Okiy, K. V., Nwabanne, J. T. (2024), JOTCSB, 7(1), 25-42. **RESEARCH ARTICLE**

Optimizing Activated Carbon Production from Waste Cashew Nut Shell with Zinc Chloride: A Box-Behnken Design and Group Method of Data Handling (GMDH) Application

Karinate Valentine Okiy and Joseph Tagbo Nwabanne

Department of Chemical Engineering, Nnamdi Azikiwe University, P.M.B. 5025, Awka 420218, Nigeria

Abstract: In this study, Response surface methodology (RSM) and innovative Group Method of Data Handling (GMDH) approaches are applied to investigate the optimal process conditions of Zinc Chloride activated cashew nut production process. The effects of activation conditions (i.e. activation temperature, activation time, and impregnation ratio) on the achievable BET surface areas were studied with the aid of Box Behnken Design (BBD) and GMDH. Comparative analysis of RSM and GMDH-type neural models were further researched. During the process, the polynomial model equations developed were modified and fine-tuned to predict the highest BET surface area(s) using regression analysis and GMDH multi-layered iterative algorithm (MIA). Analysis of Variance (ANOVA) revealed that the significant factor(s) were impregnation ratio, impregnation ratio product and the 2-way interactions (activation temperature and impregnation ratio) for $ZnCl_2$ activated cashew nut shell. The best activation conditions for producing highest BET surface area of 504 $m^2.g^{-1}$ was activation temperature (873K), activation time (60 min), and impregnation ratio (1.50).The proposed GMDH-type BET model was ascertained to be the best model with average correlation coefficient (R) and root mean square error (RMSE) of 0.925 & 32.0 respectively. Sensitivity analysis conducted for GMDH-type neural network also revealed that the activation temperature and activation time with sensitivity values of 90.6% and 74.1% respectively were the most influential parameters in the basic ($ZnCl_2$) activation process. The results of this study show that RSM and GMDH-type neural network could be applied as effective analytical tools for optimizing the ZCNS manufacturing process.

Keywords: Cheaper Agro-wastes derived Porous Carbons, Box Behnken Design (BBD), Basic activation, Group Method of Data Handling (GMDH), Process Optimization.

Submitted: January 17, 2024. **Accepted:** February 20, 2024.

Cite this: Okiy, K. V., & Nwabanne, J. T. (2024). Optimizing Activated Carbon Production from Waste Cashew Nut Shell with Zinc Chloride: A Box-Behnken Design and Group Method of Data Handling (GMDH) Application. Journal of the Turkish Chemical Society, Section B: Chemical Engineering, 7(1), 25–42. https://doi.org/10.58692/jotcsb.1421501.

Corresponding author. Email: karinateokiy@gmail.com

1. INTRODUCTION

Industrial activities are major sources of effluent discharges into rivers, water bodies, streams and reservoirs causing serious environmental pollution (Howard et al., 1986). Many industries including iron and steel, battery manufacturing, mining, petrochemicals/refineries, tanneries, microelectronics, non-ferrous metals, textile & leather, breweries, metal processing, pharmaceutical, photographic, glassware, electroplating, paints, pulp & paper, pesticides manufacturing, dyeing, breweries, ceramics and chemical manufacturing generate significant volumes of effluent containing pollutants like biodegradable organics, cyanide, dissolved inorganic solids, suspended solids, phenols, chlorinated organic compounds, refractory organics, mineral oil, polychlorinated biphenyls (PCBs), dyes and heavy metals (Shi, 2009; Dawei, 2012; Ajemba, 2014; Kulkarni et al., 2014). Heavy metals are metals with molecular weight in the range of 63.5 to 200.6 and density greater than 5 $g.cm^{-3}$ (Wang, 2011). Toxic heavy metals such as lead, chromium, mercury, arsenic & nickel are considered as major pollutants of great environmental significance (Wang., 2011; Singh et al., 2011). Wastewaters contaminated with toxic heavy metals ions require exhaustive treatment in order to remove the toxic metal ions down to minimal (trace) outlet concentrations before discharge into receiving water bodies for economic and environmental reasons (Okiy, 2006).

Numerous conventional separation techniques exist for the removal of toxic heavy metals from aqueous solutions/wastewater and these include adsorption, chemical precipitation, freeze crystallization, gamma irradiation, foam flotation, ultrafiltration, ozonation, vapour recovery, reverse osmosis, membrane separation, solvent extraction, ion exchange, electrodialysis, coagulation, evaporation, nanofiltration, flocculation, electrochemical process and biosorption methods. Whilst some of these treatment technologies have proven successful for treatment of heavy metal-contaminated wastewater, other methods have a number of downsides limiting their use such as fouling, high capital, and operational costs as well as the generation of bulk quantities of chemical sludge requiring proper disposal into the environment (Demirbas, 2008; Nwabanne and Okoye., 2013; Srivastava et al., 2015; Malik et al., 2016). Amid these, adsorption method is considered as an excellent option and is widely utilized for the treatment of effluent emanating from chemical industries (Akinbiyi, 2000; Xu et al., 2013). Activated carbon (AC) is the adsorbent of choice worldwide due to its high efficiency and versatility for adsorbing different pollutants present in municipal and industrial wastewater (Gupta et al., 2009; Malik et al., 2016). However, the price of commercial activated carbons normally produced from naturally occurring raw materials (precursors) such as lignite, peat, wood, petroleum residues, anthracite, and coal varies from 0.8 to 10 Euros per kilogram depending on process type and the activation method utilized (Stavropoulos and Zabaniotou, 2009). In order to reduce these exorbitant material costs, there is need for concerted efforts to be directed towards production of low-cost activated carbons from agro-wastes such as cashew nut shells.

Cashew nut (Anarcadium occidentale L.) is predominantly found in Asian, African, and Latin American countries such as Vietnam, India, Tanzania, Cote d'Ivoire, Nigeria, and Brazil. Nigeria is the world's second largest producer of cashew nuts with an annual production of 836,500 Metric Tonnes (Adeigbe et al., 2015). Cashew nuts are commonly eaten as a snack and utilized in preparation of meals, desserts, and confectioneries. Nigeria is also a major exporter of cashew nuts, with a significant portion of the produce being sold to international markets. In fact, Nigeria is one of the largest producers and exporters of cashew nuts in the world. The export market provides income opportunities for farmers and contributes about 24 billion naira to Nigeria's foreign exchange earnings. Nigeria is a major consumer of cashew nuts, both for domestic consumption and as an ingredient in various dishes (Adeigbe et al., 2015). Cashew nut shells are a byproduct of cashew processing and are typically discarded. The waste shells are hard and contain a toxic resin called cardol, which can cause skin irritation and other health issues if handled improperly. Improper disposal of the shells can lead to land and water pollution, as well as create breeding grounds for insects and pests. Thus, the proper disposal and management of cashew nut shells is a major area of environmental concern in Nigeria that requires attention and sustainable solutions. Efforts have been made to find alternative uses for cashew nut shells to mitigate these disposal challenges. For example, some companies and entrepreneurs have explored the utilization of cashew nut shells for energy generation, production of industrial materials like adhesives, or conversion into animal feed and organic fertilizers (Ademola et al., 2021). However, the scale of these initiatives is relatively limited. Hitherto, few studies utilizing cashew nut shells for the production of activated carbons have been conducted in the past. In this study, cashew nut shell was chosen for the production of activated carbon.

According to Dyk (2000), the predominant characteristic for classifying activated carbons industrially is BET surface area. The manufacture of activated carbon is also affected by several process variables (factors) including activation temperature, impregnation ratio, and activation time (Essa et al., 2013). The classical approach in dealing with many factors is the one factor at a time (OFAT) experimentation. This method entails studying the effect of each parameter on the response of interest by successively varying each parameter within a specified ambit, while maintaining the other parameters constant at the zero (median) level (Elibol, 2002). OFAT is painstakingly tedious and does not provide any information on the interaction (combined) effects of the process variables (factors) (Onu et al., 2021). Instead of considering each factor in isolation, it is also possible to combine the series of independent studies into one study.

Response Surface Methodology (RSM) is a useful tool for studying the interaction of two or more parameters (factors) on a response. RSM is a compendium of mathematical and statistical techniques used for process development, upgrading and optimization. RSM approach encompasses four major stages (i) the design and conduction of experiments for measuring the studied response (ii) response surface modeling via non-linear model fitting over experimental data (iii) visualization of the interactive and main effects of the independent variables on the studied response via two-dimensional (2-D) & three-dimensional (3-D) plots and, (iii) process optimization (Essa et al., 2013). A major benefit of RSM is the minimal material costs and less number of experimental runs needed to assess main and interaction effects of several parameters for the system under consideration (Montgomery, 2017). Interaction effects of the various factors could be attained using RSM with design of experiments (DoE). Box Behnken design (BBD) is one of the more established statistical experimental designs for RSM modeling (Okewale et al., 2015). BBD comprises of the factorial 2^k level designs and incomplete block designs with a number of core factorial runs in the design (n_F), number of experimental runs at the high and/or low values of the factors in the design (n_i) and the number of runs at the centre (median) point values of the factors in the design (n_0) (Okewale et al., 2015). The BBD design is more suitable when the number of factor(s) is between 3 and 4, as minimal numbers of experimental runs are needed for optimal response (Ozer et al., 2009; Montgomery., 2017). However, statistical methods such as RSM are difficult to implement for modeling complex and non-linear real-world problems. Unlike statistical approaches, other data driven models such as artificial neural network (ANN) feature exceptional ability to handle complex (noisy) non-linear problems (Palani et al., 2009; Nkurlu et al., 2020). Nonetheless, there are a number of drawbacks associated with standard ANN including poor model generalization ability (over-fitting), slow trial & error process associated with user-led specification of network architecture (i.e hidden layers, hidden nodes), and network algorithm convergence at local minimum. The group method of data handling (GMDH) is an improvement over the ANN technique, due to the GMDH algorithm's self-organizing (intelligent) control in identification of optimal network structure and process modeling (Li et al., 2017).

Group method of data handling (GMDH) is a grouping of mathematical modeling and non-linear regression algorithms (polynomial neural networks) for computational-based modeling of experimental dataset(s) characterised by fully automatic determination of model structure, hidden nodes, number of hidden levels and parametric model optimization (Madala & Ivakhneko., 1994; Voss., 2002). Complex non-linear systems with several inputs and one output are modeled with GMDH employing the Kolmogarov-Gabor polynominal given as:

$$Y = a_0 + \sum_{i=1}^{m} a_i x_i + \sum_{i=1}^{m}\sum_{j=1}^{m} a_{ij} x_i x_j + \sum_{i=1}^{m}\sum_{j=1}^{m}\sum_{i=1}^{n} a_{ijk} x_i x_j x_k \ldots \qquad \text{(Eq. 1)}$$

The GMDH method was initially proposed in 1968 at the Institute of Cybernetics, Kyiv by an Ukrainian scientist and mathematician named Professor Alexey Grigorevich Ivakhnenko (Stanley, 1981). The GMDH analysis protocol is based on an amalgamation of (1) the neural approach which uses the threshold logic (selection criterion) & network connectionism and, (2) the black box concept, which examines non-linear input-output variable relationships (Nkurlu et al., 2020). The external criterion describes the requirements for selection of a model of optimal complexity that unravels the hidden law from input data. Whilst, network connectionism defines the mapping accuracy of the region between input and output dataset(s)(Li et al., 2017).

GMDH-type neural network is computationally faster, circumvents over-fitting problems, objectively selects the optimum model, and eliminates the slow trial & error process of optimal algorithmic (i.e number of hidden layers, neurons) parameters selection required for efficient neural network design (Nkurlu et al., 2020). The novel GMDH has never been utilized before for process modeling and optimization in the field of engineering. This marks the first time that GMDH is applied to model a chemical process such as the production of activated carbon from cashew nutshell (CNS) with chemical ($ZnCl_2$) activation. In this present study, RSM (BBD) and GMDH methods will be employed to analyze the alkaline CNS activation process and predict the maximum BET surface area of $ZnCl_2$-activated cashew nut shell (CNS).

Therefore, this work aims to (i) explore the preparation of activated carbons from cashew nut shells with chemical ($ZnCl_2$) activation (ii) select optimal CNS activated carbons with maximum BET surface areas based on RSM (BBD) design (iii) utilise GMDH technique for optimizing the manufacture of $ZnCl_2$ activated CNS under different preparation conditions of impregnation ration, time, and activation temperature (iv) comparatively analyse the $ZnCl_2$-activated CNS production process using BBD with the RSM approach and a GMDH-type neural network to determine the optimal conditions for maximum achievable BET surface area.

2. MATERIALS AND METHODS

2.1. Adsorbent preparation

The raw cashew nut shells were procured from Eke-Awka market at Awka South Local Government Area, Anambra State in the Eastern part of Nigeria (N: 6^0 13′ 8″; E: 7^0 5′ 13″). Chemical activation of cashew nut shell sample with Zinc chloride as reagent was performed according to the chemical activation procedure reported by Senthil Kumar et al., (2012) and Subramaniam & Ponnusamy, (2015) with minor modification. The experimental procedure for producing the cashew nut shell activated carbon is presented in Figure 1. 300 grams of raw cashew nut shell sample was pre-treated by washing with 5000 grams of distilled water and 3945 grams of ethanol to remove dirt and other soluble impurities. Consequently, the washed avocado pear seed samples were dried in a Mermmert oven at a temperature of 343K for 24 hours. Then, the dried cashew nut shell samples were ground into fine particles utilising a Jencod grinding machine and sieved using a standard Taylor Sieve with mesh size of 300 µm. 300g of cashew nut shell sample was then carbonised at a 973K for 2 hours. Thereafter, 300 grams of carbonised cashew nut shell was mixed with 88.4 grams of 30% Zinc chloride reagent according to the impregnation ratios (0.5:1, 1:1, and 1.5:1) for 2 hours and heated in a Mermmert oven at 378K for 24 hours. The dried zinc chloride impregnated cashew nut shell was then thermally activated in a Muffle furnace at different activation temperatures (873K, 1023K, and 1173K) and times (60, 90, and 120 minutes). The $ZnCl_2$ treated sample was cooled, and repeatedly washed to remove disorganised carbon, products of decomposition and traces of sulphuric acid. The activated carbon samples were also transferred to a beaker containing 250 ml solution of HCl (about 0.1mol) for 1 hour and washed again with distilled water till a pH of 6-7 is achieved. The washed activated carbon sample was filtered with Whatman No.1 filter paper and dried in a Mermmert oven at 353K for 3 hours before usage. Textural characterisation of the Zinc chloride activated carbon sample was performed using a Quantachrome NOVA4200e BET Analyzer (Anton-Paar GmbH, Austria) to determine the texture feature (BET surface area).

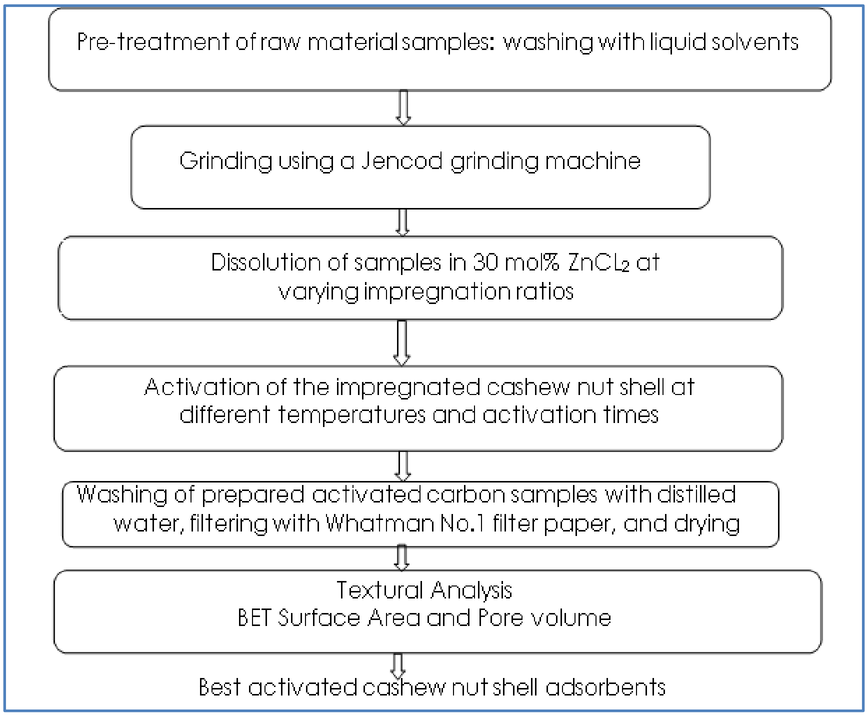


Figure 1: Flowchart showing the method of producing Zinc chloride activated carbon from cashew nut shells.

2.1.1. Design of experiment (DOE) for APS activation by BBD modeling

Box Behnken (BBD) design was chosen for accurate and precise experimental data collection in only few runs to allow for adequate approximation of the response surface (Montgomery, 2017). The effect of independent variables (activation temperature, time and impregnation ratio) on the production of optimal APS activated carbons with highest achievable BET surface area (response variable) was studied using Box Behnken (BBD) design. The lower and upper limits of the independent variables (factors) chosen for the BBD are shown in Table 1.

The ranges of the independent variables were chosen based on literature reviews and prior studies (Essa et al., 2013; Buasri et al., 2023). For the BBD design, the total number of experimental runs was determined using Eq. (2) (Melvin et al., 2015).

$$N = K^2 + K + +C_P \quad \text{Eq. (2)}$$

Where, N is the number of experimental runs, C_P is the replicate number of the central point, and K is the factor number.

Table 1: Independent Variables and Their Levels for Box Behnken Experimental design.

Independent Variable(s)	Range and Level		
	-1	0	+1
Activation temperature (A, ^0C)	600	750	900
Activation time (B, min)	60	90	120
Impregnation ratio (C, Activating agent: raw material)	0.5	1.0	1.5

To improve the effectiveness of the BBD design and minimize random errors, the factorial and incomplete block (IBD) experimental runs were performed in triplets. In accordant with 9 centre (null) points, leading to a total of 54 experimental runs were utilized in the response surface methodology (RSM) analysis. The factorial points gave an equal variation of the high and low values, while the null points gave the estimate of experimental error and ensured reproducibility of the data (Essa et al., 2013; Onu et al., 2021).

The Behnken design is an orthogonal design, with values of the experimental factors (points) at the midpoint of the edges and the centre of a multi-dimensional sphere defining the experimental domain (Douglas and Montgomery, 2007; Anderson and Whitcomb., 2016). Box Behnken Experimental Design being of spherical (cubic) design, the independent variables were coded for low, medium, and high settings, as -1, 0, and +1 and uniformly spaced (Douglas and Montgomery, 2007) as afore-presented in Table 1.

The statistical analysis was carried out using Design Expert Software version 13 (STAT-EASE Inc, USA) to predict the response of the studied system. The fifteen (15) experimental runs were performed in a random manner to avert systematic error. The analysis of variance (ANOVA) test was also used to evaluate the accuracy of the approximating polynomial model.

Furthermore, the reduction empirical model utilized in describing the $ZnCl_2$–CNS fabrication process is represented by the ensuing second-order approximating polynomial model equation, in terms of coded factors:

$$Y = b_0 + \sum_{i=1}^{n} b_i X_i + \sum_{i=1}^{n} b_{ii} X_i^2 + \sum_{i=1}^{n-1}\sum_{i=2}^{n} b_{i,j} X_i X_j + E \qquad \text{Eq. (3)}$$

Where, Y is the predicted dependent variable, b_0 is the constant coefficient (intercept), n is the number of patterns, b_i, b_{ij}, and b_{ii} are the regression coefficients of the linear, and interaction terms respectively, X_i, and X_j are the independent factors studied, i, and j are index numbers, and E is the error term.

2.1.2. GMDH-type neural network

GMDH-type neural network is structured as a feed-forward neural network with multi-layered bi-nodal polynomial activation function(s) as depicted in Figure 2. GMDH can be presumed to be a polynomial network (Ayoub et al., 2019). The GMDH analysis protocol harnesses the benefits of multi-layered neural networks and self-organizing ability to select the intrinsic affiliation(s) between input data and network (output) predictions.

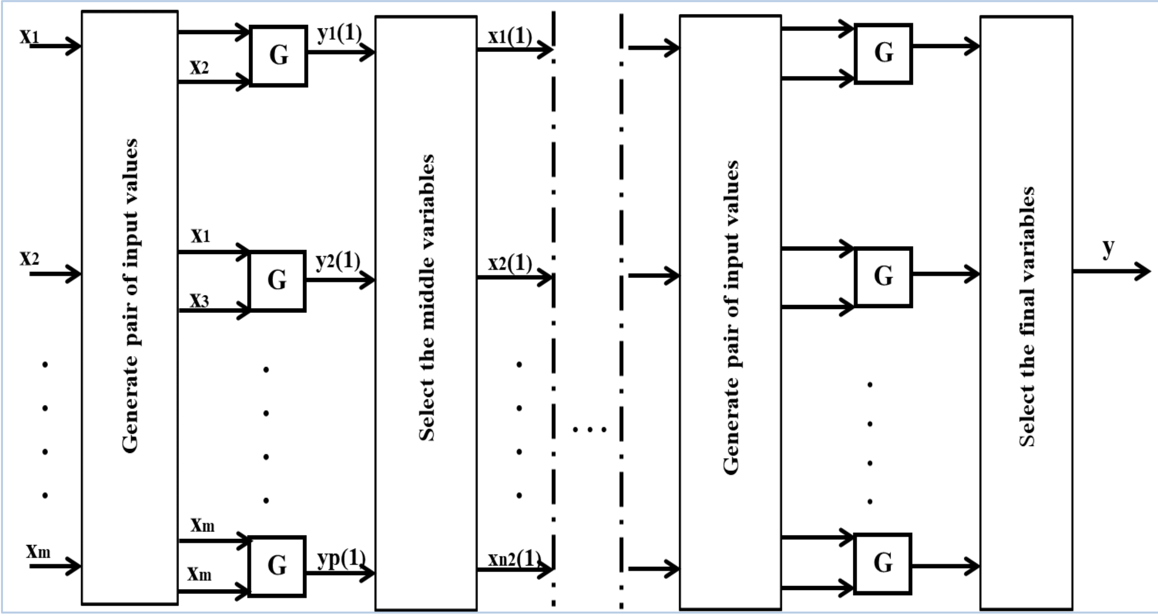


Figure 2: Schematic diagram of GMDH algorithm (Adapted from Li et al., 2017 p.9).

In Figure 2, x_i is the initial input variable, $y_i(k)$ is the output of the partial polynomials. G is the partial polynomials with quadratic polynomial for each two input variables. The partial polynomials are obtained by fitting the input data. $x_i(k)$ is the mediator, filtered by each layer criteria from $y_i(k)$, used as the input variable in the subsequent layer (Li et al., 2017).

During GMDH network design, the input dataset is split into training and test dataset(s). The training data is utilised to estimate the coefficients of the polynomial model, and the test data to choose the optimum model structure. The least squares minimization method is then applied to determine the polynomial coefficients for each of the nodal models. When all the possible outputs have been calculated, the estimated values of the neurons are crosschecked against the test data. The neurons whose mean-square-errors (MSE) are below the threshold (external criterion) were selected and the neuron with the smallest mean-square-error attained ($minMSE_1$) also retained. The retained nodes are combined together to provide another set of inputs into the succeeding layer. The new layer's nodes are checked for compliance with the external criterion (MSE). The neurons with acceptable measure of fit are kept and the remaining neurons disposed. The best neurons are combined and assigned to nodes in the subsequent layer. Thereafter, the neurons estimated values are checked against test data to evaluate measure of fit and the process is continued iteratively until the optimal output node with smallest MSE achieved is selected. The quadratic polynomial connected to the optimal neuron is the non-physical model of the system (Voss, 2002; Nkurlu et al., 2020).

The experimental data collected for GMDH analysis was a statistic of evaluated BET surface areas for CNS activated carbon produced under different alkaline activation conditions. To improve GMDH modeling, the experimental dataset(s) were quadrupled to give a total of sixty (60) data-points (Onu et al., 2021), which were utilized in the GMDH analysis. The input dataset was subdivided into training and testing dataset(s), resulting in 51 training points utilized in estimation of the nodal polynomial equations, and 9 testing points utilized for ascertaining the fitness of the estimated polynomial equations in the network model.

The neural model training was carried out using GMDH source code which automates the training process.

3. RESULTS

3.1. Response surface modeling of CNS basic ($ZnCl_2$) activation process

A total of 54 experimental runs of Box Behnken (BBD) design were conducted by using the prepared activated carbon samples and the BET surface area(s) were measured. The Sequential Model Sum of Squares (SMSS) measures the desirability of each model for the activation process based on the Sum of Squares (SS) value, where appropriateness of the model for the process increases as the SS value increases (See Table 2). Consequently, the quadratic model was suggested by Design Expert Software for optimizing the achievable BET surface areas of produced $ZnCl_2$-activated carbon(s).

Table 2: Sequential Model Sum of Squares Analysis [Type 1] for $ZnCl_2$ activated cashew nut shell.

Source	Sum of Squares	df	Mean Square	F-value	p-value	
Mean	5671257.41	1	5671257.41			
Linear	62129.44	3	20709.81	3.43	0.024	
2FI	81834.17	3	27278.06	5.82	0.00183	
Quadratic	153816.44	3	51272.15	33.85	1.69e-11	Suggested
Cubic	66652.97	3	22217.66			
Residual	0	41	0			
Total	6035690.43	54	111772.05			

The main and interaction effects were estimated by performing the Analysis of Variance (ANOVA). The significance probability value (P-value) was utilized to determine the model terms that are statistically significant. If the P-value is less than 0.05 it is safe to conclude that the effect (factor) under consideration is significant at the 95% confidence level. The final model for the response was obtained by retaining only the significant factors ($P < 0.05$) based on the probability-test. The results of the ANOVA Analysis for the ZnCl2 activated carbons are presented in Table 3.

Table 3: ANOVA Analysis for $ZnCl_2$ activated carbon.

Source	Sum of Squares	df	Mean Square	F-value	P-value
Model	2.978E+05	9	33086.67	21.84	< 0.0001
A-Activation temperature	1968.78	1	1968.78	1.30	0.2604
B-Activation time	48356.97	1	48356.97	31.92	< 0.0001
C-Impregnation ratio	8638.04	1	8638.04	5.70	0.0213
AB	774.51	1	774.51	0.5113	0.4784
AC	80599.21	1	80599.21	53.21	< 0.0001
BC	0.2139	1	0.2139	0.0001	0.9906
A^2	79093.85	1	79093.85	52.21	< 0.0001
B^2	432.94	1	432.94	0.2858	0.5956
C^2	58504.03	1	58504.03	38.62	< 0.0001
Residual	66652.97	44	1514.84		
Lack of Fit	66652.97	3	22217.66		
Pure Error	0.0000	41	0.0000		
Cor Total	3.644E+05	53			

The estimated response for BET surface area of $ZnCl_2$ activated cashew nutshell (CNS) is represented by Eq. 4:

$$\text{BETArea} = -1870.73 + 6.11\times A - 3.70\times B + 201.59\times C + 0.00155\times A\times B - 0.946\times A\times C + 0.0086\times B\times C - 0.0035\times A^2 + 0.00648\times B^2 + 271\times C^2 \quad \text{Eq. (4)}$$

From the ANOVA for $ZnCl_2$-CNS presented in Table 3, it was observed that the P-value for the response surface model was less than 0.05 (p-value < 0.0001). This indicates that the second order polynomial model is significant at the 95% confidence level. likewise, from the ANOVA results shown in Table 4, it can be concluded that the linear effects, Activation temperature (A), Impregnation ratio (C), the 2-way interactions A*C (Impregnation ratio and Activation temperature), and the square effects A*A (Activation temperature product), C*C (Impregnation ratio product) had a statistically significant effect on the response, as all their P-values were not greater than 0.1 (Montgomery., 2017). On the other hand, the linear effect A (Activation temperature), 2-way interactions A*B (Activation temperature and Activation time), B*C (Activation time and Impregnation ratio), and the square effects B*B (Activation time product) were found to be insignificant as their P-values were greater than 0.1 (Montgomery, 2017). Consequently, the conclusion can be reached that the studied response does not have any correlation with a change in these linear or combined independent variables i.e. A, A*B, B*C and B*B at specified values of the other variables (B&C), C, A, and (A&C) for the range of variables examined in Table 3. Therefore, eliminating all the insignificant terms, the final model is obtained as:

$$\text{BETArea} = -1870.73 - 3.70\times B + 201.59\times C - 0.946\times A\times C - 0.0035\times A^2 + 271\times C^2 \qquad \text{Eq. (5)}$$

The results of predicted response and experimental response are shown in Table 4.

Table 4: Model Predicted and Experimental BET Surface area for $ZnCl_2$ -activated CNS.

Point	BET Surface Area (Predicted)	BET Surface Area (Actual)	STD Error Fit	Square Residual
1	297.920	356.52	1.67387	3433.95
2	299.693	285.53	-0.40455	200.58
3	202.097	216.26	0.40455	200.58
4	231.700	173.10	-1.67387	3433.95
5	223.647	174.60	-1.40100	2405.63
6	381.285	405.00	0.67741	562.41
7	400.215	376.50	-0.67741	562.41
8	273.953	323.00	1.40100	2405.63
9	428.153	418.60	-0.27310	91.254
10	345.987	392.50	1.36911	2163.45
11	462.513	416.00	-1.36911	2163.45
12	380.863	393.60	0.37312	162.23
13	330.800	330.80	-0.00000	1.292E-26
14	330.800	330.80	-0.00000	1.292E-26
15	330.800	330.80	-0.00000	1.292E-26
16	297.920	356.52	1.67387	3433.95
17	299.693	285.53	-0.40455	200.58
18	202.097	216.26	0.40455	200.58
19	231.700	173.10	-1.67387	3433.95
20	223.647	174.60	-1.40100	2405.63
21	381.285	405.00	0.67741	562.41
22	400.215	376.50	-0.67741	562.41
23	273.953	323.00	1.40100	2405.63
24	428.153	418.60	-0.27310	91.254
25	345.987	392.50	1.36911	2163.45
26	462.513	416.00	-1.36911	2163.45
27	380.863	393.60	0.37312	162.23
28	330.800	330.80	-0.00000	1.292E-26
29	330.800	330.80	-0.00000	1.292E-26
30	330.800	330.80	-0.00000	1.292E-26
31	297.920	356.52	1.67387	3433.95
32	299.693	285.53	-0.40455	200.58
33	202.097	216.26	0.40455	200.58
34	231.700	173.10	-1.67387	3433.95
35	223.647	174.60	-1.40100	2405.63
36	381.285	405.00	0.67741	562.41
37	400.215	376.50	-0.67741	562.41
38	273.953	323.00	1.40100	2405.63

Point	BET Surface Area (Predicted)	BET Surface Area (Actual)	STD Error Fit	Square Residual
39	428.153	418.60	-0.27310	91.25
40	345.987	392.50	1.36911	2163.46
41	462.513	416.00	-1.36911	2163.46
42	380.863	393.60	0.37312	162.23
43	330.800	330.80	-0.00000	1.292E-26
44	330.800	330.80	-0.00000	1.292E-26
45	330.800	330.80	-0.00000	1.292E-26
46	297.920	356.52	1.67387	3433.95
47	299.693	285.53	-0.40455	200.58
48	202.097	216.26	0.40455	200.58
49	231.700	173.10	-1.67387	3433.95
50	223.647	174.60	-1.40100	2405.63
51	381.285	405.00	0.67741	562.41
52	400.215	376.50	-0.67741	562.41
53	273.953	323.00	1.40100	2405.63
54	428.153	418.60	-0.27310	91.25
				RMSE= 35.1

The graphical plot of the measured BET surface area against predicted BET surface area is shown in Figure 3.

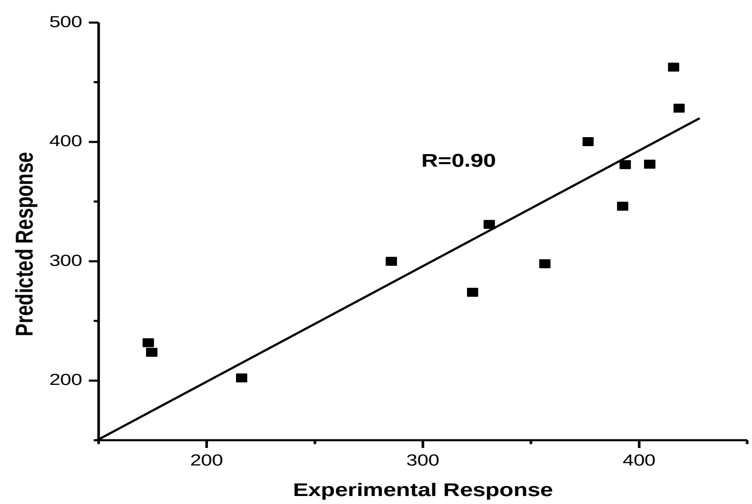


Figure 3: Plot of RSM predicted BET surface area against experimental data.

Figure 3 shows that the data points are un-evenly distributed about the 45^0 line without forming a definite pattern, indicating normality in the data values (minimum residual errors). Thus, signifying that the response surface model is adequate for predicting achievable surface area(s) of $ZnCl_2$-activated CNS. The high (correlation coefficient) R value of 0.90, confirmed that the achievable BET surface area for $ZnCl_2$-CNS activation can be predicted satisfactorily by the response model.

The normal plot of residuals depicted in Figure 4 was also utilised to check if the process data are normally distributed. The distribution of data points was similar at both the left and right portions of the plot, indicating normal distribution of the error residuals. This implies that there are no signs of problems with the process model or data (Antony, 2003).

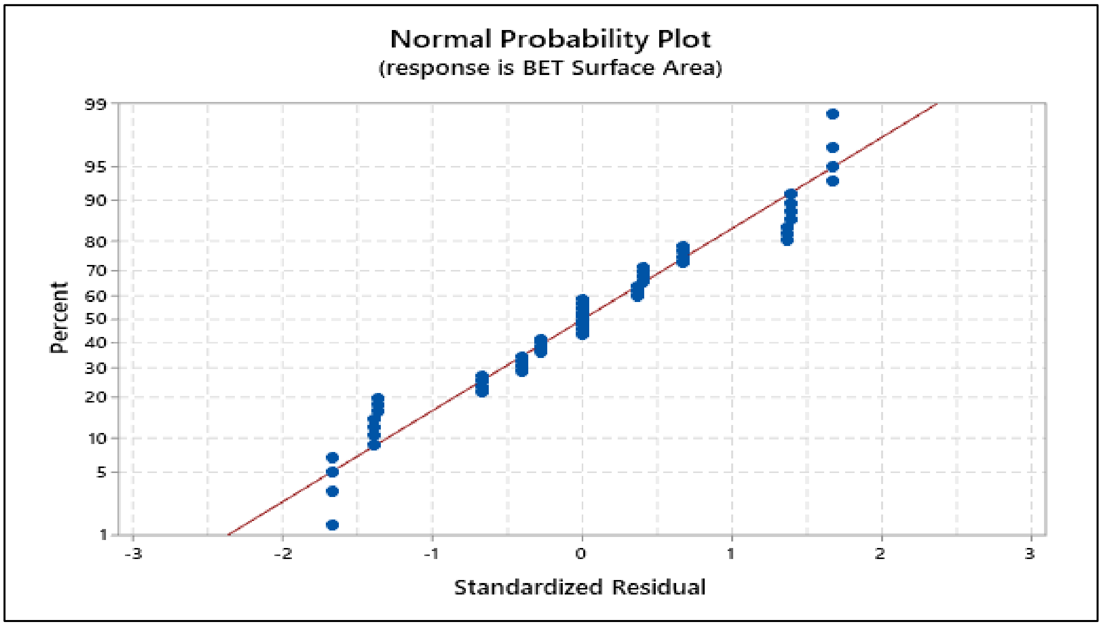


Figure 4: Normal Probability Plot of Residuals for the $ZnCl_2$ activated CNS.

The statistical significance of the estimated linear factors, interactions, and their products on the achievable BET surface area (studied response), in order of significance is represented by the Pareto diagram shown in Figure 5. The vertical (red) line indicates the magnitude of the least statistically significant effect for a 95% confidence level and the corresponding t-test value is equal to 2.015. Any factor or its interaction that transcends the vertical line is considered significant (Montgomery, 2017).

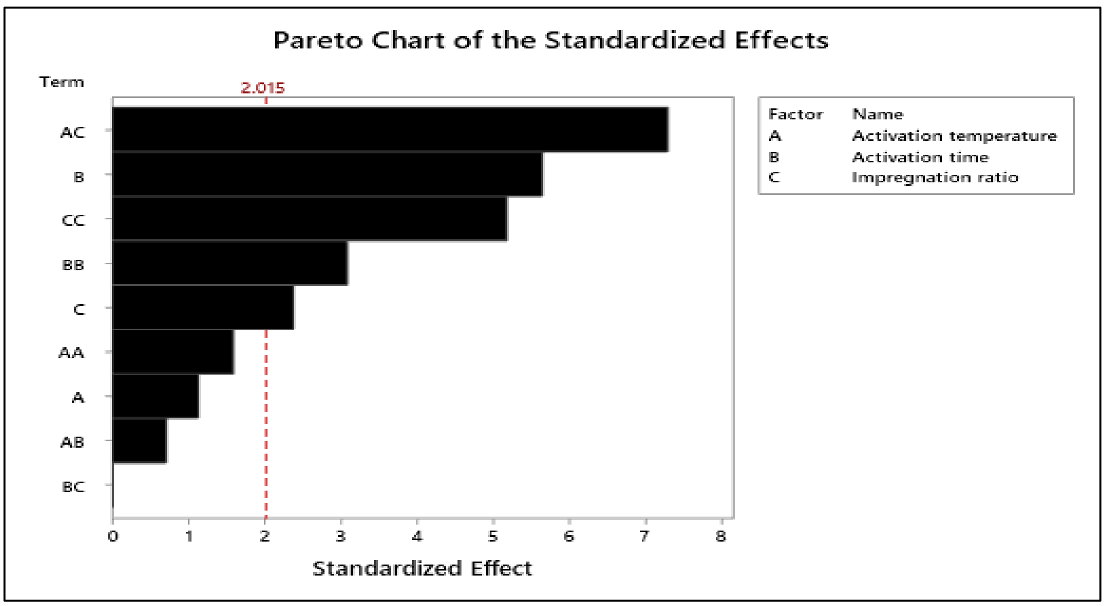


Figure 5: The Pareto plot for $ZnCl_2$ activated CNS.

Figure 5 shows that the most influential factors influencing the achievable BET surface area of $ZnCl_2$-activated CNS were the linear effect B (Activation time), 2-way interactions A*C (Activation temperature and Impregnation ratio), and the square effect C*C (Impregnation ratio product). Noteworthy, Prominent interactions are integral in attainment of overall process optimization.

The three dimensional (3-D) response surface plots generated to visualize the relationship between the independent variables (activation temperature, time, impregnation ratio) and response of interest (BET surface area) in terms of main, and interaction effects and also facilitate optimization of the $ZnCl_2$ activated CNS production process are showcased in Figure 6.

Each of the response diagrams were plotted as a function of two process variables in their respective ranges (-1 to 1), with the third independent variable(s) was maintained at median (zero) level.

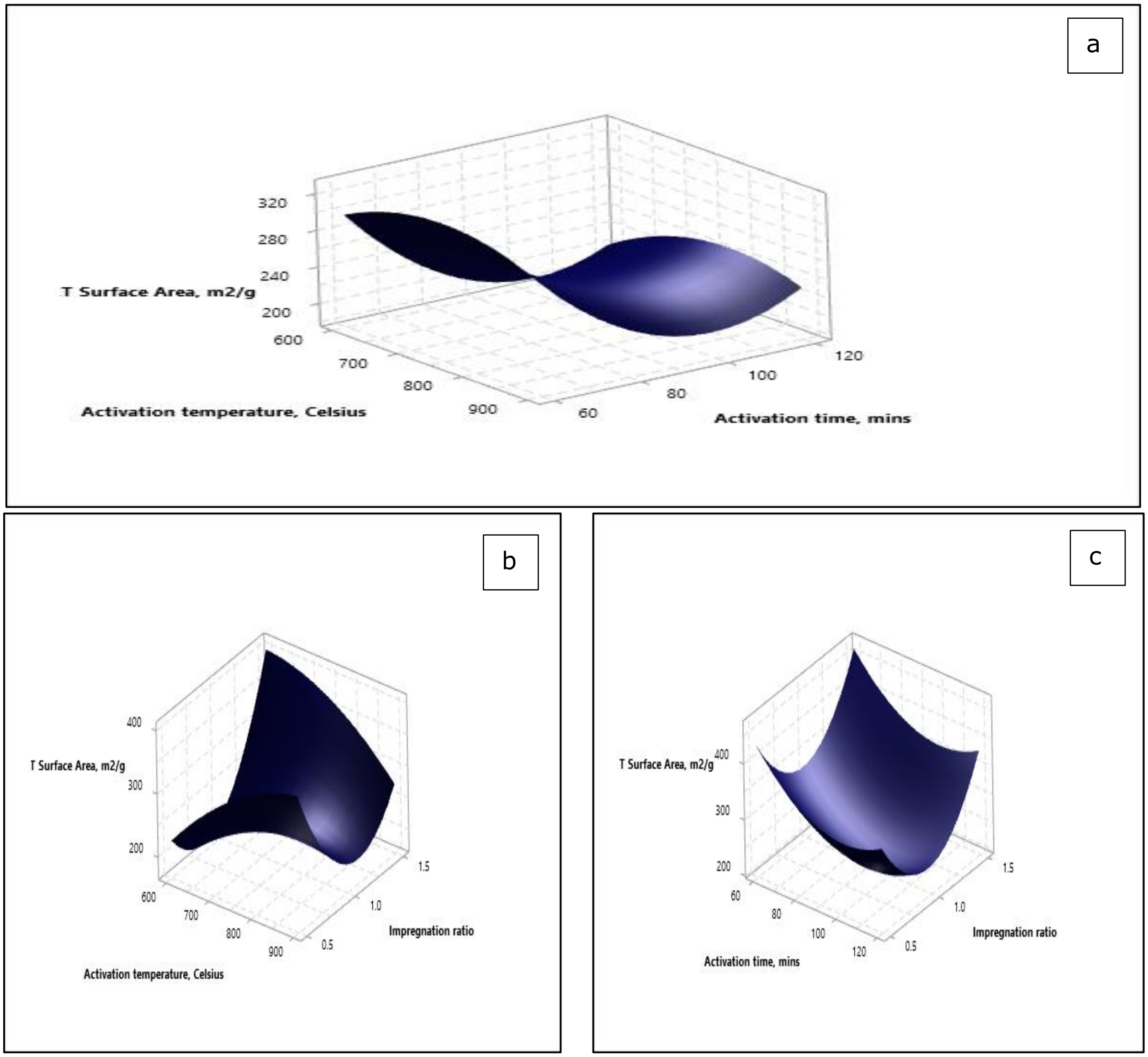


Figure 6: Response surface plots of BET surface area for (a) Impregnation ratio = 1.0 (b) Activation time = 90 min, and (c) Activation temperature = 750 ^{0}C.

From Fig 6a, it is evident that the BET surface area of $ZnCl_2$ activated carbon showed a slight decreasing trend with increase in activation temperature. Likewise, the estimated BET surface area of $ZnCl_2$ activated carbon showed a decreasing trend with increase in activation time. In Figure 6b, the BET surface area of $ZnCl_2$ activated carbon showed an increasing trend with increase in activation temperature, whilst the BET surface area of $ZnCl_2$ activated carbon initially showed decreasing trend with increasing impregnation ratio towards a minimum, and subsequently plateaued to a constant level. It is also apparent from Figure 6c, that the BET surface area of $ZnCl_2$ activated carbon showed a decreasing trend with increasing activation time. Whereas, estimated BET surface area showed an increasing trend with increase in impregnation ratio towards a maximum level. Notably, the adsorption capacity of the prepared activated carbons is contingent on impregnation ratio, which produces additional surface active (binding) sites (Marsh and Reinoso, 2006). Consequently, it is important that impregnation ratio is considered in future optimization studies.

From the statistical optimization, the optimum BET surface area attained for $ZnCl_2$ activated CNS was

$504 m^2.g^{-1}$ for activation temperature of 873K, activation time of 60min, and impregnation ratio of 1.50 respectively. Lastly, additional experimental run(s) using the optimal conditions was carried out, and the empirical BET surface area for $ZnCl_2$ activated CNS was found to be $492 m^2.g^{-1}$ (See Table 5). These results validated the polynomial model explicated with RSM.

Table 5: Predicted and Observed Value(s) of BET surface area for $ZnCl_2$ activated CNS utilizing optimal activation conditions.

Optimal Conditions				**BET Surface Area (m^2/g)**	
No of Replicates	Activation Time, min	Activation Temperature, K	Impregnation ratio	Experimental	Predicted
1	60	873	1.50	492	504

3.1.2. Modeling of the CNS base ($ZnCl_2$) activation process using GMDH

The GMDH-type neural model was implemented in MATLAB R2018a version 9.4. The architecture of the built GMDH network consists of one input layer, two hidden layers, in combination with one output layer as shown in Figure 7. A total of two neurons were included in the first layer, and one neuron in the second layer of the model. After completion of network training using the training dataset with three inputs (activation time, activation temperature and impregnation ratio), the simulated (output) result was checked with the test data to determine its computation accuracy.

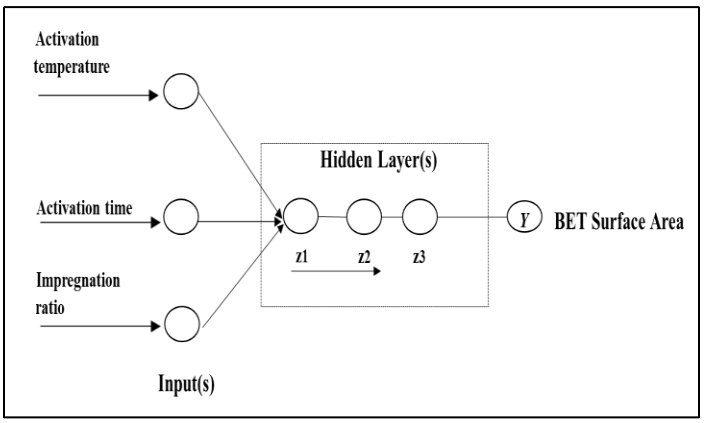
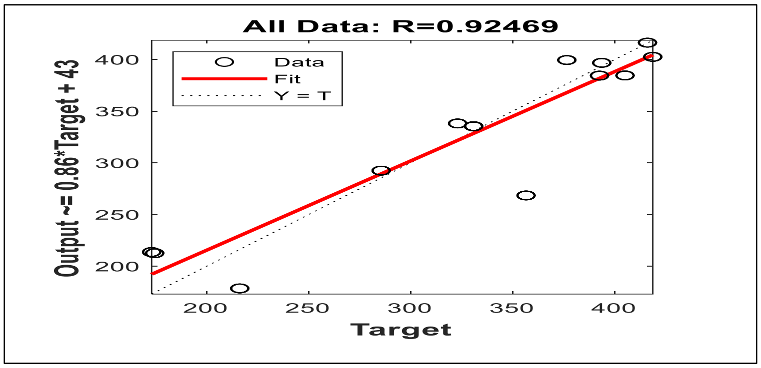


Figure 7: Architecture of proposed GMDH-type BET neural network.

The regression plots of training, overall, and evaluation are displayed in Figure 8 for the GMDH-type neural model. The optimal model for ZCNS activation corresponded to correlation coefficient (R) of 0.864 for training, and 0.932 for testing yielding an average R value of 0.925, indicating strong connection between the input values and GMDH-type neural network predictions for achievable surface area (Nkurlu et al., 2020). This finding is also corroborated by the relatively low value obtained for the performance function (RMSE = 32.0).



a

b

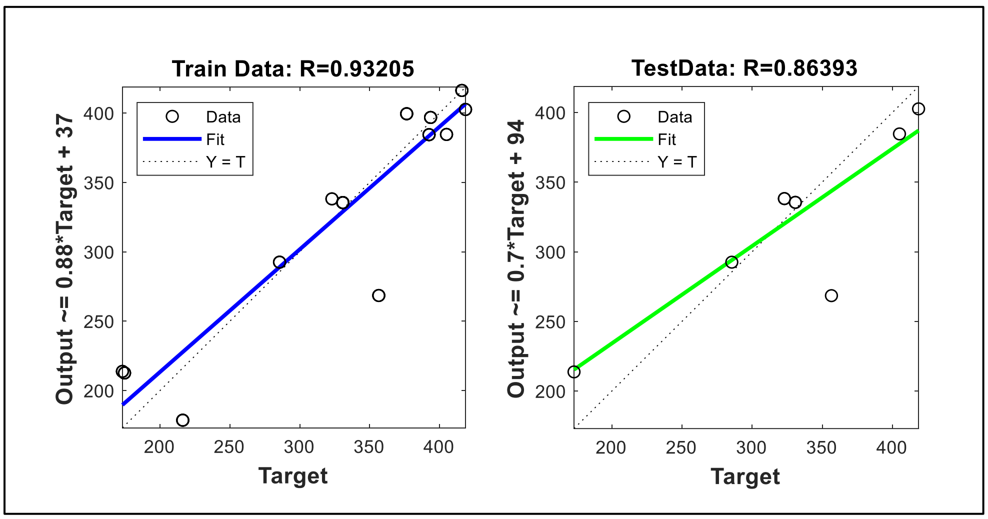


Figure 8: GMDH Regression plots for (a) overall (b) training, and testing data.

The GMDH-type neural model was further studied by comparing the GMDH prediction (outputs) and experimental outcomes (targets) for achievable BET surface area(s), as depicted in Figure 9.

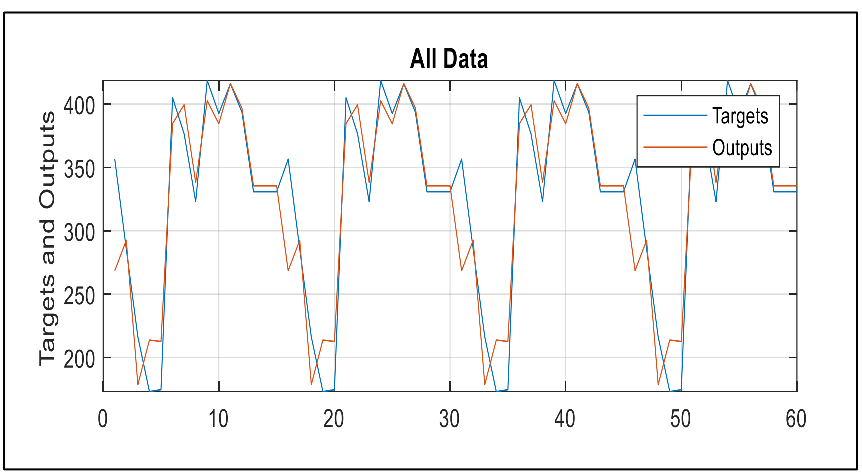


Figure 9: Cross-plot of experimental (target) and simulated (output) results for the ZCNS activation process.

From Figure 9, it is lucid that the target (experimental) and simulated (output) values match well, as confirmed by the moderately low value of the error function (RMSE=32.0) hitherto obtained. Hence, the GMDH-type neural model was able to successfully map the experimental domain, allowing for reasonably accurate estimate(s) of intermittent variations in the achievable BET surface area(s) of ZCNS.

3.1.3. Comparative Analysis of GMDH-type and RSM BET models

The efficiency of the created RSM (BBD) and GMDH-type BET neural models in predicting the achievable BET surface areas were evaluated using standard statistical indices-squared loss function and root mean squared error (RMSE) as shown in Table 6 (Ayoub et al., 2019; Chebii et al., 2022). The performance index-squared loss function is a good indicator of model predictive ability, because it is very sensitive to outliers and non-negative (Gokcesu, and Gokcesu., 2023).

Table 6: Comparison of GMDH-type and RSM BET models for ZCNS activation process.

Run No	RSM Square Residual	GMDH Square Residual
1	3433.95	7747.10
2	200.58	48.78
3	200.58	1417.92
4	3433.95	1655.27
5	2405.63	1448.49
6	562.41	417.059
7	562.41	528.18
8	2405.63	229.34
9	91.25	257.031
10	2163.45	65.191
11	2163.45	0.0767
12	162.23	10.31
13	1.29E-26	21.634
14	1.29E-26	21.634
15	1.29E-26	21.634
16	3433.95	7747.10
17	200.58	48.790
18	200.58	1417.92
19	3433.95	1655.27
20	2405.63	1448.49
21	562.41	417.06
22	562.41	528.18
23	2405.63	229.344
24	91.25	257.031
25	2163.45	65.191
26	2163.45	0.0767
27	162.23	10.307
28	1.29E-26	21.634
29	1.29E-26	21.634
30	1.29E-26	21.634
31	3433.95	7747.10
32	200.58	48.787
33	200.58	1417.92
34	3433.95	1655.27
35	2405.63	1448.49
36	562.41	417.06
37	562.41	528.178
38	2405.63	229.344
39	91.254	257.031
40	2163.45	65.191
41	2163.45	0.0767
42	162.23	10.307

Run No	RSM	GMDH
	Square Residual	**Square Residual**
43	1.29E-26	21.634
44	1.29E-26	21.634
45	1.29E-26	21.634
46	3433.95	7747.10
47	200.58	48.79
48	200.58	1417.92
49	3433.95	1655.27
50	2405.63	1448.49
51	562.41	417.06
52	562.41	528.178
53	2405.63	229.344
54	91.25	257.031
	RMSE= 35.10	**RMSE** = 32.0

The RSM and GMDH-type BET models generated squared-error residual values ranging from 1.29×10^{-26} to 3434, and 7.67×10^{-2} to 7747 respectively for achievable BET surface area(s). The lower value(s) of the loss function ascertained for RSM BET model, suggests that the RSM BET predictive model performs better than GMDH. Consequently, the comparatively higher squared-error residual values obtained for GMDH, necessitated further error analysis. From Table 7, Root mean squared error (RMSE) values of 32.0 and 35.10 respectively were also recorded for the GMDH-type and RSM BET models. According to Chebii et al., (2022), RMSE is a standard statistical index for valuating the performance of nonlinear regression models. The lower RMSE value ascertained for GMDH, confirmed that GMDH model is the more appropriate model for predicting the achievable surface area of ZCNS. Overall, RSM and GMDH techniques have been evidenced to be effective methods for delineating the $ZnCl_2$-activated CNS production process.

4. SENSITIVITY ANALYSIS

Sensitivity analysis was carried out to investigate the influence of activation parameters (impregnation ratio, activation time, & activation temperature) on the predicted BET surface experimental data conducted utilising Eq. (6) (Nkurlu et al., 2020):

$$S = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{percent\ change\ in\ output}{percent\ change\ in\ input}\right)_i \times 100 \qquad \text{Eq. (6)}$$

Where N is the total number of experimental runs, i is the respective data point, and S is sensitivity value. The lower the S value, the less the activation (process) variable affected the simulated BET surface area value. Contrariwise, the higher the S value, the greater the input activation variable influences the simulated BET surface area of ZCNS (Nkurlu et al., 2020).

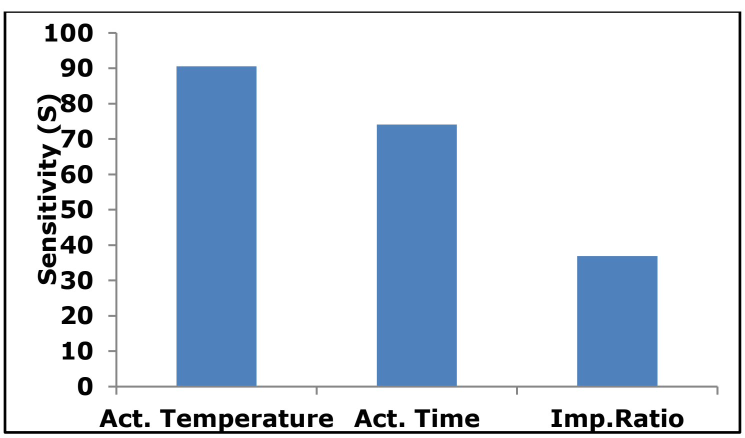


Figure 10: Effect of GMDH input variables on $ZnCl_2$-CNS activation process.

From Figure 10, both activation temperature and time had more significant impact on GMDH BET predictive model than impregnation ratio with S values of 90.6% and 74.1% respectively. The low S value of 36.9% obtained for impregnation ratio signifies that this input had little effect on the BET surface area prediction(s).

5. CONCLUSIONS

In this study, Box Behnken design (BBD) of Response Surface Methodology (RSM)) and novel Group Method of Data Handling (GMDH) approaches were employed to optimize the conditions for production of ZCNS with high BET surface areas. The RSM and GMDH-type neural models obtained via regression analysis and multi-layered iterative algorithm (MIA) respectively, predicted the response of interest ((BET surface area) fairly accurately. The highest BET surface area for ZCNS obtained from RSM optimisation was estimated to be 504 $m^2.g^{-1}$ attained at optimal process conditions of activation time (60 min), impregnation ratio (1.50), and activation temperature (873K). The optimal GMDH-type BET model was identified to consist of 3 input variables, and 2 hidden layer(s) having two and one neuron(s) respectively. The root mean square error (RMSE) and correlation coefficient (R) were chosen as statistical indices for evaluating the predictive models performance. With the least RMSE (32.0), and highest correlation coefficient (0.925), The GMDH-type BET model proved to be better compared to the RSM model. These findings confirmed that GMDH-type BET model has the best analytical performance. Lastly, the sensitivity analysis outcome unveiled that activation temperature and activation time had a predominant influence on the performance of the GMDH-type BET neural model.

6. ACKNOWLEGEMENTS

The authors will like to acknowledge Engr Christopher Obi of Polymer Engineering Department, Nnamdi Azikiwe University, Awka for proof-reading the final manuscript and suggesting a few corrections.

7. REFERENCES

Adeigbe, O. O., Olasupo, F. O., Adewale, B. D., & Muyiwa, A. A. (2015). A review on cashew research and production in Nigeria in the last four decades. *Scientific Research and Essays*, *10*(5), 196–209. https://doi.org/10.5897/SRE2014.5953

AJEMBA, R. O. (2014). Kinetics and equilibrium modeling of lead (II) and chromium (III) ions' adsorption onto clay from Kono-bowe, Nigeria. *Turkish Journal of Engineering and Environmental Sciences*, *38*(3), 455–479.

Akinbiyi, A. O. (2001). *Removal of lead from aqueous solutions by adsorption using peat moss.* Faculty of Graduate Studies and Research, University of Regina.

Anderson, M. J., & Whitcomb, P. J. (2016). *RSM Simplified: Optimizing Processes Using Response Surface Methods for Design of Experiments, Second Edition* (2nd ed.). Productivity Press.

Ayoub, Mohammed. A., Almansour, Abdullah. O., & Hassan, Anas. M. (2020). A Novel Formula for Estimating Oil Compressibility Below Bubble Point Pressure Using Group Method of Data Handling: A Comparative Approach. *Day 1 Tue, October 29, 2019*, D012S002R001. https://doi.org/10.2118/196446-MS

Buasri, A., Sirikoom, P., Pattane, S., Buachum, O., & Loryuenyong, V. (2023). Process Optimization of Biodiesel from Used Cooking Oil in a Microwave Reactor: A Case of Machine Learning and Box–Behnken Design. *ChemEngineering*, *7*(4), 65. https://doi.org/10.3390/chemengineering7040065

Chebii, S. J., Mukolwe, M. M., & Ong'or, B. I. (2022). River flow modelling for flood prediction using artificial neural network in ungauged Perkerra catchment, Baringo County, Kenya. *Water Practice and Technology*, *17*(4), 914–929. https://doi.org/10.2166/wpt.2022.034

Dawei, H. (2012). *Concise Environmental Engineering*. Ventus Publishing ApS.

Demirbas, E., Kobya, M., & Konukman, A. E. S. (2008). Error analysis of equilibrium studies for the almond shell activated carbon adsorption of Cr(VI) from aqueous solutions. *Journal of Hazardous Materials*, *154*(1–3), 787–794. https://doi.org/10.1016/j.jhazmat.2007.10.094

Elibol, M. (2002). Response surface methodological approach for inclusion of perfluorocarbon in actinorhodin fermentation medium. *Process Biochemistry*, *38*(5), 667–673. https://doi.org/10.1016/S0032-9592(02)00171-1

Essa, M. H., Zahrani, M. A. A., & Nesaratnam, S. (2013). Optimisation of activated carbon production from date pits. *International Journal of Environmental Engineering*, *5*(3), 325. https://doi.org/10.1504/IJEE.2013.054708

Farlow, S. J. (1981). The GMDH Algorithm of Ivakhnenko. *The American Statistician*, *35*(4), 210–215. https://doi.org/10.1080/00031305.1981.10479358

Gokcesu, K., & Gokcesu, H. (2023). *A Note On Nonlinear Regression Under L2 Loss*. https://doi.org/10.48550/ARXIV.2303.17745

Kehinde Ademola, Adesanya, Mutiat Oluwaseyi Agboola-Adedoja, Adejoke Adebusola Adelusi, Qudus Adebayo Ogunwolu, Chinweike Abednego Ugwu, Modinat Adewunmi Alli, & Ayodele Oladipo Akinpelu. (2021).

Opportunities in Nigerian cashew nut value chain. *World Journal of Advanced Research and Reviews*, *9*(1), 168–174. https://doi.org/10.30574/wjarr.2021.9.1.0008

Kulkarni, S. J., & Kaware, J. (2014). Removal of phenol from effluent in fixed bed: A review. *International Journal of Engineering Research and General Science*, *2*(5), 35–38.

Li, R. Y. M., Fong, S., & Chong, K. W. S. (2017). Forecasting the REITs and stock indices: Group Method of Data Handling Neural Network approach. *Pacific Rim Property Research Journal*, *23*(2), 123–160. https://doi.org/10.1080/14445921.2016.1225149

Madala, H. R., & Ivakhnenko, A. G. (1994). *Inductive learning algorithms for complex systems modeling*. CRC Press.

Malik, D. S., Jain, C. K., & Yadav, A. K. (2017). Removal of heavy metals from emerging cellulosic low-cost adsorbents: A review. *Applied Water Science*, *7*(5), 2113–2136. https://doi.org/10.1007/s13201-016-0401-8

Mathew Nkurlu, B., Shen, C., Asante-Okyere, S., Mulashani, A. K., Chungu, J., & Wang, L. (2020). Prediction of Permeability Using Group Method of Data Handling (GMDH) Neural Network from Well Log Data. *Energies*, *13*(3), 551. https://doi.org/10.3390/en13030551

Montgomery, D. C. (2017). *Design and analysis of experiments* (Ninth edition). Wiley.

Nwabanne, J. T., & Okoye, A. C. (2013). Treatment of synthetic and battery industry wastewater by electrocoagulation. *Der Chemica Sinica*, *4*(6), 32–39.

Okewale, A., Igbokwe, P., & Adesina, O. (2015). Optimization of the adsorptive dehydration of ethanol–water system. *Optimization*, *39*.

Okiy, K. V. (2006). *Scale-up of an electrochemical reactor for the removal of copper ions from wastewater* [B. Sc. Thesis]. University of Lagos.

Onu, C. E., Nwabanne, J. T., Ohale, P. E., & Asadu, C. O. (2021). Comparative analysis of RSM, ANN and ANFIS and the mechanistic modeling in eriochrome black-T dye adsorption using modified clay. *South African Journal of Chemical Engineering*, *36*, 24–42. https://doi.org/10.1016/j.sajce.2020.12.003

Özer, A., Gürbüz, G., Çalimli, A., & Körbahti, B. K. (2009). Biosorption of copper(II) ions on Enteromorpha prolifera: Application of response surface methodology (RSM). *Chemical Engineering Journal*, *146*(3), 377–387. https://doi.org/10.1016/j.cej.2008.06.041

Palani, S., Liong, S.-Y., Tkalich, P., & Palanichamy, J. (2009). Development of a neural network model for dissolved oxygen in seawater. *Indian Journal of Geo-Marine Sciences*, *38*(2), 151–159.

Peavy, H. S., Rowe, D. R., & Tchobanoglous, G. (1985). *Environmental engineering*. McGraw-Hill.

S, M. S., M, E. A. A., & Chidambaram, R. (2015). Isotherm Modelling, Kinetic Study and Optimization of Batch Parameters Using Response Surface Methodology for Effective Removal of Cr(VI) Using Fungal Biomass. *PLOS ONE*, *10*(3), e0116884. https://doi.org/10.1371/journal.pone.0116884

Senthil Kumar, P., Ramalingam, S., Abhinaya, R. V., Kirupha, S. D., Murugesan, A., & Sivanesan, S. (2012). Adsorption of Metal Ions onto the Chemically Modified Agricultural Waste. *CLEAN – Soil, Air, Water*, *40*(2), 188–197. https://doi.org/10.1002/clen.201100118

Singh, R., Gautam, N., Mishra, A., & Gupta, R. (2011). Heavy metals and living systems: An overview. *Indian Journal of Pharmacology*, *43*(3), 246. https://doi.org/10.4103/0253-7613.81505

Srivastava, S., Agrawal, S. B., & Mondal, M. K. (2015). A review on progress of heavy metal removal using adsorbents of microbial and plant origin. *Environmental Science and Pollution Research*, *22*(20), 15386–15415. https://doi.org/10.1007/s11356-015-5278-9

Stavropoulos, G. G., & Zabaniotou, A. A. (2009). Minimizing activated carbons production cost. *Fuel Processing Technology*, *90*(7–8), 952–957. https://doi.org/10.1016/j.fuproc.2009.04.002

Subramaniam, R., & Kumar Ponnusamy, S. (2015). Novel adsorbent from agricultural waste (cashew NUT shell) for methylene blue dye removal: Optimization by response surface methodology. *Water Resources and Industry*, *11*, 64–70. https://doi.org/10.1016/j.wri.2015.07.002

Van Dyk, L. D. (2000). *The production of granular activated carbon from agricultural waste products* [MSc Thesis, Stellenbosch University]. https://scholar.sun.ac.za/items/388fda93-974e-4e4d-9562-0e78cc0f8c31

Voss, M. S. (2002). *The Group Method for Cartesian Programming: A New Methodology for Complex Adaptive Functional Networks* [PhD Thesis, Marquette University]. https://epublications.marquette.edu/cgi/viewcontent.cgi?article=2797&context=dissertations_mu

Wang, Z., Wu, Q., Zhang, J., Zhang, H., Feng, J., Dong, S., & Sun, J. (2019). In situ polymerization of magnetic graphene oxide-diaminopyridine composite for the effective adsorption of Pb(II) and application in battery industry wastewater treatment. *Environmental Science and Pollution Research*, *26*(32), 33427–33439. https://doi.org/10.1007/s11356-019-06511-1

Xu, Z., Cai, J., & Pan, B. (2013). Mathematically modeling fixed-bed adsorption in aqueous systems. *Journal of Zhejiang University SCIENCE A*, *14*(3), 155–176. https://doi.org/10.1631/jzus.A1300029

Yi, Q. (2009). *Point Sources of Pollution: Local Effects and their Control—Volume II*. EOLSS Publications. https://books.google.com.tr/books?id=dq2kCwAAQBAJ