



# Towards a better understanding of sorption of persistent and mobile contaminants to activated carbon: Applying data analysis techniques with experimental datasets of limited size

Navid Saeidi<sup>a,b</sup>, Laura Lotteraner<sup>c,d</sup>, Gabriel Sigmund<sup>c,e</sup>, Thilo Hofmann<sup>c</sup>,  
Martin Krauss<sup>f</sup>, Katrin Mackenzie<sup>a</sup>, Anett Georgi<sup>a,\*</sup>

<sup>a</sup> Department of Technical Biogeochemistry, Helmholtz Centre for Environmental Research – UFZ, Leipzig 04318, Germany

<sup>b</sup> Department of Microbial Biotechnology, Helmholtz Centre for Environmental Research – UFZ, Leipzig 04318, Germany

<sup>c</sup> Department of Environmental Geosciences, Centre for Microbiology and Environmental Systems Science, University of Vienna, Josef-Holaubek-Platz 2, Vienna 1090, Austria

<sup>d</sup> Visualization and Data Analysis, Faculty of Computer Science, University of Vienna, Sensengasse 6, Vienna 1090, Austria

<sup>e</sup> Environmental Technology, Wageningen University & Research, P.O. Box 17, Wageningen 6700 AA, The Netherlands

<sup>f</sup> Department of Exposure Science, Helmholtz-Centre for Environmental Research – UFZ, Permoserstr. 15, Leipzig 04318, Germany

## ARTICLE INFO

### Keywords:

Functionalized carbon adsorbents

Sorption behavior

Persistent

mobile and toxic substances

Agglomerative clustering algorithms

Distance correlation

## ABSTRACT

The complex sorption mechanisms of carbon adsorbents for the diverse group of persistent, mobile, and potentially toxic contaminants (PMs or PMTs) present significant challenges in understanding and predicting adsorption behavior. While the development of quantitative predictive tools for adsorbent design often relies on extensive training data, there is a notable lack of experimental sorption data for PMs accompanied by detailed sorbent characterization. Rather than focusing on predictive tool development, this study aims to elucidate the underlying mechanisms of sorption by applying data analysis methods to a high-quality dataset. This dataset includes more than 60 isotherms for 22 PM candidates and well-characterized high-surface-area activated carbon (AC) materials. We demonstrate how tools such as distance correlation and clustering can be used effectively to identify the key parameters driving the sorption process. Using these approaches, we found that aromaticity, followed by hydrophobicity, are key sorbate descriptors for sorption, overshadowing steric and charge effects for a given sorbent. Aromatic PMs, although classified as mobile contaminants based on their sorption to soil, are well adsorbed by AC as engineered adsorbent via  $\pi$ - $\pi$  interactions. Non-aromatic and especially anionic compounds show much greater variability in sorption. The influence of ionic strength and natural organic matter on adsorption was considered. Our approach will help in the analysis of solute-sorption systems and in the development of new adsorbents beyond the specific examples presented here. In order to make the approach accessible, the code is freely available and described on GitHub (<https://github.com/Laura-Lotteraner/PM-Sorption>), following the FAIR data principles.

## 1. Introduction

The presence of persistent, mobile, and potentially also toxic organic contaminants (often referred to as PMs and PMTs) in aquatic systems is of increasing concern as they can overcome natural and engineered barriers, such as soil and riverbanks, and enter drinking water, posing a potential risk to human health (Arp et al., 2017; Hale et al., 2022). Quantitative parameters proposed in order to define the persistence and mobility criteria of PMs or PMTs are a freshwater half-life > 40 days

(persistence) and an organic carbon/water partition coefficient  $\log K_{OC} < 4.5$  at pH = 4–10 (mobility) (Arp et al., 2017). One of the most widely used advanced water treatment technologies to eliminate contaminants from waters is adsorptive removal via activated carbon (AC) (Hale et al., 2022; Reemtsma et al., 2016). Given the wide variety of PMs' chemical structures (including heteroatom-rich neutral, ionic, zwitterionic, and non-aromatic or aromatic substances), various sorbate-sorbent surface interactions can contribute as driving forces in addition to generally weaker hydrophobic effects: i) Electrostatic interactions, ii) Ion

\* Corresponding author.

E-mail address: [anett.georgi@ufz.de](mailto:anett.georgi@ufz.de) (A. Georgi).

<https://doi.org/10.1016/j.watres.2024.123032>

Received 22 August 2024; Received in revised form 18 December 2024; Accepted 20 December 2024

Available online 21 December 2024

0043-1354/© 2024 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

exchange, iii) Ion bridging, iv) (Charge-assisted) H-bonding, v) Anion and/or cation charge- $\pi$  bonding, and vi)  $\pi$ - $\pi$  interactions (Kah et al., 2017). Thus, in addition to high surface area, carbon adsorbent surface chemistry is considered to be more relevant to adsorbent performance for PMs than for nonpolar organic contaminants (Kah et al., 2017; Zhou et al., 2021). The complexity of molecular properties of PMs, together with associated adsorption mechanisms, makes the development of quantitative predictive tools required for adsorbent design demanding (Kah et al., 2017; Zhou et al., 2021).

High specific surface area is a well-accepted and widely used sorbent descriptor often related to adsorption performance. Beyond that, carbon adsorbents' charge distribution and surface chemistry can drive these complex interactions. They can be parametrized by: i) Point of zero charge (PZC), anion and cation exchange capacity (AEC and CEC), and content of acidic functional groups (e.g., carboxylic groups) to describe the surface charge state at a specific pH value (Saeidi et al., 2020b; Zhi and Liu, 2015; Zhou et al., 2021), ii) H/C elemental ratio as measure for aromaticity (Aumeier et al., 2023; Sigmund et al., 2020), and iii) Other heteroatom/C ratios such as O/C or (O+N)/C as measures for surface polarity (Sigmund et al., 2020; Zhi and Liu, 2016).

Recently, predictive models have been developed in order to determine relevant adsorption parameters of PM(T)s on carbon adsorbents, such as maximum adsorption capacity ( $q_m$ ) and single-point adsorption coefficient ( $K_d$ ) (Saeidi et al., 2024; Zhang et al., 2020; Zhao et al., 2019) as well as Freundlich isotherm parameters ( $K_F$  and  $n$ ) (Sigmund et al., 2020). In most of these models, molecular sorbate properties, including Abraham descriptors relating to sorbate molar volume, hydrogen bonding affinities and polarizability, as well as charge-specific descriptors, were used to create predictive tools based on either poly-parameter linear free energy relationships (pp-LFERs) (Zhao et al., 2019), neural networks (Sigmund et al., 2020) or their combination (Zhang et al., 2020). Some models also contained sorbent descriptors such as specific surface area and H/C ratio, in order to predict sorption across different carbon sorbents (Sigmund et al., 2020). However, these models typically lack the necessary reliability for use in a generalized setting, particularly when dealing with charged substances (Sigmund et al., 2022; Zhou et al., 2021), as traditional models cannot account for the additional interactions inherent to charged substances. The development of specialized predictive tools for charged substances requires large amounts of training data which can only be obtained by mining large adsorption datasets from the literature. Models trained on such diverse datasets, however, may overlook critical sorbent descriptors as well as experimental conditions due to insufficient reporting of these aspects in the literature, which hinders robust model development (Sigmund et al., 2022; Zhou et al., 2021).

Therefore, in our study we follow an alternative approach to identify key parameters for future model development for PM(T)s adsorption. We demonstrate valuable results of computational methods applied to an inherently limited amount of high-quality data in terms of defined, uniform sorption conditions (e.g., solution chemistry) and comprehensive sorbent characterization.

Our aim is to contribute to the future targeted development of predictive tools for PM adsorption in three ways, i) Presenting a set of experimental data for well-characterized high surface area ACs as carbon adsorbents with different surface chemistries, ranging from high basicity, low acidity, and low oxygen content to the opposite properties. The carbon adsorbents with the mentioned properties were developed by defunctionalization of a pristine microporous high surface area activated carbon felt at high temperature under  $H_2/N_2$  atmosphere and wet oxidation, respectively. ii) Providing relevant insights into drivers and key parameters for PM adsorption including ionizable substances using a combination of exploratory data analysis methods and data visualization. In particular, we combine distance correlation analysis with agglomerative clustering algorithms to analyze the relationships between the molecular properties of substances and their sorption behavior. Distance correlation is used to rank combinations of molecular

properties such as the number of aromatic rings, log  $D_{OW}$  (octanol-water distribution ratio), water solubility, molar volume, molecular weight, and charge according to their relationship with  $K_d$  values at environmentally relevant conditions. These results are further contextualized using agglomerative clustering. Substances are also grouped by their sorption behavior (Freundlich isotherm parameters  $K_F$  and  $n$ ) using agglomerative clustering. One large advantage of the agglomerative approach compared to other approaches, especially for small data sets, is the visual interpretability of its results (Aggarwal, 2014). iii) We share all our data and Python code via github, thus allowing other researchers to build on and extend our work and providing an example of best practice for future work.

We determined Freundlich adsorption parameters of 22 environmentally relevant PM candidate substances under controlled experimental conditions on three well-characterized AC materials. They were selected to cover a range of structural properties based on their low  $K_{OC}$  values (based on EPI Suite™ (EPA, 2024)) and their frequent detection in wastewater treatment plant effluents and rivers, indicating resistance to biodegradation. The 22 PM candidates consist of anionic aromatic (6 substances), anionic non-aromatic (3), cationic aromatic (5), cationic non-aromatic (3), zwitterionic aromatic (1), zwitterionic non-aromatic (1), neutral aromatic (2), and neutral non-aromatic (1) substances. To account for complicating factors, we additionally investigated effects of natural organic matter (NOM) and competitive inorganic ions on adsorption of PMs.

## 2. Materials and methods

### 2.1. Materials

The PMs used for adsorption batch experiments are 1,3-diphenylguanidine (1,3-DPG), 1H-benzotriazole (1H-BTZ), 2-benzothiazolesulfonic acid (2-BTSA), 2-naphthalenesulfonic acid (2-NSA), 4-hydroxyquinoline (4-HQ), 4-nitrophenol (4-NP), 7-(ethylamino)-4-methylcoumarin (7EM), acesulfame (Ace), carbendazim (CBZ), ciprofloxacin (CPFX), cyclohexylamine (CHA), diethylphthalate (DEP), gabapentin (GPN), mecoprop (MCP), N-guanyurea (GUA), perfluorobutanoic acid (PFBA), perfluorobutanesulfonic acid (PFBS), pipamperone (Pipa), sulfamethoxazole (SMX), tetraethylammonium (TEA), tramadol (Tram), and triethylphosphate (TEP). Details on the chemical standards used can be found in Table 1WS in Supporting Information (provided MS Word file, hereafter referred to as SI). Molecular properties of the PM substances can be found in Table 2ESa (in the provided MS Excel file, hereafter referred to as xls file). Suwannee River natural organic matter (SRNOM, reference number: 2R101 N), which is a well-characterized reference material (Driver and Perdue, 2014) of the International Humic Substances Society (IHSS), was used as NOM.

Viscose-based activated carbon felt (ACF) with a felt thickness of 3 mm was purchased from Qinhuangdao Zichuan Carbon Fiber Co., Ltd. The ACF was additionally modified to produce materials with different surface chemistries: In order to obtain oxygen-functionalized ACF (OXACF), the ACF (1 g) was oxidized in 5 N nitric acid (100 mL) for 6 h at 90 °C. Defunctionalized ACF (DeACF) was prepared by treating the ACF under a  $N_2/H_2$  atmosphere at 900 °C. The thermal treatment in the reducing atmosphere leads to removal of carboxylic and other oxygen-containing functional groups from AC (Saeidi et al., 2020b; Zhou et al., 2021). Prior to use, all three carbon adsorbents were washed first with methanol and then rinsed 5 times with deionized water. Finally, all carbon adsorbents were dried at 80 °C in air and kept in a desiccator for further use. We would like to refer interested readers to previous studies for more details on the surface modifications (Saeidi et al., 2020b; 2021).

## 2.2. Methods

### 2.2.1. Textural and chemical surface characterization

Specific surface area, pore volume, mean pore size, pore size distribution, and fiber diameter of the carbon adsorbents were determined following procedures reported in detail in the SI. Surface chemical properties of the carbon adsorbents were evaluated by measuring PZC (Saeidi et al., 2020a), the concentration of charge-compensating sites (mole of sites/gram of carbon adsorbent) for anions and cations at pH 7 yielding AEC and CEC at pH 7, respectively (Hsieh and Pignatello, 2017; Saeidi et al., 2020a), and temperature-programmed decomposition (TPD). A full description of the procedures can be found in the SI.

### 2.2.2. Analytical methods

Compounds were assessed via LC-MS/MS, utilizing an Agilent 1260 LC system coupled to a Sciex QTrap 6500 mass spectrometer. Positive and negative ion modes were run separately. For a comprehensive description of the analytical methodology, refer to the SI.

### 2.2.3. Adsorption experiments

Batch adsorption experiments with the targeted PMs were carried out in 25, 50, and 100 mL vessels containing tap water and certain quantities of adsorbents. Table 3WS in the SI lists the characteristics of the tap water, including its total organic carbon content (TOC). 100 µg/L of each of the PMs was then spiked from an aqueous stock solution. The adsorption experiments in presence of NOM were performed by additionally spiking 10 mg/L SRNOM from a stock solution prepared in tap water (pH 7). This concentration was selected based on the mean concentration of dissolved organic carbon (7.7 mg/L) in surface water resources in Europe (<http://weppi.gtk.fi/publ/foregsatlas/article.php?id=15>). Adsorption of the PMs in presence of competitive inorganic ions was studied by adding 8 mM CaSO<sub>4</sub>·2H<sub>2</sub>O, increasing the concentrations of Ca<sup>2+</sup> in tap water to 10 mM and SO<sub>4</sub><sup>2-</sup> to 9.3 mM in total. A full description of adsorption experiments can be found in the SI.

### 2.2.4. Data analysis

**2.2.4.1. Data preparation.** The quantity of PMs adsorbed by the carbon adsorbents (adsorbate loading) was calculated from Eq. (1):

$$q_e = \frac{V \times (C_0 - C_e)}{m} \quad (1)$$

where  $q_e$  (µg/g) denotes loading of adsorbate on the carbon adsorbent at equilibrium. The initial adsorbate concentration and its concentration in solution at equilibrium are termed  $C_0$  and  $C_e$  (µg/L), respectively.  $V$  (L) is the volume of solution and  $m$  (g) is the mass of the carbon adsorbent. Equilibrium was considered to be approached after 2 d (no significant changes in  $C_e$  observed ( $\leq 5\%$ ) at further prolonged contact times). Single-point  $K_d$  values were calculated as  $K_d = q_e/C_e$ . Unit reported for log  $K$  values pertains to the original, unaltered variable.

Freundlich fitting parameters from a linearized form of the classical Freundlich equation (Eq. (2)) (Freundlich, 1907) were determined based on a minimum of 5 concentration data points per substance and sorbent.

$$\log q_e = n \times \log C_e + \log K_F \quad (2)$$

where  $K_F$  ((µg/g)/(µg/L) <sup>$n$</sup> ) is the Freundlich constant related to the affinity between carbon adsorbent and adsorbate, and  $n$  (dimensionless) is the Freundlich exponent. Substances with  $n < 0.4$  and  $K_F$  larger than the 95 th quantile for at least one of the sorbents were excluded from the entire data set, resulting in 18 substances remaining for clustering and distance correlation analysis, where only molecular properties with a pairwise correlation coefficient below 0.75 were used to minimize co-correlation. Both molecular properties and Freundlich parameters were normalized to a value range in [0,1]. A more detailed description can be found in the SI.

**2.2.4.2. Data analysis.** To identify the property combinations most strongly related to sorption, property combinations were ranked by their distance correlation (Székely et al., 2007) with  $K_d$  calculated at  $C_e = 1$  µg/L (see Table 6ES in the xls file). The top combinations of two and three properties for each sorbent were then used in an agglomerative clustering approach (Aggarwal, 2014) and the resulting dendrograms were plotted. Find more details in the SI.

To compare sorption behavior across the three sorbents, substances were additionally clustered according to their Freundlich parameters (see Table 4ES in the xls file) and the results were then compared using a combination of Venn diagrams.

As the clustering algorithm used will partition the data into the specified number of clusters regardless of the natural structure, multiple clustering configurations (two, three, and four clusters) were tested for each sorbent. The clustering configuration that provided the most meaningful and interpretable grouping of the data was identified for each sorbent by visually comparing plots of the data points, colored by the different clustering results, and selected for further analysis.

All computations were done in python (Python version 3.8). The developed code is freely available at <https://github.com/Laura-Lottner/PM-Sorption>.

## 3. Results and discussion

### 3.1. Adsorption isotherms

In total more than 60 isotherms were generated across a wide range of environmentally relevant aqueous phase compound concentrations from the sub-nanogram to µg/L range for the three ACF types which are summarized in Fig. 1. Isotherm parameters obtained by fitting experimental data with the Freundlich equation are given in Table 1ES in the xls file.  $R^2$  values of 0.97 or higher were obtained in all cases.

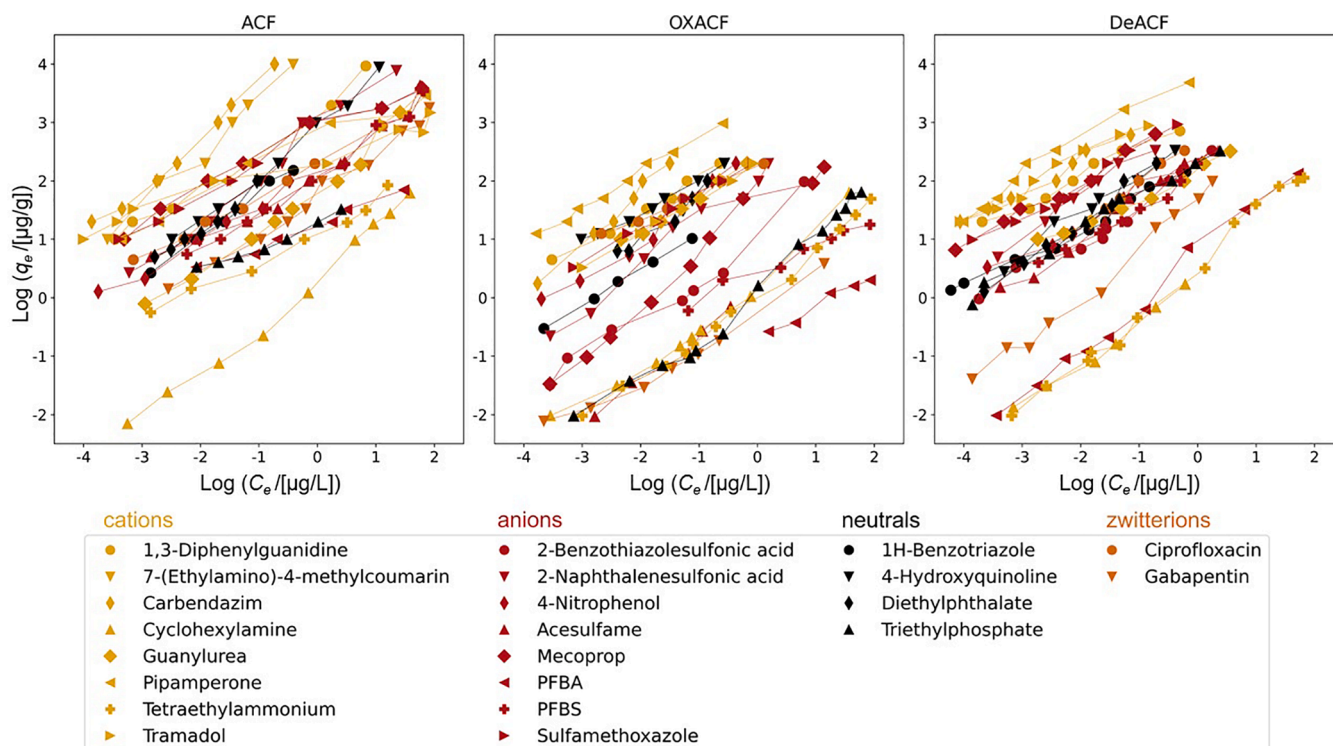
Fig. 1 provides an initial visual impression of the significant variability in adsorption affinities of the individual PMs for activated carbon (AC). The adsorption loading ( $q_e$ ) at  $C_e=100$  ng/L shows a difference of up to four orders of magnitude for one and the same PM compound. Further variability is contributed by the influence of the AC type. This diverse behavior complicates the prediction of PM adsorption and the selection of appropriate carbon adsorbents, highlighting the necessity of further investigation, which is performed in the following sections.

### 3.2. Impact of adsorbent surface modification on adsorption of PMs

#### 3.2.1. What are the variations in PMs affinities to carbon adsorbents?

In Fig. 2,  $K_F$  as the adsorption affinity parameter was plotted together with the nonlinearity proxy  $n$  for all compounds and AC types. In the three rows, different compound properties in terms of charge state, number of aromatic rings, and planarity of the molecules were highlighted by different colours. A visual inspection shows that there is no individual property that leads to a strong clustering of  $K_F$  or  $n$  values in a high or low range. There was no clear preference for neutral, cationic or anionic substances on either of the carbon adsorbents, despite their different surface net charges. Obviously, a combination of sorption drivers needs to be considered for describing PM sorption as will be explained in Sections 3.3 and 3.4.

The strongest agglomeration of  $K_F$  and  $n$  values was observed for DeACF, with a tendency towards the right upper side of the plot: 17 out of 21 substances showed high sorption affinity with  $K_F > 10^2$  ((µg/g)/(µg/L) <sup>$n$</sup> ) and  $n$  values were in all cases  $\geq 0.4$ . DeACF had the highest AEC (0.26 mmol/g vs. 0.074 (ACF) and 0.035 (OXACF)) and PZC (9.5 vs. 6.3 and 2.7) (see Table 4WS in the SI) meaning that it has a net positive charge at near neutral pH. In contrast, OXACF with a net negative charge showed the widest distribution in  $K_F$  values, with only 11 substances having  $K_F > 10^2$  ((µg/g)/(µg/L) <sup>$n$</sup> ) and 11 showing considerably lower  $K_F$ . ACF, with intermediate surface chemistries among the carbon adsorbents, also took an intermediate position in adsorption.



**Fig. 1.** Isotherm data for adsorption of PMs on ACF, OXACF, and DeACF, respectively (from left to right). Lines were added as guide to the eye. Isotherm parameters obtained by fitting to the Freundlich equation are shown in Table 1ES in the xls file. The experiments were performed in tap water at pH  $6.8 \pm 0.3$ . Detailed experimental procedures are given in the SI Section 3.

Since the carbon adsorbents had similar specific surface areas and pore size distributions (see Table 5WS and Fig. 1WS in the SI), the changes in their surface chemistries resulting from the surface modification are clearly responsible for the significantly different adsorption behavior of the PMs. This emphasizes the key role of AC surface chemistry for effective sorptive removal of PMs.

Fig. 3 shows the relationship between  $\log K_d$  (calculated at  $C_e = 1 \mu\text{g/L}$  using Eq. (2)) and  $\log K_{OC}$  values (based on EPI Suite™ (EPA, 2024), which are parameters for retardation by AC as technical adsorbent and soil organic matter as natural sorbent, respectively. In most cases,  $K_d$  values for sorption to AC are higher than  $K_{OC}$  values by  $> 1$  order of magnitude (a few cases with lower differences only for OXACF). Fig. 3 shows that substances with higher mobility (i.e., lower  $K_{OC}$  values) are highly sensitive to the type of AC used, with differences in  $K_d$  (at  $C_e = 1 \mu\text{g/L}$ ) for the same substance across various ACs exceeding three orders of magnitude. Oxygen-rich OXACF generally exhibits the lowest sorption affinity for most substances tested (down to about 200 L/kg), whereas the lowest  $K_d$  values for ACF and DeACF are still greater than 2000 L/kg.

The high variability in AC adsorption coefficients for substances with similar  $\log K_{OC}$  translates for a typical AC unit into anticipated breakthrough times (right y-axis in Fig. 3) from less than a day to hundreds of years, depending on the specific AC and substance properties. This further emphasizes the need for a better understanding and control of adsorption processes for PMs, especially for those considered to be very mobile in the environment.

### 3.2.2. Can matrix effects in water compensate for structure-related differences in sorption affinities?

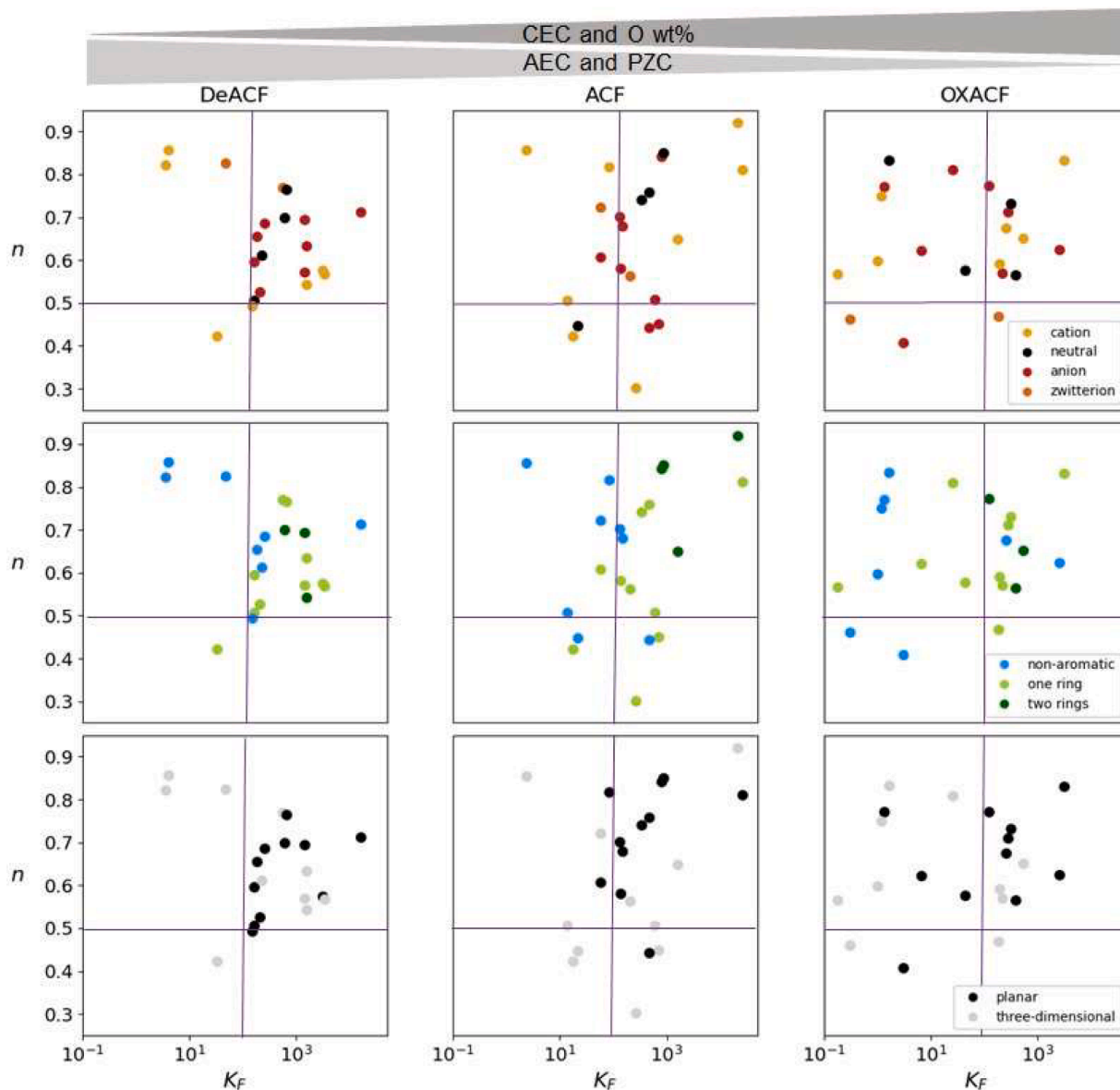
Overall, water matrix effects such as ionic strength (IS) and competition from NOM modulate adsorption do not compensate for the large differences between the various PMs and the various carbon adsorbents (Fig. 4). OXACF with its higher O-content (9.2 wt% vs. 3.9 wt% (ACF) and 1.4 wt% (DeACF)) and negative-charge density (CEC: 0.19 mmol/g

vs. 0.088 (ACF) and 0.014 (DeACF)) showed lower  $K_d$  values for most of the substances than the original ACF. The most pronounced adverse effects of surface oxidation reaching up to three orders of magnitude decline in  $K_d$  were observed for the anionic non-aromatic substances (ACE, PFBS and PFBA, see Fig. 4). This is most likely due to electrostatic repulsion. Such repulsion is slightly weakened by increased IS, resulting in slightly enhanced adsorption of anionic non-aromatic substances on OXACF ( $\text{CEC} \gg \text{AEC}$ ) and ACF ( $\text{CEC} > \text{AEC}$ ), with a stronger effect for OXACF (up to +0.7 log units or a factor ( $f$ ) = 5 in  $K_d$ ).

In the case of DeACF ( $\text{CEC} \ll \text{AEC}$  and net positive charge at neutral pH), increasing IS had the opposite effect, with a decline up to  $f = 6$  in  $K_d$  in the adsorption of anionic non-aromatic substances due to shielding of attractive interactions and/or competitive adsorption of  $\text{SO}_4^{2-}$  by positively charged surface sites. IS also had a moderate impact (up to  $f = 5$  in  $K_d$ ) on neutral and aromatic anionic substances, aligning with reduced electrostatic interactions. It slightly decreased (up to  $f = 3$  in  $K_d$ ) the adsorption of PM cations on all carbon adsorbents except DeACF, where it had no or only modest enhancing effects, altering repulsive interactions.

Our data indicate that DeACF, with its net positively charged surface, is not a poorer adsorbent for cationic substances than the two other carbon adsorbents, neither for aromatic nor for non-aromatic substances. The positive surface charge is attributed to proton adsorption on highly delocalized  $\pi$ -electron systems, which are preferentially present in carbon materials with low O-content (Boehm, 1994). Since these systems exhibit reversible proton adsorption by nature, such positive sites are much more flexible when interacting with charged organic molecules than negatively charged sites such as carboxylates and other deprotonated O-containing groups. This is due to the fact that surface-adsorbed  $\text{H}^+$  can attract organic anions, while it can also be exchanged with organic cations, allowing their direct interaction with the  $\pi$ -electron system (Kah et al., 2017; Qu et al., 2008). We hypothesize that these dual-mode interaction sites make carbon adsorbents with net positive surface charge more versatile adsorbents for a wider range of





**Fig. 2.** Freundlich parameters  $K_F$  ( $(\mu\text{g/g})/(\mu\text{g/L})^n$ ) and nonlinearity proxy  $n$  (dimensionless) plotted against each other (log scale for  $K_F$ ) for each type of sorbent, with colors and symbols indicating the substances' charge, aromaticity, and planarity. Isotherm parameters are listed in Table 1ES in the xls file. Lines (at  $n = 0.5$  and  $K_F = 10^2$ ) were added to facilitate comparison between the plots.

PMs.

In contrast, we could not observe an increased adsorption of the cationic substances to OXACF, which indicates that no increase in electrostatic attractive interactions with the more densely present carboxylate groups in OXACF was realized, possibly because enough negatively charged sites already existed on ACF or because this modification partly compromised interactions in other ways (higher surface polarity favors competition by water molecules in adsorption) (Zhou et al., 2021). Overall, anionic non-aromatic substances were substantially more susceptible to the surface charge of the carbon than cationic substances.

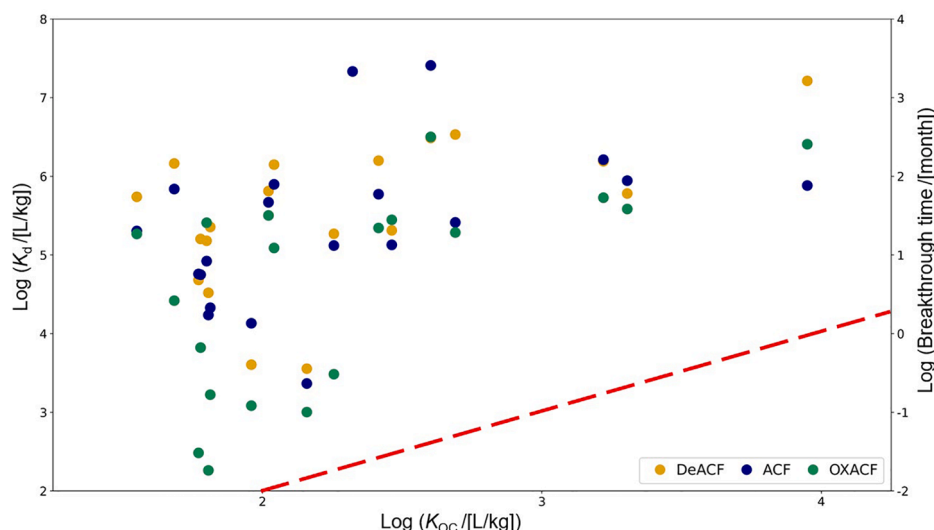
At 10 mg/L NOM, most substances showed a slight reduction in adsorption on all adsorbents (Fig. 4). The adverse effect was more pronounced for substances with higher hydrophobicity and lower mobility, such as 4-HQ, DEP, SMX, 7EM, CBZ, Tram, 1,3-DPG, and Pipa. The strongest adverse effect on adsorption (up to  $f = 10$  in  $K_d$ ) was observed for CBZ and 7EM, having very high  $K_{OC}$  (around 2800 and 6000, respectively) and comparable  $pK_a$ : 4.28 (CBZ) and 4.27 (7EM). All of

these substances are also aromatic.

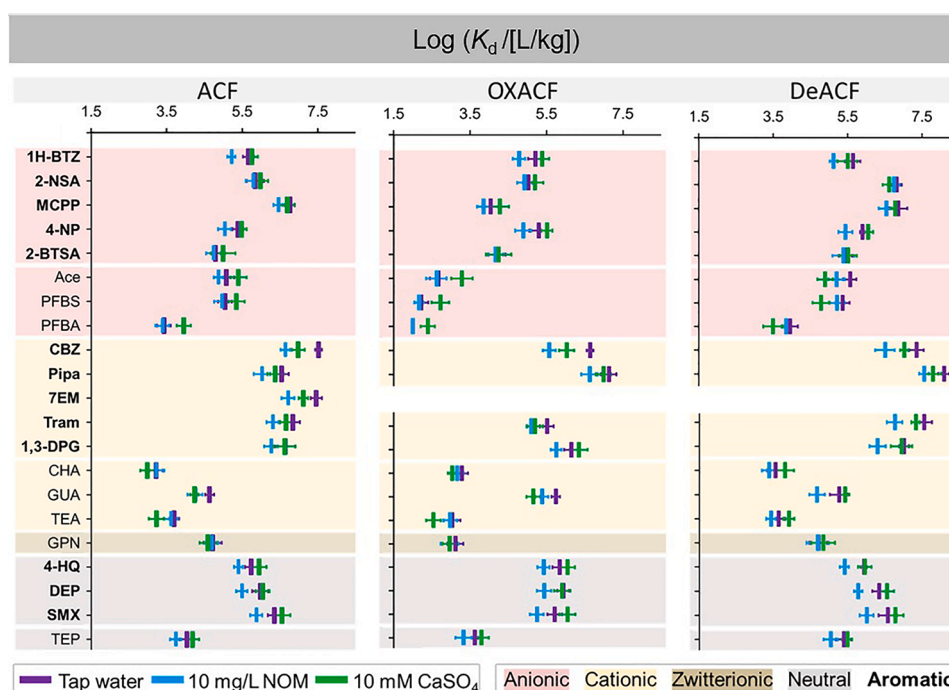
No anionic substance was impacted significantly by NOM. NOM consists of refractory macromolecules with aromatic and aliphatic regions bearing acidic functional groups such as carboxylic acid groups, thus carrying negative charges at neutral pH (Driver and Perdue, 2014).

Despite NOM's ability to sorb contaminants through hydrophobic effects and electrostatic attraction, it is unlikely to compete as sorbent effectively in our experiments due to its significantly lower concentration on a carbon basis. Therefore, the mild adverse impact of NOM on adsorption is likely due to competition for highly favorable sorption sites on the carbon adsorbents.

DeACF, with a net positive surface charge, outperforms OXACF (net negative) and ACF (intermediate charge, PZC 6.3), particularly for anionic non-aromatic substances (see Fig. 2WS in the SI), even in the presence of NOM. The microporous nature of the carbon adsorbents reduces competitive effects of NOM (Saeidi et al., 2020a; Zhou et al., 2021).



**Fig. 3.** Relationship between  $K_{OC}$  values (EPI Suite™ from Table 2ESa in the xls file), single-point AC adsorption coefficient ( $K_d$ , calculated for  $C_e = 1 \mu\text{g/L}$ ) of 22 PMs on various AC types in tap water, and breakthrough time for a hypothetical adsorbent unit (calculated by Eqs. 1WS to 3WS in the SI with approximated  $m_{\text{solid}}/V_{\text{void}} = 0.67 \text{ kg/L}$  and water residence time ( $t_w$ ) = 15 min). The red dashed line represents the  $y = x$  line, indicating where  $\log K_d$  would equal  $\log K_{OC}$ .



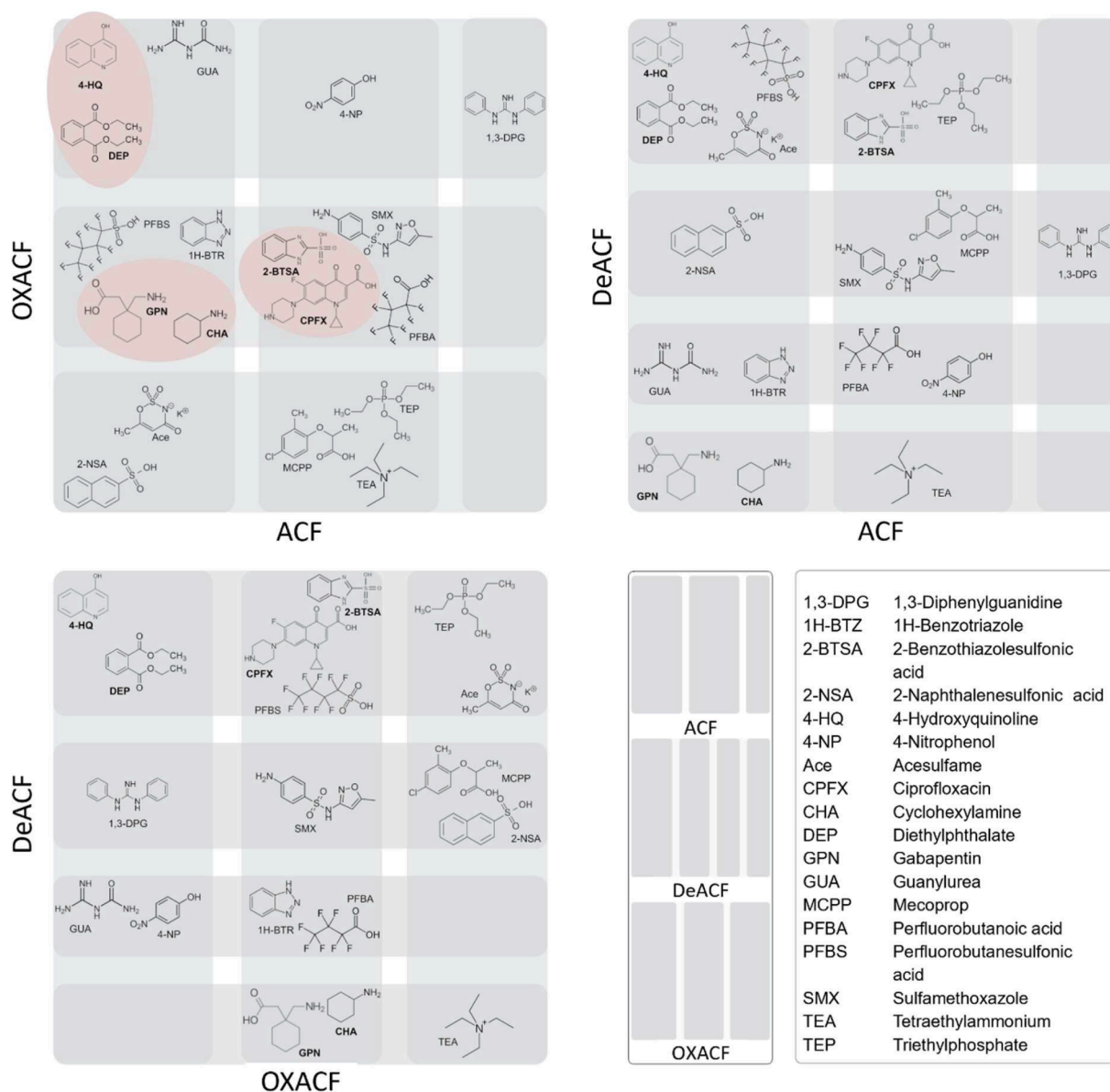
**Fig. 4.** Adsorption of PMs ( $C_0 = 100 \mu\text{g/L}$  each) in tap water and tap water with raised IS (about  $10 \text{ mM CaSO}_4$ ) or addition of  $10 \text{ mg/L NOM}$  on ACF, OXACF, and DeACF ( $1 \text{ g/L}$ ,  $\text{pH} = 7 \pm 0.3$ ). Error bars are deviation from a mean value of at least two experiments (for tap water of 3 experiments).  $K_d$  for PFBA adsorption on OXACF in tap water and with NOM are upper limit of the true  $K_d$  (insignificant depletion of the analytes  $C_e/C_0 \geq 0.85$ ). For CPFX and 7EM (except of ACF) data are not available.

### 3.3. What are the adsorption similarities across PMs?

We use a columnar Venn diagram (Fig. 5) in which 3 to 4 grouping columns from the substance interaction at each carbon adsorbent are vertically related to the grouping columns of one of the other adsorbent types. This results in overlapping clusters of substances for the two adsorbents compared, each representing a set of similar substances. As can be seen in Fig. 5, 3 pairs of the substances always clustered together, that is, they showed similar adsorption behavior, regardless of ACF type: 1) 4-HQ and DEP as two neutral aromatic substances, 2) 2-BTSA (anionic

aromatic substance) and CPFX (zwitterionic aromatic substance), and 3) CHA (non-aromatic cation) and GPN (non-aromatic zwitterionic substance). The substances in cluster 1 are both neutral and aromatic. There is another substance with these properties (1H-BTZ) which did not fall into this cluster. 4-HQ and DEP are more hydrophobic than 1H-BTZ (molar volumes:  $1.10$  and  $1.71$  vs.  $0.86 \text{ cm}^3/\text{mol}/100$  and water solubility:  $0.030$  and  $0.0047$  vs.  $0.17 \text{ mol/L}$ ).

2-BTSA and CPFX in cluster 2 both have 2 H-bond acceptor sites and aromatic rings. CHA and GPN (cluster 3) have comparable  $\text{pK}_a$  ( $10.45$  and  $9.41$ , respectively) due to primary amine groups. The substance



**Fig. 5.** Clustering of substances according to their  $K_F$  and  $n$  values visualized as Venn diagrams for all three sorbent combinations. Substances that cluster together for all three sorbents are printed bold in all panels and highlighted in red in the first. Substances that appear in the same of the three or four separate grouping columns for each carbon adsorbent show high similarity in  $K_F$  and  $n$  values for this type of carbon adsorbent. Substances that appear at crossings of the columns resulting from pairwise plots of the individual carbon adsorbents show high similarity in  $K_F$  and  $n$  for both of the carbon adsorbent types. The difference between the grouping columns is not necessarily the same.

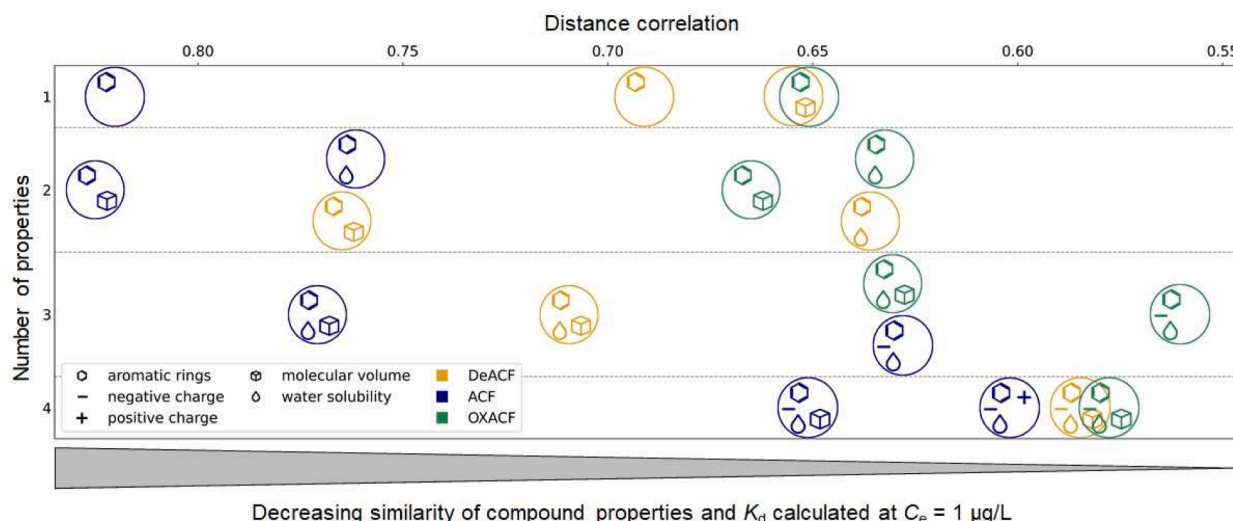
pairs of clusters 1 and 3 identified in Fig. 5 (4-HQ and DEP, CHA and GPN) were similarly influenced by increased IS and presence of NOM. This evaluation can unfortunately not be extended to cluster 2 due to the absence of CFPX data. Non-aromatic anionic substances never clustered, neither with each other nor with other substances. This means that their adsorption behavior is strongly dependent on the type of functionalization of the carbon adsorbents, which agrees with conclusions derived from evaluation of experimental data in Figs. 1 to 4. Furthermore, 1,3-DPG never clustered with any other substance in any of the carbon adsorbent combinations. This substance had the highest  $K_F$  on ACF (1627 ( $\mu\text{g/g}$ )/( $\mu\text{g/L}$ )<sup>*n*</sup> with  $n = 0.65$ ) of all substances considered, which was maintained for DeACF (1554 with  $n = 0.54$ ) but strongly declined for OXACF (535 with  $n = 0.65$ ), a behavior that was unique to this substance, which is the only cationic substance with two aromatic rings and a low water solubility (0.0048 mol/L). This therefore indicates the contribution of a hydrophobic effect in adsorption in addition to  $\pi$ - $\pi$ ,

cation- $\pi$ , and H-bonding interactions.

Overall, the DeACF-**OXACF** combination resulted in the lowest number of clusters of > 2 substances (1), followed by DeACF-ACF (3) and OXACF-ACF (4). Consequently, when the surface chemistry of the adsorbent combinations is more alike, then the number of substances that behave similarly across the adsorbents based on their Freundlich parameters increases.

### 3.4. What are the main drivers for the sorption behavior of PMs for a given sorbent?

Molar volume and aromaticity were confirmed as key sorption drivers in the similarity analysis (Fig. 6), with aromaticity being the property most strongly related to  $K_d$ . When comparing the similarity of more than one parameter with sorption, then molar volume or water solubility (as a proxy of hydrophobicity) in combination with



**Fig. 6.** Similarity between molecular properties and sorption affinity (expressed as  $K_d$  at  $C_e = 1 \mu\text{g/L}$ ). Similarity was computed for the three sorbents separately (colors), and for all combinations of molecular properties (symbols). Combinations of four or fewer molecular properties with similarity  $\geq 0.55$  are displayed. Full data in [Table 6ES](#) (the xls file).

aromaticity showed a high similarity.

The charge of the substance only came into play as the third driver in the substance descriptor-based analysis. Thus, for the group of PMs investigated here, sorbate charge appeared to make a relatively small overall contribution to the predictability of sorption behavior. As a consequence, for the parameterization of predictive models for future larger data sets for sorption of PMs with high-surface-area carbon materials, the number of aromatic rings (or a more precise speciation-dependent parameter for  $\pi$ -electron accepting and donating capacity), coupled with hydrophobicity-associated parameters such as water solubility or  $D_{OW}$ , and steric parameters such as molar volume should be considered in preference to substance-charge-associated parameters, which appear to contribute less to the overall sorption tendency for the majority of PMs for a given material.

Dendrograms are shown in Fig. 3WS in the [SI](#) for a better understanding of the clustering of the substances based on their molecular properties.

#### 4. Conclusions

The aim of this study was to identify important input parameters for model development for persistent, mobile, and potentially also toxic organic contaminants (PMs or PMTs) adsorption on carbon adsorbents such as high-surface-area activated carbon (AC). So far, the development of quantitative predictive tools for adsorbent design and performance prediction is challenged by the complex nature of adsorption of PMs on carbon adsorbents and limited experimental sorption data coupled with detailed sorbent characterization. The challenges are further compounded when dealing with charged substances as existing models tailored for neutral substances prove inadequate. Thus, by combining distance correlation and agglomeration clustering to an intermediate amount of high-quality sorption data, we further elucidated mechanisms and key drivers in adsorption of PMs, paving the way for the development of more accurate quantitative prediction models. We analyzed the relationships between the molecular properties of 18 PMs and their sorption to functionalized microporous high-surface-area AC by distance correlation and clustering. In total, we presented a set of 64 adsorption isotherms and sorbent characterization data for three well-characterized adsorbents with different surface chemistries, with widely varying basicity, acidity, and oxygen content. In general,  $K_d$  values for sorption to AC appeared to be higher and not well correlated to  $K_{OC}$  values for the tested compounds and carbon types.

The experimental data revealed that substances with at least one aromatic ring had higher affinities ( $K_F$  and  $K_d$ ) than non-aromatic substances on all three tested carbon adsorbents. Adsorption of the most hydrophobic substances was most influenced by NOM, which slightly suppressed adsorption, presumably due to competition for sorption sites. The charge of the substances appeared to play only a secondary role in the predictability of sorption within our data set, suggesting that sorption was driven by interactions with the aromatic rings of the substances and that electrostatic interactions could not overcome this dominating proportion as a driver of the overall sorption in a given adsorbent.

When the three adsorbents were analyzed individually, similarity analysis based on  $K_d$  values as a measure of adsorption confirmed that molar volume and aromaticity of the substances are the key drivers of adsorption for all three materials. Aromaticity appears to have the greatest effect on sorption.

When a second parameter was also included, molar volume or water solubility (as a proxy of hydrophobicity) paired with aromaticity showed a high similarity with sorption affinities. The charge of the substance only came into play as a third driver in the substance descriptor-based analysis.

When comparing adsorbents with different surface chemistries, the charge of the PM substance was a decisive factor in explaining the differences in sorption behavior. The adsorption behavior of the non-aromatic anionic organic PMs tested (PFBA, PFBS, and acesulfame) was strongly dependent on the AC type. Electrostatic attraction and repulsion were found to play a key role in their adsorption and the O-rich AC with negative net charge discriminated these anions most strongly. In contrast, cationic non-aromatic substances were not as sensitive to the net charge state of the carbon adsorbent. The defunctionalized carbon adsorbent with low O content and positive net charge is a good adsorbent for both anionic and cationic non-aromatic substances, and thus showed the best overall adsorption performance for all types of PMs.

Therefore, when parameterizing predictive models for larger data sets for sorption of PMs using high-surface-area carbon adsorbents, it is recommended to consider aromaticity of the sorbate, speciation-dependent determination of  $\pi$ -electron acceptor and donor capacity, parameters associated with hydrophobicity (e.g., water solubility or  $D_{ow}$ ), and steric parameters such as molar volume, which for most substances are more important than charge in describing sorption for a given adsorbent. However, charge cannot be neglected when describing the overall sorption behavior, especially when comparing adsorption at



adsorbents with different surface chemistries with a focus on non-aromatic anionic PMs such as the typical PFAS anions.

Furthermore, our experimental results show that  $K_{OC}$ , used as a mobility criterion, is not a good predictor for the retention of PMs by high-surface-area carbon adsorbents. This calls for the development of suitable predictive approaches for the adsorption of PMs in engineered systems; our work contributes to the parameterization and further development of such approaches. This study serves as a proof-of-concept, demonstrating both the potential of clustering approaches in this field and the valuable insights that small datasets can reveal when analyzed with the right methods. Future studies applying more advanced data analysis and prediction tools can build on both the results and the approach presented here.

In order to optimize the use of our data, we adhered to the FAIR data principles and provided both the clustering algorithm code (on Github: <https://github.com/Laura-Lotteraner/PM-Sorption>) and the raw data (an xls file is included).

### CRedit authorship contribution statement

**Navid Saeidi:** Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Laura Lotteraner:** Writing – review & editing, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Gabriel Sigmund:** Writing – review & editing, Validation, Supervision, Methodology, Investigation, Data curation, Conceptualization. **Thilo Hofmann:** Writing – review & editing, Validation, Supervision. **Martin Krauss:** Writing – review & editing, Validation, Investigation, Conceptualization. **Katrin Mackenzie:** Writing – review & editing, Supervision, Investigation, Conceptualization. **Anett Georgi:** Writing – review & editing, Visualization, Validation, Supervision, Methodology, Investigation, Data curation, Conceptualization.

### Declaration of competing interest

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

### Acknowledgement

Thilo Hofmann acknowledges funding from the Austrian Science Fund, Cluster of Excellence CoE7, Grant-DOI 10.55776/COE7.

### Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.watres.2024.123032](https://doi.org/10.1016/j.watres.2024.123032).

### Data availability

Data will be made available on request.

### References

- Aggarwal, C.C., Reddy, C.K., 2014. Data clustering: Algorithms and Applications, 1st ed. Chapman and Hall/CRC, New York.
- Arp, H.P.H., Brown, T.N., Berger, U., Hale, S.E., 2017. Ranking REACH registered neutral, ionizable and ionic organic chemicals based on their aquatic persistency and mobility. *Environ. Sci.: Proc. Imp.* 19 (7), 939–955.
- Aumeier, B.M., Georgi, A., Saeidi, N., Sigmund, G., 2023. Is sorption technology fit for the removal of persistent and mobile organic contaminants from water? *Sci. Total Environ.* 880, 163343.
- Boehm, H.P., 1994. Some aspects of the surface chemistry of carbon blacks and other carbons. *Carbon N Y* 32 (5), 759–769.
- Driver, S.J., Perdue, E.M., 2014. Advances in the Physicochemical Characterization of Dissolved Organic matter: Impact on Natural and Engineered Systems. American Chemical Society, pp. 75–86.
- EPA, U., 2024. Estimation Programs Interface Suite™ For Microsoft® Windows. Environmental Protection Agency, Washington, DC, USA v 4.1.
- Freundlich, H., 1907. Über die Adsorption in Lösungen. *Zeitschrift für Physikalische Chemie* 57U, 385–470.
- Hale, S.E., Neumann, M., Schliebner, I., Schulze, J., Averbeck, F.S., Castell-Exner, C., Collard, M., Drmač, D., Hartmann, J., Hofman-Caris, R., Hollender, J., de Jonge, M., Kullick, T., Lennquist, A., Letzel, T., Nödler, K., Pawlowski, S., Reineke, N., Rorije, E., Scheurer, M., Sigmund, G., Timmer, H., Trier, X., Verbruggen, E., Arp, H.P.H., 2022. Getting in control of persistent, mobile and toxic (PMT) and very persistent and very mobile (vPvM) substances to protect water resources: strategies from diverse perspectives. *Environ. Sci. Eu.* 34 (1), 22.
- Hsieh, H.-S., Pignatello, J.J., 2017. Activated carbon-mediated base hydrolysis of alkyl bromides. *Appl. Catal. B: Environ.* 211, 68–78.
- Kah, M., Sigmund, G., Xiao, F., Hofmann, T., 2017. Sorption of ionizable and ionic organic compounds to biochar, activated carbon and other carbonaceous materials. *Water Res.* 124, 673–692.
- Qu, X., Liu, P., Zhu, D., 2008. Enhanced sorption of polycyclic aromatic hydrocarbons to tetra-alkyl ammonium modified smectites via cation- $\pi$  interactions. *Environ. Sci. Technol.* 42 (4), 1109–1116.
- Reemtsma, T., Berger, U., Arp, H.P.H., Gallard, H., Knepper, T.P., Neumann, M., Quintana, J.B., Voogt, P.d., 2016. Mind the gap: persistent and mobile organic compounds—Water contaminants that slip through. *Environ. Sci. Technol.* 50 (19), 10308–10315.
- Saeidi, N., Kopinke, F.-D., Georgi, A., 2020a. Understanding the effect of carbon surface chemistry on adsorption of perfluorinated alkyl substances. *Chem. Eng. J.* 381, 122689.
- Saeidi, N., Kopinke, F.-D., Georgi, A., 2020b. What is specific in adsorption of perfluoroalkyl acids on carbon materials? *Chemosphere*, 128520.
- Saeidi, N., Kopinke, F.-D., Georgi, A., 2021. Controlling adsorption of perfluoroalkyl acids on activated carbon felt by means of electrical potentials. *Chem. Eng. J.* 416, 129070.
- Saeidi, N., Lai, A., Harnisch, F., Sigmund, G., 2024. A FAIR comparison of activated carbon, biochar, cyclodextrins, polymers, resins, and metal organic frameworks for the adsorption of per- and polyfluorinated substances. *Chem. Eng. J.* 498, 155456.
- Sigmund, G., Arp, H.P.H., Aumeier, B.M., Bucheli, T.D., Chefetz, B., Chen, W., Droge, S.T., Endo, S., Escher, B.I., Hale, S.E., Hofmann, T., Pignatello, J., Reemtsma, T., Schmidt, T.C., Schönsee, C.D., Scheringer, M., 2022. Sorption and mobility of charged organic compounds: how to confront and overcome limitations in their assessment. *Environ. Sci. Technol.* 56 (8), 4702–4710.
- Sigmund, G., Gharasoo, M., Hüffer, T., Hofmann, T., 2020. Deep learning neural network approach for predicting the sorption of ionizable and polar organic pollutants to a wide range of carbonaceous materials. *Environ. Sci. Technol.* 54 (7), 4583–4591.
- Székely, G.J., Rizzo, M.L., Bakirov, N.K., 2007. Measuring and testing dependence by correlation of distances. *Ann. Stat.* 35 (6), 2769–2794, 2726.
- Zhang, K., Zhong, S., Zhang, H., 2020. Predicting aqueous adsorption of organic compounds onto biochars, carbon nanotubes, granular activated carbons, and resins with machine learning. *Environ. Sci. Technol.* 54 (11), 7008–7018.
- Zhao, Y., Lin, S., Choi, J.-W., Bediako, J.K., Song, M.-H., Kim, J.-A., Cho, C.-W., Yun, Y.-S., 2019. Prediction of adsorption properties for ionic and neutral pharmaceuticals and pharmaceutical intermediates on activated charcoal from aqueous solution via LFER model. *Chem. Eng. J.* 362, 199–206.
- Zhi, Y., Liu, J., 2015. Adsorption of perfluoroalkyl acids by carbonaceous adsorbents: effect of carbon surface chemistry. *Environ. Pollut.* 202, 168–176.
- Zhi, Y., Liu, J., 2016. Surface modification of activated carbon for enhanced adsorption of perfluoroalkyl acids from aqueous solutions. *Chemosphere* 144, 1224–1232.
- Zhou, J., Saeidi, N., Wick, L.Y., Kopinke, F.-D., Georgi, A., 2021. Adsorption of polar and ionic organic compounds on activated carbon: surface chemistry matters. *Sci. Total Environ.* 794, 148508.