



OPEN

An insight into tetracycline photocatalytic degradation by MOFs using the artificial intelligence technique

Majedeh Gheytnazadeh¹, Alireza Baghban²✉, Sajjad Habibzadeh¹✉, Karam Jabbour³, Amin Esmaeili⁴, Ahmad Mohaddespour³ & Otman Abida³

Tetracyclines (TCs) have been extensively used for humans and animal diseases treatment and livestock growth promotion. The consumption of such antibiotics has been ever-growing nowadays due to various bacterial infections and other pathologic conditions, resulting in more discharge into the aquatic environments. This brings threats to ecosystems and human bodies. Up to now, several attempts have been made to reduce TC amounts in the wastewater, among which photocatalysis, an advanced oxidation process, is known as an eco-friendly and efficient technology. In this regard, metal organic frameworks (MOFs) have been known as the promising materials as photocatalysts. Thus, studying TC photocatalytic degradation by MOFs would help scientists and engineers optimize the process in terms of effective parameters. Nevertheless, the costly and time-consuming experimental methods, having instrumental errors, encouraged the authors to use the computational method for a more comprehensive assessment. In doing so, a wide-ranging databank including 374 experimental data points was gathered from the literature. A powerful machine learning method of Gaussian process regression (GPR) model with four kernel functions was proposed to estimate the TC degradation in terms of MOFs features (surface area and pore volume) and operational parameters (illumination time, catalyst dosage, TC concentration, pH). The GPR models performed quite well, among which GPR-Matern model shows the most accurate performance with R^2 , MRE, MSE, RMSE, and STD of 0.981, 12.29, 18.03, 4.25, and 3.33, respectively. In addition, an analysis of sensitivity was carried out to assess the effect of the inputs on the TC photodegradation by MOFs. It revealed that the illumination time and the surface area play a significant role in the decomposition activity.

In recent decades, water pollution and wastewater treatment have become a global concern^{1,2}. Notably, the release of wastewater having antibiotics can lead to severe environmental and human health issues³. The discharged antibiotics in water may contribute to the food chain and enter the living things in time. Therefore, they are known as a category of emerging organic pollutants⁴. Among the antibiotics, tetracycline (TC) can be found in different water environments, from underground waters to surface waters. It is raised from its rampant overuse, non-biodegradability, and easy dissolution in water^{5,6}. Indeed, it is challenging to degrade TC in natural conditions due to its aromatic rings with stable molecular structure⁷. The TC residual can bring antibiotic-resistant bacteria and genes into the ecological system⁸. Therefore, it is vital and urgent to control the TC concentration level in the environment.

For this purpose, many effective methods have been examined to eliminate the TCs pollution, such as adsorption^{9,10}, ozonation¹¹, ion-exchange¹², electrocoagulation¹³, and biological methods¹⁴. Nevertheless, there are undeniable shortcomings with these approaches. For example, through the physical separation methods without mineralization, the TCs need post-treatments to be degraded after separation¹⁵. The biological degradation procedures show low removal efficiency due to their antibacterial nature¹⁶. Moreover, the high treatment

¹Surface Reaction and Advanced Energy Materials Laboratory, Chemical Engineering Department, Amirkabir University of Technology (Tehran Polytechnic), Tehran, Iran. ²Chemical Engineering Department, Amirkabir University of Technology (Tehran Polytechnic), Mahshahr Campus, Mahshahr, Iran. ³College of Engineering and Technology, American University of the Middle East, Kuwait City, Kuwait. ⁴Department of Chemical Engineering, School of Engineering Technology and Industrial Trades, College of the North Atlantic-Qatar, Doha, Qatar. ✉email: Alireza_baghban@alumni.ut.ac.ir; sajjad.habibzadeh@mail.mcgill.ca

cost or generating of carcinogenic chlorinated intermediates limit the use of traditional chemical oxidation routes such as ozonation or chlorine oxidation¹⁷. Lately, semiconductor photocatalysis, an advanced oxidation process (AOP), has attracted massive attention due to its low cost and significant efficiency^{18–20}. It is based on the producing hydroxyl radical ($\cdot\text{OH}$), a highly active agent that degrades TC to a more minor or inorganic substance¹⁸. Selecting the appropriate photocatalyst is crucial to achieving excellent efficiency in photocatalysis.

Metal–organic frameworks (MOFs), a new class of porous crystalline substances, have received significant interest as a porous material in different processes^{21–25}. The MOFs are one-, two-, or three-dimensional network structures constructed by metal ions/-clusters coordinated to organic ligands (linkers). The most exciting property of MOFs is their adaptable framework which can be modified by changing the linkers. Thus, a significantly porous material with excellent surface area could be provided, which benefits surface phenomena²⁶. Furthermore, some MOFs with semiconductor properties have been utilized as photocatalysts^{27,28}. Several studies have reported the effective performance of MOFs and their derivatives in the photodegradation of TCs^{27,29,30}.

Several factors influence photocatalysis efficiency, such as the pollutant concentration, the photocatalyst dosage, and the solution pH^{31–33}. Extensive research in the photocatalysis area has been conducted to find the optimum parameters, making it possible to gather tremendous data sets^{34–36}. However, a complicated nonlinear relationship exists between the effective parameters and the degradation efficiency (the output). On the other hand, conducting extra different experiments to figure out the relationship is very time-consuming and costly, and also, it is not possible to consider all the influential experimental variables. Therefore, to save cost and experimental time, computational modeling is utilized as an alternative to assess the photocatalytic degradation of organic pollutants. On the other hand, as the complexity of the case study increases, the CPU-consuming and calculation time of the simulation increase significantly. This matter inspired the authors to use alternative methods such as machine learning (ML) techniques.

The robust ML methods can recognize the relationship between the different parameters of the process and their corresponding outcomes through passing the requirement to solve theoretical equations^{37,38}. The prognostication of nonlinear and complicated systems performance, high reliability, recognizing each influencing parameter's effect, and finding their optimal are some of the ML approaches advantages^{39–49}. Several researchers have applied different ML methods to anticipate the contaminant photocatalytic removal based on different operational variables as inputs^{37–39,50–53}.

In this work, for the first time, an effort is performed to use the gaussian process regression (GPR) model with four kernel functions to estimate the TC degradation by MOFs photocatalyst in the photocatalytic wastewater treatment process. A large experimental data bank of TC photodegradation by different MOFs is gathered for this aim. The operational parameters of TC concentration, photocatalyst amount, solution pH, irradiation time, and MOFs structural parameters of specific surface area and pore volume are chosen as input variables. Various statistical analyses were performed to study the developed GPR models. Also, sensitivity analysis was applied to show the most influential parameter(s) in the TC removal by MOFs.

Methodology

Data collection. An experimental data set with the total number of 374 of TC photocatalytic degradation using MOF as the photocatalyst, were collected from the reported studies^{54–63} and presented in the supplementary information. The assessed MOFs are MIL-88A, MIL-101(Fe), MIL-125(Ti), Cu-TCPP, PI/Uio-66-NH₂, AgI/Uio-66-NH₂, PANI/MIL-100(Fe), In₂S₃/MIL-125(Ti), Ag₂S/MIL-53(Fe), and WO₃/GO/Uio-66. Also, the catalyst dosage (g/L), the initial concentration of TC (mg/L), the solution pH, and the irradiation time (min), along with MOFs' structural features of pore volume (cm³/g) and specific surface area (m²/g), are the input parameters. To establish the most accurate model, 80% of the data set was randomly separated as a training set while the rest (20%) was considered testing data set to evaluate the preciseness of the developed model. Model accuracy examination was quantified through calculation of statistical parameters such as R², mean relative error (MRE), root-mean-square error (RMSE), mean-square error (MSE), and the standard deviation (STD), which are given as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^n [x_i^{\text{predicted}} - x_i^{\text{experimental}}]^2}{\sum_{i=1}^n [x_i^{\text{predicted}} - x_m]^2} \quad (1)$$

$$STD = \sqrt{\frac{\sum_{i=1}^n (x_i^{\text{predicted}} - x_m)^2}{n}} \quad (2)$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (x_i^{\text{predicted}} - x_i^{\text{experimental}})^2 \quad (3)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (x_i^{\text{predicted}} - x_i^{\text{experimental}})^2}{n}} \quad (4)$$

$$MRE = \frac{1}{n} \sum_{i=1}^n \frac{|x_i^{\text{predicted}} - x_i^{\text{experimental}}|}{x_i^{\text{experimental}}} \quad (5)$$

Gaussian process regression (GPR). GPR model is an effective nonparametric and probabilistic supervised ML method capable of modeling nonlinear complicated issues⁶⁴. It utilizes the Gaussian process to perform regression. One of the main attractive features of this approach is its flexible algorithm in the description of the uncertainty⁶⁵.

In general, for GPR modeling: suppose $T = \{x_{T,i}, y_{T,i}\}_{i=1}^n$ and $L = \{x_{L,i}, y_{L,i}\}_{i=1}^n$ are randomly chosen test and training data sets, respectively, where x and y are the input and the outcome variables. The GPR modeling starts from the following equation:

$$y_{L,i} = f(x_{L,i}) + \varepsilon_{L,i}, i = 1, 2, 3, \dots, n \quad (6)$$

$$\varepsilon \sim N(0, \sigma_{\text{noise}}^2 I_n) \quad (7)$$

where x_L and y_L indicate the independent variables and the outcomes of the training data points, respectively. Also, ε is the observation noise, σ_{noise}^2 represents the variance of the noise, and I_n is the unit array. Similarly, we can write for test data set:

$$y_{T,i} = f(x_{T,i}) + \varepsilon_{T,i}, i = 1, 2, 3, \dots, n \quad (8)$$

The symbols have the same definition as mentioned so far but for the test data set. Therefore, the Gaussian noise model connects every calculated y to the considered $f(x)$ function. According to the GPR model, $f(x)$ is a random function which can be determined by its corresponding mean $m(x)$ and covariance $k(x, x')$ (also called kernel) functions.

$$f(x_{L,i}) \sim GP(m(x), k(x, x')) \quad (9)$$

The $m(x)$ can be determined through utilizing the explicit basis functions, but usually, it is considered as zero in the simplified calculations^{66,67}:

$$f(x_{L,i}) \sim GP(0, k(x, x')) \quad (10)$$

Combination of Eqs. (6) and (10) give the distribution of y :

$$y \sim N(0, k(x, x') + \sigma_{\text{noise}}^2 I_n) \quad (11)$$

According to the mentioned parameters and definitions, we have:

$$\begin{bmatrix} \vec{f}_L \\ \vec{f}_T \end{bmatrix} \sim N\left(0, \begin{bmatrix} k(x_L, x_L) & k(x_L, x_T) \\ k(x_T, x_L) & k(x_T, x_T) \end{bmatrix}\right) \quad (12)$$

$$\begin{bmatrix} \vec{\varepsilon}_L \\ \vec{\varepsilon}_T \end{bmatrix} \sim N\left(0, \begin{bmatrix} \sigma_{\text{noise}}^2 I_n & 0 \\ 0 & \sigma_{\text{noise}}^2 I_n \end{bmatrix}\right) \quad (13)$$

The Gaussian expression is obtained from the summation of the recent two equations:

$$\begin{bmatrix} \vec{y}_L \\ \vec{y}_T \end{bmatrix} \sim N\left(0, \begin{bmatrix} k(x_L, x_L) + \sigma_{\text{noise}}^2 I_n & k(x_L, x_T) \\ k(x_T, x_L) & k(x_T, x_T) + \sigma_{\text{noise}}^2 I_n \end{bmatrix}\right) \quad (14)$$

The conditioning rule of Gaussians can be used to achieve the distribution of the y_T :

$$(y_T | y_L) \sim N(\mu_T, \Sigma_T) \quad (15)$$

$$\Sigma_T = k(x_T, x_T) = k(x_T, x_T) + \sigma_{\text{noise}}^2 I_n - k(x_T, x_L)(k(x_L, x_L) + \sigma_{\text{noise}}^2 I_n)^{-1} k(x_L, x_T) \quad (16)$$

$$\mu_T = m\left(\begin{matrix} \rightarrow \\ y_T \end{matrix}\right) = k(x_T, x_L)(k(x_L, x_L) + \sigma_{\text{noise}}^2 I_n)^{-1} \begin{matrix} \rightarrow \\ y_T \end{matrix} \quad (17)$$

where the Σ_T and μ_T are the covariance and the mean value, respectively. Selecting a kernel function, including a symmetric invertible matrix, can influence the strength and the robustness of the prediction ability of the final GPR model. Accordingly, four different kernel functions of Exponential, Squared exponential, Matern, and Rational quadratic are selected to find the best one. The selected kernel functions are presented as follow:

- Exponential kernel function:

$$k_E(x, x') = \sigma^2 \exp\left(-\frac{x - x'}{\downarrow}\right) \quad (18)$$

- Squared Exponential kernel function:

$$k_{SE}(x, x') = \sigma^2 \exp\left(-\frac{x - x'}{\downarrow^2}\right) \quad (19)$$

- Matern kernel function:

$$k_M(x, x') = \sigma^2 \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\sqrt{2\nu} \frac{x - x'}{\downarrow}\right)^\nu K_\nu\left(\sqrt{2\nu} \frac{x - x'}{\downarrow}\right) \quad (20)$$

- Rational quadratic kernel function:

$$k_{RQ}(x, x') = \sigma^2 \left(1 + \frac{x - x'}{2a\downarrow}\right)^{-a} \quad (21)$$

Where $\alpha > 0$, σ , σ^2 , and ℓ represent scale-mixture, the amplitude, the variance, and the length scale, respectively. Additionally, the K_ν , Γ , and ν indicate the modified Bessel function, the gamma function, and a positive parameter, respectively.

Data set outlier detection. Outlier or suspected data has different behavior compared to the rest of the data points. These data often arise from the experiment or instrumental errors. The detection of suspected data in the data set is necessary to avoid the wrong examination of the established model and make it more efficient. To do so, Leverage method was used in which the Hat matrix is defined as follows:

$$H = U(U^T U)^{-1} U^T \quad (22)$$

U is an $i \times j$ dimensional matrix, and i and j represent the parameters number and the training data number, respectively. The preciseness of the data set is assessed by plotting the standardized residuals versus Hat values, called William's plot. In this diagram, a reliable region is defined which the data out of it are suspected data. The limited area between standardized residuals of -3 to 3 and Hat values of 0 to critical leverage limit, is considered as the reliable zone. The critical leverage limit is calculated as follows^{68,69}:

$$H^* = \frac{3(j+1)}{i} \quad (23)$$

According to William's plot of the TC removal data bank (Fig. 1), most of the used data are reliable. In detail, from 374 data points, the number of outliers are only 10, 4, 6, and 10 for GPR-Matern, GPR-Exponential, GPR-Squared Exponential, and GPR-Rational quadratic models, respectively.

Results and discussion

Sensitivity analysis. Determining the effect of various operational variables on the outcome is very important for researchers and engineers to suggest an accurate model. To attain this aim, an analysis of sensitivity has been performed through which the relevancy factor of each input variable was calculated as follows⁴⁰:

$$r = \frac{\sum_{i=1}^n (X_{k,i} - \bar{X}_k)(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_{k,i} - \bar{X}_k)^2 \sum_{i=1}^n (Y_i - \bar{Y})^2}} \quad (24)$$

where $X_{k,i}$ and Y_i stand the ' k ' th input and ' i ' th output while \bar{X}_k and \bar{Y} indicate the average values of input and outputs, respectively.

Each input with a more considerable r value has a greater effect on the output quantity. The negative values denote that the corresponding parameter negatively affects the model outcome and vice versa. As depicted in Fig. 2, the illumination time is the most influential parameter in the TC degradation on various MOFs. As the time of the process increase, the degradation improves. Indeed, time is the only parameter which alters along with the degradation percentage simultaneously during the experiment.

In the case of solution pH, it affects the overall charge of the catalyst and pollutant, and consequently, the adsorption of pollutants on the surface and degradation could be influenced. Thus, it is not necessarily correct that the pollutant degrades faster at higher pH because several MOFs with various structures have different responses to pH⁵¹. For instance, Zhang et al.⁵⁷ reported that the pH 3 is the best value to degrade TC on MIL-88A, while Pan et al.⁶¹ claimed that the TC was degraded over AgI/UiO-66-NH₂ more rapidly in pH 4.3.

The surface area is the second most influential parameter with a negative effect. In this case, we cannot interpret directly without considering the following factors. First, the researchers often used MOFs composed with another compound, fabricating composites to enhance the photodegradation efficiency. In some cases, it declines the surface area and no/small change in pore sizes. In other words, the MOF pores got blocked by introducing

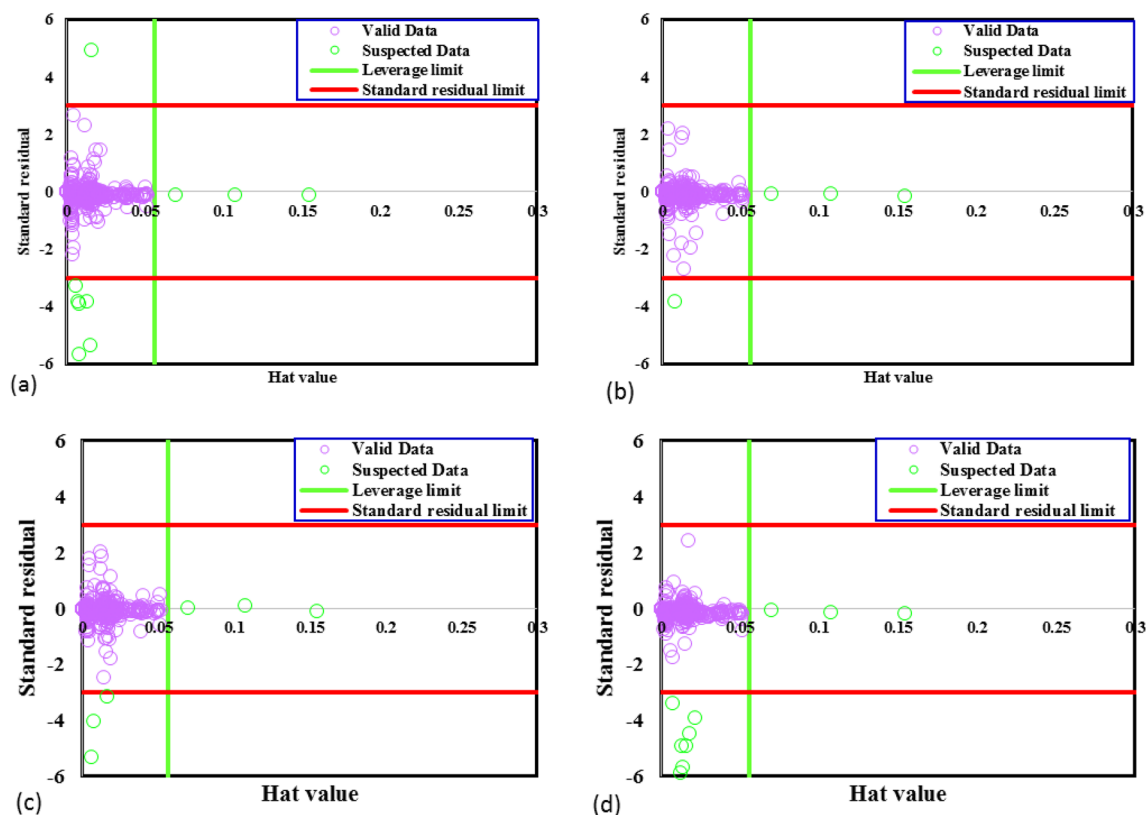


Figure 1. Detection of outliers for GPR model containing kernel function of (a) matern, (b) exponential, (c) squared exponential, (d) rational quadratic.

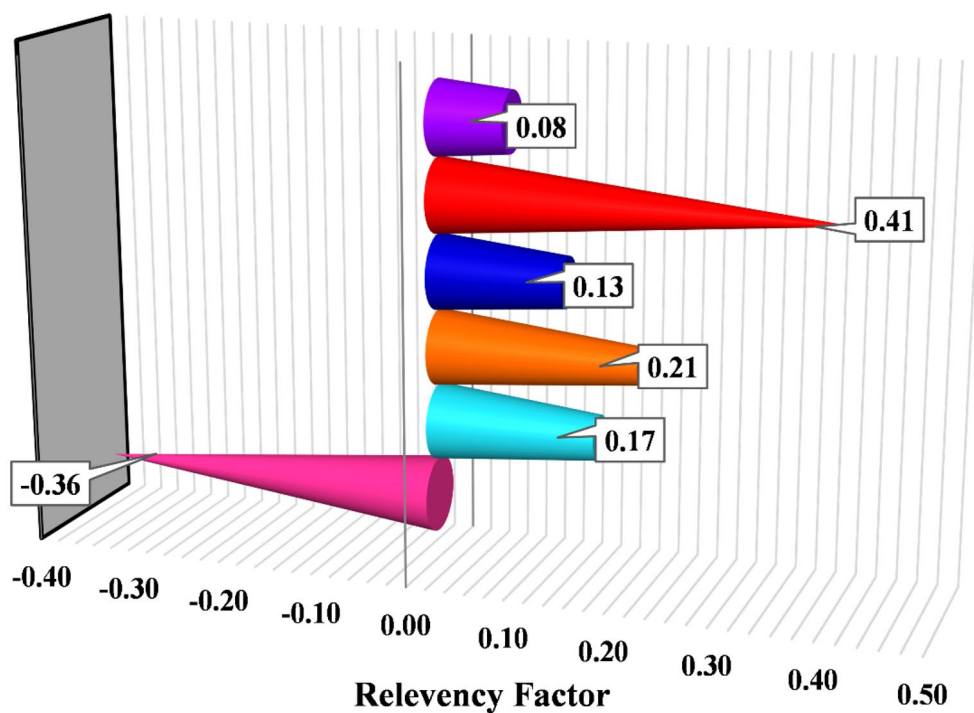
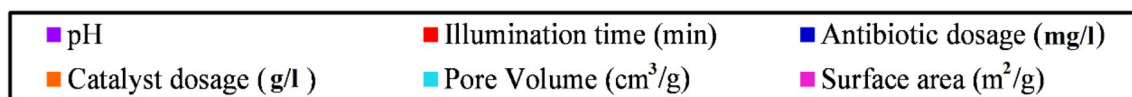


Figure 2. Sensitivity analysis of the input operational variables for TC photodegradation by various MOFs.

Model	Dataset	R ²	MRE (%)	MSE	RMSE	STD
GPR matern	Train	0.980	12.20	16.13	4.02	3.21
	Test	0.981	12.29	18.03	4.25	3.33
	Total	0.980	12.22	16.60	4.25	3.24
GPR Exp	Train	0.979	15.54	17.09	4.13	3.15
	Test	0.977	11.81	21.55	4.64	3.34
	Total	0.979	14.61	18.20	4.64	3.20
GPR SExp	Train	0.973	14.56	22.47	4.74	3.52
	Test	0.970	19.92	28.40	5.33	3.92
	Total	0.972	15.89	23.95	5.33	3.63
GPR rational quadratic	Train	0.967	19.58	27.20	5.22	3.86
	Test	0.969	28.27	28.93	5.38	4.03
	Total	0.967	21.74	27.63	5.38	3.90

Table 1. The statistical parameters of suggested GPR models.

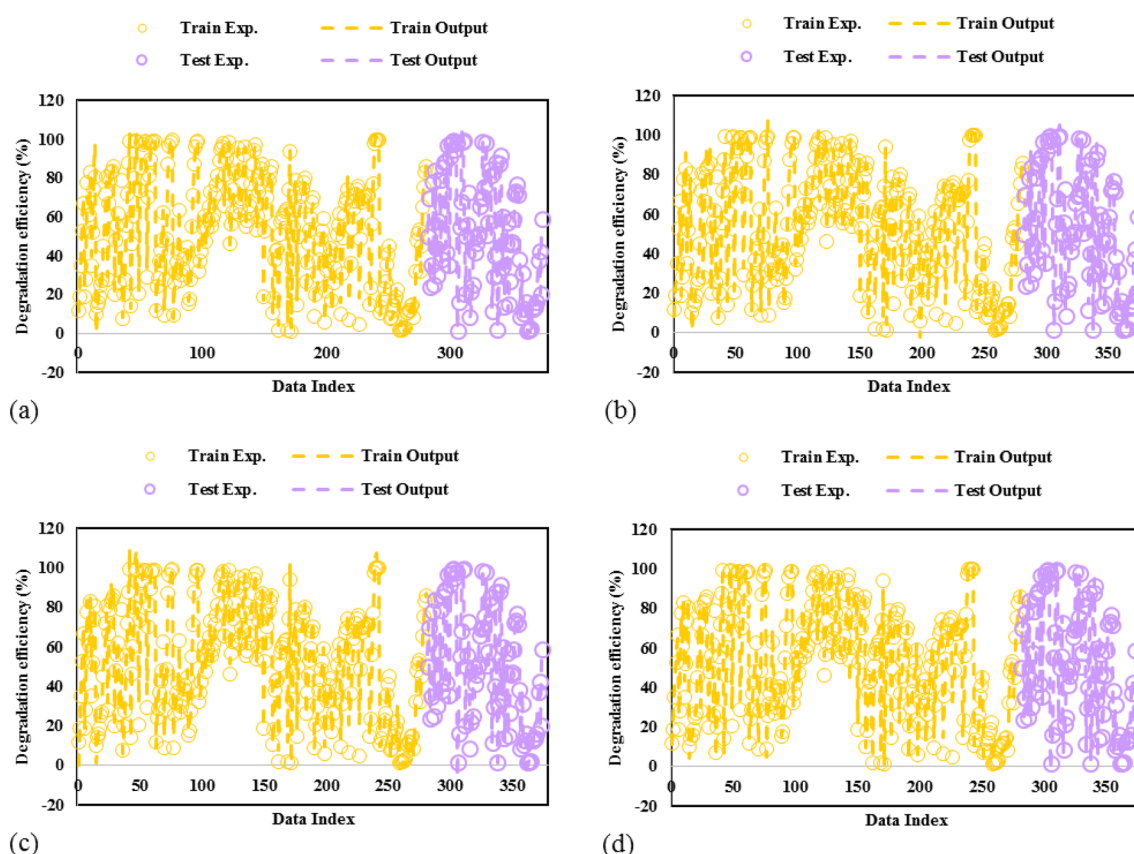


Figure 3. Comparison of actual and predicted values for GPR model with kernel function of (a) matern, (b) exponential, (c) squared exponential, (d) rational quadratic.

the second photocatalysts. For example, the 53% TC degradation by UiO-66 with 856.61 m²/g increased to 86% after composite with WO₃/graphene oxide while the surface area decreased to 379.51 m²/g⁶⁰. Secondly, about 60% of the data points (218 out of 374 data points) have a lower surface area than the mean value of the data bank. It indicates that most data points are obtained from MOFs with low surface area. Indeed, lower data points with high surface area and excellent degradation contribute to the sensitivity analysis of the surface area with the negative sign. Besides, significant changes in MOFs surface area (from 15.95 to 1548.3 m²/g) result in the severe effect of this parameter on the outcome.

Furthermore, input parameters such as catalyst dosage and antibiotic concentration have critical value to influence positively in the photocatalytic experiments. As the catalyst dosage increases, a large number of charge carriers are generated, increasing the degradation efficiency. It is until a certain amount because high amount of catalyst avoids the light from reaching the inner environment of the solution. The extra amount of TC leads to a competition between the TC and reaction-intermediates molecules for being absorbed on the catalyst surface.

Thus, the two last matters cause a decrease in the TC degradation efficiency³⁸. Considering the above discussion, the results of sensitivity analysis must interpret with enough knowledge of the under-study phenomenon. Here, some parameters of chemical features of the MOFs and the contaminant or photocatalytic properties such as the recombination rate of the photogenerated charge carriers and catalyst light absorbance were not considered while could be determining factors. Therefore, as Ayodele et al.⁵³ reported, the importance of the parameters varies based on the photocatalytic process conditions.

Modeling results and validation. In order to examine the anticipation ability of the suggested GPR models for TC photodegradation by MOFs, several assessments have been carried out and presented. The validation assessments are provided in two main approaches: statistical matching factors and graphical comparison plots. The statistical parameters indicate how much the actual and predicted TC removal values match. These parameters are listed in Table 1 for the training, testing, and total data sets. For training data, the R^2 values of 0.980, 0.979, 0.973, and 0.967 are achieved for the developed GPR models with kernel functions of Matern, Exponential, Squared Exponential, and Rational Quadratic, respectively. Also, their corresponding error values of MRE, MSE, RMSE, and STD are low, confirming that they have trained the data with adequate accuracy. In addition to the training data, the capability of the established GPR models in anticipation of the new (unseen) TC photodegradation data points is vital. Thus, the models have also been evaluated for the testing data set. As presented in Table 1, the GPR model with Matern kernel function shows the most accurate performance for the photocatalytic degradation of TC by MOFs (R^2 , MRE, MSE, RMSE, and STD are 0.981, 12.29, 18.03, 4.25, and 3.33, respectively). Note that the GPR-Exponential model also has acceptable results close to the GPR-Matern model.

The experimental and predicted values of TCs photodegradation are simultaneously depicted in Fig. 3. As shown, the predicted amounts of TC photocatalytic degradation by MOFs have excellent agreements with the actual amounts. Indeed, all the proposed models' results could follow the experimental data points in training and testing data sets. Thus, the developed GPR models present reliable anticipation performance for the TC-MOF photodegradation system.

In addition, the actual TC photodegradation efficiencies versus the GPR models' estimated ones are plotted in Fig. 4, namely, cross plot. In this plot, the bisector line in the first quarter is the prediction accuracy criterion. The closer data points to the bisector line, the more accuracy the suggested models obtain. As shown, for all the GPR models, particularly the GPR-Matern, the predicted data placed on the actual data where the fitting line of both data sets have near unit R^2 values ($R^2 = 0.98$ in the GPR-Matern model). Therefore, it is confirmed that the GPR models can predict the experimental data set in this study.

In the next step of models' results validation, the percentage of relative deviation between actual and model estimated TC removal is calculated and depicted in Fig. 5. For all the GPR models, the absolute relative deviation points are mostly lower than 30%, while for GPR with Matern kernel function, they are lower than 20%. Also, GPR-Matern gives the minimum mean relative error of 12.22% compared to GPR-Exponential (14.61%), GPR-Square exponential (15.89%), and GPR-Rational quadratic (21.74%).

Conclusion

In this study, the machine learning method of Gaussian process regression (GPR) model with four kernel functions of Exponential, Squared exponential, Matern, and Rational quadratic were investigated to anticipate the TC photodegradation from wastewater by MOFs. The MOFs features (surface area and pore volume) and operational parameters (illumination time, catalyst dosage, TC concentration, pH) were chosen as the input parameters. The established models prognosticated the experimental TC degradation very well. The GPR-Matern model shows the best performance with R^2 , MRE, MSE, RMSE, and STD of 0.981, 12.29, 18.03, 4.25, and 3.33, respectively. The analysis of sensitivity specifies that the illumination time is the most influential parameter in TC photodegradation by MOFs. The MOFs surface area was found as the second determining input in the TC-MOFs photocatalytic systems. The presented discussions in the current study could make it a supportive report for the engineers and researchers dealing with TCs contaminant elimination or control technologies.

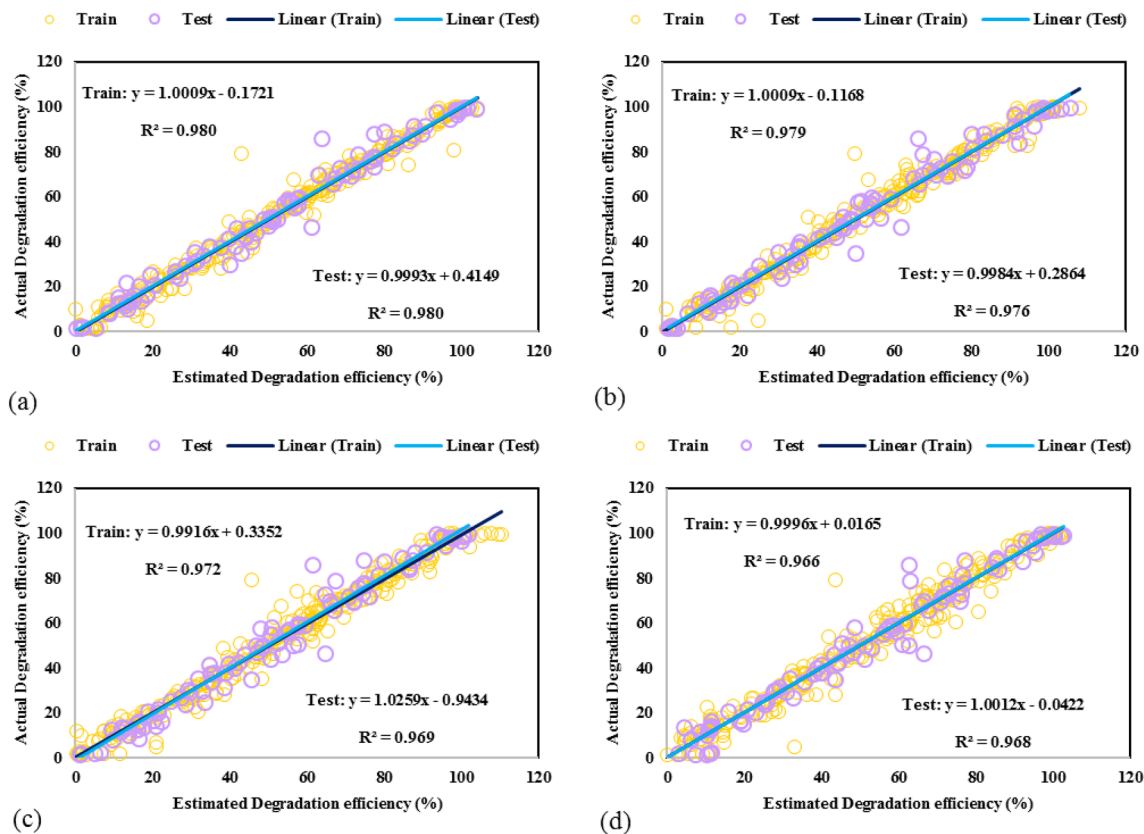


Figure 4. Cross plots for GPR model containing kernel function of (a) exponential, (b) matern, (c) squared exponential, (d) rational quadratic.

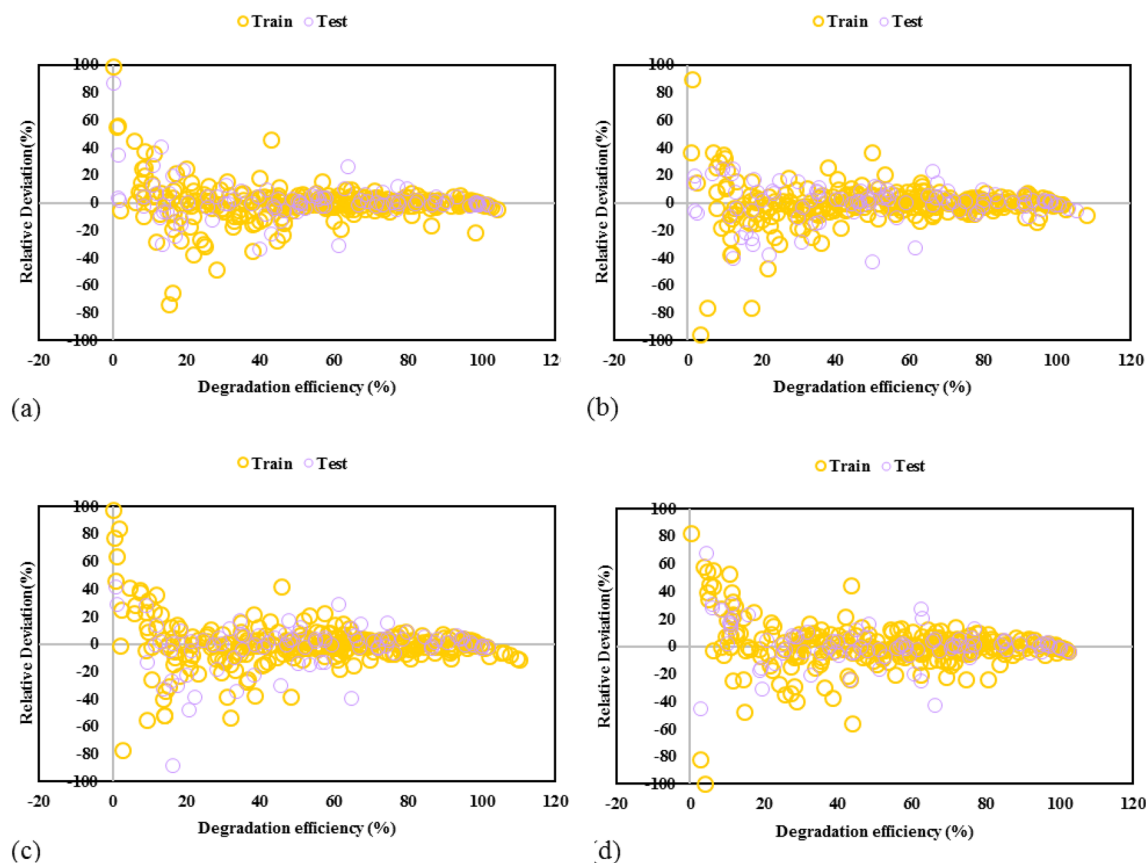


Figure 5. Comparison of actual and model predicted values for GPR model containing kernel function of (a) matern, (b) exponential, (c) squared exponential, (d) rational quadratic.

Data availability

All data generated or analysed during this study are included in this published article [and its supplementary information files].

Received: 29 January 2022; Accepted: 29 March 2022

Published online: 22 April 2022

References

- Chowdhary, P., Raj, A. & Bharagava, R. N. Environmental pollution and health hazards from distillery wastewater and treatment approaches to combat the environmental threats: A review. *Chemosphere* **194**, 229–246 (2018).
- Cheng, D. *et al.* A critical review on antibiotics and hormones in swine wastewater: Water pollution problems and control approaches. *J. Hazard. Mater.* **387** (2020).
- Lellis, B., Fávoro-Polonio, C. Z., Pamphile, J. A. & Polonio, J. C. Effects of textile dyes on health and the environment and bioremediation potential of living organisms. *Biotechnol. Res. Innov.* **3**, 275–290 (2019).
- Charuau, L., Jarde, E., Jaffrezic, A., Thomas, M. F. & Le Bot, B. Veterinary pharmaceutical residues from natural water to tap water: Sales, occurrence and fate. *J. Hazard. Mater.* **361**, 169–186 (2019).
- Wang, J., Zhi, D., Zhou, H., He, X. & Zhang, D. Evaluating tetracycline degradation pathway and intermediate toxicity during the electrochemical oxidation over a Ti/TiO₂ anode. *Water Res.* **137**, 324–334 (2018).
- Hong, Y. *et al.* Promoting visible-light-induced photocatalytic degradation of tetracycline by an efficient and stable beta-Bi₂O₃@g-C₃N₄ core/shell nanocomposite. *Chem. Eng. J.* **338**, 137–146 (2018).
- Huang, X. *et al.* Enhanced heterogeneous photo-Fenton catalytic degradation of tetracycline over γ -CeO₂/Fh composites: Performance, degradation pathways, Fe²⁺ regeneration and mechanism. *Chem. Eng. J.* **392**, 123636 (2020).
- Ao, X. *et al.* Degradation of tetracycline by medium pressure UV-activated peroxymonosulfate process: Influencing factors, degradation pathways, and toxicity evaluation. *Chem. Eng. J.* **361**, 1053–1062 (2019).
- Yuan, B. *et al.* Nanocellulose-based composite materials for wastewater treatment and waste-oil remediation. *ES Food Agrofor.* <https://doi.org/10.30919/esfaf0004> (2020).
- Peng, H. *et al.* Two-dimension N-doped nanoporous carbon from KCl thermal exfoliation of Zn-ZIF-L: Efficient adsorption for tetracycline and optimizing of response surface model. *J. Hazard. Mater.* **402**, 123498 (2021).
- Khan, M. H., Bae, H. & Jung, J. Y. Tetracycline degradation by ozonation in the aqueous phase: Proposed degradation intermediates and pathway. *J. Hazard. Mater.* **181**, 659–665 (2010).
- Lu, C. *et al.* Preparation, characterization of Fe ion exchange modified titanate nanotubes and photocatalytic activity for oxytetracycline. *Fresenius Environ. Bull.* **24**, 2348–2353 (2015).
- Elazzouzi, M., Haboubi, K., Elyoubi, M. S. & Kasmí, A. E. A developed low-cost electrocoagulation process for efficient phosphate and COD removals from real urban wastewater. *ES Energy Environ.* <https://doi.org/10.30919/esee8c302> (2019).

14. Guo, X. *et al.* Potential of Myriophyllum aquaticum for phytoremediation of water contaminated with tetracycline antibiotics and copper. *J. Environ. Manag.* **270**, 110867 (2020).
15. Ocampo-Pérez, R., Rivera-Utrilla, J., Gómez-Pacheco, C., Sánchez-Polo, M. & López-Peñalver, J. J. Kinetic study of tetracycline adsorption on sludge-derived adsorbents in aqueous phase. *Chem. Eng. J.* **213**, 88–96 (2012).
16. Yang, K. *et al.* Microbial diversity in combined UAF-UBAF system with novel sludge and coal cinder ceramic fillers for tetracycline wastewater treatment. *Chem. Eng. J.* **285**, 319–330 (2016).
17. Wang, P., He, Y. L. & Huang, C. H. Reactions of tetracycline antibiotics with chlorine dioxide and free chlorine. *Water Res.* **45**, 1838–1846 (2011).
18. Yuan, X. *et al.* In-situ synthesis of 3D microsphere-like In₂S₃/InVO₄ heterojunction with efficient photocatalytic activity for tetracycline degradation under visible light irradiation. *Chem. Eng. J.* **356**, 371–381 (2019).
19. Wang, W., Fang, J., Shao, S., Lai, M. & Lu, C. Compact and uniform TiO₂@g-C₃N₄ core-shell quantum heterojunction for photocatalytic degradation of tetracycline antibiotics. *Appl. Catal. B Environ.* **217**, 57–64 (2017).
20. Jiang, L. *et al.* Nitrogen self-doped g-C₃N₄ nanosheets with tunable band structures for enhanced photocatalytic tetracycline degradation. *J. Colloid Interface Sci.* **536**, 17–29 (2019).
21. Gong, Y. *et al.* MOF-derived nitrogen doped carbon modified g-C₃N₄ heterostructure composite with enhanced photocatalytic activity for bisphenol A degradation with peroxymonosulfate under visible light irradiation. *Appl. Catal. B Environ.* **233**, 35–45 (2018).
22. Sculley, J., Yuan, D. & Zhou, H. C. The current status of hydrogen storage in metal-organic frameworks - Updated. *Energy Environ. Sci.* **4**, 2721–2735 (2011).
23. Lv, Y. *et al.* Core/shell template-derived Co, N-doped carbon bifunctional electrocatalysts for rechargeable Zn-air battery. *Eng. Sci.* **7**, 26–37 (2019).
24. Shahmirzaee, M., Hemmati-Sarapardeh, A., Husein, M. M., Schaffie, M. & Ranjbar, M. Development of a powerful zeolitic imidazolate framework (ZIF-8)/carbon fiber nanocomposite for separation of hydrocarbons and crude oil from wastewater. *Microporous Mesoporous Mater.* **307**, 110463 (2020).
25. Wang, Y. *et al.* Significantly enhanced ultrathin NiCo-based MOF nanosheet electrodes hybridized with Ti₃C₂T_x MXene for high performance asymmetric supercapacitors. *Eng. Sci.* **9**, 50–59 (2020).
26. Abdi, J., Vossoughi, M., Mahmoodi, N. M. & Alemzadeh, I. Synthesis of amine-modified zeolitic imidazolate framework-8, ultrasound-assisted dye removal and modeling. *Ultrason. Sonochem.* **39**, 550–564 (2017).
27. Zhang, Q. *et al.* A novel Fe-based bi-MOFs material for photocatalytic degradation of tetracycline: Performance, mechanism and toxicity assessment. *J. Water Process Eng.* **44**, 102364 (2021).
28. Wu, Z. *et al.* Photocatalytic decontamination of wastewater containing organic dyes by metal-organic frameworks and their derivatives. *ChemCatChem* **9**, 41–64 (2017).
29. Wu, J. *et al.* Bimetallic silver/bismuth-MOFs derived strategy for Ag/AgCl/BiOCl composite with extraordinary visible light-driven photocatalytic activity towards tetracycline. *J. Alloys Compd.* **877**, 160262 (2021).
30. Yang, M. *et al.* Heterometallic Mg@Fe-MIL-101/TpPa-1-COF grown on stainless steel mesh: Enhancing photo-degradation, fluorescent detection and toxicity assessment for tetracycline hydrochloride. *Colloids Surf. A Physicochem. Eng. Asp.* **631**, 127725 (2021).
31. Zhu, C. *et al.* Influence of operational parameters on photocatalytic decolorization of a cationic azo dye under visible-light in aqueous Ag₃PO₄. *Inorg. Chem. Commun.* **115**, 107850 (2020).
32. Ye, Y., Feng, Y., Bruning, H., Yntema, D. & Rijnaarts, H. H. M. Photocatalytic degradation of metoprolol by TiO₂ nanotube arrays and UV-LED: Effects of catalyst properties, operational parameters, commonly present water constituents, and photo-induced reactive species. *Appl. Catal. B Environ.* **220**, 171–181 (2018).
33. Anju Chantu, L., Joychandra Singh, W., Jugeshwar Singh, K. & Nomita Devi, K. Effect of operational parameters on the photocatalytic degradation of Methylene blue dye solution using manganese doped ZnO nanoparticles. *Results Phys.* **12**, 1230–1237 (2019).
34. Gaeta, M. *et al.* Photodegradation of antibiotics by noncovalent porphyrin-functionalized tio₂ in water for the bacterial antibiotic resistance risk management. *Int. J. Mol. Sci.* **21**, 3775 (2020).
35. Vasseghian, Y. & Dragoi, E.-N. Modeling and optimization of acid Blue 193 removal by UV and peroxydisulfate process. *J. Environ. Eng.* **144**, 6018003 (2018).
36. Zilabi, S., Habibzadeh, S., Gheytnazadeh, M. & Rahmani, M. Direct sunlight catalytic decomposition of organic pollutants via Sm- and Ce-doped BiFeO₃ nanopowder synthesized by a rapid combustion technique. *Catal. Lett.* **151**, 3462–3476 (2021).
37. Reddy, B. S. *et al.* Knowledge extraction of sonophotocatalytic treatment for acid blue 113 dye removal by artificial neural networks. *Environ. Res.* **204**, 341 (2022).
38. Abdi, J., Hadipoor, M., Hadavimoghaddam, F. & Hemmati-Sarapardeh, A. Estimation of tetracycline antibiotic photodegradation from wastewater by heterogeneous metal-organic frameworks photocatalysts. *Chemosphere* **287**, 132135 (2022).
39. Tabatabai-Yazdi, F. S., EbrahimianPirbazari, A., Esmaeili Khalil Saraei, F. & Gilani, N. Construction of graphene based photocatalysts for photocatalytic degradation of organic pollutant and modeling using artificial intelligence techniques. *Phys. B Condens. Matter* **608**, 412869 (2021).
40. Gheytnazadeh, M. *et al.* Towards estimation of CO₂ adsorption on highly porous MOF-based adsorbents using gaussian process regression approach. *Sci. Rep.* **11**, 1–13 (2021).
41. Baghban, A. *et al.* Towards experimental and modeling study of heat transfer performance of water- SiO₂ nanofluid in quadrangular cross-section channels. *Eng. Appl. Comput. Fluid Mech.* **13**, 453–469 (2019).
42. Liu, Z. & Baghban, A. Application of LSSVM for biodiesel production using supercritical ethanol solvent. *Energy Sour. Part A Recover. Util. Environ. Eff.* **39**, 1869–1874 (2017).
43. Baghban, A., Bahadori, M., Lemraski, A. S. & Bahadori, A. Prediction of solubility of ammonia in liquid electrolytes using least square support vector machines. *Ain Shams Eng. J.* **9**, 1303–1312 (2018).
44. Bahadori, A. *et al.* Computational intelligent strategies to predict energy conservation benefits in excess air controlled gas-fired systems. *Appl. Therm. Eng.* **102**, 432–446 (2016).
45. Bemani, A., Baghban, A., Mohammadi, A. H. & Andersen, P. Ø. Estimation of adsorption capacity of CO₂, CH₄, and their binary mixtures in Quidam shale using LSSVM: Application in CO₂ enhanced shale gas recovery and CO₂ storage. *J. Nat. Gas Sci. Eng.* **76**, 103204 (2020).
46. Zamen, M., Baghban, A., Pourkiaei, S. M. & Ahmadi, M. H. Optimization methods using artificial intelligence algorithms to estimate thermal efficiency of PV/T system. *Energy Sci. Eng.* **7**, 821–834 (2019).
47. Baghban, A. Application of the ANFIS strategy to estimate vaporization enthalpies of petroleum fractions and pure hydrocarbons. *Pet. Sci. Technol.* **34**, 1359–1366 (2016).
48. Baghban, A., Abbasi, P. & Rostami, P. Modeling of viscosity for mixtures of Athabasca bitumen and heavy n-alkane with LSSVM algorithm. *Pet. Sci. Technol.* **34**, 1698–1704 (2016).
49. Ahmadi, M. H. *et al.* An insight into the prediction of TiO₂/water nanofluid viscosity through intelligence schemes. *J. Therm. Anal. Calorim.* **139**, 2381–2394 (2020).
50. Shang, Q. *et al.* Optimization of Bi₂O₃/TS-1 preparation and photocatalytic reaction conditions for low concentration erythromycin wastewater treatment based on artificial neural network. *Process Saf. Environ. Prot.* **157**, 297–305 (2022).

51. Jiang, Z. *et al.* A generalized predictive model for TiO₂—Catalyzed photo-degradation rate constants of water contaminants through artificial neural network. *Environ. Res.* **187**, 109697 (2020).
52. Mahmoodi, N. M. *et al.* Activated carbon/metal-organic framework nanocomposite: Preparation and photocatalytic dye degradation mathematical modeling from wastewater by least squares support vector machine. *J. Environ. Manag.* **233**, 660–672 (2019).
53. Ayodele, B. V., Alsaffar, M. A., Mustapa, S. I., Cheng, C. K. & Witoon, T. Modeling the effect of process parameters on the photocatalytic degradation of organic pollutants using artificial neural networks. *Process Saf. Environ. Prot.* **145**, 120–132 (2021).
54. Wang, H. *et al.* In situ synthesis of In₂S₃ at MIL-125(Ti) core-shell microparticle for the removal of tetracycline from wastewater by integrated adsorption and visible-light-driven photocatalysis. *Appl. Catal. B Environ.* **186**, 19–29 (2016).
55. Wang, D. *et al.* Simultaneously efficient adsorption and photocatalytic degradation of tetracycline by Fe-based MOFs. *J. Colloid Interface Sci.* **519**, 273–284 (2018).
56. Wang, J. *et al.* Fabrication of perylene imide-modified NH₂-UiO-66 for enhanced visible-light photocatalytic degradation of tetracycline. *J. Photochem. Photobiol. A Chem.* **401**, 112795 (2020).
57. Zhang, Y., Zhou, J., Chen, X., Wang, L. & Cai, W. Coupling of heterogeneous advanced oxidation processes and photocatalysis in efficient degradation of tetracycline hydrochloride by Fe-based MOFs: Synergistic effect and degradation pathway. *Chem. Eng. J.* **369**, 745–757 (2019).
58. Chen, D. D. *et al.* Polyaniline modified MIL-100(Fe) for enhanced photocatalytic Cr(VI) reduction and tetracycline degradation under white light. *Chemosphere* **245**, 40008 (2020).
59. Deng, L. *et al.* The facile boosting sunlight-driven photocatalytic performance of a metal-organic-framework through coupling with Ag₂S nanoparticles. *New J. Chem.* **44**, 12568–12578 (2020).
60. Fakhri, H. & Bagheri, H. Highly efficient Zr-MOF@WO₃/graphene oxide photocatalyst: Synthesis, characterization and photo-degradation of tetracycline and malathion. *Mater. Sci. Semicond. Process.* **107**, 104815 (2020).
61. Pan, Y. *et al.* Stable self-assembly AgI/UiO-66(NH₂) heterojunction as efficient visible-light responsive photocatalyst for tetracycline degradation and mechanism insight. *Chem. Eng. J.* **384**, 12330 (2020).
62. Wu, Q., Yang, H., Kang, L., Gao, Z. & Ren, F. Fe-based metal-organic frameworks as Fenton-like catalysts for highly efficient degradation of tetracycline hydrochloride over a wide pH range: Acceleration of Fe(II)/Fe(III) cycle under visible light irradiation. *Appl. Catal. B Environ.* **263**, 118282 (2020).
63. Zhao, S. *et al.* Microwave-assisted hydrothermal assembly of 2D copper-porphyrin metal-organic frameworks for the removal of dyes and antibiotics from water. *Environ. Sci. Pollut. Res.* **27**, 39186–39197 (2020).
64. Hoang, N. D., Pham, A. D., Nguyen, Q. L. & Pham, Q. N. Estimating compressive strength of high performance concrete with Gaussian process regression model. *Adv. Civ. Eng.* **2016**, 16 (2016).
65. Rasmussen, C. E. Gaussian Processes in machine learning. *Lect. Notes Comput. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinform.) Sci.* **3176**, 63–71 (2004).
66. CE, W. C. K. I. R. Adaptive computation and machine learning series. *Gaussian Process. Mach. Learn.* **11**, 1–272 (2005).
67. Fu, Q. *et al.* Prediction of the diet nutrients digestibility of dairy cows using Gaussian process regression. *Inf. Process. Agric.* **6**, 396–406 (2019).
68. Zhou, X., Zhou, F. & Naseri, M. An insight into the estimation of frost thermal conductivity on parallel surface channels using kernel based GPR strategy. *Sci. Rep.* **11**, 1–11 (2021).
69. Razavi, R., Bemani, A., Baghban, A., Mohammadi, A. H. & Habibzadeh, S. An insight into the estimation of fatty acid methyl ester based biodiesel properties using a LSSVM model. *Fuel* **243**, 133–141 (2019).

Author contributions

All authors contributed in writing, simulation, data collection, and conceptions.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-10563-8>.

Correspondence and requests for materials should be addressed to A.B. or S.H.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022