Chemosphere 368 (2024) 143692

Contents lists available at ScienceDirect

Chemosphere

journal homepage: www.elsevier.com/locate/chemosphere

Applications of artificial intelligence for chemical analysis and monitoring of pharmaceutical and personal care products in water and wastewater: A review

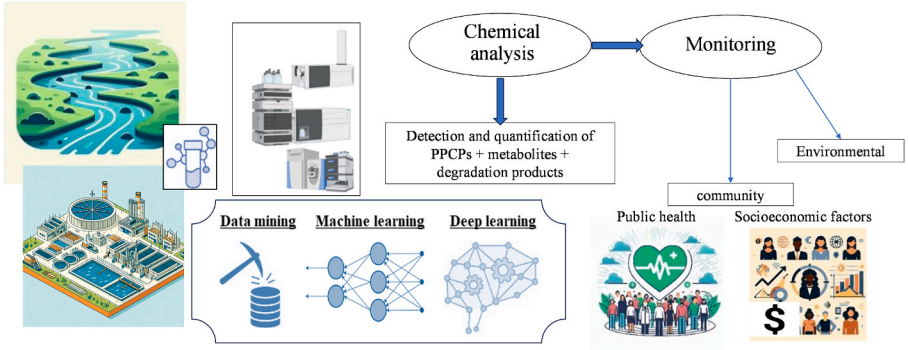
Babak Kavianpour [a], Farzad Piadeh [b,a], Mohammad Gheibi [c], Atiyeh Ardakanian [a], Kourosh Behzadian [a,d,*], Luiza C. Campos [d]

[a] School of Computing and Engineering, University of West London, St Mary's Rd, London W5 5RF, UK
[b] Centre for Engineering Research, School of Physics, Engineering and Computer Science, University of Hertfordshire, Hatfield, AL10 9AB, UK
[c] Institute for Nanomaterials, Advanced Technologies and Innovation, Technical University of Liberec, 46117, Liberec, Czech Republic
[d] Centre for Urban Sustainability and Resilience, Department of Civil, Environmental and Geomatic Engineering, University College London, London WC1E6BT, UK

HIGHLIGHTS

- AI assistance for chemical analysis and monitoring of PPCPs is reviewed.
- Typical ANNs are still dominant for retention time prediction of PPCPs.
- Spectroscopy and image processing need improvements for PPCPs real-time monitoring.
- AI-assisted CCS models can be applied for metabolites and transformation products.
- Data processing AI methods have significant potential for interpreting monitoring.

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:
Artificial intelligence
Contaminants of emerging concern
High-resolution mass spectrometry
PPCPs
Suspect and non-targeted screening
Quantitative structure retention relationship
Wastewater-based epidemiology

ABSTRACT

Specifying and interpreting the occurrence of emerging pollutants is essential for assessing treatment processes and plants, conducting wastewater-based epidemiology, and advancing environmental toxicology research. In recent years, artificial intelligence (AI) has been increasingly applied to enhance chemical analysis and monitoring of contaminants in environmental water and wastewater. However, their specific roles targeting pharmaceuticals and personal care products (PPCPs) have not been reviewed sufficiently. This review aims to narrow the gap by highlighting, scoping, and discussing the incorporation of AI during the detection and quantification of PPCPs when utilising chemical analysis equipment and interpreting their monitoring data for the first time.

In the chemical analysis of PPCPs, AI-assisted prediction of chromatographic retention times and collision cross-sections (CCS) in suspect and non-target screenings using high-resolution mass spectrometry (HRMS) enhances detection confidence, reduces analysis time, and lowers costs. AI also aids in interpreting spectroscopic analysis results. However, this approach still cannot be applied in all matrices, as it offers lower sensitivity than liquid chromatography coupled with tandem or HRMS.

* Corresponding author. School of Computing and Engineering, University of West London, St Mary's Rd, London W5 5RF, UK.
E-mail address: kourosh.behzadian@uwl.ac.uk (K. Behzadian).

https://doi.org/10.1016/j.chemosphere.2024.143692
Received 28 January 2024; Received in revised form 15 September 2024; Accepted 4 November 2024
Available online 15 November 2024

For the interpretation of monitoring of PPCPs, unsupervised AI methods have recently presented the capacity to survey regional or national community health and socioeconomic factors. Nevertheless, as a challenge, long-term monitoring data sources are not given in the literature, and more comparative AI studies are needed for both chemical analysis and monitoring. Finally, AI assistance anticipates more frequent applications of CCS prediction to enhance detection confidence and the use of AI methods in data processing for wastewater-based epidemiology and community health surveillance.

Abbreviations

AI	Artificial intelligence
ANFIS	Adaptive Neuro Fuzzy Inference System
ANN	Artificial neural network
BRANN	Bayesian regularised artificial neural network
BRR	Bayesian ridge regression
CatBoost	Categorical Boosting
CCD-RSM	Central composite design-response surface methodology
CCS	Collision cross-section
CNN	Convolutional neural network
DLM	Deep learning regression model
DNN	Deep neural network
DOC	Dissolved organic carbon
DOE	Design of experiment
ELM	Extreme learning machine
EPs	Emerging pollutants
FFNN	Feedforward neural network
FTIR	Fourier transform infrared
GA	Genetic algorithm
GNN	Graph neural network
GRNN	Generalised regression neural network
HCA	Hierarchical cluster analysis
HILIC	Hydrophilic interaction liquid chromatography
HPLC	High-performance liquid chromatography
HRMS	High-resolution mass spectrometry
KNN	k-nearest neighbourhood
Kow	Octanol-water partition coefficient
LC	Liquid chromatography
LC-MS	Liquid chromatography-mass spectrometry
LC-MS/MS	Liquid chromatography-tandem mass spectrometry
LC-QToF	Liquid chromatography Quadrupole time of flight
LIRFS	Laser-induced Raman and fluorescence spectroscopy
MAE	Mean absolute error
MARS	Multiple adaptive regression splines
MLP	Multi-layer perceptron
MLR	Multi-linear regression
MSE	Average mean square error
OLS	Ordinary least square
PCA	Principal component analysis
PLS	Partial least square
PNN	Probabilistic neural network
POCIS	Polar organic chemical integrative passive sampler
PPCPs	Pharmaceuticals and personal care products
QSRR	Quantitative Structure Retention Relationships
RBF	Radial basis function
RF	Random forest
RMSE	Root mean square error
RPLC	Reversed-phase liquid chromatography
RT	Retention time
SDG	Sustainable development goals
SMILES	Simplified Molecular Input Line Entry System
SVM	Support vector machine
WBE	Wastewater-based epidemiology
WQI	Water quality indicators
WWTP	Wastewater treatment plant
XGBoost	Extreme Gradient Boosting

1. Introduction

Pharmaceuticals and personal care products (PPCPs) have undoubtedly improved human well-being through direct consumption in agriculture or daily medical purposes (Lindim et al., 2016). However, they are categorised as emerging pollutants (EPs), contaminants of emerging concern, or environmental micro-pollutants (Cai et al., 2015). These pollutants are widespread in the environment, found in groundwater, surface water, wastewater treatment plants (WWTP) within both the influent and effluent and even in tap water and drinking water (Fick et al., 2009; Hapeshi et al., 2015). Recently, studies revealed that the presence of pharmaceutical pollution in rivers poses a global threat to the environment and human health (Wilkinson et al., 2022). For example, fish toxicity at effluent concentrations of Diclofenac found in wastewater treatment plants (WWTP) (Schwaiger et al., 2004). Additionally, the presence of pharmaceuticals, specifically antibiotics, can adversely affect bacteria related to activated sludge and nitrifying processes (Kim and Aga, 2007), algal biomass and community structures (Wilson et al., 2003), and some bacteria in the food chain (Oaks et al., 2004). Furthermore, the serious threats of antimicrobial resistance and endocrine disruption to humans provide better insight into the implications of environmental contamination (Murray et al., 2022; Puri et al., 2022), which are expected to be documented and periodically updated in a new database (The PREMIER Project, 2024). Since WWTPs are not primarily designed to remove EPs (Patel et al., 2019), effluent discharges are the main pathway of PPCP residues into the environment. This issue can be intensified, especially during heavy rain and combined sewage overflows (Munro et al., 2019; Yun et al., 2023). Subsequently, updated watchlists, monitoring campaigns and chemical investigation programs are needed to reach regulations for mitigating the implications (Gardner et al., 2022).

Regarding the contexts of chemical analysis of PPCPs, reliable equipment is essential for detecting and quantifying PPCPs in complex water and wastewater environmental samples. Although there have been breakthroughs in developing sensors for selective detection and quantification of PPCPs (Mao et al., 2021), chromatography coupled with mass spectrometry is the dominant approach, followed by spectroscopy analysers (Hameedat et al., 2022; Mohammadpoor et al., 2019). To be more specific, the most reliable analyses can be carried out via liquid chromatography coupled with tandem mass spectrometry (LC-MS/MS), including Triple Quadrupole for targeted analysis with the utmost sensitivity and high-resolution mass spectrometry (HRMS), such as Quadrupole time-of-flight (LC-QToF) or LC-MS Orbitrap for targeted, suspect, non-targeted, and retrospective analyses of compounds (Belay et al., 2022; Egli et al., 2023; Gómez-Canela et al., 2021; Van Hoi et al., 2021).

From a broader perspective, monitoring PPCPs, their metabolites,

transport or degradation byproducts in the environment and wastewater, and analysing and interpreting the monitoring data form the basis of various research areas. These directions include environmental toxicology and risk assessment, defining chemical prioritisation or biomarkers, sewage leak detection, wastewater-based epidemiology (WBE), and studies on drugs of abuse for community consumption, as well as spatial and temporal contamination analyses (Hernández et al., 2018; Kasprzyk-Hordern et al., 2023a; Tran et al., 2019; Wu et al., 2023).

Other contexts involving PPCPs but possibly relying on their chemical analysis and monitoring can include antimicrobial resistance in water and wastewater (Foroughi et al., 2024), water treatment (Chowdhury and Karanfil, 2024; Li et al., 2021), wastewater treatment (Zhang et al., 2023), and advanced treatment processes to remove pharmaceuticals (Serna-Carrizales et al., 2024). AI can optimise processes, model removal efficiencies, and predict outcomes in various treatment processes, such as ozonation (Heidari et al., 2022; Pelalak et al., 2020), photodegradation (Deylami et al., 2023; Hosseini et al., 2022; Sheydaei et al., 2022), electrochemical oxidation (Arab et al., 2022; Gholami Shirkoohi et al., 2022), Fenton-like process (Salari et al., 2022), biological degradation (Fu et al., 2023; Mojiri et al., 2022), and, more dominantly, adsorption (Abbasi et al., 2022; Ahmadi Azqhandi et al., 2022; Alam et al., 2022; Gholami et al., 2023, 2024; Naderi et al., 2022; Salehi Nasab et al., 2024). Greener treatment technologies can also offer a context for AI assistance. These technologies can include optimising the physico-chemical processes for sludge extract reuse (Badawi and Hassan, 2024), using agriculture waste-based adsorbents in treatment (Afolabi et al., 2020; Badawi et al., 2024a), eco-friendly filtration of greywater (Hassan et al., 2024), and optimising the coagulation-flocculation processes of slaughterhouse wastewater (Badawi et al., 2024b).

There are some reviews carried out recently for the contexts of identifying and monitoring PPCPs in water and wastewater. For example, Singh et al. (2023) recently compiled methodological features for AI applications in chromatography, though the specific matrices and types of analytes were not clearly defined, and the review was not tailored to any specific group of pollutants. Haddad et al. (2021) provided a comprehensive review of retention time (RT) prediction methodologies in liquid chromatography (LC), but with limited AI applications and their analyses. Notably, since Barron and McEneff (2016) who examined the methodology of the RT of EPs using artificial neural network (ANN) with a clear classification of EPs and the matrix of the samples, there is no review available for the AI-assisted chemical analysis of PPCPs. Therefore, the recent AI applications when using other chemical analysers and the progression of AI assistance for RT prediction or other parameters in suspect and non-targeted screening have remained unreviewed. The research gap widens further in the context of AI-assisted interpretation of environmental monitoring for PPCPs, as no reviews exist in this area, to the best of the authors' knowledge. Accordingly, studying where (the contexts) and how (the roles) of typical and more advanced AI methods have been implemented over the past decade in both interconnected areas -chemical analysis of PPCPs and their monitoring in water and wastewater matrices- can significantly bridge the identified gaps.

This review highlights, scopes and discusses the contexts and roles of AI techniques implemented in the chemical analysis and interpretation of the monitoring of PPCPs in environmental water and wastewater over the past decade. After conducting a concise, systematic scientometric analysis to capture trends and variations in AI methods, the contributions of AI in the chemical analysis of PPCPs are reviewed and comparatively analysed. The key research questions addressed are how AI has been applied for chromatographic condition optimisation and spectroscopic analyses, and whether they apply to all occurrence levels of PPCPs; how AI-driven models for the prediction of RT and collision cross-section (CCS) are developed and evolved for suspect and non-targeted screenings of PPCPs, followed by a comparative discussion on their reliability and challenges. The reviewed material in the monitoring of PPCPs focuses on feasibility studies for real-time monitoring, prediction of the occurrence, and classification and interpretation of monitoring data applicable to environmental, health, and WBE investigations. Finally, the future research perspectives on applying AI for chemical analysis and monitoring of PPCPs are presented.

2. Research design, bibliometrics and trends

This research is based on peer-reviewed scientific papers from the last decade (2014–2024) that study, implement, integrate, or cite AI methods for PPCPs in wastewater, water, aqueous, or lab samples, focusing solely on the chemical analysis and monitoring of PPCPs. Applications of AI in other domains, such as the treatment of PPCPs, are not within the scope of this study. However, the chemical analysis and monitoring of these compounds can be employed to assess the performance of treatment processes and new technologies for removing PPCPs.

Fig. 1 shows the flow diagram illustrating the search and screening of references. The research works were initially reviewed using the Scopus database search engine, following the guidelines suggested by (Moher et al., 2009). As illustrated in Fig. 1 in the second step (S2), the broader term EP was added to the search criteria to ensure that related papers that did not include the term "PPCPs" in their keywords, titles, or abstracts were not overlooked while searching "water" and "wastewater" only in the title of references modified the screening in S2. The investigation for AI methods included keywords such as artificial, learning, neural, mining, and statistical through the titles, abstracts, or keywords of the journal papers, representing AI, ML, neural networks, data mining, or statistical methods with the potential of coupling with AI.

Following language refinement and publication period adjustments, a more comprehensive review was conducted through the third step (S3). This involved mannually examining more recent papers that cited the key references from the early stage of the time span, such as Bade et al. (2015a) and Barron and McEneff (2016). After an in-depth review and selection of papers on both chemical analysis and monitoring aspects in S4 and S5, the contributing keywords were extracted, as shown in Fig. 1. These keywords provide insights for further studies aligned with this work.

2.1. Bibliometric analysis

A bibliometric analysis, shown in Fig. 2, examined the geographical distribution of case studies, clustering, and keyword timelines. The results highlight that 18 countries, led by China and the UK, are active in AI-based research on EPs. Despite the global importance of addressing EPs, particularly in Europe, the Americas, Australia, and Asia (Beek et al., 2016; Hawash et al., 2023), AI applications for predicting concentration, RT, or monitoring systems remain in the early stages. This emerging field offers significant potential, with room for further exploration in response to the cross-boundary challenges posed by EPs. Researchers and monitoring campaigns should embrace AI's transformative capabilities to accelerate progress.

The bibliometric analysis, illustrated in Fig. 2c–d, was conducted using VOSviewer based on the co-occurrence of key terms in keywords, titles, and abstracts. This approach reveals the interconnectedness of keywords and identifies core research themes. Cluster analysis (Fig. 2c) groups keywords into classes, with node size indicating co-occurrence frequency, links representing co-reference, and colors marking clusters. Three main clusters are identified: (1) Model inputs (red) – includes EP types and high-performance liquid chromatography (HPLC) data; (2) AI-based data processing (green) – focuses on environmental data processing using techniques like PLS, OLS, and KNN; (3) AI-based modelling (blue) – links ANNs and ML models to outputs like concentration prediction and RT forecasting.

Fig. 2d presents a timeline analysis that maps keywords over the past decade, revealing emerging trends and shifts in key areas. Most topics in

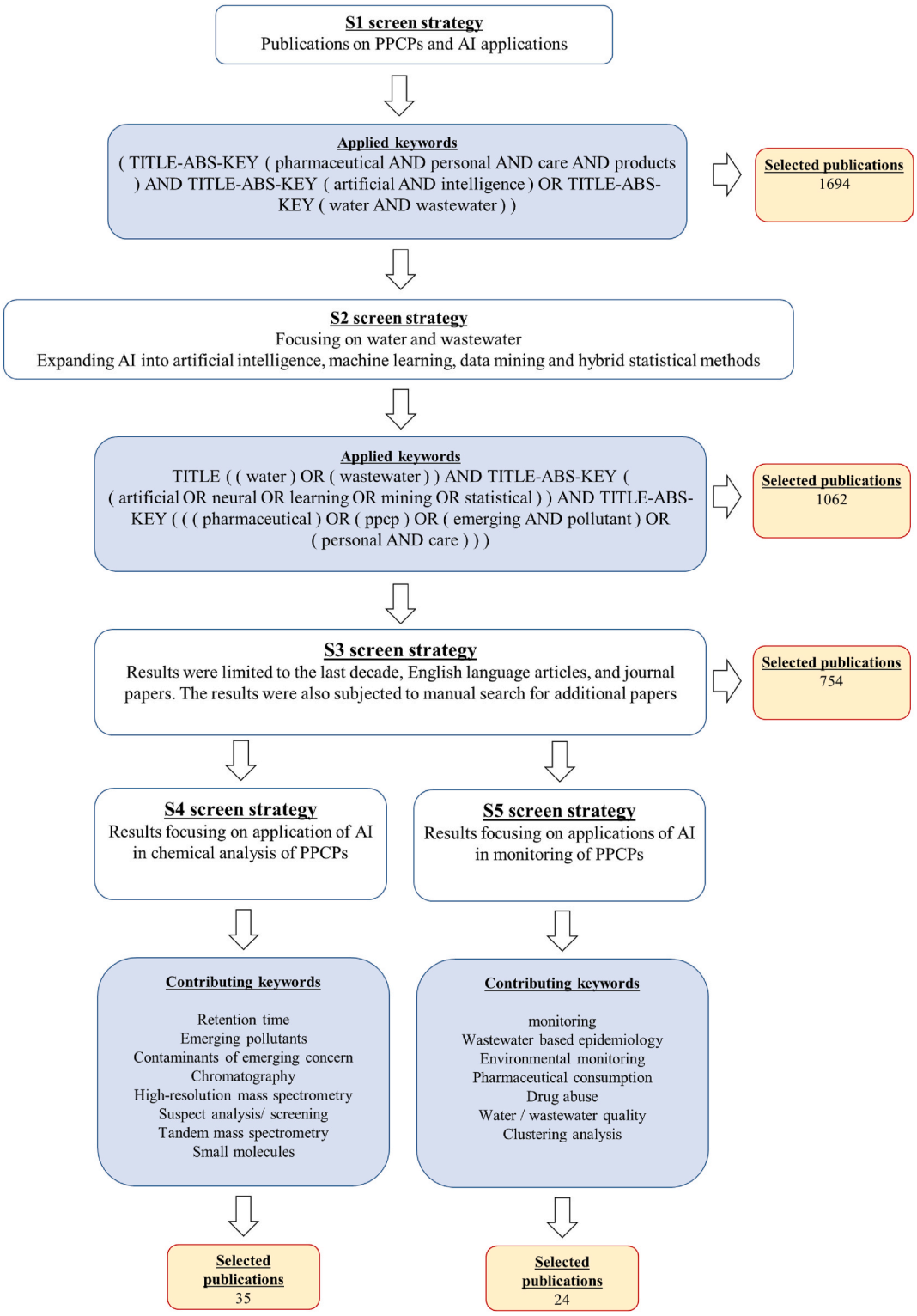


Fig. 1. Flow diagram of searching and screening of the key references.

this field are relatively recent, reflecting its dynamic nature with potentially more focus in the future. Environmental monitoring of pharmaceuticals and micropollutants has become a topical subject, emphasising the need to specify the occurrence of PPCPs in environmental water and wastewater matrices. Additionally, the growing focus on advanced ML techniques signals a shift toward more sophisticated algorithms or approaches in the chemical analysis, detection, and monitoring of PPCPs.

2.2. Modelling trends

Fig. 3 presents a scientometric analysis of AI research trends. As shown in Fig. 3a, AI-driven chemical analysis and prediction for PPCPs began around 2015, with environmental monitoring of pharmaceuticals documented from 2019. The growing interest in the studies focused on chemical analysis of real environmental samples highlights a growing interest in wastewater and surface water monitoring campaigns and case

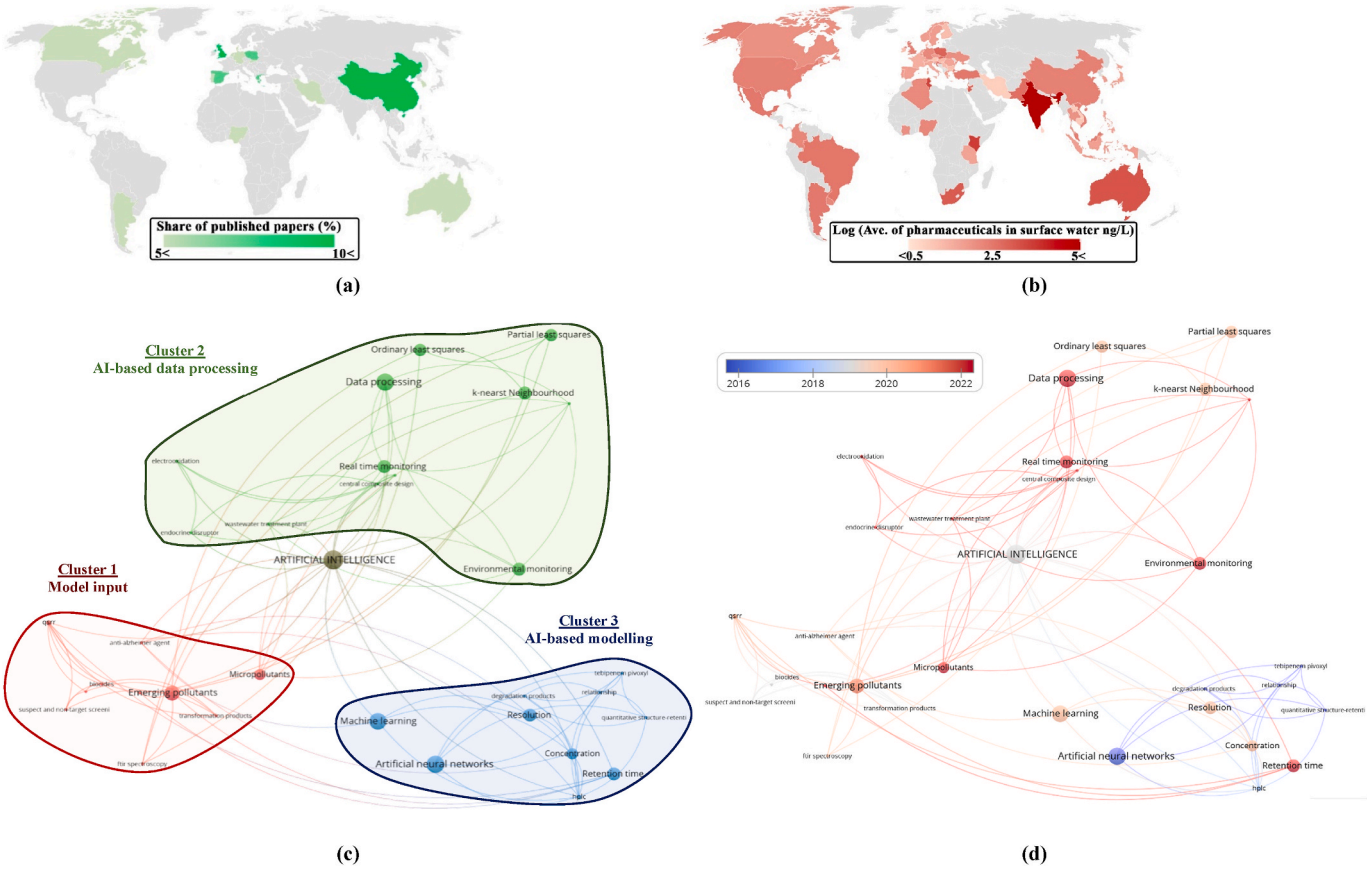


Fig. 2. Bibliometric analysis for the selected papers: **a)** geographical distribution in comparison to documented pharmaceutical concentrations, **b)** Occurrence of PPCPs in surface waters c) cluster of keywords, **d)** timeline of keywords.

studies. This might suggest more available monitoring data and a larger size of the datasets to interpret monitoring data for upcoming endeavours.

Fig. 3b provides insights into the intensity of different AI tools used over the years. ML programming emerges as a major research focus, with simple feedforward neural network (FFNN) models prioritised, as also illustrated in Fig. 3c. This observation indicates that the application of AI in this context predominantly relies on simpler models, highlighting a research gap in employing advanced models, such as recurrent or graph neural networks (GNN) or physics-informed modelling (Li et al., 2021; Ye et al., 2022). This is further supported by examining deep learning usage, revealing that only 2% of the selected papers applied models such as convolutional neural networks (CNNs). However, recent trends, as highlighted in Fig. 3b, indicate a shift from ML models towards data mining applications, with a focus on ensemble models employing various techniques such as random forest (RF) and support vector machines (SVM).

In terms of optimisation tools, the prevalent use of Genetic Algorithm (GA) for model tuning suggests the potential for exploring other nature-based techniques, especially to enhance model accuracy or reduce required data types and quantities. Fig. 3 also highlights data mining techniques led by SVM, KNN, and Hierarchical Cluster Analysis (HCA), recently applied for data preprocessing tasks such as data classification or relationship analysis. There is considerable scope for incorporating additional models in this context, especially for using ensemble modelling of weak learners or more advanced ANN models such as GRNN, ANFIS and ELM Net.

3. Chemical analysis of PPCPs in water and wastewater

AI assistance can be applied to finding the optimised chromatographic conditions, such as the suitable mobile phase and its content, buffer content, and the pH of the mobile phase to improve the signal-to-noise ratio and quantification. In practice, such methods can be applied for limited and specific PPCPs when the superior equipment in sensitivity and resolution, i.e. LC-MS/MS and LC-HRMS, are not available (Abba et al., 2020; Haddad et al., 2021; Taraji et al., 2017). Furthermore, AI assistance can be used for the quantitative concentration measurement of PPCPs via spectroscopy (Mohammadpoor et al., 2019; Quintelas et al., 2020).

Nevertheless, the majority of AI-assisted studies for the chemical analysis of PPCPs focus on the Quantitative Structure-Retention Relationships (QSRR) modelling approach to enhance the confidence of the detection and save time and financial resources. It should be noted that the other mathematical QSRR methodologies are comprehensively reviewed by Haddad et al. (2021).

According to the benchmark of confidence for HRMS (Schymanski et al., 2014), the critical step in suspect or non-targeted analysis is to tentatively yet reliably ascertain the existence of compounds using HRMS after matching the exact mass ± tolerance and the fragmentation with spectral libraries. Ideally, the detection can then be confirmed by comparing the results of the HRMS for the water sample with the reference standard substance of each compound.

However, challenges arise when the matrix of the water sample is complex, which is the case for environmental water and wastewater. This leads to many peaks on the chromatogram, which requires spending a lot of time matching the fragmentations of the candidate chemicals with close mass. Also, when reference materials are not readily available, especially in the case of transformation or degradation products or when it is not financially feasible to purchase pure chemicals for many compounds in the first place. AI applications for the prediction of RT and CCS (Mollerup et al., 2018) can significantly bolster

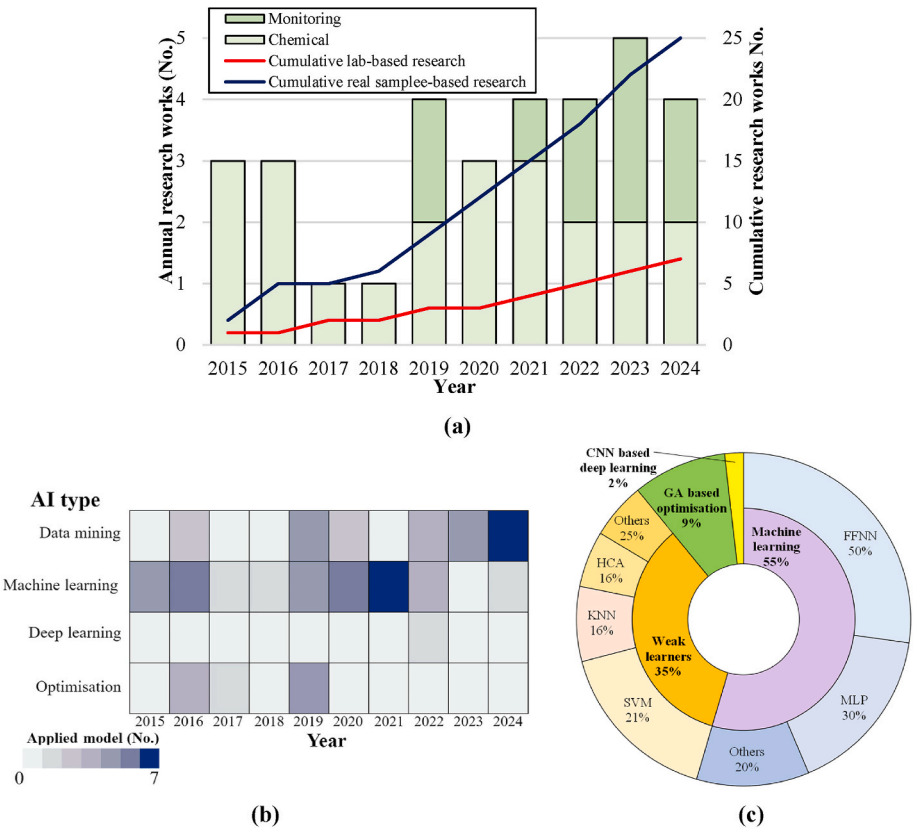


Fig. 3. Dashboard of applied AI for chemical analysis and monitoring of PPCPs (2024 is not reflected for the entire year): **a)** trends per sample types, **b)** heatmap of model types, **c)** distribution of applied algorithms.

confidence in the analysis and address the mentioned challenges (Bade et al., 2015a). It should be mentioned that the time and financial resources saved due to AI-assisted models can be per case and is dependent on the list of chemicals for the detection, the user-friendliness of the software tools of the analysers and the performance of their operators. The Network of Reference Laboratories, Research Centres and Related Organisations for Monitoring of Emerging Environmental Substances (NORMAN) has recently published valuable guidelines (Dulio et al., 2024; Hollender et al., 2023) and databases (Norman, 2024) applicable to environmental screenings. Table 1 shows the characteristics of the reviewed models for the chemical analysis of PPCP in water and wastewater, which are scoped, highlighted and discussed in the following subsections.

3.1. Quantification of PPCPs

As discussed earlier, LC-Triple Quad MS is the most reliable device for quantifying PPCPs in targeted analysis, in which analytes are selected, and their standard reference materials are readily available. However, when other types of analysers are employed or to optimise the chromatography conditions, AI assistance can contribute to quantifying the occurrence of PPCPs in water and wastewater samples.

AI methods can facilitate the systematic optimisation of chromatography (Abbasi et al., 2022) to achieve better resolution, consume less material, and run analyses faster with fewer trial-and-error attempts (Haddad et al., 2021). In this regard, Taraji et al. (2017) developed a hybrid statistical-AI QSRR model for achieving optimised pH, mobile phase (acetonitrile content), and salt concentration by Hydrophilic interaction liquid chromatography (HILIC) for analysing different groups of compounds.

This model was applied to 3 groups of pharmaceuticals: β-adrenergic agonists, benzoic acids, and nucleocides by 17 experiments designed by Central Composite Design – Response Surface Methodology (CCD-RSM). For descriptor selection, a new clustering method was developed based on compound classification and structural similarities, followed by the GA-PLS method to predict RTs and evaluate the chromatography selectivity factor as the index of desirability ($\leq$1.15), with the dataset sizes provided in Table 1.

A Monte Carlo simulation assessed the risk of prediction uncertainty based on the selectivity factor, allowing the model to identify different optimal conditions for the 3 types of pharmaceuticals. For example, β-adrenergic agonists had better RT predictions at medium pH and the highest acetonitrile content in the buffer. This methodology can be implemented to find optimised conditions for analysers when examining groups of PPCPs.

Another combined statistical-AI model was developed by Carabajal et al. (2021) to optimise LC conditions and quantify 6 veterinary drugs from kennel effluent water. The fractional factorial design method was used to specify significant factors, followed by CCD with statistical least squares and ANN to optimise the conditions, as detailed in Table 1. While several drugs were not detected, likely due to the sensitivity of the LC (LOD at 2–8.5 ng/mL), the results suggested deviations from legal environmental thresholds for several compounds. This study has potential applications for quantifying PPCPs with high occurrence levels in wastewater when the analyser is not as sensitive as tandem MS.

According to Abba et al. (2020), most methods for optimising chromatography conditions revolve around ANN with traditional FFNN and classic multi-linear regression (MLR). However, the enhancement of the Adaptive Neuro-Fuzzy Inference System (ANFIS) was studied for the first time to optimise the response surface modelling of the retention factor (Inés Toral et al., 2002) of HPLC for pharmaceuticals: methyclothiazide and amiloride, by considering the mobile phase (methanol) content and pH of the separation to achieve the desired retention factor (>2 and <10). The results showed that ANFIS was slightly more

Table 1
Reviewed models and their features for chemical analysis of PPCPs for/in water and wastewater.

Method Analyser	Datasets details	Models/layers	Applied for	Output	Selected Input	Sampling Country	External Validation Matrix	External validation Analytes	Reference
Agilent HPLC-QToF	1051	RF, XGBoost, SVR, ANN	Prediction	RT	10 descriptors	China	WWTP	70 Pharmaceuticals	Song et al. (2024)
Theoretical	490 optimised to 391	MLR, SVM, DNN	Prediction	Relative RT	10 descriptors	NA	NA	NA	(Q. He et al., 2022)
Waters UPLC-IMS-QToF	477 $RPLC^+$/169 $RPLC^-$/249 sodium adducts	MARS	Prediction	RT and CCS	10 descriptors for RT 7 for CCS ± ionisation 10 for CSS	NA	Reference standard material	25 EPs (5 PPCPs)	Celma et al. (2022)
Agilent HPLC-DAD	20 experiments	DOE (CCD-RSM) + ANN-RBF	Optimisation	Max extraction recovery of analytes	3 chromatographic conditions (2 solvents volume + vortex time)	Argentina	Wetland	6 Veterinary drugs	Carabajal et al. (2021)
Theoretical	80038 METLIN (SMRT) databank	GNN	Prediction	RT	48 molecular graph data	NA	NA	100 for model comparison	Yang et al. (2021b)
Agilent LC-QToF	70:15:15	MLP-ANN	Prediction	RT	16 descriptors	UK	Surface water	8 metabolites, 34 Pharma	Richardson et al. (2021)
Shimadzu HPLC	16 experiments	SVM, HW, MLP, SVM-E, SAE	Optimisation	Max RT+ Resolution	3 chromatographic conditions: Mobile phase (2 solvent) Column temperature	NA	Reference standard material	mixtures of a drug (alti-Alzheimer) and its derivatives	Ghali et al. (2020)
Shimadzu UHPLC ABB FTIR spectrometer	435	KNN-OLS, KNN-PLS	Prediction	Concentration	(varied for each pollutant) 4 to 10 components of absorbance curve at 200-14000 cm-1	Portugal	WWTP	8 EPs (7 pharmaceuticals)	Quintelas et al. (2020)
Waters HPLC	15 experiments	ANN, ANFIS, MLR	prediction/ Optimisation	Retention factor	Mobile phase content, pH	NA	Reference standard material	2 pharmaceuticals (Amiloride Methyclothiazide)	Abba et al. (2020)
Varian Cary UV–Vis Spectrometer	14:3:3	BRANN	Prediction	Concentration	6 components of absorbance curves at 350–750 nm	NA	Reference standard material	3 Pharmaceuticals	Mohammadpoor et al. (2019)
Thermo Fisher UHPLC QToF (Bruker)	1830 for $RPLC^+$ 308 for $RPLC^-$ 682 for $HILIC^+$	GA-MLR, GA-SVM	Prediction	RT	5 descriptors for RPLC+ 8 for RPLC-7 for HILIC+	Greece	WWTP influent effluent and sludge	10 Ozonation byproducts 28 biocides (21 PPCPs)	Aalizadeh et al. (2019)
Waters UHPLC-QToF Waters UHPLC–IMS QToF	225:48:48	MLP-ANN	Prediction	RT and CCS	24 descriptors	NA	Reference standard material	36 pharmaceuticals	Mollerup et al. (2018)
Agilent HPLC-DAD -QToF	9 drug impurities	GA + PS-ANN SPS-ANN	Prediction	RT	5 descriptors	NA	pharmaceutical (industry)	12 new impurities of another drug in the same group	Mizera et al. (2017)
POCIS passive sampling	45:14:14	MLP-GA, MLP-ANN	Prediction	Sampling rate	ANN 21 descriptors GA 24 descriptors	NA	Reference standard material	6 Pharmaceuticals	Miller et al. (2016)
Thermo Fisher HPLC	646 (17 condition × 50 Pharma)	DOE (CCD-RSM) + (GA - PLS)	Optimisation	RT	Mobile phase content, pH, buffer content	NA	Reference standard material	12 pharmaceuticals	Taraji et al. (2017)
Thermo Fisher UHPLC QToF (Bruker)	528 positive mode 298 negative mode	KNN-GA-SVM, KNN-GA-MLR, KNN-GA-ANN	Prediction	RT	7 descriptors	NA	Reference standard material	100 pharmaceuticals	Aalizadeh et al. (2016)
Waters UHPLC-QToF	344:100:100	MLP-ANN	Prediction	RT	16 descriptors	Colombia	WWTP and surface water	26 Pharmaceuticals (including 9 metabolites)	Bade et al. (2015a)
Thermo Fisher UHPLC - Qrbitrap	166	GRNN	Prediction	RT	17 descriptors	UK	WWTP	37 pharmaceuticals	Munro et al. (2015)
Waters UHPLC-QToF	625 chemicals (over 200 pharmaceuticals)	Linear regression	Prediction	RT	Kow	Spain, Germany, Colombia	WWTP and surface water	30 EPs (16 Pharma 3 drugs of abuse)	Bade et al. (2015a)

accurate than ANN, but both were superior to MLR, especially for amiloride.

For the assessment of ensemble AI methods, Ghali et al. (2020) investigated a more advanced AI methodology for optimising HPLC mobile phase content (2 solvents) and column temperature to achieve the best resolution and maximum RT for an anti-Alzheimer drug and its derivatives. The study applied multi-layer perceptron (MLP), Hammerstein-Wiener (HW), and SVM, as well as double ensemble methods: simple average ensemble (SAE) and non-linear ensemble SVM. These methods aimed to address overfitting and enhance the generalisation of AI methods by combining the outputs of SVM, HW and MLP models. The comparative performance showed SVM > MLP > HW in both resolution and RTmax predictions, with the ensemble methods achieving more accuracy than the individual AI methods. This research may also be noteworthy for analysing water samples for drugs and their partially degraded forms or metabolites analogous to drug derivatives.

Commonly ultraviolet–visible (UV–Vis) spectrometry AI-assisted modelling (Benjathapanun et al., 1998; Guo et al., 2020) is used for analysing conventional water quality parameters with some restrictions for the simultaneous detection of overlapping analytes. Nevertheless, Mohammadpoor et al. (2019) assessed the applicability of AI for the simultaneous detection and quantification of Loratadine, Naproxen and Diclofenac in pure water. They developed a Bayesian regularised ANN (BRANN) using Principal Component Analysis (PCA) to characterise inputs, as shown in Table 1. The model outperformed ANNs with Levenberg-Marquardt, scaled conjugate gradient, and resilient back-propagation techniques.

Quintelas et al. (2020) utilised Fourier transform infrared (FTIR) spectroscopy to analyse 8 EPs, including ibuprofen, carbamazepine, sulfamethoxazole, desloratadine, paracetamol, as well as β-estradiol and ethynylestradiol hormones, using samples from a WWTP spiked with pollutant-rich solutions. For dataset creation, 4 FTIR preprocessing methods and a single without preprocessing (raw) were used, generating 5 datasets for each pollutant. This resulted in 5 wavelength-absorbance curves indicating concentration. The KNN, followed by PLS and OLS, were developed for predicting the concentrations. The results showed relatively similar performance for PLS and OLS.

3.2. Tentative detection of PPCPs in suspect and non-targeted analyses

3.2.1. Overview of modelling procedure

The process of AI-assisted predictive modelling of RT and CCS for suspect and non-targeted screenings is demonstrated in Fig. 4. The initial step involves selecting a comprehensive dataset of PPCPs along with their reported measured RT or CCS values, such as the METLIN small molecule dataset for machine learning RT predictions (Domingo-Almenara et al., 2019). After obtaining the line notation of compounds, commonly using the simplified molecular-input line-entry system (SMILES), the molecular attributes of the compounds, including shape, size, geometry, class, and toxicity, can be calculated using various platforms such as RDKit: Open-source cheminformatics, Mordred (Moriwaki et al., 2018) or VCCLab (Tetko et al., 2005; VCCLAB, 2005).

Next, molecular descriptors are assessed and prioritised based on their impact and significance on RT or CSS using various statistical or AI methods such as MLR (Q. He et al., 2022), ANN (Barron and McEneff, 2016) and GA (Aalizadeh et al., 2019). The selected descriptors may vary between models. For example, He et al. (2022) identified lipophilicity, fractional polar, and negative van der Waals surface areas as the most essential descriptors, while (Barron and McEneff, 2016) highlighted lipophilicity, especially logD, and the number of oxygen and carbon atoms by analysing 10 different models. Therefore, the chosen descriptors, serving as inputs for the AI models, depend heavily on the selected dataset of PPCPs and their molecular characteristics. During descriptor selection and AI-model development and validation steps, various splitting techniques can be applied for training, testing, and validating the suitable predictive algorithms (Haddad et al., 2021), such as random splitting (Miller et al., 2013; Mollerup et al., 2018; Munro et al., 2015; Richardson et al., 2021; Yang et al. et al., 2021), k-fold cross-validation (Abba et al., 2020; Celma et al., 2022; Ghali et al., 2020), or PCA and kNN (Aalizadeh et al., 2016, 2019).

As shown in Fig. 4, a crucial evaluation step is the external validation and assessment of the predictive models with new PPCPs not included in the initial step, using analysers equipped with HRMS. This can be done by comparing predicted and measured RT and CCS values using reference standard materials of PPCPs in water or wastewater, preferably with real water and wastewater samples and a selection of standard materials. Thus, when a model demonstrates robustness and high performance in external tests, in addition to the initial training and testing, it can be used with higher confidence for suspect and non-targeted screenings.

A common approach is to exclude outlier compounds to enhance the robustness of the models. This can be done by considering a percentile of RT error, such as the 75th percentile (Munro et al., 2015), or by excluding the highest and lowest, e.g. 10% of compounds based on the median error (Barron and McEneff, 2016). While this method reduces the size of the datasets, it boosts the model's confidence. He et al. (2022) used Cook's distance method for detecting the outliers, and Aalizadeh et al. (2016, 2019) proposed a novel graphical visualisation of outliers for understanding the applicability domain of predictions. This method was modified based on the Monte Carlo sampling method in 2019, with open-source codes using LC-QToF for a large number of EPs.

3.2.2. ANN assistance progression

In the simplest form, Bade et al. (2015a) proposed a statistical linear regression RT prediction based only on the octanol-water partition coefficient (Kow) using a dataset of over 200 pharmaceuticals. The model

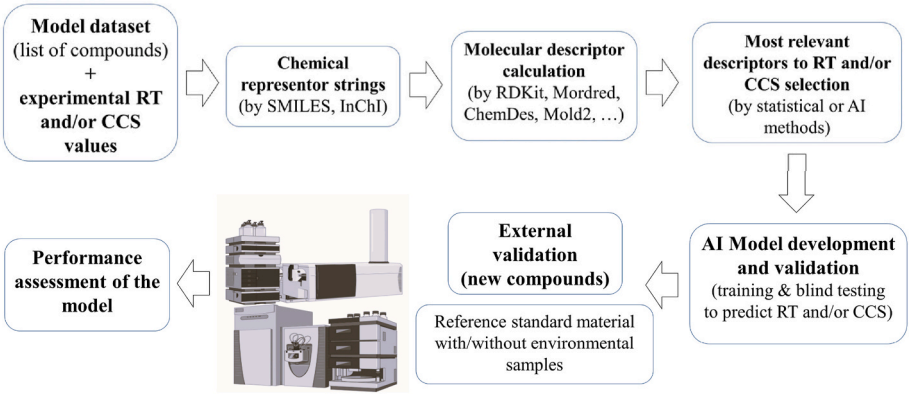


Fig. 4. Simplified schematic procedure of AI-assisted QSRR modelling for RT and/or CCS prediction.

was externally validated with real water and wastewater treatment samples, as shown in Table 1. It achieved 70% accuracy within a ±2 min tolerance for RT. In the same year, Bade et al. (2015a) published one of the most impactful studies to date on EPs in water and surface waters, including metabolites, transformation, and biodegradation products for the first time. They used a descriptor selection method originally developed for analysing doping drugs for the London Olympics 2012; Miller et al. (2013) with a large dataset of 544 compounds exhibiting diverse physicochemical properties. As summarised in Table 1, their 16-19-9-1 MLP model was more accurate than generalised regression neural networks (GRNN), RBFs, linear neural networks, and probabilistic neural networks (PNN). Of the 26 compounds tested, 23 could be screened without requiring analysis of the fragmented peaks, all peaking within ±2 min of the predicted RT.

Using a similar methodology, Munro et al. (2015) suggested that GRNN was more accurate than MLP and RBF for a dataset of 166 compounds externally validated against influent and effluent wastewater samples in the UK. Their study reported a matrix effect with a maximum shift of 1 min among 37 externally tested compounds, achieving success rates of 83% and 73% in identifying compounds in influent and effluent, respectively, and quickly eliminating up to 44% of the suspect list. These metrics provide a pragmatic performance evaluation of models for environmental suspect screenings. The same methodology was applied in a versatile monitoring programme (Richardson et al., 2021) in London rivers, using MLP-ANN. This model facilitated the tentative detection of 9 new EPs, including 7 new pharmaceuticals for the first time in rivers, and helped discard approximately 33% of false positive peaks.

Yang et al. (2021a, 2021b) proposed a new methodology using GNN for RT prediction without the need for molecular descriptors, utilising open-source code. Their approach involved creating molecular graphs from the METLIN small molecule retention time (SMRT) databank, which contains over 80,000 compounds. The simplified steps include obtaining molecular graphs from databanks using open-source chemical identifiers, updating subgraphs using GNN to generate new molecular data vectors, and summing the subgraphs in the last layer. The GNN-assisted method is robust and can be customised for other small datasets and chromatographic systems by keeping and freezing the hidden layers while modifying others. GNN outperformed RF, Bayesian ridge regression (BRR), CNN, and deep learning regression model (DLM) in RT prediction by at least 30%. However, applying this open-access GNN methodology for customised datasets of PPCPs needs further evaluation.

He et al. (2022) predicted relative RT by a deep neural network (DNN) exclusively for EPs. Initially considering 490 Eps, they streamlined the dataset to 391 compounds using a non-linearity function (ReLU) activation connected to the output singular layer, as described in Table 1. The linear type SVM was the best for training datasets, outperforming polynomial, radial basis or sigmoid SVM. The DNN model marginally outperformed MLR and SVM on the test dataset.

3.2.3. GA-integrated models

The AI-assisted models were sometimes integrated with GA as an evolutionary optimisation algorithm for fine-tuning within the training step of AI development. Miller et al. (2016) employed non-linear AI-assisted QSRR models to predict the sampling flow rate of a polar organic chemical integrative passive sampler (POCIS), thereby facilitating better chemical analysis of EPs. Their model was developed using a dataset of 73 compounds, and a GA was applied to select 24 descriptors coupled with an MLP model. Its performance was compared with a previously developed MLP-ANN model for RT predictions. The MLP was chosen based on its superior performance to RBF and GRNNs. Both models were trained using back-propagation and conjugate gradient descent algorithms. For training and verification, both models showed similar precision. For blind testing datasets, the GA-incorporated method ($R^2 = 0.8$) outperformed the ANN ($R^2 = 0.64$). However, for externally tested compounds, the absolute average error of ANN-MLP was about 4 times more accurate than the GA-incorporated method. This result highlights the importance of the external validation of models, preferably using environmental samples, as outcomes can differ significantly from blind testing when comparing multiple algorithms. Additionally, an AI-assisted molecular descriptor-based model developed for a single output, such as RT, can be assessed and applied to other outputs not necessarily limited to passive sampler flow rate. This is a successful example regarding the potential generalisability of AI-assisted QSRR models beyond the prediction of RT.

In a study relevant to the pharmaceutical industry with potential applications for PPCPs in water and wastewater, Mizera et al. (2017) integrated GA for descriptor ranking and selection, coupled with higher-order ANNs: pi-sigma (PS-ANN) and sigma-pi-sigma (SPS-ANN), as detailed in Table 1. PS-ANN showed higher robustness than SPS, although SPS outperformed in blind tests. The research suggested PS-ANN as a strong candidate for QSRR models developed with small datasets.

Aalizadeh et al. (2019, 2016) made significant contributions to the suspect and non-targeted screening of EPs by developing and testing AI-assisted models for RT prediction for different types of Reversed-Phase Liquid Chromatography (RPLC) in positive and negative ionisation modes and for HILIC in positive mode. Initially, they used a dataset of 550 compounds, extending to more than 2800 in 2019. In 2016, they used KNN machine learning and PCA statistical methods to split datasets and prevent bias. They compared stepwise statistical methods and GA for selecting descriptors, then correlated the descriptors and experimental retention data using MLR, ANN, and SVM. They proposed new model applicability methodologies and protocols to improve confidence through outlier identification. Their results showed better performance of KNN over PCA and GA compared to the stepwise statistical method. The KNN-GA-SVM ranked marginally the best, followed by KNN-GA-MLR and KNN-GA-ANN, with high accuracy and enhanced confidence from 100 externally tested compounds. In 2019, the GA-MLR and GA-SVM models were re-developed with larger datasets and externally validated with real wastewater and sludge to detect 21 PPCPs and lab-generated ozonation byproducts, showing that GA-SVM was more accurate.

3.2.4. Collision cross section prediction

Integrating LC-HRMS with ion mobility separation (IMS) has recently demonstrated significant capabilities to increase the confidence of detection for suspect and non-targeted screenings (Celma et al., 2020, 2022; Mollerup et al., 2018). This is achieved by considering CCS values, which, in contrast with RT, are regarded as system- and matrix-independent parameters (Gabelica and Marklund, 2018; Lee, 2017) and have the potential to replace complementary identification metrics such as ion fragmentation matching (Regueiro et al., 2017).

Mollerup et al. (2018) developed a combined double-output RT-CCS prediction model incorporating ANN and compared it with single-output RT and CCS models for the first time. The overall methodology of using AI was similar to Bade et al. (2015a) and Barron and McEnef (2016). The initial dataset included 827 RTs and 357 CCSs for pharmaceuticals, illicit drugs, and their metabolites. The most important descriptors for the RT models were logD, atomic logP, and the number of carbons and oxygens, while for the RT-CCS model, molecular weight, logD, and compound logP were crucial. The combined model (MLP-ANN) required more descriptors (24 inputs) compared to the RT model (MLP-ANN) and the CCS model (MLP-ANN). By defining performance metrics of 5% relative CCS error and 95th prediction error percentile for RT, the results showed around 92% of compounds (328 out of 357) met both accuracy thresholds. The combo model was accurate for internal and external testing over 36 reference materials. However, the research suggested using real samples and more external validation compounds to assess the model's performance in predicting true positives and eliminating false positives in suspect and non-targeted analyses.

Celma et al. (2022) proposed a method using the same instrument to

offset possible uncertainties due to utilising double analysers done by Mollerup et al. (2018). They used a larger dataset of 477 protonated molecules (positive ionisation mode), 169 deprotonated molecules, and 249 sodium adducts for CCS. By analysing the consistency of RT and CSS and the low correlation between these parameters, they suggested that combined prediction was unnecessary. Instead, they developed 3 models: RT, CSS for protonated and deprotonated molecules (CCS–H), and CSS for sodium adducts (CSS–Na) and considered 10, 7, and 10 different descriptors for these models, respectively (26 unshared in total). The RT prediction model had $R^2 = 0.85$ and showed ±2.3 min accuracy for a 95% confidence interval. The CCS models had an $R^2 = 0.96$ with ±6 % accuracy for 95% of the cases. These models were more accurate than the previously mentioned ANN-based model by about 13% for RT ±2 min for 95% of compounds and up to 2% for CSS, likely due to larger datasets and the better controllability of fitting methods. The use of AI to predict CCS has proven its applicability for wide-scope screening of PPCPs in the Amazon River, leading to the detection of 51 PPCPs and their metabolites (Fabregat-Safont et al., 2021).

3.3. Comparative analysis on AI modelling

The selection of the dataset (first step in Fig. 4) to develop RT or CCS predictive models is crucial for the performance and applicability of the models. Almost all the reviewed models have considered EPs rather than PPCPs, and only a few references specified the performance of the models according to the type of chemicals during external validation (Song et al., 2024). This seems to be a challenge for assessing the fairness of the models and their generalisability for PPCPs. When the datasets are for a larger group of compounds, the prioritised descriptors selected through feature selection (step 4 in Fig. 4) might differ when the dataset is only for PPCPs. This can impact the precision and accuracy of the model as the most important factor in data quality. This can lead to algorithmic bias in supervised AI model development step (step 5 in Fig. 4) regardless of the effectiveness of the common cross-validation methods for splitting the dataset during training or considering innovative outlier eliminations. Thus, it is recommended that predictive models be developed exclusively for PPCPs or that their performance be reported according to the type of EPs. Also, more research is needed to confirm the performance of advanced AI-assisted models for RT and CCS prediction derived from very large datasets of compounds, such as those studied by (Yang et al., 2021a; 2021b). Therefore, models exclusively developed for contaminants of emerging concern or PPCPs, with smaller datasets and external validation, are potentially more reliable for the tentative detection of PPCPs.

For RT prediction, the credibility of test dataset validation success has been suggested based on an acceptable error within a time window of 11% of the entire duration of the chromatography (Song et al., 2024). This criterion is more practical than relying solely on the mathematical accuracy of the models (R^2, RMSE, and MAE). The percentage of false negatives within the time window for blind test and external validation datasets can complement the reliability assessment of future RT models. While the variety of applied algorithms in chemical analysis is not diverse, weak learner models - SVM, GRNN, MLP - are used for data mining and classification. ML models, focusing on ANFIS and DNN, have outperformed other methods (Abba et al., 2020; Ghali et al., 2020; He et al., 2022). Additionally, GA's feature selection capability has proven to be a powerful tool for enhancing RT model prediction for detecting PPCPs in water and wastewater (Aalizadeh et al., 2016). However, improvements via ensemble AI techniques need further investigation (Abbasi et al., 2022).

Regarding the inputs of the RT and CCS models, the performance of models with more molecular descriptors is not necessarily superior to models with fewer inputs. Increasing the number of inputs can lead to challenges due to the non-availability of the inputs in databases or excessive calculations (Parinet et al., 2024), adding unnecessary complexity to the models and undermining their robustness (Barnard et al., 2020). The most important RT and CCS model inputs are logP and logD, followed by the number of atoms or the count of basic groups in the compounds. However, the models may select different inputs according to AI or statistical feature selection and the dataset of compounds used for model development. For a better comparison of AI-assisted models and the significance of molecular descriptors, a strategy similar to that devised for pesticides (Parinet, 2021a, 2021b) can be applied to PPCPs. This involves considering a unique dataset of PPCPs over the reviewed models and their inputs, including GA and CNN-optimised modifications, followed by experimental validation of new PPCPs.

Overall, the positives of spectroscopy include simplicity, rapidness, and cost-effectiveness compared to sophisticated tandem LC-MS. However, due to the lower sensitivity of spectroscopic analysis, the simultaneous and non-selective detection and quantification of PPCPs in complex water and wastewater samples at ng/L to μg/L levels (Hernández-Tenorio et al., 2022) present substantial challenges. Significant further modifications in methodologies and spectroscopy analyser technologies may be required. Nevertheless, AI-assisted spectroscopy can be viable for the rapid quantification or near real-time monitoring of selected PPCPs with relatively high concentrations in WWTPs.

4. AI assistance for monitoring of PPCPs

The reviewed methods for monitoring PPCPs, as detailed in Table 2, are generally proposed in 3 directions: feasibility studies for real-time monitoring, prediction of spatial or temporal occurrence, and classification and interpretation of monitoring data. The latter can be termed post-monitoring analysis, applicable for further environmental, health, and WBE investigations. As seen in Table 2, all the references are very recent, indicating that the application of AI for monitoring PPCPs is in its infancy but has a promising perspective. Kasprzyk-Hordern et al. (2022) have also highlighted the potential of AI and ML for data processing, interpretation, hazard modelling, and visualisation of community health.

For rapid and cost-effective real-time monitoring of PPCPs in wastewater, Post et al. (2022) and Yu et al. (2023) studied the feasibility of laser-induced Raman and fluorescence spectroscopy (LIRFS) and fluorescence image processing techniques, respectively. In the latter study, the method was developed by 45 experiments with 9 different oxytetracycline concentrations from 1 μM (0.44 mg/L) to 7 μM (3.11 mg/L) and tested with mixtures at 1–12 μM. The fluorescence response of luminescent metal-organic framework (MOF) material with tetracycline was captured using red, green, and blue (RGB) image saturation as inputs, followed by k-means and HCA and machine learning MLR. Only the red and blue variables showed higher importance with $R^2 \approx 0.96$. However, the model was not validated with real sewage samples and was only assessed by adding other substances as interferences to ensure satisfactory selectivity.

In addition, the model does not apply to other PPCPs and tetracyclines whose colour responses were not as notable as oxytetracycline. The results indicate that significant modifications are required for this method to enable real-time monitoring of multiple tetracyclines and, by inference, even more for other PPCPs present at lower concentrations in wastewater.

Post et al. (2022) employed spectroscopy at below 270 nm wavelengths for Raman and above 270 nm for fluorescence, integrated with CNN to classify the spectra and subsequently quantify conventional parameters such as nitrate and dissolved organic carbon (DOC), as well as concentrations of 7 pharmaceuticals, shown in Table 2, at WWTPs in Germany. The CNN had the same architecture suggested by Liu et al. (2017) with LeNet having pyramid-shaped convolutional layers and double fully connected classification layers with 23 substances with the Adam optimiser. The results showed that while the method effectively detected changes in nitrate and DOC, it lacked the precision to

Table 2
AI-based methods for monitoring and post-monitoring of PPCPs in water and wastewater.

Application	Method	Dataset	Input	Output	Pollutants	Monitoring Location	Reference
Correlating community status and well-being to PPCPs	K-means, statistical	NA	41 factors among birth control, disease burden, socioeconomic	Concentration	31 PPCPs	171 WWTPs, China	(He et al., 2024)
Community status and well-being factors due to PPCPs	RF, XGBoost, CatBoost	NA	10 factors among Economic, WWT efficiency, population	Daily mass load per capita	31 PPCPs in 7 groups	82 WWTPs, China	Chen et al. (2024)
Fluorescence imaging for tetracycline real-time monitoring in sewage	K-means, HCA + MLP	45:9	fluorescence image RGB	Concentration	Oxytetracycline	Lab tests	Yu et al. (2023)
finding patterns of EP variations during storm events	HCA + PLS	NA	Dynamic concentrations + Hyro-meteorological factors	Cluster patterns	11 Most persistent among 253 EPs monitored	Urban catchment, South Korea	Yun et al. (2023)
Online monitoring of pharmaceuticals in WWTPs	CNN (LeNet)	1700:566	fluorescence Raman spectrum	Wastewater concentration	Metformin, Carbamazepine, Hydrochlorothiazide, Acetaminophen, Naproxen, Diclofenac and Tryptophan	WWTPs, Germany	Post et al. (2022)
EPs prediction by Hybrid methods for spatial analyses	ANN, RF	366	7 water quality indicators	Concentration	Bisphenol A (BPA), N, N-diethyltoluamide (DEET)	Tropical urban reservoir, Singapore	Tong et al. (2022)
Correlating co-occurrence of drug groups	Statistical	NA	Concentration	Drug groups correlation	68 PPCPs	15 WWTPs, Slovak Republic	Bodík et al. (2021)
Mapping and prediction of the uptake of pharmaceuticals in lettuce irrigated by reclaimed water	ELMNet	NA	WWTP effluent concentration	Concentration in crops	Carbamazepine, Diclofenac	Spain	González García et al. (2019)
Location-wise patterns of EPs in surface and drinking waters	PCA + HCA	+16000 (from 229 samples)	Concentrations and locations	Location-wise clustering and EP distinction of locations	249 EPs	Lake, River, reclaimed land area, Netherlands	Houtman et al. (2019)

accurately identify individual substances. Instead, it classified groups of compounds based on their characteristic wavelengths.

The detection ranges of the pharmaceuticals, spanning from μg/L for carbamazepine, naproxen, and tryptophan to g/L for metformin, underscore an important observation: although the method demonstrates efficacy for certain pharmaceuticals, it lacks the robustness required for comprehensive monitoring in WWTPs.

For environmental prediction and monitoring of PPCPs, though not specifically in water and wastewater matrices, the accumulations of carbamazepine and diclofenac in lettuce crops irrigated with reclaimed effluent from 10 WWTPs in Spain were studied. The effluent had diverse concentrations of 12–14100 ng/L and 11–530 ng/L for carbamazepine and diclofenac, respectively (González García et al., 2019). These accumulations were predicted and regionally mapped using the extreme learning machine (ELM) algorithm with linear kernels. The accuracy of the models for both pharmaceuticals was high, with R^2 values between 0.96 and 0.98 for both training and testing datasets. Carbamazepine levels in the crops were found to be 106–180 ng/g, not harmful for human consumption, and diclofenac levels were below the lower detection limit, indicating the overall healthiness of the crops based on human daily intake standards. The method has the potential to be generalised for other crops and PPCPs.

In another study, Tong et al. (2022) demonstrated that hybrid modelling of rigorous fate and transport finite volume process-based models (Delft 3D suite models) with AI can facilitate fast predictions and early-warning system designs for PPCPs in surface water. Using monitoring data (You et al., 2015), they evaluated the performances of RF and ANN for the scenarios: (1) predicting downstream EPs (Bisphenol A and N, N-diethyltoluamide) using upstream pollutants, (2) predicting EPs at the same location using water quality indicators (WQIs), and (3) predicting downstream EPs using upstream EPs and WQIs. Results revealed that RF was distinctly superior to ANN, suggesting the capability of AI methods as effective open-source emulators or surrogates for rigorous process-based models.

This study suggests hybrid process-based/AI models should be implemented when WQIs are used to predict EPs with AI, potentially becoming a state-of-the-art method for temporal and spatial monitoring of PPCPs. The results also show that measured WQIs, especially temperature and total nutrient factors, are important for EP prediction by AI models. It should be highlighted that there is a knowledge gap focusing on the relationship between WQIs or physicochemical properties of water and the occurrence of PPCPs (Ohoro et al., 2022).

In the context of interpreting PPCP monitoring data, integrating AI and ML classifications has considerable potential for identifying the group-wise co-occurrence of PPCPs in urban wastewater without consumption records or specifying biomarkers, as commonly done in WBE. Bodík et al. (2021) classified 68 PPCPs in 15 WWTPs in Slovakia according to 10 therapeutic groups and geographical information using statistical methods. As an example, they found a meaningful correlation between the simultaneous intake of antibiotics and high blood pressure pharmaceuticals among almost all WWTPs. This study is an example of community health surveillance that relies only on interpreting the monitoring of long-term and nationwide occurrence of PPCPs in wastewater using statistical methods. This approach can be extended to include demographic factors and AI for the same purpose as suggested by the authors.

Unsupervised clustering AI methods, including HCA and K-means, and supervised classification methods, such as RF, can be applied for interpreting PPCP monitoring data (Chen et al., 2024; He et al., 2024). As reflected in Table 2, very recent studies (Chen et al., 2024; He et al., 2024) in China have attributed several socioeconomic, disease load, community health factors, and population metrics to the occurrence of PPCPs in WWTPs using AI methods, including k-means, decision tree, RF (showing the best results), XGBoost and CatBoost. RF demonstrated

higher performance. The findings revealed a moderate correlation (approximately 26%) between pharmaceutical and personal care products (PPCPs) and socioeconomic factors, and a stronger correlation (70%) with disease burden. These results suggest the potential for novel or unexpected trends, such as increased consumption of cardiovascular medications among younger populations (Chen et al., 2024) and a notable positive association between the antidepressant fluoxetine and the rates of abortion and intrauterine device (IUD) surgeries (He et al., 2024).

More advanced AI algorithms, including DNN, CNN, GNN and ELM, have demonstrated their applications for monitoring, yet further implementation and comparisons are needed to assess whether their performance against other AI methods (González García et al., 2019; Post et al., 2022).

For analysing large monitoring datasets over several years, Houtman et al. (2019) utilised PCA and HCA (listed in Table 2) to identify which compounds are most typical for 15 regions in surface water in the Netherlands. The research highlighted those nearby waterbodies can have different pollution patterns, indicating that each location may have specific priority compounds. This might be a remarkable insight for research and decision-making by including regional causative factors or industries responsible for such contamination with a river catchment-based approach (Collins et al., 2020). The approach suggests integrating and presenting data to have collaborative multilateral catchment management with the important role of community-based citizen scientists (Starkey et al., 2017).

For classifying the most persistent EPs and understanding their dynamics during stormwater events in a highly urbanised and industrialised catchment in South Korea (Yun et al., 2023) HCA and PLS were used to ascertain decisive hydro-meteorological factors. The results showed that in the first 30% volume of runoffs, pharmaceuticals such as metformin and acetaminophen (paracetamol) and pesticides reached their highest levels (high first flush effect), unlike the moderate variations of industrial EPs during storm events. The significance of antecedent dry hours for accumulation and storm flow for wash-offs of pollutants by stormwater was highlighted.

According to the future perspectives of WBE and monitoring of chemical pollution (Barron et al., 2024; Kasprzyk-Hordern et al., 2022, 2023b, 2023c, 2023d) and the available recent case studies reviewed, it is foreseeable that AI will have substantial potential in interpreting PPCP monitoring data in water and wastewater. Applying unsupervised clustering or association-based AI methods can uncover hidden patterns and reveal anomalies in community health (Chen et al., 2024; He et al., 2024). However, a major challenge is the lack of long-term monitoring data in publications, which restricts the reproducibility of interpretations and hinders comparative studies. This also brings imitations for assessing the generalisability of the methods and whether the clusters and associations identified are biased, signalling the need for improved accessibility to monitoring data in future research. The analysis of PPCP data can offer decision-makers valuable insights into community health, consumption patterns, demographic trends, stressors, and disparities.

5. Perspective calls for future research

The following points outline potential future research directions for the chemical analysis and monitoring of PPCPs in water and wastewater:

- There is a wide gap for comparative analysis of AI-integrated methods versus statistical or mathematical models in both chemical analysis and monitoring of EPs.
- AI assistance has been more extensively applied in chemical analysis than in monitoring. However, comprehensive comparative studies, particularly those involving external validation with wastewater and environmental samples, are still lacking. Therefore, future research should focus on analysing key factors affecting matrix effects, comparing models using different algorithms comprehensively, identifying significant descriptors or inputs for models categorised by class of PPCPs, examining the impact of dataset size and variety, evaluating the performance of models in HRMS analysers for eliminating false positives or rapidly shortlisting suspects, and assessing the impact of the models' accuracy tolerance.
- Implementing AI for the prediction of CCS, in addition to RT prediction, combined with routine mass-to-charge and ion fragmentation matching from databases, can significantly enhance the confidence in suspected and non-targeted detection of materials. This approach is particularly beneficial for substances with limited available reference standard materials, including pharmaceutical metabolites and their transport, degradation, or treatment byproducts in the environment and wastewater.
- AI can enhance spectroscopy and image processing methods, enabling rapid and selective chemical analysis or real-time monitoring of pharmaceuticals and PPCPs, particularly in wastewater systems with higher concentrations and better traceability than in surface waters. These methods still require substantial advancements to achieve better reliability and robustness in practice.
- AI methods have demonstrated significant versatility in data processing, particularly in clustering and classification for monitoring pharmaceuticals and PPCPs. This suggests a promising role for AI in decision-making within WBE and community health frameworks, with ample opportunities at the intersection of social sciences (Barron et al., 2024). However, the corresponding studies are still scarce.
- AI-assisted anomaly detection and assessing new inputs, including the side effects of PPCP consumption, industrial activities, and farming on community health, can be worthwhile for future studies.
- AI has the potential to be integrated with rigorous process-based models for spatiotemporal analyses of pharmaceuticals and PPCPs in surface waters. These can be hybrid methodologies or surrogate models for fate and transport mechanisms, helping predict the occurrence of PPCPs. Successful application requires more extensive monitoring data and the development or access to rigorous models.
- While AI-assisted chemical analysis and monitoring methods show potential for targeting PPCP metabolites, transformation products, or degradation products, few studies have focused on these EPs or biomarkers. Therefore, future research is likely to concentrate on these compounds increasingly.
- AI can assist the selection and optimisation of sampling types and volumes, sample preparation, chemical analysis methods, costs, energy and solvent consumption, waste management and sustainability, and environmental friendliness, being worthwhile when evaluating the degradation mechanisms, simulations, and treatment technologies for future studies.

6. Conclusions

This review examines the roles and contexts of AI in the chemical analysis, monitoring, and interpretation of PPCPs in water and wastewater matrices, focusing on studies from the past decade. These AI applications in environmental chemistry can also be incorporated into toxicology, environmental risk assessments, treatment evaluations, and community health surveillance. The trend in recent publications suggests that while the application of sophisticated AI methods for PPCP analysis and monitoring is still in its early stages, traditional neural networks have demonstrated robustness in chemical analysis.

Although AI-assisted chemical analysis and near-real-time monitoring of PPCPs using spectroscopic analysers have proven feasibility for a select few compounds with relatively high concentrations in wastewater, they fall short of the comprehensiveness and precision of chromatography combined with tandem and high-resolution mass spectrometry. Available predictive models for RT and CCS in suspect and non-targeted analyses of emerging pollutants emphasise the need to

expand the inclusion of PPCPs in modelling datasets and conduct external validations, particularly using environmental water samples in addition to standard reference materials. Group-wise chemical validation of models developed for EPs is also crucial to assess reliability and enable comparative studies.

The use of data-driven AI methods for interpreting PPCP monitoring data, especially in wastewater, shows promise for assessing community health dynamics and socioeconomic or demographic patterns at regional or national levels. However, this research direction is still in its infancy. A major challenge is the lack of long-term monitoring data, which limits the ability to draw deeper insights from existing datasets or perform comparative studies using different AI methods. Expanding collaborative efforts for open dissemination of monitoring data could significantly advance wastewater-based epidemiology and enhance the use of AI in future decision-making.

CRediT authorship contribution statement

Babak Kavianpour: Writing – review & editing, Writing – original draft, Methodology, Conceptualization. **Farzad Piadeh:** Writing – review & editing, Writing – original draft, Visualization. **Mohammad Gheibi:** Writing – review & editing. **Atiyeh Ardakanian:** Writing – review & editing, Supervision. **Kourosh Behzadian:** Writing – review & editing, Supervision, Methodology. **Luiza C. Campos:** Writing – review & editing.

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

This study is part of a PhD study at the University of West London (UWL) in collaboration with the University College London (UCL). The overall focus of this ongoing study is the investigation of PPCPs in surface waters through sampling, water quality tests, SPE sample preparation, and chemical analysis using LC-MS/MS, followed by AI-driven data analysis. The authors would like to acknowledge the UWL PhD Vice Chancellor Scholarship allocated to the first one.

Data availability

Data will be made available on request.

References

Aalizadeh, R., Nika, M.C., Thomaidis, N.S., 2019. Development and application of retention time prediction models in the suspect and non-target screening of emerging contaminants. J. Hazard Mater. 363, 277–285. https://doi.org/10.1016/j.jhazmat.2018.09.047.

Aalizadeh, R., Thomaidis, N.S., Bletsou, A.A., Gago-Ferrero, P., 2016. Quantitative structure-retention relationship models to support nontarget high-resolution mass spectrometric screening of emerging contaminants in environmental samples. J. Chem. Inf. Model. 56 (7), 1384–1398. https://doi.org/10.1021/acs.jcim.5b00752.

Abba, S.I., Usman, A.G., Işik, S., 2020. Simulation for response surface in the HPLC optimization method development using artificial intelligence models: a data-driven approach. Chemometr. Intell. Lab. Syst. 201. https://doi.org/10.1016/j.chemolab.2020.104007.

Abbasi, E., Alavi Moghaddam, M.R., Kowsari, E., 2022. A systematic and critical review on development of machine learning based-ensemble models for prediction of adsorption process efficiency. In: Journal of Cleaner Production, vol. 379. Elsevier Ltd. https://doi.org/10.1016/j.jclepro.2022.134588.

Afolabi, I.C., Popoola, S.I., Bello, O.S., 2020. Modeling pseudo-second-order kinetics of orange peel-paracetamol adsorption process using artificial neural network. Chemometr. Intell. Lab. Syst. 203. https://doi.org/10.1016/j.chemolab.2020.104053.

Ahmadi Azqhandi, M.H., Foroughi, M., Gholami, Z., 2022. Efficient removal of levofloxacin by a magnetic NiFe-LDH/N-MWCNTs nanocomposite: characterization, response surface methodology, and mechanism. Environ. Res. 215. https://doi.org/10.1016/j.envres.2022.113967.

Alam, G., Ihsanullah, I., Naushad, M., Sillanpää, M., 2022. Applications of artificial intelligence in water treatment for optimization and automation of adsorption processes: recent advances and prospects. In: Chemical Engineering Journal, vol. 427. Elsevier B.V. https://doi.org/10.1016/j.cej.2021.130011.

Arab, M., Faramarz, M.G., Hashim, K., 2022. Applications of computational and statistical models for optimizing the electrochemical removal of cephalexin antibiotic from water. Water (Switzerland) 14 (3). https://doi.org/10.3390/w14030344.

Badawi, A.K., Hassan, R., 2024. Optimizing sludge extract reuse from physico-chemical processes for zero-waste discharge: a critical review. Desalination Water Treat. 319, 100527. https://doi.org/10.1016/J.DWT.2024.100527.

Badawi, A.K., Hassan, R., Alghamdi, A.M., Ismail, B., Osman, R.M., Salama, R.S., 2024a. Advancing cobalt ferrite-supported activated carbon from orange peels for real pulp and paper mill wastewater treatment. Desalination Water Treat. 318, 100331. https://doi.org/10.1016/J.DWT.2024.100331.

Badawi, A.K., Hassan, R., Farouk, M., Bakhoum, E.S., Salama, R.S., 2024b. Optimizing the coagulation/flocculation process for the treatment of slaughterhouse and meat processing wastewater: experimental studies and pilot-scale proposal. Int. J. Environ. Sci. Technol. 21 (13), 8431–8446. https://doi.org/10.1007/S13762-024-05591-Y/METRICS.

Bade, R., Bijlsma, L., Miller, T.H., Barron, L.P., Sancho, J.V., Hernández, F., 2015a. Suspect screening of large numbers of emerging contaminants in environmental waters using artificial neural networks for chromatographic retention time prediction and high resolution mass spectrometry data analysis. Sci. Total Environ. 538, 934–941. https://doi.org/10.1016/j.scitotenv.2015.08.078.

Barnard, T., Hagan, H., Tseng, S., Sosso, G.C., 2020. Less may be more: an informed reflection on molecular descriptors for drug design and discovery. Molecular Systems Design & Engineering 5 (1), 317–329. https://doi.org/10.1039/C9ME00109C.

Barron, L.P., McEneff, G.L., 2016. Gradient liquid chromatographic retention time prediction for suspect screening applications: a critical assessment of a generalised artificial neural network-based approach across 10 multi-residue reversed-phase analytical methods. Talanta 147, 261–270. https://doi.org/10.1016/j.talanta.2015.09.065.

Barron, L.P., Richardson, A.K., Hein, W., 2024. Chemical pollution in water: scalable and intersectional opportunities for the analytical and social sciences. One Earth 7 (3), 350–354. https://doi.org/10.1016/J.ONEEAR.2024.02.008.

Beek, A.DT., Weber, F.A., Bergmann, A., Hickmann, S., Ebert, I., Hein, A., Küster, A., 2016. Pharmaceuticals in the environment-Global occurrences and perspectives. Environ. Toxicol. Chem. 35 (4), 823–835. https://doi.org/10.1002/etc.3339.

Belay, M.H., Precht, U., Mortensen, P., Marengo, E., Robotti, E., 2022. A fully automated online SPE-LC-MS/MS method for the determination of 10 pharmaceuticals in wastewater samples. Toxics 10 (3). https://doi.org/10.3390/toxics10030103.

Benjathapanun, N., Boyle, W.J.O., Grattan, K.T.V., 1998. Classification of UV-Vis spectroscopic data using principal component analysis and neural network techniques. Measurement 24.

Bodík, M., Mackuľak, T., Feher, M., Staňová, A.V., Grabicová, K., Varjúová, D., Bodík, I., 2021. Searching for the correlations between the use of different groups of pharmaceuticals from wastewaters. Ecotoxicol. Environ. Saf. 228. https://doi.org/10.1016/j.ecoenv.2021.112973.

Cai, M.-Q., Wang, R., Feng, L., Zhang, L.-Q., 2015. Determination of selected pharmaceuticals in tap water and drinking water treatment plant by high-performance liquid chromatography-triple quadrupole mass spectrometer in Beijing, China. Environ. Sci. Pollut. Control Ser. 22 (3). https://doi.org/10.1007/s11356-014-3473-8.

Carabajal, M., Teglia, C.M., Maine, M.A., Goicoechea, H.C., 2021. Multivariate optimization of a dispersive liquid-liquid microextraction method for the determination of six antiparasite drugs in kennel effluent waters by using second-order chromatographic data. Talanta 224. https://doi.org/10.1016/j.talanta.2020.121929.

Celma, A., Bade, R., Sancho, J.V., Hernandez, F., Humphries, M., Bijlsma, L., 2022. Prediction of retention time and collision cross section (CCSH+, CCSH-, and CCSNa+) of emerging contaminants using multiple adaptive regression splines. J. Chem. Inf. Model. 62 (22), 5425–5434. https://doi.org/10.1021/acs.jcim.2c00847.

Celma, A., Sancho, J.V., Schymanski, E.L., Fabregat-Safont, D., Ibáñez, M., Goshawk, J., Barknowitz, G., Hernández, F., Bijlsma, L., 2020. Improving target and suspect screening high-resolution mass spectrometry workflows in environmental analysis by ion mobility separation. Environmental Science & Technology 54 (23), 15120–15131. https://doi.org/10.1021/ACS.EST.0C05713.

Chen, X., Lu, X., Su, G., Meng, J., Li, Q., Peng, J., Hua, Y., Shi, B., 2024. Regional Disparities in PPCPs Contamination of Urban Wastewater Treatment Plants: Unveiling Influential Factors and Ecological Risk Predictions. https://doi.org/10.2139/ssrn.4688883.

Chowdhury, S., Karanfil, T., 2024. Applications of artificial intelligence (AI) in drinking water treatment processes: possibilities. Chemosphere 356, 141958. https://doi.org/10.1016/J.CHEMOSPHERE.2024.141958.

Collins, R., Johnson, D., Crilly, D., Rickard, A., Neal, L., Morse, A., Walker, M., Lear, R., Deasy, C., Paling, N., Anderton, S., Ryder, C., Bide, P., Holt, A., 2020. Collaborative water management across england – an overview of the catchment based approach. Environ. Sci. Pol. 112, 117–125. https://doi.org/10.1016/J.ENVSCI.2020.06.001.

Deylami, S., Sabzevari, M.H., Ghaedi, M., Azqhandi, M.H.A., Marahel, F., 2023. Efficient photodegradation of disulfine blue dye and Tetracycline over Robust and Green g-CN/Ag3VO4/PAN nanofibers: experimental design, RSM, RBF-NN and ANFIS modeling. Process Saf. Environ. Protect. 169, 71–81. https://doi.org/10.1016/J.PSEP.2022.10.080.

Domingo-Almenara, X., Guijas, C., Billings, E., Montenegro-Burke, J.R., Uritboonthai, W., Aisporna, A.E., Chen, E., Benton, H.P., Siuzdak, G., 2019. The METLIN small molecule dataset for machine learning-based retention time prediction. Nat. Commun. 10 (1). https://doi.org/10.1038/s41467-019-13680-7.

Dulio, V., Alygizakis, N., Ng, K., Schymanski, E.L., Andres, S., Vorkamp, K., Hollender, J., Finckh, S., Aalizadeh, R., Ahrens, L., Bouhoulle, E., Cirka, Ľ., Derksen, A., Deviller, G., Duffek, A., Esperanza, M., Fischer, S., Fu, Q., Gago-Ferrero, P., et al., 2024. Beyond target chemicals: updating the NORMAN prioritisation scheme to support the EU chemicals strategy with semi-quantitative suspect/non-target screening data. Environ. Sci. Eur. 36 (1), 1–22. https://doi.org/10.1186/S12302-024-00936-3/TABLES/5.

Egli, M., Rapp Wright, H., Oloyede, O., Francis, W., Preston-Allen, R., Friedman, S., Woodward, G., Piel, F.B., Barron, L.P., 2023. A One-Health environmental risk assessment of contaminants of emerging concern in London's waterways throughout the SARS-CoV-2 pandemic. Environ. Int. 180, 108210. https://doi.org/10.1016/J.ENVINT.2023.108210.

Fabregat-Safont, D., Ibáñez, M., Bijlsma, L., Hernández, F., Waichman, A.V., de Oliveira, R., Rico, A., 2021. Wide-scope screening of pharmaceuticals, illicit drugs and their metabolites in the Amazon River. Water Res. 200. https://doi.org/10.1016/j.watres.2021.117251.

Fick, J., Söderström, H., Lindberg, R.H., Phan, C., Tysklind, M., Larsson, D.G.J., 2009. Contamination of surface, ground, and drinking water from pharmaceutical production. Environ. Toxicol. Chem. 28 (12). https://doi.org/10.1897/09-073.1.

Foroughi, M., Arzehgar, A., Seyedhasani, S.N., Nadali, A., Zoroufchi Benis, K., 2024. Application of machine learning for antibiotic resistance in water and wastewater: a systematic review. Chemosphere 358, 142223. https://doi.org/10.1016/J.CHEMOSPHERE.2024.142223.

Fu, W., Li, X., Yang, Y., Song, D., 2023. Enhanced degradation of bisphenol A: influence of optimization of removal, kinetic model studies, application of machine learning and microalgae-bacteria consortia. Sci. Total Environ. 858, 159876. https://doi.org/10.1016/J.SCITOTENV.2022.159876.

Gabelica, V., Marklund, E., 2018. Fundamentals of ion mobility spectrometry. Curr. Opin. Chem. Biol. 42, 51–59. https://doi.org/10.1016/J.CBPA.2017.10.022.

Gardner, M.J., Comber, S.D.W., Ellor, B., 2022. Summary of data from the UKWIR chemical investigations programme and a comparison of data from the past ten years' monitoring of effluent quality. Sci. Total Environ. 832. https://doi.org/10.1016/j.scitotenv.2022.155041.

Ghali, U.M., Usman, A.G., Chellube, Z.M., Degm, M.A.A., Hoti, K., Umar, H., Abba, S.I., 2020. Advanced chromatographic technique for performance simulation of anti-Alzheimer agent: an ensemble machine learning approach. SN Appl. Sci. 2 (11). https://doi.org/10.1007/s42452-020-03690-2.

Gholami Shirkoohi, M., Tyagi, R.D., Vanrolleghem, P.A., Drogui, P., 2022. Modelling and optimization of psychoactive pharmaceutical caffeine removal by electrochemical oxidation process: a comparative study between response surface methodology (RSM) and adaptive neuro fuzzy inference system (ANFIS). Separation and Purification Technology 290. https://doi.org/10.1016/j.seppur.2022.120902.

Gholami, Z., Ahmadi Azqhandi, M.H., Hosseini Sabzevari, M., Khazali, F., 2023. Evaluation of least square support vector machine, generalized regression neural network and response surface methodology in modeling the removal of Levofloxacin and Ciprofloxacin from aqueous solutions using ionic liquid @Graphene oxide@ ionic liquid NC. Alex. Eng. J. 73, 593–606. https://doi.org/10.1016/J.AEJ.2023.04.041.

Gholami, Z., Foroughi, M., Ahmadi Azqhandi, M.H., 2024. Double ionic liquid reinforced g-CN nanocomposite for an enhanced adsorption of methylparaben: mechanism, modeling, and optimization. Chemosphere 349, 141006. https://doi.org/10.1016/J.CHEMOSPHERE.2023.141006.

Gómez-Canela, C., Edo, S., Rodríguez, N., Gotor, G., Lacorte, S., 2021. Comprehensive characterization of 76 pharmaceuticals and metabolites in wastewater by lc-ms/ms. Chemosensors 9 (10). https://doi.org/10.3390/chemosensors9100273.

González García, M., Fernández-López, C., Bueno-Crespo, A., Martínez-España, R., 2019. Extreme learning machine-based prediction of uptake of pharmaceuticals in reclaimed water-irrigated lettuces in the Region of Murcia, Spain. Biosyst. Eng. 177, 78–89. https://doi.org/10.1016/j.biosystemseng.2018.09.006.

Guo, Y., Liu, C., Ye, R., Duan, Q., 2020. Advances on water quality detection by uv-vis spectroscopy. Appl. Sci. 10 (19), 1–18. https://doi.org/10.3390/app10196874. MDPI AG.

Haddad, P.R., Taraji, M., Szücs, R., 2021. Prediction of analyte retention time in liquid chromatography. In Analytical Chemistry 93 (1), 228–256. https://doi.org/10.1021/acs.analchem.0c04190. American Chemical Society.

Hameedat, F., Hawamdeh, S., Alnabulsi, S., Zayed, A., 2022. High performance liquid chromatography (HPLC) with fluorescence detection for quantification of steroids in clinical, pharmaceutical, and environmental samples: a review. Molecules 27 (Issue 6). https://doi.org/10.3390/molecules27061807. MDPI.

Hapeshi, E., Gros, M., Lopez-Serna, R., Boleda, M.-R., Ventura, F., Petrovic, M., Barceló, D., Fatta-Kassinos, D., 2015. Licit and illicit drugs in urban wastewater in Cyprus. CLEAN - Soil, Air, Water 43 (9). https://doi.org/10.1002/clen.201400483.

Hassan, R., Alluqmani, A.E., Badawi, A.K., 2024. An eco-friendly solution for greywater treatment via date palm fiber filter. Desalination Water Treat. 317, 100163. https://doi.org/10.1016/J.DWT.2024.100163.

He, Q., Li, H., Jin, B., Li, W., Shao, B., Zhang, L., 2022. QSRR model for identification and screening of emerging pollutants based on artificial intelligence algorithms. Environmental Pollutants and Bioavailability 34 (1), 331–337. https://doi.org/10.1080/26395940.2022.2106311.

Hawash, H.B., Moneer, A.A., Galhoum, A.A., Elgarahy, A.M., Mohamed, W.A.A., Samy, M., El-Seedi, H.R., Gaballah, M.S., Mubarak, M.F., Attia, N.F., 2023. Occurrence and spatial distribution of pharmaceuticals and personal care products (PPCPs) in the aquatic environment, their characteristics, and adopted legislations. J. Water Process Eng. 52. https://doi.org/10.1016/j.jwpe.2023.103490. Elsevier Ltd.

He, R., Chen, L., Mu, H., Ren, H., Wu, B., 2024. Correlations between China's socioeconomic status, disease burdens, and pharmaceuticals and personal care product levels in wastewater. J. Hazard Mater. 463, 132867. https://doi.org/10.1016/J.JHAZMAT.2023.132867.

Heidari, Z., Pelalak, R., Eshaghi Malekshah, R., Pishnamazi, M., Rezakazemi, M., Aminabhavi, T.M., Shirazian, S., 2022. A new insight into catalytic ozonation of sulfasalazine antibiotic by plasma-treated limonite nanostructures: experimental, modeling and mechanism. Chem. Eng. J. 428, 131230. https://doi.org/10.1016/J.CEJ.2021.131230.

Hernández, F., Castiglioni, S., Covaci, A., de Voogt, P., Emke, E., Kasprzyk-Hordern, B., Ort, C., Reid, M., Sancho, J.V., Thomas, K.V., van Nuijs, A.L.N., Zuccato, E., Bijlsma, L., 2018. Mass spectrometric strategies for the investigation of biomarkers of illicit drug use in wastewater. In: In Mass Spectrometry Reviews, vol. 37. John Wiley and Sons Inc, pp. 258–280. https://doi.org/10.1002/mas.21525, 3.

Hernández-Tenorio, R., González-Juárez, E., Guzmán-Mar, J.L., Hinojosa-Reyes, L., Hernández-Ramírez, A., 2022. Review of occurrence of pharmaceuticals worldwide for estimating concentration ranges in aquatic environments at the end of the last decade. Journal of Hazardous Materials Advances 8, 100172. https://doi.org/10.1016/j.hazadv.2022.100172.

Hollender, J., Schymanski, E.L., Ahrens, L., Alygizakis, N., Béen, F., Bijlsma, L., Brunner, A.M., Celma, A., Fildier, A., Fu, Q., Gago-Ferrero, P., Gil-Solsona, R., Haglund, P., Hansen, M., Kaserzon, S., Kruve, A., Lamoree, M., Margoum, C., Meijer, J., et al., 2023. NORMAN guidance on suspect and non-target screening in environmental monitoring. Environ. Sci. Eur. 35 (1), 1–61. https://doi.org/10.1186/S12302-023-00779-4, 35(1).

Hosseini, O., Zare-Shahabadi, V., Ghaedi, M., Azqhandi, M.H.A., 2022. Experimental design, RSM and ANN modeling of tetracycline photocatalytic degradation using LDH@CN. J. Environ. Chem. Eng. 10 (5). https://doi.org/10.1016/j.jece.2022.108345.

Houtman, C.J., Kroesbergen, J., Baggelaar, P.K., van Lieverloo, J.H.M., 2019. Statistical analysis of a large set of semi-quantitative GC–MS screening data to evaluate and prioritize organic contaminants in surface and drinking water of The Netherlands. Sci. Total Environ. 697. https://doi.org/10.1016/j.scitotenv.2019.133806.

Inés Toral, M., Pope, S., Quintanilla, S., Richter, P., 2002. Simultaneous determination of amiloride and furosemide in pharmaceutical formulations by first digital derivative spectrophotometry. Int. J. Pharm. 249 (1–2), 117–126. https://doi.org/10.1016/S0378-5173(02)00482-9.

Kasprzyk-Hordern, B., Adams, B., Adewale, I.D., Agunbiade, F.O., Akinyemi, M.I., Archer, E., Badru, F.A., Barnett, J., Bishop, I.J., Di Lorenzo, M., Estrela, P., Faraway, J., Fasona, M.J., Fayomi, S.A., Feil, E.J., Hyatt, L.J., Irewale, A.T., Kjeldsen, T., Lasisi, A.K.S., et al., 2022. Wastewater-based epidemiology in hazard forecasting and early-warning systems for global health risks. Environ. Int. 161. https://doi.org/10.1016/j.envint.2022.107143.

Kasprzyk-Hordern, B., Sims, N., Farkas, K., Jagadeesan, K., Proctor, K., Wade, M.J., Jones, D.L., 2023a. Wastewater-based epidemiology for comprehensive community health diagnostics in a national surveillance study: mining biochemical markers in wastewater. J. Hazard Mater. 450. https://doi.org/10.1016/j.jhazmat.2023.130989.

Kasprzyk-Hordern, B., Sims, N., Farkas, K., Jagadeesan, K., Proctor, K., Wade, M.J., Jones, D.L., 2023b. Wastewater-based epidemiology for comprehensive community health diagnostics in a national surveillance study: mining biochemical markers in wastewater. J. Hazard Mater. 450. https://doi.org/10.1016/j.jhazmat.2023.130989.

Kasprzyk-Hordern, B., Sims, N., Farkas, K., Jagadeesan, K., Proctor, K., Wade, M.J., Jones, D.L., 2023c. Wastewater-based epidemiology for comprehensive community health diagnostics in a national surveillance study: mining biochemical markers in wastewater. J. Hazard Mater. 450. https://doi.org/10.1016/j.jhazmat.2023.130989.

Kasprzyk-Hordern, B., Sims, N., Farkas, K., Jagadeesan, K., Proctor, K., Wade, M.J., Jones, D.L., 2023d. Wastewater-based epidemiology for comprehensive community health diagnostics in a national surveillance study: mining biochemical markers in wastewater. J. Hazard Mater. 450. https://doi.org/10.1016/j.jhazmat.2023.130989.

Kim, S., Aga, D.S., 2007. Potential ecological and human health impacts of antibiotics and antibiotic-resistant bacteria from wastewater treatment plants. J. Toxicol. Environ. Health, Part A B 10 (8). https://doi.org/10.1080/15287390600975137.

Lee, J.W., 2017. Basics of Ion Mobility Mass Spectrometry 8 (No.4), 79–89. https://doi.org/10.5478/MSL.2017.8.4.79, 8(4), 79–89.

Li, L., Rong, S., Wang, R., Yu, S., 2021. Recent advances in artificial intelligence and machine learning for nonlinear relationship analysis and process control in drinking water treatment: a review. Chem. Eng. J. 405. https://doi.org/10.1016/j.cej.2020.126673.

Lindim, C., van Gils, J., Georgieva, D., Mekenyan, O., Cousins, I.T., 2016. Evaluation of human pharmaceutical emissions and concentrations in Swedish river basins. Sci. Total Environ. 572. https://doi.org/10.1016/j.scitotenv.2016.08.074.

Liu, J., Osadchy, M., Ashton, L., Foster, M., Solomon, C.J., Gibson, S.J., 2017. Deep convolutional neural networks for Raman spectrum recognition: a unified solution. Analyst 142 (21), 4067–4074. https://doi.org/10.1039/c7an01371j.

Mao, K., Zhang, H., Pan, Y., Yang, Z., 2021. Biosensors for wastewater-based epidemiology for monitoring public health. In: Water Research, vol. 191. Elsevier Ltd. https://doi.org/10.1016/j.watres.2020.116787.

Miller, T.H., Baz-Lomba, J.A., Harman, C., Reid, M.J., Owen, S.F., Bury, N.R., Thomas, K.V., Barron, L.P., 2016. The first attempt at non-linear in silico prediction of sampling rates for polar organic chemical integrative samplers (POCIS). Environ. Sci. Technol. 50 (15), 7973–7981. https://doi.org/10.1021/acs.est.6b01407.

Miller, T.H., Musenga, A., Cowan, D.A., Barron, L.P., 2013. Prediction of chromatographic retention time in high-resolution anti-doping screening data using

artificial neural networks. Anal. Chem. 85 (21), 10330–10337. https://doi.org/10.1021/AC4024878/SUPPL_FILE/AC4024878_SI_001.PDF.
Mizera, M., Krause, A., Zalewski, P., Skibiński, R., Cielecka-Piontek, J., 2017. Quantitative structure-retention relationship model for the determination of naratriptan hydrochloride and its impurities based on artificial neural networks coupled with genetic algorithm. Talanta 164, 164–174. https://doi.org/10.1016/j.talanta.2016.11.041.
Mohammadpoor, M., Kakhki, R.M., Assadi, H., 2019. A bayesian regularized artificial neural network for simultaneous determination of loratadine, naproxen and diclofenac in wastewaters. Curr. Pharmaceut. Anal. 16 (8), 1083–1092. https://doi.org/10.2174/1573412915666190618123154.
Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., Antes, G., Atkins, D., Barbour, V., Barrowman, N., Berlin, J.A., Clark, J., Clarke, M., Cook, D., D'Amico, R., Deeks, J.J., Devereaux, P.J., Dickersin, K., Egger, M., Ernst, E., Gøtzsche, P.C., et al., 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med. 6 (7), e1000097. https://doi.org/10.1371/JOURNAL.PMED.1000097.
Mojiri, A., Ozaki, N., Kazeroon, R.A., Rezania, S., Baharlooeian, M., Vakili, M., Farraji, H., Ohashi, A., Kindaichi, T., Zhou, J.L., 2022. Contaminant removal from wastewater by microalgal photobioreactors and modeling by artificial neural network. Water (Switzerland) 14 (24). https://doi.org/10.3390/w14244046.
Mollerup, C.B., Mardal, M., Dalsgaard, P.W., Linnet, K., Barron, L.P., 2018. Prediction of collision cross section and retention time for broad scope screening in gradient reversed-phase liquid chromatography-ion mobility-high resolution accurate mass spectrometry. J. Chromatogr. A 1542, 82–88. https://doi.org/10.1016/j.chroma.2018.02.025.
Moriwaki, H., Tian, Y.S., Kawashita, N., Takagi, T., 2018. Mordred: a molecular descriptor calculator. J. Cheminf. 10 (1). https://doi.org/10.1186/s13321-018-0258-y.
Munro, K., Martins, C.P.B., Loewenthal, M., Comber, S., Cowan, D.A., Pereira, L., Barron, L.P., 2019. Evaluation of combined sewer overflow impacts on short-term pharmaceutical and illicit drug occurrence in a heavily urbanised tidal river catchment (London, UK). Sci. Total Environ. 657, 1099–1111. https://doi.org/10.1016/j.scitotenv.2018.12.108.
Munro, K., Miller, T.H., Martins, C.P.B., Edge, A.M., Cowan, D.A., Barron, L.P., 2015. Artificial neural network modelling of pharmaceutical residue retention times in wastewater extracts using gradient liquid chromatography-high resolution mass spectrometry data. J. Chromatogr. A 1396, 34–44. https://doi.org/10.1016/j.chroma.2015.03.063.
Murray, C.J., Ikuta, K.S., Sharara, F., Swetschinski, L., Robles Aguilar, G., Gray, A., Han, C., Bisignano, C., Rao, P., Wool, E., Johnson, S.C., Browne, A.J., Chipeta, M.G., Fell, F., Hackett, S., Haines-Woodhouse, G., Kashef Hamadani, B.H., Kumaran, E.A.P., McManigal, B., et al., 2022. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. Lancet 399 (10325), 629–655. https://doi.org/10.1016/S0140-6736(21)02724-0.
Naderi, K., Foroughi, M., Azqhandi, M.H.A., 2022. Tetracycline capture from aqueous solutions by nanocomposite of MWCNTs reinforced with glutaraldehyde cross-linked poly (vinyl alcohol)/chitosan. Chemosphere 303, 135124. https://doi.org/10.1016/J.CHEMOSPHERE.2022.135124.
Oaks, J.L., Gilbert, M., Virani, M.Z., Watson, R.T., Meteyer, C.U., Rideout, B.A., Shivaprasad, H.L., Ahmed, S., Iqbal Chaudhry, M.J., Arshad, M., Mahmood, S., Ali, A., Ahmed Khan, A., 2004. Diclofenac residues as the cause of vulture population decline in Pakistan. Nature 427 (6975). https://doi.org/10.1038/nature02317.
Ohoro, C.R., Adeniji, A.O., Elsheikh, E.A.E., Al-Marzouqi, A., Otim, M., Okoh, O.O., Okoh, A.I., 2022. Influence of physicochemical parameters on PPCP occurrences in the wetlands. Environ. Monit. Assess. 194 (5). https://doi.org/10.1007/s10661-022-09990-x. Springer Science and Business Media Deutschland GmbH.
Parinet, J., 2021a. Predicting reversed-phase liquid chromatographic retention times of pesticides by deep neural networks. Heliyon 7 (12), e08563. https://doi.org/10.1016/j.heliyon.2021.e08563.
Parinet, J., 2021b. Prediction of pesticide retention time in reversed-phase liquid chromatography using quantitative-structure retention relationship models: a comparative study of seven molecular descriptors datasets. Chemosphere 275, 130036. https://doi.org/10.1016/J.CHEMOSPHERE.2021.130036.
Parinet, J., Makni, Y., Diallo, T., Guérin, T., 2024. Liquid chromatographic retention time prediction models to secure and improve the feature annotation process in high-resolution mass spectrometry. Talanta 267, 125214. https://doi.org/10.1016/J.TALANTA.2023.125214.
Patel, M., Kumar, R., Kishor, K., Mlsna, T., Pittman, C.U., Mohan, D., 2019. Pharmaceuticals of emerging concern in aquatic systems: chemistry, occurrence, effects, and removal methods. Chem. Rev. 119 (6), 3510–3673. https://doi.org/10.1021/acs.chemrev.8b00299.
Pelalak, R., Alizadeh, R., Ghareshabani, E., Heidari, Z., 2020. Degradation of sulfonamide antibiotics using ozone-based advanced oxidation process: experimental, modeling, transformation mechanism and DFT study. Sci. Total Environ. 734. https://doi.org/10.1016/j.scitotenv.2020.139446.
Post, C., Heyden, N., Reinartz, A., Foerderer, A., Bruelisauer, S., Linnemann, V., Hug, W., Amann, F., 2022. Possibilities of real time monitoring of micropollutants in wastewater using laser-induced Raman & fluorescence spectroscopy (LIRFS) and artificial intelligence (AI). Sensors 22 (13). https://doi.org/10.3390/s22134668.
Quintelas, C., Melo, A., Costa, M., Mesquita, D.P., Ferreira, E.C., Amaral, A.L., 2020. Environmentally-friendly technology for rapid identification and quantification of emerging pollutants from wastewater using infrared spectroscopy. Environ. Toxicol. Pharmacol. 80. https://doi.org/10.1016/j.etap.2020.103458.
Regueiro, J., Negreira, N., Hannisdal, R., Berntssen, M.H.G., 2017. Targeted approach for qualitative screening of pesticides in salmon feed by liquid chromatography coupled to traveling-wave ion mobility/quadrupole time-of-flight mass spectrometry. Food Control 78, 116–125. https://doi.org/10.1016/J.FOODCONT.2017.02.053.
Richardson, A.K., Chadha, M., Rapp-Wright, H., Mills, G.A., Fones, G.R., Gravell, A., Stürzenbaum, S., Cowan, D.A., Neep, D.J., Barron, L.P., 2021. Rapid direct analysis of river water and machine learning assisted suspect screening of emerging contaminants in passive sampler extracts. Anal. Methods 13 (5), 595–606. https://doi.org/10.1039/d0ay02013c.
Salari, M., Nikoo, M.R., Al-Mamun, A., Rakhshandehroo, G.R., Mooselu, M.G., 2022. Optimizing Fenton-like process, homogeneous at neutral pH for ciprofloxacin degradation: comparing RSM-CCD and ANN-GA. J. Environ. Manag. 317. https://doi.org/10.1016/j.jenvman.2022.115469.
Salehi Nasab, F., Ahmadi Azqhandi, M.H., Ghalami-Choobar, B., 2024. Evaluating the efficacy of recyclable nanostructured adsorbents for rapid removal of methylparaben from aqueous solutions. Environ. Res. 244, 117964. https://doi.org/10.1016/J.ENVRES.2023.117964.
Schwaiger, J., Ferling, H., Mallow, U., Wintermayr, H., Negele, R.D., 2004. Toxic effects of the non-steroidal anti-inflammatory drug diclofenac. Aquat. Toxicol. 68 (2). https://doi.org/10.1016/j.aquatox.2004.03.014.
Schymanski, E.L., Jeon, J., Gulde, R., Fenner, K., Ruff, M., Singer, H.P., Hollender, J., 2014. Identifying small molecules via high resolution mass spectrometry: communicating confidence. In Environmental Science and Technology 48 (4), 2097–2098. https://doi.org/10.1021/es5002105.
Serna-Carrizales, J.C., Zárate-Guzmán, A.I., Flores-Ramírez, R., Díaz de León-Martínez, L., Aguilar-Aguilar, A., Warren- Vega, W.M., Bailón-García, E., Ocampo-Pérez, R., 2024. Application of artificial intelligence for the optimization of advanced oxidation processes to improve the water quality polluted with pharmaceutical compounds. Chemosphere 351, 141216. https://doi.org/10.1016/J.CHEMOSPHERE.2024.141216.
Sheydaei, M., Haseli, A., Ayoubi-Feiz, B., Vatanpour, V., 2022. MoS2/N-TiO2/Ti mesh plate for visible-light photocatalytic ozonation of naproxen and industrial wastewater: comparative studies and artificial neural network modeling. Environ. Sci. Pollut. Control Ser. 29 (15), 22454–22468. https://doi.org/10.1007/s11356-021-17285-w.
Singh, Y.R., Shah, D.B., Kulkarni, M., Patel, S.R., Maheshwari, D.G., Shah, J.S., Shah, S., 2023. Current trends in chromatographic prediction using artificial intelligence and machine learning. In: Analytical Methods, vol. 15. Royal Society of Chemistry, pp. 2785–2797. https://doi.org/10.1039/d3ay00362k, 23.
Song, D., Tang, T., Wang, R., Liu, H., Xie, D., Zhao, B., Dang, Z., Lu, G., 2024. Enhancing compound confidence in suspect and non-target screening through machine learning-based retention time prediction. Environmental Pollution 347, 123763. https://doi.org/10.1016/J.ENVPOL.2024.123763.
Starkey, E., Parkin, G., Birkinshaw, S., Large, A., Quinn, P., Gibson, C., 2017. Demonstrating the value of community-based ('citizen science') observations for catchment modelling and characterisation. J. Hydrol. 548, 801–817. https://doi.org/10.1016/J.JHYDROL.2017.03.019.
Taraji, M., Haddad, P.R., Amos, R.I.J., Talebi, M., Szucs, R., Dolan, J.W., Pohl, C.A., 2017. Rapid method development in hydrophilic interaction liquid chromatography for pharmaceutical analysis using a combination of quantitative structure-retention relationships and design of experiments. Anal. Chem. 89 (3), 1870–1878. https://doi.org/10.1021/acs.analchem.6b04282.
Tetko, I.V., Gasteiger, J., Todeschini, R., Mauri, A., Livingstone, D., Ertl, P., Palyulin, V.A., Radchenko, E.V., Zefirov, N.S., Makarenko, A.S., Tanchuk, V.Y., Prokopenko, V.V., 2005. Virtual computational chemistry laboratory - design and description. J. Comput. Aided Mol. Des. 19 (6), 453–463. https://doi.org/10.1007/S10822-005-8694-Y/METRICS.
The PREMIER Project, 2024. The Database of pharmaceuticals and their environmental effects - imi-premier. https://imi-premier.eu/the-database-of-pharmaceuticals-and-their-environmental-effects/.
Tong, X., You, L., Zhang, J., He, Y., Gin, K.Y.H., 2022. Advancing prediction of emerging contaminants in a tropical reservoir with general water quality indicators based on a hybrid process and data-driven approach. J. Hazard Mater. 430. https://doi.org/10.1016/j.jhazmat.2022.128492.
Tran, N.H., Reinhard, M., Khan, E., Chen, H., Nguyen, V.T., Li, Y., Goh, S.G., Nguyen, Q.B., Saeidi, N., Gin, K.Y.H., 2019. Emerging contaminants in wastewater, stormwater runoff, and surface water: application as chemical markers for diffuse sources. Sci. Total Environ. 676, 252–267. https://doi.org/10.1016/j.scitotenv.2019.04.160.
Van Hoi, B., Vu, C.T., Phung-Thi, L.A., Thi Nguyen, T., Nguyen, P.T., Mai, H., Le, P.T., Nguyen, T.H., Thanh Duong, D., Nguyen Thi, H., Le-Van, D., Chu, D.B., 2021. Determination of pharmaceutical residues by UPLC-MS/MS method: validation and application on surface water and hospital wastewater. Journal of Analytical Methods in Chemistry 2021. https://doi.org/10.1155/2021/6628285.
VCCLAB, 2005. Virtual computational chemistry laboratory. https://vcclab.org.
Wilkinson, J.L., Boxall, A.B.A., Kolpin, D.W., Leung, K.M.Y., Lai, R.W.S., Galban-Malag, C., Adell, A.D., Mondon, J., Metian, M., Marchant, R.A., Bouzas-Monroy, A., Cuni-Sanchez, A., Coors, A., Carriquiriborde, P., Rojo, M., Gordon, C., Cara, M., Moermond, M., Luarte, T., et al., 2022. Pharmaceutical pollution of the world's rivers. Proceedings of the National Academy of Sciences of the United States of America 119 (8). https://doi.org/10.1073/PNAS.2113947119/-/DCSUPPLEMENTAL.
Wilson, B.A., Smith, V.H., deNoyelles, F., Larive, C.K., 2003. Effects of three pharmaceutical and personal care products on natural freshwater algal assemblages. Environmental Science & Technology 37 (9). https://doi.org/10.1021/es0259741.
Wu, X., Ren, J., Xu, Q., Xiao, Y., Li, X., Peng, Y., 2023. Priority screening of contaminant of emerging concern (CECs) in surface water from drinking water sources in the lower reaches of the Yangtze River based on exposure-activity ratios (EARs). Sci. Total Environ. 856. https://doi.org/10.1016/j.scitotenv.2022.159016.

Yang, Q., Ji, H., Fan, X., Zhang, Z., Lu, H., 2021a. Retention time prediction in hydrophilic interaction liquid chromatography with graph neural network and transfer learning. J. Chromatogr. A 1656, 462536. https://doi.org/10.1016/J.CHROMA.2021.462536.

Yang, Q., Ji, H., Lu, H., Zhang, Z., 2021b. Prediction of liquid chromatographic retention time with graph neural networks to assist in small molecule identification. Anal. Chem. 93 (4), 2200–2206. https://doi.org/10.1021/acs.analchem.0c04071.

Ye, J., Do, N.C., Zeng, W., Lambert, M., 2022. Physics-informed neural networks for hydraulic transient analysis in pipeline systems. Water Res. 221, 118828. https://doi.org/10.1016/J.WATRES.2022.118828.

You, L., Nguyen, V.T., Pal, A., Chen, H., He, Y., Reinhard, M., Gin, K.Y.H., 2015. Investigation of pharmaceuticals, personal care products and endocrine disrupting chemicals in a tropical urban catchment and the influence of environmental factors. Sci. Total Environ. 536, 955–963. https://doi.org/10.1016/j.scitotenv.2015.06.041.

Yu, L., Li, X., Lu, Y., Su, P., Yuan, M., Zhang, Q., Wang, L., Wang, D., Sun, M., Wang, S., 2023. Simulation monitoring of tetracyclines in wastewater based on fluorescence image processing and machine learning classifier. Sensor. Actuator. B Chem. 385. https://doi.org/10.1016/j.snb.2023.133678.

Yun, D., Kang, D., Cho, K.H., Baek, S.S., Jeon, J., 2023. Characterization of micropollutants in urban stormwater using high-resolution monitoring and machine learning. Water Res. 235. https://doi.org/10.1016/j.watres.2023.119865.

Zhang, S., Jin, Y., Chen, W., Wang, J., Wang, Y., Ren, H., 2023. Artificial intelligence in wastewater treatment: a data-driven analysis of status and trends. Chemosphere 336. https://doi.org/10.1016/j.chemosphere.2023.139163.