Development of Data and Modeling Tools for Understanding and Forecasting Indonesian Hydroclimate

by

Yanto

B.Eng., Gadjah Mada University, Yogyakarta, Indonesia, 2003
M.S.E., University of Michigan, Ann Arbor, 2011

A thesis submitted to the
Faculty of the Graduate School of the
University of Colorado in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Department of Civil, Environmental and Architectural Engineering

2016
This thesis entitled:
Development of Data and Modeling Tools for Understanding and Forecasting Indonesian Hydroclimate
written by Yanto
has been approved for the Department of Civil, Environmental and Architectural Engineering

Prof. Balaji Rajagopalan

Prof. Edith Zagona

Prof. Ben Livneh

Prof. Joseph Kasprzyk

Dr. Andrew Wood

Date __________________

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.
Indonesia routinely suffers from floods, droughts, landslides and water-borne diseases resultant from significant spatial and temporal variability of precipitation. Mitigation of these hazards and efficient management of resources require tools for understanding and forecasting Indonesian hydroclimate. To date, research efforts aimed at understanding this hydroclimate variability are few and there exists no robust hydrologic modeling tool needed to understand and forecast important variability. Motivated by these gaps and needs, this dissertation makes four unique contributions: (i) A systematic space-time analysis of seasonal precipitation over Indonesia was performed using Principal Component Analysis and Bayesian Dynamic Linear Model. El Niño Southern Oscillation (ENSO) was found to be the driver of leading modes of variability during both the wet (Oct - Mar) and dry (Apr - Sep) seasons. Furthermore, ENSO appeared to drive variability at multi-decadal timescales (8 - 16 year) especially during post 1980 period. The association between ENSO and Indonesian rainfall has strengthened in recent decades, especially during dry season. These findings suggest potential for interannual and multidecadal predictability of Indonesian rainfall. (ii) To understand the processes that drive the hydrologic variability, we built and calibrated a distributed hydrologic model, the Variable Infiltration Capacity (VIC) model for six watersheds over Java, the most populous island of Indonesia. In light of data scarcity and quality issues, model skill scores during calibration period were quite good but comparatively lower skills during validation. The magnitude and variability of baseflow and direct runoff components were found to modulate the model performance. We provided preliminary evidence that performance could be improved by refining the spatial resolution of model and input precipitation and temperature. (iii) Following this, a high resolution gridded daily meteorology - precipitation, maximum and minimum temperature -
data set at (∼14 km) resolution spanning 30 years from 1985 - 2014 was developed over Java, using an Inverse Distance Weighting method. The data set was stored in a network common data format (NetCDF) and will be publicly available, intended to support basin-scale and island-scale studies of short-term and long-term climate, hydrology and ecology. (iv) In order to provide streamflow forecasts that capture model parametric uncertainty, ensembles of model parameters are necessary. For this, we conducted a multi-objective optimization based calibration of VIC model parameters. The method generated an ensemble of model parameters optimizing six objective functions that capture key aspects of the hydrograph. This improved upon the single objective based calibration. We demonstrated the utility of this in generating skillful seasonal hydrologic forecasts conditioned on seasonal climate forecasts.

The above contributions, especially the watershed modeling tools, are unique and is a first of its kind research efforts in this region. We offered insights into the space-time variability of precipitation and a robust physically-based watershed modeling tool to understand and forecast hydrologic variability over Java in particular and Indonesia in general. Together, this research makes significant strides in providing a framework for understanding, modeling and forecasting Indonesian hydrology and climatology, that will help mitigate natural hazards and enable efficient management of Indonesia’s water and natural resources.
Dedication

To my beloved Dad in heaven for his sincere love and hard work, to my beloved Mom for her truthful love and endless climbed prayer, to my beloved wife for her heartfelt love, understanding, support and patience, to my four kids for their laughter and cheerfulness, to my big family for their prayer.
Acknowledgements

First I would like to acknowledge my advisor Balaji Rajagopalan for all his support and encouragement, without which I would not be where I am today. His enthusiasm and love of learning have been a constant inspiration - thank you. I would also like to acknowledge my co-authors, Edith Zagona, Ben Livneh and Joseph Kasprzyk for all their invaluable inputs. Finally, I would like to acknowledge the Directorate General of Higher Education, Republic of Indonesia for providing main funding support and Cooperative Institute for Research in Environmental Sciences for continuing funding support to finish my doctorate.
Contents

Chapter

Preface 1

Introduction 2

1 Space-Time Variability of Indonesian Rainfall at Inter-annual and Multi-decadal Time Scales 6
   1.1 Introduction .......................................................... 7
   1.2 Study Region and Data .............................................. 9
   1.3 Proposed Methods .................................................. 13
      1.3.1 Principal Component Analysis and Links to Large Scale Climate .... 14
      1.3.2 Spectral Analysis of Leading PCs and Coherence with Large Scale Climate 15
      1.3.3 Bayesian Dynamic Linear Model for Epochal Teleconnections .......... 16
   1.4 Results .............................................................. 18
      1.4.1 Dry Season .................................................... 18
      1.4.2 Wet Season .................................................... 27
   1.5 Summary and Discussion ......................................... 35

2 Modeling the Hydrologic Processes of Java Island, Indonesia 38
   2.1 Background .......................................................... 39
   2.2 Study Region and Data Sources ..................................... 40
   2.3 Methodology ....................................................... 43
2.3.1 Quality Control: Screening of Watershed Data ........................................ 43
2.3.2 Hydrologic Model ................................................................................... 47
2.3.3 Parameter Estimation - Calibration and Validation ................................. 48

2.4 Result .......................................................................................................... 52
2.4.1 Model performance .............................................................................. 52
2.4.2 Variability of water balance and hydrologic processes ......................... 58
2.4.3 Sensitivity to climate regimes ............................................................... 59
2.4.4 Sensitivity to spatial resolution ............................................................. 61

2.5 Summary and Discussion ........................................................................... 63

3 Development of A Meteorological Data Set over the Java Island, Indonesia 1985 - 2014 66
3.1 Background & Summary ........................................................................... 67
3.2 Methods ..................................................................................................... 69
3.2.1 Station data .......................................................................................... 69
3.2.2 Gap-infilling and gridding procedure .................................................. 70
3.2.3 Code availability .................................................................................. 73
3.3 Data Records ............................................................................................... 74
3.4 Technical Validation .................................................................................. 75
3.5 Usage Notes ............................................................................................... 78

4 Multi-objective Optimization Based Calibration of Hydrologic Model and Ensemble Hydro-
logic Forecast for Java Island, Indonesia .......................................................... 81
4.1 Background ............................................................................................... 82
4.2 Study Area and Data ................................................................................ 84
4.3 Methodology .............................................................................................. 86
4.3.1 Model setup ......................................................................................... 86
4.4 Objective functions .................................................................................. 87
4.4.1 Parameter calibration - Ensemble ........................................................ 89
### Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Hydro-climatic characteristic of Java.</td>
</tr>
<tr>
<td>2.2</td>
<td>Characteristics of the six watersheds over Java selected for this study. Land cover type A, F, S and W represent agriculture, forest, settlement and water body respectively with bold font indicating the dominant land cover.</td>
</tr>
<tr>
<td>2.3</td>
<td>Description of calibration parameters, their ranges and the values obtained from optimization for the six watersheds.</td>
</tr>
<tr>
<td>2.4</td>
<td>The change of mean rainfall and streamflow in the calibration and validation periods. The change in runoff ratio is computed as percentage change of runoff ratio in the validation period from it is in the calibration period.</td>
</tr>
<tr>
<td>2.5</td>
<td>NSE values for all watersheds in dry and wet season.</td>
</tr>
<tr>
<td>2.6</td>
<td>Water balance and calibrated soil parameters in the selected watersheds organized from west to east. The water balance was computed for calibration period. BFI is calculated as ratio of long-term baseflow to total streamflow.</td>
</tr>
<tr>
<td>2.7</td>
<td>NSE of the driest, average and wettest years in the watersheds.</td>
</tr>
<tr>
<td>3.1</td>
<td>Meteorological data products available for Java Island.</td>
</tr>
<tr>
<td>3.2</td>
<td>Summary of daily data sources used to develop present meteorological data set.</td>
</tr>
</tbody>
</table>
4.1 Characteristics of the six watersheds over Java selected for this study. Land cover type A, F, S and W represent agriculture, forest, settlement and water body respectively with bold font indicating the dominant land cover.

4.2 Summary of daily data sources used to develop present meteorological data set.

4.3 Description of calibration parameters, their ranges and the values obtained from optimization for the six watersheds.

4.4 The years when the observed streamflow data is available during the forecast period of 2001-2010 in the study watersheds.

4.5 The RPSS of ensemble forecast of seasonal flows, for dry and wet seasons, using three categories at the tercile thresholds. The last three columns show the RPSS of the extreme flows at three higher thresholds.
Figures

Figure

1  A maritime continent of Indonesia. ...................................................... 3

1.1 Indonesian archipelago. ............................................................... 10

1.2 Annual rainfall climatology at all stations (BMKG data) and grid (CRU) points shown as grey lines and their mean in a solid red line (a) BMKG (b) CRU. Indonesian seasonal average rainfall anomalies from BMKG (black) and CRU (blue) - (c) dry season and (d) wet season. ................................................................. 12

1.3 Spatial pattern of seasonal average rainfall (mm/month) over Indonesia from BMKG and CRU datasets - (a) and (c) are for dry season; (b) and (d) are for wet season. 13

1.4 Standardized time series of all Indonesian dry season rainfall (red dashed line), the black dots are flipped PC1 and the blue line is the decadal feature of dry season rainfall. ................................................................. 19

1.5 Dry season spatial and temporal pattern - (a) first spatial pattern - Eigen vector #1, (b) second spatial pattern - Eigen vector #2, (c) first temporal pattern - PC1, (d) second temporal pattern - PC2. ...................................................... 20

1.6 Dry season correlation - (a) sea surface temperature and PC1, (b) mean sea level pressure with PC1, (c) geopotential height at 500 mb with PC1, (d) sea surface temperature and PC2, (e) mean sea level pressure with PC2 and (f) geopotential height at 500 mb with PC2. The shaded areas inside thick blue and thick red lines show the correlations significant at 90% confidence level and higher. 22
1.7 Composite maps of dry season vector winds (arrows) and sea surface temperatures
anomalies (colors) during years of (a) low rainfall and, (b) high rainfall. Colored
regions are significant at 90% confidence level and higher on a t-test comparing the
mean SST between low and high rainfall years. . . . . . . . . . . . . . . . . . . . . . . 23

1.8 Wavelet spectra of dry season - (a) PC1, (b) NINO3.4, (c) wavelet spectral coherence
of PC1 and NINO3.4, (d) PC2, (e) PDO Index and (f) wavelet spectral coherence of
PC2 and PDO. The global spectra are shown on the right side of the time varying
wavelet spectra and, the blue and red lines are 95% and 90% confidence levels,
respectively. The curved lines in the spectrum and coherence figures are cones of
influence features outside of the cones are based on limited data and thus to be
interpreted cautiously. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 25

1.9 (a) Time varying coefficients of intercept and NINO3.4 from BDLM applied to PC1
and NINO3.4 (left panel). The right panel (b) is the same but from BDLM applied
to PC2 and PDO. The shaded region shows the 90% credible interval from the
posterior distribution. The blue line is the corresponding value from the best fit
linear regression - i.e., the stationary estimates, along with their 90% confidence
intervals. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 27

1.10 Standardized time series of all Indonesia wet season rainfall (red dashed line), the
black dots are flipped PC1 and the blue line is the decadal feature of wet season
rainfall. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 28

1.11 Wet season spatial and temporal pattern (a) first spatial pattern - Eigen vector #1,
(b) second spatial pattern - Eigen vector #2, (c) first temporal pattern - PC1, (d)
second temporal pattern - PC2. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 29
1.12 Wet season correlation - (a) sea surface temperature and PC1, (b) mean sea level pressure with PC1, (c) geopotential height at 500 mb with PC1, (d) sea surface temperature and PC2, (e) mean sea level pressure with PC2 and (f) geopotential height at 500 mb and PC2. The shaded areas inside thick blue and thick red lines show the correlations significant at 90% confidence level and higher. .......................... 30

1.13 Composite maps of wet season vector winds (arrows) and sea surface temperatures anomalies (colors) during years of - (a) low rainfall and, (b) high rainfall. Colored regions are significant at 90% confidence level and higher on a t-test comparing the mean SST between low and high rainfall years. ................................. 31

1.14 Wavelet spectrum of wet season - (a) PC1, (b) NINO3.4 Index and (c) wavelet spectral coherence of PC1 and NINO3.4., wavelet spectrum of - (d) PC2, (e) PDO Index and (f) wavelet spectral coherence of PC2 and PDO. The global spectra are shown on the right side of the time varying wavelet spectra and, the blue and red lines are 95% and 90% confidence levels, respectively. The curved lines in the spectrum and coherence figures are cones of influence - features outside of the cones are based on limited data and thus to be interpreted cautiously. .............................. 33

1.15 (a) Time varying coefficients of intercept and NINO3.4 from BDLM applied to PC1 and NINO3.4 (left panel). The right panel (b) is the same but from BDLM applied to PC2 and PDO. The shaded region shows the 90% credible interval from the posterior distribution. The blue line is the corresponding value from the best fit linear regression - i.e., the stationary estimates, along with their 90% confidence intervals. ......................................................... 35

2.1 Topography (a) and land cover (b) of the Java island. Mountain ranges extend across the middle of island from west to east where agricultural areas dominate. .............. 42
2.2 Spatial map of $R^2$ value between Sheffield rainfall and CRU TS3.22 rainfall in the period of 1948 - 2012. Red and blue numbers indicate $R^2 > 0.5$ and $R^2 < 0.5$ respectively. .................................................. 44

2.3 Climatology of rainfall and streamflow of 20 watersheds covering Java, each of which has at least 7 years of data. A runoff ratio (RR) is computed for each watershed as mean precipitation divided by mean flow over the entire period of record and shown in the panels. Watersheds selected for this study are indicated in red. ............... 45

2.4 Boxplot of monthly precipitation and streamflow for the 20 watersheds covering Java. The period and length of record for each watershed are different as presented in Table 2.2. Watersheds selected for this study are marked in red. ............... 46

2.5 The 20 watersheds covering Java with at least 7 years of data. The six watersheds shown in red were selected for this study after the screening process. ............... 47

2.6 Scatterplots of NSE vs soil parameter values from a Monte Carlo sampling of the parameter space. A local polynomial smoother through the scatterplot is shown as solid line. Results are used to determine sensitivity of NSE to changes in oil parameters, and choose soil parameters for model calibration. ............... 50

2.7 Time series and scatterplots of observed and simulated streamflow and rainfall hyetograph for CT, CL, PG, BG, GD and BD watersheds. The left panels show the calibration period and the right shows validation. In the time series plots, left and right y-axes are for streamflow and rainfall respectively. The boxes in the validation figures show the periods where the greatest errors occur, typically manifested in the lack of a corresponding observed hydrograph peak with precipitation input, or where there is a lag between monthly rainfall and observed streamflow peaks. ............... 53

2.8 Time series of evapotranspiration from the simulation and MOD16 in the period of 2000 - 2010. .......................................................... 57

2.9 Flow duration curve of monthly simulated streamflow in the wettest, average and driest year chosen from the calibration period of each watershed. .................. 60
2.10 Comparison of fluxes, performance metrics and discharge distribution between VIC model at 0.5° x 0.5° and 0.125° x 0.125° resolutions. In (a), N and Y represent meteorological forcings of Nijssen and Yanto respectively, on the right axis, model performance is better when the change in NSE positive, and vice versa, (b), (c) and (d) show the spatial distribution of flow at 1/2N, 1/8N and 1/8Y respectively.

3.1 Spatial and temporal distribution of precipitation data - (a) percentage of missing data of each station in the entire record period (the number of days with missing data divided by the total number of days), (b), (c) and (d) percentage of missing data for decade 1, 2 and 3 respectively and (e) percentage of stations with missing daily data (the number of stations with missing daily data divided by the number of available stations for particular day over Java).

3.2 (a) the number of stations with non-missing data within radius of influence \( r \) of 10, 25 and 50 km for each grid where each grid was assigned an index/number, the horizontal black line is a threshold of 4 stations for IDW interpolation; (b) selected locations to examine the performance of interpolation, the grid size is equivalent to 0.25° x 0.25° to capture at least one observed station; (c) model performance between the box-average gridded and observed precipitation; (d) and (e) are mean daily precipitation which gridded using \( r \) of 25 and 50 km respectively.

3.3 Scatterplot of PBIAS and precipitation magnitude for (a) temporal mean daily precipitation, computed as the average of daily precipitation in each station and (b) spatial mean daily precipitation, defined as the mean of station precipitation in each day.

3.4 (a) Q-Q plots for daily gridded precipitation and observed streamflow for five study watersheds, and (b) monthly hyetographs and hydrographs for the same watersheds, \( A \) is the basin area, km².
3.5 The probability density function (PDF) of mean daily P, Tmax and Tmin (a,e,i) respectively computed using the island-average value for each day in the record; the PDFs of mean daily P, Tmax, Tmin (b,f,j) computed across the mean daily value for each grid cell; the third column (c,g,k) shows scatterplots between daily gridded and observed P, Tmax and Tmin respectively overlaid by the 1:1 line (black) and regression line (red); and the fourth column (d,h,l) are monthly time series of P, Tmax and Tmin respectively.

4.1 Study area comprising the five watersheds over Java Island, colored in red.

4.2 Scatterplot of objective functions and model parameters for all five basins.

4.3 Scatterplot of objective functions and baseflow index (BFI) for all basins. The regression line and its respective coefficient of determination ($R^2$) was computed across basins to provide general insights of relationship between objective functions and hydrologic processes in tropical regions.

4.4 Parallel plot of objective functions of all solutions (dark gray) overlaid by the objective functions culled based on at least one of the criteria - NSE > 0.5, PBIAS < 25% and SFDCE < 25% (light gray). Objective function based on optimal parameters from a single objective function, NSE, optimization.

4.5 Same as Figure 4, but culled based on at least one of the criteria of NSE > 0.75, PBIAS < 15% and SFDCE < 15%.

4.6 Probability distribution function (PDF) of optimal model parameters (posterior) and initial population of model parameters (prior) for all the watersheds.
4.7 Monthly time series of flow ensembles (grey lines) along with observed flows (solid line) and that simulated from the parameter set with the best NSE (dotted line), for the calibration period (left column) and validation period (right column). The vertical bars are monthly precipitation, the scatterplots show the observed and simulated flows from the best NSE based parameter set with the 1:1 line (thick black line) and the best fit line through the scatter (blue line). 101

4.8 Dry and wet season precipitation forecast for dry year (2007) and wet year (2006) from IRI Columbia University. 104

4.9 Boxplot of flow ensemble for the forecast and climatology forecast of peak dry season total flow for (a) dry year and (b) wet year. The observed flow is shown as blue dots. 105

4.10 Same as Figure 4.9 but for wet season. 106
Preface

This dissertation is organized in article format. Each chapter is written as a standalone article for journal submission. As such, each chapter contains standalone introduction as well as conclusion. An aggregated bibliography has been provided for the entire dissertation due to many overlapping references.
Introduction

Indonesia spans from 11° S to 6° N and from 95° E to 142° E. It comprises five main islands: Sumatera, Java, Kalimantan (Borneo), Sulawesi (Celebes), and West Papua; two major archipelagos (Nusa Tenggara and Maluku); sixty smaller archipelagoes and nearly 17,000 small islands (Figure 1). Located in the equatorial western Pacific warm pool region, monsoonal wind and Inter Tropical Convergence Zone (ITCZ) movement characterize the climate over Indonesia, resulting in distinct dry (May - Oct) and wet seasons (Nov - Apr) (Qian et al., 2010) and clear patterns of spatial rainfall variability (Aldrian et al., 2007). The significant spatial and temporal variability of rainfall have caused various water related disasters such as floods, droughts, landslides and water-borne diseases endemic, and subsequently induced severe socio-economic impacts. To mitigate these hazards via efficient management of resources, tools for understanding, modeling and forecasting the hydroclimate over Indonesia is crucial.
Rainfall variability in this region at annual and inter-annual time scales is widely known to be driven by El Niño Southern Oscillation (ENSO) phenomena (Philander, 1983; Kirono et al., 1999; Haylock and McBride, 2001; Hendon, 2003). However, long term planning and management of natural resources and infrastructure require an understanding of rainfall variability at multi-decadal time scales and also of the hydrology. Research efforts at understanding the variability of hydroclimate are limited and there is no robust hydrologic modeling tool to understand and forecast hydrologic variability over Indonesia, which is crucial for water resources planning and natural hazard mitigation.

Motivated by these gaps and needs, this research makes four unique contributions arranged in four chapters. (i) The first chapter systematically explored the spatio-temporal rainfall variability over Indonesia at inter-annual and multi-decadal time scales. Large scale climate features modulating the space-time variability of wet and dry season rainfall at these time scales are identified. In addition, the time varying strength of these features with rainfall variability is investigated. (ii)
To understand the processes that drive the spatial and temporal hydrologic variability a physically based land surface model is important. In the second chapter, we developed a calibrated distributed Variable Infiltration Capacity (VIC) model for six watersheds over Java, the most populous island of Indonesia. Limited data and its poor quality posed significant challenge. The VIC model parameters were calibrated and validated on monthly time scale using a single objective function and, they provided insights into the hydrologic process such as surface flow, baseflow and evapotranspiration. This model represents the first such attempt to fully understand and model the hydrologic processes in Java. (iii) Data quality issue is a common problem in tropical watersheds especially in Java Island and most of Indonesia. In the third chapter, a fine resolution gridded daily meteorology - rainfall, maximum and minimum temperature - product was developed exclusively from ground-based observations. The gridded data at 0.125° x 0.125° resolution is the highest resolution product covering a 30-year period (1985 - 2014) ever created in Indonesia. The data set is stored in a network common data form (NetCDF) and intended to support basin-scale and island-scale studies of short-term and long-term climate, hydrology and ecology. (iv) Ensemble hydrologic forecasts that capture some of uncertainty in model parameters require an ensemble of model parameters. In the fourth chapter, the VIC model was set up using the high resolution daily meteorology from previous chapter and a multi-objective optimization based calibration approach. The method generates an ensemble of model parameters optimizing on a suite of objective functions designed to capture the various features of hydrograph. In addition to validating the multi-objective parameters, we demonstrated the utility of this in generating skillful ensemble seasonal hydrologic forecasts conditioned on seasonal climate forecasts.

The above contributions, especially the watershed modeling tools, are unique and are the first research efforts in this region. We offered insights into the space-time variability of precipitation and a robust physically-based watershed modeling tool to understand and forecast hydrologic variability over Java in particular and Indonesia in general. Together, this research makes significant strides in providing a good framework for understanding, modeling and forecasting Indonesian hydrology and climatology, that will help mitigate natural hazards and enable efficient management of Indonesia's
water and natural resources.
Chapter 1

Space-Time Variability of Indonesian Rainfall at Inter-annual and Multi-decadal Time Scales

This research has been published in the Climate Dynamics with the following citation:

Abstract

We investigated the space-time variability of wet (Nov - Apr) and dry (May - Oct) season rainfall over Indonesia, using monthly gridded rainfall data from the University of East Anglia Climatic Research Unit (UEA-CRU) covering the period 1901 - 2012. Three complimentary techniques were employed - (i) Principal Component Analysis to identify the dominant modes of variability, (ii) wavelet spectral analysis to identify the spectral characteristics of the leading modes and their coherence with large scale climate variables and (iii) Bayesian Dynamical Linear Model (BDLM) to quantify the temporal variability of the association between rainfall modes and climate variables. In the dry season when the Inter Tropical Convergence Zone (ITCZ) is to the north of the equator the leading two principal components (PCs) explain close to 50% of the rainfall. In the wet season the ITCZ moves to the south and the leading PCs explain close to 30% of the variance. El Niño Southern Oscillation (ENSO) is the driver of the leading modes of rainfall variability during both seasons. We find asymmetry in the teleconnections of ENSO to high and low rainfall years in the dry season. Furthermore, ENSO and the leading PCs of rainfall have spectral coherence in
the inter-annual band (2 - 8 years) over the entire period of record and in the multi-decadal (8 - 16 years) band in post-1980 years. In addition, during the 1950 - 1980 period the second mode of variability in both seasons has a strong relationship with Pacific Decadal Oscillation (PDO). The association between ENSO and the leading mode of Indonesian rainfall has strengthened in recent decades, more so during dry season. These inter-annual and multi-decadal variability of Indonesian rainfall modulated by Pacific climate drivers has implications for rainfall and hydrologic predictability important for water resources management.

Keywords : Indonesian rainfall, multi-decadal variability, PCA, Wavelet analysis, Dynamic linear models.

1.1 Introduction

Indonesia experiences significant variability of rainfall in space and time, resulting in negative socio-economic impacts. Droughts lead to reduced river inflow and consequently lower power generation and agricultural outputs, and drinking water shortages. Increased rainfall contributes to water-borne diseases such as cholera, typhoid and vector borne diseases such as malaria, dengue, etc. - which was the case during the wet epoch of 1992 - 2005 (MoE, 2007).

Rainfall over Indonesian maritime continent exhibits coherent patterns of variability at annual and inter-annual time scales (Haylock and McBride, 2001). In recent decades, the annual rainfall depth recorded in 63 stations for about 43 years (period of record varied from the earliest year 1950 and the latest 1974 until 1997) shows a decreasing trend, except for some stations in the Lesser Sunda Islands, the eastern coast of Java and the northern part of Indonesian (Aldrian, 2006; Aldrian et al., 2007; Aldrian and Djamil, 2008). Being geographically close to the tropical Pacific warm pool, it is expected that variability of rainfall over Indonesia is associated with El Niño Southern Oscillation (ENSO). The role of ENSO-induced sea surface temperatures (SSTs) variability on the seasonal and inter-annual variability of Indonesian rainfall has been widely investigated (Philander, 1983; Kirono et al., 1999; Haylock and McBride, 2001; Hendon, 2003). During El Niño (La Niña) events, or warm (cold) phases of ENSO, this region experiences lower (higher) rainfall than in
average years (Gutman et al., 2000).

Sea surface temperatures around Indonesia, more broadly in the warm pool region, play a crucial part in modulating climate around the world. Early research found the ability of SSTs in this region to forecast the Indian monsoon with longer lead times than other indices (Shukla and Paolino, 1983). Lau and Wu (2001) illustrated the significance of regional processes over Indonesia in affecting the rainfall variability of the Australian - Asian summer monsoon during the period 1979 - 1998. While ENSO SSTs account for 47% of the rainfall variability on average, the regional processes account for an additional 16%. Aldrian and Susanto (2003) found relationship between local SSTs and rainfall in the southern Sumatera, southern Kalimantan, southern Sulawesi, entire Java and Lesser Sunda Islands. These are the islands in the Indonesian throughflow region where water and heat are transported from the Pacific Ocean into the Indian Ocean (Gordon et al., 2003; Susanto et al., 2012). In addition, ENSO SSTs have teleconnection with summer drought in the midlatitude and tropics over Americas (Rajagopalan et al., 2000; Dettinger et al., 2001; Fu et al., 2013; Yin et al., 2014) and with winter climate via its correlation to Pacific North America (PNA) pattern (Renwick and Wallace, 1996; Strauss and Shukla, 2002).

To the west of Indonesia, there is an ENSO-like phenomenon called Indian Ocean Dipole (IOD), a coupled ocean and atmosphere phenomenon in the equatorial Indian Ocean indicated by anomalously low and high SSTs in the eastern and western Indian Ocean respectively, and accompanied by consistent wind and precipitation anomalies. In its active years, IOD causes severe drought in Indonesia (Saji et al., 1999). Hendon (2003) found that anomalous gradient in the equatorial Indian Ocean is strongly related to ENSO and is most prominent during the period of May - Oct. The entire Indian Ocean tends to have the same-signed SSTs anomaly as that of ENSO events during the period of Nov - Apr (Hendon, 2003).

Variability of Pacific SSTs on inter-decadal time scales which is ENSO-like have been well documented (e.g., Mantua et al., 1997; Zhang et al., 1997; Dai, 2013). Its Northern Pacific manifestation is referred as Pacific Decadal Oscillation (PDO) and the pattern covering the entire Pacific Ocean is referred as Inter-decadal Pacific Oscillation (IPO). The features of both PDO and IPO
are shown to be similar in terms of their mechanisms (Deser et al., 2004). The influence of IPO in modulating Australian rainfall has been well documented by Power et al. (1999) and more recently the role of IPO and PDO in the variability of global precipitation and temperature (Dong and Dai, 2015). These prior results motivated us to consider PDO in our analysis in this research.

Most studies in the past used data from the 1950’s to 1990’s at ~60 rainfall stations (Philander, 1983; Kirono et al., 1999; Haylock and McBride, 2001; Hendon, 2003). The time scales of analysis in previous studies are diurnal, monthly, seasonal and annual (Hamada et al., 2002; Aldrian and Susanto, 2003; Hendon, 2003; Qian et al., 2010; Teo et al., 2011) to a lesser extent on inter-annual variability. However, we found no systematic study investigating the inter-decadal variability of rainfall in this region, which is important for long term planning and management of natural resources and infrastructure. Motivated by this need, the present research sets out to investigate the space-time variability of Indonesian rainfall at inter-annual and multi-decadal time scales, which can be exploited for skillful seasonal rainfall forecasting efforts.

1.2 Study Region and Data

Indonesia spans from 11°S to 6°N and from 95°E to 142°E, extending 5,120 kilometers (3,181 mi) from east to west and 1,760 kilometers (1,094 mi) from north to south. It comprises five main islands: Sumatera, Java, Borneo (known as ”Kalimantan” in Indonesian), Sulawesi, and Papua; two major archipelagos (Nusa Tenggara and the Maluku Islands); and sixty smaller archipelagoes (Figure 1.1).
Throughout the year Indonesian climate is typically equatorial with hot and humid conditions (McSweeney et al., 2010). Warm sea surface temperatures surrounding the islands, part of the tropical Pacific warm pool mentioned above, maintain the overland temperature fairly constant with the coastal plains averaging 28°C, the inland and mountain areas averaging 26°C, and the higher mountain regions, 23°C (McSweeney et al., 2010). The rainfall is controlled by the movement of the InterTropical Convergence Zone (ITCZ). The wettest months (January and February) occur when the ITCZ is in its southern-most position, and the driest months are July to September when the ITCZ is to the north of the equator. The annual rainfall varies from 1800 - 3200 mm in the lowlands to an average of 6000 mm in some mountain areas (McSweeney et al., 2010).

Two rainfall data sets were considered for this study: (i) A monthly rainfall dataset from the University of East Anglia Climatic Research Unit (UEA-CRU). This dataset is on a 0.5° x 0.5° grid over land spanning Jan 1901 to Dec 2012 and is based on 1175 station observations with the principal source being World Meteorological Organization (WMO) (Harris et al., 2013). (ii) Daily rainfall data from 51 meteorological stations for the period of 1962 to 1996 were available from the Indonesian Agency for Meteorology, Climatology and Geophysics or Badan Meteorologi, Klimatologi dan Geofisika (BMKG). This station data is part of the WMO station network and is incorporated in the CRU data set above.

The annual rainfall climatology at all the BMKG stations and CRU grid points are shown
in Figure 1.2a and b respectively as grey lines. The mean climatology is shown in solid red line and the annual mean rainfall as dashed line. It can be seen that the mean climatological pattern is similar in both data sets, as are their annual mean rainfall. The mean climatology is above the annual mean rainfall during November to April and below during May to October. We use this distinctive feature to define the wet (Nov - Apr) and dry (May - Oct) seasons in this research, which is also consistent with the seasonal category for Indonesia suggested by Hendon (2003). The country average rainfall from these two data sets for the dry and wet seasons for the overlapping period is shown in Figure 1.2c and d respectively. The similarity between these two data sets over the time is quite clear. Spatial average rainfall for the dry season from CRU and BMKG (Figure 1.3a and c), and for the wet season (Figure 1.3b and d), suggests that these two data sets are consistent in their spatial patterns over the entire year (Figure 1.3). We computed spectral analysis of the average rainfall series from the two data sets for both the seasons and found them to be quite similar, furthermore, they both exhibited high spectral coherence at all the frequencies (figures not shown in the interest of space). Considering the close correspondence between these two data sets, we select the CRU data for our analysis in this research, as it covers a longer period (1901 - 2012) and is thus more suitable to analyzing multi-decadal variability.
Figure 1.2: Annual rainfall climatology at all stations (BMKG data) and grid (CRU) points shown as grey lines and their mean in a solid red line (a) BMKG (b) CRU. Indonesian seasonal average rainfall anomalies from BMKG (black) and CRU (blue) - (c) dry season and (d) wet season.
For large scale climate variables we selected monthly global Kaplan SSTs anomalies (Kaplan et al., 1998) and monthly circulation variables - mean sea level pressure (SLP), geopotential height at 500 mb and wind anomalies from National Centers for Environmental Prediction - National Center for Atmospheric Research (NCEP-NCAR) Reanalysis (Kalnay et al., 1996). The widely used index of ENSO, NINO3.4 (average monthly SSTs in the region 120°W - 170°W and 5°N - 5°S) was computed from Kaplan SSTs anomalies. Monthly PDO anomalies from 1900 to present were obtained from University of Washington. The PDO is calculated as the first principal component of the Northern Pacific SSTs (Mantua et al., 1997; Zhang et al., 1997). Wet and dry season values of the climate variables and indices for use in the analyses were computed from the monthly data.

1.3 Proposed Methods

Three methods were employed to (i) isolate the dominant space-time patterns in rainfall and their teleconnections to large scale climate, (ii) understand the dominant low frequency components at inter-annual and multi-decadal time scales and (iii) understand the epochal variability of rainfall-climate teleconnections. These methods are described briefly below.
1.3.1 Principal Component Analysis and Links to Large Scale Climate

The gridded rainfall is subjected to principal component analysis (PCA). This is a widely used technique (e.g. Wilks, 1995) to extract dominant spatial and temporal modes of variability from mutually correlated data, such as this. The leading principal components (PCs) are orthogonal and explain most of the variance with the first explaining greatest variance and decreasing thereafter. The leading components, capturing most of the variance, help to reduce the dimension of the original data set. The description of PCA can be found in several sources including standard books (Wilks, 1995; Stroch and Zweirs, 1999). PCA as mentioned above is widely used in climate and hydroclimate research for diagnosis (Westra et al., 2007; Gebrehiwot et al., 2011; Teo et al., 2011; Sharma et al., 2013) and for forecasting (McCabe and Dettinger, 2002; Eldaw et al., 2003; Regonda et al., 2006; Bracken et al., 2010).

The principal components are obtained by computing the eigenvectors and eigenvalues of the data covariance matrix. Suppose the data is in matrix $A$ with $m$ rows and $n$ columns (this could be space-time data, with rows being the time and columns being the space), and $\Phi$ is the Eigen vector matrix of $n$ rows and $n$ columns, such that:

$$ A\Phi = PC $$

(1.1)

where $PC$ is the matrix of orthogonal principal components, also of $m$ rows and $n$ columns. The Eigen vectors are obtained by decomposing the $n \times n$ covariance matrix, $\Sigma$, of the data $A$ as:

$$ \Sigma = \Phi \Lambda \Phi^T $$

(1.2)

where $\Lambda$ is the diagonal matrix of $n$ Eigen values in descending order. The proportion of each value to the total represents the fraction of multivariate data variance captured by each principal component. The Eigen vector matrix is orthonormal - i.e.

$$ \Phi \Phi^T = \Phi^T \Phi = I $$

(1.3)
In the case of space-time data, the Eigen vectors represent the spatial patterns and the principal components of the corresponding temporal patterns. Also:

$$\mathbf{A} = \mathbf{PC}\Phi^T$$  \hspace{1cm} (1.4)

This enables easy transformation from the principal component space to the original data space through $\Phi$.

Large scale climate features are known to drive variability of rainfall over Indonesia (Philander, 1983; Kirono et al., 1999). To identify this, we correlated the leading PCs with large scale climate variables to produce correlation maps to understand the teleconnections. We also produced composite maps of the variables for wet and dry years - together they help isolate the physical mechanisms that drive the rainfall variability. The wet and dry rainfall years were selected separately for each season based on the standardized spatial average rainfall at threshold of $\pm 0.75$.

1.3.2 Spectral Analysis of Leading PCs and Coherence with Large Scale Climate

To understand the temporal variability of rainfall, the leading PCs are subjected to spectral analysis. For this, wavelet spectral analysis is employed. This is a nonparametric spectral analysis method which provides the spectral variability over time which is important in identifying multi-decadal variability. Wavelet analysis is widely used in identifying relationships between hydroclimate variables and large scale climate features such as ENSO, PDO, North Atlantic oscillation (NAO) etc. (Torrence and Webster, 1999; Coulibaly and Burn, 2004; Grinsted et al., 2004; Zhang et al., 2007). Wavelet analysis has also been used in time series modeling and simulation (Kwon et al., 2007; Nowak et al., 2011). The method is described in detail in several places (Torrence and Compo, 1998; Nowak et al., 2011). A brief description is provided here.

A time series can be transformed using wavelet as:

$$X(a, b) = |a|^{-1/2} \int_{-\infty}^{\infty} x_t \varphi * \left( \frac{t - b}{a} \right) dt$$  \hspace{1cm} (1.5)
where $a$ is a scale parameter, $b$ is the shift parameter and $\varphi^*$ is the wavelet function (Kwon et al., 2007). The (*) denotes complex conjugate (Nowak et al., 2011). We employ the Morlet wavelet function ($\psi_0$), given by:

$$
\psi_0(\eta) = \pi^{-1/4} e^{i\omega_0 \eta} e^{-\eta/2}
$$

(1.6)

where $\omega_0$ and $\eta$ are non-dimensional frequency and time parameters respectively (Torrence and Compo, 1998). The wavelet transform, on the other hand, can be defined as the inverse Fourier transform of the product of the data and the wave function in the Fourier space. Torrence and Compo (1998) provide additional details on the wavelet-based time series spectral estimation.

The global spectrum in the wavelet analysis is used to detect significant spectral components of the wavelet transformed data. Several band passed time series are constructed from the frequencies and their residuals that correspond to statistically significant spectral peaks. This time series is then back-transformed to the original space given by:

$$
x_t = \frac{\delta_j \delta_{1/2}}{C_\delta \psi_0(0)} \sum_{j=0}^{J} \Re\{X_t(a_j)\} \frac{a_j^{1/2}}{a_j^{1/2}}
$$

(1.7)

where $\delta_j$ and $\delta_j$ are the coefficients of scale averaging and sampling period respectively. $C_\delta$ and $\psi_0$ are empirically derived reconstruction factors specific to the Morlet wavelet (Torrence and Compo, 1998). The summation term in this equation is the sum of the real part of the wavelet transform over all scales.

### 1.3.3 Bayesian Dynamic Linear Model for Epochal Teleconnections

Relationships between climate variables and rainfall undergo epochal fluctuations. Prominent examples of this are the changing strength of relationship between Indian monsoon rainfall and ENSO (Kumar and Kleeman, 1999; Torrence and Webster, 1999) and the Thailand monsoon - ENSO relationship (Singhrattna et al., 2005). Traditional linear regression methods are incapable of capturing the changing relationship, i.e., nonstationarity. Bayesian Dynamic Linear Model
(BDLM) provides an attractive alternative to modeling and understanding the nonstationarity in
the relationships (West and Harrison, 1997; Petris et al., 2009). BDLM has been applied to mod-
eling Indian summer monsoon rainfall variability (Krishnaswamy et al., 2014; Maity and Kumar, 2006) and found interesting insights and better performance than the traditional regression meth-
ods showing the strengthening influence of IOD on Indian Monsoon (IM) and extreme rainfall
events (EREs) in recent decades while the influence of ENSO is weakening. Moreover, BDLM in-
creases the predictability of Indian summer monsoon rainfall (ISMR). In this method, the regression
coefficients vary with time, unlike traditional regression where the coefficients remain fixed.

In BDLM the time series is considered as the output of a dynamical system perturbed by
random disturbances (i.e., noise) - thus, considered a nonstationary evolution. The general form of
the model is represented as:

$$y_t = x_t \theta_t + v_t \quad v_t \sim N(0, V_t)$$

$$\theta_t = \theta_{t-1} + w_t \quad w_t \sim N(0, W_t)$$

where \(y_t\) is a dependent variable (e.g. the leading PC), \(x_t\) is a vector of independent variables
(e.g., NINO34, PDO, etc.), \(\theta_t\) is coefficient of \(x_t\), \(t\) denotes time. Both observation and system
equations can have additive Gaussian errors with covariance matrices \(V_t\) and \(W_t\).

The posterior predictive distribution of model coefficients at each time \(t\) is computed from
the prior distribution of coefficients at time step \(t-1\). Using Bayes theorem, the probability of the
data \(y_t\) conditional on the model parameters at time \(t\) is defined as:

$$P(\theta_t|y_t) \propto P(y_t|\theta_t)P(\theta_t|y_{t-1})$$

The model coefficients \(\theta\) consist of PCs, NINO34, PDO, the regression coefficients and the
variances of the Gaussian errors in the above equations. In the initial step, traditional linear regres-
sion is applied to compute regression coefficient with normally distributed mean and variance - i.e.
\(\theta_0 \approx N(m_0, C_0)\) where \(m_0\) and \(C_0\) are the mean vector and the variance - covariance matrix of the
regression parameters. Using Kalman filtering approach together with Markov Chain Monte Carlo simulation approach, posterior distribution is estimated at any time $t$. The posterior distribution is then employed to generate the Bayesian confidence intervals (Petris et al., 2009).

1.4 Results

We present results for the dry and wet seasons separately - and for each, the leading PCs, correlation with climate variables, spectral analysis and BDLM results are presented and discussed.

1.4.1 Dry Season

The first two PCs explain 39.03% and 8.28% respectively, which together captures close to 50% of the total data variance. The time series of the first PC corresponds very well with the spatial average rainfall (Figure 1.4 - shown with the sign of the first PC is flipped for easy comparison). The solid line is a decadal smoother of the first PC, which will be described later. The first two Eigen vectors, i.e., the spatial components and the corresponding PCs, are shown in Figure 1.5. The first spatial pattern (Figure 1.5a) has similar magnitude and sign throughout Indonesia, indicating that the dominant pattern is coherent across space. The second Eigen vector (Figure 1.5b) shows a north-south feature consistent with the movement and location of the ITCZ. In the dry season the ITCZ is located to the north of the equator. The PCs (Figure 1.5c and d) show rich temporal variations, which are described in the spectral analysis below.
Figure 1.4: Standardized time series of all Indonesian dry season rainfall (red dashed line), the black dots are flipped PC1 and the blue line is the decadal feature of dry season rainfall.
Correlation maps of Principal Component 1 (PC1) and Principal Component 2 (PC2) with large scale climate variables are shown in Figure 1.6. As mentioned before the signs of PC1 and PC2 were flipped in these correlations so that the correlations of PC1, especially, can be directly inferred as that of all Indonesian rainfall. Figure 1.6a shows the correlation of PC1 with SSTs. Strong negative correlation (∼ -0.6) in the equatorial and eastern tropical Pacific regions with a strong positive correlation (∼ 0.6) in the western Pacific is characteristic of ENSO SSTs pattern. When the western Pacific is warmer as is the case during La Niña events, it leads to increased convection over Indonesia and surroundings resulting in enhanced rainfall - vice-versa during El Niño events (Aldrian and Susanto, 2003; Hamada et al., 2002). The correlation with sea level pressure (Figure 1.6b) is consistent with that of the SSTs above in that the eastern and
western Pacific show strong and opposite signed correlations, reminiscent of the southern oscillation (Hendon, 2003). A warmer western Pacific is consistent with a low pressure in this region and hence negative correlation. In addition, PC1 has strong correlation with SSTs and SLP in the Indonesian throughflow region indicating its significance on the Indonesian rainfall. Correlations with geopotential height at 500 mb are shown in Figure 1.6c - which shows a pattern of positive, negative and positive correlations over central north Pacific, Alaska and eastern north America, respectively. This pattern is reminiscent of the wave train PNA pattern (Wallace and Gutzler, 1981; Strauss and Shukla, 2002). Furthermore, it is known that the rainfall over Indonesia and Pacific warm pool rides a signature on the PNA pattern (Renwick and Wallace, 1996; Yu et al., 2008). Correlations of large scale climate variables with PC2 (Figure 1.6d,e,f) are weaker and barely significant. This suggests that all the variability driven by large scale climate variables present in the first mode.
Figure 1.6: Dry season correlation - (a) sea surface temperature and PC1, (b) mean sea level pressure with PC1, (c) geopotential height at 500 mb with PC1, (d) sea surface temperature and PC2, (e) mean sea level pressure with PC2 and (f) geopotential height at 500 mb with PC2. The shaded areas inside thick blue and thick red lines show the correlations significant at 90% confidence level and higher.

Composite maps of anomalous vector winds and SSTs of wet and dry years are shown in Figure 1.7. During low rainfall years (Figure 1.7a) central and eastern tropical Pacific is anomalously warmer thus moving the convection to this region from the warm pool (El Niño conditions). Consequently, the winds are anomalously from the west and blowing away from Indonesia in both the Pacific and Indian Ocean basins, thus reducing the warm waters, convection and rainfall. Interestingly, in the Indian Ocean during this phase, the winds are away from the Indian subcontinent and towards western Indian Ocean - leading to reduced Indian summer monsoon which is the standard monsoon-ENSO teleconnection mechanism. During high rainfall years (Figure 1.7b) the pattern is
almost symmetric in the winds but shows differences in SSTs. The warming in the western Pacific is conspicuous along with the cooling in the central and eastern Pacific during high rainfall years. This asymmetry in SST teleconnections between high and low rainfall years during dry season is interesting and has implications for forecasting efforts. This is consistent with recent findings of asymmetry in the eastern Australian - IPO relationship by King et al. (2013) and in the Indian summer monsoon - ENSO teleconnections of Gill et al. (2015).

Figure 1.7: Composite maps of dry season vector winds (arrows) and sea surface temperatures anomalies (colors) during years of (a) low rainfall and, (b) high rainfall. Colored regions are significant at 90% confidence level and higher on a t-test comparing the mean SST between low and high rainfall years.

To investigate the temporal variability of the leading modes of rainfall, spectral analyses are performed on PC1, PC2 and with large scale climate indices, and are shown in Figure 1.8. The global and local spectrum of PC1 (Figure 1.8a) indicates spectral peaks in the 2 - 8 year band and 16 - 32 year band - further, these periods seem to be active in recent decades. In PC2 (Figure 1.8d)
these bands are active in the middle part of 20th century. The ENSO index (NINO3.4) shows significant peaks in the 2 - 8 year period and a weaker peak in the 8 - 16 year period (Figure 1.8b). The spectral coherence between PC1 and NINO3.4 (Figure 1.8c) indicates a strong coherency in the 2 - 8 year band throughout the length of the record and stronger coherency in the decadal band of 8 - 16 year period in recent decades (post-1980). It is interesting to note that ENSO modulates Indonesian rainfall at inter-annual and multi-decadal time scales, but with different temporal variability. The coherence between PC2 and PDO appears in the period ~1950 - 1980 (Figure 1.8f). This relationship is consistent with inter-decadal modulation of precipitation by PDO over the globe including over Indonesia (e.g. Dong and Dai, 2015). Wavelet filtering of PC1 in the decadal band (8 - 16 year period) where ENSO is coherent, was made, and is shown in Figure 1.4 as a solid line. It can be seen that the decadal feature exhibits higher variability in recent decades, consistent with the features seen in the spectral coherence (Figure 1.8c).
Figure 1.8: Wavelet spectra of dry season - (a) PC1, (b) NINO3.4, (c) wavelet spectral coherence of PC1 and NINO3.4, (d) PC2, (e) PDO Index and (f) wavelet spectral coherence of PC2 and PDO. The global spectra are shown on the right side of the time varying wavelet spectra and, the blue and red lines are 95% and 90% confidence levels, respectively. The curved lines in the spectrum and coherence figures are cones of influence features outside of the cones are based on limited data and thus to be interpreted cautiously.
The temporal variability of the strength of relationship between large scale climate drivers and leading modes of rainfall is assessed using BDLM performed on PC1 and NINO3.4 and, PC2 and PDO (Figure 1.9). The horizontal lines in the panels shows the values of intercept and slopes from a best fit linear regression i.e., a stationary model along with 90% confidence interval. The time varying intercept and the coefficient of NINO3.4 are shown in Figure 1.9a. The intercept from BDLM shows a decreasing trend indicating a decrease in the mean rainfall over Indonesia, especially in recent decades. The NINO3.4 coefficient shows an increasing trend starting ∼1970, indicating that the ENSO connection to Indonesian rainfall is strengthening in recent decades. During other periods the BDLM estimates are not different from the stationary values. This is counter to the reduction in Indian monsoon - ENSO teleconnection found by Krishnaswamy et al. (2014) also using BDLM. Since ENSO has coherence in the inter-annual and decadal bands (Figure 1.8c) in the post-1980 period as mentioned above, we hypothesize this to be the potential reason for increasing ENSO strength from the BDLM analysis. For PC2 (Figure 1.9b) the intercept shows a decline during 1950 - 1970 but not significantly different from the stationary intercept. The strength of PDO is similar to the coefficient from the stationary regression with deviation during 1905 - 1980.
Figure 1.9: (a) Time varying coefficients of intercept and NINO3.4 from BDLM applied to PC1 and NINO3.4 (left panel). The right panel (b) is the same but from BDLM applied to PC2 and PDO. The shaded region shows the 90% credible interval from the posterior distribution. The blue line is the corresponding value from the best fit linear regression - i.e., the stationary estimates, along with their 90% confidence intervals.

1.4.2 Wet Season

The total data variance captured by the first two PCs is less than 50% where PC1 and PC2 explain 18.07% and 9.57% respectively which is smaller than the total variance explained by the two leading PCs of dry season rainfall. The spatial average rainfall is well captured by the first PC shown in Figure 1.10 (shown with the sign of the first PC flipped for easy comparison with the spatial average rainfall). The solid line is a decadal smoother of the first PC, which will be described later, but it can be seen to follow the low frequency temporal variations. The first two Eigen vectors - i.e., the spatial components and the corresponding PCs are shown in Figure 1.11. The first spatial pattern (Figure 1.11a) has magnitudes of same sign across the region, hence the good correlation with spatial average rainfall seen in Figure 1.10, with a slight north-south divide in the magnitude and also in sign. The second Eigen vector (Figure 1.11b) shows a north-south feature with strong positive values to the south of the equator and negative to the north, consistent with the movement and location of the ITCZ. In the wet season the ITCZ is located to the south
of the equator. The PCs (Figure 1.11c and d) show rich temporal variations, which are described in the spectral analysis below.

Figure 1.10: Standardized time series of all Indonesia wet season rainfall (red dashed line), the black dots are flipped PC1 and the blue line is the decadal feature of wet season rainfall.
Correlation maps of PC1 and PC2 with large scale climate variables are shown in Figure 1.12. As mentioned previously the signs of PC1 and PC2 were flipped in these correlations so that the correlations of PC1, especially, can be directly inferred as that of all Indonesian rainfall. Correlation of PC1 with SSTs (Figure 1.12a) shows clear ENSO pattern characterized by strong negative correlation ($\sim -0.5$) in the equatorial and eastern tropical Pacific regions and a strong positive correlation ($\sim 0.5$) in the western Pacific. The correlation with sea level pressure (Figure 1.12b) in the eastern and western Pacific show strong and opposite signed correlations which is consistent with that of the SSTs above, reminiscent of the southern oscillation (Hendon, 2003). These correlations are similar to those observed in the dry season and described above with same ENSO teleconnection mechanisms (Hamada et al., 2002; Aldrian and Susanto, 2003). In addition, correlations of local SSTs and SLP with PC1 in wet season are weaker than in dry season. The correlations are also weaker over the Indonesian throughflow region consistent with its weaker connection to Indonesian
dry season rainfall (Shinoda et al., 2012). Correlation with geopotential heights at 500 mb, shown in Figure 1.12f, corresponds to the wave train PNA pattern with lower strength than in dry season. Correlations of large scale climate variables with PC2 (Figure 1.12d,e) are weaker and barely significant except for geopotential heights at 500 mb (Figure 1.12f). This suggests that most of the variability driven by large scale climate variables is present in the first mode.

![Figure 1.12: Wet season correlation - (a) sea surface temperature and PC1, (b) mean sea level pressure with PC1, (c) geopotential height at 500 mb with PC1, (d) sea surface temperature and PC2, (e) mean sea level pressure with PC2 and (f) geopotential height at 500 mb and PC2. The shaded areas inside thick blue and thick red lines show the correlations significant at 90% confidence level and higher.](image)

Composite maps of anomalous vector winds and SSTs for high and low rainfall years based on high and low values of the seasonal spatial average rainfall are computed and shown in Figure 1.13.
During El Niño years, central and eastern tropical Pacific is anomalously warmer thus moving the convection to this region from the warm pool (Figure 1.13a) with anomalous westerlies by reducing convection and rainfall over Indonesia. During La Niña years (Figure 1.13b) the pattern is almost symmetric in the winds and SSTs in both the Pacific and Indian Ocean, unlike in the dry season.

Figure 1.13: Composite maps of wet season vector winds (arrows) and sea surface temperatures anomalies (colors) during years of - (a) low rainfall and, (b) high rainfall. Colored regions are significant at 90% confidence level and higher on a t-test comparing the mean SST between low and high rainfall years.

This finding that the wet season correlation pattern being somewhat weaker than the dry season, even though ENSO peaks during the wet season, is interesting. During the dry season the ITCZ is to the north of the equator and almost all of Indonesia is under the south easterly component of the trade winds. Furthermore, the rainfall is quite homogeneous and thus, the first principal component explains $\sim$39% of the total variance of the seasonal rainfall. During the wet season the ITCZ is to the south of the equator and the country is under the influence of north
easterly and south easterly component of the trade winds, leading to heterogeneity in the spatial pattern of seasonal rainfall. As a result, the first principal component explains $\sim 19\%$ of the total variance. Therefore, a stronger signal is likely to emerge in the correlation patterns of SLP and SST with the first PC of dry season rainfall, compared to the wet season. Thus, the correlation patterns are driven to a large extent on the strength of the first PC of the seasonal rainfall and not so much on the ENSO.

The global and local spectrum of PC1 (Figure 1.14a) indicate spectral peaks in the 2 - 8 year band and 32 - 64 year band - further, these periods seem to be active in recent decades. On the other hand, PC2 (Figure 1.14d) displays spectral peaks in the 2 - 8 year and 8 - 16 year where the first period is active in multiple time windows and the second period seem to be active in the entire 20$^{th}$ century. Similar to PC2, the ENSO index (NINO3.4) shows significant peaks in the 2 - 8 year and the 8 - 16 year period - further, both periods are active in the multiple time windows (Figure 1.14b). The spectral coherence between PC1 and NINO3.4 indicates a strong coherency in the 2 - 8 year band in the early and recent decades (pre-1940 and post-1980) and stronger coherency in the decadal band of 8 - 16 year period throughout the length of the record (Figure 1.14c). It is interesting to notice that in wet season ENSO modulates Indonesian rainfall at inter-annual and multi-decadal time scales, but with different temporal variability - which is the case in dry season. The coherence between PC2 and PDO in the 2 - 8 year and 8 - 16 year band is strong in the multiple time epochs (Figure 1.14f) - as noted with dry season before (Dong and Dai, 2015). Wavelet filtering of PC1 in the decadal band (8 - 16 year period) where ENSO is coherent, was made, and is shown in Figure 1.10 as a solid line. It can be seen that the decadal feature exhibits higher variability in recent decades, consistent with the features seen in the spectral coherence (Figure 1.14c).
Figure 1.14: Wavelet spectrum of wet season - (a) PC1, (b) NINO3.4 Index and (c) wavelet spectral coherence of PC1 and NINO3.4., wavelet spectrum of - (d) PC2, (e) PDO Index and (f) wavelet spectral coherence of PC2 and PDO. The global spectra are shown on the right side of the time varying wavelet spectra and, the blue and red lines are 95% and 90% confidence levels, respectively. The curved lines in the spectrum and coherence figures are cones of influence - features outside of the cones are based on limited data and thus to be interpreted cautiously.
To investigate the temporal variability of the strength of the teleconnections, BDLM was constructed on PC1 and NINO3.4 and; PC2 and PDO (Figure 1.15) and their 90% confidence intervals are shown. The horizontal lines in these panels shows the stationary regression slopes and their 90% confidence intervals. The intercept shows a decreasing trend in recent decades similar to that seen in the dry season, indicating a decrease in the mean rainfall over Indonesia (Figure 1.15a). The NINO3.4 coefficient in the regression shows an increasing trend during 1950 - 2000 and significantly deviates from the stationary regression estimates, with a decrease to the stationary estimates after 2000. This indicates that the ENSO connection to Indonesian rainfall strengthening over 1950 - 2000 and weakening to stationary levels thereafter (Figure 1.9a). Since ENSO has coherence in the inter-annual and decadal bands (Figure 1.14c) throughout 20th century as mentioned above, we suspect this to be the potential reason for increasing ENSO strength from the BDLM analysis. For PC2 (Figure 1.15b) the intercept and the strength of PDO are no different than that of the stationary estimates with significant deviation during 1950 - 1980, similar to the dry season features (Figure 1.9b).
Figure 1.15: (a) Time varying coefficients of intercept and NINO3.4 from BDLM applied to PC1 and NINO3.4 (left panel). The right panel (b) is the same but from BDLM applied to PC2 and PDO. The shaded region shows the 90% credible interval from the posterior distribution. The blue line is the corresponding value from the best fit linear regression - i.e., the stationary estimates, along with their 90% confidence intervals.

1.5 Summary and Discussion

Spatial and temporal variability of Indonesian rainfall is depicted using PCA on the monthly gridded rainfall data from UEA-CRU - which has been proven to be similar to the observed rainfall from BMKG in the overlapping period of 1962 - 1996 but with longer record (1901 - 1912) and larger number of stations. The PCs show significant temporal variability at inter-annual and multi-decadal time scales. The first and second Eigen vectors show the dominant rainfall pattern across Indonesia and their consistency with the movement of the ITCZ in dry season (May - Oct) and wet season (Nov - Apr). During dry season the ITCZ is to the north of equator, leading to more spatial coherence of dominant pattern and the first two PCs account for 50% of variance. In wet season the ITCZ moves to the south of equator, leading to less spatial coherence compared to dry season and the first two PCs account for 30% of variance.

Correlation maps of the leading PCs and global climate variables illustrate the atmospheric circulation over the Pacific Ocean (ENSO, PNA, NAO) in both season and composite maps of
SSTs and winds depict the physical mechanism driving rainfall over Indonesia. La Niña events are characterized by strong positive SSTs correlation (warmer temperature), a strong negative SLP correlation (low pressure) in the western Pacific, resulting in enhanced easterlies over Indonesia and rainfall - vice versa for El Niño events. Interestingly, we find asymmetry in the ENSO teleconnections between wet and dry years during the dry season. This asymmetry adds to the similar findings by King et al. (2013) in the eastern Australian rainfall and Gill et al. (2015) with the Indian summer monsoon. Moreover, the correlation maps of PCs and geopotential height at 500 mb show the association of Indonesian rainfall and the wave train PNA pattern.

Significant peak of 2 - 8 year period and decadal band of 8 - 16 year period are shown in PC1, PC2 and NINO3.4 and active in different time epochs. Spectral coherences between PC1 and NINO3.4 (pre-1940 and post-1980) and between PC2 and PDO (∼1950 - 1980) in those two bands evoke that ENSO modulates Indonesian rainfall in the inter-annual and inter-decadal time scales but with different temporal variability. These strong coherences are postulated as the potential reason for the significant strengthening influence of NINO3.4 on Indonesian rainfall in recent decades (post-1980) from the BDLM analysis. These insights into multi-decadal variability of space-time rainfall over Indonesia and the temporal variability in the strengths of ENSO and PDO links to rainfall will be of significant use in developing skillful rainfall forecasting methods and in resource planning and management. Furthermore, the BDLM provides a flexible regression approach to incorporate predictors with varying strength as the model parameters are estimated dynamically at each time thus enabling to capture time varying strengths of the predictors.

The insights from this study can be used to develop seasonal rainfall forecasting models. Especially, the asymmetry in the rainfall - ENSO teleconnections during the dry season suggests the need for different approaches during El Niño and La Niña years, which can potentially enhance the forecast skills. The wavelet analysis suggesting multi-decadal variability can be used for simulating space-time rainfall patterns (e.g., Nowak et al., 2011) that can consequently, generate hydrologic patterns. Physical and human infrastructure in Indonesia suffer from floods and droughts, thus skillful hydroclimate projections of multi-decadal space-time variability will help
public policy makers to develop effective mitigation strategies at short and long time scales.

Acknowledgment

This study was funded by The Directorate General of Higher Education, The Ministry of National Education, Indonesian (Dirjen Dikti), via a Dikti Scholarship awarded to the first author. We thank three anonymous reviewers for their comments which significantly improved the manuscript.
Chapter 2

Modeling the Hydrologic Processes of Java Island, Indonesia

This research has been submitted to Journal of Hydrology: Regional Studies and is under review.

Abstract

We modeled the hydrologic processes in six watersheds of Java Island using the Variable Infiltration Capacity (VIC) land surface model. This modeling effort focused primarily on (i) collating hydrometeorological data from the watersheds, (ii) extensive quality control procedures, and (iii) benchmarking monthly and seasonal hydrological simulation skill. Six watersheds with at least seven years of physically consistent data were selected. Due to relatively small area of the watersheds with consistent data, VIC was essentially run as a lumped model except over the largest watershed. Sensitive soil parameters of the model were identified through Monte Carlo simulation and optimized using an automated calibration procedure. In the calibration period, the model performance was generally good with Nash Sutcliffe Efficiency (NSE) scores ranging from 0.50 to 0.90. The NSE of the validation period was lower than expected, ranging from 0.09 to 0.45. Our analysis attributed this performance drop to data quality, spatial resolution and model over-fitting during the calibration period. The magnitude and variability of baseflow and direct runoff components were found to modulate the model performance. The model exhibited better performance during wet years compared to dry. We provided preliminary evidence that performance can be improved by refining model spatial resolution. This work represents an important first step
towards a framework for skillful hydrologic projection system, important for resources management in the populous island of Java.

2.1 Background

The Indonesian island of Java, where 160 million of the country’s 241 million inhabitants reside, experiences water related natural disasters - i.e., floods, droughts and landslides. In the last decade more than 50% of Indonesian hydro-climatic disasters occurred in Java (BNPB, 2015). Yet, while other main islands (Sumatera, Kalimantan, Sulawesi, West Papua, Nusa Tenggara and Maluku) experience water surplus, annual water demand on Java exceeds supply by an average of ∼69 billion cubic meters (BCM) (BAPPENAS, 2010). With a large population, these water related hazards have caused severe social and economic impacts.

Rainfall variability is the primary driver of streamflow in tropical regions and understanding the underlying streamflow generation mechanisms is essential for making predictions. Rainfall-runoff models have been used to model and forecast streamflow (e.g., Regonda et al., 2013; Wood et al., 2002, 2005). Land surface models (LSMs) have been developed to enable coupling with the atmosphere and may also be used to model hydrologic processes. These are typically physically-based models that solve the coupled energy and water balance, more recently focusing on accurate simulation of surface water budget components, particularly streamflow (Livneh et al., 2011; Niu et al., 2011; te Linde et al., 2008). These have been extensively used to model streamflow in catchments with different hydro-climatic regimes around the world (Arnold et al., 1999; Beck et al., 2013; Livneh and Lettenmaier, 2013; te Linde et al., 2008). To date, the existing studies of physically-based hydrologic models on Java have only focused on single rivers; these include lumped streamflow and sediment modeling on the Lesti River (∼381 km²) in East Java (Apip et al., 2012) and streamflow modeling on the Upper Citarum Basin (∼1821 km²) in West Java (Harlan et al., 2010; Julian et al., 2013).

With the above motivation, we apply the Variable Infiltration Capacity (VIC) LSM (Liang et al., 1994) to simulate watershed hydrology in Java in this paper. The VIC model has been
widely applied in numerous hydrological studies across a range of hydro-climatic environments (Nijssen et al., 2001, 2014; te Linde et al., 2008; Liang et al., 1994, 1996; Sheffield and Wood, 2007; Sheffield et al., 2012; Shukla et al., 2013; Zhao et al., 2011). The VIC model has also been used to simulate global soil moisture and drought severity (Sheffield and Wood, 2007; Sheffield et al., 2012). The first goal of this study is to collate and describe the available hydrometeorological data and to apply a screening procedure to deal with data scarcity and quality issues. The second goal is to evaluate VIC model simulations of seasonal hydrology and its variability, consisting of model calibration and validation in unique time periods. Given the relatively coarse spatiotemporal model implementation, we further explore the role of scale on hydrologic performance in the largest study watershed. Overall, this analysis aims to identify key processes and data issues, representing an important first step towards building a sub-seasonal and multi-decadal hydrologic projection system geared towards water resources management and natural disaster mitigation in Java.

The paper is organized as follows. The study region and data sources are first described, with essential data screening steps, resulting in a set of watersheds for analysis in this study. The model is then briefly described together with the sensitive parameters used for calibration. The results section describes model performance in calibration and validation modes; seasonal and temporal performance of the model and insights into the physical processes and model performance at higher spatial resolution. The paper concludes with summary and discussion of the results.

2.2 Study Region and Data Sources

Java is located in the southwestern part of Indonesia. Its mountains have decreasing slope from the center to the south and north coasts (Figure 2.1a). The mountain range is made up of a series of quasi-circular volcanoes in the central mountains with the exception of western Java where the mountains are continuously linked (Figure 2.1a). Consequently, precipitation over the island is drained to the north and south coastal zones through a series of basins. The spatial and temporal variability of rainfall in Java is influenced by the movement of the Inter Tropical Convergence Zone (ITCZ) across the equator, creating distinct dry (May - Oct) and wet (Nov - Apr) seasons (Qian
et al., 2010). In addition, the inter-annual variability of rainfall during the dry season in Java is comparatively larger than during the wet season (Yasunari, 1981). Moreover, the spatiotemporal variability of rainfall in this island is associated with atmospheric circulation patterns driven by the El Niño Southern Oscillation (ENSO). Anomalously low rainfall years correlate well with the warm phase of ENSO while cold phases of ENSO correlate with anomalously high rainfall years (Gutman et al., 2000). The interannual and multidecadal variability of Indonesian rainfall is documented in our recent study (Yanto et al., 2016c). The land cover of the island consists mainly of tropical forest, as well as agriculture and urban settlements with significant spatial variability (Figure 2.1b). Table 2.1 summarizes the hydro-climatic environment of Java.

To apply the VIC model for Java we specifically need soil parameters, vegetation, and meteorology inputs, each from disparate sources. Soil parameters and vegetation data were obtained from Food and Agriculture Organization (FAO, 1998) and the University of Maryland’s 1 km Global Land Cover product (Hansen et al., 2000) respectively, which are the same as those used in previous VIC modeling studies (e.g., Nijsen et al., 2014; Shukla et al., 2013). Four meteorological fields: daily total precipitation, maximum and minimum temperature, and average wind speed on a 0.5° x 0.5° grid were obtained from Nijsen et al. (2014). Forcing fields from Sheffield et al. (2006) were downscaled by Nijsen et al. (2014) from a 1.0° x 1.0° to a 0.5° x 0.5° grid and data extended up to 2012. Sheffield et al. (2006) developed the meteorological fields by combining a suite of global observation-based datasets, Climatic Research Unit Time Series 2.0 (CRU TS2.0), with the National Centers for Environmental Prediction(NCEP)-National Center for Atmospheric Research (NCAR) reanalysis. Precipitation was disaggregated in space to 1.0° x 1.0° grid by statistical relationships developed with the Global Precipitation Climatology Project (GPCP) daily product Huffman et al. (2009).

In most cases, the grid size used in this study is larger than the watershed area, which we acknowledge may oversimplify key hydrologic processes. However, this study represents a first step towards capturing the broad, island-scale hydrology and land-atmosphere response requiring a critical evaluation of available data sources and an assessment of potential model performance.
improvements through calibration. Observed streamflow at 20 locations (one per watershed) was obtained from the Center of Water Resources Research and Development, Ministry of Public Works, Indonesia.

Figure 2.1: Topography (a) and land cover (b) of the Java island. Mountain ranges extend across the middle of island from west to east where agricultural areas dominate.

Table 2.1: Hydro-climatic characteristic of Java.

<table>
<thead>
<tr>
<th>Component</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Annual rainfall</td>
<td>Spatial:</td>
</tr>
<tr>
<td></td>
<td>• 3300 mm in southwest</td>
</tr>
<tr>
<td></td>
<td>• 2000 mm in northeast</td>
</tr>
<tr>
<td></td>
<td>Temporal</td>
</tr>
<tr>
<td></td>
<td>• 655 mm in dry season</td>
</tr>
<tr>
<td></td>
<td>• 1800 mm in wet season</td>
</tr>
<tr>
<td>Temperature</td>
<td>28 °C, coastal plains, 26 °C inland and mountain area, 23 °C higher mountain region</td>
</tr>
<tr>
<td>Land cover</td>
<td>61% agricultural area, 25% forest, 10% settlement, 4% water body</td>
</tr>
</tbody>
</table>
2.3 Methodology

2.3.1 Quality Control: Screening of Watershed Data

Most tropical hydrological studies are data limited (Cook et al., 1998), often with access to only poor quality data. A key challenge in this study is that streamflow data are provisional whereby data quality information is unavailable. The first step therefore was to screen the study region for watersheds that possess physical consistency between precipitation and streamflow with at least 7 years of streamflow, from among the 20 potential watersheds shown in Figure 2.2. To this end, we applied the following quality control steps:

1. **Precipitation Consistency.** We compared the spatiotemporal consistency between monthly rainfall used in this study (Nijssen et al., 2014), with that from the CRU TS3.22 (Harris et al., 2013). At monthly time scales the CRU TS3.22 rainfall dataset has similar spatial and temporal patterns with station rainfall data from the Bureau of Meteorology, Climatology and Geophysics of Indonesia (Yanto et al., 2016c). However, CRU TS3.22 is not available at a daily time scale - which is necessary in our application here, hence we used the data of Nijssen et al. (2014). To check the consistency of Nijssen et al. (2014) data, we computed the coefficient of determination \( R^2 = \text{square of the standard Pearson correlation coefficient} \) of the monthly rainfall between these two products for each grid - shown in Figure 2.2. It can be seen that \( R^2 \) is greater than 0.5 over most of the island (Figure 2.2) indicating that the Nijssen et al. (2014) data is consistent with station data.

2. **Streamflow Length.** For stable model calibration, we selected watersheds with at least 7 years of streamflow data, as mentioned above, and less than one month of missing data.

3. **Physical Consistency Check.** The consistency between rainfall and streamflow data was investigated using the long-term water balance and relationship among climatological features. The climatology of precipitation and streamflow for each watershed is shown in Figure 2.3 along with their respective runoff ratio (RR),
\[ RR = \frac{\bar{Q}}{\bar{P}} \]  

where \( \bar{Q} \) and \( \bar{P} \) are streamflow and rainfall respectively. Boxplots of rainfall and streamflow are shown in Figure 2.4 to illustrate the long-term mass balance and data distribution. In humid tropical environments across a range of land cover types, the runoff ratio varies from \( \sim 10\% \) to \( \sim 60\% \) (Dettinger and Diaz, 2000; Dubreuil, 1985; Muñoz-Villers and McDonnell, 2012). The runoff ratio can be used to characterize the soil parameters and vegetation cover of a watershed. A watershed with impervious surfaces (e.g. settlement, pavement, roads) will have a higher runoff ratio than one with more pervious surfaces (e.g. forest, agriculture). Considering the dominant land cover of Java (Table 2.1), we selected watersheds with similar patterns of rainfall and streamflow climatology, similar long-term mass balance, and runoff ratio less than 60%. Based on these criteria, we selected CT, CL, PG, BG, GD and BD shown in color in Figure 2.5 and listed in Table 2.2. Given the small area of many of the selected watersheds in this study, the VIC model essentially represents a lumped implementation in 5 out of 6 watersheds. The largest watershed (BG) receives partial contributions from 8 grid cells, such that the application of the model in this basin can be considered spatially distributed.

![Figure 2.2: Spatial map of R² value between Sheffield rainfall and CRU TS3.22 rainfall in the period of 1948 - 2012. Red and blue numbers indicate R² > 0.5 and R² < 0.5 respectively.](image-url)
Figure 2.3: Climatology of rainfall and streamflow of 20 watersheds covering Java, each of which has at least 7 years of data. A runoff ratio (RR) is computed for each watershed as mean precipitation divided by mean flow over the entire period of record and shown in the panels. Watersheds selected for this study are indicated in red.
Figure 2.4: Boxplot of monthly precipitation and streamflow for the 20 watersheds covering Java. The period and length of record for each watershed are different as presented in Table 2.2. Watersheds selected for this study are marked in red.
Figure 2.5: The 20 watersheds covering Java with at least 7 years of data. The six watersheds shown in red were selected for this study after the screening process.

Table 2.2: Characteristics of the six watersheds over Java selected for this study. Land cover type A, F, S and W represent agriculture, forest, settlement and water body respectively with bold font indicating the dominant land cover.

<table>
<thead>
<tr>
<th>Watershed</th>
<th>Outlet Lon &amp; Lat (°)</th>
<th>Outlet Mean Elevation (m)</th>
<th>Streamflow record period (mm/yr)</th>
<th>Area (km²)</th>
<th>Mean Mean Rainfall (mm/yr)</th>
<th>Land cover</th>
<th>Runoff ratio (Q/P)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Citarum (CT)</td>
<td>107.53/-6.95</td>
<td>668</td>
<td>1351</td>
<td>1990-1997</td>
<td>1675</td>
<td>A(59), F(27), S(13), W(1)</td>
<td>0.50</td>
</tr>
<tr>
<td>Cilangla (CL)</td>
<td>108.11/-7.63</td>
<td>249</td>
<td>1875</td>
<td>1975-1984</td>
<td>177</td>
<td>A(74), F(24), S(1), W(1)</td>
<td>0.56</td>
</tr>
<tr>
<td>Progo (PG)</td>
<td>110.26/-7.67</td>
<td>180</td>
<td>1115</td>
<td>1996-2002</td>
<td>1676</td>
<td>A(15), F(79), S(5), W(1)</td>
<td>0.48</td>
</tr>
<tr>
<td>Bengawan (BG)</td>
<td>112.17/-7.10</td>
<td>11</td>
<td>885</td>
<td>1990-2002</td>
<td>16286</td>
<td>A(56), F(34), S(9), W(1)</td>
<td>0.44</td>
</tr>
<tr>
<td>Grindulu (GD)</td>
<td>111.14/-8.14</td>
<td>39</td>
<td>1094</td>
<td>1990-2002</td>
<td>556</td>
<td>A(58), F(42), S(0), W(0)</td>
<td>0.56</td>
</tr>
<tr>
<td>Bedadung (BD)</td>
<td>113.58/-8.23</td>
<td>44</td>
<td>1322</td>
<td>1991-2001</td>
<td>696</td>
<td>A(21), F(74), S(5), W(0)</td>
<td>0.50</td>
</tr>
</tbody>
</table>

2.3.2 Hydrologic Model

VIC (Liang et al., 1994) is a physically-based, fully-distributed model that independently solves the energy and water balances at the land surface. Subgrid-scale variability is user specified for land cover, soil texture, and soil moisture storage capacity. Two canopy layers represent the interaction of moisture with vegetation - i.e. evapotranspiration (ET), interception and throughfall, while the uppermost soil layer characterizes the dynamic response of soil to variable infiltration
rates of incoming rainfall. The bottom soil layer is used to coarsely represent groundwater and baseflow processes. The middle soil layer contributes diffusion of water to uppermost layer when the middle layer is wetter. Details about VIC model structure and formulations can be found in Liang et al. (1994, 1996).

In this study we used a standard implementation of 3 soil layers and 2 root zones such that upward transport of moisture from roots is represented by two top soil layers. The model was run in water balance mode using 24-hour (i.e. daily) time step on a 0.5° x 0.5° grid. The model was forced with daily precipitation, maximum and minimum temperature and wind speed.

2.3.3 Parameter Estimation - Calibration and Validation

Model calibration is a process of estimating parameters that most accurately simulate observed behavior (Refsgaard, 1997). We followed standard calibration practice, under the assumption that observations were error free (Moriasi et al., 2007). Parameter calibration and validation were conducted on a monthly time step comparing VIC simulations with historic streamflow observations. The Nash-Sutcliffe Efficiency (NSE) metric was used to evaluate the performance of model parameterizations (Equation 2). Although the NSE tends to be sensitive to extreme values (Gupta et al., 2009; Legates and McCabe, 1999), it is widely used as a representative function to quantify the overall fit of a hydrograph (Moriasi et al., 2007; Sevat and Dezetter, 1991). According to Nash and Sutcliffe (1970), NSE is defined as:

\[
NSE = 1 - \frac{\sum_{t=1}^{T} (Q_{o}^t - Q_{s}^t)^2}{\sum_{t=1}^{T} (Q_{o}^t - \bar{Q}_{o})^2}
\]  

(2.2)

where \(Q_{o}^t\) and \(Q_{s}^t\) are observed and simulated streamflow at time \(t\) respectively, \(\bar{Q}_{o}\) is the mean of observed streamflow. The range of NSE extends from \(-\infty\) to 1 and NSE < 0 indicates that the mean flow is a better predictor than the model simulations. Following Moriasi et al. (2007) the model performance at monthly time step is generally judged as satisfactory when NSE > 0.50, however "good" performance requires NSE > 0.65.

Identifying parameters is critical in hydrologic modeling. Fitting the model using too many
parameters (over-parameterization) increases predictive uncertainties associated with them (Jakeman and Hornberger, 1993). Beven (1989) suggested that 3 to 5 parameters in physically-based models should suffice to effectively model a range of hydrologic processes. Accordingly, we performed preliminary analysis to identify the most sensitive parameters from among the broader set of eight soil parameters: baseflow parameters ($Ds$, $Ds_{max}$, $Ws$), infiltration shape parameter $b$, exponent used in base flow curve $c$ and the thickness of soil layer ($thick_1$, $thick_2$ and $thick_3$). To obtain the set of most sensitive model parameters, we conducted the following steps:

1. We generated 2000 sets of the above eight parameters using a Monte Carlo approach, assuming each parameter is uniformly distributed within the respective range of values suggested by Yang et al. (2010).

2. The VIC model was driven using these soil parameter sets and the NSE was computed for each set; this was done for each of the six study watersheds.

3. We produced a scatterplot of each soil parameter and the corresponding NSE for each watershed in Figure 2.6.

Sensitivity was assessed using visual inspection of these scatterplots (Wagener et al., 2001). A sensitive parameter will exhibit noticeable changes in its NSE values as a function of changes in the soil parameter. For example, the $c$ parameter in Figure 2.6 has a "flat" response, suggesting that the NSE is not sensitive to changes in the value of $c$. In contrast, a parameter such as $Ds$ exhibits noticeable nonlinear relationship with NSE - across all the watersheds - indicating that changes in value of $Ds$ result in substantial changes in NSE. After examining Figure 2.6, parameter $b$, $Ds$, $thick_2$ and $thick_3$ were determined to be the most sensitive and hence, selected for calibration. This result is in agreement with a previous, more exhaustive sensitivity analysis (Demaria et al., 2007) that also found $b$ and $thick_2$ to be the most sensitive soil parameters. However, an important methodological limitation of this analysis is the temporally coarse evaluation metric (monthly NSE), which is likely to be largely insensitive to parameters such as $'c'$, that governs shorter duration processes such as fast runoff or baseflow. The description of calibration parameters is presented in Table 2.3.
Figure 2.6: Scatterplots of NSE vs soil parameter values from a Monte Carlo sampling of the parameter space. A local polynomial smoother through the scatterplot is shown as solid line. Results are used to determine sensitivity of NSE to changes in oil parameters, and choose soil parameters for model calibration.
Table 2.3: Description of calibration parameters, their ranges and the values obtained from optimization for the six watersheds.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>$b$</th>
<th>$D_s$</th>
<th>thick$_2$</th>
<th>thick$_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description</td>
<td>infiltration curve shape parameter</td>
<td>fraction of maximum baseflow where nonlinear baseflow begins</td>
<td>thickness of soil layer 2</td>
<td>thickness of soil layer 3</td>
</tr>
<tr>
<td>Unit</td>
<td>NA</td>
<td>NA</td>
<td>m</td>
<td>m</td>
</tr>
<tr>
<td>Range</td>
<td>$&gt;0 - 1$</td>
<td>$&gt;0 - 1$</td>
<td>0.1 - 3</td>
<td>0.1 - 3</td>
</tr>
<tr>
<td>Optimal value</td>
<td>0.28</td>
<td>0.55</td>
<td>0.40</td>
<td>2.63</td>
</tr>
<tr>
<td></td>
<td>0.52</td>
<td>0.37</td>
<td>0.30</td>
<td>1.27</td>
</tr>
<tr>
<td></td>
<td>0.43</td>
<td>0.23</td>
<td>3.00</td>
<td>3.00</td>
</tr>
<tr>
<td></td>
<td>0.44</td>
<td>0.88</td>
<td>0.61</td>
<td>0.69</td>
</tr>
<tr>
<td></td>
<td>0.34</td>
<td>0.99</td>
<td>0.30</td>
<td>1.31</td>
</tr>
<tr>
<td></td>
<td>0.32</td>
<td>0.89</td>
<td>1.71</td>
<td>1.59</td>
</tr>
</tbody>
</table>

After the sensitivity analysis was completed, we conducted the calibration using the four most sensitive parameters illustrated in Figure 2.6. When setting up the calibration procedure, we sought to use an automatic procedure that is typical of the state of the practice in hydrologic modeling: an automatic simulation-optimization approach that is connected directly to the VIC model. We selected the Borg Multiobjective Evolutionary Algorithm (Borg MOEA) (Hadka and Reed, 2013) which has been applied to calibration problems previously (Reed et al., 2013), and features diversity throughout search, adaptive population sizing and local optima escape facility. The Borg MOEA was used in a single objective mode for model calibration in this study to provide a simple and straightforward calibration procedure that yields a single best parameterization for each watershed. The four selected model parameters were estimated by optimizing the NSE over the first half ($\sim$5 years) of the streamflow record (Table 2.3). To evaluate the calibrated model in a predictive mode - i.e. validation (Refsgaard, 1997) - the model was validated against the second half of the record and the results are presented in the following section.
2.4 Result

In this section we present the model performance, infer processes using hydrologic metrics and investigate the variability in performance relative to climate regimes (wet and dry) and model spatial resolution.

2.4.1 Model performance

The observed and simulated streamflow for each watershed in the calibration (Figure 2.7, left column) and validation period (Figure 2.7, right column) are shown as time series and scatterplots along with the NSE values. In the calibration period, NSE values range from 0.50 to 0.90, with five of them exceeding 0.60, indicating that the calibrated model performance in all the watersheds ranges between satisfactory and very good (Moriasi et al., 2007). The simulations consistently fit the rising limb, peak and recession limb of the observed hydrograph during the calibration period (Figure 2.7, left panel). In all watersheds, the validation NSE values are unsatisfactory ranging from 0.09 to 0.45. Although the validation NSE values are expected to be lower, the marked drop of NSE values (especially in the four watersheds: CL, PG, BG and BD) is unexpected.
Figure 2.7: Time series and scatterplots of observed and simulated streamflow and rainfall hyetograph for CT, CL, PG, BG, GD and BD watersheds. The left panels show the calibration period and the right shows validation. In the time series plots, left and right y-axes are for streamflow and rainfall respectively. The boxes in the validation figures show the periods where the greatest errors occur, typically manifested in the lack of a corresponding observed hydrograph peak with precipitation input, or where there is a lag between monthly rainfall and observed streamflow peaks.

To reconcile the low NSE scores in validation, we compared the observed rainfall hyetograph with the observed hydrograph during those periods where they do not correspond well - shown in
boxes in Figure 2.7, right panel - particularly in watershed BD where in 3 out of 5 years in the validation period the observed streamflow peak lags the rainfall. This also occurs in watershed CT during 1990 - 1992 and in watershed CL during 1978 - 1979. These lags between monthly rainfall and streamflow could be the result of upstream flow regulation: we note that only the BG watershed has known regulation upstream of the gauge. Alternatively, disparity could result from cases where extreme events occur close to the start or end of the month - i.e. a high rainfall event may occur at the end of the month and high streamflow is recorded and reported in the following month. Lastly, these disparities could result from measurement errors. In all these cases, VIC produces high streamflow in correspondence with the peak in rainfall, but the observed flow remains low, thus resulting in poor skill.

The model soil and vegetation parameters remain unchanged over the simulation period. It is therefore expected that the runoff ratio between calibration and validation periods should be similar, i.e. constant hydrologic elasticity. However, this is not the case as demonstrated in Figure 2.7. This discrepancy certainly contributes to the significant drop in model performance in the validation period in watersheds CL, PG and BG. To understand the cause of this difference, we calculated the change in mean rainfall and streamflow between the calibration and validation periods (Table 2.4). The percent changes between rainfall and streamflow differ substantially in almost all the watersheds. In particular, watersheds CL, PG, BG and GD exhibit more than a 10% change in runoff ratio. This could be due to shifts in climate, or conversely due to errors in streamflow measurement, and/or biases in the gridded rainfall as is likely the case in watershed GD in December 2000 (Figure 2.7) where the observed streamflow exceeds the maximum rainfall input. VIC cannot adequately capture this change in elasticity through time as is apparent in the validation. Overall, the mechanism responsible for the change in elasticity through time cannot be conclusively attributed to either climate or sensor errors exclusively.
Table 2.4: The change of mean rainfall and streamflow in the calibration and validation periods. The change in runoff ratio is computed as percentage change of runoff ratio in the validation period from it is in the calibration period.

<table>
<thead>
<tr>
<th>Location</th>
<th>Monthly mean streamflow (mm)</th>
<th>Change in runoff ratio (%)</th>
<th>Monthly mean rainfall (mm)</th>
<th>Change in runoff ratio (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Calibration</td>
<td>Validation</td>
<td></td>
<td>Calibration</td>
</tr>
<tr>
<td>CT</td>
<td>116.20</td>
<td>103.65</td>
<td>-10.80</td>
<td>243.40</td>
</tr>
<tr>
<td>CL</td>
<td>150.60</td>
<td>112.90</td>
<td>-25.03</td>
<td>269.88</td>
</tr>
<tr>
<td>PG</td>
<td>99.28</td>
<td>85.97</td>
<td>-13.41</td>
<td>216.32</td>
</tr>
<tr>
<td>BG</td>
<td>71.62</td>
<td>68.70</td>
<td>-4.08</td>
<td>154.26</td>
</tr>
<tr>
<td>GD</td>
<td>85.30</td>
<td>121.61</td>
<td>42.57</td>
<td>167.46</td>
</tr>
<tr>
<td>BD</td>
<td>94.88</td>
<td>115.61</td>
<td>21.85</td>
<td>196.42</td>
</tr>
</tbody>
</table>

Soil moisture content exerts an important control on the amount of rainfall partitioned into direct runoff, baseflow and ET with higher soil moisture content corresponding with greater direct runoff. Important changes in soil moisture content have been noted from daily to annual time scales in this region (Krave et al., 2007; Kumagai et al., 2009). It is therefore of interest to understand the seasonal performance of VIC related to the state of the system, wet versus dry. We computed NSE for the dry and wet seasons separately. Dry and wet seasons were defined as the months with rainfall below and above the annual mean respectively Yanto et al. (2016c) defined as Nov - Apr for wet and May - Oct for dry season. Table 2.5 shows NSE in dry and wet seasons for the six watersheds. The model performance for the western watersheds CT, CL and PG - appears to be consistent in both the seasons, but there exists a substantial difference in model performance among the eastern watersheds (BG, GD and BD). We suspect this is due to the nature of the NSE whereby model skill can be over-estimated particularly in basins with high seasonal variability such as in Java (Gupta et al., 2009; Schaefli and Gupta, 2007).
Table 2.5: NSE values for all watersheds in dry and wet season.

<table>
<thead>
<tr>
<th>Watershed</th>
<th>Dry season NSE</th>
<th>Wet season NSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>CT</td>
<td>0.24</td>
<td>0.24</td>
</tr>
<tr>
<td>CL</td>
<td>0.89</td>
<td>0.69</td>
</tr>
<tr>
<td>PG</td>
<td>0.52</td>
<td>0.58</td>
</tr>
<tr>
<td>BG</td>
<td>0.33</td>
<td>0.86</td>
</tr>
<tr>
<td>GD</td>
<td>0.74</td>
<td>0.46</td>
</tr>
<tr>
<td>BD</td>
<td>-3.10</td>
<td>0.62</td>
</tr>
</tbody>
</table>

Although no direct measurements exist of ET, it is worthwhile exploring the characteristics of simulated ET since it is an important component of the water balance. We compared the VIC-simulated ET with the Moderate Resolution Imaging Spectroradiometer (MODIS) MOD16 global ET product (Mu et al., 2007) for the overlapping time period 2000 - 2010. Validations of 8-day MOD16 ET with flux tower measurements across climate regions have shown underestimation of high ET and overestimation of low ET values on the order of 20% (Kim et al., 2012; Ruhoff et al., 2013; Tang et al., 2015). Ruhoff et al. (2013) identified the misclassification of land cover as the major source of error in the MOD16 algorithm. Furthermore, since MOD16 ET estimates are not constrained by a water balance they can result unrealistically high values during dry conditions (Livneh and Lettenmaier, 2012; Ferguson et al., 2010). Figure 2.8 shows the standardized monthly ET anomalies from VIC and MOD16. We standardized the data by subtracting the period mean and dividing by the respective standard deviation, as it was expected that the dynamic range of the two products may differ given the lack of water (mass) balance constraint on the MOD16 product. Model simulations reproduce the temporal pattern of the MOD16 ET well except in basins CT and CL. MOD16 ET tends to be an overestimate for higher values and an underestimate for lower values compared to VIC ET. MOD16 annual ET in basins CT, CL, BG, GD and GD is higher than VIC ET by 8%, 10%, 14%, 19% and 5% respectively. This level of error in MOD16 is therefore in agreement with aforementioned findings (Kim et al., 2012; Ruhoff et al., 2013; Tang et al., 2015) limiting the ability to critically evaluate VIC ET here.
To understand the mismatch between the two ET estimates for basins CT and CL we investigated the season of minimum ET for each watershed. The minimum simulated ET consistently occurs between July and October for all watersheds which is consistent with MOD16 in PG, BG, GD and BD and also follows the temporal distribution of temperature. On the other hand, minimum ET from MOD16 in basins CL and CT frequently occurs between January and April, which is out of phase with temperature. Our analysis suggests that it is more likely this result reflects an issue with MOD16 data rather than with ET simulations from VIC which is unlikely to produce inconsistency between temperature and ET.
2.4.2 Variability of water balance and hydrologic processes

This section explores key characteristics of the simulated hydrology including the partitioning of baseflow and direct runoff. Understanding the contributions from these streamflow components and the streamflow generation process in general is important for skillful hydrologic prediction (Eckhardt, 2008; Gonzales et al., 2009).

To understand the spatial variability (i.e. west to east) of streamflow generation processes in the selected watersheds, we compared coefficient of variation (CV) of monthly streamflow, direct runoff and baseflow, shown in Table 2.6. In addition, we computed the Base Flow Index (BFI: the ratio of baseflow to total streamflow in the calibration period) to isolate the roles of soil and geology (Beck et al., 2013; Longobardi and Villani, 2008). In Table 2.6 it is shown that rainfall variability, indicated by the CV of precipitation, increases moving eastward on the island. The temporal variability of surface runoff in all the watersheds is higher than their corresponding rainfall. The higher temporal variability of runoff compared to rainfall is a consequence of streamflow dependence on soil moisture content and soil permeability (Muñoz-Villers and McDonnell, 2012). Rainfall-runoff response in humid tropical mountainous areas is expected to be dominated by baseflow from within the hillslopes (Muñoz-Villers and McDonnell, 2012). This is confirmed by lower CV values of baseflow in comparison to that of surface runoff and precipitation. The variability in rainfall, direct runoff and baseflow show a general increase from west to east over the island. The ratio of CV of direct runoff and baseflow is 2.43, 1.61, 2.37, 0.97, 1.12 and 1.50 for the watersheds CT, CL, PG, BG, GD and BD respectively. As shown in Table 2.6, this pattern is consistent with that of the calibrated thickness of the lowest soil layer, $thick_3$. Also, both the ratio of CV of direct runoff and baseflow and $thick_3$ are highly correlated, exhibiting a $R^2$ score of 0.90.
Table 2.6: Water balance and calibrated soil parameters in the selected watersheds organized from west to east. The water balance was computed for calibration period. BFI is calculated as ratio of long-term baseflow to total streamflow.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Unit</th>
<th>CT</th>
<th>CL</th>
<th>PG</th>
<th>BG</th>
<th>GD</th>
<th>BD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water balance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rainfall</td>
<td>Total (mm yr⁻¹)</td>
<td>2850</td>
<td>3217</td>
<td>2513</td>
<td>1809</td>
<td>1960</td>
<td>2264</td>
</tr>
<tr>
<td></td>
<td>Coefficient of Variation (CV)</td>
<td>0.80</td>
<td>0.77</td>
<td>0.84</td>
<td>0.90</td>
<td>0.91</td>
<td>1.06</td>
</tr>
<tr>
<td>Evaporation</td>
<td>Total (mm yr⁻¹)</td>
<td>1372</td>
<td>1434</td>
<td>1370</td>
<td>965</td>
<td>1043</td>
<td>1138</td>
</tr>
<tr>
<td></td>
<td>Coefficient of Variation (CV)</td>
<td>0.31</td>
<td>0.23</td>
<td>0.25</td>
<td>0.58</td>
<td>0.55</td>
<td>0.39</td>
</tr>
<tr>
<td>Direct runoff</td>
<td>Total (mm yr⁻¹)</td>
<td>545</td>
<td>930</td>
<td>571</td>
<td>433</td>
<td>414</td>
<td>489</td>
</tr>
<tr>
<td></td>
<td>Coefficient of Variation (CV)</td>
<td>1.02</td>
<td>1.12</td>
<td>0.96</td>
<td>1.06</td>
<td>1.18</td>
<td>1.37</td>
</tr>
<tr>
<td>Baseflow</td>
<td>Total (mm yr⁻¹)</td>
<td>909</td>
<td>775</td>
<td>560</td>
<td>404</td>
<td>504</td>
<td>705</td>
</tr>
<tr>
<td></td>
<td>Coefficient of Variation (CV)</td>
<td>0.42</td>
<td>0.69</td>
<td>0.41</td>
<td>1.09</td>
<td>1.05</td>
<td>0.91</td>
</tr>
<tr>
<td>Baseflow index</td>
<td>%</td>
<td>62</td>
<td>46</td>
<td>49</td>
<td>48</td>
<td>55</td>
<td>59</td>
</tr>
<tr>
<td>Soil parameter</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b</td>
<td></td>
<td>0.28</td>
<td>0.52</td>
<td>0.43</td>
<td>0.44</td>
<td>0.34</td>
<td>0.32</td>
</tr>
<tr>
<td>Ds</td>
<td></td>
<td>0.55</td>
<td>0.37</td>
<td>0.23</td>
<td>0.88</td>
<td>0.99</td>
<td>0.89</td>
</tr>
<tr>
<td>thick₂</td>
<td>m</td>
<td>0.40</td>
<td>0.30</td>
<td>3.00</td>
<td>0.61</td>
<td>0.30</td>
<td>1.71</td>
</tr>
<tr>
<td>thick₃</td>
<td>m</td>
<td>2.63</td>
<td>1.27</td>
<td>3.00</td>
<td>0.69</td>
<td>1.31</td>
<td>1.59</td>
</tr>
</tbody>
</table>

The distribution of BFI over the watersheds shows a low degree of variability (~0.45 - 0.60) - consistent with the fact that the watersheds all have similar geology. In addition, the BFI is highly correlated with 1/b (R² of 0.99) - a key variable in the infiltration capacity equation used by the model (Liang et al., 1994). This intuitive result confirms that the soil characteristics in the top layer have a strong influence in transferring the rainfall into both direct runoff and baseflow and consequently the BFI.

### 2.4.3 Sensitivity to climate regimes

Understanding the hydrologic response of watersheds to different climate regimes, here defined as the inter-annual variability of precipitation forcings, is of great interest especially as it would pertain to managing resources in a changing climate. We compared the simulated flow duration
curves of surface runoff and baseflow for each watershed for the driest, average and wettest years, shown in Figure 2.9. These years were selected on the basis of annual rainfall in each watershed. Runoff and baseflow are deemed sensitive if their flow duration curves in the driest and wettest year are clearly separated from the average. Figure 2.9 demonstrates that surface runoff tends to be more sensitive to climate variability (except CT) than baseflow, particularly during high flow periods. High flows occur during the wet season when the soil moisture content is typically high (Krave et al., 2007), thereby quickly saturating the soil during rainfall events and resulting in rapid overland flow. In comparison baseflow is less sensitive, with the eastern basins (BG, GD and BD) showing slightly higher sensitivity than western basins (CT, CL and PG) during the wettest year, with the opposite exhibits in the driest year. We computed the NSE scores for these three years (Table 2.7) indicating that the model performs best during the wettest and average years. During the driest year the model performance in the eastern watersheds is higher than the west.

Figure 2.9: Flow duration curve of monthly simulated streamflow in the wettest, average and driest year chosen from the calibration period of each watershed.
Table 2.7: NSE of the driest, average and wettest years in the watersheds.

<table>
<thead>
<tr>
<th>Watershed</th>
<th>Driest year</th>
<th>Average year</th>
<th>Wettest year</th>
</tr>
</thead>
<tbody>
<tr>
<td>CT</td>
<td>0.31</td>
<td>0.50</td>
<td>0.57</td>
</tr>
<tr>
<td>CL</td>
<td>0.49</td>
<td>0.80</td>
<td>0.73</td>
</tr>
<tr>
<td>PG</td>
<td>-0.08</td>
<td>0.63</td>
<td>0.48</td>
</tr>
<tr>
<td>BG</td>
<td>0.93</td>
<td>0.90</td>
<td>0.93</td>
</tr>
<tr>
<td>GD</td>
<td>0.76</td>
<td>0.60</td>
<td>0.71</td>
</tr>
<tr>
<td>BD</td>
<td>0.86</td>
<td>0.72</td>
<td>0.41</td>
</tr>
</tbody>
</table>

2.4.4 Sensitivity to spatial resolution

We reason that lower NSE scores in the validation period may be the product of the coarse model resolution of the experiments so far, which were run at a 0.5° x 0.5° spatial resolution, hereafter 1/2N. To explore this we ran VIC at a higher resolution, 0.125° x 0.125°, on the largest watershed (BG) as we felt this would provide the greatest expression of model resolution. We downscaled the 1/2N soil, and meteorology (S_{1/2}, M_{1/2}) to a 0.125° x 0.125° resolution (S_{1/8}, M_{1/8N}) following a nearest neighbor method, in which, the values of parameters S and M on the 0.125° x 0.125° grid were transferred directly from the nearest 0.5° x 0.5° grid, we refer to 1/8N. The grid elevations were obtained at the 0.125° x 0.125° resolution from a Digital Elevation Model (DEM). The 0.125° x 0.125° vegetation parameters, V_{1/8}, were gridded from the same source as Nijssen et al. (2014), the University of Maryland’s 1 km Global Land Cover product (Hansen et al., 2000). We recently created a high resolution gridded precipitation product at 0.125° x 0.125° resolution by directly gridding the station data using IDW method (Yanto et al., 2016a) referred to here as 1/8Y. We ran the VIC model with these two combinations of input soil, vegetation and meteorological forcings i.e. (S_{1/8}, V_{1/8}, M_{1/8N}) and (S_{1/8}, V_{1/8}, M_{1/8Y}) and optimized the sensitive model parameters for each combination. We computed percentage change in fluxes - precipitation, evaporation, total runoff and root zone soil moisture - and the performance metric - NSE with respect to those from the 1/2N runs.
Figure 2.10 compares the fluxes and performance metrics from the three VIC model configurations at 1/2N, 1/8N and 1/8Y configurations. The NSE values corresponding to the 1/2N are set to 0. For the 1/8N and 1/8Y configurations, the percentage change of NSE greater than 0 indicates better model performance with respect to the 1/2N and vice versa. During calibration the 1/8Y configuration shows modest improvement in performance compared to the 1/2N and 1/8N configurations. However, during validation the 1/8Y exhibits significant improvement in performance (over 150% increase in NSE). The large improvements between 1/8N and 1/8Y indicate the sensitivity to input precipitation resolution (e.g., Mobley et al., 2012; Neary et al., 2004). Total runoff in 1/8Y configuration shows a 10% increase during calibration and a considerable decrease ~150% during validation in comparison to the 1/2N configuration. The spatial distribution of annual flows from the three model runs are shown in Figure 2.10b,c,d. Figure 2.10d shows that mountainous regions in the center of the watershed contribute more flow compared than lowland areas, likely due to greater overland flow response to enhanced orographic precipitation. This feature is not seen in the 1/2N (Figure 2.10b) or 1/8N configuration (Figure 2.10c).
Figure 2.10: Comparison of fluxes, performance metrics and discharge distribution between VIC model at $0.5^\circ \times 0.5^\circ$ and $0.125^\circ \times 0.125^\circ$ resolutions. In (a), N and Y represent meteorological forcings of Nijssen and Yanto respectively, on the right axis, model performance is better when the change in NSE positive, and vice versa, (b), (c) and (d) show the spatial distribution of flow at 1/2N, 1/8N and 1/8Y respectively.

2.5 Summary and Discussion

Java is the most populous Indonesian island, plagued by flooding, water shortages, and landslides. These natural disasters underscore the need for understanding and modeling the hydrologic processes in space and time in order to aid in mitigating the events. However, there are critical challenges for hydrologic modeling in Java: complex climate dynamics near the equator, the islands diverse topography, as well as issues with scarce and limited quality hydrologic and meteorological data.

This study presented one of the first attempts to model the hydrologic processes over multiple watersheds of the Java Island, specifically using the VIC modeling system. Data were assembled for 20 watersheds and after screening for physical consistency between precipitation and streamflow as
well as missing data, six watersheds were selected for detailed study. Four sensitive soil parameters were calibrated for these watersheds. The optimal parameters were selected by maximizing NSE over the calibration period (the first half of the data period) using a single-objective optimization with the Borg MOEA. The model was validated on the second half of the data record.

In the calibration period, model performance varied from satisfactory to very good with NSE ranging from 0.50 to 0.90. The NSE of the validation period was lower than expected with skill between NSE=0.09 to NSE=0.45. Inspection of the rainfall and streamflow observations revealed data quality issues, both for precipitation and discharge, whereby streamflow peaks did not follow rainfall peaks or where peak flows greatly lagged the rainfall enough to lower the skill. Furthermore, significant differences in runoff ratio between calibration and validation periods in most of the watersheds were also observed. The marked drop in performance between calibration and validation periods is indicative of overfitting model parameters to the calibration period. The NSE for dry and wet seasons remained steady except in the watershed BD, which has a negative skill during the dry season. The dependency of model performance on the flow magnitude and soil moisture properties was deemed responsible for the seasonal variability of model performance. The BFI and 1/b, the parameter in the infiltration equation of the model, are highly correlated, indicating that the surficial soil properties play an important role in both the direct runoff and baseflow. Direct runoff is more responsive to climate variability compared to baseflow, especially in high flows. Overall, the modeling framework developed in this study offers an important first step towards systematic understanding of hydrologic processes for Java and the potential for a skillful hydrologic forecast modeling system.

Acknowledgement

This study was funded by The Directorate General of Higher Education, The Ministry of National Education, Indonesian (DirjenDikti), via a Dikti Scholarship awarded to the first author. This work utilized the Janus supercomputer, which is supported by the National Science Foundation (award number CNS-0821794) and the University of Colorado Boulder. The Janus supercomputer
is a joint effort of the University of Colorado Boulder, the University of Colorado Denver and the National Center for Atmospheric Research.
Chapter 3

Development of A Meteorological Data Set over the Java Island, Indonesia
1985 - 2014

This research is in preparation for submission to Nature Scientific Data journal.

Abstract

We describe a gridded daily meteorology data set consisting of precipitation, minimum temperature and maximum temperature over the Java Island, Indonesia at 0.125° x 0.125° (∼14 km) resolution spanning 30 years from 1985 - 2014. Importantly, this data set represents a marked improvement from existing gridded data sets over Java with higher spatial resolution, derived exclusively from ground-based observations unlike existing satellite or reanalysis-based products. Gridding was performed via Inverse Distance Weighting, IDW (radius, r, of 25 km and power of influence, α, of 3 as optimal parameters) restricted to only those stations including at least 3650 days (∼10 years) of valid data. A gap-infilling procedure was performed prior to the gridding process. We performed cross-validation on both the gap-infilling process and gridding algorithm. It shows that gridded precipitation has smaller errors than cross-validation precipitation, for both spatial and temporal performance. Visual inspection reveals an increasing performance of gridded precipitation from grid, basin to island scale. The data set, stored in a network common data form (NetCDF), is intended to support basin-scale and island-scale studies of short-term and long-term climate, hydrology and ecology.
3.1 Background & Summary

Detailed understanding of the spatiotemporal variability of natural processes (climate, hydrology, ecology) is enhanced by accurately modeling and simulating these processes (Beeson et al., 2011; Luzio et al., 2008; Yatagai et al., 2012), requiring reliable spatial and temporal estimates of meteorology (Luzio et al., 2008). Near-surface observation-based meteorological data sets are deemed as the most representative inputs for hydrologic simulations (Livneh et al., 2015; Pechlivanidis et al., 2011). However, meteorological stations are irregularly spaced (van den Besselaar et al., 2011) and often clustered around population centers, while hydrologic and ecologic simulation models require meteorological data on a quasi-continuous, regular grid, posing a challenge in representativeness. Here, we present gridded daily precipitation (P), minimum temperature (Tmin) and maximum temperature (Tmax) data for Java Island - the first of its kind, high-resolution, station-only gridded meteorology product in Indonesia.

Precipitation over Java exhibits large temporal variability in which the dry season, May - Oct, has greater variability than the wet season, Nov - Apr (Yasunari, 1981). The southwestern part of the island receives more annual rainfall than the northeastern region due to the influence of the ocean-atmospheric circulation in the Indian Ocean on the climates of southwestern Java (Yanto et al., 2016b). The existence of mountains across the island (from west to east) creates additional orographically-driven precipitation variability (Yanto et al., 2016b). The mountains form a series of small basins to the south and north of the island.

Existing gridded meteorological data products for Java include satellite-only data sets, reanalysis products, and coarse resolution gauge data sets. The details of these products are presented in Table 3.1.
Table 3.1: Meteorological data products available for Java Island.

<table>
<thead>
<tr>
<th>Source</th>
<th>Product</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>gauge data sets</td>
<td>• Climatic Research Unit Time Series (CRUTS)</td>
<td>Harris et al. (2013)</td>
</tr>
<tr>
<td></td>
<td>• Global Precipitation Climatology Centre (GPCC)</td>
<td>Schneider et al. (2014)</td>
</tr>
<tr>
<td></td>
<td>• Asian Precipitation - Highly - Resolved Observational Data Integration Towards Evaluation of Water Resources (APHRODITE)</td>
<td>Yatagai et al. (2012)</td>
</tr>
<tr>
<td>satellite-only data sets</td>
<td>• CICS High-Resolution Optimally Interpolated Microwave Precipitation from Satellites (CHOMPS)</td>
<td>Joseph et al. (2009)</td>
</tr>
<tr>
<td></td>
<td>• Precipitation Estimation from Remote Sensing Information using Artificial Neural Network (PERSIANN)</td>
<td>Sorooshian et al. (2000)</td>
</tr>
<tr>
<td></td>
<td>• Tropical Rainfall Measuring Mission Product 3B42 (TRMM3B42)</td>
<td>Liu et al. (2012)</td>
</tr>
<tr>
<td>reanalysis products</td>
<td>• Global Precipitation Climatology Project (GPCP)</td>
<td>Huffman et al. (2009)</td>
</tr>
<tr>
<td></td>
<td>• CPC Merged Analysis of Precipitation (CMAP)</td>
<td>Xie and Arkin (1997)</td>
</tr>
<tr>
<td></td>
<td>• Modern Era Retrospective Analysis for Research and Applications (MERRA)</td>
<td>Rienecker et al. (2011)</td>
</tr>
</tbody>
</table>

In spite of those products being easily accessible they all have shortcomings. Satellite products tend to have large sensor errors due to their dependence on rainfall regime, short time record and lack of local data (Ebert et al., 2007). On the other hand, reanalysis products tend to be of coarse resolution of 0.5° x 0.5° or more (Blacut et al., 2015). The International Precipitation Working Group (IPWG) showed TRMM3B42 to be relatively accurate high-resolution satellite-based precipitation estimates for operational use (Ebert et al., 2007) - although, it too suffers from the drawbacks of relatively short record and coarse spatial resolution with a 0.25° x 0.25° grid size covering the period 1998 to 2013. Conversely, APHRODITE provides daily time step data at 0.25° x 0.25° grid - for the period 1950 - 2007 (Yatagai et al., 2012). Over Java Island, APHRODITE was developed from station data of Bureau of Meteorology, Climatology and Geophysics (BMKG) (∼20 stations) (Yatagai et al., 2012).

This data set was developed at a 0.125° x 0.125° resolution, spanning 1985 - 2014. This temporal period was selected to balance available data with a temporal period long enough to enable hydroclimate research. The grid size was chosen to maximize station coverage while representing
spatial variability of basin-scale natural processes in the predominantly small basins of Java, and was gridded from a much larger number of precipitation gauges (765 stations) than APHRODITE (~20 stations). This data set is expected to be useful for basin and island scale studies - e.g., modeling of catchment hydrology, land-atmosphere interaction, and ecological processes. Yanto et al. (2016b) - *in revision* - applied an LSM over Java finding that higher spatial resolution of meteorological data yielded better model performance and hydrologic process representation than when using coarser meteorology. Below we describe the station data, our gridding methodology and results.

### 3.2 Methods

#### 3.2.1 Station data

Figure 3.1 shows the spatial and temporal availability of precipitation data over Java for each decade starting from 1985. The precipitation stations are unevenly distributed over the island, however the mean station density over the island is one station per approximately 135 km$^2$ with more than 3650 days of non-missing data over the entire record period (Figure 3.1a). East Java has the fewest non-missing data followed by Central and West Java (Figure 3.1b,c,d). The period around 2005 has the fewest missing data across the island with modest increases afterwards and marked increases before 1990.
Figure 3.1: Spatial and temporal distribution of precipitation data - (a) percentage of missing data of each station in the entire record period (the number of days with missing data divided by the total number of days), (b), (c) and (d) percentage of missing data for decade 1,2 and 3 respectively and (e) percentage of stations with missing daily data (the number of stations with missing daily data divided by the number of available stations for particular day over Java).

3.2.2 Gap-infilling and gridding procedure

Several extant methods are available for interpolating rainfall in time (i.e. gap-infilling) and space (Burgess et al., 2015; Coulibaly and Evora, 2007; Feki et al., 2012; Ly et al., 2011). The Inverse Distance Weighting (IDW) method and geostatistical Kriging method with all their respective variants are the two most widely employed and compared methods in literature (Feki et al., 2012; Ly et al., 2011; Mair and Fares, 2011). IDW is a relatively simple method and requires
relatively little input data. On the other hand, Kriging can account for spatial correlation of neighboring observations as well as several techniques to incorporate covariates to improve station data (Ly et al., 2011; Mair and Fares, 2011; Yang et al., 2015). Mair and Fares (2011) found that all Kriging variants perform better than IDW at monthly time scale. However, at daily time step, both IDW and Kriging produce comparably small errors (Ly et al., 2011). In addition, for small watersheds, IDW was shown to perform better than Kriging (Tao et al., 2009). In tropical region Yang et al. (2015) found that IDW interpolation performed slightly better than Kriging. Kriging requires estimation of the variogram to describe the degree of spatial dependence of a spatial random field or stochastic process, which is notoriously difficult to fit with short data records and it can be problematic fitting a variogram for each day. At monthly time scales, the variogram is much more stable and thus similar in performance to IDW. Given that the majority of basins in Java are small, IDW was employed here for both gap-infilling and the gridding process. In the IDW method, the interpolated value is estimated as a weighted linear combination of nearest observations, in which the weights are proportional to the inverse of the distance between neighboring observations and interpolation location (Shepard, 1968), as presented below.

\[
\hat{q} = \frac{\sum_{i=1}^{n} \frac{1}{r_i^\alpha} q_i}{\sum_{i=1}^{n} \frac{1}{r_i^\alpha}} \tag{3.1}
\]

where \( \hat{q} \) is the interpolated (gridded) value, \( q_i \) is the observed value in station \( i \), \( r_i \) is the Euclidean distance between interpolated station and station \( i \), \( \alpha \) is the power of distance and \( n \) is the total number of stations interpolated per grid.

The radius of influence (\( r \)) and the power of distance (\( \alpha \)) are parameters which can be adjusted to obtain an optimal interpolated value. Chen and Liu (2014) found that interpolated root mean square errors are minimized when \( r \) ranges from 10 - 30 km and \( \alpha \) between 0 - 5. To optimize, we considered \( r \) values of 10, 25 and 50 km to interpolate surrounding stations and \( \alpha \) of 1, 2 and 3 to weigh local and regional influences. Since missing data were numerous at times (especially before 1990) and the distances between stations in some regions are relatively large (Figure 3.1e),
we added a constraint requiring a minimum number of stations to be interpolated. Four stations with non-missing data were required to lie within the radius, otherwise the nearest next-closest stations with non-missing data were selected. Figure 3.2a shows that the number of stations with non-missing data within $r$ of 10 km was less than four, which fails to meet the threshold, thus only $r$ of 25 and 50 km were applied, with a smaller radius (25 km) preferable to avoid smoothing of data from differing precipitation events.

To obtain an $\alpha$ value which optimally captures spatial and temporal precipitation variability, we applied IDW to daily station precipitation as well as mean daily precipitation at 9 selected test locations spread relatively uniformly about Java (Figure 3.2b) using $r$ of 25 km and $\alpha$ of 1, 2 and 3. The variability of station’s elevation was incorporated by defining $r$ as three-dimensional Euclidean distance - i.e. square root of sum of squared x, y and z. For each location, the spatial average of interpolated and observed daily precipitation were calculated and the $R^2$ were computed. Figure 3.2c shows that in nearly all cases $\alpha$ of 3 achieves the highest performance. Using $\alpha$ of 3, we assessed the sensitivity of $r$ to the interpolated values. As presented in Figure 3.2d,e, IDW is less sensitive to radius of influence, as apparent by similar spatial pattern of mean daily gridded precipitation between $r$ of 25 versus 50 km. This is driven by the nature of IDW whereby the higher the value of $\alpha$, the less weight apportioned to more distant stations. Accordingly, we ultimately elected to use $\alpha = 3$ and $r = 25$ km - i.e. using fewer stations but with the ability to capture information from more distant stations - in the IDW applications for both gap-infilling and gridding.
Figure 3.2: (a) the number of stations with non-missing data within radius of influence \( (r) \) of 10, 25 and 50 km for each grid where each grid was assigned an index/number, the horizontal black line is a threshold of 4 stations for IDW interpolation; (b) selected locations to examine the performance of interpolation, the grid size is equivalent to 0.25° x 0.25° to capture at least one observed station; (c) model performance between the box-average gridded and observed precipitation; (d) and (e) are mean daily precipitation which gridded using \( r \) of 25 and 50 km respectively.

### 3.2.3 Code availability

The data was processed using a standard version of R software, R.3.2.2. The code will be publicly available alongside the data set [http://datadryad.org/review?doi=doi:10.5061/dryad.36rs0](http://datadryad.org/review?doi=doi:10.5061/dryad.36rs0). Some R packages need to be installed before implementing the code and full instructions will be provided together with the data.
3.3 Data Records

We collected meteorological data fields from various sources as presented in Table 3.2. The Center of Water Resources Development and Management, Ministry of Public Works, Indonesia provided daily precipitation data at most stations over Java. We obtained additional daily precipitation, minimum temperature and maximum temperature from the Bureau of Meteorology, Climatology and Geophysics of Indonesia. The availability of precipitation data was used for determining the data set temporal domain which was defined as the period of 1985 - 2014.

The final data set contains gridded station data for precipitation, maximum and minimum temperature at a daily time step running from 1 January 1985 - 31 December 2014 at 0.125° x 0.125° resolution. The data set, stored in network common data form (NetCDF), is archived at http://datadryad.org/review?doi=doi:10.5061/dryad.36rs0, with access details provided in the Data Citation.

Table 3.2: Summary of daily data sources used to develop present meteorological data set.

<table>
<thead>
<tr>
<th>Data type</th>
<th>Source</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precipitation</td>
<td>Center for Water Resources Research and Development, Ministry of Public Works, Indonesia</td>
<td>972 total stations and 691 stations satisfy minimum length of 3650 days</td>
</tr>
<tr>
<td></td>
<td>BMKG Indonesia</td>
<td>19 total stations and 16 stations satisfy minimum length of 3650 days</td>
</tr>
<tr>
<td>Minimum temperature</td>
<td>BMKG Indonesia</td>
<td>19 total stations and 16 stations satisfy minimum length of 3650 days</td>
</tr>
<tr>
<td>Maximum temperature</td>
<td>BMKG Indonesia</td>
<td>19 total stations and 16 stations satisfy minimum length of 3650 days</td>
</tr>
</tbody>
</table>
3.4 Technical Validation

We validated the data set on grid, basin and island scales at daily and monthly time steps using station observations as validation. We first cross-validate the gap-infilling procedure by removing daily observed values at each station and interpolate them using the IDW method - i.e. cross-validation. To validate the gridding method, we calculated "gridded-precipitation" error - i.e. an error between each station and its nearest grid cell value. Percent bias (PBIAS) - the difference between observation and interpolated estimates divided by the observation and reported as a percentage - was computed for daily precipitation to validate the goodness of fit of the gap-filling and gridding for all stations over the entire time period. Negative PBIAS indicates that the gap-filling and gridding overestimates the observation and vice versa, with an optimum value of zero. Figure 3.3 shows the scatterplot of PBIAS and its respective precipitation magnitude averaged over space (spatial mean daily precipitation) and time (temporal mean daily precipitation) for cross-validation and gridded precipitation. As shown, gridded precipitation has smaller errors than cross-validation precipitation, for both spatial and temporal performance. In addition, a shift from positive to negative PBIAS is apparent in spatial mean daily precipitation when precipitation moves from low to high magnitude, indicating that the gridded precipitation tends to underestimate small events and vice versa resulting in an overestimate of precipitation variability (Figure 3.3a). Conversely, an opposite pattern in the temporal mean daily precipitation suggesting that gridded precipitation overestimates observations in dry days while converse tendency occurs in wet days producing muted variability (Figure 3.3b).
Figure 3.3: Scatterplot of PBIAS and precipitation magnitude for (a) temporal mean daily precipitation, computed as the average of daily precipitation in each station and (b) spatial mean daily precipitation, defined as the mean of station precipitation in each day.

Figure 3.4 illustrates basin-scale variability between gridded precipitation at daily and monthly time steps. We attempt to validate the gridded precipitation with observed streamflow to underscore the importance of the water balance for understanding hydrologic processes, as discharge, $Q$, represents an integration of precipitation for an entire basin (Maurer et al., 2001; Milly and Wetherald, 2002). We compare statistical features of daily gridded precipitation and observed streamflow in Citarum (CT), Progo (PG), Bengawan (BG), Grindulu (GD) and Bedadung (BD) watersheds. These watersheds have discharge that passed quality control procedures outlined by Yanto et al. (2016b) - in revision. We use Q-Q plot to match the quantiles between the observation and interpolation and to identify the present of outliers. As displayed by Q-Q plots, both gridded precipitation and observed streamflow have similar distributions. Additionally, the monthly gridded precipitation hyetograph follows the streamflow hydrograph in all the watersheds, with a few exceptions such as the earlier years of PG, while the year 2001 at GD likely has erroneous
discharge, as it greatly exceeds precipitation. Overall, these features characterize the key hydrologic processes in Java where overland flow is readily generated after storm events, as the result of soil moisture excess (Muñoz-Villers and McDonnell, 2012, 2013), with minimal lag time between monthly precipitation and streamflow.

Figure 3.4: (a) Q-Q plots for daily gridded precipitation and observed streamflow for five study watersheds, and (b) monthly hyetographs and hydrographs for the same watersheds, $A$ is the basin area, km$^2$. 
An additional island-scale validation was conducted, comparing probability density functions (PDFs), scatterplots, and time series of station and gridded data for P, Tmin and Tmax. Both spatial Figure 3.5a,e,i and temporal Figure 3.5b,f,j PDFs were estimated, with the former computed using the mean island-wide field for each day, and the latter computed as the mean inter-grid-cell value averaged over time. In addition, the daily and seasonal variability of the data set was examined in scatterplots (Figure 3.5c,g,k) and comparing their monthly time series (Figure 3.5d,h,l).

The gridded and station-based precipitation are comparably distributed over time and space with similar spatial PDF shapes for minimum and maximum temperature. However, the temporal PDFs for temperature show fairly large differences, which we attribute to large distances between temperature stations. Overall, scatterplots and monthly time series of the spatial average of meteorological fields show that the gridded product closely matches the station data for all variables, with regression lines close to the 1:1 line (black line) (Figure 3.5c,g,k).

### 3.5 Usage Notes

We caution data users who aim to use this data set for trend analyses that many stations do not extend for the entire period 1985 - 2014. Another point of caution is that the pre-1990 period has greater than 50% of stations reporting missing data. Similarly, users of data in West Java particularly in the period of 1985 - 1994 and 2005 - 2014 should be aware that most stations have more than 70% missing data.

### Data Citation


### Acknowledgements

This study was funded by The Directorate General of Higher Education, The Ministry of National Education, Indonesian (Dirjen Dikti), via a Dikti Scholarship and extended by Graduate Student Research Award from Cooperative Institute for Research in Environmental Studies
Figure 3.5: The probability density function (PDF) of mean daily P, Tmax and Tmin (a,e,i) respectively computed using the island-average value for each day in the record; the PDFs of mean daily P, Tmax, Tmin (b,f,j) computed across the mean daily value for each grid cell; the third column (c,g,k) shows scatterplots between daily gridded and observed P, Tmax and Tmin respectively overlaid by the 1:1 line (black) and regression line (red); and the fourth column (d,h,l) are monthly time series of P, Tmax and Tmin respectively.
(CIRES) University of Colorado Boulder awarded to the first author.

**Author Contributions**

Yanto did the data collection, data quality control, gap-infilling and gridding of station data, and led the manuscript writing. Ben Livneh provided guidance in the gap-infilling and gridding procedures, as well as in making methodological decisions and data validation. Balaji Rajagopalan provided guidance in the gap-infilling and gridding process methodological decisions and analysis of the data validation.
Chapter 4

Multi-objective Optimization Based Calibration of Hydrologic Model and
Ensemble Hydrologic Forecast for Java Island, Indonesia

This research is prepared for submission in Journal of Hydrology.

Abstract

This study explores the benefits of multi-objective optimization based calibration of Variable Infiltration Capacity (VIC) model for five watersheds in Java, the most populous island in Indonesia. Six objective functions: Nash Sutcliffe Efficiency (NSE), percent bias (PBIAS), logarithmic function of root mean square error (Log-RMSE), predictive efficiency (Pe), percent errors in peak (PEP) and slope of flow duration curve error (SFDCE) were selected as evaluation metrics. These metrics were optimized by tuning the most sensitive VIC model parameters: infiltration shape parameter \( b \), fraction of maximum baseflow where nonlinear baseflow begin \( Ds \), thickness of soil layer 2 \( thick_2 \) and thickness of soil layer 3 \( thick_3 \). We employed Borg Multiobjective Evolutionary Algorithm (Borg MOEA), an automatic simulation-optimization algorithm, to search for non-dominated solutions. We identified the redundancy between NSE and Log-RMSE, Pe, and PEP through visual inspection of their sensitivity to parameters \( b \) and \( Ds \) of VIC model and to baseflow index (BFI). Accordingly, we proposed NSE, PBIAS and SFDCE as critical objective functions to represent hydrologic processes in tropical region of Java, Indonesia. Using these three objective functions, we culled the objective functions based on at least - NSE > 0.75, PBIAS < 15\% and SFDCE < 15\% - and showed that the number of optimized objective functions as well as
model parameters in the ensemble are reduced but the value ranges are maintained. We used the culled model parameters, to run the VIC model using an ensemble of conditioned seasonal climate forecast to generate an ensemble streamflow prediction in the period 2001 - 2010, the time window when the seasonal climate forecasts and observed streamflow records overlaps. We measured the skill of this seasonal forecast by computing the rank probability skill score (RPSS) of seasonal total flows and extremes at three different thresholds, for the dry and wet seasons. We showed that the RPSS of seasonal flows and the extremes are very good for both the seasons. This study, for the first time, demonstrates the utility of the multiobjective based calibration of hydrologic model in tropical regions and its applications in generating skillful seasonal ensemble hydrologic forecasts which are important for short and long term water resources planning and management.

4.1 Background

Floods, droughts and landslides visit Java, Island of Indonesia frequently, especially in recent decades, threatening more than 60% of country’s population (BNPB, 2015). To mitigate this, tools for understanding, modeling and forecasting the hydrology is crucial. Several studies have demonstrated strong teleconnection of rainfall variability and predictability over Indonesia with El Niño Southern Oscillation (ENSO) at seasonal and multi-year time scales (Philander, 1983; Kirono et al., 1999; Haylock and McBride, 2001; Hendon, 2003; Gutman et al., 2000) particularly in the post-1980 period (Yanto et al., 2016c). To translate this predictability to hydrologic processes, a robust hydrologic modeling framework is important and there have been no serious research on this for Indonesia.

Recently, Yanto et al. (2016b), in one of the first attempts, developed a Land Surface Model (LSM) using Variable Infiltration Capacity (VIC) for watersheds in Java Island (Yanto et al., 2016b). The VIC model has been widely applied for hydrologic studies at a coarse resolution of 0.5° x 0.5° on a global scale (Nijssen et al., 2001, 2014; Sheffield and Wood, 2007; Shukla et al., 2013) and at finer resolution of 0.05° x 0.05° in Germany (te Linde et al., 2008) and Australia (Zhao et al., 2011) - hence, it is selected for Java.
The LSMs have two main sources of uncertainties - (i) simplification of natural real system by model structure (Clark et al., 2008; Sulis et al., 2011) and (ii) the unobservable model parameters (Gupta et al., 1998; Houser et al., 2001). The values of these parameters need to be estimated on a trial and error basis such that the simulation accurately resembles the observation - i.e. model calibration (Gupta et al., 1998; Houser et al., 2001).

In the calibration process, the goodness of model fit is evaluated based on the value of a selected objective function (Gupta et al., 1998; Vázquez et al., 2002; Wi et al., 2015; Xia et al., 2005). Typically, these are metrics that quantify overall errors between observed and modeled flows - such as root mean squared error (RMSE) (Chai and Draxler, 2014; Gupta et al., 2009), Nash Sutcliffe Efficiency (NSE) (Gupta et al., 2009) etc. Several objective functions have been proposed to validate model performance for various applications, such as percent bias (PBIAS) to measure volumetric errors (Moriasi et al., 2007), transformed RMSE (TRMSE) to give more weights on low flow (van Werkhoven et al., 2009), slope of flow duration curve error (SFDCE) to capture the variability or flashiness of flow magnitude (van Werkhoven et al., 2009), percent error in peak (PEP) to match simulation and observation high flow (ASCE, 1993) and predictive efficiency to capture the probability distribution function of the observed flow (Santhi et al., 2001).

Optimizing on single objective function to calibrate the models have been a staple in this field (Gupta et al., 2012). However, the performance of these calibrated models depends on the selected objective function and they might miss several aspects of the flow hydrograph, which argues for considering multiple objectives. Wagener et al. (2009), using root mean square error, low-flow and water balance objective functions, demonstrated that timing-related objective functions are largely sensitive in grid cells closer to streamflow outlet, while volume-based objective function is sensitive across the watershed, showing that multiobjective functions extract more information than single objective function. Kollat et al. (2012) showed that hydroclimatic regions in Leaf River catchment, Mississippi, can be identified when the trade-offs between objective functions are at meaningful precision. Some objective functions are redundant as they are highly correlated with others. It has been suggested that 2 - 4 objective functions are sufficient to characterize most streamflow regimes.
(Olden and Poff, 2003; van Werkhoven et al., 2009). Recently, Houle (2015) applied multiobjective optimization in calibrating a snow model within VIC in western United States using four objective functions; RMSE, NSE, the differences in maximum snow water equivalent accumulation and the number of days when snow is on the ground. These results point to the utility of multiobjective optimization in model calibration.

Optimizing over multiple objectives produces an ensemble of parameter values which, collectively, can capture various aspects of the flow hydrology and has the ability to capture parameter uncertainty. Seasonal hydrologic forecasting involves translating seasonal climate forecasts to hydrologic processes. To this end, ensemble hydrologic forecasts are being widely adopted to incorporate climate uncertainties (Grantz and Rajagopalan, 2005; Regonda et al., 2006; Block and Rajagopalan, 2009). In the U.S. ensemble streamflow prediction (ESP) is widely used by River forecasting centers to issue short term (1-2 weeks) and long term (seasonal and beyond) forecasts (Grantz and Rajagopalan, 2005; Regonda et al., 2006; Bracken et al., 2010). Multiobjective optimization approach provides the prospects of incorporating parameter uncertainty via the ensemble that it generates, which motivates this research.

Here we optimized VIC model for Java Island using six objective functions to generate an ensemble of daily baseflow and surface flow. We also demonstrate the utility in generating ensemble seasonal hydrologic forecasts based on seasonal categorical forecasts of precipitation. The paper is organized as follows. The study region and data sets are briefly described. Followed by a description of the methodology which includes model set up, objective functions selected and the rationale, multi-objective optimization approach and the proposed approach for seasonal forecasting. The results are then described, followed by summary and discussion.

4.2 Study Area and Data

We considered 20 watersheds covering Java Island shown in Figure 4.1. Data quality is a major issue in Indonesia both lack of data and poor quality. After employing quality control steps (Yanto et al., 2016b) to ensure rainfall and streamflow consistency and availability of at least
7-years of quality data, six watersheds were selected. For this study we selected five watersheds - Citarum (CT), Progo (PG), Bengawan (BG), Grindulu (GD) and Bedadung (BD) - colored in red in Figure 4.1. We excluded watershed Cilangla (CL), which was included in our previous study (Yanto et al., 2016b), due to lack of overlap between the new meteorological forcing data and streamflow record. Details of the study watersheds are described in Table 4.1. High resolution, $0.125^\circ \times 0.125^\circ$ ($\sim 14$ km), daily meteorology - precipitation, maximum and minimum temperatures - dataset developed in Yanto et al. (2016b) is used in this study.

![Figure 4.1](image)

**Figure 4.1:** Study area comprising the five watersheds over Java Island, colored in red.

<table>
<thead>
<tr>
<th>Watershed</th>
<th>Outlet Lon &amp; Lat (°)</th>
<th>Outlet Mean Elevation (m)</th>
<th>Streamflow record period</th>
<th>Area (km$^2$)</th>
<th>Mean streamflow (mm/yr)</th>
<th>Mean rainfall (mm/yr)</th>
<th>Land cover</th>
<th>Runoff ratio (Q/P)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Citarum (CT)</td>
<td>107.53/-6.95</td>
<td>668</td>
<td>1351</td>
<td>1990-1997</td>
<td>1675</td>
<td>2739</td>
<td>A(59), F(27), S(13), W(1)</td>
<td>0.50</td>
</tr>
<tr>
<td>Progo (PG)</td>
<td>110.26/-7.67</td>
<td>180</td>
<td>1115</td>
<td>1996-2002</td>
<td>1676</td>
<td>2412</td>
<td>A(15), F(79), S(5), W(1)</td>
<td>0.48</td>
</tr>
<tr>
<td>Bengawan (BG)</td>
<td>112.17/-7.10</td>
<td>11</td>
<td>885</td>
<td>1990-2002</td>
<td>16286</td>
<td>1993</td>
<td>A(56), F(34), S(9), W(1)</td>
<td>0.44</td>
</tr>
<tr>
<td>Grindulu (GD)</td>
<td>111.14/-8.14</td>
<td>39</td>
<td>1094</td>
<td>1990-2002</td>
<td>556</td>
<td>2238</td>
<td>A(58), F(42), S(0), W(0)</td>
<td>0.56</td>
</tr>
<tr>
<td>Bedadung (BD)</td>
<td>113.58/-8.23</td>
<td>44</td>
<td>1322</td>
<td>1991-2001</td>
<td>696</td>
<td>2504</td>
<td>A(21), F(74), S(5), W(0)</td>
<td>0.50</td>
</tr>
</tbody>
</table>

Table 4.1: Characteristics of the six watersheds over Java selected for this study. Land cover type A, F, S and W represent agriculture, forest, settlement and water body respectively with bold font indicating the dominant land cover.
4.3 Methodology

4.3.1 Model setup

The VIC model (Liang et al., 1994) which was successfully setup and applied to model hydrology processes in Java (Yanto et al., 2016b), is used in this study. VIC is a macro-scale, physically-based and fully-distributed model that solves the energy and water balance separately (Liang et al., 1994, 1996). A standard application of VIC with 2 canopy layers, 3 soil layers and 2 root zones, is adopted. Both canopy layers characterize land surface-atmospheric interaction - i.e. evapotranspiration, interception and through-fall. The dynamic response of soil to variable infiltration rates of incoming precipitation is represented by top soil layer. This response is then transferred to the bottom soil layer via the middle soil layer, while the bottom most layer is used to model groundwater and baseflow process. With 2 root zones, upward moisture transport from roots is performed by two top soil layers. We refer the readership to Liang et al. (1994) for details about VIC model structure and formulations.

Following the preliminary result on the largest basin from Yanto et al. (2016b) where finer model resolution produced better simulation we built the model at 0.125° x 0.125° resolution for all the study watersheds. The soil parameters were downscaled from Nijssen et al. (2014) using Inverse Distance Weighting (IDW), while the vegetation covers were derived from the University of Maryland’s 1km Global Product (Hansen et al., 2000). The VIC model was forced using daily meteorology from Yanto et al. (2016a) developed exclusively from ground-based measurements Table 4.2.
Table 4.2: Summary of daily data sources used to develop present meteorological data set.

<table>
<thead>
<tr>
<th>Data type</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precipitation</td>
<td>Center for Water Resources Research and Development, Ministry of Public Works, Indonesia</td>
</tr>
<tr>
<td></td>
<td>BMKG Indonesia</td>
</tr>
<tr>
<td>Minimum</td>
<td>BMKG Indonesia</td>
</tr>
<tr>
<td>Maximum</td>
<td>BMKG Indonesia</td>
</tr>
</tbody>
</table>

4.4 Objective functions

Several objective functions have been used to evaluate hydrologic model performance (Moriasi et al., 2007). To effectively calibrate hydrologic models objective functions should capture the functional behavior of catchments (Black, 1997; Sawicz et al., 2011). In this study, we used six objective functions, together capture all aspects of the flow hydrograph: NSE, PBIAS, logarithmic root mean square error (Log-RMSE), Pe, SFDCE and PEP. These are described below.

NSE measures relative magnitude of residual variance to the observed data variance and consequently favors the high flows. It is used as a measure for the overall flow magnitude (Moriasi et al., 2007; Nash and Sutcliffe, 1970; Sevat and Dezetter, 1991). According to Nash and Sutcliffe (1970), NSE is defined as:

\[
NSE = 1 - \frac{\sum_{t=1}^{T} (Q_o^t - Q_s^t)^2}{\sum_{t=1}^{T} (Q_o^t - \bar{Q}_o)^2}
\]  

(4.1)

where in Equation 1 to 3, \(Q_o^t\) and \(Q_s^t\) are observed and simulated streamflow at time \(t\) respectively, \(\bar{Q}_o\) is the mean of observed streamflow - NSE ranges from \(-\infty\) to 1 and NSE < 0 indicates that the mean flow is a better predictor than the model simulations. Typically the model
performance is deemed acceptable when NSE > 0.50 (Moriasi et al., 2007).

Relative magnitude of long term simulated and observed values is evaluated using PBIAS, expressed in percentage (Equation 2). Negative values indicate the tendency of the simulation to be larger than the observation and vice versa and optimal value of 0 indicates accurate model simulation (Gupta et al., 1999). This metric has been expressed in different ways in literature: percent streamflow volume error (PVE) (Singh et al., 2005), prediction error (PE) (Fernandez et al., 2005), and percent deviation of streamflow volume (Dv) (ASCE, 1993).

\[
PBIAS = \frac{\sum_{t=1}^{T} (Q_t^o - Q_t^s) \times 100}{\sum_{t=1}^{T} Q_t^o}
\]  

(4.2)

To capture the low flow features, we used logarithm of the root mean square errors. This gives more weight to low flows compared to high flows and is computed as:

\[
Log - RMSE = Log \left( \frac{1}{n} \sum_{t=1}^{T} (Q_t^s - Q_t^o)^2 \right)
\]

(4.3)

Predictive efficiency, Pe, determines the likeliness between the probability distributions of observed and modeled flows. This is done by computing coefficient of determination (R²) of the ranked (descending) values of observed and simulated flows (Santhi et al., 2001). In Equation 4, and are descending rank of observed and simulated values respectively and over bar sign represents the mean value.

\[
P_e = \left( \frac{\sum_{t=1}^{T} (Q_t^o - \bar{Q}_o)(Q_t^s - \bar{Q}_s)}{\sqrt{\sum_{t=1}^{T} (Q_t^o - \bar{Q}_o)^2 \cdot \sum_{t=1}^{T} (Q_t^s - \bar{Q}_s)^2}} \right)^2
\]

(4.4)

The flashiness of a watershed's response is addressed by minimizing the error of slope of the flow duration curve computed as:
\[ SFDCE = \left( \frac{(Q_{s,67\%} - Q_{s,33\%})}{(Q_{o,67\%} - Q_{o,33\%})} - 1 \right) \times 100\% \] (4.5)

where \( Q_{s,67\%} \) and \( Q_{s,33\%} \) is the 67th and 33rd percentile of the simulated flows and subscribe \( o \) represents observed flows.

Peak flows are important, especially when simulating floods. Percent error in peak, PEP, is used to capture this. As shown in Equation 6, PEP is calculated by dividing the peak flow rate difference between simulation and observation by the measured peak flow rate and presented as percentage (ASCE, 1993; Green and Stephenson, 1986; Moriasi et al., 2007). In this study we defined peak flow as daily flow exceeds 95th percentile.

\[ PEP = \frac{(Q_{po} - Q_{ps}) \times 100}{Q_{po}} \] (4.6)

4.4.1 Parameter calibration - Ensemble

We calibrated only the most sensitive model parameters suggested by (Yanto et al., 2016b) for the study watersheds - they were, \( b, Ds, thick_2 \) and \( thick_3 \). In addition, for routing the flow on the surface, flow velocity and flow diffusivity suggested by Lohmann et al. (1998) and Nijssen et al. (1997), have to be estimated, which are included for calibration. Calibration is done on the first \(~5\) year period of the record (1990 - 1995) and validated on the second half through evaluating the objective functions at daily time scale. The description and range of values of the six parameters are presented in Table 4.3.
Table 4.3: Description of calibration parameters, their ranges and the values obtained from optimization for the six watersheds.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>b</th>
<th>Ds</th>
<th>thick_2</th>
<th>thick_3</th>
<th>Velocity</th>
<th>Diffusivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description</td>
<td>infiltration curve shape parameter controlling surface runoff</td>
<td>fraction of maximum baseflow where nonlinear baseflow begins</td>
<td>thickness of soil layer 2</td>
<td>thickness of soil layer 3</td>
<td>Flow velocity</td>
<td>Flow diffusivity</td>
</tr>
<tr>
<td>Unit</td>
<td>NA</td>
<td>NA</td>
<td>m</td>
<td>m</td>
<td>m/s</td>
<td>m²/s</td>
</tr>
<tr>
<td>Range</td>
<td>&gt; 0 - 1</td>
<td>&gt; 0 - 1</td>
<td>0.1 - 3</td>
<td>0.1 - 3</td>
<td>0.5 - 3</td>
<td>200 - 4000</td>
</tr>
</tbody>
</table>

In single objective function calibration, a single set of parameters that optimize the objective function is obtained. Typically NSE is used as the objective function (Gupta et al., 2009). In multiobjective calibration, conflict between two or more objective functions (Guo et al., 2014; Reed et al., 2013) is inherent, as result no single parameter set is produced. Instead, an ensemble of solutions (i.e. parameters) is generated, each associated with a set of objective function values. This provides several options to the user - the parameter sets can be evaluated for their tradeoffs between the objective functions; based on a desired objective function a single parameter set can be selected or, a smaller ensemble of parameters can be culled based on desired suite of objective function thresholds. We employed the Borg Multi Objective Evolutionary Algorithm (Borg MOEA), an automatic simulation-optimization algorithm, through direct connection with VIC model (Hadka and Reed, 2013). The Borg MOEA has been applied to calibration problems (Reed et al., 2013) and to calibrate the VIC model over watersheds in Java using a single objective function, NSE (Yanto et al., 2016b). It has also been implemented in various multiobjective optimization problems (e.g., Kasprzyk et al., 2012, 2013; Reed et al., 2013). Recently it has been used to calibrate snow model
within VIC in a multi objective optimization mode (Houle, 2015).

In MOEAs, epsilon dominance ($\epsilon$-dominance) is a popular algorithm used to obtain the Pareto front (Deb et al., 2002; Kollat and Reed, 2006). The front consists of ensemble of model parameters and their associated objective function values, obtained by sampling the parameter space. To overcome shortcomings in MOEAs, such as dominance resistance (Ikeda et al., 2001; Purshouse and Fleming, 2007), active diversity maintenance (Purshouse and Fleming, 2007), Borg MOEA introduced the features of epsilon box ($\epsilon$-box) dominance to maintain convergence and diversity throughout the search and epsilon progress ($\epsilon$-progress) to efficiently measure the search progression and stagnation. These values are user-specified (Hadka and Reed, 2013). Here, we used the default settings of Borg MOEA - with 20,000 function evaluations, $\epsilon$ value of 0.5 and 3 random seeds to overcome the issue of equifinality in the optimization problem.

4.4.2 Seasonal hydrologic forecasting

To demonstrate the utility of the ensemble VIC model parameters to generate seasonal ensemble hydrologic forecasting conditioned on the seasonal climate forecast. We produced hydrologic forecasts for dry season (Jul-Sep) and wet season (Dec-Feb) for the period 2001 to 2010, the period for which seasonal climate forecasts are available. The following steps are employed for each basin:

1. The historical seasonal precipitation is normalized (i.e. subtract the mean and divide by the standard deviation) and categorized as above normal, normal and below normal years based on a threshold of 0.75 and -0.75 on the normalized data. Thus each year is in one of the category. This is done separately for the dry and wet seasons.

2. Suppose the seasonal precipitation forecast from IRI, Columbia University, for an upcoming dry season is 35:25:40 - indicating 35% chance of above normal (A), 25% chance of normal (N) and 40% chance of below normal (B) categories. For each season, we used 1-month lead time seasonal climate forecast issued by IRI Columbia University (http://iri.columbia.edu/our-expertise/climate/forecasts/seasonal-climate-forecasts/) - e.g., for peak
dry season (July-August-September) we took the forecast issued in June.

(3) Select 35% of years from above normal set, 25% of years from normal and 40% from the below normal years, identified in step 2 above.

(4) The daily precipitation and temperature of the selected years are used to drive the VIC model, along with the ensemble parameter sets obtained from multiobjective optimization. Each soil parameter was combined with 20 meteorology inputs - thus with ∼300 parameter sets from the multiobjective optimization, there are ∼6,000 daily ensembles, for each basin.

(5) The ranked probability skill score (RPSS) was computed for seasonal total flow. The RPSS (Weigel et al., 2007; Wilks, 1995) computes the categorical skill relative to climatology, where a value of 1 indicates perfect forecast, value of 0 suggests the skill to be no better than climatology and negative values indicate performance poorer than climatology. This is computed as follows (Wilks, 1995):

\[
RPSS = 1 - \frac{RPS_{fct}}{RPS_{cli}}
\]

(4.7)

where and are rank probability score (RPS) for the forecast and the climatology forecast respectively. RPS is computed as (Weigel et al., 2007):

\[
RPS = \frac{1}{K-1} \sum_{k=1}^{K} (P_k - O_k)^2
\]

(4.8)

where K being the number of categories, \( P_k = \sum_{i=1}^{k} p_i \), is the kth component of the cumulative forecast with \( p_i \) being the probabilistic forecast in category \( i \), and \( O_k = \sum_{i=1}^{k} o_i \), is the kth component of the cumulative observation with \( o_i = 1 \) if the observation is in category \( i \), and 0 otherwise. Similarly, RPS of climatology is computed, replacing \( P_k \) with \( C_k \), the climatological forecasts. For three categories created at the tercile thresholds, the climatological forecast for each category is equal to 1/3.
The above approach is similar to the Ensemble Streamflow Prediction (ESP) (Dai, 1985) used by River Forecasting Centers in the US. The weighted resampling of historic meteorology is the main difference here, as opposed to the unweighted approach in traditional ESP. The above approach was used in the context of generating daily weather ensembles conditioned on seasonal climate forecast using stochastic weather generators (e.g., Verdin et al., 2015; Apipattanavis et al., 2007; Caraway et al., 2014). Caraway et al. (2014) used a conditional stochastic weather generator coupled with a physical watershed model to generate ensemble streamflow on the San Juan river, a tributary of Colorado River - and showed very good skill in the forecasts.

4.5 Result and Discussion

4.5.1 Model parameter - Objective function

It is important to identify the parameters that are sensitive to the objective functions. The multiobjective optimization, with its ensemble of parameter and objective function values, readily enables this. In Yanto et al. (2016b), using a single objective function with VIC at monthly time scale, NSE, four model parameters, described in Table 4.3, were found to be sensitive. Here, we investigated the sensitivity of these parameters to multiple objectives. Figure 4.2 shows the scatterplot of model parameters and the objective functions for all the basins to provide a general inference of model sensitivity in tropical region. From visual inspection of the scatterplots for linear and nonlinear features, it can be found that NSE is sensitive to parameters $b$ and $D_s$ but not to $thick_2$ and $thick_3$. In addition, parameters $b$ and $D_s$ are also sensitive to $Pe$, $Log$-RMSE and PEP. The sensitivity of NSE, $Pe$, $Log$-RMSE and PEP to the same model parameters, $b$ and $D_s$, indicates relationship between these objective functions and some level of redundancy - i.e. the higher $b$ correlates to higher NSE, $Pe$ and lower $Log$-RMSE and PEP meaning that increasing $b$ results in the optimizing NSE, $Pe$, $Log$-RMSE and PEP. It appears that PBIAS is consistent with PEP in their variations for the two parameters, except in watershed PG. This indicates that in most cases the model underestimates basin-scale water balance. Both soil thickness parameters,
thick_2 and thick_3 are not sensitive to most of objective functions which is not the case of monthly time step NSE (Yanto et al., 2016b).

![Figure 4.2: Scatterplot of objective functions and model parameters for all five basins.](image)

4.5.2 Critical objective functions

Partitioning of precipitation into streamflow components is important in characterizing the response of a watershed to its geology, soil distribution, relief and vegetation and also beneficial for skillful hydrologic forecasting (Eckhardt, 2008; Gonzales et al., 2009). The BFI, the ratio of baseflow to total streamflow, is used to quantify the contribution of baseflow and surface flow to total streamflow. We computed the BFI directly from the baseflow and surface flows generated by the model for each parameter set. For the estimate of BFI from the observations, we employed a digital filter technique discussed by Nathan and McMahon (1990). Specifically, we used EcoHydRology package in R.
To assess the overall relationship between the objective functions and streamflow generation mechanisms in Java, we correlated the ensemble of optimized objective functions and BFI for all basins, shown in Figure 4.3. As can be seen, NSE and BFI has a correlation $R^2$ of 0.64, thus, can be considered as the most influential objective function to portray precipitation partitioning in tropical region. Considering that good model requires NSE $> 0.65$, for which, Figure 4.3 suggests the BFI in Java ranges from 0.3 to 0.5. Using digital filter method, we found that the observed BFI is 0.48, 0.52, 0.44, 0.33 and, 0.48 for CT, PG, BG, GD and BD respectively, which is consistent with model simulations for high NSE in Figure 4.3. The objective functions, Log-RMSE, PEP and Pe also exhibit good relationship with BFI indicating some redundancy with NSE. PBIAS and SFDCE have low association with BFI, suggesting that they could complement NSE.
Figure 4.3: Scatterplot of objective functions and baseflow index (BFI) for all basins. The regression line and its respective coefficient of determination ($R^2$) was computed across basins to provide general insights of relationship between objective functions and hydrologic processes in tropical regions.

We culled the parameter ensemble using threshold values of NSE, PBIAS and SFDCE to obtain a good parameter subset. Moriasi et al. (2007) discussed that acceptable NSE, PBIAS and SFDCE values are 0.5, 25% and 25%, and higher. The model parameters that satisfy at least one of the screening criteria is selected. Figure 4.4 presents the parallel plot of all optimized objective functions without culling, in black, overlaid by objective functions fulfilling the above criteria in grey, for the five basins. The values at the top are the worst objective function values and those at the bottom are the optimum values (for NSE and Pe, the value is multiplied by -1 to give a sense of minimization). The blue lines indicate the value of the objective functions based on the optimal solution that maximizes just the NSE. The culled solutions mask the full solution set in all basins except in CT and BD. The proportion of screened solutions were 0.49, 1, 1, 1 and 0.65 for
watersheds CT, PG, BG, GD and BD respectively. For all the basins, there are solutions better on other objective functions than the single objective function solution based on NSE, indicating the importance of multiobjective optimization.

Figure 4.4: Parallel plot of objective functions of all solutions (dark gray) overlaid by the objective functions culled based on at least one of the criteria - NSE > 0.5, PBIAS < 25% and SFDCE < 25% (light gray). Objective function based on optimal parameters from a single objective function, NSE, optimization.

A good model performance requires higher NSE, lower PBIAS and SFDCE. Accordingly, we culled the solutions using the criteria of at least - NSE > 0.75, PBIAS < 15% and SFDCE < 15% and show them in Figure 4.5 (similar to Figure 4.4). The proportion of screened solutions were 0.37, 0.75, 1, 0.65, 0.49 for CT, PG, BG, GD and BD respectively.
<table>
<thead>
<tr>
<th></th>
<th>NSE</th>
<th>PBIAS</th>
<th>Pe</th>
<th>Log-RMSE</th>
<th>PEP</th>
<th>SFDCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preferred CT</td>
<td>0.319</td>
<td>43.6</td>
<td>-0.918</td>
<td>1.304</td>
<td>82.3</td>
<td>81.3530</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.102</td>
<td>22.79</td>
<td>-0.711</td>
<td>1.50</td>
<td>88.2</td>
<td>8.14e+01</td>
</tr>
<tr>
<td>Preferred PG</td>
<td>0.182</td>
<td>19.8</td>
<td>-0.900</td>
<td>0.972</td>
<td>62.0</td>
<td>55.8651</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.161</td>
<td>42.8</td>
<td>-0.707</td>
<td>1.43</td>
<td>88.1</td>
<td>83.31743</td>
</tr>
<tr>
<td>Preferred BG</td>
<td>0.058</td>
<td>8.6451</td>
<td>-0.859</td>
<td>1.0450</td>
<td>74.1</td>
<td>82.71198</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.081</td>
<td>0.9811</td>
<td>-0.457</td>
<td>18.2</td>
<td>0.9762</td>
<td></td>
</tr>
</tbody>
</table>

Figure 4.5: Same as Figure 4, but culled based on at least one of the criteria of NSE > 0.75, PBIAS < 15% and SFDCE < 15%.

4.5.3 Soil parameter distribution

Multiobjective calibration approach generates a set of model parameters, each associated with a non-dominated solution. In Borg MOEA, these model parameters are randomly selected from a population of uniformly distributed parameter values Table 4.3 - i.e. prior distribution (Hadka and Reed, 2013). It is of interest to investigate the distribution of the optimal model parameters - i.e. posterior distribution. Figure 4.6 shows the probability distribution function (PDF) of non-dominates model parameters for all basins. Note that the posterior distribution does not necessarily imply the hydrologic process inference - i.e. low dominated parameter b distribution does not mean baseflow dominated streamflow generation. Instead, it expresses the process domination in the parameter search of Borg MOEA algorithm. It can be seen that the posterior distribution differs from the prior distribution showing that each model parameter received different weights dependent on the non-dominated solution search process.
Figure 4.6: Probability distribution function (PDF) of optimal model parameters (posterior) and initial population of model parameters (prior) for all the watersheds.

4.5.4 Monthly hydrograph

The model was evaluated in the ability to capture the flow hydrograph in the calibration and validation periods. For ease of visual inspection, we showed the monthly hydrograph in Figure 4.7. Monthly time series of flow ensemble generated from VIC model using the ensemble of optimal
parameters from multiobjective optimization is shown as grey lines in Figure 4.7 and the historic flow as solid lines. The single simulated monthly flow associated with the parameter set with the best NSE, the typical objective function used to optimize model parameters (Gupta et al., 2009), is also shown as dotted line. The scatterplot of the NSE-simulated and the historic flows are shown along with the 1:1 line. The NSE-simulated flows can generally follow the hydrograph timing - i.e. rising limb, peak and flow recession - as expected from its formulation. However, given the volumetric error between simulations and observations, shown by the deviate of regression line from 1:1 line, it is difficult to estimate whether the observation should be lower or higher than the simulation for certain months. On the other hand, the generated ensemble provides a range to encapsulate the observations. The flow ensemble preserves the hydrograph shape and timing, captures low and high flow. The periods of Aug - Sep 1991 and 1994 in CT, Jul - Aug 1999 and 2001 in PG show the ability of the ensemble flows in capturing low flows in the monthly hydrograph that the NSE-simulated flows is unable. The same demonstration for high flows can be seen in the periods of Feb 1999 and Nov 2000 in PG, Feb 1991, Mar 1994 and Feb 1997 in BD. In addition, there are several periods of medium flows where the flow ensembles closely matches the observations compared to NSE-simulated flows. The frequency of observed flow captured by the flow ensemble over the entire time period for both dry and wet season, was computed. We found that the flow ensemble envelopes more than 75% and 50% of observations during dry and wet seasons respectively, indicating that the ensemble captures the observed flow variations quite well.
Figure 4.7: Monthly time series of flow ensembles (grey lines) along with observed flows (solid line) and that simulated from the parameter set with the best NSE (dotted line), for the calibration period (left column) and validation period (right column). The vertical bars are monthly precipitation, the scatterplots show the observed and simulated flows from the best NSE based parameter set with the 1:1 line (thick black line) and the best fit line through the scatter (blue line).

4.5.5 Seasonal forecast

Ensemble flow forecast, conditioned on probabilistic climate forecast, were generated for each peak wet (Dec-Feb) and dry (Jun-Aug) season for the period 2001 - 2010, for which the IRI seasonal climate forecasts are available, following the steps outlined in the previous section. Table 4.4 shows the years for which flow data is available during the 2001 - 2010 period in the five watersheds. The flow ensembles are based on ensemble of meteorology resampled from historical meteorology consistent with the seasonal climate forecast and, the parameter ensemble - together capturing the
meteorological and model parameter uncertainty. We computed the RPSS of the ensemble forecast of seasonal flows (Table 4.5) over the period when the streamflow data available in each basin (see Table 4.4), separately for wet and dry years, using the terciles as the thresholds for the three categories.

Table 4.4: The years when the observed streamflow data is available during the forecast period of 2001-2010 in the study watersheds.

<table>
<thead>
<tr>
<th>Watershed</th>
<th>Dry Season</th>
<th>Wet Season</th>
</tr>
</thead>
</table>
Table 4.5: The RPSS of ensemble forecast of seasonal flows, for dry and wet seasons, using three categories at the tercile thresholds. The last three columns show the RPSS of the extreme flows at three higher thresholds.

<table>
<thead>
<tr>
<th>Watershed</th>
<th>Dry seasonal total flow</th>
<th>Wet seasonal total flow</th>
<th>Wet seasonal total flow exceed 75th percentile</th>
<th>Wet seasonal total flow exceed 85th percentile</th>
<th>Wet seasonal total flow exceed 95th percentile</th>
</tr>
</thead>
<tbody>
<tr>
<td>CT</td>
<td>0.5992</td>
<td>0.5490</td>
<td>0.6201</td>
<td>0.9354</td>
<td>0.9852</td>
</tr>
<tr>
<td>PG</td>
<td>0.4113</td>
<td>-0.6756</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>BG</td>
<td>0.9762</td>
<td>0.8491</td>
<td>0.7468</td>
<td>0.7428</td>
<td>0.7394</td>
</tr>
<tr>
<td>GD</td>
<td>0.5494</td>
<td>0.4723</td>
<td>0.0983</td>
<td>0.5844</td>
<td>0.5928</td>
</tr>
<tr>
<td>BD</td>
<td>0.4581</td>
<td>0.9978</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

It can be seen from these scores that the RPSS values for wet and dry seasons for all the watersheds is ~0.4 and higher, except for wet season in the watershed PG. We also computed the RPSS for extreme flows at the 75th, 85th and 95th percentiles as thresholds. The skill of the ensemble forecasts in capturing the historic flows is very high, indicating that the ensemble forecasts are effective at forecasting flow extremes, which is extremely important for flood management and for water resources management.

A representative wet and dry year are selected to show the ensemble forecasts of seasonal total flow. Figure 4.8 shows the seasonal climate forecast during the dry year of 2007 and a wet year of 2006. The forecasts are quite strong in that during the dry year the forecast of seasonal rainfall being in the lower tercile to have a probability of 40% and 20% probability of being in the upper tercile. Similarly for the wet year, the forecasts show 40% - 50% probability of being in the upper tercile and 15% - 25% probability in the lower terciles during the dry and wet seasons (Figure 4.8). Translation of these seasonal climate forecasts to streamflow are shown in Figure 4.9 and Figure 4.10. The ensemble forecast flows are shown as boxplots alongside the boxplot of historic flows as climatology, the observed values in the selected years are shown as dots. In the
dry season during dry year of 2007 (Figure 4.9a) the seasonal flow ensembles for watershed CT are smaller than the climatology and the ensembles capture the observed flows. For watershed BD the ensembles capture the observed flow but the ensembles are wider compared to climatology, similarly for watershed BG. During the wet year (Figure 4.9b) of 2006 the seasonal flow ensembles in almost all the watersheds, especially in BG, GD and PG are shifted higher relative to climatology. Also the ensemble spread captures the observed flow in all watersheds. For the wet season during the dry year of 2007 (Figure 4.10a) and wet year of 2006 (Figure 4.10b) the ensembles are shifted in the right direction and they capture the observed flows quite well. These results indicate that the seasonal climate forecasts are skillfully translated to seasonal streamflow which can be further exploited for various applications.

Figure 4.8: Dry and wet season precipitation forecast for dry year (2007) and wet year (2006) from IRI Columbia University.
Figure 4.9: Boxplot of flow ensemble for the forecast and climatology forecast of peak dry season total flow for (a) dry year and (b) wet year. The observed flow is shown as blue dots.
4.6 Summary

In this study, we optimized VIC model application in Java using six objective functions: NSE, PBIAS, Pe, Log-RMSE, PEP and SFDCE to capture various aspects of watershed hydrology such as water balance, hydrograph shape, low flow, high flow and watershed's flashiness. We used parameters $b$, $Ds$, $thick_2$ and $thick_3$ for calibration which have been shown to be sensitive to NSE (Yanto et al., 2016b). Borg MOEA was employed to efficiently search model parameters that optimize the selected objective functions. Redundancy of objective functions was assessed based off their sensitivity to model parameters and baseflow index (BFI). We showed that three complimentary objective functions (NSE, PBIAS and SFDCE) are sufficient to generate model parameterization that quietly similar to those generated using six objective functions. We culled
the model parameters based on NSE > 0.75, PBIAS < 15% and SFDCE < 15% and showed that a set of model parameters satisfying at least one of the criteria yields less number of model parameters but comparable range of objective function values. Using the culled model parameters, the VIC model was used to generate ensemble of seasonal flow forecasts conditioned on seasonal climate forecast from IRI. We measured the skill of this seasonal forecast by computing the rank probability skill score (RPSS) of seasonal total flows and extremes at three different thresholds, for the dry and wet seasons. We showed that the RPSS of seasonal flows and the extremes are very good for both the seasons. Stochastic weather generators can be used to generate a rich variety of input meteorological ensembles (Verdin et al., 2015) conditioned on seasonal climate forecast, as opposed to resampling historical meteorology used here. Coupling the approach with short term weather forecasting can help with providing flood forecasts. Data issues remain a problem in this region, impacting the forecast skill. Overall, this study is important in providing multiobjective optimization framework for land surface model application in tropical regions and the utility of the framework for seasonal hydrologic forecast.
Chapter 5

Conclusion and Future Works

5.1 Summary

This research is motivated by the need for understanding and forecasting the hydroclimate over Indonesia to mitigate water-related natural hazards by efficient management of resources. To this end, this research makes four important contributions: (i) A systematic space-time analysis of seasonal precipitation over Indonesia was performed using Principal Component Analysis and Bayesian Dynamic Linear Model. (ii) A calibrated distributed VIC model was developed for six watersheds over Java, the most populous island of Indonesia. (iii) A high resolution gridded daily meteorology - precipitation, maximum and minimum temperature - data set at (14 km) resolution spanning 30 years from 1985 - 2014 over Java was developed using Inverse Distance Weighting method and, (iv) we employed multi-objective optimization based calibration of VIC model parameters, which improves upon the single objective based calibration. We demonstrated the utility of this in generating skillful seasonal hydrologic forecasts conditioned on seasonal climate forecasts.

In Chapter 1, A systematic space-time analysis of seasonal rainfall over Indonesia was performed using Principal Component Analysis and Bayesian Dynamic Linear Model. ENSO was found to be the driver of leading modes of variability during both the wet (Oct - Mar) and dry (Apr - Sep) seasons. Furthermore, ENSO appears to drive variability at multi-decadal timescales (8 - 16 year) especially during post 1980 period. During dry season the ITCZ is to the north of equator, leading to more spatial coherence of dominant pattern and the first two PCs account for 50% of variance. In wet season the ITCZ moves to the south of equator, leading to less spatial
coherence compared to dry season and the first two PCs account for 30% of variance. Interestingly, we find asymmetry in the ENSO teleconnections between wet and dry years during the dry season. The association between ENSO and Indonesian rainfall has strengthened in recent decades, especially during dry season. These findings suggest potential for interannual and multidecadal predictability of Indonesian rainfall.

In Chapter 2, to understand the processes that drive the spatial and temporal hydrologic variability a physically based land surface model was developed. Here, we developed a calibrated distributed VIC model for six watersheds over Java, the most populous island of Indonesia. Limited data and its poor quality posed significant challenge. The hydrologic processes - i.e., partitioning of rainfall into streamflow and evapotranspiration were modeled for the Java Island. The VIC model parameters were calibrated and validated on monthly time scale using a single objective function and, they provided insights into the hydrologic process such as surface flow, baseflow and evapotranspiration. Spatial and temporal variability of streamflow and key watershed attributes important for resource management such as baseflow index and evapotranspiration were analyzed to validate the model performance. This model represents the first such attempt to fully understand and model the hydrologic processes in Java. In the calibration period, the model performance in the watersheds was generally good with NSE ranging from 0.50 to 0.90. The NSE of the validation period was lower than expected, ranging from 0.09 to 0.45. Our analysis attributes this drop in performance to precipitation measurement issues and is also indicative of model over-fitting during the calibration period. Nonetheless, diagnosing errors was useful in identifying key hydrologic features of the system - i.e. the magnitude and variability of baseflow and direct runoff components, which are informative as drought and flood indices. The model exhibited better performance during the wet versus dry years, with relatively consistent performance across seasons. We provided preliminary evidence that performance can be improved by refining model spatial resolution. This work represents an important first step towards developing a platform for skillful seasonal, sub-seasonal and multi-decadal hydrologic projection systems, important for natural resources management and natural hazard mitigation in this populous island of Indonesia.
Data quality issue is a common problem in tropical watersheds especially in Java Island and most of Indonesia. In the third chapter, a fine resolution gridded daily meteorology - rainfall, maximum and minimum temperature - product was developed exclusively from ground-based observations. The gridded data at 0.125° x 0.125° resolution is the highest resolution product covering a 30-year period (1985 - 2014) ever created in Indonesia. In this, we used 765 precipitation stations to develop 707 grids over the island. IDW with radius of influence of 25 km and the power of 3 was selected as the interpolation method. This product, stored in a NetCDF will be useful for short-term and long-term studies for climate, hydrology and ecology at basin-scale and island-scale.

For skillful ensemble hydrologic forecasts, the ability to capture uncertainty in model parameters is important. In the fourth chapter, the VIC model was set up using the high resolution daily meteorology from previous chapter along with a multi-objective optimization based calibration approach, as opposed to single objective optimization. We employed six objective functions: NSE, PBIAS, Pe, Log-RMSE, PEP and SFDCE for VIC model optimization for Java to capture various aspects of watershed hydrology. Through sensitivity analysis, we found NSE, PBIAS and SFFDCE as critical objective functions. The method generates an ensemble of model parameters optimizing on these three objective functions. The parameters were culled based on strict criteria of these objective functions to reduce the ensemble size to set of robust parameters in terms of the objective functions. In addition to validating the multi-objective parameters, we demonstrated the utility of this in generating skillful ensemble seasonal hydrologic forecasts conditioned on seasonal climate forecasts. Flow ensembles were generated from these parameter ensembles along with a combination of input meteorology consistent with seasonal climate forecast. We show significant skill in the seasonal flow at all the basins, more so during dry seasons.

5.2 Discussion

The above contributions, especially the watershed modeling tools, are unique and are the first research efforts in this region. We offered insights into the space-time variability of precipitation and a robust physically-based watershed modeling tool to understand and forecast hydrologic variability
over Java in particular and Indonesia in general. Together, this research makes significant strides in providing a good framework for understanding, modeling and forecasting Indonesian hydrology and climatology, that will help mitigate natural hazards and enable efficient management of Indonesia's water and natural resources.

5.3 Future Work

This research sets stage for interesting and helpful extensions for Java and other parts of Indonesia, some of them are described below.

(i) The VIC model can be extended to other islands of Indonesia and compare their performances

(ii) Explore advanced statistical techniques such as spatial models, to improve the skill of gridded precipitation

(iii) Investigate the use of the VIC modeling framework for multi-decadal hydrologic projections. Which will be very helpful in long term water resources management and planning

(iv) Couple the developed VIC modeling framework with short term weather forecasts to develop a robust hydrologic forecasting system for floods, important for flood hazard mitigation

(v) Couple the VIC model with sediment models to understand the sedimentation problem which is significant in the rivers of Indonesia, and tropical rivers in general. This has an important bearing on water resources management
Bibliography


Ikeda et al. (2001). Failure of Pareto-based MOEAs: does non-dominated really mean near to optimal?, volume 2.


