Geoderma 403 (2021) 115257

Contents lists available at ScienceDirect

Geoderma

journal homepage: www.elsevier.com/locate/geoderma

The time it takes to reduce soil legacy phosphorus to a tolerable level for surface waters: What we learn from a case study in the catchment of Lake Baldegg, Switzerland

Cäcilia von Arb [a], Sebastian Stoll [a], Emmanuel Frossard [b], Christian Stamm [c], Volker Prasuhn [a,*]

[a] Agroscope, Water Protection and Substance Flows, Switzerland
[b] ETH Zurich, Institute of Agricultural Sciences, Switzerland
[c] Eawag, Department of Environmental Chemistry, Switzerland

ARTICLE INFO

Handling Editor: Ingrid Kögel-Knabner

Keywords:
Soil legacy P
Eutrophication
Phytoextraction
Artificial drainage
Hydrological risk
Modelling

ABSTRACT

The role of diffuse phosphorus (P) losses from agricultural land gained importance since technical and constructional measures in urban areas and industry have reached their potential in temperate regions. Phytoremediation strategies are a new suggestion to effectively reduce soil legacy P but until now, there is a lack of estimates on the time such strategies should take. With a rainfall-runoff model, spatial information on the hydrological risk of the catchment of Lake Baldegg (Switzerland) was generated and combined with a soil test P (CO_2-saturated water extraction) map. Based on these results, two potential soil target test P values (2.0 mg P (kg soil)$^{-1}$ target-P 1 and 1.6 mg P (kg soil)$^{-1}$ target-P 2) were set. A simple nonlinear mixed effects model was used to compare different balance and cropping scenarios to decrease soil test P. The relationship between P-balance (input - output) and soil test P was found to be exponential. The confidence interval for the predicted time necessary to reach target-P 2 after a cessation of P-fertilization on intensively managed grassland was 2 to 9 years depending on the initial soil test P. If fertilization is completely ceased, the predicted P-decline times were longer (8 to 32 years). The decline-time for the balance that is recommended for farmers in the catchment of a P-fertilization that covers 80% of the demand was predicted to be 11 to 47 years. The study emphasizes that P phytoextraction can be an effective and time and resource efficient mitigation strategy for catchments with high legacy P.

1. Introduction

Surface water eutrophication is a severe environmental threat that has been a policy subject in temperate regions, especially in Central Europe and North America, for the last decades (Haygarth and Jarvis, 1999; Kleinman, 2017). Diffuse P losses from agricultural soils are a major P-source for eutrophication (Withers et al., 2017).

Transfer from diffuse sources to catchment streams occurs if an area is connected to the water body and if hydrological pathways are active, allowing soil or water – as well as the dissolved and particulate constituents they contain – to be transported (Pionke et al., 2000; White et al., 2009; Sharpley et al., 2011; Alder et al., 2015; Thomas et al., 2016b). Linking soil and catchment hydrology is often constrained by the differences in scales and research questions. Identifying concepts and indicators that best predict transfer of soil P has therefore been set as a main goal in recent studies (Bai et al., 2013; Schoumans and Chardon, 2015; Cassidy et al., 2017; Couto et al., 2018).

At the catchment scale, different hydrological modelling attempts are used to predict P transfer pathways (Pionke et al., 2000; Hahn et al., 2014; Shore et al., 2014; Wang et al., 2020). Identifying the most important processes and pathways of P-losses is the first step for hydrologists to tackle the issue of P-transfer in a specific catchment. Most studies thereby focus on direct connectivity and surface runoff while neglecting P transfer through subsurface artificial drainage systems or subsurface flow (Stamm et al., 1998; Thomas et al., 2016a; Gramlich et al., 2018). While the importance of such pathways varies between catchments, subsurface or drainage-flow can be important pathways for P in catchment areas with light topsoil and low erosion risk (Thomas et al., 2016b; Zhang et al., 2017; Koch et al., 2018). Artificial drainage systems are, furthermore, often excluded from hydrological models due

* Corresponding author at: Agroscope, Water Protection and Substance Flows, Reckenholzstrasse 191, CH-8046 Zurich, Switzerland.
E-mail address: Volker.prasuhn@agroscope.admin.ch (V. Prasuhn).

https://doi.org/10.1016/j.geoderma.2021.115257
Received 16 July 2020; Received in revised form 26 April 2021; Accepted 1 May 2021
Available online 11 June 2021

to a lack of reliable and up-to-date information on their geographical distribution.

At the field scale, soil scientists investigate the relationship between P losses and P forms, including processes at the molecular scale (Schoumans and Groenendijk, 2000; Kleinman et al., 2006; Godlinski et al., 2008; Frossard et al., 2014). For this, sound soil P-testing methods and classification systems are a precondition (Kleinman et al., 2011; McDowell et al., 2017; Withers et al., 2017). Most of the soil test P methods and their classification systems have been developed to research relationships between soil P availability and crop production. Until to date, no agreement exists on a classification system that accounts for productivity as well as for the risk of losses to surface waters (de Alcantara et al., 2008; Wuenscher et al., 2015; Macintosh et al., 2019). Target P values are therefore often set according to political and economic feasibility (Schoumans and Chardon, 2003). McDowell et al. (2020), for example, defined an agronomic target and a significantly lower environmental target.

Concentrations of readily exchangeable P fractions used for soil P testing are generally said to be good indicators for actual losses compared to more stable forms of soil P (Withers et al., 2017). However, such indicators neglect the potential of less available soil P fractions to act as a future source (Daly et al., 2015; Thomas et al., 2016b).

Focusing on the mitigation of current agricultural P-inputs as a source for eutrophication, research has only recently started to recognize the importance of long-term P accumulation in soils due to excessive fertilization (Chen et al., 2018, Jarvie et al., 2019). Contrary to the perception of P as a poorly mobile element, the P-sorption capacity of a soil can be drastically reduced with long-term P-accumulation. This allows P to reach subsoil layers and get lost via subsurface flow (Koch et al., 2018). New attempts propose the accumulation of P in soils as a legacy pollution issue that should be addressed with remediation measures. Consequently, phytoremediation strategies - commonly used at sites contaminated with heavy metals (Koopmans et al., 2007) - are suggested as a strategy to drastically lower soil legacy P and its losses to surface waters (van der Salm et al., 2009). The simple concept of phytoextraction is to grow hyper-accumulator plants that take up large amounts of the pollutants. The pollutants will be removed from the plot with the harvest of the biomass.

Several studies have examined temporal changes of soil test P in field trials (van der Salm et al., 2009; Schulte et al., 2010; Dodd et al., 2012; Cassidy et al., 2017; Vadas et al., 2018). Strategies of P-mining were zero fertilization (van der Salm et al., 2009; Dodd et al., 2012) or zero P-fertilization with supply of other nutrients (Schulte et al., 2010; Cassidy et al., 2017) or reduced P-fertilization (Liu et al., 2019). Temporal decline was found to be related to original P status (Cassidy et al., 2017), soil type (Dodd et al., 2012; Vadas et al., 2018) and P-balance (Schulte et al., 2010). The studies suggested decline times between five and more than 50 years for a specific plot. However, only few studies put their findings in context with the hydrology of the total catchment (Cassidy et al., 2017; Vadas et al., 2018; Withers et al., 2019). To recommend phytoextraction strategies to policy makers and farmers, additional in-depth studies must focus on the parameters that limit such strategies and link their findings to hydrology.

In Switzerland, the eutrophication of Lake Baldegg has been subject of scientific studies since the 1980 s (Gaechter, 1987; Gaechter and Wehrli, 1998; Lazzarotto et al., 2006; Hahn et al., 2012; Frossard et al., 2014). Policy measures to reduce P transfer to the watershed are conducted since 30 years and P-loads and concentrations of available P in water and soil have been monitored in-depth from the beginning. This allows researchers to analyze long-term effects of policies and draw conclusions that are transferable to other catchments.

The aim of the study was i) to develop a model to quantify areas of high hydrological risk in the catchment of Lake Baldegg which also includes artificial drainage and ii) to evaluate time efficiency of P phytoextraction strategies for P decline.

To this end, a model based on the critical source area concept was improved by adding an artificial drainage probability map. Furthermore, a regression-based model with data from the catchment of Lake Baldegg on P status and nutrient removal by crops was generated to predict the time necessary to reach a soil test P level compatible with a low level of P losses. Different P balance and cropping scenarios were fed in the model to evaluate their efficiency for P phytoextraction.

2. Material / study area

The Lake Baldegg catchment (67.8 km^2) is located in Central Switzerland in the Canton of Lucerne (Fig. 1). In-lake measures to reduce the effects of eutrophication (oxygenation and artificial mixing) have been conducted since 1982 (Gaechter and Wehrli, 1998; Beutel and Horne, 1999). Several projects to reduce P inputs from agriculture have been realized since the late 1980 s. Consequently, the diffuse load of dissolved reactive P (DRP) to the lake decreased from 4864 kg y^{-1} (1986–1990) to 2782 kg y^{-1} (2012–2016) (uwe: state office for environment and energy of the Canton of Lucerne, 2018). The DRP concentration of the lake has decreased from 271 mg m^{-3} to 24 mg m^{-3} in the same time span. In spite of this decrease in P inputs and P concentration, no increase in O_2 contents of the lake water has been detected since 1980 (uwe: state office for environment and energy of the Canton of Lucerne, 2018).

The utilizable agricultural area (UAA) of the catchment is farmed at a high intensity based on livestock husbandry (Table 1). Grassland (including temporary ley) makes up for 67% of the UAA. The majority of the grasslands are managed intensively with up to six harvests for silage, hay or barn feeding of fresh grass and/or grazed with cattle. Such meadows are as well manured 5 to 6 times per year. The botanical composition of such meadows differs according to intensity but is generally a mixture of grasses (*Lolium multiflorum, Lolium perenne, Dactylis glomerata and others*) with clover (Trifolium pratense, Trifolium repens and others). Farmers of the area are compensated for measures to reduce P-losses from agriculture since 2000. Most farmers of the area follow the concepts of buffer strips along water bodies, appropriate soil cover and fertilization winter break. Apart from that, involved farmers have to provide calculations (method “Suisse-Bilanz”) to prove their balanced use of P on the farm level (Agridea, 2018). Average P balances at farm level are hence slightly negative, indicating that current P fertilization does not exceed plant P uptake by cultivated plants.

The soil texture (0 – 20 cm depth) is predominantly silty and the slopes are strong and often elongated. Only 13% of the UAA has a slope of less than 5% and an important share of land is probably artificially drained (Huerdler et al., 2015). The climate is temperate humid continental with a mean annual precipitation of 1140 mm (Hahn et al., 2012). Due to steep slopes, silty soils and high precipitation rates, the potential erosion risk of the area is high, but is well adapted to the area by growing predominantly grassland. Furthermore, also the connectivity to surface waters is very high (Alder et al., 2015). A detailed description of the hydrology of the catchment of Lake Baldegg can be found in Hahn et al. (2014).

For this study, the catchment was divided in subcatchments with a special focus on the five subcatchments “Ron” (R) “Staegbach” (StB), “Spittlisbach” (SpB), “Muelibach” (MB) and “Hoehibach” (HB) (Fig. 1), as they have been monitored by the state office for environment and energy of the Canton of Lucerne since 1985. Furthermore, the sub-subcatchment “Obere Ron” (OR) (indicated in green Fig. 1) of the R subcatchment was used for detailed numerical analyses, due to a nearly complete area-related dataset of soil test P.

3. Methods

Conceptual base of this study is a quantitative synthesis of hydrological catchment knowledge including hydrological risk with measured soil test P, resulting in a target soil test P level for the catchment (Fig. 2). This target level is then fed in a soil test P decline model to quantify the

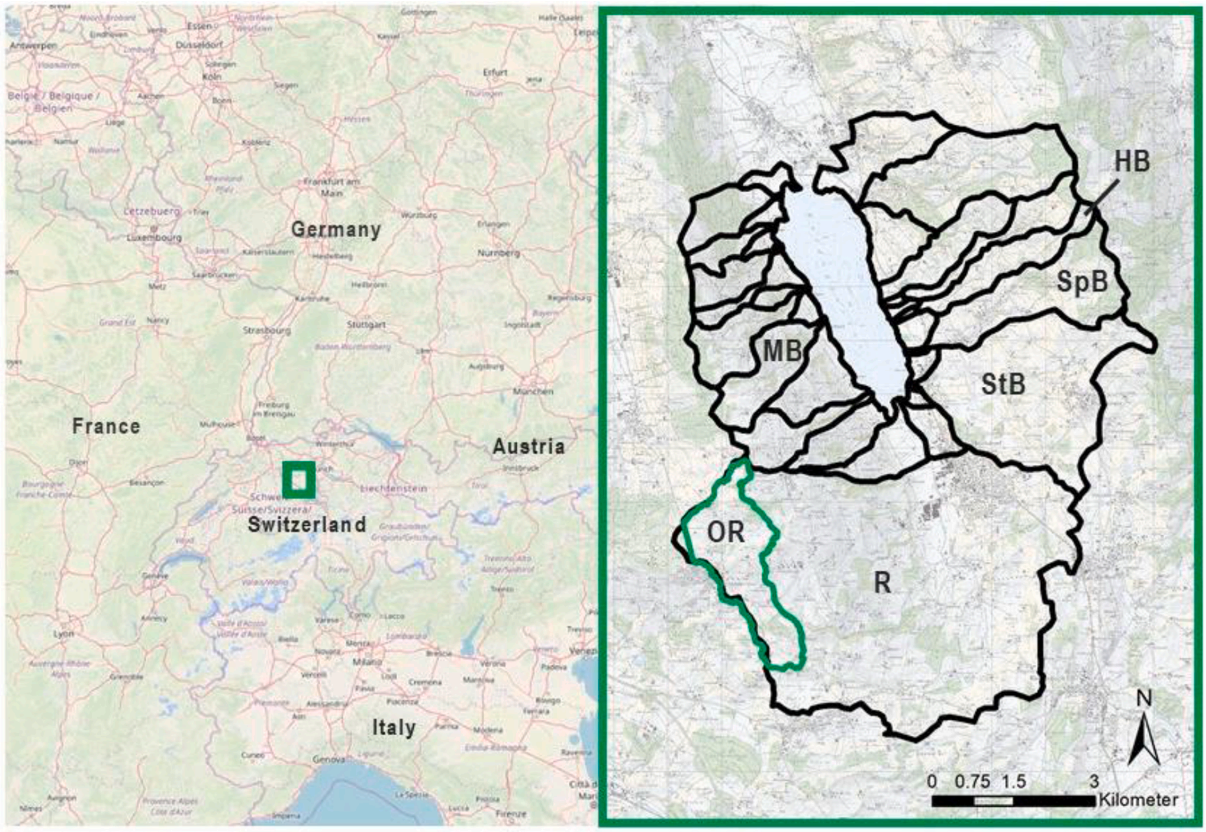

Fig. 1. Map of Switzerland with indication of the location of Lake Baldegg (OpenStreetMap contributors, 2019) and closeup of the catchment with its subcatchements "Ron" (R). "Obere Ron" (OR), "Staegbach" (StB), "Spittlisbach" (SpB), "Muelibach" (MB) and "Hoehibach" (HB).

Table 1
Agriculture in the catchment of Lake Baldegg.

Catchment area (excl. of lake) *[ha]*		6′780
Livestock number *[ ]*		12′023
Livestock density *[ha^{-1} UAA]*		2.4
	[ha]	*[%]*
Arable land	1′584	31.3
Grassland (manured)	3′023	59.8
Grassland (not manured)	370	7.3
Others	81	1.6
Total cultivated area	**5′058**	

UAA: Utilizable agricultural area

time it takes to reach the targeted DRP loads.

3.1. DRP load and concentration in streams of five subcatchments

Measurements of instream DRP concentrations and discharge at the outlet of five important streams to Lake Baldegg were provided by the state office for environment and energy of the Canton of Lucerne. Samples were taken with passive samplers from the five monitored subcatchments R, StB, SpB, MB and HB (Fig. 1) in a 22 day-rhythm and during flood events since 1985. The samples are 24 h composite samples.

From the relationship of discharge and instream DRP-concentration at the outlet of the streams, the total DRP load of each stream was calculated (Moosmann et al., 2005). For this study, average yearly DRP loads of the subcatchments R, StB, SpB, MB and HB in the period between 2001 and 2015 were calculated.

3.2. The rainfall-runoff-phosphorus (RRP) model

A parsimonious model was used to compute the hydrological risk area of the catchment of Lake Baldegg. The model simulates the transfer of DRP from intensively managed grassland soils to streams of small agricultural catchments and was developed by Lazzarotto et al. (2006) and extended by Hahn et al. (2013). Two sub-modules are combined: a semi-distributed rainfall-runoff (RR) module and a phosphorus (P) module. The model can be downloaded from GitHub (https://github.com/Rainfall-Runoff-Phosphorus-Model/RRP). A detailed description of the original RRP model can be found in Lazzarotto et al. (2006) and Hahn et al. (2013).

The rainfall-runoff submodule is based on the assumption that the hydrological behavior of soils can be grouped by a topographic wetness index (TWI) (Beven and Kirkby 1979) and the soil type. There are four different hydrological response units (HRU) defined which are each assumed to be uniform concerning runoff generation behavior.

The main modification of the original RRP model (Lazzarotto et al., 2006; Hahn et al., 2013) was the inclusion of an artificial drainage probability (P(D)). Using a machine-learning (ML) approach (Gradient Boosting Machine algorithm, H2O.ai Team, 2017) implanted in R (R Core Team, 2016) a map for the P(D) of the catchment area was generated. Existing empirical artificial drainage maps of the three Swiss cantons Zurich, Basel-Landschaft and St Gallen were put in context with corresponding maps regarding soil water properties (low soil water drainage potential (WH0), high soil water drainage potential (WH1)), wetlands (FG), geological predisposition for water logging from Szerencsits et al. (2018) (GEO), hill slope (SLOPE) and TWI to train the ML model. The P(D) was computed from the relations of the geographical information variables.

$$P(D) = f(WH0, WH1, FG, GEO, SLOPE, TWI) \quad (1)$$

To evaluate the ML model and avoid overfitting, the complete dataset was split in three independent subsets. The first subset was used to train the ML model. With the second subset, the hyperparameters

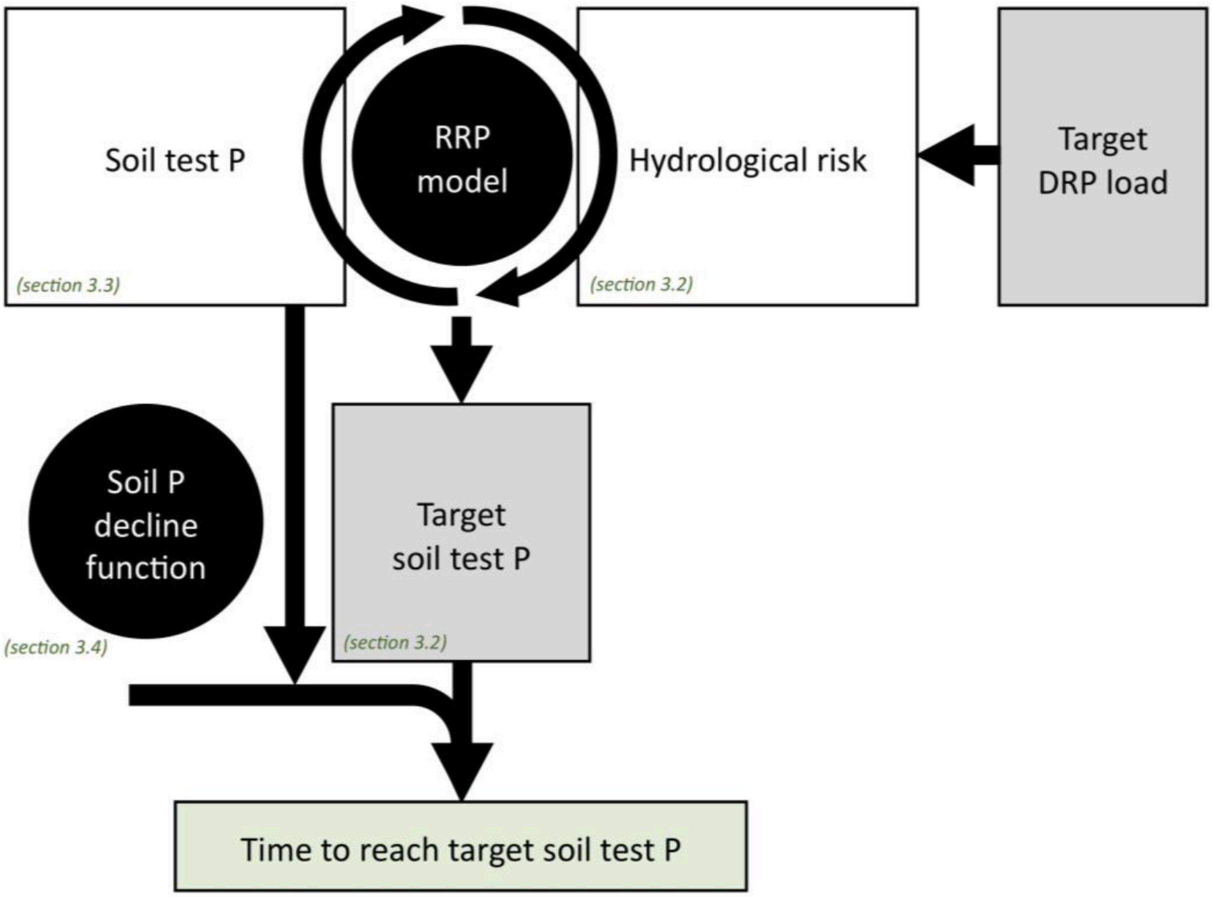

Fig. 2. Conceptual model for the development of the soil test phosphorus (P) decline function based on the rainfall-runoff-phosphorus (RRP) model and the target level of dissolved reactive P (DRP).

defining the learning process of the ML approach were evaluated and optimized. The last subset was then used to test the final optimized model with an independent dataset. The fit of the model was evaluated using the area under the receiver operating characteristic curve (0.5 – 1) after Bradley (1997), which reached a reasonable value of 0.77 (average over the three model Cantons). It was not possible to validate the generated P(D)-map quantitatively in the catchment of Lake Baldegg, as no data existed on artificially drained UAA in the area. Nevertheless, the model computed an area proportion of artificially drained land of 10.7% of the UAA. This is in line with findings by Béguin and Smola (2010) and Huerdler et al. (2015), who found artificially drained land counting for 6% and 8%, respectively, of the total UAA of the Canton of Lucerne.

The P(D) was included in the rainfall-runoff submodule by modification of the TWI. Whenever P(D) exceeded 50% for a pixel, a modified TWI_mod was used:

$$TWI\_mod = TWI^{*}(P(D)^{*}2)\ \text{for}\ P(D) > 0.5 \quad (2)$$

The 14 parameters of the rainfall-runoff submodule were calibrated using a uniform Monte-Carlo approach evaluating over 2′000′000 parameter combinations using the Nash-Sutcliffe efficiency (NSE) (Nash and Sutcliffe, 1970). The calibration was done in 2013 and 2014 in three subcatchments during the season of highest losses, namely the period March - October. Out of all these parameter combinations 1304 parameter sets reached an NSE higher than 0.55 and were accepted. These model realizations were subsequently evaluated with an independent dataset for the period March - October 2015 in five subcatchments and reached higher NSE than during the calibration period.

With the calibrated model, the hydrological risk areas were identified for a medium rainfall event with 35 mm precipitation per day. This is a typical amount to generate P losses. Such precipitation events occur averagely twice a year. Depending on the number of parameter sets indicating hydrological activity, five risk classes were defined (Table 2).

The extended model performed similar than the original model by Hahn et al. (2013). There was no direct comparison possible, as different catchments and different time periods were modeled.

To predict area-related P losses, the hydrological risk area (areas producing fast flow) of a medium rainfall event with 35 mm precipitation per day, was combined with a soil test P map of the subcatchment "Obere Ron" (OR) of the Lake Baldegg catchment. This was motivated by the fact that soil test P was known for 82% of the UAA of this subcatchment. A relationship between soil water soluble P (WSP) and DRP was determined by Hahn et al. (2012) from artificial rainfall experiments. As soil test P of the catchment area of Lake Baldegg was given in CO_2 saturated water extractable P (CO_2-P; FAL; RAC; FAW, 1996) the map was converted to WSP through the approximated relationship

$$1\ \text{mg P}(\text{kg soil})^{-1}\ \text{WSP} = 2\ \text{mg P}(\text{kg soil})^{-1}\ CO_2 - P \quad (3)$$

Table 2
Risk classes for hydrological activity according to the percentage of parameter sets (n = 1304) indicating hydrological risk from the rainfall runoff phosphorus model in the catchment of Lake Baldegg.

Parameter sets *[%]*	Hydrological risk
0.0–0.2	very low
0.2–0.4	low
0.4–0.6	medium
0.6–0.8	high
0.8–1.0	very high

with $R^2 = 0.8$ observed by Stuenzi (2006) by the analysis of 86 soil samples covering a broad range of chemical properties and textures. The P submodule was evaluated for two subcatchments.

With the RRP model, it was possible to compute different scenarios of DRP loads for OR depending on different average soil test P values of the catchment. The actual soil test P in OR with the according DRP-load was set as a reference scenario. A 50% decrease of the reference load was set as a target with its according target soil test P (target-P 1). Additionally, a target-P 2 for the entire catchment was set to make general predictions.

3.3. Soil test P

A soil test P map was generated based on a dataset of 3723 soil samples from the UAA of the catchment of Lake Baldegg. The soil samples were taken farmwise between 2009 and 2014 by an inspection body (Qualinova AG, Gunzwil) and analyzed by certified laboratories within the framework of the water protection policies in the catchment. For each sample, the area of the sampled plot and its geographical location were recorded. The samples were analyzed by the CO_2 saturated water technique, which is commonly used in Switzerland to measure soil test P (Flisch et al., 2017). For the extraction, demineralized H_2O is saturated with a p_{CO2} of 0.5 – 6.0 bar, reaching a solution pH of 3.5–4.0 (Stuenzi, 2006). Air-dried soil samples are extracted in a soil to solution ratio of 1:2.5 for one hour. Orthophosphate concentration in the extract is measured colorimetrically with the malachite green method (Ohno and Zibilske, 1991) and given as mg P (kg soil)$^{-1}$.

3.4. Soil P decline model

To evaluate the temporal scale of phytoextraction strategies, a model based on a regression with existing data from the Lake Baldegg catchment was calculated. The principle of the model was to simulate soil test P decline in relation to yearly P-balances. For this, data from a greenhouse pot-study from Frossard et al. (2014) was used. In the study, Italian ryegrass (*Lolium multiflorum*) as a typical grass species from the area was grown in soil sampled from five intensively managed grassland plots (0.1 m sampling depth) from three different sites within the catchment of Lake Baldegg. During the experiment, the effect of three P fertilization levels (0, 20 and 40 mg P (kg soil)$^{-1}$ as KH_2PO_4) after three cultivation cycles of each eight weeks were compared. The data allowed to do a regression analysis of P isotopically exchangeable within 1 min (E_{1min}) [mg P (kg soil)$^{-1}$] relative to the cumulative P balances [mg P (kg soil)$^{-1}$] of each plot.

To apply the simulations to the catchment of Lake Baldegg, the data (E_{1min}) was converted to CO_2-P [mg P (kg soil)$^{-1}$] by the equation ± the standard error of the two parameters

$$CO_2 - P = 0.08 \pm 0.01 * E_{1\ min}^{1.3\pm0.06} \tag{4}$$

The equation results from a power function regression ($R^2 = 0.93$) based on data obtained by Frossard et al. (2011). In the study, 104 soil samples from the catchment of Lake Baldegg were analyzed for both, E1min and CO_2-P. The regression was computed with the nonlinear least-squares (*nls*) function from the *stats* package in R (R Core Team, 2016; Bates and Chambers, 1992). The CO_2-P values used for the regression reached from 0.1 to 10.8 mg P (kg soil)$^{-1}$ (Frossard et al., 2011), covering the range of the values from Frossard et al. (2014) of 0.1

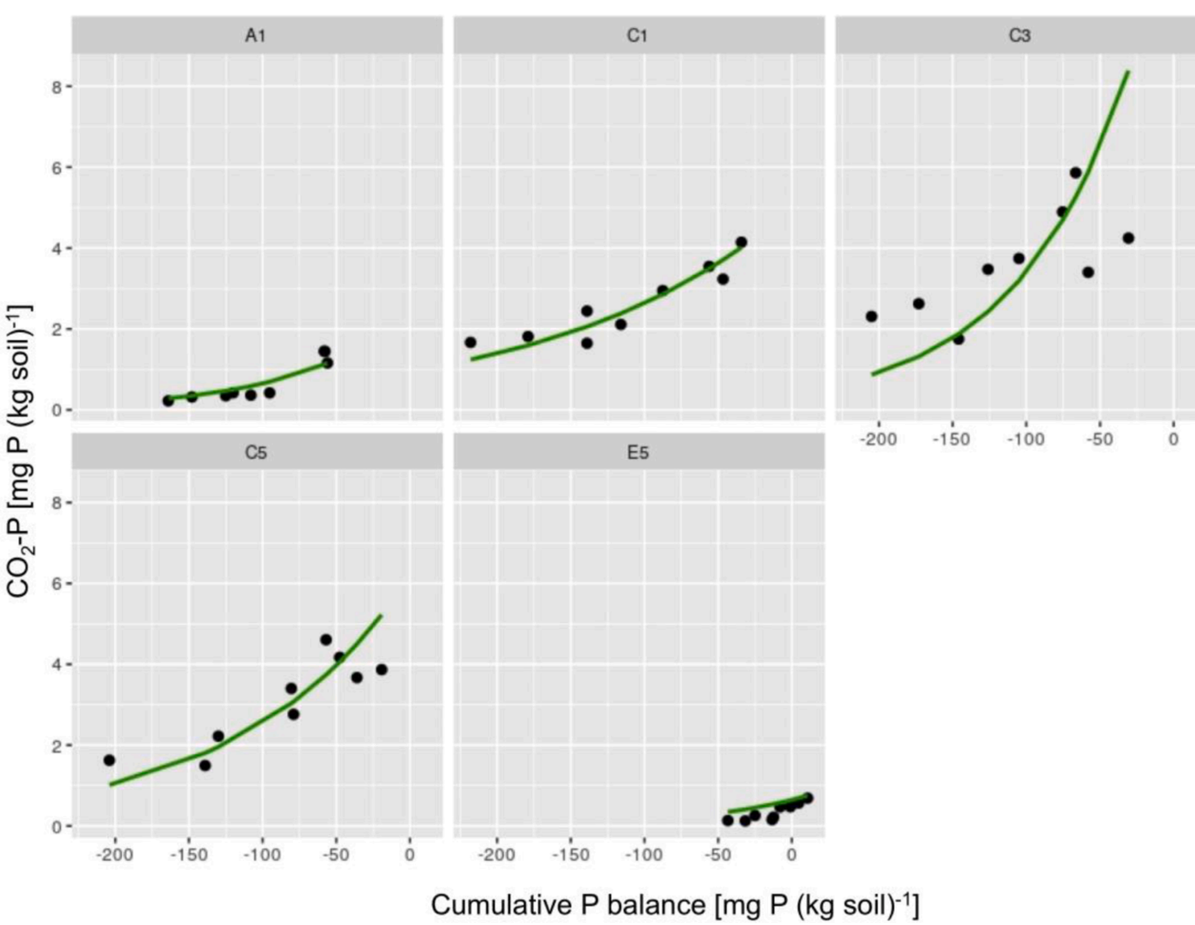

Fig. 3. Regression of soil test P, measured as CO_2-enriched water extractable P, as a function of the cumulative P-balance from Frossard et al. (2014) from the five plots A1, C1, C3, C5, E5 from the three different locations A, C, E used for the nonlinear mixed effects model of soil test P decline.

to 5.9 mg P (kg soil)$^{-1}$.

For the soil P decline model, a nonlinear mixed effects model was computed with the *nlme* function from the *nlme* package (Lindstrom and Bates, 1990). The best fitting regression function ($R^2 = 0.88$, Fig. 3) was identified to be an exponential decay curve of the following form:

$$CO_2 - P = CO_2 - P(0) * e^{(b^* \text{cumulative P} - \text{balance})} \quad (5)$$

with the two independent variables CO_2-P(0), indicating the CO_2-P at the start of the experiment [mg P (kg soil)$^{-1}$] and cumulative input–output balance [mg P (kg soil)$^{-1}$] indicating the sum of the P balances of each cycle since the start of the pot experiment. The parameter b indicates the decline rate and is estimated by the model. No fixed factor was applied to b but the grouping of the data structure was included by the two random factors "site" and "plot". The factor "plot" is nested within "site" because three of the plots were sampled within a close distance, indicated as "site". The model had to be controlled for heteroscedasticity of the within-group error. Therefore, the unequal variances were weighed with a power function (varPower()) and the stratification variables CO_2-P(0) and "site" (form = ∼ CO_2P0 | site) (Chambers and Hastie, 1992). To account for the uncertainties of the regression equation due to the unexplained variance, a 95% confidence interval for the parameter b was computed.

For the simulations with the fitted equation, a time component was included by setting a discrete time-dependent cumulative P balance. By this step, the model was converted to field conditions under the assumption that cumulative P balances after each cycle of the pot experiment correspond to cumulative P balances measured in the field. This resulted in the equation

$$\text{Cumulative P balance}(t) = \text{P balance} * t \quad (6)$$

With P balance indicating the yearly P balance [mg P (kg soil)$^{-1}$ yr^{-1}] and t the time in years. The time dependent model-equation therefore was

$$CO_2 - P(t) = CO_2 - P(0)*e^{(b^* \text{ cumulative P balance}(t))} \quad (7)$$

For the P balance, three scenarios were developed based on theoretical balances of three cropping strategies. Two scenarios represent intensive management of grassland with 5–6 cuts per year but different P input (P int), one scenario extensive use of grassland with one cut per year (P ext). The scenarios differed either in P-export (plant uptake) or in P-import (P fertilization). The two intensive scenarios (0%P int and 80%P int) yielded a maximum P-export of 46.8 kg P ha^{-1} (Table 3). Cropping assumption was an intensive management with 5–6 cuts per year and no limitation by nutrients (e.g. nitrogen (N) and potassium (K)), pests and diseases or climatic conditions. In the 0%P int scenario it was assumed, that P fertilization was stopped, while the 80%P int implied an 80% cover of P-demand. The third scenario (0%P ext) aimed to simulate a low intensity cropping strategy with no fertilization and reduced cutting frequency (1 cut per year) and limitation by N, K and other nutrients, resulting in lower P-export of 13.7 kg P ha^{-1} (Table 3). Reference numbers on yield and P uptake from the Swiss official fertilizer recommendations (Huguenin-Elie et al., 2017) were used for the scenarios and an average soil bulk-density of 1.2 t m^{-3} was assumed. The vertical system border was assumed to be 0.2 m. This was set due to a drastic decrease of available P and rooting in deeper soil layers (van der Salm et al., 2009; Fort et al., 2016). Further assumptions of the scenarios and the model application are that P is evenly distributed and that plants take up P from the entire defined soil volume. Furthermore, the P balance obtained with mineral KH_2PO_4 used in the pot experiment is assumed to be transferable to the field level.

Table 3
Yearly balances of the P balance scenarios for the P decline model: A zero P fertilization without limitation of productivity of an intensive grassland with 5–6 cuts per year (0%P int), the total cessation of fertilization of a grassland with one cut per year (0%P ext) and a 80% cover of uptake P fertilization without limitation of productivity of an intensive grassland with 5–6 cuts per year (80%P int).

	Input *[kg P ha^{-1}]*	Uptake	Balance	Balance *[mg P (kg soil)$^{-1}$]*
0%P int	0	46.8	−46.8	−39.0
0%P ext	0	13.7	−13.7	−11.4
80%P int	37.4	46.8	−9.4	−7.8

Three starting soil test P levels of the decline function were set. Besides the measured median soil test P of 1207 samples from permanent grasslands in the catchment of Lake Baldegg of 3.1 a high (upper hinge) of 4.9 was set (supplementary material, Figure SII). Due to the assumption that highly enriched soils account over-proportionally to P-losses, a very high third starting value of 6.3 mg P (kg soil)$^{-1}$ was set according to the maximum of calibration range of the model.

3.5. Correlation analysis

The DRP load of each of the five subcatchments (3.1) was used for a Spearman ranked correlation analyses with the *stats* package in R. (Becker et al., 1988). In each subcatchment, the percentage share of area of each of the five hydrological risk classes (very low, low, medium, high, very high) was paired with the respective DRP load of the catchment. Similarly, percentage area of each of the four soil test P classes (<0.7, 0.7–1.6, 1.7–2.5, >2.5 mg P (kg soil)$^{-1}$) and the percentage share of artificially drained area was computed as a function of the respective DRP load. Significance of the correlation was tested with the Wilcoxon signed-rank test due to small sample size (n = 5).

4. Results and discussion

4.1. Hydrological risk area (RRP model)

According to the RRP-model, an area covering 69% of the UAA of the catchment of Lake Baldegg shows medium to very high hydrological risk (supplementary material, Figure SI) for a medium rainfall event of 35 mm day^{-1}. The map indicates that hydrological risk is generally high in all subcatchments and evenly distributed over the entire catchment (Fig. 4). The subcatchments differ mainly in their percentage area of very high and very low hydrological risk. Correlation analysis showed that the DRP load is highest in the subcatchment with the highest share of hydrological very high risk areas (ρ = 0.8) and lowest in the subcatchment with the highest share of area with very low hydrological risk (ρ = -0.8) but the correlation was not significant (Fig. 5A). Still, this relationship is an indication that the RRP-model performs well in predicting P losses in the catchment of Lake Baldegg. Particularly, the importance of the pathway through tile drains is indicated by the correlation of percentage share of artificial drainage and the total DRP-load of the subcatchments (ρ = 0.7), although the relationship was not significant (supplementary material, Figure SIII). The lack of a significant result can mainly be explained by a generally low percentage area of artificially drained land in all subcatchments.

The P(D)-map in the model extends earlier versions of the model as also areas with a low gradient can display high hydrological risk (Hahn et al., 2013). Artificial drainage is often neglected in hydrological models (Thomas et al., 2016a; Gramlich et al., 2018) and DRP-losses are often mainly attributed to surface runoff (Pionke et al., 2000). An inclusion of artificial drainage in hydrological models is of particular importance when the critical source area concept is applied (McDowell et al., 2017). Agronomical measures targeting diffuse P losses often focus on surface connectivity. Examples of such measures are buffer strips along streams or appropriate soil cover. Such measures have limited effect on P-transfer from artificially drained land as the contributing plots often lack in surface connection.

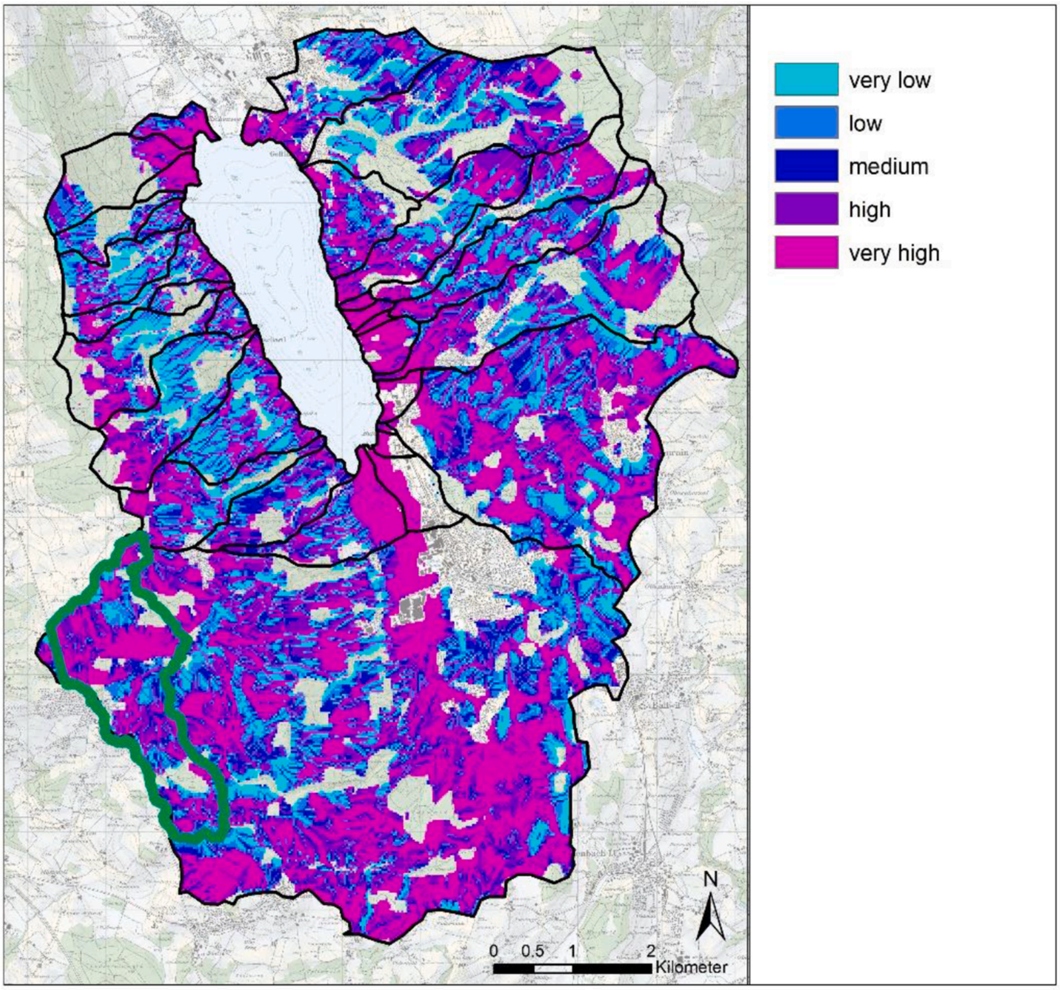

Fig. 4. Map of the hydrological risk area during a medium rainfall event of 35 mm day^{-1} in the catchment of Lake Baldegg with five risk classes based on 1304 parameter sets of the rainfall-runoff-phosphorus model.

4.2. Soil test P

Soil test P was known for an area of 63% of the entire catchment of Lake Baldegg. The area-standardized average soil test P (CO_2 saturated water) of the covered area was 2.5 mg P (kg soil)$^{-1}$ (Fig. 6). In the subsubcatchment OR a slightly higher average soil test P of 3.4 mg P (kg soil)$^{-1}$ was measured. Separate analyses of samples from permanent grassland (1207 samples) and arable land (1671 samples) resulted in a median soil test P of 3.1 and 1.75 mg P (kg soil)$^{-1}$ for grassland and arable land (supplementary material, Figure SII), respectively. In all catchments, soils with high (1.7–2.5 mg P (kg soil)$^{-1}$) and very high (greater than2.5 mg P (kg soil)$^{-1}$) soil test P accounted for more than 60% of the UAA. Differences were only observed between the percentage area of the highest two classes. Because soil test P is generally high in all catchments, catchments with larger shares of soils in the "high" class have smaller shares of "very high" area. The spearman correlation even computed a positive linear relationship of the "very high" soil test P area with the DRP load that was significant at an alpha level of 0.1 while a negative correlation was found for the "high" ($\rho = -0.9$) soil test P area (Fig. 5B). This indicates that the threshold value for excessive P-mobilization may be in between the two classes, which were set for agronomic purposes. This assumption is also supported by the fact that we could not detect any relationship between average soil test P and the DRP load of the subcatchments (data not shown).

4.3. Target soil test P (RRP Model)

With the current average soil test P of 3.4 mg P (kg soil)$^{-1}$ of the UAA in the OR an average yearly P load of 65.1 kg (km UAA)$^{-2}$ was computed by the RRP model (Fig. 7). This value is in line with the value of 76 kg (km UAA)$^{-2}$, which was computed from the concentration-discharge relationship (3.1) because the RRP model only accounts for soil-borne P inputs. To reach the targeted 50% of the current load (32.8 kg (km UAA)$^{-2}$), soil test P in the OR subcatchment must decrease to 2.0 mg P (kg soil)$^{-1}$ (target-P 1). To make general predictions, a target soil test P was set also for the entire catchment of Lake Baldegg. This target-P 2 of 1.6 mg P (kg soil)$^{-1}$ was set lower due to lower average soil test P in the total catchment (Fig. 6). Because scenarios could only be computed for OR, some uncertainty remained about the applicability of the scenario to the entire catchment of Lake Baldegg. The uncertainty derives from a different catchment hydrology in OR and because average soil test P of the entire catchment was calculated from only 63% of the area. This was the main reason why two target P values were exhibited. Target P 2 is in the range of agronomic optimum soil test P after the Swiss classification system, whereas target P 1 is higher as recommended (Flisch et al., 2017).

Most authors found detrimental losses at soil test P levels within agronomic optimum ranges of their classification system (McDowell et al., 2017; Withers et al., 2017) while Macintosh et al. (2019) stated that the relation between soil test P and P losses to surface waters is strongly dependent on soil attributes such as texture and Al- and Fe-

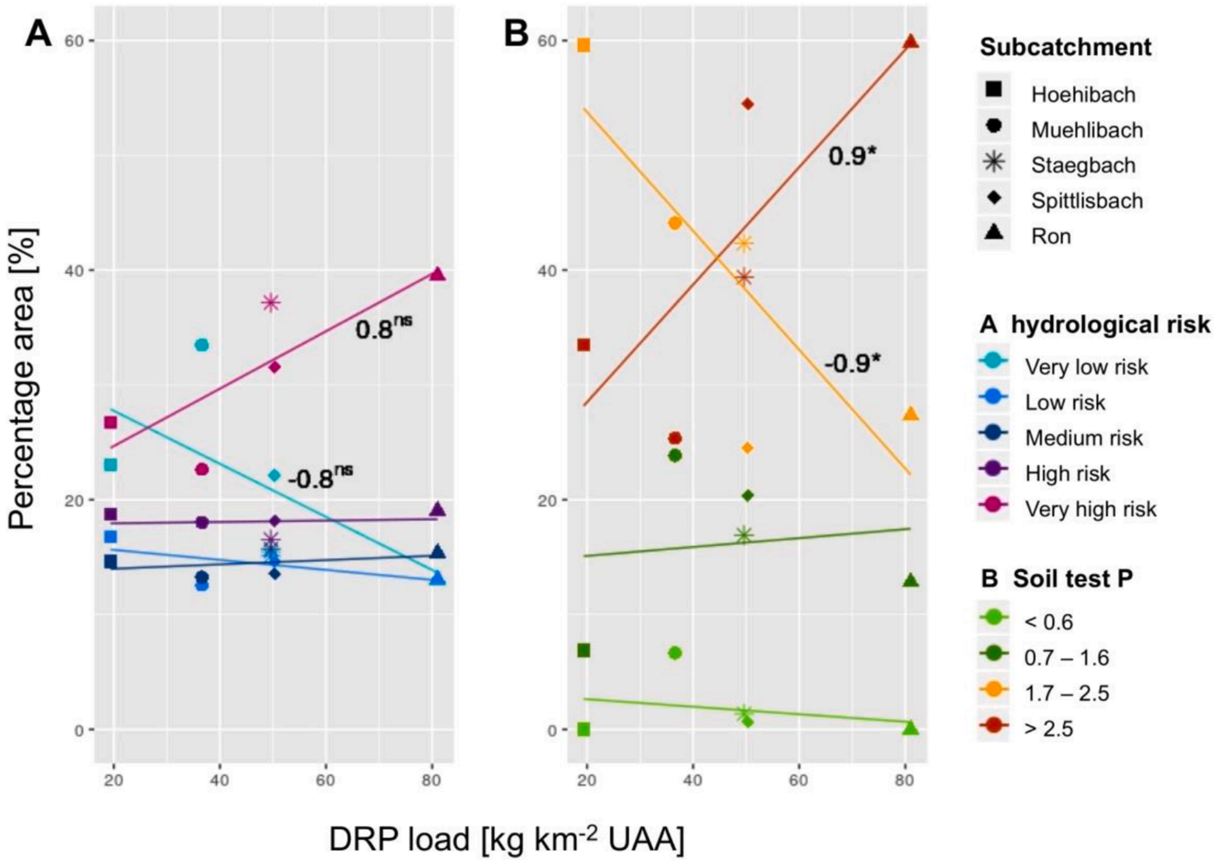

Fig. 5. Spearman rank correlation of dissolved reactive phosphorus (DRP) load of the utilizable agricultural area (UAA) in five subcatchments of Lake Baldegg (symbols) and the respective relative area of five hydrological risk classes (A) and soil test P classes, measured as CO_2-enriched water extractable P in mg P (kg soil)$^{-1}$ (B) with * indicating significance of the Wilcoxon signed-rank test at $\alpha = 0.1$ and ns indicating non-significance.

oxide concentrations. The release of reactive P from mineral soils is dependent on sorption processes. Soil P-sorption dynamics are stated to be nonlinear processes and soils release most of the reactive P once saturation occurred (Daly et al., 2015). In the RRP model, the relationship between soil test P and DRP losses is stated to be linear, while correlations rather suggest a threshold value because areas with a soil test P of more than 2.5 mg P (kg soil)$^{-1}$ seem to disproportionally contribute to the DRP load (Fig. 5B). However, the relationship between soil test P and soil-borne P-losses is still controversially discussed (Daly and Casey, 2005) as the relationship between soil test P and instream DRP is predominantly governed by the dominant hydrological transfer pathways (Chen et al., 2018). It is therefore debatable, to which extent exact threshold levels of soil test P are meaningful. However, reaching either of the two suggested target-P values demands drastic decline of test P of most farmers' soils.

4.4. Soil P decline

The estimated decline rate b of the model was 0.011 with a 95% confidence interval (CI) between 0.007 and 0.014. Estimated decline times of the 0%P int scenario were 3.1 (CI: 2.3–4.6), 5.2 (CI: 3.9 – 7.7) and 6.3 (CI: 4.7 – 9.5) years for soils with a median, an upper hinge and an extreme CO_2-P(0), respectively, until target-P 2 is reached. Correspondingly, the 0%P ext scenario reached decline time estimates of 10.4 (CI: 7.8 – 15.6), 17.6 (CI: 13.2 – 26.4) and 21.7 (CI: 16.2 – 32.4) years. If P fertilization covers 80% of demand (80% P int), respective decline time estimates are 15.3 (11.4 – 22.8), 25.8 (CI: 19.4 – 38.6), 31.7 (23.7 – 47.3) years (Fig. 8). If the higher target-P 1 is used, respective estimates are shorter.

Soil test P does not fall below the limiting level for crop production stated by Gallet et al. (2003) when either of the two target P values is used. Therefore, no yield decrease caused by P limitation is expected. However, Gallet et al. (2003) also found P uptake in grasslands to decrease below a soil P concentration of 18 mg P (kg soil)$^{-1}$ (E1min) which corresponds to a CO_2-P concentration of 3.4 mg P (kg soil)$^{-1}$. Consequently, below this concentration, P export is not only a measure of yield but also of the P-concentration in plants and varies therefore in relationship to soil test P.

Other authors also found soil test P in the field to decline exponentially (Schulte et al., 2010; Dodd et al., 2012; Johnston et al., 2016) or logarithmic (Cassidy et al., 2017) with time. The starting value of soil test P is often set as intercept (Schulte et al., 2010; Dodd et al., 2012) and the decline rate is stated to be a function of soil properties and P management (Schulte et al., 2010; Johnston et al., 2016). Schulte et al. (2010) already calculated P decline times for different P-balance scenarios and Heming (2007) found a linear relationship between P balance and resin-extractable P but did not apply this relationship to temporal changes. Until now, decline rates were mostly calculated for specific sites or soil types (Herlihy et al., 2004; Schulte et al., 2010). We created a more general model, which resulted in rather large confidence intervals but allowed to use the function for more general predictions.

Quantitative comparisons with other studies are very limited as different soil test P methods, target P values and starting soil test P values were used. Frossard et al. (2011) obtained field trials in farmers' fields in the catchment of Lake Baldegg. They found no change in CO_2-P after a cessation of fertilization of four to eight years. The lack of an effect could be explained by the CO_2-P method as reactive P forms can undergo fast fluctuations due to temperature, humidity and other physical processes related to weather and season. This high variation makes it hard to monitor and predict soil P decline (Zicker et al., 2018).

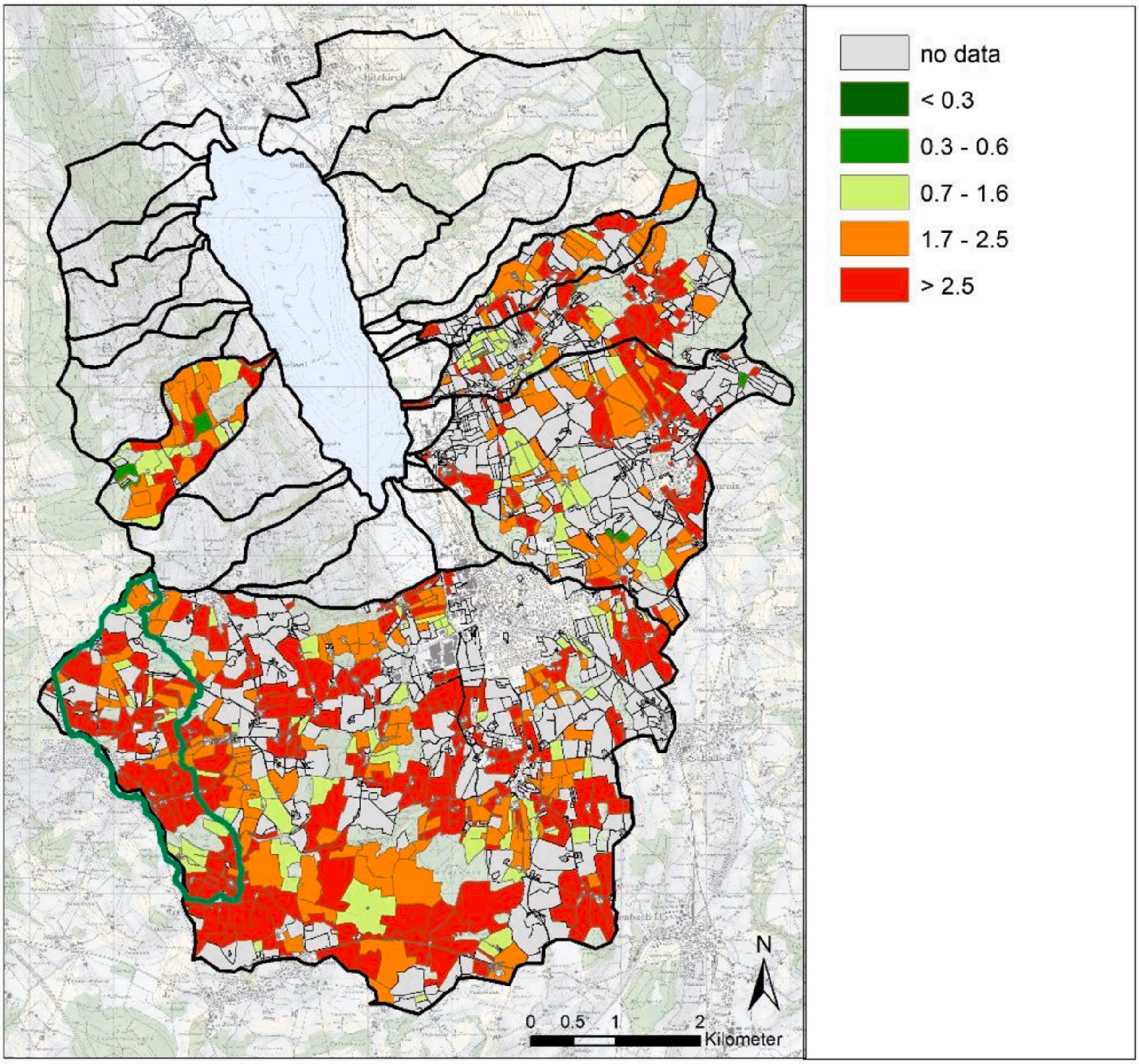

Fig. 6. Map of the soil test P, measured as CO_2-enriched water extractable P, in mg P (kg soil)$^{-1}$ in the catchment of Lake Baldegg colored according to five soil P classes.

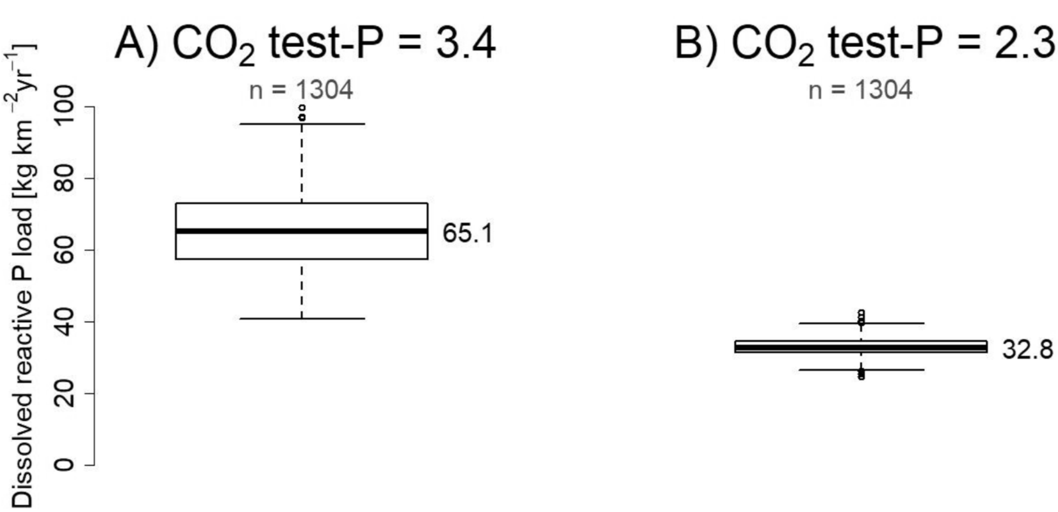

Fig. 7. Current (A) and targeted (B) load of dissolved reactive P from agricultural land in the catchment of Obere Ron derived from the 1304 parameter sets of the rainfall-runoff-phosphorus model with respective soil test P, measured as CO_2-enriched water extractable P.

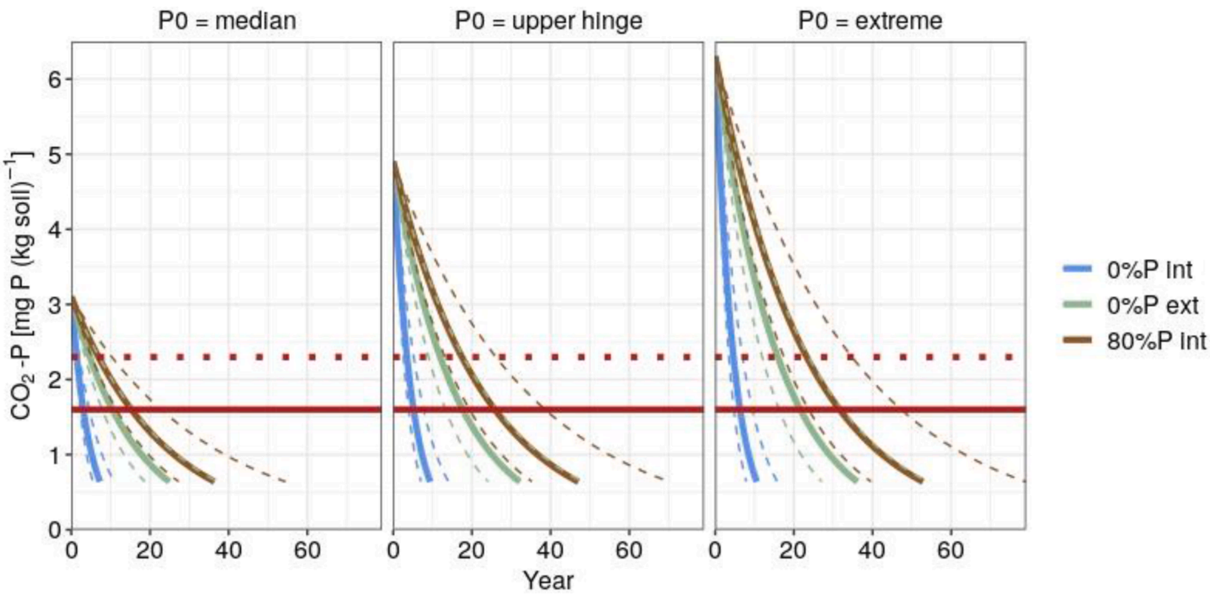

Fig. 8. P-decline scenarios with 95% confidence intervals (dashed lines) from a starting value of 3.1 (P0 = median), 4.9 (P0 = upper hinge) and 6.3 (P0 = extreme) mg P (kg soil)$^{-1}$, measured as CO_2-enriched water extractable P. Intensive grassland without P fertilization (0%P int), extensive grassland without fertilization (0% P ext), intensive grassland with P fertilization that covers 80% of P uptake (80%P int). Target values are indicated as red line (solid: target-P 2 (1.6 mg P (kg soil)$^{-1}$) and dotted: target-P 1 (2.3 mg P (kg soil)$^{-1}$)).

However, van der Salm et al. (2009) measured significant changes in P availability (water extractable P and ammonium lactate extractable P) after five years of P-phytoextraction. Schulte et al. (2010) even reached their target P in only seven years with a comparable strategy. In both studies, the plots received N-fertilization, while studies presented by Dodd et al. (2012) that did not receive N-fertilization, calculated much longer decline times of 23 – 44 years. Also Withers et al. (2019) estimated a decline time of 30–40 + years for areas in the United Kingdom, United States and Finland similarly to Vadas et al. (2018) who found slow estimated decline times of several decades with total cessation of fertilization which was also supported by their measurements.

Still, an overestimation of the decline rate by our model is possible. The model assumptions do not account for several limiting factors such as spatial and temporal heterogeneity in plant-uptake or the climatic conditions that affect P availability. The P decline model uses functions deriving from values obtained in greenhouse pot-experiments. Cumulative P balances from pot experiments after cycles of eight weeks are then extended to yearly P balances in the field. A pot is a much smaller system with a very limited soil volume for extraction while in the field, it is probable that plants take up nutrients from a much larger and less heterogeneous volume, resulting in a less negative balance per kg soil (Heming, 2007; Zicker et al., 2018). Furthermore, in the greenhouse plants are expected to be less stressed by environmental factors such as temperature and pressure by pests. We used standard yields while in the field yields normally fluctuate over the years. For the decline scenarios, a horizon of 0.2 m was used. On one hand, this does not correspond to the sampling depth of 0.1 m of Frossard et al. (2014), on the other hand, one has to consider plant uptake and vertical P transport that goes beyond 0.2 m (Rubaek et al., 2013; Koch et al., 2018; Zicker et al., 2018).

If interferences like P uptake by weeds, losses to the environment and soil biological and chemical processes are considered (Schoumans and Groenendijk, 2000; Cassidy et al., 2017), we understand that the estimated decline times are best-case scenarios.

5. Conclusion

Our findings imply that it is crucial to include artificial drainage in hydrological models to account for catchment specific pathways. The calculated P-decline scenarios, furthermore, emphasize, systems with high P-export are more efficient to decline soil legacy P compared to zero-fertilization practices.

We suggest that management tools should always consider dominant pathways and processes. If it comes to measures to reduce P transfer from artificially drained land, for example by establishing wet arable lands (Gramlich et al., 2018; Szerencsits et al., 2018), trade-offs between erosion risk and subsurface flow are to be accounted for. The use of water logged areas as arable land, where paddy rice or other wet cultures can be planted, has recently gained research interest as an innovative and environmentally friendly alternative to artificially draining such areas (Gramlich et al., 2018; Szerencsits et al., 2018).

It is crucial that political mitigation strategies not only focus on actual P application but also include legacy soil P effects. Best management practices and spatial and temporal restrictions for fertilizer and manure applications do not target the issue of soil legacy P in the long run. Soil P phytoextraction can be a promising approach to effectively reduce soil legacy P while keeping the productivity of the land. For this, it is crucial to set adequate environmental target soil P levels and to identify sound soil P-testing methods.

Still, we suggest that P decline scenarios are to be verified in the fields in long-term experiments. Especially monitoring of deeper soil horizons and different P pools is crucial to make better predictions on decline times. Furthermore, one has to accept the constraints of the phytoextraction strategy for P as for most plants, N or K are more likely to be limiting and therefore fertilization of these nutrients would be necessary (Hejcman et al., 2010). A careful identification of catch crops with multiple benefits (e.g. legumes) can, however, help to overcome such constraints.

Looking back on the history of slurry application coinciding with the actual unequal global distribution of P resources (Cordell and White, 2011; Ulrich and Frossard, 2014), legacy P use is also a strategy towards a more sustainable and resource-efficient economy (Withers et al., 2018) which at the same time emphasizes productivity. However, without an informed and empowered farmer community, taking root of such systems will proceed only very slowly.

Besides legacy P in soils, legacy nutrients in sediments of watersheds are also to be studied to avoid putting pressure on farmers if aquatic processes are underestimated (Sharpley et al., 2011; Chen et al., 2018).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This study was funded for the most part by the state office for agriculture and forest of the Canton of Lucerne (lawa Lucerne). We want to thank Sebastian Kopp and Christoph Joerg for their support during fieldwork. We thank the Canton of Lucerne for providing data, especially Robert Lovas and Ernst Butscher (uwe Lucerne) for long-term data of the catchment streams and Franz Stadelmann (lawa Lucerne) for farmers' data. Furthermore, we thank Juliane Hirte (Agroscope) for her support in statistical analyses and modelling.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.geoderma.2021.115257.

References

Agridea, 2018. Wegleitung Suisse-Bilanz. In: Bundesamt für Landwirtschaft (Ed.), Auflage 1.15, Switzerland.

Alder, S., Prasuhn, V., Liniger, H., Herweg, K., Hurni, H., Candinas, A., Gujer, H.U., 2015. A high-resolution map of direct and indirect connectivity of erosion risk areas to surface waters in Switzerland-A risk assessment tool for planning and policy-making. Land Use Policy 48, 236–249.

Bai, Z.H., Li, H.G., Yang, X.Y., Zhou, B.K., Shi, X.J., Wang, B.R., Li, D.C., Shen, J.B., Chen, Q., Qin, W., Oenema, O., Zhang, F.S., 2013. The critical soil P levels for crop yield, soil fertility and environmental safety in different soil types. Plant Soil 372, 27–37.

Bates, D. M. and Chambers, J. M., 1992. Nonlinear models. Chapter 10 of Statistical Models in S eds J. M. Chambers and T. J. Hastie, Wadsworth & Brooks/Cole.

Becker, R.A., Chambers, J.M., Wilks, A.R., 1988. The New S Language. Wadsworth & Brooks/Cole.

Béguin, J. and Smola, S., 2010. Stand der Drainagen in der Schweiz - Bilanz der Umfrage 2008. Bundesamt für Landwirtschaft, Bern, Schweiz. 24 S.

Beutel, M.W., Horne, A.J., 1999. A review of the effects of hypolimnetic oxygenation on lake and reservoir water quality. Lake Reservoir Manage. 15, 285–297.

Beven, K.J., Kirkby, M.J., 1979. A physically based, variable contributing area model of basin hydrology. Hydrological Science Bulletin 24 (1), 43–69.

Bradley, A.P., 1997. The use of the area under the ROC curve in the evaluation of machine learning algorithms. Pattern Recogn. 30 (7), 1145–1159.

Cassidy, R., Doody, D.G., Watson, C.J., 2017. Impact of legacy soil phosphorus on losses in drainage and overland flow from grazed grassland soils. Sci. Total Environ. 575, 474–484.

Chambers, J. M. and Hastie, T. J., 1992 Statistical Models in S. Wadsworth & Brooks/Cole.

Chen, D., Shen, H., Hu, M., Wang, J., Zhang, Y., Dahlgren, R.A., 2018. Legacy Nutrient Dynamics at the Watershed Scale: Principles, Modeling, and Implications. In: Sparks, D.L. (Ed.), Advances in Agronomy, Vol 149. Elsevier Academic Press Inc, 525 B Street, Suite 1900, San Diego, Ca 92101-4495 USA, 237-313.

Cordell, D., White, S., 2011. Peak Phosphorus: Clarifying the Key Issues of a Vigorous Debate about Long-Term Phosphorus Security. Sustainability 3, 2027–2049.

Couto, R.D., Martini, L.C.P., Gatiboni, L.C., Belli, P., Martins, S.R., Lazzari, C.J.R., Mueller, V., Comin, J.J., Withers, P.J.A., Schenato, R.B., Brunetto, G., 2018. Vulnerability to contamination by phosphorus in a zero-order basin with a high density of pigs and a history of slurry addition: extrapolation of an index. Environmental Earth Sciences 77, 144.

Daly, K., Casey, A., 2005. Environmental aspects of soil phosphorus testing. Irish Journal of Agricultural and Food Research 44, 261–279.

Daly, K., Styles, D., Lalor, S., Wall, D.P., 2015. Phosphorus sorption, supply potential and availability in soils with contrasting parent material and soil chemical properties. Eur. J. Soil Sci. 66, 792–801.

de Alcantara, F.A., Furtini Neto, A.E., Curi, N., de Resende, A.V., 2008. Extraction methods for phosphorus and their relationship with soils phosphorus-buffer capacity estimated by the remaining-phosphorus methodology - A pot study with maize. Commun. Soil Sci. Plant Anal. 39, 603–615.

Dodd, R.J., McDowell, R.W., Condron, L.M., 2012. Predicting the changes in environmentally and agronomically significant phosphorus forms following the cessation of phosphorus fertilizer applications to grassland. Soil Use Manag. 28, 135–147.

Flisch, R., Neuweiler, R., Kuster, T., Oberholzer, H., Huguenin-Elie, O., Richner, W., 2017. 2 / Bodeneigenschaften und Bodenanalysen. In: Grundlagen für die Düngung landwirtschaftlicher Kulturen in der Schweiz (GRUD 2017) (Ed. Richner W. und Sinaj S.). Agrarforschung Schweiz 8 (6), Spezialpublikation, 2/1 – 2/34.

FAL, RAC, FAW, 1996. Schweiz. Referenzmethoden der Eidg. landwirtschaftlichen Forschungsanstalten. Band 1, Bodenuntersuchung zur Düngeberatung. FAL Reckenholz, Zürich.

Fort, F., Cruz, P., Lecloux, E., de Oliveira, L.B., Stroia, C., Theau, J.P., Jouany, C., 2016. Grassland root functional parameters vary according to a community-level resource acquisition-conservation trade-off. J. Veg. Sci. 27, 749–758.

Frossard, E., Demaria, P., Sinaj, S., Schaerer, M., 2014. A flow-through reactor to assess potential phosphate release from agricultural soils. Geoderma 219, 125–135.

Frossard, E., Flura, T., Liebisch, F., Stadelmann, F., 2011. Effet de différentes stratégies de fertilisation phosphatée sur la disponibilité du phosphore du sol dans quelques exploitations agricoles de la région du lac Baldegg. Office fédéral de l`environnement OFEV.

Gaechter, R., 1987. Lake restoration. Why oxygenation and artificial mixing cannot substitute for a decrease in the external phosphorus loading. Swiss Journal of Hydrology 49, 170–185.

Gaechter, R., Wehrli, B., 1998. Ten years of artificial mixing and oxygenation: No effect on the internal phosphorus loading of two eutrophic lakes. Environ. Sci. Technol. 32, 3659–3665.

Gallet, A., Flisch, R., Ryser, J.P., Frossard, E., Sinaj, S., 2003. Effect of phosphate fertilization on crop yield and soil phosphorus status. J. Plant Nutr. Soil Sci. 166, 568–578.

Godlinski, F., Reiche, E.W., Lennartz, B., Meissner, R., 2008. Simulation of phosphorus losses from lysimeters. J. Plant Nutr. Soil Sci. 171, 621–633.

Gramlich, A., Stoll, S., Stamm, C., Walter, T., Prasuhn, V., 2018. Effects of artificial land drainage on hydrology, nutrient and pesticide fluxes from agricultural fields - A review. Agric. Ecosyst. Environ. 266, 84–99.

Hahn, C., Prasuhn, V., Stamm, C., Lazzarotto, P., Evangelou, M.W.H., Schulin, R., 2013. Prediction of dissolved reactive phosphorus losses from small agricultural catchments: Calibration and validation of a parsimonious model. Hydrol. Earth Syst. Sci. 17, 3679–3693.

Hahn, C., Prasuhn, V., Stamm, C., Milledge, D., Schulin, R., 2014. A comparison of three simple approaches to identify critical areas for runoff and phosphorus losses. Hydrol. Earth Syst. Sci. 18, 2975–2991.

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

Haygarth, P.M., Jarvis, S.C., 1999. Transfer of phosphorus from agricultural soils. In: Sparks, D.L. (Ed.), Advances in Agronomy, Vol 66. Elsevier Academic Press Inc, San Diego, pp. 195–249.

Hejcman, M., Schellberg, J., Pavlu, V., 2010. Long-term effects of cutting frequency and liming on soil chemical properties, biomass production and plant species composition of Lolio-Cynosuretum grassland after the cessation of fertilizer application. Appl. Veg. Sci. 13, 257–269.

Heming, S.D., 2007. Phosphorus balances for arable soils in Southern England 1986–1999. Soil Use Manag. 23, 162–170.

Herlihy, M., McCarthy, J., Breen, J., Moles, R., 2004. Effects over time of fertiliser P and soil series on P balance, soil-test P and herbage production. Irish Journal of Agricultural and Food Research 43, 147–160.

Huguenin-Elie, O., Mosimann, E., Schlegel, P., Luescher, A., Kessler, W., Jeangros, B., 2017. 9 / Düngung von Grasland. In: Grundlagen für die Düngung landwirtschaftlicher Kulturen in der Schweiz (GRUD 2017) (Ed. Richner W. und Sinaj S.). Agrarforschung Schweiz 8 (6), Spezialpublikation, 9/1 – 9/22.

Huerdler, J., Prasuhn, V., Spiess, E., 2015. Abschätzung diffuser Stickstoff- und Phosphoreinträge in die Gewässer der Schweiz. MODIFFUS 3.0. In: Bericht z.H. Bundesamt für Umwelt (BAFU), p. 117 S..

H2O.ai Team, 2017. h2o: R Interface for H2O. R package version 3.10.5.3. https://CRAN.R-project.org/package=h2o.

Jarvie, H.P., Sharpley, A.N., Flaten, D., Kleinman, P.J.A., 2019. Phosphorus mirabilis: Illuminating the Past and Future of Phosphorus Stewardship. J. Environ. Qual. 48, 1127–1132.

Johnston, A.E., Poulton, P.R., White, R.P., Macdonald, A.J., 2016. Determining the longer term decline in plant-available soil phosphorus from short-term measured values. Soil Use Manag. 32, 151–161.

Kleinman, P.J.A., 2017. The Persistent Environmental Relevance of Soil Phosphorus Sorption Saturation. Current Pollution Reports 3, 141–150.

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

Kleinman, P.J.A., Srinivasan, M.S., Dell, C.J., Schmidt, J.P., Sharpley, A.N., Bryant, R.B., 2006. Role of rainfall intensity and hydrology in nutrient transport via surface runoff. J. Environ. Qual. 35, 1248–1259.

Koch, M., Kruse, J., Eichler-Löbermann, B., Zimmer, D., Willbold, S., Leinweber, P., Siebers, N., 2018. Phosphorus stocks and speciation in soil profiles of a long-term fertilizer experiment: Evidence from sequential fractionation, P K-edge XANES, and 31P NMR spectroscopy. Geoderma 316, 115–126.

Koopmans, G.F., Römkens, P.F.A.M., Song, J., Temminghoff, E.J.M., Japenga, J., 2007. Predicting the Phytoextraction Duration to Remediate Heavy Metal Contaminated Soils. Water Air Soil Pollut. 181, 355–371.

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

Lindstrom, M.J., Bates, D.M., 1990. Nonlinear Mixed Effects Models for Repeated Measures Data. Biometrics 46, 673–687.

Liu, J., Elliott, J.A., Wilson, H.F., Baulch, H.M., 2019. Impacts of Soil Phosphorus Drawdown on Snowmelt and Rainfall Runoff Water Quality. J. Environ. Qual. 48, 803–812.

Macintosh, K.A., Doody, D.G., Withers, P.J.A., McDowell, R.W., Smith, D.R., Johnson, L. T., Bruulsema, T.W., O'Flaherty, V., McGrath, J.W., 2019. Transforming soil phosphorus fertility management strategies to support the delivery of multiple ecosystem services from agricultural systems. Sci. Total Environ. 649, 90–98.

McDowell, R.W., Withers, P.J.A., van der Weerden, T.J., 2017. The Environmental Impact of Fertiliser Nutrients on Freshwater. In: Hester, R.E., Harrison, R.M. (Eds.), Agricultural Chemicals and the Environment: Issues and Potential Solutions, 2nd Edition. Royal Soc Chemistry, Thomas Graham House, Science Park, Cambridge Cb4 4wf, Cambs, Uk, 20-44.

McDowell, R., Dodd, R., Pletnyakov, P., Noble, A., 2020. The Ability to Reduce Soil Legacy Phosphorus at a Country Scale. Front. Environ. Sci. 8, 6.

Moosmann, L., Muller, B., Gachter, R., Wuest, A., Butscher, E., Herzog, P., 2005. Trend-oriented sampling strategy and estimation of soluble reactive phosphorus loads in streams. Water Resources Research 41, 1020.

Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I -A discussion of principles. J. Hydrol. 10 (3), 282–290.

Ohno, T., Zibilske, L.M., 1991. Determination of low concentrations of phosphorus in soil extracts using malachite green. Soil Sci. Soc. Am. J. 55, 892–895.

OpenStreetMap contributors, Available at www.openstreetmap.ch 2019 accessed at September 16. 2019.

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

R Core Team, 2016. R: A language and environment for statistical computing. R Foundation for statistical computing, Vienna, Austria https://R-project.org/.

Rubaek, G.H., Kristensen, K., Olesen, S.E., Ostergaard, H.S., Heckrath, G., 2013. Phosphorus accumulation and spatial distribution in agricultural soils in Denmark. Geoderma 209, 241–250.

Schoumans, O.F., Chardon, W.J., 2003. Risk assessment methodologies for predicting phosphorus losses. J. Plant Nutr. Soil Sci. 166, 403–408.

Schoumans, O.F., Chardon, W.J., 2015. Phosphate saturation degree and accumulation of phosphate in various soil types in The Netherlands. Geoderma 237, 325–335.

Schoumans, O.F., Groenendijk, P., 2000. Modeling soil phosphorus levels and phosphorus leaching from agricultural land in the Netherlands. J. Environ. Qual. 29, 111–116.

Schulte, R.P.O., Melland, A.R., Fenton, O., Herlihy, M., Richards, K., Jordan, P., 2010. Modelling soil phosphorus decline: Expectations of Water Framework Directive policies. Environ. Sci. Policy 13, 472–484.

Sharpley, A.N., Kleinman, P.J.A., Flaten, D.N., Buda, A.R., 2011. Critical source area management of agricultural phosphorus: experiences, challenges and opportunities. Water Science & Technology 64 (4), 945–952.

Shore, M., Jordan, P., Mellander, P.E., Kelly-Quinn, M., Wall, D.P., Murphy, P.N.C., Melland, A.R., 2014. Evaluating the critical source area concept of phosphorus loss from soils to water-bodies in agricultural catchments. Sci. Total Environ. 490, 405–415.

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

Stuenzi, H., 2006. Die P-Bodenextraktionsmethoden mit Wasser und CO_2-Wasser. Agrarforschung Schweiz 13 (7), 284–289.

Szerencsits, E., Prasuhn, V., Churko, G., Herzog, F., Utiger, C., Zihlmann, U., Walter, T., Gramlich, A., 2018. Karte potentieller Feucht-(Acker-)Flächen in der Schweiz. Agroscope Science 72, 1–68.

Thomas, I.A., Jordan, P., Mellander, P.E., Fenton, O., Shine, O., Huallachain, D.O., Creamer, R., McDonald, N.T., Dunlop, P., Murphy, P.N.C., 2016a. Improving the identification of hydrologically sensitive areas using LiDAR DEMs for the delineation and mitigation of critical source areas of diffuse pollution. Sci. Total Environ. 556, 276–290.

Thomas, I.A., Mellandera, P.E., Murphy, P.N.C., Fenton, O., Shine, O., Djodjic, F., Dunlop, P., Jordan, P., 2016b. A sub-field scale critical source area index for legacy phosphorus management using high resolution data. Agric. Ecosyst. Environ. 233, 238–252.

Ulrich, A.E., Frossard, E., 2014. On the history of a reoccurring concept: Phosphorus scarcity. Sci. Total Environ. 490, 694–707.

uwe: state office for environment and energy of the Canton of Lucerne, 2018. Dataset (Microsoft-Excel) of P-load, P-concentration and flows in the Lake Baldegg catchment. (unpublished).

Vadas, P.A., Fiorellino, N.M., Coale, F.J., Kratochvil, R., Mulkey, A.S., McGrath, J.M., 2018. Estimating Legacy Soil Phosphorus Impacts on Phosphorus Loss in the Chesapeake Bay Watershed. J. Environ. Qual. 47, 480–486.

van der Salm, C., Chardon, W.J., Koopmans, G.F., van Middelkoop, J.C., Ehlert, P.A.I., 2009. Phytoextraction of Phosphorus-Enriched Grassland Soils. J. Environ. Qual. 38, 751–761.

Wang, Z., Zhang, T., Tan, C.S., Qi, Z., 2020. Modeling of phosphorus loss from field to watershed: A review. J. Environ. Qual. 49, 1203–1224.

White, M.J., Storm, D.E., Busteed, P.R., Stoodley, S.H., Phillips, S.J., 2009. Evaluating Nonpoint Source Critical Source Area Contributions at the Watershed Scale. J. Environ. Qual. 38, 1654–1663.

Withers, P.J.A., Doody, D.G., Sylvester-Bradley, R., 2018. Achieving Sustainable Phosphorus Use in Food Systems through Circularisation. Sustainability 10, 1804.

Withers, P.J.A., Hodgkinson, R.A., Rollett, A., Dyer, C., Dils, R., Collins, A.L., Bilsborrow, P.E., Bailey, G., Sylvester-Bradley, R., 2017. Reducing soil phosphorus fertility brings potential long-term environmental gains: A UK analysis. Environmental Research Letters 12, 63001.

Withers, P.J.A., Vadas, P.A., Uusitalo, R., Forber, K.J., Hart, M., Foy, R.H., Delgado, A., Dougherty, W., Lilja, H., Burkitt, L.L., Rubaek, G.H., Pote, D., Barlow, K., Rothwell, S., Owens, P.R., 2019. A Global perspective on integrated strategies to manage soil phosphorus status for eutrophication control without limiting land productivity. J. Environ. Qual. 48, 1234–1246.

Wuenscher, R., Unterfrauner, H., Peticzka, R., Zehetner, F., 2015. A comparison of 14 soil phosphorus extraction methods applied to 50 agricultural soils from Central Europe. Plant Soil and Environment 61, 86–96.

Zhang, T.Q., Tan, C.S., Wang, Y.T., Ma, B.L., Welacky, T., 2017. Soil phosphorus loss in tile drainage water from long-term conventional- and non-tillage soils of Ontario with and without compost addition. Sci. Total Environ. 580, 9–16.

Zicker, T., von Tucher, S., Kavka, M., Eichler-Lobermann, B., 2018. Soil test phosphorus as affected by phosphorus budgets in two long-term field experiments in Germany. Field Crops Research 218, 158–170.