

Imperial College London
Department of Earth Science and Engineering
MSc in Environmental Data Science and Machine Learning

Independent Research Project

Software Development for Adsorption Kinetic Data Analysis

by
Peiqi Xia

Email: peiqi.xia22@imperial.ac.uk

GitHub username: edsml-px122

Repository: <https://github.com/ese-msc-2022/irp-px122>

Supervisors:
Prof. Dominik Weiss
Dr. Jay Bullen

August 2023

Contents

1. Abstract	1
2. Introduction	1
2.1. Background	1
2.2. Literature Review	2
2.3. Aims and Objectives	3
3. Methodology	3
3.1. Data collection	3
3.2. Mathematical adsorption kinetic models	3
3.2.1. Zero-order reaction	3
3.2.2. First-order reaction	4
3.2.3. Second-order reaction	4
3.3. Reaction order analysis	4
3.4. Goodness of fit	5
3.5. Error calculation	5
3.6. Model visualisation	6
3.7. UI design and workflow	6
4. Code Metadata	6
4.1. Information	6
4.2. Code structure	7
4.3. Usage instruction	7
5. Results and Discussion	7
5.1. Software performance	7
5.2. Function validation	9
5.3. Discussion of adsorbate	12
6. Conclusion and Future Work	13
7. Acknowledgement	14
8. Reference	14
9. Appendix	16

1 Abstract

Water contamination is a significant global concern due to its serious implications on public health and the environment. Nowadays, due to the rapid increase in population, the need for safe drinking water also increased urgently. Scientists and environmentalists are working hard to develop effective water treatment strategies to remove the various contaminants from water on a large scale. Currently, they are spending much time on calculating the adsorption kinetics of the experiments in spreadsheets. In this project, a software is developed in Python aiming to help them analyse the kinetic data of adsorption more easily which is critical to the engineering process design. The models built in the software include zero-order, first-order (PFO), second-order (PSO) and the recently proposed revised PSO (rPSO). It's can both calculate single experiment kinetics by fitting the data into all the models directly and analyse multiple datasets which determine the reaction order with respect to adsorbate hence fit the particular model. This software is tested on 13 datasets with 65 experiments and gives an consistency of 68.4% by comparing parameters with that can be found in the literatures where the difference is within 10% of the literature values on a natural logarithm scale. The model performance is evaluated by coefficient of determination (R^2) values to find out that the nonlinear models always give better results than the linear models. Moreover, the accuracy of results from most of the calculations really depends on the quality of dataset. Some techniques can be adjusted to improve the reliability of this software in the future.

2 Introduction

2.1 Background

Drinking water resources always contain natural contaminants, but in today's world, the rapid urbanization and industrialization results in many possible man-made sources of contaminants in water.¹ These contamination has impacts on human health, quality of the environment such as lands and forests, and economic sustainability.²

There are mainly four types of water pollution which are inorganic, organic, biological and radiological contaminants.³ Metals including zinc, mercury, chromium, etc. and metalloids like arsenic are some of the most common chemical pollutants. Among these, the arsenic(III) removal from drinking water⁴ is one of the most concerned problem since it is highly toxic and can cause arsenicosis. Arsenic contamination has harmful effects on human including skin lesions, cardiovascular diseases, respiratory disorders, and risks of various cancers⁵ which threatens more than 150 million people around the world, especially in South and South-East-Asia (Bangladesh, China, India and so on).^{6,7} Another kind of compounds which is more often studied recently is the per- and polyfluorinated alkyl substances (PFASs). It has been used in industries since the mid-20th century. However, compounds like PFOS and PFOA were found to be carcinogenic and immunotoxic more than thirty years ago.⁸ Many PFASs can get into the environment through leaching into the soil to reach the groundwater or evaporating into the atmosphere.

As the world's population is growing rapidly, the need of access to safe drinking water is also increasing fast. Thus, scientists are working on developing the water treatment techniques for different contaminants that is time efficient and cost effective. Common methods include ion exchange, adsorption, phytoremediation, chemical precipitation, etc. where adsorption is one of the most widely used for many decades.⁹ The mechanism involves the attachment of molecules from liquid solutions onto the surface of a solid known as adsorbent. For example, $\text{TiO}_2\text{-Fe}_2\text{O}_3$ as an adsorbent for arsenic(III)⁴ removal and activated carbon as an adsorbent for PFAS¹⁰ removal are systems that quite often reported in the studies. Compared with other

methods, adsorption needs lower cost for materials and can provide higher efficiency of contaminant removal.

2.2 Literature Review

By doing batch experiments in laboratory, the adsorption study can be done from two aspects: reaction equilibrium which indicates the capacity of adsorbent and kinetics¹¹ which looks into the factors affecting the reaction rates. They are calculated by the adsorption kinetic models. The model needs to be sensitive to the reaction conditions in order to provide more applicable predictions for saving time and energy cost in engineering.¹² There are many adsorption kinetic models discovered in the literatures and pseudo-first-order (PFO) and pseudo-second-order (PSO) are the most common ones among them.

The PFO model, also known as the Lagergren model¹³, describes the adsorption progress by the first order mechanism which can be expressed as:

$$\frac{dq_t}{dt} = k_1(q_e - q_t) \quad (1)$$

where t is the time, q_e and q_t are the amount of adsorbate adsorbed at equilibrium and time t , and k_1 is rate constant of the PFO model.

Similarly, the PSO model proposed by Ho and McKay¹⁴ to show second order mechanism has the following expression:

$$\frac{dq_t}{dt} = k_2(q_e - q_t)^2 \quad (2)$$

where k_2 is the pseudo-second order kinetic rate constant.

Apart from these two, zero-order kinetics that requires easier mathematics can be used to simulate the kinetic data collected within short periods.¹⁵ The rate equation is not associated with the concentration of adsorbate, which has the form:

$$\frac{dq_t}{dt} = k_0(q_e - q_t)^0 = k_0 \quad (3)$$

where k_0 is the rate constant of zero-order model.

The concentration of adsorbate remaining at time t is given by the equation:

$$C_t = C_0 - C_s q_t \quad (4)$$

where C_0 is the initial concentration of adsorbate, C_s is the initial concentration of adsorbent.

The revised PSO model developed by Bullen¹² fixed some limitations of PSO model, predicting the adsorption kinetics changes with C_0 and C_s . The modified rate equation takes the form:

$$\frac{dq_t}{dt} = k' C_t \left(1 - \frac{q_t}{q_e}\right)^2 \quad (5)$$

where $k' = k_2 q_e^2 / C_0$ is the modified rate constant of this model, and C_t can be substituted by $C_0 - C_s q_t$ in equation 4.

Normally, the researchers use the integrated forms of these rate equations to determine the parameter k and q_e values by some sort of linear data fitting.

2.3 Aims and Objectives

At present, the data obtained from experiments is processed manually using excel, origin and so on which is time-consuming and lack of accuracy. The development of this software can provide a lot of benefits including efficiency, accuracy, automation and reduction of manual workload. This program aims to build different adsorption kinetic models and deal with all the calculations in python, and then display the results to people. An user interface is generated to make it easier employed by the engineers and scientists.

After the code construction, several datasets of metal adsorbates collected from the literature is tested to evaluate the model performance and reliability. Moreover, the software is applied in the PFAS removal case study, which is an interested part of the water treatment for our research group.

3 Methodology

3.1 Data Collection

The experimental data used to test this software are listed below, where 4 datasets have been studied by Bullen and the others are newly recorded from graphs in the literature.

Table 1. 13 sets of data from 9 literature resources

Dataset	System	Ref.	Dataset	System	Ref.
1	As(V) / Fe ₂ O ₃	17	8	PFOS / soil in 6 areas	22
2	As(V) / SiO ₂ +Al ₂ O ₃	17	9	9 PFAS / PAC (10mg/L)	23
3	As(V) / laterite	18	10	9 PFAS / PAC (30mg/L)	23
4	As(III) / HFO	19	11	PFOS / activated carbon	24
5	HPO ₄ ²⁻ / iron hydroxide	20	12	PFOS / zeolite	24
6	Cu(II) / PHC	21	13	PFOS / fluoro-sorb adsorbent	25
7	Cu(II) / GAC	21			

Datasets 1 to 7 are experiments which the initial concentration of adsorbent is unchanged for the investigation of the reaction order with respect to adsorbate. The PFAS datasets are more variable including changes in reactants (datasets 8 to 10), temperature (datasets 11 and 12) and initial concentration of adsorbent (dataset 13). In total, there are 65 batch experiments involved in these datasets.

3.2 Mathematical adsorption kinetic models

There are six models implemented in this software, which cover zero-order, first-order and second-order adsorption reactions. Except for fitting the linearised model, the non-linear fitting allows more flexible modelling and enables better capturing of underlying kinetics. Each model is fitted using the built-in function, either *LinearRegression()* for linear models or *curve_fit()* for non-linear models. The *LinearRegression()* is useful when finding the best-fitting linear line that minimizes the squared differences. However, the *curve_fit()* function requires an initial guess for the parameters and then adjust the parameters iteratively using optimisation techniques until the best-fitting curve is found. After determining the model parameters k and q_e , the q_t model data is calculated and compared with the experimental data by certain validation method which will be discussed afterwards.

3.2.1 Zero-order reaction

The zero-order rate equation is integrated for boundary conditions ($q_t = 0$ at $t = 0$ and $q_t = q_t$ at $t = t$) which define the range of definite integral to give:

$$q_t = k_0 t \quad (6)$$

which can be regarded as a linear form of $y = mx$, where $m = k_0$. Hence, the q_t and t data is fitted to the linear ZO model to calculate k_0 . (q_e is not determined for this model)

3.2.2 First-order reaction

The pseudo-first-order (PFO) rate equation can be integrated to have the form:

$$q_t = q_e(1 - e^{-k_1 t}) \quad (7)$$

which is linearized as:

$$\ln(q_e - q_t) = -k_1 t + \ln q_e \quad (8)$$

The linear PFO model fits $\ln(q_e - q_t)$ and t as $y = mx + c$, where $m = -k_1$ and $c = \ln q_e$. The value of q_e for computing $\ln(q_e - q_t)$ is temporarily determined as the maximum value of q_t data multiplied by 1.01 which adds 1% amount to it assuming the reaction progress is very close to the equilibrium.

The nonlinear PFO model will take the result parameters of linear PFO model as initial guess to fit with equation 7 directly.

3.2.3 Second-order reaction

The pseudo-second-order (PSO) rate equation after integration takes the form:

$$q_t = \frac{q_e^2 k_2 t}{1 + q_e k_2 t} \quad (9)$$

and its linearized form:

$$\frac{t}{q_t} = \frac{1}{k_2 q_e^2} + \frac{t}{q_e} \quad (10)$$

Thus, the linear PSO model fits t/q_t and t as $y = mx + c$, where $m = 1/q_e$ and $c = 1/k_2 q_e^2$. The data point at $t = 0$ is excluded to avoid error in the computation of t/q_t .

Next, similar to the PFO model, the calculated parameters k_2 and q_e are used as the initial guess for nonlinear PSO model calculation (fitting with equation 9).

The revised PSO model is implemented taking the results of nonlinear PSO model as an initial guess which $k' = k_2 q_e^2 / C_0^2$. Since the rPSO rate equation is hard to integrate, the amount of adsorbate absorbed at the n th data point is calculated by:

$$q_n = q_{n-1} + (t_n - t_{n-1})(k' C_{t(n-1)}(1 - \frac{q_{n-1}}{q_e})^2) \quad (11)$$

where $\Delta t = t_n - t_{n-1}$ is determined as the value of last time data point divided by 1000 and $C_{t(n-1)} = C_0 - C_s q_{n-1}$.

3.3 Reaction order analysis

This program allows people to analyse the order of reaction with respect to the adsorbate from multiple datasets of an adsorption system using different initial concentrations of adsorbate. Keeping the concentration of adsorbent the same, the rate equation of the reaction can be written as:

$$rate = k[adsorbate]^n \quad (12)$$

where k is the rate constant and n is the reaction order with respect to the adsorbate. Taking the natural logarithm of both sides gives:

$$\ln(rate) = \ln(k) + n \ln([adsorbate]) \quad (13)$$

which has the form of $y = mx + c$. This means that $\ln(rate)$ is proportional to $\ln([adsorbate])$, and n is the slope of its linear-regression line.

Since the reaction kinetic model has not been determined before this step, the value of rate constant cannot be calculated out. Thus, the initial rate of each experiment for this analysis is simply calculated from the slope between the origin (0,0) and first available data when $t > 0$.

3.4 Goodness of fit

In this program, the validation method to assess the goodness of fit between the fitted model and the observed data is the coefficient of determination (R^2). It has the expression:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (14)$$

where y is the experimental data, \hat{y} is the calculated model data and \bar{y} is the mean value of experimental data. Typically, the value of R^2 ranges from 0 to 1, where a higher value indicates a better fit (smaller sum of squared residuals). It tells how well the model's predictions align with the actual data.

For the calculation of one dataset, the one which gives the highest R^2 value among all the six models will be selected as the best model to simulate the adsorption system. As for multiple datasets, although the determination of model does not depend on this, the value of R^2 is still calculated for each dataset to provide people an evaluation of the chosen kinetic model.

3.5 Error calculation

The uncertainties of model parameters (k and q_e) are calculated by the Monte Carlo approach. By adding some variations to the calculated q_t model data and simulating with the same kinetic model for multiple times, the distribution of new parameter values is analysed to determine the errors. This shows the impact of variation in experimental data on the adsorption kinetic parameters.

Firstly, the sum of squared error (SSE) and the standard deviation (std) of q_t is calculated as equations below:

$$SSE = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (15)$$

$$std = \sqrt{\frac{SSE}{n-2}} \quad (16)$$

where $(n - 2)$ corresponds to the degree of freedom as the number of optimised parameters is 2 (k and q_e). Then for each simulation, random values generated from a normal distribution with the calculated standard deviation are added to the model's predicted q_t data. Through a number of 100 simulations and a confidence level of 90%, the error associated with each parameter is determined as half of this interval. Since the value of uncertainty does not affect

the model fitting results much, the number of simulations is chosen to shorten the software's execution time.

3.6 Model visualisation

To visualise the fitting results, plots are generated using the 'matplotlib' package. According to different calculation modes, the information shown on the graph varies.

For single dataset, the result image contains 2×3 subplots of six fitted models, which shows the amount of absorbate absorbed at time t verses time. All the plots have scattered points representing the experimental data and smooth lines showing the model. If multiple datasets are given, the image will only contain 1×2 subplots. The first one of them shows the natural logarithm of initial rate verses natural logarithm of initial absorbate concentration where its slope represents the order of reaction. The second subplot is similar as previous ones but contains all the datasets fitted to one chosen model based on the reaction order analysis. When the order of reaction calculated is greater than two which is beyond the software's availability, only experimental data will be shown on the graph.

3.7 UI design and workflow

The interface of this software was designed to provide scientists convenience to use the tool, which hides the coding back-end. There are input fields for entering the experimental data (t , q_t , C_0 and C_s) and output fields for showing the calculation results (both text and graph). When the 'start' button is pressed, the software will first check whether the input data matches in size, and then count the number of datasets to determine the calculation mode. The workflow of this software is shown in the figure 1 below.

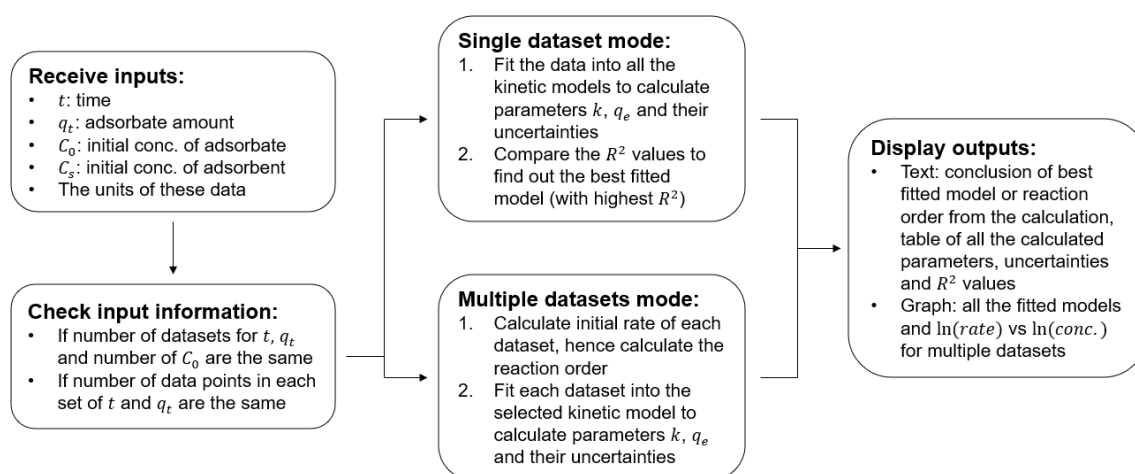


Figure 1. Workflow of the software

The text output can be saved as a txt file through the 'save' button and the generated plot will be saved automatically, both in the 'result' folder. Moreover, the 'clean' button allows the user to quickly clean up the interface for the next calculation.

4 Code Metadata

4.1 Information

- System: Windows 11
- Language: Python 3.9

- IDE: Visual Studio Code, Jupyter Notebook
- Dependencies:
 - Numpy 1.23
 - Matplotlib 3.4
 - Scipy 1.9
 - Scikit-learn 1.1
 - Pyqt5 5.15

4.2 Code structure

All the useful codes generated are in the 'tool' folder which are four files listed below. There is also an empty 'result' folder ready for saving text and image results.

- base.py – the design of the main window which is generated from the ui file created by Qt5 Designer
- funcs.py – the definitions of all the models, calculation methods and plotting functions
- main.py – the connection between the user interface and different functions, including receiving the inputs, processing through certain calculation mode, and displaying the results on the window
- run_model.ipynb – the notebook version of this program, which can be used online (in google colab) without installing python locally

4.3 Usage instruction

The program can be executed by running the main.py file in a python environment with all the required dependencies installed. Then on the main window, the user can enter experimental data and their corresponding units. If no unit is given for that variable, the default unit shown in the box will be used.

- t – time (min)
- q_t – the amount of adsorbate absorbed at time t (mg/g)
- C_0 – the initial concentration of adsorbate (mg/L)
- C_i – the initial concentration of adsorbent (g/L)

Since the calculator can take in multiple datasets, the form of input for t and q_t is to split the data points by ',' and start each dataset with a new line. For C_0 data, the values should be split by ',' where its order corresponds to the order of datasets automatically. C_i is expected to be unchanged during the analysis so a single value is expected to be input.

5 Results and Discussion

5.1 Software Performance

The user interface of this software and example usage for different calculation modes are shown in figures 2 to 4.

Adsorption Model Calculator

Enter experimental data in the first box splitting with ',' and its unit in the second box.
For t and qt, start a new line if there are multiple datasets.

t min qt mg/g

C0 mg/L Cs g/L

Results ?

Figure 2. The default window of user interface

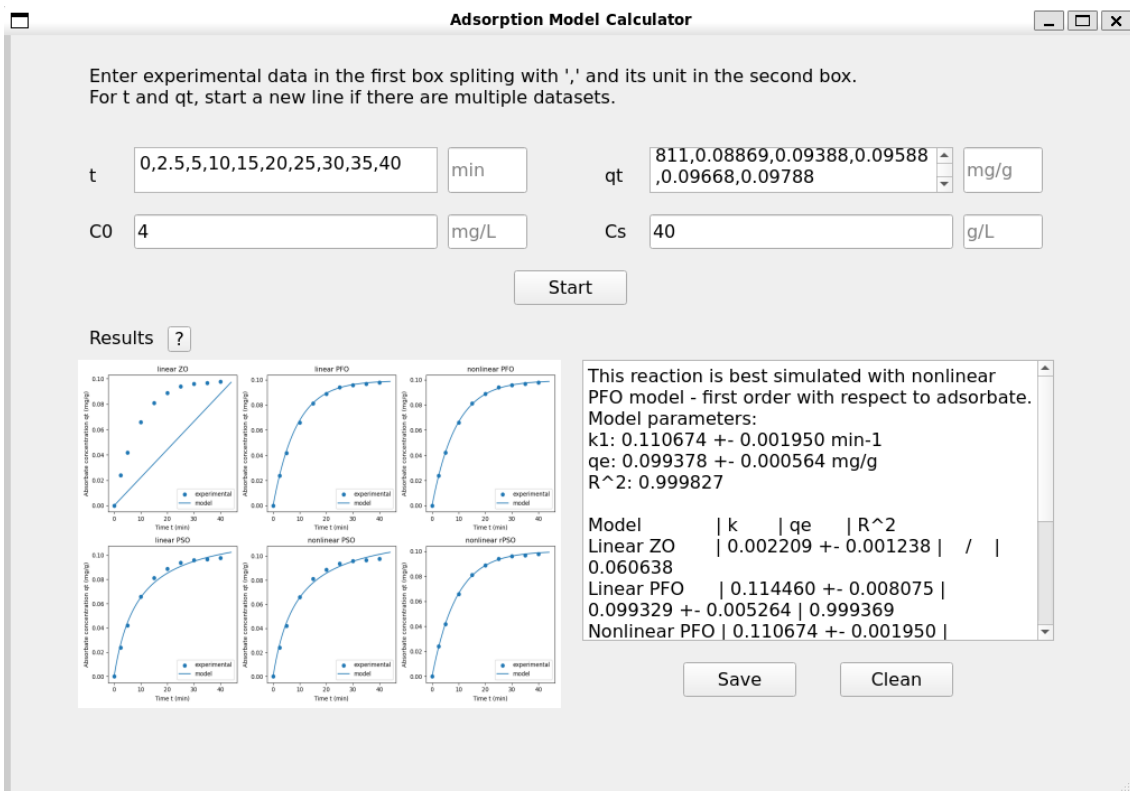


Figure 3. Example calculation of a single dataset

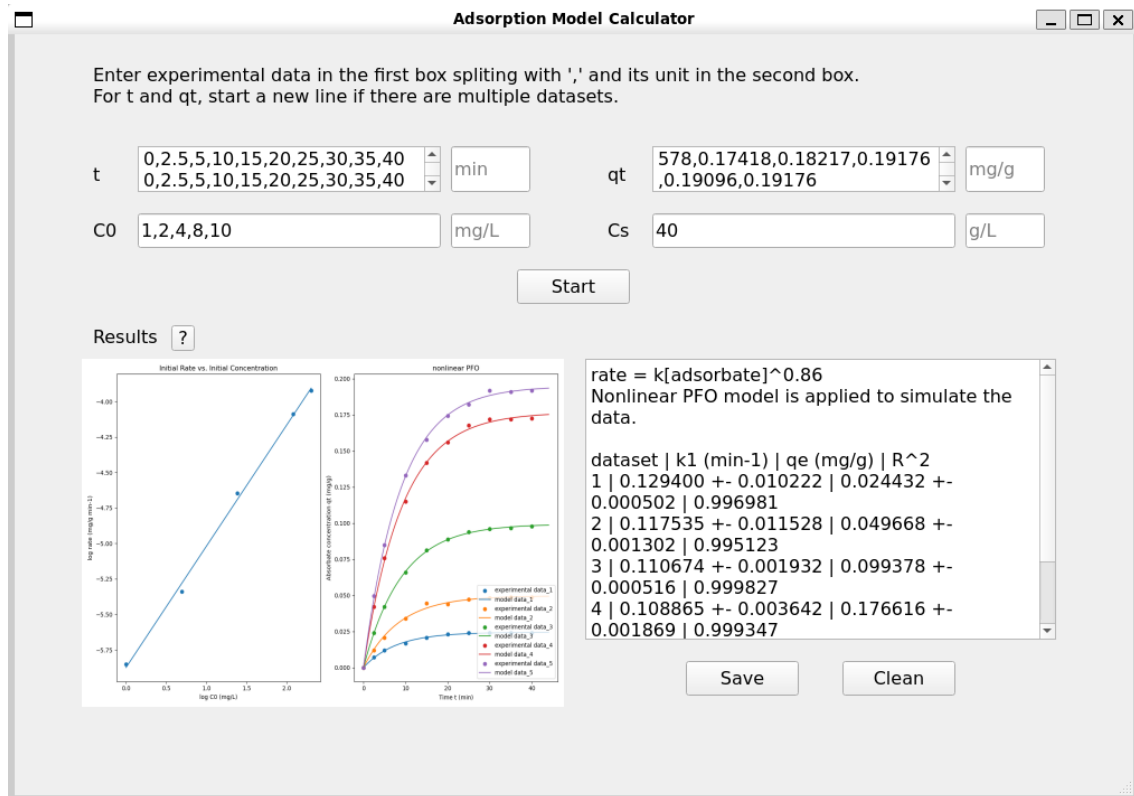


Figure 4. Example calculation of multiple datasets

The time taken to run each calculation is around 3 seconds where the *curve_fit()* function for nonlinear models is very time-consuming. The text result contains the conclusion from models and all the calculated values (rate constant k , adsorbate absorbed at equilibrium q_e , error of these parameters and predicted R^2) presented in a table. This allows the user to derive more valuable insights from the data.

5.2 Function validation

After testing the software with 13 sets of data, all the functions have been put into use and the calculated results are analysed using Excel (spreadsheet in appendix). The majority of models gives R^2 greater than 0.97 so the parameters and their uncertainties (both real value and percentage error) are recorded and the model with highest R^2 is marked. Also, the percentage progress of reaction for the first and last data point are calculated for each experiment, providing a view of the quality of dataset.

Firstly, among all the parameters calculated out in 65 experiments, 117 values can be found in the literature including k_1 of PFO, k_2 of PSO and q_e of either model. Compared with these data, the software achieves a consistency of 68.4% (80 out of 117) where the difference between calculated and literature parameters is within 10% of the literature values on the natural logarithm scale.

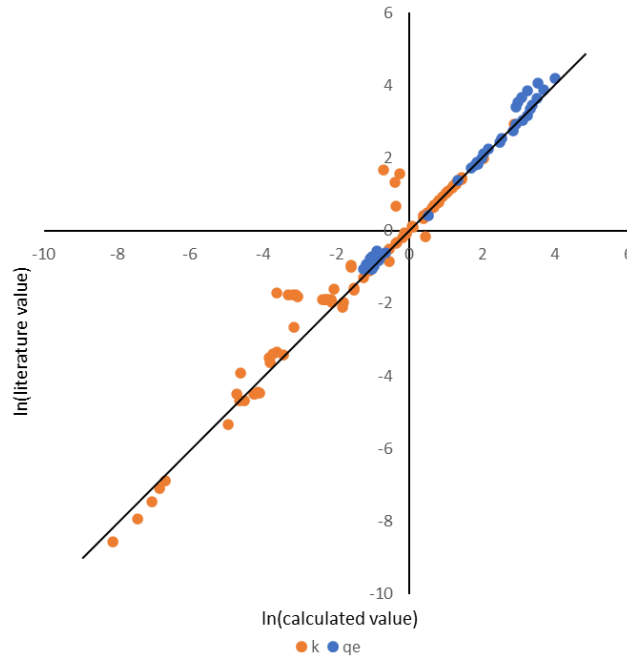


Figure 5. Comparison between calculated and literature values on natural logarithm scale

For 90% of the experiments, the R^2 value of linear ZO model is lower than 0.5 which indicates that the zero-order rate equation is not very applicable for most of the adsorption reactions. There are 6 experiments give R^2 values greater than 0.8 where it is found that the first data point is recorded within 2% of the reaction progress and final data point only reaches around 30-55% reaction completion (q_e). Moreover, the cases of both linear ZO model and linear PFO model giving particularly bad optimisation are due to the lack of data at the early stage of reaction (the first data point collected is already over 50% of the reaction progress). Thus, the confidence of these models depend on the quality of dataset to a large extent.

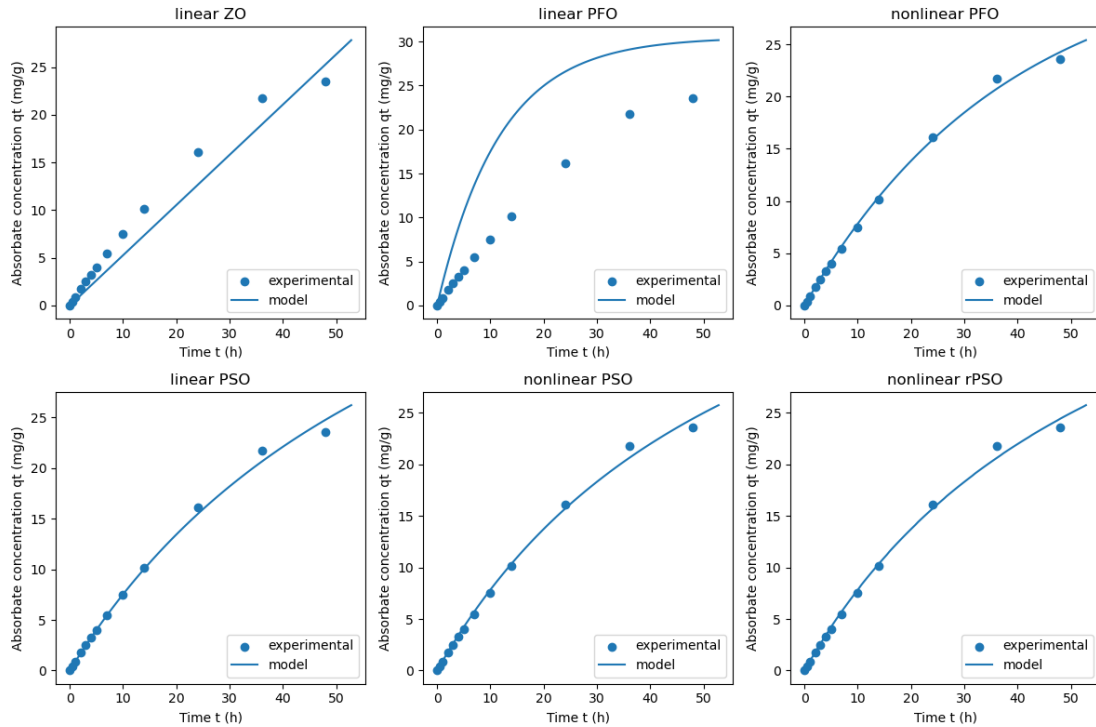


Figure 6. Fitting of the first experiment in dataset 11 which linear ZO model reaches the highest R^2 0.945

The nonlinear models fitting with *curve_fit()* always provide higher R^2 values than the linear models fitting with *LinearRegression()* in all the experiment calculations. This is why multiple datasets are fitted with the nonlinear model when the reaction order is determined. However, the difference of PFO models is more obvious than that of PSO models. There are only 13 experiments fit the linear PFO model with a R^2 value greater than 0.97, but 60 experiments fit the linear PSO model. Since the implementation of linear PFO model includes a temporary guess of q_e , it is again relying on the quality of dataset and its real kinetics. For example, if the reaction experiment is second-order in the literature but it also fits the linear PFO model, this might be due to the finishing point less than 85% of the reaction.

For the second-order like experiments, most of the results of nonlinear PSO model are quite good with R^2 values greater than 0.985 but around half of the calculations show slightly improved results for modified rPSO model. Bullen's rPSO model takes initial concentrations of adsorbate C_0 and adsorbent C_s into account which is more useful when there's limited data for early kinetics or the reaction condition varies. Furthermore, it is found that the q_e calculated by rPSO model for datasets 1 to 7 differs much from that of PSO model and sometimes gives very large uncertainty so that the significance of this parameter needs to be further studied.

Apart from these models, experiments in datasets 1 to 7 are calculated together to analyse its reaction order. There are five datasets turned out to be first-order (between 0.8 and 1.5) and two being zero-order (lower than 0.5). The R^2 values for four of them which are predicted correctly are greater than 0.987 (figure 7). As discussed above, the linear ZO model does not apply to these data, and the fact that it didn't reach second-order kinetics matters. From Bullen's study, the earlier the first kinetic data collected in the reaction progress (better within 20%), the more consistent the initial rates using tangent between the origin and the first kinetic data is with that calculated from PSO parameters. Thus, due to technical errors of initial rate calculation caused by the quality of dataset, the determination of reaction order is not fully accurate.

All the discussion about performance of models have been based on the value of coefficient of determination (R^2) which is the most widely used evaluation method in studies. However, for most of the experiments, it shows that both nonlinear PFO and nonlinear PSO models are well fitted where it's hard to distinguish if one is overfitting. The model with higher R^2 is not always the correct one. Hence, due to the limitation of this method, a combination with other error functions like the sum of square error (SSE), average relative error (ARE) and Chi-square (X^2) can be employed to validate the models.

Lastly, the percentage uncertainties of 92% calculated parameters are in the range of 0.5% to 40% where the rate constant k always comes with larger error than q_e for each model. This is because the error calculation involves random differences generated for 100 simulations, and the kinetic models are more sensitive to the change in rate constant. The parameters with huge errors can also indicate the model being bad fitted to the dataset.

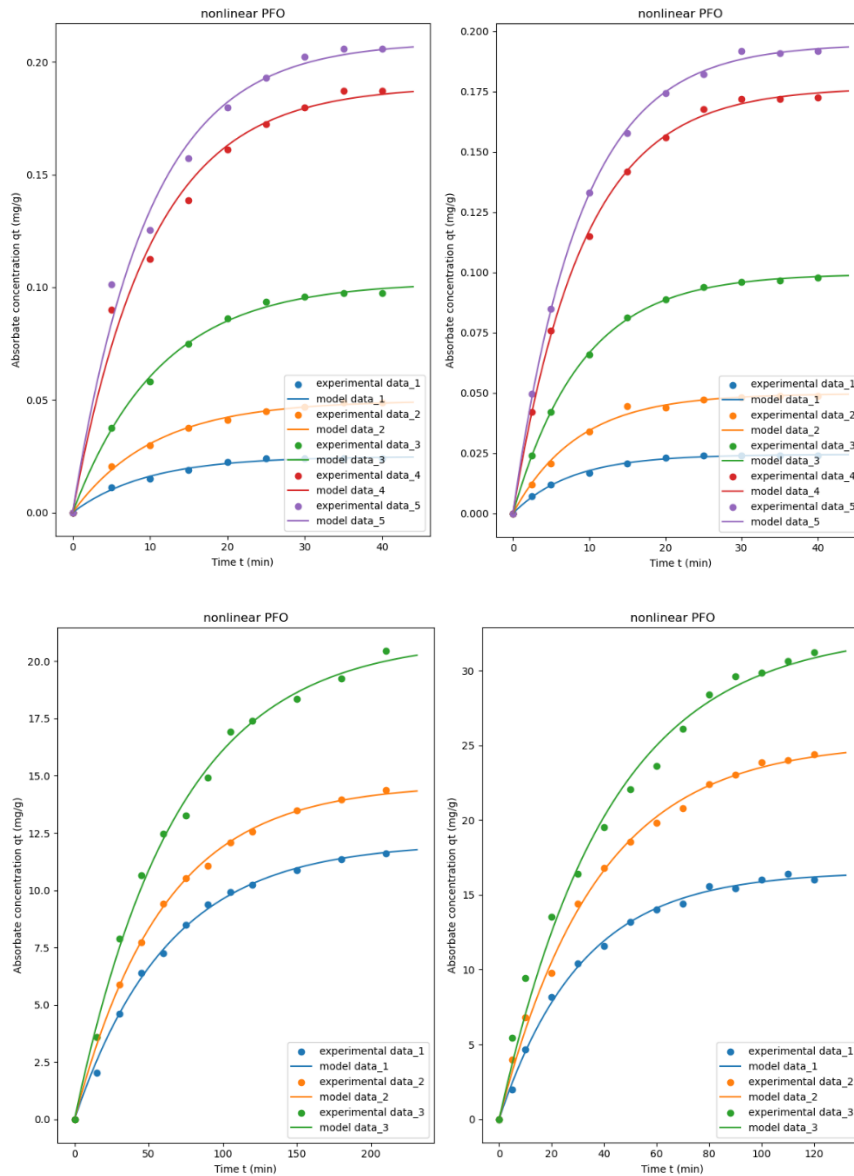


Figure 7. Datasets 1,2,4,6 fitting with nonlinear PFO model where the reaction order analysis predicted correctly

5.3 Discussion of adsorbate

The results of reaction kinetics calculated by this software can be used for many engineering purposes. For instance, an appropriate adsorbent is selected by comparing systems with the same adsorbate, such as As(V) (datasets 1 and 2) and PFOS (datasets 11 and 12). For As(V), although the q_e predicted is similar for both haematite (Fe_2O_3 , etc.) and feldspar (SiO_2 , Al_2O_3 , etc.) as adsorbent, the rate constant of system with feldspar is greater for both PFO and PSO models which means the adsorption happens faster. For PFOS, adsorption by zeolite appears to be hundreds times faster than activated carbon, but its adsorption capacity is hundreds times poorer. Hence, there is a trade-off between time efficiency and cost effectiveness.

Table 2. Parameters comparison between datasets 1 and 2 (As)

DATASET 1			DATASET 2		
PFO	C_0 (mg/L)	k_1 (min ⁻¹)	q_e (mg/g)	k_1 (min ⁻¹)	q_e (mg/g)
	1	0.104	0.025	0.129	0.024
	2	0.095	0.050	0.118	0.050
	4	0.089	0.102	0.111	0.099
	8	0.098	0.189	0.109	0.177
	10	0.103	0.209	0.114	0.195
PSO	C_0 (mg/L)	k_2 (g mg ⁻¹ min ⁻¹)	q_e (mg/g)	k_2 (g mg ⁻¹ min ⁻¹)	q_e (mg/g)
	1	3.610	0.031	4.794	0.030
	2	1.554	0.063	1.938	0.062
	4	0.654	0.131	0.900	0.124
	8	0.449	0.235	0.497	0.221
	10	0.434	0.257	0.485	0.242

Table 3. Parameters comparison between datasets 11 and 12 (PFOS)

DATASET 11			DATASET 12		
PFO	Temp.(°C)	k_1 (h ⁻¹)	q_e (mg/g)	k_1 (h ⁻¹)	q_e (mg/g)
	10	0.026	33.81	0.095	0.410
	20	0.036	25.90	0.103	0.359
	30	0.042	22.11	0.102	0.348
	40	0.045	19.76	0.111	0.309
	50	0.048	18.49	0.118	0.289
PSO	Temp.(°C)	k_2 (g mg ⁻¹ h ⁻¹)	q_e (mg/g)	k_2 (g mg ⁻¹ h ⁻¹)	q_e (mg/g)
	10	0.0003	55.16	0.167	0.522
	20	0.0006	40.37	0.221	0.450
	30	0.0009	33.53	0.224	0.436
	40	0.0011	29.48	0.287	0.382
	50	0.0013	27.37	0.339	0.353

6 Conclusion and Future Work

Knowing the adsorption kinetics of reactions is important for scientists and engineers to design the water treatment strategies on a large scale. To help people analyse the adsorption kinetic data quickly, this software is developed which includes six kinetic models (linear ZO, linear PFO, nonlinear PFO, linear PSO, nonlinear PSO and revised PSO). By entering experimental data t , q_t , C_0 and C_s on the user interface, the calculated parameters k , q_e of each model and its performance R^2 will be shown and also visualised in a graph. Generally, the parameters calculated for 65 experiments have a consistency of 68.4% compared with that in literatures, indicating that the software works and gives reasonable results. In most cases, the nonlinear PFO and nonlinear PSO models perform well with R^2 greater than 0.97 but the linear ZO model is worse. Thus, if more datasets are validated using this software, the less applicable models involved can be modified and other new useful ones can be introduced in the future. Furthermore, considered about the quality of experimental data, the calculated results should be evaluated from more aspects beside the value of R^2 . The method of reaction order analysis needs to be improved to better cover different situations.

7 Acknowledgements

I would like to express many thanks to my supervisors Prof. Dominik Weiss and Dr. Jay Bullen for their guidance, support and feedback throughout this project. The good atmosphere of this research group made me relaxed during the summer. I'm also thankful to my friends Yuyang Cheng, Xiaochong Jin, Zi Ye and my dear family who provided care and encouragement when I was anxious and confused. Thanks to my classmate Dongzi Ding for all the discussion with me about this project which brought me many useful ideas. Lastly, I'm always grateful to my idol Stray Kids and Yaowen Liu for giving me a lot of happiness while studying.

8 References

1. J. Fawell, M. J. Nieuwenhuijsen, "Contaminants in drinking water: Environmental pollution and health", *British Medical Bullentin*, vol. 68, no. 1, (2003), p. 199-208
<https://doi.org/10.1093/bmb/ldg027>
2. P. Li, D. Karunanidhi, T. Subramani, K. Srinivasamoorthy, "Sources and Consequences of Groundwater Contamination", *Archives of Environmental Contamination and Toxicology* 80, (2021), p. 1-10
<https://doi.org/10.1007/s00244-020-00805-z>
3. S. Sharma, A. Bhattacharya, "Drinking water contamination and treatment techniques", *Applied Water Science* 7, (2017), p. 1043-1067
<https://doi.org/10.1007/s13201-016-0455-7>
4. J. C. Bullen, C. Lapinee, P. Salaün, R. Vilar, D. J. Weiss, "On the application of photocatalyst-sorbent composite materials for arsenic(III) remediation: Insights from kinetic adsorption modelling", *Journal of Environmental Chemical Engineering*, vol. 8, no. 5, (2020), p. 104033
<https://doi.org/10.1016/j.jece.2020.104033>
5. G. Sun, "Arsenic contamination and arsenicosis in China", *Toxicology and Applied Pharmacology*, vol. 198, no. 3, (2004), p. 268-271
<https://doi.org/10.1016/j.taap.2003.10.017>
6. R. Singh, S. Singh, P. Parihar, V. P. Singh, S. M. Prasad, "Arsenic contamination, consequences and remediation techniques: A review", *Ecotoxicology and Environmental Safety*, vol. 112, (2015), p. 247-270
<https://doi.org/10.1016/j.ecoenv.2014.10.009>
7. A. Bhardwaj, R. Rajput, K. Misra, "Chapter 9 – Status of Arsenic Remediation in India", *Advances in Water Purification Techniques*, (2019), p.219-258
<https://doi.org/10.1016/B978-0-12-814790-0.00009-0>
8. P. Grandjean, R. Clapp, "Perfluorinated Alkyl Substances: Emerging Insights Into Health Risks", vol. 25, no. 2, (2015), p. 147-163
<https://doi.org/10.1177/1048291115590506>
9. S. Alka, S. Shahir, N. Ibrahim, M. J. Ndejiko, D. N. Vo, F. A. Manan, "Arsenic removal technologies and future trends: A mini review", *Journal of Cleaner Production*, vol. 278, (2021), p. 123805
<https://doi.org/10.1016/j.jclepro.2020.123805>

10. B. C. Crone, T. F. Speth, D. G. Wahman, S. J. Smith, G. Abulikemu, E. J. Kleiner, "Occurrence of per- and polyfluoroalkyl substances (PFAS) in source water and their treatment in drinking water", vol. 49, no. 24, (2019), p. 2359-2396
<https://doi.org/10.1080/10643389.2019.1614848>
11. K. L. Tan, B. H. Hameed, "Insight into the adsorption kinetics models for the removal of contaminants from aqueous solutions", Journal of the Taiwan Institute of Chemical Engineers, vol. 74, (2017), p. 25-48
<https://doi.org/10.1016/j.jtice.2017.01.024>
12. J. C. Bullen, S. Saleesongsom, K. Gallagher, D. J. Weiss, "A Revised Pseudo-Second-Order Kinetic Model for Adsorption, Sensitive to Changes in Adsorbate and Adsorbent Concentrations", Langmuir 37(10), (2021), 3189-3201
<https://dx.doi.org/10.1021/acs.langmuir.1c00142>
13. S. Lagergren, "Zur theorie der sogenannten adsorption geloster stoffe.", Kungliga Svenska vetenskapsakademiens. Handlingar 24, (1898), p. 1-39
14. Y. S. Ho, G. McKay, "Pseudo-second order model for sorption processes", Process Biochemistry, vol. 34, no. 5, (1999), p. 451-465
[https://doi.org/10.1016/S0032-9592\(98\)00112-5](https://doi.org/10.1016/S0032-9592(98)00112-5)
15. L. Pogliani, "Pseudo-zero-order reactions", Reaction Kinetics and Catalysis Letters 93, (2008), p. 187-191
<https://doi.org/10.1007/s11144-008-5233-6>
16. E. D. Revellame, D. L. Fortela, W. Sharp, R. Hernandez, M. E. Zappi, "Adsorption kinetic modelling using pseudo-first order and pseudo-second order rate laws: A review", Cleaner Engineering and Technology, vol. 1, (2020), p. 100032
<https://doi.org/10.1016/j.clet.2020.100032>
17. D. B. Singh, G. Prasad, D. C. Rupainwar, "Adsorption technique for the treatment of As(V)-rich effluents", Colloids and Surfaces A: Physicochemical and Engineering Aspects, vol. 111, no. 1-2, (1996), p. 49-56
[https://doi.org/10.1016/0927-7757\(95\)03468-4](https://doi.org/10.1016/0927-7757(95)03468-4)
18. A. Maiti, J. K. Basu, S. De, "Desorption kinetics and leaching study of arsenic from arsenite/arsenate-loaded Natural Laterite", International Journal of Environmental Technology and Management, vol. 12, no. 2-4, (2010), p. 294-307
<https://doi.org/10.1504/IJETM.2010.031534>
19. B. R. Manna, S. Dey, S. Debnath, U. C. Ghosh, "Removal of Arsenic from Groundwater using Crystalline Hydrous Ferric Oxide (CHFO)", Water Quality Research Journal, vol. 38, no. 1, (2003), p. 193-210
<https://doi.org/10.2166/wqri.2003.013>
20. N. Y. Mezenner, A. Bensmaili, "Kinetics and thermodynamic study of phosphate adsorption on iron hydroxide-eggshell waste", Chemical Engineering Journal, vol. 147, no. 2-3, (2009), p. 87-96
<https://doi.org/10.1016/j.cej.2008.06.024>
21. K. Periasamy, C. Namasivayam, "Removal of copper(II) by adsorption onto peanut hull carbon from water and copper plating industry wastewater", Chemosphere, vol. 32, no. 4, (1996), p. 769-789
[https://doi.org/10.1016/0045-6535\(95\)00332-0](https://doi.org/10.1016/0045-6535(95)00332-0)

22. C. Wei, X. Song, Q. Wang, Z. Hu, "Sorption kinetics, isotherms and mechanisms of PFOS on soils with different physicochemical properties", *Ecotoxicology and Environmental Safety*, vol. 142, (2017), p. 40-50
<https://doi.org/10.1016/j.ecoenv.2017.03.040>
23. H. Son, B. An, "Investigation of Adsorption Kinetics for Per- and Poly-fluoroalkyl substances (PFAS) Adsorption onto Powder Activated Carbon (PAC) in the Competing Systems", *Water,Air,&Soil Pollution* 233, (2022), 129
<https://link.springer.com/article/10.1007/s11270-022-05599-5>
24. Y. Yin, Y. Shan, D. Ma, L. Yang, M. Zhan, P. Liu, B. Lou, B. Zhang, W. Jiao, L. Yin, "Temperature-Electrokinetic Co-Driven Perfluorooctane Sulfonic Acid (PFOS) Adsorption on Geo-Adsorbents", *Processes*, vol. 11, no. 6, (2023), p.1856
<https://www.mdpi.com/2227-9717/11/6/1856>
25. B. Yan, G. Munoz, S. Sauve, J. Liu, "Molecular mechanisms of per- and polyfluoroalkyl substances on a modified clay: a combined experimental and molecular simulation study", *Water Research*, vol. 184, (2020), p. 116166
<https://doi.org/10.1016/j.watres.2020.116166>
26. W. Hu, J. Xie, H. W. Chau, B. C. Si, "Evaluation of parameter uncertainties in nonlinear regression using Microsoft Excel Spreadsheet", *Environmental Systems Research* 4, (2015)
<https://doi.org/10.1186/s40068-015-0031-4>

9 Appendix

1. The analysis spreadsheet that contains the collected data and calculated results.
[validation_data.xlsx](#)
2. The spreadsheet of comparison between calculated parameters and literature values
[comparison_data.xlsx](#)