Environmental Science & Technology

pubs.acs.org/est

Article

Characterizing Regional Variation of PFAS Profiles in Human Serum from China by Integrated Target and Nontarget Analysis with Machine Learning

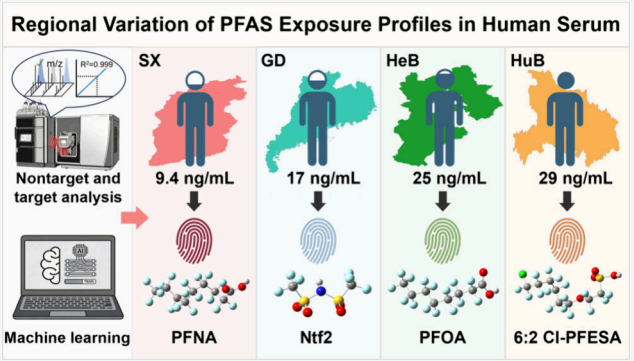
Huikang Lin, Chang He, Anna Ji, Xiaohui Yu, Xiaoping Wang, Guiying Li, and Taicheng An*

Cite This: https://doi.org/10.1021/acs.est.6c04503

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Per- and polyfluoroalkyl substances (PFASs) are globally detected in humans, yet regional variations in PFAS exposure remain insufficiently characterized. Herein, by utilizing target and nontarget analysis with machine learning, differences in PFAS exposure were assessed based on 1088 serum samples of the population from SX, GD, HeB, and HuB regions in China. Nontarget analysis identified 49 PFASs, among which bis-(trifluoromethane)sulfonimide (Ntf2), C8 per- and polyfluoroalkyl ether carboxylic acid, C12 hydrogen-substituted polyfluoroalkyl (linear) carboxylic acid, 8:2 fluorotelomer unsaturated carboxylic acid, 1-hydroxy-6:2 fluorotelomer sulfonic acid, haloxyfop, and leflunomide were reported in human serum for the first time. Target analysis indicated total PFAS concentrations of 9.4, 17, 25, and 29 ng/mL in the SX, GD, HeB, and HuB regions, respectively, with compositional profiles differing in the proportions of specific PFAS classes. Machine learning classifiers achieved an accuracy of 0.89 in distinguishing samples from the four regions, revealing significant interregional disparity. Furthermore, PFNA, Ntf2, PFOA, and 6:2 Cl-PFESA were identified as distinctive exposure biomarkers for SX, GD, HeB, and HuB, respectively, each being closely associated with local industrial activities. Overall, this study elucidates the region-specific characteristics of human PFAS exposure in China, providing a scientific basis for regionally tailored health risk assessment and pollution control strategies.



KEYWORDS: *PFASs, Human exposure, Regional difference, Machine learning, Nontarget analysis*

1. INTRODUCTION

Per- and polyfluoroalkyl substances (PFASs) constitute a class of synthetic chemicals widely used in diverse commercial and industrial applications, garnering increasing concern because of their ubiquity, persistence, bioaccumulation, and health impacts.[1] Human exposure to PFASs is pervasive, occurring through drinking water,[2] food,[3] clothing,[4] and various consumer and personal care products.[5] This widespread presence is particularly concerning given emerging evidence linking PFAS exposure to adverse health outcomes in humans, such as reproductive and developmental issues, immune and thyroid dysfunction, neurophysiological abnormality, and damage to the liver, kidney, and cardiovascular system.[6] Consequently, research on human biomonitoring of PFASs is of great significance, as it provides essential information for health risk assessment.

Exposure to PFASs is not confined to populations in contaminated areas or occupationally exposed workers but pervasively impacts the general population.[7] Given the growing awareness about universal exposure, spatial variability in PFAS exposure levels across different populations has attracted much interest. For instance, global human biomonitoring data showed pronounced regional disparities in PFAS exposure with elevated levels observed in highly industrialized regions including North America, Europe, Asia, and Oceania.[8] Similarly, a nationwide exposome study in China identified PFASs as the most regionally distinct chemicals with exposure levels varying significantly by province.[9] Despite these observations, the relationship between human PFAS exposure and regional characteristics remains poorly understood, especially the identification of internal exposure biomarkers of PFASs, defined as parent analytes capable of effectively distinguishing population groups across different regions.

There is growing interest in developing forensic tools to characterize internal exposure biomarkers across diverse human populations.[10,11] While conventional statistical meth-

Received: March 26, 2026
Revised: July 22, 2026
Accepted: July 31, 2026

ods are limited in analyzing high-dimensional environmental exposure data, machine learning (ML) offers a powerful alternative by capturing important features from complex and nonlinear interactions among multiple exposure congeners.[12,13] For example, ML applied to PFAS profiles in environmental samples has successfully discriminated between contamination sources (i.e., comparing aqueous film-forming foam impacted water with nonimpacted water) on the basis of unique PFAS signatures, showing good classification efficiency.[14−16] However, this approach has not been well applied to human biomonitoring yet. The intrinsic properties of PFASs, including source-specific uses and the resulting exposure,[17] persistence in the human body,[18] and preferential bioaccumulation in blood through lipid and protein binding,[19] enable the linkage of internal exposure to external sources. Therefore, the combination of blood PFAS biomonitoring with ML analysis is a promising strategy to differentiate regional exposure patterns and identify population-specific biomarkers.

The research on PFAS exposure in humans is bottlenecked due to the large knowledge gap of unexplored PFASs. While over 14,000 PFASs are listed in the United States Environmental Protection Agency (US EPA) CompTox Chemicals Dashboard database,[20] conventional target analysis typically detects fewer than 100 PFASs.[21] Recent advances in nontarget analysis using high-resolution mass spectrometry (HRMS) have enabled researchers to identify previously unknown PFASs in humans,[22−24] underscoring substantial gaps in our understanding of total exposure. However, the analysis of unknown PFASs in human blood is still challenging due to interference from complex mixtures of endogenous metabolites.[25] These coextracted compounds not only closely resemble the accurate mass of suspected PFASs but also occupy limited scan time in HRMS analysis, thus hindering the acquisition of tandem mass spectrometry (MS/MS) spectra of PFASs. However, common PFAS extraction methods (e.g., acetonitrile-based liquid−liquid extraction with dispersive powders) may not adequately reduce matrix effects, potentially limiting the discovery of novel PFASs. Thus, developing an improved pretreatment method for nontarget analysis is warranted to explore previously unknown PFAS exposure in humans.

In this study, a total of 1088 serum samples of the general population were collected from four regions in China, and the exposure profiles of PFASs were comprehensively investigated using target and nontarget methods integrated with ML. The four regions span southern, central, and northern China with representative industrial bases including traditional energy industry in SX, electronics in GD, heavy industry in HeB, and metallurgy in HuB. Specifically, the aims of the current work include (1) developing an optimized nontarget analysis method and discovering previously unknown PFASs in human serum samples, (2) assessing the human exposure profiles of PFASs across different regions, and (3) identifying unique PFAS exposure biomarkers in different specific regions.

2. MATERIALS AND METHODS

2.1. Chemicals

A total of 27 PFASs and their corresponding internal standards were obtained from Wellington Laboratories, Inc. (Ontario, Canada). Other chemicals used in this study are listed in Text S1.

2.2. Sample Collection

A total of 1088 human serum samples of the general population were collected from SX, GD, HeB, and HuB regions in China. All serum samples were delivered to the laboratory and stored at −80 °C until analysis. All participants were informed of the purpose of the study, and their written consent was all obtained. This study was approved by the Ethics Committee of the Guangdong University of Technology (case number: GDUTXS20260152). Detailed demographic information on the participants is given in Table S1.

2.3. Nontarget Analysis

2.3.1. Sample Preparation. A pooling strategy was conducted for the nontarget analysis to combine serum samples from the same region into a single sample. A modified fast, simple, cheap, effective, robust, and safe (QuEChERS) method was adopted for the sample pretreatment.[26] The sample was further purified by a Captiva Enhanced Matrix Removal (EMR)-Lipid cartridge (Agilent, CA, USA). Detailed procedures are provided in Text S2.

2.3.2. Instrumental Analysis. Sample analysis was performed on a Dionex U3000 ultraperformance liquid chromatography system interfaced with a Q Exactive Plus hybrid quadrupole Orbitrap (Thermo Fisher Scientific, CA, USA). The settings of the instrument were adapted from a previous study,[22] including full-scan and data-dependent acquisition (DDA) mode to screen for PFASs. A detailed description of the method is provided in Table S2.

2.3.3. PFAS Identification. PFAS identification was performed using Compound Discoverer, MS-DIAL, and FluoroMatch software.[27,28] Confidence levels of identified PFASs were assigned according to Schymanski et al.[29] Detailed methodology is presented in Figure S1 and Text S2.

2.4. Target Analysis

2.4.1. Sample Preparation. A pooling strategy was used for target analysis of PFASs based on the sampling regions. This resulted in four groups with 70 pools, which were equally contributed to by 1088 individual serum samples. The ion-pair extraction method reported in the reference was used for the sample preparation.[30] A detailed protocol is provided in Text S3.

2.4.2. Instrumental Analysis. PFAS quantification was performed using a 1260 Infinity high-performance liquid chromatography system coupled with 6470 tandem mass spectrometry (Agilent, CA, USA), and parameters of instrumental analysis are detailed in Table S3.

2.5. Machine Learning

Logistic regression (LR), random forest (RF), extreme gradient boosting (XGBoost), and support vector machine (SVC), representing both linear and tree-based algorithms, were established to classify PFAS exposure profiles and identify diagnostic PFAS biomarkers for different regional groups. The data set was split into training and testing sets in an 80:20 ratio. Model performance was evaluated using confusion matrices. Feature importance of the classifiers was assessed with the SHapley Additive exPlanations (SHAP) package. All classifiers in scikit-learn were trained and tested on a Python (v3.9.7) environment. Detailed procedures and code for the ML classifiers are provided in Text S4.

2.6. Quality Assurance and Quality Control (QA/QC)

Each batch of samples was accompanied by a procedure blank (ultrapure water) and a pooled QC sample. An instrumental blank (methanol) was injected every ten samples. Comparable concentrations of perfluoropentanoic acid (PFPeA), perfluorohexanoic acid (PFHxA), and hexafluoropropylene oxide dimer acid (HFPO−DA) were detected between instrumental blanks and serum samples, and therefore, they were excluded in this study. Apart from those, no additional PFASs were detected in the procedure blanks. PFAS levels that were below limits of quantification (LOQ) were substituted with LOQ/2. The LOQ, spike recovery, accuracy, and matrix effect were all carefully evaluated using fetal bovine serum and are provided in Table S4.

https://doi.org/10.1021/acs.est.6c04503
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

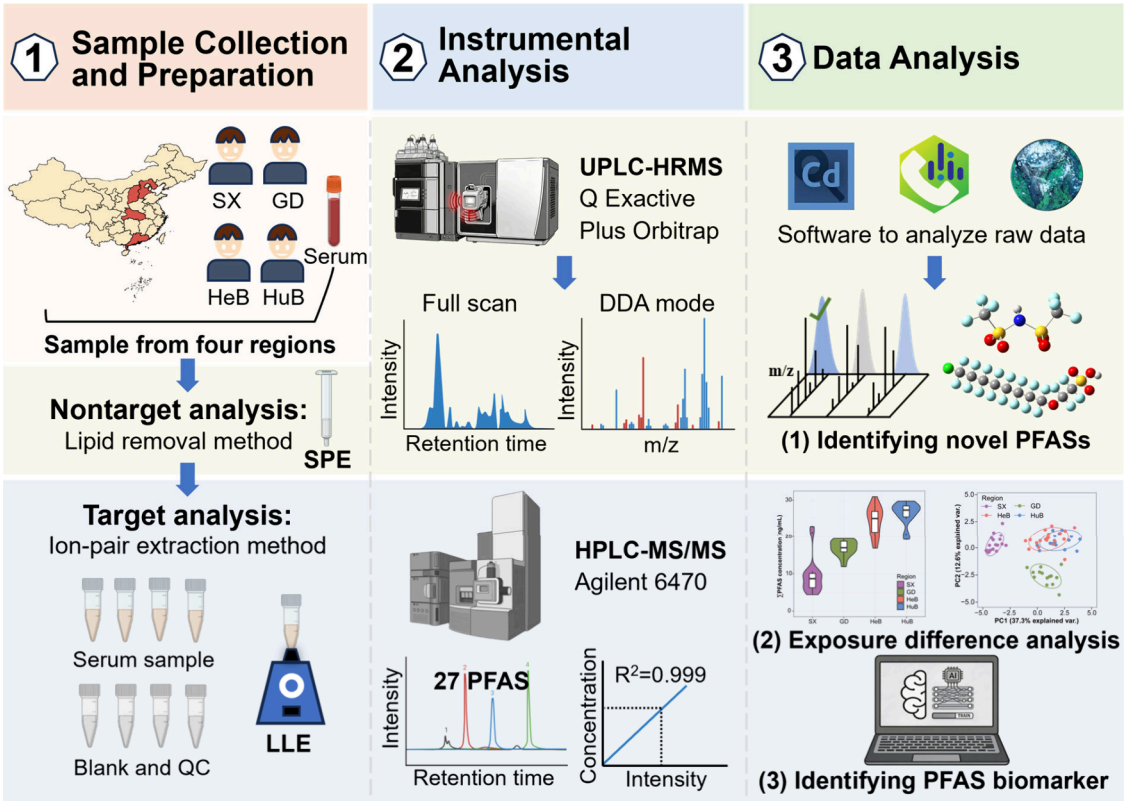


Figure 1. Schematic workflow for comprehensive PFAS analysis in human serum. It comprises lipid removal and ion-pair extraction for preparation, LC-HRMS and LC-MS/MS for detection, and data processing for novel PFAS identification, exposure difference assessment, and biomarker screening.

2.7. Statistical Analysis

Data are presented as means ± standard deviations (SDs). Normality of data was assessed using the Kolmogorov–Smirnov test, and homogeneity of variances was evaluated using Levene's test. Normally distributed data were compared using one-way analysis of variance (ANOVA), whereas the Kruskal–Wallis test was applied for skewed data. A p-value <0.05 was considered statistically significant.

2.8. Health Risk Assessment

The bioaccumulation factor (BCF), half-life ($t_{1/2}$), rat oral acute toxicity, and carcinogenicity of newly discovered PFASs were predicted using the ADMETlab 3.0 tool,[31] and the resulting health risks were ranked using the toxicological priority index (ToxPi). Hazard quotients (HQs) were calculated according to a previous study by comparing serum PFAS concentrations with the reference levels (RfLs) from the Agency for Toxic Substances and Disease Registry (ATSDR).[32]

3. RESULTS AND DISCUSSION

3.1. Nontarget Analysis Discovers Novel PFASs in Human Serum

The overall experimental procedure and data analysis workflow of this study are illustrated in Figure 1. During optimization of the sample pretreatment methodology, initial extraction of serum using acetonitrile alone yielded high recoveries of internal standards (ranging from 75% to 117%) but left a visible matrix residue after nitrogen evaporation. In order to minimize matrix interference from coextractives in subsequent PFAS screening, a solid-phase extraction (SPE) method was used for sample purification. The utilization of a weak anion-exchange cartridge led to a notable decrease in recovery rates (ranging from 15% to 75%) and resulted in the loss of specific PFASs lacking acidic functional groups, prompting the exclusion of this method. Alternatively, the application of the Captiva EMR-Lipid cartridge effectively removed lipid-related matrix components without significant PFAS loss, achieving recovery efficiencies within a range of 70% to 109%. Notably, lipid purification increased the number of identifiable PFASs by 10% compared with untreated extracts. This improvement is critical because the instrument selects a limited number of precursor ions (TopN) based on intensity for MS/MS analysis in DDA mode, meaning that low-abundance PFAS signals can be obscured.[33] Highly abundant lipid coextractives can occupy TopN scan events, thereby hindering the acquisition of quality MS/MS spectra for PFAS identification. Additionally, lipids act as significant interferents in bioanalysis by mass spectrometry, leading to matrix effects on electrospray ionization efficiency.[34] Lipid removal further permits larger injection volumes, enhancing the detectability of low-concentration PFASs. Consequently, the Captiva EMR-Lipid method was selected for subsequent nontarget analysis.

By nontarget analysis, a total of 49 PFASs belonging to 16 homologous series were identified in pooled serum samples at confidence level ≥ 3 (Table 1). These characteristic mass spectra are presented in Figures S2–S16. In a CF_2-adjusted mass defect plot, the homologous series appeared as horizontal lines (Figure S17), and retention times increased with carbon chain length (Table 1).

3.1.1. Polyfluoroalkyl Carboxylic Acids (PFCAs), Perfluoroalkyl Sulfonic Acids (PFSAs), and Their Novel Derivatives. Among the identified compounds, 19 were legacy PFCAs and PFSAs, whereas 10 comprised chlorine-, hydrogen-, and ketone-substituted or ether-inserted deriva-

https://doi.org/10.1021/acs.est.6c04503
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

Table 1. Summary of PFASs Identified by Nontarget Analysis at Confidence Level 3 or Higher

Class	Compound	Formula	m/z	RT (min)	Level	MS/MS fragments
PFCAs	TFA	$C_2HF_3O_2$	112.98504	1.1	2	$[M\text{-}CHO_2]^-$; $[C_nF_{2n+1}]^-$
	PFPA	$C_3HF_5O_2$	162.98184	1.62	2	
	PFBA	$C_4HF_7O_2$	212.97864	5.85	1	
	PFPeA	$C_5HF_9O_2$	262.97544	12.36	1	
	PFHxA	$C_6HF_{11}O_2$	312.97224	20.05	1	
	PFHpA	$C_7HF_{13}O_2$	362.96904	26.28	1	
	PFOA	$C_8HF_{15}O_2$	412.96584	31.01	1	
	PFNA	$C_9HF_{17}O_2$	462.96264	34.76	1	
	PFDA	$C_{10}HF_{19}O_2$	512.95944	37.85	1	
	PFUnA	$C_{11}HF_{21}O_2$	562.95624	40.39	1	
	PFDoA	$C_{12}HF_{23}O_2$	612.95304	42.38	1	
	PFTrDA	$C_{13}HF_{25}O_2$	662.94984	44.23	1	
	PFTeDA	$C_{14}HF_{27}O_2$	712.94664	45.39	1	
	PFODA	$C_{18}HF_{35}O_2$	912.93384	51.15	2	
PFSAs	PFBS	$C_4HF_9O_3S$	298.94243	14.67	1	$[M\text{-}H]^-$; $[SO_3F]^-$; $[SO_3]^-$
	PFHxS	$C_6HF_{13}O_3S$	398.93603	27.16	1	
	PFHpS	$C_7HF_{15}O_3S$	448.93283	31.55	1	
	PFOS	$C_8HF_{17}O_3S$	498.92963	35.05	1	
	PFDS	$C_{10}HF_{21}O_3S$	598.92323	40.46	1	
Cl-PFSAs	Cl-PFOS	$C_8HF_{16}ClO_3S$	514.90008	35.75	3	
Keto-PFSAs	Keto-PFOS	$C_8HF_{15}O_4S$	476.92775	32.49	3	
H-PFCAs	H-PFDA	$C_{10}H_2F_{18}O_2$	494.96887	31.02	3	$[M\text{-}H]^-$; $[C_nF_{2n+1}]^-$
	H-PFUnA	$C_{11}H_2F_{20}O_2$	544.96567	35.13	3	
mHPFLCAs	C12 HPFLCA	$C_{12}H_{12}F_{12}O_2$	415.05672	21.14	3	$[M\text{-}H]^-$; $[C_{11}H_nF_{n+1}]^-$
PFECAs	C8 PFECA	$C_8HF_{15}O_8$	508.93536	38.01	3	$[M\text{-}H]^-$; $[CF_3O]^-$
FTSAs	4:2 FTSA	$C_6H_5F_9O_3S$	326.97373	19.51	2	$[M\text{-}H]^-$; $[M\text{-}H_2F]^-$; $[HSO_3]^-$; $[SO_3]^-$
	6:2 FTSA	$C_8H_5F_{13}O_3S$	426.96733	30.75	1	
	8:2 FTSA	$C_{10}H_5F_{17}O_3S$	526.96093	37.76	1	
	1-Hydroxy-6:2 FTSA	$C_8H_5F_{13}O_4S$	442.96225	28.2	3	$[M\text{-}H]^-$; $[M\text{-}H_3O]^-$; $[SO_3]^-$
Cl-PFESAs	4:2 Cl-PFESA	$C_6HF_{12}ClO_4S$	430.9014	30.23	2	$[M\text{-}H]^-$; $[C_nF_{2n}ClO]^-$; $[SO_3F]^-$
	5:2 Cl-PFESA	$C_7HF_{14}ClO_4S$	480.8982	33.98	2	
	6:2 Cl-PFESA	$C_8HF_{16}ClO_4S$	530.895	36.71	1	
	8:2 Cl-PFESA	$C_{10}HF_{20}ClO_4S$	630.8886	41.53	2	
FTUCAs	8:2 FTUCA	$C_{10}H_2F_{16}O_2$	456.97207	35.56	3	$[M\text{-}H]^-$; $[C_2F_5]^-$; $[C_9F_{15}]^-$
Unsaturated PFA	C6 UPFA	$C_6HF_{11}O$	296.97732	26.28	2	$[M\text{-}H]^-$; $[C_nF_{2n-1}]^-$
	C7 UPFA	$C_7HF_{13}O$	346.97412	31.02	2	
	C8 UPFA	$C_8HF_{15}O$	396.97092	34.88	2	
	C9 UPFA	$C_9HF_{17}O$	446.96772	37.84	2	
	C10 UPFA	$C_{10}HF_{19}O$	496.96452	40.41	2	
PFSMs	PFOSA	$C_8H_2F_{17}NO_2S$	497.94561	39.56	1	$[M\text{-}H]^-$; $[SO_2N]^-$; $[SO_2]^-$
CA-PFSMs	N-MeFOSAA	$C_{11}H_6F_{17}NO_4S$	569.96675	38.99	2	$[M\text{-}H]^-$; $[C_nF_{2n+1}]^-$; $[SO_2F]^-$
	N-MeFOSA	$C_9H_4F_{17}NO_2S$	511.96126	43.76	2	
	N-EtFOSAA	$C_{12}H_8F_{17}NO_4S$	583.98239	40.19	2	
	N-EtFOSA	$C_{10}H_6F_{17}NO_2S$	525.97691	45.22	2	
bis-FASI	Ntf2	$C_2HF_6NO_4S_2$	279.91729	8.56	1	$[M\text{-}H]^-$; $[CF_3SO_2N]^-$; $[SO_2N]^-$
diPAPs	6:2 diPAP	$C_{16}H_9F_{26}O_4P$	788.97444	45.72	1	$[M\text{-}H]^-$; $[H_2PO_4]^-$; $[PO_3]^-$
Fluorinated pesticides and pharmaceuticals	Fipronil sulfone	$C_{12}H_4F_6N_4Cl_2O_2S$	450.9258	36.22	2	$[M\text{-}H]^-$; $[CF_3SO_2]^-$; $[CF_3]^-$
	Haloxyfop	$C_{15}H_{11}F_3NClO_4$	360.02505	26.38	2	$[M\text{-}H]^-$; $[C_{12}H_6F_3ClNO_2]^-$; $[C_{12}H_5F_3NO_2]^-$; $[C_6H_2F_3ClNO]^-$
	Leflunomide	$C_{12}H_9F_3N_2O_2$	269.05379	19.86	2	$[M\text{-}H]^-$; $[C_7H_5F_3N]^-$; $[C_4H_4NO]^-$

tives, including H-PFCAs, ether-PFCAs, Cl-PFSAs, keto-PFSAs, and fluorotelomer sulfonic acids (FTSAs). PFCAs and their derivatives typically exhibited the diagnostic fragment ion $[C_nF_{2n+1}]^-$. Notably, a neutral loss of [HF] (m/z = 20.00623) was identified as the characteristic fragmentation pathway for polyfluoroalkyl linear carboxylic acids (HPFLCAs), which was reported in human serum for the first time. In addition, a novel polyfluoroalkyl ether carboxylic acid (PFECA), namely, C8 PFECA, was annotated based on the dominant fragment ion $[CF_3O]^-$ (m/z = 84.99012), corresponding to the terminal ether moiety. PFSAs and their derivatives generally exhibited diagnostic fragment ions $[SO_3]^-$ (m/z = 79.95683), $[SO_3F]^-$ (m/z = 98.95523), or $[SO_3H]^-$ (m/z = 80.96466), indicating the presence of a sulfonic acid group. FTSAs showed a neutral loss of [HF] (m/z = 20.00623) from the precursor ion, whereas 1-hydroxy-6:2

https://doi.org/10.1021/acs.est.6c04503
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

FTSA exhibited a neutral loss of $[H_2O]$ (m/z = 18.01057), attributable to its hydroxyl group and adjacent methylene hydrogens. Importantly, 1-hydroxy-6:2 FTSA was identified for the first time, warranting further investigation into its environmental occurrence and health risks relative to those of 6:2 FTSA.

3.1.2. Chlorinated Polyfluorinated Ether Sulfonic Acids (Cl-PFESAs). Based on the fragment ions $[C_nF_{2n}ClO]^-$ and $[SO_3F]^-$, four X:2 Cl-PFESA congeners were identified, namely, 4:2, 5:2, 6:2, and 8:2 Cl-PFESA. Since the first report of 6:2 Cl-PFESA in metal electroplating wastewater in 2013,[35] members of this homologous series have been progressively detected in human blood using nontarget analytical approaches,[36] consistent with the present study.

3.1.3. Fluorotelomer Unsaturated Carboxylic Acids (FTUCAs) and Unsaturated Perfluoroalkyl Alcohols (UPFAs). PFASs containing unsaturated bonds were also discovered in this study. The presence of 8:2 FTUCA was confirmed by the dominant fragment ion $[C_9F_{15}]^-$ (m/z = 392.97600), generated through neutral loss of [HF] from the precursor ion, which was overlooked in previous investigations. Furthermore, based on the characteristic fragment ions $[C_nF_{2n-1}]^-$, C6−C10 unsaturated UPFAs were identified, consistent with prior reports in human blood.[22]

3.1.4. Perfluoroalkyl Sulfonamides (PFSMs), Carboxylic Acid-PFSMs (CA-PFSMs), and bis-Perfluoroalkyl Sulfonimides (bis-FASIs). Six PFASs containing sulfonamide groups were identified based on the characteristic fragment ions $[SO_2F]^-$ (m/z = 82.96031), $[SO_2N]^-$ (m/z = 77.96498), or $[SO_2]^-$ (m/z = 63.96191), including perfluorooctane sulfonamide (PFOSA), four CA-PFSM homologues, and bis(trifluoromethane)sulfonimide (Ntf2). In addition to sulfonamide-related fragments, PFOSA and CA-PFSMs exhibited a series of diagnostic fragment ions $[C_nF_{2n+1}]^-$, indicating the presence of a perfluoroalkyl backbone. Notably, Ntf2 yielded the fragment ion $[CF_3NO_2S]^-$ (m/z = 146.96018), indicative of a terminal CF_3 group attached to a sulfonamide structure. To our knowledge, this compound was identified in human serum for the first time, despite its widespread occurrence in various environmental matrices.[37,38]

3.1.5. Polyfluoroalkyl Phosphate Esters (diPAPs). A single diPAP compound, 6:2 diPAP, was identified based on its precursor ion (m/z = 788.97444) and two dominant fragment ions $[PO_4H_2]^-$ (m/z = 96.96909) and $[PO_3]^-$ (m/z = 78.95852).

3.1.6. Fluorinated Pesticides and Pharmaceuticals. Three compounds bearing terminal CF_3 groups, namely, fipronil sulfone, haloxyfop, and leflunomide, were annotated through MS/MS spectral matching. Fluorinated moieties are commonly introduced into pesticides and pharmaceuticals to enhance bioavailability and stability, representing important yet often overlooked sources of PFAS-related exposure.[39] Owing to limited signal intensity and background interference, only fipronil sulfone exhibited the diagnostic fragment ion $[CF_3]^-$ (m/z = 68.99520). These findings highlight the necessity of expanding spectral libraries for fluorinated pesticides and pharmaceuticals in future nontarget PFAS analyses.

Semiquantitative analysis revealed distinct regional patterns in PFAS exposure profiles (Figure 2). For example, PFASs containing a carboxylic acid group were most abundant in HeB samples, whereas those with a sulfonic acid group predominated in HuB samples. This exposure pattern arises from

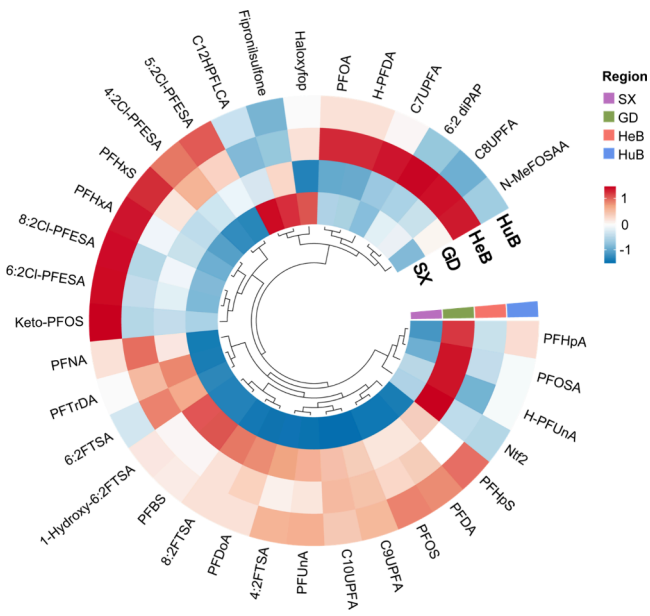


Figure 2. Relative abundance of nontarget PFAS profiles in human serum. The four circles show regional sources; each column represents each PFAS. The color of the squares represents the relative levels of the corresponding PFASs. Data of each PFAS (each column) was normalized based on Min−Max scaling.

regional distinctions in industrial bases, particularly fluoropolymer applications (e. g., coatings, food packaging, and textiles) in the HeB region that predominantly rely on PFCAs,[40,41] as well as metallurgy industries and PFSA-based fluorochemical manufacturing in the HuB region,[42] both of which contribute to differences in population exposure across PFAS classes. Furthermore, Ntf2, which is extensively used as an ionic liquid anion in electronic devices and lithium batteries,[43] exhibited a 20.17-fold higher abundance in samples from the GD region, which hosts most of China's electronics industries, than in samples from other regions. Additionally, fipronil sulfone and haloxyfop, which are applied as agricultural chemicals,[44,45] were most prevalent in SX samples with relative abundances 4.59- and 1.51-fold greater than those in other regions, respectively, consistent with the region's intensive agricultural activity. These results indicated that regional populations exhibited distinct exposure patterns characterized by elevated levels of specific PFAS classes, likely reflecting local industrial and agricultural practices, requiring further validation via target analysis.

3.2. Target Analysis Reveals Regional Differences of PFAS Exposure Profiles

The detection frequencies and concentrations of 24 target PFASs in the 70 pooled serum samples are summarized in Table S5. Five novel PFASs identified by nontarget analysis in the absence of authentic standards were subsequently detected via pseudotargeted analysis, which are provided in Table S6. Perfluorooctanoic acid (PFOA), perfluorooctanesulfonic acid (PFOS), perfluorohexanesulfonic acid (PFHxS), perfluoroheptanesulfonic acid (PFHpS), perfluorononanoic acid (PFNA), perfluorodecanoic acid (PFDA), perfluoroundecanoic acid (PFUnA), perfluorododecanoic acid (PFDoA), perfluorotridecanoic acid (PFTrDA), 6:2 FTSA, 6:2 Cl-PFESA, and Ntf2 were detected in all samples. Perfluorobutanoic acid (PFBA), perfluoroheptanoic acid (PFHpA), and perfluorobutanesul-

https://doi.org/10.1021/acs.est.6c04503
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

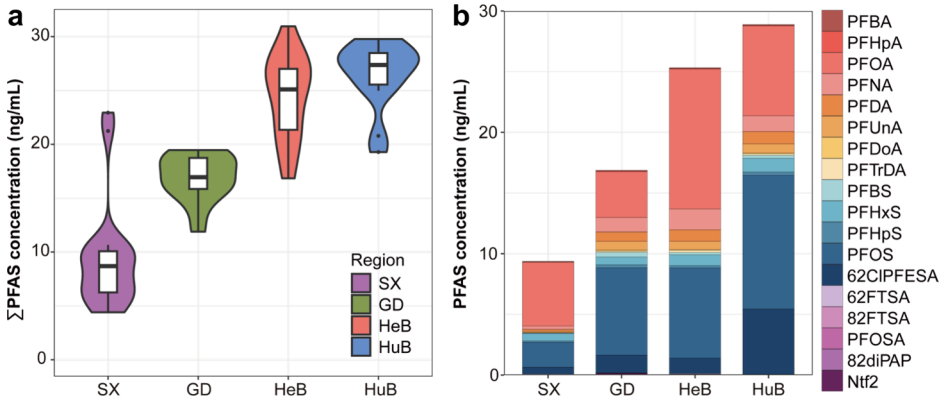


Figure 3. Regional difference of target PFAS profiles in human serum. (a) The total concentration of PFASs across the four regions. (b) The mean concentration of each PFAS in each region.

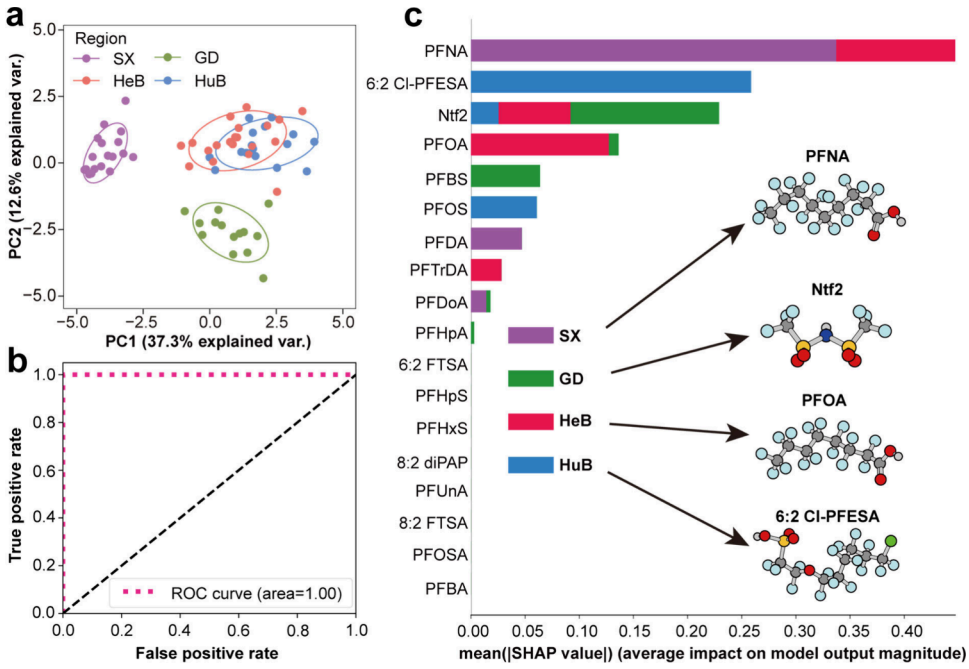


Figure 4. Performance and feature importance of machine learning classifiers to recognize exposure differences of PFASs in human serum from the four regions. (a) Principal component analysis of samples from the four regions. (b) Receiver operating characteristic of multivariate algorithms based on the mean values of LR, RF, XGBoost, and SVC classifiers. (c) SHAP summary plots of important PFASs for the XGBoost classifier. SHAP value means the contribution of PFASs to the output of classifier.

fonic acid (PFBS) were also frequently detected (>50%), whereas 8:2 FTSA, perfluorooctane sulfonamide (PFOSA), and 8:2 diPAP were infrequently detected (<50%). Among the quantified PFASs, the legacy PFOA and PFOS were the most prevalent compounds with mean concentrations of 7.3 and 6.9 ng/mL, respectively. The concentrations of other PFASs ranged from 0.0026 to 2.2 ng/mL. In terms of the compositional profile, PFOA and PFOS together contributed over 60% to the total PFAS burden (ΣPFASs), indicating their persistent dominance in human serum in China despite regulatory actions initiated in 2011.[46] With increasing regulatory restrictions on certain legacy PFASs, increasing attention has shifted to emerging alternatives, such as 6:2 Cl-PFESA, which was detected at considerable levels (mean: 2.2 ng/mL) in this study. This concentration is comparable to those reported in residents near a fluorochemical plant (2.2 ng/mL)[47] and the general population in Shandong, China (2.8 ng/mL),[48] but substantially lower than levels found in occupationally exposed metal electroplating workers (52 ng/mL).[49] Moreover, Ntf2 was found at a mean concentration of 0.039 ng/mL, filling the knowledge gap regarding human serum exposure levels of this novel PFAS. Above all, these results indicated that PFOA and PFOS remain the predominant PFASs in the Chinese population, while an expanding number of novel PFASs is being detected at lower levels, highlighting the necessity for long-term exposure monitoring.

When four regions were compared, ΣPFAS concentrations in serum ranged from 4.4 to 41 ng/mL. The highest mean level was observed in the HuB region (29 ng/mL), followed by the HeB region (25 ng/mL), while lower concentrations were recorded in the GD region (17 ng/mL) and SX region (9.4 ng/mL) (Figure 3a), suggesting significant regional variations in overall PFAS exposure. Similar evidence has revealed that PFAS exposure exhibited regional discrepancies in human blood from other cities or countries,[50−53] indicating geographic

https://doi.org/10.1021/acs.est.6c04503
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

patterns of PFAS exposure on a global scale. In this study, HuB and HeB regions showed higher PFAS exposure level, potentially attributable to the significant contribution of fluorochemical manufacturers in the HuB region and heavy industrial bases in the HeB region.[54−56] Accordingly, a cross-national scale study in China indicated that residents living in cities with heavy industries were exposed to extremely higher levels of PFASs compared to populations in other cities.[57] These results revealed that PFAS exposure in humans exhibited regional variations, while nearby industrial activities play a vital role in shaping regional exposure levels.

Specifically, the concentration of each PFAS also showed distinct regional patterns (Figure 3b). For example, several PFCAs were observed with significantly higher levels in the HeB region than other regions, including PFOA (12 ng/mL compared to 5.6 ng/mL), PFNA (1.7 ng/mL compared to 0.90 ng/mL), and PFTrDA (0.18 ng/mL compared to 0.083 ng/mL). In contrast, novel 6:2 Cl-PFESA (5.4 ng/mL compared to 1.1 ng/mL) and legacy PFSAs, including PFHxS (1.1 ng/mL compared to 0.73 ng/mL) and PFOS (11 ng/mL compared to 5.5 ng/mL), were higher in the HuB region. Furthermore, although ΣPFAS levels were lower in the GD population, PFBS and Ntf2 represented notable exceptions, with mean concentrations of 0.38 and 0.13 ng/mL, respectively, substantially exceeding the corresponding average concentrations observed in other regions (0.15 ng/mL and 0.015 ng/mL, respectively). These findings aligned with the results of nontarget analysis, which might originate from different use characteristics of specific PFAS classes across various regions. On the other hand, the relative proportions of PFASs also varied markedly across the four regions. For instance, PFOA accounted for 56% and 46% of ΣPFASs in participants from SX and HeB regions, respectively, but only 22% and 26% in the GD region and HuB region, respectively. Conversely, PFOS was the predominant PFAS in GD and HuB samples, contributing 43% and 38% to ΣPFASs, respectively (Figure S18). Collectively, these results indicated that PFAS exposure profiles differ significantly across regions, manifesting as differences in both the concentration and composition of specific PFAS classes, which are likely driven by regional differences in particular exposure sources.

3.3. Machine Learning Classifiers Identify Region-Specific PFAS Exposure Biomarker

To assess region-specific differences in PFAS exposure, principal component analysis (PCA), partial least-squares-discriminant analysis (PLS-DA), and hierarchical cluster analysis (HCA) with heatmap visualization were used. PCA and PLS-DA distinctly separated the serum samples into four major clusters corresponding to the four study regions, suggesting variations tied to region-specific sources (Figures 4a and S19). The findings were consistent with HCA results, with clear separation of PFAS profiles among different regions, which manifested in the concentrations of specific PFASs (Figure S20). These analyses revealed that the PFAS exposure profiles could be statistically classified according to four targeted regions, highlighting the predominance of regional influences on the PFAS exposure.

The subsequent application of ML classifiers, including LR, RF, XGBoost, and SVC, demonstrated satisfactory accuracy in recognizing PFAS exposure differences and pinpointing regional sources. The area under the receiver operating characteristic curve (AUROC) for the multilabel classification model reached 1.0 (Figure 4b). The confusion matrix evaluation of four classifiers, including accuracy, precision, F1 score, and balanced accuracy, also exhibited a solid classification performance (Figure S21). The classifiers were further validated using an independent internal set of 80 individual serum samples and a previously published external data set from the JS region in China.[58] When the PFAS profiles from these data sets were input into the models, the regional classifications closely matched the expected sources. The internal validation accuracies were 0.71, 0.80, 0.89, and 0.56 for LR, RF, XGBoost, and SVC classifiers, respectively (Figure S22). The performance gap arises because tree-based RF and XGBoost better capture nonlinear and high-dimensional PFAS exposure features, whereas linear LR and SVC struggle with such complexity.[59] Thus, the XGBoost model with the highest accuracy was included in the subsequent analysis. In the external evaluation, four ML classifiers assigned the PFAS exposure patterns from the JS region to the HuB region. This result can be attributed to the relatively closer distance between the JS and HuB regions, underscoring that the geographical scale is critical for characterizing regional PFAS exposure. Moreover, the JS and HuB regions are industrially developed with fluorochemical plants,[55,60] which leads to high PFAS exposure levels and similar exposure signatures in local populations. These results demonstrated that ML could serve as a reliable forensic tool for differentiating the regional origins of PFAS exposure in human populations, underscoring a strong link between regional environmental contamination and human internal exposure. Notably, ML techniques have been successfully applied to predict PFAS concentrations in groundwater,[61] soil,[62] and fish tissues[63] using environmental and geospatial variables such as known PFAS release sources. Our study indicated that PFAS exposure patterns in humans were shaped by the local regional environment, supporting the feasibility to predict PFAS body burdens. Further investigations on critical regional factors that determine PFAS concentrations in humans are needed to propose a promising prospect for prediction of PFAS levels using ML techniques.

The SHAP method was used to distinguish the PFAS exposure biomarker most indicative of each region. By discovering differences in PFAS exposure, the ML classifiers were able to assign feature importance scores to the overall PFAS exposure spectrum, thus revealing distinct PFAS biomarkers in humans from different regions. Specifically, using the XGBoost classifier, PFNA, Ntf2, PFOA, and 6:2 Cl-PFESA were identified as the most important exposure biomarkers for SX, GD, HeB, and HuB samples, respectively (Figure 4c). In the GD region, Ntf2 was identified as a characteristic exposure signature, reflecting exposure patterns associated with the extensive production of electronic products and lithium batteries in this region. Source apportionment further demonstrated that wastewater from the lithium battery industry in the GD region contained Ntf2 at concentrations as high as 12.4 mg/L, which pinpointed a critical exposure source of Ntf2 in the GD population for the first time. In the HeB region, extensive fluoropolymer production has resulted in relatively severe PFOA contamination, with surface water concentrations reaching 8397 ng/L,[40,41] thereby explaining the identification of PFOA as a representative exposure biomarker in the HeB population. In the HuB region, novel 6:2 Cl-PFESA emerged as a diagnostic exposure biomarker, which was heavily used as a PFOS alternative by the local metal electroplating industry. Correspondingly, PFOS and 6:2 Cl-

https://doi.org/10.1021/acs.est.6c04503
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

PFESA were detected at concentrations up to 35.7 and 13.4 mg/L, respectively, in electroplating wastewater,[64] which demonstrates the close link between the electroplating industry and human exposure to 6:2 Cl-PFESA in the HuB region. Above all, the identified PFAS exposure biomarkers could be tracked back to specific regional environments in which typical PFASs were used with particular relevance to industrial activities such as lithium battery manufacturing, fluoropolymer production, and metal electroplating. Therefore, future studies should prioritize concentrating on key industries and their contributions to the unique exposure characteristics of PFASs in human populations residing in these regions.

3.4. Health Risk Assessment Uncovers Regional Imbalance of PFAS Exposure Hazards

To assess the health risks of the newly identified PFASs, BCF, $t_{1/2}$, acute toxicity, and carcinogenicity were estimated and compared with those of legacy PFASs (Figures 5a and S23).

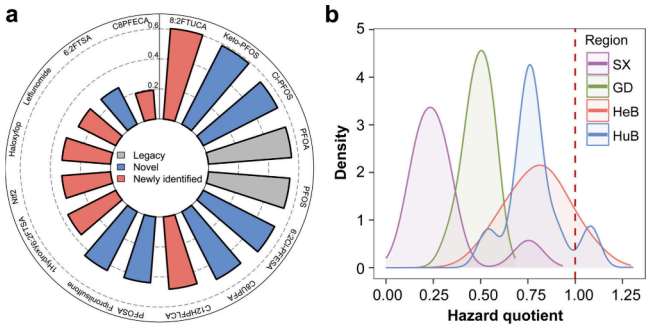


Figure 5. Summary of health risks associated with PFAS exposure. (a) Toxicological priority index of newly discovered PFASs compared with other PFASs. Higher scores indicate greater overall risk. (b) Hazard quotient profiles of PFAS exposure in populations from the four regions. Density curves show the distribution of hazard quotient values among individuals.

The results indicated that the relative risks of 8:2 FTUCA, Keto-PFOS, and Cl-PFOS were greater than those of PFOA and PFOS, suggesting that these compounds warrant priority attention. In contrast, other first discovered PFASs, including 1-hydroxy-6:2 FTSA, Ntf2, haloxyfop, leflunomide, and C8 PFECA, exhibited relatively lower health risks. To further quantify the health risks associated with PFAS exposure across the four regions, we calculated the cumulative HQs of PFOA, PFNA, PFHxS, and PFOS (Figure 5b). The results revealed that HQ values were below 1 in both SX and GD populations, ranging from 0.13 to 0.78. However, the 10th percentile HQ values exceeded 1 in the HeB and HuB populations with a maximum value of 1.1. These findings indicated substantial regional imbalance in health risks associated with PFAS exposure, highlighting the need for initiatives aimed at protecting highly exposed and vulnerable populations.

3.5. Environmental Implication

This study mapped large-scale PFAS exposure profiles in human serum across the four regions of China by using a comprehensive methodological framework. A novel lipid removal method for nontarget analysis was developed to effectively facilitate PFAS discovery, resulting in the detection of 49 PFASs. Seven PFASs were discovered for the first time, which enables a more complete understanding of human exposure to PFASs. By integrating target and nontarget analyses, significant regional differences in PFAS concentrations and compositions were observed, highlighting exposure hotspots in the HeB and HuB regions of China. The ML classifiers achieved satisfactory accuracy to allocate sample sources from the four regions and identify specific exposure biomarkers, demonstrating clear regional variations in human PFAS exposure. Specifically, PFNA, Ntf2, PFOA, and 6:2 Cl-PFESA were identified as representative exposure biomarkers for the SX, GD, HeB, and HuB regions, respectively. These biomarkers were closely associated with nearby industrial activities, including lithium battery manufacturing in GD, fluoropolymer production in HeB, and metal electroplating in HuB. Health risk assessment revealed a regional imbalance of PFAS exposure hazards, emphasizing the urgency of protecting highly exposed populations. Overall, this study provides a meaningful starting point for developing strategies to more accurately characterize regional differences of PFAS exposure in humans, track pollution sources, assess relevant health risks, and implement targeted region-specific control measures. Nevertheless, one limitation of this study was the insufficient environmental evidence directly linking specific industrial sources to regional differences in human PFAS exposures. To formulate targeted policies for controlling PFAS exposure risks across various regions, future studies are warranted to establish high-resolution causal associations between industrial sources and human-exposure disparities.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.est.6c04503.

Experimental details of chemicals, nontarget analysis, target analysis, and machine learning; supplementary tables of demographic characteristics, instrumental parameters, QA/QC, and serum PFAS levels; and supplementary figures of the mass spectrum of nontarget PFASs and health risk assessment of newly discovered PFASs as well as further results on PFASs in the serum samples (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Taicheng An − *Guangdong-Hong Kong-Macao Joint Laboratory for Contaminants Exposure and Health, Key Laboratory of Environmental Catalysis and Health Risk Control, Institute of Environmental Health and Pollution Control, Guangdong University of Technology, Guangzhou 510006, China; Guangdong Basic Research Center of Excellence for Ecological Security and Green Development, Key Laboratory for City Cluster Environmental Safety and Green Development of the Ministry of Education, School of Environmental Science and Engineering, Guangdong University of Technology, Guangzhou 510006, China;* orcid.org/0000-0001-6918-8070; Email: antc99@gdut.edu.cn

Authors

Huikang Lin − *Guangdong-Hong Kong-Macao Joint Laboratory for Contaminants Exposure and Health, Key Laboratory of Environmental Catalysis and Health Risk Control, Institute of Environmental Health and Pollution Control, Guangdong University of Technology, Guangzhou 510006, China; Guangdong Basic Research Center of*

https://doi.org/10.1021/acs.est.6c04503
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

Excellence for Ecological Security and Green Development, Key Laboratory for City Cluster Environmental Safety and Green Development of the Ministry of Education, School of Environmental Science and Engineering, Guangdong University of Technology, Guangzhou 510006, China; orcid.org/0009-0006-3378-8166

Chang He − *Guangdong-Hong Kong-Macao Joint Laboratory for Contaminants Exposure and Health, Key Laboratory of Environmental Catalysis and Health Risk Control, Institute of Environmental Health and Pollution Control, Guangdong University of Technology, Guangzhou 510006, China; Guangdong Basic Research Center of Excellence for Ecological Security and Green Development, Key Laboratory for City Cluster Environmental Safety and Green Development of the Ministry of Education, School of Environmental Science and Engineering, Guangdong University of Technology, Guangzhou 510006, China;* orcid.org/0000-0003-2345-5940

Anna Ji − *Guangdong-Hong Kong-Macao Joint Laboratory for Contaminants Exposure and Health, Key Laboratory of Environmental Catalysis and Health Risk Control, Institute of Environmental Health and Pollution Control, Guangdong University of Technology, Guangzhou 510006, China; Guangdong Basic Research Center of Excellence for Ecological Security and Green Development, Key Laboratory for City Cluster Environmental Safety and Green Development of the Ministry of Education, School of Environmental Science and Engineering, Guangdong University of Technology, Guangzhou 510006, China*

Xiaohui Yu − *Guangdong-Hong Kong-Macao Joint Laboratory for Contaminants Exposure and Health, Key Laboratory of Environmental Catalysis and Health Risk Control, Institute of Environmental Health and Pollution Control, Guangdong University of Technology, Guangzhou 510006, China; Guangdong Basic Research Center of Excellence for Ecological Security and Green Development, Key Laboratory for City Cluster Environmental Safety and Green Development of the Ministry of Education, School of Environmental Science and Engineering, Guangdong University of Technology, Guangzhou 510006, China*

Xiaoping Wang − *College of Resources and Environmental Sciences, China Agricultural University, Beijing 100193, China;* orcid.org/0000-0001-7524-7540

Guiying Li − *Guangdong-Hong Kong-Macao Joint Laboratory for Contaminants Exposure and Health, Key Laboratory of Environmental Catalysis and Health Risk Control, Institute of Environmental Health and Pollution Control, Guangdong University of Technology, Guangzhou 510006, China; Guangdong Basic Research Center of Excellence for Ecological Security and Green Development, Key Laboratory for City Cluster Environmental Safety and Green Development of the Ministry of Education, School of Environmental Science and Engineering, Guangdong University of Technology, Guangzhou 510006, China;* orcid.org/0000-0002-6777-4786

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.est.6c04503

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This study was supported by the National Natural Science Foundation of China (42530702 and 42407567) and the Introduction Innovative and Research Teams Project of Guangdong Pearl River Talents Program (2023ZT10L102).

■ REFERENCES

(1) Glüge, J.; Scheringer, M.; Cousins, I. T.; DeWitt, J. C.; Goldenman, G.; Herzke, D.; Lohmann, R.; Ng, C. A.; Trier, X.; Wang, Z. An overview of the uses of per- and polyfluoroalkyl substances (PFAS). *Environ. Sci.: Process. Impacts* **2020**, *22* (12), 2345−2373.

(2) Wee, S. Y.; Aris, A. Z. Revisiting the "forever chemicals", PFOA and PFOS exposure in drinking water. *npj Clean Water* **2023**, *6* (1), 57.

(3) Domingo, J. L.; Nadal, M. Per- and polyfluoroalkyl substances (PFASs) in food and human dietary intake: A review of the recent scientific literature. *J. Agric. Food. Chem.* **2017**, *65* (3), 533−543.

(4) Hill, P. J.; Taylor, M.; Goswami, P.; Blackburn, R. S. Substitution of PFAS chemistry in outdoor apparel and the impact on repellency performance. *Chemosphere* **2017**, *181*, 500−507.

(5) Kotthoff, M.; Müller, J.; Jürling, H.; Schlummer, M.; Fiedler, D. Perfluoroalkyl and polyfluoroalkyl substances in consumer products. *Environ. Sci. Pollut. Res. Int.* **2015**, *22* (19), 14546−14559.

(6) Sunderland, E. M.; Hu, X. C.; Dassuncao, C.; Tokranov, A. K.; Wagner, C. C.; Allen, J. G. A review of the pathways of human exposure to poly- and perfluoroalkyl substances (PFASs) and present understanding of health effects. *J. Expo. Sci. Environ. Epidemiol.* **2019**, *29* (2), 131−147.

(7) Yeoh, C. S. L.; Alrazihi, L. A.; Wong, S. T.; Wong, S. F. Per- and poly-fluoroalkyl substances (PFAS) and human health: a review of exposure routes and potential toxicities across the lifespan. *Environ. Toxicol. Chem.* **2025**, *44* (10), 2754−2786.

(8) Fan, X.; Tang, S.; Wang, Y.; Fan, W.; Ben, Y.; Naidu, R.; Dong, Z. Global exposure to per- and polyfluoroalkyl substances and associated burden of low birthweight. *Environ. Sci. Technol.* **2022**, *56* (7), 4282−4294.

(9) You, L.; Kou, J.; Wang, M.; Ji, G.; Li, X.; Su, C.; Zheng, F.; Zhang, M.; Wang, Y.; Chen, T.; Li, T.; Zhou, L.; Shi, X.; Zhao, C.; Liu, X.; Mei, S.; Xu, G. An exposome atlas of serum reveals the risk of chronic diseases in the Chinese population. *Nat. Commun.* **2024**, *15* (1), 2268.

(10) Vorkamp, K.; Castaño, A.; Antignac, J. P.; Boada, L. D.; Cequier, E.; Covaci, A.; Esteban López, M.; Haug, L. S.; Kasper-Sonnenberg, M.; Koch, H. M.; Pérez Luzardo, O.; Osite, A.; Rambaud, L.; Pinorini, M. T.; Sabbioni, G.; Thomsen, C. Biomarkers, matrices and analytical methods targeting human exposure to chemicals selected for a European human biomonitoring initiative. *Environ. Int.* **2021**, *146*, 106082.

(11) Fortino, V.; Wisgrill, L.; Werner, P.; Suomela, S.; Linder, N.; Jalonen, E.; Suomalainen, A.; Marwah, V.; Kero, M.; Pesonen, M.; Lundin, J.; Lauerma, A.; Aalto-Korte, K.; Greco, D.; Alenius, H.; Fyhrquist, N. Machine-learning-driven biomarker discovery for the discrimination between allergic and irritant contact dermatitis. *Proc. Natl. Acad. Sci. U.S.A.* **2020**, *117* (52), 33474−33485.

(12) Zhong, S.; Zhang, K.; Bagheri, M.; Burken, J. G.; Gu, A.; Li, B.; Ma, X.; Marrone, B. L.; Ren, Z. J.; Schrier, J.; Shi, W.; Tan, H.; Wang, T.; Wang, X.; Wong, B. M.; Xiao, X.; Yu, X.; Zhu, J. J.; Zhang, H. Machine learning: New ideas and tools in environmental science and engineering. *Environ. Sci. Technol.* **2021**, *55* (19), 12741−12754.

(13) Zhou, Y.; Hu, B.; Wang, X.; Fu, J.; Cousins, I. T.; Bi, X.; Wang, L.; Chen, M.; Wang, C.; Dong, H.; Ren, S.; Wang, Z.; An, T.; Li, D.; Jones, K. C.; Muir, D. C. G. Global flux of perfluoroalkyl acids from glaciers in a warming climate. *One Earth* **2025**, *8* (12), 101453.

(14) Kibbey, T. C. G.; Jabrzemski, R.; O'Carroll, D. M. Supervised machine learning for source allocation of per- and polyfluoroalkyl substances (PFAS) in environmental samples. *Chemosphere* **2020**, *252*, 126593.

https://doi.org/10.1021/acs.est.6c04503
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

(15) Joseph, N. T.; Schwichtenberg, T.; Cao, D.; Jones, G. D.; Rodowa, A. E.; Barlaz, M. A.; Charbonnet, J. A.; Higgins, C. P.; Field, J. A.; Helbling, D. E. Target and suspect screening integrated with machine learning to discover per- and polyfluoroalkyl substance source fingerprints. *Environ. Sci. Technol.* **2023**, *57* (38), 14351−14362.
(16) Stults, J. F.; Higgins, C. P.; Helbling, D. E. Integration of per- and polyfluoroalkyl substance (PFAS) fingerprints in fish with machine learning for PFAS source tracking in surface water. *Environ. Sci. Technol. Lett.* **2023**, *10* (11), 1052−1058.
(17) De Silva, A. O.; Armitage, J. M.; Bruton, T. A.; Dassuncao, C.; Heiger-Bernays, W.; Hu, X. C.; Kärrman, A.; Kelly, B.; Ng, C.; Robuck, A.; Sun, M.; Webster, T. F.; Sunderland, E. M. PFAS exposure pathways for humans and wildlife: A synthesis of current knowledge and key gaps in understanding. *Environ. Toxicol. Chem.* **2020**, *40* (3), 631−657.
(18) Rosato, I.; Bonato, T.; Fletcher, T.; Batzella, E.; Canova, C. Estimation of per- and polyfluoroalkyl substances (PFAS) half-lives in human studies: A systematic review and meta-analysis. *Environ. Res.* **2024**, *242*, 117743.
(19) Qin, W.; Escher, B. I.; Huchthausen, J.; Fu, Q.; Henneberger, L. Species difference? bovine, trout, and human plasma protein binding of per- and polyfluoroalkyl substances. *Environ. Sci. Technol.* **2024**, *58* (23), 9954−9966.
(20) U.S. Environmental Protection Agency (USEPA). *PFAS Master List of PFAS Substances (PFASMASTER)*, 2024.
(21) Wang, Z.; DeWitt, J. C.; Higgins, C. P.; Cousins, I. T. A never-ending story of per- and polyfluoroalkyl substances (PFASs)? *Environ. Sci. Technol.* **2017**, *51* (5), 2508−2518.
(22) Chen, X.; Lv, Z.; Yang, Y.; Yang, R.; Shan, G.; Zhu, L. Screening novel per- and polyfluoroalkyl substances in human blood based on nontarget analysis and underestimated potential health risks. *Environ. Sci. Technol.* **2024**, *58* (1), 150−159.
(23) Koelmel, J. P.; Lin, E. Z.; Parry, E.; Stelben, P.; Rennie, E. E.; Godri Pollitt, K. J. Novel perfluoroalkyl substances (PFAS) discovered in whole blood using automated non-targeted analysis of dried blood spots. *Sci. Total Environ.* **2023**, *883*, 163579.
(24) Yang, S.; Yi, H.; Ma, D.; Zeng, L. Non-targeted screening of per- and polyfluoroalkyl substances: Advanced methods, challenges, and environmental health. *Eco-Environ. Health* **2026**, *5* (1), 100219.
(25) Bangma, J.; Barry, K. M.; Fisher, C. M.; Genualdi, S.; Guillette, T. C.; Huset, C. A.; McCord, J.; Ng, B.; Place, B. J.; Reiner, J. L.; Robuck, A.; Rodowa, A. E. PFAS ghosts: How to identify, evaluate, and exorcise new and existing analytical interference. *Anal. Bioanal. Chem.* **2024**, *416* (8), 1777−1785.
(26) Plassmann, M. M.; Schmidt, M.; Brack, W.; Krauss, M. Detecting a wide range of environmental contaminants in human blood samples—combining QuEChERS with LC-MS and GC-MS methods. *Anal. Bioanal. Chem.* **2015**, *407* (23), 7047−7054.
(27) Koelmel, J. P.; Stelben, P.; McDonough, C. A.; Dukes, D. A.; Aristizabal-Henao, J. J.; Nason, S. L.; Li, Y.; Sternberg, S.; Lin, E.; Beckmann, M.; Williams, A. J.; Draper, J.; Finch, J. P.; Munk, J. K.; Deigl, C.; Rennie, E. E.; Bowden, J. A.; Godri Pollitt, K. J. FluoroMatch 2.0-making automated and comprehensive non-targeted PFAS annotation a reality. *Anal. Bioanal. Chem.* **2022**, *414* (3), 1201−1215.
(28) Takeda, H.; Matsuzawa, Y.; Takeuchi, M.; Takahashi, M.; Nishida, K.; Harayama, T.; Todoroki, Y.; Shimizu, K.; Sakamoto, N.; Oka, T.; Maekawa, M.; Chung, M. H.; Kurizaki, Y.; Kiuchi, S.; Tokiyoshi, K.; Buyantogtokh, B.; Kurata, M.; Kvasnička, A.; Takeda, U.; Uchino, H.; Hasegawa, M.; Miyamoto, J.; Tanabe, K.; Takeda, S.; Mori, T.; Kumakubo, R.; Tanaka, T.; Yoshino, T.; Okamoto, M.; Takahashi, H.; Arita, M.; Tsugawa, H. MS-DIAL 5 multimodal mass spectrometry data mining unveils lipidome complexities. *Nat. Commun.* **2024**, *15* (1), 9903.
(29) Schymanski, E. L.; Jeon, J.; Gulde, R.; Fenner, K.; Ruff, M.; Singer, H. P.; Hollender, J. Identifying small molecules via high resolution mass spectrometry: communicating confidence. *Environ. Sci. Technol.* **2014**, *48* (4), 2097−2098.
(30) Hansen, K. J.; Clemen, L. A.; Ellefson, M. E.; Johnson, H. O. Compound-specific, quantitative characterization of organic fluorochemicals in biological matrices. *Environ. Sci. Technol.* **2001**, *35* (4), 766−770.
(31) Dong, J.; Wang, N.-N.; Yao, Z.-J.; Zhang, L.; Cheng, Y.; Ouyang, D.; Lu, A.-P.; Cao, D.-S. ADMETlab: a platform for systematic ADMET evaluation based on a comprehensively collected ADMET database. *J. Cheminform.* **2018**, *10* (1), 29.
(32) Kuo, K.-Y.; Chen, Y.; Chuang, Y.; Lin, P.; Lin, Y.-J. Worldwide serum concentration-based probabilistic mixture risk assessment of perfluoroalkyl substances among pregnant women, infants, and children. *Ecotox. Environ. Safe.* **2023**, *268*, 115712.
(33) Amodei, D.; Egertson, J.; MacLean, B. X.; Johnson, R.; Merrihew, G. E.; Keller, A.; Marsh, D.; Vitek, O.; Mallick, P.; MacCoss, M. J. Improving precursor selectivity in data-independent acquisition using overlapping windows. *J. Am. Soc. Mass Spectrom.* **2019**, *30* (4), 669−684.
(34) Ismaiel, O. A.; Zhang, T.; Jenkins, R. G.; Karnes, H. T. Investigation of endogenous blood plasma phospholipids, cholesterol and glycerides that contribute to matrix effects in bioanalysis by liquid chromatography/mass spectrometry. *J. Chromatogr. B Analyt. Technol. Biomed. Life. Sci.* **2010**, *878* (31), 3303−3316.
(35) Wang, S.; Huang, J.; Yang, Y.; Hui, Y.; Ge, Y.; Larssen, T.; Yu, G.; Deng, S.; Wang, B.; Harman, C. First report of a Chinese PFOS alternative overlooked for 30 years: its toxicity, persistence, and presence in the environment. *Environ. Sci. Technol.* **2013**, *47* (18), 10163−10170.
(36) Li, L.; Yu, N.; Wang, X.; Shi, W.; Liu, H.; Zhang, X.; Yang, L.; Pan, B.; Yu, H.; Wei, S. Comprehensive exposure studies of per- and polyfluoroalkyl substances in the general population: Target, non-target screening, and toxicity prediction. *Environ. Sci. Technol.* **2022**, *56* (20), 14617−14626.
(37) Zhang, L.; Liu, X.; Huang, W.; Yang, Y.; Chen, D. A pilot study of bis-perfluoroalkyl sulfonimides in dust from E-waste and urban regions. *Environ. Sci. Technol. Lett.* **2024**, *11* (12), 1384−1390.
(38) Zhao, M.; Yao, Y.; Dong, X.; Baqar, M.; Fang, B.; Chen, H.; Sun, H. Nontarget identification of novel per- and polyfluoroalkyl substances (PFAS) in soils from an oil refinery in Southwestern China: A combined approach with TOP assay. *Environ. Sci. Technol.* **2023**, *57* (48), 20194−20205.
(39) Woodward, E. E.; Hladik, M. L.; Orlando, J. L.; Uychutin, M.; Gray, J. L.; Kolpin, D. W. Assessing pesticide application contributions of per- and polyfluoroalkyl substances (PFAS) in agricultural streams. *Environ. Sci. Technol. Lett.* **2026**, *13* (4), 586−593.
(40) Cui, Q.; Pan, Y.; Zhang, H.; Sheng, N.; Dai, J. Elevated concentrations of perfluorohexanesulfonate and other per- and polyfluoroalkyl substances in Baiyangdian Lake (China): Source characterization and exposure assessment. *Environ. Pollut.* **2018**, *241*, 684−691.
(41) Meng, L.; Song, B.; Lu, Y.; Lv, K.; Gao, W.; Wang, Y.; Jiang, G. The occurrence of per- and polyfluoroalkyl substances (PFASs) in fluoropolymer raw materials and products made in China. *J. Environ. Sci.* **2021**, *107*, 77−86.
(42) Liu, Z.; Lu, Y.; Song, X.; Jones, K.; Sweetman, A. J.; Johnson, A. C.; Zhang, M.; Lu, X.; Su, C. Multiple crop bioaccumulation and human exposure of perfluoroalkyl substances around a mega fluorochemical industrial park, China: Implication for planting optimization and food safety. *Environ. Int.* **2019**, *127*, 671−684.
(43) Rensmo, A.; Savvidou, E. K.; Cousins, I. T.; Hu, X.; Schellenberger, S.; Benskin, J. P. Lithium-ion battery recycling: A source of per- and polyfluoroalkyl substances (PFAS) to the environment? *Environ. Sci.: Process. Impacts* **2023**, *25* (6), 1015−1030.
(44) McMahen, R. L.; Strynar, M. J.; Dagnino, S.; Herr, D. W.; Moser, V. C.; Garantziotis, S.; Andersen, E. M.; Freeborn, D. L.; McMillan, L.; Lindstrom, A. B. Identification of fipronil metabolites by time-of-flight mass spectrometry for application in a human exposure study. *Environ. Int.* **2015**, *78*, 16−23.

https://doi.org/10.1021/acs.est.6c04503
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

dkhW_Ercy7Dm3ZL_9C3qfKAc485ysgAAALcwgbQGCSqGSIb3DQEHBqCBpjCBowIBADCBnQYJKoZIhvcNAQcBMB4GCWCGSAFlAwQBLjARBAzeOEHRnlXX0Cu-7KcCARCAclG5RdfPVdvp22M-8TCsi6U0YjYeNJTCUFY1-EdDO6ZwDkNKwueYqMZ-aB3-Bye80zGhXUTzEni3Fsh6pDjyf39fM2hol8v

(45) Ghanizadeh, H.; Mesarich, C. H.; Harrington, K. C. Molecular characteristics of the first case of haloxyfop-resistant Poa annua. *Sci. Rep.* **2020**, *10* (1), 4231.
(46) National Development and Reform Commission. *Industrial Structure Adjustment Guidance Catalogue*, 2011.
(47) Yao, J.; Pan, Y.; Sheng, N.; Su, Z.; Guo, Y.; Wang, J.; Dai, J. Novel perfluoroalkyl ether carboxylic acids (PFECAs) and sulfonic acids (PFESAs): Occurrence and association with serum biochemical parameters in residents living near a fluorochemical plant in China. *Environ. Sci. Technol.* **2020**, *54* (21), 13389−13398.
(48) Liu, M.; Zhang, G.; Meng, L.; Han, X.; Li, Y.; Shi, Y.; Li, A.; Turyk, M. E.; Zhang, Q.; Jiang, G. Associations between novel and legacy per- and polyfluoroalkyl substances in human serum and thyroid cancer: A case and healthy population in Shandong province, East China. *Environ. Sci. Technol.* **2022**, *56* (10), 6144−6151.
(49) Shi, Y.; Vestergren, R.; Xu, L.; Zhou, Z.; Li, C.; Liang, Y.; Cai, Y. Human exposure and elimination kinetics of chlorinated polyfluoroalkyl ether sulfonic acids (Cl-PFESAs). *Environ. Sci. Technol.* **2016**, *50* (5), 2396−2404.
(50) DeLuca, N. M.; Thomas, K.; Mullikin, A.; Slover, R.; Stanek, L. W.; Pilant, A. N.; Cohen Hubal, E. A. Geographic and demographic variability in serum PFAS concentrations for pregnant women in the United States. *J. Expo. Sci. Environ. Epidemiol.* **2023**, *33* (5), 710−724.
(51) Park, S. K.; Peng, Q.; Ding, N.; Mukherjee, B.; Harlow, S. D. Determinants of per- and polyfluoroalkyl substances (PFAS) in midlife women: Evidence of racial/ethnic and geographic differences in PFAS exposure. *Environ. Res.* **2019**, *175*, 186−199.
(52) Cho, C. R.; Lam, N. H.; Cho, B. M.; Kannan, K.; Cho, H. S. Concentration and correlations of perfluoroalkyl substances in whole blood among subjects from three different geographical areas in Korea. *Sci. Total Environ.* **2015**, *512*−*513*, 397−405.
(53) Bartolomé, M.; Gallego-Picó, A.; Cutanda, F.; Huetos, O.; Esteban, M.; Pérez-Gómez, B.; Castaño, A. Perfluorinated alkyl substances in Spanish adults: Geographical distribution and determinants of exposure. *Sci. Total Environ.* **2017**, *603*−*604*, 352−360.
(54) Meng, J.; Wang, T.; Song, S.; Wang, P.; Li, Q.; Zhou, Y.; Lu, Y. Tracing perfluoroalkyl substances (PFASs) in soils along the urbanizing coastal area of Bohai and Yellow Seas, China. *Environ. Pollut.* **2018**, *238*, 404−412.
(55) Yu, C. A comprehensive evaluation of industrial land intensive use in Hubei province based on typical industrial enterprises. *J. Resour. Ecol.* **2015**, *6* (3), 186−191.
(56) Xie, S.; Wang, T.; Liu, S.; Jones, K. C.; Sweetman, A. J.; Lu, Y. Industrial source identification and emission estimation of perfluorooctane sulfonate in China. *Environ. Int.* **2013**, *52*, 1−8.
(57) Yeung, L. W.; So, M. K.; Jiang, G.; Taniyasu, S.; Yamashita, N.; Song, M.; Wu, Y.; Li, J.; Giesy, J. P.; Guruge, K. S.; Lam, P. K. Perfluorooctanesulfonate and related fluorochemicals in human blood samples from China. *Environ. Sci. Technol.* **2006**, *40* (3), 715−720.
(58) Wang, J.; Pan, Y.; Cui, Q.; Yao, B.; Wang, J.; Dai, J. Penetration of PFASs across the blood cerebrospinal fluid barrier and its determinants in humans. *Environ. Sci. Technol.* **2018**, *52* (22), 13553−13561.
(59) Rajagukguk, I. M.; Hartanto, R.; Julian; Halim, R. Comparative analysis of XGBoost, random forest, and logistic regression for classifying Jakarta's air pollution index (ISPU). *Procedia Comput. Sci.* **2025**, *269*, 108−120.
(60) Liang, Z.; Lu, Y.; Cao, Z.; Huang, X.; Lei, H.; Li, J.; Wu, Z.; An, X.; Wang, P. Co-emissions of fluoride ion, fluorinated greenhouse gases, and per- and polyfluoroalkyl substances (PFAS) from different fluorochemical production processes. *Environ. Pollut.* **2024**, *360*, 124609.
(61) Dong, J.; Tsai, G.; Olivares, C. I. Prediction of 35 target per- and polyfluoroalkyl substances (PFASs) in california groundwater using multilabel semisupervised machine learning. *ACS EST Water* **2024**, *4* (3), 969−981.
(62) Moghadasi, R.; Mumberg, T.; Wanner, P. Spatial prediction of concentrations of per- and polyfluoroalkyl substances (PFAS) in European soils. *Environ. Sci. Technol. Lett.* **2023**, *10* (11), 1125−1129.
(63) DeLuca, N. M.; Mullikin, A.; Brumm, P.; Rappold, A. G.; Cohen Hubal, E. Using geospatial data and random forest to predict PFAS contamination in fish tissue in the Columbia river basin, United States. *Environ. Sci. Technol.* **2023**, *57* (37), 14024−14035.
(64) Liu, J.; Cui, Y.; Lu, M.; Lv, J.; Dong, L.; Guo, J.; Zhang, X.; Sun, Y.; Huang, Y.; Zhang, L. 6:2 Chlorinated polyfluoroalkyl ether sulfonate as perfluorooctanesulfonate alternative in the electroplating industry and the receiving environment. *Chemosphere* **2022**, *302*, 134719.

https://doi.org/10.1021/acs.est.6c04503
Environ. Sci. Technol. XXXX, XXX, XXX−XXX