PEDOSPHERE（土壤圈英文版）
Pedosphere ISSN 1002-0160/CN 32-1315/P doi:10.1016/S1002-0160(21)60012-7

Predicting phosphorus sorption isotherm parameters in soil using routine laboratory methods

Kathleen S. DUNNE[1,2], Nicholas M. HOLDEN[2] and Karen DALY[1,*]

[1]*Teagasc, Johnstown Castle, Environmental Research Centre, Co. Wexford (Ireland)*
[2]*UCD School of Biosystems and Food Engineering, University College Dublin, Belfield, Dublin 4 (Ireland)*

ABSTRACT

Knowledge of phosphorus (P) sorption dynamics across different soil types could direct agronomic and environmental management of P. The objective of this study was to predict P isotherm parameters for a national soil population, from routine laboratory methods. Langmuir and Freundlich sorption parameters were calculated from two different ranges (0–25 and 0–50 mg P L^{-1}) using an archive of representative agricultural soil types from Ireland. Multiple linear regression (MLR) identified, labile forms of Al, Fe, organic matter (OM), cation exchange capacity (CEC) and clay as significant drivers. Langmuir and Freundlich sorption capacities, Freundlich affinity constant and Langmuir buffer capacitywere predicted reliably, with R^2 of independent validation > 0.9. Sorption isotherm parameters were predicted from P sorbed at a single concentration of 50 mg P L^{-1} (S_{50}). A MLR prediction of $Smax_{50}$ was achieved, to an accurate standard, using S_{50}, OM and Mehlich-3 Fe (R^2c = 0.91 and R^2v = 0.95). Using Gile's four shapes of isotherm (C, L, H and S) L shape (non-strict) and C shape isotherm curves accounted for 64% and 27% of soils, respectively. Hierarchical clustering identified a separation of isotherm curves influenced by two ranges of Mehlich-3 Al. Soils with a low range of Mehlich-3 Al (2.5--698 mg kg^{-1}) had no incidence of rapid sorption (C shape). Single point indices, Al, or available soil data make the regression approach a feasible way of predicting Langmuir parameters that could be included with standard agronomic soil P testing.

Key Words: fertiliser, isotherm, phosphorus, soil, sorption

INTRODUCTION

Phosphorus (P) is a plant macronutrient that is widely used to fertilise soils. It must be managed efficiently, because it is a finite resource (Neset and Cordell, 2012; Chowdhury *et al*., 2017). Phosphorus fertiliser has also been shown to be a large input cost for farms, representing 15% to 20% of total variable costs of Irish dairy farms (Creamer and O'Sullivan, 2018). The mobility of P in the

*Correspondence author. E-mail: karen.daly@teagasc.ie.

soil and availability to the plant is driven by diffusion of the phosphate anion into the aggregates of particles and porous imperfections of crystals in the soil solid phase (Mikutta *et al*., 2006; Demand *et al*., 2017).

Internationally, soil P testing is used to estimate plant available P by extracting this fraction with a reagent. Values are then used to guide fertiliser applications and form the basis of nutrient management advice. These static tests consider neither detail of P sorption processes, nor other controlling factors. Langmuir isotherms have been used to describe the dynamics of P sorption for a range of Irish grassland soils derived from different parent material (Daly *et al*., 2015) (Eq. 1):

$$\frac{C}{S} = \frac{1}{kS_{max}} + \frac{C}{S_{max}} \quad (1)$$

where C is concentration of P in solution after 24-h equilibration on a shaker (mg L^{-1}), S is (initial P added (mg L^{-1}) - C (mg L^{-1})) × 15 = P sorbed (mg kg^{-1}), S_{max} is P sorption maximum (mg kg^{-1}), and k is a binding energy constant (L mg^{-1}) (Pierzynski, 2000). The Freundlich isotherm is another commonly used isotherm to describe the dynamics of P sorption in soil (Eq. 2):

$$\log S = \log K + n \log C \quad (2)$$

where *S* is the total amount of P sorbed (mg kg^{-1}), *C* is the equilibrium P concentration (mg L^{-1}), *n* is the affinity constant which is equivalent to a binding energy (L mg^{-1}) and *K* is the adsorption constant (mg kg^{-1}) (Pierzynski, 2000).

A previous study noted that the sorption maximum value determined using the Langmuir isotherm equation does not represent the P sorption maximum in reality (Wang *et al*., 2016). The shape of sorption isotherm typically falls into four categories (Giles *et al*., 1974) and a plot of solution P (mg P L^{-1}) on x-axis versus P sorbed (S, mg P kg^{-1}) can be used to interpret the sorption mechanism (Giles *et al*., 1974; Limousin *et al*., 2007). The C-shape isotherm describes a narrow range of analyte concentration or very low concentrations where the ratio between the concentration of the compound in solution and adsorbed on the solid is the same at any concentration; the L-shape isotherm (which has two subgroups in literature, named here as; L strict, where a strict asymptotic plateau is reached and there is limited sorption capacity on the solid and; L non-strict, where the curve does not reach a plateau and the solid does not show limited sorption capacity), describes a concave curve, because the ratio between the concentration of the compound in solution and adsorbed on the solid decreases as solute concentration increases; the H-shape isotherm is a variation of the L isotherm where the initial slope is greater; and the S is sigmoidal indicating two opposing mechanisms (Limousin *et al*., 2007). To save on time, cost and labour single point sorption values have been investigated in the past as indicators of isotherm properties because they reduce data collection from around 6 data points to a single one.

The amount of P sorbed after a single addition of P has been used as a predictor of isotherm properties in some studies, however it is common to use the amount of P sorbed from an added amount of P to a soil sample (X, mg P kg^{-1}) within an equation (*e.g*., X/logC) to make predictions about linearised sorption parameters (*e.g*., sorption maximum). Bache and Williams (1971) used 150 mg P per 1 g soil and the equilibrium concentration is taken into account in the quotient, X/logC when

indicating phosphate sorption isotherm parameters. Burkitt *et al*. (2002) examined the relationship between phosphorus buffer capacity, $PBC_{O\&S}$ (Ozanne and Shaw, 1968), and 11 different single point sorption indices. Burkitt *et al*. (2002) reported that when a single addition of 1 000 mg P kg^{-1} was added to soil and either the Colwell or 4.59 × Olsen extractable P were added to the amount of P sorbed, results were most closely related to $PBC_{O\&S}$. Burkitt *et al*. (2008) later reported that PBI (which was defined as the slope between an equilibrium P concentration of 0.25 and 0.35 mg P L^{-1}) without the addition of Colwell P resulted in better fertiliser management practice. Wang *et al*. (2016) indicated dissolved reactive P concentration in soil leachate using various versions of P sorption indices. They determined a single point isotherm, PSI-a, as the amount of P sorbed by a soil during 24-h shaking in 60 mg P L^{-1} solution, PSI-b as the quotient of PSI-a/log C, where C was the solution P concentration after 24-h shaking and PSI-c as the sum of PSI-a and Olsen P (Wang *et al*., 2016). They found the best prediction of P sorption maximum was PSI-c. Several studies from Brazil (Corrêa *et al*., 2011; Rogeri *et al*., 2016; Alovisi *et al*., 2020) used parameter called P remaining (Prem), which is P remaining in solution after equilibration with an initial P concentration of 60 mg P L^{-1} and correlated this with the Langmuir derived maximum adsorption capacity of P. It was also found that Prem was a better predictor than clay content for use with Mehlich measurement (Rogeri *et al*., 2016).

While predictions also have associated uncertainties, they do encourage green analytical chemistry (Pena-Pereira *et al*., 2015; Armenta *et al*., 2019), cost less, require less labour (Khaledian and Miller, 2020) and could be sufficiently reliable to allow P management to be driven by P dynamics rather than just state. Dunne *et al*. (2020) predicted P sorption parameters to a standard that was satisfactory for rough screening using a benchtop mid infrared diffuse reflectance Fourier transform (DRIFT) spectrometer and that work will be useful, in the development towards *in-situ* measurements of P in the field (which are viewed as the full utilisation potential of spectroscopic methods in soil analysis (Stenberg *et al*., 2010)), however, there is still a need for simplified wet chemical procedures for P sorption isotherms, while maintaining analytical accuracy and understanding of key sorption behaviours. Such simplified methods could be more readily adopted than isotherms, as a test to accompany a static soil P status test (Morgan's test in Ireland). A single point value, S_{50}, which is an addition of 50 mg P L^{-1}, could be combined with other soil data that are readily available in soil surveys and libraries all over the world, that are easily accessible due to widespread availability to soil databases (Padarian *et al*., 2020). A robust prediction could allow for better management of the finite P resource by agriculture. Robust predictions can be made using pedotransfer functions where these are defined as "equations or algorithms expressing relationships between soil properties different in difficulty of their measurement or their availability" (Pachepsky and van Genuchten, 2011). Pedotransfer functions that indicate P sorption dynamics have not been developed for general use. The objective of this study was to develop robust relationships between traditional P isotherm parameters which are time-consuming to determine and ancillary soil parameters that are quicker to determine, that could be used as reliable pedotransfer functions to allow P dynamics to be part of P management advice. The research used a representative sample from the national soil population of Ireland.

MATERIALS AND METHODS

Soil archive

Uppermost horizon samples (topsoil) (n = 224) were subsampled from a collection of 224 modal soil profiles representative of Irish soil types. These samples were from the archive of the Soil Information System (SIS) Ireland (Creamer *et al*., 2016) at Johnstown Castle, Wexford. The uppermost horizon of the sampled soils ranged from 0.03 to 0.55 m depth, with a mean depth of 0.20 m and represented eleven Great Groups. Fifty seven percent were Brown Earth and Surface Water Gley. Percentage organic matter (OM) was measured for all samples, as loss on ignition, of 4 g samples by muffle furnace, set to 500 °C for 16 h (adapted from the BS EN 13039:2000). Existing data from the same samples were used for hierarchical clustering and stepwise multiple regression. These data were determined prior to our study, for the Irish Soil Information System, SIS Ireland (Creamer *et al*., 2016). These were; pH in water, Mehlich-3 Fe (mg kg^{-1}), Mehlich-3 Al (mg kg^{-1}), Mehlich-3 P (mg kg^{-1}), depth to (cm), clay (< 0.002 mm % w/w), sand (2.00--0.05 mm % w/w), silt (0.002--0.05 mm % w/w), effective cation exchange capacity (CEC, cmol kg^{-1}) and exchangeable Ca (cmol kg^{-1}) The ratio of Mehlich-3 Al to Mehlich-3 P has been previously reported by Daly *et al*., 2015 as an indicator of P supply in soil and is included in this work along with ancillary soil parameters listed above and denoted as Al/P.

Phosphorus sorption isotherms

Isotherms were determined using the Paulter and Sims (2000) modification of the standardised batch technique (Nair *et al*., 1984), as previously reported by Dunne *et al*. (2020) to examine the performance of MIR spectroscopy. When raw sorption data were fit to Langmuir and Freundlich models in both ranges, this generated the following parameters: Langmuir binding energy (k_{25}), Langmuir sorption maximum ($Smax_{25}$), and Langmuir maximum buffering capacity (MBC_{25}) in the 0–25 mg L^{-1} range, Langmuir binding energy (k_{50}), Langmuir sorption maximum ($Smax_{50}$) and Langmuir maximum buffering capacity (MBC_{50}) in the 0–50 mg L^{-1} range, Freundlich affinity constant (n_{25}), Freundlich sorption maximum ($Kmax_{25}$) in the 0–25 mg L^{-1} range, Freundlich affinity constant (n_{50}) and Freundlich sorption maximum ($Kmax_{50}$) in the 0–50 mg L^{-1} range. Two ranges were compared (0–25 and 0–50 mg L^{-1} P) for the determination of P sorption isotherms using a 2-sided Wilcoxon signed-rank test (Wilcoxon, 1945) to evaluate the null hypothesis of no significant difference in parameter values from the smaller or larger range.

Prediction of 0-50 mg L^{-1} P isotherm parameters from the same parameters in 0-25 mg L^{-1} P range

Three Langmuir isotherm parameters; binding energy (k), sorption max (Smax) and maximum buffer capacity (MBC) in 0–50 mg P L^{-1} were predicted from the same parameters in 0–25 mg L^{-1} P range. Three different regression models were trialled: linear regression [a], second degree polynomial regression [b] and third degree polynomial regression [c]. For each regression model, 4 different methods of outlier removal were trialled: (1) no pre-processing, (2) removing outliers using box and whisker plots, the whiskers were the highest and lowest values in 1.5 X the interquartile range, values outside the whiskers were removed as outliers, (3) removing outliers identified through residuals and (4) identifying and removing outliers using residuals then taking the log of the data to

the base 10. Freundlich parameters in the 0–50 mg P L^{-1}; affinity constant (n_{50}) and sorption maximum ($Kmax_{50}$) were also predicted in this way.

Prediction of isotherm parameters using single point sorption

Five isotherm parameters; Langmuir binding energy (k_{50}), Langmuir sorption maximum ($Smax_{50}$), Langmuir maximum buffer capacity (MBC_{50}), Freundlich sorption maximum (K_{50}) and Freundlich affinity constant (n_{50}) were predicted from a single point P sorption (S_{50}). To make the distinction – Langmuir Smax is the value derived from the Langmuir model. However, S_{50}, is taken from the raw isotherm and represents the amount of P sorbed (S) at an initial concentration of 50 mg P L^{-1}. Phosphorus sorbed (P sorbed) was calculated as initial P added (mg L^{-1}) - final P in solution after 24-h equilibration on an end over end shaker (mg L^{-1}) × 15 (mg kg^{-1}). This method makes direct predictions of linearised sorption parameters without equation based mathematical manipulation.

Prediction of isotherm parameters in the 0-50 mg L^{-1} P range using soil information and a single point sorption

Langmuir and Freundlich parameters in the 0-50 mg L^{-1} P range were predicted using Normal Multiple Regression from formulae identified using Stepwise Multiple Regression. Predictions using this method were independently validated with 25% of the data that was set aside before modelling. The direction used was both forward and backward until Akaike Information Criterion (AIC) was lowest, except for the parameter labelled MBC_{50} (2). The formula with the lowest AIC for prediction of MBC_{50} was identified and then one parameter, CEC, was substituted with % organic matter (% OM), because it was known that the determination of % OM involves a less time consuming and laborious procedure than CEC. The substitution of a parameter was to test if the prediction performance changed, and if it did change, what was the magnitude of the change in performance. Percentage OM was used because it is known to affect P dynamics (Guppy *et al*., 2005).

Visual assessment of isotherm shapes using Gile's four (C, L, H and S) main shapes of isotherm

The category each raw isotherm (plot of P sorbed (mg kg^{-1}) v final C (mg L^{-1})) fell into was determined visually according to the classification of sorption isotherms proposed by Giles *et al*. (1974). Raw isotherms were overlaid and using hierarchical clustering (complete linkage and Euclidian distance), 20 samples from the top of the overlay diagram and 20 samples from the bottom of this diagram were investigated (RStudio Team, 2015; R Core Team, 2019). Samples were taken from the top and bottom of the overlay plot by sorting the data (descending) according to P sorbed at 50 mg P L^{-1}. A hierarchical clustering investigation of all samples with available ancillary properties (n = 175), was also carried out, to see if the same signal could be seen when a larger sample set representative of the national soil population was used. The Wilcoxon rank-sum/Mann-Whitney test was used for the hierarchical clustering because the resulting data sets from the hierarchical clustering were not paired and were independent. The Wilcoxon rank-sum/Mann Whitney test reports a test statistic, in R, labelled W.

RESULTS

Phosphorus sorption isotherms

Phosphorus sorption isotherms were derived for 224 samples (Fig. 1). Dunne *et al*. (2020) reported the sorption parameter range values for these soils and these indicated that the soils used captured a normal range of characteristics and were analysed reliably.

Fig. 1 Irish Soil Information System (SIS) soils (n = 224) depicted using Gile's main shapes of isotherm. Final C is on the x-axis and it is final concentration of P in solution after 24-h equilibration time with 2 g soil (mg L^{-1}) and P sorbed is on the y-axis and it is the amount of P sorbed to 2 g soil samples after 24-h equilibration time (mg kg^{-1}). Shapes; "C" which is coloured light blue, "L non-strict" which is coloured orange, "L strict" which is coloured yellow and "S" which is coloured navy, appeared in this sample set, H did not.

Isotherm range effects

Langmuir and Freundlich parameters determined over two ranges (0–25 and 0–50 mg L^{-1} P) had significantly different median parameter values (Supplementary Material Table SI). The 1-sided Wilcoxon signed-rank tests indicated that k_{25} had a significantly greater median than k_{50}, $Smax_{50}$ had a significantly greater median than $Smax_{25}$, Langmuir R^2_{50} had a significantly greater median than Langmuir R^2_{25}, n_{25} had a significantly greater median than n_{50}, K_{25} had a significantly greater median than K_{50} and Freundlich R^2_{25} had a significantly greater than median than Freundlich R^2_{50} (Supplementary Material Table SI).

Prediction of 0-50 mg L^{-1} P isotherm parameters from the same parameters in 0-25 mg L^{-1} P range

Examination of the laboratory values from the batch technique showed that average standard deviation in values of P sorbed varied from 2.59 to 9.56 mg kg^{-1}, from 25 to 50 mg P L^{-1}, respectively. Average standard deviation of final C varied from 0.17 to 0.64 mg P L^{-1} from 25 to 50 mg P L^{-1}. Wilcoxon signed-rank test was used to test if the difference in standard deviations, of analytical duplicates, from the larger point (50 mg P L^{-1}) to the smaller point (25 mg P L^{-1}) was significant. Standard deviations between duplicates for P sorbed at 50 mg P L^{-1} were significantly higher (V = 2824.5, P-value < 2.2e-16) than at 25 mg P L^{-1} and standard deviations between analytical duplicates for final C at 50 mg P L^{-1} are significantly higher (V = 2818, P-value < 2.2e-16) than at 25 mg P L^{-1}. Exploring prediction of the larger range from the smaller range avoids these potential errors and larger range isotherms might give a better indication of sorption plateaus. The higher the concentration range used to determine a sorption maximum, the more likely the sorption plateau will be reached.

A range of regression models was trialled, because when the data were visually examined most were linearly trending upwards, however a slight concave curve could be seen in some and a fanning effect of the scatter plotted data was also observed in some relationships. The suggestion of some non-linear relationships was addressed by reporting both Pearson and Kendall's *tau* to indicate strength of correlation. Three Langmuir isotherm parameters; binding energy (k), sorption max (Smax)

and maximum buffer capacity (MBC) in 0–50 mg P L^{-1} were predicted from the same parameters derived from a smaller range, 0–25 mg P L^{-1} (Supplementary Material Table SII). A plot of independent validation for the best Langmuir parameter that was predicted from a smaller range (MBC_{50} explained by MBC_{25}) is in Fig. 2. Maximum buffer capacity in the 0–50 mg P L^{-1} range was predicted using 2nd degree polynomial regression. Outliers were removed by determining extreme residual values and the data was logged before modelling to remove fanning effect. There was an improvement in the validation of most prediction models when the data were logged using base 10 as it removed the fanning effects seen in the raw data (Supplementary Material Fig. S1). According to Quantile-Quantile plots and Shapiro-Wilk tests, no isotherm variable in this data set was normally distributed. Freundlich parameters in the 0–50 mg P L^{-1}; affinity constant (n_{50}) and sorption maximum ($Kmax_{50}$) were also predicted (Supplementary Material Table SIII). Figure 3 shows the independent validation plot for $Kmax_{50}$.

Fig. 2 Independent validation plot depicting predicted versus actual values for the best Langmuir parameter that was predicted from a smaller range (MBC_{50} explained by MBC_{25}). This parameter was predicted using 2nd degree polynomial regression. The a, b line was set to intercept = 0 and slope = 1 for demonstration of correlation, it is not equivalent to the model.

Fig. 3 Independent validation plot depicting predicted versus actual values for the best Freundlich parameter that was predicted from a smaller range (n_{50} explained by n_{25}). This parameter was predicted using normal linear regression. The a, b line was set to intercept = 0 and slope = 1 for demonstration of correlation, it is not equivalent to the model.

Prediction of isotherm parameters using single point sorption

Three Langmuir isotherm parameters (binding energy (k), sorption max (Smax) and maximum buffer capacity (MBC)) and two Freundlich parameters (affinity constant (*n*) and sorption max (Kmax)) in 0–50 mg P L^{-1} were predicted from the raw isotherm maximum (S_{50}) and these results are presented in Table I.

TABLE I (at the end of text)

Regression based prediction of 0–50 mg L^{-1} P Langmuir and Freundlich isotherm parameters from the raw isotherm maximum of the 0–50 mg L^{-1} P range (*i.e.*, using a sorption single point as a predictor). Samples were randomly split 75% for the calibration set and 25% for the validation set. For each parameter (k_{50}, $Smax_{50}$, MBC_{50}, n_{50} and K_{50}), the model that gave the best outcome, with regard to parsimony and consistency between validation and calibration, is highlighted in bold

Prediction of isotherm parameters in the 0-50 mg L^{-1} P range using soil information and a single point sorption

The resulting formulae from forward and backward stepwise multiple regression are presented in Table II. The performance of independently validated predictions, using these formulae are also in Table II.

TABLE II (at the end of text)

Normal Multiple Regression Predictions of Langmuir parameters in the 0–50 mg L^{-1} P range using the resulting formula from Stepwise Multiple Regression (direction used was both forward and backward until Akaike Information Criterion (AIC) was lowest for all parameters except MBC_{50} (2)). Samples were randomly split 75% for the calibration set and 25% for the validation set

Visual assessment of isotherm shapes using Gile's main shapes of isotherm

Each raw isotherm was assigned a shape by visual assessment (Fig. 1) according to the general modelling of sorption isotherms proposed by Giles *et al*. (1974). Of 224 soil samples 1 could not be classified. Shapes C, S, L strict and L non-strict appeared in the sample set, H did not. There were 142 that were classified as non-strict L shape (64%), 16 were classified as strict L shape (7%), 60 were classified as C shape (27%) and 5 were classified as S shape (2%). Therefore, the majority of Irish agricultural soils were L non-strict shape with a smaller proportion described as C shape.

When raw isotherms were overlaid (Fig. 1) the strict L and S samples were clustering near the bottom half of the diagram, while non-strict L and C isotherms were clustering near the top of the diagram. Using hierarchical clustering in RStudio, 20 samples from the top of the overlay diagram and 20 samples from the bottom of this diagram were investigated and there was clear separation into two sets, with 4 out of 40 samples (10%) misclassified. The following parameters Mehlich-3 Al/Mehlich-3 P ratio, Mehlich-3 Fe, Mehlich-3 Al, Mehlich-3 P, Depth to, % OM, clay, sand, silt, pH in water, CEC and Ca were investigated to find the cause of the separation and Al was the only parameter with a complete separation between the two groups. Summary statistics for Al in set 1a (which is a set of n = 22 samples, where 19 are from the top portion of the plot and 3 are misclassified) were, range = 866.6--1 623.4 and median = 1 116.8 mg kg^{-1}. Summary statistics for Al in set 2a (which is a set of n = 18 samples, where 17 are from the bottom portion of the diagram and 1 is misclassified) were; range = 272.7--857.6 and median = 706.6. Therefore, the top and bottom sets of Gile's isotherm shapes are separated by high and low Al. In contrast to this behaviour Mehlich-3 Fe has a similar range in both sets (set 1 Fe range = 149.3--541.8 and median = 395.1, set 2 Fe range = 137.6–504.9 and median = 356.5) (Table III). A Wilcoxon signed-rank test was carried out (due to non-normality) to test for significant difference in medians and the resulting P-value (W = 396, P-value = 1.764 e-11) was very small for Mehlich-3 Al. Therefore, Mehlich-3 Al is significantly different in the top portion of these raw isotherms compared to the bottom. In set 1a, which was formulated from hierarchical clustering, and is the set with samples mostly from the top portion of the overlay plot, there were 17 C type raw isotherms, 3 L non strict, 1 L strict and 1 S shape. The 3 that were misclassified are the 1 S, 1 L strict and 1 L non strict. This L non strict sample didn't sorb as much P as the other non-strict isotherms, but also didn't reach a strict plateau. In set 2a, which was formulated from hierarchical clustering, and is the set with samples mostly from the bottom portion of the overlay plot, there were 11 L non strict and 7 L strict. The one that was misclassified was a L non strict shape. On examining parameter ranges and medians by Wilcoxon rank-sum test, there was an obvious signal from one parameter (Mechlich-3 Al), which lead to a hierarchical clustering investigation of all samples that had data for ancillary properties (n = 175), to see if the same signal

could be seen when a larger and more representative sample set was used. In set 1b, which was formulated from hierarchical clustering of all samples with available soil information (n = 175), there were 107 L non strict, 36 C, 13 L strict and 1 S shape. In set 2b, which was formulated from hierarchical clustering of all samples with available soil information, there were 17 L non strict and 1 L strict.

TABLE III (at the end of text)

Summary statistics from set 1 and set 2 of hierarchical clustering. To determine the cause of the isotherm shape difference, twenty samples from the top of the raw isotherm overlay plot and twenty samples from the bottom of the raw isotherm overlay plot (n = 40) were clustered and examined according to 13 ancillary parameters to give set 1a and 2a. All samples with available ancillary data (n = 175) were clustered to give set 1b and set 2b

Taking a parsimonious approach, by balancing the most accurate with the simplest method, the best approach for each of the 5 isotherm parameters was chosen. The best method for the prediction of each isotherm parameter are summarised in Table IV.

TABLE IV

Summary table presenting the best approach for each of 5 isotherm parameters predicted in this study. There were 3 approaches, 1) a single point sorption parameter prediction, 2) multiple linear regression or 3) prediction of larger range isotherm parameters from smaller range isotherm parameters

Predicted parameter	Best approach	Details of approach
$Smax_{50}$	MLR	S_{50}, OM and Mehlich-3 Fe
k_{50}	MLR	S_{50}, OM, Clay and Mehlich-3 Al
MBC_{50}	Single point sorption	S_{50}
n_{50}	Large range predicted from small range	$Parameters_{50} \sim Parameters_{25}$
K_{50}	Single point sorption	S_{50}

$Smax_{50}$, Langmuir maximum sorption capacity; k_{50}, Langmuir binding energy; MBC_{50}, Langmuir maximum buffer capacity; n_{50}, Freundlich binding energy; K_{50} Freundlich maximum sorption capacity; all obtained from a 0–50 mg P L^{-1} range of P addition; MLR, multiple linear regression; S_{50}, P sorbed from a single addition of 50 mg P L^{-1} after 24 h on an end over end shaker; OM, organic matter; $Parameters_{50}$, parameters arising after linearisation of data with an isotherm equation with a 0–50 mg P L^{-1} range of P addition; $Parameters_{25}$, parameters arising after linearisation of the data with an isotherm equation with a 0–25 mg P L^{-1} range of P addition; $Parameters_{50} \sim Parameters_{25}$, y is explained by X.

DISCUSSION

Phosphorus sorption isotherms

Although the Langmuir equation fit the 0–50 mg P L^{-1} range on more counts compared to the 0–25 mg P L^{-1} range, there was more analytical error in the wet chemical isotherm procedure at the higher concentration sorption range. A higher initial P concentration requires more dilution scheme steps and can contribute to cost and time when analysing hundreds of samples (Devriendt *et al*., 2019). The Wilcoxon signed-rank test confirmed that medians of all parameters were significantly different in one range compared to the other, this could be because when there is a low range in P addition this may result in the soil sorbing all of the P added and it will make it hard to identify the P sorption

maximum (plateau). Higher ranges of concentration give a more accurate indication of soil sorption mechanism, therefore one concentration range might have a significantly different median to the other.

Freundlich capacity and affinity constants were higher in the 0–25 mg P L^{-1} range, than the 0–50 mg P L^{-1} range. The Freundlich model which has been conceptualised as the sum of individual Langmuir models (Limousin *et al*., 2007), gave conflicting results when compared to the Langmuir model, because the Langmuir model is a more rudimentary model, and perhaps true to the mechanisms examined in this study, as it makes more sense for sorption maximum to be the same or higher when a higher range of P addition is used.

Prediction of 0--50 mg L^{-1} P isotherm parameters from the same parameters in 0--25 mg L^{-1} P range

The shapes of raw sorption isotherms (Fig. 1) have an upwards concave curve, therefore over a higher concentration the isotherm will give a better indication of a sorption maximum or plateau being reached. It is worth noting that there is more possibility of error being introduced when higher concentrations are used for wet chemical isotherm determination due to higher dilution factors (Devriendt *et al*., 2019). The results showed that only analysing to 25 mg P L^{-1} using wet chemistry does allow prediction of the 0–50 mg P L^{-1} linearised parameters reliably.

Prediction of isotherm parameters using single point sorption

Of Langmuir k_{50}, $Smax_{50}$ and MBC_{50} (Table I), $Smax_{50}$ was best predicted using the single isotherm point S_{50} (R^2v = 0.86; Fig. 4). This is a better prediction compared to predicted values using the 0–25 mg P L^{-1} range, therefore the single point prediction offers a better prediction for the Langmuir sorption maximum, saving on cost and labour for this parameter. A representative sample of agricultural Irish soils was used for this study, therefore this result is generalisable at least within Ireland and this single point could be combined with Morgan's P test, at a low cost, to give sorption maximum information, which could improve fertiliser P management. Maximum buffer capacity was the second best predicted parameter using a single raw isotherm data point (R^2v = 0.7) and Langmuir binding energy (k_{50}) was not predicted well (R^2v = 0.56). The maximum buffer capacity and binding energies can be predicted better using the small range isotherms (R^2v of MBC_{50} = 0.93 predicted using 0–25 mg P L^{-1} and R^2v of k_{50} = 0.83 predicted using 0–25 mg P L^{-1}. For the single point sorption predictions, removing outliers for k_{50} and MBC_{50} made a bigger difference than removing outliers for $Smax_{50}$. Therefore, the $Smax_{50}$ prediction using a single point is more robust compared to other Langmuir parameters, although, R^2c and R^2v became more consistent for $Smax_{50}$ when outliers were removed. The Freundlich binding (affinity constant, n_{50}) parameter was not well predicted, therefore this single point method of prediction proved to be better at predicting sorption maximum parameters than binding energy parameters. However, Langmuir maximum buffer capacity was predicted well using this method, which takes binding energy into account. There was a large discrepancy in Table I between Pearson's correlation (r = 0.24) and Kendall's *tau* (r = 0.70) for the MBC parameter. This can be explained, because the second degree polynomial regression method was the best method of prediction for this parameter. Kendall's *tau* was indicating (r = 0.70) there was indeed a relationship between the single point, S_{50}, predictor variable and the linearised MBC predicted variable, but Pearson's was not indicating (r = 0.24) a relationship, because Pearson's correlation is not robust to

outliers (Niven and Deutsch, 2012) and is not good for assessing non-linear relationships and Kendall's *tau* can give a good indication of correlation when there are outliers (Rosset *et al*., 2007) and when the relationship is non-linear (Liu *et al*., 2016).

Fig. 4 Independent validation plot depicting predicted v actual values for the best isotherm parameter that was predicted from a single point value (0–50 mg P L^{-1} Linearised Langmuir Smax explained by P sorbed at 50 mg P L^{-1}). This parameter was predicted using normal linear regression. The a, b line was set to intercept = 0 and slope = 1 for demonstration of correlation, it is not equivalent to the model.

Prediction of isotherm parameters in the 0-50 mg L^{-1} P range using soil information and a single point sorption

Organic matter (range: 4.3 to 87.6 OM %) was not highly correlated with binding energy in the these soils (Pearson's r = 0.02, P-value = 0.82), but appeared as a significant negative contribution (equation coefficient = -0.021, P-value < 0.001) to the prediction of binding energy using stepwise multiple linear regression (MLR) (Table II) and it is known that OM affects binding energy. High OM in soil can occlude P from binding sites primarily due to competitive inhibition (Guppy *et al*., 2005) and soils high in OM have low capacity for P storage (Daly *et al*., 2001). Clays, which include, layer aluminosilicates, amorphous aluminosilicates and sesquioxides (oxides, hydroxides, and oxyhydroxides of Al and Fe) are drivers of binding energy, because they make up the majority of sorbent surface area in soil (Sollins *et al*., 1996) and the mineralogy of clay in soils is reported (Uygur, 2009) to be critical for P retention. It has also been reported that the amount of Al in soils provides a good indicator of P retention (Ballard and Fiskell, 1974; Daly *et al*., 2015; Fan *et al*., 2019). This may indicate that single numbers such as a Pearson's correlation value on its own, cannot be used to make decisions in soil management and this work indicated that multi-parameter information is more appropriate for decision making on P management, for example Morgan's P or other agronomic P tests combined with isotherm sorption parameters or with parameters identified in this study using MLR.

Of the following parameters; P sorbed at 50 mg P L^{-1} (S_{50}), cation exchange capacity (CEC), pH in water, OM, Mehlich-3 Al and Mehlich-3 Fe the following results were obtained using forward and backward stepwise multiple linear regression (MLR). Langmuir sorption maximum in the 0–50 mg L^{-1} P range was best predicted in this study using a multiple linear regression approach (Fig. 5). The combination of parameters that best described Langmuir sorption maximum was (S_{50}) S_{50}+ OM + Mehlich-3 Fe (R^2c = 0.91, R^2v = 0.95). In a study, where various substrates were tested for the removal of P from constructed wetlands (Xu *et al*., 2006), it was reported that P sorption capacity decreased with increasing OM. As OM was increased the removal efficiency of the substrate decreased, because P sorption capacity is inhibited by OM. Sorption inhibition caused by OM happens because OM directly competes with P for sorption sites and OM can eliminate sorption sites such as Al and Fe, by causing an acidic environment (Daly *et al*., 2001; Guppy *et al*., 2005). Soluble OM can also release previously adsorbed P by complexing with surface-bound Al or Fe to form soluble organic-metal compounds (Xu *et al*., 2006). Iron oxides, mainly goethite have been found to be positively correlated with Langmuir P sorption maximum and have been described as conducive to P adsorption and desorption (Fink *et al*., 2016). The mechanism has been determined as consistently

fast, then slow adsorption due to a consistent goethite crystal face (110) (Torrent *et al*., 1992). Langmuir binding energy in the 0–50 mg L^{-1} P range was best described using S_{50} + OM + clay + Mehlich-3 Al (R^2c = 0.59, R^2v = 0.67). However, it should be noted that it has been predicted to a better R^2v in this study using prediction of a larger range from a smaller range (Supplementary Material Table SII).

Fig. 5 Independent validation plot depicting predicted *versus* actual values for the best isotherm parameter that was predicted from a normal multiple linear regression (Linearised Langmuir $Smax_{50}$ explained by P sorbed at 50 mg P L^{-1}, OM and Mehlich-3 Fe). The a, b line was set to intercept = 0 and slope = 1 for demonstration of correlation, it is not equivalent to the model.

Langmuir maximum buffer capacity (MBC) in the 0–50 mg L^{-1} P range was best described using S_{50} + cation exchange capacity + clay + Mehlich-3 Al (R^2c = 0.66, R^2v = 0.70). Maximum buffer capacity (MBC) is calculated as the product of sorption maximum and binding energy, and the MLR models predicting MBC and Langmuir binding energies k, share common variables (clay, Mehlich-3 Al) as significant drivers. Differences in P buffering and sorption capacities can significantly influence supply and availability of P in a water-soluble and plant available form (Indiati *et al*., 1999; Burkitt *et al*., 2001; Daly *et al*., 2015) and Mehlich-3 Al has been highly correlated with Langmuir MBC and binding energies and linked to P solubility (Daly *et al*., 2015). Certain clays and clay minerals are critical for P retention and cation exchange capacity is the total amount of negative charges on a surface which attract cations. In clay minerals, these charges result from internal substitution of structural elements in the clay lattice, for example, magnesium (Mg) substituting for aluminium (Al) or Al substituting for silica (Si) (Creamer and O'Sullivan, 2018). This creates a permanent and balanced negative charge at the clay surface, which may repel phosphate anions from being bound so tight and make them more available for plant uptake. Langmuir maximum buffer capacity (MBC) in the 0–50 mg L^{-1} P range was best described using S_{50} + cation exchange capacity + clay + Mehlich-3 Al, which is similar to the binding energy drivers, this indicates that binding energy power is driving MBC opposed to P sorption capacity.

Some of the ancillary parameters (*e.g.*, CEC) that were significant variables in in the multiple linear regressions are time consuming laboratory chemical tests with complex sample preparation (Wan *et al*., 2020). However, there are a lot of ancillary data easily available via internet download from soil surveys and libraries (Hallett *et al*., 2011; Creamer *et al*., 2014; Creamer *et al*., 2016; Ballabio *et al*., 2019), therefore only if the data are readily available will they have potential to improve P sorption regression outcomes, specifically the Langmuir sorption maximum parameter. The multiple regression method proposed here is an option, if these ancillary data are easily obtained. Substituting the time-consuming parameter (CEC) for a more easily measured parameter (OM determined by loss on ignition) was trialled and the results were similar (R^2c = 0.65, R^2v = 0.71) to the best equation for MBC_{50} which included CEC (R^2c = 0.66, R^2v = 0.70). Although this equation (S_{50} + OM + clay + Mehlich-3 Al) does not have the lowest Akaike information criterion (AIC) (this identifies the likely best model using an extension of the classical maximum likelihood principle) relative to other models (Akaike, 1973; Burnham and Anderson, 2002).

Visual assessment of isotherm shapes using Gile's four (C, L, H and S) main shapes of isotherm

In this study, the L non strict shape was prominent in Irish agricultural soils, this is a concave curve that does not reach a straight plateau (Supplementary Material Fig. S2); however, P absorbed on the soil decreases with increasing additions of P and the isotherm gradually plateaus. Therefore, when the L non strict shape is seen, sorption capacity decreases as P is increased, but at a slower rate compared to other mechanisms. For example, the L strict isotherm shape will suddenly lose sorption capacity (reaching a plateau), whereas in the L non strict isotherm shape, sorption capacity gradually decreases leaving the remaining P in solution after equilibration. The C shape isotherm accounted for 27% of soils and is characterised by a sharp rise in P sorbed (rapid sorption of P) without a plateau, and with low P concentrations remaining in solution after equilibration.

The visualisation of the raw isotherms in Fig. 1 highlighted differences in isotherm shapes/mechanism at the top and bottom of Fig. 1. The top and bottom sets of Gile's isotherm shapes are separated by high and low Mehlich-3 Al with a significant range separation between the two sets (*i.e.*, set 1 ranged between 866.6--1 623.4 mg kg^{-1} and set 2 ranged between 272.7--857.6 mg kg^{-1}). Once this was identified in the small sub-set of samples (n = 40), it was tested with a representative sub-set of samples (n = 175) and the same significant Al range separation was seen when a hierarchical clustering was carried out where two distinct ranges of Mehlich-3 Al within the larger set of soils were identified; 866.6--1 623.4 mg kg^{-1} and 702.5--2 460.8 mg kg^{-1} (Table III). Isotherm shapes C and S did not appear in soils with low range (2.5--698.5 mg kg^{-1}) of Mehlich-3 Al. The L non strict and L strict are the only shapes that appeared in this range, with the majority being L non strict on both occasions. Soils with between 702.5--2 460.8 mg kg^{-1} of Mehlich-3 Al, the L non-strict with the C shape appeared prominently (with the S shape on 2 occasions) highlighting the effectiveness of Al in rapidly adsorbing P. The separation of ranges of Mehlich-3 Al in the cluster analysis is supported by previous studies (Daly *et al*., 2015), where the relationship between MBC and Mehlich3-Al in a subset of Irish mineral soils was described by a broken-line regression with a change-point of 615 mg kg^{-1} within lower and upper confidence intervals of 508 and 716 mg kg^{-1}. Furthermore, a study of Portuguese soils noted extractable Al had important effects on P (Fernandes and Warren, 1994). Daly *et al*. (2015) also found that Langmuir sorption buffer capacity (MBC) and binding energy (k) were strongly correlated with soil pH and extractable Al, although the samples used by (Daly *et al*., 2015) were 0.10 m deep and the samples examined in the current study were the depth of the first horizon, which varied from 0.03 to 0.55 m, (mean depth of 0.20 m), nonetheless, the strong influence of Mehlich-3 Al is consistent in both studies.

The interpretation that Mehlich-3 Al is a separator/driver of the shapes of isotherm with and without sorption plateaus (Fig. 1, Table III), has potential to lead to the recommendation that Mehlich-3 Al can be used to inform P management if there is no access to P sorption data. This Mehlich-3 Al signal may be showing that when a soil sample is in the high or low range sets of Mehlich-3 Al, it is also exhibiting P sorption mechanisms consistent with C (rapid sorption and low P final solution P concentration) or, L non strict and L strict raw isotherms (Gradually decrease in sorption as added P increases).This interpretation is a step towards understanding the mechanism of different P sorption isotherm shapes. In contrast, Mehlich-3 Fe did not influence the shape of the isotherm curve in Irish soils. However, Fe was shown to be a significant driver of 3 isotherm parameters (Langmuir $S\text{max}_{50}$, Freundlich n_{50} and Freundlich K_{50}) when used with other parameters in multiple linear regression. This shows that Fe perhaps influences isolated sorption mechanisms, but

not the overall shape of the curve which indicates a broader mechanism. In this study, Mehlich-3 Al influenced the overall mechanism for a representative set of soils from Ireland.

CONCLUSIONS

The objective of this study was to develop relationships between P isotherm parameters and routine soil analyses that are quicker to determine, that could be used as reliable pedotransfer functions. Phosphorus sorption parameters, in the range, 0–50 mg P L^{-1}, could be predicted accurately from a short range 0–25 mg P L^{-1} isotherm parameters or single point concentration 50 mg P L^{-1}, S_{50}. A multiple linear regression prediction of the $Smax_{50}$ parameter was achieved using S_{50}, OM and Mehlich-3 Fe (R^2c = 0.91 and R^2v = 0.95). A multiple linear regression prediction of the MBC_{50} parameter was achieved with less success, but to a usable standard using S_{50}, OM, clay, Mehlich-3 Al (R^2c = 0.65 and R^2v = 0.71). These multiple regression combinations have the potential to be used as pedotransfer functions.

Simple, but accurate predictions based on wet chemistry also offered insight into P sorption behaviours of soil where the mechanisms were dominated by Al. Over a national population of Irish soils, the mechanism of P sorption was described L shape (non-strict) and C shape isotherm curves in 64% and 27% of soils, respectively. These indicate two main mechanisms for this population of soils, *i.e.*, rapid sorption of P, with low final P concentration after equilibration and no visible plateau (C shape); and, a slower sorption of P to soil defined by a concave isotherm, where sorption gradually decreases with increasing P concentrations in solution at equilibrium (L shape, non-strict). Ranges of Mehlich-3 Al were identified where these mechanisms can be defined. Soils within a low range of Mehlich-3 Al (2.5--698 mg kg^{-1}) had no incidence of C shape or rapid sorption mechanistic behaviour, whereas, soils with values between 702--2 460 mg kg^{-1} could be described by C or L shape isotherms.

Based on the data presented in this study, the following 3 recommendations emerged:

1. Single point determination, on its own, is quicker and can accurately predict the information that is usually obtained from the lengthy range type analysis (*i.e.*, 0–50 mg P L^{-1}).

2. Mehlich-3 Al has the potential to identify P sorption mechanism. There is potential for 2 different ranges of Al to help to identify mechanistic behaviour (with and without a plateau) without any prior knowledge of sorption parameters.

3. The minimum number of sorption parameters needed to gain insight into dynamics is 2, that is, one that describes sorption maximum and another that describes the strength with which the phosphate anion is bound. In order to stay true to sorption mechanisms and choose the simplest model we recommend the prediction of Langmuir parameters are included with agronomic soil testing to identify differences in P sorption behaviour in Irish soils to improve P fertiliser advice.

ACKNOWLEDGEMENTS

This research was funded by the Teagasc Walsh Fellowship Fund (RMIS 6502). The authors would like to thank Mr. Denis Brennan and Ms. Maria Pettitt for laboratory analysis of water samples and summer student Ms. Katie Lawlor for help in processing samples.

REFERENCES

Akaike, H., 1973. Information theory as an extension of the maximum likelihood principle. Á In: Petrov, BN and Csaki, F, Second International Symposium on Information Theory. Akademiai Kiado, Budapest, pp. 276Á281.

Alovisi, A.M.T., Cassol, C.J., Nascimento, J.S., Soares, N.B., da Silva Junior, I.R., da Silva, R.S., da Silva, J.A.M., 2020. Soil factors affecting phosphorus adsorption in soils of the Cerrado, Brazil. Geoderma Regional, e00298.

Bache, B., Williams, E., 1971. A phosphate sorption index for soils. Eur. J. Soil Sci. 22(3), 289-301.

Ballabio, C., Lugato, E., Fernández-Ugalde, O., Orgiazzi, A., Jones, A., Borrelli, P., Montanarella, L., Panagos, P., 2019. Mapping LUCAS topsoil chemical properties at European scale using Gaussian process regression. Geoderma 355, 113912.

Ballard, R., Fiskell, J.G., 1974. Phosphorus Retention in Coastal Plain Forest Soils: I. Relationship to Soil Properties 1. Soil Science Society of America Journal 38(2), 250-255.

Burgess, C., 2015. How Accurate are Your Dilutions? Pharmaceutical Technology 39(1).

Burkitt, L.L., Gourley, C.J.P., Sale, P.W.G., Uren, N.C., Hannah, M.C., 2001. Factors affecting the change in extractable phosphorus following the application of phosphatic fertiliser on pasture soils in southern Victoria. Aust. J. Soil Res. 39(4), 759-771.

Burkitt, L.L., Moody, P.W., Gourley, C.J.P., Hannah, M.C., 2002. A simple phosphorus buffering index for Australian soils. Aust. J. Soil Res. 40(3), 497-513.

Burkitt, L.L., Sale, P.W.G., Gourley, C.J.P., 2008. Soil phosphorus buffering measures should not be adjusted for current phosphorus fertility. Aust. J. Soil Res. 46(8), 676-685.

Burnham, K.P., Anderson, D.R., 2002. A practical information-theoretic approach. Model selection and multimodel inference, 2nd ed. Springer, New York.

Chowdhury, R.B., Moore, G.A., Weatherley, A.J., Arora, M., 2017. Key sustainability challenges for the global phosphorus resource, their implications for global food security, and options for mitigation. Journal of Cleaner Production 140, 945-963.

Cordell, D., Drangert, J.-O., White, S., 2009. The story of phosphorus: Global food security and food for thought. Global Environmental Change 19(2), 292-305.

Corrêa, R.M., Nascimento, C.W.A.d., Rocha, A.T.d., 2011. Adsorção de fósforo em dez solos do Estado de Pernambuco e suas relações com parâmetros físicos e químicos. Acta Scientiarum. Agronomy 33(1), 153-159.

Creamer, R., O'Sullivan, L., 2018. The soils of Ireland. 1st 2018 ed. Springer, Cham, Switzerland.

Creamer, R., Simo, I., Reidy, B., Carvalho, J., Fealy, R., Hallett, S., Jones, R., Holden, A., Holden, N., Hannam, J., 2014. Irish soil information system synthesis report. Teagasc Environment Research Centre, Johnstown Castle, Wexford, Co. Wexford, Ireland.

Creamer, R.E., Simo, I., O'Sullivan, L., Reidy, B., Schulte, R.P.O., Fealy, R.M., 2016. Research 204: Irish Soil Information System: Soil Property Maps. 978-1-84095-694-8, Ireland.

Daly, K., Jeffrey, D., Tunney, H., 2001. The effect of soil type on phosphorus sorption capacity and desorption dynamics in Irish grassland soils. Soil Use Manage. 17(1), 12-20.

Daly, K., Styles, D., Lalor, S., Wall, D.P., 2015. Phosphorus sorption, supply potential and availability in soils with contrasting parent material and soil chemical properties. Eur. J. Soil Sci. 66(4), 792-801.

Daly, K., Tuohy, P., Peyton, D., Wall, D.P., Fenton, O., 2017. Field soil and ditch sediment phosphorus dynamics from two artificially drained fields on poorly drained soils. Agric. Water Manage. 192, 115-125.

Demand, D., Schack-Kirchner, H., Lang, F., 2017. Assessment of diffusive phosphate supply in soils by microdialysis. J. Plant Nutr. Soil Sci. 180(2), 220-230.

Devriendt, N., Or, M., Meyer, E., Paepe, D., Vallarino, N., Bhatti, S.F., De Rooster, H., 2019. Comparative accuracy and precision of two commercial laboratory analyzers for the quantification of ammonia in cerebrospinal fluid. Veterinary clinical pathology.

Dunne, K.S., Holden, N.M., O'Rourke, S.M., Fenelon, A., Daly, K., 2020. Prediction of phosphorus sorption indices and isotherm parameters in agricultural soils using mid-infrared spectroscopy. Geoderma 358.

Fan, B., Wang, J., Fenton, O., Daly, K., Ezzati, G., Chen, Q., 2019. Strategic differences in phosphorus stabilization by alum and dolomite amendments in calcareous and red soils. Environ. Sci. Pollut. Res. 26(5), 4842-4854.

FAO, 2014. Building a common vision for sustainable food and agriculture: Principles and approaches. FAO Rome.

Fernandes, M.L.V., Warren, G.P., 1994. Exchangeable and non-exchangeable phosphate sorption in Portuguese soils. Fertiliser Research 37(1), 23-34.

Fink, J.R., Inda, A.V., Bavaresco, J., Barrón, V., Torrent, J., Bayer, C., 2016. Adsorption and desorption of phosphorus in subtropical soils as affected by management system and mineralogy. Soil and Tillage Research 155, 62-68.

Gaines, R.D., 1980. Dilution as a source of error: implications for preparation and calibration of laboratory standards and for quality control of radioimmunoassays. Clinical chemistry 26(12), 1726-1729.

Giles, C.H., Smith, D., Huitson, A., 1974. A general treatment and classification of the solute adsorption isotherm. I. Theoretical. Journal of Colloid And Interface Science 47(3), 755-765.

Guppy, C.N., Menzies, N.W., Moody, P.W., Blamey, F.P.C., 2005. Competitive sorption reactions between phosphorus and organic matter in soil: a review. Aust. J. Soil Res. 43(2), 189-202.

Hallett, S.H., Baillie, I.C., Kerr, B., Truckell, I., 2011. Development of the World Soil Survey Archive and Catalogue (WOSSAC). Commission on the History, Philosophy and Sociology of Soil Science 18, 14-17.

Holden, N.M., White, E.P., Lange, M.C., Oldfield, T.L., 2018. Review of the sustainability of food systems and transition using the Internet of Food. NPJ science of food 2(1), 1-7.

Indiati, R., Neri, U., Sharpley, A.N., Fernandes, M.L., 1999. Extractability of added phosphorus in short-term equilibration tests of portuguese soils. Communications in Soil Science and Plant Analysis 30(13-14), 1807-1818.

John, M.K., 1970. Colorimetric determination of phosphorus in soil and plant materials with ascorbic acid. Soil Sci. 109(4), 214-220.

Kendall, M.G., 1938. A new measure of rank correlation. Biometrika 30(1/2), 81-93.

Khaledian, Y., Miller, B.A., 2020. Selecting appropriate machine learning methods for digital soil mapping. Applied Mathematical Modelling 81, 401-418.

Lalor, S., Coulter, B., 2008. Major and minor micronutrient advice for productive agricultural crops. Teagasc, Oak Park, Carlow.

Lam, R., Isenhour, T., 1980. Minimizing relative error in the preparation of standard solutions by judicious choice of volumetric glassware. Analytical Chemistry 52(7), 1158-1161.

Limousin, G., Gaudet, J.P., Charlet, L., Szenknect, S., Barthes, V., Krimissa, M., 2007. Sorption isotherms: A review on physical bases, modeling and measurement. Appl. Geochem. 22(2), 249-275.

Liu, J., Tang, W., Chen, G., Lu, Y., Feng, C., Tu, X.M., 2016. Correlation and agreement: overview and clarification of competing concepts and measures. Shanghai Archives of Psychiatry 28(2), 115-120.

Mann, H.B., Whitney, D.R., 1947. On a test of whether one of two random variables is stochastically larger than the other. The annals of mathematical statistics, 50-60.

Mikutta, C., Lang, F., Kaupenjohann, M., 2006. Citrate impairs the micropore diffusion of phosphate into pure and C-coated goethite. Geochimica et Cosmochimica Acta 70(3), 595-607.

Mohr, S., Evans, G., 2013. Projections of future phosphorus production. Philicacom 2013.

Mozaffari, M., Sims, J.T., 1994. Phosphorus availability and sorption in an Atlantic Coastal Plain watershed dominated by animal-based agriculture. Soil Sci. 157(2), 97-107.

Nair, P., Logan, T., Sharpley, A., Sommers, L., Tabatabai, M., Yuan, T., 1984. Interlaboratory comparison of a standardized phosphorus adsorption procedure1. Journal of Environmental Quality 13(4), 591-595.

Neset, T.S.S., Cordell, D., 2012. Global phosphorus scarcity: Identifying synergies for a sustainable future. Journal of the Science of Food and Agriculture 92(1), 2-6.

Niven, E.B., Deutsch, C.V., 2012. Calculating a robust correlation coefficient and quantifying its uncertainty. Computers & Geosciences 40, 1-9.

Ozanne, P., Shaw, T., 1968. Advantages of the recently developed phosphate sorption test over the older extractant methods for soil phosphate. In 'Transactions of the 9th International Congress in Soil Science'. Adelaide. (Ed. JW Holmes) pp. 273–280. (International Society of Soil Science)

Padarian, J., Minasny, B., McBratney, A.B., 2020. Machine learning and soil sciences: A review aided by machine learning tools. SOIL 6(1), 35-52.

Pachepsky, Y.A., van Genuchten, M.T., 2011. Pedotransfer Functions. In: J. Gliński, J. Horabik, J. Lipiec (Eds.), Encyclopedia of Agrophysics. Springer Netherlands, Dordrecht, pp. 556-561.

Pautler, M.C., Sims, J.T., 2000. Relationships between soil test phosphorus, soluble phosphorus, and phosphorus saturation in Delaware soils. Soil Science Society of America Journal 64(2), 765-773.

Pierzynski, G.M., 2000. Methods of phosphorus analysis for soils, sediments, residuals, and waters. North Carolina State University Raleigh.

R Core Team, 2019. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Rogeri, D.A., Gianello, C., Bortolon, L., Amorim, M.B., 2016. Substitution of clay content for P-Remaining as an index of the phosphorus buffering capacity for soils of Rio Grande do Sul. Rev. Bras. Cienc. Solo 40.

Rosset, S., Perlich, C., Zadrozny, B., 2007. Ranking-based evaluation of regression models. Knowledge and Information Systems 12(3), 331-353.

RStudio Team, 2015. RStudio: Integrated Development for R. RStudio, Inc., Boston, MA. http://www.rstudio.com/.

Sollins, P., Homann, P., Caldwell, B.A., 1996. Stabilization and destabilization of soil organic matter: Mechanisms and controls. Geoderma 74(1-2), 65-105.

Steinman, A.D., Ogdahl, M.E., 2016. From wetland to farm and back again: phosphorus dynamics of a proposed restoration project. Environ. Sci. Pollut. Res. 23(22), 22596-22605.

Stenberg, B., Viscarra Rossel, R.A., Mouazen, A.M., Wetterlind, J., 2010. Visible and Near Infrared Spectroscopy in Soil Science, Advances in Agronomy, pp. 163-215.

Torrent, J., Schwertmann, U., Barrón, V., 1992. Fast and slow phosphate sorption by goethite-rich natural materials. Clays and Clay Minerals 40(1), 14-21.

Uygur, V., 2009. Phosphate Sorption in Calcareous Soils: The Role of Iron Oxide and Carbonates. Asian Journal of Chemistry 21, 3001-3009.

Wan, M., Hu, W., Qu, M., Li, W., Zhang, C., Kang, J., Hong, Y., Chen, Y., Huang, B., 2020. Rapid estimation of soil cation exchange capacity through sensor data fusion of portable XRF spectrometry and Vis-NIR spectroscopy. Geoderma 363.

Wang, Y.T., Zhang, T.Q., O'Halloran, I.P., Tan, C.S., Hu, Q.C., 2016. A phosphorus sorption index and its use to estimate leaching of dissolved phosphorus from agricultural soils in Ontario. Geoderma 274, 79-87.

Wilcoxon, F., 1945. Individual Comparisons by Ranking Methods. Biometrics Bulletin 1(6), 80-83.

Withers, P.J.A., Neal, C., Jarvie, H.P., Doody, D.G., 2014. Agriculture and eutrophication: Where do we go from here? Sustainability (Switzerland) 6(9), 5853-5875.

Xu, D., Xu, J., Wu, J., Muhammad, A., 2006. Studies on the phosphorus sorption capacity of substrates used in constructed wetland systems. Chemosphere 63(2), 344-352.

Xue, Q.Y., Lu, L.L., Zhou, Y.Q., Qi, L.Y., Dai, P.B., Liu, X.X., Sun, C.L., Lin, X.Y., 2014. Deriving sorption indices for the prediction of potential phosphorus loss from calcareous soils. Environ. Sci. Pollut. Res. 21(2), 1564-1571.

TABLE I

Regression based prediction of 0 – 50 mg l^{-1} P Langmuir and Freundlich isotherm parameters from the raw isotherm maximum of the 0 – 50 mg l^{-1} P range (i.e. using a sorption single point as a predictor). Samples were randomly split 75 % for the calibration set and 25 % for the validation set. For each parameter (k_{50}, $Smax_{50}$, MBC_{50}, n_{50} and K_{50}), the model that gave the best outcome, with regard to parsimony and consistency between validation and calibration, is highlighted in bold

method	parameter and preprocessing	n	range (unit)	median	true sd	cor (Pearson)	cor (Kendall)	R^2c	RMSEc	R^2v	RMSEv
[b]2nd	k_{50} (1)	178	0.03 - 45.44 (l mg^{-1})	0.70	3.46	0.21	0.60	0.05	3.85	0.12	1.35
	k_{50} (2)	**168**	**0.09 - 2.07**	**0.69**	**0.47**	**0.73**	0.60	**0.63**	**0.28**	**0.56**	**0.33**
	k_{50} (3)	174	0.03 - 3.15	0.70	0.59	0.68	0.61	0.58	0.37	0.40	0.50
	k_{50} (4)	174	0.03 - 3.15	0.70	0.59	0.75	0.61	0.62	0.46	0.43	0.50
[a]lm	$Smax_{50}$ (1)	178	274.4 - 1036.9 (mg kg^{-1})	656.8	149.00	0.86	0.72	0.70	80.78	0.81	66.43
	$Smax_{50}$ (2)	176	274.4 - 999.8	656.8	144.85	0.85	0.72	0.70	78.84	0.81	66.62
	$Smax_{50}$ (3)	**170**	**274.4-954.3**	**647.0**	**142.46**	**0.94**	**0.79**	**0.88**	**47.47**	**0.86**	**58.42**
	$Smax_{50}$ (4)	170	274.4-954.3	647.0	142.46	0.93	0.79	0.86	0.09	0.86	0.10
[b]2nd	MBC_{50} (1)	178	10.68-31706.70 (l kg^{-1})	457.91	2421.14	0.24	0.70	0.07	2709.00	-1.35	473.92
	MBC_{50} (2)	**167**	**39.49-1603.99**	**440.71**	**374.26**	**0.79**	**0.70**	**0.70**	**216.50**	**0.70**	**162.13**
	MBC_{50} (3)	169	10.68-1819.43	440.71	392.54	0.80	0.72	0.74	201.20	0.77	189.04
	MBC_{50} (4)	169	10.68 - 1819.43	440.71	392.54	0.87	0.72	0.76	0.43	0.83	0.35
[b]2nd	n_{50} (1)	196	0.20-1.13 (l mg^{-1})	0.48	0.15	0.34	0.35	0.35	0.12	0.01	0.18
	n_{50} (2)	192	0.20 - 0.88	0.48	0.14	0.40	0.35	0.25	0.12	0.20	0.12
	n_{50} (3)	**192**	0.20-**0.91**	**0.48**	**0.14**	**0.53**	**0.40**	**0.32**	**0.12**	**0.30**	**0.12**
	n_{50} (4)	192	0.20-0.91	0.48	0.14	0.49	0.40	0.30	0.12	0.28	0.12
[b]2nd	K_{50} (1)	196	6.35 - 858.06 (mg kg^{-1})	226.12	138.73	0.77	0.69	0.70	81.82	0.55	70.62
	K_{50} (2)	190	6.35 - 581.89	224.61	115.10	0.80	0.68	0.70	66.52	0.63	56.48
	K_{50} (3)	**190**	**6.35 - 599.88**	**224.61**	**116.82**	**0.82**	**0.70**	**0.73**	**58.93**	**0.78**	**59.34**
	K_{50} (4)	190	6.35 - 599.88	224.61	116.82	0.76	0.70	0.69	63.81	0.74	65.25

method, regression method; [a], lm, linear model; [b], 2nd, second degree polynomial; parameter and preprocessing, (1), none, (2), outliers removed by boxplot, (3), outliers removed by visual assessment of residuals, (4), outliers removed by visual assessment of residuals and the data was logged; k_{50}, Langmuir binding energy; $Smax_{50}$, Langmuir sorption maximum; MBC_{50}, maximum buffer capacity; n_{50}, Freundlich affinity constant; K_{50}, Freundlich sorption maximum; true sd, standard deviation of the 0 – 50 mg P l^{-1} parameter when determined using wet chemistry; cor (Pearson) is Pearson's correlation coefficient between the values of the 0–25 mg P l^{-1} linearised parameter and S_{50}; cor (Kendall) is Kendall's *tau* between the values of the 0–25 mg P L^{-1} linearised parameter and S_{50}.

TABLE II

Normal Multiple Regression Predictions of Langmuir parameters in the 0 – 50 mg l^{-1} P range using the resulting formula from Stepwise Multiple Regression (direction used was both forward and backward until Akaike Information Criterion (AIC) was lowest for all parameters except MBC_{50} (2)). Samples were randomly split 75 % for the calibration set and 25 % for the validation set

parameter	n	range (unit)	stepwise linear regression resulting formula and coefficients	R^2c adj	RMSEc	R^2v	RMSEv
k_{50}	167	0.09--2.07 (L mg^{-1})	Lk_{50} ~ S_{50}*** + OM** + Clay* + Mehlich-3 Al*	0.59	0.30	0.67	0.27
			Intercept = -0.613, S_{50} = 0.003, OM = -0.021, clay = -0.007, Mehlich-3 Al = 0.0001				
$Smax_{50}$	169	274.4--954.3 (mg kg^{-1})	$Smax_{50}$ ~ S_{50}*** + OM*** + Mehlich-3 Fe	0.91	40.17	0.95	35.77
			Intercept = 19.236, S_{50} = 0.972, OM = 7.922, Mehlich-3 Fe = 0.042				
MBC_{50}	168	10.68--1819.43 (L kg^{-1})	MBC_{50} ~ S_{50}*** + CEC + clay + Mehlich-3 Al	0.66	232.30	0.70	207.19
			Intercept = -693.956, S_{50} = 2.485, CEC = -6.145, clay = -4.151, Mehlich-3 Al = 0.121				
MBC_{50} (2)	168	10.68--1819.43 (L kg^{-1})	MBC_{50} ~ S_{50}*** + OM + clay * + Mehlich-3 Al	0.65	235.9	0.71	204.18
			Intercept = -733.564, S_{50} = 2.572, OM = -5.201, clay = -6.421, Mehlich-3 Al = 0.146				
n_{50}	181	0.20--0.91 (L mg^{-1})	Fn_{50} ~ S_{50}*** + CEC* + clay* + OM*** + Mehlich-3 Fe**	0.43	0.11	0.17	0.11
			Intercept = 0.242, S_{50} = 0.0004, CEC = -0.003, clay = 0.002, OM = 0.011, Mehlich-3 Fe = -0.0002				
K_{50}	185	6.35--599.88 (mg kg^{-1})	FK_{50} ~ S_{50}*** + OM** + clay + Mehlich-3 Al . + Mehlich-3 Fe	0.74	61.06	0.74	58.11
			Intercept = -167.807, S_{50} = 0.783, OM = -2.768, clay = -1.307, Mehlich-3 Al = 0.032, Mehlich-3 Fe = 0.063				

k_{50}, Langmuir binding energy; $Smax_{50}$, Langmuir sorption maximum; MBC_{50}, maximum buffer capacity; n_{50}, Freundlich affinity constant; K_{50}, Freundlich sorption maximum; $S_{50,}$ P sorbed at 50 mg l^{-1} which is the amount of P sorbed from a single addition of phosphorus at 50 mg l^{-1} concentration; OM, organic matter measured as loss on ignition; CEC, cation exchange capacity; Mehlich-3 Al, Mehlich-3 aluminium; Mehlich-3 Fe, Mehlich-3 iron; significance codes: 0 '***', 0.001 '**', 0.01 '*', 0.05 '.', 0.1 ' '.

TABLE III

Summary statistics from set 1 and set 2 of hierarchical clustering. To determine the cause of the isotherm shape difference, twenty samples from the top of the raw isotherm overlay plot and twenty samples from the bottom of the raw isotherm overlay plot (n =40) were clustered and examined according to 13 ancillary parameters to give set 1a and 2a. All samples with available ancillary data (n = 175) were clustered to give set 1b and set 2b

parameter (unit)	range	median	range	median	W	p-value
	hierarchical set 1a (n = 22, 17 C, 3 L non strict, 1 L strict and 1 S shape)		hierarchical set 2a (n = 18, 11 L non strict and 7 L strict shape)		Wilcoxon test (set 1a v set 2a)	
Mehlich-3 Al (mg kg^{-1})	866.6 - 1623.4	1116.8	272.7 - 857.6	706.6	396	0.0

Mehlich-3 Fe (mg kg^{-1})	149.3 - 541.8	395.1	137.6 - 504.9	356.5	218	0.6
Mehlich-3 P (mg kg^{-1})	5.6- 105.9	33.8	14.6 - 157.9	53.5	152	0.2
depth to (cm)	8.0 - 37.0	20.5	6.0 - 33.0	23.0	163	0.3
organic matter (%)	6.9 - 32.3	13.3	4.3 - 17.5	10.18	276	0.0
Mehlich-3 Al / P	12.7 - 166.1	39.6	1.9 - 48.0	15.0	314	0.0
sand (% w / w)	8.0 - 61.0	39.5	26.0 - 75.0	58.5	84.5	0.0
silt (% w / w)	23.0 - 64.0	33.0	14.0 - 43.0	26.0	289.5	0.0
clay (% w / w)	12.0 - 44.0	27.0	8.0 - 34.0	16.0	306.5	0.0
pH in water	5.0 - 7.3	5.6	5.0 - 7.3	6.0	147.5	0.2
CEC (cmol kg^{-1})	6.5 - 19.6	13.67	2.6 - 25.3	16.0	155	0.3
Ca (cmol kg^{-1})	3.5 - 18.4	10.6	2.2 - 22.5	12.8	138	0.1
	hierarchical set 1b (n = 157, 107 L non strict, 36 C, 13 L strict and 1 S shape)		hierarchical set 2b (n = 18, 17 L non strict and 1 L strict shape)		Wilcoxon test (set 1b v set 2b)	
Mehlich-3 Al (mg kg^{-1})	702.5 - 2460.8	1030.1	2.5 - 698.5	526.0	2826	0.0
Mehlich-3 Fe (mg kg^{-1})	36.5 - 664.4	367.2	137.6 - 504.9	352.5	1632	0.3
Mehlich-3 P (mg kg^{-1})	4.9 - 564.3	41.1	5.1 - 157.9	43.0	1490	0.7
depth to (cm)	5.0 - 55.0	21.0	6.0 - 40.0	20.0	1483	0.7
organic matter (%)	4.5 - 37.1	10.6	4.3 - 22.1	9.52	1587	0.4
Mehlich-3 Al / P	2.5 - 320.5	24.6	0.1 -99.7	13.7	2040	0.0
sand (% w / w)	8.0 - 79.0	44.0	18.0 - 85.0	54.5	1245.5	0.4
silt (% w / w)	14.0 - 64.0	33.0	10.0 - 43.0	28.5	1583.5	0.4
clay (% w / w)	7.0 - 46.0	22.0	5.0 - 46.0	18.5	1564.5	0.5
pH in water	3.9 - 7.8	5.9	5.0 - 8.3	6.6	869.5	0.0
CEC (cmol kg^{-1})	2.6 - 47.5	12.8	2.6 - 42.5	19.7	831	0.0
Ca (cmol kg^{-1})	0.2 - 41.5	10.3	2.2 - 46.4	20.1	738	0.0

Wilcoxon test, wilcox.test in RStudio where, paired = FALSE and alternative = two.sided; W, a Wilcoxon rank-sum test statistic

Figure captions

Fig. 1. Irish Soil Information System (SIS) soils (n = 224) depicted using Gile's main shapes of isotherm. Final C is on the x-axis and it is final concentration of P in solution after 24 hours equilibration time with 2 g soil (mg l^{-1}) and P sorbed is on the y-axis and it is the amount of P sorbed to 2 g soil samples after 24 hours equilibration time (mg kg^{-1}). Shapes; "C" which is coloured light blue, "L non-strict" which is coloured orange, "L strict" which is coloured yellow and "S" which is coloured navy, appeared in this sample set, H did not.

Fig. 2. Independent validation plot depicting predicted versus actual values for the best Langmuir parameter that was predicted from a smaller range (MBC_{50} explained by MBC_{25}). This parameter was predicted using 2nd degree polynomial regression. The a, b line was set to intercept = 0 and slope = 1 for demonstration of correlation, it is not equivalent to the model.

Fig. 3. Independent validation plot depicting predicted versus actual values for the best Freundlich parameter that was predicted from a smaller range (n_{50} explained by n_{25}). This parameter was predicted using normal linear regression. The a, b line was set to intercept = 0 and slope = 1 for demonstration of correlation, it is not equivalent to the model.

Fig. 4. Independent validation plot depicting predicted v actual values for the best isotherm parameter that was predicted from a single point value (0 – 50 mg P l^{-1} Linearised Langmuir Smax explained by P sorbed at 50 mg P l^{-1}). This parameter was predicted using normal linear regression. The a , b line was set to intercept = 0 and slope = 1 for demonstration of correlation, it is not equivalent to the model.

Fig. 5. Independent validation plot depicting predicted versus actual values for the best isotherm parameter that was predicted from a normal multiple linear regression (Linearised Langmuir $Smax_{50}$ explained by P sorbed at 50 mg P l^{-1}, OM and Mehlich-3 Fe). The a , b line was set to intercept = 0 and slope = 1 for demonstration of correlation, it is not equivalent to the model.

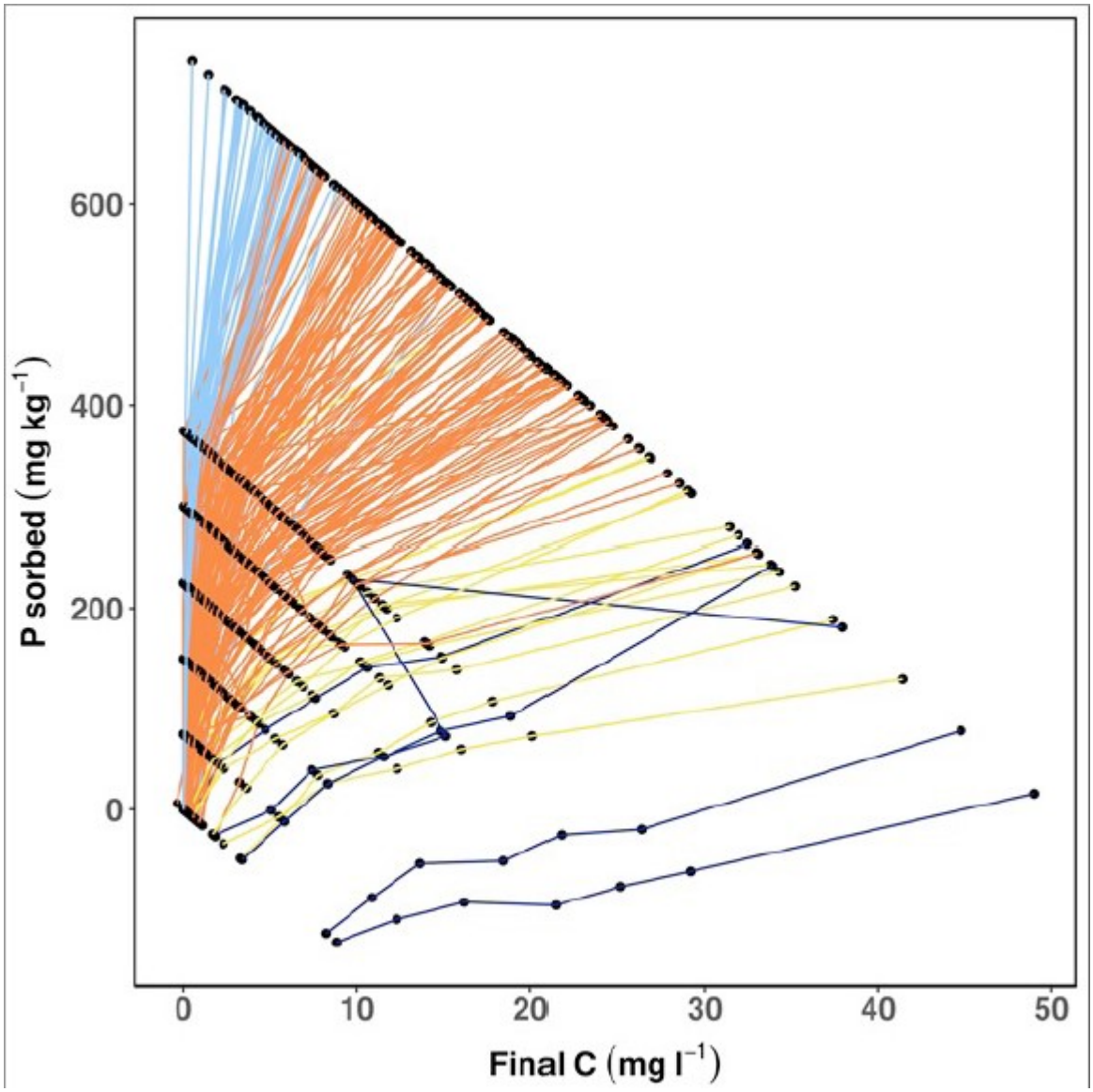


Fig. 1

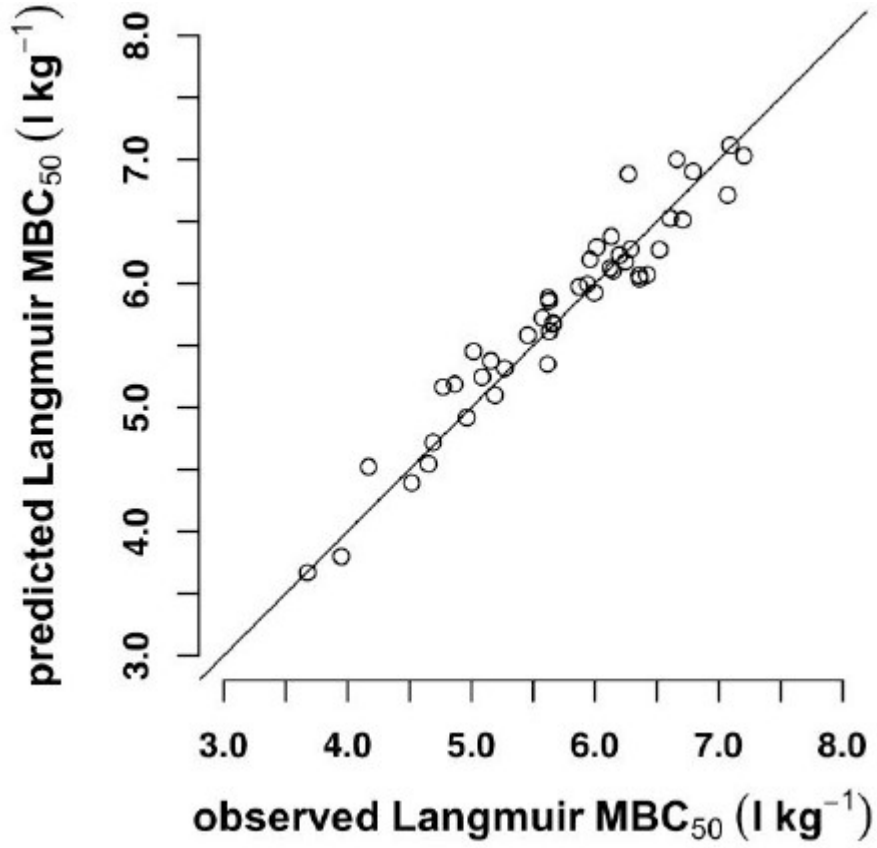

predicted Langmuir MBC50 (l kg−1)
3.0
4.0
5.0
6.0
7.0
8.0
observed Langmuir MBC50 (l kg−1)

Fig. 2

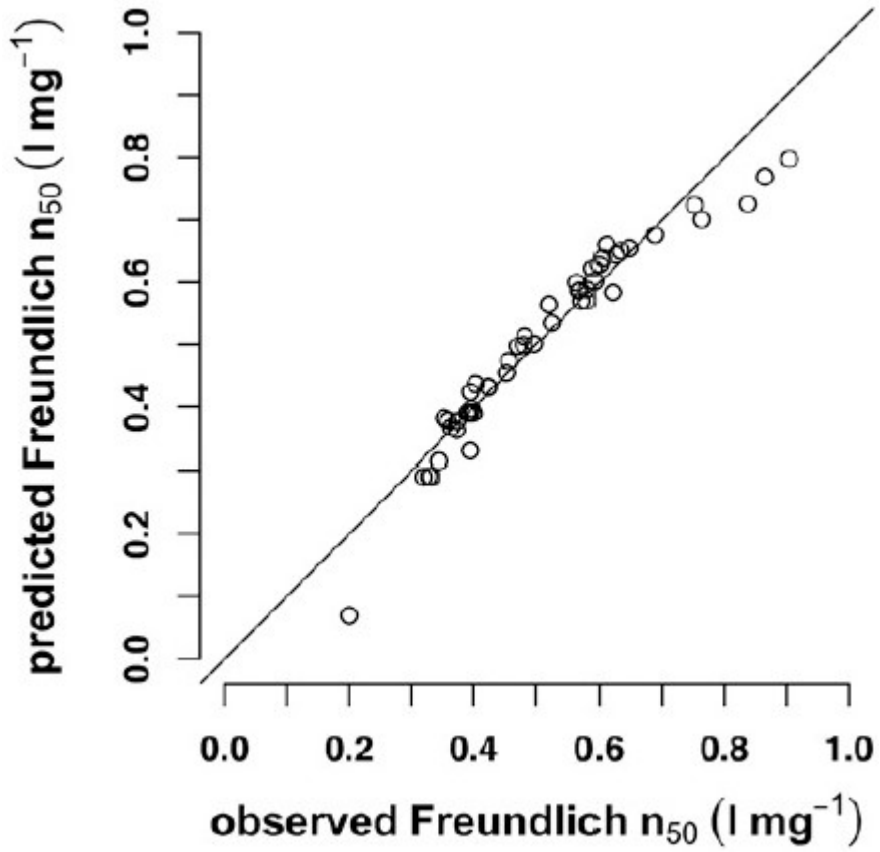

predicted Freundlich n_{50} (l mg^{-1})
0.0
0.2
0.4
0.6
0.8
1.0
observed Freundlich n_{50} (l mg^{-1})

Fig. 3

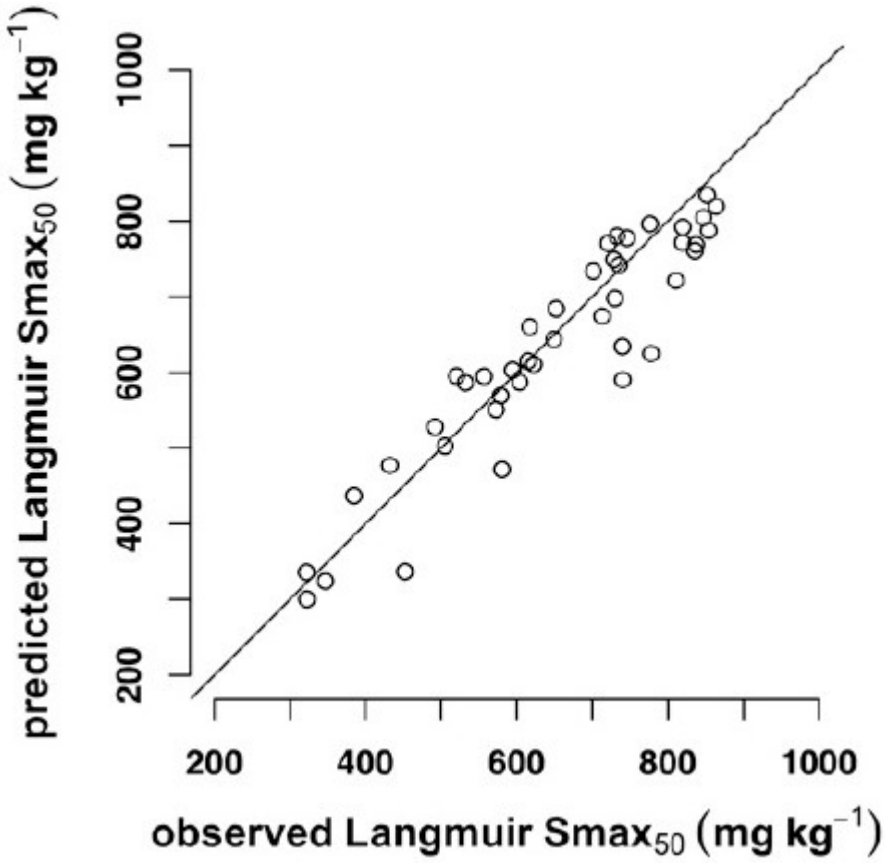
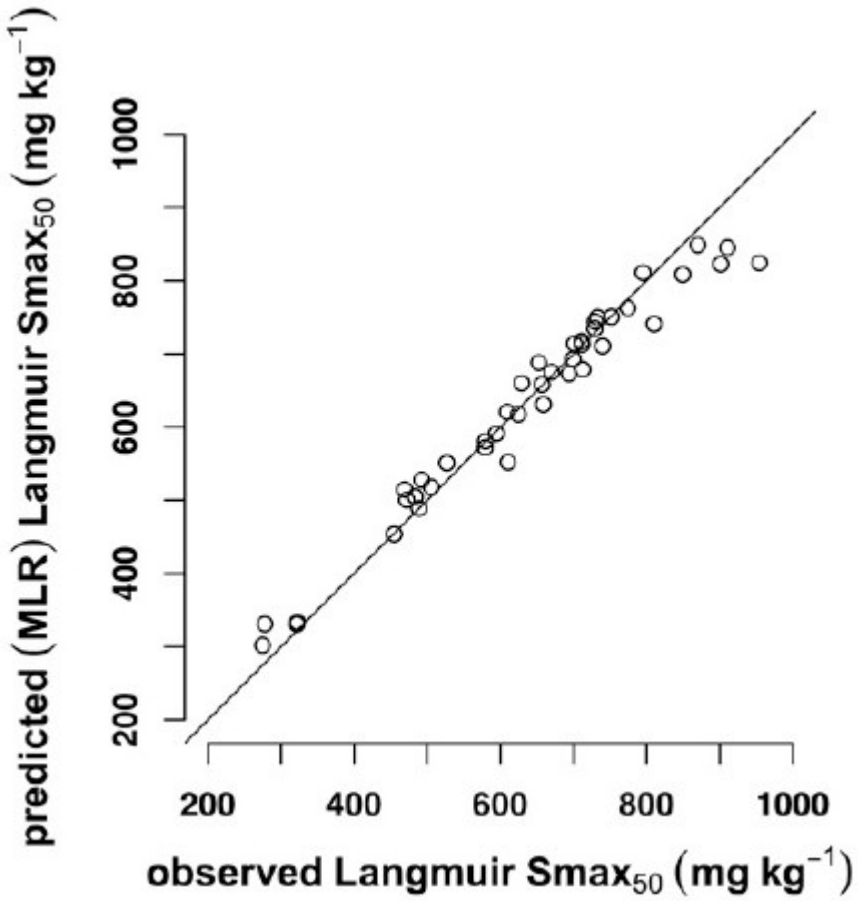

predicted Langmuir Smax50 (mg kg−1)
200
400
600
800
1000
observed Langmuir Smax50 (mg kg−1)

Fig. 4


predicted (MLR) Langmuir Smax50 (mg kg−1)
200
400
600
800
1000
observed Langmuir Smax50 (mg kg−1)
Fig. 5