Workshop – Prediction Analysis, Buckling

AN MSC NASTRAN MACHINE LEARNING WEB APP TUTORIAL

Questions? Email: christian@ the-engineering-lab.com

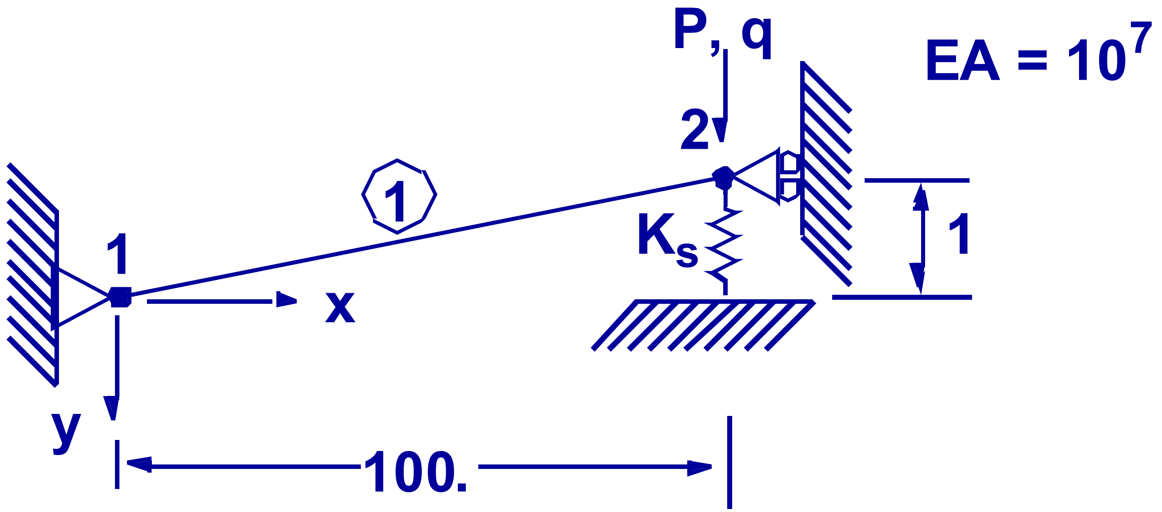
Goal: Prediction Analysis

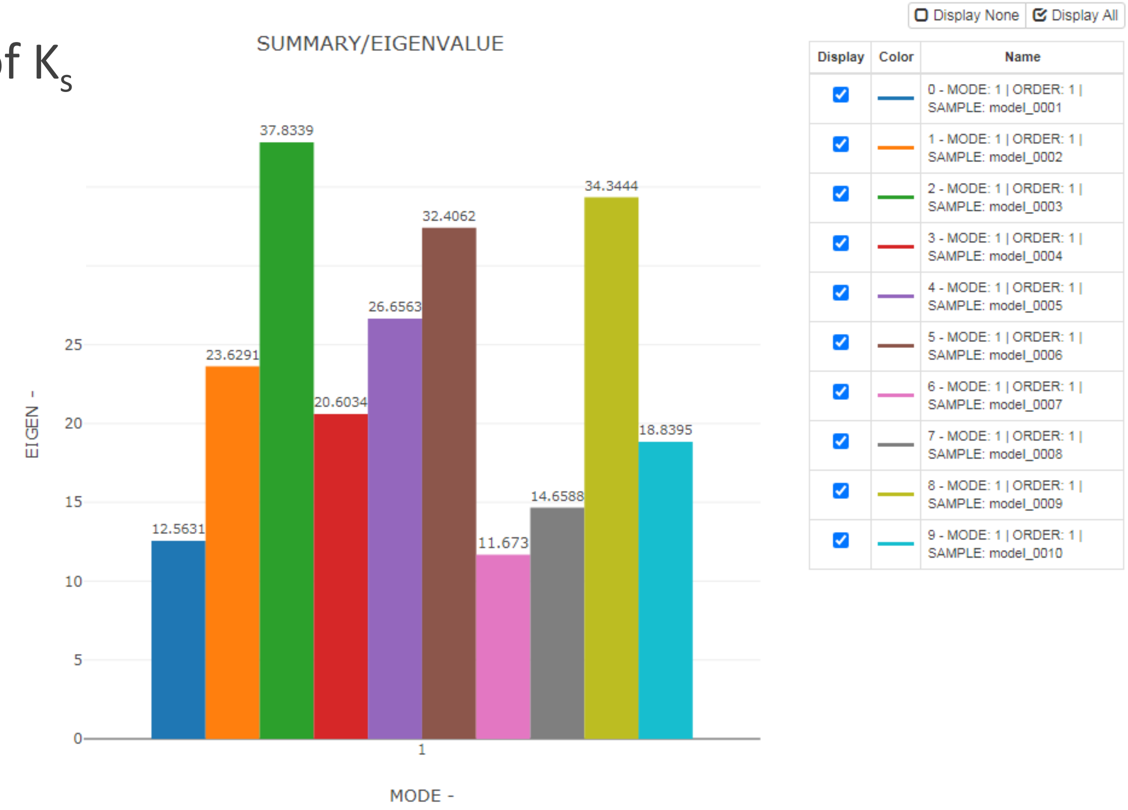
This tutorial consists of multiple parts

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 and perform predictions.
   - The prediction performance of the surrogate model is evaluated.
4. Creating Plots with the HDF5 Explorer
   - The HDF5 Explorer web app is used to create Eigenvalue vs. Mode Number plots.

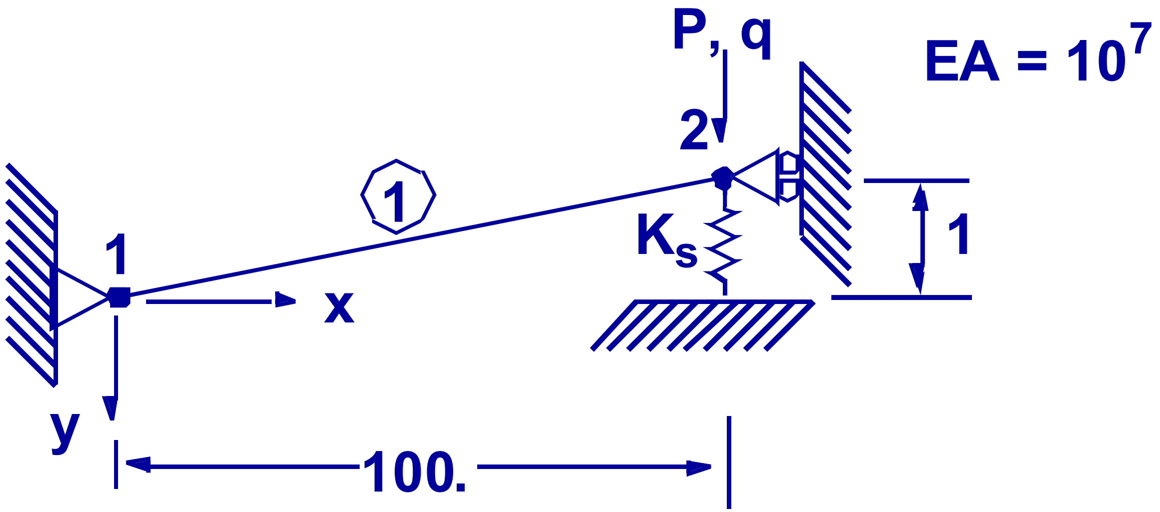
Details of the Structural Model

1. Perform a buckling analysis for different values of K_s
2. Monitor the eigenvalue response for each value of K_s





Details of the Structural Model



Problem Statement

Design Variables

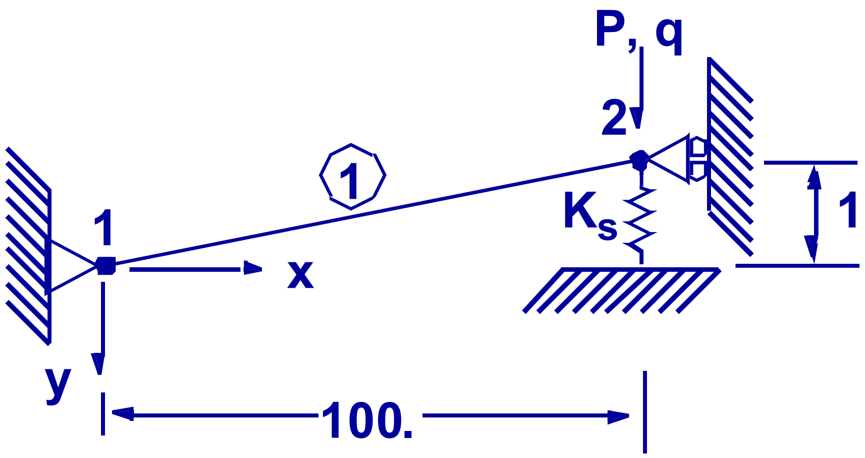
x1: K of PELAS 20

Note that K is displayed as K_s in the figure

.1 < x1 < 10.0

Samples

- Batch set 1 – 10 run LHS Design
- Batch set 2 – 20 run LHS Design



Monitored Responses

r0: Eigenvalue of Mode 1

Contact me

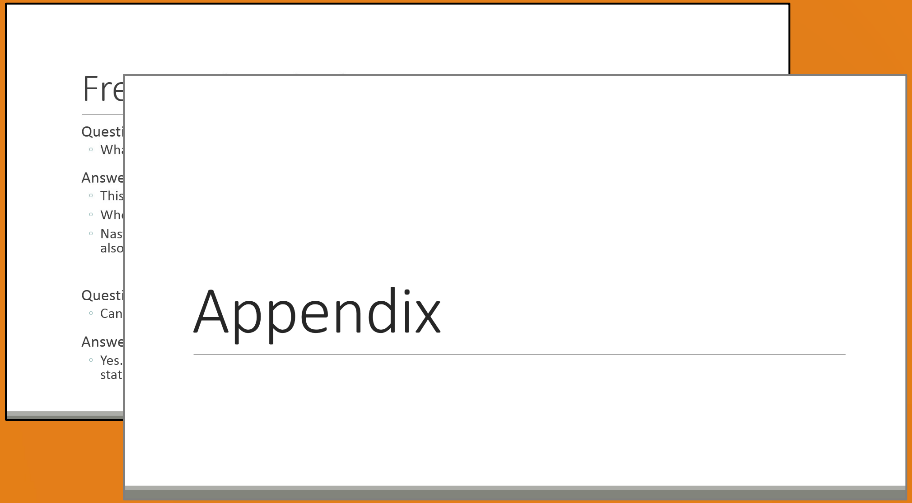
- Nastran SOL 200 training
- Nastran SOL 200 questions
- Structural or mechanical optimization questions
- Access to the SOL 200 Web App

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.

SOL 200 Web App Capabilities

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

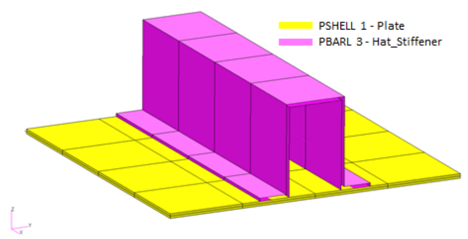
Compatibility

- 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.

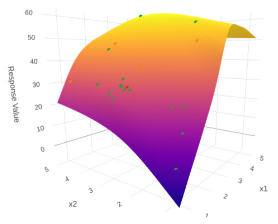
Benefits

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

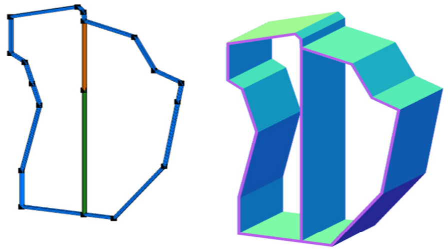
Web Apps



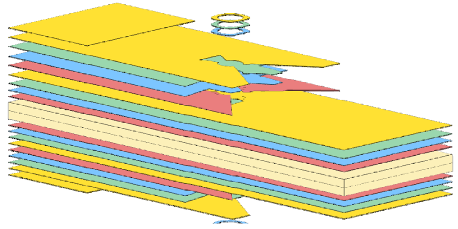
Web Apps for MSC Nastran SOL 200
Pre/post for MSC Nastran SOL 200. Support for size, topology, topometry, topography, multi-model optimization.



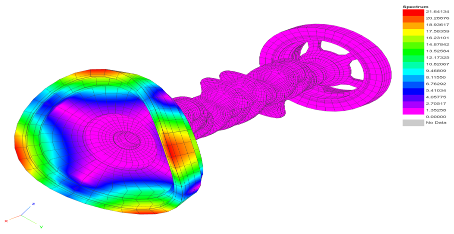
Machine Learning Web App
Bayesian Optimization for nonlinear response optimization (SOL 400)



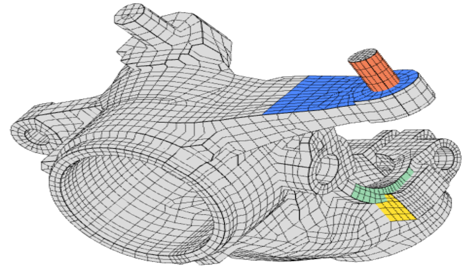
PBMSECT Web App
Generate PBMSECT and PBRSECT entries graphically



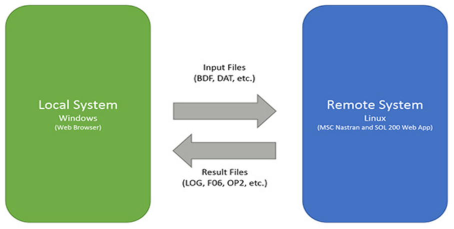
Ply Shape Optimization Web App
Optimize composite ply drop-off locations, and generate new PCOMPG entries



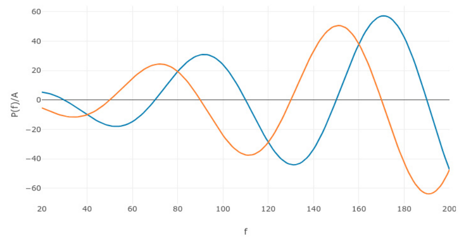
Post-processor Web App
View MSC Nastran results in a web browser on Windows and Linux



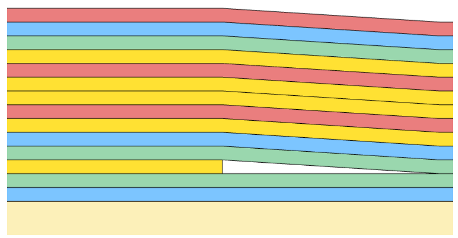
Shape Optimization Web App
Use a web application to configure and perform shape optimization.



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

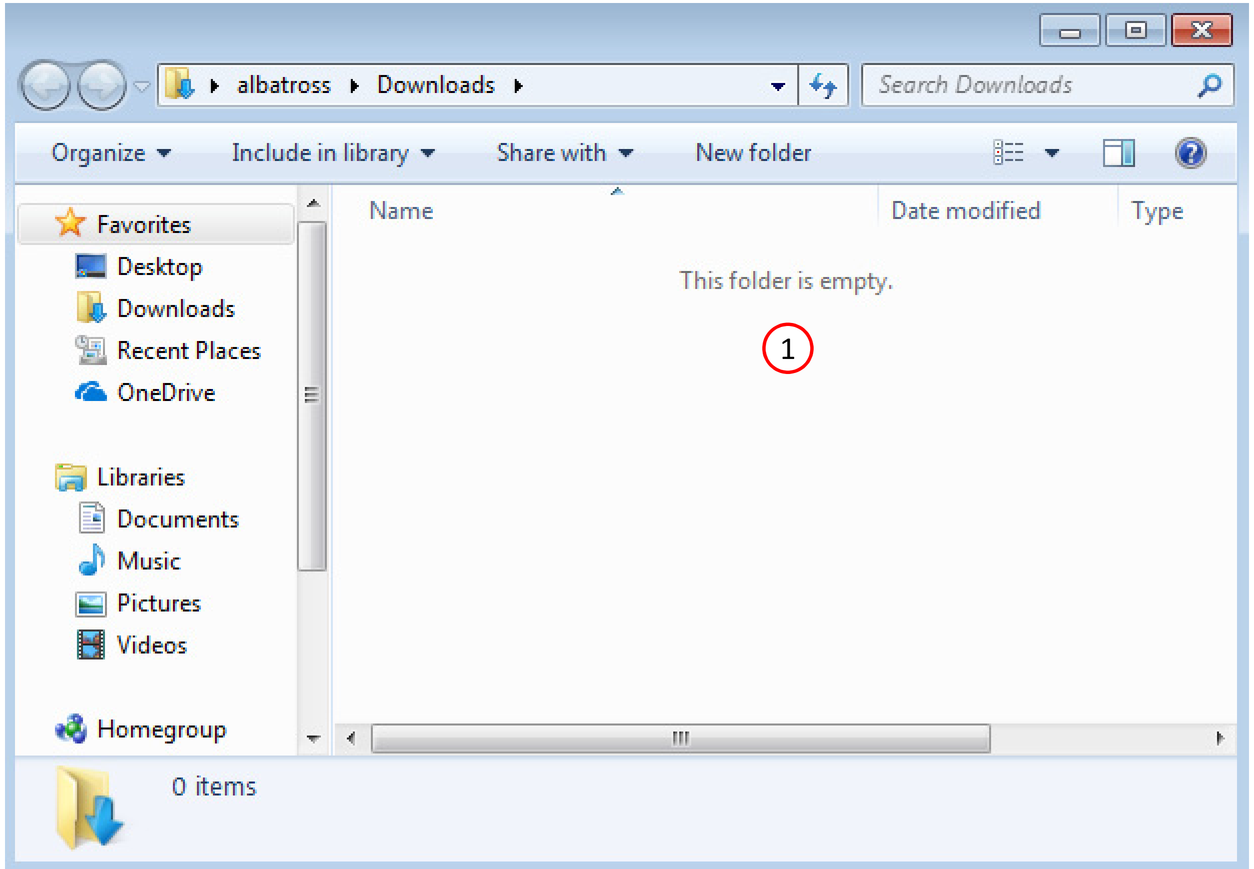
HDF5 Explorer Web App
Create graphs (XY plots) using data from the H5 file

Configuring The Problem Statement

Before Starting

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

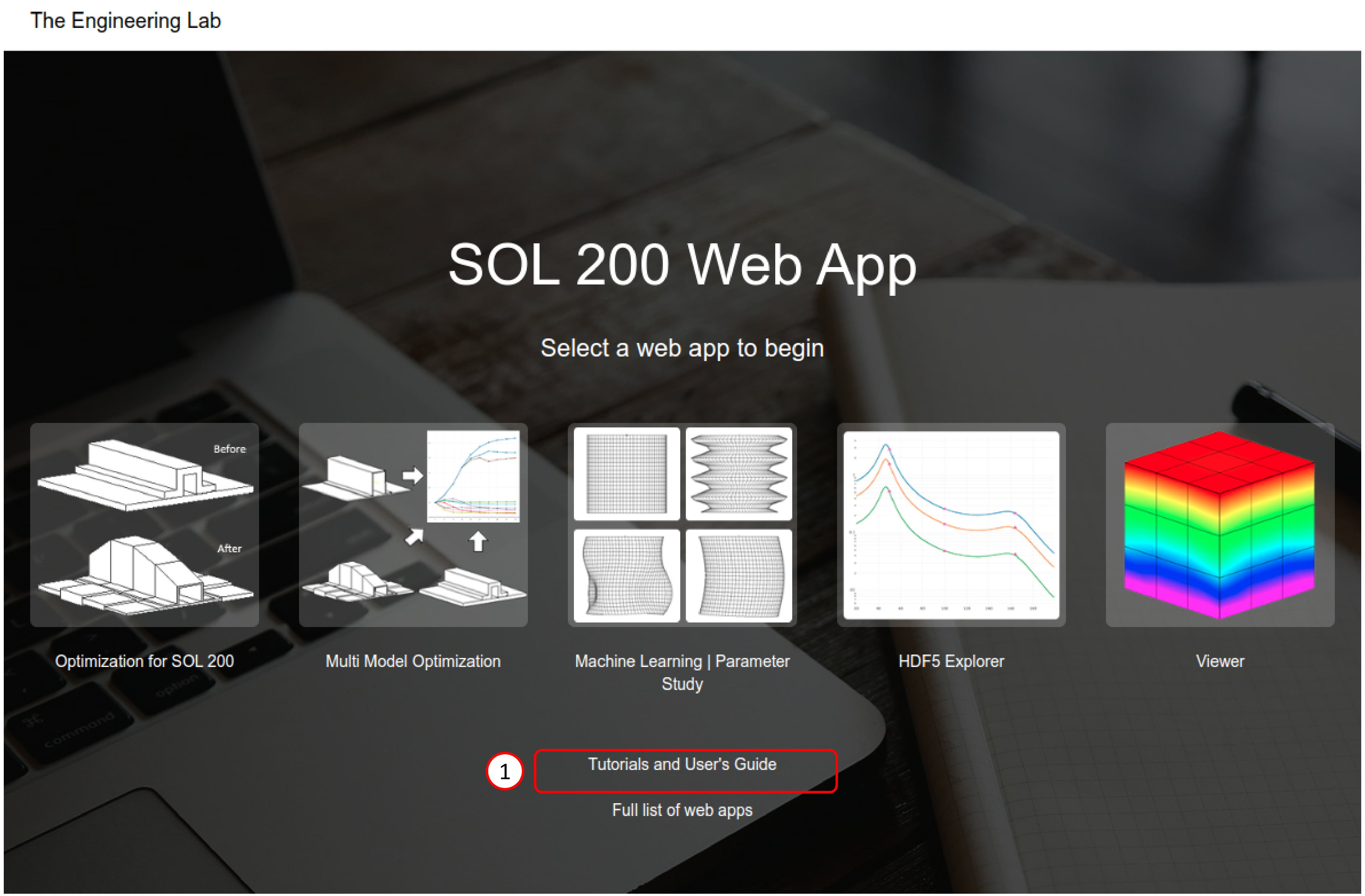
- Throughout this workshop, you will be working with multiple file types and directories such as:
  - .bdf/.dat
  - nastran_working_directory
  - .f06, .log, .pch, .h5, etc.
- To minimize confusion with files and folders, it is encouraged to start with a clean directory.



Go to the User's Guide

1. Click on the indicated link

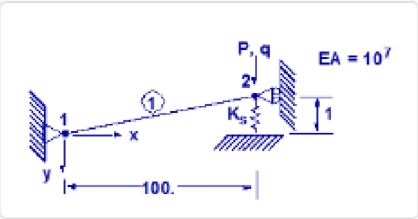
- The necessary BDF files for this tutorial are available in the Tutorials section of the User's Guide.



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.

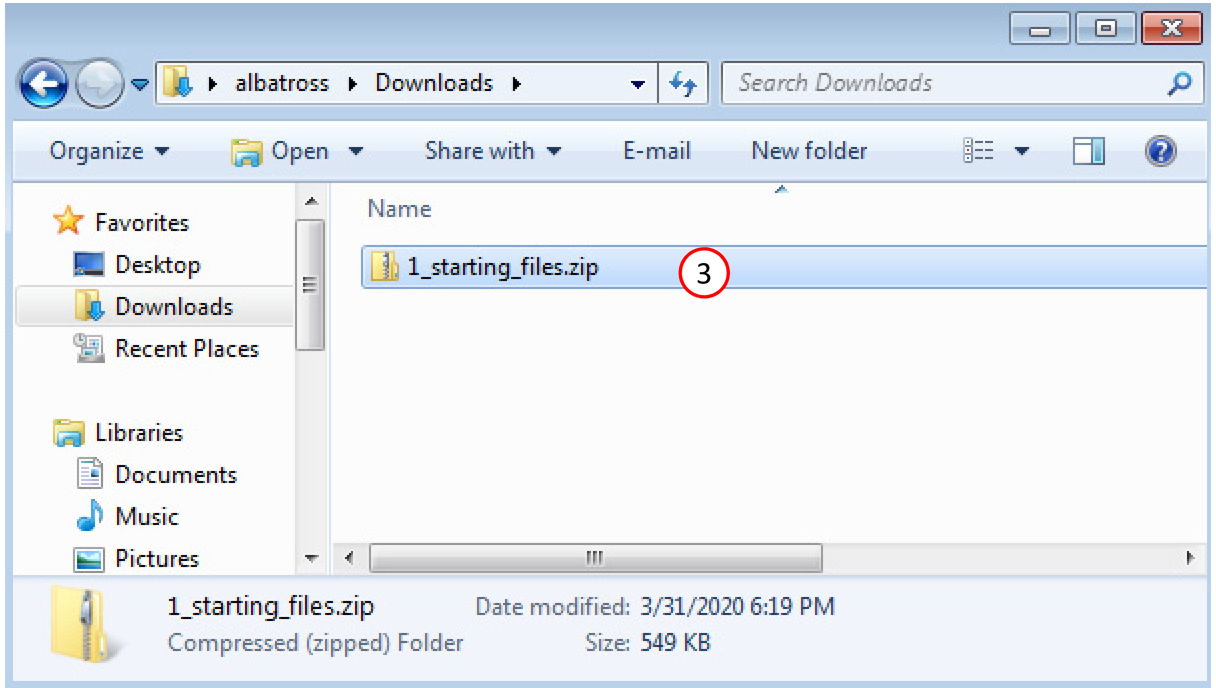


Prediction Analysis, Buckling 1

Consider a linear buckling analysis. The parameter allowed to vary is a spring constant. The response of interest is the buckling load factor.

This tutorial describes how to configure multiple MSC Nastran runs to generate training data. Gaussian process regression is used to train a surrogate model and make predictions. The prediction performance of the surrogate model is also evaluated.

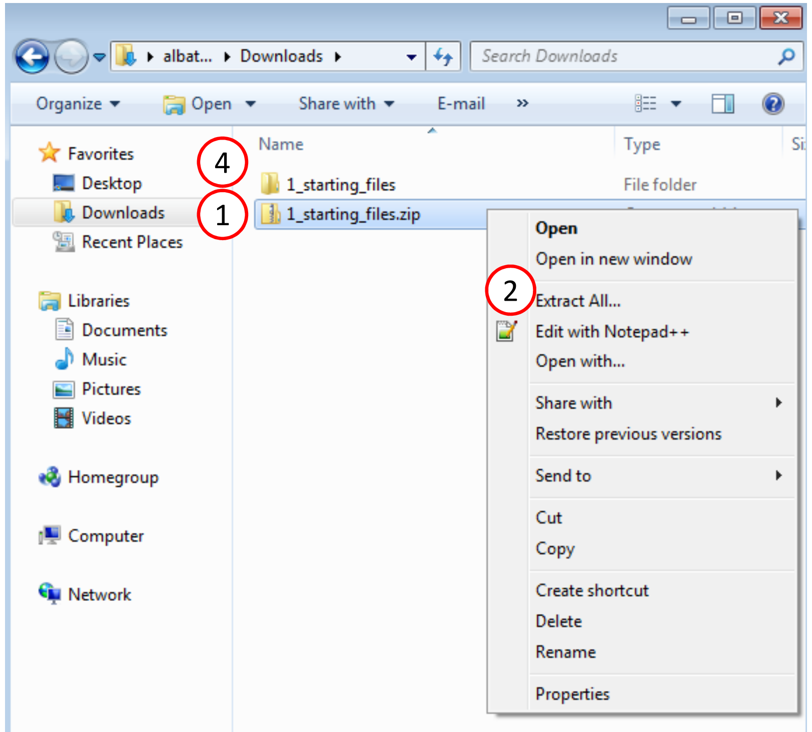
Starting Files Link 2
Solution BDF Files: Link

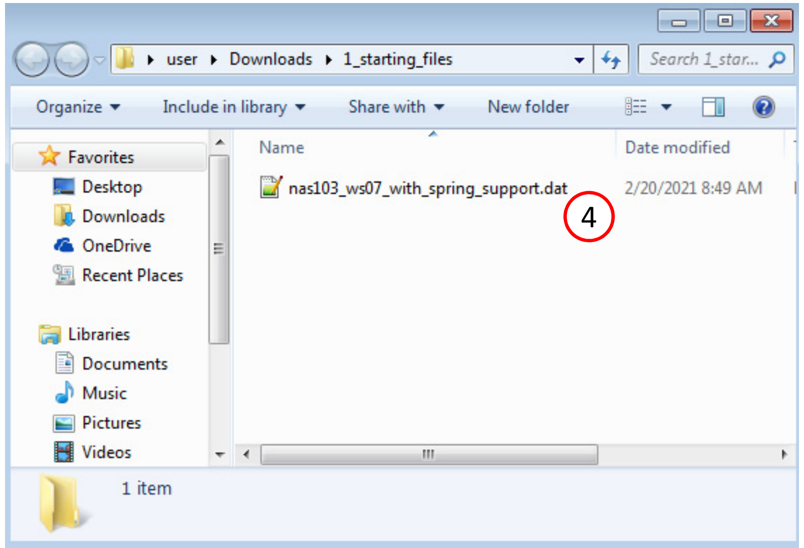


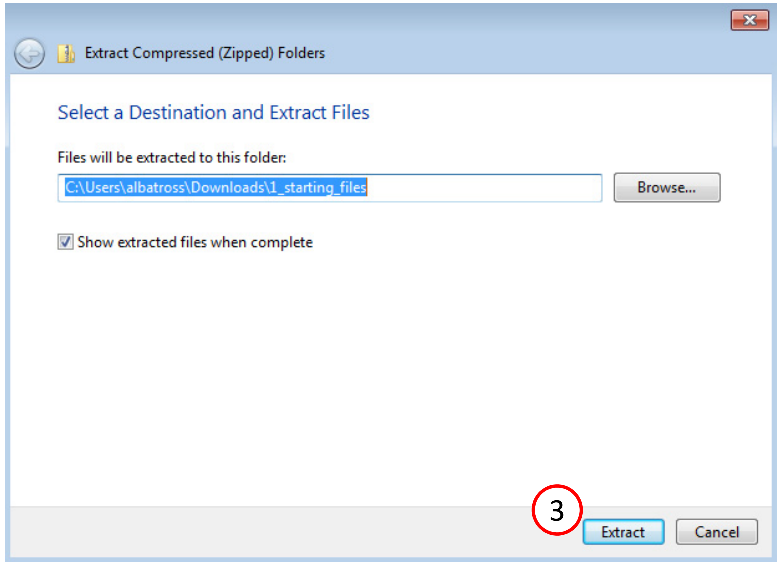
Obtain Starting Files

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.



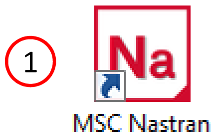


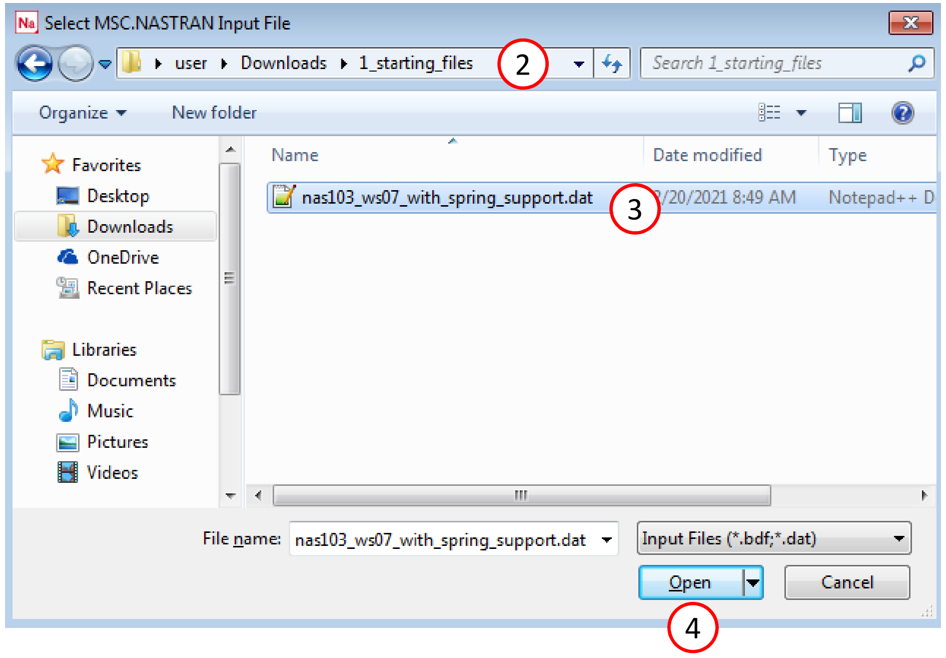


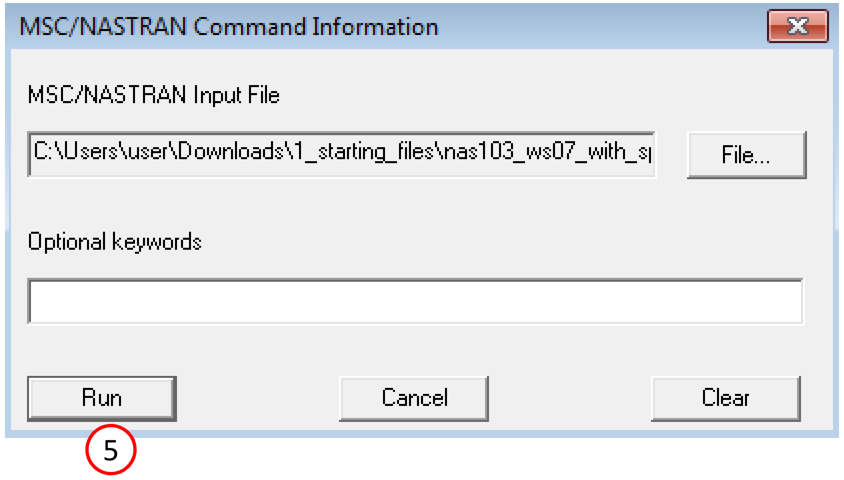
Create the Starting H5 File

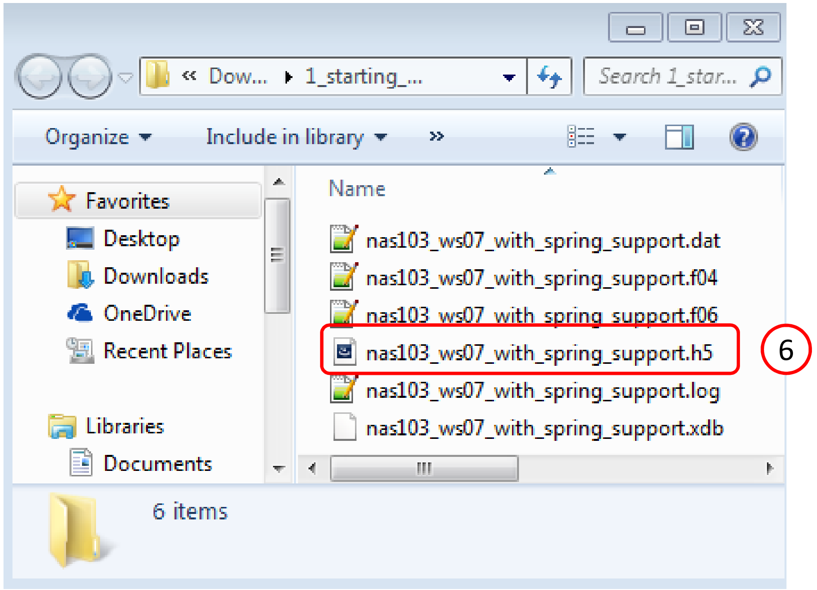
A starting H5 file must be created. This H5 file will be used to configure the responses later on.

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.

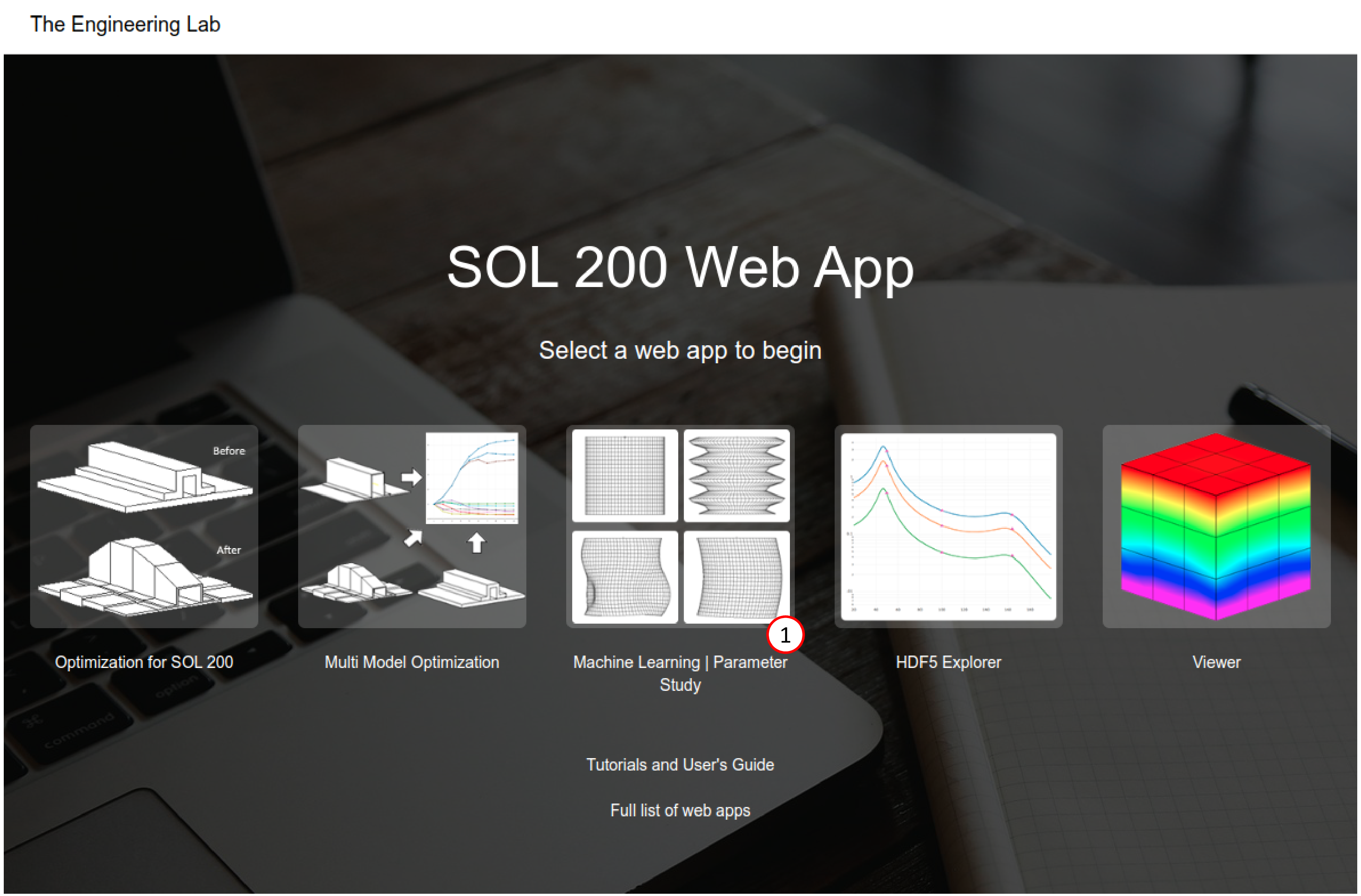
- Scenario 1 - OK
    - MSC Nastran 2021 is used to create the starting H5 file.
    - MSC Nastran 2021 is used for each run during Machine Learning or Parameter study.
- Scenario 2 – NOT OK
    - MSC Nastran 2018.2 is used to create the starting H5 file.
    - MSC Nastran 2021 is used for each run during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

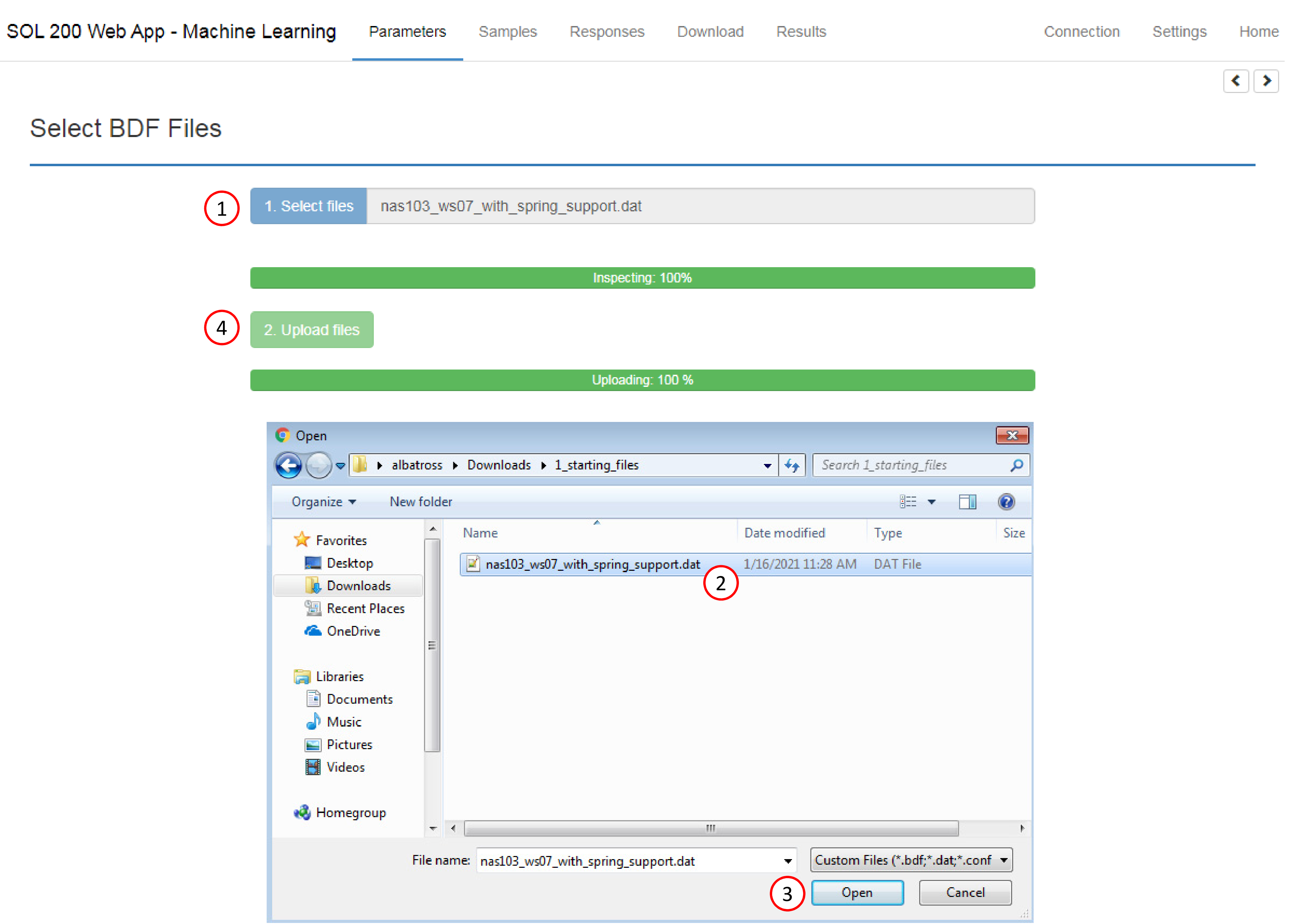
1. Click on the indicated link

- 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



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

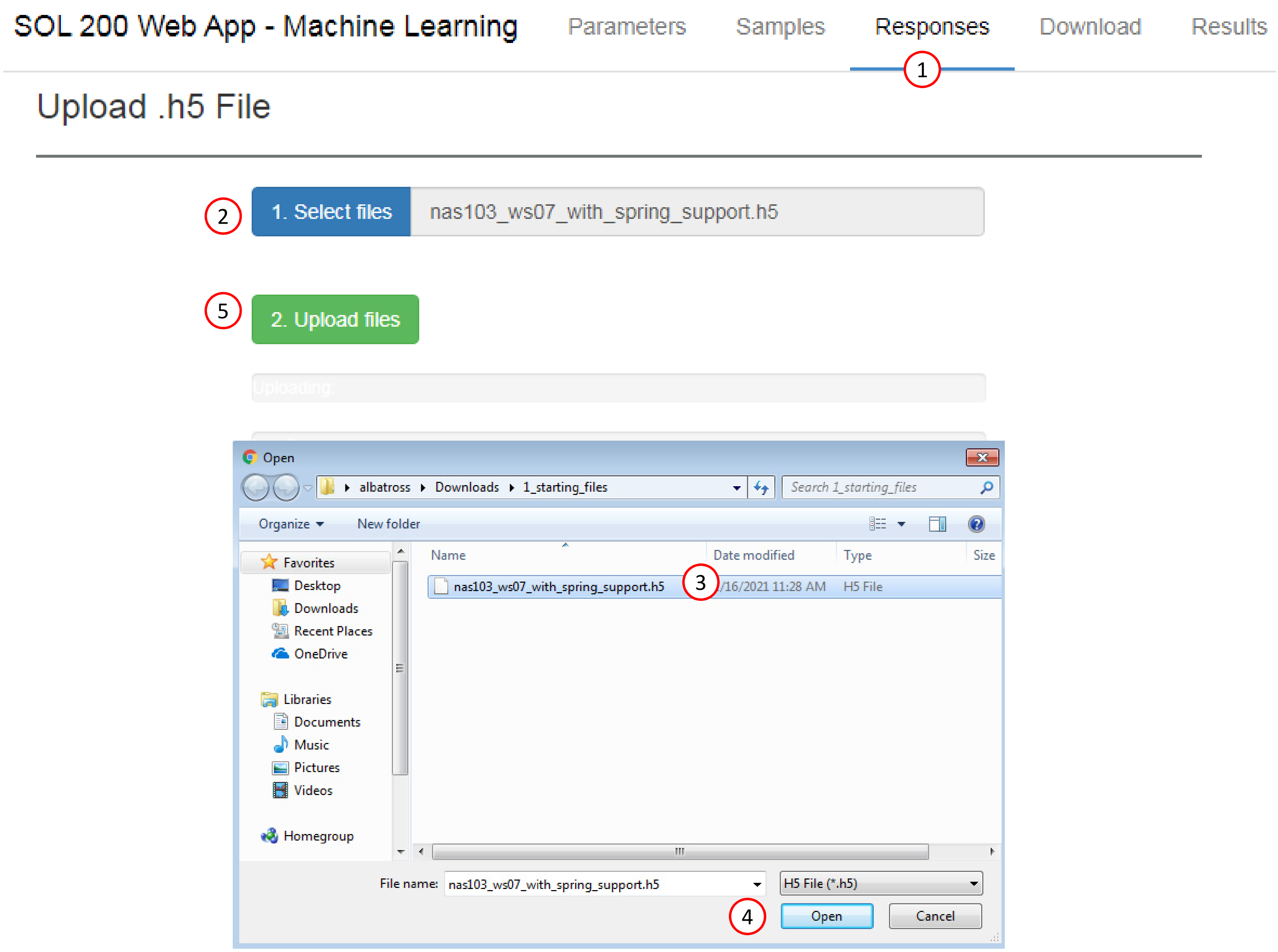
```
$__1___||__2___||__3___||__4___||__5___||__6___||__7_

EIGB    30      INV     0.      3.      20      2
FORCE   6       2               1.      0. -1.  0.
MAT1    1       10.E7
PELAS   20      %x1%
PROD    10      1       .1
```

Configure Parameters

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

Responses

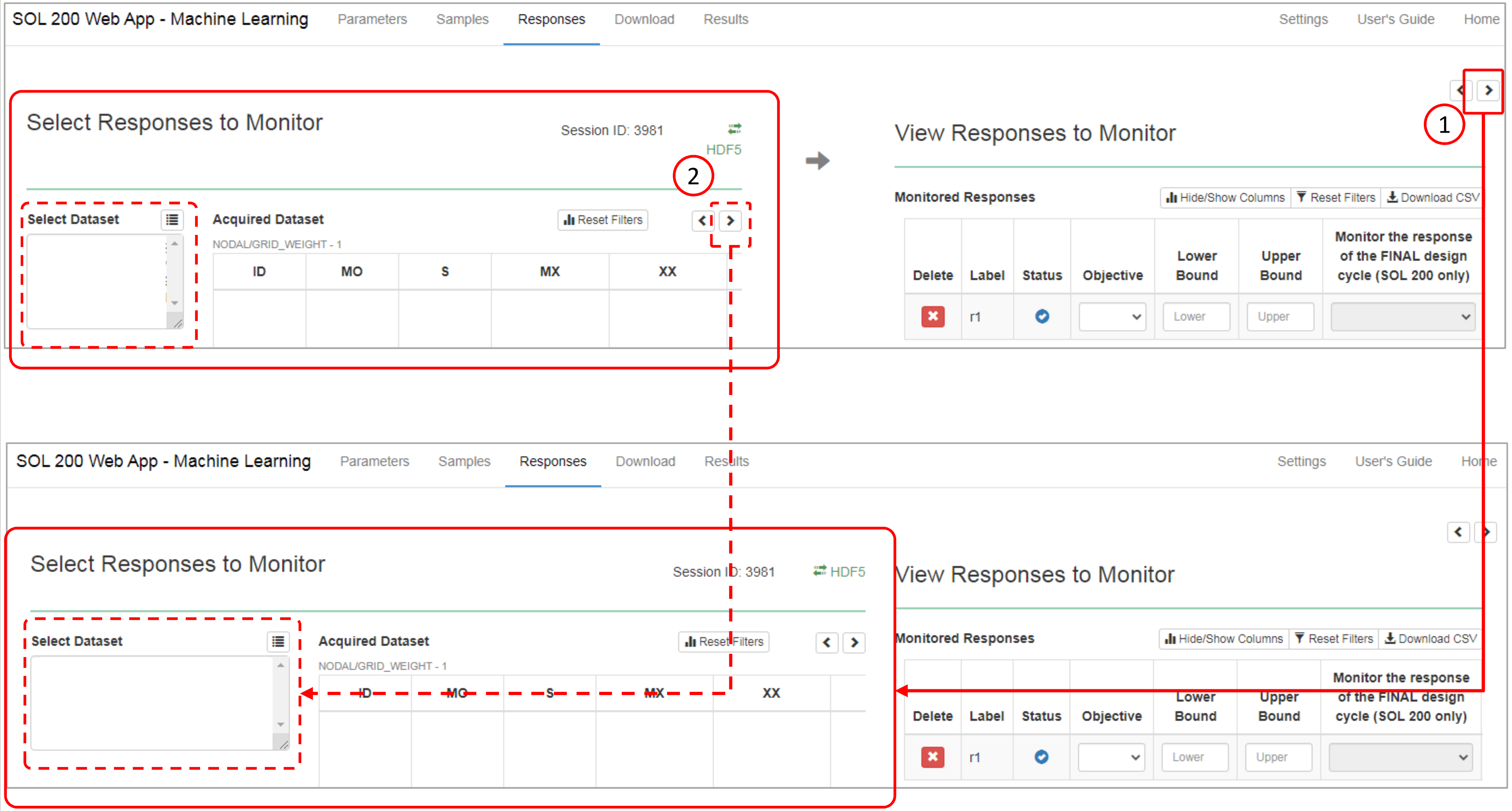
1. Click Responses
2. Click Select files
3. Select the indicated file
4. Click Open
5. Click Upload files



Adjust the Column Width

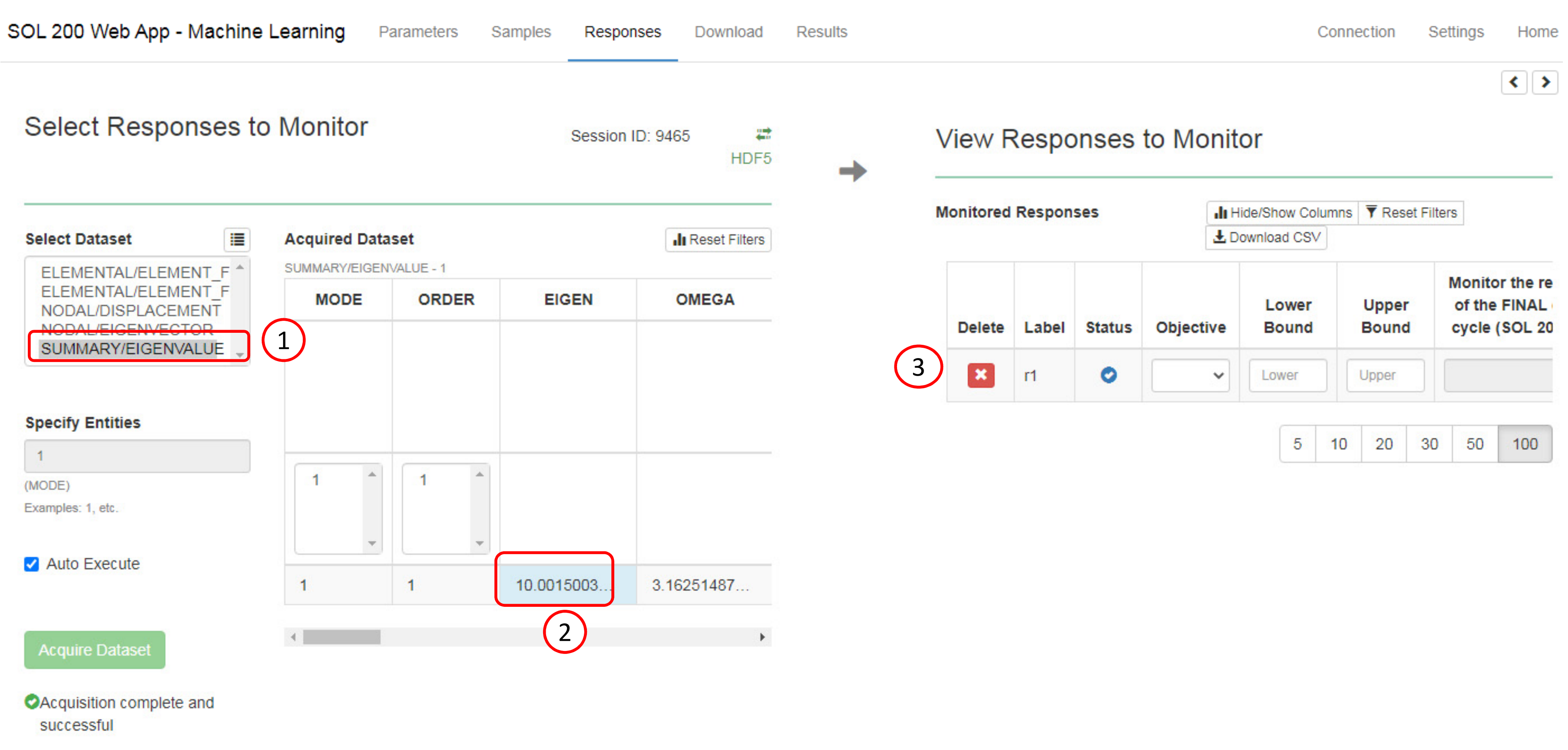
1. Optional - Use at your liking the buttons at the top right hand corner to adjust the width of the left and right columns
2. Optional – Use the indicated buttons to adjust the width of the column Select Dataset

- IMPORTANT! This image is not meant to match exactly what you see in your view. The text in this image is expected to be different from your view. The purpose of this page and image is to demonstrate how to increase the width of the indicated sections.



Select Responses

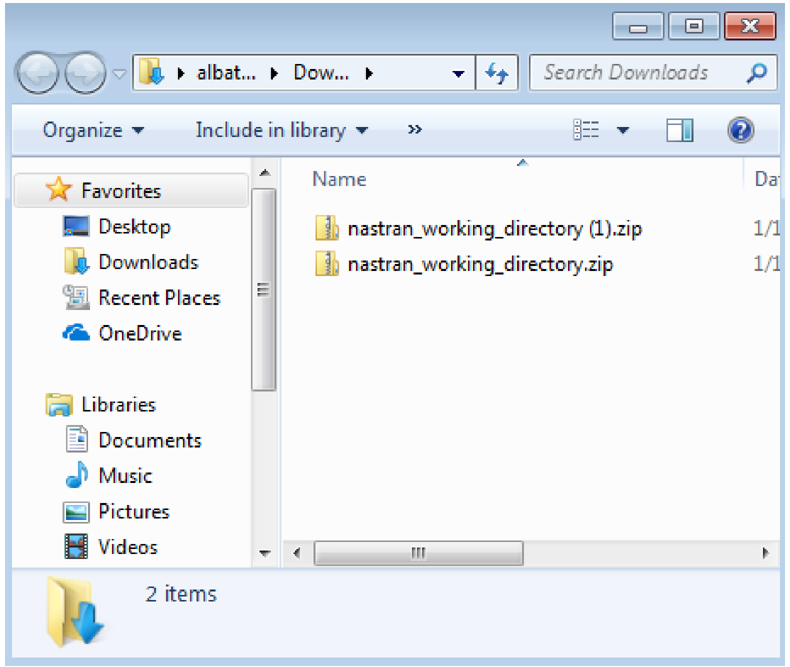
1. Select the following dataset: SUMMARY/EIGENVALUE
2. Use the horizontal scroll bar until the column TIME_FREQ_EIGR is visible
3. A new response r1 is created



Configuring Multiple Batch Runs

Samples

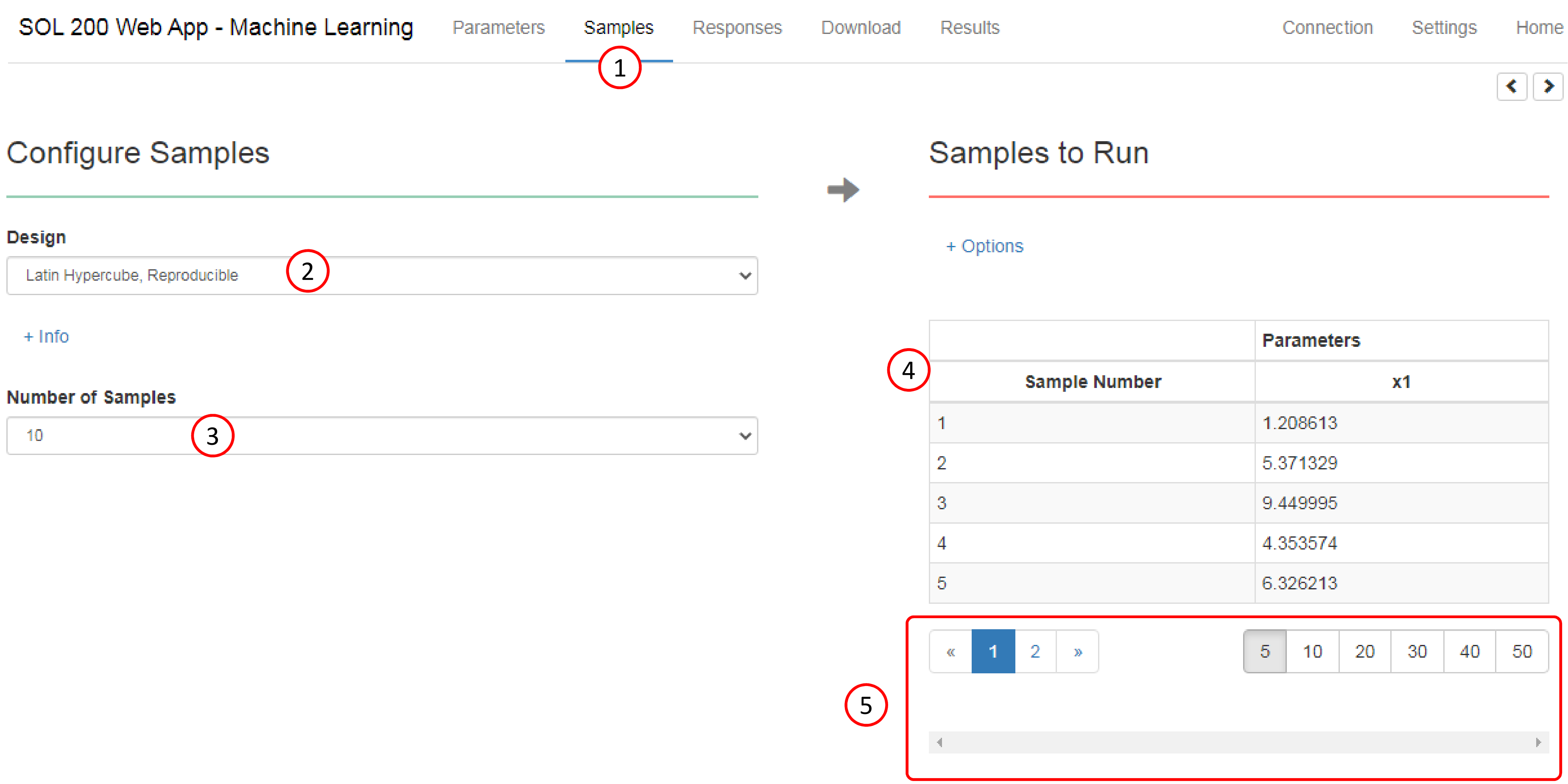
In the following slides, we will configure 2 batches to run.



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.
2	nastran_working_directory (1).zip	20	The data from these 20 runs is compared with the predictions from the surrogate model. The normalized root mean square error (NRMSE) is calculated based on these 20 runs.

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

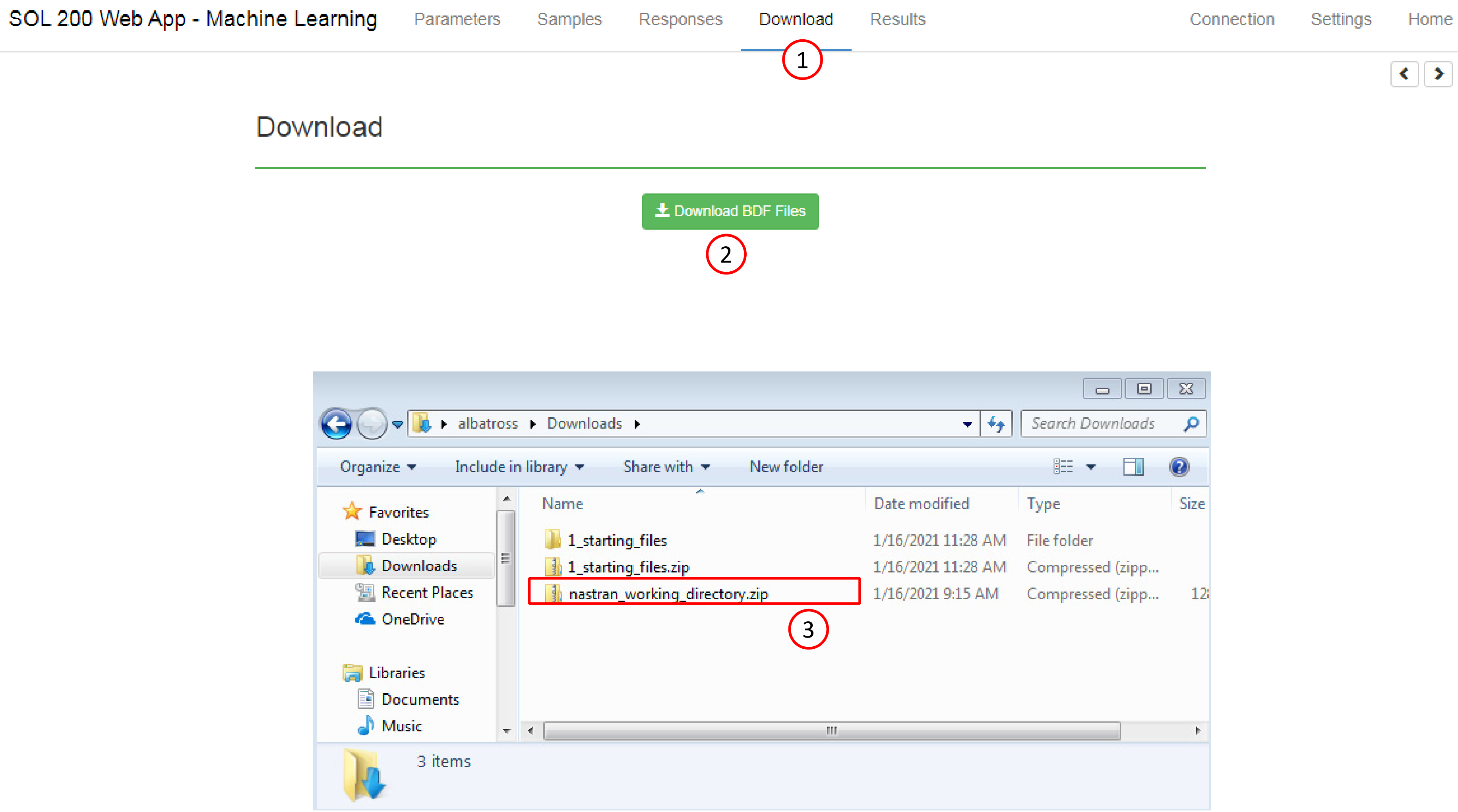


Questions? Email: christian@ the-engineering-lab.com

HEXAGON Technology Partner

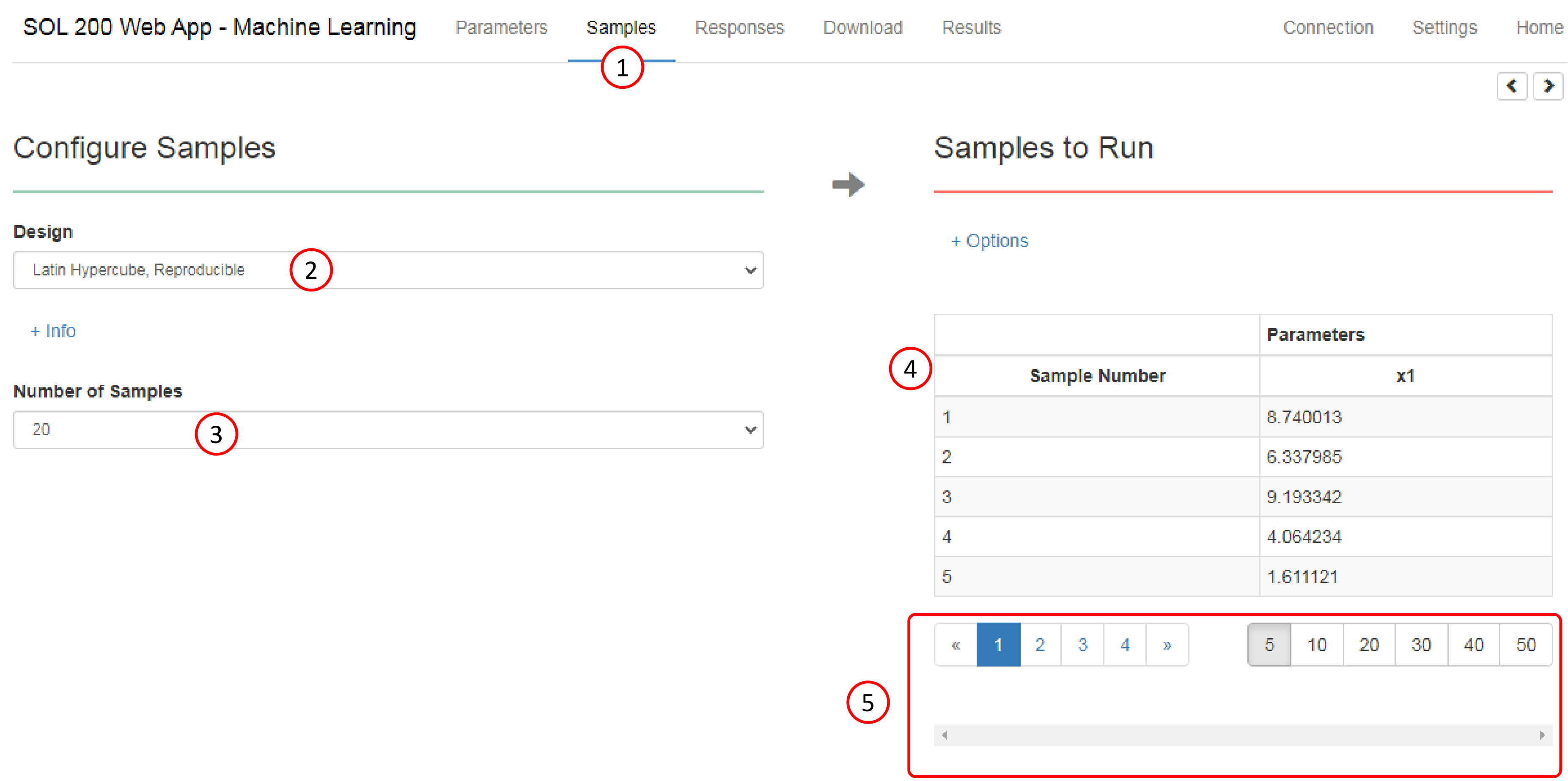
Download

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



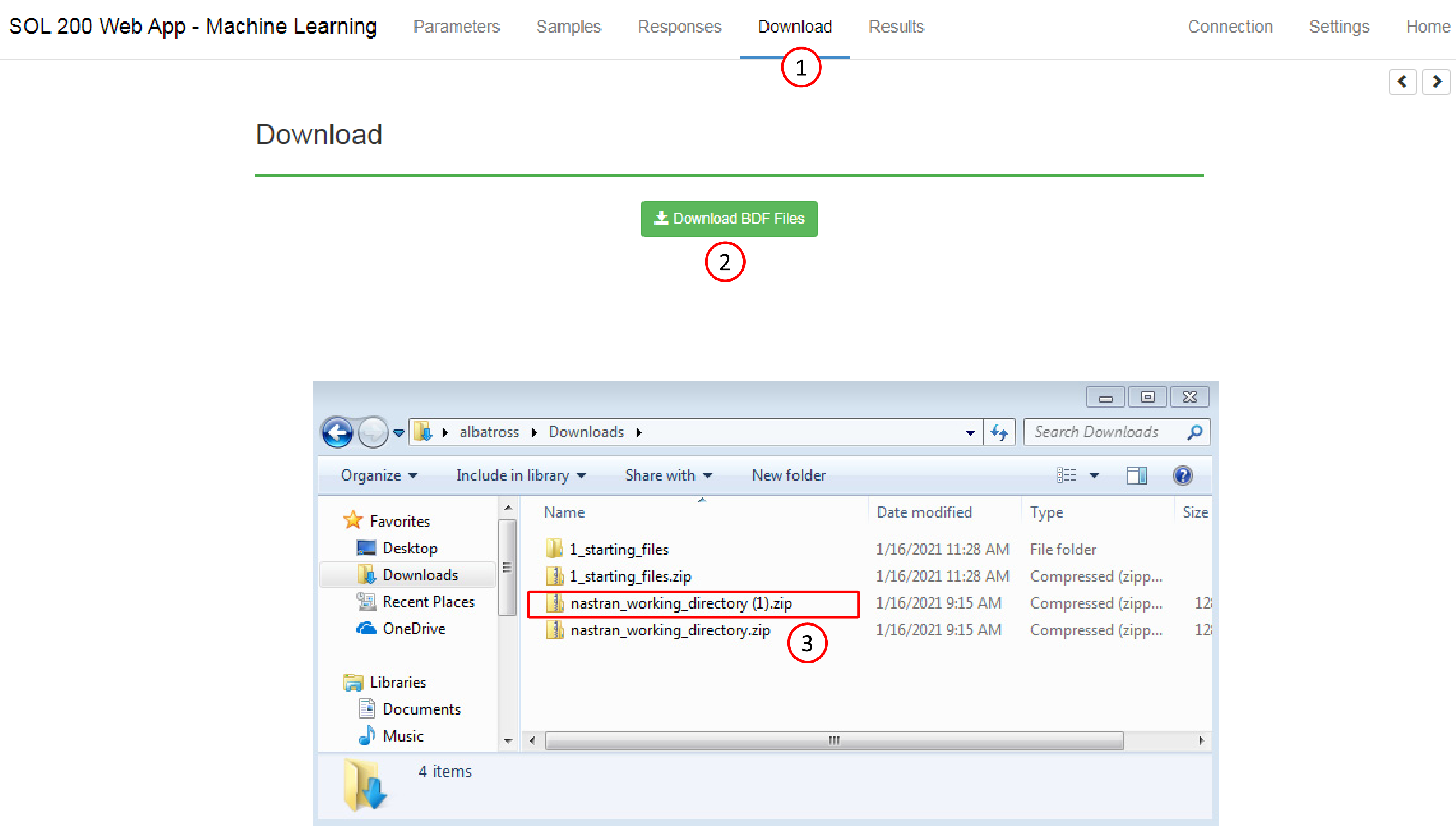
Samples

1. Click Samples
2. Ensure the following design is selected: Latin Hypercube, Reproducible
3. Set Number of Samples to 20
4. The samples have been updated, note that samples 1, 2, 3, …, 20 are visible
5. The indicated controls can be used to display the other samples



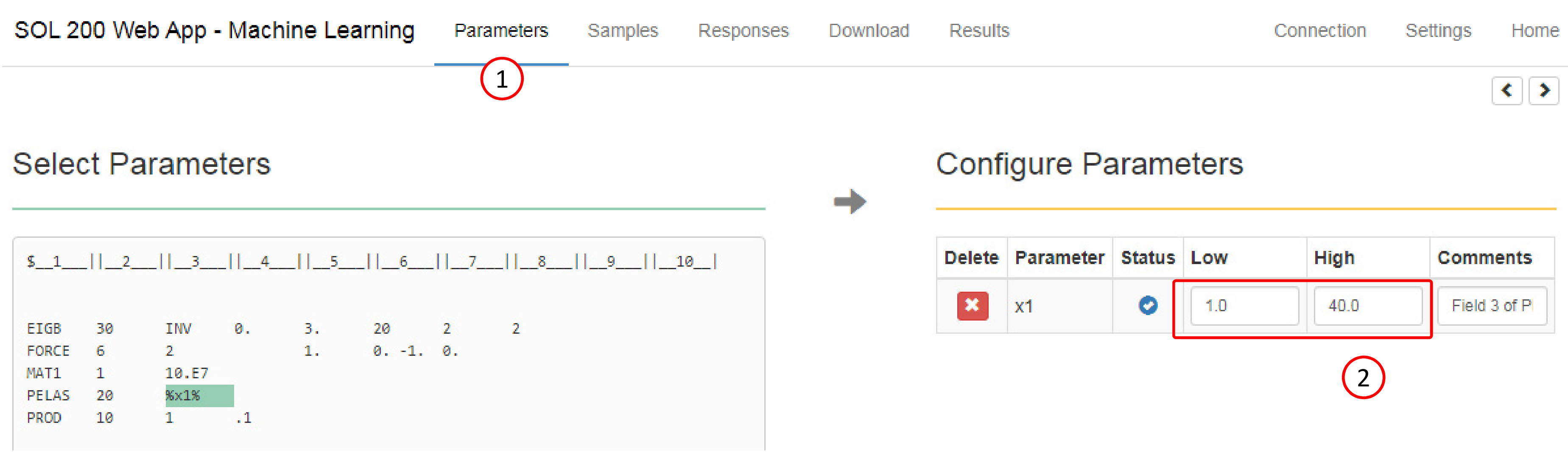
Download

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



Parameters

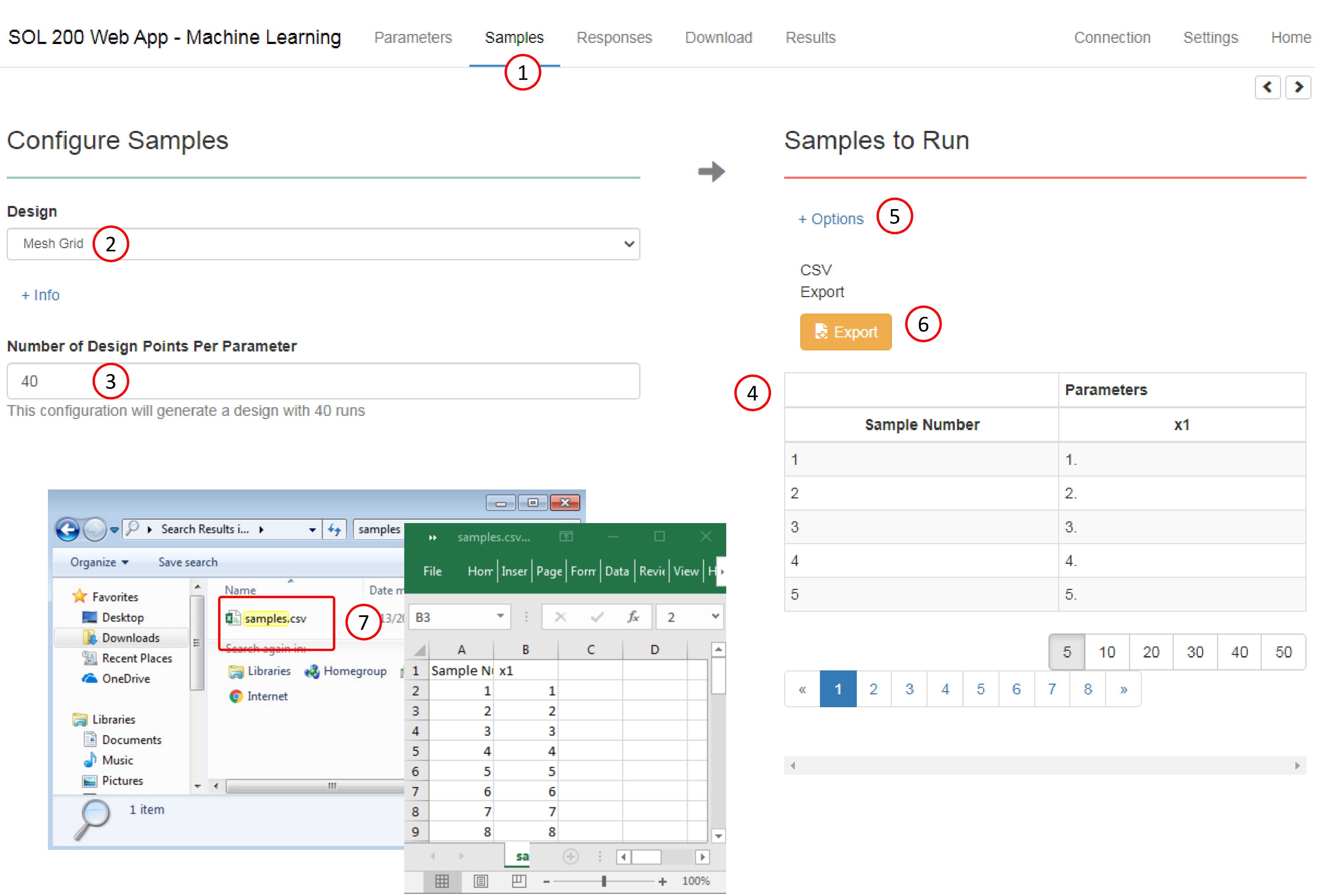
1. Click Parameters
2. For each parameter, use the following settings:
    - Low: 1.0
    - High: 40.0



Samples

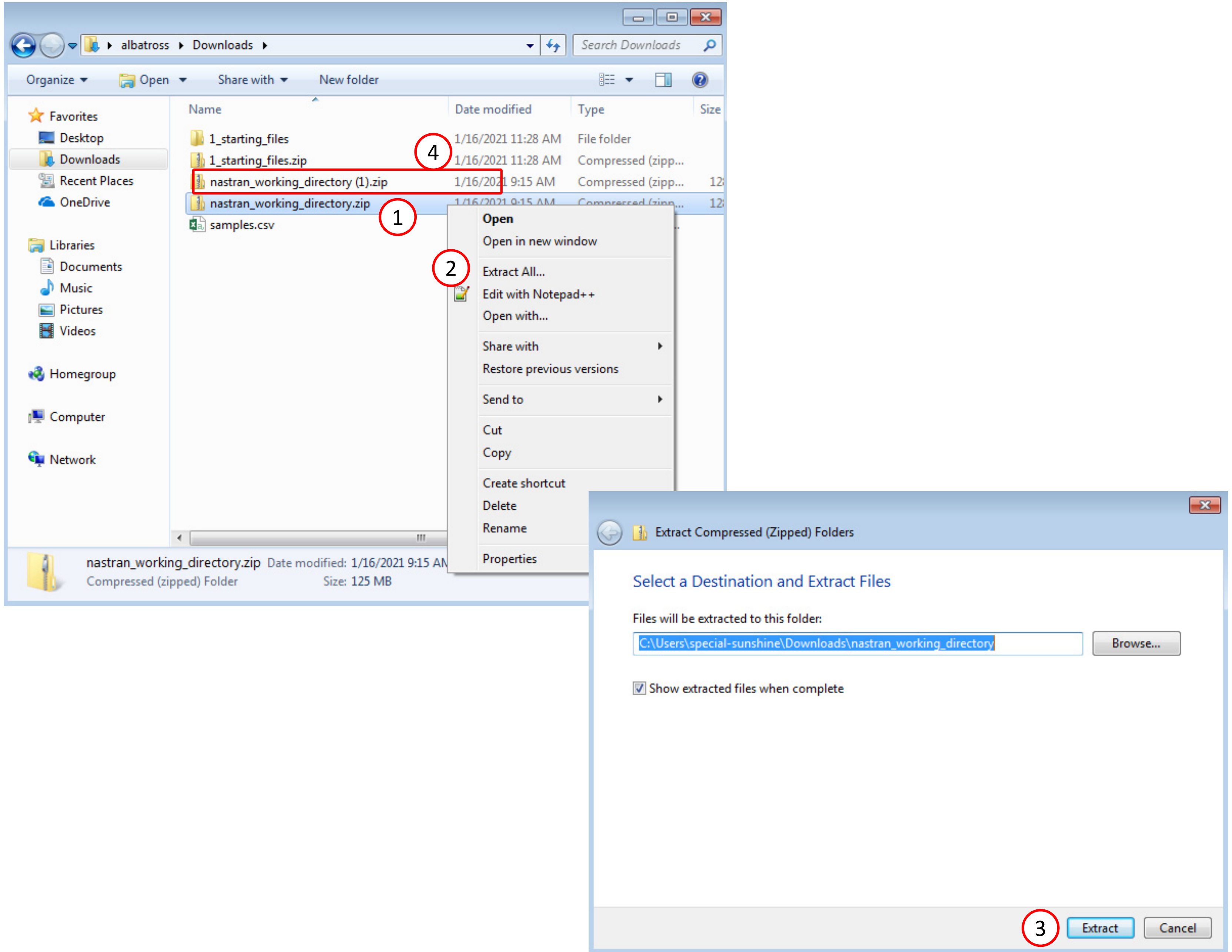
1. Click Samples
2. Set the Design as Mesh Grid
3. Set the Number of Samples as 40
4. The table now has 40 samples
5. Click +Options
6. Click Export
7. A CSV file has been download and contains the values from the table with 8 samples

- Later in the tutorial, the 40 sample points in the CSV file will be used to make predictions.



Start Desktop App

1. Right click on the indicated file
2. Click Extract All
3. Click Extract on the following window
4. Repeat steps 1-3 for the indicated files



- 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

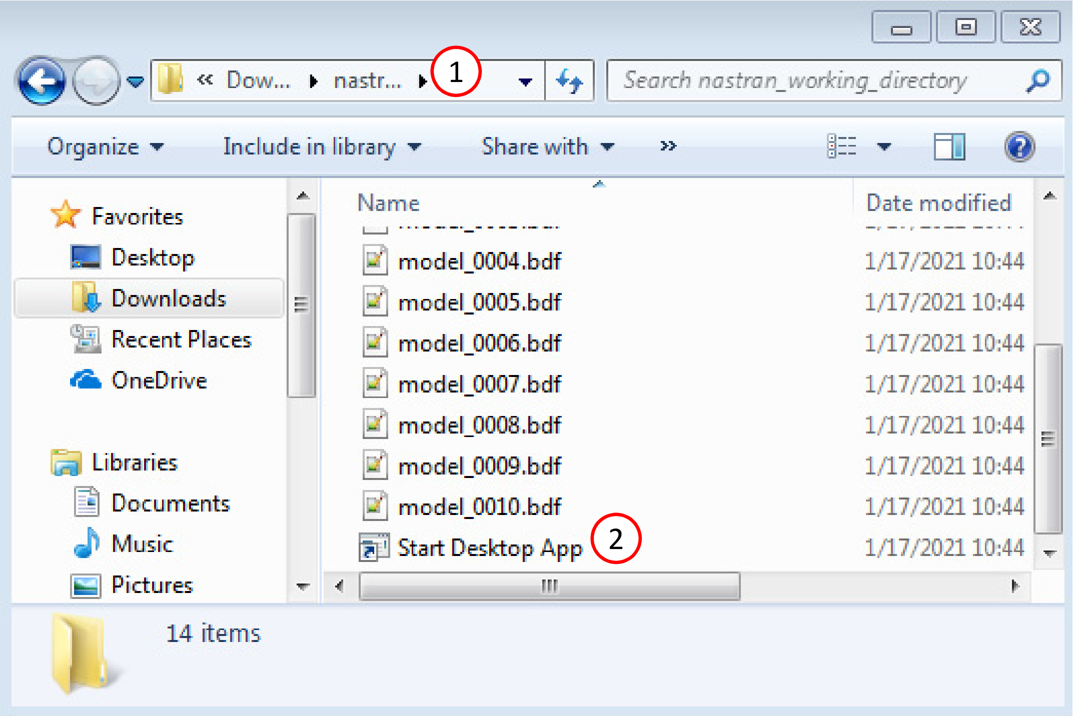
- One can run the Nastran job on a remote machine as follows:
  1) Copy the BDF files and the INCLUDE files to a remote machine. 2) Run the MSC Nastran job on the remote machine. 3) After completion, copy the BDF, F06, LOG, H5 files to the local machine. 4) Click "Start Desktop App" to display the results.

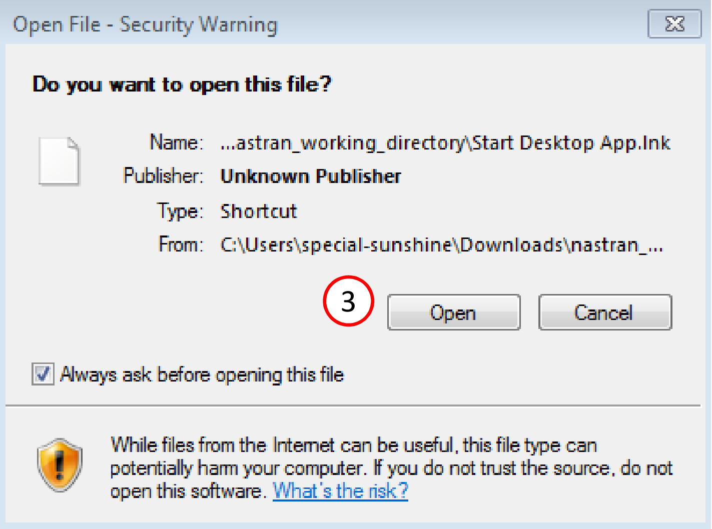
Using Linux?

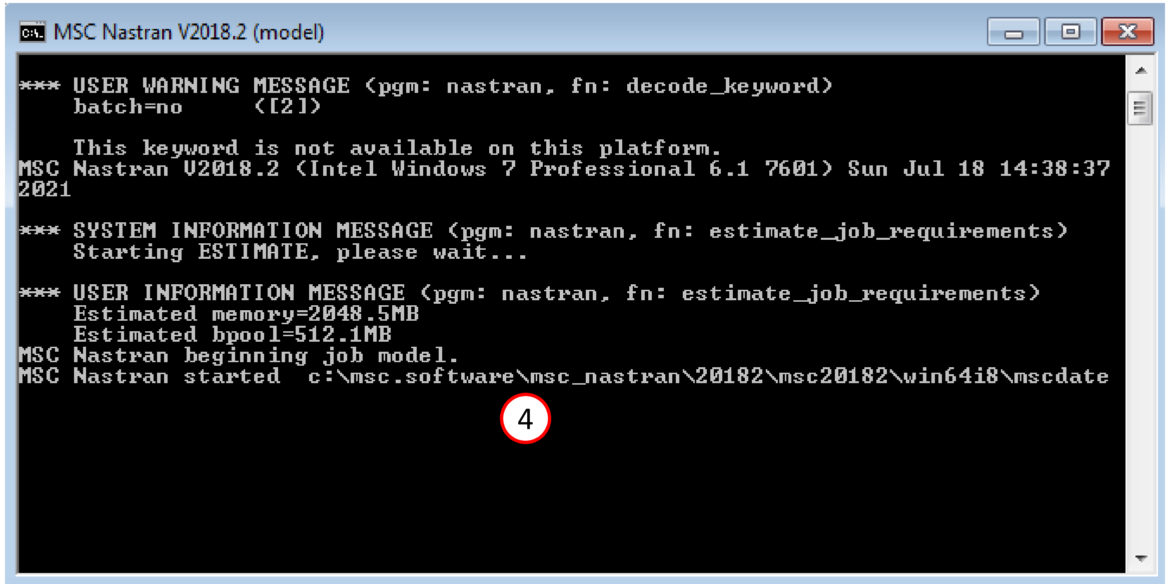
Follow these instructions:
1) Open Terminal
2) Navigate to the nastran_working_directory
   cd ./nastran_working_directory
3) Use this command to start the process
   ./Start_MSC_Nastran.sh

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

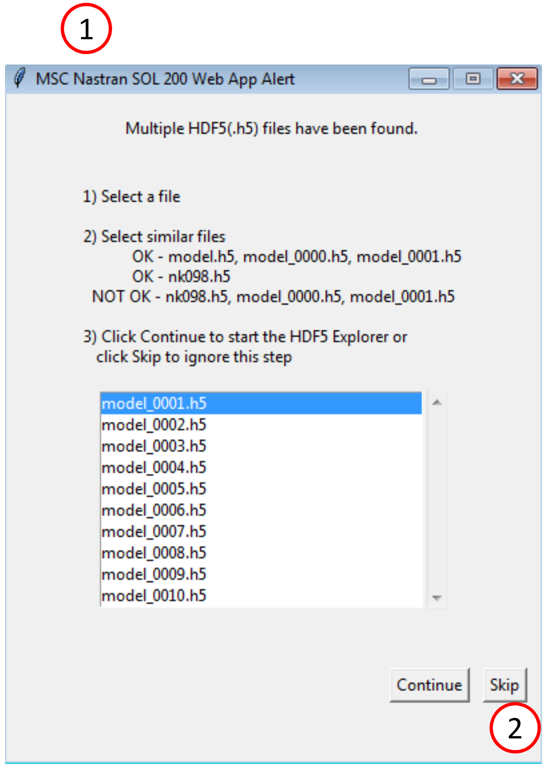
Python   MSC Nastran

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



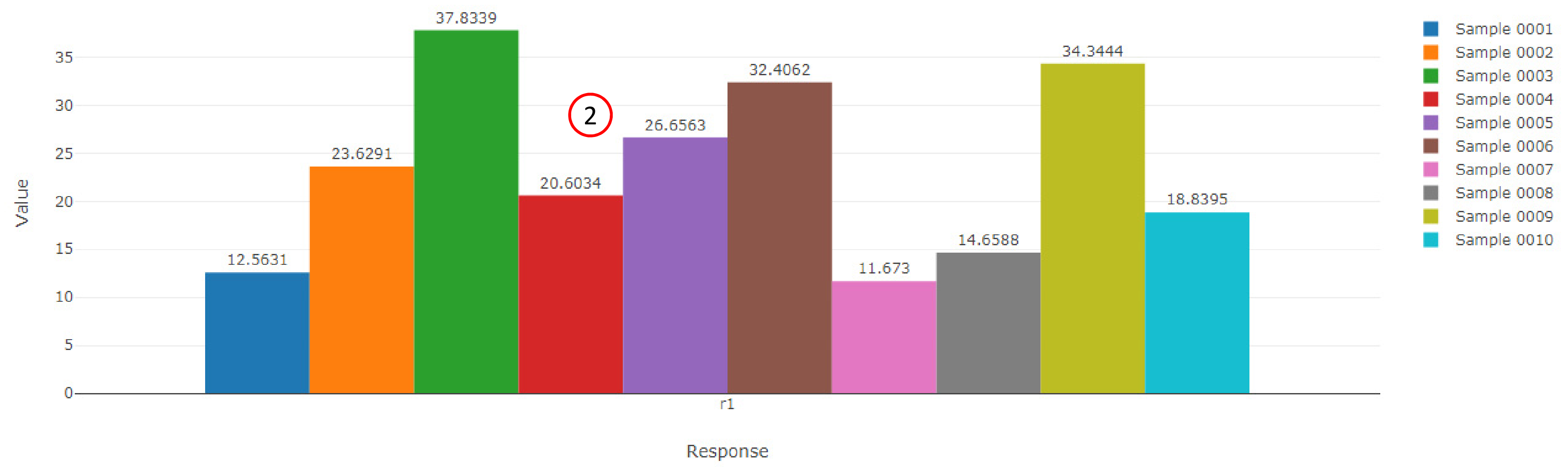
Review Results

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.

A. The table titled Monitored Response can be interacted with. Each column in the table contains filters. Once a filter is modified, the Bar Chart will instantly update.
B. Additional functions include the ability to highlight the MAX and MIN bars, download a CSV file and reset the filters.

SOL 200 Web App - Monitored Responses 1

Home



B

Monitored Responses Display MAX and MIN | Download CSV | Reset Filters

Label	Dataset Name	Field	Field Description	Current Value
r1	SUMMARY/EIGENVALUE A	EIGEN		
r1	SUMMARY/EIGENVALUE	EIGEN		10.001500337521877

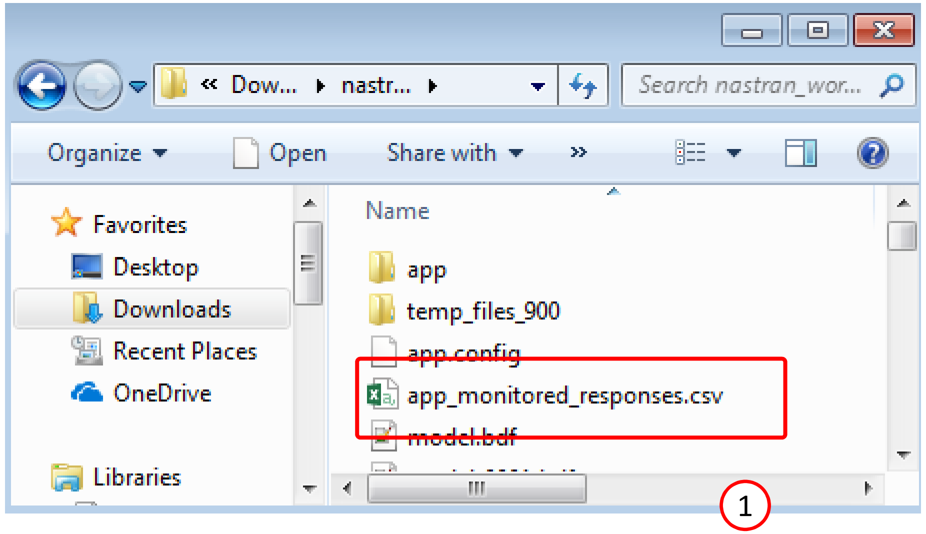
Monitored Responses from Each Sample 3

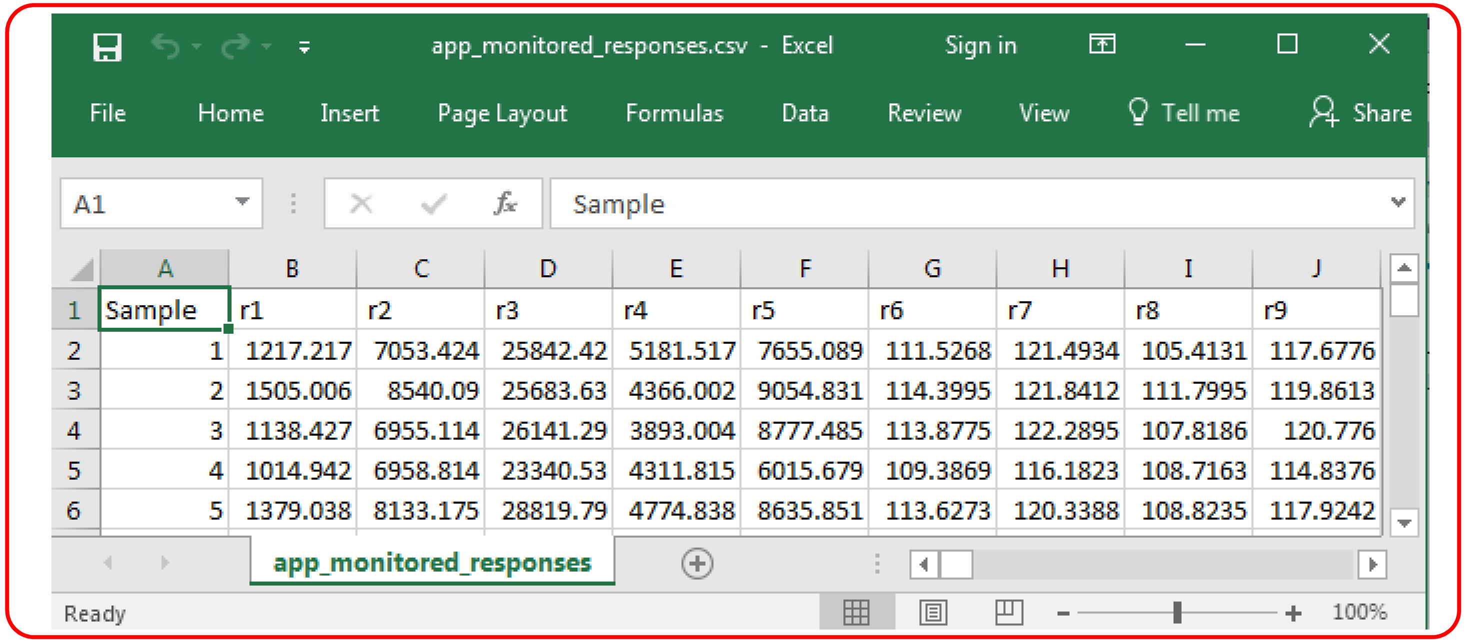
0001	0002	0003	0004	
12.563078819847316	23.62907115939849	37.8338533140817	20.60335328534742	26.6

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 use to train the surrogate model.

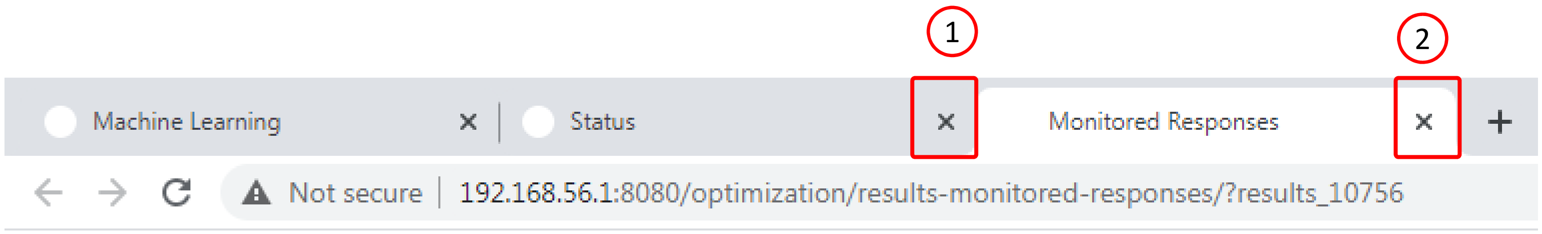




	A	B	C	D	E	F	G	H	I	J
1	Sample	r1	r2	r3	r4	r5	r6	r7	r8	r9
2	1	1217.217	7053.424	25842.42	5181.517	7655.089	111.5268	121.4934	105.4131	117.6776
3	2	1505.006	8540.09	25683.63	4366.002	9054.831	114.3995	121.8412	111.7995	119.8613
4	3	1138.427	6955.114	26141.29	3893.004	8777.485	113.8775	122.2895	107.8186	120.776
5	4	1014.942	6958.814	23340.53	4311.815	6015.679	109.3869	116.1823	108.7163	114.8376
6	5	1379.038	8133.175	28819.79	4774.838	8635.851	113.6273	120.3388	108.8235	117.9242

Close Pages

1. The Status page can be closed
2. The Monitored Responses page can be closed



Start Desktop App

1. Open this folder: nastran_working_directory (1)
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

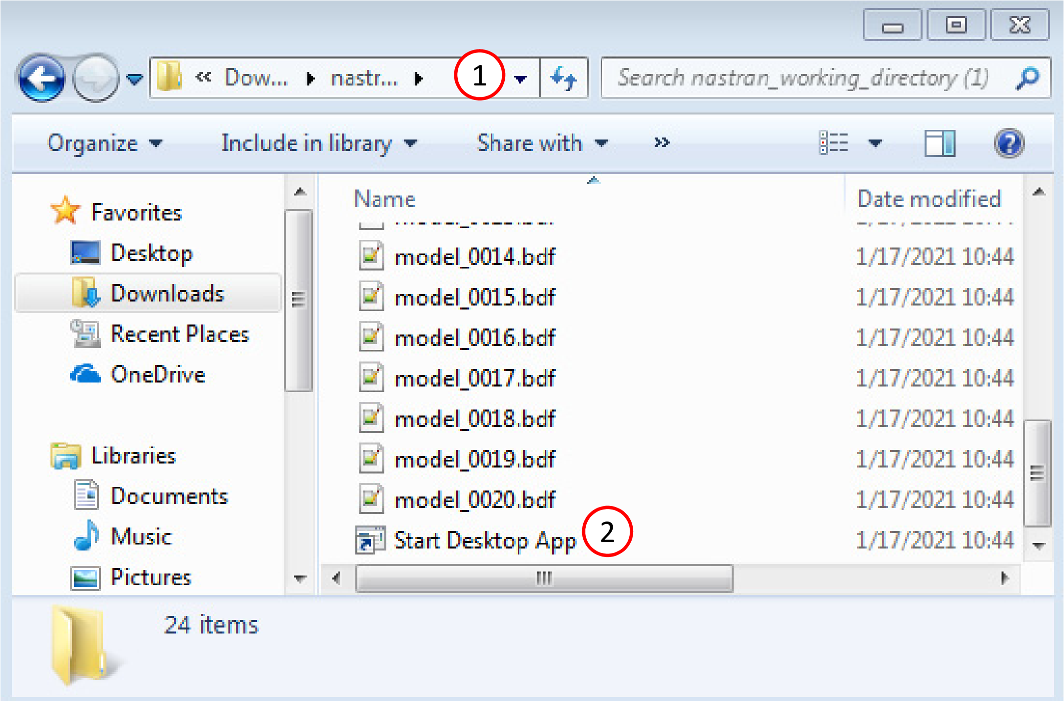
- One can run the Nastran job on a remote machine as follows: 1) Copy the BDF files and the INCLUDE files to a remote machine. 2) Run the MSC Nastran job on the remote machine. 3) After completion, copy the BDF, F06, LOG, H5 files to the local machine. 4) Click "Start Desktop App" to display the results.

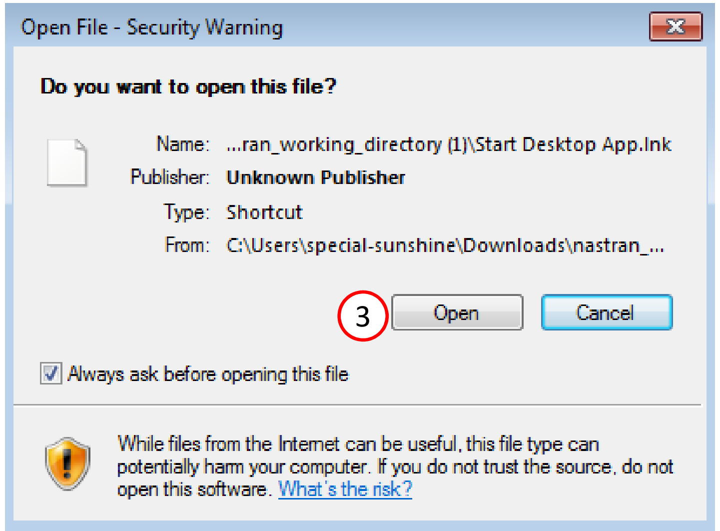
Using Linux?

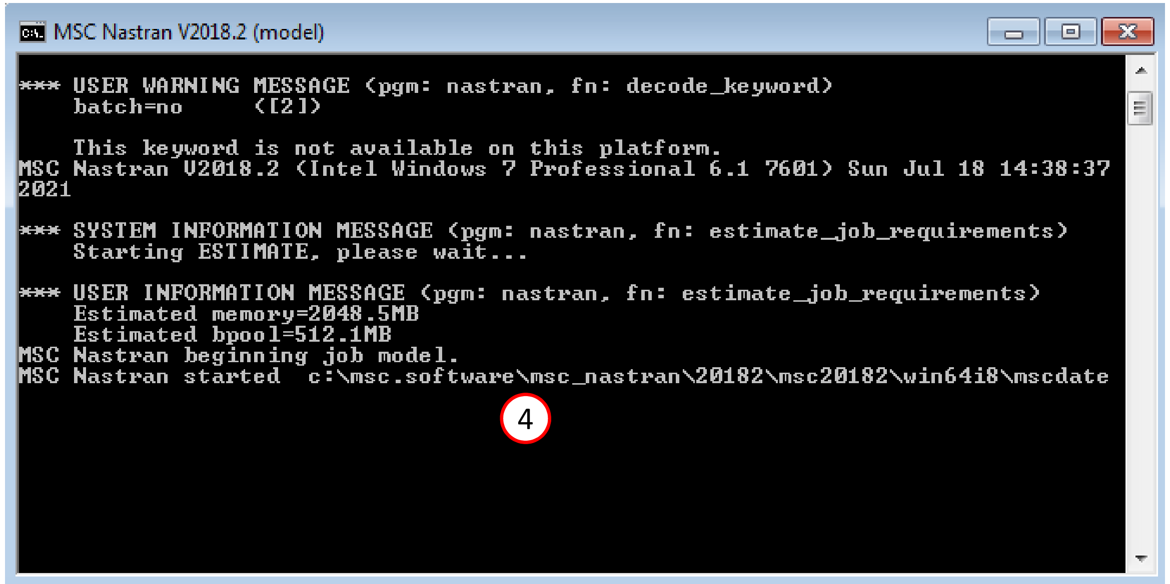
Follow these instructions:
1) Open Terminal
2) Navigate to the nastran_working_directory
cd ./nastran_working_directory
3) Use this command to start the process
./Start_MSC_Nastran.sh

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

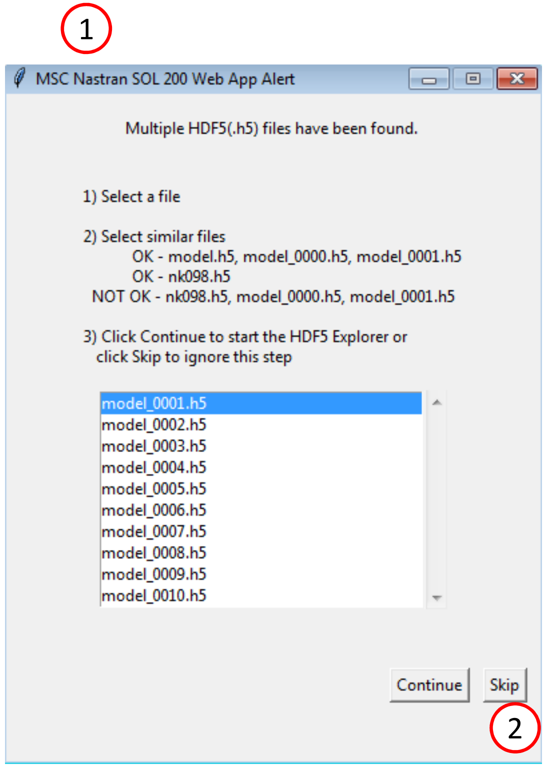
Python MSC Nastran

Status

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

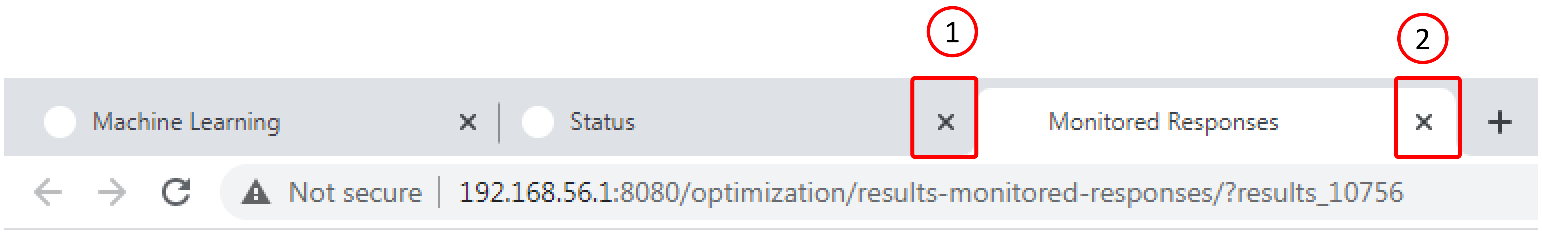
Review Results

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



Close Pages

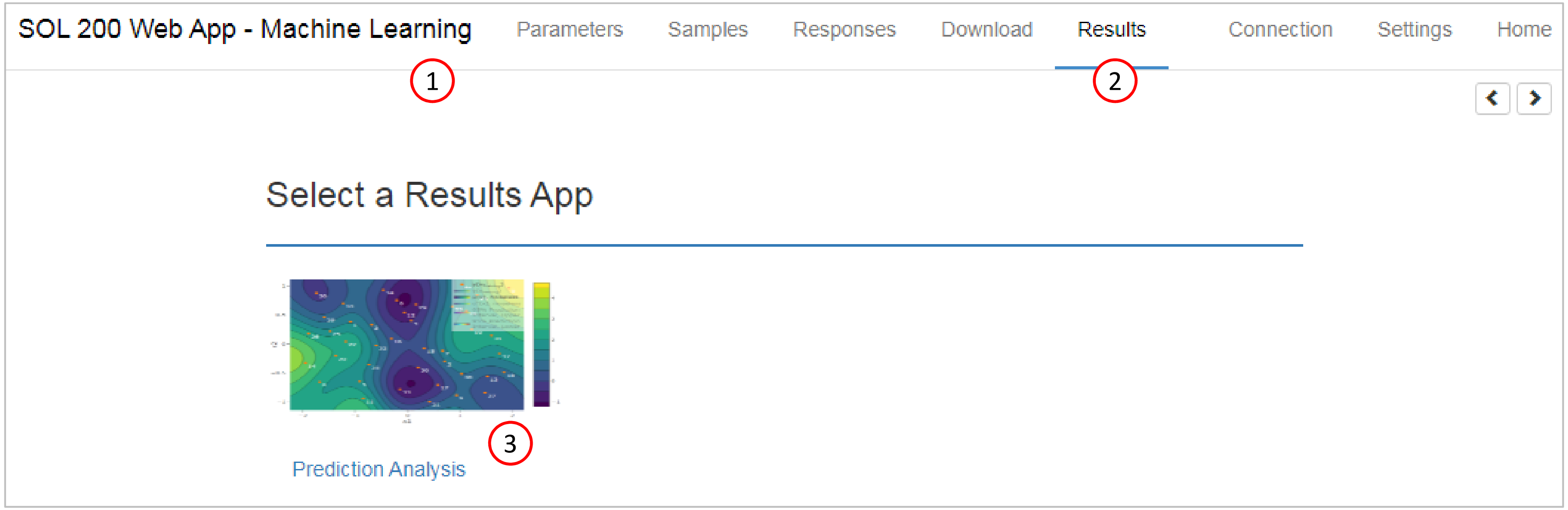
1. The Status page can be closed
2. The Monitored Responses page can be closed

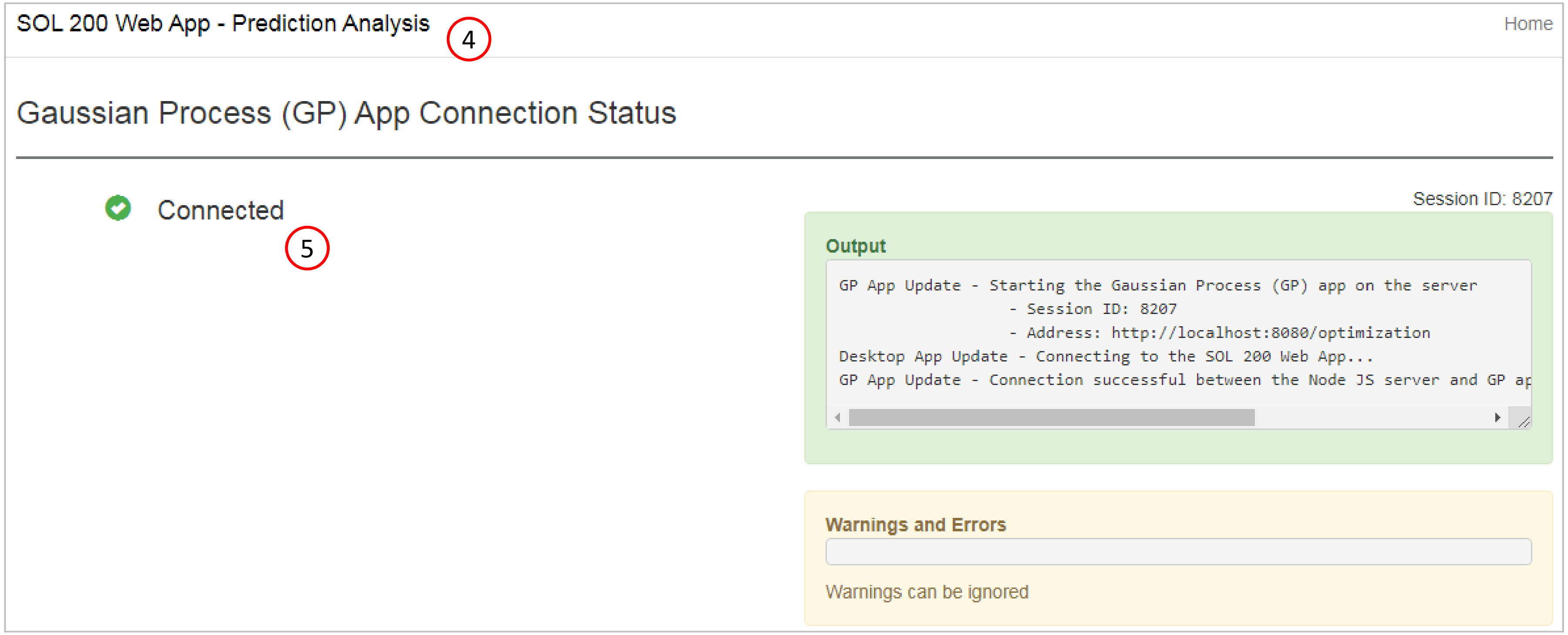


Performing Predictions

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

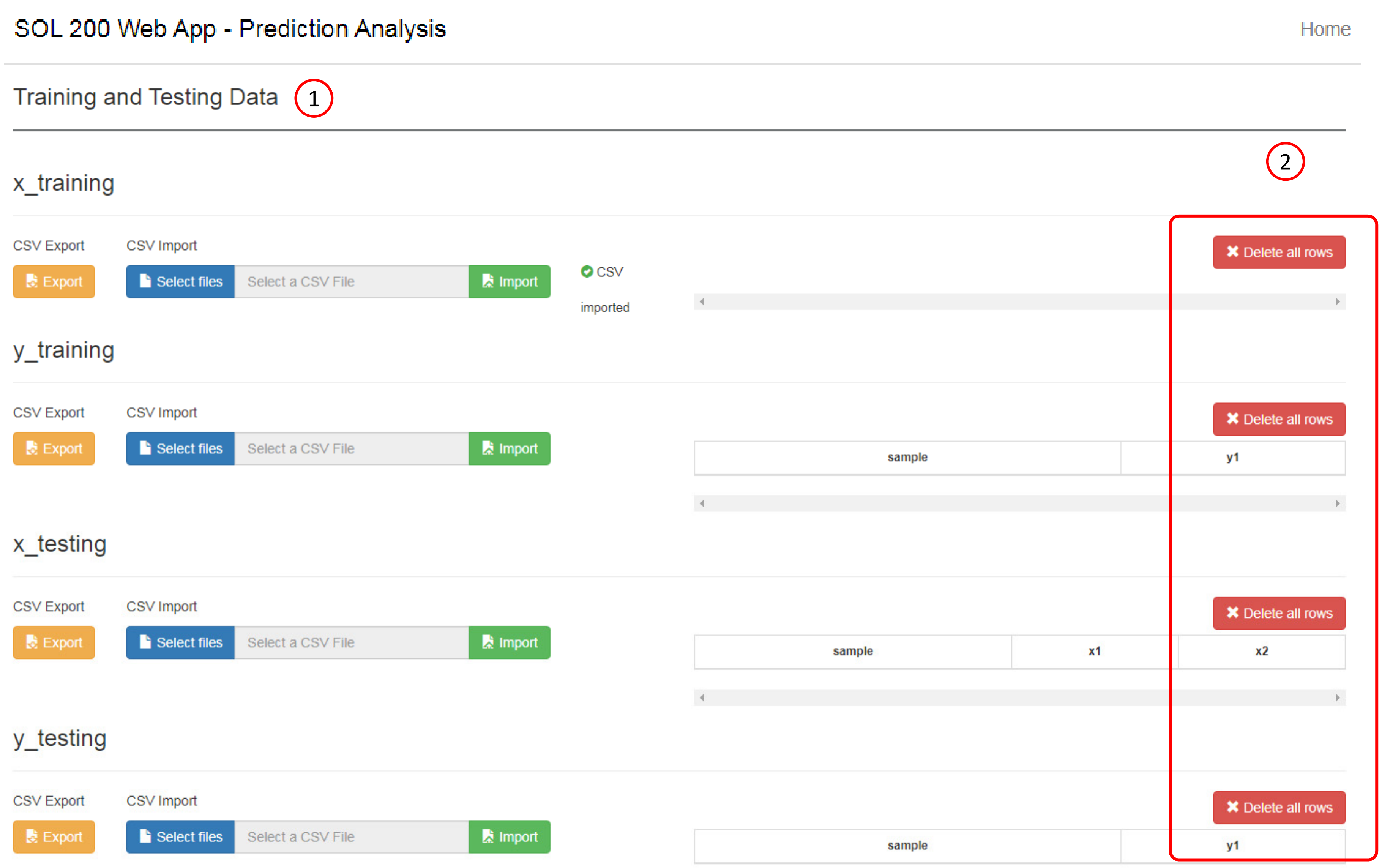




Training Data

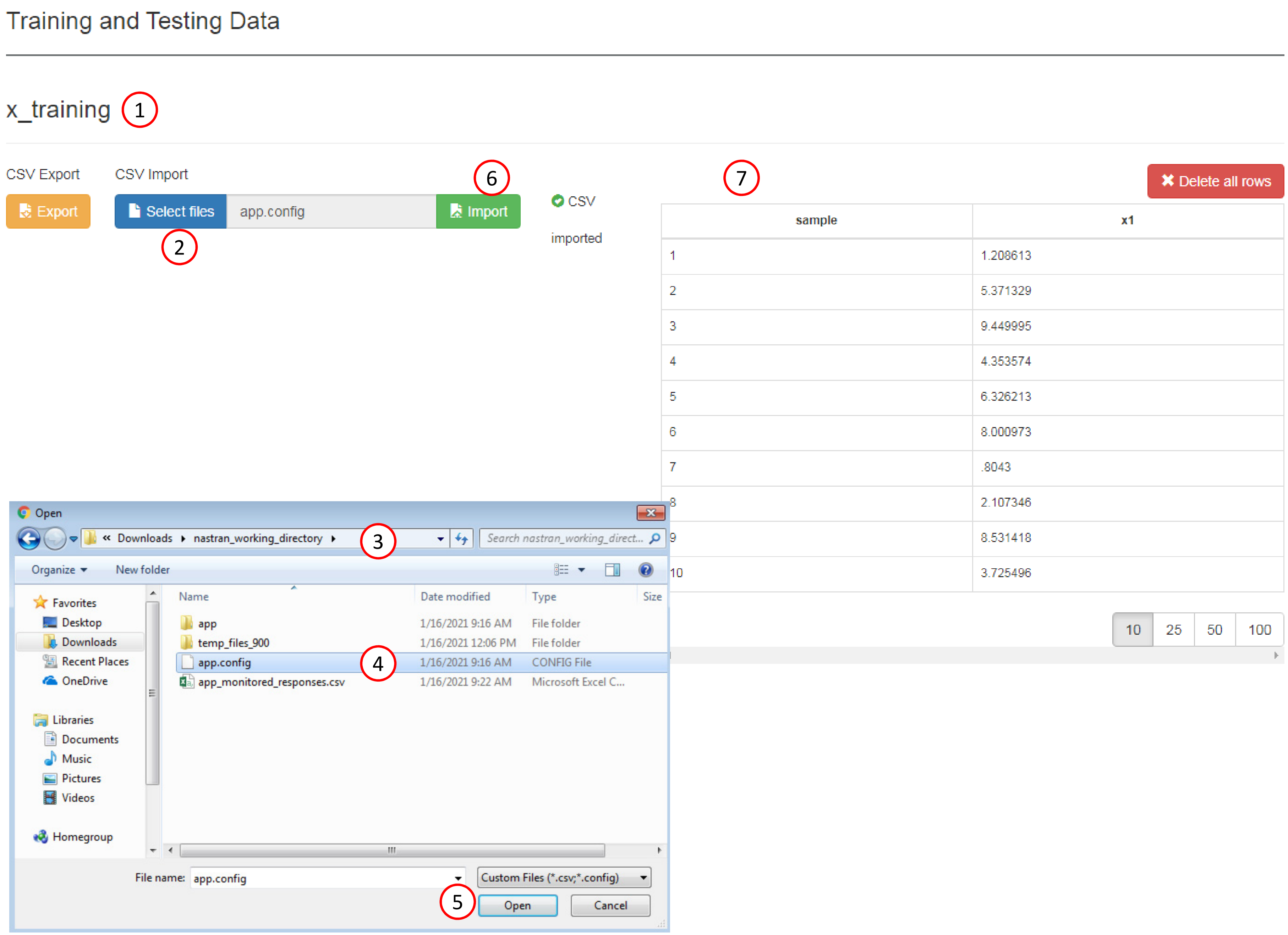
1. Navigate to the Training and Testing Data section
2. Delete any previous table data by clicking the four (4) buttons named Delete all rows

- **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.



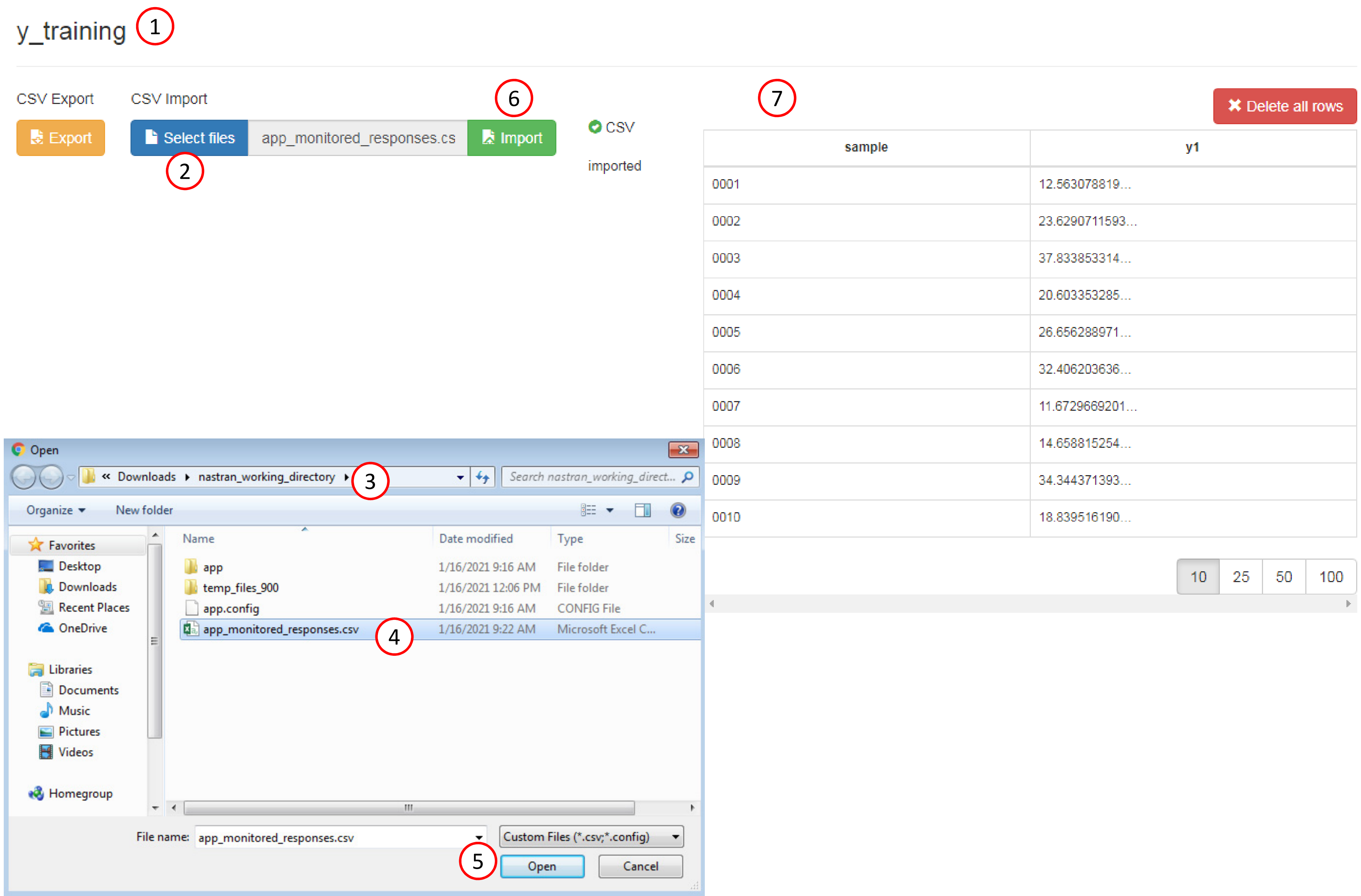
Training Data

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



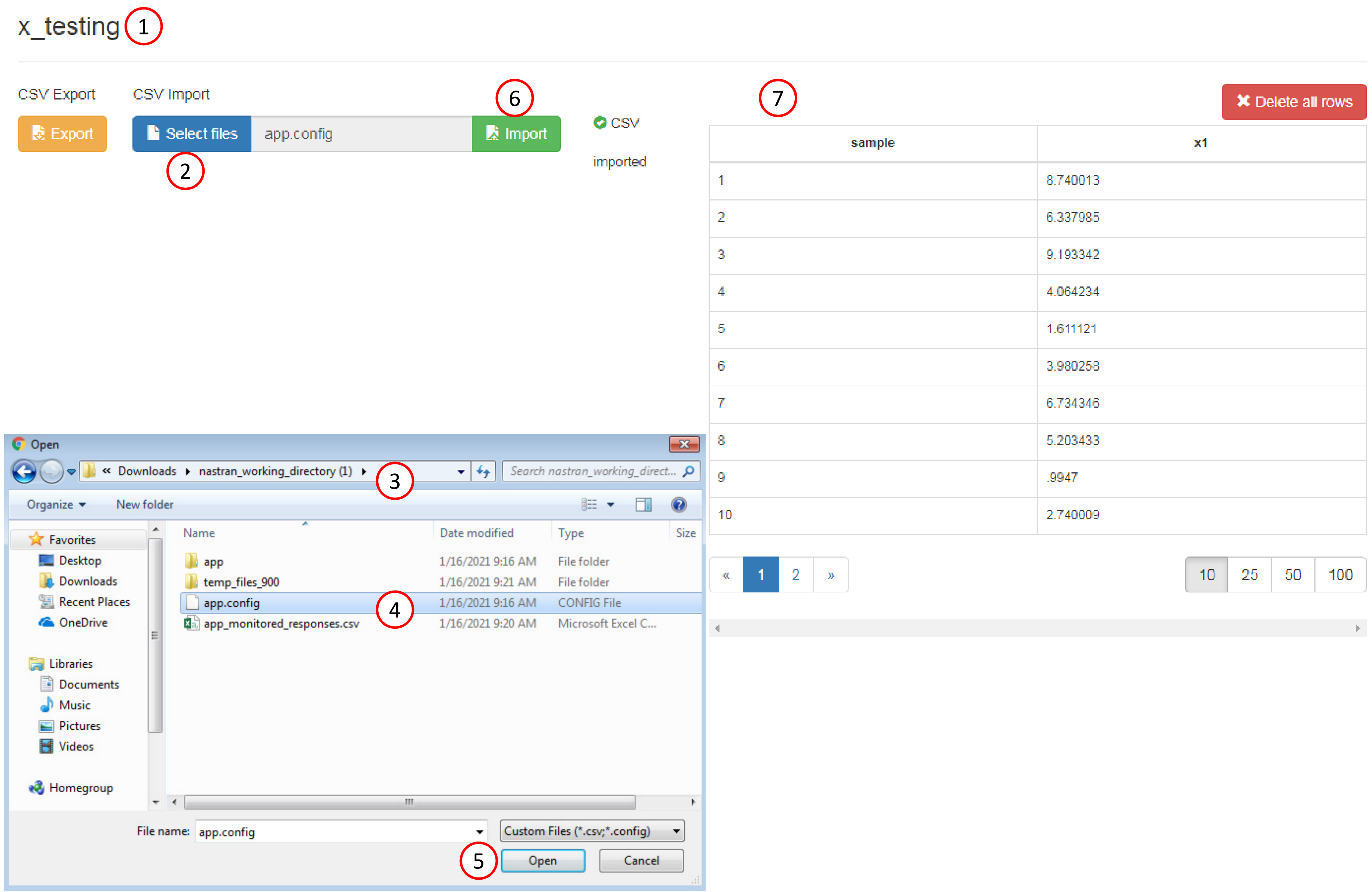
Training Data

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



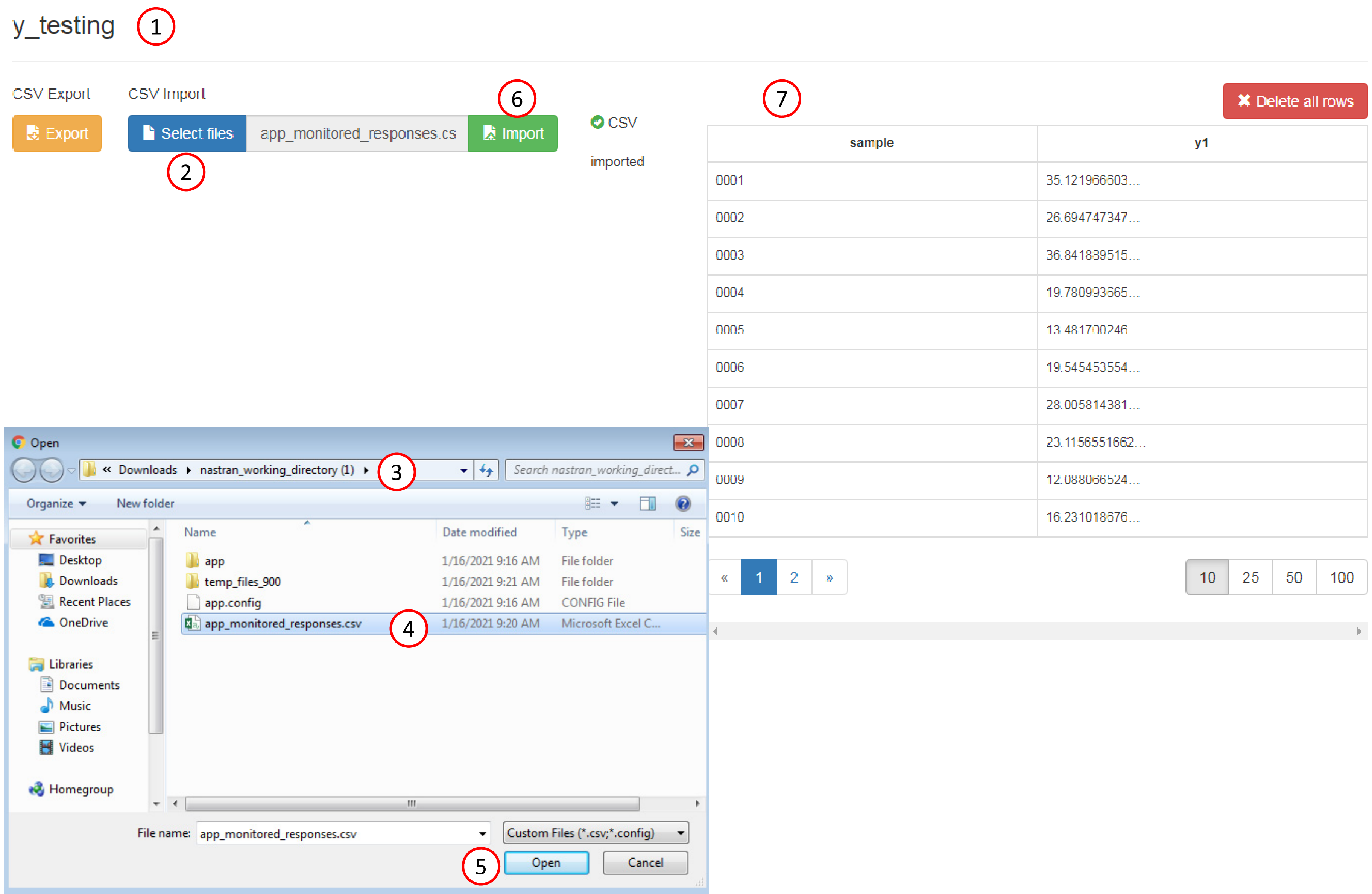
Testing Data

1. Navigate to the section titled x_testing
2. Click Select files
3. Navigate to the folder named nastran_working_directory (1) which contains data for 20 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 20 runs



Testing Data

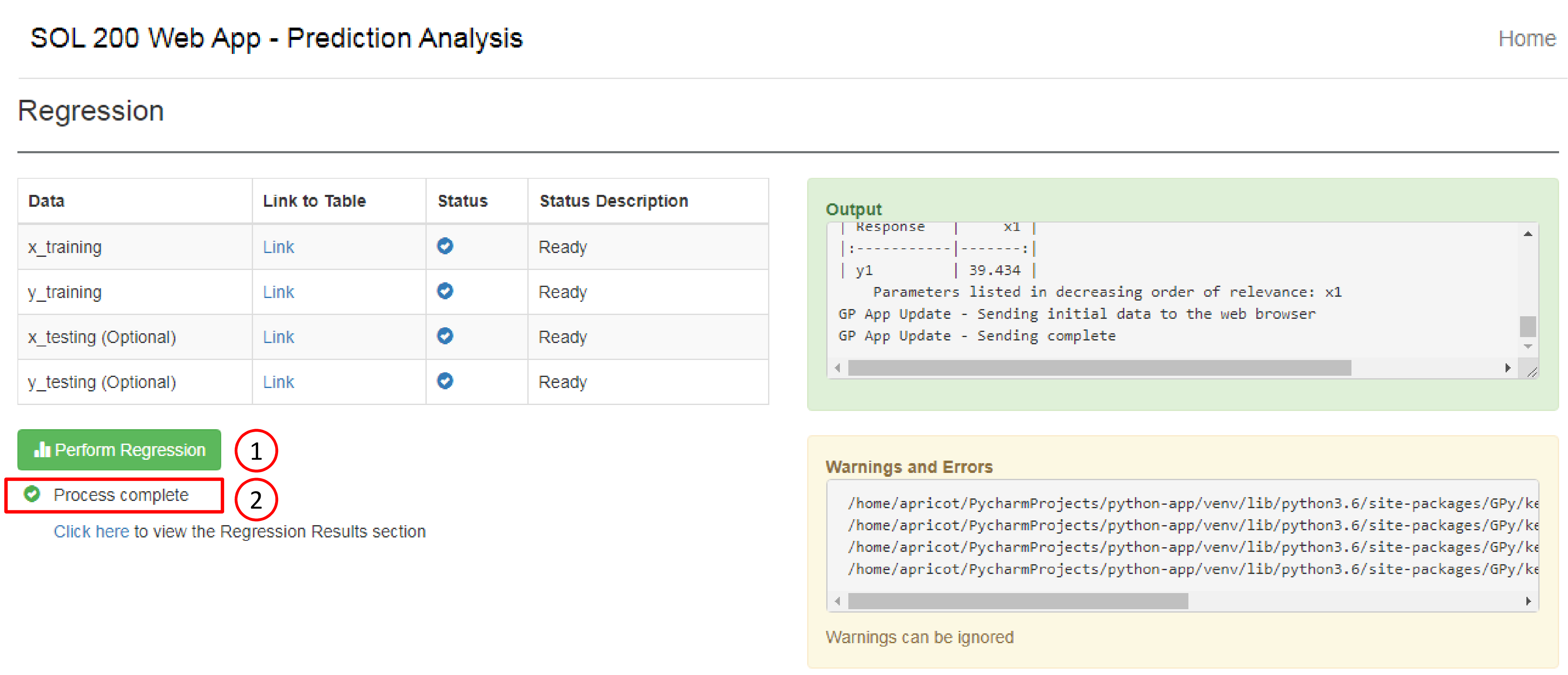
1. Navigate to the section titled y_testing
2. Click Select files
3. Navigate to the folder named nastran_working_directory (1) which contains data for 20 runs
4. Select the file app_monitored_responses.csv
5. Click Open
6. Click Import
7. The table is now loaded with the y outputs (monitored responses) for all 20 runs



Perform Regression

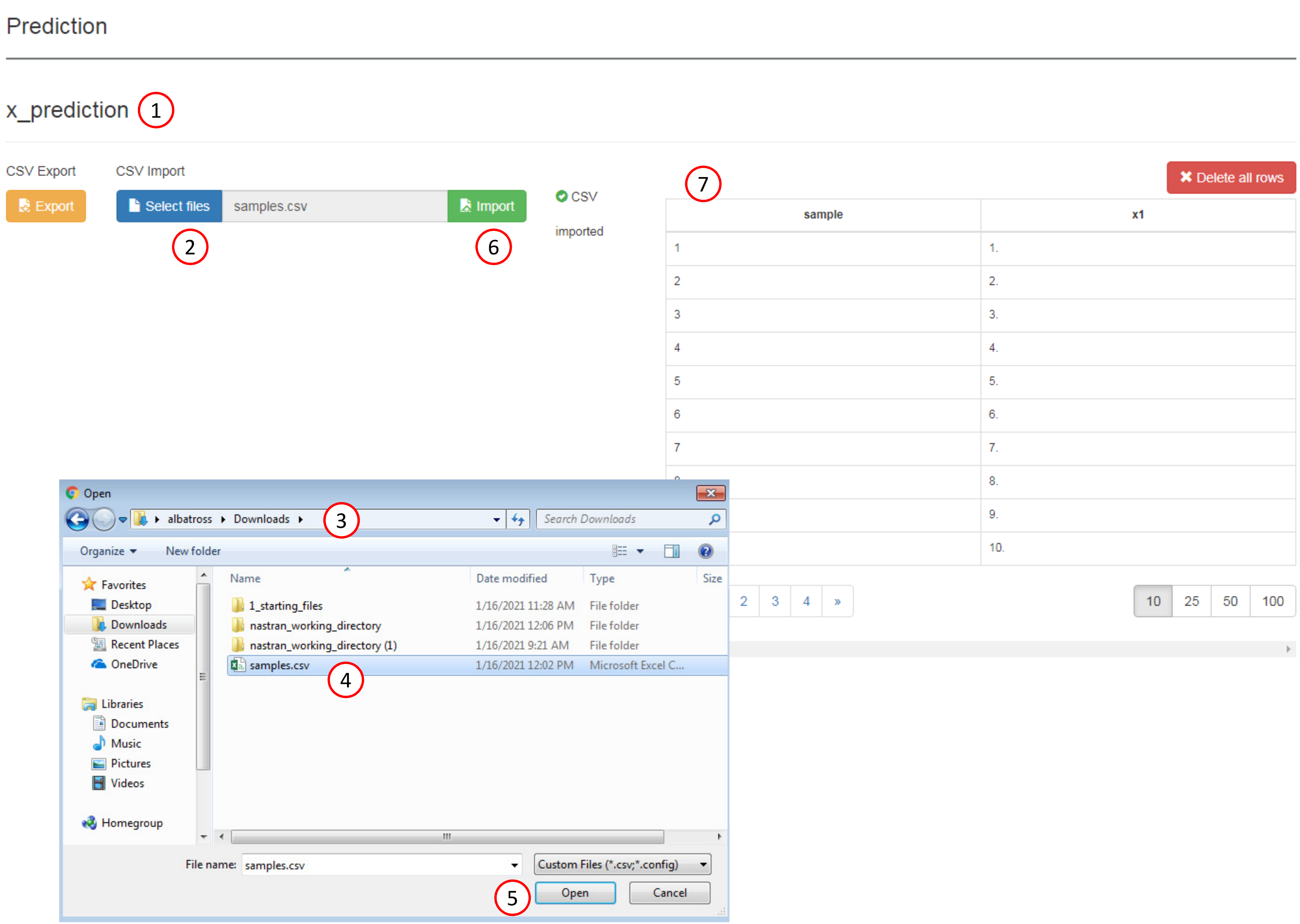
1. Click Perform Regression and the surrogate model will be fitted

2. The regression is complete when the following status message is visible:
    - Process complete



Perform Prediction

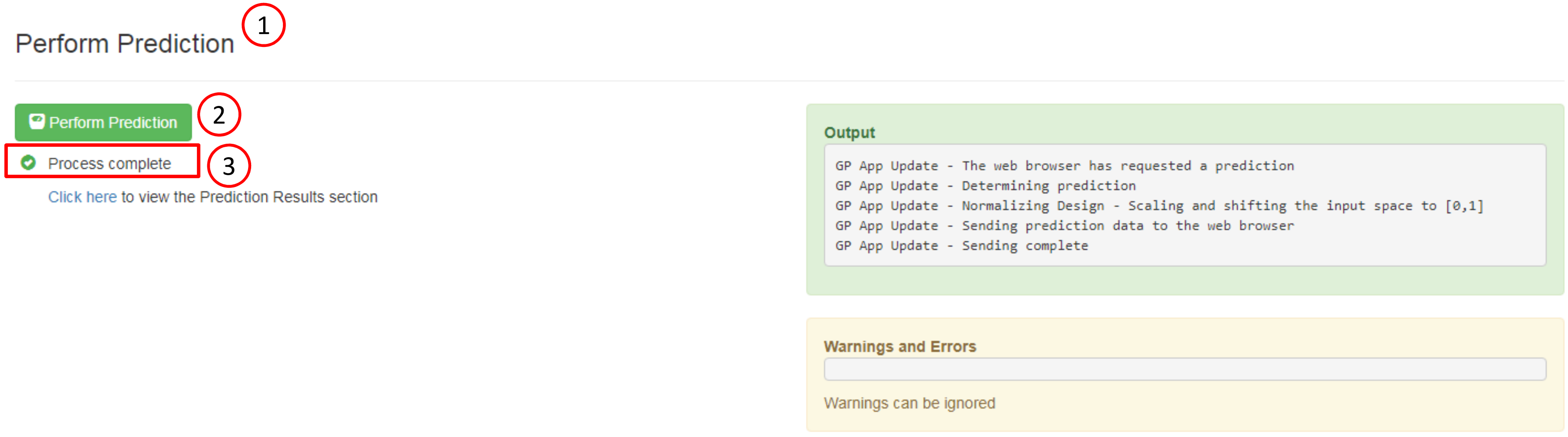
1. Navigate to the section titled x_prediction
2. Click Select files
3. Navigate to the location of the file named samples.csv
4. Select the file samples.csv
5. Click Open
6. Click Import
7. The table is now loaded with the x inputs for 40 runs



Perform Prediction

1. Navigate to the section titled Perform Prediction
2. Click Perform Prediction
3. The prediction is complete when the following status message is visible:
   - Process complete

- Note that the predictions are performed seemingly instantly



Variance

1. Navigate to the section titled Variance
2. Note the normalized root mean square error (NRMSE)
   - NRMSE values are only calculated if x_testing and y_testing are provided
   - NRMSE values less than .15 indicate the surrogate model has good prediction performance

- **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.

Variance

1

Matern52

2

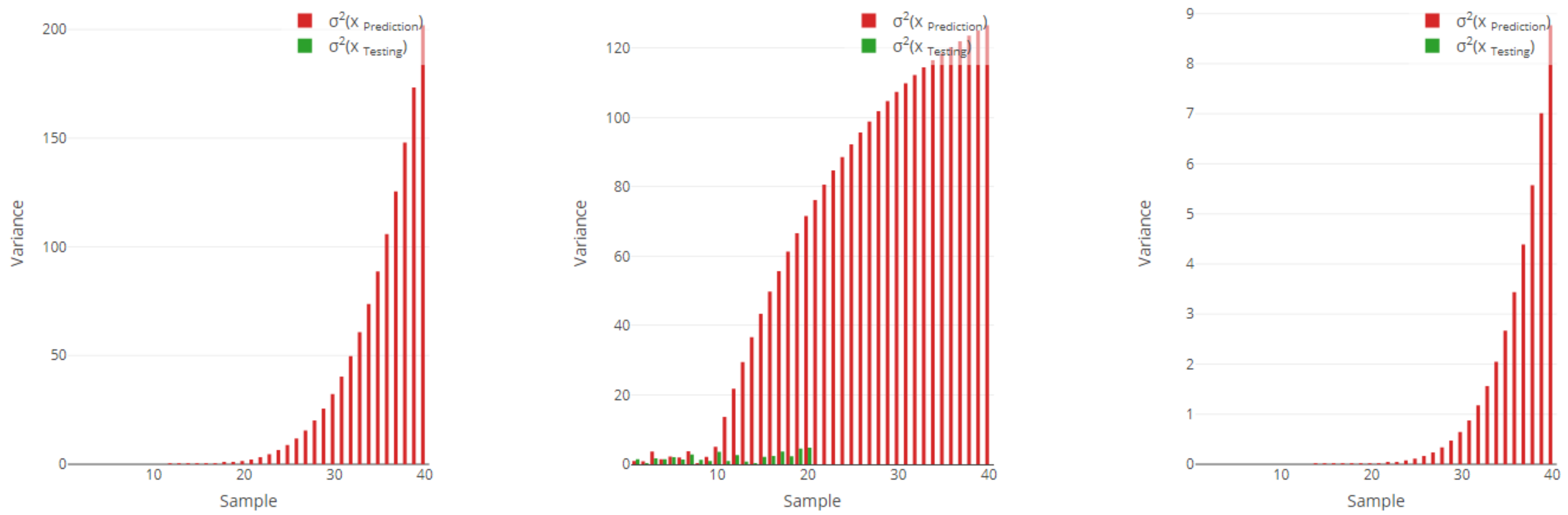
NRMSE: 8.863379504119915e-06

Exponential

NRMSE: 0.019861846267229557

RBF

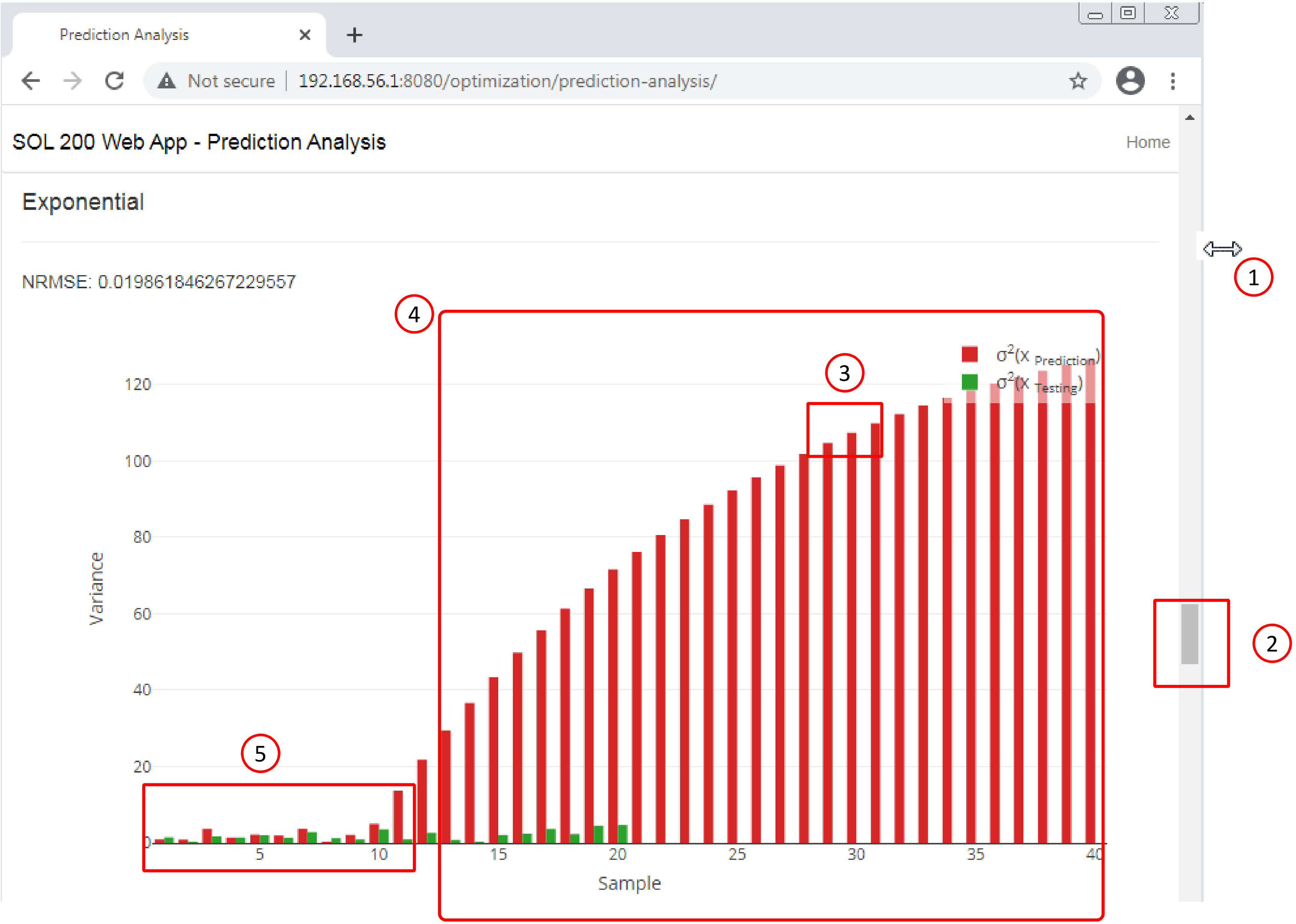
NRMSE: 1.2399877479780946e-05



Variance

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. A high bar indicates a high prediction uncertainty, or a high variance, and is indication that we do not have enough information to conclude the prediction is credible at that prediction point.
4. Note the variance (prediction uncertainty) is high for many of the prediction points (red bars). The x inputs of the original training data was within the bounds of .01 and 10.0, so the surrogate model is valid within these bounds. The prediction points are within the bounds of 1.0 and 40.0 Predictions made outside of these bounds result in high variance (high prediction uncertainty).
5. The first 10 samples of the prediction points are within the bounds of the original training data. The variance values are relatively low.

- 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

If the problem statement has 1 or 2 parameters configured, a response surface is created. For 3 or more parameters, a response surface is NOT created.

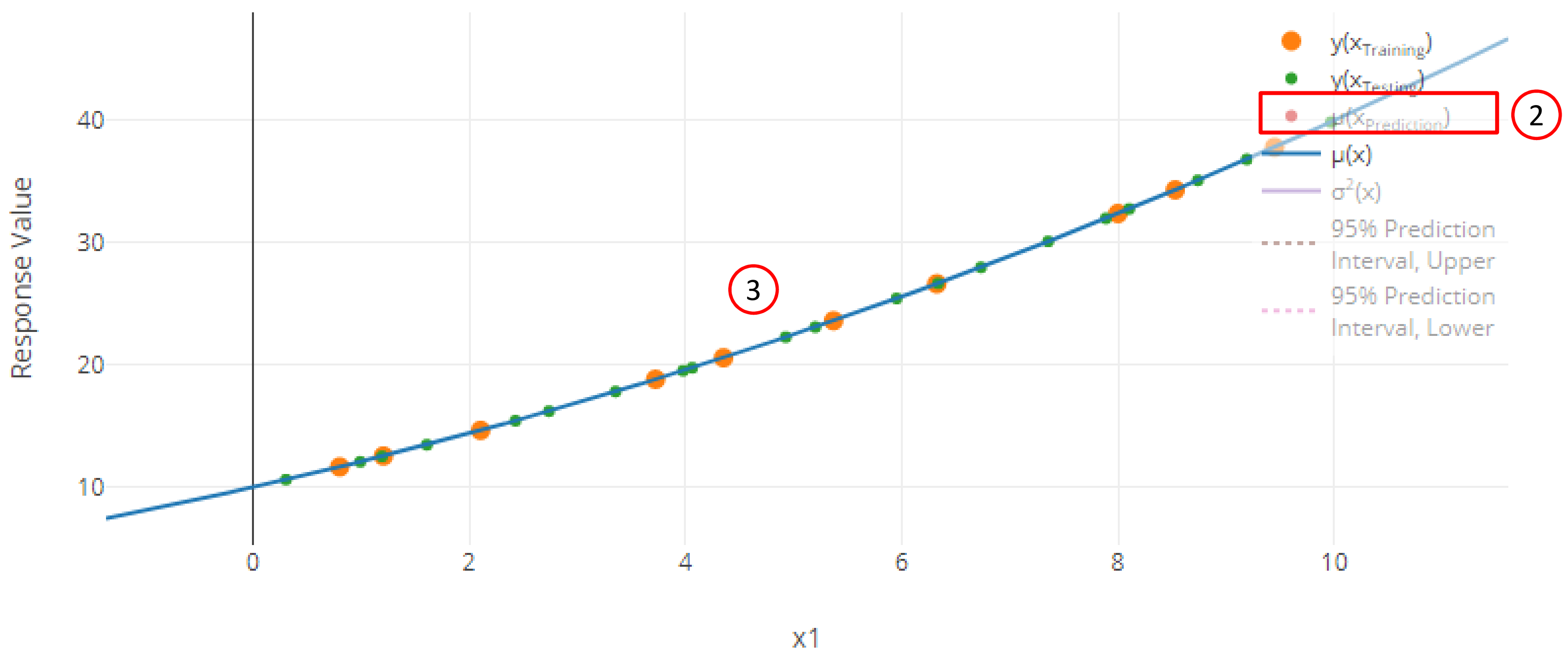
1. Navigate to the section titled Response Surface
2. Deselect the following legend entry:
   - $\mu(x_{Prediction})$
3. The plot is updated
   - $y(x_{Training})$: MSC Nastran responses used to train the surrogate model
   - $y(x_{Testing})$: MSC Nastran responses used to compute the NRMSE
   - $\mu(x_{Prediction})$: The predicted values at points x_prediction
   - $\mu(x)$: The predicted function

Regression Results

Response Surface

1

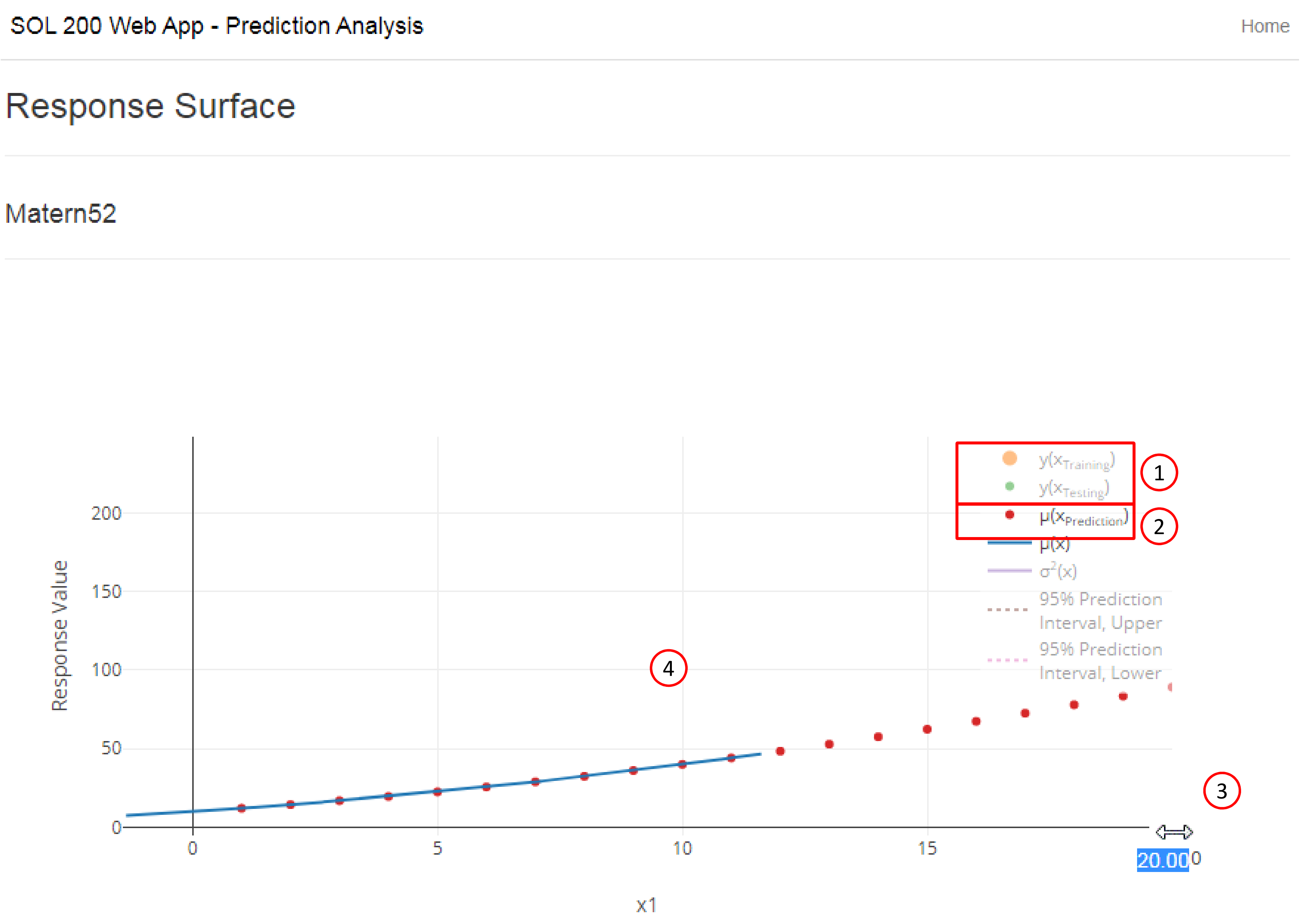
Matern52



Response Surface

1. Deselect the following legend entries
   - $y(x_{Training})$
   - $y(x_{Testing})$
2. Select the following legend
   - $\mu(x_{Prediction})$
3. Move the mouse cursor to the right most number label of the horizontal axis, single click the number, type in a value of 20.0, and press Enter. This will cause the horizontal axis to be scaled to be within 0 and 20.0.
4. The plot is updated

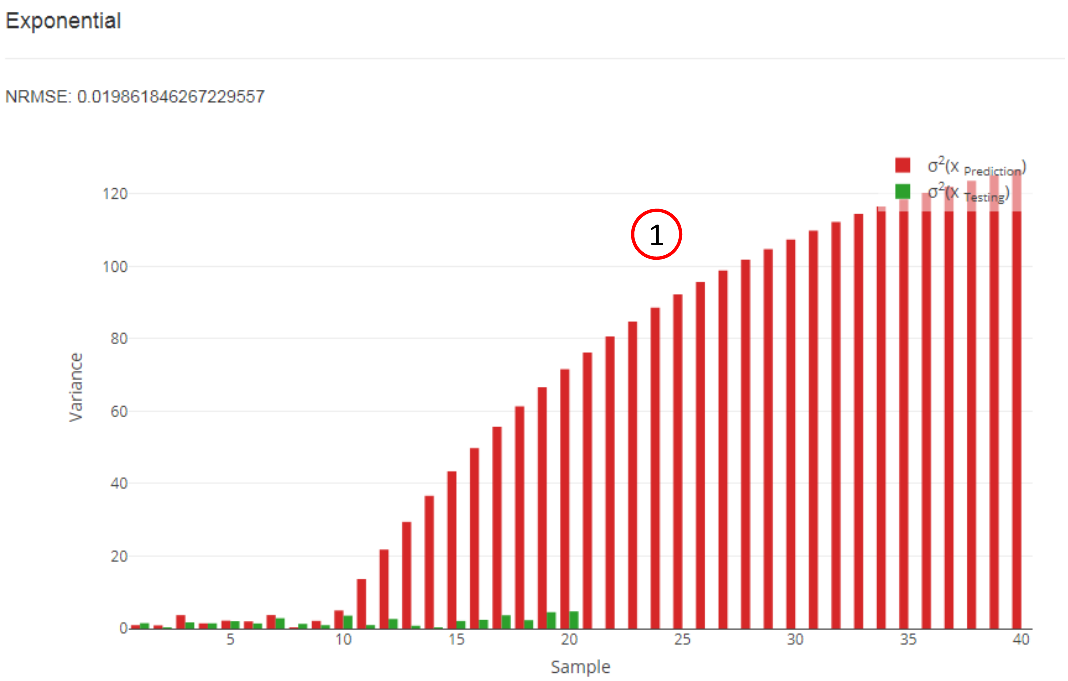
The training data used to train the surrogate model was within .1 and 10.0. Predictions should be made within these bounds. Predictions outside the bounds should NOT be performed.

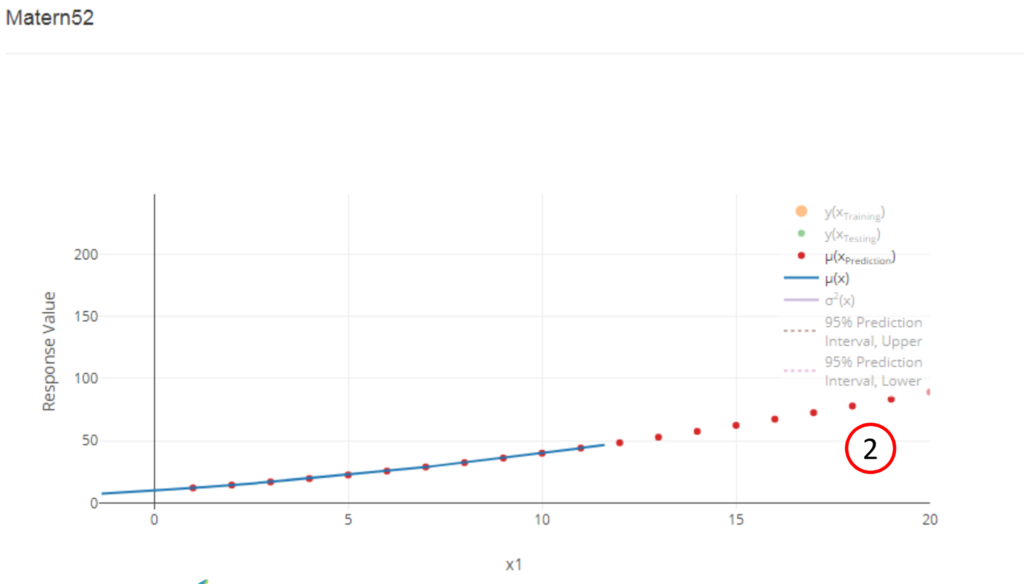


Response Surface

The training data used to train the surrogate model was within .1 and 10.0. Predictions should be made within these bounds. Predictions outside the bounds should NOT be performed.

1. The variance (prediction uncertainty) plot shows that the prediction uncertainty increases as you travel away from the bounds
2. To summarize
   - OK – Predictions are made within the bounds.
   - OK – Predictions are made close to the bounds. In some cases, the surrogate model has good prediction performance for points nearby the bounds.
   - NOT OK – Predictions are made far outside the bounds.

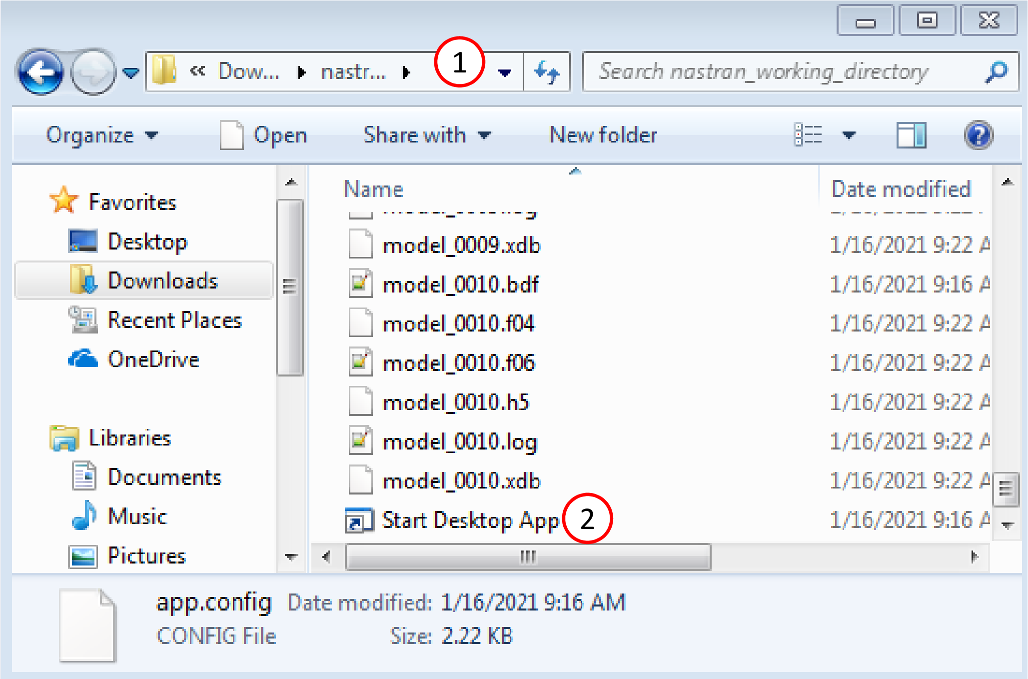


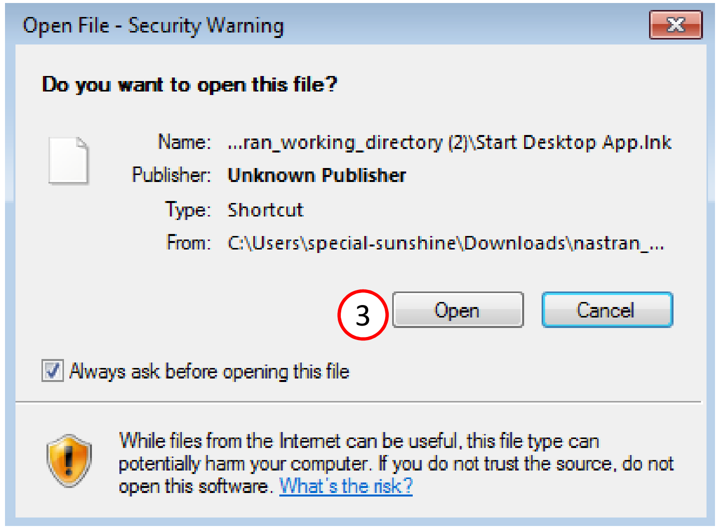


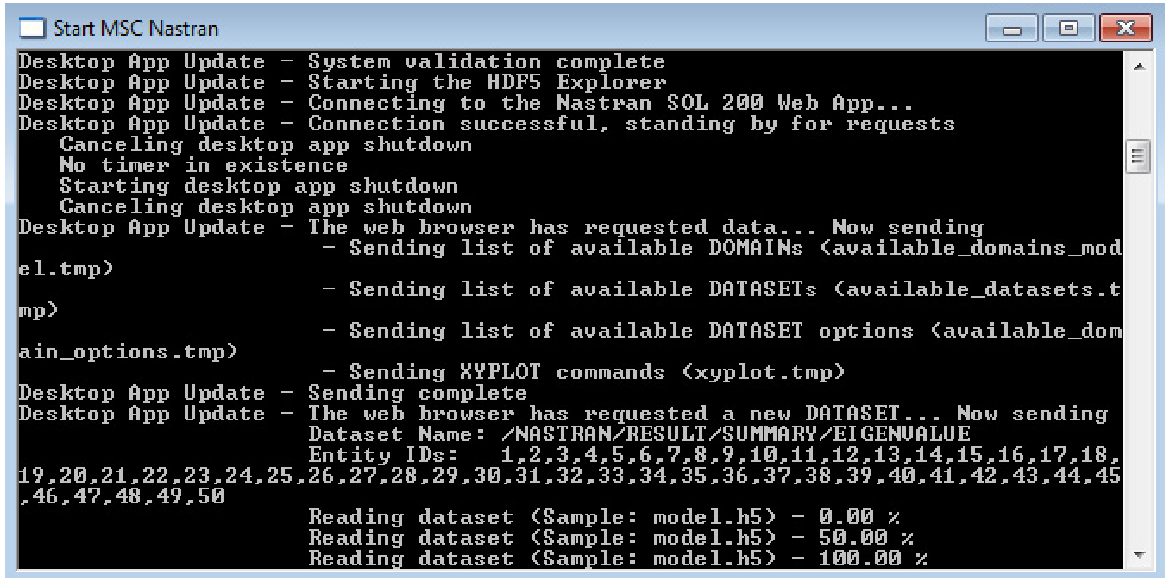
Creating Plots with the HDF5 Explorer

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

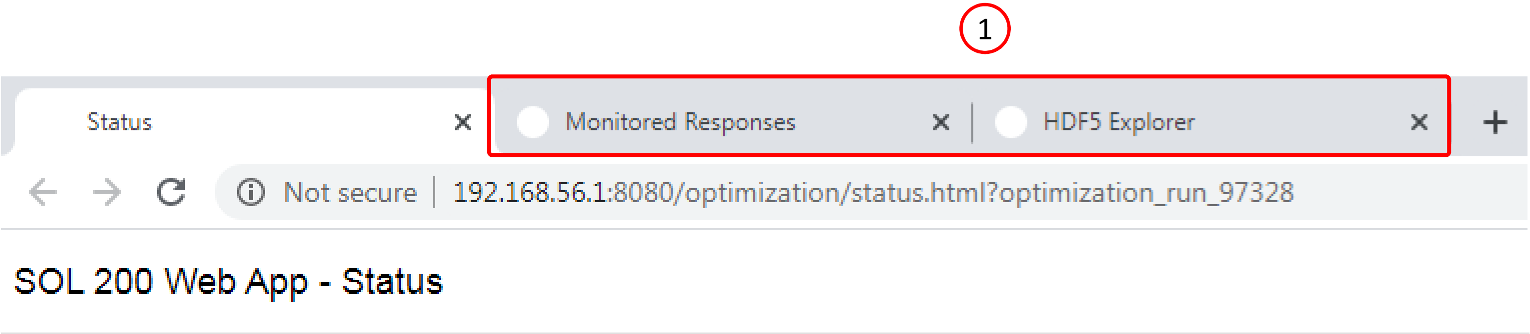
In some instances, execute permission must be granted to the directory. Use this command. This command assumes you are one folder level up.

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	• The response value from each sample can be compared.	• 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	• This web app is used to probe each H5 file and generate XY plots.	

Review Results

After MSC Nastran is finished, the results will be automatically uploaded.

1. On this page, 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.

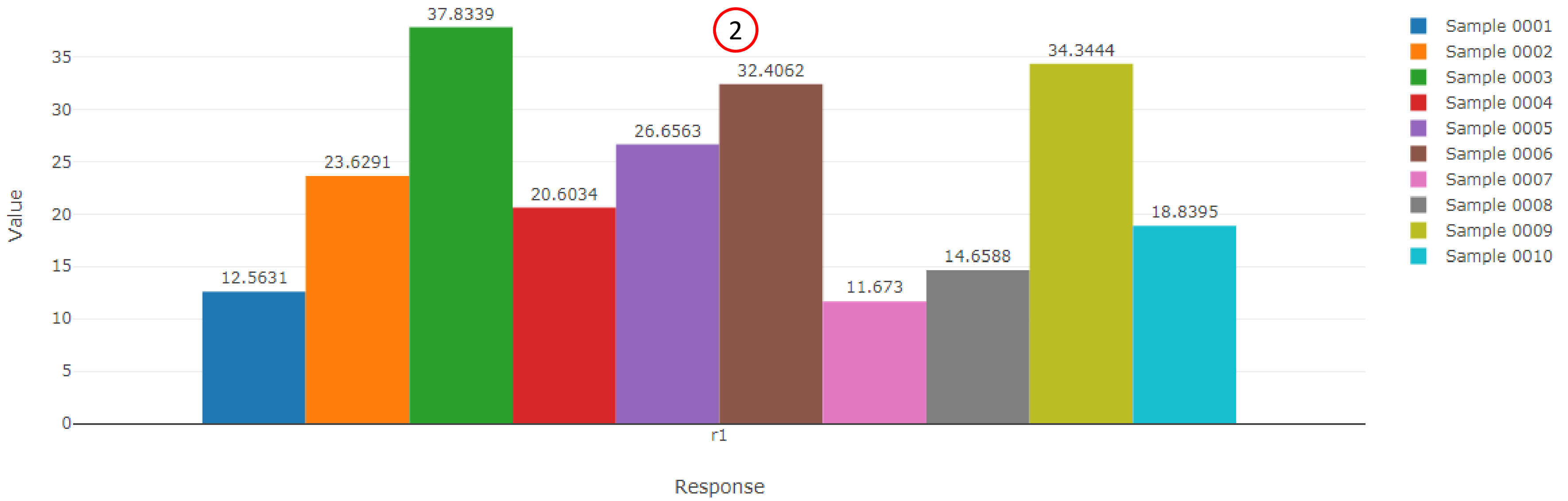
A. Additional functions include the ability to highlight the MAX and MIN bars, download a CSV file and reset the filters.

1

SOL 200 Web App - Monitored Responses

Home

2



A

Monitored Responses Display MAX and MIN | Download CSV | Reset Filters

Label	Dataset Name	Field	Field Description	Current Value
r1	SUMMARY/EIGENVALUE	EIGEN		
r1	SUMMARY/EIGENVALUE	EIGEN		10.001500337521877

3

Monitored Responses from Each Sample

0001	0002	0003	0004	
12.563078819847316	23.62907115939849	37.8338533140817	20.60335328534742	26.656

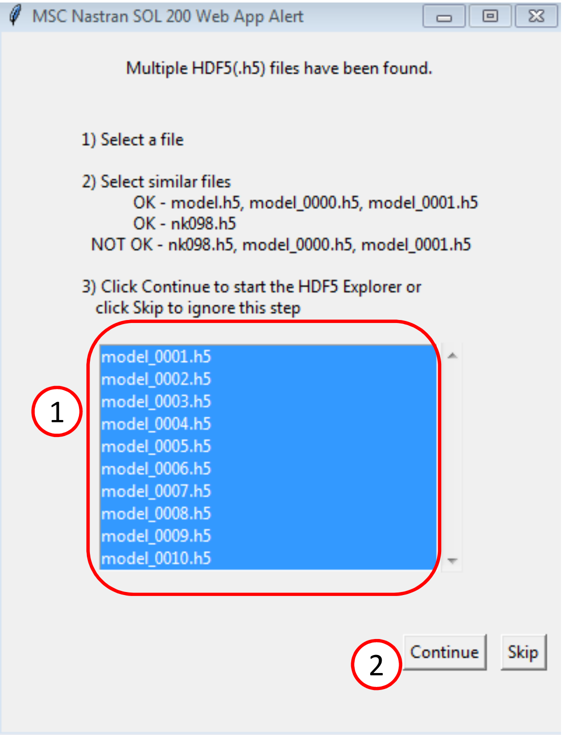
Review Results

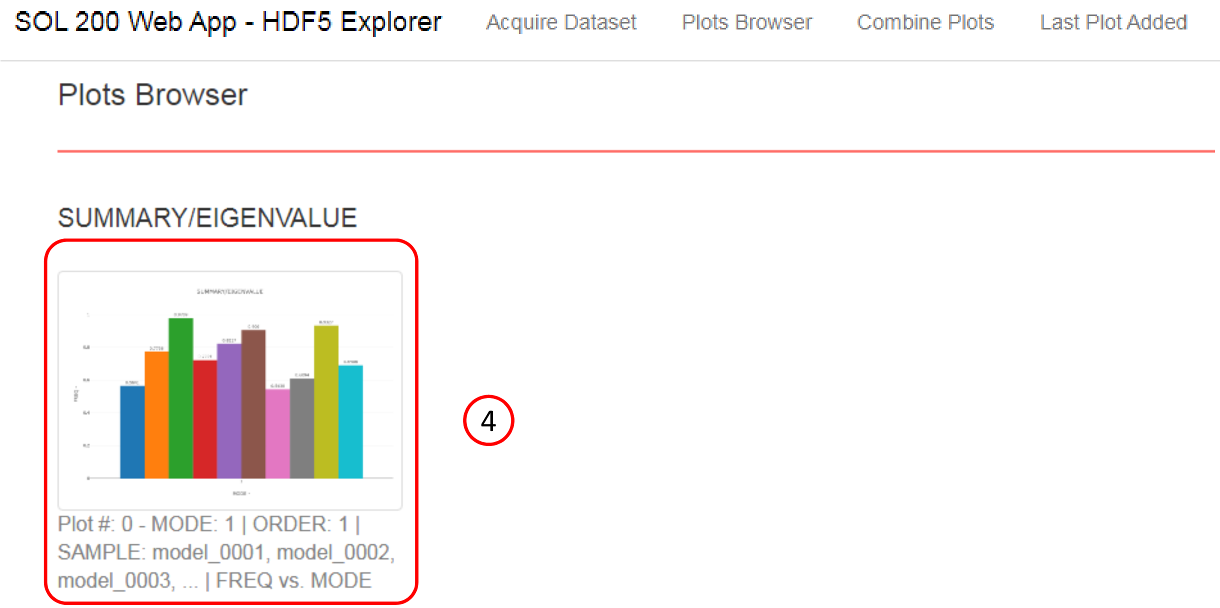
On this page, the HDF5 Explorer is opened.

1. Select all the files
2. Click Continue
3. The HDF5 Explorer is automatically opened.
4. Click the image

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.

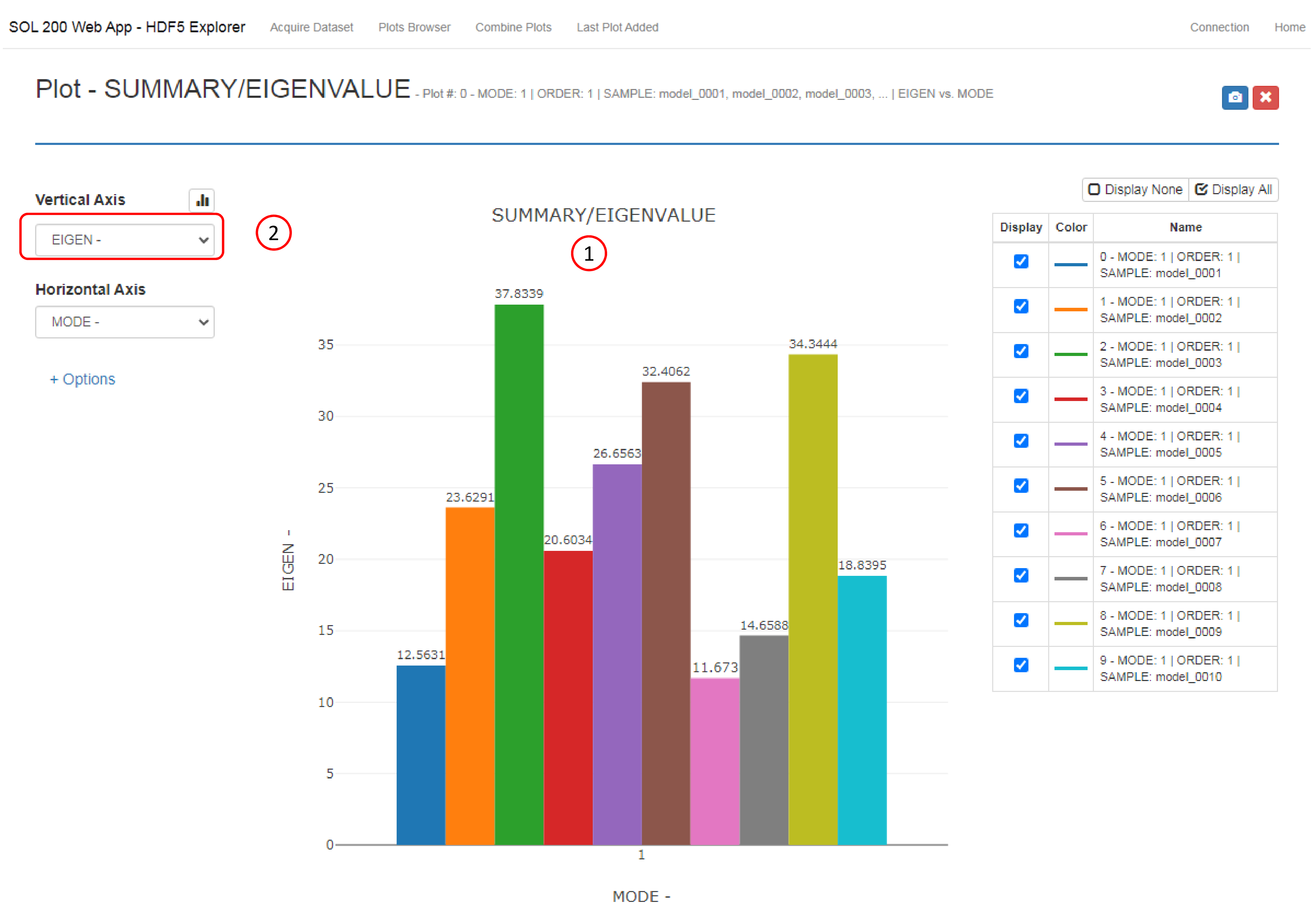




Review Results

1. After clicking on the image in the previous step, the respective plot is displayed. By default, a Frequency vs. Mode plot has been created. The plot display displays the 1st frequency for the samples.

2. An Eigenvalue vs. Mode plot is created by setting the vertical axis to EIGEN. The updated plot displays the 1st eigenvalue for the samples.



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

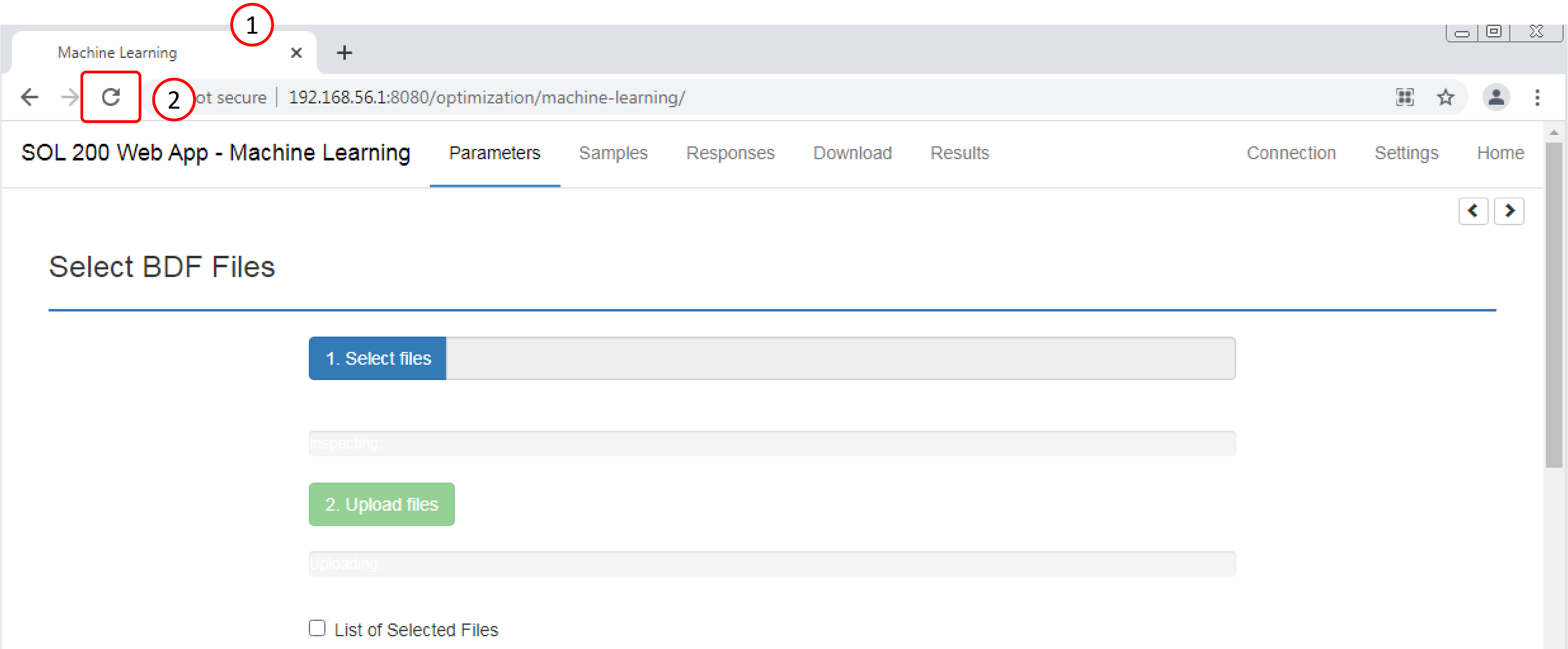
The parameters, samples and responses are contained in the following files

- 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

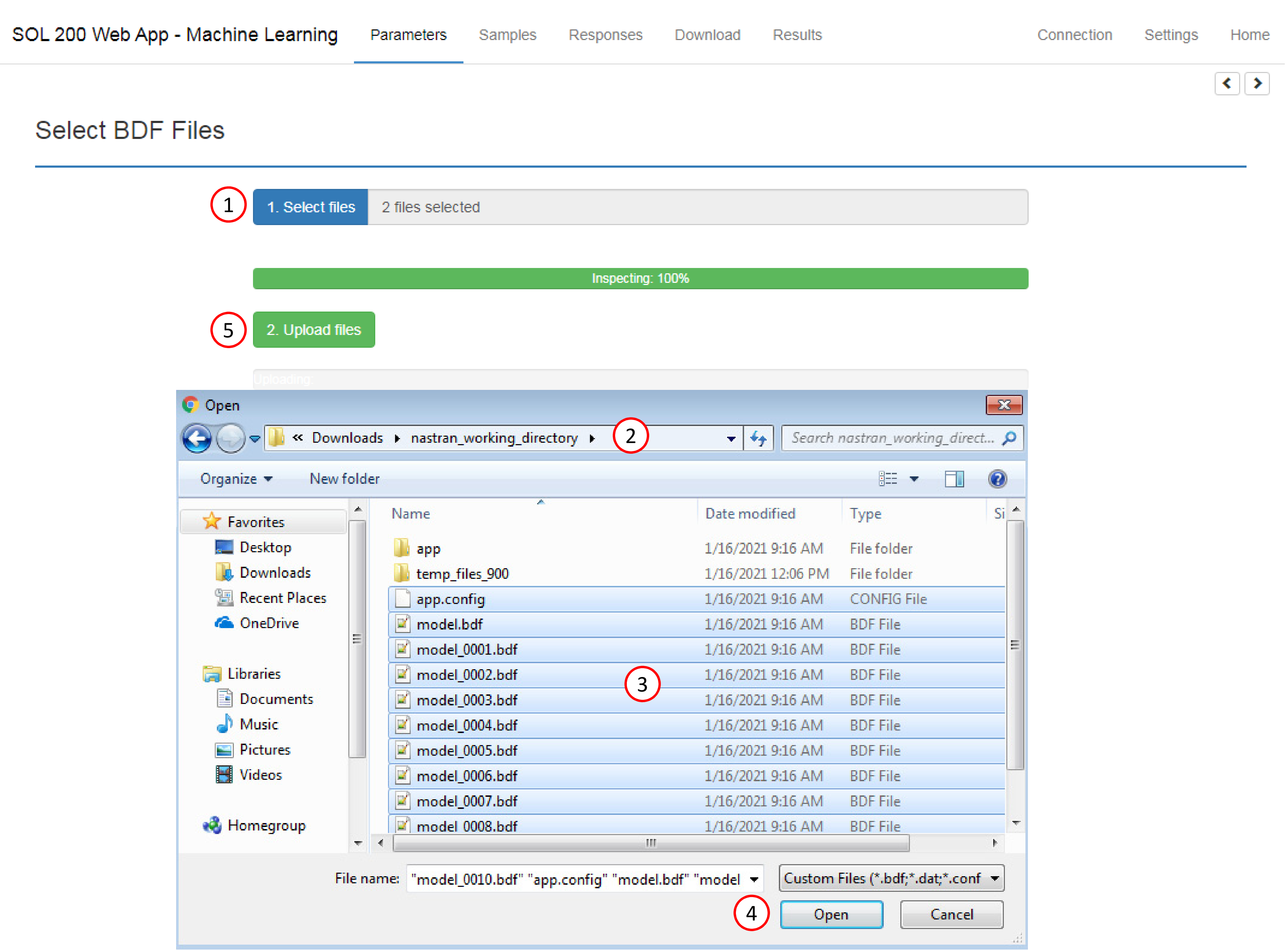


- Refreshing the page is only required when the *Select files* button is disabled.

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

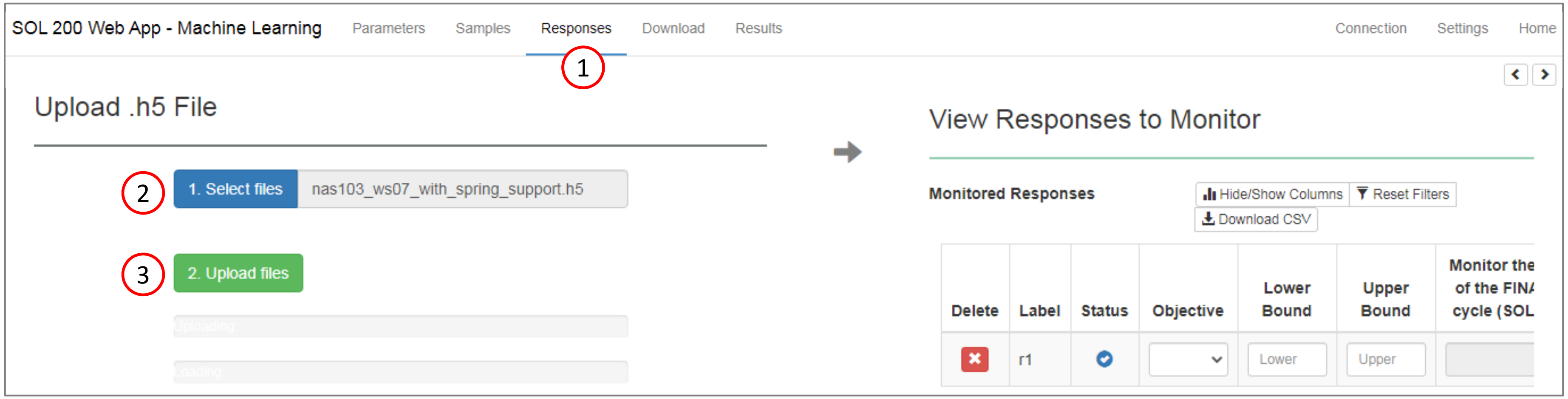
- All imports require the app.config file to be selected.

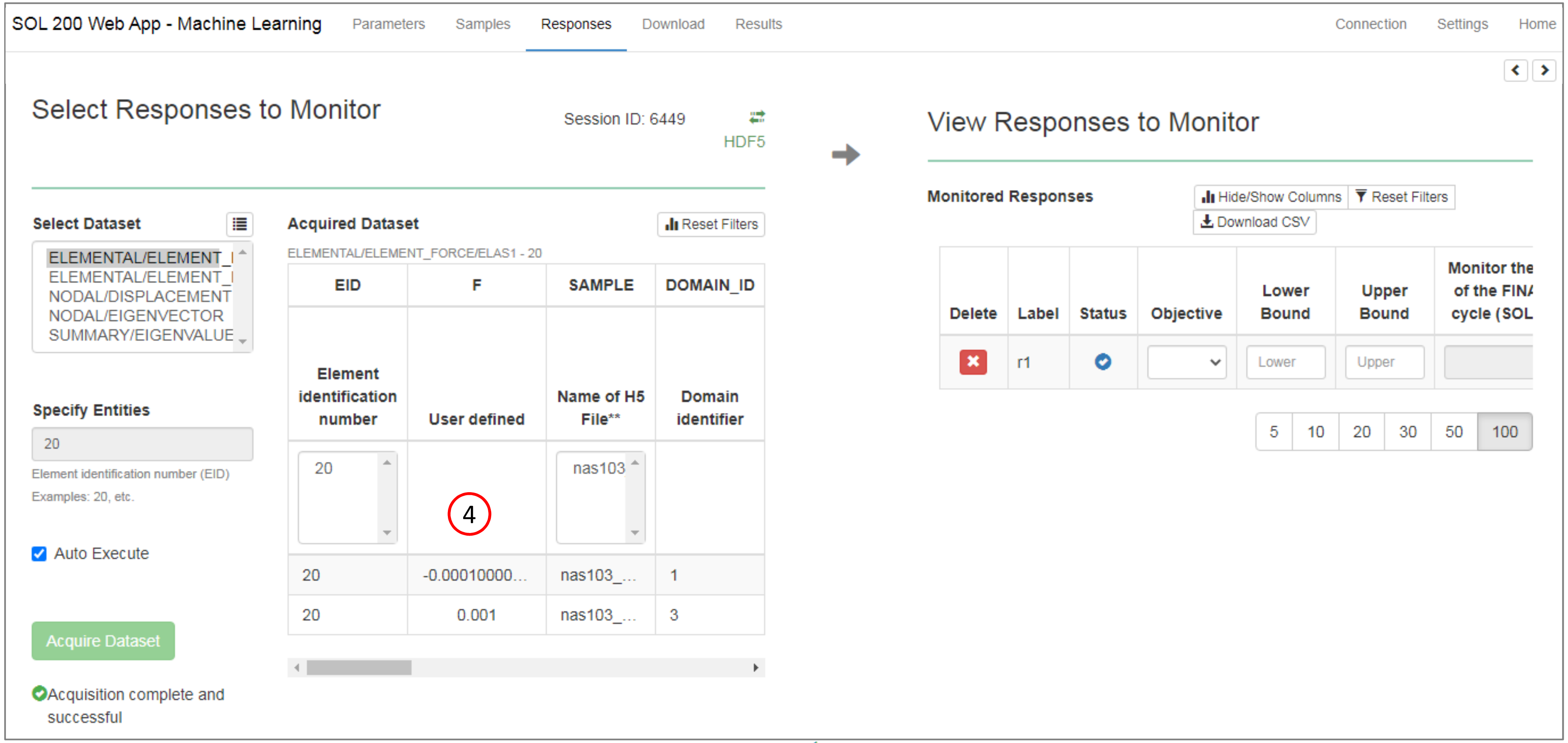


Import

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

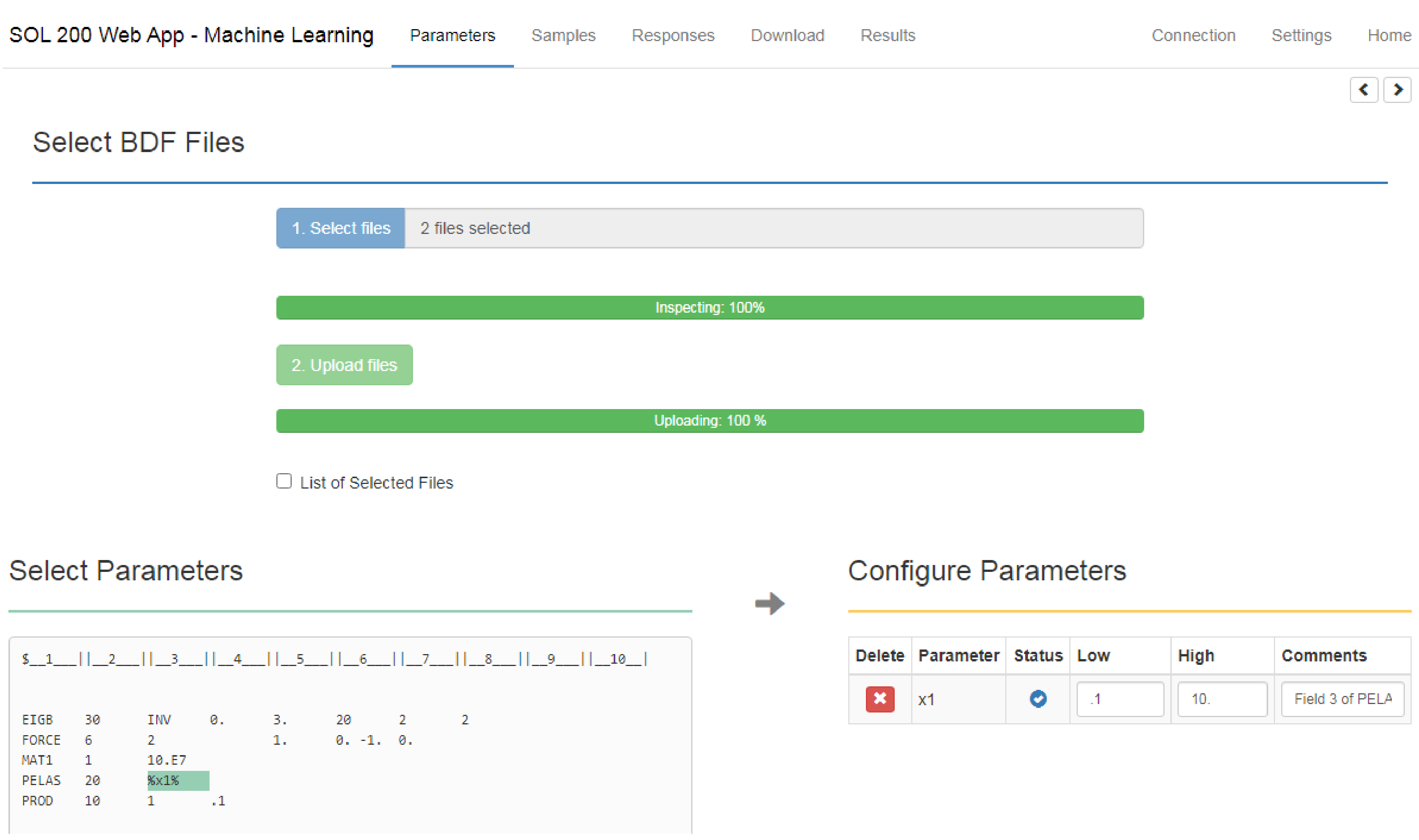
1. Click Responses
2. Select the H5 file
3. Click Upload
4. Data from the H5 is loaded and ready to use





Import

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



What is Gaussian Process Regression?

Gaussian Process Regression Overview

Training Data

D_n: Training data X_n (inputs) and Y_n (outputs)

Kernel (Covariance Function)*

$k(x, x') = \Sigma(x, x')$

MVN Conditioning Equations

$$\mu(\mathcal{X}) = \Sigma(\mathcal{X}, X_n)\Sigma_n^{-1}Y_n \qquad \text{(Prediction Model)}$$

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

Predicted Values

Mean: $\mu(x)$

Prediction Uncertainty

Variance: $\sigma^2(x)$

Sometimes expressed as $\Sigma(x)$

* 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
- $\Sigma(\chi,X_n)$: 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”

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

Prediction Model (Vary χ to make predictions)

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

Prediction Uncertainty

Example 1

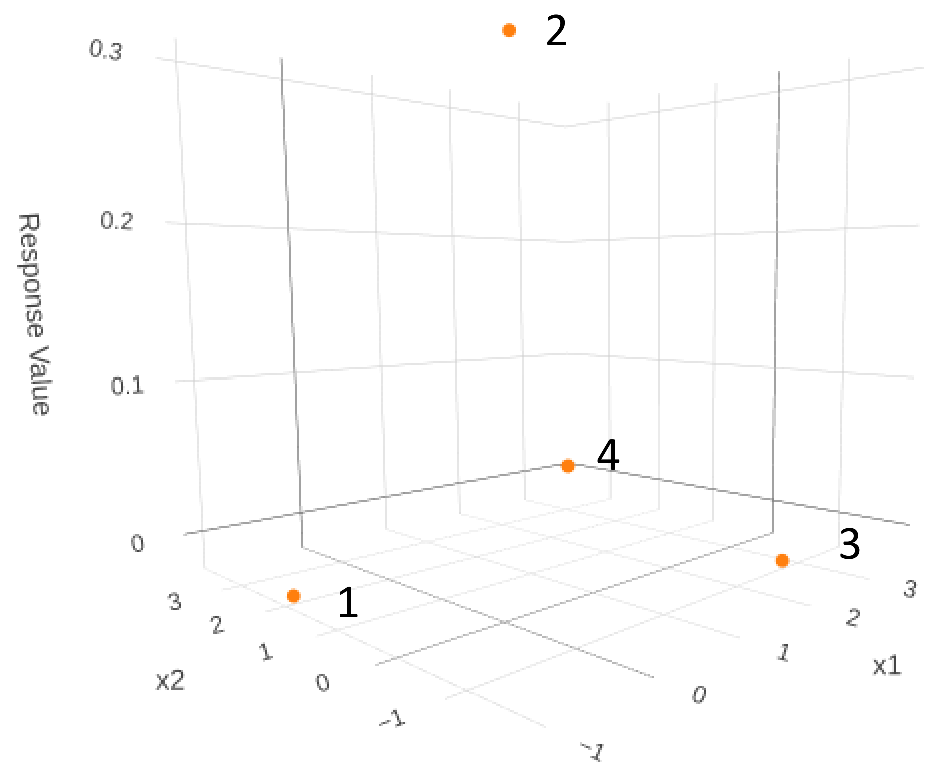
Example 1

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)

$$D = \|X - X\|^2$$

“Norm between X and X, squared”

3. Calculate Σ (Covariance Matrix)

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

Calculation of D

$$
D = \begin{bmatrix}
\begin{gathered}\sqrt{(.35\ -\ .35)^2 + (.69\ -\ .69)^2}^{2}\\ = 0\end{gathered} & \begin{gathered}\sqrt{(.35\ -\ .65)^2 + (.69\ -\ .46)^2}^{2}\\ = .1429\end{gathered} & \begin{gathered}\sqrt{(.35\ -\ -1.03)^2 + (.69\ -\ 1.76)^2}^{2}\\ = 3.0493\end{gathered} & \begin{gathered}\sqrt{(.35\ -\ .49)^2 + (.69\ -\ .49)^2}^{2}\\ = .0596\end{gathered} & \begin{gathered}\sqrt{(.35\ -\ 1.77)^2 + (.69\ -\ -1.77)^2}^{2}\\ = 8.068\end{gathered} & \begin{gathered}\sqrt{(.35\ -\ 3.62)^2 + (.69\ -\ 3.76)^2}^{2}\\ = 20.1178\end{gathered} \\
.1429 & \begin{gathered}\sqrt{(.65\ -\ .65)^2 + (.46\ -\ .46)^2}^{2}\\ = 0\end{gathered} & \begin{gathered}\sqrt{(.65\ -\ -1.03)^2 + (.46\ -\ 1.76)^2}^{2}\\ = 4.5124\end{gathered} & \begin{gathered}\sqrt{(.65\ -\ .49)^2 + (.46\ -\ .49)^2}^{2}\\ = .0265\end{gathered} & \begin{gathered}\sqrt{(.65\ -\ 1.77)^2 + (.46\ -\ -1.77)^2}^{2}\\ = 6.2273\end{gathered} & \begin{gathered}\sqrt{(.65\ -\ 3.62)^2 + (.46\ -\ 3.76)^2}^{2}\\ = 19.7109\end{gathered} \\
3.0493 & 4.5124 & \begin{gathered}\sqrt{(-1.03\ -\ -1.03)^2 + (1.76\ -\ 1.76)^2}^{2}\\ = 0\end{gathered} & \begin{gathered}\sqrt{(-1.0\ 3 -\ .49)^2 + (1.76\ -\ .49)^2}^{2}\\ = 3.9233\end{gathered} & \begin{gathered}\sqrt{(-1.03\ -\ 1.77)^2 + (1.76\ -\ -1.77)^2}^{2}\\ =20.3009\end{gathered} & \begin{gathered}\sqrt{(-1.03\ -\ 3.62)^2 + (1.76\ -\ 3.76)^2}^{2}\\ =25.6225\end{gathered} \\
.0596 & .0265 & 3.9233 & \begin{gathered}\sqrt{(.49\ -\ .49)^2 + (.49\ -\ .49)^2}^{2}\\ = 0\end{gathered} & \begin{gathered}\sqrt{(.49\ -\ 1.77)^2 + (.49\ -\ -1.77)^2}^{2}\\ = 6.746\end{gathered} & \begin{gathered}\sqrt{(.49\ -\ 3.62)^2 + (.49\ -\ 3.76)^2}^{2}\\ = 20.4898\end{gathered} \\
8.068 & 6.2273 & 20.3009 & 6.746 & \begin{gathered}\sqrt{(1.77\ -\ 1.77)^2 + (-1.7\ 7 -\ -1.77)^2}^{2}\\ = 0\end{gathered} & \begin{gathered}\sqrt{(1.77\ -\ 3.62)^2 + (-1.77\ -\ 3.76)^2}^{2}\\ = 34.0034\end{gathered} \\
20.1178 & 19.7109 & 25.6225 & 20.4898 & 34.0034 & \begin{gathered}\sqrt{(3.62\ -\ 3.62)^2 + (3.76\ -\ 3.76)^2}^{2}\\ = 0\end{gathered}
\end{bmatrix}
$$

Calculation of Σ

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

Calculation of Σ

$$\Sigma = \begin{bmatrix} e^{0}=1 & e^{-.1429}=.8668 & e^{-3.0493}=.0474 & e^{-.0596}=.9421 & e^{-8.068}=.0003 & e^{-20.1178}=1.832\text{e-}9 \\ .8668 & e^{0}=1 & e^{-4.5124}=.0110 & e^{-.0265}=.9738 & e^{-6.2273}=.0020 & e^{-19.7109}=2.8\text{e-}9 \\ .0474 & .0110 & e^{0}=1 & e^{-3.9233}=.0198 & e^{-20.3009}=1.5\text{e-}9 & e^{-25.6225}=7.5e-12 \\ .9421 & .9738 & .0198 & e^{0}=1 & e^{-6.746}=.0012 & e^{-20.4893}=1.263\text{e-}9 \\ .0003 & .0020 & 1.5\text{e-}9 & .0012 & e^{0}=1 & e^{-34.0034}=1.7e-15 \\ 1.832\text{e-}9 & 2.8\text{e-}9 & 7.5e-12 & 1.263\text{e-}9 & 1.7e-15 & e^{0}=1 \end{bmatrix}$$

$\Sigma(\chi,\chi)$

$\Sigma(\chi,X_n)$

$\Sigma(X_n,\chi)$

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

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

Calculation of Predictive Quantities

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

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

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

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

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

```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) = x1 * exp(-x1^2 - x2^2)
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

Console Terminal Jobs

```
~/
> # ################################################
> # 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

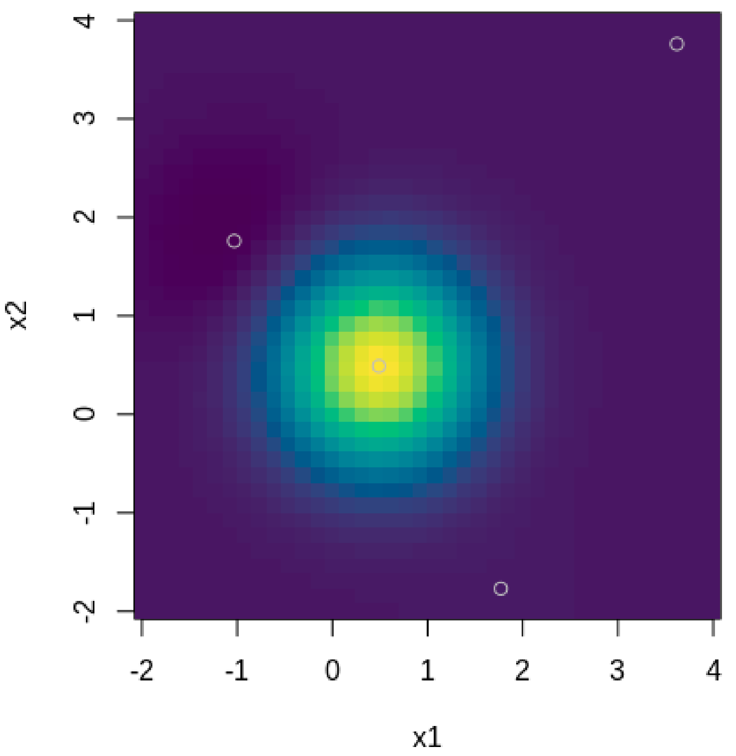
Variance

R
Code to replicate this example in R with Plots

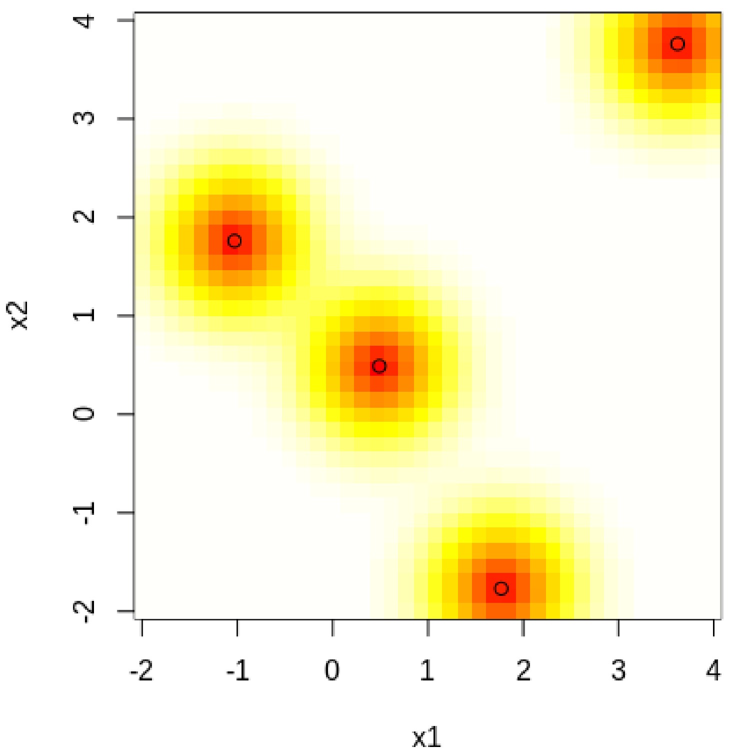
```r
library(plgp)
library(lhs)

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) = x1 * exp(-x1^2 - x2^2)
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

Mean: $\mu(\chi)$



Standard Deviation: $\sigma(\chi)$
(Square root of variance)



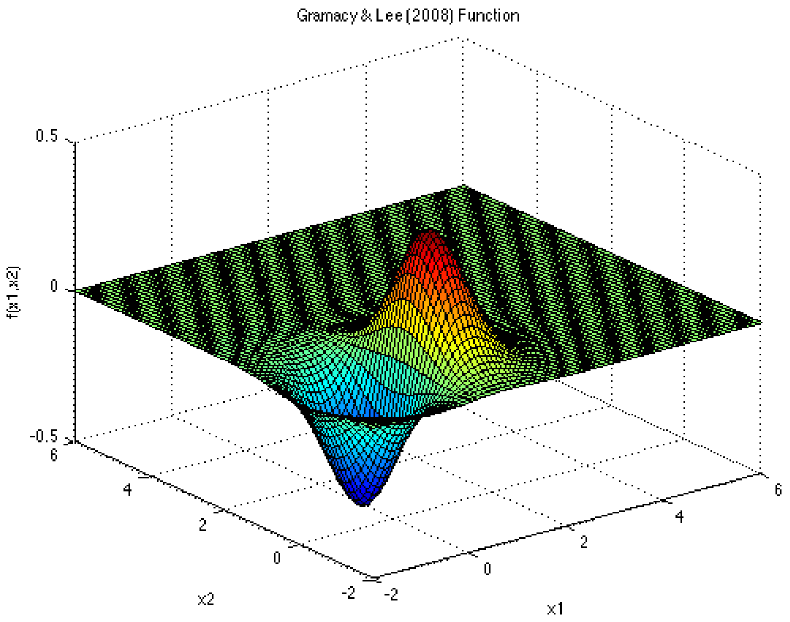
Not Good – High Standard Deviation – High Variance – High Prediction uncertainty (Prediction is not reliable)

Good – Low Standard Deviation – Low Variance – Low Prediction uncertainty (Prediction is reliable)

Prediction uncertainty is lowest near training data locations

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 ($\mu(\chi)$)