



From maps to movies: high-resolution time-varying sensitivity analysis for spatially distributed watershed models

J. D. Herman¹, J. B. Kollat², P. M. Reed¹, and T. Wagener³

¹Department of Civil and Environmental Engineering, Cornell University, Ithaca, New York, USA

²Department of Civil and Environmental Engineering, Pennsylvania State University, University Park, Pennsylvania, USA

³Department of Civil Engineering, University of Bristol, Queen's Building, Bristol, UK

Correspondence to: J. D. Herman (jdh366@cornell.edu)

Received: 5 August 2013 – Published in Hydrol. Earth Syst. Sci. Discuss.: 19 August 2013

Revised: 29 October 2013 – Accepted: 13 November 2013 – Published: 17 December 2013

Abstract. Distributed watershed models are now widely used in practice to simulate runoff responses at high spatial and temporal resolutions. Counter to this purpose, diagnostic analyses of distributed models currently aggregate performance measures in space and/or time and are thus disconnected from the models' operational and scientific goals. To address this disconnect, this study contributes a novel approach for computing and visualizing time-varying global sensitivity indices for spatially distributed model parameters. The high-resolution model diagnostics employ the method of Morris to identify evolving patterns in dominant model processes at sub-daily timescales over a six-month period. The method is demonstrated on the United States National Weather Service's Hydrology Laboratory Research Distributed Hydrologic Model (HL-RDHM) in the Blue River watershed, Oklahoma, USA. Three hydrologic events are selected from within the six-month period to investigate the patterns in spatiotemporal sensitivities that emerge as a function of forcing patterns as well as wet-to-dry transitions. Events with similar magnitudes and durations exhibit significantly different performance controls in space and time, indicating that the diagnostic inferences drawn from representative events will be heavily biased by the a priori selection of those events. By contrast, this study demonstrates high-resolution time-varying sensitivity analysis, requiring no assumptions regarding representative events and allowing modelers to identify transitions between sets of dominant parameters or processes a posteriori. The proposed approach details the dynamics of parameter sensitivity in nearly continuous time, providing critical diagnostic insights into the underlying model processes driving predictions. Furthermore,

the approach offers the potential to identify transition points between dominant parameters and processes in the absence of observations, such as under nonstationarity.

1 Introduction

Distributed rainfall–runoff models allow model parameters and forcing data to vary on a spatial grid, aiming to better represent the spatial variability of watershed processes at the cost of increasing model complexity. This added complexity poses several key challenges, most notably: (1) the difficulty of identifying appropriate parameter sets in a highly interactive, nonlinear, multimodal objective space (Gupta et al., 1998; Carpenter et al., 2001), and (2) the related difficulty of tracing the causes of desirable or undesirable model performance (i.e., diagnosing model behavior) (van Griensven et al., 2006; Gupta et al., 2008; Reusser et al., 2009). Considering the widespread operational use of distributed watershed models, there remains a need for diagnostic methods capable of studying such models at their full spatial and temporal complexity by avoiding aggregation in either dimension to the extent permitted by computational constraints.

Sensitivity analysis is a foundational diagnostic approach in the hydrologic modeling literature, (e.g., Hornberger and Spear, 1981; Franchini et al., 1996; Freer et al., 1996; Wagener et al., 2001; Muleta and Nicklow, 2005; Sieber and Uhlenbrook, 2005; Bastidas et al., 2006; Demaria et al., 2007; Cloke et al., 2008; Van Werkhoven et al., 2008a, 2009; Wagener et al., 2009a; Reusser et al., 2011; Reusser and Zehe, 2011; Herman et al., 2013a, b). However, very few studies

have performed global sensitivity analysis for spatially distributed watershed models due to the computational demands posed by the high dimension of their parameter spaces. Sensitivity analyses of distributed hydrologic and land surface models have frequently addressed this problem by aggregating parameter values across the model grid or subgrids (e.g., Carpenter et al., 2001; Hall et al., 2005; Sieber and Uhlenbrook, 2005; Zaehle et al., 2005; Alton et al., 2006; Cuo et al., 2011; Guse et al., 2013). Few studies have performed global sensitivity analysis on a full set of spatially distributed parameters. The studies that do exist have been limited to event-scale analyses, which reported highly complex spatial sensitivities arising from the interplay between forcing heterogeneity, proximity to observations, and the timescale of model performance metrics explored (e.g., Muleta and Nicklow, 2005; van Griensven et al., 2006; Tang et al., 2007; Van Werkhoven et al., 2008b; Yatheendradas et al., 2008). Although these studies suggest the potential for time-varying spatial-sensitivity analyses, computational demands limited their exploration of this issue.

More recent studies have explored time-varying sensitivities at predefined intervals throughout the model simulation, revealing the dynamics of model controls under changing conditions (e.g., Wagener et al., 2003; Van Werkhoven et al., 2008a; Reusser and Zehe, 2011; Reusser et al., 2011; Herman et al., 2013b). This approach has largely been limited to lumped models, with the exception of Reusser et al. (2011) which analyzed a spatially explicit model. The studies that have focused on event-scale spatial sensitivities (Tang et al., 2007; Van Werkhoven et al., 2008b; Wagener et al., 2009b; Herman et al., 2013a) have proposed using observations to identify representative events for a watershed, a valid concept as long as such representative events exist. However, if the dynamics of a watershed cannot be accurately restricted to one of several event classifications, the *a priori* selection of representative events introduces diagnostic biases that fail to account for the full range of process variability. In this work, we aim to extend the event-scale approach to explore the dynamic controls of a distributed watershed model at a finely resolved sub-daily time step, as well as to advance methods capable of computing and visualizing the results of this analysis.

This study proposes high-resolution time-varying sensitivity analysis for a spatially distributed rainfall–runoff model, avoiding the biases introduced by representative event selection by identifying key transitions between dominant parameters and processes *a posteriori*. These parameters dominate the performance of the model at a particular time, distinct from the true dominant watershed processes independent of our modeling efforts. Our high-resolution global sensitivity analysis employs the method of Morris (1991), which has recently been shown to attain accurate spatially distributed sensitivities at substantially lower computational expense than Sobol' variance decomposition over a temporally aggregated six-month time period (Herman et al., 2013a). The high-

resolution sensitivity analysis is applied to the Hydrology Laboratory Research Distributed Hydrologic Model (HL-RDHM) (Koren et al., 2004; Reed et al., 2004; Smith et al., 2004; Moreda et al., 2006), developed by the United States National Weather Service (NWS). The model test case focuses on the Blue River basin, Oklahoma, USA, over a six-month period using hourly time steps and spatially gridded forcing data. The sensitivity of model performance metrics is computed for the full period, the event scale, and a high-resolution moving window with a 3 h time step to demonstrate the benefit of investigating the full dynamics of spatially distributed model controls. This approach represents a novel, computationally efficient contribution to identify the dynamics of dominant model drivers under changing hydrologic conditions for highly parameterized distributed watershed models.

2 Methods

2.1 HL-RDHM model

The HL-RDHM is a distributed rainfall–runoff model with surface-connected grid cells (Koren et al., 2004; Reed et al., 2004; Smith et al., 2004; Moreda et al., 2006). The model is structured using a 4 km × 4 km grid resolution derived from the Hydrologic Rainfall Analysis Project (HRAP), corresponding to the NEXRAD precipitation products developed by the US NWS. The water balance in each grid cell is modeled with the Sacramento soil-moisture accounting (SAC-SMA) model (Burnash and Singh, 1995), which is widely used in practice by the NWS river forecast centers (Smith et al., 2004; Reed et al., 2004; Moreda et al., 2006). Figure 1c shows the water balance components of the SAC-SMA model in each grid cell, including impervious area parameters (PCTIM and ADIMP), the upper and lower storage zones (UZ– and LZ–), and the percolation functions connecting the upper and lower zones (Z_{Perc} , R_{Exp} , and P_{Free}). Routing between grid cells is modeled with a kinematic wave approximation to the St. Venant equations. This study performs sensitivity analysis on the 14 parameters defined in Table 1 for the SAC-SMA model. These 14 parameters are allowed to vary independently within each of the 78 cells of the HRAP grid shown in Fig. 1c, yielding 1092 parameters in total for the diagnostic analysis.

Herman et al. (2013b) showed that time-varying parameter sensitivity can be linked to the underlying mechanisms of a model. Here, studying the formulation of the SAC-SMA model allows the development of hypotheses regarding the expected parameter sensitivities, and how these might change in space and time. At each time step, evaporation first occurs from the additional impervious stores, both upper zone stores, and the lower zone tension store. In all cases, evaporation is proportional to the saturation level of the storage element. Next, direct runoff occurs from the impervious area,

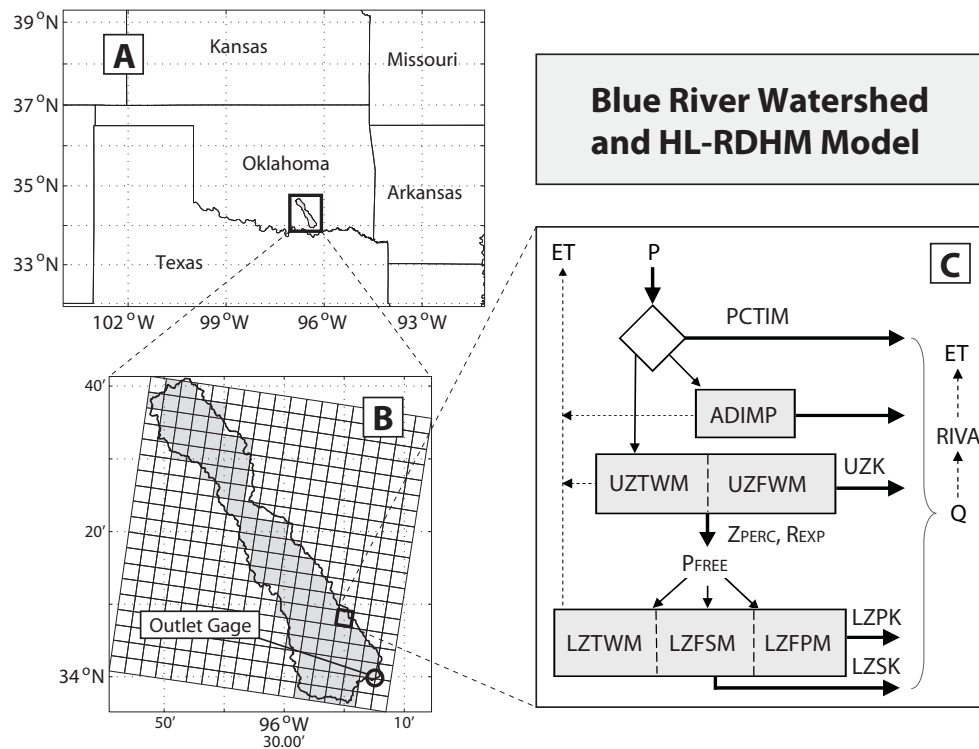


Fig. 1. (A) Location of the Blue River basin in southern Oklahoma, USA. (B) The 78 HRAP grid cells of the Blue River basin (shaded). (C) The Sacramento soil-moisture accounting (SAC-SMA) model, which simulates the water balance in each grid cell.

specified by PCTIM, and the additional impervious area due to saturation, specified by ADIMP. Precipitation not assigned to direct runoff enters the upper zone free water store. Gravity drainage occurs from the upper and lower zones according to the rate constants UZK, LZPK, and LZSK, and is linearly proportional to the amount of water in each respective store. Finally, runoff is also generated when the storage capacity of the upper zone (UZFWM) is exceeded. The same process occurs when all of the lower zone storage capacities are exceeded (LZTWM, LZFSM, and LZFPM), but otherwise excess from any of the lower zones will spill into another.

After the runoff generation mechanisms have occurred, each time step of the model concludes with a redistribution of water between stores according to their saturation levels. First, any deficiencies in the upper and lower tension stores are filled by the free water in their respective zones. Next, percolation occurs from the upper zone free water store to the lower zone based on the saturation level of the lower zone. It is important to note that the lower zone controls percolation in the SAC-SMA model, unlike many other water balance models where percolation is equivalent to spillover from the upper zone. The amount of percolation varies with the parameters Z_{Perc} , the maximum percolation rate under dry conditions, and R_{Exp} , the unitless exponent of the percolation equation (Koren et al., 2004). Finally, the parameter

P_{Free} determines the fraction of percolation that enters the primary and secondary free water stores in the lower zone.

From this description of model mechanisms, we can hypothesize which parameters might be most sensitive in space and time. During and immediately after precipitation events, the parameters associated with quick responses should be most sensitive. This includes the impervious area parameters and the upper and lower zone storage maxima, which can cause direct runoff via overflow. We might expect these sensitive parameters to be spatially concentrated near the outlet of the watershed, since only this area will have sufficient time to contribute to streamflow while the event is occurring. Between precipitation events, the primary streamflow generation mechanism will be drainage from the storage zones, controlled by the rate constants UZK, LZPK, and LZSK; we would expect these to be most sensitive in the time following an event, and with a broader spatial distribution to reflect their slower response. As found in prior work (Herman et al., 2013b), the percolation parameters are unlikely to be highly sensitive at any time, for two reasons. First, the amount of percolation is controlled by the moisture deficiency in the lower zone, so the parameter LZTWM (for example) has more influence on the magnitude of percolation than do the percolation parameters themselves. Second, the percolation parameters do not contribute directly to streamflow, so their signature may be obscured by intermediate processes.

In general, we expect the lower zone parameters to exhibit higher sensitivity over the course of the simulation than upper zone parameters, because the lower zone deficiencies are filled first during the redistribution routine. It is important to note that the spatiotemporal parameter sensitivities will depend on the metric chosen. For example, the sensitivity of the root mean squared error metric on a short timescale will emphasize transitions between quick-response processes, while a water balance error metric on a longer timescale will capture the integrated effects of interacting states and fluxes.

2.2 Study area: Blue River, Oklahoma

This study focuses on the Blue River basin in southern Oklahoma, USA, building on its inclusion in the Distributed Model Inter-comparison Project Phase 2 (DMIP2) (Smith et al., 2012). Figure 1a shows the location of the Blue River in Oklahoma. The watershed is represented by 78 HRAP grid cells, as shown in Fig. 1b, resulting in a total basin area of 1248 km². The model was forced using hourly NEXRAD precipitation data over the six-month period from 16 November 2000 to 15 May 2001, preceded by a 3 week warmup period. Figure 2 shows the hourly precipitation and stream-flow data for the Blue River during the selected simulation period. The selected period reflects a significant wet-to-dry transition during the six-month period, increasing the efficacy of the model warmup. The time period selection was also influenced by the availability of hourly NEXRAD data. The vertical axis of Fig. 2 contains the 78 HRAP grid cells of the watershed, arranged according to distance from the outlet cell. As Fig. 2 indicates, the Blue River basin experiences a series of large rainfall events early in the period before entering a long dry period in the late spring.

We begin by computing parameter sensitivity over the full simulation period. Then, in order to explore the potential consequences of event-scale diagnostics, we select *a priori* three sub-periods to represent watershed dynamics. These are highlighted in Fig. 2 for further analysis (1) a large rainfall event with the highest intensity precipitation focused in the headwaters; (2) a large rainfall event with similar cumulative precipitation but uniform intensity throughout the basin, and (3) a prolonged dry period with low flow. Figure 3 shows the spatial distribution of forcing for each of the three selected sub-periods. We utilize these three sub-periods to explore the relationship between parameter sensitivities over the full period and those derived for shorter events. We then advance this comparison by computing spatially distributed parameter sensitivities at a high-resolution moving window with a 3 h time step. In summary, the experiment consists of sensitivity analysis at three temporal resolutions: the full six-month period, three representative sub-periods, and the high-resolution moving window. We seek to understand the similarities and differences in dominant model behavior at each of these temporal resolutions. In the absence of process-level watershed data, our diagnostic analysis focuses on the transi-

tions between dominant modeled processes under changing hydrologic conditions.

2.3 Method of Morris

The method of Morris (1991) provides measures of global sensitivity from a set of local derivatives, or elementary effects, sampled on a grid defined to cover the parameter space. It is based on one-at-a-time (OAT) methods, which perturb each parameter x_i along a grid of size Δ_i to create a trajectory through the parameter space. For a model with p parameters, one trajectory will contain a sequence of p perturbations. Each of N trajectories yields one estimate of the elementary effect for each parameter (i.e., the ratio of the change in model output to the change in that parameter). Equation (1) shows the calculation of a single elementary effect for the i th parameter.

$$EE_i = \frac{f(x_1, \dots, x_i + \Delta_i, \dots, x_p) - f(\mathbf{x})}{\Delta_i}, \quad (1)$$

where $f(\mathbf{x})$ represents the function evaluation at the prior point in the trajectory. In alternative formulations, the elementary effect is normalized by multiplying by a factor of $x_i/f(\mathbf{x})$ (van Griensven et al., 2006). Using the single trajectory shown in Eq. (1), one can calculate a single elementary effect for each parameter using only $p+1$ model evaluations. However, by employing only a single trajectory, this OAT method depends strongly on the location of the initial point \mathbf{x} in the parameter space and does not account for interactions between parameters. For this reason, the method of Morris (1991) performs the OAT method over a set of N trajectories through the parameter space. This study uses the original factorial sampling approach proposed by Morris (1991), in which trajectories are generated by perturbing one factor at a time. Alternative sampling methods proposed by Campolongo et al. (2007, 2011) and Ruano et al. (2012) aim to maximize coverage of the parameter space and thus the accuracy of global sensitivity estimates. The factorial sampling approach of Morris (1991) has been successfully benchmarked against a total-order Sobol' variance decomposition (Herman et al., 2013a) and is thus suitable for this application. Once trajectories are sampled, the resulting set of elementary effects is averaged to give μ , an estimate of the global effect of each parameter. This study uses the improvement of Campolongo et al. (2007) in which an estimate of total-order sensitivity of the i th parameter, μ_i^* , is computed from the mean of the absolute values of the elementary effects over the set of N trajectories as shown in Eq. (2).

$$\mu_i^* = \frac{1}{N} \sum_{j=1}^N |EE_i^j| \quad (2)$$

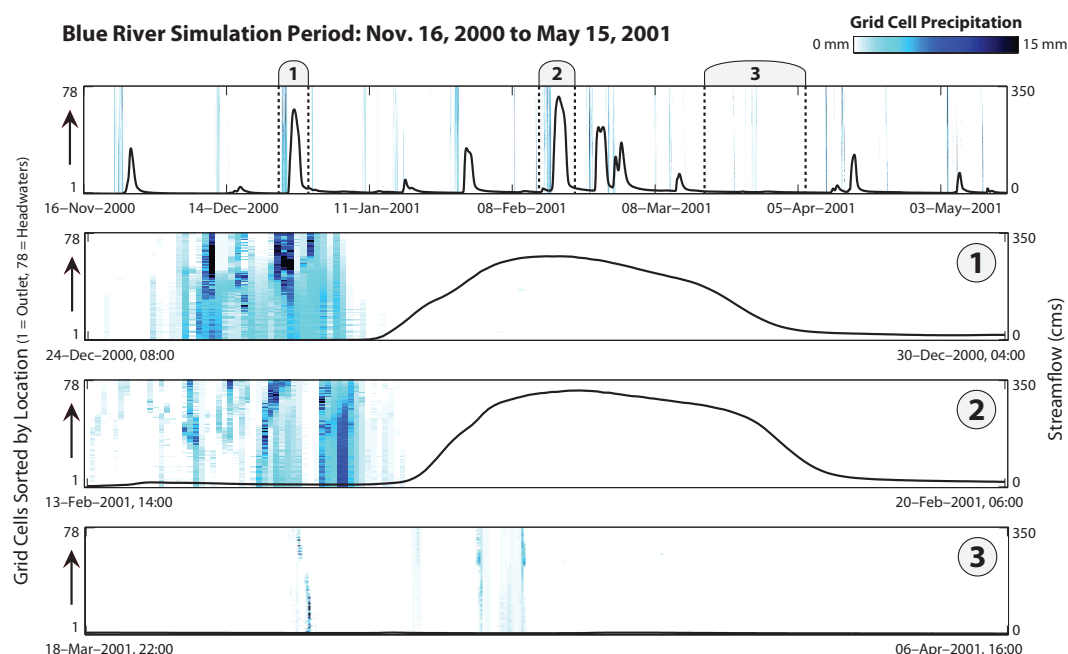


Fig. 2. Hourly streamflow and precipitation values during the six-month simulation period for the Blue River basin. The y axis on the left side of each plot represents abstracted spatial information, where the 78 grid cells of the basin model are sorted from the outlet cell (1) to the cell furthest from the outlet (78). The y axis on the right side of each plot shows the magnitude of streamflow. The color corresponds to the amount of precipitation at each hourly time step. Time periods (1), (2), and (3) are highlighted for further analysis, with (1) and (2) representing large events with different spatial distributions of precipitation, and (3) representing a low-flow period. Note that hours with high precipitation are more visible in time periods (1)–(3) than in the six-month simulation period due to the reduced width of hourly intervals when plotting over the full period.

Table 1. HL-RDHM parameters and their uniform sampling ranges for sensitivity analysis.

Parameter	Description	Units	Lower bound	Upper bound
LZFP	Lower zone primary maximum storage	mm	8.8	58.8
LZFS	Lower zone supplemental maximum storage	mm	19.2	193.2
LZPK	Lower zone primary withdrawal rate	day ⁻¹	0.0408	0.264
LZSK	Lower zone supplemental withdrawal rate	day ⁻¹	0.00168	0.0175
LZTW	Lower zone tension water maximum storage	mm	61.6	249.6
PCTIM	Permanent impervious area	%	0.0	5.0
PFREE	Percolation to lower zone	%	16.0	55.0
REXP	Percolation equation exponent	–	1.69	3.47
UZFW	Upper zone free water maximum storage	mm	8.8	64.8
UZK	Upper zone free water withdrawal rate	day ⁻¹	0.19	0.76
UZTW	Upper zone tension water maximum storage	mm	19.2	78
ZPERC	Maximum percolation rate under dry conditions	mm day ⁻¹	27.2	140.4
ADIMP	Saturated impervious area	%	0.0	20.0
RIVA	Riparian vegetation area	%	0.0	20.0

3 Computational experiment

The method of Morris was performed on the 14 SAC-SMA model parameters in each grid cell of the HL-RDHM model as indicated in Fig. 1. The uniform sampling bounds for each parameter given in Table 1 are based on the a priori gridded parameter values derived by the NWS (Koren et al.,

2004) and extended for the event-scale sensitivity analysis performed by Van Werkhoven et al. (2008b). Parameter values for each grid cell were sampled separately, resulting in a total of $78 \times 14 = 1092$ total sampled parameters. Rather than measure the sensitivity of the output streamflow directly, we measure the sensitivity of model performance metrics, calculated using the known hourly streamflow values over

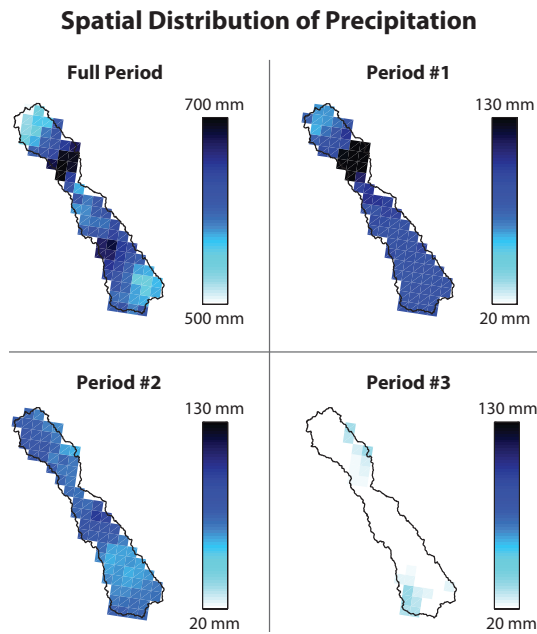


Fig. 3. Maps of precipitation in the Blue River basin over the full six-month simulation period and the three sub-periods defined in Fig. 2. Over the full period, precipitation is roughly even across the watershed. The event during period 1 is focused in the headwaters, while the event during period 2 is more evenly spread. Finally, period 3 represents a dry period with little streamflow.

the six-month simulation period. This ensures that our sensitivity indices properly incorporate measures of model accuracy, an approach strongly supported by recent literature (e.g., van Griensven et al., 2006; Demaria et al., 2007; Cloke et al., 2008; Pappenberger et al., 2008; Van Werkhoven et al., 2008a; Reusser et al., 2011; Rosolem et al., 2012). We compute sensitivity indices at the event scale using the root mean squared error (RMSE) and runoff coefficient error (ROCE) metrics. The RMSE metric represents the sum of squared residuals over a particular time window:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (Q_{s,i} - Q_{o,i})^2}, \quad (3)$$

where Q_s and Q_o are the simulated and observed flows, respectively. The ROCE metric represents the error in the water balance, calculated as a percentage bias:

$$\text{ROCE} = \frac{\left| \sum_{i=1}^n Q_{s,i} - \sum_{i=1}^n Q_{o,i} \right|}{\sum_{i=1}^n Q_{o,i}}. \quad (4)$$

The RMSE metric focuses on quick responses, while the ROCE metric highlights the long-term bias of the water balance calculated by the model (Van Werkhoven et al., 2008a).

These two metrics combine to provide a comprehensive understanding of model response at the event scale. The high-resolution sensitivity analysis is performed using a 24 h moving window with a 3 h time step; only the sensitivity of the RMSE metric is computed here, since a water balance metric would be inappropriate for such a short timescale. With this high-resolution moving window, the sensitivity indices of all 1092 parameters are calculated at a total of 1457 intervals over the course of the six-month simulation period.

We calculate sensitivity indices using a sample size of $N = 20$, corresponding to 21 860 model evaluations. This represents a significant computational savings compared to a typical global sensitivity analysis method. Herman et al. (2013a) showed that the method of Morris using $N = 20$ for the full simulation period of this study was capable of providing sensitivity results comparable to the Sobol' method using $N = 6000$, which required over 6.5 million model evaluations. The high-resolution sensitivity analysis investigated here is only computationally tractable due to the demonstrated efficiency of the method of Morris (Herman et al., 2013a).

The sensitivity analyses were performed using the NSF CyberSTAR high-performance cluster at Penn State University, which contains a combination of quad-core AMD Shanghai processors (2.7 Ghz) and Intel Nehalem processors (2.66 Ghz). An open-source implementation of the method of Morris was used from the R Sensitivity Package (Pujol et al., 2013), which includes the methodological improvement of Campolongo et al. (2007). Approximately 100 computing hours were required for the model evaluations at the $N = 20$ sample size, with an additional 100 h needed to compute the sensitivity indices for each of the nearly 1500 sub-intervals.

4 Results and discussion

Sensitivity results are presented in order of increasing temporal resolution. We begin with the full period and event-scale sensitivity indices (Figs. 4–5) before proceeding to the high-resolution results (Figs. 6–8). These results can be interpreted in the context of the precipitation patterns shown in Figs. 2–3. This sequence of results is designed to explore the potential shortcomings of the aggregated approaches along with the additional insights provided by the high-resolution approach.

4.1 Full period and event-scale sensitivity analysis

The sensitivity indices for the root mean squared error (RMSE) metric are shown in Fig. 4 for the full simulation period and the three selected events. The μ^* values from the method of Morris are normalized to the range $[0, 1]$ to facilitate comparison across experiments, from an initial range of $[0, 0.08]$. For the full six-month period, the spatial distribution of parameter sensitivity appears bimodal: a concentrated

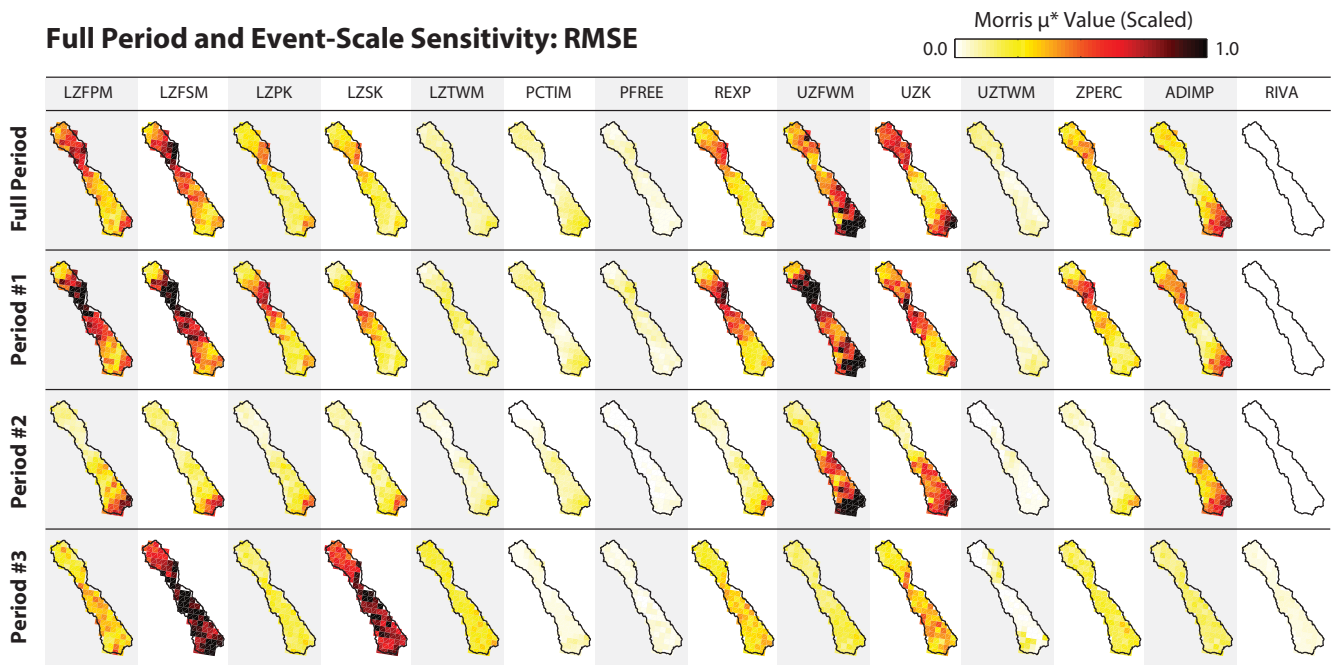


Fig. 4. Full period and event-scale sensitivities of the root mean squared error (RMSE) metric over the full six-month simulation and three selected sub-periods. The μ^* values from the method of Morris are scaled to the range [0, 1] from an initial range of [0, 0.08]. The RMSE metric focuses on the model's ability to reproduce observed streamflow peaks. The event-scale sensitivity indices differ significantly from those in the aggregated full period depending on the magnitude of the event and the spatial distribution of precipitation.

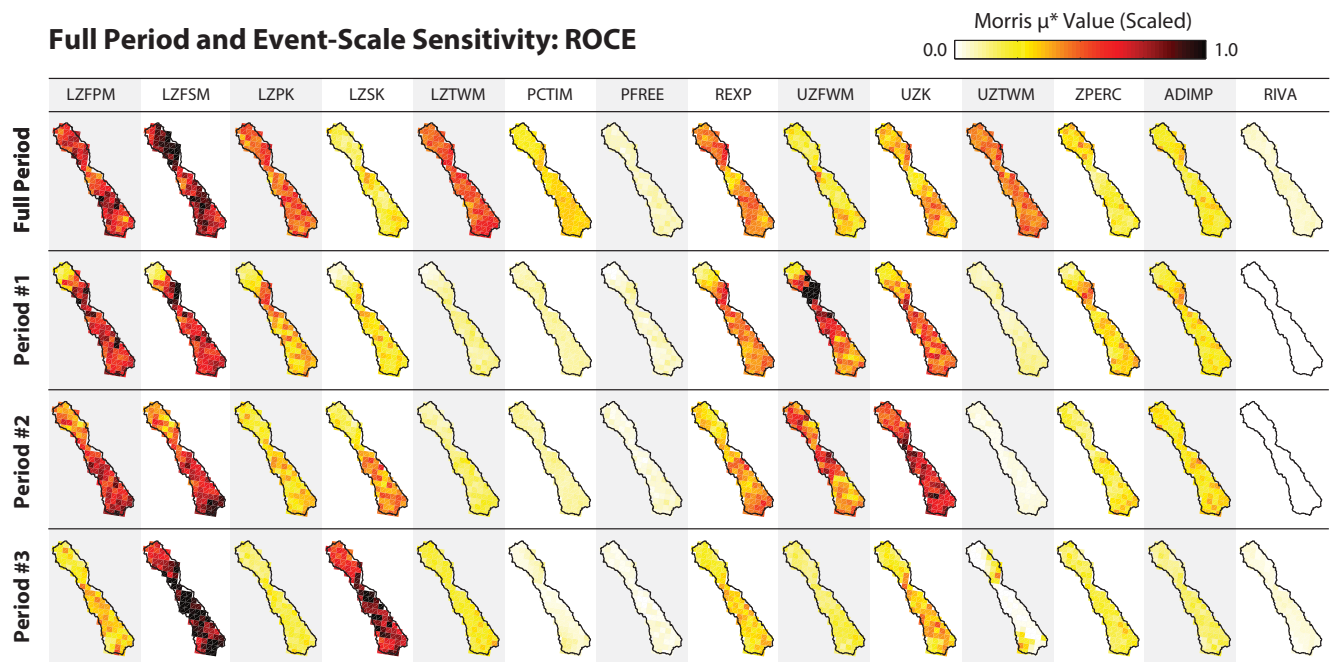


Fig. 5. Full period and event-scale sensitivities of the runoff coefficient error (ROCE) metric over the full six-month simulation and three selected sub-periods. The μ^* values from the method of Morris are scaled to the range [0, 1] from an initial range of [0, 0.08]. The ROCE metric focuses on the model's ability to reproduce the observed long-term water balance. The spatial controls on the water balance are more evenly spread across the watershed compared to the RMSE metric in Fig. 4.

high-sensitivity area occurs in the headwaters, particularly for the lower zone storage maxima LZFPM and LZFSM, and a second concentration occurs near the outlet of the watershed, particularly for the upper zone parameters UZFWM and UZK. Considering the forcing patterns shown in Fig. 3, the RMSE for the full period is likely dominated by several large events, some of which are concentrated in the headwaters of the basin such as the event during period 1. This explains the bimodal division of RMSE sensitivity between the headwaters (due to concentrated forcing during large events) and the outlet cells (due to gauge proximity). Note that both the upper zone storage maximum, UZFWM, and its associated drainage coefficient, UZK, are sensitive during this period, whereas only the storage maxima in the lower zone are sensitive. This difference indicates that flow from the upper zone is generated from a combination of storage overflow and gravity drainage, while flow from the lower zone is primarily generated by storage overflow alone. This result highlights the importance of understanding which flow generation mechanisms dominate model performance during peak events over the course of the simulation.

Period 1 exhibits a strong concentration of parameter sensitivity in the headwater cells of the basin. As illustrated in Fig. 3, the large precipitation event during this period occurs primarily in the headwaters, so this result is expected. Even the upper zone parameters are most sensitive in the headwaters during period 1 (with the exception of UZFWM), despite the typical dependence of upper zone sensitivity on gauge proximity. This result contrasts with period 2, which exhibits very little sensitivity in the headwaters. In period 2, the majority of high-sensitivity cells appear near the watershed outlet, even for the lower zone parameters. The contrast between periods 1 and 2 reveals the effect of the spatial distribution of forcing on parameter sensitivity. The headwater cells of the model are only activated when precipitation is concentrated in this region, and flow is generated by exceeding storage maxima in both the upper and lower zones. Conversely, when precipitation is distributed across the basin, model performance is dominated only by the cells near the outlet gauge where flow is generated in the upper zone by a combination of storage exceedance and gravity drainage. These differences in the responses of periods 1 and 2 are potentially complicated by internal model states, such as antecedent moisture conditions, which could alter the response signatures. However, as shown in the Supplemental, zooming in to the time-varying sensitivity of periods 1 and 2 clearly reveals the differences in the spatial distribution of precipitation between the two events. The sensitivity responses to each event begin almost immediately following the precipitation, and thus their differences may be traced primarily to the precipitation distribution over the watershed. With no information regarding the “true” watershed processes, it is worth noting that a significant portion of the model remains insensitive during both large events. The differences between sensitivity patterns in periods 1 and 2 would be very difficult

to predict in advance, and thus underscore the need for diagnostic methods that do not depend on spatial aggregation.

Finally, the dry period 3 exhibits very different sensitivity patterns from any of the other periods. Here, sensitivity is effectively concentrated in the lower zone secondary storage element, with maximum LZFSM and drainage coefficient LZSK. The lower zone secondary storage is very likely to be the last element containing water during dry periods, as it has the slowest drainage constant as shown in Table 1. Therefore, this element controls model performance after the other storage zones have drained or evaporated. The sensitivities of parameters LZFSM and LZSK are distributed across the entire watershed, unlike during periods 1 and 2 where the most sensitive parameters only occur in concentrated areas. This suggests that dry periods may provide valuable identifiability information for cells which are otherwise inactive, particularly for these slow-draining storage elements in the lower zone of the model.

At the event scale, it is valuable to assess the sensitivity of multiple diagnostic measures to obtain a more thorough understanding of controls on model performance. In general, model error can be decomposed into correlation, bias, and variability (Gupta et al., 2009). The RMSE metric, with its dependence on quick runoff response, is most closely related to correlation and variability; we also investigate the runoff coefficient error (ROCE), a water balance metric related to model bias (Van Werkhoven et al., 2008a). The event-scale sensitivity indices for the ROCE metric are shown in Fig. 5. Compared to the RMSE metric shown in Fig. 4, the sensitivity of ROCE is spread across a larger number of parameters, and more evenly distributed across the spatial extent of the watershed. Whereas RMSE is controlled by a few cells depending on their proximity to precipitation and/or the outlet gauge, the water balance error depends on the soil-moisture calculations in all cells. In general, many of the same lower and upper zone parameters dominate the ROCE and RMSE performance metrics: the storage maxima LZFPM, LZFSM, and UZFWM, and the drainage constants LZPK and UZK. Similar to RMSE, the ROCE metric depends on flow generation via storage exceedance as well as gravity drainage processes in the model. Compared to Fig. 4, the differences between periods 1 and 2 are far less pronounced in Fig. 5, indicating that the spatial distribution of precipitation does not affect the water balance error as much as it does peak flows. Period 3 shows the most similarity to the corresponding RMSE result, as its ROCE metric is still controlled primarily by the secondary storage parameters LZFSM and LZSK. The apparent independence of the ROCE metric to forcing and gauging locations suggests that this measure of performance succeeds in extracting information from a larger spatial area of the model, potentially providing benefits for identifiability. However, the ROCE metric alone will not account for the timing of flow peaks, and is therefore best applied in conjunction with a timing-based metric such as RMSE.

The event-scale sensitivity results shown in Figs. 4 and 5 provide useful diagnostic insight for the full simulation period and selected sub-periods. These findings align with previous work: spatially concentrated precipitation will cause parameter sensitivity to appear in a similar pattern as the precipitation, whereas uniformly distributed precipitation will cause sensitivity in cells near the outlet (Tang et al., 2007; Van Werkhoven et al., 2008b). However, the event-scale analysis also contains several weaknesses. First, the results are highly dependent on the choice of events to study, as illustrated by the differences in controls across the selected periods. It would be prohibitively difficult to design or select representative events which fully capture the range of model responses. Instead, it is beneficial to analyze the emergent model responses in nearly continuous time, and select sub-periods of interest *a posteriori*. Second, the event-scale results do not indicate when these parameters become sensitive relative to changing hydrologic conditions. Consequently, the sensitivity indices shown in Figs. 4 and 5 for the full simulation period are strongly influenced by only a few large events, with their dynamics obscured by aggregation. It has been noted in previous work that the value of streamflow observations for identifying distributed model parameters may be limited by the location and intensity of forcing, particularly if the period of analysis is defined to include only a single rainfall event. (Van Werkhoven et al., 2008b). We hypothesize that allowing distributed parameter sensitivity to vary in nearly continuous time will extract more value from streamflow observations by highlighting parameter activation across a much broader range of hydrologic conditions.

4.2 High-resolution distributed sensitivity

This study aims to elucidate the time-varying nature of these distributed parameter sensitivities by performing global sensitivity analysis using a high-resolution moving window with a sub-daily time step. Parameter sensitivity is calculated on a 24 h moving window with a 3 h time step, resulting in 1457 intervals over the six-month simulation period.

The choice of window size contains two competing considerations: it must be large enough to smooth out any noise in the performance metric calculations, yet also small enough to capture dominant model processes that occur on fast timescales. The moving window size and time step used in this study reflect a balance between these two issues. Since the model runs on an hourly time step, a 24 h window size (with significant overlap due to the 3 h time step) will smooth noise in the calculation of the RMSE metric. Additionally, since most large events during the simulation period are approximately 48–72 h in length, the 24 h window is also sufficiently small to capture quick responses, which is critical from a flood forecasting standpoint.

Thus, each parameter has a time series of sensitivity indices for each grid cell (i.e., the results summarize time-evolving sensitivity maps across all spatial grid cells in the

model). These sensitivity indices are shown in Figs. 6, 7, and 8, corresponding to the lower zone parameters, upper zone parameters, and remaining parameters, respectively. Each of the three plots contains the same hourly precipitation data in the top panel, as well as the same hourly hydrograph data superimposed on each subplot. The sensitivity indices are aligned at the center of each moving window interval. The μ^* values from the method of Morris are normalized to the range [0, 1] to facilitate comparison across experiments, from an initial range of [0, 0.2]. While these figures are designed for journal format, animations of time-varying sensitivity indices are available as a multimedia supplement.

In Figs. 6, 7, and 8, the two spatial dimensions of the watershed are compressed into the y axis, where the 78 grid cells are arranged according to their distance from the watershed outlet. The bottom of each subplot ($y = 1$) represents the outlet cell, while the top of each subplot ($y = 78$) represents the headwater cell furthest from the outlet. This configuration allows us to visualize both space and time on the same axes without drawing a large number of maps of the watershed. This plotting approach is particularly effective for the Blue River basin, which has a long, narrow shape.

Figure 6 shows that the lower zone parameters maintain a moderate level of sensitivity throughout the simulation. The influence of all lower zone parameters clearly recedes during large events, except for the parameters LZFSM and LZFSM in the cells nearest to the watershed outlet. This indicates that the only contribution to large streamflow events from the lower zone occurs due to exceeding the storage maxima, not due to gravity drainage. The drainage processes occur on slower timescales and would not contribute significantly to peak flows. The effect of slow drainage processes is clear from the high sensitivity of the secondary storage parameters, LZFSM and LZSK during low-flow periods, after other storage elements have been emptied. These insights largely align with those found at the event scale in Figs. 4 and 5. However, when comparing sensitivity indices across temporal resolutions, it is important to note that parameter sensitivity is measured relative to other parameters. Thus, a larger time window may cause some parameters to appear insensitive at certain locations due to the dominance of others. This phenomenon is visible, for example, for the LZSK parameter near the outlet of the watershed, where it is sensitive in Fig. 6 but not in Fig. 4. Compared to the event-scale analysis, the high temporal resolution in Fig. 6 has the advantage of clarifying the timing of parameter activation. For example, it has been noted in prior studies that the lower zone parameters frequently control the performance of the SAC-SMA model (Van Werkhoven et al., 2008a; Herman et al., 2013b), from which it might be concluded that the lower zone contributes significantly to flow peaks. However, Fig. 6 indicates that the lower zone contributes to performance primarily during non-peak periods, which, when aggregated, may yield higher levels of sensitivity depending on the period studied.

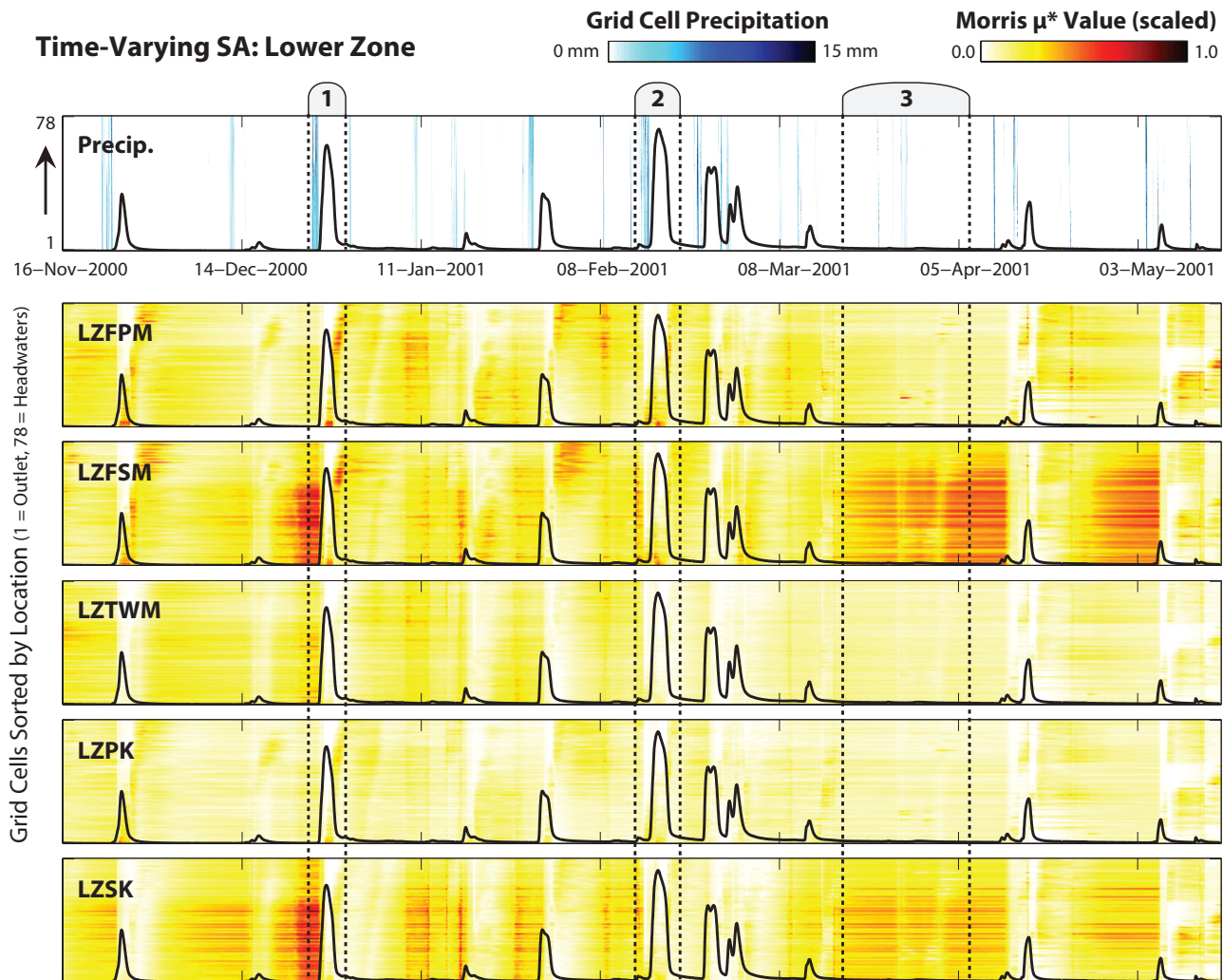


Fig. 6. Time-varying sensitivity of the RMSE metric for the five lower zone parameters of the HL-RDHM model. The indices are calculated for a 24 h moving window with a 3 h time step. The y axis arranges the 78 grid cells based on their distance from the watershed outlet, from the outlet ($y = 1$) to the furthest headwater cell ($y = 78$). The μ^* values from the method of Morris are scaled to the range $[0, 1]$ from an initial range of $[0, 0.2]$. The lower zone parameters maintain a consistent, moderate level of sensitivity throughout the simulation. Exceptions occur during large events, when the lower zone parameters are mostly insensitive. The secondary storage parameters, LZFSM and LZSK, are particularly sensitive during low-flow periods.

By contrast, the upper zone parameters are clearly activated during and after streamflow events, as shown in Fig. 7. In particular, the upper zone free water parameters (UZFWM and drainage coefficient UZK) become dominant during large events. The propagation of sensitivity upward through the watershed is clear for these parameters, starting at the outlet cells during the rising limb of each event and moving toward the headwaters during the falling limb. As expected, there is a lag between the time at which the event begins and the time at which the headwater cells begin to affect the model performance due to routing. Similarly, the timing of activation for the outlet cells depends on the event; compare the parameters UZFWM and UZK during period 1, where the outlet cells are activated midway through the event, to

period 2, where UZK is activated immediately during the rising limb of the hydrograph. The lag in parameter sensitivity during period 1 is likely due to the headwater-focused precipitation event, while in period 2 the precipitation occurs closer to the outlet. Interestingly, Fig. 7 shows that the additional impervious area parameter, ADIMP, is only sensitive for cells near the outlet during events, but this signal does not propagate to the headwaters. The impervious area only affects model performance in cells close to the outlet, since these directly control the quick response during the rising limb of each event. As Fig. 7 indicates, the upper zone parameters typically do not control model performance during low-flow periods and small events. However, there are two interesting exceptions to this. First, the drainage rate UZK

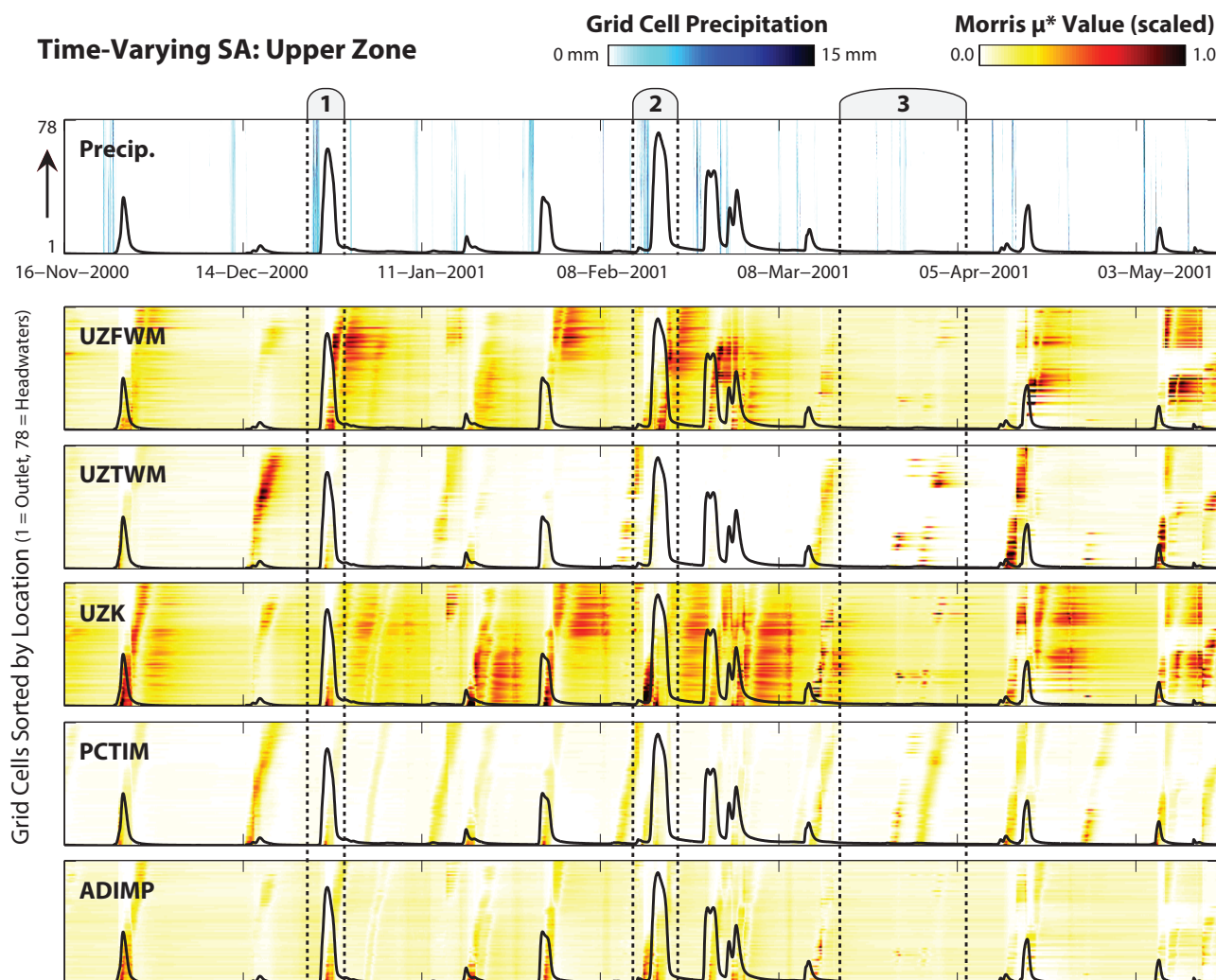


Fig. 7. Time-varying sensitivity of the RMSE metric for the five upper zone parameters of the HL-RDHM model. The indices are calculated for a 24 h moving window with a 3 h time step. The y axis arranges the 78 grid cells based on their distance from the watershed outlet, from the outlet ($y = 1$) to the furthest headwater cell ($y = 78$). The μ^* values from the method of Morris are scaled to the range $[0, 1]$ from an initial range of $[0, 0.2]$. The parameters controlling the upper zone free water element, UZFWM and UZK, are highly sensitive during large events. The high sensitivity of these parameters typically begins near the watershed outlet during the rising limb of the hydrograph, and transitions toward the headwater cells during the falling limb.

remains sensitive for several hours after each event as the upper zone drains its storage, that is, gravity drainage from the upper zone typically occurs slowly enough such that it continues to release water well into the low-flow periods. Second, the parameters UZTWM (upper zone tension storage) and PCTIM (percent impervious area) are most sensitive following rainfall events which do not lead to large streamflow events. In other words, these parameters are most important when model performance requires the absence of a response. If impervious area is too high (causing high direct runoff), or tension storage capacity too low (causing runoff via overflow), the model may overestimate streamflow and create significant errors in the RMSE metric. This phenomenon is not visible for the events analyzed in Fig. 4 because it would

be difficult to predict at the time of a priori event selection. When a large streamflow response is required, the parameters UZFWM and UZK dominate instead, as these become the primary mechanism by which large events are generated. An additional difference between Figs. 4 and 7 is that, while the UZK parameter is sensitive throughout the watershed at the event scale (e.g., during period 1), it only appears sensitive near the outlet at the high-resolution timescale. This could be due to the dominance of other parameters in the upper region of the watershed at the high-resolution timescale, or simply a difference in the sensitivity of the upper zone drainage response at different timescales. Finally, the UZFWM parameter during period 2 is only sensitive near the watershed outlet when considered at the event scale, but the high-resolution

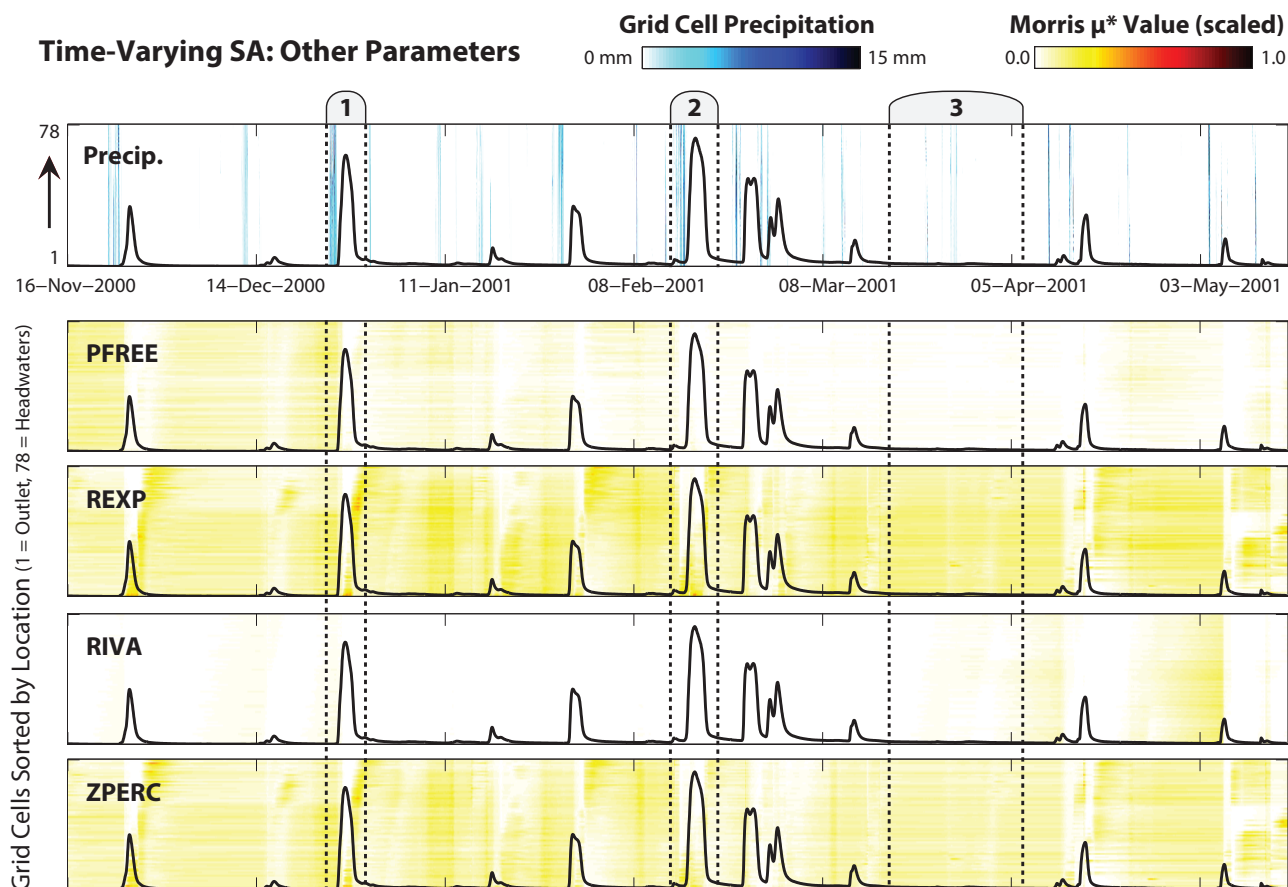


Fig. 8. Time-varying sensitivity of the RMSE metric for the four remaining parameters of the HL-RDHM model. The indices are calculated for a 24 h moving window with a 3 h time step. The y axis arranges the 78 grid cells based on their distance from the watershed outlet, from the outlet ($y = 1$) to the furthest headwater cell ($y = 78$). The μ^* values from the method of Morris are scaled to the range $[0, 1]$ from an initial range of $[0, 0.2]$. These parameters influence model performance significantly less than the lower zone parameters (Fig. 6) or the upper zone parameters (Fig. 7), and never appear highly sensitive during the course of the simulation.

results in Fig. 7 show that the sensitivity of UZFWM propagates to the headwaters during the falling limb of the event. Since the RMSE metric focuses on peak flows, the falling limb does not play a significant role in the calculation of aggregated sensitivity for each period, even though these are clearly visible at the high-resolution timescale.

Finally, Fig. 8 shows the time-varying sensitivity indices for the four remaining parameters in the model. These parameters rarely dominate the upper and lower zone parameters shown in Figs. 6 and 7. The percolation parameters Z_{Perc} and R_{Exp} maintain a low, but non-zero, level of sensitivity throughout the simulation, which decreases to zero during streamflow peaks. The percolation parameter P_{Free} follows a similar trend, but becomes inactive during the dry period 3. The riparian vegetation area, RIVA, is only activated in the late spring season when evapotranspiration becomes more pronounced, as expected. Comparing to the event-scale results in Fig. 4, these parameters are generally much less sensitive at the high-resolution timescale. For example, the R_{Exp}

parameter modifies the rate of percolation from the upper to lower zone but does not cause runoff directly, which makes it less likely to influence model performance during the 24 h moving window than over the course of an aggregated event. The parameters shown in Fig. 8 play a small role in model performance, as evidenced by their moderate but non-zero sensitivities through time. However, the sensitivities of the upper and lower zone parameters are typically much larger (see Figs. 6 and 7) and thus these remaining parameters rarely control model performance. This result suggests a potential identifiability problem for these less-sensitive parameters, as they are rarely activated in any of the model grid cells.

4.3 Discussion

The event-scale sensitivity maps shown in Figs. 4 and 5 represent a traditional approach to diagnostic analysis of a spatially distributed watershed model, on the relatively few occasions that such analyses have been performed. The

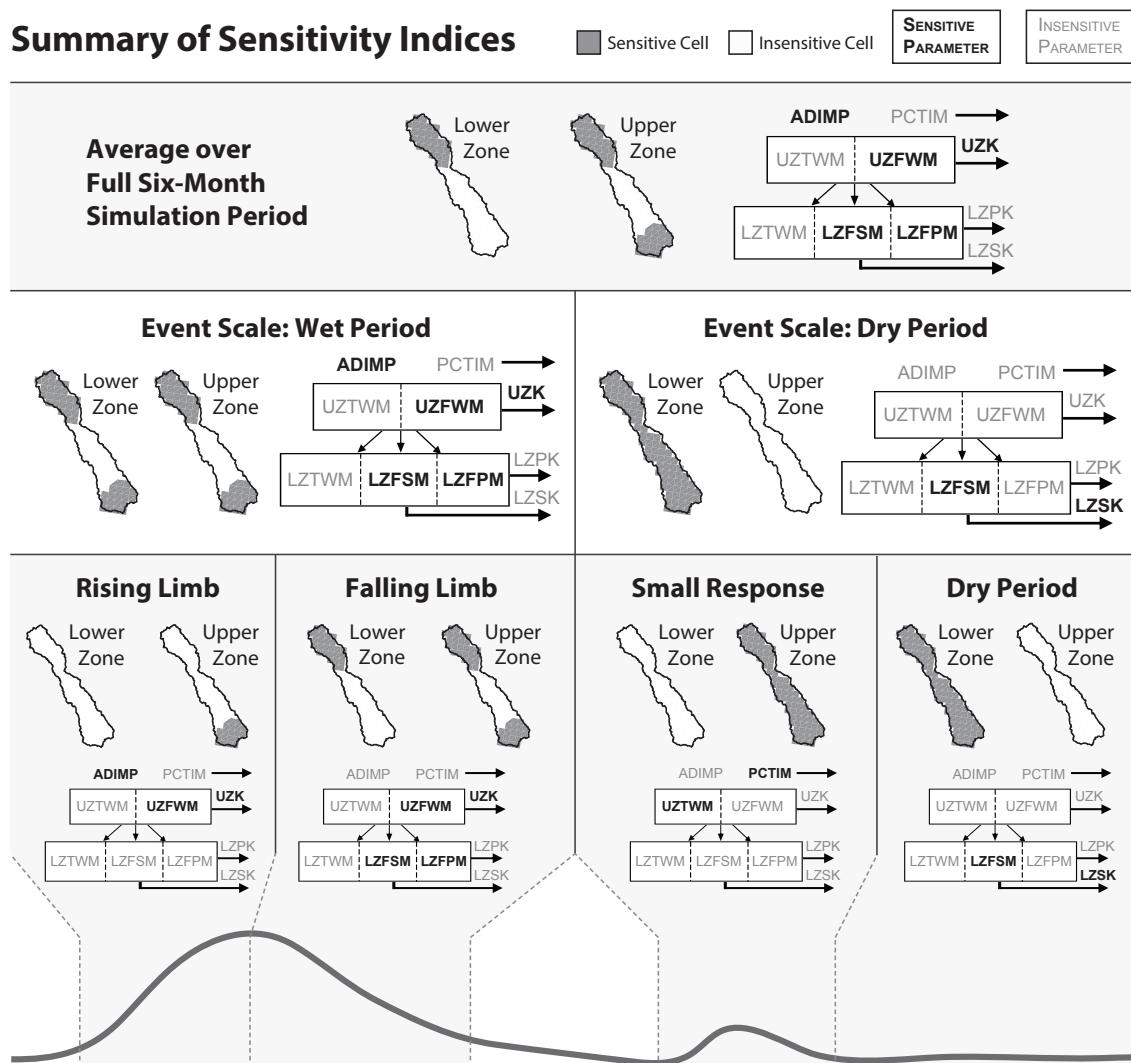


Fig. 9. Qualitative summary of sensitivity indices for the RMSE metric at increasing temporal resolutions, from the aggregated simulation period (top), to the event scale (middle), and the high-resolution moving window (bottom). Each classification includes spatial maps showing where upper zone and lower zone parameters are sensitive in the watershed, along with a model diagram highlighting the most sensitive parameters during the defined periods. Results shown for the high-resolution sensitivity analysis are a qualitative summary of insights gleaned from Figs. 6–8 and thus do not reflect the dependence of time-varying sensitivity on the spatial distribution of precipitation.

event-scale approach shows the spatial distribution of sensitivity for selected intervals, which may change dramatically from one event to another, as the results show for periods 1, 2, and 3. The choice of representative events for a diagnostic analysis can therefore strongly bias the outcome, particularly considering the complex dependencies between parameter sensitivity, the spatial distribution of forcing of an event, and the proximity of a given cell to the outlet gauge. These results suggest the severe difficulty of selecting a priori a set of precipitation events which capture the full range of potential model responses, the approach suggested in prior studies (Tang et al., 2007; Van Werkhoven et al., 2008b). For example, this difficulty is demonstrated by our inability to foresee the phenomenon in which parameters UZTWM

and PCTIM are most sensitive when modeled responses require attenuation to match observations. Furthermore, the temporal aggregation involved in the event-scale analysis obscures the underlying dynamics of parameter sensitivity. When full temporal resolution is allowed, parameter sensitivity indices show clear patterns of activation before, during, and after streamflow events, as shown in Figs. 6, 7, and 8. The high-resolution approach improves on the event-scale analysis by isolating the time and location at which individual parameters and cells are activated, allowing a larger fraction of the model to contribute to its performance measures throughout the simulation and thus making better use of the information content contained in streamflow observations. Within each event, parameter sensitivity clearly represents a

dynamic rather than static quantity, and should be analyzed accordingly as computational costs permit.

A qualitative summary of sensitivity indices at increasing temporal resolution is shown in Fig. 9. At the highest resolution (i.e., the moving window), the spatiotemporal sequence of influential parameters in Fig. 9 reflects the transitions between sets of dominant parameters and processes. In general, the highlighted sensitive parameters and cells are those with a scaled value of $\mu^* > 0.5$. For the high-resolution timescale, the four cases shown (rising limb, falling limb, small response, and dry period) are intended to reflect general insights from Figs. 4–8. These cases do not necessarily reflect all time steps in the simulation period, but rather the most interesting classifications that were observed in the results. However, since the Blue River watershed is driven primarily by infrequent large events without significant temperature or elevation effects, it is possible to broadly separate the hydrograph into these four classifications, as shown in the Supplement.

As Fig. 9 shows, the dominant controls for the full aggregated period are a combination of lower zone parameters in the headwaters of the basin, and upper zone parameters near both the headwaters and outlet. The full period sensitivities are clearly influenced by the wet periods at the event scale, which exhibit the same responses, indicating that the aggregate period is biased toward these large events (a result consistent with the focus of the RMSE metric). By contrast, dry periods at the event-scale exhibit very different sensitivity patterns, centered around slow drainage from the lower zone supplemental store. The summarized high-resolution sensitivity results in the bottom row of Fig. 9 provide a more detailed understanding of model behavior than the full period or the event scale. In general, the parameters that appear most sensitive at the event scale are also the most active for the high-resolution moving window. These primarily include the upper zone parameters UZFWM and UZK and the lower zone parameters LZFP and LZPK. This finding aligns with our initial hypotheses, since gravity drainage and overflow from exceeding storage maxima represent two of the primary runoff generation mechanisms in the model. The most sensitive cells during the rising and falling limbs of large events represent a decomposition of the event-scale sensitivity during wet period, which may be particularly valuable depending on the part of the hydrograph being analyzed. As anticipated, the upper zone and impervious area parameters dominate model performance during and immediately following large events, since these create the quick response required to reproduce observed streamflow. The high-resolution dry period exhibits largely the same sensitivities as the event scale, which would be expected considering the lack of dynamic behavior during these dry periods. Finally, the small response reflects the common scenario in which quick runoff must be avoided to achieve good performance, a behavior which remains invisible at the event scale unless a small response event is explicitly chosen for analysis a priori.

The high-resolution results in the bottom panel of Fig. 9 can also be interpreted to identify transitions between dominant parameters and processes in the model. During the rising limb of streamflow events, the dominant processes in the model are typically direct runoff from impervious area, and overflow/drainage from the upper zone free water store. As might be expected, these processes are most dominant near the outlet of the watershed, reflecting the need for a quick response to match the observed hydrograph. During the falling limb, the model transitions to a dominant process comprising slower drainage responses from the upper and lower zone. These processes are dominant in the headwaters as well in addition to the cells near the outlet, since the longer time lag allows cells further from the outlet to contribute to streamflow. During small responses, the dominant process consists of direct runoff from impervious area and overflow from upper zone tension water, both of which must be properly attenuated to avoid overshooting the observed peak. Finally, during dry periods, a dominant process consisting of slow release from the lower zone often dominates model performance. These types of insights regarding transitions between modeled processes are not attainable from a priori selection of events assumed to be broadly representative. The coarser event-scale sensitivities are typically obscured, and are not necessarily consistent even for seemingly similar events (as highlighted in Figs. 4 and 5).

It should be emphasized that even though Fig. 9 represents a qualitative aggregation of the high-resolution sensitivity patterns, this aggregation is drawn a posteriori from the full range of dynamic parameter activation characterized using the three-hour moving window. The value of the high-resolution approach, as shown in Figs. 6–8, is its ability to isolate parameter activation in space and time while avoiding the potential biases introduced by a priori event selection and aggregation. The high-resolution analysis removes these biases by reducing the size of the interval window such that peak flows do not accumulate undue influence relative to the rest of the examined interval. In the high-resolution results, the dynamic transitions between upper and lower zone sensitivity become clear: the lower zone maintains a fairly constant level of control over model performance throughout the simulation, while the upper zone dominates during large events. The upper zone storage elements are the first to receive precipitation during large events and therefore exert the most control over the timing and magnitude of the quick response. The lower zone elements release water more slowly and are most responsible for model performance in the absence of large events. In prior analysis of the SAC-SMA model (a spatially lumped version of HL-RDHM), it was found that the lower zone parameters almost exclusively control the RMSE metric at the monthly timescale (Herman et al., 2013b). By zooming in to a 3 h time step, the high-resolution method identifies the importance of the upper zone parameters for properly reproducing quick responses. This approach is able to identify sensitive cells which were not

visible at the event scale as well as the timing of their activation, making it a valuable addition to traditional diagnostic approaches.

The high-resolution sensitivity approach presented here requires several important considerations. First, as with any sensitivity analysis, the results strongly depend on the choice of performance metric. Figures 4 and 5 show the substantial differences in sensitivity indices that occur when changing from the RMSE metric to the ROCE metric. We focus on the RMSE metric in this study because its emphasis on large events is consistent with our goal of understanding model behavior in the context of quick-response flood forecasting. Interestingly, as the window size of the analysis decreases, different performance metrics begin to behave similarly (i.e., in the limit as window size approaches zero, most metrics reduce to a percent error at a single point). This leads to the second important consideration, the choice of window size. Modeled processes which dominate performance at one timescale may be invisible at another, so it is crucial to choose a window size commensurate with the purpose of the analysis. Our moving window size of 24 h (with a 3 h time step) reflects the need to capture dominant processes on a fast timescale while also containing a sufficient number of time steps to smooth out any noise in the performance metric. Finally, the visualization approach presented in Figs. 6–8 (in which the spatial dimensions are compressed into a single distance measure on the y axis) is readily applicable to the narrow Blue River watershed, but may be difficult to extend to other watersheds. This is particularly true if significant spatial heterogeneity exists in land cover or soil storage properties. In such instances, it may be preferable to represent the y axis as, for example, the soil storage capacity of each grid cell, or whichever characteristic is expected to govern grid cell sensitivity. For this case study, the primary characteristic of interest is simply the distance from the watershed outlet, but this may not hold true for all applications.

This high-resolution approach is intended to complement, rather than replace, the insights derived from an event-scale analysis. In this case, neither the event-scale results nor the high-resolution results could have been predicted from the other. The event-scale analysis provides performance controls for selected intervals of interest; the high-resolution analysis contributes clear details of parameter dynamics across the simulation period without focusing on any particular interval. The need for complementary approaches at different temporal resolutions is further highlighted by the fact that a given parameter may only be sensitive at a certain timescale, considering these sensitivity measures describe the effect of a parameter relative to the others, and these relative effects are extremely likely to change depending on the scope of the analysis.

5 Conclusions

High-resolution sensitivity analysis explores the full spatial and temporal variability of distributed watershed model controls, highlighting the importance of avoiding confounding aggregation to the extent permitted by computational constraints. The complexity of spatially distributed models typically causes a significant fraction of parameters to be inactive at any particular time, a phenomenon clearly shown in the event-scale results of this study. This sparsity of activation can lead to needless complexity and inappropriate modification of inactive parameters. However, it also presents a valuable opportunity to overcome the complexity of distributed parameter identification by restricting search to only those parameters which are active at a specific time and location. It also suggests an opportunity to identify locations and timing for optimal data collection to improve the modeled representation of hydrologic processes, particularly under non-stationary conditions in which dominant watershed processes fall outside observed ranges. For example, the results of this study indicate that large streamflow events in the model are controlled primarily by upper zone fluxes quite close to the watershed outlet; collecting flux data in only this area during a large event could provide justification to falsify the internal processes of the model, and to improve them by calibrating against the new observations. As demonstrated in this study, spatial variability can easily be visualized as a time series and provides valuable information for analyzing model behavior. In light of these opportunities, it is imperative for diagnostic analyses of distributed models to explore parameter activation at the spatial and temporal scales for which the model was designed. This study represents a novel step in this direction by visualizing spatially explicit, time-varying watershed model sensitivity. As computational power continues to increase, such methods improve the potential for efficiently isolating distributed model behaviors at high spatial and temporal resolutions, an area which remains largely unexplored relative to similar analyses of simpler lumped models.

Supplementary material related to this article is available online at <http://www.hydrol-earth-syst-sci.net/17/5109/2013/hess-17-5109-2013-supplement.zip>.

Acknowledgements. The authors of this work were partially supported by the US National Science Foundation under grant EAR-0838357. The computational resources for this work were provided in part through instrumentation funded by the National Science Foundation through grant OCI-0821527. Any opinions, findings, and conclusions are those of the authors and do not necessarily reflect the views of the US National Science Foundation.

Edited by: S. Archfield

References

- Alton, P., Mercado, L., and North, P.: A sensitivity analysis of the land-surface scheme JULES conducted for three forest biomes: Biophysical parameters, model processes, and meteorological driving data, *Global Biogeochem. Cy.*, 20, GB1008, doi:10.1029/2005GB002653, 2006.
- Bastidas, L., Hogue, T., Sorooshian, S., Gupta, H., and Shuttleworth, W.: Parameter sensitivity analysis for different complexity land surface models using multicriteria methods, *J. Geophys. Res.*, 111, 20101, doi:10.1029/2005JD006377, 2006.
- Burnash, R. and Singh, V.: The NWS River Forecast System–Catchment Modeling, in: *Computer Models of Watershed Hydrology*, 311–366, Water Resour. Publ., Littleton, Colorado, 1995.
- Campolongo, F., Cariboni, J., and Saltelli, A.: An effective screening design for sensitivity analysis of large models, *Environ. Modell. Softw.*, 22, 1509–1518, 2007.
- Campolongo, F., Saltelli, A., and Cariboni, J.: From screening to quantitative sensitivity analysis, A unified approach, *Computer Phys. Commun.*, 182, 978–988, 2011.
- Carpenter, T. M., Georgakakos, K. P., and Sperflage, J. A.: On the parametric and NEXRAD-radar sensitivities of a distributed hydrologic model suitable for operational use, *J. Hydrol.*, 253, 169–193, 2001.
- Cloke, H., Pappenberger, F., and Renaud, J.: Multi-method global sensitivity analysis (MMGSA) for modelling floodplain hydrological processes, *Hydrol. Process.*, 22, 1660–1674, 2008.
- Cuo, L., Giambelluca, T. W., and Ziegler, A. D.: Lumped parameter sensitivity analysis of a distributed hydrological model within tropical and temperate catchments, *Hydrol. Process.*, 25, 2405–2421, doi:10.1002/hyp.8017, 2011.
- Demaria, E., Nijssen, B., and Wagener, T.: Monte Carlo sensitivity analysis of land surface parameters using the Variable Infiltration Capacity model, *J. Geophys. Res.*, 112, D11113, doi:10.1029/2006JD007534, 2007.
- Franchini, M., Wendling, J., Obled, C., and Todini, E.: Physical interpretation and sensitivity analysis of the TOPMODEL, *J. Hydrol.*, 175, 293–338, 1996.
- Freer, J., Beven, K., and Ambrose, B.: Bayesian estimation of uncertainty in runoff prediction and the value of data: An application of the GLUE approach, *Water Resour. Res.*, 32, 2161–2173, 1996.
- Gupta, H., Sorooshian, S., and Yapo, P.: Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information, *Water Resour. Res.*, 34, 751–763, 1998.
- Gupta, H., Wagener, T., and Liu, Y.: Reconciling theory with observations: elements of a diagnostic approach to model evaluation, *Hydrol. Process.*, 22, 3802–3813, 2008.
- Gupta, H. V., Kling, H., Yilmaz, K. K., and Martinez, G. F.: Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, *J. Hydrol.*, 377, 80–91, 2009.
- Guse, B., Reusser, D. E., and Fohrer, N.: How to improve the representation of hydrological processes in SWAT for a lowland catchment—temporal analysis of parameter sensitivity and model performance, *Hydrol. Process.*, doi:10.1002/hyp.9777, in press, 2013.
- Hall, J., Tarantola, S., Bates, P., and Horritt, M.: Distributed sensitivity analysis of flood inundation model calibration, *J. Hydraul. Eng.*, 131, 117–126, 2005.
- Herman, J. D., Kollat, J. B., Reed, P. M., and Wagener, T.: Technical Note: Method of Morris effectively reduces the computational demands of global sensitivity analysis for distributed watershed models, *Hydrol. Earth Syst. Sci.*, 17, 2893–2903, doi:10.5194/hess-17-2893-2013, 2013a.
- Herman, J., Reed, P., and Wagener, T.: Time-varying sensitivity analysis clarifies the effects of watershed model formulation on model behavior, *Water Resour. Res.*, 49, 1400–1414, doi:10.1002/wrcr.20124, 2013b.
- Hornberger, G. and Spear, R.: Approach to the preliminary analysis of environmental systems, *J. Environ. Manage.*, 12, 7–18, 1981.
- Koren, V., Reed, S., Smith, M., Zhang, Z., and Seo, D.-J.: Hydrology laboratory research modeling system (HL-RMS) of the US national weather service, *J. Hydrol.*, 291, 297–318, 2004.
- Moreda, F., Koren, V., Zhang, Z., Reed, S., and Smith, M.: Parameterization of distributed hydrological models: learning from the experiences of lumped modeling, *J. Hydrol.*, 320, 218–237, 2006.
- Morris, M. D.: Factorial Sampling Plans for Preliminary Computational Experiments, *Technometrics*, 33, 161–174, ArticleType: research-article / Full publication date: May, 1991/Copyright© 1991 American Statistical Association and American Society for Quality, 1991.
- Muleta, M. and Nicklow, J.: Sensitivity and uncertainty analysis coupled with automatic calibration for a distributed watershed model, *J. Hydrol.*, 306, 127–145, 2005.
- Pappenberger, F., Beven, K. J., Ratto, M., and Matgen, P.: Multi-method global sensitivity analysis of flood inundation models, *Adv. Water Resour.*, 31, 1–14, 2008.
- Pujol, G., Iooss, B., and Janon, A.: Sensitivity Analysis Package, available at: <http://cran.r-project.org/web/packages/sensitivity/index.html>, R package version 1.7-0, 2013.
- Reed, S., Koren, V., Smith, M., Zhang, Z., Moreda, F., Seo, D.-J., and DMIP Participants, a.: Overall distributed model intercomparison project results, *J. Hydrol.*, 298, 27–60, 2004.
- Reusser, D. and Zehe, E.: Inferring model structural deficits by analyzing temporal dynamics of model performance and parameter sensitivity, *Water Resour. Res.*, 47, W07550, doi:10.1029/2010WR009946, 2011.
- Reusser, D., Blume, T., Schaeffli, B., and Zehe, E.: Analysing the temporal dynamics of model performance for hydrological models, *Hydrol. Earth Syst. Sci.*, 13, 999–1018, doi:10.5194/hess-13-999-2009, 2009.
- Reusser, D., Buytaert, W., and Zehe, E.: Temporal dynamics of model parameter sensitivity for computationally expensive models with the Fourier amplitude sensitivity test, *Water Resour. Res.*, 47, W07551, doi:10.1029/2010WR009947, 2011.
- Rosolem, R., Gupta, H. V., Shuttleworth, W. J., Gonçalves, L. G. G., and Zeng, X.: Towards a comprehensive approach to parameter estimation in land surface parameterization schemes, *Hydrol. Process.*, 27, 2075–2097, 2012.
- Ruano, M., Ribes, J., Seco, A., and Ferrer, J.: An improved sampling strategy based on trajectory design for application of the Morris method to systems with many input factors, *Environ. Modell. Softw.*, 37, 103–109, 2012.

- Sieber, A. and Uhlenbrook, S.: Sensitivity analyses of a distributed catchment model to verify the model structure, *J. Hydrol.*, 310, 216–235, 2005.
- Smith, M. B., Seo, D.-J., Koren, V. I., Reed, S. M., Zhang, Z., Duan, Q., Moreda, F., and Cong, S.: The distributed model intercomparison project (DMIP): motivation and experiment design, *J. Hydrol.*, 298, 4–26, 2004.
- Smith, M. B., Koren, V., Reed, S., Zhang, Z., Zhang, Y., Moreda, F., Cui, Z., Mizukami, N., Anderson, E. A., and Cosgrove, B. A.: The distributed model intercomparison project – Phase 2: Motivation and design of the Oklahoma experiments, *J. Hydrol.*, 418–419, 3–16, 2012.
- Tang, Y., Reed, P., Van Werkhoven, K., and Wagener, T.: Advancing the identification and evaluation of distributed rainfall–runoff models using global sensitivity analysis, *Water Resour. Res.*, 43, W06415, doi:10.1029/2006WR005813, 2007.
- van Griensven, A., Meixner, T., Grunwald, S., Bishop, T., Diluzio, M., and Srinivasan, R.: A global sensitivity analysis tool for the parameters of multi-variable catchment models, *J. Hydrol.*, 324, 10–23, 2006.
- Van Werkhoven, K., Wagener, T., Reed, P., and Tang, Y.: Characterization of watershed model behavior across a hydroclimatic gradient, *Water Resour. Res.*, 44, W01429, doi:10.1029/2007WR006271, 2008a.
- Van Werkhoven, K., Wagener, T., Reed, P., and Tang, Y.: Rainfall characteristics define the value of streamflow observations for distributed watershed model identification, *Geophys. Res. Lett.*, 35, L11403, doi:10.1029/2008GL034162, 2008b.
- Van Werkhoven, K., Wagener, T., Reed, P., and Tang, Y.: Sensitivity-guided reduction of parametric dimensionality for multi-objective calibration of watershed models, *Adv. Water Resour.*, 32, 1154–1169, 2009.
- Wagener, T., Boyle, D. P., Lees, M. J., Wheeler, H. S., Gupta, H. V., and Sorooshian, S.: A framework for development and application of hydrological models, *Hydrol. Earth Syst. Sci.*, 5, 13–26, doi:10.5194/hess-5-13-2001, 2001.
- Wagener, T., McIntyre, N., Lees, M., Wheeler, H., and Gupta, H.: Towards reduced uncertainty in conceptual rainfall–runoff modelling: Dynamic identifiability analysis, *Hydrol. Process.*, 17, 455–476, 2003.
- Wagener, T., Reed, P., van Werkhoven, K., Tang, Y., and Zhang, Z.: Advances in the identification and evaluation of complex environmental systems models, *J. Hydroinform.*, 11, 266–281, 2009a.
- Wagener, T., Werkhoven, K. v., Reed, P., and Tang, Y.: Multiobjective sensitivity analysis to understand the information content in streamflow observations for distributed watershed modeling, *Water Resour. Res.*, 45, W02501, doi:10.1029/2008WR007347, 2009b.
- Yatheendradas, S., Wagener, T., Gupta, H., Unkrich, C., Goodrich, D., Schaffner, M., and Stewart, A.: Understanding uncertainty in distributed flash flood forecasting for semiarid regions, *Water Resour. Res.*, 44, W05S19, doi:10.1029/2007WR005940, 2008.
- Zaehle, S., Sitch, S., Smith, B., and Hatterman, F.: Effects of parameter uncertainties on the modeling of terrestrial biosphere dynamics, *Global Biogeochem. Cy.*, 19, GB3020, doi:10.1029/2004GB002395, 2005.