Journal of Hydrology: Regional Studies 9 (2017) 127–139

Contents lists available at ScienceDirect

Journal of Hydrology: Regional Studies

journal homepage: www.elsevier.com/locate/ejrh

Hydrological model application under data scarcity for multiple watersheds, Java Island, Indonesia

Yanto [a,b,*], Ben Livneh [a,c], Balaji Rajagopalan [a,c], Joseph Kasprzyk [a]

[a] Civil, Environmental and Architectural Engineering, University of Colorado, Boulder, CO, United States
[b] Civil Engineering Department, Jenderal Soedirman University, Purwokerto, Indonesia
[c] Cooperative Institute for Research in Environmental Sciences (CIRES), University of Colorado, Boulder, CO, United States

ARTICLE INFO

Article history:
Received 25 September 2015
Received in revised form
15 September 2016
Accepted 24 September 2016

Keywords:
Hydrologic process
Physical modeling
VIC
Model calibration
Tropical watershed
Java island

ABSTRACT

Study region: Java Island, Indonesia.
Study focus: The Indonesian island of Java is home to more than half of Indonesia's population and routinely experiences water related natural disasters. This study represents a first step towards skillful hydrologic prediction and hydrologically-informed mitigation strategies. This is the first study to collate a comprehensive suite of hydrometeorological observations and systematically identify Variable Infiltration Capacity (VIC) Land Surface Model (LSM) parameters on Java to create a set of benchmark simulations.
New hydrological insights for the region: Quality control procedures revealed inconsistencies between precipitation and streamflow with only five watersheds possessing data of suitable quality. Simulations and observations confirmed that both precipitation and streamflow variability increase eastward on the island and that rainfall-runoff response was most frequently dominated by baseflow, rather than surface runoff. The most sensitive VIC parameters were identified and then calibrated with an automatic calibration procedure. In the calibration period, model performance was generally deemed satisfactory with Nash Sutcliffe Efficiency (NSE) between 0.31 to 0.89, whereas the validation period exhibited poorer performance than expected ($0.07 < NSE < 0.79$). This drop in performance was attributed to a combination of inconsistent data quality, hydrometeorological outliers during the validation period, and over-fitting parameters during the calibration period. The model indicated that direct runoff exhibits more spatial and temporal variability than both rainfall or baseflow, the latter being associated with variability of soil thickness.

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.

* Corresponding author at: Civil Engineering Department, Jendearal Soedirman University, Purwokerto, Indonesia.
E-mail address: masyanto79@gmail.com (Yanto).

http://dx.doi.org/10.1016/j.ejrh.2016.09.007

Table 1
Hydro-climatic characteristics of Java.

Components	Description
Annual rainfall	Spatial: - 3300 mm in southwest - 2000 mm in northeast Temporal: - 655 mm in dry season - 1800 mm in wet season
Temperature	28 °C coastal plains, 26 °C inland and mountain area, 23 °C higher mountain region
Land cover	61% agricultural area, 25% forest, 10% settlement, 4% water body

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., Wood et al., 2002, 2005; Regonda et al., 2013). 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 (Linde et al., 2008; Livneh et al., 2011). 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; 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 km^2) in East Java (Apip et al., 2012) and streamflow modeling on the Upper Citarum Basin (~1821 km^2) 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; Linde et al., 2008; Liang et al., 1994, 1996; Sheffield and Wood, 2007; Sheffield et al., 2006; 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., 2006). 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 daily and seasonal hydrology and their variability, consisting of model calibration and validation in unique time periods. 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. The paper concludes with summary and discussion of the results.

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 (Fig. 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 (Fig. 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 recent study by Yanto et al. (2016a). The land cover of the island consists mainly of tropical forest, as well as agriculture and urban settlements with significant spatial variability (Fig. 1b). Table 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 were downscaled from Nijssen et al. (2014) from $0.5^\circ \times 0.5^\circ$ to $0.125^\circ \times 0.125^\circ$ resolution using Inverse Distance Weighting (IDW) method, while the vegetation covers were directly derived from the University of Maryland's 1 km Global Land Cover product (Hansen et al., 2000). Four meteorological fields: daily total precipitation,

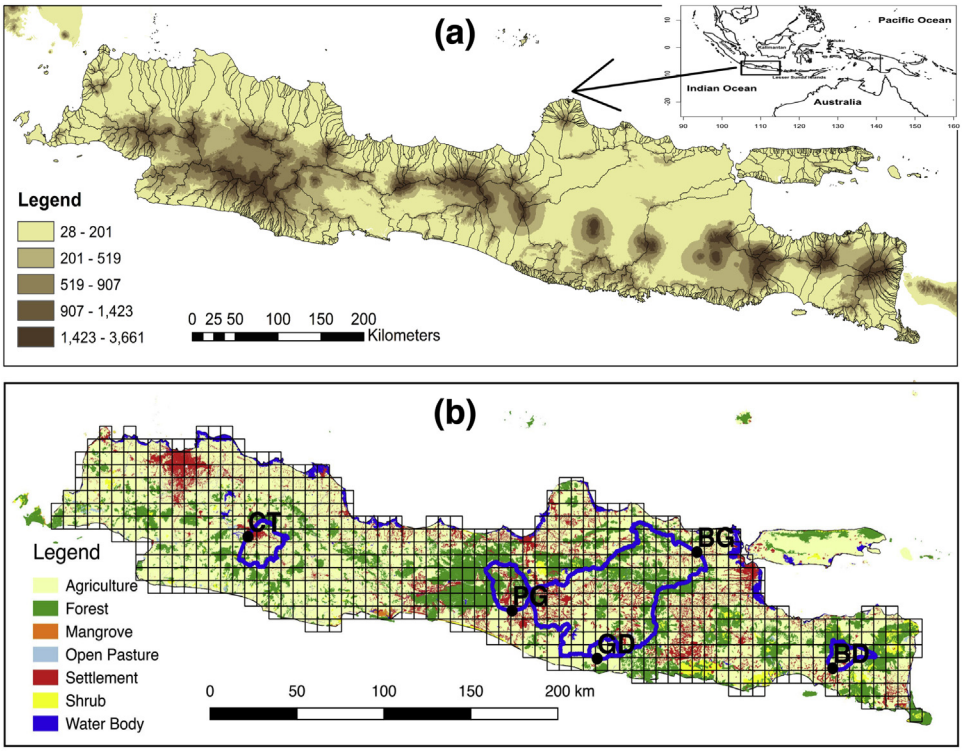

Fig. 1. Topography (a) and land cover overlaid by the model grid (b) of the Java Island. Mountain ranges extend across the middle of island from west to east where agricultural areas dominate.

maximum and minimum temperature, and average wind speed on a $0.125^\circ \times 0.125^\circ$ grid running from 1985 to 2014 were obtained from Yanto et al. (2016b). The precipitation in this dataset was exclusively developed from large number of ground-based observations. To assess and improve the model performance through calibration, observed streamflow was obtained from the Center of Water Resources Research and Development, Ministry of Public Works, Indonesia, the dominant source of ground-based precipitation used in this study.

2.1. 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 19 potential watersheds shown in Fig. 4. To this end, we applied the following quality control steps:

(1) 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.
(2) 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 Fig. 2 along with their respective runoff ratio (RR),

$$RR = \frac{\bar{Q}}{\bar{P}} \tag{1}$$

where $\bar{Q}$ and $\bar{P}$ are the mean streamflow and rainfall, respectively. Boxplots of rainfall and streamflow are shown in Fig. 3 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 ~10% to ~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 1), we selected watersheds with similar patterns of rainfall and streamflow climatology and similar long-term mass balance. Based on these criteria, we selected CT, PG, BG, GD and BD shown in color in Fig. 4 and listed in Table 2. The smallest watershed (GD) and the largest watershed

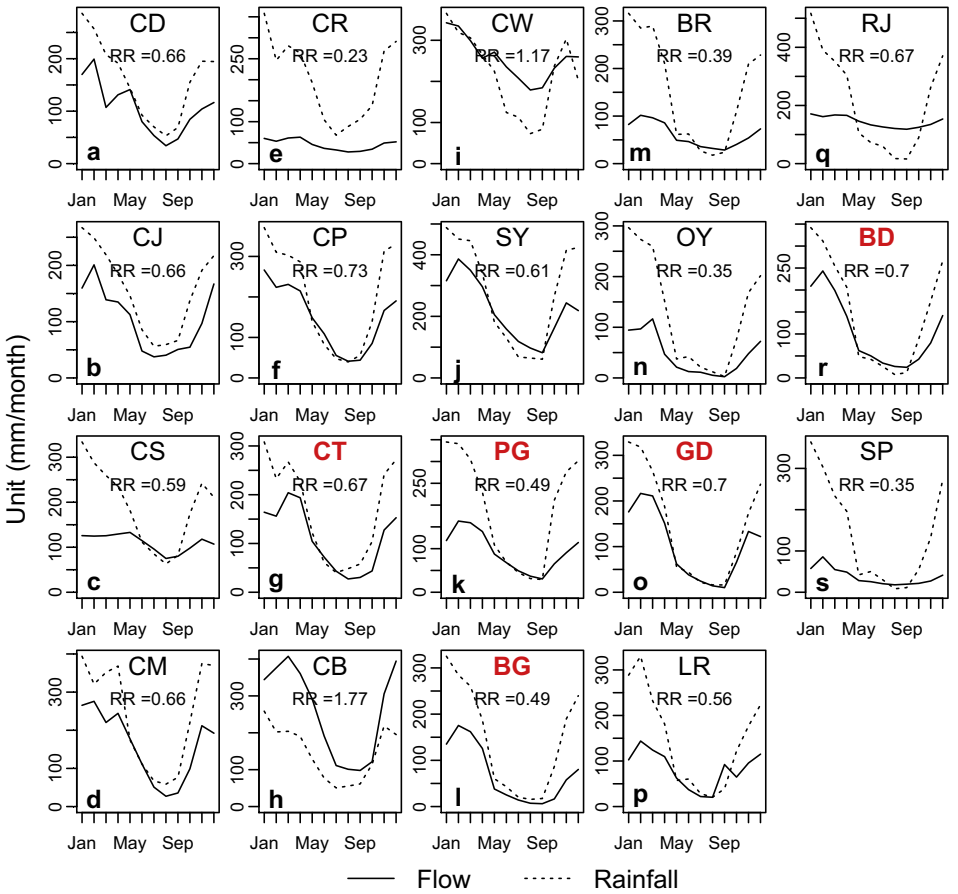

Fig. 2. Climatology of rainfall and streamflow of 19 watersheds covering Java, each of which has at least 7 years of data. A runoff ratio (RR) is computed for each watershed as mean precipitation divided by mean flow over the entire period of record and shown in the panels. Watersheds selected for this study are indicated in red. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 2
Characteristics of the five 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.

Watershed	Outlet Lon/Lat (°)	Outlet Elevation (m)	Mean stream-flow (mm/yr)	Streamflow record period	Area (km^2)	Mean rainfall (mm/yr)	Land cover	Runoff Ratio (Q/P)
Citarum (CT)	107.53/−6.95	668	1351	1990–1997	1675	2739	**A(59)**, F(27), S(13),W(1)	0.50
Progo (PG)	110.26/−7.67	180	1115	1996–2002	1676	2412	A(15), **F(79)**, S(5), W(1)	0.48
Bengawan (BG)	112.17/−7.10	11	885	1990–2002	16286	1993	**A(56)**, F(34), S(9), W(1)	0.44
Grindulu (GD)	111.14/−8.14	39	1094	1990–2002	2236	2238	**A(58)**, F(42), S(0), W(0)	0.56
Bedadung (BD)	113.58/−8.23	44	1322	1991–2001	696	2504	A(21), **F(74)**, S(5), W(0)	0.50

(BG) receive contributions from 9 and 102 grid cells respectively, such that the application of the model in this basin can be considered distributed.

3. 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 through-fall, 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).

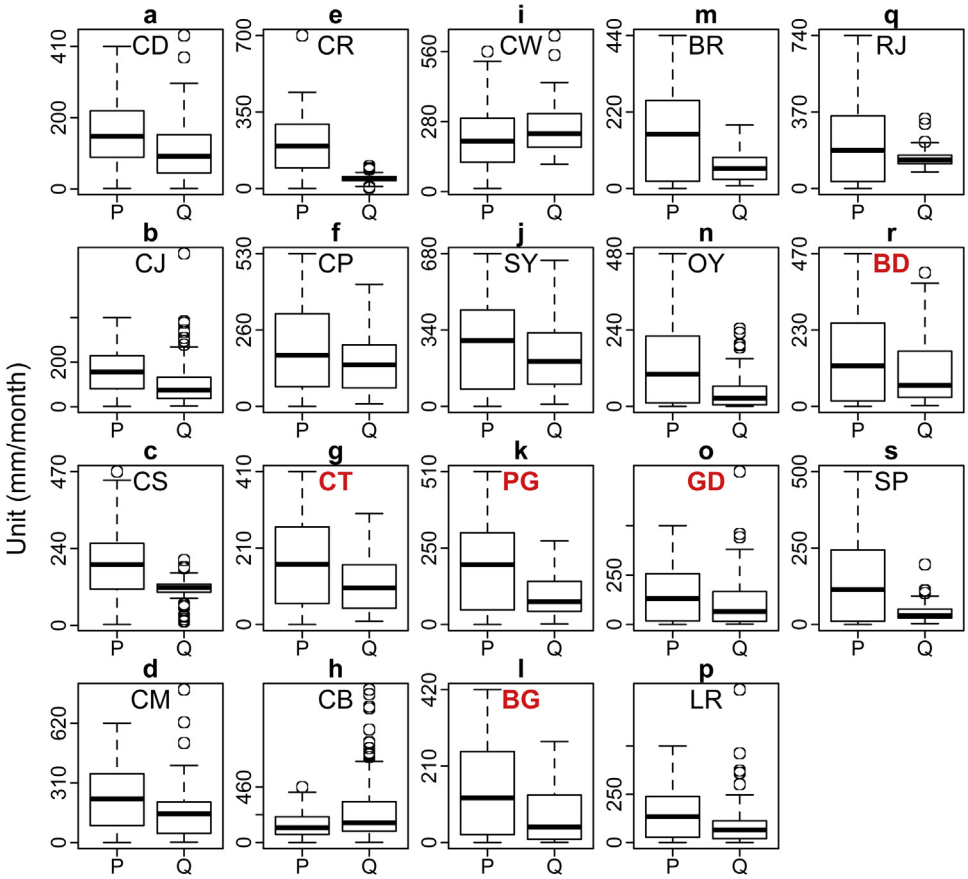

Fig. 3. Boxplot of monthly precipitation and streamflow for the 19 watersheds covering Java. The period and length of record for each watershed are different as presented in Table 2. Watersheds selected for this study are marked in red. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

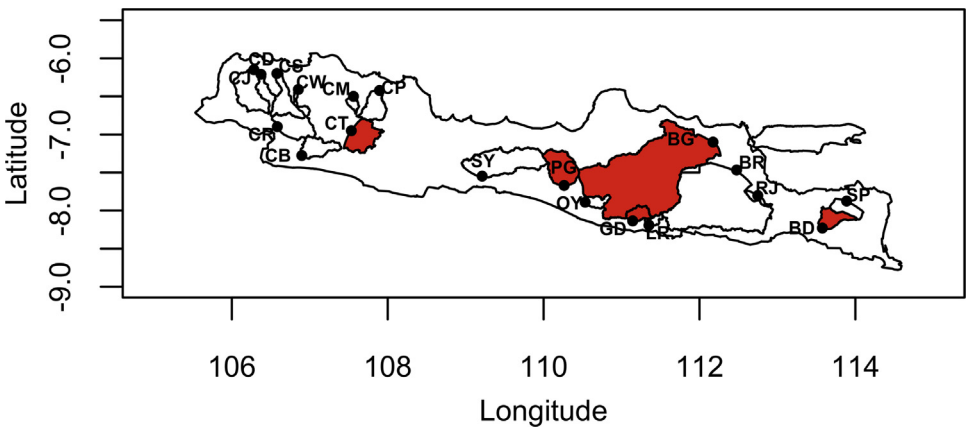

Fig. 4. The 19 watersheds covering Java with at least 7 years of data. The five watersheds shown in red were selected for this study after the screening process. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

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 a 24-h (i.e. daily) time step on a $0.125^\circ \times 0.125^\circ$ grid. The model was forced with daily precipitation, maximum and minimum temperature and wind speed.

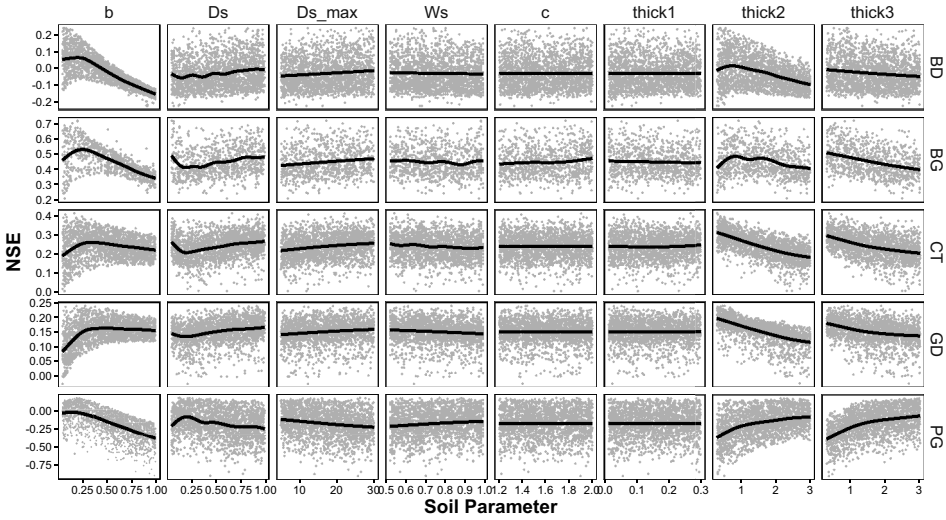

Fig. 5. Scatterplots of NSE versus soil parameter values from a Monte Carlo sampling of the parameter space. A local polynomial smoother through the scatterplot is shown as solid line. Results are used to determine sensitivity of NSE to changes in soil parameters, and choose soil parameters for model calibration.

3.1. Parameter estimation – calibration and validation

Model calibration is a process of identifying 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 daily time step comparing VIC simulations with historic streamflow observations. The Nash-Sutcliffe Efficiency (NSE) metric was used to evaluate the performance of model parameterizations (Eq. (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:

$$NSE = 1 - \frac{\sum_{t=1}^{T}\left(Q_o^t - Q_s^t\right)^2}{\sum_{t=1}^{T}\left(Q_o^t - Q_{\bar{o}}\right)^2} \tag{2}$$

where Q_o^t and Q_s^t are observed and simulated daily streamflow at time t respectively, $Q_{\bar{o}}$ 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 Motovilov et al. (1999) the model performance at daily time step is generally judged as satisfactory when NSE > 0.36, 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 Xie et al. (2001) and 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 five study watersheds.
(3) We produced a scatterplot of each soil parameter and the corresponding NSE for each watershed in Fig. 5.

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 Fig. 5 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 Fig. 5, parameter *b*, *Ds*, $thick_2$ and $thick_3$ were determined to be the most

Table 3
Description of calibration parameters, their ranges and the values obtained from optimization for the five watersheds.

Parameter	b	Ds	$thick_2$	$thick_3$	v	D
Description	infiltration curve shape parameter controlling surface runoff	fraction of maximum baseflow where nonlinear baseflow begins	thickness of soil layer 2	thickness of soil layer 3	flow velocity	flow diffusivity
Unit	NA	NA	m	m	m/s	m^2/s
Range	>0–1	>0–1	0.1–3	0.1–3	0.5–3	200–4000
Optimal values						
CT	0.99	0.99	0.30	0.30	2.99	200
PG	0.48	0.29	2.99	2.99	2.99	200
BG	0.86	0.01	0.32	2.99	2.99	204
GD	0.99	0.52	0.30	0.45	2.99	200
BD	0.55	0.99	0.30	0.87	2.99	200

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. The description of calibration parameters is presented in Table 3.

Running and assessing VIC model at daily time step require a routing scheme to take into account for dynamic partial contributions of each grid within a watershed to downstream total flow. As routing model is considered as VIC post-processing, it requires a separate parameter estimation. For this, we followed the calibration procedure outlined by Lohmann et al. (1996, 1998). Accordingly, we additionally calibrated flow velocity (v) and flow diffusivity (D) of the routing model (Table 3).

After the sensitivity analysis was completed, we conducted the calibration using the four most sensitive parameters illustrated in Fig. 5 and two routing model parameters shown in Table 3. 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 six selected model parameters were estimated by optimizing the NSE over the first half (~5 years) of the streamflow record (Table 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.

4. Results

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).

4.1. Model performance

The observed and simulated streamflow for each watershed in the calibration (Fig. 6, left column) and validation period (Fig. 6, right column) are shown as time series (monthly) and scatterplots along with the NSE values (daily). In the calibration period, NSE values range from 0.31 to 0.89, with three of them exceeding 0.36, indicating that the calibrated model performance in all the watersheds ranges between satisfactory and good (Motovilov et al., 1997). The simulations consistently fit the rising limb, peak and recession limb of the observed hydrograph during the calibration period (Fig. 6, left panel). In all watersheds, the validation NSE values are lower ranging from 0.07 to 0.79 with two basins perform satisfactorily. Although the validation NSE values are expected to be lower, the marked drop of NSE values in PG and GD are unexpected.

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 Fig. 6, right panel – particularly in watershed PG where in 2 out of 3 years in the validation period the observed streamflow peak lags the rainfall. 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 Fig. 6. This discrepancy certainly contributes to the significant drop in model performance in the validation period in watersheds PG, GD and BD. To understand the cause of this difference, we calculated the change

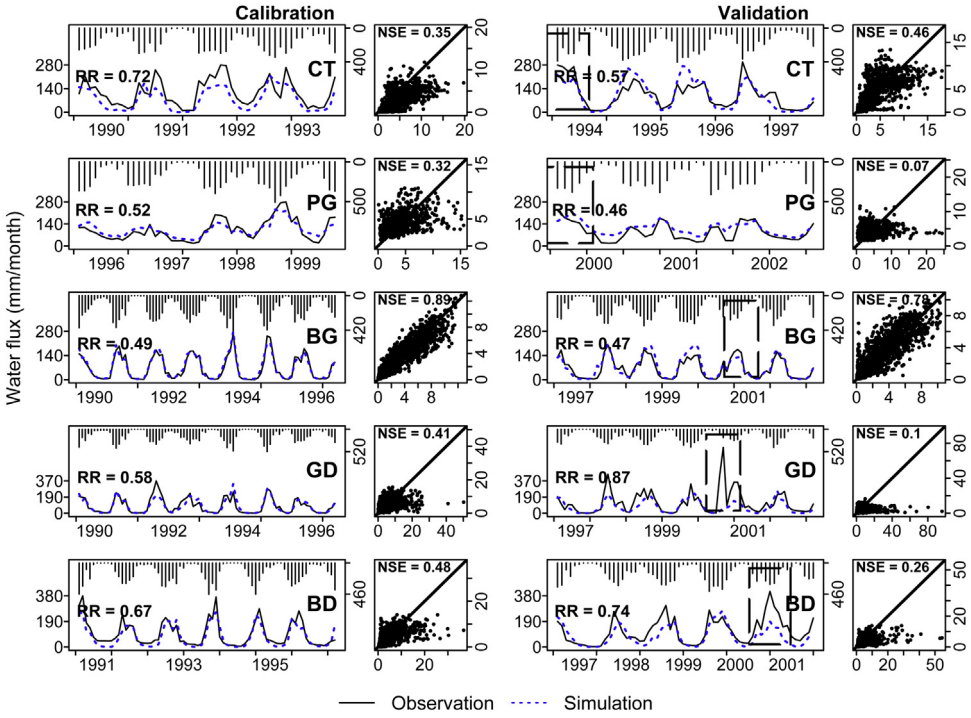

Fig. 6. Time series (monthly) and scatterplots (daily) of observed and simulated streamflow and rainfall hyetograph for CT, PG, BG, GD and BD watersheds. The left panels show the calibration period and the right shows validation. In the time series plots, left and right y-axes are for streamflow and 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.

in mean rainfall and streamflow between the calibration and validation periods (Table 4). The percent changes between rainfall and streamflow differ substantially in almost all the watersheds. In particular, nearly all study watersheds 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 (Fig. 6) 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.

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., 2016a) – defined as Nov–Apr for wet and May–Oct for dry season. Table 5 shows NSE in dry and wet seasons for the five watersheds. The substantial difference of model performance exists in all watersheds except in basin PG. 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).

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.

Table 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.

Watershed	Daily mean streamflow (mm)			Daily mean rainfall (mm)			Change in runoff ratio (%)
	Calibration	Validation	Change (%)	Calibration	Validation	Change (%)	
CT	3.87	3.44	−11.01	5.32	6.24	17.29	−24.19
PG	2.95	3.42	16.12	6.67	6.66	−0.09	16.22
BG	2.37	2.27	−4.03	5.00	5.06	1.19	−5.15
GD	2.83	4.03	42.38	5.15	5.00	−2.88	46.61
BD	3.14	3.82	21.71	4.94	5.46	10.45	10.19

Table 5
NSE values for all watersheds in dry and wet season.

Watershed	Dry Season NSE	Wet Season NSE
CT	0.12	−0.07
PG	0.18	0.18
BG	0.74	0.81
GD	0.31	0.25
BD	−0.59	0.26

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). Fig. 7 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 basin CT. 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, PG, BG, GD and GD is higher than VIC ET by 48%, 19%, 34%, 61% and 41% 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 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 basin 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.

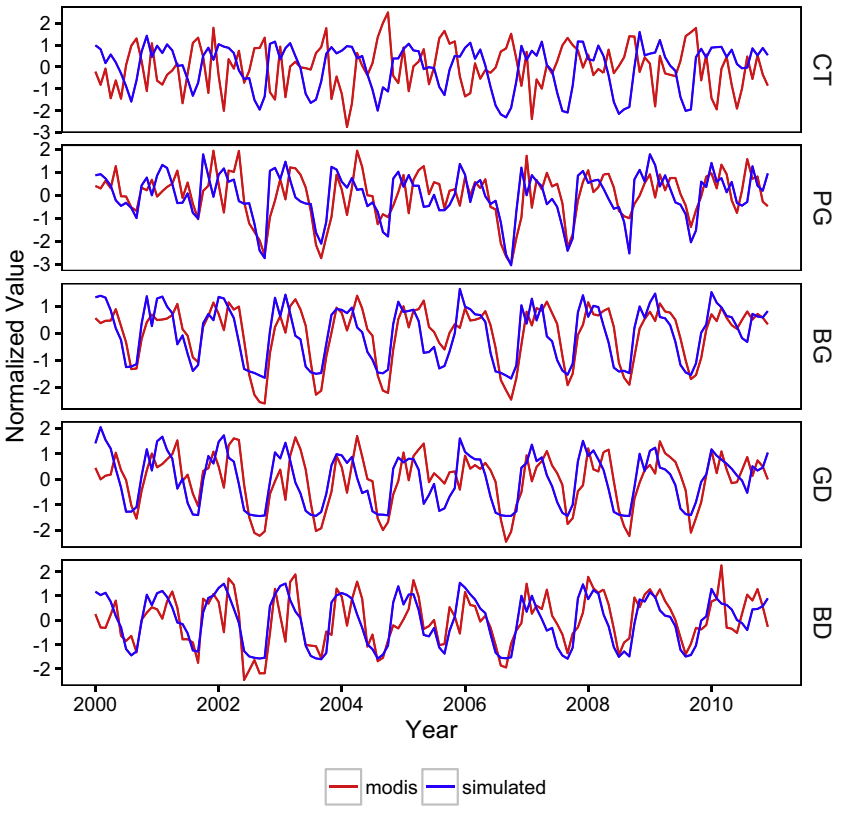

Fig. 7. Time series of normalized evapotranspiration (divide each anomaly by standard deviation of the data) from the model simulation and MOD16 for the period 2000–2010.

Table 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.

Variable	Unit	CT	PG	BG	GD	BD
Water balance						
Rainfall	Total (mm yr^{-1})	1942	2434	1826	1960	1805
	Coefficient of Variation (CV)	1.02	1.17	1.23	1.39	1.57
Evaporation	Total (mm yr^{-1})	1010	1149	926	1043	902
	Coefficient of Variation (CV)	0.54	0.35	0.54	0.79	0.84
Direct runoff	Total (mm yr^{-1})	499	569	588	414	425
	Coefficient of Variation (CV)	1.34	1.39	1.46	1.81	2.08
Baseflow	Total (mm yr^{-1})	433	666	304	504	476
	Coefficient of Variation (CV)	1.02	0.33	1.19	1.12	1.18
Baseflow Index	%	46	54	34	41	53
b		0.99	0.48	0.86	0.99	0.55
Ds		0.99	0.29	0.01	0.52	0.99
$thick_2$	m	0.30	2.99	0.32	0.30	0.30
$thick_3$	m	0.30	2.99	2.99	0.45	0.87

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 (e.g., 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 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 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 1.31, 4.21, 1.23, 1.62 and 1.76 for the watersheds CT, PG, BG, GD and BD respectively. As shown in Table 6, this pattern is consistent with that of the calibrated thickness of soil layer 2, $thick_2$. Also, both the ratio of CV of direct runoff and baseflow and $thick_2$ are highly correlated, exhibiting a R^2 score of 0.97.

The distribution of BFI over the watersheds shows a low degree of variability (~0.34–0.54) – consistent with the fact that the watersheds all have similar geology. In addition, the BFI is highly correlated with 1/b (R^2 of 0.62) – a key variable in the infiltration capacity equation used by the model (Liang et al., 1994). This intuitive result confirms that the soil characteristics

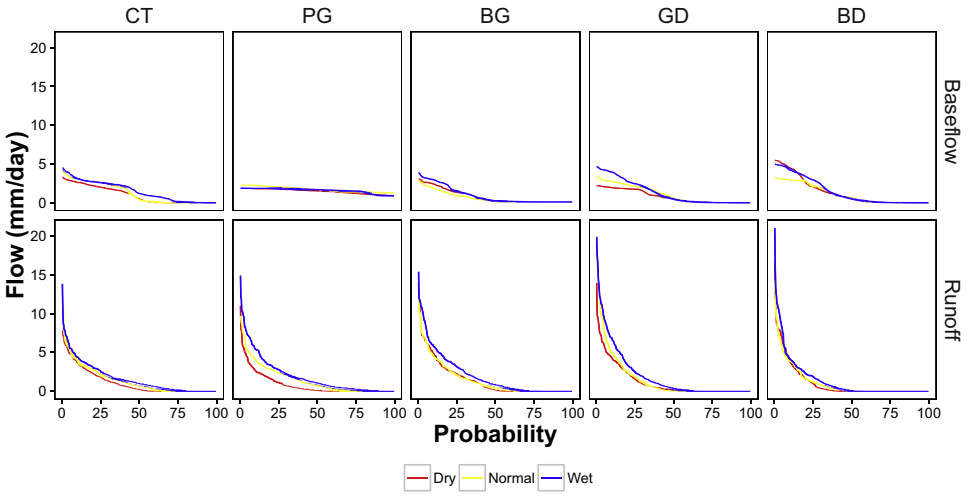

Fig. 8. Flow duration curve of monthly simulated streamflow in the wettest, average and driest year chosen from the calibration period of each watershed.

Table 7
NSE of the driest, average and wettest years in the watersheds.

Watershed	Driest year NSE	Normal year NSE	Wettest year NSE
CT	0.33	0.38	0.25
PG	0.41	−0.09	0.27
BG	0.84	0.88	0.89
GD	0.24	0.30	0.52
BD	0.62	0.30	0.48

in the top layer have a strong influence in transferring the rainfall into both direct runoff and baseflow and consequently the BFI.

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 Fig. 8. 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. Fig. 8 demonstrates that surface runoff tends to be more sensitive to climate variability 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 and PG) during the wettest year. We computed the NSE scores for these three years (Table 7) indicating that the model performances are spatially comparable during the wettest and driest years.

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 island's 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 19 watersheds and after screening for physical consistency between precipitation and streamflow as well as missing data, five 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 0.31 to 0.89. The NSE of the validation period was lower than expected with skill between NSE = 0.07 to NSE = 0.79. 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.

Conflicts of interest

None.

Acknowledgments

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. We also acknowledge the anonymous reviewers, whose comments greatly improved the paper.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ejrh.2016.09.007.

References

Apip, Sayama, T., Tachikawa, Y., Takara, K., 2012. Spatial lumping of a distributed rainfall-sediment-runoff model and its effective lumping scale. Hydrol. Process. 26, 855–871.
Arnold, J.G., Srinivasan, R., Muttiah, R.S., Allen, P.M., 1999. Continental scale simulation of the hydrologic balance. J. Am. Water Resour. Assoc. 35 (5), 1037–1051.
Badan Nasional Penanggulangan Bencana (BNPB), 2015. Data Dan Informasi Bencana Indonesia, Retrieved from http://dibi.bnpb.go.id/data-bencana.
Bappenas, 2010. Indonesian climate change sectoral roadmap. Synth. Rep. 1, 19–24.
Beck, H.E., Bruijnzeel, L.A., van Dijk, A.I.J.M., McVicar, T.R., Scatena, F.M., Schelleken, J., 2013. The impact of forest regeneration on streamflow in 12 mesoscale humid tropical catchments. Hydrol. Earth Syst. Sci. 17, 2613–2635.
Beven, K., 1989. Changing ideas in hydrology – the case of physically based models. J. Hydrol. 105, 157–172.
Cook, P.G., Hatton, T.J., Pidsley, D., Herczeg, A.L., Held, A., O'Grady, A., Eamus, D., 1998. Water balance of a tropical woodland ecosystem, Northern Australia: a combination of micro-meteorological, soil physical and groundwater chemical approaches. J. Hydrol. 210, 161–177.
Demaria, E.M., Nijssen, B., Wagener, T., 2007. Monte Carlo sensitivity analysis of land surface parameters using the variable infiltration capacity model. J. Geophys. Res. 112 (D1113), 1–15.
Dettinger, M.D., Diaz, H.F., 2000. Global characteristics of stream flow seasonality and variability. J. Hydrometeorol. 1, 289–310.
Dubreuil, P.L., 1985. Review of field observation of runoff generation in the tropics. J. Hydrol. 80, 237–264.
Eckhardt, K., 2008. A comparison of baseflow indices, which were calculated with seven different baseflow separation methods. J. Hydrol. 352, 168–173.
Ferguson, C.R., Sheffield, J., Wood, E.F., Gao, H., 2010. Quantifying uncertainty in a remote sensing based estimate of evapotranspiration over the continental United States. Int. J. Remote Sens. 31 (14), 3821–3865.
Gonzales, A.L., Nonner, J., Heijkers, J., Uhlenbrook, S., 2009. Comparison of different base flow separation methods in a lowland catchment. Hydrol. Earth Syst. Sci. 13, 2055–2068.
Gupta, H.V., Kling, H., Yilmaz, K.K., Martinez, G.F., 2009. Decomposition of the mean squared error and nse performance criteria: implications for improving hydrological modelling? J. Hydrol. 377 (1–2), 80–91.
Gutman, G., Csiszar, I., Romanov, P., 2000. Using NOAA/AVHRR products to monitor El Nino impacts: focus on Indonesian in 1997–98. Bull. Am. Meteorol. Soc. 81, 1189–1205.
Hadka, D., Reed, P., 2013. Borg: an auto-adaptive many-objective evolutionary computing frame- work? Evol. Comput. 125 (6), 333–341.
Hansen, M.C., DeFries, R.S., Townshend, J.R.G., Sohlberg, R., 2000. Global land cover classification at 1 km resolution using a classification tree approach. Int. J. Remote Sens. 21, 1331–1364.
Harlan, D., Wangsadipura, M., Munajat, C.M., 2010. Rainfall-runoff modeling of Citarum hulu river basin by using GR4J. In: Paper Presented at the World Congress on Engineering, London, June 30–July 2.
Jakeman, A.J., Hornberger, G.M., 1993. How much complexity is warranted in a rainfall-runoff model? Water Resour. Res. 29, 2637–2649.
Julian, M.M., Fink, M., Fischer, C., Krause, P., Flugel, W.-A., 2013. Implementation of J2000 hydrological model in the western part of Java Island, Indonesia. JMAS 1, 1–25.
Kim, H.W., Hwang, K., Mu, Q., Lee, S.O., Choi, M., 2012. Validation of MODIS 16 global terrestrial evapotranspiration products in various climates and land cover types in asia. KSCE J. Civil Eng. 16 (2), 229–238.
Krave, A.S., van Straalen, N.M., van Verseveld, H.W., Roling, W.F.M., 2007. Influence of El Nino and La Nina climate events and litter removal on inorganic nitrogen dynamics in pine forest soils on Central Java, Indonesia. Eur. J. Soil Biol. 43, 39–47.
Kumagai, T., Yoshifuji, N., Tanaka, N., Suzuki, M., Kume, T., 2009. Comparison of soil moisture dynamics between a tropical rain forest and a tropical seasonal forest in Southeast Asia: impact of seasonal and year-to-year variations in rainfall. Water Resour. Res. 45 (W04413), 1–14.
Legates, D.R., McCabe, G.J., 1999. Evaluating the use of goodness-of-fit measures in hydrologic and hydroclimatic model validation. Water Resour. Res. 35 (1), 233–241.
Liang, X., Lettenmaier, D.P., Wood, E.F., Burges, S.J., 1994. A simple hydrologically based model of land surface water and energy fluxes for general circulation models. J. Geophys. Res. 99 (14), 415–14,428.
Liang, X., Wood, E.F., Lettenmaier, D.P., 1996. Surface soil moisture parameterization of the VIC-2L model: evaluation and modifications. Global Planet Change 13, 195–206.
te Linde, A.H., Aerts, J.C.J.H., Hurkmans, R.T.W.L., Eberle, M., 2008. Comparing model performance of two rainfall-runoff models in Rhine basin ysing different atmospheric forcing data sets. Hydrol. Earth Syst. Sci. 12, 943–957.
Livneh, B., Lettenmaier, D.P., 2013. Regional parameter estimation for the unified land model. Water Resour. Res., http://dx.doi.org/10.1029/2012wr012220.
Livneh, B., Restrepo, P.J., Lettenmaier, P.J., 2011. Development of a unified land model for prediction of surface hydrology and land-atmosphere interactions. J. Hydrometeorol. 12, 1299–1320.
Lohmann, D., Nolte-Holube, R., Raschke, E., 1996. A large-scale horizontal routing model to be coupled to land surface parametrization schemes. Tellus A 48 (5), 708–721.
Lohmann, D., Raschke, E., Nijssen, B., Lettenmaier, D.P., 1998. Regional scale hydrology: II application of VIC-2L model to the weser river, Germany. J. Hydrol. Sci. 43 (I), 143–158.
Longobardi, A., Villani, P., 2008. Baseflow index regionalization analysis in a mediterranean area and data scarcity context: role of the catchment permeability index. J. Hydrol. 355, 63–75.
Moriasi, D.N., Arnold, J.G., Van Liew, M.W., Bingner, R.L., Darmel, R.D., Veith, T.L., 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. ASABE 50 (3), 885–900.
Motovilov, Y.G., Gottschalk, L., Engeland, K., Rodhe, A., 1999. Validation of a distributed hydrological model against spatial observations. Agric. For. Meteorol. 9899, 257–277.
Mu, Q., Heinsch, F.A., Zhao, M., Running, S.W., 2007. Development of a global evapotranspiration algorithm based on MODIS and global meteorology data. Remote Sens. Environ. 111, 519–536.

Muñoz-Villers, L.E., McDonnell, J.J., 2012. Runoff generation in a steep, tropical montane cloud forest catchment on permeable volcanic substrate. Water Resour. Res. 48 (W09528), 1–17.
Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models: part 1. a discussion of principles. J. Hydrol. 10 (3), 282–290.
Nijssen, B., O'Donnell, G.M., Lettenmaier, D.P., Lohmann, D., Wood, E.F., 2001. Predicting the discharge of global rivers. J. Climate 14, 3307–3323.
Nijssen, B., Shukla, S., Lin, C., Gao, H., Zhou, T., Ishottama Sheffield, J., Wood, E.F., Lettenmaier, D.P., 2014. A prototype global drought information system based on multiple land surface models. J. Hydrometeorol. 15, 1661–1676.
Qian, J.H., Robertson, A.W., Moron, V., 2010. Interactions among ENSO, the Monsoon, and diurnal cycle in rainfall variability over Java, Indonesia. J. Atmos. Sci. 67, 3509–3524, http://dx.doi.org/10.1175/2010JAS3348.1.
Refsgaard, J.C., 1997. Parameterisation, calibration and validation of distributed hydrological models. J. Hydrol. 198, 69–97.
Reed, P., Hadka, D., Herman, J.D., Kasprzyk, J.R., Kollat, J.B., 2013. Evolutionary multiobjective optimization in water resources: the past, present, and future. Adv. Water Resour. 51, 438–456.
Regonda, S.K., Seo, D.J., Lawrence, B., Brown, J.D., Demargne, J., 2013. Short-term ensemble streamflow forecasting using operationally-produced single-valued streamflow forecasts–a Hydrologic Model Output Statistics (HMOS) approach. J. Hydrol. 497, 80–96.
Ruhoff, A.L., Paz, A.R., Aragao, L.E.O.C., Mu, Q., Malhi, Y., Collischonn, W., Rocha, H.R., Running, S.W., 2013. Assessment of the MODIS global evapotranspiration algorithm using eddy covariance measurements and hydrological modelling in the Rio Grande basin. Hydrol. Sci. J., http://dx.doi.org/10.1080/02626667.2013.837578.
Sevat, E., Dezetter, A., 1991. Selection of calibration objective functions in the context of rainfall-runoff modeling in a Sudanese savannah area. Hydrol. Sci. J. 36 (4), 307–330.
Sheffield, J., Wood, E.F., 2007. Characteristics of global and regional drought, 1950–2000: analysis of soil moisture data from off-line simulation of the terrestrial hydrologic cycle. J. Geophys. Res. 112 (D17115), 1–21.
Sheffield, J., Goteti, G., Wood, E.F., 2006. Development of a 50-year high-resolution global dataset of meteorological for land surface modeling. J. Climate 19, 3088–3111.
Shukla, S., Sheffield, J., Wood, E.F., Lettenmaier, D.P., 2013. On the sources of global land surface hydrologic predictability. Hydrol. Earth Syst. Sci. 17, 2781–2796.
Tang, R., Shao, K., Li, Z.L., Wu, H., Tang, B.H., Zhou, G., Member, S., IEEE, Zhang, L., 2015. Multiscale validation of the 8-day MOD16 evapotranspiration product using flux data collected in china. IEEE 8 (4), 1478–1486.
Wagener, T., Boyle, D.P., Lees, M.J., Wheater, H.S., Gupta, H.V., Sorooshian, S., 2001. A framework for development and application of hydrological models. J. Geophys. Res. Atmos. 5 (1), 13–26.
Wood, A.W., Maurer, E.P., Kumar, A., Lettenmaier, D.P., 2002. Long-range experimental hydrologic forecasting for the eastern United States. J. Geophys. Res. Atmos. 107 (D20), 1–15.
Wood, A.W., Kumar, A., Lettenmaier, D.P., 2005. A retrospective assessment of National Centers for Environmental Prediction climate model-based ensemble hydrologic forecasting in the western United States. J. Geophys. Res. Atmos. 110 (D4), 1–16.
Xie, Z., Yuan, F., Duan, Q., Zheng, J., Liang, M., Chen, F., 2001. Regional parameter estimation of the VIC land surface model: methodology and application to river basins in China. J. Hydrometeor. 8 (3), 447–468.
Yang, G., Bowling, L.C., Cherkauer, K.A., Pijanowski, B.C., Niyogi, D., 2010. Hydroclimatic response of watersheds to urban intensity: an observational and modeling-based analysis for the White River Basin, Indiana. J. Hydrometeor. 11, 122–138.
Yanto, Rajagopalan, B., Zagona, E., 2016a. Space-Time variability of indonesian rainfall at inter-annual and multi-decadal time scales. Clim. Dyn., http://dx.doi.org/10.1007/s00382-016-3008-8.
Yanto, Livneh, B., Rajagopalan, B., 2016b. Development of a meteorological data set over the java island, Indonesia 1985–2014. Nat. Sci. Data (under review).
Yasunari, T., 1981. Temporal and spatial variations of monthly rainfall in Java, Indonesia. Southeast Asian Stud. 19, 170–186.
Zhao, F.F., Zhang, L., Chiew, F.H.S., Vaze, J., 2011. The effect of spatial rainfall variability on streamflow prediction for a south-eastern Australian catchment. In: Paper Presented at 19th International Congress on Modelling and Simulation, Perth, Australia, 12–16 December 2011.