American Journal of Chemical Engineering
2023; 11(5): 85-94
http://www.sciencepublishinggroup.com/j/ajche
doi: 10.11648/j.ajche.20231105.11
ISSN: 2330-8605 (Print); ISSN: 2330-8613 (Online)

Development of a Generalized Statistical Model for Hexavalent Chromium Removal Using Electrocoagulation Through SVR Regression Analysis

Mohamad Salem[1, *], Nabil Abdelmonem[1], Ehssan Nassef[2]

[1]Chemical Engineering Department, Faculty of Engineering, Cairo University, Cairo, Egypt
[2]Petrochemical Eng. Department, Faculty of Engineering, Pharos University, Alexandria, Egypt

Email address:
mg.serag@gmail.com (Mohamad Salem)
[*]Corresponding author

To cite this article:
Mohamad Salem, Nabil Abdelmonem, Ehssan Nassef. (2023). Development of a Generalized Statistical Model for Hexavalent Chromium Removal Using Electrocoagulation Through SVR Regression Analysis. *American Journal of Chemical Engineering*, *11*(5), 85-94. https://doi.org/10.11648/j.ajche.20231105.11

Received: November 7, 2023; **Accepted**: November 27, 2023; **Published**: December 8, 2023

Abstract: Hexavalent chromium Cr(VI) is a highly toxic pollutant that poses a significant threat to human health and the environment. Electrocoagulation is a promising technology for the removal of Cr(VI) from wastewater. This work reviews and evaluates statistical models developed in different studies published between 2015 and 2021 on the removal of Cr(VI) using electrocoagulation. The analysis showed that none of the models was found to be conclusive, and that they all suffer from issues such as overfitting and the inability to generalize beyond the experiment domain. These models were also highly dependent on the selection of input parameters, model selection criteria, and experimental design. An attempt to solve this problem was to utilize Machine Learning (ML) techniques to develop a more robust model that can provide generalized and accurate predictions on a broader domain. The model was developed using Support Vector Machines Regression analysis (SVR). Data compiled from previously published works were used to train and test the model using a 50:50 split ratio. The model was able to make more generalized predictions but lacked accuracy. As with all ML models, this model requires a higher volume of high-quality data to improve its accuracy. The study concluded that there is still a need for more robust statistical models that can effectively capture the complexity of the electrocoagulation process and generalize well beyond the experiment domain.

Keywords: Statistical Models, Electrocoagulation, Hexavalent Chromium, Machine Learning

1. Introduction

Today's growing population, urbanization, industry, and the agricultural sector are all contributing to a reduction in water quality and quantity. Industrial wastewater treatment has been viewed as a desirable method to safeguard water resources. Industrial effluent is produced in large quantities daily with a substantial amount of harmful trash. Water resources are important now more than ever to both human growth and the environment, especially with the increase of contaminated effluent released as a result of industrial development. [1]

Chromium is one of the problematic heavy metal pollutants that can be present in industrial wastewater. Since heavy metals cannot decompose, they will persist in the environment and accumulate in living organisms causing environmental problems and diseases. Chromium is generated from mining, tanning, electroplating, wood preservatives, paints, textile dyeing, and plants producing industrial inorganic chemicals and pigments. [2, 3]

Chromium exists in wastewater steams as trivalent Cr(III) and hexavalent Cr(VI). Cr(III) is an essential micronutrient as long as it exists in trace amounts. Cr(VI) is a harmful variant and a potentially carcinogenic compound.

Many studies have been conducted on removing chromium using different treatment methods such as electrochemical precipitation, ion exchange, membrane processing, solvent extraction, coagulation, and adsorption.

Electrocoagulation involves many physical and chemical phenomena; the primary process variables, individual and combined effects have a complex relationship with the effectiveness of removing contaminants.

The aim of this work is to investigate electrocoagulation as a method for chromium removal from wastewater using statistical modeling methods. This study will review published models that can predict the percent removal of hexavalent chromium Cr(VI) and assess the validity and accuracy of each model.

2. Literature Review

Electrocoagulation has been studied as a promising, highly efficient, cost-effective method for the removal of chromium from industrial waste waters.

Electrocoagulation involves many physical and chemical phenomena; the primary process variables, individual and combined effects have a complex relationship with the effectiveness of removing contaminants.

Experimental works aiming to study the effects of process variables on the efficiency of electrocoagulation. This is done by changing a single factor while keeping all other factors fixed at a certain set of conditions in the majority of EC research on water/wastewater treatment. A statistical regression is done to model the experimental results and use it to make predictions. This approach to EC process modelling necessitates numerous experimental runs and the selection of a suitable domain and variable ranges. Narrow experimental domain, or too few runs may result in a poorly performing model. [4]

K. Thirugnanasambandham and K. Shine investigated the influence of electrocoagulation parameters using Box-Behnken experimental design. RSM and Artificial Neural Networks (ANN) statistical methods were used for analysis and prediction. The study examined the effects of pH, current density, electrode distance, and electrolysis time on the EC process. The results show that RSM and ANN can both accurately characterize the current electrocoagulation process. Using stainless steel electrodes, initial pH of 6, current density of 25 mA/cm^2, electrode distance of 4 cm, and electrolysis period of 30 min were found to be the ideal working parameters achieving 97% chromium removal with a 0.12 kWh/m^3 electrical energy expenditure. [1]

Statistical optimization study done by Sunil R. Patel and Sachin P. Parikh using response surface methodology (RSM). The Experimental setup consists of a glass beaker and two copper plates (15 cm × 4.6 cm × 0.1 cm) serving as the electrodes. The anode and cathode were connected to a DC power supply. A wooden block was used to maintain electrode distance. Artificial waste water was created by dissolving potassium permanganate ($K_2Cr_2O_7$) in distilled water. Process variables studied are pH, Cr(VI) initial concentration, current density, Electrolyte concentration (NaCl), interelectrode distance, and treatment duration. This work also included a kinetic study to determine the rate constant for chromium removal under the experiment conditions. The experimental results demonstrated that the optimal operating parameters for achieving 93.33% removal efficiency of Cr(VI) ions from simulated waste water are current density of 41.32 A/m^2, electrode distance of 1.4 cm, initial pH of 5.65, time of electrocoagulation of 40 min, and initial conductivity 0.21 µs. [5]

A study by Rasha H. Salman and colleagues investigated the treatment of real tannery wastewater using electrocoagulation followed by Reverse Osmosis (RO). The study examined four operating parameters namely the current density, electrode distance, NaCl concentration and treatment duration. The Cr^{3+} removal percent at the optimum conditions (1.5 g/l NaCl, 25 mA/cm^2, 2h treatment duration and 20mm electrode distance) was 88.8%, which does not meet the environmental requirements in Iraq. Following the EC with an RO process increased the removal percent up to 99.89% which is well beyond the acceptable limit. [6]

Edwar Aguilar et al have studied the removal of chromium from tannery wastewater using aluminum as the electrode material. The study investigated the effect of current density, treatment time and pH. The work concluded that total chrome removal can be achieved at a current intensity of 2.9 A, a pH of 8.4, for a duration of 21 min. [7]

Nahid Genawi et al performed experimental studies on real wastewater from tanneries. The study focused on the effect of initial chromium concentration, current density and pH. The optimum conditions were determined using ANOVA analysis where maximum chromium removal was obtained at 750 ppm Cr(IV), 13 mA/cm^2 and a pH of 7. [8]

Work by Umran Tezcan Un et al studied the operational factors on electrocoagulation using an iron electrode in a stirred batch reactor. The study concluded that the highest performance was obtained a pH of 2.4, 0.05 M NaCl electrolyte, and a current density of 20 mA/cm^2 for a duration of 20 minutes at an energy cost of 2.68 kWh/m^3. The initial Cr(VI) concentration of 1000 mg/L was almost fully reduced achieving the EPA guideline of 2.77 mg/L in a single step. [9]

Ehssan Nassef and Doaa Elsayed studied the effects of pH, NaCl concentration, initial chromium concentration, Current density, and treatment time on the removal of hexavalent chromium, iron consumption, and energy consumption. The experimental results were fitted to a second-order polynomial using multiple regressions. The study concluded all parameters have a significant effect on chromium removal, and negligible effect on iron consumption. Energy consumption was only affected by current density. [10]

The influence of the type of electrode was investigated by Aji Prasetyaningrum et al. The group worked on 3 types of electrodes, (aluminum, stainless and a combination of both). The study concluded that aluminum was the best performing electrode by achieving 26% reduction of Cr(VI) in industrial waste water generated from plating industry. Stainless steel and the combination electrode showed good results in the first hour of the electrocoagulation process, but their performance deteriorates afterwards. [11]

Vishakha Gilhotra et al studied the treatment of high strength chrome bathwater. The experiments were first conducted on simulated wastewater to optimize process parameters, namely pH, current density and treatment duration. 97.5% removal was obtained at the optimum conditions (pH of 5, 6.8 mA/cm^2, and 17 minutes treatment time). It was also observed that beyond 4g/l, NaCl concentration had negligible effect on the chromium removal efficiency. When the same conditions were applied to real wastewater, negligible removal was obtained due to the high Cr(IV) concentration. Two approaches were followed to increase the removal efficiency of the real scenario. i) Dilution of chromium bathwater, ii) chemical precipitation before EC. The treatment time needed for the real scenario was much higher than that in the simulated experiment. [12]

Electrocoagulation is a very complex phenomenon. There has been a lot of work done in recent years aiming to establish a statistical model to predict the outcome of the electrocoagulation process given a set of inputs. Most models were developed considering 3 factors at 3 levels. pH, current density, and duration are the most studied parameters throughout the literature. It is very difficult to do cross comparison between the different models since the factors and levels are not the same among the different studies, and the variation in other parameters that are not considered part of the model (for example, electrode distancing, initial concentration).

The experimental setup of previous works as well as the studied parameters are summarized in Tables 1 and 2.

Table 1. *Summary of Experimental Setups.*

No.	Anode Material	Cathode Material	Reactor Dimensions	Configuration	Electrode Surface Area	Electrode Spacing	Reference
1 (A)	Stainless Steel	Stainless Steel	1.5 L volume	2 Electrodes	Not mentioned	Variable 1-5 cm	[1]
2	Aluminum	Stainless Steel	17 x 12 x 14 cm^3	2 Electrodes (Mesh Cathode)	96 cm^2	Variable 1-2 cm	[6]
3	Copper	Copper	Not mentioned	2 Electrodes	72.6 cm^2	1.4 cm	[5]
4	Aluminum	Aluminum	20 x 15 x 25 cm^3	4 Electrodes	100 cm^2	2 cm	[7]
5	Iron	Iron	Cylindrical (ID = 15 cm; H = 20 cm)	2 Electrodes	60 cm^2	4 cm	[8]
7 (B)	Iron	Iron	2 L volume	4 rods with a7 hex nuts	1427 cm^2	0.87 cm	[10]
6 (C)	Stainless Steel	Stainless Steel	Not mentioned	2 Electrodes	Not mentioned	1.5 cm	[12]

Table 2. *Summary of Studied Parameters.*

No.	pH	Current Density	Electrode Distance	NaCl Concentration	Initial Chromium Concentration	Time	Reference
1 (A)	×	×	×			×	[1]
2		×	×	×		×	[6]
3	×	×			×	×	[5]
4	×	×				×	[7]
5	×	×			×		[8]
6 (B)	×	×		×	×	×	[10]
7 (C)	×	×				×	[12]

3. Methods

Predicted values from each of the models will be compared to observed results from the different experimental data. Actual vs Predicted removal percentages will be plotted against each dataset for better visualization of the error margins. The coefficient of determination R^2 will be used to judge a model's fitness.

For Ease of reference, the models will be given code names A, B and C. The same coding applies to the experimental data that were used to develop these models. (Refer to Table 1 for experimental setup).

Model Formulas were reproduced using Design Expert 13 using quadratic regression models with 2-factor interactions or linear regression as reported in the corresponding studies.

3.1. Assessment of Existing Models

3.1.1. Model (A)

Model (A) is a second order polynomial with 2 factor interaction parameters. The R^2 value was 0.9793 against observed data (A). When tested on other studies, the model performs poorly with R^2 values of 0.0048 and 0.3395 when tested against Data (B) and (C) respectively. For data (B), there were extreme errors in the model prediction, aside from the large error values, there were also invalid predictions of negative removal percents with the highest being -180%. For Data (C), the model fails to make correct predictions, all the predictions were greater than the observed values. It is worth noting that the values for current density are below the lower limit of model (A). This can partially explain the deviation between prediction and observation but there could be other factors as well. Figures 1, 2 and 3 show the results of Model (A) predictions against data (A), (B) and (C) respectively.

When tested on other studies, the model performs poorly. Upon further analysis, it can be observed that Model (A) is not logical as it shows counter intuitive trends. For example, removal percentage decreases as current intensity or electrolysis time increases. This is seen more predominantly when testing Model (A) with Data (C).

A conclusion to the above is that Model (A) was clearly overfitted for the initial data, as such, it fails to predict results when given different inputs than used in the original study.

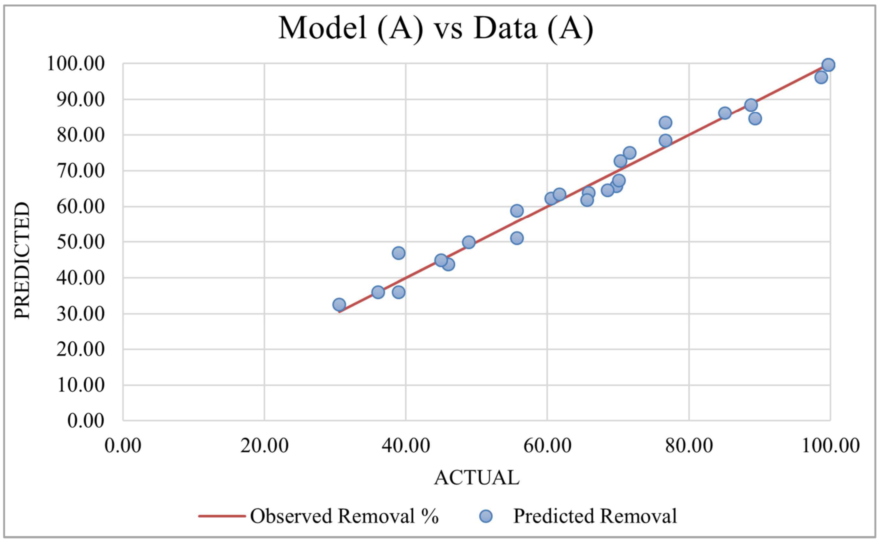

Figure 1. *Model (A) prediction vs Data (A) observation.*

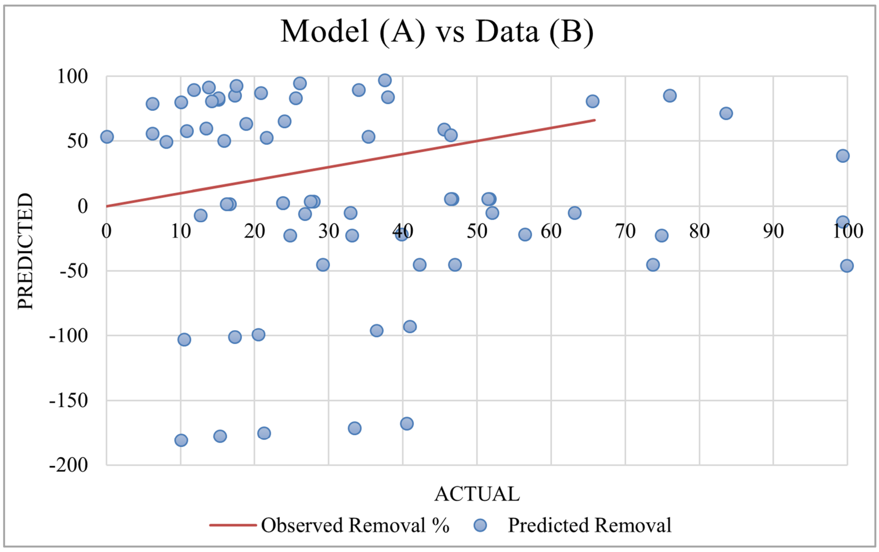

Figure 2. *Model (A) prediction vs Data (B) observation.*

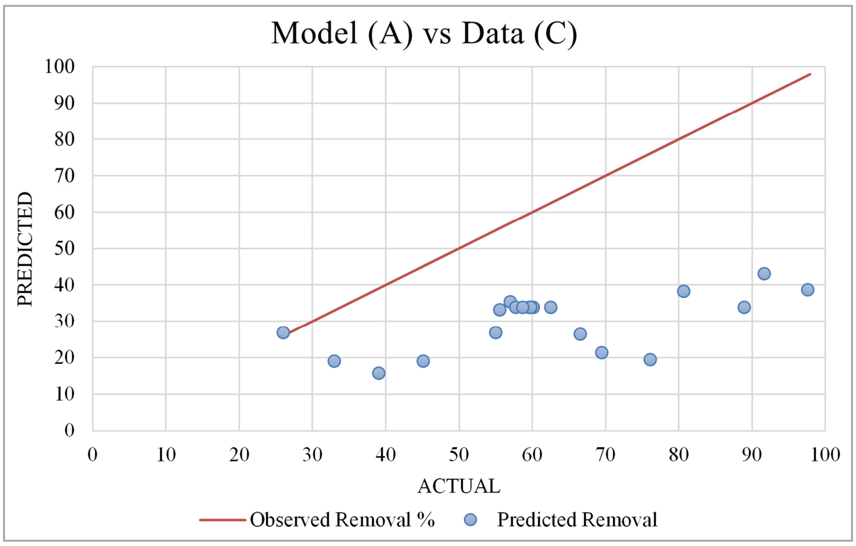

Figure 3. *Model (A) prediction vs Data (C) observation.*

3.1.2. Model (B)

Model B is a linear regression model. The results for Model (B) show a relatively bad fit when compared to observed data (B) with R^2 = 0.8282. Model (B) performs even worse when tested against data (C) with R^2 value of 0.6513. While the model describes the trend somewhat correctly, the prediction values are still off by a significant margin. It is worth noting that the data used in developing model (B) studied the effect of current density from 17.5 to 70 mA/cm^2. For data (C) all values are below 7 mA/cm^2.

Since Model (B) does not consider the electrode distance; it is not possible to test it against Data (A). Therefore, the tests will only include Data (B), Data (C). Figures 4 and 6 show the results of Model (B) predictions against data (B) and (C) respectively.

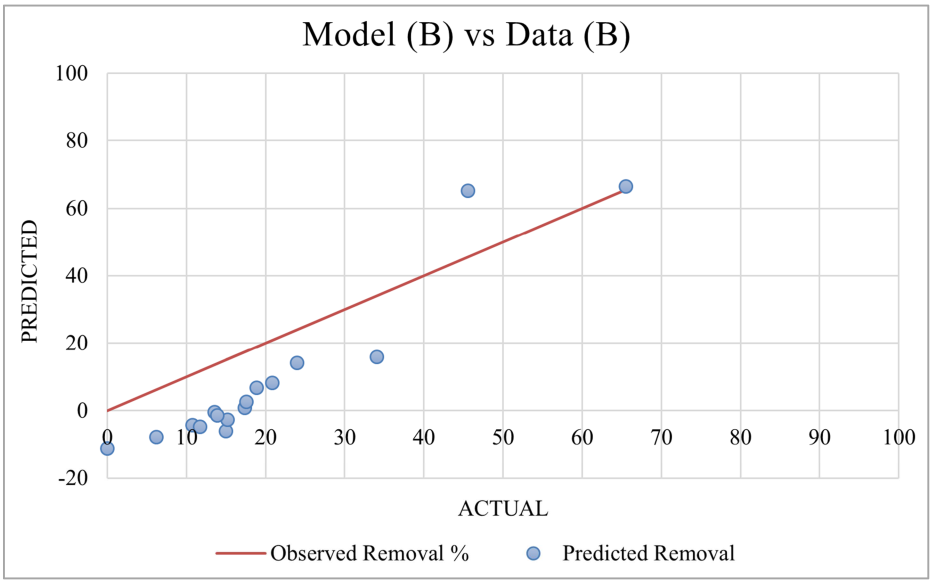

Figure 4. *Model (B) prediction vs Data (B) observation.*

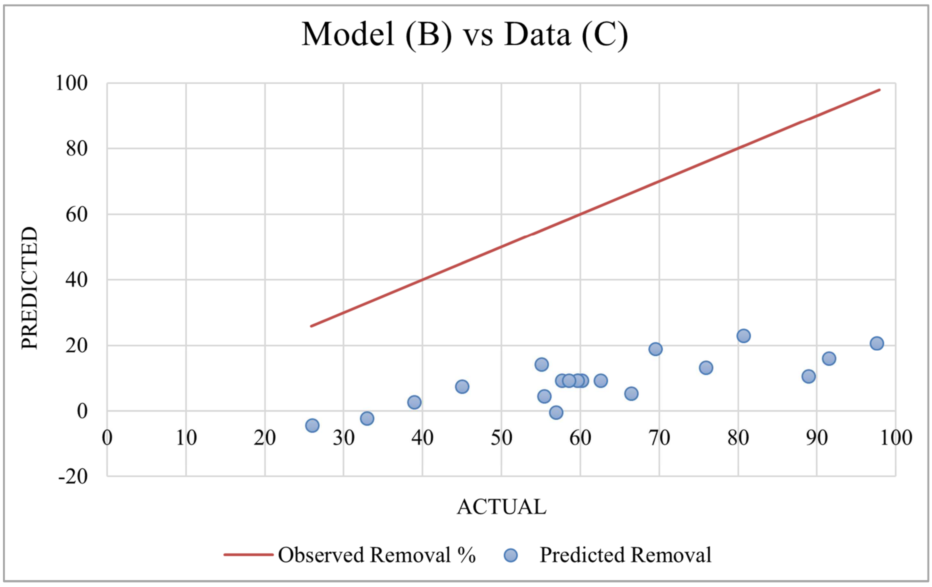

Figure 5. *Model (B) prediction vs Data (C) observation.*

3.1.3. Model (C)

While Model (C) has a very good fit with R^2 = 0.9952 against data (C), it completely fails to make logical predictions with data (B), prediction values are illogical as it exceeds 100% with the lowest being 120% and the highest being 3100%. The poor performance of Model (C) can be attributed to differences in experimental setup and domain as explained earlier. It appears that model (C) is incapable of making predictions for current densities beyond its domain.

Model (C) does not consider the electrode distance; it is not possible to test it against Data (A). Therefore, the tests will only include Data (B) and Data (C). Figures 6 and 7 show the results of Model (C) predictions against data (C) and (B) respectively.

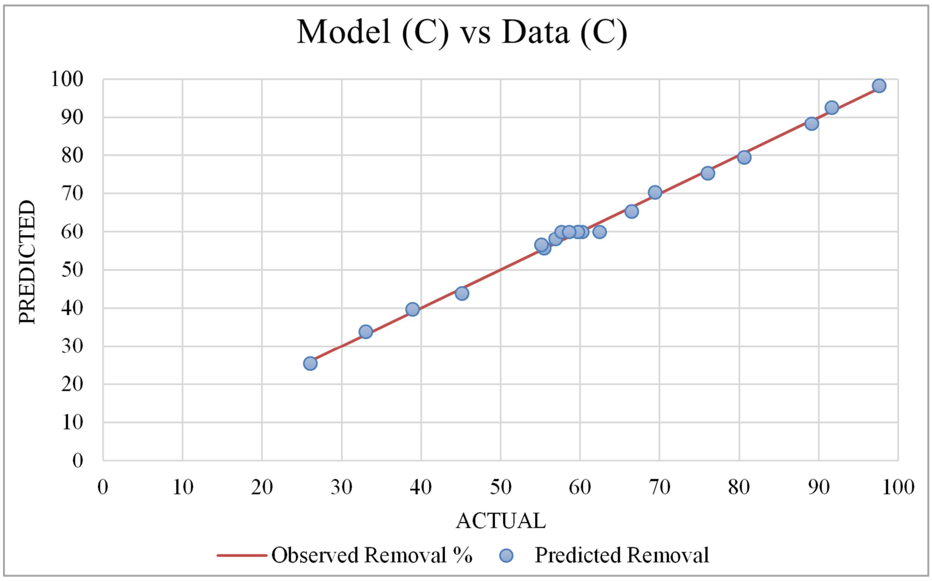

Figure 6. *Model (C) prediction vs Data (C) observation.*

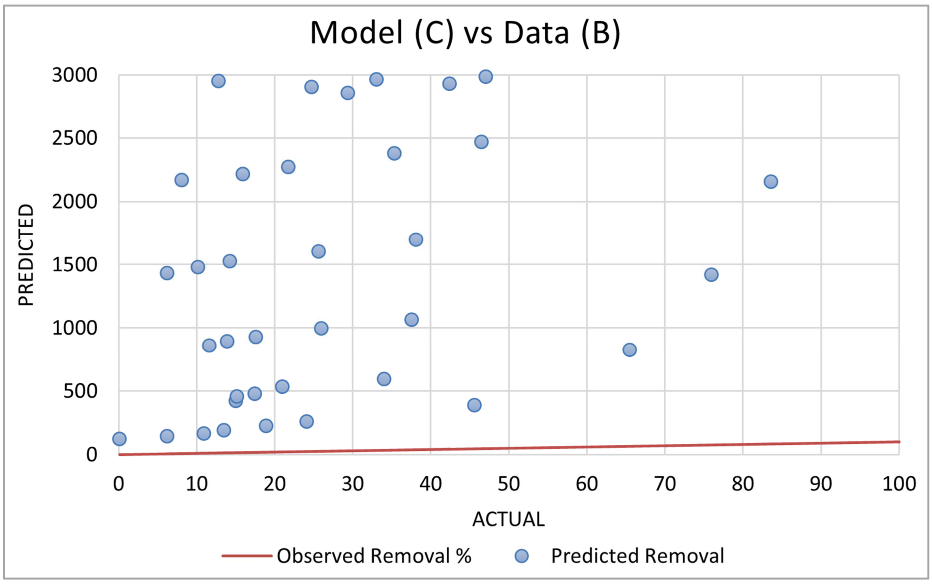

Figure 7. *Model (C) prediction vs Data (B) observation.*

The results of cross validation of the 3 models (A, B and C) indicate that these models are localized, and fail to make predictions outside the experimental domain, and hence cannot be used for real world applications. Only model E shows better results compared to the rest of the data, although its accuracy can be significantly improved if the model had a larger training set. The R^2 values for the previous plots are summarized in the Table 3 below:

Table 3. *Comparison of R^2 values for Models (A), (B) and (C).*

	Model (A)	Model (B)	Model (C)
Data (A)	0.9793	NA	NA
Data (B)	0.0048	0.8282	0.1252
Data (C)	0.3395	0.6513	0.9952

3.2. Development of the SVR Model - Model (D)

The new model presented in this study, Model (D) utilizes Machine Learning techniques, namely Support Vector Regression (SVR) analysis. The model was developed using python version 3.9.6 and the sci-kit learn (sklearn) library version 1.1.3.

SVR is a supervised learning technique, where the machine learns the regression function that maps the input variables to the output observed response value while minimizing the prediction error. The regression function is defined by a subset of support vectors that are crucial in determining the optimal hyperplane (or epsilon-insensitive tube) to fit the regression line. [13, 14]

One advantage of using an ML approach to modelling is that it aims to maximize the predictive accuracy even when the relationship between the inputs and outputs is unknown or of very complex nature, since an ML model does not need to assume a simple functional form. Which makes it very effective in modelling complex phenomena and engineering applications. [15]

SVR models are known to outperform linear, polynomial or logistical models in terms of consistency

The dataset used in developing the model was compiled from the previous experimental works that were used to develop models (A), (B) and (C).

The compiled dataset was split into a training set and a test set using a 50-50 split ratio. This allocates an equal number of data points to each set.

The training set was used to build the SVR model by mapping the input variables to the corresponding target variable. The test set was used to assess the performance and generalization ability of the SVR model. By predicting the target variable for the test data and comparing them to the actual values, the accuracy and predictive power of the model could be evaluated. This evaluation provides insights into how well the model performs on unseen data.

Model (D) considers the following parameters:

1. pH
2. Current Density
3. Electrode Distance
4. Duration

Parameters not considered in the model:

1. Initial chromium concentration
2. NaCl concentration
3. Electrode type

4. Results and Discussion

The SVR model, model (D), shows a much better generalization ability compared to previous models with R^2 equals to 0.7447 against the test data. It still can't provide good accuracy with some data points, but it can describe the trend much better than the other models. The lack of accuracy can be attributed to the size of the dataset which included only 116 points, for ML models in general, more data are required for developing high fidelity models.

The compiled dataset is shown in the Table 4 below:

Table 4. *Compiled dataset used for training and testing the SVR model.*

No.	pH	Current Density (mA/cm^2)	Electrode Distance (mm)	Time (Minute)	Removal	Data Source
1	8	40	30	30	55.70%	
2	2	40	30	30	48.90%	
3	5	40	30	10	76.60%	
4	2	25	50	30	55.60%	
5	5	25	30	30	99.70%	
6	5	40	30	50	60.60%	
7	5	25	30	30	99.70%	
8	5	25	10	50	70.40%	
9	5	10	50	30	88.70%	
10	2	10	30	30	30.60%	
11	5	40	10	30	98.70%	
12	5	10	30	50	65.70%	
13	5	40	50	30	46.00%	
14	5	25	30	30	99.70%	
15	8	10	30	30	65.60%	Data (A) [1]
16	5	25	50	50	69.70%	
17	8	25	30	50	61.70%	
18	8	25	10	30	76.60%	
19	8	25	50	30	39.00%	
20	5	10	10	30	44.90%	
21	5	25	10	10	85.00%	
22	5	25	30	30	99.70%	
23	2	25	10	30	39.00%	
24	5	10	30	10	71.60%	
25	8	25	30	10	70.10%	
26	5	25	50	10	89.30%	
27	2	25	30	10	68.60%	
28	2	25	30	50	36.00%	
29	5	25	30	30	99.70%	
30	4.66	17.5088	8.7	0	0.00%	
31	4.66	17.5088	8.7	2.0142	6.10%	
32	4.66	17.5088	8.7	3.91	10.80%	
33	4.66	17.5088	8.7	5.9242	13.50%	
34	4.66	17.5088	8.7	10.0711	18.80%	Data (B) [10]
35	4.66	17.5088	8.7	14.0403	24.00%	
36	4.66	17.5088	8.7	42.0616	45.60%	
37	4.66	28.0141	8.7	2.07346	15.00%	
38	4.66	28.0141	8.7	4.02844	15.20%	

No.	pH	Current Density (mA/cm^2)	Electrode Distance (mm)	Time (Minute)	Removal	Data Source
39	4.66	28.0141	8.7	5.98341	17.40%	
40	4.66	28.0141	8.7	9.95261	20.90%	
41	4.66	28.0141	8.7	14.0995	34.00%	
42	4.66	28.0141	8.7	42.0024	65.50%	
43	4.66	38.5193	8.7	2.07346	11.70%	
44	4.66	38.5193	8.7	4.08768	13.90%	
45	4.66	38.5193	8.7	6.04265	17.50%	
46	4.66	38.5193	8.7	9.95261	26.00%	
47	4.66	38.5193	8.7	14.0403	37.50%	
48	4.66	38.5193	8.7	42.1209	76.00%	
49	4.66	49.0246	8.7	1.95498	6.20%	
50	4.66	49.0246	8.7	3.90995	10.20%	
51	4.66	49.0246	8.7	5.98341	14.20%	
52	4.66	49.0246	8.7	9.89336	25.50%	
53	4.66	49.0246	8.7	13.9218	38.00%	
54	4.66	49.0246	8.7	42.0024	83.50%	
55	4.66	59.5299	8.7	2.07346	8.00%	
56	4.66	59.5299	8.7	3.90995	15.90%	
57	4.66	59.5299	8.7	5.98341	21.70%	
58	4.66	59.5299	8.7	9.89336	35.40%	
59	4.66	59.5299	8.7	13.981	46.50%	
60	4.66	59.5299	8.7	42.0024	99.30%	
61	4.66	70.0352	8.7	2.1327	16.50%	
62	4.66	70.0352	8.7	3.96919	23.90%	
63	4.66	70.0352	8.7	6.1019	27.90%	
64	4.66	70.0352	8.7	9.95261	46.70%	
65	4.66	70.0352	8.7	14.0403	51.70%	
66	4.66	70.0352	8.7	42	99.30%	
67	11.8	70.0352	8.7	1.99766	10.10%	
68	11.8	70.0352	8.7	3.9778	15.30%	
69	11.8	70.0352	8.7	5.99299	21.20%	
70	11.8	70.0352	8.7	10.0234	33.40%	
71	11.8	70.0352	8.7	14.0187	40.50%	
72	10.08	70.0352	8.7	1.98014	10.50%	
73	10.08	70.0352	8.7	3.96028	17.30%	
74	10.08	70.0352	8.7	5.97547	20.60%	
75	10.08	70.0352	8.7	9.97079	36.50%	
76	10.08	70.0352	8.7	13.9661	41.00%	
77	4.665	70.0352	8.7	2.05023	16.30%	
78	4.665	70.0352	8.7	4.0479	23.90%	
79	4.665	70.0352	8.7	6.04556	27.60%	
80	4.665	70.0352	8.7	10.0409	46.40%	
81	4.665	70.0352	8.7	14.0362	51.50%	
82	3.1	70.0352	8.7	1.99766	12.70%	
83	3.1	70.0352	8.7	3.99533	26.80%	
84	3.1	70.0352	8.7	5.97547	32.90%	
85	3.1	70.0352	8.7	9.98832	52.00%	
86	3.1	70.0352	8.7	14.0362	63.20%	
87	1.99	70.0352	8.7	1.98014	24.70%	
88	1.99	70.0352	8.7	3.96028	33.10%	
89	1.99	70.0352	8.7	5.97547	39.80%	
90	1.99	70.0352	8.7	9.98832	56.40%	
91	1.99	70.0352	8.7	13.9661	74.80%	
92	1	70.0352	8.7	1.98014	29.30%	
93	1	70.0352	8.7	3.9778	42.30%	
94	1	70.0352	8.7	5.99299	47.10%	
95	1	70.0352	8.7	9.98832	73.90%	
96	1	70.0352	8.7	14.0012	100.00%	
97	6.8	3.21	20	8	33.00%	
98	4.7	6.78	20	8	55.50%	
99	4.7	3.21	20	8	39.00%	
100	6.8	6.78	20	17	91.60%	
101	4.7	6.78	20	17	97.50%	Data (C) [12]
102	6.8	3.21	20	17	55.00%	
103	4.7	3.21	20	17	69.50%	
104	6.8	6.78	20	8	57.00%	
105	7.5	5	20	12.5	66.50%	

No.	pH	Current Density (mA/cm^2)	Electrode Distance (mm)	Time (Minute)	Removal	Data Source
106	5.8	5	20	5	26.00%	
107	5.8	5	20	20	80.70%	
108	4	5	20	12.5	76.00%	
109	5.8	8	20	12.5	89.00%	
110	5.8	2	20	12.5	45.10%	
111	5.8	5	20	12.5	57.70%	
112	5.8	5	20	12.5	62.50%	
113	5.8	5	20	12.5	59.60%	
114	5.8	5	20	12.5	60.20%	
115	5.8	5	20	12.5	59.70%	
116	5.8	5	20	12.5	58.60%	

The model performance against the test set is shown in Figure 8 below:

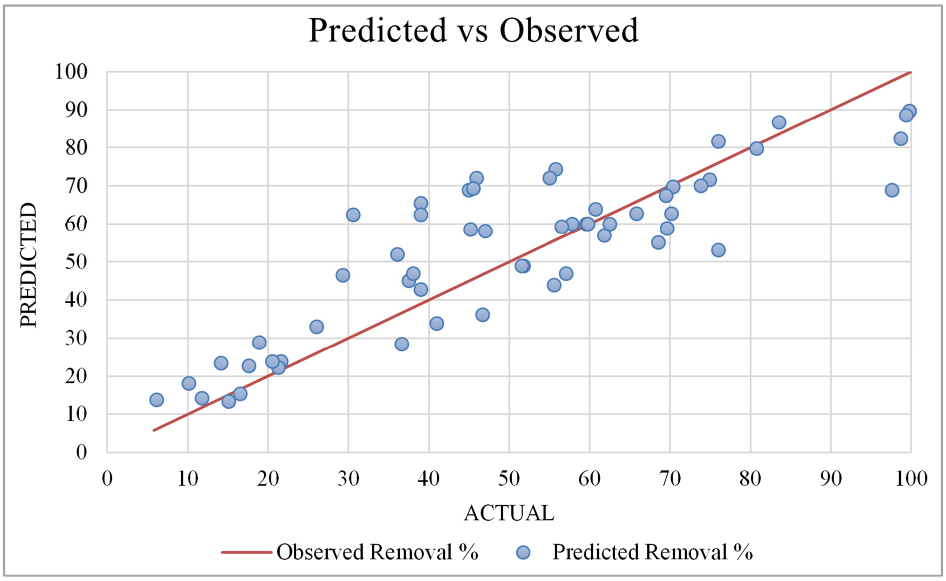

Figure 8. *Model (D) prediction vs Test Data.*

5. Conclusion

After evaluating studies published between 2015 and 2021 on the removal of Cr(VI) using electrocoagulation, it can be concluded that none of the statistical models was entirely conclusive. All models had similar problems such as overfitting, localized applicability, and inability to generalize beyond the experimental domain.

Model (A) was found to consider the most parameters, but it still falls short when evaluated against different data due to its localized nature. Furthermore, Models (C) and (B) both have significant flaws. These models produce illogical results and perform very poorly when tested with different data.

Model (D) showed much better generalization performance, but overall, it was able to approximate data from all the other experiments. The accuracy of the model while can be further improved if given a larger dataset.

It is recommended that future research focuses on the development of more robust and reliable statistical models that can capture the complexity of the electrocoagulation process and generalization beyond the experimental domain. To achieve this, researchers should reconsider the model type as quadratic regression might not be the best mode, even if it produces high R^2 value as this can be a result of overfitting and not necessarily an indication of a good model. Additionally, a more comprehensive experimental design that considers a broad range of variables should be utilized to increase the validity and reliability of statistical models and avail more data for training ML models.

This review highlights the need for further investigation and development of accurate and reliable statistical models for the removal of Cr(VI) using electrocoagulation. Careful consideration should be given to the applicability of such models in different scenarios to avoid over-reliance and misinterpretation of results.

List of Abbreviations

ANN	Artificial Neural Networks
ANOVA	Analysis of Variance
CD	Current Density
EC	Electrocoagulation
EPA	Environmental Protection Agency
ML	Machine Learning
RO	Reverse Osmosis
RSM	Response Surface Methodology
SVR	Support Vector Machines Regression

ORCID

Mohamad Salem: 0009-0004-3777-9094

Declarations

Availability of Data and Material

The datasets supporting the conclusions of this article are included within the article.

Authors' Contributions

Mohamad Salem wrote up the manuscript and conducted the statistical analysis. Prof. N. Abdelmonem and Ehssan Nassef supervised analysis and conclusion, edited, read, and approved the final manuscript. All authors read and approved the final manuscript.

Authors' Information

Prof. Nabil Mahmoud Abdelmonem is currently a professor of Chemical Engineering Department at Faculty of Engineering, Cairo University. Prof. Ehssan Nassef is a Professor of Petrochemical Engineering at Faculty of Engineering, Pharos University. Mohamad Salem is currently a Process Engineer at ENPPI (Engineering for Petroleum and Process Industries).

Acknowledgments

The authors acknowledge Cairo University, for providing support and for funding the publication of the paper.

Conflicts of Interest

The authors declare that they have no competing interests.

References

[1] K. Thirugnanasambandham and K. Shine, "Investigation on the Removal of Chromium from Wastewater using Electrocoagulation," *International Journal of Chemical Reactor Engineering,* vol. 16, 2018.

[2] Xiaokun Wang, Yingqin Wei, Shasha Wang, Lingxin Chen, "Red-to-blue colorimetric detection of chromium via Cr (III)-citrate chelating based on Tween 20-stabilized gold nanoparticles," *Colloids and Surfaces A: Physicochemical and Engineering Aspects,* pp. 57-62, 2015.

[3] S. Dhiraj, M. Garima and K. M. P., "Agricultural waste material as potential adsorbent for sequestering heavy metal ions from aqueous solutions – A review," *Bioresource Technology,* vol. 99, no. 14, pp. 6017-6027, 2008.

[4] H. Jean, G. Bouchaib, C. Mohammed, S. Youssef, V. Christophe, D. Patrick and N. Jamal, "Electrocoagulation process in water treatment: A review of electrocoagulation modeling approaches," *Desalination,* no. 404, pp. 1-21, 2017.

[5] P. R. Sunil and S. P. Parikh, "Statistical optimizing of electrocoagulation process for the removal of Cr(VI) using response surface methodology and kinetic study," *Arabian Journal of Chemistry,* vol. 13, no. 9, pp. 7032-7044, 2020.

[6] R. H. Salman, H. A. Hassan, K. M. Abed, A. F. Al-Alawy, D. A. Tuama, K. M. Hussein and H. A. Jabir, "Removal of chromium ions from a real wastewater of leather industry using electrocoagulation and reverse osmosis processes," in *AIP Conference Proceedings*, 2020.

[7] E. Aguilar-Ascón, L. Marrufo-Saldaña and W. Neyra-Ascón, "Reduction of Total Chromium Levels from Raw Tannery Wastewater via Electrocoagulation using Response Surface Methodology," *Journal of Ecological Engineering,* vol. 20, no. 11, pp. 217-224, 2019.

[8] N. M. Genawi, M. H. Ibrahim, M. H. El-Naas and A. E. Alshaik, "Chromium Removal from Tannery Wastewater by Electrocoagulation: Optimization and Sludge Characterization," *Water,* vol. 12, no. 5, 2020.

[9] U. Tezcan Un, S. Eroglu Onpeker and E. Ozel, "The treatment of chromium containing wastewater using electrocoagulation and the production of ceramic pigments from the resulting sludge," *Journal of Environmental Management,* vol. 200, pp. 196-203, 2017.

[10] Y. A. El-Taweel, E. M. Nassef, I. Elkheriany and D. Sayed, "Removal of Cr(VI) ions from waste waterby electrocoagulation using iron electrode," *Egyptian Journal of Petroleum,* vol. 24, no. 2, pp. 183-192, 2015.

[11] A. Prasetyaningrum, B. Jos, Y. Dharmawan, B. T. Prabowo, M. Fathurrazan and Fyrouzabadi, "The influence of electrode type on electrocoagulation process for removal of chromium (VI) metal in plating industrial wastewater," *Journal of Physics: Conference Series,* vol. 1025, 2018.

[12] V. Gilhotra, R. Yadav, A. Sugha, L. Das, A. Vashisht, R. Bhatti and M. S. Bhatti, "Electrochemical treatment of high strength chrome bathwater: A comparative study for best-operating conditions," *Cleaner Engineering and Technology,* vol. 2, 2021.

[13] F. Zhang and L. J. O'Donnell, "Chapter 7 - Support vector regression," in *Machine Learning*, Academic Press, 2020, pp. 123-140.

[14] X. Nong, C. Lai, L. Chen, D. Shao and C. Zhang, "Prediction modelling framework comparative analysis of dissolved oxygen concentration variations using support vector regression coupled with multiple feature engineering and optimization methods: A case study in China," *Ecological Indicators,* vol. 146, 2023.

[15] J. I. Glaser, A. S. Benjamin, R. Farhoodi and K. P. Kording, "The roles of supervised machine learning in systems neuroscience," *Progress in Neurobiology,* vol. 175, pp. 126-137, 2019.