Uncertainty and Parameter Sensitivity of the KINEROS2 Physically-Based Distributed Sediment and Runoff Model.

Menberu B. Meles, USDA-ARS, Southwest Watershed Research Center, Tucson, AZ, Menberu.Bitew@ars.usda.gov
Dave C. Goodrich, USDA-ARS, Southwest Watershed Research Center, Tucson, AZ, Dave.Goodrich@ars.usda.gov
Hoshin V. Gupta, University of Arizona, Department of Hydrology and Water Resources, Tucson, Arizona, USA, hoshin.gupta@hwr.arizona.edu
I. Shea Burns, University of Arizona, School of Natural Resources and the Environment, Tucson, Arizona, USA, isburns@email.arizona.edu
Carl L. Unkrich, USDA-ARS, Southwest Watershed Research Center, Tucson, AZ, carl.unkrich@ars.usda.gov
Saman Razavi, University of Saskatchewan, Global Institute for Water Security, Saskatoon, Canada, saman.razavi@usask.ca
D. Phillip Guertin, University of Arizona, School of Natural Resources and the Environment, Tucson, Arizona, USA, dguertin@cals.arizona.edu

Introduction

Advanced hydrologic modeling tools play an important role in developing sustainable rangeland and water resource management systems, including the implementation of real-time flood forecast and warning systems (Creutin and Borga, 2003; Kitanidis and Bras, 1980a, 1980b), and assessments of climate change adaptation strategies. In the semiarid southwestern US, the highly nonlinear nature of the rainfall-runoff relationship (Yatheendradas et al., 2008; Goodrich et al., 2000; Pilgrim et al., 1988) makes understanding the sources of modeling uncertainty especially difficult. In this region, precipitation is extremely localized, and summertime convective thunderstorms are exceptionally intense (Roeske et al., 1989, Keefer et al., 2015). The resulting flash floods in the region are considered to be the deadliest natural disaster in the US, accounting for more than 80% of all flood-related deaths; the average economic losses are about one billion U.S. dollars per year (American Meteorological Society, 1985). As a result of anthropogenic global warming the southwest is also expected to experience increased drought frequency (Seager et al., 2007; IPCC, 2013; Cook et al., 2014; Zhao and Dai, 2015; Feng et al., 2014). Short-term droughts can reduce agricultural crop yields, while longer-term droughts can lead to increased wildfire risks with subsequent mudslides and flooding.

The advanced, high-resolution, spatially distributed KINEROS2 (Kinematic Runoff and Erosion) modeling system (Smith et al., 1995; Semmens et al., 2005) is well suited to arid and semi-arid watersheds (Goodrich et al. 2012), providing a physically-based representation of the highly variable response typical to semiarid rainfall forcing and the consequent infiltration and runoff processes experienced in the southwest. Herein, we assess the relative importance of controls exerted by all the KINEROS2 parameters on hydrologic outputs. As advanced and realistic its representation of the hydrology of the southwest arid environment is, KINEROS2 (K2) results can only be as good as the quality of input and values of the parameter estimates. To explore how realistic the model results are, we must understand the relative influence of uncertainties and model output sensitivities from sources such as forcing inputs, initial conditions and model parameters.
The process by which we can understand the relative influence of parameters on the dynamics of hydrologic model behavior is termed sensitivity analysis (SA). In addition to helping elucidate the impact of different model structures, SA also helps the modeler to prepare for model parameterization, and to direct research priorities, by establishing which parameters contribute the most to uncertainty in the model response (Razavi and Gupta, 2015; Wei et al., 2007; Saltelli and Campolongo, 2000; Breshears et al., 1992). As discussed in Razavi and Gupta (2015) and Razavi and Gupta (2016a,b), several SA approaches such as the Morris method (Morris 1991; Campolongo et al. 2007), Sobol method (Sobol, 1990; Saltelli et al., 2008), PAWN method (Pianosi and Wagener 2015), DELSA method (Rakovec et al., 2014), and Moment-based method (Dell’Oca et al 2016), etc., can provide identical sensitivity rankings for situations having very different parameter sensitivity properties. In view of the limitations of many current SA methodologies, there is a need for robust, informative and computationally efficient sensitivity analysis (SA) techniques for developing, understanding, and applying the details of hydrologic modeling systems (Gupta and Razavi, 2018).

In this work, we implemented the Variogram Analysis of Response Surfaces (VARS) methodology (Razavi and Gupta, 2016 a, b) to assess the importance of KINEROS2 parameters, and to investigate the dynamics of parameter importance with varying rainfall depth/intensity on the Lucky Hills watershed at the Walnut Gulch Experimental Watershed, Arizona. Unlike other SA analysis tools, VARS accounts for the spatial correlation in model response as parameters are varied, is more efficient and cost effective, and has been reported to provide more reliable (stable) estimates in the face of sampling variability (Gupta and Razavi, 2018). We compare the results of the VARS method to that of the Sobol (Sobol’ 1990; Saltelli et al., 2008) and Morris (Morris 1991; Campolongo et al. 2007) methods, which are included in the VARS computations as by-products.

The framework followed here is known as Global SA (GSA). We investigate the strength of influence of K2 parameters on different types of model output attributes or response surfaces. We also test how the strength of influence of the parameters change based on the variability of rainfall intensities. The model output attributes studied are the predictive utilities of the K2 model applied to an event-scale simulated high-resolution runoff and sediment rate. We identified a total of 12 predictive utilities (response surfaces) from the K2 predicted runoff and sediment yield outputs. Most of these responses are categorized as performance-metric-free (Gupta and Razavi, 2018), such as flow rate, peak flow rate, flow duration, time-to-peak, time-to-start runoff, recession time, average sediment rate showing mass of sediment per given unit time, peak sediment rate, and total sediment mass. We also explore the influence of the parameters on selected performance-metric-based responses that represent goodness-of-model-fit to the observations; this includes the Nash-Sutcliffe efficiency coefficient (NSE, Nash and Sutcliffe, 1970), Bias (Observation–model prediction) and Kling-Gupta efficiency (KGE, Gupta et al., 2009).

The objectives of the paper are (i) to evaluate the strengths and weaknesses of the metric-free GSA methodology in regard to identifying parameter importance in the context of time-variant and total-period time-aggregate model responses, and (ii) to evaluate the impacts of rainfall intensity on the importance of parameters in modeling rainfall-runoff systems. The latter objective will aid understanding of the effects of parameter perturbation and their importance across simulation periods using K2 model outputs in southwestern USA.
Data and Methods

Using changing rainfall intensities to force the model, we investigated the impact of parameter perturbations on several targeted responses of the event-based, physically distributed, hydrologic model KINEROS2.

Data

Data from six, small to large rainfall depth rainfall-runoff events (Figure 1) from the Lucky Hills watershed were selected for the analysis. Event data are available in the DAP (Data Access Project) database as breakpoint formatted rainfall hyetographs and runoff hydrographs that include time and accumulated depth at slope breaks (Goodrich et al., 2008). DAP is a database for the Walnut Gulch Experimental Watershed (WGEW) that contains high-quality long-term hydro-climatic observations (https://www.tucson.ars.ag.gov/dap/, Nichols and Anson, 2008). Lucky Hills is a very small (4.8 hectare), specialized, experimental watershed within WGEW.

![Figure 1](image)

**Figure 1.** Six selected rainfall and corresponding runoff events in Lucky Hills, a) hyetograph in mm/hr (duration of the rainfall is given in the x-axis), b) event rainfall depth and 15-minute maximum intensity of rainfall for the selected events, c) the corresponding total runoff, and d) the corresponding hydrograph. The rainfall and runoff are sorted based on increasing depth of rainfall from left to right.
Modeling Using K2

K2 (Semmens et al., 2005; www.tucson.ars.ag.gov/agwa/kineros) is an event-based, physically distributed model developed to simulate runoff and sediment responses in southwestern arid environments. K2 simulates entire hydrographs and sedigraphs for a single rainfall event (Goodrich et al. 2012) at sub-minute time step resolution. The model represents interception and the dynamics of infiltration and infiltration-excess surface runoff, and flow routing is based on a finite difference solution of the one-dimensional kinematic wave equations over a basin conceptualized as a cascade of planes (hillslopes) and channels. The Automated Geospatial Watershed Assessment Tool (AGWA, Miller, et al., 2007; www.tucson.ars.ag.gov/agwa/) was used to support the K2 modeling effort in setting up the model, parametrization, and execution of the simulation. Initial estimates for the K2 parameter values were obtained through AGWA parametrization using input GIS layers such as a DEM, SSURGO soils, and USGS land use data layers. The DEM used in the modeling exercises was LiDAR-derived with a 1m resolution.

Our sensitivity analysis investigated seven of the model parameters reflecting infiltration, runoff, flow routing, and soil characteristics including the Saturated Hydraulic Conductivity (Ksat), Manning’s coefficient (n), coefficient of variability of Ksat (CV), capillary drive coefficient (G), intercepted depth (In), sediment cohesion coefficient (C), and sediment splash coefficient (S). Target responses from flow and sediment outputs of K2 were generated iteratively using a suite of AGWA-K2 multiplier files with values between 0 and 2.

Response Surfaces

Response surfaces (model output attributes) of the target model were created based on randomly generated sets of multiplier values applied to each parameter. For each randomly generated parameter set, we computed the time series of the runoff and sediment rate, from which we extracted the 12 response surfaces used to evaluate the relative importance of the K2 parameters. These include two flow-rate related, three sediment-yield related, and four flow-time related performance-metric-free responses and three goodness-of-model-fit performance-metric-based responses. Listed together, these responses are: Average flow rate (cfs), Maximum flow rate (cfs), Time-to-peak (min), Time to runoff start-time(min), Recession time (min), Average sediment yield per event runoff (kg/ha), Maximum sediment rate (kg/sec), Total sediment yield for the event (kg/ha), Nash-Sutcliffe Efficiency coefficient, Bias, and Kling-Gupta Efficiency coefficient.

VARS Methodology

The VARS framework (Razavi and Gupta, 2015a,b) enables robust and efficient global sensitivity testing of model responses based on a star-based sampling strategy (known as STAR VARS) (Razavi and Gupta, 2016b) across the full range of parameter space. The global sensitivity of a given model response y with respect to a model property \( \theta \) varying within the feasible space is characterized by its variogram \( \gamma(h) \) and covariogram \( C(h) \) functions:

\[
\gamma(h) = \frac{1}{2} V(y((\theta + h) - y(x))
\]
\[ C(h) = \frac{1}{2} \text{COV}(y((\theta + h) - y(x))) \]

where \( h = \theta^A + \theta^B \) is the length of the vector \( h = \{h_i, \ldots, h_n\} \) between any two points A and B in the factor space. Directional variograms \( \gamma(h) \) and covariograms \( C(h) \) provide sensitivity information across a full range of scales. These variogram based sensitivity metrics are also integrated to provide integrated summary metrics (IVARS: Integrated Variogram Across a Range of Scales) for particular scales \( (H_i) \) of interest.

\[ \Gamma(H_i) = \int_0^{H_i} \gamma(h)dh_i \]

For the details of the theoretical and mathematical formulation of VARS methodology, readers are referred to Razavi and Gupta, 2015; Razavi and Gupta, 2016 a, b; Gupta and Razavi, 2018. Because of the close theoretical connection between VARS directional variograms as \( h_i \to 0 \) and the expected value of the square of the ratios of changes in output to changes in inputs, VARS-SA is closely associated to the Morris-based SA. Similarly, the variogram and covariogram functions are closely related to the total-order effect Sobol variance-based SA as \( h_i \) becomes large. Based on this close relationship between VARS and the Morris and Sobol’ methods, the VARS methodology provides reliable estimates of the Sobol and Morris sensitivity rankings as by-products at no extra computational expense (Gupta and Razavi, 2018).

In this work, we conducted the SA based on both the total-period and time-variant analysis in VARS offline mode using the procedure described in the Razavi (2018) VARS user’s manual. First, sets of the K2 parameter multipliers were randomly generated using Latin-Hypercube sampling within the feasible range of the multipliers in a STAR VARS sampling strategy (Razavi and Gupta, 2016b) to extract sensitivity information across the full extent of the parameter space. Second, we ran K2 for each set of parameters and for each of the selected rainfall events. Third, we ran VARS tool in offline mode to construct and analyze the Generalized Global Sensitivity Matrix (GGSM). First, we extracted the 12 target model responses from the K2 model simulation outputs. Each of the target surfaces were considered as a single output surface. Then we applied a conventional total-period GGSM analysis using each of the single output response surfaces. Here the variogram-based algorithm computes the overall (total time-period) relative parameters importance through Directional variograms \( \gamma(h) \), covariograms \( C(h) \), and IVARS generated for this analysis. Subsequently, we implemented the time-varying GGSM analysis to conduct a time variant sensitivity assessment using the entire K2 simulation outputs for each of the parameter sets. For this time variant sensitivity analysis, we examined two types of simulated time series (sediment and flow rate) for all the selected rainfall events.

\[
\nabla_\theta Y^k(t) = \begin{bmatrix}
\frac{dY^k_t}{d\theta^1_1} & \ldots & \frac{dY^k_t}{d\theta^{N\text{pts}}_1} \\
\vdots & \ddots & \vdots \\
\frac{dY^k_t}{d\theta^1_{N\theta}} & \ldots & \frac{dY^k_t}{d\theta^{N\text{pts}}_{N\theta}}
\end{bmatrix}
\]

The global sensitivity matrix \( \nabla_\theta Y^k(t) \) is a three-dimensional array of the partial derivatives of the system output \( Y^k \) as it varies with time \( (t = 1,\ldots, T) \), with the number of parameters \( (N_{\theta}) \) and the number of points \( (N_{pts}) \) sampled across the feasible space. Accordingly, \( \nabla_\theta^j Y^k(t) \) (see Equation
4) is corresponding two-dimensional sub-matrix at a specific simulation time step \( t \). The index \( k = \{1, 2\} \) indicates the simulation output (e.g., sediment and runoff rate) of interest. In the conventional GGSM \( k = \{1, \ldots, 12\} \) represents the targeted surfaces without the simulation time dimension. For details of the analysis on the dynamic time variant VARS algorithm readers are encouraged to see Gupta and Razavi (2018).

Here we present the analysis and interpretation of both conventional GGSM of the single output response surfaces and the dynamic VARS outputs. The interpretations are based on values of the IVARS \( \Gamma(H) \), and also of the percent sensitivity across the \( h \) distance vectors, where the latter is defined as the value of \( \Gamma(H_i) \) for each parameter divided by the sum of \( \Gamma(H_i) \) for all the parameters multiplied by 100.

### Results and Discussion

#### Total Time-Period GSA

The relative importance of each K2 parameter identified using the VARS methodology was found to be intuitively consistent with the type of response investigated. In Figure 2, we show the relative importance of each of the K2 parameters to each of the responses, along with 90% confidence intervals of the parameter space. Clearly, the degree of importance of each parameter depends on the type of the response investigated. Further, we find a significant difference in the level and kind of parameter importance when evaluated for the different categories of performance-based and performance-metrics free responses. As expected given the dynamical aspect of the model, the time leading up to runoff generation is affected by all the flow parameters, with saturated hydraulic conductivity (Ksat), the capillary coefficient (G), and the variability of conductivity coefficient (CV) playing the most significant roles. However, for the flow time components responses such as duration, time to peak, and recession time, Manning's coefficient (n) appears to be the most dominant parameter. In surfaces that cover a small part of the simulation period such as the time to peak and time to start, almost all K2 flow parameters showed noticeable importance, but the number decreases to 3 (flow and time-related responses) or 4 (sediment related responses) in other response surfaces.

When using the Bias, NSE, and KGE performance-based flow rate metrics as responses, the saturated hydraulic conductivity was found to be most important, with the G value and n also having a significant impact. Using the performance-metric-free flow rate related responses, the Manning's coefficient is the most important parameter, although its impact reduces when the peak flow rate is used as the response. When examining the performance-metric-free sediment yield related responses, the hydrologic conductivity, soil surface roughness, and cohesion coefficient property related parameters play a significant role.
Figure 2. Bootstrap-based uncertainties (90% confidence interval) of the integrated variogram sensitivity metrics (y-axis) for 11 model outputs over the range of STAR-VARS (x-axis) for seven KINEROS2 parameter multipliers. Interrelated response surfaces grouped on columns of flow volume, time aspect of flow rates, optimization measures of flow rates compared to observed flow, and the sediment yield components from left to right, respectively. The inset in each panel shows the details of each the sensitivity of each of the parameters. The stacked bar chart shows the extent of sensitivity of the parameter multipliers to each response.
Figure 2 also shows that the ranges of the 90% confidence (uncertainty) intervals vary from one response to another. With the full VARS trial of 100 samples of the cross section for each factor (with a total of more than 6000 function evaluations) some responses such as Bias and KGE provide evidence for robustness against sampling variability. It is evident that the level of robustness also varies by parameter, even when compared for the same response (see flow duration, maximum sediment rate, etc).

Clearly, parameter importance can depend on the type of metric and the response investigated, as illustrated by Figure 3, where the parameter rankings vary significantly except for the sediment related responses. In the latter, the rankings do not change, but the magnitude of sensitivity shows noticeable change. One sees the same for the two performance-based metrics Bias and KGE. The difference in the percent sensitivity values and the ranking between “duration” and “recession time” were comparable for the dominant parameters with little variation in the two insensitive parameters (CV of Ksat and the intercepted depth). This small variation is attributable to the similarity between the flow duration and the recession time (the flow duration minus the time to peak) which is very small in the southwest arid environment because of the short-lived nature of intense rainfall.

![Figure 3. Comparison of sensitivity of each of the parameters across all the responses (note that the y-axis is on a log scale).](image)

The analysis also reveals significant variability in the sensitivity values of flow rate and sediment responses. The variability is minimal in the responses “time to peak” and “start time”, which share two properties in common: 1) all the parameters play an essential role, 2) the ranges of the responses are very small.

**Comparison of SA Methods**

To compare the broader view of sensitivity, we used the percent of sensitivity values of three sets of sensitivity metrics (Figure 4) found in the VARS framework. The figure showed nine selected response surfaces in three groups. The Sobol and Morris performance metrics are compared in
Figure 4 because of their similarity to the VARS method and included computation within the VARS structure (Gupta and Razavi 2018).

Figure 4. Comparison of the IVARS sensitivity to the variance-based Total-Order effect (TO-Sobol) and the derivative based Morris sensitivity metrics for K2 parameters across the different response surfaces. The vertical axis shows the proportion of each of the metrics as ratio of the sensitivity values obtained by dividing each of the sensitivity values by the sum of the sensitivity values of all the parameters. The right side plot shows a zoom-in of the left side plot. The comparison includes a) flow related, b) time related, and c) sediment related responses.
In this case study, the variance-based total-order (Sobol-TO) and derivative based Morris sensitivity metrics showed equal ranking to the resulting VARS metrics. Even though the ranking was similar across the different methods, the extent of sensitivities was different. The results show that both Ksat and n played a significant impact in flow rate related surfaces. In the time-related response surfaces, the parameter n showed dominance compared to other parameters. However, in the sediment response surfaces, n was the dominant parameter with the second dominant parameter being the soil cohesion coefficient (C).

**Time-variant GSA**

The extent of sensitivity of the parameters were smaller in the Morris method for parameters that showed essential impacts on the response surfaces such as the Ksat, n, G and C. However, among the insensitive parameters (see Figure 4, right side) the Morris metric showed significant difference compared to both IVARS and Sobol.

![Figure 5](image.png)

**Figure 5.** Time-variant sensitivity of K2 parameters for the duration flow of the event 5, a) rainfall and runoff hydrograph of event ID #5, b) the IVARS values of parameters related to flow, and c) the IVARS values of all K2 parameters. Note that the y-axis in b and c are in log scale. The pie charts on b and c represent the proportion of the sensitivity metrics at four sections of the hydrograph. These are: on the rising limb (at the time when the rainfall hit maximum intensity), peak runoff rate, time at the median of the flow rate on the recession side, and the time of the smallest flow.
Using the event #5 of July 19, 2008 (Figure 5) we showed the hyetograph and hydrograph (Figure 5a) and the corresponding high-resolution sensitivity indices of the parameters on the simulated runoff (Figure 5b) and sediment yield (Figure 5c) from the Lucky Hills watershed. In both b and c, the time-varying changes in parameter importance illustrate the integrated effects of the algorithms of K2 modeling structure and the entire system drivers (climate, surface topography, soil, and land cover properties). The GSA results were evaluated at high temporal resolution (1-minute time step) of the K2 outputs. The importance of all K2 parameters became stronger with the increasing rainfall intensity with time and decreased as the rainfall subsides. The significance of the parameter trajectories continues even after the rainfall stopped over the runoff period. The four pie charts in the subplots b and c showed the extent of the importance in percent. The relative importance of the Ksat (red) seems to increase over the event flow period on flow rate while the relative importance of the Manning’s n (green) decreases then increases over time, and the capillary drive coefficient (black) decreases over time. In the simulated sediment yield (Figure 5c), most trajectories of parameter importance showed variation across the simulation period, though Ksat (red) exhibited the same increasing trajectory as for flow rate. In the time variant analysis of parameter importance, the event-based simulation capability of K2 provides a unique opportunity to study the effects of parameters and associated model forcing on the response surfaces.

**Effects of Rainfall Intensity on Parameter Importance**

The level of parameter importance varies according to the intensity and depth of rain received in the watershed (Figure 6). The selection of the three parameters (Ksat, n, and G) and response surfaces (average runoff rate, KGE, and time to peak) was made for illustrative purposes. The relative importance of Ksat is consistently lower for high-intensity rain regardless of the type of responses, and increases with decreasing rainfall intensity. The influence of Manning’s coefficient and the capillary drive coefficient increases with increasing rainfall intensities for average runoff rate and time to peak based surfaces while its influence reverses for the performance-based KGE surface. We, therefore, conclude that rainfall intensities influence parameter importance differently on different model output attributes.
Figure 6. Comparison of the influence of rainfall intensity variation on parameter importance. Sensitivity values of selected K2 parameters (Ksat, n, and G) for a different range of event rainfall intensities. The rows indicate three selected response surfaces: a) average runoff rate, b) KGE, and c) time to peak.

Conclusions

The study revealed that the extent and rankings of parameter importance varies depending on the type of model output attributes (response surfaces) and the rainfall intensities. In flow rate and sediment related response surfaces, the VARS sensitivity metrics showed considerable variability among K2 parameters with Ksat, n, and G parameters showing the dominant impact. On the contrary, the time-based response surfaces, especially the time-to-peak and start time, showed minimum variation with all the K2 parameters showing noticeable importance. This shows, in model calibration and model parameter identification exercises, it is important to focus on the most sensitive parameters depending on the type of model attributes of interest.
In general, the coefficient of variation of the conductivity (CV), the interception (In), and Splash (S) parameters are the least sensitive parameters in K2 with an insignificant level of importance. Comparison of the variance-based total-order (Sobol-TO) and derivative based Morris sensitivity metrics showed equal ranking to the resulting VARS metrics. However, there were differences in the extent of the relative importance of the parameters given by the magnitude of VARS indices.

The time-variant global sensitivity analysis (GSA) showed the relative parameter importance and ranking varies across the flow period. The variation in the relative importance is affected by the rainfall input and continues to the end of the model period. The relative importance among the parameters also varies with the rainfall event that resulted in the runoff.
References


