A macro-scale hydrological model based on a spatial averaging approach: model structure and an application

ZHIYONG YANG, HEPING HU & FUQIANG TIAN

Department of Hydraulic Engineering, State Key Laboratory of Hydroscience and Engineering, Tsinghua University, Beijing 100084, China

huhp@tsinghua.edu.cn

Abstract The inconsistency between the scale at which the hydrological process-governing equations are applicable and the scale at which the model application are practicable lies at the heart of many current hydrological challenges. The spatial averaging approach firstly proposed by Kavvas is one of the promising methods to address such scale issues. This paper presents a macro-scale hydrological model based on the spatial averaging approach, within which the watershed is divided into model computation units (MCUs) and the hydrological processes across MCUs are described by the macro-scale equations derived from point scale equation through spatial averaging approach. As a case study for long-term continuous simulation, the model is applied to the Qin River basin with a semi-arid climate. The continuous streamflow and individual flood events simulation shows that the model can predict streamflow pretty well at lower computation time consumption, which shows the prospects of the model and its potential capabilities for application to water resources and flood management at a large-scale watershed.

Key words heterogeneity; spatial averaging; hydrological model; the Qin River basin

INTRODUCTION

Watershed-scale hydrological response is greatly influenced by the heterogeneity of landscape properties, which are intensified by the nonlinearity of hydrological processes. Such a situation leads to the so-called "scale issue" widely recognized by the hydrological community (Kalma & Sivapalan, 1995), i.e. the mismatch between the scale at which the physical equations are applicable and the scale at which the model applications are practicable. To address the scale issue, Kavvas proposed the spatial averaging approach for upscaling the point scale equations through the analysis of landscape heterogeneity (1999) and nonlinearity of hydrological processes (2003). By adopting the macro-scale equations derived by the spatial averaging approach, the WEHY (Watershed Environmental Hydrology) model was developed by Kavvas *et al.* (2004). As a first attempt, however, the WEHY model suffered from the watershed discretization scheme, as well as the over-complex groundwater simulation scheme (Yang, 2007). Also, the WEHY model has not yet been applied to continuous long-term streamflow simulation.

In this paper, a new macro-scale hydrological model (MSH model) based on the spatial averaging approach, but with a different spatial discretization scheme and simplified groundwater component, is developed and applied to simulate the continuous long-term streamflow in the Qin River basin.

MODEL DESCRIPTION

Spatial discretization

According to its river network and catchment area threshold, a watershed can be divided into some sub-watersheds, and the sub-watershed can be approximately represented by a rectangle hillslope and a river reach. The rectangle hillslope is the model computation unit (MCU) for hillslope hydrological processes simulation and its length and width can be obtained according to the river network density and river length (Yang *et al.*, 2004). The landscape heterogeneity within the MCU is quantified by the statistical approach and directly incorporated into the equations of nonlinear hydrological processes through spatial averaging, which leads to the equations applicable directly at the macro-scale, i.e. MCU. The statistical characteristic values of point scale parameters, such as

Copyright © 2008 IAHS Press

mean, variance, and covariance, are directly incorporated into macro-scale equations as parameters.

Model structure

The MCU is the computation unit for hillslope hydrological processes such as evapotranspiration, infiltration, subsurface flow, etc. and the generated runoff flows into river and routes to the watershed outlet.

Evapotranspiration Total evapotranspiration consists of four components: canopy interception evaporation (E_c), depression storage evaporation (E_d), soil surface evaporation (E_s), and vegetation transpiration (E_{tr}). They are estimated by atmosphere conditions (represented by potential evapotranspiration E_p), vegetation conditions, and available moisture (Yang *et al.*, 2002) as shown in equations (1), (2), (3) and (4):

$$E_c = K_c E_p \tag{1}$$

$$E_d = E_p \tag{2}$$

$$E_{tr} = K_c E_p f(\theta_{root}) LAI/LAI_m \tag{3}$$

$$E_s = E_p f(\theta_{surf}) \tag{4}$$

where K_c is the vegetation coefficient, LAI_m is the maximum LAI of vegetation, θ_{root} is the soil moisture in root zone, and θ_{surf} is the soil moisture at land surface.

Infiltration and unsaturated flow Infiltration and unsaturated flow determine infiltration excess runoff, soil moisture, and water recharge to saturated surface flow and groundwater. Soil characteristics vary dramatically in space, which results in complex impacts on infiltration and unsaturated flow. To account for such heterogeneity, a spatial averaging Green-Ampt model (Chen *et al.*, 1994a,b) is adopted in the MSH model. To confine the problem into a manageable level, only the horizontal heterogeneity of saturated hydraulic conductivity (K_s) is considered. The lognormal distribution is then used to quantify the probability distribution of K_s.

At a specific time t and rainfall density q_t, the wetting front depth and surface moisture are dependent with K_s. Therefore, vertical soil water flux ($q(K_s,z)$) is a function of K_s and depth z as shown in equation (5) (Kavvas *et al.*, 2004):

$$q(K_s,z) = \begin{cases} 0 & K_s < K_{sz} \\ K_s\left[1 + {1}\Big/{Y\left(\dfrac{K_s}{q_t},\tau_v\right)}\right] & K_{sz} < K_s < K_{st} \\ q_t & K_s > K_{st} \quad or \quad K_s > K_{sz} > K_{st} \end{cases} \tag{5}$$

where K_{sz} and K_{st} are two specific K_s, and the wetting front just reaches z when $K_s = K_{sz}$ and the soil surface is just saturated when $K_s = K_{st}$; τ_v is a variable indicating soil water condition; and $Y(\cdot)$ is the function of q_t, K_s, and soil water condition (Kavvas *et al.*, 2004).

The ensemble average of soil water flux ($\langle q(z)\rangle$) at z can easily be obtained by integrating equation (5) over K_s, given by equation (6), and the infiltration excess runoff ($\langle q_r\rangle$) is given in equation (7). Considering the proportion of impermeable area (β_p), the average infiltration excess runoff ($\langle q_{ex}\rangle$) is given in equation (8).

$$\langle q(z)\rangle = \int_0^{+\infty} q(K_s,z) f(K_s) dK_s \tag{6}$$

$$\langle q_r\rangle = q_t - \langle q(z)\rangle\big|_{z=0} \tag{7}$$

$$\langle q_{ex}\rangle = \langle q_r\rangle \beta_p - q_t\left(1-\beta_p\right) \tag{8}$$

where $\langle\cdot\rangle$ is the spatial average operator and $f(K_s)$ is the probability distribution function (pdf) of K_s.

Saturated subsurface flow When unsaturated flow reaches the impermeable stratum, which is assumed parallel to land surface, soil water starts accumulating on the layer and moves downhill as saturated subsurface flow. When the water table rises to the bottom of the rills, saturated soil water will recharge the rills as return flow, and when the water table rises to the land surface, saturation excess runoff will occur.

To account for the heterogeneity along the hillslope transect, the ensemble averaged model based on Darcy's law (Dogrul *et al.*, 1998) is used to model saturated subsurface flow as shown in equation (9):

$$\begin{aligned}\frac{S_s}{R}\frac{\partial\langle h\rangle}{\partial t} &= \left\{\langle K\rangle\frac{\partial}{\partial x}\left(\langle h\rangle\frac{\partial\langle h\rangle}{\partial x}\right)+\frac{\partial^2}{\partial x^2}\left[\langle h\rangle Cov\left(K,h\right)\right]\right\}\cos^2\theta \\ &+\frac{1}{R}\left[\langle K\rangle\frac{\partial\langle h\rangle}{\partial x}+\frac{\partial Cov\left(K,h\right)}{\partial x}\right]\sin\theta\cos\theta+\frac{\langle P\rangle}{R^2}\end{aligned} \tag{9}$$

where S_s is specific yield; K is saturated hydraulic conductivity; R is soil depth; h is the normalized saturated subsurface flow depth; θ is the slope of land surface; and P is the net recharge along the vertical direction.

Saturated subsurface flow (q_{sat}), return flow (q_{rf}), and saturation excess runoff (q_s) can be estimated by equations (10), (11), and (12), respectively:

$$q_{sat} = R\sin\theta\cos\theta(Kh_0 + \mathrm{cov}(h_0,K)) \tag{10}$$

$$q_{rf} = 2K\sin\theta_y\cos\theta_y R(d_r + h-1)N_{ro} \tag{11}$$

$$q_s = S_s R(h-1)/\Delta t \tag{12}$$

where h_0 is the normalized subsurface flow depth at the downstream transect; θ_y is the slope of the interrill area along the y-direction; d_r is the normalized rill depth; N_{ro} is the rill occurrence density; and Δt is time step.

Surface flow The land surface of hillslope contains irregular local topography, which is presented by a rill–interrill configuration. The influence of rill formation on soil erosion and flow dynamics is recognized in many experimental studies (Morris, 1979; Abrahams *et al.*, 1989; Govindaraju & Kavvas, 1992, 1994a,b; Parsons & Wainwright, 2006). To account for the effects of rill–interrill configuration, spatially averaged conservation equations for interacting rill–interrill area surface flow model was developed by Tayfur & Kavvas (1994) and improved by Yoon & Kavvas (2000). The Yoon's model is adopted in the MSH model with some improvements. The ensemble averaged equations for interrill areas and rills are as follows:

$$\begin{aligned}&\frac{\partial\langle\bar{h}_{ir}\rangle}{\partial t}+c_1\frac{\partial}{\partial x}\left[\left(K_x(\langle\bar{r}\rangle)+\frac{1}{2}\sum_i\sum_j \mathrm{cov}(r_i,r_j)\frac{\partial^2 K_x(\bar{r})}{\partial r_i\partial r_j}\right)\langle\bar{h}_{ir}\rangle^{5/3}\right] \\ &=\langle q_{ex}\rangle+\langle q_s\rangle-c_2\rho_{ir}\left[K_y(\langle\bar{r}\rangle)+\frac{1}{2}\sum_i\sum_j \mathrm{cov}(r_i,r_j)\frac{\partial^2 K_y(\bar{r})}{\partial r_i\partial r_j}\right]\langle\bar{h}_{ir}\rangle^{5/3}\end{aligned} \tag{13}$$

$$\begin{aligned}&\frac{\partial\langle\bar{h}_r\rangle}{\partial t}+\frac{\partial}{\partial x}\left\{K_r(\langle\bar{r}\rangle)\frac{\langle w_r\rangle^{1/2}\langle\bar{h}_r\rangle^{5/3}}{\left(\langle w_r\rangle+2\langle h_r\rangle\right)^{1/2}}+\frac{1}{2}\sum_i\sum_j \mathrm{cov}(r_i,r_j)\frac{\partial^2}{\partial r_i\partial r_j}\left[K_r(\bar{r})\frac{\langle w_r\rangle^{1/2}\langle\bar{h}_r\rangle^{5/3}}{\left(\langle w_r\rangle+2\langle h_r\rangle\right)^{1/2}}\right]\Bigg|_{\bar{r}=\langle\bar{r}\rangle}\right\} \\ &=\langle q_{ex}\rangle+\langle q_s\rangle+\frac{\langle q_{rf}\rangle}{\lambda_{rf}}+2\rho_r c_2\left[K_y(\langle\bar{r}\rangle)+\frac{1}{2}\sum_i\sum_j \mathrm{cov}(r_i,r_j)\frac{\partial^2 K_y(\bar{r})}{\partial r_i\partial r_j}\right]\langle\bar{h}_r\rangle^{5/3}\end{aligned} \tag{14}$$

where $\bar{h}_{ir}$ and $\bar{h}_r$ are average flow depths of every interrill area and rill, c_1 = 1.0944, c_2 = 2.1888 when the flow depth curve within single interrill area are assumed to be a power function; λ_{ra} is the rill surface proportion, $\rho_r = N_{ro}/\lambda_{ra}$, $\rho_{ir} = N_{ro}/(1-\lambda_{ra})$; w_r is the width of rills; l is the width of interrill area; $K_x(\langle\bar{r}\rangle) = \dfrac{\langle S_x\rangle^{1/2}}{\langle n_{ir}\rangle\left[1+\left(\dfrac{\langle S_x\rangle}{\langle S_y\rangle}\right)^2\right]^{1/4}}$, $K_y(\langle\bar{r}\rangle) = \dfrac{\langle S_y\rangle^{1/2}}{\langle n_{ir}\rangle\left[1+\left(\dfrac{\langle S_y\rangle}{\langle S_x\rangle}\right)^2\right]^{1/4}}$, $K_r(\langle\bar{r}\rangle) = \dfrac{1}{\langle n_r\rangle}$,

where S_x and S_y are bed slopes along x- and y-directions of interrill areas, respectively; S_r is bed slope of the rill; n_{ir} and n_r are the Manning roughness coefficients of interrill areas and rills respectively; and $\bar{r}$ is a vector of the hydraulic parameters mean value at hillslope scale, $\bar{r} = \{\bar{S}_x, \bar{S}_y, \bar{n}_{ir}, \bar{l}\}$ for interrill areas and $\bar{r} = \{\bar{S}_r, \bar{n}_r, \bar{w}_r\}$ for rills.

Groundwater Groundwater provides the baseflow in a stream. The water exchange between stream reaches and groundwater is estimated by equation (15):

$$Q_g = K_g(h_g^2 - h_r^2)/L \tag{15}$$

where K_g is hydraulic conductivity in an aquifer; h_g is water table of groundwater from the bed of stream reach; h_r is water depth of stream reach; and L is the width of hillslope.

River routing Stream reaches are fed by lateral inflow from neighbouring hillslopes, including surface flow, subsurface flow, and groundwater flow. River routing is simulated by 1-D kinematic wave equations (Chow, 1988).

MODEL APPLICATION

Study area and data

The Qin River is the upper stream of the Qin River basin, and it has a catchment area of 1355 km^2. The climate in the study area is typically semi-arid with annual pan-evaporation of 1700 mm and annual rainfall of 611 mm.

The input data of the model includes DEM, soil class, land use, leaf area index (LAI) data, and meteorological data. The watershed boundary and stream networks, as shown in Fig. 1, were generated from DEM with spatial resolution of 100 m. According to stream networks, the study area was divided into 35 MCUs. The geomorphologic parameters of each MCU were obtained through GIS analysis. The rainfall data of 12 raingauges (as shown in Fig. 1) were used to generate the model rainfall input. The daily stream discharge data at the outlet, i.e. Kongjiapo stream station, was used for model calibration and validation. Meteorological data such as temperature, sunshine hours, humidity, etc. were collected from a station located at the middle reach and used to estimate potential evapotranspiration. The land use of the study area was reclassified into six types: cultivated land, forest, grassland, water body, urban area, and other, as shown in Fig. 2. The cultivated land, forest, and grassland covered >99% of the study area. The soil class map was transferred from the FAO/UNESCO data set with spatial resolution of 10 km and the associated soil hydrological properties were taken from the same data set. The monthly *LAI* date is derived from the MODIS TERRA MOD15A2 product with spatial resolution of 1 km.

RESULT AND DISCUSSION

The model was applied to simulate continuous streamflow in the study area. Hourly rainfall data at 12 raingauge stations and daily discharge data at the Kongjiapo station are available from 1986 to 1989. The Thiessen method was used to assign representative areas for each station in order to determine the spatial distribution of rainfall. The period from 1986 to 1987 was selected for model warming up. The soil hydraulic parameters were estimated from the soil map and associated soil

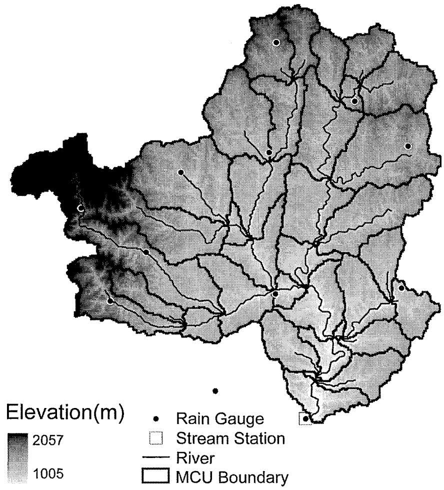

Fig. 1 DEM, MCUs, stream network and locations of rain and stream gauging station.

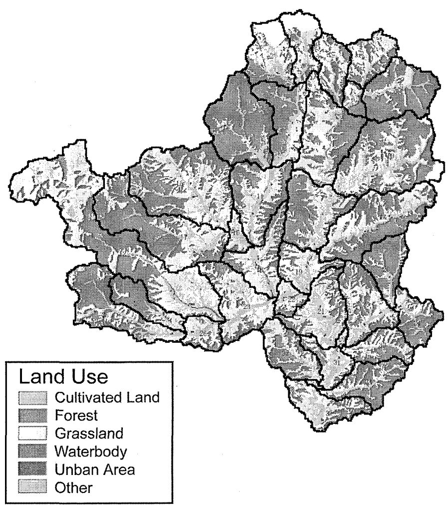

Fig. 2 Land use of the study area (overlaid with MCUs).

properties. The parameters of variance and covariance were determined with reference to former studies (Yoon & Kavvas, 2000).

The parameters such as soil depth (R), Manning's roughness of hillslope (n_s) and channel (n_r), soil anisotropy coefficient (C_a), and aquifer hydraulic conductivity (C_g), need to be estimated through model calibration. The period of 1988 was selected for model calibration. The pre-assigned range of the parameter and the calibrated results are listed in Table 1. The simulated hydrograph during the calibration period is shown in Fig. 3. The period of 1989 was selected for model validation. The hydrograph at the Kongjiapo Station during the validation period is shown in Fig. 4.

It can be seen from Figs 3 and 4 that the model-predicted hydrograph at the Kongjiapo station matched the observed data quite well. The Nash-Sutcliffe efficiency coefficient (E) (Nash & Sutcliffe, 1970) and coefficient of determination (R^2) were selected to quantify model performance (see Table 2). The results show that E in 1988 and 1989 reaches 0.88 and 0.94, respectively, and

R^2 reaches 0.89 and 0.97, respectively. Furthermore, the two coefficients during the validation period are similar to those during the calibration period (see Table. 2).

There are three flood events during calibration and validation. The model performance for a flood event is quantified simultaneously by relative error of flood peak discharge, the Nash-Sutcliffe efficiency coefficient (E), and coefficient of determination (R^2) during flood events and is listed in Table 3. The predicted flood peak and hydrograph matches the observed data very well.

Table 1 Parameter ranges and fitted values.

Parameter	R (m)	n_s	n_r	C_a	C_g
Range	0.5–5.0	0.1–1.0	0.01–0.1	1.0–10.0	0.01–1.0
Fitted Value	0.8,1.2,1.6[a]	0.5	0.05	5.0	0.05

[a] R = 0.8 ,1.2, and 1.6 for upstream, middle stream, and downstream of the study area.

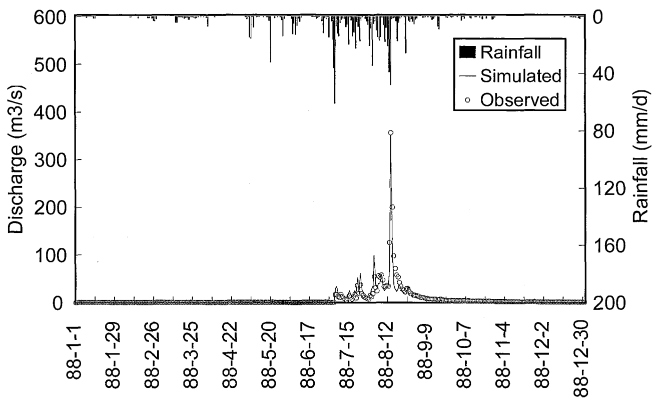

Fig. 3 Comparison of observed and simulated hydrographs during calibration period.

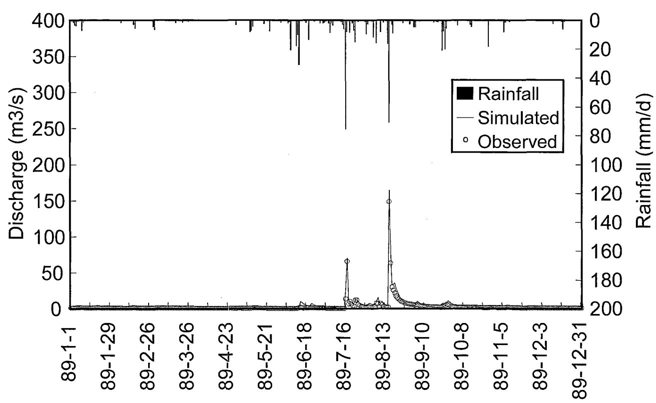

Fig. 4 Comparison of observed and simulated hydrographs during validation period.

Table 2 Model performance during calibration and validation periods.

Period	E	R^2
1988	0.88	0.89
1989	0.94	0.97

Table 3 Model performance during three flood events.

Flood events	Q_p (m^3 s^{-1}) Simulated	Observed	R_e (%)	E	R^2
7/05/1988–9/22/1988	350.63	356.00	−1.51	0.85	0.86
7/15/1989–7/22/1989	69.92	65.90	6.10	0.92	0.97
8/15/1988–9/21/1988	164.14	149.00	10.16	0.96	0.99

The high computation efficiency is obtained via coarse discretization of the watershed into MCUs, which does not lower the model performance. In the case study, 178 s was consumed for 35 064 h of hydrological modelling in a personal computer with an AMD Athlon 2500+ CPU and memory of 1GB.

SUMMARY

As a new method to address the scale issues in hydrological modelling, the spatial averaging approach is proposed to directly incorporate the micro-scale heterogeneity into the macro-scale process equations. A new macro-scale hydrological model is developed based on the spatial averaging approach. Compared to the former version of spatial averaging model, the WEHY model, our model modifies the spatial discretization scheme and simplifies the groundwater system. The model is intended to be able to simulate the long-term large scale hydrological processes at the watershed-scale with less computational burden. Our case study in the Qin River basin (1355 km^2) shows that the model can obtain the higher computation efficiency, as well as the sound performance of runoff simulation, which show the potential capabilities of the model for water resources and flood management application in semiarid regions.

Acknowledgements This research was supported by the National Natural Science Foundation of China (50721140161 and 50579031).

REFERENCES

Abrahams, A. D., Parsons, A. J. & Luk, S. H. (1989) Distribution of depth of overland flow on desert hillslopes and its implications for modelling soil erosion. *J. Hydrol.* **106**(1–2), 177–184.

Chen, Z. Q., Govindaraju, R. S. & Kavvas, M. L. (1994a) Spatial averaging of unsaturated flow equations under infiltration conditions over areally heterogeneous fields. 1. Development of models. *Water Resour. Res.* **30**(2), 523–533.

Chen, Z. Q., Govindaraju, R. S. & Kavvas, M. L. (1994b) Spatial averaging of unsaturated flow equations under infiltration conditions over areally heterogeneous fields. 2. Numerical simulations. *Water Resour. Res.* **30**(2), 535–548.

Chow, V. T. (1988) *Applied Hydrology*. McGraw-Hill, New York, USA.

Dogrul, E. C., Kavvas, M. L. & Chen, Z. Q. (1998) Prediction of subsurface stormflow in heterogeneous sloping aquifers. *J. Hydrol. Engng ASCE.* **3**(4), 258–267.

Govindaraju, R. S. & Kavvas, M. L. (1992) Characterization of the rill geometry over straight hillslopes through spatial scales. *J. Hydrol.* **130**(1–4), 339–365.

Govindaraju, R. S. & Kavvas, M. L. (1994a) A spectral approach for analyzing the rill structure over hillslopes. Part 1. Development of stochastic theory. *J. Hydrol.* **158**(3–4), 333–347.

Govindaraju, R. S. & Kavvas, M. L. (1994b) A spectral approach for analyzing the rill structure over hillslopes. Part 2. Application. *J. Hydrol.* **158**(3–4), 349–362.

Kalma, J. D. & Sivapalan, M. (eds) (1995) *Scale Issues in Hydrological Modeling.* John Wiley & Sons Inc., London, UK.

Kavvas, M. L. (1999) On the coarse-graining of hydrologic processes with increasing scales. *J. Hydrol.* **217**(3–4), 191–202.

Kavvas, M. L. (2003) Nonlinear hydrologic processes: Conservation equations for determining their means and probability distributions. *J. Hydrol. Engng ASCE* **8**(2), 44–53.

Kavvas, M. L., Chen, Z. Q., Dogrul, C., Yoon, J. Y., Ohara, N., Liang, L., Aksoy, H., Anderson, M. L., Yoshitani, J., Fukami, K. & Matsuurall, T. (2004) Watershed Environmental Hydrology (WEHY) model based on upscaled conservation equations: hydrologic module. *J. Hydrol. Engng ASCE* **9**(6), 450–464.

Morris, E. M. (1979) The effect of the small-slope approximation and lower boundary conditions on solutions of the Saint-Venant equations. *J. Hydrol.* **40**(1–2), 31–47.

Nash, J. & Sutcliffe, J. (1970) River flow forecasting through conceptual models-PT1. *J. Hydrol.* **10**(3), 282–290.

Parsons, A. J. & Wainwright, J. (2006) Depth distribution of interrill overland flow and the formation of rills. *Hydrol. Processes* **20**(7), 1511–1523.

Tayfur, G. & Kavvas, M. L. (1994) Spatially averaged conservation equations for interacting rill–interrill area overland flows. *J. Hydraul. Engng* **120**(12), 1426–1448.

Yang, D., Herath, S. & Musiake, K. (2002) A hillslope-based hydrological model using catchment area and width functions. *Hydrol. Sci. J.* **47**(1), 49–65.

Yang, D., Li, C., Ni, G. & Hu, H. (2004) Application of a distributed hydrological model to the Yellow River Basin. *Acta Geographica Sinica* **59**(1), 143–154.

Yang, Z. (2007) Study on spatial averaging watershed hydrological model at macro-scale based on probability approach. PhD Thesis, Tsinghua University, Beijing, China.

Yoon, J. & Kavvas, M. L. (2000) Spatial averaging of overland flow equations for the hillslope with rill–interrill configuration. In: *Proc. ICHE-2000, Published on CD, Seoul, Korea, 26–29 September.*