Environmental Science and Ecotechnology 28 (2025) 100625

Contents lists available at ScienceDirect

Environmental Science and Ecotechnology

journal homepage: www.journals.elsevier.com/environmental-science-and-ecotechnology/

Original Research

Accelerating bioelectrodechlorination via data-driven inverse design

Zhiling Li [a], Tianyi Huang [a], Fan Chen [b,*], Junqiu Jiang [a], Aijie Wang [a,c,**]

[a] State Key Laboratory of Urban-rural Water Resources and Environment, School of Environment, Harbin Institute of Technology, Harbin, 150090, PR China
[b] Shaanxi Key Laboratory of Qinling Ecological Intelligent Monitoring and Protection, School of Ecology and Environment, Northwestern Polytechnical University, Xi'an, 710129, PR China
[c] School of Civil & Environmental Engineering, Harbin Institute of Technology (Shenzhen), Shenzhen, 518055, PR China

ARTICLE INFO

Article history:
Received 30 March 2025
Received in revised form
25 September 2025
Accepted 26 September 2025

Keywords:
Microbial electrorespiration
Reductive dechlorination
Machine learning
Data-driven methodologies

ABSTRACT

Microbial electrorespiration harnesses bacteria to drive reductive dechlorination, offering a sustainable method to remediate environments contaminated with persistent chlorinated organic pollutants (COPs). However, aquifers' complex hydrogeological and hydrochemical conditions, combined with uneven COP distribution, create substantial spatial and temporal variability in biochemical reactions, environmental factors, and microbial communities. Traditional trial-and-error experiments are labor-intensive and slow, impeding the quick identification of conditions that accelerate dechlorination rates. Here we show that a machine learning framework, integrating experimental design with cathodic biofilm data, uncovers key interrelationships among environmental variables, dechlorination kinetics, electrochemical properties, and functional microbes, enabling rapid optimization of bio-electrodechlorination. Trained on literature-derived datasets using models such as extreme gradient boosting, random forest, and multilayer perceptron, this framework identifies temperature and cathode potential as primary drivers in experimental design while highlighting key biofilm genera, including *Clostridium*, *Desulfovibrio*, *Dehalococcoides*, *Pseudomonas*, *Dehalobacter*, *Arcobacter*, *Lactococcus*, and *Geobacter*. It supports inverse design to determine optimal parameters—such as cathode potential, temperature, and additives—for dechlorinating representative COPs, including tetrachloroethene, trichloroethene, and 1,2-dichloroethane, achieving reaction rate predictions with errors below 6%. This approach surpasses conventional methods by increasing efficiency, cutting costs, and accelerating bioremediation without extensive laboratory testing. By incorporating microbial community insights into predictive models, our data-driven strategy advances the scalable application of microbial electrorespiration for COP-contaminated water remediation and paves the way for broader bio-electrochemical uses in environmental engineering.

1. Introduction

Microbial electrode respiration is an effective approach for enhancing the microbial dechlorination process during bioremediation of water bodies contaminated with chlorinated organic compounds (e.g., soil, groundwater) [1]. However, the complexity of the hydrological and hydrochemical environments in groundwater, the heterogeneity of chlorinated organic pollutants (COPs), and the diversity of functional microorganisms significantly increase the difficulty of regulating microbial remediation activity [2,3]. Obtaining optimal system design and operational features typically involves conducting exhaustive or orthogonal experiments in the laboratory [4], which is often expensive, labor-intensive, and unsustainable [5]. In addition, finding similar scenarios with matching pollution characteristics, reaction conditions, and hydrogeological environments based on existing research and experience is challenging, making the development of comprehensive reaction models difficult.

Fortunately, decades of research on microbial electrode respiration for reductive dechlorination have laid a solid foundation for data-driven approaches, such as machine learning (ML). ML can

* Corresponding author.
** Corresponding author. State Key Laboratory of Urban-rural Water Resources and Environment, School of Environment, Harbin Institute of Technology, Harbin, 150090, PR China.
E-mail addresses: chenfanhit@163.com (F. Chen), waj0578@hit.edu.cn (A. Wang).

https://doi.org/10.1016/j.ese.2025.100625

analyze large volumes of data to extract insights and identify patterns and trends, thereby improving outcomes [6]. In the realm of organochloride migration and transformation, ML has demonstrated reliable efficacy in various applications, including chloride ion diffusion analysis [7], ball milling for enhanced dechlorination efficiency [8], and in situ electrochemical dechlorination analysis [9]. Compared to mechanistic reaction kinetics models, such as quantitative structure–activity relationships and quantitative structure–property relationships, ML can achieve more satisfactory outcomes by leveraging a broader spectrum of features, including environmental features and material structures. This ability provides support for the concept of goal-oriented optimized design. An optimized design begins with the expected performance based on the predicted model. It works backward to derive the model inputs, allowing for the use of ML to predict outcomes using rich and complex input features without requiring highly specific and in-depth foundational research.

Currently, ML-based inverse design methods have been employed to optimize process flows based on operational characteristics in areas such as electrochemical oxidation processes [5], membrane filtration [10], and biomass hydrogen production [11]. Although significant progress has been made in inverse design systems using ML, these systems rarely include microbial community data. In addition, when microbial communities are involved, bacterial communities are often used as a data-driven analytical tool rather than as features applied to inverse design [12,13]. Although Peng and Tan [7] made a preliminary attempt to apply inverse design for dechlorination adsorption in slag powder and fly ash concrete using the ultra-high lime aluminum method, no study has yet developed ML models to predict reaction rate constants incorporating microbial communities. Moreover, there is no research related to developing ML-based models for microbial electrochemical dechlorination, limiting the advancement of bioremediation technology for COPs. In addition, developing models that integrate experimental manipulation characteristics and biological features, rather than just considering biological data, will further enhance the inverse design process. The lack of in-depth exploration of these data limits the guiding value and wide promotion potential of ML in the practical applications of microbial electrode respiration.

Herein, we introduce an inverse design strategy that leverages ML to optimize bioelectrical dechlorination. This approach establishes robust correlations among operating features, cathodic biofilm characteristics, and dechlorination efficacy, thereby facilitating the identification of optimal dechlorination conditions. We first use a dataset derived from the literature to evaluate the performance of the models both with and without incorporating operational and biological features. To interpret the models, we employed the Shapley additive explanations (SHAP) methodology. Subsequently, we used particle swarm optimization (PSO) for inverse design to pinpoint the optimal combination of dechlorination features. The identified conditions were then experimentally validated through microbial electrorespiration dechlorination trials. This research provides valuable insights into the application of inverse design in microbial electrorespiration dechlorination and advocates for the advancement of inverse design models that incorporate biological data.

2. Materials and methods

2.1. Collection and processing of the dataset

The dataset was constructed from published studies on microbial electrorespiration dechlorination using keywords such as “bioelectrochemical,” “chlorinated pollutant,” and “dechlorination.” Through rigorous manual screening, 357 data points from 68 peer-reviewed research papers were selected to build the dataset for ML analysis. The dataset was categorized into three groups: experimental design, cathodic biofilm, and reaction rate data. Details of the feature categories, naming conventions, units, and additional information in the dataset are provided in Supplementary Text S1. Our datasets can be obtained in the Supplementary data (Database.xlsx).

Experimental design data describe the experimental operation characteristics and include 357 data entries. The key features include target pollutant, pollutant concentration (mmol kg^{-1}), strain, cathode chamber volume (mL), cathode electrode material, cathode electrode area (cm^2), anode chamber volume (mL), anode electrode material, anode electrode area (cm^2), electrode modification material, additives for auxiliary reactions, additive dosage, cathode potential (mV; standardized to standard hydrogen electrode [SHE]), *ISBio*, *ISEle*, and ambient temperature (°C). *ISBio* and *ISEle* are composed of 0 and 1, respectively, to distinguish whether an experiment includes bio-sludge or electrode stimulation. Target pollutant was represented by quantum descriptors [E_{gap}, E_{HOMO}, E_{LUMO}, μ, $f(-)_{min/max}$, $f(+)_{min/max}$, $f(0)_{min/max}$, $f(-)_{Cl,min/max}$, $f(+)_{Cl,min/max}$, and $f(0)_{Cl,min/max}$], which are calculated using ORCA and Multiwfn (details in Supplementary Text S2).

Cathodic biofilm data include the relative abundances of bacteria at the genus and species levels after the experimental reactor had reached a stable operating state, with 88 entries at the phylum level and 95 at the genus level. Bacterial species with low relative abundances (<0.6%) are categorized as “Others,” and unidentified species are filled with a value of 0. The reaction rate data describe the reaction rate (k, s^{-1}) of pseudo-first-order kinetics, obtained by fitting the reaction data. Owing to the wide range of k, it is transformed into a base-2 logarithm to enhance model accuracy [5].

2.2. Model development and optimization

Discrete and nonnumerical features were encoded to facilitate the analysis. In the actual bioelectrochemical system (BES), the categories for features such as cathode electrode material, anode electrode material, electrode modification material, and the additive of auxiliary reaction were more diverse than those described in the dataset. Using one-hot encoding can excessively expand feature dimensions, thereby reducing the performance of ML analysis. Therefore, in this study, nonnumerical features were described using target encoding. To prevent data leakage, encoding methods were applied only to the subset of training data. Missing values in the dataset were estimated using the k-nearest neighbors algorithm (Supplementary Text S3).

All ML analyses were performed using Python 3.8.7. Commonly used ML models were employed for designing, including extreme gradient boosting (XGBoost), random forest (RF), support vector machines, multilayer perceptron (MLP), process, decision tree, least squares linear regression, ridge regression, and stochastic gradient descent. The dataset was randomly split into a training set and a test set in a 9:1 ratio. A grid search (details in Supplementary Text S4) with tenfold cross-validation was employed for hyperparameter tuning of the models. As reaction rate data can only be obtained experimentally, k is the output of the model.

To observe and compare the effects of different inputs on ML performance, the impacts of the experimental design and cathodic biofilm data on k were examined separately. In this study, we tested five types of ML models.

(1) e2k: Experimental design data as input and reaction rate data as output.

(2) bP2k: Cathodic biofilm data (phylum level) as input and reaction rate data as output.
(3) bG2k: Cathodic biofilm data (genus level) as input and reaction rate data as output.
(4) ebP2k: Combined experimental design and cathodic biofilm data (phylum level) as input and reaction rate data as output.
(5) ebG2k: Combined experimental design and cathodic biofilm data (genus level) as input and reaction rate data as output.

The approach of predicting cathodic biofilm data using experimental design data was tested in advance (see Supplementary Text S5 for details). This method resulted in overfitting in the training set, with R^2 values < 0 in the test set. Predicting biological information through ML requires analyzing and providing a significant number of output features, increasing the dimensionality and sparsity of data, which leads to decreased model generalization ability and is the primary cause of underfitting [14]. Thus, predicting cathodic biofilm data using experimental design data was deemed infeasible for this dataset.

To interpret ML outputs with a black-box nature from both global and local perspectives, the SHAP method was used to calculate the marginal contributions of the features in the model outputs [15]. Input data were normalized to ensure consistency in the spatial range. Two common error analysis metrics, R^2 and root mean square error (*RMSE*), were used to evaluate ML performance. An ideal model typically exhibits a high R^2 value or a low *RMSE* [13]. Owing to differences in calculation methods, the best R^2 value and *RMSE* do not necessarily coincide with a set of model features. As *RMSE* is less sensitive to outliers, it was used as a metric for hyperparameter tuning. To further enhance model performance, methods such as principal component analysis (PCA), correlation analysis, and SHAP were employed to reevaluate the important features, followed by retraining to achieve the optimal model. Partial correlation analysis was used to further validate the SHAP results. A life cycle assessment (LCA; details in Supplementary Text S6) was conducted to analyze environmental impact during the model validation steps.

2.3. Construction and verification of an inverse design framework

To identify the optimal experimental features and biological distribution characteristics of the trained ML model, PSO was chosen based on its superior performance and faster convergence in continuous optimization problems [16,17]. This selection is supported by the comparative test results presented in Supplementary Text S7, where it outperformed other algorithms, such as genetic algorithms and simulated annealing. The settings for PSO are detailed in Supplementary Text S8. Currently, there are no effective means to regulate the bacterial community in biofilms directly. The distribution of bacterial communities in reactors, particularly in large-scale facilities such as sewage treatment plants, is continuously influenced by unpredictable toxic substances, including antibiotics, heavy metals, and plastics [18,19]. To address this, we first determined preliminary experimental conditions using the e2k model and PSO, as shown in the inverse design framework (Fig. 1). Subsequently, the eb2k model (including ebG2k and ebP2k) was constructed based on the cathode biofilm data obtained from actual experiments. The search algorithm was then reapplied to refine the experimental conditions. This iterative process continued until the ideal range of feature conditions was identified.

The inverse design was experimentally validated to enhance the credibility of the model predictions (Supplementary Text S9) using an H-type dual-chamber BES reactor. A saturated calomel

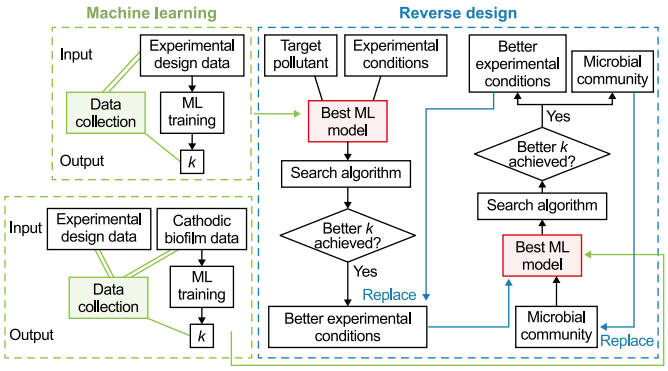

Fig. 1. Conceptual framework for the inverse design of microbial electrorespiration dechlorination using machine learning (ML). The process begins with the application of the e2k model in the absence of microorganisms. This initial step is necessary due to the inherent complexity of the microbial/sludge mixture, which hinders effective prediction. The objective is to identify a feasible, though not necessarily optimal, set of features. Subsequently, experiments are performed under these improved conditions to generate biological data. This data, combined with the corresponding experimental conditions, is then used to inform the next round of optimization using the eb2k model. The eb2k-guided search is iteratively repeated until a cost-effective and satisfactory outcome is achieved.

electrode was used as the reference electrode for monitoring electrode potential, which was converted to a SHE representation. Three chlorinated hydrocarbons were selected for reductive dechlorination experiments: tetrachloroethylene (PCE), trichloroethylene (TCE), and 1,2-dichloroethane (1,2-DCA), due to their representative environmental hazards (Supplementary Text S10). The experiment was repeated three times, and the final data were reported as average values. Further details regarding the reactor, electrodes, contaminants, and monitoring methods can be found in Supplementary Text S11.

3. Results and discussion

3.1. Descriptive analysis of the dataset

The visualization of the experimental design data distribution based on $\log_2 k$ ranking results did not reveal any obvious mathematical relationship between the features and k (Supplementary Fig. S1). This observation was quantitatively supported by the linear analysis results (Supplementary Text S12), which indicate that more complex mathematical variations exist in the data. As the dataset was compiled from various literature sources and most features did not conform to a normal distribution ($p = 0.05$, details in Supplementary Text S13), the Spearman coefficient was used for the correlation analysis. Each feature in the experimental design data showed a weak correlation with k (Supplementary Fig. S2), with absolute values < 0.350. Nevertheless, these features may still contribute differentially to the output, which means that the preliminary selection of these features is feasible.

The molecular orbital information reflected by E_{HOMO} and E_{LOMO} showed similar changes in value. Compared with the E_{HOMO} and E_{LOMO} values, E_{gap} is more relevant because it reflects the numerical difference between E_{HOMO} and E_{LOMO} and is often used to measure the ease of excitation of a molecule. Therefore, E_{HOMO} was removed from the input and replaced by E_{LOMO} and E_{gap}. As the 0 point in the cathodic potential represents no electrode stimulation, the numerical jump of the cathodic potential near the 0 point showed a similar trend to *ISEle*. To avoid model blinding caused by highly homogeneous information, μ was removed from the input. In the PCA, the positions of the experimental design data

features were far from each other (Fig. 2a), indicating a weak correlation among the features. The data points were primarily concentrated around features such as strain, cathode potential, pollutant concentration, material (anode), and the Fukui index, suggesting that these may be the primary factors influencing the dechlorination reaction. There were almost no clustered data points near f_{min} [including $f(+)_{min}$, $f(-)_{min}$, and $f(0)_{min}$]. A higher Fukui index indicates a greater likelihood of an atomic system being attacked, while f_{min}, which represents the lowest value, might lack potential analytical value in dechlorination systems. As redundant data, f_{min} could complicate the model and hinder interpretability; thus, it was removed from the input. Overall, E_{HOMO}, μ, *ISEle*, and f_{min} were considered potentially inappropriate traits and excluded from this study. The interactions between biological populations are complex, making it challenging to remove features from cathodic biofilm data reasonably. The dynamic changes in the ecological network framework suggest inseparable relationships among organisms. Hence, all features of the cathodic biofilm data were used as input.

Due to the availability of complete biological network information at each data point, no strong correlations were observed between *k* and bacterial abundance, suggesting that the dechlorination process is unlikely to be dominated by a single bacterial strain (Supplementary Text S14). The co-occurrence analysis (shown in Fig. 2c and d; Spearman's $\rho > 0.40$, $p < 0.01$) revealed that bacteria with larger node degrees exhibited lower co-occurrence with other bacteria, whereas those with smaller node degrees tended to have higher co-occurrence. This suggests that bacteria with larger node degrees may crowd out other bacteria, forcing those with smaller node degrees to collaborate. At the phylum level, bacteria such as Bacteroidetes, Chloroflexi, and Actinobacteria, and at the genus level, bacteria such as *Pseudomonas*, *Geobacter*, *Lactococcus*, and *Desulfovibrio* exhibit higher node degrees and are potential dechlorinating bacteria [19,20]. The increased data dimension in cathodic biofilm data relatively reduces the explanation of the first two principal components in the PCA compared with the experimental design data (Fig. 2b and Supplementary Fig. S3). More features were clustered near the data points, and the feature distribution became more concentrated, highlighting the direct effect of organisms on the reaction system.

The highest $\log_2 k$ values of the different pollutants (Fig. 3) were mainly concentrated in the range of −20 to −10. Hexachlorocyclohexane had the most significant dechlorination effect due to its unstable six-carbon ring and abundant chlorine substituents, followed by chloramphenicol, which had a larger molecular weight. The dechlorination rate was lowest for polychlorinated biphenyls (PCB 61), an organic matter that is extremely difficult to decompose, with a natural degradation time of approximately 50 years in soil [21]. It was better to consider using additives such as polysorbate 80, which had been reported to improve dechlorination efficiency [22,23]. Analysis of the $\log_2 k$ ranking revealed that compounds with more chlorine atoms and a more dispersed chlorine distribution tended to exhibit greater resistance to dechlorination. For example, compared with 2,4,6-trichlorophenol, 2,3,4,5-tetrachlorobiphenyl had more chlorine atoms distributed in different positions, making dechlorination more difficult. Pollutants with ring structures, such as polychlorinated biphenyls and hexachlorocyclohexane, are more challenging to dechlorinate because of their stability. In particular, compounds with greater polarity, such as trichloroacetic acid, although relatively simple in molecular structure, showed reduced dechlorination performance due to their high polarity.

3.2. Verification in ML

Preliminary tests using the ebG2k model (Supplementary Fig. S4) revealed that R^2 exceeds 0 when the dataset has more than 100 points and stabilizes at approximately 0.8 beyond 200 points. Previous ML studies on degradation with similar datasets (e.g., 215 [24] and 315 [5] points) also fall within this range, indicating that our dataset is likely sufficient to support reliable conclusions and performance. Supplementary Fig. S5 shows the performance differences of each model in the test and training sets based on the results of the hyperparameters (Supplementary Text S15). Ordinary multivariate linear estimators (e.g., ridge and least squares linear regressions) performed poorly because of the complex relationships within the data. The methods that performed exceptionally well across the different modes were XGBoost, RF, and MLP, with RF and XGBoost being particularly notable. RF achieved the best results in the e2k and bG2k modes, while XGBoost excelled in the ebP2k and ebG2k modes, displaying similar performances in both the test and training sets. The bP2k mode, which contained narrow-space bacterial data at the phylum level, enhanced MLP's ability to learn sparse information. However, neural networks generally underperform compared with tree models in small-scale datasets (data <1000) [25], resulting in poorer performance for bP2k compared with other types. XGBoost enhances performance by iteratively training multiple weak learners and continuously optimizing their combination, making it well-suited for processing high-dimensional data in ebP2k and ebG2k [5]. RF captures linear relationships effectively by constructing multiple decision trees from training samples, which was particularly advantageous for the e2k and bG2k datasets with fewer features and samples, allowing it to extract useful information more effectively. These best-performing models were used separately for training of different input frameworks.

The initial SHAP value analysis results (Supplementary Fig. S6) were used to optimize e2k. Owing to insufficient data on single bacterial strains in the dataset, strain was initially considered the least important feature. However, due to the differences in metabolic pathways between mixed and single bacteria, strain remained an important discriminative label. Features were sequentially removed in order of increasing importance. A minor improvement in e2k was achieved when $f(0)_{Cl,max}$ was excluded, resulting an optimized *RMSE* (test set) from 1.762 to 1.750. When both $f(0)_{max}$ and $f(0)_{Cl,max}$ were excluded, the *RMSE* (test set) was 1.756. Further feature deletion was not optimal. The exclusion of $f(0)_{max}$ yielded poor results in the ebP2k and ebG2k scenarios because of the changes in input. As the exclusion of $f(0)_{max}$ did not result in significant optimization and to maintain input consistency as much as possible, the original input features were retained for analysis.

Among these models with different input and output (Fig. 4), bP2k produced noticeably poorer results due to the sparse and coarse-grained nature of microbial information at the phylum level. This is primarily due to the limited resolution of microbial information at the phylum level, which tends to obscure ecologically and functionally distinct taxa under the same broad classification. For instance, both *Clostridium* and *Dehalobacter* belong to the Firmicutes phylum and play vastly different ecological roles. *Clostridium* is commonly associated with fermentative metabolism, whereas *Dehalobacter* is a known organohalide-respiring bacterium with strong dechlorination capabilities. Such key functional differences are masked when only phylum-level data are used, making it difficult for the model to capture essential biological mechanisms related to pollutant degradation. By

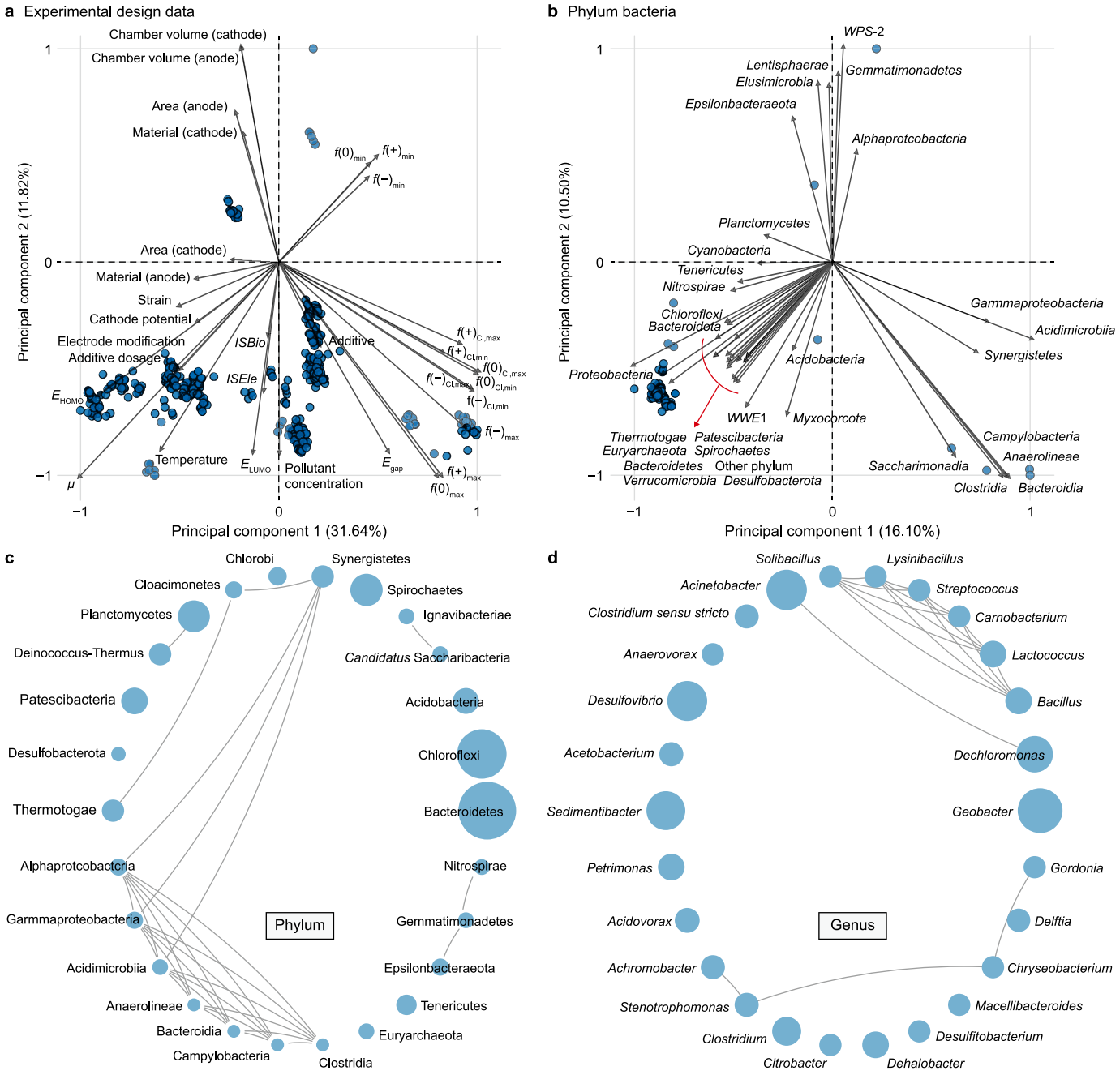

Fig. 2. Basic description of the dataset. **a**–**b**, Principal component analysis of experimental design data (**a**) and cathodic biofilm data at the phylum level (**b**). Due to the extensive characteristic information, only a subset of strains is labeled in panel **b**. **c**–**d**, Co-occurrence analysis for the frequently observed phylum (**c**) and genus (**d**) in the dataset (Spearman's $\rho > 0.4$, $p < 0.01$). Node size is scaled according to node degree.

contrast, the models that incorporated genus-level features demonstrated significantly improved performance. bG2k achieved an R^2 of 0.915, and ebG2k reached an R^2 of 0.873. Most data points clustered closely around the 1:1 line, indicating strong predictive capabilities. These results highlight the importance of finer taxonomic resolution in capturing key ecological and functional differences in microbial communities.

When quantum descriptors were not used (Supplementary Fig. S7), the average *RMSE* across all modes increased by 0.342, indicating that constructing pollutant features using quantum descriptors provides more accurate predictions. Limited by the data set, different inputs provide different numbers of data points for the model. Rich databases generally led to excellent performance, but bG2k, ebP2k, and ebG2k, which contained biological information, achieved better performance than e2k, with more data points. This suggests that microorganisms fundamentally mediate the bioelectric respiration system. In other words, the *k* value essentially depends on biological development and respiratory conditions, with operating conditions, such as electrode potential, as auxiliary indirect regulatory means. In addition, the inclusion of experimental data alleviated the sparse spatial distribution issue of gate data, improving the previously poor *RMSE* of ebP2k from 2.671 to 1.776. This demonstrates that comprehensive consideration of experimental operations and biological

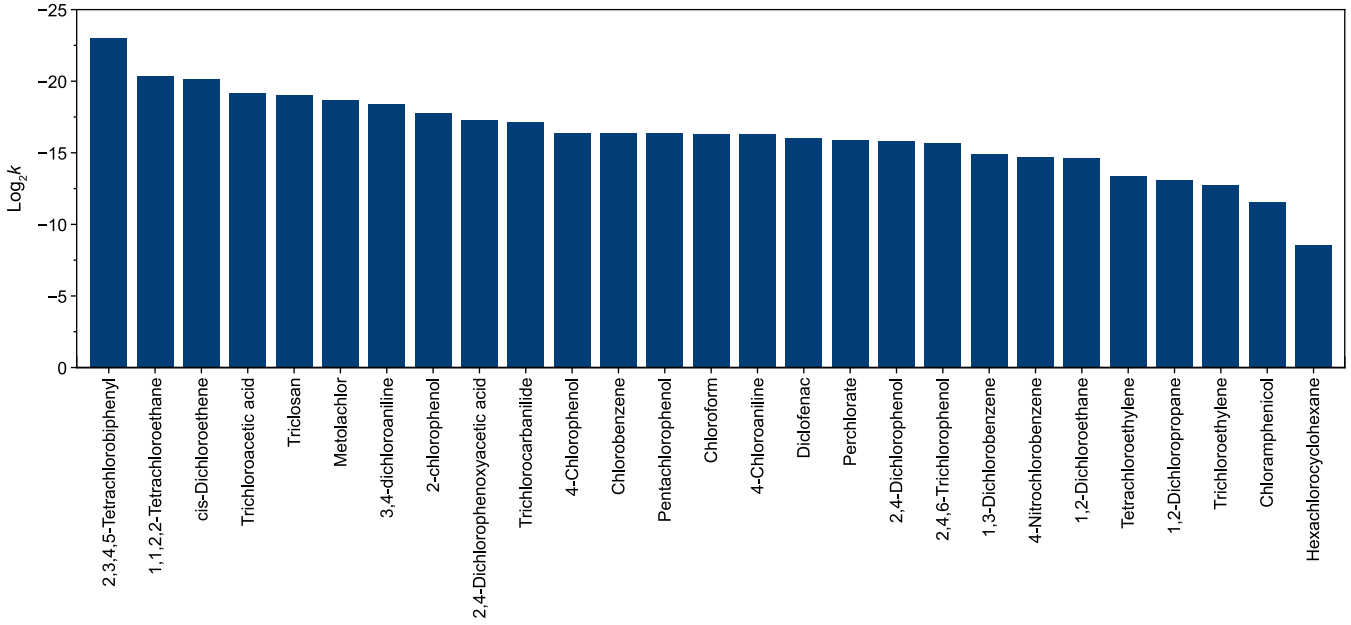

Fig. 3. The highest $\log_2 k$ for each pollutant based on the dataset, sorted in descending order.

information contributes to a more robust ML learning system. Overall, we conclude that our ML training effect is acceptable, except for bP2k.

3.3. Feature importance analysis

SHAP value was calculated to comprehensively present feature-importance for each mode (Fig. 5a–c [including e2k, bG2k, and ebG2k, respectively] and Supplementary Fig. S10 [including bP2k and ebP2k]). For the experimental operation characteristics in e2k, *ISBio* showed a bimodal SHAP value distribution, where the addition of biological elements (red data points) significantly promotes the dechlorination system. This is corroborated by correlation results (Supplementary Fig. S2). Temperature was identified as the most critical feature, with significant impacts on biological dechlorination efficacy [26,27]. This ranking also applies to ebP2k and ebG2k (Fig. 5c). In the e2k mode, the secondary experimental feature is the cathode potential. Electrodes are insoluble electron donors and can be provided sustainably and controllably to dechlorinating bacteria by applying an appropriate potential, effectively avoiding side reactions caused by organic matter addition and improving electron utilization efficiency. The SHAP data points were concentrated around the 0-axis, with a trailing effect in the negative region. High potential exhibited negative SHAP values, highlighting the bioactivation effect of low potential. Excessively low potential failed to provide sufficient electrons to microorganisms, while excessively high potential led to hydrogen-driven side reactions that hindered electron utilization by microorganisms [28]. Thus, bioelectric respiratory dechlorination usually has an optimal potential range. Other experimental features, such as electrode surface area and additive dosage, gained importance in the eb2k model, indicating stronger interactions with microorganisms. For instance, increasing the electrode surface area provides more space for biofilm attachment and has been shown to increase the proportion of electroactive microorganisms in the reaction system [29], as reflected in the SHAP distribution. As cathode materials are carriers for microbial growth, their SHAP values were generally positive (Supplementary Fig. S8a), indicating a favorable contribution to system performance. Among these materials, carbon fiber was identified as the most influential, followed by graphite (Supplementary Fig. S8b). Both materials exhibit excellent electrical conductivity, facilitating electron transfer, and feature porous structures with large specific surface areas that promote microbial adhesion and growth. However, the marginal benefit of further increasing the electrode area on degradation performance tends to diminish (Supplementary Fig. S9), possibly due to factors such as microbial distribution saturation and limitations in optimizing electron transfer pathways.

Several reported dechlorinating respirators (at the genus level) were ranked highly in our SHAP analysis, including *Clostridium*, *Desulfovibrio*, *Dehalococcoides*, *Pseudomonas*, *Dehalobacter*, *Arcobacter*, *Lactococcus,* and *Geobacter* [30,31]. The category "Other" in genus was emphasized in the biological model (bP2k and bG2k) because each sample contained these feature data. *Clostridium* and *Desulfovibrio* had clear, two-headed SHAP distributions. Although both are electroactive bacteria that can use electrodes for respiration, the average dechlorination capacity increase of the system tends toward low-abundance *Clostridium* and high-abundance *Desulfovibrio*, similar to the results reported by Lin et al. [32]. *Geobacter* was ranked higher in ebG2k, a classic dechlorinating genus that can use electrodes as the sole electron donor for reductive dichlorination [33]. Important quantum features that represent pollutant characteristics were identified as $f(+)_{\max}$ and $f(+)_{\mathrm{Cl,\ min}}$, indicating that the Fukui index is more indicative of dechlorination reaction difficulty than E_{gap}.

The SHAP heatmap (Fig. 5d, including e2k; Supplementary Fig. S10, including bP2k and ebP2k; and Supplementary Fig. S11, including ebG2k and bG2k) shows the comprehensive contribution of each feature to *k*. There were strong spatial contributions and short-lived features in the high-range *k* value interval, such as additive dosage (e2k), *Desulfovibrio* (bG2k/ebG2k), *Geobacter* (ebG2k), and material (cathode), which might be potential improvements for dechlorination systems. The total contribution of the experimental features surpassed that of the quantum descriptors and biological features (Supplementary Fig. S12),

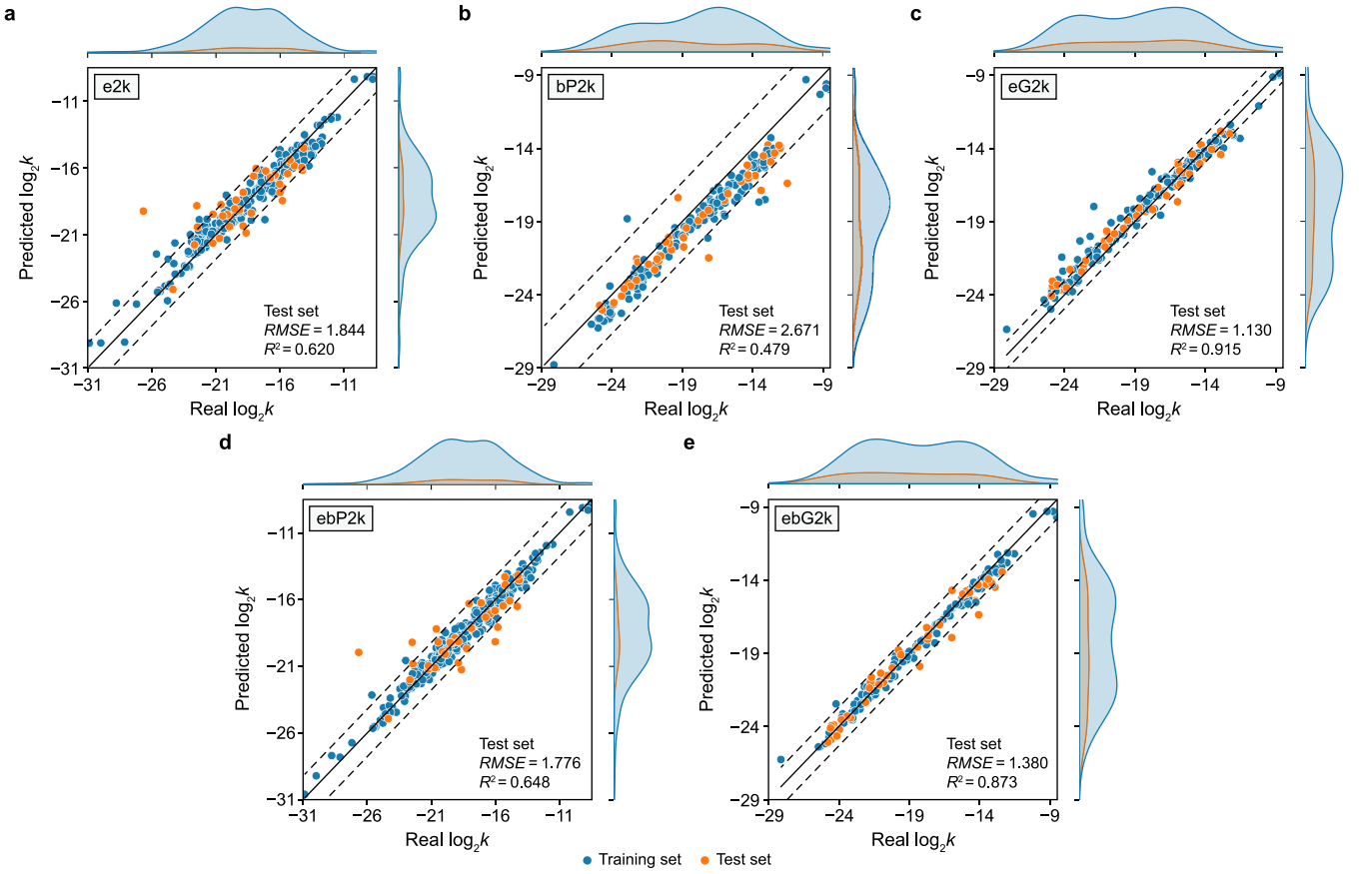

Fig. 4. Prediction performance of ML models under different configurations: e2k (**a**), bP2k (**b**), bG2k (**c**), ebP2k (**d**), and ebG2k (**e**). The solid diagonal line represents perfect prediction, while the dashed lines denote the ± root mean square error on the test set. The distribution of actual and predicted output values is shown on the upper and right sides of each panel.

emphasizing the importance of experimental strategies and methods. The one-dimensional partial dependence and SHAP feature dependence plots for e2k (Supplementary Fig. S13–S14) and ebG2k (Supplementary Fig. S15–S16) illustrate the effects of variations in experimental features on $\log_2 k$. Both analyses indicate that these effects are not statistically significant. While some differences were observed, SHAP dependence plots provide more detailed insights into data concentration than the one-dimensional partial dependence plots. Based on the SHAP value analysis results, we summarized the range of experimental features that can achieve higher k values (Supplementary Table S1). The SHAP value analysis results also provided a joint analysis perspective with cathode potential, intuitively showing the impact of potential distribution. For example, weak electrical stimulation (0–400 mV) suits pollutants with $f(+)_{\max}$ of approximately 0.150 (e.g., 2-chlorophenol and 4-nitrochlorobenzene), while a stronger potential (>400 mV) was more conducive to the removal of pollutants with an $f(+)_{\max}$ of approximately 0.100 (e.g., chloramphenicol and triclocarban). These potential conclusions have been confirmed by previous studies [34,35]. SHAP interpretation is based on ML models rather than real substances, so SHAP analysis provides functional rather than substance-based conclusions, limiting these recommendations to laboratory environments based on our dataset.

3.4. Inverse design and experimental verification

Building upon the previously mentioned experimental operating conditions (Supplementary Text S11) and the inverse design process (Fig. 1), we tested the inverse design effects of PCE, TCE, and 1,2-DCA using our trained ML model. We first fixed some basic features (Supplementary Table S2), such as the quantum description of pollutants. We then determined other features (Supplementary Table S3) to quickly narrow the adjustment range with e2k. Chamber volume tends to form a mutually reinforcing relationship with area (cathode/anode) [36], which makes the two features always converge at a larger value (chamber volume >300 mL, area of cathode/anode >200 mL). To prevent this phenomenon, we fixed the chamber volume to 100 mL and the area (cathode/anode) to 42.41 cm^2 (diameter and height, both 3 cm). The search results (Table 1) were verified through parallel experiments, with a relative error of less than 10%. Control experiments with 9 sets of random features configurations (Supplementary Table S4) prove that this result has local optimality. The rapid determination of the optimal preliminary operating factors without any preliminary experiments initially confirmed the ML-based reverse design.

Owing to the large number of features to be further adjusted, given the significant impact of the cathode potential on bioelectric respiratory dechlorination [31,37,38] and the ease of temperature control, these two features were selected as primary regulatory factors for further optimization. To address other features as outlined in Supplementary Table S2, we employed PSO to optimize the results. The search results based on the e2k model for PCE, TCE, and 1,2-DCA yielded cathode potentials of −510.8, −456.0, and −458.1 mV, respectively, with temperature results

concentrated at approximately 23.2 °C. Dechlorination performance tests at 23.0 ± 10.0 °C under the same cathodic potential (Supplementary Fig. S17) showed a lower reaction rate, which indicates that the chosen temperature may be suitable for microbial development in this range. We then set the temperature of the experimental environment to 23.0 ± 3.0 °C and the cathode potentials for PCE, TCE, and 1,2-DCA to −510.0, −450.0, and −450.0 mV, respectively, for e2k experimental verification. These features were predicted using the e2k model, with an *RMSE* of 0.977 and a relative error of 5.032% (Fig. 6a).

For a unified comparison, temperature and other conditions were fixed, and cathode potential was reoptimized using the eb2k model based on acquired cathode biofilm data. The predictive performance of the ebP2k model was inferior to that of ebG2k (Supplementary Table S5), consistent with previous model evaluations, indicating that genus-level data are more suitable for inverse design. According to the PSO results of ebG2k (Supplementary Table S6), cathode potentials were set to −260.0, −280.0, and −270.0 mV, respectively. The $\log_2 k$ prediction accuracy of ebG2k improved compared with e2k, with an *RMSE* of 0.843 and a relative error of 4.696% (Fig. 6b). The microbial abundance information at the phylum and genus levels under different potential conditions is shown in Supplementary Fig. S18 and Fig. 6d, respectively. Further optimization using ebG2k (Supplementary Fig. S19) shows limited improvement in PCE and TCE reaction rates, and a relatively low reaction rate for 1,2-DCA. This may be due to the experimental errors within the narrow adjustment range or similar solutions from the PSO random step search. The initial ebG2k potential conditions seem sufficient for achieving satisfactory dechlorination effects. Therefore, we stopped further optimization at this point.

Additional experiments at −100.0 and −600.0 mV confirmed the reliability of these results. Among the four potential settings, the ebG2k-searched potential conditions showed the optimal k value, followed by e2k, indicating that our ML model had achieved a locally optimal solution. Using the bG2k model for k prediction yielded a relative error of 1.562% (Supplementary Table S7), which suggests that preliminary inferences of k based on genus-level data are feasible. Among the five models, bG2k exhibited the best performance. However, owing to the lack of direct methods for effectively balancing suitable abundance systems, the regulatory process from e2k to eb2k remains practically significant. The relative error of bP2k was 17.923%, but the relative error of the ebP2k model, which combines experimental feature regulation, was reduced to 6.304%, which indicates that mixed data of biological and experimental conditions helped the ML model search for better results. The relative errors (Supplementary Text S16) in all experimental results for e2k and ebG2k were within 6%, comparable with the inverse design of the electrochemical oxidation process for water purification (<5%) [5] and anaerobic digestion for CH_4-rich biogas production (<9%) [39]. From an application perspective, we thought this error was acceptable. We further tested the model on newer relevant literature [40] (data not included in our dataset). The relative error of the e2bGk model was 4.553%, further demonstrating the operability of ML.

In the experimental results (Fig. 6c), the biodegradability of the three chlorinated hydrocarbons follows the order 1,2-DCA, TCE, and PCE. The numerical orders of E_{gap} and bond dissociation energy also support this view (Supplementary Fig. S20). The increased difficulty of dechlorination for TCE and PCE may be due to their double bond structures and the limited types of bacteria capable of reducing trans-1,2-dichloroethene (the dechlorination product of TCE and PCE) [41]. While the e2k model learned this pollutant information, the ebG2k model (Supplementary Fig. S21) failed to capture similar patterns, which indicates that our data scale still limits the model's spatial differentiation of different pollutants. Moreover, the complex distribution of biological information also significantly impacted the model. For example, *Geobacter* is the dominant genus in 1,2-DCA, while *Lactococcus* and *Bacillus* mainly regulate PCE and TCE (Fig. 6d). *Desulfovibrio* showed some enrichment in all three pollutants. These relationships between bacteria and cathode potential were similar to the SHAP dependency results of ebG2k (Supplementary Fig. S22). Unfortunately, the search for biological features was less

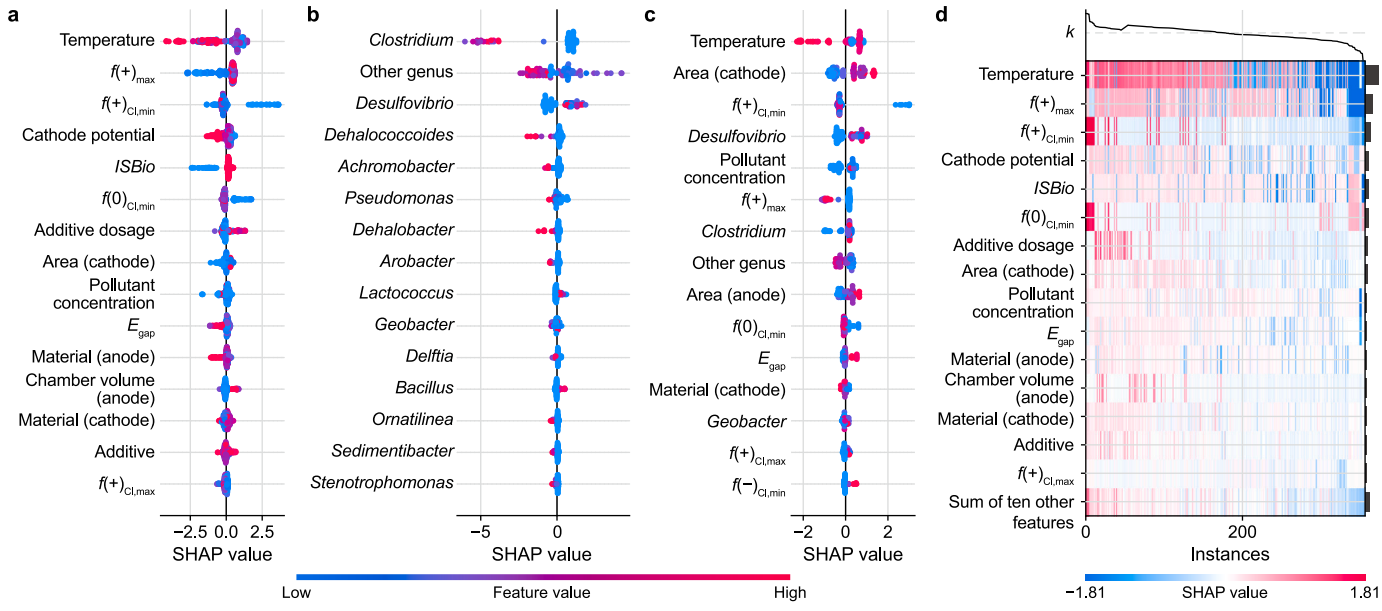

Fig. 5. Eigenvalue interpretation analysis based on Shapley additive explanations (SHAP). **a**–**c**, The distribution of SHAP values for each feature in the e2k (**a**), bG2k (**b**), and ebG2k (**c**) modes. The *x*-axis represents data points ordered by the total SHAP values of all features. **d**, The heat maps of SHAP values for each feature in the e2k mode. The learning mode of each panel highlighted the top 15 features. The bar plot adjacent to the *y*-axis indicates the mean absolute SHAP value (|SHAP|) for each feature, sorted in descending order. The value above the heatmap reflects the cumulative SHAP value for all features, indicating their overall contribution to the model's output magnitude.

Table 1
The search results for starting the experiment.

Features	PCE	TCE	1,2-DCA
ISBio	1	1	1
ISEle	1	1	1
Cathode potential (mV)	−510.8 ± 13.3	−456.0 ± 22.7	−458.1 ± 18.1
Temperature (°C)	21.4 ± 0.5	22.7 ± 0.7	21.7 ± 1.0
Pollutant concentration (mmol kg^{-1})	0.36	0.94	2.10
Additive	Acetate	Acetate	Acetate
Additive dosage (mmol kg^{-1})	3.40	4.90	4.50
Cathode material	Graphite	Graphite	Graphite
Anode material	Carbon cloth	Graphite	Graphite
Anode area (cm^2)	42.41	42.41	42.41
Ave. experimental $\log_2 k$	−14.179	−14.328	−15.089
Ave. predicted $\log_2 k$	−15.166	−14.784	−14.417

Note: Ave. represents the average value of multiple searches.

successful in our dataset. The ML results overemphasized the role of non-dehalogenating bacteria (Supplementary Fig. S23), such as *Desulfovibrio* and *Carnobacterium*, and the abundance distribution of the bacterial community was inconsistent with the findings of existing studies [32,42]. This was mainly attributed to the dataset being too small to support a drastic increase in search breadth.

Conducting traditional experiments to identify optimal conditions is economically, energetically, and time-consuming. Using ML-assisted inverse design, within the four potential ranges used, could reduce half waste with two potential experiments. The LCA results indicate that using ML in our experiments could help reduce global warming potential (GWP100) and non-renewable energy use (NREU) by 14.90 kg CO_2 eq and 29.40 kJ, respectively (Fig. 6e), representing the minimum resource savings. Next, we replaced the additive acetate with sodium bicarbonate (10.00 mmol kg^{-1}) and glucose (1.67 mmol kg^{-1}) in PCE and repeated the operations. The final search potential was similar to that when acetate was the additive. When setting the cathode potential to −270.0 mV, the relative error between the predicted results and experimental values (Supplementary Table S8) was 3.954%, indicating the impact of ranking different additives. If ML had been used for inverse design from the start — targeting all three features simultaneously — the resource savings could have reached 83.333%, calculated as 1−50% per feature divided across the three features.

4. Perspective

The ML-based inverse design method effectively overcomes the limitations of traditional trial-and-error approaches, utilizing bioinformatics to inform decision-making in experimental operations. This approach provides a novel method for rapid bioremediation of water bodies contaminated with COPs. The five inverse design frameworks (e.g., e2k and ebG2k) we constructed in

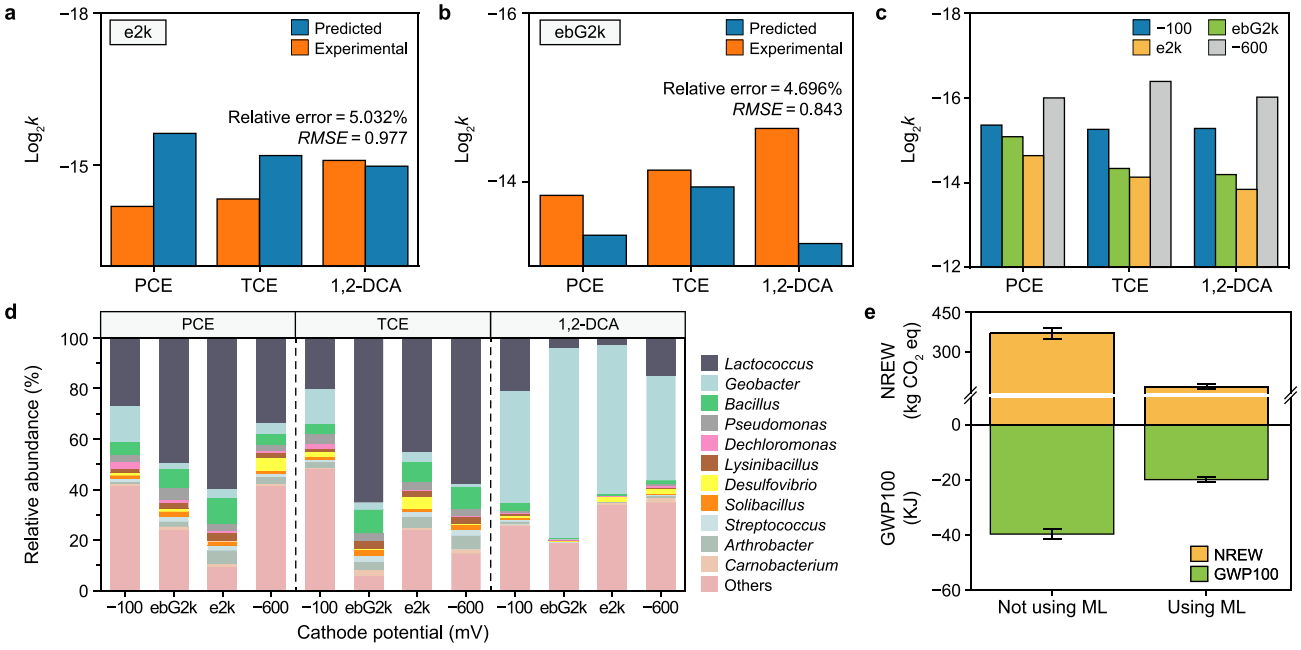

Fig. 6. Results of the inverse design and experiments. **a**–**b**, Comparison of experimental and predicted outcomes based on approximate values identified through particle swarm optimization under the e2k (**a**) and ebG2k (**b**) models. Experimental temperature was maintained at 23.0 ± 3.0 °C. For tetrachloroethylene (PCE), trichloroethylene (TCE), and 1,2-dichloroethane (1,2-DCA), the cathode potentials applied in the e2k model are −510.0, −450.0, and −450.0 mV, respectively, while those in the ebG2k model are −260.0, −280.0, and −270.0 mV. **c**, Experimental $\log_2 k$ values of pollutants at different potentials. **d**, Relative abundance of microbial taxa in the cathodic biofilm at the genus level. **e**, Comparison of life-cycle assessment outcomes for inverse design experiments conducted with and without the application of machine learning (ML), comparative indicators include lobal warming potential (GWP100) and non-renewable energy use (NREU).

Text 2.2 are suitable for different scenarios. e2k is applicable when the biological information necessary for the preliminary determination of the operating conditions of the initiation system (experimental design data) is unknown. After obtaining community data through biological information tools, either ebP2k or ebG2k can be used to integrate biological results and perform iterative optimization on the experimental design data based on the workflow (Fig. 1). bP2k or bG2k directly connects community information with *k*. bP2k and bG2k reflect the essential operations in biological systems and provide better predictive performance (e.g., bG2k in Fig. 4). However, precise regulation of biological networks is currently expensive and time-consuming, making the combination of experimental design data as input with ebP2k or ebG2k the recommended method, which was used in this study. Our ML framework (Fig. 1) is adaptable and not limited to dechlorinated metabolites, as it can also be applied to other pollutants and bioelectric respiration processes. The differences are mainly reflected in the operational characteristics of each system, and these databases still need to be sorted out; the contribution of more researchers are necessary to promote the application and development of the framework in a broader range of scenarios.

Despite these advancements, our study is limited by the quality and completeness of the dataset used, which may have affected the model's performance in certain cases. For example, when using e2k to predict the dataset from Chen et al. [40], the relative error reached 11.520% (Supplementary Fig. S24), which is likely due to the differences in experimental conditions, such as continuous flow and alternating open/closed circuits. Additionally, important features, such as agitation method and speed, are often not reported in literature, making it challenging to incorporate them into the model.

Microbial community complexity is another major challenge. The use of 16S rRNA data provides limited taxonomic resolution and often fails to capture the contribution of low-abundance but functionally critical taxa. Moreover, literature-derived datasets frequently lack detailed annotations for “Other” microbial groups, causing the model to rely overly on well-characterized species and miss interactions involving unknown taxa. To address these limitations, we proposed the following strategies for future work (1) incorporating more “negative” samples (i.e., systems without dechlorination activity): to improve the ability of the model to distinguish relevant features and reduce overfitting; (2) integrating metagenomic data to identify functional genes (e.g., *rdhA*, *pceA*, and *vcrA*) and their distributions across samples, enabling high-resolution mapping of species–enzyme–pathway relationships that supplement high-resolution mapping of community composition data; and (3) expanding the dataset through systematic data acquisition, including standardized experimental records and direct sequencing repositories (e.g., National Center for Biotechnology Information Sequence Read Archive [NCBI], the Kyoto Encyclopedia of Genes [KEGG] and Genomes, and the Gene Expression Omnibus [GEO]), to enrich the diversity and completeness of feature sets. This would enable the ML model to better generalize across systems and improve its capacity for inverse design under varying environmental and operational conditions.

Moreover, incorporating electrodynamic systems into the ML framework would facilitate the development of a unified model for biodechlorination prediction. This integration would enable cross-system comparisons between bioelectrochemical and electrodynamic setups, thereby enhancing both the interpretability and optimization of dechlorination processes through ML. In addition, the inclusion of spatial and long-term temporal data supports the generation of persistent auxiliary predictions, further improving the applicability of the framework to real-world site restoration efforts.

5. Conclusion

This study developed ML models to address a critical gap in microbial electrorespiration dechlorination and propose a machine-driven bioelectrochemical framework that integrates biological data to analyze the interrelationships between environmental factors, dechlorination efficiency, electrochemical properties, and functional microbial communities. We verified the inverse design function of the model through random experiments with three typical COPs. After iterative searches, the reaction rate prediction achieved a relative error within 6% and an *RMSE* of less than 1. This demonstrates that satisfactory operating features can be determined using ML without additional practical experiments. Furthermore, the ML framework proposed in this study is adaptable for designing and optimizing other BESs. It can also be applied to develop biosystem architectures tailored to specific needs or to lay the groundwork for creating more generalized bioinformatics metamodels.

CRediT authorship contribution statement

Zhiling Li: Methodology, Conceptualization, Writing - Original Draft. **Tianyi Huang:** Methodology, Data Curation, Visualization, Writing - Original Draft. **Fan Chen:** Investigation, Visualization. **Junqiu Jiang:** Data Curation, Supervision. **Aijie Wang:** Validation, Funding Acquisition, Software.

Data availability

Data will be made available on request.

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.

Dr. Aijie Wang, the Executive Editor of *Environmental Science and Ecotechnology*, was not involved in the editorial review or the decision to publish this article.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (No. 52370163), National Key Research and Development Program of China (No. 2022YFA0912501), and State Key Laboratory of Urban-rural Water Resource & Environment (Harbin Institute of Technology) (No. 2025DX12).

Appendix. ASupplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ese.2025.100625.

References

[1] G. Gan, S. Fan, X. Li, J. Wang, C. Bai, X. Guo, M. Tade, S. Liu, Nature of Intrinsic Defects in carbon materials for electrochemical dechlorination of 1,2-dichloroethane to ethylene, ACS Catal. 11 (2021) 14284–14292.
[2] J. Zhao, M. Rao, H. Zhang, Q. Wang, Y. Shen, J. Ye, K. Feng, S. Zhang, Evolution of interspecific interactions underlying the nonlinear relationship between active biomass and pollutant degradation capacity in bioelectrochemical systems, Water Res. 274 (2025) 123071.
[3] X. Su, M. Xie, Z. Han, Y. Xiao, R. Wang, C. Shen, M.Z. Hashmi, F. Sun, Resuscitation-promoting factor accelerates enrichment of highly active tetrachloroethene/polychlorinated biphenyl-dechlorinating Cultures, Appl. Environ. Microbiol. 89 (2023).
[4] C. Chen, G.X. Gu, Generative deep neural networks for inverse materials design using Backpropagation and active learning, Adv. Sci. 7 (2020)

1902607.
[5] Y. Sun, Z. Zhao, H. Tong, B. Sun, Y. Liu, N. Ren, S. You, Machine learning models for inverse design of the electrochemical oxidation process for water purification, Environ. Sci. Technol. 57 (2023) 17990–18000.
[6] R. Wang, S. Zhang, H. Chen, Z. He, G. Cao, K. Wang, F. Li, N. Ren, D. Xing, S.-H. Ho, Enhancing Biochar-based Nonradical Persulfate Activation using data-driven Techniques, Environ. Sci. Technol. 57 (2023) 4050–4059.
[7] J. Peng, H. Tan, Optimization of dechlorination experiment design using Lightweight deep learning model, Comput. Intell. Neurosci. 2022 (2022) 1–10.
[8] J. Lu, S. Borjigin, S. Kumagai, T. Kameda, Y. Saito, T. Yoshioka, Machine learning-based discrete element reaction model for predicting the dechlorination of poly (vinyl chloride) in NaOH/ethylene glycol solvent with ball milling, Chem. Eng. J. Adv. 3 (2020) 100025.
[9] J. Lee, J. Im, U. Kim, F.E. Löffler, A data mining approach to predict in situ detoxification potential of chlorinated Ethenes, Environ. Sci. Technol. 50 (2016) 5181–5188.
[10] C.S.H. Yeo, Q. Xie, X. Wang, S. Zhang, Understanding and optimization of thin film nanocomposite membranes for reverse osmosis with machine learning, J. Membr. Sci. 606 (2020) 118135.
[11] X. Zheng, G. Jia, Active learning based reverse design of hydrogen production from biomass fuel, Fuel 357 (2024) 129948.
[12] Y. Ma, D. Rui, H. Dong, X. Zhang, L. Ye, Large-scale comparative analysis reveals different bacterial community structures in full- and lab-scale wastewater treatment bioreactors, Water Res. 242 (2023) 120222.
[13] J. Shao, S. Huang, Y. Chen, J. Qi, Y. Wang, S. Wu, R. Liu, Z. Du, Satellite-based global Sea surface Oxygen mapping and interpretation with Spatiotemporal machine learning, Environ. Sci. Technol. 58 (2024) 498–509.
[14] Medina R. Hernández, S. Kutuzova, K.N. Nielsen, J. Johansen, L.H. Hansen, M. Nielsen, S. Rasmussen, Machine learning and deep learning applications in microbiome research, ISME COMMUN 2 (2022) 98.
[15] X. Ma, Y. Tang, C. Wang, Y. Li, J. Zhang, Y. Luo, Z. Xu, F. Wu, S. Wang, Interpretable XGBoost-SHAP model predicts Nanoparticles Delivery efficiency based on Tumor Genomic Mutations and Nanoparticle properties, ACS Appl. Bio Mater. 6 (2023) 4326–4335.
[16] B. Li, M. Khayatnezhad, Modified artificial neural network based on developed snake optimization algorithm for short-term price prediction, Heliyon 10 (2024) e26335.
[17] X. Zhao, Y. Ji, X. Ning, Accelerometer calibration based on improved particle swarm optimization algorithm of support vector machine, Sensor Actuator Phys. 369 (2024) 115096.
[18] Y. Li, J. Wang, C. Lin, M. Lian, M. He, X. Liu, W. Ouyang, Occurrence, removal efficiency, and emission of antibiotics in the sewage treatment plants of a low-urbanized basin in China and their impact on the receiving water, Sci. Total Environ. 921 (2024) 171134.
[19] L.N.A. Sackey, J. Koomson, R. Kumi, A.A. Hayford, P. Kayoung, Assessing the quality of sewage sludge: CASE study of the Kumasi wastewater treatment plant, Heliyon 9 (2023) e19550.
[20] F. Chen, B. Liang, Z.-L. Li, J.-Q. Yang, C. Huang, M. Lyu, Y. Yuan, J. Nan, A.-J. Wang, Bioelectrochemical assisted dechlorination of tetrachloroethylene and 1,2-dichloroethane by acclimation of anaerobic sludge, Chemosphere 227 (2019) 514–521.
[21] S. Subramani, A. Bagde, A. Balke, T. Chakrabarti, A. Bafana, Strategy for remediation of polychlorinated biphenyls-contaminated soil through redox Management based on Electronegativity of the contaminants, Bull. Environ. Contam. Toxicol. 112 (2024) 22.
[22] H. Wan, X. Yi, X. Liu, C. Feng, Z. Dang, C. Wei, Time-dependent bacterial community and electrochemical characterizations of cathodic biofilms in the surfactant-amended sediment-based bioelectrochemical reactor with enhanced 2,3,4,5-tetrachlorobiphenyl dechlorination, Environ. Pollut. 236 (2018) 343–354.
[23] H. Yu, H. Wan, C. Feng, X. Yi, X. Liu, Y. Ren, C. Wei, Microbial polychlorinated biphenyl dechlorination in sediments by electrical stimulation: the effect of adding acetate and nonionic surfactant, Sci. Total Environ. 580 (2017) 1371–1380.
[24] M. Yin, X. Zhang, F. Li, X. Yan, X. Zhou, Q. Ran, K. Jiang, T. Borch, L. Fang, Multitask deep learning enabling a Synergy for Cadmium and methane Mitigation with biochar Amendments in Paddy soils, Environ. Sci. Technol. 58 (2024) 1771–1782.
[25] Z. Mai, H. Shen, A. Zhang, H.Z. Sun, L. Zheng, J. Guo, C. Liu, Y. Chen, C. Wang, J. Ye, L. Zhu, T.-M. Fu, X. Yang, S. Tao, Convolutional neural networks facilitate process Understanding of Megacity Ozone temporal variability, Environ. Sci. Technol. 58 (2024) 15691–15701.
[26] D. Kong, B. Liang, D.-J. Lee, A. Wang, N. Ren, Effect of temperature switchover on the degradation of antibiotic chloramphenicol by biocathode bio-electrochemical system, J. Environ. Sci. 26 (2014) 1689–1697.
[27] B. Liang, D. Kong, J. Ma, C. Wen, T. Yuan, D.-J. Lee, J. Zhou, A. Wang, Low temperature acclimation with electrical stimulation enhance the biocathode functioning stability for antibiotics detoxification, Water Res. 100 (2016) 157–168.
[28] F. Chen, Z.-L. Li, B. Liang, J.-Q. Yang, H.-Y. Cheng, C. Huang, J. Nan, A.-J. Wang, Electrostimulated bio-dechlorination of trichloroethene by potential regulation: kinetics, microbial community structure and function, Chem. Eng. J. 357 (2019) 633–640.
[29] M. Arif, K. Wang, G. Zhu, X. Li, Y. Lv, D.-M. Piao, Q. Feng, Z. Wang, W. Qin, F. Ma, Promoting direct interspecies electron transfer for methane production in bioelectrochemical anaerobic digestion: impact of electrode surface area and switching circuit, Int. J. Hydrogen Energy 47 (2022) 21984–21996.
[30] P. Bao, Z.-Y. Hu, X.-J. Wang, J. Chen, Y.-X. Ba, J. Hua, C.-Y. Zhu, M. Zhong, C.-Y. Wu, Dechlorination of p,p′-DDTs coupled with sulfate reduction by novel sulfate-reducing bacterium Clostridium sp. BXM, Environ. Pollut. 162 (2012) 303–310.
[31] R. Verdini, F. Aulenta, F. De Tora, A. Lai, M. Majone, Relative contribution of set cathode potential and external mass transport on TCE dechlorination in a continuous-flow bioelectrochemical reactor, Chemosphere 136 (2015) 72–78.
[32] X.-Q. Lin, Z.-L. Li, B. Liang, H.-L. Zhai, W.-W. Cai, J. Nan, A.-J. Wang, Accelerated microbial reductive dechlorination of 2,4,6-trichlorophenol by weak electrical stimulation, Water Res. 162 (2019) 236–245.
[33] S.M. Strycharz, S.M. Gannon, A.R. Boles, A.E. Franks, K.P. Nevin, D.R. Lovley, Reductive dechlorination of 2-chlorophenol by *Anaeromyxobacter dehalogenans* with an electrode serving as the electron donor, Environ Microbiol Rep 2 (2010) 289–294.
[34] N. Guo, X. Ma, S. Ren, S. Wang, Y. Wang, Mechanisms of metabolic performance enhancement during electrically assisted anaerobic treatment of chloramphenicol wastewater, Water Res. 156 (2019) 199–207.
[35] J. Wang, Y. Wu, C. Zhang, A. Geng, Z. Sun, J. Yang, J. Xi, L. Wang, B. Yang, Effect of weak electrical stimulation on m-dichlorobenzene biodegradation in biotrickling filters: insights from performance and microbial community analysis, Bioresour. Technol. 390 (2023) 129881.
[36] F. Aulenta, L. Tocca, R. Verdini, P. Reale, M. Majone, Dechlorination of trichloroethene in a continuous-flow bioelectrochemical reactor: effect of cathode potential on rate, Selectivity, and electron transfer mechanisms, Environ. Sci. Technol. 45 (2011) 8444–8451.
[37] Q. Fu, Y. Kuramochi, N. Fukushima, H. Maeda, K. Sato, H. Kobayashi, Bioelectrochemical analyses of the development of a Thermophilic biocathode Catalyzing Electromethanogenesis, Environ. Sci. Technol. 49 (2015) 1225–1232.
[38] N. Guo, Y. Wang, L. Yan, X. Wang, M. Wang, H. Xu, S. Wang, Effect of bioelectrochemical system on the fate and proliferation of chloramphenicol resistance genes during the treatment of chloramphenicol wastewater, Water Res. 117 (2017) 95–101.
[39] J. Li, L. Zhang, C. Li, H. Tian, J. Ning, J. Zhang, Y.W. Tong, X. Wang, Data-driven based in-depth interpretation and inverse design of anaerobic digestion for CH_4 -rich biogas production, ACS EST Eng. 2 (2022) 642–652.
[40] S.-H. Chen, Z.-T. Li, H.-P. Zhao, Bioelectrochemical system accelerates reductive dechlorination through extracellular electron transfer networks, Environ. Res. 235 (2023) 116645.
[41] W.L. Chow, D. Cheng, S. Wang, J. He, Identification and transcriptional analysis of *trans* -DCE-producing reductive dehalogenases in *Dehalococcoides* species, ISME J. 4 (2010) 1020–1030.
[42] P. Leitão, H. Nouws, A.S. Danko, F. Aulenta, Bioelectrochemical dechlorination of 1,2-DCA with an AQDS-Functionalized cathode serving as electron donor, Fuel Cells 17 (2017) 612–617.