Polymers for Advanced Technologies

WILEY

polymers for advanced technologies

RESEARCH ARTICLE OPEN ACCESS

Adsorption Capacity Prediction and Optimization of Electrospun Nanofiber Membranes for Estrogenic Hormone Removal Using Machine Learning Algorithms

Muhammad Yasir[1] | Hamza Ul Haq[2] | Muhammad Nouman Aslam Khan[2] | Jawad Gul[2] | Mukarram Zubair[3] | Vladimír Sedlařík[1]

[1]Centre of Polymer Systems, University Institute, Tomas Bata University in Zlín, Czech Republic | [2]Laboratory of Alternative Fuel and Sustainability, School of Chemical and Materials Engineering, National University of Sciences and Technology, Islamabad, Pakistan | [3]Department of Environmental Engineering, College of Engineering, Imam Abdulrahman Bin Faisal University, Dammam, Saudi Arabia

Correspondence: Muhammad Yasir (yasir@utb.cz) | Muhammad Nouman Aslam Khan (mnouman@scme.nust.edu.pk)

Received: 4 September 2024 | **Revised:** 16 October 2024 | **Accepted:** 27 October 2024

Funding: This work was supported by Ministry of Education, Youth, and Sports of the Czech Republic—DKRVO (RP/CPS/2024-28/002).

Keywords: adsorption prediction | estrogenic hormones | machine learning | optimization | validation | wastewater treatment

ABSTRACT

This study focuses on developing four machine learning (ML) models (Gaussian process regression (GPR), support vector machine (SVM), decision tree (DT), and ensemble learning tree (ELT)) optimized and hyperparameters tuned via genetic algorithm (GA) and particle swarm optimization (PSO) to analyze and predict the adsorption capacity of four estrogenic hormones. These hormones are a serious cause of fish femininity and various forms of cancer in humans. Their adsorption via electrospun nanofibers offers a sustainable and relatively environmentally friendly solution compared to nanoparticle adsorbents, which require secondary treatment. The intricate task is to find the relationship between input parameters to obtain optimum conditions, which requires an efficient ML model. The GPR integrated GA hybrid model performed the most accurate and precise results with $R^2 = 0.999$ and RMSE = $2.4052e^{-06}$, followed by ELT (0.9976 and $4.3458e^{-17}$), DT (0.9586 and $2.4673e^{-16}$), and SVM (0.7110 and 0.0639). The 2D and 3D partial dependence plots showed temperature, dosage, initial concentration, contact time, and pH as vital adsorption parameters. Additionally, Shapley's analysis further revealed time and dosage as the most sensitive parameters. Finally, a user-friendly graphical user interface (GUI) was developed as a predictor utilizing the optimum hybrid model (GPR-GA), and the results were experimentally validated with a maximum error of < 3.3% for all tests. Thus, the GUI can legitimately work for any desired material with given input conditions to efficiently monitor the removal concentration of all four estrogenic hormones simultaneously at wastewater treatment plants.

1 | Introduction

Endocrine disruptive chemicals (EDCs), also known as synthetic estrogenic hormones, do not only have depleting effects on the health of animals but also on human beings [1–3]. Low concentrations, nanograms to micrograms, of these residual pollutants are also detected in clean water storages at the sewerage treatment plants [2, 4]. Interference of the EDCs with the functional groups of the natural hormones because of their mimicking ability has raised serious concerns in the scientific community in the world [5–7]. These hormones can spread through food chains and drinking water sources and cause serious health conditions in humans and aquatic life [7, 8]. Among EDCs, the class of estrogenic hormones includes estrone (E1), estradiol (E2), ethinylestradiol (EE2), and estriol (E3) [9]. These estrogens affect the natural hormonal functions of the human

body and the reproduction of aquatic species [8, 10]. Many studies report fertility issues in alligators and quails due to weight loss that affects the testicles and rise in the femininity of the fish life [11]. Furthermore, a decline in the male sperm count has also been reported, along with an increased risk of getting breast and ovarian cancers in females [12]. It is now widely accepted in the scientific community that the possibility of getting serious health issues such as neurological disorders, leukemia, reproductive problems, and brain cancer is high due to the consumption of EDCs [13]. European Union Directive 2020/2184 declared a threshold level for EDCs' concentration in drinking water, that is, 1 ng/L [14], whereas a few nanograms per liter for wastewater. Estrogenic hormones are present in trace quantities of 0.1–20 ng/L that can cause a threat [15]. A defined experimentation effort is needed to evaluate each hormone's removal accurately. Ranging from the experimentation design to the various threats to the environment, there are major problems that the hormone removal procedure exhibits. Trace quantities of the estrogenic hormones bear carcinogenic impacts on human and aquatic life [16]. They deteriorate soil and life on earth and in water [17].

Sedimentation in rivers can only remove 1% of EDCs, raising severe concerns about treating estrogenic hormones [18]. Conventional methods at wastewater treatment plants are considered less efficient, with greater energy cost [19]; more effective ways to treat EDCs include advanced oxidation processes, nanofiltration, UV photolysis, reverse osmosis, photocatalysis, and adsorption processes [20, 21]. Out of these processes, adsorption has proved to be more promising because of its environmentally friendly properties, cost-effectiveness, high adsorption capacity, and ease of modification [21]. Polymer-based electrospun nanofibers are thought to prove better adsorbents by avoiding secondary treatment than some high adsorbent powder materials and nanoparticles, like activated carbon, biochar, carbonaceous materials, or resins, for their higher life cycles due to higher porosity [22]. In comparison, polymer-based electrospun nanofibers can be used because of their high removal efficiency, greater surface area, controlled geometry, low production cost, and better reusability [23]. Additionally, functionalized polymeric membranes with additives have distinct abilities like selectivity and high removal efficiency [24]. The efficiency of these membranes can be enhanced by adding some additives like graphene oxide or surface functionalization with amine groups, and so forth [25–28]. So far in the literature, synthetic polymeric membranes based on cellulose acetate (CA), regenerated cellulose, polyvinylchloride, polysulfone (PSU), polypropylene, polycarbonate, and polydimethylsiloxane, and so forth, have been used for removing estrogens from the environment. On the other hand, sorbents such as rice husk silica have been used for the adsorption of E1 and E2 with removal efficiencies of 93.1% and 95.5%, as reported by M.H. Zarghi et al. [29]. In another study, A.E. Burgos Castellanos et al. used vermiculite intercalated with cetyltrimethylammonium ions to eliminate EE2 hormone [30] while Karina Bugan Debs et al. used yeast biomass from ethanol industry for its removal [31]. Moreover, Longjie Liu et al. synthesized green rGO/FeNPs nanocomposites activated peroxydisulfate for E2 and E3 adsorption [32]. Additionally, Mohammed B. Abdul-Kareem et al. prepared cement kiln dust-based beads as an effective sorbent for the removal of E1 [19]. Still, these materials require extensive experimental setup and various tangible resources that are not only time-consuming but also require a lot of financial support. An experimental procedure alone is a time-consuming process to evaluate the removal efficiency and adsorption capacity of various adsorbents. Apart from having the correct dosage and timely execution, a marginal error can result in a significant deviation in results [33]. This problem can be addressed by employing machine learning (ML) methods to predict the adsorption concentration of E1, E2, EE2, and E3 for polymer-based electrospun nanofibers [34].

In comparison to experimental studies, ML methods can predict outcomes by utilizing operational conditions and input parameters [35]. Various parameters such as material dosage, initial concentration of estrogenic hormone, experimental contact time, and solution pH are considered for their impact on adsorption efficiency and their linear/nonlinear relation. ML models craft the relationship, which is prevalent, adoptive, and sufficient to explain the behavior of various inputs related to the adsorption output. The related model can potentially simulate and predict the removal values of specific pollutants. Therefore, it is useful for predicting the adsorption concentration of polymer-based electrospun nanofibers and studying their life cycles and efficiency by developing artificial intelligence (AI) based ML models. ML integration allows the efficient and targeted removal of EDCs, such as estrogenic hormones, from wastewater. Conventional experimentation of estrogenic hormone removal via adsorption can be reduced with an active, tangible, and sustainable ML model development [36]. Moreover, the ecological deterioration and environmental hazards of experimentation and harmful waste can be minimized, and effective ML models serve as the foundation for achieving a green and sustainable future [37].

In this regard, ML models can effectively diffuse these challenges by employing an intelligent approach to pick the nonlinear trends for the prediction of the removal characteristics of various hormones via adsorption. ML methods are, at some level, viable alternatives for experimentation to achieve scientific creditability among the various parameters and their impact on output [38]. The existing experimental data serves as a foundation for data profiling in training the ML models to predict adsorption capacity with reduced experimentation [39]. Models such as regression tree, along with random forest (RF) classification, are engaged for data mining to indicate the nonlinear relationship of the variables [40, 41]. These models can also be utilized for process optimization, adsorbed concentration observation, and automation under variable conditions [42]. Such automation promises economic viability, industrial advancement, and environmental friendliness, making it a key innovation in the field. This study utilizes a dataset on the adsorption removal of estrogenic hormones using electrospun polymeric nanofibers. To the best of the authors' knowledge, no study has been reported on the use of ML in this area of estrogenic hormone removal. Hence, ML methods provide economically viable solutions to predict the removal efficiency of estrogenic hormones and offer a sustainable solution for effectively treating these complex pollutants on a larger scale.

In this study, estrogenic hormones' adsorption capacity is predicted using a hybrid model of experimentation along with ML techniques, that is, ensemble learning tree (ELT), support vector machine (SVM), gaussian process regression (GPR),

and decision tree (DT) infused with genetic algorithm (GA) and particle swarm optimization (PSO) methods. This unique methodology effectively optimizes the ML algorithms along with tuned hyperparameters to forecast the kinetic adsorption capacity of various polymeric electrospun nanofibers primarily. A comparative analysis was used to improve the performance of the prediction models, which had a high coefficient of regression (R^2) value and a low root mean square error (RMSE) value after optimization. Using the optimal model, the partial dependence plots (PDPs) were elaborated to test the impact of each input parameter, such as dosage, solution pH, time, temperature, equilibrium, and initial concentration, on both output adsorption efficiency (%) and kinetic adsorption capacity (Q_e). The Shapley and sensitivity analysis was performed to ensure the ML model's stability in terms of numerical changes in input parameters individually. Furthermore, a graphical user interface (GUI) is developed based on AI-based ML modeling to accurately predict the adsorption capacity of hormones for designed nanofiber membranes. Also, the experimental validation of the interface revealed a maximum error of 3.3%, making it suitable for monitoring the removal of all four estrogenic hormones simultaneously in wastewater treatment plants under the appropriate input conditions. This allows the ML models to help forecast and clarify estrogenic hormone adsorption mechanisms. The flowchart of the research is demonstrated in Figure 1.

2 | Methodology

2.1 | Data Collection and Preprocessing

The first step was acquiring data to build up the dataset to represent material used against pollutant removal. The data were acquired from the Web of Science and Science Direct for membranes and nanofibers fabricated by electrospinning technique for the removal of estrogenic hormones (Table S1). Nanofibers from various polymers, including PSU, waste cigarette electrospun nanofibers (WCENFs) represented as CIG, polyamide (PA), CA, polyacrylonitrile (PAN), polyether sulfone (PESu), polyurethane Elastollan (PU-E), polyurethane 918 (PU 918), PU 918 modified with polyaniline (PU-P) are reported. The key parameters include material type, BET surface area, equilibrium concentration, and kinetic adsorption capacity (Q_e). Furthermore, the conditions on which the removal was performed were also reported in the data set. The actual experimental data were preprocessed prior to ML modeling. Obtained data were subjected to data smoothing techniques, mainly involving filling in missing values and eliminating

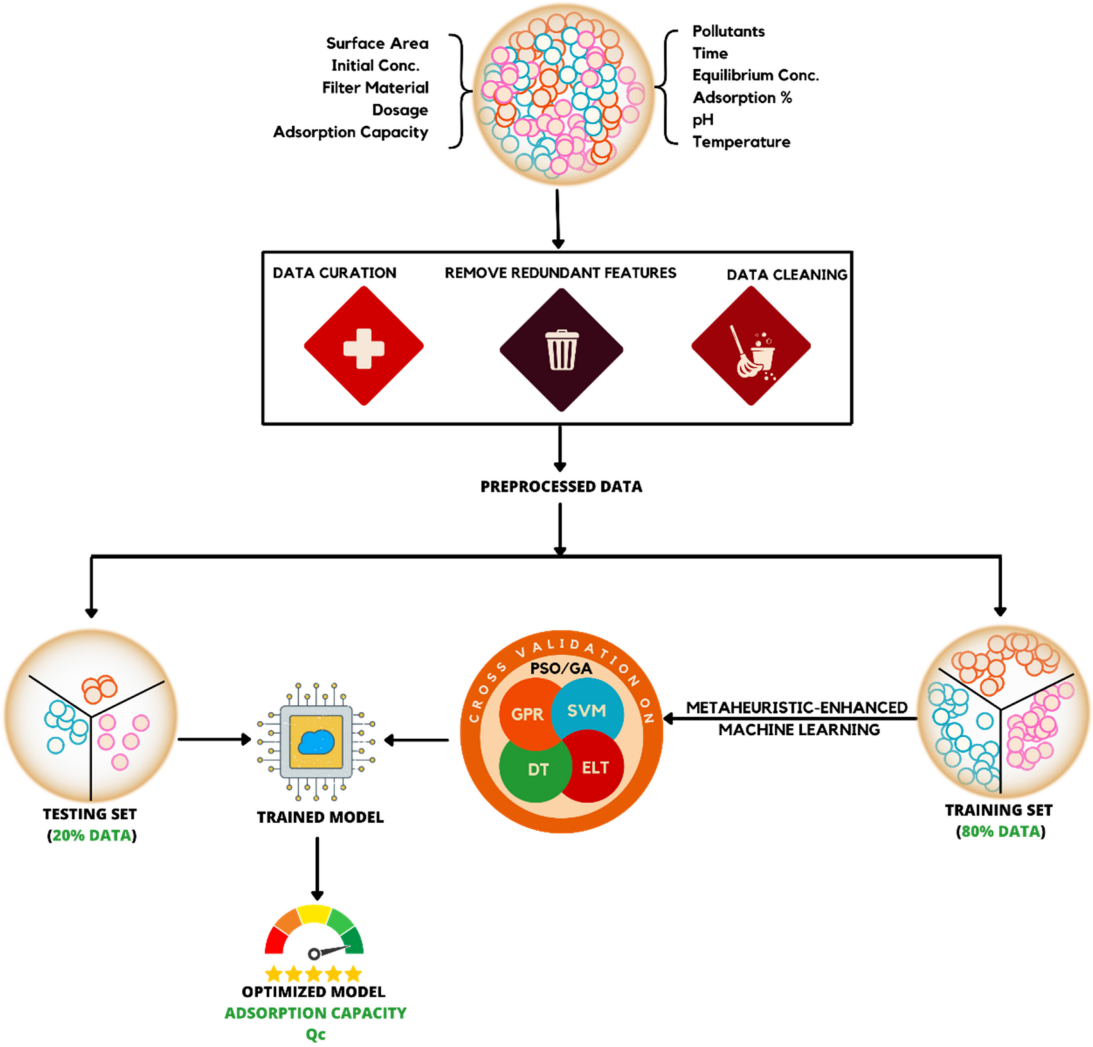

FIGURE 1 | Schematic workflow of machine learning models.

outliers in the data set. Then, the experimental data were formulated and utilized to train and test ML models.

2.2 | Data Representation

There are 10 parameters contributing to accessing the adsorption capacity for each pollutant, with a total of 449 points (Table S1). To represent data distribution, count plots and violin plots are generated (Figure 2). From the count plot, it is evident that the data points are equally distributed between the pollutants, that is, E1, E2, EE2, and E3. Data points for E2 and E3 are around 100 counts; for E1, the total data count is 120 counts, and the maximum count for the data points is for EE2, which is 140 counts (Figure 2a). Similarly, data of the absorbent fiber materials are also evenly distributed as depicted in Figure 2b.

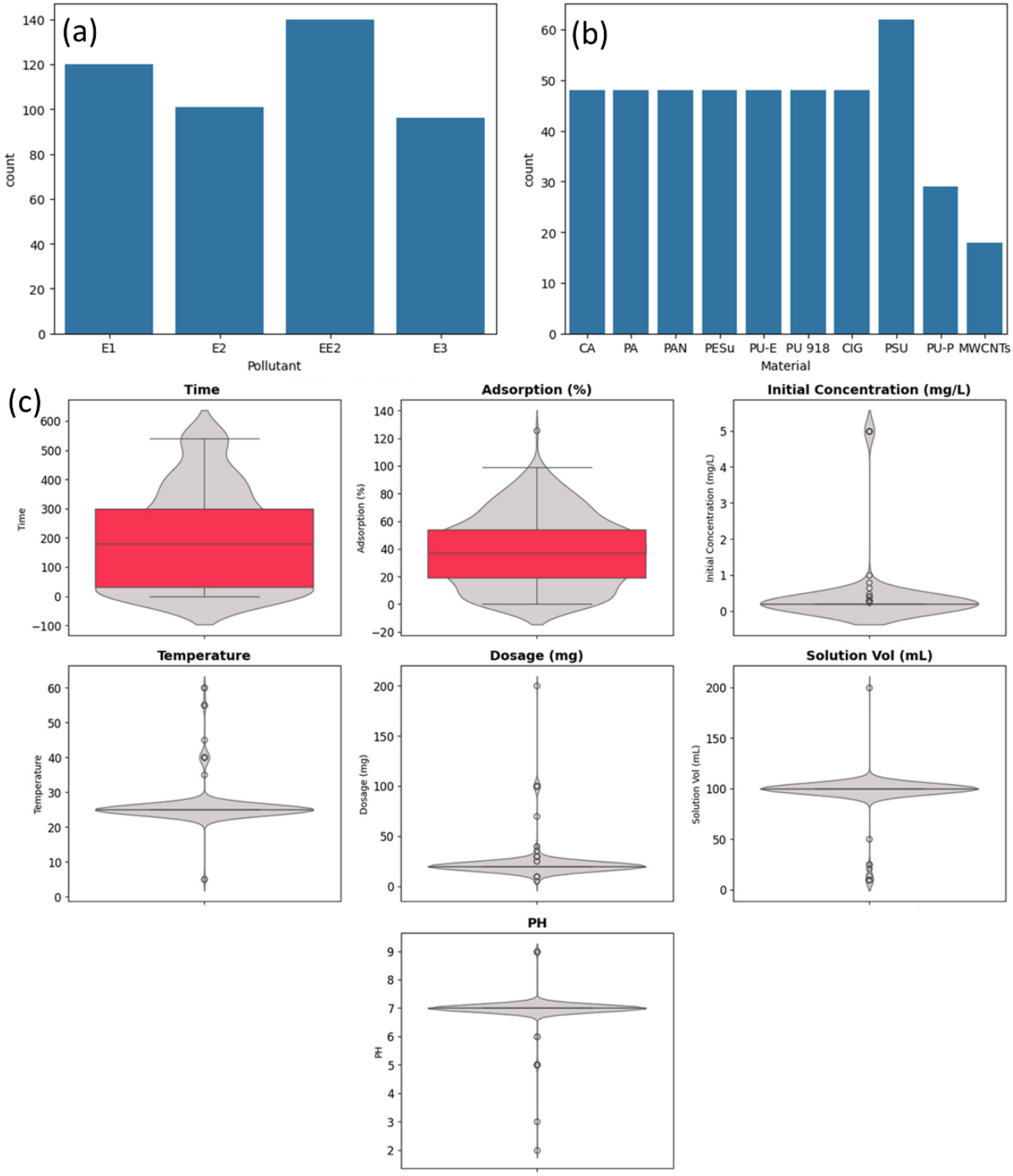

FIGURE 2 | Count plots for (a) four estrogenic hormones, (b) ten different adsorbent materials, and (c) violin plots for experimental features.

Except for PSU, PU-P, and multi-walled carbon nanotubes (MWCNTs), all other adsorbents' data points are the same and lie just below 50 points each. PSU has the highest number of data points, that is, above 60 points. The lowest data count is for MWCNTS, which is well below 20, followed by PU-P, which has counts reaching 30.

Violin plots are a very convenient way to represent available continuous data sets. The interquartile range (IQR), along with the median, maximum, and minimum range of the data, can be graphically represented [43]. The data ranges of adsorption percentage, temperature, dosage initial concentration, time, solution volume, and pH are shown in Figure 2c. The adsorption percentage is in the range of 0%–100%; the temperature is in the range of 5°C–60°C with a concentration of the data set at 25°C. The dosage of the pollutant is concentrated at 20 mg with a range of 10–200 mg, and similarly, the data set of the initial concentration of the pollutant is concentrated at 0.2 mg/L with a range of about 0.2–0.5 mg/L. Time is distributed within a range of 0–600 min, BET surface area of the adsorbent fiber is distributed in the range of 5.16–84.3 m^2/g, data of the solution volume are concentrated at 100 mL with minimum value at 10 mL and pH of the wastewater is fixed at 7 with few exceptions between 2 and 9.

2.3 | Machine Learning Models and Optimization Techniques

Due to the limitation of the dataset on the adsorption capacity, it is necessary to find the best-performing ML model. A comparative analysis of ML models ELT, GPR, SVM, and DT in integration with optimization techniques is needed to increase the predictive capability of ML models. This section briefly introduced the ML models and optimization techniques used for adsorption capacity prediction.

SVM deals with both linear and nonlinear relations of input parameters with adsorption capacity. It separates the dataset with hyperplanes drawn to separate the data into various classes [44]. After classification, a number of hyperplanes and data points can be successfully used to distinguish the various data points present in the analysis [45]. The nonlinear processes can be executed by employing kernels to SVM, which elevates the ability of SVM to adapt to the nonlinear process functions. The empirical comparisons support GPR models in deciphering interpretable and expressive optimal prediction performance while avoiding overfitting. The classification and nonlinear regressions find GPR as the most desired solution [46]. Working like a tree model, the DT helps to logically predict estrogen removal with inputs including temperature and initial dosage. Leaf nodes, intermediate nodes, and root nodes are the critical characteristics of the DT [47]. Prediction performance can be significantly enhanced by combining numerous DTs for better performance. Boosting and bagging are mainly employed techniques by the ELT, and dispersion of DT can be reduced by bootstrap aggregate [48]. Multiple data subsets are formulated by random selection to form the trained data. Every subsequent subset is used for the training of DT, which would ultimately result in an ensemble tree [49].

The optimization techniques of GA and PSO show remarkable integration with ML models to increase their efficiency. The natural selection-based ideas are used using the GA approach. Genetic structure and the chromosomes are the basis of the behavior on the basis on which the algorithm proceeds. A solution is formed by the cumulative contribution of each chromosome. A GA employs the past data for probable solution direction, providing the best fit [50]. Employing the position of the particle respective to the swarm, PSO utilizes the memories relevant to the position to present a globally optimal solution. With the ability to adapt to changing circumstances and low confinement of continuity of an objective function, PSO occupies a significant position in progressive computation algorithms.

RMSE and R^2 of the ML models were reported to evaluate the subject viability to predict and analyze. The data set obtained from the experimental data were divided into two data sets. One was 20%, used for testing, and 80% used for training for the ML models. A five-fold cross-validation was employed to avoid any overfitting exhibited on the data set.

3 | Results and Discussion

3.1 | Pearson Correlation Analysis

For establishing a suitable model, the Pearson correlation coefficients demonstrated in Figure 3 present both linear and nonlinear relations between features. The equilibrium concentration and time have a straightforward relationship due to the nature and dependency of features on the variable. The correct time corresponds to the actual condition of removing pollutants in relation to the heatmap. Similarly, the time of contact between adsorbent and adsorbate has a direct impact on the removal of the estrogenic hormone via adsorption (0.44) up to a certain specific time, after which equilibrium is achieved either due to a limited amount of adsorbate in solution or due to saturation of adsorbent material. Furthermore, the initial concentration and dosage have a high positive correlation (0.75), showing that a higher dosage of nanofiber is required for the same amount of removal efficiency if the initial concentration of the estrogenic hormone is increased. The dosage and solution volume have an indirect relation (−0.58), indicating that a higher dosage of material and a lesser solution volume will result in a higher removal of estrogenic hormones. Lastly, a −0.9 correlation value between solution volume and initial concentration of estrogenic hormone shows that increasing solution volume with increasing concentration will result in a drop in removal efficiency. This could cause early saturation of adsorbent due to a lack of available sites for adsorption.

3.2 | Performance Evaluation and Tuning

Table 1 explains after fine-tuning the ML Models under GA and PSO, the most influencing features are selected against each optimization method. Under DT, the adsorption % of the pollutant, temperature in°C of the water, solution volume in mL, and BET surface area in m^2/g are the contributing input parameters for GA. Similarly, for PSO, except adsorption % of the pollutant, pH of the wastewater, dosage in mg, solution volume in mL, and adsorbent fiber type, all other input parameters are the major contributors to the prediction of kinetic adsorption capacity (Q_e). For the discussion of the ELT,

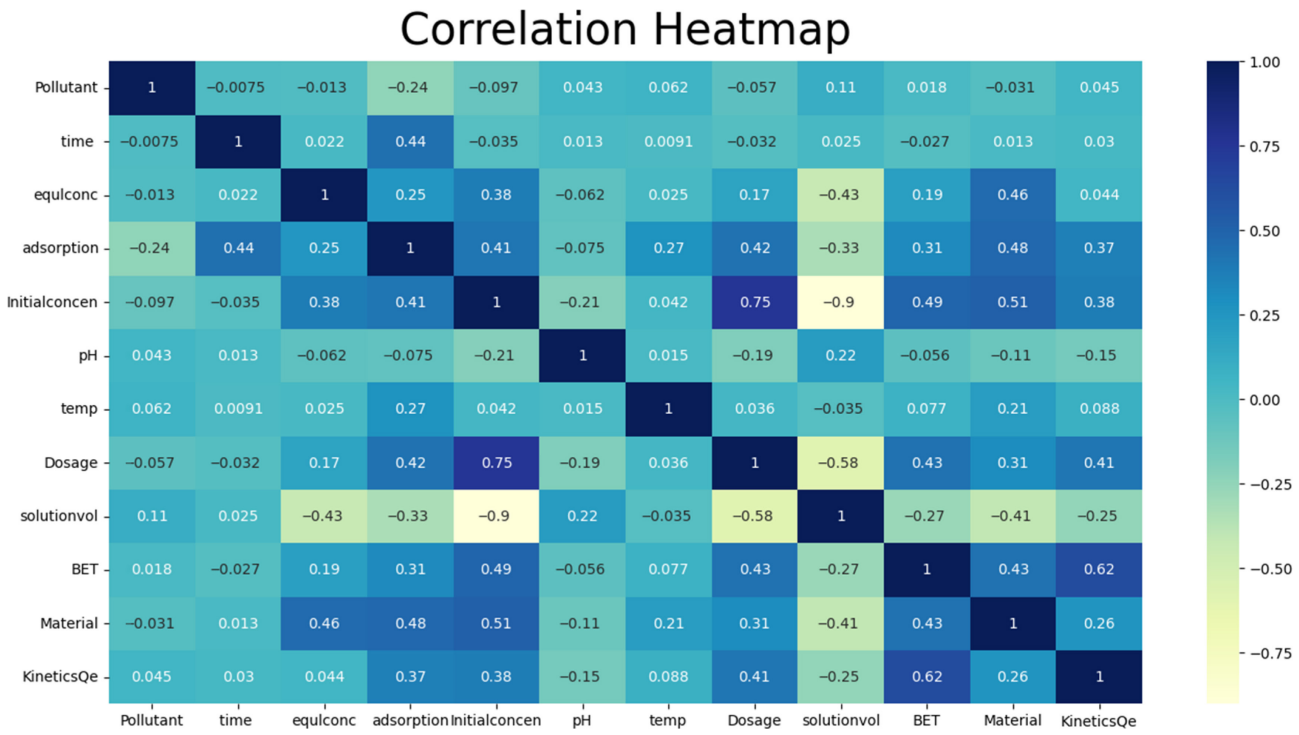

FIGURE 3 | Pearson correlation between the input parameters (experimental variables and operating conditions) and output (adsorption efficiency and capacity (kinetic Q_e)).

TABLE 1 | Fine-tuning of ML models under optimization techniques of GA and PSO.

Prediction model	GA	PSO
DT	Adsorption %, temperature in°C, solution volume in mL, BET surface area in m^2/g	Except adsorption %, pH, dosage in mg, solution volume in mL, adsorbent fiber
ELT	Except for BET surface area in m^2/g	Except temperature in°C
GPR	Except temperature in°C	Except for time in min, solution volume in mL, BET surface area in m^2/g, adsorbent fiber
SVM	Except for time in min, adsorption %, initial concentration in mg/L, dosage in mg, and BET surface area in m^2/g	Except initial concentration in mg/L

under GA (except for BET surface area in m^2/g) and under PSO (except for temperature in°C), all other input parameters are major contributors to predicting the adsorption capacity (Q_e). In GPR under GA except temperature in°C and under PSO except for time in min, solution volume in mL, BET surface area in m^2/g, and adsorbent fiber type, all input parameters influence the prediction of adsorption capacity (Q_e). Similarly, except time in min, adsorption % of the pollutant, initial concentration mg/L, dosage in mg, and BET surface area in m^2/g for SVM under GA, and except initial concentration in mg/L in the SVM under PSO, all other input parameters are contributing to the accurate prediction of the adsorption capacity (Q_e) of the adsorbent fiber.

A five-fold cross-validation allowed the selection of hyperparameters for the ML models. The GA and PSO were used to optimize the DT, GPR, SVM, and ELT hyperparameters. The optimized values for SVM were reported with a box constraint value of 964, kernel scale value of 726, epsilon 0.0269, and kernel linear function for PSO optimization. The selected hyperparameters, along with their optimized values and ranges, are illustrated in Table 2. After applying GA on SVM hyperparameters, it gives a box constraint of 1.808, epsilon of 0.319, and kernel scale of 0.744 with a Gaussian kernel function. At the same time, PSO reports a box constraint of 963.937, epsilon of 0.027, kernel scale of 726.898 with polynomial kernel function. After hyperparameter tuning under GA, DT turns the surrogate off with a minimum leaf size of 1, but in the case of PSO, it keeps the minimum leaf size the same as GA but turns the surrogate on. Hyperparameter tuning for ELT results to set the number of learning cycles of 10 for GA and 84 for PSO, learning rate to 1 for GA and 0.701 for PSO and utilizing the LSBoost method for GA. Similarly, for GPR, the exponential kernel function is applied for GA, and the squared exponential kernel function is applied for PSO. The kernel parameters are set as [125.498, 3.127] for GA and set as [0.629, 0.378] for PSO,

TABLE 2 | Selection of hyperparameters tuned with GA and PSO for ML models.

	Ranges hyperparameter	GA	PSO
DT	Surrogate [on, off]	Off	On
	Minimum leaf size [1–196]	1	1
ELT	Number of learning cycles [10–500]	10	84
	Learning rate [0.001–1]	1	0.7014534898
	Methods [LSBoost, Bag]	LSBoost	Nil
GPR	Kernel function [squared exponential, Ard exponential]	Exponential	Squared exponential
	Basic function [constant, zero, linear]	None	None
	Kernel parameters	[125.4975579404, 3.1268225463]	[0.6286711784, 0.3775129030]
	Sigma [0.001–1000]	21.4613609864	4.1205
SVM	Box constraint [0.001–1000]	1.8080103527	964.9371639470
	Epsilon [0.00032404–32.4043]	0.3199491344	0.0269844871
	Kernel scale [0.001–1000]	0.7436245213	726.8985065764
	Kernel function [polynomial, gaussian, linear]	Gaussian	Polynomial

and the sigma is recorded as 21.461 for GA and 4.121 for PSO after hyperparameter tuning.

3.3 | Performance of ML Models Under GA

After hyperparameter tuning and feature selection, all ML models accurately predicted the adsorption capacity (Q_e) except for SVM, as shown in Table 3. The R^2 and RMSE values for DT are 0.9586 and 2.4673e^{-16} for testing and 0.94 and 1.34e^{-16} for training data, respectively. ELT reports R^2 and RMSE as 0.9976 and 4.3458e^{-17} for testing and 0.9900 and 8.0748e^{-16} for training, respectively. For GPR, recorded values of R^2 and RMSE are 0.9999 and 2.4052e^{-6} for testing and 0.9999 and 2.7022e^{-6} for training, respectively. Similarly, R^2 and RMSE values reported for SVM are 0.7110 and 0.0639 for testing and 0.2484 and 13.52 for training, respectively. It is evident from the data of R^2 and RMSE that DT, ELT, and GPR performance are satisfactory for predicting the adsorption capacity of the polymeric nanofiber materials for the pollutant compared to that of SVM under GA. The nonlinear relation of adsorption capacity with the selected features is the reason behind the unsatisfactory performance of SVM. This dissatisfying behavior is evident in both testing data and getting worse in the training data set. Contrary to this, GPR performs better than all the other models in the discussion, reporting R^2 at almost 1 and RMSE at approximately 0 because of the linear relation of adsorption capacity with the corresponding input parameters. The ELT is just behind GPR in its performance in predicting adsorption capacity, followed by the DT. The relation of actual adsorption capacity and predicted adsorption capacity for each model under GA are well defined in Figure 4. The relationship between adsorption output and input parameters can be characterized by Shapely's method and GPR model. Additionally, feature attribution can elucidate the relative importance and magnitude of each feature pertaining to the output of adsorption. This was used to assess the relative importance of different input parameters on the adsorption of estrogens. The GPR performed better in testing and training on the provided data set, showing 0.999 R^2 for both training and testing data. The reported values of training and testing for GPR integrated with GA indicate high accuracy and precision that would be further used for prediction validation.

3.4 | Performance of ML Models Under PSO

The performance of ML models under PSO after hyperparameter tuning is shown in (Figure 5). The GPR performs exceptionally well as compared to other ML models, reporting R^2 and RMSE for testing as 0.9714 and 2.6290e^{-5}, and for training, it gives 0.9999 and 2.059e^{-6}, respectively (Table 3). Following GPR, DT reports R^2 and RMSE as 0.8065 and 1.6262e^{-16} for testing and 0.9587 and 4.4860e^{-17} for training, respectively. The regression plots of the actual adsorption capacity against the predicted adsorption capacity confirm the linear relations in ML models. However, the linearity index in GPR is the highest, followed by DT, ELT, and SVM for the training data. GPR performed optimally under PSO, generating R^2 of 0.9999 and RMSE of 2.059e^{-06} for training data, while it showed R^2 of 0.9714 and RMSE of 2.6290e^{-05} for testing, which is considered to be the most effective model. Based on the results, it can be concluded that the GPR model can be considered ideal for optimizing the prediction of adsorption capacity for estrogenic hormones. Moreover, the modeling parameter results of GPR optimized by GA are slightly better (less scattered data points around linear lines) than those of PSO.

3.5 | Feature Importance/Impact on Kinetic Adsorption Capacity (Q_e)

In the context of estrogenic hormone removal, a systematic investigation into the multifaceted relationships between various input parameters and the efficacy of the removal process

TABLE 3 | Parameters of ML models accuracy using GA and PSO optimization techniques.

	Before optimization	
Model	R^2	**RMSE**
DT	0.68	2.501
ELT	0.79	0.2277
GPR	0.37	0.4246
SVM	0.60	0.3372

	After Optimization			
	Training		Testing	
GA models	R^2	**RMSE**	R^2	**RMSE**
DT	0.94	$1.3458e^{-16}$	0.9586	$2.4673e^{-16}$
ELT	0.9900	$8.0748e^{-16}$	0.9976	$4.3458e^{-17}$
GPR	0.9999	$2.7022e^{-06}$	0.9999	$2.4052e^{-06}$
SVM	0.2484	13.52	0.7110	0.0639

	After Optimization			
	Training		Testing	
PSO models	R^2	**RMSE**	R^2	**RMSE**
DT	0.9587	$4.4860e^{-17}$	0.8065	$1.6262e^{-16}$
ELT	0.8702	0.1480	0.9303	0.0677
GPR	0.9999	$2.059e^{-06}$	0.9714	$2.6290e^{-05}$
SVM	0.7688	0.3621	0.9721	0.0269

is achieved through the implementation of two-way PDPs, as depicted in Figure 6. This graphical representation exposes the complex dependencies among input variables and the removal process dynamics. By utilizing two-way PDPs, a holistic examination is conducted to understand the influence of pollutants, adsorption time, initial concentration, solution pH, dosage, material characteristics, kinetics adsorption capacity (Q_e), BET surface area, and equilibrium concentration on removing estrogenic hormones. This methodological approach facilitates a comprehensive understanding of the interactions and synergies among these input adsorption parameters, thereby offering a foundational basis for optimization strategies. Optimal dosage and material properties emerge as critical factors influencing the efficiency of estrogenic hormone removal. Moreover, the plots unveil intricate relationships between temporal parameters, initial concentration, and pH, providing valuable information for optimizing the removal process.

Figure 6a shows the dependence of the adsorption capacity of the adsorbent against adsorption time from 0 to 550 min. The results indicate a sharp decrease in the adsorption capacity from 3 to nearly 1 mg/g from 0 to 200 min (about three and a half hours). However, with a further increase in adsorption time from 200 to 500 min, the adsorption capacity increased and reached a maximum value of 2.1 mg/g. Based on these results, it can be deduced that the adsorption contact time is a significant parameter affecting the removal of estrogenic hormones from water. The lower adsorption time is more feasible for effective removal, and longer contact time may be expected to promote the desorption of hormones from the adsorbent surface and reduce the adsorption capacity [51].

Figure 6b,d depict the relationship between the pollutant's average and initial concentration on the adsorbent's adsorption capacity. As shown, the increase in average and initial concentration of pollutants resulted in a linear increase in adsorption capacity. This is mainly associated with the presence of more estrogenic molecules in water, which promotes molecular diffusion and improves the sorption of pollutants on the adsorbent surface. The prediction of the adsorption percentage of the pollutant is shown in Figure 6c. As per the curve, it can be postulated that the adsorption capacity of the filter materials increases with the increase in the adsorption percentage of the pollutant. Initially, a declining trend is observed in the predicted curve till 1.5 mg/g of adsorption capacity at about 20% adsorption of the pollutant, but after that, a gradual increase in the adsorption capacity of the filter is predicted with the increasing adsorption percentage of the pollutant.

The initial pH of wastewater containing estrogenic hormone contaminants is an essential factor that directly affects the chemistry of both adsorbent and adsorbate, resulting in significant variations in adsorption efficiency. The change of pH from 2 to 9 on the adsorption capacity of absorbent is displayed in Figure 6e. It was observed that the acidic pH ranges from 2 to 5 showed low adsorption of pollutants onto the adsorbent surface. The increase in solution pH > 5 resulted in a significant rise in adsorption capacity and reached a maximum value of 1.496 mg/g at a pH of 9. The change in pH can be well described by considering the ionic state of estrogenic hormones and the surface chemistry of adsorbent. At acidic pH, the hormones exist in nonionic molecular form and may be absorbed by electrostatic interactions, which decrease with increasing pH from 2 to 5 [28]. At neutral to alkaline pH from 6 to 9, the increase in hormone sorption is mainly associated with the chemical interaction of abundant hydroxyl groups of adsorbents with hormone molecules.

In Figure 6f, it can be noted that initially, the adsorption capacity is highest, around 6.8 mg/g, and the trend has a linear fall to 1 mg/g with the rise in the dosage of the adsorbent fibers to 100 mg. This drop could be attributed to a low concentration of hormones in solution with an excess amount of adsorption sites available on the surface of nanofibers. Above this value, the ML model predicts a negligible change in the adsorption capacity with increasing the adsorbent dosage. The result is consistent with previous studies and mainly attributed to the presence of more sorption sites with increased dosage, which can lead to high hormone removal from water bodies [52].

The influence of hormone solution volume at a fixed concentration on the adsorption capacity of nanofiber adsorbents is elucidated in Figure 6g. As can be seen, there is a steep rise in the adsorption capacity with an increase in solution volume. This is attributed to the fact that more molecules are available to be attached to the adsorption sites of the nanofibers, up to 600 mL of solution volume. However, a further increase in solution volume to 700 mL shows a slight increase in adsorption capacity

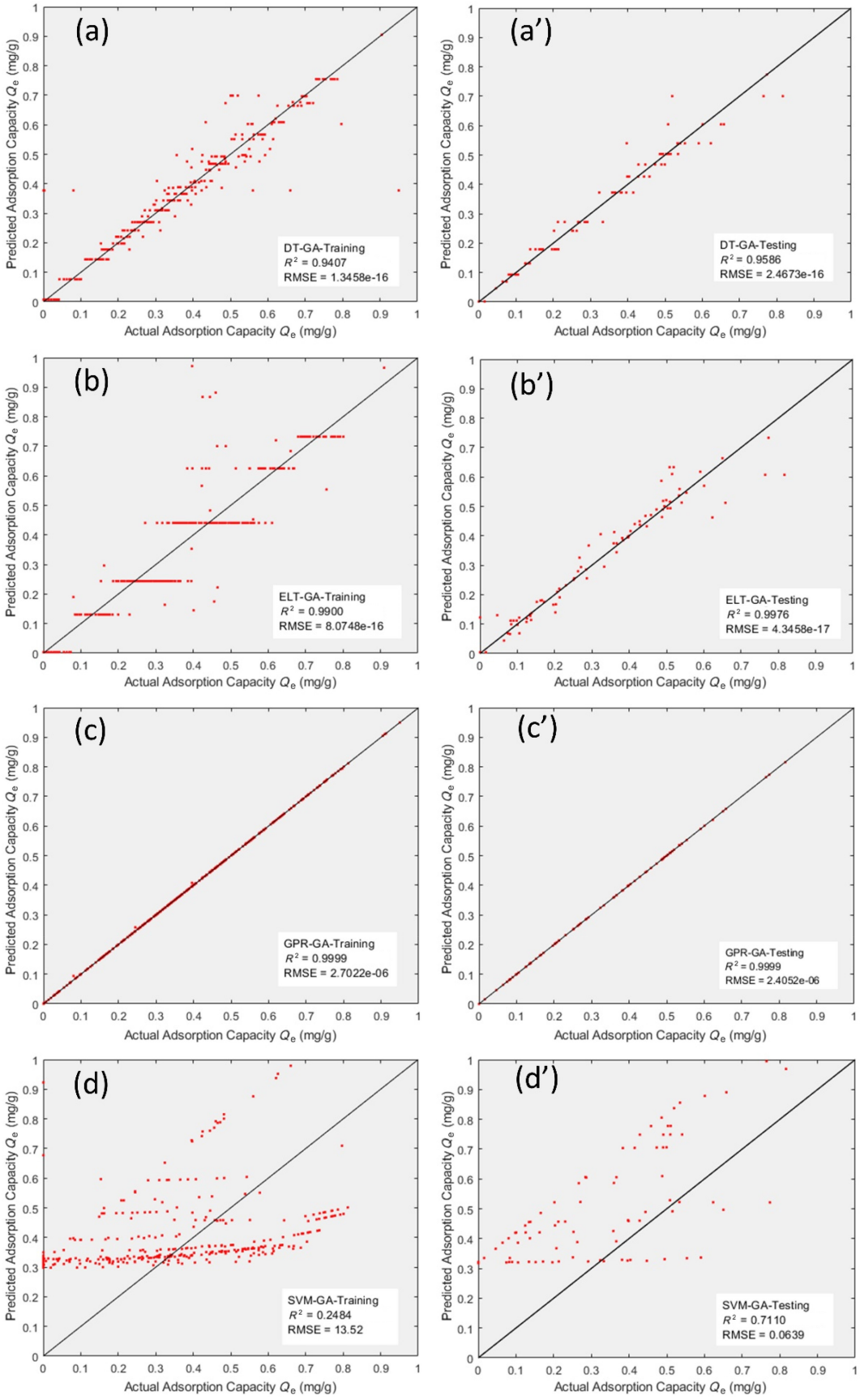

FIGURE 4 | Illustration of predicted versus actual adsorption capacity of estrogenic hormones using GA for training (left panel) and testing (right panel) data on models (a–a′) DT, (b–b′) ELT, (c–c′) GPR, and (d–d′) SVM.

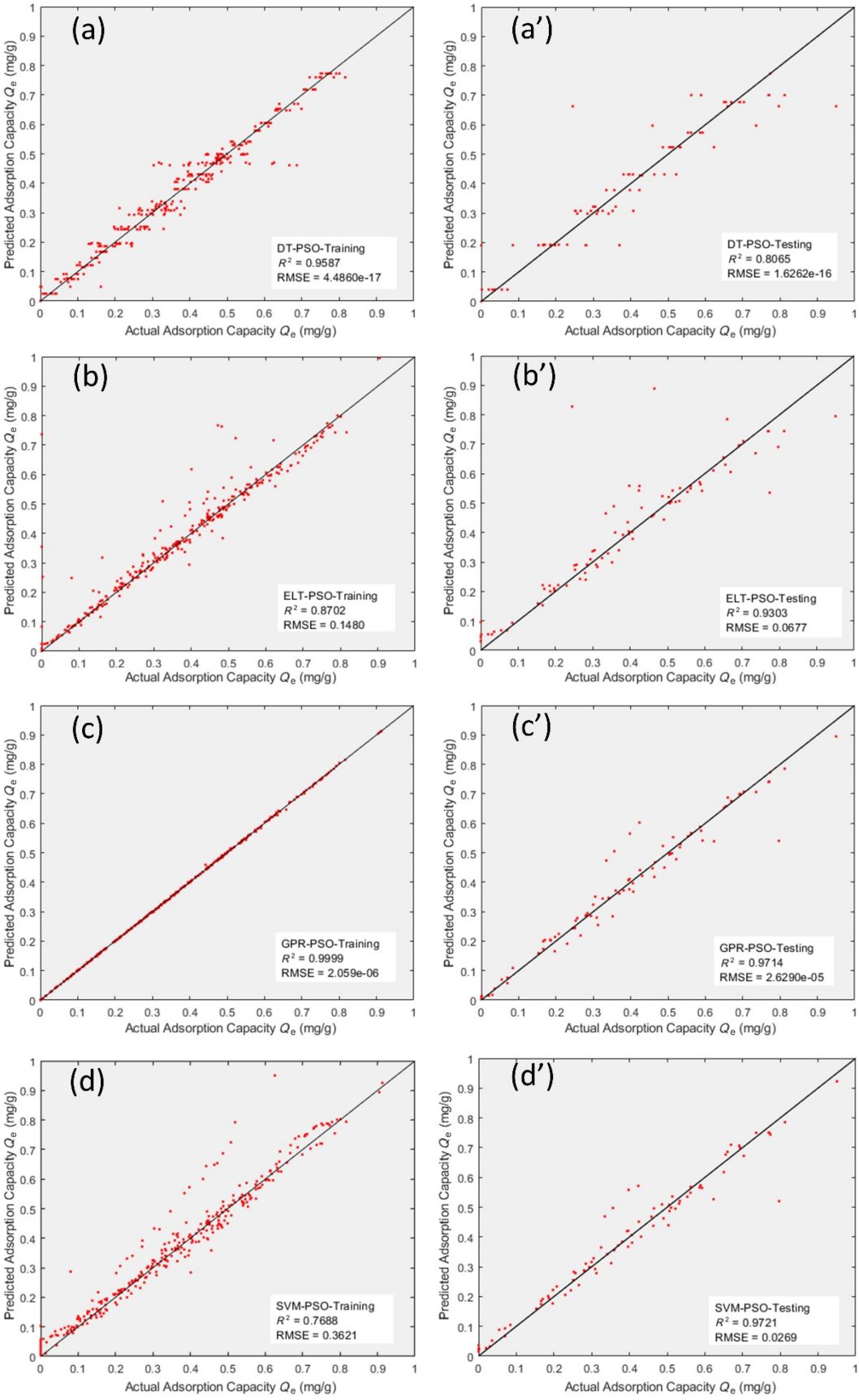

FIGURE 5 | Demonstration of predicted versus actual adsorption capacity of estrogenic hormones using PSO for training (left panel) and testing (right panel) data on models (a–a′) DT, (b–b′) ELT, (c–c′) GPR, and (d–d′) SVM.

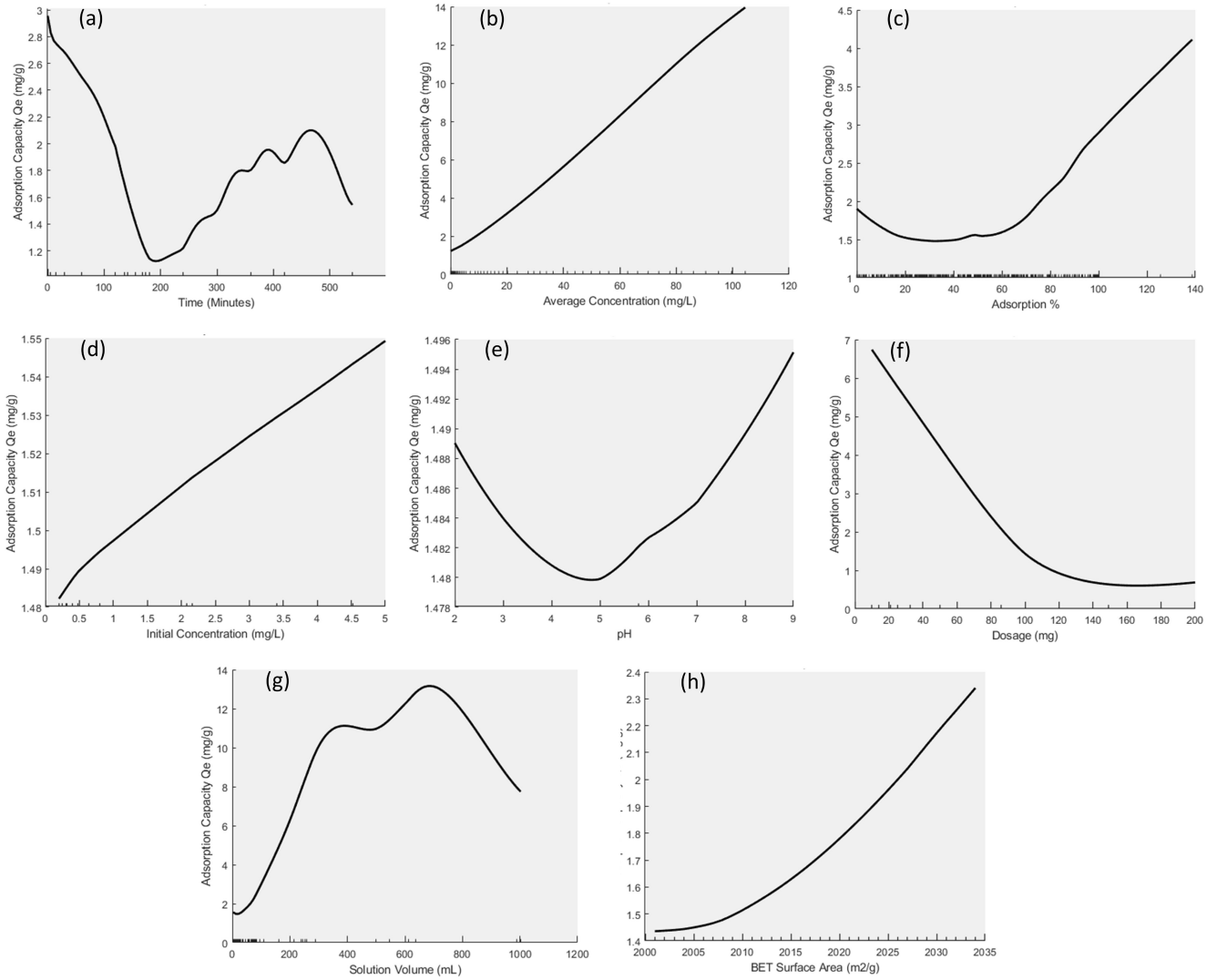

FIGURE 6 | 2D PDPs of input experimental parameters. (a) Contact time, (b) average equilibrium hormone solution concentration, (c) adsorption removal percentage, (d) initial hormone concentration, (e) solution pH, (f) dosage of nanofiber membrane, (g) solution volume, and (h) surface area (BET) of the nanofiber membrane to predict the adsorption capacity.

at a decreasing rate, which indicates that saturation might have been achieved as all adsorption sites on the given adsorbent dosage are completely occupied. After this point, a further increase in solution volume demonstrates a drop in adsorption capacity. A plausible reason could be that the concentration of hormones in the solution volume is in excess, and no more adsorption sites are available on the surface of the adsorbent, leading to a decrease in adsorption capacity.

The surface area of the absorbent is a crucial feature that directly influences the adsorption performance of any adsorbent material. A trend somewhat similar to that of solution concentration is observed with the adsorbent surface area. The adsorbent with a high surface area is expected to exhibit high adsorption capacity. As shown in (Figure 6h), there is an exponential increase in adsorption capacity from 1.45 to 2.4 mg/g with an increase in the surface area of the adsorbent. The increase in surface area indicates the presence of more active sorption sites, which are expected to adsorb more estrogenic hormones on the adsorbent nanofiber surface. A plausible reason could be that hormones are widely present in solution, so increasing the amount of nanofiber can increase the surface area. Consequently, it could also be observed that a much higher adsorbent surface area by an excess increase of nanofiber dosage showed a reduction in adsorption capacity for hormone removal. This might be associated with limiting solution concentration while using an excess nanofiber dosage. Another aspect could be the formation of micropores during adsorbent and pollutant interaction, which are expected to limit the adsorption of pollutants and lower the adsorption capacity, as investigated in another study [53].

The 3D PDPs are displayed in Figures 7 and 8 to develop a relation between the input and output parameters. Various input parameters were analyzed and discussed, including pH, temperature, adsorbent dosage, and initial concentration, which significantly impacted the hormone's output variables (adsorption capacity and % adsorption). This elaborated on the genuine relationship and the impact each input had on the output. As depicted in Figure 7a, the % adsorption of 80%–100% of hormones was achieved at an initial concentration > 4–5 mg/L and an adsorbent dosage of 100–150 mg/L. This indicates that the increased number of sorption sites at high adsorbent dosages can

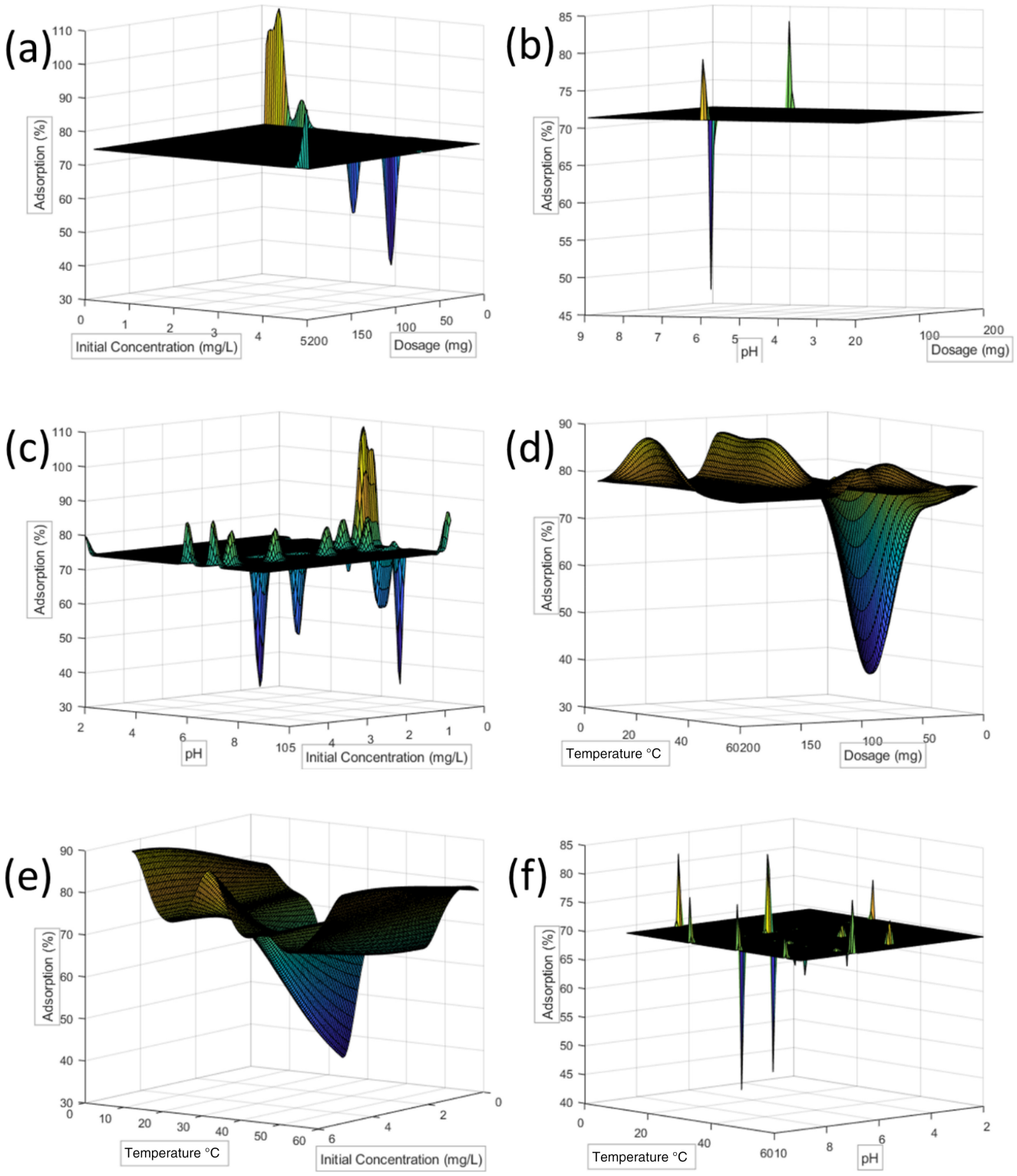

FIGURE 7 | 3D PDPs of adsorption removal (%) with input parameters.

efficiently uptake hormone molecules present in water. A lower adsorbent dosage of < 100 mg/L can remove a maximum of 70% of hormones, mainly associated with fewer active adsorption sites available on the adsorbent surface.

The impact of solution pH and dosage on the % removal of hormones is shown in Figure 7b. It was observed that the maximum % adsorption of hormones from 70% to 85% can be obtained at higher dosage > 100 mg when the solution pH value ranges 6–8. This implies that a high adsorbent dosage is required when the solution pH is in the acidic range, resulting in 80% adsorption of hormones from water. Similarly, the binary interaction of pH and initial hormones concentration (Figure 7c) showed that % adsorption of 80–100 can be obtained at alkaline pH and lower concentration (1–3 mg/L). An increase in concentration to 5 mg/L at pH (8, 9) may reduce

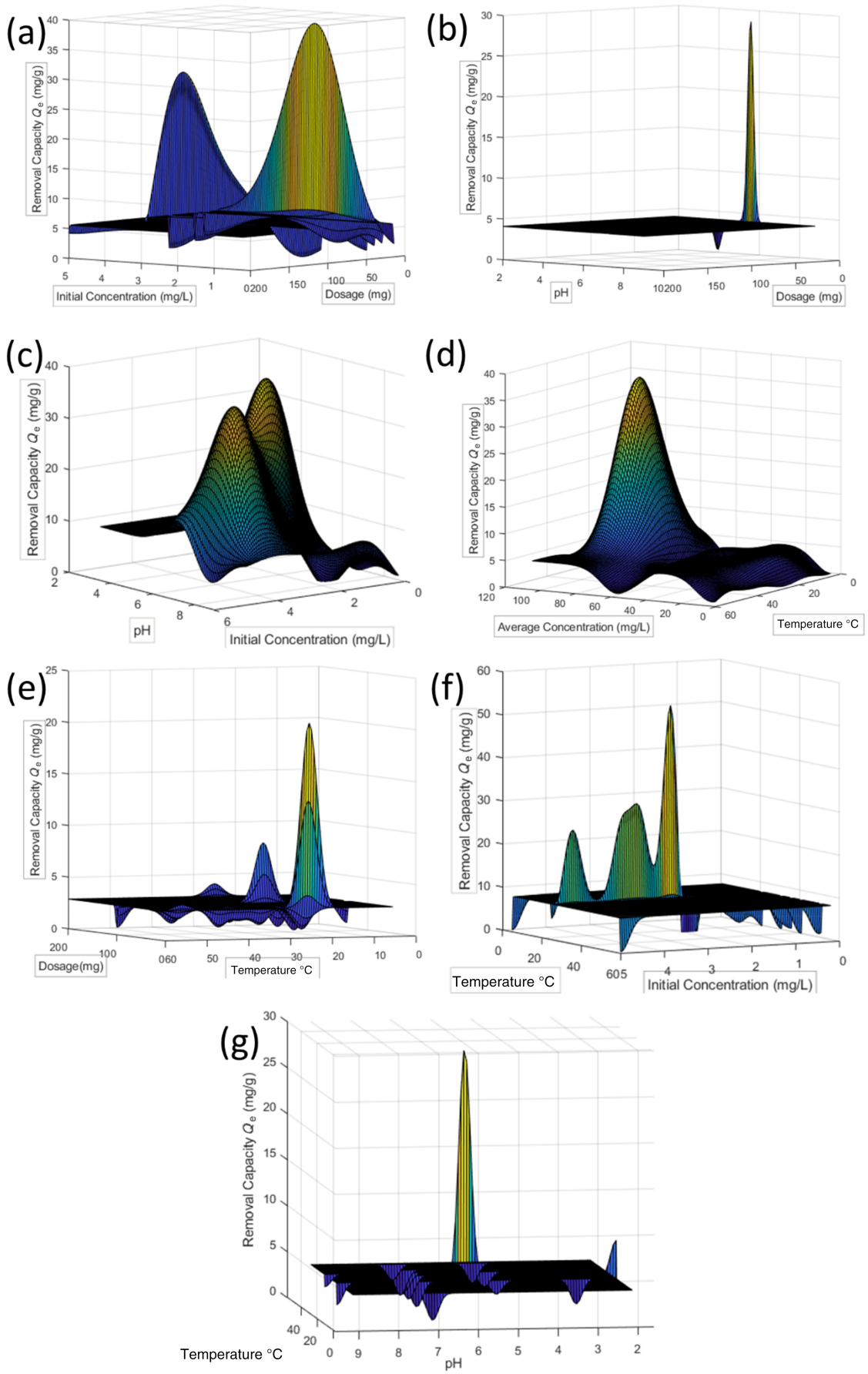

FIGURE 8 | 3D PDPs of adsorption removal capacity with input parameters.

the % adsorption at fixed adsorbent dosage due to adsorbent saturation.

The relationship between dosage and temperature depicted in Figure 7d reveals that lower dosage (0–100 mg/L) and high temperatures (40°C–60°C) are unfavorable for % removal of hormones. The maximum % adsorption of 80%–90% of hormones is achieved at high adsorbent dosage > 100 mg/L and temperature < 40°C. This indicates that the % adsorption of hormones onto adsorbent is exothermic in nature, and elevated temperature leads to a reduction in % removal. Similarly, trends were observed in the case of initial concentration and temperature, as indicated in Figure 7e. The effect of pH versus temperature plot, as shown in Figure 7f, shows that the % adsorption of hormones onto adsorbent surface is maximum at pH (6–8) and adsorption temperature < 40°C which is associated to the surface chemistry of adsorbent and ionic charge of hormones in aqueous phase as discussed in above section. The increase in adsorption temperature above 40°C showed lower sorption performance at all solution pH 2–10. Saima Farooq et al. also found in their study that adsorption removal of E1 was exothermic using copper selenide nanoflakes [54].

The investigation into the adsorption capacity of estrogenic hormones encompasses a comprehensive analysis utilizing three-dimensional PDPs in Figure 8. Three-dimensional plots elaborate on the importance of various input parameters simultaneously with the output. Keeping output as the reference, the impact of multiple inputs is considered and elucidated. In these plots, the adsorption capacity serves as the fixed *z*-axis, while various input parameters are systematically varied to discern their collective impact on the adsorption process. In general, the dosage and solution pH have a critical high influence, which impacts the adsorption capacity. Meanwhile, dosage against temperature shows a considerable range over which the capacity is altered and affected. Temperature and initial concentration develop the ranged scheme in which the capacity bears a brief region and is reduced. The capacity is achieved optimally on each temperature and initial concentration boundary condition. The dosage and initial concentration impact hormone removal crucially [55]. The dosage range of 50–150 mg reflects the highly effective zone to achieve the removal of the hormone.

As can be seen, Figure 8a explores the interaction between dosage and initial concentration and reveals the intricate relationship between these parameters. The analysis provides insights into the synergistic effects of dosage and initial concentration on the overall adsorption performance of estrogenic hormones. It is indicated that the adsorption capacity of adsorbent is maximum at an adsorbent dosage of 100–150 mg/L and initial concentration of 1–2 mg/L, and the performance tends to decrease with an increase in the hormone concentration. In addition, the exploration of the binary relationship of pH and dosage through a three-dimensional plot facilitates a visual representation of the interplay between these variables. The results shown in Figure 8b demonstrate that a pH value of 7–9 and an adsorbent dosage of 100–120 mg/L resulted in increased adsorption capacity to a maximum of 28 mg/g. As discussed in the previous section, lower solution pH values are not favorable for hormone adsorption and result in low adsorption capacity. The 3D plot in Figure 8c sheds light on the joint influence of solution pH and initial concentration on adsorption capacity. The result showed that the adsorption capacity increased with increasing pH from 2 to 6, reaching a maximum of 38 mg/g at pH 7 and a concentration range of 4–7 mg/L. A further increase in pH from 7 to 10 resulted in a decrease in adsorption capacity. The interdependence between these parameters becomes apparent in the analysis of average concentration and temperature (Figure 8d). This insight assists in comprehending how temperature influences average concentration and subsequently affects the adsorption capacity of estrogenic hormones. The maximum adsorption capacity of 38 mg/g is obtained at an average concentration of 50–60 mg/L and temperature < 40°C. The lower adsorption temperature of 20°C–40°C favors the adsorption, indicating the exothermic nature of the adsorption process.

In exploring temperature interaction with dosage, initial concentration, and pH, as shown in Figure 8e–g, The three-dimensional plots highlight the intricate relationship between these parameters, suggesting the same optimum values for desired removal capacity. This visualization assists in discerning the temperature-dependent variations in adsorption capacity with changing input variables. This analysis contributes to understanding the nuanced effects of temperature and pH variations on the adsorption process of estrogenic hormones. In conclusion, these three-dimensional PDPs collectively provide comprehensive insights into the complex interactions among dosage, initial concentration, solution pH, average concentration, and temperature and their resulting impact on adsorption capacity. This knowledge informs well-optimized strategies with suitable condition ranges for removing estrogenic hormones from environmental matrices that can appropriately be used at the monitoring wastewater treatment plants.

3.6 | Shapley and Sensitivity Analysis

SHAP (Shapley Additive Explanations) stands out as a valuable tool in the realm of ML interpretation. This method, rooted in game theory and an extension of Shapley values, offers a practical and comprehensive approach to unraveling the intricacies of model predictions. By delving into the underlying principles of game theory, SHAP provides a robust framework that empowers users to gain nuanced insights into the output of ML algorithms. This strategic application of game theory enhances interpretability and contributes to a more profound understanding of the factors influencing model predictions, fostering transparency and informed decision-making in ML.

The dependence of each input parameter on the adsorption capacity of the adsorbent fiber can also be demonstrated using the SHAP study. In Figure 9a, the Shapley value for each of the input parameters is shown. Input parameters with positive values directly impact the adsorption capacity (Q_e), and the parameters with negative Shapley values indirectly impact the adsorption capacity. Except for time and pH, all other parameters influence the Q_e indirectly and are inversely related to Q_e. Time shows a positive Shapley value, which indicates that adsorption capacity increases with an increase in contact time and vice versa. A

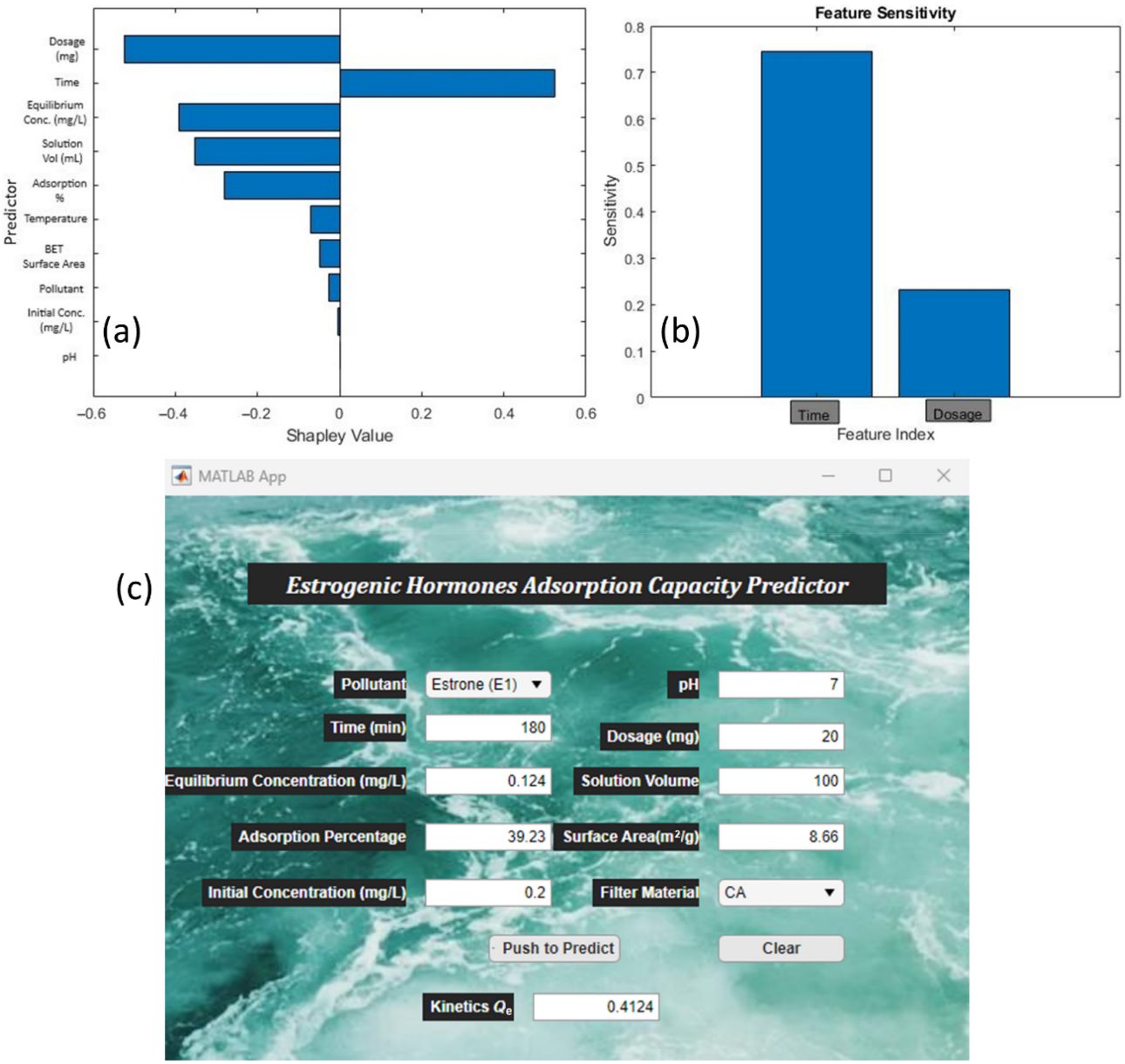

FIGURE 9 | (a) Shapley analysis, (b) sensitivity analysis, and (c) schematic representation of web-based graphical user interface.

plausible reason could be that interaction mechanisms between estrogenic hormone pollutants and adsorbent membranes can play a better part in adsorption with a sufficient time to reach equilibrium. In contrast, the Shapley value for pH is almost negligible, implying no contribution to the adsorption of estrogenic hormones on the surface of membranes is influenced by pH in the tested range. This can further be elaborated by the fact that the acid dissociation constant (pK_a value) of all estrogenic hormones is between 10.2 and 10.5, so they are stable at a lower pH, and no proton dissociation is observed in the 3–10 pH range discussed throughout the study. Therefore, change in pH has negligible dependence on capturing hormones via electrospun nanofiber membranes.

Sensitivity analysis plays a crucial role in evaluating the predictive capability of models (GPR in this case). A cross-validation employing a leave-one-out (LOO) strategy is applied by initialing a loop, and upon each successful iteration, one of the features from the dataset is excluded. The modified data are used to train the optimized GPR model. The prediction error is calculated after predicting from the formulated dataset. Once again, the same GPR model is trained with the help of the original dataset, containing all features, and the prediction error is calculated. The absolute difference between the prediction errors of the two scenarios determines the sensitivity of each feature. The value of sensitivity of each prominent feature is visualized and depicted with the help of a bar plot (Figure 9b). This analysis aids in understanding the impact of individual features on the predictive performance of the GPR model, providing insights into feature importance.

In the context of a GPR model predicting kinetic adsorption capacity (Q_e), the sensitivity analysis values of 0.74 for the "time" feature and 0.23 for the "dosage" feature indicate their respective impacts on the model's predictive accuracy. A sensitivity score of 0.74 for the "time" feature suggests that variations in

TABLE 4 | Comparison of ML predicted results to the experimental values at different conditions for removing EE2 hormone after 4 h using PU-P nanofibers.

Temperature (°C)	Solution concentration (mg/L)	pH	Dosage (mg)	Predicted Q_e (mg/g)	Experimental Q_e (mg/g)	Error (%)
40	0.3	7	20	2.170	2.110	2.84
40	0.3	7	10	3.038	3.043	0.16
25	0.2	5	10	1.639	1.587	3.28
55	0.2	9	10	2.435	2.364	3.00
25	0.4	5	10	3.809	3.932	3.13

time have a substantial effect on the predicted kinetics, implying that changes in the duration of a process or event significantly influence the outcome. On the other hand, the "dosage" feature with a sensitivity score of 0.23 indicates a moderate impact on the predicted kinetics, implying that alterations in dosage levels have a relatively lesser but still notable effect on the model's predictions. By understanding the sensitivity of these input features, stakeholders of monitoring plants can prioritize their interventions or adjustments, accordingly, focusing on factors that have the most substantial influence on the predicted kinetic adsorption capacity (Q_e) via the adsorption method. This insight derived from sensitivity analysis enhances the model's interpretability, guiding decision-making processes and potentially leading to more effective strategies in determining membranes' interaction kinetics at modern wastewater treatment plants.

3.7 | Graphical User Interface (GUI)

A GUI employs symbols, graphical icons, and user-friendly features to facilitate user interaction with electronic devices and allow users to provide their adsorption input parameters. The GUI provides a convenient platform for researchers to input and manipulate the key parameters, contributing to a comprehensive understanding of the adsorption process and streamlining the interaction with the adsorption model. The specific key parameters are time contact, initial and equilibrium concentration, temperature, solution pH, the volume of solution, the dosage of adsorbent, BET surface area, estrogenic hormone type, type of nanofiber adsorbent, and adsorption %. The software can predict the kinetic adsorption capacity (Q_e) in mg/g by employing the GPR model function optimized by GA in MATLAB 2021b from the developed application. Figure 9c depicts the GUI image of the adsorption capacity prediction calculator with a push button and input parameter conditions for the class of estrogenic hormones that include E1, E2, EE2, and E3.

3.8 | Experimental Results Validation

To validate the prediction of adsorption capacity from the most suitable model (GPR-GA), the EE2 hormone was used as a model pollutant to test the experimental adsorption capacity compared to the predicted values from GUI. The method for the detection and quantification of EE2 hormone, the fabrication of lab-synthesized polyurethane via the electrospinning process, and surface functionalized with PANI (PU-P) nanofibers is described in previous research [56]. The SEM nanofiber structure before and after functionalization is displayed in Figure S1. It can be seen from the morphology that nanofiber became dense and compact after coating with PANI (aniline was oxidized with ammonium peroxydisulfate), which improved its adsorption ability. Therefore, a set of five different conditions was randomly picked based on the generated set of experiments with varied adsorption parameters from software design expert v.11.0 using response surface methodology with a central composite design model. The experimental conditions of temperature, EE2 initial concentration, solution pH, and dosage of PU-P nanofibers are specified in Table 4, and the chromatogram of the standard mobile phase prior to analysis and EE2 calibration curve are displayed in Figures S2 and S3. Additionally, the chromatograms of each experiment for the control and tested samples are detailed in Figures S4–S8. As can be seen, for all experimental runs, the actual obtained adsorption capacity values at equilibrium time (210 min) are close to the predicted values from the optimized ML model, with the maximum reported error of only 3.28% (Experiment 3) and as low as 0.16% (Experiment 2). Moreover, it is noteworthy to mention that at the optimum removal percentage (85.5%) in Experiment 4, the adsorption capacity is reported to be 2.364 mg/L experimentally, which is firmly in agreement with the predicted value of 2.435 mg/g with an error of 3% using GUI. The experimentally verified results indicate the accuracy and precision of the implemented GPR-GA optimized ML model to create a valuable application for waste treatment plants to monitor and concurrently remove the estrogenic hormones over the recommended concentration limits.

4 | Conclusion

The adsorption capacity of estrogenic hormones via adsorption was successfully predicted by employing four different ML models optimized with PSO and GA, and the hyperparameters were fine-tuned. The features selection by GA and PSO differed for the GPR model, yet GA was better aligned to select the most effective features for the prediction function. It was found that the optimized R^2 values of all models were relatively better with GA compared to the PSO technique, except in the case of SVM, where the results obtained were better using PSO with R^2 of 0.9721 for testing data. The GA-integrated GPR model outperformed the rest of the optimized ML models in predicting the adsorption capacity with an impressive R^2 of 0.999 for both training and testing data. The values of testing data for

other models in descending order were ELT ($R^2 = 0.9976$ and RMSE = $4.3458e^{-17}$), DT ($R^2 = 0.9586$ and RMSE = $2.4673e^{-16}$), and SVM ($R^2 = 0.7110$ and RMSE = 0.0639). The GPR exhibited satisfactory performance even with optimization via PSO but remained second to the prediction results produced by GA optimization. Furthermore, the PDP demonstrated that temperature, dosage, initial concentration, contact time, and pH play vital roles as adsorption parameters chosen using GPR integrated with GA-based methodology; additionally, Shapley's analysis further revealed time and dosage were identified as the most sensitive parameters. Moreover, a GUI was developed employing the GPR-GA hybrid model. The credibility of the developed GUI was validated with five different experimental tests, and it was found that the error difference remained below 3.3% between the experimental results and predicted values of kinetic adsorption capacity. This study provided a precise and reliable prediction hybrid methodology for the simultaneous adsorption capacity of the different estrogenic hormones. Thus, the findings in this work through ML could significantly enhance water treatment processes by providing a more efficient method for removing harmful estrogenic hormones from wastewater.

Acknowledgments

The authors gratefully acknowledge the financial support from the Ministry of Education, Youth, and Sports of the Czech Republic—DKRVO (RP/CPS/2024-28/002). Muhammad Yasir also expresses his gratitude for support within the "Creativity, Intelligence and Talent for the Zlin Region" (CIT—ZK) program. The authors would also like to acknowledge the Centre of Polymer Systems (CPS) situated at Tomas Bata University in Zlin, Czech Republic, to use the available research facilities to conduct this research.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that supports the findings of this study are available in the Supporting Information of this article.

References

1. J. A. McLachlan, E. Simpson, and M. Martin, "Endocrine Disrupters and Female Reproductive Health," *Best Practice & Research. Clinical Endocrinology & Metabolism* 20, no. 1 (2006): 63–75, https://doi.org/10.1016/j.beem.2005.09.009.

2. A. I. Schäfer, K. Stelzl, M. Faghih, et al., "Poly(Ether Sulfone) Nanofibers Impregnated With β-Cyclodextrin for Increased Micropollutant Removal From Water," *ACS Sustainable Chemistry & Engineering* 6, no. 3 (2018): 2942–2953, https://doi.org/10.1021/acssuschemeng.7b02214.

3. X. Dong, S. Yu, W. Yang, L. Cheng, Y. Tang, and D. Chen, "Homogeneous Oxidation of EE2 by Ozone: Influencing Factors, Degradation Pathway, and Toxicity Assessment," *Journal of Environmental Chemical Engineering* 12, no. 2 (2024): 112360, https://doi.org/10.1016/j.jece.2024.112360.

4. Z. Tang, Z. h. Liu, H. Wang, Z. Dang, and Y. Liu, "A Review of 17α-Ethynylestradiol (EE2) in Surface Water Across 32 Countries: Sources, Concentrations, and Potential Estrogenic Effects," *Journal of Environmental Management* 292 (2021): 112804, https://doi.org/10.1016/j.jenvman.2021.112804.

5. D. Montes-Grajales and J. Olivero-Verbel, "EDCs Databank: 3D-Structure Database of Endocrine Disrupting Chemicals," *Toxicology* 327 (2015): 87–94, https://doi.org/10.1016/j.tox.2014.11.006.

6. S. Sood, S. Shekhar, and W. Santosh, "Dimorphic Placental Stress: A Repercussion of Interaction Between Endocrine Disrupting Chemicals (EDCs) and Fetal Sex," *Medical Hypotheses* 99 (2017): 73–75, https://doi.org/10.1016/j.mehy.2017.01.002.

7. Y. Chen, Y. Zhang, L. Luo, et al., "A Novel Templated Synthesis of C/N-Doped β-Bi_2O_3 Nanosheets for Synergistic Rapid Removal of 17α-Ethynylestradiol by Adsorption and Photocatalytic Degradation," *Ceramics International* 44, no. 2 (2018): 2178–2185, https://doi.org/10.1016/j.ceramint.2017.10.173.

8. J. Han, W. Qiu, and W. Gao, "Adsorption of Estrone in Microfiltration Membrane Filters," *Chemical Engineering Journal* 165, no. 3 (2010): 819–826, https://doi.org/10.1016/j.cej.2010.10.024.

9. M. Lasich and V. T. Adeleke, "Extraction of Estrogenic Pollutants in Aqueous Solution Using Poly(Lactic Acid)," *Journal of Molecular Liquids* 377 (2023): 121577, https://doi.org/10.1016/j.molliq.2023.121577.

10. A. M. Razekenari, A. E. Fereidouni, A. Movahedinia, and E. Z. Neyshabouri, "Impacts of Sublethal Concentrations of 17α-Ethinylestradiol (EE2) on Growth, Reproductive Performance, and Survival in Red Cherry Shrimp Neocaridina Davidi (Crustacea, Atyidae) During Consecutive Spawnings," *Aquatic Toxicology* 259 (2023): 106519, https://doi.org/10.1016/j.aquatox.2023.106519.

11. L. D. Nghiem and A. I. Schäfer, "Adsorption and Transport of Trace Contaminant Estrone in NF/RO Membranes," *Environmental Engineering Science* 19, no. 6 (2002): 441–451, https://doi.org/10.1089/109287502320963427.

12. G. M. Solomon and T. Schettler, "Environment and Health: 6. Endocrine Disruption and Potential Human Health Implications," *CMAJ* 163, no. 11 (2000): 1471–1476.

13. T. T. Schug, A. F. Johnson, L. S. Birnbaum, et al., "Minireview: Endocrine Disruptors: Past Lessons and Future Directions," *Molecular Endocrinology* 30, no. 8 (2016): 833–847, https://doi.org/10.1210/me.2016-1096.

14. European Commission, "Directive (EU) 2020/2184 of the European Parlliament and of the Council of 16 December 2020 on the Quality of Water Intended for Human Consumption," *Official Journal of the European Union* 435 (2020): 1–61, https://eur-lex.europa.eu/eli/dir/2020/2184/oj.

15. D. V. Cosma, M. C. Rosu, C. Socaci, et al., "Adsorption-Catalysis Synergy in the Visible-Light-Driven Removal of 17β-Estradiol by (Au) TiO_2 Nanotubes-Graphene Composites," *Journal of Environmental Chemical Engineering* 12 (2024): 112885, https://doi.org/10.1016/j.jece.2024.112885.

16. M. Grzegorzek, K. Wartalska, and R. Kowalik, "Occurrence and Sources of Hormones in Water Resources—Environmental and Health Impact," *Environmental Science and Pollution Research* 31 (2024): 37922, https://doi.org/10.1007/s11356-024-33713-z.

17. M. Adeel, X. Song, Y. Wang, D. Francis, and Y. Yang, "Environmental Impact of Estrogens on Human, Animal and Plant Life: A Critical Review," *Environment International* 99 (2017): 107–119, https://doi.org/10.1016/j.envint.2016.12.010.

18. D. V. Henley, J. Lindzey, and K. S. Korach, *Steroid Hormones BT—Endocrinology: Basic and Clinical Principles*, eds. S. Melmed and P. M. Conn (Totowa, NJ: Humana Press, 2005), 49–65, https://doi.org/10.1007/978-1-59259-829-8_4.

19. M. B. Abdul-Kareem, H. M. Rashid, W. H. Hassan, et al., "Preparation of Coated MgFe Layered Double Hydroxide Nanoparticles on Cement Kiln Dust and Intercalated With Sodium Dodecyl Sulfate as an Intermediate Layer for the Adsorption of Estrogen From Water," *Chemosphere* 344 (2023): 140338, https://doi.org/10.1016/j.chemosphere.2023.140338.

20. Y. K. K. Koh, T. Y. Chiu, A. Boobis, E. Cartmell, M. D. Scrimshaw, and J. N. Lester, "Treatment and Removal Strategies for Estrogens From Wastewater," *Environmental Technology* 29, no. 3 (2008): 245–267, https://doi.org/10.1080/09593330802099122.

21. X. Gao, S. Kang, R. Xiong, and M. Chen, "Environment-Friendly Removal Methods for Endocrine Disrupting Chemicals," *Sustainability* 12, no. 18 (2020): 7615, https://doi.org/10.3390/su12187615.

22. J. Hartmann, R. Beyer, and S. Harm, "Effective Removal of Estrogens From Drinking Water and Wastewater by Adsorption Technology," *Environmental Processes* 1, no. 1 (2014): 87–94, https://doi.org/10.1007/s40710-014-0005-y.

23. C. Cheng, X. Li, X. Yu, M. Wang, and X. Wang, "Chapter 14—Electrospun Nanofibers for Water Treatment," in *Micro and Nano Technologies*, eds. B. Ding, X. Wang, and J. Yu (Norwich, NY: William Andrew Publishing, 2019), 419–453, https://doi.org/10.1016/B978-0-323-51270-1.00014-5.

24. L. Jiang, Y. Liu, S. Liu, et al., "Adsorption of Estrogen Contaminants by Graphene Nanomaterials Under Natural Organic Matter Preloading: Comparison to Carbon Nanotube, Biochar, and Activated Carbon," *Environmental Science & Technology* 51, no. 11 (2017): 6352–6359, https://doi.org/10.1021/acs.est.7b00073.

25. C. Peiris, S. Nawalage, J. J. Wewalwela, S. R. Gunatilake, and M. Vithanage, "Biochar Based Sorptive Remediation of Steroidal Estrogen Contaminated Aqueous Systems: A Critical Review," *Environmental Research* 191 (2020): 110183, https://doi.org/10.1016/j.envres.2020.110183.

26. F. Ogata, H. Tominaga, H. Yabutani, and N. Kawasaki, "Removal of Estrogens From Water Using Activated Carbon and Ozone," *Journal of Oleo Science* 60, no. 12 (2011): 609–611, https://doi.org/10.5650/jos.60.609.

27. X. Wang, N. Liu, Y. Liu, et al., "Adsorption Removal of 17β-Estradiol From Water by Rice Straw-Derived Biochar With Special Attention to Pyrolysis Temperature and Background Chemistry," *International Journal of Environmental Research and Public Health* 14, no. 10 (2017): 1213, https://doi.org/10.3390/ijerph14101213.

28. O. Ifelebuegu and A. Removal, "Of Steriod Hormones by Activated Carbon Adsorption—Kinetic and Thermodynamic Studies," *Journal of Environmental Protection* 3, no. 6 (2012): 469–475, https://doi.org/10.4236/jep.2012.36057.

29. M. H. Zarghi, A. Roudbari, S. Jorfi, and N. Jaafarzadeh, "Removal of Estrogen Hormones (17β-Estradiol and Estrone) From Aqueous Solutions Using Rice Husk Silica," *Chemical and Biochemical Engineering Quarterly* 33, no. 2 (2019): 281–293, https://doi.org/10.15255/CABEQ.2018.1542.

30. A. E. Burgos Castellanos, T. A. Ribeiro-Santos, and R. M. Lago, "Porous Expanded Vermiculite Containing Intercalated Cetyltrimethylammonium: A Versatile Sorbent for the Hormone Ethinylestradiol From Aqueous Medium," *International journal of Environmental Science and Technology* 16, no. 7 (2019): 2877–2884, https://doi.org/10.1007/s13762-018-1901-x.

31. K. B. Debs, H. D. T. da Silva, M. de Lourdes Leite de Moraes, E. N. V. M. Carrilho, S. G. Lemos, and G. Labuto, "Biosorption of 17α-Ethinylestradiol by Yeast Biomass From Ethanol Industry in the Presence of Estrone," *Environmental Science and Pollution Research* 26, no. 28 (2019): 28419–28428, https://doi.org/10.1007/s11356-019-05202-1.

32. L. Liu, Z. Lu, W. Cai, G. Owens, and Z. Chen, "Green rGO/FeNPs Nanocomposites Activated Peroxydisulfate for the Removal of Mixed 17β-Estradiol and Estriol," *Environmental Research* 245 (2024): 118057, https://doi.org/10.1016/j.envres.2023.118057.

33. S. Mohammad, J. Fitzgerald, R. L. J. Robinson, and K. A. M. Gasem, "Experimental Uncertainties in Volumetric Methods for Measuring Equilibrium Adsorption," *Energy & Fuels* 23, no. 5 (2009): 2810–2820, https://doi.org/10.1021/ef8011257.

34. M. N. Aslam Khan, U. Ghafoor, A. Abdullah, et al., "Prediction of Thermal Diffusivity of Volcanic Rocks Using Machine Learning and Genetic Algorithm Hybrid Strategy," *International Journal of Thermal Sciences* 192 (2023): 108403, https://doi.org/10.1016/j.ijthermalsci.2023.108403.

35. M. H. Naveed, J. Gul, M. N. A. Khan, S. R. Naqvi, L. Štěpanec, and I. Ali, "Torrefied Biomass Quality Prediction and Optimization Using Machine Learning Algorithms," *Chemical Engineering Journal Advances* 19 (2024): 100620, https://doi.org/10.1016/j.ceja.2024.100620.

36. P. L. Donti and J. Z. Kolter, "Machine Learning for Sustainable Energy Systems," *Annual Review of Environment and Resources* 46, no. 1 (2021): 719–747, https://doi.org/10.1146/annurev-environ-020220-061831.

37. M. Nilashi, P. F. Rupani, M. M. Rupani, et al., "Measuring Sustainability Through Ecological Sustainability and Human Sustainability: A Machine Learning Approach," *Journal of Cleaner Production* 240 (2019): 118162, https://doi.org/10.1016/j.jclepro.2019.118162.

38. U. Kanewala, J. M. Bieman, and A. Ben-Hur, "Predicting Metamorphic Relations for Testing Scientific Software: A Machine Learning Approach Using Graph Kernels," *Software Testing, Verification and Reliability* 26, no. 3 (2016): 245–269, https://doi.org/10.1002/stvr.1594.

39. S. Zhong, K. Zhang, M. Bagheri, et al., "Machine Learning: New Ideas and Tools in Environmental Science and Engineering," *Environmental Science & Technology* 55, no. 19 (2021): 12741–12754, https://doi.org/10.1021/acs.est.1c01339.

40. Z. Jin, J. Shang, Q. Zhu, C. Ling, W. Xie, and B. Qiang, "RFRSF: Employee Turnover Prediction Based on Random Forests and Survival Analysis," in *Lecture Notes in Computer Science* (Cham, Switzerland: Springer Nature, 2020), 503–515, https://doi.org/10.1007/978-3-030-62008-0_35.

41. L. Breiman, "Random Forests," *Machine Learning* 45, no. 1 (2001): 5–32, https://doi.org/10.1023/A:1010933404324.

42. F. Khalid, M. N. Aslam, M. A. Ghani, N. Ahmad, and S. K. Abdullah, "Aging Prediction in Single Based Propellants Using Hybrid Strategy of Machine Learning and Genetic Algorithm," *Chemometrics and Intelligent Laboratory Systems* 245 (2024): 105058, https://doi.org/10.1016/j.chemolab.2023.105058.

43. N. Majaw and S. S. Ahmed, "Exploring Data Distributions Using Box and Whisker Plot Analysis," in *14th International Conference on Computing Communication and Networking Technologies (ICCCNT)* (Delhi, India: Institute of Electrical and Electronics Engineers (IEEE), 2023), 1–8, https://doi.org/10.1109/ICCCNT56998.2023.10308191.

44. M. Ngueajio, G. Washington, D. B. Rawat, and Y. Ngueabou, "Intrusion Detection Systems Using Support Vector Machines on the KDDCUP'99 and NSL-KDD Datasets: A Comprehensive Survey," in *Proceedings of SAI Intelligent Systems Conference* (Cham: Springer International Publishing, 2022), 609–629, https://doi.org/10.1007/978-3-031-16078-3_42.

45. S. Singh, K. S. Parmar, S. J. S. Makkhan, J. Kaur, S. Peshoria, and J. Kumar, "Study of ARIMA and Least Square Support Vector Machine (LS-SVM) Models for the Prediction of SARS-CoV-2 Confirmed Cases in the Most Affected Countries," *Chaos, Solitons & Fractals* 139 (2020): 110086, https://doi.org/10.1016/j.chaos.2020.110086.

46. A. G. Wilson, D. A. Knowles, and Z. Ghahramani, "Gaussian Process Regression Networks," *Proceedings of the 29th International Conference on Machine Learning* 1 (2012): 599–606.

47. A. Hashemizadeh, A. Maaref, M. Shateri, A. Larestani, and A. Hemmati-Sarapardeh, "Experimental Measurement and Modeling of Water-Based Drilling Mud Density Using Adaptive Boosting Decision Tree, Support Vector Machine, and K-Nearest Neighbors: A Case Study From the South Pars Gas Field," *Journal of Petroleum Science and Engineering* 207 (2021): 109132, https://doi.org/10.1016/j.petrol.2021.109132.

48. S. Gul, "Machine Learning Applications in Drilling Fluid Engineering: A Review," in *Proceedings of the ASME 2021 40th International Conference on Ocean, Offshore and Arctic Engineering*, OMAE2021 June 21-30, 2021, Virtual, Online, OMAE2021-63094. (American Society of Mechanical Engineers, 2021), https://doi.org/10.1115/OMAE2021-63094.

49. D. Kocev, C. Vens, J. Struyf, and S. Džeroski, "Tree Ensembles for Predicting Structured Outputs," *Pattern Recognition* 46, no. 3 (2013): 817–833, https://doi.org/10.1016/j.patcog.2012.09.023.

50. A. Arabali, M. Ghofrani, M. Etezadi-Amoli, M. S. Fadali, and Y. Baghzouz, "Genetic-Algorithm-Based Optimization Approach for Energy Management," *IEEE Transactions on Power Delivery* 28, no. 1 (2013): 162–170, https://doi.org/10.1109/TPWRD.2012.2219598.

51. M. E. Wahyuhadi, R. A. Kusumadewi, and R. Hadisoebroto, "Effect of Contact Time on the Adsorption Process of Activated Carbon From Banana Peel in Reducing Heavy Metal cd and Dyes Using a Stirring Tub (Pilot Scale)," *IOP Conference Series: Earth and Environmental Science* 1203, no. 1 (2023): 12035, https://doi.org/10.1088/1755-1315/1203/1/012035.

52. J. Zdarta, F. Ciesielczyk, M. Bilal, et al., "Inorganic Oxide Systems as Platforms for Synergistic Adsorption and Enzymatic Conversion of Estrogens From Aqueous Solutions: Mechanism, Stability and Toxicity Studies," *Journal of Environmental Chemical Engineering* 11, no. 2 (2023): 109443, https://doi.org/10.1016/j.jece.2023.109443.

53. A. Tomczyk, Z. Sokołowska, and P. Boguta, "Biochar Physicochemical Properties: Pyrolysis Temperature and Feedstock Kind Effects," *Reviews in Environmental Science and Bio/Technology* 19, no. 1 (2020): 191–215, https://doi.org/10.1007/s11157-020-09523-3.

54. S. Farooq, R. Cai, J. McGettrick, et al., "Visible-Light Induced Photocatalytic Degradation of Estrone (E1) With Hexagonal Copper Selenide Nanoflakes in Water," *Process Safety and Environment Protection* 172 (2023): 1–15, https://doi.org/10.1016/j.psep.2023.02.003.

55. C. P. Silva, M. Otero, and V. Esteves, "Processes for the Elimination of Estrogenic Steroid Hormones From Water: A Review," *Environmental Pollution* 165 (2012): 38–58, https://doi.org/10.1016/j.envpol.2012.02.002.

56. M. Yasir, F. Asabuwa Ngwabebhoh, T. Šopík, H. Ali, and V. Sedlařík, "Electrospun Polyurethane Nanofibers Coated With Polyaniline/Polyvinyl Alcohol as Ultrafiltration Membranes for the Removal of Ethinylestradiol Hormone Micropollutant From Aqueous Phase," *Journal of Environmental Chemical Engineering* 10, no. 3 (2022): 107811, https://doi.org/10.1016/j.jece.2022.107811.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.