Advances in Geosciences

Impacts of spatial data resolution on simulated discharge, a case study of Xitiaoxi catchment in South China

G. J. Zhao[1]**, G. Hörmann**[1]**, N. Fohrer**[1]**, and J. F. Gao**[2]

[1]Department of Hydrology and Water Resources Management, Ecology Centre, Kiel University, Germany
[2]Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences, Nanjing, China

Abstract. In this paper we analyse the effects of different spatial input data resolution on water balance simulation using a simple distributed hydrological model: PCR-XAJ model. A data set consisting of land use and digital elevation model at 25 m resolution of Xitiaoxi catchment in South China is used for investigation. The model was first calibrated and validated at 50m cell size, thereafter an aggregation of the digital elevation model (DEM) and land use maps at 100 m, 200 m, 300 m, 500 m and 1 km are applied to evaluate the effects of spatial data resolution on simulated discharge. The simulation results at a grid size of 50 m show a good correlation between measured and simulated daily flows at Hengtangcun station with Nash-Suttcliffe efficiency larger than 0.75 for both calibration and validation periods. In contrast, the model performs slightly worse at Fanjiacun station. The increasing grid size affects the characteristics of the slope and land use aggregation and causes important information loss. The aggregation of input data does not lead to significant errors up to a grid of 1 km. Model efficiencies decrease slightly with cell size increasing, and more significantly up to the grid size of 1 km.

1 Introduction

Advanced techniques in remote sensing, geographic information systems and computer have been widely applied to distributed hydrological models in recent years. Thereby a number of large spatial data sets are employed in spatial distributed hydrological modeling. Different spatial resolution of input data can represent the heterogeneity of landscape to some extent, which may have a significant impact on the simulation results (Blöschl et al., 1997). Thus, an appropriate spatial resolution for hydrological modeling should be considered carefully (Grayson and Blöschl, 2000).

There are numerous studies in literature, which investigated the effects of using different spatial resolution data on the results of hydrological modeling (Blöschl, 2001). Considerable research on how grid size affects the topographic characteristics, wetness index and outflow has been carried out with TOPMODEL (Beven and Kirkby, 1979; Quinn et al., 1991; Moore et al., 1993; Bormann, 2006; Wu et al., 2008). In general, spatial input data with higher resolution led better simulation results. However, the smaller the grid size, the higher is the amount of spatial information. Reduction of the grid size also means much more computational time and a tremendous increase in the work for data collection and processing. Although some authors reported that grid size can directly affect the simulation results, most research focused on the topological indices variation and their effects on discharge and sediments based on TOPMODEL. Only a few studies analyzed the impacts of spatial data resolution by using other models (Wechsler, 2007), such as the SWAT model (Chaplot, 2005; Chaubey et al., 2005; Haverkamp et al., 2005) and the Agricultural Nonpoint Source Pollution (AGNPS) (Vieux and Needham, 1993).

The objective of this paper is to develop a raster-based hydrological model: the PCR-XAJ model and assess the effects of different spatial data resolution on discharge simulation. The Xitiaoxi catchment, a humid semitropical catchment, was selected for this study. The PCR-XAJ model calculates the water balance in both the mountainous and flat sub catchments.

2 Study area

The Xitiaoxi catchment is located in the upper reaches of Tai Lake basin (Fig. 1), in South China. It covers more than 2200 km^2. In the Xitiaoxi catchment 63.4% of the mountains

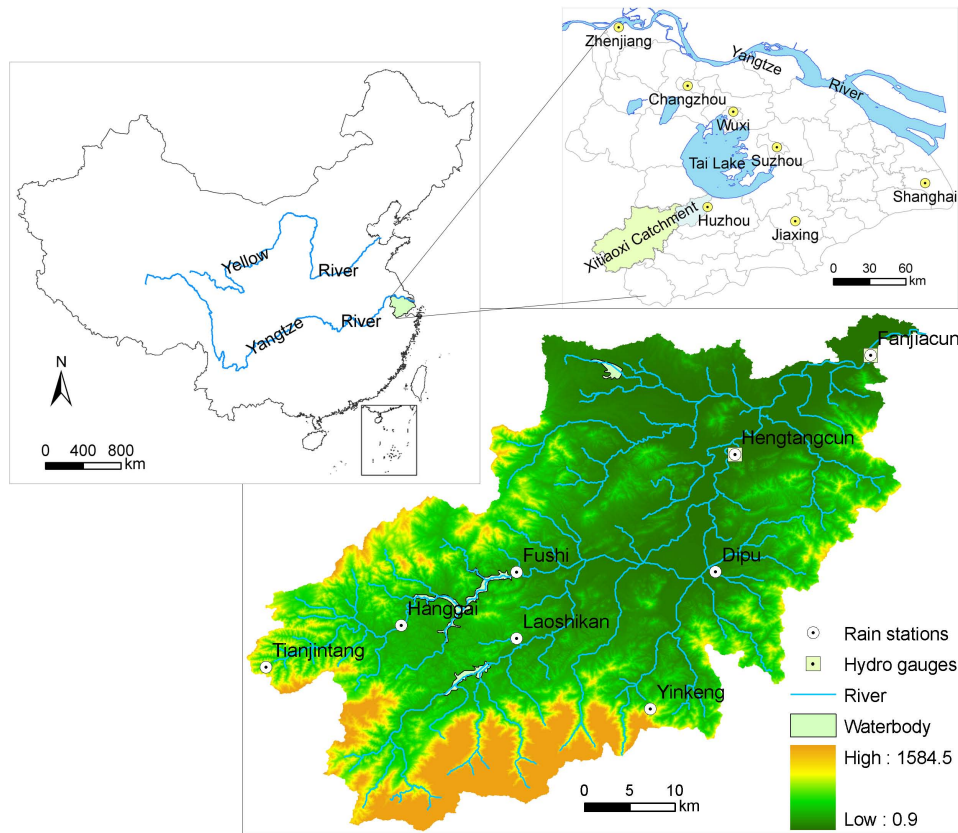*Correspondence to:* G. J. Zhao
(gzhao@hydrology.uni-kiel.de)

Fig. 1. Location of the study area and rainfall gauges.
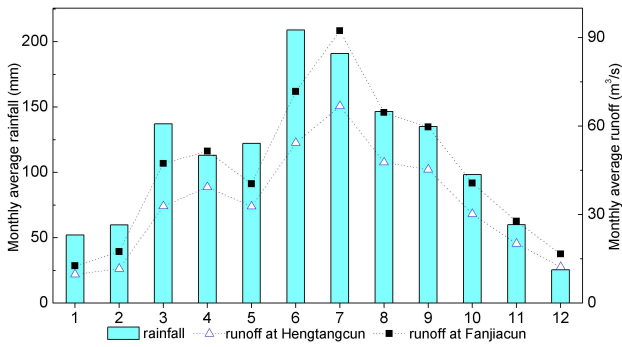

Fig. 2. Monthly precipitation and discharge at two stations in Xitiaoxi catchment (1979–1988).

in the upper reaches areas are covered with forest. About three quarters of this area is bamboos. About 25% of the catchment is paddy land, which lies in the low alluvial plains. A portion of about 4% of the area is fallow land, 1.8% is covered by urban area and the other land uses are grassland and bare land. Two reservoirs (Fushi and Laoshikan) located in the upper reaches of the catchment are primarily used for

flood control in rainy season. There are 75 polders in the lower reaches of the Xitiaoxi catchment, which are enclosed by embankments and form artificial hydrological entities.

The catchment is characterized by a semitropical climate with mean annual rainfall of about 1465 mm. The distribution of river runoff in the Xitiaoxi catchment is mainly controlled by rainfall (Fig. 2), which is dominated by the Asian summer monsoon.

Detailed climate and hydrological data collected from 1978 to 1988 for the Xitiaoxi catchment are provided by Huzhou Hydrological Bureau, Zhejiang Province. Daily precipitation data sets are available for 8 rain gauging stations within the Xitiaoxi catchment (Fig. 1). Among these stations, the Fanjiacun station and Hengtangcun station are also the streamflow gauging stations with continuous streamflow records. The land use and land cover maps for the year 2001 and DEM at 25×25 m horizontal resolution for the catchment are provided by Huzhou Bureau of surveying and mapping, Zhejiang Province.

To estimate the impact of the spatial data resolution on simulated catchment discharge, the available data set of 25m resolution was aggregated using standard GIS functions to create grid based data sets of increasing grid size: 50 m, 100 m, 200 m, 300 m, 500 m and 1 km. Thereafter, the spatial
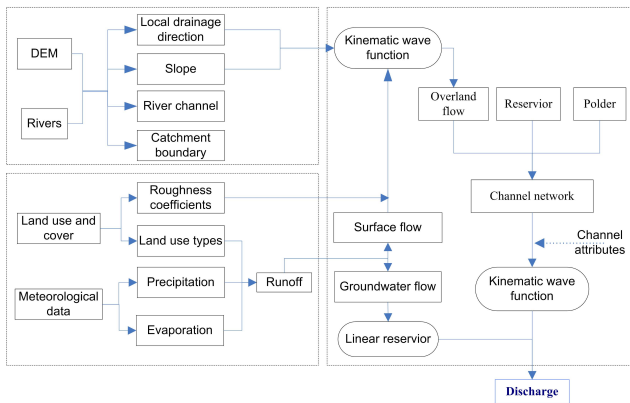
Fig. 3. Structure of the PCR-XAJ model.

data sets (e.g. land use, channel) were systematically aggregated applying the same aggregation methods. Subsequently, the PCR-XAJ model was applied to investigate the impacts of data aggregation.

3 Hydrological modelling

The simple raster-based PCR-XAJ model is so called because it is implemented within PCRaster and conceptually based on Xinanjiang model (Wesseling et al., 1996). Figure 3 shows its structure. The model simulation is based on grid calculation. Once the grid size is fixed, all the maps will be calculated at this scale for daily step. As shown in Fig. 3, DEM and river channels are used to create a local drainage direction map according to the D8 algorithm (O'Callaghan and Mark, 1984), which calculates the water flow directions. The actual evapotranspiration is calculated based on the concept of one layer evaporation method in Xinanjiang model (Zhao, 1992). Precipitation and evaporation time series are interpolated with inverse-distance weighting (IDW) to predict daily discharge. The land uses were reclassified in four types i.e. paddy, water-body, urban area and forests for runoff calculation. As a simplification, the relatively small areas of grassland and bare land are assumed to behave in a similar way as forests regarding runoff formation. The runoff generation component of the Xinanjiang model (Zhao, 1992) is used to estimate the surface runoff and groundwater in the catchment for each grid. Both the overland flow and channel flow are calculated by kinematic wave equation. The polders and reservoirs are considered as points in the simulation (van der Knijff and de Roo, 2008).

The Xinanjiang model has been widely applied in humid and semi-humid areas in South China (Zhao, 1992). The Xinanjiang model is a well-known lumped model, characterized by the concept of runoff generation on repletion of storage, which means that runoff is not generated until the soil moisture content of the aeration zone reaches maximum

capacity, and thereafter runoff equals the rainfall excess without further loss (Su et al., 2003). The runoff is separated into two components: the surface runoff and groundwater, according to their generating levels in the vertical profile. More details were described by Zhao et al. (1980, 1992).

The model was calibrated and validated for the whole catchment at both Hengtangcun (1307.6 km^2) and Fanjiacun (1913.5 km^2) stations at the 50 m grid size due to a limitation in the number of computational units by using PCRaster. Calibration period at Hengtangcun station is from 1979 to 1983, Fanjiacun from 1980 to 1983, and the validation periods are both between 1984 and 1988. For all different grid sizes (100 m, 200 m, 300 m, 500 m, 1 km) derived from the original data sets, continuous water balance simulations from 1978 to 1988 were performed without a recalibration of the simple model. All the spatial data (DEM, land uses and their derivative data such as slope, Manning's roughness) were adapted to the same resolution correspondingly.

4 Results and discussions

4.1 Model performance

The model efficiencies according to the Nash-Suttcliffe index (NS) (1970), the correlation coefficient (R^2) and the root mean square errors (RMSE) were calculated at daily resolution to evaluate the hydrological model performance.

Figure 4 shows the hydrographs of the measured and simulated discharge at Hengtangcun and Fanjiacun stations. Both of them illustrate that some peaks of the modeled values are much higher than the measured discharge, while the simulated values are lower after the peaks. This may be caused by the two reservoirs, which are located in the upper reaches of the catchment. The reservoirs are used for irrigation during dry season and flood control during rainy season. This also explains some missing peaks in the modeled data during dry season.

Table 1 shows model efficiencies for calibration and validation at the two gauges. The calibration results show a good correlation between measured and simulated daily flows at Hengtangcun station. This is demonstrated by the correlation coefficient ($R^2 = 0.82$) and Nash-Suttcliffe simulation efficiency (NS = 0.77) values. For validation, the NS was found to be 0.81 and R^2 0.86, which is in a very good agreement with the observed discharge, and the RMSE is much lower than that for the calibration period. Comparably, the simulated results at Fanjiacun station are slightly worse than at Hengtangcun station. The correlation efficiencies for daily discharge are of moderate quality (0.71 for calibration, 0.76 for validation). The NS values of 0.62 and 0.67 during the two periods, though relatively lower, are acceptable as these values are larger than 0.5 (Santhi et al., 2001a). However, the values of the RMSE are much higher during the two periods at Fanjiacun station.
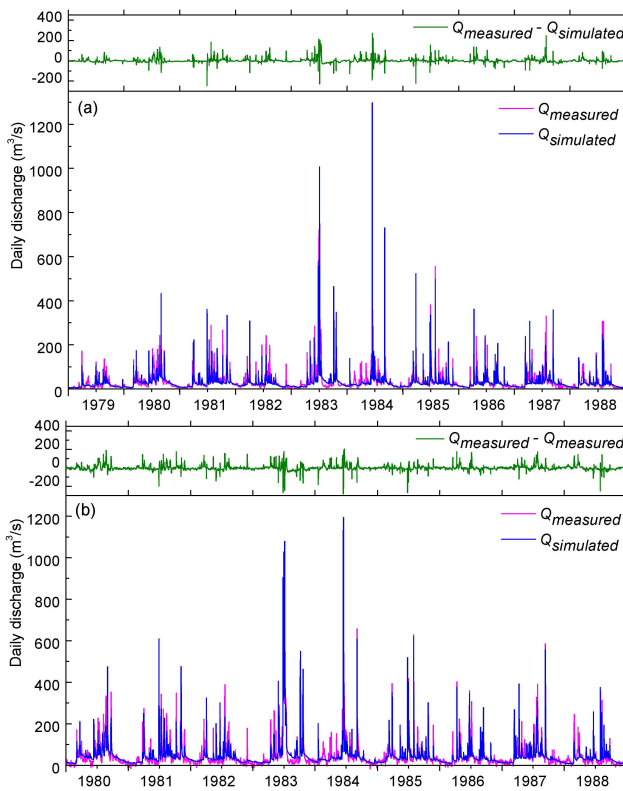
Fig. 4. Comparison of daily measured and modeled discharge in Xitiaoxi Catchment (**a**: Hengtangcun station; **b**: Fanjiacun station).

Table 1. Model performance in Xitiaoxi catchment.

Station	Used for	Period	NS	R^2	RMSE (m^3/s)
Hengtangcun	calibration	1979–1983	0.77	0.82	27.31
	validation	1984–1988	0.81	0.86	18.32
Fanjiacun	calibration	1980–1983	0.62	0.71	37.51
	validation	1984–1988	0.67	0.76	33.97

The polders connect with outside streams through man-operated devices, and the water fluxes are adjusted manually. These artificial hydrological entities are used for agricultural irrigation and flood control, which is a great challenge for water flux calculation in hydrological simulation. Additionally, the return flow from Tai Lake leads to some negative discharge values at Fanjiacun station, which can not be estimated by our hydrological model. Over the years 1980 to 1988, there is a negative discharge occurring at an average of 15.2 days per year. This may causes the higher RMSE in Fanjiacun station.

The model efficiencies at Fanjiacun station are slightly worse; however, the simulation results demonstrate that the simple hydrological model can successfully simulate water balance at regional scale in the wet subtropical area based on Xinanjiang concepts.
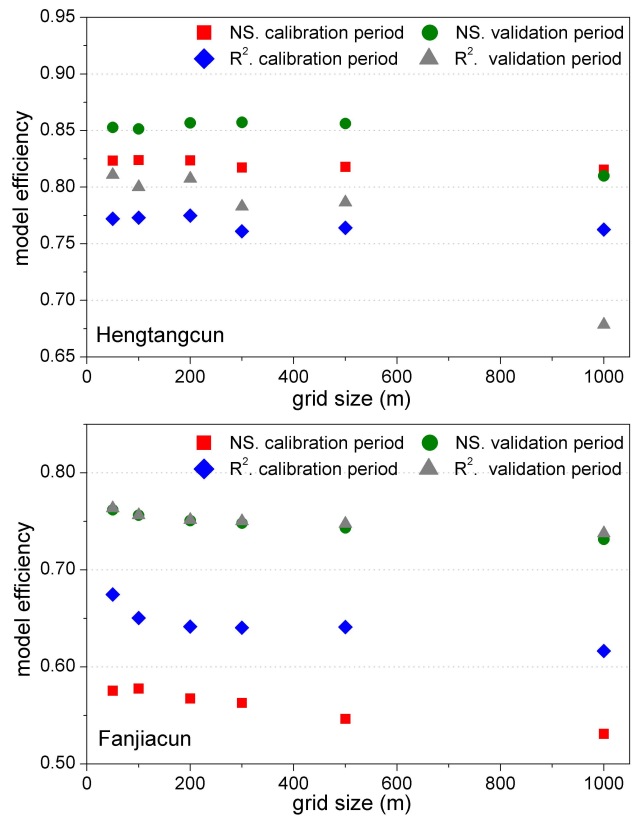

Fig. 5. Model efficiencies with different spatial resolution at two gauging stations in Xitiaoxi catchment.

4.2 Effects of spatial resolution changes on simulated discharge

The evaluation results, shown in Fig. 5, reveal that model efficiencies do not decrease significantly for most of the grid sizes at both gauges. Up to a grid size of 500m the model efficiencies remain almost constant. At a grid size of 1km the simulation results change evidently at Hengtangcun station. As shown in Fig. 5, the correlation coefficient (R^2) during the validation period for the 1km grid size (0.68) is much lower than that for 500 m (0.76). In contrast, the model efficiencies at Fanjiacun station decrease slightly with increasing grid size, but the changes are not obvious (Fig. 5). In general, the increasing resolution of spatial input data can slightly improve the model results. The comparable results were also found by other researchers (Booij, 2005; Bormann, 2006). Additionally, our simulation results indicate that the effects of spatial resolution in the hilly region are more sensitive than in flat region. This may be attributed to the slope smooth.

Figure 6 reveals the annual runoff deviations between simulated and observed discharge in dry (1985), normal (1980) and wet (1983) years. It can be obviously seen that the
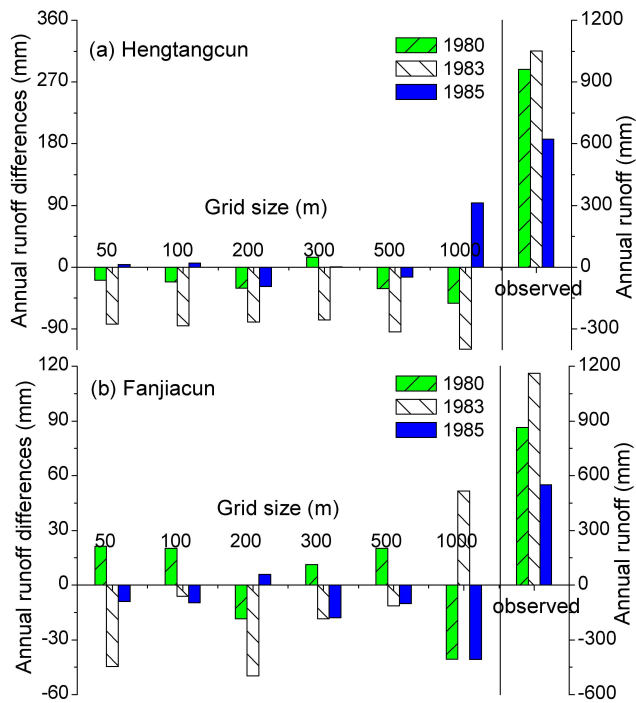
Fig. 6. Annual runoff deviations with different spatial resolution in dry (1985), normal (1980) and wet year (1983) at two stations.
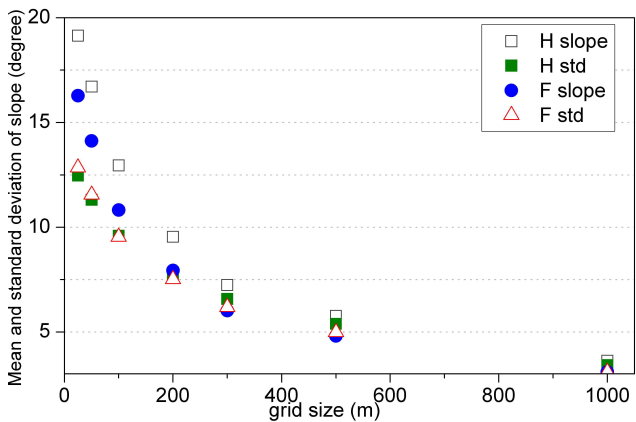

Fig. 7. Mean and standard deviation of slope in two sub catchments (H: Hengtangcun, F: Fanjiacun; std: standard deviation).

Fig. 8. Land use changes with different spatial resolution in two sub catchments (**a**: Hengtangcun; **b**: Fanjiacun).

simulated runoff is near 10% lower than the observed values at all grid size in 1983 at Hengtangcun station (Fig. 6a). In contrast, significant deviations only occur at 1km grid size in the other two years (Fig. 6a), and similar results can also be found at Fanjiacun station (Fig. 6b). In general, annual runoff deviations do not change significantly up to a grid size of 1km in both sub catchments.

The above analysis indicates that both the model efficiencies and simulated runoff changed with increasing grid size, whi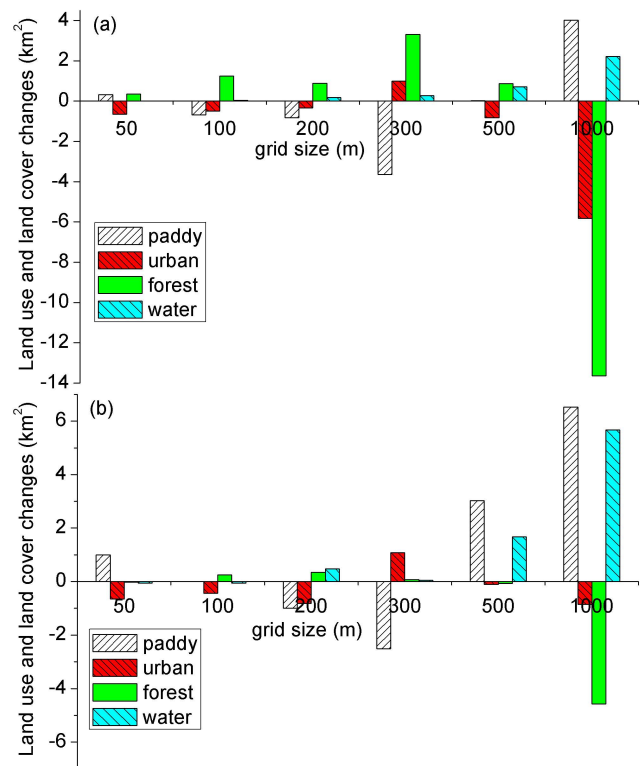ch may be caused by the slope smoothness and land use aggregation. Increasing the grid size leads to a smoothed surface of elevation and therefore to a decreased mean slope as well as the standard deviation of the slope (Fig. 7). Comparably, the slope and standard deviation in Hengtangcun sub catchment decrease much more significantly with grid size increasing. Consistent with previous finding from Chaubey et al. (2005), our research confirmed that a finer data resolution resulted in higher slope.

The effects of aggregation on the land use fractions are shown in Fig. 8. It indicates that the aggregations do not play an important role up to a grid size of 1 km in Hengtangcun sub catchment. Significant deviations for paddy and urban area, especially forest were found at the 1 km level which may have an influence on the model efficiencies (Fig. 8a). For Fanjiacun sub catchment, significant changes in land use fractions can be observed for paddy land at almost all grid sizes (Fig. 8b). In addition, the changes in most land use types are more significant at 1km grid size, which is consistent with the model efficiencies and annual runoff changes.

Although the slope smoothness and land use aggregation influence the simulation results and model efficiencies, other catchment properties (e.g. roughness coefficients, channel width) can additionally contribute to that. Thus, simulations results are related to all spatial data sets, and the data aggregation therefore would affect all the spatial data sources.

However, one difficult issue is to quantify the effects of these factors on simulation results. Bormann (2006) proposed a correlation analysis between statistics of input data and simulated annual water fluxes. He found that predominantly the correlation between catchment properties and simulated water flows varies from catchment to catchment, and catchment specific properties determine correlations between properties and fluxes, but do not influence the effect of data aggregation.

5   Conclusions

In this study, the PCR-XAJ model was applied in Xitiaoxi catchment for different spatial resolutions of all spatial input data. The simulation results present a good agreement with the observed values. The Nash-Sutcliffe index of 0.77 for calibration and 0.81 for validation period at Hengtangcun station is satisfactory. For Fanjiacun station, the results are slightly worse, which may be caused by the polders and return flow from Tai Lake in the lower reaches of Xitiaoxi catchment.

Hydrological simulation for two sub catchments was implemented for evaluation the effects of spatial data resolution. The results show that an aggregation of input data does not lead to significant errors up to a grid size of 1 km. Land use aggregation causes significant information loss from a grid size of 500 m to 1km, which leads to a large deviation in water balance simulation. Both mean slope and standard deviation decrease significantly with cell size increase in two sub catchments.

In general, this study shows that the higher the input data available, the better the simulation results can be obtained, but the trends are not always obvious. The model efficiencies do not get significantly worse up to spatial resolution at 1km cell size and the results also depends on the watershed response of interest.

Edited by: B. Schmalz, K. Bieger, and N. Fohrer
Reviewed by: two anonymous referees

References

Beven, K. J. and Kirkby, M. J.: A physically based, variable contributing area model of catchment hydrology, Hydrol. Sci., 24, 43–69, 1979.

Blöschl, G.: Scaling in hydrology, Hydrol. Proc., 15, 709–711, 2001.

Blöschl, G., Sivapalan, M., Gupta, V. K., and Beven K. J.: Scale Problems in Hydrology, Water Resour. Res., 33(12), 2881–2999, 1997.

Booij, M. J.: Impact of climate change on river flooding assessed with different spatial model resolutions, J. Hydrol., 303, 176–198, 2005.

Bormann, H.: Impact of spatial data resolution on simulated catchment water balances and mode performance of the multi-scale TOPLATS model, Hydrol. Earth Syst. Sci., 10, 165–179, 2006, http://www.hydrol-earth-syst-sci.net/10/165/2006/.

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

Chaubey, I., Cotter, A. S., Costello, T. A., and Soerens, T. S.: Effect of DEM data resolution on SWAT output uncertainty, Hydrol. Proc., 19, 621–628, 2005.

Grayson, R. B. and Blöschl, G. (Eds): Spatial Patterns in Catchment Hydrology: Observations and Modelling. Cambridge University Press, Cambridge, UK, 404 pp., 2000.

Haverkamp, S., Fohrer, N., and Frede, H. G.: Assessment of the effect of land use patterns on hydrologic landscape functions: a comprehensive GIS-based tool to minimize model uncertainty resulting from spatial aggregation, Hydrol. Proc., 19, 715–727, 2005.

Moore, I. D., Lewis, A., and Gallant, J. C.: Terrain attributes: Estimation methods and scale effects, in: Modeling change in environmental systems, edited by: Jakeman, A. J., Beck, M. B., and McAleer, M., Wiley, New York, USA, 189–214, 1993.

Nash, J. E. and Sutcliffe J. V.: River flow forecasting through conceptual models, part I: A discussion of principles, J. Hydrol., 10(3), 282–290, 1970.

O'Callaghan, J. F. and Mark, D. M.: The extraction of drainage networks from digital elevation data, Comput. Vis. Graph. Image Proc., 28, 323–344, 1984.

Quinn, P., Beven, K., Chevallier, P., and Planchon, O.: The prediction of hillslope flow paths for distributed hydrological modelling using digital terrain models, Hydrol. Proc., 5, 59–79, 1991.

Santhi, C., Arnold, J. G., Williams J. R., Dugas, W. A., Srinivasan, R., and Hauck, L. M.: Validation of the SWAT Model on a Large River Basin With Point and Nonpoint Sources, J. Am. Water Resour. Assoc., 37, 1169–1188, 2001.

Su, B. L., Kazama, S., Lu, M. J., and Sawamoto, M.: Development of a distributed hydrological model and its application to soil erosion simulation in a forested catchment during storm period, Hydrol. Proc., 17, 2811–2823, 2003.

Van der Knijff, J. and de Roo, A.: LISFLOOD Distributed Water Balance and Flood Simulation Model, Revised user manual, Office for Official Publications of the European Communities, Luxembourg, EUR 22166 EN/2, 2008.

Vieux, B. and Needham, H.: Nonpoint pollution model sensitivity to grid cell size, J. Water Resour. Plan. Manage., 119(2), 141–157, 1993.

Wechsler, S. P.: Uncertainties associated with digital elevation models for hydrologic applications: a review, Hydrol. Earth Syst. Sci, 11, 1481–1500, 2007.

Wesseling, C. G., Karssenberg, D. J., Burrough, P. A., and Van Deursen, W. P. A.: Integrated dynamic environmental models in GIS: The development of a Dynamic Modelling language, Transactions in GIS, 1-1, 40–48, 1996.

Wu, S., Li, J., and Huang, G. H.: Characterization and Evaluation of Elevation Data Uncertainty in Water Resources Modeling with GIS, Water Resour. Manage., 22, 959–972, 2008.

Zhao, R. J. Zhuang, Y. L., Fang, L. R., Liu, X. R., and Zhang, Q. S.: The Xinanjiang model. Hydological Forescasting. IAHS publication 129, IAHS press, Wallingford, 351–356, 1980.

Zhao, R. J.: The Xinanjiang model applied in China, J. Hydrol., 135, 371–381, 1992.