Imperial College London

Department of Earth Science and Engineering

MSc in Applied Computational Science and Engineering

Independent Research Project

Application of Pseudo-Second-Order kinetic Models in Software for Predicting Removal of Contaminants from Sorbents

by

Yanan Xiong

Email: yanan.xiong19@imperial.ac.uk
GitHub username: acse-yx1419
Repository: https://github.com/ese-msc-2021/irp-acse-yx1419

Supervisors:

Prof. Dominik Weiss
Dr. Jay Bullen

September 2022

Abstract

Due to the increase in the population, the pressures on water remediation today pushed researchers to study more technologies to improve efficiency. Adsorption is widely used today, measuring how quickly an adsorbent material can remove the contaminant. Specifically, Pseudo-Second-Order (PSO) model is a famous model to be applied in industry to decide how much adsorbent and time is needed. However, when it comes to obtaining the model's parameters using the experimental data, the traditional option with an Excel spreadsheet is time-consuming with a large volume of data. This project combines the linearised PSO, non-linearised PSO and revised PSO models with the python codes, making it fast to process the experimental data to obtain the parameters of the models. Not only the python code but also software for all users was generated, shortening the data processing time. Sixty-five experiments from 14 literature sources, including four categories of adsorbents, were analysed with the software. The goodness of fitness is verified by R^2. Generally, non-linearised PSO shows better goodness of fitness than linearised PSO but is less fitted than revised PSO for all categories of adsorbents except metal carbonates. For each set of experimental data, the running time of obtaining optimised parameters for every PSO model is as short as a few minutes. In the future, the software can be improved to include more kinetic models as the users' option.

Acknowledgement

I would like to thank the following people for helping with this research project: My supervisors, Professor Dominik Weiss and Dr. Jay Bullen, for providing guidance and feedback throughout this project. My dear Friends in table K: Tengteng Huang, Daniel Xing, Hansong Xiao, Lesley Zhong, Loken Yin, Haoran Cheng, Jimmy Huo, Haotong Jin, and lulu li for their support and love in both studying and learning. They bring me numerous unforgettable memories through this year, including the days we worked lately until 00am in RSM 1.47, and the bubble tea we ordered almost every day. My roommate Luna Wu for her patience and encouragement. The last trains of District Line every day throughout this summer. Especially my beloved parents for sponsoring my education.

Contents

1 Introduction and Objectives

1.1 Backgrounds

Chemical and organic contaminants have become an accessible environmental problems for a long time (Joardar et al. 2022, Rojas & Horcajada 2020). The purity and quality of water may be important threats to people's survival and the ecosystem (Sharma et al. 2015). The change in water quality can influence water use in industrial or agriculture and then harm people's health through the food chain (Staroń et al. 2019). Nowadays, many researchers have discussed the dangers of bioaccumulation and toxicity brought to life. The first famous example is heavy metals, which is defined by metals with atomic density less than 4.5 g/cm^3 (Chang et al. 2018). Heavy metals are sourced from both natural and anthropogenic sources. Once heavy metals enter the water and soil system, they can accumulate in different kinds of crops, such as rice and vegetables (Riaz et al. 2021), and then become part of the food chain. Another famous example is arsenic, which is a cosmopolitan problem in groundwater (Organization & Organization 2006). The most harmful problems lead by arsenic include but are not least attacking people's cardiovascular and respiratory systems (Singh et al. 2015).

Moreover, more than one billion people are suffering from access to safe drinking water. This number will keep growing with the problems we face, such as population growth, urbanization, and climate change. (Dev et al. 2022, Latif et al. 2020) Moreover, besides the drinking water, wastewater from the petroleum and pharmaceutical industries is also concerned (Gadipelly et al. 2014, Adham et al. 2012). Currently, we are making progress on the remediation technologies of contaminated water, such as the inactivation of toxins, the oxidative transformation of emerging contaminants, and organic synthesis (Sharma et al. 2015). However, there are still many scholars are trying to make efforts to improve the efficiency from different aspects, such as how the efficiencies related to the selection of adsorbent materials (Arslan et al. 2022) and the novel materials in different fields (Linghu et al. 2017, Saka et al. 2012). Therefore, water remediation is still a severe challenge that we are facing. To release the pressure, the improvement of efficiency in water treatment seems to be urgent.

Generally, industrial water remediation studies often regard adsorption by materials as a feasible method, measuring how quick an adsorbent material can remove the contaminants. The laboratory experiments aim to obtain results about the solutes at equilibrium states, which show the adsorption capacity, and about the kinetics of adsorption. The adsorption capacity and the kinetics are described via the empirical formulas of different models (Simonin 2016). According to recent studies, the water treatment capacity has been boosted in this century due to the dramatic increase in population, followed by the significant growth of energy consumption (Smith & Liu 2017). However, the energy costs can take up 30% of the total operation cost for a water treatment plant (Biehl & Inman 2010). Thus, the accuracy of adsorption kinetics of the utilized adsorption model is crucial to both improve efficiency and to save energy consumption (Gibelhaus et al. 2022, Bullen et al. 2021) for both batch and column or continuous flow treatments (Bullen et al. 2021, Dichiara et al. 1900), in order to predict the adsorption status via the model. The model's accuracy can be understood as the accuracy of the model's parameter, which varies with the different chemical reactions. Meanwhile, an experiment from the laboratory is limited to being operated under the condition of a certain adsorbent. In practice, the adsorbent (C_s) concentrations change with the different concentrations of treated contaminants in the influent (C_0). Hence, the adsorption models ought to be sensitive to the operating conditions, making predictions possible (Bullen et al. 2021).

1.2 Literature Review

The pseudo-second-order (PSO) equation from Ho & McKay (1999) is a widely used adsorption kinetic model (Bullen et al. 2021), taking the form

$$\frac{dq_t}{dt} = k_2(q_e - q_t)^2 \tag{1}$$

where t is the time (minutes), q_t is the amount of adsorbate adsorbed per mass of adsorbent at time t (mg g^{-1}), and q_e is the amount of adsorbate adsorbed at equilibrium (mg g^{-1}).

The concentration of aqueous adsorbate left at time t is

$$C_t = C_0 - C_s q_t \tag{2}$$

where C_t is the concentration of the aqueous adsorbate at time t (mg L^{-1}), C_0 is the initial adsorbate concentration at $t = 0$ (mg L^{-1}), and C_s is the concentration of the adsorbent (g L^{-1}).

According to Bullen et al., the PSO model has advantages in the following aspects: a) the simple mathematical form, b) adaptability for a wide range of reaction mechanisms, and c) able to be integrated and rearranged to provide linear equations so that the model parameters $k2$ and q_e can be obtained (2021). However, as mentioned above, the limitations appear to be the lack of operation concentrations of adsorbents and contaminants in the equations. Therefore, the obtained parameters $k2$ and q_e can only be valid under the same condition as the experiment. As a result, the PSO model can not provide prediction to the adsorption kinetics with any given C_0 and C_s, of course, causing limited capability in practice. (Bullen et al. 2021)

To cover the shortage of the PSO model, Bullen et al. published a revised PSO (rPSO) model which includes the C_0 and C_s in the equation (2021):

$$\frac{dq_t}{dt} = k' C_t (1 - \frac{q_t}{q_e})^2 \tag{3}$$

where the unit of rate constant k' is L $g^{-1}min^{-1}$. For including the operation conditions C_0 and C_t in the equation, the k' here substitutes the $\frac{k_2 q_e^2}{C_0^2}$, combining the eq(1) and eq(2) together.

1.3 Aim and objectives

When the kinetic data are obtained via excel (Appendix .3) (Bullen et al. 2021), requiring large amount of calculations and manipulations manually. The workload becomes extremely heavy regarding numerous data sets, especially when comparing the results from different experiments and models. Nevertheless, generating plots for visualisation brings an extra burden. This project implemented the original PSO with linear regression and nonlinear fitting, and Bullen et al.'s revised PSO into an executable python routine. Users can apply and process kinetic data by obtaining the adsorption kinetics and capacity and the visualised data to plots. What is more, software for desktops was developed with the help of QT Designer. The interface was designed to be easy to use, only requiring to enter the data obtained from the laboratory, making it friendly to all kinds of users.

2 Methodology

2.1 Dataset

In total, 14 literature sources with 65 kinetic experiments were collected, including 8 literature sources with 9 data sets and 37 experiments where C_0 is changed, and 6 literature sources with 6 data sets and 28 experiments where C_s is changed. Among these 14 literature sources, there are 5 sources with 19 experiments for metal oxides as sorbents, 2 sources with 10 experiments for activated carbon, 1 literature sources with 4 experiments for metal carbonates, and 6 sources with 32 experiments for organic sorbents (table 1). All the data collected have conversed to the uniform units (Appendix .1).

	Cs changes	**C0 changes**
Metal Oxides	Singh et al. (1996) Manna et al. (2003) Mezenner & Bensmaili (2009)	Debnath et al. (2017) Shipley et al. (2012)
Activated Carbon	Aworanti & Agarry (2019) Nandhakumar (2015)	
Metal Carbonates		Lazaridis et al. (2004)
Organics	Maiti et al. (2010) Lim & Lee (2015) Natarajan & Manivasagan (2016)	Sağ & Aktay (2000) Reategui et al. (2010) Staroń et al. (2019)

Table 1: The table of all literature sources used

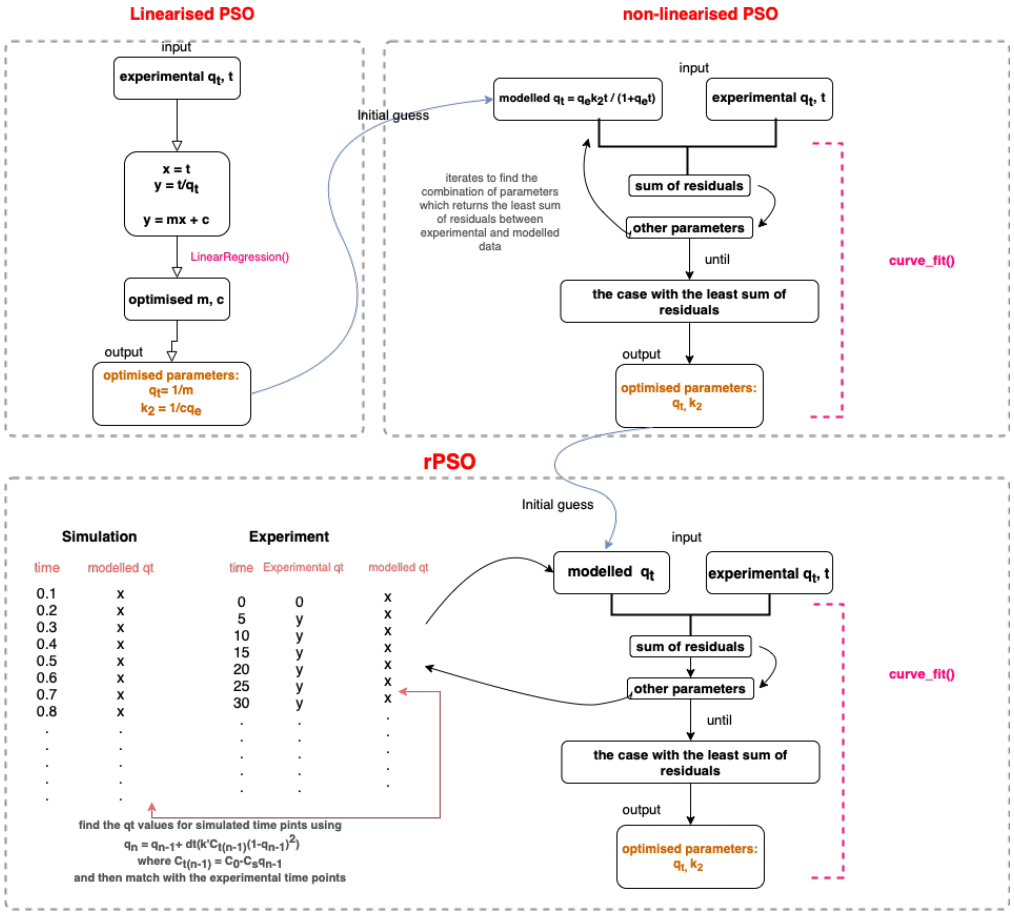


Figure 1: Workflow for obtaining q_t with PSO models

2.2 Mathematical Approaches for the Determination of Adsorption Kinetics

2.2.1 Linearised Pseudo-second-order

PSO takes the form in eq(1). When apply the boundary conditions $t = 0$ and $t = t$, $qt = 0$ and $qt = qt$ and integrate eq(1), we can have:

$$q_t = \frac{q_e^2 kt}{1 + q_e t} \tag{4}$$

The first approach toward adsorption kinetics uses the linearised form of the integrated PSO rate equation

$$\frac{t}{q_t} = \frac{1}{k_2 q_e^2} + \frac{t}{q_e} \tag{5}$$

which can be regarded as a linear form of $y = mx + c$. t/q_t is plotted as y depended on t, where m is the slope of the linear regression and c is the y-intercept. The model was optimised using the built-in function $LinearRegression()$ from $sklearn$ package, minimising the residual sum of squares between the data points and the linear regression. The slope and the y-intercept of the optimised model can be returned as outputs when call $LinearRegression()$ (*sklearn.linear*$_m odel.LinearRegression$ n.d.). Hence, the equilibrium adsorption parameters can be obtained by

$$q_e = \frac{1}{m} \tag{6}$$

and

$$k_2 = \frac{1}{c q_e^2} \tag{7}$$

The workflow chart is shown in fig 1.

2.2.2 Non-linearised PSO

Secondly, the non-linearised PSO kinetics were directly optimised via an built-in function named $curve\_solve$ from the $SciPy$ package. Starting from the initial guesses to the parameters, $curve\_solve$ iterates the input model (eq4) and calculates the sum of squared residuals between the model and experimental data for each combination of parameters. In the end, it returns the case of parameter combination resulting the least sum of squared residuals, which is the so-called "optimised parameters". This built-in function has no limit to the type of curves. The initial guess for the parameters of the model that applied for $curve\_solve$ are from the results obtained from the linear regression. (*scipy.optimize.curve*$_f it|SciPyv1.9.1Manual$ n.d.) Boundary conditions of $q_t > 0$ and $k_2 > 0$ were applied. The workflow chart is shown in fig 1.

2.2.3 Non-linearised revised PSO

Thirdly, the kinetics were obtained from the revised rate equation (eq.3). Due to the difficulty of integration of the rPSO, the adsorbed adsorbate at the nth data point was expressed as:

$$q_n = q_{n-1} + (t_n - t_{n-1})(k' C_{t(n-1)}(1 - \frac{q_{n-1}}{q_e})^2) \tag{8}$$

where $C_t(n-1) = C_0 - C_s q_{n-1}$.
The results from non-linearised PSO were used as the initial guess. As mentioned above, rate constant here changes to k' here, the initial guess for k' uses $\frac{k_2 q_e^2}{C_0^2}$.

Additionally, due to the calculation of q_t in rPSO involving iterations, which needs a defined time step, dt was input from users such as 0.1, 0,01 or 0.001 minute. Figure 1 is an example of a workflow with 0.1 minute as a time step. Taking the end of experiment time as the time range, after obtaining the q_t values for each simulated time point via the eq (6), the q_t at corresponding experimental time points were matched.

The final rPSO parameters k' and q_e were obtained by optimising the model to the experimental data via $curve_solve$ to minimise the sum of squared residuals. As mentioned above, $curve_solve$ has no limitation of the equation to be optimised, making it suitable for the unconventionally iterative equation (eq.8). Boundary conditions of $q_t > 0$ and $k' > 0$ were applied. Smaller dt leads to more accurate q_t but leads to an exponential increase in the program's running time. The workflow chart is shown in fig 1.

2.3 Filtration of data

Moreover, sometimes the experimental data need to be removed conditionally, according to the ratio of C_t to C_0. According to eq(2), when $C_s q_e$, which is equal to C_t, $\ll C_0$, the adsorbate is in excess. Ideally, when C_t/C_0 is closed to zero, the model is approximated to be the form of the original PSO model (Bullen et al. 2021). On the contrary, when C_t is close to C_0, which means that C_t/C_0 is close to 1, the adsorbent is in excess. At this time, a threshold of C_t/C_0 may be needed according to the user's requirement to remove data points unsuitable for the model. Thus, the threshold of C_t/C_0 was taken into account as an input. For example, when the users enter 1, all experimental data will be input to the program.

2.4 Evaluation of goodness of fitness

All kinetic adsorption parameters and capacities are optimised by minimising the residual sum of squares between the models and the experiments. For showing the goodness of fitness of the results, R^2 are calculated for every group of models and experiment, which is expressed as:

$$R^2 = 1 - \frac{sum\ of\ squared\ difference\ between\ model\ and\ experiment}{sum\ of\ squared\ difference\ between\ experiments\ and\ mean} \tag{9}$$

Besides, the uncertainties of obtained k_2, k' and q_e were calculated using Monte Carlo method with 200 simulations, turning the method introduced by (Hu et al. 2015) into python code. In the beginning, the modelled q_t is added with a random value:

$$q_t(new) = norm.ppf(random(), 0, std) \tag{10}$$

where the inverse of the normal cumulative distribution for the random probability, zero as the mean and standard deviation of q_t. Two hundred sets of the simulation were generated as "real experimental" data for new parameter optimisation processes. Thus, 200 optimised values can be obtained for each parameter. By calculating the 95% confidence interval for the 200 values of one parameter, the uncertainty of the optimised parameters can be regarded as half of the confidence interval.

2.5 Final Presentation

First of all, the modelled q_t values corresponding to the time points of the experiments are compared with the experimental data by calculating R^2. Secondly, for each experimental data set, the modelled results are visualised in the figures. The three approaches of PSO, as well as the linear plots of linearised PSO fitting, are presented with annotations of the goodness of the fitness, which is R^2.

Thirdly, for future applications of the models in engineering, it is necessary to have an interface for users to obtain the parameters of models and the visualised plots easily. This project managed to generate a program for users to present all the optimised data and plots for future use in predicting the status of the water treatment. The users only need to input the experimental data, the value of C_0 and C_s and the self-selection of dt and threshold of C_t/C_0.

3 Code Metadata

- System environment:
  - System: macOS Monterey
  - Chip: Apple M1 Pro
- Code language: Python 3.9.12
- IDE: Jupyter Notebook, Visual Studio Code 1.70.1
- Installation requirements and dependencies:
  - Numpy 1.21.5
  - Matplotlib 3.5.1
  - Sklearn 1.0.2
  - Scipy 1.7.3

4 Software Development Life Cycle

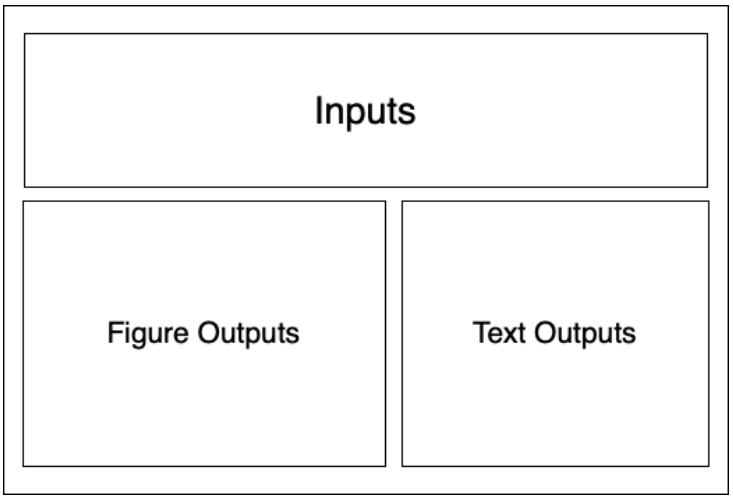


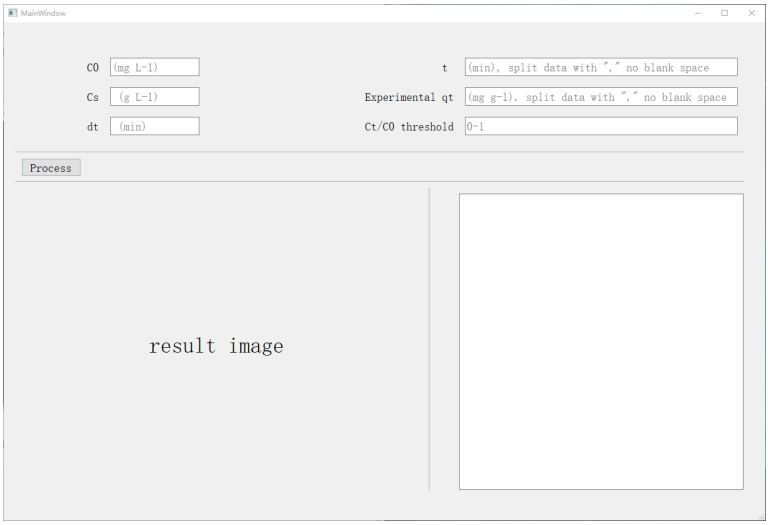
Figure 2: Interface design for the program

1. **Define requirements:**
   This software is developed for users to analyse the experimental data of adsorption easily, to obtain optimised parameters for three different kinds of PSO models for comparisons, and thus enable reaction prediction in industries. Besides, results as plots are essential for visual comparison.
2. **Design and Prototyping:**
   Generally, the interface includes the zones for inputs and outputs. Not only that, the output zone can be divided as figure output for plots and text output for parameters, uncertainties, and R^2 values (Figure 2).
3. **Software Development:**
   - System: Windows 11 21H2

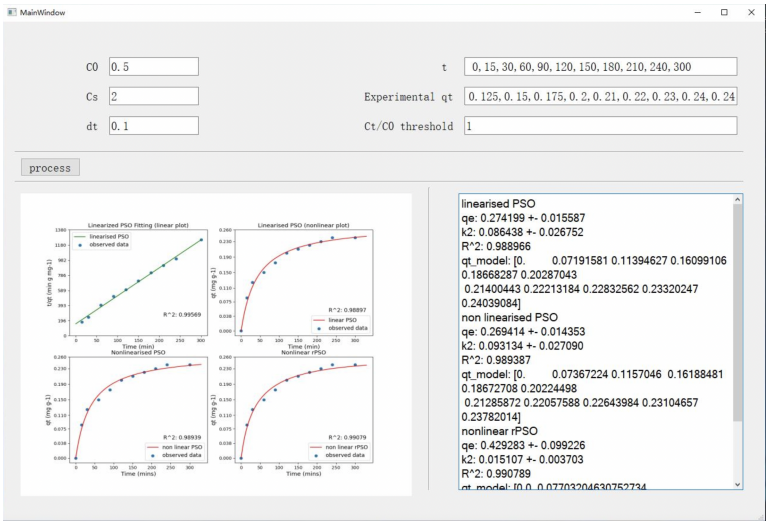
- Chip: Intel i7-12700H
- Code language: Python 3.9.12
- IDE: PyCharm
- Tools: QT Designer
- Installation requirements and dependencies:
  - PySide2 5.15.2.1
  - pyinstaller
  - Numpy 1.21.5
  - Matplotlib 3.5.1
  - Sklearn 1.0.2
  - Scipy 1.7.3

5 Results

5.1 Interface of the Software



(a) Default interface



(b) Example result interface (Maiti et al. 2010)

Figure 3

The final interface accomplished the design. From the beginning, in the default interface, users are asked to enter all the experimental data, including arrays of q_t, t, and values of C_0 and C_s. The placeholders for the array inputs notify users to split the input data with commas and without any space in between. According to users input of threshold of C_t/C_0, the input of q_t is to be filtered. After processing of filtered data, the results are presented in sections. Plots include a linear plot of linearised PSO, nonlinear plots of non-linearised PSO and rPSO, while text outputs include the calculated parameters, their Monte Carlo uncertainties, and predicted q_t at each time point for each model.

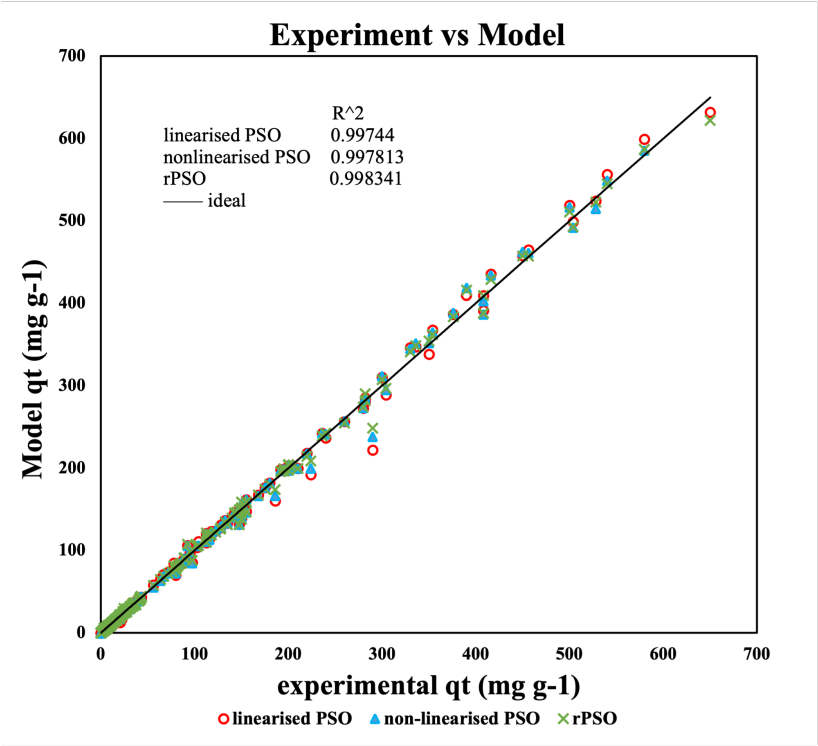
5.2 Parameter optimisation results

After analysis of experiments, parameters of three kinds of models, rate constants k (k_2 or k'), adsorbed adsorbate at equilibrium (q_e), and their uncertainties were calculated via the software. Afterwards, the obtained kinetics of the reactions were used to calculate q_t values for 655 data points from 14 literature sources (Appendix .2). On the purpose of verification to the results, examples of experiments were analysis via Excel, as described by Bullen et al. (2021) (Appendix .3). The results from software is proven for delivering the correct outputs as Excel spreadsheets, except for the uncertainties. There is a minor difference in uncertainties between the spreadsheets and the software. These differences do not mean that the uncertainty obtained by the software is wrong because the calculation of 200 simulations involves random numbers, which leads to changing 200 parameters. Thus, the confidence interval of the 200 parameters changes in cases, causing the uncertainty of a model's parameter to vary. What is important, the change in uncertainties is as minor as $< 1\%$.

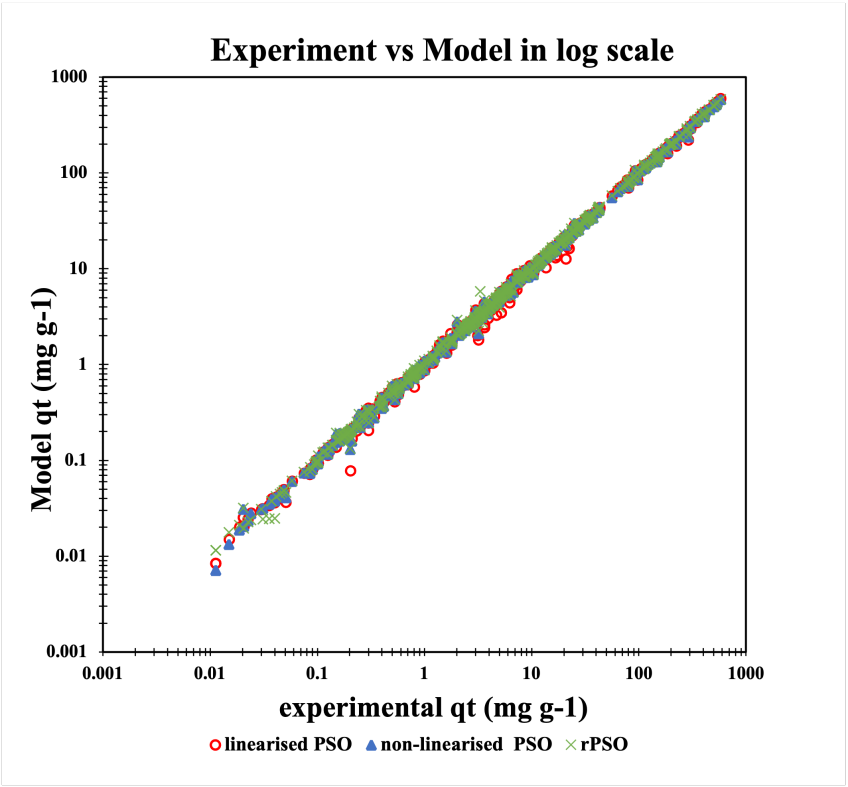
Figure 4a and figure 4b contract the 655 calculated and experimental q_t values in normal and log scale. As shown in the figures, The normal scale plot can show the points with higher q_t values better. By the contrast, the log scale plot present the points with lower q_t values clearly. Generally, the rPSO is better fitted than the original PSO models. However, even though the linearised PSO is the least fitted model, the R^2 is as high as 0.99744. The R^2 of the best performed rPSO reaches 0.998341. Furthermore, it can be seen that a few more points from linearised PSO model are off the ideal one-to-one line than the other two models, which accords with the R^2 difference among them. What's more, when q_t values are as large as more than 100 mg/g, the off-line points are less, as shown below. The R^2 of the three models show better goodness of fitness when q_t is above 100 mg/g (table 2).

range	linearised PSO	non-linearised PSO	rPSO
$q_t < 100mg/g$	0.991214	0.992055	0.994898
$q_t > 100mg/g$	0.995393	0.996791	0.996983

Table 2: The goodness of fitness of the three models when q_t is above and below 100 mg/g



(a) Experimental data verse three models



(b) Experimental data verse three models in log scale

Figure 4: 1.)The plot of three models against experimental data (the 14 literature sources, 65 experiments) 2.) Cross-calibration plot: red circles indicate q_t values calculated using the linearised original PSO model; blue triangles indicate the non-linearised PSO model, while green crosses indicated the rPSO model; black line shows the ideal 1:1.

Regarding the sorbent types, the values of R^2 are analysed separately. In general, all three models can perform well for all types of sorbents (R^2 >0.99). For all categories of sorbents, the non-linearised model data fit the experimental data better than the linearised PSO. Also, the rPSO model can deliver better results than the non-linearised model for all sorbent types except metal carbonates. The rPSO relative misfits the metal carbonates. In terms of model types, all three models deliver the best results for metal oxides and relatively bad results for metal carbonates (Appendix .2).

Sorbent Type	**linearised PSO**	**non-linearised PSO**	**rPSO**
Metal Oxides	0.999250613	0.999349098	0.999422384
Activated Carbon	0.998008084	0.998273317	0.998928766
Metal Carbonates	0.994006025	0.994518425	0.966856555
Organics	0.999017792	0.99914207	0.999194375

Table 3: The R^2 values of each type of sorbents

A box-and-whisker plot (Figure 5) shows the relative errors in percentage of all 655 data points. The plot shows the 25%, 75% percentiles and median for each model. As shown, the interquartile range of the three models is hard to be distinguished. However, the outliers of linearised PSO tend to underestimate the results, while the outliers of non-linearised PSO and rPSO tend to overestimate the results. The linearised PSO's negative value of mean relative error stands for the underestimation of the data points. The mean of relative error of rPSO is very close to zero, which means more modelled data points are almost the same as the experimental data points.

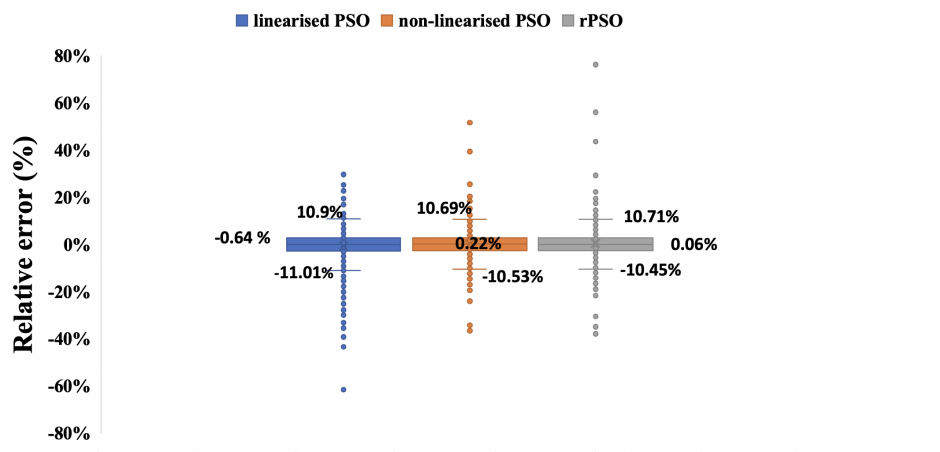


Figure 5: Box-and-whisker plot shows the relative error of q_t calculated via the linearised PSO, non-linearised PSO and rPSO model. The boxes present the 25, 75% percentiles and the median. The data points greater than the top of boxes plus 1.5×interquartile range or data points less than the bottom of boxes minus 1.5×interquartile range are defined as outliers. The whiskers present the minima and maxima of data points not including the outliers.

According to the calculations of relative error, the average relative errors of modelled q_t below and above $< 100mg/g$ are shown below. Overall, when the q_t values are greater, the relative error is greater (table 4).

range	linearised PSO	non-linearised PSO	rPSO
$< 100mg/g$	-0.24%	-0.13%	-0.14%
$> 100mg/g$	-0.69%	0.24%	0.44%

Table 4: The average relative errors of the three models when q_t is above and below 100 mg/g

By taking 0.2 as the threshold of C_t/C_0 to distinguish the end of the reactions (table 5). The points with $C_t/C_0 < 0$ are meaningless and may be recorded mistakes. After excluding these points, generally, the models show more errors at the end of the reactions.

range	linearised PSO	non-linearised PSO	rPSO
< 0.2	-0.11%	-0.26%	-0.24%
$>= 0.2$	-0.64%	0.25%	0.65%

Table 5: The average of relative errors of the three models when C_t/C_0 is above and below 0.2

6 Discussion

First of all, the quality of the results from the three kinds of the model is different, as proved by the R^2 from the comparisons of the results from each model with the 655 experimental data points. All of the three models present show good performance on fitness as $R^2 > 0.997$, including the linearised PSO, which shows the least degree of fitness. However, the difference between the three models is not as high as other literature source. Bullen et al.'s R^2 for PSO and rPSO are 0.9217 and 0.999, showing more significant difference. The distinction between this project and the literature source may be linked to the number of experimental data involved. For this project, 655 data points from 14 literature sources were calculated for R^2. By contrary, there is only 6 literature sources and 198 data points involved in Bullen et al.'s R^2 (2021). The more well-fitted sorbent types including metal oxides and organics involved may increase the R^2 value in overall data points.

Secondly, the process of calculations with rPSO involves the selection of time steps. For the purpose of lessening errors, the time step was selected as 0.001 minutes for most of the 65 experiments. Generally, when lessening the time step from 0.01 to 0.001 minutes, the result can be 0.3% more accurate, which is not significant. However, there are a few data sets that are very sensitive to the selection of time steps. For example, in the case with $C_0 = 50mg/L$ and $C_s = 2g/L$, when the selected time step is larger than the 0.001-minute scale, q_e can be a few hundred times higher than the q_e obtained in non-linearised PSO model, which is an obvious mistake.

Thirdly, due to there being more iterations involved in the rPSO, the main time-consuming is on rPSO. Even so, the running time of the program for one experiment with dt =0.1 minutes is up to 1 minute, including showing the plots on the screen. In contrast, the time consuming for obtaining parameters of one model may take 5 minutes, while the calculations of uncertainty takes longer time. Scientists may make some mistakes when manipulate with Excel spreadsheets sometimes. There is no doubt about the convenience of the program, not only because of the much shorter running time than using spreadsheets but also because of the ease of manipulating.

Fourthly, due to the Windows system being widely used in industries, the software is only designed for Windows users. However, the python codes are applicable under any system.

Lastly, even though the generated software shows good performance in terms of manipulation convenience and results, there are some possible improvements for the future. First of all, this software is limited to PSO; more kinds of models, such as the pseudo-first-order kinetic model, are possible to be solved with codes according to users' selection. Furthermore, on the interface, the used model or models can be an option for users. Last but not least, software that is applicable under the Apple OS system or an application on cellphones may be developed for the future.

7 Conclusion

Under the current stress of water use in both domestic and industries, the efficiency of water treatment is urgent to be improved. As a widely used method, pseudo-second-order kinetic models, including linearised PSO, non-linearised PSO and revised PSO, were selected to be implemented with codes for the convenience of processing adsorption experimental data. With the accurate estimations of the parameters of the models, the operating conditions and treatment process can be better predicted. With the help of the generated software, 65 experiments from 14 literature sources were analysed. When the experimental data points are as much as 655, the values of R^2 of the three kinds of models show similar numbers of > 0.997. However, rPSO model is the best fitted to the experimental data in most cases of experiments and sorbents. The three models are provided in the software. Besides the experimental data as inputs, the time step and threshold of C_t/C_0 can be customised by users. However, the software can be further improved. In terms of function, the software can be improved to include more kinds of models for users to select in the future.

References

Adham, S. S., Hussain, A., Katebah, M. & Dores, R. (2012), Using advanced water treatment technologies to treat produced water from the petroleum industry, Society of Petroleum Engineers.
URL: *https://www.onepetro.org/conference-paper/SPE-157108-MS*

Arslan, H., Eskikaya, O., Bilici, Z., Dizge, N. & Balakrishnan, D. (2022), 'Comparison of cr(vi) adsorption and photocatalytic reduction efficiency using leonardite powder', *Chemosphere* **300**, 134492. ID: 271852.
URL: *https://www.sciencedirect.com/science/article/pii/S0045653522009857*

Aworanti, O. & Agarry, S. (2019), 'Kinetics, isothermal and thermodynamic modelling studies of hexavalent chromium ions adsorption from simulated wastewater onto parkia biglobosa-sawdust derived acid-steam activated carbon'.

Biehl, W. H. & Inman, J. A. (2010), 'Energy optimization for water systems', *Journal - American Water Works Association* **102**(6), 50–55.
URL: *https://www.jstor.org/stable/41314357*

Bullen, J. C., Saleesongsom, S., Gallagher, K. & Weiss, D. J. (2021), 'A revised pseudo-second-order kinetic model for adsorption, sensitive to changes in adsorbate and adsorbent concentrations', *Langmuir* **37**(10), 3189–3201.
URL: *http://dx.doi.org/10.1021/acs.langmuir.1c00142*

Chang, L., Shen, S., Zhang, Z., Song, X. & Jiang, Q. (2018), 'Study on the relationship between age and the concentrations of heavy metal elements in human bone', *Annals of translational medicine* **6**(16), 320.
URL: *https://www.ncbi.nlm.nih.gov/pubmed/30363972*

Debnath, A., Bera, A., Chattopadhyay, K. K. & Saha, B. (2017), 'Facile additive-free synthesis of hematite nanoparticles for enhanced adsorption of hexavalent chromium from aqueous media: Kinetic, isotherm, and thermodynamic study', *Inorganic and nano-metal chemistry* **47**(12), 1605–1613.
URL: *https://www.tandfonline.com/doi/abs/10.1080/24701556.2017.1357581*

Dev, V. V., Nair, K. K., Baburaj, G. & Krishnan, K. A. (2022), 'Pushing the boundaries of heavy metal adsorption: A commentary on strategies to improve adsorption efficiency and modulate process mechanisms', *Colloid and interface science communications* **49**.
URL: *https://dx.doi.org/10.1016/j.colcom.2022.100626*

Dichiara, A. B., Weinstein, S. J. & Rogers, R. E. (1900), 'On the choice of batch or fixed bed adsorption processes for wastewater treatment', *Industrial engineering chemistry research* **54**(34), 8579–8586.
URL: *http://dx.doi.org/10.1021/acs.iecr.5b02350*

Gadipelly, C., Perez-Gonzalez, A., Yadav, G. D., Ortiz, I., Ibanez, R., Rathod, V. K. & Marathe, K. V.

(2014), 'Pharmaceutical industry wastewater: Review of the technologies for water treatment and reuse', *Industrial engineering chemistry research* **53**(29), 11571–11592.
URL: *http://dx.doi.org/10.1021/ie501210j*

Gibelhaus, A., Postweiler, P. & Bardow, A. (2022), 'Efficient modeling of adsorption chillers: Avoiding discretization by operator splitting', *International Journal of Refrigeration* . ID: 271448.
URL: *https://www.sciencedirect.com/science/article/pii/S0140700722001244*

Ho, Y. S. & McKay, G. (1999), 'Pseudo-second order model for sorption processes', *Process Biochemistry* **34**(5), 451–465.
URL: *https://dx.doi.org/10.1016/S0032-9592(98)00112-5*

Hu, W., Xie, J., Chau, H. W. & Si, B. C. (2015), 'Evaluation of parameter uncertainties in nonlinear regression using microsoft excel spreadsheet', *Environmental Systems Research* **4**(1), 1.
URL: *https://link.springer.com/article/10.1186/s40068-015-0031-4*

Joardar, M., Das, A., Chowdhury, N. R., Mridha, D., Das, J., De, A., Majumder, S., Majumdar, K. K. & Roychowdhury, T. (2022), 'Impact of treated drinking water on arsenicosis patients with continuous consumption of contaminated dietary foodstuffs: A longitudinal health effect study from arsenic prone area, west bengal, india', *Groundwater for Sustainable Development* **18**, 100786. ID: 312046.
URL: *https://www.sciencedirect.com/science/article/pii/S2352801X22000637*

Latif, A., Sheng, D., Sun, K., Si, Y., Azeem, M., Abbas, A. & Bilal, M. (2020), 'Remediation of heavy metals polluted environment using fe-based nanoparticles: Mechanisms, influencing factors, and environmental implications', *Environmental pollution (1987)* **264**, 114728.
URL: *https://dx.doi.org/10.1016/j.envpol.2020.114728*

Lazaridis, N. K., Pandi, T. A. & Matis, K. A. (2004), 'Chromium(vi) removal from aqueous solutions by mgalco3 hydrotalcite: sorptiondesorption kinetic and equilibrium studies', *Industrial engineering chemistry research* **43**(9), 2209–2215.
URL: *http://dx.doi.org/10.1021/ie030735n*

Lim, S.-F. & Lee, A. Y. W. (2015), 'Kinetic study on removal of heavy metal ions from aqueous solution by using soil', *Environmental science and pollution research international* **22**(13), 10144–10158.
URL: *https://link.springer.com/article/10.1007/s11356-015-4203-6*

Linghu, W., Yang, H., Sun, Y., Sheng, G. & Huang, Y. (2017), 'One-pot synthesis of ldh/go composites as highly effective adsorbents for decontamination of u(vi)', *ACS Sustainable Chemistry Engineering* **5**(6), 5608–5616. doi: 10.1021/acssuschemeng.7b01303.
URL: *https://doi.org/10.1021/acssuschemeng.7b01303*

Maiti, A., Basu, J. K. & De, S. (2010), 'Desorption kinetics and leaching study of arsenic from arsenite/arsenate-loaded natural laterite', *International journal of environmental technology and management* **12**(2-4), 294–307.

URL: *https://www.inderscienceonline.com/doi/10.1504/IJETM.2010.031534*

Manna, B., Dey, S., Debnath, S. & Ghosh, U. (2003), 'Removal of arsenic from groundwater using crystalline hydrous ferric oxide (chfo)', *Water Quality Research Journal* **38**(1), 193–210.
URL: *https://search.proquest.com/docview/18724454*

Mezenner, N. Y. & Bensmaili, A. (2009), 'Kinetics and thermodynamic study of phosphate adsorption on iron hydroxide-eggshell waste', *Chemical engineering journal (Lausanne, Switzerland : 1996)* **147**(2), 87–96.
URL: *https://dx.doi.org/10.1016/j.cej.2008.06.024*

Nandhakumar, V. (2015), 'Adsorption of rose bengal dye from aqueous solution onto zinc chloride activated carbon', *SOJ Materials Science and Engineering* **3**(2), 1–9.

Natarajan, R. & Manivasagan, R. (2016), 'Biosorptive removal of heavy metal onto raw activated sludge: Parametric, equilibrium, and kinetic studies', *Journal of environmental engineering (New York, N.Y.)* **142**(9).
URL: *http://ascelibrary.org/doi/abs/10.1061/(ASCE)EE.1943-7870.0000961*

Organization, W. H. & Organization, W. H. (2006), 'Regional office for south', *East Asia* **6**.

Reategui, M., Maldonado, H., Ly, M. & Guibal, E. (2010), 'Mercury(ii) biosorption using lessonia sp. kelp', *Applied biochemistry and biotechnology* **162**(3), 805–822.
URL: *https://agris.fao.org/agris-search/search.do?recordID=US201301851109*

Riaz, M., Kamran, M., Fang, Y., Wang, Q., Cao, H., Yang, G., Deng, L., Wang, Y., Zhou, Y., Anastopoulos, I. & Wang, X. (2021), 'Arbuscular mycorrhizal fungi-induced mitigation of heavy metal phytotoxicity in metal contaminated soils: A critical review', *Journal of hazardous materials* **402**, 123919.
URL: *https://dx.doi.org/10.1016/j.jhazmat.2020.123919*

Rojas, S. & Horcajada, P. (2020), 'Metal–organic frameworks for the removal of emerging organic contaminants in water', *Chemical reviews* **120**(16), 8378–8415.
URL: *http://dx.doi.org/10.1021/acs.chemrev.9b00797*

Saka, C., Ömer Şahin & Küçük, M. M. (2012), 'Applications on agricultural and forest waste adsorbents for the removal of lead (ii) from contaminated waters', *International Journal of Environmental Science and Technology* **9**(2), 379–394. ID: Saka2012.
URL: *https://doi.org/10.1007/s13762-012-0041-y*

Sağ, Y. & Aktay, Y. (2000), 'Mass transfer and equilibrium studies for the sorption of chromium ions onto chitin', *Process biochemistry (1991)* **36**(1), 157–173.
URL: *https://dx.doi.org/10.1016/S0032-9592(00)00200-4*

scipy.optimize.curve$_fit|SciPyv1.9.1Manual(n.d.)$.

URL:*https://docs.scipy.org/doc/scipy*

Sharma, V. K., Zboril, R. & Varma, R. S. (2015), 'Ferrates: Greener oxidants with multimodal action in water treatment technologies', *Accounts of chemical research* **48**(2), 182–191.
URL: *http://dx.doi.org/10.1021/ar5004219*

Shipley, H. J., Engates, K. E. & Grover, V. A. (2012), 'Removal of pb(ii), cd(ii), cu(ii), and zn(ii) by hematite nanoparticles: effect of sorbent concentration, ph, temperature, and exhaustion', *Environmental science and pollution research international* **20**(3), 1727–1736.
URL: *https://link.springer.com/article/10.1007/s11356-012-0984-z*

Simonin, J.-P. (2016), 'On the comparison of pseudo-first order and pseudo-second order rate laws in the modeling of adsorption kinetics', *Chemical engineering journal (Lausanne, Switzerland : 1996)* **300**, 254–263.
URL: *https://dx.doi.org/10.1016/j.cej.2016.04.079*

Singh, D. B., Prasad, G. & Rupainwar, D. C. (1996), 'Adsorption technique for the treatment of as(v)-rich effluents', *Colloids and Surfaces A: Physicochemical and Engineering Aspects* **111**(1), 49–56. ID: 271384.
URL: *https://www.sciencedirect.com/science/article/pii/0927775795034684*

Singh, R., Singh, S., Parihar, P., Singh, V. P. & Prasad, S. M. (2015), 'Arsenic contamination, consequences and remediation techniques: A review', *Ecotoxicology and environmental safety* **112**, 247–270.
URL: *https://dx.doi.org/10.1016/j.ecoenv.2014.10.009*

sklearn.linear$_m odel.LinearRegression(n.d.)$.
URL:*https://scikit-learn.org*

Smith, K. & Liu, S. (2017), 'Energy for conventional water supply and wastewater treatment in urban china: A review', *Global Challenges* **1**(5), 1600016–n/a.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1002/gch2.201600016*

Staroń, P., Chwastowski, J. & Banach, M. (2019), 'Sorption behavior of methylene blue from aqueous solution by raphia fibers', *International journal of environmental science and technology (Tehran)* **16**(12), 8449–8460.
URL: *https://link.springer.com/article/10.1007/s13762-019-02446-9*

8 Appendix

.1 Literature Sources

Link to google drive: https://imperiallondon-my.sharepoint.com/:x:/g/personal/yx1419_ic_ac_uk/EaKCHQSaNhdLkNm2WSnMxvQBxl4Sn0PgX3SrvVxF1JhHqA?e=mrwpfV

.2 Python Results

https://imperiallondon-my.sharepoint.com/:x:/g/personal/yx1419_ic_ac_uk/ETNrphNi1kRPr4ua8C2GmfcBczFVkVdaX5phNzkijzFDag?e=ff1drl

.3 Verification of Results

https://imperiallondon-my.sharepoint.com/:x:/g/personal/yx1419_ic_ac_uk/EdF69Qms6qtFoMfhWdlN2ssBxLs2kBRj4A4lqhw8FGdzwA?e=JMdywm