

Response to the Editor:

Dear Prof. Veldkamp,

We thank you and the two reviewers for the useful feedback to the revised version of our paper. We have addressed the comments and have uploaded an updated version.

5

In the following we respond to the individual comments (in red color).

Best regards,

10 Michael Rinderer and co-authors

Editor Comment

15 Dear Dr. Rinderer and colleagues,

I have now read your revised manuscript and the reviews provided by the two referees.

You will also see that the referees made some contrasting recommendations based on your revised manuscript. Referee #2 is very positive about the changes in the direction of the manuscript, Referee #1 expresses some concerns, although this referee continues to be very positive about the overall scientific significance and quality of your study.

20

I would like to commend you for taking the reviews seriously and making considerable changes to your original manuscript based on the earlier reviews. I think that the manuscript has improved substantially, both in focus and the message it conveys.

Nevertheless, I would recommend some additional changes which I would consider relatively minor:

25 *-Referee 1 expresses concern that the importance of nutrient flushing is not sufficiently stressed in your discussion. My impression from your results is that the nutrient flushing was not contributing enough (both in terms of waterflow and P concentration) to have a significant effect on the annual P losses from your ecosystem and so your approach to calculate annual P fluxes appeared reasonable. However, referee #1 makes a good point when referring to line 400, in which you cite another study which claims that up to 40% of the annual P flux may occur during single events. Does this mean that your study*
30 *results deviate from this earlier study? To address the concerns of Referee #1, I suggest to include in your discussion a short section in which you address what the potential error is when nutrient flushing during single events is ignored in your annual budget.*

We are aware that the P losses during the flushing phase are neglected in this simple calculation of the annual P losses. To be clear that this part of the paper is a hypothetical test based on assumptions we now write at the beginning of section 4.3: *“If one would argue that the amount of SSF during the flushing period is small compared to the remaining part of the event (see Fig. 3) one could also assume that chemotactic conditions prevail during large rainfall events. Given these simplifications, one can roughly estimate the annual P losses by knowing the amount of annual SSF and an average P concentration.”*

We have deliberately cited the study by Julich (personal communication cit. in Boll et al. (2016)) to show that we are self-critical about these simplifications and consider these as potential reasons why some other studies cited in a meta-analysis by Sohrt et al. (2017) show higher values.

And at the end of the second paragraph of section 4.3 Line 419 we write: *“Given the high P concentrations during the flushing phase of our sprinkling events it might be that relatively small amounts of SSF during the first flush can still export a significant amount of P.”*

-Referee #1 also suggests to include the effect of different tree species. Because your study sites were all dominated by beech trees I do not see that your study can contribute much new information to this interesting question. As your manuscript is already relatively long, I suggest not to explore this further in your present manuscript.

We agree that studying differences between tree species is an interesting topic but as you said, we do not have the data to contribute to it.

Both referees have also added a list of specific comments, which I would like you to address. Especially the statistics should be described if you conducted them.

We have addressed each comment by reviewer 1 and 2. In terms of the statistics we added the information requested by reviewer 2 in the Method section and repeated this information later where the reviewer asked us to do.

A few additional suggestions:

-In your response you write about ‘quantitative, hypothesis-driven ... test’ (l.49). Unfortunately, you did not write out your hypotheses in the introduction, which I strongly recommend.

We have added hypothesis at the end of the introduction section.

-Fig.1: I suggest to add error bars to the measured bulkdensity data and to connect the bulkdensity data with straight lines (unless you have a good scientific reason why you used non-linear interpolation).

We added error bats and changed the type of line connectors.

-L. 250: I suggest to use 'detection limit' instead of 'limit of quantification'. 'Detection limit' is not only used for analyses done by analytical equipment but can also include data processing steps.

We distinguish between the limit of quantitation for P_{tot} (in our case 0.009 mg/L) and the limit of detection (in our case 0.004 mg/l) based on DIN 32645 with a significance bound of 99 % for the limit of quantitation and 77 % for the limit of detection (DIN, 2008). Therefore, we need to keep the expression in figures and tables.

70

I look forward to your revisions,

Best regards,

75

Edzo Veldkamp

Thank you for handling our review process

80 Fond regards

Michael Rinderer and co-authors

Response to the Reviewers:

85

Reviewer 1

Anonymous Referee #1:

The performed revisions improve the quality and usefulness of the manuscript. From my point of view the most interesting finding is that observed chemostatic transport conditions suggest that annual P losses at the lateral and vertical boundary of a forest plot can be approximated by knowing the average P concentration and the water fluxes in forest soils. The authors now point this out nicely. Other findings, such as that annual P losses are negligible compared to P stocks are not as interesting since it is well known that P is much more retained in ecosystems than other nutrients. But I think the manuscript is ready to be published now. Just some minor corrections below.

95 **We thank reviewer 1 for his/her positive feedback and suggestions that we have included in the new version of the manuscript. Despite the fact that our rough estimation of the annual P losses from the forest ecosystem based on subsurface flow data and an average P concentration we want to emphasize that work still needs to be done to compare the P losses during the flushing period with the losses during the remaining part of an event.**

100 *Minor comments*

l. 14. Rates do not have the unit “time” but the unit “time⁻¹”. If you are talking about replenishment times please use correct wording

Indeed; we have replaced “rates” with “times”

105 *l. 28 “be-weekly” change to bi-weekly*

Thanks for pointing it out.

l. 38 Again, turnover rate has the unit time⁻¹, so please call it turnover time if you are talking about time

We have replaced “rates” with “times”

110

l. 327 Remove comma before “why”

We have removed the comma

l. 333 “Dynamics” should not be capitalized

115 Originally, we followed the rule to capitalize all nouns in headlines. We have now changed all headlines including the title to have no capitalization except for the first word.

l. 361 This sentence is strange. Since the experiment was in spring (beginning of vegetation period) and summer (full vegetation) I don't understand why you can say that the results confirm "higher P concentrations of leachate from the forest floor during the vegetation period". As far as I know there were no measurements in winter in this experiment

120 We agree that we have no measurements during winter and therefore cannot strictly compare our results with other work.

l. 364 change "form" to "from"

Thanks for pointing it out!

125 *l. 426 average P concentration is vague. Please specify concentration in what. Soil solution? Leachate?*

We added soil solution to be clear.

l. 427 Delete "quick"

We have deleted quick

130

l. 429 Add a comma after "budget"

Thanks for pointing it out!

l. 430 Replace first than with as

135 Thanks for pointing it out!

Reviewer 2

Anonymous Referee #2:

General comments

140 *The manuscript entitled „Subsurface Flow and Phosphorus Dynamics in Beech Forest Hillslopes during Sprinkling Experiments: How fast is Phosphorus replenished?“ written by Michael Rinderer, Jaane Krüger, Friederike Lang, Heike Puhmann, and Markus Weiler was rewritten according to the suggestions of the editor and two referees. The main concern was that P fluxes were not presented, even though one of the objectives was to quantify P losses with subsurface flow. The authors substantially rewrote the manuscript. They focus now on soil solution P replenishment and P buffering capacity of the*
145 *mineral soil as these are important processes for plant P nutrition. In addition, they make a rough estimation of the annual P flux from one of their research sites.*

Unfortunately, I have more concerns with the new version of the manuscript than with the first one. These are the reasons:

1. Now the authors try to meet two goals, which makes the manuscript more difficult to read. They start the introduction with plant P nutrition issues, but do not come back to them in the discussion. They should for example discuss the implications of
150 *P leaching from the forest floor and P retention in the mineral soil for plant nutrition. What about the effect of different tree species? They cite several studies from other beech forest ecosystems, but are there also results from other tree species? Nutrient flushing might be of great importance for plant nutrition. What are the processes that lead to P flushing? What role do microorganisms play?*

155 **We thank the reviewer 2 for his/her useful comments. We have included a text in the discussion section, line 421 that discusses implications of P leaching from the forest floor and P retention in the mineral soil for plant nutrition. We also end the abstract and the collusion section with this thought.**

We agree that a similar study would be interesting in forests dominated by other species but as the Editor mentioned, we have only worked in beech forest stands and have no data to contribute to this interesting toic. We therefore found it more useful to
160 **compare our results with other studies in beech forests.**

2. In order to be able to estimate the annual P flux of one of the three research sites, the authors used the P concentration after the nutrient flush and ignored the high soil solution P concentrations during the first two hours of the experiment (nutrient flush). In the first version of the manuscript the authors wrote that climate change may lead to more frequent high-intense
165 *rainfall events. Hence, nutrient flushing might become more relevant and it would be interesting how large the P flux would be during such high-intense rainfall events (not annually). In addition, the authors write in line 400 that “up to 40 % of the annual P flux might occur during single events, which would suggest that the fluxes during the first flush have an important*

share on the annual flux and cannot be neglected". The P flux during the sprinkling experiment (first flush and rest of the time) could have been calculated for all three research sites.

170 Hence, I would suggest to either focus on P replenishment and buffering or P fluxes from the ecosystem and to discuss the chosen topic in more detail.

We are aware that the P losses during the flushing phase are neglected in this simple calculation of the annual P losses. To be clear that this part of the paper is a hypothetical test based on assumptions we now write at the beginning of section 4.3: "If one would argue that the amount of SSF during the flushing period is small compared to the remaining part of the event (see Fig. 3) one could also assume that chemotactic conditions prevail during large rainfall events. Given these simplifications, one can roughly estimate the annual P losses by knowing the amount of annual SSF and an average P concentration."

175 We have deliberately cited the study by Julich (personal communication cit. in Boll et al. (2016)) to show that we are self-critical about these simplifications and consider these as potential reasons why some other studies cited in a meta-analysis by Sohr et al. (2017) show higher values.

185 And at the end of the second paragraph of section 4.3, line 419 we write: "Given the high P concentrations during the flushing phase of our sprinkling events it might be that relatively small amounts of SSF during the first flush can still export a significant amount of P."

Mentioning P replenishment/buffering and P fluxes in one paper is a result of feedback during the first round of reviews.

190 *Specific comments*

L 170-190 Still, the statistical methods used are not described. In line 247 you write "was significantly higher", but it is not explained anywhere how you tested this.

We have added the used statistical tests in section 2.3. To be very clear we repeated information on statistical tests in line 259.

195 L 205 The trench cannot yield vertical SSF.

Yes, this was a mistake. We have corrected this.

L 297 You found that especially vertical SSF is of importance for P transport. Hence, suction cups installed in a large number and spread over a research site, should yield good estimates of soil solution P concentrations, especially when soil heterogeneity is high.

200

We agree that suction caps can yield interesting information on soil solution P. They have advantages in terms of spatial coverage. As suction cups cannot be used to measure subsurface storm flow, we think that the large lysimeter plates were a good choice for the purpose of our study.

205 *L 336-339 Explain the process behind nutrient flushing.*

We have included a paragraph in section 42, line 353ff about processes leading to high P concentrations during the onset of SSF.

L 367-368 This is only right if the nutrient flush at the beginning of the experiment is not relevant. You did not calculate the flux for the flush, hence, I cannot confirm that your statement is right.

Yes, agreed. We have chosen a misleading way of writing. It was not meant to leave the impression we deliberately neglect the flushing phase. We think that the old text of section 4.3 had mentioned that by saying “A possible reason for the higher values compared to our results could be that we used mean P concentration measured after the first flush which results in rather conservative estimates.” Nevertheless, we now have improved the beginning of the section. We now write:

215 *“If one would argue that the amount of SSF during the flushing period is small compared to the remaining part of the event (see Fig. 3) one could also assume that chemotactic conditions prevail during large rainfall events. Given these simplifications, one could roughly estimate annual P losses from forest stands by knowing the amount of annual SSF and an average P concentration. (...)”.*

220 *And at the end of the second paragraph of section 4.3 we now additionally write: (...) Further research is needed to clarify the role of P-flushing on the total P loss from forest ecosystems. Given the high P concentrations during the flushing phase of our sprinkling events it might be that relatively small amounts of SSF during the first flush can still export a relatively large amount of P.”*

225 *L 397 Which ecosystems are included in the meta-analysis of Sohrt et al. (2019)?*

It is a worldwide compilation of data of studies in mainly cold temperate and boreal forest ecosystems.

Technical corrections

L 243 6 time -> 6 times

230 **Thanks for pointing this out**

L 398 either delete the colon or the parentheses

Thanks for pointing this out

235 *L 419 suggest -> suggests*

Thanks for pointing this out

L 431 ecosystem -> ecosystems

Thanks for pointing this out

240

L 432 relative -> relatively

Thanks for pointing this out

245 **Subsurface Flow and Phosphorus Dynamics in Beech Forest Hillslopes during Sprinkling Experiments: How fast is Phosphorus replenished?**

Michael Rinderer¹, Jaane Krüger², Friederike Lang², Heike Puhlmann³, Markus Weiler¹

250 ¹ Chair of Hydrology, University of Freiburg, Freiburg, Germany

² Chair of Soil Ecology, University of Freiburg, Freiburg, Germany

³ Forstliche Versuchs- und Forschungsanstalt Department of Soil and Environment, Forest Research Institute Baden-Württemberg, Freiburg, Germany

255 *Correspondence to:* Michael Rinderer (michael.rinderer@hydrology.uni-freiburg.de)

Abstract. The Phosphorus (P) concentration of soil solution is of key importance for plant nutrition. During large rainfall events, the P concentration is altered by lateral and vertical subsurface storm flow (SSF) that facilitates P mobilization, redistribution within the soil profile and potential P export from the ecosystem. These processes are not well studied under field conditions. Important factors of the replenishment of P concentrations in soil solutions are the rate of P replenishment (by biotic and abiotic processes) and the P buffering capacity of soils. Lab experiments have shown that replenishment rates can vary between minutes and months. The question remains how P concentrations in lateral and vertical SSF vary under natural field conditions. We present results of large-scale sprinkling experiments simulating 150 mm throughfall at 200 m² plots on hillslopes at three beech forests in Germany. We aimed at quantifying lateral and vertical SSF and associated P concentrations in the forest floor, the mineral soil and the saprolite during sprinkling experiments in spring and summer. The sites differed mainly in terms of soil depth, skeleton content and soil P stock (between 189 g/m² and 624 g/m² in the top 1 m soil depth). Vertical SSF in the mineral soil and in the saprolite was at least two orders of magnitude larger than lateral SSF in the same depth. Vertical and lateral SSF consisted mainly of pre-event water that was replaced by sprinkling water. Higher P concentrations in SSF in the first 1 to 2 h after onset of SSF indicated nutrient flushing, but P concentrations in the mineral soil and saprolite were nearly constant thereafter for most of the experiment despite strong increase in SSF. This suggests that P in the soil solution at all three sites was replenished fast by mineral or organic sources. The observed chemostatic transport conditions suggest that would dominate in SSF, annual P losses at the lateral and vertical boundary of a forest plot can be approximated by knowing the average P concentration and the water fluxes in forest soils. A rough estimation of the annual P loss based on this simplified assumption for one of our sites with longer SSF data, resulted in an annual P loss of 3.16 mg/m²/a from the forest plot. This P loss is similar to estimates from a previous study at the same site using bi-weekly groundwater samples. Our approximated annual P loss in SSF was in a similar order of magnitude as P input by dry and wet deposition and by mineral weathering, suggesting. Despite the fact that P losses by SSF from the ecosystem seem to be small, the

translocation of P from the forest floor to become relevant particularly the mineral soil might be of high relevance, at long time scales sites with low P stocks where the forest floor is the dominant source for the P nutrition of trees.

Formatted: English (United Kingdom)

Formatted: English (United Kingdom)

Formatted: English (United Kingdom)

1 Introduction

280 Phosphorus (P) is a major component of plant nutrition and has been reported to be a potential limiting factor of primary productivity in forest ecosystems (Achat et al., 2016; Elser et al., 2000, 2007). In the last decades a decrease of foliar P concentrations and an increase in the nitrogen to phosphorus (N:P) ratio has been observed in forests (Braun et al., 2010; Duquesnay et al., 2000; Jonard et al., 2015; Kabeya et al., 2017). From a plant nutrition perspective the replenishment of the P concentration of the soil solution is of key importance for the fertility of a soil (Lambers et al., 2008). The replenishment is a function of the exchange rate between the soil solution and the solid phase including biotic and abiotic processes. Helfenstein et al. (2018) present a worldwide compilation of P turnover times for 217 soils that range between 10^{-2} to 10^6 minutes for which the majority have a turnover time of 1 to 100 minutes. The same authors also found a negative relation between the concentration of the soil solution P and the turnover time. In our paper we use the term P *replenishment* to describe the replenishment of P by both, biotic and abiotic processes. A high rate of replenishment implies that the soil has many potential binding sites where P can be sorbed or precipitated or the soil has many microorganisms that can immobilize or mobilize P. Thus, the P replenishment is also related to the P buffering capacity which is defined as the ability of a soil to moderate changes in the concentration of soil solution P (Pypers et al., 2006; White and Beckett, 1964).

The P dynamic in subsurface storm flow (SSF) is an indicator for dilution and enrichment processes along the lateral and vertical flow paths in the soil (Bol et al., 2016; Heathwaite and Dils, 2000; Julich et al., 2017a; Steegen et al., 2001)(Bol et al., 2016; Heathwaite and Dils, 2000; Julich et al., 2017a; Steegen et al., 2001). Unlike for agricultural land only very few studies exist that quantify the dynamics of P concentrations in forest ecosystems, as concentrations are low and measuring vertical and lateral SSF is challenging. In a small number of field studies water was sampled from surface waters i.e., forest streams (Cole and Rapp, 1981; Gottselig et al., 2017; Kunimatsu et al., 2001; Schindler and Nighswander, 1970; Taylor et al., 1971; Zhang et al., 2008) or from groundwater wells in riparian zones near the stream that are easier accessible than SSF (Carlyle and Hill, 2001; Fuchs et al., 2009; Vanek, 1993). But elemental composition of stream water represents an integrated signature of the entire catchment and is therefore not appropriate for a detailed process identification within soil compartments. Stream water is also subject to in-stream retention and mobilization of P and therefore not necessarily representative of transport conditions in hillslopes (Gregory, 1978; Hill et al., 2010; Mulholland and Hill, 1997; Sohr et al., 2019; Stelzer et al., 2003). Therefore, other studies collected soil solution below forest stands using suction lysimeters (Cole and Rapp, 1981; Compton and Cole, 1998; Kaiser et al., 2000; Qualls et al., 2002). A very limited number of studies used a trench to measure P concentration in lateral SSF. Timmons et al. (1977) were among the first; they installed a 1.8 m long drainage at the intersection of the A and B horizon in 33 cm soil depth to measure water fluxes and P concentrations in an aspen-birch forest in Minnesota. Sohr et al. (2018) used 10 m wide trenches to analyze lateral subsurface flow from the forest floor, mineral soil and saprolite

at three beech forest stands in Germany during natural rainfall events. Jackson et al. (2016) performed an artificial sprinkling experiments on a 200 m² plot on a hillslope at the Savannah River Research Site (South Carolina, USA) covered by pine trees and measured water flux and P concentrations in lateral SSF at 1.25 m soil depth. They applied dye- and conservative tracers to identify dominant flow paths and fractions of event and pre-event water involved with solute transport. Makowski et al. (2020) performed artificial sprinkling experiments on 3 m² plots and reported significantly higher P concentrations in the sampled vertical SSF in the first two hours (P flushing), followed by decreasing and finally constant P concentrations for the rest of the experiment.

From field studies as described above and from lab experiments using soil columns we know that transport of P during high intensity rainfall events occurs mainly along preferential flow paths (Cox et al., 2000; Fuchs et al., 2009; Julich et al., 2017b; Missong et al., 2018a). Preferential flow paths allow subsurface flow to bypass the soil matrix that otherwise has been shown to effectively retain P (Compton and Cole, 1998; Ilg et al., 2009; Johnson et al., 2016; Qualls et al., 2002). Moreover, biopores (e.g., earthworm borrows, root channels) have been identified as biochemical hotspots that can show significantly higher P concentrations than the soil matrix, especially in fine textured soils (Backnäs et al., 2012; Hagedorn and Bundt, 2002). Preferential flow paths can extend below the rooting depth and are therefore considered an important pathway of P losses from forest ecosystems (Julich et al., 2017a; Sohr et al., 2017).

Here we present a comparative study based on hillslope-scale artificial sprinkling experiments on 200 m² plots at three beech (*Fagus sylvatica*) forest sites that differ in terms of their soil depth, skeleton content and SSF flowpaths as well as their soil P stocks. We used an experimental setup that allowed to measure soil-depth specific lateral and vertical flow of water and transported hydrochemicals (incl. P). We performed two sprinkling experiments at each site to capture potential differences in P dynamics within the vegetation period (i.e., spring and summer/fall). The rationale is that microbial activity, which is mainly responsible for P mineralization, is strongly dependent on moisture and temperature conditions (Brinson, 1977; Kirschbaum, 1995). In particular, we address the following research questions:

1. What are the main runoff generation mechanisms, flow paths and temporal delays in runoff response (lateral versus vertical, shallow versus deep) in three forest stands with different soil properties (i.e., soil texture, skeleton content, soil depth) during long, moderately intense sprinkling events?
2. What is the dynamic of P concentrations in lateral and vertical SSF, measured at different soil depths during artificial sprinkling events and does it differ seasonally and among sites with different physical soil properties and soil P availability?
3. Is P in the soil solution diluted during large sprinkling experiments or is the rate of P replenishment at all sites and all soil depths high enough to facilitate constant P concentrations (chemostatic conditions)?

We hypothesize that

1. in coarser textured soils vertical flow paths dominate and lead to faster and higher SSF at deeper depths than in fine textured soils.

2. P concentrations in SSF from the forest floor are high compared to the mineral soil and saprolite where P concentrations are lower due to adsorption processes. In summer the P concentrations in SSF are higher due to higher microbial activity and drought effects.

3. P concentrations in SSF are not diluted by the large amount of sprinkling water as P stocks are by far larger than P losses during a single event and as P replenishment is expected to be quick.

2. Methods

2.1 Study Sites

For this analysis three beech (*Fagus sylvatica*) forest sites in Germany with contrasting soil hydrological properties (i.e., skeleton content and soil depth) and P stocks were selected. Their site characteristics are summarized in Tab. 1. Mitterfels (MIT) (48° 58' 32" N; 12° 52' 37" E) is located ca. 70 km east of Regensburg in the Bavarian Forest at 1023 m a.s.l. Its mean annual precipitation is 1299 mm. The site is characterized by Hyperdystric chromic, folic cambisol with a loamy topsoil (0 – 35 cm) and a sandy-loamy subsoil (35 – 130 cm). The stone content in the top- and subsoil is 23 % and 26 % and the P stock in the upper 1 m of the soil profile is 678 g/m². The saprolite reaches a total depth of 7 m but is less weathered below 2 m depth. The parent material below the saprolite is paragneiss. Conventwald (CON) (48° 01' 16" N; 7° 57' 56" E) is located 20 km east of Freiburg in the Black Forest at 840 m a.s.l. and has a mean annual rainfall of 1749 mm. The parent material, main soil type and vegetation is similar to MIT but soils considerably differ in the skeleton content (CON: 87 % topsoil, 67 % subsoil), the depth of the saprolite (CON 17 m but less weathered below 3 m) and the P stock in the upper 1 m of the soil profile (CON: 231 g/m²). Tuttingen (TUT) (47° 58' 42" N; 8° 44' 50" E) is located 125 km south of Stuttgart at 835 m a.s.l. and has 900 mm annual rainfall. Due to its carbonatic parent material a rendzic Leptosol with a clayey top- and subsoil has developed (P stock 209 g/m²). The soil profile has a 20 - 40 cm deep topsoil and a 60 - 80 cm deep subsoil directly overlaying the fractured carbonate parent material. The stone content of the top- and subsoil is 50 % and 67 %, respectively. The site is also covered by beech but the stand is younger than at MIT and CON. Bulk soil bulk density of CON and TUT is more similar than compared to MIT, but all three soil profiles show considerably variation in the bulk density with depth (Fig. 1).

2.2. Experimental Setup and Lab Analysis

At each of the three sites we delineated an experimental plot of 200 m² (10 m by 20 m) which was separated from its uphill neighboring area by a plastic foil inserted into the soil profile. At the downhill side of the plots, a trench (TR) was dug down into the saprolite with a hydraulic shovel excavator, and drainage mats and drainage pipes were installed in three (MIT, CON) or two (TUT) depths (Fig. 2) to collect lateral flow. The actual depth of the pipes varied according to the site-specific soil profile (Tab. 2), but was chosen such that water draining from the forest floor (L, Of, Oh), the mineral soil (A and B horizon) and the saprolite (Cw) could be sampled. Plastic foil was installed across the entire 10 m width of the trench and down to the depth of the three soil compartments so that all water flowing laterally towards the trench was captured in the appropriate

drainage pipe. To measure also vertical flow, we installed zero tension lysimeters (LY) for which we used steel piling plates with a dimension of 1.0 m by 0.6 m. To install them, a trench was dug at the side of the hillslope and the steel piling plates were pushed from the side into the undisturbed soil profile with heavy duty hydraulic jacks. By this, effects on soil structure by excavation and refill were prevented and the mixing of soil P stocks from different soil horizons was avoided. We installed the LY in similar depths as the TR. At MIT and CON we installed an additional LY right below the A-horizon in 30 to 40 cm soil depth; at TUT the shallow soil depth did not allow installing a LY below 60 cm. In the following, the TR and LY are numbered as TR1B to TR3B and LY1B to LY4B with increasing soil depth (Tab. 2). ("B" indicates the sprinkling plot and allows distinction from another dataset not used in this paper but mentioned in subsequent publications). All trenches were backfilled after installation.

In the upper and lower half of the hillslope, volumetric water content and soil temperature were monitored at two soil profiles in 20, 40, 60, 80 and 120 cm soil depth (no 120 cm measurements at TUT due to shallower soil). We used SMT100 sensors (Truebner GmbH) attached to CR1000 data loggers (Campbell Scientific) and monitored volumetric water content and soil temperature in 5 min time intervals.

At each plot we performed two artificial sprinkling experiments – one in spring at the start of the growing season and one in late summer/early fall during or towards the end of the growing season but well before leaf senescence. The two periods were chosen to reflect potential seasonal differences in soil moisture and P supply, i.e., higher soil moisture content after snowmelt but less P mineralization due to colder temperature in spring, versus drier soil moisture conditions and advanced P mineralization with warmer soil temperature in summer/fall. The mean difference of the median volumetric water content over the 7 days preceding the experiment was 8 vol %, 3 vol % and 8 vol % between summer and spring for MIT, CON and TUT, respectively. We sprinkled the 200 m² plots with a mean rainfall intensity of 10 to 20 mm/h over 10 to 12 h. Although, this reflects a large rainfall event with a return period of more than 100 years for all sites (DWD Climate Data Center, 2010), the chosen rainfall intensity allowed all irrigation water to infiltrate into the soil and did not generate surface runoff. 60'000 liters of water were trucked to the site and run through an industrial deionizer (VE-300 (6x50 Liter), AFT GmbH & Co.KG) to reduce the mineral content to low levels typically found in natural rainfall. In our case the sprinkling water had an electrical conductivity of around 20 µS/cm. The sprinkling water was stored in a large water pillow (60'000 L, Sturm Feuerschutz GmbH) between 50 and 100 m above the sprinkling plot. The resulting hydrostatic gradient was sufficient to run the sprinkling without a pump. The six radial sprinklers (Xcel-wobbler and pressure regulator manufactured by Senninger) installed at a height of 2 m sprinkled 60 % of the total water onto the 200 m² plot and 40 % outside to reduce boundary effects with the otherwise dry surrounding area. 1 kg (first sprinkling experiment) and 2 kg (second experiment) of 99.96 atom% deuteriumoxid was added while filling the water pillow to elevate the natural background deuterium isotopic signature of the sprinkling water by ca. 100 permille. Water samples before and after adding the deuterium and during sprinkling (collected with totalizers on the experimental plot) were collected to measure the background isotopic composition and to check for a constant isotopic label signature over the course of the experiment.

The subsurface flow (SSF) from the TR and LY was routed outside the hillslope via a pipe system to tipping buckets that recorded the flow volume of SSF in 5 min time intervals. The pipe system had been flushed via access tubes the day before the experiments to guarantee function and cleanliness. Over the course of the sprinkling experiment (ca. 10 to 12 hours) the SSF of all TR and LY was sampled every 30 minutes into 100 ml brown glass bottles using automatic samplers (custom made by
410 the Chair of Hydrology, University of Freiburg). The sampling was continued for 12 hours after the end of the sprinkling with a sampling interval of 2 h. The water samples were transported in cooling boxes to the lab directly after the experiment for subsequent hydrochemical and stable isotope analysis.

To measure total phosphorus concentrations (P_{tot}), 50 ml of the sample was digested by adding 0.5 ml 4.5 M sulfuric acid (H₂SO₄) and processed in an autoclave. P_{tot} was analyzed by the molybdenum blue photometric method based on DIN EN
415 ISO 6878 (DIN, 2004) using a Unicam AquaMate photometer (Spectronic Unicam) with a 5 cm-cuvette at 700 nm. We determined the limit of quantitation for P_{tot} (0.026009 mg/l) and the limit of detection (0.043004 mg/l) based on DIN 32645 with a significance bound of 99 % for the limit of quantitation and 77 % for the limit of detection (DIN, 2008). The remaining 50 ml of each sample was filtered with a 0.45 µm cellulose filter (PERFECT-FLOW, WICOM) and used for analysis of ¹⁸O and ²H stable water isotopes using a Cavity Ring-Down L2130-i Isotopic Liquid Water Analyser (Picarro Inc.). Based on the
420 background isotopic signature and the isotopic signature of the sprinkling water, event and pre-event water fractions were calculated using a simple two endmember mass balance approach, also called two-component isotope hydrograph separation (Sklash and Farvolden, 1979).

2.3 Data Analysis

TR and LY volumetric measurements were scaled to 1 m² plot area and expressed as mm/h, which allows direct comparison
425 with the amount of incoming sprinkling water. We determined the time lag between the start of the sprinkling experiment and the time when 20 % of the total rise in SSF had been reached or exceeded (called t_{rise20} in SSF). In a similar way we determined the time lag between the start of the sprinkling event and the time when the event water fraction had reached 20 % (called t_{rise20} in event water fraction). The threshold of 20 % has been chosen as clear indication of first response in the change of SSF and event water fraction. If t_{rise20} in event water fraction is longer than t_{rise20} in SSF this indicates that the flow celerity is faster than
430 flow velocity. If t_{rise20} in event water fraction is shorter than ~~the~~ t_{rise20} in SSF, this indicates preferential flow as velocity is faster than celerity.

We plotted P concentrations of vertical and lateral SSF as a function of time but also as boxplots to allow a better comparison between the experiments. To investigate if the P concentrations of each flow component was higher during the first 2 h after onset of flow than during the remaining time of the experiments, we plotted our data separately for the two periods of each
435 event. As some of the flow components had only few samples in the flushing period a statistical test of significance was not meaningful. This was also the reason why we did not further differentiate between flushing period and remaining part for the following analysis.

We plotted P concentrations as a function of SSF in log-log space to investigate whether P mobilization was able to keep up with P transport. In this case, we would expect near chemostatic conditions, i.e., P concentrations would not vary much with increasing or decreasing SSF. In this case data points were expected to plot parallel to the x-axis. In case of simple dilution, P concentrations would decrease proportional with increasing SSF and the data points would be aligned on a 1:1 line. As a further measure we calculated the ratio between the range in observed SSF values and the respective range in observed P concentration. Under predominantly chemostatic conditions, the range in SSF would be much larger than the range in P concentrations.

In addition, we investigated the relation between P_{tot} concentrations and event water fractions and checked whether the slope of a linear regression based on the datapoints was significantly different from zero, using a two-sided t-test and an Alpha = 0.05. We also tested if this regression slope was significantly different from a slope describing dilution due to proportional mixing of pre-event water with our sprinkling water using a Mann-Whitney two-sided t-test. The slope describing simple dilution was determined by a regression with the best linear fit through the data points and the additional constrain of P_{tot} = 0 mg/l for event water fraction = 1.

3. Results

3.1 Lateral versus vertical SSF

In general, vertical SSF (measured by LY) dominated total water flow during all sprinkling experiments. Depending on site and soil depth, between 89 % and 99 % of total SSF percolated vertically to deeper depths and only < 1 to 11 % of total SSF was flowing laterally towards the trench (Fig. 3). At all study sites LY1B (below the forest floor) yielded steady flow with a mean rate of 10 to 15 mm/h, which is identical to the sprinkling rate. This confirms that the LY, even if positioned at the boundary of the hillslope experimental plot, were experiencing rainfall intensities that are representative for the rest of the hillslope plot. The LY at deeper soil depth showed a slower increase in SSF than LY1B, but also reached a mean flow rate of 10 to 15 mm/h towards the end of the sprinkling experiment (except LY3B and LY4B at MIT in spring and summer and LY3B at TUT in summer) (Fig. 3). Lateral SSF (measured by TR) was at least two orders of magnitude lower than vertical SSF but increased constantly towards the end of the sprinkling experiments. An exception was TR2B at TUT which reached a plateau (ca. 1 mm/h) during the experiment in spring, most probable due to wetter antecedent conditions. Maximum lateral SSF from TR2B at CON and TUT was 0.72 and 1.09 mm/h during the experiment in spring and 0.51 and 0.70 mm/h in summer. TR3B at MIT yielded neither vertical nor lateral SSF in any of the two sprinkling experiments. This is likely attributed to lower skeleton content and higher storage capacity of the soil in MIT.

3.2 Event and Pre-event Water Fractions

Vertical SSF in the topsoil, subsoil and saprolite (i.e., all LY except LY1B) at all sites and during all experiments was dominated by pre-event water (Fig. 3, Tab. 3). In contrast, the mean pre-event water fraction from the forest floor (i.e., LY1B)

470 during the sprinkling events was low (i.e., MIT: 12 % and 15 %; CON: 10 % and 16 % and TUT: 4 % and 4 % in spring and
summer, respectively). The mean pre-event water fractions of vertical SSF increased with depth at all sites and all events and
already in 35 to 40 cm soil depth (LY2B) the mean pre-event water fractions during the events in spring and summer were
83 % and 88 % at MIT, 60 % and 58 % at CON, and 64 % and 83 % at TUT. The mean pre-event water fraction in the vertical
SSF further continued to increase with soil depth (see LY3B and LY4B in Tab. 3).

475 For the lateral SSF a similar increase of the pre-event water fraction with soil depth was observed. However, the mean pre-
event water fractions of lateral SSF in the subsoil (i.e., TR2B) were typically smaller than the pre-event water fractions in
vertical SSF at similar depth (i.e., LY3B). The mean pre-event water fractions of TR2B during the events in spring and summer
were 44 % and 33 % at MIT, 63 % and 55 % at CON and 70 % and 47 % at TUT. The mean pre-event water fractions during
the events in spring and summer of LY3B were 95 % and 93 % at MIT, 63 % and 83 % at CON and 95 % and 70 % at TUT.
Mean pre-event water fractions of the lateral and vertical SSF in the saprolite (i.e., TR3B and LY4B) were similar (i.e., 78 %
480 and 74 % for TR3B and 86 % and 78 % for LY4B at CON during the experiment in spring and summer. (No flow from TR3B
at MIT and no TR3B and LY4B installed at TUT).

3.3 Differences in SSF and Response Time response time between Sspring and Ssummer

In general, SSF differed less between the sprinkling experiments in spring and in summer than between the three sites. This
was particularly true for the peak flows in SSF. An exception was TUT where peak flows for LY2B and LY3B were higher in
485 spring compared to summer which is likely due to pronounced differences in antecedent wetness between spring and summer
at TUT.

SSF between spring and summer differed mainly in respect to the response time of the individual TR and LY. This was most
pronounced at MIT and least at CON (Fig. 3, Tab. 4). At MIT the t_{rise20} in SSF was on average 1.9 times longer in summer
than in spring (except TR2B: 0.7 times shorter) (Tab. 4). At CON it was the opposite i.e., the t_{rise20} in SSF was on average 0.7
490 times shorter in summer than in spring; except for SSF in the saprolite (i.e., LY3B, LY4B, TR3B) for which t_{rise20} in SSF was
on average 1.3 times longer than in spring (Tab. 4). At TUT t_{rise20} in SSF in the top and subsoil was 1.8 times longer in summer
than in spring but t_{rise20} in SSF in the forest floor (LY1B and TR1B) was 0.3 times shorter. At CON and TUT the t_{rise20} in SSF
was not always increasing systematically with soil depth (e.g., LY3B responded earlier than LY2B at CON in spring and TUT
in summer; TR3B responded earlier than TR2B at CON in spring and summer). At MIT in summer LY4B did not respond
495 until 1.5 h after the end of the sprinkling experiment and TR3B did not respond at all in spring and summer.

The t_{rise20} of the event water fraction (Tab. 5), was typically longer than the t_{rise20} of SSF (Tab. 4). Only for TR2B and TR3B
at CON in spring and summer and TR2B at TUT in summer t_{rise20} of event water fraction was shorter than t_{rise20} of SSF. These
differences in t_{rise20} of event water fraction and t_{rise20} of SSF for these TR was in the range of 3 hours to almost 4 hours for
CON and 1 h for TUT. At MIT, SSF from LY3B and LY4B did not reach 20% event water fraction during the experiments in
500 summer and spring. The same was true for LY3B at TUT in spring.

3.4 Dynamics of P_e concentrations

Median P_{tot} concentrations of vertical SSF (LY) during the first flush (i.e., the first 1 to 2 hours after the first response of each flow component) were up to 6 times higher than during the remaining time of the sprinkling experiment (Fig. 5); some of the P_{tot} concentrations of the very first water samples were even up to one order of magnitude higher (Fig. 4a). This was particularly true for the vertical flow from the forest floor and the topsoil (LY1B and LY2B) but also apparent, albeit less distinct, for the subsoil and the saprolite (LY3B, LY4B).

The P_{tot} concentrations of vertical flow from the forest floor was significantly higher (Mann-Whitney Test, Alpha = 0.05) than in the topsoil (in 30 to 40 cm soil depth), except for CON and TUT in summer (Fig. 4a). Similarly, the P_{tot} concentrations of vertical flow in the subsoil and saprolite ~~were~~ significantly lower than in the topsoil except for MIT in spring and summer and at TUT in summer. All P_{tot} concentrations were above the limit of quantitation (i.e., 0.009 mg/l).

The P_{tot} concentrations in the same vertical flow component were not significantly different between spring and summer, except for LY1B and LY3B at MIT; LY1B at CON; LY1B, LY2B and LY3B at TUT. However, the P_{tot} concentrations of the same flow component (e.g., LY1B) at MIT was generally significantly higher than at CON and TUT, except for LY2B in spring. In a similar vein, P_{tot} concentration at CON were significantly higher than at TUT except LY3B in summer.

P_{tot} concentrations in the lateral flows (TR) also showed a sharp decline during the first 1 to 2 hours after the onset of flow of each component, except for TR1B and TR2B at MIT in spring and TR2B at CON in spring (Fig. 4b). P_{tot} concentrations of TR2B at CON in summer showed a steady decline and for TR3B the decline occurred with more delay (i.e., 5 h after first response) compared to the other experiments. At CON and TUT the P_{tot} concentrations in the lateral flow of the forest floor (TR1B) were significantly higher than in the subsoil (TR2B). Contrarily at MIT the P_{tot} concentrations in lateral flow of the forest floor (TR1B) were significantly lower than in the subsoil (TR2B), both in spring and summer (Fig. 4b). The difference in P_{tot} concentrations in the same flow component in spring and summer was not significant, except for TR1B at TUT.

3.4.1 P Concentration as a function of instantaneous flow

The range in SSF was typically several factors if not orders of magnitude larger than the range in P_{tot} concentrations during the sprinkling experiments except for those flow components that yielded little SSF in general (e.g., TR1B at CON during spring and summer, LY3B and TR2B at MIT during summer) (Fig. 6). Data points were in general aligned rather parallel to the x-axis in the log-log plots of P_{tot} concentration versus SSF and not along the 1:1 line (Fig. 6). This suggests that transport in most subsurface flow components was rather chemostatic than diluted. However, we observed weak anti-clockwise hysteresis effects i.e., median P_{tot} concentrations were higher on the rising limb than on the falling limb of the SSF hydrographs, except for TR1B at MIT in summer, LY4B at CON in spring and LY2B at TUT in summer. Most differences in P_{tot} concentrations during rising and falling limbs were, however, small (e.g., 0.007 mg/l, 0.028 mg/l and 0.081 mg/l for the 25%, 50% and 75% quantile of all differences) indicating that the hysteresis effect was small and only for LY1B this difference was consistently significant across all sites (Suppl. Tab. 1).

3.4.2 P Concentration as a Function of Event Water Fraction

P_{tot} concentrations also did not change significantly with increasing event water fraction (Fig. 7). Only in the forest floor (LY1B and TR1B), the slopes of the linear regression lines fitted to the P_{tot} concentrations as a function of event water fractions were significantly different from zero, except LY1B at CON in summer and TR1B at MIT in spring (Tab. 6). The transport conditions in the lateral and vertical SSF in the forest floor were therefore predominantly non-chemostatic. In the mineral soil and saprolite lateral and vertical SSF was dominantly chemostatic (except LY2B and LY4B at MIT in spring, LY3B at CON in summer, LY2B at CON in spring, LY2B and TR2B at TUT in summer and TR2B at TUT in spring) (Tab. 6). Most of these regression slopes were however close to zero.

For the regression slopes indicating non-chemostatic behavior, we tested if they were not significantly different from a slope describing proportional mixing. In general this was not the case, except for some flow components from the forest floor (i.e., LY1B at MIT and TUT in summer, TR1B at CON in spring; see italic values in Tab. 6) and from the mineral soil (i.e., TR2B at MIT in summer, TR2B at TUT in summer and spring and LY3B at CON in summer). At least for the latter three the regression slope was however close to zero (Tab. 6). As a rough generalization one could summarize: The regression slopes of most lateral and vertical SSF in the mineral soil and saprolite tended to be small or close to zero (the majority was chemostatic) whereas regression slopes of flow components from the forest floor were typically significantly different from zero (some indicating proportional mixing).

4. Discussion

4.1 Main SSF Flow Paths during long, moderately intense Sprinkling Events

In general, vertical SSF dominated total SSF during all sprinkling experiments and lateral SSF was at least two orders of magnitude lower than vertical SSF (Fig. 3). This finding implies that previous studies at trenched hillslopes at sites with well drained soils and moderately permeable bedrock missed out to quantify the important loss term of the water balance (e.g., Jackson et al., 2016; Sohr et al., 2018; Timmons et al., 1977). This is partly due to different research foci of these studies but mainly attributed to the technical difficulty to measure and sample vertical SSF. The use of our large zero tension lysimeters is a successful way to capture vertical flow in heterogeneous soils. They yield more representative results than traditional small size lysimeters of a few cm² or suction cups that are more likely to be affected by soil heterogeneity. The steel piling plates that we pressed into the undisturbed soil profile from the side of the hillslope also allow to preserve the natural soil profile above with its soil texture and structure and horizon-specific P stocks. However, their installation comes with high technical and man-power effort.

Vertical and lateral SSF in the mineral soil and saprolite at all sites and all experiments was predominantly pre-event water (Fig. 3, Tab. 3). This is generally in agreement with Jackson et al. (2016) who performed a tracer experiment at the Savannah River Site (South Carolina, USA) in a loamy sand topsoil overlaying a sandy clay-loam subsoil. However, their maximum pre-

565 event water fraction was 50 % while it was typically higher in our study (mean pre-event water fraction in vertical and lateral
SSF of the mineral soil was 83 % and 63 %, see also Tab. 3). Our findings suggest that SSF runoff generation was dominated
by initiation of water already stored in the soil, i.e., incoming event water ~~is~~was pushing pre-event water down into the soil
profile initiating SSF. A non-sequential onset of SSF with soil depth and shorter t_{rise20} of event water fraction than t_{rise20} of SSF
in some lateral flow components at CON and TUT suggest, however, that preferential flow occurred in parallel to matrix flow
at CON and TUT.

570 Occurrence of preferential flow is important as it allows P bypassing the soil matrix in its soluble and colloidal form and is
therefore considered to be a very prominent pathway of P loss from the ecosystem (Jardine et al., 1990; Kaiser et al., 2000;
Missong et al., 2018b). The fact that we observed indications of preferential flow predominantly at CON and TUT but not at
MIT and mainly in lateral flow and less in vertical flow may be explained by differences in soil properties, especially skeleton
content (Tab. 1) and soil bulk density (Fig. 1) of the three sites. Swelling and shrinking due to the higher clay content and more
575 biopores ~~due to~~as a result of higher earthworm abundance (due to higher soil pH) at TUT might be another mechanism leading
to preferential flowpaths. At MIT the lateral and vertical flow components in the saprolite did not respond ~~at all~~ or ~~responded~~
with strong delay to the sprinkling experiments (Fig. 3, Tab. 4) which suggests very efficient storage of the sprinkling water
in the soil. Therefore, at MIT characterized by relatively low skeleton content but high soil storage capacity, the shallow
flowpaths were most important and the deeper flowpaths did not yield much or any flow. At CON ~~it~~the opposite was the
580 ~~opposite~~case. The higher skeleton content allowed water to reach deeper soil depth and therefore deeper flowpaths yielded
more flow than shallower ones. At TUT the clay-rich topsoil led to more lateral flow at shallow depth than CON but as the
total soil depth was much smaller than at the two other sites, the storage capacity was less and therefore also the mineral soil
yielded significant amounts of flow.

The differences in SSF between the sprinkling experiment in spring and summer were smaller than the difference in SSF
585 among the three sites and mainly related to SSF response timing (Fig. 3, Tab. 4). The reason is likely due to relatively small
differences in antecedent soil moisture-~~antecedent~~ conditions between the two sprinkling experiments in summer and spring
(except TUT). This was particularly true at CON where the high skeleton content allow~~ed~~s the soils to drain quickly to field
capacity. Soil properties, such as drainable porosity, likely also explain~~r~~y seasonal differences in SSF response timing were
more pronounced at MIT than CON (Tab. 4, Tab. 5). At TUT the difference in SSF dynamics between the experiment in
590 summer and spring is likely more related to differences in antecedent wetness conditions. The 7-day-average of the median
volumetric water content of the soil profile at TUT during spring was 23 vol % compared to 15 vol % at TUT during summer.
Still, the dominance of pre-event water fractions during both events at TUT suggest that not fast preferential flow but piston
flow was the dominant process during both experiments at TUT.

4.2 P Concentration Dynamicsconcentration dynamics in vertical and lateral SSF

595 P_{tot} concentrations in vertical and lateral SSF, particularly in the forest floor and to a lesser degree in the mineral soil, were
typically significantly higher during the first 1 to 2 hours after the first flow response of each component than during the

remaining time of the sprinkling experiment (Fig. 4, Fig. 5). This so-called nutrient flushing effect (Hornberger et al., 1994) has been described also in other studies as a prominent feature of lateral export of nutrients; mainly for N (DON) and C (DOC) (Qualls and Haines, 1991; van Verseveld et al., 2008; Weiler and McDonnell, 2006) but also for P (DOP) (Burns et al., 1998; Missong et al., 2018a; Qualls et al., 2002; Sohr et al., 2018). Makowski et al. (2020), who measured P concentrations in vertical SSF, also reported nutrient-flushing in the first 2 hours of their sprinkling experiments. Various processes might contribute to this flushing effect. Drying and rewetting processes are known to induce nutrient flushing due to cell lysis due to osmotic stress and subsequent release of nutrients (Gordon et al., 2008; Turner et al., 2003) as well as due to disruption of soil aggregates allowing mineralization of former inaccessible organic substrates (Bünemann et al., 2013). Furthermore, there is evidence that with rewetting after intense drought events the microbial activity is increasing. Phosphate mobilization has been observed in forest soils as a consequence (Brödlin et al., 2019). In addition, drying of organic rich material leads to water repellence of the forest floor material and thus might enhance particle-bound nutrient export (Dincher and Calvaruso, 2020).

In our study P_{tot} concentrations in SSF from the forest floor continued to decline after the flushing phase or were relatively constant (Fig. 4). The slope of a regression line of P_{tot} concentration versus event water fraction of most flow components in the forest floor was significantly different from simple dilution that would be the case if a given amount of P was leached and diluted by an increasing amount of event water in SSF (Fig. 7, Tab. 6). A likely better explanation could be that the rate of P replenishment in the P-rich forest floor was not high enough to facilitate the high P concentrations measured at the beginning of the event. This is in line with Helfenstein et al. (2018) who found a negative relation between P concentration in the soil solution and the turnover rate based on isotopic exchange kinetic experiments on 217 soil samples collected worldwide. The P concentrations from the forest floor (except MIT during summer) were more constant towards the end of the experiment, when the P concentrations from the forest floor were low in general.

In the mineral soil the flushing effect was less distinct. The P concentrations in SSF were generally lower compared to the forest floor. The change in P_{tot} concentrations in the mineral soil were several times smaller than the change in SSF (Fig. 6) and the regression slope of P_{tot} concentration versus SSF was not significantly different from zero for most flow components in the mineral soil (Tab. 6). This suggests that P transport in SSF in the mineral soil was chemostatic. It also suggests that the rate of P replenishment in SSF of the mineral soil was high during the experiments. This was even true towards the end of the sprinkling experiments when the fraction of new water increased. For this new sprinkling water, it is clear that the contact time was short (i.e., in the order of minutes to hours) and still the P concentration was not significantly lower than for water samples with a higher pre-event water fraction. The relatively constant P concentrations in the mineral soil also suggest, that P leached from the forest floor was efficiently buffered in the mineral soil and that this buffering effect is also apparent in soils with preferential flow paths (CON and TUT) that we considered as a potential bypass of the soil matrix. We could not see significant differences between the experiments in spring and summer at the three sites. An exception were P concentrations of vertical flow from the forest floor at all sites. Our findings are therefore in line with Kaiser et al. (2003) and Sohr et al. (2019) who also reported higher P concentrations of leachate from the forest floor during the vegetation period. Sohr et al. (2019) also reported an increase in P concentrations in throughfall towards the end of the vegetation period but the average P concentration

was 0.15 mg/l and therefore more than 4 times smaller compared to the enrichment caused by leaching from the forest floor (average 0.7 mg/l). An exception were the P concentrations of vertical flow from the forest floor at all sites.

4.3 Estimates of annual P fluxes

Our finding If one would argue that P-transport in the amount of SSF after during the first flush was chemostatic would suggest flushing period is small compared to the remaining part of the event (see Fig. 3) one could also assume that chemotactic conditions prevail during large rainfall events. Given these simplifications, one can roughly estimate annual P losses from forest stands could be roughly estimated by knowing the amount of annual SSF and an average P concentration. This would be particularly relevant for long-term soil evolution modeling (Chadwick et al., 1999; Vitousek, 2004). While we lack the data to fully test this hypothesis, we performed a rough estimation of We performed this hypothetical test based on the stated assumptions and estimated the annual P fluxes in SSF at CON, which is the only site where we have SSF data for more than 1 year. We chose the period between 10.05.2018 to 09.05.2019 because the water balance is in order to not affected by include the sprinkling experiments in the water balance. The sum of SSF and the calculated P_{tot} flux for this period is shown in Tab. 7. The mean P concentration used to calculate the annual P flux is based on the P concentrations measured after the first flush during the experiment at CON in spring 2018 (i.e., 07.05.2018). This is the sprinkling experiment most recent to the one-year period of interest. Our estimation of annual P_{tot} fluxes (sum of vertical and lateral) yielded 45 mg/m²/a from the forest floor and 14 mg/m²/a from the topsoil, 6 mg/m²/a from the subsoil and 2 mg/m²/a from the saprolite. Considering only the outside boundaries of the hillslope (LY4B, TR1B, TR2B, TR3B) we estimated a total P-loss of 3.162 mg/m²/a. While we expect these values to be different for every year, the order of magnitude can be compared to values presented by Sohrt et al. (2019) who have sampled P concentrations in lateral SSF at CON on a bi-weekly interval between 01.03.2015 and 25.02.2016. Their mean P concentration from the organic layer (only lateral flow) was 0.57 mg/l, the annual water flux from the organic layer was 0.002 mm/a and the calculated annual P flux was 0.001 mg/m²/a. For the mineral soil they reported a mean P concentration of 0.043 mg/l, a total annual water flux of 446 mm/a and a resulting P flux of 20 mg/m²/a. While the annual flux from the mineral soil matches well between Sohrt et al. (2019) and our study, the annual P flux from the forest floor differs by several orders of magnitude. This is likely due to differences in the estimated water fluxes (Tab. 7). Sohrt et al. (2019) estimated the annual water flux by End Member Mixing Analysis (EMMA) based on samples from the lateral flow components (TR1B) only. Our annual water flux from the forest floor comprises both, the lateral (TR1B) and the vertical (LY1B) flow component of which the latter has been shown to be orders of magnitude larger than the lateral flow (see Fig. 3). For that reason it is likely better to compare the estimated P fluxes from the outside boundaries of the hillslope in our study (3.162 mg/m²/a) to the P fluxes in the groundwater (2.5 mg/m²/a) in the study of Sohrt et al. (2019). These two P fluxes represent the P loss from the entire hillslope. These two fluxes match well.

Our annual P fluxes are also in line with the range of fluxes published in a number of other studies or reviews. Bol et al. (2016) estimated that potential total P loss through leaching into subsoils and export by forest streams was less than 50 mg P/m²/a. But the range of published data varies more than one order of magnitude (e.g., 4 mg P_{tot}/m²/a in Benning et al. (2012); 32

mg/m²/a in Julich personal communication, cit. in Bol et al. (2016)). Other studies reported only annual fluxes of Dissolved Organic Phosphorus (DOP). While these numbers are difficult to be directly compared to our P_{tot} data, the published values are in a similar order of magnitude (e.g.: 15 to 62 mg DOP/m²/a in leachate from the organic layer and 1.7 to 38 mg DOP/m²/a from the mineral soil (Kaiser et al., 2000, 2003; Qualls et al., 2000). In another meta-analysis Sohrt et al. (2017) presented annual P fluxes that are higher: (e.g., mean: 226 mg/m²/a (sd: 389 mg/m²/a) and mean: 119 mg/m²/a (sd: 279 mg/m²/a) for water samples from the organic layers and the mineral soils). A possible reason for the higher values compared to our results could be that we used mean P concentration measured after the first flush which results in rather conservative estimates. Julich (personal communication cit. in Boll et al. (2016)) reported that up to 40 % of the annual P flux might occur during single events, which would suggest that the fluxes during the first flush have an important share on the annual flux and cannot be neglected. Further research is needed to clarify the role of P-flushing on the total P loss from forest ecosystems. Given the high P concentrations during the flushing phase of our sprinkling events it might be that relatively small amounts of SSF during the first flush can still export a significant amount of P and might be highly relevant especially at sites with low P stocks and/or P availability (e.g., acid sites developed from P poor parent material (CON) or sites developed on carbonate rock (TUT)). The forest floor at such sites serves as most important P source and efficient P recycling from the forest floor is the prerequisite for a sufficient P-supply to the ecosystems (Hauenstein et al. 2018, Lang et al. 2017). In the mineral soil P is fixed to the solid phase by sorption to sesquioxides or precipitation of Ca-phosphates and therefore P availability is very low.

Despite the considerable range of published annual P fluxes in SSF from different forest environments it becomes clear that annual P fluxes are orders of magnitude lower than soil P stocks (Achat et al., 2016; Hou et al., 2018). In the case of CON soil P stocks are estimated to be 13 g/m² and 230 g/m² for the forest floor and the mineral soil, respectively (Lang et al., 2017). Therefore, export of P from forest ecosystems by SSF becomes only relevant on the time-scale of millennia (Bol et al., 2016; Vitousek, 2004). The annual P losses are also well compensated by atmospheric deposition and mineral weathering (Aciego et al., 2017; Hartmann et al., 2014; Tipping et al., 2014). For instance, at CON the annual P fluxes by dry deposition, canopy leaching, bulk throughfall and mineral weathering are 10, 43, 60 and 76 mg/m²/a (Sohrt et al., 2017; Uhlig and von Blanckenburg, 2019) and therefore in a similar order as the estimated annual P fluxes by SSF in our study. It is also important to note that the internal ecosystem P fluxes due to forest litter fall (100 to 500 mg/m²/a, Sohrt et al. (2017)) are up to one order of magnitude larger than the annual P fluxes in SSF. This is an indication of an efficient nutrient recycling in forest ecosystems (Lang et al., 2017).

5. Conclusions

We present results of sprinkling experiments conducted at 200 m² hillslopes at three beech forest sites in Germany that differ in their soil depth, skeleton content and soil P stocks, to quantify the dynamic of vertical and lateral SSF and associated P concentrations. Vertical SSF in the mineral soil and saprolite was at least two orders of magnitude larger than lateral SSF and consisted mainly of pre-event water that was likely replaced by sprinkling water (piston flow mechanism). This suggests that

previous studies that only measured lateral flow have likely missed the major hydrological export flux from forest stands. Median P_{tot} concentrations in SSF from the forest floor, and to a lesser extent also at deeper soil depths, at all sites were up to 6 times higher in the first 1 to 2 h after the first response of each flow component (first flush). For the remaining time of the experiments (ca. 10 h), transport conditions in the mineral soil and saprolite were however close to chemostatic. This suggests that the rate of P replenishment at all three sites was high (in the order of minutes to hours). Our findings that chemostatic transport conditions were prevalent in the mineral soil for most of the experiment even for very high rainfall amounts, duration of the experiments would suggest that annual P flux from forest stands could be approximated by simply knowing the average P concentration of the soil solution and the water balance of the site. A ~~quick~~ test of this assumption in the form of a rough approximation of the annual P flux at one of our sites (CON) yielded comparable results to an earlier study at the same site (Sohr et al., 2019) and some other sites (Benning et al., 2012; Bol et al., 2016). In terms of a P budget, our approximated annual P fluxes in SSF at CON were in a similar order ~~than~~ P inputs by dry and wet deposition and mineral weathering and several orders of magnitude smaller than the total P stocks of the mineral soil. This suggest that the annual P fluxes are only relevant on long time scale and that forest ecosystem appear to be relative efficient in retaining P in their nutrient cycle. Despite the fact that P losses from the ecosystem seem to be small, the translocation of P from the forest floor to the mineral soil might be of high relevance at sites where the forest floor is the dominant source for the P nutrition of trees.

6. Data availability

Data is available under: <https://1drv.ms/u/s!AhSOi7EfJ5qmfyJQrT3O65OLBuE?e=BrwfFp>

7. Author contribution

MW wrote the grant; MR was responsible for the experimental design and field installations together with two technicians. MR planned and organized the sprinkling experiments and lab-analysis together with two technicians. MR was responsible for data pre-processing and all analysis. JK and FL provided the data on soil characterization and valuable thoughts and discussion on the soil ecological aspect of the study. HP provided meteorological data, supported the long-term monitoring of SSF at CON and provided valuable feedback on the paper manuscript. MR wrote the manuscript including all figures and tables and all other co-authors discussed the results, provided valuable feedback on the text and figures.

8. Competing interests:

The authors declare that they have no conflict of interest.

9. Acknowledgments

This project was carried out in the framework of the priority program SPP 1685 “Ecosystem Nutrition: Forest Strategies for limited Phosphorus Resources” funded by DFG (grant No. WE 4598/7-2). The article processing charge was funded by the
725 Baden-Wuerttemberg Ministry of Science, Research and Art and the University of Freiburg.

Special thanks to Jakob Sohr for parts of the field installations (TR) and very insightful discussion and helpful suggestions during the project. Many thanks to Emil Blattmann, Dominc Demand, Sandra Fritzsche, Maja Gensow, Nina Gottselig, Benjamin Grahler, Janine Heizmann, Barbara Herbstritt, Johannes Herre, Jakub Jeřábek, Britta Kattenstroth, Lara Kirn, Petra Kürfner, Hannes Leister, Marvin Lorff, Anna Missong, Regine Nitz, Felix Nyarko, Benjamin Schima, Jonas Schwarz, Stefan
730 Seeger, Jürgen Strub, Angela Thiemann, Delon Wagner, Monika Wirth-Lederer for support in the field, lab and admin. We also want to thank the [Team-team of the FVA](#): Hermann Buberl, Andrea Hölscher, Martin Maier, Melanie Strecker, Philine Raubold, Lisa Rubin for support during sampling at CON and lab analysis; Stephan Raspe (LWF Bayern) for rainfall samples and meteorological data of MIT. Thanks also to our local partners at the three sites: Hans Amman, Siegfried Fink, Harry Peintinger, Paul Steinhard, Fabian Reichel, Hubert Waldmann, Othmar Winterhalder.

735 [We also like to thank the Editor and three anonymous reviewers for their useful comments.](#)

10. References

Achat, D. L., Pousse, N., Nicolas, M., Brédoire, F. and Augusto, L.: Soil properties controlling inorganic phosphorus availability: general results from a national forest network and a global compilation of the literature, *Biogeochemistry*, 127(2–3), 255–272, doi:10.1007/s10533-015-0178-0, 2016.

740 Aciego, S. M., Riebe, C. S., Hart, S. C., Blakowski, M. A., Carey, C. J., Aarons, S. M., Dove, N. C., Botthoff, J. K., Sims, K. W. W. and Aronson, E. L.: Dust outpaces bedrock in nutrient supply to montane forest ecosystems, *Nat. Commun.*, 8, doi:10.1038/ncomms14800, 2017.

Backnäs, S., Laine-Kaulio, H. and Kløve, B.: Phosphorus forms and related soil chemistry in preferential flowpaths and the soil matrix of a forested podzolic till soil profile, *Geoderma*, 189–190, 50–64, doi:10.1016/j.geoderma.2012.04.016, 2012.

745 Benning, R., Schua, K., Schwärzel, K. and Feger, K. H.: Fluxes of Nitrogen, Phosphorus, and Dissolved Organic Carbon in the inflow of the Lehmühle reservoir (Saxony) as compared to streams draining three main land-use types in the catchment, *Adv. Geosci.*, 32, 1–7, doi:10.5194/adgeo-32-1-2012, 2012.

Bol, R., Julich, D., Brödlin, D., Siemens, J., Kaiser, K., Dippold, M. A., Spielvogel, S., Zilla, T., Mewes, D., von Blanckenburg, F., Puhlmann, H., Holzmann, S., Weiler, M., Amelung, W., Lang, F., Kuzyakov, Y., Feger, K., Gottselig, N., Klumpp, E.,

750 Missong, A., Winkelmann, C., Uhlig, D., Sohr, J., von Wilpert, K., Wu, B. and Hagedorn, F.: Dissolved and colloidal phosphorus fluxes in forest ecosystems — an almost blind spot in ecosystem research, *Plant Nutr. Soil Sci.*, 179, 425–438, doi:doi.org/10.1002/jpln.201600079, 2016.

Braun, S., Thomas, V. F. D., Quiring, R. and Flückiger, W.: Does nitrogen deposition increase forest production? The role of

- phosphorus, *Environ. Pollut.*, 158(6), 2043–2052, doi:10.1016/j.envpol.2009.11.030, 2010.
- 755 Brinson, M. M.: Decomposition and Nutrient Exchange of Litter in an Alluvial Swamp Forest, *Ecology*, 58(3), 601–609, 1977.
- [Brödlin, D., Kaiser, K., Kessler, A. and Hagedorn, F.: Drying and rewetting foster phosphorus depletion of forest soils, *Soil Biol. Biochem.*, 128\(May 2018\), 22–34, doi:10.1016/j.soilbio.2018.10.001, 2019.](#)
- [Bünemann, E. K., Keller, B., Hoop, D. and Jud, K.: Increased availability of phosphorus after drying and rewetting of a grassland soil : processes and plant use, *Plant Soil*, 370, 511–526, doi:10.1007/s11104-013-1651-y, 2013.](#)
- 760 Burns, D. A., Hooper, R. P., McDonnell, J. J., Freer, J. E., Kendall, C. and Beven, K.: Base cation concentrations in subsurface flow from a forested hillslope: The role of flushing frequency, *Water Resour.*, 34(12), 3535–3544, 1998.
- Carlyle, G. C. and Hill, A. R.: Groundwater phosphate dynamics in a river riparian zone: Effects of hydrologic flowpaths, lithology and redox chemistry, *J. Hydrol.*, 247(3–4), 151–168, doi:10.1016/S0022-1694(01)00375-4, 2001.
- Chadwick, O. A., Derry, L. A., Vitousek, P. M., Huebert, B. J. and Hedin, L. O.: Changing sources of nutrients during four million years of ecosystem development, *Nature*, 397(6719), 491–497, doi:10.1038/17276, 1999.
- 765 Cole, D. and Rapp, M.: Elemental Cycling in Forest Ecosystems, in *Dynamic Properties of Forest Ecosystems*, edited by D. Reichle, pp. 341–409, Cambridge University Press., 1981.
- Compton, J. E. and Cole, D. W.: Phosphorus cycling and soil P fractions in Douglas-fir and red alder stands, *For. Ecol. Manage.*, 110(1–3), 101–112, doi:10.1016/S0378-1127(98)00278-3, 1998.
- 770 Cox, J. W., Kirkby, C. A., Chittleborough, D. J., Smythe, L. J. and Fleming, N. K.: Mobility of phosphorus through intact soil cores collected from the Adelaide Hills, South Australia, *Aust. J. Soil Res.*, 38(5), 973–990, doi:10.1071/SR99125, 2000.
- Diaconu, D., Kahle, H. P. and Spiecker, H.: Thinning increases drought tolerance of European beech: a case study on two forested slopes on opposite sides of a valley, *Eur. J. For. Res.*, 136(2), 319–328, doi:10.1007/s10342-017-1033-8, 2017.
- DIN: DIN ES ISO 6878 - Water quality determination for phosphorus - Ammonium molybdata spectrometric method , 1–30, 775 2004.
- DIN: DIN 32645: Chemical analysis - Decision limit, detection limit and determination limit under repeatability conditions - Terms, methods, evaluation, , 1–20, 2008.
- [Dincher, M. and Calvaruso, C.: Particles in humus leaching solution influence the input – output budget of the major elements in a beech forest, *Biogeochemistry*, 151\(1\), 1–14, doi:10.1007/s10533-020-00702-6, 2020.](#)
- 780 Duquesnay, A., Dupouey, J. L., Clement, A., Ulrich, E. and Le Tacon, F.: Spatial and temporal variability of foliar mineral concentration in beech (*Fagus sylvatica*) stands in northeastern France, *Tree Physiol.*, 20(1), 13–22, doi:10.1093/treephys/20.1.13, 2000.
- DWD Climate Data Center: Raster der Wiederkehrintervalle für Starkregen(Bemessungsniederschläge) in Deutschland (KOSTRA-DWD), Version 2010R, 2010.
- 785 Elser, J. J., Fagan, W. F., Denno, R. F., Dobberfuhl, D. R., Folarin, A., Huberty, A., Interlandi, S., Kilham, S. S., McCauley, E., Schulz, K. L., Siemann, E. H. and Sterner, R. W.: Nutritional constraints in terrestrial and freshwater food webs, *Nature*, 408(6812), 578–580, doi:10.1038/35046058, 2000.

- Elser, J. J., Bracken, M. E. S., Cleland, E. E., Gruner, D. S., Harpole, W. S., Hillebrand, H., Ngai, J. T., Seabloom, E. W., Shurin, J. B. and Smith, J. E.: Global analysis of nitrogen and phosphorus limitation of primary producers in freshwater, marine and terrestrial ecosystems, *Ecol. Lett.*, 10(12), 1135–1142, doi:10.1111/j.1461-0248.2007.01113.x, 2007.
- 790 Fuchs, J. W., Fox, G. A., Storm, D. E., Penn, C. J. and Brown, G. O.: Subsurface transport of phosphorus in riparian floodplains: Influence of preferential flow paths, *J. Environ. Qual.*, 38(2), 473–484, doi:10.2134/jeq2008.0201, 2009.
- [Gordon, H., Haygarth, P. M. and Bardgett, R. D.: Drying and rewetting effects on soil microbial community composition and nutrient leaching, *Soil Biol. Biochem.*, 40, 302–311, doi:10.1016/j.soilbio.2007.08.008, 2008.](#)
- 795 Gottselig, N., Amelung, W., Kirchner, J. W., Bol, R., Eugster, W., Granger, S. J., Hernández-Crespo, C., Herrmann, F., Keizer, J. J., Korkiakoski, M., Laudon, H., Lehner, I., Löfgren, S., Lohila, A., Macleod, C. J. A., Mölder, M., Müller, C., Nasta, P., Nischwitz, V., Paul-Limoges, E., Pierret, M. C., Pilegaard, K., Romano, N., Sebastia, M. T., Stähli, M., Voltz, M., Vereecken, H., Siemens, J. and Klumpp, E.: Elemental Composition of Natural Nanoparticles and Fine Colloids in European Forest Stream Waters and Their Role as Phosphorus Carriers, *Global Biogeochem. Cycles*, 31(10), 1592–1607, doi:10.1002/2017GB005657,
- 800 2017.
- Gregory, S. V.: Phosphorus dynamics on organic and inorganic substrates in streams, *Int. Vereinigung der Angew. und Theor. Limnol.*, 20, 1340–1346, 1978.
- Hagedorn, F. and Bundt, M.: The age of preferential flow paths, *Geoderma*, 108(1–2), 119–132, doi:10.1016/S0016-7061(02)00129-5, 2002.
- 805 Hartmann, J., Moosdorf, N., Lauerwald, R., Hinderer, M. and West, A. J.: Global chemical weathering and associated P-release - The role of lithology, temperature and soil properties, *Chem. Geol.*, 363, 145–163, doi:10.1016/j.chemgeo.2013.10.025, 2014.
- Heathwaite, A. and Dils, R.: Characterising phosphorus loss in surface and subsurface hydrological pathways, *Sci. Total Environ.*, 251–252, 523–538, doi:10.1016/S0048-9697(00)00393-4, 2000.
- 810 Helfenstein, J., Jegminat, J., McLaren, T. I. and Frossard, E.: Soil solution phosphorus turnover: Derivation, interpretation, and insights from a global compilation of isotope exchange kinetic studies, *Biogeosciences*, 15(1), 105–114, doi:10.5194/bg-15-105-2018, 2018.
- Hill, B. H., McCormick, F. H., Harvey, B. C., Johnson, S. L., Warren, M. L. and Elonen, C. M.: Microbial enzyme activity, nutrient uptake and nutrient limitation in forested streams, *Freshw. Biol.*, 55(5), 1005–1019, doi:10.1111/j.1365-2427.2009.02337.x, 2010.
- 815 Hornberger, G. M., Bencala, K. E. and McKnight, D. M.: Hydrological controls on dissolved organic carbon during snowmelt in the Snake River near Montezuma, Colorado, *Biogeochemistry*, 25(3), 147–165, doi:10.1007/BF00024390, 1994.
- Hou, E., Tan, X., Heenan, M. and Wen, D.: A global dataset of plant available and unavailable phosphorus in natural soils derived by hedley method, *Sci. Data*, 5, 1–13, doi:10.1038/sdata.2018.166, 2018.
- 820 Ilg, K., Wellbrock, N. and Lux, W.: Phosphorus supply and cycling at long-term forest monitoring sites in Germany, *Eur. J. For. Res.*, 128(5), 483–492, doi:10.1007/s10342-009-0297-z, 2009.

- Jackson, C. R., Du, E., Klaus, J., Griffiths, N. A., Bitew, M. and McDonnell, J. J.: Interactions among hydraulic conductivity distributions, subsurface topography, and transport thresholds revealed by a multitracer hillslope irrigation experiment, *Water Resour. Res.*, 52(8), 6186–6206, doi:10.1002/2015WR018364, 2016.
- 825 Jardine, P. M., Wilson, G. V., McCarthy, J. F., Luxmoore, R. J., Taylor, D. L. and Zelazny, L. W.: Hydrogeochemical processes controlling the transport of dissolved organic carbon through a forested hillslope, *J. Contam. Hydrol.*, 6(1), 3–19, doi:10.1016/0169-7722(90)90008-5, 1990.
- Johnson, D. W., Trettin, C. C. and Todd, D. E.: Changes in forest floor and soil nutrients in a mixed oak forest 33 years after stem only and whole-tree harvest, *For. Ecol. Manage.*, 361, 56–68, doi:10.1016/j.foreco.2015.11.012, 2016.
- 830 Jonard, M., Fürst, A., Verstraeten, A., Thimonier, A., Timmermann, V., Potočić, N., Waldner, P., Benham, S., Hansen, K., Merilä, P., Ponette, Q., de la Cruz, A. C., Roskams, P., Nicolas, M., Croisé, L., Ingerslev, M., Matteucci, G., Decinti, B., Bascietto, M. and Rautio, P.: Tree mineral nutrition is deteriorating in Europe, *Glob. Chang. Biol.*, 21(1), 418–430, doi:10.1111/gcb.12657, 2015.
- Julich, D., Julich, S. and Feger, K. H.: Phosphorus fractions in preferential flow pathways and soil matrix in hillslope soils in the Thuringian Forest (Central Germany), *Zeitschrift für Pflanzenernährung und Bodenkd.*, 180(3), 407–417, doi:10.1002/jpln.201600305, 2017a.
- 835 Julich, D., Julich, S. and Feger, K. H.: Phosphorus in preferential flow pathways of forest soils in Germany, *Forests*, 8(1), doi:10.3390/f8010019, 2017b.
- Kabeya, D., Inagaki, Y., Noguchi, K. and Han, Q.: Growth rate reduction causes a decline in the annual incremental trunk growth in masting *Fagus crenata* trees, , 1444–1452, doi:10.1093/treephys/tpx081, 2017.
- 840 Kaiser, K., Guggenberger, G. and Zech, W.: Organically Bound Nutrients in Dissolved Organic Matter Fractions in Seepage and Pore Water of Weakly Developed Forest Soils, *Acta Hydrochim. Hydrobiol.*, 28(7), 411–419, doi:10.1002/1521-401x(20017)28:7<411::aid-aheh411>3.0.co;2-d, 2000.
- Kaiser, K., Guggenberger, G. and Haumaier, L.: Organic phosphorus in soil water under a European beech (*Fagus sylvatica* L.) stand in northeastern Bavaria, Germany: Seasonal variability and changes with soil depth, *Biogeochemistry*, 66(3), 287–310, doi:10.1023/B:BiOG.0000005325.86131.5f, 2003.
- 845 Kirschbaum, M. U. F.: The temperature dependence of soil organic matter decomposition, and the effect of global warming on soil organic C storage, *Soil Biol. Biochem.*, 27(6), 753–760, doi:10.1016/0038-0717(94)00242-S, 1995.
- Kunimatsu, T., Hamabata, E., Sudo, M. and Hida, Y.: Comparison of nutrient budgets between three forested mountain watersheds on granite bedrock, *Water Sci. Technol.*, 44(7), 129–140, doi:10.2166/wst.2001.0406, 2001.
- 850 Lambers, H., Raven, J. A., Shaver, G. R. and Smith, S. E.: Plant nutrient-acquisition strategies change with soil age, *Trends Ecol. Evol.*, 23(2), 95–103, doi:10.1016/j.tree.2007.10.008, 2008.
- Lang, F., Krüger, J., Amelung, W., Willbold, S., Frossard, E., Bünemann, E. K., Bauhus, J., Nitschke, R., Kandeler, E., Marhan, S., Schulz, S., Bergkemper, F., Schloter, M., Luster, J., Guggisberg, F., Kaiser, K., Mikutta, R., Guggenberger, G., Polle, A.,
- 855 Pena, R., Prietzel, J., Rodionov, A., Talkner, U., Meessenburg, H., von Wilpert, K., Hölscher, A., Dietrich, H. P. and Chmara,

- L.: Soil phosphorus supply controls P nutrition strategies of beech forest ecosystems in Central Europe, *Biogeochemistry*, 136(1), 5–29, doi:10.1007/s10533-017-0375-0, 2017.
- Makowski, V., Julich, S., Feger, K., Breuer, L. and Julich, D.: Leaching of dissolved and particulate phosphorus via preferential flow pathways in a forest soil : An approach using zero-tension lysimeters Leaching of dissolved and particulate phosphorus
860 via preferential flow pathways in a forest soil : An approach usi, *J. Plant Nutr. Soil Sci.*, doi:10.1002/jpln.201900216, 2020.
- Missong, A., Holzmann, S., Bol, R., Nischwitz, V., Puhlmann, H., v. Wilpert, K., Siemens, J. and Klumpp, E.: Leaching of natural colloids from forest topsoils and their relevance for phosphorus mobility, *Sci. Total Environ.*, 634, 305–315, doi:10.1016/j.scitotenv.2018.03.265, 2018a.
- Missong, A., Bol, R., Nischwitz, V., Krüger, J., Lang, F., Siemens, J. and Klumpp, E.: Phosphorus in water dispersible-colloids
865 of forest soil profiles, *Plant Soil*, 427(1–2), 71–86, doi:10.1007/s11104-017-3430-7, 2018b.
- Mulholland, P. J. and Hill, W. R.: Seasonal patterns in streamwater nutrient and dissolved organic carbon concentrations: Separating catchment flow path and in-stream effects, *Water Resour. Res.*, 33(6), 1297–1306, doi:10.1029/97WR00490, 1997.
- Pypers, P., Delrue, J., Diels, J., Smolders, E. and Merckx, R.: Phosphorus intensity determines short-term P uptake by pigeon pea (*Cajanus cajan* L.) grown in soils with differing P buffering capacity, *Plant Soil*, 284(1–2), 217–227, doi:10.1007/s11104-
870 006-0051-y, 2006.
- Qualls, G. and Haines, L.: Geochemistry of Dissolved Organic Nutrients in Water Percolating through a Forest Ecosystem, *Soil Sci. Soc. Am. JournalSci*, 1112–1123, 1991.
- Qualls, R. G., Haines, B. L., Swank, W. T. and Tyler, S. W.: Soluble Organic and Inorganic Nutrient Fluxes in Clearcut and Mature Deciduous Forests, *Soil Sci. Soc. Am. J.*, 64(3), 1068–1077, doi:10.2136/sssaj2000.6431068x, 2000.
- 875 Qualls, R. G., Haines, B. L., Swank, W. T. and Tyler, S. W.: Retention of soluble organic nutrients by a forested ecosystem, *Biogeochemistry*, 61(2), 135–171, doi:10.1023/A:1020239112586, 2002.
- Schindler, D. and Nighswander, J.: Nutrient supply and primary production in Clear Lake, Eastern Ontario, *J. Fish. Res. Board Canada*, 27(11), 2009–2036, doi:doi.org/10.1139/f70-226, 1970.
- Sklash, M. and Farvolden, R.: The role of groundwater in storm runoff, *J. Hydrol.*, 43, 45–65, 1979.
- 880 Sohr, J., Lang, F. and Weiler, M.: Quantifying components of the phosphorus cycle in temperate forests, *Wiley Interdiscip. Rev. Water*, 4(6), e1243, doi:10.1002/wat2.1243, 2017.
- Sohr, J., Puhlmann, H. and Weiler, M.: Phosphorus transport in lateral subsurface flow at forested hillslopes, *SOIL Discuss.*, (June), 1–20, doi:10.5194/soil-2018-13, 2018.
- Sohr, J., Uhlig, D., Kaiser, K., von Blanckenburg, F., Siemens, J., Seeger, S., Frick, D. A., Krüger, J., Lang, F. and Weiler, M.: Phosphorus Fluxes in a Temperate Forested Watershed: Canopy Leaching, Runoff Sources, and In-Stream Transformation,
885 *Front. For. Glob. Chang.*, 2(December), 1–14, doi:10.3389/ffgc.2019.00085, 2019.
- Steege, A., Govers, G., Takken, I., Nachtergaele, J., Poesen, J. and Merckx, R.: Factors controlling sediment and phosphorus export from two Belgian agricultural catchments, *J. Environ. Qual.*, 30(4), 1249–1258, doi:10.2134/jeq2001.3041249x, 2001.
- Stelzer, R. S., Heffernan, J. and Likens, G. E.: The influence of dissolved nutrients and particulate organic matter quality on

- 890 microbial respiration and biomass in a forest stream, *Freshw. Biol.*, 48(11), 1925–1937, doi:10.1046/j.1365-2427.2003.01141.x, 2003.
- Taylor, A., Edwards, E. and Simpson, E.: Nutrients in Streams Draining Woodland and Farmland near Coshocton, Ohio, *Water Resour. Res.*, 7, 81–89, doi:doi.org/10.1029/WR007i001p00081, 1971.
- Timmons, D. R., Verry, S., Burwell, R. E. and Holtz, R. F.: Nutrient Transport in Surface Runoff and Interflow from an Aspen-
895 Birch Fores, , 6(2), 1977.
- Tipping, E., Benham, S., Boyle, J. F., Crow, P., Davies, J., Fischer, U., Guyatt, H., Helliwell, R., Jackson-Blake, L., Lawlor, A. J., Monteith, D. T., Rowe, E. C. and Toberman, H.: Atmospheric deposition of phosphorus to land and freshwater, *Environ. Sci. Process. Impacts*, 16(7), 1608–1617, doi:10.1039/c3em00641g, 2014.
- [Turner, B. L., Driessen, J. P., Haygarth, P. M. and Mckelvie, I. D.: Potential contribution of lysed bacterial cells to phosphorus
900 solubilisation in two rewetted Australian pasture soils. *Soil Biol. Biochem.*, 35, 187–189, 2003.](#)
- Uhlig, D. and von Blanckenburg, F.: How slow rock weathering balances nutrient loss during fast forest floor turnover in montane, temperate forest ecosystems, *Front. Earth Sci.*, 7(July), doi:10.3389/feart.2019.00159, 2019.
- Vanek, V.: Transport of groundwater-borne phosphorus to Lake Bysjön, South Sweden, *Hydrobiologia*, 251(1–3), 211–216, doi:10.1007/BF00007180, 1993.
- 905 van Verseveld, W. J., McDonnell, J. J. and Lajtha, K.: A mechanistic assessment of nutrient flushing at the catchment scale, *J. Hydrol.*, 358(3–4), 268–287, doi:10.1016/j.jhydrol.2008.06.009, 2008.
- Vitousek, P.: Patterns and Processes in long-term ecosystem development, in *Nutrient Cycling and Limitation Hawaii as a Model System*, pp. 42–66, Princeton University Press, Princeton., 2004.
- Weiler, M. and McDonnell, J.: Testing nutrient flushing hypotheses at the hillslope scale: A virtual experiment approach, *J.
910 Hydrol.*, 319(1–4), 339–356, doi:10.1016/j.jhydrol.2005.06.040, 2006.
- White, R. E. and Beckett, P. H. T.: Studies on the phosphate potentials of soils - Part I - The measurement of phosphate potential, *Plant Soil*, 20(1), 1–16, doi:10.1007/BF01378093, 1964.
- Zhang, Z., Fukushima, T., Shi, P., Tao, F., Onda, Y., Gomi, T., Mizugaki, S., Asano, Y., Kosugi, K., Hiramatsu, S., Kitahara, H., Kuraji, K., Terajima, T. and Matsushige, K.: Baseflow concentrations of nitrogen and phosphorus in forested headwaters
915 in Japan, *Sci. Total Environ.*, 402(1), 113–122, doi:10.1016/j.scitotenv.2008.04.045, 2008.

Tab. 1: Characteristics of the three experimental sites Mitterfels (MIT), Conventwald (CON) and Tuttingen (TUT). ^a classified based on WRB, 2015, ^b (Lang et al., 2017), ^c (Diaconu et al., 2017), all other information is based on own data^a.

Properties	MIT	CON	TUT
Soil	Hyperdystric chromic folic cambisol ^a	Hyperdystric skeletal folic cambisol ^a	Rendzic Leptosols ^a
Humus form (thickness in cm)	Moder (8) ^b	Mor-like Moder (13) ^b	Mull (12)
Soil texture (topsoil/subsoil) ^a	Loam / Sandy Loam ^b	Loam / Sandy Loam ^b	Clay / Clay
Stone content (topsoil/subsoil) [%]	23/26	87/62	50/67
pH (H ₂ O) (forest floor/ topsoil /subsoil)	3.53 / 3.57 / 4.61 ^b	3.46 / 4.03 / 4.61 ^b	6.00 / 7.23 / 7.84
Ptot stocks forest floor [g/m ²]	7 ^b	13 ^b	19
Ptot stocks mineral soil [g/m ²]	624 ^b	230 ^b	189
isotopically exchangeable P within 1 min (Ah, BA) [mg P/kg]	9.2/3.0 ^b	4.0/5.0 ^b	- / -
Parent material	Paragneiss ^b	Paragneiss ^b	Carbonate Rock
Slope [°]	14°	28°	23°
Elevation [m a.s.l.]	1023 ^b	840 ^b	835
Dominant vegetation (mean age)	Fagus sylvatica (131 ^b)	Fagus sylvatica (132 ^b)	Fagus sylvatica (90 ^c)
Annual precipitation [mm/year]	1299 ^b	1749 ^b	900 ^c

Formatted: Font: Calibri

Formatted: Font: Calibri

Tab. 2: Depth [cm] of lysimeter (LY) and trench (TR) installations aligned with the main soil horizons for Mitterfels (MIT), Conventwald (CON) and Tuttlingen (TUT); “B” indicates the sprinkling plot and is used to be consistent with other publications from the same sites. X indicates that LY4B and TR3B could not be installed at TUT due to shallow soil depth.

Depth [cm]	MIT	CON	TUT	Lysimeter, Trench
Forest floor		0	0	0 LY1B, TR1B
Topsoil		35	40	40 LY2B
Subsoil		130	100	60 LY3B, TR2B
Saprolite		190	290	X LY4B, TR3B

925

Tab. 3: Mean pre-event water fraction [%] of the different flow components (trench: TR1B, TR2B, TR3B and lysimeter: LY1B, LY2B, LY3B, LY4B) and the 6 sprinkling experiments in Mitterfels (MIT), Conventwald (CON), Tuttlingen (TUT) in spring and summer. X indicates that this lysimeter or trench does not exist (TUT), NA indicates that [this lysimeter or trench](#) has not yielded any flow (MIT).

Event	LY1B	LY2B	LY3B	LY4B	TR1B	TR2B	TR3B
MIT Spring	12	83	95	99	9	44	NA
MIT Summer	15	88	93	98	33	33	NA
CON Spring	10	69	63	86	19	63	78
CON Summer	16	58	83	78	24	55	74
TUT Spring	4	64	95	X	14	70	X
TUT Summer	4	83	70	X	14	47	X

930

Tab. 4: Time to 20% rise in SSF [min] for all flow components and all sprinkling experiments. “-” indicates that time to 20% rise in SSF could not be calculated due to missing data at the beginning of the event (MIT in spring); X indicate that this LY or TR was not existing (TUT), NA indicates that [this lysimeter or trench](#) yielded no flow (MIT). Numbers in bold indicate cases where a flow component at deeper soil depth responded earlier than a flow component at shallower soil depth.

935

Event	LY1B	LY2B	LY3B	LY4B	TR1B	TR2B	TR3B
MIT Spring	-	95	290	515	10	80	NA
MIT Summer	30	210	535	765	20	55	NA
CON Spring	15	160	105	220	35	315	235
CON Summer	10	70	180	235	25	305	265
TUT Spring	70	85	140	X	70	145	X
TUT Summer	15	205	140	X	20	295	X

Formatted: German (Germany)

Formatted: German (Germany)

940 Tab. 5: Time to 20% event water fraction t_{20} [min] for all flow components and all sprinkling experiments; “-“ indicates that time to 20% event water fraction could not be calculated due to missing data at the beginning of the event (MIT in spring); X indicates that this LY or TR was not existing (TUT), NA indicates that **this trench** yielded no flow (MIT), “/” indicates that 20 % event water fraction was not reached during the event; numbers in bold indicate that the time to 20% event water fractions was shorter than the time to 20% rise in SSF (see Tab. 4).

Event	LY1B	LY2B	LY3B	LY4B	TR1B	TR2B	TR3B
MIT Spring	-	110	/	/	15	250	NA
MIT Summer	45	565	/	/	30	75	NA
CON Spring	25	295	325	515	45	80	65
CON Summer	15	115	285	390	35	70	80
TUT Spring	75	300	/	X	80	380	X
TUT Summer	15	480	370	X	20	235	X

Formatted: German (Germany)

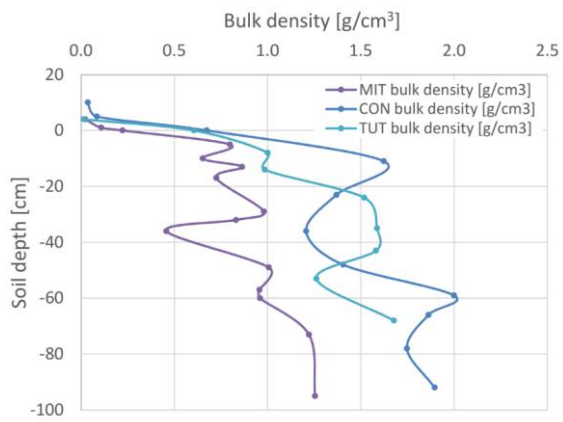
945 Tab. 6: Slope of the linear regression between Ptot concentration and fraction of event water for the lysimeters (LY) and the trenches (TR) and for all sprinkling experiments in Mitterfels (MIT), Conventwald (CON) and Tuttlingen (TUT); **Bold** values are not significantly different from a slope = zero and therefore the relation likely chemostatic (Alpha = 0.9505). *Italic* values indicate slopes that are not chemostatic and not significantly different from slopes describing simple dilution. Therefore, these relations are likely governed by proportional mixing. X indicate that this LY or TR was not existing (TUT), NA indicates that **the trench** yielded no flow (MIT).

Event	LY1B	LY2B	LY3B	LY4B	TR1B	TR2B	TR3B
MIT Spring	-0.6	0.1	0.0	1.8	-0.2	0.0	NA
MIT Summer	-2.3	-0.3	-0.6	-1.5	-0.1	-0.5	NA
CON Spring	-2.2	-1.2	0.0	0.0	-1.4	0.1	0.1
CON Summer	-0.1	0.0	-0.1	0.0	-1.0	0.1	0.1
TUT Spring	-0.1	0.0	0.0	X	-0.1	0.0	X
TUT Summer	-0.5	0.1	0.0	X	-0.6	-0.1	X

Formatted: German (Germany)

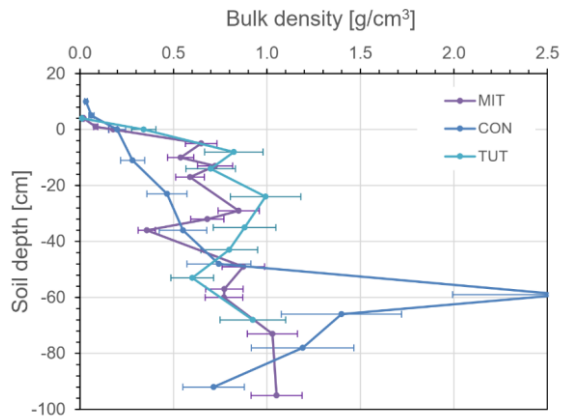
955 Tab. 7: Estimation of annual Ptot fluxes in vertical (LY) and lateral (TR) SSF at CON calculated based on the mean P concentration after the first flush during the sprinkling experiment at CON in summer (see Fig. 5) and the water flux between 10.05.2018 and 10.05.2019. This period has been chosen in order to not include the sprinkling experiments. No long-term measurements **are** available for MIT and TUT.

	LY1B	LY2B	LY3B	LY4B	TR1B	TR2B	TR3B
Water flux [mm/a]	295	239	320	102	2	4	1
Median P concentration after first flush [mg/l]	0.17	0.07	0.02	0.03	0.15	0.05	0.04
Estimated P flux [kg/ha/a]	0.49	0.16	0.07	0.03	0.00	0.00	0.00



960

	LY1B	LY2B	LY3B	LY4B	TR1B	TR2B	TR3B
Water flux [mm/a]	295	239	320	102	2	4	1
Median P concentration after first flush [mg/l]	0.17	0.07	0.02	0.03	0.15	0.05	0.04
Estimated P flux [kg/ha/a]	0.49	0.16	0.07	0.03	0.00	0.00	0.00



965 **Fig. 1:** Depth profile of soil bulk density of the three experimental sites Mitterfels (MIT), Conventwald (CON) and Tuttingen (TUT).

970 : Depth profile of soil bulk density (mass fine earth per volume fine earth) of the three experimental sites Mitterfels (MIT), Conventwald (CON) and Tuttingen (TUT). Due to the high stone contents of the soils the “quantitative soil pit” (QP) as described in (Lang et al., 2017) was used to determine bulk densities. Error bars represent standard deviation calculated based on the average coefficient of variance of the method determined at the corresponding study sites.

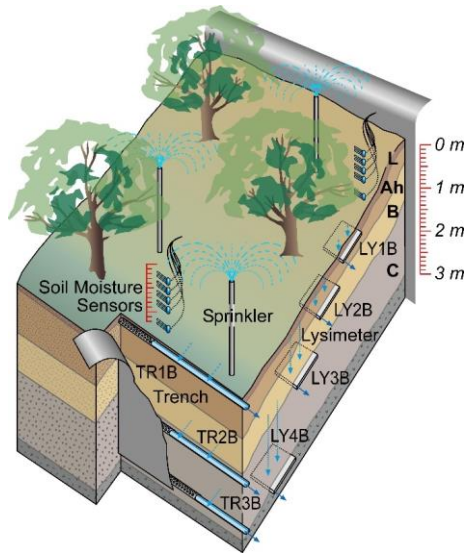
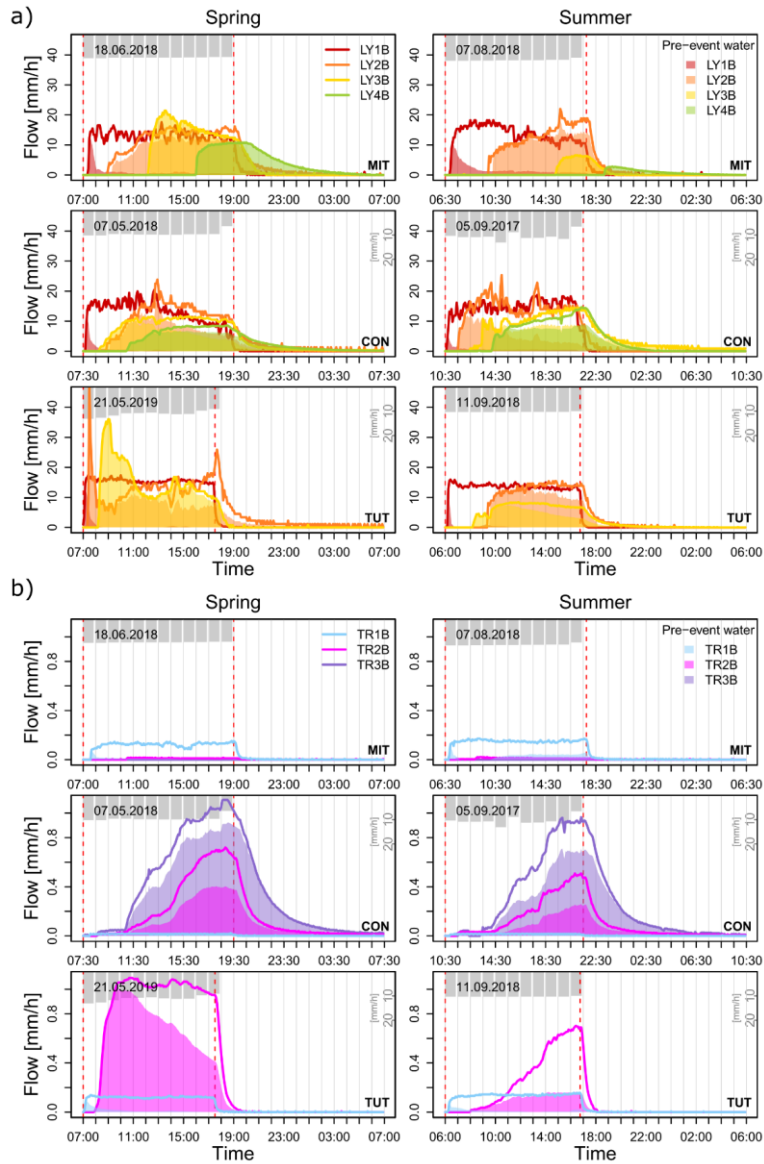
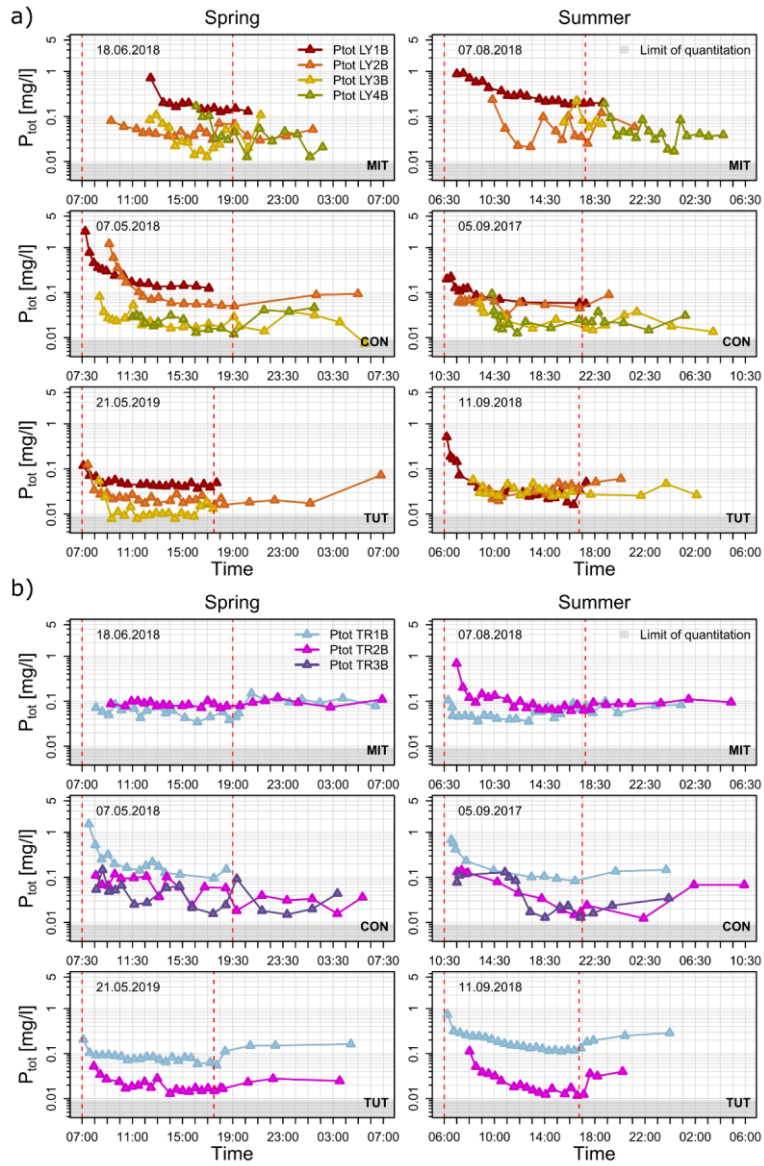


Fig. 2: Experimental setup to measure vertical and lateral subsurface flow with drainage pipes at three soil depths in a trench (TR1B, TR2B, TR3B) and zero tension lysimeters (i.e. steel piling plates pushed in the undisturbed soil profile from the side of the hillslope; LY1B, LY2B, LY3B, LY4B). "B" indicates the sprinkling plot and is used to be consistent with other publications from the same sites.

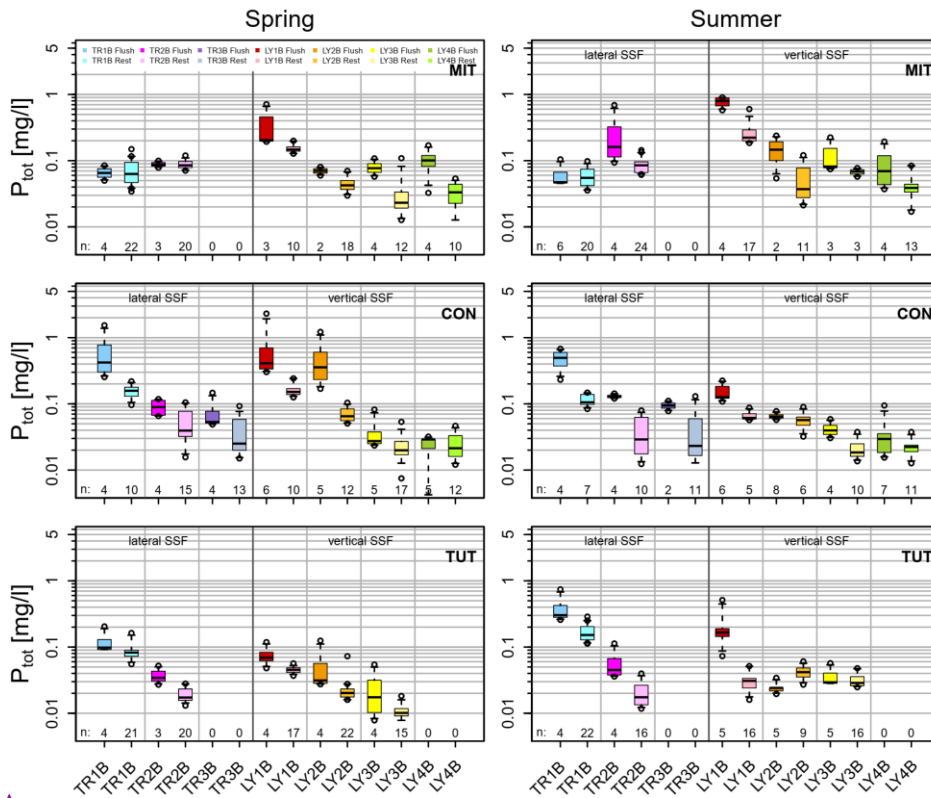


980 Fig. 3: a) Vertical subsurface flow from the lysimeters (LY1B, LY2B, LY3B, LY4B) and b) lateral subsurface flow from the trench (TR1B, TR2B, TR3B) for the sprinkling experiments in Mitterfels (MIT), Conventwald (CON) and Tuttlingen (TUT) sorted in rows and for spring (left column) and summer (right column). Color-shaded areas show the pre-event water contribution to total discharge of each component (calculated from the Deuterium tracer). Red dashed lines indicate the start and end of the sprinkling experiment. Gray bars show sprinkling input during the experiments.



985

Fig. 4: Ptot concentrations of (a) vertical subsurface flow and (b) lateral subsurface flow for the sprinkling experiments in Mitterfels (MIT), Conventwald (CON) and Tuttingen (TUT) sorted in rows and for spring (left column) and summer (right column). Red dashed lines indicate the start and end of the sprinkling experiment, light gray ~~area indicates~~ areas indicate range of Ptot below limit of quantitation.



990 Fig. 5: P_{tot} concentrations of vertical and lateral subsurface flow (SSF) during the flushing period defined as the first 2 h after onset of each flow component (darker colors) and the remaining part of the event (10 to 12 h) in lighter colors for the sprinkling experiments in Mitterfels (MIT), Conventwald (CON) and Tuttingen (TUT).

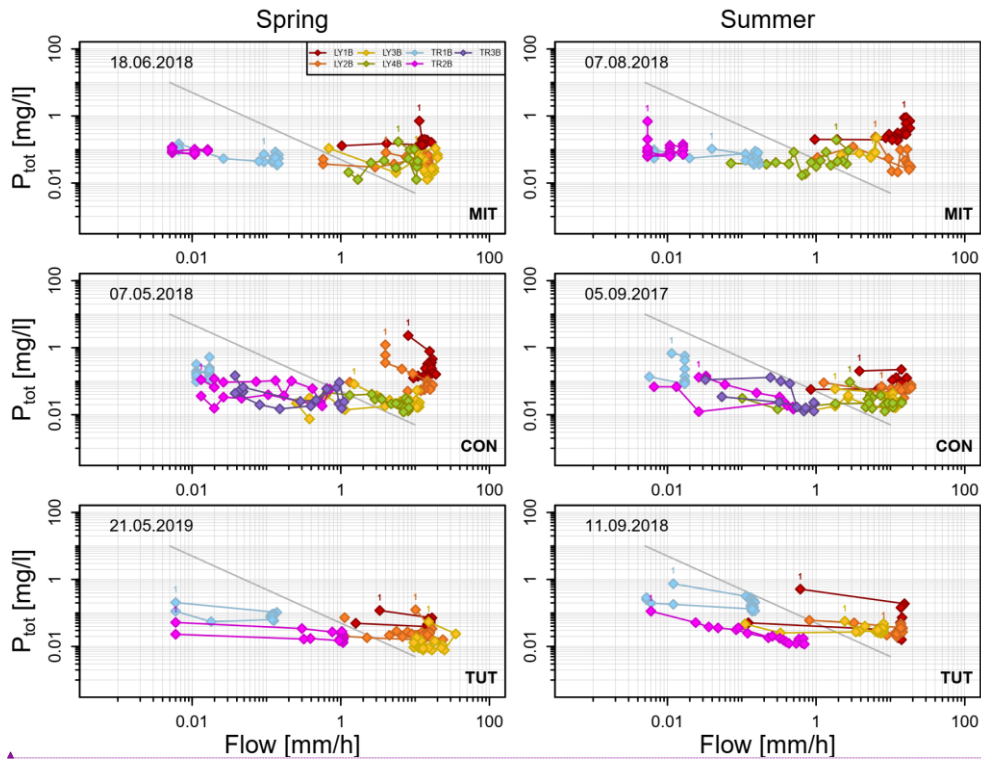
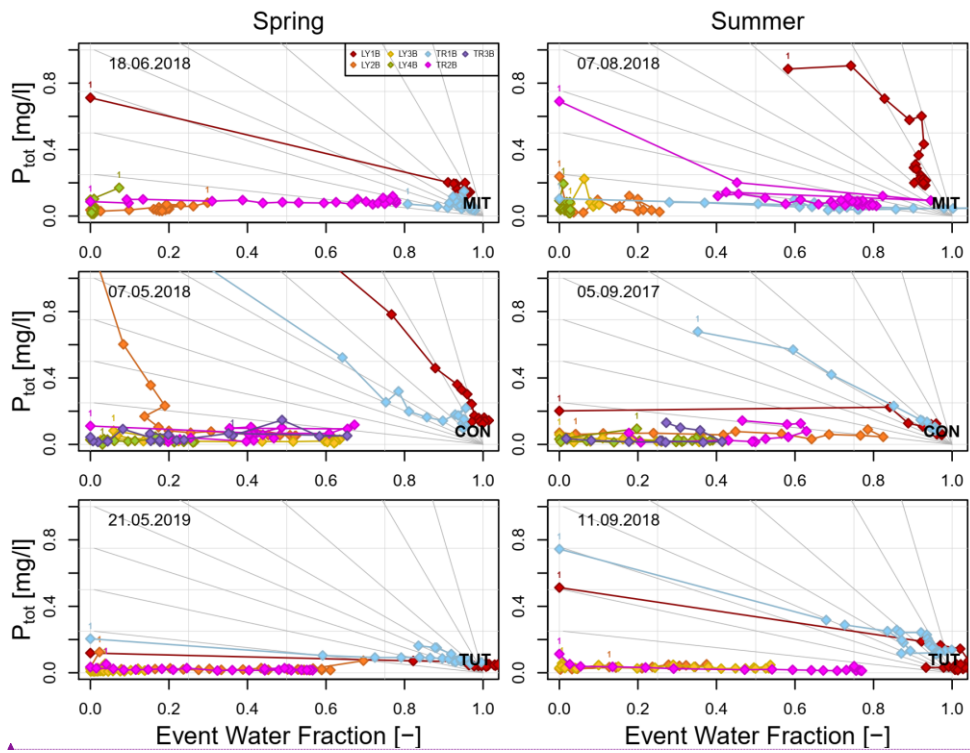


Fig. 6: P_{tot} concentrations as a function of subsurface flow in the lysimeters (LY) and the trench (TR) during the sprinkling experiments in Mitterfels (MIT), Conventwald (CON) and Tuttlingen (TUT). The gray diagonal line indicates simple dilution ($C \sim 1/Q$). Label "1" indicates the first data point of each flow component.



Formatted: German (Germany)

1000

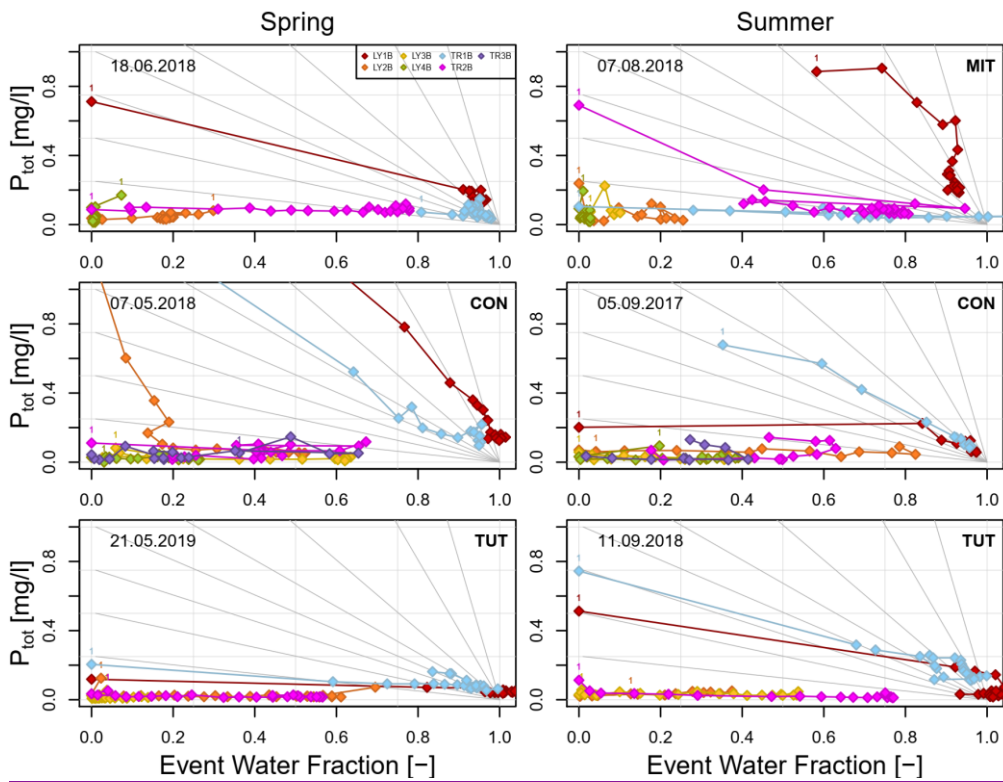


Fig. 7: P_{tot} concentrations as a function of event water fraction calculated for the sprinkling experiments. Gray lines indicate a selection of possible lines that describe the theoretical change in P_{tot} concentration assuming simple dilution (i.e., proportional mixing of event and pre-event water according to the event water fraction). Label "1" indicates the first data point of each flow component.

1005

1010 **11. Supplement**

Suppl. Tab. 1: Difference between the median P_{tot} concentration on the rising and falling limb of the sprinkling experiments in [mg/l]. Positive values indicate that P_{tot} concentrations on the rising limb are higher than on the falling limb and negative values vice versa. Small values indicate that the difference between the median P_{tot} concentration on the rising and falling limb is small which means that the hysteresis is small and vice versa. Bold values indicate that the difference between P_{tot} concentration on the rising and falling limb is significantly different. ($\alpha = 0.05$). X indicates that this lysimeter or trench does not exist (TUT), NA indicates that this trench has not yielded any water sample during the experiment (MIT).

Event	LY1B	LY2B	LY3B	LY4B	TR1B	TR2B	TR3B
MIT Spring	0.051	0.007	0.059	0.011	0.008	0.017	NA
MIT Summer	0.642	0.010	0.072	0.104	-0.007	0.117	NA
CON Spring	0.176	0.113	0.007	-0.005	0.349	0.035	0.027
CON Summer	0.063	0.004	0.012	0.000	0.488	0.059	0.062
TUT Spring	0.049	0.003	0.029	X	0.072	0.008	X
TUT Summer	0.319	-0.019	0.003	X	0.106	0.006	X

Formatted: German (Germany)