Rainfall-runoff modeling, parameter estimation and sensitivity analysis in a semiarid catchment

Yan Jiang a, *, Changming Liu b, Xuyong Li a, Lifang Liu c, Hongrui Wang c

a State Key Laboratory of Urban and Regional Ecology, Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing 100085, China
b Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China
c College of Water Sciences, Beijing Normal University, Beijing 100087, China

1. Introduction

Given the high spatiotemporal variability of precipitation and watershed characteristics, the rainfall-runoff relationship can become very complicated (Kumar et al., 2005). Conceptual hydrologic models are popular tools for simulating major rainfall-runoff processes. Many hydrologists have contributed to hydrologic model development and evaluation (Andrews et al., 2011). However, many models were developed for more humid areas (Bahat et al., 2009), and analogous work for semiarid regions is relatively rare and inadequate. There is thus an obvious need for improved understanding of the hydrology of semiarid and semi-humid regions, and development of appropriate modeling of runoff (Yu et al., 2011).

Rainfall-runoff modeling of semiarid catchments is a challenging task (Ye et al., 1997; Wheater et al., 2008), owing to a delicate hydrologic balance and different mix of hydrologic processes (Pilgrim et al., 1988). Runoff generation in such zones may involve two different mechanisms: infiltration excess and saturation excess (Scoging, 1979; Gallart et al., 1994; Taha et al., 1997; Fitzjohn et al., 1998; Martõñez-Mena et al., 1998). Infiltration excess flow results from rainfall intensity greater than the local infiltration capacity of soil (Yair and Lavee, 1985). With the increase of accumulated infiltration amount, saturation excess flow occurs. Both mechanisms may occur simultaneously on the same slope, so it may be very difficult to distinguish these forms of runoff generation (Coles et al., 1997; Pinol et al., 1997; Beven and Freer, 2001). In addition, rainfall in semiarid climate regions is typically meager and highly variable (Sharon, 1972; Goodrich et al., 1995; Ahrens, 2003). Rain intensity data for short durations are important inputs to rainfall-runoff models for the computation of infiltration and rainfall excess. However, high temporal and spatial variability of rainfall, along with sparse rainfall and runoff gauge stations (Faures et al., 1995), increase the difficulty of modeling hydrologic processes (Michaud and Sorooshian, 1994; Nouh, 2006) and uncertainty in streamflow predictions (Yatmseendradas et al., 2008). Wheater (2002) emphasized the difficulties in obtaining reliable long-term hydrologic records in arid and semiarid regions, owing to sparse populations, limited economic resources, harsh climates and infrequent but damaging hydrologic events.

Several physically-based hydrologic models are already in use for semiarid environments. Foremost are the models SCS (Michaud and Sorooshian, 1994; Symeonakis and Drake, 2004), VIC (Abdulla

© 2015 Elsevier Ltd. All rights reserved.
and Lettenmaier, 1997; Stephen et al., 2010), KINEROS (Michaud and Sorooshian, 1994; Hernandez et al., 2000; Al-Qurashi et al., 2008), TOPMODEL (Candela et al., 2005), MIKE SHE (McMichael et al., 2006), CRUM (Reaney et al., 2007), HBV (Masih et al., 2010; Jia and Sun, 2012), IHACRES (Jakeman et al., 1990; Jakeman and Hornberger, 1993; Evans and Jakeman, 1998; Dye and Croke, 2003; Croke et al., 2006; McIntyre and Al-Qurashi, 2009), HEC-HMS (Halwatura and Najim, 2013), and SWAT (Hernandez et al., 2000; Yu et al., 2011; Perrin et al., 2012). Some of these, however, share certain problems that reduce their effectiveness in semiarid regions. First, some models only consider a single overland flow generation mechanism, such as TOPMODEL, or saturation flow generation mechanism, such as VIC, so they do not perform well in semiarid areas. It has been suggested that both infiltration excess and saturation excess runoff mechanisms should be included in a rainfall-runoff model of semiarid regions (Hu et al., 2005). Second, some models, e.g., TOPMODEL and KINEROS, are complex hydrologic cycle models that require substantial hydrological, meteorological, and high-resolution watershed physical property data. It is difficult to acquire such data in semiarid regions. Given strong high rainfall variability in dry environments, a high-resolution and complex hydrologic model should be suitable to these environments. In practice, however, such models often do not provide more accurate runoff predictions than simpler ones. Michaud and Sorooshian (1994) found similar performances of complex KINEROS and SCS models for simulating runoff in a semiarid Walnut Gulch experimental watershed. Al-Qurashi et al. (2008) found that the performance of KINEROS model was poor compared with that of a simple regression model in an arid Oman watershed. In addition, increasing model complexity leads to greater uncertainty in its predictions. Finally, some models, e.g., SCS and SWAT, use the popular curve number (CN) method to determine amounts of infiltration and surface runoff. The CN model is based on the assumption that there is a unique relationship between average moisture content and CN for all hydrologic response units, and that the moisture content distribution is similar for each runoff event. This is not the case in many regions (White et al., 2011), especially semiarid ones.

Based on the energy conservation law and water balance equation, Liu et al. (Liu, 1978; Liu and Wang, 1980) formulated a rainfall infiltration model, LCM model, via analysis of data measured in a number of small experimental watersheds and various physical geographic factors in China. It was discovered that rainfall infiltration and runoff is not only related to rainfall intensity, but also to soil moisture content. Using theoretical analyses and numerical simulations, Li et al. (2014) compared the SCS and LCM models. They found that SCS model is a linear representation of LCM model, and the latter can reflect nonlinear runoff. Because LCM model contains infiltration excess and saturation excess runoff mechanisms, it has great applicability. Applying LCM model to calculate runoff, Liu et al. (2008) developed the Hydro-Informatic Modelling System (HIMS) model, which can be used for hydrological simulations, flood frequency analyses, flood warning system planning, hydraulic engineering design, stream restoration, and others. The model has been used successfully in China and Australia and been proven very efficient in simulating streamflow using a simple structure. The model has not been calibrated and validated for the Luanhe River watershed, so we must determine its suitability for semiarid basins.

Successful application of a hydrologic model depends on how well its parameters are selected. Normally, the model consists of a large number of parameters. These parameters cannot be determined from field measurement or estimated from watershed characteristics directly, but must be calibrated (Jakeman et al., 2006). Calculation of hydrologic models is used to improve their predictability by local observational data. Many optimization methods have been proposed and improved to select model parameters using calibration. Foremost are the Genetic Algorithm (GA) (Wang, 1991, 1997; Cheng et al., 2002), SCE-UA (Duan et al., 1992, 1994; Vrugt et al., 2003a, 2003b) and particle swarm optimization (PSO) (Gill et al., 2006; Zhang et al., 2009; Jiang et al., 2010; Kuok et al., 2010; Kraue et al., 2011). GA is a highly parallel and adaptive optimization algorithm, but it is easily driven into local optima even non-stationary points, owing to linear operations for generating new particles (Chu et al., 2010). PSO was proposed by Kennedy and Eberhart (1995) based on the analogy of swarming animals and is a simple and powerful heuristic method for solving nonlinear, non-differential and multimodal optimization problems. Since Gill et al. (2006) used PSO for parameter estimation in hydrology, many researchers have presented various types of PSO for hydrologic model calibration (Jiang et al., 2007, 2010, 2013; Zhang et al., 2009; Kuok et al., 2010; Kraue et al., 2011). PSO approach has many computational advantages over traditional evolutionary computing, such as rapid convergence (Jiang et al., 2006, 2007; Chau, 2007). Nonetheless, similar to general heuristic method, PSO also has the drawback of premature convergence, which degrades its performance and global search ability. Therefore, Wang et al. (2012) developed a more self-adaptive and robust PSO version (SPSO) to provide users a black-box tool for various application problems. On the basis of SPSO, Jiang et al. (2013) proposed a master-slave swarms shuffling evolution algorithm based on self-adaptive particle swarm optimization (MSSE-SPSO). Generally, performance is improved if particle velocities are selected as a time varying quantity (Chatterjee and Siarry, 2006), and search neighbor is reduced dynamically. In this paper, a novel hybrid approach is proposed by combining strengths of a self-adaptive dynamic PSO, competitive evolution and sub-swarm shuffling, which maintains population diversity and avoids premature convergence. It is called the master-slave swarm shuffling evolution algorithm, based on the self-adaptive dynamic PSO (MSSE-SDPSO). And then the approach was used for calibrating HIMS hydrological model.

The present modeling study was conducted for the Luanhe River watershed, a semiarid catchment in China. The objectives were as follows: (1) Present a novel calibration method and test its performance. (2) Check the suitability of HIMS model in the semiarid basins. (3) Analyze the effect of model structure and rainfall details on the efficiency of streamflow simulation. (4) Examine the sensitivity of HIMS model parameters.

2. Study area and data

2.1. Study area

The Luanhe River basin study area is between 115°30′–119°45′E and 39°10′–42°40′N in northern China (Fig. 1). The Luanhe River rises on the Mongolia Plateau and passes through the Yanshan Mountains, ultimately emptying into Bohai Bay. River length is about 888 km, and total basin area is 44,070 km², of which 98% is mountainous. Average annual temperature is 5–8 °C, with maximum 43.3 °C and minimum –28.4 °C. Average annual precipitation is 535 mm. Precipitation varies monthly and annually. Precipitation in flood season (June–September) generally accounts for 70–85% of annual precipitation, and concentrates in several rain events during July and August. Annual variability of precipitation is very great, with a maximum twice that of the minimum. This is attributable to unstable characteristics of the duration, intensity,
and location of the subtropical high over the northern Pacific in summer. Runoff in the flood season accounts for 60–80% of the annual total. Average annual evaporation is 1142 mm, and the aridity index is 2.2. Climate in the study area is classified as semi-arid, meaning that potential evapotranspiration is between 1.5 times and 3.49 times precipitation (Zhang and Li, 1999). Eight main tributaries were chosen as the study, the Liuhe, Baohe, Sahe, Laoniuhe, Wuliehe, Xingzhouhe, Yimatuhe, and Yixunhe Rivers. Related information is listed in Table 1.

The Luanhe River basin mainly includes three geomorphic types, plateaus, mountains and plains, with elevations decreasing from northwest to southeast. The area with loose and thin soil layers, is dominated by brown soil and cinnamon soil (more than 80%), and the rest is covered by fluvo-aquic, chestnut, skeleton, gray forest and other soils (Chen et al., 2009). Vegetation coverage is poor and soil erosion is serious, especially in Yixunhe, Wuliehe and Xingzhouhe River sub-basins. It was found that there was not a single runoff generation mechanism in the basin. Rainfall characteristics

---

**Table 1**

| Sub-basin          | Area (km²) | Mean annual P (mm) | Mean annual T (°C) | Mean annual R (mm) | Mean annual E (mm) | Mean annual R/P | Mean annual E/P | Forest coverage (%) | River slope (%) |
|--------------------|------------|--------------------|--------------------|--------------------|--------------------|----------------|----------------|-------------------|-----------------|
| Liuhe River        | 626        | 670                | 8.4                | 142.22             | 1101               | 0.21           | 1.64           | 55.3              | 3.25            |
| Baohe River        | 1661       | 509                | 10.5               | 76.54              | 1164               | 0.15           | 2.29           | 52.4              | 2.98            |
| Sahe River         | 646        | 717                | 8.8                | 151.86             | 1102               | 0.21           | 1.54           | 65.1              | 2.07            |
| Laoniuhe River     | 1015       | 499                | 10.8               | 62.39              | 1215               | 0.13           | 2.43           | 52.0              | 2.54            |
| Wuliehe River      | 2200       | 512                | 9.7                | 47.83              | 1212               | 0.09           | 2.37           | 51.4              | 4.24            |
| Xingzhouhe River   | 1378       | 470                | 8.5                | 41.74              | 1137               | 0.09           | 2.42           | 51.7              | 10.1            |
| Yimatuhe River     | 2404       | 441                | 6.7                | 25.83              | 1138               | 0.06           | 2.58           | 52.8              | 5.47            |
| Yixunhe River      | 6761       | 463                | 8.6                | 42.65              | 1066               | 0.09           | 2.30           | 48.3              | 6.80            |
and initial soil moisture have an effect on this generation. If rainfall intensity is low and initial soil moisture is high, saturation-excess runoff is dominant. If rainfall intensity is high and initial soil moisture is low, the infiltration-excess runoff dominates (Li and Feng, 2009).

During recent years, many hydrologists have investigated variations and trends of hydrological time series and proposed approaches for forecasting runoff in the Luanhe River basin (Li and Feng, 2007). Based on data from 34 meteorological stations in the basin from 1957 to 2008, the long-term trend of major climatic variables was analyzed by the Mann–Kendall (M–K) test method. The results showed that annual precipitation had a decreasing trend, especially in summer. Annual air temperature had a significant increasing tendency, especially in winter and spring (Wang et al., 2011). According to classification of precipitation frequency, rainfall series and corresponding runoff series were grouped. The evolution of various groups of both series was analyzed based on variable fuzzy set theory for six sub-basins in the Luanhe River basin (Wei and Feng, 2011). Bao et al. (2012) detected the long-term trend of hydrological time series including air mean temperature, precipitation, and streamflow in the Luanhe River basin using both nonparametric M–K test techniques. The impact of precipitation variations and human activities on annual runoff in the basin was evaluated quantitatively, based on a rainfall-runoff empirical statistical model (Yuan et al., 2012).

However, there have been few attempts to apply a rainfall-runoff model to simulate streamflow in the Luanhe River basin. The Luanhe River is the main water source for the cities of Tianjin and Tangshan in northern China. In recent years, water resources in the river basin have become heavily committed. Rapid economic development and population growth in this region have caused serious concerns over the future adequacy of quantity and quality of water withdrawn from the river. Therefore, a modeling study is urgently needed to evaluate water availability in this basin.

2.2. Data

Data for model calibration and validation included daily air temperature, rainfall and streamflow. Daily air temperature was collected from the China Meteorological Data Sharing Service System. Measurements of streamflow during 1995–2008 were from eight monitoring stations operated by the Chengde Bureau of Hydrology and Water Resources Survey of Hebei Province. Digital elevation model (DEM) data of resolution 30 m for the Luanhe River basin were downloaded from the International Scientific Data Service Platform, Chinese Academy of Sciences (http://datamirror.csdb.cn/).

Ten-minute rainfall in the Luanhe River basin was observed from 27 rainfall stations located in the river basin. These include spline (regularized and tension), inverse distance weighting, trend surface, kriging, and Thiessen polygon methods. The density and distribution of gauge stations are important for accuracy of all methods. Model input of areal rainfall was calculated from records of the 27 stations, using the Thiessen polygon method.

3. Methodology

3.1. HIMS rain-runoff model

HIMS model is a distributed hydrologic model and can simulate actual evapotranspiration, infiltration, surface runoff, interflow runoff, groundwater runoff, groundwater recharge and channel routing (Liu et al., 2006, 2008, 2010a, 2010b). Based on topological relationships of channel network and properties of soil, vegetation and land-use, a catchment is divided into different grids. Each grid includes a channel, and grids are linked throughout the stream network. The hydrological cycle of each grid or sub-basin in HIMS model is described as shown in Fig. 2 (Liu et al., 2008). The network structure of channel routing is described as shown in Fig. 3 (Liu et al., 2006). A block diagram of HIMS model operation is shown in Fig. 4 (Liu et al., 2008).

In HIMS model, daily rainfall is first subjected to an infiltration function that determines infiltration capacity; rainfall exceeding the capacity becomes surface runoff (infiltration excess runoff). Infiltrated water is then diverted to interflow, groundwater recharge and the soil moisture store. Interflow is estimated as a linear function of soil wetness (soil moisture amount divided by soil moisture capacity). Groundwater recharge is estimated, also as a linear function of soil wetness. Remaining moisture after deduction of evapotranspiration loss flows into the soil moisture store. Evapotranspiration from that store is estimated using a conceptual model, which is related to soil moisture storage, storage capacity, and potential evapotranspiration. The soil moisture store has finite capacity and overflows into the groundwater store. Baseflow from the latter store is simulated as a linear recession from it. The model therefore estimates runoff generation from three sources, i.e., surface runoff, interflow and baseflow. Channel routing is achieved by applying the Muskingum method to successive sub-reaches. Here, we briefly describe the calculations for potential evapotranspiration, actual evapotranspiration, infiltration, surface runoff, interflow runoff, groundwater runoff, groundwater recharge, and channel routing. For more details, refer to Liu et al. (2008).

HIMS model considers the underlying watershed surface as a whole, and evapotranspiration is consumed from soil water. Potential evapotranspiration in the model is calculated by the Hargreaves–Samani method (Hargreaves and Samani, 1985). Actual evaporation is calculated by a conceptual model that is related to soil water content and potential evapotranspiration. Potential evapotranspiration $ET_0$ (mm) and actual evapotranspiration $ET_a$ (mm) are described as follows.

$$ET_0 = \alpha \left( \frac{R_{\text{max}}}{T} \right) (T + 17.8)/(T_{\text{max}} - T_{\text{min}})^b$$

$$ET_a = ET_0 \left( 1 - \left( 1 - \frac{S}{S_{\text{MAX}}} \right)^{C_1} \right)$$

where $\alpha$ and $b$ are empirical parameters; $R_{\text{max}}$ is maximum radiation (MJ m$^{-2}$ d$^{-1}$); $T$, $T_{\text{max}}$ and $T_{\text{min}}$ are daily average, maximum and minimum temperatures,

![Fig. 2. Structure of runoff generation model for a hydrological unit in HIMS.](image-url)
respectively (°C); and \(f\) is latent heat of vaporization (MJ kg\(^{-1}\)); \(\epsilon\) is the evapotranspiration coefficient in need of calibration.

HIMS model considers two runoff generation mechanisms: infiltration excess and saturation excess. The infiltration simulation is modeled based on kinematic energy and continuity equations:

\[
\frac{du}{dt} = \frac{w}{\rho} \left( y + H + h_2 \right) / \gamma
\]

(3)

\[
q_t dt = w dy.
\]

(4)

where \(u\) is the speed of wetting front movement; \(\rho\) is the density of water; \(g\) is the acceleration of gravity; \(y, H\) and \(h_2\) are gravitation potential head, pressure potential, and osmotic potential head, respectively; \(\gamma\) is a coefficient; \(q_t\) is infiltration rate; and \(w\) is cross-sectional area of the porous medium.

According to Eqs. (3) and (4), we obtain.

\[
q_t = w \frac{dy}{dt} = w - w \gamma (y + H + h_1) / \gamma.
\]

(5)

For given soil type, \(w\) is a fixed value and \(\rho, g, \gamma\) are constants, \(soq_t\) is proportionate to \(H\) and \(h_2\). Because \(H\) depends on rainfall intensity \((I)\), infiltration rate is related to rainfall intensity. Eq (5) can be expressed approximately as: \(q_t = R - f_I\), which has been confirmed through analysis of data measured in a number of small experimental watersheds and various physical geographic factors in China (Liu et al., 2006). Therefore, infiltration can be modeled using an empirical relationship:

\[
f_I = R P_t,
\]

(6)

where \(f_I\) is infiltration (mm) and \(P_t\) is precipitation (mm). \(R\) and \(r\) are parameters.

Surface runoff \(R_S\) (mm) is calculated by

\[
R_S = P_I - f_I - R - R_P.
\]

(7)

According to the saturation excess mechanism and spatial variability of watershed characteristics, interflow and groundwater recharge are estimated as linear functions of soil wetness (soil moisture amount divided by soil moisture capacity). Baseflow is simulated based on the linear reservoir assumption, in which the relationship between groundwater storage and outflow is linear. Interflow\(R_t\) (mm), groundwater recharge\(REC\) (mm), baseflow\(RG\) (mm), and total runoff\(TR\) (mm) are determined by

\[
R_t = L_t \left( \frac{SMS_t}{SMSC} \right) f_I
\]

(8)

\[
REC = R_c \left( \frac{SMS_t}{SMSC} \right) \left( f_I - R_t \right)
\]

(9)

\[
RG = K_g \left( GW_t + REC \right)
\]

(10)

\[
TR = R_S + R_t + RG
\]

(11)

where \(L_t, R_c\), and \(K_g\) are efficiencies for interflow, groundwater recharge and baseflow, respectively; \(SMSC\) is the maximum value of soil moisture storage capacity (mm); \(SMS_t\) is actual soil moisture storage (mm); \(GW\) is groundwater storage (mm); \(L_t, R_c, K_g\) and \(SMSC\) are parameters in need of calibration.

Runoff generation from catchments is routed to the outlet by only considering stream channel processes. The Muskingum method (Franchini and Lamberti, 1994) is used for flow routing of river channels. The Muskingum formula is:

\[
Q_{out} = C_1 Q_{in} + C_2 Q_{out-1} + C_1 Q_{out-1} C_2 Q_{out-2}
\]

(12)

where \(C_1 + C_2 + C_3 = 1\). \(Q_{in}\), and \(Q_{out}\) are flows into the river at the beginning and end of a time step (m\(^3\)/s), \(Q_{out-1}\) and \(Q_{out-2}\) are flows out of river at the beginning and end of a time step (m\(^3\)/s), and \(C_1\) and \(C_2\) are two routing parameters in need of calibration.

Because of the difference of runoff generation and concentration characteristics between a reservoir and a natural basin, the reservoir is treated as a separate sub-basin in HIMS model. According to topological and connective relationships in the catchment network, flooding into the reservoir is calculated using runoff generation and flow concentration models in its upper catchment. Water yield of the reservoir is calculated using the runoff generation model and flow out of the reservoir is treated as inflow to the lower catchment (Liu et al., 2006).

Based on water balance theory, the reservoir storage equation is

\[
\frac{dS}{dt} = I - Q
\]

(13)

where \(S\) is reservoir storage (m\(^3\)), \(I\) is flow into the reservoir (m\(^3\)/s) and \(Q\) is flow released from the reservoir (m\(^3\)/s).

Finite difference forms of the equation can simulate multiphase regulation and storage processes.
\[ S_{t+1} - S_t = (k_{t+1} + k_t) \frac{d^2 t}{T^2} - (Q_{t+1} + Q_t) \frac{d^2 t}{T^2} \]  
(14)

The relationship between reservoir storage and outflow can be expressed as

\[ Q = f(S). \]  
(15)

Supposing that flow into the reservoir is known, according to Eqs. (14) and (15), flow release from reservoir can be obtained. Because there is no reservoir in our study area, we do not consider reservoir regulation and storage.

HIMS model inputs include DEM, rainfall and potential evapotranspiration (or maximum and minimum temperature). There are nine parameters requiring calibration. A brief description of calibration parameters is given in Table 2.

### 3.2. Model calibration method

PSO is a heuristic algorithm using speed-position search pattern. Each particle represents a candidate solution, and the superiority of solution is determined by the objective function (Kennedy and Eberhart, 1995). Supposing that the search space is D-dimensional, particle \( i \) can be represented by a D-dimensional vector \( X_i = (x_{i1}, x_{i2}, \ldots, x_{iD}) \). The velocity of this particle can be defined as \( V_i = (v_{i1}, v_{i2}, \ldots, v_{iD}) \). The best previously-visited position of particle \( i \) is taken as its individual best position \( P_i = (p_{i1}, p_{i2}, \ldots, p_{iD}) \). The best position of the entire swarm is taken as the global best position \( G = (g_1, g_2, \ldots, g_D) \). At each step, the particle updates its velocity and position according to the following two equations:

\[ v_{id}^{t+1} = v_{id}^t + a_1 r_1 (p_{id} - x_{id}) + a_2 r_2 (g_d - x_{id}) \]  
(16)

\[ x_{id}^{t+1} = x_{id}^t + v_{id}^{t+1} \]  
(17)

where \( t \) means the current iterative step and \( N \) represents the number of total populations. \( r_1 \) and \( r_2 \) are independently uniformly distributed random variables with range \([0,1]\). \( a_1 \) and \( a_2 \) are positive constant parameters called acceleration coefficients that control the maximum step size. \( a_1 \) is a constriction factor used to limit the velocity.

The influence of velocity on the optimization process can be controlled by adjusting \( a_1 \). A large \( a_1 \) results in a stronger global search ability, and a small one results in a stronger local search ability. Thus, to ensure the convergence of PSO, we adapted the following rule to update the velocity. We call this method self-adaptive dynamic particle swarm optimization (SDPSO). The rule for each particle updating its velocity is described by

\[ a_i = \begin{cases} a_{\text{max}} & \text{if } f_{\text{max}} - f_{\text{max}} \leq f_{\text{max}} - f_{\text{min}}, \\ a_{\text{min}} & \text{if } f_{\text{max}} - f_{\text{max}} > f_{\text{max}} - f_{\text{min}}. \end{cases} \]  
(18)

\[ \begin{align*} &a_1 = (d_1 - d_2) \times (T + d_1) \\ &a_2 = (d_1 - d_2) \times (T + d_2). \end{align*} \]

where \( a_{\text{max}} \) and \( a_{\text{min}} \) are initial and final inertia weight factors; \( f_{\text{max}}, f_{\text{min}}, f_{\text{max}} \) are fitness values that are determined by the current particles; \( d_1 \) and \( d_2 \) are constant factors, and \( d_1 \) is initially larger than \( d_2 \). \( T \) is the maximum generation number.

Because PSO is a one-way information sharing mechanism, it is easy to trap in local optima. Therefore, a master-slave swarms shuffling evolution algorithm based on SDPSO (MSSE-SDPSO) is proposed by introducing the concepts of population division and biological evolution, to maintain the diversity of population and avoid premature convergence. The essence of the method is as follows.

Particles are initialized in feasible region and then assigned to one master swarm and several slave swarms. Each slave swarm independently executes SDPSO. And master swarm evolves itself based on slave swarms, which is realized by incorporating a new dimension for the particle velocity in standard PSO. The resulting equation for manipulation of the master swarm is

\[ v_{id}^{t+1} = v_{id}^t + a_1 r_1 (p_{id} - x_{id}) + a_2 r_2 (g_d - x_{id}) + a_3 (y_i - x_{id}). \]  
(19)

where \( a_1 \) is a positive parameter called the migration coefficient, and \( r_1 \) is an uniform random sequence in the range \([0,1]\). \( y_i \) is the previous best position of all slave swarms. After a certain number of generations, the master swarm and all slave swarms are forced to mix, and points are then reassigned to ensure information sharing (Jiang et al., 2010, 2013). The MSSE-SDPSO strategy is described below and in Fig. 5.

**Step 1:** Initialization. Select \( M \geq 1, N \geq 1 \), where \( M \) is the number of sub-swarms (one master swarm and \( M - 1 \) slave swarms), \( N \) is the number of points in each sub-swarm. Compute the sample size \( S = M \times N \). Sample 5 points to \( x_1, \ldots, x_5 \) in the feasible space. Compute the function value, at each point \( x_i \).

**Step 2:** Ranking. Sort the points in order of increasing function value, and store them in an array \( E = \{ x_i, f_i : i = 1, \ldots, S \} \).

**Step 3:** Partitioning. Partition the points into \( M \) sub-swarms \( A^1, \ldots, A^M \), each containing points \( N \) such that \( A^t = \{ x_{id}^t : f_i < f_{id}, 1 \leq i \leq N, j = 1, \ldots, N \} \), \( k = 1, 2, \ldots, M \).

**Step 4:** Evolution. Evolve each sub-swarm \( A^t \) using SDPSO.

**Step 4.1:** Selection. Choose \( q \geq 1 \) distinct points \( x_{id}^t, y_{id}^t, \ldots, x_{id}^t \) from \( A^t \) to construct a sub-swarm. Better points in \( A^t \) have greater probability of selection; store them in a sub-swarm. Each sub-swarm independently adapts itself according to Eqs. \((16)-(18)\). The master swarm adapts itself according to Eqs. \((17)-(19)\).

**Step 4.4:** Iteration. Iterate by repeating Steps 4.2 and 4.3 \( T \) times, where \( T \) is the maximal iterated generation.

**Step 5:** Shuffling. Replace \( A^1, \ldots, A^M \) into \( E \). Sort \( E \) in order of increasing function value.

**Step 6:** Convergence checking. If the maximum number of generations \( M_{\text{MaxGen}} \) is satisfied, stop. Otherwise, return to Step 3.

MSSE-SDPSO introduces self-adaptive, dynamic, competitive evolution and the hierarchical concept into standard PSO, it can greatly enhance survivability and stability by sharing of information gained independently by each swarm.

### 3.3. Objective function, assessment criteria and calibration process

There are many objective functions that have been used in model calibration. Different objective functions describe the various characteristics of hydrologic processes. To successfully calibrate model parameters using an automatic optimization routine, it is necessary to formulate the calibration objective function comprehensively. Water balance error and degree of fit between measured and simulated discharges. Model parameters were optimized by minimizing values of the objective function (Zhang et al., 1999; Yang, 2002):

\[ f(X) = 1 - \frac{1}{N} \sum_{i=1}^{N} \left( Q_{\text{obs},i} - Q_{\text{sim},i} \right)^2 \left( 1 - \frac{Q_{\text{obs},i} - Q_{\text{sim},i}}{Q_{\text{obs},i}} \right) \]  
(20)

where the variable \( x \) represents parameters in need of calibration, and \( Q_{\text{obs},i} \) and \( Q_{\text{sim},i} \) are measured and simulated discharges, respectively. \( Q_{\text{obs}} \) and \( Q_{\text{sim}} \) are average values of measured and simulated discharges, respectively, and \( N \) is the number of the time series.

The constraint on the variable for the problem is \( X_{\text{min}} \leq X \leq X_{\text{max}} \), where \( X_{\text{min}} \) and \( X_{\text{max}} \) are the lower and upper bounds of model parameters.

Bennett et al. (2013) proposed direct and indirect methods for characterizing performance of environmental models. In hydrologic modeling, there are two commonly used criteria to evaluate models performance. The first criterion used to evaluate model performance was the Nash–Sutcliffe efficiency (NSE) criterion (Nash and Sutcliffe, 1970), defined as

\[ \text{NSE} = 1 - \frac{\sum_{i=1}^{N} (Q_{\text{obs},i} - Q_{\text{sim},i})^2}{\sum_{i=1}^{N} (Q_{\text{obs},i} - Q_{\text{obs}})^2}. \]  
(21)

NSE is always expected to approach unity for accurate simulation of the observed streamflow series.

The second efficiency criterion used was the cumulative relative error (RE) between observed streamflow and the simulated series, defined as

\[ \text{RE} = \frac{Q_{\text{obs}} - Q_{\text{sim}}}{Q_{\text{obs}}} \times 100\%, \]  
(22)
RE is expected to be near zero for accurate simulation of total volume of the observed runoff series.

The general flowchart for the calibration process using MSSE-SDPSO for HIMS hydrological model is presented in Fig. 6. The calibration can be executed via an automatic process, and this process continues until a condition is satisfied. The termination criterion for the iterations is determined according to whether the maximum iteration is reached. To avoid the influence of randomness, the calibration takes several trials and the average of all solutions is selected as the optimal parameter set.

3.4. Model comparison

To test performance of HIMS model for the Luanhe River basin, we compared it with the SimHyd and SMAR models. SimHyd model daily runoff (surface runoff, interflow and baseflow) using daily precipitation and potential evapotranspiration as input data (Chiew et al., 2002; Wang et al., 2006). SMAR model provides daily estimates of surface runoff (overland flow, saturation excess runoff and saturated throughflow from perched groundwater conditions), groundwater discharge, evapotranspiration and leakage from the soil profile for the catchment as a whole (O’Connell et al., 1970; Kachroo and Liang, 1992; Tan and O’Connor, 1996; Tuteja and Cunnane, 1999). Both SimHyd and SMAR include the infiltration excess and saturation excess mechanisms.

3.5. Impacts of model complexity and rainfall resolution on model efficiency

To analyze the impacts of model complexity and rainfall resolution on model efficiency, three forms for HIMS model were examined for the Yixunhe River basin. The first form is a lumped HIMS model with only one gauge data from the center gauge of the catchment. In the other form, precipitation data at all seven gauges are inputted to the lumped HIMS. The last form is a distributed HIMS model. The basin is divided into seven sub-basins and channel of each sub-basin joins at the end. Differences between the lumped and distributed model are attributed to routing procedures.

3.6. Sensitivity analysis

Sensitivity analysis (SA) is defined as the study of how uncertainty in model output can be apportioned to different sources of input (Saltelli et al., 2006; Foscarini et al., 2010). SA methods can be classified as either local or global (Yang, 2011). In local sensitivity analysis, each factor is perturbed in turn while keeping all others fixed at their nominal values (Baroni and Tarantola, 2014). Although local SA is widely used, it is often unable to produce meaningful results (Saltelli et al., 2004). On the other hand, Global SA (GSA) studies the effects of input variations on outputs within the entire allowable ranges of input space. GSA methods range from qualitative screening (Campolongo et al., 2007, 2011; Saltelli and Annoni, 2010) to quantitative techniques based on variance decomposition, in which the Fourier amplitude sensitivity test (FAST) (Cukier et al., 1973), response surface methodology (RSM) (McKay et al., 1979), Morris screening method (Morris, 1991; Ruano et al., 2012; Tocham et al., 2013), and Sobol’ method (Sobol’, 1993) are the most widely investigated (Sun et al., 2012; Dejong et al., 2012; Zheng et al., 2012; Zhan et al., 2013; Zhan and Zhang, 2013; Gameir et al., 2013; King and Perera, 2013). Shin et al. (2013) used the Morris and Sobol methods to analyze model parameter sensitivities and pointed out that the result of applying the Morris method is similar to that of Sobol but there is a slight difference in the ordering of the parameter sensitivities. In this work, we made global sensitivity analysis based on the Morris method. This was first proposed by Morris (1991), which is a screening method based on elementary effects to identify a subset of inputs that have the greatest influence on outputs. The method is designed to work with low computational cost to determine which parameters are (a) negligible, (b) linear and additive, and (c) nonlinear or involved in interactions with other parameters.

Fig. 5. Flowchart of MSSE-SDPSO for HIMS hydrological model calibration.

Fig. 6. Flowchart of MSSE-SDPSO for HIMS hydrological model calibration.
Consider a model for which an output y is a deterministic function of k inputs \(X_1, X_2, \ldots, X_k\) collectively by the k-element row vectors. For a given value of y, the elementary effect of the i-th input factor is defined as:

\[
d_i(x) = \frac{1}{p} \sum_{j=1}^{n} \frac{f(x_{i_j}) - f(x)}{f(x)}
\]

where \(2 = \frac{1}{(p-1)} - \frac{1}{(p-1)}\) and \(p\) is the number of levels. The \(d_i\) calculation is repeated a number of times (y). A typical setting with \(p = 4, J = 4, p = 2, J = 4\) requires a total of \(N = 10 + (k+1)\) simulations, as applied in this case study (Neumann, 2012). Morris proposed two sensitivity indices to analyze the data, the mean (\(\mu_i\)) and standard deviation (\(\sigma_i\)) values of the \(r\) samples of \(d_i\). Campolongo et al. (2007) proposed an improved measure, \(\mu_i^*\) in place of \(\mu_i\), using the following formulas:

\[
\mu_i^* = \frac{1}{r} \sum_{j=1}^{r} d(x_{i_j})
\]

\[
\sigma_i = \left(\frac{1}{r} \sum_{j=1}^{r} \frac{d(x_{i_j}) - \mu_i^*}{d(x_{i_j})}\right)^2
\]

The \(\mu_i^*\) and \(\sigma_i\) indicate the influence of each parameter on the target function. A high \(\mu_i^*\) value implies that a parameter has an important influence on the target function, and a high \(\sigma_i\) value implies that a parameter has strong interactions with other parameters or that the effect of the parameter is nonlinear (Morris, 1991; Van Grientsven et al., 2000; Neumann, 2012).

4. Results and analysis

4.1. Test of numerical simulation

In this section, four nonlinear benchmark functions that are commonly used in the literature (Trelea, 2003) are executed. Table 3 summarizes the functions, number of dimensions, admissible range of the variable, and optimum and goal values. All functions are designed to have minima at the region.

Table 4

| Parameter  | Description | Value |
|------------|-------------|-------|
| S          | The population size for all five algorithms | 10    |
| \(d_i\)    | Constant factor for MSSE-SPSO and MSSE-SDPSO | 2.8   |
| \(d_i^{\text{max}}\) | Constant factor for MSSE-SPSO and MSSE-SDPSO | 1.2   |
| \(\mu_i^{\text{max}}\) | Initial inertia weight factor for PSO, MSSE-PSO, MSSE-SPSO and MSSE-SDPSO | 0.9   |
| \(\mu_i^{\text{min}}\) | Final inertia weight factor for PSO, MSSE-PSO, MSSE-SPSO and MSSE-SDPSO | 0.4   |
| \(\beta\)  | Shrink factor for MSSE-PSPO | 0.001 |
| \(\alpha_1\) | Acceleration coefficient only for PSO | 2.0   |
| \(\alpha_2\) | Acceleration coefficient only for PSO | 2.0   |
| \(\alpha_3\) | Acceleration coefficient only for MSSE-PSO, MSSE-SPSO and MSSE-SDPSO | 2.05  |
| \(\alpha_4\) | Acceleration coefficient only for MSSE-PSO, MSSE-SPSO and MSSE-SDPSO | 2.05  |
| \(\alpha_5\) | Acceleration coefficient only for MSSE-PSO, MSSE-SPSO and MSSE-SDPSO | 2.05  |
| \(M\)      | The number of sub-swarms for SCE-UA | 4     |
| \(N\)      | The number of points in each sub-swarm for SCE-UA, MSSE-PSO, MSSE-SPSO and MSSE-SDPSO | 10    |
| \(T\)      | The maximal iteration of each sub-swarm for SCE-UA, MSSE-PSO, MSSE-SPSO and MSSE-SDPSO | 30    |
| MaxGen     | The total iterations for all four algorithms | 3000  |

To validate the feasibility of the proposed algorithm, MSSE-SDPSO was compared with SCE-UA, PSO MSSE-PSO and MSSE-SPSO for the same test set; SCE-UA is from Duan et al. (1992, 1994) and PSO is implemented with classic PSO, MSSE-PSO strategies from Jiang et al. (2010) and MSSE-SPSO strategies from Jiang et al. (2013). Relevant experimental parameters of the four algorithms for this test system are shown in Table 4. The termination criterion for the iterations is determined according to whether the maximum iteration is reached. All four algorithms required 10 trials to obtain the optimal solution for avoiding the influence of randomness. The performances of SCE-UA, PSO, MSSE-PSO, MSSE-SPSO and MSSE-SDPSO are shown in Table 5. This table shows that the average logarithmic objective values obtained by MSSE-SDPSO are smaller than that of SCE-UA, PSO, MSSE-PSO and MSSE-SPSO. Fig. 7 shows the average logarithmic objective value of the best particles found during 3000 generations by 40 particles for each function. These figures display the decreasing process of average objective value with iterations. It can be observed that for SCE-UA and PSO, the average logarithmic objective values of the four benchmark problems decrease slowly throughout optimization processes, which imply that they have poor exploitation ability. For MSSE-SDPSO, the average logarithmic objective values of all benchmark problems decrease quickly throughout optimization processes and the convergence speed is faster than SCE-UA, PSO, MSSE-PSO and MSSE-SPSO, which mean it can find optimal solutions quickly. This means that MSSE-SDPSO has better optimization performance than the other algorithms, so the proposed algorithm is efficient and effective.

4.2. Comparison of model results

A 10-year period (1 January 1995 to 31 December 2004) was used for the calibration, for which daily data of precipitation, maximum and minimum temperature, and catchment streamflow.
were available. We also used validation data covering 1 January 2005 to 31 December 2008.

Statistical and visual comparisons of observed and simulated streamflow were performed to evaluate performances of the SimHyd, SMAR and HIMS models. Table 6 gives the HIMS model optimal parameters and Table 7 summarizes statistics of simulations by the three models for the Liuhe, Baohe, Sahe, Laoniuhe, Wuliehe, Xingzhouhe, Yimatuhe and Yixunhe Rivers in the calibration and validation periods. Overall, HIMS model predictions of streamflow were superior to those of SimHyd and SMAR models. Streamflow simulations for the Liuhe, Baohe, Sahe, Laoniuhe and Wuliehe River sub-basins were successful, with NSEs 0.71–0.90 and REs –6.90% to 9.71%. However, HIMS model performances for the Xingzhouhe and Yimatuhe sub-basins were poor, as shown by NSEs 0.19–0.66. It should be emphasized that in the calibration period for the Yimatuhe River, RE reached 22.93%, outside the allowable range (no more than 20%). The main reason could be that we do not consider the influence of rubber dams along the Xingzhouhe and Yimatuhe rivers in the current HIMS model, or these observation data are insufficient or short in time series. Observed and simulated streamflow hydrographs from HIMS, SimHyd and SMAR models for the verification period 2005 are shown in Fig. 8. It is seen that simulated peak streamflows are lower than observed, but the peak streamflow from HIMS model simulation is close to the observed. As a lumped hydrological model, SimHyd model considers the watershed as a whole and usually does not account for spatial variation of geographic elements. Moreover, there is no flow concentration calculation (Chiew and McMahon, 2002), which affects simulation accuracy. SMAR model is also a lumped hydrologic model. It estimates runoff generation from two sources, surface runoff and baseflow. However, in

Table 6
HIMS model optimal parameters for different sub-basins.

| Sub-basins      | SMSC | R     | r     | La    | Rc    | C    | Kb    | C1   | C2   |
|-----------------|------|-------|-------|-------|-------|------|-------|------|------|
| Liuhe River     | 542.58 | 1.19  | 0.48  | 0.36  | 0.15  | 5.37 | 0.31  | 0.05 | 0.41 |
| Baohe River     | 885.71 | 1.15  | 0.52  | 0.61  | 0.58  | 6.18 | 0.50  | 0.78 | 0.66 |
| Sahe River      | 992.86 | 1.44  | 0.36  | 0.68  | 0.02  | 9.36 | 0.03  | 0.52 | 0.39 |
| Laoniuhe River  | 560.00 | 1.46  | 0.61  | 0.87  | 0.39  | 9.24 | 0.70  | 0.78 | 0.54 |
| Wuliehe River   | 127.33 | 1.50  | 0.89  | 0.20  | 0.52  | 7.20 | 0.02  | 0.48 | 0.64 |
| Xingzhouhe River| 254.73 | 1.41  | 0.71  | 0.10  | 0.60  | 9.72 | 0.08  | 0.71 | 0.17 |
| Yimatuhe River  | 107.13 | 1.62  | 0.99  | 0.43  | 0.66  | 9.50 | 0.01  | 0.67 | 0.22 |
| Yixunhe River   | 133.16 | 1.79  | 0.99  | 0.13  | 0.17  | 4.77 | 0.01  | 0.36 | 0.27 |

Table 7
Simulation results for all sub-basins.

| Sub-basins      | HIMS model | SimHyd model | SMAR model |
|-----------------|------------|--------------|------------|
|                 | NSE (%) RE (%) | NSE (%) RE (%) | NSE (%) RE (%) |
| Calibration     |            |              |            |
| Liuhe River     | 0.72 0.60 | 22.5 0.56    | 12.67      |
| Baohe River     | 0.74 0.71 | 39.7 0.78    | 35.01      |
| Sahe River      | 0.86 0.61 | 15.6 0.58    | 4.46       |
| Laoniuhe River  | 0.84 0.64 | 24.68 0.58   | 40.11      |
| Wuliehe River   | 0.79 0.63 | 7.71 0.66    | 12.67      |
| Xingzhouhe River| 0.66 0.50 | 20.4 0.49    | 23.27      |
| Yimatuhe River  | 0.56 0.32 | 7.98 0.35    | 32.49      |
| Yixunhe River   | 0.71 0.37 | 11.9 0.37    | 33.2       |
| Verification    |            |              |            |
| Liuhe River     | 0.80 0.68 | 10.2 0.60    | 26.44      |
| Baohe River     | 0.79 0.70 | 57.9 0.61    | 57.69      |
| Sahe River      | 0.80 0.74 | 23.6 0.76    | 18.18      |
| Laoniuhe River  | 0.90 0.78 | 3.62 0.74    | 40.37      |
| Wuliehe River   | 0.78 0.66 | 60.6 0.62    | 7.91       |
| Xingzhouhe River| 0.19 0.11 | –2.25 0.17   | 11.01      |
| Yimatuhe River  | 0.40 0.29 | –5.55 0.25   | 52.02      |
| Yixunhe River   | 0.73 0.63 | 1.08 0.60    | 31.87      |
Fig. 8. Observed and simulated streamflow hydrographs in 2005 for HIMS, SimHyd and SMAR models.
the Luanhe River basin, the main runoff form is interflow (Shi, 2013), so SMAR model cannot simulate streamflow well.

To compare model capabilities for simulating the dynamics of streamflow series, simulated streamflows were analyzed and correlated with observed values using a linear regression equation. A scatter plot and regression results are shown in Fig. 9. This figure shows that: (1) HIMS model predictions correlated well with observed values, giving squared correlation coefficient $R^2$ between 0.554 and 0.857 at a 0.01 significant level. (2) Comparison of values of slope and intercept indicates that differences among the models were not large. HIMS model had the smallest bias of regression slope and intercept.

Generally, HIMS model performed better than SimHyd or SMAR models, and is suitable for simulating streamflow in semiarid regions.

4.3. Model complexity and rainfall resolution on model efficiency

Nine optimal parameters for the three model combinations (lumped model with single-gauge data, lumped model with all-gauge data, and distributed model) were determined by the procedure described above. Fig. 10 compares observed and simulated streamflow between single-gauge and all-gauge rainfall information, for the lumped model and between the lumped and distributed models. Fig. 10a shows that for the lumped model, all-gauge information gave a better fit than single-gauge. Fig. 10b shows that the distributed model outperformed the lumped model.

The advantage of the detailed rainfall information probably results from greater accuracy of the spatially averaged rain-intensity data with computation from all gauges, relative to that from a single gauge. Table 8 shows that NSE increased from 0.53 to 0.71 in the calibration period and 0.64 to 0.73 in the verification period, for the lumped model with single-gauge data versus with all-gauge data. This demonstrates great improvement for the lumped model with all-gauge data. The spatial distribution of rainfall is important to the simulation, and rainfall biases have a great effect on streamflow prediction. The disadvantage of the lumped model with single-gauge data is underestimation of peak discharge, which contributes a relatively large part of the error and produces a generally lower simulated streamflow.

The distributed model benefits from the more detailed information on rainfall spatial distribution. Table 8 shows that NSE increased from 0.71 to 0.78 in the calibration period and 0.73 to 0.75 in the verification period, for the lumped versus distributed model using all-gauge data. Thus, the distributed model performed better than the lumped one. This is because in the latter model, the spatial variation of hydrologic elements is considered. This better represents the real physical process of rainfall runoff formation, thereby improving the accuracy of hydrologic simulation.

4.4. Sensitivity analysis

The Morris screening measures can be shown graphically by screening plots, which have as the x- and y-axes the modified means $\mu^*$ and standard deviations $\sigma_0$, respectively (Zhan et al., 2013). Tables 9 and 10 present the results of average $\mu^*$ and associated ranking of the eight sub-basins for NSE and RE objective functions respectively, considering the nine HIMS model parameters in Table 2. Figs. 11 and 12 show global sensitivity analysis results of the eight sub-basins for the objective function (21) and (22) respectively, and reveal how strongly an output variable was influenced by the parameters. From these Tables and Figures, we can see that $R$, $r$, and $C$ were the most sensitive parameters for all sub-basins except the Sahe river, and the others were relatively less sensitive for the two objective functions. The rankings for the different objective function responses are slightly different, e.g., $R$ is the most sensitive parameter for the Nash-Sutcliffe efficiency NSE and $r$ is the most sensitive parameter for the cumulative relative error RE. These results agree with the conclusions of other studies (e.g., Liu et al., 2013). Evapotranspiration affects runoff formation and is the dominant component of total runoff calculation, especially in a semiarid region where rainfall is less than evaporation capacity. Thus, evapotranspiration parameter $C$ is a sensitive one. Soil water content is very important to runoff formation and is a threshold controlling the distribution of surface runoff, interflow and baseflow. Infiltration is the core equation of the HIMS model. Increasing $R$ and $r$ can increase soil water content, subsequently increasing interflow and decreasing surface runoff, thereby affecting the peak flow and changing the shape of the flow curve. Therefore, $R$ and $r$ are also important parameters that must be calibrated carefully.

5. Discussion

Semiarid regions pose unique problems for hydrological modeling. Runoff generation in such regions is a complex process, dominated by infiltration excess and saturated excess mechanisms (Wheater, 2002; Greenbaum et al., 2006). Thus, rainfall-runoff modeling of semiarid catchments is a challenging task (Pilgrim et al., 1988; Ye et al., 1997; McIntyre et al., 2007; Wheater et al., 2008). We applied a simple hydrological model to simulate streamflow in a semiarid catchment. The degree of fit between observed and computed runoff hydrographs was assessed for the Luanhe River basin. Performances of this model in terms of objective criteria prediction for streamflow were superior to SimHyd and SMAR models (Table 3). We examined the impact of spatial rainfall resolution on the efficiency of HIMS model in modeling streamflow. Model performance was found to be affected by rainfall spatial resolution (Table 4 and Fig. 10a). NSE increased from 0.53 to 0.71 in the calibration period and 0.64 to 0.73 in the verification period, for the lumped model with single-gauge data versus with all-gauge data. Our study revealed that model performance can improve significantly with use of a lumped model with all-gauge data, versus data from a single gauge. The reason for this advantage of all-gauge data lies with the less biased rainfall input to the model, as compared to data from a single gauge. We also analyzed performance of lumped and distributed HIMS models, with the results indicating that the distributed model performed better than the lumped one (Table 4 and Fig. 10b). NSE increased from 0.71 to 0.78 and 0.73 to 0.75 in the verification period. The difference between the distributed and lumped models during calibration was also large compared to the difference between one- and seven-gauge data. However, many researchers have shown that complex models often did not provide more accurate runoff predictions than simpler and less detailed models (Michaoud and Sorooshian, 1994; Al-Qurashi et al., 2008). Bahat et al. (2009) indicated a lumped model performed better than a distributed model in a rainfall-runoff modeling study of a small, hyper-arid catchment in Israel. But the Yixunhe River sub-basin is very large and heterogeneous, so a set of model parameters cannot reflect the non-uniformity of the watershed characteristics, which degrades the performance of lumped hydrological models. Therefore, the distributed HIMS model performance significantly better than the lumped HIMS model.

Successful application of a hydrological model depends on how well its parameters are selected. This selection is determined by many factors, such as climate, soil, vegetation, terrain and others. Almost all hydrological models require calibration, and it has been long recognized that this is not a straightforward process (Sorooshian et al., 1993). In this paper, we proposed a master-slave swarms shuffling evolution algorithm based on self-adaptive...
Fig. 9. Scatter plot and regression equations of observed streamflows versus those simulated by HIMS, SIMHYD and SMAR models.
Y. Jiang et al. / Environmental Modelling & Software 67 (2015) 72–88

Fig. 10. Comparison of observed and simulated streamflows for the lumped model with all-gauge data versus single-gauge data (a), and the distributed versus lumped model (b).

dynamic particle swarm optimization (MSSE-SDPSO) to derive optimal model parameters. Optimal parameter values of HIMS model for different sub-basins are listed in Table 6. We see from that table that the SMSC value is largest in the Sahe River sub-basin and smallest in the Yimatuhe River sub-basin. SMSC reflects water-holding capacity of unsaturated soil layer. Influences on SMSC include soil texture and structure, forest coverage, precipitation, and others. In a basin where forest coverage is high and precipitation heavy, the determined SMSC will be large. Because precipitation is heaviest and forest coverage is greatest, the SMSC is a maximum in the Sahe River sub-basin. Although forest coverage in the Yimatuhe River sub-basin is larger than that in the Yixunhe river sub-basin, the maximum evaporation parameter C in the Yixunhe River calibration was determined larger than those in the Yixunhe. These results agreed with conclusions from other studies (e.g., Liu et al., 2008), in which R and r in loose sand or forest areas with heavy sticking under high antecedent moisture were greater than those in clay, shallow underground water, and rocky hilly regions under low antecedent moisture. Actual evaporation is related to soil water storage, forest coverage, temperature and precipitation. If temperature is low, then water molecules move slowly, so actual evaporation in a watershed is small. Under certain conditions, enhanced forest coverage increases actual evaporation (Liu, 1978). Because precipitation and forest coverage are greatest in the Sahe River sub-basin, the maximum evaporation parameter C was determined. In contrast, in the Yimatuhe River sub-basin, because of the least precipitation and lowest temperature, C was a minimum. In practice, we could adjust parameters, according to the relationship of parameter selection and watershed characteristics, to improve the performance of streamflow simulation.

Multi-objective global optimization (Cheng et al., 2002; Gill et al., 2006; Rajesh and Michael, 2008; Kraue et al., 2011) has been developed because any single-objective function, no matter how carefully chosen, may not adequately measure the ways in which a model fails to match important characteristics of observed data (Li et al., 2010). Considerable improvement of model performance may be achieved through parameters calibrated by such optimization. It is not known whether HIMS model prediction performance would be improved by enhanced calibration—verification strategies for parameters. Therefore, future studies should investigate calibration of HIMS model parameters by the multi-objective method, as well as appropriate algorithms for multi-objective optimization.

Table 9
Simulation results for lumped model with single-gauge and all-gauge data, and for distributed model in Yixunhe River.

| Yixunhe River | Calibration | Verification |
|---------------|-------------|--------------|
|               | NSE         | RE (%)       |
| Lumped model with one-gauge data | 0.53 | 7.83 | 0.64 | 3.73 |
| Lumped model with all-gauge data | 0.71 | 6.18 | 0.73 | 3.21 |
| Distributed model | 0.78 | 5.30 | 0.75 | 3.11 |

Table 8
Comparison of observed and simulated streamflows for the lumped model with all-gauge data versus single-gauge data (a), and the distributed versus lumped model (b).

Comparison of observed and simulated streamflows for the lumped model with all-gauge data versus single-gauge data (a), and the distributed versus lumped model (b).
Sensitivity analysis of model parameters using the Morris screening method was formulated in the present study. We concluded that $R$, $r$ and $C$ were the most sensitive parameters. Liu et al. (2013) examined HIMS model parameter uncertainty using the generalized likelihood uncertainty estimation (GLUE) method for three catchments in Australia, finding that HIMS parameters were all sensitive. We were unable to determine whether the other six parameters were sensitive. The main limitations of the Morris screening method are that it cannot provide accurate quantitative estimation of a factor’s contribution to output variability, and that it cannot distinguish the nonlinearity of a factor from its interaction with other factors, i.e., the nonlinearities and parameter interactions are confounded (Yang, 2011). Therefore, future studies should investigate which parameters are sensitive and which are not. Model accuracy not only depends on rationality of structure and reliability of data, but also on effectiveness of parameters selection. Future studies should also focus on analyzing uncertainties of model parameters, structure and rainfall input, and investigating which factor dominates uncertainty in runoff prediction, input data, model structure or parameter?

Table 10

| Parameters | Liuhe River | Baohe River | Sahe River | Laoniuhe River | Wuliehe River | Xingzhouhe River | Yimatuhe River | Yixunhe River |
|------------|-------------|-------------|------------|----------------|---------------|------------------|----------------|---------------|
| $\mu^*$ | 0.01 | 2.30 | 1.01 | 1.22 | 0.94 | 1.49 | 2.10 | 3.09 |
| R | 2.44 | 2.00 | 1.35 | 4.07 | 3.09 | 3.56 | 2.62 | 7.59 |
| $r$ | 0.50 | 0.53 | 0.52 | 0.54 | 0.57 | 0.79 | 0.81 | 0.90 |
| La | 0.45 | 0.45 | 0.45 | 0.45 | 0.45 | 0.45 | 0.45 | 0.45 |
| $C_l$ | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 |
| $C_2$ | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |

Fig. 11. Morris sensitivity index for NSE for HIMS model for eight sub-basins.
6. Conclusions

This study emphasized the following aspects of modeling: (1) calibration-validation strategies; (2) suitability of HIMS model for semiarid basins; (3) effect of model complexity and rainfall resolution; (4) analysis of model parameter sensitivities. Implications of the present results are summarized below.

(1) The proposed algorithm MSSE-SDPSO has better optimization performance than SCE-UA, PSO, MSSE-PSO or MSSE-SPSO. (2) Streamflow in the study catchment was well simulated by the HIMS rainfall-runoff model compared with the SimHyd or SMAR models. HIMS model is suitable for simulating streamflow in semiarid regions. (3) The use of rain data from several gauges improved runoff prediction, and a distributed HIMS model performed better than a lumped HIMS for the Luanhe River basin. (4) In the Luanhe River basin, R, r and C were the most sensitive parameters, and SMSC, La, Rc, Kb, C1, C2 were the second most sensitive for the objective function.

Acknowledgments

We are grateful to Prof. Anthony Jakeman and Serena Hamilton for helpful suggestions, to Dr. Yaomin Qin for his help in mapping. Funding was supported by the National Natural Science Foundation of China (Nos. 50809004, 51279006, 51479003), the Key Project of the National Natural Science Foundation of China (No. 41330529) and the National Major Science and Technology Projects for Water Pollution Control and Management (Nos. 2012ZX07203-002, 2012ZX07203-003). The used HIMS model was programmed and provided for their study by Prof. Changming Liu, Zhonggen Wang and Hongxing Zheng (Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, China).

References

Abdulla, F.A., Lettenmaier, D.P., 1997. Development of regional parameter estimation equations for a macroscale hydrologic model. J. Hydrol. 197, 230–257.
Ahrens, C.D., 2003. Meteorology Today: an Introduction to Weather, Climate, and the Environment. Thomson Learning, Brooks/Cole.
Al-Qurashi, A., McIntyre, N., Wheater, H., Unkrieh, C., 2008. Application of the Kineros2 rainfall-runoff model to an arid catchment in Oman. J. Hydrol. 355, 91–105.
Andrews, F.T., Croke, B.F.W., Jakeman, A.J., 2011. An open software environment for hydrological model assessment and development. Environ. Model. Softw. 26, 1171–1185.
Bahat, Y., Grodek, T., Lekach, J., Morin, E., 2009. Rainfall-runoff modeling in a small hyper-arid catchment. J. Hydrol. 373, 204–217.
Bao, C.X., Zhang, J.Y., Wang, G.Q., Fu, G.B., He, E.M., Yan, X.L., Jin, J.L., Liu, Y.L., Zhang, A.J., 2012. Attribution for decreasing streamflow of the Haihe River basin, northern China: climate variability or human activities? J. Hydrol. 460 (46), 117–129.
Baroni, G., Tarantola, S., 2014. A General Probabilistic Framework for uncertainty and global sensitivity analysis of deterministic models: a hydrological case study. Environ. Model. Softw. 51, 26–34.
Bennett, N.D., Croke, B.F.W., Guirguis, G., Floumaume, J.H.A., Hamilton, S.H., Jakeman, A.J., Marsili-Libelli, S., Newham, L.T.H., Norton, J.P., Perrin, C., Pierce, S.A., Robson, B., Seppelt, R., Voinov, A.A., Fath, B.D., Andreassian, V., 2013. Characterising performance of environmental models. Environ. Model. Softw. 40, 1–20.
Beven, K.J., Freer, J., 2001. A dynamic TOPMODEL. Hydrol. Process. 15, 1993–2011.
Campolongo, F., Cariboni, J., Saltelli, A., 2007. An effective screening design for sensitivity analysis of large models. Environ. Model. Softw. 22, 1509–1518.
Campolongo, F., Saltelli, A., J. 2006. From screening to quantitative sensitivity analysis. A unified approach. Comput. Phys. Commun. 182, 928–988.
Candel, A., Noto, L.V., Aronica, G., 2005. Influence of surface roughness in hydrological response of semiarid catchments. J. Hydrool. 313, 119–131.
Chatterjee, A., Sainey, P., 2008. Nonlinear inertia weight variation for dynamic adaptation in particle swarm optimization. Comput. Oper. Res. 35, 859–871.
Chau, K.W., 2007. A split-step particle swarm optimization algorithm in river stage forecasting. J. Hydrool. 346 (3–4), 131–135.
Chen, Z.F., Zhao, Y., Qiao, J., Zhang, Q., Zhu, Y.E., Xu, C.H., 2009. Retrospection of recent 30-year changes in the process of soil wind erosion in the Luhan River Source Area of North China using Cesium-137. Appl. Radiat. Isot. 67, 1785–1789.
Cheng, C.T., Ou, C.P., Chau, K.W., 2006. Combining a fuzzy optimal model with a genetic algorithm to solve multi-objective rainfall-runoff model calibration. J. Hydro. 268, 72–86.
Chiew, F.H.S., McMahon, T.A., 2002. Modelling the impacts of climate change on Australian streamflow. Hydrool. Process. 16 (5), 1233–1245.
Chiew, F.H.S., Peel, M.C., Western, A.W., 2002. Application and testing of the simple rainfall-runoff model SIMHYD. In: Singh, V.P., Frevert, D.K. (Eds.), Mathematical Models of Small Watershed Hydrology and Applications. Water Resources Publication, Littleton, Colorado, pp. 335–364.
Chu, W., Gao, X.G., Sorooshian, S., 2010. Improving the shuffled complex evolution scheme for optimization of complex nonlinear hydrological systems: application to the calibration of the Sacramento soil-moisture accounting model. Water Resour. Res. 46, W09530. http://dx.doi.org/10.1029/2009WR008222.
Coles, N.J., Hargreaves, G.H., Samani, Z.A., 1985. Reference crop evapotranspiration from temperature. Appl. Eng. Agric. 1 (1), 96–99.
Greenbaum, N., Ben-Zvi, A., Haviv, I., Enzel, Y., 2005. Hydrology and paleohydrology of the Dead Sea tributaries. In: Enzel, Y., Agnon, A., Stein, M. (Eds.), New Frontiers in Dead Sea Paleoenvironmental Research, Geological Society of America Special Paper, vol. 401, pp. 63–93. http://dx.doi.org/10.1130/2006.2401(05).
Halwatura, D., Najim, M.M.M., 2013. Application of the HECS-HMS model for runoff simulation in a tropical catchment. Environ. Model. Softw. 46, 155–162.
Hargreaves, G.H., Samani, Z.A., 1985. Reference crop evapotranspiration from temperature. Appl. Eng. Agric. 1 (1), 96–99.
Hernandez, A., Miller, C., Goodrich, D.C., Goiff, B.F., Kepner, W.G., Edmonds, C.M., Jone, B.K., 2000. Modeling runoff response to land cover and rainfall spatial variability in semi-arid watersheds. Environ. Monit. Assess. 64, 258–298.
Hu, C.H., Guo, S.L., Xiong, L.H., Peng, D.Z., 2005. A modified xiaojianjiang model and its application in northern China. Nord. Hydrool. 36 (2), 175–192.
Jakovac, A.J., Hornberger, G.M., 1993. How much complexity is warranted in a rainfall-runoff model? Water Resour. Res. 29 (8), 2637–2649.
Jiang, Y., Hu, T.S., Guo, S.L., Wu, X.N., Zeng, Z.X., 2006. Application of particle swarm optimization to parameter calibration of Xin-anjiang model. J. Wuhan. Univ. Eng. Sci. 39, 14–17 (in Chinese).
Jiang, Y., Hu, T.S., Huang, C.C., Wu, X.N., 2007. An improved particle swarm optimization algorithm. Appl. Math. Comput. 193, 231–239.
Jiang, Y., Liu, C.M., Huang, C.C., Wu, X.N., 2010. Improved particle swarm algorithm for hydrological parameter optimization. Appl. Math. Comput. 217, 3207–3215.
Jiang, Y., Li, X.Y., Huang, C.C., 2013. Automatic calibration a hydrological model using a master-slaves swarm shuffling evolution algorithm based on self-adaptive particle swarm optimization. Expert Syst. Appl. 40, 752–757.
Kachroo, R.K., Liang, G.C., 1992. River flow forecasting. Part 2. Algebraic development of linear Modelling techniques. J. Hydrool. 133, 17–40.
Kennedy, J., Eberhart, R.C., 1995. Particle swarm optimization. In: Proceedings of IEEE International Conference on Neural Networks, Piscataway, NJ, 1995, vol. 4, pp. 1942–1948.
King, D.M., Perera, B.J.C., 2013. Morris method of sensitivity analysis applied to assess the importance of input variables on urban water supply yield—a case study. J. Hydro. 477, 147–156.
Kumar, A.R.S., Sudheer, K.P., Jain, S.K., Agarwal, P.K., 2005. Rainfall-runoff modelling using artificial neural networks: comparison of network types. Hydrool. Process. 19, 1297–1304.
Kuo, K.K., Harun, S., Shamsuddin, S.M., 2010. Particle swarm optimization feed-forward neural network for modeling runoff. Int. J. Environ. Sci Technol. 7 (1), 67–78.
Kraus, T., Cullmann, J., Saile, P., Schmitz, G.H., 2011. Robust multi-objective calibration strategies-chances for improving flood forecasting. Hydrool. Earth Syst. Sci. 15, 3693–3741.
Li, J., Liu, C.M., Wang, Z.G., et al., 2014. Two universal runoff yield models: SCS vs. ICM. Acta Geogr. Sin. 69 (7), 926–932 (in Chinese).
Li, J.Z., Feng, P., 2007. Runoff variations in the Luhan River Basin during 1956–2002. J. Geogr. Sci. 33, 319–350.
Li, J.Z., Feng, P., 2009. Trend analysis of runoff generation characteristics of Luhan River Basin. J. Arid Land Resour. Environ. 23 (4), 79–85 (in Chinese).
Li, X.Y., Wheeler, D.E., Jordan, T.E., 2010. Watershed model calibration using multi-objective optimization and multi-site averaging. J. Hydrool. 380, 277–288.
Liu, C.M., 1978. Primary analysis of the effects of forests on the annual runoff in the Loess Plateau. J. Geogr. Sci. 33 (2), 112–125 (in Chinese).
Liu, C.M., Wang, C.T., 1980. The estimate of small-watershed peak flows in China. Water Resour. Res. 16 (5), 881–886.
Liu, C.M., Zheng, X.H., Wang, Z.G., et al., 2006. Distributed Simulation of Catchment Runoff and Water Quality. Science Press, Beijing.
Liu, C.M., Wang, Z.G., Zheng, H., Wu, X.F., 2008. Development and application of HIMS system. Sci. China (E) 38 (3), 350–360.
Liu, C.M., Wang, Z.G., Yang, S.T., Zheng, H.X., 2010a. Research progress of water cycle integrated simulation system (HIMS). Water Resour. Dev. Res. 8 (3), 5–15 (in Chinese).
Liu, C.M., Zheng, X.H., Wang, Z.G., Yang, S.T., 2010b. Multi-Scale integrated simulation of hydrological processes using HIMS with verified case studies. J. Beijing Univ. Nat. Univ. 46 (3), 268–273 (in Chinese).
Liu, L.F., Liu, C.M., Wang, Z.G., Jiang, Y., Zhang, Y.Q., Sang, Y.F., Wang, H., 2013. Parameter uncertainty of HIMS model and its influence factor analysis. Prog. Geog. 32 (4), 532–537 (in Chinese).
MartoAnen-Mena, M., Albaladejo, J., Castillo, V.M., 1998. Factors influencing surface runoff generation in a Mediterranean semi-arid environment Chivaco water shed, 5 E Spain. Hydrool. Process. 12, 741–754.
McMichael, C.E., Hope, A.S., Loaiciga, H.G., 2006. Distributed hydrological modelling and water quality in sewers. Water Res. 47, 4601–4611.
Miller, C.D., Goodrich, D.C., Goiff, B.F., Kepner, W.G., Edmonds, C.M., Jone, B.K., 2000. Modeling runoff response to land cover and rainfall spatial variability in semi-arid watersheds. Environ. Monit. Assess. 64, 258–298.
Naoum, S., Tsanis, I.K., 2004. Ranking spatial interpolation techniques using a GIS-based DSS. Int. J. Glob. NEST 6 (1), 1–20.
