Altair

HyperWorks

Altair HyperStudy 2019 Tutorials

HS-3000: Fit Method Comparison: Approximation on the Arm Model

HS-3000: Fit Method Comparison: Approximation on the Arm Model

In this tutorial, you will create approximations for the output responses of the arm example introduced in tutorial HS-2000: DOE Method Comparison: Arm Model Study, and review the differences between different Fit methods.

Before running this tutorial, complete tutorial HS-2000: DOE Method Comparison: Arm Model Study. You can also import the archive file `HS-2000.hstx`, available in `<hst.zip>/HS-3000/`.

In HS-2000, you learned that instead of using the nine input variables, you could continue additional studies just as effectively with six shapes since the others did not have a great influence on the output responses. This will save computational effort.

In this tutorial, you will use the six shapes variables.

Length1:	Lower Bound = -0.5, Initial Bound = 0.0, Upper Bound = 2.0
Length2:	Lower Bound = 0.0, Initial Bound = 0.0, Upper Bound = 2.0
Length3:	Lower Bound = -1.0, Initial Bound = 0.0, Upper Bound = 1.0
Length4:	Lower Bound = -1.0, Initial Bound = 0.0, Upper Bound = 1.0
Length5:	Lower Bound = -1.0, Initial Bound = 0.0, Upper Bound = 1.0
Height:	Lower Bound = -1.0, Initial Bound = 0.0, Upper Bound = 1.0

You will begin this tutorial by creating a Modified Extensible Lattice Sequence (MELS) DOE. MELS is a space filling DOE designed to equally spread out points in a space by minimizing clumps and empty spaces. The minimal required number of points to create a second order polynomial with N variables is 1.1*(N + 1)*(N + 2)/2. Using this matrix, you will then create the following Fits for both output responses: Least Square Regression (LSR), Moving Least Square (MLSM), HyperKriging (HK), and Radial Basis Function (RBF).

Step 1: Run MELS DOE Study

In order to create the approximations to be used as surrogate models, you must perform specific DOEs that will serve as the input matrix. You will need to run a DOE suitable to be used in response surface creation, such as MELS.

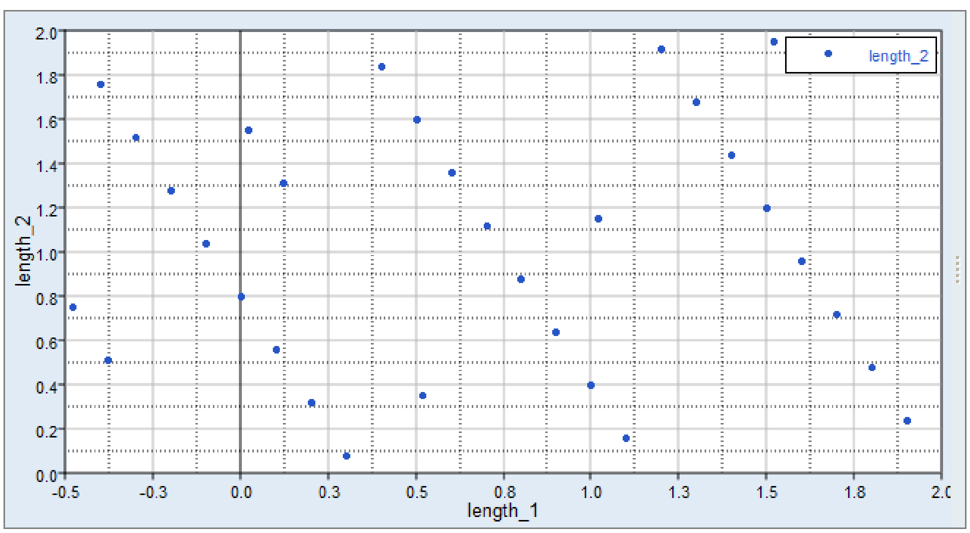
1. In the **Explorer**, right-click and select ***Add*** from the context menu.
2. In the **Add - HyperStudy** dialog, select ***DOE*** and click ***OK***.
3. Go to the **Select Input Variables** step.
4. In the work area, **Active** column, clear the **radius_1**, **radius_2** and **radius_3** checkboxes.

	Active	Label	Varname	Low
1	☑	length_1	m_1_length_1	-0.500
2	☑	length_2	m_1_length_2	0.000
3	☑	length_3	m_1_length_3	-1.000
4	☑	length_4	m_1_length_4	-1.000
5	☑	length_5	m_1_length_5	-1.000
6	☐	radius_1	m_1_radius_1	-2.000
7	☐	radius_2	m_1_radius_2	-0.500
8	☐	radius_3	m_1_radius_3	-0.500
9	☑	height	m_1_height	-1.000

5. Go to the **Specifications** step.
6. In the work area, set the **Mode** to ***Modified Extensible Lattice Sequence (MELS)***.
7. In the **Settings** tab, verify that the **Number of runs** is set to 31.
8. Click ***Apply***.
9. Go to the **Evaluate** step.
10. Click ***Evaluate Tasks***.
11. Go to the **Post-Processing** step.
12. Click the ***Scatter*** tab to review a 2D scatter plot of the results from the MELS DOE.

    The image below illustrates a typical sampling of the MELS DOE with 31 runs (length_1 vs. length_2).

    Note: This visualization is a projection of 31 points distributed in 6 dimensions onto a 2 dimensional plane.



13. **Optional**. Create a second DOE with less number of runs to be used as a Validation matrix in the Fit approach.

A Validation matrix provides information on the Fit's prediction accuracy.

Note: You should not use MELS as a Validation matrix, as it will take the same first runs from the MELS Input matrix due to its extensibility.

In this tutorial, you will use the Hammersley method to create the Validation matrix.

a. In the **Explorer**, right-click on the MELS DOE and select ***Copy*** from the context menu.
b. In the **Specifications** step, set the **Mode** to ***Hammersley***.
c. In the **Settings** tab, change the **Number of runs** to `12`.
d. Click ***Apply***.
e. In the **Evaluate** step, click ***Evaluate Tasks***.

Step 2: Setting Up a Fit Approach

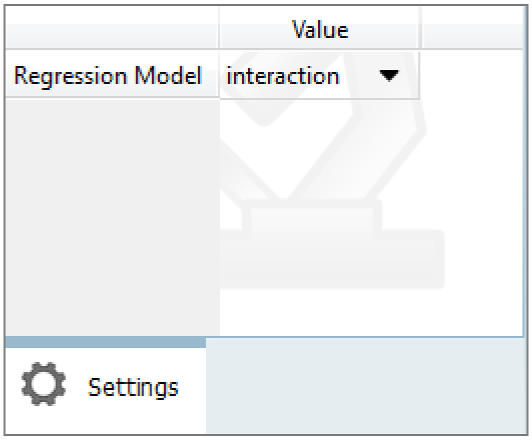
Using the 31 runs from the MELS DOE as an Input matrix and the 12 runs from the Hammersley DOE as a Validation matrix, create the following fits: Least Square Regression (LSR), Moving Least Square Method (MLSM), HyperKriging (HK), and Radial Basis Function (RBF).

1. In the **Explorer**, right-click and select ***Add*** from the context menu.
2. In the **Add - HyperStudy** dialog, select ***Fit*** and click ***OK***.
3. Go to the **Select** M**atrices** step.
4. Click ***Add Matrix*** two times.
5. Define **FitMatrix 1** and **FitMatrix 2**, by selecting the options indicated in the image below from the **Type** and **Matrix Source** columns.

	Active	Label	Varname	Type	Matrix Source	Matrix Origin	Status
1	☑	Fit Matrix 1	fitmatrix_1	Input	DOE, MELS (doe_5)	DoeDOE, MELS	Import Pending
2	☑	Fit Matrix 2	fitmatrix_2	Validation	DOE, Hammersley (doe_6)	DoeDOE, Hammersley	Import Pending

6. Click ***Import Matrix***.
7. Go to the **Specifications** step.
8. In the work area, set the **Mode** to the appropriate **Fit** method.
9. For the Least Sqaure Regressions (LSR) Fit, in the **Settings** tab, set **Regression Model** to ***Interaction***.

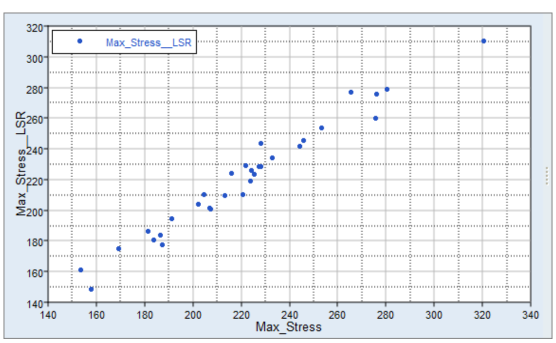
   An Interaction regression model enables linear and cross terms to be considered in the function f(x,y)=A+Bx+Cy+Dxy; where the first three terms are linear, and the last term is a cross term between the variables.



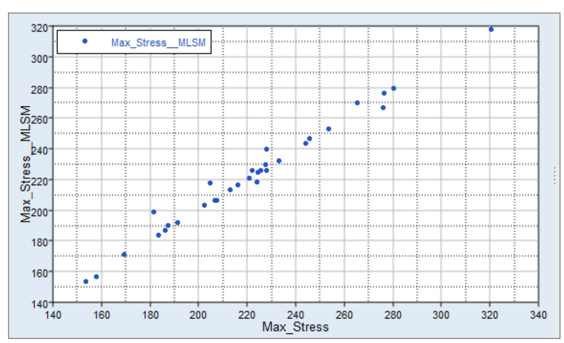
10. Click ***Apply***.
11. Go to the **Evaluate** step.
12. Click ***Evaluate Tasks***.
13. Go to the **Post-Processing** step.
14. Click the ***Scatter*** tab to compare the original Max_Stress output response to the Fit Max_Stress.

    The scatter shows the Fit accuracy. The closer together the points are along the diagonal, the better the fit. In the Max_Stress vs Max_Stress_LSR plot, you can see some dispersed points, which indicates the Fit has some inaccuracy. In comparison, the points in the Max_Stress vs Max_Stress_MLSM plot follow the diagonal more closely, which indicates it provides better Fit accuracy on Max_Stress.

    You will not compare HyperKriging and Radial Basis Function using scatter plots, because the results will be misleading. HyperKriging and Radial Basis Function go through the exact points by default, therefore the scatter plot comparing the original output response vs. the Fit output response will produce a straight line. However, this does not necessarily mean that the Fit has good predictive capability.



Max_stress and Max_stress_LSR comparison



Max_stress and Max_stress_MLSM comparison

15. Click the ***Diagnostics*** tab to review the diagnostics of the Fit study.

    The R-Square value measures how much of the variability of the response data around its mean is captured. If the model perfectly predicts the known values, R-Square will have a maximum possible value of 1.0.

	Criterion	Input Matrix	Cross-Validation Matrix	Validation Matrix
1	R-Square	0.9660680	0.6451045	-0.5578945
2	R-Square Adjusted	0.8868933	N/A	N/A
3	Multiple R	0.9828876	0.8031840	Undefined
4	Relative Average Absolute Error	0.1426813	0.4824845	0.8770736
5	Maximum Absolute Error	15.706113	47.025444	88.242350
6	Root Mean Square Error	6.8185250	22.051385	36.774466
7	Number of Samples	31	31	12

Diagnostics for Max_Stress, LSR

	Criterion	Input Matrix	Cross-Validation Matrix	Validation Matrix
1	R-Square	0.9804999	0.7722986	0.2518844
2	Relative Average Absolute Error	0.0795629	0.3524697	0.7668477
3	Maximum Absolute Error	17.634488	43.590150	48.971741
4	Root Mean Square Error	5.1689730	17.663163	25.483655
5	Number of Samples	31	31	12

Diagnostics for Max_Stress, MLSM

The R-square value for an Input Matrix in HyperKriging and Radial Basis Function has no meaning because the runs will always go through the exact data points, which will result in a value of 1.0. Although the value is 1.0, this does not mean the Fit will be accurate. In HyperKriging and Radial Basis Function, the only meaningful diagnostic values are for Cross-validation Matrix and Validation Matrix.

	Criterion	Input Matrix	Cross-Validation Matrix	Validation Matrix
1	R-Square	1.0000000	0.7584457	-6.9938493
2	Relative Average Absolute Error	1.40e-04	0.3703571	2.0746075
3	Maximum Absolute Error	0.0193176	48.369640	200.32467
4	Root Mean Square Error	0.0073919	18.192527	83.301981
5	Number of Samples	31	31	12

Diagnostics for Max_Stress, HK

	Criterion	Input Matrix	Cross-Validation Matrix	Validation Matrix
1	R-Square	1.0000000	0.7089073	0.3249540
2	Relative Average Absolute Error	1.37e-14	0.4071550	0.6727536
3	Maximum Absolute Error	1.36e-12	43.731749	48.951755
4	Root Mean Square Error	6.34e-13	19.971069	24.207172
5	Number of Samples	31	31	12

Diagnostics for Max_Stress, RBF

16. Click the ***Residuals*** tab to review the Error (and Percent Error) between the original output response and the Fit output response for each run of the Input and Validation matrices.

	Max_Stress	Max_Stress_LSR	Error	Percent Error
28	227.84496	243.53706	-15.692107	-6.8871864
5	153.32707	161.15593	-7.8288583	-5.1059856
6	265.17667	277.35118	-12.174513	-4.5910951
7	215.90872	224.17526	-8.2665369	-3.8287184
3	221.75601	229.64985	-7.8938428	-3.5596973
25	169.21870	175.00918	-5.7904743	-3.4218878
2	204.46399	210.79303	-6.3290358	-3.0954281
20	181.29922	186.85618	-5.5569586	-3.0650758
11	191.08934	194.52138	-3.4320380	-1.7960385
13	202.15817	204.38358	-2.2254038	-1.1008231
8	224.06998	226.32020	-2.2502254	-1.0042512
29	232.83224	234.09221	-1.2599674	-0.5411481
9	227.26093	228.46158	-1.2006530	-0.5283148
21	227.85101	228.87125	-1.0202365	-0.4477647
14	253.25926	253.93153	-0.6722710	-0.2654478

Input Matrix Residuals on Max_Stress, LSR

	Max_Stress	Max_Stress_LSR	Error	Percent Error
7	278.00589	189.76354	88.242350	31.741180
12	258.82968	217.22029	41.609396	16.075975
6	254.82101	230.99098	23.830037	9.3516766
5	253.55190	237.19830	16.353591	6.4498003
1	246.41943	242.40671	4.0127194	1.6284103
3	192.49901	192.07159	0.4274137	0.2220342
11	221.89412	224.24935	-2.3552289	-1.0614201
4	190.98055	197.83473	-6.8541842	-3.5889437
9	248.95694	262.80022	-13.843282	-5.5605128
10	266.74866	290.45870	-23.710042	-8.8885330
2	202.90829	223.83552	-20.927221	-10.313635
8	271.92908	339.85852	-67.929438	-24.980572

Validation Matrix Residuals on Max_Stress, LSR

The Input Matrix Residual errors are slightly smaller with Least Square Regression, than they are with Moving Least Square Method, but the Validation Matrix Residual errors are much smaller with Moving Least Square Method.

	Max_Stress	Max_Stress_MLSM	Error	Percent Error
20	181.29922	198.93371	-17.634488	-9.7267311
2	204.46399	217.80293	-13.338942	-6.5238587
28	227.84496	239.88645	-12.041492	-5.2849499
3	221.75601	226.50677	-4.7507556	-2.1423345
6	265.17667	270.28086	-5.1041893	-1.9248260
15	187.07663	190.06010	-2.9834661	-1.5947829
9	227.26093	230.27711	-3.0161824	-1.3271892
25	169.21870	171.28462	-2.0659159	-1.2208556
23	245.55453	247.27224	-1.7177041	-0.6995204
1	186.22330	187.38918	-1.1658811	-0.6260662

Input Matrix Residuals on Max_Stress, MLSM

	Max_Stress	Max_Stress_MLSM	Error	Percent Error
8	271.92908	320.90082	-48.971741	-18.009012
4	190.98055	217.61007	-26.629522	-13.943579
9	248.95694	274.61162	-25.654679	-10.304866
2	202.90829	221.15364	-18.245346	-8.9919175
10	266.74866	286.23501	-19.486357	-7.3051376
3	192.49901	203.30775	-10.808737	-5.6149572
11	221.89412	227.11995	-5.2258324	-2.3551018
1	246.41943	236.24391	10.175520	4.1293496
12	258.82968	241.29414	17.535544	6.7749355
5	253.55190	229.39311	24.158782	9.5281411
6	254.82101	230.22777	24.593242	9.6511830
7	278.00589	238.36734	39.638550	14.258169

Validation Matrix Residuals on Max_Stress, MLSM

The Input Matrix Residuals are meaningless for HyperKriging and Radial Basis Function, as indicated in the Validation Matrix Residuals below.

	Max_Stress	Max_Stress_HK	Error	Percent Error
7	278.00589	140.75357	137.25232	49.370293
12	258.82968	177.66160	81.168080	31.359649
6	254.82101	232.73639	22.084621	8.6667188
5	253.55190	234.06450	19.487398	7.6857630
4	190.98055	190.61134	0.3692090	0.1933228
9	248.95694	264.39798	-15.441043	-6.2022945
2	202.90829	226.15834	-23.250041	-11.458398
11	221.89412	249.93091	-28.036792	-12.635212
10	266.74866	308.47441	-41.725756	-15.642349
1	246.41943	332.77346	-86.354023	-35.043512
3	192.49901	270.49564	-77.996631	-40.517939
8	271.92908	472.25375	-200.32467	-73.667986

Validation Matrix Residuals on Max_Stress, HK

	Max_Stress	Max_Stress_RBF	Error	Percent Error
7	278.00589	229.05413	48.951755	17.608172
6	254.82101	222.49933	32.321685	12.684073
5	253.55190	233.47816	20.073738	7.9170137
12	258.82968	242.97057	15.859112	6.1272384
1	246.41943	243.09465	3.3247827	1.3492372
9	248.95694	249.87639	-0.91945...	-0.3693211
3	192.49901	197.92264	-5.42363...	-2.8174855
2	202.90829	216.43394	-13.5256...	-6.6658891
11	221.89412	238.39201	-16.4978...	-7.4350300
10	266.74866	289.73002	-22.9813...	-8.6153633
4	190.98055	209.21854	-18.2379...	-9.5496611
8	271.92908	311.66832	-39.7392...	-14.613826

Validation Matrix Residuals on Max_Stress, RBF

Comparison

The max percent of errors for Input and Validation matrices are as shown below:

	LSR (Interaction Regression model)	**MLSM**	**HK**	**RBF**
Max_Disp	-1.33%	-2.56%	-	-
Max_Stress	-6.88%	-9.73%	-	-

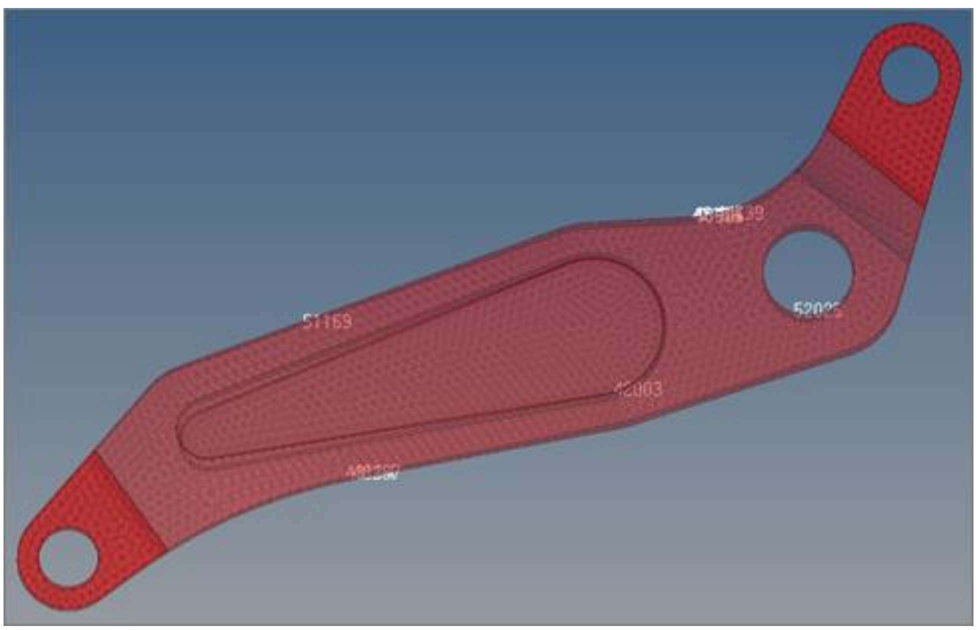
Input Matrix Residuals

	LSR (Interaction Regression model)	MLSM	HK	RBF
Max_Disp	6.23%	-3.33%	9.48%	-2.80%
Max_Stress	31.74%	-18.00%	49.37%	17.60%

Validation Matrix Residuals

It can be seen that the percent of errors for Max_Disp are smaller than Max_Stress. These results indicate the Fit approach works well for Max_Disp, but is not very efficient for Max_Stress.

These finding suggest that it is best to use the Fit model obtained from the MLSM for Max_Disp. An output response such as Max_Stress is a global envelope of localized effects. The nature of such an envelope type of output responses makes them difficult to capture accurately with a Fit. In contrast, the Max_Disp output response is not influenced by localized effects, therefore it is easier to use a Fit for such data. When proceeding in this situation, it is recommended that you either increase the number of samples, which is not guaranteed to improve the accuracy, or create a series of more localized output responses that would be simpler functions of the input variables; for example, several output responses that each capture the stress in specific regions. The image below highlights the areas of high stresses from the runs in the Input matrix.



Last modified: v2017.2 (12.1156684)