

Multi-method global sensitivity analyses – results for a rainfall-runoff model

Carolina Massmann, Hubert Holzmann

University of Natural Resources and Life Sciences, Vienna
(carolina.massmann@boku.ac.at, hubert.holzmann@boku.ac.at)

Abstract: A sensitivity study was carried out for the discharge of a small Austrian catchment using three global sensitivity methods implemented in a conceptual rainfall-runoff model: Sobol's method, the Mutual Entropy and the Regional Sensitivity Analysis (RSA). Since RSA is a graphical method, the Kolmogorov-Smirnov statistic was used for obtaining a quantitative measure of the sensitivities. It was observed that the parameter rankings as well as the temporal sensitivity dynamics agreed in general between the methods. However, the agreement was best between Sobol's method and the Mutual Entropy. The graphical RSA method, gave some additional insights about the relationship of parameter values and discharge levels, which were not supplied by the other two methods. Finally, the implications these findings have for model calibration are discussed.

Keywords: calibration; model sensitivity; parameter fixing; parameter interaction;

1 INTRODUCTION

Despite the emergence of new measurement tools and the important developments in remote sensing based applications in the last decade, hydrology can still be regarded as a data limited science [Kirchner, 2006]. This limited availability of data means that hydrologists need to rely heavily on calibration. One tool that can help in model calibration are sensitivity analyses, which describe the change in output resulting from a change in the model inputs or parameters. These methods are usually classified as local or global sensitivity analyses, where the former describe the change caused by one input parameter at the time and usually only for an optimum parameter set, while the latter consider the whole parameter input space [Mattot et al., 2009]. Sensitivity analysis can help to decide which parameters should be calibrated which is sometimes not easy to define. There might be, for instance, parameters affecting more than one process, parameters which are very important albeit for a short time, parameters which are moderately relevant during the whole period, or parameters which are only important in combination with others. Similarly, sensitivity analyses can be used for identifying unimportant parameters which can be then fixed at predefined values. Such an approach reduces the number of parameters to calibrate and results therefore in simpler models.

The objective of this paper is to compare three global sensitivity analysis methods in terms of the information that can be obtained from them as well as the computational effort they require. We selected methods that provide information about the first order indices: Sobol's method, the Mutual Entropy and the Regional Sensitivity Analysis (RSA). Sobol's method was included because it allows additionally the computation of the sensitivities to the interactions between parameters. The RSA method was chosen since it is basically a graphical method,

which can be however complemented with other approaches to provide quantitative sensitivity information.

2 METHODOLOGY

2.1 Sensitivity analysis methods

Sobol's Method

This method belongs to the family of variance decomposition approaches, characterised by the decomposition of the variance into different terms which are used for constructing sensitivity indices. If we had many Monte Carlo runs for our model in which we varied n parameters, it would be possible to calculate the total variance of the results (Var). This variance can be allocated to each parameter and combination of parameter interactions:

$$Var = \sum_i^n Var_i + \sum_i^n \sum_{j=i+1}^n Var_{ij} + \sum_i^n \sum_{j=i+1}^n \sum_{k=j+1}^n Var_{ijk} + \dots + Var_{12\dots n} \quad (1)$$

where Var_i describes the variance explained by each parameter i individually and Var_{ij} describes the variance explained by the interactions between two parameters i and j and so on. The first order (S_i) and total indices ST_i are then expressed as:

$$S_i = \frac{Var_i}{Var} \quad (2) \qquad ST_i = 1 - \frac{Var_{-i}}{Var} \quad (3)$$

where Var_{-i} stands for the variance not explained by i and is calculated by adding all the terms of Eq. (1) which do not include the variable i . From equations 2 and 3 it is seen that the first order indices describe the percentage of the total variance explained by a parameter on its own, while the total indices indicate the percentage of the variance explained by a parameter including all its interactions with other parameters. The proportion of the variance explained by the interactions involving a specific parameter is therefore calculated as the difference between the first and the total indices.

Many methods can be used for carrying out a decomposition of the variance. We used Sobol's method, since it is easy to implement. The method is based on one large matrix in which the rows are parameter sets having one value for each parameter (columns). This matrix is splitted into two matrices (M and N) with the same number of rows in each of them. The variances (denominator of eq. 2 and 3) are calculated using the model results obtained when running the model with the parameter sets in these matrices. The numerators require in addition the model results using parameter values of similar matrices in which specific columns (depending on the parameter of interest) are exchanged between the matrices N and M. For details the reader is referred to Cibin et al. [2010] and Saltelli [2002].

Mutual Entropy

Entropy can be regarded as a measure of the information content that one variable has [Mogheir et al., 2004]. It is closely related to sensitivity, since it is expected that a model will be more sensitive to a parameter carrying more information than to another one carrying less. Because the mutual information, $I(x, y)$, is the most commonly used entropy measure, it was selected for this study:

$$I(x, y) = - \sum_{i=1}^n \sum_{j=1}^m p(x_i, y_j) \ln \left[\frac{p(x_i, y_j)}{p(x_i)p(y_j)} \right] \quad (4)$$

where x_i , with $i = 1, 2, \dots, n$, and y_j , with $j = 1, 2, \dots, m$, are two discrete variables and the terms $p(x_i)$ and $p(x_i, y_j)$ stand for the probability of x_i and the joint

probability of x_i and y_j . The units of the Mutual Entropy depend on the base of the logarithm that is used. When the natural logarithm is used, as done here, the corresponding unit is 'nats'. The approach for calculating the probabilities is based on the construction of contingency tables and was taken from Mogheir et al. [2004] and Mishra et al. [2009]. After several Monte Carlo runs are carried out, the results of the dependent variable (for instance discharge) are classified into classes (columns) with increasing discharge and an equal number of observations in each class. Similarly, for the independent variable of interest, the observations are classified into equally sized classes (rows) with increasing values. The probability $p(x_i)$ is then calculated as the ratio between the number of observations in row i and the total number of observations and $p(x_i, y_j)$ as the ratio between the observations at the intersection of row i and column j and the total number of observations.

Regional Sensitivity Analysis (RSA)

This is a graphical method also based on Monte Carlo runs. In classical RSA these parameter sets are classified into two groups: behavioural and non-behavioural, depending on how well they can reproduce the observed behaviour of the system [Jakemann et al., 1990]. The approach we used classified the parameter sets into 10 equally large groups with increasing discharge according to the discharge they simulated as done by Wagener and Kollat [2007]. In each group the parameter values (and their corresponding discharge) are sorted in increasing order. The cumulative discharge is then calculated and plotted as a function of the parameter values.

The modelled system is then said to be sensitive to the parameters which have different cumulative distributions for the discharge deciles. Analogously, the system is regarded as non-sensitive to parameters with similar curves for all deciles. The Kolmogorov-Smirnoff d statistics can be used for quantifying the sensitivity. This value is a descriptor of the distance between the cumulative distribution of the first and the last deciles.

2.2 Modelling

The model used is a bucket type conceptual rainfall-runoff model. The main input into the model is the rainfall on the Rosalia catchment, located in Lower Austria. This catchment is mostly covered by forests and has a size of 2.35 km². The model output is the stream discharge (volume of water passing through the gage per time unit) catchment outlet. The sensitivity analyses presented here consider eleven parameters which will be shortly described. For additional information about the model the reader is referred to Holzmann and Nachtnebel [2002].

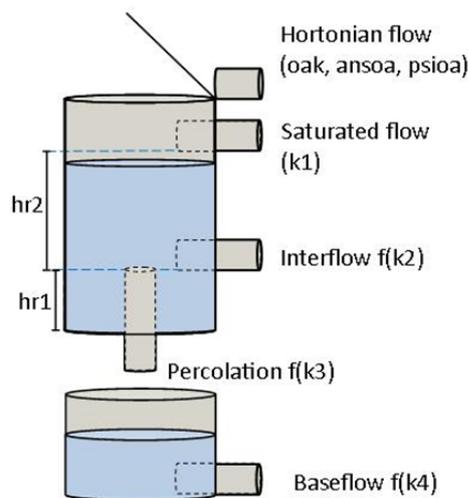


Figure 1: Structure of the rainfall runoff model used.

The model structure, consisting of a soil and a groundwater storage, is shown in Figure 1. The size of the soil storage is described by the parameters $hr1$ and $hr2$. The three outlets of the soil storage correspond to saturation flow, interflow and percolation. They are a function of the recession coefficients $k1$, $k2$ and $k3$, respectively. The outflow from the groundwater storage is released as baseflow as a function of the parameter $k4$. If the rainfall intensity is larger than the infiltration capacity of the soil ($ansoa$), Hortonian flow is activated. This means that a proportion of the rainfall ($psioa$) is routed to another linear storage with a fast

outflow described by *oak*, while the remaining part of the rainfall ($1-psioa$) is routed to the soil storage. The model also takes snow melt into account which contributes to runoff as a function of the parameters *proz* and *sk*.

For the Mutual Entropy and RSA 260.000 randomized model runs were carried out. Sobol's method required additionally 11 x 260.000 runs. The parameter ranges were defined based on previous studies. The results were calculated using daily discharges for an event between the 5th of November and 5th of December 1994.

Finally, we compared the rankings provided by the three sensitivity methods. All parameters with Sobol's indices larger than 0.01 were considered in the analysis. Parameters with smaller indices were neglected since they are regarded as unimportant and are also subject to numerical errors which are large in comparison to their sensitivity value. The parameters were ranked for each method and each day. Then, the number of places by which the ranking of all parameters differed on each day was computed. For instance, if parameter *a* had a ranking of 3 for one method and of 1 for another, then there is a difference of two places for this parameter.

3 RESULTS AND DISCUSSION

Results for the three sensitivity methods

The results for Sobol's indices are shown for a rainfall event that took place in November 1994. The measured rainfall and hydrograph are shown in the first panel of Figure 2. The first order indices (Fig. 2, second panel) indicate which proportion of the variance is explained by the respective parameter on its own. Each row shows the sensitivity for a parameter in the considered period, with a colour coding of dark blue for insensitive parameters and red colour indicating high sensitivities. It is seen that the rainfall sharply increases the importance of some parameters at the expense of others that were important before the event. Specifically, the parameters *hr2* and *k3* reduce their impact considerably with the onset of the rainfall which activates primarily the parameter *ansoa* and, to a much lesser degree, the parameter *oak*.

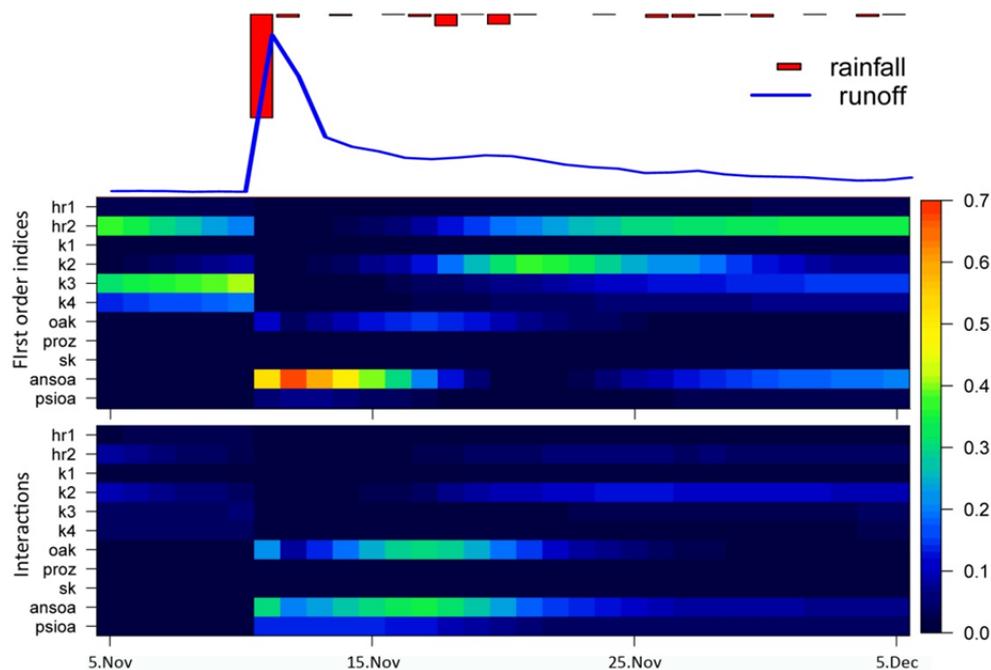


Figure 2: Precipitation and measured hydrograph for the event (first panel); Sobol's first order indices (second panel); Sobol's interactions (third panel).

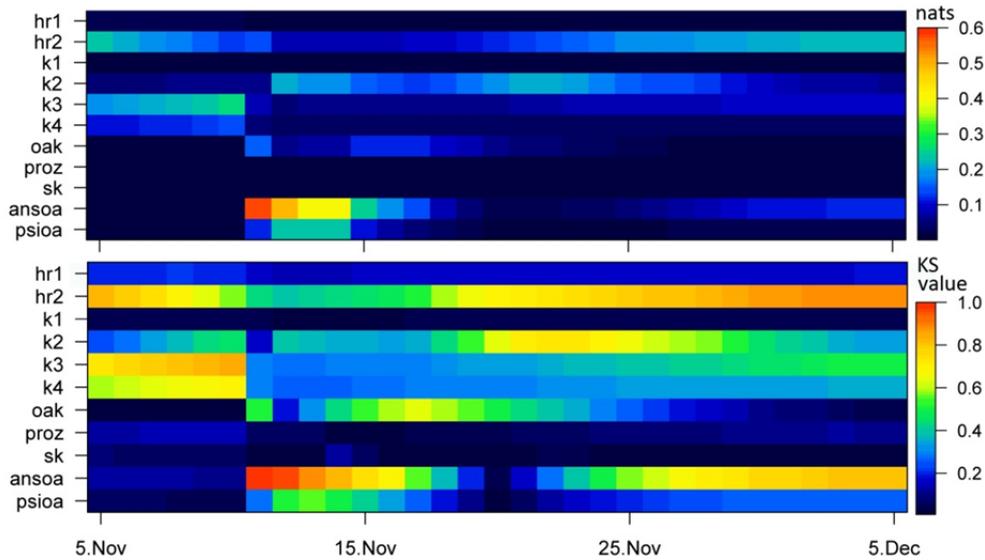


Figure 3: Mutual Entropy in nats (first panel) and RSA quantified with the Kolmogorov-Smirnov statistic (second panel) for the event.

The interactions (Fig. 2, third panel) are only important during the period with highest discharges, when the interactions involving the parameters *ansoa* and *oak* reach a value of around 0.6. The parameter *oak* has a larger impact in combination with other parameters than individually (compare the first order indices with the interactions).

Figure 3 shows the results for the Mutual Entropy in the first panel and the RSA method in the second. It can be seen that the Mutual Entropy shows a very good agreement with Sobol's first order indices. The main differences are the higher sensitivities of the Mutual Entropy during the discharge peak (around the 11th of November) for the parameters *psioa* and *k2*, which show almost no sensitivity in Sobol's method. The main characteristic of the RSA plot is that there are for each time step many parameters with a high sensitivity. This is an important difference with the other methods, for which there are in most cases a much smaller number of important parameters of which one can be clearly identified as being the best. The sensitivity trends and the temporal dynamics agree, however, with the other two methods. This method also shows a good agreement with the Mutual Entropy method during the discharge peak, since the parameters *psioa* and *k2* have also a high sensitivity (unlike Sobol's method).

Implications for model calibration

The results of a sensitivity analysis define specific windows during which the model is sensitive to each parameter. For the parameter *k2*, for instance, this period would lie between the 20th and 25th. When calibrating this parameter the modeller should therefore focus on this period. Inversely, by analysing the temporal distribution of the model errors it might be possible to identify which parameters or processes are responsible for the specific disagreement patterns [Reusser and Zehe, 2011]. Another alternative for using the results of such a sensitivity analysis for calibration is by identifying the parameters to which the system is not sensitive, in this case the parameters *hr1*, *k1*, *proz*, and *sk*. These parameters could be fixed at certain predefined value reducing in this way the number of parameters that need to be calibrated. Finally, it must be noted that the information of the sensitivities toward parameter interactions can also be useful, since it is easier to calibrate parameters having high first order indices and low interactions than parameters which have mostly on effect through interactions [Ratto et al., 2007]. This suggests that in the example above (Fig. 2) it might be beneficial to calibrate first the parameter *ansoa* and only then the parameter *oak*.

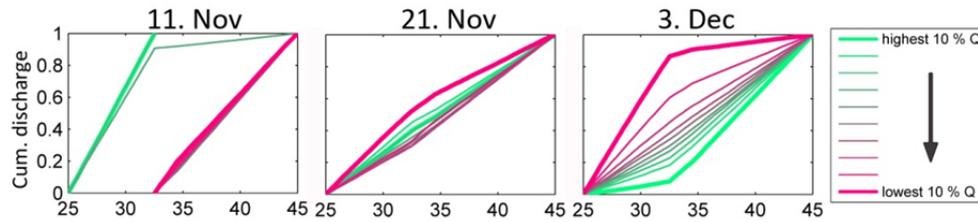


Figure 4: RSA plots for the parameter *ansoa* (x-axis) at three days during the event in November 1994.

Additional information provided by the Regional Sensitivity Analysis

As an example of the additional information that the graphical RSA can provide, we show the plots for the parameter *ansoa* at three days during the event analysed before. A first overview of the plots (Fig. 4) shows that the largest sensitivities are observed on the 11th and the smallest on the 21st of November, when there is little difference between the lines. This is consistent with the results of the other methods (e.g. Fig. 2, second panel), where it is seen that the highest sensitivities are around the 12th. From this point on they start to decline until they reach a minimum around the 20th, and start to increase again after that.

The RSA method gives, however, some additional information not provided by the other two methods, namely about the relationship between the parameter values and the discharge level as explained in the following sentences. The plot of November 11th shows that high discharges are the result of *ansoa* values smaller than 33, while low discharges are related to *ansoa* values larger than 33. This is so because Hortonian flow is only activated if *ansoa* is smaller than the effective rainfall intensity, which reached 33 mm/h for this event. Since Hortonian flow is a fast processes that routs a percentage of the rainfall directly into the stream, bypassing the soils storage (which releases the water more slowly), The activation of Hortonian flow leads to an increased discharge on the short term. The situation is different on December 3rd, where high discharges are related to low *ansoa* and low discharges to high *ansoa* values. At this time, the part of the rainwater which was directly routed to the stream when Hortonian was activated (i.e. when *ansoa* is smaller than 33) is “missing” now in the soil storage, resulting in lower discharges. On the contrary, when Hortonian flow did not take place, the soil storage has more water and thus produces a higher discharge.

Comparison of the sensitivity rankings between the different methods

A comparison of the number of places by which the rankings of the different methods differ in each day is shown in Fig. 5. It is seen that the rankings provided by Sobol’s method and the Mutual Entropy have the smallest difference in ranking. Therefore, they are the methods with the best agreement (on average there is a ranking difference of 1.14 places for each parameter with a differing ranking between both methods).

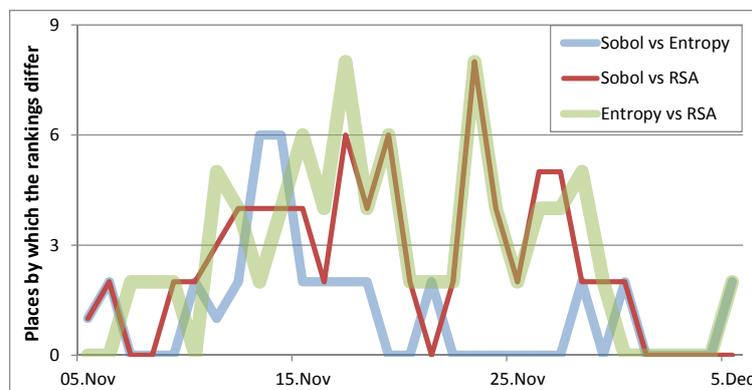


Figure 5: Number of places by which the rankings differ between the methods

A comparison between the other methods shows larger differences. The average ranking difference for the parameters not agreeing is of 1.26 and 1.23 places when comparing the Mutual Entropy with RSA and Sobol' method with RSA, respectively. It can also be seen that all methods have a similar ranking at the beginning and at the end of the analysed period, with differences between the rankings equal or less than two places. When the peak is observed there are larger differences in the rankings obtained with different methods.

4 SUMMARY AND CONCLUSIONS

Three global sensitivity methods, namely Sobol's indices, the mutual entropy and the Regional Sensitivity Analysis (RSA) approach were implemented in a conceptual rainfall-runoff model for a small catchment in Austria. It can be summarized that:

- **Sobol's** first order indices and the **Mutual Entropy** show a good agreement in most cases. The most important differences are observed during the day of peak discharge, where the Mutual Entropy shows some sensitivity for parameters that are not important for Sobol's method. Therefore, it is possible to obtain similar results using either 12 x 220.000 runs (Sobol's method) or only 260.000 runs (Mutual Entropy). This indicates that the Mutual Entropy is more suitable for models with long computing times.
- On the other side, it must be considered that Sobol's method allows in addition the computation of the effect that parameter interactions have on the model results.
- The **Regional Sensitivity Analysis**, which is a graphical method, was complemented with the Kolmogorov-Smirnoff statistics for obtaining quantitative information about the sensitivities. We carried out 260.000 runs and found that, while the general trends and the temporal dynamics of the sensitivities agreed with the other two methods, it could be observed that this method showed high sensitivities to more parameters than the other two approaches. The graphical RSA was, on the other hand, the only method that shows the relationship between the values of the model parameters and the discharge levels, allowing inferences about the mechanisms that are responsible for the impact of the considered parameter.
- The ranking between the three methods shows a good agreement. The average ranking difference for the parameters not agreeing (in their ranking placement) is smaller than 1.3 places.

With respect to the use of the information gained from global sensitivity analyses, it can be concluded that:

- The parameters to which the model is sensitive vary with the discharge level. For baseflow the most important parameters were *hr2* and *k3*, during the peak discharges *ansoa* (with some interactions with *psioa* and *oak*) are the relevant parameters and during the falling limb of the hydrograph *k2* and *hr2* are important.
- These methods are adequate for the distinguishing between parameters to which the system is sensitive to which it is insensitive. By fixing insensitive parameters to predefined values and focusing only on the sensitive parameters, less parameters need to be calibrated, thus contributing to the development of parsimonious models.
- The methods allow the identification of the temporal patterns of parameter sensitivity. This tells the modeller on which features to focus on when calibrating specific parameters. They can also be used to assess if the assigned parameter ranges give realistic results (e.g. if the periods during which certain parameters are active are consistent with the understanding of the processes they describe).

- The identification of the parameters that have mostly an effect through their interaction with other parameters (more than on its own) is important for the definition of the order in which the parameters should be calibrated.

ACKNOWLEDGMENTS

This work was funded by the Austrian Academy of Science through the project Domina_HyPro. The data for the Rosalia catchment was provided by Josef Gasch from the Forest Demonstration Centre, University of Natural Resources and Life Sciences, Vienna.

REFERENCES

- Cibin, R., K.P. Sudheer and I. Chaubey, Sensitivity and identifiability of stream flow generation parameters of the SWAT model, *Hydrological Processes*, 24, 1133-1148, doi: 10.1002/hyp.7568, 2010.
- Holzmann, H. and H.P. Nachtnebel, Abflussprognose für mittelgroße Einzugsgebiete - Methodik und Anwendungen [Runoff forecast for medium sized basins – methods and applications], *Österreichische Wasser- und Abfallwirtschaft*, 54(9-10), 142-153, 2002.
- Jakeman, A.J., F. Ghassemi and C.R. Dietrich, Calibration and reliability of an aquifer system model using generalised sensitivity analysis, *Proceedings ModelCARE 90: Calibration and Reliability in Groundwater Modelling*, IAHS publication 195, 1990.
- Kirchner J.W., Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology, *Water Resources Research*, 42, W03S04, doi:10.1029/2005WR004362, 2006.
- Matott L.S., J.E. Babendreier, and S.T. Purucker, Evaluating uncertainty in integrated environmental models: A review of concepts and tools, *Water Resources Research* 45, W06421, doi: 10.1029/2008WR007301, 2009.
- Mishra, S., N. Deeds and G. Ruskauff, Global sensitivity analysis techniques for probabilistic ground water modelling, *Ground Water*, 47, 727–744, doi: 10.1111/j.1745-6584.2009.00604.x, 2009.
- Mogheir, Y., J.L.M.P de Lima and V.P. Singh, Characterising the spatial variability of groundwater quality using the entropy theory: I. Synthetic data. *Hydrological Processes*, 18, 2165-2179, doi:10.1002/hyp.1465, 2004.
- Ratto, M., P.C. Young, R. Romanowicz, F. Pappenberger, A. Saltelli and A. Pagano, Uncertainty, sensitivity analysis and the role of data based mechanistic modeling in hydrology, *Hydrology and Earth System Sciences*, 11, 1249-1266, doi:10.5194/hess-11-1249-2007, 2007.
- Reusser, D. E., and E. Zehe, Inferring model structural deficits by analyzing temporal dynamics of model performance and parameter sensitivity, *Water Resources Research*, 47, W07550, doi:10.1029/2010WR009946, 2011
- Saltelli, A., Making best use of model evaluations to compute sensitivity indices, *Computer Physics Communications*, 145(2), 280-297, doi:10.1016/S0010-4655(02)00280-1, 2002.
- Wagener, T. and J. Kollat, Numerical and visual evaluation of hydrological and environmental models using the Monte Carlo analysis toolbox, *Environmental Modelling & Software*, 22, 1021-1033, doi:10.1016/j.envsoft.2006.06.017, 2007.