Soil metabolomics: A powerful tool for predicting and specifying pesticide sorption

Jeanne Dollinger, Pierre Pétriacq, Amélie Flandin, Anatja Samouelian

To cite this version:

Jeanne Dollinger, Pierre Pétriacq, Amélie Flandin, Anatja Samouelian. Soil metabolomics: A powerful tool for predicting and specifying pesticide sorption. Chemosphere, 2023, 337, pp.139302. 10.1016/j.chemosphere.2023.139302 . hal-04167438

HAL Id: hal-04167438
https://hal.inrae.fr/hal-04167438v1

Submitted on 7 Sep 2023

Soil metabolomics: a powerful tool for predicting and specifying pesticide sorption

Jeanne Dollinger[1], Pierre Pétriacq[2,3], Amélie Flandin[2,3], Anatja Samouelian[1]*

(1): UMR LISAH, Université Montpellier, INRAE, IRD, Institut Agro, 34060 Montpellier, France
(2): Univ. Bordeaux, INRAE, UMR1332 BFP, 33882 Villenave d'Ornon, France
(3): Bordeaux Metabolome, MetaboHUB, PHENOME-EMPHASIS, 33140 Villenave d'Ornon, France

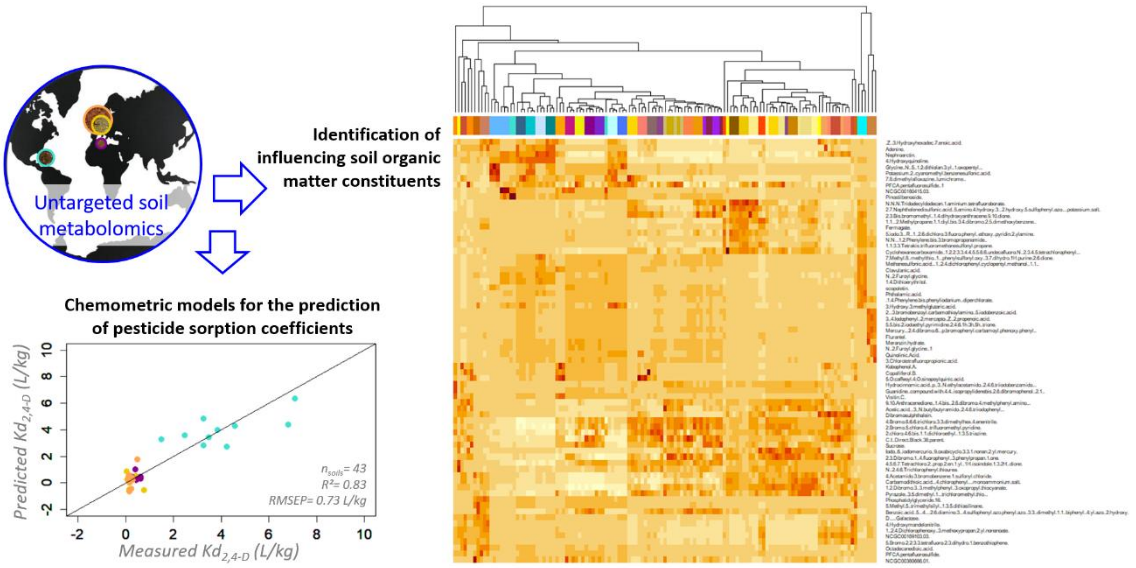
(*) corresponding author: jeanne.dollinger@inrae.fr

Abstract:

Sorption regulates the dispersion of pesticides from cropped areas to surrounding water bodies as well as their persistence. Assessing the risk of water contamination and evaluating the efficiency of mitigation measures, requires fine-resolution sorption data and a good knowledge of its drivers. This study aimed to assess the potential of a new approach combining chemometric and soil metabolomics to estimate the adsorption and desorption coefficients of a range of pesticides. It also aims to identify and characterize key components of soil organic matter (SOM) driving the sorption of these pesticides. We constituted a dataset of 43 soils from Tunisia, France and Guadeloupe (West Indies), covering extensive ranges of texture, organic carbon and pH. We performed untargeted soil metabolomics by liquid chromatography coupled with high-resolution mass spectrometry (UPLC-HRMS). We measured the adsorption and desorption coefficients of three pesticides namely glyphosate, 2,4-D and difenoconazole for these soils. We developed Partial Least Square Regression (PLSR) models for the prediction of the sorption coefficients from the RT-m/z matrix and conducted further ANOVA analyses to identify, annotate and characterise the most significant constituents of SOM in the PLSR models. The curated metabolomics matrix yielded 1213 metabolic markers. The prediction performance of the PLSR models was generally high for the adsorption coefficients Kd_{ads} ($0.3 < R^2 < 0.8$) and for the desorption coefficients Kf_{des} ($0.6 < R^2 < 0.8$) but low for n_{des} ($0.03 < R^2 < 0.3$). The most significant features in the predictive models were annotated with a confidence level of 2 or 3. The molecular descriptors of these putative compounds suggest that the pool of SOM compounds driving glyphosate sorption is reduced compared to 2,4-D and difenoconazole, and these compounds are generally more polar. This approach can provide estimates of the adsorption and desorption coefficients of pesticides, including polar pesticide, for contrasted pedoclimates.

Keywords: *Metabolomics; PLSR; UPLC-HRMS; Soil organic matter; Pesticide; Sorption coefficient*.

Graphical abstract:



Highlights

- We used soil metabolomics to predict and specify sorption for a range of pesticides
- Prediction performance of Kd_{ads} and Kf_{des} from soil metabolomics was good
- Prediction performance was lower for glyphosate than for 2,4-D and difenoconazole
- The pool of SOM compounds driving glyphosate sorption is reduced and more polar

Introduction

Over three million tons of synthetic pesticides are spread annually in the world to protect crops from pests and weeds (Sharma et al., 2019). This extensive use of pesticides in agriculture threatens the health of terrestrial and freshwater ecosystems worldwide (Sharma et al., 2019; Tang et al., 2021). Their persistence and offsite transport from agricultural plots to surrounding ecosystems have generated globalised contamination of surface and groundwater bodies used, among other anthropic usages, for drinking water production (Malla et al., 2021; Pietrzak et al., 2019; Sharma et al., 2019). Agricultural policies in several regions of the world tend to promote mitigation measures such as implementing buffer zones, innovating farming technics or restricted spraying areas near vulnerable drinking water wells (Farenhorst, 2006; Reichenberger et al., 2007; Srivastav, 2020). Assessing the risk of water contamination by pesticides and evaluating the efficiency of mitigation measures, requires implementing modelling approaches at the watershed scale (Dagès et al., 2023; Farenhorst, 2006; Gatel et al., 2019; Mottes et al., 2014).

The dispersion of pesticides by runoff or leaching is regulated mainly by sorption mechanisms (Farenhorst, 2006; Kookana et al., 2014; Tang et al., 2012). Sorption also influences their persistence as it modulates their bioavailability to degrading microorganisms (Kookana et al., 2014). Therefore, sorption coefficients are the most sensitive parameters in models simulating the fate of pesticides in cropped watersheds (Farenhorst, 2006; Wauchope et al., 2002). Yet, conventional laboratory methods for measuring sorption coefficients are extremely time-consuming and expensive (Forouzangohar et al., 2009). The current challenge is to gain insight into the sorption mechanisms to identify and design suitable mitigation measures while generating fine-resolution sorption data for accurate parametrisation of the risk assessment tools (models/indicators). This requires developing methodologies for both predicting and specifying sorption mechanisms for a range of pesticides and pedoclimates.

The estimation of sorption coefficients is traditionally based on the Koc. However, a significant discrepancy in Koc ranges for all pesticides, especially for polar pesticides, has been reported (PPDB, 2023). While soil organic carbon (SOC) is indeed a major determinant of pesticide sorption (Weber et al., 2004), not only its content but also its nature determines the extent of pesticide sorption (Farenhorst, 2006; Kookana et al., 2014). Soil organic matter (SOM) is actually a complex and very heterogeneous mixture of thousands of molecules (Longnecker and Kujawinski, 2017). Chemometrics approaches for estimating pesticide sorption coefficients based on the functional or structural characterisation of SOM such as NMR or infrared spectroscopy therefore improves the prediction performance (Forouzangohar et al., 2009; Kookana et al., 2014). We hypothesised that characterising SOM at the molecular level with untargeted metabolomics would be a step further in the prediction accuracy, especially in understanding pesticide sorption mechanisms.

Untargeted metabolomics enables the chemical profiling of biologically-derived molecules in a wide range of organisms (plants, microorganisms, algae, etc.) and environmental compartments such as soil or water (Kikuchi et al., 2018; Matich et al., 2019; Pétriacq et al., 2017; van Dam and Bouwmeester, 2016). This analytical technique aims to identify a maximum number of compounds in the 50-2000 Da range (Bell et al., 2022; Swenson et al., 2015). Metabolomics has been used to characterise biomarkers of exposure and effect of pesticides on soil microbial communities, earthworms or plants (Jones et al., 2014; Matich et al., 2019; Simpson and McKelvie, 2009). As it can provide information about the metabolic activity of soil microorganisms (Bell et al., 2022; Rodríguez et al., 2020), it has also been suggested to be a powerful approach to characterise pesticide biodegradation along with other omic approaches (Rodríguez et al., 2020). Yet, it has never been applied to predict and specify pesticide sorption mechanisms.

The objectives of this study were to 1) evaluate the predictive performance of a chemometric approaches based on untargeted metabolomics to estimate the adsorption and desorption coefficients of a range of pesticides, including polar pesticides, and 2) identify and characterise key components of SOM involved in the sorption mechanisms of these pesticides.

2. Material and Methods

2.1 Chemicals

Three pesticides among the most used worldwide to protect a variety of crops, including cereals, orchards or vineyards (Matich et al., 2019; Sharma et al., 2019) and having contrasted physico-chemical properties were selected for this study. These are: glyphosate, a hydrophilic broad-spectrum post-emergence herbicide (logP -6.28); 2,4-D, a hydrophilic selective post-emergence herbicide (logP -0.82) and difenoconazole, a hydrophobic systemic fungicide (logP 4.36) (PPDB, 2023).

Glyphosate has a very high aqueous solubility (100 g/L) and is a zwitterion under pH 10.2 (PPDB, 2023). 2,4-D also has a very high aqueous solubility (24 g/L) but is negatively charged under environmental pH ranges (PPDB, 2023). Difenoconazole has a low aqueous solubility (15 mg/L) and is uncharged under environmental pH ranges (PPDB, 2023).

Non-labeled glyphosate, 2,4-D and difenoconazole were supplied by Merck and 14C-labeled pesticides by ISOBIO (Fleurus, Belgium). Sodium azide, calcium chloride, methanol and dichloromethane were supplied by Merck. Methyl vanillate (n° CAS 3943-74-6) was supplied by Sigma Aldrich. All the chemicals used were HPLC grade.

2.2 Origin and characterisation of the soils

A set of 43 soils sampled in four locations across the word was constituted with the purpose of covering an extended range of physico-chemical properties. Ten soils were sampled in Guadeloupe, West Indies (WI), over a toposequence of volcanic ash soils. The climate in this area is tropical and the main crops are sugar cane and banana. The other soils were collected in three sites characterised by a Mediterranean climate. Six soils were sampled in the Lebna peninsula, Tunisia (TU), where crops are frequently rotating from vegetables to cereals or pasture. The other soils were sampled in two vineyard catchments in southern France, the Roujan and Rieutor watersheds (FR-RO and FR-RI), a few kilometers apart but characterised by contrasted soils due to variations of underground rocks and pedogenesis processes. Some of the TU and FR-RO soils were sampled in un-cropped areas of the sites such as fallows, hedgerows, grass strips or ditches to diversify the type and content of organic carbon.

The texture, organic carbon content (OC), pH_{H2O} and cationic exchange capacity (CEC) were measured with standardised methods at the INRAE LAS laboratory (Arras, France) for both FR-RO and FR-RI soils and at the Cirad US 49 laboratory (Montpellier, France) for WI and TU soils (specific habilitation for analysing foreign soils). These properties are displayed in Figure 1.

2.3 Measurement of the sorption coefficients

Both adsorption and desorption isotherms were characterised for all soils. The adsorption batch test procedure was designed following the OECD guidelines n°106 (OECD, 2000). 14C-labelled glyphosate, 2,4-D and difenoconazole were used for the experiments. The concentration of the solutions used were 5, 10, 50, 100 and 1000 µg/L. All these solutions were composed of 50% labelled/non-labelled pesticides. The background electrolyte was composed of 0.01M $CaCl_2$ plus 200 mg/L NaN_3 except for glyphosate for which the background electrolyte contained only NaN_3 to avoid an artificial increase of its sorption by cation bridging (Dollinger et al., 2015). The solid-to-liquid ratio for all materials was 1:10 (g/mL). Solid matrices were equilibrated for 24h with the pesticide in glass tubes at a shaking speed of 150 rpm. The tubes were then centrifuged at 3000 rpm (1770 g) for 10 min, and the supernatant was sampled and analysed by liquid scintillation (LSC). The experiments were all conducted in triplicates.

Following the adsorption phase, a five-step desorption was performed for all soils previously equilibrated with the 100 µg/L pesticide solutions. The supernatant was removed and replaced by an equivalent volume of fresh electrolyte. After 24h shaking, the tubes were centrifuged, and the supernatant was sampled for LSC analysis and replaced by fresh electrolyte.

Both linear (Equation 1) and Freundlich (Equation 2) models were fitted to the adsorption isotherms. Given the excellent linearity of the adsorption isotherms ($0.91<n_{ads}<1.01$), only the linear adsorption coefficients Kd_{ads} are used for the rest of the study. However, the desorption isotherms are non-linear. Therefore, the Freundlich Kf_{des} and n_{des} coefficients are used. The ranges of Kd_{ads}, Kf_{des} and n_{des}

are presented in Figure 2. The adsorption being linear, n_{des} provides an estimation of the desorption hysteresis that is considered significant when $H < 0.70$ ($H = n_{des}/n_{ads}$).

$$Cs = Kd * Caq \quad \text{(Equation 1)}$$

$$Cs = Kf * Caq^{n} \quad \text{(Equation 2)}$$

Where Caq is the concentration in the aqueous phase at equilibrium (µg/L), Kd the linear sorption coefficient (L/kg), K_f ([µg/kg]/[µg/L]^{n}) and n (-) are the Freundlich coefficients and Cs the concentration in the soil (µg/kg).

2.4 Extraction and UPLC-HRMS analysis of soils

Soils were air-dried to a humidity of ≈ 10% and sieved at 2 mm prior to extraction. Methanol is generally used as an extraction solvent for metabolomics fingerprinting of soils (Bell et al., 2022; Jones et al., 2014; Swenson et al., 2015). Given their contrasted polarity, both methanol and dichloromethane (DCM) were selected as extraction solvents. This aimed to enlarge the range of extractible metabolites and target polar to apolar compounds likely involved in pesticide sorption mechanisms. Two successive extractions with methanol and a third with dichloromethane were performed. Methyl vanillate was used as an internal standard and added to the extraction solvents at 25 mg/L. For each soil, five grams (equivalent dry weight) were ultrasonicated for an hour with 10 mL of methanol in glass tubes. The tubes were then centrifuged for 10 min at 3000 rpm (1770 g) and the supernatant was collected. Successively, 30 min ultrasonication with 10 mL methanol and 15 min ultrasonication with 10 mL DCM were performed. The extracts were collected, gathered, dried under nitrogen flux until dryness and suspended in 3 mL methanol. The extracts were then filtered with Nalgen 0.02 µm PTFE filters. Extracts were stored in HPLC amber vials in the freezer at -18 °C until analysis. For each soil, the extraction was performed in triplicates. Five methodological blanks were processed in the same way.

Untargeted metabolomics of soil samples was performed by liquid chromatography coupled with high-resolution mass spectrometry (UPLC-HRMS) using a protocol developed by *Bordeaux Metabolome*, as previously described (Dussarrat et al., 2022). Briefly, we used an Ultimate 3000 ultra-high-pressure liquid chromatography (UHPLC) system coupled to an LTQ-Orbitrap Elite mass spectrometer interfaced with an electrospray (ESI) ionisation source (ThermoScientific, Bremen, Germany), controlled by Thermo XCalibur v.3.0.63 software. Chromatographic separation was achieved at a flow rate of 350 µL/min using a GEMINI UHPLC C18 column (150 × 2 mm, 3 µm, Le Pecq, Phenomenex, France) coupled to a C18 SecurityGuard GEMINI pre-column (4 × 2 mm, 3 µm, Le

Pecq, Phenomenex, France). The column was maintained at 35 °C, and the injection volume was 5 µL. The mobile phase consisted of solvent (1) (0.05% (v/v) formic acid in water) and solvent (2) (acetonitrile) with the following gradient: 0–0.5 min 3% (2), 0.5–1 min 3% (2), 1–9 min 50% (2), 9–13 min 100% (2), 13–14 min 100% (2), 14–14.5 min 3% (2), 14.5–18 min 3% (2). Ionisation was performed in negative mode with the following parameters: ESI- (Heater temp: 300 ◦C, Sheath Gas Flow Rate: 45 (arb), Aux Gas Flow Rate: 15 (arb), Sweep Gas Flow Rate: 10 (arb), I Spray Voltage: 2.5 kV, Capillary Temp: 300 ◦C, S-Lens RF Level: 60%). Prior to analyses, the LTQ-Orbitrap was calibrated by infusing a solution of the calibration (Pierce© ESI Negative Ion Calibration Solution (ref: 88324). Sixteen QC samples (*i.e. a* pool of 15 µL of each sample extract) and five methodological blanks were injected to correct for mass spectrometer signal drift, and to filter out variables detected in blanks, respectively. MS2 Data Dependent Analysis (DDA) was performed on all samples, QC and blank extracts to generate fragmentation information for further annotation with the following parameters: FTMS (50 - 1500 Da) at a resolution of 60k at 200 *m/z*; activation type, CID; isolation width, 1 Da; normalised collision energy, 35 eV; activation Q, 0.250; activation time, 10 ms). In addition, high-resolution MS1 full scan detection of ions was performed for 3 QC samples by FTMS (50 - 1500 Da) at a resolution of 240k at 200 *m/z*.

2.5 Metabolomic workflow

Raw UPLC-HRMS data were processed using MS-DIAL v 4.90 (Tsugawa et al., 2015), yielding 17 770 RT-*m/z* features (parameter report available as Supplementary material, Appendix A). Briefly, MS-DIAL parameters were as follows: MS1, tolerance, 0.01 Da; MS2 tolerance, 0.025 Da; retention time begin, 0 min; retention time end, 18 min; minimum peak height, 10 000; mass slice width, 0.05 Da; smoothing filter, Linear Weighted Moving Average; smoothing level, 4 scans; minimum peak width, 5 scans; sigma window value, 0.5. After data-cleaning (blank check, SN > 10, CV QC < 30%), 3936 variables were retained for further chemometrics. MS-DIAL annotation of metabolic features was performed using the online library MSMS-Public-Neg-VS16.msp with the following parameters: retention time tolerance, 100 min; accurate mass tolerance (MS1), 0.01 Da; accurate mass tolerance (MS2), 0.05 Da; identification score cut off, 80%. Putative annotation of differentially expressed metabolites resulted from MS-DIAL screening of the MS1 detected exact HR *m/z* and MS2 fragmentation patterns against multiple online databases (http://prime.psc.riken.jp/compms/msdial/main.html#MSP) (Tsugawa et al., 2015).

Next, manual curation of the data matrix was performed. Features present in less than three soils were removed. The matrix was also screened for false positives. Therefore, features with the same retention time, same *m/z* and that correlated significantly were considered the same compound and merged. Out of the 3936 features in the curated matrix only 1213 remained after this step. The values of the matrix were then centered-reduced.

One-way ANOVA analysis was conducted with the aov function of R (R Core Team, 2021) to identify features/metabolites with peak intensities significantly varying among the soils. A heatmap of the 50 most discriminating features (lowest p-values) from ANOVA was constructed with R (Figure 3). Features that could not be annotated with MS-DIAL were annotated with METLIN (https://metlin.scripps.edu/) based on the MS1 exact HR *m/z*. The putative compounds, level of confidence and adducts are displayed in Table 1.

2.6 Chemometric prediction and specification of the pesticide sorption coefficients

Partial least square regression (PLSR) was performed with the pls package of R (R Core Team, 2021) to establish predictive models for K_{ads}, Kf_{des} and n_{des} of the three pesticides. The number of components in the PLSR was adjusted for each predictive model with cross-validation data (lowest RMSEP). The number of components varied from 4 to 9 for K_{ads}, from 5 to 25 for Kf_{des} and from 6 to 20 for n_{des}. Leave-one-out cross-validation was used to evaluate the performance of the predictions. Both the R^2 and RMSEP values were used to evaluate the performance of the models.

In order to gain insight into the adsorption mechanisms, the most discriminant features were selected from the PLSR. The top 20 features from the first component and the 10 top features from the other components were annotated (Supplementary material tables S1 to S3). Several molecular descriptors of these putative compounds, including the topological polar surface area (TPSA) (Å), the H-bound donor and acceptor count and the logP were extracted from PUBCHEM (Figure 4).

3. Results

3.1 Soil properties and pesticide sorption

The set of soils selected for this study covers most of the texture classes (Fig. 1) and an extensive range of SOC (0.46 – 6.50%), pH (4.63 – 8.68) and CEC (5.90 – 48.50 cmol/kg). The untargeted-metabolomic analyses provide information about the nature of SOC at a molecular level. This chemical profiling of SOC reveals that the extracted metabolites significantly differ in their nature and relative proportions among the 43 soils. ANOVA analysis based on the relative peak intensities of these metabolites enables identifying features having contrasted proportions among the studied soils. The ANOVA analysis shows that out of the 1213 metabolites extracted (section 2.5), 1164 had peak intensity being significantly different among the 43 soils (p-value <0.05). The heatmap of the 50 most significant features (lowest p-values) (Figure 3) shows that, except for four soils, the three replicates were grouped under the same sub-clusters. This suggests that each soil has a distinct chemical profile. The WI soils are grouped under two clusters well separated from the other soils. The soils from the other sampling sites were dispatched in several sub-clusters, especially the TU soils (Figure 3). Including more features in the clustering discriminated better the origin of the soils, especially for the FR-RO and FR-RI soils, as displayed in the top 100 features' heatmap (Figures S1) and the heatmap based on all features (data not shown).

Table 1 shows the annotation of the top 50 significant features. Most of these metabolites (49) could be annotated with MS-DIAL or METLIN (see section 2.5) with a confidence level 2 or 3. The putative annotation suggests that about 25% of these metabolites are fatty acids or other organic acids. Their accurate mass range from 114 to 1018 Da and their retention time from 0.9 to 18 min suggesting a large range of structure complexity and of polarity. Indeed, most of these putative compounds contain aromatic moieties as well as polar functional groups. Furthermore, some of these features correlated significantly with one or several other features having similar retention time, suggesting they might be a fraction of a bigger molecule.

The adsorption and desorption coefficients of glyphosate, 2,4-D and difenoconazole covered several orders of magnitude across this set of soils (Figure 2). The sorption behaviour was also contrasted among the three pesticides. Glyphosate has a moderate to high adsorption (Kd_{ads} 3.2 – 28.8 L/kg) and a very strong desorption hysteresis (Kf_{des} 263 – 4844 ([µg/kg]/[µg/L]^n) & n_{des} 0.03 – 0.25). Difenoconazole has a very high adsorption (Kd_{ads} 8.5 – 228.5 L/kg) and a strong desorption hysteresis (Kf_{des} 140 – 4116 ([µg/kg]/[µg/L]^n) & n_{des} 0.03 – 0.65). Last, exception made for the WI soils (K_{dads} 1.5 – 7.1 L/kg, Kf_{des} 189 – 624 ([µg/kg]/[µg/L]^n) & n_{des} 0.11 – 0.40), 2,4-D is weakly adsorbed (Kd_{ads} 0.02 – 0.6 L/kg) and has high to no desorption hysteresis (Kf_{des} 0 – 21 ([µg/kg]/[µg/L]^n) & n_{des} 0.03 – 1.55). 2,4-D was so weakly adsorbed on FR-RO soils and the desorption so elevate that accurate measurement of Kf_{des} and n_{des} for these soils was not possible. There was no 2,4-D left at a quantifiable level after the first desorption step so the values of Kf_{des} were set to 0 and n_{des} to 1 for these FR-RO soils. The correlation of these sorption coefficients with SOC are displayed in figure S2 (supplementary material)

3.2 Chemometric estimation of the pesticide sorption coefficients

Figure 5 displays the prediction performance of the PLSR models established to predict the adsorption and desorption coefficients of glyphosate, 2,4-D and difenoconazole. This performance, featured by the R^2 (-) and RSMEP (L/kg) values, varies across the range of coefficients and pesticides considered. The performance increase from glyphosate <difenoconazole < 2,4-D. For difenoconazole and 2,4-D the prediction performance is good for Kd_{ads} and Kf_{des} but the regressions are driven by the high values of the WI soils. It is interesting to note that the glyphosate-Kd_{ads} PLSR has only 4 components while 2,4-D and difenoconazole have 8 and 9 components, respectively. This suggests that the range of metabolites involved in the adsorption of glyphosate is reduced compared to 2,4-D and difenoconazole.

The putative annotation of the most significant features in the PLSR models (highest absolute loading weight (see section 2.6)) is presented in the supplementary material (Tables S1 to S3). 41 compounds were annotated for glyphosate, 70 for 2,4-D and 83 for difenoconazole. It is interesting to note that only one of these compounds was in these three datasets. This compound is also the only one that the glyphosate and difenoconazole datasets have in common. There are six compounds commons to the glyphosate and 2,4-D datasets and 22 commons to the 2,4-D and difenoconazole datasets. Few compounds of the 2,4-D (4) and difenoconazole (3) datasets (Tables S2 & S3) are in the top 50 compounds discriminating most the SOM molecular profiles listed in Table 1.

Figure 4 shows the density functions of six molecular descriptors of these putatively annotated datasets of compounds. Among these six descriptors only the accurate mass (*m/z*) and the retention time (RT) are independent of the annotation. The TPSA, the H-bound donor and acceptor counts and logP depend on the feature annotations that were achieved at the confidence levels 2-3. In general, the distribution of these molecular descriptors is larger for difenoconazole than for glyphosate and 2,4-D has intermediate distribution patterns. It is interesting to note that the putative metabolites with the highest TPSA (≈ 500 Å) are found only in the glyphosate dataset that also have a greater proportion of compounds in the 100-200 Å TPSA range compared to the 2,4-D and difenoconazole datasets. There is also an occurrence of compounds with very low logP in the glyphosate dataset that is not observed in the two other datasets. Last, there is a lower proportion of compounds with high RT (10-18 min) in the glyphosate dataset. This suggests that the pool of metabolite involved in the adsorption of glyphosate is more polar than those involved in the adsorption of 2,4-D and difenoconazole.

4. Discussion

The adsorption and desorption coefficients measured for the three selected pesticides cover a range of several orders of magnitude (Figure 2). This was expected from the great variability of physico-chemical properties of the 43 sampled soils (Figure 1).

For 2,4-D and difenoconazole, the measured Kd_{ads} values cover the entire ranges of Kd reported in the literature (Akyol et al., 2021; Godeau et al., 2021; Gurson et al., 2019; PPDB, 2023; Wang et al., 2020; Werner et al., 2013). The Kd_{ads} values of glyphosate are in the low-medium range of values reported in the literature (Akyol et al., 2021; Dollinger et al., 2015; Gurson et al., 2019). However, higher Kd values reported for glyphosate were measured with $CaCl_2$ as background electrolyte which significantly and artificially increases the Kd values (Cruz et al., 2007; de Jonge and Wollesen de Jonge, 1999; Dollinger et al., 2015). These extended Kd_{ads} ranges are ideal for testing the global performance of the chemometric estimation approach based on soil metabolomic profiles. However, its site-specific performance, which is the scale targeted by the risk assessment modelling, should be further evaluated.

Desorption data are usually scarce in the literature and the desorption hysteresis is actually not represented in the pesticide fate models. The Kf_{des} and n_{des} coefficients show that the desorption hysteresis is very strong for glyphosate and difenoconazole. It is also relatively strong for 2,4-D on the WI soils (Figure 2). While these coefficients cannot be implemented in the models, they help evaluate the uncertainty of model outputs.

The PLSR performance criteria (R^2 & RMSEP) indicate that for the three pesticides, the metabolomic profile explains, in part, the variability of the Kd_{ads} and Kf_{des} (Figure 5). N_{des} seems to be less influenced by the chemical characteristics of SOM (Figure 5). The prediction performance of the PLSR models is lower for the glyphosate sorption coefficients than for those of 2,4-D and difenoconazole. However, for 2,4-D and difenoconazole, the PLSR are forced by the high WI values which questions the site-specific performance for the other sampling sites. For the three pesticides the coefficient of variation of the Kd_{ads} measures is 7-8% (calculated from the batch replicates see section 2.3). For the average Kd_{ads} values, it represents a disparity of 1.4 L/kg for glyphosate, 3.8 L/kg for difenoconazole and 0.09 L/kg for 2,4-D. The RMSEP are 3 to 8 times higher than these experimental uncertainties but are quite low compared to the Kd_{ads} ranges (Figure 5).

The three pesticides selected for this study cover an extended range of polarity. Glyphosate comprises only polar functional groups and difenoconazole only hydrophobic groups, while 2,4-D contains both polar and hydrophobic moieties. The presence of polar functional groups in the pesticide structure tends to diversify the type of interaction between the pesticide and the soils compared to hydrophobic pesticides (Dollinger et al., 2015; Weber et al., 2004; Werner et al., 2013).

The adsorption coefficients of polar pesticides are generally significantly correlated to the SOM content (Akyol et al., 2021; Gurson et al., 2019; Weber et al., 2004), but also to other soil constituents, such as clay minerals (Dollinger et al., 2015; Weber et al., 2004). They are also very sensitive to soil pH, which dictates their speciation and the surface charges of edaphic constituents (Boivin et al., 2005; Kah and Brown, 2007; Wauchope et al., 2002; Weber et al., 2004).

In particular, the adsorption of glyphosate was reported to be driven chiefly by clay minerals and metal oxides with a strong influence of pH and CEC (Dollinger et al., 2015). Some studies highlighted that the sorption of glyphosate was also strongly influenced by SOM (Akyol et al., 2021; Gurson et al., 2019). However, for the set of studied soils, the influence of SOM was weak (Figure S2). This explains the low performance of the chemometric approach for estimating the adsorption and desorption coefficients of glyphosate from the SOM metabolite profile (Figure 5). Yet the correlations are significant and the RMSEP are lower than for traditional estimation approaches like pedotransfer functions (Dollinger et al., 2015). The number of components in the PLSR models (Figure 5) suggests that the pool of SOM compounds involved in the sorption of glyphosate is low compared to 2,4-D and difenoconazole. The distribution of the molecular descriptors (Figure 4) indicates that these are also more polar. However, the molecular descriptors depend on the annotation that was achieved at confidence levels 2-3 (Blaženović et al., 2018; Chaleckis et al., 2019; Creek et al., 2014).

The metabolomics profile is highly depends on the soil extraction procedure (Bell et al., 2022; Chaleckis et al., 2019; Swenson et al., 2015). Water was not used as an extraction solvent to avoid the extraction of small and very polar metabolites that constitute most of the dissolved organic matter fraction (DOM) (Swenson et al., 2015). Indeed, DOM has a complex and ambiguous role in the sorption of pesticides (Barriuso et al., 2011) that falls behind the scope of this study. With methanol and dichloromethane we targeted larger and less polar compounds. Accurate annotation of the extracted compounds is still the major bottleneck of un-targeted metabolomics (Bell et al., 2022; Blaženović et al., 2018; Chaleckis et al., 2019). The potential of metabolomics is huge to specify the mechanism involved in the fate of pesticides in soils including sorption and degradation (Rodrigues et al., 2013). However, further specification of these mechanisms would require the identification of metabolites with a confidence level of 1 or 2. This therefore points to the lack of representation of soil compounds in metabolomic databases (Pétriacq et al., 2017).

The metabolomic profiling procedure seems relatively stable, as displayed by the clustering of the replicates for 90% of the soils (Figure 3). For the present study, all extracts were injected in the same LC run. UPLC-HRMS analysis is subjected to retention time shifts which hinders the performance of the peak alignment if the extracts are not injected in the same run. Another limit to the number of extracts that can be processed is the multiplication of the detected features. Indeed, multiplying the extracts increases the risk of noise in the matrix (*i.e.* features present in less than 5-10% of the samples). The curation of the metabolomics matrix to remove this noise is time-consuming.

Conclusion

Chemometrics based on metabolomics data is a powerful approach to predict the pesticide adsorption and desorption coefficients for soils having contrasted physico-chemical properties. The prediction performance is lower for glyphosate than for 2,4-D and difenoconazole but higher than for traditional pedotransfer functions. The establishment of the PLSR models is time-consuming. Yet once this is achieved, a single extract can provide estimations for both adsorption and desorption coefficients for the whole range of pesticides tested. Therefore, it is beneficial to diversify the range of pesticides included in the risk assessment modelling. It can also help to refine the resolution of the sorption parametrisation in the models. The approach was tested for a very diverse set of soils, but its site-specific precision remains to be evaluated. It might not be the most rapid and precise estimation methods for the spatialisation of sorption coefficients. However, metabolomics is also a potential indicator of other mechanisms involved in the fate of pesticides, including biodegradation. Its ability to provide information on the biodegradation of pesticides, their presence and their effect on the environment should be further investigated to develop a global estimation approach. Metabolomics also help to gain insight into the sorption mechanisms and the fraction of SOM involved. The spectra databases’ development should help improve the accuracy of the metabolite annotation.

Acknowledgement

The authors would like to thank David Fages and Olivier Grünberger for their help with the soil sampling. We also thank Sandrine Negro and Manon Lagacherie for their help with the measurement of the sorption coefficients. Last we warmly thank Pauline Campan for providing the sorption coefficients and soil physico-chemical properties of the WI soils.

Statements & Declarations

Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author upon reasonable request.

Competing interests

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.

Authors' contributions

Jeanne Dollinger contributed a vast majority of the study conception and design with the help of Anatja Samouelian and Pierre Pétriacq. Jeanne Dollinger performed the soil extractions and the measurement of the sorption coefficients for the FR-RO, FR-RI & TU soils. Data on the WI soils were provided by Anatja Samouelian. UPLC-HRMS-based metabolomis was performed by Pierre Pétriacq and Amelie Flandin. The first draft of the manuscript was written by Jeanne Dollinger, and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding

The study was funded by a young scientist starting grant from the AgroEcoSystem department of INRAE. The authors are grateful for financial support from MetaboHUB (ANR-11-INBS-0010).

Ethical Approval

Not applicable

Consent to Participate

Not applicable

Consent to Publish

Not applicable

References

Akyol, N.H., Carroll, K.C., Ciftci Cortuk, E., Gunduz, O.C., Sahin, N., 2021. Comparison of sorption and solute transport behaviour of several herbicides in an alkaline agricultural soil. Int. J. Environ. Anal. Chem. 0, 1–19. https://doi.org/10.1080/03067319.2021.1969384

Barriuso, E., Andrades, M.-S., Benoit, P., Houot, S., 2011. Pesticide desorption from soils facilitated by dissolved organic matter coming from composts: experimental data and modelling approach. Biogeochemistry 106, 117–133. https://doi.org/10.1007/s10533-010-9481-y

Bell, M.A., McKim, U., Sproule, A., Tobalt, R., Gregorich, E., Overy, D.P., 2022. Extraction methods for untargeted metabolomics influence enzymatic activity in diverse soils. Sci. Total Environ. 828, 154433. https://doi.org/10.1016/j.scitotenv.2022.154433

Blaženović, I., Kind, T., Ji, J., Fiehn, O., 2018. Software Tools and Approaches for Compound Identification of LC-MS/MS Data in Metabolomics. Metabolites 8, 31. https://doi.org/10.3390/metabo8020031

Boivin, A., Cherrier, R., Schiavon, M., 2005. A comparison of five pesticides adsorption and desorption processes in thirteen contrasting field soils. Chemosphere 61, 668–676. https://doi.org/10.1016/j.chemosphere.2005.03.024

Chaleckis, R., Meister, I., Zhang, P., Wheelock, C.E., 2019. Challenges, progress and promises of metabolite annotation for LC–MS-based metabolomics. Curr. Opin. Biotechnol., Analytical Biotechnology 55, 44–50. https://doi.org/10.1016/j.copbio.2018.07.010

Creek, D.J., Dunn, W.B., Fiehn, O., Griffin, J.L., Hall, R.D., Lei, Z., Mistrik, R., Neumann, S., Schymanski, E.L., Sumner, L.W., Trengove, R., Wolfender, J.-L., 2014. Metabolite identification: are you sure? And how do your peers gauge your confidence? Metabolomics 10, 350–353. https://doi.org/10.1007/s11306-014-0656-8

Cruz, L.H. da, Santana, H. de, Zaia, C.T.B.V., Zaia, D.A.M., 2007. Adsorption of glyphosate on clays and soils from Paraná State: effect of pH and competitive adsorption of phosphate. Braz. Arch. Biol. Technol. 50, 385–394. https://doi.org/10.1590/S1516-89132007000300004

Dagès, C., Voltz, M., Bailly, J.-S., Crevoisier, D., Dollinger, J., Margoum, C., 2023. PITCH: A model simulating the transfer and retention of pesticides in infiltrating ditches and channel networks for management design purposes. Sci. Total Environ. 891, 164602. https://doi.org/10.1016/j.scitotenv.2023.164602

de Jonge, H., Wollesen de Jonge, L., 1999. Influence of pH and solution composition on the sorption of glyphosate and prochloraz to a sandy loam soil. Chemosphere 39, 753–763. https://doi.org/10.1016/S0045-6535(99)00011-9

Dollinger, J., Dagès, C., Voltz, M., 2015. Glyphosate sorption to soils and sediments predicted by pedotransfer functions. Environ. Chem. Lett. 13, 293–307. https://doi.org/10.1007/s10311-015-0515-5

Dussarrat, T., Prigent, S., Latorre, C., Bernillon, S., Flandin, A., Díaz, F.P., Cassan, C., Van Delft, P., Jacob, D., Varala, K., Joubes, J., Gibon, Y., Rolin, D., Gutiérrez, R.A., Pétriacq, P., 2022. Predictive metabolomics of multiple Atacama plant species unveils a core set of generic metabolites for extreme climate resilience. New Phytol. 234, 1614–1628. https://doi.org/10.1111/nph.18095

Farenhorst, A., 2006. Importance of Soil Organic Matter Fractions in Soil-Landscape and Regional Assessments of Pesticide Sorption and Leaching in Soil. Soil Sci. Soc. Am. J. 70, 1005–1012. https://doi.org/10.2136/sssaj2005.0158

Forouzangohar, M., Cozzolino, D., Kookana, R.S., Smernik, R.J., Forrester, S.T., Chittleborough, D.J., 2009. Direct Comparison between Visible Near- and Mid-Infrared Spectroscopy for Describing Diuron Sorption in Soils. Environ. Sci. Technol. 43, 4049–4055. https://doi.org/10.1021/es8029945

Gatel, L., Lauvernet, C., Carluer, N., Weill, S., Tournebize, J., Paniconi, C., 2019. Global evaluation and sensitivity analysis of a physically based flow and reactive transport model on a laboratory experiment. Environ. Model. Softw. 113, 73–83. https://doi.org/10.1016/j.envsoft.2018.12.006

Godeau, C., Morin-Crini, N., Staelens, J.-N., Martel, B., Rocchi, S., Chanet, G., Fourmentin, M., Crini, G., 2021. Adsorption of a triazole antifungal agent, difenoconazole, on soils from a cereal farm: Protective effect of hemp felt. Environ. Technol. Innov. 22, 101394. https://doi.org/10.1016/j.eti.2021.101394

Gurson, A.P., Ozbay, I., Ozbay, B., Akyol, G., Akyol, N.H., 2019. Mobility of 2,4-Dichlorophenoxyacetic Acid, Glyphosate, and Metribuzine Herbicides in Terra Rossa-Amended Soil: Multiple Approaches with Experimental and Mathematical Modeling Studies. Water. Air. Soil Pollut. 230, 220. https://doi.org/10.1007/s11270-019-4266-y

Jones, O.A.H., Sdepanian, S., Lofts, S., Svendsen, C., Spurgeon, D.J., Maguire, M.L., Griffin, J.L., 2014. Metabolomic analysis of soil communities can be used for pollution assessment. Environ. Toxicol. Chem. 33, 61–64. https://doi.org/10.1002/etc.2418

Kah, M., Brown, C.D., 2007. Prediction of the Adsorption of Ionizable Pesticides in Soils. J. Agric. Food Chem. 55, 2312–2322. https://doi.org/10.1021/jf063048q

Kikuchi, J., Ito, K., Date, Y., 2018. Environmental metabolomics with data science for investigating ecosystem homeostasis. Prog. Nucl. Magn. Reson. Spectrosc. 104, 56–88. https://doi.org/10.1016/j.pnmrs.2017.11.003

Kookana, R.S., Ahmad, R., Farenhorst, A., 2014. Sorption of Pesticides and its Dependence on Soil Properties: Chemometrics Approach for Estimating Sorption, in: Non-First Order Degradation and Time-Dependent Sorption of Organic Chemicals in Soil, ACS Symposium Series. American Chemical Society, pp. 221–240. https://doi.org/10.1021/bk-2014-1174.ch012

Longnecker, K., Kujawinski, E.B., 2017. Mining mass spectrometry data: Using new computational tools to find novel organic compounds in complex environmental mixtures. Org. Geochem. 110, 92–99. https://doi.org/10.1016/j.orggeochem.2017.05.008

Malla, M.A., Gupta, S., Dubey, A., Kumar, A., Yadav, S., 2021. Chapter 7 - Contamination of groundwater resources by pesticides, in: Ahamad, A., Siddiqui, S.I., Singh, P. (Eds.), Contamination of Water. Academic Press, pp. 99–107. https://doi.org/10.1016/B978-0-12-824058-8.00023-2

Matich, E.K., Chavez Soria, N.G., Aga, D.S., Atilla-Gokcumen, G.E., 2019. Applications of metabolomics in assessing ecological effects of emerging contaminants and pollutants on plants. J. Hazard. Mater. 373, 527–535. https://doi.org/10.1016/j.jhazmat.2019.02.084

Mottes, C., Lesueur-Jannoyer, M., Le Bail, M., Malézieux, E., 2014. Pesticide transfer models in crop and watershed systems: a review. Agron. Sustain. Dev. 34, 229–250. https://doi.org/10.1007/s13593-013-0176-3

OECD, 2000. Test No. 106: Adsorption -- Desorption Using a Batch Equilibrium Method. Organisation for Economic Co-operation and Development, Paris.

Pétriacq, P., Williams, A., Cotton, A., McFarlane, A.E., Rolfe, S.A., Ton, J., 2017. Metabolite profiling of non-sterile rhizosphere soil. Plant J. 92, 147–162. https://doi.org/10.1111/tpj.13639

Pietrzak, D., Kania, J., Malina, G., Kmiecik, E., Wątor, K., 2019. Pesticides from the EU First and Second Watch Lists in the Water Environment. CLEAN – Soil Air Water 47, 1800376. https://doi.org/10.1002/clen.201800376

PPDB, 2023. Pesticide Properties Database [WWW Document]. URL http://sitem.herts.ac.uk/aeru/ppdb/en/search.htm (accessed 1.6.23).

R Core Team, 2021. R: A Language and Environment for Statistical Computing.

Reichenberger, S., Bach, M., Skitschak, A., Frede, H.-G., 2007. Mitigation strategies to reduce pesticide inputs into ground- and surface water and their effectiveness; A review. Sci. Total Environ. 384, 1–35. https://doi.org/10.1016/j.scitotenv.2007.04.046

Rodrigues, E.T., Lopes, I., Pardal, M.Â., 2013. Occurrence, fate and effects of azoxystrobin in aquatic ecosystems: A review. Environ. Int. 53, 18–28. https://doi.org/10.1016/j.envint.2012.12.005

Rodríguez, A., Castrejón-Godínez, M.L., Salazar-Bustamante, E., Gama-Martínez, Y., Sánchez-Salinas, E., Mussali-Galante, P., Tovar-Sánchez, E., Ortiz-Hernández, Ma.L., 2020. Omics Approaches to Pesticide Biodegradation. Curr. Microbiol. 77, 545–563. https://doi.org/10.1007/s00284-020-01916-5

Sharma, A., Kumar, V., Shahzad, B., Tanveer, M., Sidhu, G.P.S., Handa, N., Kohli, S.K., Yadav, P., Bali, A.S., Parihar, R.D., Dar, O.I., Singh, K., Jasrotia, S., Bakshi, P., Ramakrishnan, M., Kumar, S., Bhardwaj, R., Thukral, A.K., 2019. Worldwide pesticide usage and its impacts on ecosystem. SN Appl. Sci. 1, 1446. https://doi.org/10.1007/s42452-019-1485-1

Simpson, M.J., McKelvie, J.R., 2009. Environmental metabolomics: new insights into earthworm ecotoxicity and contaminant bioavailability in soil. Anal. Bioanal. Chem. 394, 137–149. https://doi.org/10.1007/s00216-009-2612-4

Srivastav, A.L., 2020. Chapter 6 - Chemical fertilizers and pesticides: role in groundwater contamination, in: Prasad, M.N.V. (Ed.), Agrochemicals Detection, Treatment and Remediation. Butterworth-Heinemann, pp. 143–159. https://doi.org/10.1016/B978-0-08-103017-2.00006-4

Swenson, T.L., Jenkins, S., Bowen, B.P., Northen, T.R., 2015. Untargeted soil metabolomics methods for analysis of extractable organic matter. Soil Biol. Biochem. 80, 189–198. https://doi.org/10.1016/j.soilbio.2014.10.007

Tang, F.H.M., Lenzen, M., McBratney, A., Maggi, F., 2021. Risk of pesticide pollution at the global scale. Nat. Geosci. 14, 206–210. https://doi.org/10.1038/s41561-021-00712-5

Tang, X., Zhu, B., Katou, H., 2012. A review of rapid transport of pesticides from sloping farmland to surface waters: Processes and mitigation strategies. J. Environ. Sci. 24, 351–361. https://doi.org/10.1016/S1001-0742(11)60753-5

Tsugawa, H., Cajka, T., Kind, T., Ma, Y., Higgins, B., Ikeda, K., Kanazawa, M., VanderGheynst, J., Fiehn, O., Arita, M., 2015. MS-DIAL: data-independent MS/MS deconvolution for comprehensive metabolome analysis. Nat. Methods 12, 523–526. https://doi.org/10.1038/nmeth.3393

van Dam, N.M., Bouwmeester, H.J., 2016. Metabolomics in the Rhizosphere: Tapping into Belowground Chemical Communication. Trends Plant Sci., Special Issue: Unravelling the Secrets of the Rhizosphere 21, 256–265. https://doi.org/10.1016/j.tplants.2016.01.008

Wang, F., Cao, D., Shi, L., He, S., Li, X., Fang, H., Yu, Y., 2020. Competitive Adsorption and Mobility of Propiconazole and Difenoconazole on Five Different Soils. Bull. Environ. Contam. Toxicol. 105, 927–933. https://doi.org/10.1007/s00128-020-03034-1

Wauchope, R.D., Yeh, S., Linders, J.B.H.J., Kloskowski, R., Tanaka, K., Rubin, B., Katayama, A., Kördel, W., Gerstl, Z., Lane, M., Unsworth, J.B., 2002. Pesticide soil sorption parameters: theory, measurement, uses, limitations and reliability. Pest Manag. Sci. 58, 419–445. https://doi.org/10.1002/ps.489

Weber, J.B., Wilkerson, G.G., Reinhardt, C.F., 2004. Calculating pesticide sorption coefficients (Kd) using selected soil properties. Chemosphere 55, 157–166. https://doi.org/10.1016/j.chemosphere.2003.10.049

Werner, D., Garratt, J.A., Pigott, G., 2013. Sorption of 2,4-D and other phenoxy herbicides to soil, organic matter, and minerals. J. Soils Sediments 13, 129–139. https://doi.org/10.1007/s11368-012-0589-7

Figures

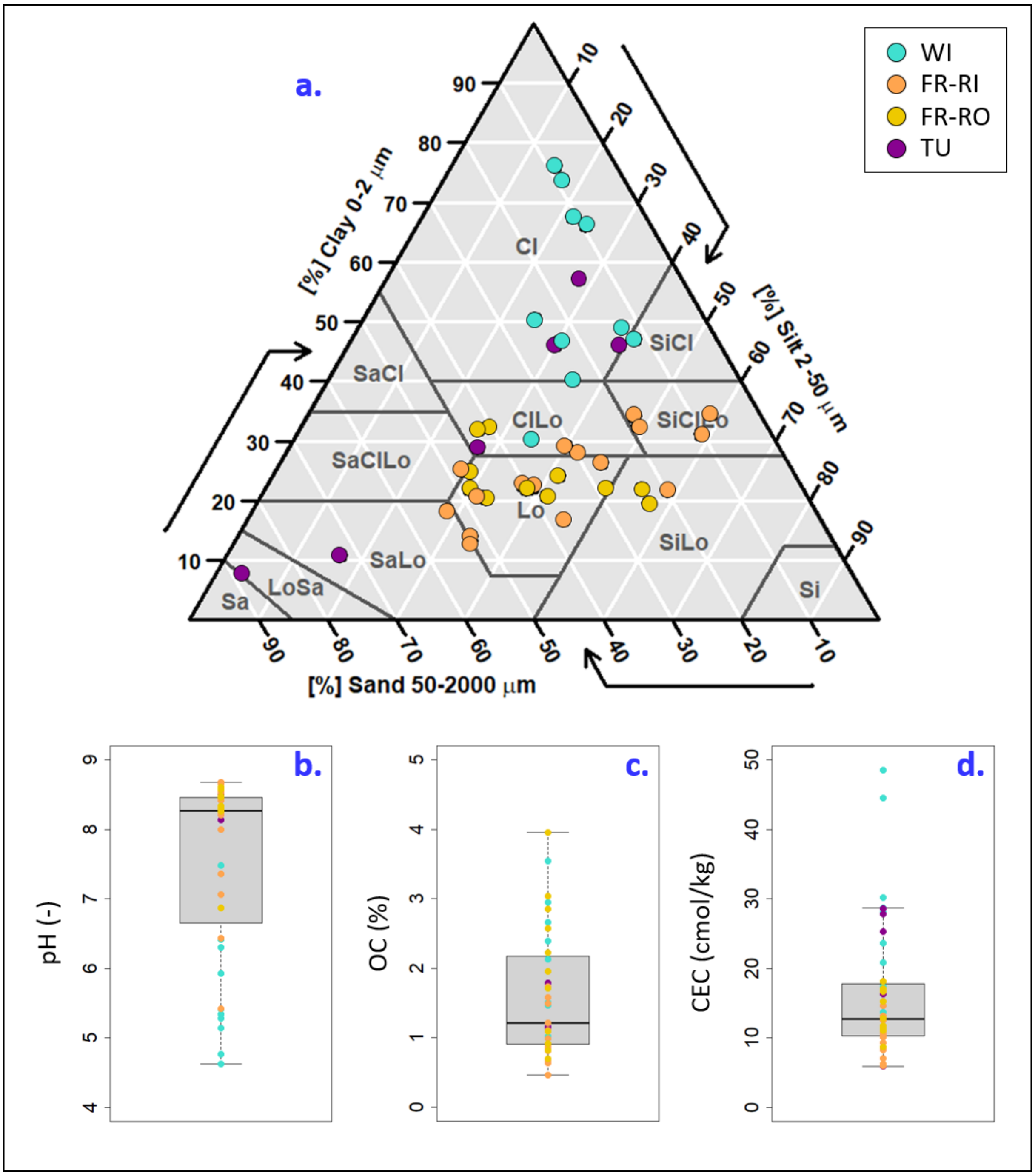


__Figure 1__: __Physico-chemical properties of the soils.__ This set of 43 soils includes soils sampled in Guadeloupe in the French West Indies (WI), in the Cap Bon Peninsula in Tunisia (TU) and two catchments from southern France (FR-RO and FR-RI). The figure displays the texture range (a.), the pH range (b.), the soil organic fraction range (c.) and the cationic exchange capacity range (d.).

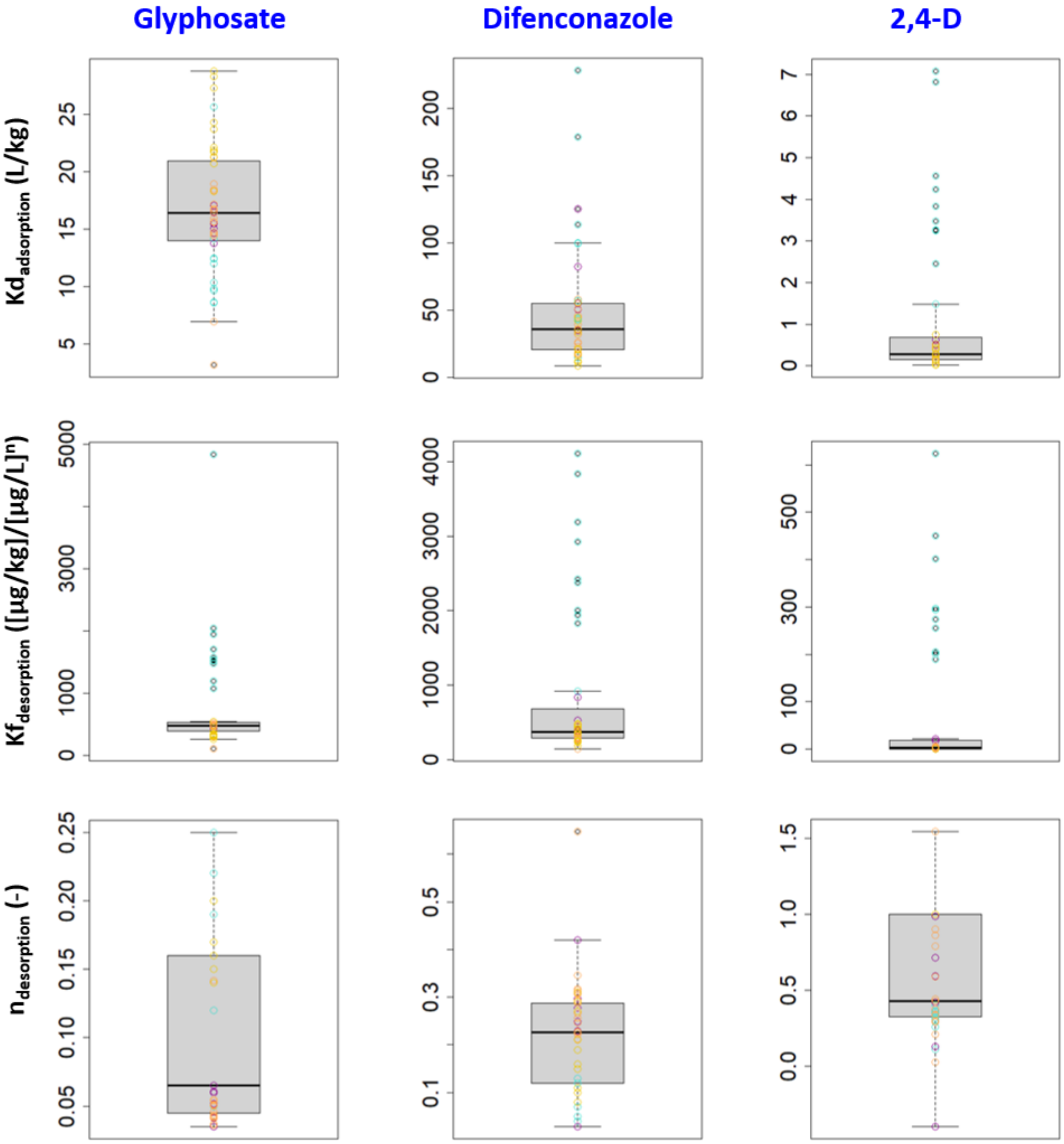


***Figure 2**: **Measured sorption coefficients**. The figure shows the distributions of adsorption (Kd_{ads}) and desorption (Kf_{des} & n_{des}) coefficients measured for the WI soils (turquoise dots), the TU soils (violet dots), the FR-RO (gold dots) and FR-RI (orange dots) soils for the pesticides glyphosate, 2,4-D and difenoconazole.*

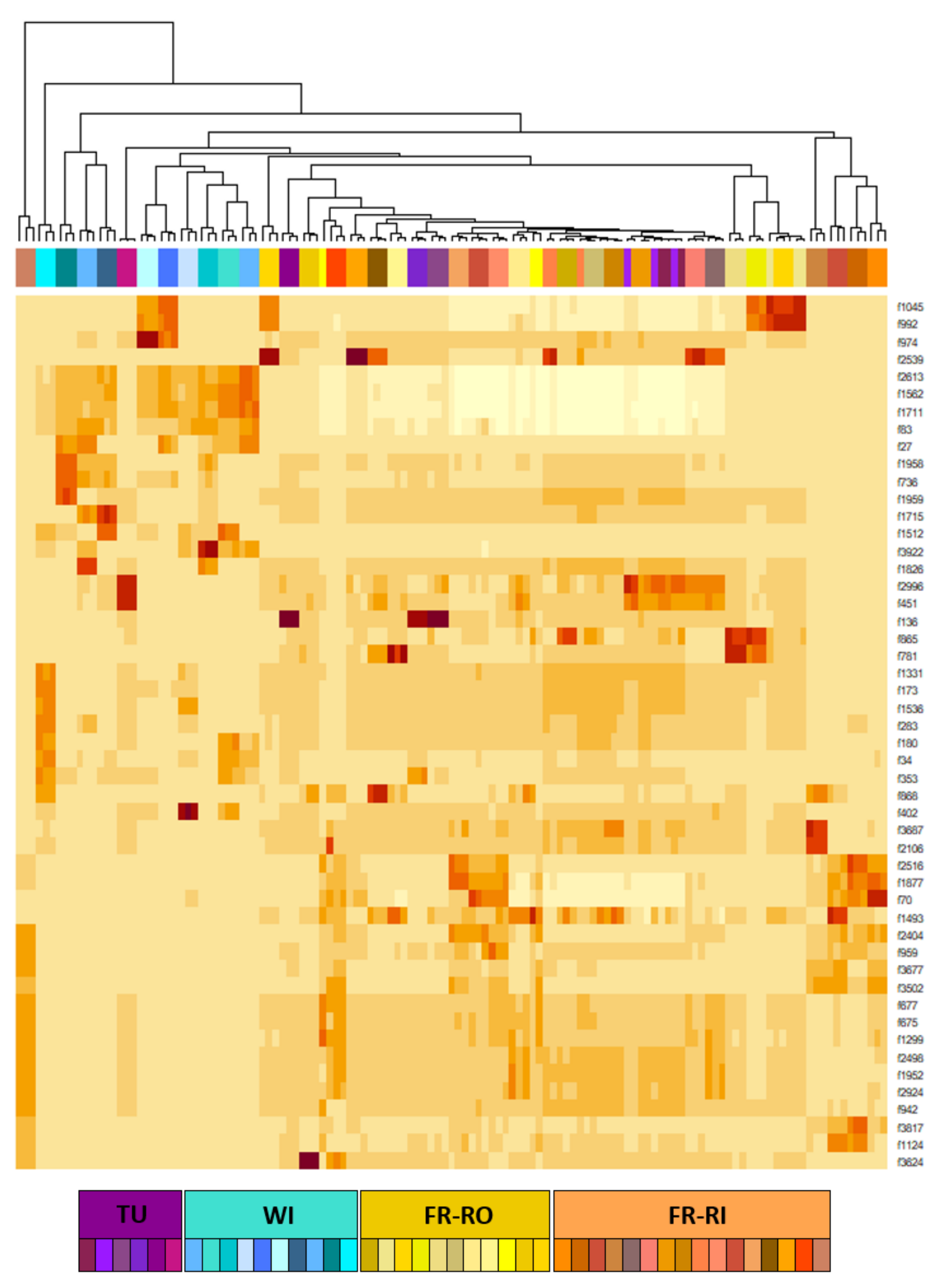


***Figure 3*: *Heatmap of the top 50 discriminant metabolites*.** *The heatmap shows the clustering of the soils according to the relative intensity of the top 50 metabolites identified by an ANOVA analysis. Each lines represent a given metabolite that is further described in Table 1.*

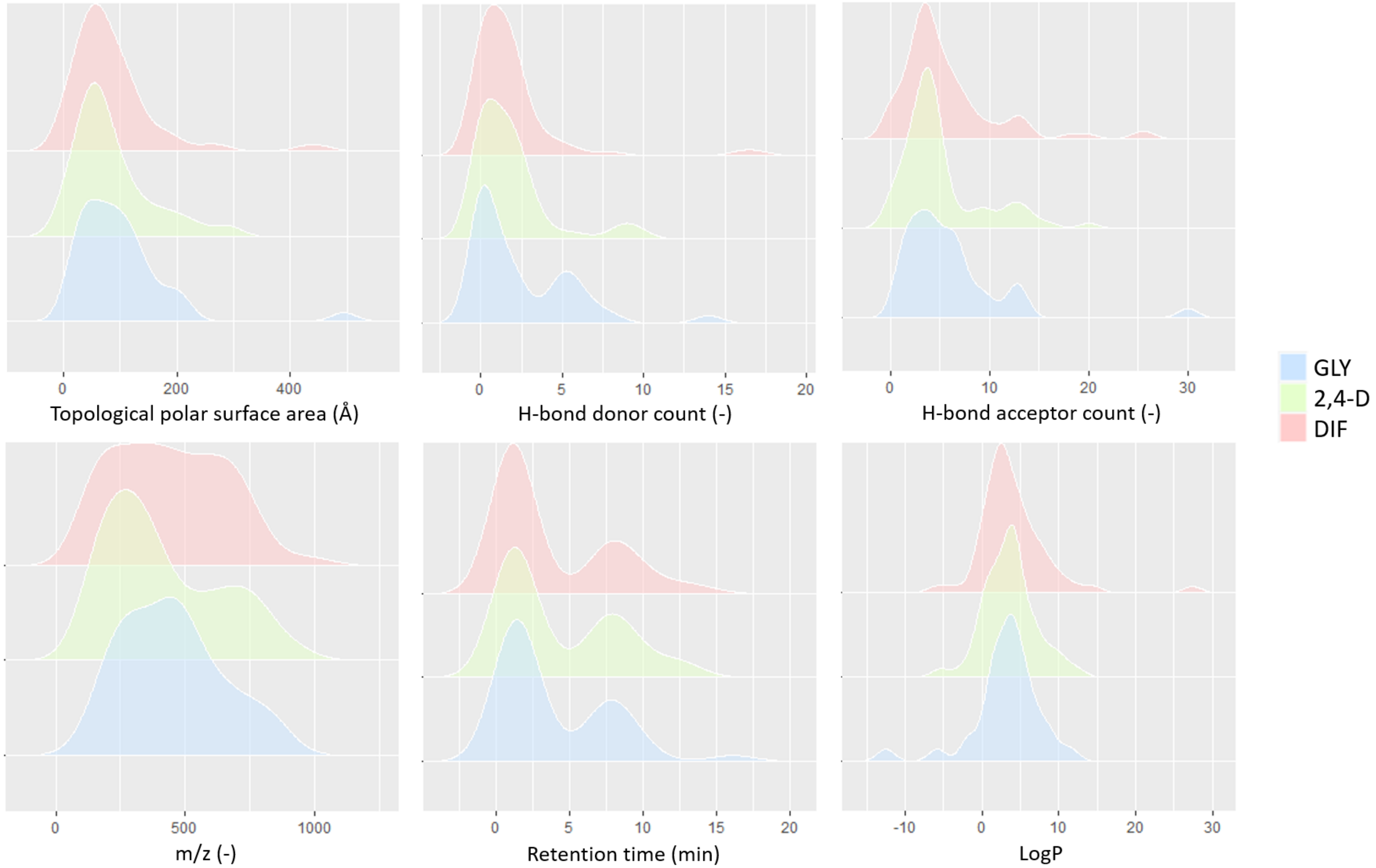


Figure 4: ***Density functions of six molecular descriptors*** *characterising the most significant features in the PLSR models established to predict the adsorption coefficients (*Kd_{ads}*) of glyphosate (blue), 2,.4-D (green) and difenoconazole (red).*

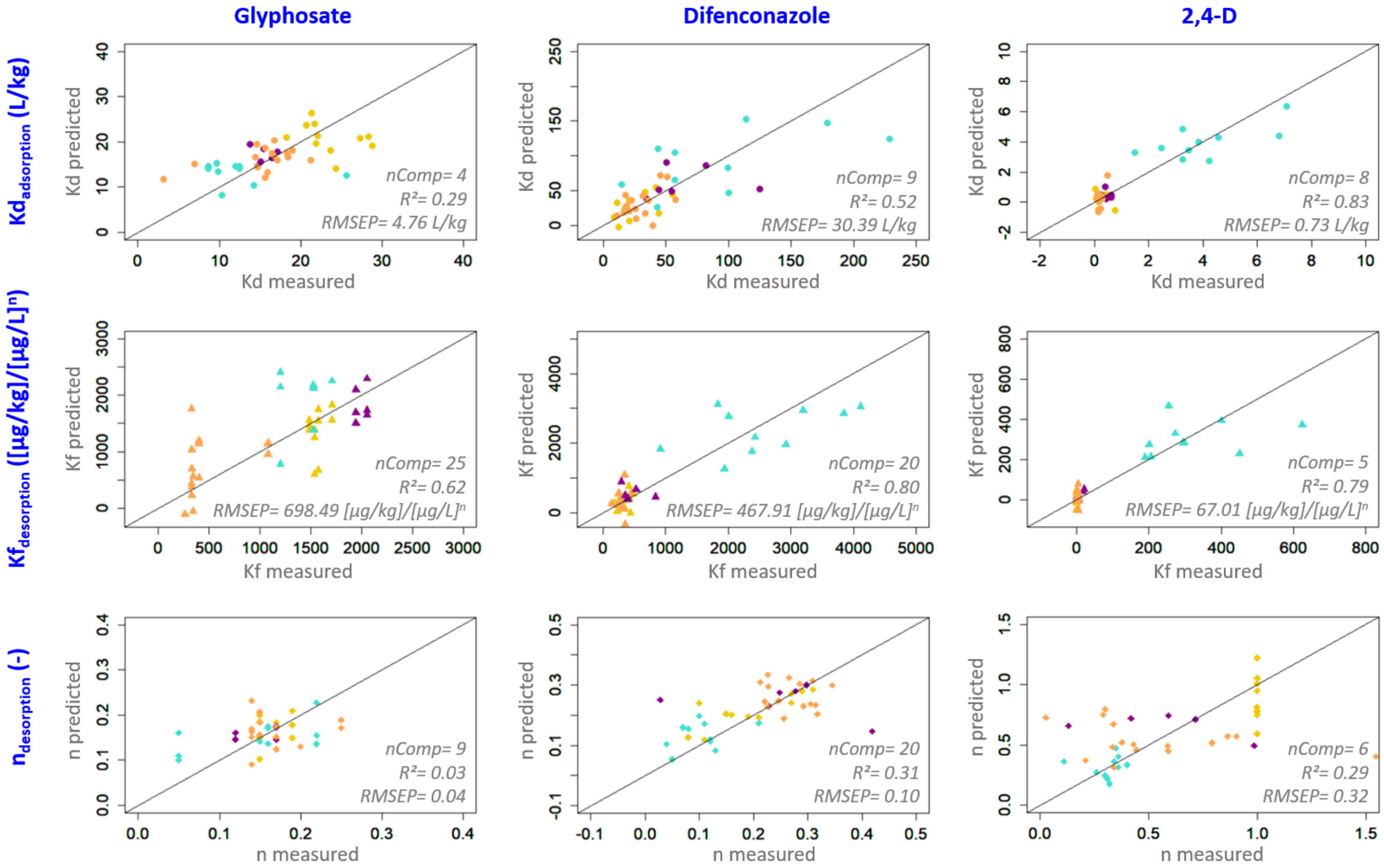


***Figure 5**: **Performance of the PLSR models for the estimation of the sorption coefficients**. The plots represent the predicted vs the measured coefficients and their position relative to the 1:1 line (plain diagonal). The performance criteria (R² & RMSEP) as well as the number of components in the PLSR are indicated for each model.*

Tables

Table 1: Top 50 features identified by One-way ANOVA discriminating most the molecular profiles of soil organic matter

Compound ID	Average m/z	Average RT (min)	Putative compound (formulation)	Putative level	ΔPPM	Adduct type	Correlation with other ions*
f173	164.03572	8.91	**Phthalamic acid** (C8H7NO3)	2*	-	[M+C2H3N+Na-2H]-	NO
f3922	1018.36053	12.98	**Heptadecanoyl CoA** (C38H68N7O17P3S)	3	7	[M-H]-	NO
f283	191.03487	7.96	**6-Methoxy-7-hydroxycoumarin** (C10H8O4)	2*	-	[M+FA-H]-	NO
f136	159.10287	6.48	**2-hydroxy caprylic acid** (C8H16O3)	3	1	[M-H]-	NO
f34	118.03001	6.67	**Furo[3,4-b]pyridine** (C7H5NO)	3	1	[M-H]-	NO
f1045	305.89166	17.12	**2-chloro-4,6-bis(1,1-dichloroethyl)-1,3,5-triazine** (C7H6Cl5N3)	3	4	[M-H]-	NO
f3624	738.77039	0.95	**Phosphoric acid--4-iodophenol (1/3)** (C18H18I3O7P)	3	4	[M-H2O-H]-	NO
f1536	367.01822	5.33	**7-Methyl-8-(methylthio)-1-((phenylsulfonyl)oxy)-3,7-dihydro-1H-purine-2,6-dione** (C13H12N4O5S2)	3	1	[M-H]-	NO
f451	228.96368	0.96	**Diethyl bromomethylphosphonate** (C5H12BrO3P)	3	1	[M-H]-	YES (m/z 702.87756 & 566.90424 & 498.9166)
f992	301.89148	17.11	**4-Bromo-6,6,6-trichloro-3,3-dimethylhex-4-enenitrile**	3	1	[M-H]-	NO

			(C8H9BrCl3N)				
f1331	339.02359	8.03	**Methanesulfonic acid--[1-(2,4-dichlorophenyl)cyclopentyl]methanol (1/1)** (C13H18Cl2O4S)	3	1	[M-H]-	YES (m/z 419.99090)
f677	264.93481	1.01	**{[(Methanesulfonyl)methanesulfinyl]methanesulfinyl}methane-SO-thioperoxol** (C4H10O5S4)	3	4	[M-H]-	NO
f1715	385.96869	3.31	**5'-Hydroxylornoxicam** (C13H10ClN3O5S2)	3	2	[M-H]-	NO
f3687	761.76508	0.97	**Mercury, (2,4-dibromo-6-((p-bromophenyl)carbamoyl)phenoxy)phenyl-** (C19H12Br3HgNO2)	3	5	[M+K-2H]-	YES (m/z 728.77887 & 598.82501 & 793.76123 & 533.84491)
f736	273.02914	2.40	**Simonyellin** (C14H10O6)	2*	-	[M-H]-	NO
f865	288.06662	10.12	**2-Naphthalenol, 8-(4-amino-2-imino-1,3,5-triazin-1(2H)-yl)-, monohydrochloride** (C13H12ClN5O)	3	2	[M-H]-	NO
f180	167.02280	5.25	**Uric acid** (C5H4N4O3)	2*	-	[2M+FA-H]-	NO
f781	277.10953	9.38	**8-(2,3-dihydroxy-3-methylbutyl)-7-methoxychromen-2-one** (C15H18O5)	2*	-	[M-H]-	NO
f942	295.22842	13.11	**(±)9-HODE** (C18H32O3)	3	1	[M-H]-	NO
f974	300.06815	12.42	**1-(azepan-1-yl)-3-(3,4-dichlorophenyl)urea** (C13H17Cl2N3O)	3	1	[M-H]-	NO
f1512	363.04669	12.26	**Quinoxaline, 2-(3-chlorophenyl)-3-[(4-chlorophenyl)methyl]-** (C21H14Cl2N2)	3	1	[M-H]-	NO
f1958	417.10638	3.15	**2-[(E)-(6-Oxocyclohexa-2,4-dien-1-ylidene)methyl]-N-[4-**	3	0	[M-H]-	NO

			(piperidine-1-sulfonyl)phenyl]hydrazine-1-carbothioamide ($C_{19}H_{22}N_4O_3S_2$)				
f1877	404.10434	1.31	**trans-Clovamide** ($C_{18}H_{17}NO_7$)	2*	-	[M+C2H3N+Na-2H]-	NO
f675	264.93469	0.98	**3-Bromo-7-(trifluoromethyl)imidazo[1,2-b][1,2,4]triazine** ($C_6H_2BrF_3N_4$)	3	1	[M-H]-	NO
f2996	566.90424	0.98	**Trisodium 3,3',3''-phosphinetriyltris(benzene-1-sulphonate)** ($C_{18}H_{12}Na_3O_9PS_3$)	3	7	[M-H]-	YES (m/z 430.92770 & 498.91660)
f959	297.24380	13.52	**(9E)-12-hydroxyoctadec-9-enoic acid** ($C_{18}H_{34}O_3$)	2*	-	[M-H]-	NO
f2498	484.87045	1.00	**3,4-Dibutyl-2,5-diiodothiophene** ($C_{12}H_{18}I_2S$)	3	0	[M+K-2H]-	NO
f1493	359.87125	1.18	**Methanesulfonic acid--[1-(2,4-dichlorophenyl)cyclopentyl]methanol (1/1)** ($C_{13}H_{18}Cl_2O_4S$)	3	1	[M-H]-	NO
f1952	416.88364	1.00	**(2,2-Diiodo-1,1-dimethoxyethyl)benzene1,1-diiodo-2,2-dimethoxy-2-phenylethane** ($C_{10}H_{12}I_2O_2$)	3	4	[M-H]-	NO
f1711	385.13504	1.40	**[3-hydroxy-2-methyl-4-(7-oxofuro[3,2-g]chromen-9-yl)oxybutan-2-yl] (Z)-2-methylbut-2-enoate** ($C_{21}H_{22}O_7$)	2*	-	[M-H]-	YES (m/z 371.11993 & 369.14026 & 501.18210)
f2613	501.18167	1.41	**NCGC00179938-02** ($C_{25}H_{28}O_8$)	2*	-	[M-H]-	YES (m/z 369.14026 & 355.12506 & 385.13504)
f3677	759.76746	0.88	**N,N'-(Hexane-1,6-diyl)didocosanamide**	3	4	[M-H]-	NO

			(C50H100N2O2)				
f1826	399.25348	14.52	**Anhydro simvastatin** (C25H36O4)	3	4	[M-H]-	NO
f402	219.96806	1.29	**N-Methyl-2-[(5-sulfanylidene-2,5-dihydro-1,2,4-thiadiazol-3-yl)sulfanyl]acetamide** (C5H7N3OS3)	3	0	[M-H]-	NO
f1562	369.14026	1.42	**NCGC00380677-01!5,7-dihydroxy-2-(4-hydroxy-3-methoxyphenyl)-6-(3-methylbut-2-enyl)-2,3-dihydrochromen-4-one** (C21H22O6)	2*	-	[M-H]-	YES (m/z 355.12506 & 385.13504 & 501.18167)
f3502	695.78937	0.90	-	4	-	-	NO
f3817	843.52698	15.57	**Phosphatidylglyceride 20:3-22:6** (C48H77O10P)	2*	-	[M-H]-	NO
f2924	552.85876	1.00	**10,11-Dibromoundecyl 2,3-dibromobutanoate** (C15H26Br4O2)	3	1	[M-H]-	YES (m/z 620.84650 & 484.87045)
f353	209.93593	11.74	**(2-Bromo-5-fluorophenoxy)acetonitrile** (C8H5BrFNO)	3	2	[M-H2O-H]-	NO
f27	114.38640	2.54	-	4	-	-	NO
f70	135.04543	6.07	**Caffeic acid** (C9H8O4)	2*	-	[M-H]-	NO
f1299	332.92239	1.00	**4-(4-Chlorobenzene-1-sulfonyl)-3-methylthiophene-2-carbonyl chloride** (C12H8Cl2O3S2)	3	0	[M-H]-	NO
f2516	486.03067	1.28	**9-Bromo-3,3-bis(4-fluorophenyl)-3,11-dihydropyrano[3,2-a]carbazole** (C27H16BrF2NO)	3	0	[M-H]-	NO
f1124	313.23865	13.17	**9,10-dihydroxy-12-octadecenoic acid** (C18H34O4)	3	0	[M-H]-	NO
f1959	417.10754	4.43	**1,3-dioxo-2-(oxolan-2-ylmethyl)-N-[4-**	3	1	[M-H]-	NO

			(trifluoromethyl)phenyl]isoindole-5-carboxamide (C21H17F3N2O4)				
f2106	434.87146	1.16	**5,5-bis(2-iodoethyl)pyrimidine-2,4,6(1h,3h,5h)-trione** (C8H10I2N2O3)	3	1	[M-H]-	YES (m/z 304.91415)
f2539	489.12576	10.85	**3-(5-{[1-Butyl-5-cyano-4-methyl-2-(morpholin-4-yl)-6-oxo-1,6-dihydropyridin-3-yl]methylidene}-4-oxo-2-sulfanylidene-1,3-thiazolidin-3-yl)propanoic acid** (C22H26N4O5S2)	3	2	[M-H]-	NO
f83	144.04565	7.69	**8-Hydroxyquinoline** (C9H7NO)	2*	-	[M-H]-	NO
f2404	471.07892	1.45	**Propanamide, N-[2-[(2-bromo-6-cyano-4-nitrophenyl)azo]-5-(diethylamino)phenyl]-** (C20H21BrN6O3)	3	1	[M-H]-	NO
f868	288.93649	1.06	**2-bromo-5-(trifluoro methyl)phenyhydrazine hydrochloride** (C7H7BrClF3N2)	3	1	[M-H]-	YES (m/z 456.86319)
f918	293.17630	11.45	**Embelin** (C17H26O4)	2*	-	[M-H]-	NO

* Compounds annotated with MS-DIAL other annotations were performed with METLIN (see section 2.5)