

A comparison of three simple approaches to identify critical areas for runoff and dissolved reactive phosphorus losses

C. Hahn^{1,2}, V. Prasuhn², C. Stamm³, D.G. Milledge⁴, and R. Schulin¹

¹ETH Zurich, Department of Environmental Systems Science, Zurich, Switzerland

²Agroscope Reckenholz-Tänikon Research Station ART, Zurich, Switzerland

³Eawag Swiss Federal Institute of Aquatic Science and Technology, Zurich, Switzerland

⁴Durham University, Department of Geography, Durham, United Kingdom

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

Abstract. Diffuse phosphorus (P) losses are the main cause for eutrophication of surface waters in many regions. Implementing mitigation measures on critical source areas (CSA) is seen to be the most effective way to reduce P-losses. Thus, tools are needed that delineate CSA on the basis of available data. We compared three models based on different approaches and sets of input data:

5 the Rainfall Runoff Phosphorus (RRP) model, the Dominant Runoff Processes (DoRP) model, and the Sensitive Catchment Integrated Modeling Analysis Platform (SCIMAP). The RRP model is a parsimonious dynamic model using the topographic index and a binary soil classification to simulate discharge and P-losses. The DoRP model distinguishes 8 soil classes based on soil and geological maps. It does not account for topography when calculating runoff. SCIMAP assesses runoff risks

10 solely on the basis of topography using the network index. Compared to surface runoff and soil moisture data available from a catchment in Switzerland, the RRP model and SCIMAP made better predictions than the DoRP model, suggesting that in our study area topography was more important for CSA delineation than soil data. Based on the results, we suggest improvements of SCIMAP to enable average risk predictions and the comparison of risk predictions between catchments.

15 1 Introduction

Diffuse phosphorus (P) losses from agricultural land continue to be a severe problem for water quality, causing eutrophication of many surface waters (Buda et al., 2012; Kleinman et al., 2011b; Schoumans et al., 2009). The European Union (EU) Water Framework Directive (WFD) calls for restoration of all water bodies to good quality by 2015 if possible, and 2027 at the latest (Hering

20 et al., 2010). Since P losses from point sources, which can be identified and targeted rather easily, were substantially reduced over the last two decades (Dubrovsky et al., 2010), diffuse P losses from agricultural areas are now the main cause of eutrophication of water bodies in many countries (Carpenter et al., 1998; Sharpley et al., 1994). Thus, in order to meet the WFD goal, the focus now needs to be directed to measures by which these diffuse inputs can effectively be reduced.

25 A wide range of different mitigation options were proposed and discussed within COST Action 869 (Schoumans et al., 2011). Some measures aim to prevent or intercept P transfer by runoff from fields to surface waters, (e.g. buffer strips (Roberts et al., 2012; Stutter et al., 2012)); others focus on the immobilization of P sources in soil and manure by adding P sorbing agents (Buda et al., 2012).

Several studies reported that most P found in runoff at the catchment outlet originated from rather
30 small areas within the catchment (Gburek and Sharpley, 1998; Pionke et al., 2000, 1997). Targeting these critical source areas (CSAs) seems to be the most efficient and cost-effective approach to mitigate the water pollution problem (Doody et al., 2012; Heathwaite et al., 2003; Rodriguez et al., 2011; Schulte et al., 2009).

Kleinman et al. (2011b) reviewed source and transport factors that need to be considered when addressing diffuse P losses and CSA delineation. Distinguishing chronic or legacy sources and acute
35 or temporary sources, the main P sources are pools of legacy P in soils due to excessive fertilizer application in the past and temporary increases in available P resulting from current applications of manure or fertilizer. Decaying crop residues can also play a role as a temporary P source (Kleinman et al., 2011b; Pote et al., 1999). While proper crop nutrition and soil management according to codes
40 of best practices can reduce acute sources very effectively, sources of legacy P are more difficult to deal with as they can persist for a long time (Kleinman et al., 2011a). In earlier studies erosion and surface runoff were considered the main or only transport mechanisms carrying P to streams. More recently, other processes have been recognized to transport substantial amounts of P to surface waters (Doody et al., 2012; Kleinman et al., 2011b) such as subsurface flow (Kleinman et al., 2009),
45 and tile drain flow (Kleinman et al., 2007; Stamm et al., 1998; Vadas et al., 2007).

Given the diversity of sources and transport processes involved in diffuse P losses, the identification of CSAs in a catchment is a challenging task. Various tools have been developed to model diffuse P losses and to delineate CSAs (Krueger et al., 2012; Radcliffe et al., 2009; Schoumans et al., 2009), ranging from rather simple site-assessment tools to complex physically based catchment models.

50 The former are usually easy to apply but are often too simplistic, whereas the latter are generally too complex for practical applications and usually require input data that are not easily available. As pointed out by various authors (Heathwaite et al., 2007; Radcliffe et al., 2009; Srinivasan and McDowell, 2007), there is a need for parsimonious process-based and easy-to-use, but nonetheless sound and reliable models that are easy to parameterize and can assist water protection policy and
55 agricultural P management in targeting effective and cost-efficient water protection and mitigation measures to the most critical areas.

Here we compare three tools that can be used to predict critical source areas in catchments based on different approaches and input data: the Rainfall Runoff Phosphorus (RRP) model, the Dominant Runoff Processes (DoRP) model, and the Sensitive Catchment Integrated Modeling Analysis Platform (SCIMAP). The Rainfall Runoff Phosphorus (RRP) model is a dynamic, parsimonious model that simulates runoff and dissolved reactive P (DRP) losses from a catchment, using the topographic index λ proposed by Beven and Kirkby (1979) and a binary soil classification by drainage capacity to simulate a site's hydrological responses to rainfall events (Hahn et al., 2013; Lazzarotto, 2005). The DoRP model, which was proposed by Schmocker-Fackel et al. (2007), distinguishes 8 soil classes by drainage and soil water storage capacities, based on soil and geological maps, but does not account for topography when calculating runoff. In contrast, SCIMAP, which was proposed by Lane et al. (2006), assesses runoff risks solely on the basis of topography, using the network index NI introduced by Lane et al. (2004). This network index is derived from the topographic index λ by correcting it for reduced topographic connectivity of locations that are not directly connected to a stream. RRP and SCIMAP provide a framework to combine these hydrological predictions with pollutant source data to identify CSAs. DoRP on the other hand solely comprises the hydrological part and does not originally provide a structure to combine the hydrological predictions with pollutant source data.

While the three models represent very different approaches, their performance has never been tested in direct comparison. Our hypothesis here was that we can get useful information from such a comparison not just about the specific models, but also on the underlying general approaches. Thus, we tested the three approaches by applying them to two catchments located in a grassland-dominated hilly region of the Swiss Plateau. First, we investigated the importance of topography relative to soil properties in defining runoff risks by comparing the ability of the RRP and DoRP models to reproduce observed runoff patterns. We then compared the RRP model to SCIMAP to test the impact of applying a simpler, time integrated model. While Lane et al. (2009) also evaluated the Network Index using a dynamic model they suggest that it is necessary to further investigate the index's potential with regards to the duration of integration (monthly, yearly). We therefore focus on that aspect during our assessment of SCIMAP. We used the RRP model as reference because it is the most comprehensive of the three models and it had already been validated against data from the catchment outlets as well as from within the experimental catchments, including observations on soil moisture, runoff generation and groundwater levels. For a detailed presentation and discussion of the validation of RRP readers are referred to (Hahn et al., 2013; Lazzarotto, 2005).

By comparing the three approaches (RRP, DoRP, SCIMAP), we investigate the importance of connectivity and detailed soil data for the delineation of CSA and identify the benefits of a parsimonious dynamic model using soil and topographical information. In addition, we assess the performance of SCIMAP regarding the duration of integration (storm event, yearly) and the comparability of SCIMAP results between catchments.

2 Materials and methods

95 2.1 Model concepts

2.1.1 The Rainfall-Runoff-Phosphorus (RRP) model

The RRP model is a parsimonious dynamic model developed by Lazzarotto (2005) to predict dissolved reactive P (DRP) losses from small agricultural catchments. The model consists of a hydrological and a P sub-model. For this study we used Version 2 of the model, which has been developed and described in detail by Hahn et al. (2013).
100

The hydrological sub-model (Lazzarotto et al., 2006), referred to here as Rainfall-Runoff model, is a semi-distributed hydrological model with an hourly time resolution that uses soil and topographic information to describe the dynamics of the system. It is based on the concept that areas with the same topographic index λ and the same soil drainage capacity have the same hydrological behavior.
105 The topographic index λ (Beven and Kirkby, 1979; Kirkby, 1975) is an indicator of the wetness of the soil at a given location within a catchment and is determined by the upslope contributing area and the local slope of a location. Soils are divided into 2 drainage classes, well drained and poorly drained soils, the ensembles of which form a poorly-drained and a well-drained hydrological response unit (HRU), respectively. The discharge from both of these two HRUs is composed of a fast and a slow flow component. The former comprises all kinds of quickly responding flow, including preferential flow, saturation excess runoff and Hortonian overland flow. The slow flow component is an approximation of the baseflow. The model was simultaneously calibrated (uniform Monte Carlo method) on discharge data from four catchments draining into Lake Sempach (Hahn et al., 2013; Lazzarotto et al., 2006). The four catchments varied in soil composition and hydrological responses.
115 The model parameters were determined by repeated random sampling from a uniform prior distribution within the range of each parameter. The performance of each parameter combination was assessed by comparing simulated discharge with measured discharge in the four catchments. While not being very long, the calibration period covered a wide range of stream flow conditions, which is more important than its actual length for obtaining reliable results (Gupta and Sorooshian, 1985; Lazzarotto et al., 2006; Yapo et al., 1996). Using the modified Nash-Sutcliffe criterion NSC (Nash and Sutcliffe, 1970) as defined by Lazzarotto et al. (2006) and a NSC threshold value of 0.6, 724 parameter sets, out of 5 million, were judged behavioral and used for model application (Hahn et al., 2013). This GLUE approach produced parameter values that gave very good predictions of the discharge for validation time periods as well as for a different catchment and thus was considered
125 satisfactory.

The P-model assigns a constant P concentration, observed during baseflow (Lazzarotto, 2005), to baseflow runoff. For areas where fast flow occurs, P loads comprise incidental P losses from manure, P losses from topsoil, which might be enriched with P due to P application in excess of crop

demands in the past (also called legacy P), and losses associated with baseflow.

130 **2.1.2 The Dominant Runoff Processes (DoRP) assessment scheme**

In contrast to the simplification in the RRP model regarding the soil data, the GIS-based DoRP model presented by Schmocker-Fackel et al. (2007) classifies the soil hydrology in a more differentiated way based on several soil attributes contained in a standard soil map and information about the parent material obtained from geological maps. The approach is based on the study of Scherrer and Naef
135 (2003), who developed decision schemes to determine the dominant runoff process (DoRP) of a soil profile. Schmocker-Fackel et al. (2007) simplified these schemes to reduce data requirements and to enable automatic GIS-based mapping of DoRPs at catchment scale. The DoRP maps were used for flood discharge simulations, but they can also be used to determine risk areas for pesticide or P losses and erosion.

140 The DoRP model distinguishes between the following runoff processes: Hortonian Overland Flow (HOF), Saturated Overland Flow (SOF), fast Subsurface Flow (SSF) and Deep Percolation (DP). For SOF and SSF to occur an impermeable layer in the soil profile is required. According to the drainable porosity of a profile, three classes of soil water storage capacity are distinguished: STO1 = very low (0 - 40 mm), STO2 = medium (40 - 100 mm), STO3 = large (100 - 200 mm), where the numbers in
145 parentheses give the total volume of drainable porosity per unit area above the impermeable layer (in mm). Schmocker-Fackel et al. (2007) used a soil map, a geological map, a forest map and a land-use map to determine the occurrence of DP and HOF and to determine the storage classes STO1, STO2, STO3. Apart from roads and other artificially sealed areas, HOF is assumed to occur on soils with very low infiltration capacity, e.g. due to soil surface sealing, water-repellency or compaction, at
150 very high rainfall intensity ($> 50 \text{ mm h}^{-1}$). At lower intensities, these areas are assigned to another runoff process. Information on slope is needed to differentiate between SOF and SSF. SOF was observed to occur below a slope of 15%, and SSF on slopes above 15% (Schmocker-Fackel et al., 2007). If drainage data is available, SSF can be further differentiated into drainage flow and natural SSF. Following the terminology of Schmocker-Fackel et al. (2007) SOF and SSF runoff from soil
155 with storage capacity STO_i ($i=1,2,3$) is denoted as SSFi and SOF_i , respectively.

Areas dominated by DP are assumed to generate no runoff; instead, rain falling on these areas infiltrates and is delivered to groundwater. As STO1 areas are saturated much faster than STO2 areas and these faster than STO3 areas under the same initial conditions, the risk of runoff is highest on areas with HOF, followed by STO1, STO2 and STO3 areas in a given event. Furthermore, SSF
160 areas are assumed to drain faster than SOF areas. Following Schmocker-Fackel et al. (2007) we used storage capacities of 5 mm for HOF and 20 mm, 70 mm and 150 mm for STO1, STO2 and STO3 areas, respectively. While the DoRP classification of soils gives a qualitative assessment of the hydrological behavior of a given location, the DoRP model also includes a simple bucket model to make quantitative discharge predictions (Schmocker-Fackel et al., 2007). Neglecting lateral flow,

165 runoff is generated, as soon as rainfall exceeds the soil's storage capacity at a given location. To
account for antecedent soil moisture conditions, all rain water is stored for 5 days in the soil, if
storage capacity is available, and is then removed.

2.1.3 SCIMAP

The Sensitive Catchment Integrated Modeling Analysis Platform (SCIMAP) estimates the risk of
170 pollutant loss from a location x to a stream or other receiving water body. This locational risk p_x^{gc}
(Reaney et al., 2011) indicates whether a pollutant is both available at a location and can be delivered
to the water body. It is calculated by multiplying two relative site indicators, the availability risk at
point x (p_x^g) and the connection risk (p_x^c), the risk that it can be delivered to a water body:

$$p_x^{gc} = p_x^g * p_x^c$$

175 In this study, which focuses on DRP losses the availability risk is defined by the P availability (source
factor). The connection risk is the likelihood that runoff is generated and delivered to a stream (trans-
port factor). The availability risk for dissolved P losses can be related to the concentration of legacy
P in the soil, which is a permanent P source. Given that the relationship between water extractable
P concentrations (WSP) in the soil and DRP in runoff was found to be linear in many studies (Hahn
180 et al., 2012; Vadas et al., 2005), we normalized the WSP data by the maximum WSP value. In line
with the original SCIMAP concept, this normalization was carried out for each catchment separately.
To allow for comparisons across catchments, we also modified this normalization procedure to ac-
count for the value range in both catchments.

Following Reaney et al. (2011), the connection risk was derived from the Network Index (NI), which
185 is a connectivity index equivalent to the topographic index λ but corrected for locations of reduced
connectivity (Lane et al., 2004). Although cells with a higher λ are more likely to be saturated than
cells with a lower λ , water from these cells might not reach the stream network, if drier cells along
the flow path enable infiltration. Assuming that the contribution of an upslope location to stream
runoff is limited by the lowest λ value further downslope along the flow path, the NI assigns the
190 lowest λ value found further downslope along that flow path to all cells further upslope that other-
wise would have a higher λ value. The NI then needs to be related to p_x^c . Based on a comparison
with a distributed hydrological model (Lane et al., 2009), Reaney et al. (2011) suggested a p_x^c of 0
for NI values below their 5% quantile and 1 for values above the 95% quantile, with a linear p_x^c to NI
relationship in between (i.e. from $p_x^c = 0$ at the 5% quantile to $p_x^c = 1$ at the 95% quantile). Again,
195 this normalization does not allow comparisons across catchments. We modified this procedure by
deriving quantiles for the pooled data sets of both catchments.

2.2 Spatial comparison of model results

Weighted kappa (Cohen, 1968) was used to compare spatial risk predictions. To calculate kappa the
model results of SCIMAP and RRP were rescaled and grouped. For this purpose, the results obtained

200 with each of the two models were divided by the respective maximum value and then grouped as follows: Locations with values ranging between 0 and 0.2 were considered to be at low risk, with values between 0.2 and 0.5 to be at medium risk, with values between 0.5 and 0.8 to be at high risk, and with values between 0.8 and 1 to be at very high risk. Weighted kappa was calculated using R (RDevelopmentCoreTeam, 2007) and the psy package.

205 **2.3 Study sites**

We applied the three approaches to the Lippenrütibach catchment (3.3 km²) and the Stägbach catchment (8.24 km²), which are both situated in the hilly area of the Swiss Plateau northwest of Lucerne (Fig. 1). The former drains into Lake Sempach, whereas the latter drains into Lake Baldegg. Both lakes have a legacy of eutrophication and are artificially aerated to avoid oxygen depletion in the lower part of the water column. The region is characterized by intensive livestock production (dairy and pig farms, 2.4 livestock units per ha (Herzog, 2005)) and intensively manured permanent grassland (4 to 6 cuts and manure applications per year). Soil P levels are elevated due to a legacy of P inputs in excess of crop demands. The average annual precipitation ranges from 1000 to 1200 mm. The catchments are characterized by hilly terrain that undulates between altitudes from 500 to 800 m above sea level. The landscape shows strong molding by the Würm glaciation. The parent material varies between upper freshwater molasse and moraines (Peyer et al., 1983). The soils are predominantly eutric and gleyic Cambisols, with loamy texture (Peyer et al., 1983; AGBA, 1993). Less than 10% of the two catchment areas are covered by settlements, roads and other constructed features. Forests account for 17% of the Lippenrütibach catchment and 8% of the Stägbach catchment. The remaining area is used for agriculture.

The Lippenrütibach catchment (LIP) is one of the four catchments that were used for the calibration of the RRP model, using discharge data and DRP analyses of runoff collected from the 7th until the 17th of July 2000. In the present study we applied the model to the period from March until November 1999, for which precipitation, ET, discharge, and manure application data were available that had not been used for model calibration. For the Stägbach catchment (Stäg) we used data collected in 2010 to run the RRP model and to validate model results. In addition to discharge and DRP measurements at the catchment outlet, permanent soil moisture measurements at four locations (S1 to S4) as well as groundwater level and Overland Flow Detector (OFD) measurements at 10 stations (S1 to S10) were available for validation. For more details the reader is referred to Hahn et al. (2013).

3 Results

3.1 Hydrology

3.1.1 RRP

The RRP model predicted the discharge dynamics at the outlet of the Lippenrütibach and the Stägbach catchment fairly well, with median NSC values of 0.5 and 0.62, respectively (Hahn et al., 2013). To
235 characterize the uncertainty of fast flow predictions, we constructed maps showing the fraction of the accepted parameter sets that predicted fast flow for each pixel at a given time (Fig. 2A). Values between 0 and 0.2 are considered to indicate 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.
240 Thus, if for example more than 80% of the model realizations predicted fast flow generation for a pixel, that pixel was assigned to the very high risk class. For both catchments the percentage of high-risk areas increased with soil moisture (Table 1). The percentage of the respective agricultural areas with very high predicted runoff risk varied between 9% for small events and 27% for large events in the Lippenrütibach catchment, and between 7 and 16% in the Stägbach catchment. Runoff measurements from OFDs installed at different locations within the Stägbach catchment were in good
245 agreement with these predictions. For example, runoff was collected at 5 out of 7 locations during the large event in June 2010, but at only 2 locations during the small event in May. The RRP model was able to capture the temporal soil moisture and runoff patterns in three out of the four locations where soil moisture was monitored continuously in 10 and 30 cm depth and runoff was collected
250 using OFDs. Only for one site (S4) the RRP model underestimated runoff risks (Hahn et al., 2013).

3.1.2 DoRP

Discharge predictions of the DoRP model are based on an extremely simple model concept of a limited storage volume characteristic for each soil type (see above). Despite its simplicity and the lack of any calibration, the predicted discharge volumes for different events correlated reasonably
255 well with the measured values. However, the slope deviates from the 1 to 1 line in both catchments with an underestimation for the Stägbach catchment and an overestimation for the Lippenrütibach (Fig. 3). In contrast to the RRP model, DoRP predicts discharge only in direct response to rainfall events, and thus, due to the small number of events, no reliable comparison with discharge measurements was possible for the Stägbach catchment. For the Lippenrütibach catchments more runoff
260 events were available to compare DoRP predictions with measurements. The model thus proved to be useful to predict the discharge of an event relative to discharge of other events from the same catchment, but cannot be used to predict discharge from another catchment without re-calibration. This is in contrast to the RRP model that yielded satisfactory results without site-specific calibration. Given the large percentages of STO2 areas in the two catchments (Table 2), the runoff contributing

265 area predicted by the DoRP model increased sharply as soon as rainfall plus antecedent moisture exceeded the storage capacity of this class of soils. The difference between simulated and measured discharge increased with increasing event size (Fig. 3) suggesting that either storage capacities need to be calibrated or more differentiation between locations in STO2 is needed.

Surface overland flow from soils of storage class 2 (SOF2) was the most prevalent runoff category in 270 both catchments (Fig. 2B). Three of the four permanent measurement stations (S2, S3 and S4) in the Stägbach catchment were situated on soils of this category, while S1 was classified as DP. Indeed, S1 was the only station where surface runoff was never detected during the whole monitoring period and where the ground water table never rose above a level of 0.5 m below the soil surface. Thus, in contrast to the RRP model, the DoRP model differentiated correctly between S1 and S4, but not 275 between S2, S3 and S4, although S2, S3 and S4 also showed substantial differences in their soil moisture and runoff regimes (Hahn et al., 2013). The fact that S2, S3 and S4 were in line with RRP predictions suggests that topography was in general a better predictor for the runoff responsiveness of a location to rainfall events than hydraulic soil properties.

3.1.3 Comparison of RRP and DoRP

280 The highest observed P losses occurred during the largest event in each of the two catchments during the study period. Thus, we used these two events (Fig. 1) to compare the DoRP assessment of the soils of the two catchments with risk classes determined by the RRP model. In general, the extreme risk classes (very high or very low) agree well between the two approaches. On the one hand, very high fast flow risks according to the RRP model were strongly associated with soils of low (STO1) 285 to medium (STO2) storage capacity (Table 3). On the other hand, the DoRP storage classes DP and STO3 primarily fell into the categories of low fast flow risks according to RRP (Table 4). The intermediate categories in each approach were distributed across the classes of the other approach in much more homogeneous manner. This holds true especially for the DoRP storage class STO2, which contains for example 45.6% of RRP low risk class areas but also 29.7% of very high risk during 290 the largest runoff event in the Lippenrütibach catchment (Table 4). The STO2 class dominates in both study areas, which explains why DoRP allows for little spatial differentiation.

One reason for the relatively good match between the two models regarding the extreme risk classes is that all STO1 soils were classified as poorly drained soils in the RRP model and that locations with low soil water storage capacity tended to have large λ values (Fig. 4). In line with this observation, 295 the Kruskal-Wallis test, which was used here because of outliers and non-normal distribution, revealed that the mean TI values of the storage classes were significantly different. These relationships can be understood as an expression of the dependence of soil formation processes on topography and position within the landscape.

3.1.4 Comparison of RRP and SCIMAP

300 In contrast to the RRP and the DoRP model, SCIMAP cannot be used to predict discharge since it is a time integrated rather than dynamic model and so does not make discharge predictions. However, it is the only one of the three approaches that accounts for connectivity when identifying runoff risks within a catchment. This aspect however, does not seem to play a major role in our study areas. A comparison between the topographic index λ and network index NI reveals only minor differences
305 (see supplementary information SI-1). Accordingly, it is not surprising that the spatial NI patterns are similar to the RRP risk classes as can be seen in Fig. 2 for the largest event during the respective monitoring periods for the two study areas.

Figure 5A shows that there were close relationships between the NI and the risk of fast flow predicted by the RRP model. The scatter is due to pixels where the topographic index λ differs from the
310 network index NI. For such locations the risk might be overestimated by the RRP model, unless there is a direct connection to a stream that does not directly depend on topography, e.g. a connection via tile drains. In Figure 5B these pixels were not displayed to avoid the influence of connectivity when comparing the RRP hydrological risk with the SCIMAP connection risk.

As should be expected, well-drained locations were predicted by the RRP model to have lower fast
315 flow risk than poorly drained locations with the same NI (Fig. 5). For the Lippenrütibach catchment this observation was even more pronounced.

Because SCIMAP yields a static risk assessment while RRP predicts risk for fast flow as a function of time it is useful to compare the SCIMAP prediction with RRP predictions for discharge events of different magnitude (Fig. 5B). For small events, the SCIMAP risk predictions are systematically
320 larger than for those resulting from RRP (all RRP results lay below the 1:1 line in Fig. 5B). For large events, the situation is more complex. According to the RRP model, a considerable areal fraction has a higher risk than according to SCIMAP. On well drained soils this relationship turns to the opposite for RRP low risk areas. For poorly drained soils, RRP yields systematically higher risk values than SCIMAP. The average RRP risk over the monitoring period does not show a risk of 1
325 because it is an average of storm flow events and low flow. Even after rescaling the RRP risk to range between 0 and 1 to enable a fair comparison the average RRP risk predictions are always lower than the SCIMAP risk predictions. The average risk integrates periods of large and small storm flows as well as very low runoff events and therefore continues to increase until quite high NI values are reached (Fig. 5A). At a SCIMAP connection risk of 1 the average RRP risk ranges between 0.2 and
330 0.6 (Fig. 5B). This is because the scaling between the 5% and 95% quantiles does not enable a fine enough differentiation between high NI values. Therefore, SCIMAP differentiates in space less than the dynamic RRP model. Hence, SCIMAP agrees better with RRP predictions for larger events.

3.1.5 Comparison of all three model predictions

Despite the differences between RRP and SCIMAP described above, their spatial predictions of areas prone to fast flow processes and hence to DRP losses are much more similar to each other than
335 to the DoRP risk predictions (Fig. 2). Many topographic features where RRP and SCIMAP predict a high risk for runoff generation are not reflected at all in the DoRP model. Despite a more refined soil classification such features may get lost in the DoRP classification. Interestingly, this is not always the case. The two SOF1 areas (DoRP) in the Stägbach catchment for example, appear very
340 clearly in both RRP and SCIMAP. In general however, DoRP seems rather coarse in its classification compared to differentiation indicated by the other two models. Interestingly, the incorporation of soil information, which is part of the RRP model concept, caused very little change in spatial patterns relative to SCIMAP, which is solely based on topographic information. Accounting for soil drainage classes in the RRP approach resulted only in two small differences between Fig. 2A and C, observed
345 in the south western part of LIP and the north eastern part on Stäg. Accounting for connectivity (SCIMAP) had little influence on the spatial pattern of predicted hydrological risk areas relative to those predicted based on the original topographic index (Fig. 2A+C).

3.2 Critical areas for phosphorus losses

3.2.1 Comparison of RRP and SCIMAP

350 As shown by Hahn et al. (2013), the RRP model produced good predictions of P loads at the outlet of the Stägbach catchment and one of its sub-catchments. Spatial RRP predictions of runoff risks were also in good (qualitative) agreement with local measurements of soil moisture, groundwater levels and surface runoff. This gives confidence that the RRP model reflects the main processes and is a valid tool to delineate CSA for P exports into the streams of the study catchments. Given the good
355 performance of the RRP model, it is interesting that SCIMAP predicted similar areas with high risk of DRP loss and that this agreement appears to be better for high-runoff events (κ Stäg: 0.54; κ LIP: 0.68) than for the average DRP load during the simulation period (κ Stäg: 0.26; κ LIP: 0.29; see also Fig. 6 and 7, and Table 5). Thus SCIMAP appears to have good potential to be used at least as a first screening tool for the identification of critical source areas.

3.2.2 Relationship between NI and the connection risk used in SCIMAP

The original SCIMAP model prescribes a static linear relationship between NI and the connection risk p_x^c from 0 at the 5% NI quantile to 1 at the 95% quantile. This relationship is considered time invariant and it is based on the assumption that the least connected 5% fraction of a catchment never connects, while the most connected 5% fraction always connects to a stream. This approach has
365 three major limitations. Firstly, the comparison with the RRP model shown above suggests that the relationship between NI and connection risk is not time invariant but that SCIMAP predictions

mainly reflect larger events in our study areas. Secondly, the assumption that 5% of the catchment is always connected and 5% is never connected makes the method insensitive to these parts of the catchment. Assuming that areas with very low NI values never connect is reasonable for single runoff events and probably also for most monitoring periods. Assuming that areas with the highest 5% of NI values always connect is appropriate for large events, but not necessarily for aggregated risks over a period of time or for small events (Fig. 5B). This can be seen in Figures 7A and 7C, which show a considerable scatter for SCIMAP locational risks. The scatter was much less when the 5% assumption was relaxed and the connection risk assumed to scale linearly with NI up to its maximum value (Fig. 8), accounting for the fact that there were significant differences in connectivity even within the most connected 5% fraction of our catchment. While areas close to the catchment outlet characterized by very high λ and NI values frequently contributed to runoff according to the RRP model, even during very small events, areas further upstream, where the λ and NI values were lower but still within the top 5%, contributed runoff much less frequently. Extending the linear NI/risk scaling up to the maximum NI enabled differentiation between these areas. A third major limitation of the original SCIMAP approach is that by normalizing the generation risk and NI values between zero and one the model can predict risks at a given location only relative to the risks at other locations within the same catchment. To enable comparisons between different catchments, we normalized the generation risk (source factor) and delivery risk (transport factor) by setting a common upper limit for all catchments. For the source factor we simply used the maximum value of all catchments for this purpose. For the transport factor it was less straightforward. The highest NI value (NI_{max}) of the two catchments studied here was 20 and the lowest (NI_{min}) was 4.7. The 5% quantile of all NI values was 6 and based on our RRP model predictions the runoff risk of cells with NI values lower than 6 can be neglected. Thus, we set the transport factor to 0 at $NI \leq 5\%$ quantile and to 1 at $NI \geq NI_{max}$ and to vary linearly with NI between these limits. The locational risk calculated with these 'globally' scaled source and transport factors ranged between 0 and 0.4. Using the RRP results as a reference, the global locational risk was in better agreement with the average DRP loads over the whole monitoring period (κ Stäg: 0.29; κ LIP: 0.45; Fig. 8) than the original locational risks with catchment-specific normalization (κ Stäg: 0.26; κ LIP: 0.30; Fig. 7). Since the catchments had similar soil P status, this improvement can be attributed to the modified relationship between NI and delivery risk.

4 Discussion

The performance of the three models indicates that a large amount of useful hydrological information can be extracted for making predictions on P export risks and CSAs from widely available data sources (soil map, geological map, topography). The results further indicate for our study area that runoff generation, and associated P export risks depend more on topography than on soil charac-

teristics, when the information is solely extracted from conventional maps. Although neglecting topographic connectivity and using a rather crude binary soil classification by drainage capacity, the RRP model predicted spatial distributions of runoff risks that were in good (qualitative) agree-
405 ment with local measurements of soil moisture, groundwater level and surface runoff (Hahn et al., 2013). The DoRP predictions showed surprisingly close relationships with measured discharges at the catchment outlets. However, the DoRP model could not sufficiently differentiate between lo-
cations, because large areas fell into the same runoff category (STO2), while there was substantial variation in runoff risks within this category due to topography, which was accounted for by the RRP
410 model and SCIMAP. The SCIMAP predictions were very similar to the RRP simulations, indicating that there arent many areas with reduced topographic connectivity in the study catchments.

While the RRP model was found to be well suited to the identification of high-risk areas in our catchments, taking account of connectivity and differentiating between more soil types may be cru-
cial in other types of catchments. This could be achieved most simply by combining predictions of
415 the three models. The DoRP model can help to better distinguish between sites assigned to the same drainage class by the RRP model, as shown for the example of the measurement locations S1 and S4. Thus, in targeting mitigation measures to areas identified as CSAs using RRP, it may be worth checking for the runoff category assigned to these areas by the DoRP model, before priorities are determined.

420 SCIMAP, which accounts for topographic connectivity along flow pathways can be used to identify areas whose connectivity with the stream may be overestimated by the other two models due to the inherent assumption of unrestricted connectivity. In case of a DP zone, it is unlikely that the area is connected to a stream or lake by an artificial drainage system, which means that the risk of P export is less than predicted by the latter models. In STO1 and STO2 zones, it may be necessary to check
425 for the presence of a drainage network before correcting the risk prediction.

4.1 Connectivity

Many studies highlight the importance of accounting for connectivity when identifying CSAs (Doody et al., 2012; Doppler et al., 2012; Frey et al., 2009; Lane et al., 2009; Srinivasan and McDowell, 2007). However, connectivity has been defined in different ways that are relevant for the interpreta-
430 tion. The Network Index NI used in SCIMAP assumes that the connectivity along a flow line breaks down if the topographic index values λ along a flow line have a local minimum. Given the definition of λ this implies that the connectivity is lost if a section of the flow path is steeper than the upslope part. Lane et al. (2009) illustrated the potential of the NI to generalize information about the hydro-
logical connectivity between locations in areas where topography is the dominant factor. Another
435 approach assumes that connectivity is only lost, if water cannot flow further downstream because it is retained in a sink area. While internal sink areas in catchments are often filled by default by conventional GIS software, detailed analysis have revealed that internal sinks may be essential ele-

ments of catchments. This may be caused by natural processes in landscapes like the pothole region in North America or may be caused by anthropogenic influence on topography in the Swiss Plateau
440 (Doppler et al., 2012; Frey et al., 2009).

In summary, the two concepts assume different mechanisms behind the loss of connectivity. The first concept (NI) assumes that re-infiltration due to an increased gradient prevents surface runoff downhill. The second approach assumes the opposite in that ponding occurs because a topographic barrier prevents any flow downhill on the soil surface. Which process is more relevant for any catch-
445 ment depends on the specific conditions. The steeper the terrain the less probable are formation of sinks. This was the reason why the NI has been used in our study area, which is steeper than other parts of the Swiss Plateau where sinks may be very relevant. Due to its simplicity, the NI has a high potential to be widely applied especially within the scope of the WFD, as shown for Ireland by (Doody et al., 2012; Wall et al., 2011). This potential is still limited by insufficient knowledge about
450 the relationship between the NI and the probability of a site to connect to the stream network of a catchment over time (Lane et al., 2009; Reaney et al., 2011). Here we related the NI to the runoff risk as predicted with the RRP model. These relationships contain information about the catchment responses during the monitoring period that can be used to refine the scaling of NI/connection-risk relationships used in SCIMAP.

The reasonable match between RRP and SCIMAP predictions of CSAs for larger events in our study catchments suggests that SCIMAP can be used as a screening tool for CSA delineation in catchments where CSAs are primarily determined by topography. This is of particular interest in regions where dynamic models cannot be applied due lack of discharge data. While SCIMAP was originally developed to predict relative risks for individual catchments, a modified version of SCIMAP can be used
460 to compare P export risks among catchments and thus help to identify the most critical catchments within watersheds, given a homogeneous distribution of rainfall. However, the transformation of NI into a connection risk needs further refinement to appropriately describe the hydrological risk. If applicable, RRP simulations can be used for this purpose in addition to field data. The stepwise linear relationship with zero risk up to the 5% NI quantile and a maximum risk level with no further
465 change at the 95% NI quantile, as proposed by Reaney et al. (2011), was found to be appropriate for storm events, but not for ensembles of different events. Based on our results, we recommend a relationship with a linear increase from the 5% quantile up to the maximum NI value in the latter case.

470 **4.2 Limitations of the model approaches**

In contrast to the other two models, SCIMAP is not a dynamic model and thus cannot be used to predict how discharge and associated P losses vary between rainfall events of different magnitudes. However, it can be very useful to identify areas with a high potential for P losses in regions where

topography governs hydrology. The DoRP model in contrast is advantageous where the drainage
475 properties of soils are the dominant factor. The RRP model is less refined than the DoRP model with
respect to differentiation between soils and does not account for topographic connectivity. However,
combining the most basic of information on soil hydrology with the topographic information within
a process-based framework seems to predict the relevant hydrological and DRP export processes
in our study catchments well. The simplifications regarding the representation of the hydrologic
480 processes in the RRP approach are an advantage in practical applications, but by necessity also
come with limitations.

4.2.1 Hydrological drivers

The generation of fast flow, which includes all kinds of quickly responding runoff in the RRP model,
is bound to λ and thus is more likely in wet areas. In reality, however, Infiltration Excess Runoff
485 (IER) can also contribute significantly to the total runoff and transport DRP and PP to a stream
(Doppler et al., 2012; Srinivasan et al., 2002), depending on rainfall patterns and soil properties.
For certain high resolution soil maps it is possible to identify areas prone to HOF using the DoRP
model (Schmocker-Fackel et al., 2007). Unfortunately, most soil maps do not contain sufficient in-
formation. In addition, one has to consider that relevant properties like the infiltration capacity of
490 soils may strongly depend on management practices and hence vary in time. Srinivasan et al. (2002)
for example reported that the occurrence of IER during field experiments was scattered, disjunct
and transient, which makes the prediction of areas prone to IER difficult. IER may also be due to
water repellent soil surfaces. While repellency is common for many land-use types with permanent
vegetation cover in humid temperate climates (Doerr et al., 2006), we assume that it did not cause a
495 lot of runoff that reached the stream network.

Another limitation of the approaches discussed here relates to the assumption that surface topog-
raphy reflects the relevant gradients controlling water fluxes. If the topography of an impermeable
layer differs significantly from surface topography runoff may occur on unexpected areas. Like the
models evaluated by Srinivasan and McDowell (2007), the models compared in our study do not en-
500 able users to recognize areas where subsurface flows can potentially emerge to the surface as seeps
and springs on steep hillsides (Srinivasan and McDowell, 2007). Zheng et al. (2004) reported that
particularly high P concentrations were associated with this type of runoff.

In principle, the RRP and DoRP models do not only account for surface but also for subsurface flow,
which can also carry substantial amounts of P to a water body (Kleinman et al., 2011b). In practice,
505 the difficulty is to account for artificial drainage systems, which can contribute substantially to DRP
losses (Stamm et al., 1998, 2002; Watson and Matthews, 2008). Drains can connect areas that ap-
pear disconnected on the basis of topographic analysis, and thus the lack of drainage data can be a
major problem for CSA determination. However, in these kind of landscapes, surface runoff and tile
drain flow are often not separate flow processes but they may occur in sequence: flow may start as

510 surface runoff and gets intercepted by e.g. macropores connected to tile drains (Stamm et al., 2002;
Doppler et al., 2012). The drained area amounts to approximately 10% of the agricultural area in
the Lippenrütibach catchment and to around 15% in the Stägbach catchment. Field inspections and
measurements in the Stägbach catchment revealed that locations predicted to be wet were indeed wet
and that surface runoff from some of these locations occurred although they were in close proximity
515 to drains. Thus, our results show that even in presence of drainage systems, topography may still
provide crucial information on runoff generation risks and CSA delineation. Because the combina-
tion of surface runoff and macropore flow to tile drains is part of the fast flow component of RRP, the
influence of tile drains is accounted for during the calibration process. To assess the potential role of
IER and artificial drainage systems, it may be helpful to complement available ground information
520 through field visits and interviews of local farmers, as suggested by Frey et al. (2011). The models
discussed here can be very useful in guiding the collection of such information.

4.2.2 Sources and types of P

This study focused on the prediction of P losses in form of DRP, because DRP can immediately be
taken up by algae (Dorioz et al., 2006; Sharpley, 1993; Sharpley et al., 1994) and thus represents the
525 main risk factor for eutrophication (Kleinman et al., 2011b). However, it should be recognized that
also particulate P (PP) can become bio-available, given specific physical-chemical dynamics (Dorioz
et al., 2006). Particulate P losses can be high, especially on arable land (Doody et al., 2012) or due
to bank erosion (Kronvang et al., 2012). There are several models that simulate sediment and PP
losses (Krueger et al., 2012). SCIMAP can also be adjusted to predict CSAs for sediment (Reaney
530 et al., 2011) and thus PP export. However, the RRP model was developed for grassland dominated
catchments and solely simulates DRP losses. Phosphorus losses from vegetation (Kleinman et al.,
2011b) and seasonal changes in P availability (Dorioz et al., 2006; Pote et al., 1999) were assumed
to be negligible. Furthermore, freshly applied manure or other fertilizers can be a relevant source
for P export (Withers et al., 2003). The RRP model is able to account for this source. However, as
535 the application of manure can be easily controlled by appropriate management, we did not consider
it for CSA delineation in this study.

5 Conclusion

The study demonstrates that a large amount of hydrological information needed for the prediction
of P export risks and CSAs can be extracted from widely available data sources. The comparison
540 of the predictions obtained with the three tools to each other and to the available experimental data
indicates that the location and extent of CSAs within the study catchments was more dependent on
topography than on variation in soil properties as they are represented in the DoRP model. SCIMAP
predicted CSAs particularly well for large storm events. It may be especially useful for screening

purposes, in particular in regions without discharge data. RRP simulations can be used to adapt
545 the relationship between NI and connection risks used in SCIMAP to the particular characteristics
of a catchment or region. The stepwise linear relationship with zero risk up to the 5% NI quantile
and a maximum risk level with no further change at the 95% NI quantile, as proposed by Reaney
et al. (2011), was found to be appropriate for storm events, while we recommend a relationship
with a linear increase from the 5% quantile up to the maximum NI value for a multitude of events.
550 Two major problems in predicting P export risks and delineating CSAs in general are the difficulty
to account for IER and the limited availability of data on tile drainage systems. To cope with this
problem, we suggest that model predictions are complemented by ground information obtained from
field visits and interviews of local farmers, as suggested by Frey et al. (2011). The models can be
very useful in guiding the collection of such information.

555 *Acknowledgements.* The study was carried out within the framework of the COST Action 869, and was fi-
nancially supported by the Swiss State Secretariat for Education and Research SER. We are thankful for the
valuable comments from the reviewers and for the financial support from the COST office for a short term
scientific mission, which gave the impulse for this model comparison.

References

- 560 AGBA: Einzugsgebiet des Sempacher Sees, Übersichtskarte Luzern, Lucerne, Switzerland, 1993.
- Beven, K. and Kirkby, M.: A physically based, variable contributing area model of basin hydrology, *Hydrol. Sci. Bull.*, 24, 43–69, 1979.
- Buda, A. R., Koopmans, G. F., Bryant, R. B., and Chardon, W. J.: Emerging technologies for removing nonpoint phosphorus from surface water and groundwater: introduction, *J. Environ. Qual.*, 41, 621–627, 2012.
- 565 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.
- Cohen, J.: Nominal scale agreement with provision for scaled disagreement or partial credit, *Psychological Bulletin*, 70, 213–220, 1968.
- Doerr, S. H., Shakesby, R. A., Dekker, L. W., and Ritsema, C. J.: Occurrence, prediction and hydrological
- 570 effects of water repellency amongst major soil and land-use types in a humid temperate climate, *Eur. J. Soil Sci.*, 57, 741–754, 2006.
- 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.
- 575 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, 10.5194/hess-16-1947-2012, 2012.
- Dorioz, J. M., Wang, D., Poulénard, J., and Trevisan, D.: The effect of grass buffer strips on phosphorus dynamics – a critical review and synthesis as a basis for application in agricultural landscapes in France,
- 580 *Agr. Ecosyst. Environ.*, 117, 4–21, 2006.
- Dubrovsky, N. M., Burow, K. R., Clark, G. M., Gronberg, J. M., Hamilton, P. A., Hitt, K. J., Mueller, D. K., Munn, M. D., Nolan, B. T., Puckett, L. J., Rupert, M. G., Short, T. M., Spahr, N. E., Sprague, L. A., and Wilber, W. G.: The quality of our Nation’s waters – nutrients in the Nation’s streams and groundwater, 1994–2004, Tech. rep., US Geological Survey Circular 1350, 2010.
- 585 Frey, M. P., Schneider, M. K., Dietzel, A., Reichert, P., and Stamm, C.: Predicting critical source areas for diffuse herbicide losses to surface waters: role of connectivity and boundary conditions, *J. Hydrol.*, 365, 23–36, 2009.
- Frey, M. P., Stamm, C., Schneider, M. K., and Reichert, P.: Using discharge data to reduce structural deficits in a hydrological model with a Bayesian inference approach and the implications for the prediction of critical
- 590 source areas, *Water Resour. Res.*, 47, W12529, doi10.1029/2010WR009993, 2011.
- Gburek, W. J. and Sharpley, A. N.: Hydrologic controls on phosphorus loss from upland agricultural watersheds, *J. Environ. Qual.*, 27, 267–277, 1998.
- Gupta, V. K., Sorooshian, S.: The relationship between data and the precision of parameter estimates of hydrologic models, *J. Hydrol.*, 81, 57–77, 1985.
- 595 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.
- Hahn, C., Prasuhn, V., Stamm, C., Lazzarotto, P., Evangelou, M. W. H., and Schulin, R.: Prediction of dissolved reactive phosphorus losses from small agricultural catchments: calibration and validation of a parsimonious

- model, *Hydrol. Earth Syst. Sci.*, 17, 3679–3693, 10.5194/hess-17-3679-2013, 2013.
- 600 Heathwaite, A. L., Reaney, S., and Lane, S.: Understanding spatial signals in catchments: linking critical areas, identifying connection and evaluating response, *Diffuse phosphorus loss: risk assessment, mitigation options and ecological effects in river basins*, The 5th International Phosphorus Workshop, Silkeborg, Denmark, 2007.
- Heathwaite, L., Sharpley, A., and Bechmann, M.: The conceptual basis for a decision support framework to
605 assess the risk of phosphorus loss at the field scale across Europe, *J. Plant Nutr. Soil Sc.*, 166, 447–458, 2003.
- Hering, D., Borja, A., Carstensen, J., Carvalho, L., Elliott, M., Feld, C. K., Heiskanen, A.-S., Johnson, R. K., Moe, J., Pont, D., Solheim, A. L., and van de Bund, W.: The European Water Framework Directive at the age of 10: A critical review of the achievements with recommendations for the future, *Sci. Total Environ.*,
610 408, 4007–4019, 2010.
- Herzog, P.: Sanierung des Baldegger Sees, Auswertung der Zufluss-Untersuchungen 2000 bis 2004, Tech. rep., Kanton Luzern, Umwelt und Energie, Lucerne, Switzerland, 2005 (in German).
- Kirkby, M.: *Hydrograph Modelling Strategies*, Heinemann, London, 69–90, 1975.
- Kleinman, P. J. A., Allen, A. L., Needelman, B. A., Sharpley, A. N., Vadas, P. A., Saporito, L. S., Folmar, G. J.,
615 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., Saporito, L. S., Buda, A. R., and Bryant, R. B.: Application of manure to no-till soils: phosphorus losses by sub-surface and surface pathways, *Nutr. Cycl. Agroecosys.*, 84, 215–227, 2009.
- 620 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.
- 625 Kronvang, B., Audet, J., Baattrup-Pedersen, A., Jensen, H. S., and Larsen, S. E.: Phosphorus Load to Surface Water from Bank Erosion in a Danish Lowland River Basin, *J. Environ. Qual.*, 41, 304–313, 2012.
- Krueger, T., Quinton, J. N., Freer, J., Macleod, C. J. A., Bilotta, G. S., Brazier, R. E., Hawkins, J. M. B., and Haygarth, P. M.: Comparing empirical models for sediment and phosphorus transfer from soils to water at field and catchment scale under data uncertainty, *Eur. J. Soil Sci.*, 63, 211–223, 2012.
- 630 Lane, S. N., Brookes, C. J., Kirkby, A. J., and Holden, J.: A network-indexbased version of TOPMODEL for use with high-resolution digital topographic data, *Hydrol. Process.*, 18, 191–201, 2004.
- Lane, S. N., Brookes, C. J., Heathwaite, A. L., and Reaney, S.: Surveillant science: challenges for the management of rural environments emerging from the new generation diffuse pollution models, *J. Agr. Econ.*, 57, 239–257, 2006.
- 635 Lane, S. N., Reaney, S. M., and Heathwaite, A. L.: Representation of landscape hydrological connectivity using a topographically driven surface flow index, *Water Resour. Res.*, 45, W08423, doi10.1029/2008wr007336, 2009.
- Lazzarotto, P.: Modeling phosphorus runoff at the catchment scale, Ph. D. thesis, ETH Zurich, Zurich, Switzerland

land, 2005.

- 640 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.
- 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.
- Peyer, K., Brunner, H., Wüst, O., and Plüss: *Bodenkarte Hochdorf, Eidg. Forschungsanstalt für landw. Pflanzen-*
645 *bau*, Zurich, Switzerland, 1983.
- Pionke, H. B., Gburek, W. J., Sharpley, A. N., and Zollweg, J. A.: Hydrological and Chemical Controls on Phos-
phorus Loss from Catchments, chap. 225–242, pp. 225–242, CAB International Press, Cambridge, workshop
September 29–31, 1995, Wexford, Ireland UK Ministry of Agriculture, Fisheries, and Food, 1997.
- Pionke, H. B., Gburek, W. J., and Sharpley, A. N.: Critical source area controls on water quality in an agricul-
650 *tural watershed located in the Chesapeake Basin, Ecol. Eng.*, 14, 325–335, 2000.
- Pote, D. H., Daniel, T. C., Nichols, D. J., Sharpley, A. N., Moore, P. A., Miller, D. M., and Edwards, D. R.:
Relationship between phosphorus levels in three ultisols and phosphorus concentrations in runoff, *J. Environ.*
Qual., 28, 170–175, 1999.
- Radcliffe, D. E., Freer, J., and Schoumans, O.: Diffuse phosphorus models in the united states and europe:
655 *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 Founda-
tion for Statistical Computing, Vienna, Austria, 2007.
- Reaney, S. M., Lane, S. N., Heathwaite, A. L., and Dugdale, L. J.: Risk-based modelling of diffuse land use
impacts from rural landscapes upon salmonid fry abundance, *Ecol. Model.*, 222, 1016–1029, 2011.
- 660 Roberts, W. M., Stutter, M. I., and Haygarth, P. M.: Phosphorus retention and remobilization in vegetated buffer
strips: a review, *J. Environ. Qual.*, 41, 389–399, 2012.
- Rodriguez, H. G., Popp, J., Maringanti, C., and Chaubey, I.: Selection and placement of best management prac-
tices used to reduce water quality degradation in Lincoln Lake watershed, *Water Resour. Res.*, 47, W01507,
doi10.1029/2009wr008549, 2011.
- 665 Scherrer, S. and Naef, F.: A decision scheme to indicate dominant hydrological flow processes on temperate
grassland, *Hydrol. Process.*, 17, 391–401, 2003.
- Schmocker-Fackel, P., Naef, F., and Scherrer, S.: Identifying runoff processes on the plot and catchment scale,
Hydrol. Earth Syst. Sci., 11, 891–906, 10.5194/hess-11-891-2007, 2007.
- Schoumans, O. F., Silgram, M., Groenendijk, P., Bouraoui, F., Andersen, H. E., Kronvang, B., Behrendt, H.,
670 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.
- Schoumans, O. F., Chardon, W. J., Bechmann, M., Gascuel-Oudou, C., Hofman, G., Kronvang, B., Litaor, M. I.,
Lo Porto, A., Newell-Price, P., and Rubaek, G.: Mitigation options for reducing nutrient emissions from
675 *agriculture. A study amongst European member states of Cost action 869, Tech. rep.*, Alterra Wageningen
UR, Wageningen, the Netherlands, 2011.
- 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.

- 680 Sharpley, A. N.: Assessing phosphorus bioavailability in agricultural soils and runoff, *Fert. Res.*, 36, 259–272, 1993.
- Sharpley, A. N., Chapra, S. C., Wedepohl, R., Sims, J. T., Daniel, T. C., and Reddy, K. R.: Managing agricultural phosphorus for protection of surface waters – issues and options, *J. Environ. Qual.*, 23, 437–451, 1994.
- Srinivasan, M. S. and McDowell, R. W.: Hydrological approaches to the delineation of critical-source areas of runoff, *New Zeal. J. Agr. Res.*, 50, 249–265, 2007.
- 685 Srinivasan, M. S., Gburek, W. J., and Hamlett, J. M.: Dynamics of stormflow generation – a hillslope-scale field study in east-central Pennsylvania, USA, *Hydrol. Process.*, 16, 649–665, 2002.
- Stamm, C., Flüehler, H., Gächter, R., Leuenberger, J., and Wunderli, H.: Preferential transport of phosphorus in drained grassland soils, *J. Environ. Qual.*, 27, 515–522, 1998.
- 690 Stamm, C., Sermet, R., Leuenberger, J., Wunderli, H., Wydler, H., Flüehler, H., and Gehre, M.: Multiple tracing of fast solute transport in a drained grassland soil, *Geoderma*, 109, 245–268, 2002.
- Stutter, M. I., Chardon, W. J., and Kronvang, B.: Riparian buffer strips as a multifunctional management tool in agricultural landscapes: introduction, *J. Environ. Qual.*, 41, 297–303, 2012.
- 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.
- 695 Vadas, P. A., Srinivasan, M. S., Kleinman, P. J. A., Schmidt, J. P., and Allen, A. L.: Hydrology and groundwater nutrient concentrations in a ditch-drained agroecosystem, *J. Soil Water Conserv.*, 62, 178–188, 2007.
- Wall, D., Jordan, P., Melland, A. R., Mellander, P. E., Buckley, C., Reaney, S. M., and Shortie, G.: Using the nutrient transfer continuum concept to evaluate the European Union Nitrates Directive National Action Programme, *Environ. Sci. Policy*, 14, 664–674, 2011.
- 700 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.
- 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.
- 705 Yapo, P. O., Gupta, H. V., and Sorooshian, S.: Automatic calibration of conceptual rainfall-runoff models: sensitivity to calibration data, *J. Hydrol.*, 181, 23–48, 1996.
- Zheng, F. L., Huang, C. H., and Norton, L. D.: Surface water quality – effects of near-surface hydraulic gradients on nitrate and phosphorus losses in surface runoff, *J. Environ. Qual.*, 33, 2174–2182, 2004.

Table 1. Spatial extent of hydrological risk classes for a large and a small runoff event in the Lippenrütibach and the Stägbach catchment relative to the total agricultural area in %

	Low	Medium	High	Very high
	[%]			
Lippenrütibach 1999				
Small event in August	71	13	7	9
Large event in May	49	12	12	27
Stägbach 2010				
Small event in May	76	12	5	7
Large event in June	44	23	17	16

Table 2. Areal percentage of each dominant runoff process (DoRP) in the Lippenrütibach catchment and the Stägbach catchment [%] - relative to the total agricultural area. STO1 - low storage capacity, STO2 - intermediate storage capacity, STO3 - high storage capacity, DP - deep drainage

	STO1	STO2	STO3	DP
	[%]			
LIP	14	69	9	8
Stäg	6	61	11	23

Table 3. Amount of RRP risk class pixel within the DoRP storage classes in percent of the total amount of pixel within one risk class [%]. The areal extent of the RRP risk classes during the large runoff events (Lippenrütibach: May 1999, Stägbach: June 2010, see Fig. 1) was taken for comparison. STO1 - low storage capacity, STO2 - intermediate storage capacity, STO3 - high storage capacity, DP - deep drainage

DoRP storage classes					
	STO1	STO2	STO3	DP	sum
RRP risk classes					
	Lippenrütibach catchment				
low	3.1	65.1	16.2	15.6	100
medium	24.8	68.1	2.7	4.3	100
high	25.8	71.5	2.3	0.4	100
very high	23.4	75.6	0.9	0.1	100
	Stägbach catchment				
low	0.6	48.1	12.3	38.9	100
medium	2.5	65.3	14.7	17.6	100
high	6.5	75.6	9.7	8.2	100
very high	21.4	70.9	5.2	2.5	100

Table 4. Amount of RRP risk class pixel within the DoRP storage classes in percent of the total area of the storage classes [%]. The areal extent of the RRP risk classes during the large runoff events (Lippenrütibach: May 1999, Stägbach: June 2010) was taken for comparison. STO1 - low storage capacity, STO2 - intermediate storage capacity, STO3 - high storage capacity, DP - deep drainage

DoRP storage classes				
	STO1	STO2	STO3	DP
RRP risk classes	Lippenrütibach catchment			
low	10.6	45.6	90.1	92.6
medium	21.6	12.0	3.8	6.5
high	22.5	12.6	3.2	0.6
very high	45.3	29.7	2.9	0.3
sum	100	100	100	100
Stägbach catchment				
low	5.2	34.7	47.8	74.5
medium	10.4	24.6	29.8	17.6
high	20.4	21.4	14.7	6.1
very high	64.1	19.2	7.6	1.8
sum	100	100	100	100

Table 5. Weighted kappa values for pair-wise comparisons of RRP and SCIMAP predictions for P loss risks in the two study catchments, showing (i) that SCIMAP risk predictions were in better agreement with RRP model predictions for the event with highest runoff than for the average DRP load over the simulation period and (ii) that average DRP load predictions obtained with RRP were in better agreement with the global than with the original catchment-specific prediction of the locational risk using SCIMAP.

	Stägbach		Lippenrütibach	
	SCIMAP	SCIMAP global	SCIMAP	SCIMAP global
	locational risk	locational risk	locational risk	locational risk
Stägbach				
RRP - average DRP load	0.26	0.29		
RRP - DRP load at highest runoff event	0.54	0.56		
Lippenrütibach				
RRP - average DRP load			0.3	0.45
RRP - DRP load at highest runoff event			0.68	0.59

Figures

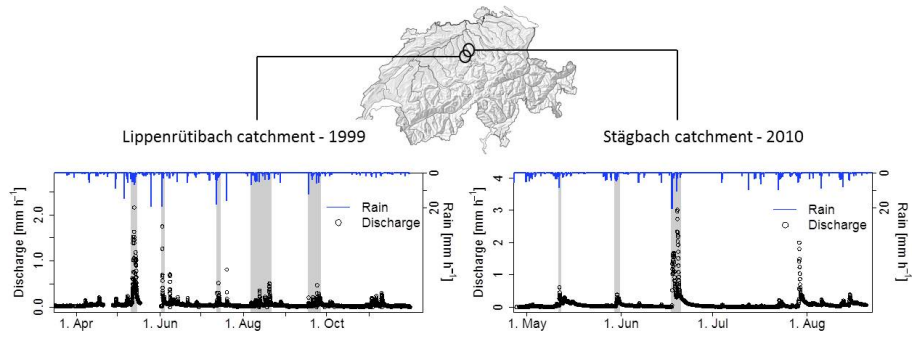


Fig. 1. Location of study areas and the rainfall and discharge characteristics for the respective monitored year. Grey shading indicates the events for which discharge was calculated using the Dominant Runoff Processes approach

RRP model - 2A

Hydrological risk areas during the largest event of the monitored year

- 0 - 0.2 low risk
- 0.2 - 0.5 medium risk
- 0.5 - 0.8 high risk
- 0.8 - 1 very high risk

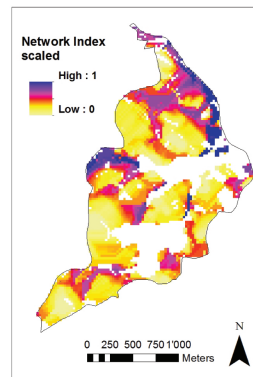
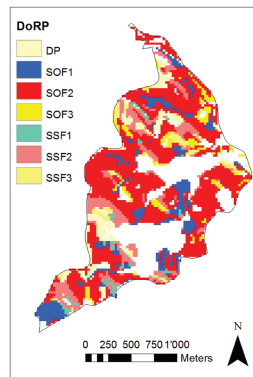
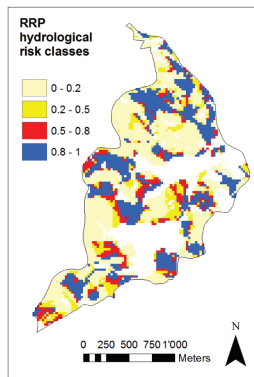
DoRP approach - 2B

- DP - deep percolation
- SOF - saturation overland flow
- SSF - subsurface flow
- 1 - 3 storage classes
- 1 - lowest storage capacity
- 3 - highest storage capacity

SCIMAP - 2C

connection risk = Network Index scaled between 5 % and 95 % quantiles

Lippenrütibach



Stägbach

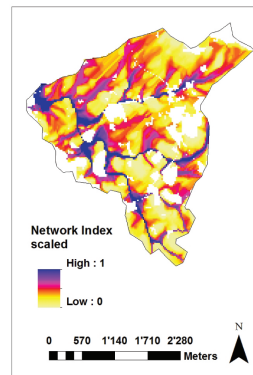
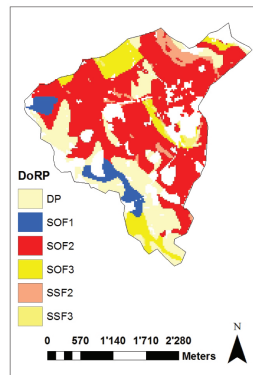
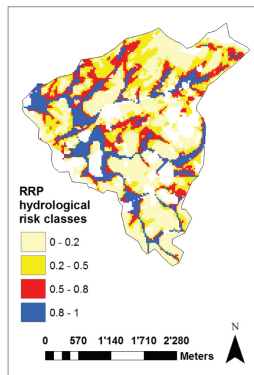


Fig. 2. Spatial distribution of hydrological risk areas as determined with the three models for the Lippenrütibach catchment and the Stägbach catchment

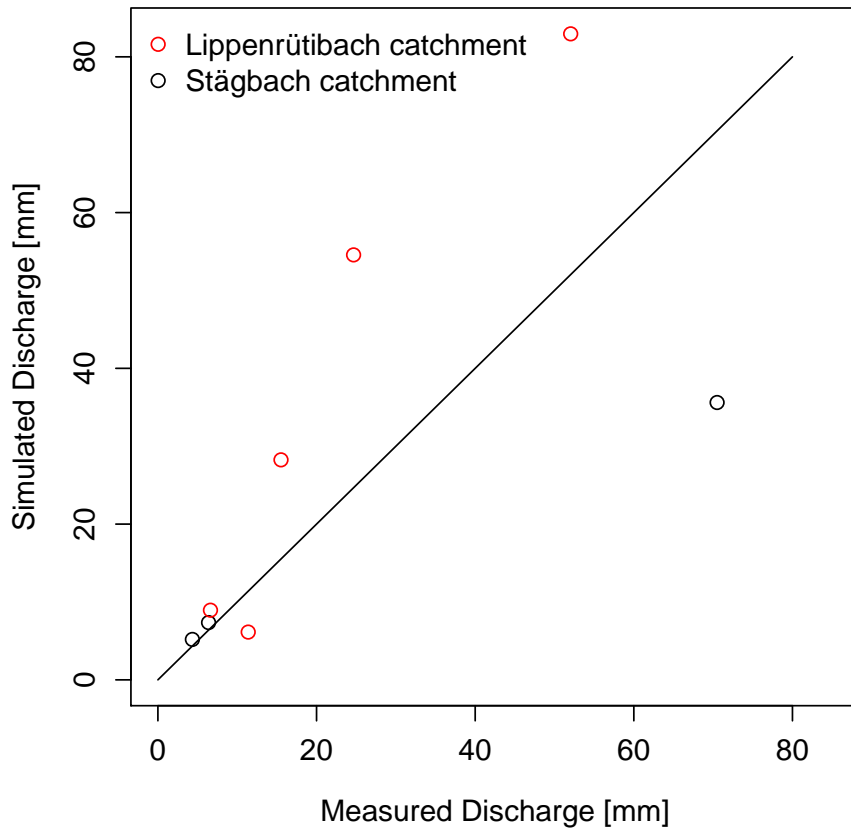


Fig. 3. Comparison of discharge estimations for runoff events in the year 1999 (Lippenrütibach catchment) and 2010 (Stägbach catchment) derived with the DoRP approach and the measured discharge. The respective runoff events are indicated in Figure 1 by grey shading.

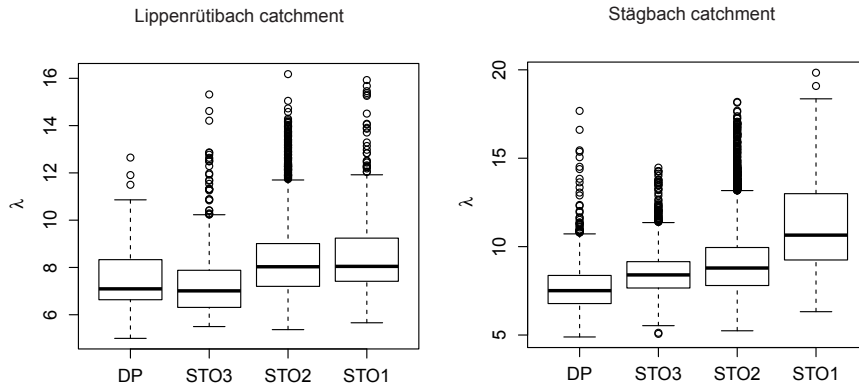


Fig. 4. Boxplots - the distribution of the Topographic Index λ within the DoRP storage classes. DP - Deep Percolation, STO1 - low storage capacity, STO2 - intermediate storage capacity, STO3 - high storage capacity

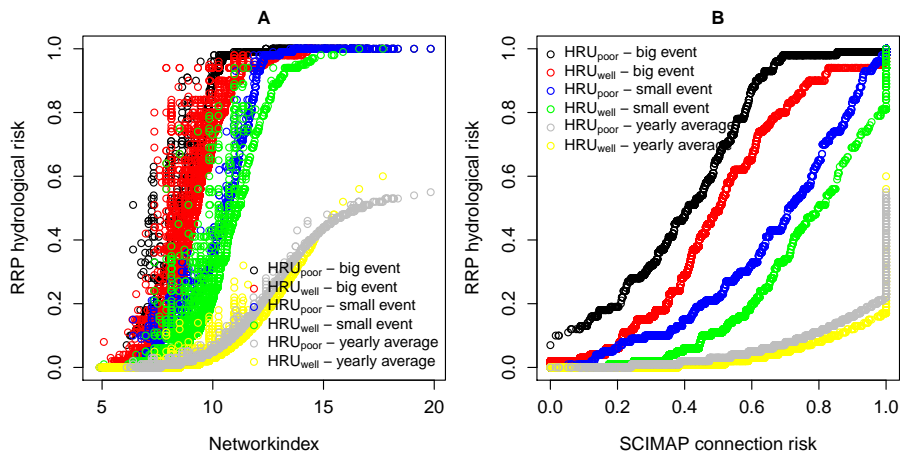


Fig. 5. RRP hydrological risk predictions for both hydrological response units (HRU_{well} and HRU_{poor}) during two different rainfall events and for the whole monitoring period versus the Network Index (5A) and versus the original SCIMAP connection risk (5B).

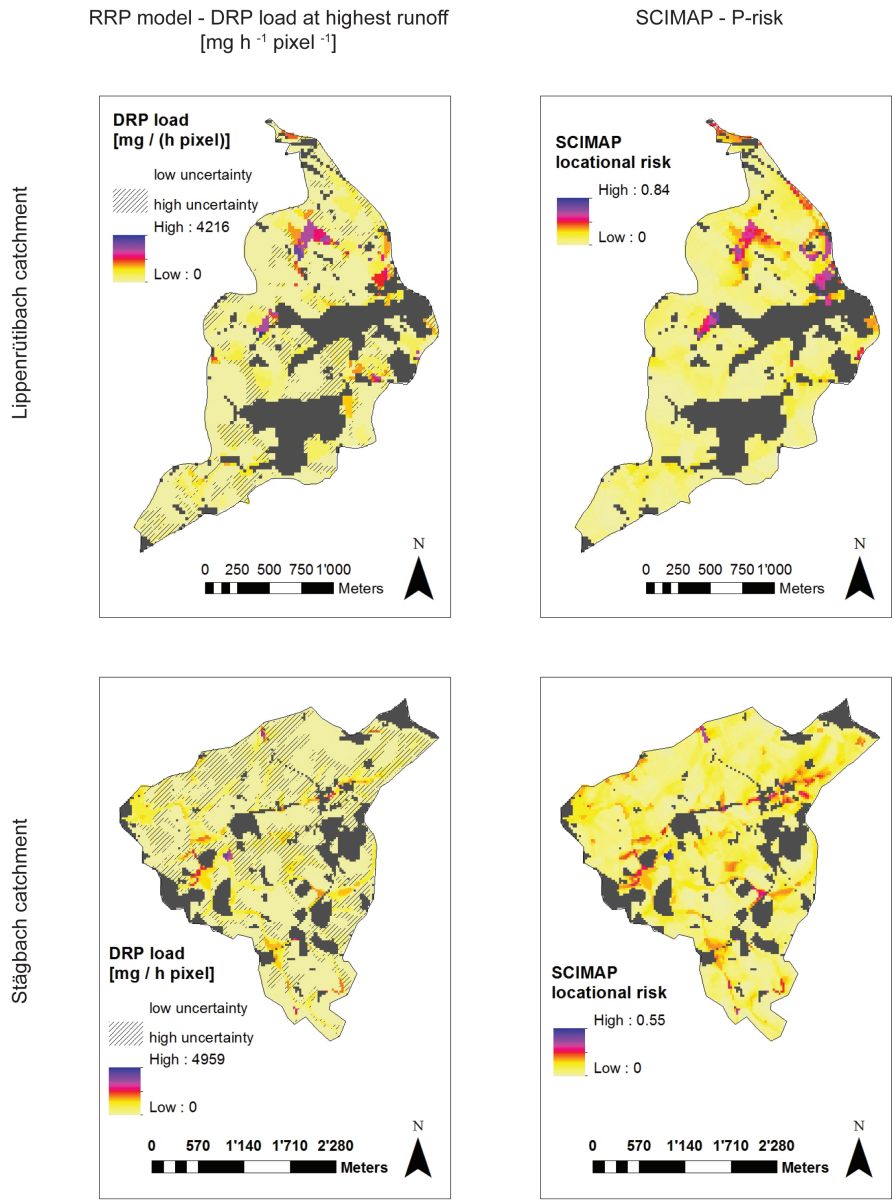


Fig. 6. Spatial model results for P losses as predicted with the RRP model and the SCIMAP model, grey areas are forested or urban areas. Areas for which less than 80% of the RRP simulations resulted in the same distribution of fast flow generation are hatched.

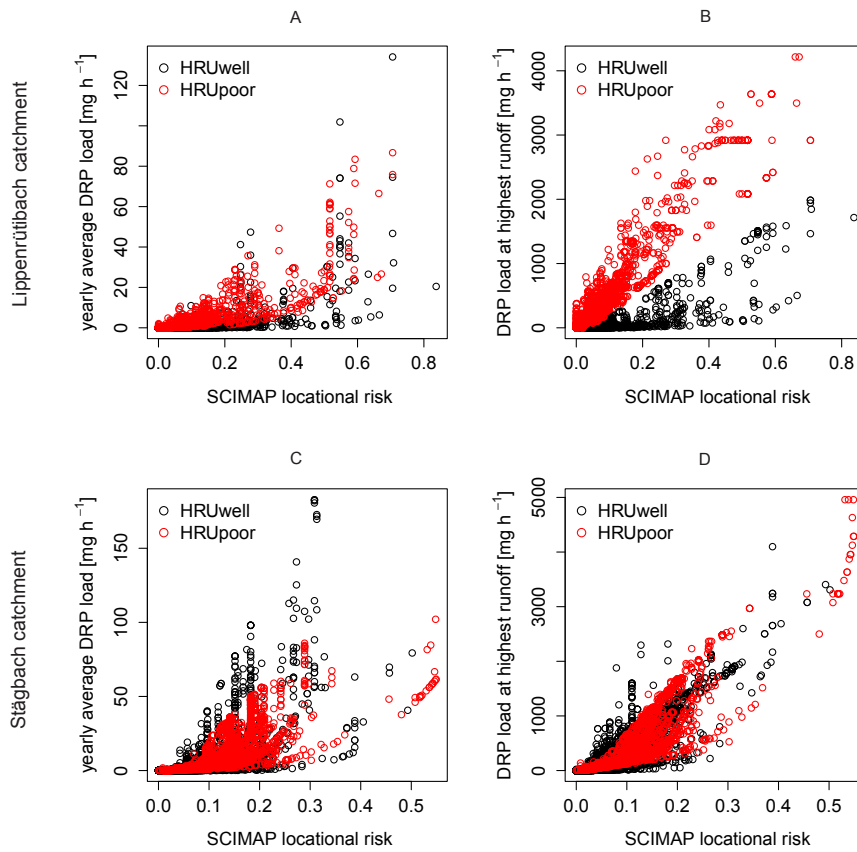


Fig. 7. Comparison of the P load calculated with the RRP model, that differentiates between well and poorly drained hydrological response units (HRU_{well} , HRU_{poor}), for the entire monitoring period and the highest runoff event with the locational risk for P losses predicted using the SCIMAP approach

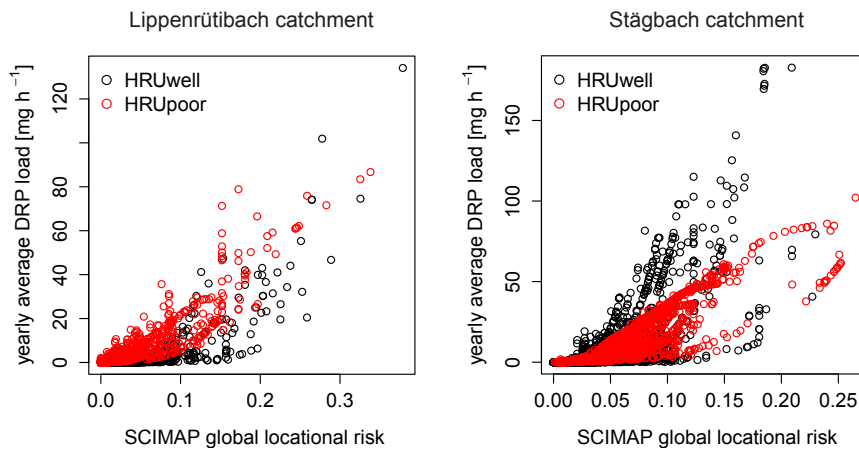


Fig. 8. Comparison between the average DRP loads for the whole monitoring periods predicted by the RRP model for both hydrological response units (HRU_{well} , HRU_{poor}) and the global SCIMAP locational risk estimations. For the global locational risk we set the delivery risk to 0 at Network Index $NI \leq 6$, the 5% quantile, and to 1 at $NI \geq 20$, the maximum NI of both catchments and scaled linearly for NI in between. The source factor (generation risk) was normalized by dividing with the maximum WSP of all catchments. In our catchments the difference to the original locational risk resulted mainly from the changed delivery risk delineation.