Science of the Total Environment 755 (2021) 143453

Contents lists available at ScienceDirect

Science of the Total Environment

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

Yield response to soil test phosphorus in Switzerland: Pedoclimatic drivers of critical concentrations for optimal crop yields using multilevel modelling

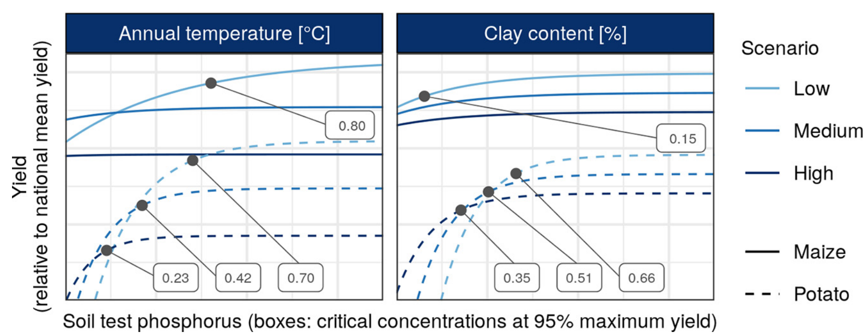
Juliane Hirte [a,*], Walter Richner [a,1], Barbara Orth [a], Frank Liebisch [a,b], René Flisch [a]

[a] *Agroscope, Agroecology and Environment, Water Protection and Substance Flows, 8046 Zurich, Switzerland*
[b] *ETH Zurich, Department of Environmental Systems Sciences, Institute of Agricultural Sciences, Crop Science Group, 8001 Zurich, Switzerland*

HIGHLIGHTS

- Critical soil test P concentrations for optimal crop yields
- 26 years of P fertilization on six sites with diverse crops, soil, and climate
- Two soil test methods and nonlinear multivariate multilevel yield response models
- Critical concentrations depend strongly on crop, temperature, and soil clay content.
- P fertilization guidelines can be improved by integrating soil and climate data.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 18 June 2020
Received in revised form 20 October 2020
Accepted 20 October 2020
Available online 9 November 2020

Editor: Charlotte Poschenrieder

Keywords:
STP
Critical concentrations
Nonlinear mixed models
Multivariate analysis
Mitscherlich function
Long-term field trials

ABSTRACT

Phosphorus (P) management in agroecosystems is driven by opposing requirements in agronomy, ecology, and environmental protection. The widely used maintenance P fertilization strategy relies on critical concentrations of soil test P (STP), which should cause the lowest possible impact on the environment while still ensuring optimal yield. While both soil P availability and crop yields are fundamentally related to pedoclimatic conditions, little is known about the extent to which soil and climate variables control critical STP. The official P fertilization guidelines for arable crops in Switzerland are based on empirically derived critical concentrations for two soil test methods (H_2O-CO_2 and AAE10). To validate those values and evaluate their relation to pedoclimatic conditions, we established nonlinear multivariate multilevel yield response models fitted to long-term data from six sites. The Mitscherlich function proved most suitable out of three functions and model fit was significantly enhanced by taking the multilevel data structure into account. Yield response to STP was strongest for potato, intermediate for barley, and lowest for wheat and maize. Mean critical STP at 95% maximum yield ranged among crops from 0.15–0.58 mg kg^{-1} (H_2O-CO_2) and 0–36 mg kg^{-1} (AAE10). However, pedoclimatic conditions such as annual temperature or soil clay content had a large impact on critical STP, entailing changes of up to 0.9 mg kg^{-1} (H_2O-CO_2) and 80 mg kg^{-1} (AAE10). Critical STP for the AAE10 method was also affected by soil pH. Our findings suggest that the current Swiss fertilization guidelines overestimate actual crop P demand on average and that site conditions account for large parts of the variation in critical STP. We propose that site-specific

Abbreviations: AAE10, ammonium-acetate-EDTA; H_2O-CO_2, carbon dioxide-saturated water; STP_{AAE10}, soil test phosphorus based on extraction with ammonium-acetate-EDTA; $STP_{H2O\text{-}CO2}$, soil test phosphorus based on extraction with carbon dioxide-saturated water.

* Corresponding author.
E-mail address: juliane.hirte@agroscope.admin.ch (J. Hirte).
[1] Present address: Canton of St. Gallen, Office for the Environment, Lämmlisbrunnenstrasse 54, 9001 St. Gallen, Switzerland.

https://doi.org/10.1016/j.scitotenv.2020.143453

fertilization recommendations could be improved on the basis of agro-climate classes in addition to soil information, which can help to counteract the accumulation of unutilized soil P by long-term P application.

1. Introduction

Phosphorus (P) dynamics in agroecosystems are highly complex and P management is still controversially discussed, driven by partially opposing requirements in agronomy, ecology, and environmental protection. Due to its outstanding role as carrier of genetic information and energy in plants, P strongly governs crop development during early vegetative growth (White and Hammond, 2008). Yet, once entering the soil, it becomes rapidly immobilized by adsorption, precipitation, or complexation and may therefore be the most limited nutrient element in its immediate bioavailability (Vance et al., 2003; Menezes-Blackburn et al., 2018). In the medium and long term, however, a significant proportion of this fixed P can be mobilized under favourable biophysical conditions and can add to the pool of plant available soil P (Rowe et al., 2016). Rate and timing of P release from soil are largely unpredictable though, which still hampers integrated P management (Menezes-Blackburn et al., 2018; Schneider et al., 2019).

As a consequence of excessive P fertilization in Europe since the mid-20th century, agricultural topsoils have accumulated vast amounts of P (Tunney et al., 2003), while surface water quality has deteriorated tremendously (Némery et al., 2005; Kleinman et al., 2011). Eutrophication-control policies have led to a strong decline in P fertilization rates in recent years but the legacy effect of soil P accounts for continuing P flows from agricultural areas to surface waters (Jarvie et al., 2013). In Switzerland, those inputs still average more than 1 Gg P $year^{-1}$ or 25% of total annual diffuse P inputs (Mehr et al., 2018). Consequently, surface waters in P accumulation hotspots often do not fulfil water quality requirements, putting agricultural P recurrently into focus of eutrophication management in Switzerland (FOAG, 2018).

The maintenance P fertilization strategy pursued in most European countries including Switzerland aims at achieving and sustaining soil P levels at which P inputs replace crop P off-takes (Flisch et al., 2017; Leinweber et al., 2018). This requires a three-step routine: testing the soil for available P by extraction of a certain P fraction, classifying soil test P (STP) according to soil P fertility classes, and calculating fertilization rates depending on fertility class and additional information on soil properties and management (Jordan-Meille et al., 2012; Flisch et al., 2017).

The second step essentially relies on class-defining critical STP concentrations to increase the agronomic efficiency and decrease the environmental strain by fertilization (Kirkby and Johnston, 2008; Rowe et al., 2016). Critical concentrations are usually established by modelling yield response to STP, i.e. fitting an appropriate response function to long-term multi-experiment data (Jordan-Meille et al., 2012). As crop yields have repeatedly been maintained at STP below the critical concentration (Schneider et al., 2019), further information on site conditions may aid to reduce established critical concentrations (Rowe et al., 2016).

Numerous response functions have traditionally been applied in agricultural research (Colwell et al., 1988; Archontoulis and Miguez, 2015), of which some are well-established to describe yield response to STP and derive critical concentrations: the Cate-Nelson split/linear-plateau function (Colomb et al., 2007; Tang et al., 2009; Recena et al., 2016; Wu et al., 2018), the Mitscherlich function (Colomb et al., 2007; Tang et al., 2009; Valkama et al., 2011; Watmuff et al., 2013; Sucunza et al., 2018; Wu et al., 2018), and quadratic, square-root quadratic, or hyperbolic functions (Valkama et al., 2011; Watmuff et al., 2013; Cadot et al., 2018).

Critical STP concentrations are usually derived at a specified percentage of maximum predicted yield (Jordan-Meille et al., 2012). Morel et al. (1992) defined three scenarios with agronomic meaning, represented by 90, 95, and 98.5% of maximum yield: A yield reduction of 10% is highly significant, while that of 5% is at the margin of statistical significance, justifying surplus and replacement P fertilization to increase or maintain crop yields, respectively. Yields ranging within the biological variation of 1.5% equal maximum yield, at which any fertilization cannot be justified.

While additional explanatory variables describing management practices, soil properties, or climate conditions have been used early to refine simulated relations between STP and yield (Colwell et al., 1988; Mallarino and Blackmer, 1992), the complexity of long-term multi-experiment data remained difficult to model appropriately (e.g. Buczko et al., 2018). Nonlinear mixed models provide powerful means to account for (nested) group-level effects, distinguish between mean effects of predictors and induced random variances, and handle unbalanced data (Yang, 2010; Schielzeth and Nakagawa, 2013). Hence, they are the foremost method to comprehensively analyse multilevel experiments (Loughin, 2006; Moore and Dixon, 2015) and have become increasingly popular for agronomic questions (Gonçalves et al., 2016; Parent et al., 2017; Sari et al., 2018).

The Principles of Agricultural Crop Fertilisation in Switzerland (PRIF) outline, among other agronomic subjects, the official P fertilization guidelines. Those are based on soil extraction by water, carbon dioxide-saturated water (Dirks and Scheffer, 1930), or ammonium-acetate-EDTA (Lakanen and Erviö, 1971). The latter two are routinely applied in crop production (Flisch et al., 2017) and will be abbreviated as H_2O-CO_2 and AAE10, respectively, from here on. The H_2O-CO_2 method has a long tradition as official soil testing method in Switzerland (Hasler, 1957; Peyer and Frei, 1971) and the AAE10 method is or has been also used for soil P testing in Belgium (Houben et al., 2011; Renneson et al., 2016) and Texas, US (Hons et al., 1990; Woodard et al., 1994). The two methods target different P fractions and thus differ in STP by up to two orders of magnitude (Demaria et al., 2005): While H_2O-CO_2 releases immediately plant-available P ($STP_{H2O\text{-}CO2}$) similar to water, AAE10 also breaks stable P bonds with calcium, iron, and aluminium similar to ammonium or calcium lactate and releases P (STP_{AAE10}) that may become plant-available in the medium to long term (Neyroud and Lischer, 2003; Stuenzi, 2006b, 2006a). In calcareous samples, however, rapid exhaustion of EDTA by calcium reduces the extraction strength of AAE10 significantly (Stuenzi, 2006b), which is why the AAE10 method is not recommended for soils with pH > 6.8 (Flisch et al., 2017).

The PRIF comprise five soil P fertility classes that depend on STP and soil clay content and indicate deficient (A), moderate (B), sufficient (C), reserve (D), or accumulated (E) soil P levels (Flisch et al., 2017). While class-defining critical concentrations for $STP_{H2O\text{-}CO2}$ were derived from empirical long-term observations, those for STP_{AAE10} were recalculated from $STP_{H2O\text{-}CO2}$, postulating linearity between $STP_{H2O\text{-}CO2}$ and STP_{AAE10}. To validate the critical concentrations and the corresponding fertility classes, Agroscope started long-term field experiments between 1989 and 1992 on six sites, which now provide the data basis for comprehensive multivariate multilevel analyses. Our objectives were to (i) establish yield response models to $STP_{H2O\text{-}CO2}$ and STP_{AAE10} and derive critical $STP_{H2O\text{-}CO2}$ and STP_{AAE10} concentrations for optimal yield of different crops, (ii) evaluate pedoclimatic effects on critical $STP_{H2O\text{-}CO2}$ and STP_{AAE10}, and (iii) assess the validity of the soil P fertility classes currently adopted in Switzerland.

2. Material and methods

2.1. Sites, crop rotations, and management practices

The study comprises data of six long-term field trials that were established between 1989 and 1992 and have been running for 20 to 30 years (Table 1). While Rümlang-Altwi (ALT), Ellighausen (ELL), Grabs (GRA), Oensingen (OEN), and Zurich-Reckenholz (REC) are situated in the Swiss Plateau north of the Alps, Cadenazzo (CAD) is located in the Magadino plain in the southern outskirts of the Alps (Supplementary Fig. 1). The sites are representative of the main crop growing regions of the eastern part of Switzerland and range in mean annual temperature from 9.0 to 11.4 °C and in mean cumulated annual precipitation from 920 to 1830 mm (Table 1). Prior to trial establishment, all sites had been used as cropland under conventional management and had undergone a transition period without fertilization of one year. Soil depth ranged from 60 cm at ALT to 80 cm at GRA and 120 cm at CAD, ELL, OEN, and REC, and calcium carbonate content (measured in 2009) from 0% at CAD, ELL, and REC to 1% at OEN, 5% at ALT, and 17% at GRA. Soil types and physico-chemical characteristics are given in Table 2.

The crop rotations at each site reflected prevailing agricultural practices in the respective sub-regions. On average, the crop rotations comprised 20% winter wheat, 20% grass-clover ley, 19% maize, 14% potato, 11% winter barley, and less than 5% soybean, rapeseed, fodder beet, sugar beet, summer wheat, spinach, and chicory each (Supplementary Table 1). Cover crops were grown in 22% of the growing seasons. All sites have been managed according to best agricultural practice until 1993, the Swiss guidelines for "Integrated Production" from 1993 to 1998, and the Swiss direct payment system "Proof of Ecological Performance" from 1998 on (Swiss Federal Council, 2013). Pests and diseases were chemically controlled and the soil on all sites was regularly ploughed before sowing to a depth of 0.18–0.25 m, with the exception of GRA, where maize was generally sown without tillage. Aboveground biomass was completely removed from the fields except for potato haulm and stubbles of main and cover crops.

On each site, the long-term treatments on fixed plots comprised zero, deficit, reduced, norm, elevated, and surplus P fertilization (Table 3), which reflected rates of recommended P inputs of 0, 33, 67, 100, 133, and 167%, respectively, according to the PRIF (Flisch et al., 2009, and preceding versions). Phosphorus was applied as superphosphate in one dose usually before tillage and sowing. The trials were arranged in completely randomized block designs with four field replications, resulting in 24 experimental plots per site. The dimensions of the plots were 4.5 by 9.25 m^2 in CAD and 4.0 by 8.25 m^2 on all other sites. Nitrogen (N), potassium, and magnesium were adequately applied in equal amounts on all plots after annual soil testing according to the PRIF (Flisch et al., 2009, and preceding versions).

2.2. Plant and soil analyses

At harvest, crop yields were determined on 2 × 7 m^2 subplots in the centre of the experimental plots. Shortly after harvest, the topsoil (0–0.2 m) was sampled by randomly taking 10–12 cores with an Edelman auger (20 mm diameter; Eijkelkamp, Netherlands) and combining those to one composite sample per plot.

After oven-drying at 40 °C and sieving through a 2-mm mesh, all samples were analysed for pH (H_2O) and H_2O-CO_2- and AAE10-extractable soil P according to the Swiss Reference Methods of the Federal Agricultural Research Stations for Analysis of Soils (Agroscope, 1996). In brief, 30 g of dry soil was agitated for 1 h in demineralized water containing 0.08 M CO_2 at a ratio of 1:2.5. Similarly, 10 g of dry soil was agitated for 1 h in 0.5 M ammonium-acetate + 0.5 M acetic acid + 0.025 M EDTA at a ratio of 1:10. Dissolved P was measured using colorimetry, by which inorganic and hydrolysable organic P forms are detected (Stuenzi, 2006a). The analytical uncertainty was 0.05 and 2.5 mg kg^{-1} $STP_{H2O\text{-}CO2}$ and STP_{AAE10}, respectively.

2.3. Data preparation and treatment effects

Crop yields are reported in Mg dry matter ha^{-1} and $STP_{H2O\text{-}CO2}$ and STP_{AAE10} in mg P kg^{-1} dry soil. In this study, yield refers to the main product only, e.g. cereal and maize grain or potato tuber. The dataset was limited to crops with multiple spatial and temporal replications, i.e. winter wheat ($n = 696$; from here on: wheat), maize ($n = 648$), potato ($n = 480$), and winter barley ($n = 360$; from here on: barley). Although frequently grown, grass-clover ley ($n = 672$) was excluded from analysis due to its biennial cultivation that entails differences in yield formation compared to arable crops. This resulted in a total of 2184 data points from 91 site-by-year combinations (Supplementary Table 1). Outliers in the data were tested using Dixon's Q test, Huber-type skipped mean (Hampel, 1985), and skewness-adjusted boxplots for multivariate skewed data (Hubert and Van der Veeken, 2008). The identified values were removed from the data set and replaced with newly measured values of the respective retained samples.

To reduce the impact of seasonal variability and historical breeding success on yield and to compare different crops in the same y-axis range, we converted absolute to relative yields [%] based on mean annual crop yields in Swiss conventional farming practice (recalculated to dry matter; Agristat, 1990-2017; Erdin, 2018). In multi-experiment studies, relative yields have traditionally been calculated based on year- and site-specific yield maxima to reduce the impact of soil and climate conditions on yield response to STP. This approach is however prone to statistical bias as error deviations are weighted differently for each site. In addition, yield maxima are often poorly defined by experimental data and their deviation is highly subjective (Colwell et al., 1988). Therefore, we deliberately allow for pedoclimatic effects on the relation of yield to STP by including a selection of variables as covariates in our models.

We retrieved the following climate variables from the Federal Office of Meteorology and Climatology MeteoSwiss: mean annual temperature [°C], sum of annual precipitation [mm], sum of annual sunshine duration [h], and deviation of temperature [°C], precipitation [%], and sunshine duration [%] from the respective 30-year climate norm (World Meteorological Organization, 2017) during the months of vegetative crop growth (wheat, barley: March/April; maize: June/July; potato: May–July; Supplementary Table 2).

Table 1
Location, climate, and management period of the long-term field trials.

Site	Location	N–E coordinates	Altitude [m]	MAT[a] [°C]	MAP[a] [mm]	Start [years]	End [years]
ALT	Altwi	47°26′19.01″–8°31′31.82″	493	9.4	1054	1989	Ongoing
CAD	Cadenazzo	46°9′37.66″–8°56′2.18″	203	11.4	1832	1990	2009
ELL	Ellighausen	47°36′35.06″–9°8′33.07″	507	9.3	947	1989	Ongoing
GRA	Grabs	47°11′39.76″–9°28′11.43″	442	10.1	923	1992	2012
OEN	Oensingen	47°17′2.95″–7°43′50.77″	454	9.0	1129	1989	Ongoing
REC	Reckenholz	47°25′50.5″–8°31′19.3″	440	9.4	1054	1989	Ongoing

[a] Mean annual temperature and mean cumulated annual precipitation for the climate norm 1981–2010.

Table 2
Soil type of the long-term field trials and soil bulk density, texture, organic C (Corg), and pH at the start of the experiments and in 2014 (averaged over fertilization treatments) in the topsoil (0–0.2 m).

Site	Soil type (WRB)	Bulk density [g cm^{-3}]	Clay [%]	Sand [%]	C_{org} [g kg^{-1}]	pH (H_2O) at start	pH (H_2O) in 2014[a]
ALT	Calcaric Cambisol	1.37	22	48	21	7.9	7.9
CAD	Eutric Fluvisol	1.22	8	40	14	6.3	5.8
ELL	Eutric Cambisol	1.22	33	31	23	6.6	6.6
GRA	Calcaric Fluvisol	1.25	17	34	16	8.3	8.1
OEN	Gleyic-calc. Cambisol	1.30	37	32	24	7.1	7.6
REC	Eutric Gleysol	1.14	39	25	27	7.4	6.7

[a] in last year of management period in CAD (2009) and GRA (2012).

Treatment effects on relative yield, $STP_{H2O\text{-}CO2}$, and STP_{AAE10} were determined separately for each crop for the entire time series. The respective data subsets were fitted to linear mixed models with treatment as fixed effect and year, site, and block as nested random effects. Differences were determined by ANOVA with Kenward-Roger approximation of degrees of freedom at a significance level of $p < 0.05$ and subsequent simultaneous multiple comparison of estimated marginal means of treatment pairs with Tukey-adjustment of p-values.

2.4. Multivariate multilevel yield response models

2.4.1. Selection of covariates

We evaluated the following pedoclimatic variables for their suitability as covariates in the yield response models: soil pH, clay content, silt content, annual temperature, precipitation, and sunshine hours, and deviation of temperature, precipitation, and sunshine hours from the 30-year norm. We tested whether they (i) were meaningful for yield response (Colwell et al., 1988), (ii) explained parts of the yield variation induced by the spatial and temporal grouping of the data (Loughin, 2006), and (iii) were minimally collinear (Bonate, 2017) using (i) linear regression and ANOVA (details: Supplementary Table 3), (ii) principal component analysis, and (iii) Pearson correlation analysis. We thus selected soil pH, clay content, annual temperature, annual precipitation, and deviation of precipitation from the norm as numerical covariates in addition to P fertilization. Although soil organic matter can affect yield response to soil P for different reasons (Johnston et al., 2014; Schneider et al., 2019), this variable could not be included as a covariate due to lack of data.

We initially applied these covariates in models fitted to data subsets of wheat, barley, maize, and potato separately to test if they were adversely related to the model parameters for different crops. As this was not the case with few exceptions, we used models fitted to the entire dataset with 'crop' as additional factorial covariate. By switching the reference level of 'crop', we extracted crop-specific intercepts for parameter means and differences in parameter intercepts between crops (based on contrasting parameter and coefficient estimates). This resulted in six models covering three functions and two soil extraction methods, each fitted to 2184 data points and including the same set of covariates.

Table 3
Fertilization treatments, mean fertilizer P inputs for wheat, barley, maize, and potato averaged over the entire time series, and cumulated P balances averaged over sites.

Fertilization treatment		Applied fertilizer P [kg ha^{-1} $year^{-1}$]				P balance[b] [kg ha^{-1}]
Description	P input [%][a]	Wheat	Barley	Maize	Potato	
Zero	0	0	0	0	0	−113 (−245–113)
Deficit	33	10	10	14	14	−12 (−110–226)
Reduced	67	20	20	28	28	117 (9–352)
Norm	100	31	30	42	41	148 (52–295)
Elevated	133	41	40	56	55	415 (216–667)
Surplus	167	51	50	70	69	408 (326–479)

[a] Relative to recommended amounts according to Swiss fertilization guidelines (PRIF).
[b] Start of the trials until 2014 or last year of management period (CAD: 2009; GRA: 2012); numbers in brackets: minimum - maximum of sites.

2.4.2. Response functions

We used three nonlinear three-parameter functions to relate $STP_{H2O\text{-}CO2}$ and STP_{AAE10} to relative yield, a (1) linear-plateau, (2) monomolecular (Mitscherlich), and (3) quadratic function (Dahnke and Olson, 1990; Archontoulis and Miguez, 2015; Parent et al., 2017):

$$Y_{rel} = intercept + slope{*}STP \rightarrow maximum \tag{1}$$

$$Y_{rel} = asymptote{*}\left(1 - e^{-rate{*}(STP + environment)}\right) \tag{2}$$

$$Y_{rel} = a{*}(STP)^2 + b{*}STP + c \tag{3}$$

where Y_{rel} is relative yield; STP is either $STP_{H2O\text{-}CO2}$ or STP_{AAE10}; *intercept* and *c* denote the intercept with the y axis and *environment* the intercept with the x axis; *slope*, *rate*, and *a* describe the steepness of the curve; and *maximum*, *asymptote*, and the vertex point of the quadratic function $((4ac - b^2) * (4a)^{-1})$ represent the maximum attainable yield.

2.4.3. Multilevel models

Linear combinations of crop and scaled (standardized for zero-mean and unit-variance) numerical covariates were applied as fixed effects to all model parameters. Year (as factor), site within years, and block within sites were handled as nested random effects and, thus, accounted for the temporal and spatial grouping of the data. To avoid overfitting of the models, random effects were applied to the intercepts of *maximum*, *asymptote*, and *b* only because we expected the largest impact of seasonal variability and site conditions on maximum attainable yield.

The linear-plateau and Mitscherlich models are sensitive to starting values. For each model, we searched a grid with multiple triplets of plausible parameters (Ritz and Streibig, 2008) and chose those triplets that led to model convergence with a positive-definite correlation matrix and returned the lowest Akaike information criterion (AIC) and upper 95% confidence limits of the standard deviations of the random effects. We evaluated the models based on residual pattern analysis, information criterion statistics, and likelihood ratio tests using maximum likelihood estimation (Pinheiro and Bates, 2006). The random effects level of block did not add to better model fits and was discarded in favour of computational time.

We eventually chose the Mitscherlich function for all further analyses (see Section 3.3 Model fit). An important asset of the Mitscherlich function is the easy interpretability of its parameters (Holford, 2015). *Asymptote* defines maximum attainable yield and *rate* is an inverse measure of the amount of STP required to increase yield (Holford et al., 1985). In a study of yield response to N fertilization, *environment* has been described as "the fertilizer-equivalent dose provided by environmental conditions", e.g. by mineralization of organic matter (Parent et al., 2017). In this study, *environment* represents the STP-equivalent additionally utilized by plants but not measured via extraction, e.g.

recently applied fertilizer P, mineralized organic P, or mobilized P through root-induced changes in the rhizosphere. *Rate* and *environment* in conjunction give information about the required amount of STP to reconcile the yield difference between zero STP and STP at maximum yield, i.e. the crop responsiveness to changes in STP (Holford et al., 1985).

The final multivariate multilevel model was composed as follows:

$$Y_{rel} = (\beta X + uZ + e_A)*\left(1-e^{-(\gamma X+e_R)*(STP+(\delta X+e_E))}\right) \quad (4)$$

where Y_{rel} is relative yield; STP is either $STP_{H2O\text{-}CO2}$ or STP_{AAE10}; β, γ, and δ are vectors of fixed effects coefficients; X is the fixed effects model matrix (same for all model parameters), u is a vector of random effects coefficients; Z is the random effects model matrix; and e_A, e_R, and e_E are the residual errors associated with *asymptote*, *rate*, and *environment*, respectively.

Parameter means and regression coefficients were estimated via restricted maximum likelihood and considered as significantly different from zero at $p < 0.05$. As correlation between covariates might have entailed false significances of the coefficients, we additionally calculated correlation-adjusted t-scores (CAT-scores) from the square root of the inverse correlation matrix and the vector of t-scores (Zuber and Strimmer, 2009).

2.5. Critical soil test phosphorus and sensitivity analyses

Critical STP was determined at 90, 95, and 98.5% of maximum predicted relative yield (Morel et al., 1992). The covariate for P fertilization was set to its minimum (i.e. 0% P fertilization) while the remaining covariates were set to their means. Hence, critical STP was calculated for the 0% P fertilization treatment only. All calculated critical concentrations < 0 mg kg^{-1} STP, e.g. when yield was already close to its predicted maximum at zero STP, are presented as 0 mg kg^{-1} STP.

We evaluated the sensitivity of critical STP to changes in the model (i) parameters for different crops and (ii) coefficients (excluding crop) at 95% of maximum predicted yield: (i) While two parameters were kept at their median, the third one was varied along its 10 to 90% quantile range among all groups of sites within years. (ii) Similarly, all coefficients but one were kept at their median (except P fertilization: minimum = 0% P) and low, intermediate, and high scenarios, corresponding to the 10, 50, and 90% quantiles, respectively, were calculated for the remaining coefficient.

2.6. Computation and data visualization

All analyses were performed in the R environment, version 3.5.3 (R Core Team, 2019), with the R packages doBy, dplyr, plyr, reshape2, and stringr for data management (Wickham, 2007, 2011; Højsgaard and Halekoh, 2018; Wickham et al., 2018; Wickham, 2019), lme4 and nlme for fitting linear and nonlinear models (Bates et al., 2015; Pinheiro et al., 2018), emmeans, lmerTest, multcomp, multcompView, pbkrtest, and psych for statistical analyses (Hothorn et al., 2008; Halekoh and Højsgaard, 2014; Graves et al., 2015; Kuznetsova et al., 2017; Lenth, 2018; Revelle, 2018) and ggplot2, GGally, ggbiplot, ggrepel, gridExtra, and gtable for data visualization (Vu, 2011; Wickham, 2016; Auguie, 2017; Schloerke et al., 2018; Slowikowski, 2018; Wickham and Pedersen, 2019).

3. Results

3.1. Crop yields and soil test phosphorus

Under norm P fertilization, dry matter yields of wheat, barley, maize, and potato averaged 5.3, 5.6, 9.2, and 8.1 Mg ha^{-1}, respectively (Table 4). Those yields corresponded to 107, 105, 112, and 91%, respectively, of Swiss mean yields in agricultural practice over the entire time series (Table 4). Relative yields of all crops were lowest under zero P fertilization and increased significantly with rising P inputs to norm fertilization, although by different increments among crops (Fig. 1). For wheat, barley, maize, and potato, respectively, relative yields were 6, 11, 6, and 11% larger in the highest-yielding treatment than under zero fertilization. Relative yields were statistically similar under norm, elevated, and surplus P fertilization, except for maize yields, which increased further towards surplus P fertilization (Fig. 1; Supplementary Fig. 3). Mean $STP_{H2O\text{-}CO2}$ and STP_{AAE10}, respectively, increased significantly with increasing P inputs from 0.44 and 25 mg kg^{-1} under zero fertilization to 1.23 and 50 mg kg^{-1} under surplus fertilization (Fig. 2).

Table 4
Wheat, barley, maize, and potato yields in Swiss agricultural practice during 1990–2015 (averaged over years) and absolute and relative yields under norm fertilization (100%) in this study (averaged over years, sites, and field replications). Mean ± standard deviation (at year level).

	CH yield [Mg ha^{-1}]	Absolute yield [Mg ha^{-1}]	Relative yield [%]
Wheat	5.1±0.4	5.3±0.5	107±11
Barley	5.4±0.5	5.6±1.0	105±24
Maize	8.2±0.9	9.2±1.8	112±20
Potato	8.1±1.1	8.1±2.7	91±26

Relative yield and STP were prominently clustered within the year-by-site combinations and varied strongly among them (Supplementary Fig. 2). Mean relative yield ranged from 72 to 141% for years and from 91 to 111% for sites. Similarly, mean $STP_{H2O\text{-}CO2}$ and STP_{AAE10}, respectively, ranged from 0.35 to 1.55 mg kg^{-1} and 16 to 56 mg kg^{-1} for years and from 0.38 to 1.05 mg kg^{-1} and 20 to 50 mg kg^{-1} for sites. The larger variation in relative yield and STP among years than sites determined the data structure for further analyses.

3.2. Pedoclimatic variables

Among the tested pedoclimatic variables only annual temperature and soil pH were significantly inversely related to relative yield independent of the crop (Supplementary Table 3). Clay content affected relative yield of maize negatively and sunshine hour deviation affected the relative yield of potato positively. Silt content, annual precipitation, and precipitation deviation were each related to relative yield of two crops but in opposite directions (Supplementary Table 3). As annual sunshine hours and temperature deviation were not meaningful for yield response, they were excluded from further analyses.

The remaining pedoclimatic variables had similar weights on two principal components and explained roughly 65% of the variation (Supplementary Fig. 3). The clustering of the principal component scores revealed strong similarities in pedoclimatic conditions among the sites ALT, ELL, OEN, and REC, while GRA differed slightly and CAD strongest from the other sites. Diverging loading vectors for several variables suggested strong negative correlations between those (Supplementary Fig. 3), which was confirmed by the Pearson correlation analysis (Supplementary Fig. 4). We considered the high correlations between clay and silt content as well as precipitation deviation and sunshine hour deviation as causal, whereas the correlation between e.g. clay content and annual temperature as coincidental. Therefore, we limited the choice of pedoclimatic variables to soil pH, clay content, annual temperature, annual precipitation, and precipitation deviation to reduce the number of highly correlated covariates in the yield response models (Fig. 2).

3.3. Model fit

Convergence was obtained for all models; however, parameter estimation was not equally straightforward. Depending on the starting values, the *intercept* and *slope* parameters of the linear-plateau model varied strongly but the AICs were largely similar (data not shown). In contrast, the parameters of the Mitscherlich model were largely

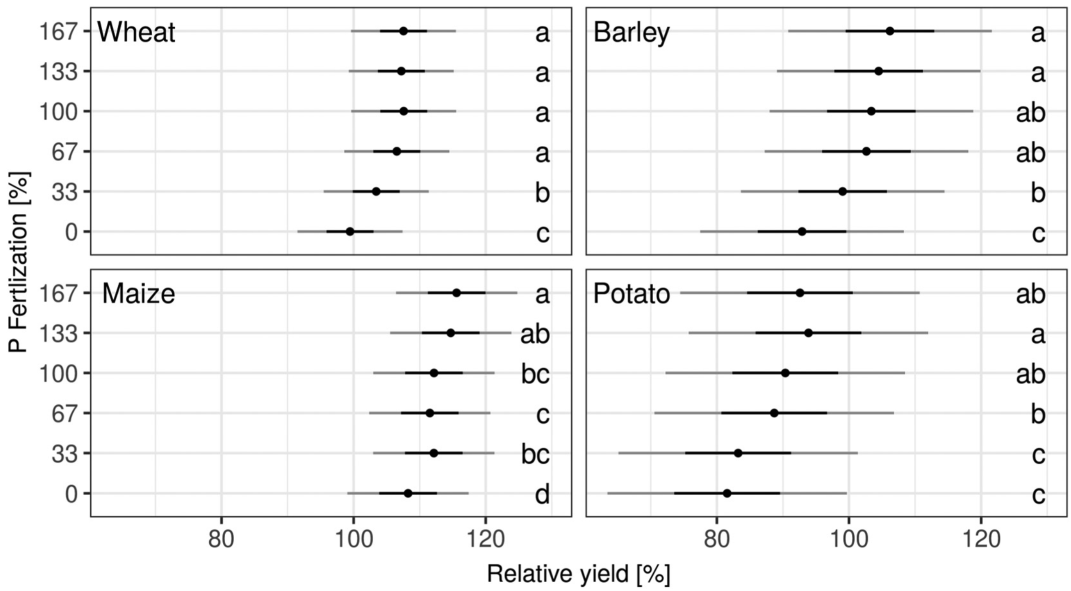


Fig. 1. Mean relative yields (proportion of mean yields in Swiss agricultural practice) of wheat, barley, maize, and potato in six P fertilization treatments during 26 years at six sites. Estimated marginal mean ± standard error (black) and 95% confidence intervals (grey) of relative yield retrieved from linear mixed models with year, site, and block as nested random effects. Different letters denote significant differences at $p < 0.05$ between treatments per crop (for linear contrasts of treatment pairs see Supplementary Fig. 3).

independent of the starting values and therefore unambiguous. While the linear-plateau and Mitscherlich functions returned similar goodness of fit measures, the random effects structure of the quadratic model for STP_{AAE10} did not cover the variance in the data at the year and site levels adequately (Supplementary Fig. 5).

Hence, we selected the Mitscherlich function for further analyses. The inclusion of random effects improved the fit considerably from, on average, $R^2 = 0.24$ at population level to $R^2 = 0.88$ at site level. The random effects had zero mean and evenly distributed standard deviations among the year, site, and unit levels of, on average, 6, 19, and 8%, respectively (Supplementary Fig. 5). Crop-specific parameter intercepts (based on the reference level of 'crop') differed significantly among crops (Table 5), entailing distinctly shaped response curves for each crop (Fig. 3). *Asymptote* was in good agreement with the respective average yield measured in the highest yielding treatment and the variation among and within *asymptotes* was small. By contrast, *rate* and *environment* varied considerably among crops and were generally contrary to each other, i.e. low *rates* were associated with large *environments* (e.g. maize) and vice versa (e.g. barley; Table 5). Compared to $STP_{H2O\text{-}CO2}$, yield response to STP_{AAE10} was less steep and shifted towards larger *environments* relative to STP (Fig. 3).

3.4. Importance of covariates

The coefficients of the covariates for *asymptote*, *rate*, and *environment*, respectively, ranged on average from −9–2%, −1.0–1.0 $(\text{mg kg}^{-1})^{-1}$, and 0.00–0.15 mg kg^{-1} for $STP_{H2O\text{-}CO2}$ and from −9–2%, −0.01–0.01 $(\text{mg kg}^{-1})^{-1}$, and −12–14 mg kg^{-1} for STP_{AAE10} (Fig. 4). All covariates but precipitation deviation had significant effects on at least two parameters and the CAT scores were mostly associated with effect size, with few exceptions. Effect variability differed most prominently between covariates for *asymptote*, being smallest for fertilization and largest for the climate variables. Effect direction and size were

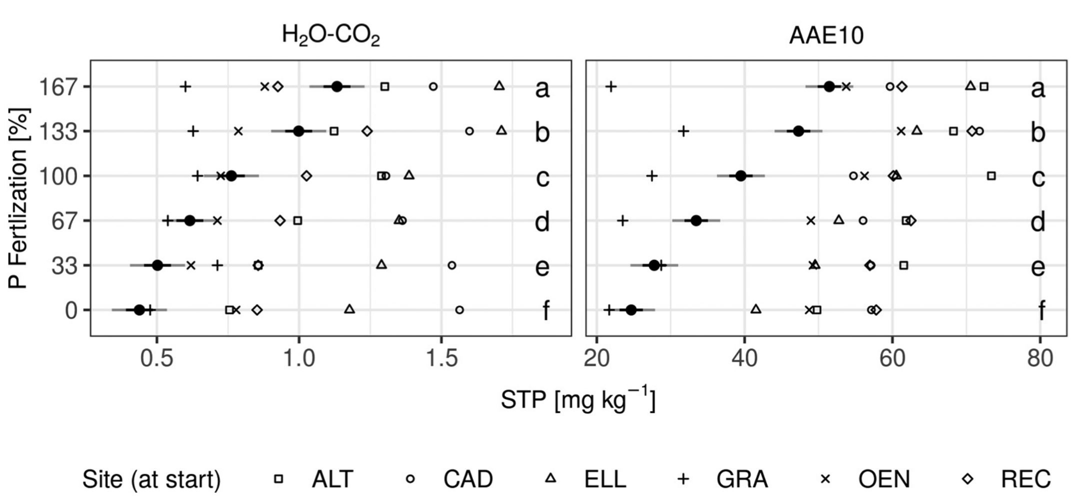


Fig. 2. Mean $STP_{H2O\text{-}CO2}$ and STP_{AAE10} in six P fertilization treatments during 26 years at six sites (solid points) and initial $STP_{H2O\text{-}CO2}$ and STP_{AAE10} at each site (open shapes). Estimated marginal mean ± standard error (black) and 95% confidence intervals (grey) of STP retrieved from linear mixed models with year, site, and block as nested random effects. Different letters denote significant differences at $p < 0.05$ between.

Table 5
Mitscherlich parameters for asymptote, rate, and environment of linear mixed models containing wheat, barley, maize, and potato, respectively, as reference level of 'crop' with year, site, and block as nested random effects. Estimated mean ± standard error. Bold font indicates significant difference from zero. Different letters denote significant differences in asymptote, rate, and environment, respectively, between crops based on contrasting parameter and coefficient estimates.

Crop	$STP_{H2O\text{-}CO2}$			STP_{AAE10}		
	Asymptote [%]	Rate [(mg kg^{-1})$^{-1}$]	Environm. [mg kg^{-1}]	Asymptote [%]	Rate [(mg kg^{-1})$^{-1}$]	Environm. [mg kg^{-1}]
Wheat	**103 ± 4** b	**4.3 ± 1.0** b	**0.68 ± 0.21** a	**104 ± 4** ab	**0.04 ± 0.01** b	**99 ± 28** a
Barley	**103 ± 5** ab	**7.7 ± 1.7** a	0.12 ± 0.10 b	**104 ± 6** ab	**0.06 ± 0.01** a	**46 ± 16** b
Maize	**116 ± 5** a	**2.6 ± 0.5** c	**1.22 ± 0.33** a	**115 ± 4** a	**0.02 ± 0.01** c	**204 ± 69** a
Potato	**95 ± 5** b	**4.9 ± 0.9** b	0.18 ± 0.10 b	**95 ± 5** b	**0.05 ± 0.01** a	**43 ± 15** b

similar for both extraction methods, except for the effect of soil pH on *environment* (Fig. 4).

Fertilization was positively related to *asymptote* and *environment* and negatively to *rate*, although partly not significantly. Clay content and soil pH were negatively related to *asymptote* and positively to *environment* and *rate*, respectively. In addition, soil pH had a strong negative relation to *environment* for STP_{AAE10} but not for $STP_{H2O\text{-}CO2}$. Annual temperature affected all parameters, *asymptote* negatively and *rate* and *environment* positively. Annual precipitation was negatively related to *rate* and positively to *environment*, while precipitation deviation had a negative effect on *rate* only.

3.5. Critical soil test phosphorus

Critical STP was generally low but varied considerably among years and sites (Fig. 5; Supplementary Table 4). Mean concentrations for critical $STP_{H2O\text{-}CO2}$ at 90, 95, and 98.5% of maximum predicted yield, respectively, were 0, 0.15, and 0.45 mg kg^{-1} for wheat, 0.24, 0.33, and 0.48 mg kg^{-1} for barley, 0, 0.09, and 0.59 mg kg^{-1} for maize, and 0.41, 0.58, and 0.87 mg kg^{-1} for potato (Fig. 5). Mean concentrations for critical STP_{AAE10} were 0, 10, and 48 mg kg^{-1} for wheat, 13, 25, and 47 mg kg^{-1} for barley, 0, 0, and 11 mg kg^{-1} for maize, and 19, 36, and 63 mg kg^{-1} for potato (Fig. 5). Critical concentrations at 95% of maximum predicted yield are given in Supplementary Table 4 for all years and in Supplementary Fig. 8 for selected years.

Critical STP was independent of *asymptote* and sensitive only to changes in *rate* and *environment* along the 10 to 90% quantile ranges (Fig. 6). While the relation was negative-linear for *environment*, it was negative-exponential for *rate*, i.e. the decrease in critical STP with increasing *rate* was more pronounced at low *rates*. Critical STP was similarly sensitive to changes in model parameters for all crops except maize, whose critical concentrations were particularly sensitive to changes in *rate* (Fig. 6). However, they became positive only at very low *rates*.

A change in coefficients altered critical STP concentrations of barley and potato more strongly than those of wheat and maize as 95% of maximum predicted yield was often already achieved at 0 STP for wheat and maize (Figs. 7–10). The most prominent change in critical STP was induced by annual temperature: 0–0.33, 0.15–0.44, 0–0.80, and 0.23–0.70 mg kg^{-1} $STP_{H2O\text{-}CO2}$ and 0–40, 0–44, 0–78, and 4–56 mg kg^{-1} STP_{AAE10} for wheat, barley, maize, and potato, respectively (Figs. 7–10). Changes in soil clay content also changed the critical concentrations for wheat, barley, and potato considerably, which ranged between 0 and 0.24, 0.20 and 0.50, and 0.35 and 0.66 mg kg^{-1} $STP_{H2O\text{-}CO2}$, respectively, and 0 and 17, 11 and 41, and 18 and 49 mg kg^{-1} STP_{AAE10}, respectively (Figs. 7, 8, and 10). Changes in fertilization and soil pH, respectively, were also associated with pronounced changes in critical STP_{AAE10} of 0–22 and 11–31 mg kg^{-1} for barley and 0–30 and 20–38 mg kg^{-1} for potato (Figs. 8 and 10). These relations were less prominent for critical $STP_{H2O\text{-}CO2}$ (Figs. 8 and 10).

4. Discussion

4.1. Effect of long-term phosphorus fertilization on yield and soil phosphorus

Both relative yield and STP are clearly affected by long-term phosphorus fertilization. Treatment effects, however, do not run parallel for the two variables: Differences in yield are limited to underfertilization and vanish under norm to overfertilization, where differences in STP are most prominent. With increasing fertilization rates, the proportion of recent fertilizer P utilized by plants decreases, while accumulated

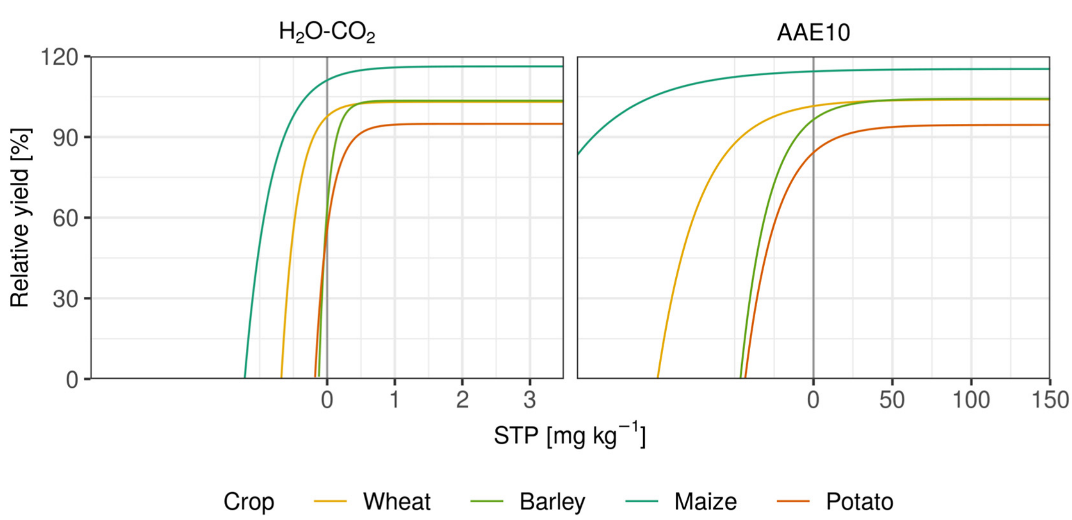


Fig. 3. Mitscherlich curves representing yield response of wheat, barley, maize, and potato to STP. Curves are based on model parameters excluding fixed effects of covariates (corresponding to values in Table 5).

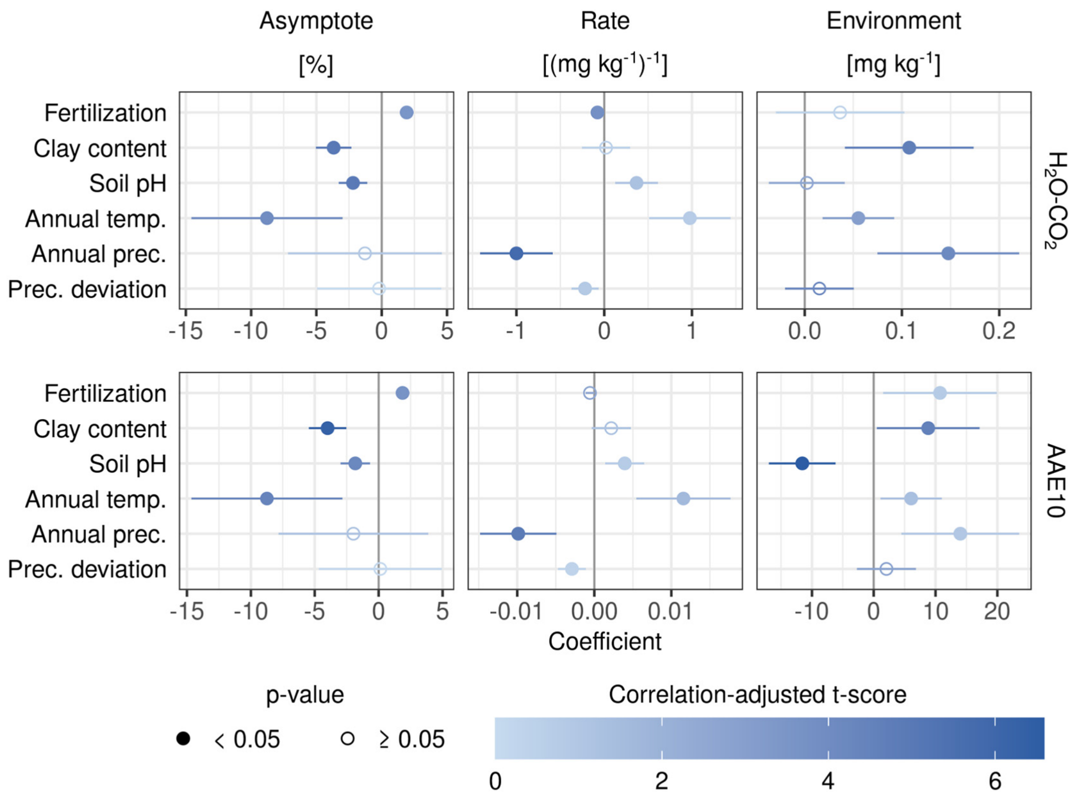


Fig. 4. Effects of scaled covariates on the Mitscherlich parameters asymptote, rate, and environment for yield response of wheat, barley, maize, and potato to STP and their significance excluding (p-value) and including (CAT score) adjustment for auto-correlation among covariates. Estimated mean ± 95% confidence intervals.

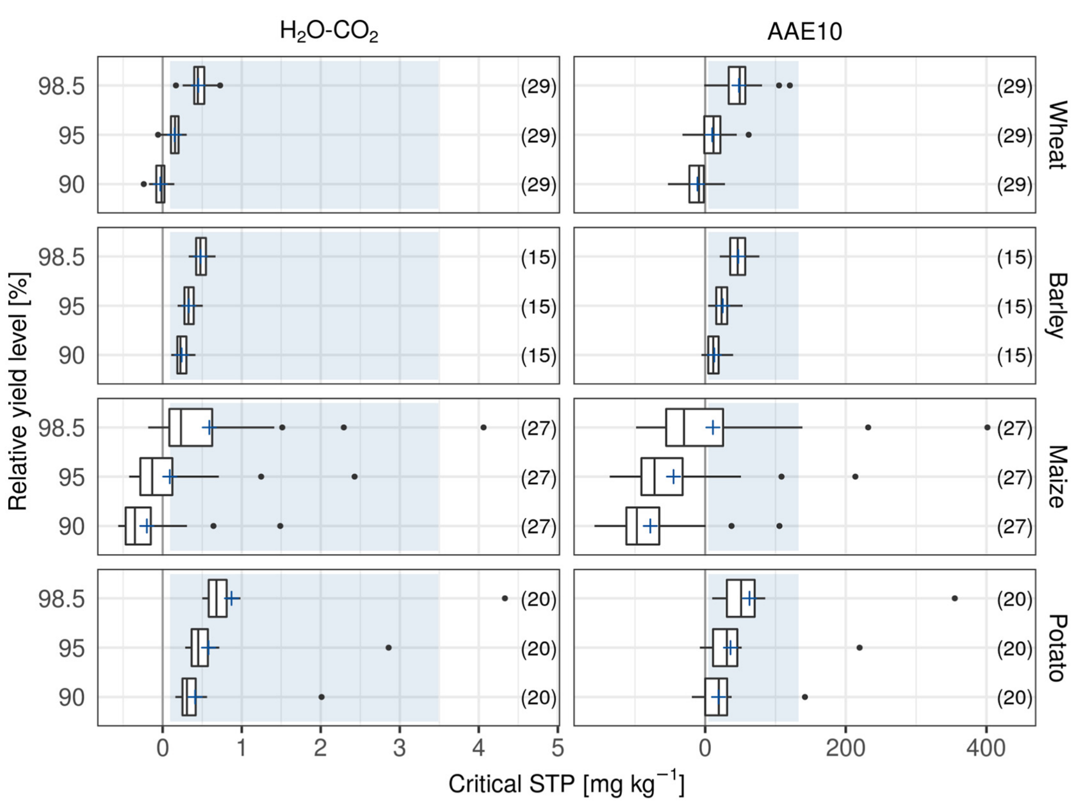


Fig. 5. Critical $STP_{H2O\text{-}CO2}$ and STP_{AAE10} for wheat, barley, maize, and potato at 90, 95, and 98.5% of maximum predicted yield. Boxplots depict ranges of critical STP for all combinations of year and site and shadings mark ranges of measured STP concentrations. Numbers in brackets are numbers of observations and crosses are mean values.

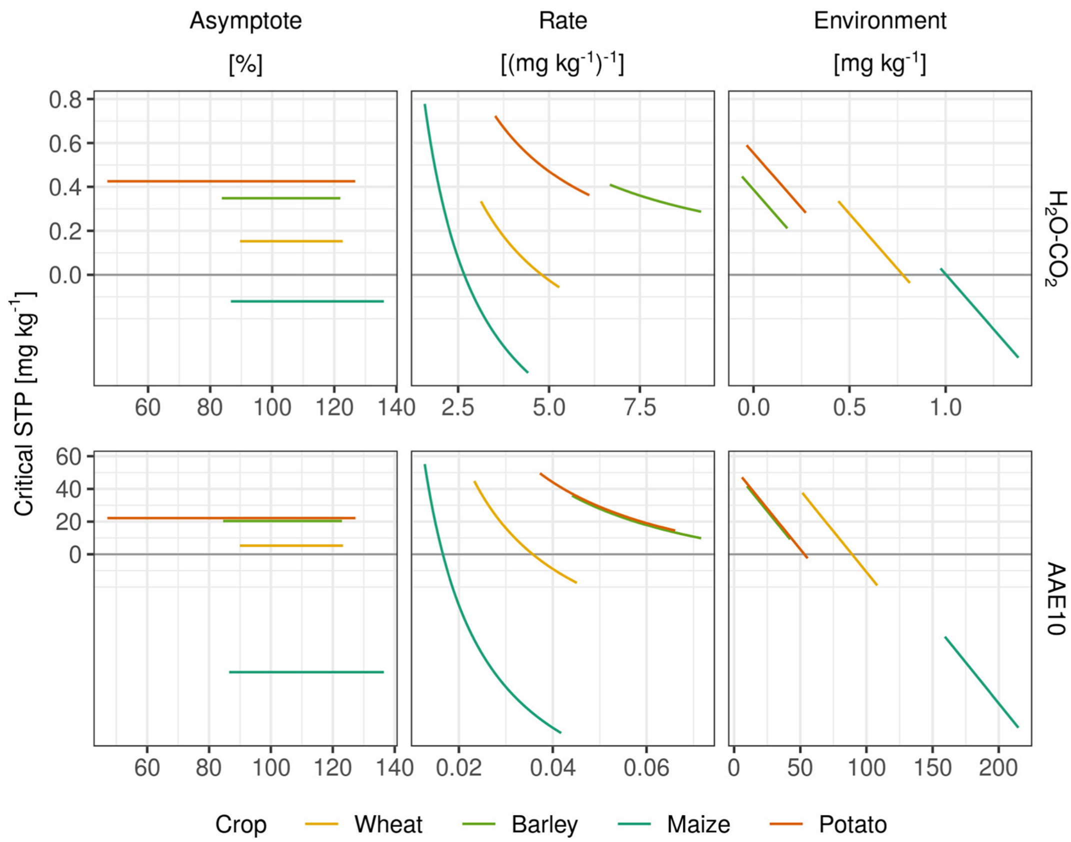


Fig. 6. Sensitivity of critical STP at 95% of maximum predicted yield to changes in model parameters when two parameters are kept at their median and the third one is varied along its 10 to 90% quantile range among all groups of sites within years.

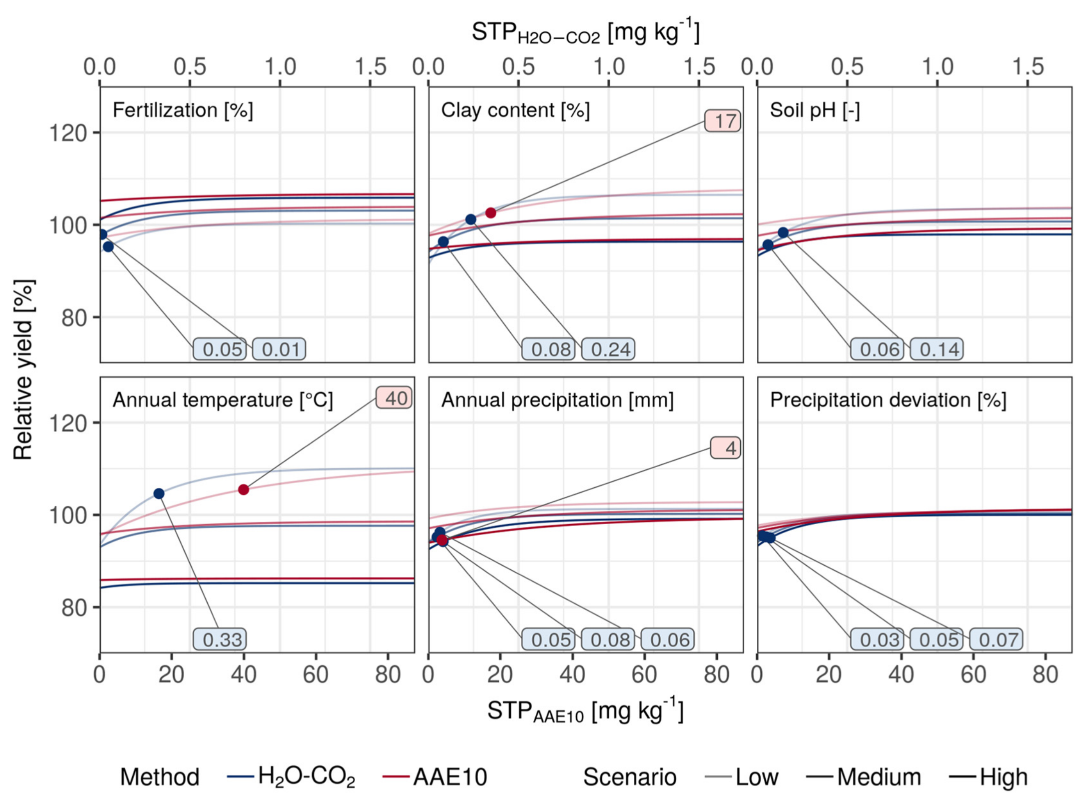


Fig. 7. Critical STP concentrations at 95% maximum predicted yield of wheat (numbers in boxes) and their sensitivity to changes in coefficients of model covariates. All coefficients but one are kept at their median (except P fertilization: minimum = 0% P input). Low, medium, and high scenarios of the remaining coefficient correspond to its 10, 50, and 90% quantiles, respectively. Only critical STP concentrations equal to or greater than zero are shown.

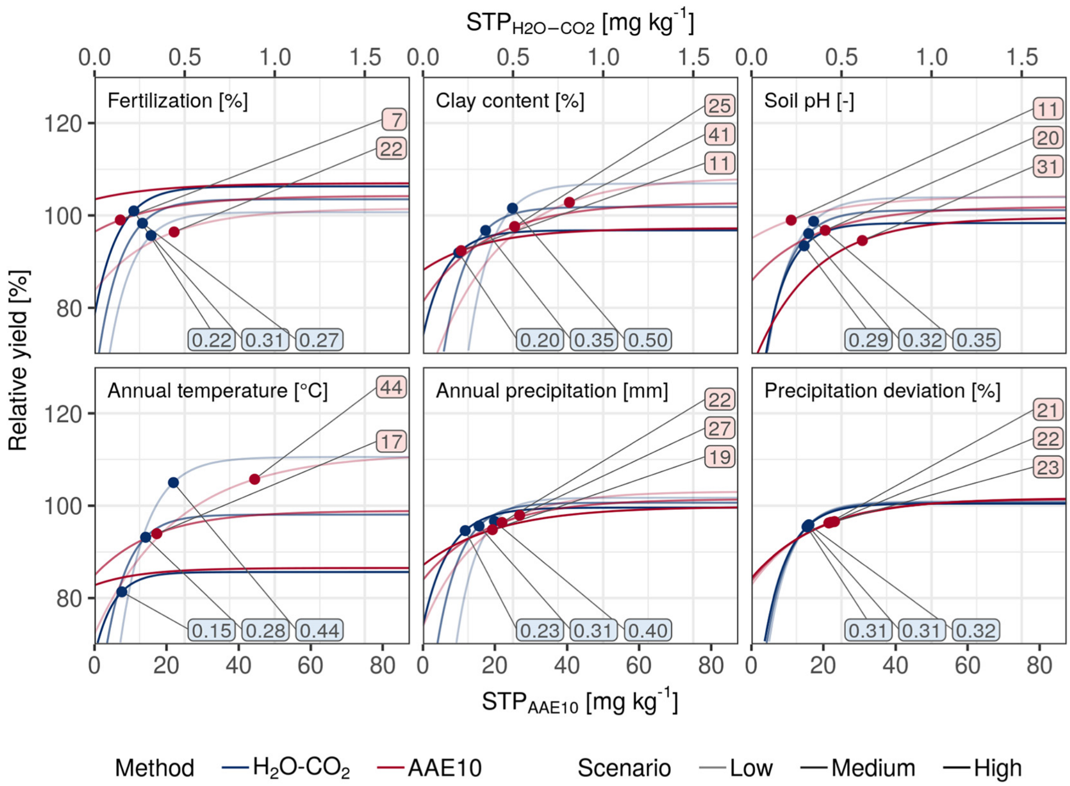


Fig. 8. Critical STP concentrations at 95% maximum predicted yield of barley (numbers in boxes) and their sensitivity to changes in coefficients of model covariates. All coefficients but one are kept at their median (except P fertilization: minimum = 0% P input). Low, medium, and high scenarios of the remaining coefficient correspond to its 10, 50, and 90% quantiles, respectively. Only critical STP concentrations equal to or greater than zero are shown.

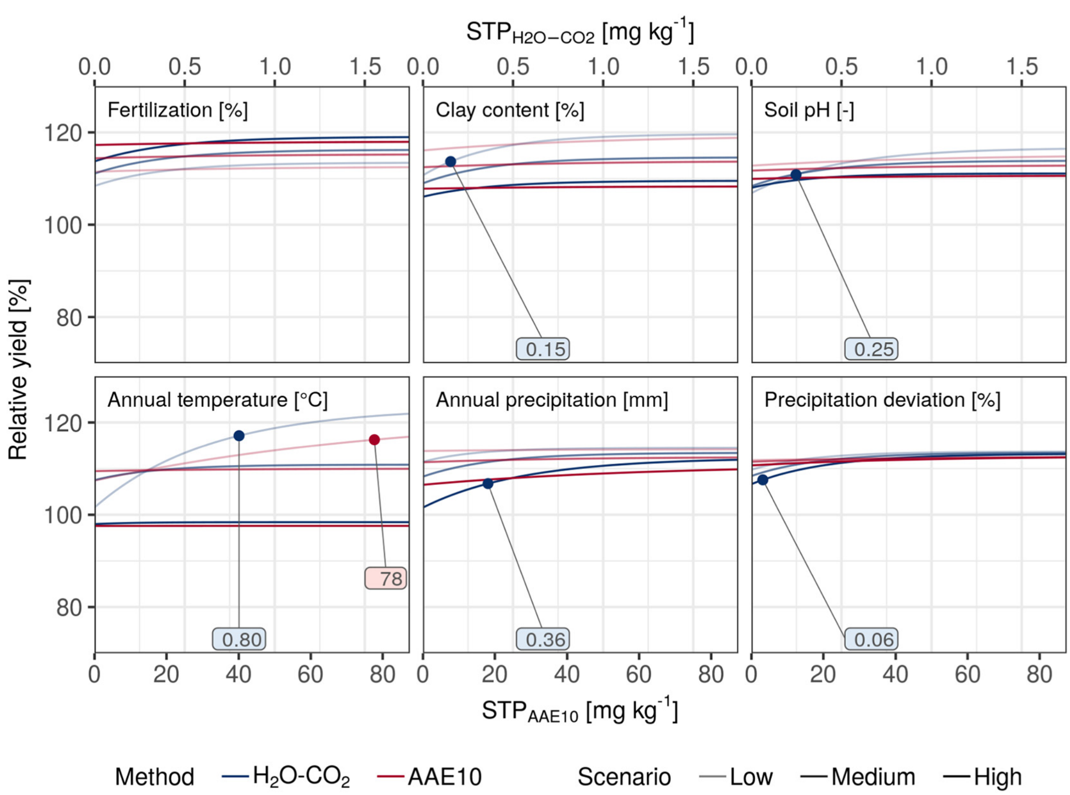


Fig. 9. Critical STP concentrations at 95% maximum predicted yield of maize (numbers in boxes) and their sensitivity to changes in coefficients of model covariates. All coefficients but one are kept at their median (except P fertilization: minimum = 0% P input). Low, medium, and high scenarios of the remaining coefficient correspond to its 10, 50, and 90% quantiles, respectively. Only critical STP concentrations equal to or greater than zero are shown.

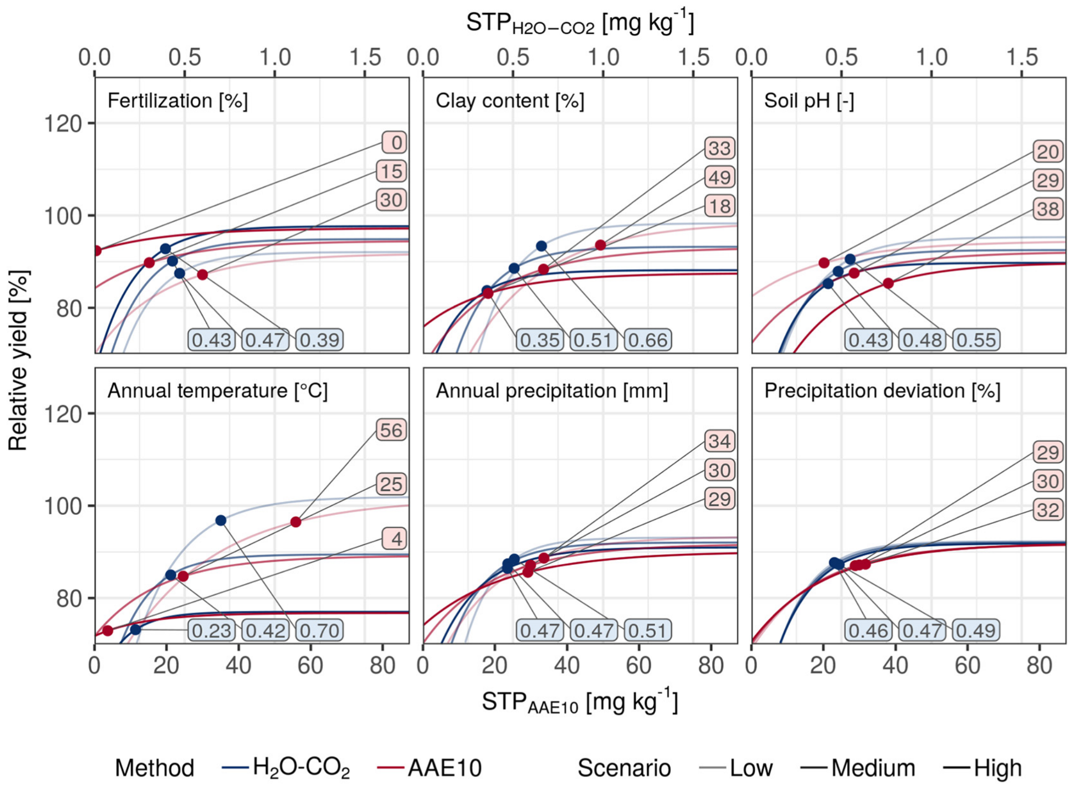


Fig. 10. Critical STP concentrations at 95% maximum predicted yield of potato (numbers in boxes) and their sensitivity to changes in coefficients of model covariates. All coefficients but one are kept at their median (except P fertilization: minimum = 0% P input). Low, medium, and high scenarios of the remaining coefficient correspond to its 10, 50, and 90% quantiles, respectively. Only critical STP concentrations equal to or greater than zero are shown.

soil P increases (Johnston and Richards, 2003; Kirkby and Johnston, 2008), building up soil P reserves (Syers et al., 2008; Yli-Halla, 2016). While maize is the only crop with significant yield increases also above norm P fertilization, the yield difference of 3% compared to surplus P fertilization is small and similar to that of barley and potato. Differences in statistical significance between crops can rather be ascribed to differences in number of data points and yield variability than differences in actual treatment effects. Crop demand for P seems to be largely met around norm P fertilization, suggesting that the availability of other resources, possibly N, determines yield response beyond norm P fertilization (Engels et al., 2012).

4.2. Suitability of models

The tested models accommodate classical nonlinear regressions of STP and yield, concomitant effects of pedoclimatic conditions, and multilevel data structures. While covariates have frequently been included in yield response models (Colwell et al., 1988; Mallarino and Blackmer, 1992; Kuchenbuch and Buczko, 2011; Buczko et al., 2018), the temporal and spatial grouping of data in multivariate nonlinear models has only recently been accounted for in agronomic studies (Gonçalves et al., 2016; Parent et al., 2017; Sari et al., 2018). Combined experiments require mixed models that integrate the data structure in random effects terms (Loughin, 2006; Moore and Dixon, 2015). Here, we consider *year* as 'global' source of yield variation, mainly affected by the seasonal variation in climate among years, and *site* as subordinate source, mainly affected by soil and microclimatic characteristics. The redundancy of *block* as random effects level suggests that experimental designs without true field replicates could be suitable for soil test calibration studies when long time series (Loughin, 2006; Payne, 2015) and diversified sites (Casler, 2015) provide seasonal and spatial variation.

The linear-plateau and Mitscherlich functions yield better model fits than the quadratic function, supporting previous studies (Mallarino and Blackmer, 1992; Hochmuth et al., 1993; Valkama et al., 2011). Yield decrease at high STP implied by the quadratic function is not reflected in the data within the calibration range of our study, which is in line with other findings (White and Hammond, 2008; Buczko et al., 2018). Additionally, the use of the quadratic function for STP calibration cannot be recommended because of the inadequate representation of the plateau at maximum yield and the overestimation of fertilizer requirements (Colwell et al., 1988; Hochmuth et al., 2011). A similar goodness-of-fit for the linear-plateau and Mitscherlich functions has also been shown by others (Dodd and Mallarino, 2005; Tang et al., 2009; Wu et al., 2018). However, the more straight-forward process of parameter estimation is the ultimate indication for the Mitscherlich function in our study.

4.3. Crop response to soil test phosphorus

The large dataset of 26 years of differential P fertilization provides an adequate basis to describe yield response to STP. Using the same data from the first nine years of the experiment, Gallet et al. (2003) could not find a suitable model to relate crop yield to STP, which underlines the long-lasting effect of accumulated soil P reserves (Yli-Halla, 2016).

In our study, the variation in crop responsiveness to STP is striking. The curves of maize and wheat are already close to their *asymptotes* at zero STP, while barley requires a larger and potato the largest increase in STP to reach their *asymptotes*. This is consistent with the observed yield increase of 6% each for wheat and maize and 11% each for barley and potato between zero P fertilization and the highest-yielding treatment. Meta-analyses of data from European P fertilization trials reveal similar differences in responsiveness to extractable soil P between crops with potato > barley > wheat but findings for maize are contradictory (Nawara et al., 2017; Buczko et al., 2018).

Potato is much less efficient than graminoids to utilize soil P because of its relatively poor architectural and physiological root characteristics (Hopkins et al., 2014) that both govern plant P uptake (Hinsinger, 2001; Lynch, 2011; Marschner and Rengel, 2012; van de Wiel et al., 2016). The entire root system including root length and root surface area is considerably smaller (Stalham and Allen, 2001; Yamaguchi, 2002), while the roots exude smaller amounts of organic acids into the rhizosphere (Wang et al., 2015) and are often less densely colonized by mycorrhizal fungi due to intensive soil management (McGonigle et al., 1999; Hopkins et al., 2014). Hence, potato plants are impaired to intercept P in the soil solution and to mobilize adsorbed P by chemical changes of the rhizosphere compared to wheat, barley, and maize.

Despite wheat and barley being alike in P removal (Sinaj et al., 2017) and rooting characteristics (Nobile et al., 2019), the two crops respond differently to STP. This may be linked to differently pronounced mechanisms of yield response or compensation strategies (McDonald et al., 2018). While P acquisition efficiency (net total P uptake per unit available P) seems to be high and not decisive for yield response in wheat, it shows a strong genetic variation and effect on yield in barley (Gahoonia and Nielsen, 1996; McDonald et al., 2018). Hence, depending on barley genotype, P uptake and consequently yield formation may be more constrained in barley than wheat under low P availability.

For maize, Nawara et al. (2017) found a similarly low responsiveness to STP, whereas Buczko et al. (2018) reported the opposite. The coarse root system and more rapid biomass increase during vegetative growth of maize compared to wheat would support a higher demand for soil P (Smolders et al., 2020). In our study, the very high yield relative to the Swiss average, however, indicates that growth conditions for maize may be exceptionally favourable, potentially attenuating yield response to STP. In addition, numerous other factors that we did not study may govern crop responsiveness to STP, such as rotational effects of legumes (Oberson et al., 2011) or the contribution of subsoil P reserves to crop uptake (Syers et al., 2008; Buczko et al., 2018).

4.4. Response curves for H_2O-CO_2 and AAE10

The Mitscherlich parameters *rate* and *environment* additionally depend on the extraction method: the response curves for $STP_{H2O\text{-}CO2}$ are steeper and shifted towards smaller *environments*, i.e. closer to the origin, than those for STP_{AAE10}. This can be attributed to the different P fractions that are measured with different methods (Neyroud and Lischer, 2003; Shwiekh et al., 2015). In accordance, Holford et al. (1985) found steeper response curves for weaker extractants and additionally linked the curvature (*rate*) to the effectiveness of the extraction method for predicting P fertilizer requirements. A response curve whose x-intercept is not significantly different from zero, i.e. close to the origin, provides a good representation of utilized crop P by STP, as in the case of barley and potato yield response to $STP_{H2O\text{-}CO2}$. This suggests that H_2O-CO_2 as extractant can generally provide better estimations of crop P requirements than AAE10.

4.5. Improvement of yield response models by pedoclimatic variables

Fitting one model to the entire dataset postulates similar pedoclimatic effects on model parameters for different crops but also facilitates the identification of general relations. We observed strong effects of several soil and climate variables on all parameters, supporting findings by Parent et al. (2017) for potato yield response to N fertilization in 93 Canadian trials. Compared to their study, we found a two-times larger impact of annual temperature and precipitation on the *rate* parameter, which might be attributed to the small dataset but large gradient in climate variables among years and sites in our data.

An increase in both clay content and soil pH is associated with lower predicted yield maxima due to their prominent effects on yield formation. Increasing clay content additionally shifts the response curves towards more negative STP values, i.e. larger *environments*, reflecting an increase in the discrepancy between soil P extractability and actual crop P utilization (Flisch et al., 2017). However, clay content is often correlated with soil organic carbon (Kleber et al., 2015), which has repeatedly been shown to have a strong impact on crop response to STP due to its effect on soil structure and, hence, penetrability and interception of P by roots (Johnston et al., 2014; Schneider et al., 2019). It is therefore not possible to identify the distinct source of effects on the *environment* parameter in this study, which would need further investigation of soils with diverging clay and soil organic carbon contents.

The effect of soil pH on *asymptote* and *rate* is similar for the two extraction methods, whereas its strong effect on *environment* appears only for STP_{AAE10} but not $STP_{H2O\text{-}CO2}$. This highlights the sensitivity of the AAE10 method to the presence of carbonates (Stuenzi, 2006b). In contrast to H_2O-CO_2, the strongly acidic AAE10 also breaks stable calcium compounds and releases large amounts of a phosphate pool that is only fractionally utilized by plants, hence, shifting the response curve towards larger STP_{AAE10} values. The AAE10 method is therefore not recommended for soils with calcium concentrations > 4 g kg^{-1} (pH > 6.8) in the current Swiss fertilization recommendation guidelines (Flisch et al., 2017). However, our results strongly suggest that yield response to STP_{AAE10} can successfully be modelled also for soils with pH > 6.8, when the horizontal position of the response curve is accordingly fitted. The calcium concentration, which can routinely be measured in AAE10 extracts, might give even more accurate information on the effect of the extractant on P recovery (Stuenzi, 2006b).

The marked negative relation of annual temperature to maximum predicted yield is in line with studies on the continental and global scale (Lobell and Field, 2007; Peltonen-Sainio et al., 2010). Elevated temperature causes yield reductions through shortened maturation and incomplete grain filling in cereals (Peltonen-Sainio et al., 2010) and secondary tuberization in potato (Rykaczewska, 2015). High temperature per se is unlikely to impair maize yields in Switzerland (Holzkämper et al., 2015) but concomitant soil water deficits and increased pest infestation during summer (Swiss Academies of Arts and Sciences, 2016) may also cause yield reductions. The strong positive effect of annual temperature on *rate* and *environment* entails a decrease in crop responsiveness to STP with increasing temperature, which is possibly linked to increased mineralization and availability of organically bound P but also overall lower P demand due to considerably lower crop biomass production.

Annual precipitation and precipitation deviation do not clearly affect maximum predicted yield despite accounting for a large part of the variation in *asymptote*. This could be due to an interaction with crop as the two variables are positively related to barley yields but negatively to potato yields. However, crop responsiveness to STP, as reflected by the combination of *rate* and *environment*, is strongly affected by annual precipitation. This highlights the impact of soil water on P availability and uptake by crops (Kirkby and Johnston, 2008). The small effect size of precipitation deviation compared to annual precipitation suggests only a minor importance of weather conditions during juvenile crop development for relative yield (Dahnke and Olson, 1990), which might differ when the response variable is absolute yield (Parent et al., 2017).

4.6. Critical soil test phosphorus

Our findings underline that critical STP strongly depends on crop, extraction method, and the defined percentage of maximum predicted yield. The large differences in mean critical concentrations among crops (0.09–0.58 mg kg^{-1} $STP_{H2O\text{-}CO2}$; 0–36 mg kg^{-1} STP_{AAE10}) at 95% of maximum predicted yield are representative of the variation in crop response to STP in general. A wide range of crop-specific critical concentrations was found in numerous studies (Johnston, 2005; Colomb et al., 2007; Cadot et al., 2018; Sucunza et al., 2018; Smolders et al., 2020), although similar values were also shown for wheat and maize (Tang et al., 2009; Wu et al., 2018) and wheat and barley (Bell et al., 2013).

Only few authors have reported critical concentrations for the H_2O-CO_2 and AAE10 methods so far. In an early work in Switzerland, Peyer and Frei (1971) studied different cereals over six years on six sites and derived critical $STP_{H2O\text{-}CO2}$ concentrations of ~0.9–1.2 mg kg^{-1}. Recently, Cadot et al. (2018) reported mean critical $STP_{H2O\text{-}CO2}$ concentrations for wheat and maize over 44 years on one Swiss site of ~0.3–0.4 mg kg^{-1}. Our results (cereals: ~0.1–0.3 mg kg^{-1}) are considerably lower than the values of the short-term study and match better with the long-term analysis. This indicates that the relation of crop P uptake to STP is likely to differ between accumulation and depletion dynamics of soil P (Maguire et al., 2005). Reported critical STP_{AAE10} concentrations for cereals are less variable among studies: 13–22 mg kg^{-1} in Switzerland (Cadot et al., 2018), 26–66 mg kg^{-1} in Belgium (Genot et al., 2011), and 17 mg kg^{-1} in Texas (pot experiment; Hons et al., 1990). The use of different response functions and site- and management-specific growth conditions may contribute to the differences (Mallarino and Blackmer, 1992).

Despite physiological differences in crop responsiveness to STP, crop-specific critical concentrations may be impractical for use in crop rotations (Nawara et al., 2017). Only a small proportion of utilized P by the crop is recent fertilizer P and the remainder originates mainly from previous P applications (Syers et al., 2008; Yli-Halla, 2016). In the long term, even single fertilizer applications once per rotation can be as efficient as annual fertilizer applications (Barber, 1980), suggesting that weighted mean values for the entire rotation rather than crop-specific critical STP concentrations can provide an adequate basis for fertilization recommendations.

The independence of critical STP of *asymptote* supports previous observations. Similar critical STP concentrations have been reported for cereals in different years although yields depended strongly on weather conditions and N supply (Johnston, 2005; Johnston et al., 2014). Even when economic aspects of yield response to nutrient supply are considered, a change in *asymptote* has only a minor effect on critical concentrations as compared to *rate* and *environment* (Parent et al., 2017). This is highly relevant for the choice of covariates for response models as they are commonly selected based on their effect on maximum yield (Colwell et al., 1988; Parent et al., 2017). However, focusing on critical STP, pedoclimatic variables should be chosen based on their influence on the other model parameters.

Recommendations of surplus, replacement, and zero fertilization as defined by Morel et al. (1992) are based on critical STP at 90, 95, and 98.5% of maximum predicted yield, respectively. The practicality of this approach is essentially governed by the curvature of the response curve, as it determines the difference in critical STP between 90 and 98.5% yield. For instance, the high *rate* of barley entails a narrow range of critical $STP_{H2O\text{-}CO2}$ between 90 and 98.5% yield of 0.24–0.48 mg kg^{-1}, whereas the low *rate* of maize results in a wide range of 0–0.59 mg kg^{-1}. When differences in critical STP between 90 and 98.5% yield become very small, analytical uncertainties may invalidate those differences. When they become very large, the expected yield increase may not justify economic and ecological downsides of high soil P (Mallarino, 2012).

Related to the Swiss fertilization guidelines PRIF (Flisch et al., 2017), the mean critical concentrations at 95% of maximum predicted yield for soils with a clay content corresponding to the mean clay content in this study (26%) are within the ranges of the soil P fertility classes A (wheat, maize) and B (barley, potato) for $STP_{H2O\text{-}CO2}$ and A (maize), B (wheat, barley), and C (potato) for STP_{AAE10}. Class C represents an adequate soil P status, while classes A and B refer to deficient and moderate soil P, respectively. Consequently, the Swiss P fertilization guidelines would recommend elevated P fertilization at 1.3–1.5 times of crop P removal for $STP_{H2O\text{-}CO2}$ and 1.0–1.5 times for STP_{AAE10}. Moreover, potato and maize are considered to be P sensitive crops and, according to the PRIF, an additional correction factor of 1.2 would be applied for calculating their definite P demands (Sinaj et al., 2017). Hence, the PRIF may overestimate fertilizer P requirements based on STP when compared to the mean critical STP concentrations found in this study, which adds to the increasing evidence that crop yields can be maintained at STP concentrations lower than those currently adopted in fertilization guidelines (Withers et al., 2014; Macintosh et al., 2019).

Due to the large variability in pedoclimatic conditions, mean critical STP is not necessarily representative for many settings (Syers et al., 2008). Similar to our study, Recena et al. (2016) and Johnston et al. (2014) reported a strong variation in critical STP with varying soil properties such as pH, clay content, and organic carbon content. In the PRIF, the pronounced negative relation of critical STP to clay content is taken into account through differentiation by five clay content classes (Flisch et al., 2017). Yet, climate conditions seem to govern critical STP to similar extent but are currently not accounted for. For instance, low temperature entails higher critical concentrations in our study, which correspond to elevated soil P according to the PRIF (Flisch et al., 2017). Fertilizer P requirements would thus be underestimated on sites with lower mean annual temperature such as in mountainous regions. Depending on site conditions, those effects may accumulate and enhance the discrepancy between fertilization recommendations and actual crop P demand.

As *rate* and *environment* differ between STP extraction methods, so do the critical concentrations. The critical concentrations are less variable among crops and pedoclimatic scenarios for $STP_{H2O\text{-}CO2}$ than for STP_{AAE10}. This highlights the need for validation of method-specific fertility classes rather than conversion factors of STP concentrations between methods.

4.7. Perspectives

Long-term experiments including a zero P fertilization treatment usually provide the platform for STP calibration (Jordan-Meille et al., 2012). Some issues with this approach comprise the confounded relation of STP to time, the resource intensity over long time periods, or the rather little relevance of long-term zero P fertilization for agricultural practice. With ongoing depletion of total P, initially fixed residual P gets mobilized and adds to the extractable soil P pool (Zhang et al., 2004; Bai et al., 2013). This requires repeated measurements over long time periods to estimate crop response to STP. In addition, long-term zero P fertilization is an unrealistic scenario in agricultural practice, which might question the applicability of critical STP derived from a zero P fertilization treatment. Hence, the need for alternative methods to estimate soil P availability and derive critical concentrations becomes increasingly obvious (van Maarschalkerweerd and Husted, 2015) and spectroscopic techniques are likely to fill this gap in the near future (Cade-Menun, 2017; Pätzold et al., 2019).

In addition to agronomic optima of STP, economic considerations are often included in yield response models. Hence, a monetary value must be assigned to STP (Morris et al., 2017; Sihvonen et al., 2018), which has recently been expanded from mere fertilizer costs for the farmer to societal benefits in the form of ecosystem services such as soil P retention for water quality (Macintosh et al., 2019). This demands far more management and site-specific information than is usually collected in routine soil testing, e.g. organic P, subsoil P reserves, or rooting patterns. Online tools such as the BFDC Interrogator in Australia (Dyson and Conyers, 2013) can then provide tailored fertilization recommendations while at the same time being easily accessible and user-friendly.

5. Conclusions

We present a novel approach based on multivariate multilevel modelling to relate crop yield to STP and concurring pedoclimatic conditions during 26 years on six sites in Switzerland. The inclusion of random effects governed model fit more profoundly than STP extraction method or choice of response function, which highlights the importance of taking the hierarchical data structure inherent to long-term multi-experiment studies into account. Crop response to $STP_{H2O\text{-}CO2}$ was better interpretable, less variable among crops and pedoclimatic scenarios, and less sensitive to changes in soil pH than crop response to STP_{AAE10}. Therefore, $STP_{H2O\text{-}CO2}$ facilitates better estimations of crop P requirements than STP_{AAE10} and soil P fertility classes in the Swiss fertilization guidelines PRIF need to be validated independently for the two methods. Our results further suggest that critical STP concentrations for optimal yield are lower than those currently adopted in the PRIF and that the empirically derived soil P fertility classes overestimate actual P demand of arable crops on average. Yet, soil and climate conditions account for a large variability in critical STP indicating that site-specific P fertilization recommendations could be improved on the basis of agro-climate classes in addition to currently included soil information. This may counteract the accumulation of unutilized soil P under the according pedoclimatic conditions and help to further reduce the environmental strain by long-term P application.

CRediT authorship contribution statement

Conceptualization: RF; Methodology: RF, JH; Data curation: RF, JH, BO; Statistical analysis and visualization: JH; Writing - original draft: JH; Writing - review and editing: RF, JH, FL, BO, WR; Project administration, funding, and resources: RF, WR.

Declaration of competing interest

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

Acknowledgement

We sincerely thank Ulrich Walther for co-conceptualizing and initiating the long-term field trials. We greatly acknowledge their continuous support by Agroscope and the meticulous maintenance by our field staff, in particular P. Bassetti, M. Bertossa, E. Brack, K. Casada, H.-R. Hunziker, K. Meier-Zimmermann, and H.-U. Zbinden. Over the years, many collaborators and civilian service workers were involved in field management, sampling and sample preparation, data analysis, and affiliated studies and we thank all of them for their commitment and engagement. We also acknowledge the service provided by MeteoSwiss, the Swiss Federal Office of Meteorology and Climatology.

Funding

This work was supported by Agroscope and did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scitotenv.2020.143453.

References

Agristat, 1990-2017. Statistische Erhebungen und Schätzungen über Landwirtschaft und Ernährung (SES). Schweizer Bauernverband, Brugg.

Agroscope, 1996. Schweizerische Referenzmethoden der Forschungsanstalten Agroscope, Band 1: Bodenuntersuchungen zur Düngeberatung.

Archontoulis, S.V., Miguez, F.E., 2015. Nonlinear regression models and applications in agricultural research. Agron. J. 107, 786–798.

Auguie, B., 2017. ggridExtra: Miscellaneous Functions for "Grid" Graphics. R Package Version 2.3.

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

Barber, S.A., 1980. Twenty-five years of phosphate and potassium fertilization of a crop rotation. Fertilizer research 1, 29–36.

Bates, D., Maechler, M., Bolker, B., Walker, S., 2015. Fitting linear mixed-effects models using lme4. J. Stat. Softw. 67, 48.

Bell, R., Reuter, D., Scott, B., Sparrow, L., Strong, W., 2013. Soil phosphorus—crop response calibration relationships and criteria for winter cereal crops grown in Australia. Crop and Pasture Science 64, 480–498.

Bonate, P.L., 2017. Effect of correlation on covariate selection in linear and nonlinear mixed effect models. Pharm. Stat. 16, 45–54.

Buczko, U., van Laak, M., Eichler-Löbermann, B., Gans, W., Merbach, I., Panten, K., Peiter, E., Reitz, T., Spiegel, H., von Tucher, S., 2018. Re-evaluation of the yield response to phosphorus fertilization based on meta-analyses of long-term field experiments. Ambio 47, 50–61.

Cade-Menun, B.J., 2017. Characterizing phosphorus forms in cropland soils with solution 31P-NMR: past studies and future research needs. Chemical and Biological Technologies in Agriculture 4, 19.

Cadot, S., Bélanger, G., Ziadi, N., Morel, C., Sinaj, S., 2018. Critical plant and soil phosphorus for wheat, maize, and rapeseed after 44 years of P fertilization. Nutr. Cycl. Agroecosyst. 112, 417–433.

Casler, M.D., 2015. Fundamentals of experimental design: guidelines for designing successful experiments. Agron. J. 107, 692–705.

Colomb, B., Debaeke, P., Jouany, C., Nolot, J.M., 2007. Phosphorus management in low input stockless cropping systems: crop and soil responses to contrasting P regimes in a 36-year experiment in southern France. Eur. J. Agron. 26, 154–165.

Colwell, J.D., Suhet, A.R., Van Raij, B., 1988. Statistical Procedures for Developing General Soil Fertility Models for Variable Regions. CSIRO Division of Soils.

Dahnke, W., Olson, R., 1990. Soil test correlation, calibration and recommendation. In: Westerman, R. (Ed.), Soil Testing and Plant Analysis. Soil Science Society of America, Madison, WI, pp. 45–71.

Demaria, P., Flisch, R., Frossard, E., Sinaj, S., 2005. Exchangeability of phosphate extracted by four chemical methods. J. Plant Nutr. Soil Sci. 168, 89–93.

Dirks, B., Scheffer, H., 1930. Der Kohlensäure-Bikarbonatauszug und der Wasserauszug als Grundlage zur Ermittlung der Phosphorsäurebedürftigkeit der Böden. Landwirtschaftliche Jahrbücher 71, 73–99.

Dodd, J.R., Mallarino, A.P., 2005. Soil-test phosphorus and crop grain yield responses to long-term phosphorus fertilization for corn-soybean rotations. Soil Sci. Soc. Am. J. 69, 1118–1128.

Dyson, C., Conyers, M., 2013. Methodology for online biometric analysis of soil test–crop response datasets. Crop and Pasture Science 64, 435–441.

Engels, C., Kirkby, E., White, P., 2012. Chapter 5 - mineral nutrition, yield and source–sink relationships. In: Marschner, P. (Ed.), Marschner's Mineral Nutrition of Higher Plants, Third edition Academic Press, San Diego, pp. 85–133.

Erdin, D., 2018. AGRISTAT Aktuell 18-06: Ertragsentwicklung im Schweizer Ackerbau. Schweizer Bauernverband, Brugg.

Flisch, R., Sinaj, S., Charles, R., Richner, W., 2009. Grudaf 2009. Grundlagen für die Düngung im Acker- und Futterbau. Agrarforschung 16, 1–100.

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

FOAG, 2018. Agricultural Report 2018. Federal Office for Agriculture FOAG Bern, Switzerland.

Gahoonia, T.S., Nielsen, N.E., 1996. Variation in acquisition of soil phosphorus among wheat and barley genotypes. Plant Soil 178, 223–230.

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

Genot, V., Renneson, M., Bock, L., Colinet, G., Goffaux, M.-J., Cugnon, T., Coutisse, P., Lizin, P., Blondiau, L.-M., Vanvyve, D., 2011. Un conseil de fumure raisonné-le cas du phosphore. Requasud.

Gonçalves, M.A.D., Bello, N.M., Dritz, S.S., Tokach, M.D., DeRouchey, J.M., Woodworth, J.C., Goodband, R.D., 2016. An update on modeling dose–response relationships: accounting for correlated data structure and heterogeneous error variance in linear and nonlinear mixed models1,2. J. Anim. Sci. 94, 1940–1950.

Graves, S., Piepho, H.-P., Selzer, L., Dorai-Raj, S., 2015. multcompView: Visualizations of Paired Comparisons. R Package Version 0.1-7.

Halekoh, U., Højsgaard, S., 2014. A Kenward-Roger approximation and parametric bootstrap methods for tests in linear mixed models - the R package pbkrtest. J. Stat. Softw. 59, 30.

Hampel, F.R., 1985. The breakdown points of the mean combined with some rejection rules. Technometrics 27, 95–107.

Hasler, A., 1957. Düngernormen in der Praxis und Ergebnisse der Bodenuntersuchung. Schweiz. Landw. Z. "Die Grüne" 85, 592–598.

Hinsinger, P., 2001. Bioavailability of soil inorganic P in the rhizosphere as affected by root-induced chemical changes: a review. Plant Soil 237, 173–195.

Hochmuth, G.J., Hanlon, E.A., Cornell, J., 1993. Watermelon phosphorus requirements in soils with low mehlich-I-extractable phosphorus. HortScience 28, 630–632.

Hochmuth, G., Hanlon, E., Overman, A., 2011. Fertilizer Experimentation, Data Analyses, and Interpretation for Developing Fertilization Recommendations—Examples With Vegetable Crop Research. University of Florida/IFAS Extension (EDIS# SL345).

Højsgaard, S., Halekoh, U., 2018. doBy: Groupwise Statistics, LSmeans, Linear Contrasts, Utilities. R package version 4.6-2.

Holford, I., 2015. Comments on papers relating to soil phosphorus testing in 'Making better fertiliser decisions for cropping systems in Australia' in Crop & Pasture Science 64, 435–441, 461–498 (2013). Crop and Pasture Science 66, 107–110.

Holford, I., Morgan, J., Bradley, J., Cullis, B., 1985. Yield responsiveness and response curvature as essential criteria for the evaluation and calibration of soil phosphate tests for wheat. Soil Research 23, 167–180.

Holzkämper, A., Fossati, D., Hiltbrunner, J., Fuhrer, J., 2015. Spatial and temporal trends in agro-climatic limitations to production potentials for grain maize and winter wheat in Switzerland. Reg. Environ. Chang. 15, 109–122.

Hons, F.M., Larson-Vollmer, L.A., Locke, M.A., 1990. NH4OAc-EDTA-extractable phosphorus as a soil test procedure. Soil Sci. 149, 249–256.

Hopkins, B.G., Horneck, D.A., MacGuidwin, A.E., 2014. Improving phosphorus use efficiency through potato rhizosphere modification and extension. Am. J. Potato Res. 91, 161–174.

Hothorn, T., Bretz, F., Westfall, P., 2008. Simultaneous inference in general parametric models. Biom. J. 50, 18.

Houben, D., Meunier, C., Pereira, B., Sonnet, P., 2011. Predicting the degree of phosphorus saturation using the ammonium acetate–EDTA soil test. Soil Use Manag. 27, 283–293.

Hubert, M., Van der Veeken, S., 2008. Outlier detection for skewed data. Journal of Chemometrics: A Journal of the Chemometrics. Society 22, 235–246.

Jarvie, H.P., Sharpley, A.N., Withers, P.J., Scott, J.T., Haggard, B.E., Neal, C., 2013. Phosphorus mitigation to control river eutrophication: murky waters, inconvenient truths, and "postnormal" science. J. Environ. Qual. 42, 295–304.

Johnston, A., 2005. Phosphorus nutrition of arable crops. In: Sims, J.T., Sharpley, A.N. (Eds.), Phosphorus: Agriculture and the Environment. Agronomy Monograph 46. American Society of Agronomy, Madison, WI, pp. 495–520.

Johnston, A., Richards, I., 2003. Effectiveness of the water-insoluble component of triple superphosphate for yield and phosphorus uptake by plants. J. Agric. Sci. 140, 267–274.

Johnston, A.E., Poulton, P.R., Fixen, P.E., Curtin, D., 2014. Chapter five - phosphorus: its efficient use in agriculture. In: Sparks, D.L. (Ed.), Advances in Agronomy. Academic Press, pp. 177–228.

Jordan-Meille, L., Rubæk, G.H., Ehlert, P., Genot, V., Hofman, G., Goulding, K., Recknagel, J., Provolo, G., Barraclough, P., 2012. An overview of fertilizer-P recommendations in Europe: soil testing, calibration and fertilizer recommendations. Soil Use Manag. 28, 419–435.

Kirkby, E.A., Johnston, A.E.J., 2008. Soil and fertilizer phosphorus in relation to crop nutrition. The Ecophysiology of Plant-Phosphorus Interactions. Springer, pp. 177–223.

Kleber, M., Eusterhues, K., Keiluweit, M., Mikutta, C., Mikutta, R., Nico, P.S., 2015. Mineral–organic associations: formation, properties, and relevance in soil environments. Adv. Agron. 1–140 Elsevier.

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

Kuchenbuch, R.O., Buczko, U., 2011. Re-visiting potassium-and phosphate-fertilizer responses in field experiments and soil-test interpretations by means of data mining. J. Plant Nutr. Soil Sci. 174, 171–185.

Kuznetsova, A., Brockhoff, P.B., Christensen, R.H.B., 2017. lmerTest package: tests in linear mixed effects models. J. Stat. Softw. 82.

Lakanen, E., Erviö, R., 1971. A comparison of eight extractants for the determination of plant available micronutrients in soils. Acta Agralia Fennica 123, 223–232.

Leinweber, P., Bathmann, U., Buczko, U., Douhaire, C., Eichler-Löbermann, B., Frossard, E., Ekardt, F., Jarvie, H., Krämer, I., Kabbe, C., Lennartz, B., Mellander, P.-E., Nausch, G., Ohtake, H., Tränckner, J., 2018. Handling the phosphorus paradox in agriculture and natural ecosystems: scarcity, necessity, and burden of P. Ambio 47, 3–19.

Lenth, R., 2018. emmeans: Estimated Marginal Means, aka Least-squares Means. R Package Version 1.2.4.

Lobell, D.B., Field, C.B., 2007. Global scale climate–crop yield relationships and the impacts of recent warming. Environ. Res. Lett. 2, 014002.

Loughin, T.M., 2006. Improved experimental design and analysis for long-term experiments. Crop Sci. 46, 2492–2502.

Lynch, J.P., 2011. Root phenes for enhanced soil exploration and phosphorus acquisition: tools for future crops. Plant Physiol. 156, 1041–1049.

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

Maguire, R., Chardon, W., Simard, R., 2005. Assessing potential environmental impacts of soil phosphorus by soil testing. In: Sims, J., Sharpley, A., Pierzynski, G., Westermann, D., Cabrera, M., Powell, J., Daniel, T., Withers, P., Al-Amoodi, L., Barbarick, K., Roberts, C., Dick, W. (Eds.), Phosphorus: Agriculture and the Environment. ASA/CSSA/SSSA, Madison WI, pp. 145–180.

Mallarino, A.P., 2012. Nutrient management for increased crop productivity and reduced environmental impacts. Actas XIX Congreso Latinoamericano-XXIII Congreso Argentino de la Ciencia del Suelo: "Latinoamérica unida protegiendo a sus suelos", pp. 16–20.

Mallarino, A., Blackmer, A., 1992. Comparison of methods for determining critical concentrations of soil test phosphorus for corn. Agron. J. 84, 850–856.

Marschner, P., Rengel, Z., 2012. Nutrient availability in soils. Marschner's Mineral Nutrition of Higher Plants. Elsevier, pp. 315–330.

McDonald, G.K., Taylor, J.D., Gong, X., Bovill, W., 2018. Responses to phosphorus among barley genotypes. Crop and Pasture Science 69, 574–586.

McGonigle, T.P., Miller, M.H., Young, D., 1999. Mycorrhizae, crop growth, and crop phosphorus nutrition in maize-soybean rotations given various tillage treatments. Plant Soil 210, 33–42.

Mehr, J., Jedelhauser, M., Binder, C., 2018. Transition of the Swiss phosphorus system towards a circular economy—part 1: current state and historical developments. Sustainability 10, 1479.

Menezes-Blackburn, D., Giles, C., Darch, T., George, T.S., Blackwell, M., Stutter, M., Shand, C., Lumsdon, D., Cooper, P., Wendler, R., Brown, L., Almeida, D.S., Wearing, C., Zhang, H., Haygarth, P.M., 2018. Opportunities for mobilizing recalcitrant phosphorus from agricultural soils: a review. Plant Soil 427, 5–16.

Moore, K.J., Dixon, P.M., 2015. Analysis of combined experiments revisited. Agron. J. 107, 763–771.

Morel, C., Plenchette, C., Fardeau, J.C., 1992. La fertilisation phosphatée raisonnée de la culture du blé. Agronomie 12, 565–579.

Morris, N., Knight, S., Philpott, H., Blackwell, M., 2017. Cost-effective phosphorus management on UK arable farms. Report on Work Package 2: Critical Levels of Soil P. Agriculture and Horticulture Development Board (AHDB).

Nawara, S., Van Dael, T., Merckx, R., Amery, F., Elsen, A., Odeurs, W., Vandendriessche, H., Mcgrath, S., Roisin, C., Jouany, C., Pellerin, S., Denoroy, P., Eichler-Löbermann, B., Börjesson, G., Goos, P., Akkermans, W., Smolders, E., 2017. A comparison of soil tests for available phosphorus in long-term field experiments in Europe. Eur. J. Soil Sci. 68, 873–885.

Némery, J., Garnier, J., Morel, C., 2005. Phosphorus budget in the Marne Watershed (France): urban vs. diffuse sources, dissolved vs. particulate forms. Biogeochemistry 72, 35–66.

Neyroud, J.A., Lischer, P., 2003. Do different methods used to estimate soil phosphorus availability across Europe give comparable results? J. Plant Nutr. Soil Sci. 166, 422–431.

Nobile, C., Houben, D., Michel, E., Firmin, S., Lambers, H., Kandeler, E., Faucon, M.P., 2019. Phosphorus-acquisition strategies of canola, wheat and barley in soil amended with sewage sludges. Sci. Rep. 9, 14878.

Oberson, A., Pypers, P., Bünemann, E.K., Frossard, E., 2011. Management impacts on biological phosphorus cycling in cropped soils. In: Bünemann, E., Oberson, A., Frossard, E. (Eds.), Phosphorus in Action: Biological Processes in Soil Phosphorus Cycling. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 431–458.

Parent, S.-É., Leblanc, M.A., Parent, A.-C., Coulibali, Z., Parent, L.E., 2017. Site-specific multilevel modeling of potato response to nitrogen fertilization. Frontiers in Environmental Science 5.

Pätzold, S., Leenen, M., Frizen, P., Heggemann, T., Wagner, P., Rodionov, A., 2019. Predicting plant available phosphorus using infrared spectroscopy with consideration for future mobile sensing applications in precision farming. Precis. Agric. 21, 737–761.

Payne, R.W., 2015. The design and analysis of long-term rotation experiments. Agron. J. 107, 772–785.

Peltonen-Sainio, P., Jauhiainen, L., Trnka, M., Olesen, J.E., Calanca, P., Eckersten, H., Eitzinger, J., Gobin, A., Kersebaum, K.C., Kozyra, J., 2010. Coincidence of variation in yield and climate in Europe. Agric. Ecosyst. Environ. 139, 483–489.

Peyer, K., Frei, E., 1971. Grenzwerte der Phosphatbedürftigkeit verschiedener Böden, ermittelt nach der schweizerischen CO2-Methode. Schweiz Landwirt Forsch 224–248.

Pinheiro, J., Bates, D., 2006. Mixed-effects Models in S and S-PLUS. Springer Science & Business Media.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., 2018. _nlme: Linear and Nonlinear Mixed Effects Models_. R Package Version 3.1-137.

R Core Team, 2019. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Recena, R., Díaz, I., del Campillo, M.C., Torrent, J., Delgado, A., 2016. Calculation of threshold Olsen P values for fertilizer response from soil properties. Agron. Sustain. Dev. 36, 54.

Renneson, M., Barbieux, S., Colinet, G., 2016. Indicators of phosphorus status in soils: significance and relevance for crop soils in southern Belgium. A review. Biotechnologie, Agronomie, Société et Environnement= Biotechnology, Agronomy, Society and Environment 20, 257–272.

Revelle, W., 2018. psych: Procedures for Personality and Psychological Research. R Package Version 1.8.10.

Ritz, C., Streibig, J.C., 2008. Nonlinear Regression With R. Springer Science & Business Media.

Rowe, H., Withers, P.J.A., Baas, P., Chan, N.I., Doody, D., Holiman, J., Jacobs, B., Li, H., MacDonald, G.K., McDowell, R., Sharpley, A.N., Shen, J., Taheri, W., Wallenstein, M., Weintraub, M.N., 2016. Integrating legacy soil phosphorus into sustainable nutrient management strategies for future food, bioenergy and water security. Nutr. Cycl. Agroecosyst. 104, 393–412.

Rykaczewska, K., 2015. The effect of high temperature occurring in subsequent stages of plant development on potato yield and tuber physiological defects. Am. J. Potato Res. 92, 339–349.

Sari, B.G., Olivoto, T., Diel, M.I., Krysczun, D.K., Lúcio, A.D.C., Savian, T.V., 2018. Nonlinear modeling for analyzing data from multiple harvest crops. Agron. J. 110, 2331–2342.

Schielzeth, H., Nakagawa, S., 2013. Nested by design: model fitting and interpretation in a mixed model era. Methods Ecol. Evol. 4, 14–24.

Schloerke, B., Crowley, J., Cook, D., Briatte, F., Marbach, M., Thoen, E., Elberg, A., Larmarange, J., 2018. GGally: Extension to 'ggplot2'. R Package Version 1.4.0.

Schneider, K.D., Thiessen Martens, J.R., Zvomuya, F., Reid, D.K., Fraser, T.D., Lynch, D.H., O'Halloran, I.P., Wilson, H.F., 2019. Options for improved phosphorus cycling and use in agriculture at the field and regional scales. J. Environ. Qual. 48, 1247–1264.

Shwiekh, R., Schick, J., Kratz, S., Rückamp, D., Schnug, E., 2015. Comparative investigation and inter-calibration of different soil P tests. J. Kult. 67, 61–72.

Sihvonen, M., Hyytiäinen, K., Valkama, E., Turtola, E., 2018. Phosphorus and nitrogen yield response models for dynamic bio-economic optimization: an empirical approach. Agronomy 8, 41.

Sinaj, S., Charles, R., Baux, A., Dupuis, B., Hiltbrunner, J., Levy, L., Pellet, D., Blanchet, G., Jeangros, B., 2017. 8/ Düngung von Ackerkulturen. In: Richner, W., Sinaj, S. (Eds.), Grundlagen für die Düngung landwirtschaftlicher Kulturen in der Schweiz (GRUD 2017). Agrarforschung Schweiz.
Slowikowski, K., 2018. ggrepel: Automatically Position Non-overlapping Text Labels With 'ggplot2'. R Package Version 0.8.0.
Smolders, E., Nawara, S., De Cooman, E., Merckx, R., Martens, S., Elsen, A., Odeurs, W., Vandendriessche, H., Santner, J., Amery, F., 2020. The phosphate desorption rate in soil limits phosphorus bioavailability to crops. Eur. J. Soil Sci. 1–13.
Stalham, M.A., Allen, E.J., 2001. Effect of variety, irrigation regime and planting date on depth, rate, duration and density of root growth in the potato (Solanum tuberosum) crop. J. Agric. Sci. 137, 251–270.
Stuenzi, H., 2006a. The methods of soil extraction with water and CO2 saturated water. Agrarforschung (Switzerland) 13 (7), 284–289.
Stuenzi, H., 2006b. The soil P extraction with ammonium acetate EDTA (AAE10). Agrarforschung (Switzerland) 13 (11–12), 488–493.
Sucunza, F.A., Gutierrez Boem, F.H., Garcia, F.O., Boxler, M., Rubio, G., 2018. Long-term phosphorus fertilization of wheat, soybean and maize on Mollisols: soil test trends, critical levels and balances. Eur. J. Agron. 96, 87–95.
Swiss Academies of Arts and Sciences, 2016. Brennpunkt Klima Schweiz. Grundlagen, Folgen und Perspektiven., Swiss Academies Reports. Akademien der Wissenschaften Schweiz.
Swiss Federal Council, 2013. Verordnung über die Direktzahlungen an die Landwirtschaft (Direktzahlungsverordnung, DZV). In: Council, S.F. (Ed.), (910.13).
Syers, J., Johnston, A., Curtin, D., 2008. Efficiency of soil and fertilizer phosphorus use. FAO Fertilizer and plant nutrition bulletin 18.
Tang, X., Ma, Y., Hao, X., Li, X., Li, J., Huang, S., Yang, X., 2009. Determining critical values of soil Olsen-P for maize and winter wheat from long-term experiments in China. Plant Soil 323, 143–151.
Tunney, H., Csatho, P., Ehlert, P., 2003. Approaches to calculating P balance at the field-scale in Europe. J. Plant Nutr. Soil Sci. 166, 438–446.
Valkama, E., Uusitalo, R., Turtola, E., 2011. Yield response models to phosphorus application: a research synthesis of Finnish field trials to optimize fertilizer P use of cereals. Nutr. Cycl. Agroecosyst. 91, 1–15.
van de Wiel, C.C., van der Linden, C.G., Scholten, O.E., 2016. Improving phosphorus use efficiency in agriculture: opportunities for breeding. Euphytica 207, 1–22.
van Maarschalkerweerd, M., Husted, S., 2015. Recent developments in fast spectroscopy for plant mineral analysis. Front. Plant Sci. 6.
Vance, C.P., Uhde-Stone, C., Allan, D.L., 2003. Phosphorus acquisition and use: critical adaptations by plants for securing a nonrenewable resource. New Phytol. 157, 423–447.
Vu, V.Q., 2011. ggbiplot: A ggplot2 Based Biplot. R Package Version 0.55.
Wang, Y.-L., Almvik, M., Clarke, N., Eich-Greatorex, S., Øgaard, A.F., Krogstad, T., Lambers, H., Clarke, J.L., 2015. Contrasting responses of root morphology and root-exuded organic acids to low phosphorus availability in three important food crops with divergent root traits. AoB Plants 7, plv097.
Watmuff, G., Reuter, D., Speirs, S., 2013. Methodologies for assembling and interrogating N, P, K, and S soil test calibrations for Australian cereal, oilseed and pulse crops. Crop and Pasture Science 64, 424–434.
White, P.J., Hammond, J.P., 2008. Phosphorus nutrition of terrestrial plants. In: White, P.J., Hammond, J.P. (Eds.), The Ecophysiology of Plant-Phosphorus Interactions. Springer Netherlands, Dordrecht, pp. 51–81.
Wickham, H., 2007. Reshaping data with the reshape package. J. Stat. Softw. 21, 20.
Wickham, H., 2011. The split-apply-combine strategy for data analysis. J. Stat. Softw. 40, 29.
Wickham, H., 2016. ggplot2: Elegant Graphics for Data Analysis. Springer.
Wickham, H., 2019. stringr: Simple, Consistent Wrappers for Common String Operations. R Package Version 1.4.0.
Wickham, H., Pedersen, T.L., 2019. gtable: Arrange 'Grobs' in Tables. R Package Version 0.3.0.
Wickham, H., Francois, R., Henry, L., Müller, K., 2018. dplyr: A Grammar of Data Manipulation. R Package Version 0.7.6.
Withers, P., Sylvester-Bradley, R., Jones, D., Healey, J., Talboys, P., 2014. Feed the crop not the soil: rethinking phosphorus management in the food chain. Environmental science & technology 48.
Woodard, H., Hossner, L., Onken, A., Clark, E., Haby, V., Hipp, B., Matocha, J., Mulkey, J., Wiedenfeld, R., 1994. A preliminary comparison of the ammonium acetate-edta soil phosphorus extraction method to the bray-1 and olsen soil phosphorus extraction methods. Commun. Soil Sci. Plant Anal. 25, 2909–2923.
World Meteorological Organization, 2017. WMO Guidelines on the Calculation of Climate Normals. World Meteorological Organization (WMO), Geneva, Switzerland, p. 29.
Wu, Q., Zhang, S., Ren, Y., Zhan, X., Xu, M., Feng, G., 2018. Soil phosphorus management based on the agronomic critical value of Olsen P. Commun. Soil Sci. Plant Anal. 49, 934–944.
Yamaguchi, J., 2002. Measurement of root diameter in field-grown crops under a microscope without washing. Soil science and plant nutrition 48, 625–629.
Yang, R.C., 2010. Towards understanding and use of mixed-model analysis of agricultural experiments. Can. J. Plant Sci. 90, 605–627.
Yli-Halla, M., 2016. Fate of fertilizer P in soils: Inorganic pathway. Phosphorus in Agriculture: 100% Zero. Springer, pp. 27–40.
Zhang, T., MacKenzie, A., Liang, B., Drury, C., 2004. Soil test phosphorus and phosphorus fractions with long-term phosphorus addition and depletion. Soil Sci. Soc. Am. J. 68, 519–528.
Zuber, V., Strimmer, K., 2009. Gene ranking and biomarker discovery under correlation. Bioinformatics 25, 2700–2707.