Application of CFD, statistical analysis and DoE for optimising H_2 generation through ammonia borane catalytic hydrolysis: A novel approach for multiobjective optimisation

The objective of this research paper is to study the hydrolysis of ammonia borane with a non-noble metal catalyst to produce H_2. Design of experiments (DoE) methodology is implemented to study the effects of the reaction parameters on the critical response variables. Afterwards, multi-objective optimization is carried out to calculate the factor values which lead the response variables to optimum results.

The factors (independent variables) examined are: X_1 = catalyst to substrate ratio, X_2 = reaction temperature (°C) and X_3 = ammonia borane concentration (mM). All the factors are continuous. The responses (dependent variables) examined are: Y_1 = reaction time (min), Y_2 = turnover frequency (sec^{-1}) and Y_3 = H_2 yield (%). The applied DoE method is Box Behnken design.

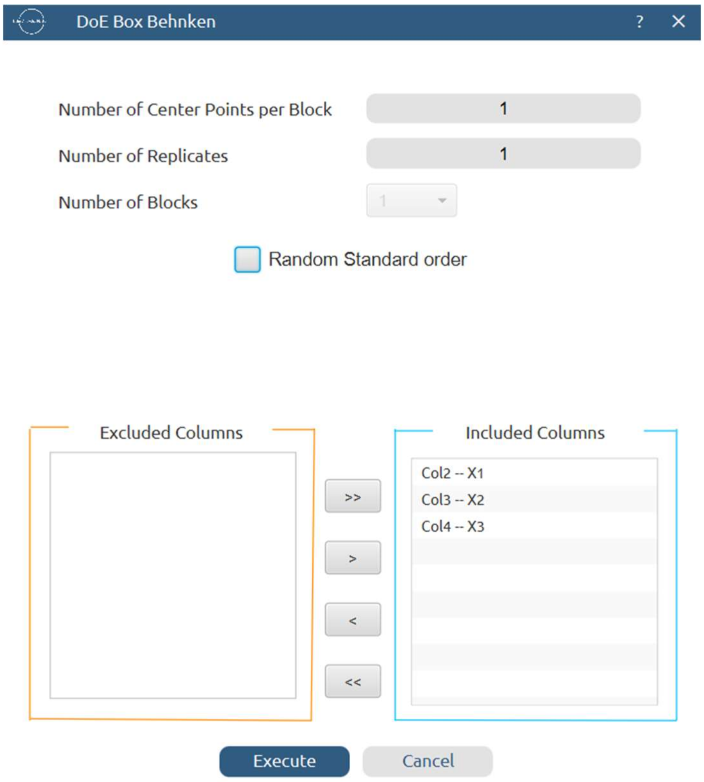
Isalos version used: 2.0.2

Scientific article: https://www.sciencedirect.com/science/article/abs/pii/S1385894725116173

Step 1: Box Behnken Design

In the first tab named "Action" define the factors in the column headers and fill each column with the low and high levels of the corresponding factors. This tab can be renamed "Box Behnken." Afterwards, apply the Box Behnken method: *DOE → Response Surface → Box Behnken*

	Col1	Col2 (I)	Col3 (I)	Col4 (I)
User Header	User Row ID	X1	X2	X3
1		1000	30	74
2		2000	50	111



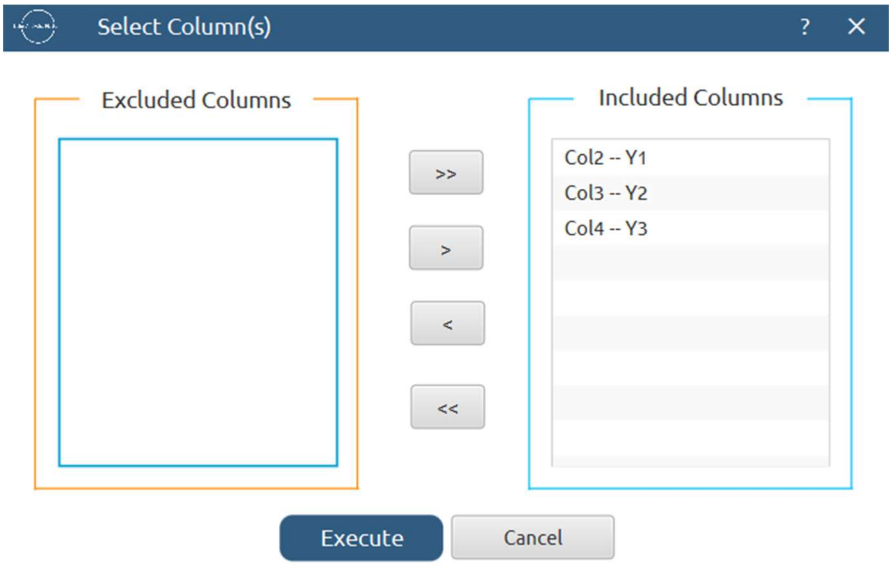
Results (right spreadsheet):

	Col2 (I)	Col3 (S)	Col4 (S)	Col5 (S)	Col6 (D)	Col7 (D)	Col8 (D)
User Header	Standard Order	Block Number	Replicate Number	Point Type	X1	X2	X3
1	1	Block: 1	Replicate: 1	Design Point	1000.0	30.0	92.5
2	2	Block: 1	Replicate: 1	Design Point	2000.0	30.0	92.5
3	3	Block: 1	Replicate: 1	Design Point	1000.0	50.0	92.5
4	4	Block: 1	Replicate: 1	Design Point	2000.0	50.0	92.5
5	5	Block: 1	Replicate: 1	Design Point	1000.0	40.0	74.0
6	6	Block: 1	Replicate: 1	Design Point	2000.0	40.0	74.0
7	7	Block: 1	Replicate: 1	Design Point	1000.0	40.0	111.0
8	8	Block: 1	Replicate: 1	Design Point	2000.0	40.0	111.0
9	9	Block: 1	Replicate: 1	Design Point	1500.0	30.0	74.0
10	10	Block: 1	Replicate: 1	Design Point	1500.0	50.0	74.0
11	11	Block: 1	Replicate: 1	Design Point	1500.0	30.0	111.0
12	12	Block: 1	Replicate: 1	Design Point	1500.0	50.0	111.0
13	13	Block: 1	----	Center Point	1500.0	40.0	92.5

Step 2: Definition of response variables

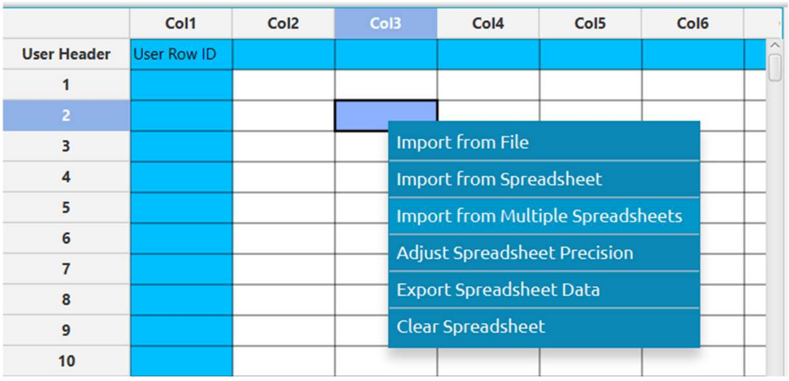
Create a new tab named “Responses” and define the responses in the column headers. Fill each column with the values of the corresponding responses that were observed and make sure the values follow the order of the experiments as given by the Box Behnken method. Then, select all columns to be transferred to the right spreadsheet: *<u>Data Transformation → Data Manipulation → Select Column(s)</u>*

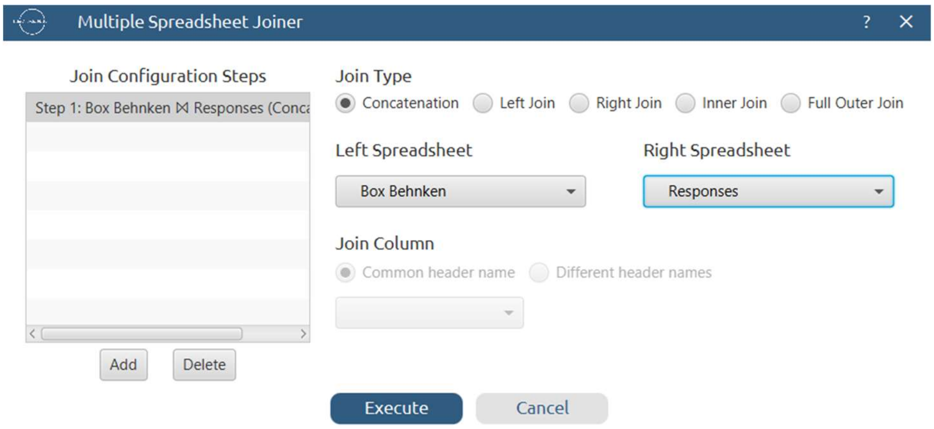
	Col1	Col2 (D)	Col3 (I)	Col4 (D)
User Header	User Row ID	Y1	Y2	Y3
1		95.35	3035	68
2		166.76	2936	55.2
3		15.96	3755	70
4		41.07	2185	67
5		42.1	2526	69
6		83.57	3480	57.7
7		31.64	3075	65.1
8		71.79	2280	54.7
9		206.6	4519	52.7
10		31.4	2754	70.2
11		141.85	3022	70.8
12		23.55	2974	67.7
13		68.8	19375	61.1

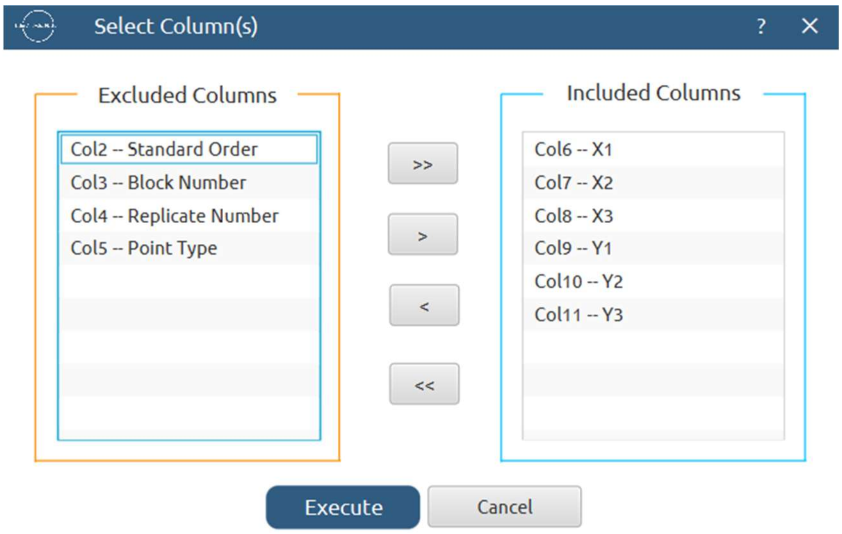


Step 3: Data isolation

Create a new tab named “Data” and import the results from the “Box Behnken” and “Responses” spreadsheets by right clicking on the left spreadsheet. Then, select only the factors and responses columns to be transferred to the right spreadsheet: _Data Transformation → Data Manipulation → Select Column(s)_





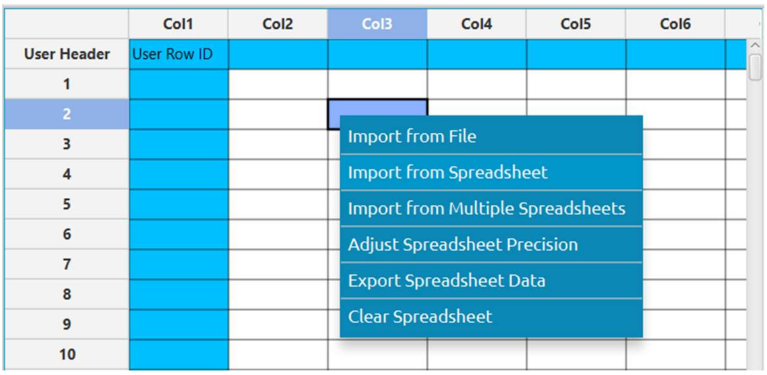


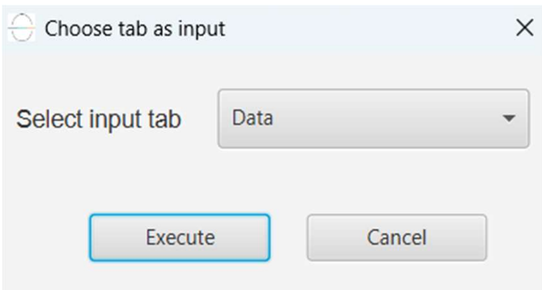
Results:

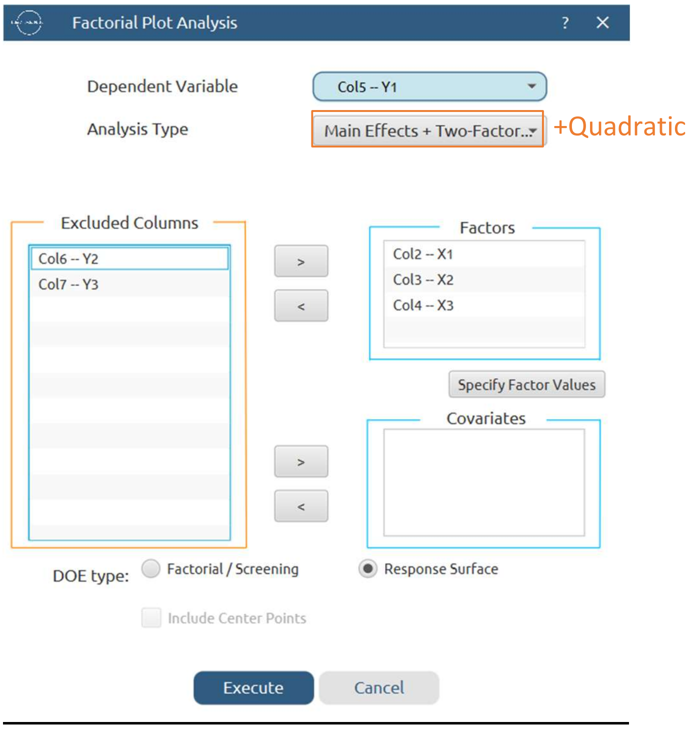
	Col1	Col2 (D)	Col3 (D)	Col4 (D)	Col5 (D)	Col6 (I)	Col7 (D)
User Header	User Row ID	X1	X2	X3	Y1	Y2	Y3
1		1000.0	30.0	92.5	95.35	3035	68
2		2000.0	30.0	92.5	166.76	2936	55.2
3		1000.0	50.0	92.5	15.96	3755	70
4		2000.0	50.0	92.5	41.07	2185	67
5		1000.0	40.0	74.0	42.1	2526	69
6		2000.0	40.0	74.0	83.57	3480	57.7
7		1000.0	40.0	111.0	31.64	3075	65.1
8		2000.0	40.0	111.0	71.79	2280	54.7
9		1500.0	30.0	74.0	206.6	4519	52.7
10		1500.0	50.0	74.0	31.4	2754	70.2
11		1500.0	30.0	111.0	141.85	3022	70.8
12		1500.0	50.0	111.0	23.55	2974	67.7
13		1500.0	40.0	92.5	68.8	19375	61.1

Step 4: Factorial Analysis

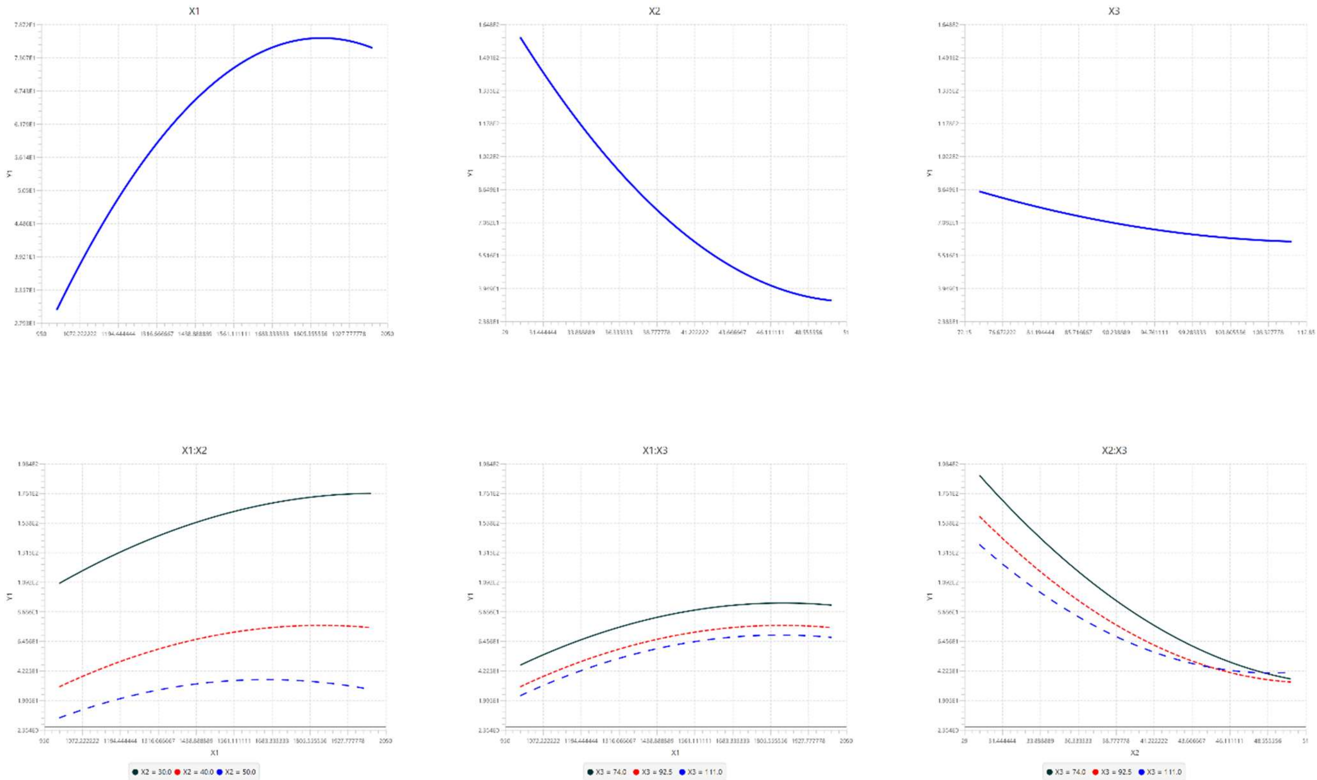
Create a new tab named "Factorial analysis – Y1" and import the results from the "Data" spreadsheet. Afterwards, conduct factorial analysis for the first response variable, Y_1, to produce main effects and interaction plots: *DOE → Post DoE Analysis → Factorial Plot Analysis*



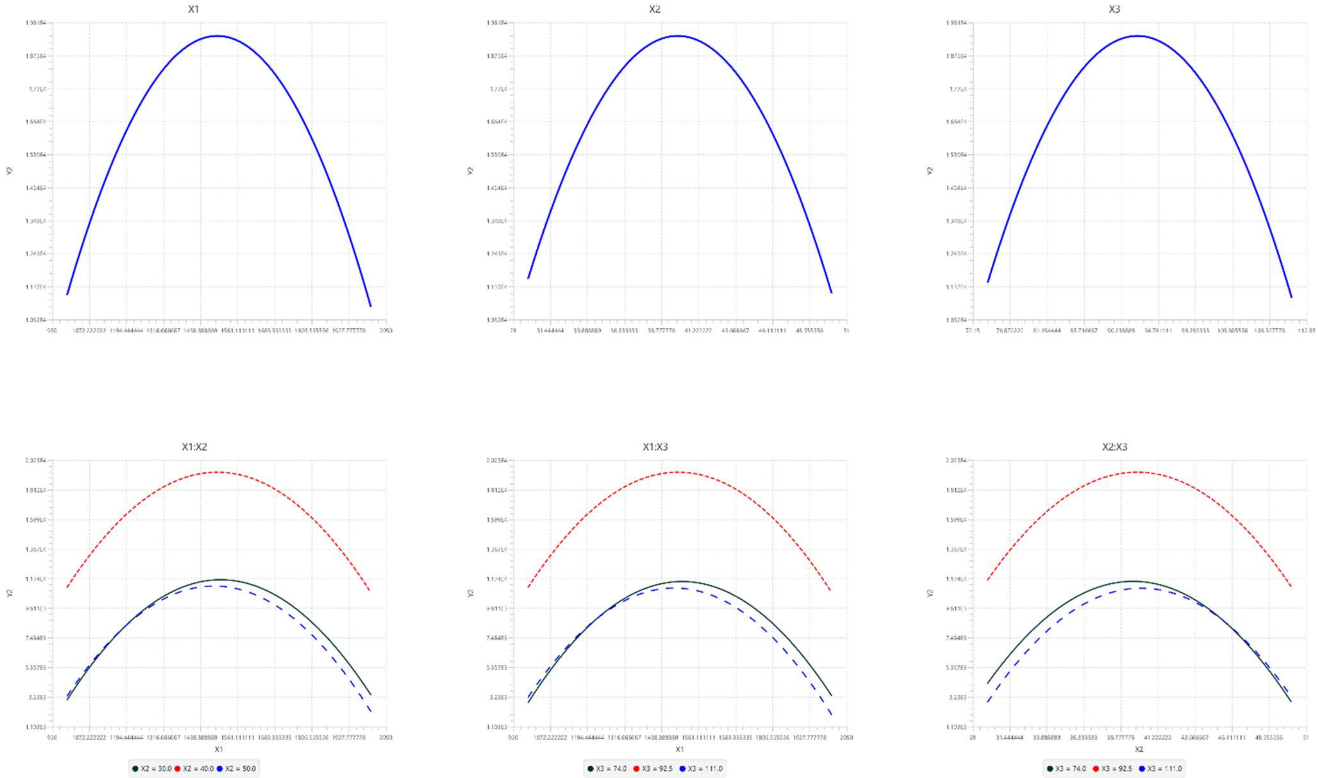




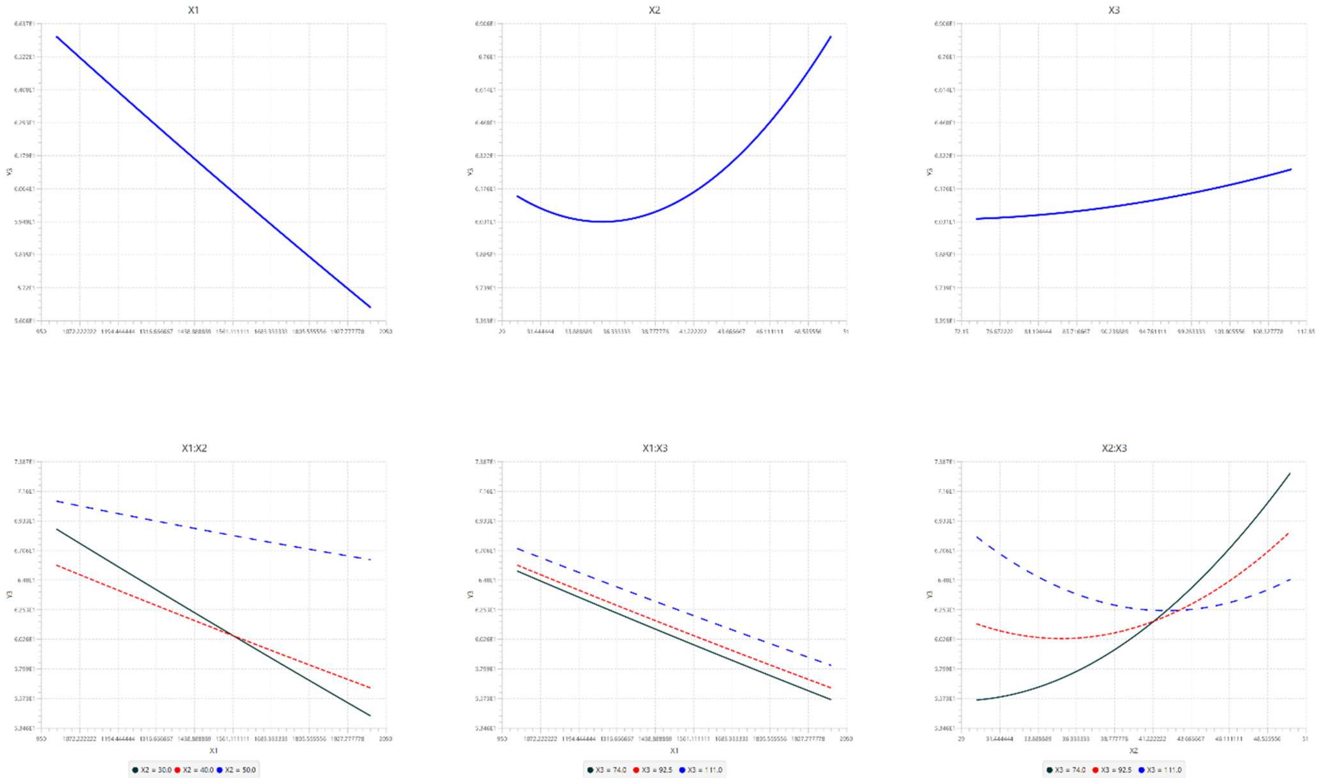
Results (also given in table format, not displayed here):



Repeat this step for the rest of the response variables. Results, Y_2:

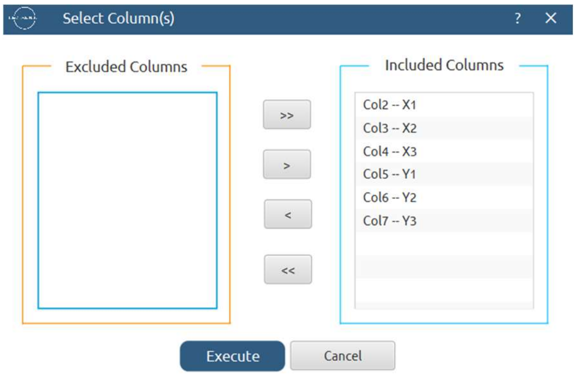


Results, Y_3:



Step 5: Define additional experimental points

Create a new tab named “Full data” and import the results from the “Data” spreadsheet. Then, add in each column the values of the additional experimental runs that were executed and select all the columns to be transferred to the right spreadsheet: *Data Transformation → Data Manipulation → Select Column(s)*

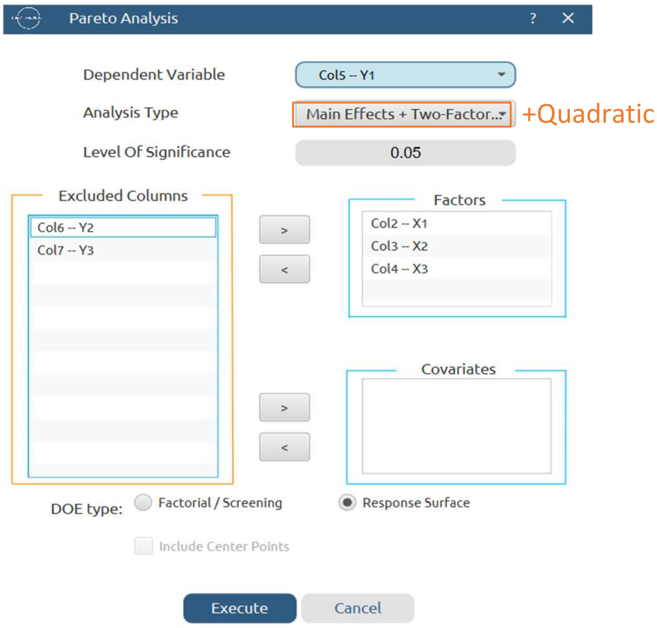


Results:

	Col1	Col2 (D)	Col3 (D)	Col4 (D)	Col5 (D)	Col6 (I)	Col7 (D)
User Header	User Row ID	X1	X2	X3	Y1	Y2	Y3
1		1000.0	30.0	92.5	95.35	3035	68
2		2000.0	30.0	92.5	166.76	2936	55.2
3		1000.0	50.0	92.5	15.96	3755	70
4		2000.0	50.0	92.5	41.07	2185	67
5		1000.0	40.0	74.0	42.1	2526	69
6		2000.0	40.0	74.0	83.57	3480	57.7
7		1000.0	40.0	111.0	31.64	3075	65.1
8		2000.0	40.0	111.0	71.79	2280	54.7
9		1500.0	30.0	74.0	206.6	4519	52.7
10		1500.0	50.0	74.0	31.4	2754	70.2
11		1500.0	30.0	111.0	141.85	3022	70.8
12		1500.0	50.0	111.0	23.55	2974	67.7
13		1500.0	40.0	92.5	68.8	19375	61.1
14		1000	50	74	20.48	2628	74
15		1000	50	111	22.4	2951	61.9
16		2000	50	74	40.97	2749	74.7
17		1000	30	74	67.02	637	65.8
18		1000	35	74	57.21	1775	76
19		1000	45	74	27.79	2080	73.5
20		2000	50	111	45.4	1008	64.7
21		1000	30	111	87.7	2913	60.8

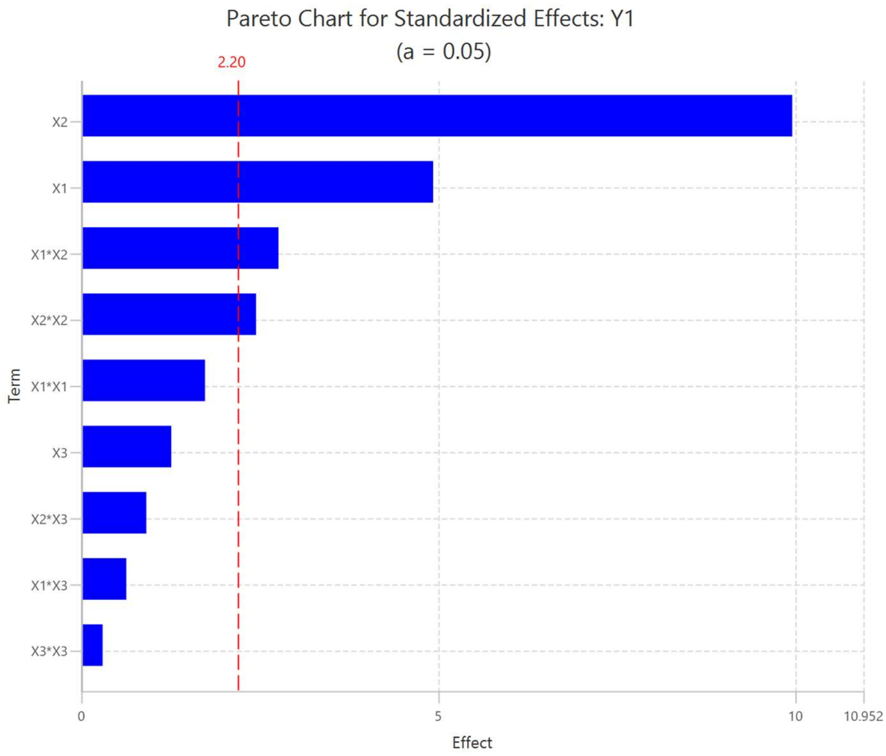
Step 6: Pareto Analysis

Create a new tab named "Pareto analysis – Y1" and import the results from the "Full data" spreadsheet. Afterwards, conduct pareto analysis for the first response variable, Y_1: *DOE → Post DoE Analysis → Pareto Analysis*



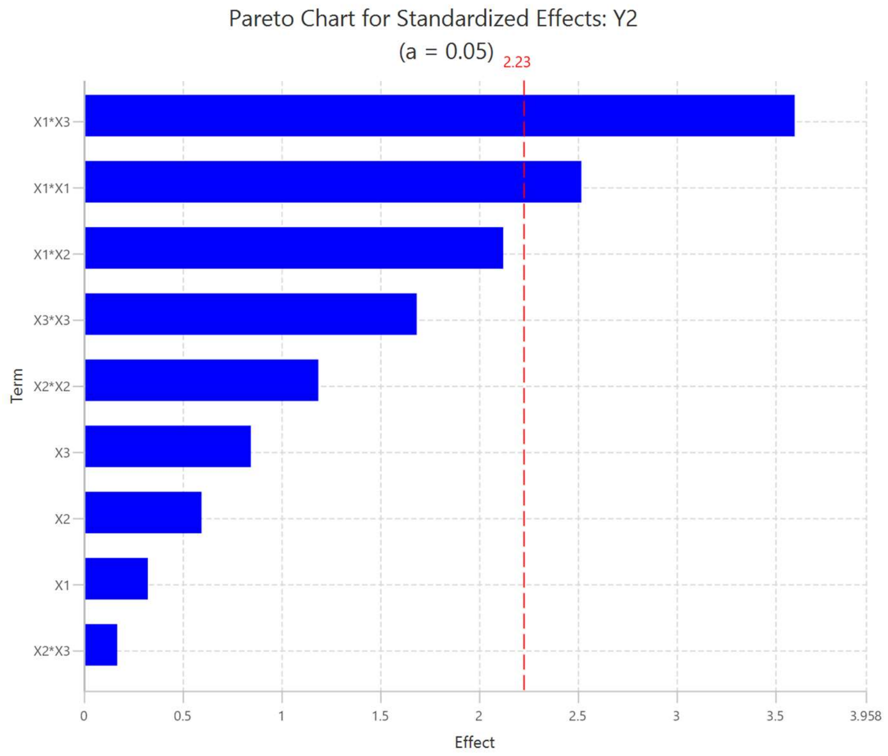
Results:

	Col1	Col2 (S)	Col3 (S)
User Header	User Row ID	Pareto Analysis of :	Standardized Effects
1		Variable	Effect
2		X2	9.9562015
3		X1	4.9272968
4		X1*X2	2.7611431
5		X2*X2	2.4477405
6		X1*X1	1.7325479
7		X3	1.2596017
8		X2*X3	0.9104352
9		X1*X3	0.6306287
10		X3*X3	0.2984454
11		Significance Value	2.2009852



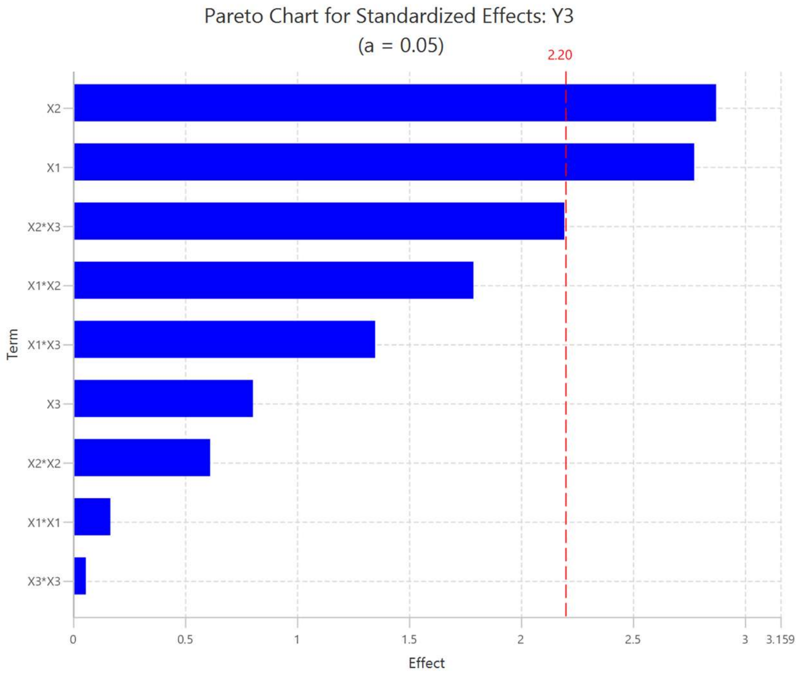
Repeat this step for the rest of the response variables. The experimental point (1500, 40, 92.5) is not included in the analysis of Y_2. Results, Y_2:

	Col1	Col2 (S)	Col3 (S)
User Header	User Row ID	Pareto Analysis of :	Standardized Effects
1		Variable	Effect
2		X1*X3	3.5978633
3		X1*X1	2.5183154
4		X1*X2	2.1227279
5		X3*X3	1.6849455
6		X2*X2	1.1867544
7		X3	0.8451740
8		X2	0.5954917
9		X1	0.3239851
10		X2*X3	0.1689221
11		Significance Value	2.2281389



Results, Y_3:

	Col1	Col2 (S)	Col3 (S)
User Header	User Row ID	Pareto Analysis of :	Standardized Effects
1		Variable	Effect
2		X2	2.8715101
3		X1	2.7737042
4		X2*X3	2.1939092
5		X1*X2	1.7877166
6		X1*X3	1.3488904
7		X3	0.8032723
8		X2*X2	0.6122096
9		X1*X1	0.1664083
10		X3*X3	0.0569245
11		Significance Value	2.2009852

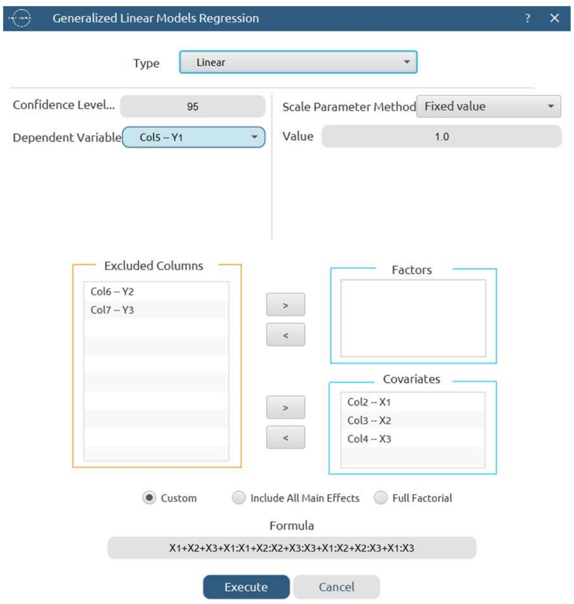


Step 7: Regression

The goal here is to produce a regression equation that includes main effects, two-factor interactions, and quadratic effects for Y_1:

$$Y = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + b_{12}X_1X_2 + b_{13}X_1X_3 + b_{23}X_2X_3 + b_{11}{X_1}^2 + b_{22}{X_2}^2 + b_{33}{X_3}^2$$

Create a new tab named "Regression – Y1" and import the results from the spreadsheet "Full data". Afterwards, fit a generalized linear model to the data: *Analytics → Regression → Statistical fitting → Generalized Linear Models*



Results:

Y1	Prediction
95.35	89.2177719
166.76	178.7443840
15.96	15.4639319
41.07	32.2093374
42.1	33.6057753
83.57	94.4561775
31.64	28.1189956
71.79	73.5406111
206.6	167.3378175
31.4	45.8110682
141.85	142.7543384
23.55	43.9922011
68.8	72.3045748
20.48	15.6115763
22.4	21.5071025
40.97	40.0713752
67.02	100.7477223
57.21	61.0332803
27.79	18.4652073
45.4	30.5381147
87.7	83.8786366

Goodness of Fit	
	Value
Deviance	4146.2045278
Scaled Deviance	4146.2045278
Pearson Chi-Square	4146.2045278
Scaled Pearson Chi-Square	4146.2045278
Log Likelihood	-2092.399973 1
Akaike's Information Criterion (AIC)	4204.7999462
Finite Sample Corrected AIC (AICC)	4226.7999462
Bayesian Information Criterion (BIC)	4215.2451705
Consistent AIC (CAIC)	4225.2451705

Parameter Estimates							
Variable	Coefficient	Std. Error	Lower CI	Upper CI	Test Statistic	df	p-value
intercept	392.3547113	18.6186338	355.8628595	428.8465630	444.0809286	1	0.0
X1	0.4529055	0.0071356	0.4389200	0.4668910	4028.6161623	1	0.0
X2	-22.5533074	0.4407998	-23.4172591	-21.6893556	2617.8074820	1	0.0
X3	-2.6350099	0.2972705	-3.2176493	-2.0523706	78.5707707	1	0.0
X1*X3	-0.0004170	0.0000341	-0.0004837	-0.0003502	149.9013141	1	0.0
X1*X2	-0.0036391	0.0000679	-0.0037721	-0.0035060	2873.6632037	1	0.0
X2*X3	0.0307630	0.0017404	0.0273519	0.0341741	312.4324160	1	0.0
X1*X1	-0.0000719	0.0000021	-0.0000761	-0.0000677	1131.4322093	1	0.0
X2*X2	0.2457387	0.0051711	0.2356037	0.2558738	2258.3371856	1	0.0
X3*X3	0.0090443	0.0015609	0.0059849	0.0121036	33.5728150	1	0E-7

Repeat this step for the rest of the response variables. The experimental point (1500, 40, 92.5) is not included in the analysis of Y_2. Results, Y_2:

Y2	Prediction
3035	2511.1317282
2936	3520.0637215
3755	3482.2284313
2185	2397.5761190
2526	2346.0969508
3480	3514.3004756
3075	3152.3004756
2280	1908.3766317
4519	3571.1317282
2754	3463.2284313
3022	3139.0637215
2974	3095.5761190
2628	2314.2881524
2628	2314.2881524
2951	3152.6995244
2749	2435.6995244
637	1375.3992965
1775	1986.0614302
2080	2455.5058582
1008	861.9835278

Goodness of Fit	
	Value
Deviance	3430502.1900374
Scaled Deviance	3430502.1900374
Pearson Chi-Square	3430502.1900374
Scaled Pearson Chi-Square	3430502.1900374
Log Likelihood	-1715269.4737894
Akaike's Information Criterion (AIC)	3430558.9475787
Finite Sample Corrected AIC (AICC)	3430583.3920232
Bayesian Information Criterion (BIC)	3430568.9049015
Consistent AIC (CAIC)	3430578.9049015

Parameter Estimates							
Variable	Coefficient	Std. Error	Lower CI	Upper CI	Test Statistic	df	p-value
intercept	-45761.6105001	27.3417421	-45815.1993300	-45708.0216703	2801239.4542347	1	0.0
X1	23.2384420	0.0087570	23.2212786	23.2556054	7042141.7481124	1	0.0
X2	546.1846697	0.5481458	545.1103237	547.2590156	992857.4414387	1	0.0
X3	488.2213882	0.3773902	487.4817170	488.9610594	1673602.7444392	1	0.0
X1*X3	-0.0651926	0.0000347	-0.0652607	-0.0651246	3527874.6310856	1	0.0
X1*X2	-0.1046792	0.0000699	-0.1048161	-0.1045423	2245067.5501817	1	0.0
X2*X3	0.0870482	0.0018882	0.0833475	0.0907490	2125.3570056	1	0.0
X1*X1	-0.0043529	0.0000028	-0.0043585	-0.0043474	2359028.1853935	1	0.0
X2*X2	-5.0125323	0.0066476	-5.0255613	-4.9995033	568576.6483036	1	0.0
X3*X3	-2.1876833	0.0020702	-2.1917408	-2.1836257	1116718.9169033	1	0.0

Results, Y_3:

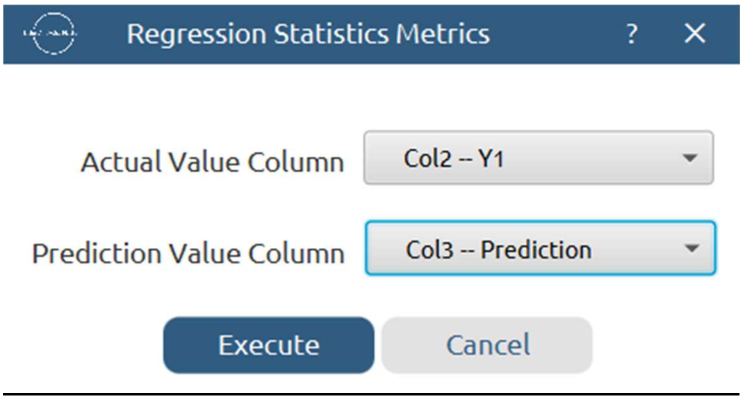
Y3	Prediction
68	67.3607889
55.2	53.2337082
70	69.5287257
67	67.8510157
69	69.9567461
57.7	57.6949867
65.1	63.3732505
54.7	59.8302191
52.7	58.0861479
70.2	73.7251030
70.8	63.1083492
67.7	64.2546384
61.1	63.3257615
74	76.2876595
61.9	62.4578309
74.7	70.2505855
65.8	66.8733898
76	68.0091233
73.5	72.7162582
64.7	65.1394849
60.8	67.5362271

Goodness of Fit	
	Value
Deviance	289.3926652
Scaled Deviance	289.3926652
Pearson Chi-Square	289.3926652
Scaled Pearson Chi-Square	289.3926652
Log Likelihood	-163.9940418
Akaike's Information Criterion (AIC)	347.9880836
Finite Sample Corrected AIC (AICC)	369.9880836
Bayesian Information Criterion (BIC)	358.4333080
Consistent AIC (CAIC)	368.4333080

Parameter Estimates							
Variable	Coefficient	Std. Error	Lower CI	Upper CI	Test Statistic	df	p-value
intercept	79.5115722	18.6186338	43.0197205	116.0034239	18.2374897	1	0.0000195
X1	-0.0491262	0.0071356	-0.0631117	-0.0351407	47.3987499	1	0E-7
X2	-0.0015113	0.4407998	-0.8654630	0.8624404	0.0000118	1	0.9972644
X3	0.4541275	0.2972705	-0.1285119	1.0367669	2.3337381	1	0.1265975
X1*X3	0.0002356	0.0000341	0.0001689	0.0003024	47.8683153	1	0E-7
X1*X2	0.0006225	0.0000679	0.0004894	0.0007555	84.0798952	1	0.0
X2*X3	-0.0195847	0.0017404	-0.0229958	-0.0161736	126.6286996	1	0.0
X1*X1	-0.0000018	0.0000021	-0.0000060	0.0000024	0.7285252	1	0.3933616
X2*X2	0.0162378	0.0051711	0.0061027	0.0263729	9.8604136	1	0.0016887
X3*X3	-0.0004558	0.0015609	-0.0035151	0.0026036	0.0852496	1	0.7703054

Step 8: Regression Metrics

Create a tab named “Metrics – Y1” and import the results from the spreadsheet “Regression – Y1”. Then, produce the regression metrics for the Y_1 regression equation: ***Statistics → Model Metrics → Regression Metrics***



Results:

	Col1	Col2 (D)	Col3 (D)	Col4 (D)	Col5 (D)
User Header	User Row ID	Mean Squared Error	Root Mean Squared Error	Mean Absolute Error	R Squared
1		197.4383108	14.0512743	9.6604150	0.9199492

Repeat this step for the rest of the response variables. Results, Y_2:

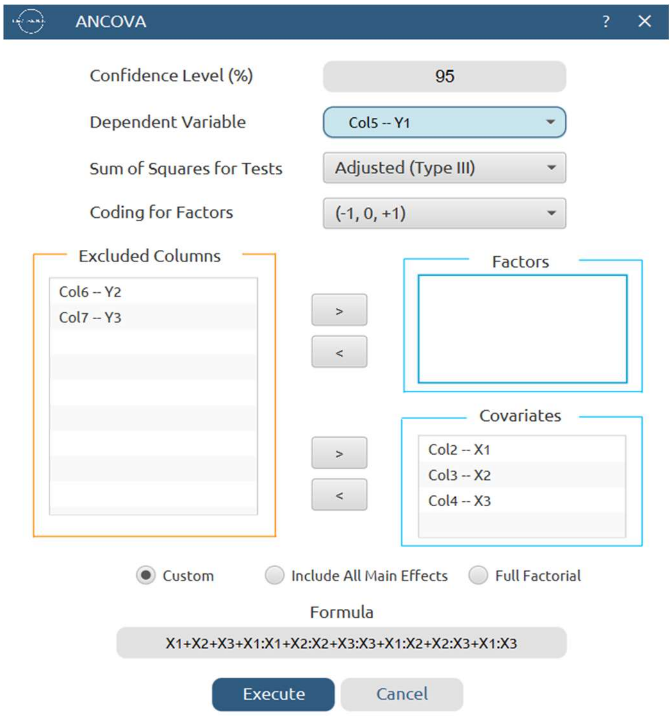
	Col1	Col2 (D)	Col3 (D)	Col4 (D)	Col5 (D)
User Header	User Row ID	Mean Squared Error	Root Mean Squared Error	Mean Absolute Error	R Squared
1		171525.1095019	414.1559000	338.2775173	0.7613835

Results, Y_3:

	Col1	Col2 (D)	Col3 (D)	Col4 (D)	Col5 (D)
User Header	User Row ID	Mean Squared Error	Root Mean Squared Error	Mean Absolute Error	R Squared
1		13.7806031	3.7122235	2.7780558	0.6883665

Step 9: Analysis of Covariance

Create a new tab named “ANCOVA – Y1” and import the results from the spreadsheet “Full data”. Afterwards perform analysis of covariance for Y_1: *Statistics → Analysis of (Co)Variance → ANCOVA*



Results:

	Col1	Col2 (S)	Col3 (I)	Col4 (D)	Col5 (D)	Col6 (D)	Col7 (D)
User Header	User Row ID	Source	DF	Adj SS	Adj MS	F-Value	P-Value
1		X1	1	4028.6161623	4028.6161623	10.6880347	0.0074751
2		X2	1	2617.8074820	2617.8074820	6.9451186	0.0231844
3		X3	1	78.5707707	78.5707707	0.2084505	0.6568679
4		X1*X1	1	1131.4322093	1131.4322093	3.0017222	0.1110818
5		X2*X2	1	2258.3371856	2258.3371856	5.9914336	0.0323753
6		X3*X3	1	33.5728150	33.5728150	0.0890696	0.7709254
7		X1*X2	1	2873.6632037	2873.6632037	7.6239112	0.0185180
8		X2*X3	1	312.4324160	312.4324160	0.8288922	0.3821102
9		X1*X3	1	149.9013141	149.9013141	0.3976925	0.5411636
10		Error	11	4146.2045278	376.9276843		
11		Total	20	51794.6763810			

Repeat this step for the rest of the response variables. Results, Y_2:

	Col1	Col2 (S)	Col3 (I)	Col4 (D)	Col5 (D)	Col6 (D)	Col7 (D)
User Header	User Row ID	Source	DF	Adj SS	Adj MS	F-Value	P-Value
1		X1	1	7049092.9333906	7049092.9333906	19.9023916	0.0012139
2		X2	1	896368.1553144	896368.1553144	2.5308036	0.1427276
3		X3	1	1613206.6235647	1613206.6235647	4.5547236	0.0586071
4		X1*X1	1	2246198.9621435	2246198.9621435	6.3419126	0.0304752
5		X2*X2	1	498826.6379497	498826.6379497	1.4083859	0.2627524
6		X3*X3	1	1005540.7389776	1005540.7389776	2.8390412	0.1229010
7		X1*X2	1	1595940.2598835	1595940.2598835	4.5059738	0.0597486
8		X2*X3	1	10106.4983269	10106.4983269	0.0285347	0.8692259
9		X1*X3	1	4584767.2486652	4584767.2486652	12.9446205	0.0048651
10		Error	10	3541832.0924196	354183.2092420		
11		Total	19	14441345.8			

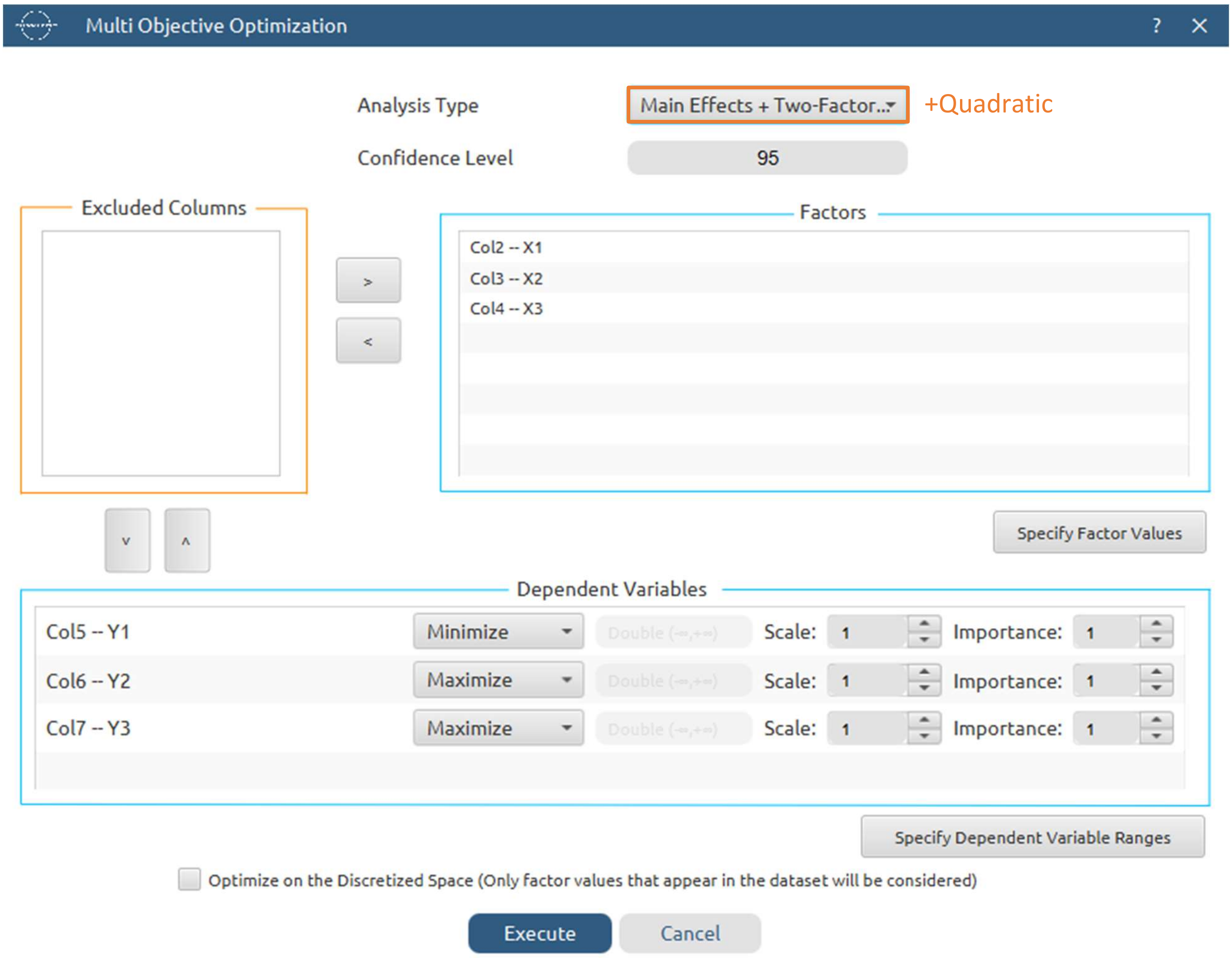
Results, Y_3:

	Col1	Col2 (S)	Col3 (I)	Col4 (D)	Col5 (D)	Col6 (D)	Col7 (D)
User Header	User Row ID	Source	DF	Adj SS	Adj MS	F-Value	P-Value
1		X1	1	47.3987499	47.3987499	1.8016568	0.2065587
2		X2	1	0.0000118	0.0000118	4E-7	0.9994786
3		X3	1	2.3337381	2.3337381	0.0887069	0.7713775
4		X1*X1	1	0.7285252	0.7285252	0.0276917	0.8708550
5		X2*X2	1	9.8604136	9.8604136	0.3748006	0.5528435
6		X3*X3	1	0.0852496	0.0852496	0.0032404	0.9556262
7		X1*X2	1	84.0798952	84.0798952	3.1959305	0.1013706
8		X2*X3	1	126.6286996	126.6286996	4.8132377	0.0506216
9		X1*X3	1	47.8683153	47.8683153	1.8195052	0.2044852
10		Error	11	289.3926652	26.3084241		
11		Total	20	928.6314286			

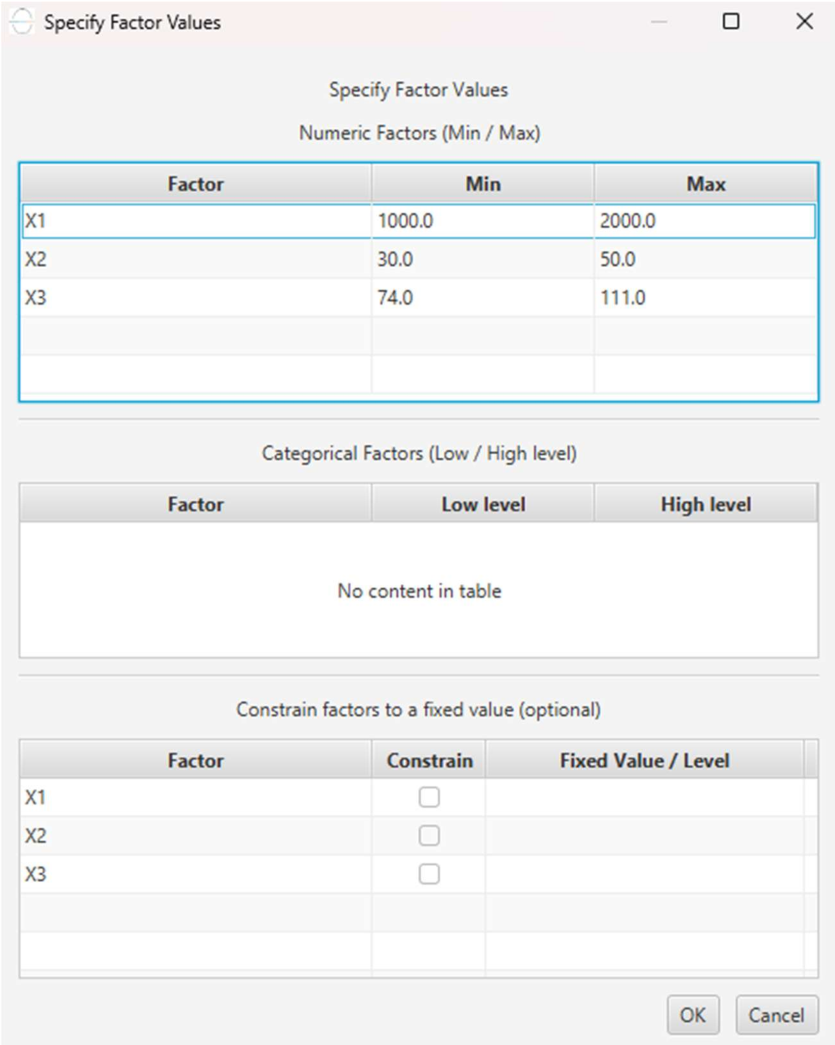
Step 10: Multi-Objective Optimization

Create a new tab named "Multi-objective optimization" and import the results from the spreadsheet "Full data". The goal of this step is to evaluate the factor values which will optimize all the response variables simultaneously. To perform multi-objective optimization, click on: *DOE → Post DoE Analysis → Multi-Objective Optimization*

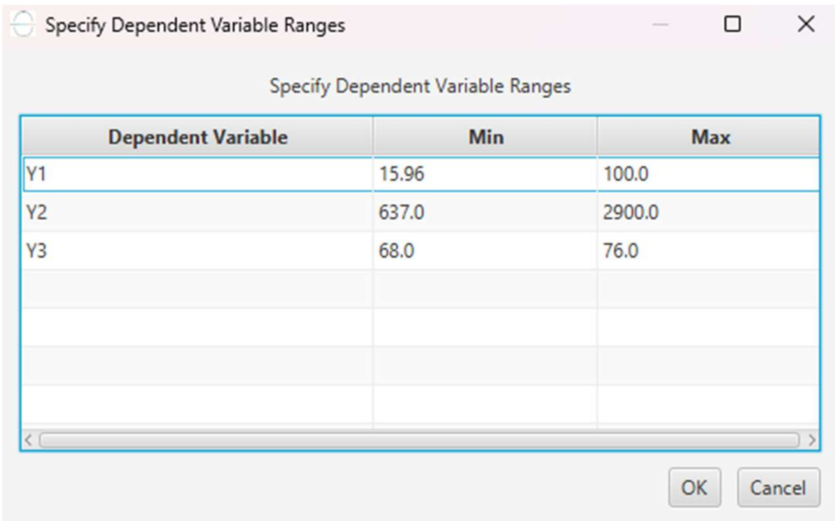
Select all three factors to be included in the optimization, and choose which response variables should be minimized and maximized. Reaction time, Y_1, is the only response variable which should be minimized. Leave scale and importance values at 1.



Specify the range for the factor values and the dependent variables as shown below.



Specify Factor Values
Specify Factor Values
Numeric Factors (Min / Max)
Factor
Min
Max
X1
1000.0
2000.0
X2
30.0
50.0
X3
74.0
111.0
Categorical Factors (Low / High level)
Factor
Low level
High level
No content in table
Constrain factors to a fixed value (optional)
Factor
Constrain
Fixed Value / Level
X1
X2
X3
OK
Cancel



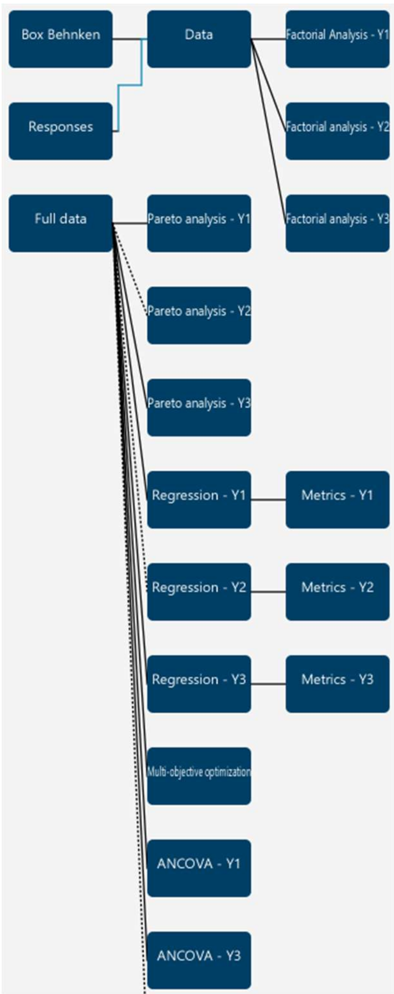
Specify Dependent Variable Ranges
Specify Dependent Variable Ranges
Dependent Variable
Min
Max
Y1
15.96
100.0
Y2
637.0
2900.0
Y3
68.0
76.0
OK
Cancel

Results:

	Col1	Col2 (S)	Col3 (S)	Col4 (S)	Col5 (S)	Col6 (S)	Col7 (S)	Col8 (S)
User Header	User Row ID							
1		Optimal Settings - Solution						
2		X1	X2	X3				
3		1113.1788258	50.0	76.0000000				
4								
5		Response Prediction						
6		Response	Predicted Value	Standard Error	(95.0%) CI Lower	(95.0%) CI Upper	(95.0%) PI Lower	(95.0%) PI Upper
7		Y1	25.1854821	11.2779124	0.3629643	50.0079998	-24.2323448	74.6033090
8		Y2	2900.0000000	1655.4345069	-743.5867833	6543.5867833	-4353.822629 4	10153.822629 4
9		Y3	75.1251399	2.9795256	68.5672484	81.6830314	62.0693834	88.1808964
10								
11		Desirabilities						
12		Y1	Y2	Y3	Overall			
13		0.8902251	1.0	0.8906425	0.9255526			

Final Isalos Workflow

The final workflow is presented below:



References

(1) Michael, A.; Adamou, P.; Harkou, E.; Christodoulou, C.; Barlocco, I.; Mintis, D.; Afantitis, A.; Delgado, J. J.; Chen, X.; Manos, G.; Dimitratos, N.; Villa, A.; Constantinou, A. Application of CFD, Statistical Analysis and DoE for Optimising H2 Generation through Ammonia Borane Catalytic Hydrolysis: A Novel Approach for Multiobjective Optimisation. *Chemical Engineering Journal* **2025**, *525*, 170772. https://doi.org/10.1016/j.cej.2025.170772.