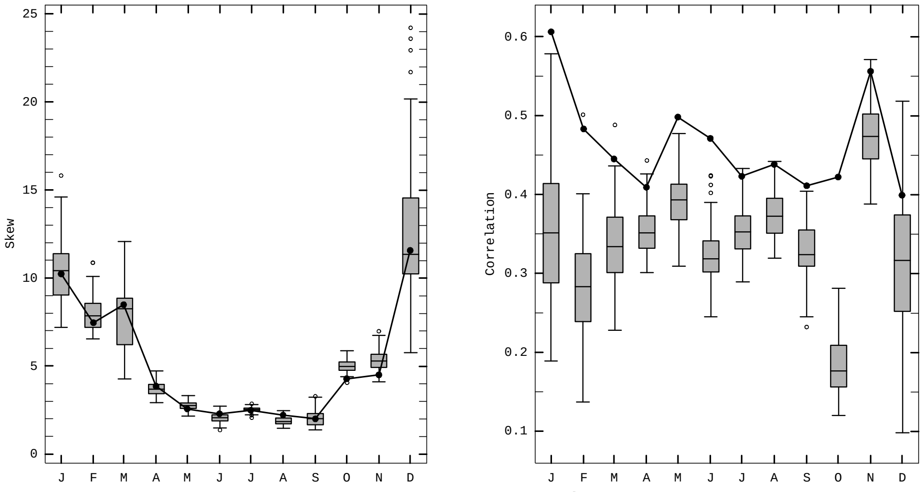
Figure 5.11: Box plots of spell statistics of 1950-2007 daily rainfall from 100 ensembles simulated by the daily rainfall generator. The dots above or below the caps of whiskers represent the outlier values. The solid lines with marks represent the spell statistics of 1950-2007 daily observations.



(a) Mean **(b)** Median



(c) SD **(d)** IQR



(e) Skew **(f)** Lag-1 autocorrelation

Figure 5.12: Box plots of basic statistics of 1950-2007 daily rainfall from 100 ensembles simulated by the daily rainfall generator. The dots above or below the caps of whiskers represent the outlier values. The solid lines with marks represent the statistics of 1950-2007 daily observations.

Figure 5.13 to 5.15 respectively show the box plots of daily means, SD and skew of 1950-2007 rainfall from 100 ensembles along with the corresponding statistics of observations at 50 rainfall stations. From Figure 5.13, the means of daily observed rainfall ranging from 2.4 to 4.7 mm d^{-1}, indicate a wide variation in the averages of daily rainfall among the 50 stations. The variability of means is well captured by the multisite generator. However, the daily means at Station 060101, 17032 and 17074 are slightly underestimated, whereas overestimated means are found at Station 400201, 07082 and 26102.

The multisite generator can also preserve the SD of observations (Figure 5.14). Moreover, at all rainfall stations, the skew of daily rainfall (Figure 5.15) is well reproduced by the multisite generator with the exception of Station 327301 and 07072, where an overestimation is observed.

As for cross-correlations among all the stations, Figure 5.16 shows the relationship between distance and spatial cross-correlations of daily observed rainfall among pairs of stations. Nonlinear relationships between distance and spatial correlations can be observed, which reflect the influence of topography on station distribution. A greater distance between two stations is associated with a smaller correlation and vice versa for stations nearer to one another. The maximally distant stations estimated to be 530.68 km apart, which are Station 07492 (19°59´47˝N latitude and 99°15´33˝E longitude) and Station 400301 (15°21´00˝N latitude and 100°30´00˝E longitude), show a spatial correlation of 0.11. On the other hand, a maximum correlation of 0.85 is found between the daily rainfall of the station 060101 (17°03´00˝N latitude and 99°04´00˝E longitude) and Station 63022 (17°02´46˝N latitude and 99°04´34˝E longitude) and it can be associated with the fact that these two are the closest stations, only 1.09 km apart from one another.

Figure 5.17 presents the comparison by month between spatial cross-correlations of historical data and the modeled results. The multisite generator performs well when preserving cross-correlations among pairs of stations, in particular from May to November. However, in the dry season (i.e. from December to April), the cross-correlations among stations cannot be well reproduced, as shown by the sparse points of correlations.

5.4 Results and discussion

5.4.1 Effects of future climate on annual variability and trends in seasonal rainfall

The modified k-nn model is adopted to simulate 300 rainfall ensembles for each season (i.e. MJJ, ASO, NDJ and FMA) from 2011 to 2100. The model is developed so as to downscale seasonal rainfall in the study basin from large-scale atmospheric variables, which are obtained from a GCM named GFDL-R30. The modified k-nn model also aims to determine the effects of future climate on seasonal rainfall under two scenarios (i.e. A2 and B2). From 300 simulated rainfall ensembles, the medians are estimated. The seasonal rainfall anomalies from 2011 to 2100 are then calculated with respect to the observed 1961-1990 average seasonal rainfall and plotted along with the anomalies of observed rainfall during 1950 to 2007.

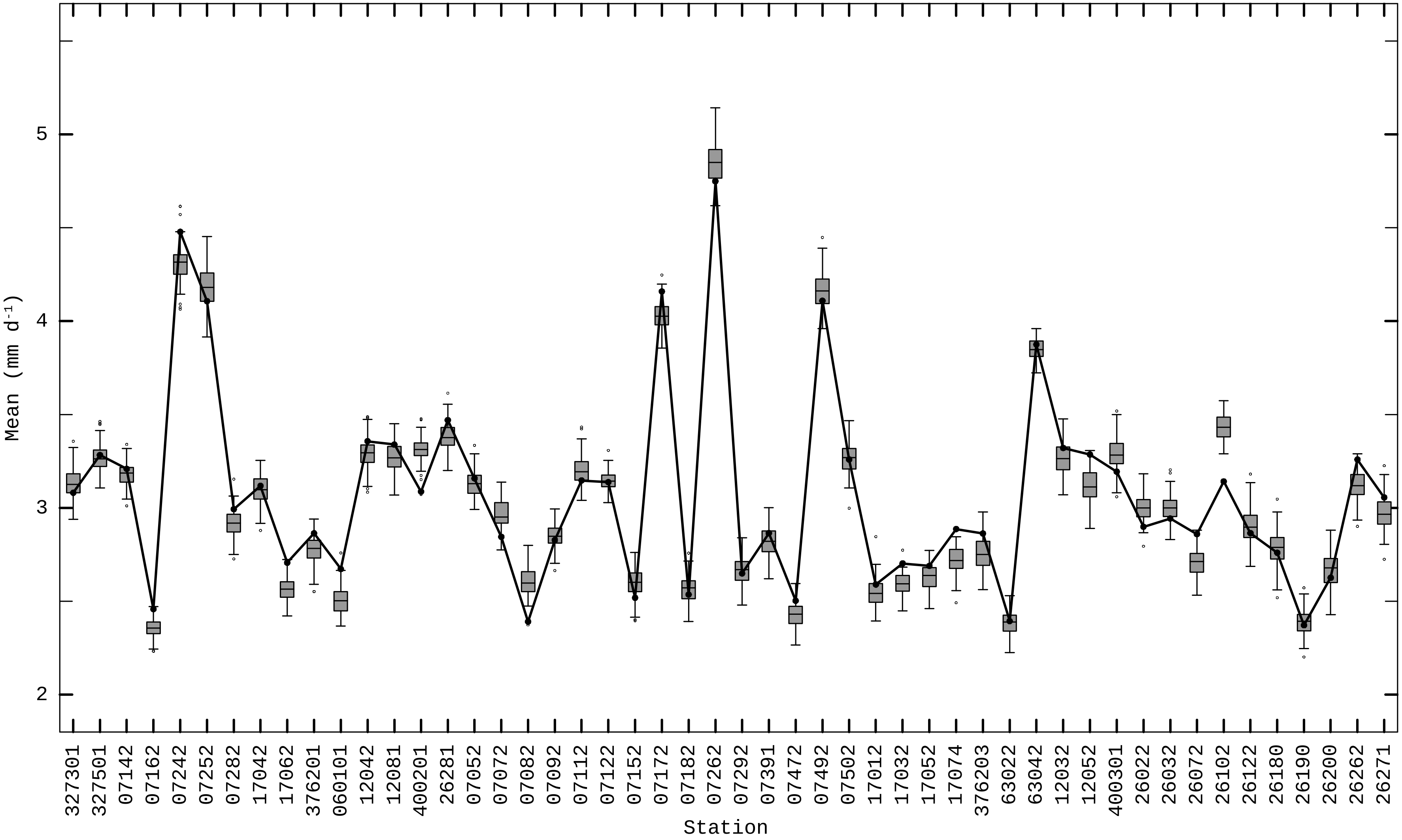


Figure 5.13: Box plots of means of 1950-2007 daily rainfall at 50 rainfall stations from 100 ensembles simulated by the multisite generator. The dots above or below the caps of whiskers represent the outlier values. The solid lines with marks represent the daily means of 1950-2007 observations.

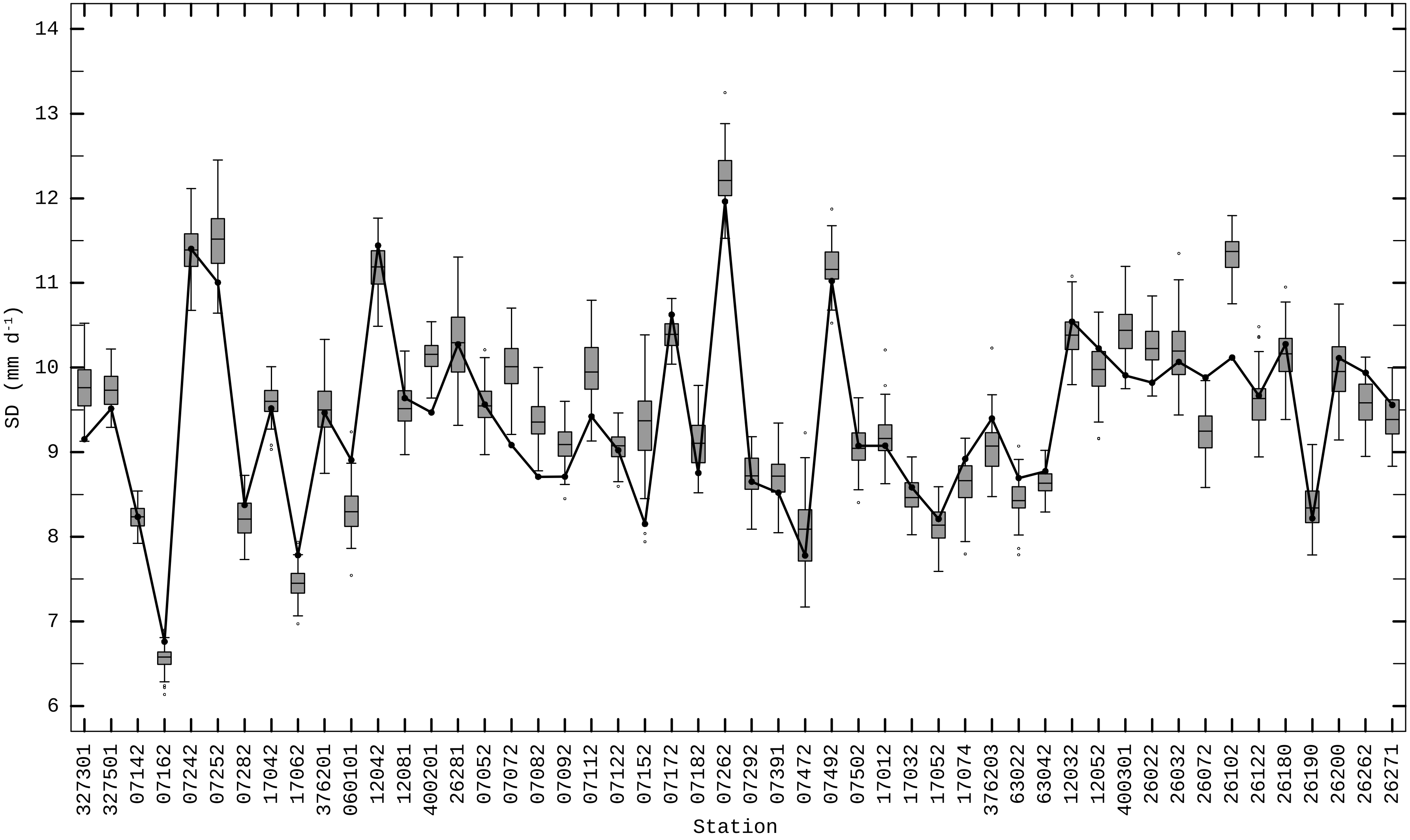


Figure 5.14: Same as Figure 5.13 but for SD.

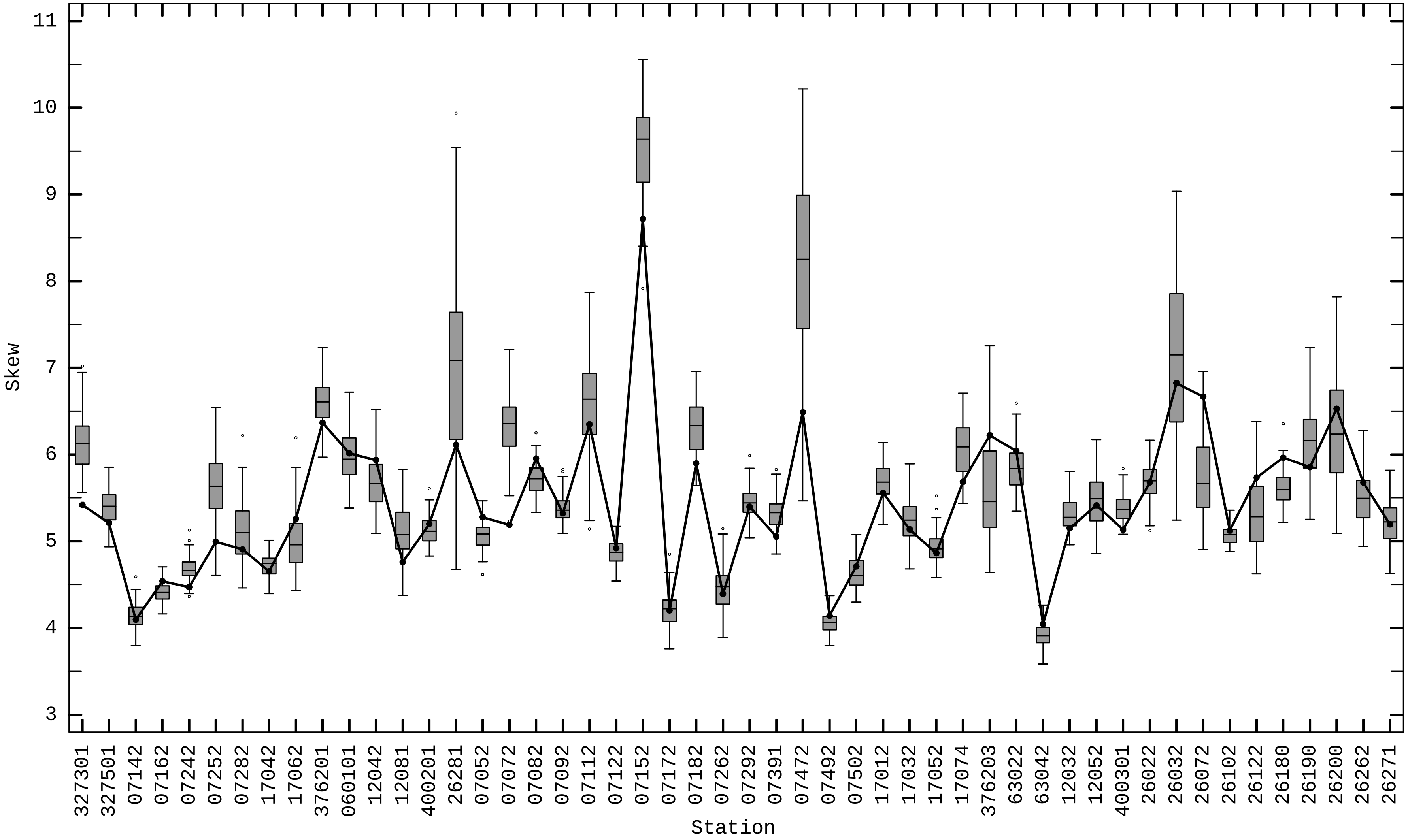


Figure 5.15: Same as Figure 5.13 but for skew.

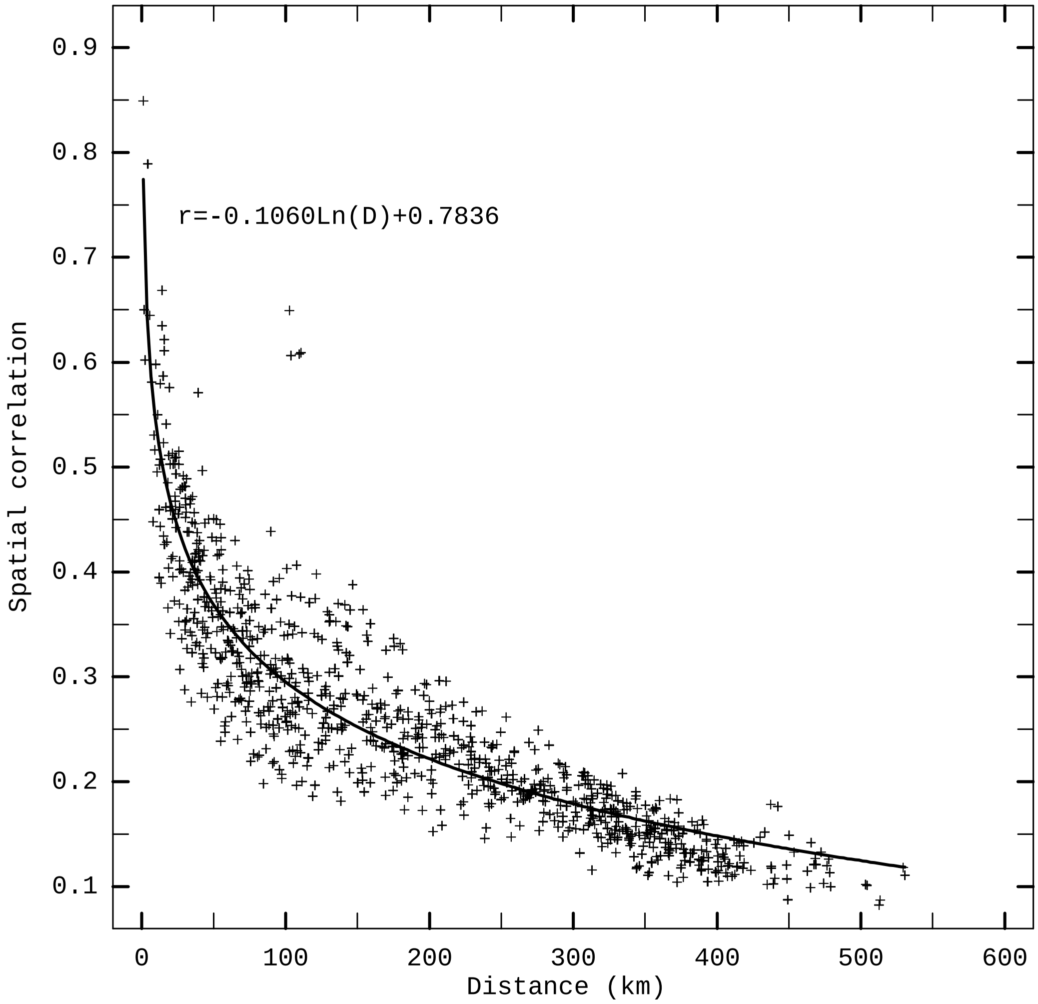


Figure 5.16: Scatter plot between distance (*D*) and spatial correlations (*r*) of 1950-2007 daily observed rainfall at the 50 selected stations.

A linear trend of 1950-2007 observed MJJ rainfall (Figure 5.18(a)) is -0.69 mm per year. The maximum and the minimum observed MJJ rainfall are 651.9 and 254.0 mm, which are seen in 1950 and 1997 respectively. The rainfall anomalies corresponding to the observed 1961-1990 average MJJ rainfall range from -2.1 to +2.5 mm. Pre-1980, the observed MJJ rainfall tended to be above-normal and vice versa post-1980. Under A2, the MJJ rainfall from 2011 to 2100 tends to decrease by 0.11 mm per year (Figure 5.18(a)). A maximum of 759.8 mm and a minimum of 279.6 mm will be observed in 2015 and 2050 respectively. Moreover, the above-normal MJJ rainfall, with respect to the observed 1961-1990 averaged MJJ rainfall, will occur during the 2020s and 2070s and vice versa for the remaining periods. On the other hand, the 2011-2100 MJJ rainfall under B2 tends to decrease with a linear trend of 0.17 mm per year. Maximum MJJ rainfall will be observed in 2022 at 600.2 mm, showing an anomaly of +1.9 mm. It is also important to note that the decreasing trends of 2011-2100 simulated MJJ rainfall are less than the decreasing trends of 1950-2007 observed MJJ rainfall. These results suggest that future climate under both scenarios influences decreasing MJJ rainfall with a slower rate in the future than it did in earlier periods. However, the results from GCMs under A1B scenario reported by NIC (2009) were different. From 2049 to 2069, the average precipitation in summer monsoon season (i.e. June-July-August: JJA) over Asia will tend to increase varying from 2.6% to 3.4%. From IPCC (2007a), JJA rainfall over the Ping River Basin under A1B will tend to increase less than 5%. Compared to precipitation in Indonesia, Boer and Faqih (2003) studied changes in future precipitation through 2080 using several GCMs such as CCSR, CSIRO and HadCM3 under A2 and B2 scenarios. The increasing and decreasing precipitation will be observed depending on model and scenario. The inconsistency in precipitation trends cannot generally conclude the effects of future climate on precipitation across Indonesia.

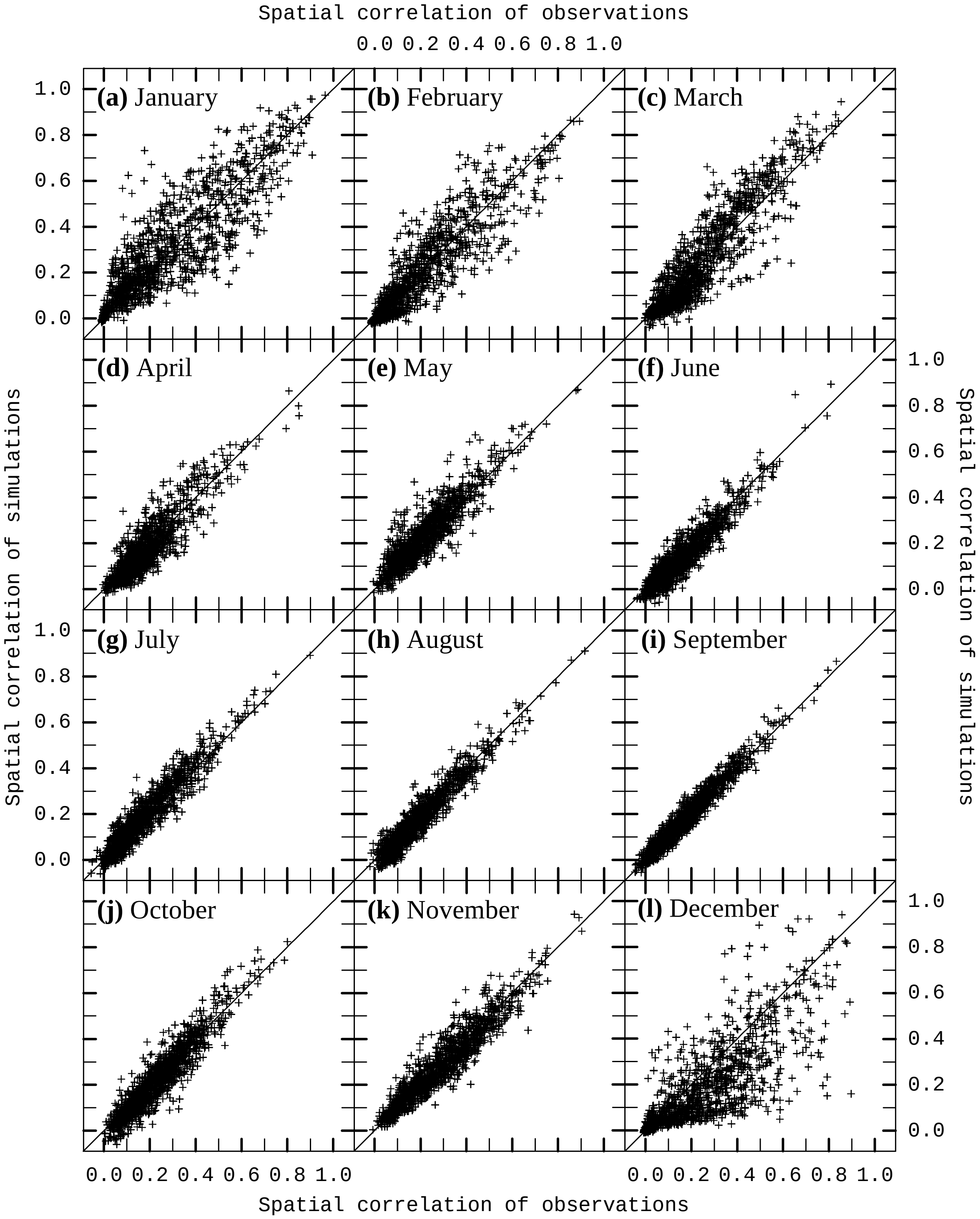


Figure 5.17: Comparison by month between spatial cross-correlations of daily observed and simulated rainfall among pairs of the 50 rainfall stations.

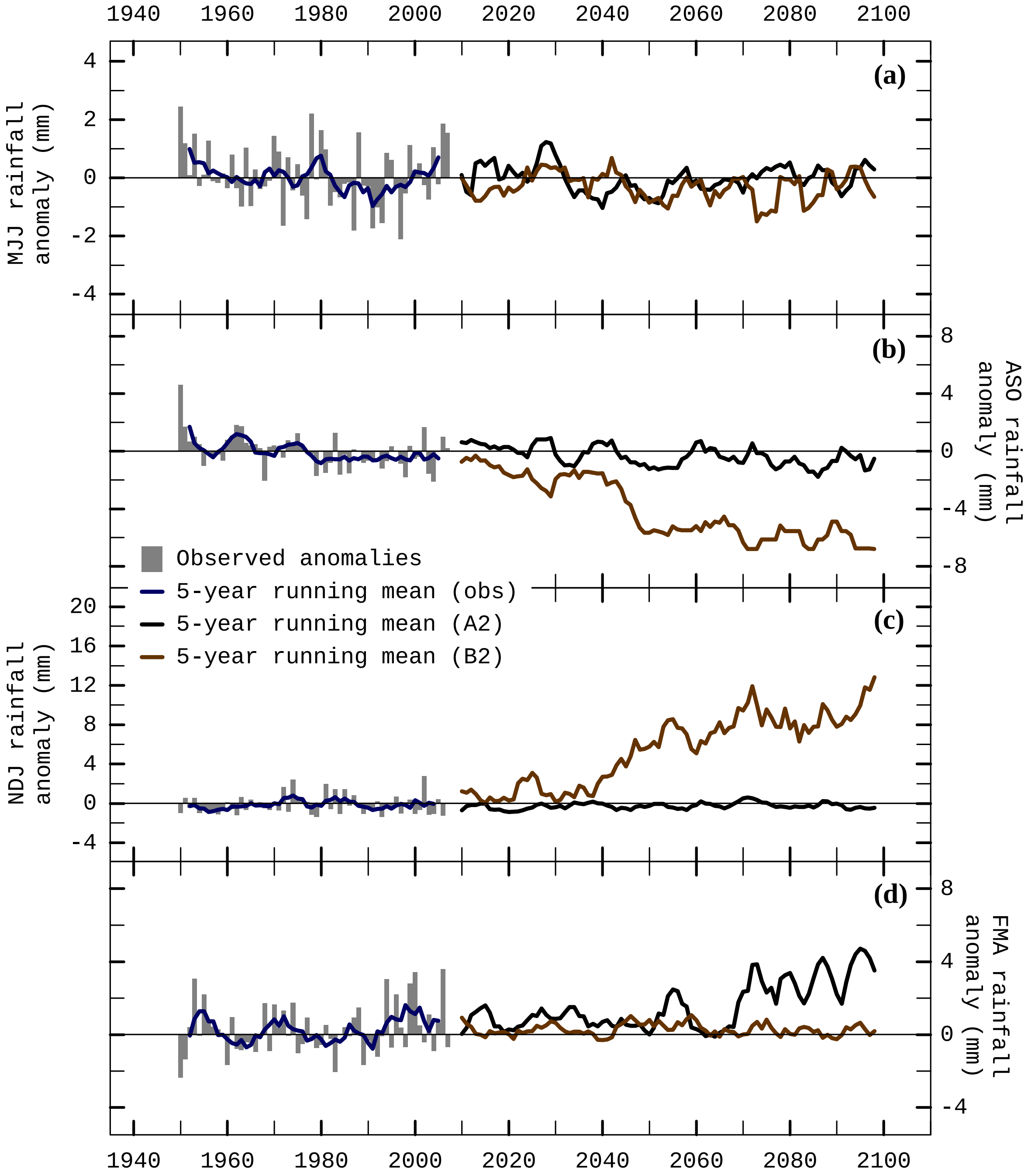


Figure 5.18: Annual anomalies of the 1950-2007 observed rainfall and the 5-year running mean from observed and modified k-nn simulated (2011-2100) rainfall during (a) MJJ; (b) ASO; (c) NDJ; and (d) FMA. The anomalies are estimated with respect to the observed 1961-1990 averaged seasonal rainfall.

The 1950-2007 observed ASO rainfall (Figure 5.18(b)) tends to decrease by 2.29 mm per year. This trend is significant at a 99.5% confidence level by the standard *t*-test (Haan, 2002). A maximum of 948.3 mm and a minimum of 387.6 mm were seen in 1950 and 2004 respectively. The positive anomalies of ASO rainfall, calculated corresponding to the observed 1961-1990 averaged ASO rainfall, are found during pre-1980, whereas negative anomalies are observed post-1980. These are consistent with the anomalies of 1950-2007 observed MJJ rainfall. The 2011-2100 ASO rainfall under A2 tends to decrease by 1.09 mm per year which is significant at 97.5% confidence levels. A maximum of 908.4 mm will be observed in 2027 with an anomaly of +4.1 mm. Minimum ASO rainfall at 346.4 mm will occur in 2095. Under B2, a decreasing linear trend of 2011-2100 ASO rainfall is 6.16 mm per year with significance at a 99.9% confidence level. In 2013, a maximum of 598.7 mm will be observed, with a rainfall anomaly of +0.4 mm. Negative rainfall anomalies are observed especially after 2040. Hence, from 2011 to 2100, the predicted effects of future climate will cause a drastic decrease of monsoon season rainfall in the study basin, in particular under Scenario B2. The decreasing trends in ASO rainfall may be caused by the reducing tropical cyclone frequency over Southeast Asia (Wu and Wang, 2004). However, the frequency and intensity of tropical cyclones are influenced by ENSO, which changes in ENSO are not consistent (IPCC, 2007a).

In contrast, from 1950 to 2007, the observed NDJ rainfall (Figure 5.18(c)) tends to increase slightly (by 0.03 mm per year). In 2002, a maximum of 184.4 mm was recorded, with an anomaly of +2.8 mm, in conjunction with the observed 1961-1990 average NDJ rainfall. A minimum of 0.0 mm was observed in 1979 and is associated with a rainfall anomaly of -1.4 mm. Under A2, an increasing trend of 2011-2100 NDJ rainfall is estimated by 0.07 mm per year. A maximum of 135.7 mm, associated with rainfall anomaly of +1.7 mm, will be seen in 2013, whereas a minimum of 3.0 mm will be observed in 2018 with an anomaly of -1.3 mm. Furthermore, with significance at a 99.9% confidence level, the 2011-2100 NDJ rainfall under B2 will tend to increase by 5.91 mm per year. A maximum of 905.9 mm and a minimum of 28.4 mm will be observed in 2098 and 2019, associated with rainfall anomalies of +18.9 and -0.7 mm respectively. High positive anomalies will be observed especially after 2040. The increasing trends of NDJ rainfall are similar to the results provided by NIC (2009). From 2049 to 2069, the winter season precipitation (i.e. December-January-February) over Asia will tend to increase by 2.9% to 3.5% under A1B emission scenario. The decreasing ASO rainfall and the increasing NDJ rainfall trends during post-2040 suggest that climate in the future will affect a shift in the monsoon season. The delay in monsoon onset date is also found from the study of Bhaskaran and Mitchell (1998) who examined the effects of changing climate on Southeast Asian monsoon precipitation from 1990 to 2100 using HadCM2. They found the 10-15 day delay in monsoon onset date over the regions of Thailand, Cambodia, Laos and Vietnam.

For the observed 1950-2007 FMA rainfall (Figure 5.18(d)), an increasing trend of 0.43 mm per year is estimated. A maximum of 187.1 mm was observed in 2006, and a minimum of 5.5 mm was found in 1950. Rainfall anomalies calculated with respect to the observed 1961-1990 average FMA rainfall range from -2.4 to +3.6 mm. During the recent decades (i.e. after 1990), positive anomalies associated with above-normal rainfall are greater than they were in earlier centuries. Under A2, the 2011-2100 FMA rainfall tends to increase (with significance at 99.9% confidence level) by 1.09 mm per year. A maximum (minimum) of 312.5 (31.5) mm will be observed in 2095 (2052), corresponding to a rainfall anomaly of +7.7 (-1.5) mm. The above-normal rainfall suggests an increase in FMA rainfall in the study basin especially after 2070. On the other hand, under B2, FMA rainfall from 2011 to 2100 tends to only moderately increase by 0.02 mm per year. A maximum of 134.1 mm and a minimum of 34.7 mm will be found in 2077 and 2076 respectively. The rainfall anomalies here range from -1.4 to +1.9 mm.

In conclusion, the effects of future climate on seasonal rainfall during 2011 to 2100 suggest decreasing trends of pre-monsoon (MJJ) and monsoon (ASO) season rainfall, and increasing trends of dry (NDJ and FMA) season rainfall. The decreasing trends of MJJ and ASO rainfall vary from 0.11 to 6.16 mm per year, whereas the increasing trends of NDJ and FMA rainfall range from 0.02 to 5.91 mm per year. Compared to A2, future climate Scenario B2 suggests more severe effects on ASO and NDJ rainfall. However, the opposite is observed for FMA rainfall under Scenario A2. In addition, it is important to note that under B2, the monsoon season in the Ping River Basin tends to shift by one season from ASO to NDJ after 2040 and this shift is associated with drastically decreasing ASO rainfall (i.e. negative anomalies) and increasing NDJ rainfall (i.e. positive anomalies).

5.4.2 Effects of future climate on anomalous weather events

From 300 ensembles simulated for each season and each year, the median has been estimated, and the PDF of median rainfall from 2011 to 2100 under A2 and B2 are plotted, overlaid by the PDF of climatology from 1950 to 2007 (Figure 5.19). Subsequently, the probabilities of anomalous weather events (i.e. dry and wet conditions) are calculated from the PDF based on the thresholds at the 20th and 80th percentile of climatology. Rainfall less than the 20th percentile is defined as below-normal rainfall or causing a dry condition, whereas rainfall greater than the 80th percentile is denoted as above-normal rainfall or as wet condition. The 20th (80th) percentiles of MJJ, ASO, NDJ and FMA rainfall during 1950 to 2007 are estimated at 381.7 (528.9), 498.2 (624.8), 15.8 (79.0) and 53.9 (114.5) mm respectively.

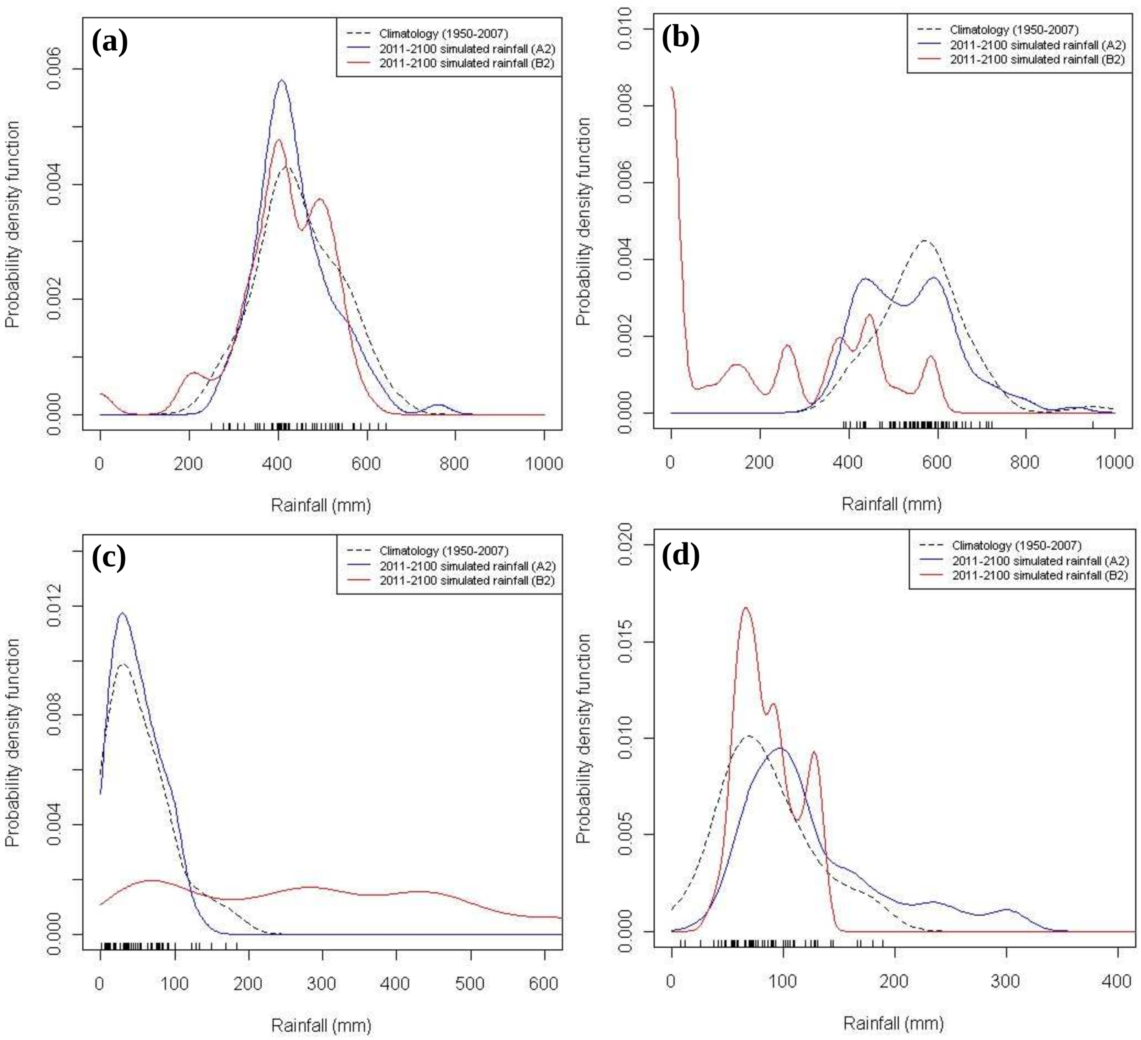


Figure 5.19: PDF of 2011-2100 median rainfall estimated from 300 rainfall ensembles during (a) MJJ; (b) ASO; (c) NDJ; and (d) FMA.

From the climatological PDF of MJJ rainfall, the chance of MJJ rainfall being less than 381.7 mm (i.e. dry condition) is 20%. Under A2 (B2), MJJ rainfall from 2011 to 2100 has higher chances of falling under the dry category, by 26.4% (33.3%). The 1950-2007 MJJ rainfall had a 20% chance of being wet, whereas the 2011-2100 MJJ rainfall shows probabilities of rainfall greater than 528.9 mm (i.e. wet condition) by 16.6% and 15.2% under A2 and B2 respectively. On the other hand, compared to a chance of 20% obtained from climatology, a probability of 39.7% (84.5%) is associated with the 2011-2100 ASO rainfall, with a higher chance of ASO rainfall being dry under A2 (B2). The probability that the 2011-2100 ASO rainfall will be above 624.8 mm (i.e. the wet condition) is estimated at 22% under A2 and 5.1% under B2. As expected, NDJ rainfall under A2 (B2) during 2011- 2100 has a lower chance of being dry by 15.9% (8.2%), versus a 20% chance of climatological NDJ rainfall. However, the 2011-2100 NDJ rainfall indicates more chances of being wet, by 21.6% and 83.3% under A2 and B2 respectively. For FMA rainfall, the PDF of climatology has a chance of being dry by 20%; however, the PDF of 2011-2100 FMA rainfall presents a chance of being dry by only 9.4% in A2 and 12.2% under B2. The probabilities of FMA rainfall greater than 114.5 mm (the wet condition) are estimated at 50.7% and 19.2% under A2 and B2 respectively.

To consider the probabilities of anomalous weather events in each year of simulation, the PDF of 300 rainfall ensembles for each simulation year are computed. The probabilities are calculated at the same thresholds as previously mentioned. A rainfall ensemble below the 20th percentile is defined as below-normal or dry, whereas a rainfall ensemble above the 80th percentile is denoted as above-normal or wet. Otherwise, rainfall falls into the normal category. Out of 300 ensembles in each simulation year, the probability of each category depends on the proportion of rainfall members falling into that category. A probability of 100% for a category shows that the all 300 ensembles fall into that category. Figure 5.20 and 5.21 show probabilities of more than 50% for all the three categories under A2 and B2 scenarios respectively.

From Figure 5.20, we can see that with a high probability of occurrence (>70%), MJJ rainfall from 2011 to 2100 will be below 381.7 mm in 10 years and will be above 528.9 mm in 8 years under A2 scenario. Out of the 8 years of being above-normal, 5 years (2015, 2029, 2043, 2073 and 2079) indicate a chance of occurrence above 90%. Also with a high probability (>70%), ASO rainfall under A2 from 2011 to 2100 shows 20 dry years (or below-normal) and 11 wet years (or above-normal). With a chance of occurrence greater than 90%, there are 7 (9) years out of these dry (wet) years of ASO rainfall. In the period of 2011-2100, dry NDJ and FMA rainfalls under A2 will be hardly observed. However, with a probability of occurrence of above 70%, above-normal NDJ and FMA rainfalls will be seen in 8 and 29 years respectively. Above-normal FMA rainfall tends to occur after 2070, which is consistent with the rainfall anomalies seen in Figure 5.18(d). There are 21 years out of 29 years of above-normal FMA rainfall obtained at a probability greater than 90% and also indicating a 4-year consecutive period of the wet condition from 2085 to 2088.

On the other hand, under B2 (Figure 5.21), 18 out of 90 simulation years show that MJJ rainfall will be below 381.7 mm with a probability greater than 70%. Out of these 18 dry years, there are 9 years showing the probability of occurrence greater than 90%. Above-normal MJJ and ASO rainfall with a probability above 60% will not be found during 2011 to 2100. However, with a probability greater than 90%, below-normal ASO rainfall will be observed in several years especially during a long consecutive period from 2046 to 2100 –

i.e. 55 years. In contrast, from 2011 to 2100, below-normal NDJ and FMA rainfall is hardly found with a high probability of occurrence (>70%). The above-normal condition will occur with high probability, especially for NDJ rainfall, which indicates a long period of the wet condition after 2045. The high probabilities of below-normal ASO rainfall and above-normal NDJ rainfall in a long consecutive period after 2046 are consistent with the annual anomalies shown in Figure 5.18(b) and (c).

Therefore, the effects of future climate on anomalous weather events during 2011- 2100 indicate more (less) chances of rainfall being below-normal (above-normal) during the wet seasons (i.e. MJJ and ASO). In contrast, for rainfall during dry seasons (i.e. NDJ and FMA), the effects of future climate suggest more (less) chances of the rainfall being above-normal (below-normal). Moreover, anomalous weather events will occur during several years from 2011 to 2100 with a probability greater than 70%. In particular under B2, monsoon rainfall will tend to be delayed by one season from ASO to NDJ, associated with a significant decreasing (increasing) trend of ASO (NDJ) rainfall and a long consecutive period of below-normal (above-normal) conditions from 2046 to 2100.

5.5 Summary

The modified k-nn model is a statistical model applied to downscale seasonal rainfall in the Ping River Basin from large-scale atmospheric variables and to assess the effects of future climate. Under the two scenarios of future climate (A2 and B2), the modified k-nn model performs moderately well when capturing the annual variability of seasonal observed rainfall from 1962 to 2007. The model also has the capability of preserving annual statistics (i.e. mean, median, SD, IQR and skew) of seasonal observed rainfall. The absolute biases of simulation are below 32%. Moreover, the median LLH score of 1962-2007 rainfall simulations is greater than +1.0, which indicates better performance of the model in capturing PDF than climatology.

The multisite daily rainfall generator is developed to resample the historical daily rainfall at several rainfall stations. It performs well when capturing transition probabilities, the dry- and wet-spell statistics and basic statistics of observations in all 12 months. The multisite generator can also preserve basic statistics and cross-correlations among pairs of rainfall stations.

Under both scenarios of future climate, the 2011-2100 rainfall in the Ping River Basin during wet seasons (MJJ and ASO) tends to decrease by 0.11-6.16 mm per year, whereas the 2011-2100 rainfall of dry seasons (NDJ and FMA) tends to increase by 0.02-5.91 mm per year. As for anomalous weather events, the effects of future climate suggest more chances of being below-normal and less chances of being above-normal for the wet season rainfall. On the other hand, the dry season rainfall shows more chances of being above-normal and less chances of being below-normal. Future climate also seems to have an effect on the monsoon schedule as the monsoon period shifts by one season from ASO to NDJ.

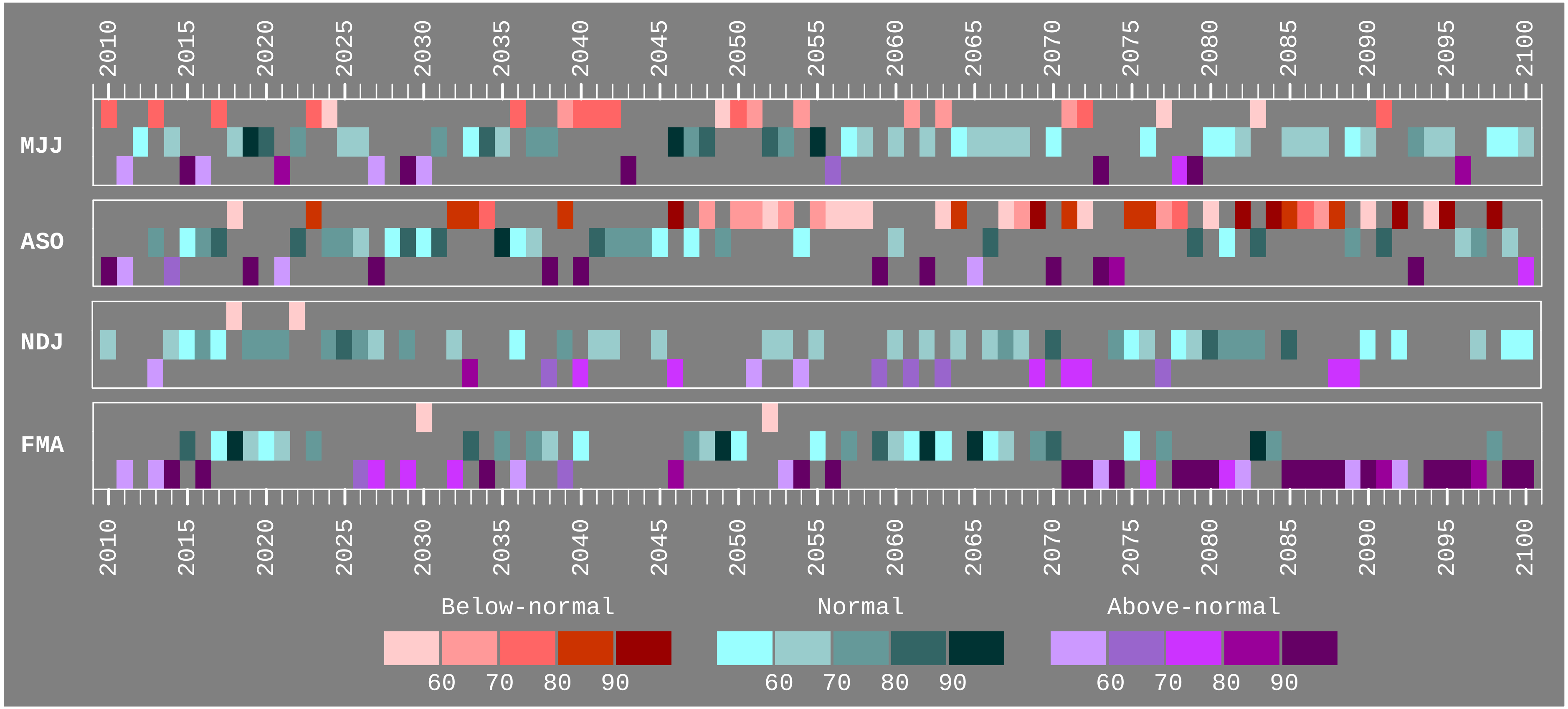


Figure 5.20: Probabilities of anomalous weather events annually in simulations from 2011 to 2100 under Scenario A2 of future climate.

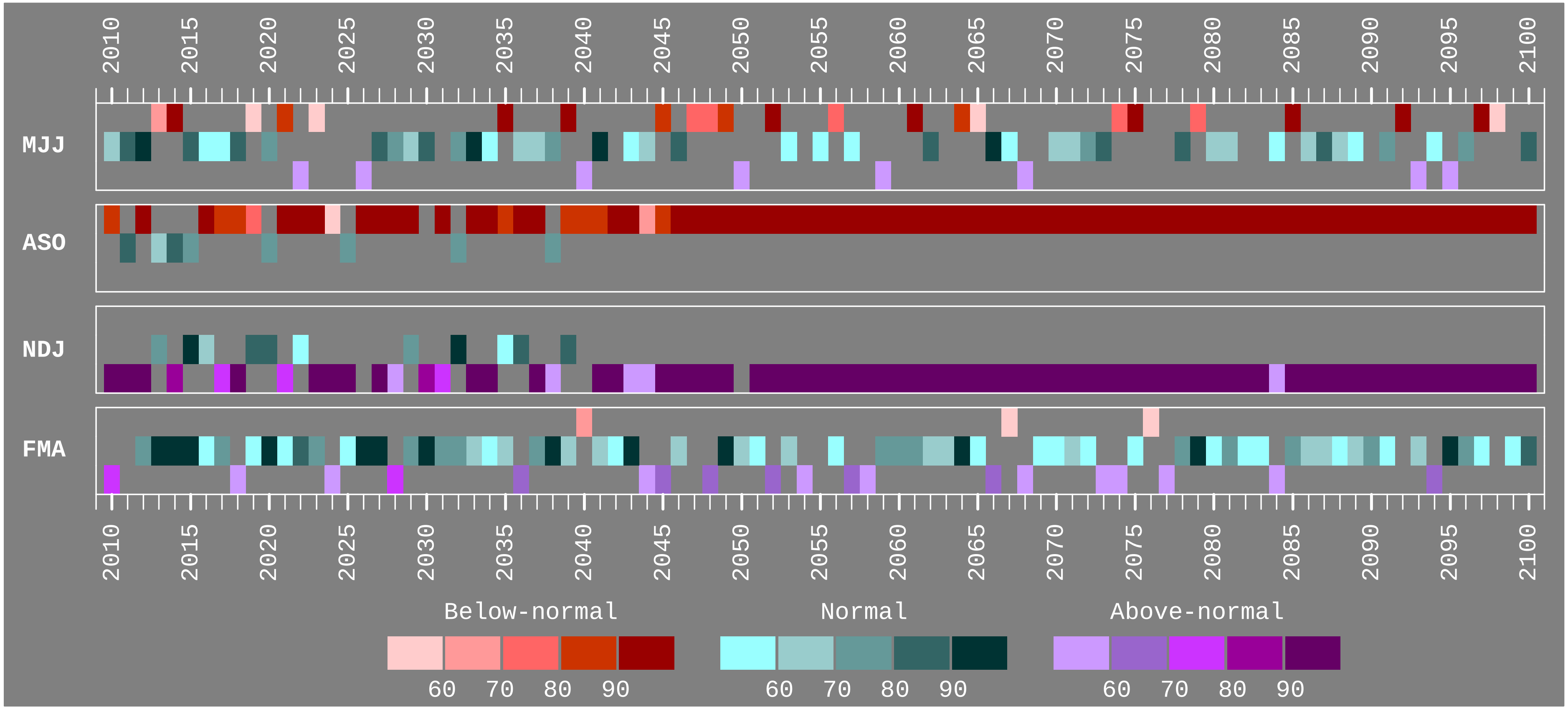


Figure 5.21: Same as Figure 5.20 but for Scenario B2 of future climate.

Chapter 6
Hydrologic Behavior of River Basins

6.1 Introduction

To show model dependency of the daily streamflow simulation for determining the effects of future climate on streamflow, this chapter compares the performance of two proposed rainfall-runoff models, i.e. the SIMHYD (Chiew et al., 1996) and the HEC-HMS (U.S. Army Corps of Engineers, 2000) models. The best model is then selected based on a better performance as per four efficiency indexes, i.e. deviation of volume, correlation coefficient, normalized root mean square error (NRMSE) and the Nash-Sutcliffe efficient index. The selected rainfall-runoff model has been adopted in Chapter 7 to simulate streamflow in the Ping River Basin using 2011-2100 daily rainfall ensembles obtained from the multisite daily rainfall generator (Chapter 5).

6.2 Data description

6.2.1 Basin data

The Ping River Basin can be divided into 20 sub-basins (see also Figure 3.4 and Table 3.1) of which 14 sub-basins are located upstream of the Bhumipol Dam. The rainfall-runoff simulation for the study is done within these 14 sub-basins and aims to apply the modeled results in the operation of the Bhumipol Dam and the management of the reservoir. The area of the 14 sub-basins ranges from 535 to 3,143 km^2 (Table 6.1). This area is covered by agriculture, urban development, forests, water resources, irrigated agriculture and miscellaneous land use activities (Figure 6.1). The forests are estimated to cover 55.60% to 97.22% of the sub-basin area. Based on the soil classification done by the Land Development Department (LDD) of Thailand (LDD, 2011), the soil type of the Ping River Basin is a slope complex (SC) that is not suitable for agriculture since the slope is more than 35% and there is a high degree of erosion of the soil surface.

Table 6.1: Description of the 14 Sub-basins Located Upstream of the Bhumipol Dam

Sub-basin code	Sub-basin name	Area (km^2)	Land use coverage (%)						Soil type
			A	U	F	W	IA	M	
0602	Upper Ping Part	2,018	9.51	0.73	89.28	0.32	0.00	0.16	SC
0603	Mae Ngad	1,260	7.18	0.79	91.38	0.38	0.00	0.28	SC
0604	Mae Taeng	1,761	10.13	0.70	89.07	0.05	0.03	0.02	SC
0605	2nd Ping Part	1,624	30.21	10.71	55.60	0.43	0.00	3.04	SC
0606	Mae Rim	584	6.11	1.28	92.51	0.11	0.00	0.00	SC
0607	Mae Kuang	1,165	19.84	5.41	73.12	0.55	0.00	1.08	SC
0608	Mae Ngan	1,711	7.23	1.62	90.53	0.11	0.00	0.51	SC
0609	Mae Li	1,956	12.84	1.63	82.80	0.11	0.00	2.61	SC
0610	Mae Klang	600	8.47	0.90	89.30	0.08	0.00	1.25	SC
0611	3rd Ping Part	3,071	3.69	0.63	92.60	1.83	0.00	1.25	SC
0612	Upper Mae Cham	1,912	2.53	0.16	97.22	0.00	0.00	0.09	SC
0613	Lower Mae Cham	1,926	4.57	0.25	94.96	0.08	0.00	0.14	SC
0614	Mae Had	535	6.25	0.61	92.29	0.39	0.00	0.45	SC
0615	Mae Tuen	3,143	3.31	0.09	96.02	0.57	0.00	0.01	SC

A: agriculture; U: urban; F: forest; W: water resource; IA: irrigated agriculture; M: miscellaneous; SC: slope complex.

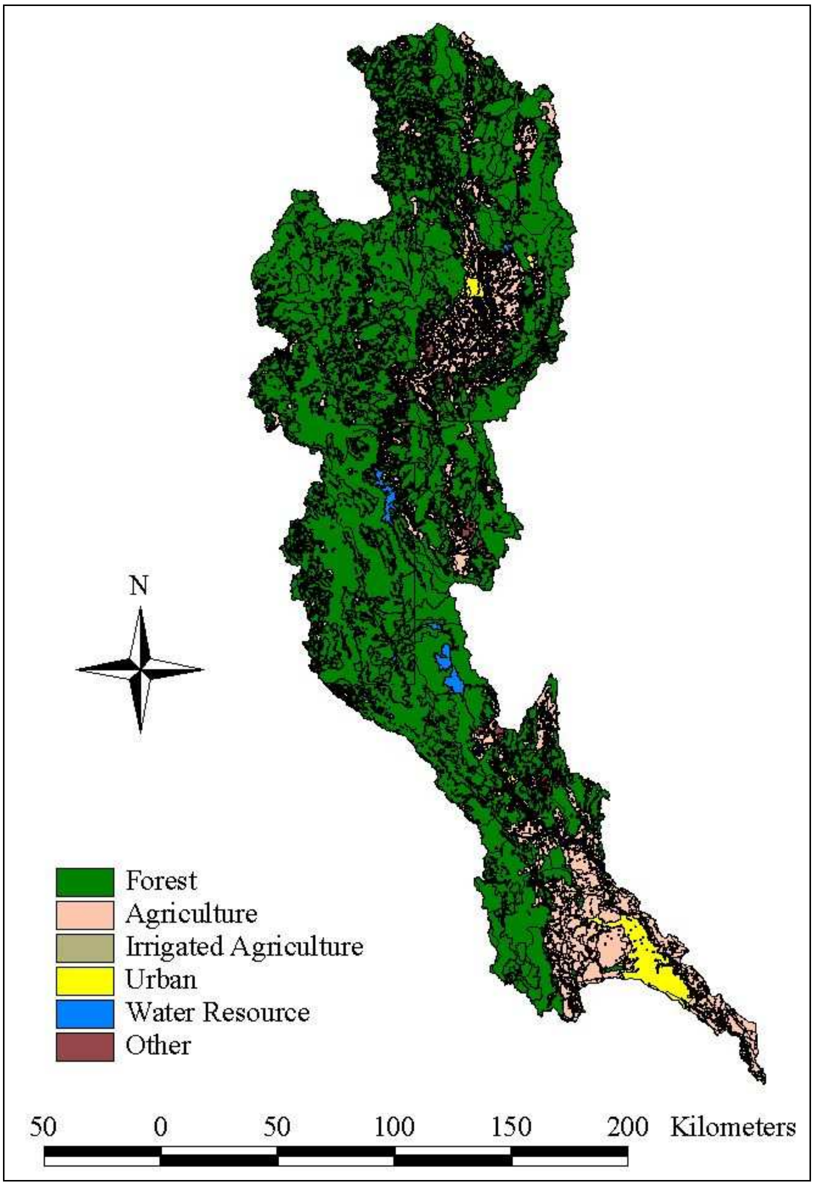


Figure 6.1: Land use map of the Ping River Basin.

6.2.2 Rainfall data

The rainfall daily data from 50 selected stations is used in the rainfall-runoff models (see also Appendix A2 and Figure 3.5). The average daily values over the area of each sub-basin are calculated using the Thiessen method (Linsley et al., 1988; Gupta, 1989). A weighting factor for each rainfall station is assigned proportionately in a representative polygon for the total sub-basin area. The dimension of the assigned polygon is based on the non-uniform distribution of rainfall stations. The average rainfall computed by the Thiessen method is more accurate than the arithmetical average; however, the Thiessen method is not flexible enough for the changes in the rainfall station network. The representative polygon for a station has to be re-assigned when a new rainfall station is installed within the basin, or if the location of an existing station is changed.

6.2.3 Evaporation data

Daily evaporation from 1969 to 2007 is measured by a class A pan in 18 meteorological stations located in and around the Ping River Basin (see also Appendix A5 and Figure 3.5). The daily and monthly evaporation are averaged over the 18 stations. Minimum evaporation is observed in December with an average of 3.2 mm d^{-1} (Figure 6.2), while the highest evaporation occurs during the summer season, i.e. from March to May. Maximum evaporation is found in April, estimated at 6.1 mm d^{-1}. To compute potential evapotranspiration (ET$_p$), Chankaew (1996) and Michalczyk (2008) suggest using a class A pan with a coefficient equal to 0.7 for Thailand.

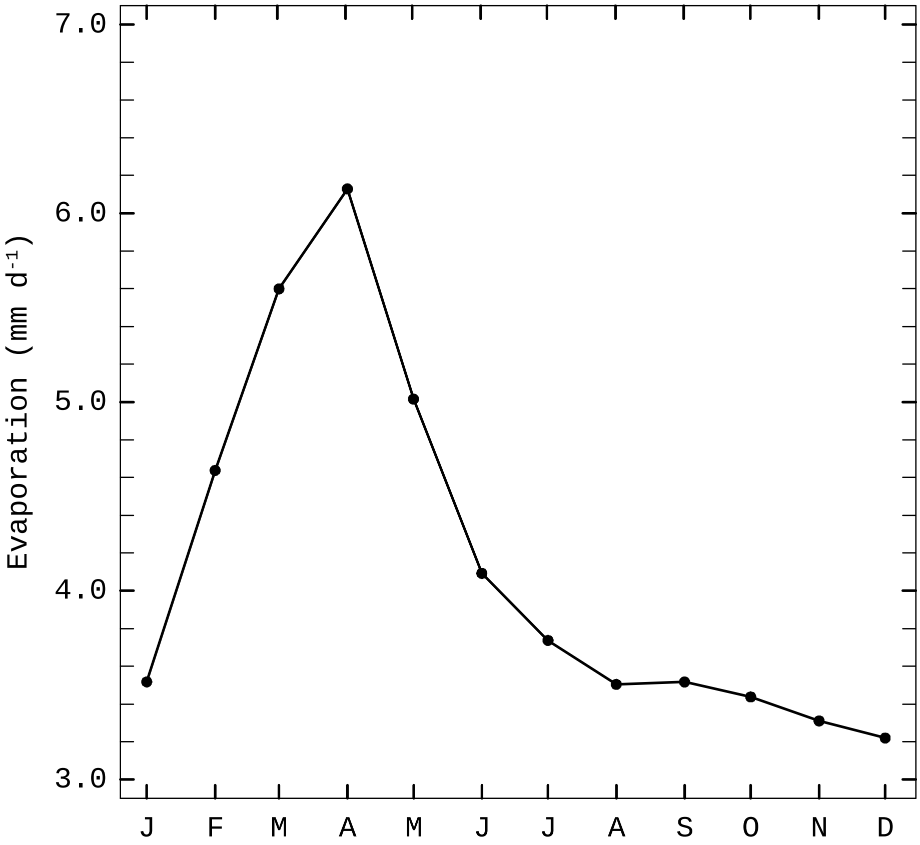


Figure 6.2: Annual cycle of evaporation from a class A pan.

6.2.4 Streamflow data

Out of 45 gauging stations located in the Ping River Basin (Appendix A3 and Figure 3.5), 12 streamflow stations have been selected based on the criteria of no incomplete daily data during a consecutive period which is consistent among all the stations. Station P12B is located at the Bhumipol Dam to measure the inflow. It is noted that an annual water cycle of streamflow in Thailand starts in the driest month (i.e. April) and ends in March of the following year. The consecutive period which has no incomplete data available from the 12 selected stations is from April 1999 to March 2007 (i.e. eight water years). The drainage area (D.A.) of the 12 selected stations ranges from 460 to 26,396 km^2 (Table 6.2). The minimum and maximum of 1999-2007 average annual runoff are estimated at 142.76 and 6,029.89 MCM at Station P21 (i.e. the Mae Rim Basin) and P12B (Bhumipol Dam Station) respectively. The peak runoff (i.e. wet season) is observed from August to November (ASON) (Figure 6.3) corresponding to the monsoon season (i.e. August-September-October: ASO). The runoff during the wet season is estimated to be 58-73% of the total annual runoff. The runoff in the dry season from December to April is calculated to be 8-23%. The remaining runoff is observed from May to July (MJJ), i.e. in the pre-monsoon season. Appendix E1 to E12 present the hydrographs of daily observed streamflow at the 12 gauging stations from April 1999 to March 2007.

Table 6.2: List of the 12 Selected Streamflow Stations

Station	Sub-basin	Stream/ River	Drainage area (km^2)	Averaged daily discharge* ($m^3 s^{-1}$)	Averaged annual runoff* (MCM)
P75	Upper Ping Part and Mae Ngad	Ping	3,080	24.01	757.93
P4A	Mae Taeng	Mae Taeng	1,902	13.42	423.66
P67	-	Ping	5,289	45.01	1,420.68
P21	Mae Rim	Mae Rim	515	4.52	142.76
P1	-	Ping	6,355	51.93	1,639.13
P71	Mae Ngan	Mae Khan	1,771	11.01	347.47
P24A	Mae Klang	Mae Klang	460	4.99	157.45
P73	-	Ping	13,353	131.83	4,160.51
061302	Upper Mae Cham	Mae Cham	1,946	22.14	698.94
P14	-	Mae Cham	3,853	36.89	1,164.40
061501	Mae Tuen	Mae Tuen	1,470	24.03	758.53
P12B	-	Ping	26,396	191.06	6,029.89

* from 1st April 1999 to 31st March 2007.

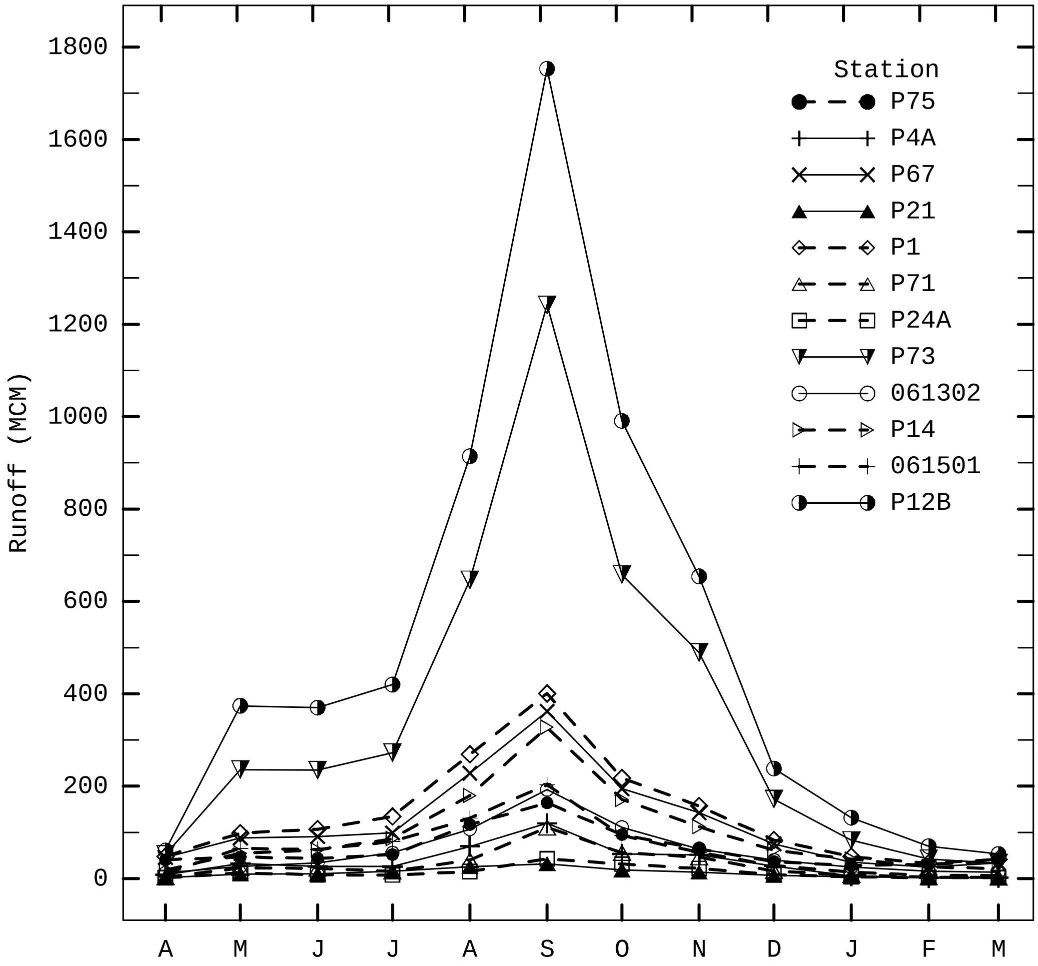


Figure 6.3: Annual cycle of streamflow at the 12 selected stations.

6.3 Methodology

To compare different algorithms of rainfall-runoff models and to show model dependency of the daily streamflow simulation, two models have been proposed in this study: the SIMHYD and HEC-HMS models. The parameters of both models are calibrated for the period of 1st April 1999 to 31st March 2003 and validated from 1st April 2003 to 31st March 2007. The model performances are evaluated using four efficiency indexes (deviation of volume (D_v), correlation coefficient (r), normalized root mean square error (NRMSE), and the Nash-Sutcliffe efficiency index (E)). The model with better performance is ultimately selected to simulate daily streamflow from 2011 to 2100 in order to determine the effects of future climate. The methodology for both models is described as follows.

6.3.1 The SIMHYD model

The SIMHYD model is a lumped conceptual rainfall-runoff model developed by the Cooperative Research Center for Catchment Hydrology (CRCCH), Australia. It is included in a software package called the Rainfall Runoff Library (RRL), which contains five rainfall-runoff models – AWBM, Sacramento, SIMHYD, SMAR and TANK. Each model has different objectives; for example, the AWBM model computes the water balance of a basin for flood hydrograph modeling. The RRL is a free software and available at http://www.toolkit.net.au/Tools/RRL.

The SIMHYD model simulates daily streamflow using the continuous time series of daily rainfall and average areal potential evapotranspiration (Podger, 2004; Chiew et al., 1996). The SIMHYD model has nine parameters: baseflow coefficient, impervious threshold, infiltration coefficient, infiltration shape, interflow coefficient, pervious fraction, rainfall interception store capacity (RISC), recharge coefficient, and soil moisture store capacity (SMSC). The default values of the nine parameters are shown in Table 6.3.

Table 6.3: Default Values of the Nine Parameters of the SIMHYD Model

Parameter	Default value	Default minimum	Default maximum
Baseflow coefficient (-)	0.3	0.0	1.0
Impervious threshold (-)	1.0	0.0	5.0
Infiltration coefficient (-)	200.0	0.0	400.0
Infiltration shape (-)	3.0	0.0	10.0
Interflow coefficient (-)	0.1	0.0	1.0
Pervious fraction (-)	0.9	0.0	1.0
RISC (mm)	1.5	0.0	5.0
Recharge coefficient (-)	0.2	0.0	1.0
SMSC (mm)	320.0	1.0	500.0

The algorithm of the SIMHYD model is presented in Figure 6.4. Rainfall and potential evapotranspiration (ET_p) are specified as the inputs for the model. The total runoff of a basin is a sum of the runoff from pervious and impervious areas. The runoff from the pervious areas involves the infiltration excess runoff, the saturation excess runoff or interflow and the baseflow. In pervious areas, the rainfall first satisfies ET_p by intercepting ET which comes from trapped water by interception storage like a canopy. Then, the throughfall moisture, which is the rainfall remaining after interception ET, infiltrates to the soil. The infiltration excess runoff is estimated depending on the infiltration coefficient and infiltration shape. Subsequently, based on the interflow coefficient, recharge coefficient and SMSC, the infiltrated moisture is diverted to the river in terms of either saturation excess runoff or interflow, depending upon the current state of soil wetness. The baseflow is calculated using a baseflow coefficient. In the impervious area, the runoff is estimated as the moisture that exceeds ET in the impervious area (i.e. impervious ET).

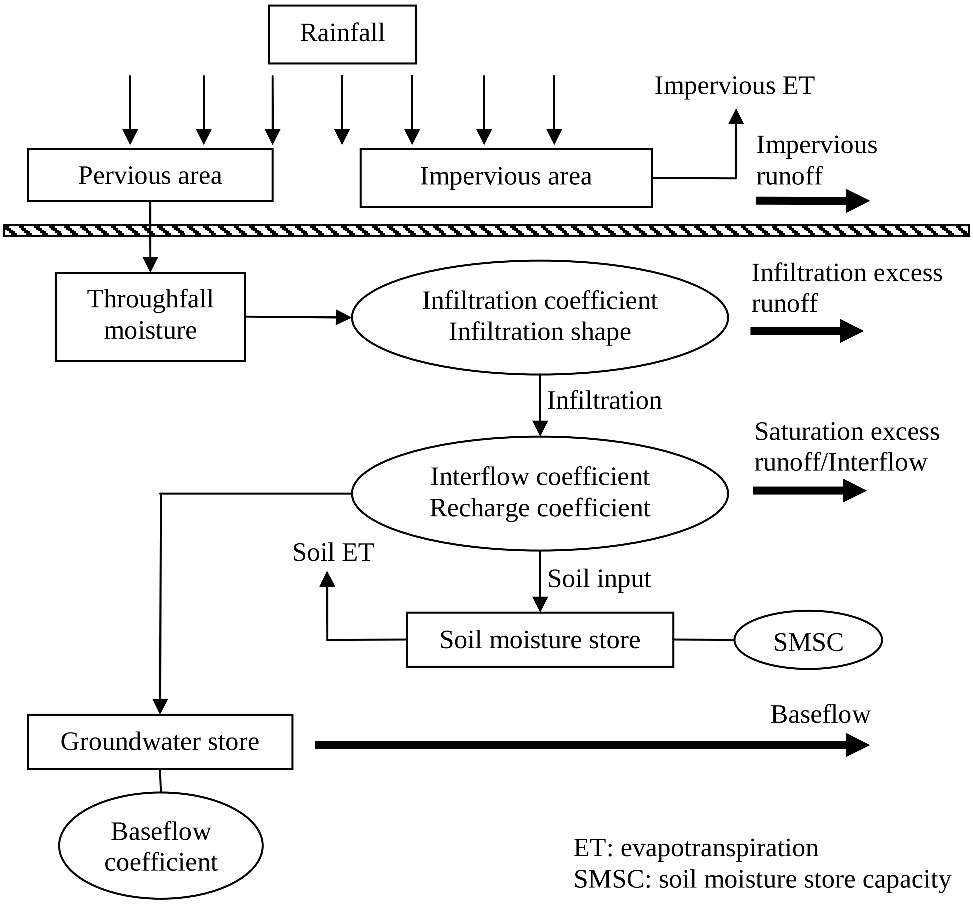


Figure 6.4: Structure of the SIMHYD model.

The SIMHYD model provides several methods of parameter optimization. Optimization is a function of automatic calibration of nine parameters that give the best value of the objective functions. The primary objective functions involve the Nash-Sutcliffe, sum of square of errors (SSE), root mean square error (RMSE), root mean square (RMS) difference about bias, absolute value of bias, the sum of square roots, the sum of the squares of the difference of square roots, and the sum of the differences of logs. Moreover, optional secondary objective functions include the runoff difference in percentage, flow duration curve and baseflow method 2. Among several optimization methods, the simplest method is uniform random sampling (Podger, 2004). Other methods are pattern search (single- and multi-start), Rosenbrock (single- and multi-start), the genetic algorithm and shuffled complex evolution (SCE-UA). In this study, the pattern search multi-start has been applied to optimize the nine parameters of the SIMHYD model. The pattern search multi-start is a simple and quick method for optimization that can overcome the problem of reaching local optimums by finding global optimums.

Due to a small number of model parameters and the provided parameter optimization, the SIMHYD model is easy to apply. Calibration and validation using SIMHYD in this study are done at six streamflow stations (P4A, P21, P71, P24A, 061302 and 061501) which are located at the outlets of certain sub-basins (Mae Taeng, Mae Rim, Mae Ngan, Mae Klang, Upper Mae Cham and Mae Tuen respectively). At other six gauging stations that measure combined flow from several sub-basins, the SIMHYD model, which is a lumped conceptual model, is not suitable for application due to its algorithm meant for a basin with homogeneous characteristics.

6.3.2 The HEC-HMS model

The Hydrologic Modeling System (HMS) is a free software developed by the Hydrologic Engineering Center (HEC) of the U.S. Army Corps of Engineers, USA and available at http://www.hec.usace.army.mil/software/hec-hms. The HEC-HMS model simulates continuous streamflow or runoff over longer periods of time using precipitation like rainfall and snow and ET_p as the input (U.S. Army Corps of Engineers, 2000). The HEC-HMS model provides both: lumped simulation and spatially distributed simulation using a grid cell depiction of the basin characteristics. In the model, the outflow of a basin is computed as the sum of the direct runoff (i.e. surface flow) and baseflow (i.e. subsurface flow). To compute direct runoff and baseflow, the HEC-HMS model has four components: loss, transform, baseflow and routing.

The loss component provides seven optional methods to calculate the total loss of a basin including actual evapotranspiration (ET_a), surface depression and soil infiltration. The seven methods are deficit constant, initial constant, exponential, Green and Ampt, SCS curve number, soil moisture accounting (SMA) and Smith Parlange. The gridded loss is also available in the deficit constant, Green and Ampt, SCS curve number and SMA. In the transform component, five methods (the Clark unit hydrograph, kinematic wave, ModClark, SCS unit hydrograph and Snyder unit hydrograph) are provided to transform excess precipitation into surface runoff. The hydrograph of the transform component can also be specified by users with two available options: (i) the user-specified S-graph; and (ii) the user-specified unit hydrograph. To estimate the baseflow of a basin, there are five methods provided in the baseflow component of the HEC-HMS model. These include bounded recession, the constant monthly baseflow, the linear reservoir, the nonlinear Boussinesq and recession. Furthermore, the routing component calculates the outflow in an open channel which combines several inflows coming from one or more elements (sub-basins) in the basin. The routing component has six optional methods: kinematic wave, lag, modified pulse, Muskingum, Muskingum Cunge and straddle stagger. Each method of the four components (loss, transform, baseflow and routing) requires different parameters (Appendix F1) which can be calibrated automatically by optimization trials. Two methods (i.e. univariate gradient, and Nelder and Mead) are available to determine the optimal parameters that give the best value of an objective function. The objective functions include the peak-weighted RMS error, sum of squared residuals, sum of absolute residuals, percentage error in peak flow, percent error in volume, RMS log error and time-weighted function (U.S. Army Corps of Engineers, 2000).

Figure 6.5 shows the algorithm of the HEC-HMS model. The total runoff is the total of the runoff from pervious and impervious surfaces. The runoff from pervious surfaces is subject to losses such as ET and infiltration, whereas the runoff from impervious surfaces is estimated without considering ET, infiltration and other losses. In this process, the precipitation on pervious surfaces is trapped by canopy interception storage on trees, shrubs and grasses. Then, the moisture in the canopy storage which cannot reach the soil surface evaporates to the atmosphere. It is also important to note that ET loss is estimated on no precipitation days and cannot exceed atmospheric ET_p. So, the canopy loss depends on maximum canopy storage and ET_p. Subsequently, precipitation which is not trapped by the canopy interception storage reaches the ground surface and is captured by surface interception storage. Based on the maximum soil storage capacity and infiltration rate, the moisture stored in the surface depressions either infiltrates the soil or evaporates into the atmosphere. The remaining depth from the total loss is defined as precipitation excess, and is ultimately transformed into surface runoff. Furthermore, the baseflow that is calculated based on a selected method is added to the surface runoff to obtain the total runoff.

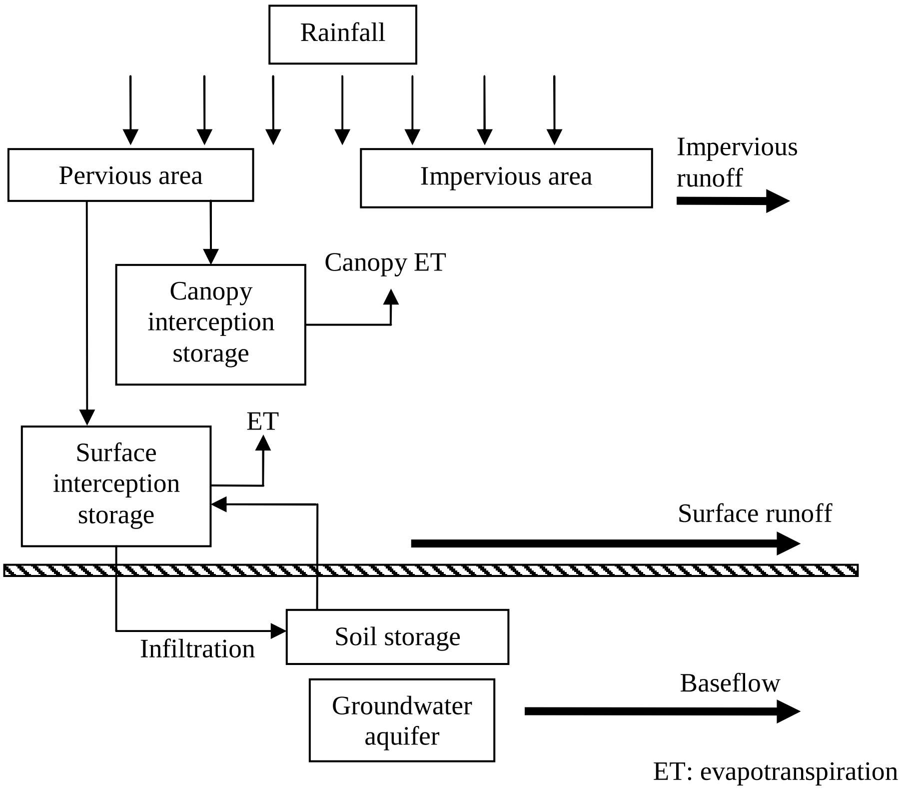


Figure 6.5: Structure of the HEC-HMS model.

Figure 6.6 presents the schematic of a HEC-HMS model for the Ping River Basin. In this study, the HEC-HMS model is calibrated and validated at 12 streamflow stations. Six stations are located at the outlets of different sub-basins to measure the streamflow of tributaries (e.g. Station P4A measures streamflow of the Mae Taeng (see also Table 6.2) before the streamflow drains into the main river, Ping. Six other stations are located on the Ping River to measure combined flow from several sub-basins (e.g. Station P75 measures combined streamflow from the Upper Ping part Basin and the Mae Ngad Basin). The inflow of the Bhumipol Dam coming from the Ping River Basin is observed at Station P12B. The streamflow simulation at this station can be used for the development of management strategies of the reservoir and operational purposes of the Bhumipol Dam in accordance with the available inflow.

In this study, the deficit and constant method is adopted to compute total loss in each sub-basin. This method has eight parameters (Table 6.4) related to canopy interception loss, surface depression loss and infiltration loss. The Clark unit hydrograph is selected to transform excess rainfall into surface runoff. Two parameters are involved in this method. The recession method is used to compute baseflow with three parameters. The lag method having one parameter is applied for the routing component.

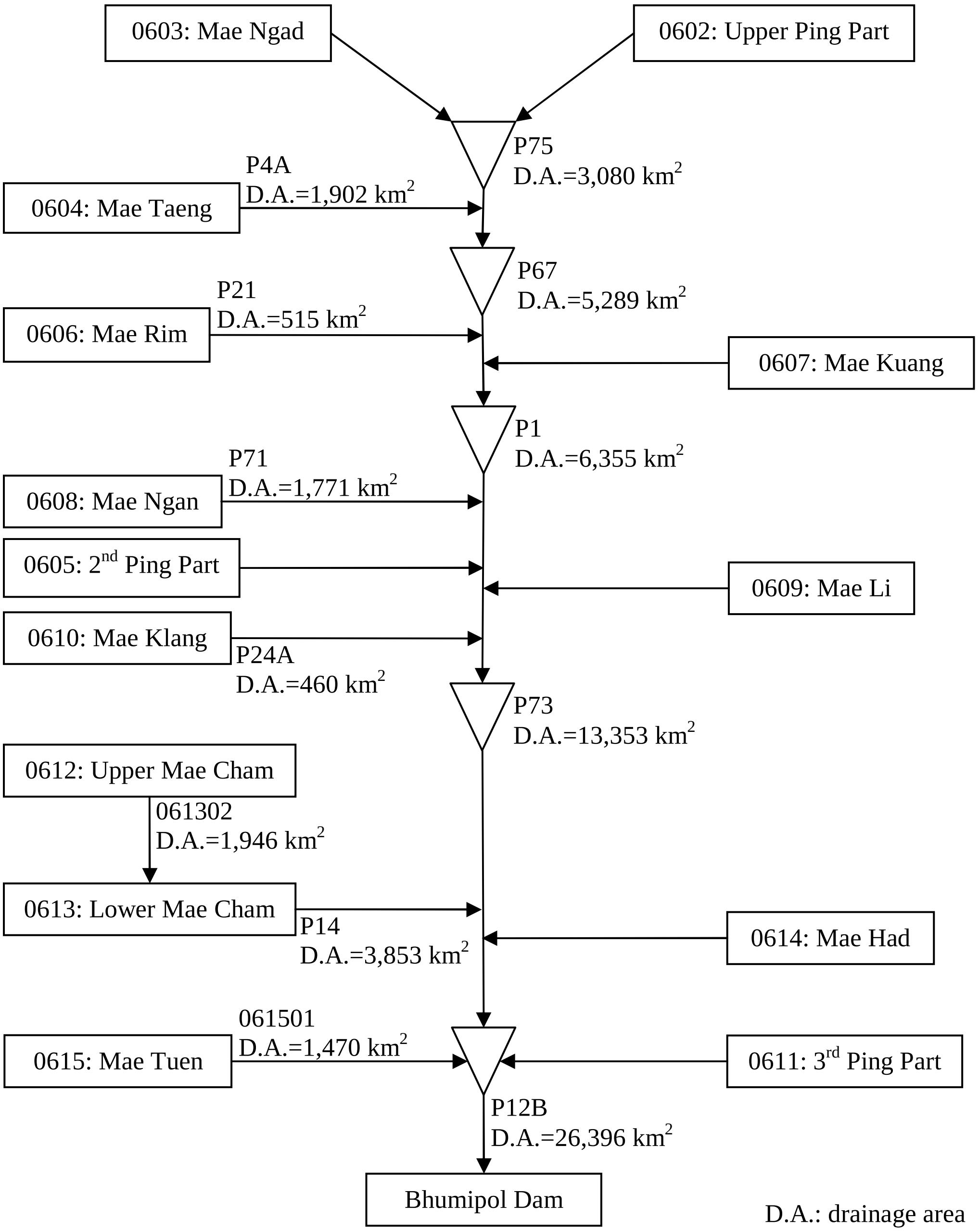


Figure 6.6: Schematic of the HEC-HMS model for the Ping River Basin.

Table 6.4: List of the Adopted Method of Each Component in the HEC-HMS Model and Its Parameters

Component/Method	Parameter
Loss component	
Deficit and constant	Initial canopy storage (%) Maximum canopy storage (mm) Initial surface storage (%) Maximum surface storage (mm) Initial deficit (mm) Maximum deficit (mm) Infiltration constant rate (mm h^{-1}) Imperviousness (%)
Transform component	
Clark unit hydrograph	Time of concentration (h) Storage coefficient (h)
Baseflow component	
Recession	Initial discharge ($m^3 s^{-1}$) Recession constant (-) Ratio to peak (-)
Routing component	
Lag	Lag time (min)

6.4 The model performances

The SIMHYD and HEC-HMS models are calibrated from 1st April 1999 to 31st March 2003 and validated from 1st April 2003 to 31st March 2007. The performance of both models is evaluated based on four efficiency indexes, including the deviation of volume (D_v), correlation coefficient (r), normalized root mean square error (NRMSE), and the Nash-Sutcliffe efficient index (E) as shown in Equation 6.1-6.4 respectively.

$$D_v = \frac{\bar{Q}_m - \bar{Q}_o}{\bar{Q}_o} * 100$$ Equation 6.1

$$r = \frac{\sum_{t=1}^{T}(Q_{m,t} - \bar{Q}_m)(Q_{o,t} - \bar{Q}_o)}{\sqrt{\sum_{t=1}^{T}(Q_{m,t} - \bar{Q}_m)^2}\sqrt{\sum_{t=1}^{T}(Q_{o,t} - \bar{Q}_o)^2}}$$ Equation 6.2

$$NRMSE = \frac{\sqrt{\left(\sum_{t=1}^{T}(Q_{m,t} - Q_{o,t})^2\right)/T}}{\bar{Q}_o}$$ Equation 6.3

$$E = 1 - \frac{\sum_{t=1}^{T}(Q_{m,t} - Q_{o,t})^2}{\sum_{t=1}^{T}(Q_{o,t} - \bar{Q}_o)^2}$$ Equation 6.4

where $Q_{m,t}$ and $Q_{o,t}$ are the daily modeled and observed discharges on day *t*, and $\bar{Q}_m$ and $\bar{Q}_o$ are the averages of daily modeled and observed discharge respectively.

D_v is applied to determine the underestimation or overestimation of average simulated streamflow compared to the observations. D_v ranges from $-\infty$ to $+\infty$; however, a value nearly 0 indicates a better performance of the model. Underestimation and overestimation are defined as the negative and positive values respectively of D_v.

r is used to test the relation between simulated and observed streamflow. It varies from -1.0 to +1.0; a value of +1.0 denotes a perfect fit between simulated and observed streamflow. However, a value of *r* greater than +0.7 can indicate satisfactory simulation (Lévesque et al., 2008).

NRMSE represents the deviation of error obtained from simulation. A smaller NRMSE with a value close to 0 denotes a small error from a model.

E can also determine standardized error. It ranges from $-\infty$ to +1.0. The value of 0 represents no difference of the model performance over the average of observed streamflow. The best fit is associated with a value nearly equal to +1.0; however, a value above +0.5 indicates a satisfactory efficiency of the model (Lévesque et al., 2008).

In this study, the model performance in capturing the variability of low and high streamflow has been evaluated. The means and medians of historical data during calibration and validation (Table 6.5) are used as the thresholds of low and high streamflow. Daily discharge that is less than the mean (median) is defined as low flow, i.e. below-mean (below-median) streamflow. Otherwise, the daily discharge is denoted as high flow, i.e. above-mean (above-median) streamflow. The efficiency indexes (D_v, NRMSE and *E*) are then adopted and computed separately for each category.

Table 6.5: Statistics of the Daily Historical Streamflow

Station	(a) Calibration				(b) Validation			
	Mean ($m\ s^{-1}$)	Median ($m\ s^{-1}$)	CV	Skew	Mean ($m\ s^{-1}$)	Median ($m\ s^{-1}$)	CV	Skew
P75	19.04	14.20	1.04	3.65	28.97	20.00	1.16	3.83
P4A	11.71	4.28	1.52	3.11	15.14	4.38	1.86	4.17
P67	34.57	21.78	1.20	3.63	55.45	25.92	1.45	3.98
P21	4.51	3.17	1.36	4.45	4.54	2.14	1.54	3.94
P1	42.95	29.80	1.07	3.44	60.91	33.20	1.36	3.81
P71	12.43	5.41	1.78	4.62	9.59	3.20	2.14	5.07
P24A	5.70	2.69	1.89	6.00	4.27	1.81	1.98	6.53
P73	128.96	63.85	1.43	3.00	134.69	59.50	1.46	2.71
061302	21.50	13.30	1.21	5.85	22.78	12.80	1.63	7.03
P14	37.33	22.74	1.12	3.56	36.45	19.90	1.78	7.54
061501	28.28	11.50	1.59	3.79	19.79	9.08	1.57	4.61
P12B	197.63	118.40	1.30	3.04	184.48	79.75	1.52	2.90

Calibration: 1st April 1999 to 31st March 2003; validation: 1st April 2003 to 31st March 2007.
CV: coefficient of variation=standard deviation/mean.

6.4.1 Calibration and validation of the SIMHYD model

The SIMHYD model is calibrated with the primary objective of preserving the variability of average monthly streamflow. For calibration (i.e. from April 1999 to March 2003), taking advantage of parameter optimization, nine parameters of the SIMHYD model are optimized by the pattern search multi-start method with the secondary objective of maximizing the Nash-Sutcliffe efficient index (*E*), which happens to be one of the objective functions provided by the model. At six gauging stations, calibration is done separately for each streamflow station. As shown in Figure 6.7, the SIMHYD model cannot capture the variability of average monthly streamflow except at Station P21 and 061302. Automatic calibration computes a large amount of actual evapotranspiration (ET_a) even during the monsoon season (i.e. ASO). Hence, the overestimation of loss is observed at all stations. In the SIMHYD model, ET_a is a sum of interception ET, impervious ET and soil ET; however, ET_a is primarily estimated by interception ET which depends upon a parameter, namely the rainfall interception store capacity (RISC (mm)). To improve ET_a calculation, manual calibration should be done together with optimization. Since ET_a is overestimated to begin with, the optimization of the SIMHYD model compensates and matches daily observed streamflow by adding baseflow. Overall, 85-97% of total simulated runoff is accounted for by the baseflow. The remaining runoff is composed of impervious runoff, infiltration excess runoff and interflow runoff.

On account of a large baseflow on the simulated discharge, the most sensitive parameters in this case include the infiltration coefficient, infiltration shape, and pervious fraction. A range of infiltration coefficients which significantly influences the model performance (i.e. *E*) is from 0 to 20. The higher the infiltration coefficient, the better the performance of the model. A value of infiltration shape above 1.0 also moderately influences the performance of the SIMHYD model. An increasing value of the infiltration shape is associated with a decreasing *E*. In contrast, a higher *E* is obtained with respect to an increasing pervious fraction. Large changes in *E* are associated with the pervious fraction varying from 0 to 0.6.

As seen in Table 6.6(a), the underestimation in the mean of daily streamflow (i.e. negative D_v) during the calibration period is obtained at Station P21, P24A, 061302 and 061501, whereas overestimation (i.e. positive D_v) is observed at other two stations (i.e. P4A and P71). D_v varies from -5% to 156%. The large D_v percentage at Station P4A and P71 is associated with the high variability of observed streamflow – i.e. the CV are estimated at 1.52 and 1.78 respectively (see Table 6.5) – in the large drainage areas (i.e. 1,902 and 1,771 km^2 respectively). Among the six streamflow stations, the best performance of the SIMHYD model is obtained at Station 061302 with the maximum *r* and *E* estimated at 0.78 and 0.60 respectively, and with a minimum NRMSE computed at 0.76. Station 061302 has a drainage area of 1,946 km^2 with 97.22% of the basin area covered by forests. On the other hand, the worst performance is observed at Station P4A which has a drainage area of 1,902 km^2 with forests covering around 89.07% of the basin area. Note that the homogeneity of land use (i.e. basin characteristics) may play a role on the model performance (Chiew and Siriwardena, 2005). Therefore, the SIMHYD model has difficulty in calibrating the high varied streamflow in large basins because of its simple mathematical computation and an assumption of homogeneity throughout the basin, which cannot well represent water balance of a large catchment. However, the model efficiency is not linearly related to the percentage of land use at all streamflow stations. The hydrographs of daily simulated streamflow from the SIMHYD model during calibration and overlaid by observations at six gauging stations are presented in Appendix G1 to G6.

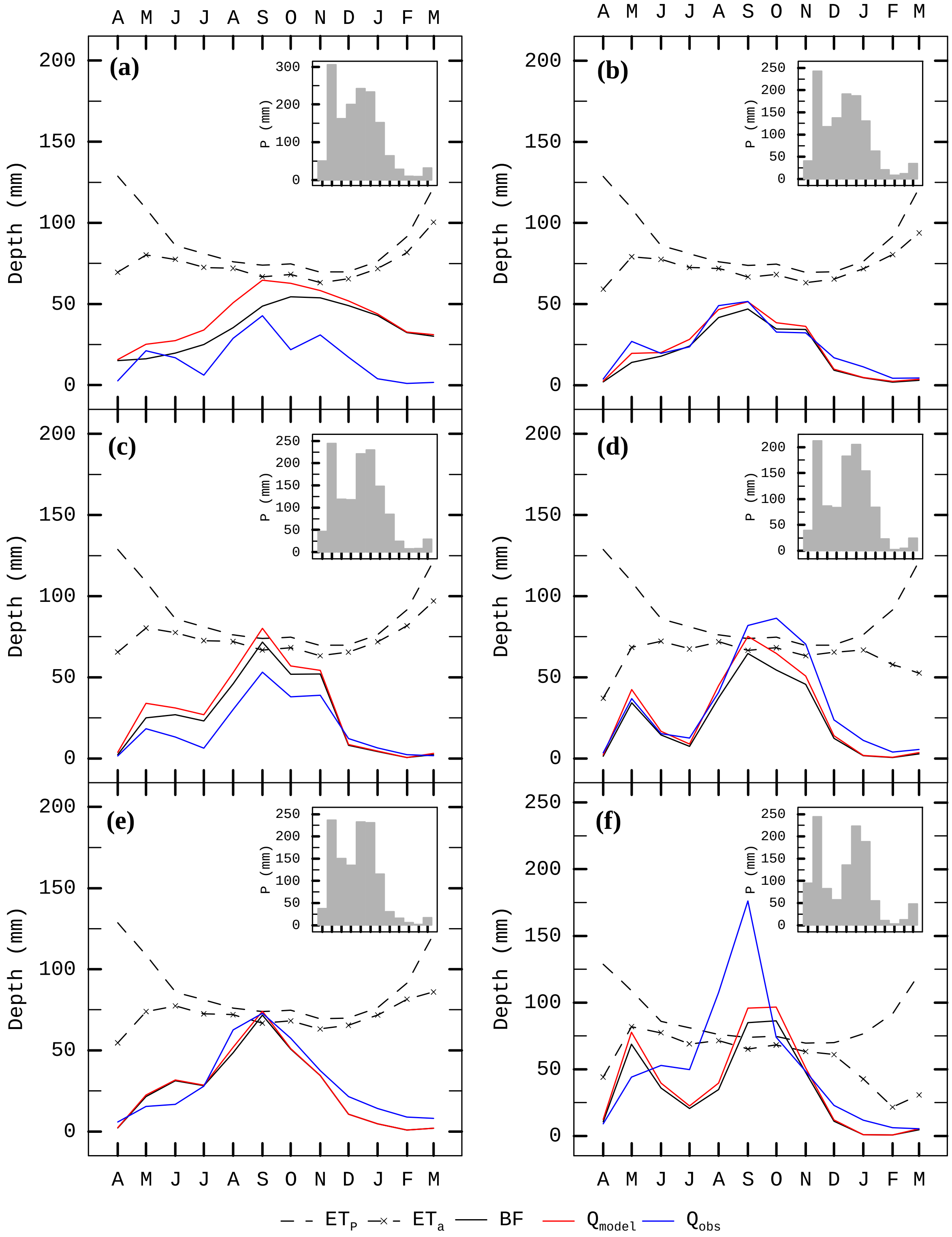


Figure 6.7: Annual variability of simulated (Q_{model}) and observed (Q_{obs}) streamflow from the calibration of the SIMHYD model at six gauging stations: (a) P4A; (b) P21; (c) P71; (d) P24A; (e) 061302; and (f) 061501.

Table 6.6: Performance of the SIMHYD Model

Station	(a) Calibration				(b) Validation			
	D_v (%)	r	NRMSE	E	D_v (%)	r	NRMSE	E
P4A	156.15	0.48	2.25	-1.19	127.59	0.57	2.08	-0.25
P21	-4.67	0.66	1.03	0.42	-12.89	0.65	1.25	0.35
P71	60.69	0.68	1.65	0.14	51.20	0.74	1.98	0.15
P24A	-16.97	0.75	1.26	0.55	-53.26	0.56	1.73	0.24
061302	-9.79	0.78	0.76	0.60	-32.36	0.72	1.21	0.45
061501	-25.37	0.59	1.31	0.32	-20.46	0.65	1.24	0.38

Calibration: 1st April 1999 to 31st March 2003; validation: 1st April 2003 to 31st March 2007.

During validation (i.e. from April 2003 to March 2007), the SIMHYD model faces difficulty in capturing the average monthly streamflow at all the six gauging stations. The underestimation of streamflow in the wet season, especially from August to October, is observed at Station P24A, 061302 and 061501, and vice versa for Station P4A, P21 and P71. Moreover, the underestimation of streamflow in the dry season (i.e. during December to March) is consistent in all stations with the exception of Station P4A. Overall, underestimation and overestimation are observed at the same stations as those where underestimation and overestimation are obtained while calibrating (Table 6.6(b)). D_v ranges from -53% to 128% and r varies from 0.56 to 0.74. A satisfactory agreement between simulations and observations (r>0.7) is found at Station P71 and 061302. The minimum and maximum NRMSE are estimated to be 1.21 and 2.08 respectively. The larger residuals correspond to higher magnitudes of observed streamflow. Furthermore, E is used to evaluate the model performance with more focus on the model capability for capturing outliers rather than the values close to average streamflow (Krause et al., 2005). The E at all stations is not satisfactory because the SIMHYD model cannot preserve the peaks of hydrographs (i.e. the outliers). However, a maximum E value of 0.45 is obtained at Station 061302. Hence, the model efficiency in validation is not significantly better than that of calibration.

6.4.2 Calibration and validation of the HEC-HMS model

The parameters of the HEC-HMS model at 12 streamflow stations were also calibrated from April 1999 to March 2003 with the primary objective of preserving the variability of average monthly streamflow and the secondary objective of attaining no bias (i.e. $D_v=0$) when calculating the mean of daily streamflow. During the calibration period, the HEC-HMS model can fairly capture the variability of average monthly streamflow at all stations (Figure 6.8). The calibration has difficulty fitting the monthly streamflow during the pre-monsoon season (i.e. MJJ), which is influenced by several parameters related to ET_a including the canopy maximum storage (mm), surface maximum storage (mm), infiltration constant rate (mm h^{-1}), and imperviousness (%). The increasing canopy and surface maximum storage, which provide more water for ET_a, influence decreasing simulated MJJ streamflow. With a higher infiltration constant rate, the modeled MJJ streamflow will also decrease due to soil infiltration loss. However, the increasing simulated MJJ streamflow is due to a greater imperviousness. For the wet season streamflow (i.e. ASON), the three most sensitive parameters are surface maximum storage, infiltration constant rate and imperviousness. Furthermore, the dominant parameters of streamflow in the dry season (i.e. from December to April) are associated with baseflow calculation (e.g. recession constant and ratio to peak).

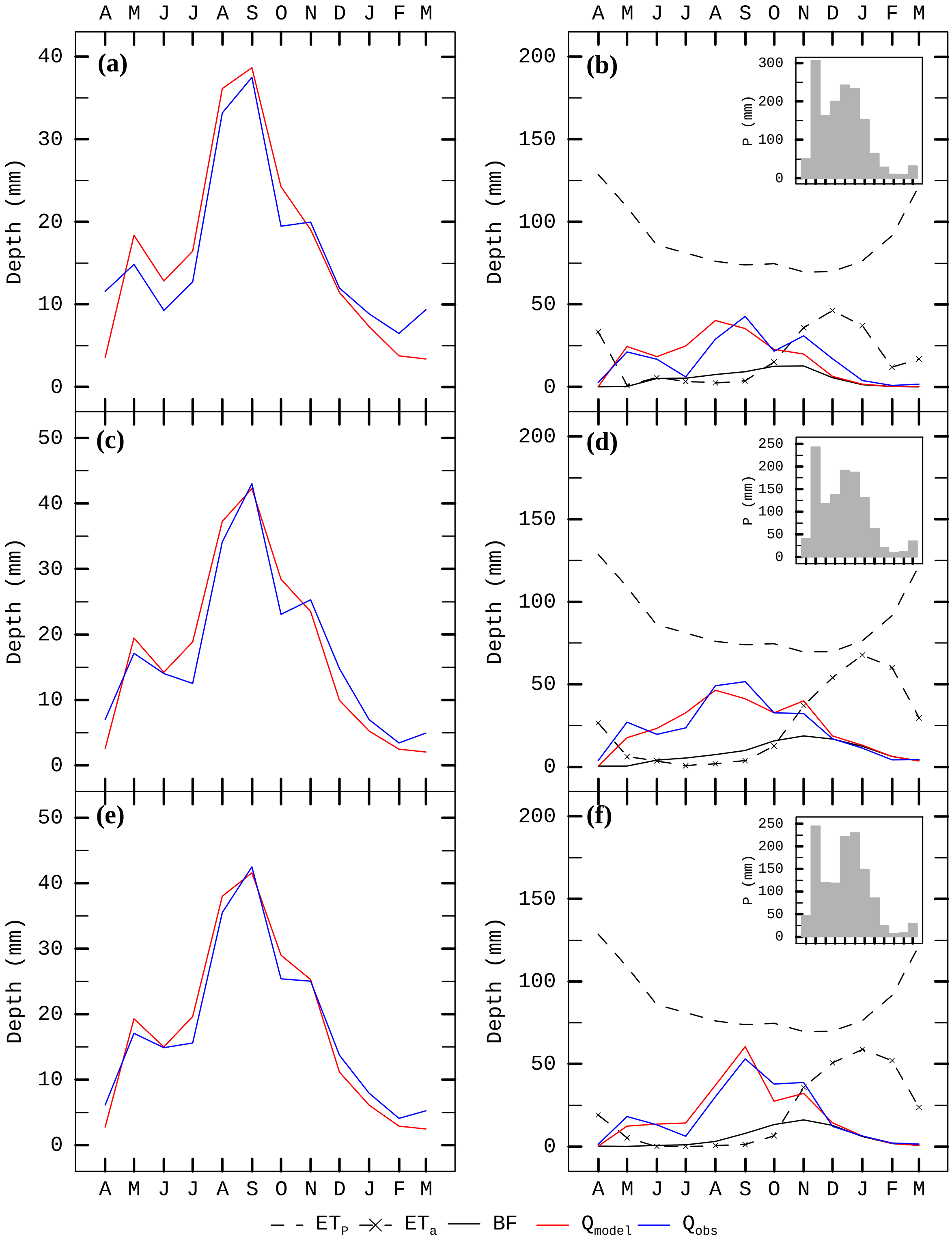


Figure 6.8: Annual variability of simulated (Q_{model}) and observed (Q_{obs}) streamflow from the calibration of the HEC-HMS model at 12 gauging stations: (a) P75; (b) P4A; (c) P67; (d) P21; (e) P1; (f) P71; (g) P24A; (h) P73; (i) 061302; (j) P14; (k) 061501; and (l) P12B.

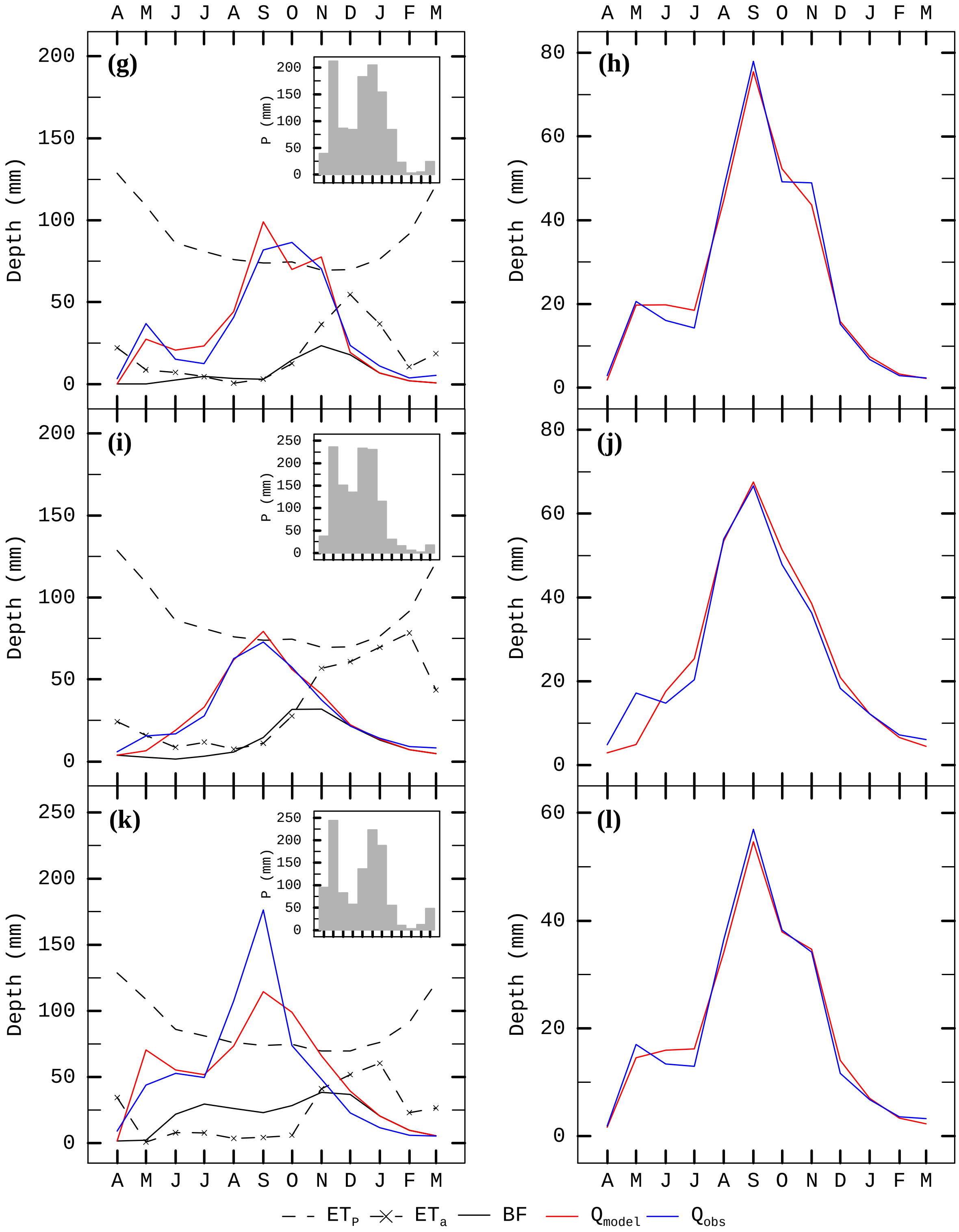


Figure 6.8 (cont).

As seen in Figure 6.8(k), the poor performance of the HEC-HMS model is found at Station 061501 as the model cannot capture September streamflow. The high peak of observed streamflow in September corresponds to positive anomalies in September 2000 and 2002 (see Figure 6.9 and Appendix H1 to H12 for more plots of the monthly anomalies at other stations). Note that the monthly anomalies of rainfall and streamflow are estimated with respect to the 1999-2007 average monthly values.

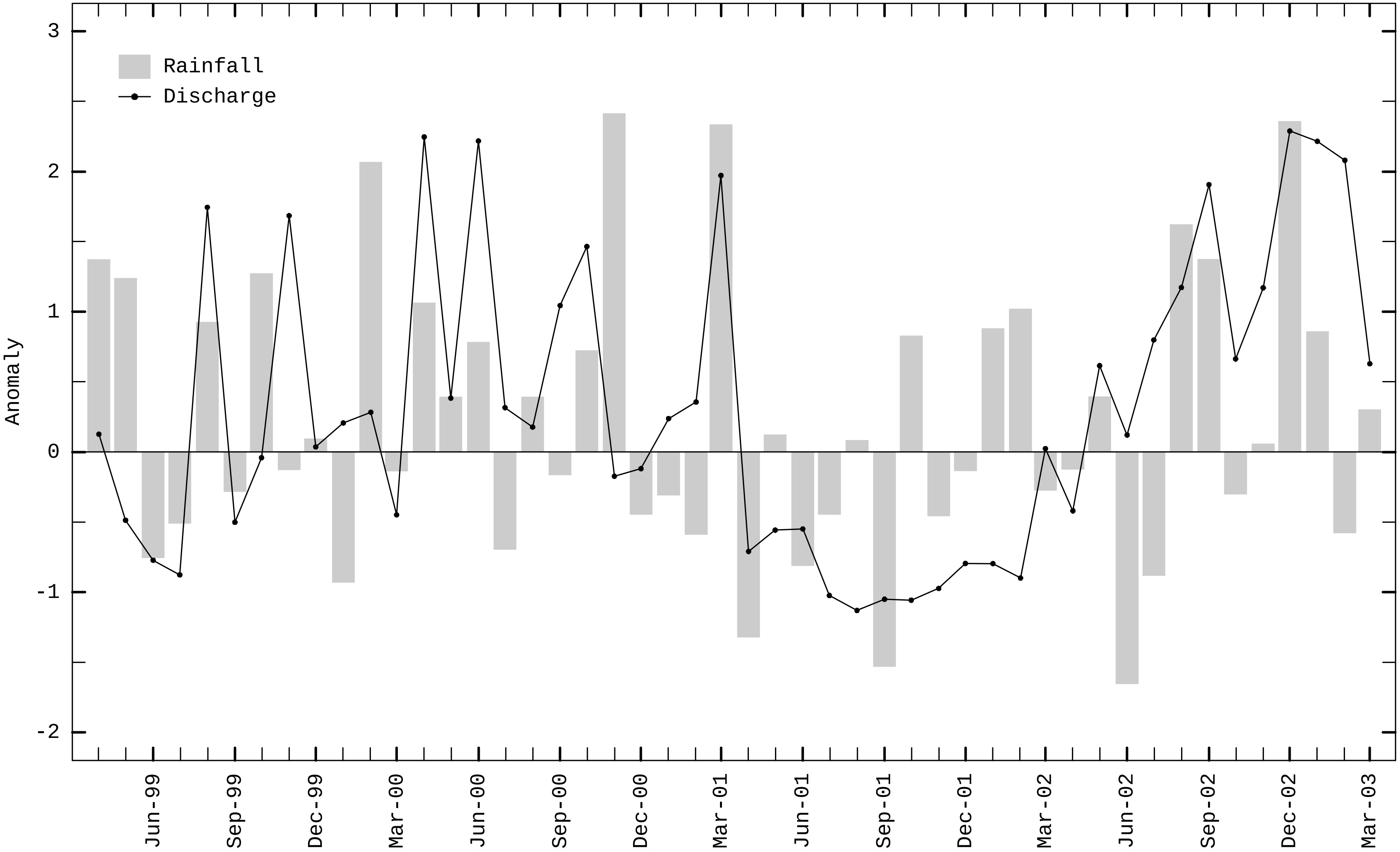


Figure 6.9: Monthly anomalies of observed rainfall and streamflow at Station 061502 from April 1999 to March 2003. The monthly anomalies are estimated with respect to the 1999-2007 averaged values of each month.

At Station 061501, the HEC-HMS model hardly captures the extremely high daily streamflow of September 2000 and 2002 (i.e. positive anomalies) as shown in the hydrographs of daily discharge (Figure 6.10). Other indexes indicating poor performance at this station (r=0.62, NRMSE=1.25, and E=0.39) have also been obtained (Table 6.7(a)). While calibration at Station 061501 can be done to capture the September peak with D_v=0, worse performances for other indexes suggests that this approach is not efficient for this particular station.

In contrast, at other streamflow stations, the overall calibration results from the HEC-HMS model are satisfactory with correlations (r) between daily simulated and observed streamflow varying from 0.67 to 0.93. The minimum NRMSE is calculated to be 0.48 at Station P1 that measures daily streamflow of the Ping River which has a drainage area of 6,355 km^2. In terms of the Nash-Sutcliffe efficient index (E), a satisfactory calibration (E>0.5) is found at seven stations (P67, P1, P71, P24A, P73, P14 and P12B) where most of the stations measure streamflow combined with the main river. The maximum E is estimated at 0.85 at Station P73 which receives the runoff from a drainage area of 13,353 km^2. From the above results, it can be inferred that the HEC-HMS model performs well when capturing the variability of daily streamflow in small or large catchments. However, the model has difficulties in preserving the anomalies of daily streamflow, especially in the wet season. The hydrographs of daily simulated streamflow from the calibration of the HEC-HMS model along with daily observations are shown in Appendix I1 to I12.

In the validation process from April 2003 to March 2007 (Table 6.7(b)), the underestimation in mean daily streamflow (i.e. negative D_v) is observed at all stations except at Station P14 and 061501. D_v varies from -30% to 4%. In terms of r, worse performance is obtained when validating than when calibrating in all stations except P4A, P21, 061302 and 061501. However, excluding Station P24A and 061501, the overall results in validation are satisfactory (r>0.7). NRMSE ranges from 0.70 at Station P12B to 1.66 at P24A. The maximum E for validation is estimated at 0.79 at Station P12B, which measures the inflow of the Bhumipol Dam. Satisfactory efficiency in both calibration and validation at this station (P12B) ensures the possibility of implementing the model results in reservoir management and planning. It can be concluded that considering r, NRMSE and E, the efficiency of the HEC-HMS model in less in validation than during calibration.

6.4.3 Performance in capturing low and high streamflows

From Figure 6.11(a), we can see that during the calibration of the SIMHYD model, overestimation in the mean of low streamflow (i.e. below-mean and below-median streamflow) is addressed in all stations excluding P24A and 061302. In contrast, the underestimation in the mean of high streamflow (i.e. above-mean and above-median streamflow) is observed at all gauging stations with the exception of P4A and P71. The overestimation (underestimation) in the mean of low (high) streamflow is caused by a large overestimation of ET_a and baseflow. Overall, the SIMHYD model better captures the variability of high discharge rather than low discharge, and this behavior is consistent with all efficiency indexes (Figure 6.11(a)-(c)). The model performance shows that parameter optimization of the SIMHYD model focuses more on matching peak discharges during the wet season more than during the low discharges in the dry season.

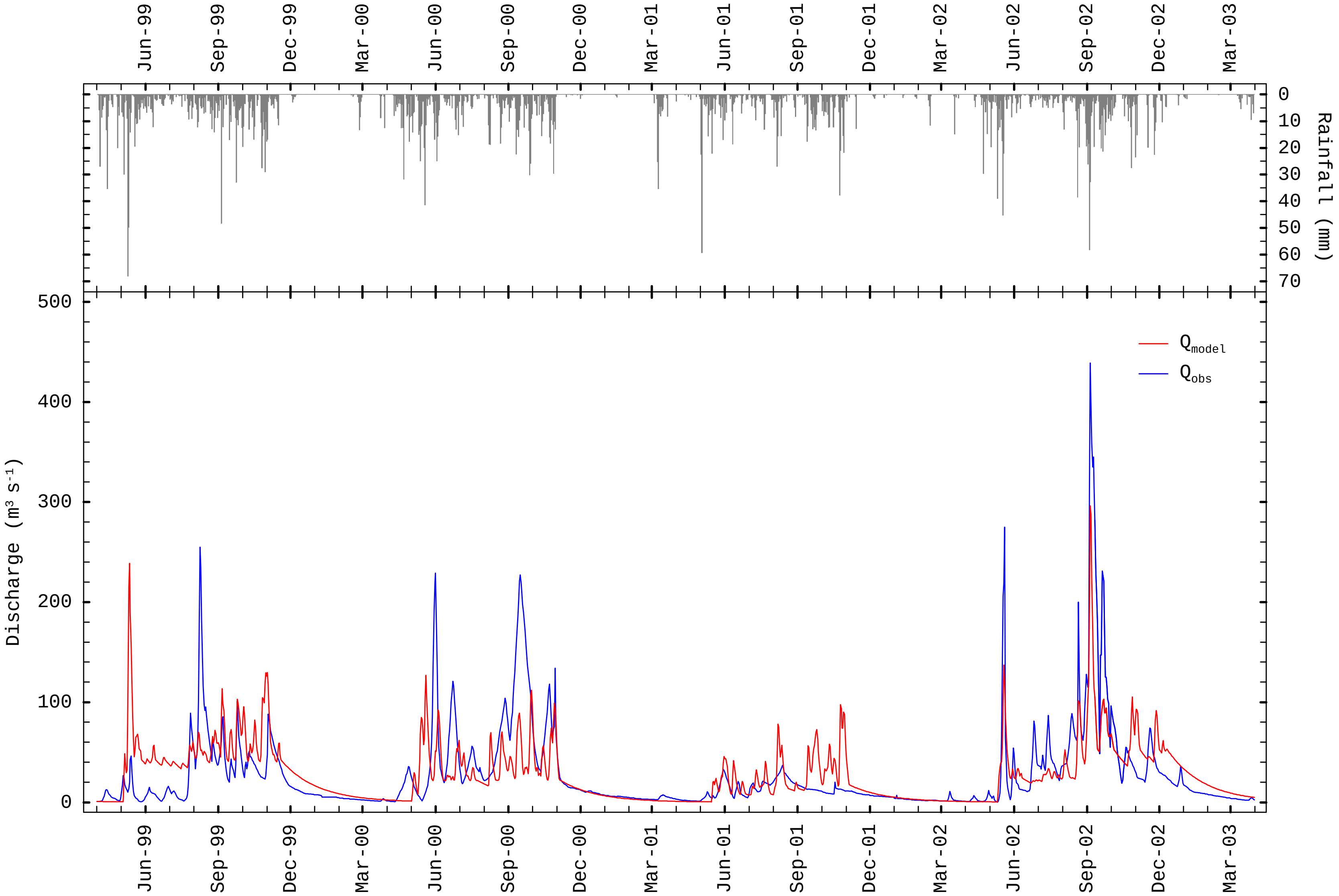


Figure 6.10: Hydrographs of daily simulated (Q_{model}) and observed (Q_{obs}) discharge from the calibration of the HEC-HMS model.

Table 6.7: Performance of the HEC-HMS Model

Station	(a) Calibration				(b) Validation			
	D_v (%)	*r*	NRMSE	*E*	D_v (%)	*r*	NRMSE	*E*
P75	0.003	0.85	0.76	0.47	-23.744	0.78	0.95	0.33
P4A	-0.001	0.70	1.12	0.46	-15.062	0.77	1.22	0.57
P67	0.002	0.90	0.58	0.77	-30.276	0.86	0.80	0.69
P21	0.001	0.67	1.10	0.34	-5.792	0.73	1.18	0.42
P1	0.004	0.92	0.48	0.80	-25.904	0.88	0.71	0.73
P71	0.000	0.77	1.22	0.53	-3.728	0.70	1.58	0.46
P24A	0.001	0.78	1.25	0.56	-17.799	0.56	1.66	0.29
P73	0.004	0.93	0.55	0.85	-23.649	0.88	0.74	0.74
061302	-0.002	0.71	0.92	0.42	-14.991	0.74	1.11	0.53
P14	-0.004	0.82	0.66	0.65	3.488	0.70	1.26	0.49
061501	-0.001	0.62	1.25	0.39	3.652	0.65	1.29	0.32
P12B	0.003	0.90	0.58	0.81	-11.556	0.90	0.70	0.79

Calibration: 1st April 1999 to 31st March 2003; validation: 1st April 2003 to 31st March 2007.

It is also observed that based on NRMSE (Figure 6.11(b)), at Station P21 and 061302, the residuals corresponding to low and high simulated streamflows are slightly different. It is also noted that when comparing all the stations in the Ping River Basin, the 1999-2003 daily observed streamflow in these stations varies slightly – as in, CV is 1.36 and 1.21 respectively (see also Table 6.5). Therefore, this can confirm that the SIMHYD model hardly simulates daily streamflow with large variability due to its simpler mathematical computation.

On the other hand, in the calibration of the HEC-HMS model, the overestimation (underestimation) in the mean of low (high) streamflow is consistent in all the stations except at Station P1 (Figure 6.12(a)). The small errors noted there are associated with the simulations of high streamflow (Figure 6.12(b)). However, at several stations (P75, P67, P21, P1, P73, 061302, P14 and P12B), the errors in low and high simulated streamflow are not significantly different. Hence, the HEC-HMS model presents the same skill in the simulation of both low and high streamflow. Based on *E*, we can infer that the HEC-HMS model shows better performance in capturing high streamflow more than low streamflow (Figure 6.12(c)). A satisfactory value of *E* (*E*>0.5) is obtained at Station P67, P1, P73 and P12B. Overall, the HEC-HMS model performs well when capturing high discharges in the wet season (i.e. from August to November) but fails to efficiently capture the variability in low streamflow. This poor performance is related to the estimation of total loss and baseflow that influences daily simulated runoff. Improving the model efficiency needs obtaining more data of the soil profile, soil property, land use coverage and imperviousness.

6.5 Comparison of models

The SIMHYD model is a lumped rainfall-runoff model. Its algorithm is a simple computation of water balance in a basin. The model uses nine parameters to estimate losses, baseflow and runoff. The nine parameters can be optimized by a given optimizer with specified objective functions. The SIMHYD model is easy to apply because of its small number of parameters and its available parameter optimizations.

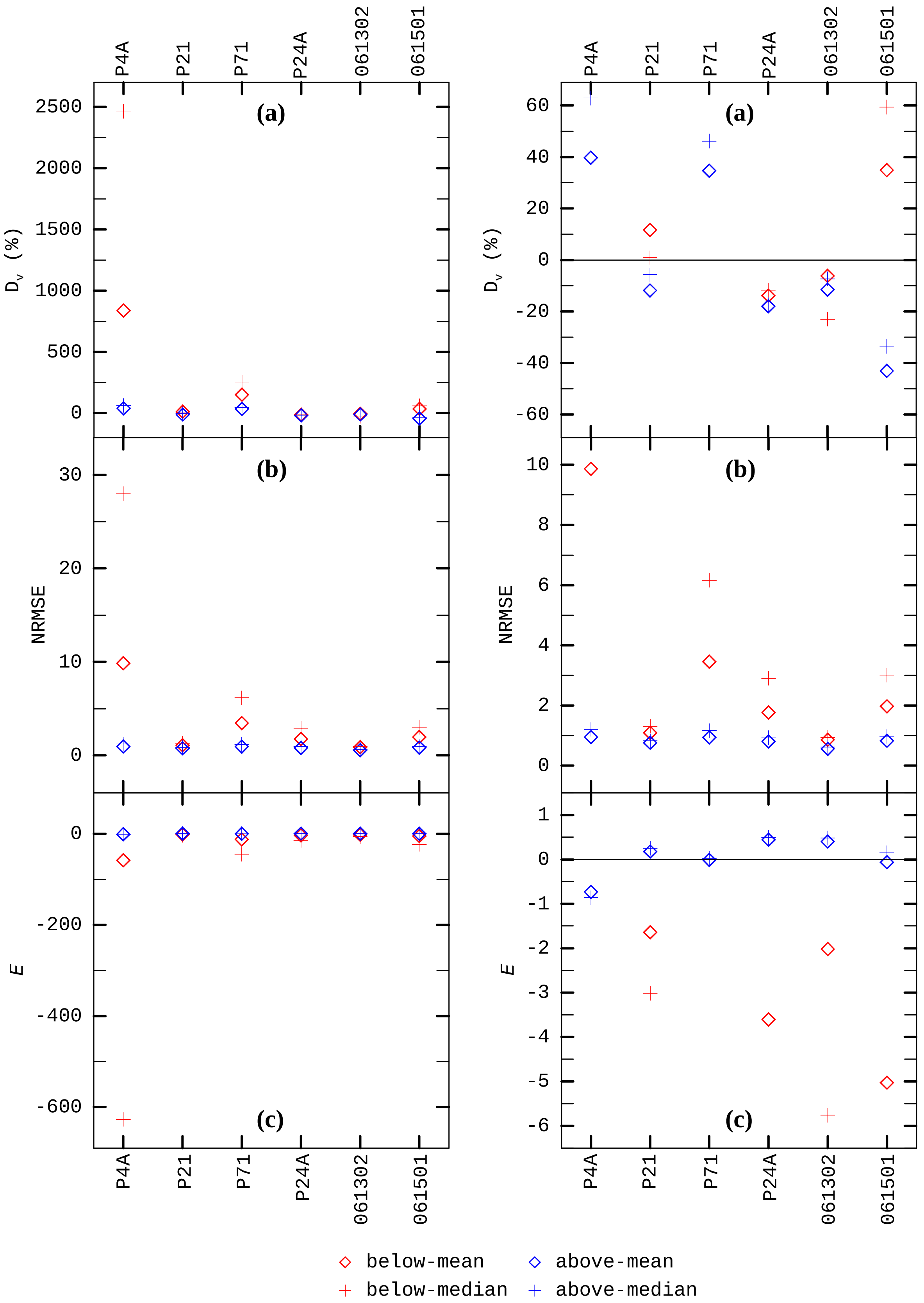


Figure 6.11: Efficiency indexes between below-mean (-median) and above-mean (-median) streamflow in the calibration of the SIMHYD model. The indexes are: (a) D_v; (b) NRMSE; and (c) *E*. The plots in the right panel are the zoomed-in plots of the ones on the left panel.

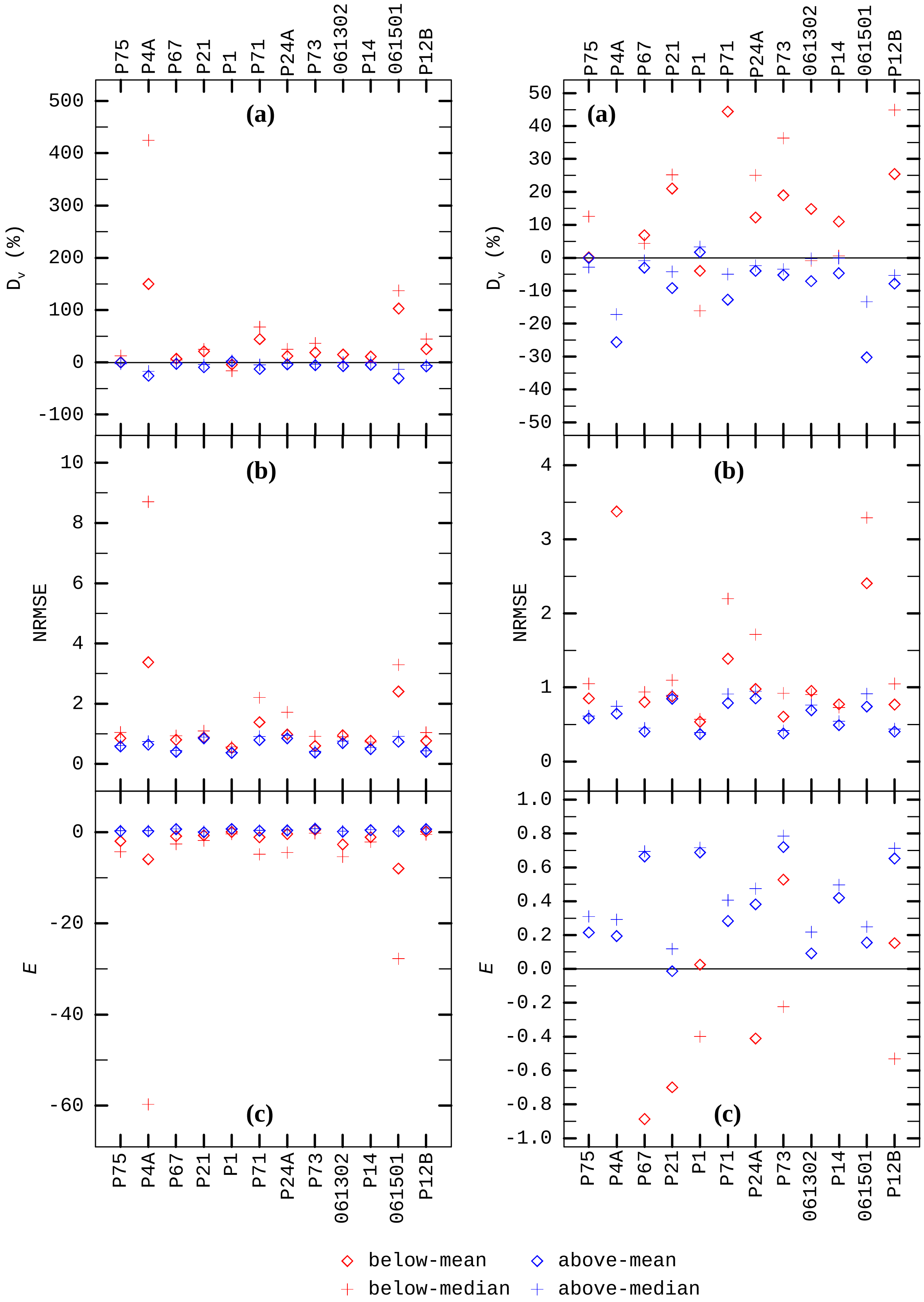


Figure 6.12: Same as Figure 6.11 but for the HEC-HMS model.

The HEC-HMS model is more complex than the SIMHYD model. The HEC-HMS model provides four components: loss, transform, baseflow and routing. Within each component, several methods of mathematical computations are used to estimate ET_a, infiltration, baseflow and runoff. The users can select a method based on the available data of model parameters. In this study, the comparison of the SIMHYD and HEC-HMS model performances has been done at six stations (P4A, P21, P71, P24A, 061302 and 061501). Depending on their efficiency, one model has been selected.

Figure 6.13(a), shows that as expected, the HEC-HMS model performs better when preserving the average of daily streamflow at all stations compared to the SIMHYD model for the ideal parameter of D_v=0 in model calibration. In terms of agreement between simulated and observed discharge (i.e. *r*) (Figure 6.13(b)), HEC-HMS performs better in calibration (validation) except in Station 061302 (P71). However, in this station (061302), the value of *r* obtained from both models is slightly different: for calibration (validation), *r*=0.78 (0.74) from SIMHYD, and *r*=0.71 (0.70) from HEC-HMS. As for calibration and validation, the significant difference in *r* between both models is observed at Station P4A which presents a better performance of the HEC-HMS model than SIMHYD. So, overall, the HEC-HMS model better captures the variability of daily streamflow as compared to the SIMHYD model.

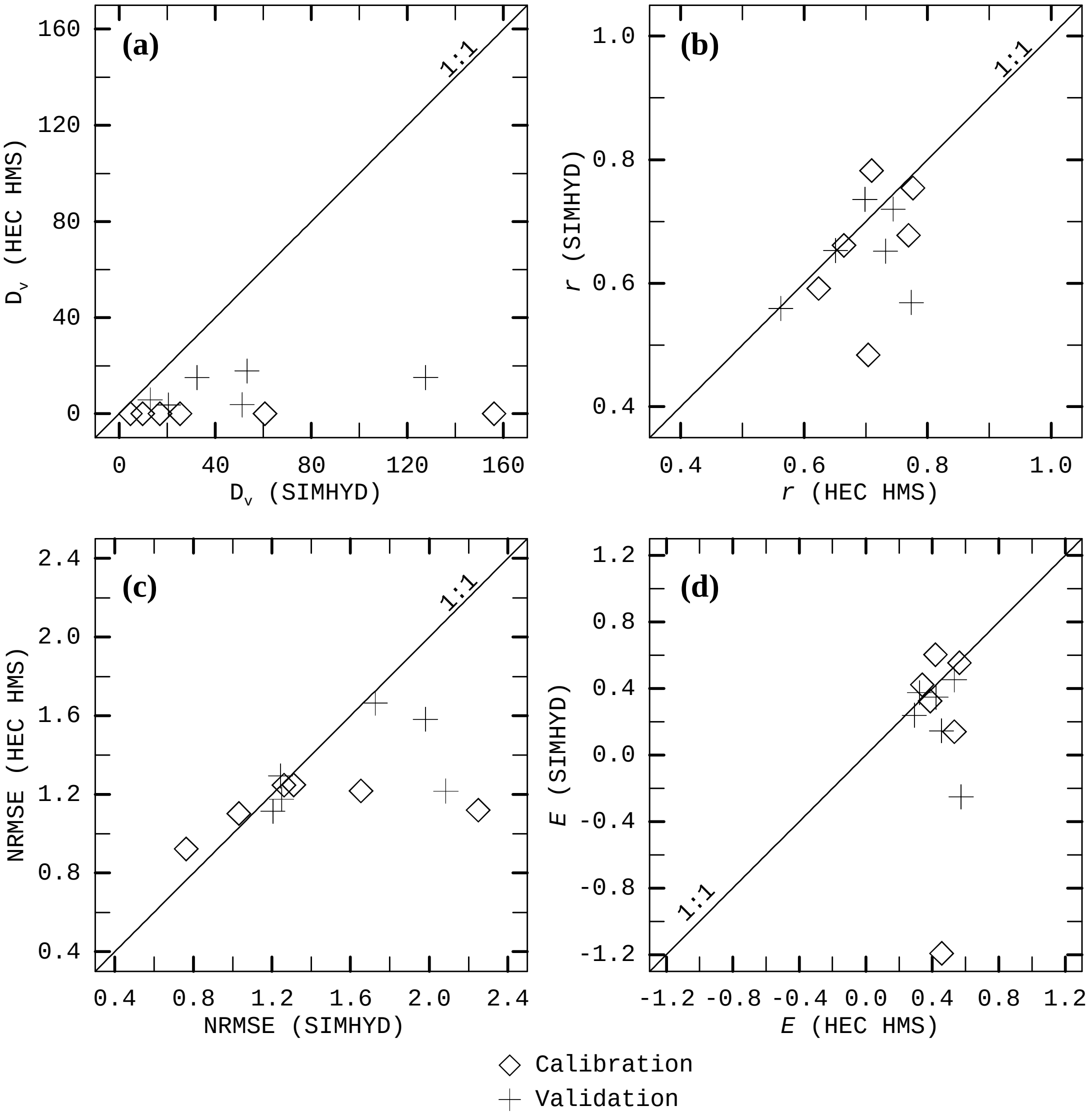


Figure 6.13: Comparison of performance between the SIMHYD and HEC-HMS models. The indexes of comparison are: (a) absolute D_v; (b) *r*; (c) NRMSE; and (d) *E*.

The residuals of simulation (i.e. NRMSE) (Figure 6.13(c)) from calibration and validation also indicate a greater efficiency of the HEC-HMS model than the SIMHYD model, in particular at Station P4A and P71. Additionally, the value of *E* from both calibration and validation (Figure 6.13(d)) is consistent with NRMSE. It is also important to note that the SIMHYD model has difficulties in streamflow simulation at Station P4A and P71 because these stations cover very large drainage areas and the heterogeneity of basin characteristics play a major role in model efficiency. At these stations, the daily observed streamflow during calibration and validation shows high variability (see also Table 6.5). Therefore, although the HEC-HMS model requires several parameters to drive its components, it can better capture the variability of daily observed streamflow than the SIMHYD model. Among several mathematical computation methods in the HEC-HMS model, the users can select a method of each component which well represents water balance in a specific basin and gives greater accuracy in the estimation of ET_a, infiltration, baseflow and runoff.

The performance corresponding to the low (i.e. below-mean) and high (i.e. above-mean) streamflow has also been evaluated, as shown in (Figure 6.14). The efficiency indexes, including D_v, NRMSE and *E* prove that the HEC-HMS model performs better when capturing the variability of low streamflow, especially in Station P4A and P71. A severe problem in the SIMHYD model found in Station P4A is in the simulation of low flow. The calibration of D_v=838%, NRMSE=9.86 and *E*=-58.31 is highly unsatisfactory. For validation, the comparison between both models shows results consistent with calibration. It can be argued that the separated and complex computation components (such as loss and baseflow) in the HEC-HMS model are more flexible and more efficient in fitting low streamflow of daily observations. Even for high streamflow, a smaller D_v from calibration and validation at all stations is obtained with HEC-HMS. However, at some stations, the NRMSE and *E* are inconsistent with D_v because D_v is a rough efficiency index used to evaluate a model performance only in capturing mean daily streamflow over a specific period, whereas NRMSE and *E* evaluate the model at each point of daily data recording. For calibration, the HEC-HMS model indicates higher efficiency at three gauging stations: P4A, P71 and 061501. As expected, since the SIMHYD model performs most inefficiently at Station P4A amongst the three stations, HEC-HMS is a better option for P4A. As for validation, although NRMSE is not significantly different between the two models, the HEC-HMS model performs better when capturing the variability of high streamflow than the SIMHYD model at four stations. *E* also confirms the better performance of the HEC-HMS model at these stations.

Hence, the SIMHYD model is more difficult to deal with heterogeneity of basin characteristics than the HEC-HMS model because of its simple mathematical computation of water balance. The SIMHYD model provides one method of each component (i.e. initial loss, infiltration and baseflow) to calculate runoff, which cannot well represent water balance of a large and heterogeneous basin. In contrast, the HEC-HMS model has several methods of mathematical computations within each component. The users can select a method that well captures water balance of a basin. In particular, compared to the SIMHYD model, the HEC-HMS model computes initial loss with a separation of the canopy interception loss and the surface interception loss, which better represents water balance of heterogeneous basin.

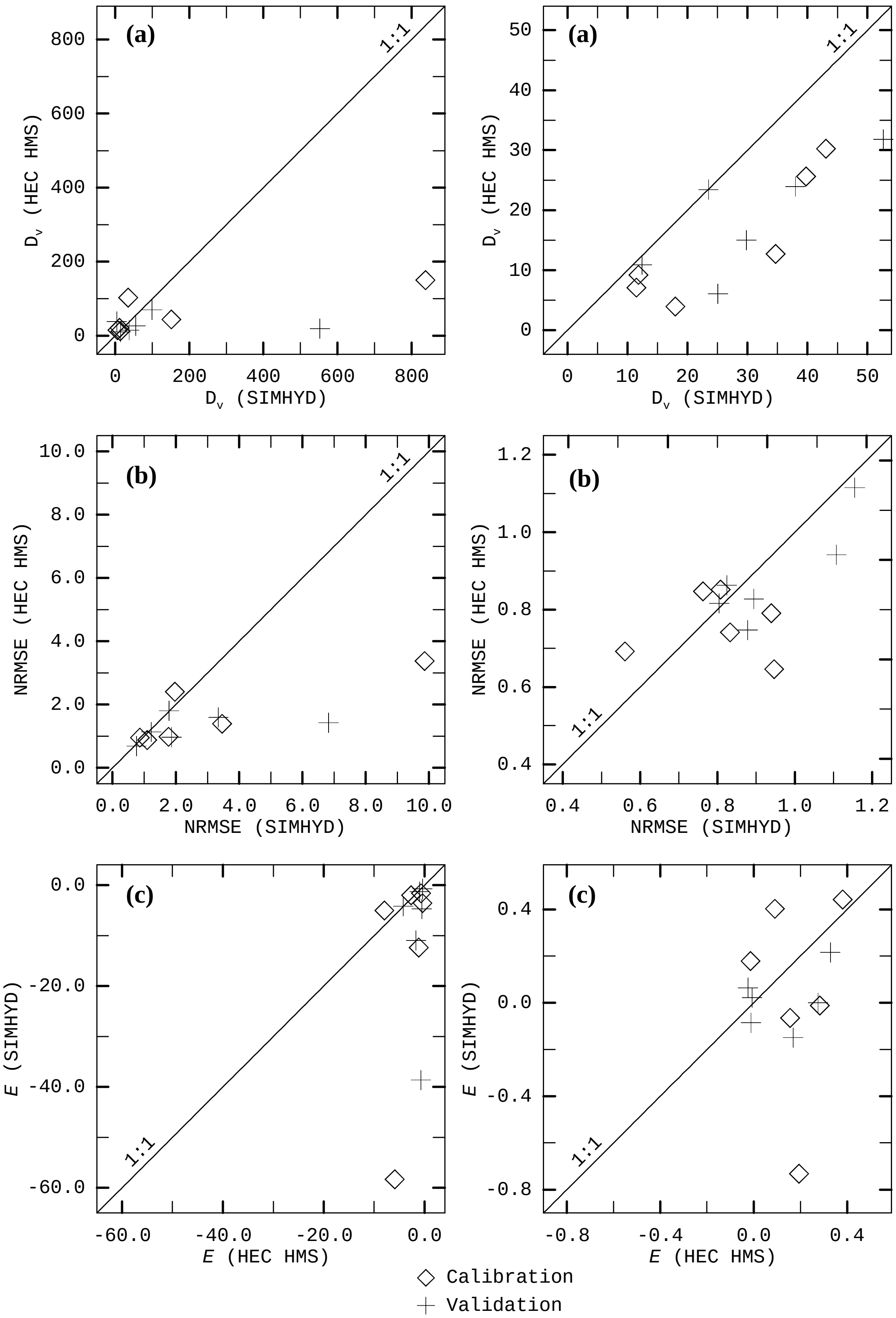


Figure 6.14: Same as Figure 6.13 but for the below-mean streamflow (left) and above-mean streamflow (right): (a) absolute D_v; (b) NRMSE; and (c) *E*.

In terms of model dependency, Figure 6.7 and 6.8 show the performance of the SIMHYD and HEC-HMS model in preserving the variability of monthly streamflow. The performance of both models is slightly different. However, the HEC-HMS model performs better in capturing daily streamflow, in particular low flow and high flow, than the SIMHYD model (Figure 6.13 and 6.14). Thus, the model dependency cannot be observed in monthly streamflow simulation, but it is found in the simulation of finer time scale. Consequently, the HEC-HMS model is selected to simulate daily streamflow to determine the effects of future changing climate.

6.6 The performance of the HEC-HMS model in capturing the frequency-duration-curve (FDC)

To confirm the performance of the HEC-HMS model as ideal for this part of the study, simulations using the 1999-2007 daily observed rainfall were done with the model parameters as obtained from calibration. At 12 gauging stations, the daily averages of simulated streamflow were calculated for a water year from April to March and plotted along with the daily average streamflow of observations (see Figure 6.15). The 1999-2007 simulation indicates that the HEC-HMS model can fairly well capture daily average streamflow at all stations. However, an underestimation of the daily average streamflow from December to April, or the dry season, is observed at Station P75 and P4A.

The frequency-duration-curve (FDC) presents the relationship between exceedence probabilities or frequencies of occurrence and magnitudes of discharges. The FDC for a specific period (like calendar years, seasons and months) can be computed using discharges at the required time scale (such as daily, monthly and seasonal). In this study, the FDC for a water year from April to March (i.e. a 365-day period) is established separately for 12 gauging stations using daily average discharges. First, the daily averages of observed discharges are computed. The 365 daily averages are then sorted in a descending order. Subsequently, a rank (*m*) is assigned for each value of the averaged discharge. Out of the 365 daily averages, the maximum discharge is associated with the first rank (*m*=1), and the minimum discharge corresponds to the last rank (*m*=365). The exceedence probability (*p*) of a discharge is then calculated using Equation 6.5.

$$p = \frac{m}{n+1} * 100\% \qquad \text{Equation 6.5}$$

where *p* is the exceedence probability of a given discharge, *m* is its rank, and *n* is the total number of discharges. The FDC provides an exceedence probability or frequency of occurrence according to a given magnitude of discharge. For example, Q_{90}=7.5 m^3 s^{-1} means that 90% of the observed discharges exceed or equal 7.5 m^3 s^{-1}. In terms of frequency also, it can be implied that discharges above or equal to 7.5 m^3 s^{-1} can be observed for 90% of the time.

The exceedence probability of a simulated discharge is also computed in this study. The FDC of 1999-2007 simulations are then plotted along with the FDC of observations in Figure 6.16. In general, the overestimation of discharge above Q_{10} can be observed, especially at Station P21, P1 and P24A. The HEC-HMS model presents a good performance in capturing the FDC of discharge below Q_{90}, excluding Station P75 where it gives an underestimation. At Station P24A (061501), the underestimation (overestimation) of discharge between Q_{10} and Q_{40} (Q_{30} and Q_{60}) has been found. As a result, it can be concluded that the HEC-HMS model can fairly capture the FDC of historical discharge at all stations.

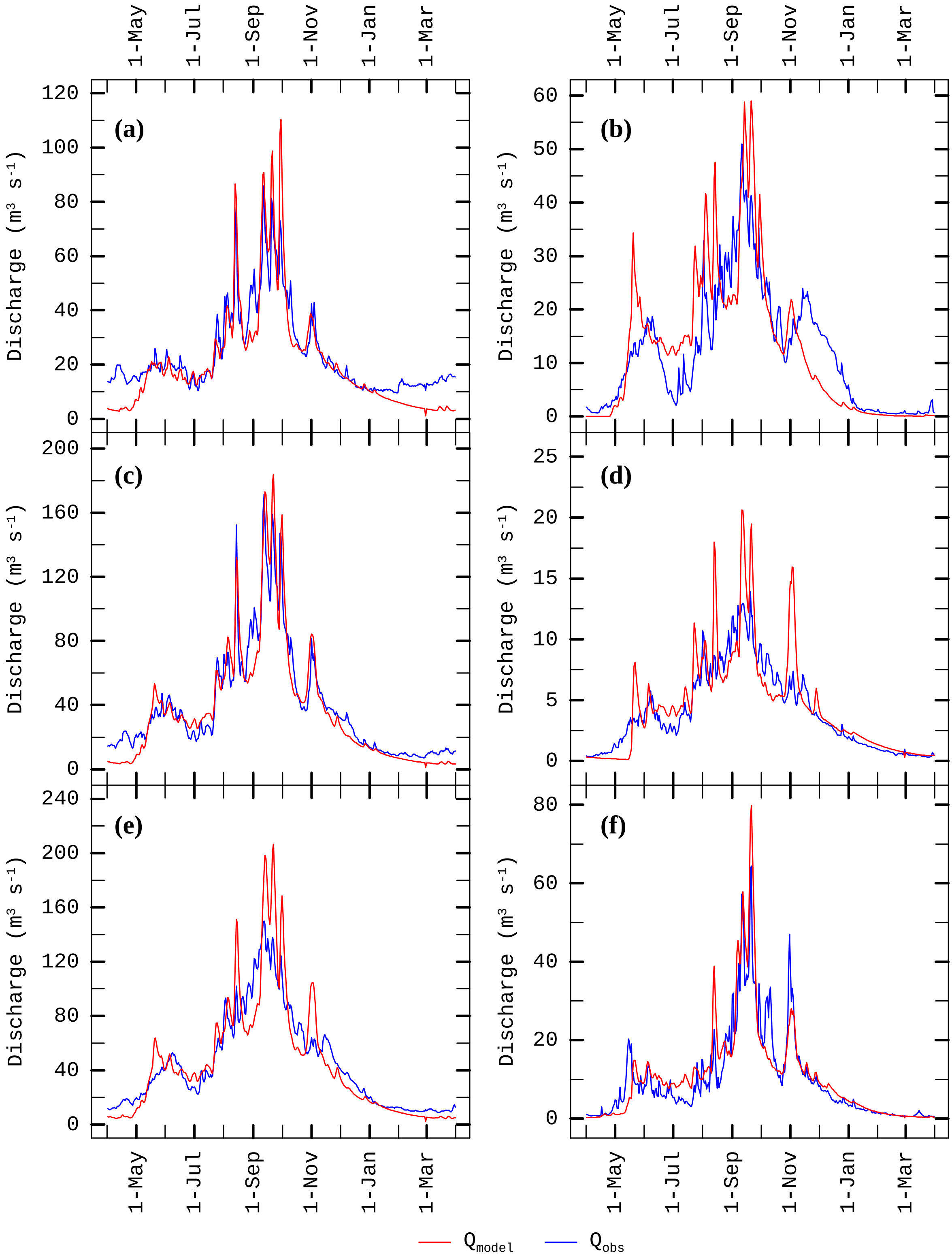


Figure 6.15: Daily averaged streamflow for a water year from 1st April to 31st March at the 12 gauging stations: (a) P75; (b) P4A; (c) P67; (d) P21; (e) P1; (f) P71; (g) P24A; (h) P73; (i) 061302; (j) P14; (k) 061501; and (l) P12B.

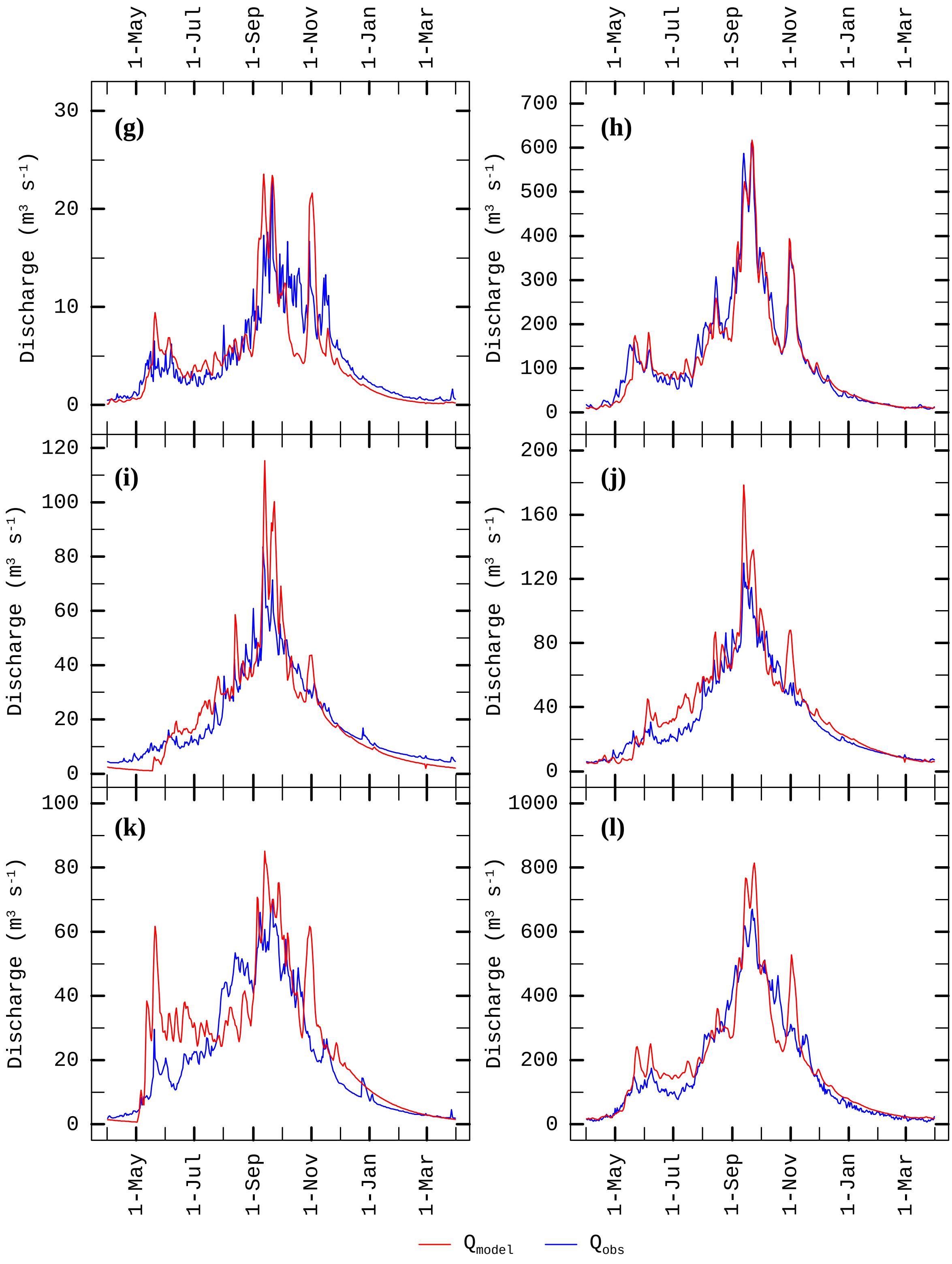


Figure 6.15 (cont).

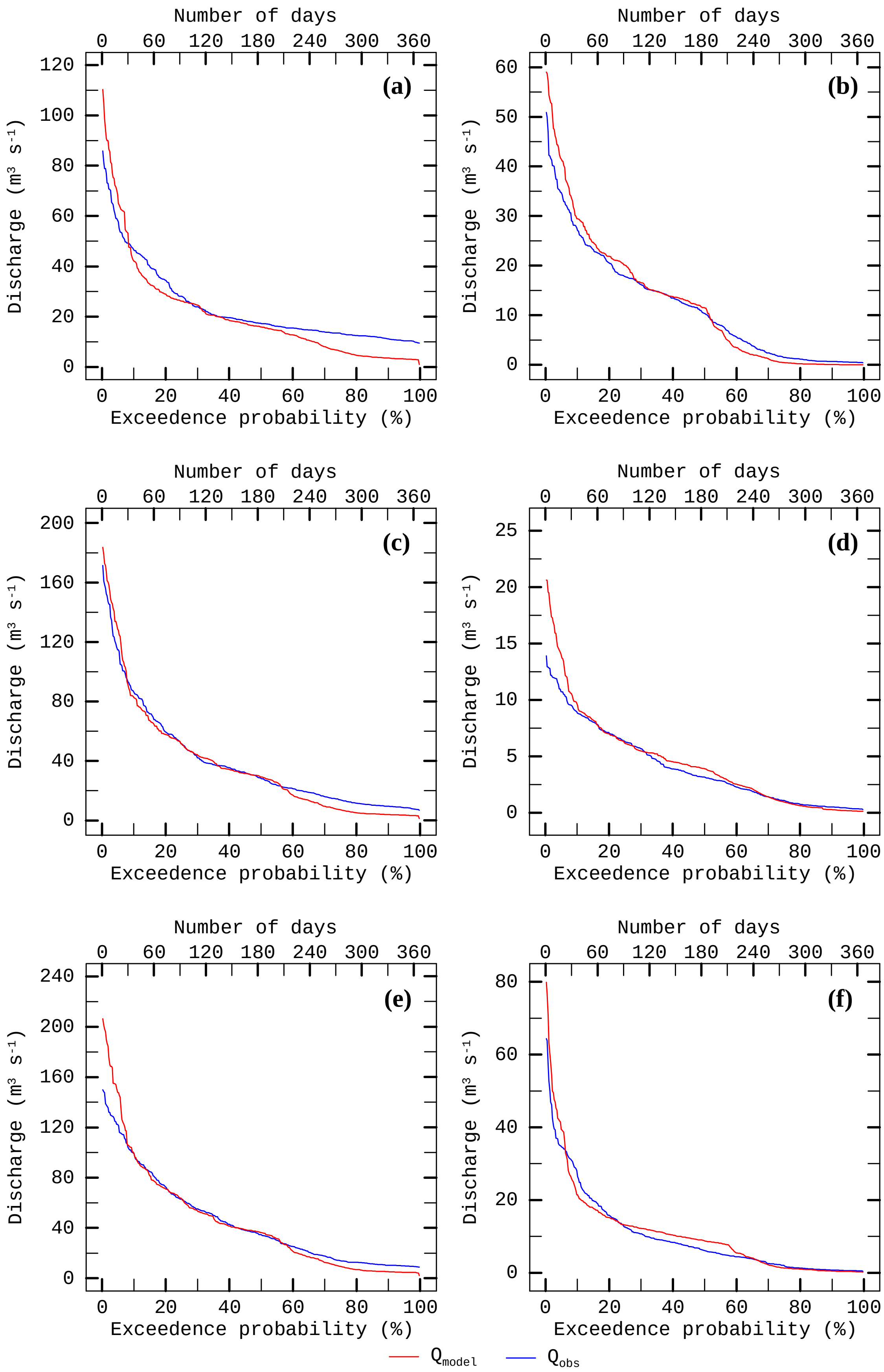


Figure 6.16: Frequency-duration-curve (FDC) for a water year of daily average discharges at the 12 gauging stations: (a) P75; (b) P4A; (c) P67; (d) P21; (e) P1; (f) P71; (g) P24A; (h) P73; (i) 061302; (j) P14; (k) 061501; and (l) P12B.

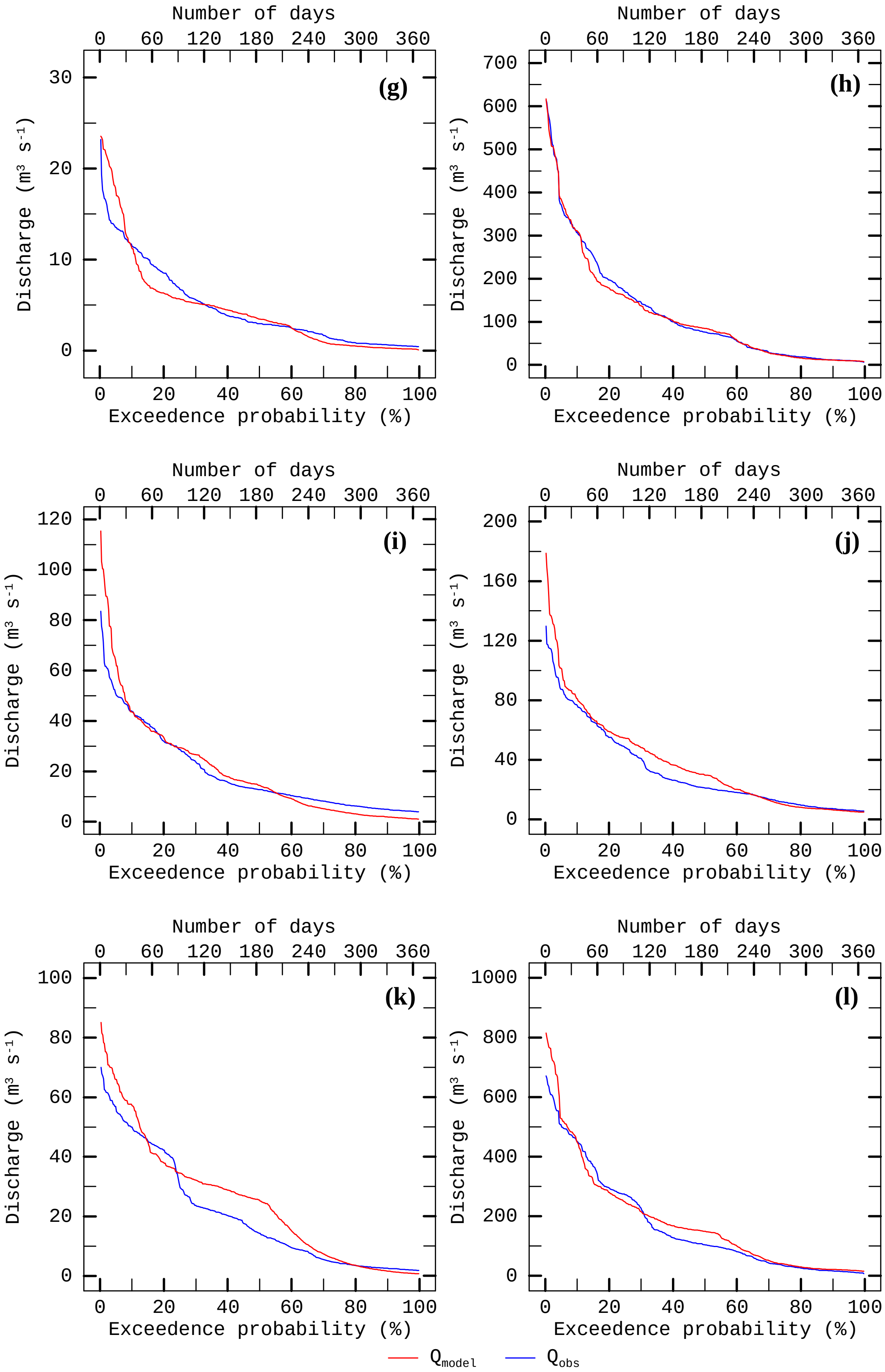


Figure 6.16 (cont).

6.7 Summary

The SIMHYD and HEC-HMS models were proposed in this study for the Ping River Basin streamflow observation. Both were calibrated from April 1999 to March 2003 and validated from April 2003 to March 2007 at six and 12 gauging stations respectively. The model performances were evaluated using four efficiency indexes, namely the deviation of volume (D_v), the correlation coefficient (r), the normalized root mean squared error (NRMSE), and the Nash-Sutcliffe efficient index (E). It can be observed that the lack of homogeneity of basin characteristics may cause the efficiency of the SIMHYD model to go down at those stations that cover large drainage areas with high variability of daily streamflow because the SIMHYD model provides a simple mathematical computation of water balance. Each component in the SIMHYD model that has one method to compute runoff cannot well represent water balance of a complex basin. In contrast, the HEC-HMS model can capture well the variability of average monthly streamflow. However, the overestimation of low flow and underestimation of high flow can be observed in the simulations of both models. Furthermore, the performance of both models in capturing high streamflow is better than when capturing low streamflow.

To select one model, the performances of both models were compared. The SIMHYD and HEC-HMS models perform well in preserving the variability of monthly streamflow; however, overall, the HEC-HMS model can better capture the daily observed streamflow. This shows model dependency of the daily streamflow simulation. Moreover, the initial loss computation of the HEC-HMS model with a separation of the canopy and surface interception can better deal with the heterogeneity of basin characteristics than the SIMHYD model. Although the efficiency corresponding to capturing low streamflow is slightly different between both models, the low flow is also better captured by the HEC-HMS model, which provides the separate component of baseflow to simulate streamflow during the dry season. For high streamflow as well, the HEC-HMS model can better preserve the average streamflow as compared to the SIMHYD model. NRMSE and E show that HEC-HMS has greater efficiency for high flow simulation, but this is not consistent at all stations. Notwithstanding minor dips in performance, the HEC-HMS model works well for this study, and hence it has been selected to simulate daily streamflow in the Ping River Basin using 2011-2100 rainfall ensembles. The simulation results are presented in the following chapter.

Chapter 7
Effects of Future Climate on Streamflow in the Ping River Basin

7.1 Introduction

As seen in the previous chapter, although the performance of the HEC-HMS and SIMHYD models in capturing monthly streamflow is slightly different (Figure 6.7 and 6.8), the HEC-HMS model performs better in preserving the variability of daily observed streamflow in particular low flow and high flow (Figure 6.14). This indicates a model dependency of the results for determining the effects of future climate on daily streamflow. Thus, the HEC-HMS model is best suited to simulate streamflow in the Ping River Basin using the 2011-2100 daily rainfall ensembles, which are obtained from the multisite daily rainfall generator presented in Chapter 5. Based on the assumption that the model parameters obtained from calibration are valid for basin characteristics in the present and future, rainfall-runoff simulation aims to determine the effects of future climate on streamflow under two scenarios of future climate, Scenario A2 and B2. Anomalous streamflow events like low and high flows are analyzed using defined thresholds. Defined thresholds are estimated from the frequency-duration-curve (FDC) of daily averages of the observed discharges. Moreover, seasonal anomalies like low flow in a wet season and high flow in a dry season are determined using thresholds from the FDC of average discharges in the wet and dry seasons respectively. The analysis of anomalous events caused by future climate can be utilized to plan adaptation strategies for agriculture and reservoir operations with the aim of reducing loss to agricultural production, reducing damages to infrastructure, cutting down the number of deaths and eliminating discontinuity of economic growth in the study basin, all of which are caused by anomalous streamflow events.

7.2 Data description

7.2.1 The 2011-2100 daily rainfall

The daily rainfall data used in this study is gained from a multisite daily rainfall generator which simulates daily rainfall at 50 selected stations based on two scenarios of future climate. The length of rainfall time series at each station is 90 years, from 1st January 2011 to 31st December 2100. Using the Thiessen method, the average daily rainfall is computed for each sub-basin in the Ping River Basin. The 2011-2100 daily rainfall is used as the input for the HEC-HMS model, whose parameter calibration and validation have been presented in Chapter 6.

7.2.2 Daily observed streamflow data

The parameters of the HEC-HMS model are calibrated at 12 gauging stations located in the Ping River and its tributaries. With no missing data, the length of daily time series at the relevant stations ranges from 8 to 31 years (Table 7.1). In this study, the daily observed streamflows are used to define thresholds for anomalous streamflow events like low flows and high flows. The daily observed data is also used in the analysis of anomalous events to determine the effects of future climate. The variability of observed streamflows is investigated and then compared to the simulated discharges under future climate scenarios A2 and B2.

7.3 Methodology

Daily streamflow is simulated by the HEC-HMS model using the 2011-2100 daily modeled rainfall under two scenarios of future climate, A2 and B2. The parameters of the HEC-HMS model at 12 gauging stations are calibrated during April 1999 to March 2003 and validated from April 2003 to March 2007, as shown in Chapter 6. Based on an assumption that the calibrated parameters are valid for both recent and future basin characteristics, the effects of future climate on daily streamflow, in particular for anomalous streamflow events, are addressed by comparing simulated and historical streamflow.

Table 7.1: List of the 12 Gauging Stations

Station	Sub-basin	Stream/ River	Length of data (year)	Start date	End date
P75	Upper Ping Part and Mae Ngad	Ping	8	1st April 1999	31st December 2007
P4A	Mae Taeng	Mae Taeng	26	1st April 1981	31st December 2007
P67	-	Ping	11	1st April 1996	31st March 2007
P21	Mae Rim	Mae Rim	26	1st April 1981	31st December 2007
P1	-	Ping	26	1st April 1981	31st December 2007
P71	Mae Ngan	Mae Khan	11	1st April 1996	31st December 2007
P24A	Mae Klang	Mae Klang	26	1st April 1981	31st December 2007
P73	-	Ping	9	1st April 1998	31st December 2007
061302	Upper Mae Cham	Mae Cham	25	1st January 1983	31st December 2007
P14	-	Mae Cham	26	1st April 1981	31st December 2007
061501	Mae Tuen	Mae Tuen	31	1st January 1977	31st December 2007
P12B	-	Ping	26	1st April 1981	31st December 2007

7.3.1 Low and high flow analysis

Anomalous streamflow events, i.e. low and high flows, are determined using thresholds that are obtained from the frequency-duration-curve (FDC) of observed streamflow. In this study, using daily averaged discharge, a separate FDC is established for each of the12 streamflow stations. Q_{90} and Q_{10} obtained from the FDC for a water year (i.e. from April to March) are the thresholds used to define what an anomalous streamflow event is. Low flow is defined by a daily discharge that is below or equal the Q_{90} obtained from the FDC of observations. Although low flow analysis depends upon defined thresholds that are not classified by any standard, several researchers have suggested using Q_{90} to determine a low flow or streamflow drought (Edossa et al., 2010; Fleig et al., 2006). On the other hand, daily discharge above Q_{10} of the observations represents a high flow. This can be defined as a flooding event, even if the flood situation in a basin also depends on the capacity of the channel and flood plain, the current wetness of soil during consecutive wet days and the water level of the aquifer. In this study, considering only daily discharges at 12 gauging stations without other information, a daily discharge that is above the observed Q_{10} is denoted as high flow or anomalous event.

7.3.2 Seasonal anomaly analysis

An anomalous event of daily discharge in a specific season is defined using thresholds obtained from the FDC for that season. In this study, the anomalous events in wet and dry seasons are examined. The FDC for a wet season is computed using daily average discharges from 1st August to 30th November (i.e. a 122-day period), whereas the FDC for a dry season is developed using daily average discharges from 1st December to 30th April (i.e. a 151-day period). The FDC for wet and dry seasons are done separately for each gauging station. The Q_{90} obtained from the FDC of observations during a wet season (i.e. $Q_{90,wet}$) and Q_{10} achieved from the FDC during a dry season (i.e. $Q_{10,dry}$) are used to define anomalous events in both seasons. A daily discharge that is below or equal to $Q_{90,wet}$ is denoted as an anomalous low flow in a wet season. On the other hand, anomalous high flow in a dry season is defined by a daily discharge that is above $Q_{10,dry}$.

7.4 Results and discussion

The 2011-2100 daily rainfall under two scenarios of future climate (A2 and B2), which is obtained from a multisite daily rainfall generator, is used in the HEC-HMS model to simulate daily discharges. Then, daily simulated streamflow under future climate scenarios are averaged over the period of 2011-2100 and plotted along with the daily average streamflow of observations (Figure 7.1). The effects of future climate on daily average streamflow from 2011 to 2100 are slightly different between the two scenarios. Under A2 and B2, at the 12 gauging stations, the daily average streamflow during the dry season, especially in December and January, decreases. However, at some stations (e.g. Station 061302, P14 and P12B), daily streamflow during the pre-monsoon season (i.e. May-June-July) shows an increase. Moreover, the effects of future climate under both scenarios on decreasing streamflow during the wet season (i.e. from August to November) can be observed at all stations. However, the magnitudes of peak discharge at Station 061302 or the Upper Mae Cham Basin and P14 or the Mae Cham Basin (Figure 7.1(i) and (j) respectively) associated with A2 are higher than the magnitudes of peak discharge under B2. The temporal shift of peak discharge from mid-September to end of September or beginning of October is also found at all stations with the exception of Station P4A (the Mae Taeng Basin), P21 (the Mae Rim Basin), P71 (the Mae Ngan Basin) and 061501 (the Mae Tuen Basin).

The FDC of 2011-2100 simulated streamflow at 12 stations is computed. From Figure 7.2, low flow (i.e. $Q_t \leq Q_{90}$) due to changes in future climate has more chances of occurrence than what observations suggest. However, a higher probability of low flow occurrence from 2011 to 2100 is not consistent among all stations. At Station 061501 (P12B), from historical records, Q_{90} is estimated at 2.57 (16.55) $m^3 s^{-1}$. Daily observations over 31 (26) years suggest that the number of low flow days is averaged at 67 (88) days per year, whereas out of the 90 projected years, a daily discharge less than 2.57 (16.55) $m^3 s^{-1}$ cannot be observed under both scenarios of future climate. At other stations, the averages of low flow days from daily observations vary from 58 to 116 days per year. Due to changes in future climate, from 2011 to 2100, the averages of low flow days under A2 (B2) scenario range from 4 to 167 days (1 to 178 days) per year.

In contrast, at the 12 gauging stations, there is less chance of a high flow (i.e. $Q_t > Q_{10}$) occurrence during the 2011- 2100 period as compared to historical records. In the Mae Taeng Basin (Station P4A), the Mae Rim Basin (Station P21) and the Mae Ngan Basin (Station P71), the Q_{10} of observed discharges is calculated at 27.25, 8.86 and 26.99 $m^3 s^{-1}$ respectively. The number of high flow days indicates an average of 43, 39 and 30 days per year respectively. However, under A2 and B2, the 2011-2100 daily streamflow above Q_{10} cannot be found at these stations. At other stations, the averages of high flow days associated with observations vary from 29 to 40 days per year, whereas due to changes in future climate, the number of high flow days averaged during 2011- 2100 ranges from only 1 to 15 days per year.

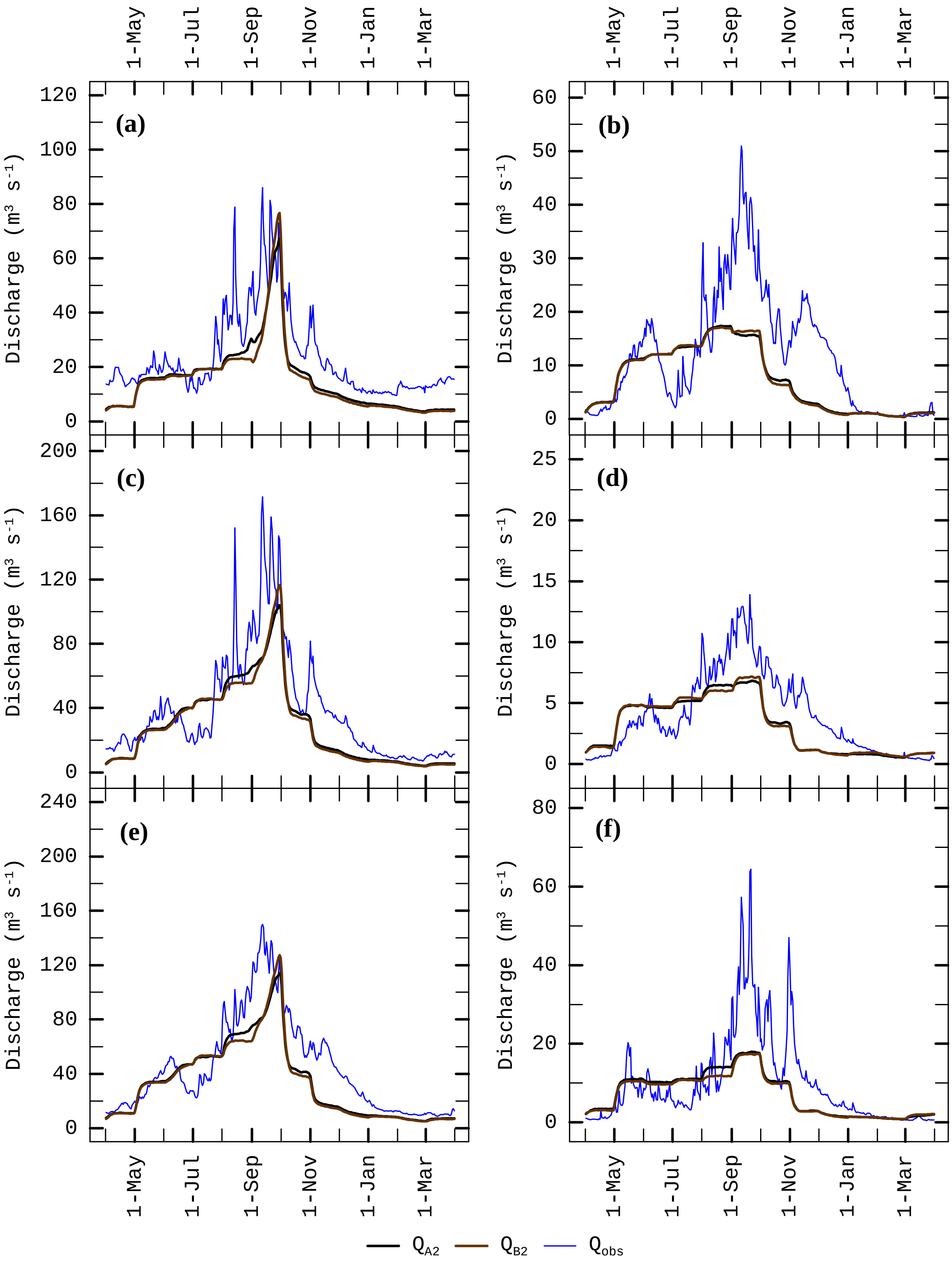


Figure 7.1: Daily averaged streamflow for a water year from 1st April to 31st March at the 12 gauging stations: (a) P75; (b) P4A; (c) P67; (d) P21; (e) P1; (f) P71; (g) P24A; (h) P73; (i) 061302; (j) P14; (k) 061501; and (l) P12B.

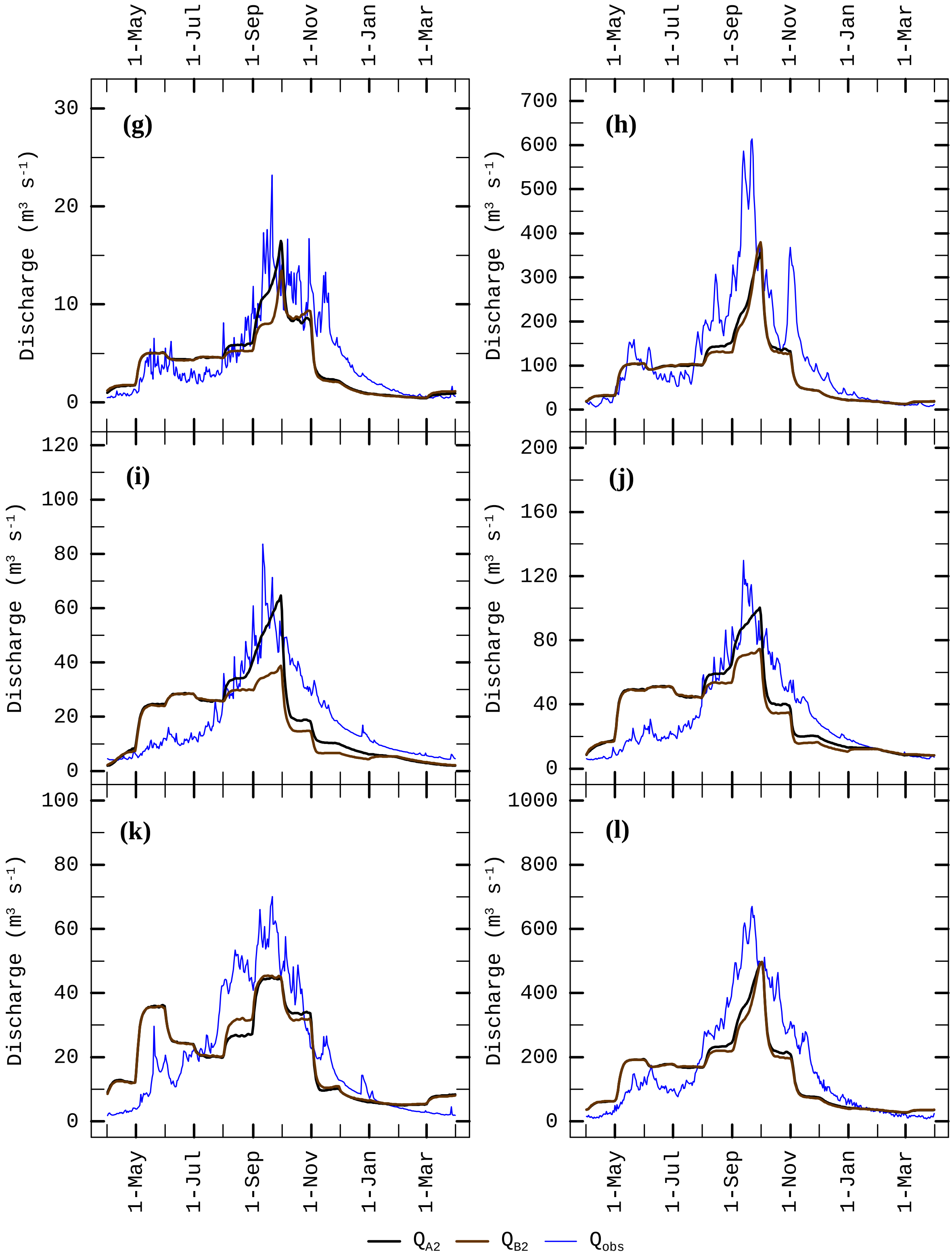


Figure 7.1 (cont).

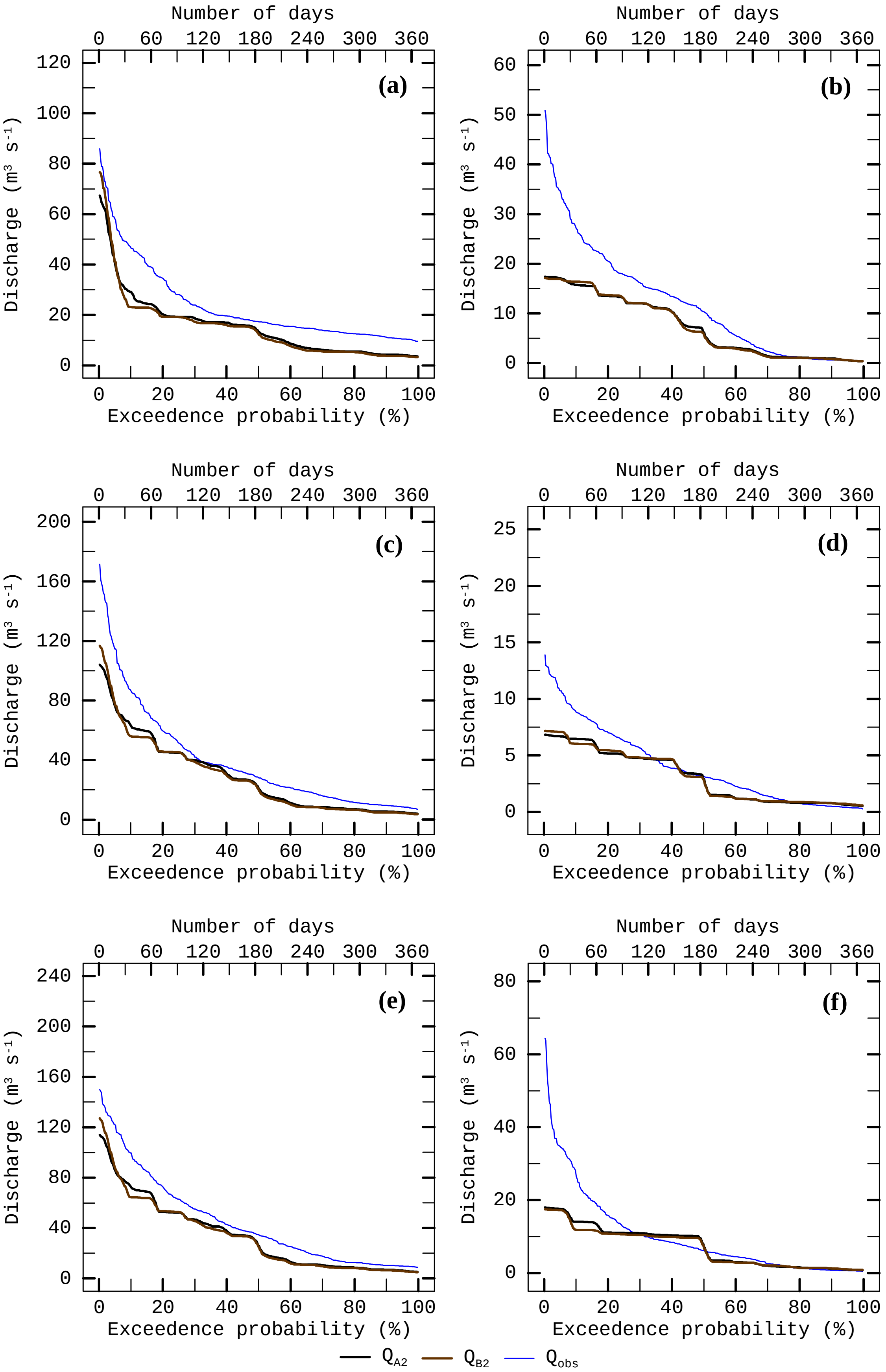


Figure 7.2: Frequency-duration-curve (FDC) for a water year of daily averaged discharges at the 12 gauging stations: (a) P75; (b) P4A; (c) P67; (d) P21; (e) P1; (f) P71; (g) P24A; (h) P73; (i) 061302; (j) P14; (k) 061501; and (l) P12B.

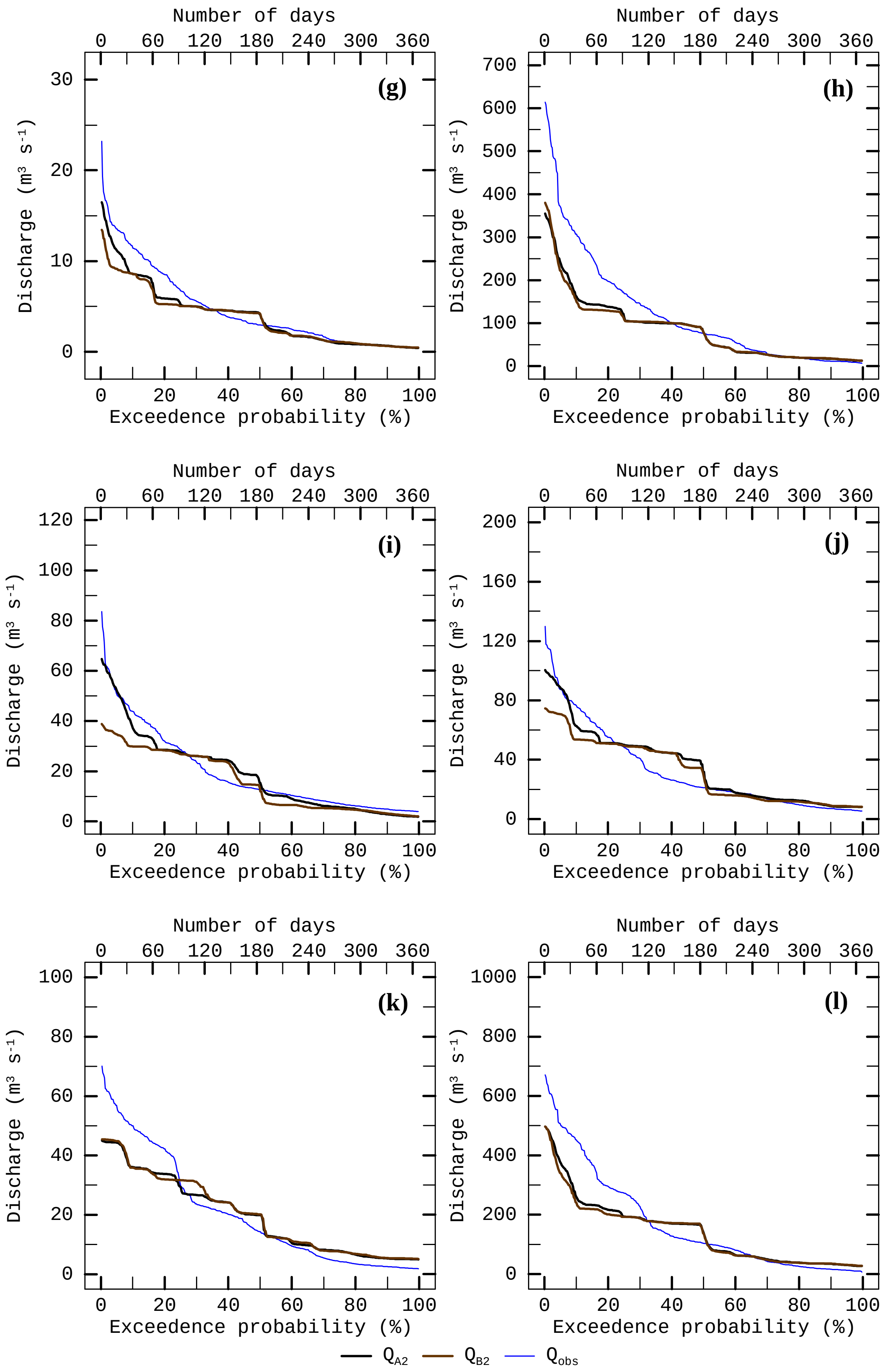


Figure 7.2 (cont).

7.4.1 Effects of future climate on low flow

A daily discharge below or equal to the observed Q_{90} is classified as low flow. The comparison between the Q_{90} magnitude of daily observed and simulated streamflow under both scenarios of future climate is presented in Table 7.2 (see also Figure 7.2). Due to changes in future climate, the magnitude of Q_{90} will decrease at Station P75, P67, P1 and 061302. Under A2 (B2), the maximum decrease in the magnitude of Q_{90}, which is observed at Station P75, is estimated to be 61.6% (65.6%) of the observed Q_{90}. However, at the remaining stations, an increasing magnitude of simulated Q_{90} is found compared to the observed Q_{90}. Moreover, the maximum increasing Q_{90} under A2 (B2) is found at the Bhumipol Dam Station (i.e. P12B) equal to 107.3% (112.9%) corresponding to the observed Q_{90}. As a result, the effects of future climate on decreasing and increasing magnitudes of Q_{90} show an inconsistency among the stations in the study basin.

Table 7.2: Q_{90} ($m^3 s^{-1}$) of the Daily Observed and Simulated Streamflow at the 12 Gauging Stations

Station	Observation	A2	B2	Station	Observation	A2	B2
P75	11.16	4.28	3.84	P24A	0.63	0.66	0.63
P4A	0.63	0.93	0.81	P73	11.59	17.35	17.93
P67	9.50	5.46	4.93	061302	4.97	2.91	3.23
P21	0.50	0.78	0.76	P14	7.17	8.75	9.12
P1	10.48	7.08	6.64	061501	2.57	5.28	5.45
P71	0.72	1.19	1.23	P12B	16.55	34.31	35.23

In terms of low flow duration, discharges below Q_{90} are observed during the dry season (from December to April). However, an event of low flow with an anomalous temporal extension may start early around the end of the wet season (i.e. in November) or may end late around the beginning of the following pre-monsoon season (i.e. in May). Historical records of low flows based on Q_{90} at the 12 stations are presented in Table 7.3. During the longest spell of low flow, the shortage volume is the difference between Q_{90} and the discharge on a specific day. The total volume of the shortage is the sum of shortage volume during the longest spell. The intensity of shortage is computed as the total volume of shortage divided by the number of days corresponding to the longest spell of low flow.

From historical records (Table 7.3), the longest spell of low flow ranges from 77 days at Station P14 (from 12th March to 27th May 1983) to 211 days at Station P1 (from 26th December 1991 to 23rd July 1992). The longest spell of low flow in the study basin is consistent with the years of El Niño (see also Table 3.2), especially so for the strongest El Niño effects seen in 1982-83, 1991-92 and 1997-98. This makes it evident that decreasing rainfall in the study basin is influenced by atmospheric anomalies (see also Chapter 3). Additionally, the maximum total shortage volume is estimated to be 208.53 MCM at the Bhumipol Dam Station (P12B) with an intensity of shortage equal to 1.345 MCM per day.

Under A2, the longest spell of low flow observed at Station P75, P67 and 061302 will be longer compared to the observations and vice versa for other stations (Table 7.4). However, low flow cannot be observed from 2011 to 2100 at Station 061501 and P12B. Among the other 10 gauging stations, the longest spell will occur during 2nd November 2096 to 2nd May 2097 (i.e. 182 days) at Station P75, with a shortage intensity of 0.505 MCM per day. Also, under A2, the severity of shortage during the longest spell of low flow will decrease compared to historical records and this is reflected in the lower intensity of the shortage. However, more intense of shortage situations can be found at two gauging stations: P67 and 061302. The shortage intensity at Station P67 will slightly increase by 1.74% more than that indicated by historical intensity. Moreover, the most severe shortage during the longest spell of low flow will be found at Station 061302 with an increasing intensity of shortage equal to 72.03% more than that of historical intensity.

Table 7.3: Low Flow Characteristics of the Historical Data at the 12 Gauging Stations

Station	Longest spell of low flow (day)	Start date	End date	Total volume of shortage (MCM)	Intensity of shortage (MCM d^{-1})
P75	121	24-Nov-1999	23-Mar-2000	70.12	0.580
P4A	128	14-Dec-1998	20-Apr-1999	5.34	0.042
P67	99	27-Dec-1996	4-Apr-1997	28.38	0.287
P21	102	3-Mar-1992	12-Jun-1992	4.10	0.040
P1	211	26-Dec-1991	23-Jul-1992	89.48	0.424
P71	138	25-Mar-1998	9-Aug-1998	7.36	0.053
P24A	145	8-Dec-2003	30-Apr-2004	5.95	0.041
P73	154	8-Dec-1998	10-May-1999	131.03	0.851
061302	105	4-Feb-2005	19-May-2005	15.01	0.143
P14	77	12-Mar-1983	27-May-1983	16.12	0.209
061501	115	15-Dec-1998	8-Apr-1999	11.22	0.098
P12B	155	30-Nov-2003	2-May-2004	208.53	1.345

MCM: million m^3.

Table 7.4: Low Flow Characteristics of the 2011-2100 Simulated Streamflow under A2 at the 12 Gauging Stations

Station	Longest spell of low flow (day)	Start date	End date	Total volume of shortage (MCM)	Intensity of shortage (MCM d^{-1})
P75	182	2-Nov-2096	2-May-2097	91.83	0.505
P4A	38	25-Jan-2027	3-Mar-2027	0.74	0.019
P67	147	6-Dec-2081	1-May-2082	42.96	0.292
P21	31	11-Feb-2021	13-Mar-2021	0.34	0.011
P1	142	6-Dec-2029	26-Apr-2030	37.39	0.263
P71	55	14-Feb-2072	8-Apr-2072	2.00	0.036
P24A	94	9-Jan-2088	11-Apr-2088	2.80	0.030
P73	26	16-Feb-2088	12-Mar-2088	4.60	0.177
		6-Feb-2097	3-Mar-2097	4.02	0.155
061302	139	12-Dec-2099	29-Apr-2100	34.17	0.246
P14	65	29-Jan-2040	2-Apr-2040	8.40	0.129
061501	0	n/a	n/a	0	0
P12B	0	n/a	n/a	0	0

MCM: million m^3.

The low flow under Scenario B2 (Table 7.5) is consistent with the results of Scenario A2. From 2011 to 2100, the longest spell of low flow at Station P75, P67 and 061302 will be extended by 185, 156 and 138 days respectively. However, the longest spell at other stations will be shortened, as compared to historical records. Daily discharge below 2.57 (16.55) m^3 s^{-1} (i.e. Q_{90} of the observations) will not be observed at Station 061501 (P12B). From 2011 to 2100, among the other 10 stations, the longest spell of 185 days will be seen at Station P75. It will last from 4th November 2078 to 7th May 2079, when a water shortage of 88.68 MCM in total volume and a shortage intensity of 0.479 MCM per day will be experienced. At Station P67, 061302 and P14, the shortage situation will be more severe than the historical situation was. Compared to historical intensity, the shortage intensity will increase by 16.72%, 81.12% and 28.23% respectively for each station mentioned above.

Table 7.5: Low Flow Characteristics of the 2011-2100 Simulated Streamflow under B2 at the 12 Gauging Stations

Station	Longest spell of low flow (day)	Start date	End date	Total volume of shortage (MCM)	Intensity of shortage (MCM d^{-1})
P75	185	4-Nov-2078	7-May-2079	88.68	0.479
P4A	50	6-Feb-2080	26-Mar-2080	1.21	0.024
P67	156	26-Nov-2086	30-Apr-2087	52.26	0.335
P21	38	18-Feb-2083	27-Mar-2083	0.51	0.013
P1	142	10-Dec-2069	30-Apr-2070	36.69	0.258
P71	32	29-Jan-2021	1-Mar-2021	0.78	0.024
		30-Jan-2030	2-Mar-2030	0.61	0.019
P24A	72	26-Dec-2049	7-Mar-2050	1.30	0.018
P73	9	22-Feb-2083	2-Mar-2083	0.27	0.031
061302	138	18-Dec-2015	3-May-2016	35.80	0.259
P14	74	25-Jan-2030	8-Apr-2030	19.81	0.268
061501	0	n/a	n/a	0	0
P12B	0	n/a	n/a	0	0

MCM: million m^3.

When comparing between A2 and B2 (Figure 7.3), the different effects of future climate on low flow in the study basin can be observed at all stations. At Station P71 (the Mae Ngan Basin), P24A (the Mae Klang Basin) and P73, a longer spell of low flow with a greater intensity of shortage is associated more closely with results under A2 than B2. However, at Station P4A (the Mae Taeng Basin), P67, P21 (the Mae Rim Basin) and P14 (the Mae Cham Basin), the effects of future climate under A2 indicate a shorter spell with a lower intensity of shortage than the effect under B2. Moreover, at the remaining stations, the effects of future climate in terms of the spell of low flow and an intensity of shortage are not consistent between the two scenarios.

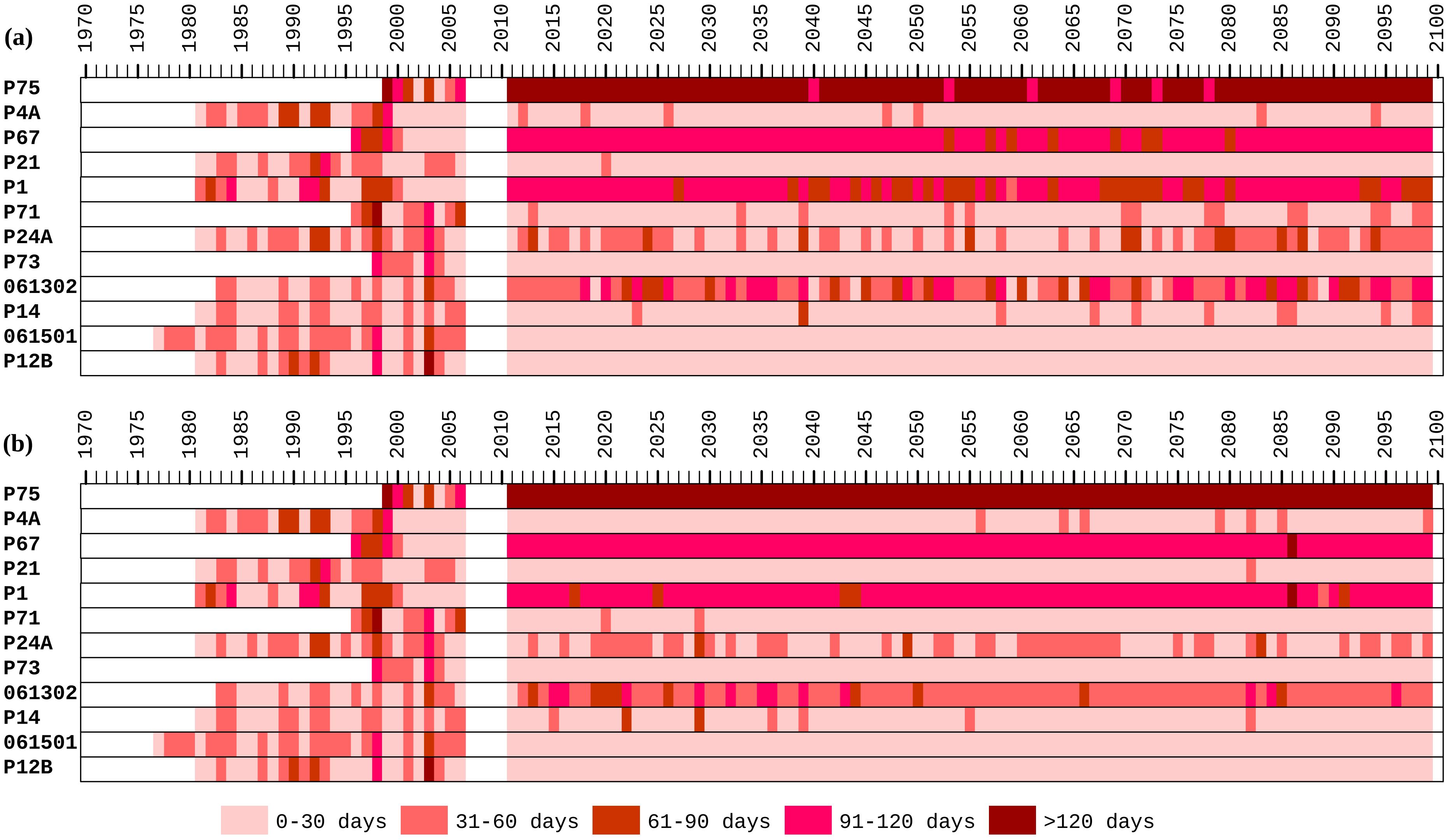


Figure 7.3: Dry spell of daily observed streamflow and 2011-2100 simulated streamflow at 12 gauging stations under scenarios: (a) A2; and (b) B2 of future climate.

7.4.2 Effects of future climate on high flow

High flow is denoted as a daily discharge that is greater than the Q_{10} of observations. Table 7.6 presents the comparison between Q_{10} magnitudes of the daily observed and simulated discharges under A2 and B2. Due to changes in climate, a decrease can be observed in Q_{10} magnitudes that are consistent among all stations. Under A2 (B2), the maximum decrease in the magnitude of Q_{10} is observed in the Mae Ngan Basin or Station P71 and corresponds to 47.9% (56.1%) of the Q_{10} from daily observed discharges. However, the minimum decrease in the magnitude of Q_{10} under A2 (B2) found at Station 061302 (P24A) is estimated to be 14.7% (25.5%) of the Q_{10} of historical data. At all stations, the Q_{10} under A2 is greater than Q_{10} under B2 except at Station P4A of the Mae Taeng Basin. Therefore, future climate influences on high flow (i.e. $Q_t > Q_{10}$) in the study basin is evident in the decreasing Q_{10} magnitude. Greater decreases in Q_{10} magnitudes are associated more closely with B2 than with A2.

Table 7.6: Q_{10} ($m^3 s^{-1}$) of the Daily Observed and Simulated Streamflow at the 12 Gauging Stations

Station	Observation	A2	B2	Station	Observation	A2	B2
P75	46.61	29.14	23.13	P24A	11.51	8.60	8.57
P4A	27.25	15.73	16.39	P73	307.15	162.75	151.72
P67	86.39	63.08	55.97	061302	43.85	37.38	29.95
P21	8.86	6.47	6.03	P14	76.55	62.82	53.60
P1	99.62	72.44	64.54	061501	50.00	36.08	35.82
P71	26.99	14.06	11.84	P12B	451.38	260.73	240.12

In terms of duration, in general, discharges above Q_{10} are observed in the wet season (i.e. from August to November). The spell of high flow can be anomalously extended from an early start around the end of the pre-monsoon season (i.e. in July). Based on Q_{10}, Table 7.7 shows the longest spell of high flow according to historical records at the gauging stations considered in this study. During the longest spell, an abundant volume is denoted as the daily discharge that is greater than the observed Q_{10}. The total abundant volume is the sum of abundant water during the longest spell of high flow. The intensity of abundance is estimated as the total abundant volume divided by the number of days corresponding to the longest spell.

From Table 7.7, in the historical records, the longest spells of high flow occur in La Niña years (see also Table 3.2), in particular during the strongest La Niña effects in 1995 and 2006. This is because the monsoon season rainfall in the study basin tends to increase during La Niña years. The longest spell of high flow is found at Station P1 and lasts from 31st July to 14th October 1995 (i.e. 76 days). The maximum abundant volume can be observed at Station P73 and is estimated to be 1,549.74 MCM with an intensity of abundance equal to 32.973 MCM per day.

Under A2 (Table 7.8), from 2011 to 2100, daily simulated discharges above the observed Q_{10} cannot be found in the Mae Taeng Basin (Station P4A), the Mae Rim Basin (Station P21) and the Mae Ngan Basin (Station P71). Among the remaining stations, the longest spell of high flow will be observed at Station 061302 or the Upper Mae Cham Basin and will last from 6th August to 8th October 2040 (i.e. 64 days) with an intensity of abundant water equal to 4.064 MCM per day. The high flow spells at Station P24A, 061302 and P14 will be of more duration than what appears in historical records and vice versa for the other stations. Considering the decreasing volume and weaker intensity of abundance, the high flow situation at all stations except Station 061302 (the Upper Mae Cham Basin) can be classified as under a more temperate condition than that seen in observations. In the Upper Mae Cham Basin, the total volume of abundant water during the longest spell of high flow will increase by 102.97% more than that of historical records. Consequently, the Upper Mae Cham Basin will have a higher tendency of flooding under A2.

Table 7.7: High Flow Characteristics of Historical Data at the 12 Gauging Stations

Station	Longest spell of high flow (day)	Start date	End date	Total abundant volume (MCM)	Intensity of abundance (MCM d^{-1})
P75	59	26-Aug-2005	23-Oct-2005	295.23	5.004
P4A	58	15-Aug-1994	11-Oct-1994	263.88	4.550
P67	49	31-Aug-2005	18-Oct-2005	692.48	14.132
P21	53	16-Aug-1994	7-Oct-1994	46.89	0.885
P1	76	31-Jul-1995	14-Oct-1995	682.78	8.984
P71	23	8-Sep-2006	30-Sep-2006	78.40	3.409
P24A	38	10-Sep-2006	17-Oct-2006	78.02	2.053
P73	47	24-Aug-2002	9-Oct-2002	1,549.74	32.973
061302	56	20-Aug-1994	14-Oct-1994	128.15	2.288
P14	45	30-Aug-1995	13-Oct-1995	313.01	6.956
061501	68	15-Jul-1994	20-Sep-1994	450.37	6.623
P12B	58	16-Aug-1995	12-Oct-1995	1,358.96	23.430

MCM: million m^3.

Table 7.8: High Flow Characteristics of the 2011-2100 Simulated Streamflow under A2 at the 12 Gauging Stations

Station	Longest spell of high flow (day)	Start date	End date	Total abundant volume (MCM)	Intensity of abundance (MCM d^{-1})
P75	46	20-Aug-2040	4-Oct-2040	208.86	4.540
P4A	0	n/a	n/a	0	0
P67	45	20-Aug-2040	3-Oct-2040	210.15	4.670
P21	0	n/a	n/a	0	0
P1	44	21-Aug-2040	3-Oct-2040	199.45	4.533
P71	0	n/a	n/a	0	0
P24A	45	31-Aug-2027	14-Oct-2027	50.64	1.125
P73	38	27-Aug-2040	3-Oct-2040	353.65	9.307
		27-Aug-2093	3-Oct-2093	400.87	10.549
061302	64	6-Aug-2040	8-Oct-2040	260.10	4.064
P14	63	6-Aug-2040	7-Oct-2040	293.58	4.660
061501	8	14-Sep-2033	21-Sep-2033	1.84	0.230
P12B	39	28-Aug-2073	5-Oct-2073	531.98	13.640

MCM: million m^3.

Under B2 (Table 7.9), the high flow at Station P4A, P21 and P71 is consistent with A2. In other words, the daily modeled discharge that is above Q_{10} of historical data cannot be observed from 2011 to 2100. However, at the remaining gauging stations, the duration of high flow spells is shorter as compared to durations in historical records. The longest spell of high flow is found in the Mae Cham Basin (Station P14) and will last from 4th September to 4th October 2013 (i.e. 31 days). At Station P75, P67 and P12B, the long spell of high flow will be observed in several years. However, the total volume and intensity of abundant water at all stations is less than those of observations. As results, from 2011 to 2100 show, the study basin has lower tendency of flooding under Scenario B2.

Table 7.9: High Flow Characteristics of the 2011-2100 Simulated Streamflow under B2 at the 12 Gauging Stations

Station	Longest spell of high flow (day)	Start date	End date	Total abundant volume (MCM)	Intensity of abundance (MCM d^{-1})
P75	29	5-Sep-2021	3-Oct-2021	81.51	2.811
		6-Sep-2049	4-Oct-2049	69.76	2.405
		5-Sep-2079	3-Oct-2079	96.94	3.343
P4A	0	n/a	n/a	0	0
P67	28	6-Sep-2021	3-Oct-2021	87.48	3.124
		7-Sep-2049	4-Oct-2049	67.63	2.416
		6-Sep-2079	3-Oct-2079	101.20	3.614
P21	0	n/a	n/a	0	0
P1	28	6-Sep-2021	3-Oct-2021	81.90	2.925
P71	0	n/a	n/a	0	0
P24A	18	18-Sep-2013	5-Oct-2013	12.44	0.691
P73	19	15-Sep-2033	3-Oct-2033	67.88	3.572
061302	21	13-Sep-2025	3-Oct-2025	72.01	3.429
P14	31	4-Sep-2013	4-Oct-2013	89.49	2.887
061501	5	14-Sep-2014	18-Sep-2014	1.66	0.332
P12B	15	20-Sep-2013	4-Oct-2013	117.38	7.825
		20-Sep-2038	4-Oct-2038	105.59	7.040
		20-Sep-2041	4-Oct-2041	76.99	5.133
		20-Sep-2052	4-Oct-2052	108.04	7.203

MCM: million m^3.

When comparing A2 and B2 (Figure 7.4), the duration of high flow spells under B2 is shorter than that under A2. At some stations (e.g. P75, P67 and P12B), the long spells of high flow under B2 will be observed more frequently from 2011 to 2100. The total volume and the intensity of abundant water associated with B2 are less than those with A2. Therefore, the effects of future climate under B2 on high flow in the study basin are temperate compared to the effects under A2.

Consequently, in terms of low flow, the effects of future climate under both scenarios indicate a decrease in the magnitude of Q_{90} compared to observations at Station P75, P67, P1 and 061302 and vice versa for other stations. In terms of low flow spells from 2011 to 2100, under both A2 and B2, the length of a spell of low flow at Station P75, P67 and 061302 will be longer than that shown in historical records and vice versa for other stations. Furthermore, low flows will not occur at two stations – 061501 and P12B. Comparing A2 and B2, the effects of future climate on daily streamflow in terms of the duration of low flow spells and an intensity of shortage are not consistent among all the streamflow stations in the study basin.

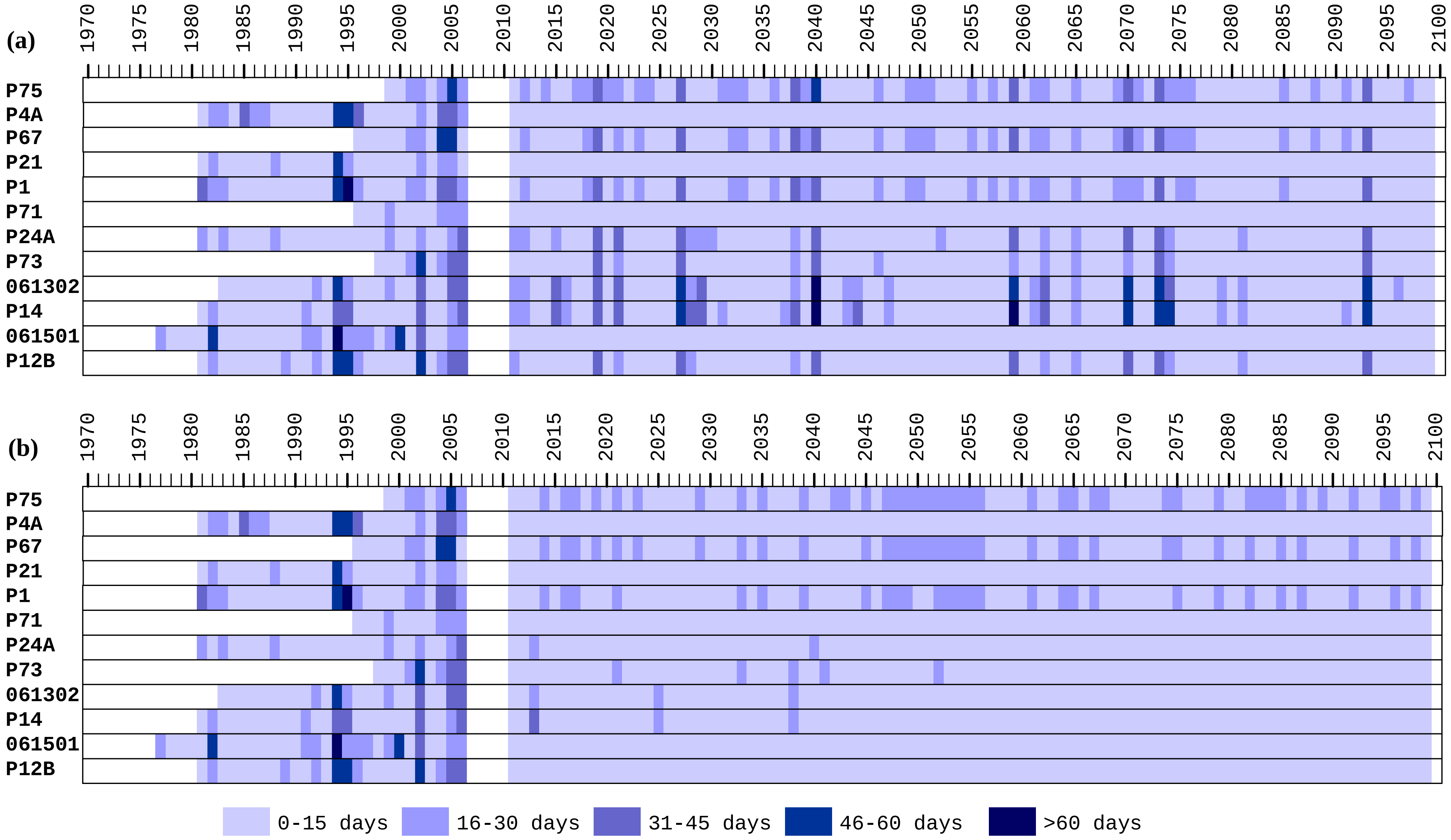


Figure 7.4: Wet spell of daily observed streamflow and 2011-2100 simulated streamflow at 12 gauging stations under scenarios: (a) A2; and (b) B2 of future climate.

As for high flow, the magnitudes of Q_{10} at all stations decrease under both scenarios of future climate. The duration of high flow spells at Station P24A, 061302 and P14 will be extended under A2 and vice versa for other stations. However, the duration of high flow spells at all stations will be shortened under B2. Moreover, under B2, consecutive days of high flow will occur more often at some stations (e.g. P75, P67 and P12B) than others. Based on larger volume and higher intensity of abundant water, the effects of future climate on high flow under A2 are more severe than the effects under B2.

7.4.3 Effects of future climate on seasonal anomalies: anomalous low flow in a wet season

An anomalous low flow in a wet season (i.e. from 1st August to 30th November) is defined as a daily discharge below or equal to $Q_{90,wet}$ of the observed streamflow. As seen in Table 7.10, the decreasing magnitudes of $Q_{90,wet}$ and $Q_{10,dry}$ at the 12 stations are influenced by changes in climate under both scenarios. Comparing A2 and B2, the effects of future climate on $Q_{90,wet}$ and $Q_{10,dry}$ are a little different from each other. However, the magnitudes of $Q_{90,wet}$ associated with A2 are slightly higher than those with B2 at all stations with the exception of Station P73 and 061501. The $Q_{10,dry}$ magnitudes under A2 are also slightly greater than those under B2, except at Station P24A, P73 and P12B.

Table 7.10: $Q_{90,wet}$ and $Q_{10,dry}$ of the Daily Observed and Simulated Streamflow at the 12 Gauging Stations

Station	$Q_{90,wet}$ ($m^3 s^{-1}$)			$Q_{10,dry}$ ($m^3 s^{-1}$)		
	Observation	A2	B2	Observation	A2	B2
P75	20.70	11.00	9.75	16.17	7.76	6.72
P4A	14.13	3.10	2.84	12.26	3.11	3.02
P67	37.06	15.15	13.75	24.64	9.74	8.64
P21	4.71	1.15	1.14	2.86	1.49	1.38
P1	52.31	17.18	15.82	31.42	11.37	11.04
P71	9.11	2.91	2.75	5.07	3.34	3.00
P24A	5.25	2.38	2.18	3.36	1.71	1.76
P73	114.52	46.50	46.94	54.01	31.80	32.72
061302	23.07	10.39	6.59	14.21	8.22	6.27
P14	40.73	20.25	16.03	21.77	16.97	16.06
061501	19.65	9.95	10.73	10.54	12.38	12.14
P12B	222.10	78.12	75.31	87.01	61.83	62.50

Of the 122 days in a wet season, the average low flow days in historical data when $Q_t \leq Q_{90,wet}$ at the 12 stations range from 38 to 57 days per season. The longest spell of anomalous low flow in a wet season varies from 28 to 83 days (Table 7.11). The longest spell from 9th September to 30th November 1998 is observed at Station 061501 with an intensity of shortage equal to 1.175 MCM per day. The maximum intensity of shortage corresponding to Station P12B is estimated to be 14.675 MCM per day.

Under A2, over 90 simulation years, a minimum (maximum) of average low flow days in a wet season is observed at Station 061501 (P4A) and is estimated to be 28 (62) days per season. Comparing this to historical records, the spell of low flow is shortened at all stations except Station P75 and P21 (Appendix J1). Due to changes in future climate, the onset date of low flow spells in a wet season, which starts around the beginning of October or November, is later than the onset date of observations (i.e. around mid-September). In addition, low flow spells during a wet season at various stations like P21, 061302 and

061501 will be observed in several years during 2011-2100 with a lower intensity of shortage than the observations indicate. Overall, under A2, the severity of anomalous shortage in the wet season is temperate compared to historical records, with the exception of Station P75.

Table 7.11: Characteristics of Anomalous Low Flow in Wet Season from Historical Data at the 12 Gauging Stations

Station	Longest spell of anomalous low flow (day)	Start date	End date	Total volume of shortage (MCM)	Intensity of shortage (MCM d^{-1})
P75	28	6-Oct-1999	2-Nov-1999	14.52	0.519
P4A	79	13-Sep-1998	30-Nov-1998	86.97	1.101
P67	78	14-Sep-1998	30-Nov-1998	187.02	2.398
P21	52	13-Sep-1998	3-Nov-1998	15.12	0.291
P1	78	14-Sep-1998	30-Nov-1998	255.38	3.274
P71	78	14-Sep-1998	30-Nov-1998	52.06	0.667
P24A	71	21-Sep-1998	30-Nov-1998	24.53	0.346
P73	75	17-Sep-1998	30-Nov-1998	513.37	6.845
061302	74	18-Sep-1998	30-Nov-1998	83.41	1.127
P14	75	17-Sep-1998	30-Nov-1998	150.65	2.009
061501	83	9-Sep-1998	30-Nov-1998	97.56	1.175
P12B	71	21-Sep-1998	30-Nov-1998	1,041.90	14.675

MCM: million m^3.

Out of 122 days in a wet season, the average low flow days over 90 simulation years at 12 streamflow stations vary from 27 to 69 days per season under B2. From 2011 to 2100, the longest spell of anomalous low flow in a wet season under B2 is consistent with A2. Comparing the observations, the duration is shortened at all stations with the exception of Station P75 and P21 (Appendix J2). Among the 12 stations, the longest spell of anomalous low flow in a wet season, corresponding to Station P4A, lasts from 1st October to 30th November (i.e. 61 days) in several years (2013, 2015, 2019 and 2020). The shortest spell of anomalous low flow in a wet season of 29 days (i.e. from 2nd November to 30th November 2060) is associated with Station 061501. The onset date of the spell shows a shift by 0.5-1.5 months from mid-September in the observed streamflow to the beginning of October or November in the simulated streamflow. However, at Station P24A, an anomalous low flow can be observed during the beginning of the wet season (i.e. from 1st August to 1st September). Anomalous low flow in the wet season under B2 is temperate, with a shortage intensity ranging from 0.030 to 7.446 MCM per day.

When comparing A2 and B2, a longer spell of anomalous low flow during the wet season under B2 corresponds to Station P14 of the Mae Cham Basin and Station 061501 of the Mae Tuen Basin. Otherwise, a shorter period is observed at other stations. The frequency of occurrence of anomalous low flow in a wet season is inconsistent among the 12 streamflow stations. From 2011 to 2100, anomalous low flow associated with A2 at Station P75, P67, P21, P1 and 061501 will be observed more frequently in the wet season than B2; however, at Station P4A, P71, P24A 061302 and P14, anomalous low flow under B2 will occur more often than under A2. In terms of the severity of shortage in the wet season, the shortage intensity is not significantly different between both scenarios.

7.4.4 Effects of future climate on seasonal anomalies: anomalous high flow in a dry season

Anomalous high flow in a dry season (i.e. from 1st December to 30th April) is defined as a daily discharge greater than $Q_{10,dry}$ of the historical streamflow. In terms of the number of high flow days in the dry season (i.e. $Q_t>Q_{10,dry}$), the average of high flow days from the observed discharges at 12 stations ranges from 12 to 49 days per dry season. At some stations, the anomalous high flow in the dry season has not been observed for several years (such as 1978-80, 1985, 1990 and 2004 at Station 061501). From historical records, the longest spell of anomalous high flow at 12 stations varies from 28 to 103 days (Table 7.12). The longest spell from 18th January to 30th April 2006 is observed at Station P75 with the intensity of abundance of 0.616 MCM per day. At Station P73, a maximum of abundant volume during consecutive high flow days in the dry season is estimated to be 362.34 MCM.

Table 7.12: Characteristics of Anomalous High Flow in Dry Season from Historical Data at the 12 Gauging Stations

Station	Longest spell of anomalous high flow (day)	Start date	End date	Total abundant volume (MCM)	Intensity of abundance (MCM d^{-1})
P75	103	18-Jan-2006	30-Apr-2006	63.43	0.616
P4A	41	1-Dec-2002	10-Jan-2003	29.70	0.724
P67	50	1-Dec-2002	19-Jan-2003	124.63	2.493
P21	54	1-Dec-2002	23-Jan-2003	14.29	0.265
P1	51	1-Dec-2002	20-Jan-2003	119.82	2.350
P71	72	1-Dec-2002	10-Feb-2003	34.45	0.479
P24A	51	1-Dec-1981	20-Jan-1982	13.86	0.272
P73	54	1-Dec-2002	23-Jan-2003	362.34	6.710
061302	40	1-Dec-1985	9-Jan-1986	15.75	0.394
P14	46	1-Dec-2002	15-Jan-2003	40.50	0.880
061501	43	1-Dec-2002	12-Jan-2003	38.82	0.903
P12B	28	1-Dec-1981	28-Dec-1981	149.34	5.333

MCM: million m^3.

Under A2, anomalous high flow in the dry season from 2011 to 2100 can only be observed at four stations (Station 061302, P14, 061502 and P12B). The average high flow days over 90 simulation years vary from 0.2 to 25 days out of the 151 days in a dry season. Among the four stations mentioned above, the longest spell of anomalous high flow in the dry season is estimated to be 33 days, from 29th March to 30th April 2090, at Station 061501 with an intensity of abundance of 0.188 MCM per day. However, the shortest spell of anomalous high flow in the dry season lasts five consecutive days, from 1st to 5th December 2040. It will be observed at Station P12B with a maximum intensity of abundance of 0.431 MCM per day.

On the other hand, the effects of future climate under B2 on anomalous high flow in the dry season from 2011 to 2100 will be observed solely at three stations (P14, 061501 and P12B). With an intensity of abundance of 0.198 MCM per day, the longest spell of anomalous high flow in the dry season covers 30 days in April 2036 at Station 061501. The shortest period of 2 consecutive days (1st and 2nd December 2079) will occur at Station

P12B. During the spell of high flow, a maximum intensity of abundance is estimated to be 0.221 MCM per day at Station P14.

7.5 Summary

The HEC-HMS model is adopted to simulate daily streamflow at 12 gauging stations using the 2011-2100 daily rainfall at 50 rainfall stations. The data is acquired from a multisite daily rainfall generator. The effects of future climate under A2 and B2 are determined using observed and modeled streamflow. In terms of magnitude, the daily average streamflow in the dry season (especially in December and January) and in the wet season from 2011 to 2100 will decrease at all stations under both scenarios. The peak discharge, which is found in mid-September from observed streamflow, will shift to the end of September or the beginning of October due to changes in future climate.

The Q_{90} and Q_{10} from the FDC of daily observed discharges are used to define low and high flows respectively. A daily discharge that is below or equal (above) to Q_{90} (Q_{10}) of the observations is defined as low (high) flow. Under A2 and B2, the magnitudes of Q_{90} of modeled streamflow are greater than those of observations at some stations; however, the magnitude of Q_{10} of simulated discharge is lower than the Q_{10} of observations at all stations. From 2011 to 2100, the spell of low flow at Station P75, P67 and 061302 will be longer than that seen in historical records. In the 90 simulation years, a low flow event will not occur at Station 061501 and P12B. More severity of shortage during low flow spells due to changes in future climate is found at Station P67 and 061302. On the other hand, the spell of high flow will be extended (shortened) at Station P24A, 061302 and P14 (all stations) under A2 (B2) compared to observations. However, from 2011 to 2100, a high flow event will not occur at Station P4A, P21 and P71 under both scenarios. The intensity of abundance due to changes in future climate is lower than that seen in historical records at all stations with the exception of Station 061302.

Daily discharge less (more) than $Q_{90,wet}$ ($Q_{10,dry}$) from the FDC of daily observed streamflow is defined as anomalous low (high) flow in the wet (dry) season. At 12 gauging stations, the magnitudes of $Q_{90,wet}$ and $Q_{10,dry}$ associated with 2011-2100 simulated discharges under two scenarios of future climate are less than those of the observed discharges. Due to changes in future climate, the spell of anomalous low flow in a wet season will be shortened at all stations except Station P75 and P21. However, anomalous low flow in the wet season will be observed in several years during the 2011-2100 periods under both scenarios. The effects of future climate also indicate a delay, from historical records, of the onset date of anomalous low flow in the wet season by 0.5-1.5 months. The severity of shortage in the wet season will decrease at all stations, as compared to observations. On the other hand, from 2011 to 2100, anomalous high flow in the dry season with a shorter duration of consecutive high flow days than the observations will only occur at four (three) gauging stations under A2 (B2). Furthermore, the intensity of abundance is less than that in historical records. Hence, based on the variability of streamflow and availability of water, agricultural practices and water resource plans will have to be developed to handle with the problems of anomalous streamflow events due to climate in the future.

Chapter 8
Conclusions and Recommendations

8.1 Conclusions

The overall objective of the dissertation was to develop a model to forecast rainfall and streamflow with the identified predictors of atmospheric variables and to determine the effects of future climate on rainfall and streamflow in the study basin. Significant relationships between rainfall in a basin of the Upper Chao Phraya River Basin, namely the Ping River Basin, and large-scale atmospheric variables were investigated to identify predictors for a statistical forecasting model. Using correlation maps, four predictors (i.e. surface air temperature, sea level pressure, and the surface zonal and meridian winds) over different regions were selected based on significant correlations with rainfall during pre-monsoon (May-June-July: MJJ), monsoon (August-September-October: ASO) and dry (November-December-January: NDJ and February-March-April: FMA) seasons. Based on the availability of data, the 1961-2100 gridded monthly data of identified predictors from a selected general circulation model (the GFDL-R30) was used to develop a statistical model and to determine the effects of future climate on rainfall in the study basin under two scenarios, A2 and B2. The optimal combination of predictors is indicated using generalized cross validation to avoid redundancy of predictors.

8.1.1 Statistical approach

A modified k-nn model has been developed to forecast rainfall and to downscale rainfall in the Ping River Basin from the large-scale atmospheric variables by an optimal combination of predictors. The model also aims to determine the effects of future climate on seasonal rainfall under Scenario A2 and B2. The model performance was evaluated from 1962 to 2007 using leave-one-out cross validation with four criteria: (i) the annual variability of rainfall; (ii) the annual statistics of seasonal rainfall – i.e. mean, median, standard deviation (SD), interquartile range (IQR) and coefficient of skew (skew); (iii) absolute bias; and (iv) the likelihood skill score (LLH). Using large-scale atmospheric variables under both scenarios of future climate as independent variables, the modified k-nn model performs fairly well in capturing the annual variability of the 1962-2007 seasonal rainfall. As seen in the study, the model can also preserve the annual statistics of seasonal observed rainfall with an absolute bias below 32%. Furthermore, the modified k-nn model shows a better performance than climatology in capturing the probability density function (PDF) with LLH scores greater than +1.0.

The multisite daily rainfall generator resamples historical daily rainfall at 50 selected rainfall stations which are located in and around the Ping River Basin. Daily rainfall, which is used as input for the multisite daily rainfall generator, is randomly selected by a conditioning daily rainfall generator. The random selection of the conditioning daily rainfall generator is done based on categorical probabilities, which are estimated from the PDF of seasonal rainfall ensembles obtained from the modified k-nn model. The performance of the multisite daily rainfall generator is evaluated by its preservation of transition probabilities, spell statistics (e.g. the dry- and wet-spell lengths) and basic statistics (e.g. mean, SD and skew) of daily observed rainfall. It is also evaluated based on how well it can capture spatial cross-correlations among the 50 rainfall stations. The multisite daily rainfall generator can accurately preserve the transition probabilities of daily observations. However, the underestimation of dry- and wet-spell lengths occurs in May and June. Overall, the multisite daily rainfall generator can capture well the basic statistics of daily observed rainfall. Furthermore, the linear relationships of daily rainfall among the

50 selected stations are also well maintained by the multisite generator. The results from the multisite daily rainfall generator are simultaneously applied in a rainfall-runoff model to simulate daily streamflow under A2 and B2 scenarios of future climate.

8.1.2 Physical approach

Two rainfall-runoff models have been proposed in this study: one is the SIMHYD model, and the other is the HEC-HMS model. The SIMHYD model is a lumped conceptual rainfall-runoff model and has nine parameters. Several parameter optimizations are provided in this approach with various objective functions. In this study, pattern search multi-start is adopted under an objective function of the Nash-Sutcliffe efficient index. The HEC-HMS model has four components (loss, transform, baseflow and routing components) to simulate streamflow using a continuous time series of precipitation and area averaged evapotranspiration. Each component provides several mathematical methods; the loss component, for instance, consists of seven optional methods to compute the total loss in a basin. In this study, the deficit and constant, the Clark unit hydrograph, recession and lag methods are adopted in the loss, transform, baseflow and routing components respectively. The SIMHYD (HEC-HMS) model is calibrated from April 1999 to March 2003 and validated from April 2003 to March 2007 at six (12) gauging stations. The model performance is evaluated based on four efficiency indexes: (i) the deviation of volume; (ii) the correlation coefficient; (iii) the normalized root mean square error; and (iv) the Nash-Sutcliffe efficient index.

The performance of the SIMHYD model is influenced by the homogeneity of basin characteristics as the model has difficulty in preserving average monthly streamflow, especially at stations that cover large drainage areas with high variability of streamflow. Based on the four efficiency indexes, the HEC-HMS model can capture well the variability of daily streamflow. Both models perform better when simulating high streamflow rather than low streamflow. Overall, the HEC-HMS model performs better than the SIMHYD model, as per the four indexes associated with the HEC-HMS model. Hence, HEC-HMS is selected to simulate daily streamflow during the 2011- 2100 period.

The simulation of daily streamflow at the 12 gauging stations located in the Ping River Basin is done using the 2011-2100 daily rainfall at the 50 selected stations, the data of which is obtained from the multisite daily rainfall generator. The HEC-HMS model is applied in this study based on an assumption that the model parameters from calibration are valid for basin characteristics in both, recent and future, periods. The objective of simulation is to determine the effects of future climate on daily streamflow in the study basin. Based on the defined thresholds at Q_{90} and Q_{10} obtained from the frequency-duration-curve (FDC) of the average daily observations, the anomalous events (low and high streamflow) are investigated under A2 and B2 scenarios of future climate. Using $Q_{90,wet}$ (Q_{90} of daily observed discharges in the wet season) and $Q_{10,dry}$ (Q_{10} of daily observed discharges in dry season) as thresholds, seasonal anomalies (i.e. low flow in a wet season and high flow in a dry season) have also been studied.

8.1.3 Effects of future climate

Under both scenarios of future climate, the 2011-2100 rainfall in the Ping River Basin in the wet (dry) season tends to decrease (increase) by 0.11-6.16 (0.02-5.91) mm per year. The effect of future climate on rainfall in the wet season indicates more chances of dryness than wetness. However, the influence of changing climate on 2011-2100 rainfall in the dry season indicates a higher chance of wetness and a lower chance of dryness. Under A2, the delay in the monsoon season by two seasons (i.e. from ASO to FMA) is also an effect of future climate. Under B2, the 1-season delay of monsoon season (i.e. from ASO to NDJ) will be observed.

The magnitudes of Q_{90} of simulated streamflow under A2 and B2 at some stations are higher than those suggested by observations; however, the magnitudes of Q_{10} of simulated discharge at all stations are less than the Q_{10} of the observations. At Station P75, P67 and 061302, the dry spell from 2011 to 2100 is longer than the duration suggested by historical records, whereas at Station 061501 and P12B, streamflow below Q_{90} will not occur at all. Compared to daily observed discharges, the low streamflow due to changes in future climate will be more severe at Station P67 and 061302 with higher intensities of shortage. At Station P24A, 061302 and P14, the wet spell from 2011 to 2100 under A2 will be extended as compared to observations, whereas the wet spell under B2 at all stations will be shortened. Under both scenarios, high streamflow will not occur at Station P4A, P21 and P71. Lower intensities of abundant water are found at all stations except at Station 061302.

Under A2 and B2, the magnitudes of $Q_{90,wet}$ and $Q_{10,dry}$ of the 2011-2100 simulated streamflow at all stations are lower than those of the observed streamflow. A shorter period of dry spell in the wet season will be observed at 12 stations with the exception of Station P75 and P21. The severity of shortage at all stations will decrease compared to historical records. Anomalous high flow in the dry season under A2 (B2) will be observed at four (three) stations, with a shorter period of wet spell and lesser intensity of abundance.

Therefore, from this study, the annual and decadal variability of hydroclimate such as temperature and rainfall were more investigated. The statistical relationships between rainfall and large-scale atmospheric variables enhanced the understanding of influences of the atmospheric circulations. The development of long-lead links aimed to identify predictors, which could be used in a forecasting model. The assessment of effects of the future changing climate on 2011-2100 rainfall in the study basin was clearly quantified by the linear trends of seasonal rainfall and the probabilities of occurrence of the anomalous weather events. The results will be a useful tool of decision making on the anomalous situations. The effects of future climate on 2011-2100 daily streamflow were also determined by the dry and wet spells with the availability of runoff. The results ultimately extended the understanding in influences of changing climate following the IPCC emission scenarios.

8.2 Recommendations

8.2.1 Adaptation and strategic plans

Since rainfall and streamflow in the Ping River Basin are influenced by an anomalous atmospheric condition that will change severely in the future, an adaptation and sustainable strategy has to be planned for water-related activities in order to deal with a pending water crisis. The analysis results show variability of rainfall and streamflow in terms of an uncertainty of water availability, a shift in the onset of monsoon season and anomalous hydrological events. Aiming to decrease tangible and intangible damages from an anomalous weather event (dry or wet), governmental departments such as the Royal Irrigation Department of Thailand, the National Disaster Warning Center and the Electricity Authority of Thailand should set up an initial long-term plan to manage water resources and water storage facilities in this basin including water allocation for all uses as the supply adaptation. Although the priority of water allocation in the Ping River Basin under a normal situation is the irrigation and hydropower, to deal with a dry condition in the future, the management of irrigated water and reallocation for other uses, in particular domestic demand will be required with a limited portion of withdrawn water for each use. Under normal and anomalous conditions, the water rights of all uses in water withdrawal from the resources should be initially conducted to meet their demand. To reduce the cost of water supply and treatment that is subsidized by the government and to raise the public awareness of water crisis, the water pricing policies and the compensation schemes should be also developed.

In terms of demand adaptation, agricultural practices need to be altered to water availability and the changing onset period of the monsoon season. The low water use crops such as onion and lettuce should be introduced to farmers in order to decrease water scarcity. However, the selection of alternative crops depends on soil type, climate in each season, technical experience of farmers, cost of the alternative crops and market opportunity, which the Ministry of Agriculture and Cooperatives has to provide and assist farmers in the essential information. The efficient agricultural practices such as drip irrigation and recycled water should be utilized to increase the irrigation efficiency. Industries and households could be encouraged to save water with reuse and recycle technologies. Moreover, the policies and plans on reduction of the water loss and leakage such as lining canals and pipe maintenance are required.

On the other hand, the mitigation strategy for a wet situation – including a warning system before the occurrence, an evacuation plan during the occurrence and an insurance policy to cover damages after the occurrence – should be addressed in the governance systems. The climate monitoring and reliable forecasting are one measure of the adaptation. The upgraded equipments of existing monitoring system, improved techniques of a forecasting model and enhanced understanding of staffs will help to deal with future changing climate. Future development of water resources and reservoirs needs to be made keeping in mind the multiple purposes of water storage during dry and wet events. The participation of stakeholders can reduce a conflict in the development of adaptation and mitigation plans. However, all defined measures need to be more studied.

8.2.2 Extension of the research

In terms of spatial coverage, this study was done only for the Ping River Basin. The extent of this dissertation should cover the investigation of all river basins in Thailand and aim to initiate an integrated national water management and development plan. This study also suggests applying the statistical method in other basins outside Thailand, which can confirm the model performance in capturing variability of hydroclimates and defining the model limitations. In terms of temporal coverage, this dissertation provided the investigation of statistical links between rainfall and large-scale atmospheric variables to identify predictors and ultimately develop a forecasting model of seasonal rainfall (i.e. MJJ, ASO, NDJ and FMA). A study of relationships between rainfall and atmospheric information in a finer time scale, e.g. monthly and daily is recommended aiming to identify predictors for a statistical model to forecast monthly or daily rainfall. However, there are some limitations of predictor identification in a finer time scale as follows:

(i) The correlation maps used to identify predictors in this study are the online analysis provided by ESRL, which are available in monthly. The daily analysis is not applicable.

(ii) The predictors of large-scale atmospheric variables may not be identified due to any significant relationships from correlation analysis in the finer time scale, particular in dry season (i.e. from November to March).

Thus, the extension of research should propose a new methodology of predictor identification in a finer time scale. Not only the predictor identification but also the forecasting model should be developed to suit the temporal coverage.

The large-scale atmospheric variables and rainfall in the study basin should have to be updated in the modified k-nn model using observed data from the meteorological stations or results from a GCM following the most recent IPCC report. Subsequently, the evaluation of updated model is suggested to clarify the improving in model performance. The changes in effects of future climate on rainfall should also be determined. Furthermore, the future research should cooperate a selected downscaling and bias correction method of the atmospheric predictors with the modified k-nn model. The evaluation and comparison of model performance will indicate a better application of simulated GCM data. A part of future study should also cover the investigation of statistical relationships between streamflow and large-scale atmospheric variables to understand the variability of streamflow under the anomalous conditions of oceanic-atmospheric circulations. The long-lead significant links will be useful in developing a model to forecast streamflow with a set of large-scale atmospheric predictors.

The effects on water-related components on natural environment, agricultural production, economic growth and living standards due to the variability of precipitation and streamflow should be determined and should also be part of future research. The studies of adaptation and mitigation (e.g. alternative crops, recycled and reused water, and water allocation) for the vulnerability of climate are suggested to address the appropriate measures for the Ping River Basin, which will deal with anomalous weather events in the future. Even without changing climate, the measures are still beneficial in the water demand and supply management. Lastly, a pilot project of selected adaptation strategies should be implemented and evaluated to account the efficiency of the research.

References and Bibliographies

An, S. I., Kug, J. S., Timmermann, A., Kang, I. S., & Timm, O. (2007). The influence of ENSO on the generation of decadal variability in the North Pacific. *J. Climate, 20*(4), 667-680.

Anderson, B. (2007). Intraseasonal atmospheric variability in the extratropics and its relation to the onset of tropical Pacific sea surface temperature anomalies. *J. Climate, 20*(5), 926-936.

Arnell, N. W. (1999). Climate change and global water resources. *Global Environ. Chang., 9*(Supplement 1), S31-S49.

Arnell, N. W. (2004). Climate change and global water resources: SRES emissions and socio-economic scenarios. *Global Environ. Chang., 14*, 31-52.

Basson, M. S., & Rooyen, J. A. (2001). Practical application of probabilistic approaches to the management of water resource systems. *J. Hydrol., 241*(1-2), 53-61.

Berri, G. S., Ghietto, M. A., & García, N. O. (2002). The influence of ENSO in the flows of the upper Paraná River of South America over the past 100 years. *J. Hydrometeorol., 3*(1), 57-65.

Bhaskaran, B., & Mitchell, J. F. B. (1998). Simulated changes in Southeast Asian monsoon precipitation resulting from anthropogenic emission. *Int. J. Climatol., 18,* 1455-1462.

Boer, R., & Faqih, A. (2003). Current and future rainfall variability in Indonesia. *AIACC semi-annual report*. Retrieved April, 2009, from http://sedac.ciesin.columbia.edu/aiacc.

Camilloni, I., & V.R., B. (2003). Extreme discharge events in the Paraná River and their climate forcing. *J. Hydrol., 278*(1-4), 94-106.

Cañón, J., González, J., & Valdés, J. (2007). Precipitation in the Colorado River basin and its low frequency associated with PDO and ENSO signal. *J. Hydrol., 333*(2-4), 252-264.

Chandimala, J., & Zubair, L. (2007). Predictability of stream flow and rainfall based on ENSO for water resources management in Sri Lanka. *J. Hydrol., 335*(3-4), 303-312.

Chang, C.-P., Wang, Z., McBride, J., & Liu, C.H. (2005). Annual cycle of Southeast Asia-Maritime continent rainfall and the asymmetric monsoon transition. *J. Climate, 18*(2), 287-301.

Chankaew, K. (1996). *Principle of watershed management* [in Thai]. Bangkok: Kasetsart University.

Chen, T.-C., & Yoon, J.-H. (2000). Interannual variation in Indochina summer monsoon rainfall: Possible mechanism. *J. Climate, 13*(11), 1979-1986.

Chiew, F. H. S., & McMahon, T. A. (2002). Global ENSO-streamflow teleconnection, streamflow forecasting and interannual variability. *Hydrolog. Sci. J., 47*(3), 505-522.

Chiew, F. H. S., Pitman, A. J., & McMahon, T. A. (1996). Conceptual catchment scale rainfall-runoff models and AGCM land-surface parameterisation schemes. *J. Hydrol., 179*, 137-157.

Chiew, F. H. S., & Siriwardena, L. (2005). Estimation of SIMHYD parameter values for application in ungauged catchments. *MODSIM05* (Vol. on CD). MSSANZ.

Chou, C., Tu, J. Y., & Yu, J. Y. (2003). Interannual variability of the western North Pacific summer monsoon: Differences between ENSO and non-ENSO years. *J. Climate, 16*(13), 2275-2287.

Chu, P. C., Edmons, N. L., & C., F. (1999). Dynamical mechanisms for the South China Sea seasonal circulation and thermohaline variabilities. *J. Phys. Oceanogr., 29*(11), 2971-2989.

Clark, M. P., Serreze, M. C., & McCabe, G. J. (2001). Historical effects of El Niño and La Niña events on the seasonal evolution of the montane snowpack in the Columbia and Colorado River basins. *Water Resour. Res., 37*(3), 741-757.

CPC (Climate Prediction Center) (2007). The tropical Pacific Ocean SST indices. Retrieved December, 2007, from http://www.cpc.noaa.gov/data/indices/.

Croke, B. F. W., Merritt, W. S., & Jakeman, A. J. (2004). A dynamic model for predicting hydrologic response to land cover changes in gauged and ungauged catchments. *J. Hydrol., 291*(1-2), 115-131.

Curtis, S., & Hastenrath, S. (1995). Forcing of anomalous sea surface temperature evolution in the tropical Atlantic during Pacific warm events. *J. Geophys. Res., 100*(C8), 15835-15847.

Delworth, T. L., Stouffer, R. J., Dixon, K. W., Spelman, M. J., Knutson, T. R., Broccoli, A. J., Kushner, P. J., & Wetherald, R. T. (2002). Review of simulations of climate variability and change with the GFDL R30 coupled climate model. *Climate Dynamics*, *19*(7), 555-574.

Dettinger, M. D., Cayan, D. R., McCabe, G. J., & Marengo, J. A. (2000). *Multiscale streamflow variability associated with El Niño/Southern Oscillation*. Cambridge: University Press.

Edossa, D. C., Babel, M. S., & Gupta, A. D. (2010). Drought analysis in the Awash River Basin, Ethiopia. *Water Resour. Manag., 24*(7), 1441-1460.

Eldaw, A. K., Salas, J. D., & Garcia, L. A. (2003). Long-range forecasting of the Nile River flows using climatic forcing. *J. Appl. Meteorol., 42*(7), 890-904.

Elsner, J. B., Tsonis, A. A., & Jagger, T. H. (2006). High-frequency variability in hurricane power dissipation and its relationship to global temperature. *B. Am. Meteorol. Soc., 87*(6), 763-768.

Emanuel, K. A. (2005). Increasing destructiveness of tropical cyclones over the pass 30 years. *Nature, 436*, doi:10.1038/nature03906 Letter, 686-688.

Enfield, D. B., & Alfaro, E. J. (1999). The dependence of Caribbean rainfall on the interaction of the tropical Atlantic and Pacific Oceans. *J. Climate, 12*(7), 2093-2103.

Enfield, D. B., & Mayer, D. A. (1997). Tropical Atlantic sea surface temperature variability and its relation to El Niño-Southern Oscillation. *J. Geophys. Res., 102*(C1), 929-945.

ESRL (Earth System Research Laboratory) (2008). Interactive plotting and analysis: Linear monthly/seasonal correlations. Retrieved September, 2008, from http://www.esrl.noaa.gov/psd/data/correlation/.

Fasullo, J., & Webster, P. J. (2002). Hydrological signatures relating the Asian summer monsoon and ENSO. *J. Climate, 15*(21), 3082-3095.

Fasullo, J., & Webster, P. J. (2003). A hydrological definition of Indian monsoon onset and withdrawal. *J. Climate, 16*(19), 3200-3211.

Fleig, A. K., Tallaksen, L. M., Hisdal, H., & Demuth, S. (2006). A global evaluation of streamflow drought characteristics. *Hydrol. Earth Syst. Sci., 10*, 535-552.

Frankignoul, C., & Sennéchael, N. (2007). Observed influence of North Pacific SST anomalies on the atmospheric circulation. *J. Climate*, *20*(3), 592-606.

Frederiksen, J. S., & Branstator, G. (2005). Seasonal variability of teleconnection patterns. *J. Atmos. Sci., 62*(5), 1346-1365.

Garreaud, R. D., & Battisti, D. (1999). Interannual (ENSO) and interdecadal (ENSO-like) variability in the southern hemisphere tropospheric circulation. *J. Climate, 12*(7), 2113-2123.

Gershunov, A. (1998). ENSO influence on intraseasonal extreme rainfall and temperature frequencies in the contiguous United States: Implications for long-range predictability. *J. Climate, 11*(12), 3192-3203.

Giannini, A., Chiang, J. C. H., Cane, M. A., Kushnir, Y., & Seager, R. (2001). The ENSO teleconnection to the tropical Atlantic Ocean: Contributions of the remote and local SSTs to rainfall variability in the tropical Americas. *J. Climate, 14*(24), 4530-4544.

Giannini, A., Kushnir, Y., & M.A., C. (2000). Interannual variability of Caribbean rainfall, ENSO, and the Atlantic Ocean. *J. Climate, 13*(2), 297-311.

Glantz, M. H. (2001). *Currents of change: Impacts of El Niño and La Niña on climate and society* (2nd ed.). Cambridge: University Press.

Gong, D. Y., & Wang, S. W. (1999). Impacts of ENSO on rainfall of global land and China. *Chinese Sci. Bull., 44*(9), 852-857.

Goswami, B. N., Wu, G., & Yasunari, T. (2006). The annual cycle, intraseasonal oscillation, and roadblock to seasonal predictability of the Asian summer monsoon. *J. Climate*, *19*(20), 5078-5099.

Grantz, K., Rajagopalan, B., Clark, M., & Zagona, E. (2005). A technique for incorporating large-scale climate information in basin-scale ensemble streamflow forecasts. *Water Resour. Res., 41*(W10410), doi:10.1029/2004WR003467.

Gupta, R. S. (1989). *Hydrology & hydraulic systems*. New Jersey: Prentice-Hall, Inc.

Gutiérrez, F., & Dracup, J. A. (2001). An analysis of the feasibility of long-range streamflow forecasting for Colombia using El Niño-Southern Oscillation indicators. *J. Hydrol., 246*(1-4), 181-196.

Gutman, G., Csiszar, I., & Romanov, P. (2000). Using NOAA/AVHRR products to monitor El Niño impacts: Focus on Indonesia in 1997-98. *B. Am. Meteorol. Soc., 81*(6), 1189-1205.

Haan, C. T. (2002). *Statistical method in hydrology* (2nd ed.). Iowa: Iowa State Press.

Hamlet, A. F., Huppert, D., & Lettenmaier, D. P. (2002). Economic value of long-lead streamflow forecasts for Columbia River hydropower. *J. Water Res. Pl.-ASCE, 128*(2), 91-101.

Hansen, J., Ruedy, R., Sato, M., & Lo, K. (2010). Global surface temperature change. Retrieved January, 2011, from http://data.giss.nasa.gov/gistemp/paper/gistemp2010_draft0803.pdf.

Harshburger, B., Ye, H., & Dzialoski, J. (2002). Observation evidence of the influence of Pacific SSTs on winter precipitation and spring stream discharge in Idaho. *J. Hydrol., 246*(1-4), 157-169.

Ho, L., & Wang, B. (2002). The time-space structure of the Asian-Pacific summer monsoon: A fast annual cycle view. *J. Climate*, *15*(15), 2001-2019.

Hu, Q., & Feng, S. (2001). Variations of teleconnection of ENSO and interannual variation in summer rainfall in the central United States. *J. Climate, 14*(11), 2469-2480.

Hu, Y., Maskey, S., Uhlenbrook, S., & Zhao, H. (2011). Streamflow trends and climate linkages in the source region of the Yellow River, China. *Hydrol. Process., 25*(22), 3399-3411.

IPCC-DDC (IPCC Data Distribution Center) (2009). Retrieved February, 2009, from http://www.ipcc-data.org.

IPCC. (2001). *Climate change 2001: Impacts, adaptation and vulnerability*. Cambridge: Cambridge University Press.

IPCC (2007a). *Climate Change 2007: Synthesis Report*. Geneva: IPCC.

IPCC (2007b). IPCC fourth assessment report: climate change 2007, Global average temperatures. Retrieved January, 2011, from http://www.ipcc.ch/publications_and_data/ar4/wg1/en/tssts-3-1-1.html.

Jain, S., & Lall, U. (2001). Floods in a changing climate: Does the past represent the future? *Water Resour. Res.*, *37*(12), 3193-3250.

Jenkins, G. J., Perry, M. C., & Prior, M. J. (2008). *The climate of the United Kingdom and recent trends*. Exeter: Met Office Hadley Center.

Jiang, T., Zhang, Q., Zhu, D. M., & Wu, Y. J. (2006). Yangtze floods and droughts (China) and teleconnections with ENSO activities (1470-2003). *Quatern. Int.*, *144*(1), 29-37.

Kanae, S., Oki, T., & Musiake, K. (2001). Impact of deforestation on regional precipitation over the Indochina Peninsula. *J. Hydrometeorol.*, *2*(1), 51-70.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W. et al. (1996). The NCEP/NCAR reanalysis 40-year project. *Bull. Am. Meteorol. Soc.* 77, 437–471.

Keenlyside, N. S., & Latif, M. (2007). Understanding equatorial Atlantic interannual variability. *J. Climate*, *20*(1), 131-142.

Kim, T. W., Valdés, J. B., Nijssen, B., & Roncayolo, D. (2006). Quantification of linkages between large-scale climatic patterns and precipitation in the Colorado River basin. *J. Hydrol.*, *321*(1-4), 173-186.

Krause, P., Boyle, D. P., & Bäse, F. (2005). Comparison of different efficiency criteria for hydrological model assessment. *Adv. Geosci.*, *5*, 89-97.

Krishna Kumar, K., Soman, M. K., & Rupakumar, K. (1995). Seasonal forecasting of Indian summer monsoon rainfall: a review. *Weather*, *50*, 449-467.

Krishnamurthy, V., & Shukla, J. (2007). Intraseasonal and seasonal persisting patterns of Indian monsoon rainfall. *J. Climate*, *20*(1), 3-20.

Krishnan, R., Mujumdar, M., Vaidya, V., Ramesh, K. V., & Satya, V. (2003). The abnormal Indian summer monsoon of 2000. *J. Climate*, *16*(8), 1177-1194.

Krishnan, R., Zhang, C., & Sugi, M. (2000). Dynamics of breaks in the Indian summer monsoon. *J. Atmos. Sci.*, *57*(9), 1354-1372.

Krzysztofowicz, R. (2001). The case for probabilistic forecasting in hydrology. *J. Hydrol.*, *249*(1-4), 2-9.

LDD (Land Development Department) (2011). Retrieved March, 2011, from http://oss101.ldd.go.th/web_thaisoils/62_soilgroup/main_62soilgroup.htm.

Lévesque, É., Anctil, F., Griensven, A. V., & Beauchamp, N. (2008). Evaluation of streamflow simulation by SWAT model for two small watersheds under snowmelt and rainfall. *Hydrolog. Sci. J.*, *53*(5), 961-976.

Li, S., Robinson, W. A., Hoerling, M. P., & Weickmann, K. M. (2007). Dynamics of the extratropical response to a tropical Atlantic SST anomaly. *J. Climate*, *20*(3), 560-574.

Linsley, R. K. Jr., Kohler, M. A., & Paulhus, J. L. H. (1988). *Hydrology for engineers* (SI metric ed.). Singapore: McGraw-Hill Book Co.

Liu, Y., Chan, J. C. L., Mao, J., & Wu, G. (2002). The role of Bay of Bengal convection in the onset of 1998 South China Sea summer monsoon. *Mon. Weather Rev.*, *130*(11), 2731-2744.

Liu, Z., & Yang, H. (2001). Regional dynamics of seasonal variability in the South China Sea. *J. Phys. Oceanogr.*, *31*(1), 272-284.

Loader, C. (1999). *Local regression and likelihood*. New York: Springer.

Mantua, N. J., Hare, S. R., Zhang, Y., Wallace, J. M., & Francis, R. C. (1997). A Pacific interdecadal climate oscillation with impacts on salmon production. *B. Am. Meteorol. Soc.*, *78*(6), 1069-1079.

Masih, I., Uhlenbrook, S., Maskey, S., & Smakhtin, V. (2011). Streamflow trends and climate linkages in the Zagros Mountains, Iran. *Climatic Change*, *104*(2), 317-338.

Masiokas, M. H., Villalba, R., Luckman, B. H., Quesne, C. L., & Aravene, J. C. (2006). Snowpack variations in the central Andes of Argentina and Chile, 1951-2005: Large-scale atmospheric influences and implications for water resources in the region. *J. Climate, 19*(24), 6334-6352.

Mason, S. J., & Goddard, L. (2001). Probabilistic precipitation anomalies associated with ENSO. *B. Am. Meteorol. Soc.*, *82*, 619-638.

McCabe, G. J., & Dettinger, M. D. (2002). Primary modes and predictability of year-to-year snowpack variation in the western United States from teleconnections with Pacific Ocean climate. *J. Hydrometeorol., 3*(1), 13-25.

Meehl, G. A., & Arblaster, J. M. (2002). Indian monsoon GCM sensitivity experiments testing tropospheric biennial oscillation transition conditions. *J. Climate, 15*(9), 923-944.

Meko, D. M., & Woodhouse, C. A. (2005). Tree-ring footprint of joint hydrologic drought in Sacramento and upper Colorado river basins, western USA. *J. Hydrol., 308*(1-4), 196-213.

Mendoza, B., Jáuregui, E., Diaz-Sandoval, R., García-Acosta, V., Velasco, V., & Cordero, G. (2005). Historical droughts in central Mexico and their relation with El Niño. *J. Appl. Meteorol., 44*(5), 709-716.

Mestas-Nuñez, A. M., & Enfield, D. B. (2001). Eastern equatorial Pacific SST variability: ENSO and non-ENSO components and their climatic associations. *J. Climate, 14*(3), 391-402.

Michalczyk, A. (2008). *Parameterization and modelling of growth and yield development of mango (Mangifera indica L.) in north Thailand with application of the WaNul CAS model*. (M.S. Humboldt -Universität zu Berlin, 2008). Berlin: Humboldt - Universität zu Berlin.

Nagura, M., & Konda, M. (2007). The seasonal development of an SST anomaly in the Indian Ocean and its relationship to ENSO. *J. Climate, 20*(1), 38-52.

NIC (National Intelligence Council) (2009). Southeast Asia and Pacific Islands: The impact of climate change 2030. *Spec Rep NIC2009-006D*. Washington DC.

Nicholls, N. (1983). Predicting Indian monsoon rainfall from sea surface temperature in the Indonesia-north Australia area. *Nature, 306*, 576-577.

Nodzu, M. I., Ogino, S. Y., Tachibana, Y., & Yamanaka, M. D. (2006). Climatological description of seasonal variations in lower-tropospheric temperature inversion layers over the Indochina peninsula. *J. Climate, 19*(13), 3307-3319.

NSO (National Statistically Office, Bangkok, Thailand) (2005). Retrieved April, 2008, from http://service.nso.go.th/nso/thailand/dataFile/01/J01W/J01W/th/0.htm.

Ogino, S. Y., Sato, K., Yamanaka, M. D., & Watanabe, A. (2006). Lower-stratospheric and upper-tropospheric disturbances observed by radiosondes over Thailand during January 2000. *J. Atmos. Sci., 63*(12), 3437-3447.

Orlanski, I. (2005). A new look at the Pacific storm track variability: sensitivity to tropical SSTs and to upstream seeding. *J. Atmos. Sci., 62*(5), 1367-1390.

Owosina, A. (1992). *Methods for assessing the space and time variability of ground water data.* (M.S. Utah State University, 1992). Utah: Utah State University.

Pavia, E. G., Graef, F., & Reyes, J. (2006). PDO-ENSO effects in the climate of Mexico. *J. Climate, 19*(24), 6433-6438.

Podger, G. (2004). *User guide RRL (Rainfall Runoff Library)*. Australia: CRC for Catchment Hydrology.

PSD (Physical Sciences Division) (2007a). Retrieved December, 2007, from http://www.esrl.noaa.gov/psd//cgi-bin/data/timeseries/timeseries1.pl.

PSD (Physical Sciences Division) (2007b). The Indian Ocean SST index. Retrieved December, 2007, from http://www.esrl.noaa.gov/psd/forecasts/sstlim/timeseries/index.html.

Rajagopalan, B., & Lall, U. (1999). A k-nearest-neighbor simulation for daily precipitation and other weather variables. *Water Resour. Res., 35*(10), 3089-3101.

Rasmusson, E. M., & Carpenter, T. (1983). The relationship between Eastern Equatorial Pacific sea surface temperature and rainfall over India and Sri Lanka. *Mon. Weather Rev., 111*(3), 517-528.

Rasmusson, E. M., & Carpenter, T. H. (1982). Variations in tropical sea surface temperature and surface wind fields associated with the Southern Oscillation/El Niño. *Mon. Weather Rev., 110*(5), 354-384.

Ropelewski, C. F., & Halpert, M. S. (1986). North American precipitation and temperature patterns associated with the El Niño/Southern Oscillation (ENSO). *Mon. Weather Rev., 114*(12), 2352-2362.

Ropelewski, C. F., & Halpert, M. S. (1987). Global and regional scale precipitation patterns associated with the El Niño/Southern Oscillation. *Mon. Weather Rev., 115*(8), 1606-1626.

Sahai, A. K., Grimm, A. M., Satyan, V. & Pant, G. B. (2003). Long-lead prediction of Indian summer monsoon rainfall from global SST evolution. *Climate Dynamics, 20*, 855-863.

Saravanan, R., & Chang, P. (2000). Interaction between tropical Atlantic variability and El Niño-Southern Oscillation. *J. Climate, 13*(13), 2177-2194.

Schöngart, J., & Junk, W. J. (2007). Forecasting the flood-pulse in Central Amazonia by ENSO-indices. *J. Hydrol., 335*(1-2), 124-132.

Shrestha, A., & Kostaschuk, R. (2005). El Niño/Southern Oscillation (ENSO)-related variability in mean-monthly streamflow in Nepal. *J. Hydrol., 308*(1-4), 33-49.

Shrestha, M. L. (2000). Interannual variation of summer monsoon rainfall over Nepal and its relation to Southern Oscillation index. *Meteorol. Atmos. Phys., 75*(1-2), 21-28.

Silverman, D., & Dracup, J. P. (2000). Artificial neural networks and long-range precipitation prediction in California. *J. Appl. Meteorol., 39*(1), 57-66.

Singhrattna, N., Rajagopalan, B., Clark, M., & Kumar, K. K. (2005a). Forecasting Thailand summer monsoon rainfall. *Int. J. Climatol., 25*(5), 649-664.

Singhrattna, N., Rajagopalan, B., Kumar, K. K., & Clark, M. (2005b). Interannual and interdecadal variability of Thailand summer monsoon season. *J. Climate, 18*(11), 1697-1708.

Smith, S. R., & O'Brien, J. J. (2001). Regional snowfall distributions associated with ENSO: Implications for seasonal forecasting. *B. Am. Meteorol. Soc., 82*(6), 1179-1191.

Sombroek, W. (2001). Spatial and temporal patterns of Amazon rainfall consequences for the planning of agricultural occupation and the protection of primary forests. *Ambio, 30*(7), 388-396.

Sourza Filho, F. A., & Lall, U. (2003). Seasonal to interannual ensemble streamflow forecasts for Cearra, Brazil: Applications of a multivariate semiparametric algorithm. *Water Resour. Res., 39*(11), doi:10.1029/2002WR001373.

Steinemann, A. C. (2006). Using climate forecasts for drought management. *J. Appl. Meteorol. Clim., 45*(10), 1353-1361.

Takahashi, H. G., & Yasunari, T. (2006). A climatological monsoon break in rainfall over Indochina—A singularity in the seasonal March of the Asian Summer Monsoon. *J. Climate, 19*(8), 1545-1556.

Tereshchenko, I., Filonov, A., Gallesgos, A., Monzón, C., & Rodríguez, R. (2002). El Niño 1997-98 and the hydrometeorological variability of Chapala, a shallow tropical lake in Mexico. *J. Hydrol., 264*(1-4), 133-146.

TMD (Thailand Meteorology Department) (2007). The natural disasters in Thailand (in Thai). Retrieved January, 2011, from http://www.tmd.go.th/info/risk.pdf.

Trenberth, K. E., Jones, P. D., Ambenje, P., Bojariu, R., Easterling, D., Tank, A. K., et al. (2007). *Observations: Surface and atmospheric climate change*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA: IPCC.

Trenberth, K. E., Moore, B., Karl, T. R., & Carlos, N. (2006). Monitoring and prediction of earth's climate: A future perspective. *J. Climate, 19*(20), 5001-5008.

U.S. Army Corps of Engineers (2000). *Hydrologic modeling system HEC-HMS: Technical reference manual*. California: Hydrologic Engineering Center.

U.S. Army Corps of Engineers (2010). *Hydrologic modeling system HEC-HMS: User's manual*. California: Institute for Water Resources.

Webster, P. J., Holland G.J., Curry, J.A., & Chang, H.R. (2005). Changes in tropical cyclone number, duration, and intensity in a warming environment. *Science, 309*, doi:10.1126/science.1116448, 1844-1846.

Whitaker, D. W., Wasimi, S. A., & Islam, S. (2001). The El Niño Southern Oscillation and long-range forecasting of the flows in the Ganges. *Int. J. Climatol., 21*(1), 77-87.

Wood, A. W., Maurer, E. P., Kumar, A., & Lettenmaier, D. (2002). Long-range experimental hydrologic forecasting for the eastern United States. *J. Geophys. Res., 107,* doi:10.1029/2001JD000659.

Woodruff, S. D., Lubker, S. J., Wolter, K., Worley, S. J., & Elms, J. D. (1993). Comprehensive Ocean-Atmosphere Data Set (COADS) Release 1a: 1980-92. *Earth System Monitor, 4*(1), 1-8.

Wu, B., & Wang, B. (2004). Assessing impacts of global warming on tropical cyclone tracks. *J. Clim., 17*(8), 1686-1698.

Wu, L., He, F., Lui, Z., & Li, C. (2007). Atmospheric teleconnections of tropical Atlantic variability: Interhemispheric, tropical-extratropical, and cross-basin interactions. *J. Climate, 20*(5), 856-870.

Wu, R., & Kirtman, B. P. (2004). Understanding the impacts of the Indian Ocean on ENSO variability in a coupled GCM. *J. Climate, 17*(20), 4019-4031.

Yang, S., Ding, X., Zheng, D., & Li, Q. (2007). Depiction of the variations of Great Plains precipitation and its relationship with tropical central-eastern Pacific SST. *J. Appl. Meteorol. Clim., 46*(2), 136-153.

Yaremchuk, M., & Qu, T. (2004). Seasonal variability of the large-scale currents near the coast of Philippines. *J. Phys. Oceanogr., 34*(4), 844-855.

Yokoi, S., & Satomura, T. (2005). An observational study of intraseasonal variations over Southeast Asia during the 1998 rainy season. *Mon. Weather Rev., 133*(7), 2091-2104.

Zehe, E., Singh, A. K., & Bárdossy, A. (2006). Modelling of monsoon rainfall for a mesoscale catchment in North-West India I: assessment of objective circulation patterns. *Hydrol. Earth Syst. Sci., 10*, 797-806.

Zhang, Q., Xu, C., Jiang, T., & Wu, Y. (2007). Possible influence of ENSO on annual maximum streamflow of the Yangtze River, China. *J. Hydrol., 333*(2-4), 265-274.

Zhang, Y., Wallace, J. M., & Battisti, D. S. (1997). ENSO-like interdecadal variability: 1990-93. *J. Climate, 10*(5), 1004-1020.

Zubair, L. (2003a). El Niño-Southern Oscillation influences on Mahaweli streamflow in Sri Lanka. *Int. J. Climatol., 23*(1), 91-102.

Zubair, L. (2003b). Sensitivity of Kelani streamflow in Sri Lanka to ENSO. *Hydrol. Process., 17*(12), 2439-2448.

Appendices

Appendix A: Lists of Meteorological Stations

A1: List of the rainfall stations located in and around the Ping River Basin

No.	Station Code	Province	Location		Period				Length (years)	Incomplete Data (days)	Data Source	Period of Incomplete Data
			Latitude	Longitude	From		To					
1	327301	CM	18.91667	99.00000	JAN	1969	DEC	2007	39	0	TMD	Sep 2007
2	327501	CM	18.79000	98.97694	JAN	1951	DEC	2007	57	0	TMD	
3	07013	CM	18.83972	98.97556	APR	1952	MAR	2006	55	153	RID	1 Jul-30 Nov 2004
4	07022	CM	18.71333	99.04139	APR	1952	MAR	2006	55	1745	RID	1-30 Jun 1956, 1 Sep-31 Oct 1957,1 May-31 Jul 1960, 1-31 May 1961, 1-31 Jul 1961, 1-31 Oct 1961, 1-31 May 1962, 1 May-30 Jun 1963, 1-31 May 1964, 1-31 Oct 1964, 1-31 May 1965, 1-31 Oct 1967, 1-31 Oct 1973, 1-31 May 1977, 1 Aug-31 Oct 1982, 1983-1984, 1-31 Oct 1987, 1 May-31 Jul 1991, 1 Sep-31 Oct 1991, 1 May-30 Jun 1992, 1-31 Jul 1993, 1-31 Oct 1994, 1 May-30 Jun 1997
5	07032	CM	18.74417	99.12444	APR	1952	MAR	2006	55	1316	RID	1-31 May 1958, 1 Sep-31 Oct 1962, 1-31 May 1963, 1 May-31 Jul 1968, 1 Aug-31 Oct 1973, 1974-1975, 1 May-31 Oct 1976, 1-31 Oct 1981, 1 Aug-31 Oct 1994, 1 Apr-30 Nov 1996, 1 May-30 Jun 1997, 1 Jun-30 Nov 1999, 1 Jul-30 Nov 2001, 1 May-30 Jun 2003
6	07042	CM	18.84750	99.04833	APR	1952	MAR	2006	55	1523	RID	1979-1982, 1-31 May 1991, 1-31 May 1992
7	07052	CM	18.86889	99.13944	APR	1952	MAR	2006	55	181	RID	1-30 Nov 1997, 1-30 Apr 1998, 1 Sep-30 Nov 2003, 1-30 Nov 2005
8	07062	CM	18.91306	99.94778	APR	1952	MAR	2006	55	977	RID	1-31 May 1991, 1-31 Oct 1993, 1-30 Sep 1994, 1 Aug-30 Nov 1996, 1 May-30 Jun 1997, 1 Oct-30 Nov 1997, 1-30 Jun 1999, 1 Aug-30 Nov 1999, 1 May-31 Aug 2001, 1 Oct-30 Nov 2002, 1 Oct-30 Nov,2003, 1 Oct-30 Nov 2004, 1 May-31 Jul 2005, 1 Sep-30 Nov 2005
9	07072	CM	18.68611	98.92194	APR	1952	MAR	2005	54	641	RID	1 May-31 Oct 1974, 1 Jun-30 Nov 1997, 1 Jun-30 Nov 1999, 1 Sep-30 Nov 2004
10	07082	CM	18.62694	98.89889	APR	1952	MAR	2006	55	548	RID	1-31 Oct 1965, 1-30 Sep 1966, 1-31 May 1972, 1-30 Jun 1975, 1980, 1-30 Nov 2003, 1-30 Nov 2005
11	07092	CM	18.19056	98.61444	APR	1952	MAR	2004	53	336	RID	1-31 May 2001, 1 Aug-30 Nov 2001, Apr 2003, 1 Jul-30 Nov 2003
12	07102	CM	19.91722	99.21667	APR	1952	MAR	2006	55	1219	RID	1-31 May 1965, 1-31 Oct 1970, 1973, 1 May-30 Jun 1974, 1984-1985
13	07112	CM	19.11889	98.94778	APR	1952	MAR	2006	55	487	RID	1-31 Oct 1977, 1-31 Oct 1979, 1 Apr-30 Jun 1997, 1 Oct-30 Nov 1997, 1-30 Nov 1999, 1 May-30 Jun 2003, 1 Sep-30 Nov 2003, 1-30 Nov 2004, 1 Jun-31 Jul 2005
14	07122	CM	19.36444	99.20472	APR	1952	MAR	2006	55	244	RID	1-31 Aug 1992, 1-30 Sep 1993, 1 Jun-30 Nov 2003
15	07132	CM	19.36472	98.96667	APR	1952	MAR	2006	55	1500	RID	1 May-31 Oct 1969, 1-31 Oct 1971, 1-31 Oct 1973, 1 Jul-30 Sep 1974, 1 Jul-31 Aug 1975, 1-31 Oct 1975, 1 Sep-31 Oct 1976, 1-31 Oct 1980, 1-31 Oct 1982, 1-31 Oct 1983, 1984-1985, 1 May-31 Oct 1986
16	07142	CM	18.84778	98.73583	APR	1952	AUG	2007	56	92	RID	1-31 Jul 1954, 1 May-30 Jun 1956

No.	Station Code	Province	Location		Period				Length (years)	Incomplete Data (days)	Data Source	Period of Incomplete Data
			Latitude	Longitude	From		To					
17	07152	CM	18.49833	98.36500	APR	1952	MAR	2006	55	429	RID	1 Sep-31 Oct 1958, 1-31 Oct 1959, 1 Jun-31 Oct 1963, 1 Sep-31 Oct 1966, 1 May-30 Jun 1968, 1-31 Oct 1976, 1-31 May 1998
18	07162	CM	17.79583	98.36000	APR	1952	AUG	2007	56	1315	RID	1 Jun-31 Oct 1953, 1 Aug-31 Oct 1954, 1-31 Oct 1955, 1956-1957, 1-30 Jun 1958, 1-31 Aug 1960, 1 Jun-31 Jul 1961, 1-31 May 1963, 1 Jul-31 Aug 1967, 1-31 Oct 1967, 1 Jul-31 Aug 1972
19	07172	CM	19.95972	99.16056	APR	1952	MAR	2006	55	93	RID	1-31 May 1952, 1-31 May 1969, 1 May-31 Aug 2003
20	07182	CM	18.41583	98.67972	APR	1952	MAR	2006	55	124	RID	1-31 May 1962, 1-31 May 1963, 1-31 Oct 1973, 1-31 Oct 1984
21	07192	CM	18.05000	98.64528	APR	1959	MAR	2005	47	887	RID	1 May-31 Aug 1959, 1 May-30 Nov 1994, 1 Jan-31 May 1995, 1-30 Nov 1995, 1 Jul-31 Aug 1997, 1-31 Oct 1998, 1 May-31 Aug 1999, 1-31 May 2003, 1 Aug-30 Nov 2004
22	07202	CM	n/a	n/a	APR	1959	MAR	1969	11	n/a	RID	< 30 years
23	07212	CM	n/a	n/a	APR	1959	MAR	1968	10	n/a	RID	< 30 years
24	07222	CM	19.85000	99.21250	APR	1959	MAR	2006	48	2411	RID	1 May-31 Oct 1959, 1961-1962, 1968, 1-31 Aug 1969, 1-31 Oct 1969, 1 May-31 Jul 1971, 1 Sep-31 Oct 1973, 1974, 1 May-30 Jun 1975, 1-31 Oct 1980, 1-31 May 1992, 1-31 Oct 1992, 1 May-31 Oct 1993, 1 Apr-30 Sep 1994, 1-30 Jun 1997
25	07232	CM	19.26000	98.92194	APR	1961	MAR	1973	13	n/a	RID	< 30 years
26	07242	CM	18.80278	98.92500	APR	1961	AUG	2007	47	123	RID	1 May-31 Aug 1961
27	07252	CM	19.26861	98.97556	APR	1964	AUG	2007	44	245	RID	1 May-30 Sep 1965, Sep-Oct 1968, 1-31 May 1975
28	07262	CM	18.80667	98.90333	APR	1965	MAR	2006	42	1524	RID	1 May-30 Sep 1965, 1 Aug-31 Oct 1968, 1969-1970, 1989, 1 May-31 Oct 1991
29	07272	CM	18.83333	98.86667	APR	1966	MAR	1978	13	n/a	RID	< 30 years
30	07282	CM	18.15028	98.39306	APR	1966	AUG	2007	42	123	RID	1 May-31 Aug 1966
31	07292	CM	18.61111	98.90056	APR	1962	MAR	2006	45	154	RID	1 May-31 Aug 1962, 1-31 Oct 1992
32	07303	CM	18.89667	99.01083	APR	1973	MAR	2006	34	1583	RID	1-30 Sep 1987, 1988-1991, 1-31 Oct 1993, 1 Sep-31 Oct 1996
33	07314	CM	18.74361	98.92222	APR	1959	MAR	1965	7	n/a	RID	< 30 years
34	07322	CM			APR	1970	MAR	1975	6	n/a	RID	< 30 years
35	07331	CM	19.21250	98.87000	APR	1952	MAR	1981	30	n/a	RID	< 30 years
36	07341	CM	18.91778	99.13056	APR	1964	MAR	2004	41	4238	RID	1 May-30 Sep 1964, 1-31 May 1972, 26 May 1981, 1-31 Oct 1990, 1 May-31 Jul 1991, 1 Sep-31 Oct 1991, 1992-1997, 1 May-31 Aug 1998, 1999-2002, 1-31 May 2003, 1 Jul-31 Aug 2003
37	07361	CM	n/a	n/a	APR	1968	MAR	1999	32	6698	RID	1980-1997, 1 May-31 Aug 1998
38	07391	CM	18.78917	99.01694	APR	1971	MAR	2006	36	215	RID	1-31 Jul 1980, 1-31 May 1997, 1-31 Aug 1999, 1-30 Jun 2000, 1 Jun-31 Aug 2002
39	07420	CM	18.99556	98.98333	APR	1952	MAR	2004	53	5309	RID	1 Jun-31 Jul 1953, 1-31 Oct 1953, 1-31 Oct 1954, 1-30 Jun 1955, 1 Jun-31 Oct 1956, 1 May-30 Jun 1957, 1 May-31 Oct 1958, 1 May-31 Oct 1959, 1 Aug-31 Oct 1964, 1-31 Aug 1965, 1 May-31 Jul 1966, 1 May-31 Oct 1968, 1-31 Oct 1979, 1 Jun-31 Jul 1990, 1 Oct-30 Nov 1990, 1 May-31 Jul 1991, 1 Sep-31 Oct 1991, 1992-1997, 1 May-31 Aug 1998, 1999-2002, 1-31 May 2003, 1 Jul-31 Aug 2003

No.	Station Code	Province	Location		Period				Length (years)	Incomplete Data (days)	Data Source	Period of Incomplete Data
			Latitude	Longitude	From		To					
40	07430	CM	18.90167	99.02056	APR	1952	MAR	2004	53	5091	RID	1 Jun-31 Jul 1953, 1-31 Oct 1953, 1-31 Oct 1954, 1-30 Jun 1955, 1 Jun-31 Oct 1956, 1 May-30 Jun 1957, 1958, 1 Aug-31 Oct 1964, 1 May-31 Jul 1966, 1-31 Oct 1966, 1-31 Oct 1980, 1 Jun-31 Jul 1990, 1 Oct-30 Nov 1990, 1 May-31 Jul 1991, 1 Sep-31 Oct 1991, 1992-1997, 1 May-31 Aug 1998, 1999-2002, 1 Jul-31 Aug 2003
41	07440	CM	19.04250	98.98111	APR	1952	MAR	2004	53	5428	RID	1 Jun-31 Jul 1953, 1-31 Oct 1953, 1-31 Oct 1954, 1-30 Jun 1955, 1 Jun-31 Oct 1956, 1 May-30 Jun 1957, 1 May-31 Oct 1958, 1 May-31 Oct 1959, 1 Aug-31 Oct 1964, 1 May-31 July 1966, 1 May-31 Oct 1968, 1 Jun-31 Jul 1990, 1-31 Oct 1990, 1 May-31 Jul 1991, 1 Sep-31 Oct 1991, 1992-2002, 1 Jul-31 Aug 2003
42	07450	CM	18.93250	99.00056	APR	1952	MAR	2004	53	5610	RID	1 Jun-31 Jul 1953, 1-31 Oct 1953, 1-31 Oct 1954, 1-30 Jun 1955, 1 Jun-31 Oct 1956, 1 May-30 Jun 1957, 1958, 1 May-31 Oct 1959, 1 Aug-31 Oct 1964, 1 May-31 July 1966, 1968, 1-31 Oct 1980, 1 Jun-31 Jul 1990, 1-31 Oct 1990, 1 May-31 Jul 1991, 1 Sep-31 Oct 1991, 1992-1997, 1 May-31 Aug 1998, 1999-2002, 1-31 May 2003, 1 Jul-31 Aug 2003
43	07460	CM	18.87778	99.08556	APR	1960	MAR	2004	45	4420	RID	1 Aug-31 Oct 1964, 1 May-31 Jul 1966, 1 Sep-31 Oct 1980, 1-31 May 1987, 1 Jun-31 Jul 1990, 1 Oct-30 Nov 1990, 1 May-31 Jul 1991, 1 Sep-31 Oct 1991, 1992-1997, 1 May-31 Aug 1998, 1999-2002, 1-31 May 2003, 1 Jul-31 Aug 2003
44	07472	CM	17.91667	98.68333	APR	1969	MAR	2006	38	549	RID	1-30 Sep 1989, 1 Jul-31 Oct 1990, 1991, 1-31 Aug 1994
45	07480	CM	19.10222	98.95583	APR	1952	MAR	2004	53	4635	RID	1 May-30 Jun 1957, 1 May-31 Oct 1958, 1 May-31 Oct 1959, 1 Aug-31 Oct 1964, 1 Jun-31 Jul 1990, 1-31 Oct 1990, 1 May-31 Jul 1991, 1 Sep-31 Oct 1991, 1992-1997, 1 May-31 Aug 1998, 1999-2002, 1-31 May 2003, 1 Jul-31 Aug 2003
46	07492	CM	19.99639	99.25917	APR	1970	MAR	2006	37	276	RID	1 May-31 Aug 1970, 1 May-30 Sep 2005
47	07502	CM	19.06667	99.21667	APR	1972	MAR	2006	35	460	RID	1 May-31 Oct 1972, 1 May-31 Aug 1984, 1 May-31 Aug 1995, 1-30 Jun 1997
48	07510	CM	18.68944	98.97222	APR	1970	MAR	2004	35	6302	RID	1-31 May 1970, 1971-1974, 1990-2002, 1 Jul-31 Aug 2003
49	07520	CM	19.15444	98.92278	APR	1974	MAR	2004	31	4599	RID	1 May-31 Oct 1974, 1-30 Jun 1977, 1-30 Nov 1977, 1 Sep-30 Nov 1990, 1 May-31 Oct 1991, 1992-2002, 1 Jul-31 Aug 2003
50	07530	CM	18.87639	99.14667	APR	1974	MAR	2004	31	3537	RID	1 May-31 Oct 1974, 1-31 Aug 1987, 1-31 Oct 1990, 1 May-31 July 1991, 1 Sep-31 Oct 1991, 1992-1997, 1 May-31 Aug 1998, 1999-2002, 1-31 May 2003, 1 Jul-31 Aug 2003
51	07540	CM	18.82139	99.17556	APR	1974	MAR	2004	31	4237	RID	1 May-31 Oct 1974, 1-31 Oct 1990, 1 May-31 July 1991, 1 Sep-31 Oct 1991, 1992-1997, 1 May-31 Aug 1998, 1999-2002, 1-31 May 2003, 1 Jul-31 Aug 2003

No.	Station Code	Province	Location		Period				Length (years)	Incomplete Data (days)	Data Source	Period of Incomplete Data
			Latitude	Longitude	From		To					
52	07550	CM	18.74056	99.16028	APR	1959	MAR	2004	46	4267	RID	1 May-31 Oct 1959, 1-30 Jun 1983, 1-31 Oct 1990, 1 May-31 July 1991, 1 Sep-31 Oct 1991, 1992-1997, 1 May-31 Aug 1998, 1999-2002, 1-31 May 2003, 1 Jul-31 Aug 2003
53	07562	CM	18.82056	98.57389	APR	1976	MAR	1979	4	n/a	RID	< 30 years
54	07574	CM	18.62306	98.51167	APR	1975	MAR	1977	3	n/a	RID	< 30 years
55	07581	CM	18.85722	99.28667	APR	1977	MAR	1986	10	n/a	RID	< 30 years
56	07591	CM	18.61667	98.74528	APR	1979	MAR	2000	22	n/a	RID	< 30 years
57	07605	CM	19.38333	98.71833	APR	1972	MAR	1995	24	n/a	RID	< 30 years
58	07614	CM	19.23000	98.81333	APR	1972	MAR	1995	24	n/a	RID	< 30 years
59	07625	CM	18.28917	98.32000	APR	1971	MAR	1980	10	n/a	RID	< 30 years
60	07634	CM	18.49778	98.36306	APR	1970	MAR	1982	13	n/a	RID	< 30 years
61	07645	CM	18.22500	98.46667	APR	1971	MAR	1992	22	n/a	RID	< 30 years
62	07652	CM	18.51667	98.82250	APR	1982	MAR	2004	23	n/a	RID	< 30 years
63	07665	CM	19.15000	99.03333	APR	1983	MAR	2003	21	n/a	RID	< 30 years
64	07670	CM	19.16944	99.05250	APR	1984	MAR	2002	19	n/a	RID	< 30 years
65	07680	CM	19.71111	99.21389	APR	1986	MAR	1993	8	n/a	RID	< 30 years
66	07695	CM	n/a	n/a	APR	1985	MAR	1998	14	n/a	RID	< 30 years
67	07702	CM	19.55750	98.64028	APR	1989	MAR	2006	18	n/a	RID	< 30 years
68	07714	CM	18.30694	98.36583	APR	1988	MAR	2006	19	n/a	RID	< 30 years
69	07722	CM	19.41667	98.96667	APR	1989	MAR	2006	18	n/a	RID	< 30 years
70	07731	CM	17.78361	98.37528	APR	1990	MAR	2006	17	n/a	RID	< 30 years
71	07740	CM	19.11944	98.94639	APR	1986	MAR	1998	13	n/a	RID	< 30 years
72	07751	CM	19.63667	98.63889	APR	1995	MAR	2006	12	n/a	RID	< 30 years
73	07760	CM	n/a	n/a	APR	2001	MAR	2006	6	n/a	RID	< 30 years
74	07770	CM	n/a	n/a	APR	2001	MAR	2006	6	n/a	RID	< 30 years
75	07780	CM	n/a	n/a	APR	2001	MAR	2006	6	n/a	RID	< 30 years
76	07792	CM	n/a	n/a	APR	2001	MAR	2006	6	n/a	RID	< 30 years
77	07801	CM	18.65222	98.69056	APR	2002	MAR	2006	5	n/a	RID	< 30 years
78	07810	CM	n/a	n/a	APR	2003	MAR	2006	4	n/a	RID	< 30 years
79	07826	CM	19.26639	98.76722	APR	2004	MAR	2005	2	n/a	RID	< 30 years
80	07846	CM	19.14000	98.65833	APR	2004	MAR	2005	2	n/a	RID	< 30 years
81	07856	CM	19.19889	99.17500	APR	2003	MAR	2005	3	n/a	RID	< 30 years
82	07866	CM	19.01417	98.88389	APR	2004	MAR	2005	2	n/a	RID	< 30 years
83	07876	CM	18.92750	98.76194	APR	2004	MAR	2005	2	n/a	RID	< 30 years
84	07886	CM	18.82667	98.57333	APR	2003	MAR	2005	3	n/a	RID	< 30 years
85	07896	CM	19.68222	98.96333	APR	2003	MAR	2005	3	n/a	RID	< 30 years
86	07906	CM	19.61667	98.96167	APR	2003	MAR	2005	3	n/a	RID	< 30 years
87	07916	CM	19.55333	99.06667	APR	2003	MAR	2005	3	n/a	RID	< 30 years
88	07936	CM	18.67833	98.38194	APR	2003	MAR	2005	3	n/a	RID	< 30 years
89	07956	CM	18.25000	98.20000	APR	2003	MAR	2005	3	n/a	RID	< 30 years
90	07966	CM	18.16667	98.45000	APR	2003	MAR	2005	3	n/a	RID	< 30 years
91	07976	CM	n/a	n/a	APR	2003	MAR	2005	3	n/a	RID	< 30 years
92	07982	CM	18.61111	98.77778	APR	2003	MAR	2006	4	n/a	RID	< 30 years
93	07992	CM	18.62306	98.51167	APR	2003	MAR	2006	4	n/a	RID	< 30 years
94	071A2	CM	18.74361	98.92222	APR	2004	MAR	2006	3	n/a	RID	< 30 years
95	060201	CM	19.32667	98.94000	OCT	1988	DEC	2004	17	0	DWR	< 30 years; Station installed on 18 October 1988
96	060301	CM	19.44833	99.21500	APR	1979	DEC	2004	26	365	DWR	< 30 years; Station installed on 1 April 1979; 1985
97	060401	CM	19.63167	98.58333	JUNE	1980	DEC	2004	25	0	DWR	< 30 years; Station installed on 20 June 1980
98	060406	CM	19.40250	98.72750	OCT	1987	DEC	2004	18	0	DWR	< 30 years; Station installed on 20 October 1987

No.	Station Code	Province	Location		Period				Length (years)	Incomplete Data (days)	Data Source	Period of Incomplete Data
			Latitude	Longitude	From		To					
99	060602	CM	19.03000	98.84000	OCT	1988	DEC	2004	17	0	DWR	< 30 years; Station installed on 7 October 1988
100	060804	CM	18.66667	98.62333	JAN	1990	DEC	2004	15	0	DWR	< 30 years; Station installed on 3 January 1990
101	061006	CM	18.28833	98.52167	JAN	1991	DEC	2004	14	0	DWR	< 30 years; Station installed on 1 January 1991
102	061202	CM	18.72833	98.40000	OCT	1987	DEC	2004	18	0	DWR	< 30 years; Station installed on 22 October 1987
103	061302	CM	18.54833	98.35833	SEP	1982	DEC	2004	23	0	DWR	< 30 years; Station installed on 1 September 1982
104	061501	CM	17.38667	98.47167	NOV	1978	DEC	2004	27	31	DWR	< 30 years; Station installed on 25 November 1978; 1-31 May 1999
105	329201	LP	18.56667	99.03333	JAN	1981	DEC	2007	27	0	TMD	< 30 years
106	17012	LP	18.57722	99.00944	APR	1952	MAR	2006	55	1737	RID	1-31 May 1954, 1-31 Oct 1955, 1957-1960, 1 May-31 Oct 1961, 1-30 Sep 1980
107	17022	LP	17.80028	98.95472	APR	1955	MAR	2006	52	1766	RID	1 May-30 Jun 1955, 1977, 1-31 Oct 1979, 1983-1984, 1-30 Jun 1989, 1 Sep-31 Oct 1989, May 1992, 1 Jun-30 Nov 1999, 1-30 Nov 2000, 1 Apr-30 Jun 2002, 1 Sep-30 Nov 2002, 1 Oct-30 Nov 2003
108	17032	LP	18.52361	98.94389	APR	1955	MAR	2006	52	945	RID	1 May-30 Jun 1955, 1956-1957, 1-31 May 1963, 1-30 Jun 1997, 1-31 Aug 2000, 1 Oct-30 Nov 2005
109	17042	LP	18.45972	99.13722	APR	1952	AUG	2007	56	245	RID	1 May-31 Aug 1953, 1-31 Oct 1957, 1 May-31 Jul 1959, 1-30 Jun 1960
110	17052	LP	18.31444	98.82250	APR	1962	MAR	2006	45	883	RID	1969, 1 Jul-31 Oct 1970, 1971, 1-30 Apr 2005
111	17062	LP	17.65556	98.77500	APR	1959	AUG	2007	49	123	RID	1 May-31 Aug 1959
112	17074	LP	17.95194	98.89917	APR	1973	MAR	2006	34	31	RID	1-31 Oct 1984
113	17081	LP	17.88750	99.08889	APR	1978	MAR	2004	27	306	RID	< 30 years; 1 May-31 Aug 1978, 1 Oct-30 Nov 1978, May 1997, 1-30 Jun 1999, 1-31 Aug 1999, 1-30 Jun 2000
114	17093	LP	18.58333	99.03333	APR	1980	MAR	2006	27	n/a	RID	< 30 years
115	17101	LP	18.58667	99.15750	APR	1983	MAR	1999	17	n/a	RID	< 30 years
116	17111	LP	18.38639	99.01028	APR	1986	MAR	1988	3	n/a	RID	< 30 years
117	17120	LP	n/a	n/a	APR	2003	MAR	2004	2	n/a	RID	< 30 years
118	17130	LP	n/a	n/a	APR	2003	MAR	2004	2	n/a	RID	< 30 years
119	17140	LP	n/a	n/a	APR	2003	MAR	2004	2	n/a	RID	< 30 years
120	17150	LP	n/a	n/a	APR	2003	MAR	2004	2	n/a	RID	< 30 years
121	17160	LP	n/a	n/a	APR	2003	MAR	2004	2	n/a	RID	< 30 years
122	17181	LP	18.13972	98.89944	APR	2003	MAR	2006	4	n/a	RID	< 30 years
123	17196	LP	18.50000	99.26667	APR	2003	MAR	2005	3	n/a	RID	< 30 years
124	17206	LP	18.41722	98.99611	APR	2003	MAR	2005	3	n/a	RID	< 30 years
125	061101	LP	17.58694	98.81111	OCT	1988	DEC	2004	17	0	DWR	< 30 years; Station installed on 10 October 1988
126	376201	TK	16.88333	98.11667	JAN	1955	JUL	2007	53	36	TMD	25-29 Aug 1959, 1-31 Aug 1988
127	376202	TK	16.65917	98.55083	JAN	1951	JUL	2007	57	33	TMD	1-2 Oct 1962, 1-31 Aug 1988
128	376203	TK	17.23333	98.05306	JAN	1961	JUL	2007	47	762	TMD	1967-1968, 1-31 Aug 1988
129	376301	TK	16.75000	98.93333	JAN	1992	JUL	2007	16	0	TMD	< 30 years
130	376401	TK	16.01583	98.86556	JAN	1977	DEC	2007	31	0	TMD	
131	63013	TK	16.88056	99.12667	APR	1921	MAR	2006	86	706	RID	1-30 Sep 1925, 1-31 May 1928, 1-31 Aug 1929, 1-31 Oct 1930, 1-31 May 1931, 1-31 May 1934, 1-31 Oct 1942, 1 Jun-31 Aug 1950, 1 Jul-31 Aug 1952, 1 Jul-31 Oct 1980, 1-30 Nov 2003, 1 Jun-30 Nov 2004

No.	Station Code	Province	Location		Period				Length (years)	Incomplete Data (days)	Data Source	Period of Incomplete Data
			Latitude	Longitude	From		To					
132	63022	TK	17.04611	99.07611	APR	1921	MAR	2006	86	1529	RID	1-31 Aug 1927, 1-31 May 1931, 1-31 May 1935, 1939, 1-30 Jun 1940, 1945, 1-30 Jun 1946, 1-31 May 1947, 1-31 Oct 1947, 1-31 May 1948, 1-31 Oct 1948, 1 May-30 Jun 1949, 1950, 1-31 May 1951, 1-31 Oct 1959, 1 Aug-31 Oct 1969, 1-31 Oct 1973, 1-31 Jul 1975, 1-31 Oct 1975, 1 May-30 Sep 2002, 1 Apr-31 May 2003
133	63033	TK	16.71194	98.57583	APR	1921	MAR	2006	86	5542	RID	1929-1932, 1937-1939, 1-31 Oct 1940, 1-31 Jul 1941, 1 Sep-31 Oct 1942, 1 Aug-31 Oct 1943, 1944-1950, 17 Nov 1998, 1-30 Nov 2003, 1 Jun-30 Nov 2004
134	63042	TK	16.01611	98.86667	APR	1923	MAR	2006	84	4663	RID	1929-1932, 1-31 Oct 1933, 1939, 1-31 Aug 1941, 1-31 Oct 1941, 1-31 Aug 1943, 1-31 Oct 1943, 1944, 1-31 Jul 1945, 1-31 Oct 1945, 1 May-31 Jul 1946, 1-31 Jul 1947, 1948-1950, 1-31 May 1951, 1 Sep-31 Oct 1968, 1-31 Oct 1969, 1974-1975, 1 Jun-30 Nov 2004
135	63052	TK	16.98056	98.52056	APR	1921	AUG	2007	87	5760	RID	1-31 May 1921, 1-30 Sep 1921, 1-31 Oct 1923, 1-31 Oct 1928, 1 Aug-31 Oct 1929, 1930-1933, 1 May-30 Jun 1934, 1-31 Oct 1936, 1939-1940, 1 Aug-31 Oct 1941, 1-31 Jul 1942, 1-30 Sep 1942, 1943-1946, 1-31 Oct 1947, 1948-1950, 1-31 May 1951, 1-31 May 1963, 1-31 Oct 1966, 1 Aug-30 Sep 1971, 1972
136	63062	TK	17.24222	99.02444	APR	1944	MAR	2006	63	2685	RID	1-31 May 1944, 1 May-31 Jul 1946, 1-31 Jul 1947, 1 Sep-31 Oct 1948, 1-31 Oct 1949, 1950, 1956-1957, 1-30 Jun 1964, 1 Sep-31 Oct 1968, 1969, 1 Aug-31 Oct 1981, 1983, 1-31 Jul 1986, 1-31 Jul 1987, 1-30 Jun 1990, 1 Jul-31 Aug 1994, 1-31 Oct 1994, 1-31 Jul 2000, 1 Oct-30 Nov 2000, 1 May-31 Aug 2004, 1-30 Nov 2005
137	63075	TK	17.24167	99.06250	APR	1959	MAR	2006	48	1190	RID	1-31 Aug 1963, 1 May-31 Oct 1967, 1-31 Oct 1968, 1999-2000, 1-30 Nov 2003, 1 Jun-30 Nov 2004
138	63082	TK	16.76667	98.93056	APR	1961	MAR	2001	41	2894	RID	1-31 May 1961, 1973-1975, 1 May-31 Oct 1976, 1996-1999, 1 May-31 Aug 2000
139	63092	TK	17.22444	98.22806	APR	1967	MAR	2006	40	122	RID	1 Sep-31 Oct 1983, 1-31 Oct 1991, 1-30 Jun 1992
140	63100	TK	n/a	n/a	APR	1955	MAR	1958	4		RID	< 30 years
141	63111	TK	17.24167	99.01250	APR	1952	MAR	1963	12		RID	< 30 years
142	63120	TK	16.92417	99.30250	APR	1971	MAR	2002	32	3775	RID	1 Jul-31 Oct 1989, 1990-1999
143	63132	TK	17.24972	98.86583	APR	1971	MAR	2006	36	1161	RID	1-30 Jun 1975, 1-31 Oct 1975, 1 May-30 Jun 1976, 1-30 Jun 1977, 1978, 1 Jun-31 Jul 1979, 1-31 Oct 1979, 1-31 Oct 1984, 1 Sep-31 Oct 1985, 1 May-31 Oct 1986, 1-31 May 1987, 1-31 May 1991, 1 Oct-30 Nov 2002, 1 May-31 Aug 2003, 1-30 Apr 2004
144	63142	TK	16.76667	98.93333	APR	1972	MAR	1995	24	n/a	RID	< 30 years
145	63152	TK	16.91667	98.11667	APR	1972	MAR	2006	35	762	RID	1 May-31 Aug 1972, 1 Sep-31 Oct 1976, 1977, 1 Oct-30 Nov 1997, 1-30 Jun 2002, 1 Aug-30 Nov 2003
146	63162	TK	17.33333	98.88333	APR	1968	MAR	2006	39	1859	RID	1-31 May 1968, 1-31 May 1969, 1977-1979, 1 Aug-31 Oct 1980, 1983, 1 Oct-30 Nov 2002, 1 May-31 Aug 2003, 1 Apr-31 May 2004

No.	Station Code	Province	Location		Period				Length (years)	Incomplete Data (days)	Data Source	Period of Incomplete Data
			Latitude	Longitude	From		To					
147	63172	TK	17.01667	98.66667	APR	1977	MAR	2006	30	428	RID	< 30 years; 1 Apr-30 Jun 1977, 1 Aug-30 Nov 1977, 1 Jun-31 Oct 1981, 1-31 Oct 1982, 1-31 Aug 2003
148	63181	TK	16.76222	98.75389	APR	1977	MAR	2006	30	92	RID	< 30 years; 1 May-30 Jun 1977, 1-31 May 1999
149	63192	TK	16.76667	99.03333	APR	1986	MAR	2006	21	n/a	RID	< 30 years
150	63202	TK	16.37444	98.68528	APR	1986	MAR	2006	21	n/a	RID	< 30 years
151	63215	TK	n/a	n/a	APR	1987	MAR	1995	9	n/a	RID	< 30 years
152	63225	TK	n/a	n/a	APR	1987	MAR	1995	9	n/a	RID	< 30 years
153	63235	TK	n/a	n/a	APR	1991	MAR	2006	16	n/a	RID	< 30 years
154	63242	TK	16.75000	98.93333	APR	1995	MAR	1996	2	n/a	RID	< 30 years
155	060101	TK	17.05000	99.06667	SEP	1971	DEC	2007	37	0	DWR	Station installed on 31 August 1971
156	380201	KP	16.48333	99.53333	JAN	1981	JUL	2007	27	0	TMD	< 30 years
157	12012	KP	16.48222	99.52389	APR	1952	MAR	2001	50	4812	RID	Jun 1957, Oct 1959, Jul 1978, May 1985, 1987-1994, Apr-Nov 1995, 1996-1999, May 2000, Aug 2000
158	12022	KP	16.21444	99.72111	APR	1952	MAR	2006	55	3237	RID	1956, May 1957, Sep-Oct 1959, Jul-Oct 1961, 1962, May-Jul 1963, May-Jul 1970, Jul & Oct 1979, May-Sep 1982, May-Oct 1986, Jul-Oct 1987, Jul-Oct 1988, 1990-1991, Oct 1992, Jul-Oct 1994, Jan-Nov 1995, Sep-Nov 2001, Jun-Aug 2002
159	12032	KP	16.66333	99.59194	APR	1953	MAR	2006	54	214	RID	May-Oct 1957, Nov 2005
160	12042	KP	16.06028	99.86361	APR	1953	MAR	2006	54	367	RID	May-Jun 1953, Jun 1956, Jun-Sep 1961, Aug & Oct 1962, May-Jul 1963
161	12052	KP	16.46667	99.65000	APR	1966	MAR	2006	41	153	RID	Sep-Oct 1979, Sep-Oct 1997, Jul 2003
162	12061	KP	16.44833	99.43250	APR	1971	MAR	1998	28	337	RID	< 30 years; May-Aug 1974, May-Jun & Oct 1993, Apr 1994, Jun-Aug 1997
163	12081	KP	15.90278	99.47917	APR	1970	MAR	2006	37	31	RID	Jul 1972
164	12091	KP	16.07278	99.40500	APR	1976	MAR	2002	27	246	RID	< 30 years; May & Oct 1976, Oct 1979, Sep 1983, Oct 1987, May 1999, Oct-Nov 2000
165	12102	KP	16.45000	99.88333	APR	1980	MAR	2006	27	214	RID	< 30 years; May-Oct 1980, Nov 2005
166	12113	KP	n/a	n/a	APR	1980	MAR	2006	27	367	RID	< 30 years; May-Oct 1980, Jun-Nov 2004
167	12121	KP	16.33417	99.27472	APR	1984	MAR	2006	23	62	RID	< 30 years; May 1989, May 1999
168	12132	KP	16.24639	99.33028	APR	1986	MAR	2005	20	489	RID	< 30 years; May-Oct 1986, Jul 1990, May-Oct 1991, Apr & Nov 2001, Nov 2004
169	12142	KP	16.70944	99.85111	APR	1986	MAR	2006	21	215	RID	< 30 years; May-Oct 1986, Oct 2001
170	12152	KP	16.45111	99.50194	APR	1994	MAR	2006	13	215	RID	< 30 years; May-Aug 1994, May & Aug-Sep 1997
171	12161	KP	n/a	n/a	APR	2000	MAR	2006	7	31	RID	< 30 years; Jul 2000
172	12176	KP	n/a	n/a	APR	2003	MAR	2006	4	215	RID	< 30 years; May-Aug 2003, May 2004, May-Jun 2005
173	12186	KP	n/a	n/a	APR	2003	MAR	2006	4	215	RID	< 30 years; May-Aug 2003, May 2004, May-Jun 2005
174	12196	KP	n/a	n/a	APR	2003	MAR	2006	4	215	RID	< 30 years; May-Aug 2003, May 2004, May-Jun 2005
175	12206	KP	n/a	n/a	APR	2003	MAR	2006	4	215	RID	< 30 years; May-Aug 2003, May 2004, May-Jun 2005
176	12216	KP	n/a	n/a	APR	2003	MAR	2006	4	215	RID	< 30 years; May-Aug 2003, May 2004, May-Jun 2005
177	400201	NK	15.80000	100.16667	JAN	1951	JUL	2007	57	6	TMD	30-31 May 1960, 1-3, 10 & 13 Jun 1960
178	400301	NK	15.35000	100.50000	JAN	1969	JUL	2007	39	123	TMD	Jul-Oct 1982
179	26013	NK	15.80000	100.16667	APR	1921	MAR	2006	86	1218	RID	1925, Sep-Oct 1942, 1944-1945, Oct 1949, 28 Oct 1998, Nov 2003

No.	Station Code	Province	Location		Period				Length (years)	Incomplete Data (days)	Data Source	Period of Incomplete Data
			Latitude	Longitude	From		To					
180	26022	NK	15.87944	100.30667	APR	1921	MAR	2006	86	5177	RID	Aug 1923, Oct 1925, May 1926, Oct 1930, May 1931, May 1934, Oct 1936, 1938, 1940-1951, Sep 1965, Jun 1996, Apr & Oct 1997, Nov 1999, Oct-Nov 2005
181	26032	NK	15.60000	100.51667	APR	1921	MAR	2006	86	3410	RID	1928-1930, Oct 1936, Oct 1938, 1939-1940, 1942, Sep 1943, 1945, May-Jul 1946, Sep-Oct 1947, Jul & Oct 1948, 1949, Oct 1950, May-Jun 1951, 2 & 4-31 Dec1998, May-Jun 1999
182	26042	NK	15.51667	100.08333	APR	1921	MAR	2006	86	2835	RID	Jul 1923, May 1926, 1930, Jun-Jul 1931, Oct 1940, Sep-Oct 1942, Aug-Oct 1944, 1945-1947, Oct 1948, Jul 1949, 1950, May-Jun 1951, Oct 1994, Oct-Nov 1995, Nov 1996, Apr-Jul & Sep-Nov 1997, May-Nov 1999, Nov 2000
183	26052	NK	15.41667	100.16667	APR	1921	MAR	2006	86	2173	RID	May 1924, Oct 1925, Jul & Oct 1927, Sep-Oct 1928, 1930, May 1934, Oct 1935, Jul 1941, Sep 1942, May & Oct 1943, Oct 1944, Sep-Oct 1945, Aug 1946, May-Oct 1948, May-Jul 1951, May 1962, Oct 1965, Oct 1969, Jul & Oct 1972, Nov 1995, Nov 1997, Oct-Nov 1999, Jun 1999, May-Aug 2002, Apr-Sep & Nov 2003, Apr-Oct 2005
184	26062	NK	15.88333	100.01667	APR	1920	MAR	2006	87	5488	RID	Jul & Oct 1925, 1927-1929, Aug 1930, May 1934, Sep 1936, 1943-1951, May-Jun 1952, Sep 1959, Oct 1979, May-Oct 1993, May-Oct 1994, Jul 1995, Oct-Nov 1996, May-Aug & Oct 1997, May-Aug 1998, Apr-Jun 1999
185	26072	NK	15.70000	99.81667	APR	1931	MAR	2006	76	1007	RID	Oct 1943, 1944, May-Jul 1945, Oct 1946, Jul 1948, 1951, Sep-Oct 1953, Nov 2005
186	26082	NK	15.21667	100.35000	APR	1921	MAR	2006	86	4092	RID	Oct 1921, Jun 1924, Oct 1925, Oct 1927, 1929, Oct 1934, Jul 1938, Jul 1941, Jun 1942, Jun & Oct 1943, 1945, May & Oct 1946, May-Jul 1951, Oct 1960, Oct 1961, Sep-Oct 1962, May-Oct 1963, 1966, May-Jul 1967, May 1977, Jun-Oct 1990, Aug-Oct 1991, May-Oct 1992, May-Oct 1993, 1994-1996, May-Aug & Oct 1997, Apr & Aug-Nov 1998, Apr-Jun 1999
187	26092	NK	15.65000	100.16667	APR	1936	MAR	2006	71	4083	RID	Oct 1938, Oct 1940, 1941-1946, Oct 1947, Sep 1948, Oct 1949, 1951-1953, Oct 1954, Oct 1979, 1991, May-Oct 1992, Jun 2002
188	26102	NK	15.86417	100.58861	APR	1952	MAR	2006	55	549	RID	1978, Oct 1979, Sep-Oct 1984, Jun & Oct 1991, Oct 1993
189	26112	NK	n/a	n/a	APR	1960	MAR	1983	24	184	RID	< 30 years; May-Oct 1960
190	26122	NK	15.56667	100.70000	APR	1967	MAR	2006	40	428	RID	Jun & Oct 1967, May-Oct 1968, May 1993, Nov 1995, Oct-Nov 1999, Jul 2003, Nov 2005
191	26134	NK	n/a	n/a	APR	1958	MAR	1976	19	3714	RID	< 30 years; 1965-1974, May & Oct 1975
192	26142	NK	15.35000	100.50000	APR	1981	MAR	2006	26	2196	RID	< 30 years; May-Oct 1981, Jun-Oct 1982, 1983-1987, May 1988, 1 Nov 1998
193	26154	NK	n/a	n/a	APR	1971	MAR	1973	3	123	RID	< 30 years; Jul-Oct 1972
194	26164	NK	n/a	n/a	APR	1935	MAR	1951	17	400	RID	< 30 years; Sep 1940, Oct 1942, May-Aug & Oct 1945, May-Jul & Oct 1946, Jul & Oct 1950
195	26170	NK	n/a	n/a	APR	1965	MAR	2006	42	793	RID	Oct 1979, Oct 1992, 1999, 2001
196	26180	NK	n/a	n/a	APR	1965	MAR	2006	42	366	RID	1999

No.	Station Code	Province	Location		Period				Length (years)	Incomplete Data (days)	Data Source	Period of Incomplete Data
			Latitude	Longitude	From		To					
197	26190	NK	n/a	n/a	APR	1965	MAR	2006	42	397	RID	May 1987, 1999
198	26200	NK	n/a	n/a	APR	1965	MAR	2006	42	397	RID	May 1987, 1999
199	26210	NK	n/a	n/a	APR	1964	MAR	2006	43	797	RID	May-Jun 1988, Jul 1994, Jul-Aug 1996, 1999, May 2002, May 2003, Jun-Aug 2004, May-Aug 2005
200	26220	NK	n/a	n/a	APR	1964	MAR	2005	42	611	RID	May-Jun 1988, Jun 1991, Jul 1994, 1999, May 2002, Jun-Aug 2004
201	26230	NK	n/a	n/a	APR	1964	MAR	2005	42	581	RID	May-Jun 1988, 1999, May 2002, May 2003, Jun-Aug 2004
202	26252	NK	n/a	n/a	APR	1970	MAR	1973	4	62	RID	< 30 years; Jul-Aug 1972
203	26262	NK	15.78333	99.68333	APR	1970	MAR	2006	37	457	RID	1991, Nov 1999, May & Nov 2005
204	26271	NK	n/a	n/a	APR	1975	MAR	2006	32	184	RID	Jun-Jul 1975, May-Jun 1996, Jul 1997, Jul 2004
205	26281	NK	n/a	n/a	APR	1975	MAR	2006	32	61	RID	Jun-Jul 1975
206	26292	NK	15.76667	100.08333	APR	1975	MAR	2006	32	1346	RID	May-Oct 1975, Oct 1979, Jul 1980, Oct 1985, Sep-Oct 1986, 1987, 1991, May-Oct 1992, Oct 1998
207	26301	NK	n/a	n/a	APR	1991	MAR	2006	16	4143	RID	< 30 years; May-Oct 1991, May-Jul & Sep 1992, May-Oct 1993, 1995-2004
208	26311	NK	n/a	n/a	APR	2000	MAR	2006	7	0	RID	< 30 years

CM: Chiang Mai; LP: Lam Phun; TK: Tak; KP: Kamphaeng Phet; NK: Nakhon Sawan.

RID: Royal Irrigation Department of Thailand; TMD: Thailand Meteorological Department; DWR: Department of Water Resources.

A2: List of the 50 selected rainfall stations

Station code	Province	Location		Period		Length (years)	Data source
		Latitude	Longitude	From	To		
327301	CM	18.91667	99.00000	1969	2007	39	TMD
327501	CM	18.79000	98.97694	1951	2007	57	TMD
07052	CM	18.86889	99.13944	1952	2006	55	RID
07072	CM	18.68611	98.92194	1952	2005	54	RID
07082	CM	18.62694	98.89889	1952	2006	55	RID
07112	CM	19.11889	98.94778	1952	2006	55	RID
07122	CM	19.36444	99.20472	1952	2006	55	RID
07142	CM	18.84778	98.73583	1956	2007	52	RID
07172	CM	19.95972	99.16056	1952	2006	55	RID
07242	CM	18.80278	98.92500	1961	2007	47	RID
07252	CM	19.26861	98.97556	1975	2007	33	RID
07262	CM	18.80667	98.90333	1965	2006	42	RID
07292	CM	18.61111	98.90056	1962	2006	45	RID
07391	CM	18.78917	99.01694	1971	2006	36	RID
07492	CM	19.99639	99.25917	1970	2006	37	RID
07502	CM	19.06667	99.21667	1972	2006	35	RID
07092	CM	18.19056	98.61444	1952	2004	53	RID
07152	CM	18.49833	98.36500	1952	2006	55	RID
07162	CM	17.79583	98.36000	1972	2007	36	RID
07182	CM	18.41583	98.67972	1952	2006	55	RID
07282	CM	18.15028	98.39306	1966	2007	42	RID
07472	CM	17.91667	98.68333	1969	2006	38	RID
17012	LP	18.57722	99.00944	1952	2006	55	RID
17032	LP	18.52361	98.94389	1955	2006	52	RID
17042	LP	18.45972	99.13722	1960	2007	48	RID
17052	LP	18.31444	98.82250	1962	2006	45	RID
17062	LP	17.65556	98.77500	1959	2007	49	RID
17074	LP	17.95194	98.89917	1973	2006	34	RID
376201	TK	16.88333	98.11667	1960	2007	48	TMD
376203	TK	17.23333	98.05306	1961	2007	47	TMD
63022	TK	17.04611	99.07611	1921	2006	86	RID
63042	TK	16.01611	98.86667	1923	2006	84	RID
060101	TK	17.05000	99.06667	1971	2007	37	DWR
12032	KP	16.66333	99.59194	1953	2006	54	RID
12042	KP	16.06028	99.86361	1964	2006	43	RID
12052	KP	16.46667	99.65000	1966	2006	41	RID
12081	KP	15.90278	99.47917	1973	2006	34	RID
400201	NK	15.80000	100.16667	1951	2007	57	TMD
400301	NK	15.35000	100.50000	1969	2007	39	TMD
26022	NK	15.87944	100.30667	1921	2006	86	RID
26032	NK	15.60000	100.51667	1921	2006	86	RID
26072	NK	15.70000	99.81667	1931	2006	76	RID
26102	NK	15.86417	100.58861	1952	2006	55	RID
26122	NK	15.56667	100.70000	1967	2006	40	RID
26180	NK	n/a	n/a	1965	2006	42	RID
26190	NK	n/a	n/a	1965	2006	42	RID
26200	NK	n/a	n/a	1965	2006	42	RID

Station code	Province	Location		Period		Length (years)	Data source
		Latitude	Longitude	From	To		
26262	NK	15.78333	99.68333	1970	2006	37	RID
26271	NK	n/a	n/a	1975	2006	32	RID
26281	NK	n/a	n/a	1976	2006	31	RID

CM: Chiang Mai; LP: Lam Phun; TK: Tak; KP: Kamphaeng Phet; NK: Nakhon Sawan.
RID: Royal Irrigation Department of Thailand; TMD: Thailand Meteorological Department; DWR: Department of Water Resources.

A3: List of the streamflow gauging stations in the Ping River Basin

No.	Station Code	Province	Location		Period				Length (years)	Data Source	Period of Incomplete Data
			Latitude	Longitude	From		To				
1	P1	CM	18.78583	99.00806	Apr	1921	Mar	2007	87	RID	Jun 1932
2	P2A	TK	16.85389	99.13056	Apr	1952	Mar	2005	54	RID	
3	P4A	CM	19.12083	98.94750	Apr	1955	Mar	2007	53	RID	Aug 1966/1-11 Feb 1974/1977/31 Jan, 31 Mar 2008
4	P5	LP	18.57556	99.01222	Apr	1954	Mar	2007	54	RID	1969-1977/1993-2004/Aug 2005/16-18, 28 Sep 2005/6,9 Mar 2008
5	P7A	KP	16.47722	99.51833	Apr	1978	Mar	2005	28	RID	
6	P12	TK	17.24167	99.01250	Apr	1952	Mar	1994	43	RID	1969-1971/14-20 Jan 1977/3-6 Jul, 4-11 Aug 1978
7	P12A	TK	17.24444	98.96472	Apr	1952	Mar	1994	43	RID	1969-1971/14-20 Jan 1977/3-6 Jul, 4-11 Aug 1978
8	P12B	TK	17.24083	99.02500	Apr	1996	Mar	2006	11	EGAT	
9	P14	CM	18.23028	98.55972	Apr	1954	Mar	2007	54	RID	31 Mar 1958/1963/Apr-May 1965
10	P19	CM	18.42083	98.69806	Apr	1958	Mar	1992	35	RID	
11	P21	CM	18.92472	98.94278	Apr	1954	Mar	2007	54	RID	Jul 1959/Jan, 1-3 Feb, 24 Apr-26 May 1976
12	P24	CM	18.38750	98.68083	Apr	1955	Mar	2004	50	RID	Sep, 1-10, 24-31 Oct, 1-14 Nov 1955/1-26 Jan, 29 Feb, May-3 Jun 1956/6-19, 21, 23-25 May 1957/31 May 1959/15 Jul 1973-31 Mar 1974
13	P24A	CM	18.41694	98.67472	Apr	1973	Mar	2007	35	RID	31 Jan, 31 Mar 2008
14	P35	KP	16.07278	99.40500	Apr	1974	Mar	2001	28	RID	
15	P47	KP	16.33417	99.27472	Apr	1983	Mar	2005	23	RID	1 Apr-25 May 1983
16	P56A	CM	19.28389	99.19028	Apr	1999	Mar	2005	7	RID	4 Jun, 15-17, 20-29 Jul, 2, 13-21, 31 Aug, 1-6, 9-30 Sep, 1-21, 28-31 Oct, 1-11 Nov, 6-11 Dec 2005
17	P64	CM	17.78361	98.37528	Apr	1990	Mar	2005	16	RID	1 Apr-11 Sep 1990/2003/Sep, 3 Oct 2005
18	P65	CM	19.63611	98.63861	Apr	1992	Mar	2004	13	RID	1 Apr-27 Jun 1992/2002-2003
19	P67	CM	19.01972	98.96167	Apr	1996	Mar	2007	12	RID	May 2007/31 Jan, 2 Feb, 31 Mar 2008
20	P70	CM	19.65222	98.66944	Apr	1995	Mar	2000	6	RID	
21	P71	CM	18.53722	98.86306	Apr	1996	Mar	2007	12	RID	31 Jan, 31 Mar 2008
22	P73	CM	18.28833	98.65306	Apr	1998	Mar	2007	10	RID	31 Jan, 31 Mar 2008
23	P75	CM	19.14778	99.01000	Apr	1999	Mar	2007	9	RID	31 Jan, 31 Mar 2008
24	P76	LP	18.13972	98.89944	Apr	2000	Mar	2005	6	RID	
25	P77	LP	18.43250	99.08333	Apr	1999	Mar	2005	7	RID	Nov 2005
26	P85	LP	18.36389	98.77556	Apr	2003	Mar	2007	5	RID	
27	060101	TK	17.05000	99.06667	Jan	1972	Dec	2004	33	DWR	1979
28	060201	CM	19.32667	98.94000	Jan	1986	Dec	2004	19	DWR	
29	060202	CM	19.48667	98.72750	Jan	1983	Dec	2004	22	DWR	
30	060301	CM	19.44833	99.21500	Jan	1977	Dec	2004	28	DWR	1-14 Jan 1977/27 Sep-13 Oct 1982
31	060302	CM	19.37333	99.24833	Jan	1986	Dec	2004	19	DWR	
32	060401	CM	19.63167	98.58333	Jan	1983	Dec	2004	22	DWR	
33	060406	CM	19.40250	98.72750	Jan	1985	Dec	2004	20	DWR	
34	060602	CM	19.03000	98.84000	Jan	1983	Dec	2004	22	DWR	1-3 Aug 1983/15-19 Sep 1994
35	060603	CM	19.02333	98.88000	Jan	1985	Dec	2004	20	DWR	
36	060701	CM	18.96500	99.27667	Jan	1983	Dec	2004	22	DWR	18 Aug-13 Nov 1987
37	060704	CM	18.98250	99.33944	Jan	1983	Dec	2004	22	DWR	
38	060804	CM	18.66667	98.62333	Jan	1983	Dec	2004	22	DWR	
39	060808	CM	18.61083	98.85444	Jan	1983	Dec	2004	22	DWR	
40	061001	CM	18.54000	98.59500	Jan	1983	Dec	2004	22	DWR	
41	061006	CM	18.28833	98.52167	Jan	1991	Dec	2004	14	DWR	
42	061101	LP	17.58694	98.81111	Jan	1984	Dec	2004	21	DWR	2000
43	061202	CM	18.72833	98.40000	Jan	1985	Dec	2004	20	DWR	
44	061302	CM	18.54833	98.35833	Jan	1983	Dec	2007	25	DWR	6-12 Sep, 23 Sep-3 Oct 1995
45	061501	CM	17.38667	98.47167	Jan	1977	Dec	2007	31	DWR	

CM: Chiang Mai; LP: Lam Phun; TK: Tak; KP: Kamphaeng Phet.

RID: Royal Irrigation Department of Thailand; DWR: Department of Water Resources; EGAT: Electricity Generating Authority of Thailand.

A4: List of the 12 selected streamflow stations

Station code	Province	Location		Period		Length (years)	Data source
		Latitude	Longitude	From	To		
P1	CM	18.78583	99.00806	1921	2007	87	RID
P4A	CM	19.12083	98.94750	1955	2007	53	RID
P12C	TK	17.24083	99.02500	1996	2006	11	EGAT
P14	CM	18.23028	98.55972	1954	2007	54	RID
P21	CM	18.92472	98.94278	1954	2007	54	RID
P24A	CM	18.41694	98.67472	1973	2007	35	RID
P67	CM	19.01972	98.96167	1996	2007	12	RID
P71	CM	18.53722	98.86306	1996	2007	12	RID
P73	CM	18.28833	98.65306	1998	2007	10	RID
P75	CM	19.14778	99.01000	1999	2007	9	RID
061302	CM	18.54833	98.35833	1983	2007	25	RID
061501	CM	17.38667	98.47167	1977	2007	31	RID

CM: Chiang Mai; TK: Tak.

RID: Royal Irrigation Department of Thailand; EGAT: Electricity Generating Authority of Thailand.

A5: List of the meteorological stations located in and around the Ping River Basin

No.	Station Code	Province	Location		Period		Length (years)	Type of Data	Data Source	Period of Incomplete Data
			Latitude	Longitude	From	To				
1	327301	CM	18.91667	99.00000	1969	2007	39	T/E	TMD	27-28 Feb 1981, 12-30 Sep 2007
2	327501	CM	18.79000	98.97694	1952	2007	56	T/E	TMD	Jan 1952, 2 Feb 1953, 18-19 Mar 1953, 14-16 Oct 1956, 10 Sep 1965, 28 Dec 1976
3	329201	LP	18.56667	99.03333	1981	2007	27	T/E	TMD	
4	376201	TK	16.88333	99.11667	1955	2007	53	T/E	TMD	25-29 Aug 1959, Aug 1988
5	376202	TK	16.65917	98.55083	1951	2007	57	T/E	TMD	21-22 Aug 1952, 20-31 Oct 1957, Nov-Dec 1957, Jan-Apr 1958, 1-25 May 1958, 27 May 1958, 30-31 May 1958, 29-30 Sep 1962, 1-2 Oct 1962, 27-29 Feb 1976, 1-9 Mar 1976, 11-31 Aug 1980, Aug 1988
6	376203	TK	17.23333	99.05306	1961	2007	47	T/E	TMD	1967-68, 9 Feb 1969, Aug 1988
7	376301	TK	16.75000	98.93333	1992	2007	16	T/E	TMD	24 Oct 2005
8	376401	TK	16.01583	98.86556	1978	2007	30	T/E	TMD	28-30 Sep 1981, Oct 1981
9	380201	KP	16.48333	99.53333	1981	2007	27	T/E	TMD	
10	400201	NK	15.80000	100.16667	1951	2007	57	T/E	TMD	
11	400301	NK	15.35000	100.50000	1969	2007	39	T/E	TMD	
12	060101	TK	17.05000	99.06667	1971	2004	34	E	DWR	
13	060401	CM	19.63167	98.58333	1980	2004	25	E	DWR	
14	060804	CM	18.66667	98.62333	1990	2004	15	E	DWR	
15	061501	CM	17.38667	98.47167	1979	2004	26	E	DWR	
16	07391	CM	18.78917	99.01694	1978	2006	29	E	RID	
17	07591	CM	18.61667	98.74528	1979	1992	14	E	RID	
18	07731	CM	17.78361	98.37528	1991	2006	16	E	RID	

CM: Chiang Mai; LP: Lam Phun; TK: Tak; KP: Kamphaeng Phet; NK: Nakhon Sawan.
T: temperature; E: Class A pan evaporation.
RID: Royal Irrigation Department of Thailand; TMD: Thailand Meteorological Department; DWR: Department of Water Resources.

Appendix B: Monthly, Seasonal and Annual Hydroclimate Diagnostics

B1: Monthly, annual and MAM temperature (°C) averaged over the 11 selected stations

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Annual	MAM
1951	24.1	25.7	29.1	31.0	30.3	27.9	27.6	27.3	28.1	27.5	27.0	24.6	27.5	30.1
1952	24.8	26.6	27.7	29.2	30.0	28.3	27.5	27.5	26.9	27.2	26.1	22.1	27.0	29.0
1953	23.7	26.4	28.3	29.2	28.6	28.1	27.6	26.0	27.3	27.3	25.5	22.9	26.7	28.7
1954	22.9	25.5	27.6	30.2	28.5	28.2	27.7	27.5	27.2	26.6	25.1	22.7	26.6	28.8
1955	21.5	25.3	27.9	29.3	29.0	27.6	27.7	27.5	27.4	26.8	24.5	21.0	26.3	28.7
1956	21.9	25.8	28.8	29.6	28.2	27.8	27.6	27.2	27.3	26.7	24.1	22.4	26.5	28.9
1957	23.2	24.3	28.1	30.6	31.3	28.4	27.9	27.7	27.6	26.9	25.6	23.9	27.1	30.0
1958	23.8	25.8	30.1	31.0	31.2	28.8	27.7	28.0	27.6	27.2	24.3	21.2	27.2	30.7
1959	23.2	25.9	28.0	30.8	29.5	28.9	27.6	27.1	27.1	26.8	25.9	25.5	27.2	29.4
1960	24.0	25.3	28.6	31.7	29.6	28.3	28.0	27.6	27.6	26.9	26.3	22.8	27.2	30.0
1961	22.7	27.1	28.7	30.6	28.7	27.6	27.3	26.9	26.8	26.5	25.6	23.7	26.8	29.3
1962	22.3	24.4	28.8	30.9	30.3	28.4	28.0	27.7	27.2	26.6	25.3	22.0	26.8	30.0
1963	20.5	24.9	28.1	30.6	31.9	28.3	27.0	27.8	27.5	26.9	26.2	23.1	26.9	30.2
1964	23.7	25.9	28.8	30.7	28.7	28.4	27.9	27.7	27.4	27.0	24.3	22.1	26.9	29.4
1965	21.4	25.9	28.0	30.8	29.8	27.4	27.9	27.7	27.3	26.8	25.5	24.8	26.9	29.5
1966	24.9	27.0	29.3	31.6	29.4	28.2	28.2	27.8	27.2	27.1	25.8	25.2	27.6	30.1
1967	23.3	25.3	28.3	30.1	29.3	29.1	28.4	27.5	27.3	26.5	25.3	16.8	26.4	29.2
1968	22.9	25.1	28.8	29.6	29.0	28.2	28.2	27.7	27.9	26.9	26.5	24.8	27.1	29.1
1969	24.8	25.9	29.4	31.0	30.1	28.7	27.6	27.1	27.3	27.0	24.3	22.3	27.1	30.1
1970	23.8	25.6	28.8	29.4	28.9	27.9	27.3	27.1	27.4	26.5	25.2	24.1	26.8	29.0
1971	22.0	24.8	27.7	29.8	29.0	27.6	27.0	27.0	27.3	25.7	23.1	23.0	26.2	28.8
1972	21.9	25.8	27.4	28.9	30.3	28.5	28.1	27.6	27.8	27.3	26.0	24.0	27.0	28.9
1973	22.9	26.5	28.3	31.1	28.9	28.6	28.1	27.3	27.3	26.9	24.3	21.3	26.8	29.4
1974	21.7	24.4	27.7	29.2	28.4	27.7	27.7	27.4	27.2	27.2	25.1	24.0	26.5	28.4
1975	23.5	25.4	28.8	31.0	29.0	28.2	27.6	27.6	27.2	27.0	24.8	20.9	26.8	29.6
1976	20.9	25.3	28.2	30.1	28.1	28.1	28.0	26.9	27.4	27.2	24.7	23.5	26.5	28.8
1977	23.5	24.6	27.5	29.2	28.7	29.3	28.3	28.1	27.1	27.1	24.5	23.6	26.8	28.5
1978	23.7	25.6	28.3	30.6	29.4	28.8	27.4	27.3	27.1	26.6	25.1	23.3	26.9	29.4
1979	25.2	26.8	29.2	30.7	29.8	28.5	28.6	27.3	27.7	26.4	24.3	23.2	27.3	29.9
1980	23.4	25.8	29.2	31.3	30.5	28.3	28.0	27.8	26.9	26.9	25.6	23.9	27.3	30.3
1981	22.6	26.1	28.4	29.7	29.0	27.5	27.3	27.2	27.9	27.4	25.3	21.8	26.7	29.0
1982	22.3	25.4	28.7	29.2	29.4	28.0	27.7	27.0	27.1	27.0	26.1	21.5	26.6	29.1
1983	22.4	26.1	28.7	31.9	30.8	28.9	29.0	28.0	27.6	26.9	23.7	21.8	27.1	30.5
1984	22.7	27.0	28.2	30.4	29.3	27.9	27.6	27.4	27.3	26.3	25.2	23.2	26.9	29.3
1985	24.6	26.4	28.9	30.5	29.1	27.6	27.3	27.6	27.3	26.6	25.4	22.7	27.0	29.5
1986	22.2	25.7	27.2	30.1	28.5	28.4	27.6	28.0	27.7	27.1	25.4	23.5	26.8	28.6
1987	23.9	25.8	28.0	30.2	30.5	28.8	28.5	28.2	27.8	27.6	26.8	19.0	27.1	29.6
1988	24.0	27.2	29.1	30.1	28.9	28.0	28.1	27.8	28.0	26.9	23.5	22.4	27.0	29.4
1989	24.7	25.6	28.2	30.9	29.7	28.1	28.2	28.0	27.8	26.8	25.3	22.0	27.1	29.6
1990	24.8	26.4	28.4	30.3	29.0	28.4	27.8	28.5	27.9	27.1	25.7	23.3	27.3	29.2
1991	24.9	25.9	29.8	31.1	30.9	28.2	28.4	27.6	28.2	27.0	24.5	23.2	27.5	30.6
1992	22.1	23.7	27.7	30.9	30.6	28.8	27.4	27.3	27.2	25.1	23.3	21.9	26.4	29.7
1993	22.5	23.7	27.9	29.6	29.3	28.8	28.5	27.2	27.1	26.4	24.8	22.6	26.5	28.9
1994	24.3	27.1	27.5	30.0	28.5	27.6	26.9	26.5	27.3	25.8	24.5	23.9	26.7	28.7
1995	23.8	24.6	29.1	30.9	28.9	28.5	27.4	27.2	27.2	26.8	24.7	22.0	26.8	29.6
1996	22.5	24.1	28.3	29.3	28.4	27.8	27.5	27.1	26.9	26.5	25.5	22.5	26.4	28.7
1997	22.3	24.5	27.7	28.1	29.9	29.1	27.6	27.3	27.0	27.1	25.6	24.7	26.7	28.6
1998	24.7	26.2	29.2	30.8	30.4	29.4	28.1	28.0	27.7	27.2	25.6	24.1	27.6	30.1
1999	24.5	26.4	28.5	28.9	27.5	27.5	27.6	27.0	27.2	26.4	25.3	20.8	26.5	28.3
2000	24.0	25.0	27.3	28.9	27.8	27.5	27.3	27.5	26.7	27.0	24.5	24.4	26.5	28.0
2001	25.0	25.9	27.3	30.9	27.8	27.8	27.4	27.5	27.5	27.1	23.5	24.0	26.8	28.7

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Annual	MAM
2002	23.3	26.2	28.0	30.1	28.6	28.2	27.5	27.1	26.9	26.6	25.5	24.8	26.9	28.9
2003	23.0	25.4	27.3	30.1	29.4	27.6	27.8	27.6	27.3	27.2	25.9	22.7	26.8	29.0
2004	23.8	25.2	28.9	31.0	28.7	27.6	27.5	27.5	27.3	26.6	25.8	22.3	26.8	29.6
2005	24.0	27.2	28.2	30.3	30.1	28.3	27.9	27.3	27.3	27.6	25.8	23.4	27.3	29.5
2006	23.8	26.3	29.1	29.5	27.7	28.1	27.3	27.2	27.4	27.0	25.9	23.2	26.9	28.8
2007	22.6	24.6	27.8	29.5	26.8	27.9	26.9	26.3	25.9	25.8	23.2	22.4	25.8	28.0
Mean	23.3	25.7	28.4	30.2	29.3	28.2	27.7	27.4	27.3	26.8	25.1	22.9	26.9	29.3

MAM: March-April-May.

B2: Monthly and annual rainfall (mm) averaged over the 50 selected stations

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Annual
1950	0.0	0.0	0.0	5.5	280.8	266.0	105.1	237.7	265.4	445.2	0.0	0.0	1605.6
1951	16.8	2.3	12.6	20.9	207.2	118.8	215.2	184.1	276.4	244.9	81.6	5.2	1386.0
1952	0.0	5.9	68.8	14.9	130.6	165.6	150.0	233.6	223.0	163.8	16.6	1.4	1174.2
1953	22.6	86.3	12.9	71.6	142.8	203.2	223.9	191.3	306.5	151.9	79.9	3.4	1496.4
1954	1.4	9.9	31.9	33.2	195.9	118.8	99.4	213.1	254.5	137.4	13.4	2.3	1111.3
1955	1.4	13.9	31.7	98.5	139.7	185.5	121.5	199.6	216.6	63.5	33.8	0.0	1105.7
1956	0.2	21.6	11.0	69.4	234.0	106.8	209.0	180.4	270.0	114.4	11.0	2.3	1230.1
1957	5.5	7.9	28.3	54.3	80.9	185.1	162.3	198.3	243.7	105.6	11.6	0.0	1083.4
1958	19.4	4.4	35.6	45.9	95.2	171.3	157.3	174.5	265.0	107.8	4.0	0.0	1080.5
1959	6.3	9.0	18.3	48.5	150.4	113.1	188.4	152.4	300.4	55.9	7.4	1.5	1051.5
1960	28.7	0.4	4.8	21.3	134.8	120.0	153.5	238.9	217.5	173.2	30.9	35.1	1159.1
1961	0.6	12.1	31.0	63.8	234.0	123.6	150.8	232.9	232.1	190.1	8.4	24.1	1303.4
1962	2.6	0.8	11.5	41.2	117.8	100.4	190.2	200.2	322.3	192.9	3.9	1.7	1185.4
1963	0.2	4.2	11.4	35.3	58.9	163.0	130.5	228.0	238.2	241.6	78.1	9.9	1199.4
1964	1.7	4.3	9.5	50.8	237.9	100.4	190.1	141.6	281.7	187.9	23.0	7.3	1236.2
1965	0.4	17.7	24.8	21.2	128.3	140.3	86.0	213.7	212.8	171.1	37.6	18.3	1072.1
1966	22.2	9.0	17.8	21.7	237.5	87.7	137.4	245.4	183.4	175.3	28.3	12.3	1178.1
1967	3.9	5.2	5.4	62.4	162.8	101.7	140.5	175.0	315.0	90.9	60.2	0.5	1123.5
1968	3.9	7.9	12.6	109.5	127.0	154.2	131.5	130.2	157.2	104.7	21.3	0.6	960.7
1969	17.8	0.0	11.3	38.3	168.9	129.6	132.3	216.5	273.7	100.6	23.7	6.8	1119.4
1970	2.2	3.8	53.9	69.6	240.4	168.1	156.0	249.8	223.4	123.2	22.5	42.6	1355.5
1971	3.2	10.6	22.5	55.8	207.1	135.3	174.8	294.1	167.0	116.1	14.3	14.1	1214.9
1972	0.6	5.2	13.5	98.3	71.1	126.8	97.6	166.7	224.6	138.2	117.7	18.6	1078.8
1973	0.0	0.7	62.7	12.2	190.3	146.2	162.9	275.3	292.2	60.2	22.1	0.1	1225.0
1974	0.4	1.8	31.8	97.3	160.1	99.2	142.7	154.8	250.1	202.3	79.4	2.5	1222.5
1975	87.7	4.8	16.4	24.9	152.5	150.0	177.4	250.5	225.3	190.2	46.2	29.2	1355.1
1976	0.1	3.6	16.0	41.4	156.9	103.4	124.3	222.5	187.8	189.8	25.6	10.6	1082.1
1977	42.9	0.6	23.0	82.3	124.8	67.9	122.5	178.4	251.5	129.8	11.6	32.4	1067.7
1978	24.1	35.1	1.5	31.6	178.6	118.8	333.9	168.8	277.1	105.7	5.6	2.6	1283.5
1979	0.8	2.6	3.0	48.9	152.4	172.0	110.1	127.2	223.0	69.7	0.0	0.0	909.8
1980	0.0	3.0	22.5	33.7	226.2	179.9	174.6	152.7	260.2	151.1	28.5	24.9	1257.3
1981	0.0	8.5	19.1	65.7	186.9	126.5	210.3	166.0	175.1	99.3	134.2	13.6	1205.2
1982	1.0	1.8	8.2	59.8	151.2	117.3	86.9	124.4	268.4	105.0	26.2	2.1	952.3
1983	6.7	0.2	2.6	11.7	130.9	123.2	143.1	211.3	207.9	252.5	113.5	11.7	1215.3
1984	0.8	20.7	5.5	52.4	115.0	131.6	133.2	113.2	191.6	125.7	5.6	0.4	895.9
1985	7.6	6.7	5.3	77.5	159.9	125.1	136.1	127.9	254.0	164.1	124.8	1.1	1190.0
1986	0.0	1.5	5.1	67.4	189.4	109.8	125.2	172.7	158.3	105.5	25.1	25.7	985.8
1987	0.8	7.5	28.7	69.5	70.9	145.9	64.7	223.9	262.6	87.3	96.8	0.0	1058.6
1988	0.3	10.3	9.9	102.7	192.0	213.7	169.6	190.6	168.1	206.5	44.6	0.3	1308.6
1989	6.3	1.4	12.1	13.0	150.9	110.2	154.5	146.0	157.5	195.3	11.1	0.0	958.5
1990	1.4	9.9	27.1	30.0	239.0	89.2	84.6	147.9	176.5	186.2	35.0	0.1	1026.8
1991	1.3	0.1	14.6	43.9	93.0	118.3	76.3	241.4	147.3	135.7	21.9	5.1	899.0
1992	3.7	23.3	0.0	17.2	55.2	102.9	192.4	178.7	215.6	166.3	13.4	57.0	1025.5
1993	0.8	0.2	22.2	51.9	135.0	67.4	101.3	122.2	229.8	112.0	0.1	0.2	843.0
1994	0.1	1.7	118.3	50.1	212.2	175.5	125.7	244.5	193.4	70.2	20.8	17.1	1229.3
1995	0.5	0.0	18.3	36.9	164.5	111.1	216.0	262.0	244.5	85.9	41.8	0.2	1181.7
1996	0.6	39.4	10.4	94.9	131.2	172.6	106.5	191.7	247.6	124.0	88.3	2.6	1209.9
1997	1.1	0.8	30.7	57.0	69.3	38.5	146.2	169.5	212.1	112.3	12.6	0.1	850.0
1998	2.9	12.1	12.8	30.9	147.0	102.3	143.8	177.1	153.8	82.6	48.0	3.0	916.3
1999	14.4	23.0	28.2	111.4	292.9	119.3	124.2	173.7	199.7	220.7	68.2	8.5	1384.1
2000	1.5	31.7	31.3	118.4	205.3	156.0	99.9	146.1	184.5	190.8	3.7	5.6	1174.6
2001	3.5	1.0	77.9	13.9	228.6	106.2	147.3	214.5	152.9	165.8	14.9	8.9	1135.3

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Annual
2002	8.3	10.3	14.2	39.8	205.0	124.1	87.8	219.9	348.4	134.7	128.9	45.2	1366.6
2003	10.3	9.2	60.9	40.6	104.4	146.1	124.1	146.1	241.2	47.0	2.7	0.0	932.6
2004	6.2	15.7	2.4	31.6	217.4	144.7	167.1	122.1	239.1	26.4	13.0	0.3	986.1
2005	0.2	2.3	27.9	67.8	91.5	172.1	155.4	141.9	304.1	115.5	62.0	16.6	1157.5
2006	0.0	8.9	15.7	162.5	226.7	182.6	190.3	224.5	280.4	144.3	5.1	0.0	1441.0
2007	0.5	0.0	6.8	49.2	317.6	145.9	110.0	159.4	251.7	171.0	39.5	0.0	1251.6
Mean	6.9	9.4	22.0	53.3	165.7	134.5	146.1	189.4	232.8	143.6	37.1	9.2	1149.9

B3: Seasonal rainfall (mm) and its classification during 1950 to 2007

Year	MJJ	Classification	ASO	Classification	NDJ	Classification	FMA	Classification
1950	651.9	W	948.3	W	16.8	N	5.5	D
1951	541.2	W	705.4	W	86.8	W	35.8	D
1952	446.2	N	620.4	N	40.6	N	89.6	N
1953	569.9	W	649.7	W	84.7	W	170.8	W
1954	414.1	N	605.0	N	17.1	N	75.0	N
1955	446.7	N	479.7	D	34.0	N	144.1	W
1956	549.8	W	564.8	N	18.8	N	102.0	N
1957	428.3	N	547.6	N	31.0	N	90.5	N
1958	423.8	N	547.3	N	10.3	D	85.9	N
1959	451.9	N	508.7	N	37.6	N	75.8	N
1960	408.5	N	629.6	W	66.6	N	26.5	D
1961	508.4	N	655.1	W	35.1	N	106.9	N
1962	408.5	N	715.4	W	5.8	D	53.5	D
1963	352.4	D	707.8	W	89.7	W	50.9	D
1964	528.4	N	611.2	N	30.7	N	64.6	N
1965	354.6	D	597.6	N	78.1	N	63.7	N
1966	462.6	N	604.1	N	44.5	N	48.5	D
1967	405.0	N	580.9	N	64.6	N	73.0	N
1968	412.7	N	392.1	D	39.7	N	130.0	W
1969	430.8	N	590.8	N	32.7	N	49.6	D
1970	564.5	W	596.4	N	68.3	N	127.3	W
1971	517.2	N	577.2	N	29.0	N	88.9	N
1972	295.5	D	529.5	N	136.3	W	117.0	W
1973	499.5	N	627.7	W	22.6	N	75.6	N
1974	402.0	N	607.2	N	169.6	W	130.9	W
1975	480.0	N	666.0	W	75.5	N	46.1	D
1976	384.6	N	600.1	N	79.1	W	61.0	N
1977	315.2	D	559.7	N	68.1	N	105.9	N
1978	631.3	W	551.6	N	9.0	D	68.2	N
1979	434.5	N	419.9	D	0.0	D	54.5	N
1980	580.7	W	564.0	N	53.4	N	59.3	N
1981	523.7	N	440.4	D	148.8	W	93.3	N
1982	355.4	D	497.8	D	35.0	N	69.8	N
1983	397.2	N	671.7	W	126.0	W	14.5	D
1984	379.8	D	430.5	D	13.6	D	78.6	N
1985	421.1	N	546.0	N	125.9	W	89.5	N
1986	424.4	N	436.5	D	51.6	N	74.0	N
1987	281.5	D	573.8	N	97.1	W	105.7	N
1988	575.3	W	565.2	N	51.2	N	122.9	W
1989	415.6	N	498.8	N	12.5	D	26.5	D
1990	412.8	N	510.6	N	36.4	N	67.0	N
1991	287.6	D	524.4	N	30.7	N	58.6	N
1992	350.5	D	560.6	N	71.2	N	40.5	D
1993	303.7	D	464.0	D	0.4	D	74.3	N
1994	513.4	N	508.1	N	38.4	N	170.1	W
1995	491.6	N	592.4	N	42.6	N	55.2	N
1996	410.3	N	563.3	N	92.0	W	144.7	W
1997	254.0	D	493.9	D	15.6	D	88.5	N
1998	393.1	N	413.5	D	65.4	N	55.8	N

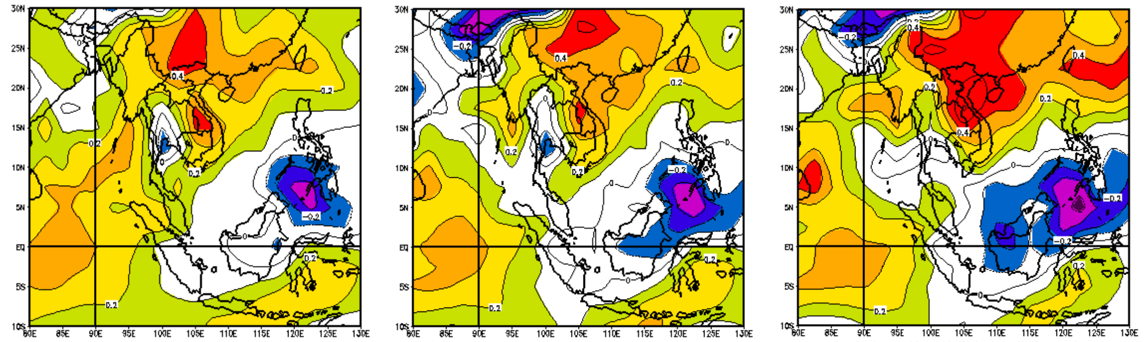
Year	MJJ	Classification	ASO	Classification	NDJ	Classification	FMA	Classification
1999	536.4	W	594.1	N	78.2	N	162.6	W
2000	461.2	N	521.4	N	12.8	D	181.4	W
2001	482.1	N	533.2	N	32.1	N	92.8	N
2002	416.9	N	703.0	W	184.4	W	64.3	N
2003	374.6	D	434.3	D	8.9	D	110.7	N
2004	529.2	W	387.6	D	13.5	D	49.7	D
2005	419.1	N	561.5	N	78.6	N	98.0	N
2006	599.6	W	649.2	W	5.6	D	187.1	W
2007	573.5	W	581.0	N	n/a	n/a	56.0	N
Mean	446.3		565.8		53.4		84.6	
P(20th)	381.7		498.2		15.8		53.9	
P(80th)	528.9		624.8		79.0		114.5	

MJJ: May-June-July; ASO: August-September-October; NDJ: November-December-January; FMA: February-March-April.

W: wet; N: normal; D: dry.

Appendix C: Correlation Maps between Rainfall and Large-Scale Atmospheric Variables

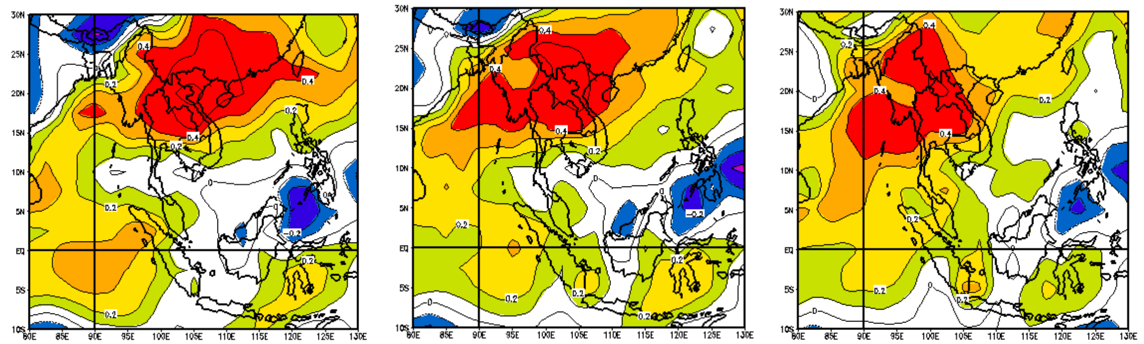
C1: Correlation maps between MJJ rainfall and SAT at lead times varying from 15 to 4 months



Feb to Apr; 1980 to 2007: Surface Air Temperature
Seasonal Correlation w/ May to Jul average50.txt (index lags by 15 months

Mar to May; 1980 to 2007: Surface Air Temperature
Seasonal Correlation w/ May to Jul average50.txt (index lags by 14 months)

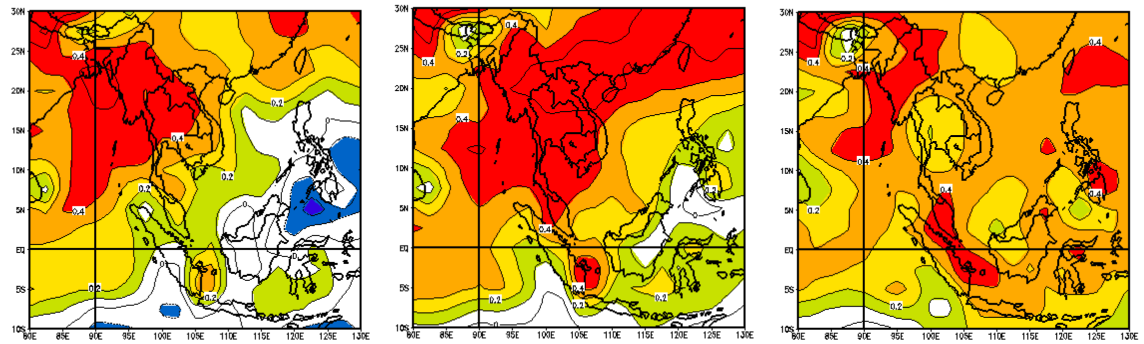
Apr to Jun; 1980 to 2007: Surface Air Temperature
Seasonal Correlation w/ May to Jul average50.txt (index lags by 13 months)



May to Jul; 1980 to 2007: Surface Air Temperature
Seasonal Correlation w/ May to Jul average50.txt (index lags by 12 months)

Jun to Aug; 1980 to 2007: Surface Air Temperature
Seasonal Correlation w/ May to Jul average50.txt (index lags by 11 months

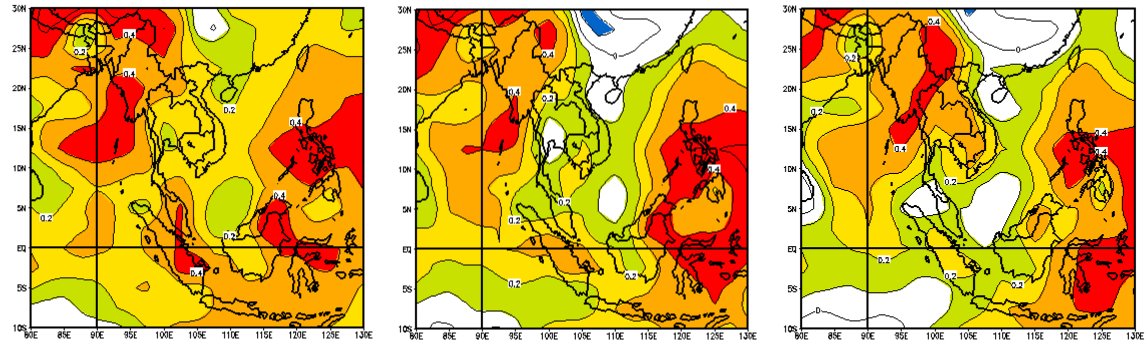
Jul to Sep; 1980 to 2007: Surface Air Temperature
Seasonal Correlation w/ May to Jul average50.txt (index lags by 10 months)



Aug to Oct; 1980 to 2007: Surface Air Temperature
Seasonal Correlation w/ May to Jul average50.txt (index lags by 9 months)

Sep to Nov; 1980 to 2007: Surface Air Temperature
Seasonal Correlation w/ May to Jul average50.txt (index lags by 8 months)

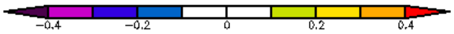
Oct to Dec; 1980 to 2007: Surface Air Temperature
Seasonal Correlation w/ May to Jul average50.txt (index lags by 7 months)



Nov to Jan; 1980 to 2007: Surface Air Temperature
Seasonal Correlation w/ May to Jul average50.txt (index lags by 6 months)

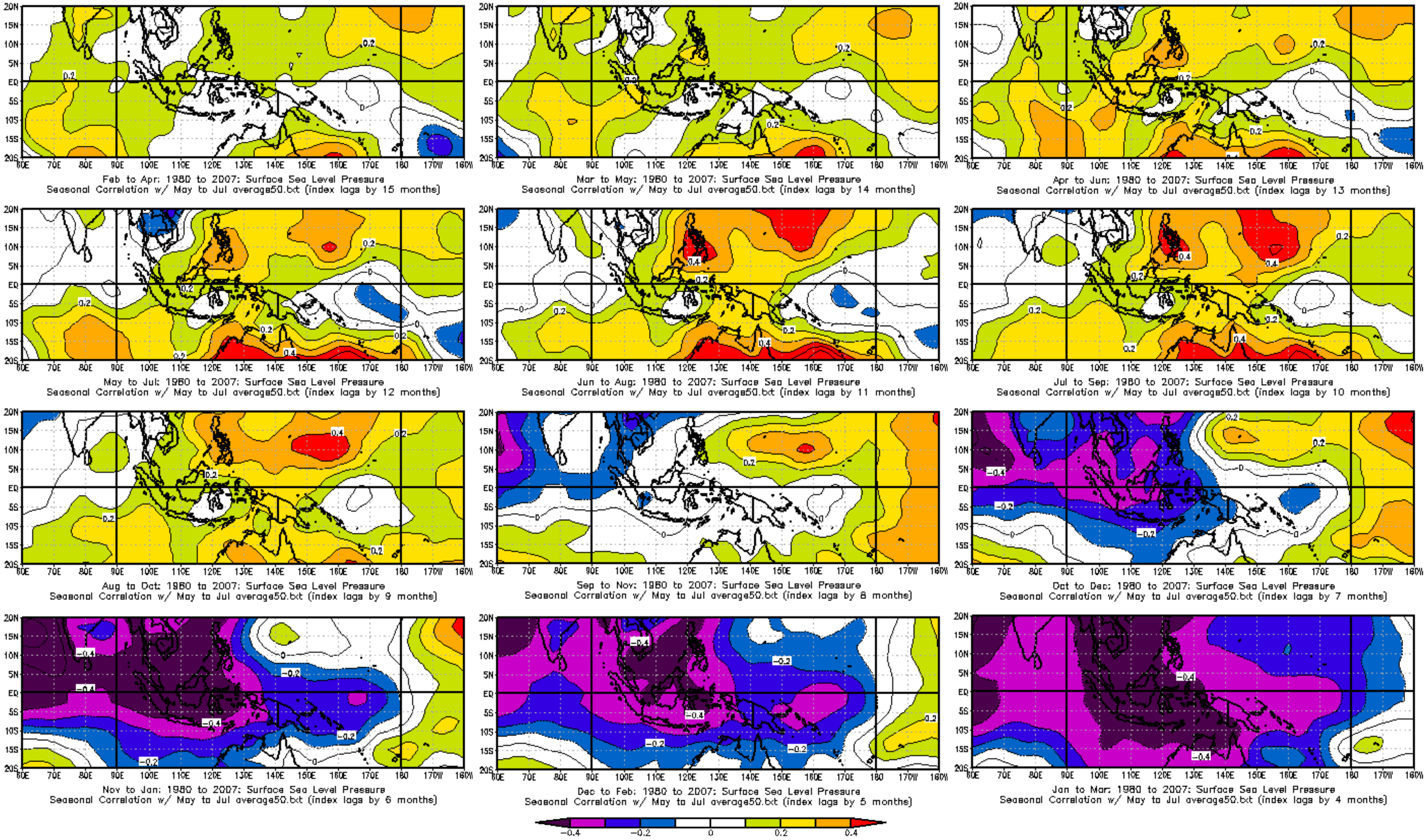
Dec to Feb; 1980 to 2007: Surface Air Temperature
Seasonal Correlation w/ May to Jul average50.txt (index lags by 5 months)

Jan to Mar; 1980 to 2007: Surface Air Temperature
Seasonal Correlation w/ May to Jul average50.txt (index lags by 4 months)



Note: The upper and lower bounds of 95% significant levels of correlation are +0.3 and -0.3, respectively.

C2: Correlation maps between MJJ rainfall and SLP at lead times varying from 15 to 4 months



Feb to Apr: 1980 to 2007: Surface Sea Level Pressure
Seasonal Correlation w/ May to Jul average50.txt (index lags by 15 months)

Mar to May: 1980 to 2007: Surface Sea Level Pressure
Seasonal Correlation w/ May to Jul average50.txt (index lags by 14 months)

Apr to Jun: 1980 to 2007: Surface Sea Level Pressure
Seasonal Correlation w/ May to Jul average50.txt (index lags by 13 months)

May to Jul: 1980 to 2007: Surface Sea Level Pressure
Seasonal Correlation w/ May to Jul average50.txt (index lags by 12 months)

Jun to Aug: 1980 to 2007: Surface Sea Level Pressure
Seasonal Correlation w/ May to Jul average50.txt (index lags by 11 months)

Jul to Sep: 1980 to 2007: Surface Sea Level Pressure
Seasonal Correlation w/ May to Jul average50.txt (index lags by 10 months)

Aug to Oct: 1980 to 2007: Surface Sea Level Pressure
Seasonal Correlation w/ May to Jul average50.txt (index lags by 9 months)

Sep to Nov: 1980 to 2007: Surface Sea Level Pressure
Seasonal Correlation w/ May to Jul average50.txt (index lags by 8 months)

Oct to Dec: 1980 to 2007: Surface Sea Level Pressure
Seasonal Correlation w/ May to Jul average50.txt (index lags by 7 months)

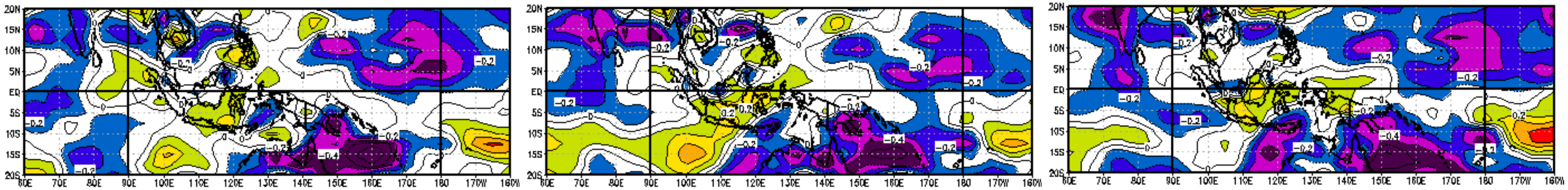
Nov to Jan: 1980 to 2007: Surface Sea Level Pressure
Seasonal Correlation w/ May to Jul average50.txt (index lags by 6 months)

Dec to Feb: 1980 to 2007: Surface Sea Level Pressure
Seasonal Correlation w/ May to Jul average50.txt (index lags by 5 months)

Jan to Mar: 1980 to 2007: Surface Sea Level Pressure
Seasonal Correlation w/ May to Jul average50.txt (index lags by 4 months)

Note: The upper and lower bounds of 95% significant levels of correlation are +0.3 and -0.3, respectively.

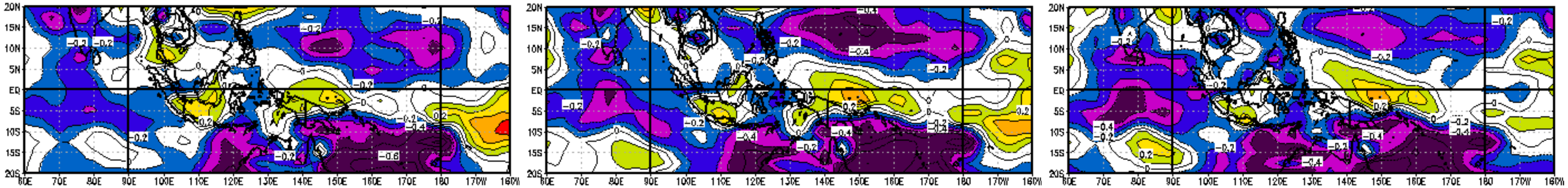
C3: Correlation maps between MJJ rainfall and SXW at lead times varying from 15 to 4 months



Feb to Apr: 1980 to 2007: Surface Zonal Wind
Seasonal Correlation w/ May to Jul average50.txt (index lags by 15 months)

Mar to May: 1980 to 2007: Surface Zonal Wind
Seasonal Correlation w/ May to Jul average50.txt (index lags by 14 months)

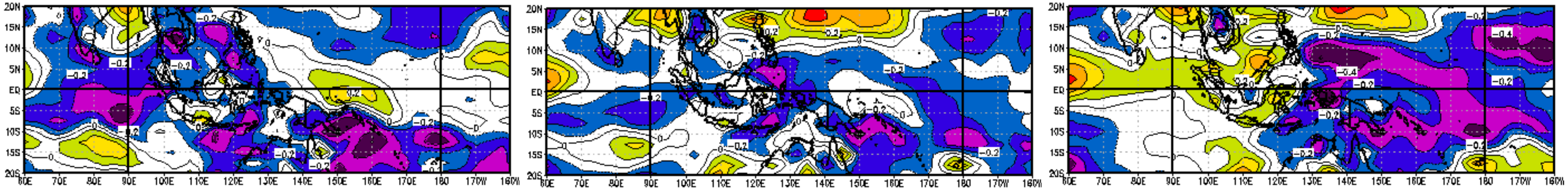
Apr to Jun: 1980 to 2007: Surface Zonal Wind
Seasonal Correlation w/ May to Jul average50.txt (index lags by 13 months)



May to Jul: 1980 to 2007: Surface Zonal Wind
Seasonal Correlation w/ May to Jul average50.txt (index lags by 12 months)

Jun to Aug: 1980 to 2007: Surface Zonal Wind
Seasonal Correlation w/ May to Jul average50.txt (index lags by 11 months)

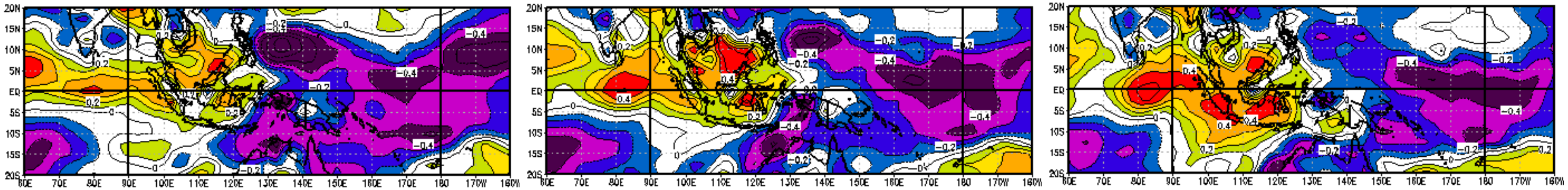
Jul to Sep: 1980 to 2007: Surface Zonal Wind
Seasonal Correlation w/ May to Jul average50.txt (index lags by 10 months)



Aug to Oct: 1980 to 2007: Surface Zonal Wind
Seasonal Correlation w/ May to Jul average50.txt (index lags by 9 months)

Sep to Nov: 1980 to 2007: Surface Zonal Wind
Seasonal Correlation w/ May to Jul average50.txt (index lags by 8 months)

Oct to Dec: 1980 to 2007: Surface Zonal Wind
Seasonal Correlation w/ May to Jul average50.txt (index lags by 7 months)



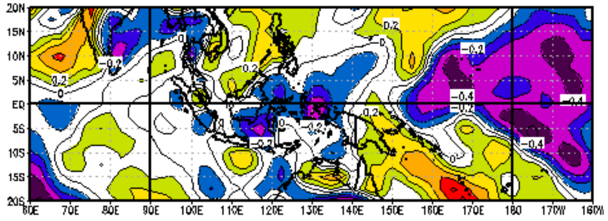
Nov to Jan: 1980 to 2007: Surface Zonal Wind
Seasonal Correlation w/ May to Jul average50.txt (index lags by 6 months)

Dec to Feb: 1980 to 2007: Surface Zonal Wind
Seasonal Correlation w/ May to Jul average50.txt (index lags by 5 months)

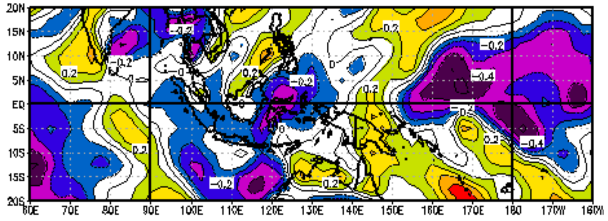
Jan to Mar: 1980 to 2007: Surface Zonal Wind
Seasonal Correlation w/ May to Jul average50.txt (index lags by 4 months)

Note: The upper and lower bounds of 95% significant levels of correlation are +0.3 and -0.3, respectively.

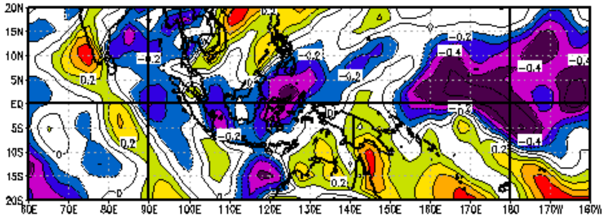
C4: Correlation maps between MJJ rainfall and SYW at lead times varying from 15 to 4 months



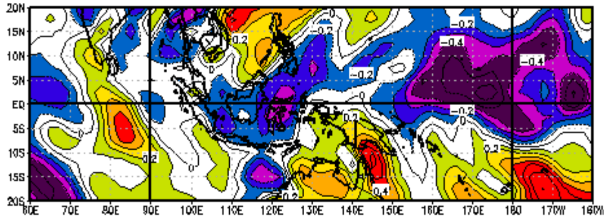
Feb to Apr: 1980 to 2007: Surface Meridional Wind
Seasonal Correlation w/ May to Jul average50.txt (index lags by 15 months)



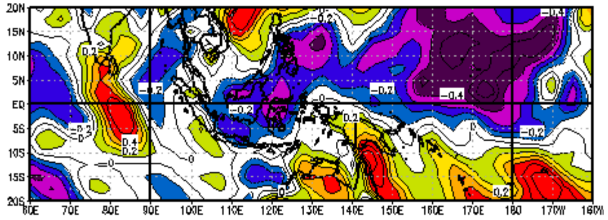
Mar to May: 1980 to 2007: Surface Meridional Wind
Seasonal Correlation w/ May to Jul average50.txt (index lags by 14 months)



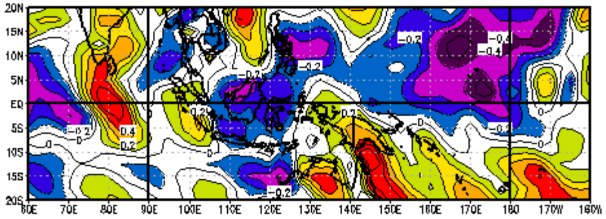
Apr to Jun: 1980 to 2007: Surface Meridional Wind
Seasonal Correlation w/ May to Jul average50.txt (index lags by 13 months)



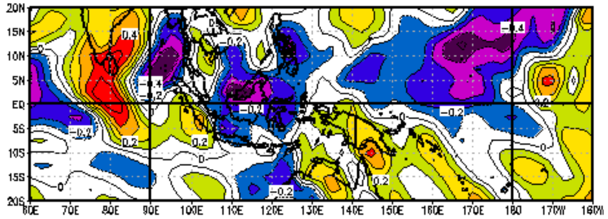
May to Jul: 1980 to 2007: Surface Meridional Wind
Seasonal Correlation w/ May to Jul average50.txt (index lags by 12 months)



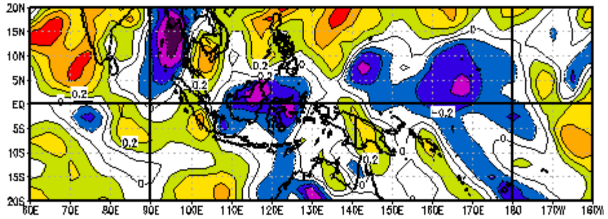
Jun to Aug: 1980 to 2007: Surface Meridional Wind
Seasonal Correlation w/ May to Jul average50.txt (index lags by 11 months)



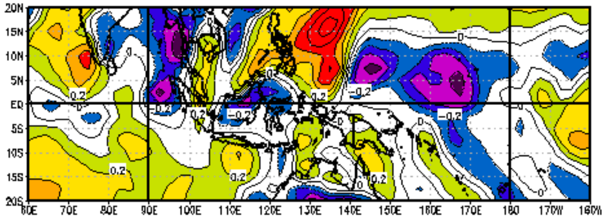
Jul to Sep: 1980 to 2007: Surface Meridional Wind
Seasonal Correlation w/ May to Jul average50.txt (index lags by 10 months)



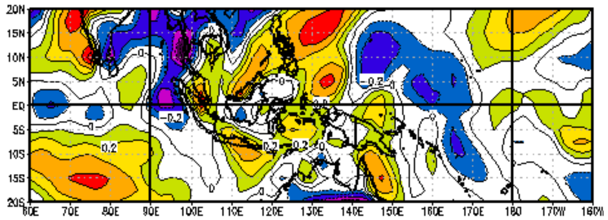
Aug to Oct: 1980 to 2007: Surface Meridional Wind
Seasonal Correlation w/ May to Jul average50.txt (index lags by 9 months)



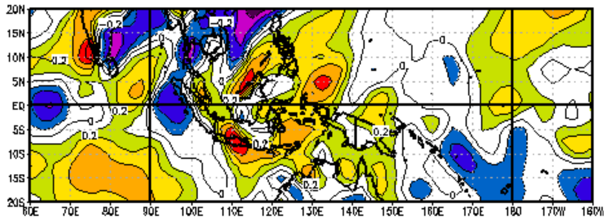
Sep to Nov: 1980 to 2007: Surface Meridional Wind
Seasonal Correlation w/ May to Jul average50.txt (index lags by 8 months)



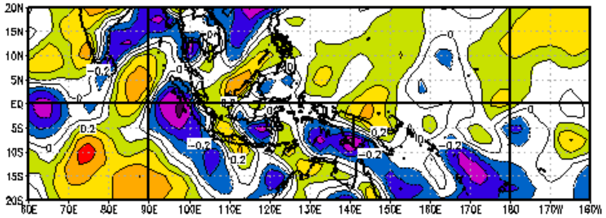
Oct to Dec: 1980 to 2007: Surface Meridional Wind
Seasonal Correlation w/ May to Jul average50.txt (index lags by 7 months)



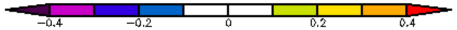
Nov to Jan: 1980 to 2007: Surface Meridional Wind
Seasonal Correlation w/ May to Jul average50.txt (index lags by 6 months)



Dec to Feb: 1980 to 2007: Surface Meridional Wind
Seasonal Correlation w/ May to Jul average50.txt (index lags by 5 months)

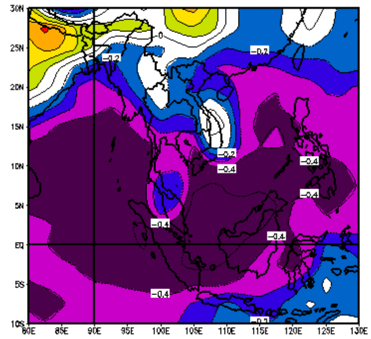


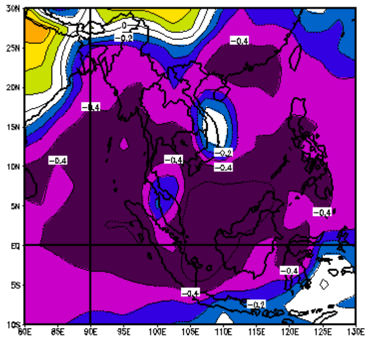
Jan to Mar: 1980 to 2007: Surface Meridional Wind
Seasonal Correlation w/ May to Jul average50.txt (index lags by 4 months)

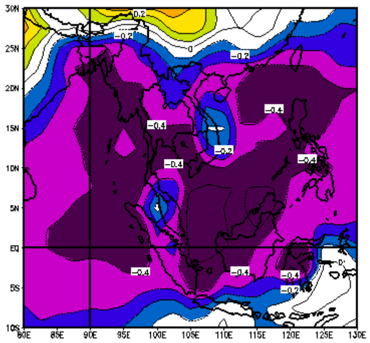


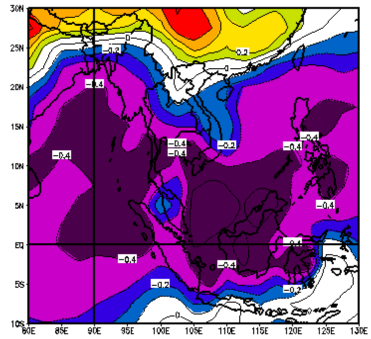
Note: The upper and lower bounds of 95% significant levels of correlation are +0.3 and -0.3, respectively.

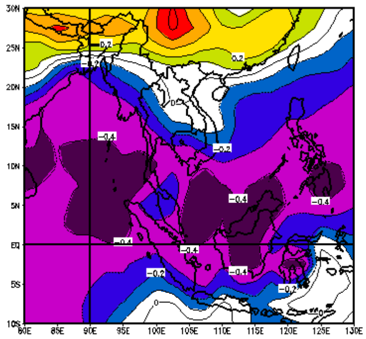
C5: Correlation maps between ASO rainfall and SAT at lead times varying from 15 to 4 months

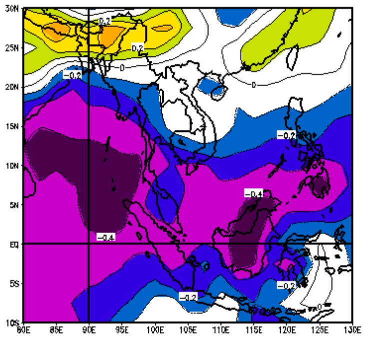

May to Jul; 1950 to 2007: Surface Air Temperature
Seasonal Correlation w/ Aug to Oct average50.txt (index lags by 15 months)

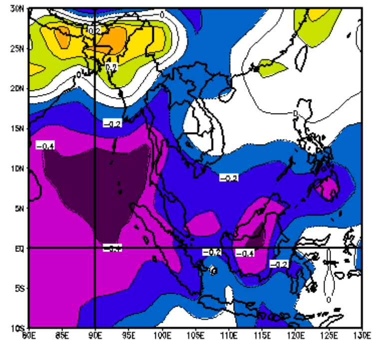

Jun to Aug; 1950 to 2007: Surface Air Temperature
Seasonal Correlation w/ Aug to Oct average50.txt (index lags by 14 months)

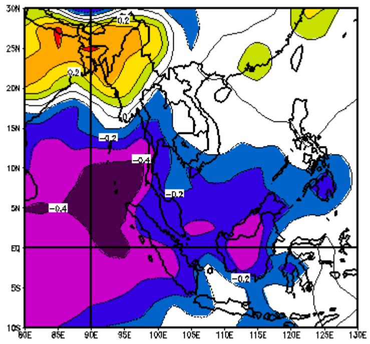

Jul to Sep; 1950 to 2007: Surface Air Temperature
Seasonal Correlation w/ Aug to Oct average50.txt (index lags by 13 months)

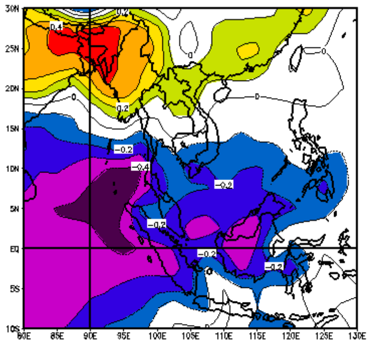

Aug to Oct; 1950 to 2007: Surface Air Temperature
Seasonal Correlation w/ Aug to Oct average50.txt (index lags by 12 months)

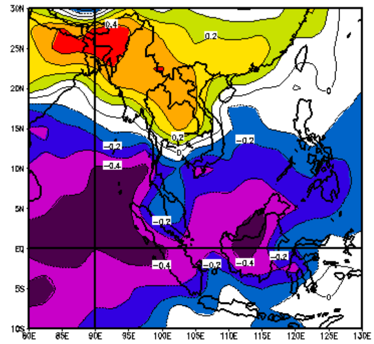

Sep to Nov; 1950 to 2007: Surface Air Temperature
Seasonal Correlation w/ Aug to Oct average50.txt (index lags by 11 months)

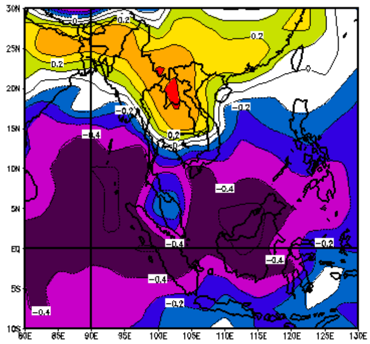

Oct to Dec; 1950 to 2007: Surface Air Temperature
Seasonal Correlation w/ Aug to Oct average50.txt (index lags by 10 months)

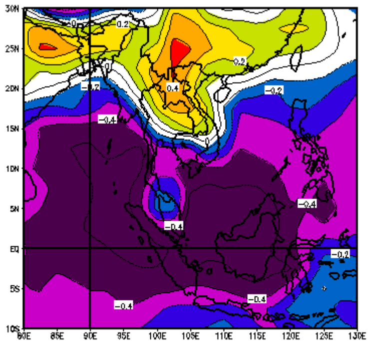

Nov to Jan; 1950 to 2006: Surface Air Temperature
Seasonal Correlation w/ Aug to Oct average50.txt (index lags by 9 months)

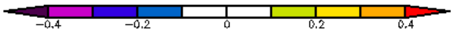

Dec to Feb; 1950 to 2006: Surface Air Temperature
Seasonal Correlation w/ Aug to Oct average50.txt (index lags by 8 months)


Jan to Mar; 1950 to 2007: Surface Air Temperature
Seasonal Correlation w/ Aug to Oct average50.txt (index lags by 7 months)


Feb to Apr; 1950 to 2007: Surface Air Temperature
Seasonal Correlation w/ Aug to Oct average50.txt (index lags by 6 months)

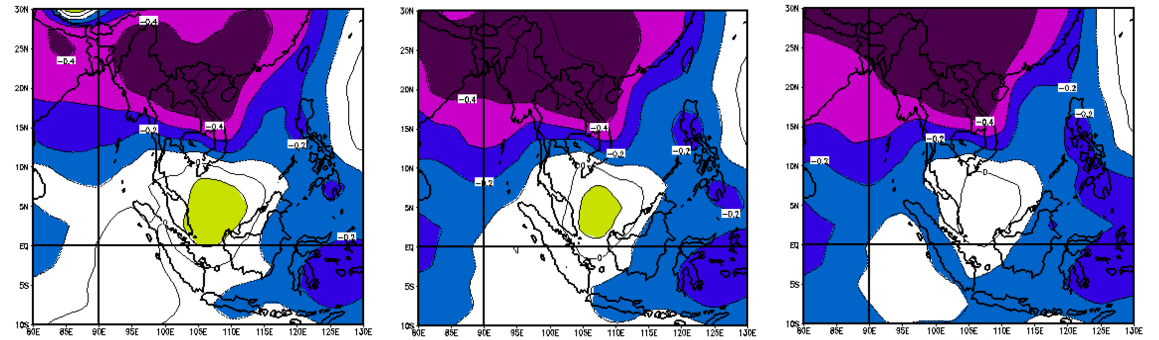

Mar to May; 1950 to 2007: Surface Air Temperature
Seasonal Correlation w/ Aug to Oct average50.txt (index lags by 5 months)


Apr to Jun; 1950 to 2007: Surface Air Temperature
Seasonal Correlation w/ Aug to Oct average50.txt (index lags by 4 months)



Note: The upper and lower bounds of 95% significant levels of correlation are +0.3 and -0.3, respectively.

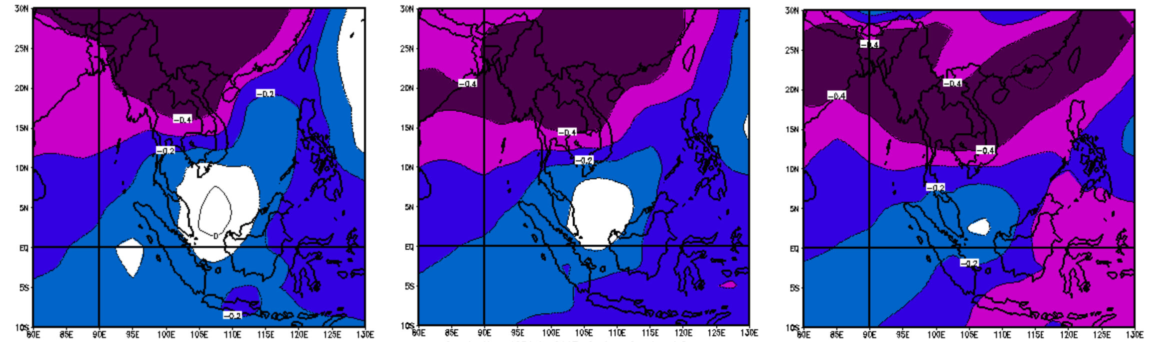
C6: Correlation maps between ASO rainfall and SLP at lead times varying from 15 to 4 months



May to Jul: 1950 to 2007: Surface Sea Level Pressure
Seasonal Correlation w/ Aug to Oct average50.txt (index lags by 15 months)

Jun to Aug: 1950 to 2007: Surface Sea Level Pressure
Seasonal Correlation w/ Aug to Oct average50.txt (index lags by 14 months)

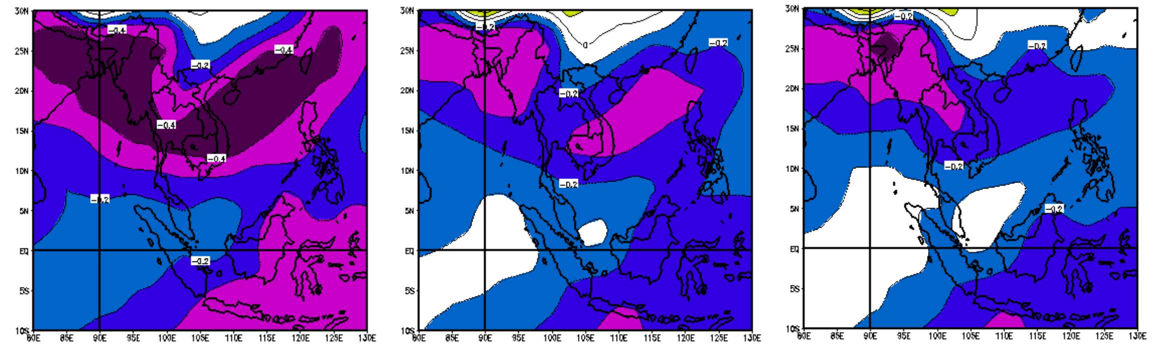
Jul to Sep: 1950 to 2007: Surface Sea Level Pressure
Seasonal Correlation w/ Aug to Oct average50.txt (index lags by 13 months)



Aug to Oct: 1950 to 2007: Surface Sea Level Pressure
Seasonal Correlation w/ Aug to Oct average50.txt (index lags by 12 months

Sep to Nov: 1950 to 2007: Surface Sea Level Pressure
Seasonal Correlation w/ Aug to Oct average50.txt (index lags by 11 months

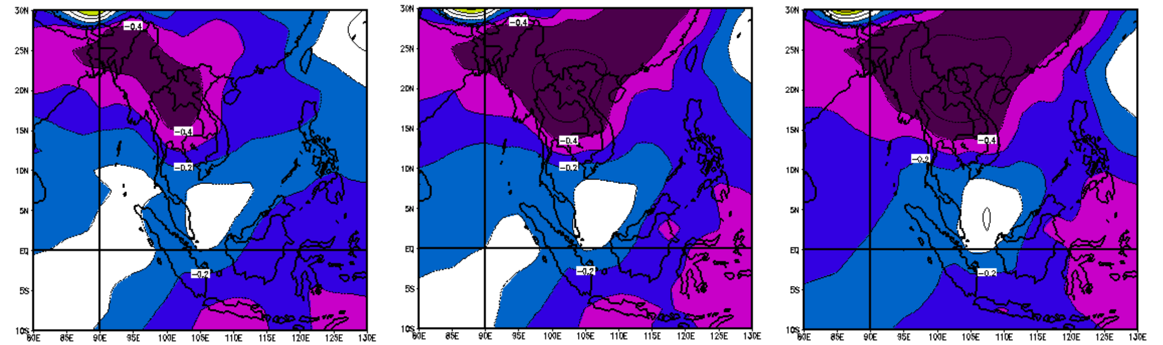
Oct to Dec: 1950 to 2007: Surface Sea Level Pressure
Seasonal Correlation w/ Aug to Oct average50.txt (index lags by 10 months



Nov to Jan: 1950 to 2006: Surface Sea Level Pressure
Seasonal Correlation w/ Aug to Oct average50.txt (index lags by 9 months)

Dec to Feb: 1950 to 2006: Surface Sea Level Pressure
Seasonal Correlation w/ Aug to Oct average50.txt (index lags by 8 months)

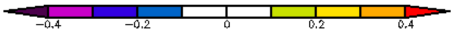
Jan to Mar: 1950 to 2007: Surface Sea Level Pressure
Seasonal Correlation w/ Aug to Oct average50.txt (index lags by 7 months)



Feb to Apr: 1950 to 2007: Surface Sea Level Pressure
Seasonal Correlation w/ Aug to Oct average50.txt (index lags by 6 months)

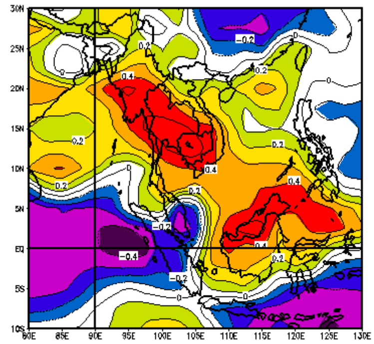
Mar to May: 1950 to 2007: Surface Sea Level Pressure
Seasonal Correlation w/ Aug to Oct average50.txt (index lags by 5 months)

Apr to Jun: 1950 to 2007: Surface Sea Level Pressure
Seasonal Correlation w/ Aug to Oct average50.txt (index lags by 4 months)

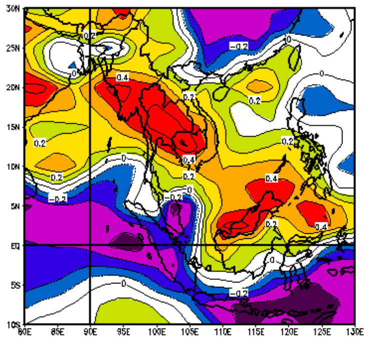


Note: The upper and lower bounds of 95% significant levels of correlation are +0.3 and -0.3, respectively.

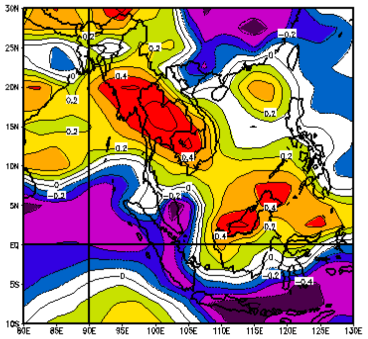
C7: Correlation maps between ASO rainfall and SXW at lead times varying from 15 to 4 months



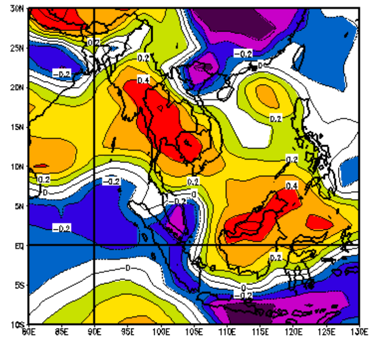
May to Jul: 1950 to 2007: Surface Zonal Wind
Seasonal Correlation w/ Aug to Oct average50.txt (index lags by 15 months



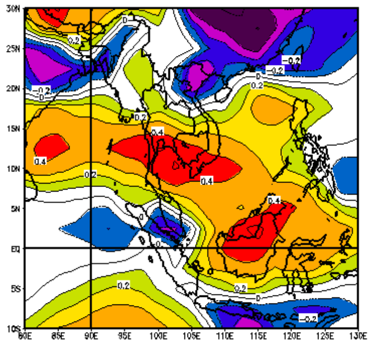
Jun to Aug: 1950 to 2007: Surface Zonal Wind
Seasonal Correlation w/ Aug to Oct average50.txt (index lags by 14 months



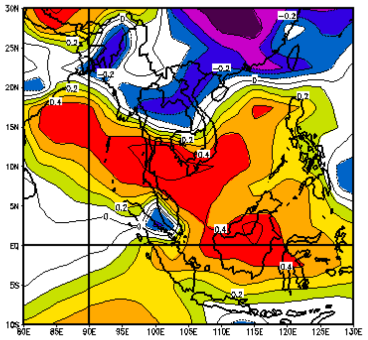
Jul to Sep: 1950 to 2007: Surface Zonal Wind
Seasonal Correlation w/ Aug to Oct average50.txt (index lags by 13 months)



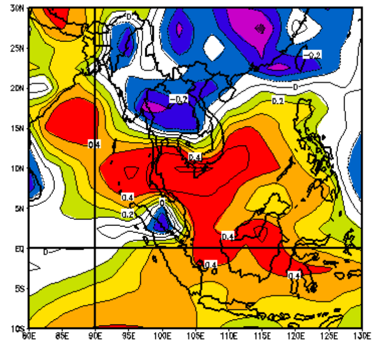
Aug to Oct: 1950 to 2007: Surface Zonal Wind
Seasonal Correlation w/ Aug to Oct average50.txt (index lags by 12 months)



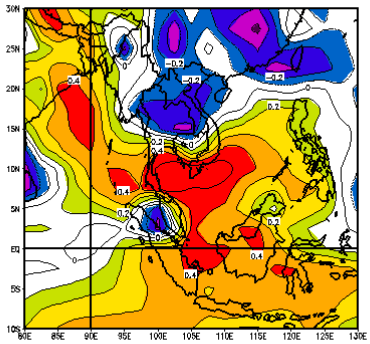
Sep to Nov: 1950 to 2007: Surface Zonal Wind
Seasonal Correlation w/ Aug to Oct average50.txt (index lags by 11 months



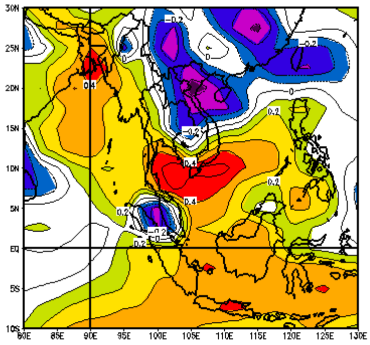
Oct to Dec: 1950 to 2007: Surface Zonal Wind
Seasonal Correlation w/ Aug to Oct average50.txt (index lags by 10 months)



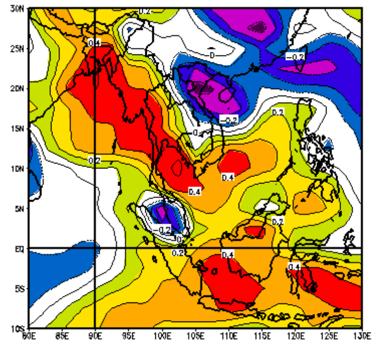
Nov to Jan: 1950 to 2006: Surface Zonal Wind
Seasonal Correlation w/ Aug to Oct average50.txt (index lags by 9 months)



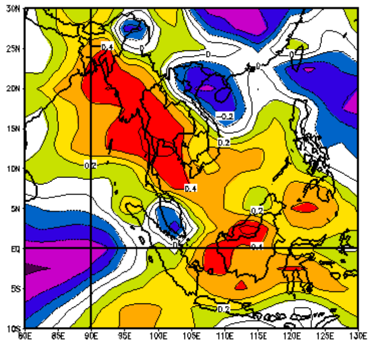
Dec to Feb: 1950 to 2006: Surface Zonal Wind
Seasonal Correlation w/ Aug to Oct average50.txt (index lags by 8 months)



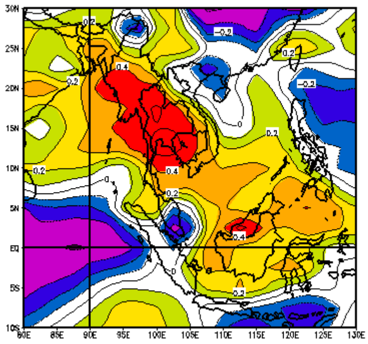
Jan to Mar: 1950 to 2007: Surface Zonal Wind
Seasonal Correlation w/ Aug to Oct average50.txt (index lags by 7 months)



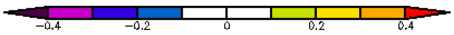
Feb to Apr: 1950 to 2007: Surface Zonal Wind
Seasonal Correlation w/ Aug to Oct average50.txt (index lags by 6 months)



Mar to May: 1950 to 2007: Surface Zonal Wind
Seasonal Correlation w/ Aug to Oct average50.txt (index lags by 5 months)

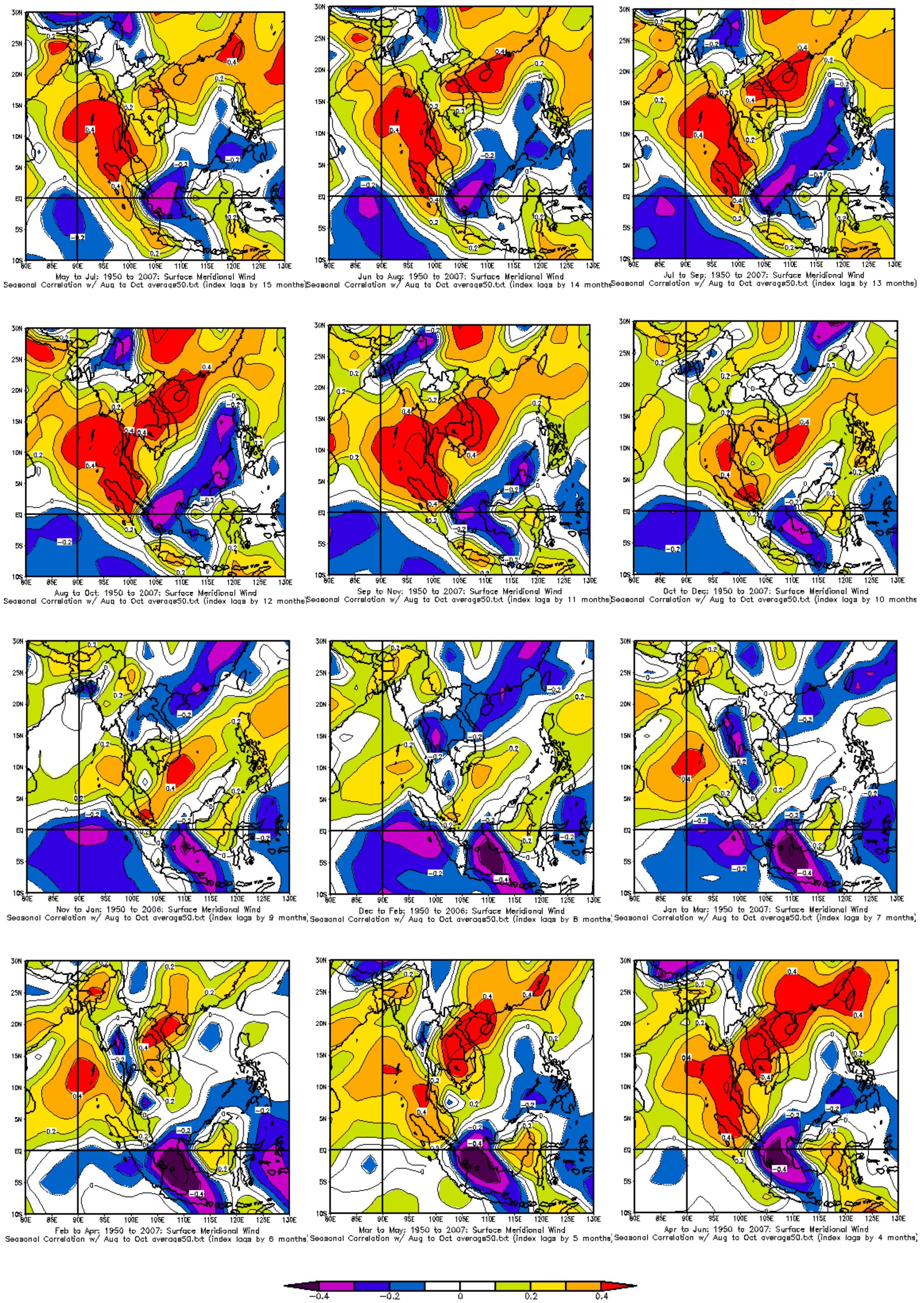


Apr to Jun: 1950 to 2007: Surface Zonal Wind
Seasonal Correlation w/ Aug to Oct average50.txt (index lags by 4 months)



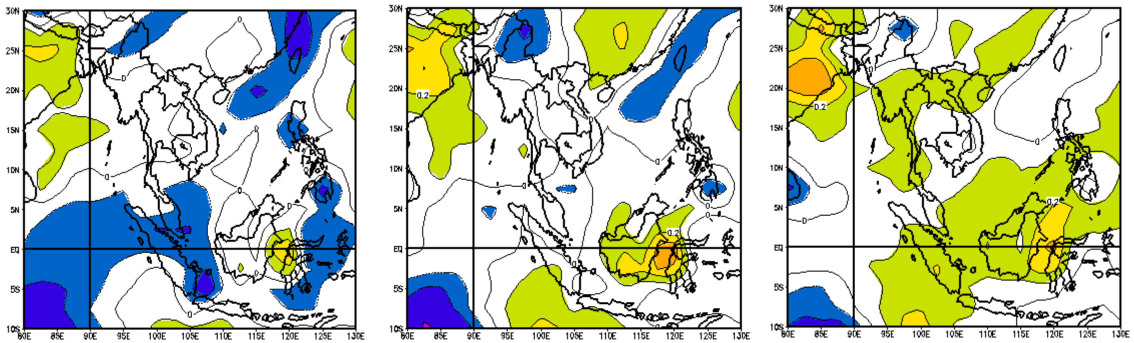
Note: The upper and lower bounds of 95% significant levels of correlation are +0.3 and -0.3, respectively.

C8: Correlation maps between ASO rainfall and SYW at lead times varying from 15 to 4 months



Note: The upper and lower bounds of 95% significant levels of correlation are +0.3 and -0.3, respectively.

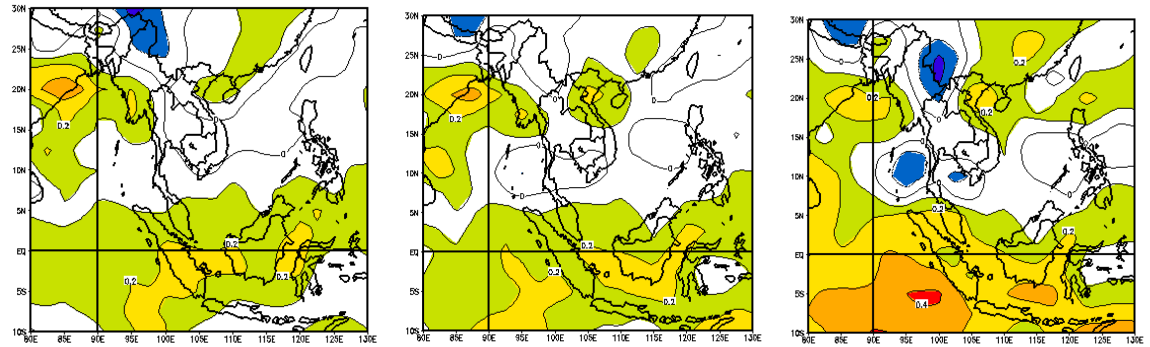
C9: Correlation maps between NDJ rainfall and SAT at lead times varying from 15 to 4 months



Aug to Oct: 1980 to 2007: Surface Air Temperature
Seasonal Correlation w/ Nov to average50.txt (index lags by 15 months)

Sep to Nov: 1980 to 2007: Surface Air Temperature
Seasonal Correlation w/ Nov to average50.txt (index lags by 14 months)

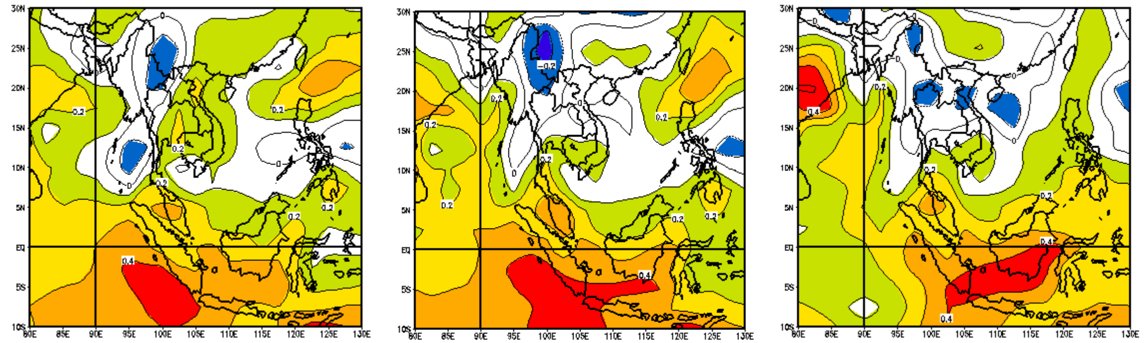
Oct to Dec: 1980 to 2007: Surface Air Temperature
Seasonal Correlation w/ Nov to average50.txt (index lags by 13 months)



Nov to Jan: 1980 to 2007: Surface Air Temperature
Seasonal Correlation w/ Nov to Jan average50.txt (index lags by 12 months)

Dec to Feb: 1980 to 2007: Surface Air Temperature
Seasonal Correlation w/ Nov to Jan average50.txt (index lags by 11 months)

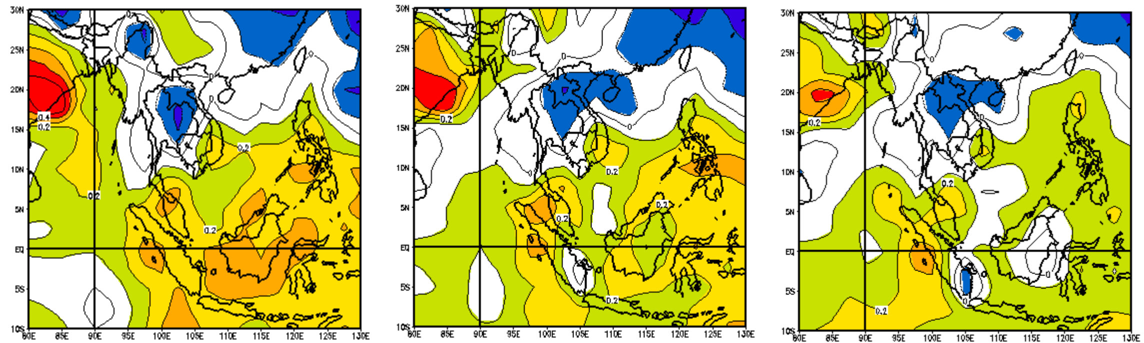
Jan to Mar: 1980 to 2007: Surface Air Temperature
Seasonal Correlation w/ Nov to Jan average50.txt (index lags by 10 months)



Feb to Apr: 1980 to 2007: Surface Air Temperature
Seasonal Correlation w/ Nov to Jan average50.txt (index lags by 9 months)

Mar to May: 1980 to 2007: Surface Air Temperature
Seasonal Correlation w/ Nov to Jan average50.txt (index lags by 8 months)

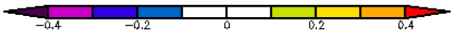
Apr to Jun: 1980 to 2007: Surface Air Temperature
Seasonal Correlation w/ Nov to Jan average50.txt (index lags by 7 months)



May to Jul: 1980 to 2007: Surface Air Temperature
Seasonal Correlation w/ Nov to Jan average50.txt (index lags by 6 months)

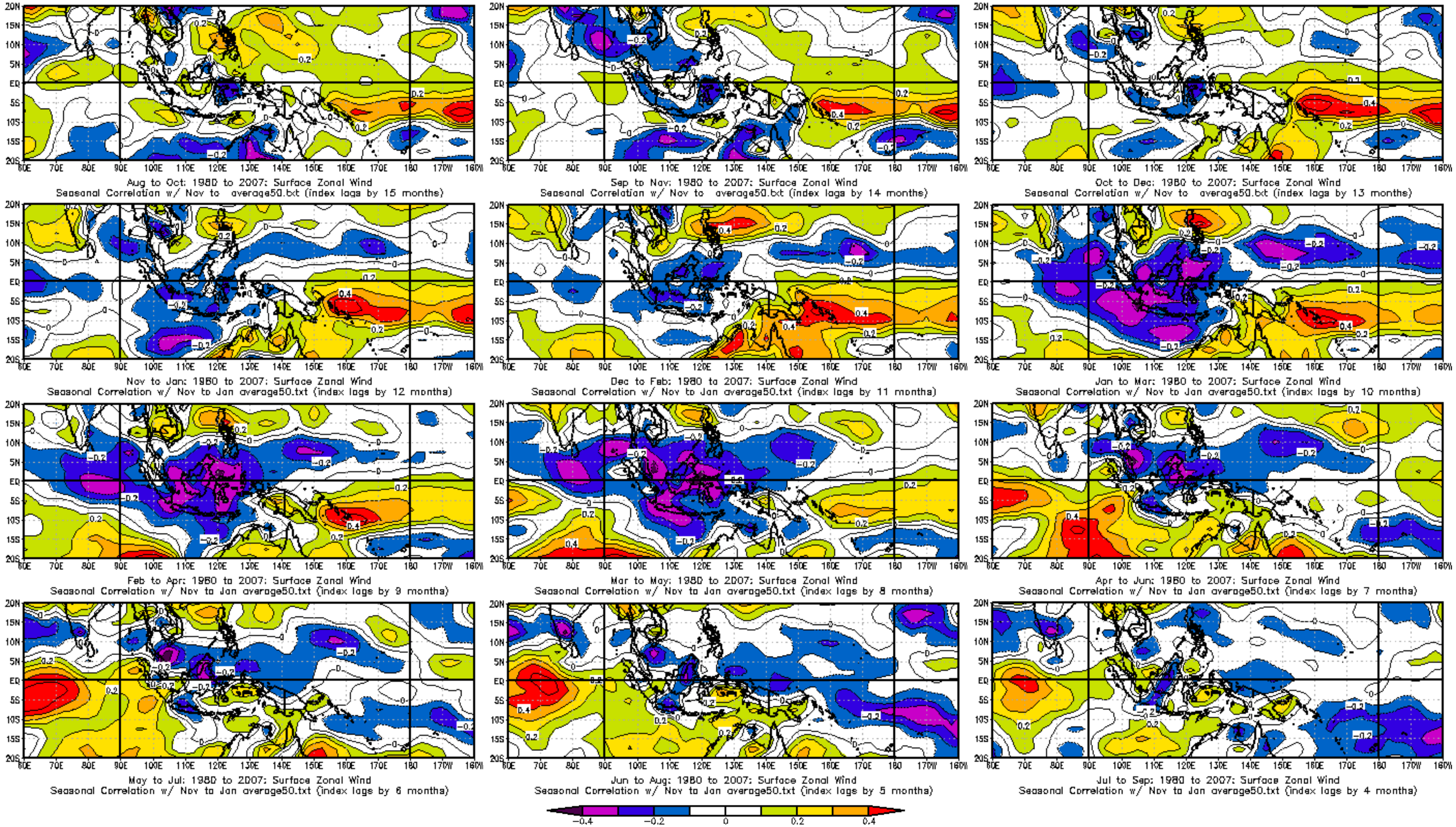
Jun to Aug: 1980 to 2007: Surface Air Temperature
Seasonal Correlation w/ Nov to Jan average50.txt (index lags by 5 months)

Jul to Sep: 1980 to 2007: Surface Air Temperature
Seasonal Correlation w/ Nov to Jan average50.txt (index lags by 4 months)



Note: The upper and lower bounds of 95% significant levels of correlation are +0.3 and -0.3, respectively.

C10: Correlation maps between NDJ rainfall and SXW at lead times varying from 15 to 4 months



Aug to Oct: 1980 to 2007: Surface Zonal Wind
Seasonal Correlation w/ Nov to average50.txt (index lags by 15 months)

Sep to Nov: 1980 to 2007: Surface Zonal Wind
Seasonal Correlation w/ Nov to average50.txt (index lags by 14 months)

Oct to Dec: 1980 to 2007: Surface Zonal Wind
Seasonal Correlation w/ Nov to average50.txt (index lags by 13 months)

Nov to Jan: 1980 to 2007: Surface Zonal Wind
Seasonal Correlation w/ Nov to Jan average50.txt (index lags by 12 months)

Dec to Feb: 1980 to 2007: Surface Zonal Wind
Seasonal Correlation w/ Nov to Jan average50.txt (index lags by 11 months)

Jan to Mar: 1980 to 2007: Surface Zonal Wind
Seasonal Correlation w/ Nov to Jan average50.txt (index lags by 10 months)

Feb to Apr: 1980 to 2007: Surface Zonal Wind
Seasonal Correlation w/ Nov to Jan average50.txt (index lags by 9 months)

Mar to May: 1980 to 2007: Surface Zonal Wind
Seasonal Correlation w/ Nov to Jan average50.txt (index lags by 8 months)

Apr to Jun: 1980 to 2007: Surface Zonal Wind
Seasonal Correlation w/ Nov to Jan average50.txt (index lags by 7 months)

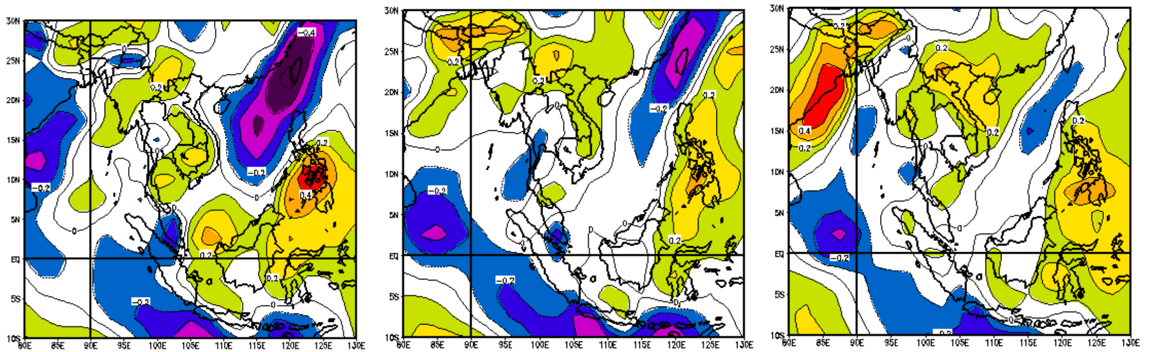
May to Jul: 1980 to 2007: Surface Zonal Wind
Seasonal Correlation w/ Nov to Jan average50.txt (index lags by 6 months)

Jun to Aug: 1980 to 2007: Surface Zonal Wind
Seasonal Correlation w/ Nov to Jan average50.txt (index lags by 5 months)

Jul to Sep: 1980 to 2007: Surface Zonal Wind
Seasonal Correlation w/ Nov to Jan average50.txt (index lags by 4 months)

Note: The upper and lower bounds of 95% significant levels of correlation are +0.3 and -0.3, respectively.

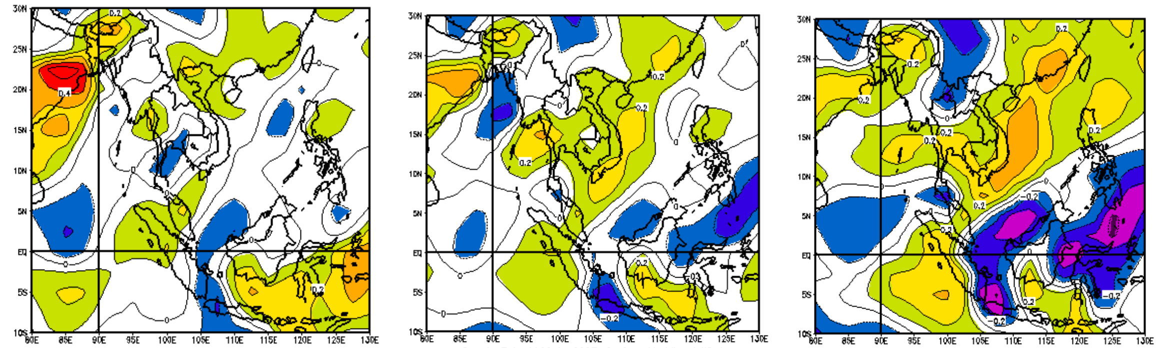
C11: Correlation maps between NDJ rainfall and SYW at lead times varying from 15 to 4 months



Aug to Oct: 1980 to 2007: Surface Meridional Wind
Seasonal Correlation w/ Nov to average50.txt (index lags by 15 months)

Sep to Nov: 1980 to 2007: Surface Meridional Wind
Seasonal Correlation w/ Nov to average50.txt (index lags by 14 months)

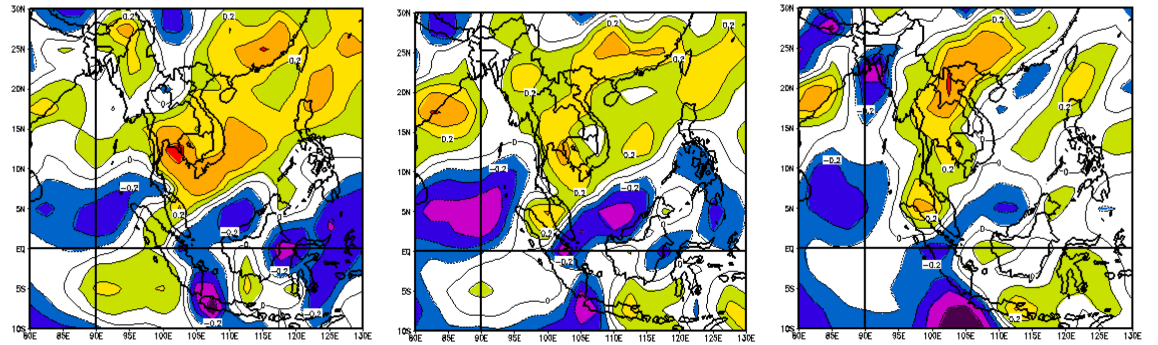
Oct to Dec: 1980 to 2007: Surface Meridional Wind
Seasonal Correlation w/ Nov to average50.txt (index lags by 13 months)



Nov to Jan: 1980 to 2007: Surface Meridional Wind
Seasonal Correlation w/ Nov to Jan average50.txt (index lags by 12 months)

Dec to Feb: 1980 to 2007: Surface Meridional Wind
Seasonal Correlation w/ Nov to Jan average50.txt (index lags by 11 months)

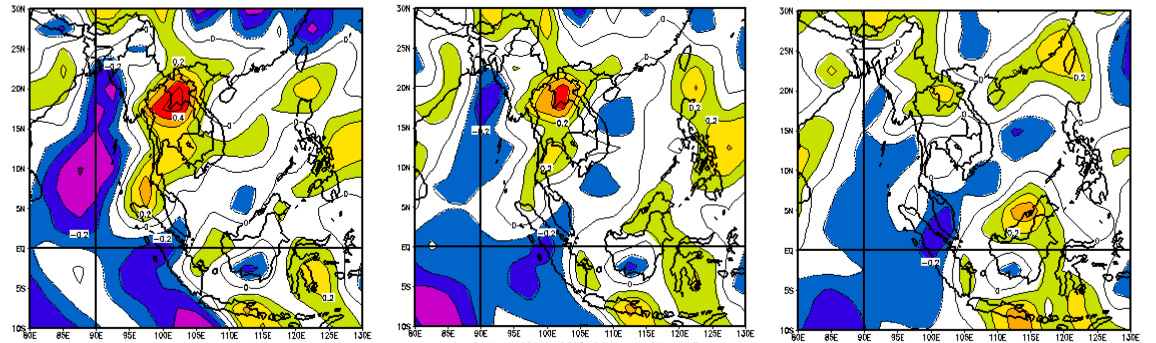
Jan to Mar: 1980 to 2007: Surface Meridional Wind
Seasonal Correlation w/ Nov to Jan average50.txt (index lags by 10 months)



Feb to Apr: 1980 to 2007: Surface Meridional Wind
Seasonal Correlation w/ Nov to Jan average50.txt (index lags by 9 months)

Mar to May: 1980 to 2007: Surface Meridional Wind
Seasonal Correlation w/ Nov to Jan average50.txt (index lags by 8 months)

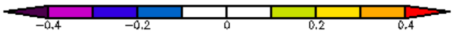
Apr to Jun: 1980 to 2007: Surface Meridional Wind
Seasonal Correlation w/ Nov to Jan average50.txt (index lags by 7 months)



May to Jul: 1980 to 2007: Surface Meridional Wind
Seasonal Correlation w/ Nov to Jan average50.txt (index lags by 6 months)

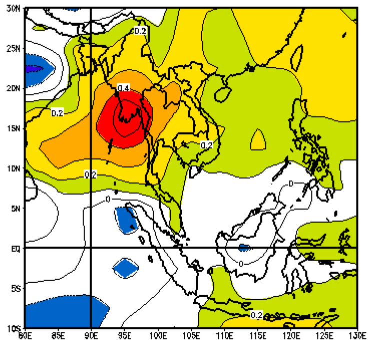
Jun to Aug: 1980 to 2007: Surface Meridional Wind
Seasonal Correlation w/ Nov to Jan average50.txt (index lags by 5 months)

Jul to Sep: 1980 to 2007: Surface Meridional Wind
Seasonal Correlation w/ Nov to Jan average50.txt (index lags by 4 months)

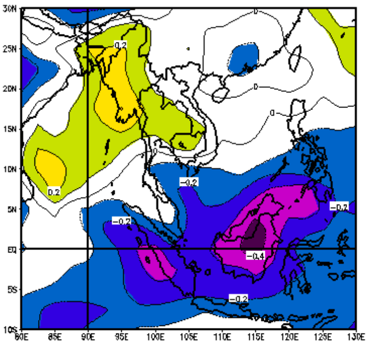


Note: The upper and lower bounds of 95% significant levels of correlation are +0.3 and -0.3, respectively.

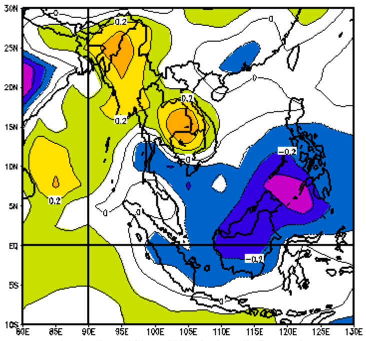
C12: Correlation maps between FMA rainfall and SAT at lead times varying from 15 to 4 months



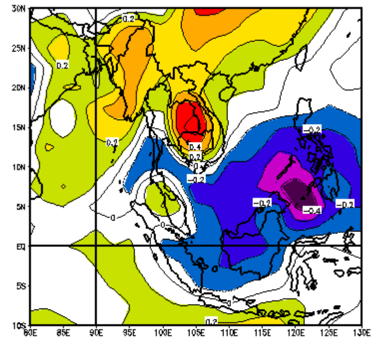
Nov to Jan: 1980 to 2007: Surface Air Temperature
Seasonal Correlation w/ to Apr average50.txt (index lags by 15 months



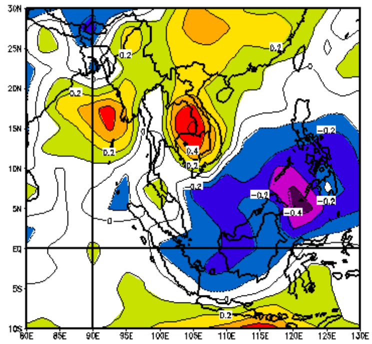
Dec to Feb: 1980 to 2007: Surface Air Temperature
Seasonal Correlation w/ to Apr average50.txt (index lags by 14 months



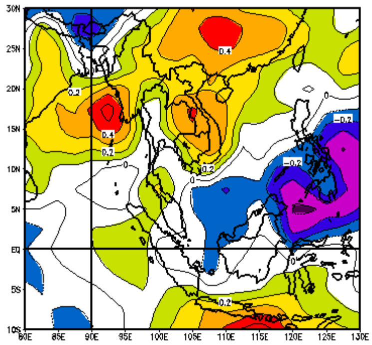
Jan to Mar: 1980 to 2007: Surface Air Temperature
Seasonal Correlation w/ Feb to Apr average50.txt (index lags by 13 months)



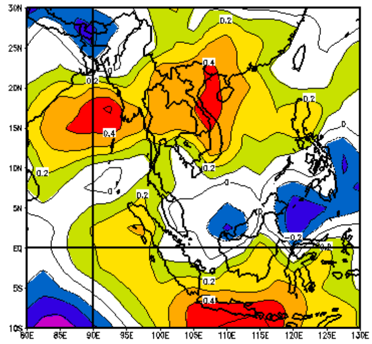
Feb to Apr: 1980 to 2007: Surface Air Temperature
Seasonal Correlation w/ Feb to Apr average50.txt (index lags by 12 months)



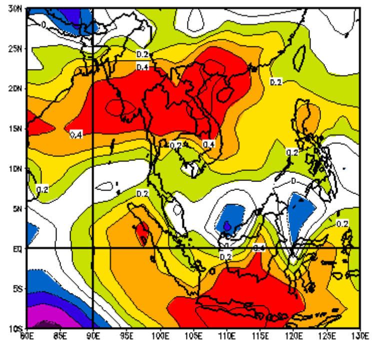
Mar to May: 1980 to 2007: Surface Air Temperature
Seasonal Correlation w/ Feb to Apr average50.txt (index lags by 11 months)



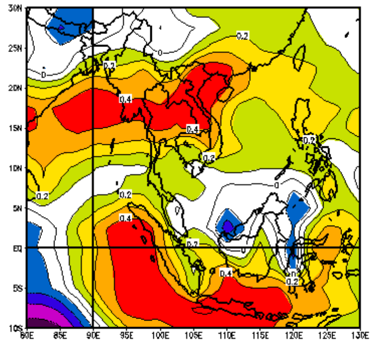
Apr to Jun: 1980 to 2007: Surface Air Temperature
Seasonal Correlation w/ Feb to Apr average50.txt (index lags by 10 months)



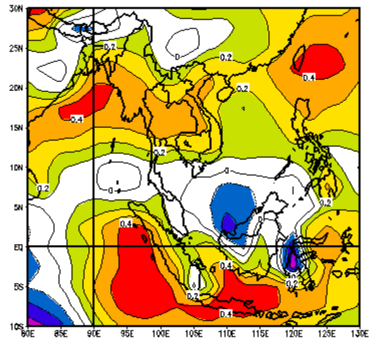
May to Jul: 1980 to 2007: Surface Air Temperature
Seasonal Correlation w/ Feb to Apr average50.txt (index lags by 9 months)



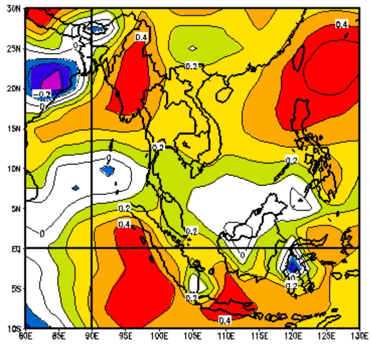
Jun to Aug: 1980 to 2007: Surface Air Temperature
Seasonal Correlation w/ Feb to Apr average50.txt (index lags by 8 months)



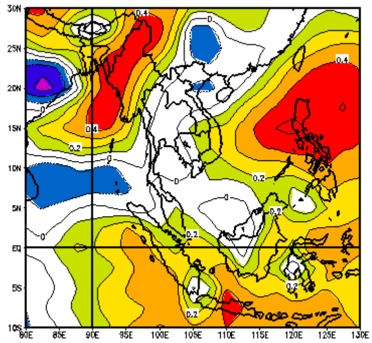
Jul to Sep: 1980 to 2007: Surface Air Temperature
Seasonal Correlation w/ Feb to Apr average50.txt (index lags by 7 months)



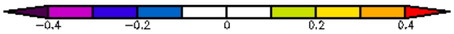
Aug to Oct: 1980 to 2007: Surface Air Temperature
Seasonal Correlation w/ Feb to Apr average50.txt (index lags by 6 months)



Sep to Nov: 1980 to 2007: Surface Air Temperature
Seasonal Correlation w/ Feb to Apr average50.txt (index lags by 5 months)

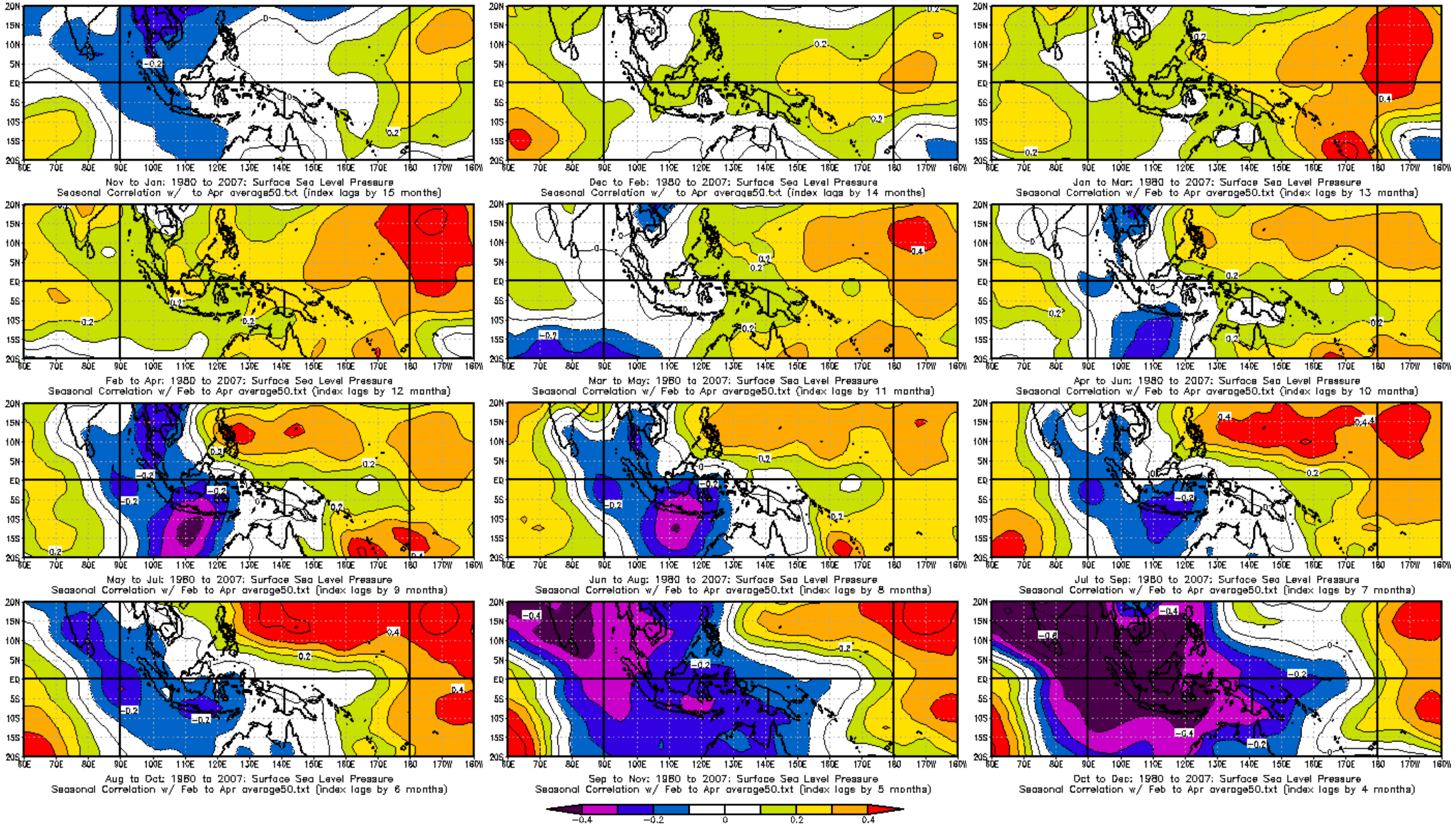


Oct to Dec: 1980 to 2007: Surface Air Temperature
Seasonal Correlation w/ Feb to Apr average50.txt (index lags by 4 months)



Note: The upper and lower bounds of 95% significant levels of correlation are +0.3 and -0.3, respectively.

C13: Correlation maps between FMA rainfall and SLP at lead times varying from 15 to 4 months



Nov to Jan: 1980 to 2007: Surface Sea Level Pressure
Seasonal Correlation w/ to Apr average50.txt (index lags by 15 months)

Dec to Feb: 1980 to 2007: Surface Sea Level Pressure
Seasonal Correlation w/ to Apr average50.txt (index lags by 14 months)

Jan to Mar: 1980 to 2007: Surface Sea Level Pressure
Seasonal Correlation w/ Feb to Apr average50.txt (index lags by 13 months)

Feb to Apr: 1980 to 2007: Surface Sea Level Pressure
Seasonal Correlation w/ Feb to Apr average50.txt (index lags by 12 months)

Mar to May: 1980 to 2007: Surface Sea Level Pressure
Seasonal Correlation w/ Feb to Apr average50.txt (index lags by 11 months)

Apr to Jun: 1980 to 2007: Surface Sea Level Pressure
Seasonal Correlation w/ Feb to Apr average50.txt (index lags by 10 months)

May to Jul: 1980 to 2007: Surface Sea Level Pressure
Seasonal Correlation w/ Feb to Apr average50.txt (index lags by 9 months)

Jun to Aug: 1980 to 2007: Surface Sea Level Pressure
Seasonal Correlation w/ Feb to Apr average50.txt (index lags by 8 months)

Jul to Sep: 1980 to 2007: Surface Sea Level Pressure
Seasonal Correlation w/ Feb to Apr average50.txt (index lags by 7 months)

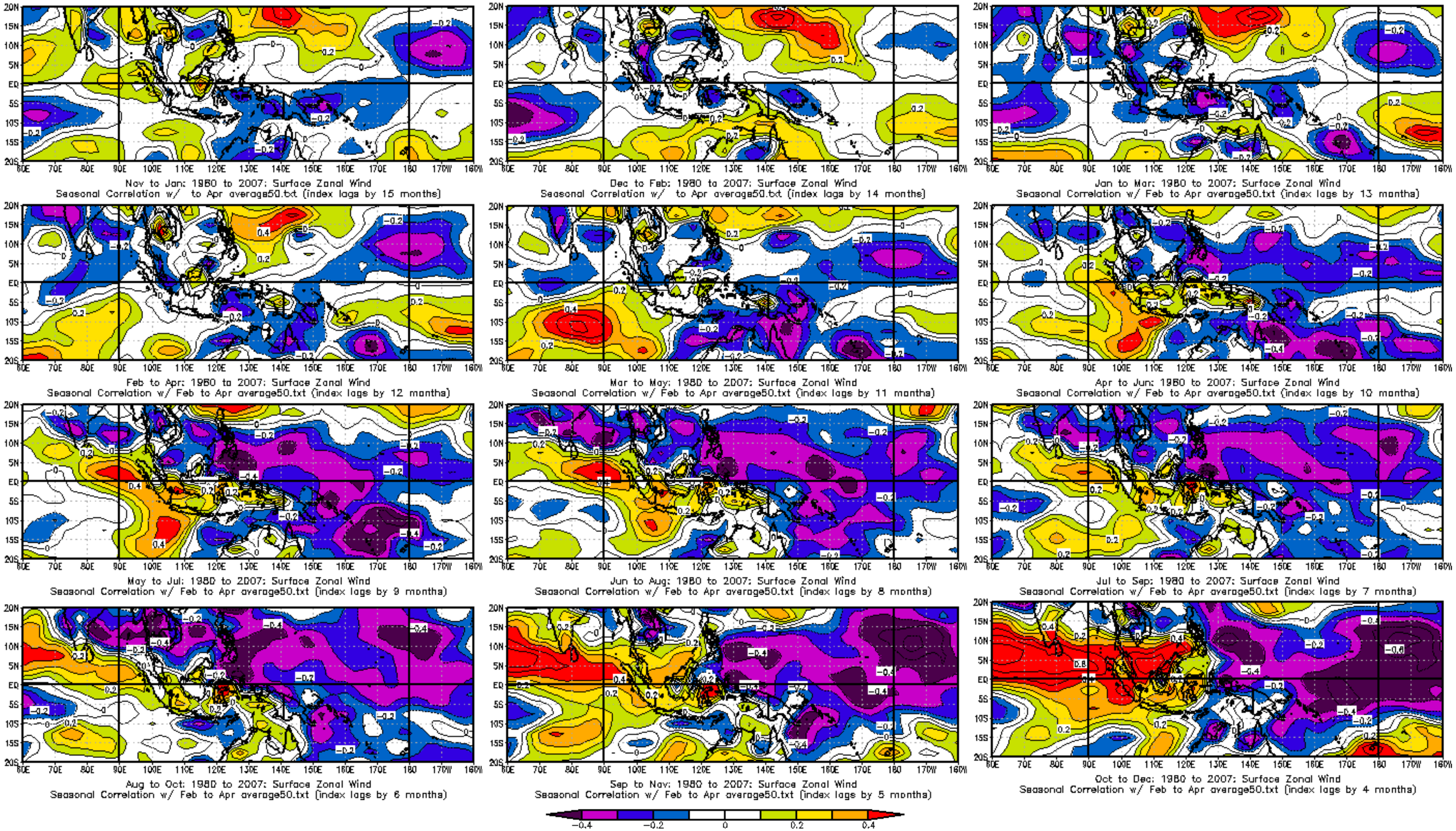
Aug to Oct: 1980 to 2007: Surface Sea Level Pressure
Seasonal Correlation w/ Feb to Apr average50.txt (index lags by 6 months)

Sep to Nov: 1980 to 2007: Surface Sea Level Pressure
Seasonal Correlation w/ Feb to Apr average50.txt (index lags by 5 months)

Oct to Dec: 1980 to 2007: Surface Sea Level Pressure
Seasonal Correlation w/ Feb to Apr average50.txt (index lags by 4 months)

Note: The upper and lower bounds of 95% significant levels of correlation are +0.3 and -0.3, respectively.

C14: Correlation maps between FMA rainfall and SXW at lead times varying from 15 to 4 months



Nov to Jan: 1980 to 2007: Surface Zonal Wind
Seasonal Correlation w/ to Apr average50.txt (index lags by 15 months)

Dec to Feb: 1980 to 2007: Surface Zonal Wind
Seasonal Correlation w/ to Apr average50.txt (index lags by 14 months)

Jan to Mar: 1980 to 2007: Surface Zonal Wind
Seasonal Correlation w/ Feb to Apr average50.txt (index lags by 13 months)

Feb to Apr: 1980 to 2007: Surface Zonal Wind
Seasonal Correlation w/ Feb to Apr average50.txt (index lags by 12 months)

Mar to May: 1980 to 2007: Surface Zonal Wind
Seasonal Correlation w/ Feb to Apr average50.txt (index lags by 11 months)

Apr to Jun: 1980 to 2007: Surface Zonal Wind
Seasonal Correlation w/ Feb to Apr average50.txt (index lags by 10 months)

May to Jul: 1980 to 2007: Surface Zonal Wind
Seasonal Correlation w/ Feb to Apr average50.txt (index lags by 9 months)

Jun to Aug: 1980 to 2007: Surface Zonal Wind
Seasonal Correlation w/ Feb to Apr average50.txt (index lags by 8 months)

Jul to Sep: 1980 to 2007: Surface Zonal Wind
Seasonal Correlation w/ Feb to Apr average50.txt (index lags by 7 months)

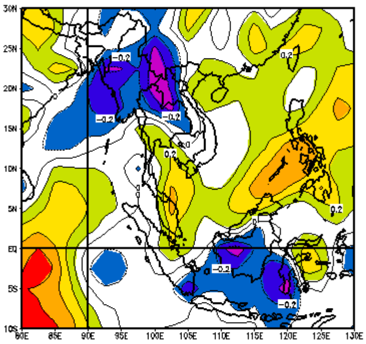
Aug to Oct: 1980 to 2007: Surface Zonal Wind
Seasonal Correlation w/ Feb to Apr average50.txt (index lags by 6 months)

Sep to Nov: 1980 to 2007: Surface Zonal Wind
Seasonal Correlation w/ Feb to Apr average50.txt (index lags by 5 months)

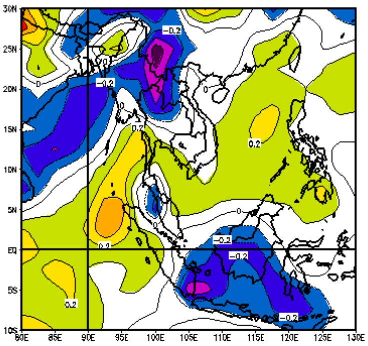
Oct to Dec: 1980 to 2007: Surface Zonal Wind
Seasonal Correlation w/ Feb to Apr average50.txt (index lags by 4 months)

Note: The upper and lower bounds of 95% significant levels of correlation are +0.3 and -0.3, respectively.

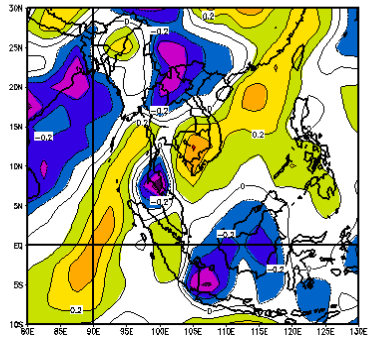
C15: Correlation maps between FMA rainfall and SYW at lead times varying from 15 to 4 months



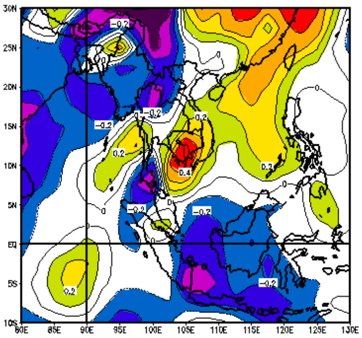
Nov to Jan; 1980 to 2007; Surface Meridional Wind
Seasonal Correlation w/ to Apr average50.txt (index lags by 15 months)



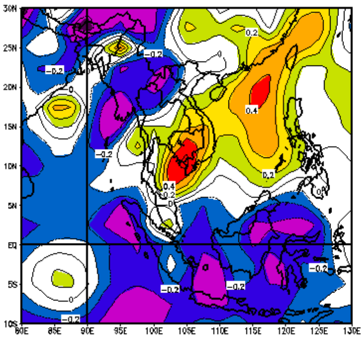
Dec to Feb; 1980 to 2007; Surface Meridional Wind
Seasonal Correlation w/ to Apr average50.txt (index lags by 14 months)



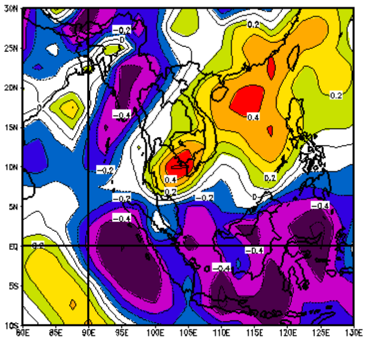
Jan to Mar; 1980 to 2007; Surface Meridional Wind
Seasonal Correlation w/ Feb to Apr average50.txt (index lags by 13 months)



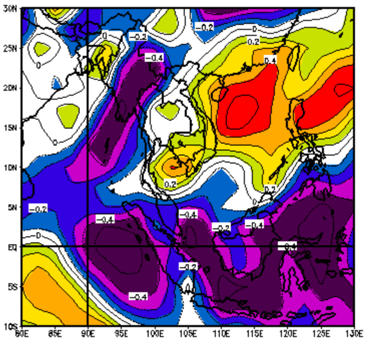
Feb to Apr; 1980 to 2007; Surface Meridional Wind
Seasonal Correlation w/ Feb to Apr average50.txt (index lags by 12 months)



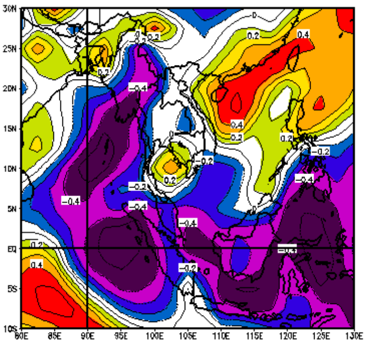
Mar to May; 1980 to 2007; Surface Meridional Wind
Seasonal Correlation w/ Feb to Apr average50.txt (index lags by 11 months)



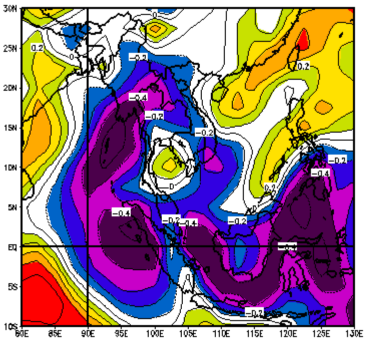
Apr to Jun; 1980 to 2007; Surface Meridional Wind
Seasonal Correlation w/ Feb to Apr average50.txt (index lags by 10 months)



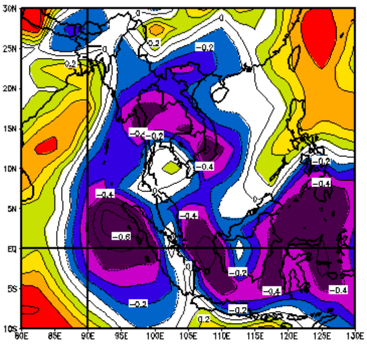
May to Jul; 1980 to 2007; Surface Meridional Wind
Seasonal Correlation w/ Feb to Apr average50.txt (index lags by 9 months)



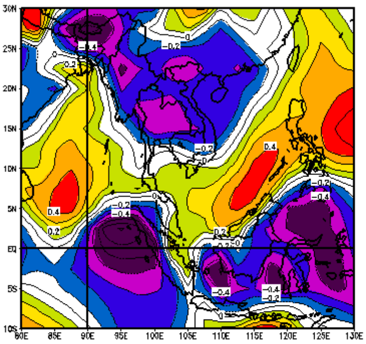
Jun to Aug; 1980 to 2007; Surface Meridional Wind
Seasonal Correlation w/ Feb to Apr average50.txt (index lags by 8 months)



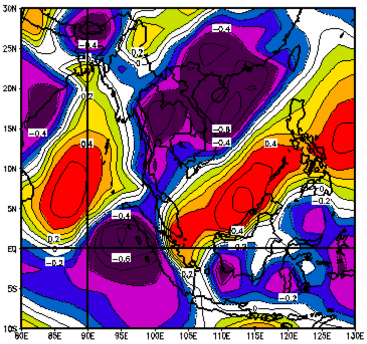
Jul to Sep; 1980 to 2007; Surface Meridional Wind
Seasonal Correlation w/ Feb to Apr average50.txt (index lags by 7 months)



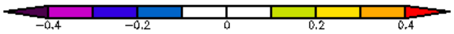
Aug to Oct; 1980 to 2007; Surface Meridional Wind
Seasonal Correlation w/ Feb to Apr average50.txt (index lags by 6 months)



Sep to Nov; 1980 to 2007; Surface Meridional Wind
Seasonal Correlation w/ Feb to Apr average50.txt (index lags by 5 months)



Oct to Dec; 1980 to 2007; Surface Meridional Wind
Seasonal Correlation w/ Feb to Apr average50.txt (index lags by 4 months)



Note: The upper and lower bounds of 95% significant levels of correlation are +0.3 and -0.3, respectively.

Appendix D: Combination Cases of the Predictors

D1: GCV of the combination cases of predictors for MJJ rainfall under A2

CASE	Variable					Season of predictors (Lead time in months)											
	SAT	SLP	SXW	SYW	RAIN	JFM (4)	DJF (5)	NDJ (6)	OND (7)	SON (8)	ASO (9)	JAS (10)	JJA (11)	MJJ (12)	AMJ (13)	MAM (14)	FMA (15)
CASE 1	X					9573.7	9595.0	9564.7	10670.8	10064.2	10004.7	10270.6	10385.1	9859.4	9601.8	9309.0	11093.9
CASE 2		X				8569.1	8731.9	8653.0	9728.7	9048.3	8875.0	9388.0	9085.5	9910.2	10091.3	9308.9	9549.3
CASE 3			X			9085.2	8874.9	8154.0	10111.4	9453.4	9646.6	10258.2	9522.3	10355.7	8689.8	10068.4	10088.8
CASE 4				X		8689.6	9792.2	9691.3	9698.6	8624.3	8239.2	8777.8	8917.3	9389.1	9407.8	10808.1	9811.4
CASE 5					X	8885.4	8943.9	9362.8	11199.3	9904.7	9793.2	9428.6	9621.4	9383.9	9264.9	9046.2	9613.8
CASE 6	X	X				10017.9	10993.3	10570.9	12837.7	11588.5	11249.7	11313.3	8603.9	13304.9	12462.9	11755.0	13194.2
CASE 7	X		X			11955.8	15850.7	8923.1	11268.5	11702.4	12142.2	12375.7	13079.3	11987.1	10011.4	8913.3	12293.4
CASE 8	X			X		10382.7	11326.4	11606.4	17485.1	12061.4	9475.6	10096.5	9302.0	9829.3	9655.2	11745.3	15873.3
CASE 9	X				X	10327.5	11294.0	10802.7	13194.8	12988.5	11448.9	11800.7	14032.7	11781.5	10451.3	9690.5	13085.8
CASE 10		X	X			9146.3	9572.5	7692.5	11119.3	10485.4	12686.0	10121.4	11401.1	12989.2	9959.4	11587.8	10646.1
CASE 11		X		X		9317.8	11186.6	10548.9	11381.3	9167.5	8290.6	10463.3	10493.7	12180.4	13301.9	13769.0	11173.4
CASE 12		X			X	9808.1	8450.2	9820.2	12860.4	9538.1	10875.3	8744.7	11552.8	13462.9	12000.3	11433.0	11617.1
CASE 13			X	X		9661.2	10793.6	10191.9	13220.0	11226.9	10588.9	11055.5	10427.1	10376.3	8924.2	10705.3	11542.9
CASE 14			X		X	9730.4	9562.4	8939.5	11582.6	11640.5	16270.1	9896.5	11509.8	11645.4	10460.6	11326.2	10982.1
CASE 15				X	X	9069.9	10312.3	13814.3	15009.2	10181.4	10064.8	10187.5	10034.8	10241.7	10851.6	11118.2	11088.7
CASE 16	X	X	X			13886.0	17517.0	8525.5	13027.2	14658.3	17023.2	16136.7	12067.5	22671.6	13125.5	10374.4	14321.4
CASE 17	X	X		X		15039.4	15129.7	13389.7	21888.8	12420.5	11178.6	13067.1	11452.2	12818.4	13922.8	17889.9	18944.3
CASE 18	X	X			X	11538.7	14893.1	11161.0	15635.6	16429.7	13080.3	10988.8	14125.0	17204.9	15878.6	13358.5	19481.4
CASE 19	X		X	X		15241.4	15953.2	13269.8	18678.7	17003.6	15975.9	13165.9	11735.2	13669.5	11824.3	12763.6	16765.9
CASE 20	X		X		X	13009.5	22125.6	9890.5	14942.0	17043.9	25623.6	16904.4	21015.7	15194.7	12822.7	11167.4	16176.7
CASE 21	X			X	X	11199.0	13294.9	12517.1	22041.6	13329.3	13223.8	12190.9	14724.2	12087.7	12274.8	15278.1	17050.0
CASE 22		X	X	X		10831.3	12935.4	9210.5	13636.9	13902.9	23185.6	14226.2	12830.5	16769.6	10847.9	15765.4	12615.9
CASE 23		X	X		X	12969.2	10108.3	8886.6	12166.3	11822.5	13694.7	10329.4	14242.2	17919.2	14903.5	15276.8	14702.5
CASE 24		X		X	X	11695.3	9527.5	10801.2	14055.1	10775.5	13768.2	9510.9	13769.2	15301.7	16988.7	18147.9	13123.9
CASE 25			X	X	X	10945.0	10205.8	13592.6	14035.4	12872.3	10068.2	15654.3	12909.5	12824.0	13937.6	16484.0	13172.1
CASE 26	X	X	X	X		23239.5	22370.2	12327.2	15010.4	17954.2	24662.0	16813.9	14362.3	29265.4	14261.1	17930.3	20995.2
CASE 27	X	X	X		X	18527.6	23859.3	9107.0	20996.2	23670.5	24500.3	11855.7	22300.5	22499.1	20608.3	14427.5	17938.1
CASE 28	X	X		X	X	20705.9	12934.9	12301.8	21414.9	17808.9	18253.4	12217.6	19942.9	19943.4	19298.3	24640.1	23451.3
CASE 29	X		X	X	X	16425.4	18075.8	25202.1	21304.6	17404.3	17724.4	13676.0	14810.0	14380.4	13605.8	20105.4	23922.7
CASE 30		X	X	X	X	14117.2	10866.0	10099.3	13490.7	13782.9	15305.1	18198.5	17426.4	21215.3	20014.6	25617.7	16117.9
CASE 31	X	X	X	X	X	22428.5	19355.3	13503.6	20435.5	26312.3	32715.7	23268.4	19672.8	31467.6	23356.7	27810.8	29105.3
Minimum GCV						8569.1	8450.2	7692.5	9698.6	8624.3	8239.2	8744.7	8603.9	9383.9	8689.8	8913.3	9549.3
Selected Case						CASE 2	CASE 12	CASE 10	CASE 4	CASE 4	CASE 4	CASE 12	CASE 6	CASE 5	CASE 3	CASE 7	CASE 2

D2: GCV of the combination cases of predictors for ASO rainfall under A2

CASE	Variable					Season of predictors (Lead time in months)											
	SAT	SLP	SXW	SYW	RAIN	AMJ (4)	MAM (5)	FMA (6)	JFM (7)	DJF (8)	NDJ (9)	OND (10)	SON (11)	ASO (12)	JAS (13)	JJA (14)	MJJ (15)
CASE 1	X					7930.4	6729.2	6629.3	6765.6	6274.8	7351.1	9789.1	9399.6	8891.3	7993.5	7939.0	8406.5
CASE 2		X				7263.5	7768.2	7623.4	7011.4	5558.5	5860.7	6683.3	7830.6	7813.9	7664.3	6991.3	7578.9
CASE 3			X			6449.8	7225.3	6861.4	6360.4	7860.8	7844.6	8266.3	8161.9	7590.0	7348.9	7652.8	7318.2
CASE 4				X		7483.4	7279.7	6802.9	7157.3	8141.9	7550.4	8419.4	7778.3	8365.0	7916.5	8217.3	7512.8
CASE 5					X	7987.4	7595.8	7646.4	7547.1	7923.2	7850.8	8824.5	8742.7	7868.6	8027.1	7733.6	7963.3
CASE 6	X	X				8997.2	9946.5	10195.5	7807.7	7615.1	7471.7	8170.5	9800.9	10148.2	9849.1	11364.8	9748.8
CASE 7	X		X			6940.0	7479.0	5909.9	8705.2	8985.2	8390.1	14421.4	13058.7	10598.5	8138.5	9217.4	9944.4
CASE 8	X			X		8637.4	7985.5	7298.2	6502.4	7789.8	8961.1	12270.4	8102.4	9364.3	9497.5	18455.1	9633.5
CASE 9	X				X	9600.9	6801.7	7421.8	7185.7	8059.1	9269.7	11095.6	7141.0	8844.4	9969.8	9688.6	10381.0
CASE 10		X	X			7450.7	9716.7	7477.4	7744.1	8345.7	9829.8	7912.9	9071.4	9506.3	8946.8	7657.0	8799.4
CASE 11		X		X		8671.8	10836.2	7357.0	8918.1	11539.6	6540.2	8904.9	9196.9	9018.7	8992.0	8824.0	9130.4
CASE 12		X			X	7622.6	12045.8	9307.6	8186.8	6350.9	8802.5	10244.6	10352.7	8791.5	8889.4	7584.2	8586.8
CASE 13			X	X		5404.6	7414.4	7033.8	6945.8	11643.9	9861.5	11214.7	10402.4	8896.4	7803.7	8924.8	8504.8
CASE 14			X		X	6226.8	7497.9	7567.8	7621.5	9544.8	10868.9	10838.9	9076.9	8945.9	9867.3	8649.6	9064.2
CASE 15				X	X	8325.0	9625.3	7826.0	8871.4	9332.0	9355.9	10777.3	9908.1	9232.4	10129.1	8103.0	8181.1
CASE 16	X	X	X			9924.6	10806.0	9401.2	9202.6	13272.8	11981.9	12669.3	11519.3	7820.8	13283.8	13194.1	12381.5
CASE 17	X	X		X		12091.8	9776.3	11271.5	8988.3	14365.3	9704.9	11862.3	11562.8	11091.2	13020.8	21136.0	11562.3
CASE 18	X	X			X	9631.4	11037.4	12823.6	9726.7	10033.7	11841.5	10107.4	8716.0	12937.1	11823.0	13837.3	12922.3
CASE 19	X		X	X		6009.3	9250.1	6986.3	6859.7	11009.0	13404.3	13482.7	11070.1	9976.5	9301.0	14962.5	11865.6
CASE 20	X		X		X	7575.3	7267.3	7018.6	10610.2	12247.7	16960.0	14299.3	7244.3	9916.7	10620.6	11006.0	13188.9
CASE 21	X			X	X	9782.4	9545.0	9548.8	8314.5	10856.0	12378.2	14037.1	7897.2	13214.5	13380.8	14827.3	10784.2
CASE 22		X	X	X		8974.1	11008.7	9266.6	9925.6	12046.8	11946.7	9187.6	14060.2	11956.2	14762.8	12017.7	13113.6
CASE 23		X	X		X	9037.7	19891.4	10831.5	10353.6	10453.8	10397.3	12388.3	11159.6	10461.0	10919.5	9101.3	11692.5
CASE 24		X		X	X	9127.3	18160.7	12215.0	11569.4	8728.3	8903.9	11987.2	12972.6	11548.4	13928.2	8772.1	11403.8
CASE 25			X	X	X	5478.8	9875.1	8110.8	8955.9	11739.9	13680.8	14009.2	11735.6	11391.4	10104.1	10140.2	11617.4
CASE 26	X	X	X	X		14622.7	12423.0	13562.3	12552.8	16173.9	17266.0	12204.5	17113.2	10499.4	31329.8	20507.5	21757.2
CASE 27	X	X	X		X	13509.2	15307.7	15909.7	12809.3	19531.5	18142.1	9658.1	9237.5	8364.2	17705.8	14504.5	18230.8
CASE 28	X	X		X	X	13100.3	18938.7	19232.0	16986.0	16529.9	14112.4	13598.6	12098.7	13584.6	20764.4	20684.4	22072.0
CASE 29	X		X	X	X	7354.2	10063.7	8509.9	11832.1	15112.3	26037.6	15219.3	9899.2	15181.9	24067.4	21869.2	25894.8
CASE 30		X	X	X	X	12397.8	17286.2	14982.5	13859.8	11971.8	14269.0	17268.0	16007.7	13450.2	17823.4	12561.3	25346.3
CASE 31	X	X	X	X	X	21200.4	16384.0	23401.8	20824.8	34220.4	33152.6	17459.2	15942.0	12350.6	52413.5	30538.0	58057.5
Minimum GCV						5404.6	6729.2	5909.9	6360.4	5558.5	5860.7	6683.3	7141.0	7590.0	7348.9	6991.3	7318.2
Selected Case						CASE 13	CASE 1	CASE 7	CASE 3	CASE 2	CASE 2	CASE 2	CASE 9	CASE 3	CASE 3	CASE 2	CASE 3

D3: GCV of the combination cases of predictors for NDJ rainfall under A2

CASE	Variable				Season of predictors (Lead time in months)											
	SAT	SXW	SYW	RAIN	JAS (4)	JJA (5)	MJJ (6)	AMJ (7)	MAM (8)	FMA (9)	JFM (10)	DJF (11)	NDJ (12)	OND (13)	SON (14)	ASO (15)
CASE 1	X				3114.5	3723.8	2698.4	2671.9	2872.3	2834.0	2638.9	2988.7	2967.1	2516.3	2632.5	2958.1
CASE 2		X			2811.6	2951.6	2883.7	2239.0	2490.8	3125.4	3196.1	2793.4	2349.7	2771.0	4119.2	1937.2
CASE 3			X		2642.7	3650.2	2926.3	2828.2	4234.6	2919.5	3027.3	3246.1	3778.0	5685.0	3997.0	3534.4
CASE 4				X	4086.6	2704.0	3218.2	2849.6	3394.3	2537.5	2880.9	2879.0	2954.7	6138.7	2625.1	2437.5
CASE 5	X	X			4163.4	4253.6	3389.1	4876.7	3206.3	4896.2	3997.4	3533.0	2696.4	2710.4	3247.8	9276.4
CASE 6	X		X		4076.2	4513.5	2817.2	3923.8	6359.4	4635.0	4198.5	4977.6	5031.6	7971.3	3131.3	4334.7
CASE 7	X			X	4667.8	3952.6	4218.3	3899.2	4235.7	2729.8	3430.9	4187.0	3727.2	6119.9	3305.8	4329.6
CASE 8		X	X		3967.5	4614.8	4569.3	2618.1	5695.3	3397.1	2649.3	5503.5	6673.2	6923.5	5766.1	2582.3
CASE 9		X		X	5572.3	3713.0	4163.5	4950.2	3849.1	3404.1	3770.1	3903.6	12681.7	6227.7	6178.7	6555.7
CASE 10			X	X	8102.6	3350.7	4870.4	3802.4	5159.9	3110.2	4439.7	3833.5	5302.9	10444.7	3718.5	3648.2
CASE 11	X	X	X		6737.4	7826.8	17419.2	4333.0	8661.8	4562.1	6310.7	8961.0	11156.4	10930.3	4792.2	11428.8
CASE 12	X	X		X	7977.3	7422.9	7362.1	5485.5	4799.0	5446.5	4656.5	8146.6	4941.6	5086.2	3107.7	6802.3
CASE 13	X		X	X	9370.2	6817.6	11374.0	8547.1	9195.9	4144.1	6678.5	5741.8	8757.3	8412.7	4783.6	5756.5
CASE 14		X	X	X	16416.4	6690.0	8244.0	12384.1	7332.7	4898.1	4217.2	6357.9	10180.0	14136.0	5011.3	7641.0
CASE 15	X	X	X	X	33322.1	15103.0	33631.5	12010.0	16551.9	25460.3	8201.3	19729.6	33368.3	17780.4	6719.9	21246.1
Minimum GCV					2642.7	2704.0	2698.4	2239.0	2490.8	2537.5	2638.9	2793.4	2349.7	2516.3	2625.1	1937.2
Selected Case					CASE 3	CASE 4	CASE 1	CASE 2	CASE 2	CASE 4	CASE 1	CASE 2	CASE 2	CASE 1	CASE 4	CASE 2

D4: GCV of the combination cases of predictors for FMA rainfall under A2

CASE	Variable					Season of predictors (Lead time in months)											
	SAT	SLP	SXW	SYW	RAIN	OND (4)	SON (5)	ASO (6)	JAS (7)	JJA (8)	MJJ (9)	AMJ (10)	MAM (11)	FMA (12)	JFM (13)	DJF (14)	NDJ (15)
CASE 1	X					2274.5	2623.6	4407.1	2287.9	2959.0	4062.9	3029.0	2608.0	2735.1	2642.0	2135.6	1847.9
CASE 2		X				2645.9	2436.8	2831.0	2592.6	2629.5	2421.5	2756.7	1827.9	2699.2	3120.1	2703.5	2706.1
CASE 3			X			2621.3	2750.9	2802.1	2687.8	2543.5	2678.9	1999.9	2308.0	3298.5	4756.2	2765.7	3003.0
CASE 4				X		2410.3	2290.2	2883.0	2783.6	2435.3	2686.8	2735.0	2543.4	2362.3	2550.8	2690.2	3212.0
CASE 5					X	2811.1	2473.4	2892.0	2825.2	2141.1	2017.6	1920.8	2676.2	3190.2	2455.7	2630.6	10860.5
CASE 6	X	X				3562.0	6018.2	4746.5	3910.2	5487.4	5788.3	3805.0	3433.6	5413.0	5211.0	2193.8	1569.0
CASE 7	X		X			3184.0	5319.0	3224.7	4378.6	4127.4	4482.3	4322.8	3395.6	3471.2	9203.7	2984.9	2801.0
CASE 8	X			X		3930.6	3573.1	8850.0	3351.5	4953.3	6534.9	4723.3	4439.6	3318.4	3169.7	2442.6	2972.6
CASE 9	X				X	3326.6	3456.9	6174.6	3702.6	4125.2	3564.9	2871.0	3040.2	3037.9	3660.2	3064.6	8566.8
CASE 10		X	X			3716.8	3304.8	4581.3	3293.2	2974.5	3280.0	3049.6	5179.4	5383.9	5171.1	3597.2	2685.9
CASE 11		X		X		5438.2	4347.3	3944.3	3166.7	3011.0	3335.9	3903.9	3413.9	3028.4	3567.7	3940.4	3816.2
CASE 12		X			X	3378.1	2924.7	5681.3	2910.9	2927.8	2386.2	2314.1	2968.3	5555.1	4004.9	3407.3	16065.8
CASE 13			X	X		3864.1	6043.3	3364.2	3448.5	2923.4	3450.2	2643.8	2971.7	4445.1	6075.3	4251.1	3967.4
CASE 14			X		X	4579.8	3072.3	4728.7	3585.3	2283.9	2727.2	2475.7	3088.1	5118.4	4494.8	3073.2	14712.1
CASE 15				X	X	3056.8	5277.0	4056.9	3280.4	3196.8	2575.4	3354.0	11236.6	3160.1	2692.2	3505.2	12756.1
CASE 16	X	X	X			5723.3	17330.8	4935.4	6527.0	8611.5	6684.5	4000.5	7871.0	5859.6	18257.2	5773.2	2292.6
CASE 17	X	X		X		5100.5	8869.7	30991.4	3594.6	10774.6	10549.2	5559.6	7171.9	6806.7	7379.0	2874.1	4641.8
CASE 18	X	X			X	7820.2	7173.3	6968.3	3879.2	5759.7	4302.2	4375.0	4975.6	5047.1	7967.8	4186.5	10670.0
CASE 19	X		X	X		5623.9	4031.5	7880.0	6098.2	5960.1	8006.6	6257.7	7737.8	7094.8	15939.7	4517.9	5368.9
CASE 20	X		X		X	5333.1	5945.2	7881.1	5101.5	5301.4	4155.8	4155.1	4603.3	4988.9	8943.4	3899.9	14040.0
CASE 21	X			X	X	3752.6	6421.0	10199.2	5366.6	6161.1	5966.0	5560.7	9942.5	4802.1	8372.2	4938.2	5878.1
CASE 22		X	X	X		8047.7	11152.3	6282.5	4365.9	5838.6	5190.2	3704.3	9014.0	8586.8	9848.1	12929.5	4212.0
CASE 23		X	X		X	6534.5	4614.5	9105.8	5175.8	3812.1	3680.5	3656.8	5613.9	11920.3	7462.6	4373.0	7301.2
CASE 24		X		X	X	5341.2	22894.8	4600.1	3525.5	4076.4	3327.7	7426.8	9520.3	4080.1	5640.3	5479.7	15829.5
CASE 25			X	X	X	7488.5	14780.0	6096.1	4187.1	4815.0	3832.2	7710.8	18995.7	7252.6	6653.3	5088.6	16579.2
CASE 26	X	X	X	X		18980.1	27510.6	67689.7	13406.6	52440.1	28236.6	11994.1	12360.2	19896.2	25680.2	8129.8	4480.6
CASE 27	X	X	X		X	19605.5	25344.4	16318.8	8564.6	12917.7	5633.3	11461.9	13744.5	16496.3	21502.1	6844.8	11229.0
CASE 28	X	X		X	X	14344.7	23269.1	19204.3	5666.9	14008.4	9393.8	11144.4	15084.5	8993.2	13489.1	9808.8	19352.1
CASE 29	X		X	X	X	11805.9	15141.9	20782.5	8977.3	17205.6	13131.4	17215.1	35125.6	5735.7	19000.4	5859.0	14865.8
CASE 30		X	X	X	X	18854.0	17268.0	11022.6	10185.2	8442.7	8041.4	7790.6	14769.3	13920.0	12636.5	8732.0	24878.4
CASE 31	X	X	X	X	X	60464.4	68964.3	131736.8	18214.6	39704.8	38007.9	18159.2	23913.3	26362.0	64537.2	28863.1	144228.4
Minimum GCV						2274.5	2290.2	2802.1	2287.9	2141.1	2017.6	1920.8	1827.9	2362.3	2455.7	2135.6	1569.0
Selected Case						CASE 1	CASE 4	CASE 3	CASE 1	CASE 5	CASE 5	CASE 5	CASE 2	CASE 4	CASE 5	CASE 1	CASE 6

D5: GCV of the combination cases of predictors for MJJ rainfall under B2

CASE	Variable					Season of predictors (Lead time in months)											
	SAT	SLP	SXW	SYW	RAIN	JFM (4)	DJF (5)	NDJ (6)	OND (7)	SON (8)	ASO (9)	JAS (10)	JJA (11)	MJJ (12)	AMJ (13)	MAM (14)	FMA (15)
CASE 1	X					9280.8	9618.9	9734.1	11091.4	10457.4	9685.6	9308.9	10868.5	9536.9	8998.3	8815.4	9499.3
CASE 2		X				9250.5	9987.9	7544.0	9128.9	10155.1	9630.2	11019.8	9259.9	9581.4	10448.1	10469.6	10052.4
CASE 3			X			9835.2	7872.0	9152.7	8915.3	9700.3	11670.6	8488.1	8753.9	11136.9	11111.3	9779.4	9583.1
CASE 4				X		8286.0	9333.5	9319.7	10082.6	9742.3	9639.3	9295.2	9550.5	9649.4	9905.8	9165.5	10312.9
CASE 5					X	9611.1	8685.9	8386.1	10266.0	10342.7	9721.8	9542.1	8240.1	9667.9	9473.2	9933.9	10215.0
CASE 6	X	X				12783.7	11868.1	9446.2	12028.7	13877.7	11941.5	13363.0	13133.7	10932.3	12825.1	11751.1	11910.0
CASE 7	X		X			11696.5	9428.7	10311.7	11559.6	11330.9	13323.1	8974.8	15813.2	13110.1	12157.4	10458.2	11697.1
CASE 8	X			X		9891.9	10836.2	11572.4	14997.4	12738.5	11539.4	11680.8	12364.8	10623.0	8982.9	9565.6	15304.0
CASE 9	X				X	10830.8	10350.7	14571.3	14381.7	12368.5	11332.6	12530.2	11617.3	11761.7	11048.7	9895.3	11998.2
CASE 10		X	X			11997.2	11239.4	8426.9	12505.0	13120.6	13992.9	12204.7	16113.5	15061.9	14792.8	12616.2	11155.4
CASE 11		X		X		10372.0	11662.0	10198.5	10299.9	12076.1	12720.8	13417.6	10722.8	11976.1	12504.3	11305.6	14508.7
CASE 12		X			X	15351.6	10840.6	7724.6	11122.8	14058.2	15909.1	12311.0	10353.7	10989.3	11136.0	12148.7	14208.3
CASE 13			X	X		8629.1	9633.2	10009.1	12428.5	11457.5	12284.9	8416.5	9682.4	14842.9	14018.7	11889.7	12573.7
CASE 14			X		X	11514.9	8223.9	8772.4	11722.4	11815.8	12957.9	25399.5	12286.9	13014.1	11502.7	11214.5	16630.9
CASE 15				X	X	11115.7	9767.0	24890.9	12691.9	11176.8	10895.3	11886.1	8602.3	10224.9	11590.7	10531.4	26656.1
CASE 16	X	X	X			17349.2	11311.4	10958.7	13774.9	23652.8	16314.6	14258.6	22970.8	18121.1	19474.0	14707.1	14615.6
CASE 17	X	X		X		13079.2	15308.1	12545.6	16720.6	18643.0	18106.8	17084.9	15063.6	13330.5	12887.5	11370.4	18102.0
CASE 18	X	X			X	23953.2	13630.7	27678.8	14663.9	17935.6	19978.4	14630.6	13182.4	13838.9	15207.9	12479.5	17555.2
CASE 19	X		X	X		11133.2	12358.2	13987.9	14588.8	15274.8	14470.8	18561.8	14621.5	19659.3	11361.3	11565.2	18213.5
CASE 20	X		X		X	14429.8	9675.0	12449.3	17796.0	14094.6	15596.2	42959.6	23687.5	17751.4	14147.0	13176.2	21546.0
CASE 21	X			X	X	17358.2	12370.1	32261.2	19806.3	16052.9	14146.9	19544.6	12049.3	10686.2	10549.6	11125.6	29817.4
CASE 22		X	X	X		11014.0	16106.0	11665.3	15425.4	18017.6	16692.6	13468.8	15785.5	21387.6	18942.0	15946.6	15455.2
CASE 23		X	X		X	21193.0	10806.9	16641.1	13186.7	18899.3	23104.4	32682.0	22064.9	19740.7	16505.4	16005.5	27167.7
CASE 24		X		X	X	15181.7	13794.2	21032.9	11326.3	16754.7	27043.8	18265.5	10932.2	11989.8	15011.4	11904.1	38481.5
CASE 25			X	X	X	14124.5	11673.1	20844.8	13650.7	12598.2	16481.4	40342.0	13570.1	13718.5	15924.4	13245.0	51660.1
CASE 26	X	X	X	X		18978.0	18440.7	21362.6	19158.7	28208.1	20635.6	29519.1	22165.4	26243.7	22483.7	16126.1	26637.9
CASE 27	X	X	X		X	28946.4	13180.3	42100.4	18782.5	24019.8	30733.7	34686.5	29759.1	29572.1	22659.5	18959.3	39377.2
CASE 28	X	X		X	X	28240.1	17057.5	60572.7	19829.4	26857.5	37322.9	23969.3	12297.9	16763.1	16632.6	13917.8	21086.5
CASE 29	X		X	X	X	23556.6	20400.5	46738.0	24108.0	18606.2	27619.4	75304.3	25870.6	15749.6	14088.0	17328.2	58995.7
CASE 30		X	X	X	X	24556.3	18195.4	29669.9	19921.5	22504.3	36172.8	52733.0	20165.8	26073.7	23235.4	16692.4	82917.4
CASE 31	X	X	X	X	X	60277.2	58534.9	61959.0	29019.1	38552.1	68569.9	62640.2	18059.3	36902.7	25460.0	24750.4	57253.9
Minimum GCV						8286.0	7872.0	7544.0	8915.3	9700.3	9630.2	8416.5	8240.1	9536.9	8982.9	8815.4	9499.3
Selected Case						CASE 4	CASE 3	CASE 2	CASE 3	CASE 3	CASE 2	CASE 13	CASE 5	CASE 1	CASE 8	CASE 1	CASE 1

D6: GCV of the combination cases of predictors for ASO rainfall under B2

CASE	Variable					Season of predictors (Lead time in months)											
	SAT	SLP	SXW	SYW	RAIN	AMJ (4)	MAM (5)	FMA (6)	JFM (7)	DJF (8)	NDJ (9)	OND (10)	SON (11)	ASO (12)	JAS (13)	JJA (14)	MJJ (15)
CASE 1	X					7862.3	8080.5	7496.8	7458.7	7712.7	7616.9	7325.0	7404.3	7305.5	6632.6	6192.0	6965.2
CASE 2		X				7795.2	6450.8	8489.2	7896.5	7685.8	7676.5	7930.0	7733.4	13058.2	7865.6	8285.8	8072.7
CASE 3			X			7810.5	7834.7	7139.1	7786.7	7698.7	7680.9	7080.9	7481.7	7655.7	7715.0	7583.0	7555.9
CASE 4				X		7976.2	8091.9	7411.8	7720.0	8053.1	8011.9	7781.7	7796.1	7971.3	7936.0	8311.2	6811.3
CASE 5					X	7964.1	12919.2	6828.4	8030.8	7804.2	7138.4	7944.7	8655.0	8641.7	8673.4	7008.4	8555.0
CASE 6	X	X				7732.4	8191.0	9966.4	8165.7	9594.0	7694.9	8872.3	8988.7	13153.5	7663.3	5998.3	9945.2
CASE 7	X		X			9090.3	8491.6	7890.8	7950.3	9710.9	8894.5	8044.5	8045.4	8042.4	7806.5	7397.1	8291.2
CASE 8	X			X		9154.3	9248.9	8384.0	8266.8	8368.6	8239.5	8740.6	8800.3	9264.6	7319.3	7342.7	7257.4
CASE 9	X				X	10217.2	12439.8	7580.5	8861.5	8470.4	9085.5	9915.2	9324.8	9214.2	8329.0	6783.3	9882.6
CASE 10		X	X			9011.8	7927.2	9269.2	9709.2	10518.7	8565.5	8874.9	8681.0	11402.2	9566.0	11054.9	10388.2
CASE 11		X		X		9336.8	7863.3	9313.4	8892.0	10957.4	9266.4	9278.5	9025.1	14540.6	9182.5	9366.3	8074.2
CASE 12		X			X	9833.7	11182.0	7496.9	9557.2	9176.4	8103.8	9574.2	9769.2	14609.2	11740.0	8044.4	8926.3
CASE 13			X	X		8760.1	9418.6	8183.5	9567.7	9152.7	8601.7	7872.9	8074.7	8602.3	9891.4	11197.1	9012.7
CASE 14			X		X	10862.0	16861.1	7037.8	9464.8	8676.6	8040.0	9027.2	11383.7	10740.4	9134.0	9318.0	8623.9
CASE 15				X	X	9381.8	17354.9	8109.8	9628.9	11495.2	8191.5	9308.2	10260.4	9885.6	10379.6	8140.7	7782.2
CASE 16	X	X	X			9779.7	15446.5	14585.6	9364.1	12413.2	10121.6	10504.6	11564.3	11531.1	11920.8	8638.6	12226.6
CASE 17	X	X		X		10293.0	10507.8	11621.2	10197.2	16113.5	9746.4	12721.7	12607.7	14827.0	9872.3	7699.6	7968.2
CASE 18	X	X			X	10279.3	11204.1	9889.1	11749.4	11880.0	9620.5	11823.8	11478.7	13584.4	9235.2	7461.0	13281.3
CASE 19	X		X	X		10750.1	12188.9	9985.2	9968.8	10847.4	11349.7	10984.5	10146.3	9965.9	10622.8	10021.4	9333.5
CASE 20	X		X		X	14234.8	15478.2	13952.1	13956.5	11859.2	10520.6	12156.0	11402.2	8917.5	8928.5	9562.5	9314.2
CASE 21	X			X	X	13037.2	14680.8	11354.6	11781.0	11949.8	10439.2	14078.2	12136.5	10638.2	8617.5	8081.9	9930.7
CASE 22		X	X	X		11460.8	10282.6	11595.1	12311.0	13588.3	11498.5	11243.2	12658.5	16909.8	13002.1	14009.3	11029.7
CASE 23		X	X		X	17711.7	12745.9	12313.9	12106.7	14333.6	10762.5	12935.5	11352.6	16098.7	12893.2	12054.4	12892.9
CASE 24		X		X	X	11775.3	13066.7	8709.0	17175.6	18749.6	10620.7	11834.1	12252.7	16212.1	13965.9	9404.7	8730.9
CASE 25			X	X	X	13946.1	18199.7	8215.7	16650.4	25057.6	13764.3	11843.5	13457.8	11652.8	11743.4	12359.3	13461.8
CASE 26	X	X	X	X		15303.7	23836.9	18941.1	12888.8	17234.4	13171.2	19950.2	24295.2	18420.1	14146.3	12942.0	11048.2
CASE 27	X	X	X		X	16637.9	21190.5	22170.6	17140.5	13936.9	14138.5	17240.4	13802.0	15477.4	11824.6	11281.8	12862.1
CASE 28	X	X		X	X	14709.9	13694.8	14402.8	23101.4	17710.3	13225.5	18062.1	17416.0	16228.3	9823.6	9290.9	10038.5
CASE 29	X		X	X	X	22211.2	14269.5	26282.1	25930.1	27336.8	17292.1	24718.7	14337.7	14811.6	8622.5	13937.3	12046.0
CASE 30		X	X	X	X	23912.6	17541.8	13863.2	28067.8	28751.2	27630.4	18848.6	16892.2	20870.7	15632.6	15575.2	19770.3
CASE 31	X	X	X	X	X	27925.2	30668.1	43395.6	44569.0	24641.4	13615.5	29415.6	32271.3	26811.9	12703.7	16215.8	15745.1
Minimum GCV						**7732.4**	**6450.8**	**6828.4**	**7458.7**	**7685.8**	**7138.4**	**7080.9**	**7404.3**	**7305.5**	**6632.6**	**5998.3**	**6811.3**
Selected Case						**CASE 6**	**CASE 2**	**CASE 5**	**CASE 1**	**CASE 2**	**CASE 5**	**CASE 3**	**CASE 1**	**CASE 1**	**CASE 1**	**CASE 6**	**CASE 4**

D7: GCV of the combination cases of predictors for NDJ rainfall under B2

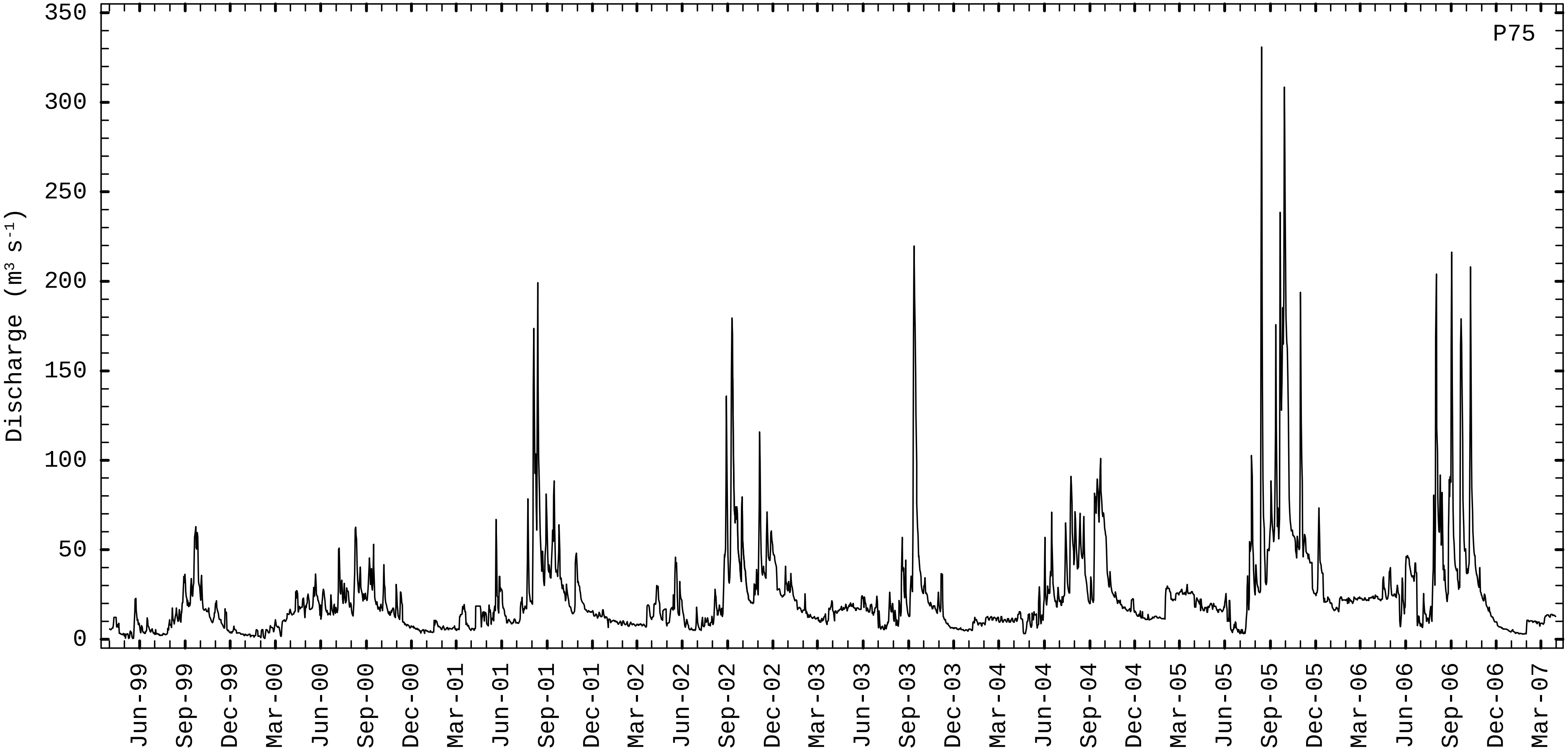
CASE	Variable				Season of predictors (Lead time in months)											
	SAT	SXW	SYW	RAIN	JAS (4)	JJA (5)	MJJ (6)	AMJ (7)	MAM (8)	FMA (9)	JFM (10)	DJF (11)	NDJ (12)	OND (13)	SON (14)	ASO (15)
CASE 1	X				2849.3	3829.4	3583.3	2780.9	2087.3	2754.9	2844.9	2569.9	2781.8	2815.2	2775.5	2574.8
CASE 2		X			3446.0	2692.1	2559.8	3092.8	2672.8	3120.2	2512.7	2987.6	2711.0	2581.9	3366.1	2543.2
CASE 3			X		2647.1	3546.6	3055.3	2596.7	2974.7	3435.0	3129.8	2843.8	3241.0	2905.8	4901.0	3135.7
CASE 4				X	3003.3	3077.7	3335.5	3088.9	2949.8	2404.8	3091.1	3923.4	2611.3	2970.2	3061.0	2762.3
CASE 5	X	X			2926.5	3018.9	2865.1	4086.0	3606.5	4750.2	3766.0	3743.2	2642.3	3770.3	3982.6	4324.7
CASE 6	X		X		3711.2	6321.6	4854.8	3743.7	2582.8	4882.9	3213.5	3681.7	4889.0	4254.2	11848.3	3808.3
CASE 7	X			X	3055.9	4417.2	6461.8	3247.3	3317.1	5286.1	4735.6	5408.8	4039.1	4788.7	4192.5	3113.9
CASE 8		X	X		2452.2	5416.0	4835.3	4726.2	3785.3	5566.4	4414.2	4615.9	5115.4	3164.0	9836.8	25336.4
CASE 9		X		X	4140.2	3513.3	4440.0	3952.4	3372.3	4354.8	4531.2	5214.2	3089.8	3588.2	4754.6	4122.5
CASE 10			X	X	5953.0	3976.1	4416.6	4625.8	2671.8	2632.5	3837.7	5074.3	3678.5	1482.3	3563.4	5598.0
CASE 11	X	X	X		3659.7	4241.1	7404.6	8200.0	7178.4	7737.2	7233.2	6761.8	6911.5	11739.5	9962.8	64315.0
CASE 12	X	X		X	2890.8	3943.2	6110.5	5377.8	3955.2	5398.2	6040.9	10307.5	8612.2	6159.9	10563.5	4865.7
CASE 13	X		X	X	8292.3	3657.3	10035.4	4868.1	3997.4	4667.3	7605.1	7012.4	5588.7	2720.6	9094.4	5257.3
CASE 14		X	X	X	9446.3	6492.7	9185.6	9002.2	3256.2	4144.7	10146.0	7757.4	4790.8	2309.7	28998.6	5569.2
CASE 15	X	X	X	X	12367.0	5440.9	16964.9	17369.2	6685.7	20371.4	18723.1	11619.5	19341.1	13194.3	10739.4	39863.6
Minimum GCV					2452.2	2692.1	2559.8	2596.7	2087.3	2404.8	2512.7	2569.9	2611.3	1482.3	2775.5	2543.2
Selected Case					CASE 8	CASE 2	CASE 2	CASE 3	CASE 1	CASE 4	CASE 2	CASE 1	CASE 4	CASE 10	CASE 1	CASE 2

D8: GCV of the combination cases of predictors for FMA rainfall under B2

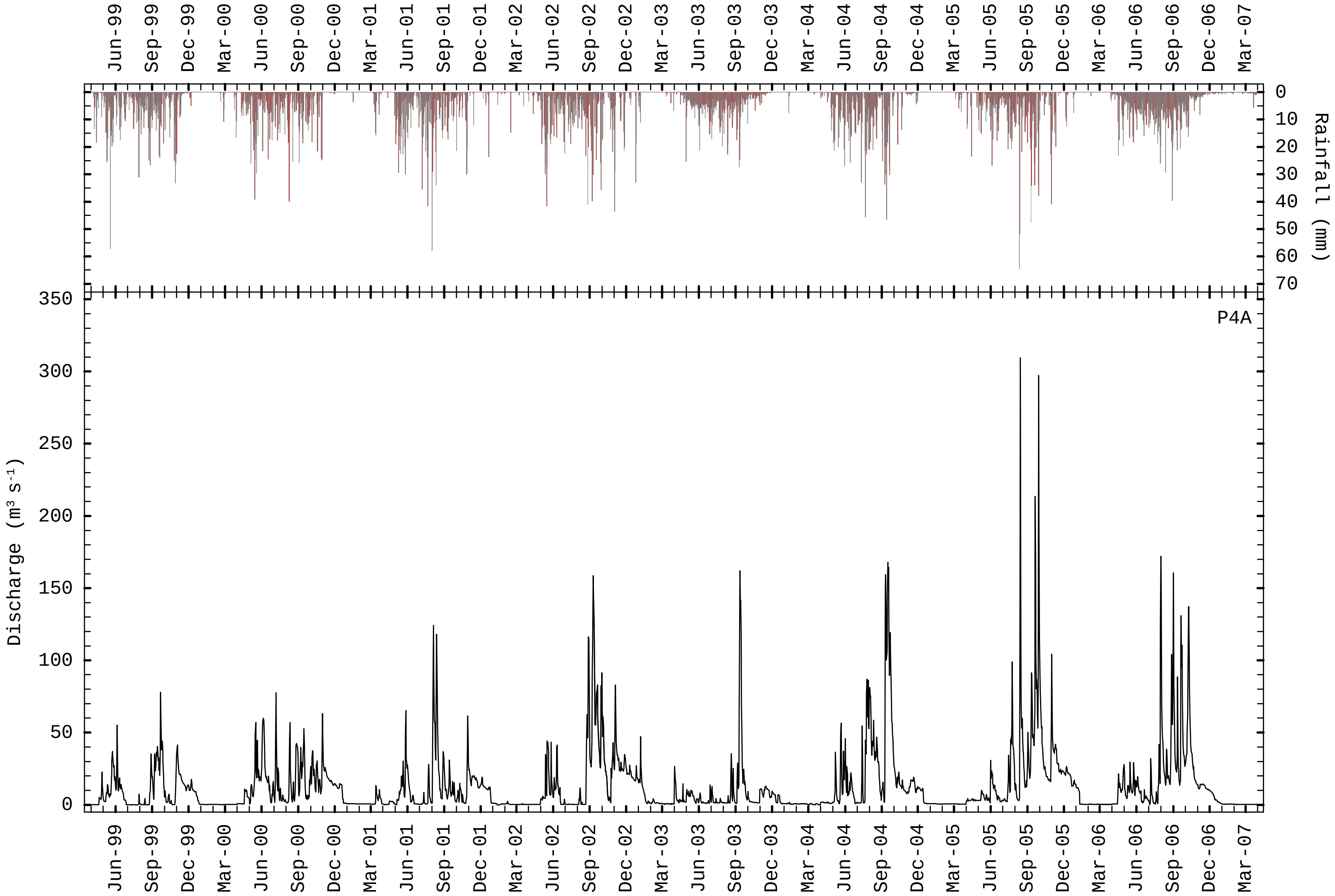
CASE	Variable					Season of predictors (Lead time in months)											
	SAT	SLP	SXW	SYW	RAIN	OND (4)	SON (5)	ASO (6)	JAS (7)	JJA (8)	MJJ (9)	AMJ (10)	MAM (11)	FMA (12)	JFM (13)	DJF (14)	NDJ (15)
CASE 1	X					2440.5	2616.1	3023.9	2541.9	2532.6	3396.6	3756.4	3183.3	3001.9	2636.1	2602.3	2676.5
CASE 2		X				2565.3	4742.5	2508.8	2625.8	2493.2	2668.9	2525.0	2402.8	2766.2	2946.5	2651.4	2651.3
CASE 3			X			2406.5	2300.1	1974.6	2390.3	2054.4	3030.1	2461.8	2142.0	3508.2	4460.8	2735.4	2879.5
CASE 4				X		2496.1	2475.4	2349.1	2979.5	2170.1	2115.8	2404.3	2333.3	2572.1	2785.4	2303.9	2324.4
CASE 5					X	2610.1	2903.3	3371.3	2554.1	3159.6	2934.3	2609.9	2803.6	2465.3	2661.3	2692.6	2442.1
CASE 6	X	X				19345.6	10649.2	8403.0	4050.3	3248.5	4914.0	5492.2	9873.5	6752.3	3963.0	3438.6	3504.7
CASE 7	X		X			4001.7	2783.7	4331.3	2985.0	2215.2	7024.1	3886.9	4824.6	5453.5	8217.4	3565.1	6071.7
CASE 8	X			X		3128.2	4867.1	4993.5	4550.1	2729.2	4765.0	4233.9	7296.3	7311.0	3602.8	4241.6	6094.9
CASE 9	X				X	3046.0	4692.4	8597.7	3038.9	5013.8	4050.2	2644.2	5446.1	4346.2	3535.4	4522.2	3205.0
CASE 10		X	X			3138.8	8432.8	5576.4	4047.3	2591.9	3566.8	6675.3	2941.2	3871.3	4670.7	2499.3	4217.5
CASE 11		X		X		3017.7	5442.6	3500.2	3884.2	2852.9	3026.4	3239.6	3202.0	3331.9	4473.1	3045.5	4579.7
CASE 12		X			X	3089.9	3404.0	4565.1	2395.7	3258.0	3459.0	3464.5	3382.7	4107.5	6479.5	8344.9	3127.4
CASE 13			X	X		2857.0	3029.3	2478.3	4622.5	2230.6	3897.5	3435.3	2696.0	4181.7	5503.2	2429.4	4170.2
CASE 14			X		X	3277.9	4497.8	3467.3	3667.7	4207.7	4642.5	3087.8	4840.3	2998.0	6265.9	6192.1	3485.1
CASE 15				X	X	3520.6	3090.3	5227.7	5024.8	3980.0	3488.2	3105.8	4561.6	5603.0	3985.9	3364.8	15983.1
CASE 16	X	X	X			48777.7	10546.5	6758.1	7602.4	3765.6	10564.7	4291.5	12328.2	7373.5	6558.3	5176.6	4642.8
CASE 17	X	X		X		20842.5	12639.9	16688.0	6018.5	6539.7	6666.3	4918.0	12429.4	13527.5	3447.3	6726.0	5523.8
CASE 18	X	X			X	11864.1	8629.3	6660.1	4730.1	4108.9	6301.3	5993.8	16607.6	14935.3	7023.7	7831.7	7183.1
CASE 19	X		X	X		4885.6	6650.3	6499.8	5889.7	3861.6	7409.2	8131.1	10152.0	8938.1	11081.1	4508.2	12780.8
CASE 20	X		X		X	5273.6	8085.9	12308.7	4556.2	7052.4	4637.4	3570.8	8643.1	8321.1	6404.9	8697.2	7020.6
CASE 21	X			X	X	6279.3	5930.4	20098.3	5761.7	9094.6	5558.6	6148.0	12194.0	7567.6	6191.0	6327.1	25609.7
CASE 22		X	X	X		5891.1	7344.7	10859.9	8117.7	3239.3	4941.3	22630.9	6141.7	9244.8	7630.5	7089.9	4952.3
CASE 23		X	X		X	3283.1	6249.0	4699.7	3577.6	4148.9	6762.5	5008.5	7681.5	5981.9	7724.9	9350.4	8787.7
CASE 24		X		X	X	4489.7	13767.2	6052.3	4004.8	6872.3	4580.7	3707.7	8037.3	22867.4	7497.9	18987.0	22336.5
CASE 25			X	X	X	3129.9	5181.7	10218.6	6184.1	8200.9	5913.2	5123.4	8578.3	4370.7	10589.3	8059.3	47839.5
CASE 26	X	X	X	X		46448.0	26455.8	12762.3	14757.5	11504.1	12206.7	13646.8	9853.5	23842.7	28885.2	9994.5	10242.6
CASE 27	X	X	X		X	24409.0	20787.9	4221.5	6052.8	7506.1	15200.2	21004.9	17570.8	21023.6	8630.8	12434.3	27032.9
CASE 28	X	X		X	X	15856.2	28403.0	24147.7	7996.8	15346.3	11742.5	16296.7	40681.1	38539.1	16379.9	29518.7	17472.5
CASE 29	X		X	X	X	5565.0	15083.0	37622.8	7134.5	34829.8	10510.4	16477.2	8163.2	71410.4	15531.6	13039.9	71335.5
CASE 30		X	X	X	X	3837.0	27190.3	6705.0	11205.1	13140.9	6966.6	7361.1	27733.1	35118.7	18060.8	20075.0	119612.3
CASE 31	X	X	X	X	X	53132.9	293039.7	12729.2	26376.5	73322.5	38773.8	39523.7	10564.5	346095.5	61829.1	53327.2	70349.4
Minimum GCV						2406.5	2300.1	1974.6	2390.3	2054.4	2115.8	2404.3	2142.0	2465.3	2636.1	2303.9	2324.4
Selected Case						CASE 3	CASE 3	CASE 3	CASE 3	CASE 3	CASE 4	CASE 4	CASE 3	CASE 5	CASE 1	CASE 4	CASE 4

Appendix E: Hydrographs of the Daily Observed Streamflow at 12 Stations from April 1999 to March 2007

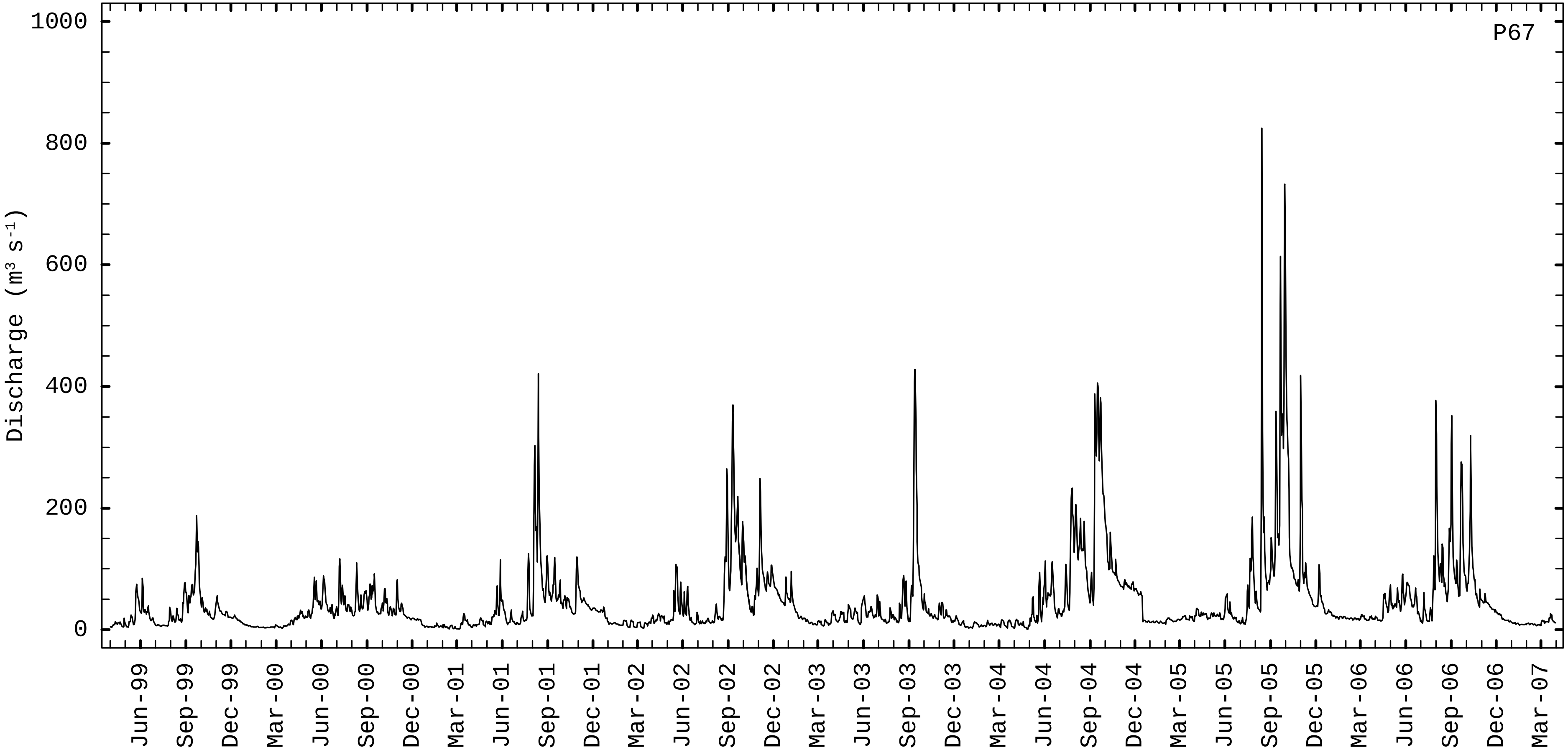
E1: Station P75



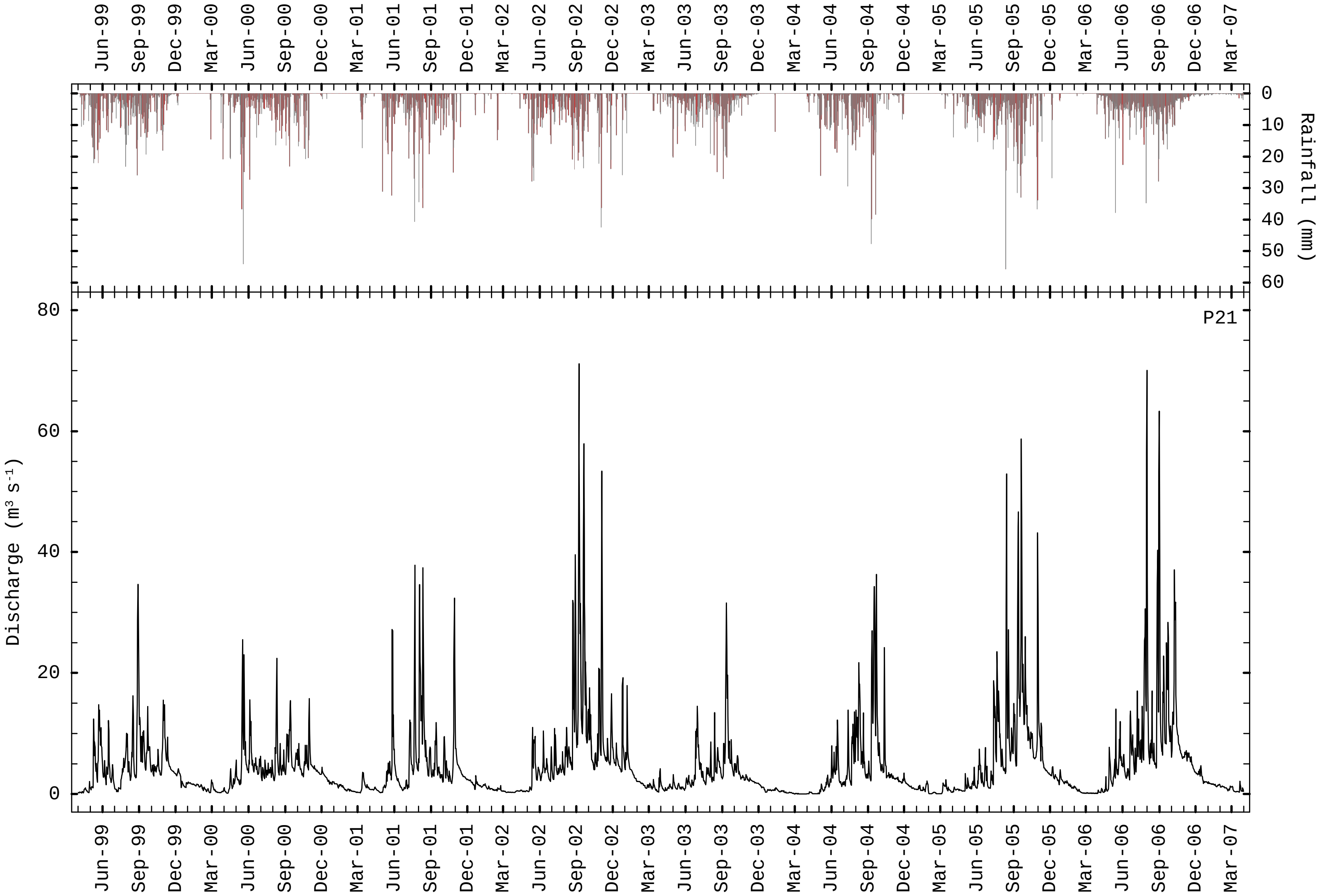
E2: Station P4A



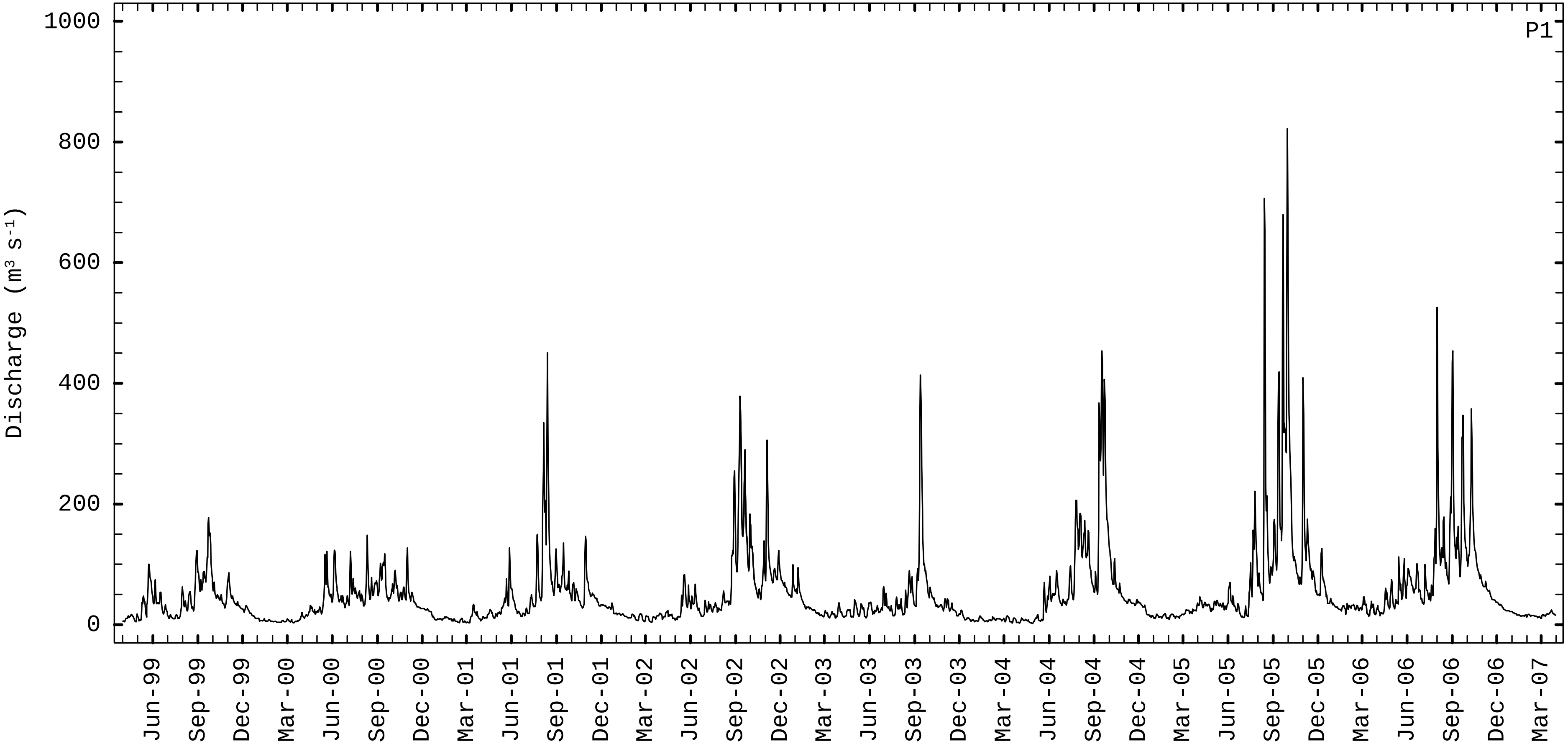
E3: Station P67



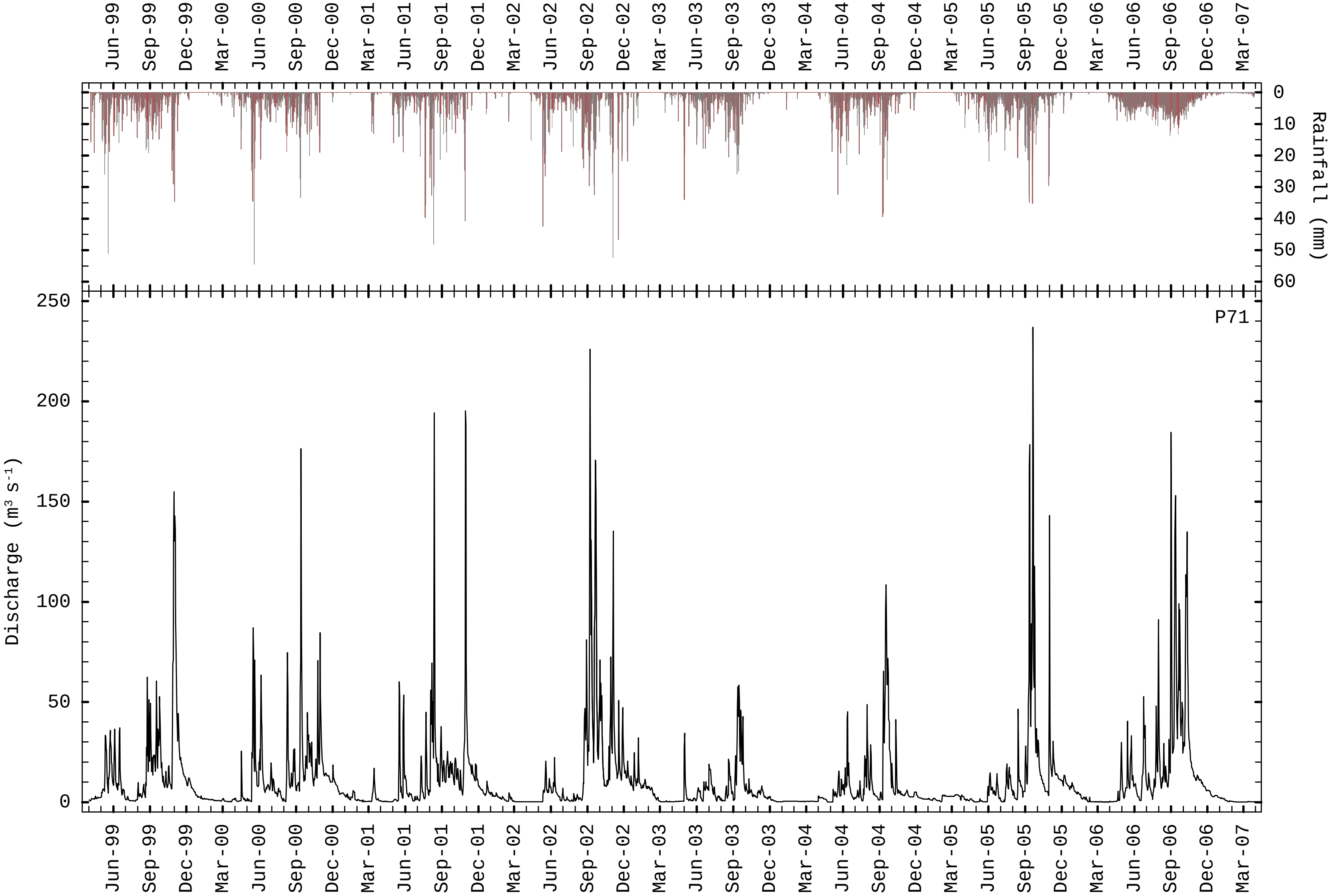
E4: Station P21



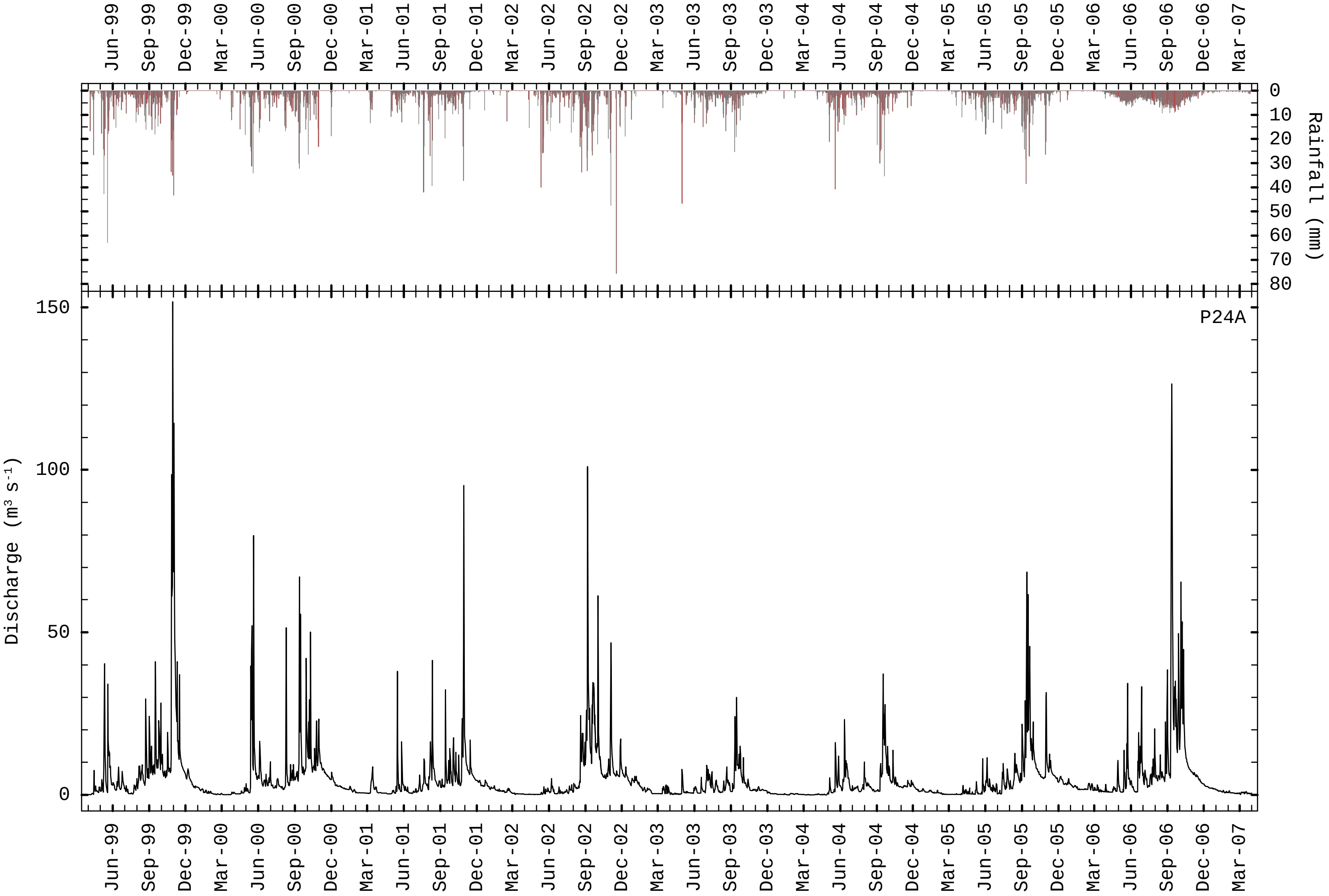
E5: Station P1



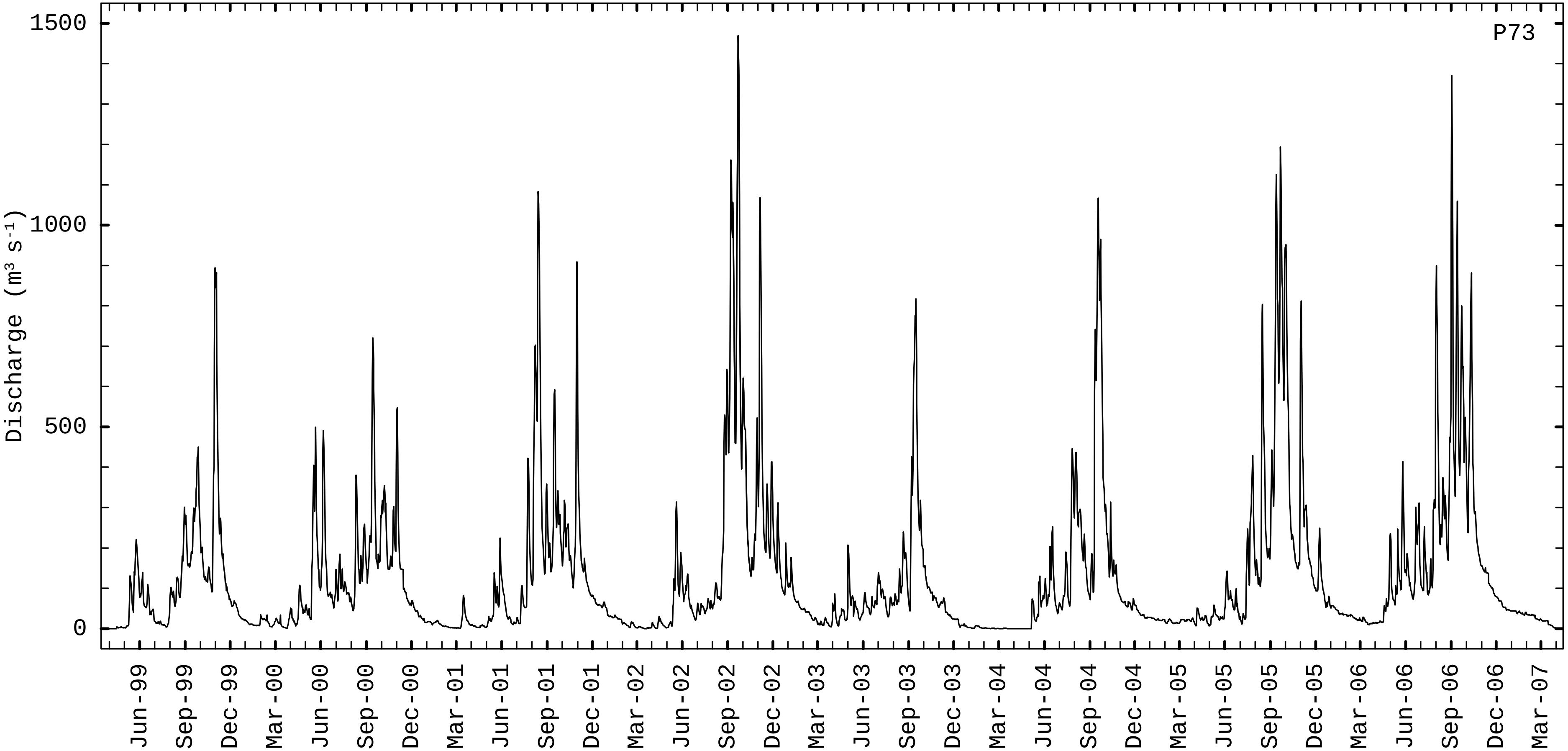
E6: Station P71



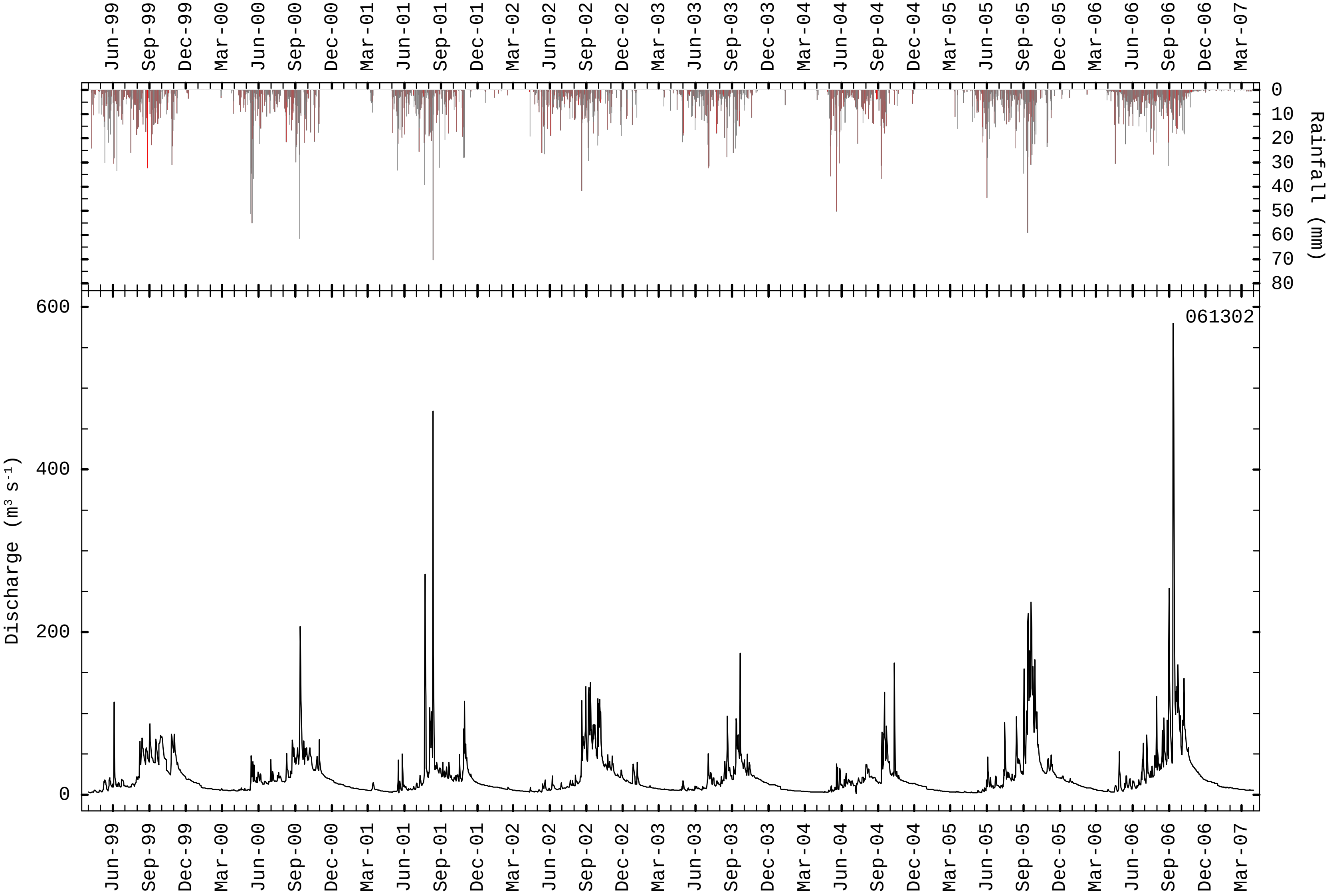
E7: Station P24A



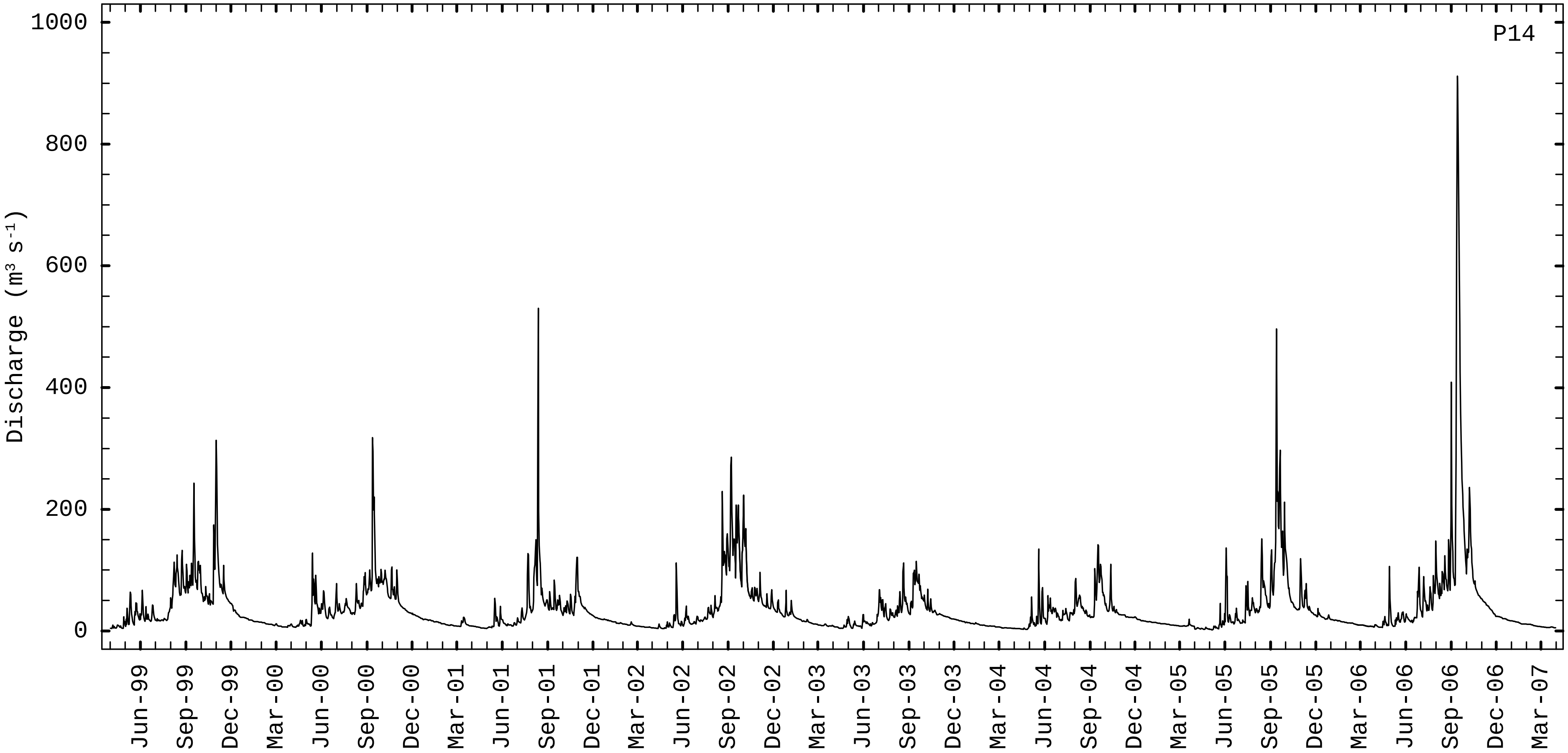
E8: Station P73



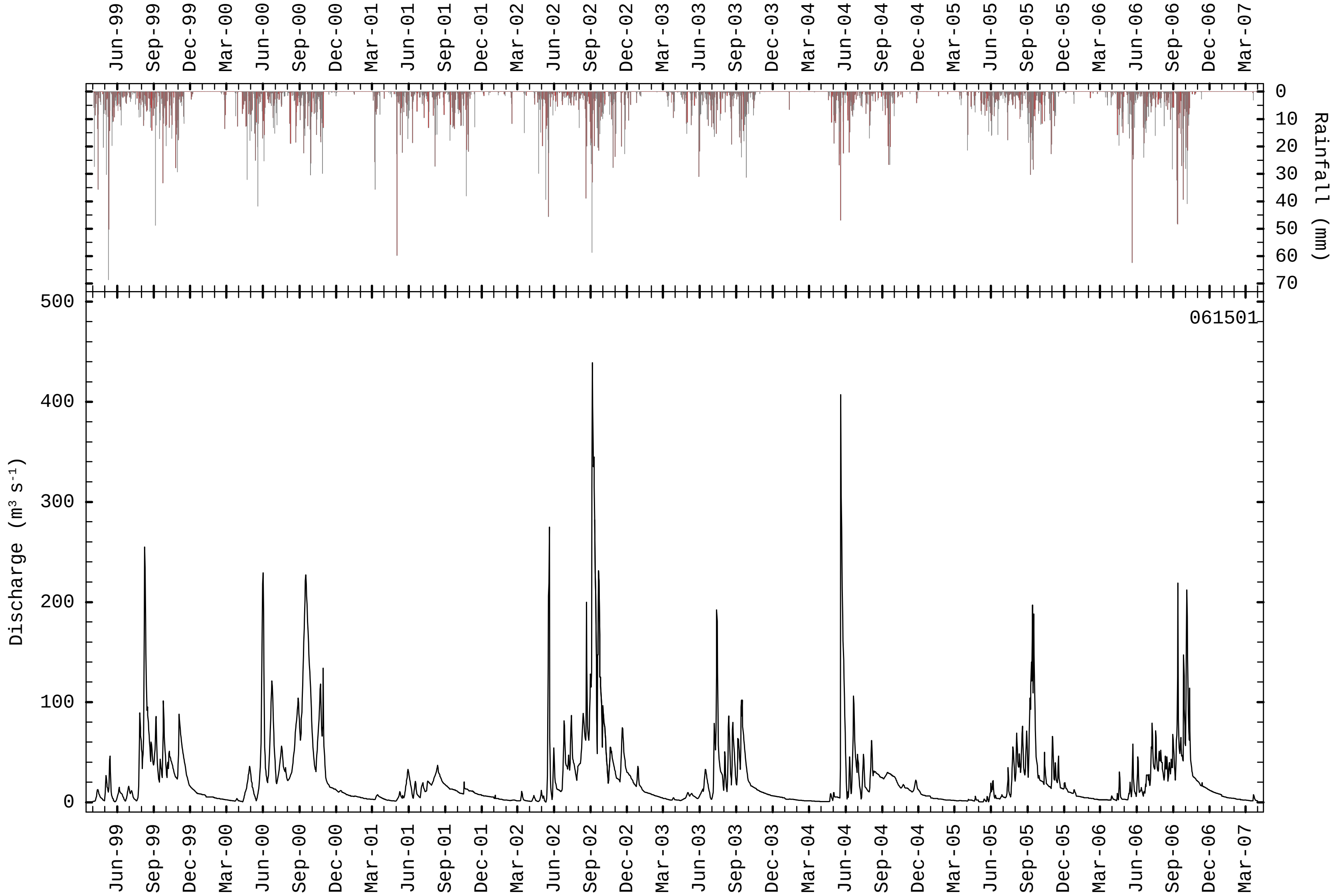
E9: Station 061302



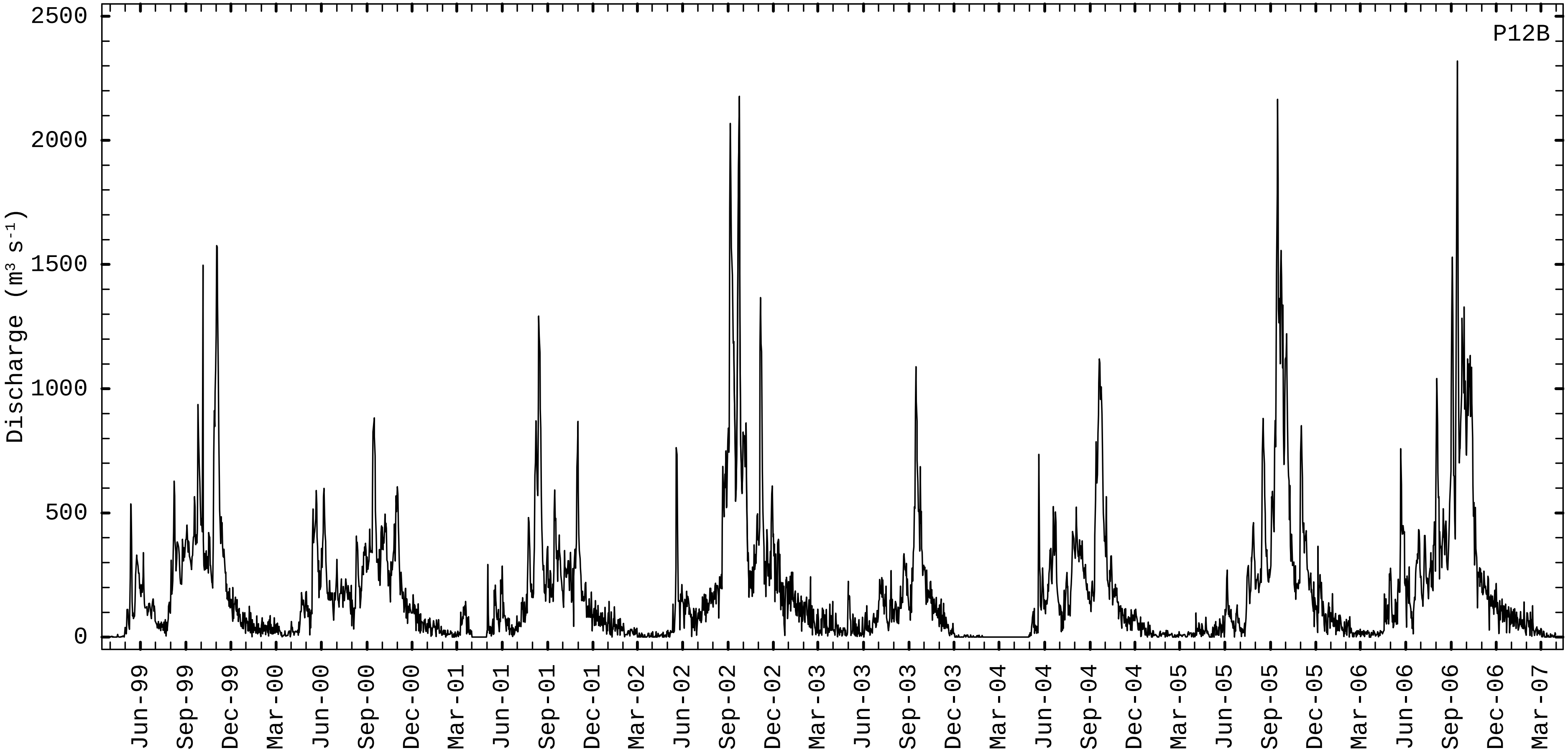
E10: Station P14



E11: Station 061501



E12: Station P12B



Appendix F: Model Parameters

F1: List of the parameters of HEC-HMS model

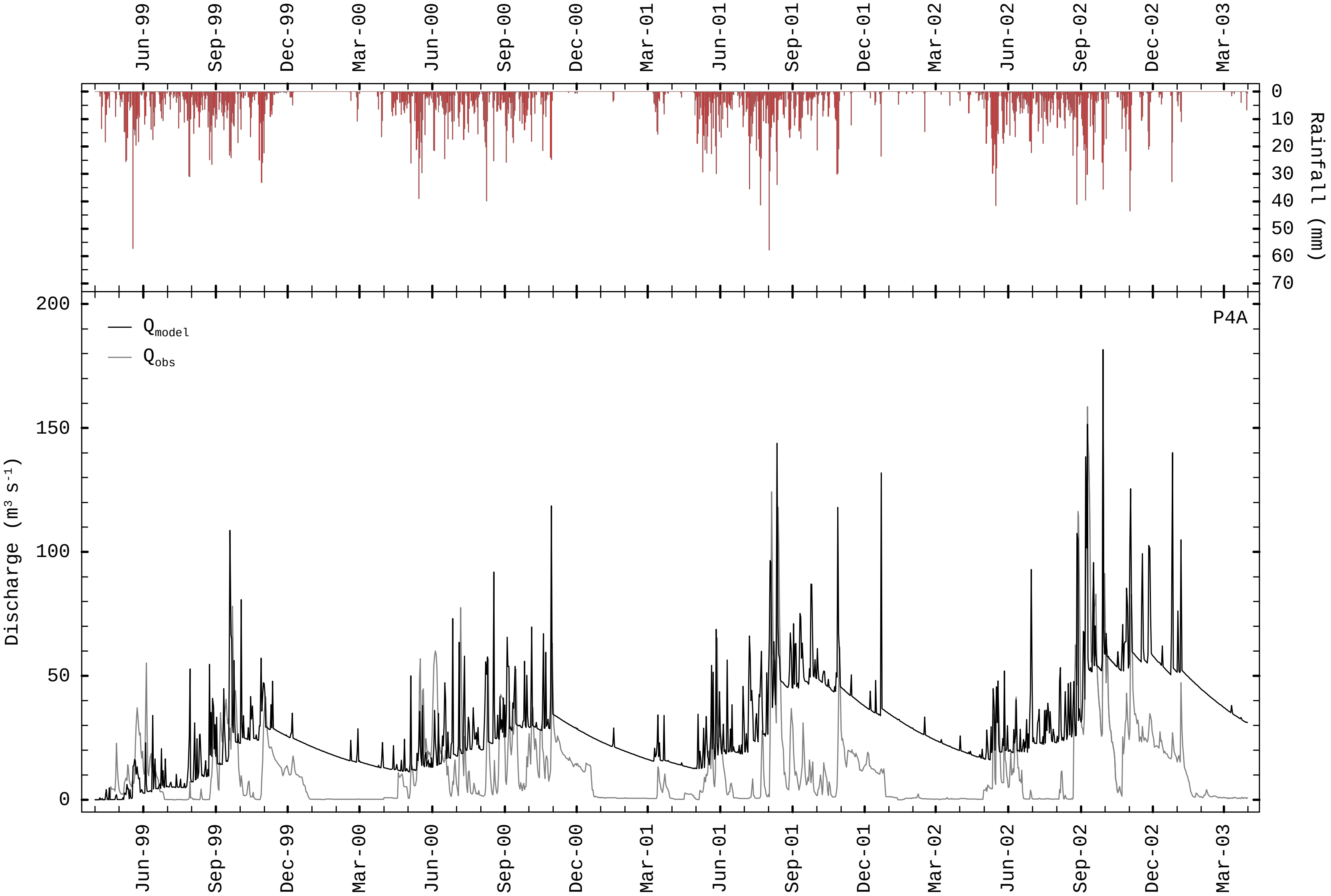
Component	Method	Parameter
Loss	Deficit and constant	Initial deficit (mm) Maximum deficit (mm) Constant rate (mm h^{-1}) Impervious (%)
	Exponential	Initial range (mm) Initial coefficient ((mm $h^{-1})^{1-x}$) Coefficient ratio (-) Exponent (-) Impervious (%)
	Green and Ampt	Initial content (-) Saturated content (-) Suction (mm) Conductivity (mm h^{-1}) Impervious (%)
	Initial and constant	Initial loss (mm) Constant rate (mm h^{-1}) Impervious (%)
	SCS curve number	Initial abstraction (mm) Curve number (-) Impervious (%)
	Smith Parlange	Initial content (-) Residual content (-) Saturated content (-) Bubbling pressure (mm) Pore distribution (-) Conductivity (mm h^{-1}) Impervious (%) Time series of temperature
	SMA	Soil (%) GW#1, 2 (%) Maximum infiltration (mm h^{-1}) Impervious (%) Soil storage (mm) Tension storage (mm) Soil percolation (mm h^{-1}) GW#1, 2 storage (mm) GW#1, 2 percolation (mm h^{-1}) GW#1, 2 coefficient (h)
Transform	Clark unit hydrograph	Time of concentration (h) Storage coefficient (h)
	Kinematic wave	Length (m) Slope (m m^{-1}) Manning’s n (-) Number of sub-reaches (-) Shape (-) Bottom width (m) Side slope (H:V)

Component	Method	Parameter
Transform	Kinematic wave	Area of collector and sub-collector (km^2) Roughness (-) Number of routing steps (-)
	ModClark	Time of concentration (h) Storage coefficient (h)
	SCS unit hydrograph	Graph type: standard or Delmarva Lag time (min)
	Snyder unit hydrograph	Method: standard, Ft worth district or Tulsa district Standard lag (h) Peaking coefficient (-)
Baseflow	Bounded recession	Initial type: discharge or discharge per area Initial value (m^3 s^{-1} or m^3 s^{-1} km^{-2}) Recession constant (-) Monthly values (m^3 s^{-1})
	Constant monthly	Monthly values (m^3 s^{-1})
	Linear reservoir	Initial type: discharge or discharge per area GW#1, 2 initial value (m^3 s^{-1} or m^3 s^{-1} km^{-2}) GW#1, 2 coefficient (-) GW#1, 2 number of reservoirs (-)
	Nonlinear Boussinesq	Initial type: discharge or discharge per area Initial value (m^3 s^{-1} or m^3 s^{-1} km^{-2}) Threshold type: ratio to peak or threshold discharge Threshold value: ratio (-) or flow (m^3 s^{-1}) Length (m) Conductivity (mm h^{-1}) Porosity (-)
	Recession	Initial type: discharge or discharge per area Initial value (m^3 s^{-1} or m^3 s^{-1} km^{-2}) Recession constant (-) Threshold type: ratio to peak or threshold discharge Threshold value: ratio (-) or flow (m^3 s^{-1})
Routing	Kinematic wave	Length (m) Slope (m m^{-1}) Manning's n (-) Invert (m) Number of sub-reaches (-) Shape: circle, deep, rectangle, trapezoid, triangle Bottom width (m) Side slope (H:V)
	Lag	Lag (min)
	Modified puls	Storage-discharge function (-) Number of sub-reaches (-) Initial condition: discharge or inflow=outflow Initial value (m^3 s^{-1})

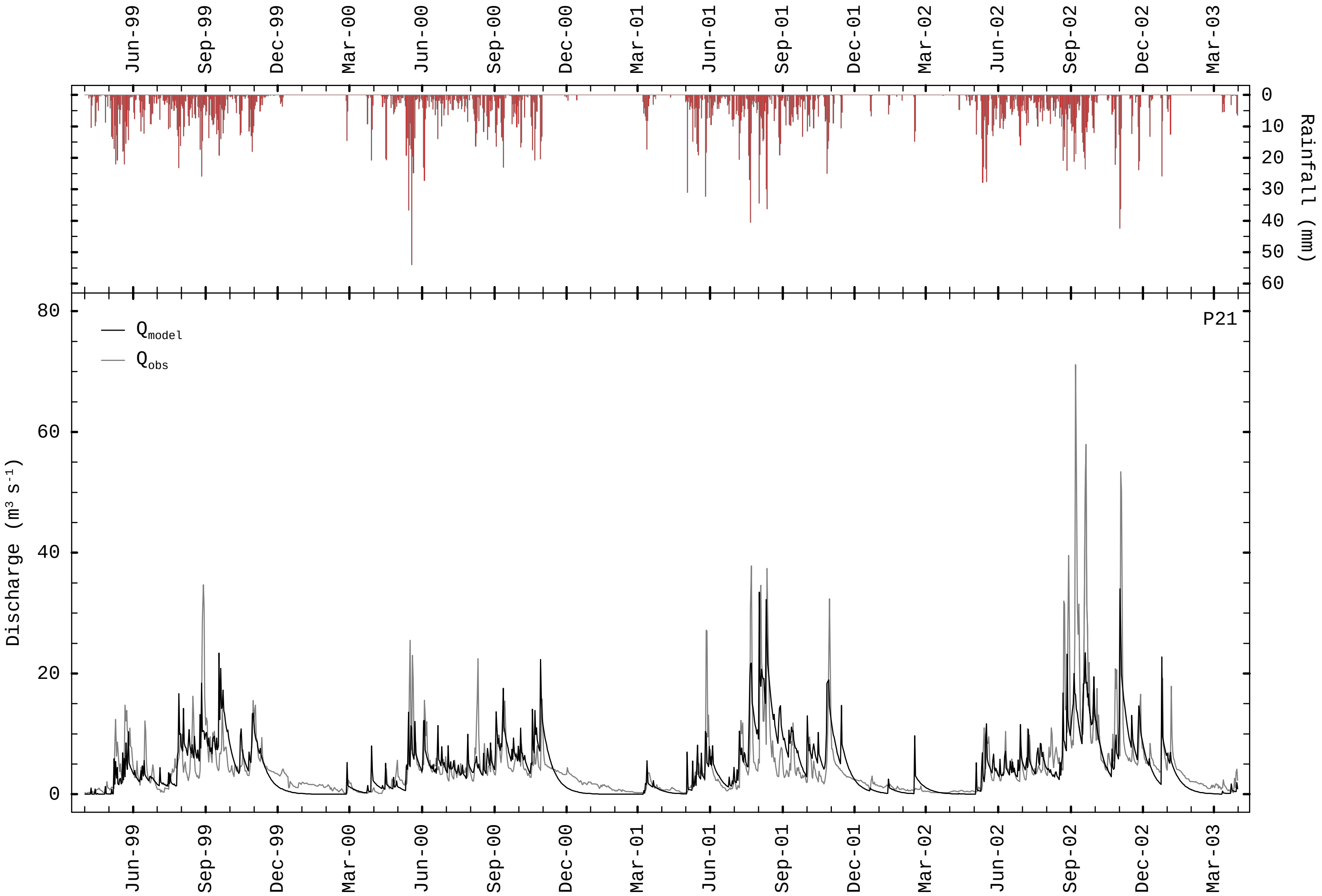
Component	Method	Parameter
Routing	Modified puls	Elevation-discharge function (-) Invert (m)
	Muskingum	Muskingum K (h) Muskingum X (-) Number of sub-reaches (-)
	Muskingum-Cunge	Length (m) Slope (m m^{-1}) Manning’s n (-) Invert (m) Shape: circle, deep, rectangle, trapezoid, triangle Bottom width (m) Side slope (H:V)
	Straddle stagger	Lag (min) Duration (min)

Appendix G: Hydrographs of Daily Simulated Streamflow at Six Stations from April 1999 to March 2003 by SIMHYD Model

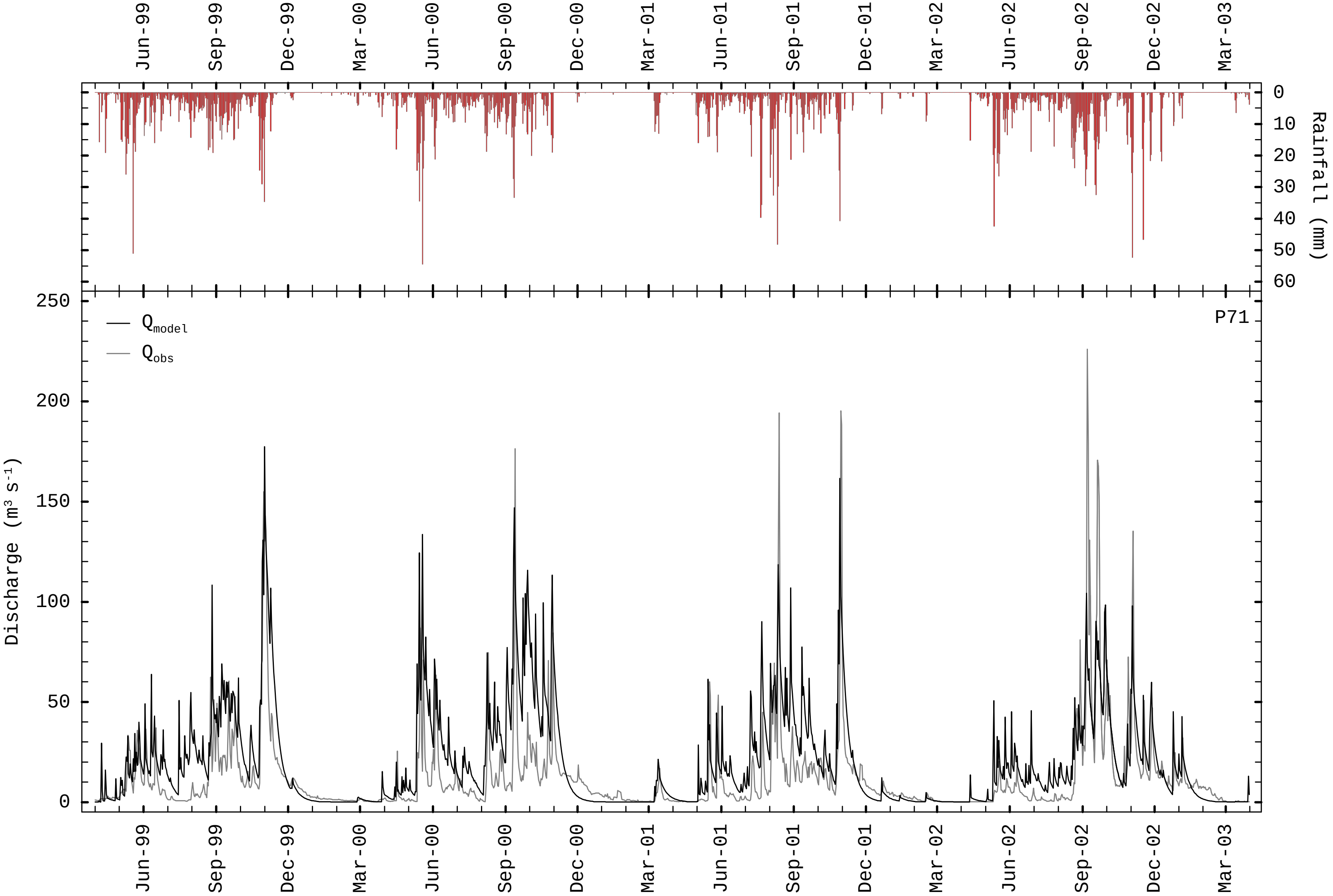
G1: Station P4A



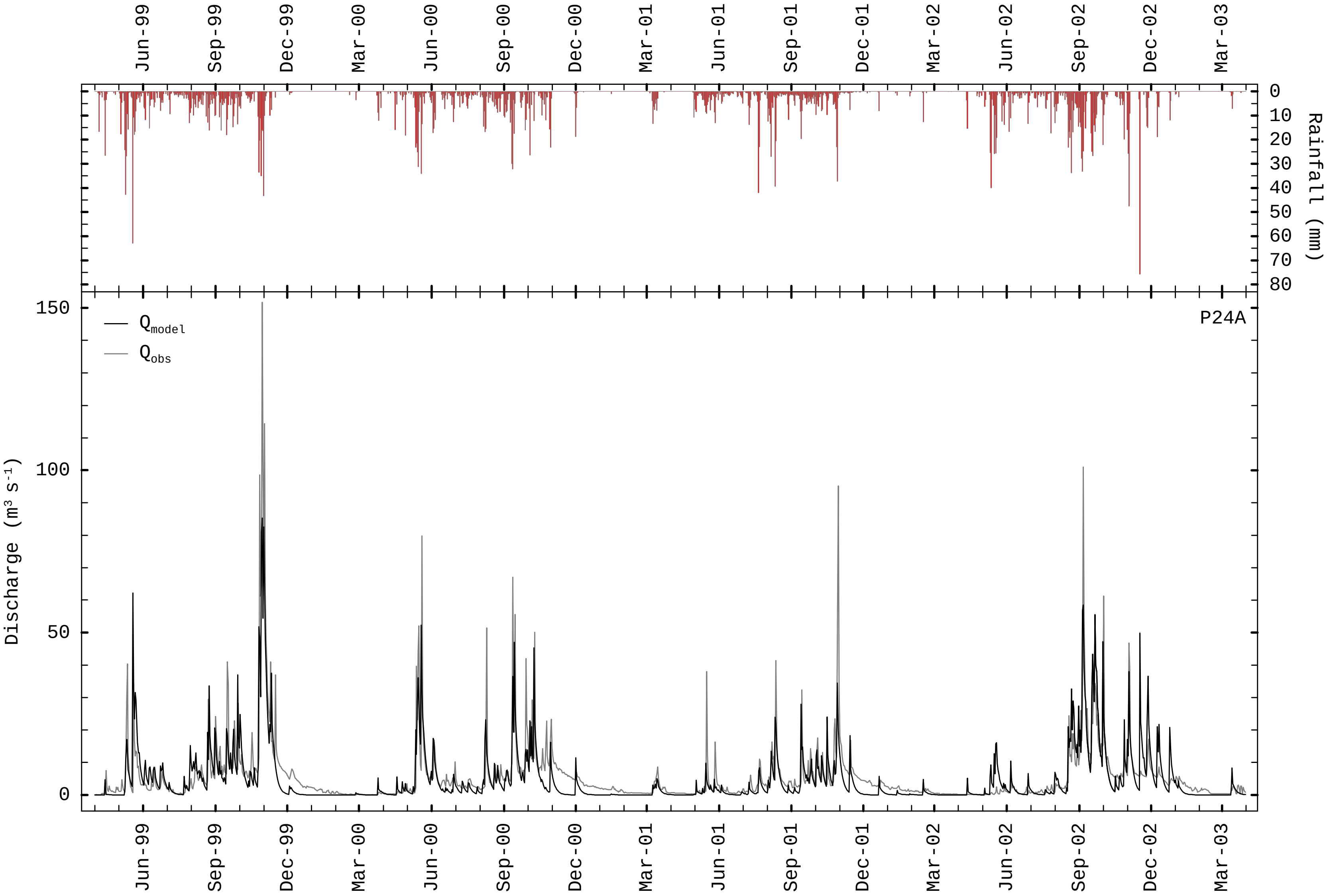
G2: Station P21



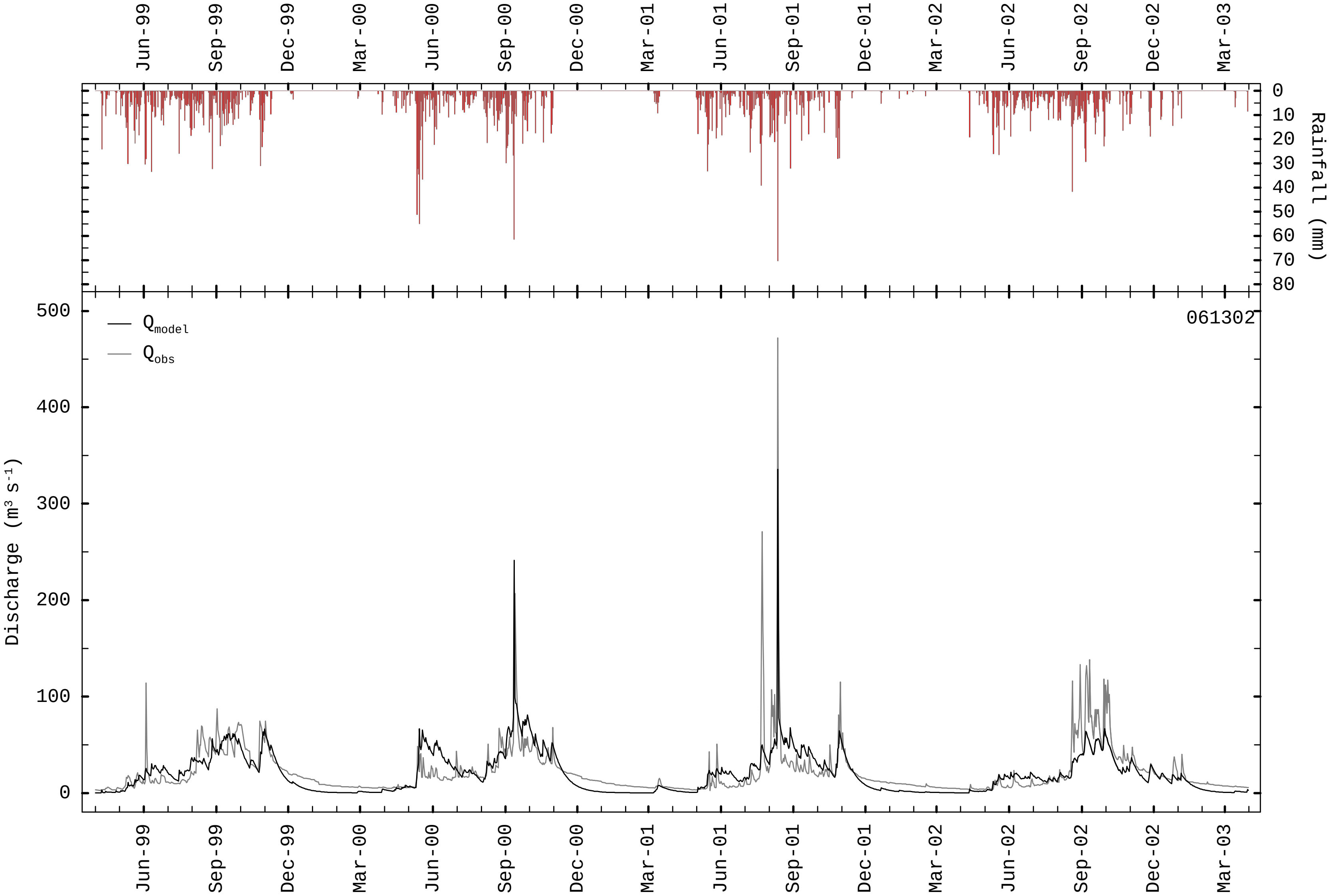
G3: Station P71



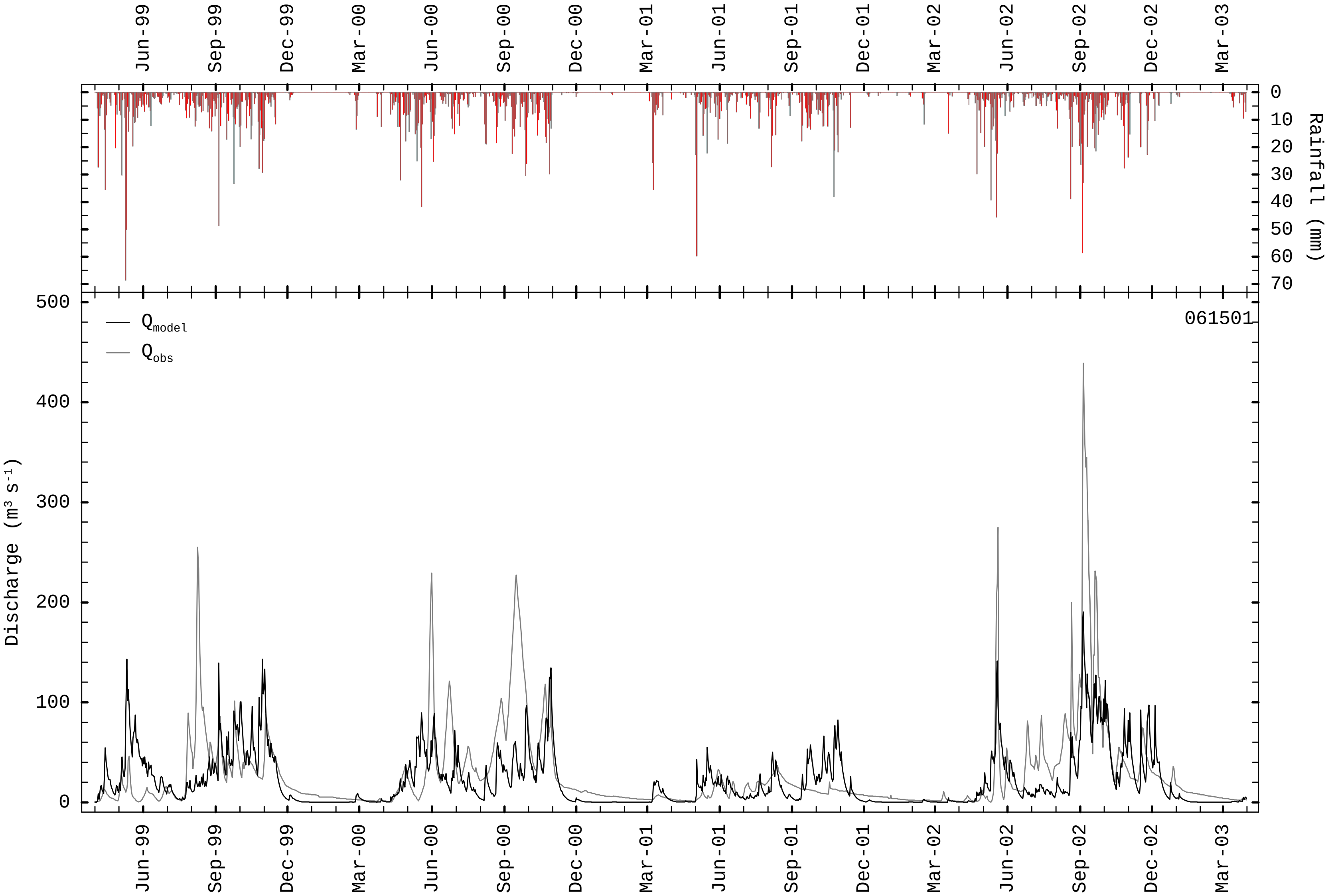
G4: Station P24A



G5: Station 061302

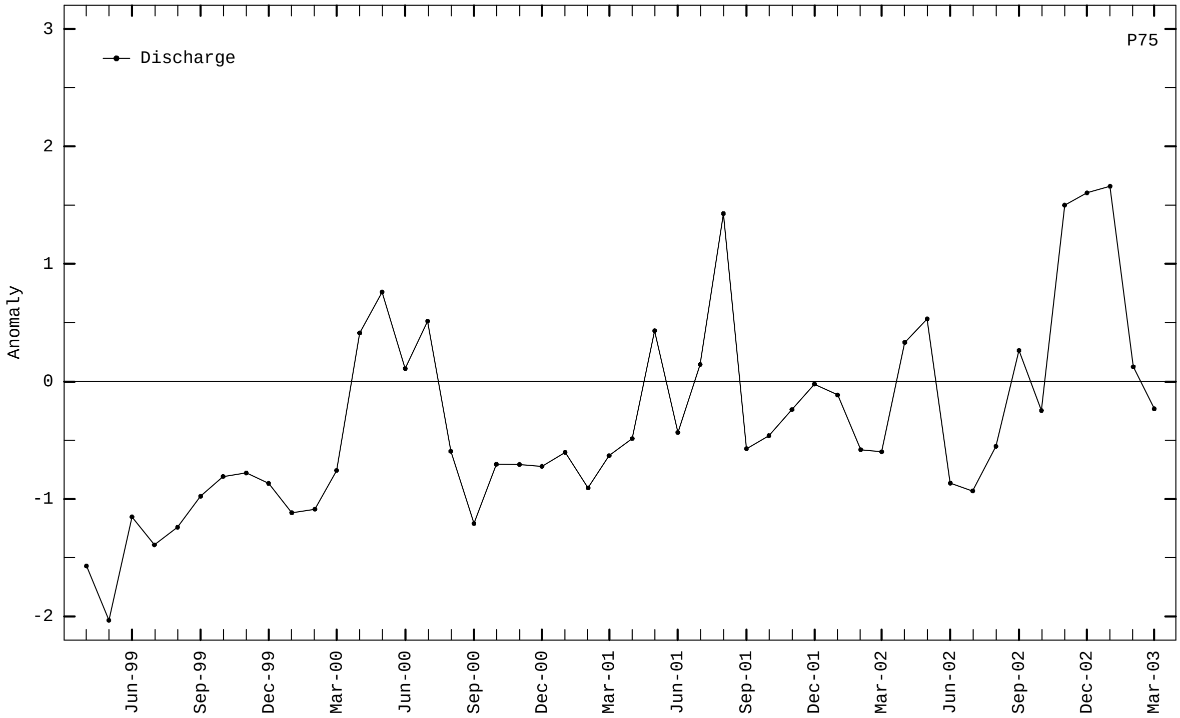


G6: Station 061501

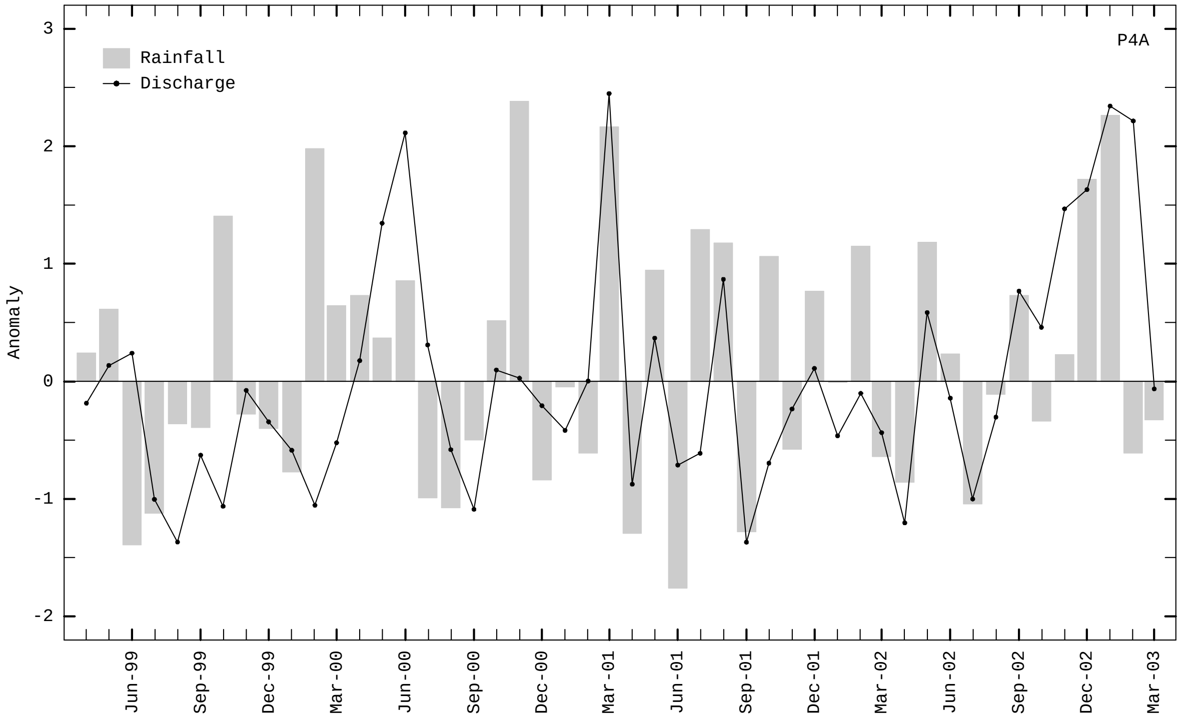


Appendix H: Monthly Anomalies of Observed Rainfall and Streamflow from April 1999 to March 2003

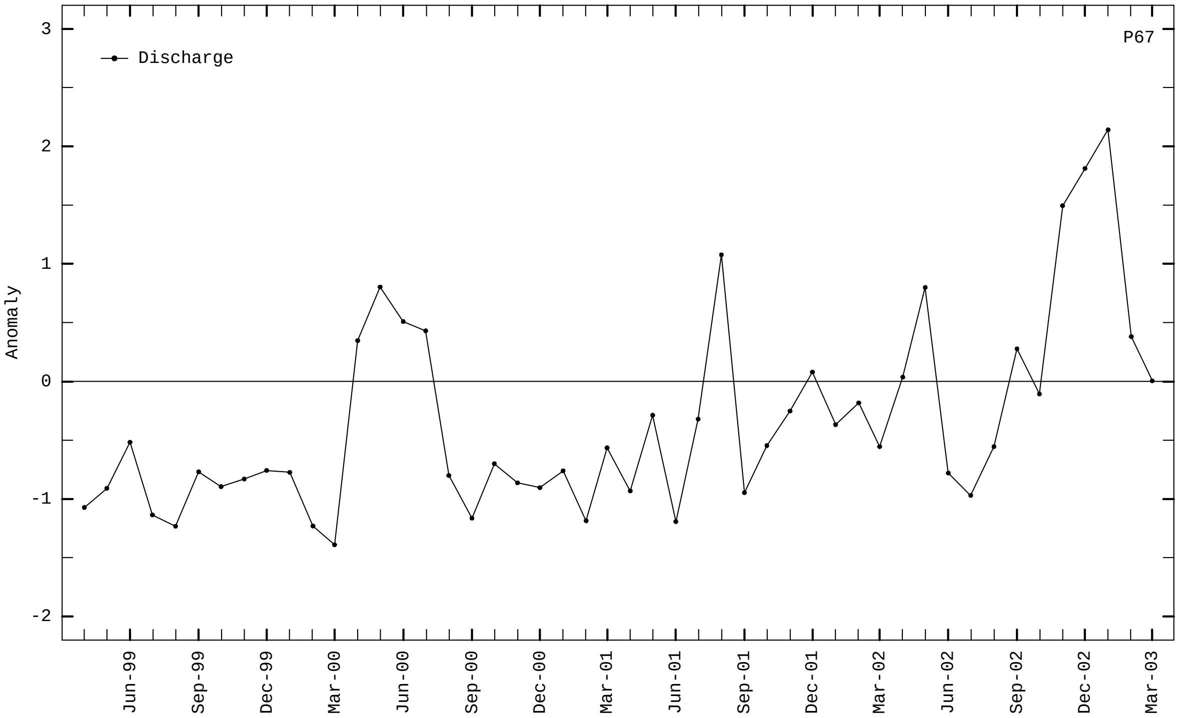
H1: Station P75



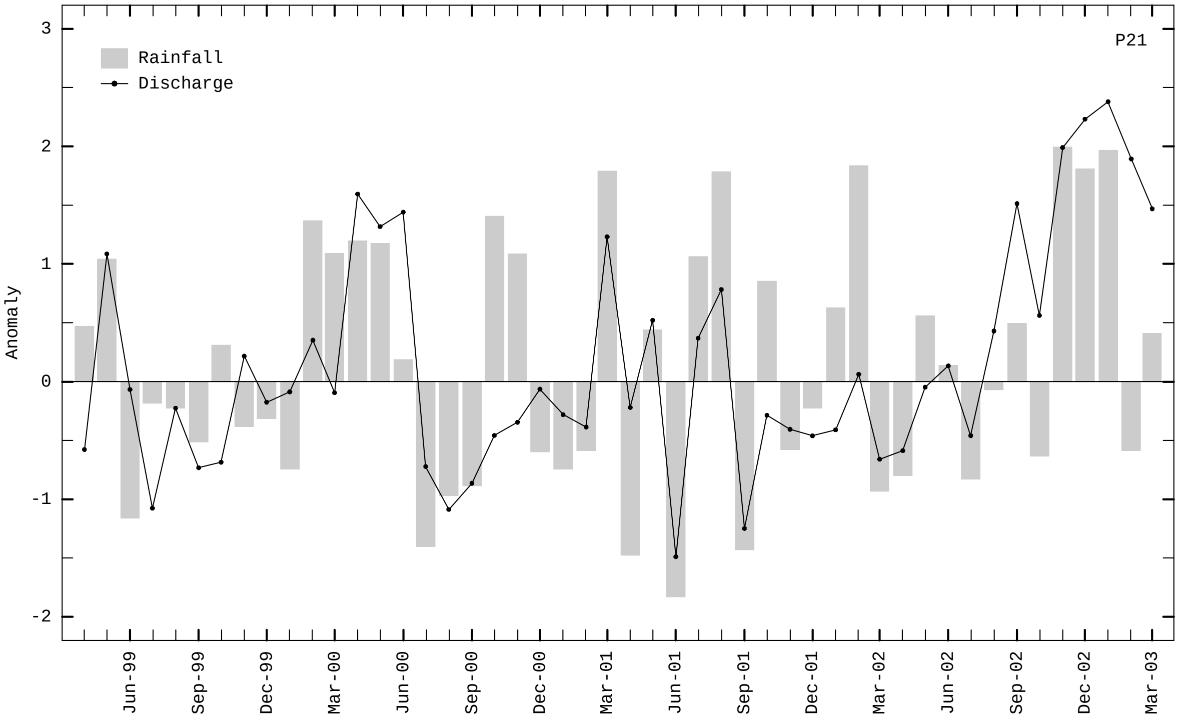
H2: Station P4A



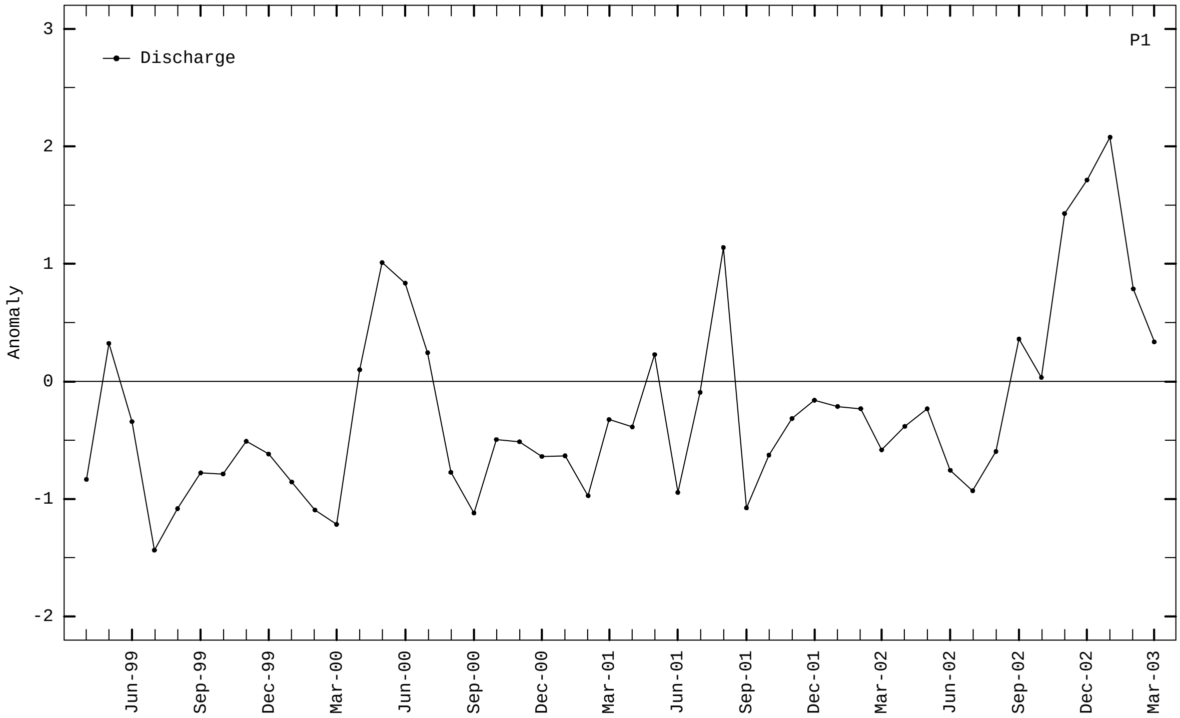
H3: Station P67



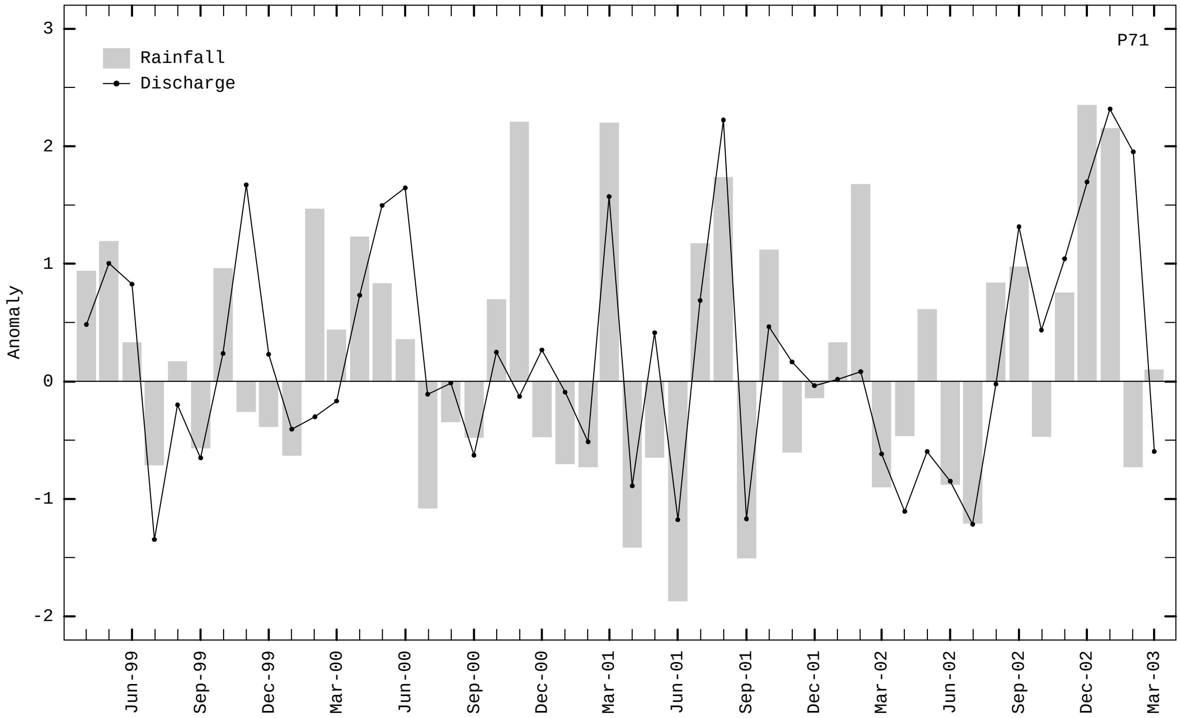
H4: Station P21



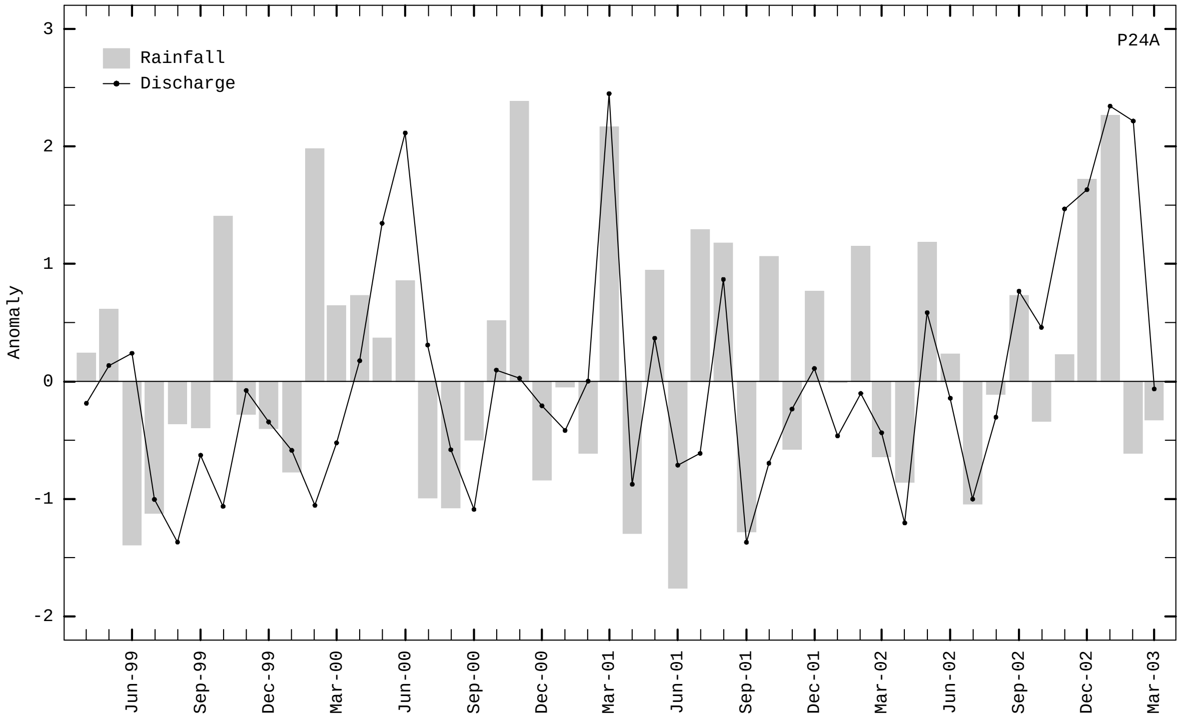
H5: Station P1



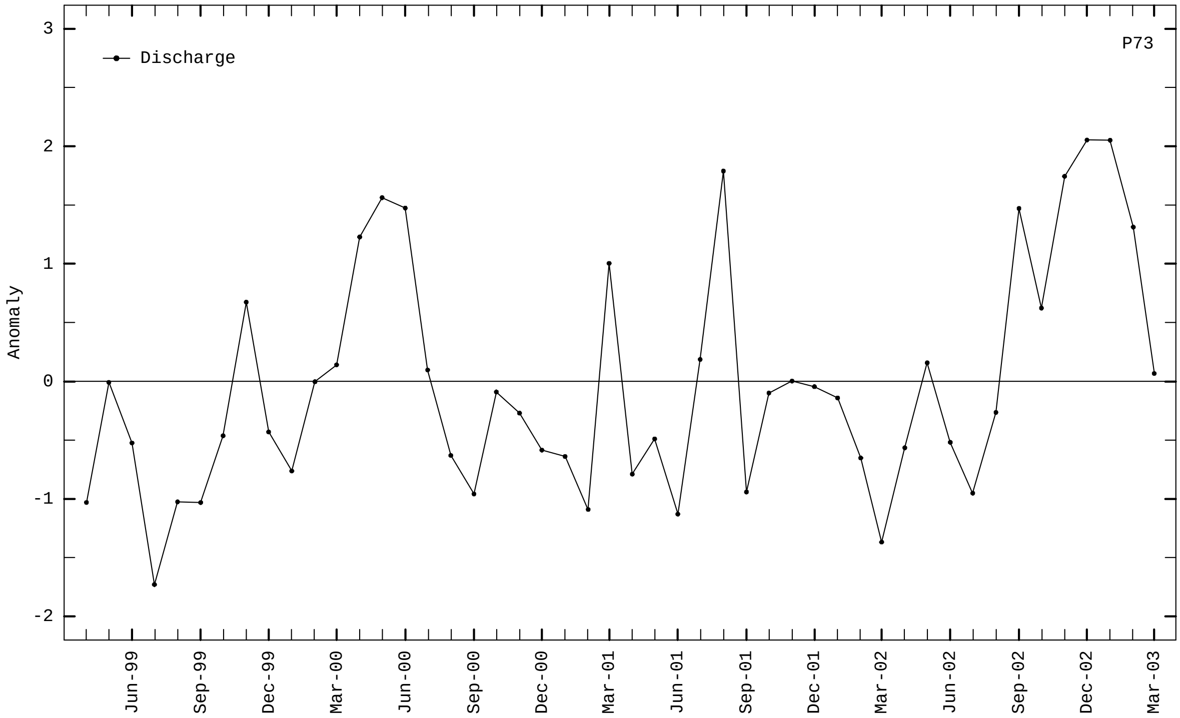
H6: Station P71



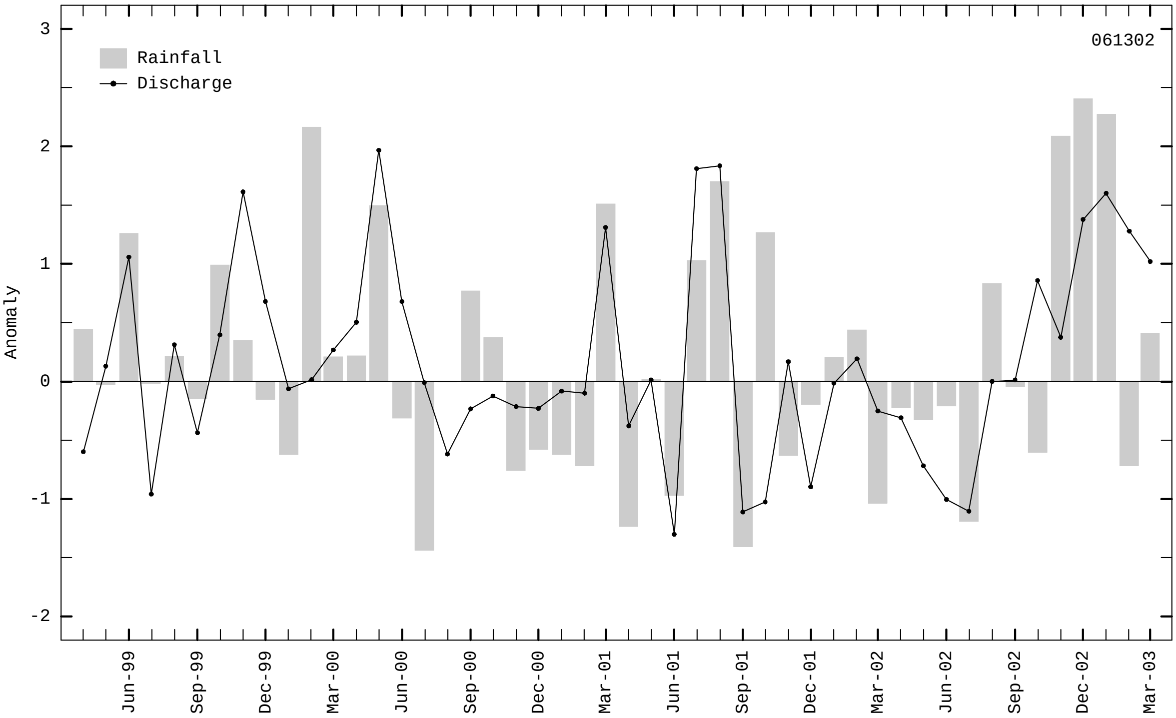
H7: Station P24A



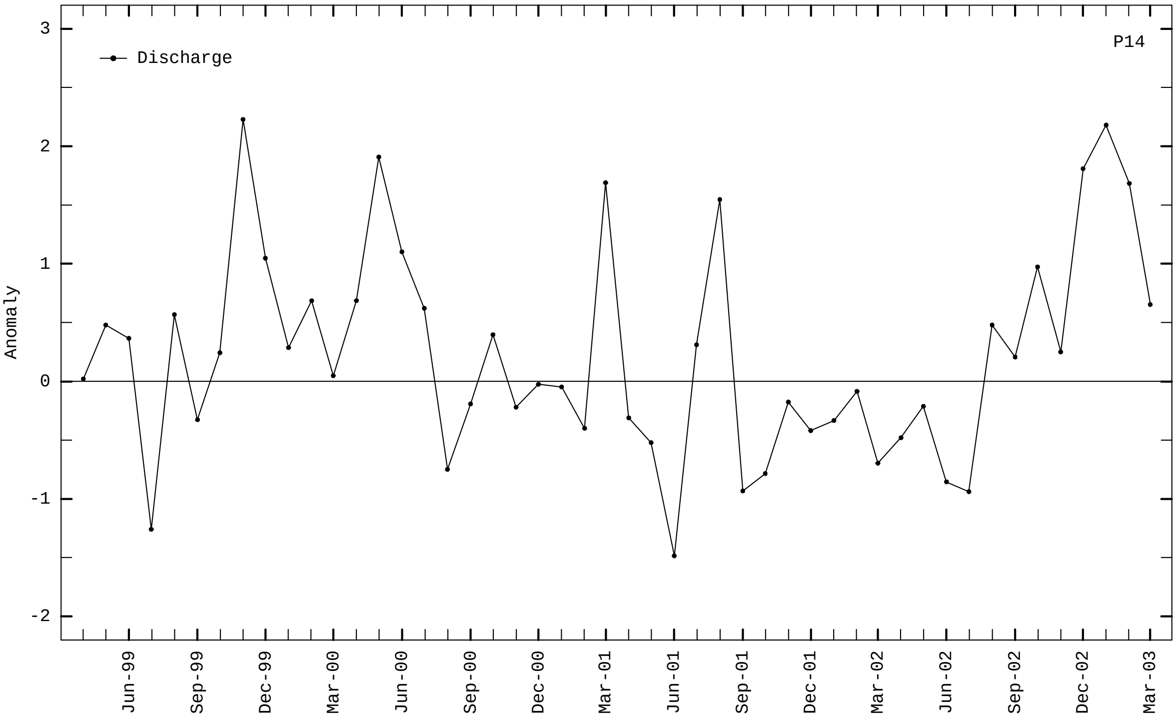
H8: Station P73



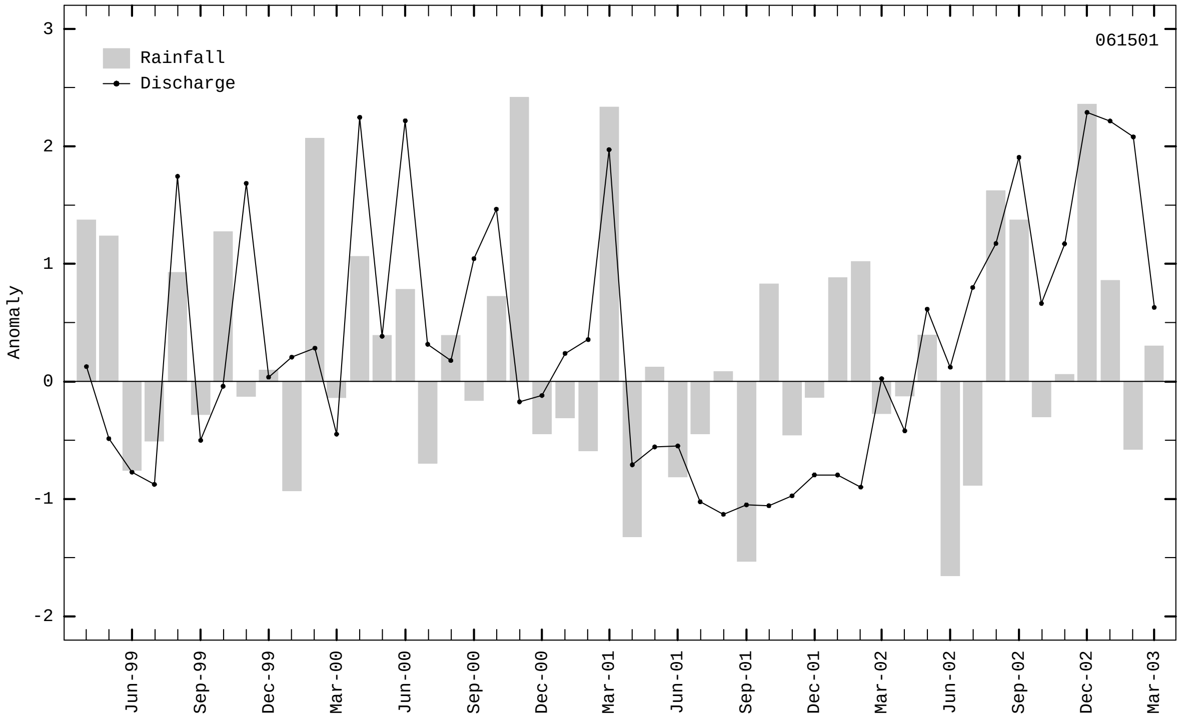
H9: Station 061302



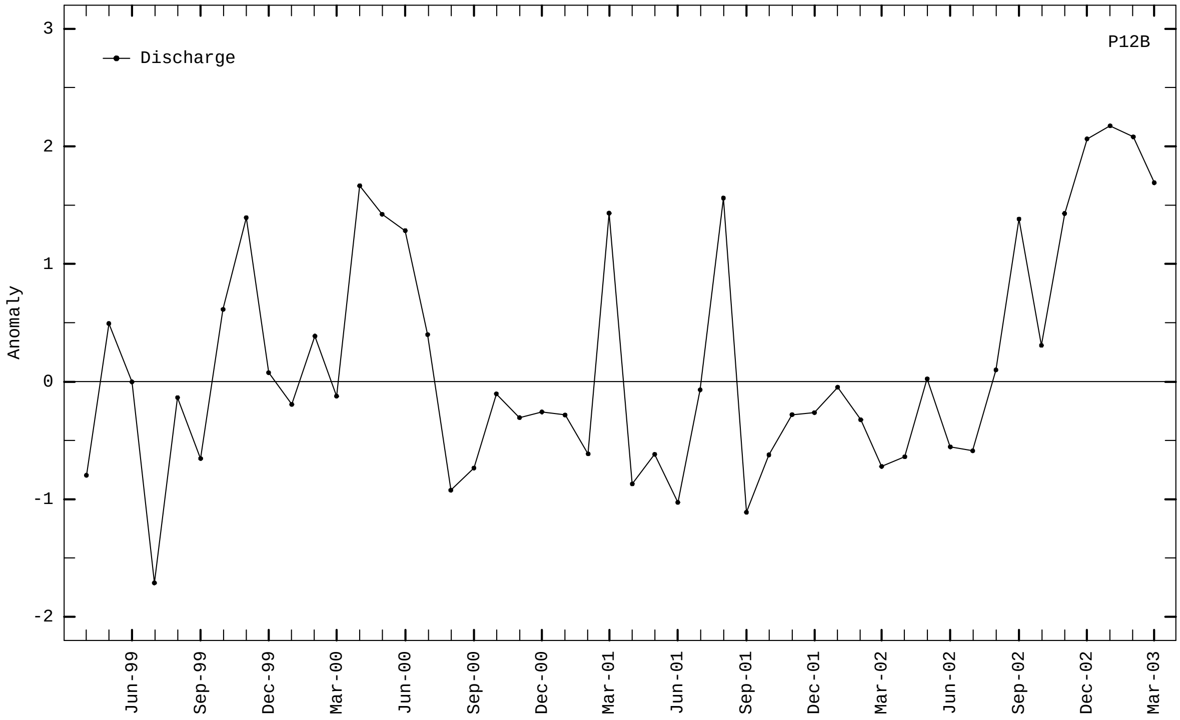
H10: Station P14



H11: Station 061501

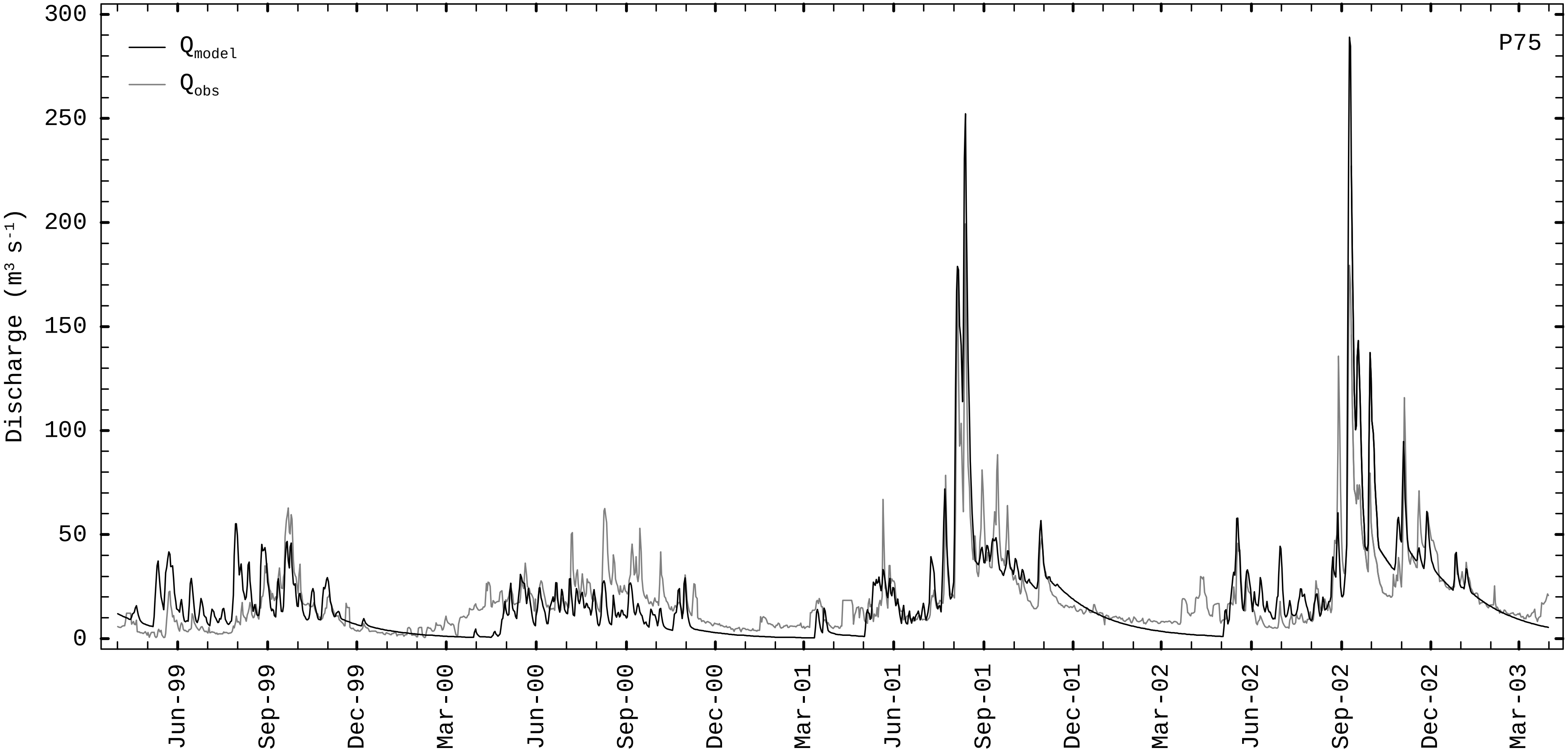


H12: Station P12B

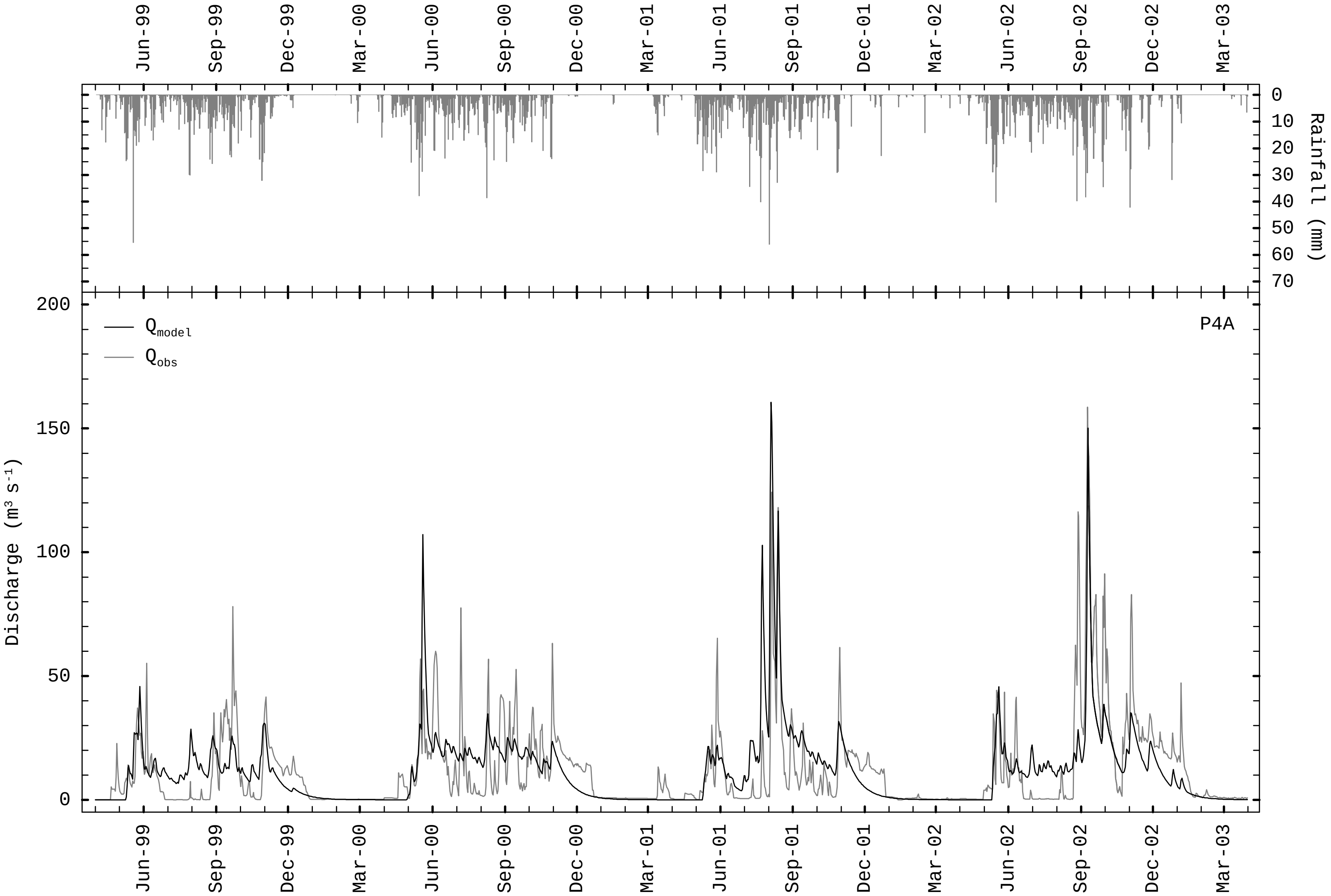


Appendix I: Hydrographs of Daily Simulated Streamflow at 12 Stations from April 1999 to March 2003 by HEC-HMS Model

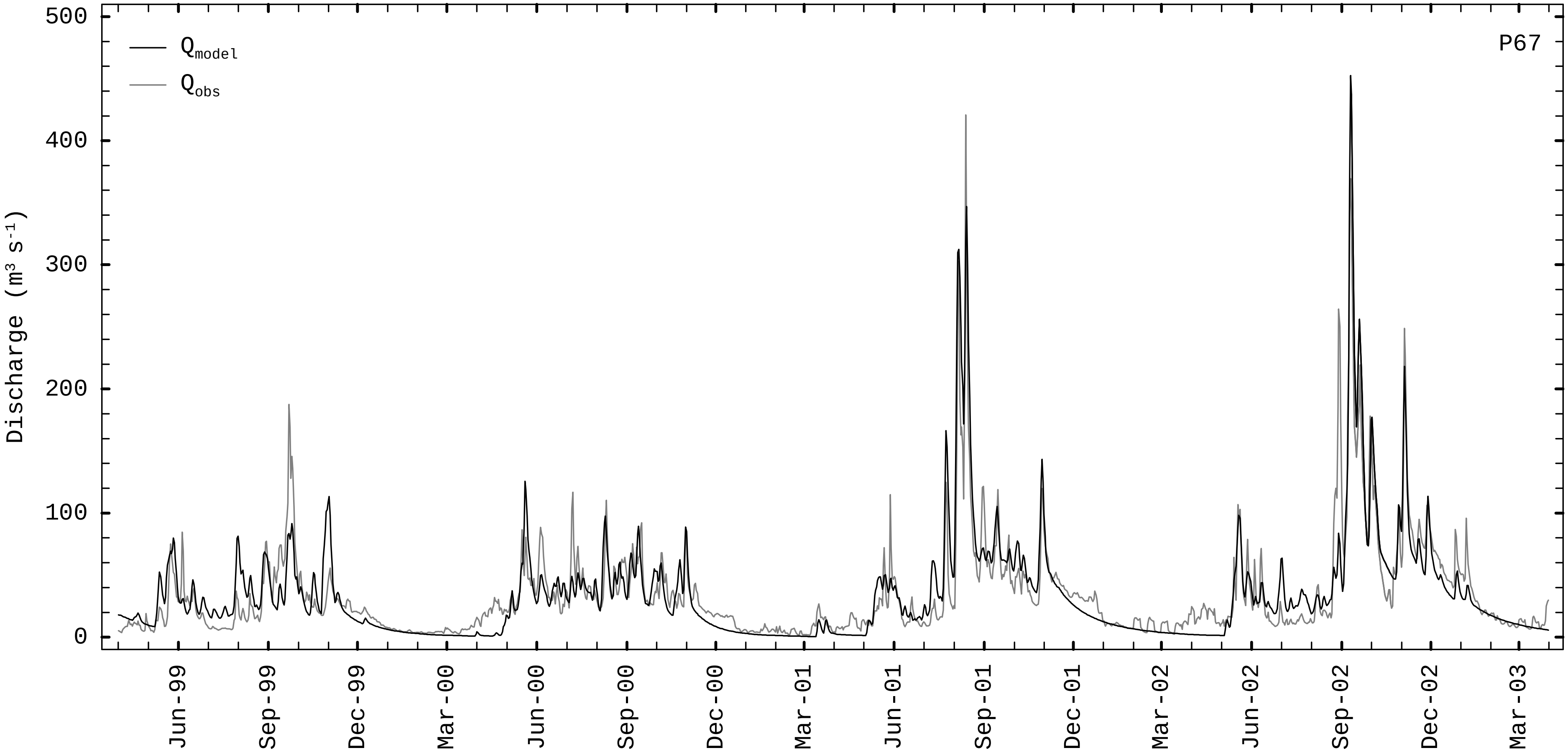
I1: Station P75



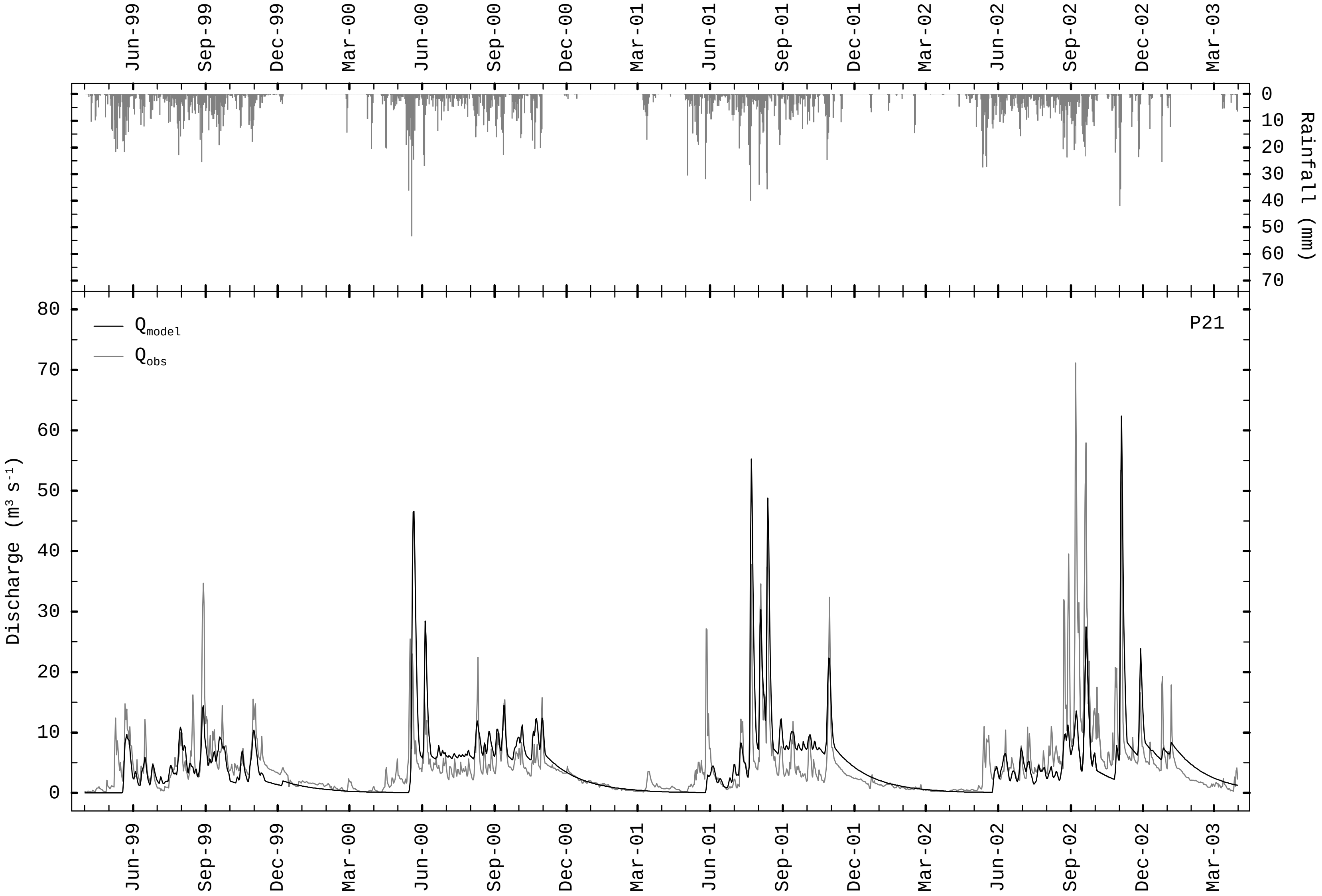
I2: Station P4A



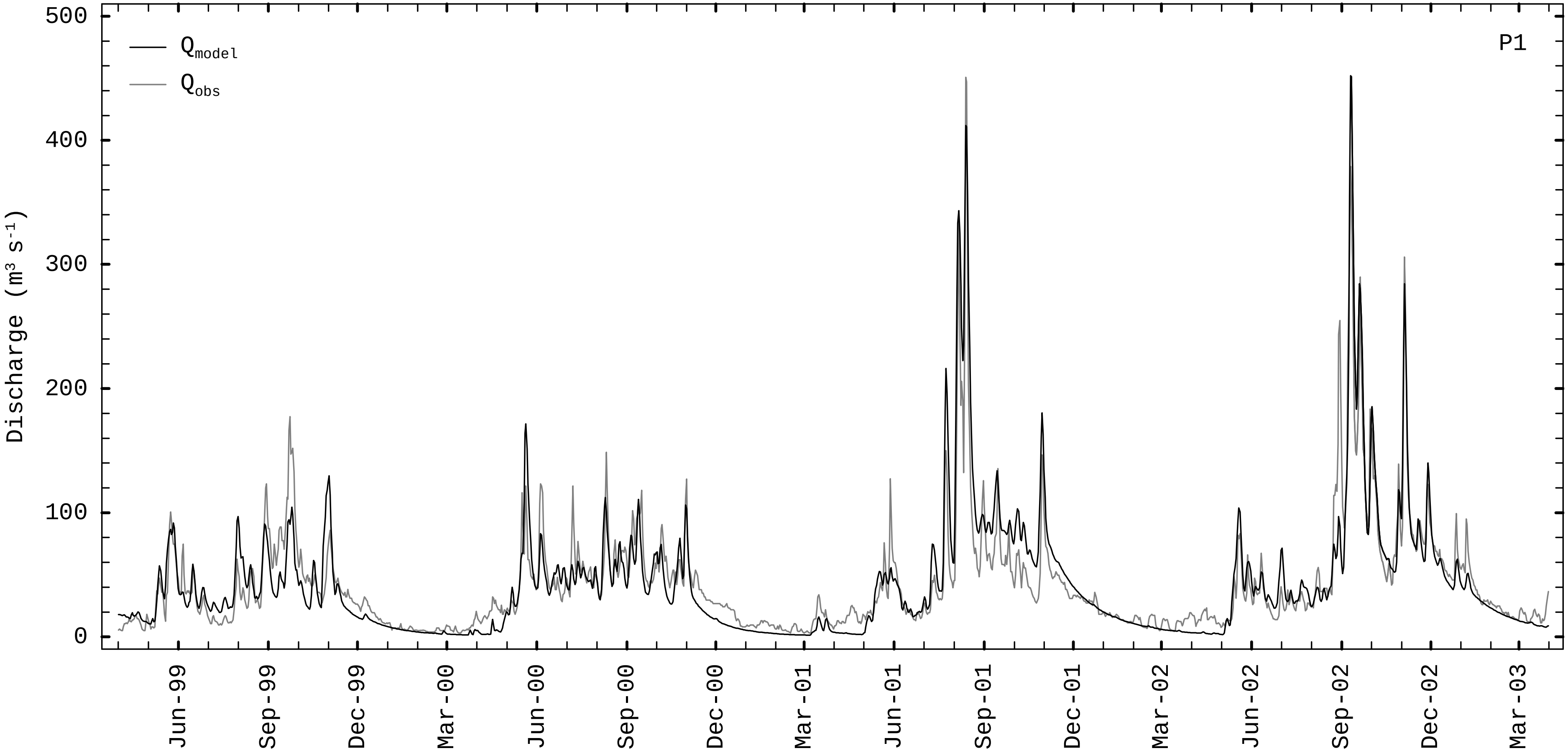
I3: Station P67



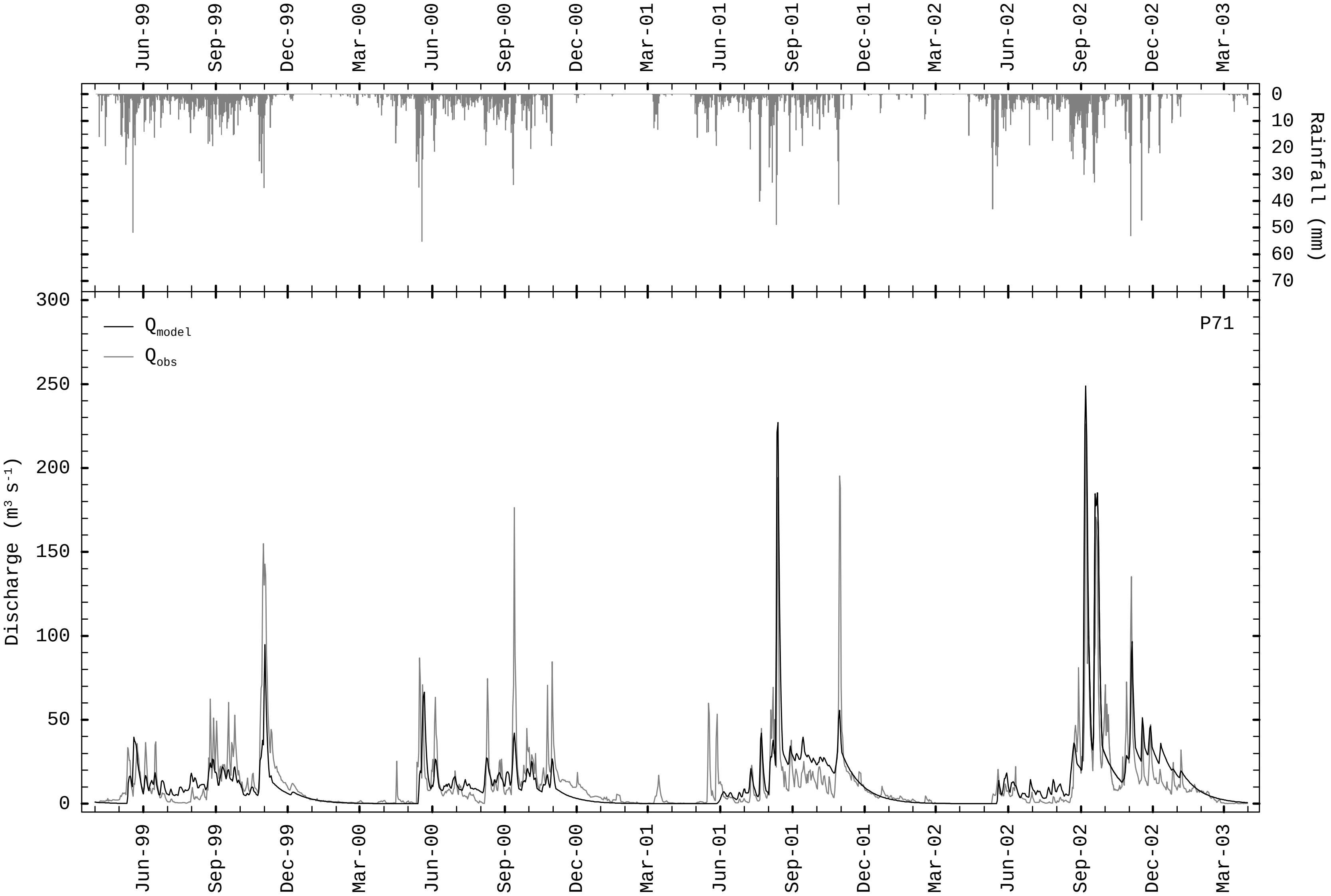
I4: Station P21



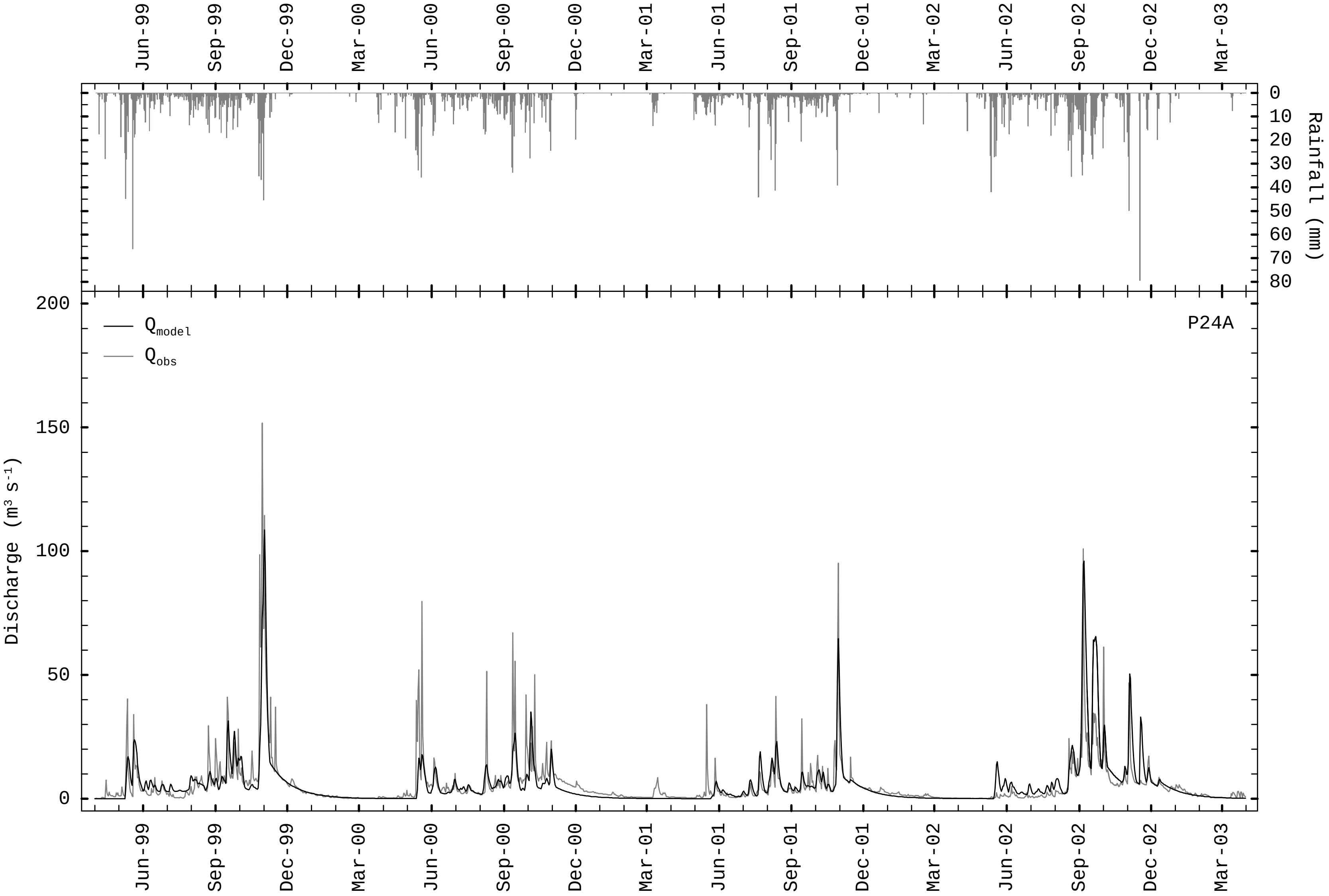
I5: Station P1



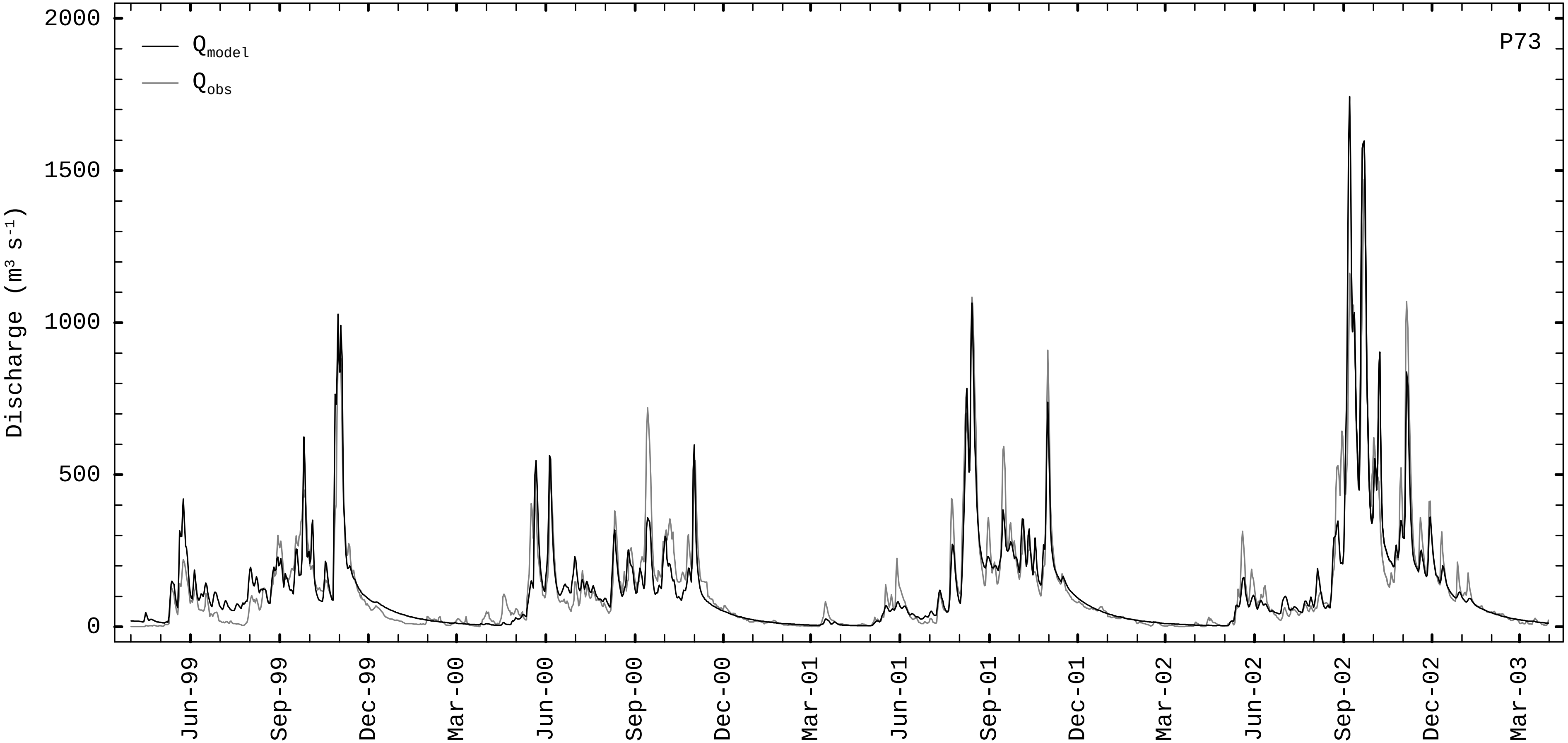
I6: Station P71



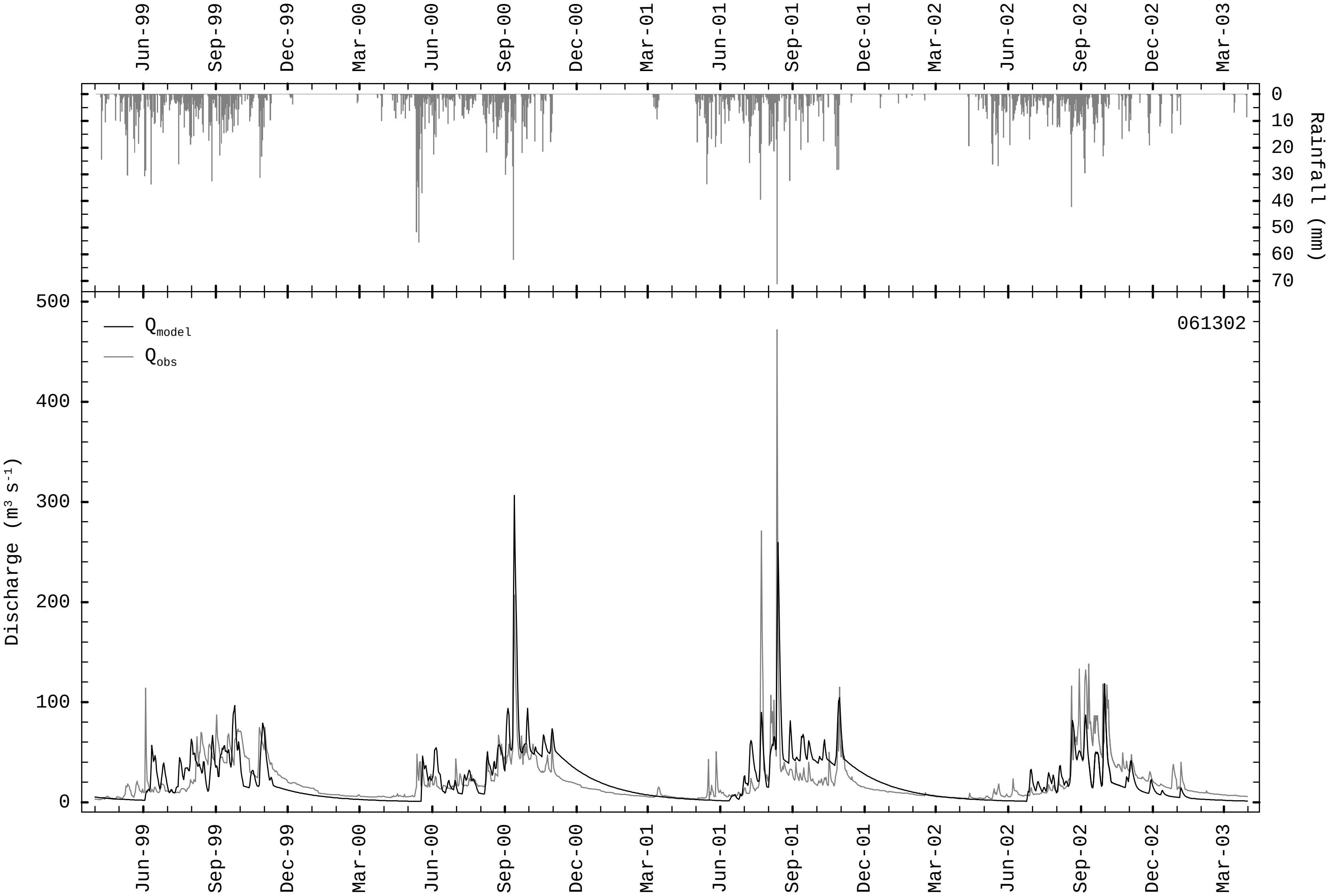
I7: Station P24A



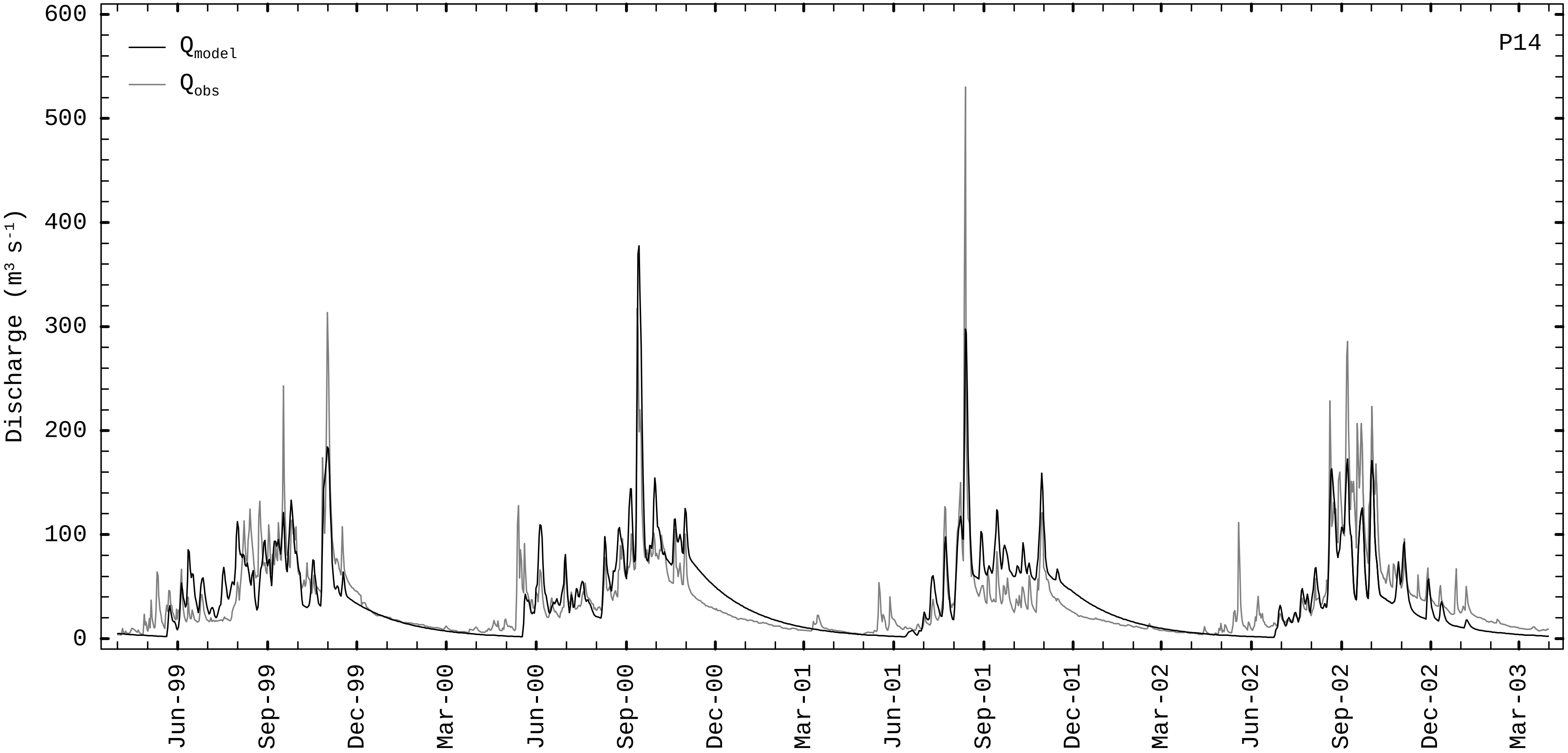
I8: Station P73



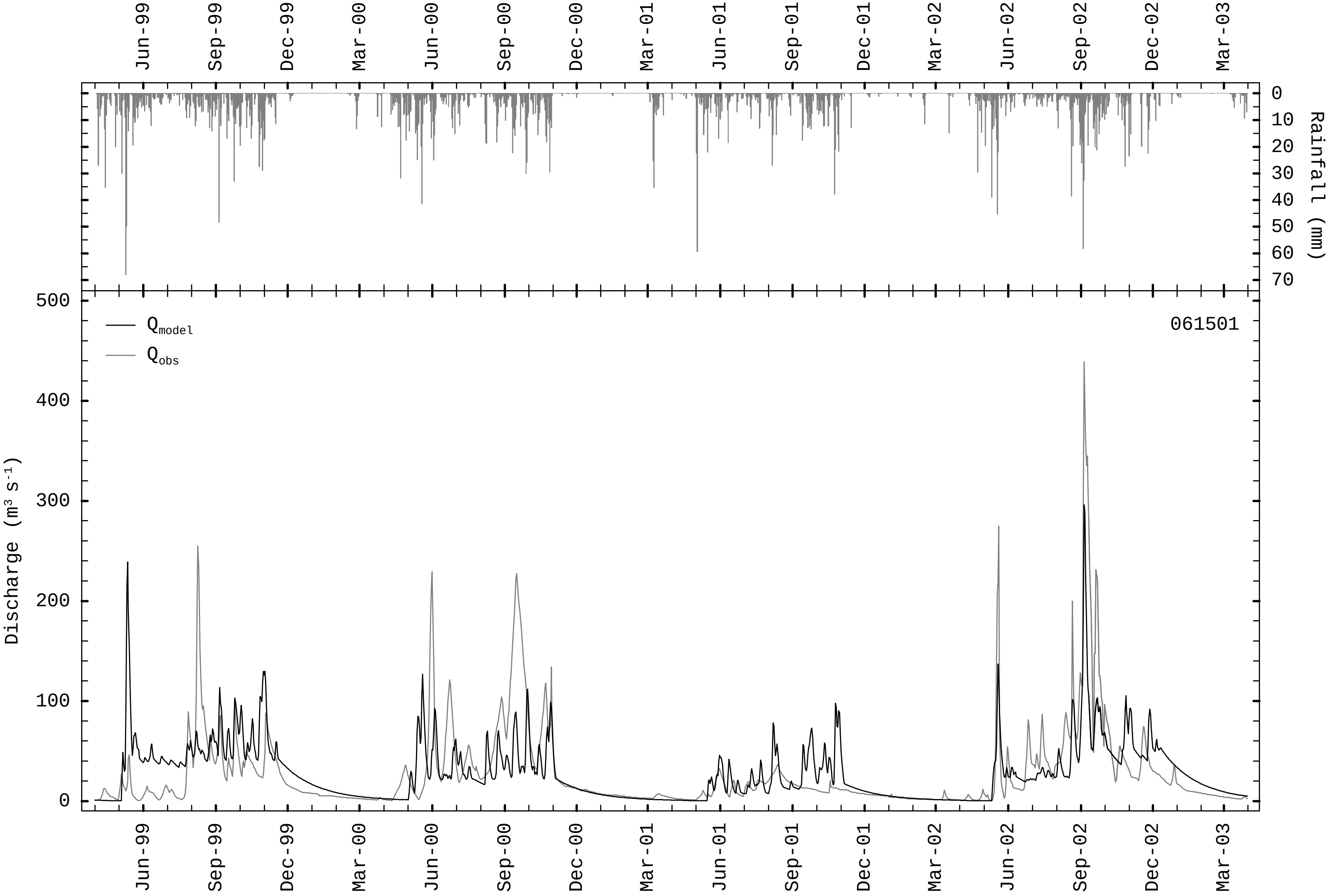
I9: Station 061302



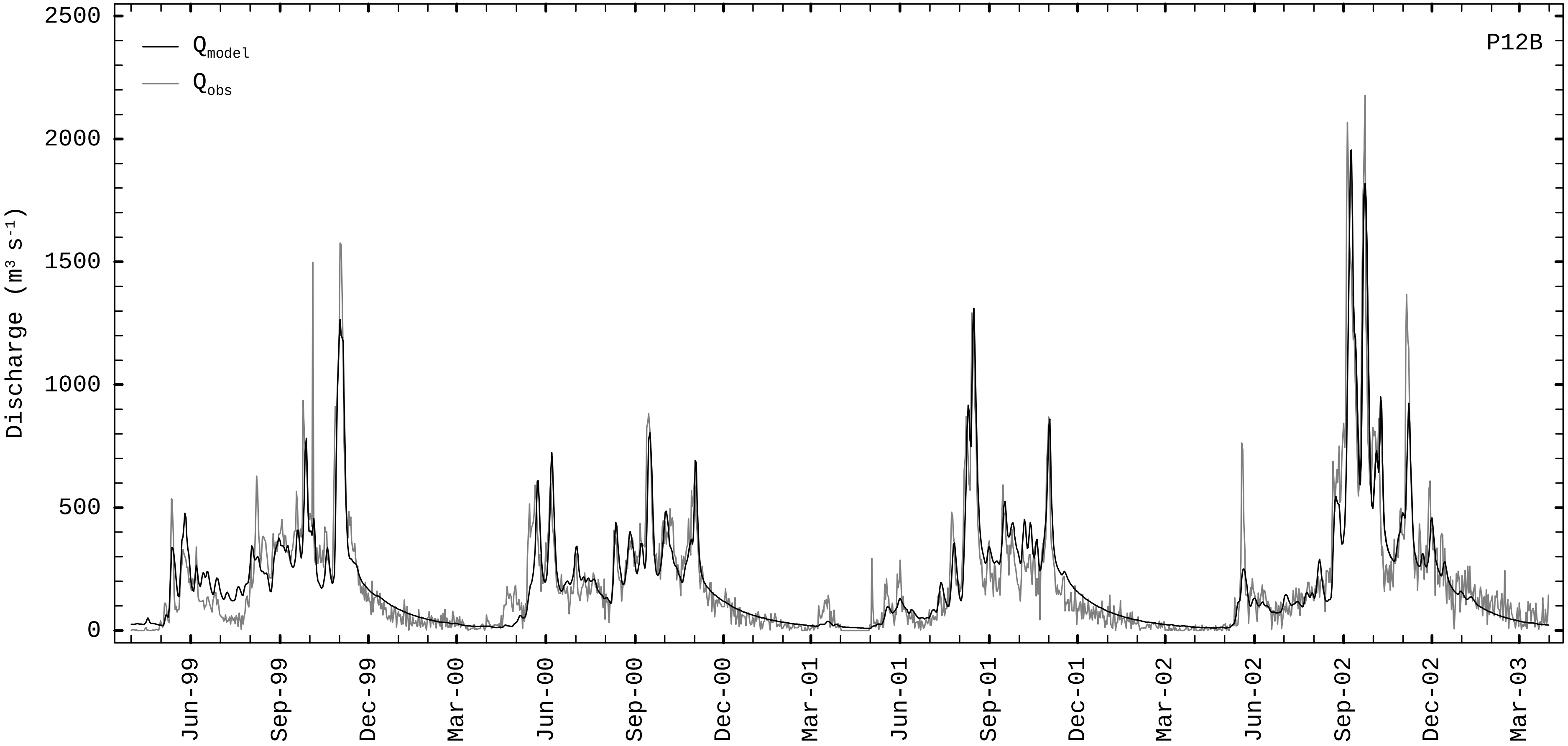
I10: Station P14



I11: Station 061501



I12: Station 12B



Appendix J: Effects of Future Climate on Anomalous Events

J1: Characteristics of the anomalous low flow in wet season from the 2011-2100 simulated streamflow under A2 at the 12 gauging stations

Station	Longest spell of anomalous low flow (day)	Start date	End date	Total volume of shortage (MCM)	Intensity of shortage (MCM d^{-1})
P75	62	30-Sep-2029	30-Nov-2029	51.06	0.824
		30-Sep-2096	30-Nov-2096	49.23	0.794
P4A	71	21-Sep-2014	30-Nov-2014	45.85	0.646
P67	58	4-Oct-2029	30-Nov-2029	75.40	1.300
		4-Oct-2096	30-Nov-2096	70.07	1.208
P21	59	3-Oct-2014	30-Nov-2014	12.29	0.208
		3-Oct-2015	30-Nov-2015	12.42	0.210
		3-Oct-2016	30-Nov-2016	11.42	0.194
		3-Oct-2021	30-Nov-2021	12.36	0.209
		3-Oct-2022	30-Nov-2022	11.09	0.188
		3-Oct-2029	30-Nov-2029	12.17	0.206
		3-Oct-2037	30-Nov-2037	11.63	0.197
		3-Oct-2039	30-Nov-2039	11.36	0.193
		3-Oct-2070	30-Nov-2070	11.09	0.188
		3-Oct-2071	30-Nov-2071	11.96	0.203
		3-Oct-2074	30-Nov-2074	13.03	0.221
		3-Oct-2091	30-Nov-2091	12.21	0.207
		3-Oct-2099	30-Nov-2099	10.21	0.173
P1	59	3-Oct-2029	30-Nov-2029	133.59	2.264
		3-Oct-2096	30-Nov-2096	127.15	2.155
P71	34	28-Oct-2026	30-Nov-2026	13.97	0.411
		28-Oct-2088	30-Nov-2088	14.80	0.435
P24A	55	7-Oct-2054	30-Nov-2054	10.13	0.184
P73	46	16-Oct-2071	30-Nov-2071	190.70	4.146
061302	59	3-Oct-2071	30-Nov-2071	60.01	1.017
		3-Oct-2083	30-Nov-2083	54.64	0.926
		3-Oct-2095	30-Nov-2095	54.70	0.927
P14	56	6-Oct-2046	30-Nov-2046	72.86	1.301
		6-Oct-2082	30-Nov-2082	69.50	1.241
		6-Oct-2098	30-Nov-2098	76.20	1.361
061501	28	3-Nov-2011	30-Nov-2011	22.97	0.820
		3-Nov-2023	30-Nov-2023	23.01	0.822
		3-Nov-2031	30-Nov-2031	22.11	0.790
		3-Nov-2035	30-Nov-2035	21.66	0.774
		3-Nov-2037	30-Nov-2037	21.37	0.763
		3-Nov-2039	30-Nov-2039	23.16	0.827
		3-Nov-2042	30-Nov-2042	23.12	0.826
		3-Nov-2051	30-Nov-2051	20.71	0.740
061501	28	3-Nov-2052	30-Nov-2052	22.53	0.805
		3-Nov-2054	30-Nov-2054	22.19	0.792
		3-Nov-2055	30-Nov-2055	22.54	0.805
		3-Nov-2056	30-Nov-2056	21.85	0.780
		3-Nov-2057	30-Nov-2057	21.52	0.769
		3-Nov-2058	30-Nov-2058	21.56	0.770

Station	Longest spell of anomalous low flow (day)	Start date	End date	Total volume of shortage (MCM)	Intensity of shortage (MCM d^{-1})
061501	28	3-Nov-2060	30-Nov-2060	21.21	0.757
		3-Nov-2064	30-Nov-2064	22.82	0.815
		3-Nov-2069	30-Nov-2069	19.01	0.679
		3-Nov-2075	30-Nov-2075	21.38	0.764
		3-Nov-2078	30-Nov-2078	22.86	0.816
		3-Nov-2080	30-Nov-2080	23.14	0.826
		3-Nov-2087	30-Nov-2087	21.08	0.753
		3-Nov-2088	30-Nov-2088	20.14	0.719
		3-Nov-2090	30-Nov-2090	21.52	0.769
		3-Nov-2095	30-Nov-2095	21.73	0.776
		3-Nov-2096	30-Nov-2096	19.52	0.697
		3-Nov-2097	30-Nov-2097	21.44	0.766
P12B	55	7-Oct-2096	30-Nov-2096	424.79	7.723

MCM: million m^3.

J2: Characteristics of the anomalous low flow in wet season from the 2011-2100 simulated streamflow under B2 at the 12 gauging stations

Station	Longest spell of anomalous low flow (day)	Start date	End date	Total volume of shortage (MCM)	Intensity of shortage (MCM d^{-1})
P75	60	2-Oct-2086	30-Nov-2086	53.56	0.893
P4A	61	1-Oct-2013	30-Nov-2013	44.06	0.722
		1-Oct-2015	30-Nov-2015	40.81	0.669
		1-Oct-2019	30-Nov-2019	41.62	0.682
		1-Oct-2020	30-Nov-2020	41.51	0.681
		1-Oct-2023	30-Nov-2023	44.92	0.736
		1-Oct-2032	30-Nov-2032	44.65	0.732
		1-Oct-2069	30-Nov-2069	47.94	0.786
P67	55	7-Oct-2086	30-Nov-2086	76.65	1.394
P21	59	3-Oct-2011	30-Nov-2011	10.66	0.181
		3-Oct-2015	30-Nov-2015	10.79	0.183
		3-Oct-2069	30-Nov-2069	13.09	0.222
P1	57	5-Oct-2086	30-Nov-2086	132.55	2.325
P71	34	28-Oct-2013	30-Nov-2013	15.40	0.453
		28-Oct-2051	30-Nov-2051	15.54	0.457
		28-Oct-2088	30-Nov-2088	14.57	0.428
P24A	32	1-Aug-2047	1-Sep-2047	1.29	0.040
		1-Aug-2080	1-Sep-2080	1.23	0.039
		1-Aug-2090	1-Sep-2090	0.95	0.030
P73	42	20-Oct-2029	30-Nov-2029	168.75	4.018
061302	59	3-Oct-2023	30-Nov-2023	60.49	1.025
		3-Oct-2028	30-Nov-2028	59.54	1.009
		3-Oct-2039	30-Nov-2039	53.90	0.914
		3-Oct-2057	30-Nov-2057	65.40	1.109
		3-Oct-2059	30-Nov-2059	63.31	1.073
		3-Oct-2071	30-Nov-2071	64.48	1.093
		3-Oct-2073	30-Nov-2073	65.56	1.111
		3-Oct-2077	30-Nov-2077	63.70	1.080
		3-Oct-2078	30-Nov-2078	68.06	1.154
		3-Oct-2094	30-Nov-2094	61.71	1.046
		3-Oct-2095	30-Nov-2095	62.25	1.055
		3-Oct-2097	30-Nov-2097	62.87	1.066
P14	57	5-Oct-2023	30-Nov-2023	78.62	1.379
		5-Oct-2028	30-Nov-2028	74.02	1.299
		5-Oct-2043	30-Nov-2043	85.34	1.497
		5-Oct-2063	30-Nov-2063	76.41	1.340
		5-Oct-2071	30-Nov-2071	82.46	1.447
		5-Oct-2073	30-Nov-2073	84.59	1.484
		5-Oct-2077	30-Nov-2077	77.39	1.358
		5-Oct-2078	30-Nov-2078	87.30	1.532
		5-Oct-2082	30-Nov-2082	81.74	1.434
		5-Oct-2085	30-Nov-2085	80.74	1.416
		5-Oct-2095	30-Nov-2095	75.72	1.328
061501	29	2-Nov-2060	30-Nov-2060	16.53	0.570
P12B	54	8-Oct-2023	30-Nov-2023	402.08	7.446

MCM: million m^3.