Supporting information

Eliminating luck and chance in the reactivation process: a systematic and quantitive study of the thermal reactivation of activated carbons

Karthik Rathinam[1*], Volker Mauer[2], Christian Bläker[2], Christoph Pasel[2], Lucas Landwehrkamp[1], Dieter Bathen[2,3], Stefan Panglisch[1*]

[1]Chair for Mechanical Process Engineering and Water Technology, University of Duisburg-Essen, Duisburg, 47057 Germany

[2]Chair for Thermal Process Engineering, University of Duisburg-Essen, Duisburg, 47057 Germany

[3]Institute for energy and environmental technology (IUTA e.V.), Duisburg, 47229 Germany

*Corresponding author:

stefan.panglisch@uni-due.de (Stefan Panglisch)

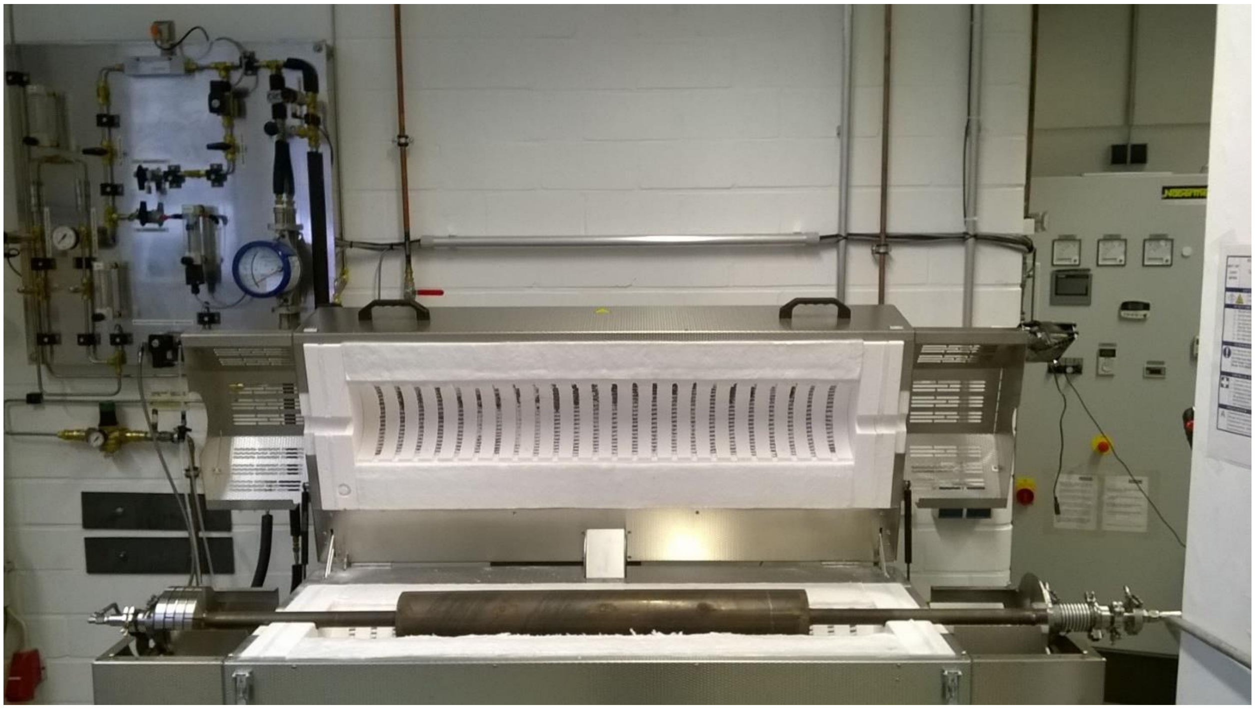

Figure S1: Photo image of a bench-scale rotary kiln furnace set-up used for the thermal reactivation process

Table S1. Calculated isotherm parameters for DA adsorption onto RACs and fresh AC.

	Freundlich			Langmuir		
	K_F (mg/g)/(mg/L)^n	n	R^2	K_L 1/(mg/L)	q_{max} (g/kg)	R^2
Fresh AC	48.7	0.19	0.977	5.6	60.9	0.999
1	19.9	0.1	0.592	5.5	23.4	0.989
2	34.3	0.16	0.792	4.6	42.7	0.986
3	30.3	0.18	0.868	3.4	40.8	0.990
4	43.6	0.11	0.664	35.9	47.1	0.982
5	25.9	0.16	0.902	3.5	34.1	0.996
6	42.5	0.21	0.945	2.8	60	0.983
7	32.9	0.06	0.500	2.7	35.4	0.990
8	52.4	0.16	0.971	8.7	62.4	0.998
9	19.2	0.24	0.973	1.7	30.3	0.995
10	32.7	0.15	0.864	4.3	41.3	0.989
11	29.5	0.27	0.907	1.6	47.7	0.972
12	44.9	0.12	0.947	10.5	51.5	0.998
13	22.6	0.25	0.98	1.7	36.2	0.996
14	41.1	0.22	0.965	2.8	58.6	0.990
15	34.2	0.22	0.927	2.1	50.6	0.979
16	56.9	0.17	0.987	7.2	69.7	0.994
17	23.2	0.13	0.740	4.5	29.2	0.986
18	52.2	0.18	0.989	6.1	65	0.996
19	21.5	0.21	0.779	1.8	32	0.963
20	41.1	0.22	0.965	2.8	58.6	0.990
21	38.5	0.19	0.907	3.4	52.1	0.986
22	37.9	0.19	0.866	2.9	53.1	0.979
23	34.2	0.1	0.539	11.6	38.8	0.977
24	37.9	0.07	0.371	96.8	39.7	0.971
25	35.8	0.17	0.506	6.1	41	0.980
26	35.7	0.14	0.956	5.4	43.3	0.999

S1 Analysis of Variance using the F-Test of the overall significance[1]

In general, an F-test in regression compares the fits of different models. Unlike t-tests that can assess only one regression coefficient at a time, the F-test can assess multiple coefficients simultaneously. The F-test of the overall significance is a specific form of the F-test. It compares a model with no predictors to the determined model with predictors. A regression model that contains no predictors is also known as an intercept-only model.

The hypotheses for the F-test of the overall significance are as follows:

- Null hypothesis: The fit of the intercept-only model and the determined model are equal (all coefficient besides the constant coefficient are zero).
- Alternative hypothesis: The fit of the intercept-only model is significantly reduced compared to the determined model.

If the P value for the F-test of overall significance is less than the choosen significance level (here 5 %), the null-hypothesis can be rejected and thus be concluded that the determined model provides a better fit than the intercept-only model.

Degree of freedom (DF)

The total degrees of freedom (DF) are the amount of information in the data. The analysis uses that information to estimate the values of unknown population parameters. The total DF is determined by the number of observations in the sample. The DF for a term show how much information that term uses. Increasing the sample size provides more information about the population, which increases the total DF. Increasing the number of terms in the model uses more information, which decreases the DF available to estimate the variability of the parameter estimates.

If the experimental design has replications (i.e., multiple runs with identical levels for all model terms), there are DF for the pure error. Each set of replications (r) contributes r - 1 degrees of freedom to the pure error. The total DF for the (residual) error are the total number of runs minus the number of estimated parameters.

Adjusted sums of squares

Adjusted sums of squares (Adj SS) are measures of variation for different components of the model. The order of the predictors in the model does not affect the calculation of the adjusted sums of squares. In the Analysis of Variance table, Minitab separates the sums of squares into different components that describe the variation due to different sources.

[1] https://www.minitab.com/en-us/support/ accessed on 26 November 2023

Minitab uses the adjusted sums of squares to calculate the p-value for a term. Minitab also uses the sums of squares to calculate the R^2 statistic. Usually, the p-values and the R^2 statistic are interpreted instead of the sums of squares.

Adj SS Term (Model)

The adjusted sum of squares for a term is the increase in the regression sum of squares compared to a model with only the other terms. It quantifies the amount of variation in the response data that is explained by each term in the model. The adjusted sum of squares for the model is the sum of the term sum of squares.

Adj SS Error

The error sum of squares is the sum of the squared residuals. It quantifies the variation in the data that the predictors do not explain.

Adj SS Total

The total sum of squares is the sum of the term sum of squares and the error sum of squares. It quantifies the total variation in the data.

Adjusted mean squares

Adjusted mean squares (Adj MS) measure how much variation a term or a model explains, assuming that all other terms are in the model, regardless of the order they were entered. Unlike the adjusted sums of squares, the adjusted mean squares consider the degrees of freedom:

$$ADJ\ MS = \frac{ADJ\ SS}{DF}$$

The adjusted mean square of the error (also called MSE or s^2) is the variance around the fitted values.

Minitab uses the adjusted sums of squares to calculate the p-value for a term. Minitab also uses the sums of squares to calculate the R^2 statistic. Usually, the p-values and the R^2 statistic are interpreted instead of the sums of squares.

F-value

Minitab uses the F-value to calculate the p-value, which is used to make decision about the statistical significance of the terms and model. The p-value is a probability that measures the evidence against the null hypothesis. Lower probabilities provide stronger evidence against the null hypothesis:

A sufficiently large F-value indicates that the term or model is significant. If the F-value is used to determine whether to reject the null hypothesis, the F-value is compared to the critical value. In the present case, however, the P-Value is used.

An F-value appears for each term in the Analysis of Variance table:

F-value for the model or the terms is the test statistic used to determine whether the term is associated with the response:

$$F - Value = \frac{ADJ\ MS\ (Term)}{ADJ\ MS\ (Error)}$$

F-value for the lack-of-fit test is used to determine whether the model is missing higher-order terms that include the predictors in the current model:

$$F - Value = \frac{ADJ\ MS\ (Lack - of - Fit)}{ADJ\ MS\ (Error)}$$

Lack-of-fit

If the experimental design has replications and the model is not saturated, some of the DF relate to the missing fit. The DFs for the missing fit are determined by subtracting the DFs for the pure error from the DFs for the residual errors. The sum of squares for the missing fit is calculated by subtracting the sums of squares for the pure error from the sum of squares of the residual errors. The sum of squares for the missing fit represents the total effect of all estimable interaction terms that were omitted from the model.

P-value

The p-value is a probability that measures the evidence against the null hypothesis. Lower probabilities provide stronger evidence against the null hypothesis. The p-value is calculated from the F-distribution.

P-value – Regression

To determine whether the model explains variation in the response, the p-value for the model is compared to the significance level to assess the null hypothesis. The null hypothesis for the overall regression is that the model does not explain any of the variation in the response. Usually, a significance level (denoted as α or alpha) of 0.05 works well. A significance level of 0.05 indicates a 5% risk of concluding that the model explains variation in the response when the model does not.

P-value ≤ α: The model explains variation in the response

If the p-value is less than or equal to the significance level, it is concluded that the model explains variation in the response.

P-value > α: There is not enough evidence to conclude that the model explains variation in the response

If the p-value is greater than the significance level, one cannot conclude that the model explains variation in the response. A new model might have to be fitted.

P-Value – Term

To determine whether the association between the response and each term in the model is statistically significant, the p-value for the term is compared to the significance level to assess the null hypothesis. The null hypothesis is that there is no association between the term and the response. Usually, a significance level (denoted as α or alpha) of 0.05 works well. A significance level of 0.05 indicates a 5% risk of concluding that an association exists when there is no actual association.

P-value ≤ α: The association is statistically significant

If the p-value is less than or equal to the significance level, you can conclude that there is a statistically significant association between the response variable and the term.

P-value > α: The association is not statistically significant

If the p-value is greater than the significance level, you cannot conclude that there is a statistically significant association between the response variable and the term. You may want to refit the model without the term.

If there are multiple predictors without a statistically significant association with the response, one can reduce the model by removing terms one at a time.

If a model term is statistically significant, the interpretation depends on the type of term. The interpretations are as follows:

- If a continuous predictor is significant, one can conclude that the coefficient for the predictor does not equal zero.
- If a categorical predictor is significant, one can conclude that not all the level means are equal.
- If an interaction term is significant, one can conclude that the relationship between a predictor and the response depends on the other predictors in the term.

- If a polynomial term is significant, one can conclude that the data contain curvature.

P-value – Lack-of-fit

Minitab automatically performs the pure error lack-of-fit test when the data contain replicates, which are multiple observations with identical x-values. Replicates represent "pure error" because only random variation can cause differences between the observed response values.

To determine whether the model correctly specifies the relationship between the response and the predictors, the p-value for the lack-of-fit test is compared to the significance level to assess the null hypothesis. The null hypothesis for the lack-of-fit test is that the model correctly specifies the relationship between the response and the predictors. Usually, a significance level (denoted as alpha or α) of 0.05 works well. A significance level of 0.05 indicates a 5% risk of concluding that the model does not correctly specify the relationship between the response and the predictors when the model does specify the correct relationship.

P-value ≤ α: The lack-of-fit is statistically significant

If the p-value is less than or equal to the significance level, one can conclude that the model does not correctly specify the relationship. To improve the model, one may need to add terms or transform the data.

P-value > α: The lack-of-fit is not statistically significant

If the p-value is larger than the significance level, the test does not detect any lack-of-fit.

Standard deviation

S represents the standard deviation of the distance between the data values and the fitted values. S is measured in the units of the response.

$$S = \sqrt{ADJ\ MS\ (Error)}$$

S is used to assess how well the model describes the response. S is measured in the units of the response variable and represents how far the data values fall from the fitted values. The lower the value of S, the better the model describes the response. However, a low S value by itself does not indicate that the model meets the model assumptions.

Percentage of variation in the response

R^2 is the percentage of variation in the response that is explained by the model. It is calculated as 1 minus the ratio of the error sum of squares (which is the variation that is not explained by model) to the total sum of squares (which is the total variation in the model).

$$R^2 = 1 - \frac{ADJ\ SS\ (Error)}{ADJ\ SS\ (Total)}$$

R^2 is used to determine how well the model fits the data. The higher the R^2 value, the better the model fits the data. R^2 is always between 0% and 100%.

Table S2: Analysis of Variance for R1

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Model	4	46150.0	11537.5	26.23	0.000
Linear	3	44444.7	14814.9	33.68	0.000
X_1	1	27990.8	27990.8	63.64	0.000
X_2	1	13194.5	13194.5	30	0.000
X_3	1	3259.4	3259.4	7.41	0.013
Square	1	2832.4	2832.4	6.44	0.019
X_3*X_3	1	2832.4	2832.4	6.44	0.019
Error	21	9236.6	439.8		
Lack-of-Fit	18	8877.6	493.2	4.12	0.135
Pure Error	3	359	119.7		
Total	25	55386.6	2215.5		
S = 20.97, R^2 = 83.3%					

Table S3: Analysis of Variance for R2

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Model	5	34.0597	8.5149	1217.79	0.000
Linear	3	32.3809	16.1905	2315.53	0.000
X_1	1	10.7369	10.7369	1535.57	0.000
X_2	1	21.6440	21.6440	3095.49	0.000
Square	1	0.0532	0.0532	7.60	0.012
X_2*X_2	1	0.0532	0.0532	7.60	0.012
2-Way Interaction	1	1.6256	1.6256	232.49	0.000
X_1*X_2	1	1.6256	1.6256	232.49	0.000
Error	21	0.1468	0.0070		
Lack-of-Fit	18	0.1268	0.0070	1.06	0.561
Pure Error	3	0.0200	0.0067		
Total	25	34.2065			
S = 0.078, R^2 = 99.64%					

Table S4: Analysis of Variance for R3

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Model	4	85.79	21.4476	86.89	0.000
Linear	3	81.285	27.0951	109.76	0.000
X_1	1	38.183	38.1829	154.68	0.000
X_2	1	40.803	40.8031	165.3	0.000
X_3	1	2.299	2.2993	9.31	0.006
2-Way Interaction	1	4.505	4.505	18.25	0.000
X_1*X_2	1	4.505	4.505	18.25	0.000
Error	21	5.184	0.2468		
Lack-of-Fit	18	2.999	0.1666	0.23	0.982
Pure Error	3	2.185	0.7282		
Total	25	90.974	3.6390		
S = 0.497, R^2 = 94.30%					

Table S5: Analysis of Variance for R4

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Model	5	3243.74	648.75	34.88	0.000
Linear	4	3061.28	765.32	41.15	0.000
X_1	1	1749.66	1749.66	94.08	0.000
X_2	1	697.01	697.01	37.48	0.000
X_3	1	467.74	467.74	25.15	0.000
X_4	1	149.69	149.69	8.05	0.010
2-Way Interaction	1	182.46	182.46	9.81	0.005
X_1*X_3	1	182.46	182.46	9.81	0.005
Error	20	371.96	18.6		
Lack-of-Fit	17	360.23	21.19	5.42	0.095
Pure Error	3	11.73	3.91		
Total	25	3615.7			
S = 4.312, R^2 = 89.71%					

S2 Standardized Effects[2].

In this study, all standardized effects correspond to the absolute value of the t-statistic for the respective coefficient of the predictors:

$$t - value = \frac{\sqrt{ADJ\ SS\ (Term)}}{S}$$

[2] https://www.minitab.com/en-us/support/ accessed on 26 November 2023