https://www.docs.isalos.novamechanics.com/

The objective of this practical session is to investigate how the physicochemical properties of per- and polyfluoroalkyl substances (PFAS) influence their environmental fate and potential human exposure. We will apply a structured modelling workflow that integrates environmental fate simulations, statistical analysis, and machine-learning modelling using three computational tools: SimpleBox4Planet (https://www.enaloscloud.novamechanics.com/proplanet/simplebox4planet/) and Isalos Analytics Platform (https://enaloscloud.novamechanics.com/novamechanicssystem/userregistration/).

In addition to evaluating the influence of key physicochemical properties, the practical will aim to identify combinations of property values and levels that minimize environmental mass, thereby supporting the design of safer alternative PFAS within a Safe-and-Sustainable-by-Design (SSbD) framework. Furthermore, a machine-learning surrogate model will be developed to predict environmental mass directly, reducing the need to repeatedly run the underlying mechanistic mass-balance and differential-equation models and thus lowering the computational cost of screening.

Please, follow the steps presented in the following link to download and install Isalos Analytics on your local PCs: https://www.docs.isalos.novamechanics.com/installation.html.

Isalos version used: 2.0.2

Step-by-step guide for implementing in Isalos

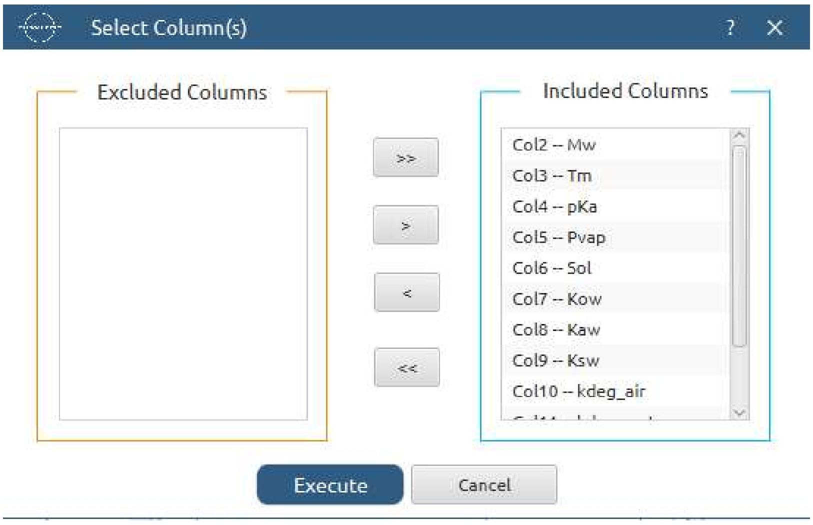
Design of Experiment using Isalos

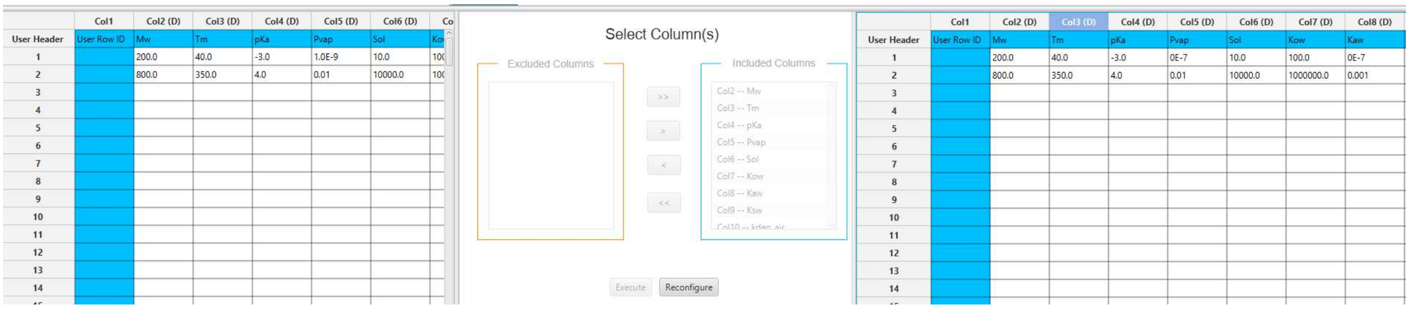
1. **Manually insert data** as shown in the table below into Isalos on the input spreadsheet (left).

Mw	Tm	pKa	Pvap	Sol	Kow	Kaw	Ksw	kdeg_air	kdeg_water	kdeg_sed	kdeg_soil
200.0	40.0	-3.0	1.0E-9	10.0	100.0	1.0E-8	0.1	1.0E-7	1.0E-8	1.0E-11	1.0E-10
800.0	350.0	4.0	0.01	10000.0	1000000.0	0.001	100.0	1.0E-6	1.0E-7	1.0E-9	1.0E-8

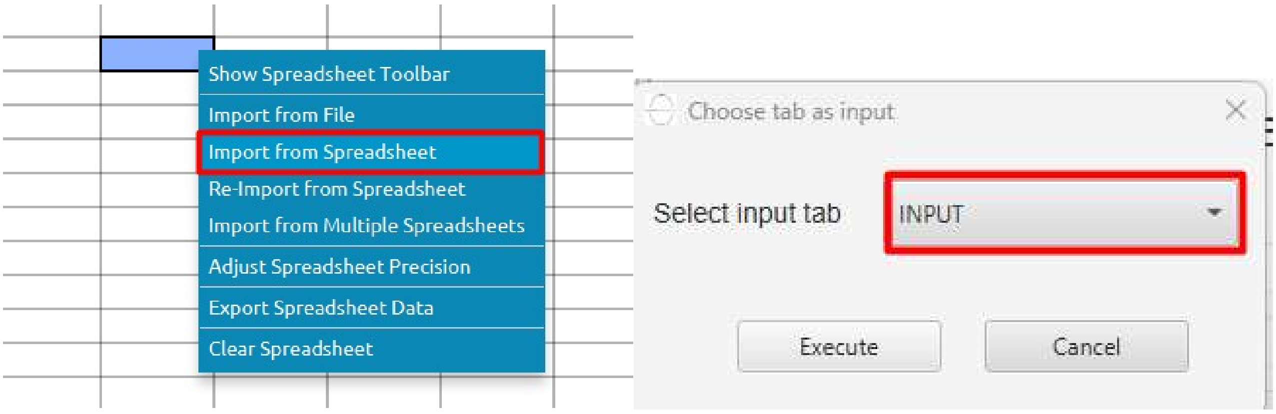
	Col1	Col2 (D)	Col3 (D)	Col4 (D)	Col5 (D)	Col6 (D)	Col7 (D)	Col8 (D)	Col9 (D)	Col10 (D)	Col11 (D)	Col12 (D)	Col13 (D)
User Header	User Row ID	Mw	Tm	pKa	Pvap	Sol	Kow	Kaw	Ksw	kdeg_air	kdeg_water	kdeg_sed	kdeg_soil
1		200.0	40.0	-3.0	1.0E-9	10.0	100.0	1.0E-8	0.1	1.0E-7	1.0E-8	1.0E-11	1.0E-10
2		800.0	350.0	4.0	0.01	10000.0	1000000.0	0.001	100.0	1.0E-6	1.0E-7	1.0E-9	1.0E-8

2. **Rename** the sheet to 'INPUT' by right-clicking on the node and selecting 'Rename'. Select all columns by navigating *Data Transformation > Data Manipulation > Select Column(s)* and click on 'Execute' button.

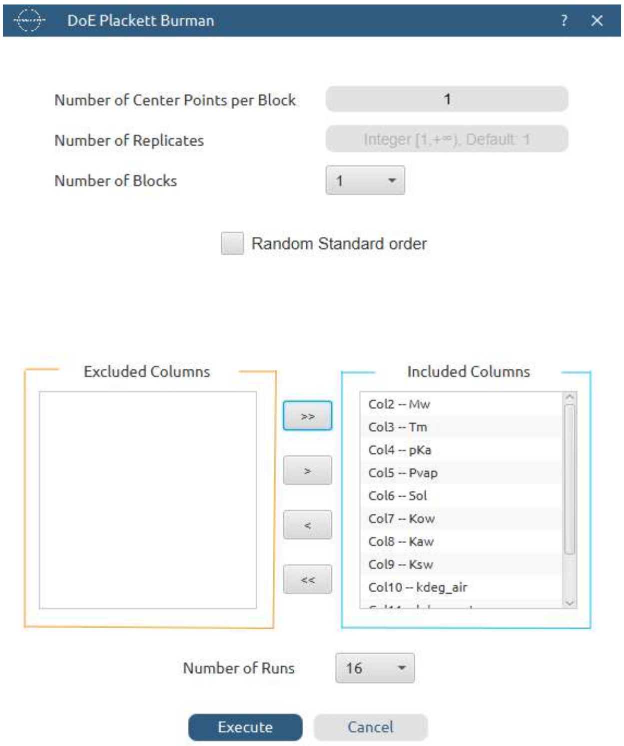




3. **Create a new sheet** and name it as ‘PLACKETT BURMAN’. Right-click on the input (left) spreadsheet and select ‘Import from Spreadsheet’. Then select ‘INPUT’.

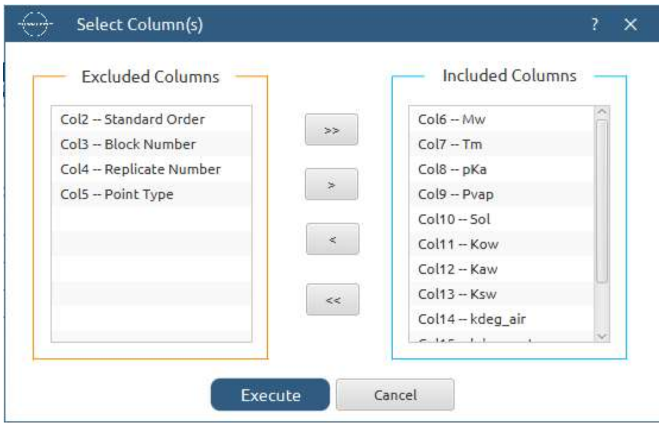


4. **Design Plackett Burman** by selecting from the tabs _DOE > Screening > Plackett Burman_. Once the configuration window appears, make the selections as indicated below and proceed by clicking the ‘Execute’ button.

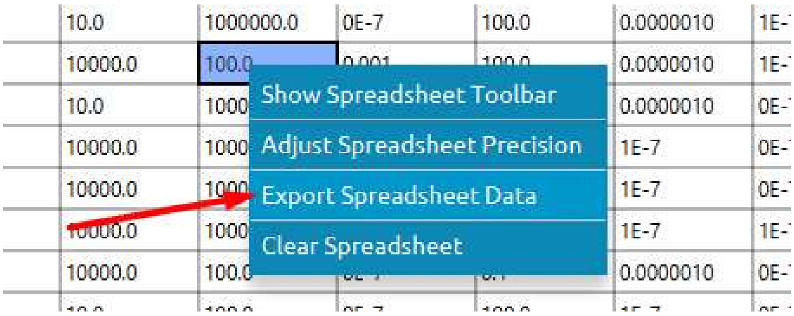


Hint:
The analysis began with a screening design, specifically a Plackett–Burman design, to identify the most influential factors affecting the environmental fate response and to provide an initial indication of whether the system could be described using a linear model. Since the Plackett–Burman design is primarily intended for screening purposes, it can only estimate main effects and does not capture interaction or quadratic terms.

5. **Create a new sheet** and name it as 'PREPROCESS'. Import data from the output of the 'PLACKETT BURMAN' sheet. Then select from tabs *Data Transformation > Data Manipulation > Select Column(s)*. Ensure to replicate the selection of columns as indicated below, and then click the 'Execute' button.

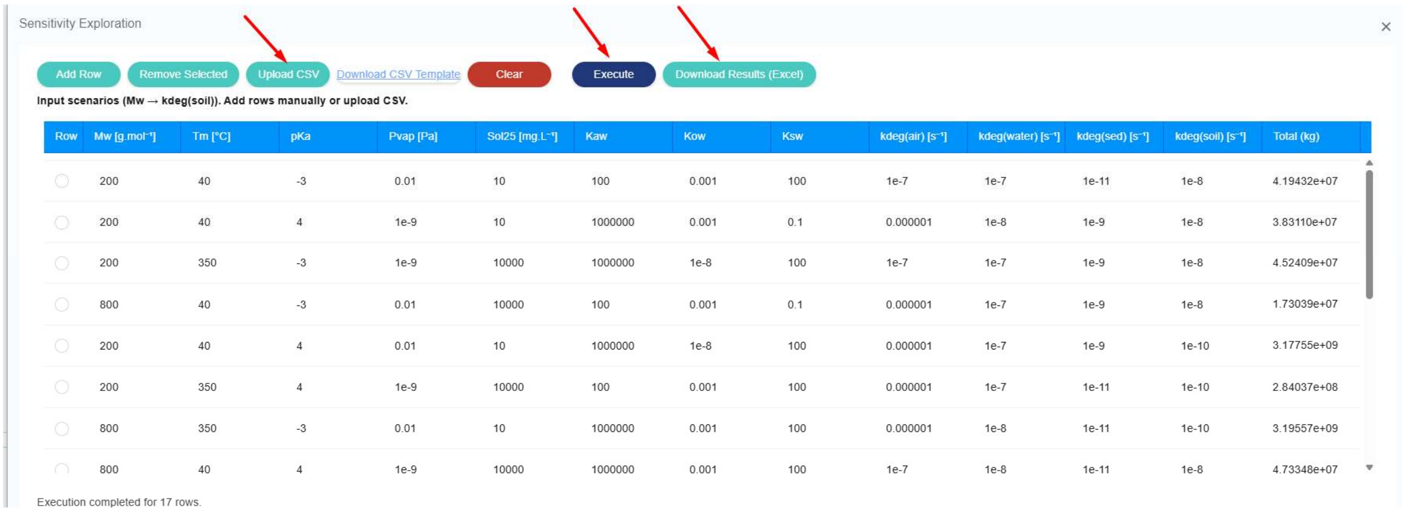


6. **Export the output** to a CSV file by right-clicking on the output spreadsheet (on the right) and selecting 'Export Spreadsheet Data.' Choose 'CSV' as the file extension and click the 'Execute' button. Name the file 'ScreeningDesign' and save it. Subsequently, open the file, delete the first column, and save the file again.



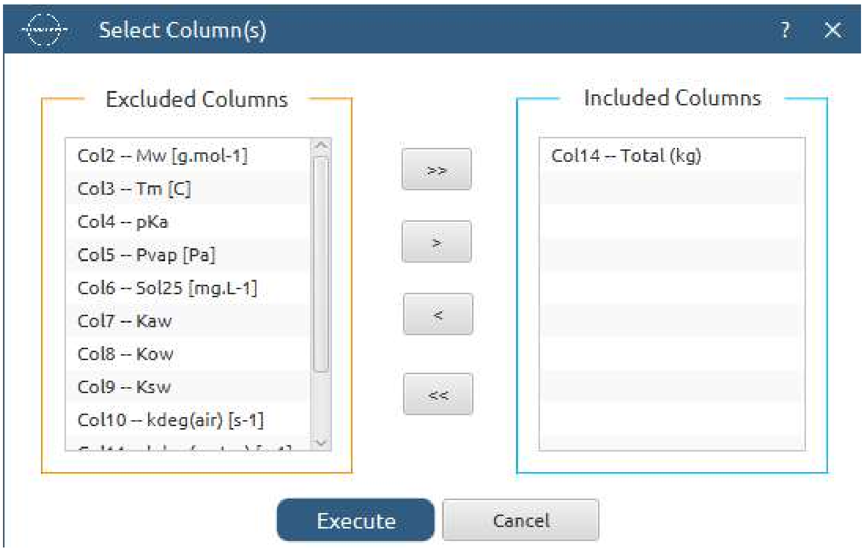
Compute environmental mass using SimpleBox4Planet

7. **Open** the following URL link: https://www.enaloscloud.novamechanics.com/proplanet/simplebox4planet/ and navigate to the bottom of the page. Select 'Sensitivity Exploration' button. When a new window appears, click on 'Upload CSV' button and select the file 'ScreeningDesign.csv' which created in the previous step. After the data is successfully loaded, click on the 'Execute' button to generate the total mass produced (in kg). Finally, click on the 'Download Results (Excel)' button to download the results, naming the file 'ScreeningDesignResults.xlsx'.

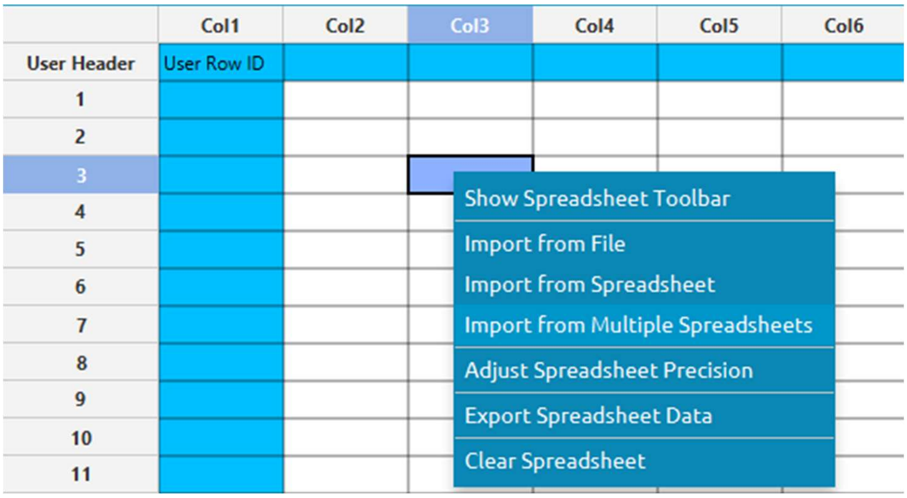


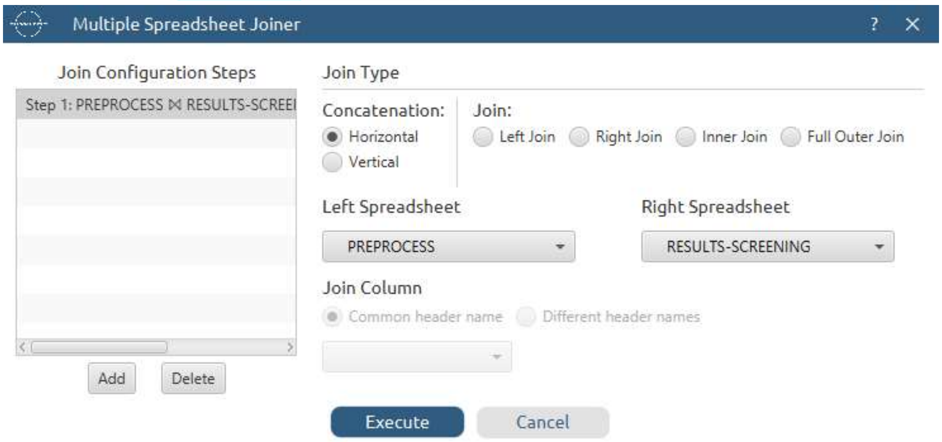
Statistical Analysis using Isalos

8. **Create** a new sheet in Isalos Analytics and name it as 'RESULTS-SCREENING'. Right-click on the input (left) spreadsheet and select 'Import from File' to import the ScreeningDesignResults.xslx created in the previous step. Then navigate to *Data Transformation > Data Manipulation > Select Column(s)* and select only the column corresponding to the generated output, as shown below.

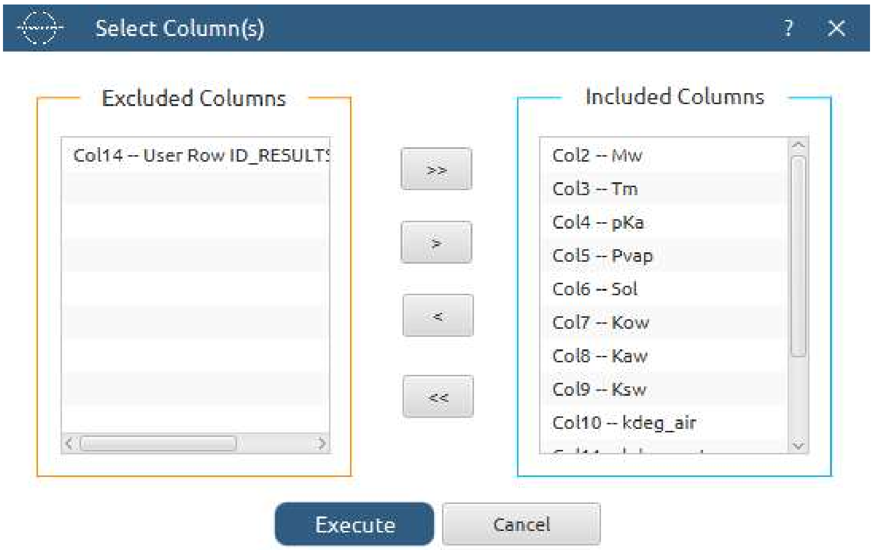


9. **Create** a new sheet and name it as ‘PREPROCESS-v2’. Right-click on the input (left) spreadsheet and select ‘Import from Multiple Spreadsheets’. Select the ‘PREPROCESS’ as the Left Spreadsheet and ‘RESULTS-SCREENING’ as the Right Spreadsheet, then click on ‘Execute’ button.

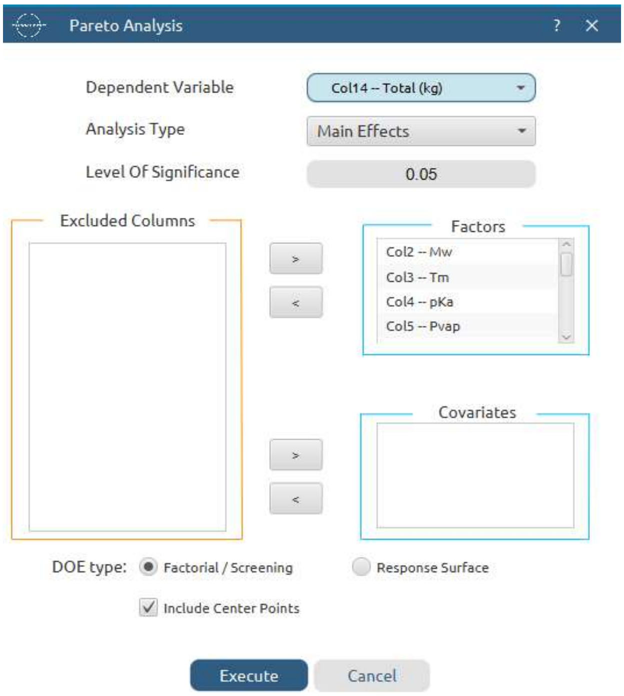




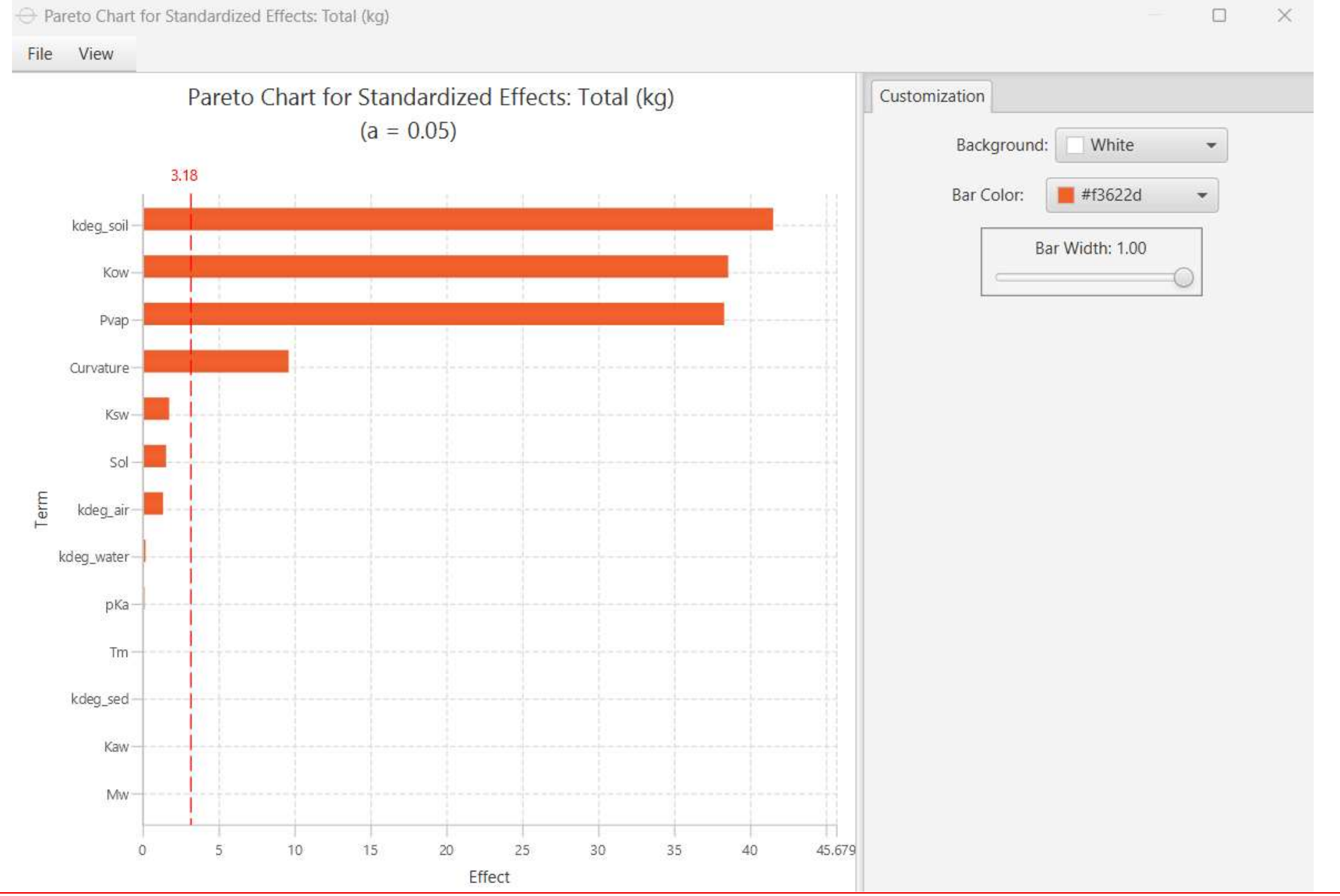
Navigate to *Data Transformation > Data Manipulation > Select Column(s)* and select the columns shown below.



10. **Create** a new sheet and name it as ‘PARETO’. Right-click on the input (left) spreadsheet and select ‘Import from Spreadsheet’ and select the ‘PREPROCESS-v2’. Navigate to *DOE > Post DoE Analysis > Pareto Analysis* and select the configuration as shown below.

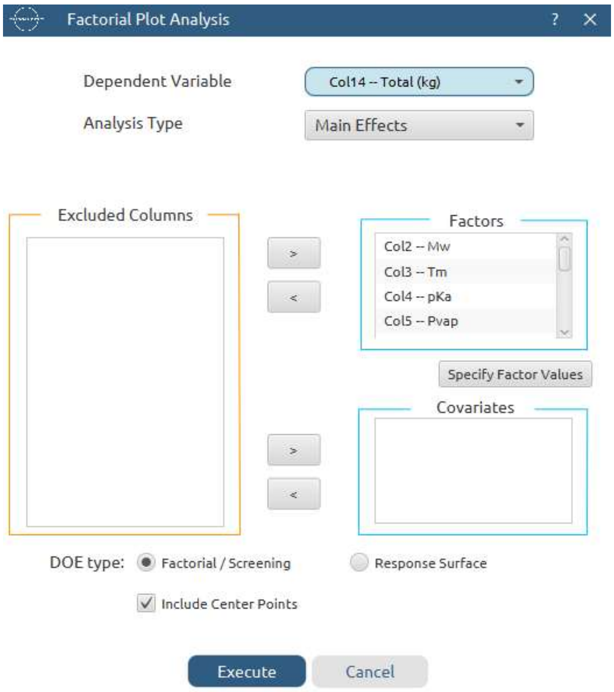


Output:

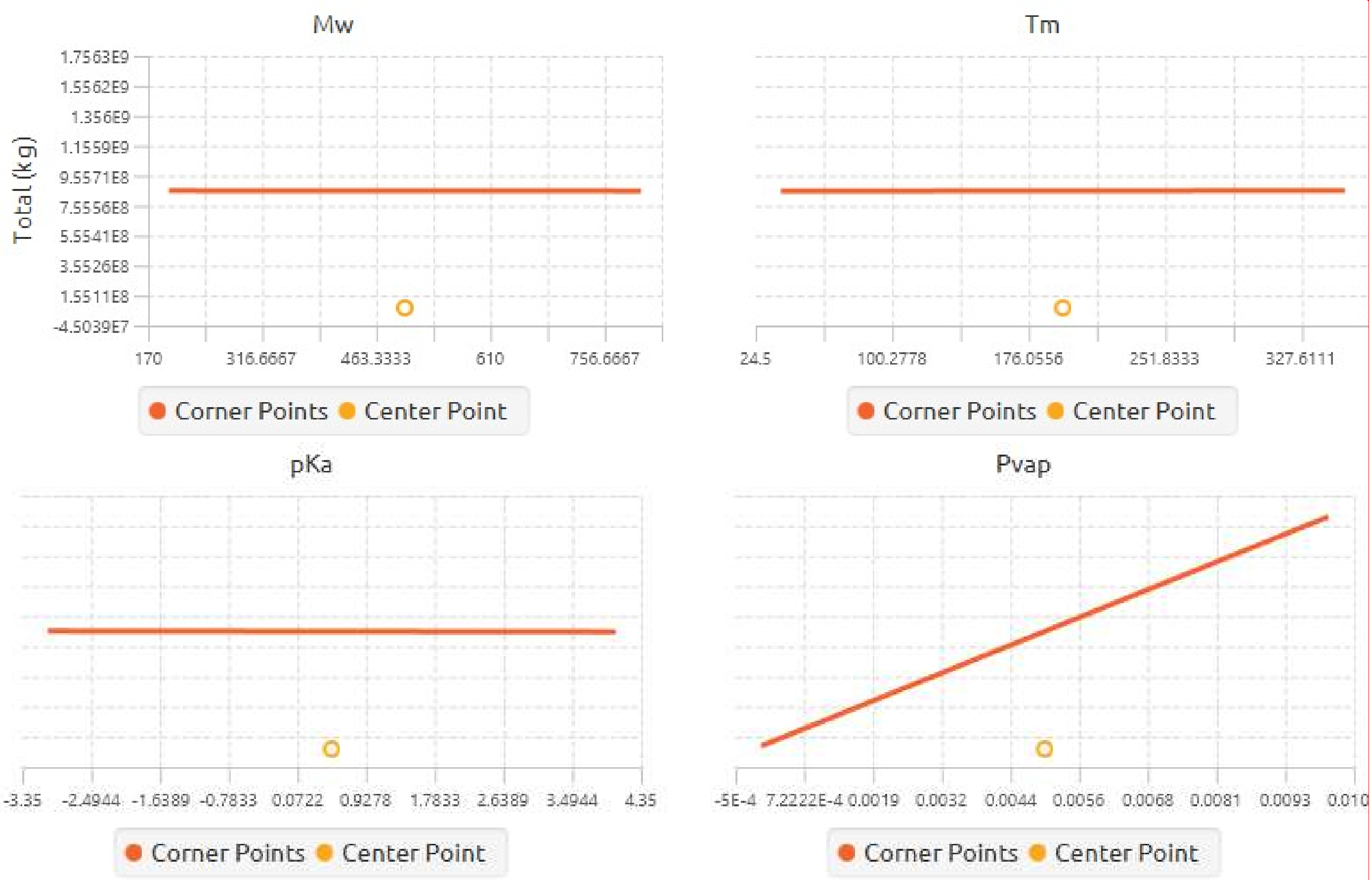


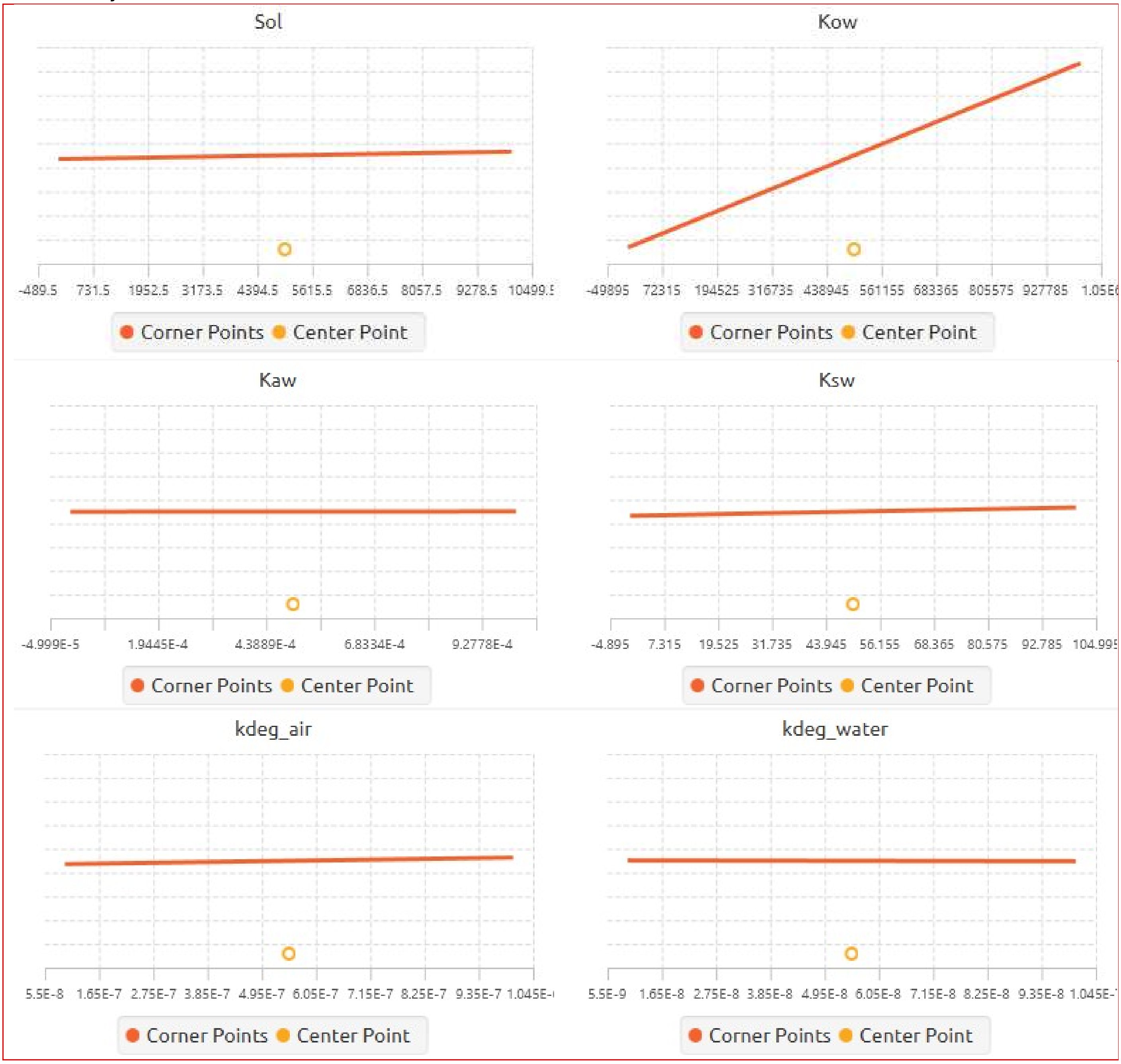
Observation:
The analysis indicates that kdeg(soil), Kow and Pvap effects exceed the statistical significance threshold and are therefore identified as the dominant factors influencing the environmental fate response. All remaining variables fall below the threshold and are considered non-significant within the explored design space and thus are excluded from further analysis.

11. **Create** a new sheet and name it as 'FACTORIAL-PLOTS'. Right-click on the input (left) spreadsheet and select 'Import from Spreadsheet' and select the 'PREPROCESS-v2'. Navigate to *DOE > Post DoE Analysis > Factorial Plot Analysis* and select the configuration as shown below.


Factorial Plot Analysis
Dependent Variable
Col14 -- Total (kg)
Analysis Type
Main Effects
Excluded Columns
Factors
Col2 -- Mw
Col3 -- Tm
Col4 -- pKa
Col5 -- Pvap
Specify Factor Values
Covariates
DOE type:
Factorial / Screening
Response Surface
Include Center Points
Execute
Cancel

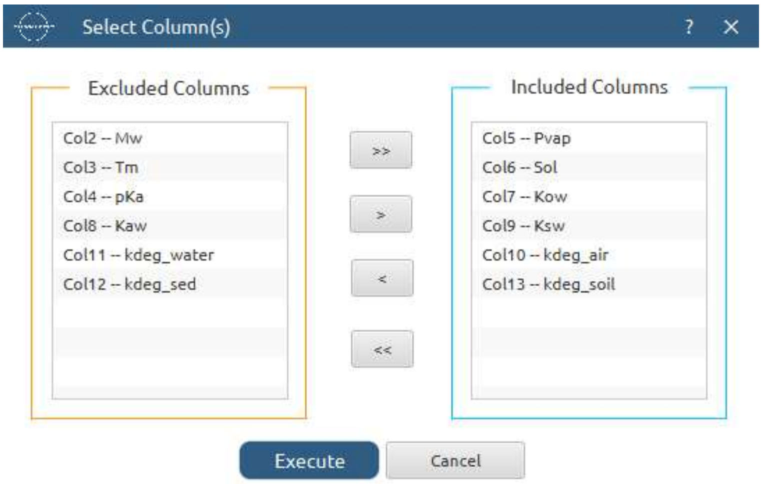
Output:


Mw
Tm
pKa
Pvap
Total (kg)
1.7563E9
1.5562E9
1.356E9
1.1559E9
9.5571E8
7.5556E8
5.5541E8
3.5526E8
1.5511E8
-4.5039E7
170
316.6667
463.3333
610
756.6667
24.5
100.2778
176.0556
251.8333
327.6111
-3.35
-2.4944
-1.6389
-0.7833
0.0722
0.9278
1.7833
2.6389
3.4944
4.35
-5E-4
7.2222E-4
0.0019
0.0032
0.0044
0.0056
0.0068
0.0081
0.0093
Corner Points
Center Point


Sol
-489.5 731.5 1952.5 3173.5 4394.5 5615.5 6836.5 8057.5 9278.5
Corner Points Center Point
Kow
-49895 72315 194525 316735 438945 561155 683365 805575 927785
Corner Points Center Point
Kaw
-4.999E-5 1.9445E-4 4.3889E-4 6.8334E-4 9.2778E-4
Corner Points Center Point
Ksw
-4.895 7.315 19.525 31.735 43.945 56.155 68.365 80.575 92.785
Corner Points Center Point
kdeg_air
5.5E-8 1.65E-7 2.75E-7 3.85E-7 4.95E-7 6.05E-7 7.15E-7 8.25E-7 9.35E-7
Corner Points Center Point
kdeg_water
5.5E-9 1.65E-8 2.75E-8 3.85E-8 4.95E-8 6.05E-8 7.15E-8 8.25E-8 9.35E-8
Corner Points Center Point

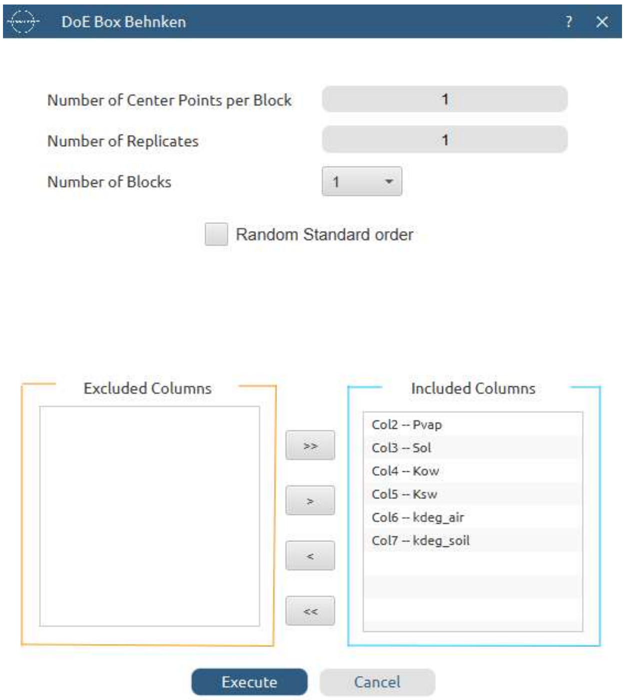
Observation:
The factorial plot analysis indicates that a linear model is not sufficient to describe the system behavior. For all examined factors, the center points do not lie on the straight line connecting the low and high factor levels, indicating the presence of curvature in the response. This deviation suggests that the relationship between the physicochemical properties and the environmental fate response is non-linear, thereby supporting the need for an RSM design to adequately capture these curvature effects.

12. **Create** a new sheet and name it as 'PREPROCESS-v3'. Import data from the 'INPUT' sheet. Navigate to *Data Transformation > Data Manipulation > Select Column(s)* and select columns according to the specified selections below.



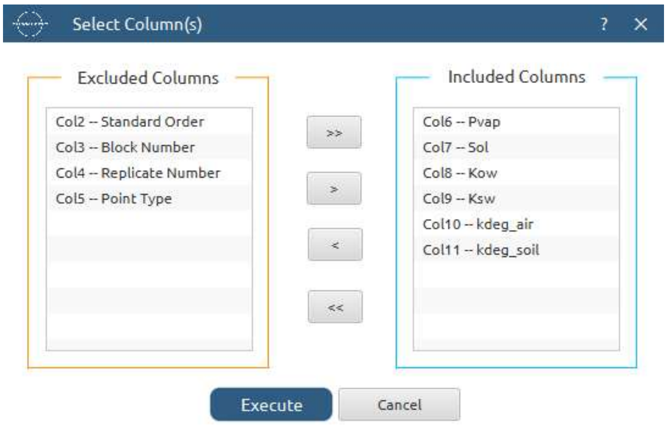
Hint:
Since the significant factors had already been identified (step 10) during the screening stage, the least influential variables were excluded from the subsequent RSM analysis in order to reduce model complexity and focus on the most relevant physicochemical drivers. Accordingly, kdeg(soil), Kow, and Pvap were retained as clearly significant factors. In addition, Ksw, Sol, and kdeg(air) were also included in the RSM study, although they were slightly below the significance threshold, because their effects were close enough to the cutoff to suggest potential relevance in the presence of curvature or interaction effects.

13. **Create** a new sheet and name it as 'BOX BEHNKEN'. Import data from the 'PREPROCESS-v3' sheet. Navigate to *DOE > Response Surface > Box Behnken*. Follow the specified selections outlined below and click on 'Execute'.



Hint:
The Box–Behnken design was selected because it efficiently estimates quadratic models while requiring fewer experimental runs compared to full factorial designs, and avoids extreme corner points, making it suitable for stable environmental modelling studies.

14. **Create** a new sheet and name it as 'PREPROCESS-v4'. Import data from the 'BOX BEHNKEN' sheet. Navigate to *Data Transformation > Data Manipulation > Select Column(s)* and select columns according to the specified selections below.

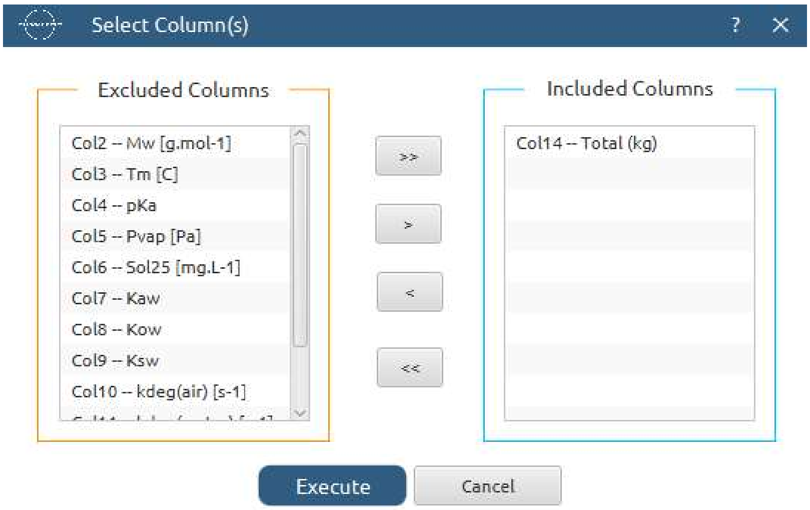


15. **Create** a new sheet and name it as 'RESULTS-RSM'. Copy and paste the table below on the input (left) spreadsheet.

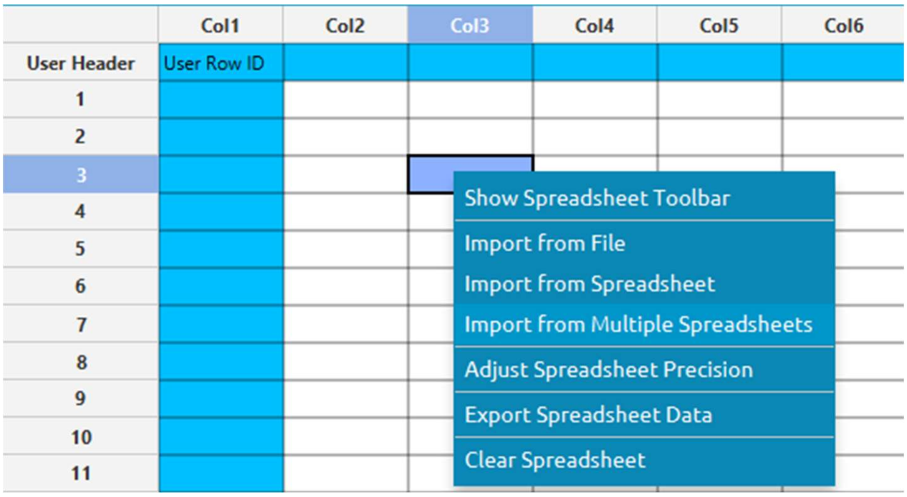
Mw	Tm	pKa	Pvap	Sol	Kow	Kaw	Ksw	kdeg_air	kdeg_water	kdeg_sed	kdeg_soil	Total (kg)
500	195	0.5	1E-09	10	500050	0.0005	0.1	5.5E-07	5.5E-08	5.05E-10	5.05E-09	7.37001e+07
500	195	0.5	0.01	10	500050	0.0005	0.1	5.5E-07	5.5E-08	5.05E-10	5.05E-09	7.37655e+07
500	195	0.5	1E-09	10000	500050	0.0005	0.1	5.5E-07	5.5E-08	5.05E-10	5.05E-09	7.37001e+07

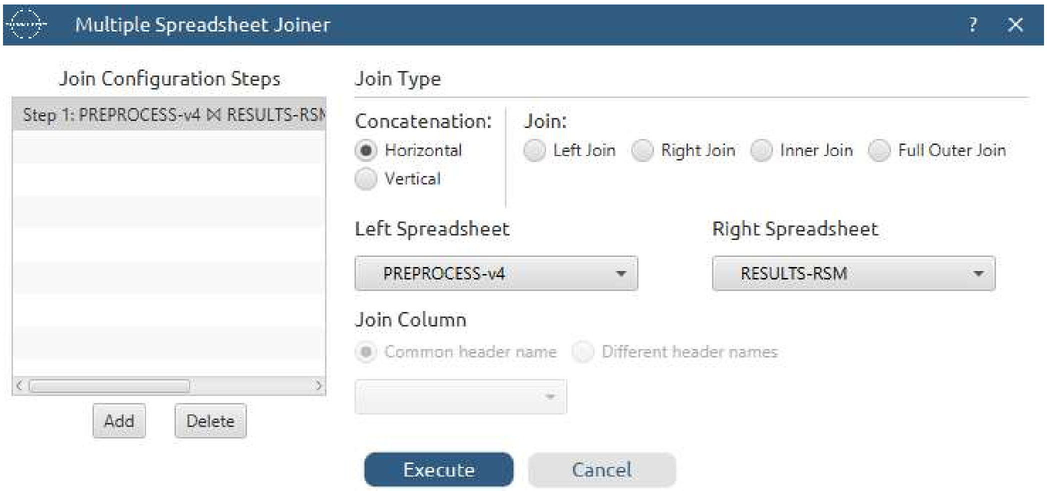
500	195	0.5	0.01	10000	500050	0.0005	0.1	5.5E-07	5.5E-08	5.05E-10	5.05E-09	7.37655e+07
500	195	0.5	1E-09	10	500050	0.0005	100	5.5E-07	5.5E-08	5.05E-10	5.05E-09	7.38901e+07
500	195	0.5	0.01	10	500050	0.0005	100	5.5E-07	5.5E-08	5.05E-10	5.05E-09	7.39547e+07
500	195	0.5	1E-09	10000	500050	0.0005	100	5.5E-07	5.5E-08	5.05E-10	5.05E-09	7.38901e+07
500	195	0.5	0.01	10000	500050	0.0005	100	5.5E-07	5.5E-08	5.05E-10	5.05E-09	7.39547e+07
500	195	0.5	0.005	10	100	0.0005	50.05	1E-07	5.5E-08	5.05E-10	5.05E-09	6.12784e+07
500	195	0.5	0.005	10000	100	0.0005	50.05	1E-07	5.5E-08	5.05E-10	5.05E-09	6.12784e+07
500	195	0.5	0.005	10	1000000	0.0005	50.05	1E-07	5.5E-08	5.05E-10	5.05E-09	8.08496e+07
500	195	0.5	0.005	10000	1000000	0.0005	50.05	1E-07	5.5E-08	5.05E-10	5.05E-09	8.08496e+07
500	195	0.5	0.005	10	100	0.0005	50.05	0.000001	5.5E-08	5.05E-10	5.05E-09	5.31330e+07
500	195	0.5	0.005	10000	100	0.0005	50.05	0.000001	5.5E-08	5.05E-10	5.05E-09	5.31330e+07
500	195	0.5	0.005	10	1000000	0.0005	50.05	0.000001	5.5E-08	5.05E-10	5.05E-09	7.31556e+07
500	195	0.5	0.005	10000	1000000	0.0005	50.05	0.000001	5.5E-08	5.05E-10	5.05E-09	7.31556e+07
500	195	0.5	0.005	5005	100	0.0005	0.1	5.5E-07	5.5E-08	5.05E-10	1E-10	3.53034e+07
500	195	0.5	0.005	5005	1000000	0.0005	0.1	5.5E-07	5.5E-08	5.05E-10	1E-10	3.18569e+09
500	195	0.5	0.005	5005	100	0.0005	100	5.5E-07	5.5E-08	5.05E-10	1E-10	3.00012e+08
500	195	0.5	0.005	5005	1000000	0.0005	100	5.5E-07	5.5E-08	5.05E-10	1E-10	3.18593e+09
500	195	0.5	0.005	5005	100	0.0005	0.1	5.5E-07	5.5E-08	5.05E-10	1E-08	1.99702e+07
500	195	0.5	0.005	5005	1000000	0.0005	0.1	5.5E-07	5.5E-08	5.05E-10	1E-08	3.84344e+07
500	195	0.5	0.005	5005	100	0.0005	100	5.5E-07	5.5E-08	5.05E-10	1E-08	3.51830e+07
500	195	0.5	0.005	5005	1000000	0.0005	100	5.5E-07	5.5E-08	5.05E-10	1E-08	3.86233e+07
500	195	0.5	1E-09	5005	500050	0.0005	0.1	1E-07	5.5E-08	5.05E-10	5.05E-09	8.06886e+07
500	195	0.5	0.01	5005	500050	0.0005	0.1	1E-07	5.5E-08	5.05E-10	5.05E-09	8.07525e+07
500	195	0.5	1E-09	5005	500050	0.0005	100	1E-07	5.5E-08	5.05E-10	5.05E-09	8.08734e+07
500	195	0.5	0.01	5005	500050	0.0005	100	1E-07	5.5E-08	5.05E-10	5.05E-09	8.09367e+07
500	195	0.5	1E-09	5005	500050	0.0005	0.1	0.000001	5.5E-08	5.05E-10	5.05E-09	7.29903e+07
500	195	0.5	0.01	5005	500050	0.0005	0.1	0.000001	5.5E-08	5.05E-10	5.05E-09	7.30557e+07
500	195	0.5	1E-09	5005	500050	0.0005	100	0.000001	5.5E-08	5.05E-10	5.05E-09	7.31807e+07
500	195	0.5	0.01	5005	500050	0.0005	100	0.000001	5.5E-08	5.05E-10	5.05E-09	7.32455e+07
500	195	0.5	0.005	10	500050	0.0005	50.05	1E-07	5.5E-08	5.05E-10	1E-10	3.16038e+09
500	195	0.5	0.005	10000	500050	0.0005	50.05	1E-07	5.5E-08	5.05E-10	1E-10	3.16038e+09
500	195	0.5	0.005	10	500050	0.0005	50.05	0.000001	5.5E-08	5.05E-10	1E-10	3.15254e+09
500	195	0.5	0.005	10000	500050	0.0005	50.05	0.000001	5.5E-08	5.05E-10	1E-10	3.15254e+09
500	195	0.5	0.005	10	500050	0.0005	50.05	1E-07	5.5E-08	5.05E-10	1E-08	4.54982e+07
500	195	0.5	0.005	10000	500050	0.0005	50.05	1E-07	5.5E-08	5.05E-10	1E-08	4.54982e+07
500	195	0.5	0.005	10	500050	0.0005	50.05	0.000001	5.5E-08	5.05E-10	1E-08	3.78057e+07
500	195	0.5	0.005	10000	500050	0.0005	50.05	0.000001	5.5E-08	5.05E-10	1E-08	3.78057e+07
500	195	0.5	1E-09	5005	100	0.0005	50.05	5.5E-07	5.5E-08	5.05E-10	1E-10	1.55376e+08
500	195	0.5	0.01	5005	100	0.0005	50.05	5.5E-07	5.5E-08	5.05E-10	1E-10	1.74744e+08
500	195	0.5	1E-09	5005	1000000	0.0005	50.05	5.5E-07	5.5E-08	5.05E-10	1E-10	3.12269e+09
500	195	0.5	0.01	5005	1000000	0.0005	50.05	5.5E-07	5.5E-08	5.05E-10	1E-10	3.18736e+09
500	195	0.5	1E-09	5005	100	0.0005	50.05	5.5E-07	5.5E-08	5.05E-10	1E-08	3.24302e+07
500	195	0.5	0.01	5005	100	0.0005	50.05	5.5E-07	5.5E-08	5.05E-10	1E-08	3.27370e+07
500	195	0.5	1E-09	5005	1000000	0.0005	50.05	5.5E-07	5.5E-08	5.05E-10	1E-08	3.85112e+07
500	195	0.5	0.01	5005	1000000	0.0005	50.05	5.5E-07	5.5E-08	5.05E-10	1E-08	3.85193e+07
500	195	0.5	0.005	5005	500050	0.0005	50.05	5.5E-07	5.5E-08	5.05E-10	5.05E-09	7.38488e+07

16. **Navigate** to _Data Transformation > Data Manipulation > Select Column(s)_ and select only the column corresponding to the generated output, as shown below.

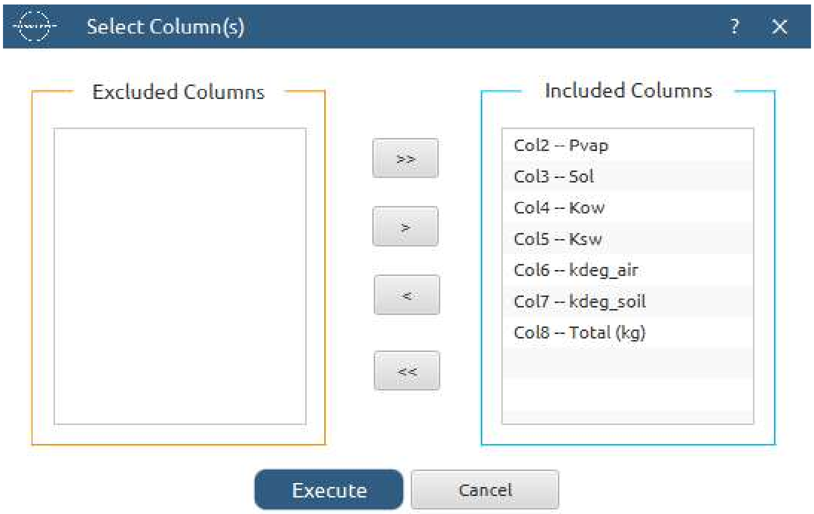


17. **Create** a new sheet and name it as ‘PREPROCESS-v5’. Right-click on the input (left) spreadsheet and select ‘Import from Multiple Spreadsheets’. Select the ‘PREPROCESS-v4’ as the Left Spreadsheet and ‘RESULTS-RSM’ as the Right Spreadsheet, then click on ‘Execute’ button.

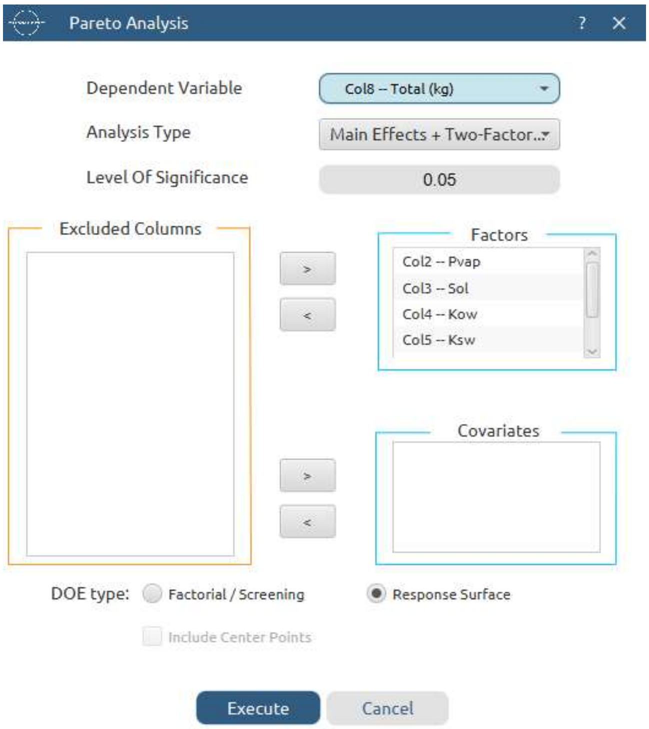




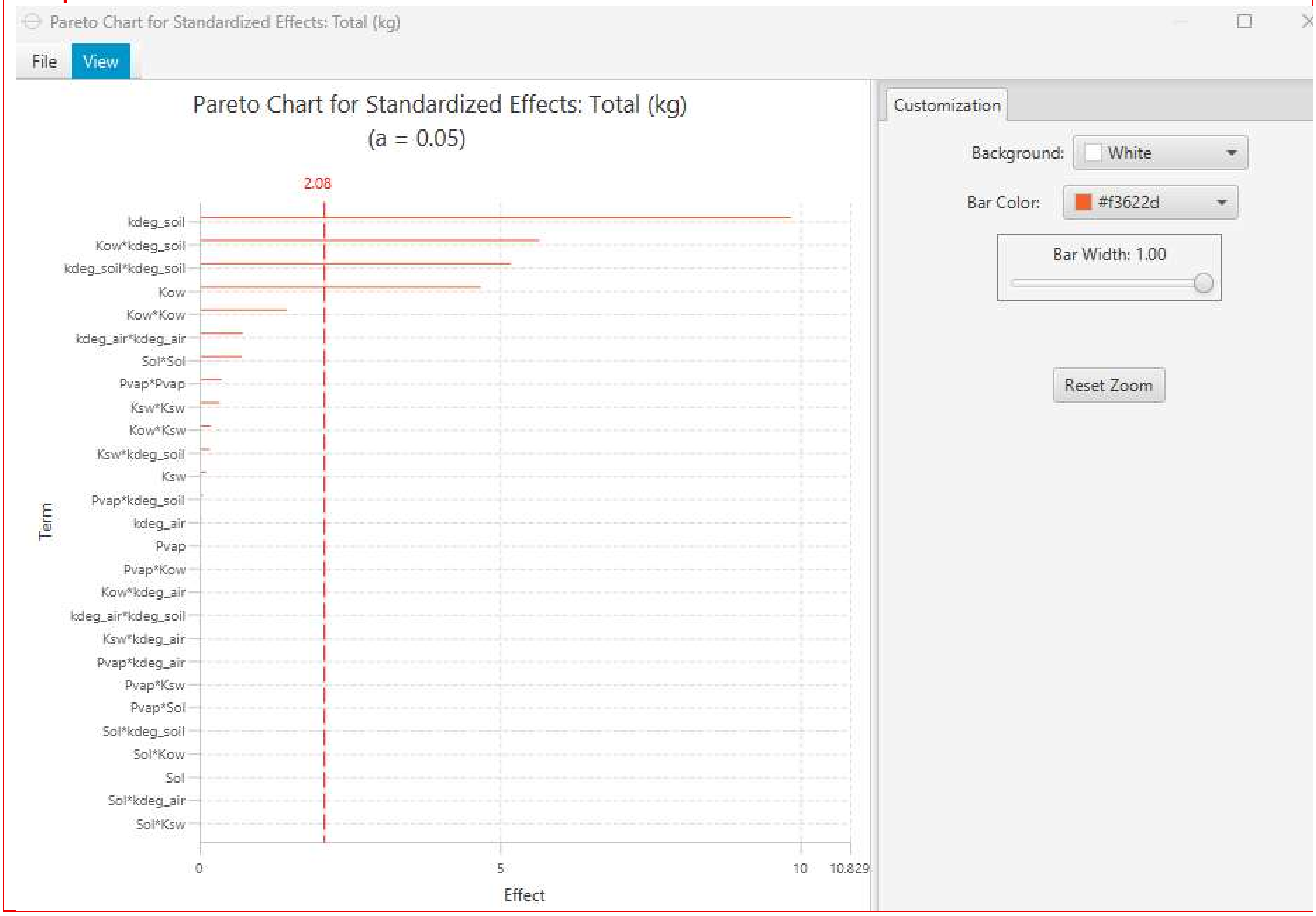
Navigate to Data Transformation > Data Manipulation > Select Column(s) and select the columns shown below.



18. **Create** a new sheet and name it as 'PARETO-RSM'. Right-click on the input (left) spreadsheet and select 'Import from Spreadsheet' and select the 'PREPROCESS-v5'. Navigate to _DOE > Post DoE Analysis > Pareto Analysis_ and select the configuration as shown below.

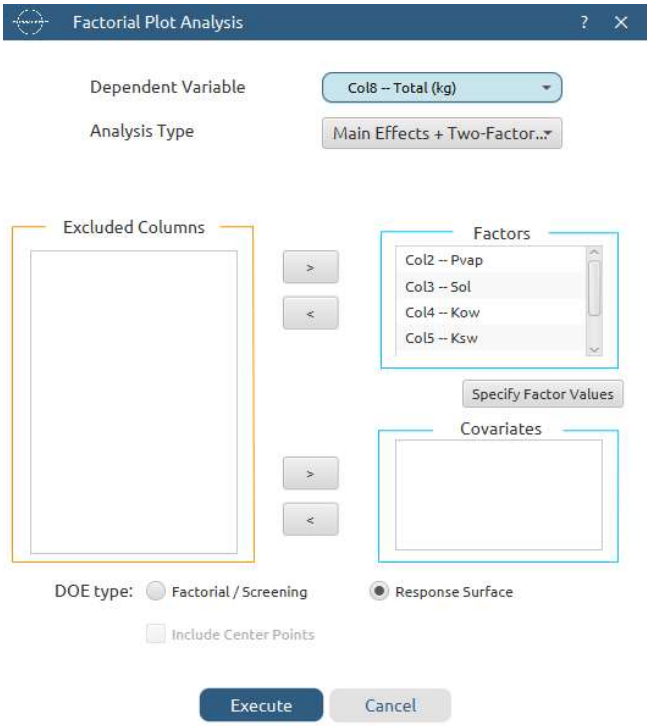


Output:

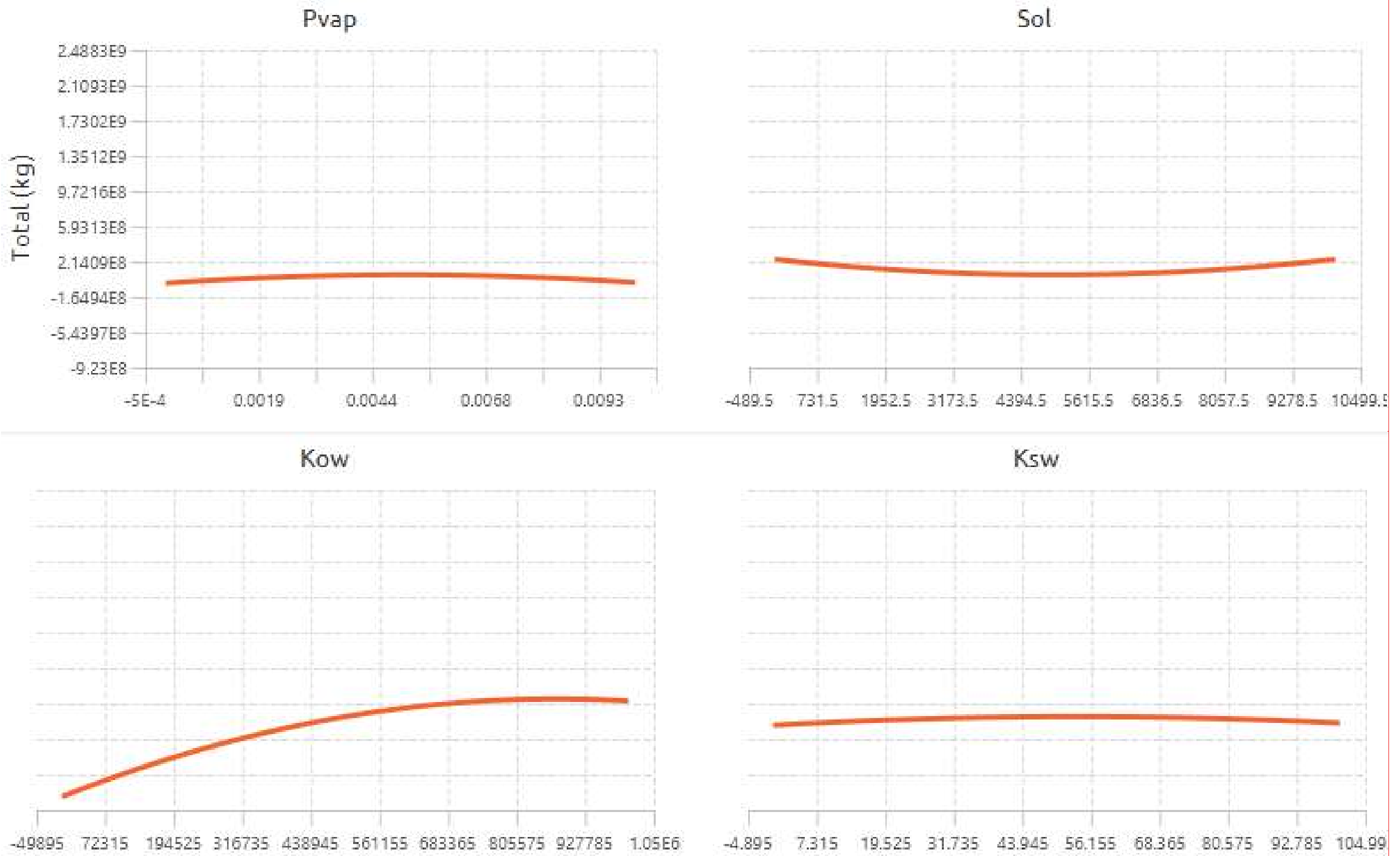


Observation:
The Pareto analysis of the Box Behnken results showed that kdeg(soil), the interaction term Kow × kdeg(soil), the quadratic term kdeg(soil)^2, and Kow were the most important statistically significant contributors to the environmental fate response.

19. **Create** a new sheet and name it as 'FACTORIAL-PLOTS-RSM'. Right-click on the input (left) spreadsheet and select 'Import from Spreadsheet' and select the 'PREPROCESS-v5'. Navigate to *DOE > Post DoE Analysis > Factorial Plot Analysis* and select the configuration as shown below.

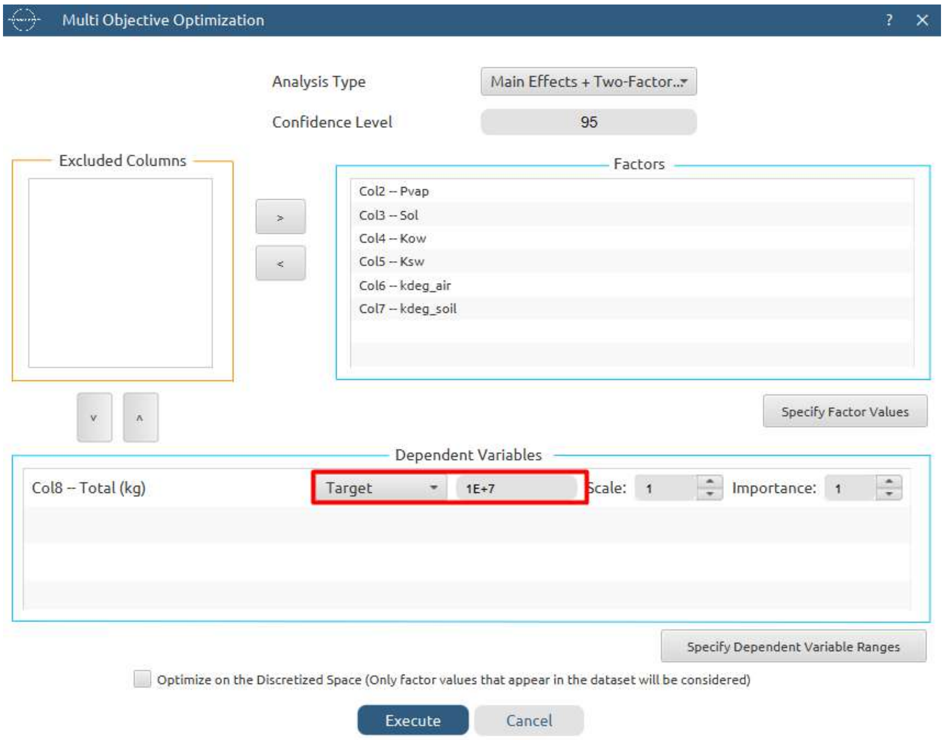

Factorial Plot Analysis
Dependent Variable
Col8 -- Total (kg)
Analysis Type
Main Effects + Two-Factor..
Excluded Columns
Factors
Col2 -- Pvap
Col3 -- Sol
Col4 -- Kow
Col5 -- Ksw
Specify Factor Values
Covariates
DOE type:
Factorial / Screening
Response Surface
Include Center Points
Execute
Cancel

Output:


Pvap
Sol
Kow
Ksw
Total (kg)
2.4883E9
2.1093E9
1.7302E9
1.3512E9
9.7216E8
5.9313E8
2.1409E8
-1.6494E8
-5.4397E8
-9.23E8
-5E-4 0.0019 0.0044 0.0068 0.0093
-489.5 731.5 1952.5 3173.5 4394.5 5615.5 6836.5 8057.5 9278.5
-49895 72315 194525 316735 438945 561155 683365 805575 927785 1.05E6
-4.895 7.315 19.525 31.735 43.945 56.155 68.365 80.575 92.785 104.99

Observation:
The factorial plots obtained from the Box-Behnken design further highlight the presence of curvature effects.

20. **Create** a new sheet and name it as 'OPTIMIZATION'. Right-click on the input (left) spreadsheet and select 'Import from Spreadsheet' and select the 'PREPROCESS-v5'. Navigate to *DOE > Post DoE Analysis > Multi-Objective Optimization* and select the configuration as shown below.



Output:

User Row ID							
	Optimal Settings - Solution						
	Pvap	Sol	Kow	Ksw	kdeg_air	kdeg_soil	
	0.0015009	1957.0324061	579703.6465934	59.6863679	5E-7	0E-7	
	Response Prediction						
	Response	Predicted Value	Standard Error	(95.0%) CI Lower	(95.0%) CI Upper	(95.0%) PI Lower	(95.0%) PI Upper
	Total (kg)	10000000.0000001	351317420.5930987	-720604571.7594274	740604571.7594277	-1311298072.0422013	1331298072.0422013
	Desirabilities						
	Total (kg)	Overall					
	1.0	1.0					

Observation:
Following the Box–Behnken design, a multi-objective optimization was performed with the target of minimizing the total environmental mass to approximately 10,000,000 kg. The optimal solution suggests that low degradation in soil (kdeg(soil) = 1.0×10^{-8} s^{-1}) combined with relatively high Kow and moderate Pvap and Sol values leads to the desired outcome. The predicted total mass is close to the target, with an overall desirability of 1.0, indicating an excellent optimization result. However, the wide confidence and prediction intervals highlight a high level of uncertainty in the prediction, suggesting that while the model captures the trend, variability in the system remains significant and should be considered when interpreting the results.
These optimal conditions can be interpreted within an SSbD framework, as they suggest combinations of physicochemical properties that minimize environmental persistence and accumulation, thereby guiding the design of safer PFAS alternatives.

Final Isalos Workflow

Following the above-described steps, the final workflow on Isalos will look like this: