water

MDPI

Article

Modeling and Optimizing of NH_4^+ Removal from Stormwater by Coal-Based Granular Activated Carbon Using RSM and ANN Coupled with GA

Aixin Yu, Yuankun Liu *, Xing Li, Yanling Yang, Zhiwei Zhou and Hongrun Liu

Faculty of Urban Construction, Beijing University of Technology, Beijing 100124, China; yuaixin809@163.com (A.Y.); lixing@bjut.edu.cn (X.L.); yangyanling@bjut.edu.cn (Y.Y.); hubeizhouzhiwei@163.com (Z.Z.); liuhongrun7768@163.com (H.L.)

* Correspondence: liuyuankun@bjut.edu.cn; Tel.: +86-152-1097-9027

Abstract: As a key parameter in the adsorption process, removal rate is not available under most operating conditions due to the time and cost of experimental testing. To address this issue, evaluation of the efficiency of NH_4^+ removal from stormwater by coal-based granular activated carbon (CB-GAC), a novel approach, the response surface methodology (RSM), back-propagation artificial neural network (BP-ANN) coupled with genetic algorithm (GA), has been applied in this research. The sorption process was modeled based on Box-Behnben design (BBD) RSM method for independent variables: Contact time, initial concentration, temperature, and pH; suggesting a quadratic polynomial model with p-value < 0.001, R^2 = 0.9762. The BP-ANN with a structure of 4-8-1 gave the best performance. Compared with the BBD-RSM model, the BP-ANN model indicated better prediction of the response with R^2 = 0.9959. The weights derived from BP-ANN was further analyzed by Garson equation, and the results showed that the order of the variables' effectiveness is as follow: Contact time (31.23%) > pH (24.68%) > temperature (22.93%) > initial concentration (21.16%). The process parameters were optimized via RSM optimization tools and GA. The results of validation experiments showed that the optimization results of GA-ANN are more accurate than BBD-RSM, with contact time = 899.41 min, initial concentration = 17.35 mg/L, temperature = 15 °C, pH = 6.98, NH_4^+ removal rate = 63.74%, and relative error = 0.87%. Furthermore, the CB-GAC has been characterized by Scanning electron microscopy (SEM), X-ray diffraction (XRD) and Brunauer-Emmett-Teller (BET). The isotherm and kinetic studies of the adsorption process illustrated that adsorption of NH_4^+ onto CB-GAC corresponded Langmuir isotherm and pseudo-second-order kinetic models. The calculated maximum adsorption capacity was 0.2821 mg/g.

Keywords: response surface methodology (RSM); back-propagation artificial neural network (BP-ANN); genetic algorithm (GA); coal-based granular activated carbon (CB-GAC); stormwater

Citation: Yu, A.; Liu, Y.; Li, X.; Yang, Y.; Zhou, Z.; Liu, H. Modeling and Optimizing of NH_4^+ Removal from Stormwater by Coal-Based Granular Activated Carbon Using RSM and ANN Coupled with GA. *Water* **2021**, *13*, 608. https://doi.org/10.3390/w13050608

Academic Editor: Margaritis Kostoglou

Received: 6 February 2021
Accepted: 23 February 2021
Published: 26 February 2021

1. Introduction

In recent years, with the development of intensive urbanization, population growth, land use transformation, pollution, and changing climate patterns, fresh water availability has become one of the main issues facing humankind [1]. In this context, stormwater is being considered in water-stressed cities as an alternative water resource [2]. Stormwater utilization is also an ecological and sustainable method of water management, resulting in the reduction of urban runoff and flooding [3].

Despite its potential for augmenting water supplies and reducing runoff flood, urban stormwater contains contaminants that pose risks to human health and aquatic ecosystems. Stormwater is often contaminated by organic matter, nutrients (i.e., phosphorus and nitrogen) and heavy metals [4,5] (Some studies suggest it should be replaced with the term potentially toxic elements [6]). Dissolved nitrogen species include nitrate (NO_3^-), nitrite

(NO_2^-), ammonia and ammonium (NH_3 and NH_4^+), and organic N. The distribution between NH_3 versus NH_4^+ is pH dependent (pK_a = 9.25), but NH_4^+ dominates in stormwater (typical pH values are 6–8) [7]. The high level of ammonia in water imparts taste and odor problems, can cause symptoms in aquatic organisms such as hypoxia, coma, and decreased immunity, resulting in slow growth and even a large numbers of deaths [8,9].

Over the years, various techniques have been used to remove ammonia nitrogen from water, such as biological process [10], stripping [11], breakpoint chlorination [12], chemical precipitation [13], and adsorption [14]. As the most commonly used nitrogen removal technology, however, biological processes are critical to the environment. Low temperature or high ammonia concentration may affect the activity of nitrobacteria and reduce the treatment efficiency [10]. Nitrifying bacteria activity is reduced below 15 °C, and life activities are inhibited below 5 °C [14]. Therefore, adsorption is considered to be a feasible method to remove ammonia nitrogen from water, since adsorption is a more direct, stable, and low-cost method for contaminants removal from water. Ideally, sorbents for contaminants in stormwater should exhibit high selectivity and rapid sorption kinetics, as well as sufficient sorption capacity [15]. Activated carbon is a kind of cheap, large specific surface area, easy to obtain, and rich pore structure adsorption material, which is broadly used in the water treatment process [16], but currently has limited application for other types of contaminated water, including stormwater [17]. Coal-based granular activated carbon is the most productive of all kinds of activated carbons due to its wide source of raw materials, mature production process, and good adsorption effect, and 70% of it is used for water treatment [18].

Removal rate is a key parameter in the adsorption process, but it is not available under most operating conditions due to the time and cost of experimental testing. Therefore, it is important to optimize experimental conditions and obtain the maximum removal rate by using modeling and optimization. The adsorption process is a complex process, which is difficult to describe by statistical model due to the complexity of the relationship between input parameters and output. Computational intelligence models are generally more flexible than statistical models in modeling complex data sets that may contain nonlinear or missing data [19]. In order to reveal the influence of the interaction between variables in the adsorption process on the dependent variables, analyze the relative influence degree of different variables, and obtain the optimal conditions of the adsorption process, more reasonable and accurate methods need to be developed.

Response surface methodology (RSM) is widely used in the optimization of chemical and physical processes. It is an optimization method that integrates mathematical modeling and experimental design [20]. Through experiments on representative local points, the functional relationship between factors in the global scope and experimental results is fitted by regression, and the optimal value of each factor is obtained. The difference between the response surface method and the traditional orthogonal experimental design method is that the response surface method has high precision of regression equation, short experimental period, and the ability to study the interaction between multiple factors at the same time [21].

In addition, a number of data analysis tools have evolved into sophisticated modeling techniques, such as fuzzy logic (FL) and artificial intelligence (AI). As one of the main tools of artificial intelligence, artificial neural networks (ANNs) have been widely used due to their general ability to simulate nonlinear changes and process incomplete data. ANN does not require mathematical description of the phenomena in the process, so the simulation of the complicated systems could be performed more efficiently [22–24]. Compared with the RSM focusing on the statistical importance of the individual process variables and their interactions via ANOVA, AAN is more reliable in capturing the nonlinear relationship between the removal rate and process variables [25]. Although, ANN is an effective tool to predict and optimize any complex process parameters, it cannot guarantee the global optimal solution [26]. Genetic algorithm (GA) can be used to minimize (or maximize) the global search and to obtain feasible solutions with a known fitness function [19]. GA is a

kind of adaptive heuristic optimization search algorithm based on a Darwinian genetic evolution principle, taking "survival of the fittest" as the concept, and using genetic operators such as selection, mutation, and crossover to find the optimal solution of the problem [27]. Therefore, the combination of artificial neural network and genetic algorithm can generate the optimal operating variables of the process under study.

This study was aimed at modeling and optimizing the NH_4^+ removal process from stormwater by coal-based granular activated carbon (CB-GAC). RSM experimental design was selected to establish the regression model. The model involved four parameters (i.e., contact time, initial concentration, temperature, and pH) as independent variables and NH_4^+ removal rate as dependent variable. The interaction effect of the independent variables with the response using the response surface plots was illustrated. Meanwhile, the back-propagation artificial neural network BP-ANN model was also developed and the optimal number of hidden neurons was determined by trial and error method. Coupled with GA, the ANN-GA model was used for optimization of the operating conditions to determine the maximum NH_4^+ removal rate. The predictive capabilities and modelling efficiencies of the two models are compared and verified. Furthermore, the CB-GAC has been characterized by scanning electron microscopy (SEM), X-ray diffraction (XRD), and Brunauer-Emmett-Teller (BET). Finally, adsorption isotherm model (Langmuir, Freundlich, and Temkin) and adsorption kinetics model (pseudo-first-order, pseudo-second-order) had been investigated to explore the mechanisms of the NH_4^+ removal process.

2. Materials and Methods

2.1. Materials and Characterization

All the chemicals (NH_4Cl, NaOH, HCl) used were analytical grade and were purchased from Beijing Yili Fine Chemicals Co., Ltd. (Beijing, China). The coal based granular activated carbon (CB-GAC) is columnar and has a diameter of 2–3 mm, washed three times with deionized water, then drying at 105 °C for 24 h before use. All the solutions were prepared by the deionized water.

Scanning electron microscopy (SEM) (HITACHI, SU8020, Tokyo, Japan) was performed to study the surface characteristics and morphology of the coal based activated carbon. X-ray diffraction (XRD) was carried out on a Bruker Analytical X-ray (D8 Advance, Bruker, Germany) to indicate crystallinity and phases of the CB-GAC. Surface area, pore volume, and pore size distribution were determined through the adsorption/desorption isotherms of N_2 at 77 K in a Surface Area and Pore Size Analyzer (Micromeritics Instruments, ASAP 2460, Norcross, GA, USA).

2.2. Batch Adsorption Experiments

The adsorption experiments were conducted by adding 10 g of the CB-GAC to 100 mL ammonium chloride solution in a 250 mL conical flask, which was shaken in a constant-temperature shaker (Guohua Electric Appliance Co. Ltd., THZ-82, Changzhou, China) at 150 rpm for 24 h. Batch experiments were carried out to examine the effect of contact time (300–900 min), initial concentration (10–20 mg/L), temperature (15–35 °C), and pH (5–9) on the removal rate of NH_4^+. The initial pH was adjusted to the desired value by using 0.1 mol/L HCl or 0.1 mol/L NaOH. After reaching the designed time, an aliquot of the sample was withdrawn and filtered through 0.45 μm poly tetra fluoroethylene (PTFE) filter (Anpel Co. Ltd., Shanghai, China) to remove suspended solids. The concentration of NH_4^+ were measured by UV-visible Multi-parameter Analyzer (Lianhua Tech Co. Ltd., LH-3BA, Beijing, China). The removal rate and adsorption amount of NH_4^+ were calculated using the following equations.

$$R = (C_0 - C_t)/C_0 \times 100\% \quad (1)$$

$$q_t = (C_0 - C_t) \times V/m \quad (2)$$

where R, q_t (mg/g) are the removal rate and adsorption amount of NH_4^+, respectively. C_0 (mg/L) and C_t (mg/L) are the concentrations at time 0 and t, respectively. V (L) is the volume of the solution, and m (g) is the mass of adsorbent.

2.3. Isotherm and Kinetics Study

As shown in Table 1, three isothermal models (Langmuir, Freundlich, and Temkin models) and two adsorption kinetic models (pseudo-first order, pseudo-second order) were used to evaluate the adsorption of NH_4^+ onto the CB-GAC. All of this part of the experiments were performed under the following conditions: 10 g of CB-GAC was mixed with 100 mL of ammonium chloride solution (5–30 mg/L), at pH 7.0 and room temperature (25 ± 1 °C).

Table 1. Adsorption isotherm and kinetic models used in this study.

Type	Models	Equations
Isotherm models	Langmuir	$\frac{C_e}{Q_e} = \frac{C_e}{Q_{max}} + \frac{1}{K_L Q_{max}}$
	Freundlich	$lgQ_e = \frac{1}{n} lgC_e + lgK_F$
	Temkin	$Q_e = \frac{RT}{b_T} lnK_T + \frac{RT}{b_T} lnC_e$
Kinetic models	Pseudo-first-order	$\ln(Q_e - Q_t) = \ln Q_e - k_1 t$
	Pseudo-second-order	$\frac{t}{Q_t} = \frac{1}{k_2 \times Q_e^2} + \frac{t}{Q_e}$

Where, C_e (mg/L) = equilibrium concentration; Q_e (mg/g) = equilibrium sorption amount; Q_{max} (mg/g) = maximum adsorption capacity; K_L (L/g) = Langmuir adsorption constant; n, K_F (L/g) = Freundlich adsorption constants; R (8.314 J/mol) = universal gas constant; T = temperature in terms of Kelvin; b_T = Temkin constant; K_T = equilibrium bond constant related to the maximum energy of bond; t = time (min); Q_t (mg/g) = adsorption capacity at time t (min); k_1, k_2 is the pseudo-first-order rate constant and pseudo-second-order rate constants, respectively.

2.4. Experimental Design and Mathematical Models

2.4.1. Response Surface Methodology

RSM is the most efficient way to determine the best combination of experimental conditions, which reduces the research workloads, and provides an appropriate model for preparation technology optimization than the conventional variable control approaches [28,29]. In the present work, the experiment design of Box-Behnben design-response surface methodology (BBD-RSM) was completed by using Design-Expert 12 software, so as to investigate the effect of different variables on NH_4^+ removal rate and obtain an accurate model to predict the removal rate. The Box-Behnken design involves 4 variables and 3 levels. There were 29 runs, including 5 zeros, which usually represented repeated experiments, and were used to estimate experimental errors. Specifically, contact time, initial concentration, temperature, and pH were selected as independent variables, their levels were coded as: −1, 0, +1, and the NH_4^+ removal rate as the dependent variable. The relationship between the dependent and the independent variables can be represented by the Equation (3). The variables and the levels for BBD-RSM used in this study are shown in Table 2.

$$Y = \alpha_0 + \alpha_1 A + \alpha_2 B + \alpha_3 C + \alpha_4 D + \alpha_{12} AB + \alpha_{13} AC + \alpha_{14} AD + \alpha_{23} BC + \alpha_{24} BD + \alpha_{34} CD + \alpha_{11} A^2 + \alpha_{22} B^2 + \alpha_{33} C^2 + \alpha_{44} D^2 \quad (3)$$

where Y represents the removal rate of NH_4^+; α_0 is a constant offset term; α_1, α_2, α_3, α_4, α_{12}, α_{13}, α_{14}, α_{23}, α_{24}, α_{34}, α_{11}, α_{22}, α_{33}, α_{44} are the estimated coefficients, respectively; A, B, C, D are contact time, initial concentration, temperature, and pH, respectively.

Table 2. Experimental ranges and levels of variables.

Factors	Variables	Unit	Level		
			Low (−1)	Middle (0)	High (+1)
A	Contact time	min	300	600	900
B	Initial concentration	mg/L	10	15	20
C	Temperature	°C	15	25	35
D	pH	/	5	7	9

2.4.2. BP-ANN Modeling and Optimization

In the present study, the three-layer back-propagation ANN was trained by back-propagation gradient-descendent algorithm. BP-ANN is composed of an input, output layer, and one or more hidden layers. It has been verified theoretically that a 3-layer BP neural network can approach arbitrary complex mappings [30]. The simple illustration of the ANN structure was shown in Figure 1.

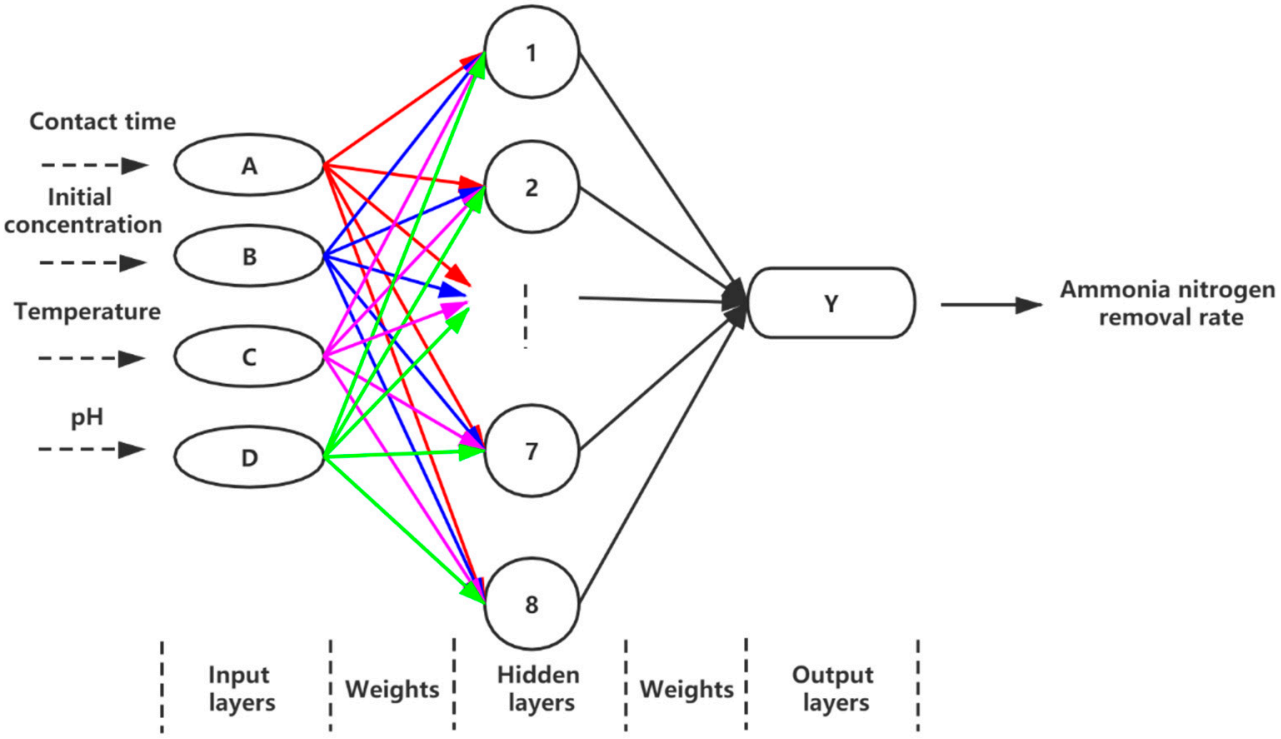

Figure 1. Simple illustration of the artificial neural network (ANN) structure.

In this study, 29 experimental points were used to feed the neural network structure. These sites were divided into three training groups. It contains 16, 4, 4, 5 samples for training, testing, validation and simulation, respectively. Weights and deviations were updated with training data through Levenberge Marquardt algorithm, and the network generalization ability after training was evaluated with test data. In addition, the error of validation data was monitored during training to avoid overfitting [31]. The network used in this study consists of four input nodes (contact time, initial concentration, temperature, pH) in the first layer, and one output node in the third layer (NH_4^+ removal rate).

Normalization can accelerate the convergence of the training network and reduce the impact of different orders of magnitude of the input data set. For increasing the training rate of the network [19], all the input and output variables were normalized to 0–1 by

the following Equation (4). The normalized data will be used to train and test the neural network, and then the output data will be unnormalized.

$$X_i = (X - X_{min})/(X_{max} - X_{min}) \quad (4)$$

where X_i stands for normalized values, and X, X_{min}, and X_{max} are the original, minimum, and maximum values of variables, respectively.

The number of neurons in the hidden layer due to their substantial impact on the performance of the network is used as a design parameter of the model. Therefore, to determine the optimum number of neurons between 1–14 in this layer, various topologies were examined by the mean square error (MSE), MSE was calculated by Equation (5). The relationship between the predicted value of the model and the experimental value is expressed by the root mean square error (RMSE), RMSE was calculated by Equation (6).

$$MSE = \sum_{i=0}^{n} (y_{i,cal} - y_{i,exp})^2 / n \quad (5)$$

$$RMSE = \sqrt{\sum_{i=0}^{n} (y_{i,cal} - y_{i,exp})^2 / n} \quad (6)$$

where, $y_{i,cal}$ and $y_{i,exp}$ are predicted and experimental values of the response, respectively, and n represents the number of data points. Each topology was repeated ten times to prevent random correlation because of random initialization of the weights and biases [32].

Other details of the parameter in the process of training and testing of the BP-ANN model were listed in the Table 3.

Table 3. Summary of the parameters used in construction of ANN model.

Type	Description
Input layer	4 neurons (contact time, initial concentration, temperature, pH)
Hidden layer	1 layer; 8 neurons
Output layer	1 neuron (NH_4^+ removal rate)
Learning rate	0.01
Epoch	1000
MSE goal	0.001
Algorithms	Levenberg-Marquardt (trainlm)
Function	Sigmoid (tansig): Between input and hidden layers Linear: Between hidden and output layers

Sensitivity analysis was conducted by Garson algorithm with the connected weights obtained by BP-ANN, so as to calculate the relative influence degree of different input variables on output variables [33]. The equation below is proposed by Garson for this type of analysis [34].

$$Q_{ik} = \frac{\sum_{j=1}^{L} \left(\left| w_{ij} v_{jk} \right| / \sum_{r=1}^{N} \left| w_{rj} v_{jk} \right| \right)}{\sum_{i=1}^{N} \sum_{j=1}^{L} \left(\left| w_{ij} v_{jk} \right| / \sum_{r=1}^{N} \left| w_{rj} v_{jk} \right| \right)} \quad (7)$$

where, N, L, and M are the number of neurons in the input layer, hidden layer, and output layer, respectively. w, v are the connection weights between the input layer and the hidden layer, and hidden layer and the output layer, respectively.

2.4.3. Genetic Algorithms

Using genetic algorithms (GA), as the artificial intelligence-based stochastic non-linear optimization formalism, the optimum values of variables were found. For this purpose, the developed ANN models were utilized as a fitness function. In the genetic algorithms, evolutionary operators including reproduction, crossover, and mutation are used to find the best answer in a large search space [35]. At first, it randomly provides an initial population of individuals called chromosome and evaluates it using fitness function. Then

low fitness chromosomes are eliminated, and by employing evolutionary operators, the next generation is produced. This process is continued until convergence forms the solution, and a suitable result is obtained. The flow chart of the combination of genetic algorithm and neural network was shown in Figure 2.

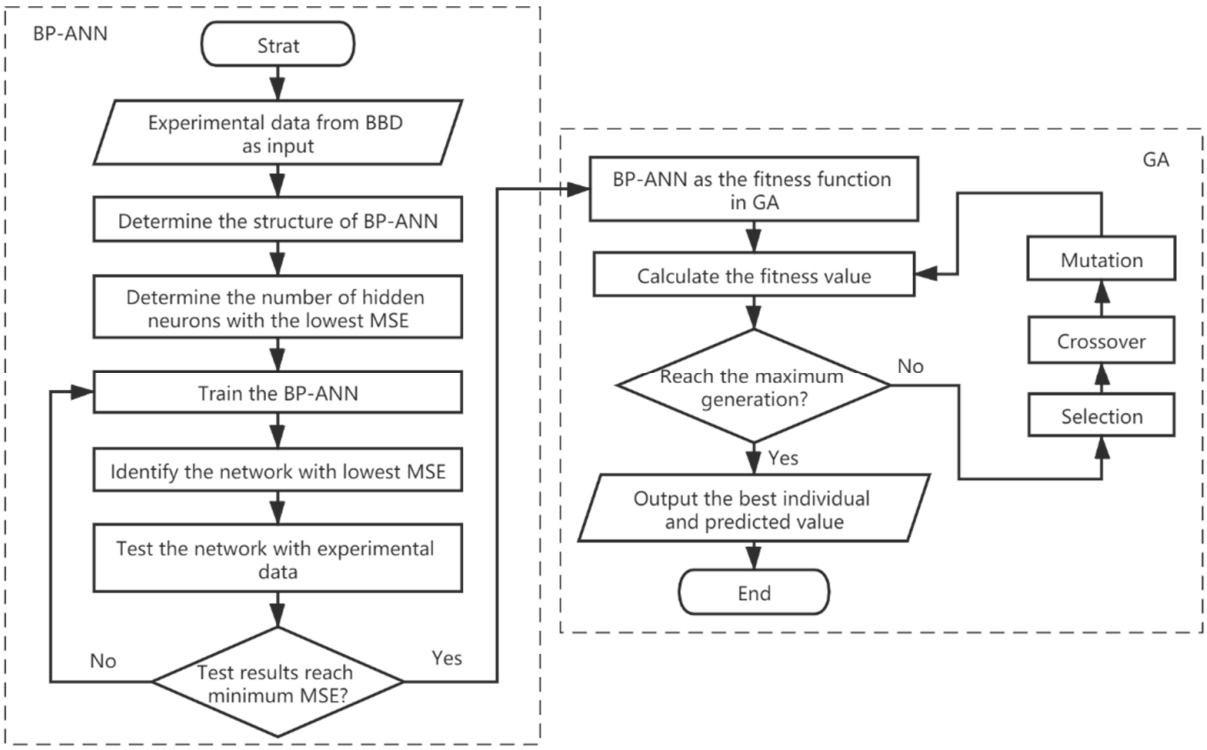

Figure 2. Flow chart of back-propagation artificial neural network (BP-ANN) coupled with genetic algorithms (GA).

3. Results and Discussion

3.1. Characterization of the CB-GAC

The surface electron microscopy (SEM) images (1000× and 5000×) of the CB-GAC used in this study were shown in Figure 3a,b. It can be seen that the surface of the activated carbon is very rough, with many small particles on the surface and many pores of different sizes. In addition, the surface of CB-GAC can be observed with both micropores and slit mesoporous pores under the magnification of 5000×. X-ray diffraction studies help in determining the degree of crystalline or amorphous nature of activated carbon. The X-ray diffraction patterns of CB-GAC was shown in Figure 3c. An increase in sharpness of the ~26° peak is seen with the rise, indicating that the CB-GAC consists of graphite-like microcrystallites [36]. The peak width is relatively narrow, indicating that its laminar crystal structure is good. Nitrogen adsorption-desorption isotherms were illustrated in Figure 3d. The physisorption isotherms of CB-GAC are classified as Type IV isotherm with H3 shaped hysteresis loops according to the International Union of Pure and Applied Chemistry (IUPAC) classification. Type IV isotherms are characteristic of mesoporous materials. When gas is adsorbed on the mesoporous adsorbent below the critical temperature, the monomolecular adsorption layer is formed first, and the multi-molecule adsorption layer begins to occur when the adsorption reaches saturation. The isotherm of CB-GAC reached saturation shortly after the adsorption began (p/p° = 0.1). With the increase of relative pressure (p/p° = 0.4~1.0), the adsorption capacity increased rapidly, which was due to capillary condensation, and also indicated that the mesoporous size was relatively uniform. This may be related to dopants in the preparation process of columnar activated carbon, such as metals and their compounds. On the one hand, metals occupy a certain space when they are mixed with raw materials. After pickling, metals are removed to form a

hole space. On the other hand, when metal compounds are added to the raw material, it can catalyze the reaction of water vapor or carbon dioxide with carbon, thus forming or enlarging pores [18]. Yahya et al. also obtained this type of curve in Ni-Co modified activated carbon [27]. The information of the specific surface area, pore volume, and pore diameter was shown in Table 4. It can be seen that micropore and mesoporous pore volumes account for 40% and 60%, respectively, and the average pore diameter is 7.9 nm, indicating that the CB-GAC is a mesoporous material.

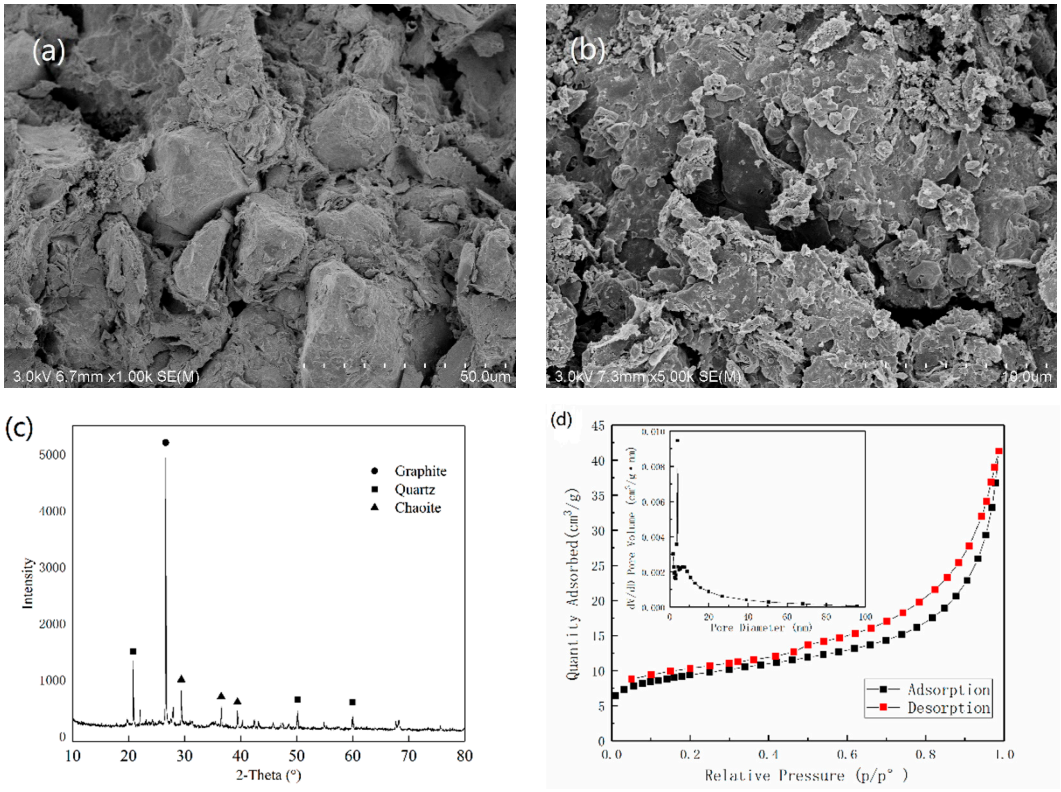

Figure 3. The characteristics of coal-based granular activated carbon (CB-GAC): (**a**) SEM image 1000×; (**b**) SEM image 5000×; (**c**) X-ray diffraction patterns; (**d**) N_2 adsorption desorption isotherm and pore size distribution.

Table 4. Characteristics determined by Brunauer-Emmett-Teller (BET) method for composite samples.

Parameters	BET Surface Area (m^2/g)	Langmuir Surface Area (m^2/g)	t-Plot Micropore Area (m^2/g)	t-Plot External Surface Area (m^2/g)	Total Pore Volume of Pores (cm^3/g)	t-Plot Micropore Volume (cm^3/g)	Average Pore Diameter (nm)
CB-GAC	32.2108	59.304	12.9979	19.2129	0.063856	0.006107	7.92976

3.2. Isotherm and Kinetics Studies

3.2.1. Adsorption Isotherm

As shown in Figure 4, Langmuir, Freundlich, and Temkin models were used to fit the adsorption equilibrium data. The fitting parameter values and errors of the adsorption isotherm model were shown in Table 5. It was observed that the R^2 value of Langmuir isotherm model (0.9951) was higher than that of the Freundlich (0.9843) and Temkin models (0.9925), indicating that the adsorption behavior of CB-GAC for NH_4^+ was more consistent with the Langmuir model. The Langmuir isotherm model assumes the number of active sites distributed homogeneously on the surface of the adsorbent followed by monolayer adsorption (physical adsorption) having high adsorptive power [37]. This suggested that the adsorption of NH_4^+ takes place on the surface of the CB-GAC until a monolayer

coverage was formed, after which the driving force of the sorption process decreases drastically [38]. The Q_{max} calculated from the Langmuir model was 0.2821 mg/g.

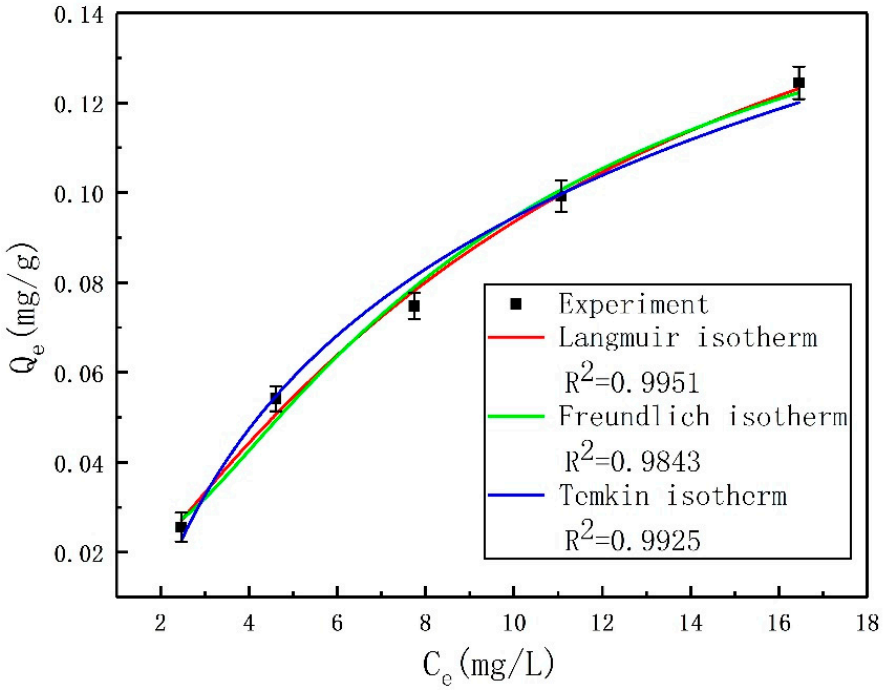

Figure 4. Langmuir, Freundlich, and Temkin isotherms for the adsorption of NH_4^+ onto CB-GAC. (Experimental conditions: Initial pH: 7; CB-GAC dose: 10 g/100 mL; temperature: 25 °C; contact time: 24 h).

Table 5. Parameters of isotherm models for the adsorption of NH_4^+ onto CB-GAC.

Model	Parameters	Values
Langmuir	Q_{max} (mg/g)	0.2821
	K_L (L/mg)	0.0481
	R^2	0.9951
	RSS ($\times 10^{-4}$)	2.9047
Freundlich	1/n	0.7165
	K_F (L/mg)	0.0171
	R^2	0.9843
	RSS ($\times 10^{-4}$)	9.2406
Temkin	b_T	48,263.81
	K_T	0.6299
	R^2	0.9925
	RSS ($\times 10^{-4}$)	4.3932

3.2.2. Adsorption Kinetics

The adsorption kinetic models and the trend of NH_4^+ removal rate with time were illustrated in Figure 5. At the started phase, the adsorption efficiency was very high, and the amount of adsorbed ammonia nitrogen increased rapidly. After the started phase, the adsorption efficiency gradually slowed down and reached equilibrium at around 16 h. The fitting parameter values and errors of the adsorption kinetic models were shown in Table 6. It can be seen from the R^2 value that, compared with the pseudo-first-order kinetic model (0.9535), the adsorption kinetics behavior of NH_4^+-CB-GAC is more consistent with the pseudo-second-order kinetic model (0.9868), which suggested that the adsorption might depend on the availability of the adsorption sites. This is similar to the kinetic adsorption characteristics of many carbon materials [21,39–41]. The calculated value Q_e

from pseudo-second-order kinetic was 0.0956 mg/g, which is close to the experimental value of 0.0927 mg/g.

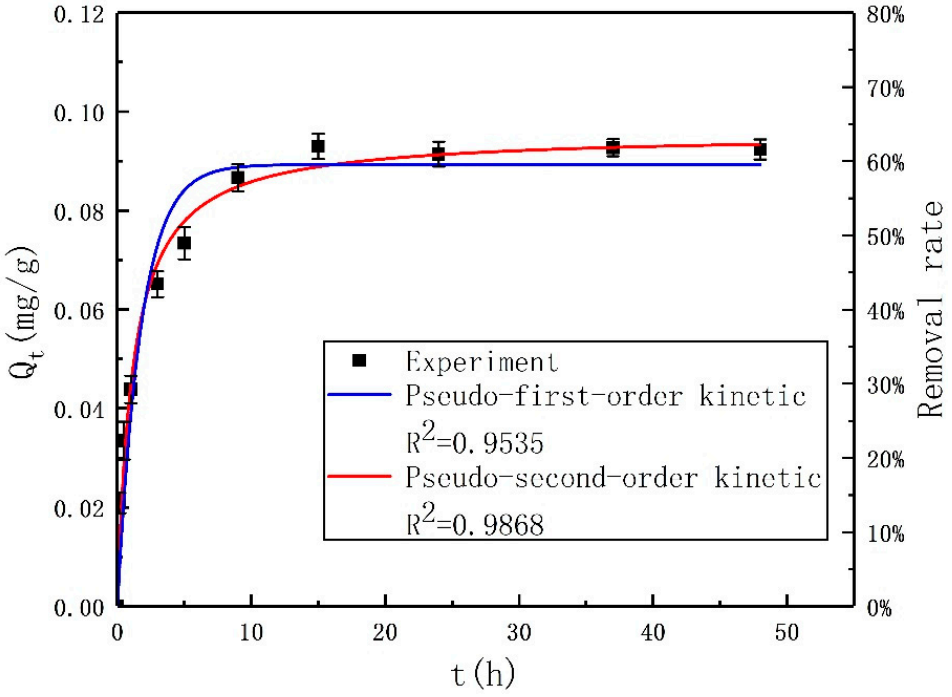

Figure 5. Pseudo-first-order, pseudo-second-order kinetics and removal rate for the adsorption of NH_4^+ onto CB-GAC. (Experimental conditions: Initial pH: 7; CB-GAC dose: 10 g/100 mL; temperature: 25 °C; initial concentration: 15 mg/L).

Table 6. Parameters of kinetic models for the adsorption of NH_4^+ onto CB-GAC.

Model	Parameters	Values
Pseudo-first-order	k_1	0.5682
	Q_e (mg/g)	0.0893
	R^2	0.9535
	RSS ($\times 10^{-4}$)	5.2933
Pseudo-second-order	k_2	9.1682
	Q_e (mg/g)	0.0956
	R^2	0.9868
	RSS ($\times 10^{-4}$)	1.5052

3.3. Modeling and Optimization by BBD-RSM

3.3.1. Modeling

The experimental results obtained according to the design of BBD-RSM were shown in Table 7, and the experimental values of the 5 repeated groups are relatively close, indicating that the data has good reproducibility. The predicted value and the experimental value have a good coincidence, and the relative error is less than 0.2, indicating that the model has a good predictive ability. The experiment was fitted according to the second-order model of polynomial regression analysis. A quadratic regression model was established by taking contact time (A), initial concentration (B), temperature (C), and pH (D) as independent variables, and NH_4^+ removal rate Y as response value.

Table 7. RSM (response surface methodology) and ANN (artificial neural network) predicted results and errors, along with experimental values of the response.

Run	Variables				NH_4^+ Removal Rate				
	A/min	B/mg·L^{-1}	C/°C	D	Experiment	RSM	Error	ANN	Error
1 (Tra.)	600	15	35	9	54.37	53.92	0.0083	54.3700	0.0000
2 (Tra.)	300	20	25	7	52.65	52.6	0.0009	52.6500	0.0000
3 (Tra.)	600	15	25	7	59.85	59.7	0.0025	59.6175	0.0039
4 (Tra.)	600	10	15	7	61.27	60.49	0.0127	61.2700	0.0000
5 (Tra.)	900	10	25	7	61.44	61.97	0.0086	61.4400	0.0000
6 (Tra.)	600	10	25	9	56.36	56.45	0.0016	56.3600	0.0000
7 (Tra.)	600	15	25	7	59.47	59.7	0.0039	59.6175	0.0025
8 (Tra.)	600	15	25	7	60.27	59.7	0.0095	59.6175	0.0108
9 (Tra.)	600	20	25	9	56.16	55.95	0.0037	56.1600	0.0000
10 (Tra.)	600	15	25	7	58.88	59.7	0.0139	59.6175	0.0125
11 (Tra.)	300	15	15	7	54.67	55.02	0.0064	54.6700	0.0000
12 (Tra.)	600	10	25	5	53.38	54.16	0.0146	53.3800	0.0000
13 (Tra.)	600	20	35	7	55.97	55.7	0.0048	55.9700	0.0000
14 (Tra.)	900	15	25	5	57.47	56.69	0.0136	57.4700	0.0000
15 (Tra.)	600	15	35	5	51.34	50.86	0.0093	51.3400	0.0000
16 (Tra.)	300	15	25	5	49.26	48.33	0.0189	49.2600	0.0000
17 (Val.)	300	15	25	9	50.27	56.18	0.1176	51.0862	0.0162
18 (Val.)	900	15	25	9	59.62	59.61	0.0002	58.9691	0.0109
19 (Val.)	600	15	25	7	60.03	49.37	0.1776	59.6175	0.0069
20 (Val.)	900	15	15	7	63.32	60.08	0.0512	62.2006	0.0177
21 (Tes.)	600	10	35	7	56.44	50	0.1141	56.4456	0.0000
22 (Tes.)	900	15	35	7	59.39	59.5	0.0019	59.9995	0.0103
23 (Tes.)	300	15	35	7	48.12	59.7	0.2406	49.0211	0.0187
24 (Tes.)	600	20	15	7	60.88	62.64	0.0289	60.2450	0.0104
25 (Pre.)	600	20	25	5	53.29	53.77	0.0090	53.2720	0.0003
26 (Pre.)	300	10	25	7	53.19	52.84	0.0066	52.9502	0.0045
27 (Pre.)	600	15	15	9	56.47	57.44	0.0172	56.3900	0.0014
28 (Pre.)	900	20	25	7	60.48	61.32	0.0139	61.5586	0.0178
29 (Pre.)	600	15	15	5	55.08	56.02	0.0171	55.1653	0.0015

Empirical models are generated from coding factors (standardized equation) and actual factors (non-standardized equation). The coded equation (Equation (8)) is useful for identifying the relative impact of the variables by comparing the variables coefficients, while the actual equation (Equation (9)) can be used to make predictions about the response for given levels of each factor. The model demonstrates quadratic coefficient of contact time (A) had the main obverse effect on NH_4^+ removal rate for its highest positive coefficient value, which is consistent with the conclusion of kinetic model that the NH_4^+ removal rate increases with time. The second important variable is temperature with negative sign. This suggested that the adsorption capacity of the adsorbent decreased with the increase of temperature. Similar results were obtained by Ren et al. in the experiment of the adsorption of ammonia nitrogen by iron-loaded activated carbon [42].

$$Y = 59.7 + 4.46A - 0.2208B - 2.17C + 1.12D - 0.105AB + 0.655AC + 0.285AD - 0.02BC - 0.0275BD + 0.41CD - 1.99A^2 - 0.5317B^2 - 1.06C^2 - 4.09D^2 \quad (8)$$

$$y = -8.64802 + 0.033617a + 0.665083b + 0.042042c + 14.10667d - 0.00007ab + 0.000218ac + 0.000475ad - 0.0004bc - 0.00275bd + 0.0205cd - 0.000022a^2 - 0.021267b^2 - 0.010554c^2 - 1.02167d^2 \quad (9)$$

3.3.2. Analyzing

ANOVA method was used to assess the adequacy and validity of the generated regression models. This is to determine the significant effect of process variables to response variable as well as to fit the second-order polynomial models to the experimental data.

The analysis of variance (ANOVA) and the fit statistics were shown in Tables 8 and 9, respectively. Montgomery points out that the following conditions should be met when testing the model: F value > 0.1, R^2 > 0.95, $R_{pred}{}^2$ > 0.7, $R_{adj}{}^2$-$R_{pred}{}^2$ < 0.2, C.V. < 10%, Adeq Precision > 4 [20]. F-value and *p*-value was used to determine the statistical significance of the model. It can be seen that the F value of this model is 40.94 > 0.1, $p < 0.0001$, showing that the model is reliable and fits well in the whole regression area. $R_{adj}{}^2 - R_{pred}{}^2 = 0.0785 < 0.2$, C.V. = 1.54% < 10%, indicating high reliability and accuracy of the experiment. Adeq Precision is the ratio of effective signal to noise. The experimental model Adeq Precision = 22.8713 > 4, indicating that the model is reliable and has enough signals to respond to the design. As shown in Figure 6, the high determination coefficient (R^2 = 0.9762) indicated a strong correlation between the predicted and actual values. Hence, the obtained model provided a good estimation of the predicted response within the studied range. The parameters were considered significant if *p*-value (Prob > F) is lower than 0.05. From the ANOVA (Table 8), the coded parameters A, C, D, A^2, C^2, D^2 are significant parameters, i.e., $(p > F) < 0.05$.

Table 8. Analysis of variance (ANOVA) of the second-order polynomial equation.

Source	Sum of Squares	Degree of Freedom	Mean Square	F-Value	*p*-Value	
Model	433.70	14	30.98	40.94	<0.0001	significant
A-Contact time	239.06	1	239.06	315.92	<0.0001	
B-Initial concentration	0.5852	1	0.5852	0.7734	0.3940	
C-Temperature	56.59	1	56.59	74.79	<0.0001	
D-pH	15.03	1	15.03	19.86	0.0005	
AB	0.0441	1	0.0441	0.0583	0.8127	
AC	1.72	1	1.72	2.27	0.1543	
AD	0.3249	1	0.3249	0.4294	0.5229	
BC	0.0016	1	0.0016	0.0021	0.9640	
BD	0.0030	1	0.0030	0.0040	0.9505	
CD	0.6724	1	0.6724	0.8886	0.3618	
A^2	25.57	1	25.57	33.79	<0.0001	
B^2	1.83	1	1.83	2.42	0.1419	
C^2	7.23	1	7.23	9.55	0.0080	
D^2	108.33	1	108.33	143.16	<0.0001	
Residual	10.59	14	0.7567			
Lack of Fit	9.41	10	0.9412	3.19	0.1377	not significant
Pure Error	1.18	4	0.2954			
Cor Total	444.30	28				

Table 9. Fit statistics of ANOVA.

Std. Dev.	Mean	C.V. %	R^2	Adjusted R^2	Predicted R^2	Adeq Precision
0.8699	56.53	1.54	0.9762	0.9523	0.8738	22.87

The interaction effect of contact time and pH on $NH_4{}^+$ removal rate was displayed in Figure 7. The results showed that with the increase of pH, the removal of $NH_4{}^+$ increased first and then decreased, reaching the maximum value around 7.2. For the contact time, as the contact time increases, the $NH_4{}^+$ removal rate also increased, and the trend slowed down after reaching a certain value, indicating that the contribution of the contact time to the $NH_4{}^+$ removal rate gradually tends to be saturated. This is consistent with the kinetic results. The research showed that in response surface analysis, if the contour shape is elliptic, it means that the interaction between factors is significant, while the circle means that the interaction between factors is not significant [43]. As can be seen intuitively from the contour plot, the interaction between contact time and pH is relatively significant. When

pH is close to the optimal value, the removal rate of NH_4^+ gradually reaches the limit value with the increase of contact time, and the value is closer to the optimal removal rate. This can be explained by the morphology transform of NH_4^+: in the acidic environment, high amount of H^+ caused a strong competition with NH_4^+, while in the alkaline environment, NH_4^+ was converted into the $NH_3 \cdot H_2O$ molecular form [42]. Both the conditions can result in a reduction in NH_4^+ removal rate. The interaction between contact time and pH indicated that there is an optimal NH_4^+ removal area, that is, the area with a contact time of 700~900 min and pH of 7.2, with NH_4^+ removal rate of over 62%.

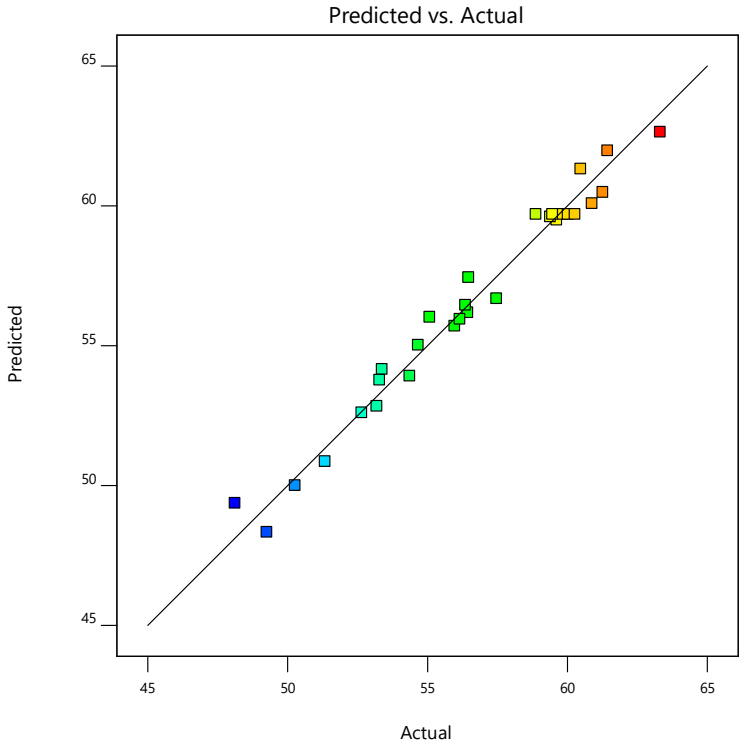

Figure 6. The relationship between the predicted and actual values for the Box-Behnken design, $R^2 = 0.9762$.

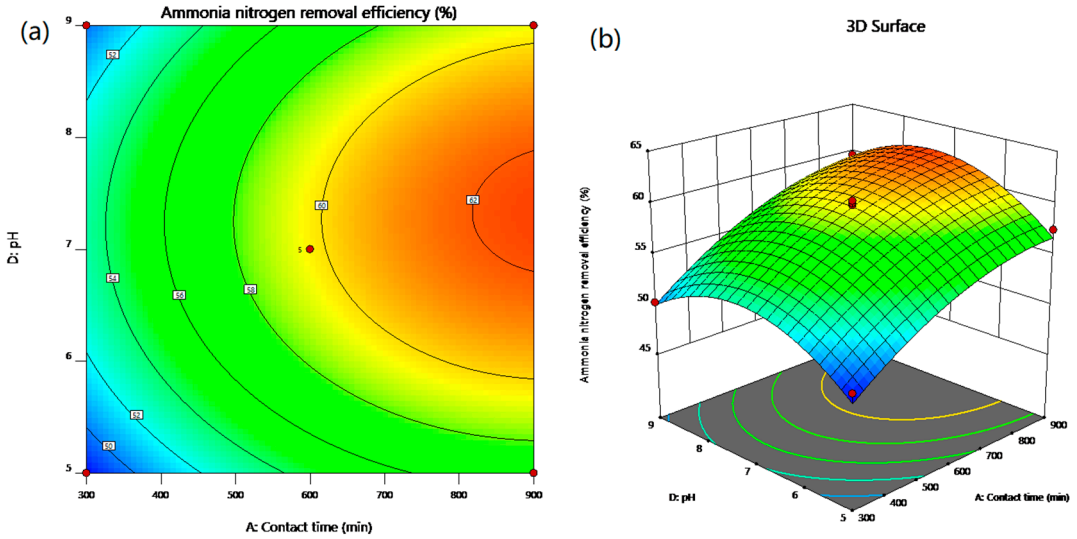

Figure 7. 2D contour plot (**a**) and 3D response surface plot (**b**) showing the effect of contact time and pH on the removal rate of NH_4^+.

The interaction effect of contact time and temperature on NH_4^+ removal rate was displayed in Figure 8. The results showed that with the increase of contact time and decrease of temperature, NH_4^+ removal rate increased, then the trend was gradually slowed. In the contour plot, the contour was thinning, which suggested that contact time and temperature on the contribution of NH_4^+ removal rate were both gradually tending to saturation. It can be intuitively seen from the contour plot that the interaction between contact time and temperature is relatively significant, which is reflected in that when the temperature approaches the optimal value, with the increase of contact time, the removal rate of NH_4^+ gradually reaches the limit value, and the value is closer to the optimal removal rate. The interaction between contact time and temperature indicated that there is an optimal NH_4^+ removal area, that is, the area with a contact time of 700~900 min and temperature of 25~35 °C, and the NH_4^+ removal rate is above 62%. It can also be seen from the figure that contact time has a greater influence on NH_4^+ removal rate than temperature.

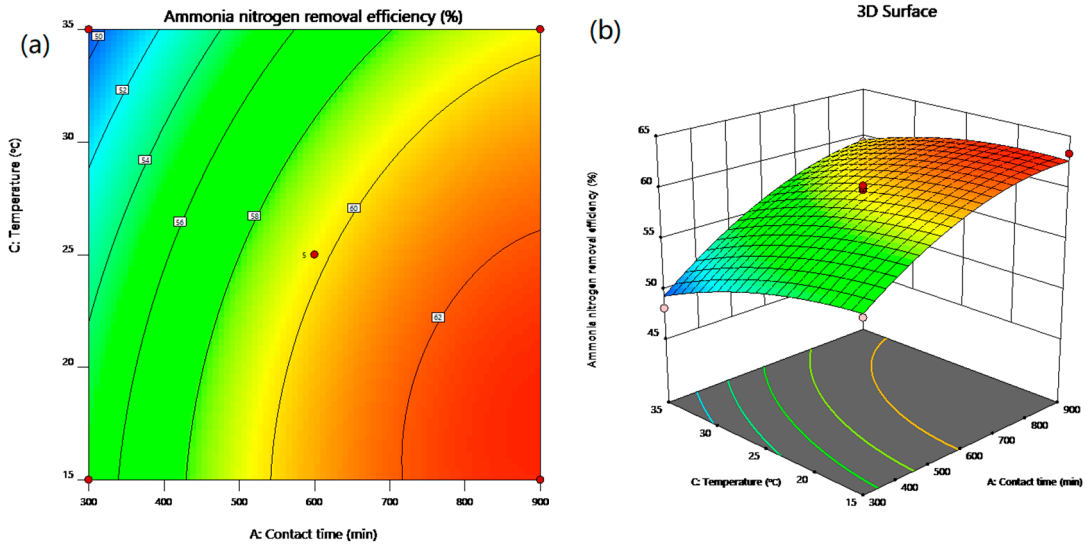

Figure 8. 2D contour plot (**a**) and 3D response surface plot (**b**) showing the effect of contact time and temperature on the removal rate of NH_4^+.

3.3.3. Determination of Optimal Conditions for RSM

The actual regression equation can be written as follows:

$$Y = a + X^T b + X^T cX \tag{10}$$

where, $X = [A, B, C, D]^T$.

According to,

$$\frac{\partial Y}{\partial X} = \left[\frac{\partial Y}{\partial A}, \frac{\partial Y}{\partial B}, \frac{\partial Y}{\partial C}, \frac{\partial Y}{\partial D}\right] = 0 \tag{11}$$

Take the first partial derivative of the equation:

$$\begin{bmatrix} 4.4 \times 10^{-5} & 7 \times 10^{-5} & -2.8 \times 10^{-4} & -4.75 \times 10^{-4} \\ 7 \times 10^{-5} & 4.2534 \times 10^{-2} & 4 \times 10^{-4} & 2.75 \times 10^{-3} \\ -2.18 \times 10^{-4} & 4 \times 10^{-4} & 2.1108 \times 10^{-2} & -2.05 \times 10^{-2} \\ -4.75 \times 10^{-4} & 2.75 \times 10^{-3} & -2.05 \times 10^{-2} & 2.04334 \end{bmatrix} \cdot \begin{bmatrix} A \\ B \\ C \\ D \end{bmatrix} = \begin{bmatrix} 3.3617 \times 10^{-2} \\ 0.665083 \\ 4.2042 \times 10^{-2} \\ 14.10667 \end{bmatrix} \tag{12}$$

Solution of equation: A = 911.42min, B = 13.49 mg/L, C = 18.22 °C, D = 7.28. The optimal process conditions for NH_4^+ removal by CB-GAC: Contact time = 911.42 min, initial concentration = 13.49 mg/L, temperature = 18.22 °C, pH = 7.28, removal rate = 62.83%.

3.4. BP-ANN

3.4.1. Determination of the Number of Hidden Neurons

Error minimization within the networks is comprised of the appropriate selection of the number of neurons in the hidden layer. The calculated network error is compared with the output continuously until the network reaches the minimum error by adjusting weights and biases. In order to obtain the minimum MSE of the training network, a trial-and-error method was used to produce the optimum neurons with lowest MSE value. The MSE plot for different number of neurons (1–14) in the hidden layer for the response of NH_4^+ removal rate was depicted in Figure 9. It is observed that the lowest MSE is obtained with 8 neurons for the NH_4^+ removal rate. Therefore, the best network structure of 4-8-1 is used for process optimization, which represents 4 inputs in the first layer, followed by 8 neurons in the hidden layer and one output in the last layer.

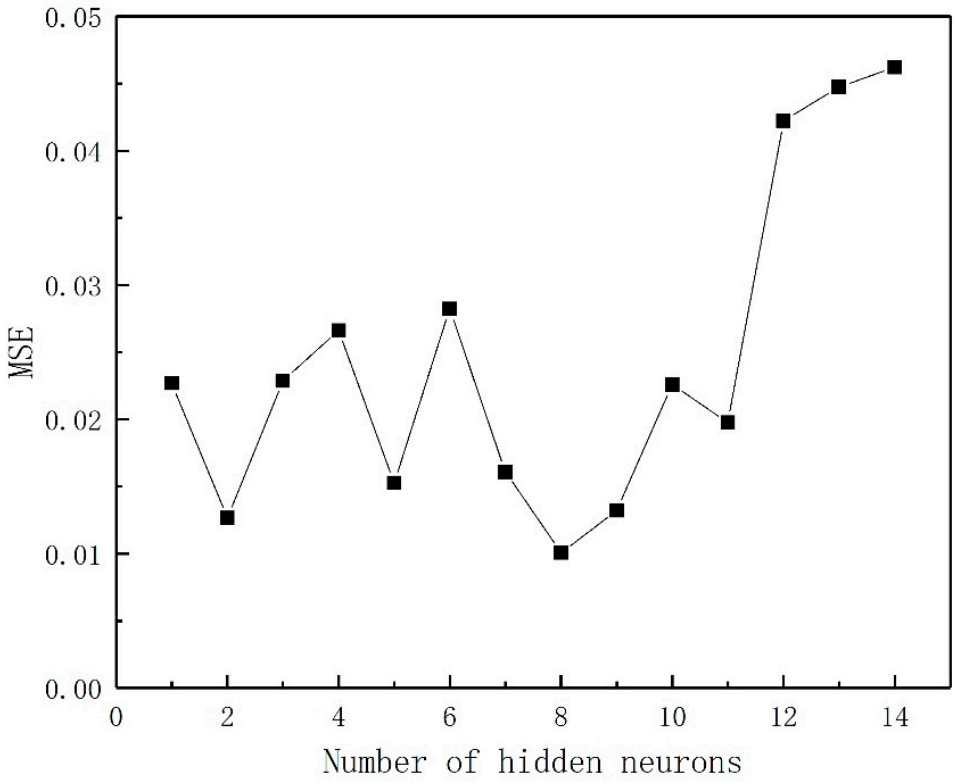

Figure 9. MSE (mean square error) plot for different numbers of neurons in the hidden layer (1–14) for the response of NH_4^+ removal rate.

3.4.2. Evaluation of Model

The R^2 values for training, validation, and test and all data, which evaluate the relationship between experimental and predicted values, have been shown in Figure 10. It is seen that approximately the whole values have located around the 45° line with R^2 values of 0.99738, 0.99965, 0.99584, and 0.9951 for training, validation, test, and all data. This indicated excellent compatibility between the experimental and predicted results by the ANN model. As shown in Figure 11, after the first iteration, the MSE of the system reaches the preset value, and the system stops training. The trained neural network was tested by the experimental data, and the R^2 between the predicted and actual data was 0.99589 (Figure 12), which indicated that the BP-ANN model has good predictive ability.

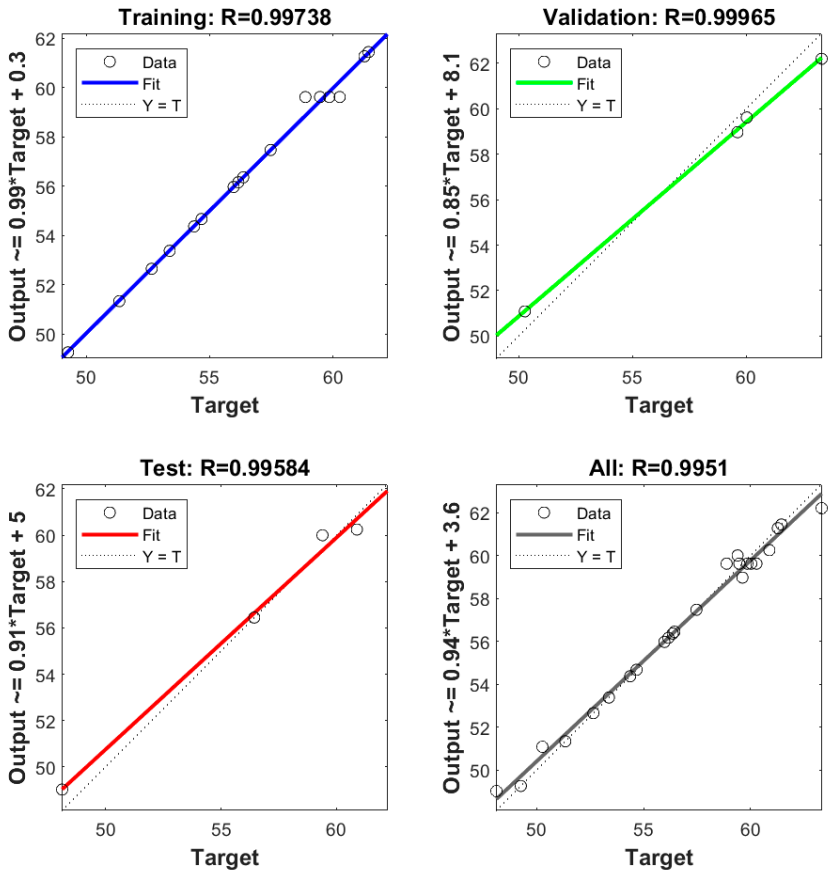

Figure 10. Regression plot of experimental data and BP-ANN model simulated values.

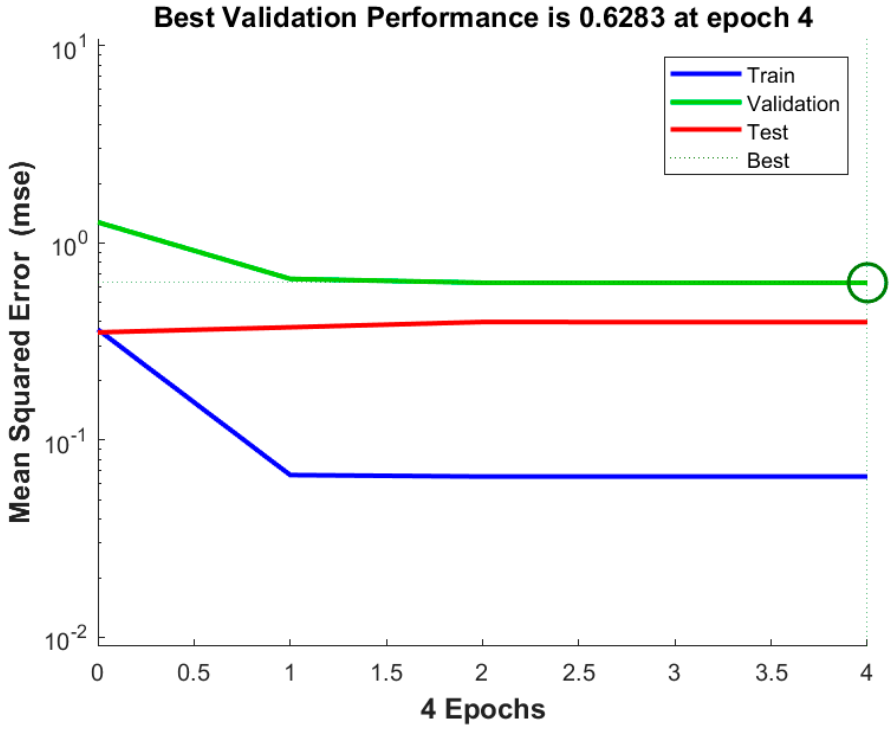

Figure 11. Plot of validation performance by MSE.

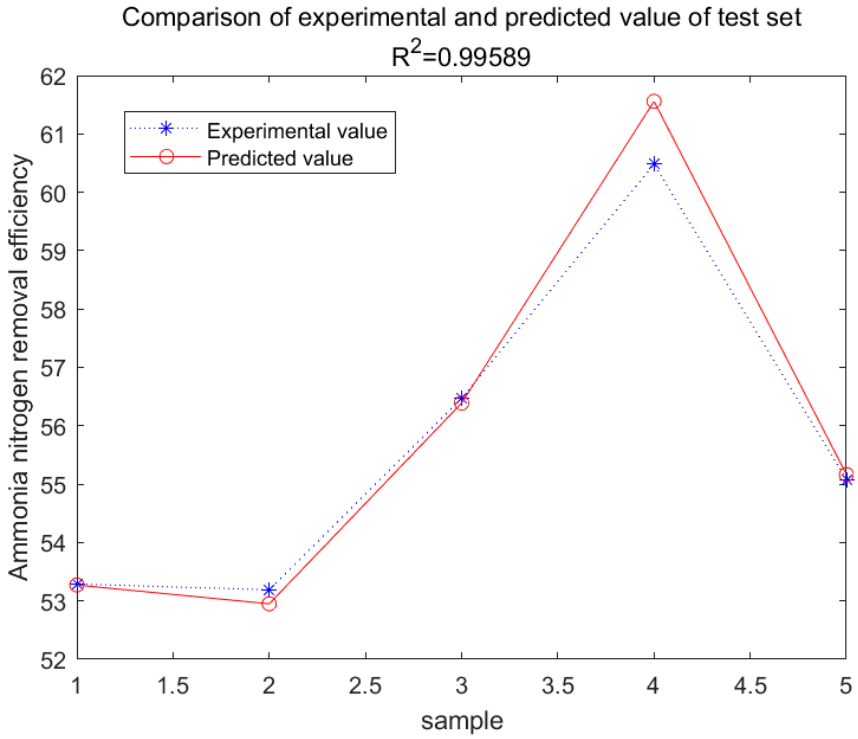

Figure 12. Comparison of experimental and predicted value of test set.

3.4.3. Sensitivity Analysis

A sensitivity analysis was performed to analyze the "cause-and-effect" relationship between the input variables and the modeling outputs. The values of weights and biases in input-hidden layer (w_i and b_i) and hidden-output layer (w_j and b_j) were listed in Table 10. In order to investigate the contributions of the independent variables on the performance of the system, their relative importance was calculated using Garson algorithm. It can be seen from the calculated results (Table 11) that the contact time appears to be the most influential variables, followed by pH and temperature, initial concentration has the least influence on the removal rate of NH_4^+. This result is not completely consistent with the RSM result. They agreed that contact time was the most important factor of the four variables, and initial concentration was the least important. The difference is that the quadratic equation in RSM considers that the influence of temperature is greater than the influence of pH. In addition, the results of sensitivity analysis can concretize the proportion of relative influence degree, but the quadratic equation can reflect the positive and negative correlation of factors. Combining these two results, increasing the contact time, lowering the temperature, and keeping the pH near 7 can effectively improve the adsorption removal rate of NH_4^+.

Table 10. The weights and biases of BP-ANN in input-hidden layer (w_i and b_i) and hidden-output layer (w_j and b_j).

Number of Neurons	w_i				b_i	w_j	b_j
	Contact Time	Initial Concentration	Temperature	pH			
1	0.2513	0.7242	2.6647	0.6973	2.6315	0.3643	−0.7167
2	0.5408	0.1291	0.5256	2.2940	−1.5721	0.0955	
3	−0.9756	−2.1045	0.5286	−1.0694	0.1463	0.1405	
4	1.3235	−0.5437	0.8937	−0.3642	0.7363	0.7143	
5	1.2157	1.4659	−0.4634	−1.3170	−0.2739	−0.0235	
6	−1.9242	0.6120	−1.8276	1.3909	−1.2270	0.0557	
7	0.7446	1.5502	−1.9754	−0.6814	1.8150	0.2089	
8	−1.1734	0.5911	2.5940	−0.6891	−1.7490	−0.3516	

Table 11. Relative significance of input variables.

Input Variables	Relative Significance (%)	Ranking
Contact time	31.23	1
Initial concentration	21.16	4
Temperature	22.93	3
pH	24.68	2

3.5. Genetic Algorithm (GA)

GA method was employed to optimize the input space of the optimal network with the intention of maximizing the NH_4^+ adsorbed in the adsorption procedure. The optimization objective was determined via searching for the optimum points of the process variables between lower and upper bounds. Variables ranges was set as follows: Contact time 300–900 min, initial concentration 10–20 mg/L, temperature 15–35 °C, pH 5–9. The number of iterations of the genetic algorithm is set to 500. As shown in Figure 13, the system has shown a good convergence effect after about 70 iterations, while after 327 iterations, the system does not change, indicating the optimal results has been found. The results showed that the maximum removal rate was 63.74% under the optimal conditions of contact time = 899.41 min, initial concentration = 17.35 mg/L, temperature = 15 °C, pH = 6.98. The higher prediction accuracy of the ANN-GA model is attributed to the general ability of ANN-GA to estimate the nonlinear behavior of the system, while the response surface model is limited by second-order polynomial regression [44]. Therefore, these results confirm the advantages of ANN-GA model as an alternative to RSM model in prediction.

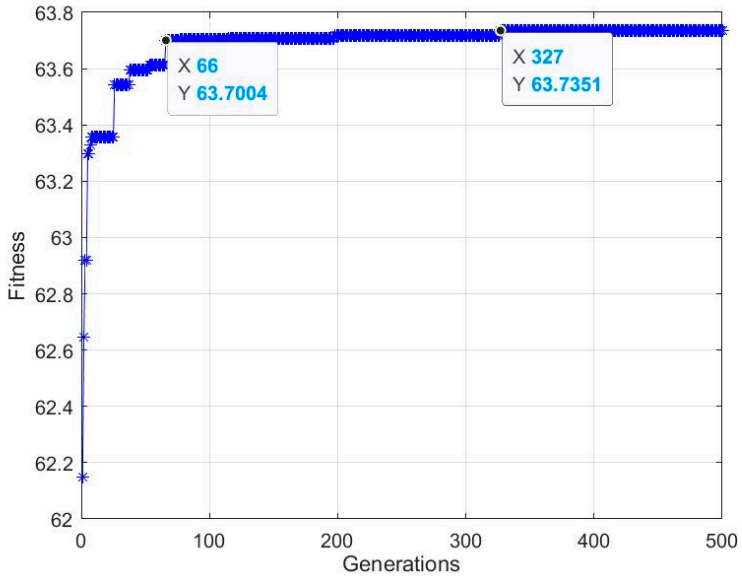

Figure 13. Evolvement of fitness with 500 generations.

3.6. Comparison between ANN-GA and RSM

The coefficient of determination is usually considered as the most common statistical metric to determine the goodness of prediction and measured by fitting a straight line. This approach focuses on linear relationships between the experimental results and model predictions in direction and strength, but does not provide any information for error distribution and nonlinear relationship. It can be seen from Table 12 that the R^2 values between the experimental and predicted values are calculated as 0.9762 and 0.9959 for the RSM and ANN models, respectively, indicating that the predictions resulted by the ANN model are closer to experimental values. The RMSE were found lower values for predictions given by ANN rather than the RSM model, confirming that the less error deviation resulted from the ANN predictions. In comparison, the prediction and statistical

metrics for the ANN model were relatively better than the RSM model, and the difference was obvious.

Table 12. Comparison of experimental validation results and predicted results of BBD-RSM and BP-ANN-GA.

Variables	BBD-RSM		BP-ANN-GA	
	Predicted Parameters	Experimental Parameters	Predicted Parameters	Experimental Parameters
Contact time (min)	911.42	910	899.41	900
Initial concentration (mg/L)	13.49	13.5	17.35	17.5
Temperature (°C)	18.22	18	15	15
pH	7.28	7.3	6.98	7.0
Removal rate (%)	62.83	61.32	63.74	63.19
Relative error (%)	2.46		0.87	
R^2	0.9762		0.9959	
RMSE	3.4509		0.4690	

To confirm the predicted-response reliability by the obtained fitted model, three tests at optimum level of the independent variables were examined. The experimental removal rate and absolute error were presented in Table 12. The experimental and predicted removal efficiencies revealed that both models have a promising ability to predict the values remarkably close to experimental values; however, it seemed that the ANN model compared to RSM model is more robust to provide the predictions closer to the experimental efficiency. It can be observed that both models predicted a similar value for the optimum contact time and pH. The predicted percentage error of NH_4^+ removal rate is 0.87% and 2.46% for ANN-GA and RSM models, respectively. These results suggested greater accuracy and higher reliability of ANN-GA in modelling and optimizing the parameter interaction related to the NH_4^+ removal rate.

In conclusion, both models have their advantages. The RSM showed the influence of the interaction among various factors on the removal rate of NH_4^+, and provided graphs to intuitively explain the relationship between independent variables and response values. Additionally, this method only needs a few experiments to produce more information, reducing time and cost. However, the disadvantage is that it can only provide a first or second order polynomial model. Artificial neural network can simulate any form of nonlinearity, because it is considered as a black box model, which does not need experimental design to achieve a clear relationship. Therefore, it overcomes the difficulty of experimental design and is a more unlimited method. Finally, according to the excellent results obtained from both modeling processes, the modeling approaches in real-scale stormwater treatment systems can be developed to benefit from their application in modeling, optimizing, and recognizing the relationship among variables.

4. Conclusions

In the present work, the statistical modeling and optimization of process parameters (contact time, initial concentration, temperature, pH) for NH_4^+ removal from stormwater by CB-GAC was carried out using BBD-RSM and GA-ANN. In terms of prediction, neural network had better prediction accuracy than response surface method, with R^2 of 0.9959 and 0.9762, respectively. The ANOVA and response surface plots in RSM confirmed that contact time was the most significant parameter of NH_4^+ removal, and the relative influence order of the factors according to the coefficients of the code equation is as follows: Contact time > temperature > pH > initial concentration. The best network structure of 4-8-1 was utilized in BP-ANN modeling. The results of sensitivity analysis showed that the factors of NH_4^+ removal rate were in the order of: Contact time (31.23%) > pH (24.68%) > temperature (22.93%) > initial concentration (21.16%). The process input factors were optimized by GA-ANN and BBD-RSM for the optimum NH_4^+ removal rate. The predicted

results were verified by experiments. According to the results, the predicted values of GA-ANN were in better agreement with the experimental values. The optimum level of contact time, initial concentration, temperature, pH is 899.41 min, 17.35 mg/L, 15 °C, 6.98, respectively, under which condition, the maximum NH_4^+ removal rate is achieved 63.74%. The proposed method is effective for optimizing the process parameters of NH_4^+ removal from stormwater by CB-GAC, and is helpful to reduce the time and cost of experiments. In the future research, this method can be applied to the parameter optimization and efficiency prediction of the actual stormwater treatment process.

Author Contributions: A.Y. were responsible for the experimental and software work. H.L. helped to build the model. X.L., Y.Y. and Z.Z. supervised the laboratory work. Y.L. led the research. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the The National Key Research and Development Program of China (Grant No. 2018YFC0406203).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data used to support the findings of this study are available from the corresponding author upon requests.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Lv, H.; Yang, L.; Zhou, J.; Zhang, X.; Wu, W.; Li, Y.; Jiang, D. Water Resource Synergy Management in Response to Climate Change in China: From the Perspective of Urban Metabolism. *Resour. Conserv. Recycl.* **2020**, *163*, 105095. [CrossRef]
2. Wei, F.; Zhang, X.; Xu, J.; Bing, J.; Pan, G. Simulation of Water Resource Allocation for Sustainable Urban Development: An Integrated Optimization Approach. *J. Clean. Prod.* **2020**, *273*, 122537. [CrossRef]
3. Jiang, Y.; Zevenbergen, C.; Ma, Y. Urban Pluvial Flooding and Stormwater Management: A Contemporary Review of China's Challenges and "Sponge Cities" Strategy. *Environ. Sci. Policy* **2018**, *80*, 132–143. [CrossRef]
4. Ray, J.R.; Shabtai, I.A.; Teixidó, M.; Mishael, Y.G.; Sedlak, D.L. Polymer-Clay Composite Geomedia for Sorptive Removal of Trace Organic Compounds and Metals in Urban Stormwater. *Water Res.* **2019**, *157*, 454–462. [CrossRef]
5. LeFevre, G.H.; Paus, K.H.; Natarajan, P.; Gulliver, J.S.; Novak, P.J.; Hozalski, R.M. Review of Dissolved Pollutants in Urban Storm Water and Their Removal and Fate in Bioretention Cells. *J. Environ. Eng.* **2015**, *141*, 04014050. [CrossRef]
6. Pourret, O.; Hursthouse, A. It's Time to Replace the Term "Heavy Metals" with "Potentially Toxic Elements" When Reporting Environmental Research. *Int. J. Environ. Res. Public Health* **2019**, *16*, 4446. [CrossRef]
7. Urban Stormwater Toxic Pollutants: Assessment, Sources, and Treatability—Pitt—1995—Water Environment Research—Wiley Online Library. Available online: https://onlinelibrary.wiley.com/doi/epdf/10.2175/106143095X131466 (accessed on 16 December 2020).
8. Krakat, N.; Anjum, R.; Dietz, D.; Demirel, B. Methods of Ammonia Removal in Anaerobic Digestion: A Review. *Water Sci. Technol.* **2017**, *76*, 1925–1938. [CrossRef] [PubMed]
9. Wang, M.; Xie, R.; Chen, Y.; Pu, X.; Jiang, W.; Yao, L. A Novel Mesoporous Zeolite-Activated Carbon Composite as an Effective Adsorbent for Removal of Ammonia-Nitrogen and Methylene Blue from Aqueous Solution. *Bioresour. Technol.* **2018**, *268*, 726–732. [CrossRef] [PubMed]
10. Liu, Y.; Ngo, H.H.; Guo, W.; Peng, L.; Wang, D.; Ni, B. The Roles of Free Ammonia (FA) in Biological Wastewater Treatment Processes: A Review. *Environ. Int.* **2019**, *123*, 10–19. [CrossRef]
11. Provolo, G.; Perazzolo, F.; Mattachini, G.; Finzi, A.; Naldi, E.; Riva, E. Nitrogen Removal from Digested Slurries Using a Simplified Ammonia Stripping Technique. *Waste Manag.* **2017**, *69*, 154–161. [CrossRef] [PubMed]
12. Stefán, D.; Erdélyi, N.; Izsák, B.; Záray, G.; Vargha, M. Formation of Chlorination By-Products in Drinking Water Treatment Plants Using Breakpoint Chlorination. *Microchem. J.* **2019**, *149*, 104008. [CrossRef]
13. Two-Sectional Struvite Formation Process for Enhanced Treatment of Copper–Ammonia Complex Wastewater. *Trans. Nonferrous Met. Soc. China* **2017**, *27*, 457–466. [CrossRef]
14. Cheng, H.; Zhu, Q.; Xing, Z. Adsorption of Ammonia Nitrogen in Low Temperature Domestic Wastewater by Modification Bentonite. *J. Clean. Prod.* **2019**, *233*, 720–730. [CrossRef]
15. Ko, D.; Mines, P.D.; Jakobsen, M.H.; Yavuz, C.T.; Hansen, H.C.B.; Andersen, H.R. Disulfide Polymer Grafted Porous Carbon Composites for Heavy Metal Removal from Stormwater Runoff. *Chem. Eng. J.* **2018**, *348*, 685–692. [CrossRef]
16. Xiao, F.; Bedane, A.H.; Mallula, S.; Sasi, P.C.; Alinezhad, A.; Soli, D.; Hagen, Z.M.; Mann, M.D. Production of Granular Activated Carbon by Thermal Air Oxidation of Biomass Charcoal/Biochar for Water Treatment in Rural Communities: A Mechanistic Investigation. *Chem. Eng. J. Adv.* **2020**, *4*, 100035. [CrossRef]

17. Björklund, K.; Li, L.Y. Adsorption of Organic Stormwater Pollutants onto Activated Carbon from Sewage Sludge. *J. Environ. Manag.* **2017**, *197*, 490–497. [CrossRef] [PubMed]
18. Xie, Q.; Zhang, X.; Liang, D.; Cao, J.; Liu, J. Directional preparation of coal-based activated carbon principles, approaches and applications. *Coal Sci. Technol.* **2020**, 1–28. [CrossRef]
19. Ghaedi, A.M.; Vafaei, A. Applications of Artificial Neural Networks for Adsorption Removal of Dyes from Aqueous Solution: A Review. *Adv. Colloid Interface Sci.* **2017**, *245*, 20–39. [CrossRef] [PubMed]
20. Montgomery, D. *Design and Analysis of Experiments*, 2nd ed.; John Wiley & Sons: Hoboken, NJ, USA, 2000; Volume 16, ISBN 978-0-471-31649-7.
21. Roy, S.; Sengupta, S.; Manna, S.; Das, P. Chemically Reduced Tea Waste Biochar and Its Application in Treatment of Fluoride Containing Wastewater: Batch and Optimization Using Response Surface Methodology. *Process Saf. Environ. Prot.* **2018**, *116*, 553–563. [CrossRef]
22. Mourabet, M.; Rhilassi, A.E.; Bennani-Ziatni, M.; Taitai, A. Comparative Study of Artificial Neural Network and Response Surface Methodology for Modelling and Optimization the Adsorption Capacity of Fluoride onto Apatitic Tricalcium Phosphate. *Univers. J. Appl. Math.* **2014**, *2*, 84–91. [CrossRef]
23. Aghaeinejad-Meybodi, A.; Ebadi, A.; Shafiei, S.; Khataee, A.; Kiadehi, A.D. Degradation of Fluoxetine Using Catalytic Ozonation in Aqueous Media in the Presence of Nano-γ-Alumina Catalyst: Experimental, Modeling and Optimization Study. *Sep. Purif. Technol.* **2019**, *211*, 551–563. [CrossRef]
24. Fan, M.; Hu, J.; Cao, R.; Ruan, W.; Wei, X. A Review on Experimental Design for Pollutants Removal in Water Treatment with the Aid of Artificial Intelligence. *Chemosphere* **2018**, *200*, 330–343. [CrossRef]
25. Ghasemzadeh, K.; Ahmadnejad, F.; Aghaeinejad-Meybodi, A.; Basile, A. Hydrogen Production by a PdAg Membrane Reactor during Glycerol Steam Reforming: ANN Modeling Study. *Int. J. Hydrogen Energy* **2018**, *43*, 7722–7730. [CrossRef]
26. Rajendra, M.; Jena, P.C.; Raheman, H. Prediction of Optimized Pretreatment Process Parameters for Biodiesel Production Using ANN and GA. *Fuel* **2009**, *88*, 868–875. [CrossRef]
27. Yahya, H.S.M.; Abbas, T.; Amin, N.A.S. Optimization of Hydrogen Production via Toluene Steam Reforming over Ni–Co Supported Modified-Activated Carbon Using ANN Coupled GA and RSM. *Int. J. Hydrogen Energy* **2020**. [CrossRef]
28. Zulfiqar, M.; Samsudin, M.F.R.; Sufian, S. Modelling and Optimization of Photocatalytic Degradation of Phenol via TiO2 Nanoparticles: An Insight into Response Surface Methodology and Artificial Neural Network. *J. Photochem. Photobiol. Chem.* **2019**, *384*, 112039. [CrossRef]
29. Van Pham, T.; Van Tran, T.; Duy Nguyen, T.; Thi Hong Tham, N.; Thanh Tri Quang, P.; Thi To Uyen, D.; Thi Hong Le, N.; Vo, D.-V.N.; Trung Thanh, N.; Giang Bach, L. Development of Response Surface Methodology for Optimization of Congo Red Adsorption Utilizing Exfoliated Graphite As An Efficient Adsorbent. *Mater. Today Proc.* **2020**, *22*, 2341–2350. [CrossRef]
30. Hornik, K.; Stinchcombe, M.; White, H. Multilayer Feedforward Networks Are Universal Approximators. *Neural Netw.* **1989**, *2*, 359–366. [CrossRef]
31. Artificial Neural Networks for Small Dataset Analysis—PubMed. Available online: https://pubmed.ncbi.nlm.nih.gov/26101654/ (accessed on 15 December 2020).
32. Zarei, M.; Khataee, A.R.; Ordikhani-Seyedlar, R.; Fathinia, M. Photoelectro-Fenton Combined with Photocatalytic Process for Degradation of an Azo Dye Using Supported TiO_2 Nanoparticles and Carbon Nanotube Cathode: Neural Network Modeling. *Electrochim. Acta* **2010**, *55*, 7259–7265. [CrossRef]
33. Zhang, Y.; Pan, B. Modeling Batch and Column Phosphate Removal by Hydrated Ferric Oxide-Based Nanocomposite Using Response Surface Methodology and Artificial Neural Network. *Chem. Eng. J.* **2014**, *249*, 111–120. [CrossRef]
34. Chang, M.-W.; Chung, C.-C.; Chern, J.-M.; Chen, T.-S. Dye Decomposition Kinetics by UV/H2O2: Initial Rate Analysis by Effective Kinetic Modelling Methodology. *Chem. Eng. Sci.* **2010**, *65*, 135–140. [CrossRef]
35. Bhatti, M.S.; Kapoor, D.; Kalia, R.K.; Reddy, A.S.; Thukral, A.K. RSM and ANN Modeling for Electrocoagulation of Copper from Simulated Wastewater: Multi Objective Optimization Using Genetic Algorithm Approach. *Desalination* **2011**, *274*, 74–80. [CrossRef]
36. Supong, A.; Bhomick, P.C.; Baruah, M.; Pongener, C.; Sinha, U.B.; Sinha, D. Adsorptive Removal of Bisphenol A by Biomass Activated Carbon and Insights into the Adsorption Mechanism through Density Functional Theory Calculations. *Sustain. Chem. Pharm.* **2019**, *13*, 100159. [CrossRef]
37. Foo, K.Y.; Hameed, B.H. Insights into the Modeling of Adsorption Isotherm Systems. *Chem. Eng. J.* **2010**, *156*, 2–10. [CrossRef]
38. Lucaci, L.; Bulgariu, D.; Ahmad, I.; Lisa, G.; Mocanu, A.; Bulgariu, L. Potential Use of Biochar from Various Waste Biomass as Biosorbent in Co (II) Removal Processes. *Water* **2019**, *11*, 1565. [CrossRef]
39. Peiris, C.; Gunatilake, S.R.; Mlsna, T.E.; Mohan, D.; Vithanage, M. Biochar Based Removal of Antibiotic Sulfonamides and Tetracyclines in Aquatic Environments: A Critical Review. *Bioresour. Technol.* **2017**, *246*, 150–159. [CrossRef]
40. Teong, C.Q.; Setiabudi, H.D.; El-Arish, N.A.S.; Bahari, M.B.; Teh, L.P. Vatica Rassak Wood Waste-Derived Activated Carbon for Effective Pb (II) Adsorption: Kinetic, Isotherm and Reusability Studies. *Mater. Today Proc.* **2020**. [CrossRef]
41. Bello, M.O.; Abdus-Salam, N.; Adekola, F.A.; Pal, U. Isotherm and Kinetic Studies of Adsorption of Methylene Blue Using Activated Carbon from Ackee Apple Pods. *Chem. Data Collect.* **2021**, *31*, 100607. [CrossRef]
42. Ren, Z.; Jia, B.; Zhang, G.; Fu, X.; Wang, Z.; Wang, P.; Lv, L. Study on Adsorption of Ammonia Nitrogen by Iron-Loaded Activated Carbon from Low Temperature Wastewater. *Chemosphere* **2021**, *262*, 127895. [CrossRef]

43. Muralidhar, R.V.; Chirumamila, R.R.; Marchant, R.; Nigam, P. A Response Surface Approach for the Comparison of Lipase Production by Candida Cylindracea Using Two Different Carbon Sources. *Biochem. Eng. J.* **2001**, *9*, 17–23. [CrossRef]
44. Shin, Y.; Kim, Z.; Yu, J.; Kim, G.; Hwang, S. Development of NOx Reduction System Utilizing Artificial Neural Network (ANN) and Genetic Algorithm (GA). *J. Clean. Prod.* **2019**, *232*, 1418–1429. [CrossRef]