Patterns of predictability in hydrological threshold systems

E. Zehe,1 H. Elsenbeer,1 F. Lindenmaier,1 K. Schulz,2 and G. Blöschl1

Received 4 October 2006; revised 5 February 2007; accepted 14 March 2007; published 20 July 2007.

Observations of hydrological response often exhibit considerable scatter that is difficult to interpret. In this paper, we examine runoff production of 53 sprinkling experiments on the water-repellent soils in the southern Alps of Switzerland; simulated plot scale tracer transport in the macroporous soils at the Weiherbach site, Germany; and runoff generation data from the 2.3-km² Tannhausen catchment, Germany, that has cracking soils. The response at the three sites is highly dependent on the initial soil moisture state as a result of the threshold dynamics of the systems. A simple statistical model of threshold behavior is proposed to help interpret the scatter in the observations. Specifically, the model portrays how the inherent macrostate uncertainty of initial soil moisture translates into the scatter of the observed system response. The statistical model is then used to explore the asymptotic pattern of predictability when increasing the number of observations, which is normally not possible in a field study. Although the physical and chemical mechanisms of the processes at the three sites are different, the predictability patterns are remarkably similar. Predictability is smallest when the system state is close to the threshold and increases as the system state moves away from it. There is inherent uncertainty in the response data that is not measurement error but is related to the observability of the initial conditions.


1. Introduction

[1] Observations of hydrological response often exhibit considerable scatter that is difficult to interpret. In this paper, we examine runoff production of 53 sprinkling experiments on the water-repellent soils in the southern Alps of Switzerland; simulated plot scale tracer transport in the macroporous soils at the Weiherbach site, Germany; and runoff generation data from the 2.3-km² Tannhausen catchment, Germany, that has cracking soils. The response at the three sites is highly dependent on the initial soil moisture state as a result of the threshold dynamics of the systems. A simple statistical model of threshold behavior is proposed to help interpret the scatter in the observations. Specifically, the model portrays how the inherent macrostate uncertainty of initial soil moisture translates into the scatter of the observed system response. The statistical model is then used to explore the asymptotic pattern of predictability when increasing the number of observations, which is normally not possible in a field study. Although the physical and chemical mechanisms of the processes at the three sites are different, the predictability patterns are remarkably similar. Predictability is smallest when the system state is close to the threshold and increases as the system state moves away from it. There is inherent uncertainty in the response data that is not measurement error but is related to the observability of the initial conditions.

Spatial probability distribution of soil moisture at the catchment scales [Deeks et al., 2004; Western et al., 2004]. Even if the probability density function (pdf) were known, the detailed spatial pattern will still elude full quantification.

[2] Understanding predictive model uncertainty (its quantification and its reduction) is a major concern in hydrological science [Sivapalan et al., 2003]. This includes an understanding of how the uncertainties of model parameters, model structure, and inputs affect the uncertainty of the model predictions. However, as pointed out by Beven [1996, p. 260], one cannot expect process models to be more accurate than the repeatability of nature itself. Hence it is similarly important to understand how well one may repeat observations of cause and effect. Repeatability of experiments or observations is a prerequisite to the predictability of that system. A system one hopes to predict with perfect accuracy needs to give exactly the same response if an experiment is repeated under identical conditions, i.e., when the forcings and the initial states of the repeated trials are exactly the same. However, distributed hydrological measurements are never exhaustive and therefore uncertain. Their support, spacing, and extent scales tend to be incompatible with the scale of state variables and processes in natural systems [Blöschl and Sivapalan, 1995]. For instance, the space-time pattern of soil moisture is usually estimated from a set of point observations, for example, using time domain reflectometry (TDR). Typically, these measurements do not even suffice to fully characterize the spatial probability distribution of soil moisture at the catchment scales [Deeks et al., 2004; Western et al., 2004]. Even if the probability density function (pdf) were known, the detailed spatial pattern will still elude full quantification.

[3] Hydrological states and, in particular, their spatial patterns are hence not fully observable and inherently uncertain [Rundel et al., 2006]. Because of this, one can never repeat an observation under truly identical conditions albeit under apparently identical conditions. Observations of hydrological system response, such as tracer transport depths, residence times, and runoff volumes, often exhibit considerable scatter, even if the initial and boundary conditions are apparently identical [Lischeid et al., 2000; Zehe et al., 2005; Zehe and Blöschl, 2004]. This implies a lack of repeatability and hence poor predictability. Unobserved (and indeed unobservable) small-scale variability in the states along with nonlinear system characteristics may explain this lack of repeatability [Pitman and Stouffer, 2006; Rundel et al., 2006; Zehe and Blöschl, 2004]. If we want to understand the limits to the accuracy of model predictions, we have to understand how inherent uncertainties of state observations limit the repeatability of observations of the hydrological system response, especially in the presence of strongly nonlinear behavior such as threshold controlled dynamics.

[4] Pitman and Stouffer [2006] define threshold behavior as abrupt changes in system dynamics that occur at time-scale that are much smaller than the usual timescales of the system. Threshold behavior can hence be considered as an extreme type of nonlinear dynamics. Threshold-type nonlinearities occur when state variables switch from zero to nonzero values [Blöschl and Zehe, 2005] as is the case for...
water levels of overland flow, rainfall rates, or flow in macropore systems [Vogel et al., 2005; Zehe and Flühler, 2001b]. Threshold behavior may be further enhanced by emerging and vanishing features as in the case of swelling and cracking soils [Návar et al., 2002; Lindenmaier et al., 2006] and hydrophobic soils [e.g., DeJonge et al., 1999; Doerr and Thomas, 2000], and during phase transitions.

[5] Zehe and Blöschl [2004] drew from the concept of macrostates and microstates of gases in statistical mechanics [Tolman, 1979] to characterize the observability of state variables in environmental systems. The macrostate is represented by the values of the kinetic energy of each of the individual molecules in a volume of gas which are never observable. The macrostate, in contrast, is defined as the average kinetic energy (or gas temperature) of the gas molecules in that volume, and it is the macrostate that is observable. A very large number of possible microstates are consistent with the same macrostate. The term “identical state” of two volumes of gas can only refer to their macrostates, as the microstates cannot be observed and hence cannot be compared.

[6] In analogy, the microstate of initial soil moisture can be defined as the true spatial pattern of soil moisture of, say, a field plot. Clearly, this microstate is not observable as there will always be fine-scale detail that cannot be resolved by the measurements. The macrostate can, for example, be defined as the first two statistical moments, the variogram and the probability density function of soil moisture, as these are the variables that are observable in a typical field study. An infinite number of possible soil moisture patterns (i.e., microstates) will be consistent with the same observed macrostate. This will introduce uncertainty that will propagate through the system and cause uncertainty in the system response such as streamflow and infiltration patterns.

[7] Zehe and Blöschl [2004] conducted numerical simulations to illustrate the effect of lack of knowledge of the microstate on hydrologic response in the presence of threshold behavior. They found that, under apparently identical conditions, repeated trials produced uncertainty in system response of more than 100% of the average response. So far, it has, however, been unclear whether these effects are consistent with the observed scatter of hydrological field experiments and whether they apply to other threshold processes.

[8] The objective of this paper is to contribute to a better understanding of the repeatability of experiments and hence the scatter of observed hydrological response by analyzing the transformation of the inherent uncertainties of observed states into uncertainties in hydrological response. We will examine the transformation for three different processes with threshold behavior at the plot and catchment scales.

[9] This paper goes beyond that of Zehe and Blöschl [2004] in a number of ways. First, we use field data from three case studies to examine three different processes that involve threshold behavior, whereas Zehe and Blöschl [2004] used simulations for a single process (matrix/macropore flow transitions). Second, we propose a statistical model that helps interpret field observations for a limited number of trials relative to a hypothetical ensemble of an infinite number of observations.

[10] This paper is organized as follows. Section 2 reviews the three threshold processes for three field experiments and gives details on the field data. Section 3 presents the statistical model and the methods to quantify the repeatability and hence predictability of the response. Section 4 gives the results, which are discussed in section 5.

2. Threshold Processes, Study Sites, and Data

2.1. Plot Scale Runoff Generation in a Water-Repellent Soilscape

2.1.1. Water Repellency and Runoff Generation

[11] The first process examined is runoff generation on water-repellent soils. Water repellency was reported as early as 1910 by Schreiner and Shorey [1910] for soils in California that could be wetted neither by infiltration nor by the rise of groundwater tables. Water repellency of soils is related to degraded organic matter [Krammes and DeBano, 1965] and sometimes to forest fires [de Bano and Rice, 1973]. Water repellency is often associated with sandy soils and occasionally with clay soils. The main controls are the type of organic matter, the occurrence of the dry season, and the soil moisture [DeJonge et al., 1999]. de Bano and Rice [1973] suggested that water repellency occurs after soil moisture drops below a certain threshold. Bond [1964] emphasized the role of organic matter coatings on coarse grains, whose hydrophobicity forces water to act as a nonwetting fluid that tries to minimize the interface to the water-repellent environment. The most important concepts of quantifying the degree of hydrophobicity/water repellency are the “Water Drop Penetration Time” [Letey et al., 2000], which is the time it takes for a water droplet to infiltrate into the soil, and the “Molarity of Aqueous Ethanol Droplets” [King, 1981], which is the concentration of an ethanol/water mix that infiltrates within 10 s into the soil sample. The higher the molarity of the ethanol/water mix, the higher will be the degree of water repellency. A recent bibliography on water repellency [Dekker et al., 2005] highlights the global occurrence of this phenomenon. Although its effects on runoff generation have been studied globally since the 1960s [e.g., Krammes and DeBano, 1965; Burch et al., 1989; Keizer et al., 2005], the role of antecedent soil moisture over a range of scales has received much less attention [e.g., Doerr and Thomas, 2000; Doerr et al., 2003].

2.1.2. Sprinkling Experiments in the Southern Swiss Alps

[12] To shed light on the control of antecedent soil moisture on surface runoff generation in a landscape with hydrophobic soils, we performed sprinkling experiments in a landscape prone to occasional droughts and wildfires. The research area is located in southern Switzerland at 46°30’N, 9°01’E at an elevation of about 720 m. Aspects range from S to SW and slopes range from 20° and 37°. The area is covered by mixed deciduous forest (Castanea sativa, Fagus silvatica, Quercus sp.). The dominant soil type is Dystric Cambisol [International Society of Soil Sciences/International Soil Reference and Information Centre/Food and Agriculture Organization (ISSS-ISRIC-FAO), 1998]. Mean annual precipitation is 1800 mm, with a pronounced dry season between December and March.
Fifty-three sprinkling experiments were performed using an automated sprinkling device [Ritschard, 2000]. We applied artificial rainfall with an intensity of 50 mm/h, which occurs with a return interval of 5 years. Each of the 53 test plots was 1 m\(^2\) in size. Figure 1 shows a typical field plot. To study the effect of dry periods, we covered the plots for a period ranging from 10 to 100 days. Runoff from the plots was measured manually, and initial soil moisture was measured at two points in each plot using time domain reflectometry. Water repellency was estimated before each sprinkling experiment by determining the water drop penetration time and by performing the molarity-of-ethanol-droplet test on air-dried soil samples [Letey et al., 2000]. According to King’s [1981] classification, all plots with the exception of one were extremely hydrophobic at the onset of the sprinkling experiments. Figure 2 shows examples of observed runoff at the field plot shown in Figure 1 for wet and dry antecedent soil moisture conditions (0.30 and 0.07 m\(^3\)/m\(^3\), respectively). Because of the high saturated hydraulic conductivities of about 5 \(\times\) 10\(^{-5}\) m/s, a large part of the water infiltrates when the soil is wet. In contrast, when the soil is dry, surface runoff triples as a result of water repellency.

2.2. Plot Scale Tracer Transport in Macroporous Heterogeneous Soils

2.2.1. Preferential Flow and Solute Transport

The second process we examined is plot scale transport in heterogeneous soils. When dealing with contamination of shallow groundwater, preferential flow often is a key process [Stamm et al., 1998; Flury, 1996; Zehe and Flühler, 2001a]. Typical transport distances of preferential flow events in macroporous soils are on the order of 50–100 cm [Flury, 1996]. The main control on the occurrence of preferential flow in macropores is the presence of interconnected preferential pathways that link the soil surface and the subsoil [Vogel et al., 2005; Zehe and Flühler, 2001b]. Other important controls are soil moisture and rainfall magnitude as found, for example, by the dye tracer experiments in the Weiherbach catchment [Zehe and Blöschl, 2004].

2.2.2. Simulated Replicates of Plot Scale Transport Experiments in a Macroporous Soil

In this example, we use results from Monte Carlo simulations that are based on plot scale tracer experiments on a 1-m\(^2\) field site in a highly macroporous Colluvisol [ISSS-ISRIC-FAO, 1998]. The study site is located in the rural Weiherbach valley at 49°01’N, 9°00’E which is a typical catchment for this landscape named “Kraichgau.” Preferential flow events at this study site were observed several times: Tracer transport depths observed during plot scale tracer experiments (Figure 3) were much larger than predicted when assuming that matrix flow was the only relevant process [Zehe and Blöschl, 2004]. Furthermore, observed tracer breakthrough into tile drains located at 1.2 m below surface was faster by 2 to 3 orders of magnitude than simulations based on matrix flow and transport assumptions [Zehe and Flühler, 2001a].

Geologically, the Weiherbach valley consists of Keuper and Loess layers of up to 15-m thickness. The climate in the Weiherbach valley is humid with an average annual precipitation of 750–800 mm, average annual runoff of 150 mm, and annual potential evapotranspiration of 775 mm. Most of the Weiherbach hillslopes exhibit a typical loess catena with the moist but drained Colluviosols at the hill foot and drier Calcaric Regosols [ISSS-ISRIC-FAO, 1998] at the top and midslope sector. Interconnected preferential pathways are apparent in the Colluvisol and are a result of earthworm activities (aneic earthworms such as Lumbicus terrestris).

Zehe and Blöschl [2004] simulated repeated trials of the transport experiments by the physically based CATFLOW model. The model was extensively tested against field data of that site. On the basis of the findings of Zehe and Blöschl, the model was used to simulate repeated experiments by tuning the model parameters to the specific site conditions. The results showed that the model was able to accurately predict the observed tracer breakthrough times and the spatial distribution of the tracer within the soil profile. This indicates that the model can be used to simulate preferential flow and solute transport in heterogeneous soils under realistic field conditions.
they represented preferential flow and transport in the macropore system as a threshold process that starts when soil water saturation exceeds field capacity. This is plausible as free gravity water is present in this case that may percolate into coarse pores and macropores. In a first step, they statistically generated a macroporous medium with the properties observed in the field based on the results of a detailed macropore and soil mapping (details are explained in the work of Zehe and Flühler [2001a]). They defined the initial soil moisture macrostate by the first two moments of soil moisture obtained from two sets of 25-point measurements collected at two nearby plots of 4 m². Soil moisture was measured using a portable two-rod TDR sensor that integrated over the upper 15 cm of the soil. They then generated 40 realizations of soil moisture patterns, each pattern representing one possible microstate that was consistent with the observed macrostate. Using these soil moisture patterns as initial conditions, they simulated events of two-dimensional transport of a conservative tracer. The plot scale simulations of Zehe and Blöschl [2004] are used here to analyze the threshold dynamics in the context of the other field studies.

2.3. Catchment Scale Runoff Generation in a Soilscape With Cracking Soils

2.3.1. Cracking Soils and Runoff Generation

The third process is runoff generation on cracking soils at the small catchment scale. Shrinkage and swelling of clay soils can lead to significant changes in the infiltrability and the soil volume and is usually a result of changes in soil moisture. Internal shrinkage and swelling of clay minerals such as smectites or vermiculites may influence shrinkage characteristics of soils or clay pastes [Kariuki and van der Meer, 2004]. Especially the size and structure of the clay minerals determine the soil porosity [Chertkov, 2000]. The shrinkage characteristic curve represents the volume change of soils in relation to soil moisture [McGarry and Malafant, 1987]. Four zones of volume change are usually distinguished: zero, residual, normal, and structural shrinkage [Braudeau et al., 1999]. The “normal” shrinkage zone, generally, is the most important one in terms of the volume change in soils [Peng and Horn, 2005], as volume loss is approximately proportional to the water loss. In this zone, a deformation of pores occurs instead of a penetration or displacement of air in the pores [Bronswijk, 1988; Chertkov, 2000]. Because of this, the volume change of the pores may influence the hydraulic characteristics of the matrix in nonrigid soils [Kim et al., 1999; Peng and Horn, 2005; Chertkov, 2004] and may even lead to the development of cracks in dry soils which will then allow preferential infiltration into lower soil horizons [Bronswijk, 1988; Wells et al., 2003]. If the soil wets up above a certain threshold, the cracks tend to close successively and surface runoff generation will occur more often. This is, for example, the case of the vertisols of northern Mexico [Návar et al., 2002]. For temperate climates, the normal and structural shrinkage zones are the most important ones, as they occur within the typical ranges of soil moisture variability in these climates.

2.3.2. Observed Rainfall-Runoff Response at the Tannhausen Catchment

This example uses runoff data from the 2.3-km² Tannhausen catchment that has cracking clay soils (Figure 4). It is located in the headwaters of the Danube River (49°10’N, 10°01’E) and is typical for this landscape named “Hohenloher Ebene.” Mean annual precipitation is 660 mm. The geology consists of clayey and marly sediments of Lower Jurassic age from which Luvisols, stagnic Gleysols, and Regosols [ISRIC-FAO, 1998] have developed. Hydraulic conductivities range from 1.5 × 10⁻⁶ to 2 × 10⁻⁷ m/s for the upper soil horizons (0–40 cm) to 2 × 10⁻⁷ to 2 × 10⁻⁹ m/s in the lower soil horizons (>40 cm). Because of their high clay content of 50–75%, the Regosol soils show characteristic shrinking and crack patterns during the dry periods as well as swelling and vanishing of the cracks during wet periods. Opening and closing of the cracks as a result of soil moisture changes modifies the infiltrability of the soils which in turn leads to a strong seasonality in runoff response [Lindenmaier et al., 2006].

In the present study, we determined the runoff coefficients for a total of 260 rainfall-runoff events observed in the period from 1994 to 2004. Figure 5 shows the rainfall-runoff response observed in the Tannhausen catchment (2.3 km²) for a summer and a winter event. Although event precipitation and antecedent precipitation [Mosley, 1979] within the last 10 days were very similar for these two events, runoff response was vastly different. In summer, almost all the precipitation infiltrated because of the open cracks (see Figure 4, right panel). During the winter season, the cracks closed and the catchment responded with significant surface runoff production. As no soil moisture observations were available in this catchment, we used the CATFLOW model [Zehe and Blöschl, 2004; Lindenmaier et al., 2006] to estimate soil moisture in the upper 30 cm of the profile.
the soil. On the basis of a digital elevation model, soil hydraulic functions for the soils, as well as observed meteorological data, the water balance of the catchment was simulated for the period of 1994–2004. For each time step, the simulated soil moisture field was averaged over the catchment yielding a time series of catchment average soil moisture as an estimator for the initial states.

3. A Seven-Parameter Statistical Model of Threshold Behavior

[21] A simple statistical model of threshold behavior is proposed here to help interpret how inherent macrostate uncertainty translates into the scatter of the observed system response. This model is then used to explore the asymptotic pattern of predictability when increasing the number of observations/simulations, which is normally not possible in a field study.

[22] We assume that threshold behavior in a hydrological system produces two stable dynamic regimes. These regimes are defined by finite ranges of a control variable, in the present case, the soil moisture macrostate \( \theta_{\text{macro}} \). In each of the two regimes, system response can be represented by a stationary probability density function \( f_{\text{res}} \), which we assume to be Gaussian. The response functions differ with respect to their expectation \( m_{\text{res}} \) and their standard deviation \( \sigma_{\text{res}} \) [compare to scheme in Figure 6, equation (1)]:

\[
\begin{align*}
    f_{\text{res}}^1 &= N(m_{\text{res}}^1, \sigma_{\text{res}}^1) \quad \text{for } \theta_{\text{macro}} < \theta_{\text{macro}}^1 \\
    f_{\text{res}}^2 &= N(m_{\text{res}}^2, \sigma_{\text{res}}^2) \quad \text{for } \theta_{\text{macro}} > \theta_{\text{macro}}^2
\end{align*}
\]

In the unstable range of the control variable (i.e., between \( \theta_{\text{macro}}^1 \) and \( \theta_{\text{macro}}^2 \)), the system switches between the two response functions. This transition is conceptualized as an error function that relates the expectation of the response function to the control variable. The error function is centered at the midpoint between \( \theta_{\text{macro}}^1 \) and \( \theta_{\text{macro}}^2 \). The standard deviation of the error function, i.e., the width of the transition, is chosen as one eighth of the difference \( \theta_{\text{macro}}^2 - \theta_{\text{macro}}^1 \), which assures that the error function is close to zero and unity at the range limits:

\[
    m_{\text{res}}(\theta_{\text{macro}}) = m_{\text{res}}^1 + (m_{\text{res}}^2 - m_{\text{res}}^1) \int_{\theta_{\text{macro}}^1}^{\theta_{\text{macro}}^2} \frac{\exp(-0.5(\theta_{\text{macro}} - \theta_{\text{macro}}^1)^2) \left( \frac{\theta_{\text{macro}} - 0.5(\theta_{\text{macro}}^2 - \theta_{\text{macro}}^1)}{0.25(\theta_{\text{macro}}^2 - \theta_{\text{macro}}^1)} \right)^2}{\theta_{\text{macro}}^2 - \theta_{\text{macro}}^1} \, d\theta_{\text{macro}}
\]
The transition of the standard deviation is represented by an error function in a similar way:

\[ \sigma_{\text{res}}(\theta_{\text{macro}}) = \sigma^1_{\text{res}} + (\sigma^2_{\text{res}} - \sigma^1_{\text{res}}) \int_{\theta_{\text{macro}}}^{\bar{\theta}_{\text{macro}}} \exp \left( \frac{[\theta_{\text{macro}} - 0.5(\bar{\theta}^2_{\text{macro}} - \bar{\theta}^1_{\text{macro}})]^2}{0.25(\bar{\theta}^2_{\text{macro}} - \bar{\theta}^1_{\text{macro}})} \right) d\theta_{\text{macro}} \]  

(3)

The model parameters are the mean and the standard deviations of the system response functions in the stable ranges, \( m^i_{\text{res}} \) and \( \sigma^i_{\text{res}} \) \( (i = 1, 2) \), as well as the upper limit of the first stable range \( \theta^1_{\text{macro}} \) and the lower limit of the second range \( \theta^2_{\text{macro}} \).

[23] In the examples of this paper, the control variable is the average antecedent soil moisture of a field site or a catchment. This is considered as a macrostate estimated from a limited number of measurements and is hence uncertain. In the proposed model, we assume that the estimated spatial average soil moisture obeys a Gaussian distribution, as suggested by the central limit theorem. The standard deviation \( \sigma_{\text{macro}} \) is set to the standard deviations of the distributed observations divided by the square root of the sample size.

[24] If (1) the response functions in the different regimes are known, for example, estimated from field observations, (2) the width of the unstable range is known, and (3) the uncertainty of the macrostate \( \sigma_{\text{macro}} \) is known, the behavior of the threshold system can be simulated for different initial soil moisture macrostates. We did this in two steps. First, we generated a normally distributed random number that represents a possible soil moisture macrostate consistent with the observed macrostate. Second, we drew a second random number with a distribution according to \( f_{\text{res}} \), which represents the system response to a rainfall input. By repeating this procedure, we were able to examine the propagation of the macrostate uncertainty through the hydrological system as a function of average initial soil moisture (different soil moisture macrostates). It is intuitively clear that the macrostate uncertainty is not important in the stable ranges but will affect system response (and therefore its predictability) when the nonstationary range of the system response function and the uncertainty of the initial soil moisture macrostates overlap.

[25] To quantify hydrological predictability for a given initial soil moisture macrostate, we used the scaled range proposed by Zehe and Blöschl [2004]. The scaled range \( r \) is defined as the difference of the maximum and the minimum response divided by the average response for a given initial soil moisture macrostate.

\[ r = \frac{\text{max(response)} - \text{min(response)}}{\text{ave(response)}} \]  

(4)

[26] In the case of a limited sample size, we consider this measure as appropriate because it relates the extreme deviations to the average systems response, which is an intuitive concept. In the case of a very large sample size (representing the limiting case as the sample size approaches infinity), the range is not well defined. We hence used the scaled range of the 97.5th and 2.5th percentiles instead,

\[ r = \frac{\text{response}_{0.975} - \text{response}_{0.025}}{\text{ave(response)}} \]  

(5)

which represents an uncertainty band. In both cases, \( r \) is a measure of repeatability, i.e., of how precisely one can observe the system response under apparently identical conditions. The inverse of \( r \) is a measure of predictability;
for example, small values of the range \( r \) indicate large (or favorable) predictability.

4. Results

4.1. Plot Scale Runoff Generation in a Water-Repellent Soilscape

[27] Figure 7 (upper left panel) presents the total runoff volumes (expressed as runoff depths) observed at the 53 field sites in the southern Swiss Alps plotted against average initial soil moisture. Two stable regimes are apparent: runoff response with an average of 45 mm for initial soil moisture less than 0.11 m\(^3\) m\(^{-3}\), and a much weaker runoff response in average of 9.5 mm for soil moisture greater than 0.21 m\(^3\) m\(^{-3}\). The difference between the average runoff values observed in the two regimes is significant at a level of 95%. In the unstable range between 0.11 and 0.21 m\(^3\) m\(^{-3}\) initial soil moisture, the runoff response exhibits a much larger variability than in each of the stable ranges. From these data, we calculated the scaled range as a function of average initial soil moisture [equation (4)]. After several tests, we chose a width of 0.05 m\(^3\) m\(^{-3}\) for the soil moisture classes, which is sufficiently fine to detect the uncertainty peak and sufficiently coarse to give robust estimates of the range for this data set. The resulting scaled ranges peak in the class centered at 0.225 m\(^3\) m\(^{-3}\) initial soil moisture (Figure 7, upper right panel). A peak value of 1.5 means that the total range of the runoff depths in this class is 150% of the average response. Clearly, this would be considered poor repeatability and therefore poor predictability.

[28] In a second step, we obtained the parameters of the statistical model by estimating the mean and standard deviations of the observed runoff depths in the two stable ranges (Table 1). The uncertainty of the initial soil moisture macrostate was set to the measurement error of the time domain reflectometry device with a standard deviation of \( \sigma_{\text{macro}} = 0.02 \) m\(^3\) m\(^{-3}\). We then used the statistical model to simulate runoff response of the sprinkling experiments and varied the macrostate of average initial soil moisture. The number of simulated experiments in each soil moisture class was equal to the number of sprinkling experiments in that class. As can be seen from Figure 7 (left panels), the simulated pattern of the runoff volumes is very similar to the observed pattern. Also, the patterns of the scaled ranges of the simulations are quite similar to those of the observations, although the peaks do not exactly coincide (Figure 7, right panels). For this relatively low number of trials at each macrostate, the pattern of the scaled range does not fully reflect the patterns of the underlying statistical model.

4.2. Plot Scale Tracer Transport in Macroporous Heterogeneous Soils

[29] Figure 8 shows the simulated transport depths one day after tracer application plotted against the average initial soil moisture (upper left panel). There are, again, two stable regimes for initial soil moisture, either smaller...
than 0.15 m$^3$ m$^{-3}$ or larger than 0.25 m$^3$ m$^{-3}$. In these two ranges, the macrostate uncertainty translates into a small scatter of the average tracer transport depths. Transport dynamics is stable with an expected transport depth of 0.04 and 0.12 m, respectively, for the two regimes. In the unstable range between 0.15 and 0.25 m$^3$ m$^{-3}$, macrostate uncertainty is amplified and the average transport depths fill the entire range between the two stable regimes. The scaled ranges (Figure 8, upper right panel) peak at an average initial soil moisture of 0.18 m$^3$ m$^{-3}$ with a value of 1.2. The maximum uncertainty of the transport depth is hence 120% of the average value.

The statistical model was, again, used to mimic the behavior of the threshold system. The parameters were estimated from the simulated output with the full dynamic model and are given in Table 1. The uncertainty of the soil moisture macrostate was estimated from the field data as 0.02 m$^3$ m$^{-3}$. By generating 40 realizations for a given average initial soil moisture, we obtained the scatterplot of average transport depths and the scaled ranges (Figure 8, bottom panels). They are very similar to those of the full dynamic model, although the statistical model is much simpler. The similarity in the patterns suggests that the statistical model can indeed be used to help interpret the scatter in hydrological response measurements in case of threshold behavior.

### 4.3. Catchment Scale Runoff Generation in a Soilscape With Cracking Soils

Figure 9 (upper left panel) presents the runoff coefficients observed in the Tannhausen catchment plotted against average initial soil moisture macrostate, $i_{sm}$ (upper left panel, circles). For comparison, the crosses show the transport depths for a similar soil but without macropores. Scaled ranges of the transport depths plotted against soil moisture macrostate, $i_{sm}$ (upper right panel). Lower panels show the corresponding graphs simulated by the statistical model.
against the average initial soil moisture. Similar to the other process examples, there are two stable ranges of system response: weak runoff response with an average runoff coefficient of 0.06 for soil moisture less than 0.29 m³ m⁻³, and significant runoff response with an average runoff coefficient of 0.66 for average monthly soil moisture of more than 0.35 m³ m⁻³. Clearly, in the drier regime, the soil cracks are open, so most of the water infiltrates and does not produce surface runoff. In contrast, in the wet regime, the cracks are closed, so the contribution of rainfall to runoff is much larger. In the unstable range between 0.29 and 0.35 m³ m⁻³, the scatter is large. The scaled range of the observed runoff coefficients peaks at 0.32 m³ m⁻³ average soil moisture with a maximum of almost 4. This means that total uncertainty in the runoff coefficient is almost 400% of its average.

The parameters of the statistical model were obtained in a similar way as in the other examples. The macrostate uncertainty was estimated as 0.02 m³ m⁻³ (Table 1) using the standard deviation within the model grid according to Zeh and Blöschl [2004]. The simulation results are shown in the bottom panels of Figure 9. Again, the patterns are very similar to those obtained from the measurements, although the model yields a lower peak for the scaled range and therefore a better repeatability than that suggested by the data.

It is interesting that the patterns of runoff generation on water-repellent soils (Figure 7) are similar to those of plot scale transport in macroporous soils (Figure 8) and similar to the response patterns of the cracking soils (Figure 9). More specifically, in all cases, the repeatability of the observations is poorest in the vicinity of the threshold, which points to a more generic pattern of predictability of hydrologic response in the presence of threshold processes.

Asymptotic Simulations of the Statistical Model

As the conceptual model of threshold behavior is of a statistical nature, it is clear that, for a small number of trials, the simulated system response and the scaled ranges will strongly depend on the realization and exhibit a certain amount of randomness. This resembles the situation with field data that are inherently difficult to interpret, as the number of trials will always be limited. The statistical model proposed here can be used to shed light on the role of the number of trials in observed system response. Specifically, we used the model to examine the asymptotic behavior, i.e., the hypothetical case of having an infinite number of trials available for the same site. The model was used to simulate system response with the parameters listed in Table 1, but the number of realizations was set to 10,000 for each initial soil moisture state to obtain the results for the asymptotic behavior.

Figure 10 shows the results of this exercise. The upper panels present the results for plot scale runoff generation on the water-repellent soils in the southern Swiss Alps, the center panels those for plot scale tracer transport in the macroporous soils at the Weiherbach site, and the lower panels those for runoff generation on the cracking soils in the Tannhausen catchment. The corresponding scaled 95% uncertainty range and the coefficients of variation are shown in the right panels. The scaled ranges indicate that the overall pattern of predictability in the three case studies is
remarkably similar, although the physical and chemical mechanisms of the processes are different. There appears to exist inherent uncertainty in the measurements that is not measurement error but related to the observability of the initial conditions.

Figure 10 also indicates that the uncertainty ranges in the asymptotic case differ from those apparent in the observations of plot scale experiments (Figures 7 and 9) and the process model simulations (Figure 8). The maximum scaled 95% uncertainty range in Figure 10 are 3.5, 1.8, and 2.5 as compared to about 1.6, 1.2, and 4, respectively, in Figures 7–9. From a hydrological perspective, this means that, as the number of trials increases, the maximum range of the scatter will also increase. However, the distribution of the response, for a given macrostate initial soil moisture, can be estimated much more robustly as the number of trials increases.

5. Discussion and Conclusions

The analyses of the response data of the three sites suggest that threshold mechanisms are operative in each of the three cases. In the observations of plot scale surface runoff generation in the southern Swiss Alps, threshold behavior stems from the switch between hydrophobic and hydrophilic conditions, the latter resulting in substantially reduced runoff production. Although hydrophobicity is a result of soil chemical processes [DeJonge et al., 1999; Doerr and Thomas, 2000; Letey et al., 2000], antecedent soil moisture is clearly the main control at the macroscopic level. Soil moisture could hence be used as a predictor of soil hydrophobicity and runoff response characteristics in the stable ranges. In the unstable range, there is small-scale variability of soil moisture as well as other processes not captured that translate into significant scatter in runoff response for antecedent soil moisture macrostates that are apparently identical.

In the case of tracer transport simulated by a process model at the Weiherbach site, soil moisture controls the onset of macropore flow. However, the macrostates of initial soil moisture do not account for the details of the microstate. One of the important pieces of information that is lost in the macrostate description is the correlation between local soil saturation and local macroporosity at the plot. Zehe and Bloschl [2004] showed that if the soil was prone to switching from matrix to macropore flow, the soil moisture patterns (i.e., the microstates) were positively correlated with the local macroporosity on the plot that caused fast infiltration and transport associated with preferential flow. If, in contrast, soil moisture and flow patterns were negatively correlated, slow matrix flow and transport dominated. Information about this type of correlation is critically important for tracer transport but is rarely observable in the field. It is interesting that similar correlations between forcing and states are present in other parts of the hydro-
logical cycle and at larger scales. For example, Woods and Sivapalan [1999] noted that the spatial correlations of storm rainfall and soil saturation might substantially increase storm runoff from a catchment. Clearly, there is an interplay of mechanisms that may or may not be observable at the scale of interest.

[39] In the case of rainfall-runoff response observed in the Tannhausen catchment, the dynamics of the soil cracks are the main mechanism of threshold behavior which cause a marked seasonality in the runoff response [Lindenaier et al., 2006] that has also been observed at other sites [e.g., Návar et al., 2002]. As suggested by Figure 9 (upper left panel), the transition range of the system is narrow. The catchment switches from a fully developed crack system to almost no cracks in the range of 0.25–0.35 m m−3 average catchment scale soil moisture. The catchment scale soil moisture in the top 30 cm of the soil (simulated with a physically based catchment model) is a useful predictor of the average rainfall-runoff response in the Tannhausen catchment within the stable ranges.

[40] These three examples illustrate that threshold dynamics can indeed have a very significant effect on runoff response and tracer transport. There are, of course, numerous other threshold systems in hydrology that are likely to show a similar behavior. One example is the switch between vertical and lateral flow processes as a function of soil moisture observed in a small Australian catchment [Grayson et al., 1997], and other examples are given in the works of McGrath et al. [2006], Rundle et al. [2006], and Pitman and Stoffel [2006].

[41] Part of the observed scatter, one could argue, may be due to the limited number of experimental trials for the three examples. A statistical model was hence used to explore the system response one would obtain if a much larger number of experimental trials for each soil moisture macrostate were available. In other words, the model was used to analyze the asymptotic pattern of repeatability and predictability when increasing the number of observations, which is normally not possible in a field study. The model results suggest that, in all three cases, the inherent uncertainty in initial soil moisture combined with the threshold dynamics of the system results in predictability patterns similar to those observed.

[42] Although the physical and chemical mechanisms of the processes at the three sites are different, the predictability patterns are remarkably similar. In all three cases, the scatter is largest and hence the predictability is smallest when the system state is close to the threshold. As the system state moves away from it, the scatter gets smaller and the predictability increases. In all three cases, the changes in predictability are related to the interplay between the inherently uncertain initial state of the system and the threshold dynamics. Hence there is inherent uncertainty in the response data that is not measurement error but is related to the observability of the initial conditions. To some degree, it may be possible to increase the predictability by obtaining more accurate soil moisture estimates, but there is a limit to what can be obtained in the field. The macrostate descriptions chosen here were simple but typical of the amount of information available at experimental field sites. In a related study, Zehe and Blöschl [2004] used CATFLOW to simulate repeated trials of runoff observations in the 3.6-km² Weiherbach catchment for apparently identical initial soil moisture conditions based on data from 61 TDR stations. They found similar patterns of scaled ranges of peak discharges indicating that this type of unstable behavior may exist over a wide range of spatial scales.

[43] It is also interesting that the simple statistical model captures the predictability patterns well, although it has been formulated at a much more general level of conceptualization than the process model of, say, the Weiherbach site. The model could be used to explore the inherent uncertainty in observed system response to assess the limits of predictability of distributed hydrological models. It is straightforward to extend the statistical model, for example, by incorporating skewed probability distribution functions if needed. Whatever distribution is used, it is clear that predictive uncertainty may not only result from less than perfect model structures and model parameters but may be an inherent characteristic of the system under study. The range of system states for model calibration and validation hence needs to be carefully chosen both to allow for inherent uncertainty in any unstable range and to capture different regimes of the hydrological system by the model. In a similar vein, the statistical model could be used to trace the limits of confidence for planned field experiments to better understand the sources of observed scatter. A small number of trials may suffice to find the parameters of the model that could then be used to mimic the measurement process. Clearly, careful planning of field campaigns and judicious interpretation of the data are necessary if the response function strongly depends on the system state, as in the cases shown in this paper.

References


Zehe, E., and H. Flühler (2001a), Preferential transport of isoproturon at a plot scale and a field scale tile-drained site, *J. Hydrol.*, 247, 100–115.


G. Blöschl, Institute of Hydraulic and Water Resources Engineering, Vienna University of Technology, A-1040, Vienna, Austria. H. Elsenbeer, F. Lindennmaier, and E. Zehe, Institute of Geocology, University of Potsdam, 14415 Potsdam, Germany. (zehe@iws.uni-stuttgart.de)

K. Schulz, UFZ-Center for Environmental Research, Leipzig, Germany.