Appl Water Sci (2013) 3:545–557
DOI 10.1007/s13201-013-0102-5

ORIGINAL ARTICLE

Applicability of global public domain data versus local detailed data for distributed hydrological modelling: a study form Gin river basin Sri Lanka

T. N. Wickramaarachchi · H. Ishidaira · T. M. N. Wijayaratna

Received: 1 December 2012 / Accepted: 2 April 2013 / Published online: 26 April 2013

Abstract This study investigates the applicability of global public domain data versus local detailed data for distributed hydrological modelling using a case study approach. Major hydrological characteristics in Gin river basin are simulated in the study by applying the distributed hydrological model, YHyM/BTOPMC (University of Yamanashi Distributed Hydrological Model with Block-wise use of TOPMODEL and Muskingum-Cunge method) utilizing the global public domain data sets (Case 1) and local detailed data sets (Case 2). Evaluation of the model outputs for Case 1 and Case 2 shows that the overall hydrological behavior of the Gin river basin is adequately simulated by the model for both Case 1 and Case 2. The simulated average annual discharge volumes in Case 1 and Case 2 at Agaliya during 2002–2006, vary from the observed average annual discharge volume by +4.25 and +1.31 %, respectively. In general, simulated daily discharge in Case 1 shows slightly higher value than that of Case 2 resulting a difference of 0.9 m^3/s during 2002–2006, on average. The relative differences between the simulated daily discharges in Case 1 and Case 2 become higher during the recession limbs of the flow hydrographs at Agaliya. Reasons for these variations are being discussed in the study. The results of the study give motivation towards the use of global public domain data for hydrologic simulations in data-poor (limited availability of local data) basins.

Keywords Distributed hydrological modelling · YHyM/BTOPMC · Gin river · Global public domain data · Local detailed data

T. N. Wickramaarachchi (✉) · T. M. N. Wijayaratna
Department of Civil and Environmental Engineering, Faculty of Engineering, University of Ruhuna, Hapugala, Galle, Sri Lanka
e-mail: thusharanw@yahoo.com; thushara@cee.ruh.ac.lk

H. Ishidaira
Interdisciplinary Graduate School of Medicine and Engineering, University of Yamanashi, 4-3-11, Takeda, Kofu, Yamanashi 400-8511, Japan

Introduction

Hydrologic simulation has become a powerful technique in sustainable land and water resources planning and management. Hydrological models can be classified according to the process description and spatial representation. According to the process description, models can be classified into three categories: metric, conceptual and physically based models. According to the spatial representation, the hydrological models are either lumped or distributed. Choice of a suitable model structure relies heavily on the function that the model needs to serve. Distributed models in hydrology are usually physically based, in that they are defined in terms of theoretically acceptable continuum equations. Distributed hydrological models can fulfill the necessities of describing basin heterogeneity, and assess the impact of natural and human-induced changes (Virtual Academy 2010). However, distributed hydrological model applications are partly limited due to the requirement of large amount of spatial data which are not always available and difficulties in obtaining such data. To overcome this problem, it is appropriate to make use of global public domain datasets available on the internet, the quality of which is rapidly increasing.

The objective of this study was to investigate the applicability of global public domain data versus local detailed data for distributed hydrological modelling using a

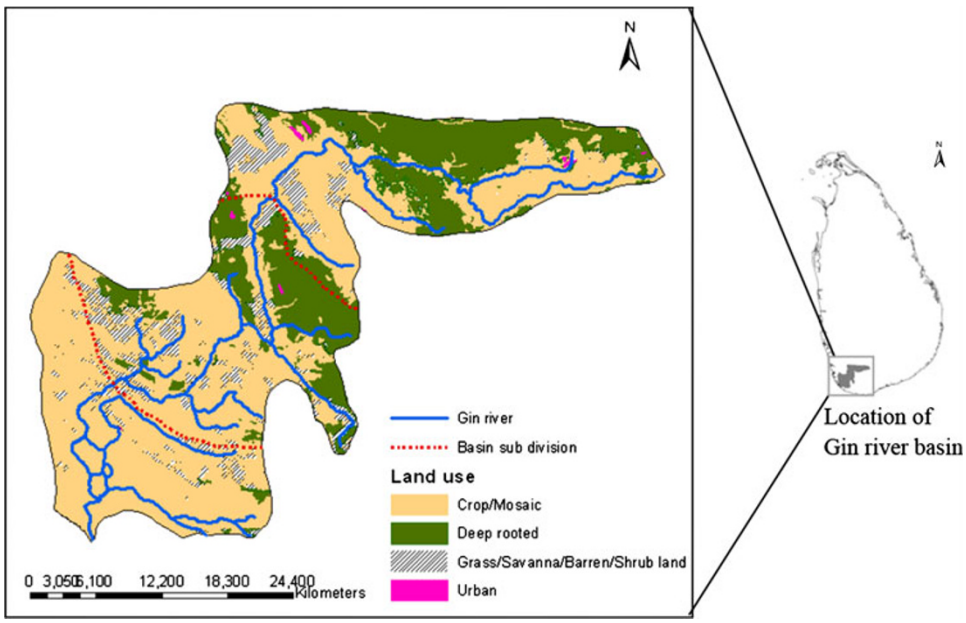

Fig. 1 Gin river, its catchment, and location, basin subdivision, and key land use types

case study. To reach this objective, the distributed hydrological model, YHyM/BTOPMC is applied to Gin river basin using the global public domain data sets (Case 1) and local detailed data sets (Case 2), and the simulation results for the two cases are evaluated. YHyM/BTOPMC was developed to cover most of the requirements for modelling hydrological responses of a basin, and has already been successfully applied to many catchments around the world (Takeuchi et al. 2008). Most of the parameters to be identified in the YHyM/BTOPMC are related to physical basin features of land cover and soil.

Study area and background

Gin river is one of the main sources of water supply in southern region of Sri Lanka. It is located roughly between longitudes 80°08″E and 80°40″E, and latitudes 6°04″N and 6°30″N. Gin river having a catchment area of about 932 km^2 includes Galle (83 % of the basin area), Matara (9 % of the basin area), Rathnapura (7 % of the basin area), and Kalutara (1 % of the basin area) administrative districts. Gin river originates from the Gongala mountains in Deniyaya having an elevation of over 1,300 m and flows to the Indian Ocean in Ginthota area of Galle District. Rainfall pattern in the catchment is of bi-modal, falling between May and September (south-west monsoon, which is the major rainfall event), and again between November and February (north-east monsoon) followed by the inter-monsoon rains during the remaining months of the year. Rainfall varies with altitude with mean annual rainfall above 3,500 mm in the upper reaches to <2,500 mm in the lower reaches of the catchment. River Gin annually discharges about 1,268 million cubic meters to sea (National Atlas 2007). Average temperature in the catchment varies from 24 to 32 °C with high-humidity levels, and the dominant soil texture is sandy clay loam. Gin river basin is rather a natural catchment in Sri Lanka which entirely lies within the wet zone of the country and having a natural rain forest covering considerable area in its upper catchment. Figure 1 shows Gin river, its catchment and location, basin subdivision, and key land use types.

Since most of the low-lying areas in the Galle District frequently subjected to flooding during the rainy seasons, problem of Gin river flooding is considered as leading environmental hazard of the district. Due to rapidly growing population and development activities in some parts of the catchment, competition for water is likely to increase. Galle is the capital city in southern Sri Lanka and the city's main pipe-borne water supply system depends on the water resources in Gin river basin. Hence, it is vital to comprehend the hydrology of the river basin in order to gain knowledge on current and future hydrological conditions.

Model description

The University of Yamanashi Distributed Hydrological Model (YHyM) with Block-wise use of TOPMODEL and Muskingum-Cunge method (BTOPMC) is a grid-based distributed hydrological model developed at the University of Yamanashi, Japan. In the YHyM, runoff is generated based on the TOPMODEL concept (Beven and Kirkby 1979) and flow routing is carried out using the Muskingum Cunge method (Cunge 1969; Ao et al. 2003a, b). The

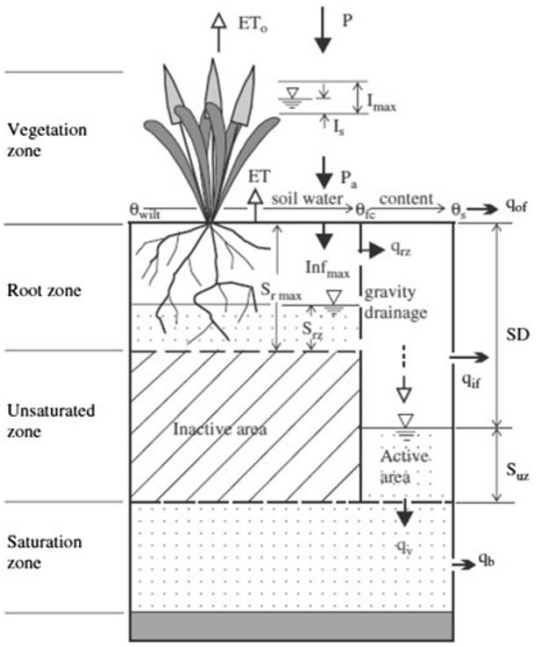

Fig. 2 Runoff generation in a grid cell in the BTOP model (the *vertical* profile). P is the gross rainfall, ET_0 is the interception evaporation, I_{max} is the interception storage capacity, I_s is the interception state, Inf_{max} is the infiltration capacity, P_a is the net rainfall on the land surface, ET is the actual evapotranspiration, $S_{r\,max}$ is the storage capacity of the root zone, S_{rz} is the soil moisture state in root zone, SD is the soil moisture deficit in unsaturated zone, S_{uz} is the soil moisture state in unsaturated zone, q_{of} is the overland runoff, q_{if} is the saturation excess runoff, q_v is the groundwater recharge, and q_b is groundwater release. θ_{wilt}, θ_{fc}, and θ_s are soil water content at wilting point, field capacity, and saturation, respectively

hydrological processes in a grid cell in the BTOP model are illustrated in Fig. 2 (Takeuchi et al. 2008).

The watershed is described by drainage networks extracted from digital elevation models (DEMs), in which all pits are filled with calculated small elevation increments (Ao 2001, Ao et al. 2003b). The topographic index γ_i for the i-th grid cell is calculated using the equation (1).

$$\gamma_i = \ln(\alpha_i / \tan\beta_i) \tag{1}$$

where α_i is the drainage area per unit length of contour, tan β_i denotes the slope of grid cell, i.

The runoff from a grid cell to the local schematic stream reach is the sum of saturation excess overland flow (q_{of}) and groundwater discharge (q_b) per unit length of contour line:

$$q_{of}(i,t) = \{S_{uz}(i,t) - \mathrm{SD}(i,t)\} \tag{2}$$

where S_{uz} is the unsaturated zone storage and SD is the saturation deficit for the i-th grid cell at time t.

$$q_b(i,t) = T_0(i)\exp\left(\frac{-\mathrm{SD}(i,t)}{m(k)}\right)\tan\beta_i \tag{3}$$

where SD indicates the saturation deficit, T_0 is the transmissivity, and $m(k)$ is the discharge decay factor in subbasin k.

In flow routing calculation using the Muskingum Cunge method, the river cross-section is assumed to be rectangular and river width B (meters) is approximated by

$$B(i) = C\sqrt{A(i)} \tag{4}$$

where constant $C = 10$ and A is the drainage area in square kilometers (Lu et al. 1989).

The equivalent Manning's roughness coefficient of a grid cell is estimated as

$$n_i = n_0(k)[\tan\beta_i / \tan\beta_0(k)]^{1/3} \tag{5}$$

where n_0 and tan β_0 are the equivalent roughness coefficient and slope at the outlet of sub-catchment k, respectively and n_0 is a model parameter to be calibrated.

The generated overland flow and groundwater flow of each cell will be added to the stream and then routed to the basin outlet. The maximum root zone storage is calculated considering the distribution of land cover and rooting depth. The spatial variation of soil transmissivity (T_0) over the catchment is considered based on the percentage of sand, silt, and clay present in each soil type. In the BTOPMC, spatially distributed monthly average potential evapotranspiration is calculated using the Shuttleworth–Wallace (S–W) method (Shuttleworth and Wallace 1985).

This article does not include a detailed description of the model and further references can be proposed as Ishidaira et al. (2000) and Takeuchi et al. (1999, 2008).

Hydrologic simulation

Data sets

In the case of YHyM/BTOPMC run with the global public domain data sets (Case 1), soil map and land cover data acquired from global public domain were input to the model, whereas in the case of YHyM/BTOPMC run with local detailed data sets (Case 2), local detailed soil map and land cover data were input to the model. In both Case 1 and Case 2, DEM data and data for S–W Simulation acquired from global public domain were used together with locally available rainfall and discharge data. Local detailed soil map and land cover data in Case 2 based on the high-resolution local maps presenting detailed classifications are more accurate in comparison to Case 1. Table 1 shows the basic data input to run YHyM/BTOPMC for Case 1 and Case 2, and the sources of data. Global public domain data shown in Table 1 will be available for a long time and it is indicated in Table 1, when these data became available.

Table 1 Basic data input to run YHyM/BTOPMC for Case 1 and Case 2, and the sources of data

Data set	Source
Digital elevation map (DEM) acquired from global public domain[a, b]	Shuttle Radar Topography Mission (SRTM) (Available since 2004)
Soil map acquired from global public domain[a]	Harmonized world soil database (HWSD) V 1.1 by Food and Agricultural Organization (FAO) (Available since 2008 with regular updates)
Local detailed soil map[b]	Department of survey, Sri Lanka
Land cover map acquired from global public domain[a]	United States geological survey—International Geosphere Biosphere Programme (USGS—IGBP) Global land cover characteristics data base version 2.0
Local detailed land cover map[b]	(Available since 1997 with regular updates) Department of Survey, Sri Lanka
Data for Shuttleworth and Wallace (S–W) simulation acquired from global public domain	
• Normalized Difference Vegetation Index (NDVI)[a,b]	Advanced very high resolution radiometer-global inventory modelling and mapping studies (AVHRR—GIMMS) (Available since 2004)
• Mean daily temperature[a,b] • Diurnal temperature range[a,b] • Vapor pressure[a,b] • Cloud cover[a,b] • Wind speed[a,b]	Intergovernmental Panel on Climate—Climate research unit (IPCC-CRU) 2.0 (Available since 1999)
Locally available daily discharge data (From 1997 to 2006) at Tawalama and Agaliya gauging stations[a,b]	Department of irrigation, Sri Lanka
Locally available daily rainfall data (From 1997 to 2006) at Anninkanda, Natagala, Pallegama, Baddegama, Labuduwa and Galle gauging stations[a,b]	Department of meteorology, Sri Lanka

[a] Model input data for Case 1

[b] Model input data for Case 2

DEM data

DEM data, which had spatial resolution of 3″ × 3″ (Jarvis et al. 2008), were extracted from the SRTM data set available in http://srtm.csi.cgiar.org/SELECTION/inputCoord.asp. It was scaled up to 30″ × 30″ when input to the YHyM/BTOPMC due to the computational limitations. While inputting to the model, these scaled-up DEM data were further compared with the DEM generated using the locally available contour data of 1: 50,000 scale. Input of the scaled-up DEM data (from SRTM data set) to the model performed better in generating the stream network than the locally available data and hence used as the topography data in this study. Figure 3 shows the stream network generated by the model using SRTM data set.

Discharge data

Daily discharge data were obtained from the Department of Irrigation, Sri Lanka for the two discharge gauging stations located in Gin river basin (Fig. 3). Agaliya station was established in lower reaches while Tawalama station was located in upper reaches of the river basin. The basin delineation was done by the YHyM/BTOPMC based on the locations of the discharge gauging stations. The whole

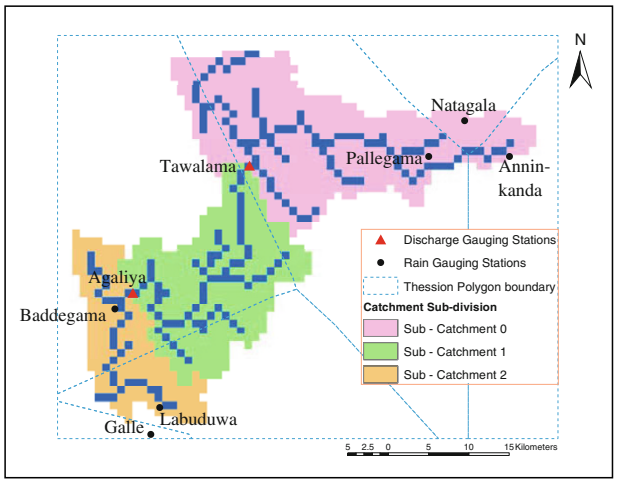

Fig. 3 Locations of discharge and rain gauging stations, catchment sub-divisions, stream network, and Thiessen polygons generated by YHyM/BTOPMC in the Gin river basin

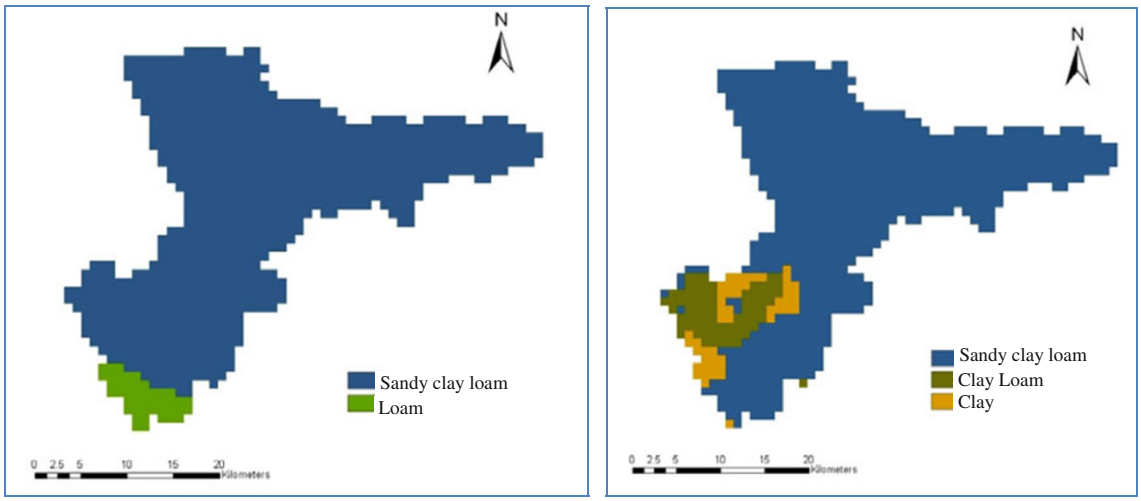

Fig. 4 **a** Soil map acquired from global public domain (Case 1) and (**b**) Local detailed soil map (Case 2)

catchment was divided to three sub-catchments based on that (Fig. 3). This study considered only the sub-catchment 0 and the sub-catchment 1, which includes the upstream zone above Agaliya. The total land area of the upstream zone above Agaliya and Tawalama stations are 780 and 470 km^2, respectively.

Rainfall data

Six rain gauging stations were selected within the river basin and nearby locations considering the availability of long-term data. Daily rainfall data were obtained from the Department of Meteorology, Sri Lanka. Rainfall data were checked for missing data and it was noted that missing of few data, which could not be recorded during certain days. Since there existed a significant variation in the normal rainfall data of the above stations, according to Das (2009), the normal ratio method was applied to estimate the missing data. Spatial distribution of the rainfall data was done by the YHyM/BTOPMC using the Thiessen polygon method (Fig. 3).

Soil data

Soil map acquired from global public domain is shown in Fig. 4a. The map was extracted from HWSD V1.1 produced by FAO which was having a resolution of 30″ × 30″ (FAO/IIASA/ISRIC/ISSCAS/JRC 2009). Local detailed soil map (Fig. 4b) was created by digitizing the Soils of Ceylon map which was having a scale of 1:1,000,000 produced by the Department of Survey, Sri Lanka. Particle size distribution of local soils was identified according to Moormann and Panabokke (1961) and Panabokke (1996). Variation of percentage area covered

Table 2 Variation of percentage area covered by different soil textures between Case 1 and Case 2

	Case 1 (% area)		
	Sandy Clay Loam	Loam	Total
Case 2 (% area)			
Sandy clay loam	83.32	3.63	86.94
Clay loam	7.88	0	7.88
Clay	4.25	0.93	5.18
Total	95.44	4.56	

Table 3 Soil properties and distribution of soil textures

	Soil ID/soil type	Texture	θ^a_{fc}	θ^b_{wilt}	(% area)	
					Sub-catchment 0	Sub-catchment 0 and sub-catchment 1
Soil map acquired from global public domain (Case 1)	FAO ID_3641 FAO ID_3645 FAO ID_3778	Sandy Clay Loam	0.255	0.068	100	95.44
	FAO ID_3654	Loam	0.270	0.027	0	4.56
Local detailed soil map (Case 2)	Red Yellow Podzolic soils	Sandy clay loam	0.255	0.068	100	86.94
	Alluvial soils	Clay loam	0.318	0.075	0	7.88
	Bog and Half-bog soils	Clay	0.396	0.090	0	5.18

[a] Field capacity

[b] Soil moisture content at wilting point

Table 4 Variation of land cover types between Case 1 and Case 2

		Case 1 (% area)			
		C/M[a]	G/S/B/S[b]	Deep rooted	Total
Case 2 (% area)	C/M[a]	42.59	3.32	1.55	47.46
	G/S/B/S[b]	12.75	0.31	0	13.06
	Deep rooted	36.48	2.49	0.41	39.38
	Urban	0.10	0	0	0.10
	Total	91.92	6.11	1.97	

[a] Crop/Mosaic

[b] Grass/Savanna/Barren/Shrub

by different soil textures between Case 1 and Case 2 is shown in Table 2.

For the soil types shown in both Fig. 4a and b, related soil properties and soil textures (Table 3) were identified in accordance with USDA soil triangle (Rawls et al. 1982, Rawls and Brakensiek 1985).

Land cover data

Land cover map acquired from global public domain included 30″ × 30″ resolution IGBP V 2.0 (From April 1992 to March 1993) developed by USGS (2008). Based on the hydrological view point, the original eight IGBP land cover classes were reclassified to four classes (Fig. 5a; Table 5). 1:50,000 scale, digital land cover map produced by the Department of Survey, Sri Lanka was used as the local detailed land cover map. Ten land cover types available in the area were categorized to the re-classified four IGBP classes (Fig. 5b; Table 5), so as to make the comparison easy. Variation of land cover types between Case 1 and Case 2 is shown in Table 4.

In the upper most and middle catchment areas of the Gin river, land cover types in Case 1 do not match properly with the land cover types in Case 2, while they match fairly well in the downstream area and in some parts of the upstream area (Fig. 5a, b). Land cover data in Case 2 based on the local detailed map are remarkably smooth and hence accurate in comparison to Case 1. Thus, the above mismatch could be attributed to the relatively low spatial resolution of Case 1 data compared to Case 2 data resulting majority of forest covers including forest plantation areas in the Case 2 to be classified as crop cultivations in the Case1.

Distribution of different land cover types and their root depths is shown in Table 5. Root depths for land cover data acquired from global public domain were based on Sellers et al. (1994, 1996), while root depths for local detailed land cover data determined in accordance with Allen et al. (1998) and local knowledge.

Data for Shuttleworth and Wallace (S–W) simulation

0.50 × 0.50 resolution, mean monthly climatology for 1961–1990 (30 year mean of the IPCC-CRU data) was used for the S–W potential evapotranspiration simulation in the YHyM/BTOPMC. These data acquired from global public domain included mean daily temperature, diurnal temperature range, vapor pressure, cloud cover, and wind speed. Also extraterrestrial radiation and daylight duration were derived from the above data by using the YHyM/BTOPMC Preprocessors. NDVI data used for S–W potential evapotranspiration module included monthly data from 1981 to 2006 with 4′ × 4′ resolution (Tucker et al. 2005).

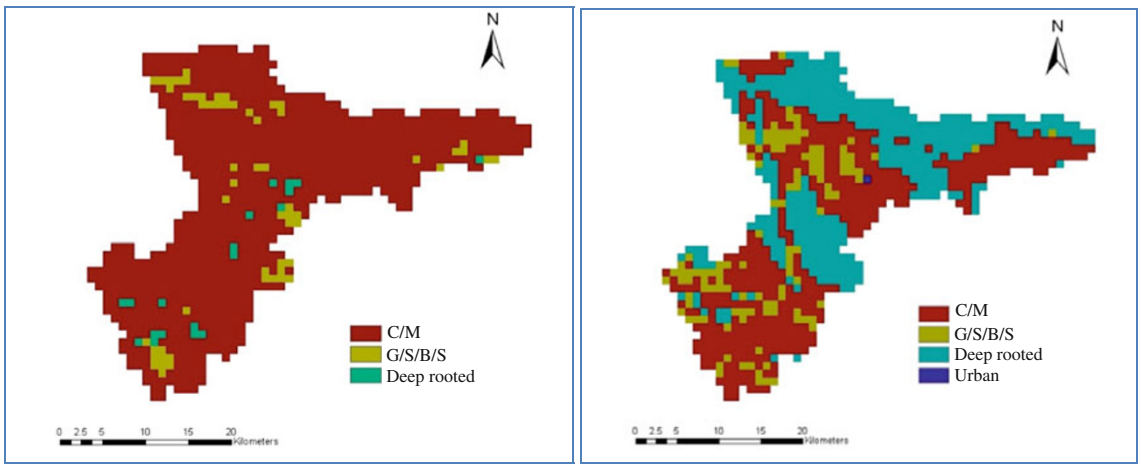

Fig. 5 **a** Land cover map acquired from global public domain (Case 1) and (**b**) Local detailed land cover map (Case 2)

Table 5 Distribution of different land cover types according to re-classified IGBP classes and root depths

	Re-classified IGBP classes	Land cover description	Root depth (m)	(% area)	
				Sub-catchment 0	Sub-catchment 0 and sub-catchment 1
Land cover data set acquired from global public domain (Case 1)	Deep rooted forest/ woodland	Evergreen broadleaf forest	2.5	1.21	1.97
		Deciduous broadleaf forest	2.5		
		Mixed forest	2.0		
	G/S/B/S[a]	Closed shrublands	1.0	6.57	6.11
		Woody savannas	1.0		
	C/M[b]	Croplands	0.7	92.21	91.92
		Cropland/natural vegetation mosaic	1.0		
	Urban	Urban and Built-up	0.001	0	0
Local detailed land cover data set (Case 2)	Deep rooted Forest/ Woodland	Forest-Unclassified	2.5	46.54	39.38
	G/S/B/S[a]	Scrub land	1.0	9.86	13.06
	C/M[b]	Rubber	2.0	43.43	47.46
		Coconut	1.5		
		Paddy	0.7		
		Tea	1.0		
		Homesteads/Garden	1.0		
		Other Cultivation	1.0		
		Chena	1.0		
	Urban	Rock	0.001	0.17	0.1

[a] Grass/savanna/barren/shrub

[b] Crop/mosaic

Model application

Most of the parameters to be identified in the YHyM/BTOPMC are related to physical basin features of land cover and soil. Parameters have been determined manually using the model parameter identification sub-module. Decay factor of transmissivity (m), Block average roughness coefficient (n_0), and Saturated transmissivity (T_0) were identified through trial-and-error calibration while the maximum storage capacities at root zone were presumed based on literature values.

The effects of the actual soil properties of each grid cell are included by the T_0 value which is assigned to each grid cell based on the following equation:

$$T_0 = U_{clay} \times T_{0\_clay} + U_{sand} \times T_{0\_sand} + U_{silt} \times T_{0\_silt} \quad (6)$$

where U_{clay}, U_{sand}, and U_{silt} are the percentages of clay, sand, and silt present in each grid. It is assumed that the soil texture inside each grid cell is homogeneous (Hapuarachchi et al. 2004a).

The maximum storage capacity of the root zone ($S_{\text{rz max}}$) is assigned to each grid cell based on the soil properties and the root depths according to the land cover maps (Hapuarachchi et al. 2004b).

$$S_{\text{rz max}} = (\theta_{\text{fc}} - \theta_{\text{wilt}}) \times \text{Root depth} \quad (7)$$

where θ_{fc} (m/m) is the field capacity and θ_{wilt} (m/m) is the moisture content at wilting point of the top soil layer in each grid. Soil properties in the Table 3 and root depths in the Table 5 were used for the $S_{\text{rz max}}$ calculations.

The calibrated parameter set was selected in such a way that it is to be valid for both Case 1 and Case 2, with observed discharge data at both discharge gauging stations. Model performance was further improved by fine-tuning the parameter values. The calibrated parameter set used by the YHyM/BTOPMC is shown in Table 6.

Daily discharge data from 1997 to 2001 and from 2002 to 2006 were used for calibrating and validating the model, respectively. Model performance was evaluated by the Nash–Sutcliffe Efficiency (E) and the volume ratio of total simulated discharge to total observed discharge (V_{r}).

$$E = 1 - \frac{\sum_{i=1}^{n} \left(Q_{\text{obs}_i} - Q_{\text{sim}_i}\right)^2}{\sum_{i=1}^{n} \left(Q_{\text{obs}_i} - \overline{Q_{\text{obs}}}\right)^2} \quad (8)$$

where Q_{obs_i} is the observed discharge, Q_{sim_i} is the simulated discharge, Q_{obs} is the average observed discharge, and n is the number of time step.

Table 6 Calibrated parameter set

Parameter	Value
Decay factor of transmissivity (m)	0.067 m
Block average roughness coefficient (n_0)	0.4
Saturated transmissivity (T_0)	T_0_sand: 12 m^2/h T_0_silt: 5 m^2/h T_0_clay: 1 m^2/h

$$V_r = \frac{\sum_{i=1}^{n} Q_{sim}}{\sum_{i=1}^{n} Q_{obs}} \quad (9)$$

where Q_{sim} is simulated runoff volume and Q_{obs} is observed runoff volume.

Results and discussion

Model performance

Table 7 shows the YHyM/BTOPMC performance during the calibration and validation. Nash–Sutcliffe efficiencies in both calibration and validation at Agaliya (located downstream with 780 km^2 drainage area) are better than Tawalama (located upstream with 470 km^2 drainage area) for both Case 1 and Case 2. Nash–Sutcliffe efficiency increases with the effective drainage area and this is further supported by Nawarathna et al. (2001).

Simulation of major hydrological characteristics

The YHyM/BTOPMC validation results for Case 1 and Case 2 at Agaliya during 2002–2006 are shown in Fig. 6 for which the Nash–Sutcliffe efficiency is 62.73 and 60.73 %, respectively (Table 7).

According to Fig. 6a, the hydrographs show a good agreement between the observed and simulated discharges in both Case 1 and Case 2 except for few extreme events. In particular, the low flows are simulated very well. Except for few years during which the simulated peaks are similar to the observed ones, most of the peak flows are underestimated. The reliability of the simulation results depends also upon the availability and quality of the input data. The deviation of simulated discharge from the observed is believed to be due to the overall error in the basin rainfall estimation and uncertainty associated with the river discharge observations during peak flows. Basin rainfall varies with altitude with mean annual rainfall above 3,500 mm in the upper reaches, 2,500–3,500 mm in the middle reaches, and <2,500 mm in the lower reaches of the basin. Most of the rain gauging stations lie in the upper and the lower reaches of the basin. Rainfall estimations in the middle reaches are based on the approximations using the nearest stations. This leads to form large Thiessen polygons in the middle reaches of the basin (Fig. 3) resulting less accurate rainfall estimations. The main storm runoff generation process considered by the YHyM/BTOPMC is surface runoff due to saturation excess overland flow. Hence, the peak flow estimations during the monsoon seasons with high rainfall intensities might have affected due to the lack of ability of the model to incorporate the infiltration excess runoff mechanism.

Considering Fig. 6a and c, it can be shown that how the model output simulates variation of the soil moisture condition of the basin with respect to the rainfall. Just after a dry season, with the start of rainfall (shown by rising limbs of the hydrographs in Fig. 6a), SD begins to decrease from its maximum. With the continuation of rainfall, S_{rz} is increasing and S_{uz} also begins to increase. During a peak flow event, both S_{rz} and S_{uz} achieve their maximum values. At the end of the peak flow event, with the decrease of rainfall, S_{uz} decreases with a gradual increase of SD. When S_{uz} achieves its minimum, S_{rz} begins to decrease and SD further increases. ET is considered to occur from the root zone according to the EP and the availability of water in the root zone. The variation of ET (Fig. 6b) follows the pattern of S_{rz} change (Fig. 6c) in accord with the fact that the evaporation takes place basically from the root zone. When there is enough rainfall, ET reaches nearly to its potential value which is EP.

Above suggests the applicability of YHyM/BTOPMC to Gin river basin in simulating the major hydrological characteristics, utilizing the global public domain data sets as well the local detailed data sets. Although the global public domain data sets are easily available through electronic data archives on the internet, these data sets are

Table 7 Model performance

	Case 1				Case 2			
	Calibration		Validation		Calibration		Validation	
	Agaliya	Tawalama	Agaliya	Tawalama	Agaliya	Tawalama	Agaliya	Tawalama
Nash–Sutcliffe efficiency (E) %	67.63	53.75	62.73	48.31	66.43	54.03	60.73	48.18
Ratio of total simulated discharge to total observed discharge (V_r) %	93.15	105.50	84.94	104.24	90.63	102.77	82.31	101.31

hardly used for integrated water resources management in river basins since most of the water resource managers are not aware of the existence of such data. The emerging trends in geographical information systems and their applications coupled with hydrological modelling should be oriented towards raising awareness and thus maximizing usage of global public domain data. Key limitations of using public domain data could be summarized as coarse spatial and temporal resolution and missing data during cloudy conditions. Accuracy of future data sets can be improved by incorporating more ground-truthing as well as better interpolation techniques during cloudy conditions.

Evaluation of the simulation results

Table 8 summarizes the YHyM/BTOPMC simulated water balance for Case 1 and Case 2 at Agaliya during 2002–2006. During 2002–2006, the simulated average

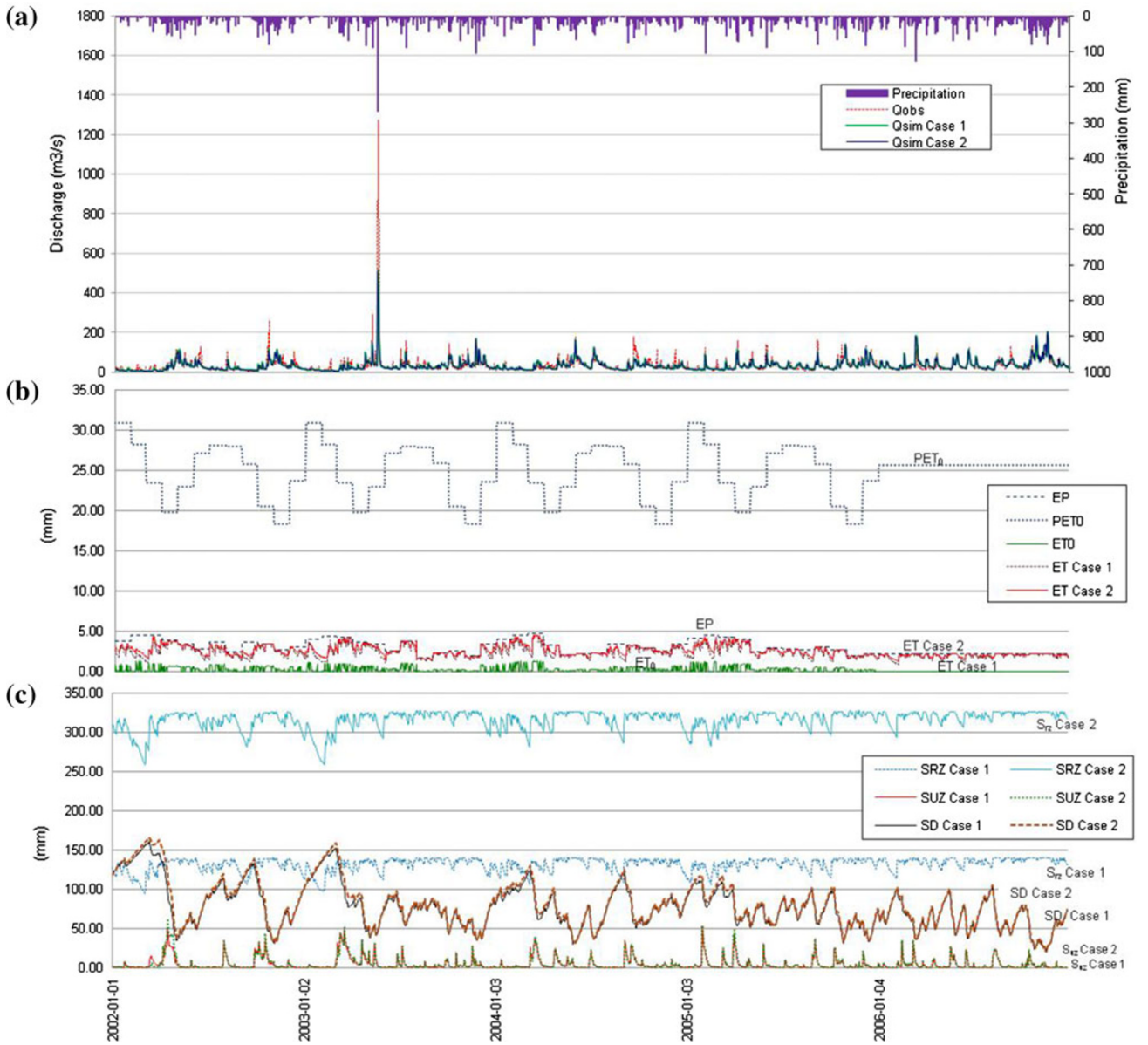

Fig. 6 Model validation results for both Case 1 and Case 2 at Agaliya discharge gauging station during 2002–2006. **a** Observed and simulated discharge hydrographs (*Qobs* is the observed discharge and Q_{sim} is the simulated discharge), (**b**) simulated basin average evapotranspiration (PET_O is the potential evaporation from interception, EP is the potential evapotranspiration from root zone, ET_O is the actual evaporation from interception, and ET is the actual evapotranspiration from root zone). (**c**) variation of simulated soil moisture (SD is the average saturation deficit, S_{rz} is the average storage in the root zone, and S_{uz} is the average storage in unsaturated zone)

Table 8 Summary of the hydrological characteristics simulated by YHyM/BTOPMC in Case 1 and Case 2 at Agaliya during 2002–2006

	Observed precipitation (mm)	Observed discharge volume × 10^6 (m^3)	Simulated evapotranspiration (ET) (mm)		Simulated interception evaporation (ET_o) (mm)	Simulated Ground water recharge (q_v) (mm)		Simulated discharge volume X 10^6 (m^3)		Simulated discharge (mm)[a]	
			Case 1	Case 2	Case 1 and Case 2	Case 1	Case 2	Case 1	Case 2	Case 1	Case 2
2002	2,868.25	807.58	919.65	1,009.66	106.76	1,082.74	1,036.88	772.71	738.00	990.65 (0.34)	946.15 (0.32)
2003	3,395.87	1,058.50	878.39	948.69	93.87	1,291.47	1,256.60	1,012.61	981.85	1,298.21 (0.38)	1,258.78 (0.37)
2004	3,211.67	975.12	935.42	1,013.42	105.45	1,287.52	1,253.43	932.77	900.90	1,195.85 (0.37)	1,155.00 (0.36)
2005	3,336.48	968.78	897.82	967.89	100.13	1,359.97	1,327.76	980.02	949.46	1,256.44 (0.38)	1,217.26 (0.36)
2006	3,885.14	1,017.44	692.67	726.81	0.09	1,659.28	1,654.30	1,334.32	1,320.64	1,710.66 (0.44)	1,693.12 (0.43)
Average (2002–2006)	3,339.48	965.48	864.79	933.29	81.26	1,336.20	1,305.80	1,006.48	978.17	1,290.36 (0.38)	1,254.06 (0.37)

[a] Simulated discharge/observed precipitation is shown within brackets

annual discharge volumes in Case 1 and Case 2 vary from the observed average annual discharge volume by +4.25 and +1.31 %, respectively. The difference between the simulated average annual discharges in the two cases is 36.30 mm. The simulated average annual actual evapotranspiration in Case 1 and Case 2 differs by 68.5 mm. The difference between the simulated average annual ground water recharge in the two cases is 30.4 mm. It is evident that the water balance components in Case 1 and Case 2 do not differ much from each other, but most of the time, simulated discharge in Case 1 is slightly greater than that of Case 2.

The primary differences between the Case 1 and the Case 2 are associated with the variations in soil type composition and land cover. Variations in vegetation composition result in change in root depth which leads to a change in amount of water restraining around the plant roots. Crop/mosaic is the main land cover type, covering 91.92 % of the area in sub-catchment 0 and sub-catchment 1 in Case 1 (Table 5). In Case 2, most parts of the headwaters in sub-catchment 0 and sub-catchment 1 (39.38 % of the area) are covered by deep rooted vegetation which are natural forests, while 47.46 % of the area is covered by crop/mosaic (Table 5). According to Table 5, the land cover in Case 2 includes greater area with deep rooted vegetation, and hence, more water restraining around the plant roots than Case 1. Further the simulation results show that, with more deep rooted area, S_{rz} Case 2 is always having a greater value than S_{rz} Case 1 (Fig. 6c), and hence higher value for ET Case 2 than ET Case 1 (Fig. 6b; Table 8). In Case 2, sub-catchment 0 and sub-catchment 1 are covered by sandy clay loam, clay, and clay loam whereas in Case 1, both sub-catchments are covered by sandy clay loam, and loam. Percentage areas covered by the different soil groups in sub-catchment 0 and sub-catchment 1 are shown in Table 3. Sandy soils covering substantial area in Case 1 incorporate more ground water recharge than Case 2 (Table 3). Soils in the Case 2 include more clayey soils, which hold lower infiltration rate resulting a higher runoff potential than the soils in Case 1. But this is outweighed by the more runoff generated in the Case 1 due to the lower root zone storage capacities resulting greater simulated discharge in Case 1 than that of Case 2.

Table 9 shows the simulated maximum, minimum, and average daily discharges for Case 1 and Case 2 and their differences.

According to Table 9, the differences of simulated maximum, minimum, and average daily discharges between the two cases range from 0.5 to 7.2, 0.2 to 0.6 and 0.4 to 1.1 m^3/s, respectively. Of the simulated discharges in both Case 1 and Case 2, the differences are larger for the maximum daily discharges than the minimum daily discharges.

Table 9 Simulated daily discharges (maximum, minimum, average values and their differences) in Case 1 and Case 2

	Q_{sim} Case 1(m^3/s)[a]			Q_{sim} Case 2 (m^3/s)[b]			Q_{sim} (Case 1–Case 2) m^3/s		
	Maximum	Minimum	Average	Maximum	Minimum	Average	Maximum	Minimum	Average
2002	119.2v (Oct. 20)	4.5 (March 12)	24.5	111.0 (Oct. 20)	4.2 (March 12)	23.4	7.2	0.3	1.1
2003	514.9 (May 17)	5.0 (March 3)	32.1	508.0 (May 17)	4.7 (March 3)	31.1	6.9	0.3	1.0
2004	165.0 (May 29)	7.9 (March 9)	29.5	163.3 (May 29)	7.5 (March 9)	28.5	1.7	0.4	1.0
2005	141.2 (Oct. 27)	10.0 (Jan 25)	31.1	139.5 (Oct. 27)	9.4 (Jan. 25)	30.1	1.7	0.6	1.0
2006	201.3 (Nov. 18)	11.2 (Aug 9)	42.4	200.8 (Nov. 18)	11.0 (Aug. 9)	42.0	0.5	0.2	0.4
Average (2002–2006)	228.3	7.7	31.9	224.7	7.4	31.0	3.6	0.3	0.9

[a] Simulated daily discharge in Case 1

[b] Simulated daily discharge in Case 2

Figure 7 shows the flow duration curves for Q_{sim} Case 1 and Q_{sim} Case 2 at Agaliya during the period 2002–2006. Comparison of the flow duration curves shows that there exist only slight variations between the Q_{sim} Case 1 and Q_{sim} Case 2. Such variations are mostly associated with the higher (less than 20 % of the time) and lower (greater than 80 % of the time) discharges. Most of the time, Q_{sim} Case 1 is greater than that of Q_{sim} Case 2 showing a difference of 0.9 m^3/s during 2002–2006, on average (Table 9).

Table 10 shows the flow change over different exceedance levels with related percentile values. As shown in Table 10, the variability is large for the simulated daily discharges greater than 80th percentile and <20th percentile.

To evaluate the variability between the discharges further, it is defined that the high flows as flows that exceed 20 % and low flows as being 80 %. Simulated daily discharges greater than 80th percentile (high flows) and <20th percentile (low flows) for Case 1 in relation to Case 2 are shown in Fig. 8a and b, respectively.

Considering the simulated daily discharges at Agaliya during 2002–2006, the statistical relations between Q_{sim} Case 1 and Q_{sim} Case 2 are compared based on the regression method (Fig. 8). A better correlation is found between Q_{sim} Case 1 and Q_{sim} Case 2 for the high flows with a correlation coefficient (R^2) of 0.9985 than that of the low flows with a R^2 of 0.9524.

Relative differences between the simulated discharges in Case 1 and Case 2 are shown in Fig. 9 with Q_{sim} Case 1 and Q_{sim} Case 2 at Agaliya. The relative differences between the Q_{sim} Case 1 and Q_{sim} Case 2 are higher during the recession limbs of the hydrographs. The recession limb of a hydrograph is the result of the gradual release of water from the catchment which influenced more by the storage characteristics of the catchment. Hence, the reason has been identified as the more contribution from the ground water storage to the Q_{sim} Case 1 due to the high permeable sandy soils covering substantial area in Case 1 than that of Case 2. The maximum relative difference noted is 0.53.

Conclusions

According to the results of the study, YHyM/BTOPMC simulation adequately represents the major hydrological characteristics in Gin river basin including runoff volume,

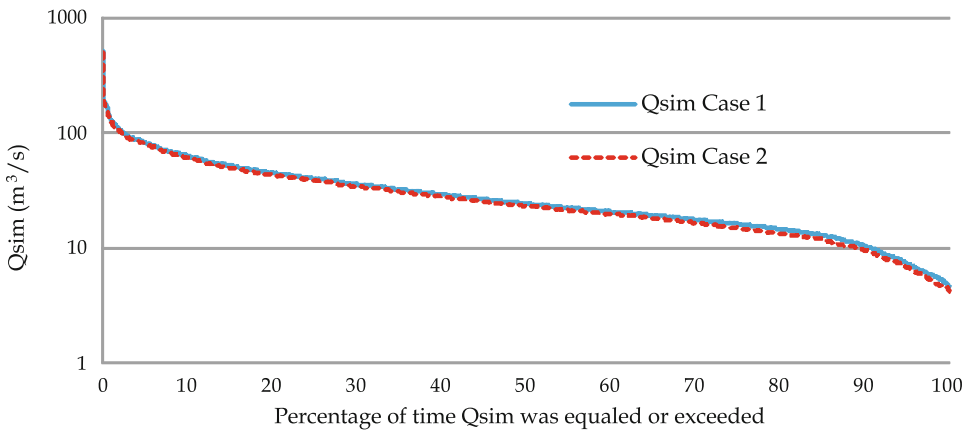

Fig. 7 Flow duration curves for Q_{sim} Case 1 and Q_{sim} Case 2

Table 10 Flow change over different exceedance levels and related percentile values

% Exceedance (percentile)	Q_{sim} Case 1 (m^3/s)	Q_{sim} Case 2 (m^3/s)	Difference (m^3/s)
15 (85th)	51.2	49.7	1.5
20 (80th)	44.3	43.3	1.0
40 (60th)	28.8	28.1	0.7
50 (50th)	23.9	23.2	0.7
60 (40th)	20.4	19.7	0.7
80 (20th)	14.4	13.4	1.0
85 (15th)	12.8	12.0	0.8

evapotranspiration, and soil moisture states of the catchment, utilizing the global public domain data sets (Case 1) as well the local detailed data sets (Case 2). The primary differences between the model input data in Case 1 and Case 2 are associated with the variations in soil type composition and land cover. The water balance components in Case 1 and Case 2 do not differ much from each other, but most of the time, simulated daily discharge in Case 1 is slightly greater than that of Case 2 showing a difference of 0.9 m^3/s during 2002–2006, on average. This difference is due to the higher runoff potential incorporates with more clayey soils in Case 2 which is outweighed by the more runoff generated in Case 1 due to the lower root zone storage capacities.

Evaluation of the simulation results in Case 1 and Case 2 including the flow duration curve plots, correlation comparisons and relative difference calculations further shows that the differences between global public domain data and local detailed data seem to be acceptable as input for the distributed hydrological model applications. The overall results of the study give motivation towards the use of global public domain data for hydrologic simulations in the basins where there are no local data available or where the available local data are too difficult to obtain or where the available detailed local data could not be used shortly for quick water resources assessments. It would be interesting to continue this study to evaluate YHyM/BTOPMC simulation results by using only local detailed data (evaporation, particularly) as the model inputs to see what difference could be obtained.

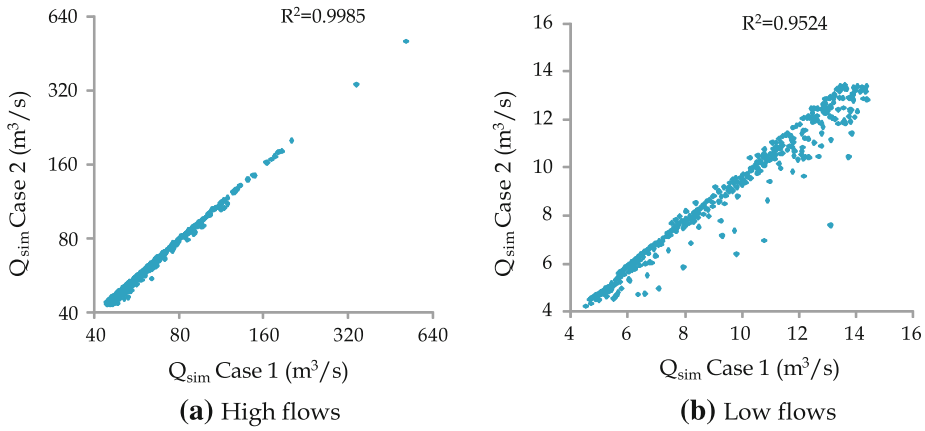

Fig. 8 Correlation of Q_{sim} Case 1 and Q_{sim} Case 2 for (**a**) greater than 80th percentile (*high flows*) and (**b**) less than 20th percentile (*low flows*)

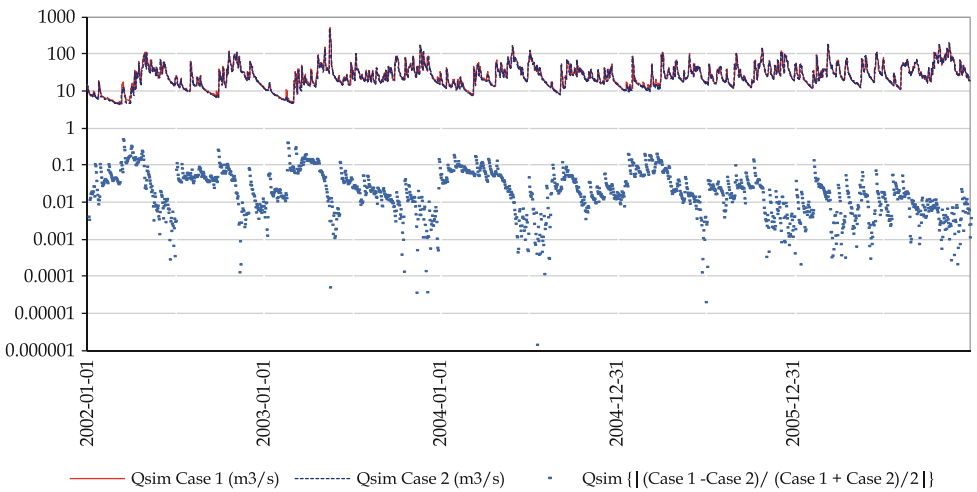

Fig. 9 Relative differences between Q_{sim} Case 1 and Q_{sim} Case 2 [Qsim {|(Case 1 − Case 2)/(Case 1 + Case 2)/2|}] with Q_{sim} Case 1 and Q_{sim} Case 2 at Agaliya

Acknowledgments The authors gratefully acknowledge the Virtual Academy, GCOE (Global Centre of Excellence) Program, University of Yamanashi, Japan and JSPS (Japan Society for Promotion of Science) for providing necessary support for the study including the financial support. Sincere thanks are extended to Dr GHAC Silva for his guidance in making this study a success.

References

Allen RG, Pereira LS, Raes D, Smith M (1998) Crop evapotranspiration: Guidelines for computing crop water requirements. FAO Irrigation and Drainage Paper No. 56. FAO, Rome, Italy

Ao T (2001) Development of a distributed hydrological model for large river catchments and its application to Southeast Asian rivers. Ph.D. Thesis, Department of civil and environmental engineering, University of Yamanashi

Ao TQ, Yoshitani J, Takeuchi K, Fukami K, Mutsura T, Ishidaira H (2003a) Effects of sub-basin scale on runoff simulation in distributed hydrological model: BTOPMC. In: Tachikawa Y, Vieux BE, Georgakakos KP, Nakakita E (eds) Weather radar information and distributed hydrological modeling. IAHS Publ. no. 282, pp 227–234

Ao TQ, Takeuchi T, Ishidaira H, Yoshitani J, Fukami K, Matsuura T (2003b) Development and application of a new algorithm for automated pits removal for grid DEMs. Hydrol Sci J 48:985–997

Beven KJ, Kirkby MJ (1979) A physically based, variable contributing area model of hydrology. Hydrol Sci Bull 24:43–69

Cunge JA (1969) On the subject of a flood propagation computation method (Muskingum method). J Hydraul Res 7:205–230

Das G (2009) Hydrology and soil conservation engineering, 2nd edn. PHI, pp 12–13

FAO/IIASA/ISRIC/ISSCAS/JRC (2009) Harmonized world soil database (version 1.1). FAO, Rome

Hapuarachchi HAP, Kiem AS, Ishidaira H, Magome J, Takeuchi K (2004a) Eliminating uncertainty associated with classifying soil types in distributed hydrologic modeling. In: Proceedings of 2nd APHW Conference, Singapore

Hapuarachchi HAP, Kiem AS, Takeuchi K, Ao T, Magome J, Zhou M (2004b) Applicability of the BTOPMC model for predictions in ungauged basins. In: Proceedings of International Conference sustainable water resources management in changing environment of monsoon region, Colombo, 17–19 Nov 2004

Ishidaira H, Takeuchi K, Ao T (2000) Hydrological simulation of large river basins in Southeast Asia. In: Proceedings of fresh perspectives on hydrology and water resources in southeast Asia and the Pacific, IHP-V Technical Document in Hydrology No. 7, Christ Church, 21–24 Nov 2000, pp 53–54

Jarvis A, Reuter HI, Nelson A, Guevara E (2008) Hole-filled seamless SRTM data V4. International centre for tropical agriculture (CIAT). http://srtm.csi.cgiar.org. Accessed 29 Aug 2010

Lu MJ, Koike T, Hayakawa N (1989) Development of a precipitation-runoff model corresponding to distributed hydrological information (in Japanese). PhD Thesis, Nagaoka University, pp 34–35

Moormann F, Panabokke CR (1961) Soils of Ceylon. The Government Press, Ceylon

National Atlas (2007) National Atlas. 2nd edn. Survey department of Sri Lanka

Nawaratne NMNSB, Ao T, Kazama S, Sawamoto M, Takeuchi K (2001) Influence of Human Activities on the BTOPMC Model Runoff Simulations in Large-Scale Watersheds. In: XXIX IAHR Congress Proceedings, Theme a, pp 93–99

Panabokke CR (1996) Soils and Agro-ecological Environments of Sri Lanka. Natural Resources, Energy and Science Authority of Sri Lanka

Rawls WJ, Brakensiek DL (1985) Prediction of soil water properties for hydrologic modeling. In: Jones EB, Ward TJ (eds) Watershed management in the eighties. Proceedings of a symposium ASCE, Denver, Colorado. 30 April–2 May 1985. ASCE, New York, pp 293–299

Rawls WJ, Brakensiek DL, Saxton KE (1982) Estimation of soil water properties. Trans ASAE 25:1316–1320

Sellers PJ, Tucker PJ, Collatz GJ, Los SO, Justice CO, Dazlich DA, Randall DA (1994) A global 1 degree by 1 degree NDVI data set for climate studies. Part 2: the generation of global fields of terrestrial biophysical parameters from NDVI. Int J Remote Sens 15:3519–3546

Sellers PJ, Los SO, Tucker CJ, Iustice CO, Dazlich DA, Collatz GJ, Randall DA (1996) A revised land surface parameterization (SiB2) for atmospheric GCMs. Part II: the generation of global fields of terrestrial biophysical parameters from satellite data. J Clim 9:706–737

Shuttleworth WJ, Wallace JS (1985) Evaporation from sparse crops—an energy combination theory. Q J R Meteorol Soc 111:839–855

Takeuchi K, Ao T, Ishidaira H (1999) Introduction of block-wise use of TOPMODEL and Muskingum-Cunge method for the hydro-environmental simulation of a large ungauged basin. Hydrol Sci J 44:633–646

Takeuchi K, Hapuarachchi P, Zhou M, Ishidaira H, Magome J (2008) A BTOP model to extend TOPMODEL for distributed hydrological simulation of large basins. Hydrol Process 22:3236–3251

Tucker CJ, Pinzon JE, Brown ME, Slayback D, Pak EW, Mahoney R, Vermote E, Saleous NE (2005) An Extended AVHRR 8-km NDVI Data Set Compatible with MODIS and SPOT Vegetation NDVI Data. Int J Remote Sens 26:4485–5598

USGS (2008) United States Geological Survey—International Geosphere Biosphere Programme (USGS—IGBP) Global Land Cover Characteristics Data Base Version 2.0. http://edc2.usgs.gov/glcc/tabgeo_globe.php. Accessed 02 September 2010

Virtual Academy (2010) http://www.coe.yamanashi.ac.jp/. Accessed 02 September 2010