Set Up and Calibration of a Spatial Tool for Simulating River Discharge of Western Java in Recent Decades: Preliminary Results and Assessments

Poerbandono[1]*, Philip J. Ward[2] & Miga M. Julian[3]

[1]Hydrographic Science and Engineering Research Division
Faculty of Earth Sciences and Technology, Institut Teknologi Bandung
*email: poerbandono@gd.itb.ac.id
[2]Institute for Environmental Studies
Faculty of Earth and Life Sciences, VU University Amsterdam
[3]Geodesy and Geomatics Engineering Study Program
Faculty of Earth Sciences and Technology, Institut Teknologi Bandung

Abstract. This paper discusses a study of the application of global spatio-temporal climate datasets and the hydrological model STREAM (Spatial Tools for River Basin Environmental Analysis and Management Options). In the study, set up and calibration of STREAM for the reconstruction of monthly discharge for several locations in the western part of Java, Indonesia, for the period 1983-2002 are carried out. The set up includes the preparation of monthly precipitation and temperature datasets, a digital elevation model of the domain being studied, and maps of land cover and soil water holding capacity. Discharge observations from six stations located mostly in the upper parts of major watersheds in the domain are used to calibrate the model by comparing simulated and observed discharge variables. The model performs reasonably well. Comparison between computed and observed mean monthly discharges yield correlation coefficients ranging from 0.72 to 0.93. The computed mean annual discharge in five out of six observation stations ranges between -8 and 5% with respect to the mean annual observed discharge. This study offers a tool which can be used for reconstructing historical discharge.

1   Introduction

Historical records of water balance parameters, such as precipitation, evaporation, and run-off, provide invaluable information for detecting trends of environmental changes, as well as return periods of possible environmental disasters. Unfortunately, such records are usually very limited in quantity or, if sufficient, are not designed to meet the needs of scientific exploration. They may lack adequate temporal and spatial coverage, not meet acceptable accuracy, or not be accompanied by reliable documentation. In fact, with increasing awareness of global climate change, better insights into what has been

happening in recent decades are required. This could provide essential information for detecting long-term trends. In order to contribute to improving the understanding of recent long-term environmental trends, the water balance is simulated in this paper on decadal timescales. The objectives of this study are to investigate tools and their corresponding input datasets for simulating monthly discharge focusing on major watersheds flowing into the Jakarta Bay (Figure 1).
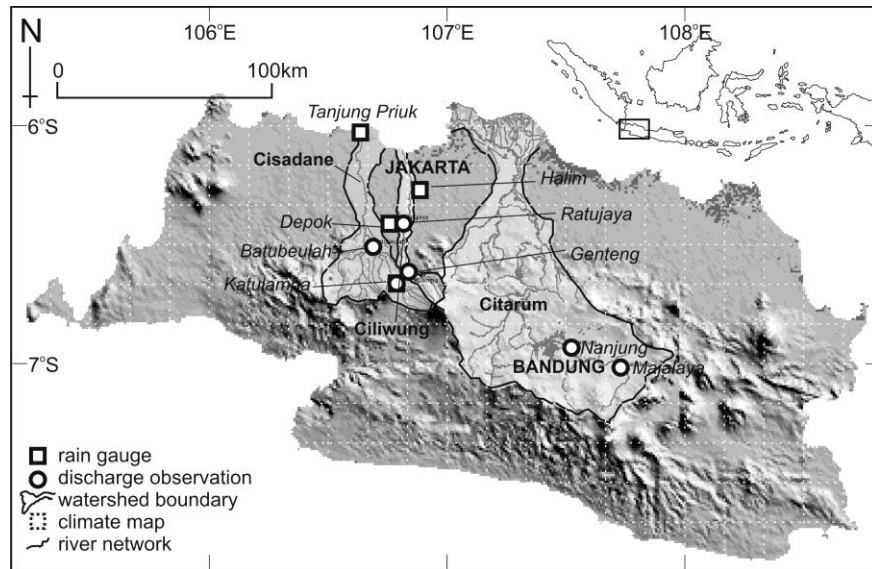

Figure 1 Study site with the watersheds flowing into the Jakarta Bay.

The results presented in this paper outline the set up and calibration of a spatial tool for simulating mainland discharge to the Jakarta Bay. The spatial tool used here is STREAM (Spatial Tools for River Basins and Environment and Analysis of Management Options) [1]. The set up includes the preparation of monthly precipitation and temperature datasets at a 10'×10' resolution. Historical observation data from the recent period, consisting of rainfall monitoring stations, are considered to evaluate whether the global datasets used here fit the local observation. At this initial stage, the geographic setting is left static and the reconstruction period is limited to the maximum length of the available observation data. The study site is situated in the north-western part of Java (see Figure 1). Input data for the model include a Digital Elevation Model (DEM), a land use map, a map of soil water holding capacity and maps of monthly climate (precipitation and temperature). Observed discharge data from stations located mainly in the upper and additionally in the middle parts of the

Cisadane, Ciliwung, and Citarum watersheds (see Figure 1) are used for carrying out the calibration of the model.

2   Discharge Computation

STREAM applies Thornthwaite-Mather's water balance approach [2] to calculate flow discharges along drainage networks derived from a DEM. The main inputs to the model are climate data (maps of precipitation and temperature), a DEM, land use maps and *WHC* maps. Potential and actual evapotranspiration are calculated in STREAM at the so-called soil compartment represented by a grid cell. Following this, storage of water in a grid cell is estimated according to the difference between evapotranspiration and precipitation. Finally, discharge per time step is calculated according to excess of water in each grid cell and baseflow from groundwater storage. The governing equations are described as follows. Potential evapotranspiration (ET_{pot}) is calculated using the Thornthwaite equations [3] and is defined as:

$$ET_{pot} = ET_{ref} \cdot CropF \cdot CropF_c \tag{1}$$

where, if $T \leq 26.5$ then $ET_{ref} = 16\left(10\frac{T}{H}\right)^A \tag{2}$

or, if $T > 26.5$ then $ET_{ref} = -415.85 + 32.24T - 0.43T^2 \tag{3}$

and $A = 0.49239 + 0.01792H - 0.000077177H^2 + 0.000000675H^3 \tag{4}$

where $CropF$ = crop factor, $CropF_c$ = a calibration parameter, T = mean temperature (°C) and H = HEAT parameter of Thornthwaite [3], defined by:

$$H = \sum_{jan}^{dec}\left(\frac{T_m}{5}\right)^{1.514} \tag{5}$$

where T_m = long-term average monthly temperature (°C). The actual evapotranspiration is calculated based on [2]:

if $P_{eff} \geq 0$ then $AE = ET_{pot} \tag{6}$

or, if $P_{eff} < 0$ then $AE = P + MELT + \left(SS_{t-1} - WHC^{\left(-\frac{APWL}{WHC}\right)}\right) \tag{7}$

$$\text{where } P_{eff} = P + MELT - ET_{pot} \tag{8}$$

$$SS = SS_{t-1} + P_{eff} - SO \tag{9}$$

$$SO = SP + AE \tag{10}$$

$$\text{and, if } SS \geq WHC \text{ then } APWL = 0 \tag{11}$$

$$\text{or, if } SS \geq WHC \text{ then } APWL = WHC \cdot \ln\left(\frac{WHC}{SS_{t-1}}\right) - \left(P + MELT - ET_{pot}\right) \tag{12}$$

where P = rainfall (mm), P_{eff} = effective rainfall (mm), AE = actual evapotranspiration, $MELT$ = amount of snowmelt water (mm), SS_{t-1} = soil storage (SS) in the previous iteration (mm), WHC = soil water holding capacity (mm/m), and SP is soil seepage. The groundwater storage (GW) is calculated as follows:

$$GW_a = GW_{t-1} + SP \tag{13}$$

$$\text{where, if } P_{eff} \geq 0 \text{ then } SP = (1 - TOGW_c)(SS_{t-1} + P_{eff} - WHC), \tag{14}$$

$$\text{and, if } P_{eff} < 0 \text{ then } SP = 0 \tag{15}$$

$$GW = GW_a - Q_{base} \tag{16}$$

where GW_{t-1} = groundwater storage in the previous iteration, $TOGW_c$ = calibration factor which separates between direct runoff and seepage to groundwater, C = calibration parameter based on cell topography and Q_{base} = baseflow, defined as:

$$Q_{base} = \frac{GW_a}{C} \tag{17}$$

Overland flow (Q_{over}) is defined as:

$$\text{if } P_{eff} \geq 0 \text{ then } Q_{over} = TOGW_c \cdot (SS_{t-1} + P_{eff} - WHC) \tag{18}$$

$$\text{or if } P_{eff} < 0 \text{ then } Q_{over} = 0 \tag{19}$$

Total discharge per grid cell (Q) is defined as:

$$Q = Q_{base} + Q_{over} \tag{20}$$

3   Set Up of Morphologic Setting and Historical Climate Dataset

A DEM is used to set up the geographical domain of the model. The DEM used in this study is that of SRTM (Shuttle RADAR Topography Mission) from 2003 [4], which has a spatial resolution of 90m×90m. This dataset was resampled to a resolution of 1km×1km (see Figure 1). The DEM is used to derive slope, and the direction of flow between grid cells, based on the steepest decent. The land use map used in this study is assumed to be constant over the simulation period. The land use map is based on interpreted LANDSAT imagery from 2001 [5], resampled to a resolution of 1km×1km, and converted to crop factors ($CropF$). A $CropF$ map is used in STREAM to calculate potential evapotranspiration (ET_{pot}). The crop factor is a dimensionless factor by which the reference evapotranspiration (ET_{ref}) is multiplied in order to account for the difference in ET_{pot} over different land use types. The land use maps were reclassed to $CropF$ maps based on values in [6] and [7]. The land use maps are also used to generate maps of soil water holding capacity (WHC) by reclassing to standard values of WHC. In Figure 2, models of morphologic setting of the domain comprising of river network and water holding capacity respectively resulted from the DEM from 2003 and the land use from 2001 are shown. Table 1 summarizes characteristics of catchments considered in this study.
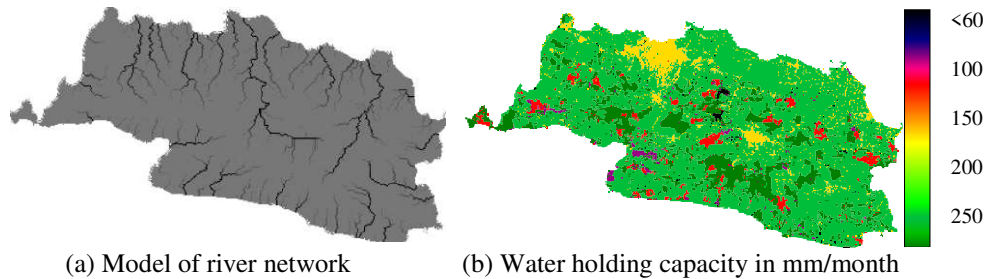

(a) Model of river network   (b) Water holding capacity in mm/month

Figure 2 Model of morphologic setting.

Table 1   Landscape characteristics of selected catchments.

Catchment Name	Area (km^2)	Average Elevation	Average Slope	% Urban	% Agricultural	% Forest
Citarum	7,046	605m	3°	12	72	16
Cisadane	1,551	383m	3°	17	65	18
Ciliwung	485	398m	2°	54	34	12

Global climate (precipitation and temperature) time-series datasets covering a 100-year record (i.e. 1901-2002) are made available from the Climatic Research Unit (CRU), University of East Anglia, United Kingdom [8]. In this study the CRU dataset is used, which provides a gridded set of monthly climate reanalysis

data for the entire globe at the 30'×30' resolution. Only the latest 20 years of available precipitation and temperature data (i.e. 1981-2002) are considered here. In addition to that, climatology data from the same source (i.e. CRU) are also used. These show mean monthly temperature and precipitation for the same period at a higher spatial resolution (10'×10'), but are not available as time series [9]. The low resolution (30'×30') climate time-series datasets are statistically downscaled to the higher 10'×10' resolution (see dashed lines in Figure 1). This involves two steps. Firstly, the low resolution data are simply resampled onto a 10'×10' grid. Then, for each grid cell, the time-series data are statistically downscaled, such that [10]:

$$T_t' = T_t + \left(\overline{T_c} - \overline{T_t} \right) \tag{21}$$

$$P_t' = P_t \left(\overline{P_c} \Big/ \overline{P_t} \right) \tag{22}$$

with T_t' = downscaled temperature time-series data, T_t = original temperature time-series data (resampled to 10'×10'), $\overline{T_c}$ = mean monthly temperature from 20-year high resolution climatology, $\overline{T_t}$ = mean monthly temperature calculated from 20-year low resolution time-series, P_t' = downscaled precipitation time-series data, P_t = original precipitation time-series data (resampled to 10'×10'), $\overline{P_c}$ = mean monthly precipitation from 20-year high resolution climatology, $\overline{P_t}$ = mean monthly precipitation calculated from 20-year low resolution time-series.

The agreement between the downscaled global precipitation time-series dataset during the calibration period and the observations carried out in Tanjung Priuk, Halim, Katulampa and Depok (see Figure 1) from 1989 to 2002 are evaluated. The correlation coefficient (r) between these datasets, as well as the total annual model precipitation expressed as a % of the total annual observed precipitation are given as:

$$r = \frac{\sum (P_t' - \overline{P_t'})(P_o - \overline{P_o})}{\sqrt{\sum (P_t' - \overline{P_t'})^2 \sum (P_o - \overline{P_o})^2}} \tag{23}$$

$$\frac{P_{tot}^t}{P_{tot}^o} \times 100\% \tag{24}$$

with $P_{tot}^t = \sum P_t'$, $P_{tot}^o = \sum P_o'$, P_o = monthly observed precipitation and $\overline{P_o}$ = mean observed precipitation. r and % hence indicate the agreement in trend and

magnitude, respectively. Figure 3 shows comparison between global datasets and local observations of precipitation.
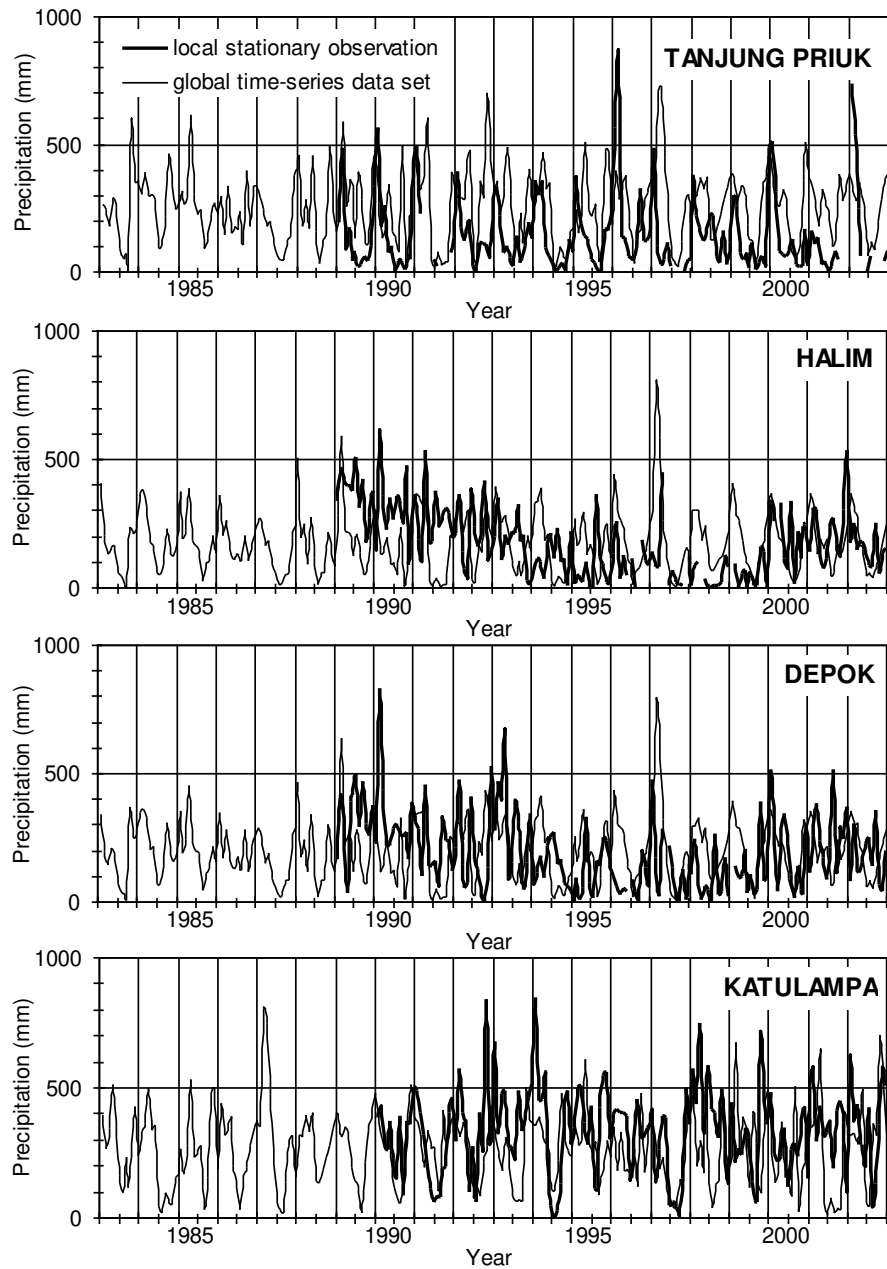

Figure 3 Global precipitation datasets and local station observations.

In Figure 4 comparisons between mean monthly values of the downscaled global dataset and local observation stations are shown. Summary of the agreement is given in Table 2, showing the *r* and %. Higher correlation is seen in Tanjung Priuk and Katulampa. In terms of percentage, the downscaled global datasets agree well with the local station observations at two stations (i.e. Halim and Depok), and reasonably well for the station at Katulampa.
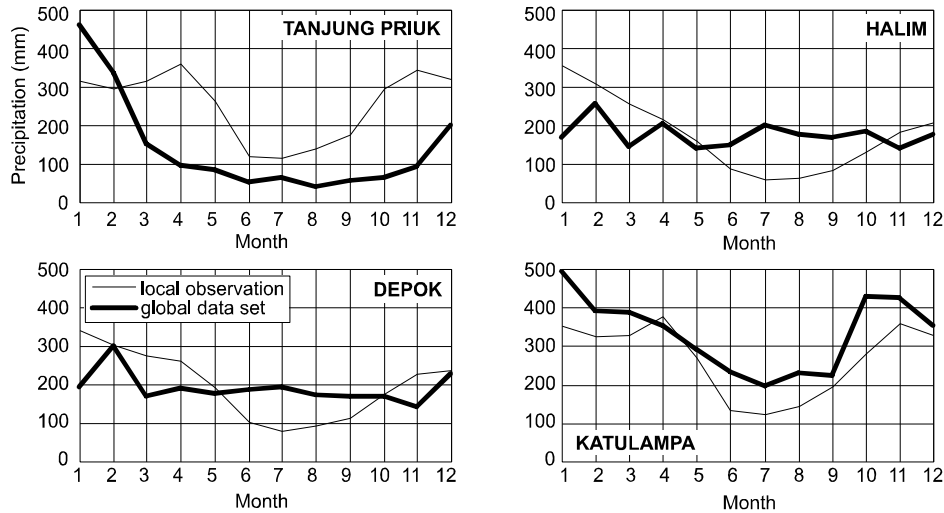

Figure 4 Comparison between monthly average precipitation from the global climate dataset and local observation stations.

Table 2   Agreement between monthly average global precipitation datasets and local station observations.

Station	*r*	*%*
Tanjung Priuk	0.54	184
Halim	0.27	101
Depok	0.38	106
Katulampa	0.87	80

4   Calibration of Discharge Computation

The STREAM model is calibrated in order to determine the optimum setting of model parameters for simulating monthly average discharge with acceptable agreement with observation data. Agreement between observed and computed discharge is assessed using the correlation coefficient (r), ratio between computed and observed total annual average values given in %, and Nash-Sutcliffe coefficient (E) [11] as:

$$E = 1 - \frac{\sum_{1}^{n}(Q'_i - Q_i)^2}{\sum_{1}^{n}(Q_i - \overline{Q})^2} \tag{25}$$

with Q'_i = computed monthly discharge, Q_i = observed monthly discharge, $\overline{Q}$ = average of observed discharge and n = number of data.

STREAM is a water balance model which simulates the water balance in a simplified manner, and hence only a small number of calibration parameters are required. This is advantageous over process-based distributed-parameter models, in which hydrology is simulated based on more thorough equations describing the physical processes involved in the hydrological cycle. However, as the physical reality of such models increases, so too does the number of parameters which must be estimated. Since empirical data on which to base such parameter estimations are usually scarce, and sometimes non-existent, over-parameterisation introduces many uncertainties into these models [12,13]. This high number of parameters may be necessary when the purpose of a model is to simulate the detailed processes of the hydrological cycle, such as in models for water quality management [14]. However, unnecessary over-parameterisation often leads to little improvement in the performance of models to simulate the water balance of a basin [12,15].

In calibrating the model, care is taken to select parameters which are physically meaningful. Calibration is carried out for six discharge gauging stations in three river basins (Table 3; see Figure 1) to reduce the problem of equifinality of parameter estimation. Prior to the calibration, sensitivity analysis of the calibration parameters is reviewed. This provides measures of the changes of computed discharge magnitude due to the changes of a particular calibration parameter. The parameters used for the calibration are: $CropF_c$, WHC, HEAT (used in the Thornthwaite-Mather approach for calculating potential evapotranspiration [3]), $TOGW_c$ multiplier (determines the proportion of surplus water per grid cell that runs off directly or that seeps to the groundwater) and C factor (determines the proportion of groundwater that contributes to baseflow, based on slope).

It is found that discharge calculation is sensitive to $CropF_c$ and fairly sensitive to WATERH and the C factor, as also reported by [17]. In our study, the model is also sensitive to the HEAT factor. Several sets of calibration parameters are used for input to run several simulations. The resulting simulated discharges are compared to the observed values. The optimum set of calibration parameters is

taken according to the fitness of computed discharges to the observed ones evaluated from the correlation coefficient (r), ratio between computed and observed total annual average values given in %, and Nash-Sutcliffe coefficient (E). The final calibration parameters chosen for those parameters governing the overall water balance required no or little calibration: for WATERH and HEAT we used the standard uncalibrated parameters (i.e. 1.0), whilst the value used for $CropF_c$ (i.e. 0.9) shows little change from the standard uncalibrated value of 1.0. Hence, the water balance was simulated well using the standard parameters based on the empirical equations by which the model is driven, so that very little calibration was required. The parameters $TOGW_c$ and C, which influence the peaks and troughs in discharge, are also very much standard values (see, for example, [17]). The resulting optimum set of calibration parameters for discharge computation is shown in Table 4.

Table 3   Discharge characteristics at the observation stations.

Catchment	Station	Duration	Length of record	q_{min} (m³/s)	q_{max} (m³/s)	q_{mean} (m³/s)	Q_{tot} (m³/s)	Missing data
Citarum	Majalaya	1988-2002	9 years	9.7	51.7	11.6	138.9	1989-1991
	Nanjung	1983-2002	18 years	2.2	210.6	74.6	894.7	1989, 1998
Cisadane	Katulampa	1990-2002	9 years	0.1	40.8	9.1	109.4	1991, 1999
	Batubeulah	1984-2002	18 years	7.7	254.4	96.7	1160.1	-
Ciliwung	Genteng	1991-2002	10 years	0.2	40.2	11.5	138.4	1993, 1999
	Ratujaya	1980-1998	9 years	0.4	31.4	11.3	135.7	1982-1990, 1996

Table 4   Setting of optimum calibration parameters.

Symbol	Parameter	Value
$CropF_c$	Crop factor	0.9
C	Based on slope	3.0
WATERH	Water holding capacity factor	1.0
H	HEAT factor	1.0
$TOGW_c$	Ratio of direct and delayed run off	0.5

In Figure 5, time series comparisons are shown between observed and computed discharge. It is seen that the spatial tool can simulate discharge events with low and high magnitudes. In Batubeulah, the computed values do not follow the increasing observed values in the last half of the simulation period resulting in poor agreement. We are challenged by limited knowledge of the increasing observed discharge in Batubeulah in the last half of the observation period and unable to elaborate how such an increase occurred. However, better agreement is seen in the other observation stations (i.e. Majalaya, Nanjung, Katulampa, Genteng and Ratujaya).
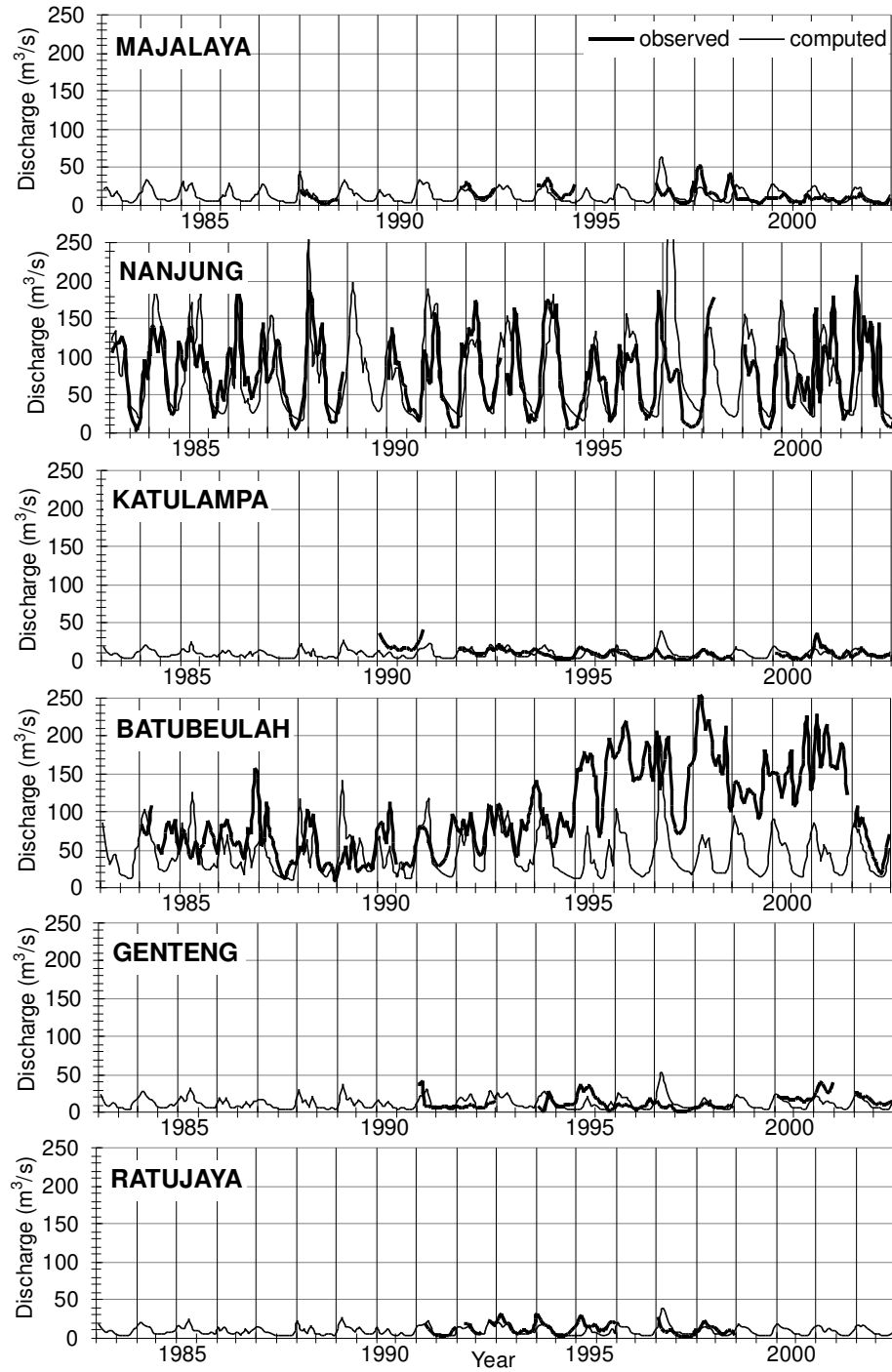
Figure 5 Comparison between computed and observed discharge time-series.

In Figure 6, hydrographs are shown comparing modelled and observed mean monthly discharge at various stations. The corresponding statistics are shown in Table 5. The computations agree quite well with the observations, with correlation coefficients (r) ranging from 0.72 to 0.93. Excluding Batubeulah, the computation shows accuracy of total mean annual discharge ranging between 92 and 105%. These correspond to deviations between modelled and observed mean annual discharges ranging from -8 to +5%. Among the other observation stations, good agreements are found in Nanjung, Katulampa and Genteng.
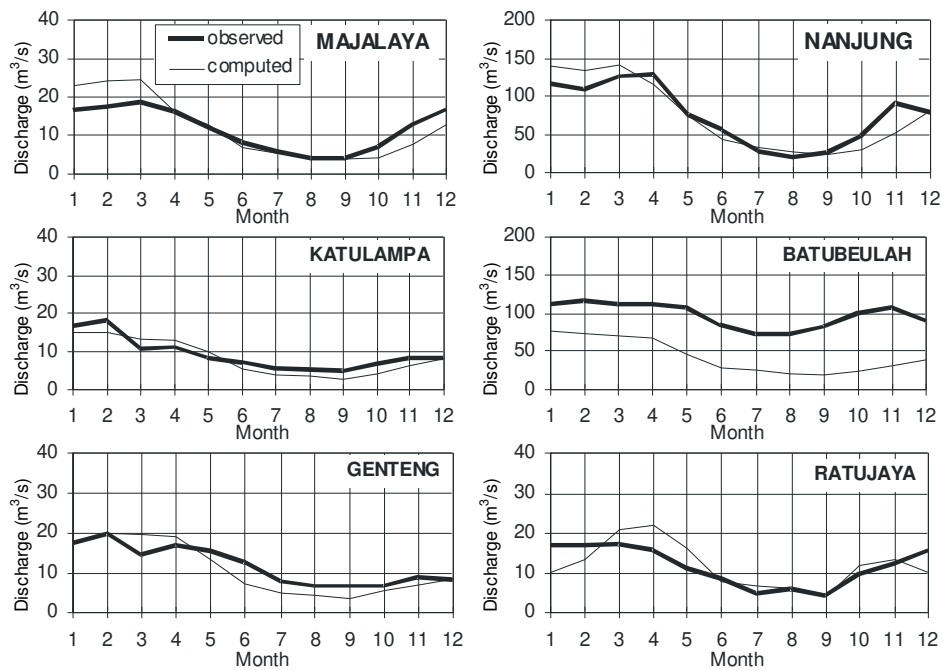
Figure 6 Comparison between computed and observed mean monthly discharge.

Table 5   Agreement between computed and observed discharges.

No	Station	Catchment	Length	r	%	E
1	Majalaya	Citarum	9 years	0.89	105	0.46
2	Nanjung	Citarum	18 years	0.93	100	0.80
3	Katulampa	Cisadane	9 years	0.91	92	0.76
4	Batubeulah	Cisadane	18 years	0.81	45	-11.41
5	Genteng	Ciliwung	10 years	0.93	94	0.66
6	Ratujaya	Ciliwung	9 years	0.72	105	0.35

5   Further Concerns

STREAM has been applied for simulating the water balance parameters of some of the world's major watersheds [1,16,17,18]. Here, STREAM shows its capability in simulating discharges across multiple micro watersheds. Bearing in mind that generalised land use maps are used, the land use data are held constant over the simulation period, and the climate data are of a fairly low spatial resolution, with minor exceptions the spatial tool studied here provides computed discharges with acceptable agreement with the observed values. Still, further elaboration on this is required for further study. Upon elaboration for improvement of the agreement between computed and observed discharges, the model could be used for simulating longer time-series of discharge in the past, and could be improved by using more representative (time-varying) land use data.

Acknowledgement

References

[1]   Aerts, J.C.J.H., Kriek, M. & Schepel, M., *STREAM, Spatial Tools for River Basins and Environment and Analysis of Management Options: 'Set Up and Requirements'*, Physics and Chemistry of the Earth Part B, **24**(6), 591-595, 1999.

[2]   Thornthwaite, C.W. & Mather, J.R., *Instructions and Tables for Computing Potential Evapotranspiration and The Water Balance*, Publications in Climatology, **10**(3), 183-243, 1957.

[3]   Thornthwaite, C.W., *An Approach Toward A Rational Classification of Climate*, Geographical Review, **38**, 55-94, 1948.

[4]   Rodriguez, E., Morris, C.S., Belz, J.E., Chapin, E.C., Martin, J.M., Daffer, W., & Hensley, S., *An Assessment of the SRTM Topographic Products*, Technical Report JPL D-31639, Jet Propulsion Laboratory, Pasadena, California, USA, 143 pp, 2005.

[5]   Poerbandono & Harto, A.B., *Reconstruction and Prediction of Physical Environmental Quality Degradation due to The Changes of Erosional Behaviour of Citarum Watershed Using Spatial Model*, Unpublished Report, Institute for Research and Community Services, Institute of Technology, Bandung, Indonesia, 2008.

[6]   Kwadijk, J.C.J., *The Impact of Climate Change on The Discharge Of The River Rhine*, Ph.D. Thesis, Universiteit Utrecht, Utrecht, The Netherlands, 201 pp, 1993.

[7]   Aerts, J.C.J.H. & Bouwer, L.M., *STREAM Krishna, A Hydrological Model for the Krishna River in India*, RIKZ/Coastal Zone Management Centre, The Hague, The Netherlands, 34 pp., www.geo.vu.nl/users/ivmstream/public/aerts&bouwer2002a.pdf, 2002.

[8]   Mitchell, T.D. & Jones, P.D., *An Improved Method of Constructing A Database of Monthly Climate Observations and Associated High-Resolution Grids*, International Journal of Climatology, **25**(6), 693-712, 2005.

[9]   New, M., Lister, D., Hulme, M. & Makin, I., *A High-Resolution Dataset of Surface Climate Over Global Land Areas,* Climate Research, **21**, 1-25, 2002.

[10] Bouwer, L.M, Aerts, J.C.J.H, van de Coterlet, G.M., van de Giesen, N, Gieske, A. & Mannaerts, C., *Evaluating Downscaling Methods for Preparing Global Circulation Model (GCM) Data for Hydrological Impact Modelling*, In Aerts, J. & Droogers, P. (Eds.). Climate change in contrasting river basins: adaptation strategies for water, food, and environment. CABI Publishing. Oxford, UK, 2004.

[11]   Huang, M. & Liang, X., *On The Assessment of The Impact of Reducing Parameters and Identification of Parameter Uncertainties for A Hydrologic Model with Applications to Ungauged Basins*, Journal of Hydrology, **320**, 37-61, 2006.

[12]   Jakeman, A.J. & Hornberger, G.M., *How much complexity is warranted in a rainfall-runoff model?*, Water Resources Research, **29**(8), 2637-2649, 1993.

[13]   Krajewski, W.F., Lakshmi, V., Georgakakos, K.P., & Jain, S.C., *A Monte Carlo Study of Rainfall Sampling Effect on A Distributed Catchment Model*, Water Resources Research, **27**(1), 119-128, 1991.

[14]   Zheng, Y., & Keller, A.A., *Understanding Parameter Sensitivity and Its Management Implications in Watershed-Scale Water Quality Modeling*, Water Resources Research, **42**, W05402, 2006.

[15]   Beven, K., *Changing Ideas in Hydrology-The Case of Physically-Based Models*, Journal of Hydrology, **105**(1-2), 157-172, 1989.

[16]   Ward, P.J., Renssen, H., Aerts, J.C.J.H., van Balen, R.T. & Vandenberghe, J., *Strong Increases in Flood Frequency and Discharge of The River Meuse Over The Late Holocene: Impacts of Long-Term Anthropogenic Land Use Change and Climate Variability*, Hydrology and Earth System Sciences, **12**, 159-175, 2008.

[17]   Ward, P.J., Aerts, J.C.J.H., de Moel, H. & Renssen, H., *Verification of A Coupled Climate-Hydrological Model Against Holocene Palaeohydrological Records*, Global and Planetary Change, **57**, 283-300, 2007.

[18]   Aerts, J.C.J.H., Renssen, H., Ward, P.J., De Moel, H., Odada, E., Bouwer, L.M. & Goosse, H., *Sensitivity of Global River Discharges Under Holocene and Future Climate Conditions*, Geophysical Research Letters, L19401, 2006.