Hydrology and Earth System Sciences

Open Access

Prediction of dissolved reactive phosphorus losses from small agricultural catchments: calibration and validation of a parsimonious model

C. Hahn[1,2], **V. Prasuhn**[2], **C. Stamm**[3], **P. Lazzarotto**[1], **M. W. H. Evangelou**[1], **and R. Schulin**[1]

[1]ETH Zurich, Department of Environmental Systems Science, Zurich, Switzerland
[2]Agroscope Reckenholz-Tänikon Research Station ART, Zurich, Switzerland
[3]Eawag Swiss Federal Institute of Aquatic Science and Technology, Zurich, Switzerland

Correspondence to: C. Hahn (claudia.hahn@env.ethz.ch)

Abstract. Eutrophication of surface waters due to diffuse phosphorus (P) losses continues to be a severe water quality problem worldwide, causing the loss of ecosystem functions of the respective water bodies. Phosphorus in runoff often originates from a small fraction of a catchment only. Targeting mitigation measures to these critical source areas (CSAs) is expected to be most efficient and cost-effective, but requires suitable tools.

Here we investigated the capability of the parsimonious Rainfall-Runoff-Phosphorus (RRP) model to identify CSAs in grassland-dominated catchments based on readily available soil and topographic data. After simultaneous calibration on runoff data from four small hilly catchments on the Swiss Plateau, the model was validated on a different catchment in the same region without further calibration. The RRP model adequately simulated the discharge and dissolved reactive P (DRP) export from the validation catchment. Sensitivity analysis showed that the model predictions were robust with respect to the classification of soils into "poorly drained" and "well drained", based on the available soil map. Comparing spatial hydrological model predictions with field data from the validation catchment provided further evidence that the assumptions underlying the model are valid and that the model adequately accounts for the dominant P export processes in the target region. Thus, the parsimonious RRP model is a valuable tool that can be used to determine CSAs. Despite the considerable predictive uncertainty regarding the spatial extent of CSAs, the RRP can provide guidance for the implementation of mitigation measures. The model helps to identify those parts of a catchment where high DRP losses are expected or can be excluded with high confidence. Legacy P was predicted to be the dominant source for DRP losses and thus, in combination with hydrologic active areas, a high risk for water quality.

1 Introduction

Eutrophication of surface waters due to diffuse phosphorus (P) inputs continues to be a severe water quality problem worldwide (Carpenter et al., 1998; Kleinman et al., 2011b), causing, for example, algal blooms, oxygen shortage, fish death and loss of water bodies for recreation and drinking. It has been observed that the majority of P found in the runoff at the outlet of a catchment may originate from a small fraction of the catchment only (Gburek and Sharpley, 1998; Pionke et al., 2000, 1997). Thus, targeting mitigation options to these critical source areas (CSAs) is seen to be particularly efficient and cost-effective (Heathwaite et al., 2003; Schulte et al., 2009; Strauss et al., 2007; White et al., 2009). Critical source areas are characterized by a direct transport connection of available P sources to a receiving water body (Gburek and Sharpley, 1998). Originally, erosion and surface runoff were assumed to be the only relevant transport mechanisms, but now it is recognized that also subsurface flow can significantly contribute to P export (Doody et al., 2012; Kleinman et al., 2007, 2011b; Stamm et al., 2002; Watson and Matthews, 2008). Important sources of such P exports

are (1) freshly applied fertilizers or manure (Shigaki et al., 2007; Smith et al., 2001; Vadas et al., 2011), and (2) soils that are enriched with P due to excessive fertilizer application in the past (Kleinman et al., 2011a; Vadas et al., 2005). To a much smaller extent, also plants can contribute that are freshly grazed, trampled or in decay (Kleinman et al., 2011b). Runoff from locations with freshly applied manure or high soil P concentrations bear particularly high risks for P export. Buda et al. (2009) demonstrated that even sites with relatively low soil P concentrations can deliver very high P loads when runoff is large. However, due to the complexity of the processes controlling diffuse P losses, the identification of CSAs is still difficult (Doody et al., 2012; Kleinman et al., 2011a, b).

Various tools exist to describe water and P transport from non-point sources and to identify CSAs (Radcliffe et al., 2009; Schoumans et al., 2009; Sharpley et al., 2003), ranging from site assessment tools such as the P index (Weld and Sharpley, 2007) to process-based dynamic models such as SWAT (Arnold et al., 1998), INCA-P (Wade et al., 2002), and ANSWERS-2000 (Beasley et al., 1980). While static models are not able to account for the temporal and spatial variability of runoff and P losses, spatially distributed dynamic models are often over-parameterized (Radcliffe et al., 2009) and require many input data that are often not available. Therefore, as pointed out by Radcliffe et al. (2009), there is a need for parsimonious models that can be used to assess the spatial distribution of P export risks in a catchment. Irrespective of which type of model is used, a model requires validation for the purpose for which it is used. A major problem in validating spatially localized predictions of P export from a catchment is that P export risks depend on processes that are subject to high local spatial variability and fluctuation in time.

A parsimonious model developed to predict runoff and P losses at the outlets of small agricultural catchments is the Rainfall-Runoff-Phosphorus (RRP) model (Lazzarotto, 2005; Lazzarotto et al., 2006). The RRP model is based on the concept of spatially distributed CSAs that vary in size with hydrological conditions. It describes the export of dissolved reactive phosphorus (DRP). This form is immediately available for algal uptake (Sharpley, 1993; Sharpley et al., 1994) and thus has a direct impact on eutrophication (Kleinman et al., 2011b). The RRP model gave a good description of discharge and DRP losses at the outlet of experimental catchments (Lazzarotto, 2005; Lazzarotto et al., 2006).

In the model it is assumed that two sites with the same topographic position belonging to the same soil type behave the same. In order to keep the number of model parameters low, the model only distinguishes between two soil types (i.e., well and poorly drained soils). This allowed for parameterizing the soil types by simultaneously calibrating the model to four catchments of different soil composition (Lazzarotto, 2005; Lazzarotto et al., 2006). Accordingly, the model should be transferable to other sites without calibration if the topographic and soil information is

available. Because the moisture regime is a continuum, assigning the soils to these two classes may be somewhat arbitrary in some cases.

In this study we investigated the validity of RRP model predictions and in particular their sensitivity on the binary classification of soils by water regime classes. First, we calibrated the model simultaneously on runoff data of four small catchments in an agricultural area of Switzerland and then used it to predict runoff and P export from a neighboring catchment. Aside from testing the validity of these model predictions, we investigated the sensitivity of the model predictions on the soil grouping and assessed the spatial performance of various model versions using field data on soil moisture, groundwater table, runoff volumes and P concentrations in runoff.

2   Materials and methods

2.1   The Rainfall-Runoff-Phosphorus (RRP) model

The Rainfall-Runoff-Phosphorus (RRP) model is a parsimonious model for continuous simulations of DRP transport from intensively managed grassland soils into streams in small agricultural catchments. It consists of two sub-models: the semi-distributed rainfall-runoff model and the phosphorus (P) model.

2.1.1   Rainfall-runoff sub-model

The rainfall-runoff sub-model is a soil-type-based semi-distributed model (Lazzarotto et al., 2006). It is based on the assumptions that (1) areas with the same topographic index λ and class of soil have the same hydrological behavior, and that (2) soils can be divided into two classes (i.e., well and poorly drained soils) having the same hydrologic characteristics within each class. The topographic index λ (Beven and Kirkby, 1979; Kirkby, 1975) is defined as

$$\lambda = \ln(A_{\text{upstream}}/\tan\beta), \tag{1}$$

where A_{upstream} is the upslope area draining through the respective location (multiple flow direction algorithm of Quinn et al. (1991)) and β is the local slope at that location. It is an indicator for the wetness of the soil at a given location within the catchment. Catchments are divided into four types of hydrological response units (HRUs) differing in runoff dynamics: well-drained soils (HRU$_1$), poorly drained soils (HRU$_2$), urban areas (HRU$_3$), and forests (HRU$_4$). Soil moisture is assumed to be uniform within each HRU. Changes in water storage S_i in HRU$_i$ are calculated in hourly time steps (Δt) from the mass balance equation:

$$S_i(t + \Delta t) = S_i(t) + [\text{rain}(t) - \text{ET}(t) - \text{runoff}_i(t)]\Delta t, \tag{2}$$

where rain(t) [mm], ET(t) [mm] and runoff$_i(t)$ [mm] are the respective rates of rainfall, evapotranspiration and simulated

runoff from HRU_i during the time interval Δt. For HRU_1 and HRU_2 the model considers two types of runoff: fast flow $q_{i,\text{fast}}$ [mm] and slow flow $q_{i,\text{slow}}$ [mm]. The slow flow component, which is given by

$$q_{i,\text{slow}}(t) = \Theta_i(t)c_i, \tag{3}$$

depends (i) on the parameter c_i [mm] determining how much water from HRU_i contributes to baseflow and (ii) on the degree of soil saturation $\Theta_i(t)$ [-], which is defined as the ratio between soil water storage $S_i(t)$ [mm] and the maximum soil water storage capacity $S_{i,\text{max}}$ [mm]:

$$\Theta_i(t) = \frac{S_i(t)}{S_{i,\text{max}}}. \tag{4}$$

The fast flow component includes all types of quickly responding flow, such as preferential flow, saturation excess and Hortonian overland flow. It is the sum of an autoregressive part describing the recession of fast flow and of a part representing the fraction of rain directly converted into fast flow:

$$q_{i,\text{fast}}(t + \Delta t) = a_i q_{i,\text{fast}}(t) + b_i \text{rain}(t - \text{dt}_i) \frac{A_{i,\text{fast}}(t)}{A_i}. \tag{5}$$

The parameter a_i [–] is the fast flow decline rate, b_i [–] the proportion of rain that is directly converted into fast flow, dt_i [h] the time delay between rainfall and runoff in HRU_i, and $A_{i,\text{fast}}A_i^{-1}$ the areal fraction of HRU_i that contributes to fast flow. The latter depends on the soil moisture status at time t. For every time step a threshold value $\lambda_{0,i}(t)$ [–] is determined for the topographic index λ of HRU_i:

$$\lambda_{0,i}(t) \propto (1 - \Theta_i(t))^{n_i}. \tag{6}$$

Locations with a topographic index higher than this threshold value are attributed to $A_{i,\text{fast}}$ [m²]. The parameter n_i [–] is determined by calibration. In contrast to HRU_1 and HRU_2, all runoff is assumed to occur as fast flow in urban areas (HRU_3):

$$q_{3,\text{fast}}(t + \Delta t) = a_3 q_{3,\text{fast}}(t) + b_3 \text{rain}(t - \text{dt}_3). \tag{7}$$

The total catchment response results from the sum of all flow components weighted with their respective areal fractions A_i/A_{total}, with $A_{\text{total}} = \sum A_i$. Neglecting runoff from forest areas due to their limited size in the study catchments, this sum was

$$Q(t) = (q_{1,\text{slow}}(t) + q_{1,\text{fast}}(t)) \frac{A_1}{A_{\text{total}}}$$
$$+ (q_{2,\text{slow}}(t) + q_{2,\text{fast}}(t)) \frac{A_2}{A_{\text{total}}} + q_{3,\text{fast}}(t) \frac{A_3}{A_{\text{total}}} \tag{8}$$

in our case.

2.1.2 Calibration of the rainfall-runoff sub-model

Using uniform Monte Carlo simulations, the soil parameters (Table 1) were determined by simultaneous calibration of the model on four catchments (see Sect. 2.2) that differed in their soil composition and their hydrological response (Lazzarotto et al., 2006). The calibration period extended from 7 to 17 July 2000. This short calibration period proved to be sufficient (Lazzarotto et al., 2006), as conditions varied between very wet and dry. Different parameter combinations were generated using random sampling within the domain of each parameter. The following Nash–Sutcliffe criterion (NSC) (Nash and Sutcliffe, 1970), calculated for the four catchments together, was used to assess model performance of each parameter combination:

$$\text{NSC} = 1 - \frac{\sum_{k=1}^{4} \sum_{t=t_0}^{t_e} (Q_{\text{obs}}^k(t) - Q_{\text{sim}}^k(t))^2}{\sum_{k=1}^{4} \sum_{t=t_0}^{t_e} (Q_{\text{obs}}^k(t) - \bar{Q}_{\text{obs}}^k)^2}, \tag{9}$$

where $Q_{\text{obs}}(t)$ [mm] is the observed runoff at time t, $Q_{\text{sim}}(t)$ [mm] the simulated runoff at time t, and $\bar{Q}_{\text{obs}}$ [mm] the mean observed runoff for the whole time period in catchment k. The evaluated parameter sets were classified as either "behavioral" (or "accepted") for $\text{NSC} > \text{NSC}_{\text{threshold}}$ or "nonbehavioral" for $\text{NSC} < \text{NSC}_{\text{threshold}}$ (Hornberger and Spear, 1981). Behavioral parameter sets were used for model application. Thus, the number of accepted parameter sets (mc) defines the number of simulation results. The 10 % quantiles and 90 % quantiles of these simulations were used to characterize the uncertainty of the model predictions.

For more information on the hydrological model, the reader is referred to Lazzarotto et al. (2006). Here, we converted the model from FORTRAN77 to FORTRAN95 in order to make a few modifications (such as corrections of some coding errors and removal of parameter constraints). We will refer to this version of this model in which all soil parameters were calibrated simultaneously as Version 1. In a second model version (Version 2) the urban parameters a_3 and b_3 were separately calibrated using discharge data from six small runoff events in July 2010 recorded in the Stägbach catchment, which is located in the vicinity of the calibration catchments (see Sect. 2.2). As soil moisture was low prior to these six events, runoff from agricultural land could be neglected. The resulting parameter values were $a_3 = 0.0968$ and $b_3 = 0.0894$. The third model version (Version 3) was identical to Version 2, but used a different soil classification (see Sect. 2.3.1).

For each of the three model versions, more than 500 accepted parameter sets were determined. For each of these sets, a prediction of runoff was calculated for a given catchment and time period and fed into the P sub-model to calculate a prediction of P export using the P sub-model (see Sect. 2.1.3).

www.hydrol-earth-syst-sci.net/17/3679/2013/

Hydrol. Earth Syst. Sci., 17, 3679–3693, 2013

Table 1. Parameters of the hydrological response units (HRU$_i$ = 1, 2, 3) that need to be determined during calibration (adopted from Table 2 in Lazzarotto et al., 2006). HRU$_1$ = well drained, HRU$_2$ = poorly drained, HRU$_3$ = urban.

Global parameter	Minimum value	Maximum value	Property	Used HRU
$S_{i,\mathrm{max}}$ [mm]	0	800	Maximum soil water storage capacity	$i = 1, 2$
a_i [–]	0	1	Fast flow decline rate	$i = 1, 2, 3$
b_i [–]	0	1	Proportion of rainfall converted into fast flow on the contributing areas	$i = 1, 2, 3$
c_i [mm]	0	1	Flow rate between the scaled soil water storage and the slow flow components	$i = 1, 2$
n_i [–]	1	10	Expansion control of areas contributing to fast flow	$i = 1, 2$

2.1.3 The phosphorus model

The Phosphorus (P) sub-model was developed to predict DRP losses at catchment outlets and CSAs within catchments in combination with the rainfall-runoff sub-model (Lazzarotto, 2005). The model was developed for the Lippenrütibach catchment, a catchment on the Swiss Plateau, which was also used for calibration of the hydrological sub-model. Previous studies in the study region had shown that DRP concentrations in runoff were strongly correlated with runoff volume (Lazzarotto et al., 2005; Pacini and Gächter, 1999; Stamm et al., 1998), indicating that high rates of P losses were associated with fast runoff. To account for the elevated P concentrations of fast runoff as compared to slow runoff, fast flow is assumed to be composed of "old" and "new" water, while slow flow is assumed to consist of "old" water only. While $q_{i,\mathrm{slow}}(t)$ and $q_{i,\mathrm{fast}}(t)$ are average values that apply to all cells within an HRU$_i$, the P sub-model distinguishes between grid cells within the respective HRU$_i$ that actually contribute to fast flow in a given event and cells that do not, assuming that total fast flow is equally distributed among the cells that contribute. Thus, for cells that contribute fast flow $q_{i,\mathrm{fast}}(t, x, y)$ is calculated by dividing $q_{i,\mathrm{fast}}(t)$ by the areal fraction $(A_{i,\mathrm{fast}}(t)/A_i)$ of HRU$_i$ that contributes to fast flow, while fast flow $q_{i,\mathrm{fast}}(t, x, y)$ from cells that are not contributing is zero. The new water component, $q_{i,\mathrm{new}}(t, x, y)$ [mm], is assumed to be a constant fraction η [–] of the total fast flow from the contributing area:

$$q_{i,\mathrm{new}}(t, x, y) = \eta\, q_{i,\mathrm{fast}}(t) \frac{A_i}{A_{i,\mathrm{fast}}(t)} F(t, x, y), \qquad (10)$$

where $F(t, x, y)$ is 0 for cells not contributing to fast flow, and 1 for cells contributing to fast flow at time t, and x and y are the central coordinates of the respective cell. The fraction η was estimated from nitrate dilution data collected during runoff events and baseflow conditions as 0.25 ± 0.05 (Lazzarotto, 2005). The flow of old water [mm] is the sum of the remaining fast flow and the slow flow of the respective cell:

$$q_{i,\mathrm{old}}(t, x, y) = (1 - \eta)\, q_{i,\mathrm{fast}}(t) \frac{A_i}{A_{i,\mathrm{fast}}(t)} F(t, x, y) + q_{i,\mathrm{slow}}(t). \quad (11)$$

The DRP loss with old water flow is calculated for every grid cell as

$$L_{i,\mathrm{old}}(t, x, y) = \mathrm{DRP}_{\mathrm{baseflow}} q_{i,\mathrm{old}}(t, x, y) \mathrm{gridsize}, \qquad (12)$$

assuming that the concentration of DRP in old water is the same as the DRP concentration of the baseflow, $\mathrm{DRP}_{\mathrm{baseflow}}$ ($0.05\,\mathrm{mg\,L^{-1}}$). DRP losses associated with new water flow include incidental P losses from freshly applied manure ($\mathrm{DRP}_{\mathrm{IPL}}$ [mg L^{-1}]) and P losses from soil ($\mathrm{DRP}_{\mathrm{soil}}$ [mg L^{-1}]) enriched in P due to excessive manure applications in the past. $\mathrm{DRP}_{\mathrm{soil}}$ concentrations were calculated for every pixel from water-soluble soil P (WSP) concentrations. The WSP–DRP relationship was taken from artificial rainfall experiments carried out in the catchment area of Lake Baldegg (Hahn et al., 2012).

$$\mathrm{DRP}[\mathrm{mgL^{-1}}] = 0.0852\,\mathrm{WSP}[\mathrm{mgkg^{-1}}] - 0.3039, \qquad (13)$$

with the condition that no negative DRP values can occur. The WSP concentrations (and thus also the $\mathrm{DRP}_{\mathrm{soil}}$ concentrations) were assumed to remain constant over the simulation period in the present study.

In contrast, $\mathrm{DRP}_{\mathrm{IPL}}(t, x, y)$ concentrations in runoff were considered to vary in time. Based on the studies of Braun et al. (1993) and Von Albertini et al. (1993), $\mathrm{DRP}_{\mathrm{IPL}}(t, x, y)$ is assumed to decrease exponentially with increasing time lag $\Delta t_{\mathrm{m}} = t_{\mathrm{r}} - t_{\mathrm{a}}$ between manure application t_{a} and onset of runoff t_{r}:

$$\mathrm{DRP}_{\mathrm{IPL}}(t, x, y) = \mathrm{DRP}^0_{\mathrm{IPL}}(t, x, y) \exp(-\Delta t_{\mathrm{m}} h) \qquad (14)$$

The time t of runoff onset is the time when the respective soil pixel starts to contribute to fast flow (Lazzarotto, 2005). The parameter h was assumed to be the same for well and poorly drained soils: 0.007 ± 0.004. The value range of parameter h was determined preliminary, by fitting the $\mathrm{DRP}_{\mathrm{IPL}}$ function to observed data (Lazzarotto, 2005). With each application of manure Δt_{m} is set to zero, and the $\mathrm{DRP}_{\mathrm{IPL}}(t, x, y)$ concentration is increased immediately to the new value of $\mathrm{DRP}^0_{\mathrm{IPL}}(t, x, y)$ resulting from the addition of the new DRP to the $\mathrm{DRP}_{\mathrm{IPL}}$ remaining from the prior applications.

The total DRP load [mg] associated with new water is the sum of DRP_{soil} and DRP_{IPL} loss at each pixel:

$$L_{i,new}(t, x, y) = (DRP_{soil}(x, y)$$
$$+ DRP_{IPL}(t, x, y)) q_{i,new}(t, x, y) \text{gridsize} \qquad (15)$$

while the total DRP loss from a pixel at time t is the sum of $L_{i,old}(t, x, y) + L_{i,new}(t, x, y)$, and the total loss of DRP from the catchment is the sum of DRP loss from all soil pixels.

We used Gaussian error propagation to account for uncertainty in the model parameters η and h and in the WSP–DRP relationship. Thus, for each mc model run and time step we obtained an error estimate. These were combined with the 10 % and 90 % quantiles of the hydrological predictions to give the uncertainty of the DRP export predictions.

2.2 Study area

The study area was situated on the Swiss Plateau in the vicinity of Lucerne. It is characterized by undulating terrain, ranging between 500 and 800 m altitude above sea level and covered by glacial tills (Lazzarotto et al., 2006). The soils are generally loamy and of low permeability (Peyer et al., 1983). Average amounts of annual precipitation in the region range between 1000 and 1200 mm, depending primarily on altitude.

The four catchments used for model calibration (Lippenrütibach (LIP), Greuelbach (GRB), Rotbach (RTB), Meienbach (MEI)) drain into Lake Sempach (Lazzarotto, 2005), whereas the catchment (Stägbach catchment (Stäg)) used for model validation drains into Lake Baldegg (Fig. 1). Both lakes have serious eutrophication problems and are artificially aerated. The region is characterized by intensive animal husbandry (dairy and pig farms, 2.4 livestock units per ha (Herzog, 2005)), which in the past has resulted in highly increased soil P stocks (Stamm et al., 1998).

In addition to the Stägbach catchment as a whole, we also used a sub-catchment of the Stägbach catchment, denoted as Stäg2, for validation (Fig. 1). Table 2 shows that the percentages of urban area, forest, and agricultural area in the validation catchments were in the range of the calibration catchments. Agriculture is the dominating land use in all catchments, whereas the area classified as urban covered less than 10 %. The latter consisted of a few villages and some isolated farms. While the Stäg2 sub-catchment was comparable in size to the calibration catchments, the Stägbach catchment as a whole (8.24 km^2) was larger than all four calibration catchments. More information on the calibration catchments is given by Lazzarotto et al. (2006). Small differences between the HRU percentiles given here and those by Lazzarotto et al. (2006) are due to the fact that the data had to be processed anew.
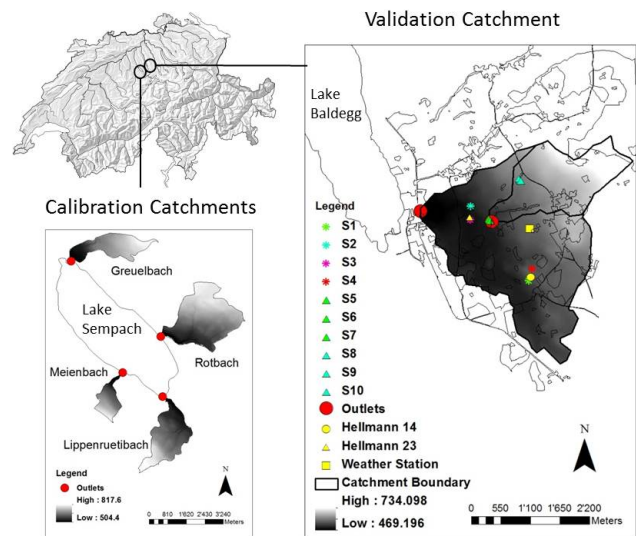

Fig. 1. Locations of calibration and validation catchments and the installed measurement devices.

2.3 Model validation

2.3.1 Model input data

Precipitation and evapotranspiration

From April till October 2010 a weather station was installed in the center of the Stägbach catchment to obtain representative precipitation data for the Stägbach catchment. The station was equipped with a R102/R102H tipping bucket rain gauge. Data were recorded every 15 min. For two short time periods (28 May–8 June 2010 and 21 July–1 August 2010), no data were recorded at this weather station, due to technical problems. For these periods we used precipitation data from the nearest weather station (Hochdorf, data from uwe in the canton of Lucerne), which is located less than 2 km away from the Stägbach catchment. All other data gaps were filled with mean precipitation data from the three closest weather stations (Buchs, Lucerne, Cham, provided by the Swiss Federal Office of meteorology and Climatology) surrounding the catchment. Two Hellmann rain gauges were installed in the catchment to check for spatial variability in rainfall. For the global radiation data, we used evapotranspiration data from the three weather stations Buchs, Lucerne and Cham (provided by the Swiss Federal Office of meteorology and Climatology). These data are based on the Primault formula. They were available at daily resolution, but using mean global radiation data from the same three MeteoSwiss stations, we derived estimates of hourly evapotranspiration.

Topographic index and HRU determination

The topographic index λ (Quinn et al., 1991) was determined on a 25 m resolution digital elevation model (DEM),

Table 2. Areal fraction of each hydrological response unit (HRU) on the total catchment area in [%] – roman font = model Version 2, bold font = model Version 3.

	Calibration catchments								Validation catchments			
	LIP	**LIP**	RTB	**RTB**	MEI	**MEI**	GRB	**GRB**	Stäg	**Stäg**	Stäg2	**Stäg2**
Urban [%]	8.1	**8.1**	9	**8.7**	2.5	**2.5**	8	**7.6**	9	**9**	6	**6**
Forest	16.7	**16.7**	16	**16.4**	7.7	**7.7**	16	**16.3**	8	**8**	9.5	**9.5**
Well	38.6	**13.1**	56	**31.5**	74	**40.6**	61	**34.6**	66	**42**	67.2	**41.2**
Poor	36.6	**62.1**	19	**43.4**	15.8	**49.2**	15	**41.5**	17	**41**	17.3	**43.3**
Area [km^2]	3.3		6.0		1.2		2.6		8.24		2.27	

which is available for the whole of Switzerland (provided by the Swiss Federal Office of Topography), using the open source GIS software Saga 2.0. The convergence coefficient was set to 1. Urban areas and forests were identified using aerial photographs (provided by the Swiss Federal Office of Topography). The data were processed and prepared for model input using ArcGIS (ArcGIS Desktop 10 Service Pack 2, ESRI) and the software package R (RDevelopment-CoreTeam, 2007).

Soil classification into drainage classes

The assignment of soils to the two classes of well and poorly drained soils was based on the local soil map (Peyer et al., 1983). In model Version 1 and 2 we followed Lazzarotto et al. (2006), who classified Eutric and Dystric Cambisols and Eutric Regosols as well-drained soils and Gleyic Cambisols and Eutric Gleysols as poorly drained soils. To investigate the sensitivity of the model to this classification, we compared Version 2 with Version 3. In the latter we also assigned soils considered well drained by Lazzarotto et al. (2006), although showing signs of temporary water stagnation or water-logging according to the soil map, to the poorly drained soils. Accordingly the areal fraction occupied by the poorly drained HRU was larger in Version 3 than in Version 1 and 2 (Table 2).

Soil P status and manure application

A map of the spatial distribution of soil P concentrations (see Fig. S1) was constructed from data of soil P analyses farmers have to provide to local authorities every 5 yr. With the help of the farmers, the available data on soil P status were assigned to individual fields. Some farmers did not cooperate. In these cases we used P data obtained from the environmental protection agency of the canton of Lucerne and attributed area-weighted mean P values to the respective management units.

Some farmers also provided detailed data on the amounts, locations and times of manure application on their farms. For the other farms, which covered more than 80 % of the area, the manure P pool was neglected. In contrast, manure application data were complete for the Lippenrütibach catchment, one of the calibration catchments, in the year 1999.

2.3.2 Model validation

Discharge measurements

At the outlet of the Stägbach catchment, a 6712 Full-size Portable Sampler (ISCO, USA) was used to determine discharge and collect water samples. In addition, the water level was recorded every minute by means of a Bubbler Flow Module. Further flow and water level measurements (dilution method) were taken by a consulting company (Büro für Wasser und Umwelt, BWU) working for the cantonal environmental protection agency. They provided us also with the level–discharge data necessary to calculate the discharge from the level data.

The discharge at the outlet of Stäg2 was estimated from the discharge at the outlet of the entire catchment using a relationship that was determined on the basis of eight manual measurements of flow velocity profiles and water levels at the outlet of Stäg2 between the beginning of June and the end of July 2010, using a current meter (MiniAir2) and a measuring rod. From these measurements we calculated the discharge across the entire flow profile for each of these eight occasions and related it to the discharge from the entire catchment. The discharge estimates for Stäg2 based on this relationship were validated by measurements with a 6712 Full-size Portable Sampler equipped with a 750 Area Velocity Module (ISCO, USA) installed at the outlet of sub-catchment Stäg2. Unfortunately, no continuous automatic measurements were available because the instruments were dislocated during the large rain event in June 2010 and partly damaged. Discharge estimations based on the Stäg catchment were very similar to discharge values deduced from the relationship between the manual discharge measurements and the automatic flow velocity data. Only during the high runoff event at the end of July and afterwards, the two graphs differed. This period was therefore not taken into account for model assessment.

Water samples

Using the aforementioned 6712 Full-size Portable Sampler, flow-proportional water samples were collected automatically at the outlet of the Stägbach catchment and the Stäg2 sub-catchment. A pre-defined water level (Stäg) or flow velocity (Stäg2 sub-catchment) threshold was set, and when it was reached, samples were taken automatically every 15 min. Four subsequent samples were collected in the same bottle, resulting in one composite sample every hour, as long as the water level (or the velocity, respectively) was above the threshold. After a runoff event, samples were collected and stored at 4 °C till analysis. In addition, we took grab samples each time we went into the field, at least once a week. Dissolved reactive phosphorus (DRP) was analyzed by means of the molybdate colorimetry method (Vogler, 1965) after filtration (< 450 nm) of sample solution. In order to determine total phosphorus (TP), unfiltered samples were digested in potassium persulfate before they were analyzed for P using the molybdate colorimetry method. Electrical conductivity (EC) was measured using a Metrohm Conductometer 712.

Soil moisture measurements

On four grassland sites (Table 3) soil water content was monitored at 10 and 30 cm depth using six horizontally inserted 2-rod TDR probes at each depth. The signal was recorded by means of a TDR100 and stored by a data logger (CR10X, Campbell Scientific, Inc.). The volumetric soil water content ($m^3 m^{-3}$) was calculated using the equation given by Topp et al. (1980). Volumetric soil samples were taken at each of the four soil water monitoring locations using steel cylinders to determine soil bulk density and porosity.

Piezometer and overland flow detectors

Furthermore, we installed a piezometer equipped with a light plummet and an overland flow detector (OFD; see Doppler et al., 2012) at each soil water monitoring station and 6 other locations. Readings of these instruments were taken approximately once a week normally and more often after rainfall events.

3 Results

3.1 Model performance at the catchment outlet

3.1.1 The rainfall-runoff model

Model calibration with data from the year 2000

Without separate calibration of the urban parameters (Version 1), the model performed poorly. Out of 7 million Monte Carlo (MC) simulations, no parameter set achieved a NSC value > 0.5; 661 parameter sets yielded a NSC > 0.4. Separate calibration of the urban parameters (Version 2) improved the model results substantially and resulted in 724 accepted parameter sets from 5 million MC runs when the threshold value was set to 0.6, with 25 %, 50 %, and 75 % quantiles of 0.61, 0.61 and 0.63, respectively. Changing the classification of the soils (Version 3) decreased the performance for the calibration period, so that the NSC threshold had to be reduced to 0.5 to obtain 606 accepted parameter sets, with 25 %, 50 % and 75 % NSC quantiles of 0.51, 0.52 and 0.53 respectively.

## Comparison of predictions for the Lippenrütibach catchment

Before we applied the calibrated model to the Stägbach catchment, we compared hydrological predictions for the Lippenrütibach catchment (LIP), one of the calibration catchments, for the year 1999. The same data had been used for validation by Lazzarotto et al. (2006). Figure 2a shows a fair agreement between simulations (Version 2) and measurements. Predictions were again better for the model version with separate calibration of the urban HRU parameters a_3 and b_3 (Version 2) than for the corrected original version of the model (Version 1) (Table 4). This improvement was in particular due to better prediction of small peaks, which were overestimated by the original model (Lazzarotto et al., 2006). However, two other problems, which had already been identified by Lazzarotto et al. (2006), remained unsolved: (1) some high runoff peaks were still underestimated, and (2) baseflow declined too fast after long periods with no rainfall (Fig. 2b).

Model validation – Stägbach 2010

To test how well the model performs when applied outside the watersheds used for calibration, we applied the calibrated Version 2 model to the Stägbach catchment for a forward prediction of discharge during the year 2010 and compared predictions with measurements. Figures 3 and 4 show that the model performed well for the entire catchment as well as for the Stäg2 sub-catchment. The median NSC values were 0.62 and 0.72, respectively (Table 4). With the global parameter c_{well} ranging mainly between 0.7 and 0.92 (25 % and 75 % quantiles), and c_{poor} ranging between 0.33 and 0.61, more baseflow was predicted to come from the well-drained than from the poorly drained HRU. Thus, the relatively high amount of well-drained soils within the validation catchments as compared with the calibration catchments (Table 2) led to a baseflow overestimation in both validation catchments. Due to the general overestimation of baseflow, accelerated baseflow decline as observed for the Lippenrütibach catchment in 1999 was only observed during the very dry period in summer. The underestimation of discharge during the large event end of July ($h = 5000$) was probably due to the fact that no rain data were available from the Stägbach weather station for this event, while spatial variability of rainfall was very high in the study area, as indicated by the Hellmann rain gauges. One Hellmann rain gauge collected

Table 3. Site characteristics of the four permanent measurement stations in the Stägbach catchment.

	S1		S2		S3		S4	
HRU	well drained		well drained		poorly drained		well drained	
Topographic index	7.16		10.65		11.13		7.27	
Soil map	vertically permeable soil		vertically permeable soil		ground-/slope water-dominated soil		vertically permeable soil, partly ground- or slope water-influenced	
Soil	Calcaric Cambisol		Eutric Cambisol		Eutric Cambisol		Eutric Cambisol	
Texture (FAQ)	loam		sandy clay loam		loam		loam	
Soil depth	10 cm	30 cm	10 cm	30 cm	10 cm	30 cm	10 cm	30 cm
Clay [%]	20.94 ± 0.66	22.14 ± 1.03	25.63 ± 1.26	26.52 ± 2.11	25.25 ± 0.13	19.62 ± 0.35	17.80 ± 1.39	18.99 ± 2.27
Silt [%]	32.99 ± 1.24	38.98 ± 0.91	36.27 ± 3.41	39.78 ± 0.09	46.39 ± 0.82	44.03 ± 0.56	32.40 ± 0.47	35.03 ± 0.2
Sand [%]	46.07 ± 0.58	38.88 ± 0.12	38.10 ± 3.75	33.71 ± 2.02	28.36 ± 0.95	36.34 ± 0.21	49.80 ± 0.92	45.97 ± 2.07
pH	7.02	7.16	6.05	6.26	5.32	5.45	5.89	6.59
pore volume [%]	52	47	54	49	41	41	53	44
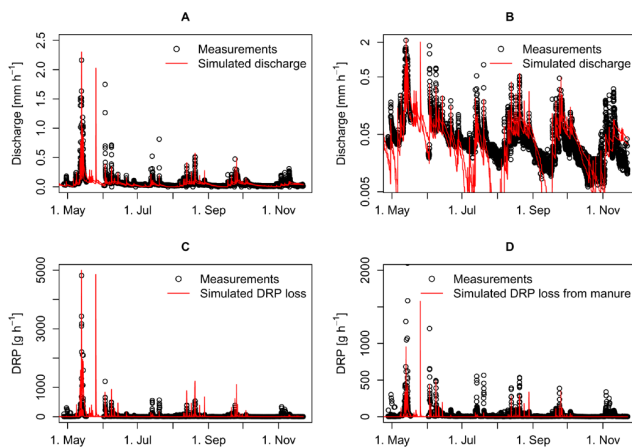
Fig. 2. Simulations (lines, 10 % and 90 % quantiles) using RRP Version 2 versus measured (points) discharge and DRP loss from the Lippenrütibach catchment in 1999. The y axes in figures on the right are in logarithmic scale (**B**) or focus on a certain part of the value range (**D**). (**C**) shows the total range of the DRP loads measured and simulated on a non-log scale.

126 mm rain while the other one only collected 88 mm within the same time frame.

Influence of soil classification – model Version 3

As in the calibration, Version 3 did not perform as well as Version 2 also in the validation for the Lippenrütibach catchment (Table 4), as runoff peaks were slightly lower in simulations with Version 3 than with Version 2. The higher value of c_{poor} and the higher areal percentage of HRU$_{poor}$ (62.1 %) resulted in higher slow flow from poorly drained soils, which led to lower soil moisture and thus to lower peak flows. In contrast, the change in soil classification from Version 2 to Version 3 improved model predictions for the entire Stägbach catchment and the Stäg2 sub-catchment (Table 4). The improvement was due to better simulations of baseflow and of the large runoff event in June. This can be attributed to
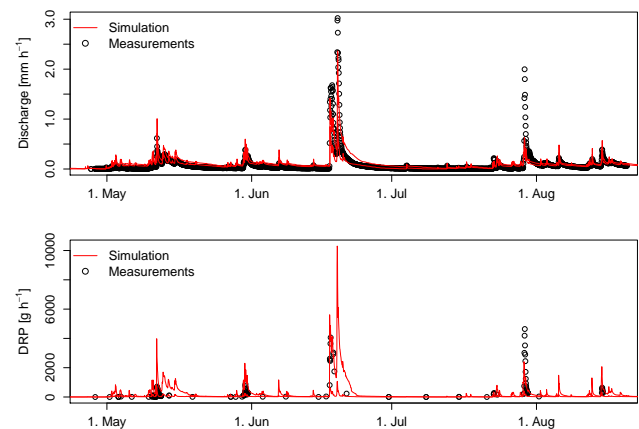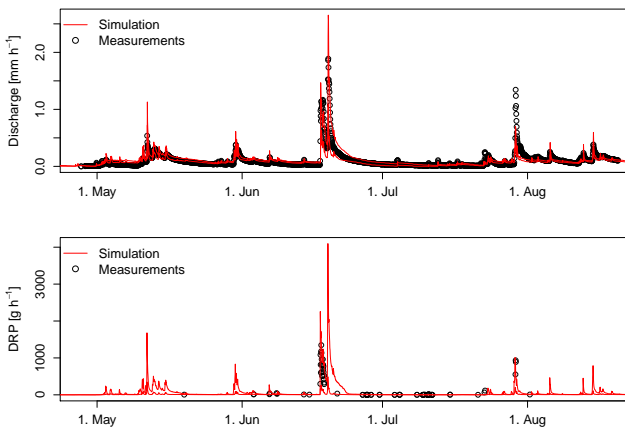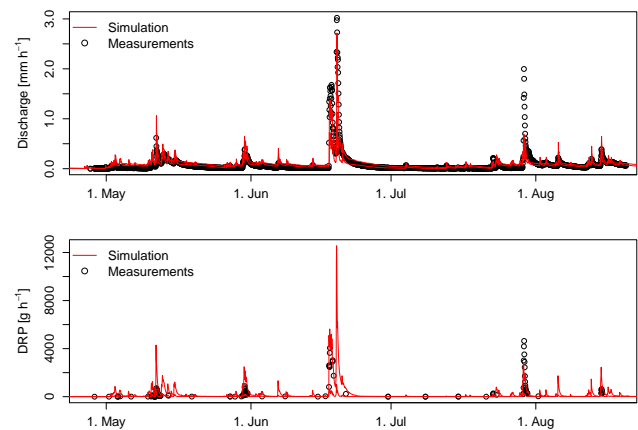

Fig. 3. Simulations (lines, 10 % and 90 % quantiles) using RRP Version 2 versus measured (points) discharge and DRP loss from the Stägbach catchment in 2010.

the lower value of c_{well} (25 % and 75 % quantiles: 0.31, 0.75) and the lower areal fraction of well-drained soils in Version 3 (Table 2), resulting in higher soil moisture and consequently also in higher peak flows. The larger area of poorly drained soils also led to steeper decline of the hydrographs (Fig. 5), due to a larger contribution of the poorly drained HRU to fast flow. Apart from these rather small differences, both versions of the model simulated the discharge dynamics of the study catchments quite well.

3.1.2 The phosphorus model

The simulated DRP losses for the Lippenrütibach catchment in the year 1999 and the Stägbach catchments in the year 2010 are in fair agreement with the measurements (Figs. 2c, 3, 4, Table S1). There was little difference between Version 2 and 3 of the model. The predictions of DRP loads mainly depended on runoff prediction. For example, DRP losses from the Lippenrütibach catchment were underpredicted for the events at the beginning of June, in July and November 1999,

Table 4. Performance of different model versions in three catchments (Lippenrütibach catchment (LIP), Stägbach catchment (Stäg), Stägbach sub-catchment (Stäg2)), measured with the Nash–Sutcliffe criterion (NSC) (Nash and Sutcliffe, 1970). Version 1 – corrected original model, Version 2 – separate urban parameter calibration, Version 3 – separate urban parameter calibration + different soil classification.

Model version	Lip			Stäg			Stäg2			Calibration
	NSC quantiles									
	25 %	50 %	75 %	25 %	50 %	75 %	25 %	50 %	75 %	$NSC_{threshold}$
Version 1	0.42	0.44	0.46	0.50	0.61	0.70	0.65	0.71	0.78	0.4
Version 2	0.48	0.50	0.52	0.53	0.62	0.71	0.66	0.72	0.80	0.6
Version 3	0.44	0.46	0.48	0.62	0.68	0.75	0.68	0.74	0.80	0.5

Fig. 4. Simulations (lines, 10 % and 90 % quantiles) using RRP Version 2 versus measured (points) discharge and DRP loss from the Stägbach sub-catchment (Stäg2) in 2010.

Fig. 5. Simulations (lines, 10 % and 90 % quantiles) using RRP Version 3 versus measured (points) discharge and DRP loss from the Stägbach catchment in 2010.

for which runoff was underestimated as well. On the other hand discharge and DRP load were well predicted for the large events in May 1999 (Lippenrütibach) and in June 2010 (Stägbach). As the simulated water fluxes match the observed water fluxes most of the time quite well, a good match of the observed DRP losses implies that also the DRP concentrations are matched well. Unfortunately, in the Stägbach catchments no samples were collected during the second peak of the extreme event because the sampling device was either clogged (Stäg) or dislocated (Stäg2).

The simulated loss of DRP from the Lippenrütibach catchment that was attributable to recently applied manure (Fig. 2d) was about 1/5 of the total DRP loss (Fig. 2c) during the large event in May, and less than half of the total simulated DRP load in most of the other events. Thus, most DRP lost with runoff came from the soils according to the model. In the Stägbach catchments, a good fit between simulations and measurements was obtained despite the limited availability of manure application data, suggesting again that soil P was the main source for the DRP losses with runoff.

3.2 Spatial model performance

3.2.1 Hydrological risk areas

For each time step, we constructed maps showing for each pixel the fraction of accepted parameter sets (out of a total of 724 accepted sets for Version 2 and 606 accepted sets for Version 3) that resulted in fast flow in that pixel at the respective time. These maps give a picture of the uncertainty in the prediction of fast flow at the specific time across the catchment for the respective model version. For simplicity, we refer to the fraction of accepted parameter sets predicting fast flow as "risks" of fast flow. This measure reflects how sensitive the fast flow prediction is towards changes of the parameter sets. We introduce four classes and denote values ranging between 0 and 0.2 a low risk, values between 0.2 and 0.5 a medium risk, values between 0.5 and 0.8 a high risk, and values between 0.8 and 1 a very high risk of fast flow.

The spatial extent of risk areas changes with time. For the small runoff event of 14 May 2010, 7 % of the agricultural area in the Stägbach catchment was classified as very high risk area, whereas for the large event in June (19 June 2010)

16 % (Version 2) or even 21 % (Version 3) of the agricultural area was very high risk area. Also the percentage of high and medium risk areas within the catchment increased during this event (Table 5). On the other hand, large fractions of the catchment were considered at low risk during the small event by both model versions (76 and 48 %, respectively). However, during the June event, model Version 3 predicted a low risk for only 13 % of the catchment, while this percentage was higher (44 %) for model Version 2. Hence, based on the model results one cannot exclude the risk for DRP losses from a considerable fraction of the area.

The spatial patterns of predicted fast flow risk areas were very similar for model versions 2 and 3 (Table 5, Fig. 6). The major difference was that the medium risk was more prevalent and the low risk class less frequent in Version 3 than in Version 2. This can be attributed to the lower overall runoff in Version 3 simulations, which led to higher soil moisture predictions and thus lower topographical threshold values.

3.2.2 Spatial predictions of DRP losses from soil

In the RRP model, the risk of P loss depends on the combination of runoff risk and the presence of DRP at a given location. While manure is a DRP source that decreases rapidly after application and can be managed, soil DRP has much slower dynamics and is always present as a source (Kleinman et al., 2011a). Areas with high simulated DRP loads were mainly distributed along the stream network, or in flat areas with high soil P concentrations a bit farther away from the stream (Fig. 7). There was little difference between the two model versions regarding the area that is expected to contribute the most. The extent of the hatched area in Fig. 7 however was larger for Version 3 than for Version 2. The hatched area illustrates where less than 80 % of the simulations resulted in the same distribution of fast flow generation and thus indicates where model predictions were fairly uncertain. Accounting for all model predictions, we calculated the average DRP load for each pixel. For 90 % of the agricultural area in the Stägbach catchment, the average DRP load calculated over the whole simulation period was below $14.9\,\mathrm{mg\,h^{-1}\,pixel^{-1}}$ for Version 2 and below $13.7\,\mathrm{mg\,h^{-1}\,pixel^{-1}}$ for Version 3. The remaining 10 % of the agricultural area delivered more than half of the total load exported from agricultural land (Version 2: 52 %, Version 3: 54 %). Neglecting winter months, the estimated yearly DRP loads from 10 % of the agricultural area averaged $3.4\,\mathrm{kg\,ha^{-1}}$ (Version 2) and $3.1\,\mathrm{kg\,ha^{-1}}$ (Version 3). During the large runoff event in June 2010, much higher loads per hour were simulated. Again, 10 % of the agricultural area delivered more than 50 % of the DRP load from the total agricultural area. The estimated load per hectare for 10 % of the area averaged $24\,\mathrm{g\,ha^{-1}\,h^{-1}}$ (Version 2) and $29\,\mathrm{g\,ha^{-1}\,h^{-1}}$ (Version 3) during this event.

Table 5. Spatial extent of risk classes in the Stägbach catchment for different model versions and two runoff events in 2010 – relative to the total agricultural area in %.

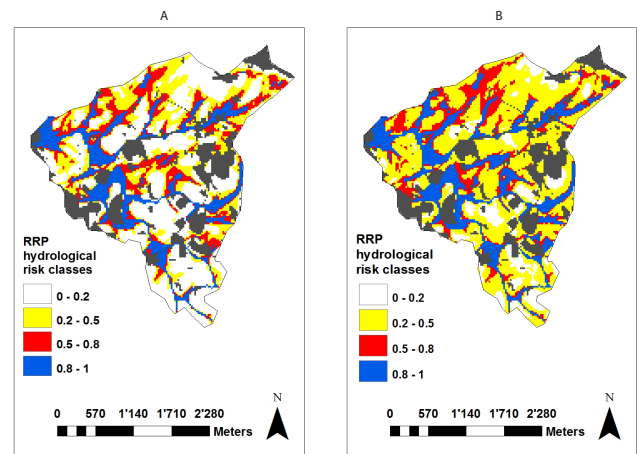| Risk classes Risk values	Low 0 to 0.2	Medium 0.2 to 0.5	High 0.5 to 0.8	Very high 0.8 to 1
Version 2				
small event in May	76	12	5	7
large event in June	44	23	17	16
Version 3				
small event in May	48	41	4	7
large event in June	13	50	16	21

Fig. 6. Risk maps for the large event of June 2010 in the Stägbach catchment, obtained with model versions 2 (left) and 3 (right). Grey shading denotes forested and urban areas

3.3 Spatial model performance and field measurements

3.3.1 Test of model assumptions

The data from the four permanent measurement stations shown in Fig. 8 supported the assumptions underlying the model that (1) soil water saturation increases with topographic index λ, and that (2) well-drained soils are drier than poorly drained soils. The location of station S3, which was situated in the poorly drained HRU of the Stägbach catchment, had the highest λ (11.13) and showed the highest water saturation over the whole measurement period. In contrast, the location of station S1, which was situated in the well-drained HRU, had the lowest λ value (7.16) and always showed the lowest soil water saturation. Station S2, which was also situated in the well-drained HRU but at a location with a higher λ value (10.65) than S1, showed a soil water saturation between that of S1 and S3 and similar to that of station 4. The latter also had a similar topographic index ($\lambda = 7.27$) as station 2, while it had an intermediate position with respect to the classification by drainage classes. Station 4 was situated in the poorly drained HRU according to the soil classification used in Version 3 of the model but in the
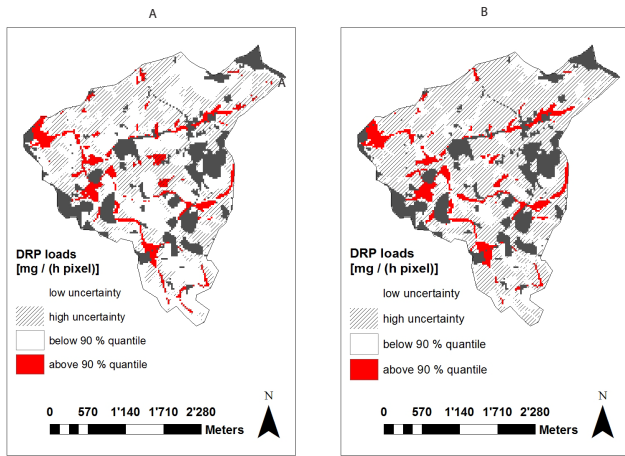
Fig. 7. Simulated distribution of DRP loads in the Stägbach catchment during the large event in June 2010 obtained by averaging over all Monte Carlo simulations. Red color shows the area (10 % of the total agricultural area) where according to the simulations more than 50 % of the total DRP loss occurred. Areas for which less than 80 % of the simulations resulted in the same distribution of fast flow generation are hatched. Grey shading denotes forested and urban areas.

well-drained HRU according to the classification used in Version 2. Thus, the results suggest that soil moisture was more closely related to topographic index than to soil drainage category.

3.3.2 Model predictions and soil moisture measurements

Figure 8 furthermore shows that the predicted risk of fast flow was closely related to measured soil water saturation, confirming the validity of the hydrological simulations presented before. At the two stations with high λ values (S2, S3), the predicted risk of fast flow strongly increased when soil moisture approached full saturation, while there was generally a very low risk of fast flow with comparatively little response to variations in soil moisture at the two stations with low λ values (S1, S4). Version 3 consistently predicted higher risks of fast flow than Version 2, in line with the results presented in Sect. 3.2.1.

3.3.3 Model predictions, OFD and groundwater measurements

The model predictions of fast flow risks were also in reasonable agreement with runoff data recorded by the OFD (Fig. 8). Surface runoff occurred at sites S2 and S3 when both model versions predicted a risk of fast flow above 0.75. No runoff was collected when the predicted risk was below 0.5. On the other hand, runoff was never observed at station S1, for which the predicted risk values were always below 0.05 for model Version 2 and 0.225 for model Version 3. Some
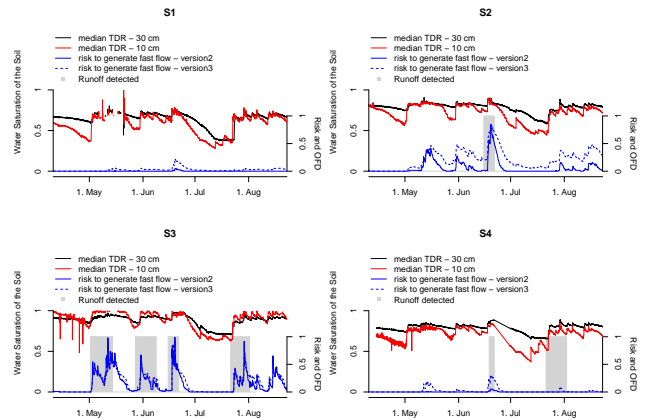

Fig. 8. Comparison of soil moisture measurements, runoff measurements (with overland flow detectors, OFDs) and model predictions of fast flow risks at the four permanent soil moisture measurement stations in the Stägbach catchment for the year 2010.

over-prediction of runoff risks may be due to the fact that OFDs only collect surface runoff, whereas predicted fast flow also includes preferential flow in the RRP model. This may in particular have been the case at station S7, which was one of the six other measurement stations that were not permanently operated. For this location both model versions often predicted high fast flow risks, sometimes even in all simulations, but runoff was collected only once with the installed OFD. This station was located close to a brook where a large amount of the simulated runoff may actually have been due to subsurface flow. In contrast to stations S1, S2, S3 and S7, the risk of runoff from station S4 was underestimated. Surface runoff was collected at S4 during the large event in June, while model Version 2 predicted fast flow only in 6 % of the simulations. Similarly, no elevated risk was predicted for the event at the end of July, when 10 mL of runoff were collected (Fig. 8). Using model Version 3 substantially higher risks of fast flow were predicted for S4 than by Version 2, but even for the extreme event in June the predicted risk still did not exceed a value of 0.3. Similar under-predictions of runoff risks were also obtained for one event at sites S5, S8 and S10, where runoff was collected by the OFD, while the predicted risks remained below 0.1 for model Version 2 and below 0.3 for model Version 3. At two of the three locations, infiltration excess runoff or runoff from a street farther upslope may have had some influence.

While OFD-recorded runoff data only showed a rather loose relationship with the prediction of fast flow events, there were close relationships between groundwater levels recorded by the piezometers and the fraction of accepted parameter sets that resulted in the prediction of fast flow (Fig. S2). Even changes in groundwater table at relatively low levels were associated with changes in risk predictions, in particular with model Version 3.

4 Discussion

Despite the low amount of input data required, the predictions of the RRP model were in good agreement with the measurements, especially after separate calibration of the urban HRU parameters. The latter improved the prediction of small runoff peaks, which were overestimated by the original model (Lazzarotto et al., 2006). The fact that the model adequately predicted discharge and DRP export at the outlet of a catchment and sub-catchment that had not been used for calibration is evidence for the validity of the underlying concept and assumptions. The comparison of hydrological model predictions with measurements of soil moisture, surface runoff and groundwater levels at various locations within the Stägbach catchment provides further support for this conclusion. Of course, model application is always limited to situations that fulfill the assumptions on which a model is based (Radcliffe et al., 2009; Schoumans et al., 2009).

One of the inherent assumptions of the RRP model is that soil can be represented as a single compartment (Lazzarotto et al., 2006), which is a linear storage for the slow flow component. These simplifications lead to an accelerated baseflow decline during dry conditions. This may limit the usefulness of the model for areas dominated by highly permeable soils and for long dry periods. Limited performance during dry periods may furthermore be explained with (i) the short dry period in the calibration data set and with (ii) the role of the topographic index λ in the model. Western et al. (1999) found that the spatial organization of soil moisture could be well described by topography during wet periods, when surface and subsurface lateral redistribution of water occurs. During dry periods they observed little spatial organization of soil moisture. Thus, triggering fast flow by a threshold based on the topographic index is assumed to perform better for wet than for dry soil conditions. Furthermore, the λ does not account for differences in soil moisture caused by aspect (Kopecky and Cizkova, 2010). This might partially explain the limited differentiation between S1 and S4. To test how the linearity of the slow flow storage affects the model results, we also implemented a version where flow is proportional to the square of the storage. Although the model version resulted in slightly better NSC values, the basic problem of drying out too quickly remained and the spatial CSA predictions were rather similar (see Fig. S3). It seems that a better model performance for dry conditions would probably need a better depth resolution of the model at the cost of higher model complexity. Another drawback of the RRP model is that it disregards connectivity. Although the measurements showed that the model did not identify all locations of high fast flow risk, we do not consider the soil representation and disregard of connectivity a major problem for the target region because in the hilly areas of the Swiss Plateau, soils are generally of low permeability and often directly connected to streams or lakes through artificial subsurface drains. Since baseflow simulations are less important for DRP losses than high flow

conditions, the use of λ is justifiable. It however restricts fast flow generation, including infiltration excess runoff (IER), to potentially wet areas and therefore may underestimate the extent of IER generating areas. This as well as the occurrence of local thunderstorms, for which the amount of rain may not be recorded correctly, can lead to underestimations of runoff peaks.

Due to the crucial role of hydrology in P losses, which was also pointed out by Kleinman et al. (2011a), the accuracy of predictions of DRP loads at the catchment outlets mainly depended on the quality of the hydrological simulations, an observation also made by Hively et al. (2006). The good agreement between predictions and measurements obtained when discharge was described well indicates that the model adequately captured all relevant processes in our catchments. Thus, there was no need to incorporate further processes, such as those proposed by Vadas et al. (2011). They suggested relating DRP concentrations to the runoff-to-rain ratio. According to them, a higher runoff-to-rain ratio leads to higher DRP concentrations in runoff from manured soils. Based on data from 9 studies, Vadas et al. (2011) showed that a high runoff-to-rain ratio often means that runoff starts in an earlier phase of an event than in events with a low ratio. This was considered important because concentrations of P released from manure decrease with time during an event (Sharpley and Moyer, 2000). While the RRP model worked well in our study without such a refinement, it would be easy to incorporate this relationship if deemed appropriate for other applications.

According to our simulations, most DRP lost with runoff originated from P-enriched soils. This holds despite the uncertainty of the representation of manure-derived losses. Varying parameter h, which describes the temporal decay of the manure-P pool, between 0.003 and 0.011 changed the fraction attributed to manure between 20 and 30 %. However, soil P was mostly the dominant fraction. Using a fully distributed model, Hively et al. (2006) came to a similar conclusion for a rural watershed in the New York State. These findings support the conclusions of Kleinman et al. (2011a), who announced that legacy P continues to represent a high and permanent risk of P export into waters that needs to be reduced. Also manure application can lead to substantial P loads in runoff (Shigaki et al., 2007; Withers et al., 2003).

Unfortunately, the capability of models to determine the sources of P in catchment discharge has still not been tested by direct measurements. This lack of validation also includes the capability of models to allocate the spatial origin of P losses from a catchment (White et al., 2009). Till now validation has been mainly based on the comparison of P concentrations in different stream segments (Gburek and Sharpley, 1998). Further development of isotopic methods such as that of Tamburini et al. (2010) is needed to determine the sources of P (manure, legacy P) found in runoff. While tracers have been used to determine source areas of eroded sediments (Stevens and Quinton, 2008) and pesticides (Doppler et al.,

2012; Leu et al., 2004), the identification of source areas for P losses remains challenging.

Our model predictions were quite robust with respect to the two schemes of binary soil classification by drainage capacity compared in our study. This applied not only to discharge predictions but also to the delineation of CSAs within the catchment, which makes the model valuable for the identification of CSAs within catchments. Our findings suggest that soil drainage capacity was less important for soil moisture status, and thus also for the risk of fast flow generation in our study area, than topography.

In accordance with the "variable source area" concept (Ward, 1984) and observations of Gburek and Sharpley (1998), the RRP model predicted an increase in runoff generating areas with increasing soil moisture. If enriched with P sources that can be easily mobilized, these hydrologically active areas can be a severe threat for water quality (Gburek and Sharpley, 1998). In the Stägbach catchment, areas with high simulated DRP loads, averaged over the monitoring period, were mainly situated along the stream network. Nevertheless, there are also some high risk areas located farther away from streams, which can contribute to catchment runoff and P export via drainage systems. According to our simulations, the 10 % of the area contributing the most delivered more than 50 % of the total DRP export from the Stägbach catchment. Pionke et al. (2000) and White et al. (2009) obtained similar results. Pionke et al. (2000) calculated that the majority of the DRP exported from the Brown catchment into the Chesapeake Bay derived from 11 % of the catchment area, while simulations of P export from 6 catchments in Oklahoma by White et al. (2009) predicted that on average 5 % of the area yielded 34 % of the exported P loads. However, the model results and their uncertainty demonstrate also that one cannot exclude the possibility that large fractions (40–50 %) of the catchment may contribute (see Fig. 7).

Our findings provide further support for suggestions of previous authors that management strategies to reduce P transfer from agricultural areas into surface water bodies should focus on the prevention and reduction of P accumulation in soils close to streams and in particular restrict fertilizer and manure applications in these areas. In addition, high risk areas located farther away from streams but connected via drainage systems also need to be considered.

5 Conclusions

Our results demonstrate that the RRP model is able to make useful predictions of discharge and DRP losses from grassland-dominated catchments. The validity of the underlying concept is further supported by the agreement between spatial predictions of runoff generation risks with ground measurements of soil moisture, surface runoff and groundwater levels. The predictions were sufficiently robust with respect to the binary classification of soil drainage capacity

to allow the use of conventional soil maps to assign the soils of the simulated catchment to these classes. The hydrological predictions were in line with the CSA concept and highlight the dominant role of topography. While the model suggests that the 10 % of the catchment area contributing the most delivered more than 50 % of the total DRP load, the result also reveals a considerable risk that larger fractions of the catchments contribute as well. For practical applications this means that targeting the 10 % of high risk areas will most probably reduce DRP losses. However, more areal options may be needed to reduce them to a sufficient degree. According to the model, the actual measures should focus on legacy P as it was the dominant source for DRP losses. These findings confirm conclusions of previous authors that P enrichment in soils of hydrological active areas presents a high risk for water quality and needs to be reduced. The parsimonious RRP model is a suitable tool to delineate risk areas and guide the implementation of mitigation measures.

Supplementary material related to this article is available online at http://www.hydrol-earth-syst-sci.net/17/3679/2013/hess-17-3679-2013-supplement.pdf.

References

Arnold, J. G., Srinivasan, R., Muttiah, R. S., and Williams, J. R.: Large area hydrologic modeling and assessment – Part 1: Model development, J. Am. Water Resour. As., 34, 73–89, 1998.

Beasley, D. B., Huggins, L. F., and Monke, E. J.: ANSWERS – a model for watershed planning, T. ASAE, 23, 938–944, 1980.

Beven, K. J. and Kirkby, M. J.: A physically based, variable contributing area model of basin hydrology, Hydrolog. Sci. Bull., 24, 43–69, 1979.

Braun, M., Hurni, P., and Von Albertini, N.: Abschwemmung von Phosphor auf Grasland an zwei verschiedenen Standorten im Einzugsgebiet des Sempacher Sees, Landwirtschaft Schweiz, 6, 615–620, 1993 (in German).

Buda, A. R., Kleinman, P. J. A., Srinivasan, M. S., Bryant, R. B., and Feyereisen, G. W.: Effects of hydrology and field management on phosphorus transport in surface runoff, J. Environ. Qual., 38, 2273–2284, 2009.

Carpenter, S. R., Caraco, N. F., Correll, D. L., Howarth, R. W., Sharpley, A. N., and Smith, V. H.: Nonpoint pollution of surface waters with phosphorus and nitrogen, Ecol. Appl., 8, 559–568, 1998.

Doody, D. G., Archbold, M., Foy, R. N., and Flynn, R.: Approaches to the implementation of the Water Framework Directive: Targeting mitigation measures at critical source areas of diffuse phosphorus in Irish catchments, J. Environ. Manage., 93, 225–234, 2012.

Doppler, T., Camenzuli, L., Hirzel, G., Krauss, M., Lück, A., and Stamm, C.: Spatial variability of herbicide mobilisation and transport at catchment scale: insights from a field experiment, Hydrol. Earth Syst. Sci., 16, 1947–1967, doi:10.5194/hess-16-1947-2012, 2012.

Gburek, W. J. and Sharpley, A. N.: Hydrologic controls on phosphorus loss from upland agricultural watersheds, J. Environ. Qual., 27, 267–277, 1998.

Hahn, C., Prasuhn, V., Stamm, C., and Schulin, R.: Phosphorus losses in runoff from manured grassland of different soil P status at two rainfall intensities, Agr. Ecosyst. Environ., 153, 65–74, 2012.

Heathwaite, L., Sharpley, A., and Bechmann, M.: The conceptual basis for a decision support framework to assess the risk of phosphorus loss at the field scale across Europe, J. Plant Nutr. Soil Sc., 166, 447–458, 2003.

Herzog, P.: Sanierung des Baldegger Sees, Auswertung der Zufluss-Untersuchungen 2000 bis 2004, 2005 (in German).

Hively, W. D., Gérard-Marchant, P., and Steenhuis, T. S.: Distributed hydrological modeling of total dissolved phosphorus transport in an agricultural landscape, part II: dissolved phosphorus transport, Hydrol. Earth Syst. Sci., 10, 263–276, doi:10.5194/hess-10-263-2006, 2006.

Hornberger, G. M. and Spear, R. C.: An approach to the preliminary-analysis of environmental systems, J. Environ. Manage., 12, 7–18, 1981.

Kirkby, M.: Hydrograph modelling strategies, 69–90, Heinemann, London, 1975.

Kleinman, P. J. A., Allen, A. L., Needelman, B. A., Sharpley, A. N., Vadas, P. A., Saporito, L. S., Folmar, G. J., and Bryant, R. B.: Dynamics of phosphorus transfers from heavily manured Coastal Plain soils to drainage ditches, J. Soil Water Conserv., 62, 225–235, 2007.

Kleinman, P. J. A., Sharpley, A. N., Buda, A. R., McDowell, R. W., and Allen, A. L.: Soil controls of phosphorus in runoff: Management barriers and opportunities, Can. J. Soil Sci., 91, 329–338, 2011a.

Kleinman, P. J. A., Sharpley, A. N., McDowell, R. W., Flaten, D. N., Buda, A. R., Tao, L., Bergstrom, L., and Zhu, Q.: Managing agricultural phosphorus for water quality protection: principles for progress, Plant Soil, 349, 169–182, 2011b.

Kopecky, M. and Cizkova, S.: Using topographic wetness index in vegetation ecology: does the algorithm matter?, Appl. Veg. Sci., 13, 450–459, 2010.

Lazzarotto, P.: Modeling phosphorus runoff at the catchment scale, Ph.D. thesis, 2005.

Lazzarotto, P., Prasuhn, V., Butscher, E., Crespi, C., Fluhler, H., and Stamm, C.: Phosphorus export dynamics from two Swiss grassland catchments, J. Hydrol., 304, 139–150, 2005.

Lazzarotto, P., Stamm, C., Prasuhn, V., and Fluhler, H.: A parsimonious soil-type based rainfall-runoff model simultaneously tested in four small agricultural catchments, J. Hydrol., 321, 21–38, 2006.

Leu, C., Singer, H., Stamm, C., Muller, S. R., and Schwarzenbach, R. P.: Variability of herbicide losses from 13 fields to surface water within a small catchment after a controlled herbicide application, Environ. Sci. Technol., 38, 3835–3841, 2004.

Nash, J. E. and Sutcliffe, J. V.: River flow forecasting through conceptual models part I – A discussion of principles, J. Hydrol., 10, 282–290, 1970.

Pacini, N. and Gächter, R.: Speciation of riverine particulate phosphorus during rain events, Biogeochemistry, 47, 87–109, 1999.

Peyer, K., Brunner, H., Wüst, O., and Plüss: Bodenkarte Hochdorf, Eidg. Forschungsanstalt für landw. Pflanzenbau, 1983.

Pionke, H. B., Gburek, W. J., Sharpley, A. N., and Zollweg, J. A.: Hydrological and chemical controls on phosphorus loss from catchments, 225–242, CAB International Press, Cambridge, 1997.

Pionke, H. B., Gburek, W. J., and Sharpley, A. N.: Critical source area controls on water quality in an agricultural watershed located in the Chesapeake Basin, Ecol. Eng., 14, 325–335, 2000.

Quinn, P., Beven, K., Chevallier, P., and Planchon, O.: The prediction of hillslope flow paths for distributed hydrological modeling using digital terrain models, Hydrol. Process., 5, 59–79, 1991.

Radcliffe, D. E., Freer, J., and Schoumans, O.: Diffuse phosphorus models in the united states and europe: Their usages, scales, and uncertainties, J. Environ. Qual., 38, 1956–1967, 2009.

RDevelopmentCoreTeam: R: A language and environment for statistical computing, release 2.6.1, R Foundation for Statistical Computing, Vienna, Austria, 2007.

Schoumans, O. F., Silgram, M., Groenendijk, P., Bouraoui, F., Andersen, H. E., Kronvang, B., Behrendt, H., Arheimer, B., Johnsson, H., Panagopoulos, Y., Mimikou, M., Lo Porto, A., Reisser, H., Le Gall, G., Barr, A., and Anthony, S. G.: Description of nine nutrient loss models: capabilities and suitability based on their characteristics, J. Environ. Monitor., 11, 506–514, 2009.

Schulte, R. P. O., Doody, D. G., Byrne, P., Cockerill, C., and Carton, O. T.: Lough Melvin: Developing cost-effective measures to prevent phosphorus enrichment of a unique aquatic habitat, Tearmann, 7, 211–228, 2009.

Sharpley, A. N.: Assessing phophorus bioavailability in agricultural soils and runoff, Fert. Res., 36, 259–272, 1993.

Sharpley, A. and Moyer, B.: Phosphorus forms in manure and compost and their release during simulated rainfall, J. Environ. Qual., 29, 2053–2053, 2000.

Sharpley, A. N., Chapra, S. C., Wedepohl, R., Sims, J. T., Daniel, T. C., and Reddy, K. R.: Managing agricultural phophorus for protection of surface waters - issues and options, J. Environ. Qual., 23, 437–451, 1994.

Sharpley, A. N., Weld, J. L., Beegle, D. B., Kleinman, P. J. A., Gburek, W. J., Moore, P. A., and Mullins, G.: Development of phosphorus indices for nutrient management planning strategies in the United States, J. Soil Water Conserv., 58, 137–152, 2003.

Shigaki, F., Sharpley, A., and Prochnow, L. I.: Rainfall intensity and phosphorus source effects on phosphorus transport in surface runoff from soil trays, Sci. Total Environ., 373, 334–343, 2007.

Hydrol. Earth Syst. Sci., 17, 3679–3693, 2013

www.hydrol-earth-syst-sci.net/17/3679/2013/

Smith, K. A., Jackson, D. R., and Withers, P. J. A.: Nutrient losses by surface run-off following the application of organic manures to arable land. 2. Phosphorus, Environ. Pollut., 112, 53–60, 2001.

Stamm, C., Flühler, H., Gächter, R., Leuenberger, J., and Wunderli, H.: Preferential transport of phosphorus in drained grassland soils, J. Environ. Qual., 27, 515–522, 1998.

Stamm, C., Sermet, R., Leuenberger, J., Wunderli, H., Wydler, H., Flühler, H., and Gehre, M.: Multiple tracing of fast solute transport in a drained grassland soil, Geoderma, 109, 245–268, 2002.

Stevens, C. J. and Quinton, J. N.: Investigating source areas of eroded sediments transported in concentrated overland flow using rare earth element tracers, Catena, 74, 31–36, 2008.

Strauss, P., Leone, A., Ripa, M. N., Turpin, N., Lescot, J. M., and Laplana, R.: Using critical source areas for targeting costeffective best management practices to mitigate phosphorus and sediment transfer at the watershed scale, Soil Use Manage., 23, 144–153, 2007.

Tamburini, F., Bernasconi, S. M., Angert, A., Weiner, T., and Frossard, E.: A method for the analysis of the delta O-18 of inorganic phosphate extracted from soils with HCl, Eur. J. Soil Sci., 61, 1025–1032, 2010.

Topp, G. C., Davis, J. L., and Annan, A. P.: Electromagnetic determination of soil water content: measurements in coaxial transmission lines, Water Resour. Res., 16, 574–582, 1980.

Vadas, P. A., Kleinman, P. J. A., Sharpley, A. N., and Turner, B. L.: Relating soil phosphorus to dissolved phosphorus in runoff: A single extraction coefficient for water quality modeling, J. Environ. Qual., 34, 572–580, 2005.

Vadas, P. A., Jokela, W. E., Franklin, D. H., and Endale, D. M.: The effect of rain and runoff when assessing timing of manure application and dissolved phosphorus loss in runoff, J. Am. Water Resour. As., 47, 877–886, 2011.

Vogler, P.: Beiträge zur Phosphatanalytik in der Limnologie. II. Die Bestimmung des gelösten Ortophosphates, Fortschritte der Wasserchemie und ihrer Grenzgebiete, Vol. 2, 1965 (in German).

Von Albertini, N., Braun, M., and Hurni, P.: Oberflächenabfluss und Phosphorabschwemmung von Grasland, Landwirtschaft Schweiz Band 6, 575–582, 1993 (in German).

Wade, A. J., Whitehead, P. G., and Butterfield, D.: The Integrated Catchments model of Phosphorus dynamics (INCA-P), a new approach for multiple source assessment in heterogeneous river systems: model structure and equations, Hydrol. Earth Syst. Sci., 6, 583–606, doi:10.5194/hess-6-583-2002, 2002.

Ward, R. C.: On the response to precipitation of headwater streams in humid areas, J. Hydrol., 74, 171–189, 1984.

Watson, C. J. and Matthews, D. I.: A 10-year study of phosphorus balances and the impact of grazed grassland on total P redistribution within the soil profile, Eur. J. Soil Sci., 59, 1171–1176, 2008.

Weld, J. and Sharpley, A. N.: Phosphorus indices, 3–19, CRC Press, Boca Raton, 2007.

Western, A. W., Grayson, R. B., Bloschl, G., Willgoose, G. R., and McMahon, T. A.: Observed spatial organization of soil moisture and its relation to terrain indices, Water Resour. Res., 35, 797–810, 1999.

White, M. J., Storm, D. E., Busteed, P. R., Stoodley, S. H., and Phillips, S. J.: Evaluating nonpoint source critical source area contributions at the watershed scale, J. Environ. Qual., 38, 1654–1663, 2009.

Withers, P. J. A., Ulén, B., Stamm, C., and Bechmann, M.: Incidental phosphorus losses – can they be predicted?, J. Plant Nutr. Soil Sc., 166, 459–468, 2003.