



Application of Analytical Quality by Design concept for bilastine and its degradation impurities determination by hydrophilic interaction liquid chromatographic method

The objective of this research paper is to develop a hydrophilic interaction liquid chromatographic method for the analysis of bilastine and its degradation impurities following Analytical Quality by Design approach. Two Design of Experiments (DoE) methodologies are implemented for method optimization and robustness testing.

The DoE method applied for optimization is Box Behnken design. The factors (independent variables) examined in this stage are: X_1 = acetonitrile content in the mobile phase (%), X_2 = ammonium acetate concentration in the aqueous phase (mmol/L) and X_3 = pH of the aqueous phase. All the factors are continuous. The responses (dependent variables) examined are: k_3 = retention factor of Impurity 2 and α_2 = selectivity factor of critical peak pair.

The DoE method applied for robustness testing is fractional factorial design. The factors examined in this stage are: A = acetonitrile content in the mobile phase (%), B = buffer concentration in aqueous phase (mM), C = pH of the aqueous phase, D = mobile phase flow rate (mL/min) and E = column temperature (°C). All the factors are continuous and the responses examined are the same as those of the first stage.

Isalos version used: 2.1.2

Scientific article: <https://www.sciencedirect.com/science/article/abs/pii/S0731708516302084>

Optimization stage

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 (D)
User Header	User Row ID	X1	X2	X3
1		90	40	4
2		94	80	5.5

 DoE Box Behnken ? ×

Number of Center Points per Block

3

Number of Replicates

1

Number of Blocks

1

☐ Random Standard order

Excluded Columns

>>

>

<

<<

Included Columns

Col2 -- X1

Col3 -- X2

Col4 -- X3

Execute

Cancel

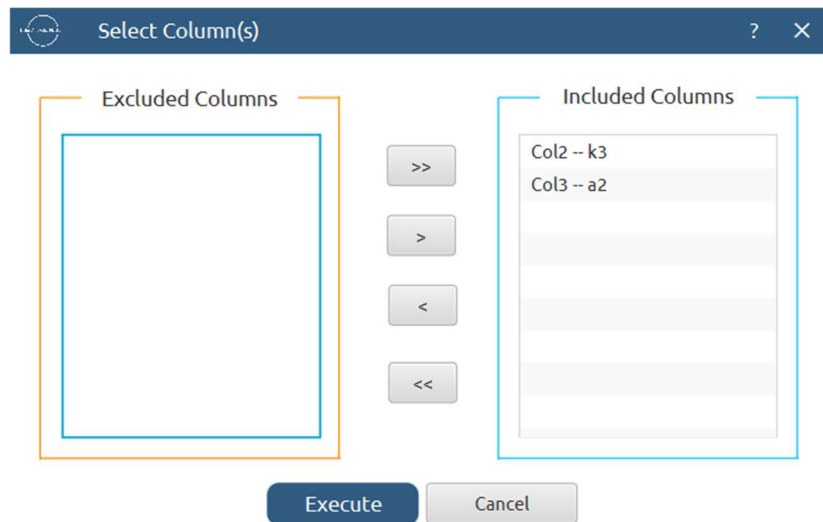
Results (right spreadsheet):

	Col1	Col2 (I)	Col3 (S)	Col4 (S)	Col5 (S)	Col6 (D)	Col7 (D)	Col8 (D)
User Header	User Row ID	Standard Order	Block Number	Replicate Number	Point Type	X1	X2	X3
1		1	Block: 1	Replicate: 1	Design Point	90.0	40.0	4.75
2		2	Block: 1	Replicate: 1	Design Point	94.0	40.0	4.75
3		3	Block: 1	Replicate: 1	Design Point	90.0	80.0	4.75
4		4	Block: 1	Replicate: 1	Design Point	94.0	80.0	4.75
5		5	Block: 1	Replicate: 1	Design Point	90.0	60.0	4.0
6		6	Block: 1	Replicate: 1	Design Point	94.0	60.0	4.0
7		7	Block: 1	Replicate: 1	Design Point	90.0	60.0	5.5
8		8	Block: 1	Replicate: 1	Design Point	94.0	60.0	5.5
9		9	Block: 1	Replicate: 1	Design Point	92.0	40.0	4.0
10		10	Block: 1	Replicate: 1	Design Point	92.0	80.0	4.0
11		11	Block: 1	Replicate: 1	Design Point	92.0	40.0	5.5
12		12	Block: 1	Replicate: 1	Design Point	92.0	80.0	5.5
13		13	Block: 1	----	Center Point	92.0	60.0	4.75
14		14	Block: 1	----	Center Point	92.0	60.0	4.75
15		15	Block: 1	----	Center Point	92.0	60.0	4.75

Step 2: Definition of response variables

Create a new tab named “Responses - BBD” 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: *Data Transformation* → *Data Manipulation* → *Select Column(s)*

	Col1	Col2 (D)	Col3 (D)
User Header	User Row ID	k3	a2
1		3.53	1.5
2		23.65	1.65
3		3.47	1.38
4		20.19	1.5
5		1.34	1.26
6		8.78	1.36
7		4.24	1.43
8		28.96	1.64
9		3.63	1.33
10		2.9	1.23
11		9.92	1.72
12		10.53	1.61
13		7.34	1.47
14		7.08	1.46
15		7.86	1.51



The "Select Column(s)" dialog box features two main panels: "Excluded Columns" on the left and "Included Columns" on the right. The "Included Columns" panel contains a list with two entries: "Col2 -- k3" and "Col3 -- a2". Between the panels are four arrow buttons: ">>", ">", "<", and "<<". At the bottom, there are "Execute" and "Cancel" buttons.

Step 3: Data selection

Create a new tab named "Data - BBD" and import the results from the "Box Behnken" and "Responses - BBD" 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)*

	Col1	Col2	Col3	Col4	Col5	Col6
User Header	User Row ID					
1						
2						
3						
4						
5						
6						
7						
8						
9						
10						

Import from File

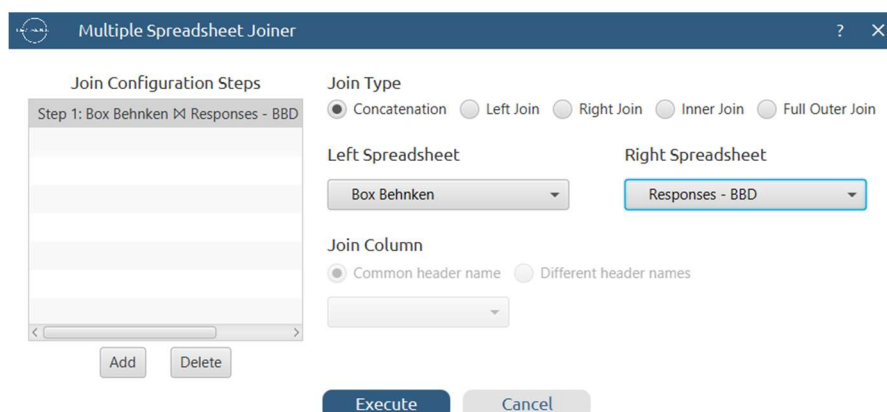
Import from Spreadsheet

Import from Multiple Spreadsheets

Adjust Spreadsheet Precision

Export Spreadsheet Data

Clear Spreadsheet



The "Multiple Spreadsheet Joiner" dialog box shows "Join Configuration Steps" on the left with "Step 1: Box Behnken X Responses - BBD". On the right, under "Join Type", "Concatenation" is selected. Below, "Left Spreadsheet" is set to "Box Behnken" and "Right Spreadsheet" is set to "Responses - BBD". Under "Join Column", "Common header name" is selected. At the bottom are "Execute" and "Cancel" buttons.

Select Column(s)

Excluded Columns

- Col2 -- Standard Order
- Col3 -- Block Number
- Col4 -- Replicate Number
- Col5 -- Point Type

Included Columns

- Col6 -- X1
- Col7 -- X2
- Col8 -- X3
- Col9 -- k3
- Col10 -- a2

>> > < <<

Execute Cancel

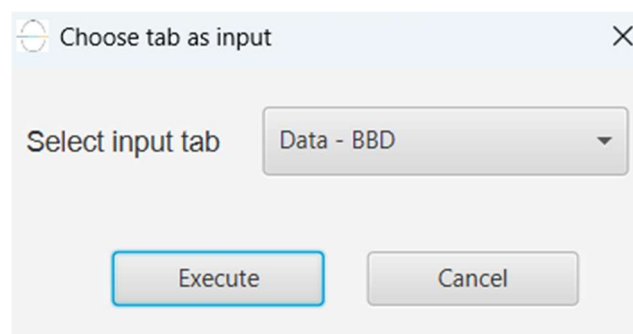
Results:

	Col1	Col2 (D)	Col3 (D)	Col4 (D)	Col5 (D)	Col6 (D)
User Header	User Row ID	X1	X2	X3	k3	a2
1		90.0	40.0	4.75	3.53	1.5
2		94.0	40.0	4.75	23.65	1.65
3		90.0	80.0	4.75	3.47	1.38
4		94.0	80.0	4.75	20.19	1.5
5		90.0	60.0	4.0	1.34	1.26
6		94.0	60.0	4.0	8.78	1.36
7		90.0	60.0	5.5	4.24	1.43
8		94.0	60.0	5.5	28.96	1.64
9		92.0	40.0	4.0	3.63	1.33
10		92.0	80.0	4.0	2.9	1.23
11		92.0	40.0	5.5	9.92	1.72
12		92.0	80.0	5.5	10.53	1.61
13		92.0	60.0	4.75	7.34	1.47
14		92.0	60.0	4.75	7.08	1.46
15		92.0	60.0	4.75	7.86	1.51

Step 4: Normalization

Create a new tab named “Normalized data - BBD” and import the results from the “Data - BBD” spreadsheet. Afterwards, normalize the factor columns to take values in the range $[-1, 1]$: *Data Transformation* → *Normalizers* → *Min-Max*

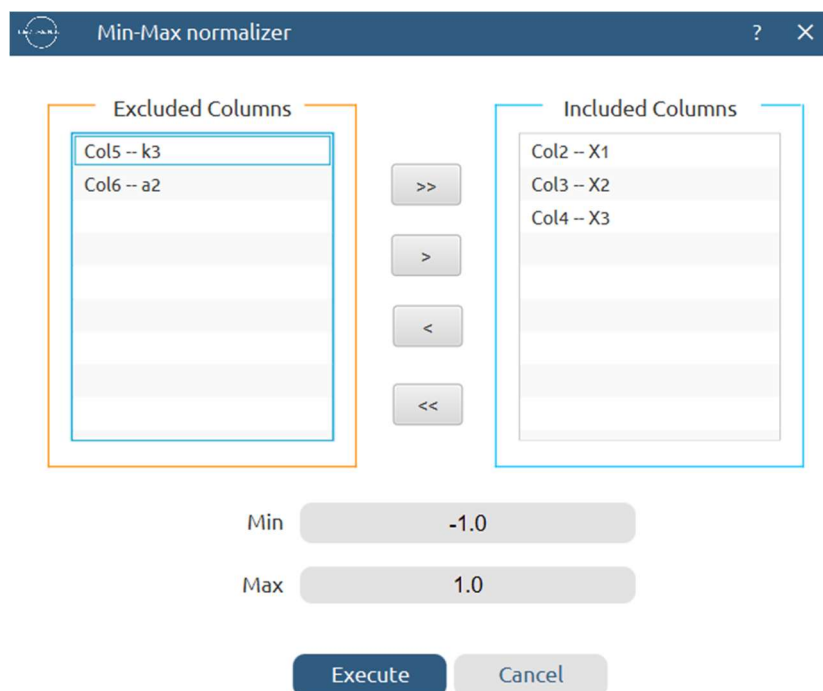
	Col1	Col2	Col3	Col4	Col5	Col6
User Header	User Row ID					
1						
2						
3						
4						
5						
6						
7						
8						
9						
10						



Choose tab as input

Select input tab: Data - BBD

Execute Cancel



Min-Max normalizer

Excluded Columns

- Col5 -- k3
- Col6 -- a2

Included Columns

- Col2 -- X1
- Col3 -- X2
- Col4 -- X3

Min: -1.0

Max: 1.0

Execute Cancel

Results:

	Col1	Col2 (D)	Col3 (D)	Col4 (D)	Col5 (D)	Col6 (D)
User Header	User Row ID	X1	X2	X3	k3	a2
1		-1.0	-1.0	0.0	3.53	1.5
2		1.0	-1.0	0.0	23.65	1.65
3		-1.0	1.0	0.0	3.47	1.38
4		1.0	1.0	0.0	20.19	1.5
5		-1.0	0.0	-1.0	1.34	1.26
6		1.0	0.0	-1.0	8.78	1.36
7		-1.0	0.0	1.0	4.24	1.43
8		1.0	0.0	1.0	28.96	1.64
9		0.0	-1.0	-1.0	3.63	1.33
10		0.0	1.0	-1.0	2.9	1.23
11		0.0	-1.0	1.0	9.92	1.72
12		0.0	1.0	1.0	10.53	1.61
13		0.0	0.0	0.0	7.34	1.47
14		0.0	0.0	0.0	7.08	1.46
15		0.0	0.0	0.0	7.86	1.51

Step 5: Regression

The goal here is to produce a regression equation that includes main effects, two-factor interactions, and quadratic effects for k_3 :

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

Create a new tab named “Regression – k_3 - BBD” and import the results from the spreadsheet “Normalized data - BBD”. Afterwards, fit a generalized linear model to the data: *Analytics* → *Regression* → *Statistical fitting* → *Generalized Linear Models*

The screenshot displays the 'Generalized Linear Models Regression' software interface. The main window shows settings for a Linear model with a Confidence Level of 95% and a Scale Parameter Method of Fixed value (1.0). The Dependent Variable is set to 'Col5 -- k3'. Below these settings, there are sections for 'Excluded Columns' (Col6 -- a2), 'Factors', 'Covariates' (Col2 -- X1, Col3 -- X2, Col4 -- X3), and 'Formula Type' (Custom selected). A 'Specify Custom Model Formula' button is present, and the formula field contains: $X1 + X2 + X3 + X1:X2 + X1:X3 + X2:X3 + X1:X1 + X2:X2 + X3:X3$.

The 'Custom Formula Creation' window is also visible, showing 'Included Columns' (X1, X2, X3) and 'Included Terms' (X1, X2, X3, X1:X2, X1:X3, X2:X3, X1:X1, X2:X2, X3:X3). It includes buttons for 'Add interactions', 'Add terms', 'Add Custom Interaction', 'Delete Selected', and 'Reset Selections'. The resulting formula is displayed at the bottom: $X1 + X2 + X3 + X1:X2 + X1:X3 + X2:X3 + X1:X1 + X2:X2 + X3:X3$.

Results:

k3	Prediction
3.53	3.6900000
23.65	22.6400000
3.47	4.4800000
20.19	20.0300000
1.34	1.9000000
8.78	10.5100000
4.24	2.5100000
28.96	28.4000000
3.63	2.9100000
2.9	1.3300000
9.92	11.4900000
10.53	11.2500000
7.34	7.4266667
7.08	7.4266667
7.86	7.4266667

Goodness of Fit	
	Value
Deviance	14.9864667
Scaled Deviance	14.9864667
Pearson Chi-Square	14.9864667
Scaled Pearson Chi-Square	14.9864667
Log Likelihood	-21.2773113
Akaike's Information Criterion (AIC)	62.5546227
Finite Sample Corrected AIC (AICC)	117.5546227
Bayesian Information Criterion (BIC)	69.6351247
Consistent AIC (CAIC)	79.6351247

Parameter Estimates							
Variable	Coefficient	Std. Error	Lower CI	Upper CI	Test Statistic	df	p-value
intercept	7.4266667	0.5773503	6.2950809	8.5582524	165.4661333	1	0.0
X1	8.6250000	0.3535534	7.9320481	9.3179519	595.1250000	1	0.0
X2	-0.4550000	0.3535534	-1.1479519	0.2379519	1.6562000	1	0.1981172
X3	4.625	0.3535534	3.9320481	5.3179519	171.1250000	1	0.0
X1*X3	4.32	0.5	3.3400180	5.2999820	74.6496	1	0.0
X1*X2	-0.8500000	0.5	-1.8299820	0.1299820	2.8900000	1	0.0891309
X2*X3	0.3350000	0.5	-0.6449820	1.3149820	0.4489000	1	0.5028578
X1*X1	4.6841667	0.5204165	3.6641691	5.7041643	81.0144641	1	0.0
X2*X2	0.5991667	0.5204165	-0.4208309	1.6191643	1.3255410	1	0.2496000
X3*X3	-1.2808333	0.5204165	-2.3008309	-0.2608357	6.0573564	1	0.0138485

Repeat this step for the second response variable. Results, a_2 :

a_2	Prediction
1.5	1.4875000
1.65	1.6475000
1.38	1.3825000
1.5	1.5125000
1.26	1.2250000
1.36	1.315
1.43	1.4750000
1.64	1.6750000
1.33	1.3775000
1.23	1.2625000
1.72	1.6875000
1.61	1.5625000
1.47	1.4800000
1.46	1.4800000
1.51	1.4800000

Goodness of Fit	
	Value
Deviance	0.0148500
Scaled Deviance	0.0148500
Pearson Chi-Square	0.0148500
Scaled Pearson Chi-Square	0.0148500
Log Likelihood	-13.7915030
Akaike's Information Criterion (AIC)	47.5830060
Finite Sample Corrected AIC (AICC)	102.5830060
Bayesian Information Criterion (BIC)	54.6635080
Consistent AIC (CAIC)	64.6635080

Parameter Estimates							
Variable	Coefficient	Std. Error	Lower CI	Upper CI	Test Statistic	df	p-value
intercept	1.4800000	0.5773503	0.3484143	2.6115857	6.5712000	1	0.0103642
X1	0.0725000	0.3535534	-0.6204519	0.7654519	0.0420500	1	0.8375245
X2	-0.0600000	0.3535534	-0.7529519	0.6329519	0.0288000	1	0.8652416
X3	0.1525000	0.3535534	-0.5404519	0.8454519	0.1860500	1	0.6662247
X1*X3	0.0275000	0.5	-0.9524820	1.0074820	0.0030250	1	0.9561385
X1*X2	-0.0075000	0.5	-0.9874820	0.9724820	0.0002250	1	0.9880322
X2*X3	-0.0025000	0.5	-0.9824820	0.9774820	0.0000250	1	0.9960106
X1*X1	-0.0112500	0.5204165	-1.0312476	1.0087476	0.0004673	1	0.9827532
X2*X2	0.0387500	0.5204165	-0.9812476	1.0587476	0.0055442	1	0.9406447
X3*X3	-0.0462500	0.5204165	-1.0662476	0.9737476	0.0078981	1	0.9291843

Step 6: Data selection for Regression Metrics

Create a tab named “Metrics Data – k3-BBD” and import the results from the spreadsheet “Regression – k3-BBD”. Next, assign names to all unnamed columns containing data, using names such as Z1–Z9. Then, select only the response column (k3) and the “Prediction” column to be transferred to the right spreadsheet: *Data Transformation* → *Data Manipulation* → *Select Column(s)*

	Col1	Col2 (D)	Col3 (D)	Col4	Col5 (S)	Col6 (S)	Col7 (S)	Col8 (S)	Col9 (S)	Col10 (S)	Col11 (S)	Col12 (S)
User Header	User Row ID	k3	Prediction	Z1	Z2	Z3	Z4	Z5	Z6	Z7	Z8	Z9
1		3.53	3.6900000		Goodness of Fit							
2		23.65	22.6400000			Value						
3		3.47	4.4800000		Deviance	14.9864667						
4		20.19	20.03		Scaled Deviance	14.9864667						
5		1.34	1.9000000		Pearson Chi-Square	14.9864667						
6		8.78	10.5100000		Scaled Pearson Chi-Square	14.9864667						
7		4.24	2.5100000		Log Likelihood	-21.2773113						
8		28.96	28.4000000		Akaike's Information Criterion (AIC)	62.5546227						
9		3.63	2.9100000		Finite Sample Corrected AIC (AICC)	117.5546227						
10		2.9	1.3300000		Bayesian Information Criterion (BIC)	69.6351247						
11		9.92	11.4900000		Consistent AIC (CAIC)	79.6351247						
12		10.53	11.2500000									
13		7.34	7.4266667									
14		7.08	7.4266667		Parameter Estimates							
15		7.86	7.4266667		Variable	Coefficient	Std. Error	Lower CI	Upper CI	Test Statistic	df	p-value
16					intercept	7.4266667	0.5773503	6.2950809	8.5582524	165.4661333	1	0.0

Results:

Select Column(s)

Select Columns Manually

Select Matching Columns from Node

Select Columns Manually and from Node

Select Node

Select Node

Excluded Columns

Col5 -- Z2

Col6 -- Z3

Col7 -- Z4

Col8 -- Z5

Col9 -- Z6

Col10 -- Z7

Col11 -- Z8

Col12 -- Z9

>>

>

<

<<

Included Columns

Col2 -- k3

Col3 -- Prediction

Execute

Cancel

User Header	Col1	Col2 (D)	Col3 (D)
1		3.53	3.6900000
2		23.65	22.6400000
3		3.47	4.4800000
4		20.19	20.03
5		1.34	1.9000000
6		8.78	10.5100000
7		4.24	2.5100000
8		28.96	28.4000000
9		3.63	2.9100000
10		2.9	1.3300000
11		9.92	11.4900000
12		10.53	11.2500000
13		7.34	7.4266667
14		7.08	7.4266667
15		7.86	7.4266667

Repeat this step for a2 as well.

Step 7: Regression Metrics

Create a tab named “Metrics – k_3 - BBD” and import the results from the spreadsheet “Metrics Data – k_3 - BBD”. Then, produce the regression metrics for the k_3 regression equation: *Statistics* → *Model Metrics* → *Regression Metrics*

Regression Statistics Metrics
?
×

Actual Value Column

Col2 – k_3

Prediction Value Column

Col3 – Prediction

Execute

Cancel

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		0.9990978	0.9995488	0.8244444	0.9842641

Repeat this step for the second response variable. Results, a_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		0.0009900	0.0314643	0.0273333	0.9487931

Step 8: Analysis of Covariance

Create a new tab named “ANCOVA – k3 - BBD” and import the results from the spreadsheet “Normalized data - BBD”. Afterwards perform analysis of covariance for k_3 : *Statistics* → *Analysis of (Co)Variance* → *ANCOVA*

ANCOVA

Confidence Level (%)

95

Dependent Variable

Col5 -- k3

Sum of Squares for Tests

Adjusted (Type III)

Coding for Factors

(-1, 0, +1)

Excluded Columns

Col6 -- a2

Factors

Covariates

Col2 -- X1
Col3 -- X2
Col4 -- X3

Formula Type:

☒ Custom
 ☐ Main Effects
 ☐ Full Factorial

Formula:

X1 + X2 + X3 + X1:X2 + X1:X3 + X2:X3 + X1:X1 + X2:X2 + X3:X3

Execute

Cancel

Custom Formula Creation

Included Columns

X1
X2
X3

Add interactions

up to order 2

Add terms

up to order 2

Add Custom Interaction

Delete Selected

Reset Selections

Included Terms

X1
X2
X3
X1:X2
X1:X3
X2:X3
X1:X1
X2:X2
X3:X3

Move up

Move down

X1 + X2 + X3 + X1:X2 + X1:X3 + X2:X3 + X1:X1 + X2:X2 + X3:X3

Save

Cancel

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	595.1250000	595.1250000	198.5541400	0.0000324
2		X2	1	1.6562000	1.6562000	0.5525652	0.4907024
3		X3	1	171.1250000	171.1250000	57.0931774	0.0006436
4		X1*X1	1	81.0144641	81.0144641	27.0292077	0.0034700
5		X2*X2	1	1.3255410	1.3255410	0.4422460	0.5354806
6		X3*X3	1	6.0573564	6.0573564	2.0209421	0.2144031
7		X1*X2	1	2.8900000	2.8900000	0.9642033	0.3712248
8		X2*X3	1	0.4489000	0.4489000	0.1497685	0.7146804
9		X1*X3	1	74.6496000	74.6496000	24.9056704	0.0041381
10		Error	5	14.9864667	2.9972933		
11		Total	14	952.3729733			

Repeat this step for the second response variable. Results, a_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	0.0420500	0.0420500	14.1582492	0.0131204
2		X2	1	0.0288000	0.0288000	9.6969697	0.0264310
3		X3	1	0.1860500	0.1860500	62.6430976	0.0005183
4		X1*X1	1	0.0004673	0.0004673	0.1573427	0.7079790
5		X2*X2	1	0.0055442	0.0055442	1.8667444	0.2300879
6		X3*X3	1	0.0078981	0.0078981	2.6592852	0.1638757
7		X1*X2	1	0.0002250	0.0002250	0.0757576	0.7941430
8		X2*X3	1	0.0000250	0.0000250	0.0084175	0.9304615
9		X1*X3	1	0.0030250	0.0030250	1.0185185	0.3591866
10		Error	5	0.0148500	0.0029700		
11		Total	14	0.2900000			

Robustness testing stage

Step 9: Fractional Factorial Design

Create a new tab named “Fractional Factorial” and define the factors of this stage in the column headers. Afterwards, fill each column with the low and high levels of the corresponding factors and apply the fractional factorial method: *DOE* → *Screening* → *Fractional Factorial*

	Col1	Col2 (I)	Col3 (I)	Col4 (D)	Col5 (D)	Col6 (I)
User Header	User Row ID	A	B	C	D	E
1		90	45	5.1	0.9	25
2		91	55	5.5	1.1	35

DoE Fractional Factorial

?

×

Available Designs

Custom

Fraction Relationship

a b c ab ac

Number of Center Points per Block

3

Number of Replicates

1

Number of Blocks

1

☐ Random Standard order

Excluded Columns

>>

>

<

<<

Included Columns

Col2 -- A
Col3 -- B
Col4 -- C
Col5 -- D
Col6 -- E

Execute

Cancel

Results:

	Col1	Col2 (I)	Col3 (S)	Col4 (S)	Col5 (S)	Col6 (D)	Col7 (D)	Col8 (D)	Col9 (D)	Col10 (D)
User Header	User Row ID	Standard Order	Block Number	Replicate Number	Point Type	A	B	C	D	E
1		1	Block: 1	Replicate: 1	Design Point	90.0	45.0	5.1	1.1	35.0
2		2	Block: 1	Replicate: 1	Design Point	91.0	45.0	5.1	0.9	25.0
3		3	Block: 1	Replicate: 1	Design Point	90.0	55.0	5.1	0.9	35.0
4		4	Block: 1	Replicate: 1	Design Point	91.0	55.0	5.1	1.1	25.0
5		5	Block: 1	Replicate: 1	Design Point	90.0	45.0	5.5	1.1	25.0
6		6	Block: 1	Replicate: 1	Design Point	91.0	45.0	5.5	0.9	35.0
7		7	Block: 1	Replicate: 1	Design Point	90.0	55.0	5.5	0.9	25.0
8		8	Block: 1	Replicate: 1	Design Point	91.0	55.0	5.5	1.1	35.0
9		9	Block: 1	----	Center Point	90.5	50.0	5.3	1.0	30.0
10		10	Block: 1	----	Center Point	90.5	50.0	5.3	1.0	30.0
11		11	Block: 1	----	Center Point	90.5	50.0	5.3	1.0	30.0

Step 10: Definition of response variables

Create a new tab named “Responses - FF” 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 fractional factorial method. Then, select all columns to be transferred to the right spreadsheet: *Data Transformation* → *Data Manipulation* → *Select Column(s)*

	Col1	Col2 (D)	Col3 (D)
User Header	User Row ID	k3	a2
1		3.32	1.55
2		6.18	1.51
3		4.4	1.49
4		4.97	1.54
5		3.6	1.53
6		5.73	1.55
7		4.5	1.47
8		5.17	1.59
9		4.52	1.49
10		4.6	1.49
11		4.29	1.5

Select Column(s) ? X

Excluded Columns

>>

>

<

<<

Included Columns

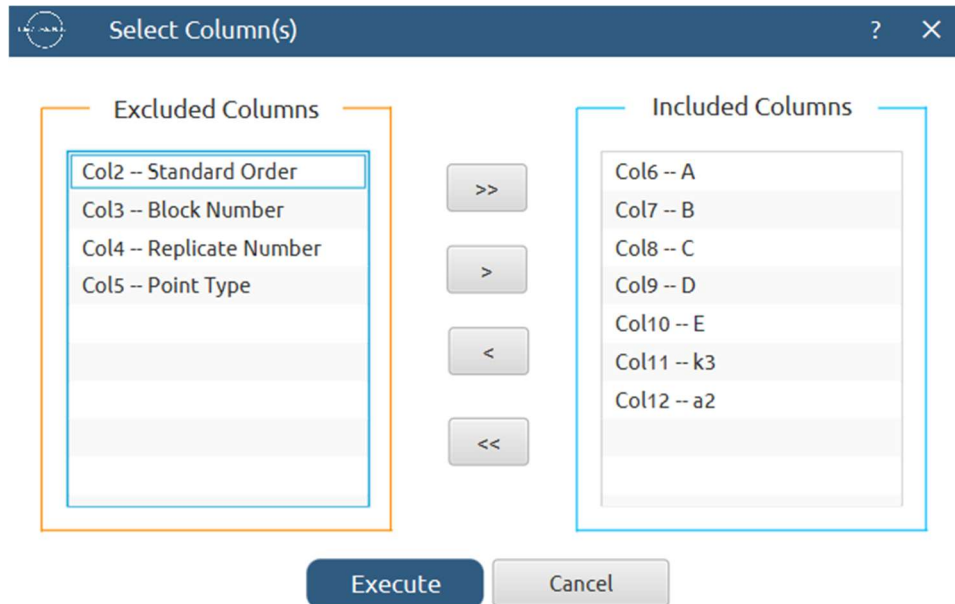
Col2 -- k3
Col3 -- a2

Execute

Cancel

Step 11: Data selection

Create a new tab named “Data - FF” and import the results from the “Fractional Factorial” and “Responses - FF” 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 (D)	Col7 (D)	Col8 (D)
User Header	User Row ID	A	B	C	D	E	k3	a2
1		90.0	45.0	5.1	1.1	35.0	3.32	1.55
2		91.0	45.0	5.1	0.9	25.0	6.18	1.51
3		90.0	55.0	5.1	0.9	35.0	4.4	1.49
4		91.0	55.0	5.1	1.1	25.0	4.97	1.54
5		90.0	45.0	5.5	1.1	25.0	3.6	1.53
6		91.0	45.0	5.5	0.9	35.0	5.73	1.55
7		90.0	55.0	5.5	0.9	25.0	4.5	1.47
8		91.0	55.0	5.5	1.1	35.0	5.17	1.59
9		90.5	50.0	5.3	1.0	30.0	4.52	1.49
10		90.5	50.0	5.3	1.0	30.0	4.6	1.49
11		90.5	50.0	5.3	1.0	30.0	4.29	1.5

Step 12: Normalization

Create a new tab named “Normalized data - FF” and import the results from the “Data - FF” spreadsheet. Afterwards, normalize the factor columns to take values in the range [-1, 1]: *Data Transformation* → *Normalizers* → *Min-Max*

Min-Max normalizer

Excluded Columns

Col7 -- k3

Col8 -- a2

>>

>

<

<<

Included Columns

Col2 -- A

Col3 -- B

Col4 -- C

Col5 -- D

Col6 -- E

Min

-1.0

Max

1.0

Execute

Cancel

Results:

	Col1	Col2 (D)	Col3 (D)	Col4 (D)	Col5 (D)	Col6 (D)	Col7 (D)	Col8 (D)
User Header	User Row ID	A	B	C	D	E	k3	a2
1		-1.0	-1.0	-1.0	1.0	1.0	3.32	1.55
2		1.0	-1.0	-1.0	-1.0	-1.0	6.18	1.51
3		-1.0	1.0	-1.0	-1.0	1.0	4.4	1.49
4		1.0	1.0	-1.0	1.0	-1.0	4.97	1.54
5		-1.0	-1.0	1.0	1.0	-1.0	3.6	1.53
6		1.0	-1.0	1.0	-1.0	1.0	5.73	1.55
7		-1.0	1.0	1.0	-1.0	-1.0	4.5	1.47
8		1.0	1.0	1.0	1.0	1.0	5.17	1.59
9		0.0	0.0	0.0	0E-7	0.0	4.52	1.49
10		0.0	0.0	0.0	0E-7	0.0	4.6	1.49
11		0.0	0.0	0.0	0E-7	0.0	4.29	1.5

Step 13: Regression

The goal here is to produce a regression equation that includes main effects for k_3 :

$$Y = b_0 + b_1A + b_2B + b_3C + b_4D + b_5E$$

Create a new tab named “Regression – k_3 - FF” and import the results from the spreadsheet “Normalized data - FF”. Afterwards, fit a generalized linear model to the data: *Analytics* → *Regression* → *Statistical fitting* → *Generalized Linear Models*

Generalized Linear Models Regression

Type: Linear

Confidence Level: 95

Scale Parameter Method: Fixed value

Value: 1.0

Dependent Variable: Col7 -- k3

Excluded Columns: Col8 -- a2

Factors:

Covariates: Col2 -- A, Col3 -- B, Col4 -- C, Col5 -- D

☒ Custom ☐ Include All Main Effects ☐ Full Factorial

Formula: A+B+C+D+E

Execute Cancel

Results:

k3	Prediction
3.32	3.2930682
6.18	5.9455682
4.4	4.2830682
4.97	5.0605682
3.6	3.4830682
5.73	5.8205682
4.5	4.4730682
5.17	4.9355682
4.52	4.6618182
4.6	4.6618182
4.29	4.6618182

Goodness of Fit	
	Value
Deviance	0.3173011
Scaled Deviance	0.3173011
Pearson Chi-Square	0.3173011
Scaled Pearson Chi-Square	0.3173011
Log Likelihood	-10.2669744
Akaike's Information Criterion (AIC)	32.5339489
Finite Sample Corrected AIC (AICC)	53.5339489
Bayesian Information Criterion (BIC)	34.9213205
Consistent AIC (CAIC)	40.9213205

Parameter Estimates							
Variable	Coefficient	Std. Error	Lower CI	Upper CI	Test Statistic	df	p-value
intercept	4.6618182	0.3015113	4.0708668	5.2527696	239.0580364	1	0.0
A	0.7787500	0.3535534	0.0857981	1.4717019	4.8516125	1	0.0276203
B	0.0262500	0.3535534	-0.6667019	0.7192019	0.0055125	1	0.9408145
C	0.0162500	0.3535534	-0.6767019	0.7092019	0.0021125	1	0.9633406
D	-0.4687500	0.3535534	-1.1617019	0.2242019	1.7578125	1	0.1848976
E	-0.0787500	0.3535534	-0.7717019	0.6142019	0.0496125	1	0.8237389

Repeat this step for the second response variable. Results, a_2 :

a2	Prediction
1.55	1.5403409
1.51	1.4978409
1.49	1.4803409
1.54	1.5328409
1.53	1.5203409
1.55	1.5428409
1.47	1.4603409
1.59	1.5778409
1.49	1.5190909
1.49	1.5190909
1.5	1.5190909

Goodness of Fit	
	Value
Deviance	0.0028284
Scaled Deviance	0.0028284
Pearson Chi-Square	0.0028284
Scaled Pearson Chi-Square	0.0028284
Log Likelihood	-10.1097381
Akaike's Information Criterion (AIC)	32.2194761
Finite Sample Corrected AIC (AICC)	53.2194761
Bayesian Information Criterion (BIC)	34.6068478
Consistent AIC (CAIC)	40.6068478

Parameter Estimates							
Variable	Coefficient	Std. Error	Lower CI	Upper CI	Test Statistic	df	p-value
intercept	1.5190909	0.3015113	0.9281395	2.1100423	25.3840091	1	5E-7
A	0.0187500	0.3535534	-0.6742019	0.7117019	0.0028125	1	0.9577056
B	-0.0062500	0.3535534	-0.6992019	0.6867019	0.0003125	1	0.9858960
C	0.0062500	0.3535534	-0.6867019	0.6992019	0.0003125	1	0.9858960
D	0.0237500	0.3535534	-0.6692019	0.7167019	0.0045125	1	0.9464423
E	0.0162500	0.3535534	-0.6767019	0.7092019	0.0021125	1	0.9633406

Step 14: Data selection for Regression Metrics

Create a tab named “Metrics Data – k3-FF” and import the results from the spreadsheet “Regression – k3-FF”. Next, assign names to all unnamed columns containing data, using names such as Z1–Z9. Then, select only the response column (k3) and the “Prediction” column to be transferred to the right spreadsheet: *Data Transformation* → *Data Manipulation* → *Select Column(s)*

	Col1	Col2 (D)	Col3 (D)	Col4	Col5 (S)	Col6 (S)	Col7 (S)	Col8 (S)	Col9 (S)	Col10 (S)	Col11 (S)	Col12 (S)
User Header	User Row ID	k3	Prediction	Z1	Z2	Z3	Z4	Z5	Z6	Z7	Z8	Z9
1		3.32	3.2930682		Goodness of Fit							
2		6.18	5.9455682			Value						
3		4.4	4.2830682		Deviance	0.3173011						
4		4.97	5.0605682		Scaled Deviance	0.3173011						
5		3.6	3.4830682		Pearson Chi-Square	0.3173011						
6		5.73	5.8205682		Scaled Pearson Chi-Square	0.3173011						
7		4.5	4.4730682		Log Likelihood	-10.2669744						
8		5.17	4.9355682		Akaike's Information Criterion (AIC)	32.5339489						
9		4.52	4.6618182		Finite Sample Corrected AIC (AICC)	53.5339489						
10		4.6	4.6618182		Bayesian Information Criterion (BIC)	34.9213205						
11		4.29	4.6618182		Consistent AIC (CAIC)	40.9213205						
12												
13												
14					Parameter Estimates							
15					Variable	Coefficient	Std. Error	Lower CI	Upper CI	Test Statistic	df	p-value
16					intercept	4.6618182	0.3015113	4.0708668	5.2527696	239.0580364	1	0.0

Results:

Select Column(s)

Select Columns Manually

Select Matching Columns from Node

Select Columns Manually and from Node

Select Node

Select Node

Excluded Columns

Col5 -- Z2

Col6 -- Z3

Col7 -- Z4

Col8 -- Z5

Col9 -- Z6

Col10 -- Z7

Col11 -- Z8

Col12 -- Z9

>>

>

<

<<

Included Columns

Col2 -- k3

Col3 -- Prediction

Execute

Cancel

User Header	Col1	Col2 (D)	Col3 (D)
1	User Row ID	k3	Prediction
2		3.32	3.2930682
3		6.18	5.9455682
4		4.4	4.2830682
5		4.97	5.0605682
6		3.6	3.4830682
7		5.73	5.8205682
8		4.5	4.4730682
9		5.17	4.9355682
10		4.52	4.6618182
11		4.6	4.6618182
12		4.29	4.6618182

Repeat this step for a2 as well.

21

Step 15: Regression Metrics

Create a tab named “Metrics – k_3 - FF” and import the results from the spreadsheet “Metrics Data – k_3 - FF”. Then, produce the regression metrics for the k_3 regression equation: *Statistics* → *Model Metrics* → *Regression Metrics*


Regression Statistics Metrics
?
×

Actual Value Column

Col2 – k_3

Prediction Value Column

Col3 – Prediction

Execute

Cancel

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		0.0288456	0.1698398	0.1375620	0.9545672

Repeat this step for the second response variable. Results, a_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		0.0002571	0.0160352	0.0140496	0.7805889

Step 14: Analysis of Covariance

Create a new tab named “ANCOVA – k3 - FF” and import the results from the spreadsheet “Normalized data - FF”. Afterwards perform analysis of covariance for k₃: *Statistics* → *Analysis of (Co)Variance* → *ANCOVA*

ANCOVA

Confidence Level (%) 95

Dependent Variable Col7 -- k3

Sum of Squares for Tests Adjusted (Type III)

Coding for Factors (-1, 0, +1)

Excluded Columns

Col8 -- a2

Factors

Covariates

Col2 -- A
Col3 -- B
Col4 -- C
Col5 -- D

Custom ☒ Include All Main Effects ☐ Full Factorial

Formula

A+B+C+D+E

Execute Cancel

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		A	1	4.8516125	4.8516125	76.4512311	0.0003242
2		B	1	0.0055125	0.0055125	0.0868654	0.7800455
3		C	1	0.0021125	0.0021125	0.0332886	0.8623953
4		D	1	1.7578125	1.7578125	27.6994359	0.0032906
5		E	1	0.0496125	0.0496125	0.7817889	0.4170786
6		Error	5	0.3173011	0.0634602		
7		Total	10	6.9839636			

Repeat this step for the second response variable. Results, a_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		A	1	0.0028125	0.0028125	4.9718763	0.0761870
2		B	1	0.0003125	0.0003125	0.5524307	0.4907525
3		C	1	0.0003125	0.0003125	0.5524307	0.4907525
4		D	1	0.0045125	0.0045125	7.9770992	0.0369181
5		E	1	0.0021125	0.0021125	3.7344315	0.1111383
6		Error	5	0.0028284	0.0005657		
7		Total	10	0.0128909			

References

(1) Terzić, J.; Popović, I.; Stajić, A.; Tumpa, A.; Jančić-Stojanović, B. Application of Analytical Quality by Design Concept for Bilastine and Its Degradation Impurities Determination by Hydrophilic Interaction Liquid Chromatographic Method. *Journal of Pharmaceutical and Biomedical Analysis* 2016, 125, 385–393. <https://doi.org/10.1016/j.jpba.2016.04.022>.