Int. J. Environ. Sci. Technol. (2014) 11:1563–1570
DOI 10.1007/s13762-013-0342-9

ORIGINAL PAPER

Effects of spatial data resolution on runoff predictions by the BASINS model

C.-L. Chang · Y.-C. Chao

Received: 22 February 2013 / Revised: 23 April 2013 / Accepted: 26 May 2013 / Published online: 11 June 2013

Abstract The BASINS model, developed by the United States EPA, is a popular simulation tool for predicting watershed responses, such as runoff, pollution exports, and water quality. It requires large amounts of data to set parameters. Many studies state that model input is a major source of model uncertainty. Thus, improvements to the quality and completeness of the data will improve the certainty of the model. The objective of this study is to discuss the effects of spatial data, including digital elevation models (DEMs) and spatial rainfall records, on predictions of runoff from the BASINS model. The result shows that both DEMs and rainfall data can significantly influence peak flow and runoff volume. Rainfall input has more influence on the curve shape of hydrograph than DEM resolution. DEM resolution can have more impact on peak flow predictions than rainfall input. Because the model uncertainties from DEMs and rainfall records influence each other, the prediction error does not always decrease when DEM resolution increases. The present results show that the BASINS model produces reliable answers in the case area when the grid size is less than 100 m $\times$ 100 m and the precipitation records from the Bihu Rainfall Station are correct and complete.

Keywords BASINS · Runoff · Spatial data · Uncertainty

C.-L. Chang (✉) · Y.-C. Chao
Department of Water Resources Engineering and Conservation, Feng Chia University, No. 100 Wenhwa Rd., Seatwen, Taichung 40724, Taiwan
e-mail: f89541201@ntu.edu.tw

Introduction

Simulation plays an important role in environmental management. Most water resource management strategies depend on hydrologic simulations. Runoff prediction can help us understand the risks of floods and droughts. To avoid wasteful and misguided environmental strategies, we must improve the reliability of our models (Lung 2001). The sources of model uncertainty include model structure, observations, model inputs, initial values, and boundary conditions (Troutman 1983; Klepper 1997). Model structure should enable the description of relevant environmental phenomena. Initial values and boundary conditions are usually determined by calibration and validation of parameters. Model inputs should accurately represent environmental properties. This study considers model uncertainties caused by model inputs; in particular, this study considers a model of runoff in the Fei-tsui reservoir watershed.

Most inputs for environmental models involve spatial variation, and hydrologic models are particularly sensitive to spatial variation of properties such as rainfall. The reliability of simulation results for an environmental model depends on how well it represents the spatial variability of environmental properties (Lopes 1996; Chaubey et al. 1999; Chaplot 2005). When measured records are limited, data interpolation plays a significant role in improvement of model inputs and presentation of the spatial variability of environmental properties (Bartier and Keller 1996; Nalder and Wein 1998; Price et al. 2000). Numerous precipitation interpolation methods, such as the Kriging method, the optimal interpolation method, and the weighted method, have been developed and are commonly applied to estimate precipitation and to present spatial variation of rainfall when actual rainfall records are limited

(Dirks et al. 1998; Chang et al. 2005, 2006). Interpolation errors can seriously impair the simulation results of spatially distributed models (Donald and Danny 1996).

Digital elevation models (DEMs) can represent topography and stream networks in a watershed. The resolution of DEMs varies with the accuracy of watershed delineation. To map a single area, a low-resolution DEM must have a small number of large subbasins, whereas a high-resolution DEM must have a large number of small subbasins. Typical simulation practice is to assume that all properties are homogeneous in a subbasin. If a single subbasin covers a wide area, however, this assumption is incorrect because terrain varies within that subbasin (Jha et al. 2004). The effect of subbasin scaling on a watershed simulation is related to the sources of heterogeneity, which include rainfall characteristics, land-use, soil, and topography (Arnold et al. 1998).

Although interpolation methods can be applied to estimate spatial data when measured data are insufficient, they cannot avoid the uncertainty caused by interpolation errors. On the other hand, simplified watershed delineation can also result in simulation errors due to assumptions of homogeneity (Kalin et al. 2003; Chaubey et al. 2005; Chang 2009). This study focuses on DEMs and spatial rainfall data in order to assess the effects of spatial data on the prediction of hydrologic responses in a watershed. Small uncertainties from spatial data tend to exacerbate each other. A high-resolution DEM can divide a watershed into many subbasins. It can minimize the uncertainty due to assumptions of homogeneity. However, because each subbasin within a high-resolution DEM requires complex data preparation and model calculation, a high-resolution DEM necessarily has high uncertainty from its model inputs (Bingner et al. 1997; Gandolfi and Bischetti 1997). The Daiyuku Creek and the Qupoliao Creek in the Fei-tsui reservoir watershed provided the setting for this case study.

Materials and methods

Site description

There are six rainfall gauging stations in the Fei-tsui reservoir watershed, but only one rainfall station—the Bihu station—is located in the case area. The area immediately around the Bihu station, labeled P4 in Fig. 1, is the case area. The Jiuqionggen station (labeled P2) and the Pinglin station (labeled P3) are outside the case area. The Daiyuku Creek and the Qupoliao Creek are both within the Fei-tsui reservoir watershed of Taiwan. The Fei-tsui reservoir supplies most domestic water used in northern Taiwan. The site area is about 79 km^2. There is more than 90 % forest in this area. One must choose some resolution for one's DEM before one can divide the case area into subbasins; the precipitation in a subbasin is assumed to be uniform and is given by a single value. The average precipitation for each subbasin is determined by the nearest rainfall station.

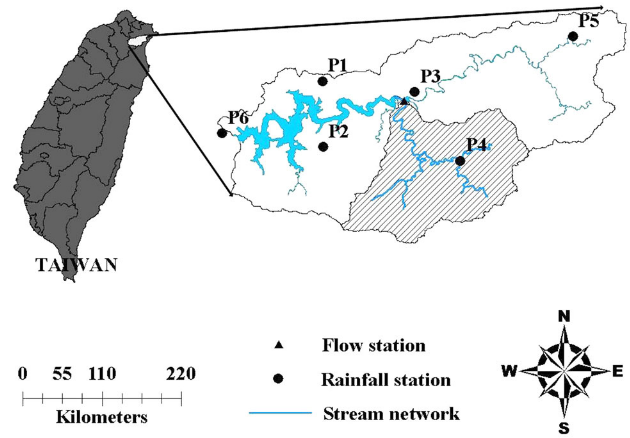


Fig. 1 Case area: Daiyuku Creek and Qupoliao Creek. *Note*: P1, Shisangu Rainfall Station; P2, Jiuqionggen Rainfall Station; P3, Pinglin Rainfall Station; P4, Bihu Rainfall Station; P5, Taiping Rainfall Station; P6, Feitsui Rainfall Station

BASINS model

The United States Environmental Protection Agency (EPA) developed the BASINS model as a geographically based watershed assessment tool that can simulate such things as stream flows, pollution exports, water quality, and the like. This study applies BASINS to predict hydrologic responses. BASINS is also an integrated decision-making system (Laroche et al. 1996; Jacomino and Fields 1997; Whittemore and Beebe 2000; Luzio et al. 2002; Albek et al. 2004; Hsieh and Yang 2006). Because most environmental data involve spatial variability, many studies use geographic information systems (GISs) (Helmlinger et al. 1993; Montgomery and Foufoula-Georgiou 1993; Chang and Lo 2006). The BASINS model itself can be integrated with a GIS, and most of the data can be analyzed within and extracted from that GIS. The BASINS software integrates several modules in a single-window interface. HSPF, one of these modules, can predict hydrologic responses and pollution exports in a watershed. This study applies HSPF to predict hydrologic responses, such as daily runoff, peak flow, and hydrographic response, in the case area and uses meteorological and geographical data from 2007 and 2008 for model calibration and validation, respectively.

Scenario design

Input-related model uncertainties are always caused by incorrect descriptions of the spatial properties of the

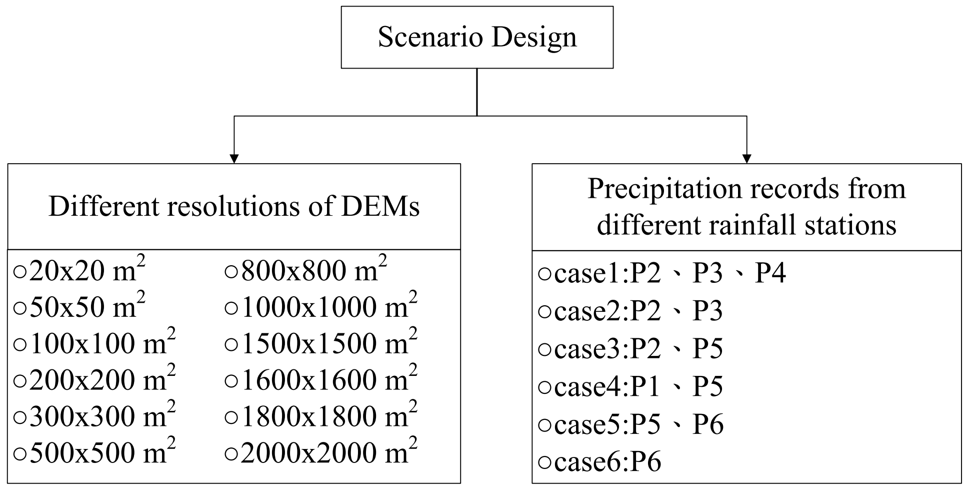


Fig. 2 Scenario design. *Note*: P1, Shisangu Rainfall Station; P2, Jiuqionggen Rainfall Station; P3, Pinglin Rainfall Station; P4, Bihu Rainfall Station; P5, Taiping Rainfall Station; P6, Feitsui Rainfall Station

environment. To assess the model uncertainties due to improper model inputs, this study discusses 72 scenarios with various DEM resolutions and rainfall inputs. DEM resolution is described in terms of grid size, from 20 m × 20 m to 2,000 m × 2,000 m. Twelve levels of DEM resolution and six levels rainfall input accuracy are combined to produce 72 scenarios. Figure 2 illustrates the treatment levels that are combined to make each scenario. The grid size influences the resolution of DEMs. The more land area a grid square contains, the lower the DEM resolution is. The completeness of rainfall records decreases from case 1 to case 6. Over-simplification of DEM resolution introduces one kind of model uncertainties; incomplete rainfall gauging records introduce different uncertainties, and the mutual influence of these two types of uncertainties is a noteworthy phenomenon.

Results and discussion

Model calibration and validation

The typical BASINS model requires many parameters, and various studies have discussed the sensitivity of parameters in BASINS models. Simulated hydrologic responses are more sensitive to the parameters INFILT (index to the infiltration capacity of the soil), LSUR (length of the assumed overland flow plane), UZSN (upper zone nominal storage), and LZS (initial lower zone storage) than to other parameters (Jacomino and Fields 1997; Al-Abed and Whiteley 2002). The scenario with grid size of 20 m × 20 m and rainfall records from the Jiuqionggen station, the Pinglin station, and Bihu station (case 1) were used for model calibration and validation. Our model was calibrated and validated as shown in Fig. 3. The process of parameter calibration decreases differences between predicted values and observed values. The R-squared (R^2) between predicted and observed runoff is about 0.7, and the absolute value of relative error of peak flow prediction is less than 7 %. Runoff predictions from this BASINS model are reliable when the quality and completeness of model inputs have been calibrated.

Uncertainty due to simplified DEMs

Table 1 shows the results of scenarios with complete rainfall records and various DEM resolutions. The differences in DEM resolution are the only factor that can cause different hydrologic responses in these scenarios. The results show that DEM resolution can greatly affect the predicted peak flow and the predicted runoff volume. For grid sizes smaller than 300 m × 300 m, the absolute value of relative error of peak flow prediction is less than 30 %. However, the prediction error of peak flow is very large when the grid size is larger than 500 m × 500 m.

Figure 4 displays stream networks in the case area, as depicted by models with different DEM resolutions. A large grid necessarily delineates the watershed in a very different fashion than a small grid. The boundaries of subbasins and stream networks cannot be identified by watershed models with grids larger than 500 m × 500 m. Models with large grids alter the boundary of the case area

Fig. 3 Results of model calibration and validation

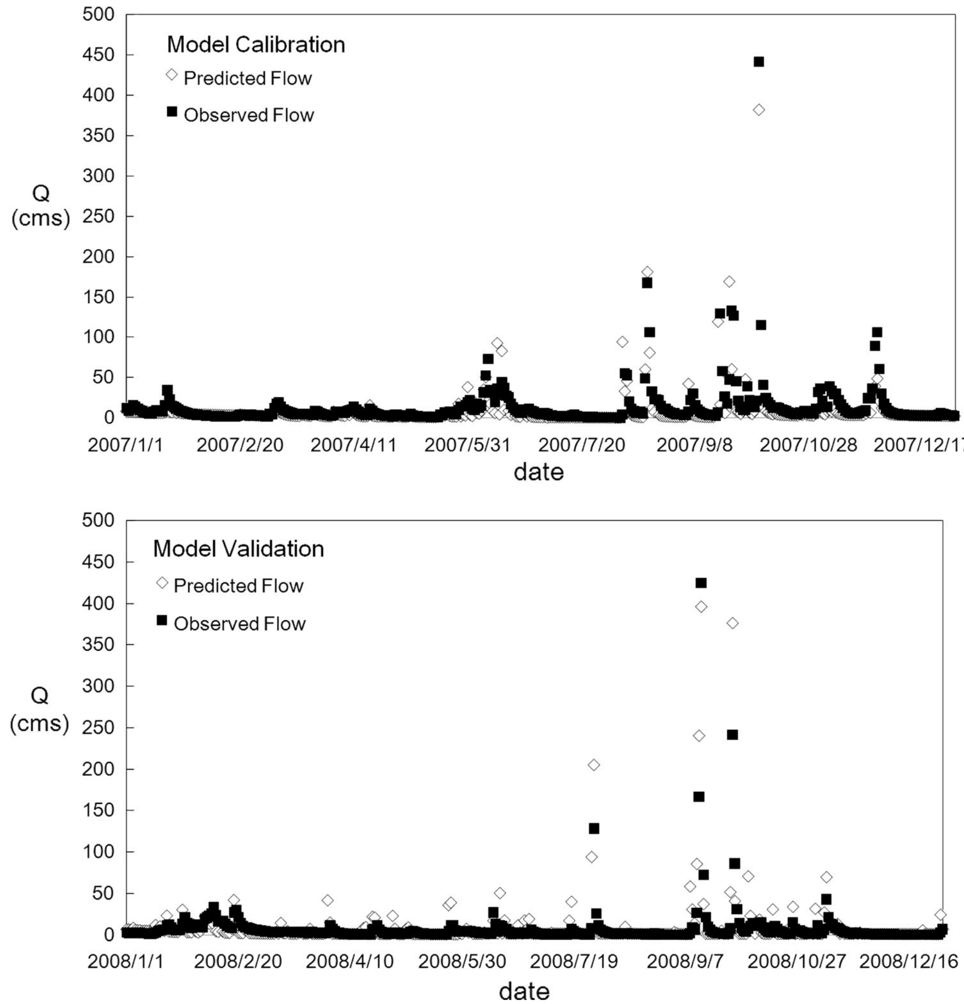


Table 1 Runoff predictions under scenarios with different DEM resolutions

Grid size (m^2)	Predicted peak flow (cm)	Predicted runoff volume (m^3)	Absolute value of relative error of peak flow prediction (%)	R^2 between predicted and observed runoff
20 × 20	396.43	274,466,487	6.67	0.84
50 × 50	382.27	264,752,206	10.00	0.85
100 × 100	353.96	270,007,890	16.67	0.80
200 × 200	297.32	210,291,459	30.00	0.82
300 × 300	317.15	225,548,160	25.33	0.82
500 × 500	154.33	110,106,115	63.67	0.82
800 × 800	12.86	9,009,674	96.97	0.83
1,000 × 1,000	13.73	9,112,919	96.77	0.87
1,500 × 1,500	41.91	27,494,357	90.13	0.87
1,600 × 1,600	20.93	14,926,172	95.07	0.82
1,800 × 1,800	41.63	29,245,354	90.20	0.83
2,000 × 2,000	33.41	24,234,081	92.13	0.81

and underestimate both the peak flow and the runoff volume. The predicted peak flows for simulations with grid sizes of 20 m × 20 m and 2,000 m × 2,000 m are about 396 and 33 cm, respectively. These predicted peak flows differ by more than an order of magnitude.

A large R^2 value between predicted and observed runoff values indicates that the model accurately predicts the trend of hydrologic responses. The R^2 between predicted and observed runoff is always larger than 0.8 when the rainfall inputs are complete. Our BASINS model can simulate hydrograph curves accurately, although the error of predicted peak flow is large for scenarios with low DEM resolutions.

Uncertainty due to incomplete rainfall records

When DEM resolution is held constant at 20 m × 20 m and different rainfall levels are considered, striking hydrologic responses can be seen, as shown in Table 2. Rainfall records can also influence the predicted peak flow and the predicted runoff volume. For peak flow predictions, however, the effect of rainfall is less than that of DEM resolution.

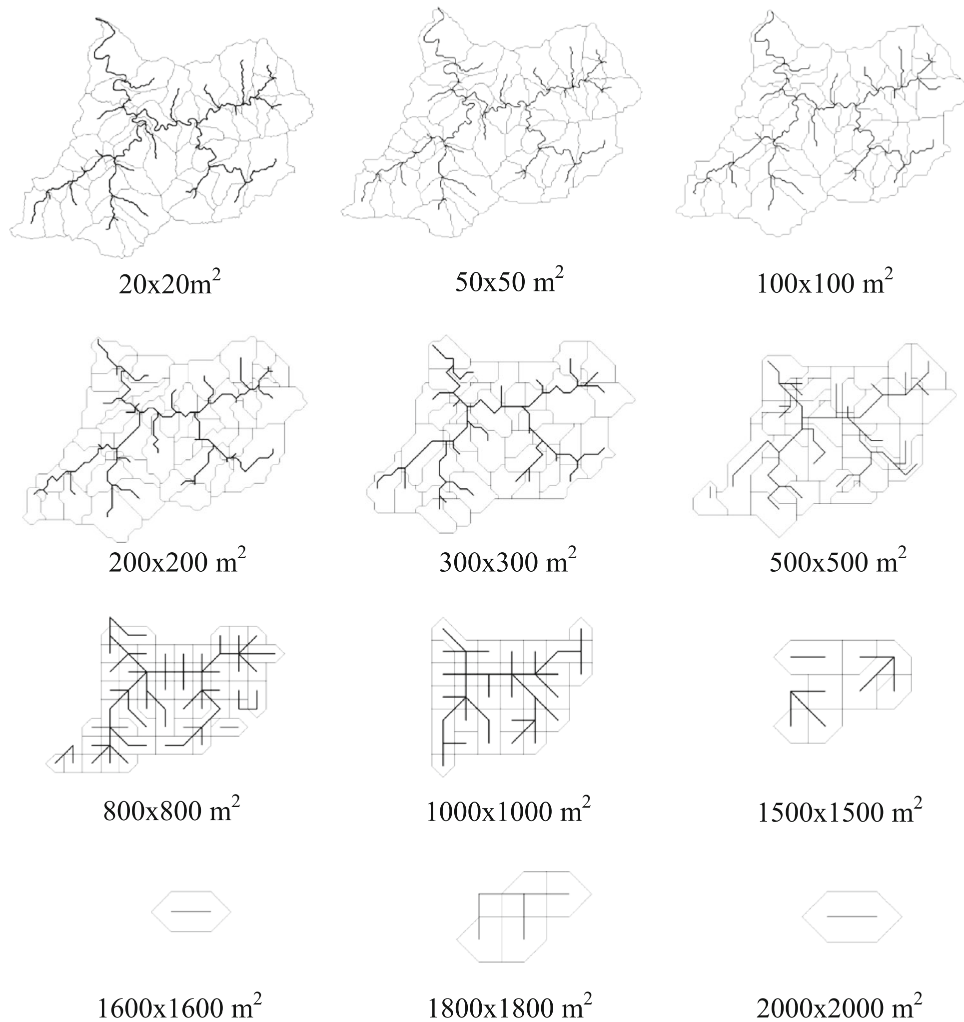


Fig. 4 Stream networks delineated by different DEM resolutions

Table 2 Runoff predictions under scenarios with different rainfall inputs

Case	Rainfall inputs	Predicted peak flow (cm)	Predicted runoff volume (m^3)	Absolute value of relative error of peak flow prediction (%)	R^2 between predicted and observed runoff
Case 1	P2, P3, P4	396.4	274,466,487	6.67	0.84
Case 2	P2, P3	294.5	248,138,402	30.67	0.77
Case 3	P2, P5	194.3	237,718,046	54.27	0.63
Case 4	P1, P5	230.2	245,984,458	45.80	0.61
Case 5	P5, P6	188.9	359,328,340	55.53	0.51
Case 6	P6	192.6	387,131,941	54.67	0.65

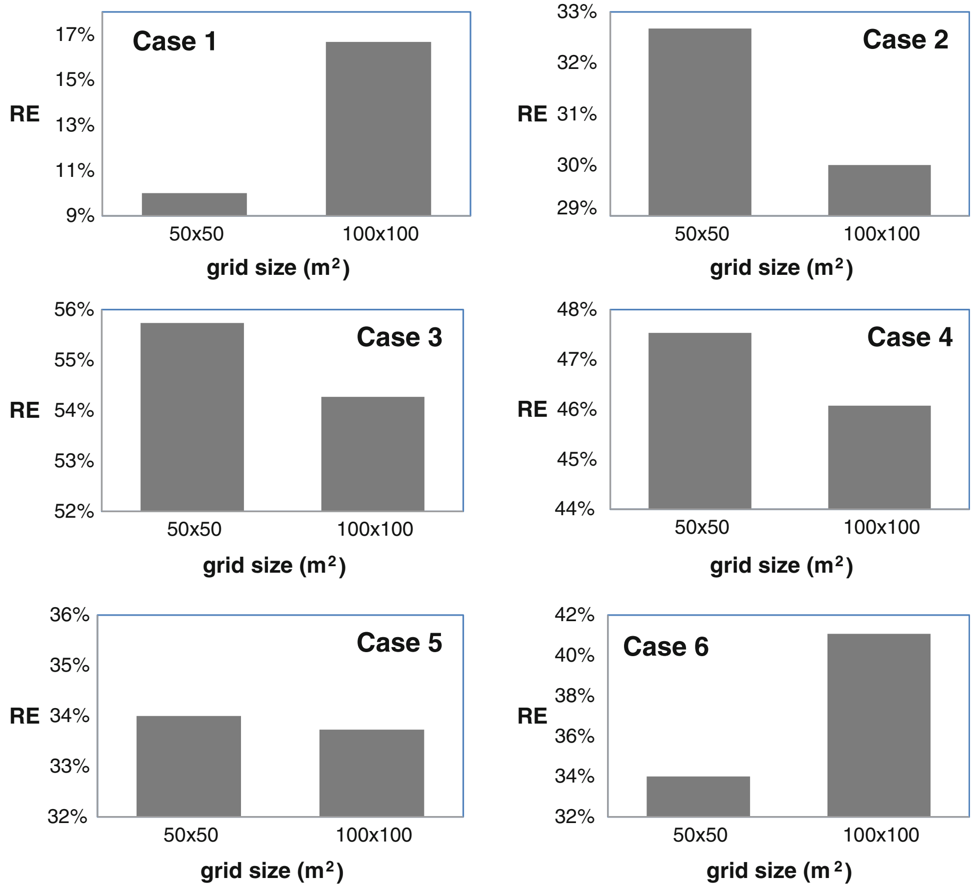
P1, Shisangu Rainfall Station; P2, Jiuqionggen Rainfall Station; P3, Pinglin Rainfall Station; P4, Bihu Rainfall Station; P5, Taiping Rainfall Station; P6, Feitsui Rainfall Station

The rainfall records at the Bihu Rainfall Station are very important. When the Bihu Rainfall Station records are not used, the absolute value of relative error of predicted peak flow is larger than 30 %. When the records from the Bihu and Pinglin stations are both omitted, the absolute value of relative error of predicted peak flow is larger than 50 %, due to interpolation errors. The R^2 between predicted and observed runoff is also influenced by rainfall inputs. When rainfall inputs cannot represent rainfall properties in the case area, both peak flow and the hydrograph curve cannot be simulated accurately by our BASINS model.

Mutual uncertainties from model inputs

Uncertainties from inadequate rainfall inputs and DEM resolutions influence each other. Figure 5 shows mutual uncertainties from inputs that include spatial rainfall data and DEMs. When the rainfall data are complete and can sufficiently represent the rainfall properties in the case area (case 1), the absolute values of relative error of predicted peak flow are about 10 and 17 % for the scenarios with grid sizes of 50 m × 50 m and 100 m × 100 m, respectively. This result shows that the simulation error of hydrologic responses can be lowered when the DEM resolution is increased.

Fig. 5 Mutual uncertainties from model inputs



When the rainfall records from some stations are omitted, the inputs cannot describe the spatial variability of rainfall in the case area (case 2 to case 6). The scenarios without sufficient rainfall inputs have such complex uncertainties that reliability levels are not improved even when DEM resolutions are increased. For example, the absolute values of relative error of predicted peak flow are about 33 and 30 % for scenarios with 50 m × 50 m and 100 m × 100 m grids, respectively, when the rainfall records from the Bihu station are omitted (case 2).

Conclusion

Researchers must understand the sources of uncertainty in models and simulations. This study discusses the influences of spatial data, including DEMs and spatial rainfall records, on simulated hydrologic responses. Both rainfall inputs and DEM resolution can be sources of model uncertainty. If DEM resolution is low or if rainfall information is incomplete, the BASINS model cannot accurately simulate peak flows and runoff volumes.

Rainfall input has more influence on the curve shape of hydrograph than DEM resolution. DEM resolution can have more impact on peak flow predictions than rainfall input. The uncertainties from insufficient DEM resolution and rainfall inaccuracy tend to influence each other; thus, a simulation with high DEM resolution can still have gross errors. For reliable hydrologic predictions with the BASINS model in this case area, it is recommended that the grid squares should be smaller than 100 m × 100 m, and complete precipitation records from the central rainfall station (the Bihu station) should be included.

Acknowledgments The authors would like to thank the National Science Council of the Republic of China for financially supporting this research under Contract No. NSC 97-2221-E-035-092-MY3.

References

Al-Abed NA, Whiteley HR (2002) Calibration of the hydrological simulation program fortran (HSPF) model using automatic calibration and geographical information systems. Hydrol Process 16:3169–3188

Albek M, Ogutveren UB, Albek E (2004) Hydrological modeling of Seydi Suyu watershed (Turkey) with HSPF. J Hydrol 285:260–271

Arnold JG, Srinivasan R, Muttiah RS, Williams JR (1998) Large area hydrologic modeling and assessment. Part I: model development. J Am Water Resour Assoc 34(1):73–89

Bartier PM, Keller CP (1996) Multivariate interpolation to incorporate thematic surface data using inverse distance weighting (IDW). Comput Geosci 22(7):795–799

Bingner RL, Garbrecht J, Arnold JG, Srinvasan R (1997) Effect of watershed subdivision on simulation runoff and fine sediment yield. Trans ASAE 40(5):1329–1335

Chang CL (2009) The impact of watershed delineation on hydrology and water quality simulation. Environ Monit Assess 148:159–165

Chang CL, Lo SL (2006) Role of geographic information system (GIS) in watershed simulation by WinVAST model. Environ Monit Assess 121(1–3):287–299

Chang CL, Lo SL, Yu SL (2005) Applying fuzzy theory and genetic algorithm to interpolate precipitation. J Hydrol 314:92–104

Chang CL, Lo SL, Yu SL (2006) The parameter optimization in the inverse distance method by genetic algorithm for estimating precipitation. Environ Monit Assess 117(1–3):145–155

Chaplot V (2005) Impact of DEM mesh size and soil map scale on SWAT runoff, sediment, and NO_3–N loads predictions. J Hydrol 312:207–222

Chaubey I, Haan CT, Grunwald S, Salisbury JM (1999) Uncertainty in the model parameters due to spatial variability of rainfall. J Hydrol 220:48–61

Chaubey I, Cotter AS, Costello TA, Soerens TS (2005) Effect of DEM data resolution on SWAT output uncertainty. Hydrol Process 19:621–628

Dirks KN, Hay JE, Stow CD, Harris D (1998) High-resolution studies of rainfall on Norfolk island, part I: interpolation of rainfall data. J Hydrol 208:187–193

Donald LP, Danny GM (1996) Spatial uncertainty analysis: propagation of interpolation errors in spatially distributed models. Ecol Model 91:213–229

Gandolfi C, Bischetti GB (1997) Influence of the drainage network identification method on geomorphological properties and hydrological response. Hydrol Process 11:353–375

Helmlinger KR, Kumar P, Foufoula-Georgiou E (1993) On the use of digital elevation model data for Hortonian and fractal analyses of channel networks. Water Resour Res 29:2599–2613

Hsieh CD, Yang WF (2006) Study of total maximum daily load and nonpoint source pollution control strategies for reservoir watershed. Pract Period Hazard Toxic Radioact Waste Manag 10(3):134–143

Jacomino VMF, Fields DE (1997) A critical approach to the calibration of a watershed model. J Am Water Resour Assoc 33(1):143–154

Jha M, Gassman PW, Secchi S, Gu R, Arnold J (2004) Effect of watershed subdivision on SWAT flow, sediment, and nutrient predictions. J Am Water Resour Assoc 40(3):811–825

Kalin L, Govindaraju RS, Hantush MM (2003) Effect of geomorphologic resolution on modeling of runoff hydrograph and sedimentograph over small watersheds. J Hydrol 276:89–111

Klepper O (1997) Multivariate aspects of model uncertainty analysis: tools for sensitivity analysis and calibration. Ecol Model 101:1–13

Laroche AM, Gallichand J, Lagace R, Pesant A (1996) Simulating Atrazine Transpot with HSPF in an Agricultural Watershed. J Environ Eng 122:622–630

Lopes VL (1996) On the effect of uncertainty in spatial distribution of rainfall on catchment modeling. Catena 28:107–119

Lung WS (2001) Water quality modeling for wasteload allocations and TMDLs. Wiley, New York

Luzio MD, Srinivasan R, Arnold JG (2002) Integration of watershed tools and SWAT model into BASINS. J Am Water Resour Assoc 38(4):1127–1141

Montgomery DR, Foufoula-Georgiou E (1993) Channel network source representation using digital elevation models. Water Resour Res 25:1907–1918

Nalder IA, Wein RW (1998) Spatial interpolation of climatic normals: test of a new method in the Canadian boreal forest. Agric For Meteorol 92(4):211–225

Price DT, Mckenny DW, Nalder IA, Hutchinson MF, Kesteven JL (2000) A comparison of two statistical methods for spatial interpolation of Canadian monthly mean climate data. Agric For Meteorol 101:81–94

Troutman BM (1983) Runoff prediction errors and bias in parameter estimation induced by spatial variability of precipitation. Water Resour Res 19(3):791–810

Whittemore RC, Beebe J (2000) EPA's BASINS model: good science of Serendipitious modeling. J Am Water Resour Assoc 36(3): 493–499