Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-46
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 26 February 2018


Hydrology and Earth System Sciences
Open Access
Discussions

# The effect of input data complexity on the uncertainty in simulated streamflow in a humid, mountainous watershed

Linh Hoang[1,2,*], Rajith Mukundan[2] , Karen E. B. Moore[2], Emmet M. Owens[2] and Tammo S. Steenhuis[3]

[1] Hunter College, City University of New York, 695 Park Avenue, New York, NY 10065, USA

[2] New York City Department of Environmental Protection, 71 Smith Avenue, Kingston, NY 12401, USA

[3] Department of Biological and Environmental Engineering, Cornell University, Ithaca, NY 14853, USA

***Correspondence to:*** Linh Hoang (hnklinh@yahoo.com)

**Abstract**

Uncertainty in hydrological and water quality modelling is of significant concern due to its effects on prediction and subsequent application in watershed management. Similar to other distributed hydrological models, model uncertainty is an issue in applying the Soil and Water Assessment Tool (SWAT). Previous research has shown how SWAT predictions are affected by uncertainty in parameter estimation and input data resolution. Nevertheless, little information is available on how parameter uncertainty and output uncertainty are affected by input data of varying complexity. In this study, SWAT-Hillslope (SWAT-HS), a modified version of SWAT capable of predicting saturation excess runoff was applied to assess the effects of input data with varying degrees of complexity on parameter uncertainty and output uncertainty. Four digital elevation model (DEM) resolutions (1, 3, 10 and 30 m) were tested for their ability to predict streamflow and saturated areas. In a second analysis, three soil maps and three land use maps were used to build nine SWAT-HS setups from simple to complex (fewer to more soil types/ land use classes), which were then compared to study the effect of input data complexity on model prediction/output uncertainty. The case study was the Town Brook watershed in the upper reaches of the West Branch Delaware River in the Catskill Region, New York, USA. Results show that DEM resolution did not impact

Hydrology and Earth System Sciences
Discussions

parameter uncertainty or affect the simulation of streamflow at the watershed outlet but significantly affected the spatial pattern of saturated areas, with 10 m being the most appropriate grid size to use for our application. The comparison of nine model setups revealed that input data complexity did not affect parameter uncertainty. Model setups using intermediate soil/land use specifications were slightly better than the ones using simple information, while the most complex setup did not show any improvement from the intermediate ones. We conclude that increasing spatial input details may not necessarily improve model performance or reduce parameter and output uncertainty, but using multiple temporal and spatial observations can aid in finding the appropriate parameter sets and in reducing prediction/output uncertainty.

**Keywords**: Input data complexity, parameter uncertainty, output uncertainty, SWAT-HS, Catskill region

## 1. Introduction

Uncertainty in hydrological and water quality modelling is of significant concern due to its effects on prediction and subsequent decision making (Van Griensven *et al.*, 2008; Sudheer *et al.*, 2011). The uncertainty of a model ~~has components associated with~~: (i) model structure, (ii) input data, and (iii) model parameters (Lindenschmidt *et al.*, 2007). Uncertainty due to model structure ~~uncertainty~~ results from assumptions or simplifications made in the formulation of the model, and in application of the model under conditions that are not consistent with those assumptions or simplifications (Tripp and Niemann, 2008). Input data uncertainty is caused by changes in natural conditions, limitations of measurement, and lack of data (Beck, 1987). Parameter uncertainty results from the non-linear response of predictions to parameter changes and parameter interdependence leading to the possibility that changes in some parameters may be compensated for by changes in others, so that different parameter sets may produce the same simulated results (Bárdossy and Singh, 2008).

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-46
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 26 February 2018




This so called equifinality is very common in hydrological models and is the cause for uncertainties in model predictions (Beven and Freer, 2001).

Our previous study (Hoang *et al.*, 2017) demonstrated how model structure uncertainty can be addressed by applying SWAT-Hillslope (SWAT-HS), a modified version of the widely applied Soil and Water Assessment Tool (SWAT) (Arnold *et al.*, 1998) developed to improve predictions of saturation-excess runoff. Initial testing of SWAT-HS was carried out in the Town Brook watershed, a 37 km² headwater watershed in the upper reaches of the West Branch Delaware River in the Catskill Mountains of New York. The West Branch Delaware River drains into Cannonsville Reservoir, part the New York City (NYC) water supply system which supplies high quality drinking water to over 9 million people in NYC and nearby communities. In this region, rainfall intensities rarely exceeds infiltration rates and saturation-excess runoff is common (Walter *et al.*, 2003). Results showed good agreement between measured and modeled streamflow at both daily and monthly time steps. More importantly, the model predicted correctly the occurrence of saturated areas on specific days for which observations are available, which was not achieved with application of the standard SWAT model. Consequently, SWAT-HS performs well for our region and shows promise as a good model for humid vegetated areas where saturation-excess runoff is dominant. The model modification is relatively new and research into its proper application is ongoing. Here SWAT-HS is applied to evaluate the effect of complexity of input data on parameter uncertainty and model prediction/output uncertainty.

In previous SWAT studies, parameter uncertainty has received the most attention among the three types of model uncertainty (Shen *et al.*, 2008; Cibin *et al.*, 2010; Shen *et al.*, 2010; Sexton *et al.*, 2011). These studies confirmed limited identifiability of SWAT parameters and equifinality in calibrating discharge at the outlet of the watershed. Sexton et al. (2011) found that the model output uncertainty is not only caused by uncertainty of sensitive parameters but also contributed by non-sensitive parameters, and thus, suggested considering non-sensitive parameters in calibration and uncertainty analysis. Parameter uncertainty caused the least uncertainty for runoff (Shen *et al.*, 2008; Shen *et al.*, 2010) and greatest uncertainty for sediment (Sexton *et al.*, 2011) among streamflow, sediment, nitrogen and phosphorus outputs. Moreover, the effect of parameter uncertainty can be temporally and spatially different. Temporally, parameter uncertainty causes higher output uncertainty in high-flow periods (Shen *et al.*, 2008; Sexton *et al.*, 2011; Shen *et al.*, 2012). Spatially, SWAT generally predicted streamflow with less uncertainty in watersheds in humid climates relative to arid or semi-arid climates (Veith *et al.*, 2010). The source of uncertainty is mainly influenced by parameters associated with runoff (Shen *et al.*, 2008). However, soil properties can also contribute to uncertainty (Shen *et al.*, 2010).

Effects of input data uncertainty have been evaluated in several SWAT applications by exploring the sensitivity of required input data for SWAT model set up, including the DEM, soil, and land use, on model outputs. While most studies focused on the sensitivity of predictions to DEM resolution, a few studies focused on the effects of soil and land use with varying spatial scales. Cotter et al. (2003), found that DEM resolution is the most sensitive input variable, while soil and land use resolution have insignificant impacts on the simulation of streamflow, sediment, nitrate, and total phosphorus. They suggested that the minimum DEM resolution should range from 30 to 300 m, and minimum land use and soils data resolution should range from 300 to 500 m. Chaubey et al. (2005) showed the significant impact of DEM resolution not only on watershed delineation, stream network and subbasin classification, but also on streamflow and nitrate load predictions. Based on SWAT application to a 21.8 $km^2$ watershed in Lower Walnut Creek, central Iowa USA, Chaplot (2005) proposed an upper limit of ~~DEM at 50 m~~ for watershed simulation after determining that coarser grid sizes do not substantially affect the runoff but result in significant errors for nitrogen and sediment yields. Geza and McCray (2008) and Mukundan et al. (2010) compared SWAT streamflow simulations using a low resolution State Soil Geographic database (STATSGO) and a high resolution Soil Survey Geographic database (SSURGO) soil database. While Geza and McCray (2008) found that STATSGO performed better than SSURGO before calibration and the opposite was observed after calibration, Mukundan et al. (2010) found insignificant differences between the two datasets in simulating streamflow.

Most previous SWAT studies focused on how SWAT predictions are affected by uncertainty of parameter estimation and different input data. Limited information is available on how the parameter uncertainty and output uncertainty are affected by different input data with the exception of Kumar and Merwade (2009) who tested the impact of watershed subdivision and the use of two soil datasets (STATSGO and SSURGO) on streamflow calibration and parameter uncertainty. Although there have been numerous studies on the effect of DEM resolutions on SWAT predictions, none have discussed its effects on model uncertainty and specifically on parameter uncertainty. Moreover, these studies on model uncertainty used an integrated response of the watershed (i.e., discharge at the outlet) for assessing complex processes inside the watershed and have not used additional spatial datasets that may reduce model uncertainty.

The two main objectives of this paper are to evaluate: (i) the effect of DEMs of various spatial resolution (1, 3, 10, and 30 m) on the uncertainty of streamflow and saturated area predictions, and (ii) the impact of combinations of soil and land use data with various degrees of complexity on the uncertainty in model output. In both analyses, we not only investigate the effect on model prediction/output uncertainty but also discuss their effect on the uncertainty in parameter estimation. This study seeks to answer three research questions: (i) what is the suitable DEM resolution in order to get good model performance, (ii) what is the appropriate complexity of the distributed input data, and (iii) how does input complexity affect parameter uncertainty and model output uncertainty. The answers to these research questions will be the basis for reducing decision uncertainty on model input selection in our future applications of SWAT-HS in the NYC water supply system.

## 2. Material and methods

### 2.1. Study area: Town Brook watershed, New York

The 37 km$^2$ Town Brook watershed is located in the Catskill Mountains, Delaware County, New York State (Fig. 1). Elevation ranges from 493 to 989 m. The area is humid with an average temperature of eight (8) °C and average annual precipitation of 1123 mm/year. Most soils are either silt loam or silty clay loam. The upper terrain of the watershed has shallow soils (average thickness 80 cm) overlaying fractured bedrock and steep slopes (average slope 29%), while deeper soils (average thickness 180 cm) underlain by a dense fragipan restricting layer and more gentle slope (average slope 14%) are common in the lower terrain. Deciduous and mixed forests predominate in the upper terrain, covering more than half of the land area. In the lowland area, the principle land uses are agriculture (32%) that includes dairy and beef farms with cropland and pastures; brushland (9%); and residential areas (4%).

### 2.2. Brief description of SWAT-HS

SWAT-HS is a modified version of the SWAT model version 2012 (SWAT2012) that is capable of predicting saturation-excess runoff. Two main modifications made in SWAT-HS include: (i) adding information of topography and soil water storage capacity to the modelling unit of SWAT, i.e. Hydrological Response Unit (HRU); and (ii) introducing a surface aquifer that allows lateral exchange of subsurface water from upslope to downslope areas.

Similar to SWAT, SWAT-HS divides the watershed into subbasins. Additionally, the watershed is divided into maximum 10 wetness classes, each of which consists of areas in the subbasin with similar topographic indices (TI). Subsequently, subbasin is further divided into HRUs which are a unique combination of soil, land use, and slope as in SWAT plus an additional component: wetness class. The topographic index (TI) is defined as:

$$TI = ln\left(\frac{\alpha}{tan(\beta)K_s D}\right) \tag{1}$$

where *TI* is the soil topographic index [with units of ln(d m $^{-1}$)], $\alpha$ is the upslope contributing area per unit contour length (m), $\tan(\beta)$ is the local surface topographic slope, $K_s$ is the mean saturated hydraulic conductivity of the soil (m d$^{-1}$), and D is the soil depth (m).

The soil water storage capacity of the wetness classes is defined as the amount of water in the rootzone between field capacity and saturation. This was assumed to vary across the soil wetness classes following a Pareto distribution (Hoang *et al.*, 2017). Wetness classes that are

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-46
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 26 February 2018




located in the downslope areas have lower storage capacities, which means they are 'wetter' than wetness classes in the upslope areas with smaller TI values and higher storage capacities. The wetter the wetness class, the faster the runoff response is during a rainstorm. A surface aquifer is introduced to connect all wetness classes across the hillslope, and transmits subsurface flow that is generated from this aquifer (known as lateral flow in SWAT) laterally through the hillslope from "drier" (upslope) to "wetter" (downslope) wetness classes.

SWAT-HS removes the original curve number method of SWAT in predicting total surface runoff. Instead, it simulates ~~separately~~ infiltration-excess runoff and saturation-excess runoff with different methods. Infiltration-excess runoff is predicted using the Green-Ampt method built in SWAT. Saturation-excess runoff in SWAT-HS is generated in the "wetter" (downslope) wetness classes by two processes: (i) rain falls in wet areas with limited storage capacities where the excess water becomes runoff, and (ii) water from the upland areas is transported laterally to the lowland areas and the water exceeding soil storage capacity becomes runoff (see Hoang et al., 2017 for details).

## 2.3. Methodology

### 2.3.1. Effect of DEM resolution

Four DEMs from fine to coarse resolution were used to set up the SWAT-HS model for the Town Brook watershed. The resolutions employed were 1m, 3m, 10m, and 30m. The 1m DEM (DEM1m) was derived from 2009 aerial LiDAR data acquired by New York City Department of Environmental Protection (NYCDEP). This was resampled to create 3m, 10m and 30m resolution DEMs (DEM3m, DEM10m and DEM30m).

DEMs were used to delineate the watershed, calculate flow paths, slopes, drainage areas, and compute gridded values of topographic index (TI). Based on TI values, the watershed was divided into 10 wetness classes. Wetness class 1 covering very small fraction of the watershed (0.59%) is actually fit to the perennial stream network and is the "wettest" wetness class. We grouped 50% of the watershed with lowest TI values in the upland as the "driest" wetness

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-46
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 26 February 2018

Hydrology and Earth System Sciences
Discussions

class (wetness 10) because saturated areas never exceeded 50% of the watershed based on observations (Harpold *et al.*, 2010) and predictions by other watershed models (SMR (Agnew *et al.*, 2006), SWAT-VSA (Easton *et al.*, 2008) and SWAT-WB (White *et al.*, 2011). Subsequently, we divided the remaining areas into 8 wetness classes (wetness class 2 – 9) with approximately equal areas (~ 6% each) based on TI values.

HRUs were created based on 10 wetness classes, 17 soil types, and 11 land use types. A single time series of daily precipitation and temperature data were interpolated from a 4km x 4km gridded PRISM climate dataset (Daly *et al.*, 2008) using the inverse distance weighting method. Solar radiation was averaged from Albany and Binghamton airport data. Relative humidity and wind speed were generated by the built-in weather generator in SWAT. The procedure outlined above is similar to the SWAT-HS setup used by Hoang *et al.* (2017).

Four SWAT-HS setups were run ~~for~~ a daily time step from 1998 – 2012. The first 3 years were used as the warming up period and the model was calibrated and validated for the periods 2001-2007 and 2008-2012, respectively. We excluded the year 2011 from the validation period because there were two extreme events (Hurricane Irene and Tropical Storm Lee) in August 2011 that the model could not capture well. We used a Monte Carlo sampling method to calibrate all four models in two stages: snowmelt calibration (with 5 snowmelt-related parameters) and streamflow calibration (with 9 flow parameters). The details of parameters and their value ranges are shown in Table 1. In each stage, 10,000 parameter sets were randomly generated within their value ranges and then run with the SWAT-HS model. The optimal parameter set from the snowmelt calibration was used in the final streamflow calibration.

We evaluated the effect of DEM resolution on representing topographical characteristics of the watershed by comparing the statistical distributions of elevation, slope angle, upslope contributing area, and TI using DEMs with various spatial resolutions (1m, 3m, 10m and 30m). Subsequently, to evaluate the effect of DEM resolution on model uncertainty, we compared the four SWAT-HS setups with different DEM resolutions based on: (i) the uncertainty in streamflow predictions using *"good"* performance parameter sets, (ii) predictions of saturated areas and their uncertainties, and (iii) uncertainty in parameter estimation. We used the Generalized Likelihood Uncertainty Estimation (GLUE) approach (Beven and Binley, 1992) to estimate the uncertainty in streamflow and saturated area predictions caused by parameter uncertainty. For each model setup, "*good*" simulations were identified as those with a Nash-Sutcliffe Efficiency (NSE) greater than 0.65 for use in uncertainty estimation of streamflow. Subsequently, from these "*good*" simulations, we compared predictions of saturated areas with our available field observations of saturated areas to re-select the "*good*" parameter sets for both simulated streamflow and saturated areas to estimate the uncertainty in predicted saturated areas. Six observations of saturated areas (28, 29, 30 April 2006, 12 April 2007, 7 June 2007, and 2 August 2007) are available in a small area in the headwater of the Town Brook watershed.

### 2.3.2. Effect of soil and land use complexity

We built nine SWAT-HS setups ranging from simple (fewer soil types/land use classes/fewer HRUs) to complex (more soil types/ land use classes, more HRUs) based on three soil maps and three land use maps. In all nine setups, the 10m DEM was used based on its performance as the best predictor of saturated areas (see discussion).

Three soil maps were created with increasing levels of complexity (Fig. 2). The simplest map (*TBsoil_1*) had a homogenous soil type, which was created using area-weighted average soil data from the 4 dominant soil types (*Hcc, LhB, OeB, WmB*) in Town Brook. The second soil map *TBsoil_2* has a unique soil type for each wetness class and was created by area-weighted averaging of dominant soil properties in the corresponding wetness class. The most complex soil map *TBsoil_3* consisted of all 17 soil types.

Three land use maps with increasing levels of complexity were created (Fig. 2). The simplest land use map (*TBlanduse_1*), had agriculture as the representative land use for the watershed because it is one of the dominant land uses and potentially has a more significant impact on water quality than other land use types. The more complex land use map (*TBlanduse_2*) classifies Town Brook into 3 diverse land use types: agriculture, forest and urban areas. The most complex one (*TBlanduse_3*) contains all 11 land use types.

HRUs were generated based on a wetness map (10 classes), soil map, land use and slope maps. In order to simplify the setup, we assumed that slope does not have an impact on HRU discretization. We also set the threshold of 1% for soil and 1% for land use to eliminate minor soil types/ land uses that cover only less than 1% of the subbasin.

The nine setups are categorized in 3 groups: (i) *simple*: the setups that use either the simplest soil or land use (TB1-TB5), (ii) *intermediate*: the setups that use the average complexity for maps of either soil or land use (TB6 – TB8); and (iii) *complex*: the setup that uses the most complex maps (TB9) (Table 2).

To evaluate the effect of soil and land use complexity on model uncertainty, we compared the nine SWAT-HS setups using the same methodology used to evaluate the effect of DEM resolution on model uncertainty that is described above.

## 3. Results

### 3.1. The effect of DEM resolution on model uncertainty

#### 3.1.1. Effect on topographic characteristics

DEM resolution has varying effects on the distribution of elevation, slope angle, upslope contributing area, and topographic index values. ~~While~~ the distributions of elevation are similar using different DEMs indicating no effect from DEM resolution (Fig. 3a). The finer resolution DEMs (DEM1m and DEM3m) are able to give more precise slope values. Therefore, coarser DEM resolutions produce slightly narrower slope distributions, lower mean slope angles, lower probability for steep slopes and higher probability for gentle slopes than the finer DEM resolutions because of the smoothing of topography and loss of topographic details (Figure 3b). DEM resolution has a significant effect on the calculated values of upslope contributing areas (Figure 3c). With the finer spatial resolutions, grids in DEM1m and DEM3m have smaller contributing areas than the ones in coarser resolution

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-46
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 26 February 2018


Hydrology and Earth System Sciences
Open Access
Discussions

DEM10m and DEM30m. This results in significant differences in the distribution of TI in that the finer resolution DEMs provide lower values of TI (Fig. 3d). The impact of DEM grid size on TI distribution is mainly due to its impact on upslope contributing area rather than slope. Our results are consistent with previous studies on the effect of DEM resolution on topographic attributes and topographic wetness index (Zhang and Montgomery, 1994; Thompson *et al.*, 2001; Sørensen and Seibert, 2007; Gillin *et al.*, 2015).

Depending on the DEM used, the four wetness maps formed by grouping areas of similar TI into 10 wetness classes showed remarkable differences (Fig. 4). It should be noted here that the differences are on spatial distribution of wetness classes while the areal percentage of each wetness class is approximately similar irrespective of the DEM used. In Figure 4, we show the wetness maps for the headwater area where observations of saturated areas are available. It can be clearly seen that the spatial patterns of wetness classes in coarser resolution DEMs (10m and 30m) are quite similar but they are very different from the finer resolution DEMs (1m and 3m). At a very fine resolution, DEM1m can capture the detailed drainage network, and the wetness classes mirror the drainage network, making it difficult to see the boundary of each wetness class. These detailed drainage features are obscured in coarser grid DEMs, therefore, the boundaries of the wetness classes are easier to distinguish. Our results are consistent with previous studies on the effect of grid size on spatial patterns of topographic wetness index that have been reported by Thomas *et al.* (2017), Erskine *et al.* (2006), Zhang and Montgomery (1994).

### 3.1.2. Effect on the prediction of streamflow

To evaluate the effect of DEM on the uncertainty of streamflow predictions, we compared streamflow outputs from 10,000 Monte Carlo simulations of four model setups with DEMs of different resolution (Fig. 5a). Subsequently, we evaluated and compared streamflow estimates in the validation period based on only "*good*" parameter sets (Fig. 5b). Statistical criteria for uncertainty evaluation are shown in Table 3. In all setups, more than 50% of the parameter sets gave "*satisfactory*" performances (NSE ≥ 0.5) (Fig. 5). Of the total randomly

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-46
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 26 February 2018




generated parameter sets, 9-23% gave "*good*" streamflow performance in the four setups, with higher percentages in coarser resolution setups (DEM10m and DEM30m) (Table 3). For the calibration period, the maximum NSE values were equivalent (around 0.69) in the four setups. However, the median NSE values were slightly higher and the mean NSE were significantly higher in coarser resolution setups (DEM10m and DEM30m) than the higher resolution ones (DEM1m and DEM3m). In the finer resolution setups, there are higher percentages of parameter setups that gave poor fit to observed streamflow (NSE is negative) which causes lower mean values of NSE. For the validation period, the "*good*" parameter sets in four setups all give above satisfactory to good fit to observations and four setups gave relatively similar performance to each other. Generally, there are only slight differences in SWAT-HS performance on streamflow using different DEMs implying the insignificant effect of DEM resolution on streamflow simulation and the uncertainty of streamflow outputs.

Although the effect of DEMs is minor, the setups using coarser resolution DEM10m and DEM30m are slightly better. These two setups gave higher NSE value ranges and significantly higher mean NSE values resulted from all random combinations of parameters. These two setups also have more "*good*" parameter sets indicating higher probability to get "*good*" representation of the modeled watershed. This implies better streamflow prediction by these two setups even without calibration.

### 3.1.3. Effect on the prediction of saturated areas

The probabilities of saturation in 10 wetness classes were compared among four DEM resolution setups using only "*good*" parameter sets for both streamflow and saturated area predictions (Fig. 6). The probability of saturation, which indicates the number of days in the calibration period when the wetness class is saturated, showed no significant difference among the four setups indicating that DEM resolution does not have an impact on the probability of saturation. In all four setups, wetness class 7 to 10 are predicted to be mostly dry, implying that almost 70% of the watershed is rarely saturated. Wetness class 1 has a high probability of saturation (80-100%) because its soil water storage capacity is very low, i.e., the wetness class is prone to saturation whenever there is precipitation. The probability of saturation decreases in the more upslope wetness classes: 60-80% in wetness class 2, 30 – 50% in class 3, 5 – 22% in class 4, 1 – 9% in class 5, 0-3% in class 6, 0-1% in class 7, 0-0.3% in class 8, 0-0.08% in class 9, and 0% in class 10. We also observed that the uncertainty of saturation probability of the more upslope wetness classes is lower because they only respond to high rainfall events.

The results of probability of saturation correspond well with the uncertainty of percentage of saturated areas shown in Figure 7. The four model setups did not have significant differences in the percentage of saturated areas in the watershed. The maximum, minimum and the interquartile range indicated by the top and bottom values of the four box plots are slightly different because of minor differences in division of wetness classes in the watershed. For the majority of the time, no more than approximately 25% of the total watershed area is saturated. The watershed can be saturated up to more than 50% in extreme events that are shown as outliers in the boxplots. The median percentage of saturated areas in the watershed is only around 7-8%.

Although the statistical distributions of saturated areas in four DEM setups are relatively similar, the spatial distributions of saturated areas simulated in a small headwater area (Fig. 1) on specific days (28-30 April 2006) when observations are available appeared to be different as shown in Figure 8. In Figure 8, the saturated areas simulated in four DEM setups correspond to the saturation of wetness classes 1, 2 and 3. Saturated areas cover approximately equal areas of the watershed for the different DEM resolutions, but differ significantly in spatial distribution. With the smallest cell size, DEM1m captures the most detailed drainage pattern, which results in saturated areas following detailed flow paths. The saturated areas resulting from DEM3m are scattered, not well connected, and broadly distributed. For coarser resolution DEM10m and DEM30m, saturated areas connect well with each other and with the areas concentrated near the streams. The percentages of simulated saturated areas that intersect with observations increase with coarser resolution DEMs: 34%

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-46
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 26 February 2018


(DEM1m), 53% (DEM3m), 85% (DEM10m) and 90% (DEM30m). Therefore, based on visual comparison with observations and our calculation, the coarser resolution DEMs gave better fits to observed saturated areas than the higher resolution DEMs. Among the four DEMs, DEM10m provided the most realistic representation of saturated areas and reasonable fit to observations.

#### 3.1.4. Effect on parameter uncertainty

Figure 9 shows the comparison between the distribution of *"good"* for streamflow (in light green) and the distribution of *"good"* parameters for both streamflow and saturated areas (in green) in four SWAT-HS model setups with different resolution DEMs. Only two parameters distributions (*latb* and *Smax*) are plotted in Figure 9 because they are the most sensitive parameters (Hoang et al., 2017). Although the number of good parameters for streamflow varied in four setups, the ranges of good parameter values and the shape of their distributions are alike for all calibrated parameters. Using multiple observations (both streamflow and saturated areas) helped to reduce a great number of *"good"* parameters in all 4 setups, but did not significantly narrow down the value ranges of good parameters. The similarity in distribution of good parameters in four setups with different DEM resolution implies that DEM resolution has a negligible impact on parameter uncertainty for this watershed.

### 3.2. Effect of soil and land use input complexity on model uncertainty

#### 3.2.1. Effect on uncertainty in streamflow predictions

All nine SWAT-HS setups with different degrees of complexity were able to obtain good model performance and were comparable to one another (Fig. 10 and Table 4). More than 50% of the total simulations in each setup produced NSE greater than 0.5, which corresponds to "satisfactory" performance. All setups also had high percentages of "*good*" performance (12.5 – 22.6%), with TB1 and TB8 having the lowest and highest percentages, respectively. The maximum NSE obtained from nine setups are relatively equivalent. The mean and median NSE values are slightly different, except for the TB3 setup with the lowest mean NSE

(-0.08). Applying only the "*good*" parameter sets in the validation period, we still observe insignificant differences among the nine setups. Additionally, all these "*good*" parameter sets gave above '*satisfactory*' to "*good*" fit to observations in the validation periods implying that all nine setups are reasonable setups to use for flow predictions. In spite of minor differences, from all the evaluation criteria, TB3 gives the poorest performance among nine setups followed by the simplest setup TB1. Setups TB6 to TB9 give equally good performance and are better than the remaining ones.

Grouping the nine setups into three groups: (i) *simple* (TB1 – TB5), (2) *intermediate* (TB6 – TB8); and (iii) *complex* (TB9), we observed that the model performance of setups in *intermediate* groups are slightly better than the *simple* one although the differences were small. The *intermediate* group has a higher number of "*good*" parameter sets, a higher mean NSE in the calibration period, as well as consistently better performance in the validation period. The most complex setup (TB9) gives equally good performance as setups in the *intermediate* group with no improvement in any statistical metric.

#### 3.2.2. Effect on the prediction of saturated areas

We did not observe any significant difference in the prediction of saturated areas in nine setups. All setups used the same DEM with 10m resolution and have the same distribution of wetness classes; therefore, the distributions of saturated areas, which are based on wetness maps in SWAT-HS, are similar. Consequently, with SWAT-HS, the prediction of saturated areas is highly dependent on the division of wetness classes and is not affected by soil and land use information as long as the model is well calibrated for both streamflow and saturated areas.

#### 3.2.3. Effect on parameter uncertainty

We tested the effect of soil and land use complexity on parameter uncertainty by comparing the distribution of good parameters among nine setups with different degrees of complexity as in Figure 11. We only showed the distribution of one calibrated parameter *latb* as an

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-46
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 26 February 2018

Hydrology and Earth System Sciences
Open Access
Discussions

example because we observed the same behavior in the remaining calibrated parameters. Similar to the comparison of four setups using different DEMs, the nine setups with different degrees of complexity produced different numbers of good parameters for streamflow and saturated areas, but were similar in the shape of their distributions and value ranges. Accordingly, soil and land use complexity have negligible effect on parameter uncertainty.

## 4. Discussion

The objective of this study was to estimate uncertainty in ~~parameter estimation~~, and predictions of streamflow and saturated areas due to the effects of DEM resolution and complexity in model setup, specifically combinations of land use and soils. The following sections discuss the ~~proposed~~ research questions based on the obtained results.

### 4.1. What is the most suitable DEM resolution to use in SWAT-HS?

Our results show that randomly generated parameter values from coarser resolution DEMs (DEM10m and DEM30m) perform better for streamflow prediction. However, after calibration, the effect of DEM resolution on the uncertainty of streamflow prediction is very minor. This result is in agreement with Liu et al. (2005) using the Wetspa model with 50 – 800 m cell sizes, Molnar and Julien (2000) using the CASC2D model with 127 – 914 m cell sizes, and Chaplot (2005) using SWAT with 20-500m DEMs. These studies found that discharge was simulated equally well irrespective of DEM resolution as long as parameters are calibrated properly. Kuo et al. (1999) found only minor effects of grid size ranging from 10-400m on discharge during dry years.

For models calibrated for both streamflow and saturated areas, DEM resolution has very limited impact on probability of saturation in wetness classes and percentage of saturated areas in the watershed. However, it greatly influences the spatial distribution of saturated areas. SWAT-HS simulates the saturation-excess runoff coming from saturated areas based on a statistical soil water distribution assigned to wetness classes. The *"wettest"* wetness classes downslope with lowest soil water storage capacity is saturated first followed by *"drier"* adjacent wetness classes located more upslope. Therefore, the distribution of saturated areas follows the distribution of wetness classes categorized by the values of TI. In our analysis of effect of DEM resolution on topographic characteristics, we observed that the statistical distribution of TI is very sensitive to DEM resolution (Fig. 3d), which results in considerable differences in spatial distribution of wetness classes (Fig. 4). This explains why the distribution of simulated saturated areas by SWAT-HS is also very sensitive to DEM resolution.

Our results show that DEM10m is the best choice among four DEMs tested because of its slightly better performance on streamflow and more importantly, its good fit to observations of saturated areas. Although DEM30m also gives very good results for streamflow and distribution of saturated areas, we did not choose DEM30m because its coarse cell size may overestimate the extent of actual saturated areas. Therefore, DEM10m is the preferred choice to scale-up the application of SWAT-HS to larger watersheds in the New York City water supply system for future applications. Our choice of DEM10m is in agreement with Kuo et al. (1999) who evaluated DEM grid size ranging from 10-400m on runoff and soil moisture for a variable-source area hydrology model and observed that by using the 10x10m grid cells, the overall pattern of simulated wet areas showed a close correspondence with the poorly drained areas defined in the soil survey. Zhang and Montgomery (1994), in a study that evaluated grid size effect using TOPMODEL, also suggested that a 10m grid size presents a rational compromise between increasing resolution and data volume for simulating geomorphic and hydrological processes. In contrast, Thomas *et al.* (2017) indicated that LiDAR DEM 1-2 m is optimal for modelling hydrologically sensitive areas (runoff generating areas) and is far better than the radar based DEM 5m. However, their case study is a complex agricultural catchment dominated by micro-topographic features, which can only be captured using high resolution DEMs. In the Town Brook watershed, saturation occurring in moist wetness classes is not only caused by the local saturation-excess runoff, but mainly by the lateral transport of subsurface flow from the upslope to the downslope areas that keeps the downslope areas saturated (Hoang *et al.*, 2017). With the dominance of lateral flow that typically follows the general topography of the landscape, fine resolution DEMs that capture small-scale surface variations may not be necessary and appropriate. Our choice of DEM10m is in contrast to Buchanan *et al.* (2014) who preferred DEM3m rather than DEM10m because of the better fit with the observed patterns of soil moisture collected in five different agricultural field sites. The difference in scale of case studies (field scale vs. watershed scale) and characteristics of case studies (agricultural fields vs. a mixture of forest and agriculture) between Buchanan *et al.* (2014) and our study may have resulted in different conclusions on choice of the appropriate DEM resolution.

It should be noted here that all four DEMs in this study are derived from the same source of 2009 aerial LiDAR data with 1 meter resolution. The coarser DEMs (DEM3m, DEM10m and DEM30m) are resampled products from DEM1m. Therefore, the four different DEM resolutions carry similar information, but differ in topographic smoothing. A comparison of various resolution DEMs from different sources may not yield the same results.

### 4.2. What is the appropriate complexity of the distributed soil and land use inputs?

From our comparison of nine SWAT-HS setups in three groups of complexity (*simple*, *intermediate* and *complex*), we found that with all randomly generated parameter values, the *intermediate* and *complex* groups are better than the *simple* group based on slightly higher mean NSE values and higher probability of good performance based on randomly generated parameter values. The TB3 setup, which was built from the most complex soil maps (17 soil types) and the simplest land use maps (1 land use) and the simplest setup TB1 are the two poorest setups in the *simple* group. Additionally, compared to the *intermediate* group, the *complex* group does not gain any improvement from using inputs that are more detailed. However, with proper calibration, all nine models are able to provide good performances and their "*good*" parameter sets continue to perform equally well in the validation period. In addition to streamflow, all nine setups are able to capture saturated areas correctly on specific days where observations are available. We conclude that increasing spatial input details does not necessarily give better results for streamflow simulation as long as the model is properly calibrated. However, over-simplification like the simple setups TB1, TB3 with only one land use type may have greater impacts on water quality modelling. We recommend using *intermediate* inputs for the SWAT-HS setup that adequately represent the spatial distribution of dominant soils and land use types. It should be noted here that in this paper, hydrological response is the main focus of this study, and streamflow may not be very sensitive to the details of land use. However, water quality modelling may need more detailed classification of land uses. For example, agriculture land use may have to be divided into croplands and pasture because nutrient inputs and management practices are different in these two sub-classes of land use.

Our results are in agreement with previous studies on the effect of model input complexity on streamflow simulation. Using an urban hydrological distributed model, Petrucci and Bonhomme (2014) show that the inclusion of some basic geographical information, particularly on land use, improves the model performance, but further refinements are less effective. Finger *et al.* (2015) compared different setups with increasing detail in input information using the HBV model and three observational data sets. They found that enhanced model input complexity does not lead to significant increase in overall performance, but suggested that the availability and use of different datasets to calibrate hydrological models might be more important than model input data complexity to achieve realistic estimations of runoff composition. Muleta *et al.* (2007) also showed that streamflow simulated by SWAT is relatively insensitive to spatial scale when comparing multiple watershed delineations from different soil and land use input data details.

### 4.3. How does input complexity affect parameter uncertainty and model output uncertainty?

Our results show that regardless of the level of detail of input data, we obtained numerous sets of parameter values that give equally good performance for streamflow and saturated area predictions. Modifying the level of detail in input data changed the number of "*good*" parameter sets, but the ranges of "*good*" parameter values and the shape of their distributions remained the same. Therefore, we conclude that for this case study and the particular model SWAT-HS, using higher resolution DEM or adding complex information on soil or land use did not reduce parameter uncertainty or solve the equifinality problem. This statement may not be valid for other areas that are characterized by numerous land use and complex variations in topography and soil types. This is also not valid for physically based models which require detailed soil and land use information and a minimum number of parameters for calibration.

Combining different observations sets (temporal observations of streamflow and spatial observations of saturated areas in multiple days) in calibration will help to reduce the number of "*good*" parameter sets and choose the appropriate parameter sets that give good representation of hydrological processes in the watershed. The importance of using multiple data sets have been addressed in Finger *et al.* (2015), McMillan *et al.* (2011) and Kirchner (2006). Our study is not aimed at solving the equifinality problem, but rather reduces the number of solutions considered when using SWAT-HS to predict streamflow and water quality for decision-making. The outcome of this study directly reduces the decision uncertainty with regard to selecting the optimum combination of input datasets for model setup that gives the best model results both spatially and temporally. This has implications on watershed modelling by reducing model run time as we scale-up the application of SWAT-HS and to other larger watersheds within the NYC water supply system.

## 5. Summary and conclusions

This paper is a follow-up to our previous study using the SWAT-HS model, investigating the effect of input data complexity on the uncertainty in predictions of streamflow and saturated areas. The input data include DEMs with different resolutions and the combinations of simple to complex soil and land use maps. Major objectives were to explore whether using more complex spatial data yield better and more robust results, and choose the most appropriate input data for future applications of SWAT-HS in other watersheds or larger watersheds within the New York City water supply system.

We chose DEM10m that is resampled from LiDAR DEM1m as the most appropriate resolution to use, and recommended the use of an intermediate soil and land use map for our future applications of SWAT-HS. Our results showed that streamflow is not sensitive to both DEM resolution and soil and land use complexity as long as proper calibration is carried out. However, DEM resolution has a significant impact on the spatial distribution of predicted saturated areas because of its major effect on the division of wetness classes. The prediction of saturated areas is not sensitive to soil and land use inputs when the same DEM resolution is used.

Regardless of the level of detail for input data, the equifinality problem can cause uncertainty in modeled results when using different SWAT-HS setups. Increasing input data complexity does not help to reduce parameter uncertainty and the uncertainty of model predictions. However, using multiple types of observed datasets such as spatial observations in addition to the conventional temporal observations can eliminate a high number of unsuitable parameter sets and guide selection of the appropriate parameter sets that give good temporal and spatial predictions for streamflow and saturated areas.

## References

Agnew LJ, Lyon S, Gérard-Marchant P, Collins VB, Lembo AJ, Steenhuis TS, Walter MT. 2006. Identifying hydrologically sensitive areas: Bridging the gap between science and application. Journal of Environmental Management, **78**: 63-76. DOI: http://dx.doi.org/10.1016/j.jenvman.2005.04.021.

Arnold JG, Srinivasan R, Muttiah RS, Williams JR. 1998. Large area hydrologic modeling and assessment part 1: Model development. JAWRA Journal of the American Water Resources Association, **34**: 73-89. DOI: 10.1111/j.1752-1688.1998.tb05961.x.

Bárdossy A, Singh SK. 2008. Robust estimation of hydrological model parameters. Hydrology and Earth System Science, **12**: 1273-1283. DOI: 10.5194/hess-12-1273-2008.

Beck MB. 1987. Water quality modeling: a review of the analysis of uncertainty. Water Resources Research, **23**: 1393-1442.

Beven K, Binley A. 1992. The future of distributed models: Model calibration and uncertainty prediction. Hydrological Processes, **6**: 279-298. DOI: 10.1002/hyp.3360060305.

Beven K, Freer J. 2001. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. Journal of hydrology, **249**: 11-29. DOI: http://dx.doi.org/10.1016/S0022-1694(01)00421-8.

Buchanan BP, Fleming M, Schneider RL, Richards BK, Archibald J, Qiu Z, Walter MT. 2014. Evaluation topograpic wetness indices across central New York agricultural landscapes. Hydrology and Earth System Science, **18**: 3279-3299.

Chaplot V. 2005. Impact of DEM mesh size and soil map scale on SWAT runoff, sediment, and NO 3–N loads predictions. Journal of hydrology, **312**: 207-222.

Chaubey I, Cotter A, Costello T, Soerens T. 2005. Effect of DEM data resolution on SWAT output uncertainty. Hydrological Processes, **19**: 621-628.

Cibin R, Sudheer K, Chaubey I. 2010. Sensitivity and identifiability of stream flow generation parameters of the SWAT model. Hydrological processes, **24**: 1133-1148.

Cotter AS, Chaubey I, Costello TA, Soerens TS, Nelson MA. 2003. Water quality model output uncertainty as affected by spatial resolution of input data. JAWRA Journal of the American Water Resources Association, **39**: 977-986. DOI: 10.1111/j.1752-1688.2003.tb04420.x.

Daly C, Halbleib M, Smith JI, Gibson WP, Doggett MK, Taylor GH, Curtis J, Pasteris PP. 2008. Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. International Journal of Climatology, **28**: 2031-2064. DOI: 10.1002/joc.1688.

Easton ZM, Fuka DR, Walter MT, Cowan DM, Schneiderman EM, Steenhuis TS. 2008. Re-conceptualizing the soil and water assessment tool (SWAT) model to predict runoff from variable source areas. Journal of Hydrology, **348**: 279-291. DOI: http://dx.doi.org/10.1016/j.jhydrol.2007.10.008.

Erskine RH, Green TR, Ramirez JA, MacDonald LH. 2006. Comparison of grid-based algorithms for computing upslope contributing area. Water Resources Research, **42**: n/a-n/a. DOI: 10.1029/2005WR004648.

Finger D, Vis M, Huss M, Seibert J. 2015. The value of multiple data set calibration versus model complexity for improving the performance of hydrological models in mountain catchments. Water Resources Research, **51**: 1939-1958.

Geza M, McCray JE. 2008. Effects of soil data resolution on SWAT model stream flow and water quality predictions. Journal of environmental management, **88**: 393-406.

Gillin CP, Bailey SW, McGuire KJ, Prisleyt SP. 2015. Evaluation of LiDAR-derived DEMs through terrain analysis and field comparison. Photogrammetric Engineering & Remote Sensing, **81**: 387-396.

Harpold AA, Lyon SW, Troch PA, Steenhuis TS. 2010. The hydrological effects of lateral preferential flow paths in a glaciated watershed in the Northeastern USA. Vadose Zone Journal, **9**: 397-414.

Hoang L, Schneiderman EM, Moore KE, Mukundan R, Owens EM, Steenhuis TS. 2017. Predicting saturation-excess runoff distribution with a lumped hillslope model: SWAT-HS. Hydrological Processes.

Hoang L, Schneiderman EM, Moore KEB, Mukundan R, Owens EM, Steenhuis TS. 2017. Predicting saturation-excess runoff distribution with a lumped hillslope model: SWAT-HS. Hydrological Processes, **31**: 2226-2243. DOI: 10.1002/hyp.11179.

Kirchner JW. 2006. Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology. Water Resource Research, **42**: W03S04. DOI: DOI: 10.1029/2005wr004362.

Kumar S, Merwade V. 2009. Impact of watershed subdivision and soil data resolution on SWAT model calibration and parameter uncertainty1. Wiley Online Library.

Kuo WL, Steenhuis TS, McCulloch CE, Mohler CL, Weinstein DA, DeGloria SD. 1999. Effect of grid size on runoff and soil moisture for a variable-source-area hydrology model. Water Resources Research, **35**: 3419-3428.

Lindenschmidt K-E, Fleischbein K, Baborowski M. 2007. Structural uncertainty in a river water quality modelling system. Ecological Modelling, **204**: 289-300.

Liu YB, Li Y, Batelaan O, De Smedt F. 2005. Assessing grid size effects on runoff and flow response using a GIS-Based hydrologic model. In: Proceeding of the 13th International Conference on Geoinformatics, Li SN, Tao V (eds.).

McMillan HK, Clark MP, Bowden WB, Duncan M, Woods RA. 2011. Hydrological field data from a modeller's perspective: Part 1. Diagnostic tests for model structure. Hydrological Processes, **25**: 511-522.

Molnar DK, Julien PY. 2000. Grid size effects on surface runoff modeling. Journal of hydrologic engineering, **5**: 8 - 16.

Mukundan R, Radcliffe D, Risse L. 2010. Spatial resolution of soil data and channel erosion effects on SWAT model predictions of flow and sediment. Journal of Soil and Water Conservation, **65**: 92-104.

Muleta MK, Nicklow JW, Bekele EG. 2007. Sensitivity of a distributed watershed simulation model to spatial scale. Journal of Hydrologic Engineering, **12**: 163-172.

Petrucci G, Bonhomme C. 2014. The dilemma of spatial representation for urban hydrology semi-distributed modelling: Trade-offs among complexity, calibration and geographical data. Journal of hydrology, **517**: 997-1007.

Sexton A, Shirmohammadi A, Sadeghi A, Montas H. 2011. Impact of parameter uncertainty on critical SWAT output simulations. Transactions of the ASABE, **54**: 461-471.

Shen Z-y, Hong Q, Yu H, Niu J-f. 2010. Parameter uncertainty analysis of non-point source pollution from different land use types. Science of the total environment, **408**: 1971-1978.

Shen Z, Chen L, Chen T. 2012. Analysis of parameter uncertainty in hydrological and sediment modeling using GLUE method: a case study of SWAT model applied to Three Gorges Reservoir Region, China. Hydrology and Earth System Sciences, **16**: 121-132.

Shen Z, Hong Q, Yu H, Liu R. 2008. Parameter uncertainty analysis of the non-point source pollution in the Daning River watershed of the Three Gorges Reservoir Region, China. Science of the total environment, **405**: 195-205.

Sørensen R, Seibert J. 2007. Effects of DEM resolution on the calculation of topographical indices: TWI and its components. Journal of Hydrology, **347**: 79-89.

Sudheer K, Lakshmi G, Chaubey I. 2011. Application of a pseudo simulator to evaluate the sensitivity of parameters in complex watershed models. Environmental Modelling & Software, **26**: 135-143.

Thomas I, Jordan P, Shine O, Fenton O, Mellander P-E, Dunlop P, Murphy P. 2017. Defining optimal DEM resolutions and point densities for modelling hydrologically sensitive areas in agricultural catchments dominated by microtopography. International Journal of Applied Earth Observation and Geoinformation, **54**: 38-52.

Thompson JA, Bell JC, Butler CA. 2001. Digital elevation model resolution: effects on terrain attribute calculation and quantitative soil-landscape modeling. Geoderma, **100**: 67-89. DOI: http://dx.doi.org/10.1016/S0016-7061(00)00081-1.

Tripp DR, Niemann JD. 2008. Evaluating the parameter identifiability and structural validity of a probability-distributed model for soil moisture. Journal of hydrology, **353**: 93-108.

Van Griensven A, Meixner T, Srinivasan R, Grunwald S. 2008. Fit-for-purpose analysis of uncertainty using split-sampling evaluations. Hydrological sciences journal, **53**: 1090-1103.

Veith T, Van Liew M, Bosch D, Arnold J. 2010. Parameter sensitivity and uncertainty in SWAT: A comparison across five USDA-ARS watersheds. Transactions of the ASABE, **53**: 1477-1486.

Walter M, Mehta V, Marrone A, Boll J, Gérard-Marchant P, Steenhuis T, Walter M. 2003. Simple Estimation of Prevalence of Hortonian Flow in New York City Watersheds. Journal of Hydrologic Engineering, **8**: 214-218. DOI: 10.1061/(ASCE)1084-0699(2003)8:4(214).

White ED, Easton ZM, Fuka DR, Collick AS, Adgo E, McCartney M, Awulachew SB, Selassie YG, Steenhuis TS. 2011. Development and application of a physically based landscape water balance in the SWAT model. Hydrological Processes, **25**: 915-925. DOI: 10.1002/hyp.7876.

Zhang W, Montgomery DR. 1994. Digital elevation model grid size, landscape representation, and hydrologic simulations. Water resources research, **30**: 1019-1028.

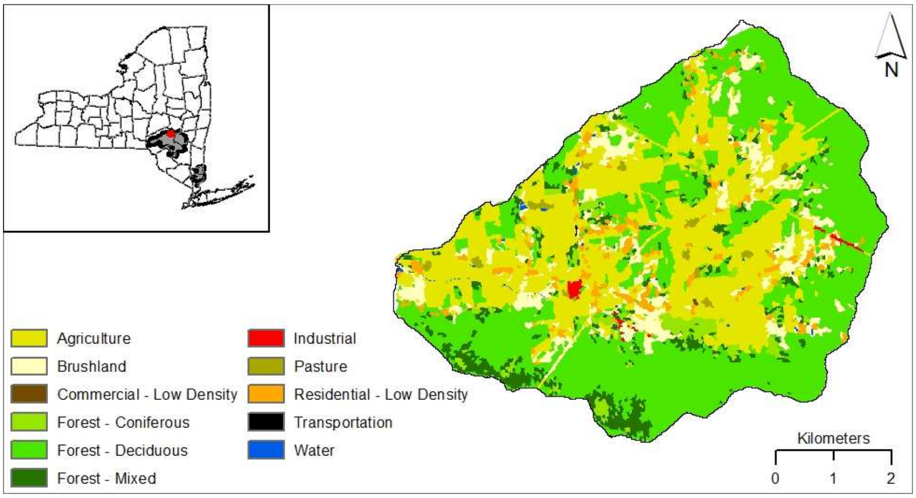
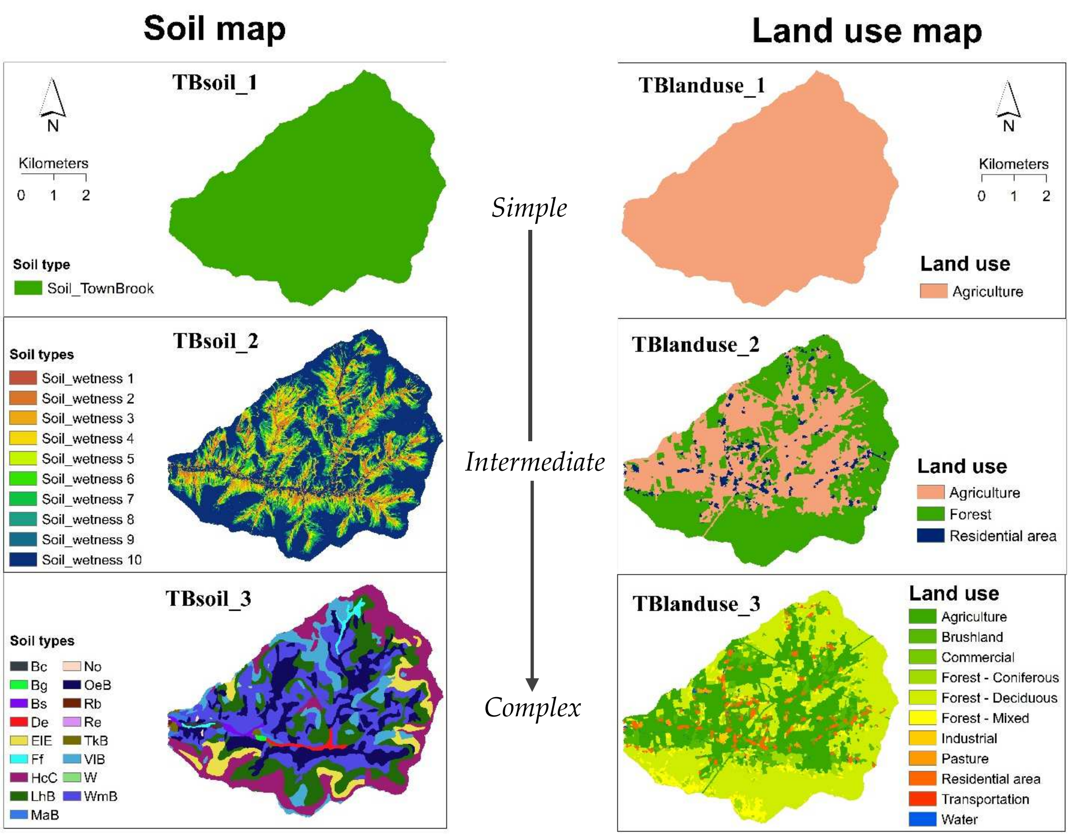



**Figure 1: Town Brook watershed, Delaware County, New York**

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-46
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 26 February 2018




1 


2 **Figure 2: Soil and land use maps with increasing levels of complexity to build SWAT-HS**
3 **model setups**

4

5

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-46
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 26 February 2018


Hydrology and Earth System Sciences
Open Access
Discussions

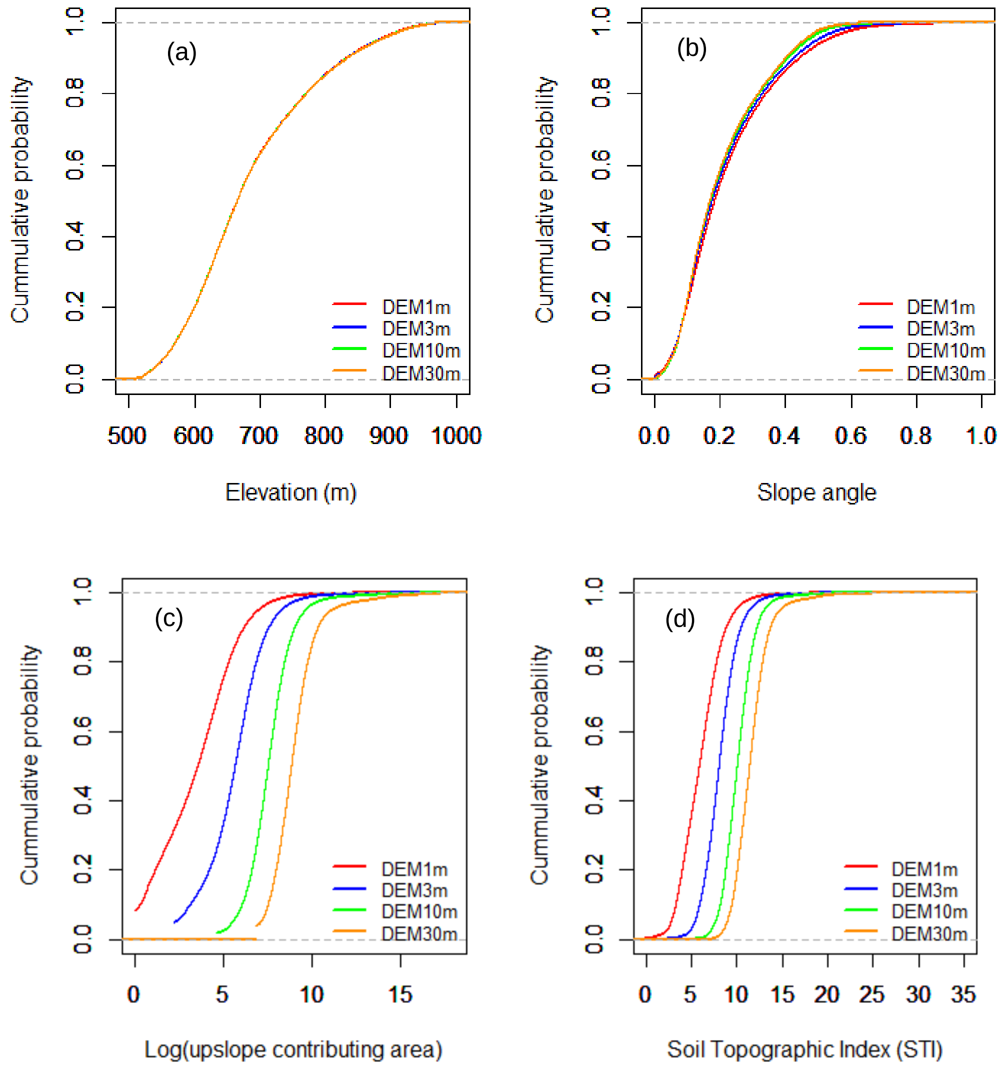



**Figure 3: Difference in cumulative probability distribution of elevation, slope, upslope contributing area and topographic index between different DEM resolutions**

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-46
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 26 February 2018



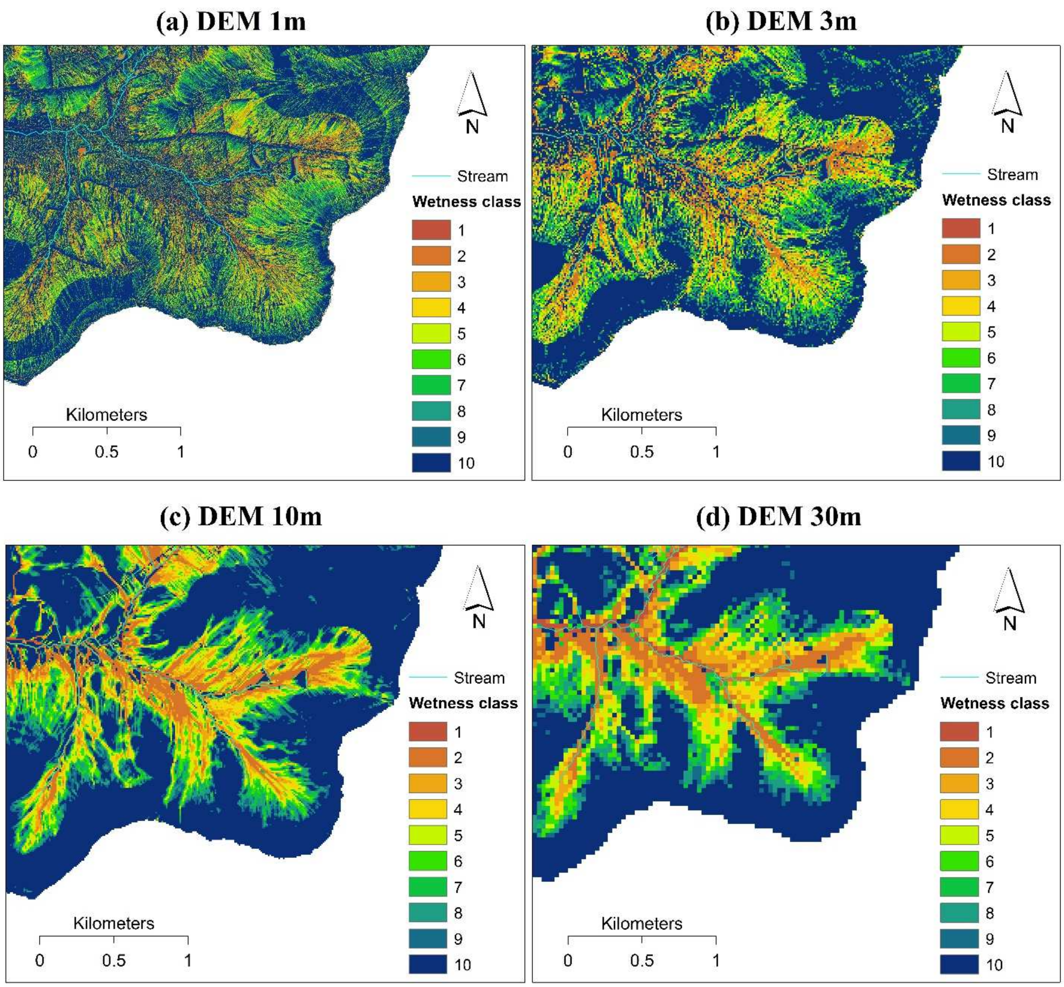



**Figure 4: Wetness maps created from DEMs with different resolutions**

Hydrology and Earth System Sciences
Open Access
Discussions

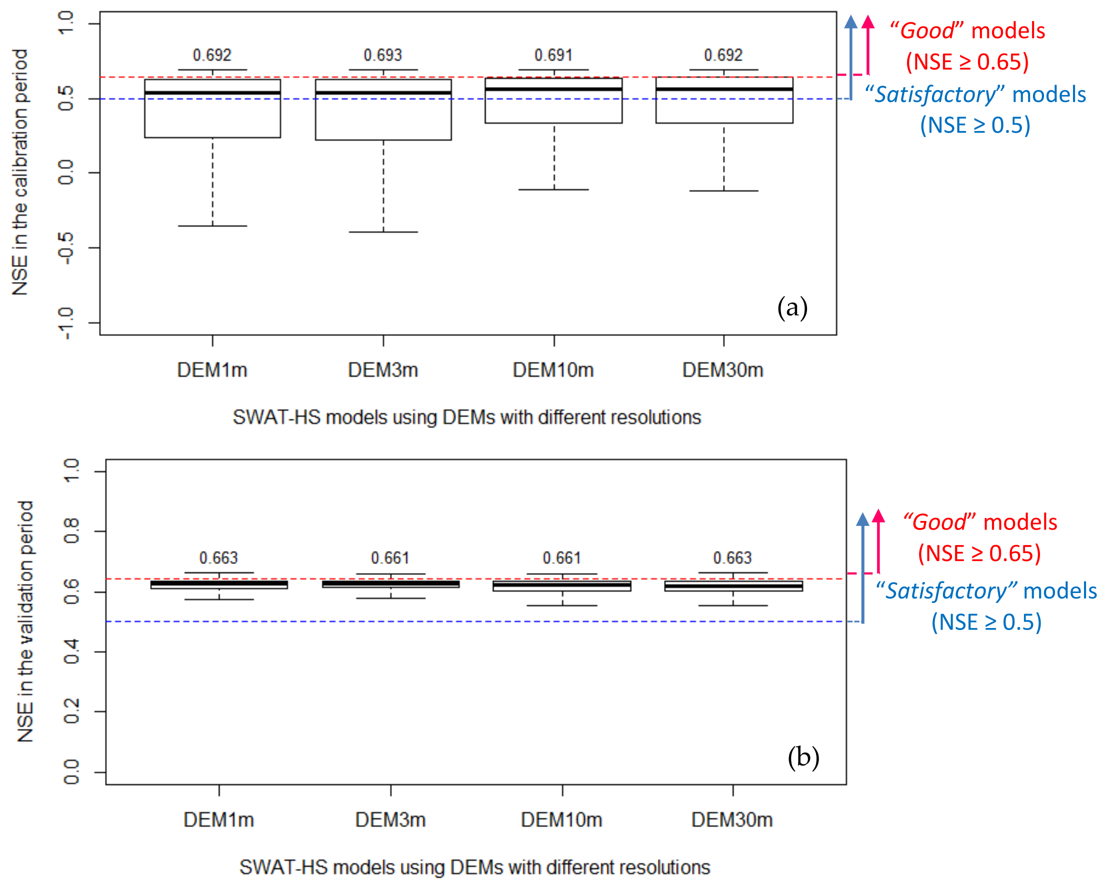



**Figure 5: Boxplots of NSE values in SWAT-HS set ups with different DEM resolutions in: (a) calibration period for 10,000 Monte Carlo parameter sets, and (b) validation period based on *"good"* performing parameter sets (*the number above each boxplot indicates the maximum NSE value for each setup*)**

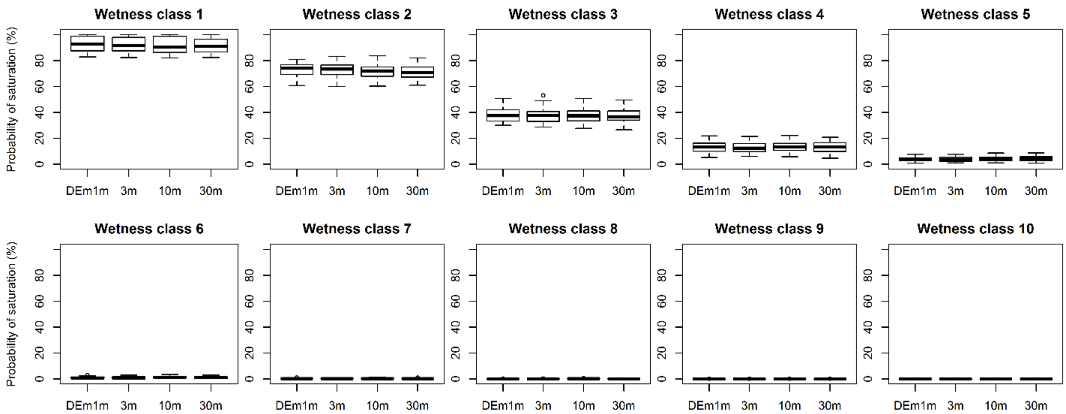



**Figure 6: Probability of saturation of wetness classes in SWAT-HS set ups with different DEM resolutions**

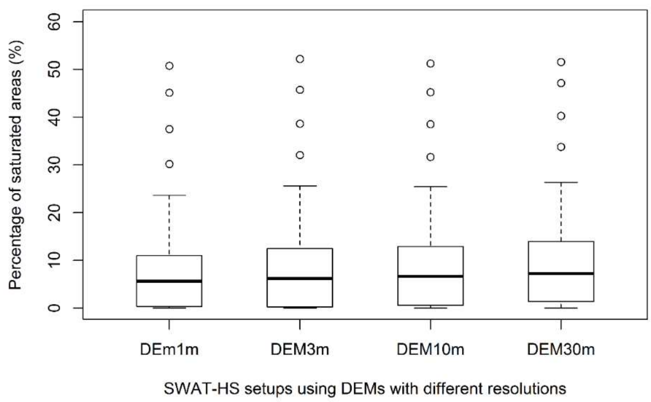



**Figure 7: Percentage of saturated areas taking into account parameter uncertainty in the calibration period in SWAT-HS setups using DEMs with different resolutions**

Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-46
Manuscript under review for journal Hydrol. Earth Syst. Sci.
Discussion started: 26 February 2018



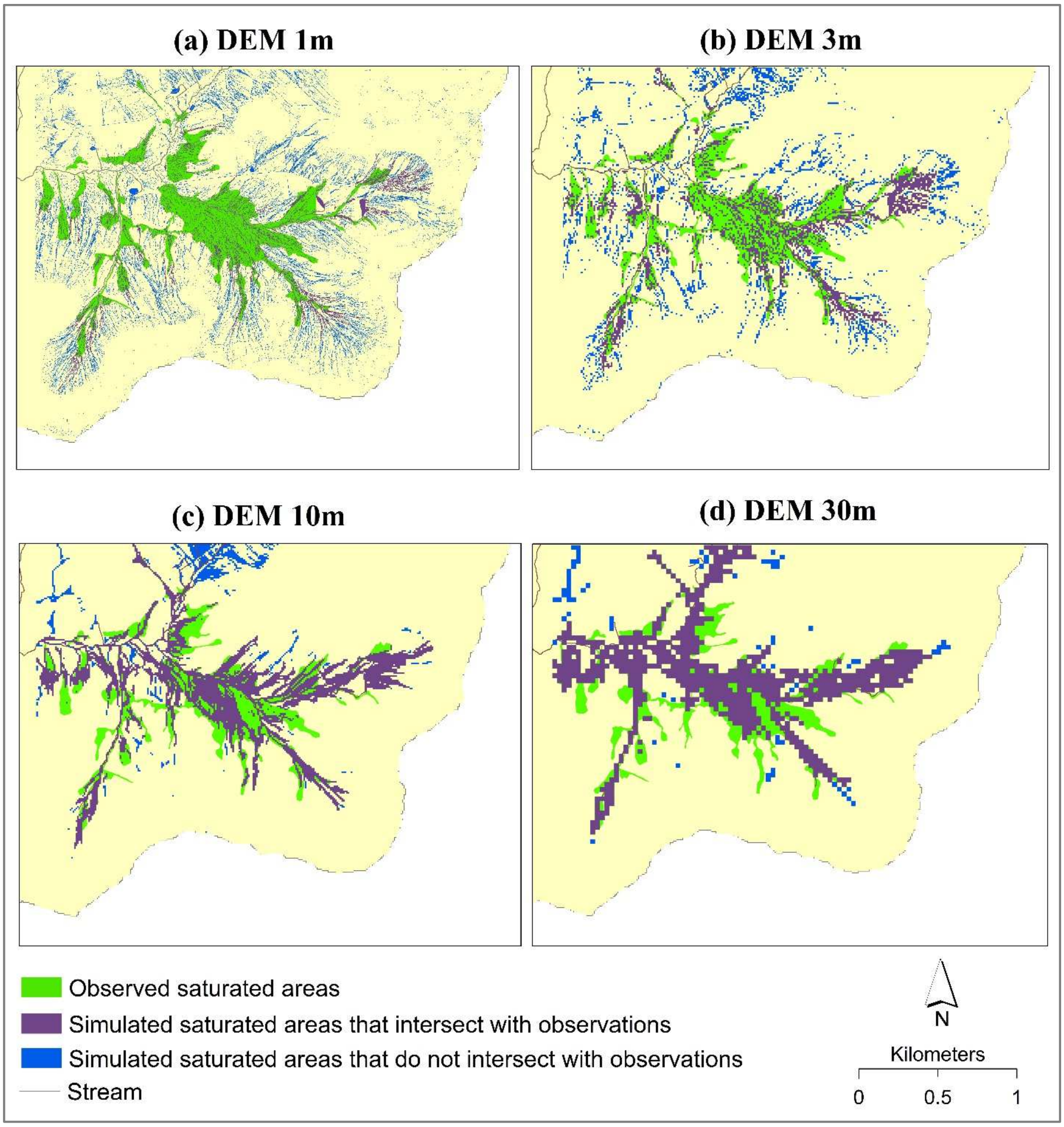



**Figure 8: Simulated and observed saturated areas from four SWAT-HS setups using different DEMs, 28-30 April 2006**

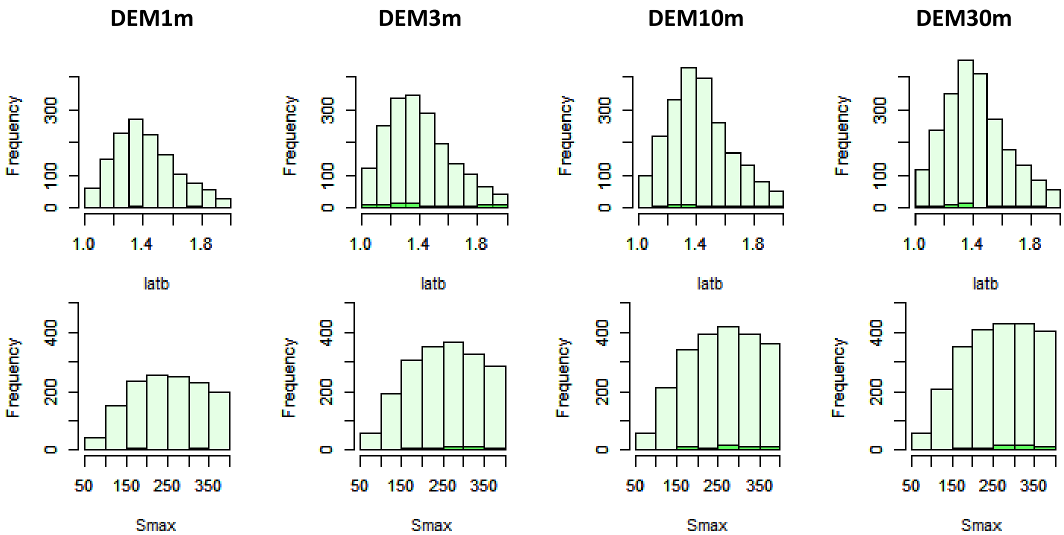
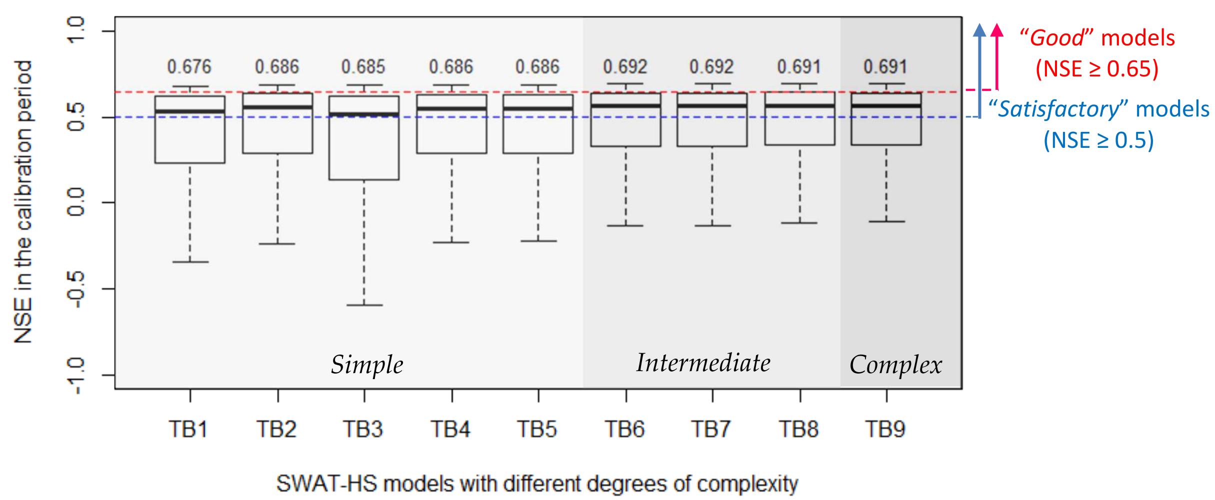
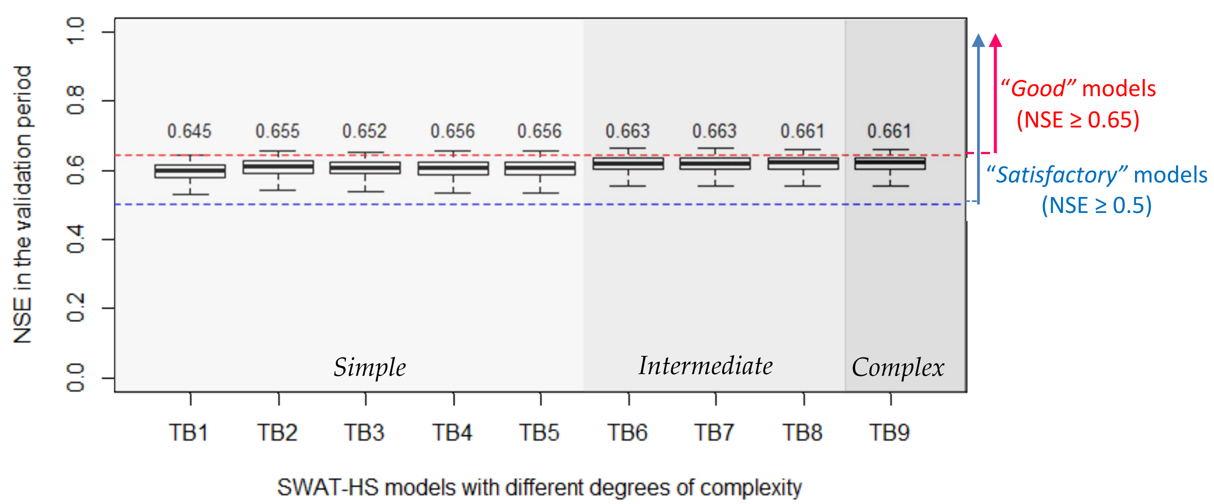



**Figure 9: Distribution of *"good"* parameters for streamflow (in light green) and for both streamflow and saturated areas (in dark green) in four SWAT-HS setups using different DEM resolutions**

Hydrology and Earth System Sciences
Open Access
Discussions




(a) **Calibration period** (based on 10,000 Monte Carlo parameter sets)




(b) **Validation period** (based on *"good"* Monte Carlo parameter sets)

**Figure 10: Boxplots of NSE values in SWAT-HS set ups with different degrees of complexity for calibration and validation periods** *(The texts above the boxplot is the maximum NSE of each setup)*

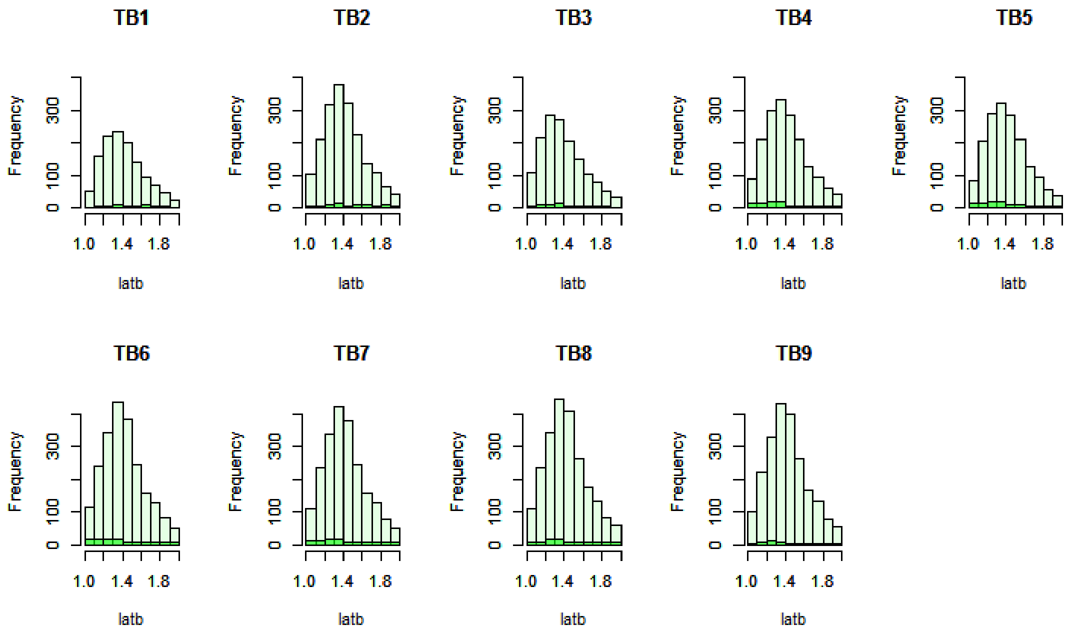



**Figure 11: Distribution of good parameter values (parameter *latb*) for streamflow (in light green) and for both streamflow and saturated areas (in dark green) in nine SWAT-HS setups with different degrees of complexity**

**Table 1: SWAT-HS parameters for streamflow calibration**

| Name | Unit | Definition | Range |
|---|---|---|---|
| *Group (i): Snowmelt calibration* | | | |
| SFTMP | °C | Snowfall temperature | -5 - 5 |
| SMTMP | °C | Snowmelt temperature | -5 - 5 |
| SMFMX | mm/°C | Maximum snowmelt factor | 5 - 10 |
| SMFMN | mm/°C | Minimum snowmelt factor | 0 - 5 |
| TIMP | - | Snow pack temperature lag factor | 0 – 1 |
| *Group (ii): Flow calibration* | | | |
| *RCHRG_PAF* | mm | Fraction of root zone percolation that recharges the surface aquifer | 0-1000 |
| *latA* | | Surface aquifer non-linear reservoir coefficient | 0 - 1 |
| *latB* | | Surface aquifer non-linear reservoir coefficient | 1 - 3 |
| *ALPHA_BF* | $days^{-1}$ | Base flow recession constant | 0 - 1 |
| *EFFPORFACTOR* | | Fraction of effective porosity that can hold water under saturated conditions | 0 - 1 |
| *EPCO* | | Plant water uptake compensation factor | 0 - 1 |
| *ESCO* | | Soil evaporation compensation factor | 0 - 1 |
| *Smax* | mm | Maximum soil water storage capacity in the watershed | 100-400 |
| *b* | | Shape parameter defining the distribution of soil water storage capacity | 0.1-3 |

**Table 2: SWAT-HS model set ups with increasing levels of complexity**

| SWAT-HS setups | Wetness classes | Soil map | Land use map | Number of HRUs | Degree of complexity |
|---|---|---|---|---|---|
| TB1 | 10 | TBsoil_**1** | TBlanduse_**1** | **10** | |
| TB2 | 10 | TBsoil_**2** | TBlanduse_**1** | **10** | |
| TB3 | 10 | TBsoil_**3** | TBlanduse_**1** | **26** | *Simple* |
| TB4 | 10 | TBsoil_**1** | TBlanduse_**2** | **30** | |
| TB5 | 10 | TBsoil_**1** | TBlanduse_**3** | **60** | |
| TB6 | 10 | TBsoil_**2** | TBlanduse_**2** | **30** | |
| TB7 | 10 | TBsoil_**2** | TBlanduse_**3** | **60** | *Intermediate* |
| TB8 | 10 | TBsoil_**3** | TBlanduse_**2** | **80** | |
| TB9 | 10 | TBsoil_**3** | TBlanduse_**3** | **146** | *Complex* |

TBsoil_**1**: homogeneous soil

TBsoil_**2**: 10 soil types (unique soil type for each wetness class)

TBsoil_**3**: 17 soil types

TBlanduse_**1**: homogenous land use (Agriculture)

TBlanduse_**2**: 3 land use types (Agriculture, Forest, and Urban)

TBlanduse_**3**: 11 land use types

1 **Table 3: Statistical criteria to compare the effect of DEM resolution on model uncertainty**

| | **DEM1m** | **DEM3m** | **DEM10m** | **DEM30m** |
|---|---|---|---|---|
| ***Calibration period: based on 10,000 Monte Carlo parameter sets*** | | | | |
| Number of *"good"* parameter sets (%) for streamflow | 1362 | 1890 | 2180 | 2293 |
| Number of *"good"* parameter sets (%) for both streamflow and saturated areas | 27 | 49 | 66 | 67 |
| Max NSE | 0.69 | 0.69 | 0.69 | 0.69 |
| Mean NSE | 0.09 | 0.05 | 0.33 | 0.34 |
| Median NSE | 0.54 | 0.54 | 0.57 | 0.57 |
| ***Validation period: based on "good" parameter sets from calibration*** | | | | |
| Max NSE | 0.66 | 0.66 | 0.66 | 0.66 |
| Mean NSE | 0.60 | 0.62 | 0.62 | 0.62 |
| Median NSE | 0.62 | 0.63 | 0.62 | 0.62 |

2

3

**Table 4: Statistical criteria to compare the effect of input complexity on model uncertainty**

| **Statistical criteria/Setup** | *Simple* | | | | | *Intermediate* | | | *Complex* |
|---|---|---|---|---|---|---|---|---|---|
| | **TB1** | **TB2** | **TB3** | **TB4** | **TB5** | **TB6** | **TB7** | **TB8** | **TB9** |
| ***Calibration period: based on 10,000 Monte Carlo parameter set*** | | | | | | | | | |
| Number of *"good"* parameter sets (%) for streamflow | 1245 | 1917 | 1510 | 1753 | 1722 | 2194 | 2144 | 2258 | 2180 |
| Number of *"good"* parameter sets (%) for both streamflow and saturated areas | 76 | 99 | 88 | 60 | 61 | 64 | 61 | 59 | 66 |
| Max NSE | 0.68 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 | 0.69 |
| Mean NSE | 0.26 | 0.30 | -0.08 | 0.30 | 0.30 | 0.34 | 0.33 | 0.34 | 0.33 |
| Median NSE | 0.53 | 0.55 | 0.51 | 0.55 | 0.55 | 0.56 | 0.56 | 0.57 | 0.57 |
| ***Validation period: based on "good" parameter sets from calibration*** | | | | | | | | | |
| Max NSE | 0.65 | 0.66 | 0.65 | 0.66 | 0.66 | 0.66 | 0.66 | 0.66 | 0.66 |
| Mean NSE | 0.60 | 0.61 | 0.60 | 0.60 | 0.60 | 0.62 | 0.62 | 0.62 | 0.62 |
| Median NSE | 0.60 | 0.61 | 0.61 | 0.59 | 0.61 | 0.62 | 0.62 | 0.62 | 0.62 |