

1 **Quantifying dynamic sensitivity of optimization algorithm parameters**  
2 **to improve hydrological model calibration**

3

4 **Wei Qi<sup>a,b</sup>, Chi Zhang<sup>a\*</sup>, Guangtao Fu<sup>b</sup> and Huicheng Zhou<sup>a</sup>**

5 <sup>a</sup> School of Hydraulic Engineering, Dalian University of Technology, Dalian 116024, China

6 <sup>b</sup> Centre for Water Systems, College of Engineering, Mathematics and Physical Sciences,  
7 University of Exeter, North Park Road, Harrison Building, Exeter EX4 4QF, UK

8 \* Corresponding author: email: czhang@dlut.edu.cn; Tel: +86 411 84708517; Fax: +86 411  
9 84708517

10

11 **Abstract** It is widely recognized that optimization algorithm parameters have significant  
12 impacts on algorithm performance, but quantifying the influence is very complex and  
13 difficult due to high computational demands and dynamic nature of search parameters. The  
14 overall aim of this paper is to develop a global sensitivity analysis based framework to  
15 dynamically quantify the individual and interactive influence of algorithm parameters on  
16 algorithm performance. A variance decomposition sensitivity analysis method, Analysis of  
17 Variance (ANOVA), is used for sensitivity quantification, because it is capable of handling  
18 small samples and more computationally efficient compared with other approaches. The  
19 Shuffled Complex Evolution method developed at the University of Arizona algorithm  
20 (SCE-UA) is selected as an optimization algorithm for investigation, and two criteria, i.e.,  
21 convergence speed and success rate, are used to measure the performance of SCE-UA.  
22 Results show the proposed framework can effectively reveal the dynamic sensitivity of  
23 algorithm parameters in the search processes, including individual influences of parameters  
24 and their interactive impacts. Interactions between algorithm parameters have significant  
25 impacts on SCE-UA performance, which has not been reported in previous research. The

26 proposed framework provides a means to understand the dynamics of algorithm parameter  
27 influence, and highlights the significance of considering interactive parameter influence to  
28 improve algorithm performance in the search processes.

29

30 **Keywords** Algorithm; Optimization; SCE-UA; Sensitivity; TOPMODEL; Variance  
31 decomposition

32

### 33 **1 Introduction**

34 Many optimization algorithms have been proposed to solve hydrological model optimization  
35 problems, such as the Shuffled Complex Evolution algorithm developed at the University of  
36 Arizona (SCE-UA) (Duan et al., 1992; Duan et al., 1993; Duan et al., 1994), various Genetic  
37 algorithms (Deb et al., 2002; Kollat and Reed, 2006; Tang et al., 2006; Fu et al., 2012), and  
38 the dynamically dimensioned search algorithm (Tolson and Shoemaker, 2007; Tolson et al.,  
39 2009; Asadzadeh and Tolson, 2013). Many studies have been carried out to investigate the  
40 strengths and weaknesses of various algorithms, because algorithm performance is of  
41 significant concern to users (Duan et al., 1992; Duan et al., 1993; Sorooshian et al., 1993;  
42 Bäck, 1996; Thyer et al., 1999; Kollat and Reed, 2006; Tolson and Shoemaker, 2007; Zhang  
43 et al., 2008; van Werkhoven et al., 2009; Wang et al., 2010; Fu et al., 2012; Arsenault et al.,  
44 2014; Chao et al., 2015; Qi et al., 2015).

45

46 It is widely recognized that algorithm parameters have a significant influence on algorithm  
47 performance, but quantifying the influence is very complex and difficult due to high  
48 computational demands and dynamic nature of search parameters (Giorgos et al., 2015).  
49 Many optimization applications use trial and error to determine parameter values, or simply  
50 use default parameter values without investigating their influence on algorithm performance

51 (Deb et al., 2002; Tolson and Shoemaker, 2007). However, attempts have been made to find  
52 optimal parameter combinations. For example, Duan et al. (1994) analyzed the performance  
53 of SCE-UA under different parameter combinations for a hydrological model calibration  
54 problem, and suggested that many combinations could produce good performance in terms of  
55 success rate which was defined as the ratio of success among a number of algorithm runs.  
56 However, it has been pointed out that the parameter values suggested by Duan et al. (1994)  
57 may be inefficient, when other algorithm performance criteria: for example, convergence  
58 speed, are considered (Behrangi et al., 2008a; Tolson and Shoemaker, 2008). More  
59 importantly, Duan et al. (1994) did not considered the interactions among parameters, that is,  
60 only the individual impacts of algorithm parameters were considered.

61

62 Hadka and Reed (2011) proposed a framework to assess the influence of multi-objective  
63 algorithm parameters based on Sobol's global sensitivity analysis method (Sobol', 2001).  
64 However, the proposed framework has a huge computational demand, due to the use of  
65 Sobol's method. In the study of Hadka and Reed (2011), 280 million algorithm runs were  
66 executed on a CyberStar computing cluster which consists of 512 2.7 GHz processors and  
67 1536 2.66 GHz processors. This huge computational burden is not affordable with commonly  
68 available computational resources. Further, Hadka and Reed (2011) did not show the dynamic  
69 sensitivity of optimization algorithm parameters which is particularly useful to understand the  
70 convergence speed in hydrological model calibration.

71

72 A variance decomposition-based method - Analysis of Variance (ANOVA) has been used to  
73 quantify the influence of uncertain contributors in a process in many studies. It allows for the  
74 analysis of individual and interactive impacts of contributors, and therefore allows for the  
75 identification of influential contributors and the understanding of parameter interactions. For

76 instance, it has been used to quantify the influence of climate models, statistical downscaling  
77 approaches and hydrological models on projected future flows (Bosshard et al., 2013). This  
78 method has also been used to investigate the influence of climate change scenarios on water  
79 resources, the influence of climate change uncertainties on projected future flows, and the  
80 impacts of climate changes on flow frequency (Köplin et al., 2013; Addor et al., 2014;  
81 Giuntoli et al., 2015). In these investigations, respective contributions of various uncertainty  
82 sources to the overall output variance have been compared, and ANOVA has shown good  
83 performance.

84

85 Dynamic sensitivity analysis can reveal the changes of the influences of individual  
86 parameters and their interactions during a search process. Most recently, it has gained  
87 increasing attention in the field of hydrological modeling. For example, the dynamic  
88 sensitivity of hydrological model parameters has been studied to understand the variations of  
89 modelled hydrological processes, and to verify the modifications of hydrological models  
90 (Pfannerstill et al., 2015). In addition, advancements have been made in studying the dynamic  
91 effects of hydrological model formulations, dynamic performance of hydrological models  
92 and dynamic tuning of algorithm parameters (Rolf, 1982; Sandip et al., 2009; van Werkhoven  
93 et al., 2009; Eiben and Smit, 2011; Reusser et al., 2011; Reusser and Zehe, 2011; Garambois  
94 et al., 2013; Herman et al., 2013). However, to the best of our knowledge, few studies have  
95 been carried out to investigate the dynamic sensitivity of optimization algorithm parameters.

96

97 The overall aim of this paper is to provide a global sensitivity analysis-based framework to  
98 dynamically quantify individual and interactive impacts of algorithm parameters on  
99 optimization performance. ANOVA was employed to quantify the impacts, because it is more  
100 computationally efficient compared with Sobol's approach. The SCE-UA algorithm was

101 selected as an optimization algorithm to demonstrate the framework. The proposed  
102 framework was first tested on five benchmark test functions, with up to 12 dimensions, and  
103 then applied to a TOPMODEL hydrological model calibration problem, representing different  
104 problems of various levels of difficulty. Two algorithm performance criteria - convergence  
105 speed and success rate - were compared in terms of parameter influence. The framework  
106 provides an improved understanding of the significant roles of algorithm parameters in the  
107 optimization processes, and highlights the importance of considering interactive influence  
108 among parameters, which is beyond the information that can be provided by conventional  
109 approach. Thus it can assist hydrological model calibration by selecting more appropriate  
110 algorithm parameter values to improve calibration efficiency, which is particularly important  
111 for a computationally intensive model.

112

## 113 **2 Algorithm and materials**

### 114 **2.1 SCE-UA algorithm**

115 SCE-UA algorithm was investigated because the influence of its parameters had been  
116 investigated in many studies (Duan et al., 1994; Behrangi et al., 2008a; Tolson and  
117 Shoemaker, 2008). The SCE-UA has four main features: (1) combination of deterministic and  
118 probabilistic approaches; (2) systematic evolution of complex points; (3) complex shuffling;  
119 and (4) competitive evolution. These characteristics stand for a combination of several  
120 approaches, including the simplex method (Nelder and Mead, 1965), the control random  
121 search (Price, 1987) and evolutionary algorithms (Holland, 1975). The introduction of  
122 complex shuffling in SCE-UA is an advanced technique which successfully ensures that the  
123 information of all populations is shared by each individual complex. Initially, a set of  
124 individuals are randomly sampled from the parameter space, and then selected individuals are  
125 divided into several complexes. Each complex evolves using a competitive evolutionary

126 algorithm. All individuals are shuffled and reassigned to new complexes to enable  
127 information sharing. As the search progresses, the entire population moves to global optimal  
128 solutions. A detailed description of SCE-UA can be found in Duan et al. (1993).

129

130 The SCE-UA performance is affected by objective functions, dimensions of decision  
131 variables and data used for calibration (Duan et al., 1994; Tolson and Shoemaker, 2007;  
132 Behrangi et al., 2008a; Tolson and Shoemaker, 2008). Thus five benchmark test functions,  
133 with up to 12 dimensions, and a hydrological model for flood simulations were employed to  
134 represent different levels of complexities.

135

## 136 **2.2 Benchmark test functions**

137 The five benchmark test functions were Rastrigin, Ackley, Levy and Montalvo 1 (LM1),  
138 Levy and Montalvo 2 (LM2) and Levy. These functions are characterized by a large number  
139 of local minima and a large search space, and have been chosen by many researchers to  
140 evaluate optimization algorithms (Ali et al., 2005; Deep and Thakur, 2007; Tolson and  
141 Shoemaker, 2007; Behrangi et al., 2008a; Tolson and Shoemaker, 2008; Chia et al., 2011).  
142 The equations of these benchmark test functions were listed in Appendix A.

143

## 144 **2.3 Hydrological model calibration problem**

145 The Biliu river basin (2814 km<sup>2</sup>), located in a peninsula region between the Bohai Sea and  
146 the Huanghai Sea, China, was used for the TOPMODEL calibration. It covers longitudes  
147 from 122.29 °E to 122.92 °E and latitudes from 39.54 °N to 40.35 °N. This basin is  
148 characterized by a monsoon climate, and summer (July to September) is the main rainfall  
149 period. The average annual temperature is 10.5 °C, and the lowest and the highest temperature  
150 is -4.7 °C in January and 24 °C in August, respectively. The major land cover types are forest

151 and farmland. There are eleven rainfall gauges and one discharge gauge. The basin average  
152 rainfall was calculated using the Thiessen method, and six flood data with different flood  
153 magnitudes were used in calibration to represent the influence of data on SCE-UA  
154 performance.

155

156 TOPMODEL is a physically based, variable contributing area model which combines the  
157 advantages of a simple lumped parameter model with distributed effects (Beven and Kirkby,  
158 1979). Fundamental of TOPMODEL's parameterization are three assumptions: (1)  
159 saturated-zone dynamics can be approximated by successive steady-state representations; (2)  
160 hydrological gradients of the saturated zone can be approximated by the local topographic  
161 surface slope; and (3) the transmissivity profile whose form exponentially declines along the  
162 vertical depth of the water table or storage, is spatially constant. On the basis of above  
163 mentioned assumptions, the index of hydrological similarity is represented as the topographic  
164 index  $\ln(a / \tan \beta)$  where  $a$  is the area per unit contour length and  $\beta$  is local slope angle.  
165 The greater upslope contributing areas and lower gradient areas are more likely to be  
166 saturated. More detailed description of TOPMODEL and its mathematical formulations can  
167 be found in Beven and Kirkby (1979). TOPMODEL has been widely used, because of its  
168 relatively simple model structure (Blazkova and Beven, 1997; Cameron et al., 1999; Hossain  
169 and Anagnostou, 2005; Bastola et al., 2008; Gallart et al., 2008; Bouilloud et al., 2010; Qi et  
170 al., 2013). TOPMODEL consists of six parameters, and their ranges and brief descriptions  
171 were given in Table 1.

172

173 The Nash-Sutcliffe Efficiency (NSE) was selected as a performance metric for TOPMODEL  
174 calibration:

175

$$NSE = 1 - \frac{\sum_{t=1}^T (Q_{st} - Q_{mt})^2}{\sum_{t=1}^T (Q_{mt} - Q_m^m)^2} \quad (1)$$

176 where  $Q_{st}$  ( $\text{m}^3/\text{s}$ ) and  $Q_{mt}$  ( $\text{m}^3/\text{s}$ ) are the simulated and measured flows at time  $t$ ;  $T$  is the total  
 177 number of flood data points and  $Q_m^m$  ( $\text{m}^3/\text{s}$ ) is the average of measured flows. The best  
 178 theoretical value of  $NSE$  is 1.0. As SCE-UA was set up for minimization problems in this  
 179 study, the following objective function was used in the TOPMODEL calibration

180

$$f = 1 - NSE \quad (2)$$

181 The best theoretical value of  $f$  is 0.0, while its true minimum value is unknown for real  
 182 calibration problems since model and data errors exist.

183

### 184 **3 Methodology**

185 Fig. 1 shows the flowchart of the proposed framework. The framework includes three main  
 186 components for an investigated algorithm: (1) selection of concerned parameter values and  
 187 random combinations (Fig. 1a); (2) selection of performance metrics which should reflect the  
 188 concerns of algorithm users: for example, convergence speed and success rate, which are  
 189 illustrated in Fig. 1b; and (3) use of ANOVA to decompose the contributions of parameters  
 190 and their interactions to reveal the influence of parameters on algorithm performance, as  
 191 shown in Fig. 1c where the influence on convergence speed and success rate is shown as a  
 192 three parameter case. It should be noted that the sample number for each parameter can be  
 193 different, that is,  $m_1$ ,  $m_i$  and  $m_n$  are not required to be equal in Fig. 1a.

194

195 The remainder of this section will illustrate the framework using SCE-UA algorithm and  
 196 selected calibration problems.

197

### 198 3.1 SCE-UA parameters and performance metrics

199 Three parameters of SCE-UA were investigated: (1) complex number ( $P$ ), (2) reflection  
200 parameter (alpha) and (3) contraction parameter (beta), as suggested by several studies  
201 (Tolson and Shoemaker, 2007; Behrangi et al., 2008a; Tolson and Shoemaker, 2008). The  
202 selected SCE-UA parameters  $P$ , alpha and beta are in the ranges of [1, 40], [0.1, 3.0] and  
203 [0.05, 1], respectively. It should be noted that  $P$  must be an integer. The parameter ranges  
204 were defined based on the following studies: Duan et al. (1994), Tolson and Shoemaker  
205 (2007) and Tolson and Shoemaker (2008).

206

207 In this paper, 11 values for each selected parameter were randomly selected from parameter  
208 ranges considering the computational burdens. Fig. 2 depicts the random combinations of  
209 algorithm parameters, and every combination was used to optimize objective functions  $f$ . In  
210 each box of Fig. 2, the number is the selected parameter values, and three values out of the 11  
211 values were shown.

212

213 Two algorithm performance criteria, convergence speed and success rate, were studied. These  
214 two criteria are of concern for researchers (Duan et al., 1994; Behrangi et al., 2008a; Tolson  
215 and Shoemaker, 2008). Convergence speed is assessed by averaging the best objective  
216 function value  $f$  over several random seed trial runs at every function evaluation (Tolson  
217 and Shoemaker, 2007; Tolson and Shoemaker, 2008). In this study, 30 and 10 random seed  
218 trial runs were used in benchmark function and TOPMODEL calibration, respectively.  
219 Success rate measures the ability to find global optimal solutions (Duan et al., 1994), and was  
220 evaluated as

$$221 \quad Success\ rate = \frac{1}{N} \left\{ number\ of\ f_{end}\ such\ that\ \left| f_{end} - f_{optimal} \right| \leq e \right\} \quad (3)$$

222 where  $f_{end}$  is a best objective function value obtained at the end of optimization;  $f_{optimal}$   
223 is a known optimal objective function value which can be a theoretical value or a specified  
224 value if theoretical value is unknown;  $e$  is an error limit and specified by algorithm users;  
225  $N$  is the number of algorithm runs: for example, 30 and 10 runs were used in benchmark  
226 function and TOPMODEL calibration problems respectively. The reasons why these numbers  
227 of runs were used are explained in Section 4. Each parameter combination in Fig. 2  
228 corresponds to a convergence speed and a success rate, and therefore  $11 \times 11 \times 11$  convergence  
229 speed data at every function/model evaluation and success rates can be obtained, where  
230 number 11 represents the number of selected parameter values. ANOVA was used to  
231 decompose the convergence speed and success rate variances resulted from 1331 parameter  
232 combinations into contributions of individual SCE-UA parameters and parameter interactions.  
233 To relate performance criteria ( $M$ ) to algorithm parameters, superscripts  $j$ ,  $k$  and  $l$  in  $M^{j,k,l}$   
234 were used to represent  $P$ , alpha and beta, respectively, in the equations below.

235

### 236 **3.2 Influence quantification**

237 It has been argued that ANOVA approach is based on a biased variance estimator that  
238 underestimates the variance when a small sample size is used (Bosshard et al., 2013). To  
239 reduce the effects of the biased estimator on contribution quantification, Bosshard et al. (2013)  
240 proposed a subsampling method, which was also used in this study. This subsampling  
241 approach does not need extra optimization trials; therefore it can reduce the computational  
242 burden. In each subsampling iteration  $i$ , we selected two  $P$  values out of all  $P$  values, and the  
243 superscript  $j$  in calculating  $M^{j,k,l}$  was replaced with  $g(h,i)$ . The total number of  
244 2-combination is 55 in this study, and correspondingly, the superscript  $g$  is a  $2 \times 55$  matrix as  
245 follows

246 
$$g = \begin{pmatrix} 1 & 1 & \dots & 1 & 2 & 2 & \dots & 8 & 8 & 8 & 9 & 9 & 10 \\ 2 & 3 & \dots & 11 & 3 & 4 & \dots & 9 & 10 & 11 & 10 & 11 & 11 \end{pmatrix} \quad (4)$$

247 Based on ANOVA, the total sum of squares (*SST*) can be divided into sums of squares due to  
 248 the individual and interactive effects:

249 
$$SST = SSA + SSB + SSC + SSI \quad (5)$$

250 where *SSA* is the contribution of *P*; *SSB* is the contribution of alpha; *SSC* is the contribution  
 251 of beta; and *SSI* is the contribution of their interactions.

252

253 The terms can be estimated using the subsampling procedure as follows (Bosshard et al.,  
 254 2013):

255 
$$SST_i = \sum_{h=1}^H \sum_{k=1}^K \sum_{l=1}^L \left( M^{g(h,i),k,l} - M^{g(o,i),o,o} \right)^2 \quad (6)$$

256 
$$SSA_i = K \cdot L \cdot \sum_{h=1}^H \left( M^{g(h,i),o,o} - M^{g(o,i),o,o} \right)^2 \quad (7)$$

257 
$$SSB_i = H \cdot L \cdot \sum_{k=1}^K \left( M^{g(o,i),k,o} - M^{g(o,i),o,o} \right)^2 \quad (8)$$

258 
$$SSC_i = H \cdot K \cdot \sum_{l=1}^L \left( M^{g(o,i),o,l} - M^{g(o,i),o,o} \right)^2 \quad (9)$$

259 
$$SSI_i = \sum_{h=1}^H \sum_{k=1}^K \sum_{l=1}^L \left( M^{g(h,i),k,l} - M^{g(h,i),o,o} - M^{g(o,i),k,o} - M^{g(o,i),o,l} + 2 \cdot M^{g(o,i),o,o} \right)^2 \quad (10)$$

260 where symbol ° indicates the averaging over the particular index. Then the contribution of  
 261 each influential source  $\eta^2$  is calculated as follows:

262 
$$\eta_P^2 = \frac{1}{I} \sum_{i=1}^I \frac{SSA_i}{SST_i} \quad (11)$$

263 
$$\eta_{alpha}^2 = \frac{1}{I} \sum_{i=1}^I \frac{SSB_i}{SST_i} \quad (12)$$

264 
$$\eta_{beta}^2 = \frac{1}{I} \sum_{i=1}^I \frac{SSC_i}{SST_i} \quad (13)$$

265 
$$\eta_{interaction}^2 = \frac{1}{I} \sum_{i=1}^I \frac{SSI_i}{SST_i} \quad (14)$$

266  $\eta^2$  has a value between 0 and 1, which represents the respective contribution to the overall  
267 variations of  $M$ .

268

## 269 **4 Results and discussion**

### 270 **4.1 Benchmark functions**

271 In the simulations, SCE-UA algorithm was stopped when the total number of function  
272 evaluations reached a prescribed value. In the following subsections, the contributions of  
273 individual SCE-UA parameters and parameter interactions to the variance of convergence  
274 speed at every function evaluation and success rate are quantified for the selected benchmark  
275 functions.

276

#### 277 **4.1.1 Convergence speed analyses**

278 Fig. 3 shows the contributions of individual SCE-UA algorithm parameters and their  
279 interactions in terms of convergence speed in benchmark function calibration, where average  
280 best function values over 30 random seed trial runs were used. The 30 random seed trial runs  
281 were used considering computational burden, and were the same as many other studies: for  
282 example, Deep and Thakur (2007), Tolson and Shoemaker (2007) and Chia et al. (2011). The  
283 benchmark functions were optimized under 6, 8, 10 and 12 dimensions. The contributions of  
284 individual parameters and their interactions are represented by color strips varying with the  
285 function evaluation number shown in the x-axis.

286

287 For the 6-dimensional Rastrigin function, the influence of  $P$  increases and then decreases,  
288 while the impacts of beta and alpha increase with an increase in function evaluation number.  
289 The influence of alpha is larger than beta, and the influence of  $P$  at early stages is larger than

290 alpha and beta. The interactions among  $P$ , beta and alpha have significant influence,  
291 decreasing with an increase in function evaluations. Interactive impacts are larger than those  
292 from any individual parameter at initial search stages, and have approximately the same  
293 influence as  $P$  and alpha, but have a slightly larger influence than beta at later optimization  
294 stages. For other 6-dimensional functions, similar results are shown; except that, for LM1,  
295 LM2 and Levy at later stages, the influence of beta becomes larger than  $P$ , alpha and  
296 interactions, and that the influence of alpha becomes the smallest. The differences result from  
297 differences in benchmark functions, which implies that objective functions have influence on  
298 algorithm performance and that using several test functions is necessary.

299

300 Comparing different dimensions at later stages, with a dimension increase, influence of  $P$   
301 increases but influence of beta decreases, whilst alpha influence and interactive influence  
302 remain approximately the same, which indicates with an increase in dimensions the  
303 importance of  $P$  increases but the importance of beta decreases. This information implies that  
304 dimensions have influence on the performance of parameters, and that optimal parameter  
305 values derived from low dimensional problems may not have optimal performance for high  
306 dimensional problems. All results show that the contributions from various sources become  
307 almost constant at the end of the search process, indicating that 1000 function evaluations are  
308 sufficient.

309

#### 310 **4.1.2 Success rate analyses**

311 The contributions to success rate based on the 30 random seed trial runs are shown in Fig. 4  
312 under an error level of 0.001 in terms of benchmark function calibration. The error level  
313 represents the absolute differences between an optimal objective function value found at the  
314 end of the optimization and a real optimal value, and is subjectively selected: for example,

315 Duan et al. (1994) has used 0.001 as an error limit, and Deep and Thakur (2007) and Chia et  
316 al. (2011) have used 0.01 as error limits.

317

318 For the 6-dimensional Rastrigin function,  $P$ , beta and alpha all have significant contributions,  
319 and alpha contributes more than  $P$  and beta, while interactions account for the majority.

320 Comparing different dimensions, with a dimension increase, the contributions of  $P$ , beta and

321 alpha decrease, but interactive contribution increases. Compared with other functions, similar

322 results can be obtained, except that the contribution of alpha is smaller than beta for

323 6-dimensional LM1 function. The differences may result from the limited number of random

324 trials. These results indicate that, for the success rate, interactions among parameters are most

325 important, and good combinations of parameters are more important than individual

326 parameters. The results are different from convergence speed analyses. This difference

327 indicates parameters have a different influence when algorithm performance criteria change.

328 It should be noted that the contributions actually includes influence of initial random seeds,

329 but this influence should be very small after many function evaluations (Wang et al., 2010).

330 In addition, the success rate is influenced by the number of function evaluations, but in our

331 study the investigations of convergence speed and success rate used the same number of

332 function evaluations: thus the comparison results are free of influence. Another error limit

333 0.005 was also analyzed, and similar results are obtained.

334

## 335 **4.2 TOPMODEL**

336 Every parameter combination can generate a convergence speed line and a success rate in

337 TOPMDOEL calibration, and therefore 1331 convergence speed lines and success rates were

338 obtained. They are shown in Fig. 5 using flood 1984-06-15 as an examples. Different colors

339 are used to distinguish lines in Fig. 5a. Because the theoretically optimal objective function

340 values were not known, the optimal values obtained from all the  $1331 \times 10$  optimization runs  
341 were used. The variations of histogram heights represent the variance of success rate, as  
342 shown in Fig. 5b. Fig. 5a shows there are many vertical lines before 500 function evaluations  
343 which are resulted from larger  $P$  values. This information implies that  $P$  has larger influence  
344 before 500 function evaluations. Significant differences exist in convergence speed and  
345 success rate, which can be attributed to the variations of parameter values. Thus it is  
346 necessary to analyze the parameter influence. In the flowing subsections, the contributions of  
347 individual SCE-UA parameters and parameter interactions are quantified for all six flood  
348 events.

349

#### 350 **4.2.1 Convergence speed analyses**

351 Fig. 6 shows the contributions of SCE-UA algorithm parameters and interactions in terms of  
352 convergence speed in TOPMODEL calibration, where average best function values over 10  
353 random seed trial runs were used considering the computational burdens. The number of  
354 random seed trial runs are similar to the study by Duan et al. (1994). Each panel represents  
355 the results from a flood event. The contributions of individual parameters and their  
356 interactions are represented by strips varying with the model evaluations shown in the x-axis.

357

358 Fig. 6a shows the influence of  $P$  increases first and then decreases. However, the influence of  
359 alpha grows with an increase in model evaluations, and the contribution of beta slightly  
360 increases. Interactions among  $P$ , alpha and beta have significant contributions, decreasing  
361 with an increase in model evaluations. Interactive impacts are larger than beta in all model  
362 evaluations, while significantly higher than  $P$  at initial stages and at later stages. Compared  
363 with alpha, interactive influence is larger at initial stages, and is a little smaller at later stages.  
364 This information implies that without considering interactions the calibration of parameters

365 may not be effective for improving algorithm performance. For other flood events, similar  
366 results can be obtained. These results are consistent with the convergence speed variance in  
367 Fig. 5a where  $P$  has greater influence at early stage. This is because that larger  $P$  values can  
368 slow the information exchange among different complexes. Consequently, larger  $P$  values  
369 have few positive efforts in improving convergence speed at early optimization stage. This  
370 information implies that larger influence does not suggest greater convergence speed.

371

372 Comparing results in each panel, differences can be attributed to the different roles that  
373 parameters play in the SCE-UA calibration processes, while differences among panels result  
374 from the influence of data. The complex number  $P$  controls information exchange among  
375 complexes; with an increase in model evaluations, information exchange among complexes  
376 doesn't provide more positive influence in searching for optimal solutions compared with  
377 early stages, which implies the complex number has significant influence on the searching  
378 speed at early stage. However, for alpha, much more positive influence arises with an  
379 increase in model evaluations. Comparing Fig. 3 and Fig. 6, the influence of beta is the  
380 smallest in Fig. 6, which is different from the results of the 6-dimensional functions in Fig. 3.  
381 This difference results from objective functions and errors in data used in Fig. 6, which  
382 implies that objective functions and data have significant influence besides variable  
383 dimensions. All results show patterns are clearly revealed at the end of optimization, and thus  
384 1000 model evaluations are sufficient.

385

#### 386 **4.2.2 Success rate analyses**

387 The contributions to success rate in the 10 runs under an error level of 0.001 in terms of  
388 TOPMODEL calibration are shown in Fig. 7.

389

390 Fig. 7a shows the contribution of beta is the smallest, and the contribution of alpha is the  
391 largest among individual parameter contributions. However, interactions among parameters  
392 contribute the most. Similarly, other five cases show the contributions of alpha are the  
393 greatest among individual parameter contributions, or at least not smaller than individual  
394 parameter contributions. Comparing the differences among different flood data, Figs. 7d, 7e  
395 and 7f show the contribution of beta is larger than  $P$ , and Fig. 7b shows contributions of beta  
396 and  $P$  are equal. These differences may result from different flood data and optimal objective  
397 function values: for example, the optimal objective function value is 0.0223 for Fig. 7a, and  
398 is 0.193 for Fig. 7d. This implies that calibration data have impacts on the parameter  
399 influence, and therefore using several flood data sets is necessary. Compared with Fig. 4,  
400 similar results can be obtained, which indicates that the results could be applicable to other  
401 calibration problems.

402

403 In Fig. 5b, the success rate has several peaks, and these peaks are the results of some good  
404 parameter combinations that have relatively small  $P$  values (smaller than 5), which may be  
405 because the smaller dimension 6 and limited model evaluations (Duan et al., 1994). When  
406 dimension increases, required  $P$  and model evaluation number should increase to obtain high  
407 success rate (Duan et al., 1994). This information implies  $P$  has large influence on greater  
408 success rate, which is different from Fig. 7a where interactions contribute the majority of the  
409 variance. This difference is resulted from the differences in definitions of success rate and  
410 variance: success rate measures the ability of finding optimal results, but variance measures  
411 the changes of this ability along the variations of parameter values. This information implies  
412 that larger influence does not guarantee greater success rate. Another error limit 0.005 was  
413 also analyzed, and similar results are obtained.

414

### 415 **4.3 Discussion**

416 There has been a trend to develop parsimonious algorithms and adaptive parameter control  
417 schemes for users' convenience and reduction in algorithm complexity (Gao et al., 2014; Wu  
418 et al., 2014; Yu et al., 2014; Goldman and Punch, 2015). However, the proposed framework  
419 in this study provides a means to understand the performance of optimization algorithms by  
420 revealing the dynamics of parameter sensitivity in the search processes. In addition, the  
421 dynamic sensitivity can provide information to set dynamic algorithm parameter values,  
422 which could provide a method to improve algorithm efficiency (Eiben and Smit, 2011; Rui et  
423 al., 2015). Furthermore, the dynamic sensitivity information could provide evidence for  
424 assigning appropriate parameter values in different optimization stages to improve the fitness  
425 of optimization algorithms (Giorgos et al., 2015).

426

427 In the study by Tolson and Shoemaker (2007), the convergence speed of the SCE-UA was  
428 assessed based on adjustments of parameter  $P$ , and the results were problematic because other  
429 parameters: such as, beta and alpha, were not considered, as was pointed out by Behrangi et  
430 al. (2008a). Although Behrangi et al. (2008a) realized the influence of other parameters, they  
431 did not quantitatively show the influence nor explicitly indicated interactions among  
432 parameters. In contrast, the results of this study do quantitatively compare the influence of  
433 parameters and explicitly show the dynamic impacts of interactions along the number of  
434 function evaluations. This information could guide algorithm development and applications:  
435 for example, if an algorithm parameter is not sensitive, it would be helpless to tune this  
436 parameter to change algorithm performance; if a parameter has greater sensitivity than the  
437 sum of other parameters and interactions, the calibration efficiency may be mainly  
438 determined by this parameter and calibrating other parameters may be ineffective to change  
439 algorithm performance.

440

441 In the study by Duan et al. (1994), the importance of  $P$  was stressed, and it suggested that  $P$   
442 should increase with an increase in the difficulty of model calibration problems to obtain a  
443 high success rate. However, our study reveals that alpha could have a larger influence than  $P$   
444 on success rate, and more importantly, the interactions could play an important role in success  
445 rate. This information will help optimization algorithm parameter selections in hydrological  
446 model calibration, and promote further development in searching for optimal parameters for  
447 SCE-UA given consideration of parameter interactions.

448

449 It should be noted that the success rate is influenced by the number of function evaluations  
450 and error limits. There are several parameter combinations that are failed to success within  
451 1000 function/model evaluations under an error limit 0.001. More function/model evaluations  
452 are needed if it is needed to make sure all parameter combinations are successful. In addition,  
453 the SCE-UA parameter ranges and the random seed trial runs could also have influence on  
454 results. However, the case study of this research shows that  $P$  is not always the most  
455 influential parameter; the developed framework can provide a means to quantify the  
456 influences of function evaluation number, error limits, parameter ranges and random seed  
457 trial runs on the parameter sensitivity, which can be done by comparing the sensitivity  
458 differences of several numbers of these influential variables. It should also be noted that the  
459 variance decomposition results reveal the variations of convergence speed and success rates  
460 along parameter variations, but larger influence does not guarantee faster convergence speed  
461 and greater success rate. Larger influence just suggests convergence speed and success rate  
462 can be significantly changed when parameter values are altered.

463

464 Convergence speed and success rate have to be considered carefully in the model calibration  
465 process in practice. Essentially, the selection of algorithm parameter values is based on  
466 modellers' preference to convergence speed or success rate, and the computational demand of  
467 a hydrological model also plays a key role. Duan et al. (1994) provided guidance for model  
468 calibration but it can be applied to success rate only (Behrangi et al., 2008b). However, in this  
469 study, we showed how the convergence speed is affected by the parameters and a need to  
470 balance convergence and success rate. The value of  $P$  should be carefully selected to improve  
471 convergence speed at an early stage during optimization; the values of  $\beta$  and  $\alpha$  should  
472 have more attention in order to improve the convergence speed at a later stage. For success  
473 rate,  $\alpha$  can be more influential than  $P$ .

474

475 Using the Rastrigin function with up to 12 dimensions as an example, Fig. 8 shows  
476 comparison of the convergence speed curves (black bold line) from a set of default parameter  
477 values suggested by Duan et al. (1994) and the lower convergence speed boundray curves  
478 (red bold line) from the 1331 parameter combinations. Three points (A, B and C) from the  
479 lower convergence speed boundray lines are selected and corresponding parameter values are  
480 shown as well. Points A, B and C correspond to 100, 400 and 700 function evaluations,  
481 respectively.

482

483 In the three cases of varying dimensions, the best combination of paramter values is different  
484 at different function evaluation numbers, implying that one combination of parameter values  
485 can not maintain good performance during the search process. It should be noted that, in the  
486 cases of 6- and 8-dimensions, although the *alpha* values are the same at points A, B and C,  
487 the *P* and *beta* values are different: thus the paramter value combinations are different at  
488 points A, B and C. Because the best parameter values that have the best convergence speed

489 vary at different function evaluation numbers, it is difficult to provide a set of parameter values  
490 that can maintain the best convergence speed during the search process. However, in this  
491 study, we provide useful information on the parameter influence on convergence speed in the  
492 search process, including interactive influences of parameter values, and therefore we provide  
493 an enhanced understanding of SCE-UA algorithm parameter value setting. Future research is  
494 encouraged to develop dynamic parameter values in the search process to improve the  
495 convergence speed.

496

497 In Fig. 8 it can be seen that there is a gap between the two bold convergence lines, indicating  
498 that an improvement can be achieved by changing the default parameter values. In addition, it  
499 can be seen that the gaps become wider with an increase in the dimension, and this implies  
500 that higher gains in the convergence speed improvements can be obtained for high dimension  
501 optimization problems compared with low dimension problems. Thus, quantifying dynamic  
502 sensitivity of parameters reveals useful information for model calibration.

503

504 It should be noted that hydrological models such as TOPMODEL have the equifinality  
505 problem, which is defined as that many sets of different parameter values are acceptable and  
506 result in the same objective function values (Beven and Binley, 1992; Beven and Freer, 2001).  
507 However, the equifinality problem does not include the influence on the variations of  
508 objective function values, and therefore its influence is negligible in algorithm performance  
509 assessment (Tolson and Shoemaker, 2007; Tolson and Shoemaker, 2008; Zhang et al., 2008;  
510 Arsenault et al., 2014).

511

## 512 **5 Conclusions**

513 The diverse control mechanisms of algorithm parameters in algorithm performance should be

514 investigated, which can provide users with the information on which parameter is most  
515 influential and on how influence changes along function evaluation number and algorithm  
516 performance criteria. This study developed a new framework to quantify dynamic sensitivity  
517 of optimization algorithm parameters and their interactions based on ANONA, and  
518 investigated the influence of the parameters of SCE-UA using a suite of benchmark functions  
519 and a hydrological model calibration problem. The major findings are as follows.

520

521 First, the proposed framework can effectively reveal the dynamic sensitivity of algorithm  
522 parameters in the search process, including individual influences of parameters and their  
523 interactive impacts on algorithm performance. This provides an effective tool to gain an  
524 improved understanding of the significant roles of algorithm parameters.

525

526 Second, the value of  $P$  should be carefully selected to improve convergence speed at early  
527 optimization stage; beta and alpha should draw much more attention to improve the  
528 convergence speed at later optimization stage. For success rate, alpha can be more influential  
529 than  $P$ .

530

531 Third, parameter combinations could have significant influence on algorithm performance,  
532 which highlights the importance of considering interactive influence among parameters.

533

534 The proposed framework can guide efforts to calibrate algorithm parameters to improve  
535 computational efficiency in hydrological model calibration processes. In the future, a  
536 sensitivity-based parameter auto-adjusting approach will be studied for SCE-UA.

537

538 **Acknowledgements:**

539 This study was supported by the National Natural Science Foundation of China (Grant No.  
 540 51320105010 and 51279021). The first author gratefully acknowledges the financial support  
 541 provided by the China Scholarship Council. The data from Biliu basin were obtained from  
 542 the Biliu reservoir administration.

543

544 **Appendix A: Benchmark functions**

No.	Function	Definition	Parameter Space	Optimal Value
1	Rastrigin (1974)	$f(x) = \sum_{i=1}^D [x_i^2 - \cos(2\pi x_i)]$	$[-2, 2]^D$	$f^* = -D$
2	Ackley (1987)	$f(x) = -20 \exp \sqrt{\frac{1}{D} \sum_{i=1}^D x_i^2} - \exp \left[ \frac{1}{D} \sum_{i=1}^D (\cos(2\pi x_i)) \right]$	$[-1, 3]^D$	$f^* = -20 - e$
3	Levy and Montalvo 1 (LM1)	$f(x) = \frac{\pi}{n} \left( 10 \sin^2(\pi y_1) + \sum_{i=1}^{n-1} (y_i - 1)^2 \left[ 1 + 10 \sin^2(\pi y_{i+1}) \right] + (y_n - 1)^2 \right), y_i = 1 + \frac{1}{4}(x_i + 1)$	$[-10, 10]^D$	$f^* = 0$
4	Levy and Montalvo 2 (LM2)	$f(x) = 0.1 \left( \sin^2(3\pi x_1) + \sum_{i=1}^{n-1} (x_i - 1)^2 \left[ 1 + \sin^2(3\pi x_{i+1}) \right] + (x_n - 1)^2 \left[ 1 + \sin^2(2\pi x_n) \right] \right)$	$[-5, 5]^D$	$f^* = 0$
5	Levy	$f(x) = \left( \sin^2(\pi y_1) + \sum_{i=1}^{n-1} (y_i - 1)^2 \left[ 1 + 10 \sin^2(\pi y_{i+1} + 1) \right] + (y_n - 1)^2 \left( 1 + 10 \sin^2(2\pi y_n) \right) \right), y_i = 1 + \frac{1}{4}(x_i + 1)$	$[-10, 10]^D$	$f^* = 0$

545

546 **References**

547 Addor, N., Rössler, O., Köplin, N., Huss, M., Weingartner, R., Seibert, J., 2014. Robust  
 548 changes and sources of uncertainty in the projected hydrological regimes of Swiss  
 549 catchments. *Water Resources Research*, 50(10): 7541-7562.

550 Ali, M.M., Khompatraporn, C., Zabinsky, Z.B., 2005. A numerical evaluation of several  
 551 stochastic algorithms on selected continuous global optimization test problems. *Journal of*  
 552 *Global Optimization*, 31(4): 635-672.

553 Arsenault, R., Poulin, A., Côté P., Brissette, F., 2014. Comparison of Stochastic Optimization  
554 Algorithms in Hydrological Model Calibration. *Journal of Hydrologic Engineering*, 19(7):  
555 1374-1384.

556 Asadzadeh, M., Tolson, B., 2013. Pareto archived dynamically dimensioned search with  
557 hypervolume-based selection for multi-objective optimization. *Engineering Optimization*,  
558 45(12): 1489-1509.

559 Bäck, T., 1996. *Evolutionary Algorithms in Theory and Practice: Evolution Strategies,*  
560 *Evolutionary Programming, Genetic Algorithms.* Oxford University Press.

561 Bastola, S., Ishidaira, H., Takeuchi, K., 2008. Regionalisation of hydrological model  
562 parameters under parameter uncertainty: A case study involving TOPMODEL and basins  
563 across the globe. *Journal of Hydrology*, 357(3-4): 188-206.

564 Behrangi, A., Khakbaz, B., Vrugt, J.A., Duan, Q., Sorooshian, S., 2008a. Comment on  
565 “Dynamically dimensioned search algorithm for computationally efficient watershed  
566 model calibration” by Bryan A. Tolson and Christine A. Shoemaker. *Water Resources*  
567 *Research*, 44(12): W12603.

568 Behrangi, A., Khakbaz, B., Vrugt, J.A., Duan, Q., Sorooshian, S., 2008b. Comment on  
569 “Dynamically dimensioned search algorithm for computationally efficient watershed  
570 model calibration” by Bryan A. Tolson and Christine A. Shoemaker. *Water Resources*  
571 *Research*, 44(12).

572 Beven, K., Binley, A., 1992. The future of distributed models: Model calibration and  
573 uncertainty prediction. *Hydrological Processes*, 6(3): 279-298.

574 Beven, K.J., Freer, J.E., 2001. Equifinality, data assimilation, and uncertainty estimation in  
575 mechanistic modelling of complex environmental systems using the GLUE methodology.  
576 *Journal of Hydrology*, 249(1-4): 11–29.

577 Beven, K.J., Kirkby, M.J., 1979. A physically based, variable contributing area model of  
578 basin hydrology. *Hydrological Sciences Bulletin*, 24(1): 43-69.

579 Blazkova, S., Beven, K., 1997. Flood frequency prediction for data limited catchments in the  
580 Czech Republic using a stochastic rainfall model and TOPMODEL. *Journal of Hydrology*,  
581 195(1-4): 256-278.

582 Bosshard, T., Carambia, M., Goergen, K., Kotlarski, S., Krahe, P., Zappa, M., Schär, C., 2013.  
583 Quantifying uncertainty sources in an ensemble of hydrological climate-impact  
584 projections. *Water Resources Research*, 49(3): 1523-1536.

585 Bouilloud, L., Chancibault, K., Vincendon, B., Ducrocq, V., Habets, F., Saulnier, G.-M.,  
586 Anquetin, S., Martin, E., Noilhan, J., 2010. Coupling the ISBA Land Surface Model and  
587 the TOPMODEL Hydrological Model for Mediterranean Flash-Flood Forecasting:  
588 Description, Calibration, and Validation. *Journal of Hydrometeorology*, 11(2): 315-333.

589 Cameron, D.S., Beven, K.J., Tawn, J., Blazkova, S., Naden, P., 1999. Flood frequency  
590 estimation by continuous simulation for a gauged upland catchment (with uncertainty).  
591 *Journal of Hydrology*, 219(3-4): 169-187.

592 Chao, Z., Ru-bin, W., Qing-xiang, M., 2015. Calibration of Conceptual Rainfall-Runoff  
593 Models Using Global Optimization. *Advances in Meteorology*, 2015.

594 Chia, J.Y., Goh, C.K., Tan, K.C., Shim, V.A., 2011. Memetic informed evolutionary  
595 optimization via data mining. *Memetic Computing*, 3(2): 73-87.

596 Deb, K., Pratap, A., Agarwal, S., Meyarivan, T., 2002. A fast and elitist multiobjective genetic  
597 algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, 6(2).

598 Deep, K., Thakur, M., 2007. A new crossover operator for real coded genetic algorithms.  
599 *Applied Mathematics and Computation*, 188(1): 895-911.

600 Duan, Q., Gupta, V.K., Sorooshian, S., 1993. Shuffled Complex Evolution Approach for  
601 Effective and Efficient Global Minimization. *Journal of optimization theory and*  
602 *Applications*, 73(3).

603 Duan, Q., Sorooshian, S., Gupta, V., 1992. Effective and efficient global optimization for  
604 conceptual rainfall-runoff models. *Water Resources Research*, 28(4): 1015-1031.

605 Duan, Q., Sorooshian, S., Gupta, V.K., 1994. Optimal use of the SCE-UA global optimization  
606 method for calibrating watershed models. *Journal of Hydrology*, 158(3-4): 265-284.

607 Eiben, A.E., Smit, S.K., 2011. Parameter tuning for configuring and analyzing evolutionary  
608 algorithms. *Swarm and Evolutionary Computation*, 1(1): 19-31.

609 Fu, G., Kapelan, Z., Reed, P., 2012. Reducing the Complexity of Multiobjective Water  
610 Distribution System Optimization through Global Sensitivity Analysis. *Journal of Water*  
611 *Resources Planning and Management*, 138(3): 196-207.

612 Gallart, F., Latron, J., Llorens, P., Beven, K.J., 2008. Upscaling discrete internal observations  
613 for obtaining catchment-averaged TOPMODEL parameters in a small Mediterranean  
614 mountain basin. *Physics and Chemistry of the Earth*, 33(17-18): 1090-1094.

615 Gao, W., Yen, G.G., Liu, S., 2014. A cluster-based differential evolution with self-adaptive  
616 strategy for multimodal optimization. *IEEE transactions on cybernetics*, 44(8):  
617 1314-1327.

618 Garambois, P.A., Roux, H., Larnier, K., Castaings, W., Dartus, D., 2013. Characterization of  
619 process-oriented hydrologic model behavior with temporal sensitivity analysis for flash  
620 floods in Mediterranean catchments. *Hydrology and Earth System Sciences*, 17:  
621 2305-2322.

622 Giorgos, K., Mark, H., Eiben, A.E., 2015. Parameter Control in Evolutionary Algorithms:  
623 Trends and Challenges. *IEEE Transactions on Evolutionary Computation*, 19(2):  
624 167-187.

625 Giuntoli, I., Vidal, J.P., Prudhomme, C., Hannah, D.M., 2015. Future hydrological extremes:  
626 the uncertainty from multiple global climate and global hydrological models. *Earth*  
627 *System Dynamics Discussions*, 6(1): 1-30.

628 Goldman, B.W., Punch, W.F., 2015. Fast and Efficient Black Box Optimization using the  
629 Parameter-less Population Pyramid. *Evolutionary Computation*.

630 Hadka, D., Reed, P., 2011. Diagnostic Assessment of Search Controls and Failure Modes in  
631 Many-Objective Evolutionary Optimization. *Evolutionary Computation*, 20(3): 423-452.

632 Herman, J.D., Reed, P.M., Wagener, T., 2013. Time - varying sensitivity analysis clarifies the  
633 effects of watershed model formulation on model behavior. *Water Resources Research*,  
634 49: 1400-1414.

635 Holland, J.H., 1975. *Adaptation in Natural and Artificial Systems*. University of Michigan  
636 Press, Ann Arbor.

637 Hossain, F., Anagnostou, E.N., 2005. Assessment of a stochastic interpolation based  
638 parameter sampling scheme for efficient uncertainty analyses of hydrologic models.  
639 *Computers & Geosciences*, 31(4): 497-512.

640 Kollat, J.B., Reed, P.M., 2006. Comparing state-of-the-art evolutionary multi-objective  
641 algorithms for long-term groundwater monitoring design. *Advances in Water Resources*,  
642 29(6): 792-807.

643 Köplin, N., Schädler, B., Viviroli, D., Weingartner, R., 2013. The importance of glacier and  
644 forest change in hydrological climate-impact studies. *Hydrology and Earth System*  
645 *Sciences*, 17(2): 619-635.

646 Nelder, J.A., Mead, R., 1965. A simplex method for function minimization. *The Computer*  
647 *Journal*, 7(4).

648 Pfannerstill, M., Guse, B., Reusser, D., Fohrer, N., 2015. Temporal parameter sensitivity  
649 guided verification of process dynamics. *Hydrology and Earth System Sciences*  
650 *Discussions*, 12: 1729-1764.

651 Price, W.L., 1987. Global optimization algorithms for a CAD workstation. *Journal of*  
652 *Optimization Theory and Applications*, 55(1).

653 Qi, W., Zhang, C., Chu, J., Zhou, H., 2013. Sobol's sensitivity analysis for TOPMODEL  
654 hydrological model: A case study for the Biliu River Basin, China. *Journal of Hydrology*  
655 *and Environment Research*, 1: 1-10.

656 Qi, W., Zhang, C., Fu, G., Zhou, H., 2015. Global Land Data Assimilation System data  
657 assessment using a distributed biosphere hydrological model. *Journal of Hydrology*, 528:  
658 652-667.

659 Reusser, D.E., Buytaert, W., Zehe, E., 2011. Temporal dynamics of model parameter  
660 sensitivity for computationally expensive models with the Fourier amplitude sensitivity  
661 test. *Water Resources Research*, 47: W07551.

662 Reusser, D.E., Zehe, E., 2011. Inferring model structural deficits by analyzing temporal  
663 dynamics of model performance and parameter sensitivity. *Water Resources Research*.

664 Rolf, I., 1982. Parameter adaptive control algorithms—A tutorial. *Automatica*: 513528.

665 Rui, W., Maszatul, M.M., Robin, C.P., Peter, J.F., 2015. An analysis of parameter sensitivities  
666 of preference-inspired co-evolutionary algorithms. *International Journal of Systems*  
667 *Science*: 1-14.

668 Sandip, A., Rajeev, K., Chakrabarti, P.P., 2009. Adaptive parameter control of evolutionary  
669 algorithms to improve quality-time trade-off. *Applied Soft Computing*, 9(2): 527-540.

670 Sobol', I.M., 2001. Global sensitivity indices for nonlinear mathematical models and their  
671 Monte Carlo estimates. *Mathematics and Computers in Simulation*, 55(1-3): 271-280.

672 Sorooshian, S., Duan, Q., Gupta, V.K., 1993. Calibration of rainfall-runoff models:  
673 Application of global optimization to the Sacramento Soil Moisture Accounting Model.  
674 Water Resources Research, 29(4): 1185-1194.

675 Tang, Y., Reed, P., Wagener, T., 2006. How effective and efficient are multiobjective  
676 evolutionary algorithms at hydrologic model calibration? Hydrology and Earth System  
677 Sciences, 10(2): 289-307.

678 Thyer, M., Kuczera, G., Bates, B.C., 1999. Probabilistic optimization for conceptual  
679 rainfall-runoff models: A comparison of the shuffled complex evolution and simulated  
680 annealing algorithms. Water Resources Research, 35(3): 767-773.

681 Tolson, B.A., Asadzadeh, M., Maier, H.R., Zecchin, A., 2009. Hybrid discrete dynamically  
682 dimensioned search (HD-DDS) algorithm for water distribution system design  
683 optimization. Water Resources Research, 45(12): W12416.

684 Tolson, B.A., Shoemaker, C.A., 2007. Dynamically dimensioned search algorithm for  
685 computationally efficient watershed model calibration. Water Resources Research, 43(1):  
686 W01413.

687 Tolson, B.A., Shoemaker, C.A., 2008. Reply to comment on “Dynamically dimensioned  
688 search algorithm for computationally efficient watershed model calibration” by Ali  
689 Behrangi et al. Water Resources Research, 44(12).

690 van Werkhoven, K., Wagener, T., Reed, P., Tang, Y., 2009. Sensitivity-guided reduction of  
691 parametric dimensionality for multi-objective calibration of watershed models. Advances  
692 in Water Resources, 32(8): 1154-1169.

693 Wang, Y.-C., Yu, P.-S., Yang, T.-C., 2010. Comparison of genetic algorithms and shuffled  
694 complex evolution approach for calibrating distributed rainfall-runoff model.  
695 Hydrological Processes, 24(8): 1015-1026.

696 Wu, X., Rozycki, P., Wilamowski, B.M., 2014. A Hybrid Constructive Algorithm for  
697 Single-Layer Feedforward Networks Learning. IEEE transactions on neural networks and  
698 learning systems.

699 Yu, W.-J.J., Shen, M., Chen, W.-N.N., Zhan, Z.-H.H., Gong, Y.-J.J., Lin, Y., Liu, O., Zhang, J.,  
700 2014. Differential evolution with two-level parameter adaptation. IEEE transactions on  
701 cybernetics, 44(7): 1080-1099.

702 Zhang, X., Srinivasan, R., Zhao, K., Liew, M.V., 2008. Evaluation of global optimization  
703 algorithms for parameter calibration of a computationally intensive hydrologic model.  
704 Hydrological Processes, 23(3): 430-441.

705

706

Table 1 TOPMODEL parameters

Name (units)	Description	Lower bound	Upper bound
$SZM$ (m)	parameter of exponential decline in conductivity	0.005	0.04
$LNT0$ ( $m^2 h^{-1}$ )	effective lateral saturated transmissivity	-25	10
$RV$ ( $m^2 h^{-1}$ )	hill slope routing velocity	3500	8000
$SR_{max}$ (m)	maximum root zone storage	0.001	0.01
$SR_0$ (m)	initial root zone deficit	0	0.01
$TD$ ( $m h^{-1}$ )	unsaturated zone time delay per unit deficit	0.5	5

707

708

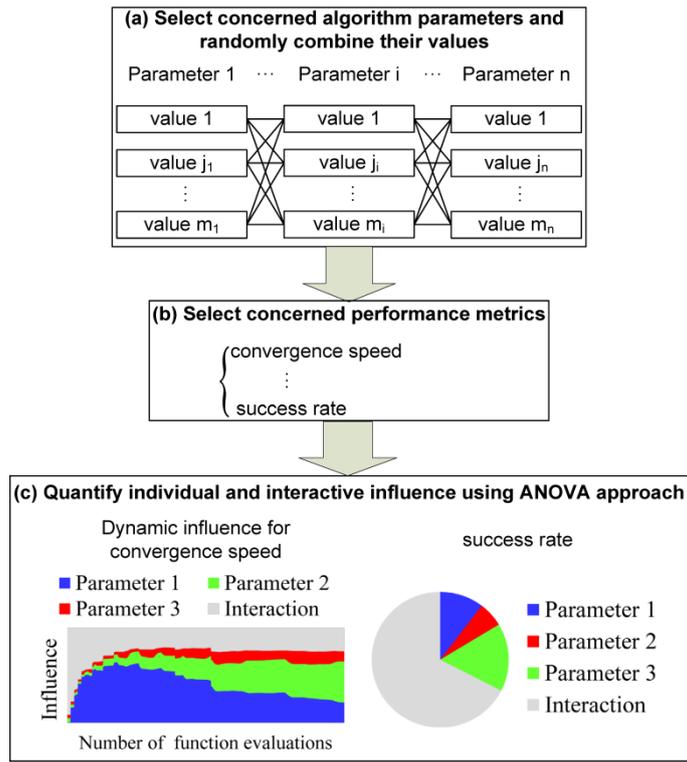
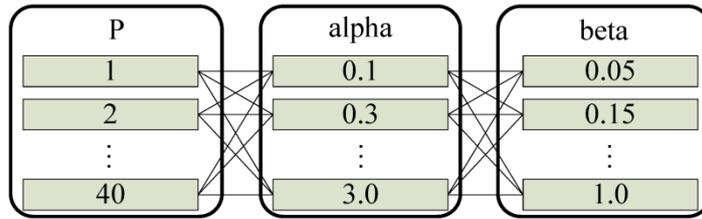


Fig. 1 Flowchart of the proposed framework.

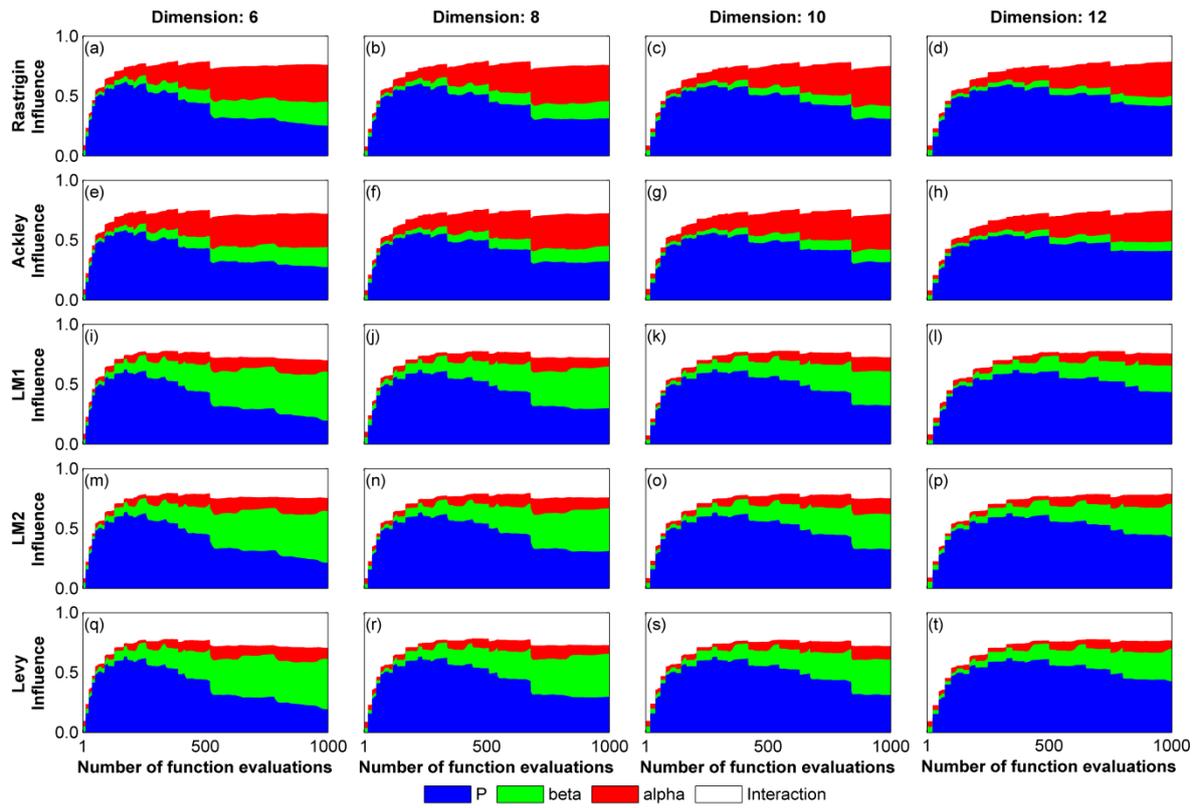
709  
710  
711



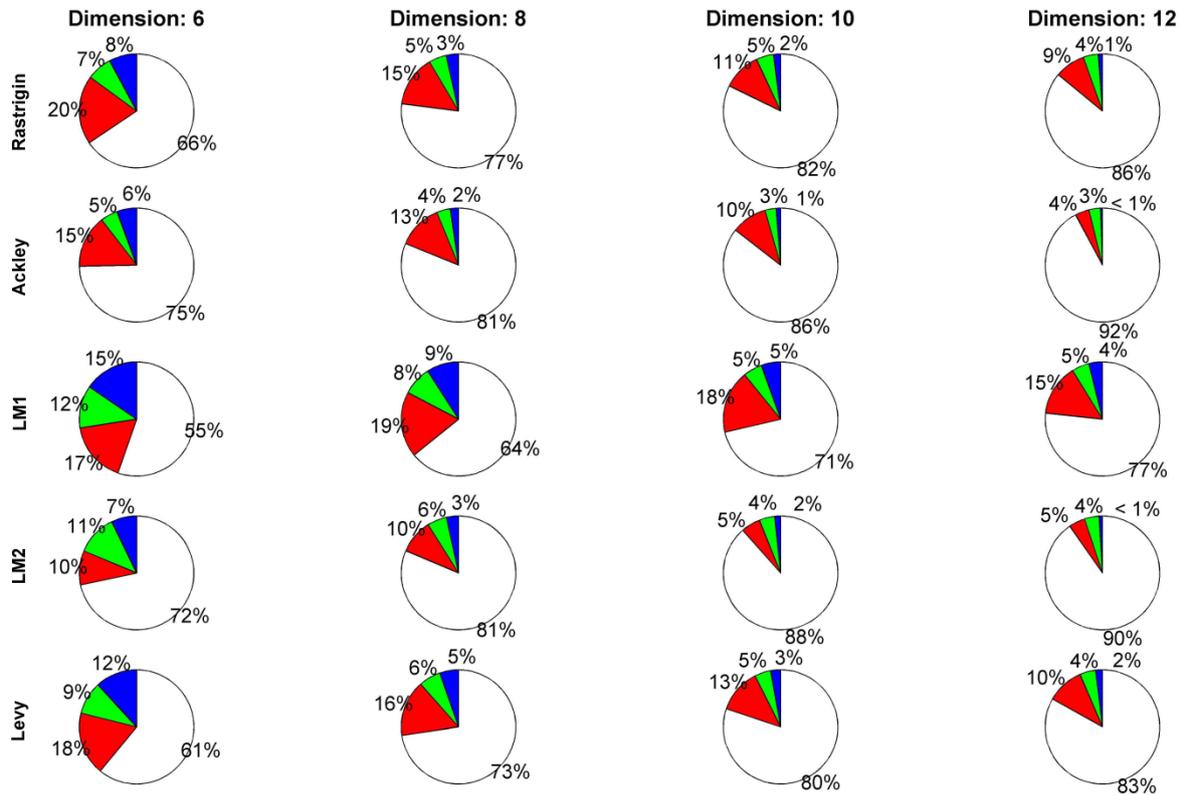
712  
713

Fig. 2 Combinations of the SCE-UA algorithm parameters:  $P$ ,  $\alpha$  and  $\beta$ .

714



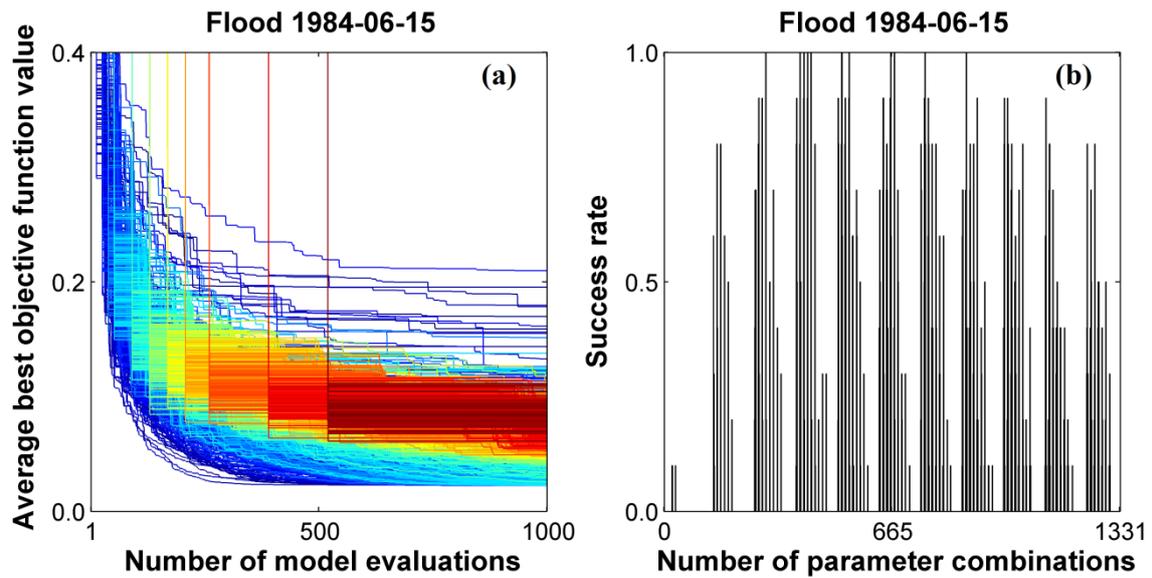
715  
 716 Fig. 3 Contributions of individual SCE-UA parameters and their interactions in terms of  
 717 convergence speed in benchmark function calibration. Each row represents a benchmark  
 718 function with 6, 8, 10 and 12 dimensions.  
 719



■ P 
 ■ beta 
 ■ alpha 
  Interaction

720 Fig. 4 Contributions of individual SCE-UA parameters and interactions in terms of success  
 721 rate in benchmark function calibration. Each row represents a benchmark function with 6, 8,  
 722 10 and 12 dimensions.  
 723

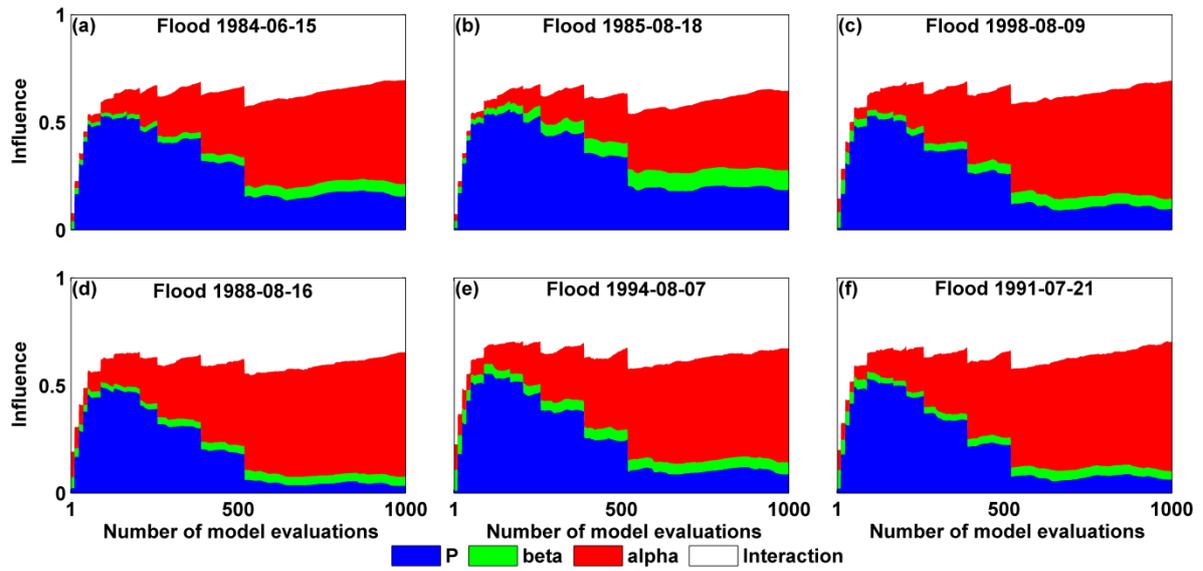
724  
 725



726

727 Fig. 5 Convergence speed and success rate variances for flood 1984-06-15, which are  
 728 generated from 1331 parameter combinations. Different convergence speed lines are  
 729 represented using different colors in Fig. 5a. The variations of the histogram heights in Fig.5b  
 730 represent the variance of success rate.

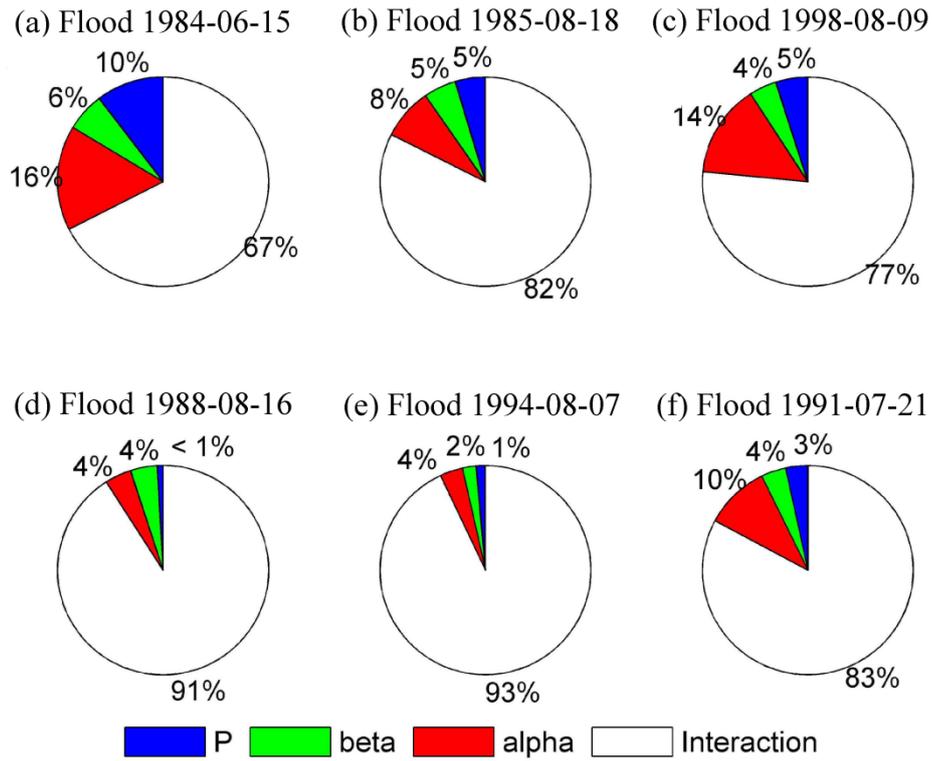
731



732  
 733 Fig. 6 Contributions of individual SCE-UA parameters and their interactions in terms of  
 734 convergence speed in TOPMODEL calibration. Each figure represents a flood calibration  
 735 problem.

736

737

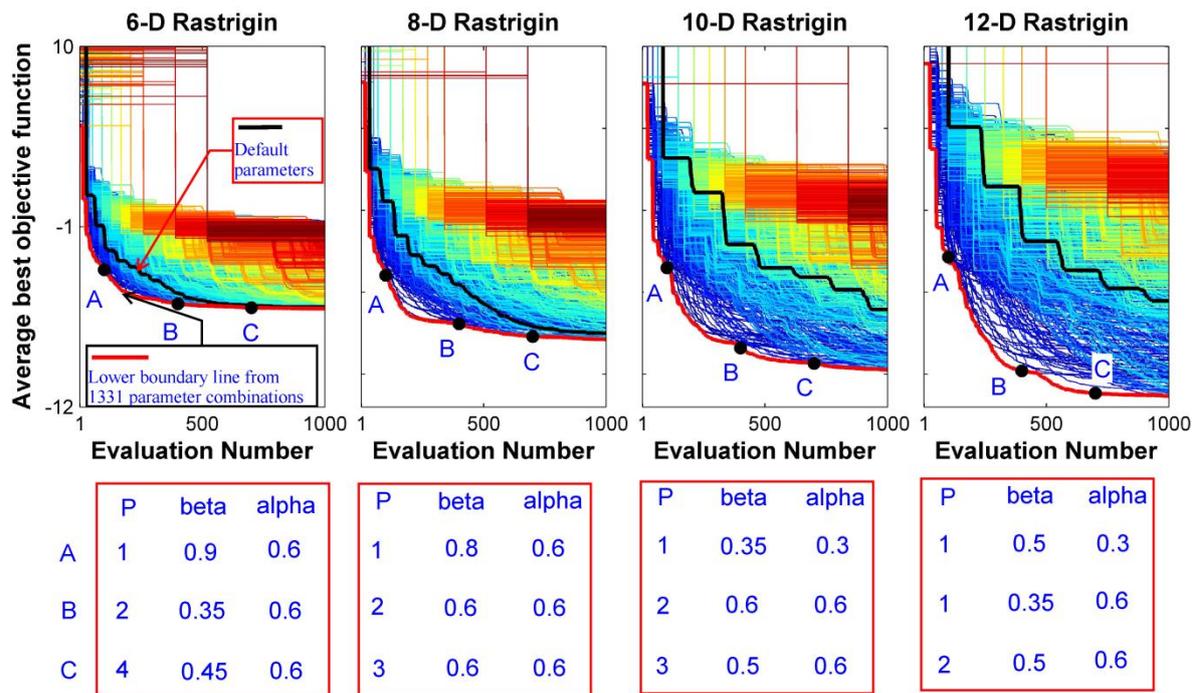


738

739 Fig. 7 Contributions of individual SCE-UA parameters and their interactions in terms of

740 success rate in TOPMODEL calibration. Each figure represents a flood calibration problem.

741



Parameter value combinations of points A, B and C from the lower boundray lines

742

743 Fig. 8 Comparison of the convergence speed curves (black bold line) from a set of default  
 744 parameter values suggested by Duan et al. (1994) and the lower convergence speed boundray  
 745 curves (red bold line) from the 1331 parameter combinations. Three points (A, B and C) from  
 746 the lower convergence speed boundray lines and their corresponding parameter values are  
 747 shown.