Dear Reviewer

Thank you very much for your comment of January 18, 2016, informing us of valuable suggestions to improve our manuscript 'A comprehensive evaluation of input data-induced uncertainty in nonpoint source pollution modeling' (hess-2015-377). Our point-by-point responses are as follows.

===

Response to the Reviewer 2

1) ***Your comment***: While the results need to be interpreted carefully so that readers understand how the attributes of the study site, the authors typically guide this interpretation. For instance, the authors attribute the low uncertainty due to fertilizer inputs to the low levels of fertilizers applied in the study site.

Our respond: I agree with the reviewer's idea that the attributes of the study site should be provided. To benefit our readers, we have checked the manuscript very carefully, while revised the methodology and result sections accordingly. Pleased find our revised manuscript.

Specifically, the following sentences related to the fertilizer input have been added as: "Traditional potato-sweet potato rotation was the most popular cropping system in the agricultural areas under the slope of 15-degree, while the duration of rotations were typically half year-half year. Besides, most of the growers on the higher areas (>15-degree) panted corn, which is becoming more and more popular due to higher returns under recent market conditions. In our analysis, we studied the impacts of fertilizer and did not attempt to change the rotation pattern or introduce alternative crops. Attribute data, including crop planting time, irrigation, fertilization, and tillage, were mainly obtained from the agricultural bureau and local farmers; therefore, these data only reflect the average information at an average level. In this sense, there were inevitable differences in management practices among farmers; therefore, the use of this average information might result in fertilizer amount errors. In this analysis, the errors in the recorded amount of fertilizer applied was also treated as input uncertainty. Based on our limited local investigation, the initial annual applied urea and compound

fertilizer was set as 450kg/ha and 300kg/ha for the potato-sweet potato rotation, while 150kg/ha and 225 kg/ha for the corn system, respectively. A survey conducted by local agricultural administration revealed that the error or standard deviation in the record fertilizer amount was ±5%, which was based on a statistical analysis of historical fertilizer data."

2) ***Your comment***: I did have comments for the authors regarding their literature review and methods, which are detailed below. The general synopsis is that more work is needed to situate this study in the literature, and that the methods needs to be better described. However, I do not believe any of these criticisms are fatal to the paper itself. I believe they can all be addressed with a major revision.

Our respond: I agree with the reviewer's idea that more work is needed to situate this study in the literature, and also the methods needs to be better described. As mentioned above, we have checked the manuscript very carefully and more sentences have been added to benefit our readers. For example, in the introduction section, we added "First, there is relatively more uncertainty research about hydrological processes (Beven, 2006; Balin et al., 2010; Vrugt et al., 2008) but less on NPS pollution (Chaplot et al., 2005a; Chaplot, 2005b; Gassman et al., 2007; Wellen et al., 2015). These studies have showed the input uncertainty is propagated through the watershed model, to some extent, to sediment modeling and then carry-over and magnify into pollutant simulation. Uncertainty is currently considered as one of the core dilemmas in watershed studies, especially in the field of NPS modeling. Second, the sensitivity of watershed models also depends on how well attribute data aggregation describes the relevant characteristics of human management. For example, the SWAT assumed P could be added onto the soil in the form of fertilizer or manure, and specific input data include the timing of fertilization, the type and amount of fertilizer/manure, and the distribution of the soil layer. Thus, it is useful to understand the assumptions of these attribute data and how these assumptions will likely impact the model results. Third, previous studies have not evaluated the relative contribution of each input data set so a strategy on how to reduce input uncertainty cannot be

formulated in a cost-effective manner (Munoz-Carpena et al., 2006)."

Besides, the following references have been added:

References

Balin, D., Lee, H., and Rode, M.: Is point uncertain rainfall likely to have a great impact on distributed complex hydrological modeling?, Water Resour. Res., 2010, 46, W11520.

Beven, K.: A manifesto for the equifinality thesis. J. Hydrol. 2006, 320 (1−2), 18−36.

Chaplot, V.: Impact of DEM mesh size and soil map scale on SWAT runoff, sediment, and NO3-N loads predictions, J. Hydrol. 2005, 312, 207−222.

Chaplot, V., Saleh, A., Jaynes, D. B.: Effect of the accuracy of spatial rainfall information on the modeling of water, sediment, and NO3−N loads at the watershed level, J. Hydrol. 2005, 312, 223−234.

Cibin, R., Sudheer, K. P., Chaubey, I.: Sensitivity and identifiability of stream flow generation parameters of the SWAT model. Hydrol. Process. 2010, 24, 1133−1148.

Gassman, P., Reyes, M., Green, C., Arnold, J.: The soil and water assessment tool: Historical development, applications, and future research directions, Trans. ASABE 2007, 50 (4), 1211−1250.

Vrugt, J. A., ter Braak, C. J. F. , Clark, M. P. , Hyman, J. M., and Robinson, B. A.: Treatment of input uncertainty in hydrologic modeling: Doing hydrology backward with Markov chain Monte Carlo simulation, Water Resour. Res., 2008, 44, W00B09

Wellen, C., Kamran-Disfani, A., and Arhonditsis, G.B.: Evaluation of the current state of distributed watershed-water quality modeling, Environ. Sci. Technol. 2015, 49: 3278-3290.

*3) **Your comment**: I did have a number of questions for the authors regarding their methods, which I believe are not described in sufficient detail. The primary conceptual issue I had with their approach lies with the empirical nature of many of the parameters of SWAT. For instance, earlier studies have found that much of*

the parametric uncertainty of SWAT lies with the curve number parameters (Cibin et al., 2010). The curve numbers are very empirical, and their optimal values probably serve to compensate somewhat for the input uncertainties. If the model had been re-calibrated to each perturbed input set, the calibrated parameters would likely have compensated somewhat for the perturbed inputs in an effort to reproduce the observed data. By perturbing inputs but not re-calibrating the model to them, the authors may be overestimating the uncertainty due to the inputs. I understand that recalibrating to each perturbed input set would be quite computationally intensive, and beyond the scope of this study. I do not know if the overall results would not change significantly if the model were re-calibrated to the perturbed inputs, though the authors should mention this possible shortcoming in the methods or the discussion.

Our respond: I agree with the reviewer that the calibrated parameters would likely have compensated somewhat for the perturbed inputs. In fact, we have focused on this in one previous study (Shen et al., 2012). In that paper, we have divided the model parameters into the conceptual group and physical group. The conceptual parameters such as CN2 in the SCS curve method are defined as the conceptualization of non-quantifiable process, and determined by the process of model calibration. Conversely, physical parameters could be measured or estimated based on watershed characteristic when intensive data collection is possible. As the unknown spatial heterogeneity of studied area and expensive experiments involved, the physical parameters are usually determined by calibrating the model against the measured data.

However, when the number of parameters is large either due to the large number of sub-processes being considered or due to the model structure itself, the calibration process becomes complex and calibration uncertainty issues surround. Nevertheless, parameter identification is a non-linear problem and there might be numerous possible solutions obtained by optimization algorithms. Thus, the parameters could not be identified easily. Additionally, different parameter sets may result in similar prediction known as the phenomenon of equifinality. It has been proved that parameter uncertainty is inevitable in watershed modeling. In this sense, models are not

re-calibrated to show the differences in model predictions because calibration masks the differences that may occur as a result of different input datasets. In addition, the un-recalibrated model results can show how good each dataset predicts stream flow and NPS before recalibration, which would indicate the effort required for calibration when using each dataset.

We agree with the reviewer's idea that input uncertainty may be amplified through the calibration process. In fact, we did calibrate the SWAT using different input datasets. In fact, we did calibrate the SWAT when different datasets were used. Compared to the result of the un-recalibrated model, the relative error between predicted and observed data become smaller after recalibration, while E_{NS} and R^2 values have increased slightly. This can be due to the compensation mechanism of calibration process. Besides, we have been conducting researches on the interaction between soil data error propagation and parameter uncertainty amplified through calibration with uncertain input data.

As suggested, we discussed this in section 4.1, which is as follows: "If the model had been re-calibrated to each perturbed input set, the calibrated parameters would likely have compensated somewhat for the perturbed inputs in an effort to reproduce the observed data. However, even with the best calibration process, there is always parameter uncertainty in the model predictions due to the imprecise representation of parameter ranges and distributions; therefore, recalibration was not conducted in this study (Van Griensven et al., 2006). It should be noted that comparison using un-recalibrated models is useful to evaluate the differences in model predictions because calibration masks the differences that may occur as a result of the input data sets. In addition, the un-recalibrated model results can show how good each dataset predicts stream flow before calibration, which would indicate the effort required for calibration when using each data set."

4) ***Your comment***: The authors also mention that when they calculated the uncertainty due to all of the inputs (presented in Figure 3), they retained only behavioral inputs, which they defined as those leading to a Nash-Sutcliffe

Efficiency of greater than 0.5. This is a reasonable calibration approach for model parameters, and is used by the GLUE methodology. However, I don't see how this approach translates well to model inputs. If a (perturbed) model input gives a poor fit, but is within the uncertainty envelope of the inputs, doesn't a poor fit suggest that the model is sensitive to that input? The authors need to explain their rationale and approach better in the methodology section.

Our respond: Sorry for this confusion. In fact, the sensitivity of simulated TP to each input data was quantified in the form of summary statistics, such as the *SD* and the coefficient of variation (*CV*). Specifically, the *CV*, which is a normalized measure of dispersion of a probability distribution, is defined as a dimensionless number by quantifying the ratio of the *SD* to the *MV*. Compared to *SD*, the *CV* is more appropriate for comparing different data sets; therefore, it was used as the main approach for expressing sensitivity in this study. During this process, the *Ens* is not used and all input datasets and related simulated data are retained. In this sense, if a (perturbed) model input gives a poor fit, the higher *SD* and *CV* values of simulated data would suggest that the model is sensitive to that input.

After this sensitivity analysis, the GLUE methodology was then used to determine the prediction uncertainty by focusing on different input datasets implicitly through the likelihood measure. The key of this step is use the likelihood function to evaluate SWAT outputs against observed values. In our study, Nash–Sutcliffe coefficient (E_{NS}) was picked because it's the most frequently used likelihood measure for GLUE based on literature. To provide a static state instead of subjective personal judgment, the performance ratings typically applied to the E_{NS} by Arabi et al. (2007) were adopted: very good (0.75-1), good (0.65-0.75), satisfactory (0.50-0.65), and unsatisfactory ($\leqslant$0.5). Compared with other applications, the SWAT model was judged to be 'very good' for flow and sediment prediction and 'good' for TP prediction. Thus, 0.5 was selected to retain only behavioral inputs. However, it should be noted that the choice of 0.5 is subjective. Previously, we have used different E_{NS} values (from 0 to 0.6) as likelihood thresholds, and quantified the impacts of these values on prediction uncertainty (Gong et al., 2011). Based on the results, we

highlighted higher threshold values to increase the modeler's confidence in model reliability. Thus, 0.5 was selected.

Another question is the SWAT model was not re-calibrated. As the reviewer mentioned, the calibrated parameters might likely have compensated somewhat for the perturbed inputs in an effort to reproduce the observed data. However, when the SWAT was calibrated at the WX station for the period from 2000 to 2007, we use the best available input datasets, which contains all rainfall stations and high-resolution GIS maps. Thus, behavior input data ($E_{NS} \geq 0.5$), which refer to the phenomenon of equifinality and can be representative of a watershed system ($E_{NS} \geq 0.5$), were grouped to express the prediction uncertainty. Finally, input-induced model uncertainty was generated via sampling from the output distributions that are generated from these effective input datasets.

5) ***Your comment***: Regarding the introduction and discussion, more work is needed to situate this study in the relevant literature. A number of key statements are made with no attribution at all.

Our respond: Thank you for this valuable suggestion. I agree with the reviewer's idea that more reference are needed. As mentioned above, we have checked the manuscript very carefully and added more sentences to benefit our readers. Please find the attached manuscript. Specifically, the following references have been added:

References

Balin, D., Lee, H., and Rode, M.: Is point uncertain rainfall likely to have a great impact on distributed complex hydrological modeling?, Water Resour. Res., 46, W11520, 2010

Beven, K.: A manifesto for the equifinality thesis. J. Hydrol. 320 (1−2), 18−36 2006

Chaplot, V.: Impact of DEM mesh size and soil map scale on SWAT runoff, sediment, and NO3-N loads predictions, J. Hydrol. 312, 207−222, 2005

Chaplot, V., Saleh, A., Jaynes, D. B.: Effect of the accuracy of spatial rainfall information on the modeling of water, sediment, and NO3−N loads at the watershed level, J. Hydrol. 312, 223−234, 2005

Cibin, R., Sudheer, K. P., Chaubey, I.: Sensitivity and identifiability of stream flow generation parameters of the SWAT model. Hydrol. Process. 24, 1133−1148, 2010

Gassman, P., Reyes, M., Green, C., Arnold, J.: The soil and water assessment tool: Historical development, applications, and future research directions, Trans. ASABE 50 (4), 1211−1250, 2007

Vrugt, J. A., ter Braak, C. J. F. , Clark, M. P. , Hyman, J. M., and Robinson, B. A.: Treatment of input uncertainty in hydrologic modeling: Doing hydrology backward with Markov chain Monte Carlo simulation, Water Resour. Res., 44, W00B09, 2008

Wellen, C., Kamran-Disfani, A., and Arhonditsis, G.B.: Evaluation of the current state of distributed watershed-water quality modeling, Environ. Sci. Technol. 49: 3278-3290, 2015.

6) ***Your comment****:* P.5 Can you provide a reference where readers can find documentation of the study area's soil types? Many readers will not be familiar with these soils.

Our respond: Thank you for this valuable suggestion. Soil type is a key factor for understanding the complex and interdependent geophysical processes in the near surface. In previous studies, researchers have carried out soil samplings and measurements to build up site-specific soil databases. These databases are then extrapolated with the support of Remote Sensing and Geographic Information System (GIS) techniques to simulate surface hydrology and NPS pollutant transport in larger watersheds. However, we have provide soil information in our previous papers (Shen et al., 2012a, 2013a, b). Thus, P.5 has been revised as: "The primary land uses in this watershed are forest (61.8%), arable land (25.3%), and pasture (12.5%), and yellow-brown earths (26.5%), yellow-cinnamon soils (16.9%) and purplish soils (14.5%) are the dominant soil types. More information about the study area are referred to Shen et al., (2012a, 2013a, b)."

*7) **Your comment***: P.5. The authors should clarify in the methodology whether they refer to total phosphorus load, concentration, or flow-weighted concentration.

Our respond: As suggested, the following sentences have been added in P.5: "The model outputs were simulated flow amount, sediment load, and TP load, which were predicted at a monthly step because only monthly measured TP were available in this area."

In this study, the SWAT model was run on a daily time-step on the basis of the daily rainfall input and other daily meteorological data. The daily flow sediment, and TP loads were simulated and the monthly outputs were the sum of daily loads of the simulated daily flow, sediment, and NPS-TP. Our brief literature review indicates that may studies have shown that SWAT simulations carried out in monthly time step generally provides better prediction outputs than those in daily step. It is thus considered to be more meaningful to analyze prediction uncertainty and model accuracy on the basis of monthly time step. Thus, monthly TP load was used in this study.

*8) **Your comment***: P.7 I think most readers won't understand exactly how the authors perturbed the land use data to simulate their contribution to model uncertainty. More clarity is required here.

Our respond: As suggested, we have clarified this in the revised paper. Typically, land use and land cover changes are regarded as one of the major causal factors in altering the watershed system. The hydrological responses to different periods of land use inputs have been therefore covered by many studies. As discussed above, land use data available for the modeling effort will likely come from numerous sources; therefore, an assessment of available land use data and the time period covered by these data should be made.

Thus, P.7 has been revised as "In this study, land use data were obtained from the 1980s (1980–1989), 1995, 2000, and 2007. Specifically, maps from the 1980s, 1995 and 2000 were interpreted from MSS/TM/ETM images by the Chinese Academy of Sciences, whereas the land use map for 2007 was created from a TM

image. To substantiate the impacts of land use maps, an analytical framework was developed in two steps. Firstly, the land use distribution characteristics during each period was analyzed according to use type of each map. The land use statistics are shown in **Table 2**. Second, these four land use maps were used as model inputs and their impacts were estimated respectively using the calibrated SWAT model. In our previous study (Shen et al., 2013a), the resolution of land use data was shown to have only a slight influence on simulated NPS-P for the study region; therefore, the land use map was not resampled in this study."

9) *Your comment*: P.8. It sounds like the standard deviation is the variability in the amount of fertilizer applied. This is not necessarily the same as the uncertainty. The value of the standard deviation is not given − it would help to situate this study in the literature. Further, can the values of the mean and standard deviation be given in amounts of phosphorus applied?

Our respond: In fact, the standard deviation and CV values are referred to the errors in the recorded amount of fertilizer. In this study, attribute data, including crop planting time, irrigation, fertilization, and tillage, were mainly obtained from the agricultural bureau and local farmers; therefore, these data only reflect the average information at an average level. In this sense, there were inevitable differences in management practices among farmers. A survey conducted by local agricultural administration revealed the error or averaged deviation in the record fertilizer amount, which was based on a statistical analysis of historical fertilizer data.

To benefit our reader, P.8. has been revised as "Based on our limited local investigation, the initial annual applied urea and compound fertilizer was set as 450kg/ha and 300kg/ha for the potato-sweet potato rotation, while 150kg/ha and 225 kg/ha for the corn system, respectively. Because there was not enough information available regarding the distribution of the fertilizer, normal distribution was used in this study. Using the Monte Carlo technique, these errors were generated by sampling stochastically from a normal distribution expressed as $X \sim N(\mu, \sigma^2)$, where μ and σ

are the recorded amount of fertilizer and the standard deviation (*SD*), respectively. The Latin Hypercube sampling technique, which employs a constrained sampling scheme instead of random sampling, was applied to ensure a sufficient precision of sampling. To cover 99.7% of the error range, the sampling range was designated as ±15% from the initial amount of fertilizer and 5,000 model runs were conducted."

10) Your comment: P.9, line 13. There are also calibration data uncertainty and parameter uncertainty. A citation would strengthen this line.

Our respond: Sorry for those confusing statements and these sentences have been revised accordingly. In fact, we have distinguished calibration data uncertainty and parameter uncertainty in one previous study (Chen et al., 2014). Based on our works, we found appreciable inherent errors exist in the calibration data even when following strict quality assurance and quality control (QA/QC) guidelines. Harmel (2005) has mentioned calibration data uncertainty may stem from errors in flow measurements, water quality sample collection, the processes of preservation, storage, transport and laboratory analysis. Given the river discharge data, errors from different sources such as river stage measurement or the interpolation of the rating curve, affect the measured data. In a thorough review (Harmel et al., 2006), all possible errors in the H/WQ measured data were compiled.

In comparison, the cause of parameter uncertainty may be due to the value of the parameter being case-specific. The process of parameter calibration is a complex and subjective task. For these reasons, decisions regarding modeling should be based on available knowledge about the range and associated probability distribution function (PDF) for each parameter. Model calibration is the process of estimating model parameters using a pair-wise comparison between the predicted and measured points, while the validation process involves running the well-calibrated model to check its performance. Nevertheless, parameter identification is a complex non-linear problem, so the parameters could not be identified easily due to the numerous possible solutions obtained by optimization algorithms. Additionally, different parameter sets may result in similar pre-dictions in a phenomenon known as equifinality. In this

study, models are not re-calibrated to show the differences in model predictions because calibration masks the differences that may occur as a result of different input datasets.

Thus, we have designed an interval-deviation approach (IDA) and incorporated it into likelihood functions with the support of interval theory (Chen et al., 2014). The proposed IDA was validated in a real application of the Soil and Water Assessment Tool (SWAT) and Generalized Likelihood Uncertainty Estimation (GLUE) in the Three Gorges Reservoir Area (TGRA), China. Compared with the traditional point estimates of observations and predictions, the IDA incorporated both prediction and measurement uncertainty into the process of model evaluation. This proposed IDA can be extrapolated into other forms of error indices and model to provide a substitute method of model evaluation within an uncertainty framework.

11) ***Your comment:*** P.10 The coefficient of variability (CV) as expressed in Eq. 2. assumes a normal distribution. However, the distributions used to estimate the predictive uncertainty may be highly skewed, in which case the CV would need to be calculated with a different equation.

Our respond: Thank for this valuable suggestion. In fact, we did compare different probability distribution functions (PDFs) for each fertilizer data and quantify their impacts on the model outputs. Based on the results, we found that the choice of PDFs had little impacts on the predicted NPS-TP so we only assumed that the fertilizer data were identically chosen from normal distribution spanning the feasible error range due to their simplicity.

It should be noted when intensive data collection is possible, the range and distribution of each data should be measured or estimated according to local detailed agriculture practices. However, such experiments are often impossible for many reasons, such as the unknown spatial heterogeneity of watershed, high costs and time constraints, as well as the experience of the personnel involved. Alternatively, we collect and estimate the PDFs of data by the means of a statistical process. Then from the perspective of statistical estimation, the question could be rephrased as the

problem of selecting the proper PDFs. According to Sohrabi et al. (2003), PDFs could be assigned by professional judgment and literature information. In common, theoretical and empirical PDFs can be adopted instead of data collection and monitoring. It is usually to adopt classic PDFs such as uniform, normal, and logarithmic normal (Bobba et al. 1996) to evaluate the propagation of input uncertainty o on model outputs (Vrugt et al. 2003).

12) *Your comment*: P.11 The number of gauges beyond which improvements to the model predictions are not found should be normalized to the area of the study site.

Our respond: I agree with the reviewer's idea that the number of gauges beyond which improvements to the model predictions are not found should be normalized to the area of the study site. For this study area (2,421 km^2), the optimal number of gauges were identified as 6 beyond which improvements to the model predictions would not be found. It is encouraging that a small number of gauges distributed more optimally and perform well for logistical reasons. In reality, there might not be many dense rain gauge networks similar to those used for this study; therefore, the fact that spatial rainfall variation is a function of key gauges rather than all gauges would indicate a wider range of applicability.

However, we also found the response of hydrology models to rainfall input is scale dependent. Traditionally, the rain station is the fundamental tool for representing spatial distribution of rainfall within a watershed (Andréassian et al., 2001). Designing the proper location, number and density of rain-gauge stations is important to hydrological research. Fu et al (2011) find that the effect of rainfall input on discharge modeling is relatively low for catchment sizes above 250 km^2, and even negligible for watersheds larger than 1000 km^2. In general, a well-located station might be sufficient for watersheds up to about 50 ha, while 20 km is demonstrated as the threshold distance between stations for a reliable hydrology modeling (Vischel and Lebel, 2007).

Thank you very much for your wonderful job. Hope that our responses are satisfactory, and look forward to hearing from you.

Zhenyao Shen

Professor

School of Environment,

Beijing Normal University

Beijing,

China, 100875

Tel/Fax: +86-10-5880 0398

E-mail: zyshen@bnu.edu.cn

References

Bobba AG, Bourbonniere RA, Singh VP, Bengtsson L. Numerical simulation model of fatty acids in lake sediments. Water Air Soil Poll. 1996, 89:77-90.

Chen L, Shen ZY*, Yang XH, Liao Q, Shaw L. Yu. An Interval-Deviation Approach for Hydrology/Water Quality Model Evaluation. Journal of Hydrology. 2014, 509C: 207-214.

Harmel, R.D., Cooper, R.J., Slade, R.M., Haney, R.L., Arnold, J.G. cumulative uncertainty in measured streamflow and water quality data for small watersheds. T. ASABE 2006, 49, 689-701.

Harmel, R.D., King, K.W. Uncertainty in measured sediment and nutrient flux in runoff from small agricultural watersheds. T. ASABE 2005, 48, 1713-1721.

Fu, S.H., Sonnenborg, T.O., Jensen, K., He X. Impact of Precipitation Spatial Resolution on the Hydrological Response of an Integrated Distributed Water Resources Model. Vadose Zone J. 2011, 10: 25-36.

Gong Y.W., Shen Z.Y., Hong Q., Liu R.M., Liao Q. Parameter uncertainty analysis in watershed total phosphorus modeling using. Agriculture, Ecosystems and Environment. 2011, 142: 246– 255.

Harmel, R.D., Cooper, R.J., Slade, R.M., Haney, R.L., Arnold, J.G.. cumulative uncertainty in measured streamflow and water quality data for small watersheds. T. ASABE 2006, 49, 689-701.

Harmel, R.D., King, K.W.. Uncertainty in measured sediment and nutrient flux in runoff from small agricultural watersheds. T. ASABE 2005, 48, 1713-1721.

Shen ZY, Chen L, Chen T. 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. 2012, 16: 121-132.

Sohrabi TM, Shirmohammadi A, Chu TW, Montas H, Nejadhashem AP. Uncertainty Analysis of Hydrologic and Water Quality Predictions for a Small Watershed Using SWAT2000. Environ. Forensics 2003, 4: 229-238.

Vischel, T., Lebel, T. Assessing the water balance in the Sahel: Impact of small scale

rainfall variability on runoff. Part 2: Idealized modeling of runoff sensitivity. J. Hydrol. 2007, 333:340-355.

Vrugt JA, Gupta HV, Bouten W, Sorooshian S. A shuffled complex evolution metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters. Water Resour. Res. 2003, 39:1201-1216.