Meta-analysis and machine learning to explore soil-water partitioning of common pharmaceuticals

Andrea-Lorena Garduño-Jiménez [a], Juan-Carlos Durán-Álvarez [b], Rachel Louise Gomes [a,*]

[a] *Food Water Waste Research Group, Faculty of Engineering, University of Nottingham, NG7 2RD, United Kingdom*
[b] *Instituto de Ciencias Aplicadas y Tecnología, Universidad Nacional Autónoma de México, Circuito Exterior S/N, Ciudad Universitaria, Coyoacan, Ciudad de México 04510, Mexico*

HIGHLIGHTS

- First pharmaceuticals soil/water partitioning meta-dataset modelling
- Modelling of meta-dataset shows outcomes not identified in individual studies.
- More environmentally relevant conditions recommended for batch-sorption studies.

GRAPHICAL ABSTRACT
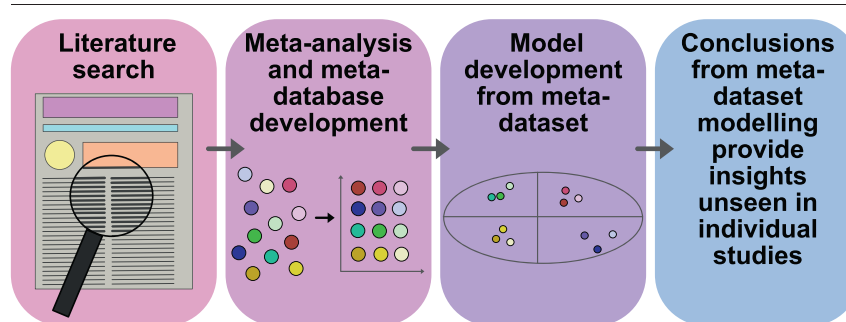
ARTICLE INFO

ABSTRACT

The first meta-analysis and modelling from batch-sorption literature studies of the soil/water partitioning of pharmaceuticals is presented. Analysis of the experimental conditions reported in the literature demonstrated that though batch-sorption studies have value, they are limited in evaluating partitioning under environmentally-relevant conditions. Recommendations are made to utilise environmental relevant pharmaceutical concentrations, perform batch-sorption studies at temperatures other than 4, 20 and 25 °C to better reflect climate diversity, and utilise the Guideline 106 methodology as a benchmark to enable comparison between future studies (and support modelling and prediction). The meta-dataset comprised 82 data points, which were modelled using multivariate analysis; where K_d (soil/water partitioning coefficient) was the independent variable. The dependent variables fit into three categories: 1) pharmaceutical studied (including physical-chemical properties), 2) soil characteristics and 3) experimental conditions. The pharmaceutical solubility, the soil/liquid equilibration time (prior to adding the pharmaceutical), the soil organic carbon, the soil sterilisation method and the liquid phase were found to be significantly important variables for predicting K_d.

1. Introduction

The higher food demand due to increasing human population (UN-WWAP, 2015), as well as poverty and unequal distribution of resources (Holt-Giménez et al., 2012) highlights the necessity for more efficient and sustainable agricultural practices. Wastewater reuse for irrigation can relieve environmental and economic pressures of agriculture and increase year-round water accessibility (Holt-Giménez et al., 2012; Jovanovic, 2008; Jaramillo and Restrepo, 2017; Jimenez-Cisneros, 2006). Considered widespread practice, 65% of irrigated croplands are located in wastewater-dependent catchments, of which 82% are located in countries where <75% of wastewater is treated (Thebo et al., 2017). At the global scale, 80% of wastewater is not treated adequately and is used to irrigate 11% of croplands on average, with higher proportions for areas such as Latin America, South Asia and Africa (Kookana et al., 2020).

Wastewater contains a variety of pollutants, with pharmaceuticals occurring at unregulated levels in both treated and untreated wastewater (Petrie et al., 2015). The presence and environmental fate of pharmaceuticals in

soils, following irrigation with treated and intreated wastewater has implications to leaching to ground/surface waters (Paz et al., 2016; Llamas et al., 2020), bioavailability (Goulas et al., 2018; Li et al., 2019) and uptake by plants, development of antimicrobial resistance (AMR) (Cen et al., 2020; Marti et al., 2014), and entering the wider food network (Azanu et al., 2018; de Santiago-Martín et al., 2020; Kodesova et al., 2019). Given the widespread practice of wastewater irrigation, as well as the negative effects from pharmaceuticals in wastewaters, there is a pressing need to understand the fate of pharmaceuticals in agro-ecosystems.

There exist discrepancies when comparing sorption studies in the literature. For example, the variability of sulfonamides sorbing to soil (Martínez-Hernández et al., 2016; Andriamalala et al., 2018a) has been attributed to the extent that these antibiotics engage in covalent cross-coupling, which depends on the abundance and availability of phenol-oxidase enzymes and reactive manganese oxide surfaces, as well as the nature of the soil organic matter (Bialk et al., 2005). There are sorption behaviours which can only be observed at field or column level experiments, such as the sorption affinity of sulfamethoxazole and carbamazepine being observed to decrease with soil depth (correlating to decreasing organic matter content) (Kočárek et al., 2016). Discrepancies will always be reported when comparing laboratory-scale studies and field observations as it is not possible to fully emulate field conditions in the laboratory. However, it is important to identify the experimental conditions that have the most effect on pharmaceuticals soil/water partitioning in order to make laboratory-scale studies as environmentally relevant as possible. One way of prioritising variables important for further study which may help to bridge the gap between laboratory and field studies is to model a meta-dataset encompassing a range of experimental conditions.

Modelling of pharmaceutical sorption behaviour in soil has to date only been undertaken where results were also obtained within the same research group (Li et al., 2020; Conde-Cid et al., 2020). These modelling studies on their own data offer value through the mechanistic insights they provide on sorption. However, a meta-dataset comprising data from across multi-lab studies can give insights into wider trends, identify literature gaps, and unveil experimental conditions, which may be disproportionally impacting on the sorption trends observed. Production and utilising met-datasets has been undertaken for the sorption of per and polyfluoroalkyl substances (PFASs) in soil but yet to be carried out for pharmaceuticals in soil/water (Li et al., 2018a).

To model a meta-dataset, it is necessary to have sufficient and comparable data, which also aligns to modelling needs, especially with predictive capability. Most sorption studies evaluating the partitioning of pharmaceuticals to soil refer to batch sorption experiments that follow the Organisation for Economic Co-operation and Development (OECD) 106 Guideline (Table 1). The similar experimental conditions and standardisation of variables reported (e.g., soil texture) allows for comparability across the differing pharmaceuticals (type, concentration) and/or soils used in these studies. Column percolation experiments have been shown to be more reproducible and comparable to real world environmental conditions than batch studies (Wojsławski et al., 2019). Moreover, there is a wide variety of column experiments setups, which include transport of pharmaceuticals through undisturbed soil columns with no plants being grown (Durán-Álvarez et al., 2014) and lysimeters (i.e., a soil column where a plant is being grown) (Paz et al., 2016; Koroša et al., 2020). Some of the variations in soil column or lysimeter experiments are quantitative such as the lysimeter dimensions (e.g., from 0.1 m (Paz et al., 2016) to 5 m depth (Koroša et al., 2020)), and these quantitative variations are easy to model. In contrast, many other variations are qualitative and more challenging to model. For example, the types of plants being grown in a lysimeter (e.g., grass surrounded by mixed forest (Koroša et al., 2020), individually

Table 1

Literature used for meta-analysis and numbering system for the references used for model development.

#	Reference study obtained from literature search	Stated experiments are based on organisation for economic co-operation and development (OECD) Guideline 106
1	Martínez-Hernández, V, Meffe, R, Herrera López, S & de Bustamante, S, 2016, 'The role of sorption and biodegradation in the removal of acetaminophen, carbamazepine, caffeine, naproxen and sulfamethoxazole during soil contact: A kinetics study', *Science of the Total Environment,* vol. 559, 232–241. DOI: https://doi.org/10.1016/J.SCITOTENV.2013.11.036	Yes
2	Durán-Álvarez, J C, Prado-Pano, B & Jiménez-Cisneros, B, 2012, 'Sorption and desorption of carbamazepine, naproxen and triclosan in a soil irrigated with raw wastewater: Estimation of the sorption parameters by considering the initial mass of the compounds in the soil', *Chemosphere,* vol. 88, pp. 84–90. doi: https://doi.org/10.1016/j.chemosphere.2012.02.067	Yes
3	Liang, C, Lan, Z, Zhang, X & Liu, Y, 2016 'Mechanism for the primary transformation of acetaminophen in a soil/water system', *Water Research,* vol.98, pp. 215–224. doi: https://doi.org/10.1016/J.WATRES.2016.04.027	No
4	Paz, A, Tadmor, G, Malchi, T, Blotevogel, J, Borch, T, Polubesova, T & Chefetz, B, 2016, 'Fate of carbamazepine, its metabolites, and lamotrigine in soils irrigated with reclaimed wastewater: Sorption, leaching and plant uptake', *Chemosphere,* vol. 160, pp. 22–29. doi: https://doi.org/10.1016/J.CHEMOSPHERE.2016.06.048	No
5	Li, J., Carter, L J & Boxall, A B A, 2020, 'Evaluation and development of models for estimating the sorption behaviour of pharmaceuticals in soils', *Journal of Hazardous Materials*, vol. 392, pp. 122469. doi: https://doi.org/10.1016/j.jhazmat.2020.122469	Yes
6	Conde-Cid, M, Ferreira-Coelhob, G, Núñez-Delgadob, A, Fernández-Calviñoa, D, Arias-Estéveza, M, Álvarez-Rodríguezb, E & Fernández-Sanjurjob, M J, 2019, 'Competitive adsorption of tetracycline, oxytetracycline and chlortetracycline on soils with different pH value and organic matter content', *Environmental Research*, vol. 178, pp. 108669.doi: https://doi.org/10.1016/j.envres.2019.108669	No
7	Uddin, M, Chen, J, Qiao, X, Tian,R & Zhu, M, 2020, 'Insight into dynamics and bioavailability of antibiotics in paddy soils by in situ soil moisture sampler', *Science of the Total Environment,* vol. 703, pp. 135562.doi: https://doi.org/10.1016/j.scitotenv.2019.135562.	Yes
8	Hu, S, Zhang, Y, Shena,G, Zhanga, H, Yuand, Z & Zhang, W, 2019 'Adsorption/desorption behaviour and mechanisms of sulfadiazine and sulfamethoxazole in agricultural soil systems', *Soil and Tillage Research,* vol. 186, pp.233–241. doi: https://doi.org/10.1016/j.still.2018.10.026	Yes
9	Andriamalala, A, Vieublé-Gonod,L, Dumeny,V & Cambier, P, 2018, 'Fate of sulfamethoxazole, its main metabolite N-ac-sulfamethoxazole and ciprofloxacin in agricultural soils amended or not by organic waste products', *Chemosphere,* vol 191, pp. 607–615.doi: https://doi.org/10.1016/J.CHEMOSPHERE.2017.10.093	No
10	Wojslawski, J, Białk-Bielińska, A, Stepnowski, P & Dołzonek, J, 2019 'Leaching behaviour of pharmaceuticals and their metabolites in the soil environment', *Chemosphere,* 231, pp. 269–275.doi: https://doi.org/10.1016/j.chemosphere.2019.05.031	Yes
11	Chefetz, B, Mualem, T, & Ben-Ari, J, 2008, 'Sorption and mobility of pharmaceutical compounds in soil irrigated with reclaimed wastewater', *Chemosphere*, vol. 73, pp. 1335–1343.doi: https://doi.org/10.1016/j.chemosphere.2008.06.070.	No
12	Kiecaka, A, Sassine, L, Boy-Roura, M, Elsner, M, Mas-Pla, J, Le Gal La Salle, C, & Stumpp, C, 2019, 'Sorption properties and behaviour at laboratory scale of selected pharmaceuticals using batch experiments', *Journal of Contaminant Hydrology*, vol. 225, pp. 103500.doi: https://doi.org/10.1016/j.jconhyd.2019.103500	Yes
13	Dalkmann, P, Willaschek, E, Schiedung, H, Bornemann, L, Siebe, C & Siemens, J, 2014, 'Long-term Wastewater Irrigation Reduces Sulfamethoxazole Sorption, but Not Ciprofloxacin Binding, in Mexican Soils', *Journal of Environmental Quality*, vol. 43(3), pp. 964–970.doi: https://doi.org/10.2134/jeq2013.11.0473	Yes

grown carrot, sweet potato and wheat (Malachi et al., 2014), individually grown spinach, arugula and radish (Kodesova et al., 2019) or comparing between rice and no plants (Kodesova et al., 2019)). There are also variations in terms of different irrigation regimes (e.g., daily (Uddin et al., 2020), approximately every two days (Kodesova et al., 2019); for 35 min at a time (Koroša et al., 2020) or until column saturation (Yu et al., 2020)). Column experimental setups also contact the water with the soil via different methods, like drip irrigation (Paz et al., 2016), sprinkle irrigation (Koroša et al., 2020), or saturation of the soil column (Uddin et al., 2020; Yu et al., 2020). It is possible to enter qualitative variables as quantitative variables in a model but only as binary ('yes or no') variables (Agudelo-Jaramillo et al., 2016). It is not easy to adapt such a wide range of different qualitative variables to quantitative variables. Therefore, the wider availability of comparable batch sorption data following the OECD 106 Guideline (or similar batch-sorption setups) gives preferences to these batch studies over column studies for meta-analysis and modelling. The focus for data collection or this meta-analysis was on carbamazepine and sulfamethoxazole, due to their persistence and prevalence in the environment (Petrie et al., 2015) and given their status as proposed markers of raw wastewater contamination in natural waters and soil (Hai et al., 2018; Thiebault, 2020). Other pharmaceuticals were subsequently considered to encompass a wider range of physical-chemical properties.

Machine learning models, such as principal component analysis (PCA) and partial least squares (PLS)can quantitatively identify the independent variables influencing a dependent variable (Li et al., 2020; Syms, 2018). Both PCA and PLS have been used to identify variables that have a significant effect on the sorption behaviour of twenty-one pharmaceuticals to different types of soil (Li et al., 2020). This work is the first-time multivariate analysis (PCA and PLS) models are developed for a meta-dataset of batch-sorption studies of pharmaceuticals. Multivariate analysis is used to assess the effect of independent variables, such as different experimental conditions and soil characteristics, on the extent of pharmaceutical sorption to soil. Furthermore, the visualization of the PLS score plot is very useful in identifying underlying similarities and differences between studies (Li et al., 2020; Syms, 2018). The aim is to zoom out from several sorption studies in order to gain insights from meta-analysis, which can help to strike a balance between the advantage of comparability across studies from different research groups when performing laboratory-scale batch sorption experiments at standardised conditions (OECD, 2000) and the advantage of experiments performed at more environmentally relevant conditions. Therefore, the objectives of this work were to:

1. Identify the variables and ranges studied in the batch sorption literature from a meta-dataset to inform on the environmental relevance of experiments in the literature.
2. Model data from a meta-dataset to identify and prioritise specific variables which could be used to guide future laboratory studies and ensure environmental relevance.

2. Experimental

2.1. Meta-analysis and building the meta-dataset

Using Science Direct, Web of Science and Google Scholar search engines, batch-sorption studies in journal publications were identified. The following search method was used on Web of Science advanced search function: {'sorption' OR 'partitioning'} AND {'soil' OR 'solid phase' OR 'liquid phase'}, AND {'sulfamethoxazole' OR 'carbamazepine'} OR {'bioactive chemical pollutant' OR 'emerging pollutants' OR 'emerging contaminants' OR 'pharmaceutical' OR 'personal care product' OR 'antibiotics' OR 'pharmaceutical and personal care products'}. Sulfamethoxazole and carbamazepine were prioritised due to their environmental prevalence and persistence (Petrie et al., 2015; Hai et al., 2018; Thiebault, 2020), but data from other pharmaceuticals was also collected to account for the diverse chemical properties of this kind of pollutants (Table S1). Combinations of these keywords were also entered on the Science Direct search

engine as well as Google Scholar. The reference papers obtained from this literature search were screened to ensure that the necessary data for modelling was provided meeting the following criteria:

a) Batch sorption studies either based on the OECD Guideline 106 or using a similar experimental set-up.
b) An iterative screening process was undertaken in which the most commonly reported experimental variables (as well as soil properties) were noted and then those reports which did not report at least 90% of the experimental variables of interest were withdrawn from the pool of references to be revised.

References were screened to ensure that sorption was either i) reported as K_d for a specific concentration (Eq. (1)); ii) K_d had been derived from a linear isotherm (the isotherm was deemed to be linear when the corresponding Freundlich N value was between 0.8 and 1 (Tseng and Wu, 2008)) iii) if the isotherm was not linear, then the equilibrium concentrations of the pharmaceutical in the soil and liquid phase could be used to derive K_d at this concentration level using Eq. (1).

$$K_d = \frac{m}{(1-m)} \frac{v}{s} \tag{1}$$

where m is the fraction of pharmaceutical sorbed onto soil particles, v is the volume of liquid phase and s the amount of soil (USEPA, 1999).

When K_d was reported as K_{oc}, it was normalised using the soil organic matter content as specified in Eq. (2).

$$K_{oc} = \frac{K_d}{f_{oc}} \tag{2}$$

where K_{oc} is the linear approximation of sorption behaviour normalised by the fraction of organic carbon (f_{oc}) in the soil (Chefetz et al., 2008).

The K_d was used as the independent variable and starting point for the model, as opposed to Freundlich or Langmuir coefficients, because it is not a fitted curve but rather a number describing the partition at a given concentration. Furthermore, K_d was used since the sorption capacity of soil was not surpassed in the range of concentrations of sorbate reported (i.e., sorption was linear). Table S2 details how each K_d value was obtained or derived from each reference.

The references remaining after the selection process are given in Table 1, along with the numbering system with which they will be referred to throughout the text. The thirteen studies selected were analysed to identify the independent and dependent variables. Batch experiments consist of contacting a liquid phase, containing the target pollutant(s), with the soil and after an equilibration period measuring the amount of pollutant in the solid and/or liquid phase (mostly determined by difference in the liquid phase). The dependent variables identified from the studies analysed were the target pharmaceutical (Table 2), the soil properties (Table 3) and the experimental conditions (Table 2). To build the meta-dataset and models, the workflow described in Fig. 1 was undertaken. A table was filled, where each row was a data point, given by a K_d value obtained under a particular set of conditions (i.e., a particular pharmaceutical at a specific concentration in a particular soil under certain experimental conditions). Each of the K_d values used for the model can be seen in Table S2. The dependent variables were divided into three categories (Fig. 1, step b). The first category of dependent variables is the pharmaceutical studied, along with its physical-chemical properties (pK_a, water solubility and octanol-water portioning coefficient (LogK$_{ow}$)). These properties were all obtained from DrugBank (Wishart et al., 2017). The second category was the soil properties (Table 3). The third category was the experimental conditions (detailed in Table 2). From Tables 2 and 3, it was possible to determine which soil parameters were universally reported in the meta-dataset, as well as identifying which dependent variables were controlled and the range they were studied.

2.1.1. Conversion of data for comparability and uniformity of the meta-dataset

Following collation of all data points from the 13 reference studies, it was necessary to ensure the data-set was uniform and therefore comparable (same units and same dependent variables reported for each data point, Fig. 1c and d). The data points from references () had to be removed from the meta-dataset because it was not possible to

Table 2

Experimental conditions (dependent variables) found reported across the references used for the models developed and the ranges in which they were studied.

Ref.	No. data points (distinct K_d values)	Pharmaceutical [initial concentration ($\mu g\,L^{-1}$)]	Soil type, source, and preparation	Liquid phase	Vessel volume and material	Liquid (mL):soil (g)	pH	Temp. (°C)	Sterilisation method	Equilibration time (h)
1	3[a]	Acetaminophen Carbamazepine Sulfamethoxazole [100][1]	Loamy sand from aquifer zone in Spain collected at 10 cm. Air dried and sieved (2 mm).	Synthetic reclaimed water[S]	3 L glass	1200:300	7.98	20	Autoclaved at 121 °C for 15 min.	12
2	6[*]	Carbamazepine Naproxen Triclosan [0.1][1]	Silty clay and clay from wastewater irrigated agricultural zone in Mexico at 10 and 40 cm depth. Air dried and sieved (2 mm).	Aqueous calcium chloride (10 mM)	50 mL amber glass	15:3	Not stated	25	Frozen at −10 °C and air dried for 12 h.	48
4	3[b]	Carbamazepine [500][1]	Loamy sand, sandy loam, sandy clay from an agricultural region in Iran. Preparation not stated.	Aqueous calcium chloride (5 mM)	50 mL Teflon	1:1 (stated as ratio, not as a volume and mass)	7.6–7.8 (at the end)	25	Analytes dissolved in a solution of 100 mg L^{-1} sodium azide before being added to the soil.	168
5	5[c]	Carbamazepine [5000-25,000][1]	Sandy clay loam, sandy loam and clay loam bought from Landlook (Midlands, UK). Preparation not stated.	Aqueous calcium chloride (10 mM).	50 mL poly-propylene tubes.	10–1200:1–5 (depends on soil and pollutant, not stated for each test)	Not stated	4	Autoclaved at 105 °C for 3 h.	12
7	19[*]	Clarithromycin Florfenicol Sulfadiazine Sulfamethoxazole Trimethoprim [200][1]	Three silt loams from paddy soil in China collected at 0–20 cm depth. Air dried and sieved (2 mm).	Aqueous calcium chloride (10 mM).	Not stated	20:10	Not stated	Not stated	Autoclaved (conditions not stated).	24
8	4[c]	Sulfamethoxazole [500-10,000][1]	Three silt loams from agricultural fields in China collected at 0–20 cm depth. Preparation not stated.	Aqueous calcium chloride (10 mM).	30 mL glass vessel.	20:4	Not stated	20	Autoclaved (conditions not stated).	24
9	3[*]	Sulfamethoxazole (Thebo et al., 2017)[2]	Three silt loams, one amended with organic waste, one with compost and sewage sludge and one with farmyard manure. Collected at 0–25 cm depth. Preparation not stated.	Aqueous calcium chloride (10 mM).	Not stated.	Volume of liquid phase not specified and 3 g of soil.	Not stated	20	Not specified.	24
10	4[c]	Carbamazepine [250–10,000][1] Tramadol [625-80,000][1]	Loamy sand and clay bought from a German institute. Preparation not stated.	Aqueous calcium chloride (10 mM).	Poly-propylene tubes, volume not stated.	20–100:1	Not stated	20	100 mgL^{-1} sodium azide in the liquid phase.	24
11	23[d]	Carbamazepine Diclofenac Naproxen [0.5, 5.0][1]	Sand and sandy loam from a field irrigated with wastewater for 25 years from Israel. Collected from 0 to 5 and 15–25 cm depth. Preparation not stated.	Aqueous calcium chloride (5 mM) and secondary treated wastewater	Teflon, volume not specified.	Not stated.	Not stated	25	100 mg L^{-1} sodium azide in the liquid phase.	72
12	3[c]	Ofloxacine, Caffeine [500–7500]	Technical coarse sand. Sandy loam, sandy clay loam and clay from aquifer material in France and Spain. Preparation not stated.	Aquifer water	120 mL glass	50:10	7–7.5	20	50 mg L^{-1} sodium azide in the liquid phase.	96
13	8[*]	Ciprofloxacin [0.02][1] Sulfamethoxazole [0.02][1]	Vertisol from agricultural fields in Mexico under wastewater irrigation for 0, 14, 35 and 100 years. Collected from 0 to 30 cm depth. Preparation not stated.	Aqueous calcium chloride (10 mM).	Borosilicate vessel, volume not specified.	25:0.5 for ciprofloxacin 25:1 for sulfamethoxazole	Not stated	20	Not stated	48

[a] K_d was calculated based on reported concentrations at equilibrium of a batch-sorption experiment with initial aqueous concentration of 100 $\mu g\,L^{-1}$ using Eq. (1).

[b] K_d was calculated by extracting the concentration at equilibrium in the liquid and soil phase reported in linear sorption-desorption isotherms for the sedon-lowest initial concentration (More details in Table S2).

[c] Reported in the reference for an isotherm, only K_d values corresponding to a linear isotherm were used (full details in Table S2).

[d] Given in the reference as K_{oc} for a specific concentration and converted to K_d using Eq. (2) and the organic carbon content provided in the paper.

[1] Mass balance was carried out to calculate the concentration per mass of soil, as it was given per liquid volume.

[2] Unit conversion from $\mu g\,g^{-1}$ to $\mu g\,kg^{-1}$.

[*] No conversions done, data used as reported for a specific concentration. (see Table S2 for details on whether data is reported in the main text or in the supplementary information of the paper).

[S] Synthetic reclaimed water was composed of: NH_4Cl (0.07 g L^{-1}), $MgSO_4$ (0.1 g L^{-1}), $CaCl_2$ (0.01 g L^{-1}), K_2HPO_4 (0.02 g L^{-1}), $NaHCO_3$ (0.25 g L^{-1}), peptone (0.01 g $^{-1}$), and meat extract (0.01 g L^{-1}) (Martínez-Hernández et al., 2016).

Table 3

Soil properties reported in the batch-sorption references from which the meta-dataset was built.

Reference number to literature study	1	2	4	5	7	8	9	10	11	12	13
Texture (% sand, silt, and clay)	X	X	X	X	X	X	X	X	X	X	X
Predominant clay type		X					X				
pH measured in a calcium chloride solution (pH$_{cacl}$)				X			X		X		
pH measured in water (pH$_{H2O}$)			X		X	X		X			
pH measured in a potassium chloride solution (pH$_{KCL}$)						X					
pH measurement method not specified	X	X					X				X
Fraction of organic content			X	X	X	X				X	X
Total organic carbon			X	X	X		X	X	X	X	X
Total dissolved organic carbon					X						
Total organic matter		X			X						
Total carbon							X			X	
Total nitrogen					X		X				X
Moisture						X					
Point of zero charge		X									
Electrical conductivity	X	X	X		X				X		
Cation exchange capacity	X	X		X		X	X	X		X	X
Exchangeable sodium (Na$^+$)			X	X					X		
Exchangeable magnesium (Mg^{+2})			X	X					X		
Exchangeable calcium (Ca^{+2})			X	X					X		
Exchangeable potassium (K$^+$)			X	X					X		
Exchangeable aluminium						X					
Exchangeable iron						X					
Exchangeable manganese						X					
Exchangeable acidity						X					
Aluminium oxide content		X									
Iron oxide content		X									
Total potassium content					X						
Total phosphorus content											
Total calcium content					X	X					
Total sodium content					X	X					
Total aluminium content					X	X					
Total iron content					X	X					
Total magnesium content					X	X					
Total manganese content					X	X					
Calcium carbonate content				X	X				X		X
Soil bulk density				X							
Specific surface area				X					X	X	
Endogenous sorbate concentration				X							
Sampling location	X	X	X	X	X	X	X	X	X	X	X
Sampling depth	X	X	–		X	X	X	–		X	X
Air dried	X	X		X		X	X	X		X	X
Mesh size of soil sieving (mm)	X	X		X		X	X	X	X	X	X

extract the linear K$_d$ accurately or because they reported values for systems of mixtures of pharmaceuticals (versus individual pharmaceuticals) and these were not comparable to the other data points. However, references () still formed part of the meta-analysis carried out. Dependent variables which were not reported across all references were also removed. These were all soil properties which are not in bold in Table 3, as well as five experimental set-up variables: pH during experiment, vessel material, vessel volume, liquid volume, and soil mass (Table 2). A total of 82 data points from 11 references remained as the meta-dataset for the model.

Cation exchange capacity (CEC) of soil was not always reported in the same units and so all CEC data was converted to cmol kg^{-1}. All except one study reported organic carbon content of the soil and one study reported organic matter (Uddin et al., 2020) (Table 3), a factor of 1.72 was used to convert total organic carbon to organic matter (Fenton et al., 2008). The soil pH, pollutant's pK_a and LogK_{ow} were entered as LogD (Schaffer et al., 2012), based on Eqs. (3) to (5), to account for multiple pK_a points of some pharmaceuticals:

$$LogD = \log_{10}\left(\frac{\alpha_0}{\alpha_1}\right) \sum_{i=1}^{n} i = \frac{1}{1+10^{pH-pka^n}} \tag{3}$$

$$\alpha_0 = \sum_{pka=1}^{n} \frac{1}{1+10^{pH-pka^n}} \tag{4}$$

$$\alpha_1 = 1 - \alpha_0 \tag{5}$$

where α_1 is the fraction of pharmaceutical ionised in the soil, α_0 is the fraction of neutral pharmaceutical in the soil, pK_a^n refers to the pharmaceutical pK_a values (Schaffer et al., 2012). The pH values of the soil in the literature are often obtained in different ways (Minasny et al., 2011), with the 13 references either obtaining pH by preparing a paste with soil and pure water (pH$_{H2O}$) or using 0.01 M CaCl$_2$ (pH$_{CaCl2}$). Measurements in CaCl$_2$ solution have been shown to be less affected by soil electrolyte concentration and therefore more stable (Minasny et al., 2011). The pH values were converted from pH$_{H2O}$ to pH$_{CaCl2}$ or pH$_{CaCl2}$ to pH$_{H2O}$ using a conversion table from the literature, which was built from comparing 7844 pH measurements obtained in pure water to the same soil measured using 0.01 M CaCl$_2$ (Henderson and Bui, 2002). Most of the data used to build this conversion table was in the pH range of 3.5 to 9.5 and the authors state that their model is more reliable within this range (Henderson and Bui, 2002). This pH range is adequate for this meta-dataset, as 99.5% of the pH values were between 6.8 and 7.8. More accurate conversion models of pH$_{H2O}$ to pH$_{CaCl2}$ could not be used as not all references reported the soil electrical conductivity (Minasny et al., 2011), as can be seen in Table 3. When the solution used to measure the pH was not specified (Martínez-Hernández et al., 2016; Andriamalala et al., 2018a; Uddin et al., 2020; Dalkmann et al., 2014a), it was assumed that it was measured using pure water. To assess the accuracy of this assumption, three models were developed, one using pH values as reported (no conversion), the second using all pH$_{H2O}$ values converted to pH$_{CaCl2}$ and a third model with all pH$_{CaCl2}$ values converted to pH$_{H2O}$.

The qualitative variables were turned into quantitative variables using the quantification matrix method (Agudelo-Jaramillo et al., 2016). The type of liquid phase was turned from a qualitative variable to a quantitative variable by having four variables called a) liquid phase: aqueous calcium chloride, b) liquid phase: aqueous calcium chloride plus other salts, c) liquid phase: secondary treated water, and d) liquid phase: well/aquifer water (Table 4). For each of these variables, 1 or 0 was entered depending on whether the liquid phase was used, following guidance on how to model qualitative data (Agudelo-Jaramillo et al., 2016). This approach was also applied for the sterilisation method, using two variables: Sterilisation by autoclaving or Sterilisation by sodium azide addition, and when no sterilisation method was stated in the study a zero was entered for both variables. All data was normalised using unit-variance to avoid over-sensibility towards anyone variable.

2.2. Multivariate analysis of the data collected from the thirteen references

Multivariate analysis was undertaken using the SIMCA Multivariate Data Analysis Software tool (Umetrics-Sartorius, Sweden, version 16.0.1). The independent variable (y) was K_d, the dependent variables (x) were: soil pH$_{CaCl2}$, soil sand content (%), soil silt content (%), soil clay content (%), soil organic carbon, liquid phase: aqueous calcium chloride, liquid phase: aqueous calcium chloride plus other salts, liquid phase: secondary treated water, liquid phase: well/aquifer water, sterilisation by autoclaving, sterilisation with sodium azide, equilibration time (just soil and liquid), time to equilibrium once pollutant was added, initial pharmaceutical concentration, pharmaceutical water solubility, pharmaceutical dispersion coefficient (LogD), and pharmaceutical octanol-water partitioning coefficient (LogK_{ow}) (Table 4).

The pharmaceutical name, reference number and soil type were entered as secondary identifiers in SIMCA. Secondary identifiers can be used to label individual data point in the visualization plots provided by SIMCA; e.g., all data points belonging to reference number 12 would have a '12' label next to them. This allows observing if the data points were grouped according to pollutant, soil type or reference. If the data points were grouped based on reference number, it would indicate a dependence on K_d based on experimental conditions, which was not accounted for in the model. Using this data three different models were built (Table 4), where
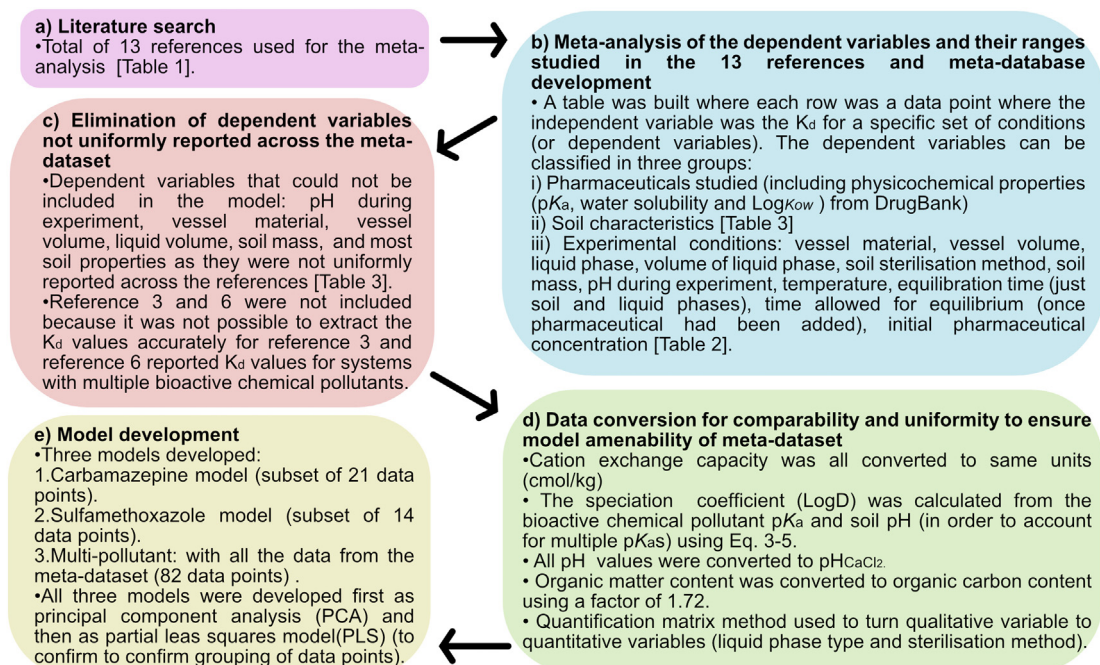
a) Literature search
•Total of 13 references used for the meta-analysis [Table 1].

b) Meta-analysis of the dependent variables and their ranges studied in the 13 references and meta-database development
• A table was built where each row was a data point where the independent variable was the K_d for a specific set of conditions (or dependent variables). The dependent variables can be classified in three groups:
i) Pharmaceuticals studied (including physicochemical properties (pK_a, water solubility and Log_{Kow}) from DrugBank)
ii) Soil characteristics [Table 3]
iii) Experimental conditions: vessel material, vessel volume, liquid phase, volume of liquid phase, soil sterilisation method, soil mass, pH during experiment, temperature, equilibration time (just soil and liquid phases), time allowed for equilibrium (once pharmaceutical had been added), initial pharmaceutical concentration [Table 2].

c) Elimination of dependent variables not uniformly reported across the meta-dataset
•Dependent variables that could not be included in the model: pH during experiment, vessel material, vessel volume, liquid volume, soil mass, and most soil properties as they were not uniformly reported across the references [Table 3].
•Reference 3 and 6 were not included because it was not possible to extract the K_d values accurately for reference 3 and reference 6 reported K_d values for systems with multiple bioactive chemical pollutants.

d) Data conversion for comparability and uniformity to ensure model amenability of meta-dataset
•Cation exchange capacity was all converted to same units (cmol/kg)
• The speciation coefficient (LogD) was calculated from the bioactive chemical pollutant pK_a and soil pH (in order to account for multiple pK_as) using Eq. 3-5.
• All pH values were converted to $pH_{CaCl2.}$
• Organic matter content was converted to organic carbon content using a factor of 1.72.
• Quantification matrix method used to turn qualitative variable to quantitative variables (liquid phase type and sterilisation method).

e) Model development
•Three models developed:
1.Carbamazepine model (subset of 21 data points).
2.Sulfamethoxazole model (subset of 14 data points).
3.Multi-pollutant: with all the data from the meta-dataset (82 data points) .
•All three models were developed first as principal component analysis (PCA) and then as partial leas squares model(PLS) (to confirm to confirm grouping of data points).

Fig. 1. Process undertaken to develop the meta-dataset and models from the 13 references studied.

Model 1 and Model 2 were for carbamazepine (21 data points) and sulfamethoxazole (14 data points) data points, respectively. Model 3 included other pollutants and comprised all 82 data points, with the proportion of pollutants presented in Fig. 2.

Principal component analysis (PCA) was applied to the data for all three models, before applying partial least squares (PLS) regression, as the latter is an unsupervised model. PLS models have been shown to generate class separation, even with random data (Worley and Powers, 2013). Therefore, the PCA model and the PLS were visually compared to ensure similar grouping or class separation, indicating the PLS was not over-fitting the model using the data. The PCA score plot of the models can be seen in Figs. S1 and S3. PCA clusters Y-variables (K_ds in this case) on how similar they are depending on how the X-variables (the experimental conditions in this case) varied within them. Partial least square (PLS) model was then used on the same data to identify the variables which have a larger weigh on the PCA grouping (Umetrics-Sartorious, 2019). PLS is a regression extension analysis of PCA, where PLS can differentiate between which X variables predict Y and which do not, and identify the X-variables important for prediction of the Y-variable (K_d in this case) (Umetrics-Sartorious, 2019). Variables which have a significant importance to the prediction of the dependent variable have a value >1 (Umetrics-Sartorious, 2019). For PLS, there are two important statistics which are indicators of the quality of the model, RX^2 is a measure of how well the model fits the data and Q^2 cumulative is a measure of how well Y can be predicted using X (Umetrics-Sartorious, 2019). Only models with RX^2 and $Q^2 > 0.5$ were considered as valid (Umetrics-Sartorious, 2019). Model validity was also assessed based on the permutation test ($n = 100$), where the random RX^2 and Q^2 values must be lower than the real model R^2 and Q^2 values. Permutations for all models can be found in Figs. S2 and S4.

Table 4

Models developed using the meta-dataset, dependent variables included, number of data points and references use.

Model no.		1	2	3
Pharmaceutical		Carbamazepine	Sulfamethoxazole	Multi-pollutant
No. of data points		21	14	82
References used[a]		1,2,4,5,10–12	1,7–9, 13	1,2,4,5,7–13
Dependent variables included in model	Pharmaceutical octanol-water partition coefficient (LogK$_{ow}$)			X
	Pharmaceutical water solubility *(mg L^{-1})*			X
	Pharmaceutical speciation depending on soil pH *(LogD)*	X	X	X
	Soil sand content *(%)*	X	X	X
	Soil silt content *(%)*	X	X	X
	Soil clay content *(%)*	X	X	X
	Soil fraction of organic carbon *(%)*	X	X	X
	Liquid phase: Aqueous calcium chloride *(Yes or no, entered as 1 or 0)*	X	X	X
	Liquid phase: Aqueous calcium chloride plus other salts *(Yes or no, entered as 1 or 0)*	X	X	X
	Liquid phase: Secondary treated water *(Yes or no, entered as 1 or 0)*	X	X	X
	Liquid phase: Well/aquifer water *(Yes or no, entered as 1 or 0)*	X	X	X
	Sterilisation by autoclaving *(Yes or no, entered as 1 or 0)*	X	X	X
	Sterilisation with sodium azide *(Yes or no, entered as 1 or 0)*	X	X	X
	Equilibration time, just soil and liquid *(h)*	X	X	X
	Time to equilibrium once pollutant was added *(h)*	X	X	X
	Initial pharmaceutical concentration *(μg kg^{-1})*	X	X	X

[a] Numbering system for the 13 references used for the meta-analysis is provided in Table 1.
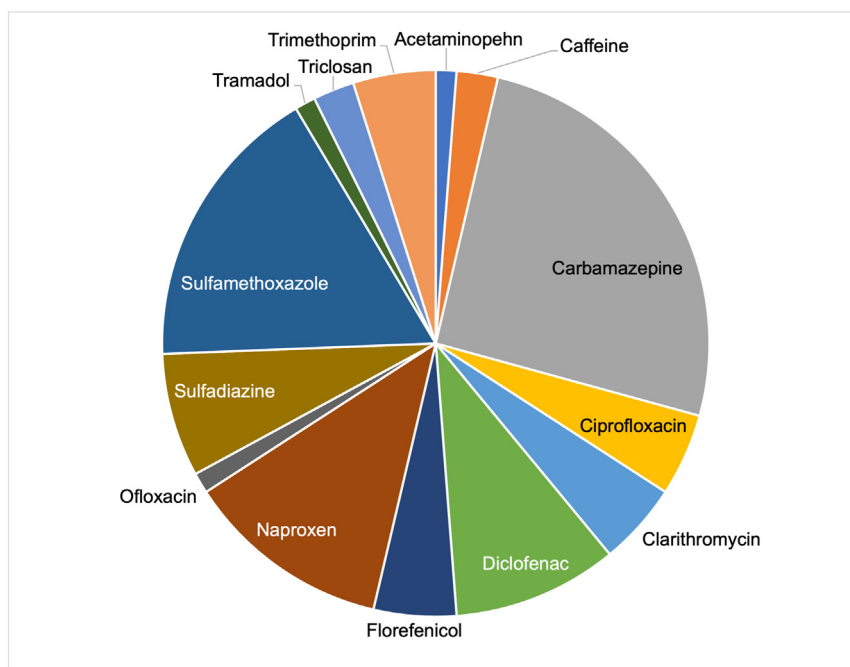
Fig. 2. Proportion of pollutants from the literature search that were used for the multi-pollutant model.

3. Results and discussion

3.1. Parameters and ranges of experimental conditions (dependent variables) used for batch sorption experiments in the meta-dataset

The studied parameters that can influence K_d and were reported by all or most papers included in the meta-dataset were divided into pollutant studied, soil properties, and experimental conditions (Tables 2 and 4). The K_d is a snapshot of the partitioning of a sorbate in a soil/water system at a particular set of conditions and, as such, it is amenable for modelling where each data point is a K_d value at particular experimental conditions (dependent variables).

3.1.1. Soil properties reported in the meta-dataset

Different references report different soil characteristics (Table 3). Only the soil characteristics reported across all the references could be included in the models. These were soil pH, organic carbon content and soil texture (percentage of sand, clay, and silt) (Table 3). According to the OECD Guidelines, these are the most important soil characteristics determining the sorption (OECD, 2000). However, other soil characteristics that affect sorption are cation exchange capacity (CEC), content of amorphous iron and aluminium oxides, and specific surface area of the soil particles (OECD, 2000). It would be beneficial for studies to consistently report other soil properties to allow for more comprehensive meta-analysis in the future.

Soil texture was reported uniformly across all references (Table 3), displaying a wide range of characteristics (e.g., sand <10 to <97%, silt <5 to 75% and clay <5 to 53%, see Fig. 3a). Seven out of the thirteen references used soils taken from agricultural areas (Table 2). Two references used soil from the Tula Valley, an agricultural area in Mexico where treated and untreated wastewater is used for irrigation (Dalkmann et al., 2014a; Durán-Álvarez et al., 2012). In one study, soil that had been irrigated with untreated water for approximately 85 years was used (Durán-Álvarez et al., 2012), whilst another study had soils irrigated with untreated wastewater for 0, 14, 35 or 100 years (Dalkmann et al., 2014a). Two references used
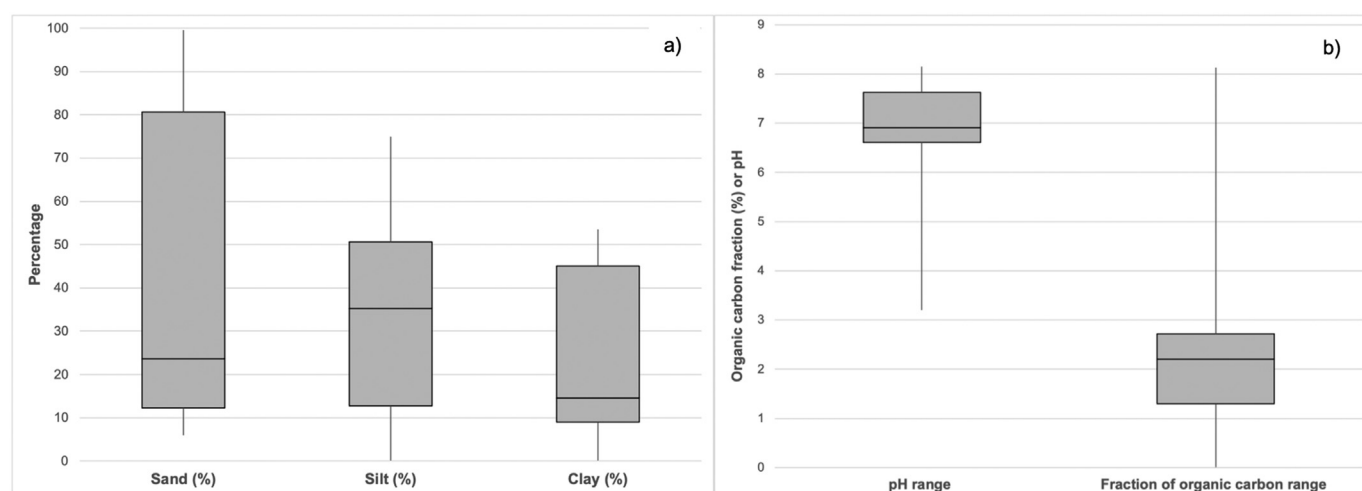
Fig. 3. Range of properties in the soils studied in the literature consulted (Table 1) a) soil texture (percentage content of each component); b) pH and fraction of organic carbon.

agricultural soil form Iran (Paz et al., 2016; Chefetz et al., 2008), in one of them the soil was from a field irrigated with treated wastewater for over 25 years (Chefetz et al., 2008). Two references used Chinese agricultural soils (Uddin et al., 2020; Hu et al., 2019). One reference used soil from an experimental field in France to compare soil never amended with organic waste, soil amended with compost and sewage sludge, and soil amended with manure (Andriamalala et al., 2018a). Soils were also taken from an aquifer zone in Spain for two studies (Martínez-Hernández et al., 2016; Kiecak et al., 2019). Reference soils (Wojsławski et al., 2019) and technical sand with low clay and organic matter contents were used as a comparison to aquifer soils in one study (Kiecak et al., 2019). The soil was mostly collected from the top layer (0 to 30 cm) (Table 2), with several references comparing sorption using different soil depths: 10 cm was compared to 40 cm in (Durán-Álvarez et al., 2012). Some aquifer soils were collected from deeper layers (80–300 cm) due to a shallow aquifer (Kiecak et al., 2019). The depth of soil collection was not always specified (Paz et al., 2016). This wide variety of soils (wide texture range, different soil depths, different geographic locations, and different amendment using either organic solids or wastewater) provides richness to the meta-dataset. Where conclusions from the model are only relevant to the dataset used to develop it, and this wide range of soils allows for comparison of different soils (in terms of texture, geography and irrigation or amendment history).

The range of soil pH_{CaCl2} studied was mostly from 6.8 to 7.8 (Fig. 3b). This is on the higher end of the optimum range for most crops, which is between 5.5 and 7.5 (Oshunsanya, 2018). Seven out of 13 references used soil which had been under wastewater irrigation, and wastewater irrigation is known to affect soil pH (Abd-Elwahed, 2018). In some instances soil pH has been reported to increase with wastewater irrigation (Abd-Elwahed, 2018; Qian and Mecham, 2005), and in other cases a decrease has been reported (Abd-Elwahed, 2018; Angin et al., 2005). This means the conclusions drawn from the model will not be applicable to the whole range of pH presented by agricultural soils. However, data availability for soils that have and have not been under wastewater irrigation provides environmental relevance. The pH during the experiment was not monitored in most references (Table 2), although in one paper it was controlled (Martínez-Hernández et al., 2016). Therefore pH could not be accounted for in the models as most references did not evaluate pH variations during the experiment. Any changes of pH during the experiment could affect sorption, therefore this variable should be monitored or kept constant.

3.1.2. Experimental conditions (dependent variables) reported in the literature

Important variables in batch experiments are the mass of soil and the volume of liquid used, as well as the initial concentration of the pollutant studied (OECD, 2000). These were not always specified or were reported as a range (Table 2), where correlation of results to initial conditions is not possible (Li et al., 2020). This means that these variables could not be included in the model. Liquid volume and soil mass ranged from 30 to 3000 mL and 0.5 to 300 g, respectively (Table 2). A recommendation would be for future studies to state these variables (soil mass, liquid volume, and initial concentration of the sorbate) corresponding to each K_d value reported. This would allow for more comprehensive meta-analysis and provide an understating of the effect of scale (soil mass and volume liquid) on mass transfer of the pharmaceuticals from the liquid phase to the soil.

Pharmaceutical concentrations were reported by some papers in terms of pharmaceutical mass per soil mass and sometimes in terms of pharmaceutical mass per liquid volume (Table 2). In the model, all pharmaceutical concentrations were converted to concentrations in a soil mass basis, which was achieved through a mass balance using the soil mass and liquid volumes reported. Carbamazepine initial concentrations in the soil phase ranged from 0.05 to 3,000,000 μg kg^{-1} (Table 2), while for sulfamethoxazole it ranged from 20 to 22,083.3 μg kg^{-1} (Table 2). Environmental concentrations in agricultural soils for sulfamethoxazole have been reported as: 0.2 to 54.5 μg kg^{-1} (Christou et al., 2019; Cycoń et al., 2019), while for carbamazepine concentrations from 1.2 to 98.3 μg kg^{-1} are presented (Christou et al., 2019). These concentrations in soil equate to concentrations in the liquid

phase of 0.1 to 1,000,000 μg L^{-1} for carbamazepine and 1000 to 1,000,000 μg L^{-1} for sulfamethoxazole. The lowest sulfamethoxazole concentrations are still 3 orders above that reported in wastewater influent (Petrie et al., 2015; Thiebault, 2020). For carbamazepine, the lowest concentration reported is environmentally relevant (Petrie et al., 2015; Hai et al., 2018), however most studies are above this range. The model conclusions are only valid within the concentration ranges in the meta-dataset and they are mostly higher than those measured in the environment. This discrepancy between environmental concentrations and those studied in batch-sorption studies is a significant gap in the literature because the percentage of pollutant sorption to soil has been shown to be concentration dependent (Kiecak et al., 2019; Singh, 2016).

The type of liquid phase used was specified in 9 out of 13 studies, using an aqueous calcium chloride (CaCl$_2$) solution (Table 2), with concentrations of either 5 or 10 mM (Table 2). Other liquid phases used were synthetic reclaimed water (Martínez-Hernández et al., 2016), groundwater (Kiecak et al., 2019) and secondary treated wastewater (Chefetz et al., 2008). It has been shown that K_d is dependent on the liquid phase utilised for batch-sorption experiments (Chefetz et al., 2008). This is a gap in the literature, as most studies perform partitioning experiments in aqueous calcium chloride, yet agricultural soil is irrigated with either treated/untreated wastewater or surface water (Table 2), meaning that the meta-dataset is not fully representative of the real environmental conditions.

The method used to sterilise the soil (ruling out biodegradation) was not specified for 2 out of the 13 references (Table 2). The methods reported in the meta-dataset were autoclaving and addition of sodium azide (NaN$_3$) (Table 2). The soil sterilisation method can have a significant effect on sorption by changing soil properties. Autoclaving has been shown to increase the content coarse-size particles attributed to aggregation (Lotrario et al., 1995). It can also change the quality of soil organic matter as the content of the bacterial cells can be released at high temperature and pressure (Berns et al., 2008). Whilst sodium azide has been reported to increase soil pH (Skipper and Westermann, 1973). When autoclaving was used, the conditions were not always specified, and differing temperatures and time lengths were used among the references (Table 2), which can impact on the extent to which the soil is affected. The soil sterilisation method was included as two dependent variables (sterilisation by autoclaving and sterilisation by sodium azide) in the models as a quantitative binary variable (yes, 1 or no, 0 adapted from a qualitative variable). For the two references in which the sterilisation method was not specified, 0 was entered for both options.

The temperature at which the experiment was conducted has an effect on sorption (OECD, 2000; Li et al., 2018b) and this was not specified in all the references of the meta-dataset (Table 2) (Uddin et al., 2020). When it was specified, most experiments were conducted at 20 (six out of eleven) or 25 °C (three out of eleven) (Table 2), except for one performed at 4 °C using carbamazepine (Li et al., 2020). The OECD 106 Guideline specify that experiments should be conducted between 20 and 25 °C (OECD, 2000) thus providing comparable results across independent studies, but this does not account for the full range of environmentally relevant temperatures. While the overall temperature range studied (4 to 25 °C) is environmentally relevant, the distribution of temperatures within 4 and 20 °C is not adequate in terms of granularity. It is worth noting both from the perspective of environmentally relevant conditions in laboratory experiments and from the relevance of the conclusions drawn from modelling this data that this does not account for temperature variations, which will occur seasonally and geographically in agro-ecosystems.

The linear soil partitioning coefficient (K_d) is measured at equilibrium, i.e., when the sorbate concentration in both the liquid and solid phase is steady (OECD, 2000). This time period is determined prior to the kinetics and isotherm experiments and depends on the pollutant and soil studied and on the initial pharmaceutical concentration, liquid volume and soil amount (OECD, 2000). In the references revised, equilibrium times ranged from 12 h to 7 days (Table 2). Only two references included an equilibration period between the liquid and solid phase before adding the pollutant, to ensure the soil is fully wetted, where one was for 12 h (Kiecak et al.,

2019) and the other 24 h (Dalkmann et al., 2014a). These variables were included in the model and so their effect can be assessed.

Methods to prevent photodegradation were mostly stated and included amber glass (Li et al., 2020; Durán-Álvarez et al., 2012) and wrapping the containers with aluminium foil (Martínez-Hernández et al., 2016; Uddin et al., 2020; Hu et al., 2019; Kiecak et al., 2019). However, it was not always specified (Wojsławski et al., 2019; Chefetz et al., 2008; Dalkmann et al., 2014b) and subsequently unable to include in the model. Provided photodegradation has been controlled, the methods to achieve this are not expected to have an impact on the K_d. Another parameter which was not always specified is the type of test vessel used (Table 2). The OECD 106 Guideline state that checks should be carried out to rule out adsorption of the pollutant onto the vessel (OECD, 2000), and previous studies have shown that vessel material has an effect on pharmaceutical sorption to soil (Li et al., 2020). Most common materials used are glass and polypropylene tubes (Table 2). Nevertheless, if preliminary experiments have been carried out to check for vessel sorption, this variable should not impact K_d.

3.2. Model 1: subset of meta-dataset for carbamazepine

Model 1 was developed for carbamazepine only (Fig. 4). The variable with the most effect on grouping, of those in the carbamazepine PLS model, was the organic matter content (VIP = 2.0 ± 0.5). This is consistent with literature, which has shown increased carbamazepine sorption with increasing organic matter and previously found strong positive correlations between both parameters (Paz et al., 2016; Evyatar et al., 2018; Gibson et al., 2010). Soil clay content had a higher VIP (1.0 ± 0.4) than silt content (0.9 ± 0.5) and sand (0.8 ± 0.5). This has not been previously observed, though authors have speculated that clay content in the soil has a significant effect on carbamazepine sorption (Gibson et al., 2010).

Using aqueous calcium chloride as the liquid phase appears to have a more defined impact on K_d (1.25 ± 0.25) compared to using well/aquifer water or treated wastewater, where the error of the VIP was higher than the VIP score itself. There are more datapoints using aqueous calcium chloride as the liquid phase (12 out of 21), while for aqueous $CaCl_2$ with other salts there were only 5 datapoints and 4 datapoints for secondary treated water. It is possible that more datapoints are required with these last two liquid phases to discern how important liquid phase is for prediction of K_d for carbamazepine. The high variance (coefficients of variation >100%) in variable importance for prediction value for well/aquifer and treated water as liquid phase may be due to unknowns, such as competitive

sorption, which may be taking place due to other pollutants and dissolved organic matter present in the liquid phase. Untargeted analysis, when dealing with realistic environmental liquid phases (treated or untreated wastewater), could provide more environmentally relevant results as it would allow the elucidation of competitive sorption.

Sterilisation using sodium azide was a variable which was important for prediction (VIP = 0.95 ± 0.20) and sterilisation by autoclaving had a value of 0.6 ± 0.4, indicating that there may be some effect of the sterilisation method on the sorption behaviour of carbamazepine. The equilibration time between the liquid and soil phase (prior to adding the pollutant) had a VIP of 0.75 ± 0.30, indicating this step could have some effect on the carbamazepine sorption.

3.3. Model 2: subset of meta-dataset for sulfamethoxazole

The sulfamethoxazole model was not valid as it failed on the permutation test, while the RX^2 and Q^2 values were both below 0.5. In the PCA and PLS score plot, reference (Marti et al., 2014) was outside the model area, so this reference was removed but the model was still not valid. This could indicate that data points for sulfamethoxazole were not sufficient for a valid model or the granularity of the data did not allow for valid model building. The K_d values were between 0.12 and 25.78 L kg^{-1} with a mean of 3.46 L kg^{-1} and a standard deviation of 4.45 L kg^{-1} (Uddin et al., 2020; Dalkmann et al., 2014a; Hu et al., 2019; Kiecak et al., 2019; Andriamalala et al., 2018b), so the limited spread could have affected data granularity. Correlations between K_d values and percentage of clay, silt, and sand, pH$_{CaCl2}$, organic carbon content and initial concentration were explored but nothing significant was found. Sulfamethoxazole has been reported to have variable behaviour in soil (Martínez-Hernández et al., 2016; Andriamalala et al., 2018a) and inconsistencies between lab-scale and field observations have been reported (Kiecak et al., 2019). All the references studied for this pollutant used a liquid phase of aqueous $CaCl_2$ or aqueous $CaCl_2$ plus other salts. The failing of this model indicates more batch-sorption studies are needed to understand the sorption behaviour of sulfamethoxazole, with preference at more environmentally relevant conditions.

3.4. Model 3: multi-pollutant (full meta-dataset) model

The multipollutant model was subsequently undertaken to explore the effect of different pharmaceuticals with varied physical-chemical properties on
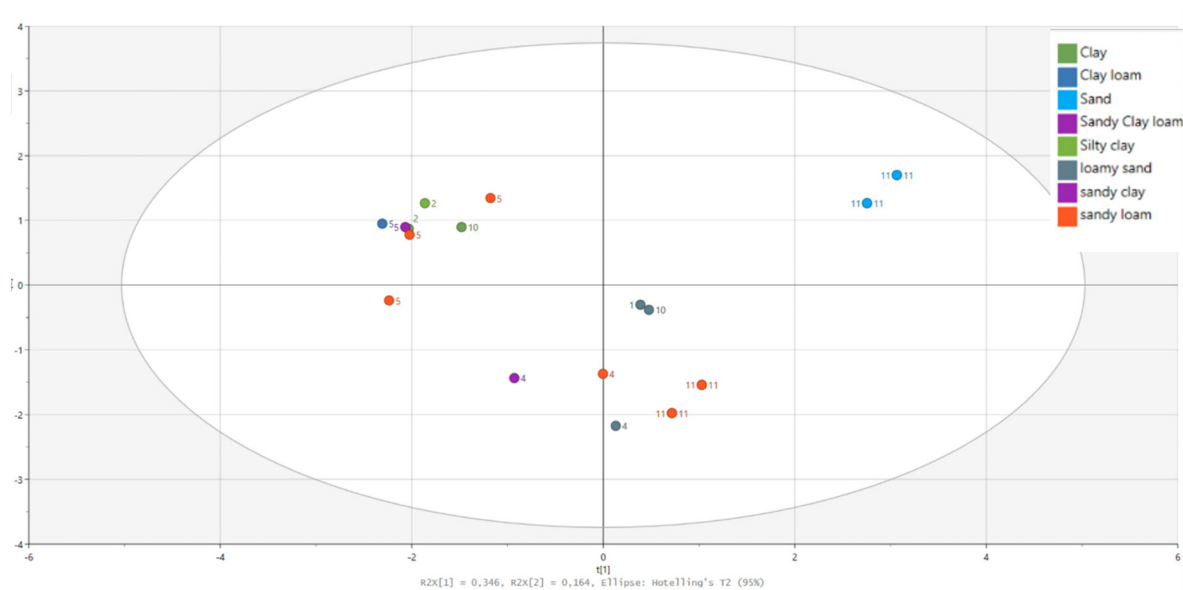

Fig. 4. Carbamazepine partial least squares scores plot: visualization of principal component 1 (x-axis) and principal component 2 (y-axis), datapoint marks based on soil type and labels based on references.

the outcomes of the PCA analysis. The proportion of pollutants studied in the meta-dataset is summarised in Fig. 2 and their main physical-chemical properties are given in Table S1. Most data points for each pollutant were obtained from only one reference, with the exceptions of carbamazepine (seven references), sulfamethoxazole (six references) and two references each for ciprofloxacin, diclofenac, and naproxen (Table 2). It is important to be aware of this proportion, as the conclusions drawn from Model 3 may be more relevant to the pollutants which had a larger proportion of datapoints, rather than to all of them.

Model 3 was evaluated using soil pH in three forms, first as stated in the reference (regardless of whether it was measured in water or in aqueous calcium chloride), second all pH_{CaCl2} converted to pH_{H2O}, and third pH_{H2O} converted to pH_{CaCl2}. The results were very similar with non-discernible data grouping and RX^2 values of 0.640, 0.590 and 0.651, respectively, indicating the buffer in which the pH was measured and the conversion between them does not have a significant impact on the results. Q^2 values were of 0.703, 0.698 and 0.712 for the three assessed conditions (Table S3). Given the higher RX^2 and Q^2 values obtained for the pH_{CaCl2}, this model was used for analysis. According to the PLS score plot (Fig. 5), there was no clear grouping regarding to pharmaceutical or to reference, there is a need for more data from a wider range of pharmaceuticals.

3.4.1. Variables of importance for prediction of the grouping in Model 3

From the variable importance of prediction scores (Fig. 6) the most important is the pharmaceutical water solubility, which highlights the importance of pharmaceutical-specific data and the need for more data amenable for modelling from a wide range of pharmaceuticals. After that, the variables for importance of prediction, with values above 1 making them significant for predicting K_d (Umetrics-Sartorious, 2019) were equilibration time between solid and liquid phase prior to pollutant addition > soil fraction of organic carbon > sterilisation by autoclaving. These results showed that performing an equilibration time between the solid and liquid phase prior to adding the pollutant influences K_d and should therefore be considered during batch sorption experimental design. This means that it makes a difference if the soil pores are fully wetted before the pollutant is added, as this can change the mass transfer of the pollutant from the liquid to soil. This finding is important as it is an unexplored variable to date in the literature, and may account for some of the discrepancies observed between lab-scale and field conditions. This is because in field studies, the agricultural soils will likely l be wetted (from precipitation or prior irrigation events) prior to being introduced to the pollutant. Soils are typically dried after collection and prior to batch-sorption studies, and it would be interesting to further explore the effect of an equilibration period between the liquid and soil prior to introducing the pharmaceuticals.

The sterilisation method was also a variable of importance, and requires further attention as may affect conclusions drawn from experiments using
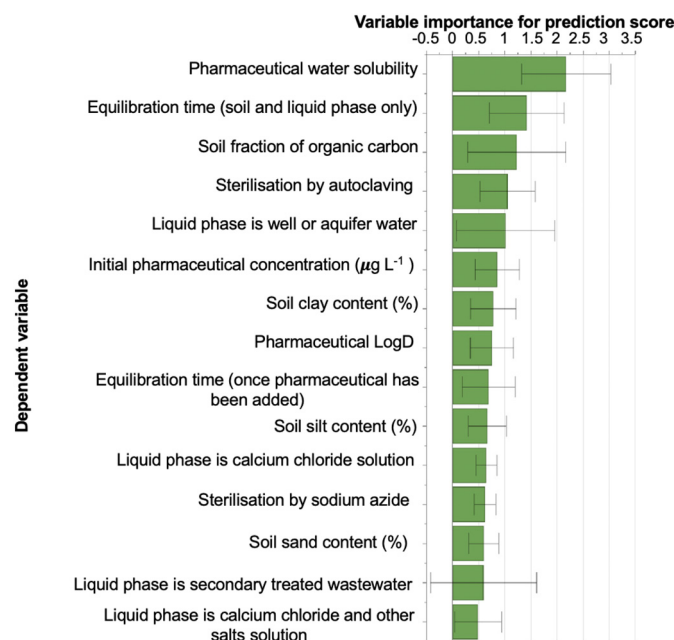
Fig. 6. Variables of importance >1 for the multi-pollutant partial least squares mode.

different types of sterilisation. Furthermore, it may also contribute to discrepancies observed between lab-scale and field conditions, given that agricultural soil will not have undergone any sterilisation processes. The VIP of the type of liquid phase used had high error associated so it is not possible to ascertain their importance for this model. However this should be further explored to ensure environmental relevance of batch sorption studies. This is of high relevance for discrepancies between lab-scale and field observations as in the environment the aqueous phase will be different in ionic strength to the $CaCl_2$ buffer as well as having other components, including dissolved organic matter. The initial concentration of the pharmaceutical phase had presented a VIP just below 1 (Fig. 5). However, considering the error associated, this variable may be above 1, which further highlights the importance of batch studies performed at environmentally relevant concentrations.

It is important to highlight that obtaining mechanistic insights into the sorption of the pollutants in the multi-pharmaceutical model is outside of the scope of this paper, but rather zooming out of specific studies in order to observe experimental trends that can inform on how best to increase the environmental relevance of lab-scale batch sorption studies. A limitation of the model is that it was not possible to include all the variables known to affect sorption and important sorption mechanisms, such as cation exchange,
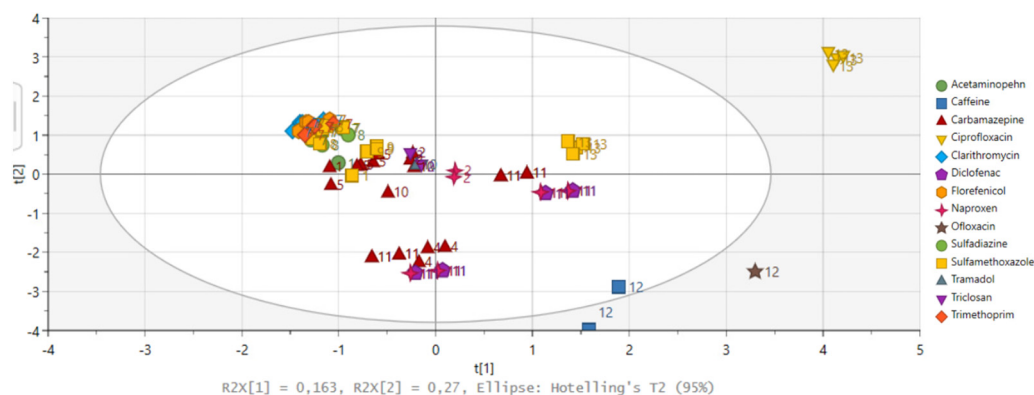
Fig. 5. Multi-pollutant PLS scores plot: visualization of principal component 1, t (UN-WWAP, 2015), and principal component 2, t (Holt-Giménez et al., 2012), with datapoints shaped and coloured according to pollutant and labelled according to reference number [as detailed in Table 1]. R^2X (UN-WWAP, 2015) refers to the proportion of variability explained by principal component 1 (t (UN-WWAP, 2015)) and R^2X (Holt-Giménez et al., 2012) refers to the proportion of variability explained by the second principal component.

cation bridging on clay surfaces, surface complexation, and hydrogen bonding (Tolls, 2001) are not accounted for in the model. In terms of cation exchange, even though LogD was accounted for, the charge at which the pharmaceuticals would be found was not accounted for in the model.

3.5. Recommendations for future research

Based on the analysis of the meta-dataset, the following recommendations are proposed to increase the environmental relevance of batch-scale pharmaceutical sorption studies, and therefore support 'upscaling' from batch to column to field studies. Furthermore, the recommendations also align with obtaining results more amenable for meta-analysis:

1. More comprehensive soil characteristics consistently reported, such as CEC, content of amorphous iron and aluminium oxides and specific surface area of soil particles so that future meta-analysis can consider these variables.
2. Performing an equilibration time between the solid and liquid phase prior to adding the pollutant, to emulate field conditions in which soil is likely have been in contact with irrigation water prior to being in contact with the pharmaceutical. Also, comparing with and without this equilibration stage to understand sorption of pharmaceuticals after draught, compared to when soil is wet.
3. Reporting volume of vessel, liquid phase volume and soil mass so that future meta-analysis may assess the effect of scale on mass transfer and its relatability to environmental conditions.
4. Monitor pH during kinetics sorption experiment to account for pH changes due to equilibration with the liquid phase.
5. Specify the method used for soil sterilisation and, when autoclaving is used, specify the conditions (length and temperature) as this would allow to understand how soil characteristics could have been changed.
6. As well as following OECD Guideline 106 as a baseline or benchmark, it would be useful to add one or more set of experiments at more environmentally relevant conditions for example:

   • More (often lower) environmentally relevant initial concentrations of the pharmaceuticals.
   • Comparing sorption from calcium chloride buffer to sorption when using an environmentally relevant liquid phase (e.g., surface water, wastewater effluent or influent) since $CaCl_2$ buffer has greater ionic strength and no organic matter, compared to wastewater reused for irrigation.
   • Compare sorption in standard temperature conditions (i.e., 20 or 25 °C) to other specific temperatures which could be informed by the local climate where the study is carried out.

4. Conclusion

A meta-analysis and meta-dataset machine learning (PCA and PLS modelling) of batch-sorption studies of pharmaceuticals in soil/water systems for the first time elucidated important considerations that can inform on specific variables to target to make batch-sorption experiments more environmentally relevant. The balance between the advantages of the comparability across laboratory-scale experiments based on standardised guidelines, such as the OECD 106 Guideline, needs to be weighed against the importance of laboratory-scale experiments which better emulate environmental conditions. In terms of the environmental relevance, the three main findings were: i) Pharmaceutical concentrations in batch studies being 1 to 5 orders of magnitude above those measured in waters used for soil irrigation; ii) experimental temperatures did not cover the whole environmental range in terms of granularity; iii) more studies are necessary using environmentally relevant aqueous phases. Furthermore, more information could be gained from meta-analysis and modelling of meta-datasets if more references consistently reported soil parameters, such as cation exchange capacity and specific surface area of the soil particles. The four main dependent variables which were shown to have a significant impact on K_d through PLS modelling of

the meta-dataset were i) pharmaceutical water solubility, ii) the equilibration time between soil and liquid phase prior to adding the pharmaceuticals; iii) the soil organic carbon fraction, and iv) the soil sterilisation method. The equilibration time between the soil and liquid phase prior to pollutant addition is likely affecting K_d because soil pores are fully wetted before the addition of the pollutant which can have an impact on mass transfer of the pollutant through the soil pores. In terms of the sterilisation methods, they are necessary, and reporting specific conditions at which they were carried out (e.g., temperature and time) would be beneficial for future meta-analysis. A further step would be to find relationships between batch, column, and real field pharmaceutical sorption studies to be able to adequately model real-life agro-ecosystems and draw conclusions with less experiments needed.

CRediT authorship contribution statement

Andrea-Lorena Garduño-Jiménez: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Writing – review & editing, Visualization, Project administration. **Juan-Carlos Durán-Álvarez:** Writing – review & editing. **Rachel Louise Gomes:** Conceptualization, Methodology, Validation, Resources, Writing – review & editing, Visualization, Supervision, Project administration, Funding acquisition.

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.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scitotenv.2022.155675.

References

Abd-Elwahed, M.S., 2018. Influence of long-term wastewater irrigation on soil quality and its spatial distribution. Ann. Agric. Sci. 63, 191–199. https://doi.org/10.1016/j.aoas.2018.11.004.

Agudelo-Jaramillo, S., Ochoa-Munoz, M., Zuluaga-Diaz, F.I., 2016. Principal Component Analysis for Mixed Quantitative and Qualitative Data Proposal Report Research Practise, Medellin.

Andriamalala, A., Vieublé-Gonod, L., Dumeny, V., Cambier, P., 2018. Fate of sulfamethoxazole, its main metabolite N-ac-sulfamethoxazole and ciprofloxacin in agricultural soils amended or not by organic waste products. Chemosphere 191, 607–615. https://doi.org/10.1016/J.CHEMOSPHERE.2017.10.093.

Andriamalala, A., Vieuble-Gonod, L., Dumeny, V., Cambier, P., 2018. Fate of sulfamethoxazole, its main metabolite N-ac-sulfamethoxazole and ciprofloxacin in agricultural soils amended or not by organic waste products. Chemosphere 191, 607–615. https://doi.org/10.1016/j.chemosphere.2017.10.093.

Angin, I., Yaganoglu, A.V., Turan, M., 2005. Effects of long-term wastewater irrigation on soil properties. J. Sustain. Agric. 26, 31–42. https://doi.org/10.1300/J064v26n03_05.

Azanu, D., Styrishave, B., Darko, G., Weisser, J.J., Abaidoo, R.C., 2018. Occurrence and risk assessment of antibiotics in water and lettuce in Ghana. Sci. Total Environ. 622–623, 293–305. https://doi.org/10.1016/j.scitotenv.2017.11.287.

Berns, A.E., Philipp, H., Narres, H.D., Burauel, P., Vereecken, H., Tappe, W., 2008. Effect of gamma-sterilization and autoclaving on soil organic matter structure as studied by solid state NMR, UV and fluorescence spectroscopy. Eur. J. Soil Sci. 59, 540–550. https://doi.org/10.1111/j.1365-2389.2008.01016.x.

Bialk, H.M., Simpson, A.J., Pedersen, J.A., 2005. Cross-coupling of sulfonamide antimicrobial agents with model humic constituents. Environ. Sci. Technol. 39, 4463–4473. https://doi.org/10.1021/es0500916.

Cen, T., Zhang, X., Xie, S., Li, D., 2020. Preservatives accelerate the horizontal transfer of plasmid-mediated antimicrobial resistance genes via differential mechanisms. Environ. Int. 138, 105544. https://doi.org/10.1016/j.envint.2020.105544.

Chefetz, B., Mualem, T., Ben-Ari, J., 2008. Sorption and mobility of pharmaceutical compounds in soil irrigated with reclaimed wastewater. Chemosphere 73, 1335–1343.

Christou, A., Papadavid, G., Dalias, P., Fotopoulos, V., Michael, C., Bayona, J.M., Piña, B., Fatta-Kassinos, D., 2019. Ranking of crop plants according to their potential to uptake and accumulate contaminants of emerging concern. Environ. Res. 170, 422–432. https://doi.org/10.1016/J.ENVRES.2018.12.048.

Conde-Cid, M., Fernández-Calviño, D., Núñez-Delgado, A., Fernández-Sanjurjo, M.J., Arias-Estévez, M., Álvarez-Rodríguez, E., 2020. Estimation of adsorption/desorption Freundlich's affinity coefficients for oxytetracycline and chlortetracycline from soil properties: experimental data and pedotransfer functions. Ecotoxicol. Environ. Saf. 196, 110584. https://doi.org/10.1016/j.ecoenv.2020.110584.

Cycoń, M., Mrozik, A., Piotrowska-Seget, Z., 2019. Antibiotics in the soil environment—degradation and their impact on microbial activity and diversity. Front. Microbiol. 10. https://doi.org/10.3389/fmicb.2019.00338.

Dalkmann, P., Willaschek, E., Schiedung, H., Bornemann, L., Siebe, C., Siemens, J., 2014. Long-term wastewater irrigation reduces sulfamethoxazole sorption, but not ciprofloxacin binding, in mexican soils. J. Environ. Qual. 43, 964–970. https://doi.org/10.2134/jeq2013.11.0473.

Dalkmann, P., Siebe, C., Amelung, W., Schloter, M., Siemens, J., 2014. Does long-term irrigation with untreated wastewater accelerate the dissipation of pharmaceuticals in soil? Environ. Sci. Technol. 48, 4963–4970. https://doi.org/10.1021/es501180x.

de Santiago-Martín, A., Meffe, R., Teijón, G., Martínez Hernández, V., López-Heras, I., Alonso Alonso, C., Arenas Romasanta, M., de Bustamante, I., 2020. Pharmaceuticals and trace metals in the surface water used for crop irrigation: risk to health or natural attenuation? Sci. Total Environ. 705, 135825. https://doi.org/10.1016/J.SCITOTENV.2019.135825.

Durán-Álvarez, J.C., Prado, B., Ferroud, A., Juayerk, N., Jiménez-Cisneros, B., 2014. Sorption, desorption and displacement of ibuprofen, estrone, and 17β estradiol in wastewater irrigated and rainfed agricultural soils. Sci. Total Environ. 473–474, 189–198. https://doi.org/10.1016/J.SCITOTENV.2013.12.018.

Durán-Álvarez, J.-C., Prado-Pano, B., Jiménez-Cisneros, B., 2012. Sorption and desorption of carbamazepine, naproxen and triclosan in a soil irrigated with raw wastewater: estimation of the sorption parameters by considering the initial mass of the compounds in the soil. Chemosphere 88, 84–90. https://doi.org/10.1016/j.chemosphere.2012.02.067.

Evyatar, B.M., Tarchitzky, J., Chen, Y., Shenker, M., Chefetz, B., 2018. Composted biosolids and treated wastewater as sources of pharmaceuticals and personal care products for plant uptake: a case study with carbamazepine. Environ. Pollut. 232, 164–172. https://doi.org/10.1016/J.ENVPOL.2017.09.029.

Fenton, M., Albers, C., Ketterings, Q., 2008. Soil organic matter agronomy fact sheet series. Agron. Fact Sheet Ser. 41, 1–2. https://doi.org/10.1016/0378-4290(80)90012-X.

Gibson, R., Durán-Álvarez, J.C., Estrada, K.L., Chávez, A., Jiménez Cisneros, B., 2010. Accumulation and leaching potential of some pharmaceuticals and potential endocrine disruptors in soils irrigated with wastewater in the Tula Valley, Mexico. Chemosphere 81, 1437–1445. https://doi.org/10.1016/J.CHEMOSPHERE.2010.09.006.

Goulas, A., Sabourin, L., Asghar, F., Haudin, C.S., Benoit, P., Topp, E., 2018. Explaining the accelerated degradation of ciprofloxacin, sulfamethazine, and erythromycin in different soil exposure scenarios by their aqueous extractability. Environ. Sci. Pollut. Res. 25, 16236–16245. https://doi.org/10.1007/s11356-018-1834-4.

Hai, F.I., Yang, S., Asif, M.B., Sencadas, V., Shawkat, S., Sanderson-Smith, M., Gorman, J., Xu, Z.Q., Yamamoto, K., 2018. Carbamazepine as a possible anthropogenic marker in water: occurrences, toxicological effects, regulations and removal by wastewater treatment technologies. Water (Switzerland) 10, 1–32. https://doi.org/10.3390/w10020107.

Henderson, B.L., Bui, E.N., 2002. An improved calibration curve between soil pH measured in water and CaCl2. Aust. J. Soil Res. 40, 1399–1405.

Holt-Giménez, E., Shattuck, A., Altieri, M., Herren, H., Gliessman, S., 2012. We already grow enough food for 10 billion people... and still can't end hunger. J. Sustain. Agric. 36, 595–598. https://doi.org/10.1080/10440046.2012.695331.

Hu, S., Zhang, Y., Shen, G., Zhang, H., Yuan, Z., Zhang, W., 2019. Adsorption/desorption behavior and mechanisms of sulfadiazine and sulfamethoxazole in agricultural soil systems. Soil Tillage Res. https://doi.org/10.1016/j.still.2018.10.026.

Jaramillo, M.F., Restrepo, I., 2017. Wastewater reuse in agriculture: a review about its limitations and benefits. Sustainability 9. https://doi.org/10.3390/su9101734.

Jimenez-Cisneros, B., 2006. Irrigation in developing countries using wastewater. IGES Int. Rev. Environ. Strateg. 6, 229–250.

Jovanovic, N.Z., 2008. The use of treated effluent for agricultural irrigation: current status in the Bottelary catchment (South Africa). WIT Trans. Ecol. Environ. 112, 371–380. https://doi.org/10.2495/SI080361.

Kiecak, A., Sassine, L., Boy-Roura, M., Elsner, M., Mas-Pla, J., La Salle, C.L.G., Stumpp, C., 2019. Sorption properties and behaviour at laboratory scale of selected pharmaceuticals using batch experiments. J. Contam. Hydrol. 225. https://doi.org/10.1016/j.jconhyd.2019.103500.

Kočárek, M., Kodešová, R., Vondráčková, L., Golovko, O., Fér, M., Klement, A., Nikodem, A., Jakšík, O., Grabic, R., 2016. Simultaneous sorption of four ionizable pharmaceuticals in different horizons of three soil types. Environ. Pollut. 218, 563–573. https://doi.org/10.1016/j.envpol.2016.07.039.

Kodesova, R., Klement, A., Golovko, O., Fer, M., Nikodem, A., Kocarek, M., Grabic, R., 2019. Root uptake of atenolol, sulfamethoxazole and carbamazepine, and their transformation in three soils and four plants. Environ. Sci. Pollut. Res. 26, 9876–9891. https://doi.org/10.1007/s11356-019-04333-9.

Kookana, R.S., Drechsel, P., Jamwal, P., Vanderzalm, J., 2020. Urbanisation and emerging economies: issues and potential solutions for water and food security. Sci. Total Environ. 732, 139057. https://doi.org/10.1016/j.scitotenv.2020.139057.

Koroša, A., Brenčič, M., Mali, N., 2020. Estimating the transport parameters of propyphenazone, caffeine and carbamazepine by means of a tracer experiment in a coarse-gravel unsaturated zone. Water Res. 175. https://doi.org/10.1016/j.watres.2020.115680.

Li, J., Carter, L.J., Boxall, A.B.A., 2020. Evaluation and development of models for estimating the sorption behaviour of pharmaceuticals in soils. J. Hazard. Mater. 392, 122469. https://doi.org/10.1016/j.jhazmat.2020.122469.

Li, Y., Oliver, D.P., Kookana, R.S., 2018. A critical analysis of published data to discern the role of soil and sediment properties in determining sorption of per and polyfluoroalkyl substances (PFASs). Sci. Total Environ. 628–629, 110–120. https://doi.org/10.1016/j.scitotenv.2018.01.167.

Li, Y., Oliver, D.P., Kookana, R.S., 2018. A critical analysis of published data to discern the role of soil and sediment properties in determining sorption of per and polyfluoroalkyl substances (PFASs). Sci. Total Environ. 628–629, 110–120. https://doi.org/10.1016/J.SCITOTENV.2018.01.167.

Li, Y., Sallach, J.B., Zhang, W., Boyd, S.A., Li, H., 2019. Insight into the distribution of pharmaceuticals in soil-water-plant systems. Water Res. 38–46. https://doi.org/10.1016/j.watres.2018.12.039.

Llamas, M., Vadillo-Pérez, I., Candela, L., Jiménez-Gavilán, P., Corada-Fernández, C., Castro-Gámez, A.F., 2020. Screening and distribution of contaminants of emerging concern and regulated organic pollutants in the heavily modified Guadalhorce river basin, southern Spain. Water (Switzerland) 12, 1–19. https://doi.org/10.3390/w12113012.

Lotrario, J.B., Stuart, B.J., Lam, T., Arands, R.R., O'Connor, O.A., Kosson, D.S., 1995. Effects of sterilization methods on the physical characteristics of soil: implications for sorption isotherm analyses. Bull. Environ. Contam. Toxicol. 54, 668–675. https://doi.org/10.1007/BF00206097.

Malachi, T., Maor, Y., Tadmor, G., Shenker, M., Chefetz, B., 2014. Irrigation of root vegetables with treated wastewater: evaluating uptake of pharmaceuticals and the associated human health risks. Environ. Sci. Technol. 48 (16), 9325–9333. https://doi.org/10.1021/es5017894.

Marti, E., Variatza, E., Balcazar, J.L., 2014. The role of aquatic ecosystems as reservoirs of antibiotic resistance. Trends Microbiol. 22, 36–41. https://doi.org/10.1016/j.tim.2013.11.001.

Martínez-Hernández, V., Meffe, R., Herrera López, S., de Bustamante, I., 2016. The role of sorption and biodegradation in the removal of acetaminophen, carbamazepine, caffeine, naproxen and sulfamethoxazole during soil contact: a kinetics study. Sci. Total Environ. 559, 232–241. https://doi.org/10.1016/J.SCITOTENV.2016.03.131.

Minasny, B., Mcbratney, A.B., Brough, D.M., Jacquier, D., 2011. Models relating soil pH measurements in water and calcium chloride that incorporate electrolyte concentration. Eur. J. Soil Sci. 62, 728–732. https://doi.org/10.1111/j.1365-2389.2011.01386.x.

OECD, 2000. Guideline 106 for the Testing of Chemicals: Adsorption - Desorption Using a Batch Equilibrium Method, pp. 1–44.

Oshunsanya, S.Odutola, 2018. Introductory chapter: relevance of soil pH to agriculture. Soil PH Nutr. Availab. Crop Perform. IntechOpen, pp. 3–6 https://doi.org/10.5772/intechopen.82551.

Paz, A., Tadmor, G., Malchi, T., Blotevogel, J., Borch, T., Polubesova, T., Chefetz, B., 2016. Fate of carbamazepine, its metabolites, and lamotrigine in soils irrigated with reclaimed wastewater: sorption, leaching and plant uptake. Chemosphere 160, 22–29. https://doi.org/10.1016/J.CHEMOSPHERE.2016.06.048.

Petrie, B., Barden, R., Kasprzyk-Hordern, B., 2015. A review on emerging contaminants in wastewaters and the environment: current knowledge, understudied areas and recommendations for future monitoring. Water Res. 72, 3–27. https://doi.org/10.1016/j.watres.2014.08.053.

Qian, Y.L., Mecham, B., 2005. Long-term effects of recycled wastewater irrigation on soil chemical properties on golf course fairways. Agron. J. 97, 717–721. https://doi.org/10.2134/agronj2004.0140.

Schaffer, M., Boxberger, N., Börnick, H., Licha, T., Worch, E., 2012. Sorption influenced transport of ionizable pharmaceuticals onto a natural sandy aquifer sediment at different pH. Chemosphere 87, 513–520. https://doi.org/10.1016/j.chemosphere.2011.12.053.

Singh, A.K., 2016. Chapter 8 - nanoparticle ecotoxicology. In: Singh, A.K.B.T.-E.N. (Ed.)Academic Press, Boston, pp. 343–450 https://doi.org/10.1016/B978-0-12-801406-6.00008-X.

Skipper, H.D., Westermann, D.T., 1973. Comparative effects of propylene oxide, sodium azide, and autoclaving on selected soil properties. Soil Biol. Biochem. 5, 409–414. https://doi.org/10.1016/0038-0717(73)90067-9.

Syms, C., 2018. Principal components analysis. Encycl. Ecol., 566–573 https://doi.org/10.1016/B978-0-12-409548-9.11152-2.

Thebo, A.L., Drechsel, P., Lambin, E.F., Nelson, K.L., 2017. A global, spatially-explicit assessment of irrigated croplands influenced by urban wastewater flows. Environ. Res. Lett. 12. https://doi.org/10.1088/1748-326/aa75d1.

Thiebault, T., 2020. Sulfamethoxazole/Trimethoprim ratio as a new marker in raw wastewaters: a critical review. Sci. Total Environ. 715, 136916. https://doi.org/10.1016/j.scitotenv.2020.136916.

Tolls, J., 2001. Sorption of veterinary pharmaceuticals in soils: a review. Environ. Sci. Technol. 35, 3397–3406. https://doi.org/10.1021/es0003021.

Tseng, R.L., Wu, F.C., 2008. Inferring the favorable adsorption level and the concurrent multistage process with the freundlich constant. J. Hazard. Mater. 155, 277–287. https://doi.org/10.1016/j.jhazmat.2007.11.061.

Uddin, M., Chen, J., Qiao, X., Tian, R., Zhu, M., 2020. Insight into dynamics and bioavailability of antibiotics in paddy soils by in situ soil moisture sampler. Sci. Total Environ. 703, 135562. https://doi.org/10.1016/j.scitotenv.2019.135562.

Umetrics-Sartorious, 2019. Webniar: an OPLS Multivariate Solver. https://umetrics.com/services/webinars/oplsr-based-multivariate-solver. (Accessed 20 June 2020).

UN-WWAP, 2015. The United Nations World Water Development Report 2015: Water for a Sustainable World doi:978-92-3-100071-3.

USEPA, 1999. Understanding variation in partition coefficient, Kd, values. Volume I: The Kd Model, Methods of Measurement, and Application of Chemical Reaction Codes. I. Environ. Prot. Agency, p. 212. https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=4&ved=0ahUKEwi045ai9v3NAhVMQSYKHV9wD0QQFgg0MAM&url=https%3A%2F%2Fwww.epa.gov%2Fsites%2Fproduction%2Ffiles%2F2015-05%2Fdocuments%2F402-r-99-004a.pdf&usg=AFQjCNEMons5lWXRBeJ0TNYK7FnQt8tGpA&sig2=d.

Wishart, D., Feunang, Y., Guo, A., Lo, E., Marcu, A., Grant, J., Sajed, T., Johnson, D., Li, C., Sayeeda, Z., Assempour, N., Iynkkaran, I., Liu, Y., Maciejewski, A., Gale, N., Wilson, A., Chin, L., Cummings, R., Le, D., Pon, A., Knox, C., Wilson, M., Wishart, D., Feunang, Y., Guo, A., Lo, E., Marcu, A., Grant, J., Sajed, T., Johnson, D., Li, C., Sayeeda, Z., Assempour, N., Iynkkaran, I., Liu, Y., Maciejewski, A., Gale, N., Wilson, A., Chin, L., Cummings, R., Le, D., Pon, A., Knox, C., Wilson, M., 2017. DrugBank 5.0: a major update to the DrugBank database for 2018. Nucleic Acids Res. https://doi.org/10.1093/nar/gkx1037.

Wojsławski, J., Białk-Bielińska, A., Stepnowski, P., Dołżonek, J., 2019. Leaching behavior of pharmaceuticals and their metabolites in the soil environment. Chemosphere 231, 269–275. https://doi.org/10.1016/j.chemosphere.2019.05.031.

Worley, B., Powers, R., 2013. Multivariate analysis in metabolomics. Curr. Metabolomics 1, 92–107. https://doi.org/10.2174/2213235x11301010092.

Yu, W., Du, B., Fan, G., Yang, S., Yang, L., Zhang, M., 2020. Spatio-temporal distribution and transformation of 17α- and 17β-estradiol in sterilized soil: a column experiment. J. Hazard. Mater. 389, 122092. https://doi.org/10.1016/j.jhazmat.2020.122092.