

Workshop – Prediction Analysis, Post-Buckling

AN MSC NASTRAN MACHINE LEARNING WEB APP TUTORIAL

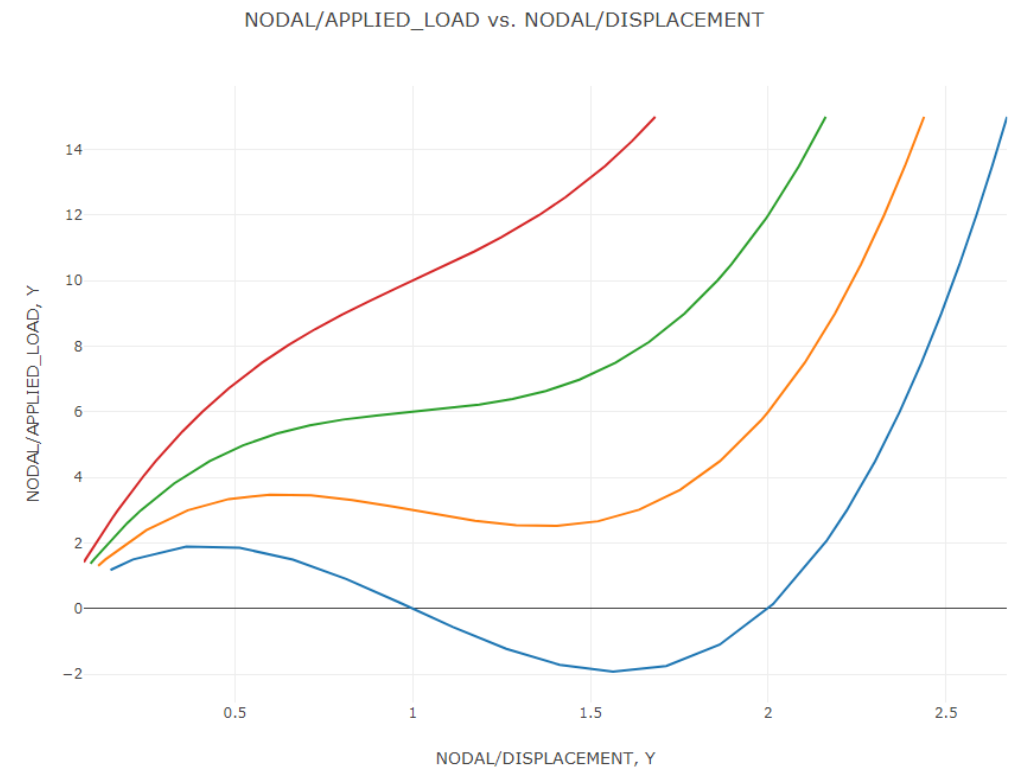
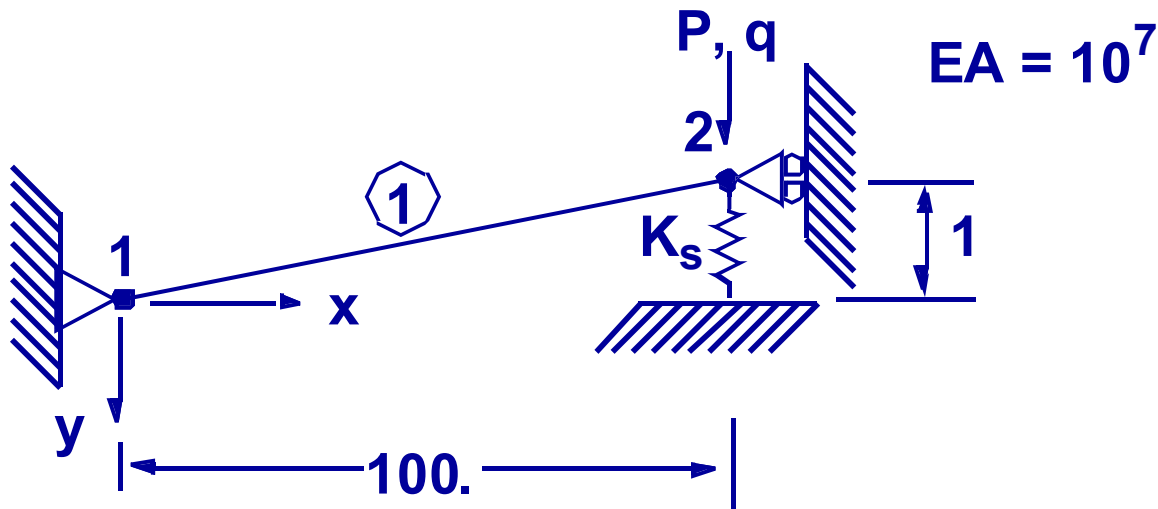
Goal: Prediction Analysis

This tutorial consists of multiple parts

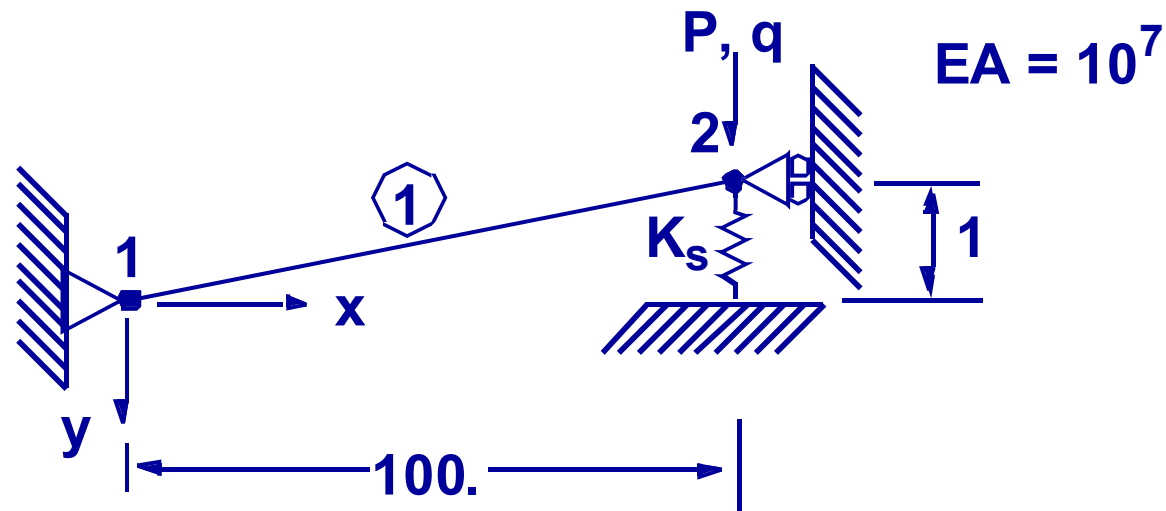
1. Configuring The Problem Statement
 - In this tutorial, we configure the parameters and the responses to monitor.
2. Configuring Multiple Batch Runs
 - This section discusses how to configure and execute multiple MSC Nastran runs.
3. Performing Predictions
 - Gaussian process (GP) regression is used to train a surrogate model.
 - Two (2) artificial samples are manually added to the training data, which results in a “spiked” surrogate model. A solution to the “spiked” surrogate model is demonstrated.
 - Excel is used to manually customize the training data.
4. Creating Plots with the HDF5 Explorer
 - The HDF5 Explorer web app is used to create load vs. displacement plots.

Details of the Structural Model

1. Perform a post-buckling analysis for different values of K_s
2. Create a Load vs. Displacement plot for node 2



Details of the Structural Model



Problem Statement

Design Variables

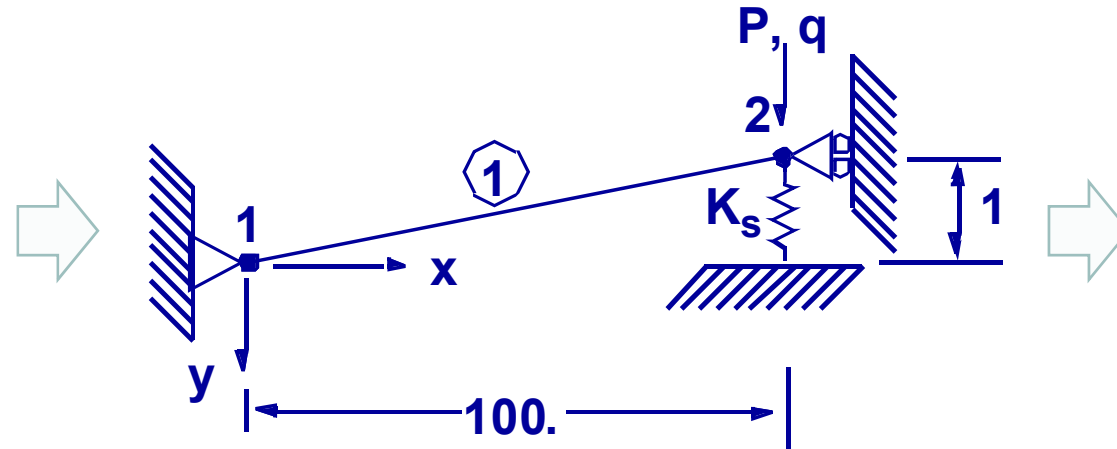
x1: K of PELAS 20

Note that K is displayed as K_s in the figure

$$.1 < x1 < 10.0$$

Samples

- Batch set 1 – 10 run LHS Design



Monitored Responses

r1: Displacement at node 2, y component, max displacement for time steps

r2: Displacement at node 2, y component, at time step .1

r3: Load at node 2, y component, max load for time steps

r4: Load at node 2, y component, at time step .1

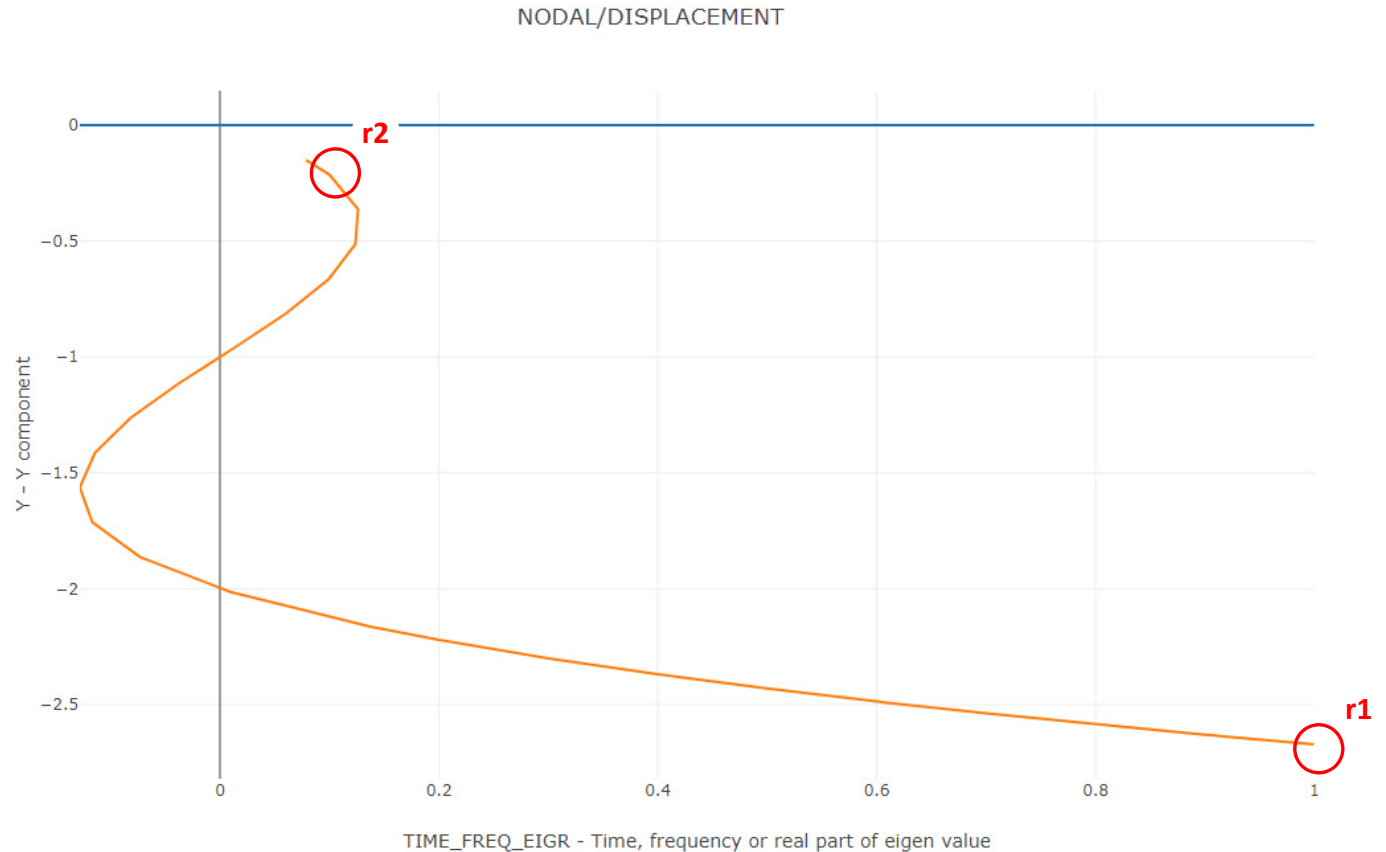
Problem Statement, Continued

The responses defined in this tutorial correspond to points on the displacement plot

- r1: The minimum point of displacement is tracked
- r2: The displacement at time step .1 is tracked

Surrogate Models

- When the surrogate model is created for r1, the surrogate model is used to predict the max displacement.
- When the surrogate model is created for r2, the model is used to predict the displacement only at time step .1.



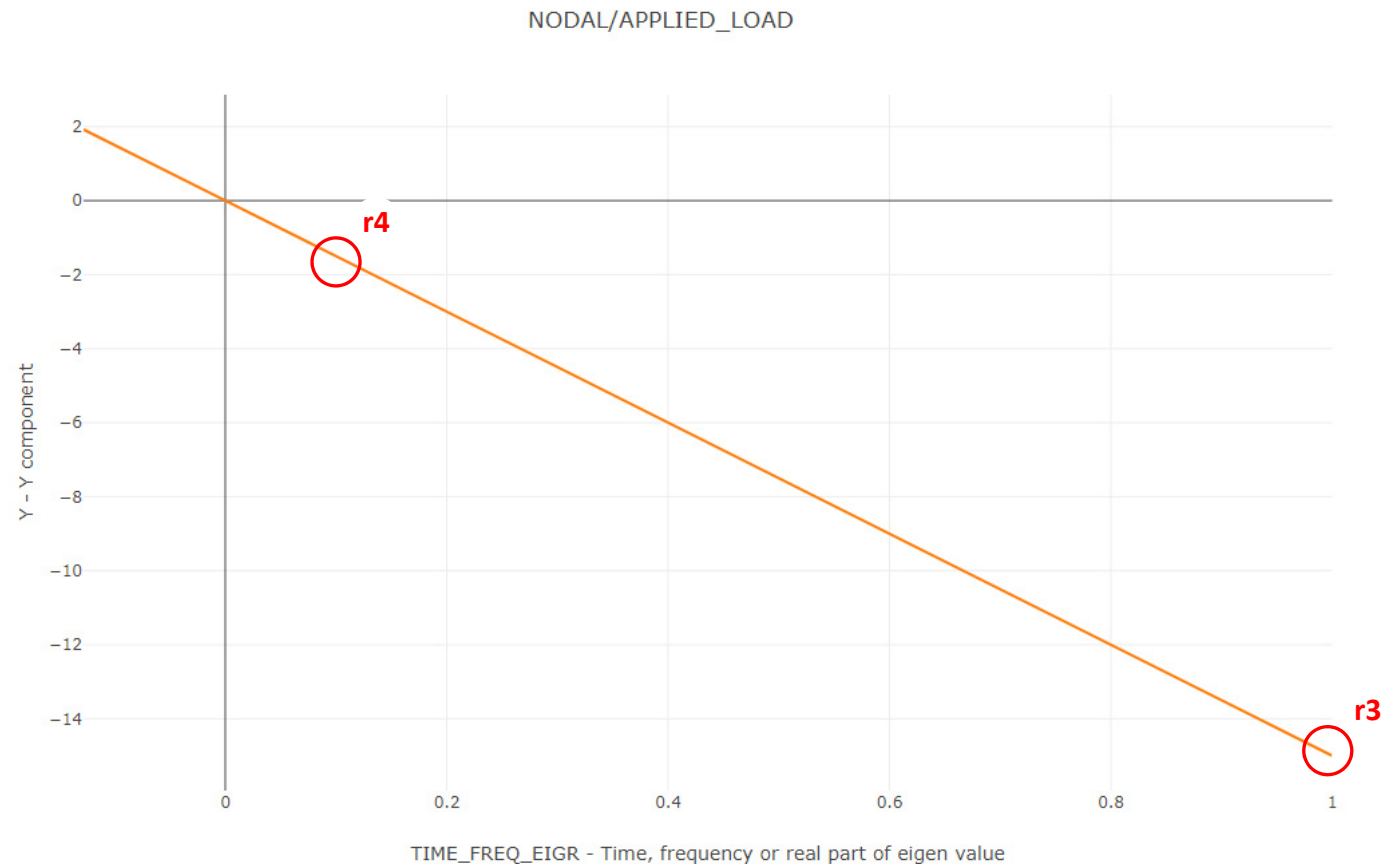
Problem Statement, Continued

The responses defined in this tutorial correspond to points on the applied load plot

- r3: The minimum point of applied load is tracked
- r4: The applied load at time step .1 is tracked

Surrogate Models

- A surrogate model is not necessary for the applied load since the loading is the same for all samples.
- The applied load responses are monitored so that the HDF5 Explorer automatically builds the applied load plot.
- With separate displacement and load plots, the HDF5 Explorer is used to create a combined Load vs. Displacement plot.



Contact me

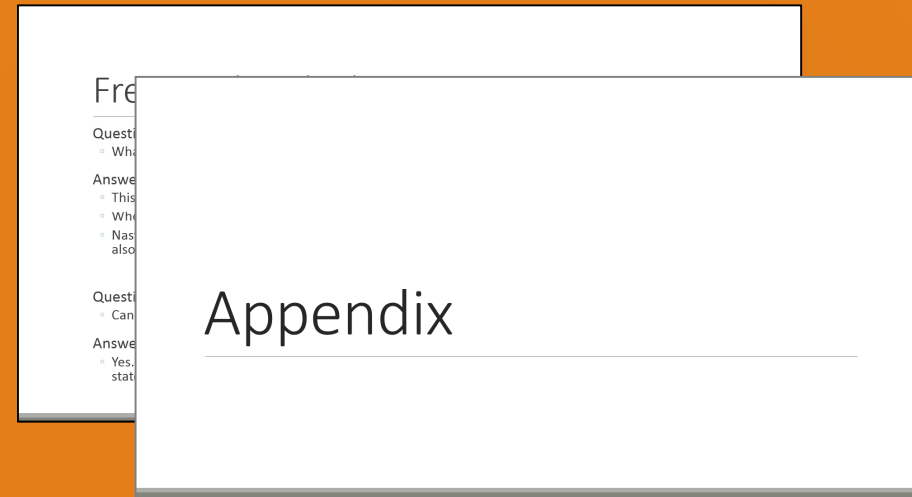
- Nastran SOL 200 training
- Nastran SOL 200 questions
- Structural or mechanical optimization questions
- Access to the SOL 200 Web App

christian@ the-engineering-lab.com

More Information Available in the Appendix

The Appendix includes information regarding the following:

- How to import and edit previous files
- What is Gaussian Process Regression?



Tutorial

Tutorial Overview

1. Start with a .bdf or .dat file
2. Use the Machine Learning web app to:
 1. Configure the problem statement
 2. Configure multiple batch runs
3. Use the Prediction Analysis web app to:
 1. Perform predictions
4. Use the HDF5 Explorer to:
 1. Create plots

Special Topics Covered

Training Data – The training data consists of the parameter inputs and respective output responses for multiple MSC Nastran runs. This tutorial describes how to configure multiple MSC Nastran runs, each with different parameter inputs, and how to monitor each response.

Gaussian process regression – This tutorial describes the procedure to use Gaussian process regression to train a surrogate model and make predictions.

Automatic Response Extraction – Often responses are manually or automatically extracted from the F06 file. This becomes challenging when extracting responses from multiple F06 files. This tutorial highlights the web app's ability to automatically extract responses from multiple H5 files with minimal user effort.

Automatic Plots - Multiple plots must be created. This traditionally involves extracting data from result files, for example the F06 file, and using Excel to create the plots. This tutorial highlights the automatic generation of these plots that require minimal user effort.

SOL 200 Web App Capabilities

The Post-processor Web App and HDF5 Explorer are free to MSC Nastran users.

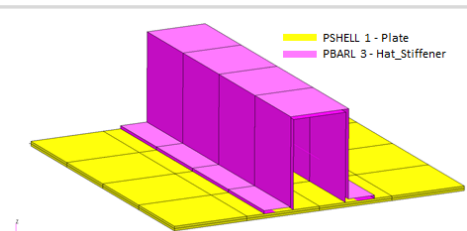
Compatibility

- Google Chrome, Mozilla Firefox or Microsoft Edge
- Windows and Red Hat Linux
- Installable on a company laptop, workstation or server. All data remains within your company.

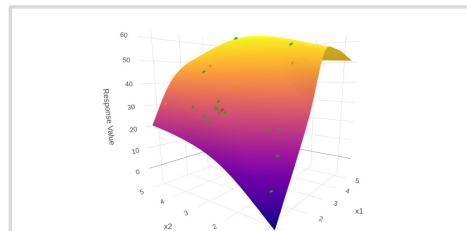
Web Apps

Benefits

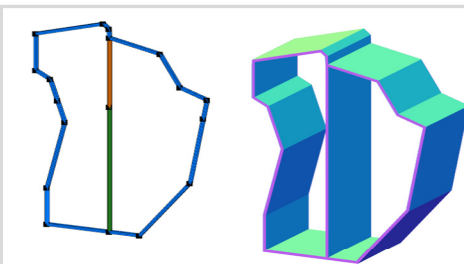
- REAL TIME error detection. 200+ error validations.
- REAL TIME creation of bulk data entries.
- Web browser accessible
- Free Post-processor web apps
- +80 tutorials



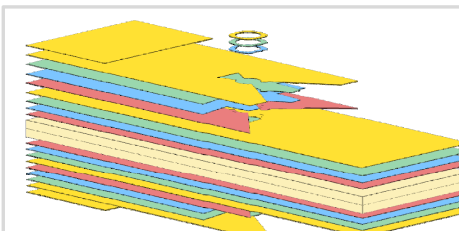
Web Apps for MSC Nastran SOL 200
Pre/post for MSC Nastran SOL 200.
Support for size, topology, topometry, topography, multi-model optimization.



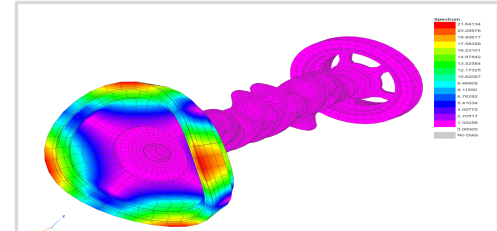
Machine Learning Web App
Bayesian Optimization for nonlinear response optimization (SOL 400)



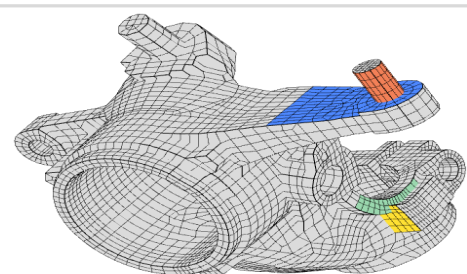
PBMSECT Web App
Generate PBMSECT and PBRSECT entries graphically



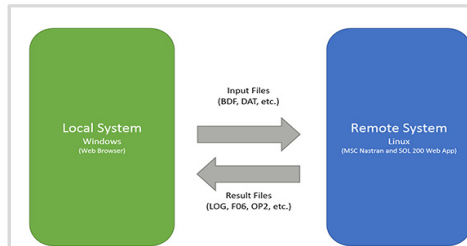
Ply Shape Optimization Web App
Optimize composite ply drop-off locations, and generate new PCOMP entries



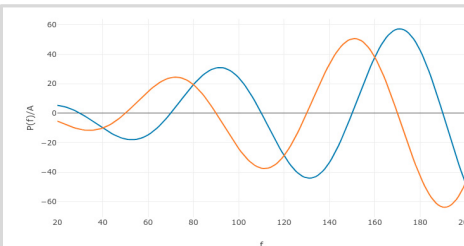
Post-processor Web App
View MSC Nastran results in a web browser on Windows and Linux



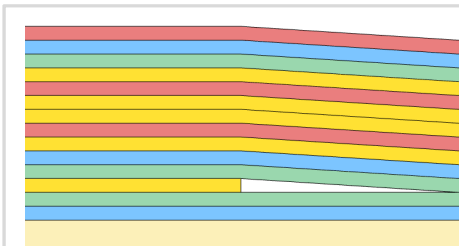
Shape Optimization Web App
Use a web application to configure and perform shape optimization.



Remote Execution Web App
Run MSC Nastran jobs on remote Linux or Windows systems available on the local network



Dynamic Loads Web App
Generate RLOAD1, RLOAD2 and DLOAD entries graphically



Stacking Sequence Web App
Optimize the stacking sequence of composite laminate plies



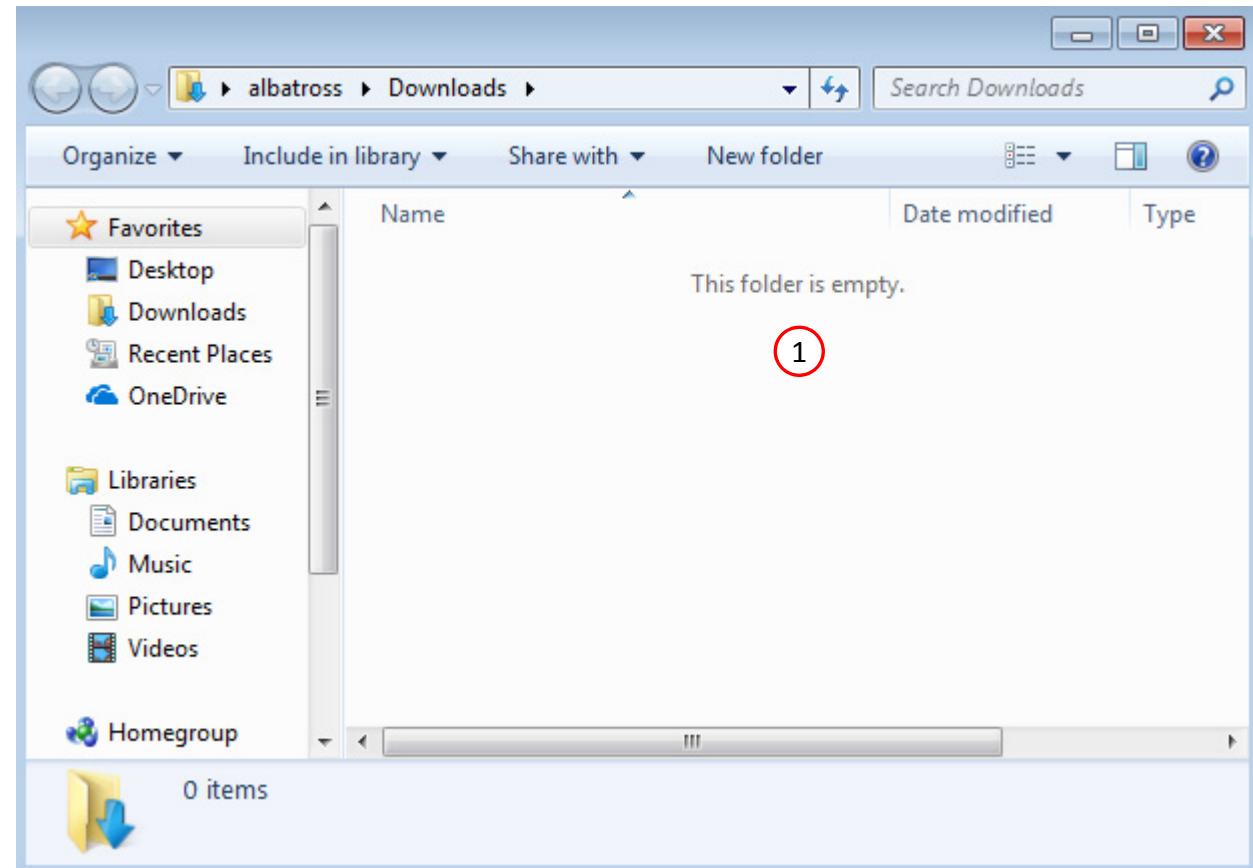
HDF5 Explorer Web App
Create graphs (XY plots) using data from the H5 file

Configuring The Problem Statement

Before Starting

1. Ensure the Downloads directory is empty in order to prevent confusion with other files

- Throughout this workshop, you will be working with multiple file types and directories such as:
 - .bdf/.dat
 - nastran_working_directory
 - .f06, .log, .pch, .h5, etc.
- To minimize confusion with files and folders, it is encouraged to start with a clean directory.



Go to the User's Guide

1. Click on the indicated link

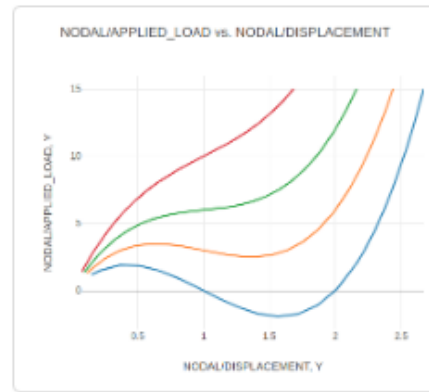
- The necessary BDF files for this tutorial are available in the Tutorials section of the User's Guide.



Obtain Starting Files

1. Find the indicated example
2. Click Link
3. The starting file has been downloaded

- When starting the procedure, all the necessary BDF, or DAT, files must be collected and uploaded together. Relevant INCLUDE files must also be collected and uploaded.

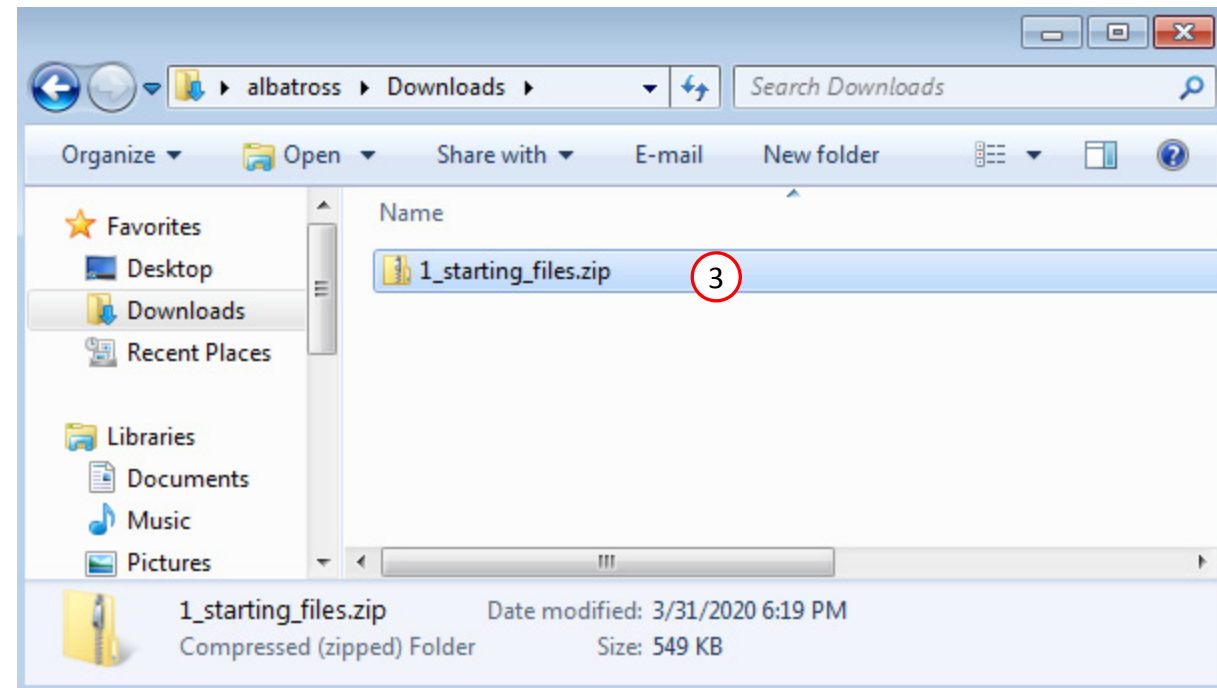


Prediction Analysis, Post-Buckling ①

Consider a post-buckling analysis. The parameter allowed to vary is a spring constant. The responses of interest include the applied load and displacements.

This tutorial is similar to the previous tutorial named Prediction Analysis, Buckling. Like the previous tutorial, this tutorial discusses the use of Gaussian process (GP) regression to train a surrogate model. To expand your experience with GP regression, this tutorial purposely demonstrates a scenario where a poorly fitted model is obtained and what the procedure is to remedy this issue. Also discussed are instructions to create load vs. displacement plots.

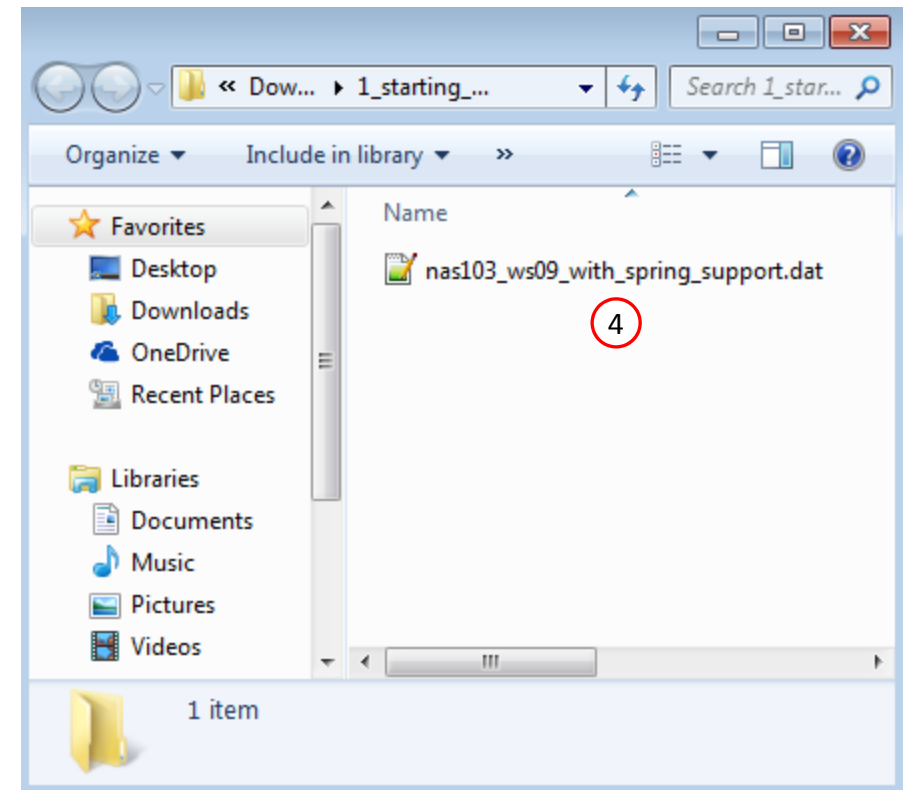
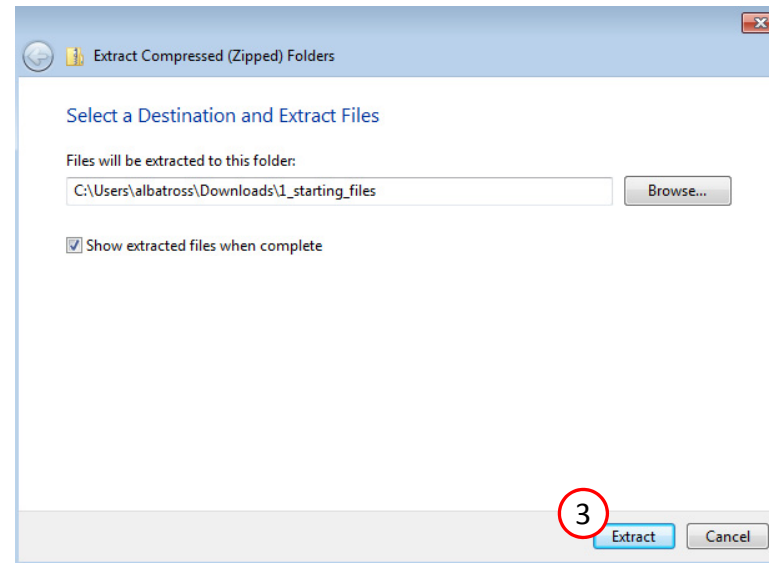
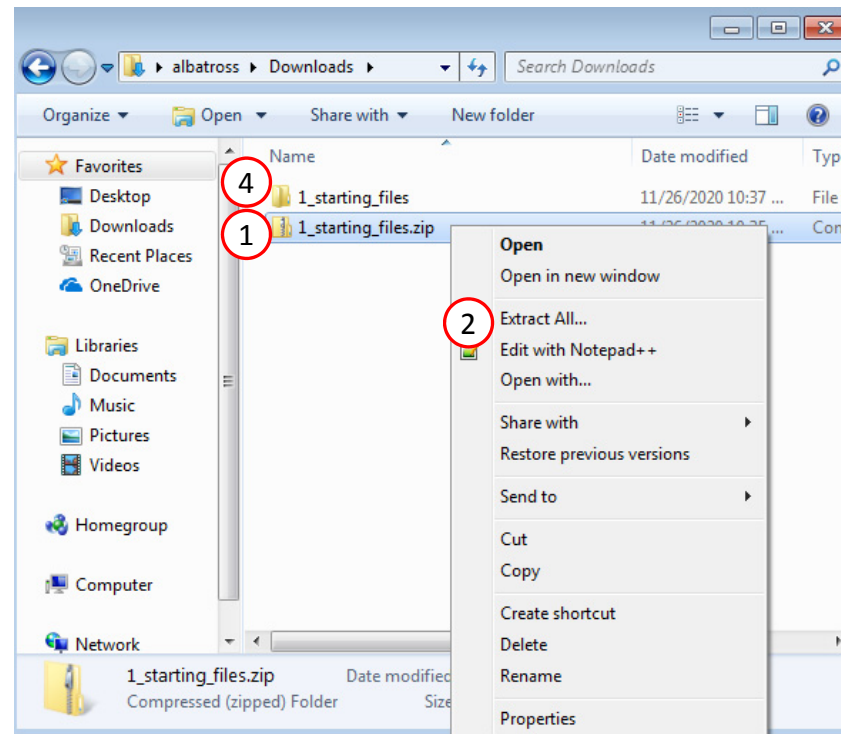
Starting Files: [Link](#) ②
Solution BDF Files: [Link](#)



Obtain Starting Files

1. Right click on the zip file
2. Select Extract All...
3. Click Extract
4. The starting files are now available in a folder

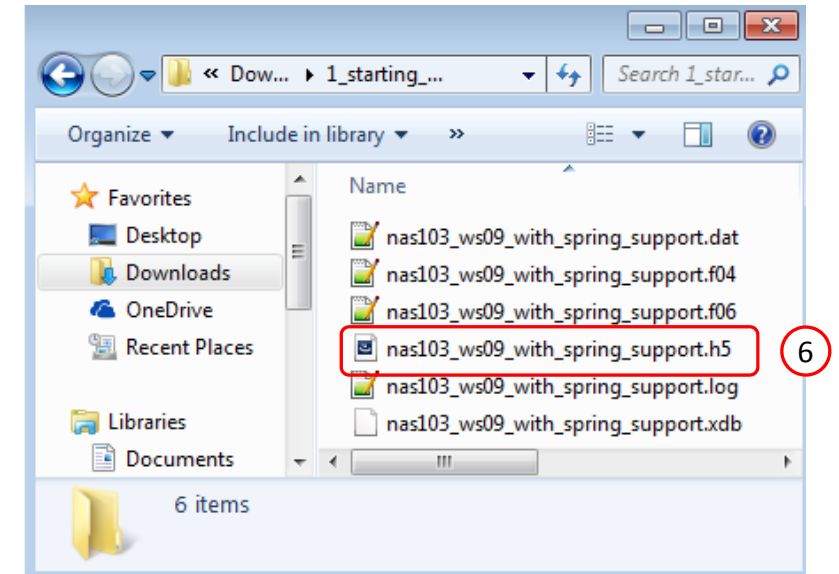
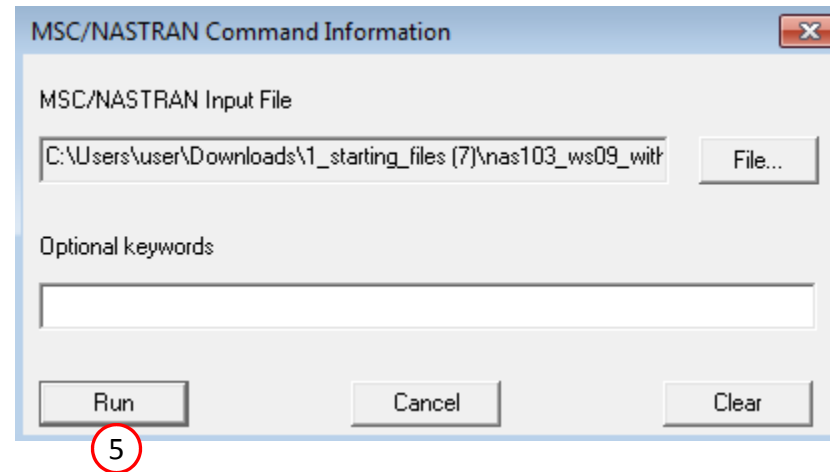
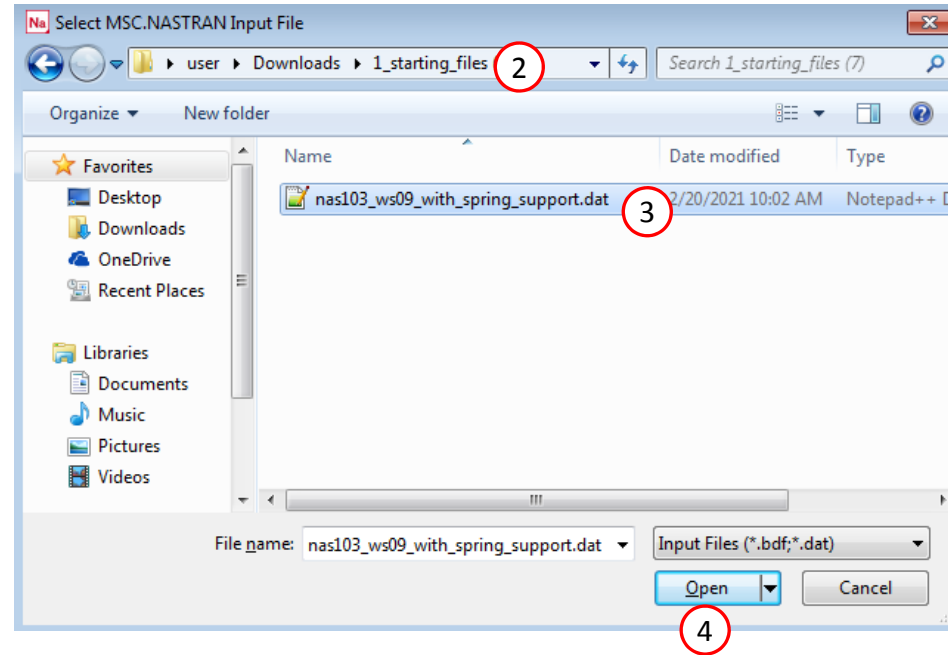
- The starting files for this tutorial are contained in a ZIP file and must be extracted as shown.



Create the Starting H5 File

A starting H5 file must be created. This H5 file will be used to configure the responses later on.

1. Double click the MSC Nastran desktop shortcut
2. Navigate to the directory named 1_starting_files
3. Select the indicated file
4. Click Open
5. Click Run
6. The starting H5 file is created



Use the same MSC Nastran version throughout this exercise

The following applies if you have multiple versions of MSC Nastran installed.

To ensure compatibility, use the same MSC Nastran version throughout this exercise.
For example, scenario 1 is OK but scenario 2 is NOT OK.

- Scenario 1 - OK
 - MSC Nastran 2021 is used to create the starting H5 file.
 - MSC Nastran 2021 is used for each run during Machine Learning or Parameter study.
- Scenario 2 – NOT OK
 - MSC Nastran 2018.2 is used to create the starting H5 file.
 - MSC Nastran 2021 is used for each run during Machine Learning or Parameter study.

Using the same MSC Nastran version is critical for consistent response extraction from the H5 file. A response configured for Nastran version X may not match in Nastran version Y, which leads to unsuccessful response extraction from the H5 files. The goal is to make sure all H5 files generated are from the same MSC Nastran version.

Open the Correct Page

1. Click on the indicated link

- MSC Nastran can perform many optimization types. The SOL 200 Web App includes dedicated web apps for the following:
 - Optimization for SOL 200 (Size, Topology, Topometry, Topography, Local Optimization, Sensitivity Analysis and Global Optimization)
 - Multi Model Optimization
 - Machine Learning
- The web app also features the HDF5 Explorer, a web application to extract results from the H5 file type.



Select BDF Files

1. Click Select files
2. Select the indicated files
3. Click Open
4. Click Upload files

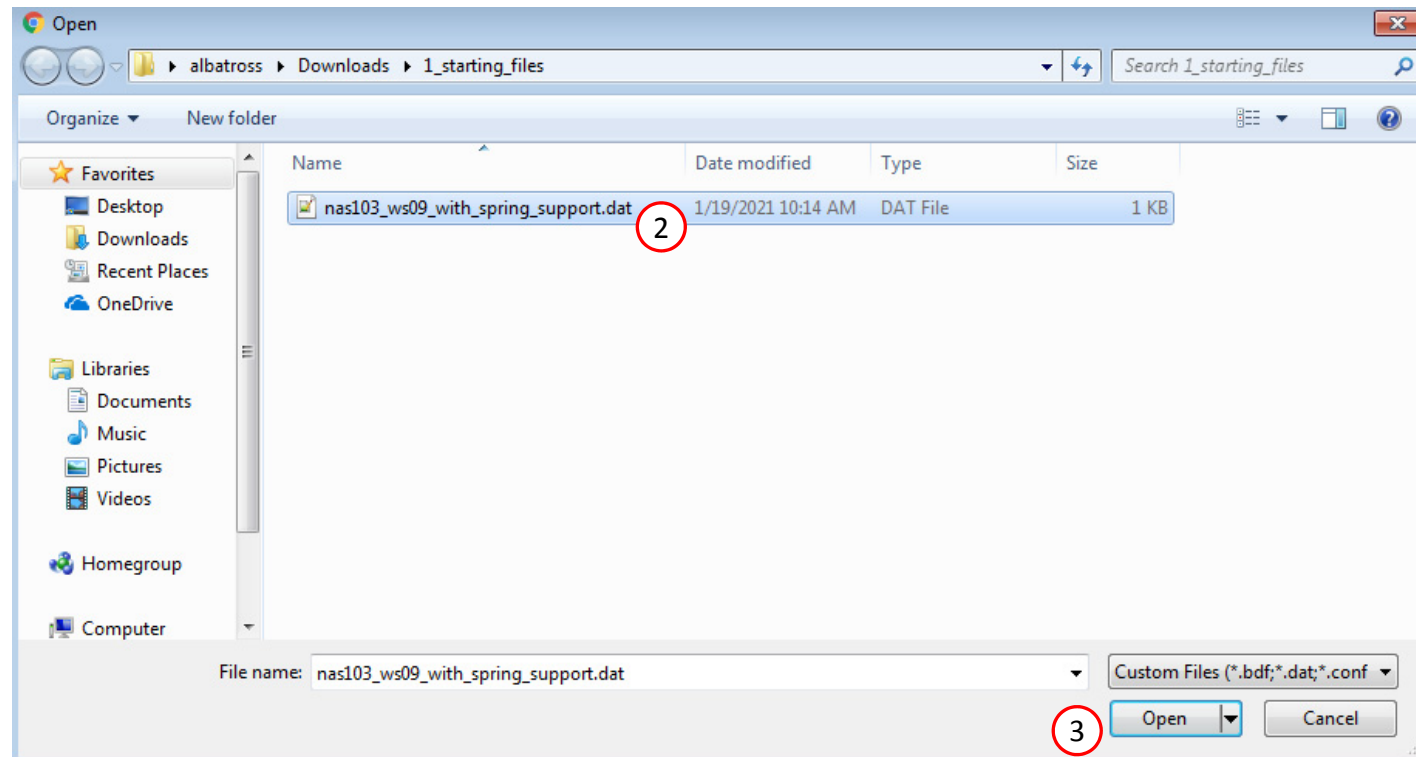
Select BDF Files

1. Select files nas103_ws09_with_spring_support.dat

Inspecting: 100%

4. Upload files

Uploading: 100 %



Parameters

1. Set the following fields as parameters
 - x1: The spring stiffness of PELAS 20
2. Parameters have been created for the selected fields
3. For each parameter, use the following settings:
 - Low: .1
 - High: 10.

- Bulk data entries will always be displayed in the small field format.
- Only fields that have real or integer data entries may be selected as parameters. If the field is blank or contains only characters, the field may not be selected.

Select Parameters

\$ _1_ _2_ _3_ _4_ _5_ _6_					
FORCE	15	2		15.	0. -1.
MAT1	1	10.E7			
NLPCI	20	CRIS	1.	1.	
PELAS	20	%x1%			
PROD	10	1	.1		

Configure Parameters

Delete	Parameter	Status	Low	High	Comments
2 ✖	x1	✔	.1	10.	Field 3 of F

Responses

1. Click Responses
2. Click Select files
3. Select the indicated file
4. Click Open
5. Click Upload files

- On this page, the H5 file is uploaded to the web app.

1

Upload .h5 File

2

1. Select files

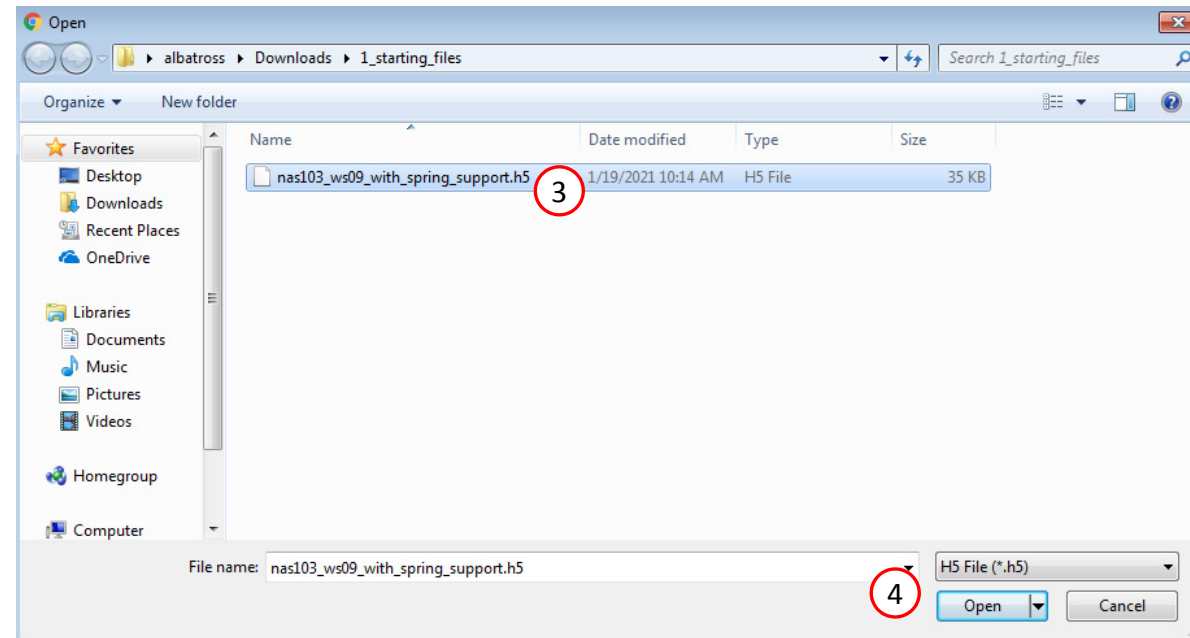
nas103_ws09_with_spring_support.h5

5

2. Upload files

Uploading

Loading



Adjust the Column Width

- 1. Optional - Use at your liking the buttons at the top right hand corner to adjust the width of the left and right columns

SOL 200 Web App - Machine LearningParametersSamplesResponsesDownloadResultsConnectionSettingsHome

Select Responses to MonitorSession ID: 3710HDF5

Select Dataset

NODAL/ACCELERATION
NODAL/DISPLACEMENT
NODAL/GRID_WEIGHT
NODAL/VELOCITY_CPL
SUMMARY/EIGENVALUE

Acquired Dataset

NODAL/GRID_WEIGHT - 0

Reset Filters

SAMPLE	DOMAIN_ID	SUBCASE	STEP	ANA
--------	-----------	---------	------	-----

View Responses to Monitor

Monitored Responses

Hide/Show ColumnsReset FiltersDownload CSV

Delete	Label	Status	Objective	Lower Bound	Upper Bound	Monitor th of the Fil cycle (SO
	r1			Lower	Upper	

SOL 200 Web App - Machine LearningParametersSamplesResponsesDownloadResultsConnectionSettingsHome

Select Responses to MonitorSession ID: 3710HDF5

Select Dataset

NODAL/ACCELERATION_CPLX
NODAL/DISPLACEMENT_CPLX
NODAL/GRID_WEIGHT
NODAL/VELOCITY_CPLX
SUMMARY/EIGENVALUE

Specify Entities

0

Examples: 0, etc.

Acquired Dataset

NODAL/GRID_WEIGHT - 0

Reset Filters

SAMPLE	DOMAIN_ID	SUBCASE	STEP	ANALYSIS	TIME_FREQ_EIGR	
						Im eig api
lame of H5 File**		Subcase number	Step number	Analysis type	Time, frequency or real part of eigen value	
model		0	0	0	0	

View Responses to Monitor

Monitored Responses

Hide/Show ColumnsReset FiltersDownload CSV

Delete	Label	Status	Objective	L B
	r1			Lc

510203050100

Select Responses

1. Select the following dataset:
NODAL/DISPLACEMENT
2. For the Grid Identifier, select 2
3. Select the indicated cells
4. The newly created Response to Monitor are listed as r1 and r2
5. Adjust the horizontal bar to make the following column visible: Monitor the maximum or minimum [...]
6. Configure the following setting for response r1
 1. Monitor the maximum or minimum [...]: Yes
7. Configure the following setting for response r2
 1. Monitor the maximum or minimum [...]: No

- Any cell that includes a single decimal point can be set as a response to monitor.
- For this example, the displacement, y component, at grid 2 has been set as a monitored response.

Select Responses to Monitor

Select Dataset

ELEMENTAL/ELEMENT_FORCE/ELA
ELEMENTAL/ELEMENT_FORCE/ROI
NODAL/APPLIED_LOAD
NODAL/DISPLACEMENT 1

Specify Entities

1, 2

Grid identifier (ID)
Examples: 1, 2, etc.

☒ Auto Execute

Acquire Dataset

Acquisition complete and successful

Acquired Dataset

NODAL/DISPLACEMENT - 1, 2

ID	X	Y
1		
2		
2	0	-0.15000749...
2	0	-0.21331443...
2	0	-0.33333333...
2	0	-0.51332942...
2	0	-0.66333692...
2	0	-0.81334442...
2	0	-0.96335192...
2	0	-1.11335941...
2	0	-1.26336691...
2	0	-1.41337441...

View Responses to Monitor

Monitored Responses

Hide/Show Columns Reset Filters Download CSV

Label	Status	Objective	Lower Bound	Upper Bound	Monitor the response of the FINAL design cycle (SOL 200 only)	Monitor the maximum or response, whichever is greatest absolute v
r1	<input checked="" type="checkbox"/>	<input type="text"/>	Lower	Upper	6	Yes - Monitor the maximum
r2	<input checked="" type="checkbox"/>	<input type="text"/>	Lower	Upper	7	No - Monitor response of cor

5

Select Responses

1. Select the following dataset:
NODAL/APPLIED_LOAD
2. For the Grid Identifier, select 2
3. Select the indicated cells
4. The newly created Response to Monitor are listed as r3 and r4
5. Adjust the horizontal bar to make the following column visible: Monitor the maximum or minimum [...]
6. Configure the following setting for response r3
 1. Monitor the maximum or minimum [...]: Yes
7. Configure the following setting for response r4
 1. Monitor the maximum or minimum [...]: No

- Any cell that includes a single decimal point can be set as a response to monitor.
- For this example, the applied load, y component, at grid 2 has been set as a monitored response.

Select Responses to Monitor

Select Dataset

ELEMENTAL/ELEMENT_FORCE/ELA
 ELEMENTAL/ELEMENT_FORCE/RO
NODAL/APPLIED_LOAD 1
 NODAL/DISPLACEMENT

Specify Entities

1, 2

Grid identifier (ID)
Examples: 1, 2, etc.

☒ Auto Execute

Acquire Dataset

Acquisition complete and successful

Acquired Dataset
 NODAL/APPLIED_LOAD - 1, 2

ID	X	Y
Grid identifier	X component	Y component
1		
2		
2	0	-1.17941642...
2	0	-1.5
2	0	-1.89313835...
2	0	-1.85735858...
2	0	-1.49306371...
2	0	-0.90150563...
2	0	-0.18394389...
2	0	0.55835769...
2	0	1.22413448...
2	0	1.71212441...

View Responses to Monitor

Monitored Responses

Hide/Show Columns Reset Filters Download CSV

Label	Status	Objective	Lower Bound	Upper Bound	Monitor the response of the FINAL design cycle (SOL 200 only)	Monitor the maximum or response, whichever is greatest absolute v
r1	<input checked="" type="checkbox"/>		Lower	Upper		Yes - Monitor the maximum
r2	<input checked="" type="checkbox"/>		Lower	Upper		No - Monitor response of co
r3	<input checked="" type="checkbox"/>		Lower	Upper		Yes - Monitor the maximum
r4	<input checked="" type="checkbox"/>		Lower	Upper		No - Monitor response of co

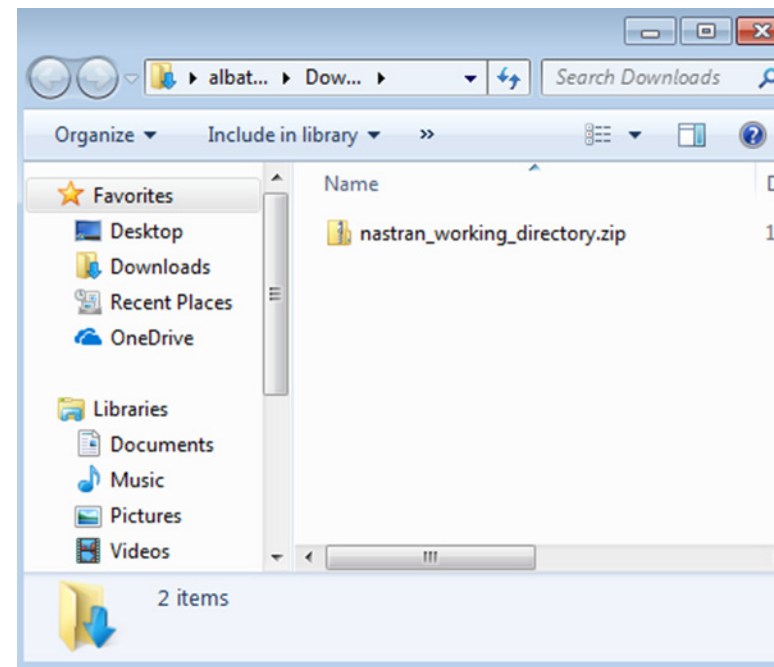
5 10 20 30 50 100

5

Configuring Multiple Batch Runs

Samples

In the following slides, we will configure 1 batch to run.



Batch	File Name	Number of Runs	Purpose
1	nastran_working_directory.zip	10	The data from these 10 runs is used to train the surrogate model.

- For problems with 1-5 parameters, usually 10 runs per parameter is sufficient. This example purposely uses 20 runs (20 runs per parameter) to demonstrate an issue. The issue is discussed later in this tutorial.

Samples

1. Click Samples
2. Ensure the following design is selected:
Latin Hypercube, Reproducible
3. Set Number of Samples to 10
4. The samples have been updated, note that
samples 1, 2, 3, ..., 10 are visible
5. The indicated controls can be used to
display the other samples

SOL 200 Web App - Machine Learning Parameters **Samples** Responses Download Results Connection Settings Home

1

Configure Samples

Design

Latin Hypercube, Reproducible 2

+ Info

Number of Samples

10 3

Samples to Run

+ Options

4

	Parameters
Sample Number	x1
1	1.208613
2	5.371329
3	9.449995
4	4.353574
5	6.326213

5

« 1 2 » 5 10 20 30 40 50

◀ ▶

Download

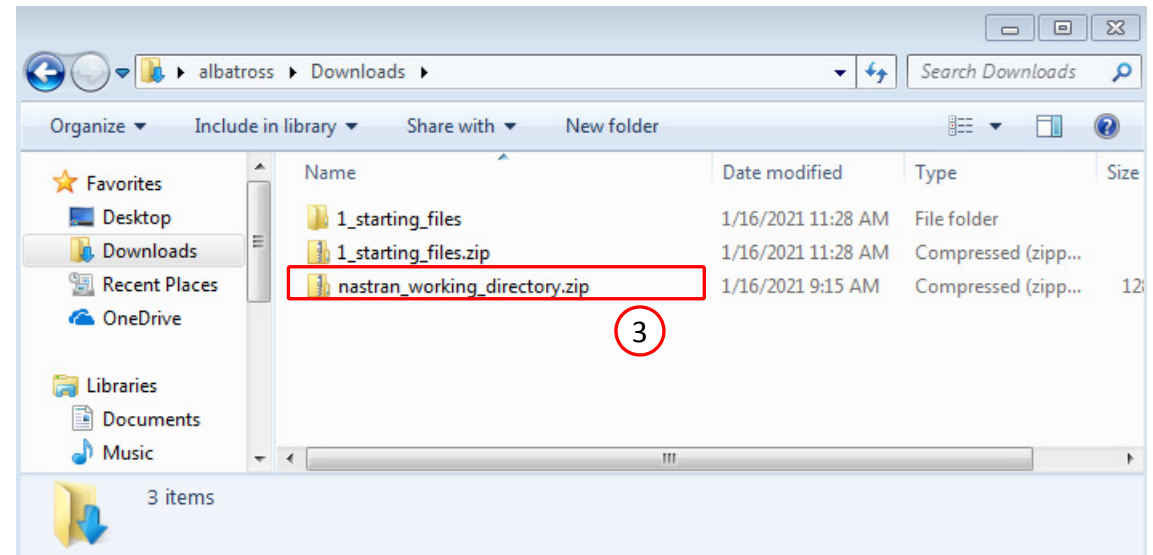
1. Click Download
2. Click Download BDF Files
3. A new ZIP file has been downloaded

1

Download

Download BDF Files

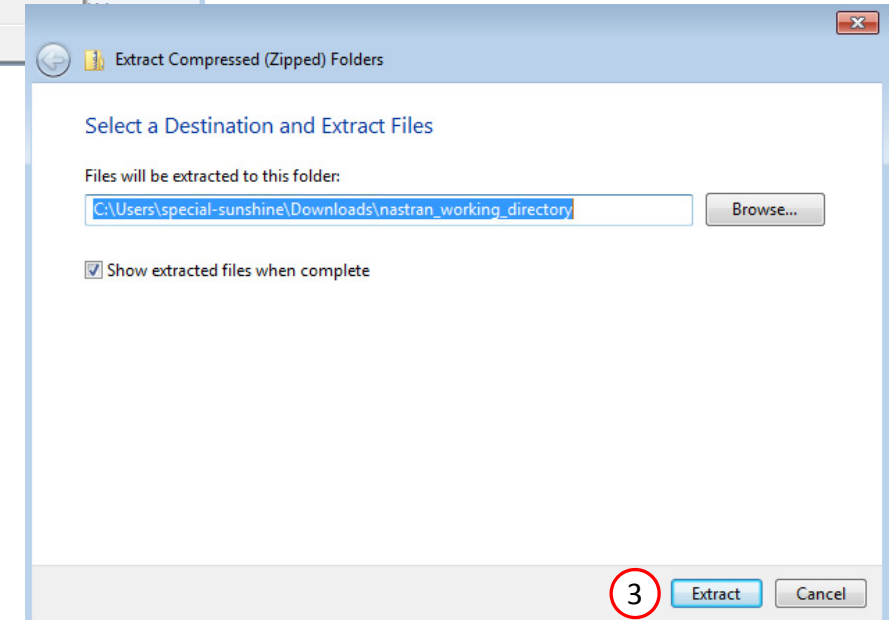
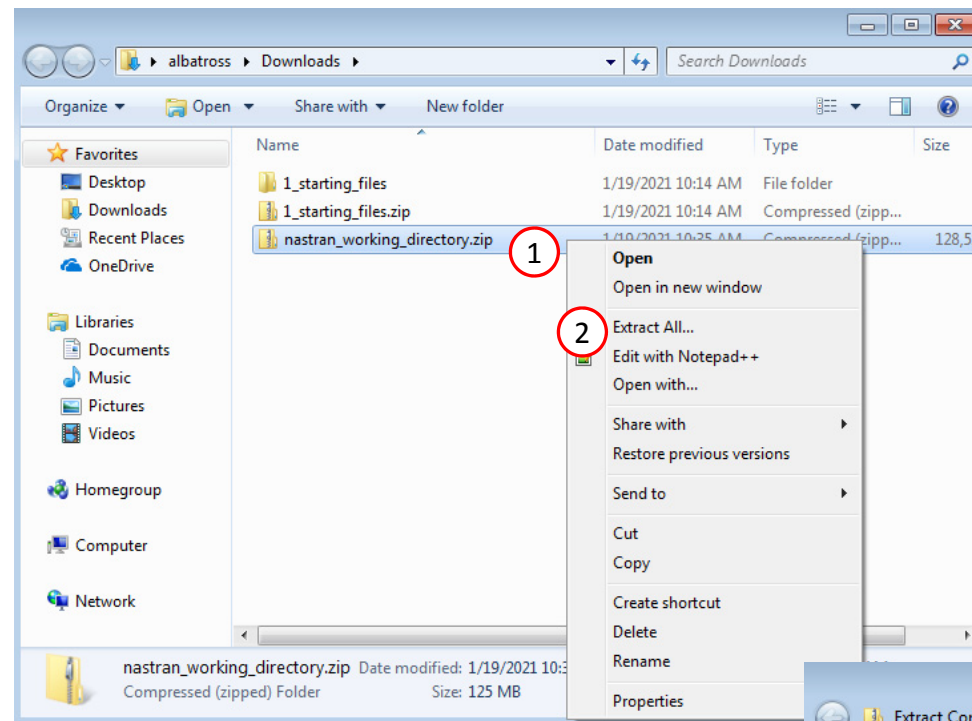
2



Start Desktop App

1. Right click on the indicated file
2. Click Extract All
3. Click Extract on the following window

- Always extract the contents of the ZIP file to a new, empty folder.



Start Desktop App

1. Open this folder:
nastran_working_directory
2. Inside of the new folder, double click on
Start Desktop App
3. Click Open, Run or Allow Access on any
subsequent windows
4. MSC Nastran will now start

- One can run the Nastran job on a remote machine as follows:
1) Copy the BDF files and the INCLUDE files to a remote machine. 2) Run the MSC Nastran job on the remote machine. 3) After completion, copy the BDF, F06, LOG, H5 files to the local machine. 4) Click "Start Desktop App" to display the results.

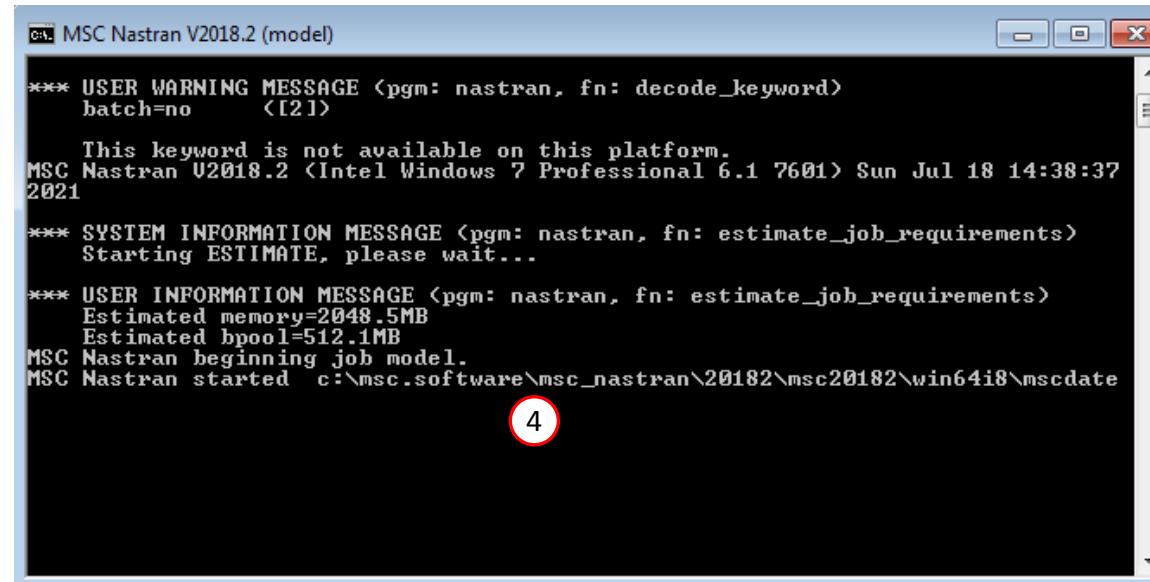
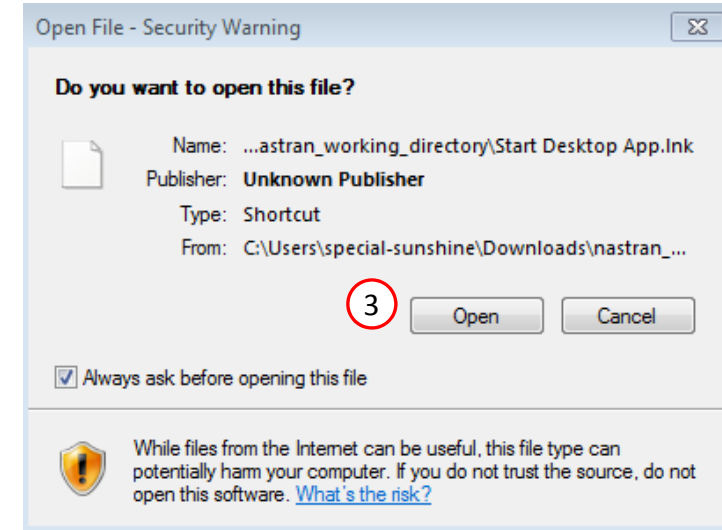
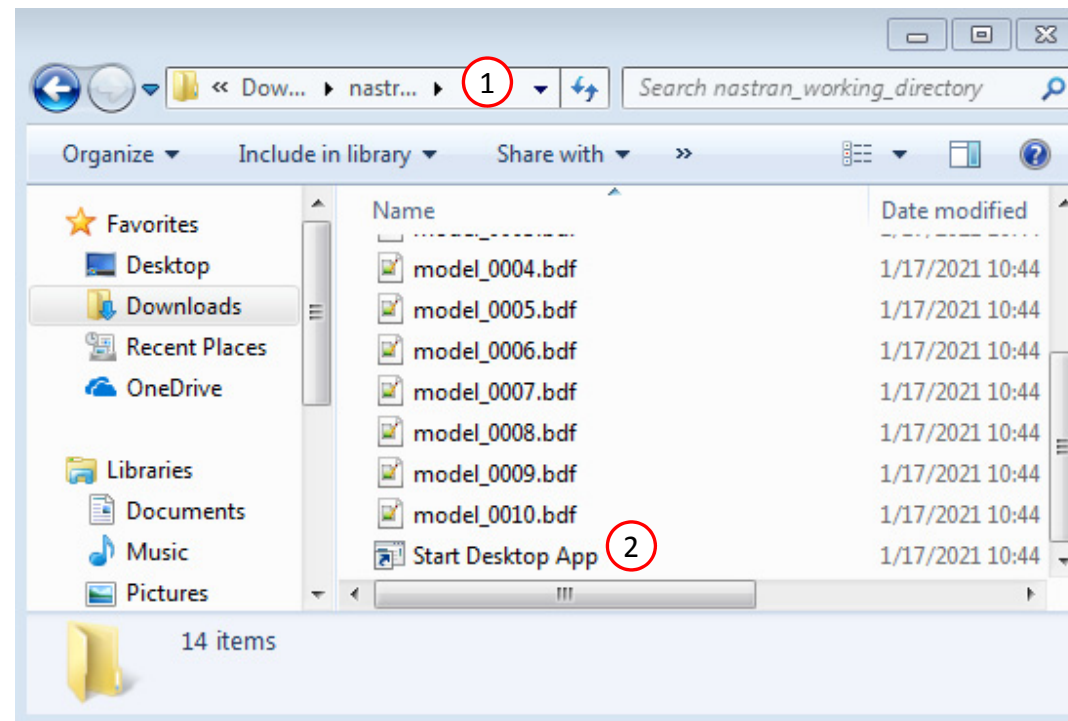
Using Linux?

Follow these instructions:

- 1) Open Terminal
- 2) Navigate to the nastran_working_directory
`cd ./nastran_working_directory`
- 3) Use this command to start the process
`./Start_MSC_Nastran.sh`

In some instances, execute permission must be granted to the directory. Use this command. This command assumes you are one folder level up.

```
sudo chmod -R u+x ./nastran_working_directory
```



Status

- While MSC Nastran is running, a status page will show the current state of MSC Nastran

SOL 200 Web App - Status

 Python

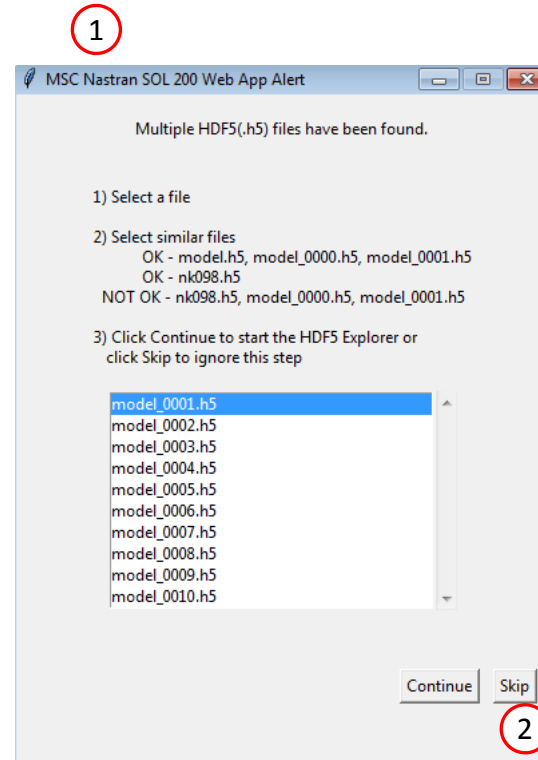
 MSC Nastran

Status

Name	Status of Job	Design Cycle	RUN TERMINATED DUE TO
model.bdf	Running	None	

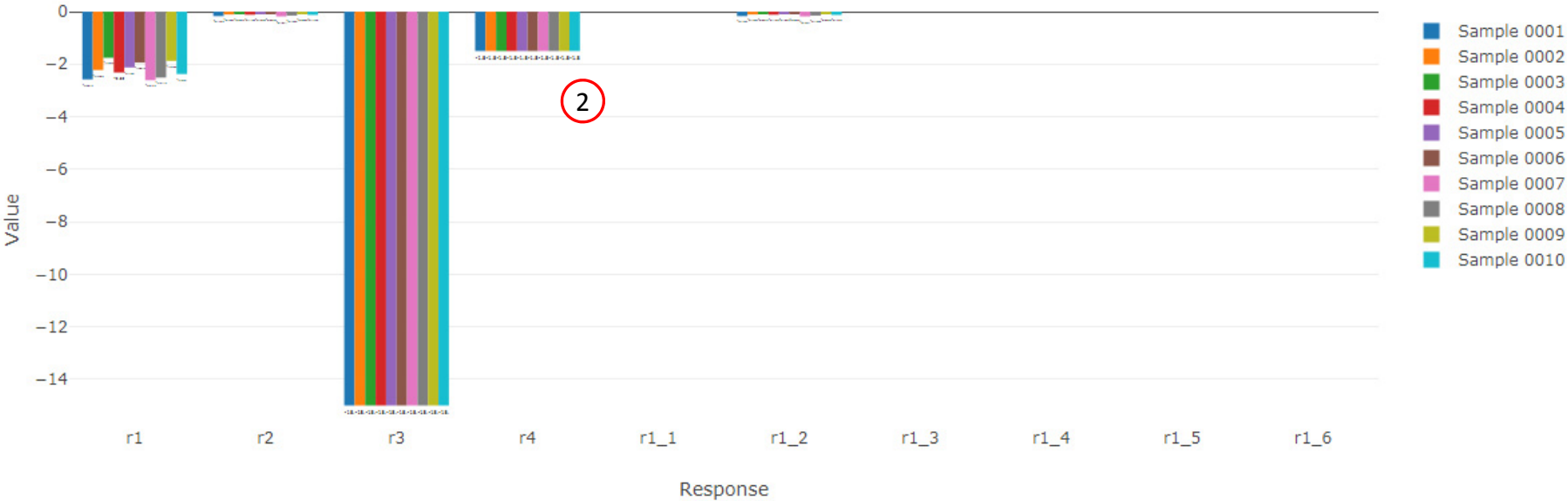
Review Results

1. A window appears asking to start the HDF5 Explorer
2. Click Skip to not open the HDF5 Explorer



Review Results

1. The Monitored Responses web app is opened
2. The value of each response and for each sample is displayed in a bar chart
3. A table lists the values for each response and sample.



Monitored Responses

Label	Dataset Name	Field	Field Description	Current Value	ID	SUBCASE	STEP
r1	NODAL/APPLIED_LOAD	Y	Y component	-0.15000749756256565	2	10	0
r2	NODAL/APPLIED_LOAD	Y	Y component	-0.2133144335494699	2	10	0
r3	NODAL/APPLIED_LOAD	Y	Y component	-1.1794164253032222	2	10	0
r4	NODAL/APPLIED_LOAD	Y	Y component	-1.5	2	10	0
r1_1	NODAL/APPLIED_LOAD	Y	Y component	-0.15000749756256565	2	10	0
r1_2	NODAL/APPLIED_LOAD	Y	Y component	-0.2133144335494699	2	10	0
r1_3	NODAL/APPLIED_LOAD	Y	Y component	-0.36332193111203553	2	10	0
r1_4	NODAL/APPLIED_LOAD	Y	Y component	-0.5133204286745012	2	10	0

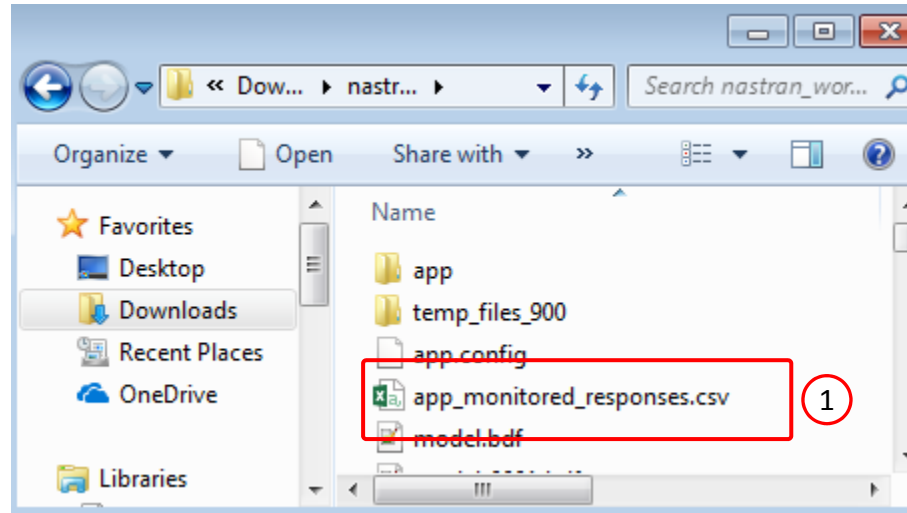
Monitored Responses from Each Sample

0001	0002
-2.5818811152279864	-2.224257497640053
-0.1705460155888461	-0.10869148213921452
-15.0	-15.0
-1.5	-1.5
None - The response could not be found in the H5 file(DOMAIN_ID)	None - The response could not be found in the H5 file(DOMAIN_ID)
-0.1705460155888461	-0.10869148213921452
None - The response could not be found in the H5 file(DOMAIN_ID)	None - The response could not be found in the H5 file(DOMAIN_ID)
None - The response could not be found in the H5 file(DOMAIN_ID)	None - The response could not be found in the H5 file(DOMAIN_ID)

Review Results

1. The monitored responses are contained in the CSV file named `app_monitored_responses.csv`

The responses in this CSV file will be used to train the surrogate model.



A screenshot of an Excel spreadsheet titled 'app_monitored_responses.csv - Excel'. The spreadsheet contains a table with columns labeled 'Sample', 'r1', 'r2', 'r3', 'r4', 'r1_1', 'r1_2', 'r1_3', 'r1_4', and 'r1_5'. The data rows show numerical values for 'r1' through 'r4' and 'None - Th' for 'r1_1' through 'r1_5'.

Sample	r1	r2	r3	r4	r1_1	r1_2	r1_3	r1_4	r1_5
1	-2.58188	-0.17055	-15	-1.5	None - Th	-0.17055	None - Th	None - Th	None - Th
2	-2.22426	-0.10869	-15	-1.5	None - Th	-0.10869	None - Th	None - Th	None - Th
3	-1.75794	-0.08219	-15	-1.5	None - Th	-0.08219	None - Th	None - Th	None - Th
4	-2.32003	-0.11861	-15	-1.5	None - Th	-0.11861	None - Th	None - Th	None - Th
5	-2.12808	-0.10092	-15	-1.5	None - Th	-0.10092	None - Th	None - Th	None - Th
6	-1.94152	-0.08986	-15	-1.5	None - Th	-0.08986	None - Th	None - Th	None - Th
7	-2.61253	-0.18206	-15	-1.5	None - Th	-0.18206	None - Th	None - Th	None - Th
8	-2.51154	-0.15059	-15	-1.5	None - Th	-0.15059	None - Th	None - Th	None - Th
9	-1.87684	-0.08688	-15	-1.5	None - Th	-0.08688	None - Th	None - Th	None - Th
10	-2.37612	-0.12584	-15	-1.5	None - Th	-0.12584	None - Th	None - Th	None - Th

Performing Predictions

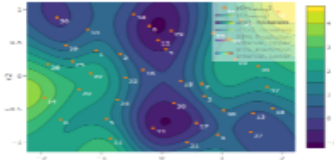
Prediction Analysis Web App

1. Return to the Machine Learning web app
2. Click Results
3. Click Prediction Analysis
4. The Prediction Analysis web app is now open
5. Ensure it says Connected

SOL 200 Web App - Machine Learning Parameters Samples Responses Download **Results** Connection Settings Home

1 2

Select a Results App

 3

Prediction Analysis

SOL 200 Web App - Prediction Analysis 4 Home

Gaussian Process (GP) App Connection Status

✓ Connected 5

Session ID: 8207

Output

```
GP App Update - Starting the Gaussian Process (GP) app on the server
- Session ID: 8207
- Address: http://localhost:8080/optimization
Desktop App Update - Connecting to the SOL 200 Web App...
GP App Update - Connection successful between the Node JS server and GP ap
```

Warnings and Errors

Warnings can be ignored

Training Data

1. Navigate to the Training and Testing Data section
2. Delete any previous table data by clicking the four (4) buttons named Delete all rows

- **x_training, y_training** - This specifies the x inputs and y outputs used to train the surrogate model.
- **x_testing, y_testing** - This specifies the x inputs and y outputs used to calculate the Normalized Root Mean Square Error (NRMSE) between the predicted values and actual MSC Nastran responses. This testing data is option.
- **x_prediction** – The x inputs at which to make predictions.

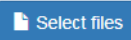
Training and Testing Data 1

x_training

CSV Export



CSV Import



Select a CSV File



CSV

imported

2

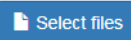
✕ Delete all rows

y_training

CSV Export



CSV Import



Select a CSV File



✕ Delete all rows

sample

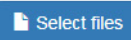
y1

x_testing

CSV Export



CSV Import



Select a CSV File



✕ Delete all rows

sample

x1

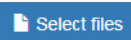
x2

y_testing

CSV Export



CSV Import



Select a CSV File



✕ Delete all rows

sample

y1

Training Data

1. Navigate to the section titled x_training
2. Click Select files
3. Navigate to the folder named nastran_working_directory which contains data for 10 runs
4. Select the file app.config
5. Click Open
6. Click Import
7. The table is now loaded with the x inputs for all 10 runs

Training and Testing Data

x_training 1

CSV Export

Export

CSV Import

Select files

app.config

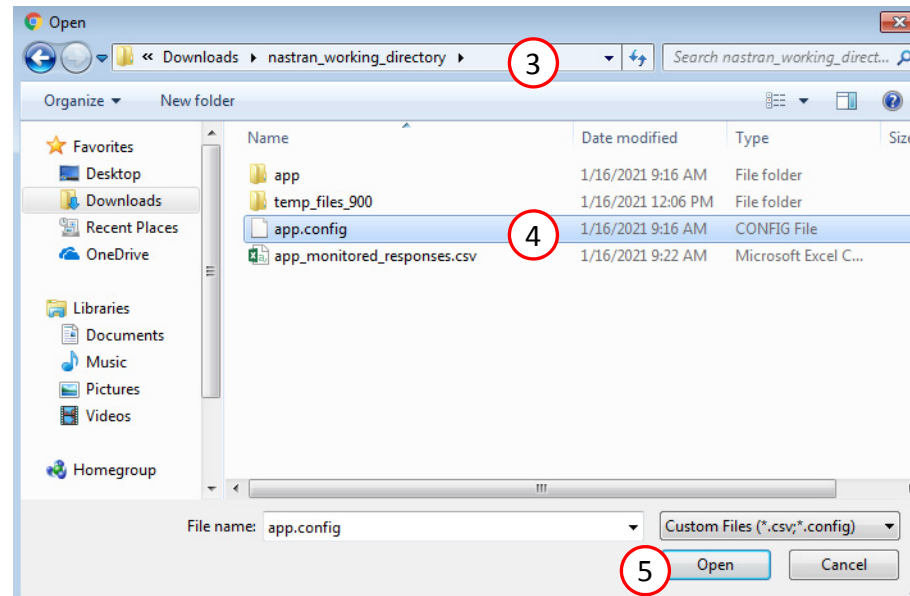
Import

✓ CSV
imported

Delete all rows

sample	x1
1	1.208613
2	5.371329
3	9.449995
4	4.353574
5	6.326213
6	8.000973
7	.8043
8	2.107346
9	8.531418
10	3.725496

10 25 50 100



Training Data

1. Navigate to the section titled y_training
2. Click Select files
3. Navigate to the folder named nastran_working_directory which contains data for 10 runs
4. Select the file app_monitored_responses.csv
5. Click Open
6. Click Import
7. The table is now loaded with the y outputs (monitored responses) for all 10 runs
8. Note that responses r1_1, r1_2, ... r3_24 contain characters instead of numbers. These characters will fail the regression and must be removed from the training data.

- **Why do responses r_1, r1_2, ..., r3_24 contain characters?** The original FE model uses an time step procedure that varies for each sample. For example, if you are monitoring the response at time step 1.0, sample 1 may have only time step 1.1 and sample 2 may have only time step .9. Since sample 1 and 2 do not have a response corresponding time step 1.0, the extracted response is reported as "None - The response could not be found in the H5." Training data should only contain number values. Non-number data should be removed from the training data.

y_training 1

CSV Export

Export

CSV Import

Select files

app_monitored_responses.csv

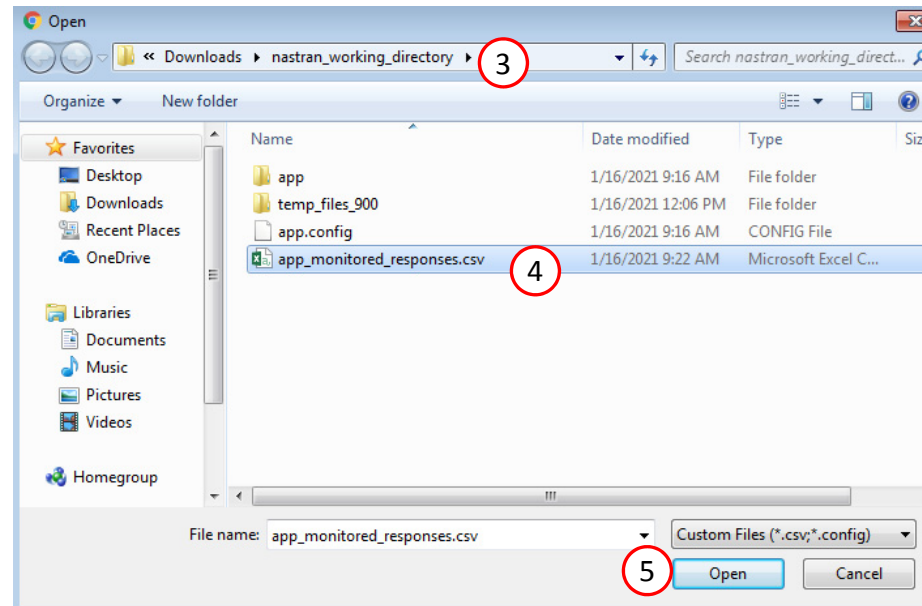
Import

CSV
imported

Delete all rows

sample	y1	y2	y3	y4	y1_
0001	-2.5818811152...	-0.1705460155...	-15.0	-1.5	None-The
0002	-2.2242574976...	-0.1086914821...	-15.0	-1.5	None-The
0003	-1.7579419057...	-0.0821920070...	-15.0	-1.5	None-The
0004	-2.3200339884...	-0.1186101317...	-15.0	-1.5	None-The
0005	-2.1280809974...	-0.1009203946...	-15.0	-1.5	None-The
0006	-1.9415194735...	-0.0898593593...	-15.0	-1.5	None-The
0007	-2.6125297609...	-0.1820618814...	-15.0	-1.5	None-The
0008	-2.5115376439...	-0.1505872377...	-15.0	-1.5	None-The
0009	-1.8768434132...	-0.0868802372...	-15.0	-1.5	None-The
0010	-2.3761151118...	-0.1258383787...	-15.0	-1.5	None-The

10 25 50 100



Training Data

1. For x_training click Export to download a CSV file
2. For y_training click Export to download a CSV file

x_training

CSV Export

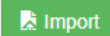


1

CSV Import



app.config



✓ CSV

imported

✕ Delete all rows

sample	x1
1	1.208613
2	5.371329
3	9.449995
4	4.353574
5	6.326213
6	8.000973
7	.8043
8	2.107346
9	8.531418
10	3.725496

10 25 50 100

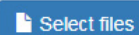
y_training

CSV Export

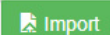


2

CSV Import



app_monitored_responses.cs



✓ CSV

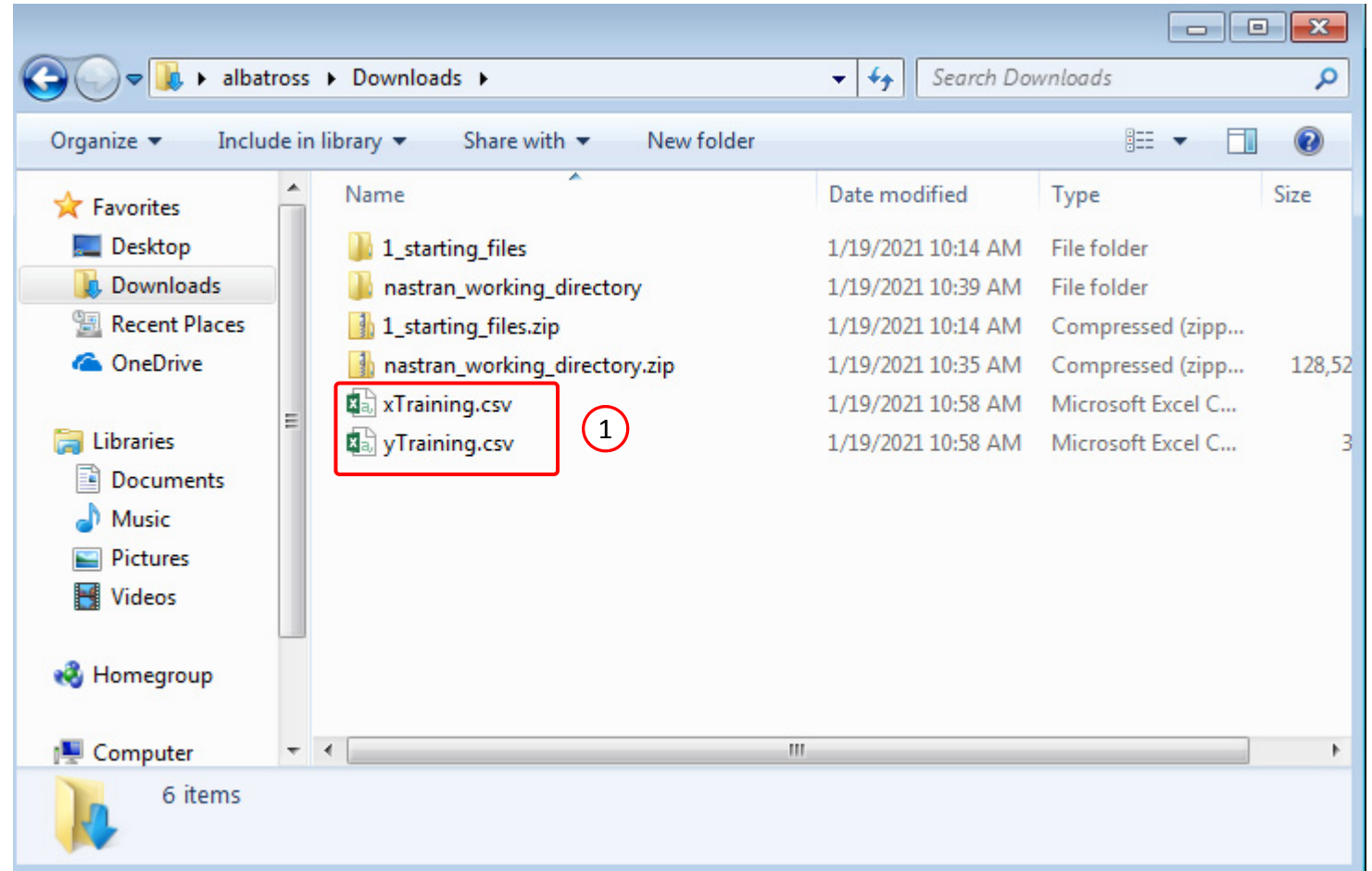
imported

✕ Delete all rows

sample	y1	y2	y3	y4	y1_
0001	-2.5818811152...	-0.1705460155...	-15.0	-1.5	None-The

Custom Training Data with Excel and CSV Files

1. CSV files have been download for the training data (x inputs and y outputs)



Custom Training Data with Excel and CSV Files

Two (2) artificial samples are added manually with Excel.

1. Open the file yTraining.csv in Excel
2. Delete the following response columns: r1_1, r1_2, ..., r3_24. The only column that should remain are: sample, y1, y2, y3, y4.
3. Add sample 11 with the following response values
 - y1: -100
 - Y2: -1
 - y3: -15
 - y4: -1.5
4. Add sample 12 with the following response values
 - y1: -90
 - Y2: -0.9
 - y3: -15
 - y4: -1.5

5. Click Save

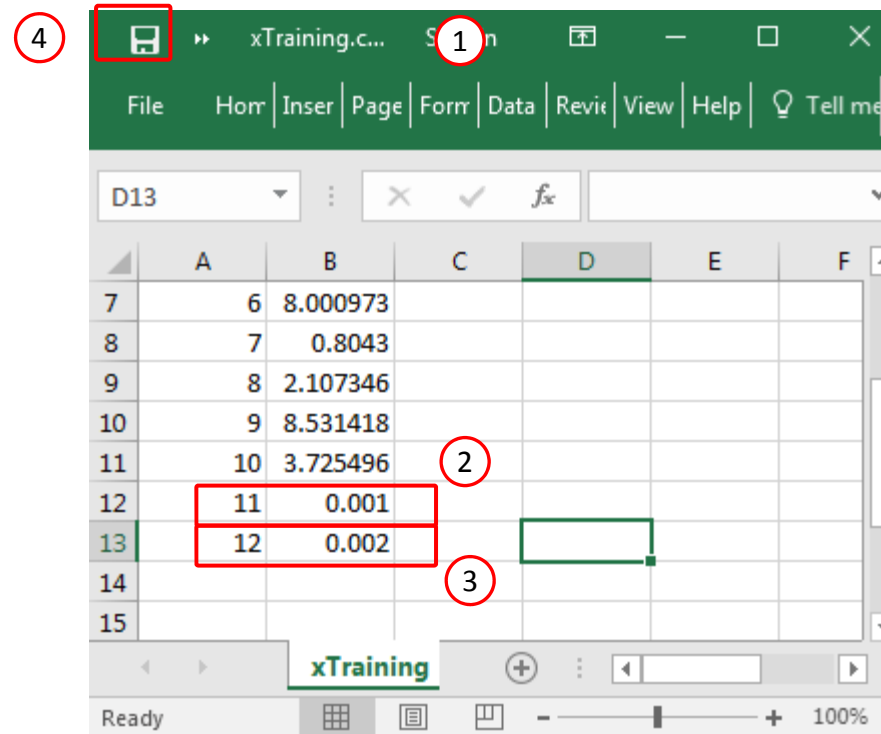
Excel window titled "yTraining.csv - Excel". The ribbon shows File, Home, Insert, Page Layout, Formulas, Data, Review, View, Help, and Tell me what you want to do. The formula bar shows "F1". The spreadsheet data is as follows:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	sample	y1	y2	y3	y4											
2	1	-2.58188	-0.17055	-15	-1.5											
3	2	-2.22426	-0.10869	-15	-1.5											
4	3	-1.75794	-0.08219	-15	-1.5											
5	4	-2.32003	-0.11861	-15	-1.5											
6	5	-2.12808	-0.10092	-15	-1.5											
7	6	-1.94152	-0.08986	-15	-1.5											
8	7	-2.61253	-0.18206	-15	-1.5											
9	8	-2.51154	-0.15059	-15	-1.5											
10	9	-1.87684	-0.08688	-15	-1.5											
11	10	-2.37612	-0.12584	-15	-1.5											
12	11	-100	-1	-15	-1.5											
13	12	-90	-0.9	-15	-1.5											
14																
15																
16																

The status bar shows "Ready" and "100%".

Custom Training Data with Excel and CSV Files

1. Open the file xTraining.csv in Excel
2. Add a new sample and x1 value
 - sample: 11
 - x1: .001
3. Add a new sample and x1 value
 - sample: 12
 - x1: .002
4. Click Save



Training Data

1. Navigate to the Training and Testing Data section
2. Delete any previous table data by clicking the four (4) buttons named Delete all rows

- **x_training, y_training** - This specifies the x inputs and y outputs used to train the surrogate model.
- **x_testing, y_testing** - This specifies the x inputs and y outputs used to calculate the Normalized Root Mean Square Error (NRMSE) between the predicted values and actual MSC Nastran responses. This testing data is optional.
- **x_prediction** – The x inputs at which to make predictions.

Training and Testing Data 1

x_training

CSV Export



CSV Import



Select a CSV File



CSV

imported

2

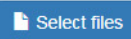
✕ Delete all rows

y_training

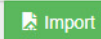
CSV Export



CSV Import



Select a CSV File



✕ Delete all rows

sample

y1

x_testing

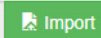
CSV Export



CSV Import



Select a CSV File



✕ Delete all rows

sample

x1

x2

y_testing

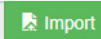
CSV Export



CSV Import



Select a CSV File



✕ Delete all rows

sample

y1

Training Data

1. Navigate to the section titled x_training
2. Click Select files
3. Navigate to the folder that contains the CSV files
4. Select the file xTraining.csv
5. Click Open
6. Click Import
7. The table is now loaded with the x inputs for 12 samples
8. Click Page 2
9. Ensure samples 11 and 12 are present

Training and Testing Data

x_training¹

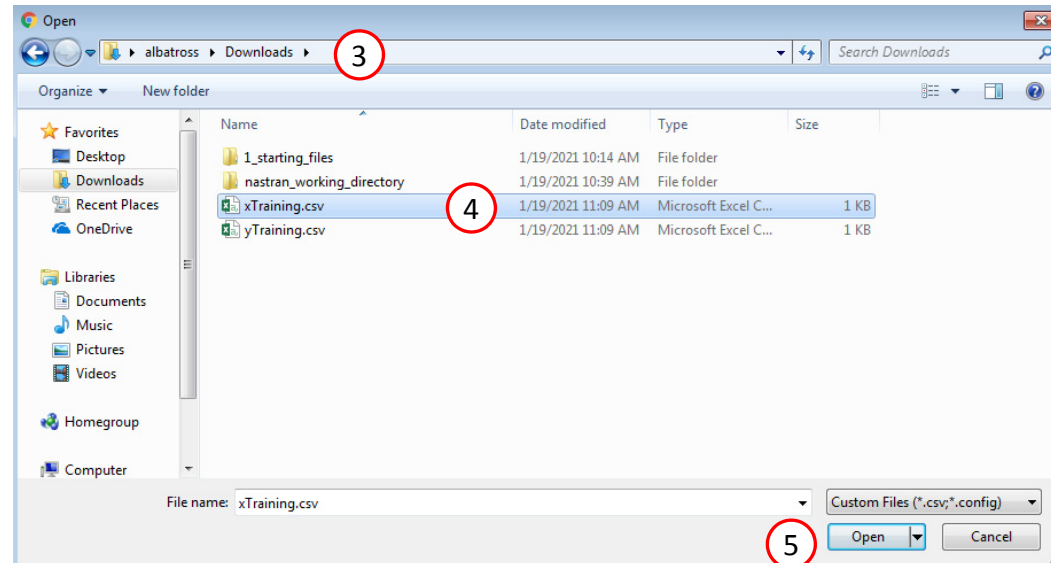
CSV Export CSV Import⁶

² Export Select files xTraining.csv Import⁶ CSV imported

⁷ Delete all rows

sample	x1
11	0.001
12	0.002

⁹ « 1 **2** » ⁸ 10 25 50 100



Training Data

1. Navigate to the section titled y_training
2. Click Select files
3. Navigate to the folder that contains the CSV files
4. Select the file yTraining.csv
5. Click Open
6. Click Import
7. The table is now loaded with the y outputs (monitored responses) for 12 samples
8. Click Page 2
9. Ensure samples 11 and 12 are present and responses y1, y2, y3 and y4 are present

y_training 1

CSV Export CSV Import 6

Export Select files yTraining.csv Import 2

CSV imported 7

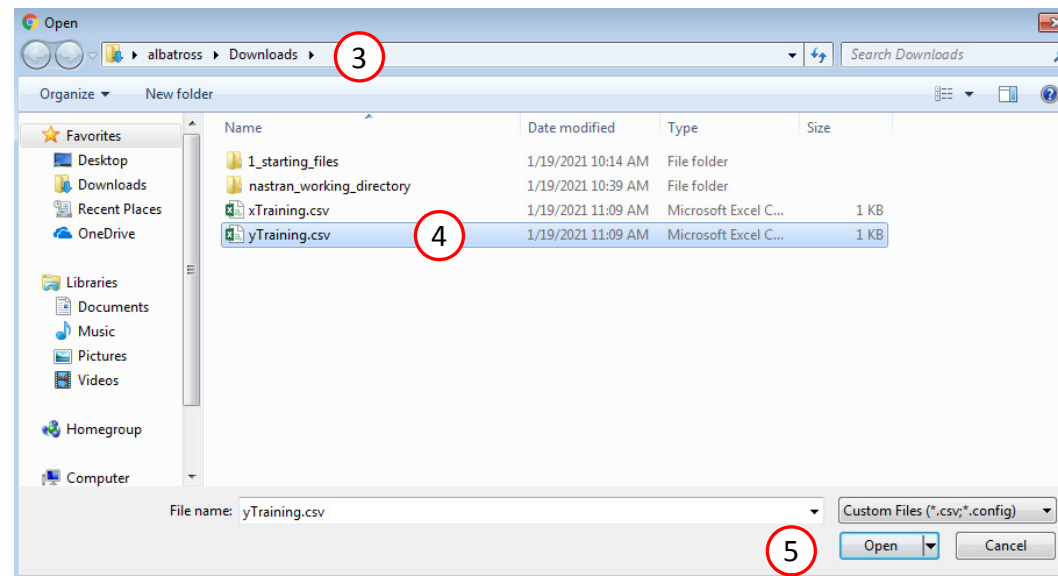
9

sample	y1	y2	y3	y4
11	-100	-1	-15	-1.5
12	-90	-0.9	-15	-1.5

« 1 2 » 10 25 50 100

8

Delete all rows




Perform Regression

1. Click Perform Regression and the surrogate model will be fitted
2. The regression is complete when the following status message is visible:
 - Process complete

Regression

Data	Link to Table	Status	Status Description
x_training	Link	✓	Ready
y_training	Link	✓	Ready
x_testing (Optional)	Link	!	Ready
y_testing (Optional)	Link	!	Ready

 Perform Regression

1

✓ Process complete

2

[Click here](#) to view the Regression Results section

Output

```
| y2 | 5.87183e+07 |
| y3 | 1 |
| y4 | 1 |

Parameters listed in decreasing order of relevance: x1
GP App Update - Sending initial data to the web browser
GP App Update - Sending complete
```

Warnings and Errors

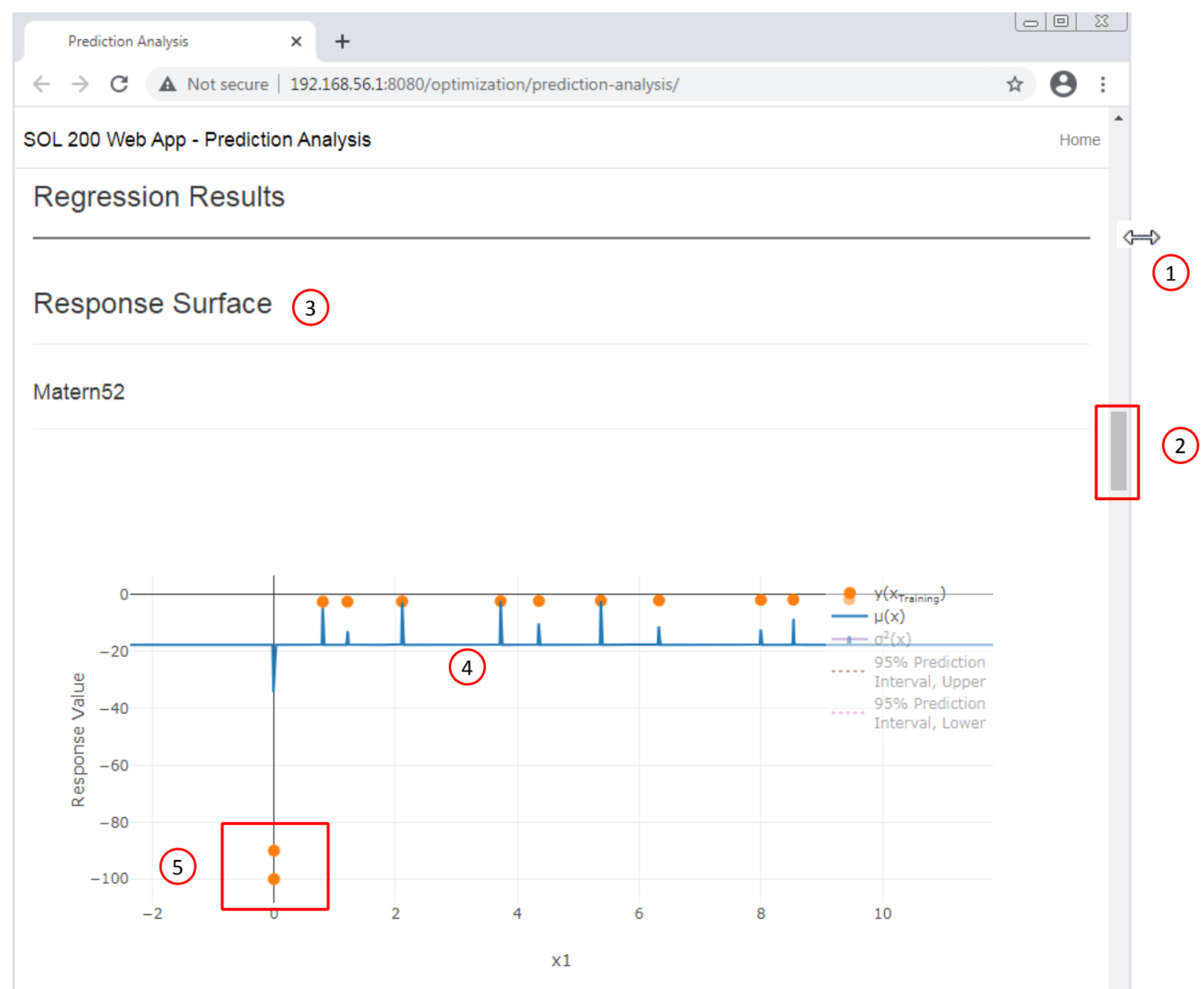
```
/home/apricot/PycharmProjects/python-app/venv/lib/python3.6/site-packages/GP
```

Warnings can be ignored

Response Surface

1. Resize the window to fit half the screen, this causes the plot to be increased in size
2. Adjust the scroll bar to make the plot visible
3. Navigate to the section titled Response Surface
4. The predicted function (surrogate model) appears “spiked”
5. The 2 artificial samples have response values of approximately -100 but the other samples 1-10 have response values between 1 and 3. The 2 artificial samples are causing a poor fit.

- Note that the surrogate model $\mu(x)$ goes through each training point, but the displayed plot does not show this. This is due to the resolution of the plot. 1,000 x points were used to generate the indicated plot. If 10,000 x points were used, the surrogate model $\mu(x)$ would better appear to go through each training point.



Spike Behavior

The “spike” behavior can occur in the following scenario.

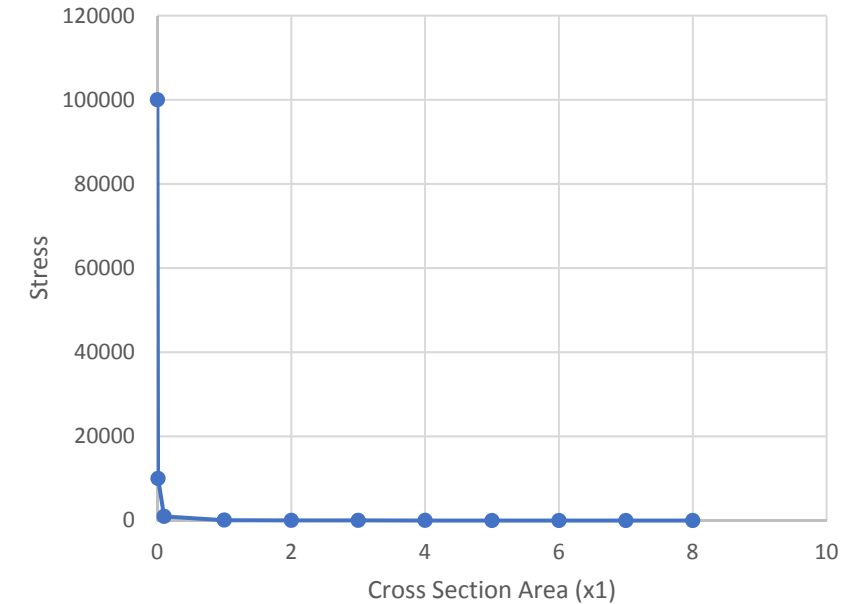
- Suppose there is a rod with cross section area A and is loaded axially with load P=100. The axial stress is simply P/A . As the area A goes to 0 and with a fixed load P, the stress goes to infinity. The order of magnitude of the stress values can drastically change for different values of A. Care should be taken to avoid samples with drastic variations in monitored responses. Responses that are significantly larger than other responses may result in a poor fit.

There are 2 actions that can be taken to address this issue:

- Adjust the bounds of the parameters to avoid a sample that produces drastically large response values.
- Manually remove samples with large response values. For example, in the table 1, samples 1 and 2 are candidates for removal.

Table 1

Sample	P	A	Stress
1	100	0.001	100000
2	100	0.01	10000
3	100	0.1	1000
4	100	1	100
5	100	2	50
6	100	3	33.33333333
7	100	4	25
8	100	5	20
9	100	6	16.66666667
10	100	7	14.28571429
11	100	8	12.5



1 Table 2: Example of bounds that are suitable and unsuitable

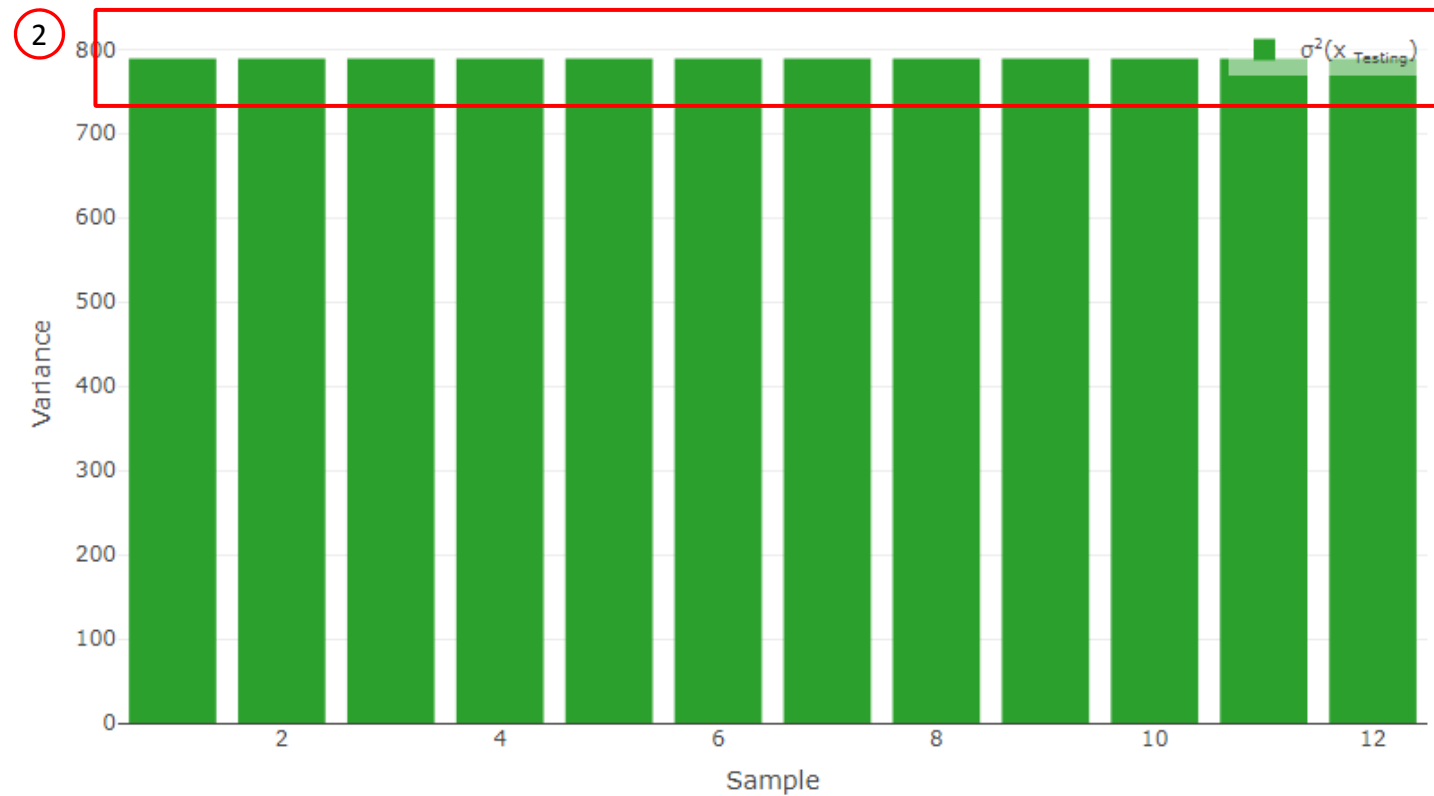
OK?	Parameter	Low	High
OK	x1	0.1	8.0
NOT OK	X1	.001	8.0

Variance

1

Matern52

NRMSE:



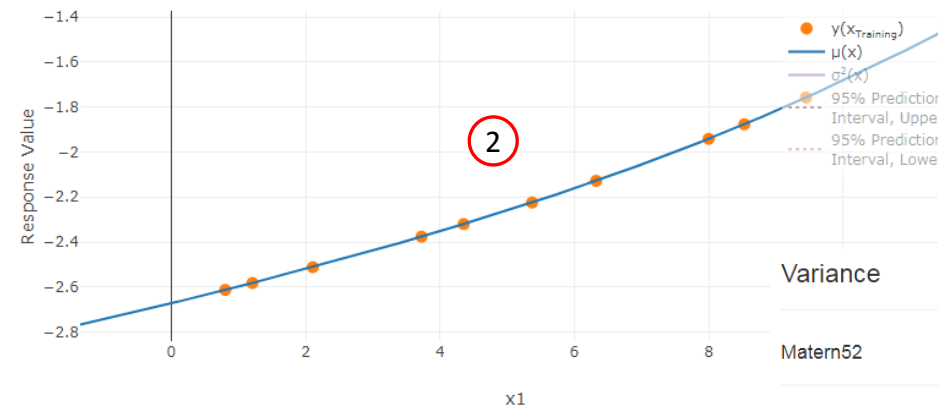
Variance

1. Navigate to the section titled Variance
2. The “spike” behavior can be inferred from the variance plot. Very tall and leveled bars is indication of the “spike” behavior.

- In this tutorial, variance (σ^2) is used to gauge the prediction uncertainty. Sometimes, you will see this prediction uncertainty expressed as the standard deviation (σ).

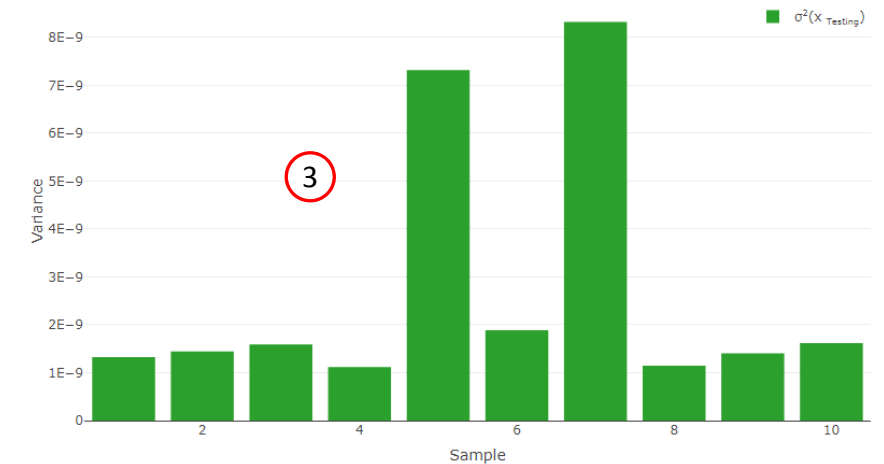
Response Surface

Matern52



Fix the Spike Behavior

1. Edit the CSV files (xTraining.csv and yTraining.csv) and remove the artificial samples 11 and 12. Reupload the CSV files to the Prediction Analysis web app and perform the regression again. (Not Shown)
2. Inspect the response surface plot. The “spike” behavior is no longer present.
3. Inspect the variance bar chart. Most bars take on unique heights and are not tall.



Creating Plots with the HDF5 Explorer

Start Desktop App

1. Open this folder:
nastran_working_directory
2. Inside of the new folder, double click on
Start Desktop App
3. Click Open, Run or Allow Access on any
subsequent windows

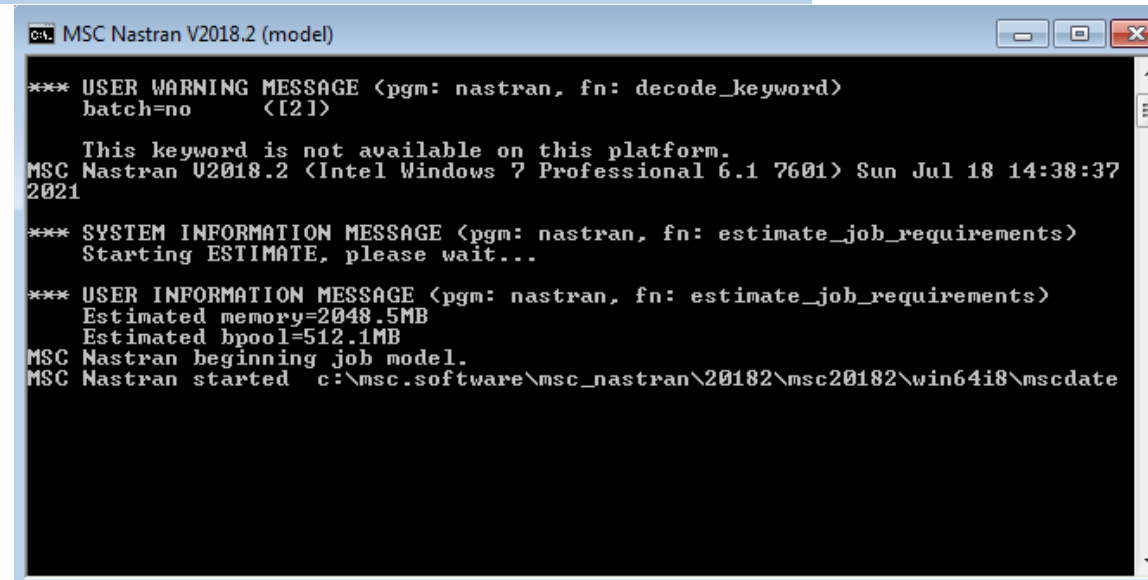
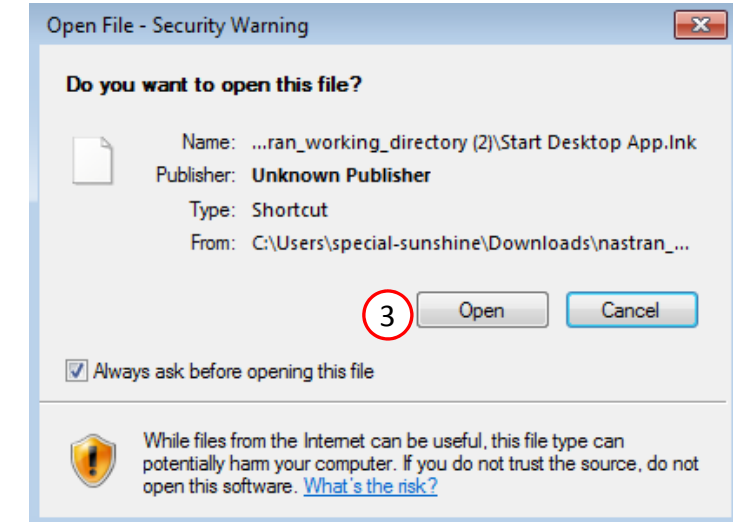
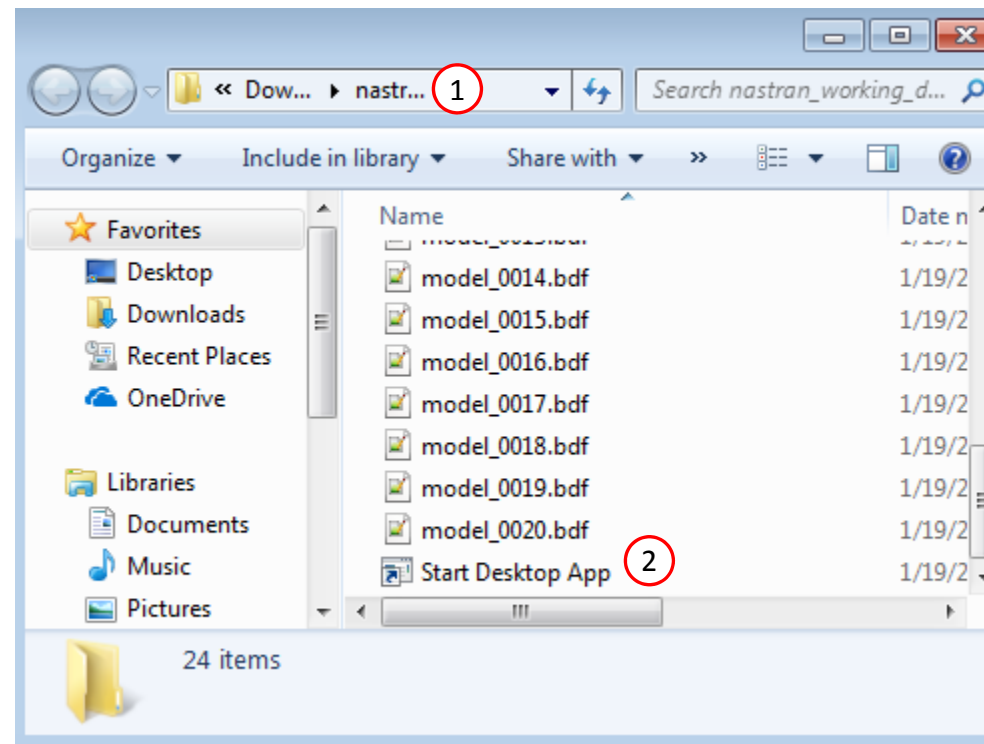
Using Linux?

Follow these instructions:

- 1) Open Terminal
- 2) Navigate to the nastran_working_directory
`cd ./nastran_working_directory`
- 3) Use this command to start the process
`./Start_MSC_Nastran.sh`

In some instances, execute permission must be granted to the directory. Use this command. This command assumes you are one folder level up.

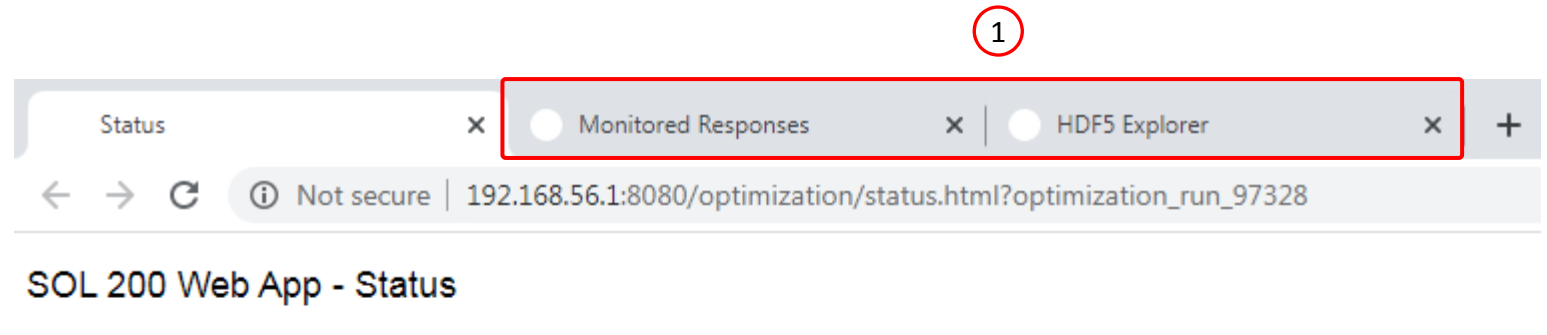
```
sudo chmod -R u+x ./nastran_working_directory
```



Results

Multiple web apps are automatically opened to display the results.

1. Use the tabs to switch between each web app
2. A description of each web app is given in the table.



2

Name of Web App	Purpose	Description
Monitored Responses	<ul style="list-style-type: none">• The response value from each sample can be compared.	<ul style="list-style-type: none">• After each MSC Nastran analysis, the response values are extracted from the H5 file and contained in a file named app_monitored_responses.csv. The Monitored Responses web app is used to create a bar chart of the values contained in this CSV file.
HDF5 Explorer	<ul style="list-style-type: none">• This web app is used to probe each H5 file and generate XY plots.	

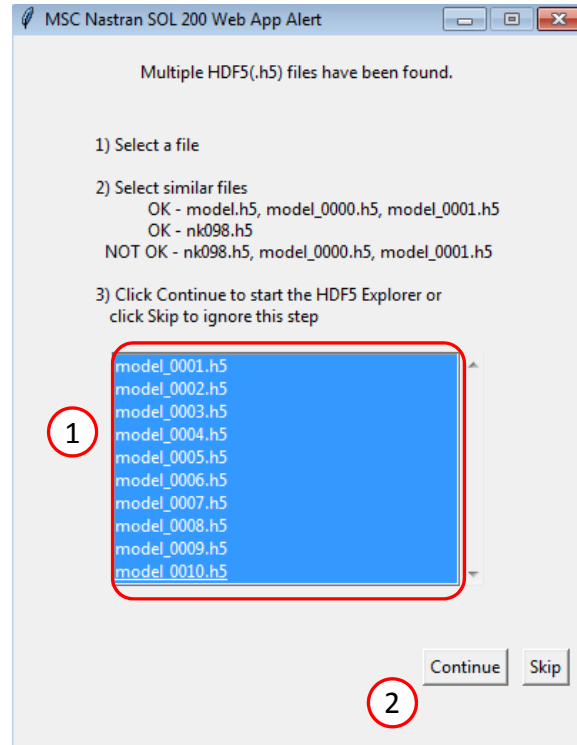
Review Results

On this page, the HDF5 Explorer is opened.

1. Select all 10 H5 files
2. Click Continue
3. The HDF5 Explorer is automatically opened.

The HDF5 Explorer is broken into sections.

- Acquire Dataset – Specific datasets from the H5 file can be extracted in this section.
- Plots Browser – Use this section to navigate every plot created.
- Combine Plots – This section allows you to combine multiple plots. For example, you can create Load vs. Displacement plots in this section.
- Last Plot Added – This display the last plot that was created.



SOL 200 Web App - HDF5 Explorer

3

Acquire Dataset

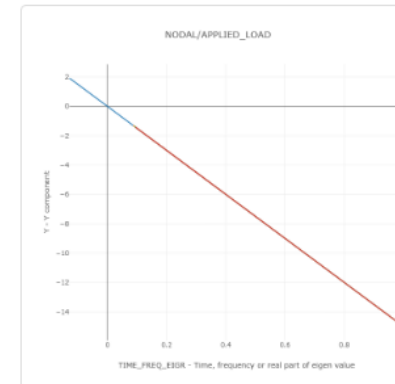
Plots Browser

Combine Plots

Last Plot Added

Plots Browser

NODAL/APPLIED_LOAD



Plot #: 0 - ID: 2 | SAMPLE: model_0000, model_0001, model_0002, ... | SUBCASE: 10 | Y vs. TIME_FREQ_EIGR

Review Results

1. Click Combine Plots
2. For Dataset A, Plot Number, select 0
3. For Dataset B, Plot Number, select 1
4. Click Update Plot
5. Mark both checkboxes
6. Click Display None
7. Mark the checkbox for the first 4 samples
8. The corresponding load vs. displacement plot is created

Combine Plots

Dataset A

Plot Number

0

Dataset	NODAL/APPLIED_LOAD
Field, Vertical Axis	Y
Field, Horizontal Axis	TIME_FREQ_EIGR

Dataset B

Plot Number

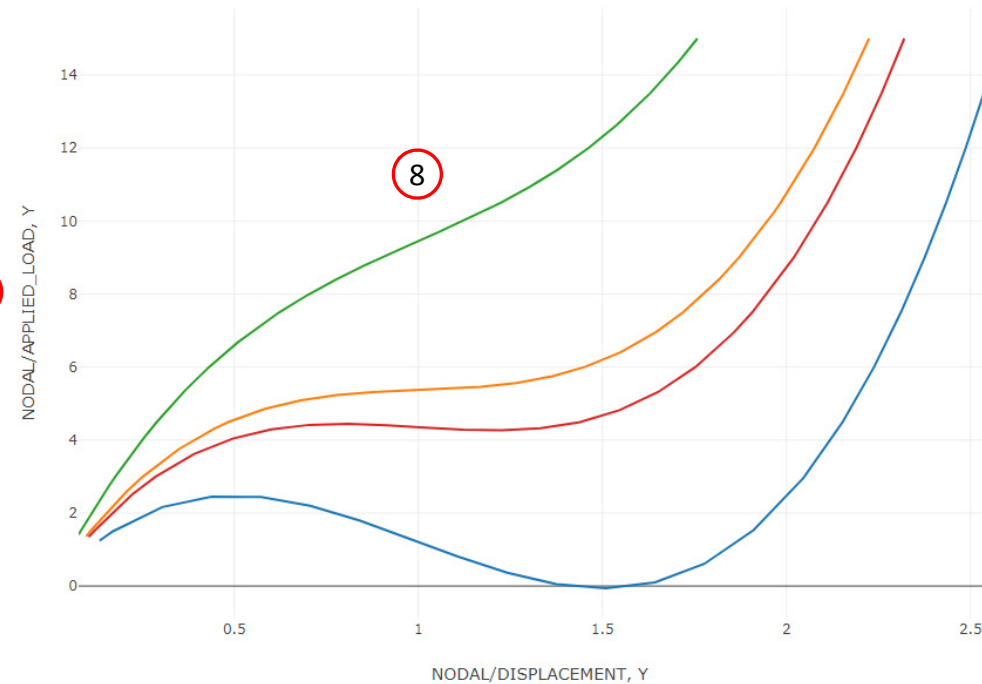
1

Dataset	NODAL/DISPLACEMENT
Field, Vertical Axis	Y
Field, Horizontal Axis	TIME_FREQ_EIGR

Update Plot

- ☒ Reverse sign of horizontal values
- ☒ Reverse sign of vertical values

NODAL/APPLIED_LOAD vs. NODAL/DISPLACEMENT



☐ Display None ☒ Display All

Display	Color	Name
<input checked="" type="checkbox"/>	Blue	0 - Dataset A - ID: 2 SAMPLE: model_0001 SUBCASE: 10 Dataset B - ID: 2 SAMPLE: model_0001 SUBCASE: 10
<input checked="" type="checkbox"/>	Orange	1 - Dataset A - ID: 2 SAMPLE: model_0002 SUBCASE: 10 Dataset B - ID: 2 SAMPLE: model_0002 SUBCASE: 10
<input checked="" type="checkbox"/>	Green	2 - Dataset A - ID: 2 SAMPLE: model_0003 SUBCASE: 10 Dataset B - ID: 2 SAMPLE: model_0003 SUBCASE: 10
<input checked="" type="checkbox"/>	Red	3 - Dataset A - ID: 2 SAMPLE: model_0004 SUBCASE: 10 Dataset B - ID: 2 SAMPLE: model_0004 SUBCASE: 10
<input type="checkbox"/>	Purple	4 - Dataset A - ID: 2 SAMPLE: model_0005 SUBCASE: 10 Dataset B - ID: 2 SAMPLE: model_0005 SUBCASE: 10
<input type="checkbox"/>	Brown	5 - Dataset A - ID: 2 SAMPLE: model_0006 SUBCASE: 10 Dataset B - ID: 2 SAMPLE: model_0006 SUBCASE: 10
<input type="checkbox"/>	Pink	6 - Dataset A - ID: 2 SAMPLE: model_0007 SUBCASE: 10 Dataset B - ID: 2 SAMPLE: model_0007 SUBCASE: 10
<input type="checkbox"/>	Grey	7 - Dataset A - ID: 2 SAMPLE: model_0008 SUBCASE: 10 Dataset B - ID: 2 SAMPLE: model_0008 SUBCASE: 10
<input type="checkbox"/>	Yellow	8 - Dataset A - ID: 2 SAMPLE: model_0009 SUBCASE: 10 Dataset B - ID: 2 SAMPLE: model_0009 SUBCASE: 10
<input type="checkbox"/>	Cyan	9 - Dataset A - ID: 2 SAMPLE: model_0010 SUBCASE: 10 Dataset B - ID: 2 SAMPLE: model_0010 SUBCASE: 10

End of Tutorial

Appendix

Appendix Contents

- How to import and edit previous files
- What is Gaussian Process Regression?

How to import and edit previous files

How to import and edit previous files

The parameters, samples and responses are contained in the following files

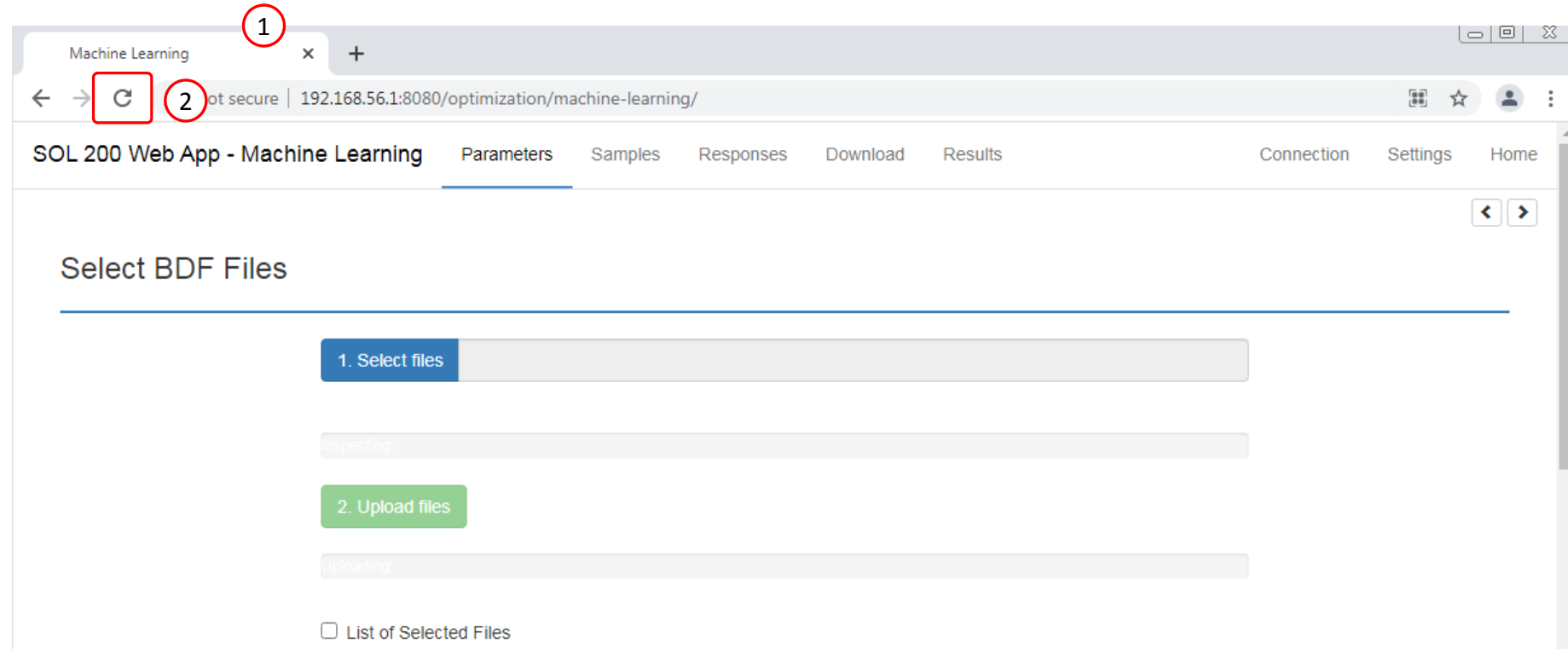
- app.config
- BDF files

These files may be imported back to the Parameter Study web app, and any parameters, samples and responses can be reconfigured

Import

1. Return to the window or tab that has the Parameter Study web app opened
2. Refresh the web page to start a new session

- Refreshing the page is only required when the *Select files* button is disabled.



Import

1. Click Select Files
2. Navigate to the folder named nastran_working_directory
3. Select all the BDF files AND the app.config file.
4. Click Open
5. Click Upload files

- All imports require the app.config file to be selected.

Select BDF Files

1

1. Select files

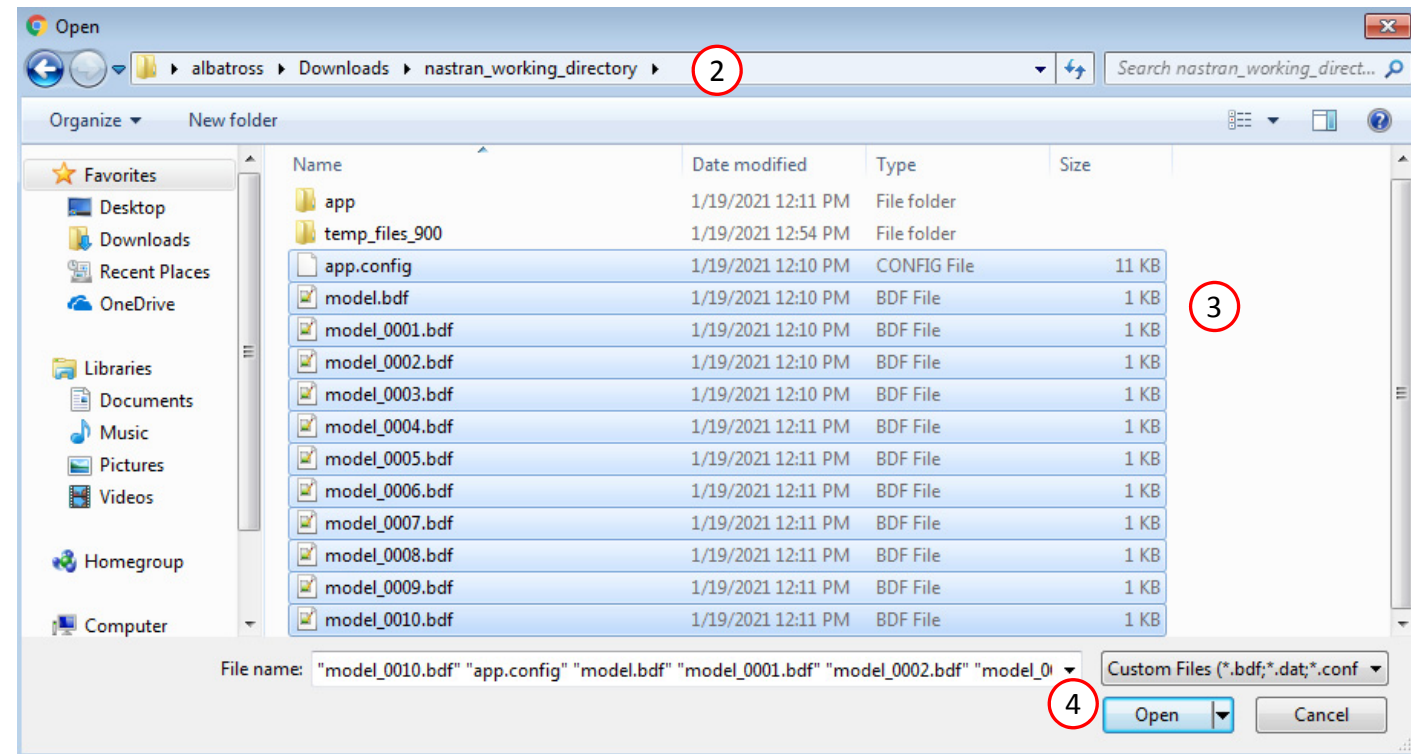
2 files selected

Inspecting: 100%

5

2. Upload files

Uploading:



Import

For the Response section, the H5 file will need to be re-uploaded.

1. Click Responses
2. Select the H5 file
3. Click Upload
4. Data from the H5 is loaded and ready to use

SOL 200 Web App - Machine Learning

Parameters

Samples

Responses

Download

Results

1

Upload .h5 File

2

1. Select files

nas103_ws09_with_spring_support.h5

3

2. Upload files

Uploading

Uploading

Select Responses to Monitor

Session ID: 5628

HDF5

View Responses to Monitor

Monitored Responses

Hide/Show Columns

Reset Filters

Download CSV

Delete	Label	Status	Objective	Lower Bound	Upper Bound	Monitor the response of the FINAL design cycle (SOL 200 on)
	r1 r2 r3 r4					
✕	r1	✓		Lower	Upper	
✕	r2	✓		Lower	Upper	

SOL 200 Web App - Machine Learning

Parameters

Samples

Responses

Download

Results

Session ID: 3736

HDF5

Select Responses to Monitor

Select Dataset

ELEMENTAL/ELEMENT_FORCE

ELEMENTAL/ELEMENT_FORCE

NODAL/APPLIED_LOAD

NODAL/DISPLACEMENT

Specify Entities

20

Element identification number (EID)

Examples: 20, etc.

Auto Execute

Acquire Dataset

Acquisition complete and successful

Acquired Dataset

Reset Filters

ELEMENTAL/ELEMENT_FORCE/ELAS1 - 20

EID	F	SAMPLE	DOMAIN_ID	SUBC.
Element identification number	User defined	Name of H5 File**	Domain Identifier	Subc. num1
20	4	nas103		10
20	-0.00015000...	nas103_...	1	10
20	-0.00021331...	nas103_...	2	10
20	-0.00036332...	nas103_...	3	10
20	-0.00051332...	nas103_...	4	10
20	-0.00066333...	nas103_...	5	10

View Responses to Monitor

Monitored Responses

Hide/Show Columns

Reset Filters

Download CSV

Delete	Label	Status	Objective	Lower Bound	Upper Bound	Monitor the response of the FINAL design cycle (SOL 200 on)
	r1 r2 r3 r4					
✕	r1	✓		Lower	Upper	
✕	r2	✓		Lower	Upper	
✕	r3	✓		Lower	Upper	
✕	r4	✓		Lower	Upper	

5 10 20 30 50 100

Import

After import, any Parameter, Samples or Responses can be modified.



Select BDF Files

1. Select files

2 files selected

Inspecting: 100%

2. Upload files

Uploading: 100 %

☐ List of Selected Files

Select Parameters

\$ _1 _ || _2 _ || _3 _ || _4 _ || _5 _ || _6 _ || _7 _ || _8 _ || _9 _ || _10 _ |

FORCE	15	2		15.	0.	-1.	0.			
MAT1	1		10.E7							
NLPCI	20		CRIS	1.	1.				40	
PELAS	20		%x1%							
PROD	10	1		.1						

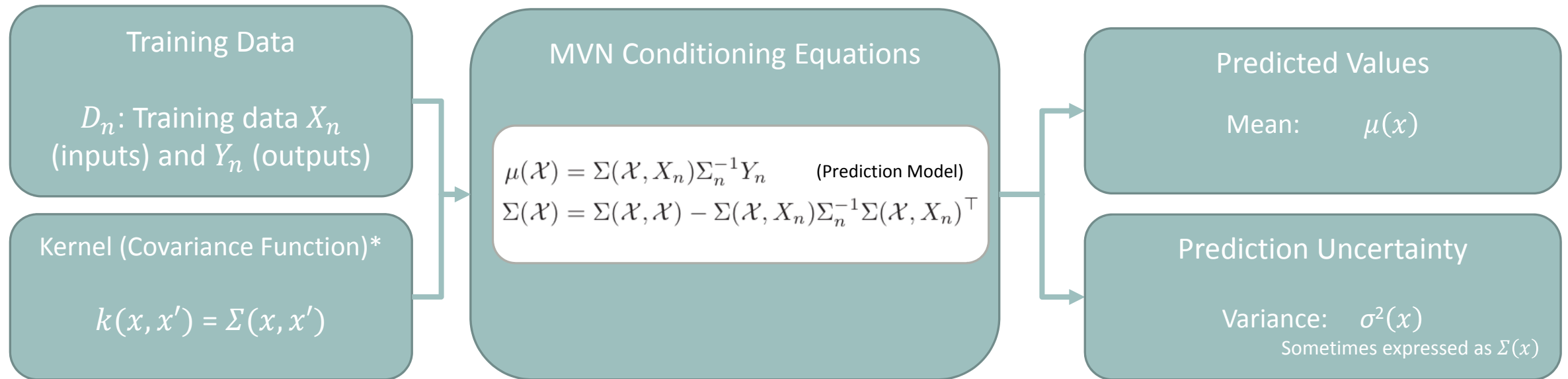


Configure Parameters

Delete	Parameter	Status	Low	High	Comments
	x1		.1	10.	Field 3 of PELA

What is Gaussian Process Regression?

Gaussian Process Regression Overview



* Hyperparameter optimization is part of the procedure but not covered in this presentation

** $\mu(x)$: This function corresponds to the mean function or kriging model. This function is the prediction model, also known as the surrogate model, meta model or emulator.

Multivariate Normal (MVN) Conditioning Equations

The following must be calculated: Covariance Matrix, Mean and Variance

Covariance Matrix

$$\Sigma = \begin{pmatrix} \Sigma(\chi, \chi) & \Sigma(\chi, X_n) \\ \Sigma(X_n, \chi) & \Sigma_n = \Sigma(X_n, X_n) \end{pmatrix}$$

X_n : Training locations
 χ : Testing (predictive) locations

Apply the covariance function $\Sigma(x, x')$ (kernel $k(x, x')$)

- $\Sigma(\chi, \chi)$: Covariance between testing (predictive) locations and themselves
- $\Sigma(\chi, X_n)$: Covariance between testing (predictive) and training locations
- $\Sigma(X_n, \chi)$: Covariance between training and testing (predictive) locations, which is the transpose of $\Sigma(\chi, X_n)$
- $\Sigma_n = \Sigma(X_n, X_n)$: Covariance between training locations and themselves

MVN Conditioning Equations (Mean and Variance)

Also referred to as “Gaussian process regression,” “kriging” or “kriging equations”

mean $\mu(\chi) = \Sigma(\chi, X_n) \Sigma_n^{-1} Y_n$ Prediction Model (Vary χ to make predictions)

and variance $\Sigma(\chi) = \Sigma(\chi, \chi) - \Sigma(\chi, X_n) \Sigma_n^{-1} \Sigma(\chi, X_n)^\top$ Prediction Uncertainty

Example 1

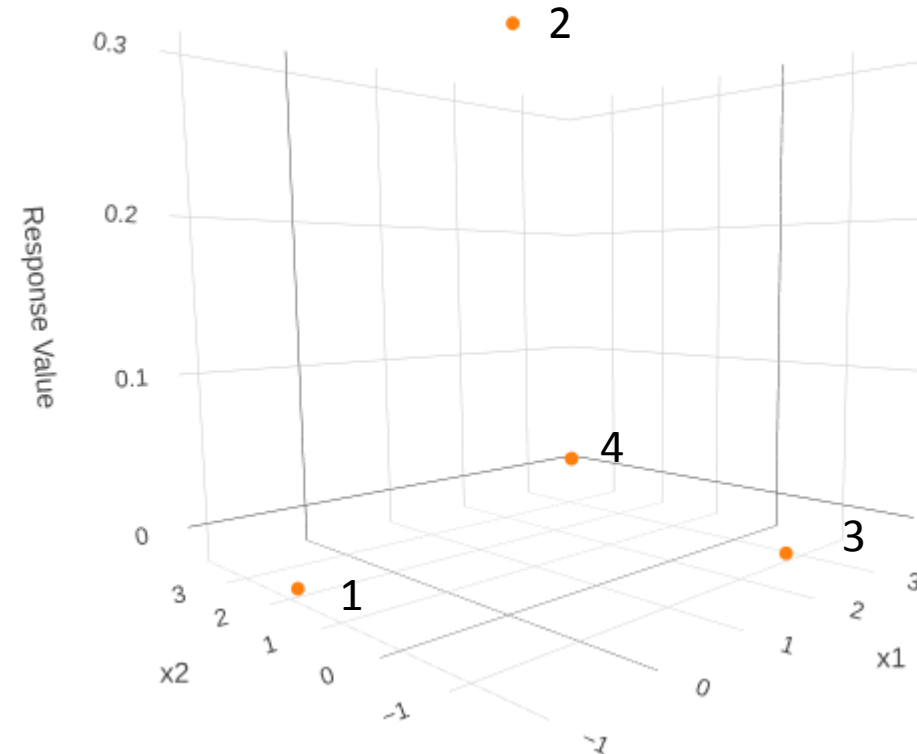
Example 1

Suppose a black box function was executed at 4 different samples (x1, x2 combinations)

With limited data (x and y), what does the response surface look like?

Training Data

Sample	x1	x2	y
1	-1.03	1.76	-1.56E-02
2	.49	.49	3.04E-01
3	1.77	-1.77	3.38E-03
4	3.62	3.76	5.43E-12



Training Data and Testing (Predictive) Locations

Suppose you have the following training data (X_n and Y_n) and testing locations (χ)

- X_n : The training design consists of 4 points
- χ : The test design (locations to make predictions) consists of 2 points

$$X = \begin{bmatrix} \chi \\ X_n \end{bmatrix} = \begin{bmatrix} .35 & .69 \\ .65 & .46 \\ -1.03 & 1.76 \\ .49 & .49 \\ 1.77 & -1.77 \\ 3.62 & 3.76 \end{bmatrix}$$

$$\begin{bmatrix} y^* \\ Y_n \end{bmatrix} = \begin{bmatrix} ? \\ ? \\ -1.56e-02 \\ 3.04e-01 \\ 3.38e-03 \\ 5.43e-12 \end{bmatrix}$$

The goal is make predictions (y^*) for points in χ

Note

- X_n : inputs of the training data
- Y_n : outputs of the training data
- χ or x : inputs of the testing data (predictive locations, i.e. points to make predictions)
- y^* : predicted outputs
- D_n : Training data X_n and Y_n

X : upper case of Greek letter chi (pronounced kai in English)
 χ : lower case of Greek letter chi

Calculation of the Covariance Matrix

1. Select a covariance (kernel) function

- Many covariance functions (kernels) exist: Radial Basis Function (RBF), Matern 5/2, 3/2, Exponential, ...
- For this example, a form of the RBF covariance function is used. This covariance function is described as the “inverse exponentiated squared Euclidean distance”

$$k(x, x') = \Sigma(x, x') = \exp\{-\|x - x'\|^2\} = e^{-\|x - x'\|^2}$$

2. Calculate D (Distance Matrix)

$$D = \|X - X\|^2 \quad \text{“Norm between } X \text{ and } X, \text{ squared”}$$

3. Calculate Σ (Covariance Matrix)

$$\Sigma = e^{-D}$$

Calculation of D

$D =$

$\sqrt{(.35 - .35)^2 + (.69 - .69)^2}$ = 0	$\sqrt{(.35 - .65)^2 + (.69 - .46)^2}$ = .1429	$\sqrt{(.35 - -1.03)^2 + (.69 - 1.76)^2}$ = 3.0493	$\sqrt{(.35 - .49)^2 + (.69 - .49)^2}$ = .0596	$\sqrt{(.35 - 1.77)^2 + (.69 - -1.77)^2}$ = 8.068	$\sqrt{(.35 - 3.62)^2 + (.69 - 3.76)^2}$ = 20.1178
.1429	$\sqrt{(.65 - .65)^2 + (.46 - .46)^2}$ = 0	$\sqrt{(.65 - -1.03)^2 + (.46 - 1.76)^2}$ = 4.5124	$\sqrt{(.65 - .49)^2 + (.46 - .49)^2}$ = .0265	$\sqrt{(.65 - 1.77)^2 + (.46 - -1.77)^2}$ = 6.2273	$\sqrt{(.65 - 3.62)^2 + (.46 - 3.76)^2}$ = 19.7109
3.0493	4.5124	$\sqrt{(-1.03 - -1.03)^2 + (1.76 - 1.76)^2}$ = 0	$\sqrt{(-1.03 - .49)^2 + (1.76 - .49)^2}$ = 3.9233	$\sqrt{(-1.03 - 1.77)^2 + (1.76 - -1.77)^2}$ = 20.3009	$\sqrt{(-1.03 - 3.62)^2 + (1.76 - 3.76)^2}$ = 25.6225
.0596	.0265	3.9233	$\sqrt{(.49 - .49)^2 + (.49 - .49)^2}$ = 0	$\sqrt{(.49 - 1.77)^2 + (.49 - -1.77)^2}$ = 6.746	$\sqrt{(.49 - 3.62)^2 + (.49 - 3.76)^2}$ = 20.4898
8.068	6.2273	20.3009	6.746	$\sqrt{(1.77 - 1.77)^2 + (-1.77 - -1.77)^2}$ = 0	$\sqrt{(1.77 - 3.62)^2 + (-1.77 - 3.76)^2}$ = 34.0034
20.1178	19.7109	25.6225	20.4898	34.0034	$\sqrt{(3.62 - 3.62)^2 + (3.76 - 3.76)^2}$ = 0

Calculation of Σ

$$\Sigma = \begin{bmatrix} e^0 = 1 & e^{-.1429} = .8668 & e^{-3.0493} = .0474 & e^{-.0596} = .9421 & e^{-8.068} = .0003 & e^{-20.1178} = 1.832e-9 \\ .8668 & e^0 = 1 & e^{-4.5124} = .0110 & e^{-.0265} = .9738 & e^{-6.2273} = .0020 & e^{-19.7109} = 2.8e-9 \\ .0474 & .0110 & e^0 = 1 & e^{-3.9233} = .0198 & e^{-20.3009} = 1.5e-9 & e^{-25.6225} = 7.5e-12 \\ .9421 & .9738 & .0198 & e^0 = 1 & e^{-6.746} = .0012 & e^{-20.4898} = 1.263e-9 \\ .0003 & .0020 & 1.5e-9 & .0012 & e^0 = 1 & e^{-34.0034} = 1.7e-15 \\ 1.832e-9 & 2.8e-9 & 7.5e-12 & 1.263e-9 & 1.7e-15 & e^0 = 1 \end{bmatrix}$$

Calculation of Σ

$\Sigma =$

$$\Sigma(\chi, \chi)$$

$e^0 = 1$	$e^{-1.429} = .8668$
.8668	$e^0 = 1$

$$\Sigma(\chi, X_n)$$

$e^{-3.0493} = .0474$	$e^{-.0596} = .9421$	$e^{-8.068} = .0003$	$e^{-20.1178} = 1.832e-9$
$e^{-4.5124} = .0110$	$e^{-0.265} = .9738$	$e^{-6.7273} = .0020$	$e^{-19.7109} = 2.8e-9$

$$\Sigma(X_n, \chi)$$

.0474	.0110
.9421	.9738
.0003	.0020
1.832e-9	2.8e-9

$$\Sigma_n = \Sigma(X_n, X_n)$$

$e^0 = 1$	$e^{-3.9233} = .0198$	$e^{-20.3009} = 1.5e-9$	$e^{-25.6225} = 7.5e-12$
.0198	$e^0 = 1$	$e^{-6.746} = .0012$	$e^{-20.4893} = 1.263e-9$
1.5e-9	.0012	$e^0 = 1$	$e^{-34.0034} = 1.7e-15$
7.5e-12	1.263e-9	1.7e-15	$e^0 = 1$

Since Σ is symmetric, note that $\Sigma(X_n, \chi) = \Sigma(\chi, X_n)^T$

Calculation of Predictive Quantities

The MVN conditioning equations are used to determine the predictive quantities mean and variance

mean $\mu(\mathcal{X}) = \Sigma(\mathcal{X}, X_n) \Sigma_n^{-1} Y_n$

$$\mu(\chi) = y * = \begin{pmatrix} 0.2849657 \\ 0.2954011 \end{pmatrix} \quad \text{Predicted values for locations in } \chi$$

and variance $\Sigma(\mathcal{X}) = \Sigma(\mathcal{X}, \mathcal{X}) - \Sigma(\mathcal{X}, X_n) \Sigma_n^{-1} \Sigma(\mathcal{X}, X_n)^\top$

$$\Sigma(\chi) = \begin{pmatrix} 0.11154162 & -0.05042265 \\ -0.05042265 & 0.05155061 \end{pmatrix} \quad \text{Prediction Uncertainty}$$

The diagonal terms are the variances at prediction points 1 and 2

$$\sigma^2(\chi) = \begin{pmatrix} 0.11154162 \\ 0.05155061 \end{pmatrix}$$

R

Code to replicate this example in R

```
library(plgp)

eps = sqrt(.Machine$double.eps)

# Training points
X = rbind(c(-1.03,1.76), c(.49,.49), c(1.77,-1.77), c(3.62,3.76))

# The goal is to fit this function:  $y(x) = x_1 * \exp(-x_1^2 - x_2^2)$ 
y = X[,1] * exp(-X[,1]^2 - X[,2]^2)

# Test points
XX = rbind(c(.35, .69), c(.65, .46))
XX

# Sigma 22 (Sigma) and its inverse (Si)
# #####
# Distance among the Training Data
D = distance(X)
Sigma = exp(-D)
Si = solve(Sigma)

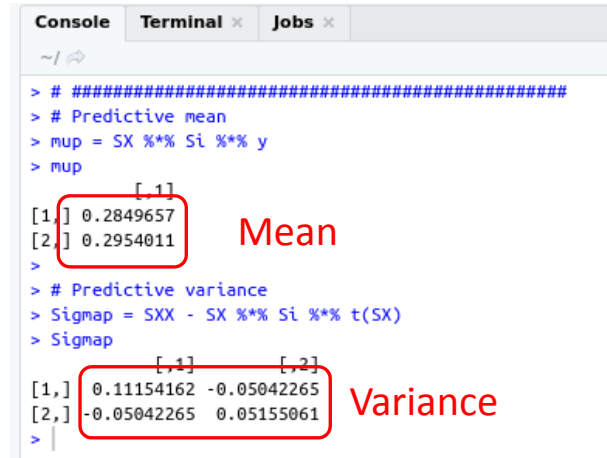
# Sigma 11
# #####
# Distance among the Testing Data
DXX = distance(XX)
SXX = exp(-DXX)

# Sigma 12 and Sigma 21 (Transpose of Sigma 12)
# #####
# Distance between training and testing data
DX = distance(XX, X)
SX = exp(-DX)

# Calculate the predictive mean and predictive variance
# #####
# Predictive mean
mup = SX %*% Si %*% y
mup

# Predictive variance
Sigmap = SXX - SX %*% Si %*% t(SX)
Sigmap
```

Output



```
> # #####
> # Predictive mean
> mup = SX %*% Si %*% y
> mup
      [,1]
[1,] 0.2849657
[2,] 0.2954011
>
> # Predictive variance
> Sigmap = SXX - SX %*% Si %*% t(SX)
> Sigmap
      [,1] [,2]
[1,] 0.11154162 -0.05042265
[2,] -0.05042265 0.05155061
> |
```

Mean

Variance

R

Code to replicate this example in R with Plots

```
library(plgp)
library(lhs)

eps = sqrt(.Machine$double.eps)

# Training Data
# #####
# Training points
number_of_sample_points = 4
X = rbind(c(-1.03,1.76), c(.49,.49), c(1.77,-1.77), c(3.62,3.76))

# Observed values
# The goal is to fit this function:  $y(x) = x_1 * \exp(-x_1^2 - x_2^2)$ 
y = X[,1] * exp(-X[,1]^2 - X[,2]^2)

# Testing Data
# #####
# Test points
number_of_test_points_per_axis = 40
xx = seq(-2, 4, length=number_of_test_points_per_axis)
XX = expand.grid(xx, xx)

# Sigma 22 (Sigma) and its inverse (Si)
# #####
# Distance among the Training Data
D = distance(X)
Sigma = exp(-D) + diag(eps, nrow(X))
Si = solve(Sigma)

# Sigma 11
# #####
# Distance among the Testing Data
```

```
DXX = distance(XX)
SXX = exp(-DXX)

# Sigma 12 and Sigma 21 (Transpose of Sigma 12)
# #####
# Distance between training and testing data
DX = distance(XX, X)
SX = exp(-DX)

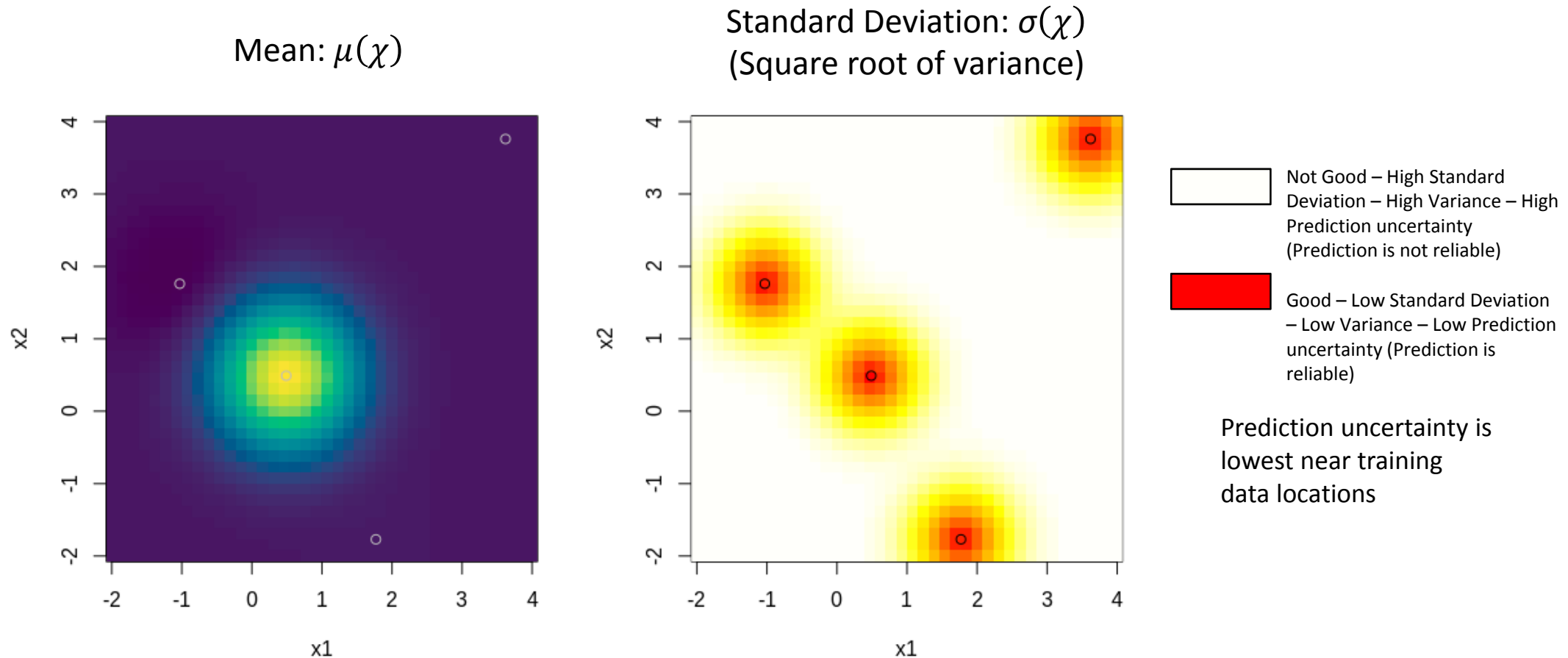
# Calculate the predictive mean and predictive variance
# #####
mup = SX %*% Si %*% y
Sigmap = SXX - SX %*% Si %*% t(SX)

# Predictive standard deviation
diag(Sigmap)
sdp = sqrt(diag(Sigmap))

# Figure 5.5
par(mfrow=c(1, 2))
cols_a = hcl.colors(128, palette = "viridis")
cols_b = heat.colors(128)
image(xx, xx, matrix(mup, ncol=length(xx)), xlab='x1', ylab='x2', col=cols_a)
points(X[,1], X[,2])
image(xx, xx, matrix(sdp, ncol=length(xx)), xlab='x1', ylab='x2', col=cols_b)
points(X[,1], X[,2])

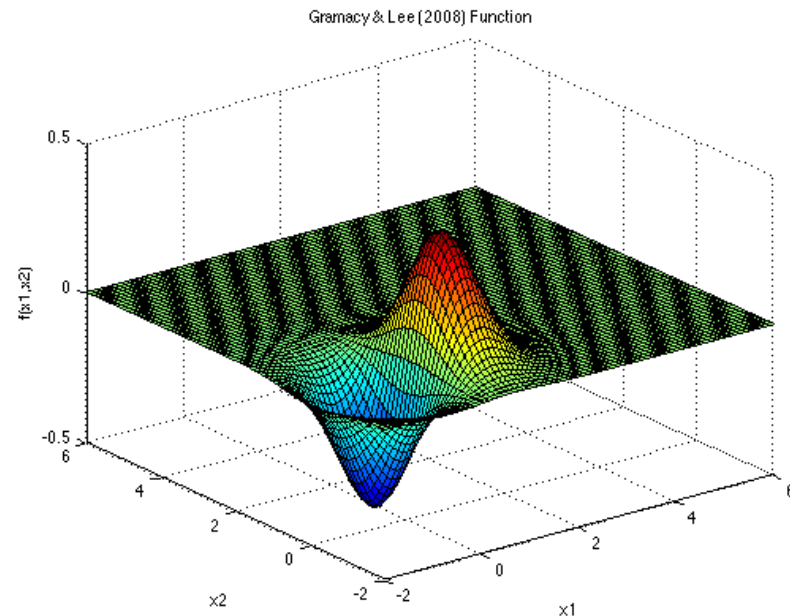
# Figure 5.6
persp(xx, xx, matrix(mup, ncol=number_of_test_points_per_axis), theta=-30, phi=30,
xlab='x1', ylab='x2', zlab='y', zlim = c(-.5,.5))
```


Predictive Quantities Mean and Standard Deviation



Comparison of True Function and Prediction Model

True Function



$$f(\mathbf{x}) = x_1 \exp(-x_1^2 - x_2^2)$$

Source: <https://www.sfu.ca/~ssurjano/grlee08.html>

Prediction Model ($\mu(\chi)$)

