Doctoral Thesis

# Teleconnections of El Niño–Southern Oscillation (ENSO) and Indian Ocean Dipole (IOD) to Streamflow in Java Island, Indonesia

エルニーニョ南方振動およびインド洋ダイポール
モード現象とインドネシアジャワ島での
河川流量とのテレコネクション

**Adam Rus Nugroho (1183915002)**

**Advisor: Ichiro TAMAGAWA**

**Basin Hydrology Dynamics Laboratory**

**Graduate School of Engineering**

**Gifu University**

**2022 June**

# Acknowledgements

The doctoral program has become quite a long journey for me, especially with a one-and-a-quarter-year extension from the normal three-year study. The extension was mainly caused by a lengthy process for finding the research topic and the rejection of the first paper at the beginning of the third year. After the rejection, the study was narrowed to a climatic study. Finally, in the fourth year, two research papers could be published successfully, constructing this thesis titled *Teleconnections of ENSO and IOD to Streamflow in Java Island, Indonesia*.

This doctoral thesis could not be done without the contribution of many people. Before acknowledging those people, I would like first to praise Allah, our almighty God in Islam. Without His blessing, I am sure I could not survive my doctoral study.

Firstly, I would like to sincerely thank my advisor, Prof. Ichiro TAMAGAWA sensei of the Basin Hydrology Dynamics (BHD) Laboratory, for his continuous support, guidance, motivation, and patience during my study. I also would like to sincerely thank the associate professor of the BHD Laboratory, Morihiro HARADA sensei, for his guidance, insightful comments, and interesting discussions. Besides those two teachers of mine, my sincere gratitude also goes to Prof. Tomonao KOBAYASHI sensei for being on a thesis committee along with Harada sensei and giving a constructive review of my thesis.

Secondly, I would like to express my gratitude to the Advanced Global Program (AGP) of the Graduate School of Engineering Gifu University for the scholarship and to Dr. Joni Aldilla Fajri as the mediator from my institution, Universitas Islam Indonesia (UII). The gratitude also goes to the UII's rectorate and all related staff for the support, especially for the financial aid.

Thirdly, I would like to thank all my friends with whom I interacted and got along during my stay in Gifu from 2018 to 2022: 1) the Indonesian community, 2) the Islamic community of Gifu Mosque, 3) fellow international students, and 4) all

Japanese friends (especially Nakamura-san / Katsu-san). I hope we can still connect to each other in the future, even after being separated by distance.

Last but not least, I would like to express huge gratitude to 1) my mother, father, and sisters for their support in various forms; 2) my dear wife, Baiq Raudatul Jannah, for her willingness to marry me during my second-year study, following me to Gifu, and supporting me all this time; and 3) my son who was born in Gifu, Cidris Khoirus Sora, for all happiness you had given to me all this time. Special for you, my son: do you know that your last name came from the theme of this dissertation?

# Abstract

Predicting streamflow regimes by considering climate dynamics is important in the future of water resource management. Many studies have already studied the relationship between climate and rainfall in Indonesia, but those targeting the Indonesian climate and streamflow relationship are still rare. Only a single previous study was found on this topic for Indonesian streamflow. The previous study stated that developing a statistical prediction model for the climate–streamflow relationship is difficult due to its non-linear nature. Addressing this gap, the studies in this thesis are devoted to developing a statistical model for predicting streamflow by climate indices as predictors. El Niño–Southern Oscillation (ENSO) and Indian Ocean Dipole (IOD) are the global climate phenomena used as predictors. The multiple regression (MR) models are developed in two studies using polynomial functions to accommodate the non-linearity of the relationship. In the first study, using data from a river located in south-central Java, namely Code River, good model skills were found, especially when using 3- and 6-month moving average data sets. The moving average sets could increase the accuracy of explaining the variance of ENSO and IOD because they can avoid the effect of Madden Julian Oscillation (MJO) and other short-term phenomena. It was also found that the IOD had a modulation effect on ENSO–streamflow relationship, increasing the high flow in the La Niña phases. The MR models developed from six-month leading climate indices were found to explain the flow regimes' variability better than those with same-time climate indices. The time-lagged MR models have relatively good skills with high values of correlation coefficient (mostly higher than 0.7) and moderate Root Mean Square Error (44–51%). It suggests that the forecasting of flow regime may be possible at Code River. The second study extended the first by investigating the spatial and temporal pattern of the climate–streamflow relationship in eight Javanese rivers. The seasonal arrangement of MR models using data from 1970 to 2018 showed that September–November (SON) season has the best model skill. The evaluation of MR models from the eight rivers indicated that

the eastern rivers seem more predictable than the western rivers. However, the result of the next analysis changed this temporary conclusion. A further temporal investigation found that the predictability in 1970–1989 tended to be higher than in 1999–2018, suggesting that the relationship between the climate modes and streamflow in Java changed over time. This tendency can be seen in all rivers, even the western rivers. The model skill of western rivers in the earlier period was very high, then became much lower in the later period. Therefore, it can be concluded that no clear spatial distribution patterns of the predictability were found, suggesting similar effects of ENSO and IOD occurred in all parts of Java Island. The physical changes of a river basin during the urban development might be the reason for this relationship change. Ultimately, this thesis successfully confirmed and developed the climate–streamflow relationship using quantitative statistical models while also giving new insight into the Indonesian teleconnection models by elucidating their behaviors.

# Table of Contents

# List of Tables

# List of Figures

# Chapter 1
# Introduction

## 1.1. Background

Understanding the pattern of streamflow regimes is important in ecosystem conservation and disaster anticipation. Keeping up and managing the alteration of the natural flow regime then become critical for the health of the river environment [1]. A recent study suggested that the perspective of flow regime management should be shifted toward the climate and ecological dynamics in the future [2]. Hence, the studies connecting flow regimes to climate phenomena are important for the future.

Integrating the prediction system into the streamflow management gives a significant advantage. The river managers who can predict the streamflow regimes should be able to improve the riverine ecosystem protection and water conservation. Not only that, but an accurate streamflow forecast would also reduce the impact of water-related disasters (i.e., floods and drought). Controlling floods and mitigating drought would become more effective using an accurate river flow prediction. This prediction would be even more efficient by using teleconnection models.

A teleconnection is a causal link or correlation between meteorological or other environmental events that happen apart at a great distance. In atmospheric sciences, the term teleconnection is often used to explain climate connections between geographically separated regions [3]. Many teleconnection studies tried to connect a global climate phenomenon to another global or a local climate phenomenon, such as rainfall. Not only to rainfall, recently, the term teleconnection has also been used to link a climate phenomenon to streamflows [4–7].

One of the global climate phenomena widely used in teleconnection studies is El Niño–Southern Oscillation (ENSO), a periodical variation in winds and sea surface temperatures over the tropical Pacific Ocean. ENSO consists of El Niño as the warming phase and La Niña as the cooling phase [8]. Indonesian rainfall is

known to be affected by the ENSO. El Niño correlates to less rainfall and drought in Indonesia, while La Niña correlates to rainfall increase and floods in Indonesia [9,10]. The ENSO events were found to be correlated well to the rainfall in southern Indonesia [11,12]. Indonesian rainfall is also being affected by the Indian Ocean Dipole (IOD), a seasonal oscillation of sea surface temperature in the Indian Ocean [13]. The high activity of the IOD has been recognized to cause drought in Indonesia [14]. A study in northwestern Java showed that IOD events affect the interannual rainfall variation in the dry season [15].

The effect of ENSO and IOD is studied not only on rainfall but also on streamflow. The effect of ENSO on streamflows has been studied in many regions of the earth [16–22]. During this study, only one study related to the ENSO and IOD teleconnection to the streamflow in Indonesia was found. Using streamflow data of the Citarum River in the northwestern Java, that study found the association between: 1) extremely high flow events with La Niña and negative IOD, and 2) the extremely low flow events with El Niño and positive IOD [23]. The study also argued that developing a statistical prediction model of the climate–streamflow relationship is difficult due to the non-linear relationship caused by climate phenomena.

Indeed, finding good streamflow predictability by statistical models is difficult due to various climate phenomena affecting hydrological processes. Maybe because of this, very limited studies are found related to statistical models for predicting streamflow in Indonesia. The authors of the study mentioned in the previous paragraph only discussed using only correlation and probability analyses. A correlation study can only show the relationship tendency and cannot provide the output of streamflow amount. However, the model output in streamflow values seems practically better. On the other side, a stochastic (statistical) method is likely preferable, recalling that many Indonesian river basins are difficult to model physically due to the current data scarcity. Therefore, this thesis attempted to develop a statistical model to elucidate the predictability of streamflows by the climate indices in an Indonesian region.

The main part of this thesis is two studies in which both applied multiple regression models for streamflow regime indices using ENSO and IOD indices as

predictors. As a feasibility study, the first study focused on a river located in southern central Java, namely the Code River. The model developed using Code River data interestingly showed good model skills. The second study extended the exploration of the model predictability by investigating their spatial and temporal patterns using data from 8 rivers on Java Island.

## 1.2. Objectives

The studies in this thesis aimed to elucidate the teleconnection of ENSO and IOD to the streamflow regime in the rivers of Java, Indonesia, using multiple polynomial regression analysis. The objectives of the studies are:

a) to understand the connection of ENSO and IOD indices to the Javanese rivers' flow regime indices
b) to develop a suitable regression model for predicting the streamflow regimes using ENSO and IOD indices
c) to elucidate the spatial variation predictability of ENSO and IOD teleconnection to streamflow in Javanese rivers
d) to elucidate the temporal variation predictability of ENSO and IOD teleconnection to streamflow regimes in Javanese rivers

## 1.3. Outline

The research workflow of this thesis' studies—from background until conclusion—is shown in Figure 1.1. Meanwhile, the writing of this book is arranged in the following order:

Chapter 1. Introduction
Chapter 2. Literature Review of ENSO and IOD
Chapter 3. Modelling Methods
Chapter 4. Regression of ENSO and IOD to Code River's Flow Regime
Chapter 5. Spatial and Temporal Pattern of the Javanese Streamflow Predictability by ENSO and IOD
Chapter 6. Conclusion

Chapter 1 introduces the background and objectives of the thesis studies. Chapter 2 provides summarized literature reviews on the two climate indices used as predictors in this thesis' studies. It includes the general knowledge of the ENSO

and IOD, along with an introduction to a few previous studies related to their influence on rainfall and streamflow. Chapter 3 provides the methods of the modelling along with the model evaluation. Chapter 4 explains the regression models prediction skill of the climate indices to the streamflow of Code River. Chapter 5 investigates temporal and spatial patterns of the streamflow predictability by the ENSO and IOD indices in Java. The last chapter, Chapter 6, concludes the finding of the studies by summarizing the results of the previous two chapters, mentioning the significance of the thesis result, and suggesting some related future works. The List of Publications can be found after Chapter 6, showing the scientific publications which constructed this thesis. Chapter 4 corresponds to the first paper, and Chapter 5 corresponds to the second paper listed in the List of Publications.

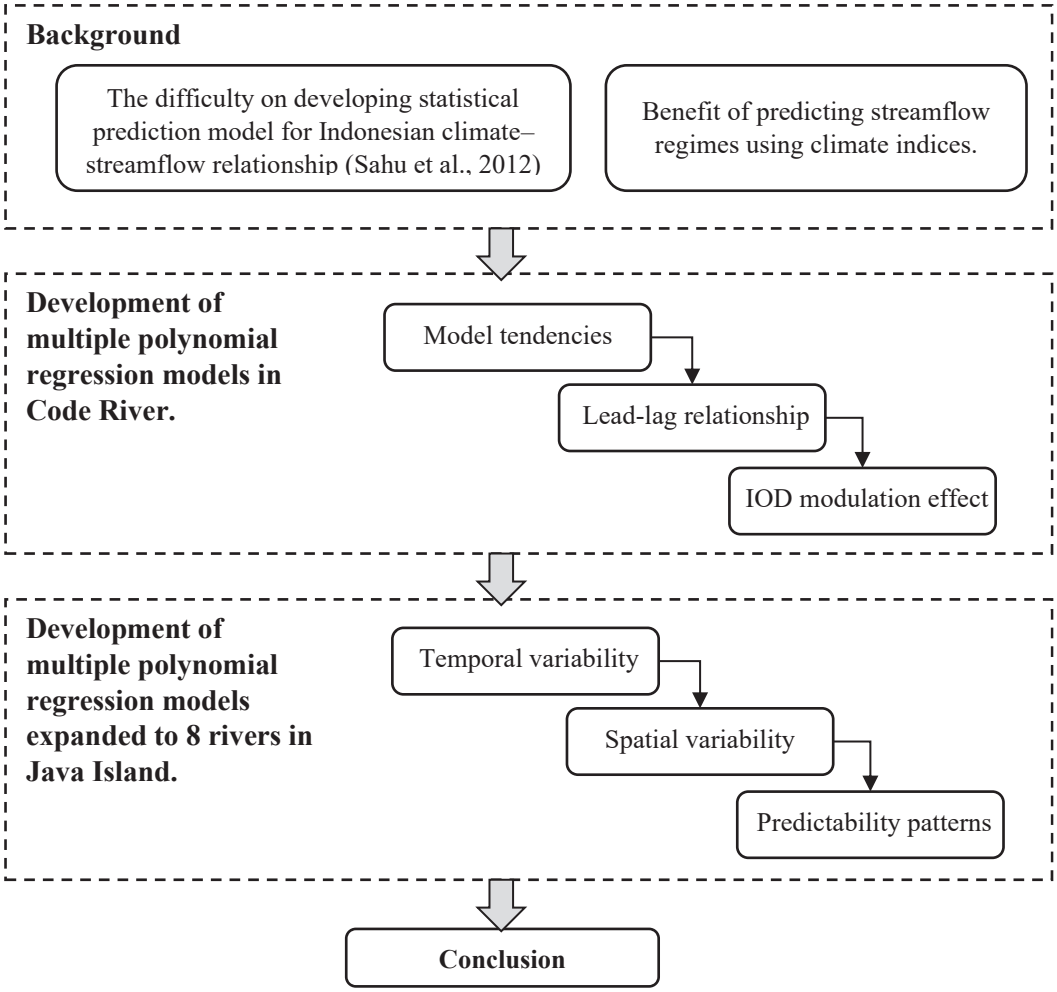


Figure 1.1 Research flow

# Chapter 2
# Literature Review of ENSO and IOD

## 2.1. El Nino Southern Oscillation (ENSO)

### 2.1.1. Overview of ENSO

El Niño Southern Oscillation (ENSO) is a periodical variation in sea surface temperatures over the Pacific Ocean caused by interactions between the ocean and the atmosphere. Originally, El Niño was observed as unusually warm ocean temperatures that occur every 2–7 years around the end of the year along the Peruvian coast, equatorial eastern, and the central Pacific Ocean [24]. Then, there is Southern Oscillation which refers to a 'seesaw' of the atmospheric pressure between the Pacific and Indian Oceans. Then, in the 1960s, scientists realized that these two phenomena were related. The acronym ENSO which unites the two phenomena then has been widely used to describe this interannual climate variation until now [25].

ENSO can be divided into three phases: normal, El Niño, and La Niña. In the normal phase, equatorial winds warm as they flow westward across the Pacific. Cold water is pulled up along the west coast of South America, and warm water is pushed toward the western Pacific, making a warm sea surface in the west. Hot air rises in the western Pacific, travel eastward, and cool air descends on South America. In El Niño state, the sea surface is warm in the central and eastern Pacific, making less cold water being pulled up along the west coast of South America. Hot air rises in the central Pacific, travel east and west, then cools and descends. In La Niña state, warm water accumulates in the far western Pacific, making equatorial water cooler than in the normal state. Figure 2.1 explains the physical process of the three ENSO phases.

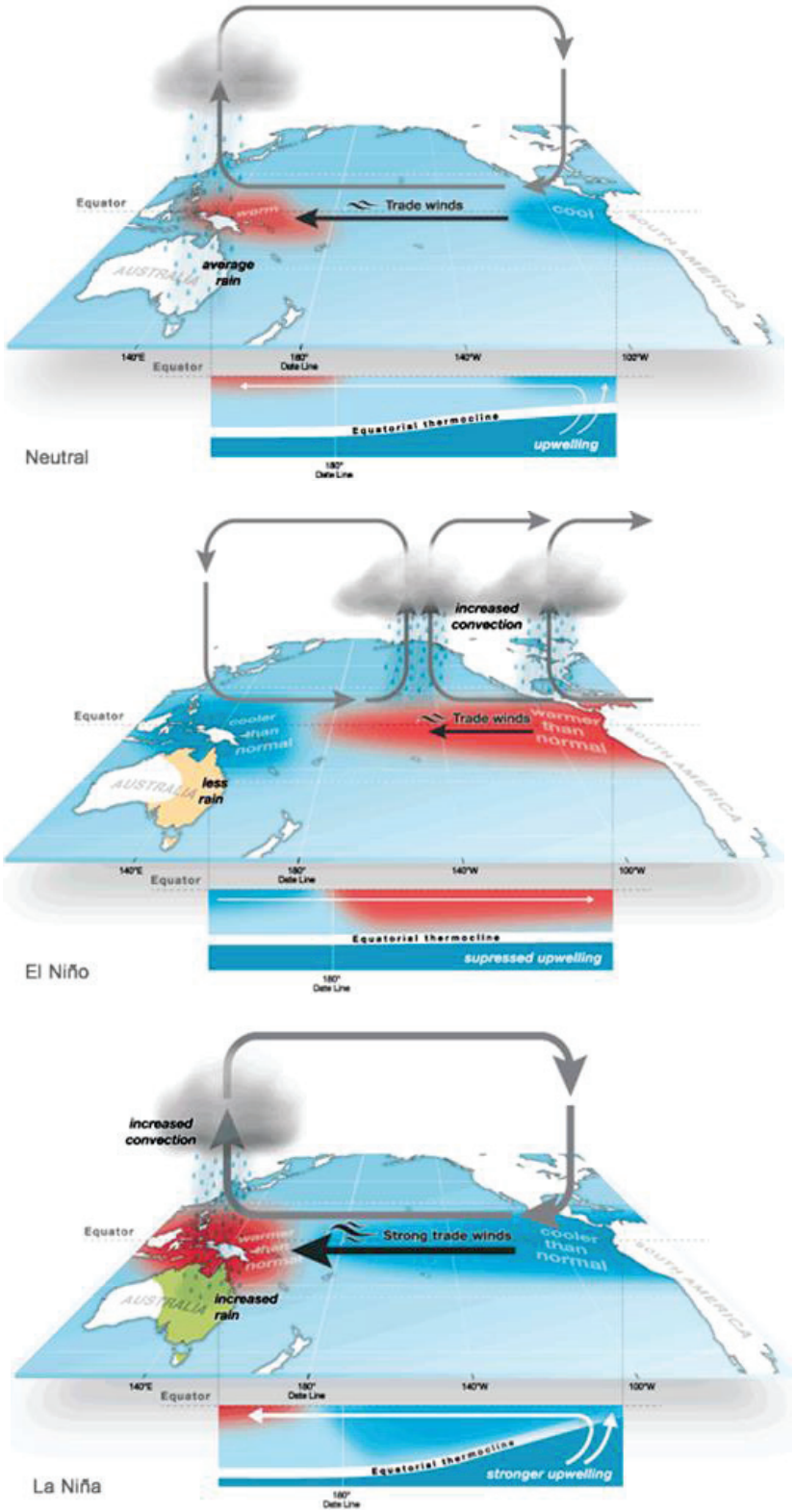


Figure 2.1 Illustration of physical processes of ENSO in three phases
(http://www.bom.gov.au/climate/enso/history/ln-2010-12/three-phases-of-ENSO.shtml)

### 2.1.2. ENSO indices

An index, a scale of a single number derived from some factors describing a phenomenon, has been used to understand the status of ENSO. The ENSO index is observed over time to see its trend [26]. The two main indices of ENSO are derived from sea level pressure and sea surface temperature.The oldest indicator of ENSO is the Southern Oscillation Index (SOI). The measure of SOI is based on the difference between the atmospheric sea level pressures (SLP) at Darwin and Tahiti [27] (Figure 2.2). A seesaw in pressure at those locations reflects the atmospheric component of ENSO, which was discovered in the early 1990s. During El Niño, the pressure becomes below average in Tahiti and above average in Darwin, resulting in a negative SOI. During La Niña, it is the opposite, resulting in a positive SOI. Because SOI is based on the SLP at two individual stations, it can be affected by shorter-term (day-to-day or week-to-week) fluctuations unrelated to ENSO. However, averaging the index values over months or seasons helps isolate more sustained deviations [26].

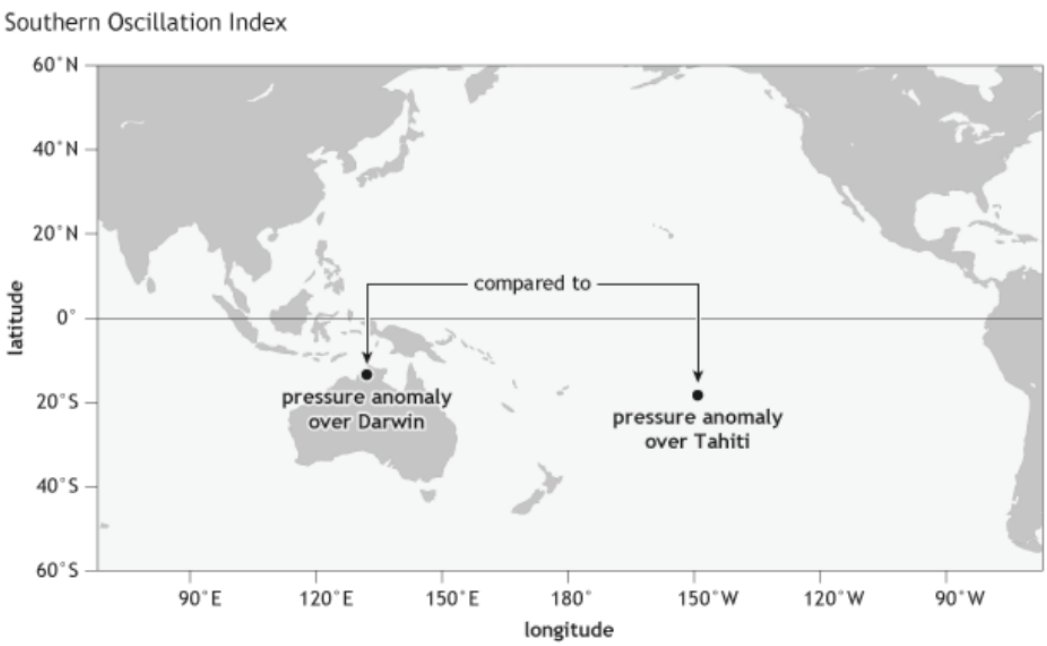


Figure 2.2 Location of the two stations where sea level pressure is observed for SOI (https://www.climate.gov/sites/default/files/styles/full_width_620_alternate_image/public/Fig1_ENSOindices_SOI_610.png) by Fiona Martin, NOAA Climate.gov.

After the finding of SOI, sea surface temperature (SST) data were increasingly used for observing ENSO because the ocean was recognized as the key player in ENSO. SST changes more slowly than SLP, making it easier to identify ENSO, even when averaging over a period as short as a month or less. Initially, certain regions were defined for measurements—namely Niño 1, Niño 2, Niño 3 and Niño 4—because of consistently available data coming from ships passing through those areas [26]. Later, an area overlapping Niño 3 and Niño 4 was identified as a better representative of ENSO and named Niño 3.4. The Niño 3.4 was identified as the most representative index for ENSO since its publication in 1997 [28] and was frequently used in recent studies (e.g., [29–31]). Added with Niño 3.4, there are five regions of Niño. The five Niño regions cover the following coordinates (Figure 2.3):

- Niño 1 : 5–10°S, 80–90°W
- Niño 2 : 0–5°S, 80–90°W
- Niño 3 : 5°N–5°S, 150–90°W
- Niño 3.4 : 5°N–5°S, 120–170°W
- Niño 4 : 5°N–5°S, 160°E–150°W

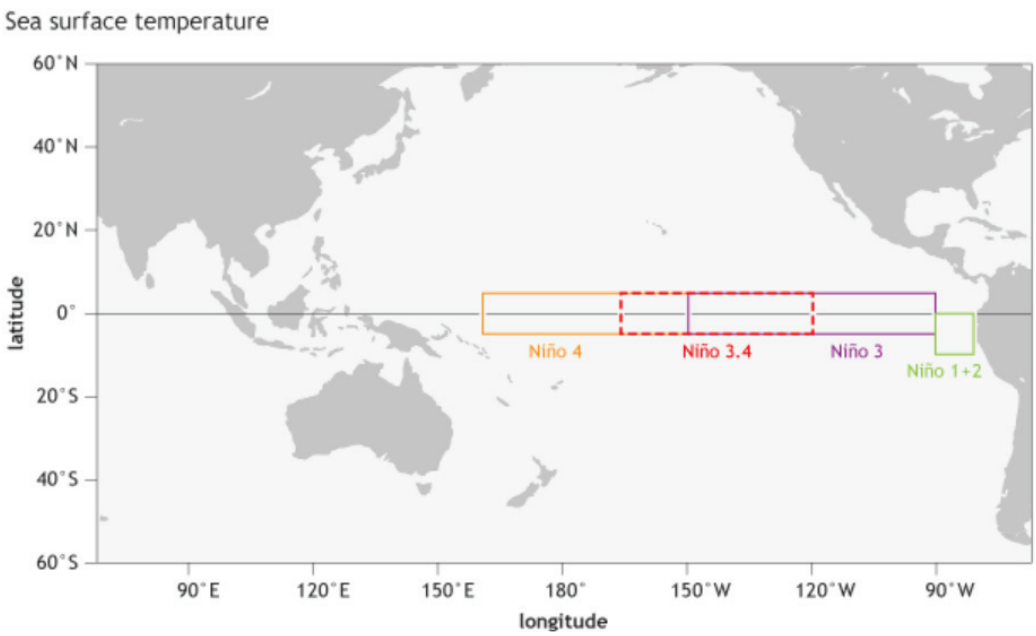


Figure 2.3 Parts of the Pacific Ocean which sea surface temperature monitored for ENSO Indices Niño 1+2, Niño 3, Niño 4, and Niño 3.4 (https://www.climate.gov/sites/default/files/styles/full_width_620_alternate_image/public/Fig3_ENSOindices_ SST_610.png by Fiona Martin, NOAA Climate.gov).

SOI and NINO 3.4 were well-correlated and not statistically different in most seasons [28]. Figure 5.3 in chapter 5 shows the similar tendency of both indices and high correlation ($R$ = 0.71) between those two indices.

### 2.1.3. Impact of ENSO

ENSO phenomenon originates in the tropical Pacific, but its impact does not only affect the regions around the Pacific Ocean. ENSO is considered the most important climate phenomenon on interannual timescales due to its teleconnection, which can potentially disturb weather patterns worldwide [25]. Ropelewski and Halpert (1987) found that the most extensive areas of the ENSO-precipitation relationships are located in the western and central equatorial Pacific, while the consistent variations were found in northern South America, eastern equatorial Africa, and southern India [9]. El Niño impacts in South America are mentioned in a recent paper (2020) as floods in the west (Ecuador, Peru, and Colombia) and drought in the northeast (including the Amazon) [32].

Indeed, the highest relationships between ENSO and precipitation are located around the equator line. However, it was found that ENSO is also linked to the variability of temperature and precipitation in Europe. El Niño was found to affect cold anomaly in northern Europe, warm anomaly in southern Europe, negative precipitation anomaly in Scandinavia, and positive precipitation anomaly in southern Europe [33]. It has been indicated that ENSO mainly affects the precipitation in Europe through the modulation of low-level moisture advection [34]. Also, a recent study (2020) found that the extremely strong El Niño was related to the dryness in western Europe [35].

Over most parts of Indonesia, El Niño is associated with less rainfall and drought, while La Niña is associated with more rainfall and floods [9,10]. Many studies have found a high correlation between ENSO and Indonesian rainfall. Hendon (2003) stated that Indonesian rainfall has a high correlation with SST during the dry season (May–October) and a low correlation with SST during the peak of the wet season (January–March) [10]. Aldrian and Susanto (2003) found that the strongest and most significant influence of ENSO was shown in north-eastern Indonesia (Maluku and northern Sulawesi), followed by southern Indonesia

(from south Sumatra to Timor Island, southern Kalimantan, Sulawesi and a part of Irian Jaya) [11]. Meanwhile, the ENSO-related signal in northwest Indonesia (northwestern Sumatra to northwestern Kalimantan) is weak [11]. Lee (2015) confirmed the positive correlation of La Niña to the rainfall in eastern Indonesia [12]. Hendrawan et al. (2019) found that the ENSO events peaked from November to January, and there are significant positive correlations between SST (Niño 3.4) anomaly and Indonesian rainfall on timescales shorter than 3 years and between 5.5–6.5 years [29].

## 2.2. Indian Ocean Dipole (IOD)

### 2.2.1. Overview of IOD

Indian Ocean Dipole (IOD) is a seasonal oscillation of SSTs in the Indian Ocean. The IOD is based on the SST anomalies on two poles in the tropical Indian Ocean, the western Indian Ocean (WIO) and the eastern Indian Ocean (EIO). The activity level of IOD is measured by the Dipole Mode Index (DMI) [14], which is calculated by the difference between SST anomalies located in WIO (10°S–10°N and 50°E–70°E) and EIO (10°N–0°S, 90°E–110°E). The finding of an unusual relationship between the ocean and atmosphere in the Indian Ocean [14,36,37] introduced the existence of IOD to the public in the year 1999.

In the active phase of IOD, SSTs in EIO, near Sumatra (west Indonesia), are cooler than usual, while the SSTs in the WIO are warmer than usual, making the surface wind bring more chance of rain around the eastern coast of Africa. This active phase of IOD typically peaks in September–November [38]. In the neutral phase of IOD, the SSTs are normal in the EIO and WIO; easterly surface wind brings more moisture to Indonesia and, thus, normal precipitation in Indonesia and Australia. In the negative phase of IOD, the SSTs in both Indian Ocean poles behave just the opposite of those in the active phase, making a reduced chance of rain in eastern Africa and increased rain events around Indonesia and Australia. Figure 2.4 shows the physical process of the three IOD phases. IOD has a tendency of biennial [14] or quasi-biennial periodicity [39].

The IOD is a coupled ocean-atmosphere phenomenon just like the ENSO. Traditionally, many studies believe that IOD is independent of ENSO [14,29,36,39,40], while some researchers argue that the two climate modes have a relationship [41–49]. This link between IOD and ENSO is thought of as the extension of Walker Circulation [42–45] and Indonesian Throughflow [46] to the west. The positive and negative IOD events are often associated with El Niño and La Niña, respectively.

### 2.2.3. Impact of IOD

The impact of IOD can be seen as a significant factor affecting rainfall variability in Indonesia, Australia, and other regions surrounding the Indian Ocean. The change in SST across the Indian Ocean affects moisture and air dynamics over surrounding regions [50]. In Australia, IOD has a significant impact on the rainfall in the western and southern of the continent, with negative partial correlations of IOD-rainfall extending south-eastward from Indonesia to south-east Australia [51]. Saji and Yamagata (2003) found that IOD is well correlated to the variability of temperature and rain in the countries around the Indian Ocean; and it is also strongly correlated to the temperature and rain over Europe, northeast Asia, North America, South America, and South Africa, with the association of positive IOD to the warm anomalies and reduced rainfall [40]. Saji and Yamagata also previously (1999) had stated that the active phase of IOD causes strong rainfall in eastern Africa and droughts in Indonesia [14].

In Indonesia, IOD events have been recognized to influence interannual rainfall variation in the dry season (May–October). Jun-ichi et al. (2012) found that, in northwestern Java, the dry season's drought condition is associated with positive IOD and El Niño events, while the dry season's wet condition is associated more with negative IOD events rather than La Niña events [15]. Lee (2015) confirmed that in the southern Sumatra and southern Java Islands (which area faces the Indian Ocean), the rainfall was positively correlated with the negative IOD; meanwhile, the rainfall in northwestern Sumatra was positively correlated with the positive IOD [12]. Hidayat et al. (2016) found that most of the interannual variation of the rainfall in the northwestern Java (Makassar) is related to IOD (ENSO) event [52].

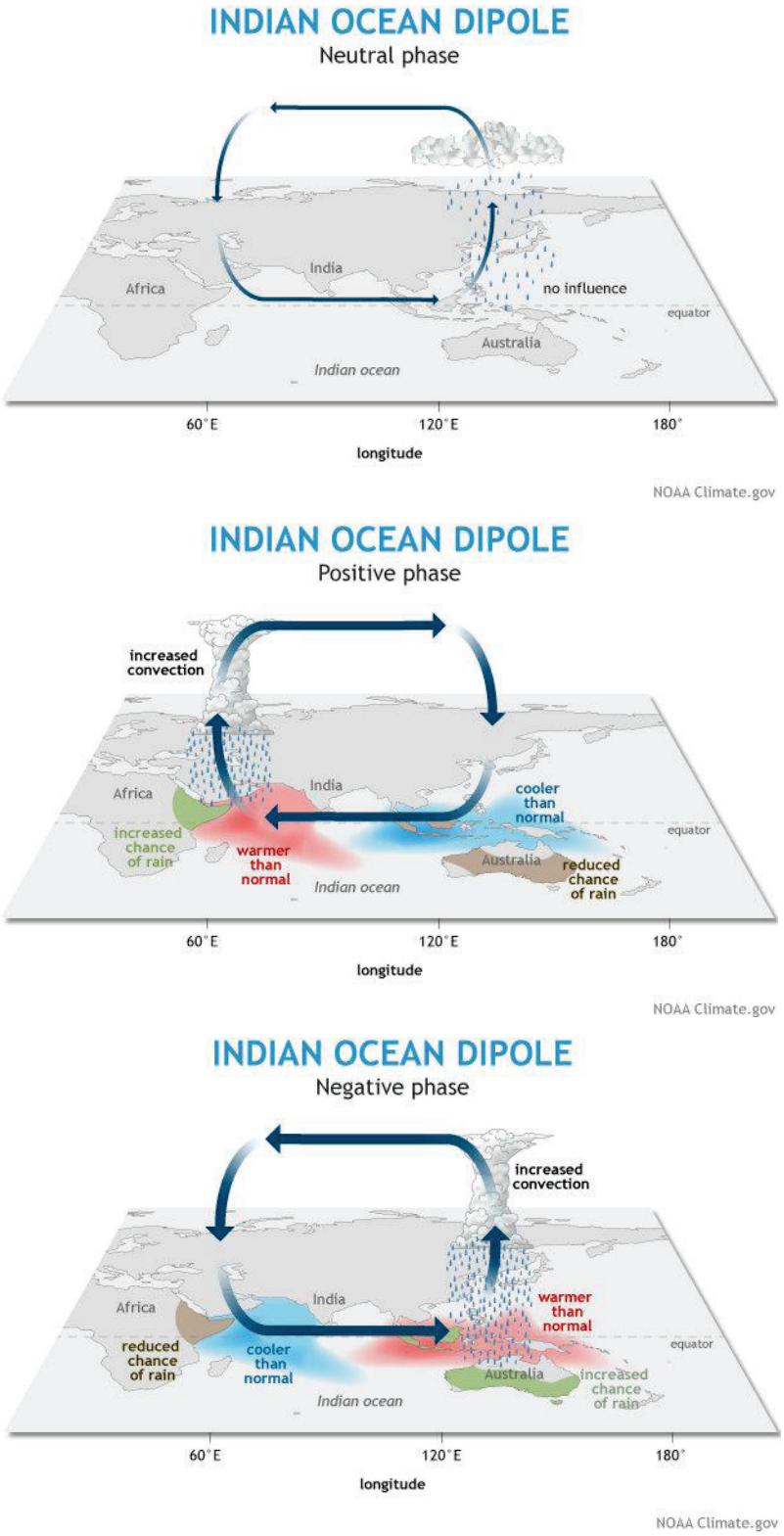


Figure 2.4 Illustration of physical processes of IOD in three phases (https://www.climate.gov/news-features/blogs/enso/meet-enso's-neighbor-indian-ocean-dipole)

Hendrawan et al. (2019) found that a significant positive correlation between IOD and Indonesian rainfall exists on the timescales of 2.5–4 years [29].

In the most recent study, Kurniadi et al. (2021) confirmed three points related to IOD and ENSO impacts on rainfall extremes over Indonesia. Firstly, the impacts of ENSO and IOD on rainfall extremes are the strongest during the dry seasons (June–November) and weaker during the wet seasons (December–May) [53]. Secondly, a positive (negative) IOD is associated with more extreme dry (wet) conditions, while El Niño (La Niña) is associated with drier (wetter) rainfall extremes [53]. Thirdly, the IOD is more related to rainfall extremes in the southern and western regions of Indonesia, while ENSO is more related to rainfall extremes in Indonesia's northern and eastern regions [53].

### 2.3. ENSO and IOD relationship to streamflow

The teleconnection studies of ENSO and IOD were not only for rainfall but also for streamflow. The effect of ENSO on streamflows has been studied in many regions worldwide [16–18,20,21]. Kahya and Dracup (1993) demonstrated the response of the mid-latitude region to the tropical ENSO phenomena by finding a coherent and significant streamflow response to ENSO in four regions of the United States [16]. Cluis and Laberge (2002) detected the El Niño-related signal in most areas of the Asia-Pacific region using 78 stations and 25 years of data sets of large rivers [17]. Cardoso and Silva Dias (2006) developed a satisfying regression model for the relationship between ENSO and the flows of the Paraná River in South America [18]. Chandimala and Zubair (2007) investigated the streamflow predictability of the Kelani River in Sri Lanka by ENSO and Indian Ocean SST data [19]. Zhang et al. (2007) successfully detected the relationships between the annual maximum streamflow of the Yangtze River—the longest river in China—and ENSO by cross-wavelet and wavelet coherence methods [20]. Mohsenipour et al. (2013) revealed the relationship between El Niño events and the discharge of the Kor River in Iran using Fourier analysis [21]. Sahu et al. (2020) linked ENSO, IOD, and two other climate phenomena (El Niño Modoki Index and Indian monsoon) to the river flow of the Mahanadi River in India to understand the impact of climate variability on the basin’s hydrological regime [22].

Only one study explicitly discussed the ENSO and IOD effects on the Indonesian streamflow. Using streamflow data from the Citarum River in West Java, Sahu et al. (2012) found that high streamflow events were associated with La Niña and negative IOD, while the low streamflow events were associated with positive IOD and El Niño [23]. Those relationships were well shown in the boreal fall season (September–November), the same timing as the peak of IOD. The study identified non-linear relationships between the two climate modes and the extreme streamflows in boreal winter and spring seasons. The study also acknowledged the difficulty in developing statistical prediction models due to the complexity of various climate phenomena, such as El Niño Modoki and anomalous Walker Circulation [23].

# Chapter 3
# Modelling Methods

## 3.1. Polynomial Regression

The two studies conducted in this thesis both used multiple polynomial regression techniques to develop models for predicting the indices of average flow ($Q_{50}$), high flow ($Q_{10}$), and low flow ($Q_{90}$). The multiple regression models were developed using two predictors: SOI or Nino 3.4 as the ENSO index; and DMI as the IOD index. The first study used only SOI, while the second provided the results of both SOI and Nino 3.4.

The predicted value of the regression models was calculated by taking the model term coefficients (β) to the equation of the regression models. The first-, second-, and third-order polynomial regression equations are shown by Equation 3.1, Equation 3.2, and Equation 3.3, respectively. The first study used all three polynomial orders, while the second only used the second- and third-orders.

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon \qquad \text{(Equation 3.1)}$$

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 {x_1}^2 + \beta_4 {x_2}^2 + \beta_5 x_1 x_2 + \varepsilon \qquad \text{(Equation 3.2)}$$

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 {x_1}^2 + \beta_4 {x_2}^2 + \beta_5 x_1 x_2 + \beta_6 {x_1}^3 + \beta_7 {x_2}^3 + \beta_8 {x_1}^2 x_2 + \beta_9 x_1 {x_2}^2 + \varepsilon \qquad \text{(Equation 3.3)}$$

with:

$y$ = flow regime index ($Q_{50}$ or $Q_{10}$ or $Q_{90}$),

$x_1$ = ENSO index (SOI or Nino 3.4)

$x_2$ = IOD index (DMI),

$\beta$ = model term coefficient, and

$\varepsilon$ = model error.

## 3.2. Model Evaluation

In this thesis, the quality of the model is referred to as *model skill*. The model skill of the regression models was generally evaluated by comparing and correlating the predicted and the observed values. Besides that, the models are also being

evaluated by the values of these evaluation metrics: the Root Mean Square Error (RMSE), the adjusted $R^2$, and the King-Gupta Efficiency (KGE) coefficient. The first study used the RMSE and adjusted $R^2$, while the second used the KGE and adjusted $R^2$ to evaluate the models.

### 3.2.1. Root Mean Square Error (RMSE)

The standard error of the model's predicted values was measured by the root mean square error (RMSE) using Equation 3.4 [54]. The RMSE metric is used in the first study and displayed as a percentage of the observed values mean.

$$RMSE = \sqrt{\frac{\Sigma(y - y')^2}{N - (k + 1)}} \qquad \text{(Equation 3.4)}$$

with:

$y$ = observed values,

$y'$ = estimated values,

$N$ = number of observations,

$k$ = number of parameters (β, excluding the intercept) to be estimated, and

$\Sigma$ denotes the summation for the monthly set of the $x_1$, $x_2$, and y.

### 3.2.2. Adjusted $R^2$

$R^2$ is a metric to measure the goodness of fit of a model. The coefficient of $R^2$ is a statistical measure of how close the predicted values of a regression model are to the observed data. The adjusted $R^2$ is a better evaluation metric than the ordinary $R^2$ due to the adjustment done by the number of observations and variables. The adjustment is hoped to overcome the bias issue caused by the degree of freedom problem in the ordinary $R^2$. The adjusted $R^2$ is calculated by Equation 3.5. The value of adjusted $R^2$ closer to 1 indicates better model skill.

$$R^2_{adj} = 1 - \left(1 - R^2\right)\frac{n - 1}{n - p - 1} \qquad \text{(Equation 3.5)}$$

with:

$n$ = number of observations, and

$p$ = number of predictors.

### 3.2.3. Kling-Gupta Efficiency (KGE)

The Kling–Gupta efficiency (KGE) coefficient is increasingly used in recent hydrological modeling studies because it has been acknowledged to improve the famous model efficiency metric in hydrology, Nash–Sutcliffe Efficiency (NSE) [55,56]. The NSE is said to have been criticized for varying over an unbounded range; therefore, Gupta and Kling offered the KGE metric to overcome this problem [57].

The KGE coefficient is composed of three components: correlation ($r$), variability (standard deviation, $\sigma$), and bias (mean, $\mu$). The values which can be considered to show any model skill are within $-0.41 < \mathrm{KGE} \leq 1$ [58], in which the model skill becomes better when the value is closer to 1. The metric of KGE will be more useful for future validation studies due to having a bias component. The study of this thesis used the modified KGE [59], in which it replaced the standard deviation (of the variability component) with the coefficient of variation (CV, ratio of standard deviation $\sigma$ to mean $\mu$) of each variable, as displayed by Equation 3.6 below.

$$KGE = 1 - \sqrt{(r-1)^2 + \left(\frac{CV_{sim}}{CV_{obs}} - 1\right)^2 + \left(\frac{\mu_{sim}}{\mu_{obs}} - 1\right)^2} \qquad \text{(Equation 3.6)}$$

$$CV = \frac{\sigma}{\mu}$$

# Chapter 4
# Predictability of Code River's Flow Regime by ENSO and IOD

## 4.1. Framework of the First Study

This study attempted to investigate the non-linear relationship between ENSO and IOD on streamflow by applying multiple polynomial regression analyses. The objective of this study is to understand the streamflow regimes predictability by ENSO and IOD indices in Code River, southern Java Island, Indonesia.

This study consisted of four parts (Figure 4.1). The first part is the correlation analysis between the climate indices and the streamflow regime indices to know their raw relationship. The second part is the development of multiple polynomial regression models to predict the streamflow regime indices using ENSO and IOD indices as predictors. The third part investigates the forecasting ability of time-lagged regression models. The final part is a simulation using dummy variables to understand the modulation effect of IOD on the ENSO–streamflow relationship in the models developed in the previous part.

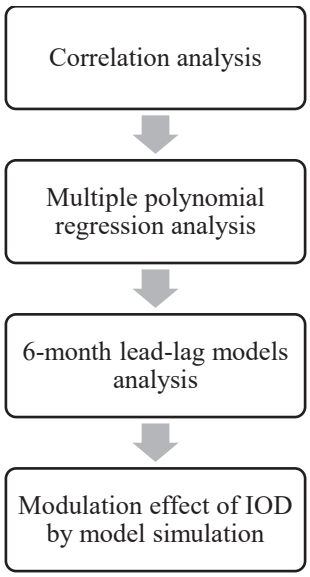


Figure 4.1 Research flow of the first study

## 4.2. Code River Overview

The Special Region of Yogyakarta is a small Indonesian province located in the south-central region of Java Island, occupying an area of 3179 km$^2$. One of the main rivers in the province is the Opak River. Many streams of Opak River tributaries cross in the center of the province in Yogyakarta City. One of the tributaries is Code River, a well-known stream that often represents Yogyakarta City. Code River basin occupies an area of approximately 40 km$^2$. Code River joins Opak River after traveling south for 41 km from its spring on Mount Merapi, north of Yogyakarta (Figure 3.1). As the stream is situated in a strategic location, Code River holds cultural value, aesthetic value, and tourism potential [60].

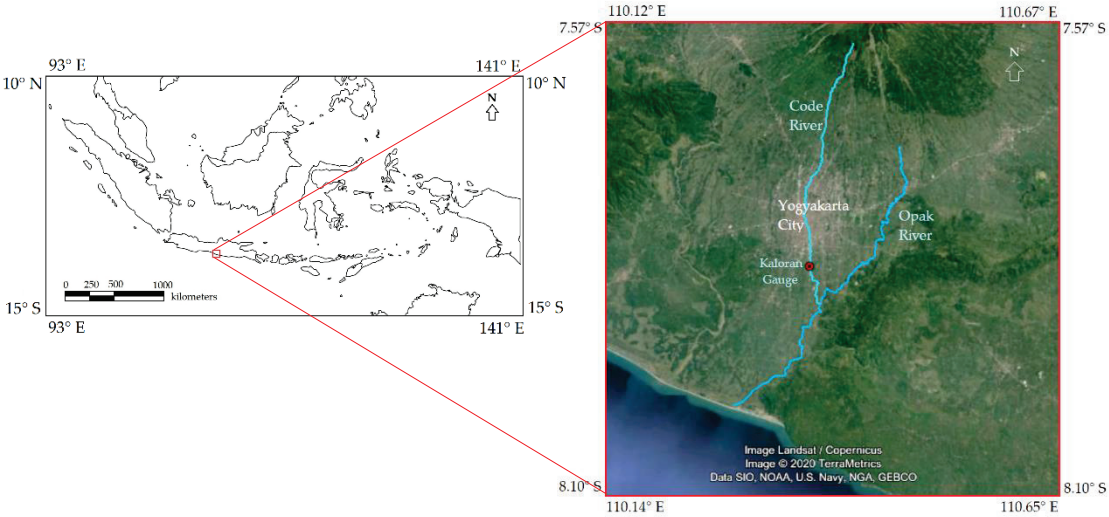


Figure 4.322 Location of Code River in Yogyakarta, southern Indonesia

Based on the Köppen climate type, the Yogyakarta region typically has a tropical savanna ($A_w$) climate. Sometimes, this region has a tropical monsoon ($A_m$) climate type. The region has dry months from May to October (MJJASO)—with August as the driest—and wet months from November to April (NDJFMA)—with January or February as the wettest. The annual rainfall varies between 400 to 3600 mm. The average temperature varies from 22 °C in the upstream area to 27 °C in the downstream area. The lowest temperatures occur in the dry months (June to August). The dry months have lower temperatures than the wet months owing to the cold wind from the southern hemisphere (Australia) winter [61].

## 4.3. Data

The streamflow data used in this study were from the gauge station Kaloran in the downstream part of Code River (Figure 4.2). The streamflow data were obtained from the Indonesian Ministry of Public Works and Housing. The daily streamflow data of Kaloran station has been available for 21 years, covering 1994 to 2018 (Figure 4.3). The river flow regime parameters used in this study are the average, high, and low flows defined by the flow duration curve. The high flow index is important for flood prevention, while the low flow index is useful for drought anticipation and environmental flow setting.

The climate variables used in this study are the Southern Oscillation Index (SOI) and Dipole Mode Index (DMI). The DMI is the common index for IOD and was easy to obtain. The SOI was chosen because it is an easily defined traditional ENSO index and was frequently used by many researchers before to show the relationship between ENSO and precipitation (e.g., [29,41]). The SOI and DMI data were obtained from the USA's National Oceanic and Atmospheric Administration (NOAA).

The data sets used for the analysis were in the monthly time step. While the climate indices were already in the monthly set, the monthly flow regime indices were obtained by extracting the average flow ($Q_{50}$, 50% percentile flow), the high flow ($Q_{10}$, 10% percentile flow), and the low flow ($Q_{90}$, 90% percentile flow) from the daily data of one month.

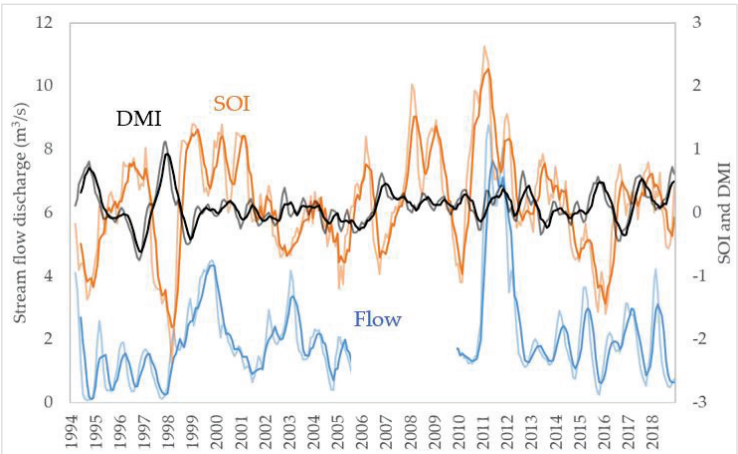


Figure 4.3 Time series of Code River streamflow, SOI, and DMI data displayed in 3-month moving average (light colors) and 6-month moving average (bold colors).

Time series of 3- and 6-month averaged data of streamflow, SOI, and DMI are shown in Figure 4.3. The missing data of streamflow in 2006–2009 were mainly during the time of La Niña (high SOI) events in the wet seasons of 2008 and 2009, indicating missing information on La Niña variance in the analysis. However, the strong La Niña events in 2011 can cover that variance information. For additional information, the strong El Niño (low SOI) events captured in 1998/1999 and 2015/2016 were two of the three super El Niño events identified since 1980 [30].

## 4.4. Results

### 4.4.1. Correlation analysis

This analysis identified the correlation between the observed flow regime indices ($Q_{50}$, $Q_{10}$, and $Q_{90}$) and the observed climate indices (SOI and DMI) to understand their raw relationships. The calculation of correlation was in the monthly time step. Pearson correlation was used to find the correlation coefficient *R*. The significance of the correlation was then evaluated using a two-tailed t-test at the 95% confidence level.

The correlation coefficients as the result of this analysis are shown in Table 4.1. The correlation to the flow regime indices of SOI (ENSO) was distinctively higher than those of the DMI. Although the correlation coefficients of SOI seem low, they are statistically significant. Meanwhile, DMI (IOD) seems uncorrelated with the flow regime indices due to the near-zero *R*. The previous studies showed that IOD correlated with rainfall [15] and low flow [23]. The difference in the results seems to be caused by different timescales used. It is further discussed in Section 4.5.1.

The 3- and 6-month moving averaged data sets were also analyzed to investigate the temporal effect. Six months was used as it is the longest period due to the Indonesian biannual seasons. For the SOI, the 6-month moving average had a higher correlation than the 3-month moving average and the original data set. The 6- month moving average increased the original data set's *R* by 0.167, 0.136, and 0.157 for $Q_{50}$, $Q_{10}$, and $Q_{90}$, respectively. These results indicate that the ENSO effect has a timescale longer than one month and peaked at the 6-month scale. The moving average sets of DMI did not show a noticeable change from the original data sets.

Table 4.1 The correlation coefficients between the observed flow regime indices and climate indices (SOI and DMI) in the monthly data sets. The values in italic are statistically significant at a 95% confidence level with a two-tailed *t*-test.

| **Correlation Variables** | **No Moving Average** | **3-month Moving Average** | **6-month Moving Average** |
|---|---|---|---|
| $Q_{50}$ with SOI | *0.342* | *0.436* | *0.509* |
| $Q_{10}$ with SOI | *0.330* | *0.407* | *0.466* |
| $Q_{90}$ with SOI | *0.347* | *0.413* | *0.504* |
| $Q_{50}$ with DMI | -0.031 | -0.047 | -0.029 |
| $Q_{10}$ with DMI | -0.052 | -0.046 | -0.015 |
| $Q_{90}$ with DMI | -0.099 | -0.109 | -0.104 |

The correlation coefficient becomes higher at the longer timescale of the moving average. It is because the moving average eliminated the variations shorter than the averaging period. This tendency may be related to eliminating variation under a 3- or 6-month timescale; for example, Madden–Julian Oscillation (MJO). The active phase of MJO, which lasts about 1 or 2 months [62], has been known to increase the extreme precipitation events over western and central Indonesia [63]. This kind of noise is wanted to be eliminated because this study focuses on capturing the variances of ENSO and IOD, which variations are in a longer timescale than MJO's variation.

The effect of ENSO on the flow regime was clearly more dominant than the effect of IOD in Code River (Table 4.1). Based on the two-tailed *t*-test, the correlations of SOI to the flow regime indices were all significant, as opposed to the DMI correlations to the flow regime indices, which were all not significant. With low and insignificant correlation values, IOD had no association with the observed streamflow of Code River. On the other hand, ENSO had a positive association with the streamflow—and was the highest in a 6-month timescale. This effect of ENSO was essentially the same with $Q_{50}$, $Q_{10}$, and $Q_{90}$.

The moving average setting could reduce the degree of freedom of the data and may change the result of statistical analysis. We checked the worst case by calculating the correlation coefficient using one datum to resolve this issue. For the 3-month (6-month) moving average, it was one datum every 3 (6) months, making

the total number of data used in each analysis become one-third (one-sixth) of the original moving average data sets. The results found are similar to the values included in the original method of this study (Table 4.1). For example, the correlation coefficient of $Q_{50}$ in this calculation's 3- and 6-month moving averages are 0.440 and 0.516, respectively. The status of the significance also did not change.

### 4.4.2. Multiple regression analysis

Equations 3.1–3.3 were used to develop multiple polynomial regression (MR) models to estimate the $Q_{50}$, $Q_{10}$, and $Q_{90}$ using SOI and DMI as predictors, with the original data set and 3- and 6-month moving average data sets. The regression models were evaluated by correlation coefficient ($R$) between the estimated and the observed data, Root Mean Square Error (RMSE), and adjusted coefficient of determination ($R^2$).

Higher $R$ values between the estimated and observed flow regimes were identified on higher order regressions and longer timescales (Table 4.2). The third-order MR achieved the highest correlation ($R$ = 0.746) for $Q_{50}$ in the 6-month moving average timescale. A better model skill (lower RMSE and higher $R^2$) for the higher order MRs and a longer timescale of moving average was identified (Figure 4.4). The results suggest that the best model's skill was achieved by the 6-month moving average set with third-order MR. Because of the poor model's skill achieved by first-order MR, from now on, it will be omitted from the analysis.

Table 4.2 The Pearson correlation coefficient ($R$) between the estimated and observed values of flow regimes in multiple regression models.

| **Flow Regime Index** | **Multiple Regression Order** | ***R* in different moving average data sets** | | |
|---|---|---|---|---|
| | | **No Moving Average** | **3-month Moving Average** | **6-month Moving Average** |
| $Q_{50}$ | 1st-order | 0.350 | 0.459 | 0.564 |
| | 2nd-order | 0.509 | 0.633 | **0.708** |
| | 3rd-order | 0.553 | 0.650 | **0.746** |
| $Q_{10}$ | 1st-order | 0.334 | 0.429 | 0.522 |
| | 2nd-order | 0.474 | 0.586 | **0.644** |
| | 3rd-order | 0.495 | 0.608 | **0.701** |
| $Q_{90}$ | 1st-order | 0.347 | 0.418 | 0.527 |
| | 2nd-order | 0.474 | 0.563 | **0.652** |
| | 3rd-order | 0.542 | 0.594 | **0.688** |

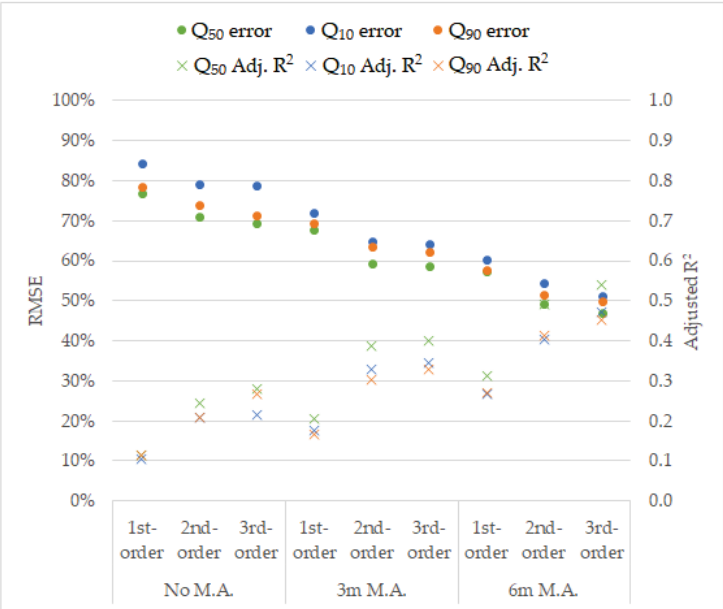


Figure 4.4 The root mean square error (RMSE) and the adjusted $R^2$ of multiple regression (MR) models in 1st-, 2nd-, and 3rd-orders. The 3m M.A. and 6m M.A. are the 3-month and 6-month moving average sets, respectively.

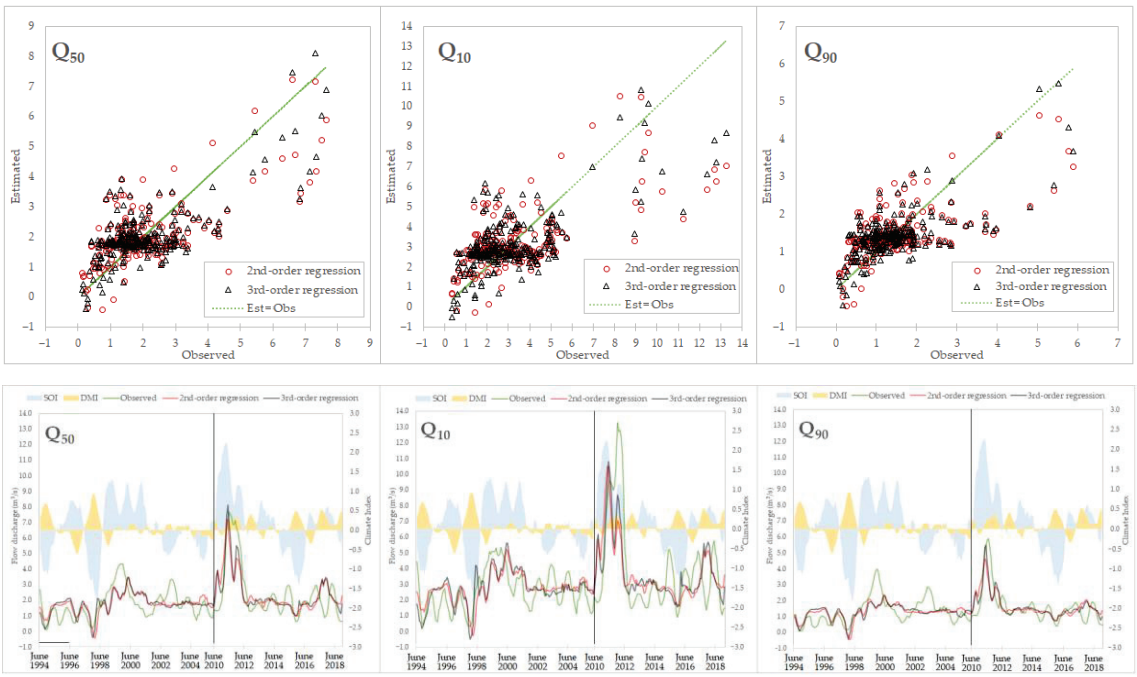


Figure 4.5 Comparison between the observed variables and the estimated flow regimes by multiple polynomial regression model with 6-month moving average data set. The upper figures show the scatterplots of the estimated versus observed values. Lower figures show the time series of observed and estimated values in 1994–2005 (left side of the vertical line) and 2010–2018 (right side of the vertical line).

Figure 4.5 shows the estimated results of the 6-month moving average MR models in scatterplots and time-series graphs. The scatterplot of the estimated versus the observed flow values in Figure 4.5 showed that $Q_{50}$ obtained the most accurate estimates while $Q_{10}$ obtained the least accurate estimates. For all flow indices, both second- and third-order MR models have an underestimate tendency in higher flows. On all flow indices, the third-order MR estimate tended to have a higher value compared with the second-order estimate. The MR models in time series (Figure 4.5) showed that the models had the highest accuracy, mostly in the extreme values, corresponding to strong (positive and negative) SOI values, such as in late 1994, mid-1998, early 2011, and late 2015.

### 4.4.3. Lagged multiple regression analysis

The objective of MR analysis with variables at the same period was to investigate the tendencies of the relationships and cannot be used for forecasting the future without dynamic model prediction. A lagged MR with different periods of response and predictor variables can be used to forecast the response variables. The climate indices in this study may be able to forecast streamflow in the future with a few months' lags applied in the MR. The possibility of forecasting streamflow was attempted in this study through lagged MR analysis with climate indices leading the streamflow data set by six months.

The MR with the same period variables (the MR used in Section 4.4.2) is named lag0, and the name lag6 is used for MR with 6-month lagged variables. The lag0 regression result is shown in the previous section (Section 4.4.2). It shows that the 6-month moving average was the best timescale to reduce variation noises and achieve the highest model skill (highest correlation with the observed flow, lowest standard error, and highest coefficient of determination). Hence, in this lagged MR, we used only the 6-month moving average data set with second- and third-order MR. Additionally, the lag6 MR did not temporally overlap with the use of the 6-month moving average data set. The model evaluations are shown in Table 4.3 and Figure 4.6.

Table 4.3 Multiple regression (MR) model evaluations for a 6-month moving average data set in not-lagged (lag0) and six-month lagged (lag6) MR.

| Flow Regime Indices | Multiple Regression Order | Lag0 | | | Lag6 | | |
|---|---|---|---|---|---|---|---|
| | | ***R*** | **RMSE** | $R^2$ | ***R*** | **RMSE** | $R^2$ |
| $Q_{50}$ | 2nd-order | 0.708 | 49% | 0.490 | 0.763 | 45% | 0.572 |
| | 3rd-order | 0.746 | 47% | 0.539 | 0.779 | 44% | 0.590 |
| $Q_{10}$ | 2nd-order | 0.644 | 55% | 0.402 | 0.701 | 51% | 0.480 |
| | 3rd-order | 0.701 | 51% | 0.472 | 0.724 | 50% | 0.505 |
| $Q_{90}$ | 2nd-order | 0.652 | 52% | 0.413 | 0.694 | 49% | 0.470 |
| | 3rd-order | 0.688 | 50% | 0.452 | 0.742 | 46% | 0.532 |

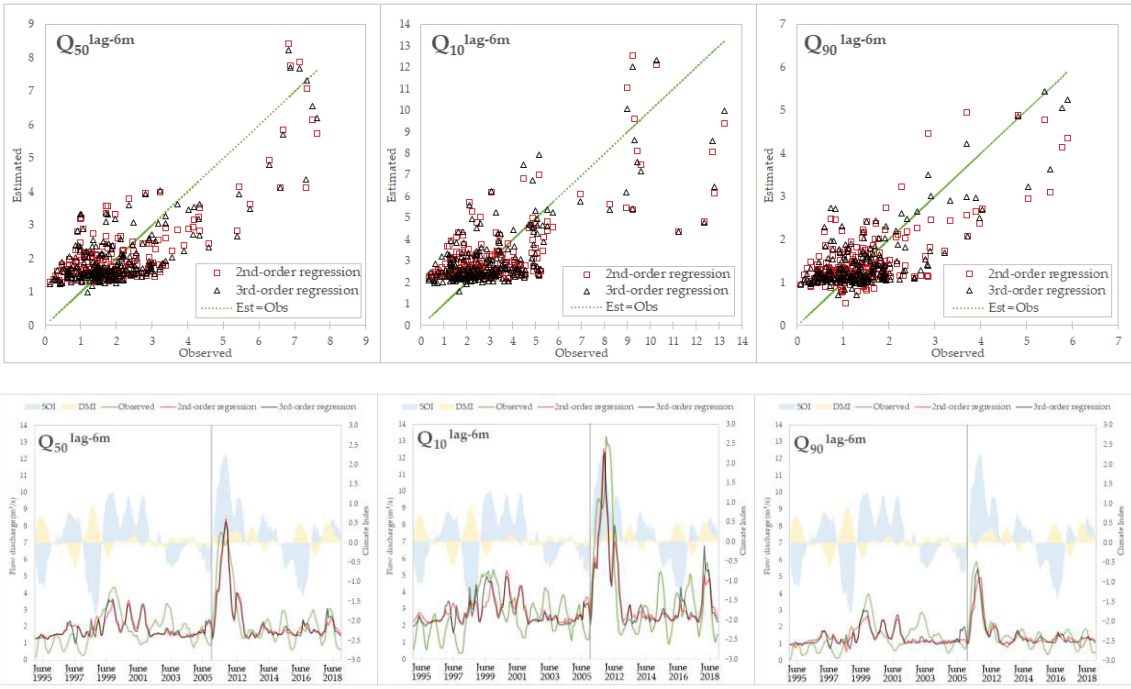


Figure 4.6 Comparison between the observed variables and the estimated flow regimes by 6-month lagged multiple polynomial regression model with 6-month moving average data set. The upper figures show the scatterplots of the estimated versus observed values. The lower figures show the time series of observed and estimated values in 1994–2005 (left side of the vertical line) and 2010–2018 (right side of the vertical line). The months shown on the x-axis are for the streamflow only. The months for the climate indices are six months before the streamflow months.

Based on the correlation coefficient ($R$) of the estimated and observed values, lag6 performed well with a higher $R$ compared with lag0. The third-order regressions had a higher correlation to the observed values than the second-order regressions. However, both orders showed similar patterns (Figure 4.6). The RMSE and adjusted $R^2$ evaluation showed that the lag6 regressions had a lower error and higher $R^2$ compared with lag0, but not by much. This indicates that lagged-time regressions had a slightly better accuracy over the same-time regressions in predicting the flow regimes using SOI and DMI. With these evaluation results, the lag6 MR can make forecasts of streamflow.

## 4.5. Discussion

### 4.5.1. IOD climate effects

The correlation analysis showed that SOI was a prominent indicator of streamflow while DMI was not. Despite its low and insignificant correlation in this study, DMI has been acknowledged for its effect on Indonesia's rainfall [12,14,15]. A previous study in West Java identified a positive IOD (with DMI) in the low flow period [23]. The timescale of the previous study was three months, focusing on the September–November (SON) period. In contrast, this study considered all months. The correlation analysis results (Section 4.4.1) captured the correlation of the ENSO-streamflow but could not capture the correlation of the IOD-streamflow.

Despite the low correlation of IOD to the flow regime indices (Section 4.4.1), the time series of the observed and estimated flows showed that the period of high positive IOD mostly appeared in events of extremely low flow, such as in the year 1994 (Figure 4.3 and 4.6). This agrees with the previous study, which concluded that positive IOD was associated with low flow periods. Considering that the regression models in this study presented relatively good evaluations, IOD can be confirmed to modulate the ENSO–streamflow relationship.

To further investigate the IOD contribution to the ENSO–streamflow relationship in the model developed in this study, the second- and third-order regression models (6-month moving average) were applied for simulations using dummy values of SOI and DMI. The range of SOI (DMI) values in the 6-month moving average data set was −1.80 to 2.27 (−0.61 to 0.93). Therefore, these

simulations' SOI (DMI) variation was set from −2 to 2 (−0.8 to 0.8). The correlation of SOI–streamflow is already known to be positive; thus, a higher SOI should give a higher streamflow. In this case, simulations were done to know whether positive DMI (pDMI) and negative DMI (nDMI) changed the tendency of the positive relationship.

The behavior of MR functions is illustrated by simulation using the dummy data in Figure 4.7. The simulations showed that in a higher positive SOI (pSOI), a higher pDMI produced a more significant increase in the streamflow, and a lower nDMI produced a relatively constant decrease. The second-order MR produced more physically appropriate estimates than the third-order MR, which allowed the estimate values to be below zero, especially on a negative SOI (nSOI). It is due to the lack of variance of nSOI-nDMI combination in the data set for developing the regression models.

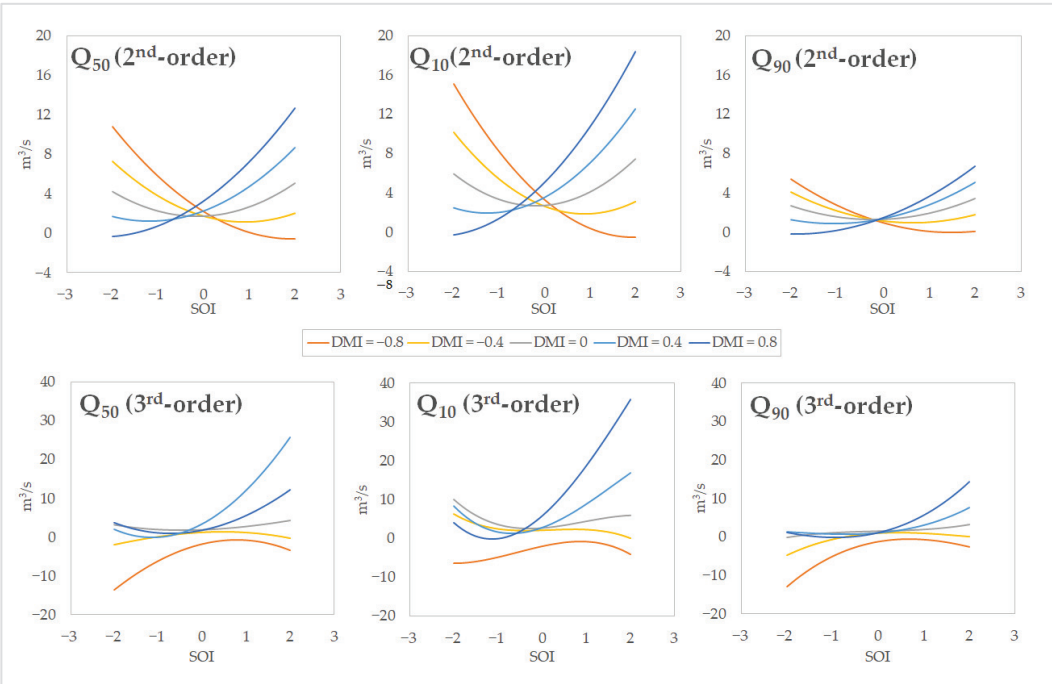


Figure 4.7 Dummy simulations of 2nd- and 3rd-order MR models (6-month moving average) to show the modulation effect of DMI to SOI–streamflow relationship.

In this dummy simulation, the effect of DMI on the SOI–streamflow relationship in this simulation was more apparent at a higher flow to strongly increase the flow in positive SOI events. This modulation of SOI dependency due

to the DMI cannot be expressed in linear (first-order) MR because the crossing terms containing both SOI and DMI included in the second- and third-order MR equations are necessary to express this modulation.

### 4.5.2. Flow regimes forecasting

The forecasted $Q_{50}$ can be used for managing water resources in general, while the forecasted $Q_{10}$ and $Q_{90}$ can be used for anticipating flood and drought events, respectively. Six months in advance is good timing to forecast the flow regime because it may provide a sufficient period for planning the countermeasures to water-related disasters. The evaluation of this study's 6-month lagged multiple regression models showed a good correlation and moderate accuracy (Table 4.3). It indicates that the last 6-month average of SOI and DMI would be able to predict the $Q_{50}$, $Q_{10}$, and $Q_{90}$ in the next six months using the regression model coefficients developed in this study (Table 4.4) with some notes on the accuracy.

A simple and usable model is important for practical uses like forecasting. The goal of developing multiple regression equations is to include meaningful variables that explain most data variability [64]. Therefore, the second-order polynomial equation is generally recommended for forecasting since it has a sufficient model skill that does not fall far behind the third-order MR models (Table 4.4). Furthermore, a lower order polynomial MR keeps the model equation simple and prevents extreme results in the case of inputs outside the considered range.

Table 4.4 Model coefficients of multiple regression (MR) developed from 6-month moving average sets and lagged regressions with 6-month leading climate indices.

| **Flow Regime Indices** | **MR Order** | **Coefficients** | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | **$\beta_0$** | **$\beta_1$** | **$\beta_2$** | **$\beta_3$** | **$\beta_4$** | **$\beta_5$** | **$\beta_6$** | **$\beta_7$** | **$\beta_8$** | **$\beta_9$** |
| | | **Inter-cept** | **SOI** | **DMI** | **$SOI^2$** | **$DMI^2$** | **SOI. DMI** | **$SOI^3$** | **$DMI^3$** | **$SOI^2$. DMI** | **SOI. $DMI^2$** |
| $Q_{50}$ | 2nd | 1.449 | 0.225 | -0.457 | 1.133 | 1.588 | 2.189 | | | | |
| | 3rd | 1.452 | -0.004 | -0.993 | 1.252 | -0.203 | 1.293 | -0.002 | 8.920 | -0.133 | 4.938 |
| $Q_{10}$ | 2nd | 2.210 | 0.164 | -0.173 | 1.614 | 5.990 | 5.247 | | | | |
| | 3rd | 2.182 | -0.001 | -1.406 | 1.777 | 0.699 | 1.978 | -0.135 | 19.498 | 1.865 | 15.137 |
| $Q_{90}$ | 2nd | 1.141 | 0.316 | -0.263 | 0.616 | -1.181 | -0.080 | | | | |
| | 3rd | 1.041 | 0.114 | -0.276 | 1.091 | -0.181 | 1.111 | -0.140 | 2.738 | -1.974 | -0.958 |

The $Q_{50}$ MR models had similar accuracy in second- and third-order MR (Figure 4.5), indicating that both regression orders can be used to forecast $Q_{50}$ with a tendency to underestimate. The only noticeable difference between the two MR orders was in the event of a combination of strong negative IOD (less than −0.5) and strong positive ENSO (greater than 0.5). In this event, which was captured only once in this study (in 1997), the third-order MR estimated an immediately reduced flow, while the second-order MR did not (Figure 4.6, bottom row). For the general purpose of river flow information, a simpler equation with second-order regression may be better.

The $Q_{10}$ MR models have the least accuracy among the three flow indices (Figure 4.5). The most accurate estimation of the model was only found in two events with a combination of positive SOI and positive DMI, in 2011–2012 and 2018. Although both accuracies were similar, the third-order MR was better than the second-order regression in forecasting $Q_{10}$ because the estimation product of the third-order MR tended to be higher than the second-order MR (especially in the event of positive SOI and DMI). Note that a higher estimation of high flow is usually favored to set a safety margin for anticipating floods.

The $Q_{90}$ MR models had better accuracy in the third-order MR than in the second-order MR (Figure 4.6). The MR models of $Q_{90}$ had almost equal numbers in producing underestimated and overestimated values. However, the underestimated (overestimated) values tended to be in the higher (lower) level of $Q_{90}$, while an underestimation is more favorable for anticipating drought and conserving ecosystems in rivers.

### 4.5.3. Model validation

The relationships among variables found in this study are expected to be useful for predicting streamflow using the climate indices. The regression equations in this study (Table 4.4) can be used accurately for forecasting after the models have been validated. Unfortunately, model validation could not be done in this study due to a lack of enough data. The streamflow data obtained by this study covered only 21 years, including only four or five ENSO cycles. All 21 years of data were required to capture the complete variations of ENSO and IOD. With more data for

at least two additional ENSO cycle periods, capturing the modulation effect and validating the model may be possible.

### 4.6. Summary of the First Study

In this study, multiple regression (MR) models have been developed to predict the flow regimes of Code River using the ENSO and IOD indices with second- and third-order polynomial functions. First, the SOI (ENSO) and DMI (IOD) relationship to the Code River streamflow was investigated. Good and significant correlations of SOI to the flow regime indices were found in the 6-month moving average data set. However, monthly data used in this study could not capture the effect of IOD in a simple correlation analysis.

Despite the insignificant low DMI correlation to the flow regime indices, the result of this study mostly agreed with the previous study (in Citarum River), that the extremely low flows appeared in strong positive IOD periods. Considering the relatively good skill obtained by the MR models in this study, the IOD was confirmed to modulate the ENSO–streamflow relationship in further investigation. This effect could be captured by second- and third-order polynomial MRs. The effect of IOD on the streamflow was identified to greatly increase the higher flows (of $Q_{50}$ and $Q_{10}$) in the positive ENSO phases.

The MR analysis identified the 6-month moving average as the highest correlated timescale. It was also identified that the third-order regression model has better skills than the second-order regression model. However, the difference between their estimation results was not significant. Therefore, both regression orders can be used to make a prediction. A time-lagged MR analysis was conducted using the 6-month moving average data set and the two regression orders. It was found that 6-month lagged (lag6) MR was better in explaining more flow regime variability than the same-time (lag0) MR. The lag6 MR model showed a relatively good evaluation with moderate accuracy: most $R$ values with the observed data were over 0.700, 44–51% RMSE, and 0.470–0.590 adjusted $R^2$.

It is beneficial for water resource managers to be able to forecast the flow indices six months in advance using climate indices. Flow regime forecasting using 6-month lagged (lag6) multiple polynomial regression was possible with some

caution regarding the model's accuracy. The average flow index $Q_{50}$ had the most stable and accurate estimates; thus, it does not require a higher order than the second-order MR to forecast. Oppositely, the high flow index $Q_{10}$ and low flow index $Q_{90}$ had quite low accurate model estimates; thus, using the third-order MR for forecasting is recommended.

The lag6 MR model is expected to be useful for streamflow forecasting after a proper model validation in the future. In this study, the model validation could not be done due to a lack of data; thus, it becomes a future work. Nevertheless, this study demonstrated non-linear climate–streamflow relationships and streamflow predictability with the climate indices using multiple polynomial regressions.

# Chapter 5
# Spatial and Temporal Pattern of the Predictability of Streamflow Regimes in Java

## 5.1. Framework of the Second Study

The first study has successfully demonstrated the use of multiple polynomial regression models for explaining the predictability of streamflow regimes by the climate indices of ENSO (SOI) and IOD (DMI) on a river in the south of Java Island. Although the first study has developed a quite good statistical model, it was still unknown whether the method could be applied to develop similar good models for other rivers in the same region. Even though the areas within Java Island are fallen in the same rainfall region [11], the different characteristics of rivers (e.g., catchment area, land use, and gauge location) could affect the relationship with climates.

This study extended the previous achievement of the first study by applying the multiple polynomial regression approach to several rivers in Java. This study aimed to investigate the predictability of streamflow regimes by the ENSO and IOD indices in the Javanese rivers, which are not limited to natural rivers. The investigation was conducted by finding temporal and spatial patterns of the flow regimes' predictability. The different timescales and seasonal periods are investigated to find the highest predictability skill. Since ENSO and IOD originated from different sides of Indonesia, different effects of them on the western and eastern sides of Java maybe exist.

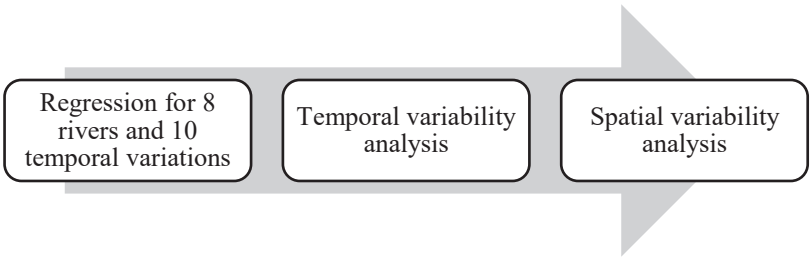


Figure 5.1 Research flow of the second study

The regression analysis in this study used one of the ENSO indices (SOI or Niño 3.4) and DMI as predictors to predict $Q_{50}$ (median flow), $Q_{10}$ (high flow), and $Q_{90}$ (low flow) for each river. Multiple regression (MR) models were developed using second-order (Equation 3.2) and third-order (Equation 3.3) polynomial functions to minimize the model's error term. All regression analyses and statistics in this study were carried out by using R programming [65,66]. The multiple polynomial regression in this study was evaluated by the adjusted coefficient of determination (adjusted-$R^2$) and the Kling–Gupta Efficiency (KGE) coefficient. Both evaluation metrics were used so it can be easily compared to other studies which use the common metric ($R^2$) and the recently popular metric (KGE). The two evaluation metrics were calculated using the R package named hydroGOF [67].

## 5.2. Data

### 5.2.1. Streamflow data

The streamflow data used in this study were obtained from the Research Center of Water Resource under the Ministry of Public Works Indonesia. The daily streamflow data from 1970–2018 of eight Indonesian rivers on Java Island were used. The eight rivers were Ciujung River (Cu), Cisadane River (Cs), Bodri River (Bo), Progo River (Pr), Bengawan Solo River (BS), Madiun River (Ma), Baru River (Ba), and Code River (Co). The profile and location of those rivers can be found in Table 5.1 and Figure 5.2.

The river selection is based on three criteria: temporal data quality, geographical location, and catchment area. Meanwhile, Co River is taken from the first study for comparison. The data quality was first controlled by sorting the river gauges with minimum flow records of 30 years. Years with too many missing daily flows, also zero and interpolated flows, were omitted. To compare the effect of ENSO, which originated from the Pacific Ocean (east of Java), and IOD, which originated from the Indian Ocean (west of Java), we selected at least two rivers for each of the eastern, western, and central regions of Java. Finally, rivers with catchment areas as similar as possible were selected. However, due to the limitation of the other criteria, the selected rivers based on the catchment area became these

Table 5.1 Profiles of the streamflow gauge stations

| Code-name | River Name | Flow Records Used (year) | Catchment Area ($km^2$) | Region in Java | Location Coordinates |
|---|---|---|---|---|---|
| Cu | Ciujung | 42 | 1562.7 | West | -6.133, 106.300 |
| Cs | Cisadane | 44 | 850.2 | West | -6.514, 106.678 |
| Bo | Bodri | 31 | 522.3 | Central | -7.011, 110.137 |
| Pr | Progo | 38 | 423.4 | Central | -7.340, 110.209 |
| BS | Bengawan Solo | 42 | 11125.0 | East | -7.150, 111.599 |
| Ma | Madiun | 30 | 2126.0 | East | -7.645, 111.513 |
| Ba | Baru | 33 | 454.0 | East | -8.413, 114.094 |
| Co | Code | 21 | 40.0 | Central | -7.849, 110.375 |

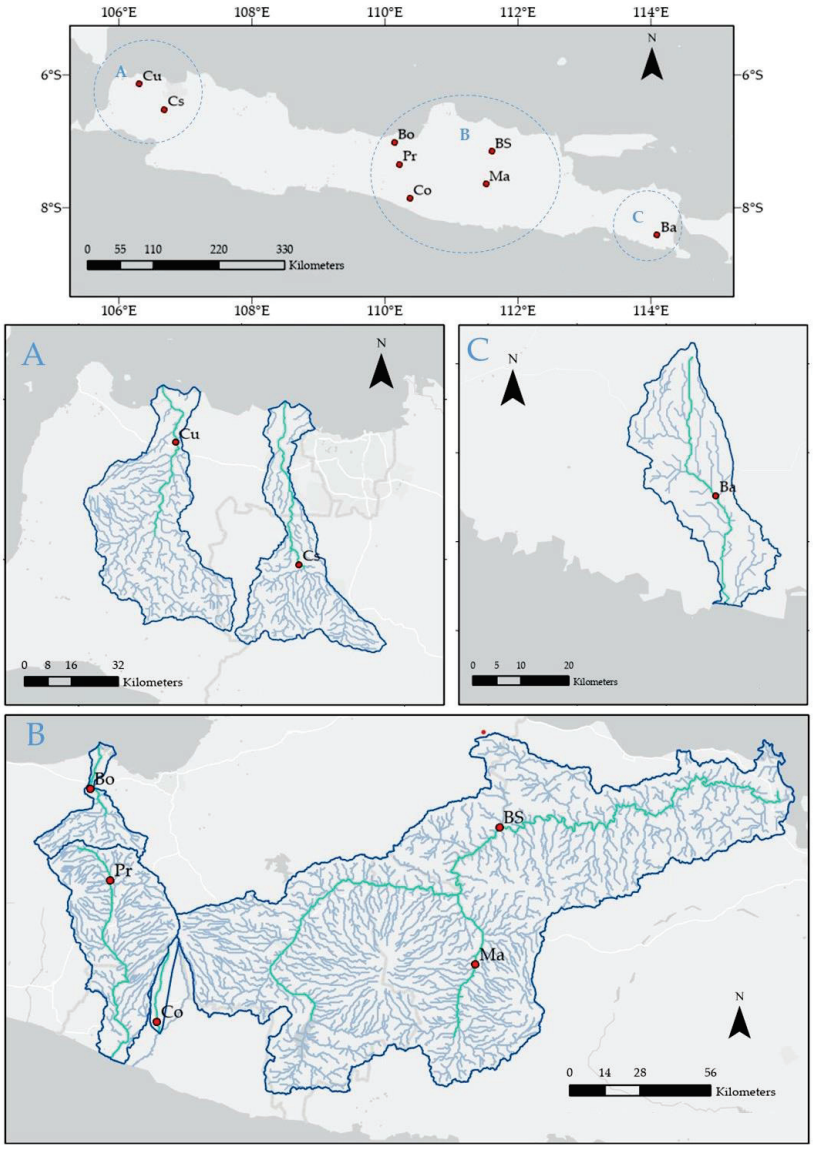


Figure 5.2 Location of the selected rivers in Java Island, Indonesia

three categories: <1000 $km^2$ (Cs, Bo, Pr, Ba), >1000 $km^2$ (Cu, Ma), and >10,000 $km^2$ (BS). For information, the BS River is the largest river basin in Java which flows from Central Java to East Java.

The streamflow data analyzed in this study are in the form of flow regime indices based on the flow duration curve (FDC) definition. The FDC is a cumulative frequency curve that depicts the percentage of time that specified discharges were equaled or exceeded in a particular period [68]. This study used three flow regime indices, which are medium flow ($Q_{50}$), high flow ($Q_{10}$), and low flow ($Q_{90}$). With the definition of the FDC, $Q_{10}$ ($Q_{90}$) is the flow equaled or exceeded at 10 (90) percent of the time. Meanwhile, in data distribution, the definition of $Q_{10}$ ($Q_{90}$) is the $90^{th}$ ($10^{th}$) percentile due to the data ranking from lowest to highest. The three flow regimes were calculated using the R package fasstr [69].

### 5.2.2. Climate data

The ENSO indices used in this study were the Southern Oscillation Index (SOI) and Niño 3.4, while Dipole Mode Index (DMI) was used for the IOD index. The DMI is the most commonly used index for IOD. SOI is the traditional ENSO index based on different sea level pressure between Darwin and Tahiti. It has been generally used in the early [70] and recent [53] research. The Niño 3.4 is a sea surface temperature (SST) index, located in 5° N – 5° S and 170° W – 120° W, that has been acknowledged as the most representative index for ENSO [28] and was frequently used in many recent studies [29–31]. The SOI, Niño 3.4 and DMI data sets were obtained from the USA National Oceanic and Atmospheric Administration (NOAA).

The first study used SOI as the ENSO index, while this study compared two ENSO indices, SOI and Niño 3.4. The correlation between the two ENSO indices using monthly data from 1970–2018 (Figure 5.3) was highly negative ($R = -0.71$), with high positive DMI mostly appearing in the El Niño phase (negative SOI, positive Niño 3.4). The SOI and Niño 3.4 did not differ by much in showing the dynamics of ENSO. Therefore, the results of regression models from both climate indices were expected to be similar.

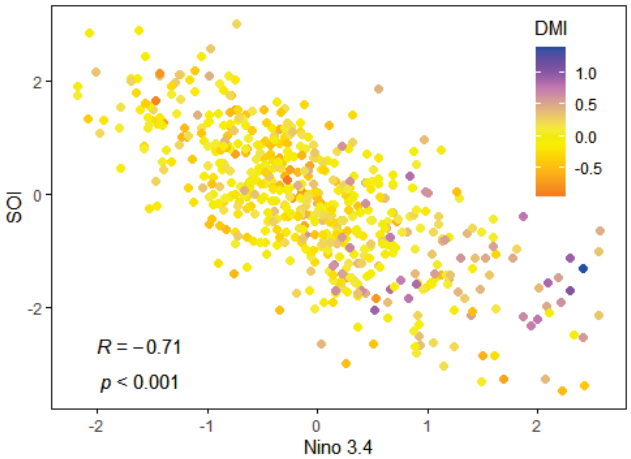


Figure 5.3 Scatterplot of SOI and Niño 3.4 with DMI levels (1970-2018)

### 5.2.3. Temporal variation sets

This study attempted to find the best regression to explain ENSO and IOD variance on the streamflow data with four sets. The first, second, and third sets are monthly series, while the fourth set is an annual series.

The first set is a series of monthly flows calculated from the distribution of daily flows. The second set is a six-month moving average series of monthly flow regime indices calculated from the average of the first series. The purpose of averaging over six months of monthly data is to reduce the effect of short timescale variation such as Madden–Julian Oscillation (MJO) [71]. The third set is a series of seasonal monthly flows calculated as the first series in the three-monthly season window (explained below). The fourth set is a series of the seasonal average calculated by averaging the third set for each year. Table 5.2 shows the detail of the streamflow data arrangement on each set. ENSO and IOD indices data sets followed the same temporal arrangement.

There are four seasons: March-April-May (MAM), June-July-August (JJA), Sep-September-October-November (SON), and December-January-February (DJF). The DJF series started in the year 1970, which consists of 1970 December, 1971 January, and 1971 February. This study strictly uses the 1970–2018 period; hence no 2019 data at all, even for DJF. Therefore, the period of the DJF series is 1970–2017, slightly different from the other three seasons, whose period is 1970–2018.

Table 5.2 Temporal variation sets calculation method

| Set Name | Codename | Timescale | Calculation Method |
|---|---|---|---|
| Ordinary monthly | MON | monthly flow | $Q_{50}$: median of daily flow distribution in a month.<br>$Q_{90}$: 10% percentile of daily flow distribution in a month.<br>$Q_{10}$: 90% percentile of daily flow distribution in a month.<br>Thus, there are 12 data sets for each year. |
| Six-month moving average | MA6 | monthly flow with six-month information | Mean of each flow regime monthly series for six months.<br>Thus, there are 12 data sets for each year, except for one year. |
| Seasonal monthly (SM) | MAM/ JJA/ SON/ DJF | monthly flow | The same as with the ordinary monthly flow series, except the series is limited only to three months per season (e.g., $Q_{50}$ of the MAM series includes the monthly $Q_{50}$ of only March, April, and May, from 1970 to 2018).<br>Thus, there are three data sets for each year. |
| Seasonal average ($SM_a$) | $MAM_a$/ $JJA_a$/ $SON_a$/ $DJF_a$ | three-month flow | Mean of the seasonal monthly series (e.g., the first datum in the $MAM_a$ series is the mean of three months: March, April, and May, in 1970). Thus, there is only one datum for each year. |

## 5.3. Results

### 5.4.1. Temporal variability

Multiple regression (MR) models with second- and third-orders polynomial functions were developed using ten different temporal variation sets. The objective is to find the best temporal set for estimating the flow regime indices using ENSO and IOD indices. The MR models were first developed by all possible variables.

The model evaluation using adjusted-$R^2$ and KGE coefficient (Figure 5.3) showed that the regression of streamflow by ENSO and IOD indices is best achieved by the seasonal three-month average ($SM_a$) set in September–November ($SON_a$). The September–November (SON) months in the seasonal monthly (SM) set also performed better among the SM sets. The results indicate two things.

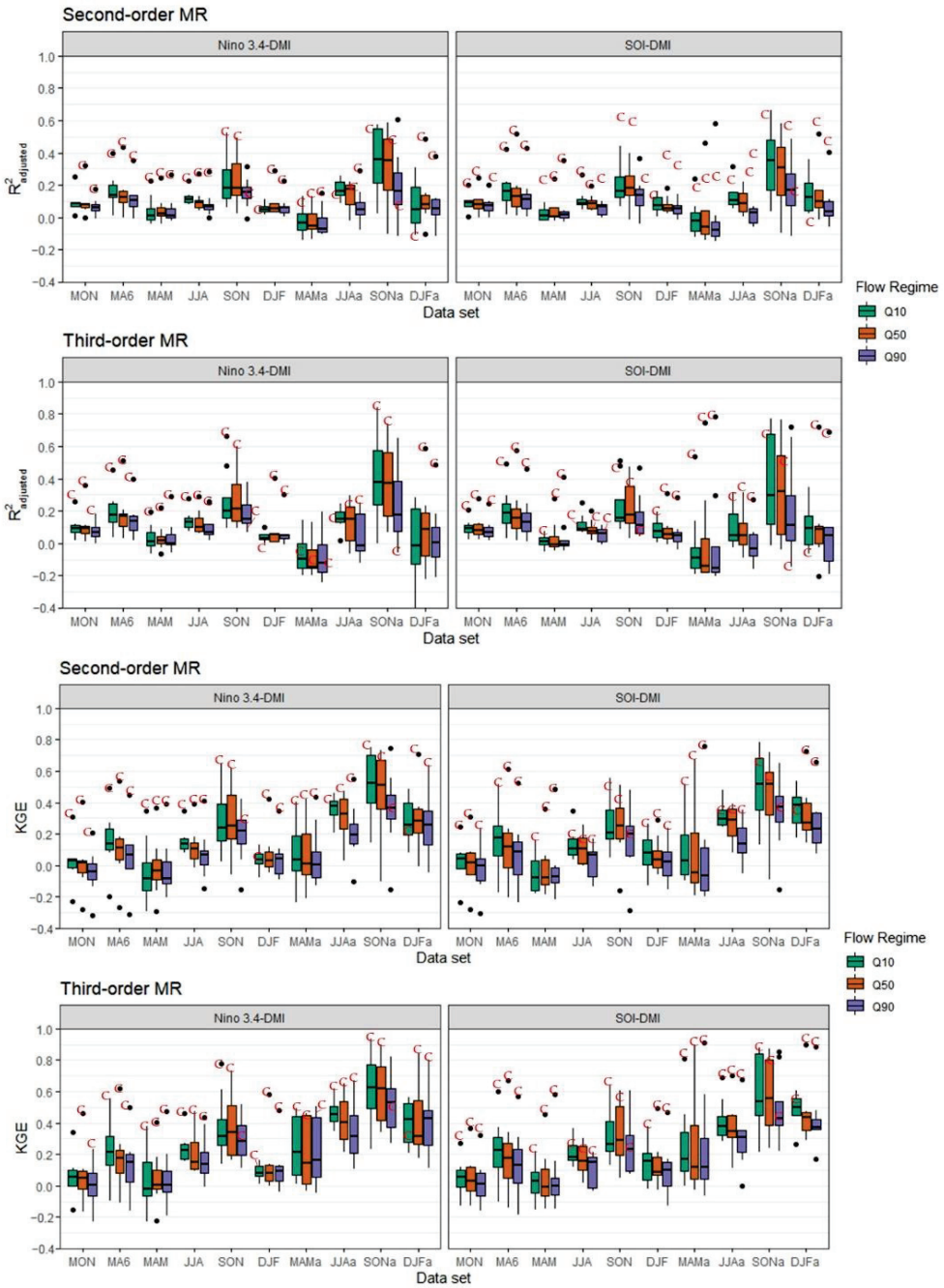


Figure 5.4 Boxplot series showing model's evaluation metrics of adjusted $R^2$ and KGE coefficient. The x-axis shows the temporal variation sets (MON: monthly flow, MA6: six-month moving average flow, MAM-DJF: monthly flow in the window of each respective season, MAMa–DJFa: three-month average flow in the window of each respective season). The lines in the middle of the box are median. The top of the box is the third quartile (Q3), and the bottom is the first quartile (Q1). The vertical lines above and below the box show the data range, excluding outliers. Outliers, values farther than 1.5 times their interquartile ranges (IQR = Q3−Q1), are shown as the black dots. The red C points are Code River data.

First, the streamflow predictability by ENSO and IOD was good only in the September–November season, when the transition from dry season to wet season usually occurs in the Java region [72]. Many previous studies have found a high correlation of ENSO and IOD during May–October or September–November to the rainfall [10,11,15,52,53,72] and streamflow [23] in the Java region.

Second, the $SM_a$ sets generally have better model skills than the SM sets. It indicates that even with monthly data series in three-month windows, just ENSO and IOD are not enough to satisfy the need to describe the variance of the streamflows. The comparison between the monthly set (MON) and the six-month moving average set (MA6) also showed the same behavior. MA6 obtained a better model skill than MON due to MA6 carrying over the information of streamflow and climate indices for six months. Thus, it could filter the noises from other factors, such as another climate phenomenon with a shorter timescale, such as Madden–Julian Oscillation.

In general, the $Q_{50}$ and $Q_{10}$ models achieved better model skills than the $Q_{90}$. The overall distribution of $Q_{90}$ model skills is always lower than the other two regimes in SON and $SON_a$. It corresponds to higher flow tending to be more sensitive in the regression with ENSO and IOD indices, as shown in a simulation in the first study (Section 4.5.1, Figure 4.7). A previous study on a river in northwestern Java showed that low streamflow events are more related to positive IOD based on a probability analysis [23]. However, the result of this study cannot be compared directly to that study due to the different methods. While this study used ENSO and IOD indices together in multiple regression, that study treated them independently in simple correlation analyses.

As expected, there were no significant differences between the model predicted by SOI-DMI and those by Niño 3.4-DMI. Only a few insignificant differences appeared in the maximum or minimum values of adjusted $R^2$ and KGE coefficients (Figure 5.4). It may also be worth noting that, in the averaged seasonal sets ($MAM_a$, $JJA_a$, $SON_a$, $DJF_a$) of the third-order models, the Niño 3.4 variant showed no KGE outliers. In contrast, the SOI variant has more than ten outliers of KGE values. Meanwhile, the second-order models had no significant difference in

KGE sets. This may indicate that the quality of the models developed in third-order MR is better when predicted by Niño 3.4 than SOI.

The results of the third-order MR were better than those of the second-order MR. The second (third)-order MR models for $Q_{10}$, $Q_{50}$, and $Q_{90}$, which Niño 3.4-DMI predicted in $SON_a$, have the maximum KGE of 0.75 (0.94), 0.74 (0.90), and 0.55 (0.82), respectively. Additionally, in the prediction by Niño 3.4-DMI in $SON_a$, the $R^2$ and KGE's outliers are found only in the second-order MR models. Overall, the third-order MR improved the model skills after the second-order MR, not only for the maximum but also for the minimum values of the evaluation metrics (Figure 5.4). This improvement is more apparent in the higher flows, as shown by the example of the Ma River in Figure 5.5.

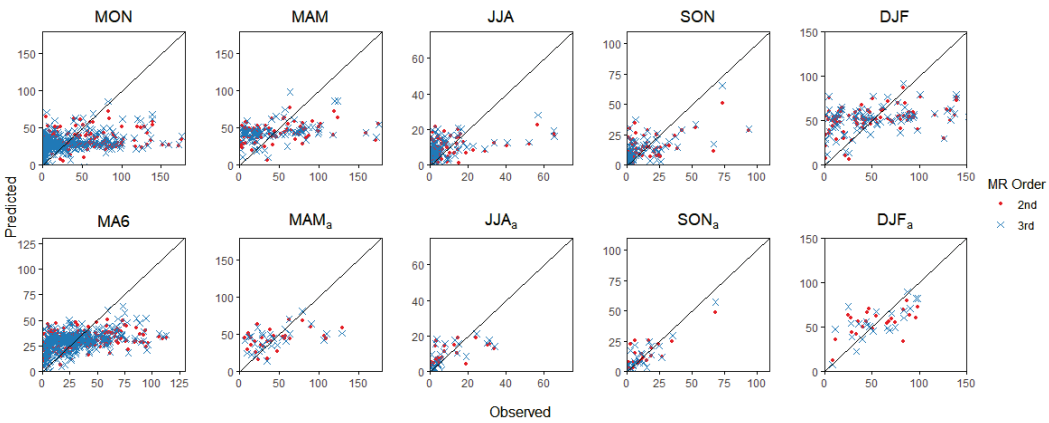


Figure 5.5 Scatterplots between the predicted (2nd- and 3rd-order MR with Niño 3.4-DMI as predictors) and the observed $Q_{50}$ of Madiun (Ma) River in all four temporal variants: monthly (MON), 6-month moving average (MA6), monthly seasonal (MAM-DJF), and averaged monthly seasonal ($MAM_a$–$DJF_a$).

### 5.4.2. Spatial variability

In the previous section, it has already been known that the best model skill is found on the $SON_a$ set. In Figure 5.4, the boxplots of the $SON_a$ evaluation metrics are generally longer relative to the other sets, indicating a higher spatial variability among the temporal variation sets for all rivers. This section explains the spatial variability among the rivers by analyzing the goodness-of-fit metrics (Table 5.3–5.5). Generally, the west and the central rivers (Cu, Cs, Bo, Pr) have lower predictability than the eastern rivers (BS, Ma, Ba). The western river models also

tend to underestimate the higher flows, while the eastern river models' estimations are spread evenly (Figure 5.6).

Ma River showed the best model skills (Table 5.3–5.5) by consistently achieving excellent evaluation for all flow regimes and all MR orders. The scatterplots of second-order MR in Figure 5.6 showed that Ma River has some far-spread overestimation in the lower $Q_{50}$ and $Q_{90}$, even though their correlation coefficients (*R*) were higher than that of the $Q_{10}$. In that case, it can be considered that another river in the same region, the Ba River, has a higher prediction skill due to having more consistent lower error distances. Only the highest flow of each flow regime in Ba River showed a relatively too far error distance, which underestimates about 50% of the observed values. Those error distances were lowered in the third-order MR models (Figure 5.7).

Meanwhile, the Pr River in central Java consistently showed exceptionally poor model skills among all rivers in this study. With the negative values on KGE for all flow regimes, Pr becomes the only river that could not satisfy the minimum standard of KGE (−0.41) by the second-order MR modeling. The values of *R* (Table 5.3), adjusted-$R^2$ (Table 5.4), and KGE (Table 5.5) all agreed upon the conclusion that Pr streamflow regimes cannot sufficiently develop a relationship model to the climate indices.

Table 5.3 Correlation coefficient (*R* Pearson) between observed and predicted flow regimes by Niño 3.4 and DMI in the $SON_a$ set. Values higher than 0.7 are bolded, indicating good model skill.

| **River Name** | | ***R* of 2nd-order MR** | | | ***R* of 3rd-order MR** | | |
|---|---|---|---|---|---|---|---|
| **Original** | **Initial** | **$Q_{50}$** | **$Q_{10}$** | **$Q_{90}$** | **$Q_{50}$** | **$Q_{10}$** | **$Q_{90}$** |
| Ciujung | Cu | 0.60 | 0.63 | 0.52 | 0.61 | 0.65 | 0.54 |
| Cisadane | Cs | 0.49 | 0.43 | 0.44 | 0.53 | 0.46 | 0.48 |
| Bodri | Bo | 0.56 | 0.62 | 0.55 | 0.67 | **0.73** | 0.69 |
| Progo | Pr | 0.22 | 0.40 | 0.19 | 0.54 | 0.62 | 0.56 |
| Bengawan Solo | BS | **0.71** | **0.70** | 0.58 | **0.79** | **0.75** | **0.72** |
| Madiun | Ma | **0.81** | **0.79** | **0.82** | **0.88** | **0.84** | **0.87** |
| Baru | Ba | **0.76** | **0.80** | 0.69 | **0.81** | **0.84** | **0.75** |
| Code | Co | **0.78** | **0.83** | 0.56 | **0.93** | **0.96** | 0.65 |

Table 5.4 Adjusted $R^2$ for the models using Niño 3.4 and DMI as predictors in the $SON_a$ set. Values higher than 0.5 are bolded, indicating good model skill.

| River Name | | $R^2_{adj.}$ of 2nd-order MR | | | $R^2_{adj.}$ of 3rd-order MR | | |
|---|---|---|---|---|---|---|---|
| **Original** | **Initial** | $Q_{50}$ | $Q_{10}$ | $Q_{90}$ | $Q_{50}$ | $Q_{10}$ | $Q_{90}$ |
| Ciujung | Cu | 0.28 | 0.17 | 0.31 | 0.19 | 0.10 | 0.26 |
| Cisadane | Cs | 0.14 | 0.09 | 0.08 | 0.09 | 0.03 | 0.00 |
| Bodri | Bo | 0.18 | 0.16 | 0.26 | 0.21 | 0.26 | 0.33 |
| Progo | Pr | −0.10 | −0.12 | 0.02 | 0.07 | 0.09 | 0.19 |
| Bengawan Solo | BS | 0.43 | 0.25 | 0.41 | **0.53** | 0.38 | 0.43 |
| Madiun | Ma | **0.59** | **0.61** | **0.54** | **0.67** | **0.65** | **0.57** |
| Baru | Ba | 0.50 | 0.37 | **0.58** | **0.53** | 0.40 | **0.60** |
| Code | Co | 0.48 | 0.08 | **0.57** | **0.74** | −0.06 | **0.84** |

Table 5.5 KGE for the models using Niño 3.4 and DMI as predictors in the $SON_a$ set. Values higher than 0.5 are bolded, indicating good model skill.

| River Name | | KGE of 2nd-order MR | | | KGE of 3rd-order MR | | |
|---|---|---|---|---|---|---|---|
| **Original** | **Initial** | $Q_{50}$ | $Q_{10}$ | $Q_{90}$ | $Q_{50}$ | $Q_{10}$ | $Q_{90}$ |
| Ciujung | Cu | 0.24 | −0.04 | 0.30 | 0.24 | 0.05 | 0.35 |
| Cisadane | Cs | −0.15 | −0.40 | −0.44 | 0.00 | −0.19 | −0.31 |
| Bodri | Bo | 0.11 | 0.06 | 0.28 | 0.41 | 0.46 | **0.54** |
| Progo | Pr | −2.65 | −3.48 | −0.64 | 0.04 | 0.09 | 0.28 |
| Bengawan Solo | BS | 0.49 | 0.17 | 0.47 | **0.67** | **0.52** | **0.58** |
| Madiun | Ma | **0.70** | **0.72** | **0.65** | **0.82** | **0.80** | **0.75** |
| Baru | Ba | **0.61** | 0.44 | **0.68** | **0.70** | **0.59** | **0.76** |
| Code | Co | **0.64** | 0.08 | **0.73** | **0.89** | 0.35 | **0.94** |

Scatterplots of the second-order MR in Figure 5.6 showed in more detail that the predicted-observed correlation of Pr River is poor compared to those of the other rivers. Moreover, the Pr River has two flow regimes ($Q_{50}$ and $Q_{90}$) with statistically insignificant $R$. Although in the third-order MR its KGE and $R$ values become above the minimum standards (KGE > −0.41 and $R$ > 0.50), the models are still weak with the adjusted $R^2$ values of only 0.07–0.19. The third-order MR's prediction-observed scatterplots (Figure 5.7) showed that the prediction skill in the Pr River is more improved than in the second-order MR. Even so, the spread of the estimations could not make it a strong model.

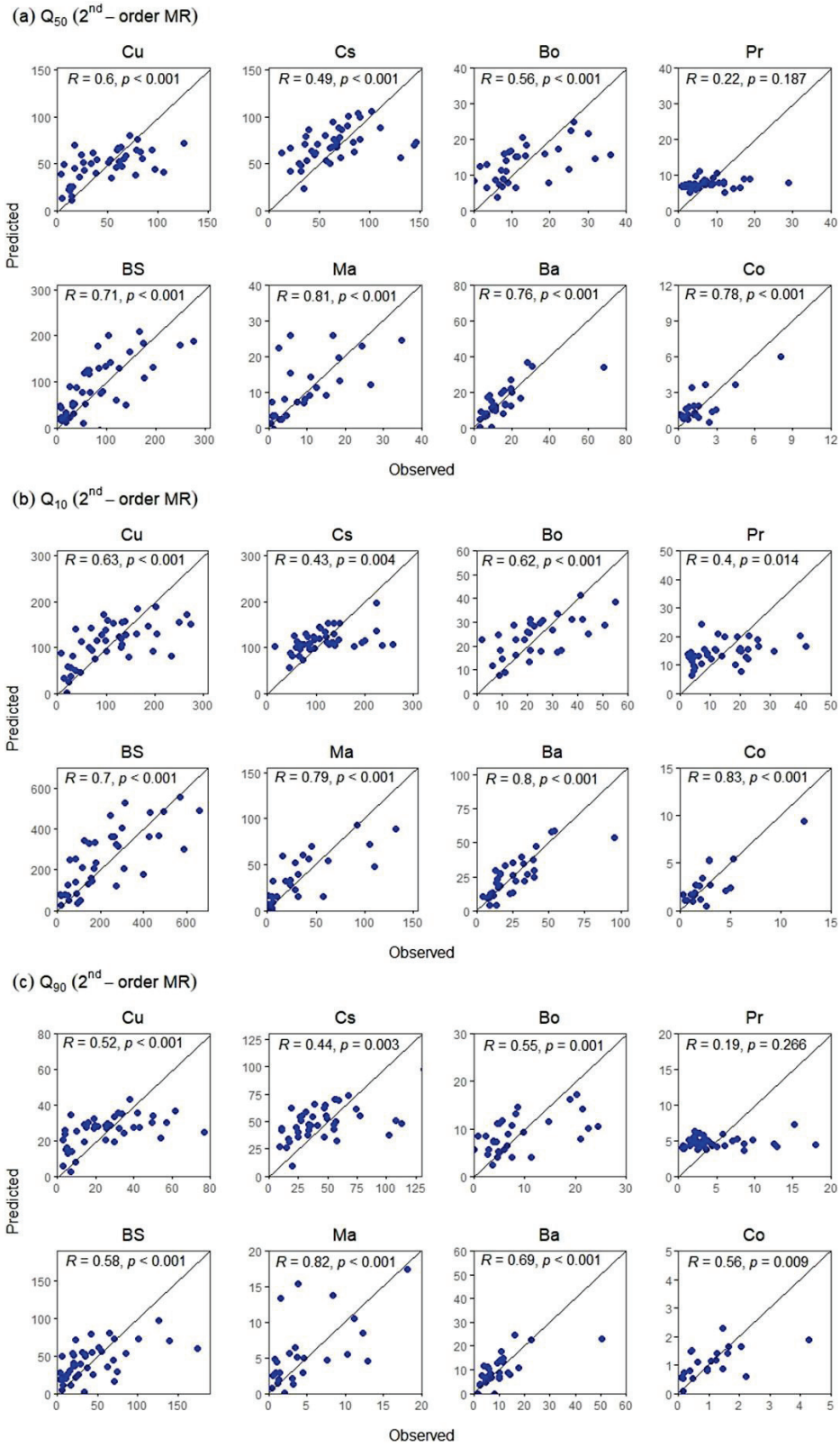


Figure 5.6 Scatterplots between the predicted (Niño 3.4-DMI, 2nd-order MR, $SON_a$ set) and the observed values (in $m^3/s$) of the flow regimes (a) $Q_{50}$, (b) $Q_{10}$, and (c) $Q_{90}$ for all rivers in this study.

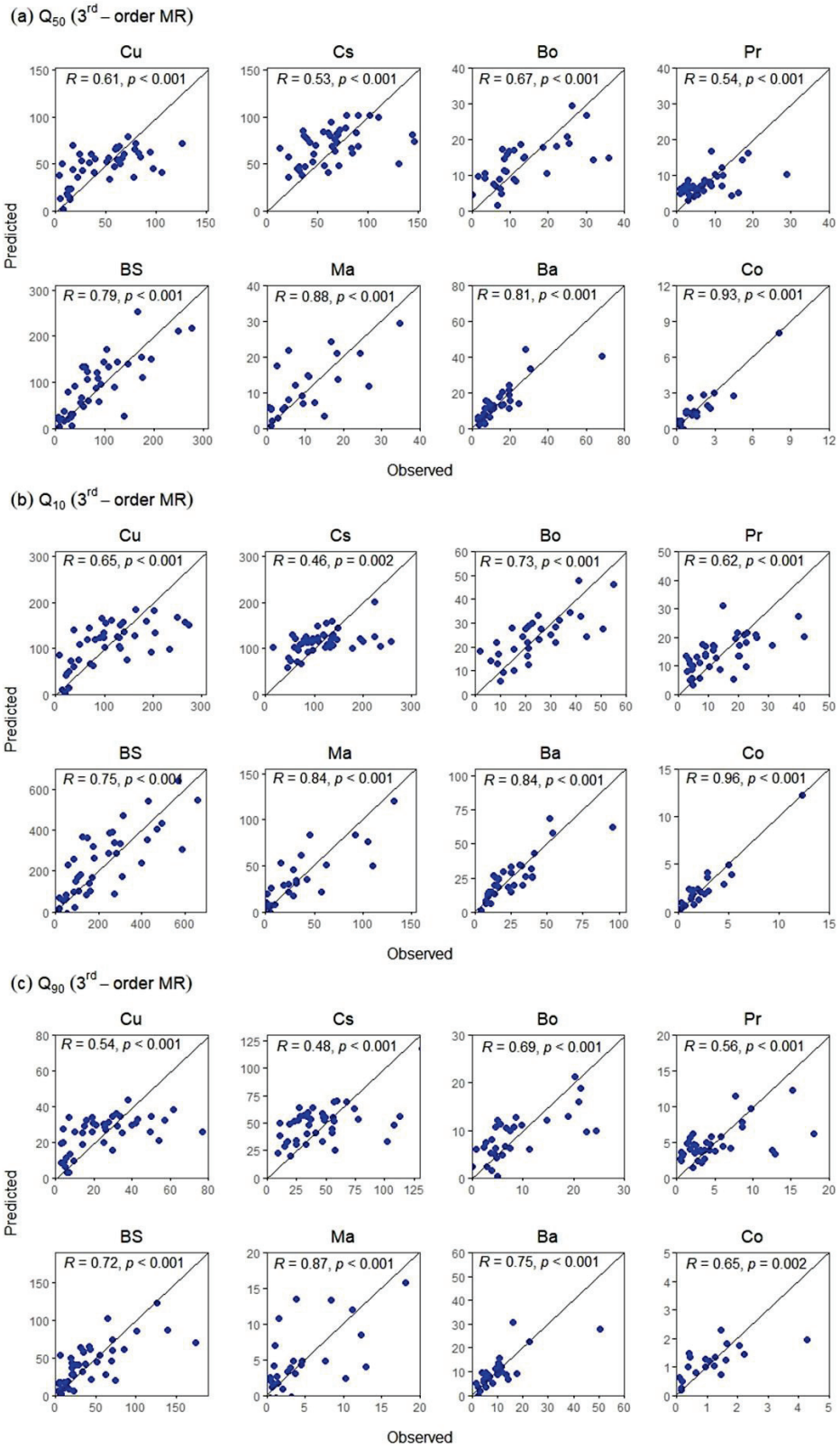


Figure 5.7 Scatterplots between the predicted (Niño 3.4-DMI, 3rd-order MR, $SON_a$ set) and the observed values (in $m^3/s$) of the flow regimes (a) $Q_{50}$, (b) $Q_{10}$, and (c) $Q_{90}$ for all rivers in this study.

Until this part, it can be concluded that: 1) the eastern rivers could be predicted better than the western and central rivers, 2) the Ma River in eastern Java achieved the highest model skill, and 3) the Pr River in central Java has the worst model skill. Although a minor point, it is also noted that the Bo River did not show a similar feature to the Pr River, although they are in the same region. Instead, it showed more similar quality to the western rivers, especially the Cu River. Many possible factors could affect the difference in predictability among the rivers. The tendency is investigated further in Section 5.4 to find the reason for the different predictabilities.

## 5.4. Discussion

### 5.4.1. Predictability tendency by the number of observations

The rivers of Cu and Cs in western Java, along with the BS River in eastern Java, have the highest number of observations among the eight rivers in this study (Table 5.1). The model evaluation metrics (Table 5.3–5.5) showed that the rivers of Cu and Cs have lower model skills than the eastern rivers. Typically, this phenomenon may be addressed on the problem of degree of freedom. A lower degree of freedom could lead to overfitting a regression model when dealing with higher order polynomial functions in a regression.

The plots between the number of observations used on the models in $SON_a$ and the value of model evaluation metrics (adjusted $R^2$, KGE, and $R$ Pearson) in Figure 5.8 showed the same tendency for all the metrics. The BS River, which has a large amount of data, like the rivers of Cu and Cs, could achieve a higher skill than the average, nearing the highest skills among all achieved by the rivers of Ma and Ba (eastern rivers). Meanwhile, Bo River, which has a similarly low number of observations as the Ma and Ba Rivers, had poor model skills. Pr River still had the lowest skills despite having more data than Ma and Ba Rivers.

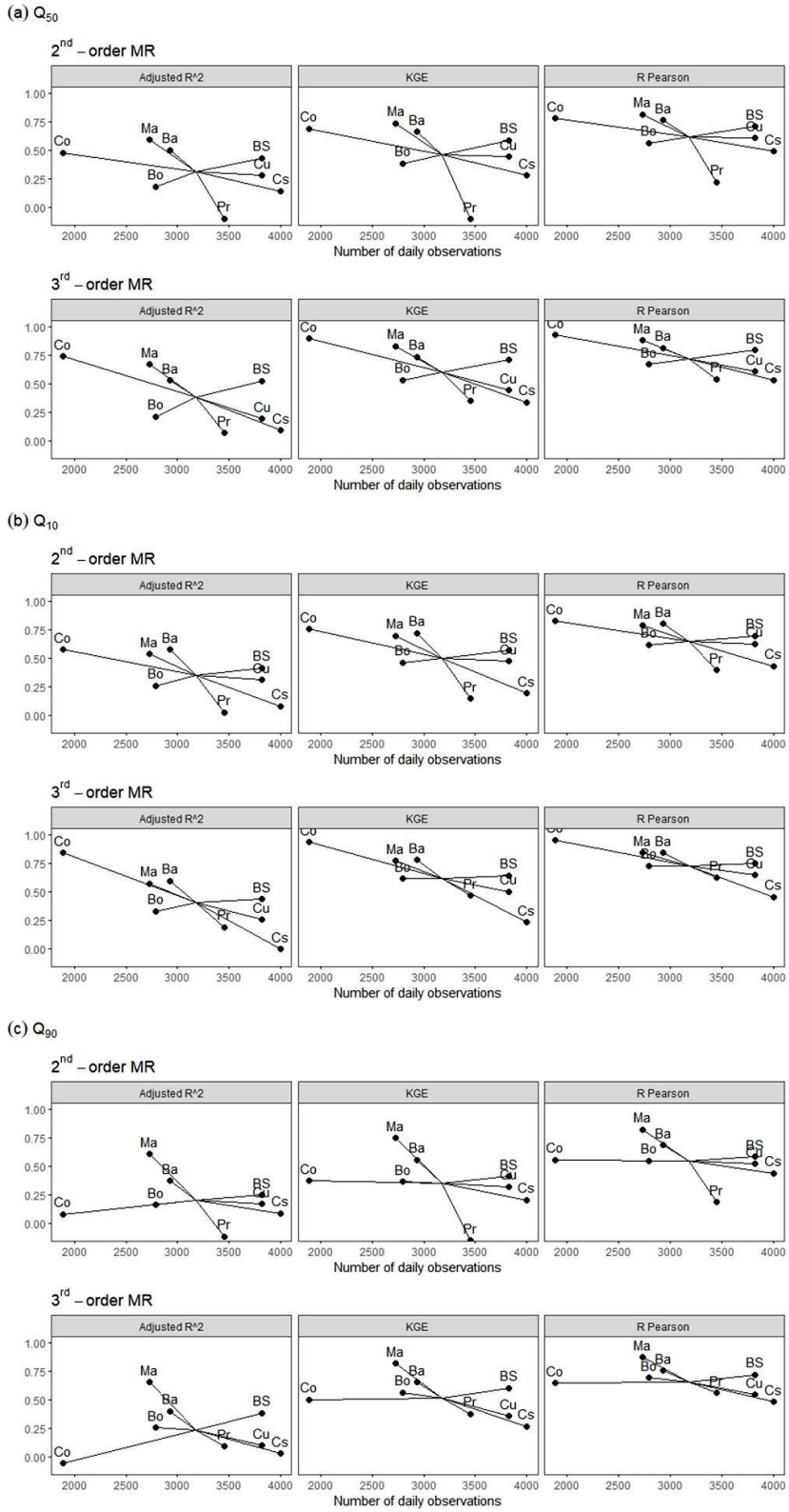


Figure 5.8 Plots between the number of daily observations (x-axis) and their model skills (y-axis) for developing MR models of (a) $Q_{50}$, (b) $Q_{10}$, and (c) $Q_{90}$ in the $SON_a$ set series. R Pearson indicates correlation between predicted and observed values.

### 5.4.2. Change of predictability over periods

To resolve the degree of freedom issue, the $SON_a$ data sets have been reanalyzed in two different periods: 1970–1989 and 1990–2018. This analysis may reduce the number of observations become half of the original. Moreover, the number of parameters in the multiple polynomial regressions could further reduce the degree of freedom. Therefore, the second-order MR was used in this reanalysis due to having fewer parameters than third-order MR.

Figure 5.9 shows the result of the models' evaluation for two periods. The results showed that the models in 1970–1989 (first period) generally have higher skills than those in 1990–2018 (second period). It is true for all flow regimes. Most rivers achieved good model skills in the first period, while in the second period, their model skills became poor mostly. Meanwhile, the model skills of the original period (1970–2018) are between the two periods. The complete adjusted $R^2$ and KGE values of the second- and third-order $SON_a$ MR models for the three periods are provided in Tables 5.6–5.9.

The scatterplots of predicted-observed rivers of Cu and Cs in Figure 5.9 clearly showed that the first period has better model skills. Cu and Cs are two examples of rivers in the lower model skills among the seven rivers in this study,

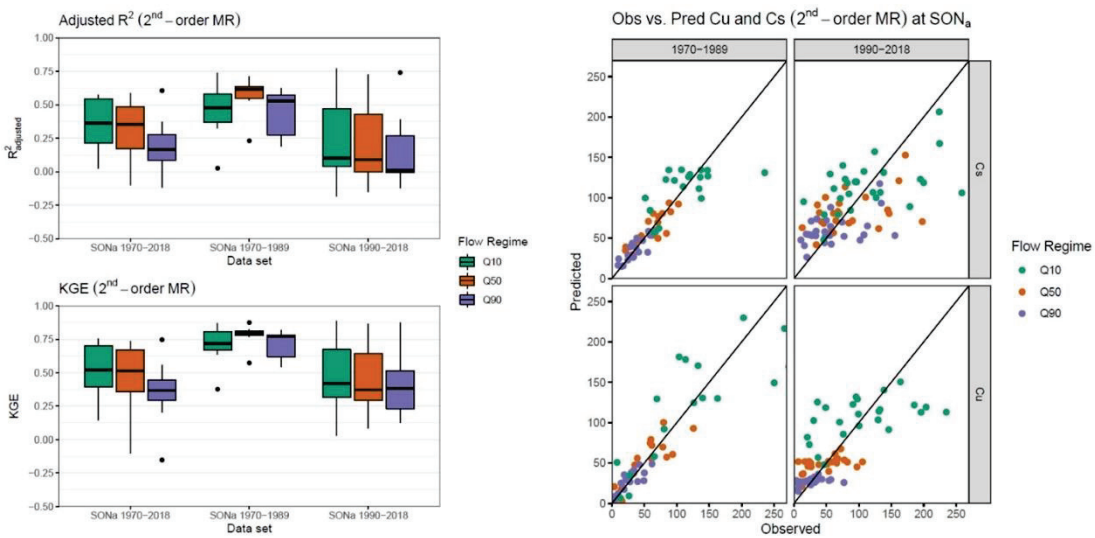


Figure 5.9 Evaluation metrics boxplots and predicted-observed scatterplots of MR models ($2^{nd}$-order, Niño 3.4-DMI predictors, $SON_a$ set) in three different periods: 1970–2018, 1970–1989, and 1990–2018. The boxplots show the metrics for all rivers, while the scatterplots only show the result from Cs and Cu Rivers.

placing them both in the lower parts of the boxplots, which shows the model skills among the rivers. The lower parts of the boxplots in Figure 5.9 showed that the first period achieved much better model skills than the second period, especially for the $Q_{50}$. Moreover, the model skill of the original period with all the data from both periods naturally became lower than in the first period and higher than in the second period. These results indicate that the rivers of Cu and Cs did not always have a poor relationship with the ENSO and IOD. Instead, it suggests that the relationship in the first period was changed in the second period.

Table 5.6 Adjusted $R^2$ of the 2nd-order MR models predicted by Niño 3.4 and DMI in the $SON_a$ set on three different periods. Values ≥ 0.5 are bolded, indicating a good model skill.

| River | 1970-2018 | | | 1970-1989 | | | 1990-2018 | | |
|---|---|---|---|---|---|---|---|---|---|
| | $Q_{50}$ | $Q_{10}$ | $Q_{90}$ | $Q_{50}$ | $Q_{10}$ | $Q_{90}$ | $Q_{50}$ | $Q_{10}$ | $Q_{90}$ |
| Cu | 0.277 | 0.309 | 0.170 | **0.618** | 0.478 | **0.568** | 0.090 | 0.100 | -0.001 |
| Cs | 0.142 | 0.078 | 0.085 | **0.615** | 0.027 | **0.579** | 0.029 | 0.055 | 0.011 |
| Bo | 0.181 | 0.260 | 0.162 | **0.533** | 0.411 | **0.626** | -0.037 | 0.022 | -0.009 |
| Pr | -0.101 | 0.025 | -0.117 | 0.231 | 0.329 | 0.189 | -0.149 | -0.184 | -0.122 |
| BS | 0.431 | 0.415 | 0.248 | **0.653** | **0.738** | 0.361 | 0.362 | 0.398 | 0.147 |
| Ma | **0.590** | **0.537** | **0.606** | **0.568** | **0.508** | 0.188 | **0.727** | **0.771** | **0.741** |
| Ba | **0.502** | **0.576** | 0.372 | **0.711** | **0.653** | 0.527 | **0.502** | **0.548** | 0.389 |
| Co | - | - | - | - | - | - | 0.479 | **0.575** | 0.078 |

Table 5.7 Adjusted $R^2$ of the 3rd-order models predicted by Niño 3.4 and DMI in the $SON_a$ set on three different periods. Values ≥ 0.5 are bolded, indicating a good model skill.

| River | 1970-2018 | | | 1970-1989 | | | 1990-2018 | | |
|---|---|---|---|---|---|---|---|---|---|
| | $Q_{50}$ | $Q_{10}$ | $Q_{90}$ | $Q_{50}$ | $Q_{10}$ | $Q_{90}$ | $Q_{50}$ | $Q_{10}$ | $Q_{90}$ |
| Cu | 0.192 | 0.257 | 0.099 | **0.563** | 0.294 | **0.505** | 0.160 | 0.284 | -0.048 |
| Cs | 0.093 | -0.002 | 0.030 | **0.564** | -0.326 | **0.560** | -0.178 | -0.148 | -0.216 |
| Bo | 0.212 | 0.328 | 0.255 | **0.525** | **0.580** | **0.645** | -0.222 | -0.297 | -0.060 |
| Pr | 0.065 | 0.188 | 0.091 | **0.540** | **0.648** | 0.318 | -0.189 | -0.253 | 0.040 |
| BS | **0.526** | 0.433 | 0.383 | **0.887** | **0.876** | **0.655** | 0.361 | 0.353 | 0.086 |
| Ma | **0.670** | **0.571** | **0.652** | **0.643** | 0.419 | -0.254 | **0.863** | **0.893** | **0.839** |
| Ba | **0.530** | **0.596** | 0.401 | **0.625** | **0.514** | 0.300 | **0.708** | **0.674** | **0.653** |
| Co | - | - | - | - | - | - | **0.741** | **0.841** | -0.058 |

Table 5.8 KGE of the 2nd-order MR models predicted by Niño 3.4 and DMI in the $SON_a$ set on three different periods. Values ≥ 0.5 are bolded, indicating a good model skill.

| River | 1970-2018 | | | 1970-1989 | | | 1990-2018 | | |
|---|---|---|---|---|---|---|---|---|---|
| | $Q_{50}$ | $Q_{10}$ | $Q_{90}$ | $Q_{50}$ | $Q_{10}$ | $Q_{90}$ | $Q_{50}$ | $Q_{10}$ | $Q_{90}$ |
| Cu | 0.441 | 0.323 | 0.473 | **0.800** | **0.771** | **0.718** | 0.334 | 0.230 | 0.345 |
| Cs | 0.281 | 0.205 | 0.194 | **0.792** | **0.771** | 0.377 | 0.254 | 0.232 | 0.284 |
| Bo | 0.383 | 0.362 | 0.462 | **0.769** | **0.818** | **0.701** | 0.371 | 0.395 | 0.420 |
| Pr | -0.105 | -0.152 | 0.146 | **0.573** | **0.544** | **0.637** | 0.085 | 0.125 | 0.028 |
| BS | **0.587** | 0.410 | **0.572** | **0.826** | **0.657** | **0.870** | **0.575** | 0.383 | **0.604** |
| Ma | **0.735** | **0.747** | **0.696** | **0.797** | **0.586** | **0.767** | **0.865** | **0.872** | **0.887** |
| Ba | **0.663** | **0.555** | **0.719** | **0.875** | **0.789** | **0.849** | **0.711** | **0.634** | **0.741** |
| Co | - | - | - | - | - | - | **0.689** | 0.371 | **0.753** |

Table 5.9 KGE of the 3rd-order MR models predicted by Niño 3.4 and DMI in the $SON_a$ set on three different periods. Values ≥ 0.5 are bolded, indicating a good model skill.

| River | 1970-2018 | | | 1970-1989 | | | 1990-2018 | | |
|---|---|---|---|---|---|---|---|---|---|
| | $Q_{50}$ | $Q_{10}$ | $Q_{90}$ | $Q_{50}$ | $Q_{10}$ | $Q_{90}$ | $Q_{50}$ | $Q_{10}$ | $Q_{90}$ |
| Cu | 0.441 | 0.356 | **0.502** | **0.858** | **0.837** | **0.761** | **0.560** | 0.417 | **0.637** |
| Cs | 0.337 | 0.268 | 0.230 | **0.847** | **0.845** | 0.453 | 0.287 | 0.251 | 0.314 |
| Bo | **0.533** | **0.564** | 0.615 | **0.875** | **0.907** | **0.890** | **0.597** | **0.659** | **0.567** |
| Pr | 0.351 | 0.375 | 0.464 | **0.863** | **0.792** | **0.897** | 0.387 | 0.536 | 0.339 |
| BS | **0.708** | **0.604** | **0.642** | **0.968** | **0.899** | **0.964** | **0.673** | 0.497 | **0.668** |
| Ma | **0.829** | **0.819** | **0.773** | **0.920** | **0.694** | **0.867** | **0.961** | **0.954** | **0.969** |
| Ba | **0.737** | **0.653** | **0.777** | **0.932** | **0.870** | **0.911** | **0.887** | **0.865** | **0.873** |
| Co | - | - | - | - | - | - | **0.895** | **0.500** | **0.937** |

The change of relationship in the western Java rivers was probably caused by the human perturbation to the river basin, such as land-use change and river development. The start of the river development era in Indonesia might be related to this phenomenon. River development means that rivers were physically changed—such as meander cutoff, riverbank narrowing, adding irrigation weir, and dam building—thus, it becomes more unnatural and changes the hydrological processes of streamflow. In Indonesia, these structural modifications to rivers started to grow rapidly around 1990, especially in the region of Java [73]. The construction of several dams could be identified around that time, such as the

Pongkor Dam on the Cs River (1996) [74] and the Wonogiri Dam on the BS River (1981) [75]. At that time, the purpose of the structural changes was mostly for agriculture use. Such physical changes may alter the hydrological processes in the watershed, thus giving a different relationship between climate and streamflow. Therefore, one reason the relationship between ENSO and IOD to the streamflow could be better predicted before 1990 might be because the rivers were still more natural than those in the next period (1990–2018).

### 5.4.3. Good predictability in Code River

In the first study, good predictability of the Code (Co) River streamflow was obtained by SOI-DMI predictors. In this study, the Code River data were reanalyzed further with the seasonal and the averaged seasonal sets, achieving higher model skills than the six-month moving average data set (MA6). The result of spatiotemporal analysis in this study indicates that Code River has lower seasonal variability than the other rivers. Code River's model skills are consistently at the maximum point in all rivers' evaluation metrics distribution (Figure 5.4), sometimes even being shown as an outlier.

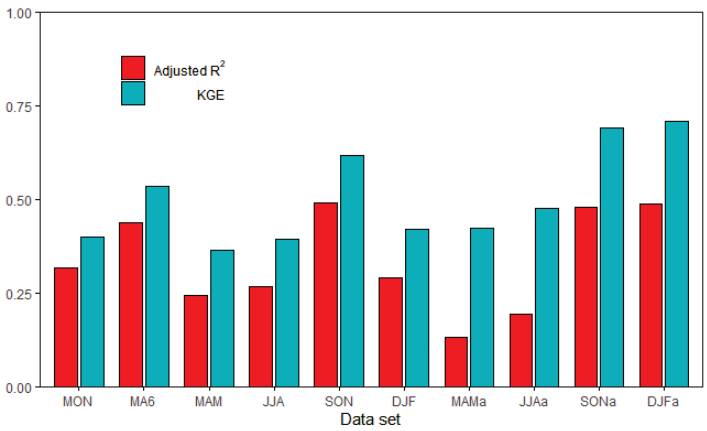


Figure 5.10 The adjusted $R^2$ and KGE of Code River's $Q_{50}$ 2nd-order MR models in different temporal data sets.

The intra-annual predictability of Code River is relatively low. It can be seen from the facts of the MA6 set's model skill. The MA6 data set represents intra-annual variation because it takes an average of overlapped six months. It has already been known that only the SON season gives a distinguished model skill among the

four seasons for all rivers. If the model skill of MA6 is high, then the predictability of the other three seasons is not too far different from the SON season. The Code River's predictabilities in $SON_a$ and $DJF_a$ are similarly high, while its predictabilities in $MAM_a$ and $JJA_a$ are about 0.25 lower (Figure 5.10). With this, it concludes that the excellent model skill obtained by Code River in the previous study seems to be owing to its seasonal characteristic. The Code River could achieve higher model skills among the eight rivers due to its lower intra-annual (month-to-month) streamflow variation than the other rivers.

### 5.4.4. Streamflow prediction by ENSO and IOD

The results of this second study showed that the ENSO and IOD teleconnection to flow regimes in Java rivers were only good in the SON season (the onset of the wet season [72]), with the averaged set ($SON_a$) being better than the monthly set (SON). The result of this temporal analysis emphasizes that intra-annual variation is crucial to be considered in studying the relationship between climate indices and the streamflow or rainfall in Indonesia. Despite Indonesia having only two seasons (wet and dry), intra-annual variation is still important to be considered owing to many interrelated climate phenomena around the regions in different periods and timescales.

The spatial variability analysis results suggest that the teleconnections for climate–rainfall and climate–streamflow are not necessarily similar. The physical processes in climate–streamflow teleconnection are more complex than in climate–rainfall teleconnection. In the climate–streamflow relationship, physical processes are mainly related to the hydrological processes, such as runoff generation and evaporation. It has been known that anthropogenic activities could directly affect the hydrological processes [76,77]. Therefore, different physical conditions of a watershed due to human perturbation could create a large variation in the climate–streamflow relationship, even though the rivers are in the same region. In this case, not only spatial variability but temporal variability could also be related to the hydrological changes of a river basin, as discussed in Section 5.4.2.

Good predictability of flow regimes by ENSO and IOD would be useful in forecasting the streamflow. This study did not forecast the streamflow by the two

climate indices. Instead, this study successfully elucidated the connection between the indices of two climate phenomena and the indices of streamflow in simultaneous seasons diagnostically. As good progress was shown in the prediction of ENSO [78] and IOD [79], now we can forecast them more accurately. If ENSO and IOD can be forecasted accurately, then the streamflow regimes could also be forecasted accurately using a good skilled teleconnection model. The knowledge of climate teleconnections to the streamflow is very useful for river managers, as the forecast of extreme high and low flow events could lead to better mitigation of water-related disasters.

However, this study found a difference in predictability skills between the high and low flows. Based on the study results, a high flow index ($Q_{10}$) and a medium flow index ($Q_{50}$) could be predicted better than a low flow index ($Q_{90}$). While the $Q_{10}$ predictability level seemed similar among rivers in Java, $Q_{90}$ was the opposite. The different predictability of $Q_{90}$ from one river to another is probably caused by the low flow dependency on the physical condition of a river's watershed. Anthropogenic drought problems, such as land-use change and groundwater exploitation, make $Q_{90}$ prediction by climate harder. Meanwhile, high flow situations are usually linear to the amount of rainfall (hence climate), making the prediction of $Q_{10}$ by climate better than $Q_{90}$.

### 5.5. Summary of the Second Study

This study developed the relationship between river flow regimes in Java and climate indices of ENSO and IOD by multiple regression (MR) models in ten different temporal data sets. Good predictability of the flow regimes was obtained using the ENSO and IOD indices in September–November (SON), especially the averaged SON set ($SON_a$). The high predictability in the SON season is consistent with previous studies on ENSO and IOD indices' connections to rainfall [52,53] and streamflow [23] in Indonesia. This study advanced the simple correlation analysis in the past studies to multiple polynomial regression and achieved new insights into the prediction of Indonesian streamflow by both ENSO and IOD indices.

Spatial variability analysis showed no clear distribution pattern of predictability over the Javanese rivers. The difference in model skills among the

eight rivers in Java does not seem to be due to regional distribution effects of ENSO and IOD; instead, the characteristics of each river are suggested to be the main cause. This conclusion comes after the comparison analysis of teleconnection between the two different periods, 1970–1989 and 1990–2018. While the previous analysis showed that the western rivers show very poor predictability, this comparison analysis showed a different insightful result. Interestingly, the first period bestowed better prediction skills than the second period. This is most likely due to physical changes caused by river development, thereby weakening the climate–streamflow relationship over the years.

The difference between climate–rainfall and climate–streamflow relationships may be seen from their spatial variabilities caused by various occurrences in river basins. The hydrological processes in each river basin often varied significantly, caused mainly by anthropogenic disturbance. This disturbance could also be detected in the temporal variations by comparing the predictability of different periods. In this study, the different predictability skills were found in the 1970–1989 and the 1990–2018 periods. Therefore, it can be assumed that the difference was due to the relationship change before and after the disturbance.

The Code River, the main feature of the first study, was included in this second study to compare and better understand the teleconnection of rivers in the Java region. The Code River achieved good model skills in the first study (with a six-month moving average set) and this second study (with seasonal data sets). However, the other seven rivers in this study showed varied quality in the model skills. It was caused by the lower intra-annual variation and higher inter-annual variation that the Code River had compared to the other rivers. Therefore, the results of the first study was a fortunate finding of a good river flow regime predictability.

This study contributes to the new understanding of ENSO and IOD's relationship to the streamflow in Indonesia. This study showed new insights into the relationship using flow regimes and spatiotemporal analysis. Varied predictability skills were found among the three flow regimes and the eight rivers of Java. The predictability of the high flow regime tends to be higher than that of the low flow regime over many rivers in Java. From the eight rivers, it was found that the rivers of Ma, Ba, and Co could be better predicted than the rest of the rivers.

# Chapter 6
# Conclusion

## 6.1. Summary of the Two Studies

The studies of this thesis were conducted using multiple polynomial regression models (MR) to predict high ($Q_{10}$), average ($Q_{50}$), and low ($Q_{90}$) streamflows by climate indices of ENSO and IOD. By gathering eight rivers' data sets of the period between 1970 to 2018 on Java Island, thorough discussions were established in this thesis for developing a good quantitative climate–streamflow model to face the limited availability of Indonesian data. These studies are the first to develop a quantitative statistical model for determining the streamflow values using climate indices in Indonesia.

The first study found good model skills in Code River, especially when using 3- and 6-month moving average data sets. The moving average data sets could increase the accuracy of explaining the effect of ENSO and IOD because it can avoid the effect of Madden Julian Oscillation (MJO) or other short-term phenomena. The time-lagged MR models have relatively good skills with a high Pearson correlation coefficient (mostly higher than 0.7) and moderate Root Mean Square Error (44–51%). Unfortunately, a model validation could not be established due to the small number of data available.

The good model skill found in Code River could not be reproduced using the monthly data set for other Javanese rivers. In the second study, temporal variability is investigated by arranging the data set of eight Javanese rivers into four three-monthly sets (March-May MAM, June-August JJA, September-November SON, and December-February DJF). It was then found with −0.10–0.74 (0.34–0.89) KGE for the second-order (third-order) $Q_{50}$ MR in the $SON_a$ set. While DJF in the monthly set suffers from very low model skill, along with JJA and MAM, DJF in the three-month averaged set ($DJF_a$) gives only a bit lower skill than $SON_a$ set, making it significantly more reliable when the monthly scale effect is removed. However, even in the three-month averaged set, JJA and MAM still do not have

good skills, with MAM being the lowest with −0.21–0.42 (−0.03–0.46) KGE for the second-order (third-order) $Q_{50}$.

Further temporal investigation showed a tendency that the predictability of the past period (1970-1989) was better than those in the more recent period (1990-2018). This tendency is the same for all rivers, suggesting that the relationships between streamflow and climates were changed over periods. Hence, even though the models from rivers in eastern Java are found to have higher predictability skills, there is no clear spatial distribution pattern of the predictability over the studied eight rivers in Java.

The followings are the summaries of the main findings from the first and second studies. During the first study using Code River data, it is found that:

a.) high and significant correlations of ENSO to the flow regime indices were found in the 6-month moving average data set, while no effect of IOD on the flow regime indices can be identified in this simple correlation analysis,

b.) the modulation of the ENSO–streamflow relationship by IOD is identified as a significant increase of high flow in the La Niña phase,

c.) the third-order MR generally had better skill than the second-order MR, but their estimations were not significantly different,

d.) 6-month leading climate indices in MR models achieved a relatively good model estimation ($R$ of estimated-observed > 0.70) with moderate RMSE (44–51%) and adjusted $R^2$ (0.47–0.59),

e.) the average flow regime index $Q_{50}$ generally had better model skills than the flow regime indices of high ($Q_{10}$) and low ($Q_{90}$) flows.

During the second study using eight Javanese rivers data, it is found that:

a.) good predictability skills were found in the September–November (SON) season, in which the seasonal average set had better skills than the monthly set,

b.) it was found that the predictability in the 1970-1989 period tended to be higher than that of the 1990-2018 period in seven Javanese rivers, suggesting that it is caused by physical changes in the watershed over the period,

c.) there was no clear spatial distribution pattern of predictability found over the eight Javanese rivers studied,

d.) Code River has an exceptional flow regimes predictability among the studied eight rivers in Java due to its low seasonal variability,
e.) high flow predictability tended to be higher than the low flow predictability,
f.) the highest predictability skill among the eight rivers studied is achieved by the Madiun River (East Java), Baru River (East Java), and Code River (Central Java).

## 6.2. Final Remarks

This thesis showed that the predictability of streamflow in Java, Indonesia, seems to have a similar pattern to the previously known teleconnection studies in Indonesia. Previous studies have found that in September–November (SON), ENSO and IOD have a high correlation to rainfall [10,11,15,52,53,72] and streamflow [23]. It has been known that IOD begins to form in JJA and peaks in SON [14,36]. Therefore, the result of this thesis confirmed that the predictability of streamflow regimes in Java is highly dependent on the peak of IOD. Additionally, ENSO begins in JJA and peaks in DJF [80]. It may explain why DJF has the second-highest model skill among the four seasons during the second study of this thesis.

This thesis has elucidated the relationship between the Javanese river flow regimes and the climate indices of ENSO and IOD using quantitative statistical models. Compared to the previous studies, which provide more qualitative results, the studies of this thesis could provide quantitative results in the sense of streamflow values. Using the MR equation, the values for high flow ($Q_{10}$), median flow ($Q_{50}$), and low flow ($Q_{90}$) can be reported quickly. With the quantitative results, the climate-streamflow MR models can get more direct access to environmental or disaster mitigation studies.

## 6.3. Future Studies

This thesis successfully confirmed the climate-streamflow relationship using quantitative statistical models. Having thoroughly elucidated the behaviors of the models, the results of this thesis' studies can give new insight into Indonesian teleconnection in future works.

Future studies can be aimed toward forecasting the streamflow regimes, as it will be useful for river managers. A forecasting investigation could be done by looking more into the lead-lag relationship between the climate indices and the streamflow regimes.

Other works for the future should be on validating and improving the predictability skill of the teleconnection model. The improvement of the models could be attempted by some approaches, such as:

a. To synchronize the climate-streamflow models to the climate forecast. ENSO and IOD forecast nowadays has been much improved. Therefore, it is necessary to investigate further how much the climate–streamflow model can predict with the current skill of the forecast of ENSO and IOD.
b. To improve the predictability of a specific flow index. In order to increase the accessibility of the climate-streamflow model to environmental studies, it would be preferable to improve the poor predictability of the low flow index. Some advanced transforming, filtering, or modeling techniques could be tested to improve predictability. For example: transforming the low flow observation data into a logarithmic data set.
c. To add other climate phenomena into the model. Other climate phenomena, such as Madden–Julian Oscillation (MJO) and ENSO Modoki, could add more dynamics to the climate-streamflow relationship. By incorporating another climate phenomenon into consideration, whether as a predictor or as part of a method to filter the data sets, the climate–streamflow relationships can be elucidated more, and a better model for streamflow management can be expected.

# List of Publications

The published studies which are the cores of this doctoral thesis are:

1. A. R. Nugroho, I. Tamagawa, M. Harada. 2021. The Relationship between River Flow Regimes and Climate Indices of ENSO and IOD on Code River, Southern Indonesia. *Water*, 13(10), 1375.
2. A. R. Nugroho, I. Tamagawa, M. Harada. 2022. Spatiotemporal Analysis on the Teleconnection of ENSO and IOD to the Stream Flow Regimes in Java, Indonesia. *Water*, 14(2), 168.

# References


1. Sofi, M.S.; Bhat, S.U.; Rashid, I.; Kuniyal, J.C. The natural flow regime: A master variable for maintaining river ecosystem health. *Ecohydrology* **2020**, *13*, 1–12, doi:10.1002/eco.2247.

2. Poff, N.L. Beyond the natural flow regime? Broadening the hydro-ecological foundation to meet environmental flows challenges in a non-stationary world. *Freshw. Biol.* **2018**, *63*, 1011–1021, doi:10.1111/fwb.13038.

3. Nigam, S.; Baxter, S. GENERAL CIRCULATION OF THE ATMOSPHERE | Teleconnections. In *Encyclopedia of Atmospheric Sciences*; North, G.R., Pyle, J., Zhang, F., Eds.; Elsevier, 2015; Vol. 3, pp. 90–109.

4. Chiew, F.H.S.; McMahon, T.A. Global ENSO-streamflow teleconnection, streamflow forecasting and interannual variability. *Hydrol. Sci. J.* **2002**, *47*, 505–522, doi:10.1080/02626660209492950.

5. Kingston, D.G.; McGregor, G.R.; Hannah, D.M.; Lawler, D.M. River flow teleconnections across the northern North Atlantic region. *Geophys. Res. Lett.* **2006**, *33*, 1–5, doi:10.1029/2006GL026574.

6. Degefu, M.A.; Bewket, W. Variability, trends, and teleconnections of stream flows with large-scale climate signals in the Omo-Ghibe River Basin, Ethiopia. *Environ. Monit. Assess.* **2017**, *189*, doi:10.1007/s10661-017-5862-1.

7. Canchala, T.; Cerón, W.L.; Francés, F.; Carvajal-Escobar, Y.; Andreoli, R.V.; Kayano, M.T.; Alfonso-Morales, W.; Caicedo-Bravo, E.; de Souza, R.A.F. Streamflow variability in colombian pacific basins and their teleconnections with climate indices. *Water (Switzerland)* **2020**, *12*, doi:10.3390/w12020526.

8. Nicholls, N. TROPICAL METEOROLOGY & CLIMATE | El Niño and the Southern Oscillation: Observation. In *Encyclopedia of Atmospheric Sciences*; North, G.R., Pyle, J., Zhang, F., Eds.; Elsevier, 2015; Vol. 6, pp. 91–96 ISBN 9780123822260.

9. Ropelewski, C.F.; Halpert, M.S. Global and Regional Scale Precipitation Patterns Associated with the El Niño/Southern Oscillation. *Mon. Weather Rev.* **1987**, *115*, 1606–1626, doi:10.1175/1520-0493(1987)115<1606:GARSPP>2.0.CO;2.

10. Hendon, H.H. Indonesian rainfall variability: Impacts of ENSO and local air-sea interaction. *J. Clim.* **2003**, *16*, 1775–1790, doi:10.1175/1520-0442(2003)016<1775:IRVIOE>2.0.CO;2.

11. Aldrian, E.; Susanto, R.D. Identification of three dominant rainfall regions within Indonesia and their relationship to sea surface temperature. *Int. J. Climatol.* **2003**, *23*, 1435–1452, doi:10.1002/joc.950.

12. Lee, H.S. General Rainfall Patterns in Indonesia and the Potential Impacts of Local Seas on Rainfall Intensity. *Water* **2015**, *7*, 1751–1768, doi:10.3390/w7041751.

13. Aldrian, E.; Dümenil Gates, L.; Widodo, F.H. Seasonal variability of Indonesian rainfall in ECHAM4 simulations and in the reanalyses: The role of ENSO. *Theor. Appl. Climatol.* **2007**, *87*, 41–59, doi:10.1007/s00704-006-0218-8.

14. Saji, N.; Goswami, B.; Vinayachandran, P.; Yamagata, T. A dipole mode in the Tropical Ocean. *Nature* **1999**, *401*, 360–363.

15. Jun-Ichi, H.; Mori, S.; Kubota, H.; Yamanaka, M.D.; Haryoko, U.; Lestari, S.; Sulistyowati, R.; Syamsudin, F. Interannual rainfall variability over northwestern Jawa and its relation to the Indian Ocean Dipole and El Niño-Southern Oscillation events. *Sci. Online Lett. Atmos.* **2012**, *8*, 69–72, doi:10.2151/sola.2012-018.

16. Kahya, E.; Dracup, J.A. U.S. streamflow patterns in relation to the El Niño/Southern Oscillation. *Water Resour. Res.* **1993**, *29*, 2491–2503, doi:10.1029/93WR00744.

17. Cluis, D.; Laberge, C. Analysis of the El Niño Effect on the Discharge of Selected Rivers in the Asia-Pacific Region. *Water Int.* **2002**, *27*, 279–293, doi:10.1080/02508060208687002.

18. Cardoso, A.. O.; Silva Dias, P.L. The relationship between ENSO and Paraná River flow. *Adv. Geosci.* **2006**, *6*, 189–193, doi:10.5194/adgeo-6-189-2006.

19. Chandimala, J.; Zubair, L. Predictability of stream flow and rainfall based on ENSO for water resources management in Sri Lanka. *J. Hydrol.* **2007**, *335*, 303–312, doi:10.1016/j.jhydrol.2006.11.024.

20. Zhang, Q.; Xu, C. yu; Jiang, T.; Wu, Y. Possible influence of ENSO on annual maximum streamflow of the Yangtze River, China. *J. Hydrol.* **2007**, *333*, 265–274, doi:10.1016/j.jhydrol.2006.08.010.

21. Mohsenipour, M.; Shahid, S.; Nazemosadat, M.J. Effects of El Nino Southern Oscillation on the Discharge of Kor River in Iran. *Adv. Meteorol.* **2013**, *2013*, 1–7, doi:10.1155/2013/846397.

22. Sahu, N.; Panda, A.; Nayak, S.; Saini, A.; Mishra, M.; Sayama, T.; Sahu, L.; Duan, W.; Avtar, R.; Behera, S. Impact of indo-pacific climate variability on high streamflow events in Mahanadi River Basin, India. *Water (Switzerland)* **2020**, *12*, doi:10.3390/w12071952.

23. Sahu, N.; Behera, S.K.; Yamashiki, Y.; Takara, K.; Yamagata, T. IOD and

ENSO impacts on the extreme stream-flows of Citarum River in Indonesia. *Clim. Dyn.* **2012**, *39*, 1673–1680, doi:10.1007/s00382-011-1158-2.

24. Trenberth, K.E. The Definition of El Niño. *Bull. Am. Meteorol. Soc.* **1997**, *78*, 2771–2777, doi:10.1175/1520-0477(1997)078<2771:TDOENO>2.0.CO;2.

25. Chang, P.; Zebiak, S.E. TROPICAL METEOROLOGY & CLIMATE | El Niño and the Southern Oscillation. In *Encyclopedia of Atmospheric Sciences*; North, G.R., Pyle, J., Zhang, F., Eds.; Elsevier, 2015; Vol. 6, pp. 97–101 ISBN 9780123822260.

26. Barnston, A. Why are there so many ENSO indexes, instead of just one? Available online: https://www.climate.gov/news-features/blogs/enso/why-are-there-so-many-enso-indexes-instead-just-one (accessed on Mar 22, 2022).

27. Trenberth, K.E. Signal Versus Noise in the Southern Oscillation. *Mon. Weather Rev.* **1984**, *112*, 326–332, doi:10.1175/1520-0493(1984)112<0326:SVNITS>2.0.CO;2.

28. Barnston, A.G.; Chelliah, M.; Goldenberg, S.B. Documentation of a Highly ENSO-Related SST Region in the Equatorial Pacific: Research Note. *Atmos. - Ocean* **1997**, *35*, 367–383, doi:10.1080/07055900.1997.9649597.

29. Hendrawan, I.G.; Asai, K.; Triwahyuni, A.; Lestari, D.V. The interanual rainfall variability in Indonesia corresponding to El Niño Southern Oscillation and Indian Ocean Dipole. *Acta Oceanol. Sin.* **2019**, *38*, 57–66, doi:10.1007/s13131-019-1457-1.

30. Rao, J.; Ren, R. Parallel comparison of the 1982/83, 1997/98 and 2015/16 super El Niños and their effects on the extratropical stratosphere. *Adv. Atmos. Sci.* **2017**, *34*, 1121–1133, doi:10.1007/s00376-017-6260-x.

31. Qian, H.; Xu, S.-B. Prediction of Autumn Precipitation over the Middle and Lower Reaches of the Yangtze River Basin Based on Climate Indices. *Climate* **2020**, *8*, 53, doi:10.3390/cli8040053.

32. Cai, W.; McPhaden, M.J.; Grimm, A.M.; Rodrigues, R.R.; Taschetto, A.S.; Garreaud, R.D.; Dewitte, B.; Poveda, G.; Ham, Y.G.; Santoso, A.; et al. Climate impacts of the El Niño–Southern Oscillation on South America. *Nat. Rev. Earth Environ.* **2020**, *1*, 215–231, doi:10.1038/s43017-020-0040-3.

33. Brönnimann, S.; Xoplaki, E.; Casty, C.; Pauling, A.; Luterbacher, J. ENSO influence on Europe during the last centuries. *Clim. Dyn.* **2007**, *28*, 181–197, doi:10.1007/s00382-006-0175-z.

34. Shaman, J. The seasonal effects of ENSO on European precipitation: Observational analysis. *J. Clim.* **2014**, *27*, 6423–6438, doi:10.1175/JCLI-D-14-00008.1.

35. King, M.P.; Yu, E.; Sillmann, J. Impact of strong and extreme El Niños on European hydroclimate. *Tellus, Ser. A Dyn. Meteorol. Oceanogr.* **2020**, *72*, 1–10, doi:10.1080/16000870.2019.1704342.

36. Webster, P.J.; Moore, A.M.; Loschnigg, J.P.; Leben, R.R. Coupled oceanic-atmoshperic dynamics in the Indian Ocean during 1997-1998. *Nature* **1999**, *401*, 356–360.

37. Behera, S.K.; Krishnan, R.; Yamagata, T. Unusual ocean-atmosphere conditions in the tropical Indian Ocean during 1994. *Geophys. Res. Lett.* **1999**, *26*, 3001–3004, doi:10.1029/1999GL010434.

38. Black, E.; Slingo, J.; Sperber, K.R. An observational study of the relationship between excessively strong short rains in coastal East Africa and Indian ocean SST. *Mon. Weather Rev.* **2003**, *131*, 74–94, doi:10.1175/1520-0493(2003)131<0074:AOSOTR>2.0.CO;2.

39. Ashok, K.; Guan, Z.; Yamagata, T. A look at the relationship between the ENSO and the Indian Ocean Dipole. *J. Meteorol. Soc. Japan* **2003**, *81*, 41–56, doi:10.2151/jmsj.81.41.

40. Saji, N.H.; Yamagata, T. Possible impacts of Indian Ocean Dipole mode events on global climate. *Clim. Res.* **2003**, *25*, 151–169, doi:10.3354/cr025151.

41. Xie, S.P.; Annamalai, H.; Schott, F.A.; McCreary, J.P. Structure and mechanisms of South Indian Ocean climate variability. *J. Clim.* **2002**, *15*, 864–878, doi:10.1175/1520-0442(2002)015<0864:SAMOSI>2.0.CO;2.

42. Lau, N.C.; Nath, M.J. Atmosphere-ocean variations in the Indo-Pacific sector during ENSO episodes. *J. Clim.* **2003**, *16*, 3–20, doi:10.1175/1520-0442(2003)016<0003:AOVITI>2.0.CO;2.

43. Behera, S.K.; Luo, J.J.; Masson, S.; Rao, S.A.; Sakuma, H.; Yamagata, T. A CGCM study on the interaction between IOD and ENSO. *J. Clim.* **2006**, *19*, 1688–1705, doi:10.1175/JCLI3797.1.

44. Kug, J.S.; Li, T.; An, S. Il; Kang, I.S.; Luo, J.J.; Masson, S.; Yamagata, T. Role of the ENSO-Indian Ocean coupling on ENSO variability in a coupled GCM. *Geophys. Res. Lett.* **2006**, *33*, doi:10.1029/2005GL024916.

45. Luo, J.J.; Zhang, R.; Behera, S.K.; Masumoto, Y.; Jin, F.F.; Lukas, R.; Yamagata, T. Interaction between El Niño and extreme Indian Ocean dipole. *J. Clim.* **2010**, *23*, 726–742, doi:10.1175/2009JCLI3104.1.

46. Zhou, Q.; Duan, W.; Mu, M.; Feng, R. Influence of positive and negative Indian Ocean Dipoles on ENSO via the Indonesian Throughflow: Results from sensitivity experiments. *Adv. Atmos. Sci.* **2015**, *32*, 783–793, doi:10.1007/s00376-014-4141-0.

47. Kajtar, J.B.; Santoso, A.; England, M.H.; Cai, W. Tropical climate variability: interactions across the Pacific, Indian, and Atlantic Oceans.

*Clim. Dyn.* **2017**, *48*, 2173–2190, doi:10.1007/s00382-016-3199-z.

48. Stuecker, M.F.; Timmermann, A.; Jin, F.F.; Chikamoto, Y.; Zhang, W.; Wittenberg, A.T.; Widiasih, E.; Zhao, S. Revisiting ENSO/Indian Ocean Dipole phase relationships. *Geophys. Res. Lett.* **2017**, *44*, 2481–2492, doi:10.1002/2016GL072308.

49. Yuan, D.; Hu, X.; Xu, P.; Zhao, X.; Masumoto, Y.; Han, W. The IOD-ENSO precursory teleconnection over the tropical Indo-Pacific Ocean: dynamics and long-term trends under global warming. *J. Oceanol. Limnol.* **2018**, *36*, 4–19, doi:10.1007/s00343-018-6252-4.

50. Bureau of Meteorology Record-breaking La Niña events. *Bur. Meteorol.* **2012**, 26.

51. Ashok, K.; Guan, Z.; Yamagata, T. Influence of the Indian Ocean Dipole on the Australian winter rainfall. *Geophys. Res. Lett.* **2003**, *30*, 3–6, doi:10.1029/2003GL017926.

52. Hidayat, R.; Ando, K.; Masumoto, Y.; Luo, J.J. Interannual Variability of Rainfall over Indonesia: Impacts of ENSO and IOD and Their Predictability. *IOP Conf. Ser. Earth Environ. Sci.* **2016**, *31*, 012043, doi:10.1088/1755-1315/31/1/012043.

53. Kurniadi, A.; Weller, E.; Min, S.K.; Seong, M.G. Independent ENSO and IOD impacts on rainfall extremes over Indonesia. *Int. J. Climatol.* **2021**, *41*, 3640–3656, doi:10.1002/joc.7040.

54. Pardoe, I. *Applied Regression Modelling*; Wiley: Hoboken, New Jersey, 2012; ISBN 978-1-1180-9728-1.

55. Nash, J.E.; Sutcliffe, J.V. River flow forecasting through conceptual models part I — A discussion of principles. *J. Hydrol.* **1970**, *10*, 282–290, doi:10.1016/0022-1694(70)90255-6.

56. Gupta, H. V.; Kling, H.; Yilmaz, K.K.; Martinez, G.F. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *J. Hydrol.* **2009**, *377*, 80–91, doi:10.1016/j.jhydrol.2009.08.003.

57. Gupta, H.V.; Kling, H. On typical range, sensitivity, and normalization of Mean Squared Error and Nash-Sutcliffe Efficiency type metrics. *Water Resour. Res.* **2011**, *47*, 2–4, doi:10.1029/2011WR010962.

58. Knoben, W.J.M.; Freer, J.E.; Woods, R.A. Technical note: Inherent benchmark or not? Comparing Nash-Sutcliffe and Kling-Gupta efficiency scores. *Hydrol. Earth Syst. Sci.* **2019**, *23*, 4323–4331, doi:10.5194/hess-23-4323-2019.

59. Kling, H.; Fuchs, M.; Paulin, M. Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios. *J. Hydrol.* **2012**, *424–425*, 264–277, doi:10.1016/j.jhydrol.2012.01.011.

60. Soemardiono, B.; Gusma, A.F. The Development of Code River Area in Yogyakarta as a Sustainable Urban Landscape Asset acknowledging Local Traditional Knowledge. *Int. Rev. Spat. Plan. Sustain. Dev.* **2014**, *2*, 4–18, doi:10.14246/irspsd.2.4_4.

61. Smith, P.J.; Krishnamurti, T.N.; Gentili, J. Malaysian-Australian monsoon Available online: https://www.britannica.com/science/Malaysian-Australian-monsoon.

62. Madden, R.A.; Julian, P.R. Detection of a 40–50 Day Oscillation in the Zonal Wind in the Tropical Pacific. *J. Atmos. Sci.* **1971**, *28*, 702–708, doi:10.1175/1520-0469(1971)028<0702:DOADOI>2.0.CO;2.

63. Muhammad, F.R.; Lubis, S.W.; Setiawan, S. Impacts of the Madden–Julian oscillation on precipitation extremes in Indonesia. *Int. J. Climatol.* **2021**, *41*, 1970–1984, doi:10.1002/joc.6941.

64. Gordon, N.D.; McMahon, B.L.; Gippel, C.J.; Nathan, R.J. *Stream Hydrology: An Introduction for Ecologists*; 2004; Vol. 2; ISBN 0470843578.

65. R Core Team R: A Language and Environment for Statistical Computing 2021.

66. Wickham, H. modelr: Modelling Functions that Work with the Pipe. 2020.

67. Zambrano-Bigiarini, M. hydroGOF: Goodness-of-fit functions for comparison of simulated and observed hydrological time series 2020.

68. Searcy, J.K. *Flow-Duration Curves*; United States Government Printing Office: Washington, 1959;

69. Goetz, J.; Schwarz, C.J. fasstr: Analyze, Summarize, and Visualize Daily Streamflow Data 2021.

70. Ropelewski, C.F.; Halpert, M.S. Precipitation Patterns Associated with the High Index Phase of the Southern Oscillation. *J. Clim.* **1989**, *2*, 268–284, doi:10.1175/1520-0442(1989)002<0268:PPAWTH>2.0.CO;2.

71. Hidayat, R.; Kizu, S. Influence of the Madden-Julian Oscillation on Indonesian rainfall variability in austral summer. *Int. J. Climatol.* **2010**, *30*, 1816–1825, doi:10.1002/joc.2005.

72. Hamada, J.I.; Yamanaka, M.D.; Matsumoto, J.; Fukao, S.; Winarso, P.A.; Sribimawati, T. Spatial and temporal variations of the rainy season over Indonesia and their link to ENSO. *J. Meteorol. Soc. Japan* **2002**, *80*, 285–310, doi:10.2151/jmsj.80.285.

73. Maryono, A. *River Restoration*; 4th ed.; Gadjah Mada University Press: Yogyakarta, 2020; ISBN 978-979-420-667-6.

74. Kali Cisadane. In *Catalogue of Rivers for Southeast Asia and the Pacific-Volume V*; Tachikawa, Y., James, R., Abdullah, K., Desa, M.N. bin M.,

Eds.; UNESCO-IHP, 2004; pp. 45–56 ISBN 4-902712-00-8.

75. Bengawan Solo. In *Catalogue of Rivers for Southeast Asia and the Pacific-Volume I*; Takeuchi, K., Jayawardena, A.W., Takahasi, Y., Eds.; UNESCO-IHP, 1995; pp. 91–102.

76. Gornitz, V.; Rosenzweig, C.; Hillel, D. Effects of anthropogenic intervention in the land hydrologic cycle on global sea level rise. *Glob. Planet. Change* **1997**, *14*, 147–161, doi:10.1016/S0921-8181(96)00008-2.

77. Vorosmarty, C.J.; Sahagian, D. Anthropogenic disturbance of the terrestrial water cycle. *Bioscience* **2000**, *50*, 753–765, doi:10.1641/0006-3568(2000)050[0753:ADOTTW]2.0.CO;2.

78. Tang, Y.; Zhang, R.H.; Liu, T.; Duan, W.; Yang, D.; Zheng, F.; Ren, H.; Lian, T.; Gao, C.; Chen, D.; et al. Progress in ENSO prediction and predictability study. *Natl. Sci. Rev.* **2018**, *5*, 826–839, doi:10.1093/nsr/nwy105.

79. Feba, F.; Ashok, K.; Collins, M.; Shetye, S.R. Emerging Skill in Multi-Year Prediction of the Indian Ocean Dipole. *Front. Clim.* **2021**, *3*, 1–8, doi:10.3389/fclim.2021.736759.

80. Wang, C.; Fiedler, P.C. ENSO variability and the eastern tropical Pacific: A review. *Prog. Oceanogr.* **2006**, *69*, 239–266, doi:10.1016/j.pocean.2006.03.004.